



# Riemannian fast-marching on cartesian grids using Voronoi's first reduction of quadratic forms

Jean-Marie Mirebeau

## ► To cite this version:

Jean-Marie Mirebeau. Riemannian fast-marching on cartesian grids using Voronoi's first reduction of quadratic forms. 2017. hal-01507334v4

**HAL Id: hal-01507334**

**<https://hal.science/hal-01507334v4>**

Preprint submitted on 14 Jun 2019

**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.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

# Riemannian fast-marching on cartesian grids, using Voronoi's first reduction of quadratic forms

Jean-Marie Mirebeau\*

June 14, 2019

## Abstract

We address the numerical computation of distance maps with respect to Riemannian metrics of strong anisotropy. For that purpose, we solve generalized eikonal equations, discretized using adaptive upwind finite differences on a Cartesian grid, in a single pass over the domain using a variant of the fast marching algorithm. The key ingredient of our PDE numerical scheme is Voronoi's first reduction, a tool from discrete geometry which characterizes the interaction of a quadratic form with an additive lattice. This technique, never used in this context, which is simple and cheap to implement, allows us to efficiently handle Riemannian metrics of eigenvalue ratio  $10^2$  and more.

Two variants of the introduced scheme are also presented, adapted to sub-Riemannian and to Rander metrics, which can be regarded as degenerate Riemannian metrics and as Riemannian metrics perturbed with a drift term respectively. We establish the convergence of the proposed scheme and of its variants, with convergence rates. Numerical experiments illustrate the effectiveness of our approach in various contexts, in dimension up to five, including an original sub-Riemannian model related to the penalization of path torsion.

**Keywords:** Riemannian metric, Sub-Riemannian metric, Rander metric, Eikonal equation, Viscosity solution, Fast Marching Method, Voronoi Reduction.

## 1 Introduction

In this paper, we develop a new and efficient numerical method for the computation of distance maps with respect to anisotropic Riemannian metrics, sub-Riemannian metrics and Rander metrics. For that purpose we discretize generalized eikonal equations, also called static first-order Hamilton Jacobi Bellman (HJB) Partial Differential Equations (PDEs), on a Cartesian grid. The novelty of our approach lies on a special representation of the Hamiltonian, via upwind finite differences on an adaptive stencil, which is designed using Voronoi's first reduction of quadratic forms [67] - a tool from discrete geometry mostly known for its applications in the study sphere packings and in number theory. This construction yields a discretized system of equations representing a Riemannian eikonal equation, that can be solved in *a single pass over the domain*, a property that ensures quasi-linear computation times -  $\mathcal{O}(N \ln N)$  where  $N$  is the problem size - and that is notoriously hard to achieve for anisotropic HJB equations, see [13] and the discussion below. For this reason, the method is referred to as Fast-Marching using Voronoi's First Reduction (FM-VR1).

---

\*University Paris-Sud, CNRS, University Paris-Saclay, 91405, Orsay, France  
This work was partly supported by ANR research grant MAGA, ANR-16-CE40-0014

Before entering the details of the addressed PDEs and of their discretizations, let us mention some of the potential applications. The standard eikonal equation reads  $\|du\| = f$ , with suitable boundary conditions, where  $du$  denotes the differential of a function  $u$  defined on a domain of  $\mathbb{R}^d$ , and the r.h.s.  $f$  is given and positive. This PDE characterizes distance maps with respect to an *isotropic* metric, defined locally as  $f$ -times the Euclidean metric, see [69] for a study and an overview of its numerous applications. Once the distance map is computed, globally optimal minimal paths (geodesics) w.r.t. the metric can be extracted by gradient descent, with again numerous applications in e.g. image processing [58] or motion planning. This paper is devoted to the numerical solution of generalized eikonal equations, characterizing distance maps with respect to *anisotropic* metrics, of the three following types.

- A *Riemannian metric* on a domain of  $\mathbb{R}^d$  is described by a field  $\mathcal{M}$  of positive definite tensors, and gives rise to the generalized eikonal equation  $\|du\|_{\mathcal{M}^{-1}} = 1$ . Numerical methods for Riemannian distance computation have applications in geometry processing [64], optics [37], statistics with the Fisher-Rao distance, ... In image processing and segmentation, anisotropic Riemannian metrics are often used to favor paths aligned with tubular structures of interest [38, 7, 18].
- A *sub-Riemannian metric* can be regarded as a degenerate Riemannian metric, whose tensors have some infinite eigenvalues [51]. As a result, motion is only possible along a subspace of the tangent space, depending on the current position. This property is referred to as non-holonomy, and models for instance a robotic system with fewer controls than degrees of freedom. A fundamental instance is the Reeds-Shepp car model, posed on the configuration space  $\mathbb{R}^2 \times \mathbb{S}^1$ , which can move forward and backward, rotate, but not translate sideways, see [66, 28] for a numerical study with applications to image segmentation and motion planning. A variant, presented in this paper §3.2, is related to the penalization of path torsion.
- A *Rander metric* is defined locally as the sum of a Riemannian metric  $\mathcal{M}$  and of a sufficiently small co-vector field  $\hat{\eta}$ , see [61] and §1.3. These metrics are non-symmetric, thus define asymmetric distances, and give rise to the inhomogeneous generalization of the eikonal equation  $\|du - \hat{\eta}\|_{\mathcal{M}^{-1}} = 1$ . The travel-time of a boat subject to a drift due to water currents can be measured by integrating a Rander metric, see §3, and its optimization is called Zermelo's navigation problem [2, 15]. In image segmentation, the Chan-Vese energy of a region can be reformulated as the length of its contour measured w.r.t a Rander metric [17].

Our numerical approach has its limitations: it cannot address more general anisotropic metrics than the above ones, such as those arising in seismic imaging [63], and it cannot handle domains discretized using triangulations or unstructured point sets [41]. Indeed, the algorithmic tools that we leverage [67] limit the scope of our method to eikonal equations whose Hamiltonian has a quadratic structure, and to domains discretized using a Cartesian grid. As often, efficiency is at the cost of specialization. In order to better describe the advantages and the specificities of our approach, let us formally state the addressed problem and review the existing methods.

This paper is devoted to the construction and analysis of a numerical scheme for computing the arrival times  $u : \Omega \rightarrow \mathbb{R}$  of a front starting from the boundary of a domain  $\Omega$ , and propagating at unit speed w.r.t. a given metric  $\mathcal{F} : T\Omega \cong \Omega \times \mathbb{R}^d \rightarrow [0, \infty]$  of one of the above three classes. Several classes of methods can be distinguished in the literature for such purposes.

- *Eulerian schemes*, such as the one presented in this paper, rely on a characterization of the arrival times as the unique viscosity solution [22] to the eikonal PDE, which reads

$$\forall \mathbf{p} \in \Omega, \mathcal{H}_{\mathbf{p}}(du(\mathbf{p})) = 1/2, \quad \forall \mathbf{p} \in \partial\Omega, u(\mathbf{p}) = 0, \quad (1)$$

where  $\mathcal{H}$  denotes the Hamiltonian associated with the metric, i.e. the Legendre-Fenchel transform of the Lagrangian  $\frac{1}{2}\mathcal{F}^2$ . Finite differences are typically used for discretization [80, 65, 71], although discontinuous Galerkin methods have recently been considered [44].

A number of finite difference schemes have been developed for anisotropic eikonal equations, associated with Riemannian metrics [27, 77], with various classes of Finslerian anisotropy characterizing for instance the travel-time of seismic waves [29, 45, 39, 60, 75, 43, 35], or with sub-Riemannian anisotropy [5]. Some of these works achieve high order accuracy [35]. Note, however, that there is a fundamental difference distinguishing these Eulerian schemes on one side, and [71] and the scheme proposed in this paper on the other side. Indeed the later obey a property, here referred to as causality<sup>1</sup>, see Definition 2.1, which allows to solve the discretized system of equations in a single pass, using the fast marching method as in [71, 78, 41, 72, 81, 1]. In contrast, the former are typically solved numerically using the iterative fast sweeping method, see the discussion below.

- *Semi-Lagrangian schemes* rely on a self-consistency property of the arrival times, referred to as Bellman’s optimality principle: for any point and neighborhood  $\mathbf{p} \in V(\mathbf{p}) \subseteq \Omega$ , one has

$$u(\mathbf{p}) = \min_{\mathbf{q} \in \partial V(\mathbf{p})} \left( u(\mathbf{q}) + d_{\mathcal{F}}(\mathbf{q}, \mathbf{p}) \right), \quad (2)$$

where  $d_{\mathcal{F}}$  denotes the path-length distance associated with the metric  $\mathcal{F}$ . In the semi-Lagrangian paradigm, a discrete counterpart of (2) is implemented numerically, by constructing polygonal stencils  $V(\mathbf{p})$  whose vertices lie among the discretization points, interpolating the unknown  $u$  on the facets of  $\partial V(\mathbf{p})$ , and locally approximating the distance with the metric  $d_{\mathcal{F}}(\mathbf{q}, \mathbf{p}) \approx \mathcal{F}_{\mathbf{p}}(\mathbf{p} - \mathbf{q})$ . A natural choice for  $V(\mathbf{p})$  is the union of the triangles containing the vertex  $\mathbf{p}$  in a given mesh of the domain  $\Omega$  [10]. Following the discovery [41, 81] that a generalized *acuteness* property obeyed by the facets of  $V(\mathbf{p})$  enables solving the discretized system in an efficient, single pass manner, see below, a number of more complex designs have been proposed [11, 72, 38, 1]. Constructions based on algorithmic geometry, introduced by the author in [47, 48, 49], allow to satisfy the acuteness property for strongly anisotropic metrics while limiting the size and the cardinality of the stencils.

- *Heat related methods* solve a diffusion equation on a short time interval [24], or an elliptic equation with a small parameter, and exploit the relationship between the geodesic distance and the short time asymptotics of the heat kernel [79]. This approach is limited to Riemannian metrics, either isotropic or anisotropic [84]. Its efficiency is tied to the numerical cost of solving sparse linear systems discretizing a Laplacian, which is often favorable over alternative methods, especially in dimension  $d = 2$ , thanks to the existence of highly optimized linear algebra libraries. The method requires some parameter tuning, since it involves two small scales (in time and space), and loses accuracy or degenerates to a graph distance if their relative magnitude is incorrectly set [24].
- *Path based techniques* compute minimal geodesics directly, rather than extracting them from the front arrival times. Ray-tracing techniques solve Hamilton’s Ordinary Differential Equations (ODE) of motion from a point  $\mathbf{p}$  of interest, adjusting the initial velocity

---

<sup>1</sup>In other works, causality has been attributed other meanings, such as upwindness, see Remark 2.2.

direction until the desired target is reached [16]. Path bending methods progressively deform a path joining two endpoints of interest, so as to obey Hamilton’s ODEs [63]. Path-based methods can be very accurate, by using high order ODE integration schemes, and do not suffer from the curse of dimensionality, since the computational domain needs not be discretized. However, they lack robustness, have difficulty handling obstacles, and one usually cannot guarantee that the path found is globally the shortest one.

Among the first two classes of methods, Eulerian and semi-Lagrangian discretization schemes, a further distinction must be made depending on the numerical solver of the coupled, non-linear system of equations resulting from the discretization of (1) or (2). In this paper, we put an emphasis on *single pass methods*, which solve this system of equations deterministically, in a finite number of steps, visiting each point of the discretization grid a bounded number of times that is fixed in advance and independent of the problem instance. Single-pass methods guarantee short and predictable computation times, but are notoriously harder to develop than iterative methods when the metric is anisotropic [13].

- *Single pass methods*, such as the one proposed in this paper<sup>2</sup>, are referred to as Fast Marching Methods (FMM), and can be regarded as variants of Dijkstra’s algorithm. This approach is computationally efficient, with complexity  $\mathcal{O}(\lambda N \ln N)$  where  $N$  is the number of discretization points, and  $\lambda$  the average number of neighbors of each discretization point in the numerical scheme, see §A. (Complexity can be improved to  $\mathcal{O}(N)$  in certain circumstances [78, 40].) However, it is only applicable if the discretization scheme obeys a property referred to as *causality*, see Definition 2.1. Among Eulerian schemes, this property holds for the natural discretization of the isotropic eikonal equation [71], but could not be extended to anisotropic metrics until the FM-VR1 presented here. In the case of semi-Lagrangian schemes, causality is related to a geometrical property of the stencil, see [41, 81, 11, 72, 38, 1, 47, 48, 49] and the above discussion.

Alternatively, Dijkstra’s method can also be used directly to compute shortest paths on a network with sufficiently fine connectivity within the domain [52, 76], whose edge-lengths defined according to the metric. The resulting approximation of the front arrival times typically lacks accuracy, but adequate heuristics may improve it [14].

- *Iterative methods* apply Gauss-Siedel updates to the numerical solution, of the system of equations discretizing the problem of interest, until it meets a convergence criterion. The ordering of the discretization points at which the updates are applied can be more or less sophisticated:
  - a simultaneous update of all the grid points is used in [34, 31, 65, 5].
  - fast sweeping methods alternate sweeps along the  $2d$  directions of the grid [85, 46, 77, 43, 35, 44]. In other words updates are applied line after line of the grid, then column after column, then in reverse order, repeatedly.
  - adaptive Gauss Siedel iterations (AGSI) are considered in [10], whose ordering is determined using a queue.
  - the buffered fast marching method, introduced in [25], applies the updates within a buffer which monotonically progresses similar to a front, and whose width is determined by the dynamics of the equation - the buffer may cover the whole domain in the worst case. See also the related progressive fast marching method [12].

---

<sup>2</sup>Except for the numerical scheme devoted to Rander metrics

The complexity of the AGSI is  $\mathcal{O}(N^{1+1/d}/\ln \varepsilon)$  according to [10], where  $N$  is the discrete domain cardinality,  $d$  is the domain dimension, and  $\varepsilon > 0$  is the desired accuracy of the solution. The complexity of the fast sweeping method is  $\mathcal{O}(N)$  according to [85], but this favorable estimate requires the problem instance to be nice enough, see the discussion in §3.1. In contrast with single-pass methods, the constant hidden in these complexity estimates depends on the problem instance and may increase substantially if the metric is strongly anisotropic or if the minimal paths change direction often (due to obstacles or to non-uniformities in the metric) [7, 47].

Yet another approach is to introduce a time variable and solve the time dependent PDE  $\partial_t u + \mathcal{H}u - 1/2 = 0$ , whose asymptotic steady state obeys the static equation (1), under suitable assumptions and with suitable boundary conditions [44, 5].

Recapitulating, the FM-VR1 scheme introduced in this paper is the first numerical solver of eikonal equations that is simultaneously (i) *Eulerian*, (ii) solvable in a *single pass*<sup>3</sup>, and (iii) compatible with several classes of *anisotropic* metrics. For these reasons it is simple to implement, fast to solve numerically independently of the problem instance, and has a wide application scope and generalization potential, see e.g. [50] for a variant devoted to the global optimization of path energies involving curvature.

*Contributions.* We describe numerical schemes devoted to the computation of Riemannian, sub-Riemannian, and Rander distances, by solving the corresponding generalized eikonal equations. We prove convergence rates, based on the doubling of variables technique, see chapter 10 of [30] or [73], which applies rather directly in the Riemannian case but requires non-trivial adaptations in the sub-Riemannian and Rander cases. Numerical experiments illustrate the efficiency of our numerical schemes, in dimension  $2 \leq d \leq 5$ , and their potential applications in image segmentation and motion planning.

*Outline.* The rest of this introduction is devoted to general notations, to the description of Voronoi's first reduction which is a key ingredient of our discretization, and to elements of optimal control. The impatient reader may jump to §1.1, §1.2 and §1.3, where the numerical schemes are described and the convergence results stated, in the Riemannian, sub-Riemannian and Rander cases respectively. Convergence proofs are provided in §2, and in Appendices D and E respectively. Numerical experiments are presented in §3.

**General notations.** The ambient space dimension is fixed and denoted by  $d$ . The Euclidean space and the Cartesian grid are respectively denoted

$$\mathbb{E} := \mathbb{R}^d, \quad \mathbb{L} := \mathbb{Z}^d.$$

Let  $\Omega \subseteq \mathbb{E}$  be a domain, assumed throughout this paper to be bounded; additional geometrical assumptions are required in some results. For any grid scale  $h > 0$  we let

$$\Omega_h := \Omega \cap h\mathbb{L}, \quad \partial\Omega_h := (\mathbb{E} \setminus \Omega) \cap h\mathbb{L}. \quad (3)$$

Geometric points are denoted  $\mathbf{p} \in \Omega$ , vectors  $\dot{\mathbf{p}} \in \mathbb{E}$ , and co-vectors  $\hat{\mathbf{p}} \in \mathbb{E}^*$ . The symbol  $\gamma$  is reserved for paths within  $\bar{\Omega}$  and has the special convention that  $\dot{\gamma}(t) := \frac{d}{dt}\gamma(t)$  denotes time derivation. We denote by  $\|\dot{\mathbf{p}}\|$  the Euclidean norm, by  $(\dot{\mathbf{p}} \cdot \dot{\mathbf{q}})$  the scalar product, and by  $\langle \hat{\mathbf{p}}, \dot{\mathbf{q}} \rangle$  the duality bracket, where  $\dot{\mathbf{p}}, \dot{\mathbf{q}} \in \mathbb{E}$  are vectors and  $\hat{\mathbf{p}} \in \mathbb{E}^*$  is a co-vector. Denote by  $\text{GL}(\mathbb{E}) \subseteq \text{L}(\mathbb{E}, \mathbb{E})$  the group of invertible linear transformations, and by  $\text{GL}(\mathbb{L}) \subseteq \text{GL}(\mathbb{E})$  the subgroup of those which leave the Cartesian grid  $\mathbb{L}$  invariant - equivalently their matrix has

---

<sup>3</sup>Except the variant devoted to Rander metrics.

integer coefficients and determinant  $\pm 1$ . Denote by  $S(\mathbb{E}) \subseteq L(\mathbb{E}, \mathbb{E}^*)$  the space of symmetric linear maps, by  $S^+(\mathbb{E})$  the subset of semi-definite ones, and by  $S^{++}(\mathbb{E})$  the positive definite ones. We adopt the notations

$$\|\dot{\mathbf{p}}\|_M := \sqrt{\langle M\dot{\mathbf{p}}, \dot{\mathbf{p}} \rangle}, \quad \hat{\mathbf{p}} \otimes \hat{\mathbf{p}} \in S^+(\mathbb{E}),$$

for the norm of  $\dot{\mathbf{p}} \in \mathbb{E}$  induced by  $M \in S^{++}(\mathbb{E})$ , and for the self outer product of  $\hat{\mathbf{p}} \in \mathbb{E}^*$ . The dual vector space  $\mathbb{E}^*$  and the dual lattice<sup>4</sup>  $\mathbb{L}^*$  can be naturally identified with their primal counterparts:  $\mathbb{E}^* \cong \mathbb{E}$  and  $\mathbb{L}^* \cong \mathbb{L}$  using the Euclidean structure, but the distinction is kept for clarity.

**Tensor decomposition based on Voronoi's first reduction of quadratic forms.** In the design of our numerical scheme, an essential ingredient is the ability to decompose an arbitrary tensor  $D \in S^{++}(\mathbb{E})$  in the form

$$D = \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \rho^{\dot{\mathbf{e}}}(D) \dot{\mathbf{e}} \dot{\mathbf{e}}^T, \quad \text{where } \forall \dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}, \rho^{\dot{\mathbf{e}}}(D) \geq 0, \quad (4)$$

and only finitely many of these coefficients are non-zero (in practice  $d(d+1)/2$  coefficients). The tensor decomposition (4) resembles the eigenvalue-eigenvector decomposition of the matrix  $D$ , except that it has more than  $d$  terms, and that the vectors belong to the lattice  $\mathbb{L}$  and are thus suitable for the construction of a finite difference scheme on the Cartesian grid.

For each positive definite tensor  $D \in S^{++}(\mathbb{E})$ , there are infinitely many decompositions in the form (4), hence we need a suitable criterion to construct and select one. Our choice is far from arbitrary: we choose a decomposition which maximizes the sum of the coefficients:

$$\sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \rho^{\dot{\mathbf{e}}}(D). \quad (5)$$

A possible motivation for this selection criterion is that large coefficients are typically associated with small offsets (and thus a local numerical scheme), as evidenced by the identity

$$\text{Tr}(D) = \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \rho^{\dot{\mathbf{e}}}(D) \|\dot{\mathbf{e}}\|^2.$$

However, the genuine reason for maximizing (5), subject to the constraints (4), is that the dual linear program is well known in the field of lattice geometry, and referred to as Voronoi's first reduction of the quadratic form defined by the matrix  $D$ . This problem was introduced by Voronoi [83] with the purpose of classifying the equivalence classes of positive quadratic forms, i.e. elements of  $S^{++}(\mathbb{E})$ , under the action of the group  $\text{GL}(\mathbb{L})$ , following a line of research dating back to Lagrange [42]. We refer to [67] for a modern and extensive presentation of Voronoi's theory.

Thanks to ingredients from this literature, the above matrix decomposition can be computed extremely efficiently, for each point of the discretized PDE domain, with an overall numerical cost that is dominated by the incompressible cost of maintaining the priority-queue inherent to the Fast-Marching (Dijkstra-like) method. We provide in Appendix B some additional detail on Voronoi's theory, a simple algorithm for computing the above decomposition in dimension  $d \leq 3$ , as well as the proof of the following result.

---

<sup>4</sup>Formally,  $\mathbb{L}^* = \{\hat{\mathbf{p}} \in \mathbb{E}^*; \forall \dot{\mathbf{p}} \in \mathbb{L}, \langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle \in \mathbb{Z}\}$ .

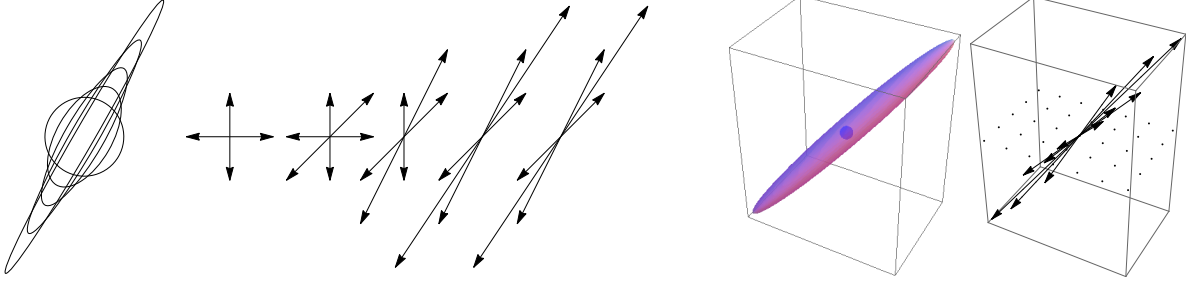


Figure 1: Ellipsoid  $\{\|\mathbf{p}\|_M \leq 1\}$ , and offsets appearing in the decomposition of  $D := M^{-1}$  by Proposition 1.1, for some  $M \in S^{++}(\mathbb{R}^d)$ , in dimension  $d = 2$  (left) or  $d = 3$  (right).

**Proposition 1.1.** *Let  $D \in S_d^{++}$ . Then the linear program of maximizing (5) subject to (4) admits at least one solution. Any such solution can be written in the form*

$$D = \sum_{1 \leq i \leq I} \rho_i \dot{\mathbf{e}}_i \dot{\mathbf{e}}_i^T,$$

where  $I$  is a (finite) integer,  $\rho_i > 0$  and  $\dot{\mathbf{e}}_i \in \mathbb{Z}^d \setminus \{0\}$  for all  $1 \leq i \leq I$ . One can always select a solution such that  $I \leq d(d+1)/2$ . Furthermore, one has

$$\|\dot{\mathbf{e}}_i\| \leq C(d)\mu(D)^{d-1}, \quad \text{where } \mu(D) := \sqrt{\|D\|\|D^{-1}\|}, \quad (6)$$

for all  $1 \leq i \leq I$ , with the improved estimate  $\|\dot{\mathbf{e}}_i\| \leq C\mu(D)$  in dimension  $d = 3$ .

Figure 1 illustrates, in dimension two and three, the close relationship between the anisotropy of the tensor  $D$  and the directions of its decomposition, which locally define the stencil points of our adaptive discretization (16) of the eikonal PDE. The quantity  $\mu(D)$ , which is the square root of the condition number of  $D$ , is referred to as its *anisotropy ratio*.

In addition to Eikonal equations, which are the object of the present paper, the tensor decomposition of Proposition 1.1 is used in [32] to design numerical schemes for two and three-dimensional anisotropic diffusion. An equivalent but two-dimensional only concept is applied in [6] to Monge-Ampere equations, and in [9] to HJB PDEs of stochastic control. See [49] for estimates related to (6) in the average case upon random rotations of the tensor  $D$ , in dimension two.

**Elements of optimal control.** We refer to [3] for an overview of optimal control theory and its PDE formulations, and only introduce here the notations and definitions required for our purposes. Let  $\mathfrak{C}(\mathbb{E})$  be the collection of compact and convex subsets of  $\mathbb{E}$  containing the origin, equipped with the Hausdorff distance. Denote  $\text{Lip}(X, Y)$  the class of Lipschitz maps, with arbitrary Lipschitz constant, from a metric space  $X$  to a metric space  $Y$ .

**Definition 1.2.** *A family of controls is an element  $\mathcal{B}$  of  $\mathfrak{B} := C^0(\overline{\Omega}, \mathfrak{C}(\mathbb{E}))$ , which continuously associates to each point  $\mathbf{p} \in \overline{\Omega}$  a control set  $\mathcal{B}(\mathbf{p})$ . A path  $\gamma \in \text{Lip}([0, T], \overline{\Omega})$ , where  $T \geq 0$ , is said  $\mathcal{B}$ -controllable iff for almost every  $t \in [0, T]$*

$$\dot{\gamma}(t) \in \mathcal{B}(\gamma(t)), \quad \text{where } \dot{\gamma}(t) := \frac{d}{dt} \gamma(t). \quad (7)$$

The minimal control time from  $\mathbf{p} \in \overline{\Omega}$  to  $\mathbf{q} \in \overline{\Omega}$ , is defined as

$$T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) := \inf\{T \geq 0; \exists \gamma \in \text{Lip}([0, T], \overline{\Omega}), \mathcal{B}\text{-controllable}, \gamma(0) = \mathbf{p}, \gamma(T) = \mathbf{q}\}. \quad (8)$$

The control sets corresponding to Riemannian, sub-Riemannian and Rander geometry are respectively ellipsoids, degenerate ellipsoids (with empty interior), and ellipsoids centered off the origin, see the illustrating figure. One easily shows that a minimal path from  $\mathbf{p}$  to  $\mathbf{q}$  exists as soon as  $T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) < \infty$ , using Arzela-Ascoli's compactness theorem and the fact that  $\Omega$  is bounded. See the appendices of [19, 28] for details, as well as related results such as the convergence of the control times and of the minimal paths associated with a converging family of controls under suitable assumptions. The above concepts can be rephrased in the framework of a local metric defined on the tangent space  $\mathcal{F} : \overline{\Omega} \times \mathbb{E} \rightarrow [0, \infty]$ : given controls  $\mathcal{B} \in \mathfrak{B}$ , define for all  $\mathbf{p} \in \overline{\Omega}$ ,  $\dot{\mathbf{p}} \in \mathbb{E}$ , and any path  $\gamma \in \text{Lip}([0, 1], \overline{\Omega})$

$$\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) := \inf\{\lambda > 0; \dot{\mathbf{p}}/\lambda \in \mathcal{B}(\mathbf{p})\}, \quad \text{Length}_{\mathcal{F}}(\gamma) := \int_0^1 \mathcal{F}_{\gamma(t)}(\dot{\gamma}(t)) dt.$$

Note that these quantities can be infinite if the control sets have empty interior, such as in the sub-Riemannian case, and can be asymmetric ( $\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) \neq \mathcal{F}_{\mathbf{p}}(-\dot{\mathbf{p}})$ ) if the control sets are not centered on the origin, as in the Rander case, see the illustrating figure and §1.2, §1.3. Conversely, the metric  $\mathcal{F}$  uniquely determines the control sets  $\mathcal{B}(\mathbf{p}) = \{\dot{\mathbf{p}} \in \mathbb{E}; \mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) \leq 1\}$ , and by time reparametrization the control time  $T_{\mathcal{B}}(\mathbf{p}, \mathbf{q})$  from  $\mathbf{p}$  to  $\mathbf{q} \in \Omega$  is shown equal to the (quasi-)distance

$$d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) := \inf\{\text{Length}_{\mathcal{F}}(\gamma); \gamma \in \text{Lip}([0, 1], \overline{\Omega}), \gamma(0) = \mathbf{p}, \gamma(1) = \mathbf{q}\}. \quad (9)$$

This paper is concerned with the exit time optimal control problem, whose value function is defined for all  $\mathbf{p} \in \overline{\Omega}$  by

$$u(\mathbf{q}) := \inf_{\mathbf{p} \in \partial\Omega} T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) \quad \left( = \inf_{\mathbf{p} \in \partial\Omega} d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) \right). \quad (10)$$

The numerical computation of the function  $u$  is the main topic of this paper. Under suitable assumptions [3], the function  $u$  is the unique viscosity solution to the following HJB PDE involving the dual metric  $\mathcal{F}^* : \overline{\Omega} \times \mathbb{E}^* \rightarrow \mathbb{R}_+$ : for all  $\mathbf{p} \in \Omega$

$$\mathcal{F}_{\mathbf{p}}^*(du(\mathbf{p})) = 1, \quad \text{where } \mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}}) := \sup_{\dot{\mathbf{p}} \neq 0} \frac{\langle \hat{\mathbf{p}}, \dot{\mathbf{p}} \rangle}{\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}})}, \quad (11)$$

and  $u(\mathbf{p}) = 0$  for all  $\mathbf{p} \in \partial\Omega$ . The formulations (11, left) and (1, left) of the eikonal PDE are equivalent, in view of the relation  $\mathcal{H} = \frac{1}{2}(\mathcal{F}^*)^2$  between the Hamiltonian and the dual metric. Once  $u$  is known, the shortest path from  $\partial\Omega$  to  $\mathbf{p} \in \Omega$  can be extracted by solving *backwards in time* the Ordinary Differential Equation (ODE)

$$\dot{\gamma}(t) := V(\gamma(t)), \quad \text{where } V(\mathbf{p}) := d\mathcal{F}_{\mathbf{p}}^*(du(\mathbf{p})), \quad (12)$$

with final condition  $\gamma(T) = \mathbf{p}$  where  $T = u(\mathbf{p})$ , see e.g. appendix C in [28]. In (12, right) the dual metric  $\mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}})$  is differentiated w.r.t. the variable  $\hat{\mathbf{p}}$ . Note that  $d\mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}}) \in (\mathbb{E}^*)^* \cong \mathbb{E}$ . For robust numerical geodesic backtracking it is essential to use an upwind estimation of the vector field  $V(\mathbf{p})$ , see Appendix C.

