A linear finite-difference scheme for approximating Randers distances on Cartesian grids

Frédéric Bonnans, Guillaume Bonnet, Jean-Marie Mirebeau

▶ **To cite this version:**

Frédéric Bonnans, Guillaume Bonnet, Jean-Marie Mirebeau. A linear finite-difference scheme for approximating Randers distances on Cartesian grids. 2021. hal-03125879v2

HAL Id: hal-03125879

https://hal.science/hal-03125879v2

Preprint submitted on 9 Jun 2021 (v2), last revised 21 Mar 2022 (v3)

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.

A linear finite-difference scheme for approximating Randers distances on Cartesian grids

J. Frédéric Bonnans[*], Guillaume Bonnet[†], Jean-Marie Mirebeau[‡]

June 9, 2021

Abstract

Randers distances are an asymmetric generalization of Riemannian distances, and arise in optimal control problems subject to a drift term, among other applications. We show that Randers eikonal equation can be approximated by a logarithmic transformation of an anisotropic second order linear equation, generalizing Varadhan's formula for Riemannian manifolds. Based on this observation, we establish the convergence of a numerical method for computing Randers distances, from point sources or from a domain's boundary, on Cartesian grids of dimension 2 and 3, which is consistent at order 2/3, and uses tools from low-dimensional algorithmic geometry for best efficiency. We also propose a numerical method for optimal transport problems whose cost is a Randers distance, exploiting the linear structure of our discretization and generalizing previous works in the Riemannian case. Numerical experiments illustrate our results.

1 Introduction

A Randers metric is the sum of a Riemannian metric and of an anti-symmetric perturbation, suitably bounded and defined by linear form. By construction, a Randers metric is in general non-symmetric, and so is the associated path-length distance, see Remark 1.3 on terminology. Such metrics can account, in a very simple manner, for the fact that moving a vehicle uphill, or advancing a boat against water currents, costs more than the opposite operation. The asymmetry embedded in Randers metrics opens up numerous applications which cannot be addressed with the simpler Riemannian metrics, ranging from general relativity [28] to image segmentation [13], through quantum vortices [1] and path curvature penalization [11], see Remark 1.1.

In this paper, we present a numerical scheme for computing Randers distances by solving a *linear second order* Partial Differential Equation (PDE). Our approach is based on a generalization of Varadhan's formula [32], which is commonly used to compute Riemannian distances [17]. Let us emphasize that Randers distances also obey a *non-linear first order* PDE [3], which can be solved directly numerically [24, 25], yet the linear structure of the PDE formulation considered in this paper has a number of computational advantages, including easier numerical implementation, faster computation in some cases, and smoothness of the numerical solution, see Remark 1.2. Some of our results, such as the identification of the optimal scaling of the relaxation parameter ε

[*]Inria-Saclay and CMAP, École Polytechnique, Palaiseau, France

[†]LMO, Université Paris-Saclay, Orsay, France, and Inria-Saclay and CMAP, École Polytechnique, Palaiseau, France

[‡]Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, Gif-sur-Yvette, France

The first author was partially supported by the FiME Lab Research Initiative (Institut Europlace de Finance).

w.r.t. the grid scale h, and the proof of convergence in the case of point sources, are new as well in the special cases of isotropic and Riemannian metrics. We present an application to numerical optimal transportation, enabled by the linear structure of the discretization, with an asymmetric cost function defined as the Randers distance between the source and target, generalizing previous works limited to Riemannian costs [19].

In order to make our statements more precise, we need to introduce some notations. Throughout this paper, $\Omega \subset \mathbb{R}^d$ denotes a smooth bounded and connected open domain, equipped with a metric $\mathcal{F} : \overline{\Omega} \times \mathbb{R}^d \to [0, \infty[$, whose explicit form is discussed below (2). The corresponding path-length and distance are defined by

$$\text{length}_{\mathcal{F}}(\gamma) := \int_0^1 \mathcal{F}_{\gamma(t)}(\gamma'(t))\,\mathrm{d}t, \qquad \text{dist}_{\mathcal{F}}(x,y) := \inf_{\gamma \in \Gamma_x^y} \text{length}_{\mathcal{F}}(\gamma). \tag{1}$$

We denoted by γ an element of the collection $\Gamma := \text{Lip}([0,1], \overline{\Omega})$ of locally Lipschitz paths within the domain closure, and by $\Gamma_x^y \subset \Gamma$ the subset of paths from $x \in \overline{\Omega}$ to $y \in \overline{\Omega}$. We assume in this paper that $\mathcal{F}$ has the structure of a Randers metric: there exists a field $M : \overline{\Omega} \to S_d^{++}$ of symmetric positive definite matrices, and a vector field $\omega : \overline{\Omega} \to \mathbb{R}^d$, both having Lipschitz regularity, and such that for all $x \in \overline{\Omega}$ and all $v \in \mathbb{R}^d$ one has

$$\mathcal{F}_x(v) := |v|_{M(x)} + \langle \omega(x), v \rangle, \qquad \text{where } |\omega(x)|_{M(x)^{-1}} < 1. \tag{2}$$

We denoted by $\langle \cdot, \cdot \rangle$ the standard Euclidean scalar product, and by $|v|_M := \sqrt{\langle v, Mv \rangle}$ the anisotropic (but symmetric) norm on $\mathbb{R}^d$ defined by a symmetric positive definite matrix M. The smallness constraint (2, right) ensures that $\mathcal{F}_x(v) > 0$ for all $v \neq 0$, $x \in \overline{\Omega}$. Randers metrics include Riemannian metrics as a special case, when the vector field ω vanishes identically over the domain. See Fig. 2 for an illustration of their unit balls, distance maps, and minimal paths.

Our approach to the computation of Randers distances goes through the solution to a linear second order PDE, depending on a small parameter $\varepsilon > 0$, and some boundary data $g \in C^0(\partial\Omega, \mathbb{R})$

$$u_\varepsilon + 2\varepsilon \langle b, \nabla u_\varepsilon \rangle - \varepsilon^2 \,\text{Tr}(A_b \nabla^2 u_\varepsilon) = 0 \text{ in } \Omega, \qquad u_\varepsilon = \exp(-g/\varepsilon) \text{ on } \partial\Omega, \tag{3}$$

where A_b is a field of symmetric positive definite matrices, and b is a vector field, depending in a simple algebraic manner on the Randers metric parameters M and ω, see Lemma 2.6 and (9). In the Riemannian special case one has $A_b = M^{-1}$ and $b = \omega = 0$, consistently with [32]. We establish in Theorem 2.12, following [4], that for all $x \in \Omega$

$$\mathbf{u}(x) := \lim_{\varepsilon \to 0} -\varepsilon \ln u_\varepsilon(x) \qquad \text{exists and satisfies} \quad \mathbf{u}(x) = \min_{p \in \partial\Omega} g(p) + d_{\mathcal{F}}(p,x). \tag{4}$$

In other words, $\mathbf{u}$ is the Randers distance from the boundary $\partial\Omega$, with initial time penalty g, see §4 for the case of point sources. Note that one often considers the opposite problem, of reaching a boundary point $p \in \partial\Omega$ from x, which is equivalent up to replacing the vector field ω with its opposite in (2), see Definition 2.3 and the discussion below. The distance map $\mathbf{u}$ also obeys the first order non-linear Hamilton-Jacobi-Bellman equation

$$|\nabla \mathbf{u} - \omega|_{M^{-1}} = 1 \text{ in } \Omega, \qquad \mathbf{u} = g \text{ on } \partial\Omega, \tag{5}$$

in the sense of viscosity solutions (possibly with discontinuous boundary conditions) [3], which is numerically tractable [24, 25] as well. The point of this paper is however to study the linear approach (3) which has a number of advantages, see Remark 1.2. We present a finite differences

discretization of (3) on the Cartesian grid $\Omega_h := \Omega \cap h\mathbb{Z}^d$, of dimension $d \in \{2,3\}$, denoting by $h > 0$ the grid scale, reading

$$u + 2\varepsilon \sum_{1\le i\le I} \rho_i \langle A_b^{-1} b, e_i\rangle \; \overline{\delta}_h^{e_i} u - \varepsilon^2 \sum_{1\le i\le I} \rho_i \Delta_h^{e_i} u = 0 \quad \text{ on } \Omega_h, \tag{6}$$

where $\overline{\delta}_h^e$ and Δ_h^e denote standard centered and second order finite differences (26), modified close to $\partial\Omega$ to account for the Dirichlet boundary conditions, see (27) and (28). We denoted by $\rho_i(x) \ge 0$ and $e_i(x) \in \mathbb{Z}^d$, $1 \le i \le I = d(d+1)/2$ the weights and offsets of Selling's decomposition [29, 15] of the matrix $A_b(x)$, a tool from lattice geometry which is convenient for the design of anisotropic finite differences schemes [21, 23, 25, 9] in dimension $d \in \{2,3\}$, see Appendix B. Denoting by u_ε^h the solution of (6) we prove in Theorem 3.18 that $-\varepsilon \ln u_\varepsilon^h \to \mathbf{u}$ as $(\varepsilon, h/\varepsilon) \to 0$. The case of point sources also requires $\varepsilon \ln h \to 0$, see Theorem 4.1. The optimal consistency order is achieved when $\varepsilon = h^{\frac{2}{3}}$, see Corollary 3.14.

Finally we present in §5 an application to the optimal transport problem

$$\inf_{\gamma\in\Pi(\mu,\nu)} \int_{\Omega\times\Omega} c(x,y)\,\mathrm{d}\gamma(x,y), \qquad \text{with } c(x,y) := \mathrm{dist}_{\mathcal{F}}(x,y), \tag{7}$$

where μ and ν are given probability measures on Ω, and $\Pi(\mu,\nu)$ is the set of probability measures on $\Omega\times\Omega$ whose first and second marginals coincide respectively with μ and ν. The proposed implementation relies on Sinkhorn's matrix scaling algorithm [30], and the linear structure of (3). We emphasize that the matrix of costs $(c(x,y))_{x,y\in\Omega_h}$ is never numerically constructed, and would not fit in computer memory, but instead that the adequate matrix-vector product are evaluated by solving finite difference equations similar to (6), in an efficient manner thanks to a preliminary sparse matrix factorization. Let us acknowledge here that, in contrast with the Riemannian case [31], our approach does not extend to the quadratic cost $c(x,y) = \mathrm{dist}_{\mathcal{F}}(x,y)^2$. Indeed, this specific cost is handled in the Riemannian case using the short time asymptotic estimates of the diffusion equation, which becomes non-linear in the case of Randers geometry, see Remark 4.5, in contrast with the Poisson equation (3).

Contributions. We establish that the solution to a linear second order PDE converges, as a relaxation parameter $\varepsilon \to 0$ and after a logarithmic transformation, to the Randers distance from the domain boundary. We propose a finite differences discretization of that linear PDE, on a Cartesian grid of scale h, and establish convergence of the numerical solutions as $\varepsilon \to 0$ and $h/\varepsilon \to 0$, with optimal consistency when $\varepsilon = h^{\frac{2}{3}}$. We extend the approach to the case of point sources, under the additional condition $\varepsilon \ln h \to 0$. We propose a computational method for optimal transport with Randers distance as cost. Numerical experiments illustrate our results.

Outline. We recall in §2 the definition of Randers distances and introduce an extension of Varadhan's formula to Randers manifolds. We describe the coefficients of (3) in terms of the Randers metric (2), and prove the convergence result (4).

We study in §3 the linear finite-difference scheme (6). We show that a logarithmic transformation of the solution (6) solves another nonlinear scheme, for which we prove convergence and consistency with the non-linear PDE (5). We also discuss heuristic techniques introduced in [17] to improve the numerical accuracy of the geodesic distance approximation, and extend them to Randers metrics.

We address in §4 the computation of the geodesic distance from a point source, and in §5 the discretization of the optimal transportation problem (7), extending [31] which is limited to Riemannian distance costs.

Finally, we illustrate in §6 our results with numerical experiments, devoted to the computation of Randers distances and of the corresponding geodesic paths, and to the solution of the optimal transport problem (7) on a Randers manifold.

Remark 1.1 (Applications of Randers metrics)**.** *Randers metrics are, arguably, the simplest model of a* non-symmetric metric*, often referred to as a quasi-metric, see Remark 1.3. They play a key role in Zermelo's problem [2] of path planning subject to a drift, see §2.2, but also have numerous independent applications, of which we can only give a glimpse here. The common feature of these applications is that the paths are naturally endowed with an orientation.*

The boundary of an image region, oriented trigonometrically, minimizes the classical Chan-Vese segmentation functional iff it is a minimal geodesic for a suitable Randers metric, leading to a robust numerical optimization method [13]. The Euler-Mumford elastica minimal path model, whose cost is defined by integrating the squared curvature (plus a constant), is a limiting case of a Randers model, which allows the numerical computation of global minimizers with applications to tubular structure extraction in images [11]. Quantum vortex filaments, in a suitable limit and under appropriate assumptions, follow Randers geodesics, where the asymmetric part of the metric is derived from the magnetic field [1]. Finally, let us mention that Randers metrics were introduced in the context of general relativity, where the trajectory orientation stems from the time coordinate induced by the Minkowski space-time quadratic form [28].

Remark 1.2 (Advantages of linear schemes for distance map computation)**.** *Distance maps are ubiquitous in mathematics and their applications, and a variety of approaches have been proposed for their numerical computation [18], including Randers distances [24, 25]. The use of a linear PDE* (3)*, is here largely motivated by its application to the optimal transport problem* (7)*, but this approach has other advantages, see [17] for a more detailed discussion:*

- *(Ease of implementation) While we limit here our attention to domains discretized on Cartesian grids, geodesic distance computation also makes sense on manifolds presented as triangulations [17], patch based surfaces, etc. In that context, discretizing the non-linear PDE* (5) *can be challenging, whereas standard tools are often available for linear PDEs such as* (6)*.*

- *(Computation speed) Factorization techniques for sparse linear systems of equations are a subject of continued research, including non-symmetric Laplacian-like operators [14]. Once the linear system* (6) *is factored, it can be solved for a modest cost with a large number of right-hand sides, for instance to compute all pairwise Randers distances within a set of points, or when addressing the optimal transport problem* (7) *using Sinkhorn's matrix scaling algorithm as described in §5.*

- *(Solution smoothness) The distance map* $\mathbf{u}$ *defined by* (4) *is non-differentiable across the cut-locus*[1]*, and numerical solvers [24, 25] of the generalized eikonal PDE* (5) *produce non-smooth approximations of it. In contrast, the solution to the linear equation* (3) *is smooth and yields a natural regularizations* $\mathbf{u}_\varepsilon := -\varepsilon \ln u_\varepsilon$*, for any* $\varepsilon > 0$*, of the limit distance map* $\mathbf{u}$*.*

Remark 1.3 (*quasi-* prefix and asymmetric geometry)**.** *Non-symmetric norms, metrics and distances are often referred to as* quasi-*norms,* quasi-*metrics and* quasi-*distances. However, we drop the prefix "quasi" in this paper for the sake of readability and uniformity.*

[1]The cut locus is the set of points where the minimum (4, right) is attained by several minimal paths from the boundary.

Conventions and notations. We denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^d$, and by $\mathcal{S}_d$, $\mathcal{S}_d^+$, and $\mathcal{S}_d^{++}$ the sets of symmetric, symmetric positive semidefinite, and symmetric positive definite matrices of size $d \times d$ respectively. For any $A, B \in \mathcal{S}_d$, the relation $A \succeq B$ stands for $A - B \in \mathcal{S}_d^+$ (resp. $A \succ B$ stands for $A - B \in \mathcal{S}_d^{++}$), which is the Loewner order on symmetric matrices. For any $A \in \mathcal{S}_d^{++}$ and $b \in \mathbb{R}^d$, we define

$$\|A\| := \sup_{|x| \leq 1} |Ax|, \qquad\qquad |b|_A := \langle b, Ab \rangle^{1/2}.$$

From now on, we consider an open, bounded, connected, and nonempty domain $\Omega \subset \mathbb{R}^d$ with a $W^{3,\infty}$ boundary. The unknowns to the partial differential equations, and to their discretization schemes, are distinguished by typography: u for the linear second order PDEs (3) or numerical scheme (6) and variants, and $\mathbf{u}$ for the non-linear PDE (5) and related.

2 Elements of Randers geometry

A Randers metric is defined as the sum of a Riemannian metric, and of a suitably bounded linear term (2). We present §2.1 these geometrical objects in more detail, making connections with Zermelo's navigation problem §2.2. The eikonal equation (5) is discussed in §2.3, and its linear variant (3) in §2.4. We establish in Theorem 2.12 the existence of a solution u_ε to the linear PDE (3), and the convergence of $\mathbf{u}_\varepsilon = -\varepsilon \ln u_\varepsilon$ to the value function of the arrival time problem (4) as the relaxation parameter $\varepsilon > 0$ vanishes. The proof, based on the theory of viscosity solutions to degenerate elliptic PDEs, is postponed to Appendix A.

Before specializing to the case of Randers geometry, we briefly recall here the generic or *Finslerian* definition of a *non-symmetric* norm, dual-norm, metric, and path-length distance, and some of their elementary properties.

Definition 2.1. *A function $F : \mathbb{R}^d \to \mathbb{R}_+$ is a norm iff it is convex, 1-homogeneous, and vanishes only at the origin. The dual norm $F^* : \mathbb{R}^d \to \mathbb{R}_+$ is defined for all $v \in \mathbb{R}^d$ by*

$$F^*(v) := \max_{F(w) \leq 1} \langle v, w \rangle. \tag{8}$$

Equivalently, by homogeneity of F, one has $F^*(v) := \max\{\langle v, w \rangle / F(w);\ |w| = 1\}$. Conventionally, the above defines a *quasi*-norm, whereas a norm is subject to the additional symmetry axiom $F(v) = F(-v)$ for all $v \in \mathbb{R}^d$. However the prefix "quasi" before norms, metrics and distances is dropped in this paper for readability, as already mentioned in Remark 1.3. The following facts, stated without proof, are standard results of convex analysis and Finsler geometry.

Lemma 2.2 (Norm duality)**.** *Any norm F on $\mathbb{R}^d$ is Lipschitz continuous on $\mathbb{R}^d$, and as a result the extremum in* (8) *is indeed attained, for any $w \in \mathbb{R}^d$. The dual norm F^* is also a norm, and furthermore one has $F^{**} = F$ identically on $\mathbb{R}^d$.*

Definition 2.3. *A metric on a domain $\Omega \subset \mathbb{R}^d$ is a continuous map $\mathcal{F} : \overline{\Omega} \times \mathbb{R}^d \to \mathbb{R}_+$, denoted $(x, v) \mapsto \mathcal{F}_x(v)$, such that $\mathcal{F}_x$ is a norm on $\mathbb{R}^d$ for all $x \in \overline{\Omega}$. The dual metric $\mathcal{F}^*$ is defined pointwise from the dual norms. The related path length and distance are defined from* (1) *and denoted* $\mathrm{length}_{\mathcal{F}}$ *and* $\mathrm{dist}_{\mathcal{F}}$.

Let us emphasize that $\mathrm{dist}_{\mathcal{F}}(x, y) \neq \mathrm{dist}_{\mathcal{F}}(y, x)$ in general, for $x, y \in \overline{\Omega}$, since norms and metrics are not assumed here to be symmetric. In the special case where $\mathcal{F}_x = F$ is a constant metric, and $[x, y] \subset \overline{\Omega}$, one has $\mathrm{dist}_{\mathcal{F}}(x, y) = F(y - x)$.

Lemma 2.4 (Path-length distance)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded connected domain with smooth boundary and equipped with a metric $\mathcal{F}$. Then the extremum* (1) *defining* $\mathrm{dist}_{\mathcal{F}}(x,y)$ *is attained, for any $x, y \in \overline{\Omega}$, and defines a distance over $\overline{\Omega}$. Furthermore there exists $0 < c \leq C$ such that $c|x-y| \leq \mathrm{dist}_{\mathcal{F}}(x,y) \leq C|x-y|$ for all $x, y \in \overline{\Omega}$.*

2.1 Algebraic structure of Randers metrics

Randers norms are defined by analogy to Randers metrics (2), as the sum of a symmetric part defined from a symmetric positive definite matrix, and of an anti-symmetric linear part.

Definition 2.5. *A Randers norm $F : \mathbb{R}^d \to \mathbb{R}^d$ takes the form $F(v) = |v|_M + \langle \omega, v \rangle$, where $M \in \mathcal{S}_d^{++}$, and $\omega \in \mathbb{R}^d$ is subject to $|\omega|_{M^{-1}} < 1$.*

The dual to a Randers norm also is a Randers norm, as shown by the following lemma, whose proof can be found in Proposition 4.1 of [24] and Appendix C of [25].

Lemma 2.6 (Randers dual norm [24, 25])**.** *The dual to a Randers norm F of parameters M, ω is also a Randers norm, of parameters A, b characterized by the following relation, where $\alpha > 0$:*

$$\begin{pmatrix} A & b \\ b^\top & 1 \end{pmatrix} = \alpha \begin{pmatrix} M & \omega \\ \omega^\top & 1 \end{pmatrix}^{-1}.$$

Note that α in (2.6) is determined as the inverse of the bottom right coefficient in $\begin{pmatrix} M & \omega \\ \omega^\top & 1 \end{pmatrix}^{-1}$. In the special case where $\omega = 0$, one obtains $A = M^{-1}$, $b = 0$, and $\alpha = 1$, recovering the well known fact that the dual to a Riemmanian norm is also a Riemannian norm, defined by the inverse symmetric matrix. The duality formula in Lemma 2.6 is only the first of a family of algebraic identities associated with Randers norms, presented in Lemma 2.7 below, and used to reformulate the PDEs (3) and (5). For that purpose, we need to introduce some notation. For any $A \in \mathcal{S}_d$, and any $b \in \mathbb{R}^d$ we let

$$A_b := A - bb^\top. \tag{9}$$

The Schur complement formula yields the following positive-definiteness equivalences:

$$A_b \succ 0 \quad \Leftrightarrow \quad \begin{pmatrix} A & b \\ b^\top & 1 \end{pmatrix} \succ 0 \quad \Leftrightarrow \quad (A \succ 0 \text{ and } |b|_{A^{-1}} < 1). \tag{10}$$

If $M : \overline{\Omega} \to S_d^{++}$ and $\omega : \overline{\Omega} \to \mathbb{R}^d$ are Lipschitz fields obeying $|\omega|_{M^{-1}} < 1$ pointwise on $\overline{\Omega}$ (which is compact by assumption), then the fields $A : \overline{\Omega} \to S_d^{++}$ and $b : \overline{\Omega} \to \mathbb{R}^d$ defined by Lemma 2.6 as the dual Randers parameters are also Lipschitz, since matrix inversion is differentiable, and obey the equivalent properties (10) pointwise on $\overline{\Omega}$, thus $|b|_{A^{-1}} < 1$. The following lemma provides several equivalent characterizations of the unit ball associated with a Randers norm, and ends this subsection.

Lemma 2.7. *Let M, ω denote the parameters of a Randers norm, and A, b the parameters of the dual Randers norm, see Lemma 2.6. Then for all $v \in \mathbb{R}^d$*

$$\Big[|v|_M + \langle \omega, v \rangle - 1\Big] \quad \propto \quad \Big[|v|^2_{M_\omega} + 2\langle \omega, v \rangle - 1\Big] \quad \propto \quad \Big[|v - b|_{A^{-1}} - 1\Big] \tag{11}$$

$$\Big[|v|_A + \langle b, v \rangle - 1\Big] \quad \propto \quad \Big[|v|^2_{A_b} + 2\langle b, v \rangle - 1\Big] \quad \propto \quad \Big[|v - \omega|_{M^{-1}} - 1\Big], \tag{12}$$

where $x \propto y$ means that $\mathrm{sign}(x) = \mathrm{sign}(y)$, with $\mathrm{sign} : \mathbb{R} \to \{-1, 0, 1\}$ the sign function.

Proof. Note that the second line can be deduced from the first one, by exchanging the role of the Randers norm and of its dual norm. The positive definiteness of A_b and M_ω follows from (10) and Definition 2.5. Under the assumptions of the lemma, one has the equivalences

$$|v|_M + \langle \omega, v\rangle - 1 \leq 0 \;\Leftrightarrow\; |v|_M \leq 1 - \langle \omega, v\rangle \;\Leftrightarrow\; |v|_M^2 \leq (1 - \langle \omega, v\rangle)^2 \;\Leftrightarrow\; |v|_{M_\omega} + 2\langle \omega, v\rangle - 1 \leq 0,$$

and likewise with strict inequalities, which implies (11, left equivalence). The only difficulty lies in the reverse implication of the second equivalence: we must exclude the case where $|v|_M \leq \langle \omega, v\rangle - 1$, and indeed this is in contradiction with $|\langle \omega, v\rangle| \leq |\omega|_{M^{-1}}|v|_M < |v|_M + 1$ since $|\omega|_{M^{-1}} < 1$ by assumption.

Denoting by F the Randers norm of parameters M, ω, and by F^* the dual norm, one has

$$|v - b|_{A^{-1}} \leq 1 \Leftrightarrow \big(\forall w, \langle w, v - b\rangle \leq |w|_A\big) \Leftrightarrow \big(\forall w, \langle w, v\rangle \leq |w|_A + \langle b, w\rangle := F^*(w)\big) \Leftrightarrow F(v) \leq 1,$$

where implicitly $w \in \mathbb{R}^d$. In the last equivalence we used $F(v) = F^{**}(v) = \max\{\langle v, w\rangle; F^*(w) \leq 1\}$. A similar equivalence can be obtained with strict inequalities for any $w \neq 0$, which concludes the proof of (11, right equivalence) and of this lemma. □

2.2 Zermelo's navigation problem

Zermelo [2] considers a vehicle able to move at speed at most $c(x)$ *relative* to a given medium, which itself is subject to a drift $\eta(x)$, where $x \in \overline{\Omega}$ is the position. Typically, the vehicle is described as a boat subject to water currents, or as a flying object subject to air currents.

