# Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem

Jean-David Benamou[1] and Vincent Duval[1]

[1]INRIA Paris (MOKAPLAN) and CEREMADE, CNRS, Université Paris-Dauphine, PSL Research University, 75016 PARIS, FRANCE
{jean-david.benamou,vincent.duval}@inria.fr

October 30, 2017

### Abstract

We present an adaptation of the MA-LBR scheme to the Monge-Ampère equation with second boundary value condition, provided the target is a convex set. This yields a fast adaptive method to numerically solve the Optimal Transport problem between two absolutely continuous measures, the second of which has convex support. The proposed numerical method actually captures a specific Brenier solution which is minimal in some sense. We prove the convergence of the method as the grid stepsize vanishes and we show with numerical experiments that it is able to reproduce subtle properties of the Optimal Transport problem.

## 1    Introduction

Given two bounded open domains $X$ and $Y$ in $\mathbb{R}^2$ and (strictly positive) probability densities $f$ and $g$, defined respectively on $X$ and $Y$, our goal is to numerically solve the quadratic Monge-Kantorovich problem

$$\inf_{\{T_\sharp f = g\}} \int_X \|x - T(x)\|^2 f(x)\, dx \tag{1}$$

where $T : X \to Y$ is a Borel map and $T_\sharp f = g$ is a mass conservation property, that is (see also (5))

$$\int_{T^{-1}(A)} f(x)\, dx = \int_A g(y)\, dy, \text{ for all Borel subset } A \text{ of } Y. \tag{2}$$

This problem has been extensively studied, we refer the reader to the classical monograph of Villani [Vil09] and also the more recent book by Santambrogio [San15] for a comprehensive review of its mathematical theory and applications.

**Numerical approaches to optimal transport.**    From the numerical point of view, the oldest approach is the linear programming formulation of Kantorovich [Kan42] which relaxes the problem in the product space $X \times Y$. This approach however does not scale with the size of the discretization. The so called "Benamou-Brenier" approach [BB00] is based on a different convex relaxation in a time extended space. It is difficult to assess exactly its efficiency but its many numerical implementations suggest it does no better than $O(N^3)$.

Significant progress has been achieved this last decade and new algorithms are now available which can reach almost linear complexity. They can be classified in two groups : first, alternate projection methods (a.k.a. Split Bregman, Sinkhorn, IPFP) [Cut13, Gal16, BCC$^+$15] which are based on the entropic regularization of the Kantorovich problem. These methods are extremely flexible, they apply to a much wider range of Optimal Transport problems and are easy to parallelize. The entropic regularization however blurs the transport map. Decreasing this effect requires more sophisticated tools, see [CPSV16] and references therein.

The second class relies on the Monge-Ampère (MA) interpretation of problem (1). Because the optimal solution satisfies $T = \nabla u$, u convex (see theorem 2.1) below), one seeks to solve

$$\begin{cases} \det(\nabla^2 u) = \dfrac{f}{g \circ \nabla u} & \text{on } X, \\ \nabla u(x) \in \overline{Y} & \text{a.e. } x \in X, \\ u \text{ convex.} \end{cases} \tag{3}$$

Classical solutions of this problem have been studied in [Urb90]. They are unique up to a constant which can be fixed, for example by adding $\int_X u \, dx$ to the equation which has to vanish because of the densities balance. Weak solutions can also be considered either in the Aleksandrov sense or in the regularity framework developed after Cafarelli in the 90s using Brenier weak solutions, that is solutions of (1). *Section 2* briefly reviews these notions as we will work with $\mathscr{C}^1$ solutions.

**Numerical resolution of the Monge-Ampère equation.** Numerical methods based on Monge-Ampère subdivide again in two branches, the semi-discrete approach where $g$ is an empirical measure with a finite number $N$ of Dirac masses [OP88, CP84, Mér11, Lév15a], and finite-difference methods (FD) where $f$ is discretized on a grid of size $N$ [LR05, SAK15, BFO14]. Efficient semi-discrete algorithms rely on the fast computation of the measure of Laguerre cells which correspond to the subgradient of the dual OT map at the Dirac locations. In this paper we focus on the second approach, but we will also use that finite differences solutions yield an approximation of this subgradient at grid points.