## 1.1 Riemannian metrics

A Riemannian metric on the bounded domain  $\Omega \subseteq \mathbb{E}$  is described via a field of symmetric positive definite tensors  $\mathcal{M} \in C^0(\overline{\Omega}, S^{++}(\mathbb{E}))$ . The metric function  $\mathcal{F} : \overline{\Omega} \times \mathbb{E} \rightarrow \mathbb{R}_+$  has the expression

$$\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) := \|\dot{\mathbf{p}}\|_{\mathcal{M}(\mathbf{p})}. \quad (13)$$

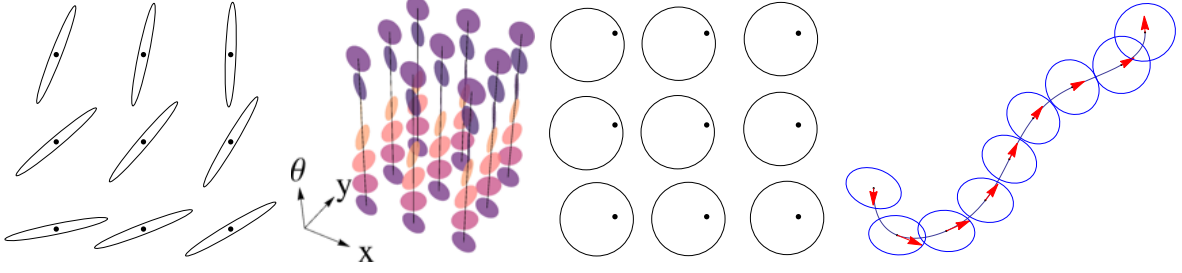


Figure: Examples of control sets for a (i) Riemannian, (ii) sub-Riemannian, and (iii) Rander metric. (iv) An admissible path, with tangents shown in red, w.r.t to some controls. (Illustration absent from journal version.)

Our objective is to compute the Riemannian distance  $u : \bar{\Omega} \rightarrow \mathbb{R}$  to the boundary of  $\Omega$ , see (10), which is known to be the unique viscosity solution [23] to the Riemannian eikonal equation: for all  $\mathbf{p} \in \Omega$

$$\|du(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} = 1 \quad \text{where } \mathcal{D}(\mathbf{p}) := \mathcal{M}(\mathbf{p})^{-1},$$

and  $u(\mathbf{p}) = 0$  for all  $\mathbf{p} \in \partial\Omega$ . Indeed, the dual to the Riemannian metric (13) reads  $\mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}}) = \|\hat{\mathbf{p}}\|_{\mathcal{D}(\mathbf{p})}$ . For each  $\mathbf{p} \in \bar{\Omega}$ , let  $(\rho_i(\mathbf{p}), \dot{\mathbf{e}}_i(\mathbf{p}))_{i=1}^{d'} \in (\mathbb{R}_+ \times \mathbb{L})^{d'}$  be weights and offsets such that

$$\mathcal{D}(\mathbf{p}) = \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \dot{\mathbf{e}}_i(\mathbf{p}) \otimes \dot{\mathbf{e}}_i(\mathbf{p}). \quad (14)$$

In this paper, we advocate the use of Voronoi's first reduction of quadratic forms for obtaining the decomposition (14), see Proposition 1.1. Our convergence result however only requires to control the maximal stencil radius

$$r_* := \max\{\|\dot{\mathbf{e}}_i(\mathbf{p})\|; \mathbf{p} \in \bar{\Omega}, 1 \leq i \leq d'\}. \quad (15)$$

If Proposition 1.1 is used for the stencil construction, then  $r_*$  is bounded in terms of the maximal condition number of the metric, and the number of terms in (14) is  $d' = d(d+1)/2$ . For the sake of readability, we omit in the rest of the paper to write the dependence of the offset  $\dot{\mathbf{e}}_i = \dot{\mathbf{e}}_i(\mathbf{p})$  on the point  $\mathbf{p} \in \bar{\Omega}$ . In the following, by  $\max\{0, a, b\}^2$  we mean  $(\max\{0, a, b\})^2$ .

**Theorem 1.3.** *Let  $\mathcal{M} \in C^0(\bar{\Omega}, S^{++}(\mathbb{E}))$  be a Riemannian metric, and for all  $\mathbf{p} \in \bar{\Omega}$  let  $\mathcal{D}(\mathbf{p}) := \mathcal{M}(\mathbf{p})^{-1}$  and  $(\rho_i(\mathbf{p}), \dot{\mathbf{e}}_i(\mathbf{p}))_{i=1}^{d'}$  be as in (14). Then for any  $h > 0$  there exists a unique solution  $U_h : h\mathbb{L} \rightarrow \mathbb{R}$  to the following discrete problem: for all  $\mathbf{p} \in \Omega_h$*

$$h^{-2} \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \max\{0, U_h(\mathbf{p}) - U_h(\mathbf{p} + h\dot{\mathbf{e}}_i), U_h(\mathbf{p}) - U_h(\mathbf{p} - h\dot{\mathbf{e}}_i)\}^2 = 1 \quad (16)$$

and  $U_h(\mathbf{p}) = 0$  for all  $\mathbf{p} \in \partial\Omega_h$ . The solution  $U_h$  can be computed via the fast-marching algorithm with complexity  $\mathcal{O}(d'N_h \ln N_h)$ , where  $N_h = \#(\Omega_h)$ . If in addition the domain  $\Omega$  satisfies an exterior cone condition, and if  $\mathcal{M} \in \text{Lip}(\bar{\Omega}, S^{++}(\mathbb{E}))$ , then for some constant  $C = C(\mathcal{M}, \Omega)$  one has for all  $h > 0$

$$\max_{\mathbf{p} \in \Omega_h} |U_h(\mathbf{p}) - u(\mathbf{p})| \leq C\sqrt{r_*h}. \quad (17)$$

This numerical scheme (16) is a direct generalization of the classical discretization [65, 71], which is however limited to isotropic metrics, locally proportional to the Euclidean norm.

The estimate (17) outlines the importance of the stencil radius  $r_*$ , since it determines the *effective* scale  $r_*h$  of the discretization and thus the accuracy of the numerical method. The construction of Proposition 1.1 is shown in [49] to minimize  $r_*$ , in dimension  $d = 2$ . A convergence rate similar to (17) is obtained in [73] for the Ordered Upwind Method [72], a semi-Lagrangian solver of anisotropic eikonal equations. Note that the dependency of the constant  $C = C(\Omega, \mathcal{M})$  in (17) with respect to the metric  $\mathcal{M}$  is not made explicit in Theorem 1.3. This point is analyzed in detail in the next sub-section, where we consider a family of increasingly anisotropic Riemannian metrics converging to a degenerate sub-Riemannian model.

**Remark 1.4.** *The numerical scheme (16) relies on upwind finite differences, which are first order consistent with the absolute value of a directional derivative: for any  $U \in C^2(\Omega)$ ,  $\mathbf{p} \in \Omega$ ,  $\dot{\mathbf{e}} \in \mathbb{E}$ , and any sufficiently small  $h > 0$*

$$\max\{0, U(\mathbf{p}) - U(\mathbf{p} - h\dot{\mathbf{e}}), U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}})\} / h = \max\{0, \langle dU(\mathbf{p}), \dot{\mathbf{e}} \rangle, -\langle dU(\mathbf{p}), \dot{\mathbf{e}} \rangle\} + \mathcal{O}(h), \\ = |\langle dU(\mathbf{p}), \dot{\mathbf{e}} \rangle| + \mathcal{O}(h).$$

The presence of “0” in the max, which may seem superfluous in view of the consistency analysis, is required for the degenerate ellipticity and the causality of the numerical scheme, see Definition 2.1. In the related literature [65, 71], the l.h.s. in the above equation is often written in the equivalent form  $\max\{\delta_e^- U(\mathbf{p}), \delta_e^+ U(\mathbf{p})\}$  where denoting  $a_{\pm} := \max\{0, \pm a\}$  one has

$$\delta_e^- U(\mathbf{p}) := \left( \frac{U(\mathbf{p}) - U(\mathbf{p} - h\dot{\mathbf{e}})}{h} \right)_+, \quad \delta_e^+ U(\mathbf{p}) := \left( \frac{U(\mathbf{p} + h\dot{\mathbf{e}}) - U(\mathbf{p})}{h} \right)_-.$$

## 1.2 Sub-Riemannian metrics

We introduce a numerical approach to the computation of sub-Riemannian distances and geodesics, based on solving the eikonal equations associated with a sequence of increasingly anisotropic approximate Riemannian metrics. This approach is related to [66], which however uses a different numerical scheme for the Riemannian problems, and does not establish a convergence rate. More precisely our results apply to the slightly more general class of pre-Riemannian models.

**Definition 1.5.** *A pre-Riemannian model on  $\Omega$  is a finite family of vector fields  $\dot{\omega}_1, \dots, \dot{\omega}_n \in \text{Lip}(\overline{\Omega}, \mathbb{E})$ . The control sets  $\mathcal{B} \in \text{Lip}(\overline{\Omega}, \mathfrak{C}(\mathbb{E}))$ , and the semi-definite tensor field  $\mathcal{D} \in \text{Lip}(\overline{\Omega}, \mathbb{S}^+(\mathbb{E}^*))$ , for this model are defined for all  $\mathbf{p} \in \overline{\Omega}$  by*

$$\mathcal{B}(\mathbf{p}) := \left\{ \sum_{1 \leq i \leq n} \alpha_i \dot{\omega}_i(\mathbf{p}); \alpha \in \mathbb{R}^n, \sum_{1 \leq i \leq n} \alpha_i^2 \leq 1 \right\}, \quad \mathcal{D}(\mathbf{p}) := \sum_{1 \leq i \leq n} \dot{\omega}_i(\mathbf{p}) \otimes \dot{\omega}_i(\mathbf{p}).$$

A sub-Riemannian model [51] of step  $k \geq 1$  is a pre-Riemannian model with the additional properties that the vector fields  $(\dot{\omega}_i)_{1 \leq i \leq n}$  are smooth and that, together with their iterated commutators up to depth  $k$ , they span the tangent space  $\mathbb{E}$  at each point  $\mathbf{p} \in \overline{\Omega}$ . The minimal control time  $T_{\mathcal{B}}(\mathbf{p}, \mathbf{q})$  for a sub-Riemannian model is called the Carnot-Theodory distance, and by Chow’s theorem it obeys  $T_{\mathcal{B}}(\mathbf{p}, \mathbf{q}) \leq C \|\mathbf{p} - \mathbf{q}\|^{\frac{1}{k}}$ , as  $\mathbf{q} \rightarrow \mathbf{p} \in \Omega$ . The distance  $u$  to  $\partial\Omega$  is the unique viscosity solution to the sub-Riemannian eikonal equation:  $\|du(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} = 1$  for all  $\mathbf{p} \in \Omega$ , and  $u = 0$  on  $\partial\Omega$ .

For better or worse, we do not use any techniques or results from sub-Riemannian geometry in this paper but stick instead to the simpler pre-Riemannian concept. We do however make a further assumption.

**Assumption 1.6.** We fix a pre-Riemannian model  $(\dot{\omega}_i)_{i=1}^n$ , and assume that the exit time value function  $u$ , defined in (10), is bounded on  $\Omega$ . We further assume that the domain admits outward normals  $\dot{\mathbf{n}}(\mathbf{p})$  with Lipschitz regularity on  $\partial\Omega$ , and that for each  $\mathbf{p} \in \partial\Omega$  there exists  $1 \leq i \leq n$  such that  $\dot{\mathbf{n}}(\mathbf{p}) \cdot \dot{\omega}_i(\mathbf{p}) \neq 0$ .

The finiteness of  $u$  on  $\Omega$  is a *global controllability* assumption, and it is obviously required if one intends to prove convergence rates of discrete approximations of  $u$ . The second assumption is related to *short time local controllability at the boundary* [3]. Together, these assumptions imply the Lipschitz regularity of  $u$ , see §D.1.

**Definition 1.7.** A completion of a pre-Riemannian model  $(\dot{\omega}_i)_{i=1}^n$  is a second finite family of vector fields  $\dot{\omega}_1^*, \dots, \dot{\omega}_{n^*}^* \in \text{Lip}(\bar{\Omega}, \mathbb{E})$ , such that  $\dot{\omega}_1(\mathbf{p}), \dots, \dot{\omega}_n(\mathbf{p}), \dot{\omega}_1^*(\mathbf{p}), \dots, \dot{\omega}_{n^*}^*(\mathbf{p})$  spans  $\mathbb{E}$  for each  $\mathbf{p} \in \bar{\Omega}$ .

For each  $0 < \varepsilon \leq 1$  the augmented pre-Riemannian model  $(\omega_1, \dots, \omega_n, \varepsilon\omega_1^*, \dots, \varepsilon\omega_{n^*}^*)$  is equivalent (i.e. has the same control sets) to the Riemannian model of metric  $\mathcal{M}_\varepsilon := \mathcal{D}_\varepsilon^{-1}$ , where pointwise on  $\bar{\Omega}$

$$\mathcal{D}_\varepsilon = \mathcal{D} + \varepsilon^2 \mathcal{D}^*, \quad \text{with } \mathcal{D}^* := \sum_{1 \leq i \leq n^*} \dot{\omega}_i^* \otimes \dot{\omega}_i^*. \quad (18)$$

In order to solve numerically the pre-Riemannian exit time problem, our strategy is to apply the scheme of Theorem 1.3 to the positive definite (but strongly anisotropic) Riemannian metric  $\mathcal{M}_\varepsilon$ , for small  $\varepsilon > 0$ . Convergence towards the pre-Riemannian exit times  $u : \Omega \rightarrow \mathbb{R}$  is established in the next theorem, when the relaxation parameter  $\varepsilon$  and grid scale  $h$  tend to 0 suitably.

**Theorem 1.8.** Consider a pre-Riemannian model  $\dot{\omega}_1, \dots, \dot{\omega}_n \in \text{Lip}(\bar{\Omega}, \mathbb{E})$  obeying Assumption 1.6, and a completion  $\dot{\omega}_1^*, \dots, \dot{\omega}_{n^*}^*$ . For each  $0 < \varepsilon \leq 1$  let  $u_\varepsilon$  denote the distance to  $\partial\Omega$  for the Riemannian metric  $\mathcal{M}_\varepsilon$ , and let  $U_{h,\varepsilon}$  be the discrete solution of (16) with scale  $h > 0$ . Then

$$\max_{\bar{\Omega}} |u - u_\varepsilon| \leq C\varepsilon, \quad \max_{\Omega_h} |u_\varepsilon - U_{h,\varepsilon}| \leq C' \sqrt{r_\varepsilon h}, \quad (19)$$

where  $r_\varepsilon$  denotes the maximal stencil radius for  $\mathcal{M}_\varepsilon$ , see (15), and where  $C, C'$  only depend on  $\Omega$ ,  $(\dot{\omega}_i)_{i=1}^n$ , and  $(\dot{\omega}_i^*)_{i=1}^{n^*}$ . In particular  $U_{h,\varepsilon} \rightarrow u$  uniformly as  $\varepsilon \rightarrow 0$  and  $h r_\varepsilon \rightarrow 0$ .

By construction the anisotropy ratio of the tensors  $\mathcal{M}_\varepsilon$  is  $\mathcal{O}(\varepsilon^{-1})$ , hence  $r_\varepsilon \leq C\varepsilon^{-(d-1)}$  if Proposition 1.1 is used for the stencil construction. The convergence rate  $\max_{\Omega_h} |U_{h,\varepsilon} - u| \leq Ch^{\frac{1}{d+1}}$  is thus ensured by choosing  $\varepsilon = h^{\frac{1}{d+1}}$ .

### 1.3 Rander geometry

Rander metrics are asymmetric metrics<sup>5</sup>, defined as the sum of a symmetric Riemannian part and of an anti-symmetric linear part [61]. A Rander metric is thus described by a tensor field  $\mathcal{M} \in C^0(\bar{\Omega}, \text{S}^{++}(\mathbb{E}))$ , and a co-vector field  $\hat{\eta} \in C^0(\bar{\Omega}, \mathbb{E}^*)$ , subject to a compatibility condition:

$$\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) := \|\dot{\mathbf{p}}\|_{\mathcal{M}(\mathbf{p})} + \langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{p}} \rangle, \quad \text{where } \|\hat{\eta}(\mathbf{p})\|_{\mathcal{M}(\mathbf{p})^{-1}} < 1. \quad (20)$$

The smallness constraint (20, right) ensures the positivity of the asymmetric norm  $\mathcal{F}_{\mathbf{p}}(\cdot)$ . The distance induced by a Rander metric is oriented:  $d_{\mathcal{F}}(\mathbf{p}, \mathbf{q}) \neq d_{\mathcal{F}}(\mathbf{q}, \mathbf{p})$  in general.

<sup>5</sup>A metric lacking the symmetry property, i.e.  $\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) \neq \mathcal{F}_{\mathbf{p}}(-\dot{\mathbf{p}})$  for some point  $\mathbf{p} \in \bar{\Omega}$  and vector  $\dot{\mathbf{p}} \in \mathbb{E}$ , is usually referred to as a quasi-metric. The *quasi*- prefix is however dropped in this paper, for consistency with the Riemannian and sub-Riemannian cases.

**Proposition 1.9.** *The distance  $u$  to  $\partial\Omega$ , see (10), is the unique viscosity solution to the inhomogeneous static first order HJB PDE*

$$\|du(\mathbf{p}) - \hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} = 1, \quad \text{where } \mathcal{D}(\mathbf{p}) = \mathcal{M}(\mathbf{p})^{-1}, \quad (21)$$

for all  $\mathbf{p} \in \Omega$ , and  $u = 0$  on  $\partial\Omega$ .

*Proof.* It is known that  $u$  obeys the eikonal PDE  $\mathcal{F}_{\mathbf{p}}^*(du(\mathbf{p})) = 1$ , where  $\mathcal{F}^*$  is the dual metric, see (11). Now for any  $\hat{\mathbf{p}} \in \mathbb{E}^*$  observe the sequence of equivalences:

$$\begin{aligned} \mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}}) = 1 &\Leftrightarrow \exists \dot{\mathbf{p}} \in \mathbb{E} \setminus \{0\}, \hat{\mathbf{p}} = d\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) \\ &\Leftrightarrow \exists \dot{\mathbf{p}} \in \mathbb{E} \setminus \{0\}, \hat{\mathbf{p}} = \mathcal{M}(\mathbf{p})\dot{\mathbf{p}}/\|\dot{\mathbf{p}}\|_{\mathcal{M}(\mathbf{p})} + \hat{\eta}(\mathbf{p}) \\ &\Leftrightarrow \|\hat{\mathbf{p}} - \hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} = 1. \end{aligned}$$

The first equivalence follows from convex duality  $\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) = \sup\{\langle \dot{\mathbf{p}}, \dot{\mathbf{p}} \rangle; \mathcal{F}_{\mathbf{p}}^*(\dot{\mathbf{p}}) = 1\}$  and the envelope theorem, and the second one from the explicit expression (20) of  $\mathcal{F}$ .  $\square$

**Theorem 1.10.** *Let  $(\mathcal{M}, \hat{\eta})$  be a Rander metric, and for all  $\mathbf{p} \in \bar{\Omega}$  let  $\mathcal{D}(\mathbf{p}) := \mathcal{M}(\mathbf{p})^{-1}$  and  $(\rho_i(\mathbf{p}), \dot{\mathbf{e}}_i(\mathbf{p}))_{i=1}^{d'}$  be as in (14). Then for any  $h > 0$  there exists a unique solution  $U_h : h\mathbb{L} \rightarrow \mathbb{R}$  to the following discrete problem: for all  $\mathbf{p} \in \Omega_h$*

$$\sum \rho_i(\mathbf{p}) \max\{0, U_h(\mathbf{p}) - U_h(\mathbf{p} + h\dot{\mathbf{e}}_i) + h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle, U_h(\mathbf{p}) - U_h(\mathbf{p} - h\dot{\mathbf{e}}_i) - h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle\}^2 = h^2, \quad (22)$$

and  $U_h(\mathbf{p}) = 0$  for all  $\mathbf{p} \in \partial\Omega_h$ . If in addition  $\Omega$  obeys an exterior cone condition, and  $\mathcal{M}$  and  $\eta$  have Lipschitz regularity, then for some  $C = C(\Omega, \mathcal{M}, \hat{\eta})$  one has for all  $h > 0$

$$\max_{\mathbf{p} \in \Omega_h} |U_h(\mathbf{p}) - u(\mathbf{p})| \leq C\sqrt{r_*h}.$$

The numerical scheme (22) is degenerate elliptic, see Definition 2.1, and can therefore be solved using an iterative method such as fast sweeping [85] or the AGSI [10], which is used in our numerical experiments §3. However, it *cannot be solved in a single pass* using the Fast-Marching algorithm, in contrast with the Riemannian (16) and sub-Riemannian discretizations, because the expression (22) is not a function of the positive parts only of the finite differences  $U(\mathbf{p}) - U(\mathbf{p} + \dot{\mathbf{e}})$  when  $\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}} \rangle > 0$ , in contradiction with Definition 2.1 of causality. Note that, in dimension  $d = 2$ , Rander distances can be computed in a single pass using an entirely different semi-Lagrangian discretization [48].

## 2 Convergence in the Riemannian case

This section is devoted to the proof of Theorem 1.3, which contains two parts: a claim of well-posedness for the system of equations discretizing the Riemannian eikonal PDE, and an error analysis as the grid scale is refined. For that purpose, two general and classical results are stated in §2.1, and later specialized in §2.2 for the model of interest.

### 2.1 Two general results

We formally introduce the concepts of degenerate elliptic and causal finite difference schemes, and present (reformulations of) two classical results. Theorem 2.3 states that degenerate elliptic schemes possess a unique solution, in the spirit of [78, 71, 54] and under adequate assumptions,

which can be computed in a single pass under the additional assumption of causality. Theorem 2.4 introduces a strategy for the numerical analysis, referred to as the doubling of variables argument and adapted from [30].

**Definition 2.1.** A (finite differences) scheme on a finite set  $X$  is a continuous map  $\mathfrak{F} : \mathbb{R}^X \rightarrow \mathbb{R}^X$ , written as

$$(\mathfrak{F}U)(\mathbf{x}) := \mathfrak{F}(\mathbf{x}, U(\mathbf{x}), (U(\mathbf{x}) - U(\mathbf{y}))_{\mathbf{y} \neq \mathbf{x}}).$$

The scheme is said:

- Degenerate elliptic iff  $\mathfrak{F}$  is non-decreasing w.r.t. the second and (each of) the third variables.
- Causal iff  $\mathfrak{F}$  only depends on the positive part of (each of) the third variables.

A discrete map  $U \in \mathbb{R}^X$  is called a sub- (resp. strict sub-, resp. super-, resp. strict super-) solution of scheme  $\mathfrak{F}$  iff  $\mathfrak{F}U \leq 0$  (resp.  $\mathfrak{F}U < 0$ , resp.  $\mathfrak{F}U \geq 0$ , resp.  $\mathfrak{F}U > 0$ ) pointwise on  $X$ .

When the scheme  $\mathfrak{F}$  is obvious from context, we simply speak of a sub- and super-solution. The numerical schemes (16) and (22), discretizing Riemannian and Rander eikonal PDEs respectively, are both degenerate elliptic, and the Riemannian one (16) is also causal.

**Remark 2.2** (Causality). In this paper, following [72, 81, 1] we define causality as the property ensuring that a solution of the scheme  $F$  can be computed in a single pass over the domain, using the fast marching algorithm, see Theorem 2.3. For the Gauss-Siedel operator  $\Lambda$  associated with the scheme, this property becomes

$$\Lambda U(x) \text{ may depend on } U(y) \text{ only if } \Lambda U(x) > U(y), \quad (23)$$

see Definition A.1 and Proposition A.4 in §A. This notion of causality is also defined in §3 of [1], in equation (5) of [81], and in Properties 3.6 and A.1 of [72]. It is a key ingredient, although not given a name, in [78] (see Lemma 3.1) and in [41].

In the literature related to fast sweeping methods, causality is often given an entirely different (non-equivalent) meaning, related with upwindness and presented as equivalent to stability [39]. In Definition 2.1, the property related with upwindness is Degenerate ellipticity, not Causality. Degenerate ellipticity also implies stability through the discrete comparison principle, see Theorem 2.3. Definition 2.1 of Causality is in contrast irrelevant to upwindness or stability, and Causality is not used in the proof of convergence (which holds for the Rander scheme lacking this property), but only to show in §A that the discretized PDE can be solved in a single pass.

**Theorem 2.3.** Let  $\mathfrak{F}$  be a degenerate elliptic scheme on a finite set  $X$  s.t.

- (i) There exists a sub-solution  $U^-$  and a super-solution  $U^+$  to the scheme  $\mathfrak{F}$ .
- (ii) Any super-solution to  $\mathfrak{F}$  is the limit of a sequence of strict super-solutions.

Then there exists a unique solution  $U \in \mathbb{R}^X$  to  $\mathfrak{F}U = 0$ , and it satisfies  $U^- \leq U \leq U^+$ . If in addition the scheme is causal, then this solution can be computed in a single pass over the domain using the Fast-Marching algorithm, with complexity  $\mathcal{O}(M \ln N)$  where

$$N = \#(X), \quad M = \#(\{(x, y) \in X \times X; \mathfrak{F}U(\mathbf{x}) \text{ depends on } U(\mathbf{y})\}). \quad (24)$$

*Proof.* We provide for completeness the proof of existence and uniqueness, see [54] and [4] for closely related arguments in the discrete and continuous settings respectively. The description of the fast marching algorithm and the related claims are postponed to §A.

*Proof of uniqueness, via the comparison principle.* Let  $U^+$  be a strict super-solution, and  $U^-$  a sub-solution. Let  $\mathbf{p} \in X$  be such that  $U^-(\mathbf{p}) - U^+(\mathbf{p})$  is maximal, so that  $U^-(\mathbf{p}) - U^-(\mathbf{q}) \geq U^+(\mathbf{p}) - U^+(\mathbf{q})$  for any  $\mathbf{q} \in X$ . Assuming for contradiction that  $U^-(\mathbf{p}) \geq U^+(\mathbf{p})$  we obtain  $0 \geq \mathfrak{F}U^-(\mathbf{p}) \geq \mathfrak{F}U^+(\mathbf{p}) > 0$  by degenerate ellipticity of the scheme and by definition of a sub- and strict super-solution. This is a contradiction, hence  $U^- \leq U^+$ . Next using assumption (ii) we obtain that  $U^- \leq U^+$  still holds for any sub-solution  $U^-$  and any (possibly non-strict) super-solution  $U^+$ . The uniqueness of the solution to  $\mathfrak{F}U = 0$  follows.

*Proof of existence, by Perron's method.* We prove that  $U : X \rightarrow \mathbb{R}$ , defined by  $U(\mathbf{p}) := \sup\{\tilde{U}(\mathbf{p}); \tilde{U} \text{ sub-solution}\}$  for all  $\mathbf{p} \in X$ , is a solution to the scheme  $\mathfrak{F}$ . By the previous argument one has  $\tilde{U} \leq U^+$  for any sub-solution  $\tilde{U}$ , and up to considering the sub-solution  $\max\{\tilde{U}, U^-\}$  we may as well assume  $\tilde{U} \geq U^-$ . Thus  $U^- \leq \tilde{U} \leq U^+$  and therefore  $U^- \leq U \leq U^+$  by taking the pointwise supremum. Consider an arbitrary  $\mathbf{p} \in X$ , and let  $\tilde{U}$  be a sub-solution such that  $U(\mathbf{p}) = \tilde{U}(\mathbf{p})$ , which exists by continuity of  $\mathfrak{F}$  and a compactness argument. By construction  $U \geq \tilde{U}$ , hence  $\mathfrak{F}U(\mathbf{p}) \leq \mathfrak{F}\tilde{U}(\mathbf{p}) \leq 0$  by monotony of the scheme, hence  $U$  is a sub-solution by arbitrariness of  $\mathbf{p} \in X$ . Furthermore, assume for contradiction that there exists  $\mathbf{p}_0 \in X$  such that  $\mathfrak{F}U(\mathbf{p}_0) < 0$ , and define  $U_\varepsilon(\mathbf{p}_0) := U(\mathbf{p}_0) + \varepsilon$  and  $U_\varepsilon(\mathbf{p}) := U(\mathbf{p})$  for all  $\mathbf{p} \in X \setminus \{\mathbf{p}_0\}$ . Then  $U_\varepsilon$  is a sub-solution for any sufficiently small  $\varepsilon > 0$ , by monotony and continuity of the scheme  $\mathfrak{F}$ , thus  $U(\mathbf{p}_0) \geq U_\varepsilon(\mathbf{p}_0)$  by construction which is a contradiction. Finally we obtain  $\mathfrak{F}U = 0$  identically on  $X$ , as announced.  $\square$

The following result is a general strategy for proving convergence rates for discretizations of first order HJB PDEs, adapted from [30]. For completeness, the proof is presented in §F. The Cartesian grid  $h\mathbb{L}$  could be replaced with an arbitrary  $h$ -net of  $E$ , in other words a discrete set such that union of all balls of radius  $h$  centered at the points of this set covers  $\mathbb{E}$ .