The set admissible *absolute* velocities v at the point x is characterized by the following relation

$$|v - \eta(x)| \leq c(x). \tag{13}$$

Given two endpoints $x, y \in \overline{\Omega}$, Zermelo's navigation problem asks for the smallest time $T = T_c^\eta(x, y) \geq 0$ such that there exists $\gamma \in \mathrm{Lip}([0, T], \overline{\Omega})$ obeying $|\gamma'(t) - \eta(\gamma(t))| \leq c(\gamma(t))$ for a.e. $t \in [0, T]$, and $\gamma(0) = x$, $\gamma(T) = y$. In other words, $T_c^\eta(x, y)$ is the minimal time from x to y subject ot the velocity constraints (13).

The vehicle described by Zermelo's problem is locally controllable at $x \in \Omega$ iff $|\eta(x)| < c(x)$, in other words iff the drift velocity norm is smaller than the maximum relative vehicle speed. In that case, it can be reformulated as a Randers minimal path problem.

Proposition 2.8. *Let $c : \overline{\Omega} \to \mathbb{R}$, and $\eta : \overline{\Omega} \to \mathbb{R}^d$ be continuous and obey $c > 0$ and $|\eta| < c$, pointwise on $\overline{\Omega}$. Consider the Randers metric $\mathcal{F}^*$ of parameters $A = c^2 I_d$ and $b = \eta$ on Ω, as well as its dual $\mathcal{F}^{**} = \mathcal{F}$. Then for all $x, y \in \overline{\Omega}$*

$$T_c^\eta(x, y) = \mathrm{dist}_{\mathcal{F}}(x, y).$$

Proof. Let $M : \overline{\Omega} \to S_d^{++}$ and $\omega : \overline{\Omega} \to \mathbb{R}^d$ be parameters of the Randers metric $\mathcal{F}$. The distance $\mathrm{dist}_{\mathcal{F}}(x, y)$ is the smallest time T for which there exists a path $\gamma \in \mathrm{Lip}([0, T], \overline{\Omega})$ obeying $1 \geq \mathcal{F}_{\gamma(t)}(\gamma'(t)) := |\gamma'(t)|_{M(\gamma(t))} + \langle \omega(\gamma(t)), \gamma'(t)\rangle$ for a.e. $t \in [0, T]$, and $\gamma(0) = x$, $\gamma(T) = y$. Indeed, this follows from the definition (1) and by reparametrization of any Lipschitz path at unit speed w.r.t. the metric $\mathcal{F}$. From this point, the announced result follows from the equivalence of $1 \geq \mathcal{F}_x(v) := |v|_{M(x)} + \langle \omega(x), v\rangle$ with (13), established in (11). □

2.3 The Eikonal equation

Consider a domain Ω, equipped with a Randers metric $\mathcal{F}$ with Lipschitz coefficients on Ω, and penalty function $g \in C^0(\partial\Omega, \mathbb{R})$. We are interested in the following value function $\mathbf{u} : \overline{\Omega} \to \mathbb{R}$, corresponding to the minimal time to reach $x \in \overline{\Omega}$ from a boundary point $p \in \partial\Omega$, with initial time penalty $g(p)$:

$$\mathbf{u}(x) := \min_{p \in \partial\Omega} g(p) + \mathrm{dist}_{\mathcal{F}}(p, x). \tag{14}$$

We prove in Theorem A.9 that (14) is a viscosity solution, see Definition A.2, to the first order non-linear PDE

$$\mathcal{F}_x^*(\nabla \mathbf{u}(x)) = 1 \text{ for all } x \in \Omega, \qquad \mathbf{u}(x) = g(x) \text{ for all } x \in \partial\Omega. \tag{15}$$

The boundary condition $\mathbf{u} = g$ on $\partial\Omega$ is satisfied in a strong sense if $g(x) \leq g(p) + \mathrm{dist}_{\mathcal{F}}(p, x)$ for all $x, p \in \partial\Omega$, but in the weak sense of Definition A.2 otherwise. The comparison principle Theorem A.8 implies that the viscosity solution is uniquely determined in Ω.

Corollary 2.9. *If $\mathcal{F}$ is a Randers metric of parameters M, ω, and dual parameters A, b, then the eikonal PDE (15, left) admits the following three equivalent formulations in Ω in the sense of viscosity solutions*

$$|\nabla \mathbf{u}|_A + \langle \nabla \mathbf{u}, b \rangle = 1, \qquad |\nabla \mathbf{u}|_{A_b}^2 + 2\langle \nabla \mathbf{u}, b \rangle = 1, \qquad |\nabla \mathbf{u} - \omega|_{M^{-1}} = 1. \tag{16}$$

Proof. The equation $\mathcal{F}_x^*(\nabla \mathbf{u}(x)) = 1$ is a shorthand for (16, left) at $x \in \Omega$, see Definition 2.5 of a Randers norm. It is equivalent to (16, center and right) by (12). □

In applications, computing the value function (14) is often only a means to obtain the globally optimal path γ from $\partial\Omega$ to an arbitrary point $x_* \in \Omega$. This path can be extracted by solving, backwards in time, the following Ordinary Differential Equation (ODE), see e.g. [20, Appendix C]

$$\gamma'(t) = V(\gamma(t)), \qquad \text{where } V(x) := \mathrm{d}\mathcal{F}_x^*(\nabla \mathbf{u}(x)) \tag{17}$$

for all $x \in \overline{\Omega}$. The ODE needs to be solved on the interval $[0, T]$ where $T = \mathbf{u}(x_*)$, with terminal condition $\gamma(T) = x_*$. By $\mathrm{d}\mathcal{F}_x^*(w)$ we denote the derivative of $\mathcal{F}_x^*$ w.r.t. the variable w, where $x \in \overline{\Omega}$ is fixed.

Corollary 2.10. *The following expressions are positively proportional to the geodesic flow V defined by (17, right), at all points where $\mathbf{u}$ is differentiable*

$$\frac{A\nabla \mathbf{u}}{|\nabla \mathbf{u}|_A} + b, \qquad A_b \nabla \mathbf{u} + b, \qquad M^{-1}(\nabla \mathbf{u} - \omega). \tag{18}$$

Proof. Fix a point $x \in \Omega$ where $\mathbf{u}$ is differentiable, and denote $v := \nabla \mathbf{u}(x)$. Introduce the Randers norm $F^* = \mathcal{F}_x^*$ whose parameters are denoted $A \in S_d^{++}$ and $b \in \mathbb{R}^d$, in such way that $F^*(v) = 1$ by (15). Differentiating $F^*(v) = |v|_A + \langle b, v \rangle$ we obtain $\mathrm{d}F(v) = Av/|v|_A + b$ which yields (18, left). The three expressions (12) vanish, and their respective gradients w.r.t. v are $g_1 := Av/|v|_A + b$, $g_2 := 2(A_b v + b)$ and $g_3 := M^{-1}(v - \omega)/|v - \omega|_{M^{-1}}$. These gradients are non-zero since $\langle v, g_1 \rangle = F^*(v) = 1$, $\langle v, g_2 \rangle = 1 + |v|_{A_b}^2 \geq 1$ and $\langle v - \omega, g_3 \rangle = |v - \omega|_{M^{-1}}^2 = 1$. Since g_1, g_2 and g_3 are orthogonal to the same level set, and point outward of it, they are positively proportional. The result follows. □

2.4 Varadhan's formula

Varadhan's formula is based on a logarithmic transformation of the unknown [32], which turns the linear PDE (19) into the non-linear PDE (20). The point of this transformation is that, with a proper scaling of the unknown and the PDE coefficients, a relaxation parameter $\varepsilon > 0$ is eliminated from the boundary conditions and from all the PDE coefficients except one, of principal order.

Lemma 2.11 (Logarithmic transformation)**.** *Let $\varepsilon > 0$, and let u_ε be a viscosity solution to*

$$u + 2\varepsilon\langle \nabla u, b\rangle - \varepsilon^2 \operatorname{Tr}(A_b \nabla^2 u) = 0 \text{ in } \Omega, \qquad u = \exp(-g/\varepsilon) \text{ on } \partial\Omega, \tag{19}$$

where $\Omega \subset \mathbb{R}^d$ is a smooth bounded domain, $A_b : \overline{\Omega} \to S_d^{++}$ and $b : \overline{\Omega} \to \mathbb{R}^d$ are Lipschitz, and $\varepsilon > 0$. Then $\mathbf{u}_\varepsilon := -\varepsilon \ln u_\varepsilon$ is a viscosity solution to the PDE

$$|\nabla \mathbf{u}|^2_{A_b} + 2\langle \nabla \mathbf{u}, b\rangle - \varepsilon \operatorname{Tr}(A_b \nabla^2 \mathbf{u}) = 1 \text{ in } \Omega, \qquad \mathbf{u} = g \text{ on } \partial\Omega. \tag{20}$$

Lemma 2.11 is an immediate consequence of Corollary A.5 established in Appendix A. For later convenience, we introduce the following PDE operators on the domain Ω

$$\mathcal{L}^\varepsilon u = u + 2\varepsilon\langle \nabla u, b\rangle - \varepsilon^2 \operatorname{Tr}(A_b \nabla^2 u), \quad \mathcal{S}^\varepsilon \mathbf{u} = |\nabla \mathbf{u}|^2_{A_b} + 2\langle \nabla \mathbf{u}, b\rangle - \varepsilon \operatorname{Tr}(A_b \nabla^2 \mathbf{u}) - 1, \tag{21}$$

and observe that, formally, one has $\mathcal{S}^\varepsilon \mathbf{u} = -e^{\frac{\mathbf{u}}{\varepsilon}} \mathcal{L}^\varepsilon (e^{-\frac{\mathbf{u}}{\varepsilon}})$. The following result relies on the framework of viscosity solutions to take the limit $\varepsilon \to 0$ in $\mathcal{S}^\varepsilon$, letting the second order "viscous" term $-\varepsilon \operatorname{Tr}(A_b \nabla^2 \mathbf{u})$ vanish, and recovering in the limit a first order non-linear equation equivalent to the Randers eikonal equation, see Corollary 2.9.

Theorem 2.12 (Vanishing viscosity limit)**.** *The PDE* (19) *admits a unique viscosity solution in Ω. In addition $\mathbf{u}_\varepsilon := -\varepsilon \ln u_\varepsilon$ converges locally uniformly in Ω to the value function* (14)*, associated with the Randers metric $\mathcal{F}$ whose dual metric $\mathcal{F}^*$ has parameters A, b.*

The elements of proof relying on the concept of viscosity solutions are postponed to Appendix A. In particular, uniqueness of the solution to (19) follows from the comparison principle Proposition A.7, see [4]. Convergence as $\varepsilon \to 0$ is established in Theorem A.12. We limit our attention here to the existence of a solution to (19), which is based on the interpretation of u_ε as an expectation of a cost associated with a stochastic process. Fix $\varepsilon > 0$, and introduce the stochastic process $(X_t^{x,\varepsilon})_{t \geq 0}$

$$dX_t^{x,\varepsilon} = 2\varepsilon b(X_t^{x,\varepsilon})\, dt + \varepsilon\sqrt{2A_b}(X_t^{x,\varepsilon})\, dW_t, \qquad X_0^{x,\varepsilon} = x, \tag{22}$$

where $(W_t)_{t \geq 0}$ is a d-dimensional Wiener process. Define also the exit time $\tau^{x,\varepsilon}$ by

$$\tau^{x,\varepsilon} := \inf\{t \geq 0;\, X_t^{x,\varepsilon} \notin \Omega\}.$$

Since Ω is bounded, and A_b is positive definite, the exit time $\tau^{x,\varepsilon}$ is almost surely finite. Thus $X_t^{x,\varepsilon}$ is a Brownian motion starting at x, with drift $2\varepsilon b$, and whose fluctuations are scaled by $\varepsilon\sqrt{2A_b}$. According to the Feynman-Kac formula, see Theorem A.11 in Appendix A, the following expectation is the unique solution to the PDE (19)

$$u_\varepsilon(x) = \mathbb{E}\left[\exp\left(-\tau^{x,\varepsilon} - \frac{g(X_{\tau^{x,\varepsilon}}^{x,\varepsilon})}{\varepsilon}\right)\right]. \tag{23}$$

In particular, u_ε is positive. In the framework of the stochastic approach, Theorem 2.12 expresses the convergence of the following soft-mininimum

$$\mathbf{u}_\varepsilon(x) = -\varepsilon \log\left(\mathbb{E}\left[\exp\left(-\tilde{\tau}^{x,\varepsilon} - \frac{g(\tilde{X}^{x,\varepsilon}_{\tilde{\tau}^{x,\varepsilon}})}{\varepsilon}\right)\right]\right), \tag{24}$$

towards the minimum (14) defining the value function $\mathbf{u}$.

Remark 2.13 (Divergence form Laplacian)**.** *One may replace in* (19) *the non-divergence form anisotropic Laplacian with the divergence form variant*

$$\operatorname{div}(A_b \nabla u) = \operatorname{Tr}(A_b \nabla^2 u) + \langle \operatorname{div}(A_b), \nabla u \rangle,$$

where $\operatorname{div}(A_b)$ *denotes column-wise divergence, assuming that* A_b *is continuously differentiable. Indeed, this amounts to replacing in* (19) *the vector field* b *defining the first order term with* $b_\varepsilon := b - \frac{\varepsilon}{2}\operatorname{div}(A_b)$*. This small perturbation is easily handled in the setting of viscosity solutions, and the same limit* (4) *is obtained as* $\varepsilon \to 0$*.*

The divergence form Laplacian is often preferred in applications [17] since it is simpler to implement numerically on some geometries, such as triangulated surfaces using finite elements. Finite element methods may however lack the discrete comparison principle Lemma 3.4 used to establish the convergence of the numerical scheme in this paper.

3 The numerical scheme

We present a numerical implementation of the linear second order PDE (3) based on discrete degenerate elliptic finite differences, on a Cartesian discretization grid. This approach is chosen for the simplicity of its implementation and of the convergence analysis. Alternative discretizations may also be considered, for instance using finite elements on triangulated manifolds, see [17] and Remark 2.13.

Throughout this section, we denote by $h > 0$ the grid scale of the Cartesian discretization grid, which is fixed unless otherwise specified, and we define the discrete domain as

$$\Omega_h := \Omega \cap h\mathbb{Z}^d, \qquad\qquad \overline{\Omega}_h := \Omega_h \cup \partial\Omega. \tag{25}$$

In our application, the values of u on $\partial\Omega$ are given by the Dirichlet boundary conditions, and the numerical implementation does not treat them as unknowns. For any $u : \Omega_h \to \mathbb{R}$, any $x \in \Omega_h$ and any $e \in \mathbb{Z}^d$, we define the first order and second order centered finite differences operators as follows: assuming $[x - he,\, x + he] \subset \Omega$

$$\overline{\delta}^e_h u(x) := \frac{u(x+he) - u(x-he)}{2h}, \qquad \Delta^e_h u(x) := \frac{u(x+he) - 2u(x) + u(x-he)}{h^2}. \tag{26}$$

If x is adjacent to $\partial\Omega$, then (26) may involve values outside the domain Ω_h, and thus be ill defined. In order to address this issue, we consider $u : \overline{\Omega}_h \to \mathbb{R}$ which is also defined on the domain boundary. The following finite difference expressions make sense for arbitrary $x \in \Omega_h$, $e \in \mathbb{Z}^d$, and they reduce to (26) if $[x - he,\, x + he] \subset \Omega$:

$$\overline{\delta}^e_h u(x) := \frac{1}{2}\Big(\frac{u(x + h^e_x e) - u(x)}{h^e_x} - \frac{u(x - h^{-e}_x e) - u(x)}{h^{-e}_x}\Big), \tag{27}$$

$$\Delta^e_h u(x) := \frac{2}{h^e_x + h^{-e}_x}\Big(\frac{u(x + h^e_x e) - u(x)}{h^e_x} + \frac{u(x - h^{-e}_x e) - u(x)}{h^{-e}_x}\Big), \tag{28}$$

where we denoted

$$h_x^e := \min\{\eta > 0;\, x + \eta e \in \overline{\Omega}_h\}. \tag{29}$$

Note that $h_x^e \in]0, h]$ by construction. If $u \in C^4(\overline{\Omega})$ then one has the consistency relation

$$\overline{\delta}_h^e u(x) = \langle \nabla u(x), e\rangle + \mathcal{O}(h^r), \qquad \Delta_h^e u(x) = \langle e, \nabla^2 u(x) e\rangle + \mathcal{O}(h^r),$$

where $r = 2$ if $[x - he, x + he] \subset \overline{\Omega}$, and $r = 1$ otherwise. In the next proposition we obtain, by linear combination, consistent finite differences approximations of linear PDE operators of order one and two.

Proposition 3.1. *Let $D \in S_d$, and let $\omega \in \mathbb{R}^d$. Consider weights ρ_i and offsets $e_i \in \mathbb{Z}^d$, for all $1 \le i \le I$, such that*

$$D = \sum_{1 \le i \le I} \rho_i e_i e_i^\top. \tag{30}$$

Then for $u \in C^4(\overline{\Omega})$ and $x \in \Omega_h$ one has

$$\sum_{1 \le i \le I} \rho_i\, \overline{\delta}_h^{e_i} u(x)\, e_i = D\nabla u(x) + \mathcal{O}(h^r), \qquad \sum_{1 \le i \le I} \rho_i \Delta_h^{e_i} u(x) = \mathrm{Tr}(D\nabla^2 u(x)) + \mathcal{O}(h^r), \tag{31}$$

where $r = 2$ if $[x - he_i, x + he_i] \subset \overline{\Omega}$ for all $1 \le i \le I$, and $r = 1$ otherwise.

As an immediate application, we define a finite differences discretization $\mathcal{L}_h^\varepsilon$ of the linear operator $\mathcal{L}^\varepsilon$ defined in (21). For any $u : \overline{\Omega}_h \to \mathbb{R}$ we let

$$\mathcal{L}_h^\varepsilon u = u + 2\varepsilon \sum_{1 \le i \le I} \rho_i \langle A_b^{-1} b, e_i\rangle\, \overline{\delta}_h^{e_i} u - \varepsilon^2 \sum_{1 \le i \le I} \rho_i \Delta_h^{e_i} u. \tag{32}$$

with boundary condition $u = \exp(-g/\varepsilon)$ on $\partial\Omega$. The weights $\rho_i = \rho_i(x)$ and offsets $e_i = e_i(x)$, $1 \le i \le I$, provide a decomposition of the matrix $A_b = A_b(x)$ in the sense of (30). Note that for the schemes (31) to be well defined, it is crucial that the offsets involved in (30) have integer coordinates, and therefore the similar looking eigenvalue-eigenvector decomposition typically cannot be used since it involves arbitrary unit vectors. Obtaining a suitable decomposition is thus non-trivial in general, and it is also not unique. We rely in this paper on Selling's decomposition, which is defined in dimension $d \in \{2, 3\}$, and has the additional benefit of producing non-negative weights $(\rho_i)_{1 \le i \le I}$ and thus a discrete degenerate elliptic scheme, see §3.1 below.

Remark 3.2 (Approximation of the gradient, improved reconstruction, following [17])**.** *An approximate gradient $V_h^\varepsilon : \Omega_h \to \mathbb{R}^d$ of the solution u_h^ε of* (32) *can be estimated using (31, left):*

$$V_h^\varepsilon(x) := A_b(x)^{-1} \sum_{1 \le i \le I} \rho_i\, \overline{\delta}_h^{e_i} u_h^\varepsilon(x), \qquad \mathbf{V}_h^\varepsilon(x) := \frac{-V_h^\varepsilon(x)}{|\nabla V_h^\varepsilon(x)|_{A(x)} - \langle b(x), V_h^\varepsilon(x)\rangle}, \tag{33}$$

The negated and normalized vector field $\mathbf{V}_h^\varepsilon$ is meant to approximate the gradient of Randers distance $\mathbf{u}$ from the boundary (4)*: it is negatively proportional to V_h^ε, reflecting the fact that logarithmic transformation is decreasing, and is normalized consistently with Randers eikonal equation* (15)*. An empirical observation of [17], in the context of isotropic and Riemannian metrics which are special cases of Randers metrics (and using a different discretization), is that $\mathbf{V}_h^\varepsilon$ is for suitable parameters h, ε an excellent approximation of $\nabla \mathbf{u}$. In particular, it can be used*

for geodesic backtracking via (17) *and* (18). *Following [17] we may also obtain an empirically improved reconstruction* $\mathbf{v}_h^\varepsilon : \Omega_h \to \mathbb{R}$ *of the Randers distance by minimizing*

$$\sum_{x\in\Omega_h} \sum_{1\le|i|\le I} \rho_i |\delta_h^{e_i}\mathbf{v}(x) - \langle e_i, \mathbf{V}_h^\varepsilon(x)\rangle|^2, \tag{34}$$

which is consistent with $\int_\Omega |\nabla \mathbf{v} - \mathbf{V}_h^\varepsilon|^2_{A_b}$, *where* $\rho_{-i} := \rho_i$ *and* $e_{-i} := e_i$ *for all* $1 \le i \le I$, *and where the first order upwind finite difference* δ_h^e *is defined in* (35). *Equations (33, left) and* (34) *also make sense if one replaces the weights and offsets* $(\rho_i, e_i)_{i=1}^I$ *and matrix* A_b *used in the numerical scheme* (32), *with unit weights and the canonical basis and the identity matrix. However, the latter (and simpler) choice yields slightly less accurate results empirically as evidenced in our numerical experiments §6. In Figs. 3 and 4 we refer to these post-processed distance maps as* $\mathbf{u}_h^{A_b}$ *and* $\mathbf{u}_h^{I_2}$ *respectively.*

3.1 Discrete degenerate ellipticity

Discrete degenerate ellipticity is a counterpart to the degenerate ellipticity property of Hamilton-Jacobi-Bellman PDE operators [16, 26], which is at the foundation of the theory of viscosity solutions, see Definition A.1.

Definition 3.3 (Discrete degenerate ellipticity [26])**.** *Let* X *be a finite set, and let* $\mathbb{U} := \mathbb{R}^X$. *A (finite difference) scheme on* X *is a function* $F : \mathbb{U} \to \mathbb{U}$. *Such a function can be written in the form*

$$Fu(x) := \tilde{F}(x, u(x), (u(x) - u(y))_{y\in X\setminus\{x\}}),$$

and the scheme is said discrete degenerate elliptic (DDE) if $\tilde{F}$ *is non-decreasing w.r.t. the second variable, and w.r.t. the third variable componentwise. The scheme is said elliptic if* $u \mapsto Fu - \lambda u$ *is degenerate elliptic for some* $\lambda > 0$.

Similarly to its continuous counterpart, discrete ellipticity implies a comparison principle, used in the proof of the existence and uniqueness of solutions to discretized PDEs, and of their convergence to the continuous solutions as the grid scale is refined §3.3. For completeness we present the proof of two basic but fundamental properties of discrete elliptic operators, see e.g. [26] for additional discussion.

Lemma 3.4 (Discrete comparison principle)**.** *Let* F *be an elliptic finite differences scheme on a finite set* X, *and let* $u, v : X \to \mathbb{R}$. *If* $Fu \le Fv$ *on* X, *then* $u \le v$ *on* X.

Proof. Let $x_* \in X$ be such that $u(x_*) - v(x_*)$ is maximal, so that $u(x_*) - u(y) \ge v(x_*) - v(y)$ for all $y \in X$. Assume for contradiction that $u(x_*) > v(x_*)$, otherwise the result is proved. Then, by discrete degenerate ellipticity of $F - \lambda\,\mathrm{Id}$, we obtain $Fu(x_*) - \lambda u(x_*) \ge Fv(x_*) - \lambda v(x_*)$, thus $0 < \lambda(u(x_*) - v(x_*)) \le Fu(x_*) - Fv(x_*) \le 0$, which proves the result by contradiction. □

We say that u is a sub-solution (resp. super-solution, resp. solution) of the scheme F, if $Fu \le 0$ (resp. $Fu \ge 0$, resp. $Fu = 0$) on X.

Corollary 3.5 (Solution to elliptic linear operators)**.** *If* F *is an affine (i.e. linear plus constant) and elliptic scheme on a finite set* X, *then there exists a unique solution* $u : X \to \mathbb{R}$ *to* $Fu = 0$.

Proof. If $Fu = Fv$ on X then $u = v$, by Lemma 3.4. Thus $F : \mathbb{R}^X \to \mathbb{R}^X$ is injective, hence by linearity it is bijective, and there exists a unique solution to $Fu = 0$. □

The finite difference schemes considered in this paper (27), (28), and (32) formally involve a function defined on the uncountable set $\overline{\Omega}_h = \Omega_h \cup \partial\Omega$, which does not comply with the finiteness assumption in Definition 3.3. This obstruction is only superficial, since only finitely many boundary values of u are actually involved these schemes, for any given $h > 0$. Alternatively, one may consider the Dirichlet boundary values of u as given constants rather than unknown variables in the scheme.