The second boundary value conditions (BV2)

$$\nabla u(x) \in \overline{Y} \text{ a.e. } x \in X \tag{4}$$

is a non local condition and a difficulty for the Monge-Ampère finite differences approach. Under a convexity hypothesis for $Y$, an equivalent local non-linear boundary condition is given in [BFO14] which preserves the monotonicity (aka "degenerate ellipticity" after Oberman [Obe06]) of the scheme. In particular a Newton method can be applied for the solution of the discretized system as or periodic BC [LR05, SAK15] or for Dirichlet BC [FO13, Mir15]. We provide in *Section 3* a new interpretation of these BC in term of an infinite domain "minimal" convex extension. It can be used to build the same boundary conditions as in [BFO14], it shows how to extend the FD scheme outside of possiby non convex supports of $f$ and also provides a suitable continuous interpolation tool for the convergence proof in *Section 5*.

**Convergence of the discretizations.** Existing convergence proofs for FD methods in [FO13, Mir15] rely on the viscosity solutions of Crandall and Lions [CIL92] and the abstract convergence theory of Barles and Souganidis [BS91]. This is a powerful framework which can be applied, in particular, to general degenerate elliptic non-linear second order equations. Our problem however has two specificities. First, the Monge-Ampère operator is degenerate elliptic only on the cone of convex function and this constraint must somehow be satisfied by the discretization

2

and preserved in the convergence process. Second, the theory requires a uniqueness principle stating that viscosity sub-solution are below super-solution. The first problem or more generally the approximation of convex functions on a grid, has attracted a lot of attention, see [Mir16a] and the references therein. The Lattice Basis Reduction (LBR) technique in particular was applied to the MA Dirichlet problem in [BCM16]. The second issue (uniquess principle) is more delicate and, even though the BV2 reformulation in [BFO14] clearly belongs to the family of oblique boundary conditions (see [CIL92] for references) there is, to the best of our knowledge, no treatment of the specific (MA-BV2) and convexity constraint in the viscosity theory literature.

We follow a different path in this paper. We build a "minimal" convex extension interpolation of the discrete solutions and interpret it as an Aleksandrov solution of an adapted semi-discrete problem. We can then use classical optimal transport theory to prove convergence of our finite difference discretization. Instead of monotonicity and consistency of the scheme, the proof relies on three ingredients : specific properties of the LBR discretization of the Monge-Ampère operator (*Section 4.5*), the volume conservation enforced by the BV2 boundary condition and the uniqueness and the $\mathscr{C}^1$ regularity of the limit problem. This last condition also requires the convexity of the target $Y$. We borrow here some of the techniques used in [CMOB15] to prove convergence of semi-discrete approximation of JKO gradient steps problems. In the case where $g$ the target density is not constant, we show that simple centered FD is sufficient for the discrete gradient. In summary, we provide a convergent finite difference method for the optimal transport problem which applies to Lebesgue integrable source dentsities $f$ and Lipschitz target densities $g$ with convex support.

**Newton solvers.** A common feature of the semi-discrete and FD approaches is the successful use of a damped Newton method to solve the discrete set of equation. It results in a numerically observed linear complexity of these methods. Mérigot, Kitagawa and Thibert [KMT16] have proven in the semi discrete case that the Jacobian of the discrete non-linear system is strictly positive definite, they show convergence of the Newton method and provide speed convergence. For finite difference similar results are available for the periodic and Dirichlet problems [LR05, FO13, Mir15]. The convergence of the Newton method relies on the invertibility of the Jacobian of the non-linear scheme. It remains open in the BV2 BC case. We provide a numerical study in *Section 6* indicating convergence and that the method has linear complexity.

## 2 The Monge-Ampère problem

### 2.1 Weak solution theory for the Monge-Ampère equation

#### 2.1.1 Brenier Solutions

The connection between the Monge-Ampere equation and the optimal transport problem is made clear by the following theorem.

**Theorem 2.1** ( [Bre91]). *Let $\mu$ and $\nu$ be two compactly supported probability measures on $\mathbb{R}^n$, and assume that $\mu$ and $\nu$ are absolutely continuous with respect to the Lebesgue measure, with respective densities $f$ and $g$. Then,*

(i) *There exists a unique solution $T$ to the optimal transport problem* (1),

(ii) *There exists a convex lower semi-continuous function $u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ such that the optimal map $T$ is given by $T(x) = \nabla u(x)$ for $\mu$-a.e. $x \in \mathbb{R}^n$.*

*Moreover, T is differentiable $\mu$-a.e. and*

$$|\det(\nabla T(x))| = \frac{f(x)}{g(T(x))} \; \text{for } \mu\text{-a.e. } x \in \mathbb{R}^n. \tag{5}$$

Hence, the theorem implies that $u$ is a (possibly weak) solution of (3). We shall call any convex lower semi-continuous (l.s.c.) function $u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ which satisfies (5) and (4) a *Brenier solution to the Monge-Ampère problem* (3).