**Theorem 2.4** (Doubling of variables argument). *Let  $u : \mathbb{E} \rightarrow \mathbb{R}$  be supported on a bounded domain  $\bar{\Omega}$  and  $C_{\text{Lip}}^u$ -Lipschitz, and let  $U_h : h\mathbb{L} \rightarrow \mathbb{R}$  be supported on  $\Omega_h := \Omega \cap h\mathbb{L}$ . Given  $\lambda \in [1/2, 1[$  and  $\delta > 0$ , define*

$$\overline{M}_{\lambda,\delta} := \sup_{(\mathbf{p},\mathbf{q}) \in h\mathbb{L} \times \mathbb{E}} \lambda U_h(\mathbf{p}) - u(\mathbf{q}) - \frac{1}{2\delta} \|\mathbf{p} - \mathbf{q}\|^2, \quad \widetilde{M}_{\lambda,\delta} := \sup_{(\mathbf{p},\mathbf{q}) \in h\mathbb{L} \times \mathbb{E}} \lambda u(\mathbf{q}) - U_h(\mathbf{p}) - \frac{1}{2\delta} \|\mathbf{p} - \mathbf{q}\|^2, \Delta$$

and denote by  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in (h\mathbb{L}) \times \mathbb{E}$  the point pairs where the maxima are respectively attained. Then  $\max\{\|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|, \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\|\} \leq 4C_{\text{Lip}}^u \delta$ . Assume furthermore that for some  $C_{\text{bd}}^U, C_{\text{bd}}^U$  and  $c_{\text{bd}}^U \geq 4C_{\text{Lip}}^u \delta$  the following holds:

(i) None of the two maximal pairs  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$  and  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$  belongs to  $\Omega_h \times \Omega$ .

(ii)  $|U_h(\mathbf{p})| \leq C_{\text{bd}}^U d_{\partial\Omega}(\mathbf{p}) + C_{\text{bd}}^U h$ , for all  $\mathbf{p} \in \Omega_h$  such that  $d_{\partial\Omega}(\mathbf{p}) \leq c_{\text{bd}}$ .

Then one has with  $C_0 = 4C_{\text{Lip}}^u \max\{C_{\text{Lip}}^u, C_{\text{bd}}^U\}$

$$\max_{\mathbf{p} \in \Omega_h} |u(\mathbf{p}) - U_h(\mathbf{p})| \leq 2 \left( C_0 \delta + C_{\text{bd}}^U h + (1 - \lambda) \max_{\Omega} |u| \right). \quad (25)$$

When applying Theorem 2.4, to one of our specific models, Property (i) follows from the consistency and the degenerate ellipticity of the discretization, while Property (ii) requires to establish a discrete counterpart of short-time local controllability at the boundary. In the Riemannian case, Property (i) is established in Lemma 2.8, and Property (ii) in Proposition 2.11. Some

adaptations of these arguments are required to establish property (ii) in the sub-Riemannian case, see Proposition D.7, and property (i) in the Rander case, see Lemma E.3.

Explicit expressions of the constants  $C_{\text{Lip}}^u$ ,  $C_{\text{bd}}^U$ ,  $c_{\text{bd}}^U$ , are provided in terms of the model parameters, namely  $(\Omega, \mathcal{M})$  in the Riemannian case. The constant  $C_{\text{bd}}^U$ , also given explicitly, depends linearly on the stencil maximal radius:  $C_{\text{bd}}^U = C_{\text{bd}}^U r_*$ , and property (i) is shown to hold provided

$$\lambda \leq 1 - C_1 \delta - C_2 r_* \frac{h}{\delta},$$

where  $C_1, C_2$  are again explicit constants depending only on the model parameters. Choosing  $\lambda$  equal to this upper bound, and defining  $\delta = \sqrt{r_* h}$ , one gets the error estimate

$$\max_{\mathbf{p} \in \Omega_h} |u(\mathbf{p}) - U_h(\mathbf{p})| \leq 2C_{\text{bd}}^U r_* h + 2(C_0 + (C_1 + C_2)\|u\|_\infty)\sqrt{r_* h}. \quad (26)$$

as announced in Theorems 1.3, 1.8 and 1.10 for Riemannian, sub-Riemannian, and Rander metrics respectively.

## 2.2 Application to the Riemannian case

We establish Theorem 1.3, on the discretization of Riemannian exit time problems, by specializing the general results of §2.1. For that purpose we consider a discretization scheme  $\mathfrak{F}_h$ , on the finite domain  $\Omega_h$ , see (3), of the following form: for any  $U : \Omega_h \rightarrow \mathbb{R}$  and any  $\mathbf{p} \in \Omega_h$

$$(\mathfrak{F}_h U(\mathbf{p}))^2 := h^{-2} \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \max\{0, U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}}_i), U(\mathbf{p}) - U(\mathbf{p} - h\dot{\mathbf{e}}_i)\}^2, \quad (27)$$

where  $U$  is extended by zero on  $h\mathbb{L} \setminus \Omega_h$ . The next proposition implies, by Theorem 2.3, the existence and uniqueness of a solution to the equation  $\mathfrak{F}_h U - 1 \equiv 0$ , and the applicability of the Fast-Marching algorithm to compute it, as announced in Theorem 1.3.

Recall that  $r_*$  is the maximal stencil radius, as defined in (15). The square root of the largest eigenvalue among all tensors of a tensor field  $\mathcal{M} \in C^0(\bar{\Omega}, S^{++}(\mathbb{E}))$  is denoted by

$$\lambda^*(\mathcal{M}) := \max_{\mathbf{p} \in \bar{\Omega}, \|\dot{\mathbf{p}}\|=1} \|\dot{\mathbf{p}}\|_{\mathcal{M}(\mathbf{p})}. \quad (28)$$

**Proposition 2.5.** *Let  $\mathcal{M} \in C^0(\bar{\Omega}, S^{++}(\mathbb{E}))$  be a Riemannian metric, and for all  $\mathbf{p} \in \bar{\Omega}$  let  $\mathcal{D}(\mathbf{p}) := \mathcal{M}(\mathbf{p})^{-1}$  and  $(\rho_i(\mathbf{p}), \dot{\mathbf{e}}_i(\mathbf{p}))_{i=1}^{d'}$  be as in (14). Then the scheme  $\mathfrak{F}_h$  defined by (27) is degenerate elliptic and causal. In addition:*

- (i) *The null map  $U = 0$  satisfies  $\mathfrak{F}_h U \equiv 0$ , hence is a sub-solution to the scheme  $\mathfrak{F}_h - 1$ .*
- (ii) *Let  $R > 0$  be such that  $\Omega$  is contained in the ball of radius  $R - hr_*$  and centered at the origin, and let  $U(\mathbf{p}) := R - \|\mathbf{p}\|$ , for all  $\mathbf{p} \in \Omega_h$ . Then for all  $\lambda \geq 0$ , and all  $\mathbf{p} \in \Omega_h$*

$$\mathfrak{F}_h U(\mathbf{p}) \geq \|\mathbf{p}/\|\mathbf{p}\|\|_{\mathcal{D}(\mathbf{p})},$$

*where  $\mathbf{p}/\|\mathbf{p}\|$  can be replaced with an arbitrary unit vector in the case  $\mathbf{p} = 0$ . As a result,  $\lambda U$  is a super-solution to the scheme  $\mathfrak{F}_h - 1$  for any  $\lambda \geq \lambda^*(\mathcal{M})$ .*

- (iii) *If  $U$  is a super-solution to  $\mathfrak{F}_h - 1$ , then  $(1 + \varepsilon)U$  is a strict super-solution for any  $\varepsilon > 0$ .*

*Proof.* The monotony and causality of the scheme  $\mathfrak{F}_h$  immediately follow from its expression (27). Point (i) is trivial, and point (iii) follows from the 1-homogeneity property  $\mathfrak{F}_h(\lambda U) = \lambda \mathfrak{F}_h U$ . In the rest of this proof, the point  $\mathbf{p} \in \Omega$  is regarded both as a vector in  $\mathbb{E}$  and as a co-vector in  $\mathbb{E}^*$ , thanks to the Euclidean structure of  $\mathbb{E}$ . For point (ii), we obtain by the convexity of the Euclidean norm, for any  $\mathbf{p}, \dot{\mathbf{e}} \in \mathbb{E}$ .

$$U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}}) = \|\mathbf{p} + h\dot{\mathbf{e}}\| - \|\mathbf{p}\| \geq h\langle \mathbf{p}/\|\mathbf{p}\|, \dot{\mathbf{e}} \rangle,$$

where  $\mathbf{p}/\|\mathbf{p}\|$  can be replaced with any unit vector if  $\mathbf{p} = 0$ . Hence for all  $\mathbf{p} \in \Omega_h$ , as announced,

$$(\mathfrak{F}_h U(\mathbf{p}))^2 \geq \frac{1}{\|\mathbf{p}\|^2} \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \langle \mathbf{p}, \dot{\mathbf{e}}_i \rangle^2 = \frac{\|\mathbf{p}\|_{\mathcal{D}(\mathbf{p})}^2}{\|\mathbf{p}\|^2}.$$

Finally  $\mathfrak{F}_h(\lambda U)(\mathbf{p}) \geq 1$  if  $\lambda \geq \lambda^*(\mathcal{M})$  by the 1-homogeneity of  $\mathfrak{F}_h$ , and the observation that the least eigenvalue of  $\mathcal{D}(\mathbf{p})$  is inverse of the largest eigenvalue of  $\mathcal{M}(\mathbf{p})$ .  $\square$

In the rest of this section, we establish the properties required to apply the doubling of variables argument, Theorem 2.4, to prove the second part of Theorem 1.3. The following proposition immediately implies that the exit time value function, denoted hereafter by  $u$ , is  $C_{\text{Lip}}^u$ -Lipschitz, with  $C_{\text{Lip}}^u := \lambda^*(\mathcal{M})$ .

**Proposition 2.6.** *Let  $\Omega \subseteq \mathbb{E}$  be an arbitrary bounded domain, equipped with a metric  $\mathcal{F} : \overline{\Omega} \times \mathbb{E} \rightarrow \mathbb{R}$  such that  $\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) \leq C_0 \|\dot{\mathbf{p}}\|$  for any  $\mathbf{p} \in \overline{\Omega}$ ,  $\dot{\mathbf{p}} \in \mathbb{E}$ . Then the distance  $u$  from  $\partial\Omega$  is  $C_0$ -Lipschitz.*

*Proof.* Let  $\mathbf{p}, \mathbf{q} \in \overline{\Omega}$ , and let us prove that  $u(\mathbf{q}) \leq u(\mathbf{p}) + C_0 \|\mathbf{p} - \mathbf{q}\|$ . Let  $\gamma : [0, 1] \rightarrow \mathbb{E}$  be the parametrization of the line segment  $[\mathbf{p}, \mathbf{q}]$  at constant speed. If this segment intersects  $\partial\Omega$ , then denoting  $T \in [0, 1]$  the largest time such that  $\gamma(T) \in \partial\Omega$  one has  $u(\mathbf{q}) \leq \text{Length}_{\mathcal{F}}(\gamma|_{[T, 1]}) \leq C_0 \|\mathbf{p} - \mathbf{q}\|$  as announced. Otherwise, denoting by  $\tilde{\gamma}$  a minimal path from  $\partial\Omega$  to  $\mathbf{p}$  one has by path concatenation  $u(\mathbf{q}) \leq \text{length}_{\mathcal{F}}(\tilde{\gamma}) + \text{length}_{\mathcal{F}}(\gamma) \leq u(\mathbf{p}) + C_0 \|\mathbf{p} - \mathbf{q}\|$  as announced.  $\square$

The rest of this section is split into two parts, devoted to proving assumptions (i) and (ii) of Theorem 2.4 (the doubling of variables argument) for the Riemannian discretization scheme, thus by (26) concluding the proof of Theorem 1.3. For that purpose, we adopt the notations and other assumptions of Theorems 1.3 and 2.4, and we denote by  $U_h$  the solution to the scheme  $\mathfrak{F}_h - 1$ . In particular  $\lambda \in [1/2, 1]$  and  $\delta > 0$  are parameters from Theorem 2.4, and  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in \mathbb{E} \times h\mathbb{L}$  are points pairs where the maxima  $\bar{M}_{\lambda, \delta}, \tilde{M}_{\lambda, \delta}$  are attained.

**Establishing assumption (i) of Theorem 2.4.** The proof consists of two lemmas, Lemma 2.7 and Lemma 2.8, which respectively rely on the degenerate ellipticity and the consistency of the proposed numerical scheme  $\mathfrak{F}_h$ .

**Lemma 2.7.** *Let  $\bar{\mathbf{w}} := (\bar{\mathbf{p}} - \bar{\mathbf{q}})/\delta$ , and let  $\bar{U}(\mathbf{p}) := \langle \bar{\mathbf{w}}, \mathbf{p} \rangle + \frac{1}{2\delta} \|\mathbf{p} - \bar{\mathbf{p}}\|^2$  for all  $\mathbf{p} \in h\mathbb{L}$ . Then*

$$\mathfrak{F}_h \bar{U}(\bar{\mathbf{p}}) \leq \lambda, \quad \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{q}})} \geq 1. \quad (29)$$

*Let  $\tilde{\mathbf{w}} := (\tilde{\mathbf{p}} - \tilde{\mathbf{q}})/\delta$ , and let  $\tilde{U}(\mathbf{p}) := \langle \tilde{\mathbf{w}}, \mathbf{p} \rangle - \frac{1}{2\delta} \|\mathbf{p} - \tilde{\mathbf{p}}\|^2$  for all  $\mathbf{p} \in h\mathbb{L}$ . Then*

$$\mathfrak{F}_h \tilde{U}(\tilde{\mathbf{p}}) \geq 1, \quad \|\tilde{\mathbf{w}}\|_{\mathcal{D}(\tilde{\mathbf{q}})} \leq \lambda. \quad (30)$$

Here and below we regard  $\bar{\mathbf{w}}$  and  $\tilde{\mathbf{w}}$  as co-vectors, using the Euclidean structure of  $\mathbb{E}$ .

*Proof.* Note that the scheme  $\mathfrak{F}_h$  is here (slightly abusively) applied to the functions  $\bar{U}$ ,  $\tilde{U}$ , which are non-zero over  $h\mathbb{L} \setminus \Omega_h$ . We focus on the proof of (29), the case of (30) being similar. By definition of  $\bar{M}_{\lambda,\delta}$ , the function

$$\mathbf{p} \in h\mathbb{L} \mapsto \lambda U_h(\mathbf{p}) - u(\bar{\mathbf{q}}) - \frac{1}{2\delta} \|\mathbf{p} - \bar{\mathbf{q}}\|^2 = \lambda U_h(\mathbf{p}) - \bar{U}(\mathbf{p}) - K$$

attains its maximum at  $\bar{\mathbf{p}}$ , where  $K = \bar{u}(\bar{\mathbf{q}}) - \frac{1}{2} \langle \bar{\mathbf{w}}, \bar{\mathbf{p}} + \bar{\mathbf{q}} \rangle$  is independent of the variable  $\mathbf{p}$ . Hence for all  $\mathbf{p} \in h\mathbb{L}$

$$\lambda U_h(\bar{\mathbf{p}}) - \bar{U}(\bar{\mathbf{p}}) - K \geq \lambda U_h(\mathbf{p}) - \bar{U}(\mathbf{p}) - K, \quad \text{equivalently } U_h(\bar{\mathbf{p}}) - U_h(\mathbf{p}) \geq U(\bar{\mathbf{p}})/\lambda - \bar{U}(\mathbf{p})/\lambda.$$

By degenerate ellipticity of the scheme  $\mathfrak{F}_h$ , see Definition 2.1, we obtain  $\mathfrak{F}_h(\bar{U}/\lambda)(\bar{\mathbf{p}}) \leq \mathfrak{F}_h U_h(\bar{\mathbf{q}}) = 1$ , hence (29, left) by the homogeneity of  $\mathfrak{F}_h$ . Likewise, defining  $\bar{u}(\mathbf{q}) := \langle \bar{\mathbf{w}}, \mathbf{q} \rangle - \frac{1}{2\delta} \|\mathbf{q} - \bar{\mathbf{q}}\|^2$  for all  $\mathbf{q} \in \mathbb{E}$ , the function

$$\mathbf{q} \in \mathbb{E} \mapsto \lambda U_h(\bar{\mathbf{p}}) - u(\mathbf{q}) - \frac{1}{2\delta} \|\mathbf{p} - \bar{\mathbf{q}}\|^2 = \bar{u}(\mathbf{q}) - u(\mathbf{q}) - K'$$

attains its minimum at  $\bar{\mathbf{q}}$ , where  $K'$  is the adequate constant. Since  $u$  is by definition the viscosity solution of the PDE (1), hence also a super-solution, this implies  $1 \leq \|\mathrm{d}\bar{u}(\bar{\mathbf{q}})\|_{\mathcal{D}(\bar{\mathbf{q}})} = \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{q}})}$ , which concludes the proof.  $\square$

The following lemma assumes *for contradiction* that  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in \Omega_h \times \Omega$  and obtains estimates *contradicting* Lemma 2.7 established above, provided  $\lambda$  is above a certain bound (which is assumed). Therefore, arguing by contradiction, one must have  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \notin \Omega_h \times \Omega$ , and likewise  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \notin \Omega_h \times \Omega$  by a similar argument, which establishes assumption (i) of Theorem 2.4. Let  $C_{\text{Lip}}^{\mathcal{D}}$  be a constant such that for all  $\mathbf{p}, \mathbf{q} \in \bar{\Omega}$  and all  $\hat{\mathbf{p}} \in \mathbb{E}^*$

$$|\|\hat{\mathbf{p}}\|_{\mathcal{D}(\mathbf{p})} - \|\hat{\mathbf{p}}\|_{\mathcal{D}(\mathbf{q})}| \leq C_{\text{Lip}}^{\mathcal{D}} \|\mathbf{p} - \mathbf{q}\| \|\hat{\mathbf{p}}\|. \quad (31)$$

Such a constant exists by the Lipschitz regularity of the metric  $\mathcal{M}$ , assumed in Theorem 1.3.

**Lemma 2.8.** *Assume that  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in \Omega_h \times \Omega$  and define  $\bar{\mathbf{w}}$  and  $\bar{U}$  as in Lemma 2.7. Then*

$$|\mathfrak{F}_h \bar{U}(\bar{\mathbf{p}}) - \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})}| \leq C_1 r_* \frac{h}{\delta}, \quad \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{q}})} - \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})} \leq C_2 \delta, \quad (32)$$

with  $C_1 := \lambda^*(\mathcal{D})\sqrt{d}$  and  $C_2 := (4C_{\text{Lip}}^u)^2 C_{\text{Lip}}^{\mathcal{D}}$ . This contradicts (29) unless  $\lambda \geq 1 - C_1 r_* h / \delta - C_2 \delta$ . The same estimates and conclusion hold for  $\tilde{\mathbf{w}}$  and  $\tilde{U}$  if  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in \Omega_h \times \Omega$ .

*Proof.* We focus on the case of  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ , the second case of  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$  being similar, and begin with the proof of (32, left) which contains the key technical points. By definition of the quadratic function  $\bar{U}$ , one has

$$\max\{0, \bar{U}(\bar{\mathbf{p}}) - \bar{U}(\bar{\mathbf{p}} + h\dot{\mathbf{e}}_i), \bar{U}(\bar{\mathbf{p}}) - \bar{U}(\bar{\mathbf{p}} - h\dot{\mathbf{e}}_i)\} = h|\langle \bar{\mathbf{w}}, \dot{\mathbf{e}}_i \rangle| + \frac{h^2}{\delta} \|\dot{\mathbf{e}}_i\|^2, \quad (33)$$

for any  $1 \leq i \leq d'$ , where  $(\rho_i(\bar{\mathbf{p}}), \dot{\mathbf{e}}_i)_{i=1}^{d'}$  are the weights and offsets of the discretization scheme at  $\bar{\mathbf{p}}$ , see (14). Denote by  $w, e \in \mathbb{R}^{d'}$  the vectors of components, respectively,  $w_i := |\langle \bar{\mathbf{w}}, \dot{\mathbf{e}}_i \rangle|$ , and  $e_i := \|\dot{\mathbf{e}}_i\|^2$ , for all  $1 \leq i \leq d'$ . Introduce also the semi-norm  $\|z\|_{\bar{\mathbf{p}}} := \sqrt{\sum_{i=1}^{d'} \rho_i(\bar{\mathbf{p}}) z_i^2}$  on  $\mathbb{R}^{d'}$ . Then by (33) and the consistency relation (45) one has

$$\mathfrak{F}_h \bar{U}(\bar{\mathbf{p}}) = \|w + \frac{h}{\delta} e\|_{\bar{\mathbf{p}}}, \quad \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})} = \|w\|_{\bar{\mathbf{p}}},$$

and therefore  $|\mathfrak{F}_h \bar{U}(\bar{\mathbf{p}}) - \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})}| \leq \frac{h}{\delta} \|e\|_{\bar{\mathbf{p}}}$  by the triangular inequality. Finally observe that

$$\|e\|_{\bar{\mathbf{p}}}^2 = \sum_{1 \leq i \leq d'} \rho_i(\bar{\mathbf{p}}) (\|\dot{\mathbf{e}}_i\|^2)^2 \leq \sum_{1 \leq i \leq d'} \rho_i(\bar{\mathbf{p}}) \|\dot{\mathbf{e}}_i\|^2 \max_{1 \leq i \leq d'} \|\dot{\mathbf{e}}_i\|^2 = \text{Tr}(\mathcal{D}(\bar{\mathbf{p}})) r_*^2,$$

where  $\text{Tr}$  denotes the trace of a matrix. The announced result (32, left) then follows from  $\text{Tr}(\mathcal{D}(\bar{\mathbf{p}})) \leq d(\lambda^*(\bar{\mathbf{p}}))^2$ .

The second estimate (32, right) follows from the Lipschitz regularity of the metric (31), together with the upper bound  $\|\bar{\mathbf{p}} - \bar{\mathbf{q}}\| \leq 4C_{\text{Lip}}^u \delta$  established in Theorem 2.4, which implies  $\|\bar{\mathbf{w}}\| \leq 4C_{\text{Lip}}^u$ . Combining these estimates with Lemma 2.7 yields

$$\lambda + C_1 r_* h / \delta \geq \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})} \geq 1 - C_2 \delta,$$

which implies the announced lower bound for  $\lambda$ . The same estimates can be derived in the second case, and with Lemma 2.7 they imply  $1 - C_1 r_* h / \delta \leq \|\tilde{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})} \leq \lambda + C_2 \delta$  which yields the same lower bound for  $\lambda$ .  $\square$

**Establishing assumption (ii) of Theorem 2.4.** The reste of this section is devoted to estimating on the growth of the discrete solution  $U_h$  close to  $\partial\Omega$ , see Proposition 2.11, thereby establishing assumption (ii) of Theorem 2.4, and concluding the proof of Theorem 1.3. A natural strategy would be to prove a global Lipschitz type estimate for the discrete solution  $U_h$ , as in e.g. [10, 47], but unfortunately the assumptions of Theorem 1.3 are too weak for that purpose, and actually we cannot exclude a staggered grid effect (never observed in practice) far from  $\partial\Omega$ . Instead, the idea underlying our proof is to construct from any point in  $\mathbf{p}_0 \in \Omega_h$  sufficiently close to  $\partial\Omega$ , a short chain of neighbors  $\mathbf{p}_1, \dots, \mathbf{p}_n$  ending in  $\partial\Omega_h$  and connected by offsets of the numerical scheme:  $\mathbf{p}_{i+1} = \mathbf{p}_i + h\dot{\mathbf{e}}_i(\mathbf{p}_i)$ , where the associated weight  $\rho_i(\mathbf{p}_i)$  is positively bounded below. This chain is the discrete counterpart of a short time local control to the to boundary [3].

Our first step is to provide a precise definition to the *exterior cone condition* assumed in the statement of Theorem 1.3. It is followed with a technical lemma comparing the Euclidean norm with its first order Taylor expansion.

**Definition 2.9** (Exterior cone condition). *The domain  $\Omega \subseteq \mathbb{E}$  obeys an exterior cone condition iff there exists constants  $C_\Omega$  and  $c_\Omega > 0$  such that for all  $h \leq c_\Omega$ ,*

$$\forall \mathbf{p} \in \partial\Omega, \exists \mathbf{q} \in B(\mathbf{p}, C_\Omega h), \text{ such that } B(\mathbf{q}, h) \subseteq \mathbb{E} \setminus \Omega,$$

where  $B(\mathbf{q}, h)$  denotes the open ball of center  $\mathbf{q}$  and radius  $h$ .

**Lemma 2.10.** *For any  $\mathbf{p}, \dot{\mathbf{e}} \in \mathbb{E}$  with  $\mathbf{p} \neq 0$ , one has  $\|\mathbf{p} + \dot{\mathbf{e}}\| \leq \|\mathbf{p}\| + (\mathbf{p} \cdot \dot{\mathbf{e}}) / \|\mathbf{p}\| + \|\dot{\mathbf{e}}\|^2 / (2\|\mathbf{p}\|)$ .*

*Proof.* Multiplying both sides by  $\|\mathbf{p}\|$  and rearranging terms the statement is found equivalent to  $\|\mathbf{p}\| \|\mathbf{p} + \dot{\mathbf{e}}\| \leq \frac{1}{2}(\|\mathbf{p}\|^2 + \|\mathbf{p} + \dot{\mathbf{e}}\|^2)$ , equivalently to  $0 \leq (\|\mathbf{p}\| - \|\mathbf{p} + \dot{\mathbf{e}}\|)^2$  which holds true.  $\square$

The following proposition concludes the proof of Theorem 1.3. We let  $\mu(\mathcal{D}) := \max\{\mu(\mathcal{D}(\mathbf{p})); \mathbf{p} \in \bar{\Omega}\}$ , where the anisotropy ratio of a symmetric matrix is defined in (6). condition number of a symmetric matrix is defined in (46).

**Proposition 2.11.** *Let  $\mathbf{p} \in \Omega_h$ , and let  $\mathbf{q} \in \mathbb{E}$  be such that  $\|\mathbf{p} - \mathbf{q}\| \geq C_0 r_* h$ , with  $C_0 := \text{Cond}(\mathcal{D})\sqrt{d'}$ . Then there exists  $1 \leq i \leq d'$  and a sign  $s \in \{-1, 1\}$  such that*

$$U_h(\mathbf{p}) \leq U_h(\mathbf{p} + hs\dot{\mathbf{e}}_i) + hC_1 \|\dot{\mathbf{e}}_i\|, \quad \text{and } \|\mathbf{p} + hs\dot{\mathbf{e}}_i - \mathbf{q}\| \leq \|\mathbf{p} - \mathbf{q}\| - hc_2 \|\dot{\mathbf{e}}_i\|, \quad (34)$$

with  $C_1 := \lambda^*(\mathcal{M})\sqrt{d'}$ ,  $c_2 := 1/(2 \text{Cond}(\mathcal{D})\sqrt{d'})$ . This implies assumption (ii) of Theorem 2.4, with the constants  $C_{\text{bd}}^U = C_1/c_2$ ,  $C_{\text{bd}}^U := C_{\text{bd}}^U C_\Omega C_0 r_*$ , and  $c_{\text{bd}} = +\infty$ .

*Proof.* Denote by  $(\lambda_*)^2$  and  $(\lambda^*)^2$  the smallest and largest eigenvalue of  $\mathcal{D}(\mathbf{p})$  respectively. Let  $\mathbf{n} := (\mathbf{q} - \mathbf{p})/\|\mathbf{q} - \mathbf{p}\|$ , regarded as a co-vector thanks to the Euclidean structure of  $\mathbb{E}$ . Then

$$\lambda_*^2 \leq \|\mathbf{n}\|_{\mathcal{D}(\mathbf{p})}^2 = \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \langle \mathbf{n}, \dot{\mathbf{e}}_i \rangle^2. \quad (35)$$

Fix  $1 \leq i \leq d'$  such that  $\rho_i(\mathbf{p}) \langle \mathbf{n}, \dot{\mathbf{e}}_i \rangle^2 \geq \lambda_*^2/d'$ . Denote  $\rho^2 := \rho_i(\mathbf{p})$ , and  $\dot{\mathbf{e}} := s\dot{\mathbf{e}}_i$  where  $s$  is the sign of  $\langle \mathbf{n}, \dot{\mathbf{e}}_i \rangle$ . One has, using that  $\rho^2 \dot{\mathbf{e}} \otimes \dot{\mathbf{e}} \preceq \mathcal{D}(\mathbf{p}) \preceq (\lambda^*)^2 \text{Id}$  for the second inequality

$$\rho \langle \mathbf{n}, \dot{\mathbf{e}} \rangle \geq \lambda_*/\sqrt{d'} \quad \rho \|\dot{\mathbf{e}}\| \leq \lambda^* \quad (36)$$

By definition of the discretization scheme (27) one has  $h^{-2} \rho^2 \max\{0, U_h(\mathbf{p}) - U_h(\mathbf{p} + h\dot{\mathbf{e}})\}^2 \leq (\mathfrak{F}_h U_h(\mathbf{p}))^2 = 1$ , hence using (36, left) we obtain (34, left):

$$U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}}) \leq \frac{h}{\rho} \leq h \frac{\|\dot{\mathbf{e}}\| \sqrt{d'}}{\lambda_*}. \quad (37)$$

By (36) and  $\text{Cond}(\mathcal{D}) \geq \lambda^*/\lambda_*$  one has  $\langle \mathbf{n}, \dot{\mathbf{e}} \rangle \geq \|\dot{\mathbf{e}}\|/(\text{Cond}(\mathcal{D})\sqrt{d'})$ . Denote  $\mathbf{r} := (\mathbf{p} - \mathbf{q})/h$  and observe that  $\|\mathbf{r}\| \geq C_0 r_* \geq \text{Cond}(\mathcal{D})\sqrt{d'}\|\dot{\mathbf{e}}\|$ , by assumption and by definition of the max stencil radius  $r_*$ , see (15). Using Lemma 2.10 we obtain

$$\|\mathbf{r} + \dot{\mathbf{e}}\| \leq \|\mathbf{r}\| - \langle \mathbf{n}, \dot{\mathbf{e}} \rangle + \frac{\|\dot{\mathbf{e}}\|^2}{2\|\mathbf{r}\|} \leq \|\mathbf{r}\| - \frac{\|\dot{\mathbf{e}}\|}{\text{Cond}(\mathcal{D})\sqrt{d'}} + \frac{\|\dot{\mathbf{e}}\| \|\dot{\mathbf{e}}\|}{2C_0 r_*} \leq \|\mathbf{r}\| - \frac{\|\dot{\mathbf{e}}\|}{2\text{Cond}(\mathcal{D})\sqrt{d'}},$$

which is (34, right).