The simplest DDE operator is the negated upwind finite difference $-\delta_h^e$ on Ω_h, where $h > 0$ and $e \in \mathbb{Z}^d$, which is defined as

$$\delta_h^e u(x) := \frac{u(x+he) - u(x)}{h}. \tag{35}$$

The operator δ_h^e is modified similarly to (27) and (28) if $[x, x+he] \not\subset \overline{\Omega}$, and is first order consistent with a directional derivative: for any $u : \overline{\Omega}_h \to \mathbb{R}$ and any $x \in \Omega_h$

$$\delta_h^e u(x) := \frac{u(x+h_x^e e) - u(x)}{h_x^e}, \qquad \delta_h^e u(x) = \langle e, \nabla u(x)\rangle + \mathcal{O}(h). \tag{36}$$

The negated second order finite difference operator $-\Delta_h^e$ is also DDE. The centered finite difference operator $\overline{\delta}_h^e$ is not DDE, but linear combinations with Δ_h^e whose coefficients have suitable signs and obey suitable bounds satisfy this property, as shown in the next lemma. For that purpose, we observe the relations

$$\Delta_h^e u(x) = \frac{2}{h_x^e + h_x^{-e}} \left(\delta_h^e u(x) + \delta_h^{-e} u(x)\right), \qquad \overline{\delta}_h^e u(x) = \frac{1}{2}\left(\delta_h^e u(x) - \delta_h^{-e} u(x)\right). \tag{37}$$

Lemma 3.6. *Let $e \in \mathbb{Z}^d$, and $h > 0$. The finite difference scheme $-\Delta_h^e$ is unconditionally DDE, and the linear combination $\mu\overline{\delta}_h^e - \lambda\Delta_h^e$ is DDE when $h|\mu| \le 2\lambda$.*

Proof. In view of (37) one has the equality of schemes $\mu\overline{\delta}_h^e - \lambda\Delta_h^e = -\alpha\delta_h^e - \beta\delta_h^{-e}$, where $\alpha : X \to \mathbb{R}$ is defined by $\alpha(x) := 2\lambda/(h_x^e + h_x^{-e}) - \mu/2$ which is non-negative if $h|\mu| \le 2\lambda$, since $0 < h_x^{\pm e} \le h$. Likewise $\beta(x) := 2\lambda/(h_x^e + h_x^{-e}) + \mu/2 \ge 0$ if $h|\mu| \le 2\lambda$. We conclude by observing that DDE schemes form a cone: linear combinations with non-negative coefficients of DDE schemes are DDE. □

Corollary 3.7. *The finite difference scheme $\mathcal{L}_h^\varepsilon$ defined by (32) is elliptic, with $\lambda = 1$, if $\rho_i \ge 0$ and $h|\langle A_b^{-1} b, e_i\rangle| \le \varepsilon$ for all $1 \le i \le I$.*

Proof. Under these assumptions, the finite difference scheme $u \mapsto \mathcal{L}_h^\varepsilon u - u$ is the sum of the finite difference operators $\varepsilon\rho_i(2\mu_i\overline{\delta}_h^{e_i} u - \varepsilon\Delta_h^{e_i})$ where $\mu_i = \langle A_b^{-1} b, e_i\rangle$, for all $1 \le i \le I$. By Lemma 3.6, which applies regardless of the fact that ρ_i and e_i depend on the point $x \in \Omega_h$, each of these elementary operators is DDE when $\rho_i \ge 0$ and $h|\mu_i| \le \varepsilon$. Hence $\mathcal{L}_h^\varepsilon - \mathrm{Id}$ is DDE, and therefore $\mathcal{L}_h^\varepsilon$ is elliptic with $\lambda = 1$ by Definition 3.3. □

As announced in the introduction of this section, and in order to benefit from Lemma 3.4 and Corollary 3.5, we do want the discrete operator $\mathcal{L}_h^\varepsilon$ to be DDE. For that purpose, we introduce Selling's decomposition [29, 15] of a positive definite matrix $D \in S_d^{++}$, where $d \in \{2, 3\}$, which is efficiently computable numerically via Selling's algorithm. In view of their key role in our numerical scheme, Selling's constructions and some of their properties are presented in more detail in Appendix B.

Theorem 3.8 (Selling [29], this version [23]). *Let $D \in S_d^{++}$, where $d \in \{2,3\}$. Then there exists non-negative weights $\rho_i \geq 0$, and offsets $e_i \in \mathbb{Z}^d$, where $1 \leq i \leq I := d(d+1)/2$, such that*

$$D = \sum_{1 \leq i \leq I} \rho_i e_i e_i^\top, \qquad\qquad |e_i| \leq 2C_d\, \mu(D),\ \forall 1 \leq i \leq I,$$

where $C_2 = 2$, $C_3 = 2\sqrt{3}$, and $\mu(D) := \sqrt{\|D\|\|D^{-1}\|}$ is the anisotropy ratio of D.

In the rest of this section, we assume that the weights and offsets $(\rho_i(x), e_i(x))_{i=1}^I$ used to define the scheme $\mathcal{L}_h^\varepsilon$, see (32), are obtained from Selling's decomposition of the matrix $A_b(x)$, for all $x \in \Omega_h$. For the sake of readability, the dependency of ρ_i and e_i w.r.t. the base point x is often left implicit in the equations. The following proposition, stated without proof, immediately follows from Corollary 3.7 and Theorem 3.8.

Proposition 3.9. *The scheme $\mathcal{L}_h^\varepsilon$ is elliptic provided that $Ch \leq \varepsilon$, where*

$$C := 2C_d \max_{x \in \overline{\Omega}} \mu(A_b(x))\, |A_b^{-1}(x) b(x)|.$$

The construction of finite difference schemes for linear and quasi-linear PDEs using Selling's algorithm, and the compatibility conditions ensuring the DDE property, are discussed in more detail in [7]. Finally, let us mention an alternative discretization of the PDE operator $\mathcal{L}^\varepsilon$ defined in (21), using upwind finite differences for the first order term, which is unconditionally stable but has a lower consistency order

$$\mathcal{L}_h^{\varepsilon,+} u = u - 2\varepsilon \sum_{1 \leq j \leq d} |\langle b, f_j \rangle| \delta_h^{-\sigma_j f_j} u - \varepsilon^2 \sum_{1 \leq i \leq I} \rho_i \Delta_h^{e_i} u, \tag{38}$$

where $(f_j)_{j=1}^d$ is the canonical basis of $\mathbb{R}^d$, and σ_j is the sign of $\langle b, f_j \rangle$.

3.2 Logarithmic transformation

We use a logarithmic transformation of the unknown to study the convergence of the solutions to the discrete schemes (32) and (38) as the relaxation parameter ε and the gridscale h tend to zero suitably, mimicking the approach used in the continuous case, see §2.4. Our first step is to describe the effect of the logarithmic/exponential transformation on a finite difference scheme.

Proposition 3.10. *Let $h > 0$ and $\varepsilon > 0$. Let F be a DDE scheme on $\overline{\Omega}_h$, such that $Fu(x)$ is a* linear *function of u for all $x \in \Omega_h$, with boundary condition $u = \exp(-g/\varepsilon)$ on $\partial\Omega$, where $u : \overline{\Omega}_h \to \mathbb{R}$. We define the exponentially transformed scheme $\boldsymbol{F}^\varepsilon$ as follows:*

$$\begin{aligned} \boldsymbol{F}^\varepsilon \mathbf{u}(x) &:= -e^{\frac{\mathbf{u}(x)}{\varepsilon}} \left[F e^{\frac{-\mathbf{u}}{\varepsilon}} \right](x) \\ &= \tilde{F}\Big(x, -1, \Big[\exp\big(\frac{\mathbf{u}(x) - \mathbf{u}(y)}{\varepsilon}\big) - 1 \Big]_{y \in X \setminus \{x\}} \Big), \end{aligned} \tag{39}$$

for any $x \in \Omega_h$, with boundary condition $\mathbf{u} = g$ on $\partial\Omega$, where $\mathbf{u} : \overline{\Omega}_h \to \mathbb{R}$. The scheme $\boldsymbol{F}^\varepsilon$ is DDE, and furthermore if $\mathbf{u}$ is a sub-solution (resp. super-solution) of $\boldsymbol{F}^\varepsilon$, then $u := \exp(-\mathbf{u}/\varepsilon)$ is a super-solution (resp. sub-solution) of F.

Proof. The two expressions of $\boldsymbol{F}^\varepsilon \mathbf{u}(x)$ given in (39), where $x \in \Omega_h$, are equivalent in view of the linearity of $\tilde{F}$. The discrete degenerate ellipticity of $\boldsymbol{F}^\varepsilon$ follows from the same property of F, and from the fact that $t \in \mathbb{R} \mapsto \exp(t/\varepsilon) - 1$ is non-decreasing. □

We use the scheme unknown transformation $u = \exp(-\mathbf{u}/\varepsilon)$, which is classical in the study of relations between the heat, Poisson, and eikonal equations [32, 17]. However, since the mapping $t \mapsto \exp(-t/\varepsilon)$ is decreasing, it exchanges the notions of sub-solutions and super-solutions, see Proposition 3.10. The exponentially transformed upwind finite difference is denoted $\boldsymbol{\delta}_h^{e,\varepsilon}$, and reads

$$\boldsymbol{\delta}_h^{e,\varepsilon}\mathbf{u}(x) = \frac{1}{h}\Big(1 - \exp\big(\frac{\mathbf{u}(x) - \mathbf{u}(x+he)}{\varepsilon}\big)\Big), \tag{40}$$

where $x \in \Omega_h$, $e \in \mathbb{Z}^d$, and assuming $[x, x+he] \subset \overline{\Omega}$. Otherwise replace h with h_x^e in the above expression, see (29). The next lemma approximates (40) in terms of the derivatives of $\mathbf{u}$.

Lemma 3.11. *Let $\mathbf{u} \in C^3(\overline{\Omega})$ and $0 < h \le \varepsilon \le 1$. Then for any $x \in \Omega_h$, and bounded $e \in \mathbb{Z}^d$,*

$$\boldsymbol{\delta}_h^{e,\varepsilon}\mathbf{u}(x) = \frac{1}{\varepsilon}\langle \nabla \mathbf{u}(x), e\rangle + \frac{h}{2\varepsilon}\langle e, \nabla^2\mathbf{u}(x) e\rangle - \frac{h}{2\varepsilon^2}\langle \nabla \mathbf{u}(x), e\rangle^2 + \frac{h^2}{6\varepsilon^3}\langle \nabla \mathbf{u}(x), e\rangle^3 + \mathcal{O}\big(\frac{h^2}{\varepsilon^2} + \frac{h^3}{\varepsilon^4}\big), \tag{41}$$

assuming $[x, x+he] \subset \overline{\Omega}$. Otherwise, replace h with h_x^ε in the above expression.

Proof. The announced result immediately follows from (40) and the Taylor expansion $1 - \exp(-s) = s - \frac{1}{2}s^2 + \frac{1}{6}s^3 + \mathcal{O}(s^4)$, where s is defined by $\varepsilon s = \mathbf{u}(x+he) - \mathbf{u}(x) = h\langle \nabla \mathbf{u}(x), e\rangle + \frac{1}{2}h^2\langle e, \nabla^2\mathbf{u}(x) e\rangle + \mathcal{O}(h^3)$. □

The exponentially transformed second order and first order centered finite difference operators are denoted $\boldsymbol{\Delta}_h^{e,\varepsilon}$ and $\overline{\boldsymbol{\delta}}_h^{e,\varepsilon}$, and their Taylor expansion is deduced from that of $\boldsymbol{\delta}_h^{e,\varepsilon}$. The assumption $0 < h \le \varepsilon \le 1$ of Lemma 3.11 serves to eliminate spurious negligible terms in the Taylor expandion, and is asymptotically satisfied in convergence analysis Theorem 3.18 which requires $\varepsilon \to 0$ and $h/\varepsilon \to 0$. Note that if $\varepsilon = \mathcal{O}(\sqrt{h})$, as considered in Corollary 3.14 below, then the remainder in (41) (resp. (42) and (44) below) simplifies to $\mathcal{O}(h^3/\varepsilon^4)$ (resp. $\mathcal{O}(h^r/\varepsilon^{2+r})$ and $\mathcal{O}(h^r/\varepsilon^r)$).

Corollary 3.12. *Under the assumptions of Lemma 3.11, one has*

$$\begin{aligned} \boldsymbol{\Delta}_h^{e,\varepsilon}\mathbf{u}(x) &= \frac{1}{\varepsilon}\langle e, \nabla \mathbf{u}(x) e\rangle - \frac{1}{\varepsilon^2}\langle \nabla \mathbf{u}(x), e\rangle^2 + \mathcal{O}\big(\frac{h}{\varepsilon^2} + \frac{h^r}{\varepsilon^{2+r}}\big), \\ \overline{\boldsymbol{\delta}}_h^{e,\varepsilon}\mathbf{u}(x) &= \frac{1}{\varepsilon}\langle \nabla \mathbf{u}(x), e\rangle + \mathcal{O}\big(\frac{h^r}{\varepsilon^{1+r}}\big). \end{aligned} \tag{42}$$

where $r = 2$ if $[x-he, x+he] \subset \overline{\Omega}$, and $r = 1$ otherwise.

Proof. The operators $\boldsymbol{\Delta}_h^{e,\varepsilon}$ and $\overline{\boldsymbol{\delta}}_h^{e,\varepsilon}$ can be expressed in terms of the corresponding upwind finite difference operators $\boldsymbol{\delta}_h^{\pm e,\varepsilon}$, similarly to their original counterparts (37). The announced result follows by inserting the Taylor expansion obtained in Lemma 3.11. In the case where $[x-he, x+he] \subset \overline{\Omega}$, the expansion of $\boldsymbol{\Delta}_h^{e,\varepsilon} = \frac{1}{h}(\boldsymbol{\delta}_h^{e,\varepsilon} + \boldsymbol{\delta}_h^{-e,\varepsilon})$ benefits form the cancellation of the term $\langle \nabla \mathbf{u}(x), e\rangle^3$ in (41) which is anti-symmetric w.r.t. e, and likewise the expansion of $\overline{\boldsymbol{\delta}}_h^{e,\varepsilon} = \frac{1}{2}(\boldsymbol{\delta}_h^{e,\varepsilon} - \boldsymbol{\delta}_h^{-e,\varepsilon})$ benefits from the cancellation of the terms $\langle \nabla \mathbf{u}, e\rangle^2$ and $\langle e, \nabla^2 \mathbf{u}\, e\rangle$ in (41) which are symmetric w.r.t. e. □

Consistently with the continuous case (21), we denote by $\mathcal{S}_h^\varepsilon$ the exponential transformation of the finite differences scheme $\mathcal{L}_h^\varepsilon$ defined by (32). In other words, following Proposition 3.10

$$\mathcal{S}_h^\varepsilon \mathbf{u} := -e^{\frac{\mathbf{u}}{\varepsilon}} \mathcal{L}_h^\varepsilon e^{-\frac{\mathbf{u}}{\varepsilon}} \tag{43}$$

on Ω_h, with boundary condition $\mathbf{u} = g$ on $\partial\Omega$.

Proposition 3.13 (Consistency with the regularized eikonal equation)**.** *For any* $\mathbf{u} \in C^3(\overline{\Omega})$*, any* $0 < h \le \varepsilon \le 1$*, and any* $x \in \Omega_h$ *one has*

$$\mathcal{S}_h^\varepsilon \mathbf{u}(x) = \mathcal{S}^\varepsilon \mathbf{u}(x) + \mathcal{O}(h + h^r/\varepsilon^r), \quad \text{where} \quad \mathcal{S}^\varepsilon \mathbf{u} := |\nabla \mathbf{u}|^2_{A_b} + 2\langle b, \nabla \mathbf{u}\rangle - 1 - \varepsilon \operatorname{Tr}(A_b \nabla^2 \mathbf{u}), \tag{44}$$

and where $r = 2$ *if* $[x - he_i, x + he_i] \subset \overline{\Omega}$ *for all* $1 \le i \le I$*, and* $r = 1$ *otherwise.*

Proof. Denoting $\mu_i := \rho_i \langle A_b^{-1} b, e_i\rangle$ we obtain as announced,

$$\begin{aligned}
\mathcal{S}_h^\varepsilon \mathbf{u}(x) &= 1 + 2\varepsilon \sum_{1\le i\le I} \mu_i \overline{\boldsymbol{\delta}}_h^{e_i} \mathbf{u}(x) - \varepsilon^2 \sum_{1\le i\le I} \rho_i \boldsymbol{\Delta}_h^{e_i} \mathbf{u}(x) \\
&\approx 1 + 2 \sum_{1\le i\le I} \mu_i \langle e_i, \nabla \mathbf{u}(x)\rangle + \sum_{1\le i\le I} \rho_i \langle e_i, \nabla \mathbf{u}(x)\rangle^2 - \varepsilon \sum_{1\le i\le I} \rho_i \langle e_i \nabla^2 \mathbf{u}(x) e_i\rangle \\
&= 1 + 2\Big\langle \sum_{1\le i\le I} \mu_i e_i,\, \nabla \mathbf{u}(x)\Big\rangle + \operatorname{Tr}\Big(\big(\nabla \mathbf{u}(x) \nabla \mathbf{u}(x)^\top - \varepsilon \nabla^2 \mathbf{u}(x)\big) \sum_{1\le i\le I} \rho_i e_i e_i^\top\Big) \\
&= 1 + 2\langle b, \nabla \mathbf{u}(x)\rangle + |\nabla \mathbf{u}(x)|^2_{A_b(x)} - \varepsilon \operatorname{Tr}(A_b(x) \nabla^2 \mathbf{u}(x)),
\end{aligned}$$

where $\approx$ denotes equality up to an $\mathcal{O}(h + h^r/\varepsilon^r)$ error. □

We obtain a consistency order of $2/3$ in the domain interior, and $1/2$ close to the boundary, by choosing ε as an optimal power of h.

Corollary 3.14 (Consistency with the eikonal equation)**.** *For any* $u \in C^3(\overline{\Omega})$*, any* $0 < h \le \varepsilon \le 1$*, and any* $x \in \Omega_h$ *one has*

$$\mathcal{S}_h^{h^\alpha} \mathbf{u}(x) = \mathcal{S}\mathbf{u}(x) + \mathcal{O}(h^\alpha), \qquad \text{where} \quad \mathcal{S}\mathbf{u} := |\nabla \mathbf{u}|^2_{A_b} + 2\langle b, \nabla \mathbf{u}\rangle - 1,$$

and where $\alpha = 2/3$ *if* $[x - he_i, x + he_i] \subset \overline{\Omega}$ *for all* $1 \le i \le I$*, and* $\alpha = 1/2$ *otherwise.*

Proof. One has $\mathcal{S}^\varepsilon \mathbf{u} = \mathcal{S}\mathbf{u} - \varepsilon \operatorname{Tr}(A_b \nabla^2 u)$, and therefore $\mathcal{S}_h^\varepsilon \mathbf{u}(x) = \mathcal{S}\mathbf{u} + \mathcal{O}(\varepsilon + h + h^r \varepsilon^{-r})$, where r is defined pointwise as in Proposition 3.13. Observing that $\alpha = r/(1+r)$, and inserting $\varepsilon = h^\alpha$ in this expression, one obtains the announced result. □

The upwind scheme $\mathcal{L}_h^{\varepsilon,+}$ obeys Proposition 3.13 but with $r = 1$ over all Ω_h, and likewise Corollary 3.14 but with $\alpha = 1/2$ over all Ω_h.

Note that the choice $\varepsilon = h^\alpha$ with $\alpha = \frac{r}{1+r}$, considered in Corollary 3.14, minimizes the error term $\sigma(h, \varepsilon) := \varepsilon + h + h^r \varepsilon^{-r}$ up to a fixed multiplicative constant. Indeed $\sigma(h, h^\alpha) = \mathcal{O}(h^\alpha)$ whereas $\sigma(h, \varepsilon) = \varepsilon + h + h^r \varepsilon^{-r} \ge \alpha\varepsilon + (1-\alpha) h^r \varepsilon^{-r} \ge \varepsilon^\alpha (h^r/\varepsilon^r)^{1-\alpha} = h^\alpha$, where the concavity of the logarithm was used for the second inequality. The parameter scaling $h = c\varepsilon$, where $c > 0$ is a small but fixed positive constant, is commonly considered in applications [17] and appears to produce usable results in practice, but is not consistent asymptotically since $\sigma(h, ch) \to c^r$. In the simplified setting where $d = 1$, $A = 1$ and $b = 0$, one easily checks that $\mathcal{S}_h^\varepsilon$ admits the solution $\mathbf{u}(x) = \lambda x$ (with suitable boundary conditions) where the slope λ obeys

$$e^{c\lambda} + e^{-c\lambda} = 2 + c^2 \qquad \text{thus } |\lambda| = 1 - c^2/24 + \mathcal{O}(c^4), \tag{45}$$

where $c = h/\varepsilon$. The correct slope $|\lambda| = 1$ is thus only obtained as $c = h/\varepsilon \to 0$.

3.3 Convergence

We establish the convergence of the logarithmically transformed solution to the numerical scheme $\mathcal{L}_h^\varepsilon$, towards the solution of Randers eikonal equation as $\varepsilon \to 0$ and $h/\varepsilon \to 0$, see Theorem 3.18 which was announced in the introduction. The proof follows the lines of [6, Theorem 2.1], and requires some preliminary steps establishing the stability and consistency of the proposed scheme. The arguments apply without modification to the less accurate but unconditionally stable $\mathcal{L}_h^{\varepsilon,+}$.

Note that, formally, the schemes $\mathcal{S}_h^\varepsilon$ and $\mathcal{L}_h^\varepsilon$ are defined over $\overline{\Omega}_h := \Omega_h \cup \partial\Omega$. In particular $\mathcal{S}_h^\varepsilon \mathbf{u}(x) = \mathbf{u}(x) - g(x)$ and $\mathcal{L}_h^\varepsilon u(x) = u(x) - \exp(-g(x)/\varepsilon)$ for all $x \in \partial\Omega$ and $u, \mathbf{u} : \overline{\Omega}_h \to \mathbb{R}$.

Lemma 3.15. *The scheme $\mathcal{S}_h^\varepsilon$ admits a constant sub-solution $\overline{\mathbf{u}} : \overline{\Omega}_h \to \mathbb{R}$ defined as*

$$\overline{\mathbf{u}}(x) := g_{\min}, \qquad \text{where} \quad g_{\min} := \min_{y \in \partial\Omega} g(y),$$

and for any $p \in \mathbb{R}^d$ with $|p|$ sufficiently large and $(\varepsilon, h/\varepsilon)$ small enough, a super-solution $\underline{\mathbf{u}} : \overline{\Omega}_h \to \mathbb{R}$ defined as the affine map

$$\underline{\mathbf{u}}(x) := \langle p, x\rangle + c_{\max}, \qquad \text{where} \quad c_{\max} := \max_{y \in \partial\Omega} \big(g(y) - \langle p, y\rangle\big).$$

Proof. Case of the sub-solution. One has $\mathcal{S}_h^\varepsilon \overline{\mathbf{u}}(x) = -1$ for all $x \in \Omega_h$, in view of (32) and (39). In addition $\mathcal{S}_h^\varepsilon \overline{\mathbf{u}}(x) = g_{\min} - g(x) \le 0$ for all $x \in \partial\Omega$, hence $\overline{\mathbf{u}}$ is a sub-solution of $\mathcal{S}_h^\varepsilon$.

Case of the super-solution. If $|p|$ is sufficiently large, then for all $x \in \overline{\Omega}$

$$|p|^2_{A_b(x)} + 2\langle b(x), p\rangle - 1 \ \ge c_0 > 0. \tag{46}$$

Indeed, recall that the matrix field $A_b : \overline{\Omega} \to S_d^{++}$ is pointwise positive definite (10), and continuous. Then by Proposition 3.13, $\mathcal{S}_h^\varepsilon \underline{\mathbf{u}}(x) \ge c_0 + \mathcal{O}(h + h^r/\varepsilon^r)$ for all $x \in \Omega_h$, which is non-negative for $(\varepsilon, h/\varepsilon)$ small enough. In addition $\mathcal{S}_h^\varepsilon \underline{\mathbf{u}}(x) = c_{\max} + \langle p, x\rangle - g(x) \ge 0$ for all $x \in \partial\Omega$, hence $\underline{\mathbf{u}}$ is a sub-solution of $\mathcal{S}_h^\varepsilon$. □

As a consequence, we prove in the next lemma that the scheme $\mathcal{S}_h^\varepsilon$ admit a unique solution, uniformly bounded as $(\varepsilon, h/\varepsilon) \to 0$.