*Remark* 2.2. In Theorem 2.1, $T = \nabla u$ may be regarded as a map $T : \mathbb{R}^n \to \mathbb{R}^n$ defined almost everywhere. It is actually the uniqueness of the *restriction of T (or $\nabla u$) to* $\text{supp}(f)$ which is asserted. It is obviously possible to modify $T$ (or $u$) on $\mathbb{R}^n \setminus \text{supp}(f)$ without changing the optimality of $T$. In section 3, we use this fact to choose a solution with a "good" behavior outside $X$.

*Remark* 2.3. Given $u$ such that $\nabla u_\sharp \mu = \nu$, the mass equality between $\mu$ and $\nu$ implies that the condition (4) is equivalent to $\nabla u(X) \subseteq \overline{Y}$ up to Lebesgue negligible sets.

### 2.1.2 Aleksandrov Solutions and Semi-Discrete OT

**Definition 2.4** (Aleksandrov solution). Let $\mu$ and $\nu$ be two compactly supported probability measures on $\mathbb{R}^n$, A convex function $u$ is an *Aleksandrov solution* of (3) if for every measurable set $E \in \mathbb{R}^2$,

$$\nu(\partial u(E)) = \mu(E).$$

When $\mu$ and $\nu$ are absolutely continuous with respect to the Lebesgue measure with respective densities $f$ and $g$ as in (2.1) then Aleksandrov and Brenier solutions coincide.

Semi-Discrete Optimal Transport corresponds to the situation when one of the two measures is an empirical measure. If the *target* measure $\nu$ has only atoms, for instance $\nu = \sum_{i=1}^N g_i \, \delta_{y_i}$ is an empirical measure here $(g_i)s$ are positive weigths and $(y_i)s$ the location of dirac masses in the plane, the Brenier map is still well defined and takes values on the finite set $(y_i)$.

Conversely, if the *source* measure is an empirical measure $\mu = \sum_{i=1}^N f_i \, \delta_{x_i}$ then it is not possible to define a Brenier map, but the Aleksandrov solution $u$ still makes sense and satisfies

$$\int_{\partial u(E)} g(y) \, dy = f(E) = \sum_{x_i \in E} f_i \tag{6}$$

or equivalently for all $i$

$$\int_{\partial u(x_i)} g(y) \, dy = f_i \tag{7}$$

The mass concentrated at the Dirac locations is mapped to cells corresponding to the sub-gradients. The Semi-Discrete numerical approach is based on solving system (7) using a Newton method and fast computations of the subgradients $\partial u(x_i)$. These subgradients are known as Laguerre cells tesselation in computational geometry [Mér11, Lév15a]. It should be noted that $u^*$, its Legendre Fenchel transform is still a Brenier solution for the reverse mapping $\nu$ to $\mu$. For more on the duality properties of Semi-Discrete OT see [BFO14].

## 2.2 Cafarelli regularity

Based on Brenier solutions Caffarelli has developed a regularity theory for MA. In particular, when $Y$ is convex, the following result holds :

**Theorem 2.5** ( [Caf92, Caf96]). *Let $X, Y \subseteq \mathbb{R}^n$ be two bounded open sets, $f : X \to \mathbb{R}_+$, $g : Y \to \mathbb{R}_+$ be two probability densities, bounded away from zero and infinity respectively on $X$ and $Y$, i.e.*

$$\exists \lambda > 0, \quad \frac{1}{\lambda} \leq f \leq \lambda \text{ a.e. on } X, \quad \frac{1}{\lambda} \leq g \leq \lambda \text{ a.e. on } Y \tag{8}$$

*and let $T = \nabla u : X \to Y$ be the optimal transport maps sending $f$ to $g$. If $Y$ is convex, then*

*(i) $T \in C_{loc}^{0,\alpha}(X)$ for some $\alpha > 0$.*

*(ii) If in addition $f \in C_{loc}^{k,\beta}(X)$ and $g \in C_{loc}^{k,\beta}(Y)$ for some $\beta \in (0, 1)$, then $T \in C_{loc}^{k+1,\beta}(X)$.*

*(iii) If $f \in C_{loc}^{k,\beta}(\overline{X})$, $g \in C_{loc}^{k,\beta}(\overline{Y})$ and both $X$ and $Y$ are smooth and uniformly convex, then $T : \overline{X} \to \overline{Y}$ is a global diffeomorphism of class $C^{k+1,\beta}$.*

In the following, we shall assume that $f$ and $g$ are bounded away from zero and infinity. The first conclusion of the theorem thus implies that $u \in C_{loc}^{1,\alpha}(X)$.