Finally, we conclude the proof of assumption (ii). Let  $\mathbf{p}_0 \in \Omega_h$ . Let  $\mathbf{q}_* \in \partial\Omega$  be the closest point to  $\mathbf{p}_0$ , and let  $\mathbf{q} \in B(\mathbf{q}_*, C_\Omega C_0 r_* h)$  be such that  $B(\mathbf{q}, C_0 r_* h) \subseteq \mathbb{E} \setminus \Omega$ . By the above argument, there exists a finite sequence of points  $\mathbf{p}_1, \dots, \mathbf{p}_{k-1} \in \Omega_h$ ,  $\mathbf{p}_k \in \partial\Omega_h$ , such that  $U(\mathbf{p}_i) \leq U(\mathbf{p}_{i+1}) + C_1 \delta_i$  and  $\|\mathbf{p}_{i+1} - \mathbf{q}\| \leq \|\mathbf{p}_i - \mathbf{q}\| - c_2 \delta_i$ , denoting  $\delta_i := \|\mathbf{p}_{i+1} - \mathbf{p}_i\|$ , for all  $0 \leq i < k$ . Since  $U(\mathbf{p}_k) = 0$  we obtain  $U(\mathbf{p}_0) \leq C_1(\delta_0 + \dots + \delta_{k-1})$ , and since  $\|\mathbf{p}_k - \mathbf{q}\| \geq 0$  we obtain  $c_2(\delta_0 + \dots + \delta_{k-1}) \leq \|\mathbf{p}_0 - \mathbf{q}\| \leq \|\mathbf{p}_0 - \mathbf{q}_*\| + C_\Omega C_0 r_* h$ . Hence finally, as announced

$$U(\mathbf{p}_0) \leq (C_1/c_2)(d_{\partial\Omega}(\mathbf{p}_0) + C_\Omega C_0 r_* h). \quad \square$$

### 3 Numerical results

We illustrate the numerical methods introduced in this paper with a series of numerical experiments, involving Riemannian, sub-Riemannian and Rander metrics, in §3.1, §3.2 and §3.3 respectively. Open source numerical codes for the Riemannian and sub-Riemannian models<sup>6</sup> are available on the author's webpage<sup>7</sup>.

#### 3.1 Riemannian examples

We validate our algorithm on several two and three-dimensional Riemannian test cases, which are split into two groups. The problems of the first group - related to differential geometry and seismic imaging - feature smooth Riemannian metrics with pronounced yet bounded anisotropy, and accuracy is the main concern. The problems of the second group - related to tubular structure segmentation in medical image data - feature discontinuous Riemannian metrics and extreme anisotropies, so that robustness is the main concern.

<sup>6</sup>Numerical codes for Rander metrics, which are more experimental, are available on demand.

<sup>7</sup><https://github.com/Mirebeau>

**Smooth Riemannian metrics.** The first test, two dimensional and introduced in [82], is the computation of the distance from the origin on a parametric surface w.r.t. the Riemannian metric induced by the Euclidean metric on  $\mathbb{R}^3$ . The surface is described by the height map

$$z(x, y) := (3/4) \sin(3\pi x) \sin(3\pi y),$$

hence the Riemannian metric is  $\mathcal{M}(x, y) = \text{Id} + \nabla z(x, y) \nabla z(x, y)^T$ , which maximum condition number (46) is  $\approx 5.1^2$ . The parametrization domain is the unit square  $[-0.5, 0.5]^2$ , rotated<sup>8</sup> by the angle  $\pi/6$ .

The second test, two dimensional and introduced in [72], is inspired by seismic imaging applications. Note that this application often involves more complex types of Finslerian anisotropies [63]. The Riemannian metric tensor  $\mathcal{M}(x, y)$  has eigenvector  $(1, (\pi/2) \cos(4\pi x))$  with eigenvalue  $0.8^{-2}$ . The second eigenvalue is  $0.2^{-2}$ , hence the condition number is  $4^2$ . The parametrization domain is  $[-0.5, 0.5]^2$ , and the distance is computed from the origin.

The third test, introduced here for the first time, extends the seismic imaging inspired second test to three dimensions. The Riemannian metric tensor  $\mathcal{M}(x, y, z)$  has eigenvector  $(\cos(3\pi(x+y)), \sin(3\pi(2x-y)), 0.5)$ , with eigenvalue  $0.2^{-2}$ . The two other eigenvalues are equal to  $0.8^{-2}$ , hence the condition number is  $4^2$ . The domain is  $[-0.5, 0.5]^3$  and the distance is computed from the origin  $(0, 0, 0)$ .

Figure 2 illustrates the level sets of the distance maps associated with these three tests, computed numerically using the FM-VR1, and a number of the corresponding minimal geodesics. The accuracy and computation time in the two-dimensional test cases are compared in §3.4 with several alternative numerical methods.

**Discontinuous Riemannian metrics with extreme anisotropy.** Anisotropic fast marching methods have shown their relevance for image segmentation methods based on minimal paths [7, 18]. In these applications, the metric often varies quickly, if not discontinuously, both in orientation and aspect ratio. For instance, the Riemannian metric is often designed so as to favor paths that remain close and tangent to a collection of thin tubular structures in the image.

We present two numerical experiments inspired by these applications, in two and three dimensions, which first appeared in [7] and [47] respectively. The Riemannian metric is Euclidean (identity matrix) except in the neighborhood of a curve  $\Gamma$  embedded in the domain, where the metric is extremely anisotropic, with eigenvalues  $(1, 1/100^2)$  or  $(1, 1, 1/50^2)$  in the two and three dimensional experiments respectively, and the tangent vector to the curve  $\Gamma$  is an eigenvector for the small eigenvalue. See [47] for a complete description. Figure 3 illustrates the level sets of the distance maps associated with these two test cases, computed numerically using the FM-VR1, and some of the corresponding minimal geodesics.

In these extreme test cases, the FM-VR1 behaves particularly well in terms of CPU time and accuracy, comparable to the FM-LBR which similarly uses an adaptive discretization strategy. In contrast, iterative numerical methods such as the AGSI [10], and fast marching methods based on less sophisticated stencil constructions such as [1], have been shown to fail on these types of benchmarks [7, 47].

### 3.2 Sub-Riemannian models

We consider several sub-Riemannian models, posed on the configuration space  $\mathbb{M} := \mathbb{R}^d \times \mathbb{S}^{d-1}$  of positions and orientations. Such configurations are denoted  $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$ , and their tangent

---

<sup>8</sup>Without this rotation, the Riemannian metric anisotropy is mostly aligned with the coordinate axes, which makes the problem significantly easier numerically, see the numerical experiments in [47].

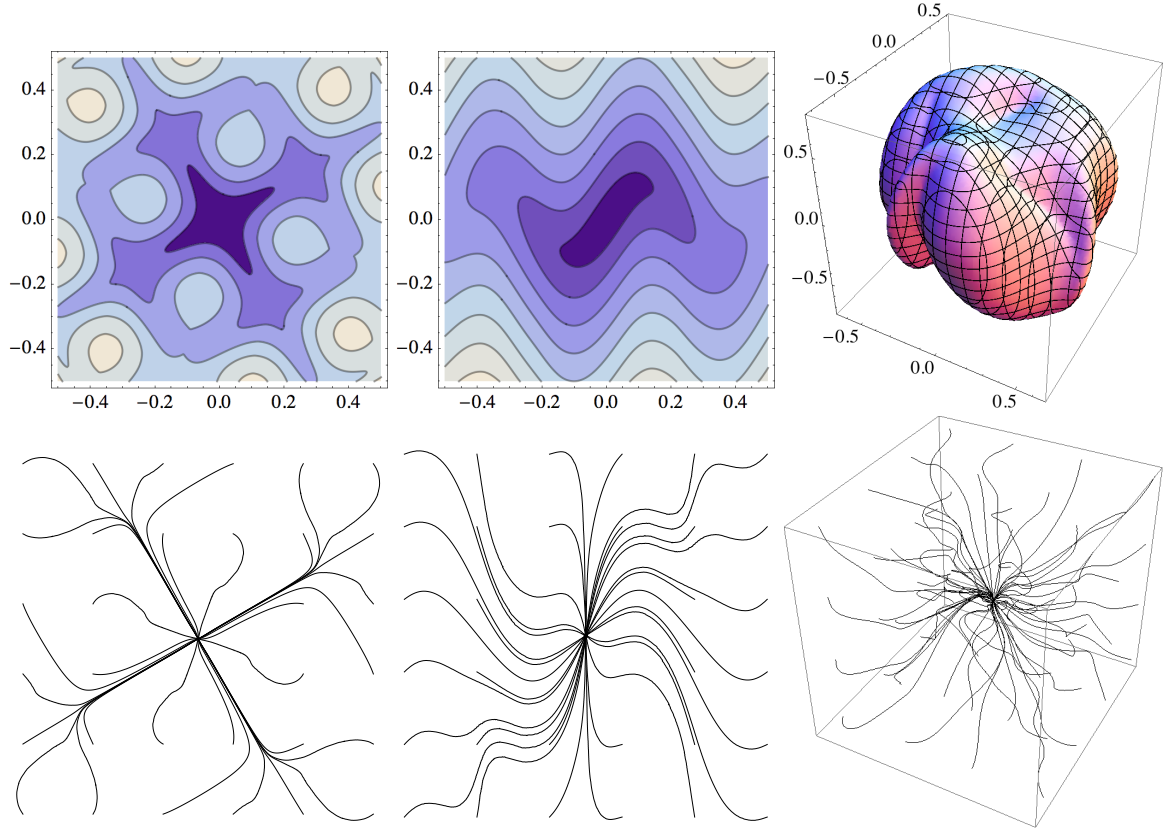


Figure 2: Numerics, illustrating §3.1, for the three test cases involving smooth anisotropic Riemannian metrics. Top: level lines of the numerically computed distance map from the domain center. Bottom: backtracked minimal geodesics, from the domain center to various endpoints. Left: geodesic distance on a parametric surface,  $283^2$  grid, 0.1s CPU time. Center: test inspired by seismic data analysis,  $193^2$  grid, 0.04s CPU time. Right: likewise in 3D, on a  $101^3$  grid, CPU time 5.02s. *All CPU times measured on a 2.7Ghz laptop using a single thread.*

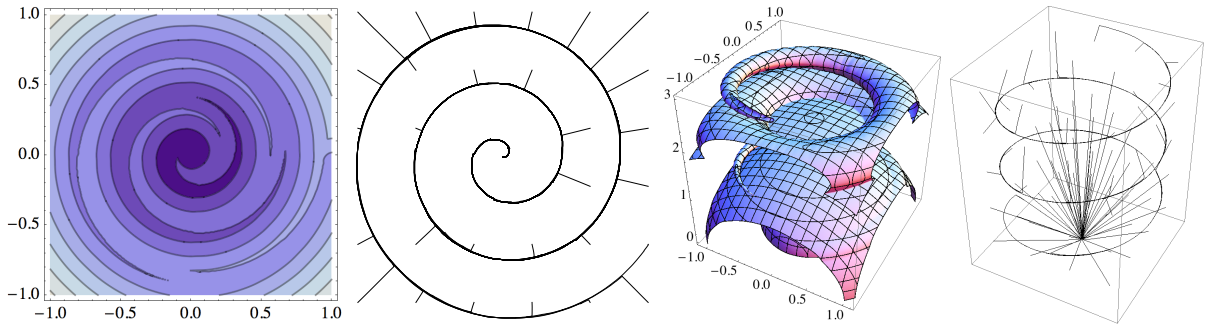


Figure 3: Numerics, illustrating §3.1, for test cases involving discontinuous and extremely anisotropic Riemannian metrics, in dimension  $d \in \{2, 3\}$ , inspired by applications to tubular structure segmentation in medical data [7]. (i,iii) Level lines of the numerically computed distance from the image center. (ii,iv) Backtracked minimal geodesics, from the image center to various endpoints. Because of the chosen metric, these paths are concatenations of straight lines (Euclidean geodesics), and of portions adjacent to a given spiraling curve  $\Gamma$  (along which paths are favored by the metric anisotropy). Left:  $201 \times 201$  grid, 0.03s CPU time. Right:  $201 \times 201 \times 272$  grid, 25s CPU time.

vectors are denoted  $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_{\mathbf{p}}\mathbb{M}$ . For the simplicity of the exposition we regard  $\mathbf{n}$  as a genuine unit vector in  $\mathbb{R}^d$ , so that  $\langle \mathbf{n}, \dot{\mathbf{n}} \rangle = 0$ , although our numerical implementation relies on an angular parametrisation of the sphere  $\mathbb{S}^{d-1}$ .

We choose to describe the sub-Riemannian models of interest via an approximating family of Riemannian metrics  $(\mathcal{M}_\varepsilon)_{\varepsilon>0}$ , where  $\varepsilon$  is a relaxation parameter. The orthogonal projection of a vector  $\dot{\mathbf{x}} \in \mathbb{R}^d$ , onto the hyperplane orthogonal to a given unit vector  $\mathbf{n} \in \mathbb{S}^{d-1}$ , is denoted

$$P_{\mathbf{n}}(\dot{\mathbf{x}}) := \dot{\mathbf{x}} - \langle \mathbf{n}, \dot{\mathbf{x}} \rangle \mathbf{n}.$$

**The Reeds-Shepp model.** This model, defined on  $\mathbb{R}^2 \times \mathbb{S}^1$ , describes a car<sup>9</sup>, which state is described by a position  $\mathbf{x} \in \mathbb{R}^2$ , and an orientation  $\mathbf{n} = (\cos \theta, \sin \theta) \in \mathbb{S}^1$ , see Figure 4. The car can move forward and backward, and rotate right and left, but cannot move sideways. This model also plays a central role in the study of the visual cortex organization and function, in which case it is referred to as the Petitot-Citti-Sarti model [57]. Recently, data-driven variants of the Reeds-Shepp model and of its higher dimensional counterparts have been considered for tubular structure segmentation in medical image data [5, 28]. The Riemannian relaxations of this model's metric read, for each  $\varepsilon > 0$ ,

$$\|\dot{\mathbf{p}}\|_{\mathcal{M}_\varepsilon(\mathbf{p})}^2 := \mathcal{S}(\mathbf{p})^{-2} (\langle \mathbf{n}, \dot{\mathbf{x}} \rangle^2 + \varepsilon^{-2} \|P_{\mathbf{n}}(\dot{\mathbf{x}})\|^2 + \xi^2 \|\dot{\mathbf{n}}\|^2) \quad (38)$$

where  $\mathcal{S} : \mathbb{M} \rightarrow ]0, \infty[$  is a point dependent speed function, with physical units [length]/[time], and  $\xi$  is a parameter which has the dimension [length] of a radius of curvature. Parameters  $\mathcal{S}$  and  $\xi$  may be constant or variable over the domain, possibly dictated by the considered application in a data-driven manner [5, 28].

The Reeds-Shepp model is related to curvature penalization for the following reason: consider a smooth path  $\mathbf{x} : [0, T] \rightarrow \mathbb{R}^d$ , with non-vanishing speed. Then there exists a unique  $\mathbf{n} : [0, T] \rightarrow \mathbb{S}^{d-1}$ , up to a global change of sign, such that the lifted path  $t \in [0, T] \mapsto \gamma(t) = (\mathbf{x}(t), \mathbf{n}(t))$  has finite length with respect to the sub-Riemannian metric  $\mathcal{M}_0$ . Indeed, one must set  $\mathbf{n}(t) := \pm \dot{\mathbf{x}}(t) / \|\dot{\mathbf{x}}(t)\|$ , so that  $P_{\mathbf{n}(t)}(\dot{\mathbf{x}}(t)) = 0$  for all  $t \in [0, T]$ . Then, denoting by  $\kappa(t) := \|\dot{\mathbf{n}}(t)\| / \|\dot{\mathbf{x}}(t)\|$  the curvature of the path  $\mathbf{x}$ , one obtains

$$\int_0^T \|\dot{\gamma}(t)\|_{\mathcal{M}_0(t)} dt = \int_0^T \sqrt{1 + \xi^2 \kappa(t)^2} \frac{\|\dot{\mathbf{x}}(t)\| dt}{\mathcal{S}(\gamma(t))}.$$

Note that (contrary to what this discussion may suggest) the physical projections of geodesic paths for the sub-Riemannian metric  $\mathcal{M}_0$  are only piecewise smooth typically, because they feature *cusps*, see Figures 4, 6, and the discussion in [28].

Some experiments involving two and three dimensional physical paths are presented in Figures 4, 5 and 6. Let us emphasize that we are here solving strongly anisotropic PDEs on three and five dimensional domains respectively. The control sets for the Reeds-Shepp model posed on  $\mathbb{R}^2 \times \mathbb{S}^1$  are illustrated on page 9. For the model posed on  $\mathbb{R}^3 \times \mathbb{S}^2$ , the sphere  $\mathbb{S}^2$  is parametrized using the Euler angles  $(\theta, \varphi) \mapsto (\cos \theta, \sin \theta \cos \varphi, \sin \theta \sin \varphi)$ , from the flat domain  $[0, \pi] \times [0, 2\pi]$  equipped with the adequate Riemannian metric and boundary conditions.

The numerical results are similar to those obtained in [28] using the semi-lagrangian FM-LBR, but computation times are shorter by a factor 5 typically for the five-dimensional test cases, see the discussion in §3.1. Figure 4 illustrates the spatial projections in  $\mathbb{R}^2$  of the minimal geodesics associated with the classical Reeds-Shepp model posed on  $\mathbb{R}^2 \times \mathbb{S}^1$ , with and without obstacles, and in the latter case comparison with high accuracy solutions obtained with an ODE shooting method.

---

<sup>9</sup>Actually, the Reeds-Shepp model more faithfully describes a wheelchair.

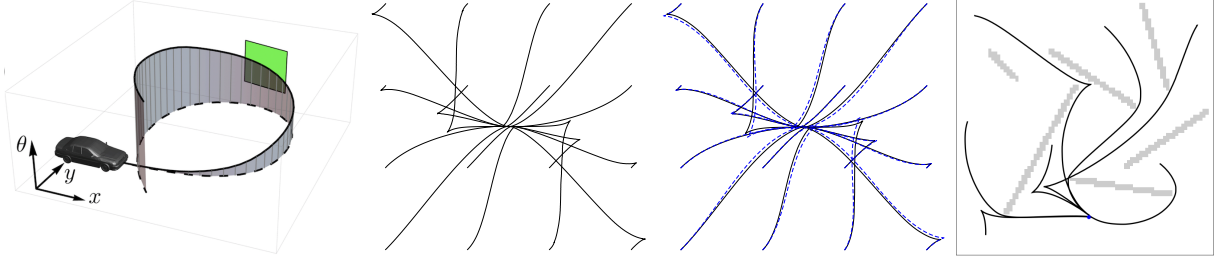


Figure 4: Numerics for the Reeds-Shepp sub-Riemannian model, illustrating §3.2. (i) Explanation of the model: the parameter space is three dimensional, and the sub-Riemannian structure forces the car ground speed  $\dot{\mathbf{x}}(t)$  to remain aligned with the direction  $\mathbf{n}(t) = (\cos \theta(t), \sin \theta(t))$  defined by the third coordinate. In the next sub-figures, only the planar projection  $\mathbf{x}(t)$  of the minimal paths  $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t)) \in \mathbb{R}^2 \times \mathbb{S}^1$  is shown. (ii) Projections of minimal paths in  $[-0.5, 0.5]^2 \times \mathbb{S}^1$ , from  $(0, 0, 0)$  to various endpoints with imposed orientation  $\theta = \pi/4$ , with model parameters  $\varepsilon = 0.1$ ,  $\xi = 0.3$ . (iii) Comparison of the numerically backtracked paths (solid) with those obtained using a high order ODE shooting method based on the Hamilton equations of geodesics (40) (dashed blue). (iv) Some minimal paths (black) for the Reeds-Shepp model in the presence of obstacles (grayed). Domain  $[0, 1]^2 \times \mathbb{S}^1$  discretized on a  $90^2 \times 60$  grid, CPU time 0.36s. Model parameters  $\varepsilon = 0.1$ , and  $\xi = 0.4$ . Some orientations, arbitrary, are imposed at the geodesics endpoints.

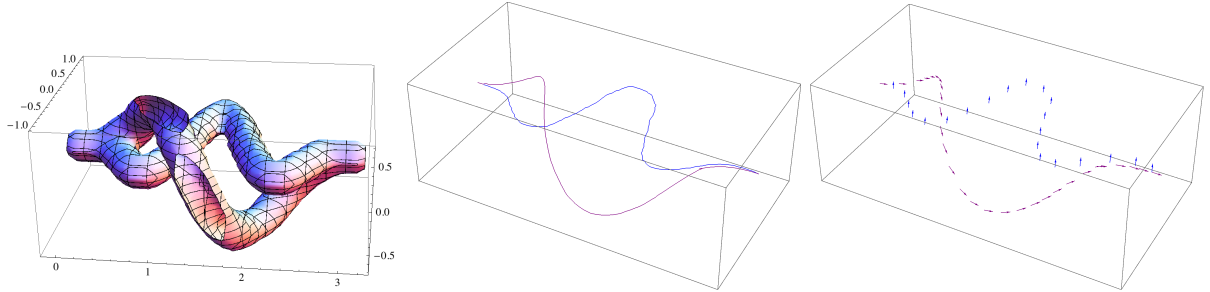


Figure 5: Numerics, illustrating §3.2, for sub-Riemannian models posed on  $\mathbb{R}^3 \times \mathbb{S}^2$ . Test case inspired by 3D tubular structure segmentation. (i) Contour plot of the speed function, which is high in the neighborhood of two curves of small curvature and small torsion respectively. (ii, iii) Position  $\mathbf{x}(t)$  and orientation  $\mathbf{n}(t)$ , respectively, of the sub-Riemannian geodesics  $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t))$  extracted with the Reeds-Shepp model (purple), and its torsion related variant (blue). Distinct paths are selected, along the appropriate vessel centerline. Domain  $([0, \pi] \times [-1, 1] \times [-1, 0.5]) \times \mathbb{P}^2$  discretized using a  $(40 \times 20 \times 16) \times (5 \times 20)$  grid,  $\varepsilon = 0.2$ . CPU time 6.6 s.

**A variant related to torsion penalization.** We introduce a new sub-Riemannian model, which relaxed metric is defined for all  $\mathbf{p} = (\mathbf{x}, \mathbf{n}) \in \mathbb{M}$  and  $\dot{\mathbf{p}} = (\dot{\mathbf{x}}, \dot{\mathbf{n}}) \in T_{\mathbf{p}}\mathbb{M}$  by

$$\|\dot{\mathbf{p}}\|_{\mathcal{M}_\varepsilon(\mathbf{p})}^2 := \mathcal{S}(\mathbf{p})^{-2} (\langle \mathbf{n}, \dot{\mathbf{x}} \rangle^2 + \varepsilon^{-2} \|P_{\mathbf{n}}(\dot{\mathbf{x}})\|^2 + \xi^2 \|\dot{\mathbf{n}}\|^2), \quad (39)$$

where again  $\mathcal{S} : \mathbb{M} \rightarrow ]0, \infty[$  is the speed function, and  $\xi$  has the dimension of a length. The model (39) favors paths which are possibly non-smooth but are embedded in smooth surfaces, a property that is relevant for certain tasks in medical data segmentation [76]. Indeed the physical velocity  $\dot{\mathbf{x}}$  is constrained by the cost of  $\varepsilon^{-2} \langle \mathbf{n}, \dot{\mathbf{x}} \rangle^2$  to remain (approximately if  $\varepsilon > 0$ ) in the plane orthogonal to the vector  $\mathbf{n}$ , whose variation is itself controlled by the cost of  $\|\dot{\mathbf{n}}\|^2$ . Note also that the most natural way to lift a physical curve  $\mathbf{x} : [0, T] \rightarrow \mathbb{R}^3$  into  $\gamma = (\mathbf{x}, \mathbf{n}) : [0, T] \rightarrow \mathbb{R}^3 \times \mathbb{S}^2$  obeying the orthogonality constraint  $\langle \dot{\mathbf{x}}(t), \mathbf{n}(t) \rangle = 0$  for all  $t \in [0, T]$ , is to define  $\mathbf{n}(t) := (\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)) / \|\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)\|$ . Then denoting by  $\tau(t) := \|\dot{\mathbf{n}}(t)\| / \|\dot{\mathbf{x}}(t)\|$  the torsion of the path  $\mathbf{x}$  one obtains

$$\int_0^T \|\dot{\gamma}(t)\|_{\mathcal{M}_0(t)} dt = \int_0^T \sqrt{1 + \xi^2 \tau(t)^2} \frac{\|\dot{\mathbf{x}}(t)\| dt}{\mathcal{S}(\gamma(t))}.$$

Nevertheless our model is only related to torsion penalization, and not equivalent to it, because there exists other lifts  $\gamma = (\mathbf{x}, \mathbf{n}) : [0, T] \rightarrow \mathbb{R}^d \times \mathbb{S}^{d-1}$  of the curve  $\mathbf{x} : [0, T] \rightarrow \mathbb{R}^d$  obeying the required orthogonality constraint, and which energy could be smaller than the torsion based one.

On the experiment presented Figure 5, the speed function  $\mathcal{S} : \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow ]0, \infty[$  only depends on the physical position  $\mathbf{x}$ , and is small away from two curves  $\Gamma_1$  and  $\Gamma_2$  of interest

$$\mathcal{S}(\mathbf{x}, \mathbf{n}) := \max\{s, \exp\left(-\frac{\text{dist}(\mathbf{x}, \Gamma_1 \cup \Gamma_2)^2}{2\sigma^2}\right)\},$$

where  $s = 1/6$  and  $\sigma = 0.15$ . The curves  $\Gamma_1$  and  $\Gamma_2$  are parametrized by  $t \in [0, \pi]$  as follows

$$\Gamma_1(t) := (t, \sin(t)^2 \cos(4t), 0), \quad \Gamma_2(t) := (t, \sin(t)^3 \cos(2t), \sin(t)^3 \sin(2t)).$$

Hence  $\Gamma_1$  has large curvature but no torsion, whereas  $\Gamma_2$  has small curvature but some torsion. Using our anisotropic fast marching method, we compute the shortest path between the common endpoints  $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^3$  of these curves, among all possible tangent directions  $\mathbf{n}_0, \mathbf{n}_1 \in \mathbb{S}^2$  at these endpoints. Figure 5 shows the level lines of the cost function  $\mathcal{S}$ , and the minimal geodesics corresponding to the two models (38) and (39), numerically computed using the FM-VR1. As could be expected, the torsion related model selects a path along  $\Gamma_1$ , whereas the Reeds-Shepp model selects a path along  $\Gamma_2$ .

**Validation of the approach.** We present in Figure 6 two empirical validations of our numerical approach to computing globally optimal geodesics for the five-dimensional Reeds-Shepp model and its torsion related variant. We first show that the sub-Riemannian constraint, of collinearity  $P_{\mathbf{n}}(\dot{\mathbf{x}})$  and orthogonality  $\langle \mathbf{n}, \dot{\mathbf{x}} \rangle$  are approximately satisfied, despite their relaxation in (38) and (39), with  $\varepsilon = 0.1$ . We then compare the obtained minimal paths with solutions of the Hamilton equations of geodesics

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial \mathcal{H}}{\partial \dot{\mathbf{p}}}, \quad \frac{d\dot{\mathbf{p}}}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \text{where } \mathcal{H}(\mathbf{p}, \dot{\mathbf{p}}) := \frac{1}{2} \langle \dot{\mathbf{p}}, \mathcal{D}_0(\mathbf{p}) \dot{\mathbf{p}} \rangle. \quad (40)$$

Here  $\mathcal{D}_0$  denotes the inverse tensor to the sub-Riemannian metric (38) or (39), which is well defined when  $\varepsilon = 0$ , in contrast to  $\mathcal{M}_0$  itself. This ODE is solved using a fourth order Runge-Kutta method, and the initial conditions are adjusted using a Newton method so as to meet the desired endpoints.

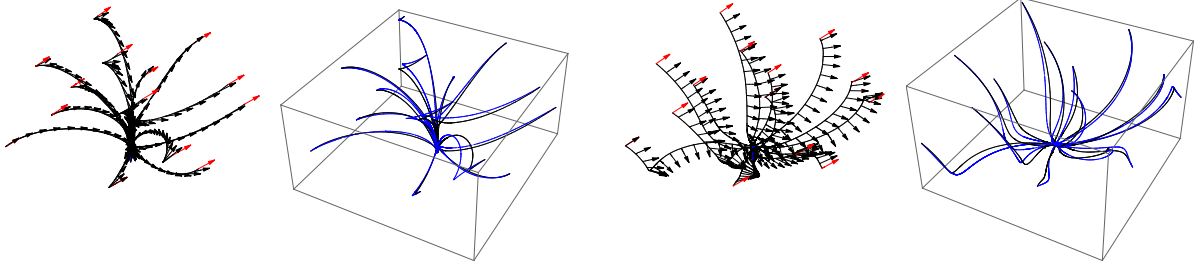


Figure 6: Validation of the FM-VR1 numerical method applied to the sub-Riemannian Reeds-Shepp models (i, ii), and its torsion related variant (iii, iv), posed on  $\mathbb{R}^3 \times \mathbb{S}^2$ . Parameters  $\varepsilon = 0.1$ ,  $\xi = 0.5$ , constant cost function  $\mathcal{S}$ . (i, iii) The angular component  $\mathbf{n}(t)$  of the minimal paths  $\gamma(t) = (\mathbf{x}(t), \mathbf{n}(t)) \in \mathbb{R}^3 \times \mathbb{S}^2$ , illustrated with arrows, satisfies (approximately, as expected) the sub-Riemannian constraint (i)  $\mathbf{n}(t) \times \dot{\mathbf{x}}(t) = 0$  or (iii)  $\langle \mathbf{n}(t), \dot{\mathbf{x}}(t) \rangle = 0$ . (ii, iv) Comparison of the backtracked geodesics (black) with the results of an ODE shooting method (blue) based on Hamilton's equations of geodesics.

### 3.3 Rander models

We consider some instances of Zermelo's navigation problem, which models a boat navigating on a body of water [15, 2]. The (motor) boat is capable of a certain maximum speed, in any direction, and inertia is not taken into account. However, the boat is also subject to a drift due to current or wind, which in our experiments is variable over the domain (and constant in time). The goal is to move from one given point to another in minimal time.