Corollary 3.16 (Stability)**.** *For sufficiently small $(\varepsilon, h/\varepsilon)$, the scheme $\mathcal{L}_h^\varepsilon$ admits a unique solution u_h^ε, which is positive, and $\mathcal{S}_h^\varepsilon$ admits a unique solution $\mathbf{u}_h^\varepsilon$, which obeys $\mathbf{u}_h^\varepsilon = -\varepsilon \ln u_h^\varepsilon$ and satisfies $\overline{\mathbf{u}} \le \mathbf{u}_h^\varepsilon \le \underline{\mathbf{u}}$ on $\overline{\Omega}_h$, where $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$ are from Lemma 3.15.*

Proof. By Proposition 3.10, the maps $\underline{u}^\varepsilon := \exp(-\overline{\mathbf{u}}/\varepsilon)$ and $\overline{u}^\varepsilon := \exp(-\underline{\mathbf{u}}/\varepsilon)$, where $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$ are from Lemma 3.15, are respectively a super-solution and a sub-solution to the scheme $\mathcal{L}_h^\varepsilon$, which is elliptic by Proposition 3.9. Since that scheme is also linear, it admits a unique solution u_h^ε by Corollary 3.5, obeying $\overline{u}^\varepsilon \le u_h^\varepsilon \le \underline{u}^\varepsilon$ by Lemma 3.4. Note that Corollary 3.5 and Lemma 3.4 apply here regardless of the fact that the domain $\overline{\Omega}_h = \Omega_h \cup \partial\Omega$ is infinite, because the finite difference scheme $\mathcal{L}_h^\varepsilon$ only uses finitely many boundary values. We conclude that u_h^ε is positive since $\underline{u}^\varepsilon$ is positive, that $\mathbf{u}_h^\varepsilon := -\varepsilon \ln u_h^\varepsilon$ is the unique solution to $\mathcal{S}_h^\varepsilon$ by Proposition 3.10, and that $\overline{\mathbf{u}} \le \mathbf{u}_h^\varepsilon \le \underline{\mathbf{u}}$ on $\overline{\Omega}_h$ by monotony of the logarithm. The result follows. □

Lemma 3.17 (Consistency up to the boundary)**.** *For any $\varphi \in C^3(\overline{\Omega})$ and any $x \in \overline{\Omega}$ one has*

$$\limsup_{\substack{(\varepsilon,h/\varepsilon)\to 0, \xi \to 0 \\ y \in \overline{\Omega}_h,\, y \to x}} \mathcal{S}_h^\varepsilon[\varphi + \xi](y) \le \begin{cases} \mathcal{S}\varphi(x) & \text{if } x \in \Omega, \\ \max\{\mathcal{S}\varphi(x), \varphi(x) - g(x)\} & \text{if } x \in \partial\Omega. \end{cases}$$

$$\liminf_{\substack{(\varepsilon,h/\varepsilon)\to 0, \xi \to 0 \\ y \in \overline{\Omega}_h,\, y \to x}} \mathcal{S}_h^\varepsilon[\varphi + \xi](y) \ge \begin{cases} \mathcal{S}\varphi(x) & \text{if } x \in \Omega, \\ \min\{\mathcal{S}\varphi(x), \varphi(x) - g(x)\} & \text{if } x \in \partial\Omega. \end{cases}$$

Proof. For any $h > 0$, $x \in \Omega_h$, and $\xi \in \mathbb{R}$, one has by Proposition 3.13

$$\mathcal{S}_h^\varepsilon[\varphi + \xi](x) = \mathcal{S}_h^\varepsilon \varphi(x) = \mathcal{S}\varphi(x) + \mathcal{O}(\varepsilon + h + (h/\varepsilon)^r),$$

where $r \in \{1, 2\}$. In particular $r \geq 1$ and therefore $\varepsilon + (h/\varepsilon)^r \to 0$ as $h \to 0$. The announced result follows from this observation, and from the uniform continuity of the mapping $x \in \overline{\Omega} \mapsto \mathcal{S}\varphi(x) := |\nabla\varphi(x)|^2_{A_b(x)} + 2\langle b, \nabla\varphi(x)\rangle - 1$. □

Theorem 3.18 (Convergence)**.** *As $\varepsilon \to 0$ and $h/\varepsilon \to 0$ the quantity $\mathbf{u}_h^\varepsilon := -\varepsilon \ln u_h^\varepsilon$, where $\mathcal{L}_h^\varepsilon u_h^\varepsilon = 0$, converges uniformly on compact subsets of Ω to the viscosity solution $\mathbf{u}$ of (15).*

Proof. Define for all $x \in \overline{\Omega}$

$$\overline{\mathbf{v}}(x) := \limsup_{(\varepsilon, h/\varepsilon) \to 0,\, y \to x} \mathbf{u}_h^\varepsilon(x) \qquad \Big(= \sup\big\{ \limsup_{n \to \infty} u_{h_n}^{\varepsilon_n}(y_n);\ (\varepsilon_n, h_n/\varepsilon_n) \to 0,\ y_n \to x,\ y_n \in \overline{\Omega}_{h_n} \big\} \Big),$$

and likewise $\underline{\mathbf{v}}(x) := \liminf \mathbf{u}_h(x)$ as $(\varepsilon, h/\varepsilon) \to 0$ and $y \to x$. By Corollary 3.16, $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$ are well defined and bounded : $\overline{\mathbf{u}} \leq \underline{\mathbf{v}} \leq \overline{\mathbf{v}} \leq \underline{\mathbf{u}}$ on $\overline{\Omega}$ where $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$ are from Lemma 3.15. By Lemma 3.17 and following the proof of [6, Theorem 2.1], $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$ are respectively a sub-solution and a super-solution to the operator $\mathcal{S}$, or equivalently to (15).

By the continuous comparison principle Theorem A.8, one has $\overline{\mathbf{v}} \leq \mathbf{u}_* \leq \mathbf{u}^* \leq \underline{\mathbf{v}}$ on Ω, where $\mathbf{u}_*(x) := \liminf_{y \to x} \mathbf{u}(y)$ and $\mathbf{u}^*(x) := \limsup_{y \to x} \mathbf{u}(y)$ are the lower and upper semi-continuous envelopes of the solution $\mathbf{u}$ of (15). By definition $\overline{\mathbf{v}} \geq \underline{\mathbf{v}}$ on $\overline{\Omega}$, thus $\overline{\mathbf{v}} = \mathbf{u} = \underline{\mathbf{v}}$ on Ω, and the locally uniform convergence follows from the definitions of $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$. □

4 Randers distance from a point

In this section, we adapt the numerical scheme presented in §3 so as to compute Randers distance from a point source, instead of the distance to the boundary. Point sources appear to be the most common setting in applications [17, 34, 33]. However the convergence of the numerical method in this case did not appear to be backed by theory, not least because the corresponding PDE is ill posed, see Remark 4.4. To our knowledge, the convergence results of this section Theorems 4.1 and 4.2 are also new for isotropic and Riemannian metrics, which are special cases of Randers metrics of the form $\mathcal{F}_x(v) = c(x)|v|$ and $\mathcal{F}_x(v) = |v|_{M(x)}$, where $c : \overline{\Omega} \to \mathbb{R}_{++}$ and $M : \overline{\Omega} \to S_d^{++}$, and thus validate previous numerical practice.

We assume that the domain Ω is connected, and contains the origin which w.l.o.g. is the point source of interest, in addition to the previously assumed boundedness and $W^{3,\infty}$ boundary. For all $\varepsilon > 0$, $h > 0$, and $u : \overline{\Omega}_h \to \mathbb{R}$ we let

$$\tilde{\mathcal{L}}_h^\varepsilon u(x) = \begin{cases} \mathcal{L}_h^\varepsilon u(x) & \text{if } x \in \Omega_h \setminus \{0\}, \\ u(x) - 1 & \text{if } x = 0, \\ u(x) & \text{if } x \in \partial\Omega. \end{cases} \tag{47}$$

The main result of this section, Theorem 4.1 below, justifies the use of the Poisson method, i.e. solving the linear scheme $\tilde{\mathcal{L}}_h^\varepsilon$, to approximate Randers geodesic distance from the origin.

Theorem 4.1. *The solution to $\tilde{\mathcal{L}}_h^\varepsilon u_h^\varepsilon = 0$ obeys, locally uniformly in $\Omega \ni x$*

$$-\varepsilon \ln u_h^\varepsilon(x) \to d_{\mathcal{F}}(0, x), \qquad \text{as } (\varepsilon, h/\varepsilon, \varepsilon \ln h) \to 0.$$

Note that $\tilde{\mathcal{L}}_h^\varepsilon$ is a discrete degenerate elliptic operator when h/ε is sufficiently small, see Proposition 3.9, hence it does admit a unique solution by Corollary 3.5. Under the same conditions, the matrix of $\mathcal{L}_h^\varepsilon$ is invertible.

Theorem 4.2. *Denote by $L_h^\varepsilon \in \mathbb{R}^{\Omega_h \times \Omega_h}$ the matrix of the linear operator $\mathcal{L}_h^\varepsilon$ on Ω_h, with null boundary conditions on $\partial\Omega$. Then locally uniformly on $\Omega \times \Omega \ni (x, y)$ one has*

$$-\varepsilon \ln[(L_h^\varepsilon)^{-1}_{xy}] \to d_{\mathcal{F}}(x, y), \qquad \text{as } (\varepsilon, h/\varepsilon, \varepsilon \ln h) \to 0.$$

As evidenced by the constraint $\varepsilon \ln h \to 0$, Theorems 4.1 and 4.2 have no immediate continuous counterparts, see also Remark 4.4. Contrast this with the smooth boundary case, where Theorem 2.12 corresponds to Theorem 3.18 with $h = 0$. The proofs of are presented in the rest of this section. In the case of Theorem 4.1, it consists in building sub-solutions and a super-solutions to the operator $\tilde{\mathcal{L}}_h^\varepsilon$, on disk or ring domains around the origin depending on the problem scales h, ε and r, where the radius $r > 0$ is fixed but small, see §4.1 to 4.3. Sub-solutions (resp. super-solutions) over these sub-domains are glued together using the following lemma, which immediately follows from the DDE property Definition 3.3.

Lemma 4.3. *Let F be a DDE scheme on a finite set X, let $x \in X$, and let $u, v : X \to \mathbb{R}$. If $Fu(x) \le 0$ and either ($u(x) \ge v(x)$ or $Fv(x) \le 0$), then $F[\max\{u, v\}](x) \le 0$. Likewise if $Fu(x) \ge 0$ and either ($u(x) \le v(x)$ or $Fv(x) \ge 0$), then $F[\min\{u, v\}](x) \ge 0$.*

Remark 4.4 (Continuous setting)**.** *The numerical scheme* (47) *does not discretize a well posed PDE. Indeed, Dirichlet boundary conditions cannot be enforced at isolated points of elliptic PDEs in dimension $d \ge 2$. The most closely related well posed PDE is*

$$\mathcal{L}^\varepsilon u(x) = \delta_0(x) \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega,$$

where δ_0 denotes the Dirac mass at the origin. This equation admits a solution [10, Theorem 4] in the Sobolev space $W^{1,s}(\Omega)$ where $s < d/(d-1)$, in dimension $d \in \{2, 3\}$. The solution is unbounded near 0. We do not further discuss this approach, which belongs to a framework distinct from the setting of viscosity solutions considered in this paper.

Remark 4.5 (Heat method)**.** *In the Riemannian case ($\omega = 0$) an alternative approach to geodesic distance computation from a point source relies on the short time asymptotics of the heat kernel*

$$-4t \ln u(t, x) = \mathrm{dist}_{\mathcal{F}}(x_*, x)^2 + o(1), \qquad \text{where } \partial_t u = \mathrm{div}(D\nabla u), \tag{48}$$

and $u(0, \cdot) = \delta_{x_*}$ is the Dirac mass at the source point [32]. Numerically, the heat equation is solved over a short time interval, using a series of implicit time steps, each of which is equivalent to a Poisson equation [17]. To the extent of our knowledge, solving a single Poisson equation is the preferred over the heat method in applications, since it is computationally less expensive, and less susceptible to raise floating point underflow errors, in addition to being more general in view of the extension Randers metrics presented in this paper. An advantage of the heat equation is however that it allows efficient implementations of optimal transport with quadratic cost [31] in the spirit of §5.

A natural generalization of (48, right) to Finsler manifolds [27] is

$$\partial_t u(t, x) = \mathrm{div}(\partial_v H(x, \nabla u(t, x))), \qquad \text{where } H(x, v) = \tfrac{1}{2} F_x^*(v)^2, \tag{49}$$

with again $u(0, \cdot) = \delta_{x_*}$. This PDE can be reformulated as a gradient flow, in two different manners [27]. In this setting and under suitable assumptions, the heat kernel asymptotics (48, left) extend to Finsler manifolds, see [27, Example 5.5]. However, discretizing the non-linear and time dependent PDE (49) *is non-trivial, and also defeats the purpose of this paper which is to consider* linear *schemes for Randers distance computation. (If non-linear PDEs are considered, then one may as well solve Randers eikonal PDE* (5) *directly, see [24, 25].)*

Notations. The Euclidean ball, its boundary the sphere, and its intersection with the grid, defined for each center $x \in \mathbb{R}^d$, radius $r > 0$ and grid scale $h > 0$, are denoted

$$\mathbb{B}(x,r) := \{y \in \mathbb{R}^d; |y-x| < r\}, \qquad \mathbb{S}(x,r) := \partial\mathbb{B}(x,r), \qquad \mathbb{B}_h(x,r) := \mathbb{B}(x,r) \cap h\mathbb{Z}^d,$$

with the convention $\mathbb{B}(r) := \mathbb{B}(0,r)$, $\mathbb{S}(r) := \mathbb{S}(0,r)$, $\mathbb{B}_h(r) := \mathbb{B}_h(0,r)$. We introduce constants $0 < c_{\mathcal{F}} \le C_{\mathcal{F}}$ and $R_{\mathcal{F}}$, which exist by Lemma 2.4, such that for all $x, y \in \overline{\Omega}$

$$c_{\mathcal{F}}|x-y| \le \mathrm{dist}_{\mathcal{F}}(x,y) \le C_{\mathcal{F}}|x-y|, \qquad \mathrm{dist}_{\mathcal{F}}(x,y) \le R_{\mathcal{F}}. \tag{50}$$

Recall that the numerical scheme $\mathcal{L}_h^\varepsilon$ is defined in terms of a Lipschitz symmetric matrix field A and vector field b which are the parameters of the dual Randers metric. Selling's decomposition of $A_b := A - bb^\top$, see (9), which is uniformly positive definite, is denoted

$$A_b(x) = \sum_{1 \le i \le I} \rho_i(x) e_i e_i^\top, \qquad \text{where } |e_i| \le R_{\mathcal{S}},\ 1 \le i \le I, \tag{51}$$

where the bound $R_{\mathcal{S}}$ on the offsets exists in view of Theorem 3.8, and I is a suitable integer. The shorthand "$C = C(M_{\mathcal{F}})$" means that a constant C, appearing in an estimate, can be expressed or bounded in terms of the following the problem parameters

$$M_{\mathcal{F}} := \max\{c_{\mathcal{F}}^{-1}, C_{\mathcal{F}}, R_{\mathcal{F}}, R_{\mathcal{S}}, \|A\|_\infty,\ \|b\|_\infty,\ \|A_b^{-1}\|_\infty,\ \mathrm{Lip}(A_b)\},$$

where $\|A\|_\infty := \sup\{\|A(x)\|;\ x \in \overline{\Omega}\}$, and $\mathrm{Lip}(A_b)$ is the Lipschitz regularity constant of the matrix field A_b.

4.1 Viscosity regime

We construct a solution to the scheme (47) far enough from the point source singularity, at points $x \in \overline{\Omega}_h$ such that $|x| \ge r$, where r is independent of ε and h, by using the results developed §3. For that purpose, a radius $r > 0$ is fixed in the rest of this section, unless otherwise specified, and such that $\mathbb{B}(6r) \subset \Omega$. The *erosion* with radius r of the domain Ω, and its intersection with the grid, are defined as

$$\mathrm{int}(\Omega, r) := \{x \in \Omega;\ \mathbb{B}(x,r) \subset \Omega\}, \qquad \mathrm{int}_h(\Omega, r) := \mathrm{int}(\Omega, r) \cap h\mathbb{Z}^d.$$

Lemma 4.6. *For each $\varepsilon > 0$ and $h > 0$ let u_h^ε be the solution to*

$$\mathcal{L}_h^\varepsilon u = 0 \text{ on } \Omega_h \setminus \overline{\mathbb{B}}(r), \qquad u = 1 \text{ on } \mathbb{S}(r) \qquad u = \exp(-R_{\mathcal{F}}/\varepsilon) \text{ on } \partial\Omega. \tag{52}$$

Then for $(\varepsilon, h/\varepsilon)$ sufficiently small, and denoting $\mathbf{u}_h^\varepsilon := -\varepsilon \ln u_h^\varepsilon$, one has with $C = C(M_{\mathcal{F}})$

$$|\mathbf{u}_h^\varepsilon(x) - \mathrm{dist}_{\mathcal{F}}(0,x)| \le Cr \qquad \text{on } \mathrm{int}_h(\Omega, r) \setminus \mathbb{B}(2r). \tag{53}$$

Proof. Applying Theorem 3.18 to the domain $\Omega \setminus \mathbb{B}(r)$ we obtain that $\mathbf{u}_h^\varepsilon$ converges uniformly over the relatively compact subset $\mathrm{int}(\Omega, r) \setminus \mathbb{B}(2r)$ as $(\varepsilon, h/\varepsilon) \to 0$, to the limit

$$\mathbf{u}(x) = \min\big\{ \min_{p \in \mathbb{S}(r)} d_{\mathcal{F}}(p,x),\ R_{\mathcal{F}} + \min_{q \in \partial\Omega} d_{\mathcal{F}}(q,x)\big\} = \min_{p \in \mathbb{S}(r)} d_{\mathcal{F}}(p,x),$$

where the second equality follows from (50, right). Observing that $|d_{\mathcal{F}}(p,x) - d_{\mathcal{F}}(0,x)| \le C_{\mathcal{F}}|p| \le C_{\mathcal{F}}\, r$ for all $p \in \mathbb{S}(r)$, see (50, left), we conclude the proof. □

Corollary 4.7. *For $(\varepsilon, h/\varepsilon)$ sufficiently small, there exists $\underline{u}_h^\varepsilon : \overline{\Omega}_h \to \mathbb{R}$ such that $\tilde{\mathcal{L}}_h^\varepsilon \underline{u}_h^\varepsilon \geq 0$ and $\overline{\mathbf{u}}_h^\varepsilon(x) := -\varepsilon \ln \underline{u}_h^\varepsilon(x) \geq \mathrm{dist}_{\mathcal{F}}(0,x) - Cr$ on $\mathrm{int}_h(\Omega, r)$, where $C = C(M_{\mathcal{F}})$.*

Proof. From Lemma 4.6 introduce $\mathbf{u}_h^\varepsilon = -\varepsilon \ln u_h^\varepsilon$ obeying (53) for sufficiently small $(\varepsilon, h/\varepsilon)$, with constant $C_0 = C_0(M_{\mathcal{F}})$. Then let

$$\underline{u}_h^\varepsilon(x) := \begin{cases} 1 & x \in \mathbb{B}_h(2r), \\ \min\{1, u_h^\varepsilon(x) \exp(C_1 r/\varepsilon)\} & x \notin \overline{\Omega}_h \setminus \mathbb{B}_h(2r), \end{cases} \tag{54}$$

where $C_1 = C_0 + 3C_{\mathcal{F}}$. By construction one has $\underline{u}_h^\varepsilon(0) = 1$, and $\underline{u}_h^\varepsilon(x) \geq 0$ on $\partial\Omega$, so that $\tilde{\mathcal{L}}_h^\varepsilon u_h^\varepsilon \geq 0$ at these boundary points. By choice of the constant C_1 and in view of (53), one has $1 \leq u_h^\varepsilon(x) \exp(Cr/\varepsilon)$ on $\mathbb{B}_h(3r) \setminus \mathbb{B}_h(2r)$. Note that provided $h \leq r/R_{\mathcal{S}}$ the expression of $\mathcal{L}_h^\varepsilon \underline{u}_h^\varepsilon(x)$ at any $x \in \overline{\Omega}_h \setminus \mathbb{B}(3r)$ only involves values of $\underline{u}_h^\varepsilon$ in $\overline{\Omega}_h \setminus \mathbb{B}(2r)$. By Lemma 4.3, and since the constant 1 is a super-solution to $\mathcal{L}_h^\varepsilon$, we obtain that $\mathcal{L}_h^\varepsilon \underline{u}_h^\varepsilon \geq 0$, as announced. Finally, one has $\overline{\mathbf{u}}_h^\varepsilon(x) \geq \mathbf{u}_h^\varepsilon(x) - C_1 r \geq \mathrm{dist}_{\mathcal{F}}(0,x) - (C_0 + C_1)r$ on $\mathrm{int}_h(\Omega, r) \setminus \mathbb{B}_h(2r)$, and $\overline{\mathbf{u}}_h^\varepsilon(x) \geq 0 \geq \mathrm{dist}_{\mathcal{F}}(0,x) - 2C_{\mathcal{F}} r$ on $\mathbb{B}_h(2r)$, which concludes the proof. □

4.2 Taylor expansion regime

We construct explicit sub-solutions to the scheme (47), at points $h \lesssim |x| \lesssim \varepsilon$ and $\varepsilon \lesssim |x| \lesssim r$, which are radial functions with respectively a power and exponential profile. For that purpose, we need to estimate the derivatives of such functions.

Lemma 4.8. *Let $f \in C^2(\mathbb{R}_{++}, \mathbb{R})$, let $\mu \in \mathbb{R}$, and let $u(x) := \exp(-\mu f(|x|))$ for all $x \in \mathbb{R}^d \setminus \{0\}$. Then one has with $n(x) := x/|x|$, omitting the arguments of f, f', f'', f''' and n*

$$\frac{\nabla u(x)}{u(x)} = -\mu f' n, \qquad \frac{\nabla^2 u(x)}{u(x)} = \mu^2 f'^2 n n^\top + \mathcal{O}\big(\mu |f''| + \frac{\mu |f'|}{|x|}\big),$$

$$\frac{\nabla^3 u(x)}{u(x)} = \mathcal{O}\big(\mu^3 |f'|^3 + \mu^2 |f'||f''| + \frac{\mu^2 |f'|^2}{|x|} + \frac{\mu |f''|}{|x|} + \mu |f'''| + \frac{\mu |f'|}{|x|^2}\big),$$

with absolute constants underlying the $\mathcal{O}$ notation.

Proof. The expression of $\nabla u(x)$ follows from the standard rules for the differentiation of an exponential function $\nabla(\exp \circ g) = (\exp \circ g) \nabla g$, and of a radial function $\nabla g(|x|) = g'(|x|) n(x)$. The full expression of $u^{-1} \nabla^2 u(x) = \mu^2 f'^2 n n^\top - \mu f'' n n^\top - \mu f' (\mathrm{Id} - n n^\top)/|x|$ can be obtained using the Liebniz rule for the differentiation of a product, and recalling that the Jacobian matrix of $n(x)$ is $(\mathrm{Id} - n n^\top)/|x|$. The expression of $\nabla^3 u$ follows likewise. □

Corollary 4.9. *Define $u(x) := \exp(-\lambda |x|/\varepsilon)$ where $\lambda \geq 1$, $\varepsilon > 0$. If $x \in \Omega_h$, $\varepsilon \leq |x| \leq 5r$ and $2R_{\mathcal{S}} h \leq \varepsilon$ then*

$$u(x)^{-1} \mathcal{L}_h^\varepsilon u(x) \leq 1 - \lambda^2 |n(x)|^2_{A_b(x)} + C_0(\lambda + \lambda^3 h/\varepsilon). \tag{55}$$

In particular, $\mathcal{L}_h^\varepsilon u(x) \leq 0$ if $\lambda \geq C_1$ and $\lambda h/\varepsilon \leq c_2$, where $C_0, C_1, c_2 > 0$ only depend on $M_{\mathcal{F}}$.

Proof. Applying Lemma 4.8 to the identity function $f : r \in \mathbb{R}_{++} \mapsto r$, and parameter $\mu := \lambda/\varepsilon$ (note that $\mu \geq 1/\varepsilon$), we obtain whenever $|x| \geq \varepsilon/2$

$$\frac{\nabla u(x)}{u(x)} = \mathcal{O}(\mu), \qquad \frac{\nabla^2 u(x)}{u(x)} = \mu^2 n n^\top + \mathcal{O}(\frac{\mu}{\varepsilon}), \qquad \frac{\nabla^3 u(x)}{u(x)} = \mathcal{O}(\mu^3).$$

If $|x| \geq \varepsilon$ and $|e| \leq R_{\mathcal{S}}$, then any $y \in [x - he, x + he]$ obeys $|y| \geq \varepsilon/2$. Therefore

$$\frac{\overline{\delta}_h^e u(x)}{u(x)} = \mathcal{O}(\mu R_{\mathcal{S}} + h\mu^2 R_{\mathcal{S}}^2), \qquad \frac{\Delta_h^e u(x)}{u(x)} = \mu^2 \langle n, e\rangle^2 + \mathcal{O}(\frac{\mu}{\varepsilon} R_{\mathcal{S}}^2 + h\mu^3 R_{\mathcal{S}}^3),$$

with again absolute constants underlying the $\mathcal{O}$ notation. Inserting these estimates in the scheme expression we obtain omitting the argument of ρ_i, $A_b^{-1} b$ and n

$$\frac{\mathcal{L}_h^\varepsilon u(x)}{u(x)} \leq 1 + 2\varepsilon C \sum_{1 \leq i \leq I} \rho_i |\langle A_b^{-1} b, e_i\rangle| (\mu + h\mu^2) + \varepsilon^2 \sum_{1 \leq i \leq I} \rho_i [-\mu^2 \langle n, e_i\rangle^2 + C(\frac{\mu^2}{\varepsilon} + h\mu^3)],$$

where C depends only on $R_{\mathcal{S}}$. This establishes (55) observing that $\sum_{i=1}^I \rho_i \langle n, e_i\rangle^2 = \mathrm{Tr}(A_b n n^\top) = |n|_{A_b}^2$, and that $\sum_{i=1}^I \rho_i \leq \|\mathrm{Tr}(A_b)\|_\infty$. Since A_b is uniformly positive definite over $\overline{\Omega}$ and n is a unit vector, one has $|n|_{A_b}^2 \geq c_0 = c_0(M_{\mathcal{F}}) > 0$, and the result follows with $C_1 = \max\{4C_0/c_0, \sqrt{2/c_0}\}$ and $c_2 = c_0/4C_0$. □

Corollary 4.10. *Define $u(x) := |x|^{-\mu}$, where $\mu \geq 1$. If $x \in \Omega_h$ and $2R_{\mathcal{S}} h \leq |x| \leq 4\varepsilon$ then*

$$\frac{\mathcal{L}_h^\varepsilon u(x)}{u(x)} \leq 1 - \frac{\varepsilon^2 \mu^2}{|x|^2} |n(x)|_{A_b(x)}^2 + C_0(\frac{\varepsilon^2 \mu}{|x|^2} + \frac{h\varepsilon^2 \mu^3}{|x|^3}). \tag{56}$$

In particular $\mathcal{L}_h^\varepsilon u(x) \leq 0$ if $\mu \geq C_1$ and $\mu h/\varepsilon \leq c_2$, where $C_0, C_1, c_2 > 0$ only depend on $M_{\mathcal{F}}$.