We finally recall a slightly more general regularity result due to Figalli and Loeper wich does not require lower bounds on the source density $f$ and which holds in the plane ($n = 2$) :

**Theorem 2.6** ( [FL09, Theorem 2.1]). *Let $X, Y \subseteq \mathbb{R}^2$ be two bounded open sets, $Y$ convex, $f : X \to \mathbb{R}_+$, $g : Y \to \mathbb{R}_+$ be two probability densities, such that there exist $\lambda > 0$ with $f \leq \frac{1}{\lambda}$ in $\mathbb{R}^2$ and $\lambda \leq g$ in $Y$. Let $u : \mathbb{R}^2 \to \mathbb{R}$ be a Brenier solution to*

$$\begin{cases} \det(\nabla^2 u) = \dfrac{f}{g \circ \nabla u} & \text{on } X, \\ \det(\nabla^2 u) = 0 & \text{on } \mathbb{R}^2 \setminus \overline{X}, \\ \nabla u(x) \in \overline{Y} & \text{a.e. } x \in X, \\ u \text{ convex.} \end{cases} \tag{9}$$

*Then $u \in \mathscr{C}^1(\mathbb{R}^2)$.*

# 3   Minimal Brenier Solutions in $\mathbb{R}^2$

## 3.1   Definition

From now on, we consider $X \subseteq \mathbb{R}^2$, $Y \subseteq \mathbb{R}^2$ two bounded open domains, and we assume that $Y$ is convex. We assume that the probability densities $f$, $g$ are bounded away from zero and infinity respectively on $X$ and $Y$ (see (8)). We consider a specific Brenier solution to the Monge-Ampère problem which is characterized by the following Proposition.

**Proposition 3.1.** *Assume that $Y$ is convex. Then there is a unique (up to an additive constant) convex lower semi-continuous function $\tilde{u} : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ which is a Brenier solution to (3) and which satisfies $\partial \tilde{u}(\mathbb{R}^2) \subseteq \overline{Y}$.*
*Moreover, $\tilde{u} \in \mathscr{C}^1(\mathbb{R}^2)$.*

*Remark* 3.2 (Uniqueness). As the Brenier map is unique, note that uniqueness of the potential $\tilde{u}$ up to a constant carries over to the other notions of solution recalled in Section 2.

*Proof.* We first prove uniqueness. Let $u_1$, $u_2$ be two such functions. Then the gradients of their Legendre-Fenchel conjugates $u_1^*$ and $u_2^*$ solve the optimal transport problem from $\nu$ to $\mu$, and by Theorem 2.1, $\nabla u_1^*(y) = \nabla u_2^*(y)$ for a.e. $y \in \overline{Y}$. As a result, there is some $C \in \mathbb{R}$ such that those two convex functions satisfy $u_1^*(y) = u_2^*(y) + C$ for all $y \in Y$. Let us prove that the equality also holds on $\partial Y$. Let $y_0 \in \partial Y$, $y_1 \in Y$. By lower semi-continuity and convexity

$$
\begin{aligned}
u_1^*(y_0) &\leq \liminf_{\substack{y \to y_0 \\ y \in Y}} u_1^*(y) \\
&= \liminf_{\substack{y \to y_0 \\ y \in Y}} (u_2^*(y) + C) \\
&\leq \liminf_{t \to 0^+} ((1-t)u_2^*(y_0) + tu_2^*(y_1)) + C = u_2^*(y_0) + C.
\end{aligned}
$$

The converse inequality is obtained by swapping the role of $u_1^*$ and $u_2^*$, and the equality is proved.

Now, since $u_1$, $u_2$ are proper convex lower semi-continuous, we know from [Roc83] that

$$
\forall i \in \{1, 2\}, \qquad \mathrm{ri}\,(\mathrm{dom}\,u_i^*) \subseteq \mathrm{dom}\,(\partial u_i^*) \subseteq \mathrm{dom}\,u_i^*.
$$

Since $Y \subseteq \mathrm{ri}(\mathrm{dom}\,u_i^*)$ and $\mathrm{dom}\,(\partial u_i^*) = \partial u_i(\mathbb{R}^2) \subseteq \overline{Y}$, we get $\mathrm{dom}\,u_i^* = \overline{Y}$. The double conjugate reconstruction formula yields

$$
\forall x \in \mathbb{R}^2, \quad u_1(x) = \sup_{y \in \overline{Y}} (\langle y, x \rangle - u_1^*(y)) = \sup_{y \in \overline{Y}} (\langle y, x \rangle - (u_2^*(y) + C)) = u_2(x) - C.
$$