Formally, let us denote by  $\Omega \subseteq \mathbb{R}^d$  the domain, denote by  $\dot{\eta} : \bar{\Omega} \rightarrow \mathbb{R}^d$  the drift, and assume that the maximum speed is 1 (in the Euclidean norm). The boat starts from anywhere on  $\partial\Omega$ , and all points of  $\Omega$  are regarded as potential target points. The objective is thus to find for each  $\mathbf{p} \in \Omega$  the minimal time  $T = u(\mathbf{p}) \geq 0$  for which there exists a path  $\gamma : [0, T] \rightarrow \bar{\Omega}$  such that  $\gamma(0) \in \partial\Omega$ ,  $\gamma(T) = \mathbf{p}$ , and

$$\|\dot{\gamma}(t) - \dot{\eta}(\gamma(t))\| \leq 1, \quad \text{for a.e. } t \in [0, T].$$

We assume that  $\|\dot{\eta}(\mathbf{x})\| < 1$  for all  $\mathbf{x} \in \mathbb{R}^d$ , otherwise the system would not be locally controllable (the drift speed being larger than or equal to the maximum boat speed). Following [2] we reformulate this problem as a shortest path problem with respect to a Rander metric, of parameters  $(\mathcal{D}^{-1}, \hat{\eta})$  specified in the next proposition, see also (21).

**Proposition 3.1.** *The value function for this problem obeys  $\|du(\mathbf{p}) - \hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} = 1$  on  $\Omega$  in the sense of viscosity solutions, and  $u = 0$  on  $\partial\Omega$ , where*

$$\mathcal{D}(\mathbf{p}) := (1 - \|\dot{\eta}(\mathbf{p})\|^2)(1 - \dot{\eta}(\mathbf{p}) \otimes \dot{\eta}(\mathbf{p})), \quad \hat{\eta}(\mathbf{p}) = \frac{-\dot{\eta}(\mathbf{p})}{1 - \|\dot{\eta}(\mathbf{p})\|^2}.$$

*Proof.* The value function differential  $\hat{\mathbf{p}} = du(\mathbf{p})$ , where defined, obeys the equivalent constraints

$$\sup_{\dot{\mathbf{v}} \in B} \langle \hat{\mathbf{p}}, \dot{\mathbf{v}} + \dot{\eta} \rangle \leq 1 \Leftrightarrow \|\hat{\mathbf{p}}\| + \langle \hat{\mathbf{p}}, \dot{\eta} \rangle \leq 1 \Leftrightarrow \|\hat{\mathbf{p}}\|^2 \leq (1 - \langle \hat{\mathbf{p}}, \dot{\eta} \rangle)^2 \Leftrightarrow \|\hat{\mathbf{p}} - \hat{\eta}\|_{\mathcal{D}}^2 \leq 1,$$

where  $B$  denotes the Euclidean unit ball. The dependency of  $\dot{\eta}, \hat{\eta}, \mathcal{D}$  to the base point  $\mathbf{p}$  was omitted for readability. The leftmost identity follows from Bellman's optimality principle. The first equivalence is trivial, the second equivalence follows from the impossibility of  $\|\hat{\mathbf{p}}\| = \langle \hat{\mathbf{p}}, \dot{\eta} \rangle - 1$  (since  $\|\dot{\eta}\| < 1$ ), and the third equivalence results from direct computations using e.g. that  $\mathcal{D}\hat{\eta} = -(1 - \|\dot{\eta}\|^2)\dot{\eta}$ .  $\square$

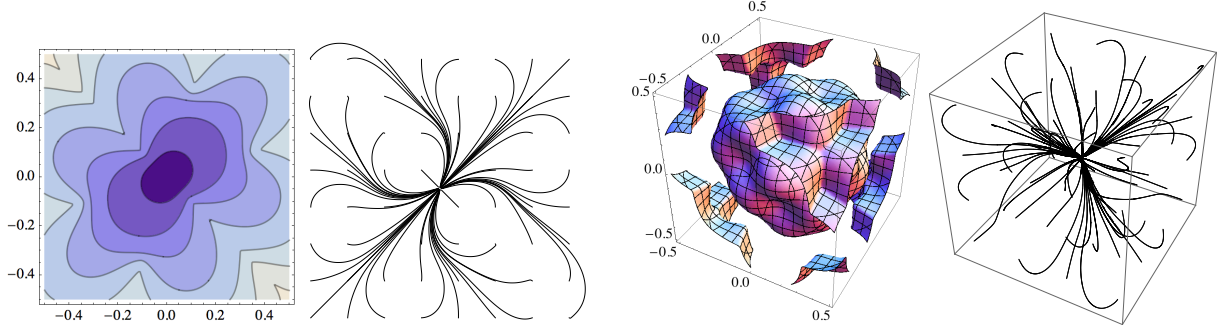


Figure 7: Numerics, Illustrating §3.3, of two instances of Zermelo's navigation problem in dimensions two and three. Level lines of the distance map from the domain center (minimal travel time), and minimal geodesics to various endpoints, computed with the variant of the FM-VR1 adapted to Rander metrics. (i, ii) Grid size  $201 \times 201$ , CPU time 0.14s. (iii, iv) Grid size  $101^3$ , CPU time 7.8s.

|                  | FM-VR1                                                |           | FM-LBR | FM-8  | FE   | MAOUM |
|------------------|-------------------------------------------------------|-----------|--------|-------|------|-------|
|                  | 1st order                                             | 2nd order |        |       |      |       |
|                  | Embedded surface distance test, $293 \times 293$ grid |           |        |       |      |       |
| CPU time         | 0.10*                                                 | 0.12*     | 0.20   | 0.21  | 1.44 | 1.31  |
| $L^\infty$ error | 5.8                                                   | 0.22*     | 5.52   | 12.5  | 9.45 | 8.56  |
| $L^1$ error      | 1.6                                                   | 0.066     | 1.46   | 3.42  | 2.51 | 2.52  |
|                  | Seismic inspired test, $193 \times 193$ grid          |           |        |       |      |       |
| CPU time         | 0.042*                                                | 0.048*    | 0.076  | 0.079 | 0.77 | 0.36  |
| $L^\infty$ error | 4.5                                                   | 0.15*     | 2.90   | 3.03  | 3.67 | 7.66  |
| $L^1$ error      | 1.5                                                   | 0.056     | 1.03   | 1.30  | 1.40 | 2.3   |

Figure 8: Comparison of the CPU time and accuracy of the proposed FM-VR1 with several alternatives in two Riemannian test cases, see §3.1 and §3.4. All errors multiplied by 100 for readability. Asterix \*, see Remark 3.2.

We present a two dimensional experiment on  $\Omega = ]-1/2, 1/2[^2$ , first introduced in [72], and a three dimensional generalization in  $\Omega = ]-1/2, 1/2[^3$ . The drift has the explicit expression

$$\dot{\eta}(\mathbf{x}) := \alpha \sin(4\pi\mathbf{x}_1) \sin(4\pi\mathbf{x}_2) \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad \left( \text{resp. } \dot{\eta}(\mathbf{x}) := \alpha \sin(4\pi\mathbf{x}_1) \sin(4\pi\mathbf{x}_2) \sin(4\pi\mathbf{x}_3) \frac{\mathbf{x}}{\|\mathbf{x}\|} \right)$$

with  $\alpha = 0.9$ . We use Adaptive Gauss Siedel Iterations, in the spirit of [10], to solve our numerical scheme, which are applicable thanks to the degenerate ellipticity property of the scheme, see Definition 2.1. Note that the single pass Fast Marching method is not applicable since it requires the additional causality property. Figure 7 illustrates the level lines of the distance map  $u$ , and some of the corresponding minimal geodesics, in the two and three dimensional test cases. The computation time and the  $L^\infty$  and  $L^1$  errors obtained with the two dimensional problem are presented in §3.4, and compared with several alternative semi-lagrangian methods [10, 1, 48].

### 3.4 Comparison with alternative methods

We compare the accuracy and computation time of the FM-VR1 with several alternative methods proposed in the literature for solving anisotropic eikonal equations [78, 70, 1, 10, 47, 48]. As discussed in the introduction, these numerical methods can be divided into two groups: causal

|                  | RD-VR1                                            |             | FM-ASR | FM-8 | FE   | MAOUM |
|------------------|---------------------------------------------------|-------------|--------|------|------|-------|
|                  | 1st order                                         | 2nd order   |        |      |      |       |
|                  | Zermelo navigation problem, $285 \times 285$ grid |             |        |      |      |       |
| CPU time         | 1.03(0.48*)                                       | 1.39(0.81*) | 0.29   | 0.16 | 1.08 | 0.69  |
| $L^\infty$ error | 0.84                                              | 0.23        | 0.64   | 0.64 | 1.05 | 2.8   |
| $L^1$ error      | 0.44                                              | 0.0095      | 0.13   | 0.11 | 0.41 | 0.17  |

Figure 9: Comparison of the CPU times and accuracy of the proposed RD-VR1 method with several alternatives in a test case involving a Rander metric, see §3.3 and §3.4. All errors multiplied by 100 for readability. When testing second-order accuracy, the seed point  $(0, 0)$  was replaced with a precomputed solution on a small disk illustrated on Figure 12. Asterix \* first time using the AGSI as described in [10], second time with a variant which limits the front width to 10 pixels. See also Remark 3.2.

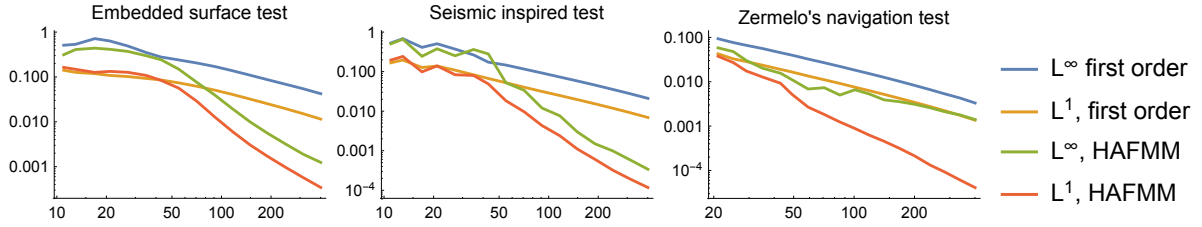


Figure 10: Numerical error as a function of gridsize for the two dimensional test cases. Second order convergence is achieved in the  $L^1$  norm, and in the  $L^\infty$  norm except for the Zermelo's navigation problem. See remark 3.2 for the experiment setup.

| First order |         |            |         |            |         |            | Second order |         |            |         |            |         |            |
|-------------|---------|------------|---------|------------|---------|------------|--------------|---------|------------|---------|------------|---------|------------|
| n           | Surface |            | Seismic |            | Zermelo |            | n            | Surface |            | Seismic |            | Zermelo |            |
|             | $L^1$   | $L^\infty$ | $L^1$   | $L^\infty$ | $L^1$   | $L^\infty$ |              | $L^1$   | $L^\infty$ | $L^1$   | $L^\infty$ | $L^1$   | $L^\infty$ |
| 51          | 7.73    | 24.9       | 5.54    | 15.4       | 2.21    | 4.63       | 51           | 6.60    | 17.9       | 2.28    | 14.3       | 0.474   | 0.953      |
| 101         | 4.65    | 15.6       | 2.87    | 8.42       | 1.20    | 2.42       | 101          | 0.946   | 2.98       | 0.373   | 1.22       | 0.0887  | 0.657      |
| 201         | 2.40    | 8.28       | 1.46    | 4.35       | 0.62    | 1.21       | 201          | 0.149   | 0.527      | 0.050   | 0.142      | 0.0214  | 0.319      |
| 401         | 1.15    | 4.23       | 0.692   | 2.11       | 0.303   | 0.575      | 401          | 0.034   | 0.124      | 0.012   | 0.034      | 0.0044  | 0.162      |
| $\sigma$    | 1.06    | 0.97       | 1.08    | 1.04       | 1.05    | 1.08       | $\sigma$     | 2.12    | 2.09       | 2.10    | 2.07       | 2.28    | 0.98       |

Figure 11: Numerical error observed with the proposed schemes, for the three test cases, for several grid sizes  $n \times n$ . Last line: exponent such that  $\text{err}(n) \approx n^{-\sigma}$ , obtained using  $n \in \{201, 401\}$ . All errors multiplied by 100 for readability.

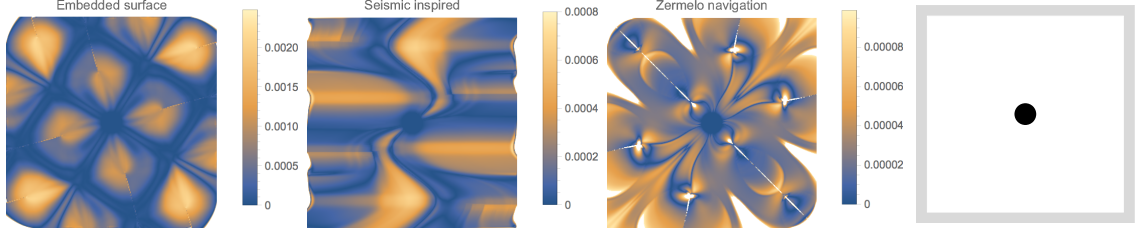


Figure 12: (I,II,III) Numerical error using the second order scheme for the three tests, with resolution  $n = 273$ ,  $273$  and  $375$  respectively. The numerical error is  $\mathcal{O}(h^2)$ , except on the pixels displayed white where it is typically  $\mathcal{O}(h)$ . These are located close to the boundary in the Riemannian cases, as well as along the cut locus and in other regions in the Rander case, which uses a non-causal scheme. (IV) Computational domain  $[-0.5, 0.5]^2$ . For the tests involving second-order finite differences, the exact solution is provided within the black disk of radius  $0.05$ , so as to mitigate the numerical error related to the point source singularity, and the gray region of width  $0.05$  along the boundary is excluded from the computation of the  $L^\infty$  error.

discretizations, solved via the single pass fast marching algorithm, and non-causal discretizations, solved via iterative methods. See Definition 2.1 of causality. These groups are analogous, in the context of distance computation on graphs, to Dijkstra’s algorithm and Bellman-Ford’s algorithm respectively, in which case the counterpart of causality is the positivity of the edge weights, see §1.

The test cases, taken from [70], are two dimensional and involve smooth metrics of Riemannian and Rander type with pronounced anisotropy. The good performance of the introduced discretization is confirmed, in terms of accuracy and CPU time, although it does not outperform e.g. the previous adaptive semi-Lagrangian schemes [47, 48] of the author.

Before proceeding, we would like to attract the attention of the reader to other qualities of the FM-VR1 scheme introduced in this paper, which are not put into light by this specific benchmark. Indeed, our discretization reliably handles much more extreme anisotropies than those considered here, see §3.1, to the point that we can approximate sub-Riemannian metrics, see §3.2. This is made possible by its full adaptivity, local and anisotropic. In addition, the Eulerian nature of the FM-VR1 makes it particularly simple to implement and cheap numerically, especially as dimension increases, in comparison with semi-Lagrangian methods which require to handle the complex combinatorics of the polyhedral neighborhood of each point, see (2). The numerical cost reduction is particularly evident in the five-dimensional sub-Riemannian experiments, see §3.2, which run approximately five times faster<sup>10</sup> using the FM-VR1 than with the semi-Lagrangian implementation described in [28].

The benchmarks presented here extend previous works of the author published in [48, 47].

### Causal discretizations.

- Fast Marching using Voronoi’s first reduction (FM-VR1), introduced in this paper.
- Fast-Marching using Lattice Basis Reduction (FM-LBR), and Fast-Marching using Adaptive Stencil Refinement (FM-ASR), introduced by the author in [47] and [48]. Like the FM-VR1, these are single-pass methods, which require a Cartesian grid and achieve their efficiency by the use of adaptive stencils built using techniques from lattice geometry. In

<sup>10</sup>Two dimensional instances do not show similar CPU time reduction, because numerical cost is in that case dominated by the handling of a priority queue, rather than by geometrical computations.

contrast with the FM-VR1, these are semi-Lagrangian discretizations. The FM-LBR applies to two and three-dimensional Riemannian metrics, whereas the FM-ASR applies to two-dimensional Finslerian metrics.

- The Monotone Acceptance Ordered Upwind method (MAOUM) [1] is a single pass semi-Lagrangian method using adaptive stencils. It differs from the FM-LBR and FM-ASR by its less sophisticated stencil construction, which produces large isotropic stencils, often at the expense of accuracy and complexity.
- Fast-Marching using 8-point stencils (FM-8) is the original semi-Lagrangian scheme [78] instantiated with non-adaptive stencils consisting of the 8 closest grid neighbors, see [38] for a three dimensional extension. This method is *non-consistent* for Riemannian metrics whose condition number exceeds  $(1 + \sqrt{2})^2$  in dimension  $d = 2$ , because its stencils lack the acuteness property [70]. Hence convergence towards the continuous problem solution *fails* as the grid is refined in the problem instances considered here, which involve more pronounced anisotropies. Nevertheless, the FM-8 is fast and its accuracy is surprisingly competitive at low grid sizes.

**Non causal discretizations.** We use Adaptive Gauss Siedel Iterations (AGSI<sup>11</sup>) to solve the following discretizations of the eikonal equation. Indeed, the single pass fast marching method is not applicable, since they lack the causality property, see Definition 2.1. We also report some computation times obtained by limiting the front width to 5 pixels.

- Rander Distances using Voronoi’s First Reduction (RD-VR1), introduced in this paper.
- The Finite Element discretization (FE) of [10], a semi-Lagrangian discretization using non-adaptive stencils extracted from a triangulation of the domain, here by half-squares.

**Remark 3.2** (Experiment setup, possible sources of bias). *All CPU times obtained using a single thread. CPU times are empirical data, only indicative of general performance. CPU times for the FM-VR1 and RD-VR1 obtained on a 2.7GHz core i7 laptop, whereas CPU times for the other methods were copied from previous works [47, 48] and obtained using a 2.4 GHz core 2 duo laptop. Numerical errors are with respect to a solution computed with the proposed algorithms on a fine grid of resolution  $2001 \times 2001$ , and then bilinearly interpolated. In the experiments involving the second order scheme, the boundary data is provided on a small disk of radius 0.05 rather than a single point source, and a boundary layer of width 0.05 is excluded from the  $L^\infty$  error computation, as illustrated on Figure 12 (iv).*

**Discussion of accuracy.** The numerical scheme introduced in this paper, the FM-VR1, belongs to the category of degenerate elliptic [54] discretizations of PDEs. This ensures excellent stability and robustness properties, by the discrete comparison principle see Theorem 2.3, but limits the accuracy that can be achieved: monotone finite difference schemes are at most first-order consistent for first order equations (such as the eikonal equation here considered), and at most second order for second order equations, see Theorem 4 in [54]. In this paragraph, we discuss some techniques aiming at improving the accuracy of eikonal equation solvers, that have been proposed in the literature.

---

<sup>11</sup>Note that in [47, 48] the acronym AGSI refers to discretization FE of [10]. This denomination is not proper, since it ties the PDE discretization with the numerical solver used for the discrete system of equations.

First, let us point out that the observed numerical error is  $\mathcal{O}(h)$  in typical instances, see Figures 10 and 11, despite the established rate of convergence being only  $\mathcal{O}(\sqrt{h})$ , see Theorem 1.3. In order to further reduce the numerical error, one possibility that we consider here is to introduce second order accurate approximations of the directional derivatives of the solution, in the spirit of the Higher Accuracy Fast Marching Method (HAFMM) introduced in [69]:

$$\langle du(\mathbf{p}), \dot{\mathbf{e}} \rangle = \frac{1}{h}(u(\mathbf{p} + h\dot{\mathbf{e}}) - u(\mathbf{p})) - \frac{1}{2h}(u(\mathbf{p}) - 2u(\mathbf{p} + h\dot{\mathbf{e}}) + u(\mathbf{p} + 2h\dot{\mathbf{e}})) + \mathcal{O}(h^2). \quad (41)$$

This expression replaces the first order finite differences used in the FM-VR1 (16), in some occurrences only and with enough caution and failsafe policies. Indeed, it breaks the monotony of the scheme. In order to preserve the convergence guarantees, the numerical method should fall back to the original first order upwind difference  $(u(\mathbf{p} + h\dot{\mathbf{e}}) - u(\mathbf{p}))/h$  whenever the second order corrective term exceeds the expected  $\mathcal{O}(h)$  magnitude<sup>12</sup>. This ensures that the solution  $U_h^*$  to the modified scheme obeys the original monotone scheme  $\mathcal{F}_h$  defined in (16) up to an  $\mathcal{O}(h)$  error:  $\mathcal{F}_h U_h^* = 1 + \mathcal{O}(h)$ . This implies global estimate  $U_h^* = U_h + \mathcal{O}(h)$ , and therefore the convergence of  $U_h^*$ , by the discrete comparison principle, see Theorem 2.3. The improved convergence rate of  $U_h^*$ , in comparison with the original scheme solution  $U_h$ , is only (often) observed numerically, and is not established by the analysis. We refer to [33] for a detailed analysis of related techniques in the (more complex) context of second-order PDEs, in particular of the Monge-Ampere equation, using the concepts of filtered-scheme and of quasi-monotone discretization.

The solutions to eikonal equations typically have non-smooth singularities, which require special treatment if second-order convergence is to be achieved, especially in the  $L^\infty$  norm.

- The *cut locus* is the collection of points  $\mathbf{p} \in \Omega$  reached by more than one minimal geodesic. The solution  $u$  is non-differentiable on this set, denoted by  $\mathcal{C}$  and which is typically a union of  $(d-1)$ -dimensional manifolds. Fortunately, because the minimal paths for the addressed problem do not cross  $\mathcal{C}$  (geodesics loose optimality when they do so), the numerical error associated with this singularity does not excessively pollute the rest of the numerical solution.

An  $\mathcal{O}(h)$  numerical error is in principle to be expected at discretization points whose stencil goes accross the non-differentiability set  $\mathcal{C}$  of the solution. This is observed in the Rander test case, but surprisingly not in the Riemannian test cases where an  $\mathcal{O}(h^2)$  only error is observed, see Figures 10 and 12. This empirical good surprise can be attributed to the limited diffusivity of the chosen discretization, see the discussion in [71].

- *Point source singularities.* In the considered test cases, as in many applications, the null boundary condition  $u(\mathbf{p}_*) = 0$  is imposed at an isolated seed point, rather than on a full domain's boundary as in our theoretical results §1. This produces a non-differentiable singularity with dominant term  $\mathbf{q} \mapsto \|\mathbf{p}_* - \mathbf{q}\|_{\mathcal{M}(\mathbf{p}_*)}$  at the source point  $\mathbf{p}_*$ . Similar effects are encountered at the corners of obstacles [59]. Factoring techniques, additive or multiplicative, incorporate corrective terms in the numerical scheme based on the analytic expression of the singularity [46]. Another approach, used in our experiments on second-order accuracy, is to pre-compute the solution  $u$  on a finer grid in a small neighborhood of the singularity.

---

<sup>12</sup>The implementation details of the fast marching method suggest additional restrictions, discussed in [69], such as requiring upwinding  $u(\mathbf{p} + 2\dot{\mathbf{e}}) \leq u(\mathbf{p})$ , and limiting the use of second order finite differences to a recomputation stage executed when a point of the front propagation is *accepted* and its value is frozen.

- *Outflow boundary solutions.* In the considered test cases, as in many applications, outflow boundary conditions are imposed on part of the domain’s boundary, rather than null Dirichlet conditions as in our theoretical results §1. As a result, some of the problem’s minimal paths are the (non-smooth) concatenation of usual geodesics and of parts of the domain’s boundary. A similar issue occurs with non-constant Dirichlet boundary conditions unless they are  $K$ -Lipschitz with respect to the metric [10], with  $K < 1$ . This phenomenon limits the smoothness of the problem solution and raises numerical difficulties.

In our experiments using the second order scheme, we exclude a small band along  $\partial\Omega$  from the  $L^\infty$  error computation, where this phenomenon causes the error to be  $\mathcal{O}(h)$  instead of  $\mathcal{O}(h^2)$ . This is sufficient in the Riemannian test cases, but not in the Rander case where the non-causality of the scheme causes the error to diffuse more inside the domain, see Figure 12.

Finally, let us mention multi-stencil schemes [36] and other variants [26] of the fast marching method, which improve accuracy by the use of stencils wider than strictly necessary, featuring in particular the diagonal directions of the discretization grid.

## 4 Conclusion

In this paper, we introduced a new discretization of anisotropic eikonal equations on Cartesian grids. The discretization makes use of adaptive stencils built using a tool from discrete geometry: Voronoi’s first reduction of quadratic forms.

A convergence proof is provided, with convergence rates, in the setting of Riemannian metrics, but also of sub-Riemannian and of asymmetric Rander metrics. Numerical experiments show that the method is particularly suitable for problems involving strong anisotropy, such as Riemannian tensor condition numbers of  $\approx 10^2$  and more, on domains of dimension  $2 \leq d \leq 5$ .

Future directions of research include designing causal discretizations for non-quadratic Hamiltonians, addressing point sets more general than Cartesian grids, and developing applications to motion planning and image segmentation.

**Acknowledgement.** The author thanks Da-Chen<sup>13</sup>, Jorg Portegies<sup>14</sup> and Erik Bekkers<sup>15</sup>, for careful testing, bug-fixing, and feedback on the numerical codes.

## References

- [1] Ken Alton and Ian M Mitchell. An ordered upwind method with precomputed stencil and monotone node acceptance for solving static convex Hamilton-Jacobi equations. *Journal of Scientific Computing*, 51(2):313–348, 2012.
- [2] David Bao, Colleen Robles, Zhongmin Shen, and others. Zermelo navigation on Riemannian manifolds. *Journal of Differential Geometry*, 66(3):377–435, 2004.
- [3] M Bardi and I Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhauser, 1997.

---

<sup>13</sup>Post-Doctoral researcher at University Paris-Dauphine.

<sup>14</sup> Ph.D. student under the direction of R. Duits at TU/e University, Eindhoven.

<sup>15</sup>Post-Doctoral researcher at TU/e University, Eindhoven.

- [4] Guy Barles. An Introduction to the Theory of Viscosity Solutions for First-Order Hamilton–Jacobi Equations and Applications. In *Hamilton–Jacobi Equations: Approximations, Numerical Analysis and Applications*, pages 49–109. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [5] Erik J Bekkers, Remco Duits, Alexey Mashtakov, and Gonzalo R Sanguinetti. A PDE approach to data-driven sub-riemannian geodesics in SE (2). *SIAM Journal on Imaging Sciences*, 8(4):2740–2770, 2015.
- [6] Jean-David Benamou, Francis Collino, and Jean-Marie Mirebeau. Monotone and Consistent discretization of the Monge-Ampere operator. *Mathematics of computation*, September 2015.
- [7] Fethallah Benmansour and Laurent D. Cohen. Tubular structure segmentation based on minimal path method and anisotropic enhancement. *International Journal of Computer Vision*, 92(2):192–210, 2011.
- [8] Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 2. Athena scientific Belmont, MA, 1995.
- [9] J Frederic Bonnans, Elisabeth Ottenwaelter, and Hasnaa Zidani. A fast algorithm for the two dimensional HJB equation of stochastic control. Technical report, 2004.
- [10] 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.
- [11] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. Weighted distance maps computation on parametric three-dimensional manifolds. *Journal of Computational Physics*, 225(1):771–784, July 2007.
- [12] Simone Cacace, Emiliano Cristiani, and Maurizio Falcone. A local ordered upwind method for Hamilton–Jacobi and Isaacs equations. *IFAC Proceedings Volumes*, 44(1):6800–6805, 2011.
- [13] Simone Cacace, Emiliano Cristiani, and Maurizio Falcone. Can Local Single-Pass Methods Solve Any Stationary Hamilton–Jacobi–Bellman Equation? *SIAM Journal on Scientific Computing*, 36(2):A570–A587, 2014.
- [14] Marcel Campen, Martin Heistermann, and Leif Kobbelt. Practical Anisotropic Geodesy. *Computer Graphics Forum*, 32(5):63–71, August 2013.
- [15] Constantin Carathéodory. Partial differential equations of the first order. *Calculus of Variations and Partial Differential Equations of the First Order*, 1965.
- [16] Vlastislav Cervený. *Seismic ray theory*. Cambridge university press, 2005.
- [17] Da Chen. *New Minimal Path Models for Tubular Structure Extraction and Image Segmentation*. PhD thesis, University Paris-Dauphine, September 2016.
- [18] Da Chen, Laurent D. Cohen, and Jean-Marie Mirebeau. Vessel extraction using anisotropic minimal paths and path score. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1570–1574, paris, France, 2014. IEEE.

- [19] Da Chen, Jean-Marie Mirebeau, and Laurent D. Cohen. Global Minimum for a Finsler Elastica Minimal Path Approach. *International Journal of Computer Vision*, 122(3):458–483, 2017.
- [20] J H Conway and N J A Sloane. Low-Dimensional Lattices. III. Perfect Forms. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 418(1854):43–80, July 1988.
- [21] 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.
- [22] Michael G Crandall, Lawrence C Evans, and P L Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 282(2):487–502, 1984.
- [23] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1):1–42, 1983.
- [24] 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.
- [25] Emiliano Cristiani. A fast marching method for Hamilton-Jacobi equations modeling monotone front propagations. *Journal of Scientific Computing*, 39(2):189–205, 2009.
- [26] Per-Erik Danielsson and Qingfen Lin. A modified fast marching method. *Image Analysis*, 2003.
- [27] Joe Dellinger and William Symes. Anisotropic finite-difference traveltimes using a Hamilton-Jacobi solver. In *SEG Technical Program Expanded Abstracts 1997*, pages 1786–1789. Society of Exploration Geophysicists, 1997.
- [28] Remco Duits, Stephan PL Meesters, Jean-Marie Mirebeau, and Jorg M Portegies. Optimal paths for variants of the 2D and 3D Reeds-Shepp car with applications in image analysis. *Journal of Mathematical Imaging and Vision*, pages 1–33, 2018.
- [29] David W S Eaton. Finite Difference Traveltime Calculation For Anisotropic Media. *Geophysical Journal International*, 114(2):273–280, August 1993.
- [30] Lawrence C Evans. *Partial Differential Equations*. American Mathematical Soc., 2010.
- [31] Maurizio Falcone. A numerical approach to the infinite horizon problem of deterministic control theory. *Applied Mathematics and Optimization*, 15(1):1–13, 1987.
- [32] Jérôme Fehrenbach and Jean-Marie Mirebeau. Sparse Non-negative Stencils for Anisotropic Diffusion. *Journal of Mathematical Imaging and Vision*, pages 1–25, 2013.
- [33] Brittany D Froese and A M Oberman. Convergent Filtered Schemes for the Monge-Ampère Partial Differential Equation. *SIAM Journal on Numerical Analysis*, 51(1):423–444, January 2013.