Proof. We apply Lemma 4.8 to the logarithm function $f = \ln$, obtaining

$$\frac{\nabla u(y)}{u(y)} = \mathcal{O}(\frac{\mu}{|y|}), \qquad \frac{\nabla^2 u(y)}{u(y)} = \frac{\mu^2 n n^\top}{|y|^2} + \mathcal{O}(\frac{\mu}{|y|^2}), \qquad \frac{\nabla^3 u(y)}{u(y)} = \mathcal{O}(\frac{\mu^3}{|y|^3}).$$

If $|x| \geq 2R_{\mathcal{S}} h$ and $|e| \leq R_{\mathcal{S}}$, then any $y \in [x - he, x + he]$ obeys $|y| \geq |x|/2$. Therefore

$$\frac{\overline{\delta}_h^e u(x)}{u(x)} = \mathcal{O}(\frac{\mu}{|x|} + \frac{h\mu^2}{|x|^2}), \qquad \frac{\Delta_h^e u(x)}{u(x)} = \frac{\mu^2 \langle n, e\rangle^2}{|x|^2} + \mathcal{O}(\frac{\mu}{|x|^2} + \frac{h\mu^3}{|x|^3}).$$

Inserting these estimates in the scheme expression (32), we conclude similarly to Corollary 4.9. □

4.3 Finite neighborhood regime

We produce a sub-solution to the scheme $\tilde{\mathcal{L}}_h^\varepsilon$ which is useful in the immediate neighborhood of the origin, where $|x| \lesssim h$. The construction is not based on the approach of viscosity solutions, or on a Taylor expansion, but on the discrete structure of the scheme. For that purpose, we establish additional properties of its coefficients (51), suitably normalized: the first d offsets form a basis of $\mathbb{Z}^d$, and the corresponding weights are bounded below in a neighborhood of the source point. This implies that the stencils of our numerical scheme are locally connected, and allows to construct a subsolution in Corollary 4.13. The proof is based on the *spanning property* of Selling's decomposition, see Proposition B.8, which is used here for the first time in the context of PDE numerical analysis.

Proposition 4.11. *Up to reordering the terms $(\rho_i, e_i)_{i=1}^I$ of Selling's decomposition (51) of the matrix field A_b, and grouping duplicate and opposite offsets $(e_i)_{i=1}^I$, one has for all $|x| \leq r_{\mathcal{S}}$*

$$\min\{\rho_1(x), \cdots, \rho_d(x)\} \geq \rho_{\mathcal{S}}, \qquad \det(e_1, \cdots, e_d) = 1, \tag{57}$$

where the constants $\rho_{\mathcal{S}} > 0$ and $r_{\mathcal{S}} > 0$ only depend on $M_{\mathcal{F}}$.

Proof. Up to grouping duplicates and opposites, we may assume that the vectors $\pm e_1, \cdots, \pm e_I$ are pairwise distinct. Thus by Proposition B.5 one has for all $x, y \in \overline{\Omega}$ and all $1 \leq i \leq I$

$$|\rho_i(x) - \rho_i(y)| \leq C|x - y|, \tag{58}$$

where $C = C(M_{\mathcal{F}})$. Then by Proposition B.8, and up to reordering $(\rho_i, e_i)_{i=1}^I$, one has $\det(e_1, \cdots, e_d) = 1$ and $\rho_i(0) \geq 2\rho_{\mathcal{S}}$ for all $1 \leq i \leq d$, where $\rho_{\mathcal{S}}$ only depends on $\|A_b(0)\|$ and $\|A_b(0)^{-1}\|$. The announced result follows, by choosing $r_{\mathcal{S}} := \rho_{\mathcal{S}}/C$. □

In the rest of this section, we assume that $(\rho_i, e_i)_{i=1}^I$ are ordered in such way that (57) holds. We also denote $\rho_{-i} := \rho_i$ and $e_{-i} := -e_i$ for all $1 \leq i \leq I$. Hence for any $x \in \Omega_h$ such that $\mathbb{B}(x, R_{\mathcal{S}}h) \subset \Omega_h$

$$\mathcal{L}_h^\varepsilon u(x) = \alpha_h^\varepsilon(x)u(x) - \sum_{1 \leq |i| \leq I} \beta_{h,i}^\varepsilon(x) u(x + he_i),$$

where the coefficients are

$$\alpha_h^\varepsilon(x) := 1 + 2\frac{\varepsilon^2}{h^2} \sum_{1 \leq i \leq I} \rho_i(x), \qquad \beta_{h,i}^\varepsilon(x) := \rho_i(x)\Big(\frac{\varepsilon^2}{h^2} - \frac{\varepsilon}{h}\langle A_b(x)^{-1}b(x), e_i\rangle\Big). \tag{59}$$

Note that $\alpha_h^\varepsilon(x) \leq 1 + 2(\varepsilon/h)^2 \| \operatorname{Tr}(A_b)\|_\infty$, since $\sum_{i=1}^I \rho_i(x) \leq \sum_{i=1}^I \rho_i(x)|e_i|^2 = \operatorname{Tr}(A_b(x))$. We denote by $|x|_1$ the sum of the absolute values of the coefficients of a vector $x \in \mathbb{R}^d$.

Lemma 4.12. *Let $G \in \mathrm{GL}(\mathbb{Z}^d)$ be the matrix of columns $e_1, \cdots, e_d$, and let $N(x) := |G^{-1}x|_1$. Then for any $z \in \mathbb{Z}^d \setminus \{0\}$ there exists $1 \leq |i| \leq d$ such that $N(z + e_i) = N(z) - 1$. In addition $c|x| \leq N(x) \leq C|x|$ where the constants $C, c > 0$ only depend on $M_{\mathcal{F}}$.*

Proof. The matrix G has integer coefficients by construction, and $\det(G) = 1$ by (57, left) hence its inverse is the adjugate matrix $G^{-1} = \mathrm{co}(G)^\top$ which also has integer coefficients, thus $G \in \mathrm{GL}(\mathbb{Z}^d)$ as announced. Since the coefficients of G are bounded by $R_{\mathcal{S}}$, those of the adjugate matrix G^{-1} are bounded by $(d-1)!R_{\mathcal{S}}^{d-1}$, and the equivalence of N with the Euclidean norm follows.

Let $z \in \mathbb{Z}^d \setminus \{0\}$, and let $\lambda_1, \cdots, \lambda_d \in \mathbb{Z}$ be the coordinates of z in the basis $e_1, \cdots, e_d$, in other words $(\lambda_1, \cdots, \lambda_d)^\top = G^{-1}z$. Since $z \neq 0$, one at least of these coordinates is non-zero. We thus assume w.l.o.g. that $\lambda_1 > 0$, up to a change of sign and permutation of the axes. Then $N(z - e_1) = |\lambda_1 - 1| + |\lambda_2| + \cdots + |\lambda_d| = -1 + |\lambda_1| + \cdots + |\lambda_d| = N(z) - 1$, which concludes the proof. □

Corollary 4.13. *Define $u(x) := \exp(-\nu N(x)/h)$. Then $\tilde{\mathcal{L}}_h^\varepsilon u(x) \leq 0$ on $\mathbb{B}_h(r_{\mathcal{S}})$, provided $\nu \geq \nu_0 = \nu_0(M_{\mathcal{F}})$, $\mathbb{B}(x, R_{\mathcal{S}}h) \subset \Omega$, and h/ε is sufficiently small.*

Proof. Note that $\beta_{h,i}^\varepsilon(x) \geq \rho_i(x)\varepsilon^2/(2h^2) \geq 0$, for all $1 \leq i \leq I$, when $h/\varepsilon \leq c := 1/(2\|A_b^{-1}b\|_\infty R_{\mathcal{S}})$. In particular $\beta_{h,i}^\varepsilon(x) \geq \rho_{\mathcal{S}}\varepsilon^2/(2h^2)$ if $|x| \leq r_{\mathcal{S}}$ and $1 \leq |i| \leq d$. By Lemma 4.12 there exists $1 \leq |i| \leq d$ such that $N(x + he_i) = N(x) - h$, and therefore $u(x + he_i) \geq e^\nu u(x)$. Thus

$$\frac{\mathcal{L}_h^\varepsilon u(x)}{u(x)} \leq \alpha_h^\varepsilon(x) - \beta_{h,i}^\varepsilon(x)\frac{u(x + he_i)}{u(x)} \leq 1 + 2\| \operatorname{Tr}(A_b)\|_\infty \frac{\varepsilon^2}{h^2} - e^\nu \frac{\rho_{\mathcal{S}}}{2}\frac{\varepsilon^2}{h^2}. \tag{60}$$

The result follows, by assuming in addition that $h \leq \varepsilon$ and choosing ν_0 such that $e^{\nu_0} := 2(1 + 2\| \operatorname{Tr}(A_b)\|_\infty)/\rho_{\mathcal{S}}$. □

4.4 Gluing the sub-solutions

In the previous subsections, we have produced four sub-solutions to the operator $\tilde{\mathcal{L}}_h^\varepsilon$, on different subsets of the domain $\overline{\Omega}_h$ defined according to the distance to the origin, see Lemma 4.6 and Corollaries 4.9, 4.10, and 4.13. We glue here these partial sub-solutions using Lemma 4.3, to produce a global sub-solution on $\overline{\Omega}_h$ and conclude the proof of Theorem 4.1. For that purpose, we introduce four mappings $u_h^{\varepsilon,i}$ defined on adequate subdomains $\Omega_h^{\varepsilon,i} \subset \overline{\Omega}_h$, $1 \le i \le 4$, and depending on the scale parameters (ε, h) as well as constants (λ, μ, ν, ξ) specified later.

- $u_h^{\varepsilon,0}(x) := v_h^\varepsilon(x) - \exp(-R_{\mathcal{F}}/\varepsilon)$, and $\Omega_h^{\varepsilon,0} := \overline{\Omega}_h \setminus \mathbb{B}_h(2r)$, where v_h^ε solves (52).
- $u_h^{\varepsilon,1}(x) = \exp(-\lambda|x|/\varepsilon)$, and $\Omega_h^{\varepsilon,1} := \mathbb{B}_h(5r) \setminus \mathbb{B}_h(\varepsilon)$.
- $u_h^{\varepsilon,2}(x) = |x|^{-\mu}$, and $\Omega_h^{\varepsilon,2} = \mathbb{B}_h(4\varepsilon) \setminus \mathbb{B}_h(2R_{\mathcal{S}}h)$.
- $u_h^{\varepsilon,3}(x) = \exp(-\nu N(x)/h)$, and $\Omega_h^{\varepsilon,3} = \mathbb{B}_h(\xi h)$, where N is from Lemma 4.12.

Proposition 4.14. *For any $(\varepsilon, h/\varepsilon)$ sufficiently small one has $\tilde{\mathcal{L}}_h^\varepsilon \overline{u}_h^\varepsilon \le 0$ on $\overline{\Omega}_h$, where*

$$\overline{u}_h^\varepsilon(x) := \max\{u_h^{\varepsilon,3}(x),\, \alpha_2 h^\mu\, u_h^{\varepsilon,2}(x),\, \alpha_1(\tfrac{h}{\varepsilon})^\mu\, u_h^{\varepsilon,1}(x),\, \alpha_0(\tfrac{h}{\varepsilon})^\mu e^{-3\lambda\frac{r}{\varepsilon}}\, u_h^{\varepsilon,0}(x)\}, \tag{61}$$

for all $x \in \overline{\Omega}_h$, and where the quantity $u_h^{\varepsilon,i}(x)$ is only considered in the maximum if $x \in \Omega_h^{\varepsilon,i}$. The constants $(\lambda, \mu, \nu, \xi, \alpha_0, \alpha_1, \alpha_2)$ only depend on $M_{\mathcal{F}}$.

Proof. By Corollaries 4.9, 4.10, and 4.13 one may choose the constants λ, μ, ν such that $\tilde{\mathcal{L}}_\varepsilon^h u_h^{\varepsilon,i} \le 0$ on $\Omega_h^{\varepsilon,i}$ for all $1 \le i \le 3$ and $(\varepsilon, h/\varepsilon)$ sufficiently small. Furthermore, this property is preserved if λ, μ or ν is increased. Also $\tilde{\mathcal{L}}_\varepsilon^h u_h^{\varepsilon,0} \le 0$ on $\Omega_h^{\varepsilon,0}$, by noting that the positive constant $\exp(-R_{\mathcal{F}}/\varepsilon)$ subtracted in its definition accounts for the null boundary conditions of $\tilde{\mathcal{L}}_h^\varepsilon$, compare (47) with (52). Since the operator $\tilde{\mathcal{L}}_h^\varepsilon$ is linear on $\overline{\Omega}_h \setminus \{0\}$, see (47), the product of a sub-solution with a positive constant remains a sub-solution (outside of the origin). Hence (61) is a maximum of 4 sub-solutions on their respective domains.

We next proceed to prove estimates of the following form: for any $x \in \Omega_h^{\varepsilon,i} \cap \Omega_h^{\varepsilon,i+1}$

$$m_h^{\varepsilon,i} u_h^{\varepsilon,i}(x) \le (\text{resp. } \ge)\, u_h^{\varepsilon,i+1}(x) \quad \text{when} \quad \mathbb{B}_h(x, R_{\mathcal{S}}h) \not\subset \Omega_h^{\varepsilon,i}\ (\text{resp. } \Omega_h^{\varepsilon,i+1}), \tag{62}$$

where $m_h^{\varepsilon,i}$ is a suitable function of the scale parameters. Thus by Lemma 4.3,

$$u_h^\varepsilon(x) := \max\{u_h^{3,\varepsilon}(x),\ m_h^{\varepsilon,2} u_h^{2,\varepsilon}(x),\ m_h^{\varepsilon,2} m_h^{\varepsilon,1} u_h^{1,\varepsilon}(x),\ m_h^{\varepsilon,2} m_h^{\varepsilon,1} m_h^{\varepsilon,0} u_h^{1,\varepsilon}(x)\}$$

is a sub-solution, which is the announced result. Indeed one has $\mathcal{L}_h^\varepsilon u_h^{\varepsilon,i}(x) \le 0$ if $\mathbb{B}_h(x, R_{\mathcal{S}}h) \subset \Omega_h^{\varepsilon,i}$, but $\mathcal{L}_h^\varepsilon u_h^{\varepsilon,i}(x)$ may not make sense if $\mathbb{B}_h(x, R_{\mathcal{S}}h) \not\subset \Omega_h^{\varepsilon,i}$ since it could involve values of $u_h^{\varepsilon,i}$ outside $\Omega_h^{\varepsilon,i}$; in that case however, (62) shows that $\overline{u}_h^\varepsilon(x)$ is not defined from $u_h^{\varepsilon,i}(x)$.

The estimates (62) follow from basic upper and lower bounds of the involved functions, and of the norms of the relevant points x. Namely

$$\begin{aligned} &u_h^{\varepsilon,0}(x) \le 1, && u_h^{\varepsilon,1}(x) \ge \exp(-3\lambda r/\varepsilon), && \text{when } 2r \le |x| \le 3r. \\ &u_h^{\varepsilon,0}(x) \ge \exp(-Cr/\varepsilon), && u_h^{\varepsilon,1}(x) \le \exp(-4\lambda r/\varepsilon), && \text{when } 4r \le |x| \le 5r. \end{aligned}$$

The upper bound on $u_h^{\varepsilon,0}$ is derived from the maximum principle, and the lower bound from Lemma 4.6, with $C = C(M_{\mathcal{F}})$ and for sufficiently small $(\varepsilon, h/\varepsilon)$. This establishes (62, $i = 0$) with $m_h^{\varepsilon,0} = \exp(-3\lambda r/\varepsilon)$, up to increasing λ so that $\lambda \ge C$. Likewise

$$\begin{aligned} &u_h^{\varepsilon,1}(x) \le \exp(-\lambda), && u_h^{\varepsilon,2}(x) \ge (2\varepsilon)^{-\mu}, && \text{when } \varepsilon \le |x| \le 2\varepsilon. \\ &u_h^{\varepsilon,1}(x) \ge \exp(-4\lambda), && u_h^{\varepsilon,2}(x) \le (3\varepsilon)^{-\mu}, && \text{when } 3\varepsilon \le |x| \le 4\varepsilon. \end{aligned}$$

This establishes (62, $i=1$) with $m_h^{\varepsilon,1} = e^{\lambda}(2\varepsilon)^{-\mu}$, up to increasing μ so that $(3/2)^{\mu} \geq e^{3\lambda}$. Lastly

$$u_h^{\varepsilon,2}(x) \leq (2R_{\mathcal{S}}h)^{-\mu}, \quad u_h^{\varepsilon,3}(x) \geq \exp(-3R_{\mathcal{S}}C_N\nu), \qquad \text{when } 2R_{\mathcal{S}}h \leq |x| \leq 3R_{\mathcal{S}}h.$$
$$u_h^{\varepsilon,2}(x) \geq (\xi R_{\mathcal{S}}h)^{-\mu}, \quad u_h^{\varepsilon,3}(x) \leq \exp(-(\xi - R_{\mathcal{S}})c_N\nu), \quad \text{when } (\xi - R_{\mathcal{S}})h \leq |x| \leq \xi R_{\mathcal{S}}h,$$

where c_N and C_N are the equivalence constants in Lemma 4.12. We define ξ by $(\xi - R_{\mathcal{S}})c_N - 3R_{\mathcal{S}}C_N = 1$. This establishes (62, $i=2$) with $m_h^{\varepsilon,2} = e^{-3R_{\mathcal{S}}C_N\mu}(2R_{\mathcal{S}}h)^{\mu}$, up to increasing ν so that $e^{\nu} \geq (\xi/(2R_{\mathcal{S}}))^{\mu}$, in view of the expression of ξ, which concludes the proof. □

Corollary 4.15. *For $(\varepsilon, h/\varepsilon)$ sufficiently small, there exists $\overline{u}_h^{\varepsilon} : \overline{\Omega}_h \to \mathbb{R}$ such that $\tilde{\mathcal{L}}_h^{\varepsilon}\overline{u}_h^{\varepsilon} \leq 0$ and $\underline{\mathbf{u}}_h^{\varepsilon}(x) := -\varepsilon \ln \overline{u}_h^{\varepsilon}(x) \leq \mathrm{dist}_{\mathcal{F}}(0,x) + C(r + \varepsilon\ln(\varepsilon/h))$ on $\mathrm{int}_h(\Omega, r)$, where $C = C(M_{\mathcal{F}})$.*

Proof. We distinguish two cases. (i) If the maximum in (61) is attained by the last term, then the announced result follows Lemma 4.6 and the expression of the multiplicative factor $\alpha_0(h/\varepsilon)^{\mu}\exp(-3\lambda r/\varepsilon)$. (ii) If the maximum in (61) is attained by one of the first three terms, then $|x| \leq 5r$ and the announced result follows from the explicit expressions of $u_h^{\varepsilon,1}, u_h^{\varepsilon,2}, u_h^{\varepsilon,3}$ as well as $\mathrm{dist}_{\mathcal{F}}(0,x) \leq 5C_{\mathcal{F}}r$. □

Proof of Theorem 4.1. For sufficiently small $(\varepsilon, h/\varepsilon)$, we obtain from the comparison principle Lemma 3.4 and with the mappings $\overline{\mathbf{u}}_h^{\varepsilon}$ and $\underline{\mathbf{u}}_h^{\varepsilon}$ of Corollaries 4.7 and 4.15 respectively that

$$\mathrm{dist}_{\mathcal{F}}(0,x) - Cr \leq \overline{\mathbf{u}}_h^{\varepsilon}(x) \leq \mathbf{u}_h^{\varepsilon}(x) \leq \underline{\mathbf{u}}_h^{\varepsilon}(x) \leq \mathrm{dist}_{\mathcal{F}}(0,x) + C(r + \varepsilon\ln(\varepsilon/h)), \tag{63}$$

on $\mathrm{int}_h(\Omega, r)$, where $C = C(M_{\mathcal{F}})$. Since the parameter $r > 0$ is arbitrary[2], except for the constraint $\mathbb{B}(6r) \subset \Omega$, we conclude as announced that $\mathbf{u}_h^{\varepsilon}(x) \to \mathrm{dist}_{\mathcal{F}}(0,x)$ locally uniformly on Ω as $(\varepsilon, h/\varepsilon, \varepsilon\ln(\varepsilon/h)) \to 0$. The result follows, noting that $\varepsilon\ln(\varepsilon/h) \leq \varepsilon|\ln h|$ when $0 < h \leq \varepsilon \leq 1$. □

4.5 Convergence on $\Omega \times \Omega$ and inverse matrix

We establish Theorem 4.2, which relates the Randers distance with the inverse matrix of our finite differences scheme. For that purpose, we use the following convention: if $U(x; x_*)$ if a bivariate discrete mapping, defined for all $(x, x_*) \in \overline{\Omega}_h \times \Omega_h$, and if F is a finite differences scheme of the form of Definition 3.3, then $FU(x; x_*) := \tilde{F}(x, U(x; x_*), [U(x; x_*) - U(y; x_*)]_{y \in X \setminus \{x\}})$. In other words, the numerical scheme sees U as a function of its first variable x only.

Lemma 4.16. *For any $(\varepsilon, h/\varepsilon)$ sufficiently small, and any $x_* \in \mathbb{B}_h(r/2)$, one has $\tilde{\mathcal{L}}_h^{\varepsilon}U_h^{\varepsilon}(x; x_*) \leq 0$ on $\Omega_h \setminus \{x_*\}$, where for all $x \in \overline{\Omega}_h$*

$$\overline{U}_h^{\varepsilon}(x; x_*) := \max\{u_h^{\varepsilon,3}(x),\ \alpha_2 h^{\mu}\, u_h^{\varepsilon,2}(x - x_*),\ \alpha_1(h/\varepsilon)^{\mu}\, u_h^{\varepsilon,1}(x - x_*),\ \alpha_0(h/\varepsilon)^{\mu}e^{-3\lambda r/\varepsilon}\, u_h^{\varepsilon,0}(x - x_*)\},$$

and where the quantity $u_h^{\varepsilon,i}(x - x_*)$ is only considered in the maximum if $x - x_* \in \Omega_h^{\varepsilon,i}$. The constants $(\lambda, \mu, \nu, \xi, \alpha_0, \alpha_1, \alpha_2)$ only depend on $M_{\mathcal{F}}$. In addition $\underline{\mathbf{U}}_h^{\varepsilon}(x; x_*) := -\varepsilon \ln \overline{U}_h^{\varepsilon}(x; x_*) \leq \mathrm{dist}_{\mathcal{F}}(0,x) + C(r + \varepsilon\ln(\varepsilon/h))$ for all $(x, x_*) \in \mathrm{int}_h(\Omega, r) \times \mathbb{B}_h(r/2)$, where $C = C(M_{\mathcal{F}})$.

Proof. The proofs of Proposition 4.14 and Corollary 4.7 adapt in a straightforward manner to a point source x_* sufficiently close to the origin, as here. □

Proposition 4.17 (Convergence in the product space)**.** *Denote by $U_h^{\varepsilon} : \overline{\Omega}_h \times \Omega_h \to \mathbb{R}$ the solution to*

$$\mathcal{L}_h^{\varepsilon}U_h^{\varepsilon}(x;\, x_*) = 0, \forall x \in \Omega_h \setminus \{x_*\}, \qquad U_h^{\varepsilon}(x_*;\, x_*) = 1 \qquad U_h^{\varepsilon}(x; x_*) = 0, \forall x \in \partial\Omega. \tag{64}$$

Then locally uniformly on $\Omega \times \Omega$ one has $-\varepsilon \ln U_h^{\varepsilon}(x; x_*) \to d_{\mathcal{F}}(x_*, x)$ as $(\varepsilon, h/\varepsilon, \varepsilon\ln h) \to 0$.

[2]Note nevertheless that (63) holds when $\varepsilon \leq \delta$ and $h/\varepsilon \leq \delta$, where δ depends on $M_{\mathcal{F}}$ and r.

Proof. First note that $x \in \overline{\Omega}_h \mapsto U(x;x_*)$, for any given $x_* \in \Omega_h$, solves a linear problem which is elliptic when h/ε is sufficiently small, hence has a unique solution, see Corollary 3.5 and Proposition 3.9.

Let $r > 0$ be such that $\mathbb{B}(6r) \subset \Omega$. Then for $(\varepsilon, h/\varepsilon)$ sufficiently small and for all $(x, x_*) \in \mathrm{int}_h(\Omega, r) \times \mathbb{B}_h(r/2)$ one has by Corollary 4.7 and Lemma 4.16 and for some constant $C = C(M_{\mathcal{F}})$

$$\mathrm{dist}_{\mathcal{F}}(0,x) - Cr \le \overline{\mathbf{U}}_h^{\varepsilon}(x;x_*) \le \mathbf{U}_h^{\varepsilon}(x;x_*) \le \underline{\mathbf{u}}_h^{\varepsilon}(x) \le \mathrm{dist}_{\mathcal{F}}(0,x) + C(r + \varepsilon \ln(\tfrac{\varepsilon}{h})), \tag{65}$$

and therefore $|\mathbf{U}(x;x_*) - \mathrm{dist}_{\mathcal{F}}(x_*,x)| \le (2C + C_{\mathcal{F}})r$ when in addition $\varepsilon \ln(\varepsilon/h) \le r$, noting that $|\,\mathrm{dist}_{\mathcal{F}}(x_*,x) - \mathrm{dist}_{\mathcal{F}}(0,x)| \le C_{\mathcal{F}} r$.