- [34] Roberto Gonzalez and Edmundo Rofman. On deterministic control problems: An approximation procedure for the optimal cost I. The stationary problem. *SIAM Journal on Control and Optimization*, 23(2):242–266, 1985.
- [35] Song Han, Wei Zhang, and Jie Zhang. Calculating qP-wave traveltimes in 2-D TTI media by high-order fast sweeping methods with a numerical quartic equation solver. *Geophysical Journal International*, 210(3):1560–1569, 2017.
- [36] M Sabry Hassouna and A 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.
- [37] Chia-Yu Hu and Chun-Hung Lin. Reverse ray tracing for transformation optics. *Optics Express*, 23(13):17622–17637, 2015.
- [38] S Jbabdi, P Bellec, R Toro, J Daunizeau, M Pélérini-Issac, and H Benali. Accurate anisotropic fast marching for diffusion-based geodesic tractography. *Journal of Biomedical Imaging*, 2008(1-2):2–12, January 2008.
- [39] Seongjai Kim. On eikonal solvers for anisotropic traveltimes. In *SEG Technical Program Expanded Abstracts 1999*, pages 1875–1878. Society of Exploration Geophysicists, 1999.
- [40] Seongjai Kim. An  $O(N)$  Level Set Method for Eikonal Equations. *SIAM Journal on Scientific Computing*, 22(6):2178–2193, 2001.
- [41] R Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences*, 95(15):8431–8435, July 1998.
- [42] J L Lagrange. *Recherches d’arithmétique*. C.F. Voss, 1775.
- [43] Haiqiang Lan, Jingyi Chen, and Zhongjie Zhang. A fast sweeping scheme for calculating P wave first-arrival travel times in transversely isotropic media with an irregular surface. *Pure and Applied Geophysics*, 171(9):2199–2208, 2014.
- [44] 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.
- [45] I Lecomte. Finite difference calculation of first traveltimes in anisotropic media 1. *Geophysical Journal International*, 113(2):318–342, 1993.
- [46] 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.
- [47] Jean-Marie Mirebeau. Anisotropic Fast-Marching on cartesian grids using Lattice Basis Reduction. *SIAM Journal on Numerical Analysis*, 52(4):1573–1599, January 2014.
- [48] Jean-Marie Mirebeau. Efficient fast marching with Finsler metrics. *Numerische Mathematik*, 126(3):515–557, 2014.
- [49] Jean-Marie Mirebeau. Minimal stencils for discretizations of anisotropic PDEs preserving causality or the maximum principle. *SIAM Journal on Numerical Analysis*, 54(3):1582–1611, 2016.

- [50] Jean-Marie Mirebeau. Fast-marching methods for curvature penalized shortest paths. *Journal of Mathematical Imaging and Vision*, pages 1–32, 2017.
- [51] Richard Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. American Mathematical Soc., August 2006.
- [52] T J Moser. Shortest path calculation of seismic rays. *Geophysics*, 56(1):59–67, 1991.
- [53] Phong Q. Nguyen and Damien Stehlé. Low-dimensional lattice basis reduction revisited. In Ducan Buell, editor, *ANTS*, pages 338–357. Springer, 2004.
- [54] 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.
- [55] Stanley Osher. A level set formulation for the solution of the Dirichlet problem for Hamilton–Jacobi equations. *SIAM Journal on Mathematical Analysis*, 24(5):1145–1152, 1993.
- [56] Marco A Perez and John C Bancroft. Finite-difference anisotropic traveltimes and raypaths. *CREVES research report*, (13):445–461, 2001.
- [57] Jean Petitot. The neurogeometry of pinwheels as a sub-Riemannian contact structure. *Journal of Physiology-Paris*, 97(2-3):265–309, March 2003.
- [58] Gabriel Peyré, Mickael Péchaud, Renaud Keriven, and Laurent D. Cohen. *Geodesic Methods in Computer Vision and Graphics*.
- [59] Dongping Qi and Alexander Vladimirsky. Corner cases, singularities, and dynamic factoring. *arXiv preprint arXiv:1801.04322*, 2018.
- [60] Jianliang Qian and William W Symes. Finite-difference quasi-P traveltimes for anisotropic media. *Geophysics*, 67(1):147–155, 2002.
- [61] Gunnar Randers. On an Asymmetrical Metric in the Four-Space of General Relativity. *Physical Review*, 59(2):195–199, January 1941.
- [62] Anand Rangarajan and Karthik Gurumoorthy. A fast eikonal equation solver using the schrödinger wave equation. Technical report, 2011.
- [63] Nicholas Rawlinson, Jurg Hauser, and Malcolm Sambridge. Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in Geophysics*, 49:203–273, 2008.
- [64] Mael Rouxel-Labbé, Mathijs Wintraecken, and J D Boissonnat. Discretized riemannian delaunay triangulations. *Procedia engineering*, 163:97–109, 2016.
- [65] Elisabeth Rouy and Agnès Tourin. A Viscosity Solutions Approach to Shape-From-Shading. *SIAM Journal on Numerical Analysis*, 29(3):867–884, July 1992.
- [66] Gonzalo Sanguinetti, Erik Bekkers, Remco Duits, Michiel HJ Janssen, Alexey Mashtakov, and Jean-Marie Mirebeau. Sub-Riemannian fast marching in SE (2). In A Pardo and J Kittler, editors, *Iberoamerican Congress on Pattern Recognition*, pages 366–374, Cham, 2015. Springer, Springer, Cham.

- [67] A Schürmann. Computational geometry of positive definite quadratic forms. *University Lecture Series*, 2009.
- [68] Eduard Selling. Ueber die binären und ternären quadratischen Formen. *Journal für die Reine und Angewandte Mathematik*, 77:143–229, 1874.
- [69] J. A. Sethian. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. 1999.
- [70] J. A. Sethian and A. Vladimirsky. Ordered upwind methods for static Hamilton-Jacobi equations. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20):11069–11074 (electronic), 2001.
- [71] 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.
- [72] James A. Sethian and Alexander Vladimirsky. Ordered upwind methods for static Hamilton-Jacobi equations: theory and algorithms. *SIAM Journal on Numerical Analysis*, 41(1):325–363, 2003.
- [73] Alex Shum, Kirsten Morris, and Amir Khajepour. Convergence Rate for the Ordered Upwind Method. *arXiv.org*, (3):889–913, January 2016.
- [74] Mathieu Sikirić, Achill Schürmann, and Frank Vallentin. Classification of eight-dimensional perfect forms. *Electronic Research Announcements of the American Mathematical Society*, 13(3):21–32, 2007.
- [75] Alexey Stovas and Tariq Alkhalifah. A new traveltime approximation for TI media. *Geophysics*, 77(4):C37–C42, 2012.
- [76] Petter Strandmark, Johannes Ulen, Fredrik Kahl, and Leo Grady. Shortest Paths with Curvature and Torsion. In *2013 IEEE International Conference on Computer Vision (ICCV)*, pages 2024–2031. IEEE, 2013.
- [77] YHR Tsai, LT Cheng, and S Osher. Fast sweeping algorithms for a class of Hamilton-Jacobi equations. *SIAM Journal on Numerical Analysis*, 2004.
- [78] J.N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. 40(9):1528–1538, September 1995.
- [79] S R S Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Communications on Pure and Applied Mathematics*, 20(2):431–455, May 1967.
- [80] John Vidale. Finite-difference calculation of travel times. *Bulletin of the Seismological Society of America*, 78(6):2062–2076, 1988.
- [81] Alexander Vladimirsky. Label-setting methods for Multimode Stochastic Shortest Path problems on graphs. *Mathematics of Operations Research*, 33(4):821–838, 2008.
- [82] Alexander Boris Vladimirsky. *Fast methods for static Hamilton-Jacobi Partial Differential Equations*. PhD thesis, May 2001.

- [83] G Voronoi. *Sur quelques propriétés des formes quadratiques positives parfaites*. J. reine angew. Math, 1908.
- [84] Fang Yang and Laurent D. Cohen. Geodesic distance and curves through isotropic and anisotropic heat equations on images and surfaces. *Journal of Mathematical Imaging and Vision*, 55(2):210–228, 2016.
- [85] H Zhao. A fast sweeping method for eikonal equations. *Mathematics of computation*, 2005.

## A Single pass solution to causal schemes

We prove in this appendix that the system of equations associated with a degenerate elliptic and causal scheme, in the sense of Definition 2.1, can be solved in a single pass over the domain (under the mild additional assumption that the Gauss-Siedel updates are well defined), using a variant of Dijkstra’s algorithm. This property guarantees fast and predictable run times for the numerical solver, quasi-linear in the number of unknowns, independently of the problem instance. It is the sole purpose of the causality property from Definition 2.1, which in particular is *not* used in the convergence analysis of the numerical scheme solutions §2, §D, §E.

Concepts and arguments closely related with Definition A.1 and Proposition A.2 can be found in [78, 41, 72, 8]. See in particular the definition of the causality property in [72], and the proof that a stochastic shortest path problem on a graph can be solved in a single pass if there exists an optimal consistently improving policy in [8] chapter 2.

Throughout this section, we fix a finite set  $X$ , and denote by

$$\mathbb{U} := ]-\infty, +\infty]^X$$

the set of functions on  $X$  with either real or  $+\infty$  values. For any  $u \in \mathbb{U}$ , and any  $t \in [-\infty, +\infty]$ , we define the function  $u^{<t} \in \mathbb{U}$  (resp.  $u^{\leq t} \in \mathbb{U}$ ) by replacing all values above (resp. strictly above) the threshold  $t$  with  $+\infty$ . Explicitly, for all  $x \in X$ ,

$$u^{<t}(x) := \begin{cases} u(x) & \text{if } u(x) < t, \\ +\infty & \text{otherwise.} \end{cases}$$

We introduce, in the next definition, the concept of Monotone and Causal operator on the set  $X$ . The fast marching method, presented in abstract form in Algorithm 1, computes in a finite number of steps a fixed point of such an operator, see Proposition A.2. Note that we distinguish the concept of numerical scheme  $F$ , where the objective is to find a root:  $Fu = 0$ , and of operator  $\Lambda$ , where the objective is to find a fixed point:  $\Lambda u = u$ , denoting by  $u$  the unknown.

**Definition A.1.** *An operator on the set  $X$  is a map  $\Lambda : \mathbb{U} \rightarrow \mathbb{U}$ . It is said*

- Monotone, *iff for all  $u, v \in \mathbb{U}$  such that  $u \preceq v$  one has  $\Lambda u \preceq \Lambda v$ .*
- Causal, *iff for all  $u, v \in \mathbb{U}$  and  $t \in [-\infty, +\infty]$  s.t.  $u^{<t} = v^{<t}$  one has  $(\Lambda u)^{\leq t} = (\Lambda v)^{\leq t}$ .*

The Causality property of Definition A.1 is equivalent to (23), only stated in a more formal manner. Here and below, for any  $u, v \in \mathbb{U}$ , the notation  $u \preceq v$  means  $u(x) \leq v(x)$  for all  $x \in X$ . By convention, the minimum of the empty set is defined as  $+\infty$ .

---

**Algorithm 1** Abstract fast marching method

---

**Initialization**  $n \leftarrow 0$ ,  $t_0 \leftarrow -\infty$ , and  $u_0 \leftarrow +\infty$  on  $X$ .

**While**  $t_n < \infty$  **do**

$$u_{n+1} \leftarrow \Lambda u_n^{\leq t_n}, \quad t_{n+1} := \min\{u_{n+1}(x); x \in X, u_{n+1}(x) > t_n\}, \quad n \leftarrow n + 1.$$

---

**Proposition A.2.** *Let  $\Lambda$  be a monotone and causal operator on a finite set  $X$ , to which the fast marching method is applied, see Algorithm 1. Then for some  $N < \#(X)$  one has  $t_N = +\infty$ , so that the algorithm terminates, and one has  $\Lambda u_N = u_N$ . Furthermore,  $u_n \succeq u_{n+1}$  and  $u_n^{\leq t_n} = u_{n+1}^{\leq t_n}$  for each  $0 \leq n \leq N$ .*

*Proof.* Note that  $t_n \leq t_{n+1}$  for all  $0 \leq n < N$ , by definition of  $t_{n+1}$ . We begin with the proof of the last two announced properties, by induction over  $n \in \{0, \dots, N\}$ .

*Claim:*  $u_n \succeq u_{n+1}$ . Initialization holds since  $u_0 \equiv +\infty \succeq u_1$ . Then, for  $n \geq 1$ , we obtain by induction and Monotony of the operator  $\Lambda$

$$(u_{n-1} \succeq u_n \text{ and } t_{n-1} \leq t_n) \Rightarrow u_{n-1}^{\leq t_{n-1}} \succeq u_n^{\leq t_n} \Rightarrow \Lambda u_{n-1}^{\leq t_{n-1}} \succeq \Lambda u_n^{\leq t_n} \Leftrightarrow u_n \succeq u_{n+1}.$$

*Claim:*  $u_n^{\leq t_n} = u_{n+1}^{\leq t_n}$ . Initialization follows from  $t_0 = -\infty$ . Then, for  $n \geq 1$ , we obtain

$$u_{n-1}^{\leq t_{n-1}} = (u_{n-1}^{\leq t_{n-1}})^{< t_n}, \quad u_n^{\leq t_{n-1}} = u_n^{\leq t_n} = (u_n^{\leq t_n})^{< t_n},$$

by definition of  $t_n$ . Therefore, by induction and Causality of the operator  $\Lambda$ , we obtain

$$(\Lambda u_{n-1}^{\leq t_{n-1}})^{\leq t_n} = (\Lambda u_n^{\leq t_n})^{\leq t_n} \Leftrightarrow u_n^{\leq t_n} = u_{n+1}^{\leq t_n}.$$

Conclusion of the proof. The sets  $X_n := \{x \in X; u_n(x) \leq t_n\}$  are strictly increasing for inclusion as  $n$  increases, by the above, and are included in the finite domain  $X$ . The algorithm therefore terminates, in at most  $\#(X)$  iterations. In the last step we have  $t_N = +\infty$  and thus  $u_N = u_N^{\leq t_N} = u_{N+1}^{\leq t_N} = u_{N+1} = \Lambda u_N$ .  $\square$

**Remark A.3** (Terminology). *The literature [78] classically distinguishes several classes of points  $x \in X$ , at each step  $n$  of the fast marching algorithm: Frozen if  $u_n(x) \leq t_n$ , Considered if  $t_n < u_n(x) < \infty$ , and Far if  $u_n(x) = +\infty$ . A point  $x \in X$  is Accepted in step  $n$  if it is Frozen in step  $n$  but was not Frozen in step  $n - 1$ .*

The next proposition defines an operator  $\Lambda$  as the local Gauss-Siedel update associated with a numerical scheme  $F$ . If the scheme is Degenerate elliptic then the resulting operator is Monotone, and Causality also transfers from the scheme to the operator. This property is known regarding Degenerate ellipticity/Monotony [54], but appears to be new for Causality.

**Proposition A.4.** *Let  $F$  be a numerical scheme on  $X$ , which is degenerate elliptic in the sense of Definition 2.1. For each  $u \in \mathbb{U}$ , and each  $x \in X$ , the mapping*

$$\lambda \in \mathbb{R} \mapsto f(\lambda) := F(x, \lambda, (\lambda - u(y))_{y \neq x}) \tag{42}$$

*is non-decreasing. We assume that either (i) there exists a unique  $\lambda \in \mathbb{R}$  such that  $f(\lambda) = 0$ , and we set  $\Lambda u(x) := \lambda$ , or (ii) for all  $\lambda \in \mathbb{R}$  one has  $f(\lambda) < 0$ , and we set  $\Lambda u(x) := +\infty$ .*

*Then  $\Lambda$  is a monotone operator on  $X$ , in the sense of Definition A.1. In addition, if the scheme  $F$  is causal, then the operator  $\Lambda$  is causal, in the sense of Definitions 2.1 and A.1 respectively.*

*Proof.* The fact that (42) is non-decreasing directly follows from the degenerate ellipticity of the scheme  $F$ .

*Monotony of the operator  $\Lambda$ .* Let  $u, v \in \mathbb{U}$ , and let  $x \in X$  be arbitrary. If  $u \preceq v$ , then

$$F(x, \lambda, (\lambda - u(y))_{y \neq x}) \geq F(x, \lambda, (\lambda - v(y))_{y \neq x})$$

for all  $\lambda \in \mathbb{R}$ , by monotony of scheme  $F$ . Thus  $\Lambda u(x) \leq \Lambda v(x)$ , as announced

*Causality of the operator  $\Lambda$ .* Let  $u, v \in \mathbb{U}$ , let  $t \in [-\infty, \infty]$ , and let  $x \in X$ . If  $u^{<t} \equiv v^{<t}$ , then

$$F(x, \lambda, \max\{0, \lambda - u(y)\}_{y \neq x}) = F(x, \lambda, \max\{0, \lambda - v(y)\}_{y \neq x})$$

for all  $\lambda \leq t$ , since indeed  $\max\{0, \lambda - u(y)\} = \max\{0, \lambda - v(y)\}$ . Assuming the Causality of scheme  $F$ , this implies  $F(x, \lambda, (\lambda - u(y))_{y \neq x}) = F(x, \lambda, (\lambda - v(y))_{y \neq x})$  for all  $\lambda \leq t$ . Thus these two functions either (i) have a common root  $\lambda \leq t$ ,  $\lambda = \Lambda u(x) = \Lambda v(x)$ , or (ii) have (possibly distinct) roots  $\Lambda u(x) > t$ ,  $\Lambda v(x) > t$ . Finally  $(\Lambda u)^{\leq t} = (\Lambda v)^{\leq t}$ , as announced.  $\square$

**Remark A.5** (Complexity of the fast marching algorithm). *The fast marching algorithm requires to evaluate  $\Lambda u_n^{\leq t_n}$  and  $\Lambda u_{n+1}^{\leq t_{n+1}}$  in consecutive iterations of rank  $n$  and  $n+1$ . Observe that the maps  $u_{n+1}^{\leq t_{n+1}}$  and  $u_n^{\leq t_n} = u_{n+1}^{\leq t_n}$  differ only at the points  $A_{n+1} = \{x \in X; u_{n+1}(x) = t_{n+1}\}$  which have been Accepted in this iteration (typically a single point). Therefore the recomputation of the operator  $\Lambda$  can be restricted to the points  $x \in X$  such that  $\Lambda u(x)$  depends on  $u(y)$  for some  $y \in A_{n+1}$ . Since each point is Accepted only once, and assuming that the operator  $\Lambda$  is defined as the Gauss-Siedel update of a numerical scheme with a bounded stencil, the overall number of elementary local evaluations of the operator  $\Lambda$  (regarded as of unit complexity) in the course of the fast marching algorithm is  $\mathcal{O}(KM)$  where  $K = \#(X)$  and  $M$  is the average stencil size.*

*The overall optimal complexity of the fast marching algorithm is  $\mathcal{O}(KM + K \ln K)$ , where the second term comes from the cost of maintaining a Fibonacci heap of all the Considered points, see e.g. Remark 1.7 in [48].*

*A linear complexity  $\mathcal{O}(KM)$  can be achieved under a strengthened and quantitative variant of the causality property, using a variant of Dial's algorithm, see [78, 81].*

## B Tensor decomposition based on Voronoi's first reduction

In this appendix, we present a tool from the field of lattice geometry - Voronoi's first reduction [83] - which allows to efficiently compute the tensor decomposition (4) maximizing (5), and is thus a key ingredient of our numerical scheme. §B.1 presents a glimpse of the general theory, and the proof of Proposition 1.1. §B.2 describes Selling's algorithm, a practical tensor decomposition algorithm in dimension  $d \in \{2, 3\}$ . Python and C++ codes for computing this tensor decomposition in dimension  $d \in \{2, 3, 4, 5\}$  are available on the author's webpage<sup>16</sup>.

### B.1 Voronoi's first reduction of a quadratic form

The modern presentation of Voronoi's theory [67] describes the reduction of a quadratic form  $D \in S^{++}(\mathbb{E}^*)$  as a linear program  $\mathcal{L}(D)$  over Ryskov's convex and locally finite polyhedron  $\mathcal{P} \subseteq S^{++}(\mathbb{E})$ . These objects are defined as follows:

$$\mathcal{P} := \{M \in S^{++}(\mathbb{E}); \forall \mathbf{e} \in \mathbb{L} \setminus \{0\}, \|\mathbf{e}\|_M \geq 1\}, \quad \mathcal{L}(D) := \inf_{M \in \mathcal{P}} \text{Tr}(MD). \quad (43)$$

---

<sup>16</sup>[github.com/Mirebeau/](https://github.com/Mirebeau/)

Introducing the duality bracket  $\langle\langle M, D \rangle\rangle := \text{Tr}(MD)$  between  $S(\mathbb{E})$  and  $S(\mathbb{E}^*)$ , and observing that  $\|\dot{\mathbf{e}}\|_M^2 = \langle\langle M, \dot{\mathbf{e}} \otimes \dot{\mathbf{e}} \rangle\rangle$ , one can rephrase the linear program  $\mathcal{L}(D)$  as follows

$$\text{minimize } \langle\langle M, D \rangle\rangle \quad \text{subject to } \langle\langle M, \dot{\mathbf{e}} \otimes \dot{\mathbf{e}} \rangle\rangle \geq 1 \text{ for all } \dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}. \quad (44)$$

The vertices (resp. edges) of a polyhedron are its 0-dimensional (resp. 1-dimensional) facets.

**Theorem** (Voronoi, see [67]). *The linear problem  $\mathcal{L}(D)$  is feasible, for any  $D \in S^{++}(\mathbb{E}^*)$ , in the sense that the set of minimizers is non-empty and compact. In addition, the convex polytope  $\mathcal{P} \subseteq S(E)$  has a finite number of equivalence classes of vertices under the action of  $\text{GL}(\mathbb{L})$ .*

In dimension  $d \leq 3$ , one can solve (44) using Selling's algorithm as described in §B.2. In dimension  $d > 3$  one can implement a simplex-like method relying on the classification of the vertices of the convex polyhedron  $\mathcal{P}$ , which is known in dimension  $d \leq 8$  [20, 74]. These vertices are classically referred to as perfect quadratic forms and are studied for their relationship with the densest periodic sphere packings [67].

The Karush-Kuhn Tucker optimality relations for the linear program (44) determine a decomposition of the input tensor  $D$ , see the next proposition, involving directions with integer entries corresponding to the active constraints.

**Proposition B.1.** *The Kuhn-Tucker optimality conditions for the linear optimization problem  $\mathcal{L}(D)$ ,  $D \in S^{++}(\mathbb{E}^*)$ , imply that there exists  $(\rho_i, \dot{\mathbf{e}}_i)_{i=1}^{d'} \in (\mathbb{R}_+ \times \mathbb{L})^{d'}$ , where  $d' := d(d+1)/2$ , such that*

$$D = \sum_{1 \leq i \leq d'} \rho_i \dot{\mathbf{e}}_i \otimes \dot{\mathbf{e}}_i. \quad (45)$$

Furthermore there exists  $C = C(d)$  such that for any  $1 \leq i \leq d'$

$$\|\dot{\mathbf{e}}_i\| \leq C\mu(D)^{d-1}, \quad \text{where } \mu(D) := \sqrt{\|D\|\|D^{-1}\|}. \quad (46)$$

In dimension  $d = 3$ , one has the improved estimate  $\|\dot{\mathbf{e}}_i\| \leq C\mu(D)$ .

*Proof.* The first point (45) follows from the feasibility of the linear program  $\mathcal{L}(D)$ , and its formulation (44). The number  $d(d+1)/2$  of contributions in this decomposition is the number of independent entries in a  $d \times d$  symmetric matrix.

For proving the second point, the Euclidean space  $\mathbb{E}$  is identified with its dual, which gives meaning to the trace  $\text{Tr}(M)$  and determinant  $\det(M)$  of any  $M \in S(\mathbb{E})$ . Denote by  $(M_k)_{k=1}^K$  a representative of each equivalence class of vertices of  $\mathcal{P}$  under the action of  $\text{GL}(\mathbb{L})$ , see Theorem B.1. Let  $D \in S^{++}(\mathbb{E}^*)$  be arbitrary, and let  $M$  be the minimizer of  $\mathcal{L}(D)$ . Then  $M = A^T M_k A$  for some  $1 \leq k \leq K$  and some  $A \in \text{GL}(\mathbb{L})$ . Thus  $\det(M) \geq \min_{k=1}^K \det(M_k) =: c_\Delta > 0$ . On the other hand,  $\text{Tr}(M)\|D^{-1}\|^{-1} \leq \text{Tr}(MD) \leq \text{Tr}(\text{Id } D) \leq d\|D\|$ , by sub-optimality of  $\text{Id} \in \mathcal{P}$ , hence  $\text{Tr}(M) \leq d\mu(D)^2$ . For any  $1 \leq i \leq d'$  one has  $\|\dot{\mathbf{e}}_i\|_M = 1$ , since the corresponding constraint of the linear problem  $\mathcal{L}(D)$  is active, hence as announced

$$\|\dot{\mathbf{e}}_i\|^2 \leq \|M^{-1}\| \|\dot{\mathbf{e}}_i\|_M^2 = \|M^{-1}\| \leq \frac{\|M\|^{d-1}}{\det(M)} \leq \frac{\text{Tr}(M)^{d-1}}{c_\Delta} \leq \frac{(d\mu(D)^2)^{d-1}}{c_\Delta}.$$

The three dimensional improved estimate is established in [50], see Theorem 4.11, using dimension specific techniques.  $\square$

In order to conclude the proof of Proposition 1.1, we need to show that the linear program (44) is in duality with the problem of finding the tensor decomposition (4) which maximizes (5). We sketch the argument, which is based on the following computations

$$\mathcal{L}(D) = \inf_{M \succeq 0} \left( \langle D, M \rangle + \sup_{\lambda \geq 0} \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \lambda(\dot{\mathbf{e}})(1 - \langle \dot{\mathbf{e}} \otimes \dot{\mathbf{e}}, M \rangle) \right) \quad (47)$$

$$= \sup_{\lambda \geq 0} \left( \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \lambda(\dot{\mathbf{e}}) + \inf_{M \succeq 0} \langle D - \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \lambda(\dot{\mathbf{e}}) \dot{\mathbf{e}} \otimes \dot{\mathbf{e}}, M \rangle \right), \quad (48)$$

where the optimization variables are  $\lambda : \mathbb{L} \setminus \{0\} \rightarrow [0, \infty[$  a finitely supported and non-negative map, and  $M \in \mathbf{S}^+(\mathbb{E})$  a symmetric positive semi-definite matrix. Note that the “sup” in (47) equals 0 iff the constraint  $\langle \dot{\mathbf{e}} \otimes \dot{\mathbf{e}}, M \rangle \geq 1$  is satisfied for all  $\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}$ , and equals  $+\infty$  otherwise. Likewise the “inf” in (48) equals 0 iff the constraint  $D \succeq \sum_{\dot{\mathbf{e}} \in \mathbb{L} \setminus \{0\}} \lambda(\dot{\mathbf{e}}) \dot{\mathbf{e}} \otimes \dot{\mathbf{e}}$  is satisfied (equality holds at the optimum, see Proposition B.1), and equals  $-\infty$  otherwise. Voronoi’s theorem presented above shows that the linear program (44) is well posed, hence one can exchange the “sup” and “inf” in (47) and (48), and there is no duality gap.

## B.2 Selling’s algorithm

Selling’s algorithm [68, 21] allows in dimension  $d \leq 3$  to compute Voronoi’s first reduction (44), and the related tensor decomposition (45) which is at the foundation of our numerical scheme. This algorithm is used in all the numerical experiments presented in this paper. For that purpose, we need to introduce a few concepts from lattice geometry. The vertices (resp. edges) of a polyhedron are its 0-dimensional (resp. 1-dimensional) facets.

**Definition B.2.** A superbase of  $\mathbb{Z}^d$  is a  $(d+1)$ -tuple  $b = (\mathbf{v}_0, \dots, \mathbf{v}_d) \in (\mathbb{Z}^d)^{d+1}$  such that  $|\det(\mathbf{v}_1, \dots, \mathbf{v}_d)| = 1$  and  $\mathbf{v}_0 + \dots + \mathbf{v}_d = 0$ . One also defines  $M_b := \frac{1}{2} \sum_{i=0}^d \mathbf{v}_i \mathbf{v}_i^T \in \mathbf{S}^{++}(\mathbb{E})$ .

In dimension  $d \leq 3$ , the vertices of Ryskov’s polyhedron  $\mathcal{P} \subseteq \mathbf{S}^{++}(\mathbb{R}^d)$ , defined in (43), are the matrices  $M_b$ , where  $b = (\mathbf{v}_0, \dots, \mathbf{v}_d)$  is an arbitrary superbase of  $\mathbb{Z}^d$ . See [20, 74] for a similar classification, in dimension  $d \leq 8$ , of these vertices under the name of *perfect forms*. The vertices joined to  $M_b$  by an edge of the polyhedron  $\mathcal{P}$  are in dimension  $d = 2$  (resp.  $d = 3$ ) associated with the superbases

$$b' = (-\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_i - \mathbf{v}_j) \quad \left( \text{resp. } b' = (-\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k + \mathbf{v}_i, \mathbf{v}_l + \mathbf{v}_i) \right), \quad (49)$$

where  $\{i, j, k\}$  (resp.  $\{i, j, k, l\}$ ) is an arbitrary permutation of  $\{0, \dots, d\}$ . In addition one has

$$\text{Tr}(M_{b'} D) = \text{Tr}(M_b D) - C_d \langle \mathbf{v}_i, D \mathbf{v}_j \rangle, \quad (50)$$

where  $C_2 = 2$  and  $C_3 = 1$ . Selling’s algorithm, given an input matrix  $D$  and a superbase  $b = (\mathbf{v}_0, \dots, \mathbf{v}_d)$ , looks for indices  $0 \leq i < j \leq d$  such that  $\langle \mathbf{v}_i, D \mathbf{v}_j \rangle > 0$ . If any exists, then the superbase is replaced with (49) and the process is repeated. Otherwise the algorithm stops, and yields what is called a  $D$ -obtuse superbase, i.e.  $\langle \mathbf{v}_i, D \mathbf{v}_j \rangle \leq 0$  for all  $i \neq j$ .

Clearly, this algorithm is equivalent to a walk on the graph defined by the vertices and the edges of Ryskov’s polyhedron  $\mathcal{M}_d$ , in which the next vertex is a neighbor selected so as to reduce the objective function  $M \mapsto \text{Tr}(MD)$ . Since Voronoi’s first reduction (44) is a well-posed linear program, the process eventually ends, similarly to the classical simplex algorithm. The

vertex  $M_b$  associated with the eventual  $D$ -obtuse superbase  $b = (\mathbf{v}_0, \dots, \mathbf{v}_d)$  is optimal, and the corresponding Karush-Kuhn-Tucker relations read

$$D = - \sum_{0 \leq i < j \leq d} \langle \mathbf{v}_i, D\mathbf{v}_j \rangle \mathbf{e}_{ij} \otimes \mathbf{e}_{ij},$$

where  $\mathbf{e}_{ij} := \mathbf{v}_k^\perp$  in dimension  $d = 2$  and with  $\{i, j, k\} = \{0, 1, 2\}$ , and  $\mathbf{e}_{ij} := \mathbf{v}_k \times \mathbf{v}_l$  in dimension  $d = 3$  and with  $\{i, j, k, l\} = \{0, 1, 2, 3\}$ . See e.g. Lemma 4.4 in [50] for this matrix identity.

The numerical cost of Selling's algorithm, in our applications which involve condition numbers  $\lesssim 10^2$ , is low enough to be neglected. For applications involving (much) higher condition numbers, it is natural to perform a preliminary basis reduction [53], with cost  $\mathcal{O}(\ln \mu(D))$ , after what Selling's algorithm terminates in a single step, see Proposition 1 in [32].

## C Reconstruction of an upwind gradient, for geodesic extraction

Solving eikonal equations is, in many applications, only a means to extract minimal geodesics for the corresponding optimal control problem. These paths are the integral lines of the *intrinsic* gradient (defined w.r.t. the metric) of the eikonal PDE solution:

$$\dot{\gamma}(t) := V(\gamma(t)), \quad \text{where } V(\mathbf{p}) := d\mathcal{F}_{\mathbf{p}}^*(du(\mathbf{p})). \quad (51)$$

See the discussion §1, or Appendix C of [28]. We describe in this appendix a robust estimator of the vector field  $V$ , defined directly from the PDE scheme, and in terms of upwind quantities.

In the case of Riemannian metrics one has  $V(\mathbf{p}) = \mathcal{D}(\mathbf{p})du(\mathbf{p})$ , as follows from the dual metric expression  $\mathcal{F}_{\mathbf{p}}^*(\hat{\mathbf{p}}) = \|\hat{\mathbf{p}}\|_{\mathcal{D}(\mathbf{p})}$  and the eikonal equation. The following proposition, generalizing a similar result in the isotropic case in [65], extracts this direction from the FM-VR1 scheme.

**Proposition C.1.** *For each  $\mathbf{p} \in \Omega_h$  one has the first-order approximation*

$$\mathcal{D}(\mathbf{p})du(\mathbf{p}) = h^{-1} \sum_{1 \leq i \leq d'} \varepsilon_i \rho_i(\mathbf{p}) \max\{0, u(\mathbf{p}) - u(\mathbf{p} + h\dot{\mathbf{e}}_i), u(\mathbf{p}) - u(\mathbf{p} - h\dot{\mathbf{e}}_i)\} \dot{\mathbf{e}}_i + \mathcal{O}(h), \quad (52)$$

where for each  $1 \leq i \leq d'$  one defines  $\varepsilon_i = 0$  (resp.  $\varepsilon_i = -1$ , resp.  $\varepsilon_i = 1$ ) if the  $i$ -th maximum is 0 (resp.  $u(\mathbf{p}) - u(\mathbf{p} + h\dot{\mathbf{e}}_i)$ , resp.  $u(\mathbf{p}) - u(\mathbf{p} - h\dot{\mathbf{e}}_i)$ ).

*Proof.* We may assume that  $u$  is linear,  $u(\mathbf{p}) = \langle \hat{\eta}, \mathbf{p} - \mathbf{p}_0 \rangle$ , in which case (52) becomes by (14)

$$h^{-1} \sum_{1 \leq i \leq d'} \varepsilon_i \rho_i(\mathbf{p}) |\langle \hat{\eta}, h\dot{\mathbf{e}}_i \rangle| \dot{\mathbf{e}}_i = \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \langle \hat{\eta}, \dot{\mathbf{e}}_i \rangle \dot{\mathbf{e}}_i = \mathcal{D}(\mathbf{p})\hat{\eta} = \mathcal{D}(\mathbf{p})du(\mathbf{p}). \quad \square$$

Rander metrics take the general form  $\mathcal{F}(\dot{\mathbf{p}}) = \|\dot{\mathbf{p}}\|_{\mathcal{M}} + \langle \hat{\eta}, \dot{\mathbf{p}} \rangle$ , omitting for readability the dependency of  $\mathcal{F}$ ,  $\mathcal{M}$  and  $\hat{\eta}$  on the base point  $\mathbf{p}$ . Interestingly, the algebraic structure of the dual metric is similar:  $\mathcal{F}^*(\hat{\mathbf{p}}) = \|\hat{\mathbf{p}}\|_{\mathcal{M}^*} + \langle \hat{\mathbf{p}}, \hat{\eta}^* \rangle$ , where  $(\mathcal{M}^*, \hat{\eta}^*)$  is related to  $(\mathcal{M}, \hat{\eta})$  by

$$\begin{pmatrix} \mathcal{M} & \hat{\eta}^T \\ \hat{\eta} & 1 \end{pmatrix} \begin{pmatrix} \mathcal{M}^* & \hat{\eta}^* \\ (\hat{\eta}^*)^T & 1 \end{pmatrix} = \alpha^2 \text{Id}, \quad (53)$$

for some  $\alpha = \alpha(\mathcal{M}, \hat{\eta})$ . See also [48] for equivalent explicit formulas. As a result, the intrinsic gradient (51, right) reads

$$V(\mathbf{p}) = \frac{\mathcal{M}^*(\mathbf{p})du(\mathbf{p})}{\|du(\mathbf{p})\|_{\mathcal{M}^*(\mathbf{p})}} + \hat{\eta}^*(\mathbf{p}).$$

Finally, an estimate of the differential  $du(\mathbf{p})$  is deduced as in the Riemannian case from the numerical scheme, inferred from the analogous first-order approximation:  $\mathcal{D}(\mathbf{p})(du(\mathbf{p}) - \hat{\eta}(\mathbf{p})) =$

$$h^{-1} \sum_{1 \leq i \leq d'} \varepsilon_i \rho_i(\mathbf{p}) \max\{0, u(\mathbf{p}) - u(\mathbf{p} + h\dot{\mathbf{e}}_i) + h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle, u(\mathbf{p}) - u(\mathbf{p} - h\dot{\mathbf{e}}_i) - h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle\} \dot{\mathbf{e}}_i + \mathcal{O}(h).$$

## D Convergence in the sub-Riemannian case

This section is devoted to the proof of Theorem 1.8, namely the numerical analysis of the FM-VR1 scheme introduced in this paper in the sub-Riemannian (or pre-Riemannian) setting. The estimates of  $u - u_\varepsilon$  and  $u_\varepsilon - U_{h,\varepsilon}$ , respectively related to the model relaxation and to its discretization, are presented separately in §D.1 and §D.2. The arguments used to prove the Lipschitz regularity of the solution  $u$  to the continuous problem, and to control the growth close to  $\partial\Omega$  of the solution  $U_{h,\varepsilon}$  to the relaxed and discretized problem, differ substantially from those used in the Riemannian case.

Before turning to these proofs, we recall two basic results on the regularity of the orthogonal projection onto a set, assumed respectively to be convex or to have a smooth boundary. Let  $d_B(\mathbf{p}) := \min_{\mathbf{q} \in B} \|\mathbf{p} - \mathbf{q}\|$  denote the distance to a non-empty closed set  $B \subseteq \mathbb{E}$ , and let  $\mathbf{P}_B(\mathbf{p})$  denote the minimizer  $\mathbf{q} \in B$  for  $d_B(\mathbf{p})$ , when it is unique. The Hausdorff distance between two closed subsets of  $\mathbb{E}$  is denoted  $\mathcal{H}(\cdot, \cdot)$ .

**Proposition D.1.** *Let  $\mathbf{p}, \mathbf{p}' \in \mathbb{E}$ , and let  $B, B' \subseteq \mathbb{E}$  be non-empty closed and convex. Then*

$$\|\mathbf{P}_B(\mathbf{p}) - \mathbf{P}_B(\mathbf{p}')\| \leq \|\mathbf{p} - \mathbf{p}'\|, \quad \|\mathbf{P}_B(\mathbf{p}) - \mathbf{P}_{B'}(\mathbf{p})\| \leq \sqrt{\mathcal{H}(B, B')} \sqrt{d_B(\mathbf{p}) + d_{B'}(\mathbf{p})}. \quad (54)$$

*Proof.* The uniqueness and Lipschitz regularity of the projection onto a closed convex set (54, left) are extremely classical, hence their proof is omitted. *Proof of (54, right).* Let  $\mathbf{q} := \mathbf{P}_B(\mathbf{p})$  and  $\mathbf{q}' := \mathbf{P}_{B'}(\mathbf{p})$ . We first assume that  $\mathbf{p} \neq \mathbf{q}$ , and regard  $\mathbf{q} - \mathbf{p}$  as a co-vector by Riez duality. Observe that  $B$  is contained in the half space  $H := \{\mathbf{r} \in \mathbb{E}; \langle \mathbf{q} - \mathbf{p}, \mathbf{r} - \mathbf{q} \rangle \geq 0\}$ , hence

$$\mathcal{H}(B, B') \geq d_B(\mathbf{q}') \geq d_H(\mathbf{q}') = \max \left\{ 0, \left\langle \frac{\mathbf{q} - \mathbf{p}}{\|\mathbf{q} - \mathbf{p}\|}, \mathbf{q} - \mathbf{q}' \right\rangle \right\} \geq \langle \mathbf{q} - \mathbf{p}, \mathbf{q} - \mathbf{q}' \rangle / d_B(\mathbf{p}).$$

Thus  $d_B(\mathbf{p})\mathcal{H}(B, B') \geq \langle \mathbf{q} - \mathbf{p}, \mathbf{q} - \mathbf{q}' \rangle$ , and this inequality also holds without the assumption  $\mathbf{q} \neq \mathbf{p}$ . Summing this identity with the similar one obtained exchanging the roles of  $(B, B')$  and  $(\mathbf{q}, \mathbf{q}')$  we obtain  $(d_B(\mathbf{p}) + d_{B'}(\mathbf{p}))\mathcal{H}(B, B') \geq \langle (\mathbf{q} - \mathbf{p}) + (\mathbf{p} - \mathbf{q}'), \mathbf{q} - \mathbf{q}' \rangle = \|\mathbf{q} - \mathbf{q}'\|^2$  which is the announced result.  $\square$

Here and below, slightly abusively, we regard normal vectors to  $\partial\Omega$  as co-vectors.

**Proposition D.2.** *Assume that the domain boundary  $\partial\Omega$  admits outward normals  $\mathbf{n}(\mathbf{q})$ ,  $\mathbf{q} \in \partial\Omega$ , which have  $1/R_\Omega$ -Lipschitz regularity w.r.t.  $\mathbf{q}$ . Then  $\mathbf{P}_{\partial\Omega}(\mathbf{p})$  is uniquely defined for all  $\mathbf{p} \in \mathbb{E}$  such that  $d_{\partial\Omega}(\mathbf{p}) < R_\Omega$ . Furthermore  $d_{\partial\Omega}(\mathbf{p} + \dot{\mathbf{e}})$ , for  $\mathbf{p}, \dot{\mathbf{e}} \in \mathbb{E}$ , is either zero or obeys*

$$d_{\partial\Omega}(\mathbf{p} + \dot{\mathbf{e}}) \leq d_{\partial\Omega}(\mathbf{p}) + \langle \mathbf{n}(\mathbf{p}), \dot{\mathbf{e}} \rangle + \frac{\|\dot{\mathbf{e}}\|^2}{2R_\Omega}, \quad \text{where } \mathbf{n}(\mathbf{p}) := \mathbf{n}(\mathbf{P}_B(\mathbf{p})).$$

*Proof.* The Lipschitz assumption on the normals implies, for any  $\mathbf{q} \in \partial\Omega$ , the inclusions  $B(\mathbf{q} - R_\Omega \mathbf{n}(\mathbf{q}), R_\Omega) \subseteq \bar{\Omega}$  and  $B(\mathbf{q} + R_\Omega \mathbf{n}(\mathbf{q}), R_\Omega) \subseteq \mathbb{E} \setminus \Omega$ . Fix  $\mathbf{p} \in \mathbb{E}$ , and let  $\mathbf{q} \in \partial\Omega$  be an arbitrary closest point to  $\mathbf{p}$ . The first inclusion implies the announced uniqueness when  $d_{\partial\Omega}(\mathbf{p}) < R_\Omega$ , and the second inclusion, together with Lemma 2.10 applied to  $\|(\mathbf{q} + R_\Omega \mathbf{n}(\mathbf{q}) - \mathbf{p}) - \dot{\mathbf{e}}\|$ , implies the distance estimate.  $\square$

### D.1 Estimating $u_\varepsilon - u$

In this subsection we bound, in the uniform norm, the difference between the value function  $u$  of the pre-Riemannian problem, and the one  $u_\varepsilon$  associated to the Riemannian approximation (18),

for any  $0 < \varepsilon \leq 1$ . Assumption 1.6, on global controllability and short time local controllability at the boundary, is central in the proof. Related arguments can be found in [3], but the proof is provided for completeness and because in the process we establish estimates used in §D.2. We use the notations and assumptions of Theorem 1.8. Let us introduce the control sets of the Riemannian relaxation

$$\mathcal{B}_\varepsilon(\mathbf{p}) := \{\dot{\mathbf{p}} \in \mathbb{E}; \|\dot{\mathbf{p}}\|_{\mathcal{M}_\varepsilon(\mathbf{p})} \leq 1\}$$

for each  $0 < \varepsilon \leq 1$ ,  $\mathbf{p} \in \overline{\Omega}$ , with the convention that  $\mathcal{B}_0$  denotes the pre-Riemannian control sets of Definition 1.5. Note the inclusion  $\mathcal{B}_\varepsilon(\mathbf{p}) \subseteq \mathcal{B}_{\varepsilon'}(\mathbf{p})$  for any  $0 \leq \varepsilon \leq \varepsilon' \leq 1$ , which implies the pointwise inequalities  $T_{\mathcal{B}_\varepsilon}(\mathbf{p}, \mathbf{q}) \geq T_{\mathcal{B}_{\varepsilon'}}(\mathbf{p}, \mathbf{q})$  for the control times, and  $u_\varepsilon(\mathbf{p}) \geq u_{\varepsilon'}(\mathbf{p})$  for the exit times, for any  $\mathbf{p}, \mathbf{q} \in \overline{\Omega}$ .

Our first lemma establishes the Lipschitz regularity of the control sets  $\mathcal{B}_\varepsilon(\mathbf{p})$  and of the tensors  $\mathcal{D}_\varepsilon(\mathbf{p})$  with respect to the position  $\mathbf{p} \in \overline{\Omega}$  and the relaxation parameter  $\varepsilon \in [0, 1]$ . The control sets regularity allows to apply Gronwall's Lemma in the proof of Proposition D.4, whereas the tensors regularity is used in §D.2 for establishing Assumption (i) of the doubling of variables argument Theorem 2.4. Denote by  $A_\varepsilon(\mathbf{p})$  the matrix of columns  $\dot{\omega}_1, \dots, \dot{\omega}_n, \varepsilon \dot{\omega}_1^*, \dots, \varepsilon \dot{\omega}_{n^*}^*$ , for all  $0 \leq \varepsilon \leq 1$ ,  $\mathbf{p} \in \overline{\Omega}$ . Since these vector fields are Lipschitz and bounded, the matrix field  $A_\varepsilon$  obeys for some constant  $C_{\text{Lip}}^{\mathcal{D}}$  the following Lipschitz regularity property: for all  $\mathbf{p}, \mathbf{q} \in \overline{\Omega}$  and all  $\varepsilon, \varepsilon' \in [0, 1]$

$$\max\{\|A_\varepsilon(\mathbf{p}) - A_{\varepsilon'}(\mathbf{q})\|, \|A_\varepsilon^T(\mathbf{p}) - A_{\varepsilon'}^T(\mathbf{q})\|\} \leq C_{\text{Lip}}^{\mathcal{D}}(\|\mathbf{p} - \mathbf{q}\| + |\varepsilon - \varepsilon'|). \quad (55)$$

Recall that the operator norm of an  $m \times n$  matrix  $A$  is defined by  $\|A\| := \sup\{\|A\dot{\mathbf{r}}\|; \dot{\mathbf{r}} \in \mathbb{R}^n, \|\dot{\mathbf{r}}\| \leq 1\}$ . One easily checks that  $\|A\| = \|A^T\|$  for any matrix  $A$ , hence the l.h.s. of (55) could be slightly simplified, but we prefer to emphasize the fact that both the regularity of the matrix field  $A_\varepsilon$  and of its transpose are used.

**Lemma D.3.** *One has the Lipschitz regularity: for all  $\mathbf{p}, \mathbf{q} \in \overline{\Omega}$ ,  $\varepsilon, \varepsilon' \in [0, 1]$ , and all  $\hat{\mathbf{r}} \in \mathbb{E}^*$*

$$\mathcal{H}(\mathcal{B}_\varepsilon(\mathbf{p}), \mathcal{B}_{\varepsilon'}(\mathbf{q})) \leq C_{\text{Lip}}^{\mathcal{D}}(\|\mathbf{q} - \mathbf{p}\| + |\varepsilon - \varepsilon'|) \quad (56)$$

$$|\|\hat{\mathbf{r}}\|_{\mathcal{D}_\varepsilon(\mathbf{p})} - \|\hat{\mathbf{r}}\|_{\mathcal{D}_{\varepsilon'}(\mathbf{q})}| \leq C_{\text{Lip}}^{\mathcal{D}}(\|\mathbf{q} - \mathbf{p}\| + |\varepsilon - \varepsilon'|)\|\hat{\mathbf{r}}\|. \quad (57)$$

*Proof.* For any maps  $\varphi, \psi$  from an arbitrary space  $X$  to  $\mathbb{E}$ , one has  $\mathcal{H}(\varphi(X), \psi(X)) \leq \sup_{x \in X} \|\varphi(x) - \psi(x)\|$ . Observing that  $\mathcal{B}_\varepsilon(\mathbf{p}) = (A_\varepsilon(\mathbf{p}))(B)$ , where  $B$  is the unit ball of  $\mathbb{R}^{n+n^*}$ , we obtain

$$\mathcal{H}(\mathcal{B}_\varepsilon(\mathbf{p}), \mathcal{B}_{\varepsilon'}(\mathbf{q})) \leq \sup_{x \in B} \|A_\varepsilon(\mathbf{p})x - A_{\varepsilon'}(\mathbf{q})x\| = \|A_\varepsilon(\mathbf{p}) - A_{\varepsilon'}(\mathbf{q})\| \leq C_{\text{Lip}}^{\mathcal{D}}(\|\mathbf{q} - \mathbf{p}\| + |\varepsilon - \varepsilon'|).$$

which is (56). Observing that  $\mathcal{D}_\varepsilon(\mathbf{p}) = A_\varepsilon(\mathbf{p})A_\varepsilon(\mathbf{p})^T$ , see Definition 1.5, we obtain (57), since

$$|\|\hat{\mathbf{r}}\|_{\mathcal{D}_\varepsilon(\mathbf{p})} - \|\hat{\mathbf{r}}\|_{\mathcal{D}_{\varepsilon'}(\mathbf{q})}| = |\|A_\varepsilon^T(\mathbf{p})\hat{\mathbf{r}}\| - \|A_{\varepsilon'}^T(\mathbf{q})\hat{\mathbf{r}}\|| \leq \|A_\varepsilon^T(\mathbf{p}) - A_{\varepsilon'}^T(\mathbf{q})\|\|\hat{\mathbf{r}}\|. \quad \square$$

**Proposition D.4.** *Let  $0 \leq \varepsilon_0 \leq 1$ , let  $\gamma_0 : [0, T_0] \rightarrow \overline{\Omega}$  be a  $\mathcal{B}_{\varepsilon_0}$ -admissible path, see Definition 1.2, and let  $\mathbf{p}_0 := \gamma_0(0)$ . Let  $0 \leq \varepsilon_1 \leq 1$ , and let  $\gamma_1 : [0, T_1] \rightarrow \overline{\Omega}$  be a solution to the ODE*

$$\dot{\gamma}_1(t) := P_{\mathcal{B}_{\varepsilon_1}(\gamma_1(t))}(\dot{\gamma}_0(t)),$$

*with initial condition  $\gamma_1(0) = \mathbf{p}_1 \in \Omega$ , where the final time  $T_1$  is either  $T_0$  or the time where  $\gamma_1$  reaches  $\partial\Omega$ . Then  $\gamma_1$  is  $\mathcal{B}_1$ -admissible, and for any  $0 \leq t \leq T_1$  one has*

$$|\gamma_0(t) - \gamma_1(t)| \leq (\|\mathbf{p}_0 - \mathbf{p}_1\| + |\varepsilon_0 - \varepsilon_1|) \exp(C_{\text{Lip}}^{\mathcal{B}} t) - |\varepsilon_0 - \varepsilon_1|.$$

*Proof.* The orthogonal projection  $P_B(\mathbf{p})$ , of a given  $\mathbf{p} \in \mathbb{E}$ , depends continuously (in fact with  $(1/2)$ -Holder regularity) on the closed and convex set  $B$ , see Proposition D.1. The right-hand side of (D.4) therefore depends continuously on  $\gamma_1(t)$ , hence this ODE admits solutions by Peano's existence theorem. Note that the Picard-Lindelof/Cauchy-Lipschitz uniqueness theorem does *not* apply since it requires Lipschitz regularity of the r.h.s., but the lack of uniqueness is fortunately not an issue in this proof.

The  $\mathcal{B}_{\varepsilon_1}$ -admissibility of  $\gamma_1$  holds by construction, and since  $\dot{\gamma}_0(t) \in \mathcal{B}_{\varepsilon_0}(\gamma_0(t))$  one has

$$|\dot{\gamma}_0(t) - \dot{\gamma}_1(t)| \leq \mathcal{H}(\mathcal{B}_{\varepsilon_0}(\gamma_0(t)), \mathcal{B}_{\varepsilon_1}(\gamma_1(t))) \leq C_{\text{Lip}}^{\mathcal{B}}(|\varepsilon_0 - \varepsilon_1| + |\gamma_0(t) - \gamma_1(t)|),$$

for any  $0 \leq t \leq T_1$ . The announced estimate then follows from Gronwall's lemma.  $\square$

The following lemma makes use of the transversality property in Assumption 1.6 to upper bound the exit time function  $u_\varepsilon$  close to the domain boundary  $\partial\Omega$ . This property is equivalent to:  $\|\mathbf{n}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})}^2 = \sum_{i=1}^n \langle \mathbf{n}(\mathbf{p}), \dot{\omega}_i(\mathbf{p}) \rangle^2 \neq 0$  for all  $\mathbf{p} \in \partial\Omega$ , where  $\mathbf{n}(\mathbf{p})$  denotes the outward normal to  $\partial\Omega$ . Hence denoting  $\mathbf{n}(\mathbf{p}) := \mathbf{n}(P_{\partial\Omega}(\mathbf{p}))$  for  $\mathbf{p}$  close enough to  $\partial\Omega$ , as in Proposition D.2, there exists by continuity positive constants  $c_{\mathcal{D}}, c_\Omega$  such that

$$\|\mathbf{n}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} \geq c_{\mathcal{D}}, \text{ for all } \mathbf{p} \in \bar{\Omega} \text{ such that } d_{\partial\Omega}(\mathbf{p}) \leq c_\Omega. \quad (58)$$

In the next lemma and proposition we construct paths from an arbitrary point  $\mathbf{p} \in \bar{\Omega}$  to  $\partial\Omega$ , whereas the original problem (8) is to find a path from  $\partial\Omega$  to  $\mathbf{p}$ . This change of orientation is only used for notational simplicity, and is valid since the paths can be reverse parametrized, and since the control sets are symmetric:  $\mathcal{B}_\varepsilon(\mathbf{p}) = \{-\dot{\mathbf{p}}; \dot{\mathbf{p}} \in \mathcal{B}_\varepsilon(\mathbf{p})\}$ , for all  $\mathbf{p} \in \bar{\Omega}$ ,  $\varepsilon \in [0, 1]$ .

**Lemma D.5.** *For all  $\mathbf{p} \in \bar{\Omega}$  such that  $d_{\partial\Omega}(\mathbf{p}) \leq c_\Omega$ , one has  $u(\mathbf{p}) \leq d_{\partial\Omega}(\mathbf{p})/c_{\mathcal{D}}$ .*

*Proof.* Let  $D\Omega := \{\mathbf{q} \in \bar{\Omega}; d_{\partial\Omega}(\mathbf{q}) \leq c_\Omega\}$ . Define a vector field  $\dot{\mathbf{v}} : D\Omega \rightarrow \mathbb{E}$  by

$$\dot{\mathbf{v}}(\mathbf{q}) := \frac{\mathcal{D}(\mathbf{q})\mathbf{n}(\mathbf{q})}{\|\mathbf{n}(\mathbf{q})\|_{\mathcal{D}(\mathbf{q})}} = A_0(\mathbf{q}) \frac{A_0^T(\mathbf{q})\mathbf{n}(\mathbf{q})}{\|A_0^T(\mathbf{q})\mathbf{n}(\mathbf{q})\|},$$

for all  $\mathbf{q} \in D\Omega$ , and note that  $\dot{\mathbf{v}}(\mathbf{q}) \in \mathcal{B}(\mathbf{q})$ . Consider the solution to the ODE  $\dot{\gamma}(t) := \dot{\mathbf{v}}(\gamma(t))$ , with initial condition  $\gamma(0) = \mathbf{p} \in D\Omega$ , stopping at the time  $T \in [0, \infty]$  when  $\gamma$  leaves  $D\Omega$ . By construction,  $\gamma$  is a  $\mathcal{B}_\varepsilon$ -admissible path, and for all  $0 \leq t \leq T$  one has by Proposition D.2

$$\frac{d}{dt} d_{\partial\Omega}(\gamma(t)) = -\langle \mathbf{n}(\gamma(t)), \dot{\mathbf{v}}(\gamma(t)) \rangle = -\|\mathbf{n}(\gamma(t))\|_{\mathcal{D}(\gamma(t))} \leq -c_{\mathcal{D}}.$$

Therefore  $T \leq d_{\partial\Omega}/c_{\mathcal{D}}$  and  $\gamma(T) \in \partial\Omega$ , hence  $u(\mathbf{p}) \leq T$  and the announced result follows.  $\square$

The following proposition establishes the Lipschitz regularity of  $u_\varepsilon(\mathbf{p})$  with respect to both  $\varepsilon \in [0, 1]$  and  $\mathbf{p} \in \bar{\Omega}$ . Regularity w.r.t.  $\varepsilon$  proves (19, left) in Theorem 1.8, which was the aim of this section. Regularity w.r.t.  $\mathbf{p}$  is used in the next subsection to apply the doubling of variables techniques.

**Proposition D.6.** *One has the Lipschitz regularity property*

$$|u_{\varepsilon_0}(\mathbf{p}_0) - u_{\varepsilon_1}(\mathbf{p}_1)| \leq C_{\text{Lip}}^u(\|\mathbf{p}_0 - \mathbf{p}_1\| + |\varepsilon_0 - \varepsilon_1|). \quad (59)$$

for all  $\mathbf{p}_0, \mathbf{p}_1 \in \bar{\Omega}$  and all  $\varepsilon_0, \varepsilon_1 \in [0, 1]$ , where  $C_{\text{Lip}}^u := \exp(C_{\text{Lip}}^{\mathcal{B}}\|u\|_\infty)/c_{\mathcal{D}}$ .

*Proof.* Assume w.l.o.g. that  $u_{\varepsilon_0}(\mathbf{p}_0) \leq u_{\varepsilon_1}(\mathbf{p}_1)$ . Let  $\gamma_0 : [0, T_0] \rightarrow \bar{\Omega}$  be an optimal  $\mathcal{B}_{\varepsilon_0}$ -admissible path from  $\mathbf{p}_0$  to  $\partial\Omega$ , where  $T_0 := u_{\varepsilon_0}(\mathbf{p}_0)$ . Let  $\gamma_1 : [0, T_1] \rightarrow \bar{\Omega}$  be as in Proposition D.4. If  $\gamma_1(T_1) \in \partial\Omega$ , then  $u_1(\mathbf{p}_1) \leq T_1 \leq T_0$  and the result follows.

Otherwise by Proposition D.4 we get  $d_{\partial\Omega}(\gamma_1(T_0)) \leq \|\gamma_1(T_0) - \gamma_0(T_0)\| \leq C(|\varepsilon_0 - \varepsilon_1| + \|\mathbf{p}_0 - \mathbf{p}_1\|)$  with  $C := \exp(C_{\text{Lip}}T_0)$ . Therefore

$$u_{\varepsilon_1}(\mathbf{p}_1) \leq T_0 + d_{\partial\Omega}(\gamma_1(T_0))/c_{\mathcal{D}} \leq u_{\varepsilon_0}(\mathbf{p}_0) + C_{\text{Lip}}^u(\|\mathbf{p}_0 - \mathbf{p}_1\| + |\varepsilon_0 - \varepsilon_1|)$$

as announced, using Lemma D.5 and recalling that  $u_{\varepsilon_1} \leq u_0 = u$  on  $\bar{\Omega}$ , and assuming that  $d_{\partial\Omega}(\gamma_1(T_0)) \leq c_{\Omega}$ . Thus (59) holds for all  $(\mathbf{p}_0, \varepsilon_0), (\mathbf{p}_1, \varepsilon_1) \in \bar{\Omega} \times [0, 1]$  such that  $\|\mathbf{p}_0 - \mathbf{p}_1\| + |\varepsilon_0 - \varepsilon_1| \leq c_{\Omega}/C$ . By defining  $u_{\varepsilon}(\mathbf{p}) = 0$  for all  $\mathbf{p} \in \mathbb{E} \setminus \Omega$ ,  $\varepsilon \in [0, 1]$ , the result extends to all  $(\mathbf{p}_0, \varepsilon_0), (\mathbf{p}_1, \varepsilon_1) \in \mathbb{E} \times [0, 1]$  subject to the same closeness constraint, which finally can be removed since Lipschitz regularity on a convex set is a local property.  $\square$

## D.2 Estimating $U_{h,\varepsilon} - u_{\varepsilon}$

In this subsection we complete the proof of Theorem 1.8 by estimating the difference  $|U_{h,\varepsilon} - u_{\varepsilon}|$  on  $\Omega_h$ , for any  $h > 0$ ,  $\varepsilon \in ]0, 1]$ , using the doubling of variables technique Theorem 2.4. We use the notations and assumptions of Theorems 1.8 and 2.4, except of course assumptions (i) and (ii) of the latter which we intend to prove.