Now let $K_* \subset \Omega$ be a compact set. Up to reducing r one can find a finite cover $K_* \subset \cup_{j=1}^{J} \mathbb{B}(y_j, r/2)$ such that $\mathbb{B}(y_j, 6r) \subset \Omega$ for all $1 \le j \le J$. Applying the above reasoning to each ball $\mathbb{B}_h(y_j, r/2)$, $1 \le j \le J$, instead of $\mathbb{B}_h(r/2)$, we obtain $|\mathbf{U}(x;x_*) - \mathrm{dist}_{\mathcal{F}}(x_*,x)| \le (2C + C_{\mathcal{F}})r$ for all $(x, x_*) \in \mathrm{int}_h(\Omega, r) \times (K_* \cap h\mathbb{Z}^d)$, when $(\varepsilon, h/\varepsilon, \varepsilon \ln h)$ is small enough. Since r can be chosen arbitrarily small, the result follows. □

Lemma 4.18. *If h/ε is sufficiently small, then for all $x_* \in \Omega_h$ such that $\mathbb{B}(x_*, R_{\mathcal{S}} h) \subset \Omega$ one has $1 \le \mathcal{L}_h^{\varepsilon} U_h^{\varepsilon}(x_*;x_*) \le 1 + C\frac{\varepsilon^2}{h^2}$ where $C = 2\|\,\mathrm{Tr}(A_b)\|_{\infty}$.*

Proof. We assume that $C_0 h \le \varepsilon$ where $C_0 = \|A_b^{-1} b\|_{\infty} R_{\mathcal{S}}$, and obtain by Proposition 3.9 that $\mathcal{L}_h^{\varepsilon}$ is DDE. By the comparison principle, one has $0 \le U_h^{\varepsilon}(x;x_*) \le 1$ for all $x \in \Omega_h$. Thus $1 \le \mathcal{L}_h^{\varepsilon} U_h^{\varepsilon}(x_*,x_*) \le a_h^{\varepsilon}(x_*)$, with the notations (59), since $\beta_{h,i}^{\varepsilon}(x_*) \ge 0$ for all $1 \le i \le I$. The result follows. □

Proof of inverse matrix convergence, Theorem 4.2. By definition of L_h^{ε} and U_h^{ε}

$$(L_h^{\varepsilon})^{-1}_{x_* x} = \frac{U_h^{\varepsilon}(x;x_*)}{\mathcal{L}_h^{\varepsilon} U_h^{\varepsilon}(x_*;x_*)}.$$

Thus $\varepsilon|\ln[(L_h^{\varepsilon})^{-1}_{x_* x}] - \ln U_h^{\varepsilon}(x;x_*)| \le \varepsilon \ln(1 + C\varepsilon^2/h^2)$, under the conditions of Lemma 4.18. Noting that $\varepsilon \ln(1 + C\varepsilon^2/h^2) \to 0$ as $(\varepsilon, h/\varepsilon, \varepsilon \ln h) \to 0$, and that $-\varepsilon \ln U_h^{\varepsilon}(x;x_*) \to \mathrm{dist}_{\mathcal{F}}(x_*,x)$ locally uniformly by Proposition 4.17, we conclude the proof. □

5 Application to regularized optimal transport

In this section, we describe a numerical approach to the 1-Wasserstein optimal transport problem, with cost defined as a Randers distance, and with entropic relaxation. Given probability measures $\mu, \nu \in \mathcal{P}(\Omega)$, the addressed problem reads

$$W_{\varepsilon}(\mu,\nu) := \inf_{P \in \Pi(\mu,\nu)} \int_{\Omega\times\Omega} C(x,y)\,\mathrm{d}P(x,y) - \varepsilon\,\mathrm{Ent}(P), \tag{66}$$

where $\varepsilon \ge 0$ is the entropic relaxation parameter, and where $\Pi(\mu,\nu)$ is the set of probability measures on $\Omega \times \Omega$ whose first and second marginals coincide respectively with μ and ν, known as *transport plans* between μ and ν. The transport cost and entropy are defined as

$$C(x,y) := \mathrm{dist}_{\mathcal{F}}(x,y), \qquad \mathrm{Ent}(P) := -\int_{\Omega\times\Omega} \ln\Big(\frac{\mathrm{d}P(x,y)}{e\,\mathrm{d}P_0(x,y)}\Big)\mathrm{d}P(x,y)$$

where $\mathcal{F}$ is a Randers metric on the domain Ω, subject to the well posedness assumptions listed in the last paragraph of §1, and P_0 is a reference measure on $\Omega \times \Omega$. The Euler constant e appearing

in $\mathrm{Ent}(P)$ only changes the entropy by an additive constant, since P has total mass one, and allows to simplify later calculations.

As mentioned in the introduction, our approach extends [19] from Riemannian to non-symmetric Randers metrics. However, the quadratic cost $\mathrm{dist}_{\mathcal{F}}(x,y)^2$ corresponding to the 2-Wasserstein distance cannot be addressed in our setting, see Remark 4.5. Let us also acknowledge that the effect of entropic relaxation cannot be ignored in the numerical implementation of this class of methods: indeed, empirically, the transport plan is blurred over a radius $\sqrt{\varepsilon}$, while ε itself must be substantially larger than the discretization grid scale, see Theorem 3.18. Nevertheless such as smoothing is not necessarily an issue in applications [19], and the estimation of the Wasserstein distance itself as $\varepsilon \to 0$ can be accelerated by suitable techniques [12].

5.1 Kantorovich duality, and Sinkhorn's algorithm

We assume in the following that μ and ν are supported on a finite set $X \subset \Omega$, and present in this setting Kantorovich's dual formulation of the optimal transport problem (66), and its numerical solution by Sinkhorn's algorithm. With a slight abuse of notation, we identify a measure μ on the finite set X (resp. P on $X \times X$), which is a weighted sum of Dirac masses $\mu = \sum_{x\in X} \mu_x \delta_x$, with the corresponding non-negative vector $(\mu_x)_{x\in X}$ (resp. matrix $(P_{xy})_{x,y\in X}$). With this convention, the set of probability measures on X, and of transport plans between two such probabilities, are defined as

$$\mathcal{P}(X) := \{\mu \in \mathbb{R}_+^X;\ \mu^\top \mathbb{1} = 1\}, \qquad \Pi(\mu,\nu) := \{P \in \mathbb{R}_+^{X\times X};\ P\mathbb{1} = \mu,\ P^\top \mathbb{1} = \nu\}, \tag{67}$$

where $\mathbb{R}_+ := [0,\infty[$ denotes the set of non-negative reals, and $\mathbb{1} = (1,\cdots,1)^\top \in \mathbb{R}^X$. In this discrete setting, the optimal transport problem (66) reads

$$W_\varepsilon(\mu,\nu) = \inf_{P\in\Pi(\mu,\nu)} \langle\!\langle P, C\rangle\!\rangle + \varepsilon \,\langle\!\langle P, \ln\big(\frac{P}{eP_0}\big)\rangle\!\rangle, \tag{68}$$

where $\langle\!\langle A, B\rangle\!\rangle := \mathrm{Tr}(A^\top B) = \sum_{x,y\in X} A_{xy}B_{xy}$. In (68) and below, the fraction bar, the logarithm and the exponential function apply componentwise to vectors and matrices. We assume that the reference measure $P_0 = (P^0_{xy})$ has positive entries, and use the standard convention $0 \times \infty = 0$ in the definition of the entropic term if some entries of $P \in \Pi(\mu,\nu)$ vanish. Noting that $s \in \mathbb{R}_{++} \mapsto s\ln s$ is convex and has a vertical tangent at the origin, we find that the minimization problem (68) is convex and that the optimal P has positive entries whenever $\varepsilon > 0$.

Kantorovich duality introduces potentials $\varphi, \psi \in \mathbb{R}^X$ to account for the equality constraints in (67), and uses Sion's minimax theorem [22] to re-order the sup and inf:

$$\begin{aligned}
W_\varepsilon(\mu,\nu) &= \inf_{P\in\mathbb{R}_+^{X\times X}} \Big(\langle\!\langle P, C\rangle\!\rangle + \varepsilon \,\langle\!\langle P, \ln\big(\frac{P}{eP_0}\big)\rangle\!\rangle + \sup_{\varphi,\psi\in\mathbb{R}^X} \langle\varphi, \mu - P\mathbb{1}\rangle + \langle\psi, \nu - P^\top\mathbb{1}\rangle \Big) \\
&= \sup_{\varphi,\psi\in\mathbb{R}^X} \Big(\langle\varphi,\mu\rangle + \langle\psi,\nu\rangle + \inf_{P\in\mathbb{R}_+^{X\times X}} \langle\!\langle P,\, C + \varepsilon\ln\big(\frac{P}{eP_0}\big) - \varphi\mathbb{1}^\top - \mathbb{1}\psi^\top\rangle\!\rangle \Big) \\
&= \sup_{\varphi,\psi\in\mathbb{R}^X} \langle\varphi,\mu\rangle + \langle\psi,\nu\rangle - \varepsilon\,\langle\!\langle P_0,\, \exp\big(\frac{\varphi\mathbb{1}^\top + \mathbb{1}\psi^\top - C}{\varepsilon}\big)\rangle\!\rangle.
\end{aligned} \tag{69}$$

The third line was obtained by solving, component-wise and in closed form, the minimization w.r.t. P. Namely, the convex one dimensional mapping $p \in \mathbb{R}_{++} \mapsto p\big(C_{xy} + \varepsilon\ln\big[p/(eP^0_{xy})\big] - \varphi_x - \psi_y\big)$ attains its minimum for

$$P_{xy} = P^0_{xy}\exp[(\varphi_x + \psi_y - C_{xy})/\varepsilon]. \tag{70}$$

Using the change of variables $\Phi = \exp(\varphi/\varepsilon)$ and $\Psi := \exp(\psi/\varepsilon)$ we conclude that

$$W_\varepsilon(\mu,\nu) = \varepsilon \max_{\Phi,\Psi\in\mathbb{R}^X_{++}} \langle \ln\Phi, \mu\rangle + \langle \ln\Psi, \nu\rangle - \langle \Phi^\top, K_\varepsilon \Psi\rangle, \tag{71}$$

where $K_\varepsilon = (K^\varepsilon_{xy})_{x,y\in X}$ with $K^\varepsilon_{xy} := P^0_{xy}\exp(-C_{xy}/\varepsilon)$. Note that the maximization problem (69) is strictly concave. The equivalent form (71) can be numerically solved using alternate maximization, in other words successively solving w.r.t. the unknown Φ with Ψ fixed (resp. w.r.t. Ψ with Φ fixed). This approach is known as Sinkhorn's algorithm [30], and is particularly simple and efficient since the optimal value w.r.t. either of these variables has a closed form, when the other variable is fixed. More precisely, given an arbitrary $\Psi_0 \in \mathbb{R}^X_{++}$ one defines for all $n \geq 0$

$$\Phi_n := \frac{\mu}{K_\varepsilon \Psi_n}, \qquad\qquad \Psi_{n+1} := \frac{\nu}{K_\varepsilon^\top \Phi_n}, \tag{72}$$

and the sequence $(\Phi_n, \Psi_n)_{n\geq 0}$ converges geometrically to a maximizer of (71), see [30].

5.2 Efficient computation

The more computationally intensive part of Sinkhorn's algorithm (72) is to repeatedly compute the matrix-vector products $K_\varepsilon\Phi_n$ and $K_\varepsilon^\top\Psi_n$ in (72), since the matrix K_ε is dense and large. An efficient way to approximate those products using Varadhan's formula was proposed in [31], in the case of Riemannian manifolds. We adapt here this approach to Randers manifolds, thus specializing to the case $K^\varepsilon_{xy} := \exp(-\operatorname{dist}_{\mathcal{F}}(x,y)/\varepsilon)$ where the reference measure $P_0 \equiv 1$ is the uniform probability, the transport cost is defined as a Randers distance $\operatorname{dist}_{\mathcal{F}}$, and where $X = \Omega_h$ is a domain discretized on a Cartesian grid of scale $h > 0$.

Under these conditions, denoting by L^ε_h the matrix of our linear discretization scheme (32) with null boundary conditions, one has by Theorem 4.2

$$[L^\varepsilon_h]^{-1}_{xy} = \exp(-\frac{\operatorname{dist}_{\mathcal{F}}(x,y) + o(1)}{\varepsilon}), \qquad\qquad \text{as } (\varepsilon, h/\varepsilon, \varepsilon\ln h) \to 0, \tag{73}$$

locally uniformly on $\Omega\times\Omega$. Therefore the dense matrix product $\Phi' = K_\varepsilon\Phi$ can be approximated by solving the sparse linear system $\Phi = L^\varepsilon_h \Phi'$, which is considerably less memory intensive, and has a lower complexity along the iterations especially if a sparse pre-factorization of the matrix L^ε_h is used.

6 Numerical results

We illustrate the numerical methods presented in this paper, for Randers distance computation and numerical optimal transport, with synthetic numerical experiments in dimension $d = 2$. Geodesic distance computation based on solving the heat or Poisson PDEs has already numerous applications [17, 34, 33] and is part of established algorithmic geometry libraries such as CGAL®. Likewise Wasserstein distance computation based on entropic relaxation is an established numerical approach [19, 31, 12]. The contributions of this paper are thus mostly theoretical, see §7.

The approach presented in this paper for Randers distance computation is applied in [33] to image segmentation problems, using numerical codes provided by the last author and with due acknowledgement[3]. Optimal transport w.r.t. Randers geometry and the present numerical method

[3]However [33, §2.2] attempts to relate the numerical method with the Finsler heat equation (49). This is incorrect to our belief, and was published without the knowledge of the authors of this paper.

is yet to find a concrete application, but let us nevertheless mention the following motivation which was recently presented to us: monitoring forest fires using a fleet of small drones, which requires spreading the agents over a large specified area, and involves strongly asymmetrical displacement costs depending on the winds and terrain.

In this numerical section, we compare in several occasions the results of the centered scheme $\mathcal{L}_h^\varepsilon$ (32) emphasized in this paper, with those of the upwind scheme $\mathcal{L}_h^{\varepsilon,+}$ (38) which is unconditionally stable but is also less accurate. We limit our experiments to two dimensional problems, consistently with the literature, and although our theoretical results apply in dimension three as well, due to the overwhelming cost of solving three dimensional Laplacian-like linear systems at the considered grid scales.

The PDE domain for the experiments presented in this section is the two dimensional unit ball $\Omega = \{x \in \mathbb{R}^2; |x| \leq 1\}$, which is discretized on a regular Cartesian grid, using finite differences modified as in (28) to account for the (null) boundary conditions on $\partial\Omega$. The grid scale $h = 0.00625$ commonly used in the experiments below corresponds to a grid of size 320×320 (intersected with the ball). In the first two problems we numerically approximate

$$\mathbf{u}(x) := \min_{y \in Y} \mathrm{dist}_{\mathcal{F}}(x, y), \tag{74}$$

where Y is a finite set of target points, and $\mathcal{F}$ is a Randers metric on Ω which is described in terms of the parameters A, b of its dual, see Lemma 2.6. From the convergence analysis standpoint, the case of finitely many isolated point sources is a straightforward generalization of the case of a single one considered §4, and considering targets instead of sources amounts to a change of sign in the asymmetric part of the metric as discussed below (4).

In our experiments, the largest contributor to computation time is the factorization of the sparse linear systems, using the SuperLU routine provided with the scipy Python package. In contrast, the preliminary step of scheme construction (including Selling's algorithm to decompose the matrix $A_b(x)$ at each point $x \in \Omega_h$, and sparse matrix assembly) only accounts for fraction of this cost, and the subsequent solve operation is approximately $10\times$ faster than matrix factorization. In the application to optimal transport, which is based on Sinkhorn's algorithm (72), the same linear system needs to be solved multiple times, and thus a single matrix factorization is followed by 13 to 54 solve operations. The SuperLU factorization time when using a 320×320 discretization grid (thus $\approx 10^5$ unknowns) ranges from $1s$ to $1.6s$ depending on the test case, on a laptop equipped with a 2.3 GHz Intel Core i5 dual core processor.

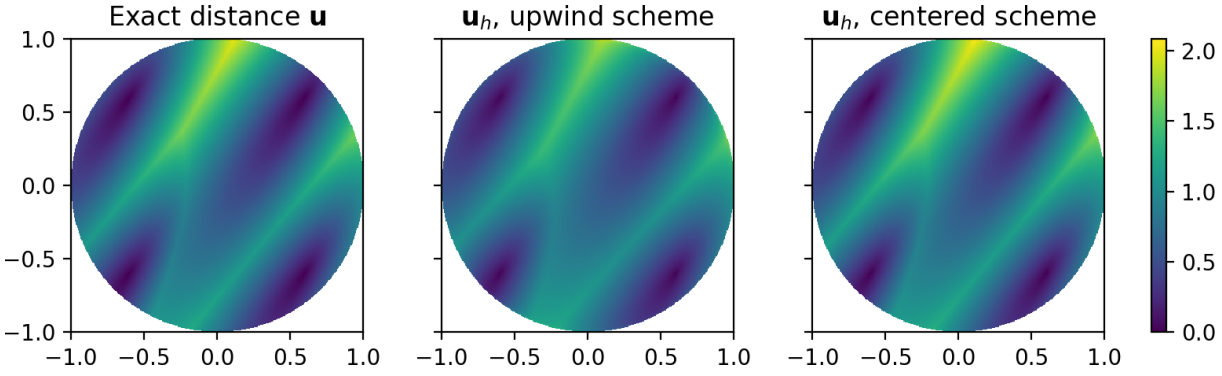

Figure 1: Randers distance with parameters (75). Left: exact solution. Center : solution based on the upwind scheme $\mathcal{L}_h^{\varepsilon,+}$ (38). Right : more accurate solution based on the centered scheme $\mathcal{L}_h^\varepsilon$ (32). In all cases $h = 0.00625$, $\varepsilon = 0.5h^{2/3}$.

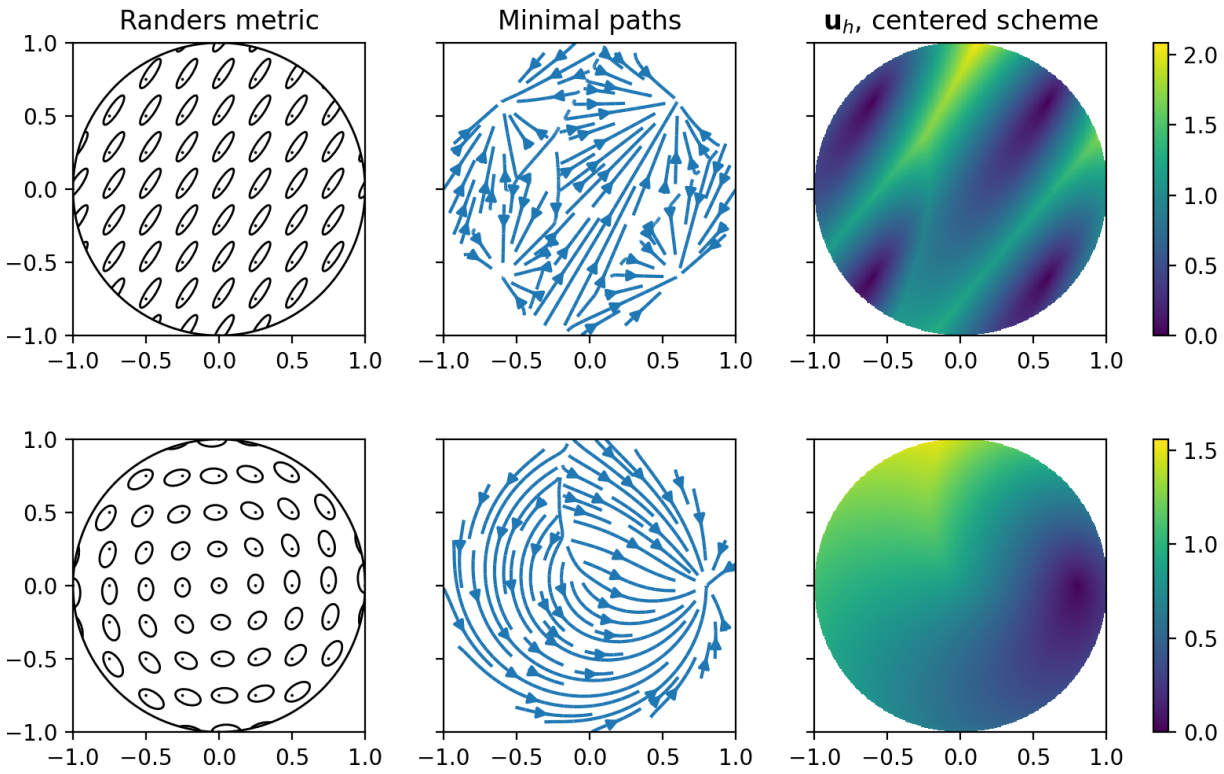

Figure 2: Representation of the Rander metric and approximations of minimal paths and of the Rander distance for parameters (75) (top), and (77) (bottom), with $h = 0.00625$ and $\varepsilon = 0.5h^{2/3}$.

Randers metric with constant coefficients. We consider a finite set Y of target points and a Randers metric whose dual $\mathcal{F}^*$ is defined by the following coefficients A, b

$$A := \begin{pmatrix} 0.5 & 0.6 \\ 0.6 & 1.0 \end{pmatrix}, \qquad b := \begin{pmatrix} 0.3 \\ 0.4 \end{pmatrix}, \qquad Y := \left\{ \begin{pmatrix} -0.6 \\ 0.6 \end{pmatrix}, \begin{pmatrix} -0.6 \\ -0.6 \end{pmatrix}, \begin{pmatrix} 0.6 \\ -0.6 \end{pmatrix}, \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix} \right\}. \tag{75}$$

Since the metric is constant and the domain is convex, the geodesic distance is explicit: $\mathrm{dist}_{\mathcal{F}}(x, y) = F(y - x)$ where $\mathcal{F}_x(v) = F(v)$ for all $x \in \Omega$, and the minimal paths are straight lines, see the discussion below Definition 2.3. In particular (74) can be evaluated exactly, which allows to estimate convergence rates.

The exact Randers distance from Y, and its approximation produced using the centered scheme (32) and the upwind scheme (38), are illustrated on Fig. 1. We present on Figure (2, top left) Tissot's indicatrix of the metric $\mathcal{F}$, which is a representation of the sets

$$\{x + v;\, v \in \mathbb{R}^2,\, \mathcal{F}_x(v) = r\}, \tag{76}$$

at a number of points $x \in \Omega$ and for a suitable radius $r > 0$. In Randers case, the set (76) is an ellipse which is *not* centered on the point x, and admits several equivalent characterizations see Lemma 2.7. The numerical approximation of Randers distance obtained with the centered scheme is illustrated on Figure (2, top right), while the numerical approximations of minimal paths from Y obtained by solving the ODE (17) are shown Figure (2, top center).

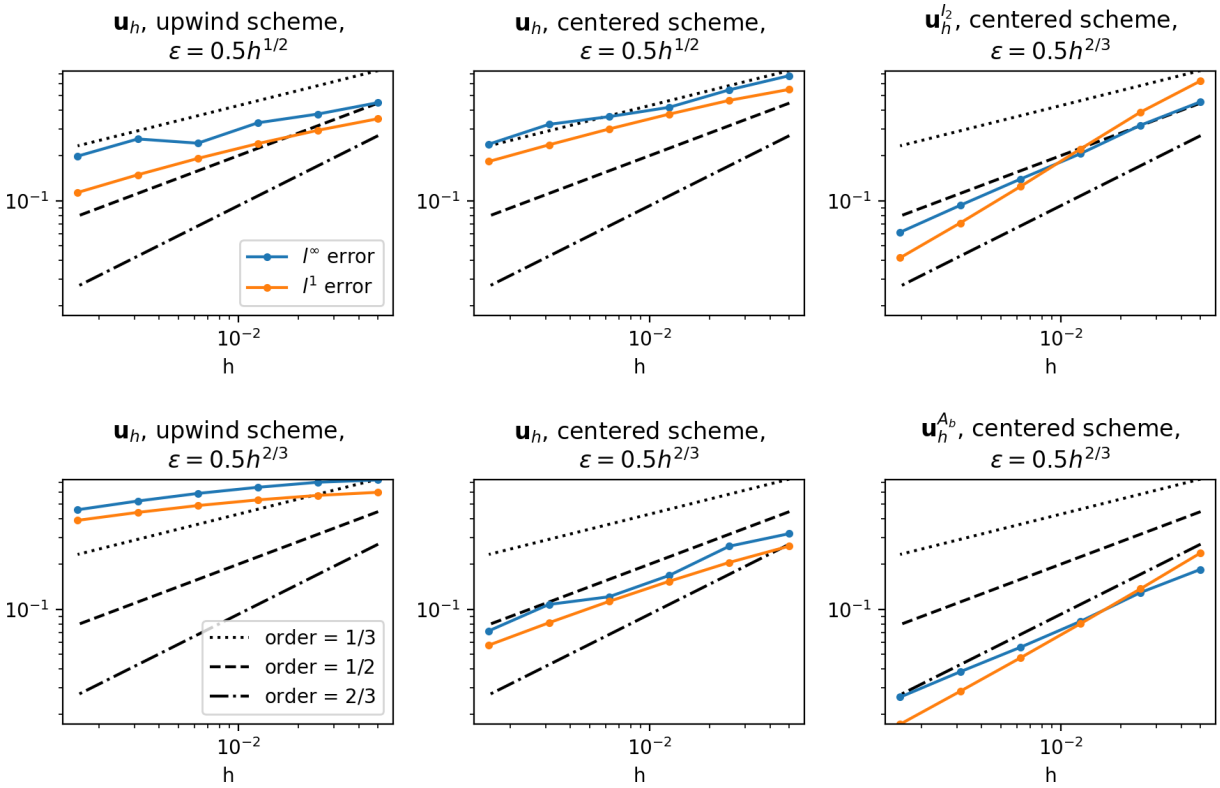

Figure 3: l^1 and l^∞ error between the exact distance $\mathbf{u}$, with parameters (75), and its numerical approximation, as a function of the grid scale h. Left : the upwind scheme $\mathcal{L}_h^{\varepsilon,+}$ (38) works best with $\varepsilon \approx h^{1/2}$. Center : the centered scheme is more accurate and works best with $\varepsilon \approx h^{2/3}$. The accuracy of the centered scheme solution is improved with a post-processing step, see Remark 3.2, which works best using the same stencil as the finite difference scheme (right, bottom), rather than an axis aligned stencil (right, top).

Randers metric with variable coefficients. A single target point is considered $Y = \{(0.8, 0)\}$, and the dual metric parameters are defined at $x = (x_1, x_2) \in \overline{\Omega}$ as

$$A(x) := \begin{pmatrix} 1 + \frac{2x_2^2}{|x|} + x_2^2 & -\frac{2x_1x_2}{|x|} - x_1x_2 \\ -\frac{2x_1x_2}{|x|} - x_1x_2 & 1 + \frac{2x_1^2}{|x|} + x_1^2 \end{pmatrix}, \qquad b(x) := x^\perp = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}, \tag{77}$$

where A is extended by continuity at the origin. Numerical results are shown Figure (2, bottom).