The first assumption of Theorem 2.4 is the Lipschitz regularity of the value function  $u_{\varepsilon}$ , which is established in Proposition D.6 above, with a constant  $C_{\text{Lip}}^u$  independent of  $\varepsilon$ . Note that, in contrast, naively adapting the Riemannian argument of Proposition 2.6 yields the Lipschitz constant  $\lambda(\mathcal{M}_{\varepsilon}) \approx \varepsilon^{-1}$  exploding as  $\varepsilon \rightarrow 0$ , and thus unsuitable for proving Theorem 1.8.

The next step is to establish assumption (i) of Theorem 2.4. Lemma 2.7 from the Riemannian case applies without modification to  $u_{\varepsilon}$  and  $U_{h,\varepsilon}$  since it does not involve any quantitative properties of the Riemannian metric field  $\mathcal{M}_{\varepsilon}$ . Lemma 2.8 from the Riemannian case also applies, with constants independent of  $\varepsilon \in ]0, 1]$ . Indeed the dual tensors have the expression  $\mathcal{D}_{\varepsilon} = \mathcal{D} + \varepsilon^2 \mathcal{D}^*$ , see (18), and therefore their max norm (28) is bounded:  $\lambda^*(\mathcal{D}_{\varepsilon}) \leq \lambda^*(\mathcal{D}_1)$  independently of  $\varepsilon \in [0, 1]$ . The last ingredient used to prove assumption (i) in the Riemannian case is the Lipschitz regularity of the dual norms (31), which is established in (57) for the pre-Riemannian model with a constant  $C_{\text{Lip}}^{\mathcal{D}}$  independent of  $\varepsilon \in [0, 1]$ .

The following proposition establishes assumption (ii) of Theorem 2.4, a discrete counterpart of short time local controllability at the boundary, by adapting the arguments developed in the Riemannian case, see Proposition 2.11. Note the use of Assumption 1.6, which is required due to the lack of uniform definiteness of the the tensors  $\mathcal{D}_{\varepsilon}$ ,  $0 < \varepsilon \leq 1$ . The weights and offsets used in the decomposition (14) of  $\mathcal{D}_{\varepsilon}(\mathbf{p})$  are denoted  $(\rho_{i,\varepsilon}(\mathbf{p}), \dot{\mathbf{e}}_{i,\varepsilon}(\mathbf{p}))_{i=1}^{d'}$ ,  $\mathbf{p} \in \bar{\Omega}$ ,  $0 < \varepsilon \leq 1$ .

**Proposition D.7.** *Let  $\mathbf{p} \in \Omega_h$  be such that  $d(\mathbf{p}, \partial\Omega) \leq c_{\Omega}$ , and let  $\varepsilon \in ]0, 1]$ . Then there exists  $1 \leq i \leq d'$  and a sign  $s \in \{-1, 1\}$  such that either  $\mathbf{p} + hs\dot{\mathbf{e}}_{i,\varepsilon} \notin \Omega_h$  or*

$$U_{h,\varepsilon}(\mathbf{p}) \leq U_{h,\varepsilon}(\mathbf{p} + hs\dot{\mathbf{e}}_{i,\varepsilon}) + hC_1\|\dot{\mathbf{e}}_{i,\varepsilon}\|, \quad d(\mathbf{p} + hs\dot{\mathbf{e}}_{i,\varepsilon}, \partial\Omega) \leq d(\mathbf{p}, \partial\Omega) - hc_2\|\dot{\mathbf{e}}_{i,\varepsilon}\|,$$

with  $C_1 := \sqrt{d'}/c_{\mathcal{D}}$  and  $c_2 := c_{\mathcal{D}}/(2\lambda^*(\mathcal{D}_1)\sqrt{d'})$ . This implies assumption (ii) of Theorem 2.4, with the constants  $C_{\text{bd}} = C_1/c_2$ ,  $C'_{\text{bd}} := C_{\text{bd}}r_{\varepsilon}$ , and  $c_{\text{bd}} = c_{\mathcal{D}}$ .

*Proof.* Using Assumption 1.6 and (58), one obtains a counterpart for (35)

$$c_{\mathcal{D}}^2 \leq \|\mathbf{n}(\mathbf{p})\|_{\mathcal{D}_{\varepsilon}(\mathbf{p})}^2 = \sum_{1 \leq i \leq d'} \rho_{i,\varepsilon}(\mathbf{p}) \langle \mathbf{n}(\mathbf{p}), \dot{\mathbf{e}}_{i,\varepsilon} \rangle^2,$$

where  $\mathbf{n}(\mathbf{p}) := \mathbf{n}(P_{\partial\Omega}(\mathbf{p}))$ . The proof is then similar to the one of Proposition 2.11, up to the replacement of Lemma 2.10 with Proposition D.2.  $\square$

## E Convergence in the Rander case

This section is devoted to the proof of Theorem 1.10, which establishes the convergence of the solutions to the Rander metric variant of our PDE discretization scheme. We use the notations and assumptions of this theorem. For that purpose we consider the scheme  $\mathfrak{F}_h$  on the discrete domain  $\Omega_h$  defined for any  $U : \Omega_h \rightarrow \mathbb{R}$  and any  $\mathbf{p} \in \Omega_h$  by the expression  $(\mathfrak{F}_h U(\mathbf{p}))^2 :=$

$$h^{-2} \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \max\{0, U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}}_i) + h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle, U(\mathbf{p}) - U(\mathbf{p} - h\dot{\mathbf{e}}_i) - h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle\}^2. \quad (60)$$

By convention,  $U$  is extended by 0 outside  $\Omega_h$ . Note that this scheme is degenerate elliptic, see Definition 2.1, but non-causal as soon as  $\hat{\eta}(\mathbf{p}) \neq 0$ , in contrast with (27).

**Proposition E.1.** *The scheme (60) is degenerate elliptic. In addition:*

- (i) *The null map  $U = 0$  satisfies  $\mathfrak{F}_h U(\mathbf{p}) = \|\hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})}^2 < 1$  for all  $\mathbf{p} \in \Omega_h$ , hence it is a sub-solution for  $\mathfrak{F}_h - 1$ .*
- (ii) *Let  $R > 0$  be such that  $\Omega$  is contained in the ball of radius  $R - hr_*$ , and let  $U(\mathbf{p}) := R - \|\mathbf{p}\|$ , for all  $\mathbf{p} \in \Omega_h$ . Then for all  $\lambda \geq 0$ , and all  $\mathbf{p} \in \Omega_h$ ,*

$$\mathfrak{F}_h(\lambda U)(\mathbf{p}) \geq \|\lambda \mathbf{p} / \|\mathbf{p}\| - \hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})}^2,$$

*where  $\mathbf{p} / \|\mathbf{p}\|$  can be replaced with an arbitrary unit vector in the case  $\mathbf{p} = 0$ . Thus  $\lambda U$  is a super-solution for all sufficiently large  $\lambda$ .*

- (iii) *Let  $U$  be a super-solution for  $\mathfrak{F}_h - 1$ , and let  $\mathbf{p}_1, \dots, \mathbf{p}_N$  be the points of  $\Omega_h$  ordered in such way that  $U(\mathbf{p}_1) \leq \dots \leq U(\mathbf{p}_N)$ . For each  $\varepsilon > 0$  let  $V_\varepsilon : \Omega_h \rightarrow \mathbb{R}$  be defined by  $V_\varepsilon(\mathbf{p}_i) = U(\mathbf{p}_i) + \varepsilon - \varepsilon^{1+i}$ . Then  $V_\varepsilon$  is a strict super-solution to  $\mathfrak{F}_h - 1$  for all sufficiently small  $\varepsilon$ .*

*Proof.* Point (i) follows from the identity  $\sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle^2 = \|\hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})}^2$ , and the smallness assumption (20, right) on the co-vector field  $\hat{\eta}$ . Point (ii) is proved as in Proposition 2.5. Point (iii) is in contrast non-trivial. Let  $U$  be a super-solution for  $\mathfrak{F}_h - 1$  and let  $1 \leq k \leq N$ . Denote by  $m_i(\mathbf{p})$  the  $i$ -th maximum of three terms appearing in (60), so that  $\mathfrak{F}_h U(\mathbf{p}) = \sum_{i=1}^{d'} \rho_i(\mathbf{p}) m_i(\mathbf{p})^2$  for each  $\mathbf{p} \in \Omega_h$ . Then one has the Taylor expansion

$$\mathfrak{F}_h V_\varepsilon(\mathbf{p}_k) - \mathfrak{F}_h U(\mathbf{p}_k) = 2 \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}_k) m_i(\mathbf{p}_k) (\varepsilon^{1+k_i} - \varepsilon^{1+k}) + \mathcal{O}(\varepsilon^{2+2k_*}), \quad (61)$$

where  $k_i$  is an integer depending on  $i \in \{1, \dots, d'\}$  and  $k$ , and chosen so that  $\mathbf{p}_{k_i} = \mathbf{p}_k$ , (resp.  $\mathbf{p}_{k_i} = \mathbf{p}_k + h\dot{\mathbf{e}}_i$ , resp.  $\mathbf{p}_{k_i} = \mathbf{p}_k - h\dot{\mathbf{e}}_i$ ) if the maximum defining  $m_i(\mathbf{p}_k)$  is achieved for the first (resp. second, resp. third) term. (If this point is outside  $\Omega_h$ , we let  $k_i = 0$ . In case of a tie, the point with the smallest index is chosen.) We also denoted  $k_* := \min\{k_i; 1 \leq i \leq d'\}$ .

We prove below that  $k_i < k$  for some  $1 \leq i \leq d'$ , which by (61) implies that  $\mathfrak{F}_h V_\varepsilon(\mathbf{p}_k) > \mathfrak{F}_h U(\mathbf{p}_k) \geq 1$  for all sufficiently small  $\varepsilon > 0$  as announced. Assume for contradiction that  $k_i \geq k$  for all  $1 \leq i \leq d'$ , hence that  $U(\mathbf{p}_{k_i}) \geq U(\mathbf{p}_k)$  and therefore that  $m_i(\mathbf{p}_k) \leq |\langle \hat{\eta}(\mathbf{p}_k), \dot{\mathbf{e}}_i \rangle|$  for all  $1 \leq i \leq d'$ . Then denoting  $\mathbf{p} := \mathbf{p}_k$  one obtains

$$\mathfrak{F}_h U(\mathbf{p}) = \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) m_i(\mathbf{p})^2 \leq \sum_{1 \leq i \leq d'} \rho_i(\mathbf{p}) \langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}}_i \rangle^2 = \|\hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})}^2 < 1,$$

in contradiction with our assumption that  $U$  is a super-solution for  $\mathfrak{F}_h - 1$ . The result follows.  $\square$

In the rest of this section, we establish the properties required to apply the doubling of variables argument Theorem 2.4, using its notations. The first ingredient is the Lipschitz regularity of the exit time value function  $u$ . By construction the Rander metric (20) satisfies for all  $\mathbf{p} \in \bar{\Omega}$  and all  $\dot{\mathbf{p}} \in \mathbb{E}$ .

$$\mathcal{F}_{\mathbf{p}}(\dot{\mathbf{p}}) := \|\dot{\mathbf{p}}\|_{\mathcal{M}(\mathbf{p})} + \langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{p}} \rangle \leq (\lambda^*(\mathcal{M}) + \|\hat{\eta}\|_{\infty}) \|\dot{\mathbf{p}}\|.$$

Hence  $u$  is  $C_{\text{Lip}}^u$ -Lipschitz by Proposition 2.6, as desired, with constant  $C_{\text{Lip}}^u := \lambda^*(\mathcal{M}) + \|\hat{\eta}\|_{\infty}$  where the max-norm  $\lambda^*(\mathcal{M})$  of a Riemannian metric is defined in (28).

**Establishing assumption (i) of Theorem 2.4.** We proceed similarly to the Riemannian case §2.2, starting with an extension of Lemma 2.7 to Rander metrics. The proof, left to the reader, is similar up to fact that the scheme  $\mathfrak{F}_h U$  and the PDE operator  $\|du - \hat{\eta}\|_{\mathcal{D}}$  are not homogeneous w.r.t. their respective argument  $U$  and  $u$ . The variables  $\lambda \in [1/2, 1[$ ,  $\delta > 0$ ,  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in \Omega_h \times \Omega$  are from Theorem 2.4.

**Lemma E.2.** *Let  $\bar{\mathbf{w}} := (\bar{\mathbf{p}} - \bar{\mathbf{q}})/\delta$  and let  $\bar{U}(\mathbf{p}) := \langle \bar{\mathbf{w}}, \mathbf{p} \rangle + \frac{1}{2\delta} \|\mathbf{p} - \bar{\mathbf{p}}\|^2$ , for all  $\mathbf{p} \in \mathbb{L}_h$ . Then*

$$\mathfrak{F}_h(\bar{U}/\lambda)(\bar{\mathbf{p}}) \leq 1 \quad \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{q}})\|_{\mathcal{D}(\bar{\mathbf{q}})} \geq 1.$$

*Let  $\tilde{\mathbf{w}} := (\tilde{\mathbf{p}} - \tilde{\mathbf{q}})/\delta$  and let  $\tilde{U}(\mathbf{p}) := \langle \tilde{\mathbf{w}}, \mathbf{p} \rangle - \frac{1}{2\delta} \|\mathbf{p} - \tilde{\mathbf{p}}\|^2$ , for all  $\mathbf{p} \in \mathbb{L}_h$ . Then*

$$\mathfrak{F}_h \tilde{U}(\tilde{\mathbf{p}}) \geq 1 \quad \|\tilde{\mathbf{w}}/\lambda - \hat{\eta}(\tilde{\mathbf{q}})\|_{\mathcal{D}(\tilde{\mathbf{q}})} \leq 1.$$

Let  $C_{\text{Lip}}^{\mathcal{D}}$  be the Lipschitz regularity constant of the tensor field  $\mathcal{D}$ , in the sense of (31). Let also  $C_{\text{Lip}}^{\hat{\eta}}$  and  $c_{\mathcal{D}, \hat{\eta}} > 0$  be such that for all  $\mathbf{p}, \mathbf{q} \in \bar{\Omega}$

$$\|\hat{\eta}(\mathbf{p}) - \hat{\eta}(\mathbf{q})\| \leq C_{\text{Lip}}^{\hat{\eta}} \|\mathbf{p} - \mathbf{q}\|, \quad \|\hat{\eta}(\mathbf{p})\|_{\mathcal{D}(\mathbf{p})} \leq 1 - c_{\mathcal{D}, \hat{\eta}}.$$

Similarly to the Riemannian case, we argue by contradiction to establish assumption (i). Indeed, a contraposition of the following lemma shows that, if parameter  $\lambda$  is below the given bound, then  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \notin \Omega_h \times \Omega$  and  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \notin \Omega_h \times \Omega$ , which is assumption (i) of Theorem 2.4.

**Lemma E.3.** *Assume that  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in \Omega_h \times \Omega$  and define  $\bar{\mathbf{w}}$  and  $\bar{U}$  as in Lemma 2.7. Then*

$$\left| \mathfrak{F}_h(\bar{U}/\lambda)(\bar{\mathbf{p}}) - \|\bar{\mathbf{w}}/\lambda - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})} \right| \leq C_0 r_* \frac{h}{\delta}, \quad \left| \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{q}})\|_{\mathcal{D}(\bar{\mathbf{q}})} - \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})} \right| \leq C_1 \delta, \quad (62)$$

with  $C_0 := \lambda^*(\mathcal{D})\sqrt{d}$  and  $C_1 := C_{\text{Lip}}^{\mathcal{D}}(4C_{\text{Lip}}^u)(4C_{\text{Lip}}^u + C_{\text{Lip}}^{\hat{\eta}}) + \lambda^*(\mathcal{D})C_{\text{Lip}}^{\hat{\eta}}$ . Assuming  $\delta \leq c_{\mathcal{D}, \hat{\eta}}/(2C_1)$  this implies  $\lambda \geq 1 - \frac{2}{c_{\mathcal{D}, \hat{\eta}}}(C_0 r_* \frac{h}{\delta} + C_1 \delta)$ . The same estimates and conclusion hold for  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$ .

*Proof.* We focus on the case of  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ , the second case of  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$  being similar, and we begin with the proof of (62, left). By definition of the quadratic function  $\bar{U}$ , one has for any  $1 \leq i \leq d'$

$$\begin{aligned} & \max\{0, \bar{U}(\bar{\mathbf{p}})/\lambda - \bar{U}(\bar{\mathbf{p}} + \dot{\mathbf{e}}_i)/\lambda + h\langle \hat{\eta}(\bar{\mathbf{p}}), \dot{\mathbf{e}}_i \rangle, \bar{U}(\bar{\mathbf{p}})/\lambda - \bar{U}(\bar{\mathbf{p}} - \dot{\mathbf{e}}_i)/\lambda - h\langle \hat{\eta}(\bar{\mathbf{p}}), \dot{\mathbf{e}}_i \rangle\} \\ &= h|\langle \bar{\mathbf{w}}/\lambda - \hat{\eta}(\bar{\mathbf{p}}), \dot{\mathbf{e}}_i \rangle| + \frac{h^2}{\delta} \|\dot{\mathbf{e}}_i\|^2. \end{aligned} \quad (63)$$

From this point, the arguments developed in the Riemannian case apply without modification. The second estimate (62, right) follows from

$$\left| \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{q}})\|_{\mathcal{D}(\bar{\mathbf{q}})} - \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})} \right| \leq C_{\text{Lip}}^{\mathcal{D}} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\| (\|\bar{\mathbf{w}}\| + \|\hat{\eta}(\bar{\mathbf{q}})\|) + \lambda^*(\mathcal{D}) C_{\text{Lip}}^{\hat{\eta}} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|,$$

combined with  $\|\bar{\mathbf{p}} - \bar{\mathbf{q}}\| \leq 4C_{\text{Lip}}^u \delta$ , see Theorem 2.4, hence  $\|\bar{\mathbf{w}}\| \leq 4C_{\text{Lip}}^u$ . Combining the two estimates (32) with Lemma E.2, and using the convexity of the norm, we obtain

$$1 + C_0 r_* \frac{h}{\delta} \geq \|\bar{\mathbf{w}}/\lambda - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})} \geq \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})} + \frac{1/\lambda - 1}{\|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\bar{\mathbf{p}})}} \langle \bar{\mathbf{w}}, \mathcal{D}(\mathbf{p})(\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})) \rangle$$

The scalar product in the above r.h.s. is bounded below as follows

$$\begin{aligned} 2\langle \bar{\mathbf{w}}, \mathcal{D}(\mathbf{p})(\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})) \rangle &= \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\mathbf{p})}^2 + \|\bar{\mathbf{w}}\|_{\mathcal{D}(\bar{\mathbf{p}})}^2 - \|\hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\mathbf{p})}^2 \\ &\geq \|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\mathbf{p})}^2 - (1 - c_{\mathcal{D},\eta})^2 \end{aligned}$$

Using  $\|\bar{\mathbf{w}} - \hat{\eta}(\bar{\mathbf{p}})\|_{\mathcal{D}(\mathbf{p})} \geq 1 - C_1 \delta$ , and assuming  $2C_1 \delta \leq c_{\mathcal{D},\eta}$  for the second inequality, we obtain

$$1 + C_0 r_* \frac{h}{\delta} \geq (1 - C_1 \delta) + (1/\lambda - 1) \left( (1 - C_1 \delta) - \frac{(1 - c_{\mathcal{D},\eta})^2}{1 - C_1 \delta} \right) \geq (1 - C_1 \delta) + (1 - \lambda) \frac{c_{\mathcal{D},\eta}}{2}.$$

We used the elementary inequalities  $1/\lambda - 1 \geq 1 - \lambda$  and  $(1 - c/2) - (1 - c)^2/(1 - c/2) \geq c/2$ . This implies the announced lower bound for  $\lambda$ .  $\square$

**Establishing assumption (ii) of Theorem 2.4.** The case of Rander metrics only requires a minor adaptation of the Riemannian argument, presented in Proposition 2.11.

**Proposition E.4.** *Let  $\mathbf{p} \in \Omega_h$ , and let  $\mathbf{q} \in \mathbb{E}$  be such that  $\|\mathbf{p} - \mathbf{q}\| \geq C_0 r_* h$ , where  $C_0 := \mu(\mathcal{D})\sqrt{d'}$ . Then there exists  $1 \leq i \leq d'$  and a sign  $s \in \{-1, 1\}$  such that*

$$U_h(\mathbf{p}) \leq U_h(\mathbf{p} + hs\dot{\mathbf{e}}_i) + hC_1 \|\dot{\mathbf{e}}_i\|, \quad \|\mathbf{p} + hs\dot{\mathbf{e}}_i - \mathbf{q}\| \leq \|\mathbf{p} - \mathbf{q}\| - hc_2 \|\dot{\mathbf{e}}_i\|,$$

with  $C_1 := \lambda^*(\mathcal{M})\sqrt{d'} + \|\hat{\eta}\|_\infty$ ,  $c_2 := 1/(2\mu(\mathcal{D})\sqrt{d'})$ . This implies assumption (ii) of Theorem 2.4, with the constants  $C_{\text{bd}}^U = C_1/c_2$ ,  $C_{\text{bd}}'^U := C_{\text{bd}}C_\Omega C_0 r_*$ , and  $c_{\text{bd}}^U = +\infty$ .

*Proof.* The arguments developed in the Riemannian case apply with the following adaptation of (37): by definition of the discretization scheme (60) one has  $h^{-2}\rho^2 \max\{0, U_h(\mathbf{p}) - U_h(\mathbf{p} + \dot{\mathbf{e}}) - h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}} \rangle\}^2 \leq (\mathfrak{F}_h U_h(\mathbf{p}))^2 = 1$ , hence

$$U(\mathbf{p}) - U(\mathbf{p} + h\dot{\mathbf{e}}) \leq \frac{h}{\rho} + h\langle \hat{\eta}(\mathbf{p}), \dot{\mathbf{e}} \rangle \leq h\|\dot{\mathbf{e}}\| \left( \frac{\sqrt{d'}}{\lambda_*} + \|\hat{\eta}\|_\infty \right). \quad \square$$

## F Doubling of variables

We establish in this section the doubling of variables argument, presented in Theorem 2.4 and adapted from [30]. Since the domain  $\Omega$  is by assumption bounded, its closure  $\bar{\Omega}$  is compact, and its sampling  $\Omega_h := \Omega \cap h\mathbb{L}$  is finite. Since the functions  $u$  and  $U_h$  are supported on these sets, and since  $u$  is  $C_{\text{Lip}}$ -Lipschitz hence continuous, the maxima  $\bar{M}_{\lambda,\delta}$  and  $\widetilde{M}_{\lambda,\delta}$  are well defined and attained, at some point pairs  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}), (\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in (h\mathbb{L}) \times \mathbb{E}$ . Our first step is to establish the closeness of  $\bar{\mathbf{p}}$  with  $\bar{\mathbf{q}}$ , and of  $\tilde{\mathbf{p}}$  with  $\tilde{\mathbf{q}}$ , as announced in Theorem 2.4.

**Lemma F.1.** *Under the assumptions of Theorem 2.4, one has  $\max\{\|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|, \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\|\} \leq 4C_{\text{Lip}}\delta$ .*

*Proof.* Using the optimality properties defining  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$  and  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$ , see Theorem 2.4, and comparing with the alternative point pairs  $(\bar{\mathbf{p}}, \bar{\mathbf{p}})$  and  $(\tilde{\mathbf{p}}, \tilde{\mathbf{p}})$  respectively, one obtains

$$\begin{aligned}\lambda U_h(\bar{\mathbf{p}}) - u(\bar{\mathbf{q}}) - \frac{1}{2\delta} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|^2 &= \bar{M}_{\lambda,\delta} \geq \lambda U_h(\bar{\mathbf{p}}) - u(\bar{\mathbf{p}}), \\ \lambda u(\tilde{\mathbf{q}}) - U_h(\tilde{\mathbf{p}}) - \frac{1}{2\delta} \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\|^2 &= \widetilde{M}_{\lambda,\delta} \geq \lambda u(\tilde{\mathbf{p}}) - U_h(\tilde{\mathbf{p}}).\end{aligned}$$

The next line is obtained by first rearranging the terms of these inequalities, eliminating in particular the instances of  $U_h$ , and then using the Lipschitz regularity of  $u$

$$\frac{1}{2\delta} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|^2 \leq u(\bar{\mathbf{p}}) - u(\bar{\mathbf{q}}) \leq C_{\text{Lip}} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|, \quad \frac{1}{2\delta} \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\|^2 \leq \lambda(u(\tilde{\mathbf{q}}) - u(\tilde{\mathbf{p}})) \leq \lambda C_{\text{Lip}} \|\tilde{\mathbf{q}} - \tilde{\mathbf{p}}\|.$$

The announced result follows from these estimates and the upper bound  $1/\lambda \leq 2$ .  $\square$

The rest of this section is devoted to the proof of (25), and begins with an estimate of  $u - U_h$  in terms of the suprema  $\bar{M}_{\lambda,\delta}$ ,  $\widetilde{M}_{\lambda,\delta}$  and of the max norm  $\|u\|_\infty := \sup_{\mathbb{E}} |u|$ , which is well defined by continuity of  $u$  and compactness of its support.

**Lemma F.2.** *Under the assumptions of Theorem 2.4, one has*

$$\sup_{\mathbf{p} \in h\mathbb{L}} |u(\mathbf{p}) - U_h(\mathbf{p})| \leq 2 \left( \max\{\bar{M}_{\lambda,\delta}, \widetilde{M}_{\lambda,\delta}\} + (1 - \lambda) \|u\|_\infty \right).$$

*Proof.* By the optimality properties of  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$  and  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$  one obtains for any  $\mathbf{p} \in \Omega_h$ , respectively,

$$\begin{aligned}\bar{M}_{\lambda,\delta} &\geq \lambda U_h(\mathbf{p}) - u(\mathbf{p}) \geq \lambda(U_h(\mathbf{p}) - u(\mathbf{p})) - (1 - \lambda) \|u\|_\infty, \\ \widetilde{M}_{\lambda,\delta} &\geq \lambda u(\mathbf{p}) - U_h(\mathbf{p}) \geq (u(\mathbf{p}) - U_h(\mathbf{p})) - (1 - \lambda) \|u\|_\infty.\end{aligned}$$

The announced result follows from these one-sided estimates on  $u - U_h$ , and from  $1/\lambda \leq 2$ .  $\square$

The next paragraph establishes some conditional estimates on  $\bar{M}_{\lambda,\delta}$  and  $\widetilde{M}_{\lambda,\delta}$ , depending on the location of the points  $\bar{\mathbf{p}}, \bar{\mathbf{q}}, \tilde{\mathbf{p}}, \tilde{\mathbf{q}}$ . If  $\tilde{\mathbf{p}} \in h\mathbb{L} \setminus \Omega_h$ , then  $U_h(\tilde{\mathbf{p}}) = 0$  and we obtain

$$\widetilde{M}_{\lambda,\delta} = \lambda u(\tilde{\mathbf{q}}) - \frac{1}{2\delta} \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\|^2 \leq \lambda u(\tilde{\mathbf{q}}) \leq \lambda C_{\text{Lip}} d_{\partial\Omega}(\tilde{\mathbf{q}}) \leq C_{\text{Lip}} \|\tilde{\mathbf{p}} - \tilde{\mathbf{q}}\| \leq 4C_{\text{Lip}}^2 \delta. \quad (64)$$

We used successively the negativity of the quadratic term, the Lipschitz regularity of  $u$  and the fact that it vanishes outside  $\Omega$ , the fact that  $\tilde{\mathbf{p}} \in h\mathbb{L} \setminus \Omega_h \subseteq \mathbb{E} \setminus \Omega$ , and the previously obtained estimate on  $\tilde{\mathbf{p}} - \tilde{\mathbf{q}}$ . Likewise, if  $\bar{\mathbf{p}} \in h\mathbb{L} \setminus \Omega_h$  then  $\bar{M}_{\lambda,\delta} \leq -u(\bar{\mathbf{q}}) \leq 4C_{\text{Lip}}^2 \delta$ .

Next if  $\bar{\mathbf{q}} \in \mathbb{E} \setminus \Omega$ , then  $u(\bar{\mathbf{q}}) = 0$  and we obtain similarly to (64)

$$\bar{M}_{\lambda,\delta} = \lambda U_h(\bar{\mathbf{p}}) - \frac{1}{2\delta} \|\bar{\mathbf{p}} - \bar{\mathbf{q}}\|^2 \leq U_h(\bar{\mathbf{p}}) \leq C_{\text{bd}} d_{\partial\Omega}(\bar{\mathbf{p}}) + C'_{\text{bd}} h \leq 4C_{\text{Lip}} C_{\text{bd}} \delta + C'_{\text{bd}} h.$$

We used the same arguments as in (64), except for the second inequality which is based on assumption (ii) of Theorem 2.4. Likewise, if  $\tilde{\mathbf{q}} \in \mathbb{E} \setminus \Omega$  then  $\widetilde{M}_{\lambda,\delta} \leq -U_h(\tilde{\mathbf{p}}) \leq C_{\text{bd}}(h + 4C_{\text{Lip}}\delta)$ .

Following assumption (i) of Theorem 2.4, we assume that  $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \notin \Omega_h \times \Omega$ , thus either  $\bar{\mathbf{p}} \notin \Omega_h$  or  $\bar{\mathbf{q}} \notin \Omega$ , which yields by the above arguments

$$\bar{M}_{\lambda,\delta} \leq \max\{4C_{\text{Lip}}^2 \delta, 4C_{\text{Lip}} C_{\text{bd}} \delta + C'_{\text{bd}} h\}.$$

Likewise for  $\widetilde{M}_{\lambda,\delta}$  using the assumption that  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \notin \Omega_h \times \Omega$ . The announced result (25) follows from Lemma F.2 and the these bounds on  $\bar{M}_{\lambda,\delta}$  and  $\widetilde{M}_{\lambda,\delta}$ .