Numerical convergence rates. We discuss the convergence of some approximations of the exact distance function $\mathbf{u}$, defined by the metric parameters and target points (75). The l^∞ and l^1 errors between $\mathbf{u}$ and one of its approximations $\mathbf{u}_h^\varepsilon$ are respectively defined as

$$\max_{x \in \Omega_h} |\mathbf{u}_h^\varepsilon(x) - \mathbf{u}(x)|, \qquad h^2 \sum_{x \in \Omega_h} |\mathbf{u}_h^\varepsilon(x) - \mathbf{u}(x)|.$$

We display on Fig. 3 the convergence curves for the centered $\mathcal{L}_h^\varepsilon$ (32) and the (unconditionally stable but less accurate) upwind scheme $\mathcal{L}_h^{\varepsilon,+}$ (38), and for $\varepsilon = \frac{1}{2}h^\alpha$ where $\alpha \in \{1/2, 2/3\}$. Empirically, the centered scheme works best when $\alpha = 2/3$, and the upwind scheme when

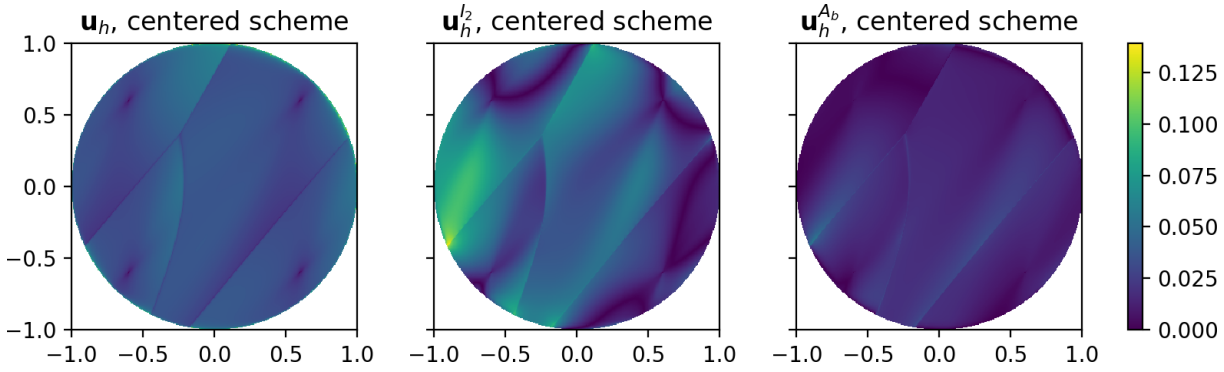

Figure 4: Absolute difference between the exact distance map $\mathbf{u}$ associated with the parameters (75) and its numerical approximation $\mathbf{u}_h^\varepsilon$ (left), the improved reconstruction using an axis-aligned stencil (center), or using the stencil of the finite difference scheme (right), see Remark 3.2. Grid scale $h = 0.0015625$ and $\varepsilon = 0.5h^{2/3}$.

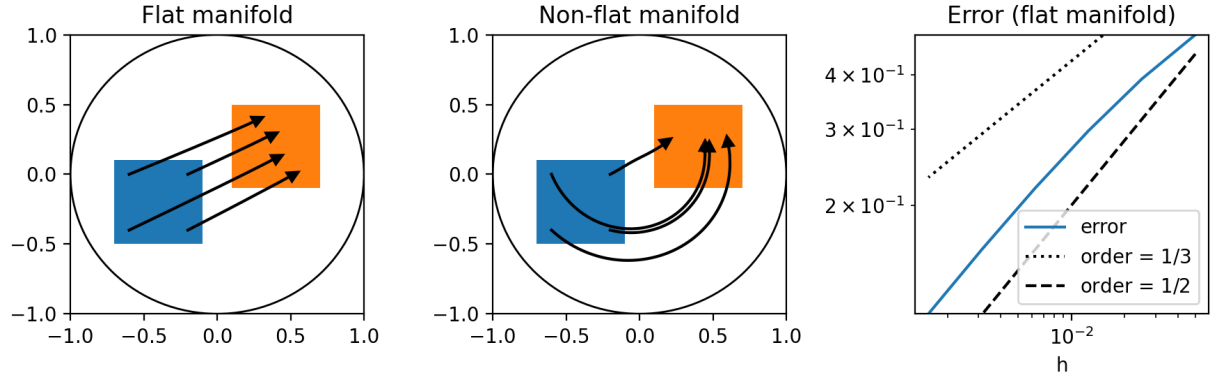

Figure 5: Numerical solution of the optimal transport problem (66). Left : manifold parameters (75), grid scale $h = 0.00625$. Middle : parameters (77), grid scale $h = 0.00625$. Right : convergence toward the exact Wasserstein distance as $h \to 0$, with parameters (75). In all cases: $\varepsilon = 0.5h^{\frac{2}{3}}$.

$\alpha = 1/2$. This experiment illustrates and empirically confirms Corollary 3.14, which establishes that the minimal consistency error with the eikonal equation is achieved when $\varepsilon \approx h^\alpha$, where $\alpha = 2/3$ for the centered scheme, and $\alpha = 1/2$ for the upwind scheme. Note however that the empirical solution error appears to be higher than the scheme consistency error, which is $\mathcal{O}(h^\alpha)$, see Corollary 3.14.

The post-processing step discussed in Remark 3.2, and adapted from [17], allows to improve the accuracy of our numerical solution of the Randers eikonal equation solution, as illustrated on Fig. 3 and 4. This post-processing works best when using the stencil of the finite difference scheme, as opposed to a basic axis-aligned stencil, see Fig. 4 and the last sentence of Remark 3.2.

Optimal transport problems. On Fig. 5, we solve numerically the optimal transport problem (66), where μ and ν are uniform probability measures on $[-0.7, -0.1] \times [-0.5, 0.1]$ and $[0.1, 0.7] \times [-0.1, 0.5]$ respectively. We use Sinkhorn's algorithm (72) to numerically approximate the exponential Kantorovitch potentials $\Phi, \Psi \in \mathbb{R}_+^{\Omega_h}$ maximizing (71), using the efficient approximation (73) of the product with the kernel $K_\varepsilon = \exp(-\operatorname{dist}_{\mathcal{F}}(x, y)/\varepsilon)$. The arrows on the figure follow Randers geodesics and illustrate a numerical approximation of the mapping $\sigma : \Omega_h \to \Omega$ defined

by

$$\sigma(x) := \sum_{y \in \Omega_h} P_{xy} y, \tag{78}$$

where $(P_{xy})_{x,y\in\Omega_h}$ is the optimal coupling measure (70) for the optimal transport problem (68). Thus $\sigma(x)$ is the barycenter of the image by the transport plan of the Dirac mass at x. The numerical evaluation of σ involves a product with the kernel K_ε which again is efficiently approximated using (73). Note that the coupling measure P is typically not supported on a graph, not even approximately, and that σ is not a one to one mapping. In particular, σ does not approximate a translation in Figure (5, right). This behavior reflects the specific properties of the 1-Wasserstein distance, as opposed to the p-Wasserstein distance for $p > 1$, and it is not related to our numerical approximation procedure. Figure (5, right) displays the error between the approximation $W_h^\varepsilon(\mu, \nu)$ of the Wasserstein distance obtained with grid scale $h > 0$ and entropic relaxation $\varepsilon = \frac{1}{2}h^{\frac{2}{3}}$, and the exact optimal transport cost corresponding to the continuous problem without relaxation $\varepsilon = h = 0$.

7 Conclusions

In this paper, we introduced and studied a numerical scheme for approximating geodesic distances by solving a linear finite differences scheme, with an application to Schrödinger's entropic relaxation of the optimal transport problem. The approach builds on previous works [32, 17, 19, 31, 34, 33], and brings the following contributions: (i) justification of the distance computation method in the case of point sources, which is a common setting in applications, (ii) identification of the optimal parameter scaling $\varepsilon = h^{\frac{2}{3}}$, in contrast with the commonly used scaling $h = c\varepsilon$ which is inconsistent asymptotically (45), (iii) extension of these methods to asymmetric geometries defined by Randers metrics.

Our numerical scheme obeys the discrete degenerate ellipticity property, and thus benefits from comparison principles, numerical stability, and a convergence proof in the setting of viscosity solutions. For that purpose we use adaptive finite differences offsets depending on the PDE parameters and obtained via a tool from discrete geometry known as Selling's decomposition of positive definite matrices [29, 15]. Our convergence proof (in the case of a point source) exploits fine properties of Selling's decomposition: uniqueness, Lipschitz regularity, and spanning property (which implies the local connectivity of the stencils derived from it), for the first time in the context of PDE analysis [21, 9, 7, 23, 25]. Future work will be devoted to investigating their relevance in other applications to numerical analysis, and possible substitutes in dimension $d \geq 4$ where Selling's decomposition does not apply.

References

[1] Stan Alama, Lia Bronsard, and J Alberto Montero. On the Ginzburg-Landau model of a superconducting ball in a uniform field. In *Annales de l'IHP Analyse non linéaire*, pages 237–267, 2006.

[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*. Modern Birkhäuser Classics. Birkhäuser, Basel, 1997.

[4] G. Barles and B. Perthame. Exit time problems in optimal control and vanishing viscosity method. *SIAM J. Control Optim.*, 26(5):1133–1148, 1988.

[5] G. Barles and E. Rouy. A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its applications. *Comm. Partial Differential Equations*, 23(11–12):1995–2033, 1998.

[6] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[7] Frédéric Bonnans, Guillaume Bonnet, and Jean-Marie Mirebeau. Second order monotone finite differences discretization of linear anisotropic differential operators. *Arxiv preprint*, 2020.

[8] Frédéric Bonnans and Stéphane Gaubert. *Recherche opérationnelle. Aspects mathématiques et applications*. Ellipse, 2016.

[9] Joseph Bonnans, Guillaume Bonnet, and Jean-Marie Mirebeau. Monotone and second order consistent scheme for the two dimensional Pucci equation. 2020.

[10] Eduardo Casas. Control of an elliptic problem with pointwise state constraints. *SIAM Journal on Control and Optimization*, 24(6):1309–1318, 1986.

[11] 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.

[12] Lenaic Chizat, Pierre Roussillon, Flavien Léger, François Xavier Vialard, and Gabriel Peyré. Faster Wasserstein Distance Estimation with the Sinkhorn Divergence. *Advances in Neural Information Processing Systems*, 33, 2020.

[13] Laurent D. Cohen, Da Chen, and Jean-Marie Mirebeau. Finsler Geodesics Evolution Model for Region based Active Contours. In Edwin R Hancock Richard C Wilson and William A P Smith, editors, *Proceedings of the British Machine Vision Conference (BMVC)*, pages 22.1–22.12. BMVA Press, September 2016.

[14] Michael B Cohen, Jonathan Kelner, Rasmus Kyng, John Peebles, Richard Peng, Anup B Rao, and Aaron Sidford. Solving directed laplacian systems in nearly-linear time through sparse LU factorizations. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 898–909. IEEE, 2018.

[15] J. H. Conway and N. J. A. Sloane. Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices. *Proc. R. Soc. A*, 436(1896):55–68, 1992.

[16] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67, 1992.

[17] K. Crane, C. Weischedel, and M. Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Trans. Graph.*, 32(5):152:1–152:11, 2013.

[18] Keenan Crane, Marco Livesu, Enrico Puppo, and Yipeng Qin. A Survey of Algorithms for Geodesic Paths and Distances. *arXiv preprint arXiv:2007.10430*, 2020.

[19] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Proc. 26th International Conference on Neural Information Processing Systems — Volume 2*, pages 2292–2300, 2013.

[20] 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.

[21] Jérôme Fehrenbach and Jean-Marie Mirebeau. Sparse non-negative stencils for anisotropic diffusion. *Journal of Mathematical Imaging and Vision*, 49(1):123–147, 2014.

[22] Hidetoshi Komiya. Elementary proof for Sion's minimax theorem. *Kodai mathematical journal*, 11(1):5–7, 1988.

[23] J.-M. Mirebeau. Fast-marching methods for curvature penalized shortest paths. *J. Math. Imaging Vision*, 60(6):784–815, 2018.

[24] Jean-Marie Mirebeau. Efficient fast marching with Finsler metrics. *Numerische Mathematik*, 126(3):515–557, 2014.

[25] Jean-Marie Mirebeau. Riemannian Fast-Marching on Cartesian Grids, Using Voronoi's First Reduction of Quadratic Forms. *SIAM Journal on Numerical Analysis*, 57(6):2608–2655, 2019.

[26] 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.

[27] Shin-ichi Ohta and Karl-Theodor Sturm. Heat flow on Finsler manifolds. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 62(10):1386–1433, 2009.

[28] Gunnar Randers. On an Asymmetrical Metric in the Four-Space of General Relativity. *Physical Review*, 59(2):195–199, January 1941.

[29] E. Selling. Uber die binären und ternären quadratischen Formen. *J. Reine Angew. Math.*, 77:143–229, 1874.

[30] Richard Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The Annals of Mathematical Statistics*, 35(2):876–879, 1964.

[31] J. Solomon, F. de Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas. Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Trans. Graph.*, 34(4):66:1–66:11, 2015.

[32] S. R. S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20(2):431–455, 1967.

[33] Fang Yang, Li Chai, Da Chen, and Laurent D. Cohen. Geodesic via Asymmetric Heat Diffusion Based on Finsler Metric. In *Asian Conference on Computer Vision*, pages 371–386. Springer, 2018.

[34] 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.

A Viscosity solutions

In this appendix, we establish the existence, uniqueness, comparison principles and convergence properties announced in §2 for the following three PDEs:

$$u + 2\varepsilon\langle\nabla u, b\rangle - \varepsilon^2 \operatorname{Tr}(A_b\nabla^2 u) = 0 \text{ in } \Omega, \qquad u - \exp(-g/\varepsilon) = 0 \text{ on } \partial\Omega, \tag{79}$$

$$|\nabla \mathbf{u}|_A + \langle\nabla \mathbf{u}, b\rangle - 1 = 0 \text{ in } \Omega, \qquad \mathbf{u} - g = 0 \text{ on } \partial\Omega, \tag{80}$$

$$|\nabla \mathbf{u}|^2_{A_b} + 2\langle\nabla \mathbf{u}, b\rangle - \varepsilon \operatorname{Tr}(A_b\nabla^2 \mathbf{u}) - 1 = 0 \text{ in } \Omega, \qquad \mathbf{u} - g = 0 \text{ on } \partial\Omega, \tag{81}$$

The linear PDE (79), introduced in (3), is the foundation of our approach to Randers distance computation. The Randers eikonal PDE (80), which can be rephrased in many equivalent forms, see (5) and Corollary 2.9, characterizes Randers distance from the domain boundary with initial time penalty g. Finally (81) makes the link between the first two equations, being equivalent for any $\varepsilon > 0$ to (79) up to a logarithmic transformation of the unknown, and being equivalent for $\varepsilon = 0$ to (80). We recall that, by assumption, Ω is a bounded, connected and open domain with a $W^{3,\infty}$ boundary and $g \in C(\partial\Omega)$. The fields $A : \overline{\Omega} \to S_d^{++}$ and $b : \overline{\Omega} \to \mathbb{R}^d$ are Lipschitz, and $A_b := A - bb^\top$ is pointwise positive definite over $\overline{\Omega}$.

The content of this section is presented in the appendix because it often mirrors similar results presented in the discrete setting of §3 which we have chosen to emphasize, and because several key results are obtained by specialization of [4, 5, 3, 16]. We present in Appendix A.1 the concepts of degenerate elliptic operator and of viscosity solution to a PDE, and we justify the change of unknown known as the logarithmic transformation. The comparison principle, established in Appendix A.2 for the PDEs of interest, implies the uniqueness and boundedness of their solutions in Ω. We prove in Appendix A.3 the validity of the explicit solutions to (79) and (80) defined as a distance map (14) and as the expectation (23) of the stochastic process (22), and we establish convergence as $\varepsilon \to 0$.

A.1 Degenerate ellipticity, change of unknowns

The PDEs considered in this appendix (79) to (81) benefit from a common structure, known as degenerate ellipticity [16, 26], introduced in Definition A.1 below and whose discrete counterpart is presented in Definition 3.3.

Definition A.1 (Degenerate ellipticity)**.** *An operator $F\colon \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$, denoted $F(x,t,p,X)$, is said* degenerate elliptic[4] *if it is (i) non-decreasing w.r.t. the second variable t, and (ii) non-increasing w.r.t. the last variable X for the Loewner order. The operator F is said elliptic if $F(x,t,p,X) - \delta t$ is degenerate elliptic for some constant $\delta > 0$.*

The Dirichlet problem for a degenerate elliptic equation writes as

$$F(x, u(x), \nabla u(x), \nabla^2 u(x)) = 0 \text{ in } \Omega, \qquad u(x) - \psi(x) = 0 \text{ on } \partial\Omega, \tag{82}$$

where $\psi\colon \partial\Omega \to \mathbb{R}$. For example when considering equation (80), one should choose

$$F(x,t,p,X) = |p|_{A(x)} + \langle p, b(x)\rangle - 1, \qquad \psi(x) = g(x).$$

This specific operator F is degenerate elliptic, since $F(x,t,p,X)$ does not depend on either t or X, and thus obeys the required monotony conditions. Equation (81) is likewise defined by a

[4] Or *proper degenerate elliptic* in the wording of [16]. For consistency with the discrete case Definition 3.3, and following [26], we drop the 'proper' qualifier.

degenerate elliptic operator, because the matrix field A_b is positive semi-definite. Equation (79) is elliptic thanks to the additional zeroth order term.

In the discrete setting, a comparison principle can be directly derived from the definition of ellipticity, see Lemma 3.4, and the related notions of sub-solution and super-solution are straightforward. Some additional care is however needed in the continuous case, see Definition A.2, Proposition A.7 and Theorem A.8 below. For any bounded function $u\colon \overline{\Omega} \to \mathbb{R}^d$, we denote respectively by $u^*\colon \overline{\Omega} \to \mathbb{R}$ and $u_*\colon \overline{\Omega} \to \mathbb{R}$ its upper semicontinuous and lower semicontinuous enveloppes, defined by

$$u^*(x) := \limsup_{y \in \overline{\Omega},\, y \to x} u(y), \qquad\qquad u_*(x) := \liminf_{y \in \overline{\Omega},\, y \to x} u(y). \tag{83}$$

Definition A.2 (Viscosity solution)**.** *Let $F\colon \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$ be a continuous degenerate elliptic operator and let $\psi \in C(\partial\Omega)$. A bounded function $u\colon \overline{\Omega} \to \mathbb{R}$ is a* viscosity sub-solution *to* (82) *if for any $\varphi \in C^2(\overline{\Omega})$ and any local maximum $x \in \overline{\Omega}$ of $u^* - \varphi$,*

$$\begin{cases} F(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \le 0 & \text{if } x \in \Omega, \\ \min\{u^*(x) - \psi(x),\, F(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)\} \le 0 & \text{if } x \in \partial\Omega. \end{cases}$$

It is a viscosity super-solution *if for any $\varphi \in C^2(\overline{\Omega})$ and any local minimum $x \in \overline{\Omega}$ of $u_* - \varphi$,*

$$\begin{cases} F(x, u_*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \ge 0 & \text{if } x \in \Omega, \\ \max\{u_*(x) - \psi(x),\, F(x, u_*(x), \nabla\varphi(x), \nabla^2\varphi(x)\} \ge 0 & \text{if } x \in \partial\Omega. \end{cases}$$

It is a viscosity solution *if it is both a viscosity sub-solution and super-solution.*

Definition A.2 encompasses discontinuous solutions u, obeying the boundary conditions in a weak sense, which allows to implement outflow boundary conditions in the case of the eikonal equation (80) by using large enough boundary data g. A well-known property of viscosity solutions is their stability under monotone changes of variables.

Proposition A.3. *Let $F\colon \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$ be a continuous degenerate elliptic operator, let $\psi \in C(\partial\Omega)$, let I, $J \subset \mathbb{R}$ be open intervals, let $\eta\colon I \to J$ be a strictly increasing C^2-diffeomorphism, and let $v\colon \overline{\Omega} \to I$ be bounded away from ∂I. Define the continuous degenerate elliptic operator $G\colon \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$ and boundary condition $\chi : \partial\Omega \to \mathbb{R}$ by*

$$G(x, t, p, X) := F(x,\, \eta(t),\, \eta'(t)p,\, \eta''(t)p \otimes p + \eta'(t)X), \qquad\qquad \chi(x) := \eta^{-1}(\psi(x)).$$

Then $u := \eta \circ v$ is a viscosity sub-solution (respectively super-solution) to (82) *if and only if v is a viscosity sub-solution (respectively super-solution) to*

$$G(x, v(x), \nabla v(x), \nabla^2 v(x)) = 0 \text{ in } \Omega, \qquad\qquad v(x) - \chi(x) = 0 \text{ on } \partial\Omega. \tag{84}$$

Proof. We only show the result for sub-solutions, since the case of super-solutions is similar. We assume that v is a sub-solution to (84) and prove that u is a sub-solution to (82). The proof of the converse is the same, using that

$$F(x, t, p, X) = G(x, \eta^{-1}(t), (\eta^{-1})'(t)p, (\eta^{-1})''(t)p \otimes p + (\eta^{-1})'(t)X).$$

The assumption that v is bounded away from ∂I implies that v^* and v_* are valued in I, hence $u^* = (\eta \circ v)^* = \eta \circ v^*$ is valued in J and likewise for u_*, by continuity of η. Let $\varphi \in C^2(\overline{\Omega})$ and

$x \in \overline{\Omega}$ be a local maximum of $u^* - \varphi$. Without loss of generality, we may assume that $\varphi(\overline{\Omega}) \subset J$. Let $\tilde{\varphi} := \eta^{-1} \circ \varphi$. Using that η is strictly increasing, and $\varphi = \eta \circ \tilde{\varphi}$, we deduce that x is a local maximum of $v^* - \tilde{\varphi}$. We conclude the proof by noticing that for all $x \in \Omega$

$$F(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)) = G(x, v^*(x), \nabla\tilde{\varphi}(x), \nabla^2\tilde{\varphi}(x)).$$

In addition, if $x \in \partial\Omega$, then $u^*(x) - \psi(x)$ and $v^*(x) - \eta^{-1}(\psi(x))$ have the same sign. □

Remark A.4. *Sign changes exchange the notions of sub-solution and super-solution. More precisely, $u = -v$ is a viscosity sub-solution (resp. super-solution) to* (82) *iff v is a viscosity super-solution (resp. sub-solution) to* (84) *with*

$$G(x, t, p, X) := -F(x, -t, -p, -X), \qquad \chi(x) = -\psi(x).$$

Combining Proposition A.3 and Remark A.4 allows to address the decreasing change of unknown $u = \exp(-\mathbf{u}/\varepsilon)$ considered by Varadhan [32], see Lemma 2.11. Note the discrete counterpart Proposition 3.10 of this result.

Corollary A.5. *Let $\mathbf{u} : \overline{\Omega} \to \mathbb{R}$, and let $u := \exp(-\mathbf{u}/\varepsilon)$. Then $\mathbf{u}$ is a sub-solution (resp. super-solution) to* (81) *iff u is a super-solution (resp. sub-solution) to* (79).

Proof. The PDE (79) corresponds to (82) with the following operator and boundary conditions

$$F(x, t, p, X) = t + 2\varepsilon\langle p, b(x)\rangle - \varepsilon^2 \operatorname{Tr}(A_b(x)X), \qquad \psi(x) = \exp(-g(x)/\varepsilon).$$

Applying successively Proposition A.3 with the increasing diffeomorphism $\eta(t) := -\exp(-t/\varepsilon)$, and Remark A.4, yields the boundary conditions $\chi(x) = -\eta^{-1}(\psi(x)) = g(x)$ and the operator

$$\begin{aligned} G(x, t, p, X) &= -F(x, -\eta(t), -\eta'(t)p, -\eta''(t)p \otimes p - \eta'(t)X) \\ &= -F\big(x, e^{-\frac{t}{\varepsilon}}, \ -\tfrac{1}{\varepsilon}e^{-\frac{t}{\varepsilon}}p, \ \tfrac{1}{\varepsilon^2}e^{-\frac{t}{\varepsilon}}p \otimes p - \tfrac{1}{\varepsilon}e^{-\frac{t}{\varepsilon}}X\big) \\ &= -e^{-\frac{t}{\varepsilon}}\big(1 - 2\langle p, b(x)\rangle - \langle p, A_b(x)p\rangle + \varepsilon \operatorname{Tr}(A_b(x)X)\big). \end{aligned}$$

Simplifying by the positive factor $e^{-\frac{t}{\varepsilon}}$, and distributing the minus sign, we recognize (81). □

A.2 The comparison principle

The linear PDE (79) and Randers eikonal equation (80) admit a *strong comparison principle*, which in particular implies that their viscosity solutions are uniquely determined on Ω — though not on $\partial\Omega$. The proofs, presented in Proposition A.7 and Theorem A.8 below, are obtained as a specialization of [5]. For that purpose, we reformulate the first order term of (80) in Bellman form, based on the following identity: for all $x \in \overline{\Omega}$ and all $w \in \mathbb{R}^d$

$$|w|_{A(x)} + \langle w, b(x)\rangle = \sup_{\alpha \in \mathbb{B}^d} -\langle w, \mathfrak{b}(x, \alpha)\rangle, \qquad \text{where} \quad \mathfrak{b}(x, \alpha) := A^{\frac{1}{2}}(x)\alpha - b(x), \tag{85}$$

where $\mathbb{B}^d := \{x \in \mathbb{R}^d;\ \|x\| \leq 1\}$ denotes the closed unit ball.

Lemma A.6. *The mappings $A^{\frac{1}{2}}, A_b^{\frac{1}{2}} : \overline{\Omega} \to S_d^{++}$ are Lipschitz continuous. The mapping $\mathfrak{b} : \overline{\Omega} \times \mathbb{B}^d \to \mathbb{R}^d$ defined by (85, right) is Lipschitz continuous. In addition, for each $x \in \overline{\Omega}$ and $p \in \mathbb{R}^d \setminus \{0\}$ there exists $\alpha \in \mathbb{B}^d$ such that $\langle \mathfrak{b}(x, \alpha), p\rangle > 0$.*

Proof. Recall that the mappings $A, A_b : \overline{\Omega} \mapsto S_d^{++}$ are Lipschitz continuous, and note that their lower eigenvalues are bounded away from zero by compactness. Since the matrix square root $\sqrt{\cdot} : S_d^{++} \to S_d^{++}$ is C^∞, as follows from holomorphic functional calculus[5], we obtain that $A^{1/2}$ and $A_b^{1/2}$ also are Lipschitz continuous on $\overline{\Omega}$. The announced regularity of $\mathfrak{b}$ follows.

Regarding the last property, we observe that choosing $\alpha = A^{\frac{1}{2}}(x)p/|A^{\frac{1}{2}}(x)p|$ yields

$$\langle \mathfrak{b}(x,\alpha),\, p\rangle = \langle \alpha - A^{-\frac{1}{2}}(x)b(x),\, A^{\frac{1}{2}}(x)p\rangle \geq \big(1 - |A^{-\frac{1}{2}}(x)b(x)|\big)|A^{\frac{1}{2}}(x)p| > 0, \tag{86}$$

since $|A^{-\frac{1}{2}}(x)b(x)| = |b(x)|_{A(x)^{-1}} < 1$ over $\overline{\Omega}$ by assumption. □

The comparison principle established in [5, Theorem 2.1] encompasses both the second order linear PDE (79), and the first order non-linear PDE (80) considered in this paper, although a reformulation is needed in the latter case.

Proposition A.7. *Let $\overline{u}$ and $\underline{u}$ be respectively a sub-solution and a super-solution of the linear PDE (79), for some $\varepsilon > 0$. Then $\overline{u}^* \leq \underline{u}_*$ in Ω.*

Proof. The announced result is a direct application of [5, Theorem 2.1], using that $A_b^{1/2} \colon \mathbb{R}^d \to \mathcal{S}_d^{++}$ and $b \colon \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz continuous, $\partial\Omega$ is of class $W^{3,\infty}$, and $g \in C(\partial\Omega)$. □

Theorem A.8. *Let $\overline{\mathbf{u}}$, $\underline{\mathbf{u}} : \overline{\Omega} \to \mathbb{R}$ be respectively a sub-solution and a super-solution of (80). Then $\overline{\mathbf{u}}^* \leq \underline{\mathbf{u}}_*$ in Ω.*

Proof. Since (80) involves an operator which is degenerate elliptic but not elliptic, see Definition A.1, we perform the Kruzhkov exponential change of variables and define $\overline{\mathbf{v}} := -\exp(-\overline{\mathbf{u}})$ and $\underline{\mathbf{v}} := -\exp(-\underline{\mathbf{u}})$. By Proposition A.3, $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$ are respectively a viscosity sub-solution and super-solution to

$$|\nabla \mathbf{v}(x)|_{A(x)} + \langle \nabla \mathbf{v}(x), b(x)\rangle + \mathbf{v}(x) = 0 \text{ in } \Omega, \qquad \mathbf{v}(x) + \exp(-g(x)) = 0 \text{ on } \partial\Omega.$$

The boundary $\partial\Omega$ is of class $W^{3,\infty}$, and the boundary data $-\exp(-g) \in C(\partial\Omega)$, consistently with the framework of [5]. Furthermore, the PDE can be rewritten as $\sup_{\alpha\in\mathbb{B}^d} -\langle \mathfrak{b}(x,\alpha), \nabla\mathbf{v}(x)\rangle + \mathbf{v}(x) = 0$ in Ω, and the required regularity properties of $\mathfrak{b}$ are established in Lemma A.6, as well as the additional condition which amounts to a local controllability property. Then by [5, Theorem 2.1], we obtain $\overline{\mathbf{v}}^* \leq \underline{\mathbf{v}}_*$ in Ω, and therefore $\overline{\mathbf{u}}^* \leq \underline{\mathbf{u}}_*$ in Ω as announced. □

A.3 Explicit solutions, and convergence

We establish that viscosity solutions to Randers eikonal equation (80) and to the linear PDE (79) may be explicitely obtained as the distance from the boundary (4) with suitable penalty term, and as the expectation of a stochastic process (23). We also prove bounds for these solutions, see Theorems A.9 and A.11, and conclude the proof of Varadhan's formula for Randers metrics in Theorem A.12.

Theorem A.9. *Denote by $\mathcal{F}$ the Randers metric of parameters (M,ω) dual to (A,b), see Lemma 2.6. Then $\mathbf{u} : x \in \overline{\Omega} \mapsto \min_{p\in\partial\Omega} d_{\mathcal{F}}(p,x) + g(p)$ is a bounded viscosity solution to (80).*

[5] More directly, if the eigenvalues of $A \in S_d^{++}$ lie in $]0, 2r[$, then one has the series expansion $\sqrt{A} = \sqrt{r}\sum_{k\geq 0} a_k (A/r - \mathrm{Id})^k$, where $\sqrt{1+t} = \sum_{k\geq 0} a_k t^k$ for all $t \in]-1, 1[$.

Proof. The boundedness of $\mathbf{u}$ follows from the equivalence of the Randers distance with the Euclidean distance, see Lemma 2.4. Since $g \in C(\partial\Omega)$ and the control function $\mathfrak{b}$ is Lipschitz continuous [3, Theorem V.4.13 and Remark V.4.14] yields a viscosity solution $\mathbf{v}$ to (80) in the form

$$\mathbf{v}(x) = \inf\{T + g(\gamma_x^\alpha(T));\, T \geq 0,\, \alpha : [0,T] \to \mathbb{B}^d, \gamma_x^\alpha(T) \in \partial\Omega\} \tag{87}$$

where $\gamma = \gamma_x^\alpha$ is defined by $\gamma(0) = x$ and $\gamma'(t) = \mathfrak{b}(\gamma(t), \alpha(t))$ for all $0 \leq t \leq T$, and where α is implicitly assumed to be measurable. Now, for any $v \in \mathbb{R}^d$ one obtains, omitting the argument x in $M(x)$, $\omega(x)$, $A(x)$ and $b(x)$ for readability

$$\mathcal{F}_x(v) \leq 1 \Leftrightarrow |v|_M + \langle \omega, v\rangle \leq 1 \Leftrightarrow |v-b|_{A^{-1}} \leq 1 \Leftrightarrow \exists \tilde{\alpha} \in \mathbb{B}^d, v - b = A^{\frac{1}{2}}\tilde{\alpha} \Leftrightarrow \exists \alpha \in \mathbb{B}^d, v = -\mathfrak{b}(x,\alpha),$$

where the first equivalence holds by definition, the second is established in Lemma 2.7, the third follows from $|A^{\frac{1}{2}}\tilde{\alpha}|_{A^{-1}} = |\tilde{\alpha}|$ for any $\tilde{\alpha} \in \mathbb{R}^d$, and the last is obtained by choosing $\alpha = -\tilde{\alpha}$. Thus

$$\begin{aligned}\mathbf{v}(x) = \inf\{T + g(\gamma(T)); T \geq 0, \exists \gamma \in \mathrm{Lip}([0,T], \overline{\Omega}), \gamma(0) = x, \gamma(T) \in \partial\Omega,\\ \mathcal{F}_{\gamma(t)}(-\gamma'(t)) \leq 1, \text{ for a.e. } t \in [0,T]\}.\end{aligned}$$

Noting that any Lipschitz path can be reparametrized at constant speed w.r.t. the metric $\mathcal{F}$, and have its orientation reversed (from x to $\partial\Omega$), we obtain that $\mathbf{v}(x) = \mathbf{u}(x)$, which concludes the proof. □

We obtain a sub-solution and a super-solution to the PDE (81), independent of the relaxation parameter, similarly to the discrete case in Lemma 3.15

Lemma A.10. *The PDE* (81) *admits, for any $\varepsilon \geq 0$, the constant sub-solution $\overline{\mathbf{u}} : x \in \overline{\Omega} \mapsto g_{\min}$, where $g_{\min} := \min\{g(y);\, y \in \partial\Omega\}$. It also admits the affine super-solution $\underline{\mathbf{u}} : x \in \overline{\Omega} \mapsto \langle p, x\rangle + c_{\max}$, for any $p \in \mathbb{R}^d$ such that $|p|$ is sufficiently large, where $c_{\max} := \max\{g(y) - \langle p, y\rangle;\, y \in \partial\Omega\}$.*

Proof. Denote $\mathcal{S}^\varepsilon\mathbf{u} := |\nabla\mathbf{u}|^2_{A_b} + 2\langle\nabla\mathbf{u}, b\rangle - \varepsilon\,\mathrm{Tr}(A_b\nabla^2\mathbf{u}) - 1$ the operator of (81). Clearly $\mathcal{S}^\varepsilon\overline{\mathbf{u}} = -1 < 0$ in $\overline{\Omega}$, whereas $\mathcal{S}^\varepsilon\underline{\mathbf{u}}(x) = |p|^2_{A_b(x)} + 2\langle p, b(x)\rangle - 1 \geq c_0 > 0$ for all $x \in \overline{\Omega}$, provided $|p|$ is sufficiently large, since A_b and b are bounded over $\overline{\Omega}$, and A_b is uniformly positive definite. The constants $g_{\min}$ and $c_{\max}$ are chosen so as to comply with the boundary conditions. □

Theorem A.11. *For any $\varepsilon > 0$, the function $u_\varepsilon \colon \overline{\Omega} \to \mathbb{R}_-$ defined by* (23) *is a viscosity solution to* (79)*. In addition, u_ε is positive, and $\overline{\mathbf{u}} \leq \mathbf{u}_\varepsilon \leq \underline{\mathbf{u}}$ in Ω, where $\mathbf{u}_\varepsilon := -\varepsilon\ln(u_\varepsilon)$ and $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$ are from Lemma A.10.*

Proof. Since $A_b^{1/2} \colon \mathbb{R}^d \to \mathcal{S}_d^{++}$ and $b \colon \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz continuous, $\partial\Omega$ is of class $W^{3,\infty}$, and $g \in C(\partial\Omega)$, [5, Theorem 3.1] implies that u_ε is a viscosity solution to (79).

By Corollary A.5, $\overline{u}_\varepsilon := \exp(-\underline{\mathbf{u}}/\varepsilon)$ and $\underline{u}_\varepsilon := \exp(-\overline{\mathbf{u}}/\varepsilon)$ are respectively a sub-solution and a super-solution to (79). Thus $\overline{u}_\varepsilon \leq (u_\varepsilon)_* \leq u_\varepsilon \leq (u_\varepsilon)^* \leq \underline{u}_\varepsilon$ in Ω by Theorem A.8. Therefore u_ε is positive, as announced, and we conclude using the monotony of the logarithm. □

We are able to complete the proof of formula (24) by making rigorous the passing to the limit between problems (81) and (80). Note that we follow a standard sketch of proof, already used in [4, Proposition II.6] for example.

Theorem A.12. *With the notations of Theorem A.11, and denoting by $\mathbf{u}$ the solution to* (15)*, one has $\mathbf{u}_\varepsilon \to \mathbf{u}$ uniformly on compact subsets of Ω, as $\varepsilon \to 0$.*

Proof. By Theorem A.11, $\mathbf{u}_\varepsilon$ is bounded above and below, uniformly on $\overline{\Omega}$ and uniformly w.r.t. $\varepsilon > 0$. Therefore the following limit is well defined, for any $x \in \overline{\Omega}$

$$\overline{\mathbf{v}}(x) := \limsup_{\varepsilon \to 0, y \to x} \mathbf{u}_\varepsilon(y) \quad \Big(= \lim_{\delta \to 0} \sup \big\{ \mathbf{u}_\eta(y);\, 0 < \eta \leq \delta,\, |y - x| \leq \delta \big\} \Big),$$

and likewise $\underline{\mathbf{v}}(x) := \liminf \mathbf{u}_\varepsilon(y)$ as $\varepsilon \to 0$ and $y \to x$. Thus we can apply [16, Lemma 6.1 and Remark 6.3] to functions $(\mathbf{u}_\varepsilon)_*$ and $(\mathbf{u}_\varepsilon)^*$, and deduce that $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$ are respectively a viscosity subsolution and supersolution to (81) with $\varepsilon = 0$, or equivalently to (80) by Corollary 2.9. Hence by Theorem A.8, $\overline{\mathbf{v}} \leq \mathbf{u}_* \leq \mathbf{u}^* \leq \underline{\mathbf{v}}$ on Ω. By definition, $\overline{\mathbf{v}} \geq \underline{\mathbf{v}}$ on $\overline{\Omega}$. Therefore $\overline{\mathbf{v}} = \mathbf{v} = \underline{\mathbf{v}}$ on Ω. The locally uniform convergence of $\mathbf{u}_\varepsilon$ to $\mathbf{u}$ on Ω follows from the definitions of $\overline{\mathbf{v}}$ and $\underline{\mathbf{v}}$. □

B Selling's decomposition of positive definite matrices

This appendix is devoted to a brief description of Selling's decomposition of symmetric positive definite matrices [29, 15] of dimension $d \in \{2, 3\}$, a tool from algorithmic geometry which is convenient when discretizing anisotropic PDEs on Cartesian grids [21, 23, 25, 7], here used §3.1. Selling's formula and algorithm are presented in Lemma B.2 and Proposition B.3. Two properties of the resulting normalized decomposition (91), established in Propositions B.5 and B.8, are used in §4.3 for the first time in the context of PDE numerical analysis.

Definition B.1. *A superbase of $\mathbb{Z}^d$ is a family $(v_0, \cdots, v_d) \in (\mathbb{Z}^d)^{d+1}$ such that $v_0 + \cdots + v_d = 0$ and $|\det(v_1, \cdots, v_d)| = 1$. It is said D-obtuse, where $D \in S_d^{++}$, iff $\langle v_i, Dv_j \rangle \leq 0$ for all $0 \leq i < j \leq d$.*

To each superbase $(v_0, \cdots, v_d)$ of $\mathbb{Z}^d$, we associate the family of vectors $e_{ij} \in \mathbb{Z}^d$, $0 \leq i < j \leq d$ defined by the linear relations

$$\langle e_{ij}, v_k \rangle = \delta_{ik} - \delta_{jk}, \tag{88}$$

for all $0 \leq k \leq d$, where δ_{ij} denotes Kroenecker's symbol. In dimension $d = 2$ (resp. $d = 3$), if $\{i, j, k\} = \{0, 1, 2\}$ (resp. $\{i, j, k, l\} = \{0, 1, 2, 3\}$), one easily checks that $e_{ij} = \pm v_k^\perp$ (resp. $e_{ij} = \pm v_k \times v_l$). Selling's formula and algorithm are classical [29, 15, 23], yet their (short) proofs are presented for completeness, since they are core elements of our numerical scheme.

Lemma B.2 (Selling's formula)**.** *Let $D \in S_d$ and let $(v_0, \cdots, v_d)$ be a superbase of $\mathbb{Z}^d$. Then*

$$D = -\sum_{0 \leq i < j \leq d} \langle v_i, Dv_j \rangle e_{ij} e_{ij}^\top. \tag{89}$$

Proof. By (88) we obtain $\langle v_i, Dv_j \rangle = \langle v_i, D'v_j \rangle$ for all $0 \leq i < j \leq d$, where D' denotes (89, rhs). Thus $\langle v_i, Dv_i \rangle = \langle v_i, D'v_i \rangle$ by linearity and since $v_i = -(v_0 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_d)$. The result follows since $(v_1, \cdots, v_d)$ is a basis. □

If $D \in S_d^{++}$ and $(v_0, \cdots, v_d)$ is D-obtuse, then (89) is known as Selling's decomposition of D. Selling's algorithm provides a constructive proof of existence of such a D-obtuse superbase, in dimension $d \in \{2, 3\}$.

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

$$b \leftarrow (-v_0, v_1, v_0 - v_1) \text{ if } d = 2, \qquad b \leftarrow (-v_0, v_1, v_2 + v_0, v_3 + v_0) \text{ if } d = 3. \tag{90}$$

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

Proof. Define $\mathcal{E}(b) = \sum_{i=0}^{d} \|v_i\|_D^2$. If $b = (v_0, \cdots, v_d)$ is such that $\delta := \langle v_0, Dv_1 \rangle > 0$, and if b' is defined by (90) then one easily checks that b' also is a superbase and that $\mathcal{E}(b') = \mathcal{E}(b) - C_d\delta$, where $C_2 = 4$ and $C_3 = 2$. There are only finitely many superbases of $\mathbb{Z}^d$ whose energy $\mathcal{E}$ is below any given bound, since their elements have integer coordinates and since D is positive definite. Hence Selling's algorithm must terminate, which happens when the iteration condition fails, i.e. when a D-obtuse superbase b is obtained. This concludes the proof. □

The elements of a D-obtuse superbase, and the corresponding offsets in Selling's formula, are bounded in terms of the anisotropy ratio $\mu(D) := \sqrt{\|D\|\|D^{-1}\|}$.

Proposition B.4. *Let $D \in S_d^{++}$, and let $b = (v_0, \cdots, v_d)$ be a D-obtuse superbase, where $d \in \{2, 3\}$. Then $|v_i| \leq C\mu(D)$, $0 < i < d$, and $|e_{ij}| \leq 2C\mu(D)$, $0 \leq i < j \leq d$, where $C = 2$ if $d = 2$ (resp. $C = 2\sqrt{3}$ if $d = 3$). In fact, one has the slightly stronger estimates $|v_i|_D \leq C\|D\|^{\frac{1}{2}}$ and $|e_{ij}|_{D^{-1}} \leq 2C\|D^{-1}\|^{\frac{1}{2}}$.*

Proof. The bounds $|v_i| \leq C\mu(D)$ and $|e_{ij}| \leq 2C\mu(D)$ are established in [23, Proposition 4.8 and Theorem 4.11]. Inspecting the proof of these results, one obtains the other announced estimates. Specifically, $|v_i|_D \leq C\|D\|^{\frac{1}{2}}$ is established in the last line of [23, Proposition 4.8]. Using this refined estimate (instead of $|v_i| \leq C\mu(D)$) in the proof of [23, Theorem 4.11] yields $|e_{ij}|_{D^{-1}} \leq 2C\|D^{-1}\|^{\frac{1}{2}}$ (instead of $|e_{ij}| \leq 2C\mu(D)$). The result follows. □

Selling's decomposition of a matrix $D \in S_d^{++}$, $d \in \{2, 3\}$, is obtained by applying Selling's formula Lemma B.2 to a D-obtuse superbase, whose existence is ensured by Selling's algorithm Proposition B.3. This description is constructive and used in all our numerical experiments, since it is efficient enough for the moderately ill-conditioned matrices encountered in our applications. We normalize Selling's decomposition as follows, up to replacing some offsets with their opposites:

$$D = \sum_{e \in \mathcal{Z}^d} \rho(e; D)\, ee^\top, \qquad \text{where } \mathcal{Z}^d := \{e \in \mathbb{Z}^d;\ e \succ_{\text{lex}} 0\}, \tag{91}$$

where $\succ_{\text{lex}}$ stands for the lexicographic ordering. (Note that exactly one of $e \succ_{\text{lex}} 0$ or $-e \succ_{\text{lex}} 0$ holds for each $e \in \mathbb{Z}^d \setminus \{0\}$.) The weights $[\rho(e; D)]_{e \in \mathcal{Z}^d}$ are known as Selling parameters [15], and depend on D but *not* on the choice of D-obtuse superbase, see e.g. [7, Remark 2.13] for a proof. In view of Selling's formula (89), one has $\rho(e; D) = 0$ except for at most $d(d+1)/2$ offsets $e \in \mathcal{Z}^d$. In addition, $\rho(e; D) = 0$ if $|e| > 2C\mu(D)$, by Proposition B.4.

Proposition B.5 (Lipschitz regularity)**.** *For any $e \in \mathcal{Z}^d$, $d \in \{2, 3\}$, the mapping $D \in S_d^{++} \mapsto \rho(e; D)$ is locally Lipschitz with constant $C^2\mu(D)^2$, where C is from Proposition B.4.*

Proof. Let $b = (v_0, \cdots, v_d)$ be a superbase of $\mathbb{Z}^d$, and define $S_b := \{D \in S_d^{++};\ b \text{ is } D\text{-obtuse}\}$. For each $0 \leq i < j \leq d$ let $\tilde{e}_{ij} := \pm e_{ij}$, where the sign is chosen so that $\tilde{e}_{ij} \in \mathcal{Z}^d$. By (89) one has $\rho(D; \tilde{e}_{ij}) = -\langle v_i, Dv_j \rangle$ for all $D \in S_b$, which is a linear function of D with Lipschitz constant at most $|v_i||v_j| \leq C^2\mu(D)^2$ by Proposition B.4. In addition, $\rho(D; e) = 0$ for all $D \in S_b$ and all $e \in \mathcal{Z}^d \setminus \{\tilde{e}_{ij}; 0 \leq i < j \leq d\}$, thus $D \mapsto \rho(e; D)$ is Lipschitz with the announced constant over the set S_b. The announced result follows since S_d^{++} is the union of the closed and convex sets S_b associated to superbases b of $\mathbb{Z}^d$, by Proposition B.3, and since this union is locally finite by Proposition B.4 □

We conclude this appendix by establishing, in Proposition B.8, that some offsets of Selling's decomposition, associated with weights suitably bounded below, span the integer lattice $\mathbb{Z}^d$ by linear combinations with integer coefficients. This implies that the stencils of our numerical

scheme (32) define a locally connected graph, a property used in §4.3 to control its solution in the neighborhood of a point source.

Lemma B.6. *Let $(v_0, \cdots, v_d)$ be a superbase of $\mathbb{Z}^d$, and let $(i_k, j_k)_{k=1}^d$ be such that $0 \leq i_k < j_k \leq d$ for all $0 \leq k \leq d$. Then* $\det(e_{i_1 j_1}, \cdots, e_{i_d j_d}) \in \{-1, 0, 1\}$.

Proof. By Definition B.1, $(v_1, \cdots, v_d)$ is a basis of $\mathbb{Z}^d$. We may thus assume that $(v_1, \cdots, v_d)$ is the canonical basis of $\mathbb{Z}^d$, up to a change of basis, so that $v_0 = (-1, \cdots, -1)^\top$. Then $e_{0j} = -v_j$ for all $1 \leq j \leq d$, and $e_{ij} = v_i - v_j$ for all $1 \leq i < j \leq d$. Each of the vectors e_{ij}, $0 \leq i < j \leq d$, thus features at most once the coefficient 1, and at most once the coefficient -1, the other coefficients being 0. The announced result then follows from [8, Proposition 2.37]. □

Lemma B.7. *Let $D \in S_d^{++}$, and let $e_1, \cdots, e_I \in \mathbb{R}^d$ be such that $D = \sum_{i=1}^I e_i e_i^\top$. Then there exists $1 \leq i_1 < \cdots < i_d \leq I$ s.t. $\sum_{k=1}^d e_{i_k} e_{i_k}^\top \geq cD$, where $c = c(d, I) > 0$.*

Proof. Without loss of generality, up to a linear change of coordinates, one may assume that $D = \mathrm{Id}$ is the $d \times d$ identity matrix. Define

$$\Xi := \big\{(e_i)_{i=1}^I \in (\mathbb{R}^d)^I;\ \sum_{1 \leq i \leq I} v_i v_i^\top = \mathrm{Id}\,\big\}, \qquad \Lambda\big((e_i)_{i=1}^I\big) = \max_{i_1 < \cdots < i_d} \lambda_{\min}\Big(\sum_{1 \leq k \leq d} e_{i_k} e_{i_k}^T\Big),$$

where $\lambda_{\min}$ denotes the smallest eigenvalue. Any family $(e_i)_{i=1}^I \in \Xi$ spans $\mathbb{R}^d$, thus a basis $(e_{i_1}, \cdots, e_{i_d})$ can be extracted from it, and therefore $\Lambda((e_i)_{i=1}^I) \geq \lambda_{\min}(\sum_{k=1}^d e_{i_k} e_{i_k}^\top) > 0$. Denoting by $c(I, d)$ the lower bound of Λ over Ξ, which is positive since Ξ is compact and since Λ is continuous and positive over Ξ, we conclude the proof. □

Proposition B.8 (Spanning property)**.** *For any $D \in S_d^{++}$, $d \in \{2, 3\}$, there exists $e_1, \cdots, e_d \in \mathcal{Z}^d$ such that, for some absolute constant $c > 0$*

$$\det(e_1, \cdots, e_d) = 1, \qquad \min_{1 \leq i \leq d} \rho(e_i; D) \geq c \|D^{-1}\|^{-1}. \tag{92}$$

Proof. From (91) and Lemma B.7 there exists $e_1, \cdots, e_d \in \mathcal{Z}^d$ such that $\sum_{i=1}^d \rho_i e_i e_i^\top \geq cD$, where $\rho_i := \rho(e_i; D)$ and $c = c(d, I) > 0$ is an absolute constant since $d \in \{2, 3\}$ and $I = d(d+1)/2$. Let v be a non-zero vector orthogonal to $e_2, \cdots, e_d$. Then $c|v|_D^2 \leq \rho_1 \langle v, e_1 \rangle^2 \leq \rho_1 |v|_D^2 |e_1|_{D^{-1}}^2 \leq (2C)^2 \rho_1 |v|_D^2 \|D^{-1}\|$ by Proposition B.4. Thus $\rho_1 \geq (c/(2C)^2)\|D^{-1}\|^{-1}$, and likewise for $\rho_2, \cdots, \rho_d$, which concludes the proof. □