To prove the existence of $\tilde{u}$, let $u_0$ be a Brenier solution to the Monge-Ampère problem (3). We define

$$
\forall x \in \mathbb{R}^2, \quad \tilde{u}(x) \overset{\text{def.}}{=} \sup_{y \in \overline{Y}} \left\{ \langle x, y \rangle - u_0^*(y) \right\}, \tag{10}
$$

where $u_0^*$ is the Legendre-Fenchel conjugate of $u_0$.

$$
\forall y \in \overline{Y}, \quad u_0^*(y) \overset{\text{def.}}{=} \sup_{x \in \mathbb{R}^2} \left\{ \langle x, y \rangle - u_0(x) \right\}. \tag{11}
$$

It is immediate that $\tilde{u}$ is convex lower semi-continuous. From Theorem 2.5, we know that $u_0$ is a convex function which is $\mathscr{C}_{loc}^{1,\alpha}$ in $X$. Moreover, $u_0$ is continuous up to $\partial X$ since, for all $x \in X$, $\nabla u_0(x) \in Y$ which is a bounded set. As the supremum of a (finite) upper semi-continuous function on the compact set $\overline{Y}$, $\tilde{u}$ is finite on $\mathbb{R}^2$.

Now, we prove that $\tilde{u}(x) = u_0(x)$ for all $x \in X$. Since $u_0$ is proper convex l.s.c.,

$$
u_0(x) = \langle x, \nabla u_0(x) \rangle - u_0^*(\nabla u_0(x)) = \sup_{y \in \mathbb{R}^2} (\langle x, y \rangle - u_0^*(y)) \geq \sup_{y \in \overline{Y}} (\langle x, y \rangle - u_0^*(y)) = \tilde{u}(x).
$$

But since $\nabla u_0(x) \in Y$, $\langle x, \nabla u_0(x) \rangle - u_0^*(\nabla u_0(x)) \leq \sup_{y \in \overline{Y}} (\langle x, y \rangle - u_0^*(y))$, and the above inequality is in fact an equality.

To prove that $\partial \tilde{u}(\mathbb{R}^2) \subseteq \overline{Y}$, let $x \in \mathbb{R}^2$, $p \in \partial \tilde{u}(x)$. Since for all $h \in \mathbb{R}^2$, $\tilde{u}(x + h) \geq \tilde{u}(x) + \langle p, h \rangle$,

$$
\sup_{y' \in \overline{Y}} (\langle x, y' \rangle - u_0^*(y')) + \sup_{y \in \overline{Y}} \langle h, y \rangle \geq \sup_{y' \in \overline{Y}} (\langle x + h, y' \rangle - u_0^*(y')) \geq \sup_{y' \in \overline{Y}} (\langle x, y' \rangle - u_0^*(y')) + \langle h, p \rangle.
$$

As a result, $\sup_{y \in \overline{Y}} \langle h, y \rangle \geq \langle h, p \rangle$ for all $h \in \mathbb{R}^2$, hence $p \in \overline{Y}$, and $\partial \tilde{u}(\mathbb{R}^2) \subseteq \overline{Y}$.

We conclude that $\tilde{u}$ is also a Brenier solution to (3), and by Theorem 2.6, we deduce that $\tilde{u}$ is $\mathscr{C}^1(\mathbb{R}^2)$. □

6

*Remark* 3.3. From (10), one may observe that $\tilde{u}(x) \leq u_0(x)$ for all $x \in \mathbb{R}^2$. It is the *minimal convex extension* of $u_0$ outside $X$ in the sense that it is the smallest convex function defined on $\mathbb{R}^2$ which coincides with $u_0$ on $X$. Additionally, it is minimal among all Brenier solutions in the sense that its subdifferential is the smallest possible.

*Remark* 3.4. Standard arguments from convex analysis show that any $y_0 \in \overline{Y}$ is a subgradient (hence the gradient) of $\tilde{u}$ at $x \in \mathbb{R}^2$ if and only if $\langle y_0, x \rangle - u_0^*(y_0) = \tilde{u}(x)$.

## 3.2 The affine ray property

The aim of this section is to give some insight on the behavior of $\tilde{u}$ outside $X$, which helps motivate the discrete scheme of Section 4. In the following, $\overline{\text{co}}(X)$ denotes the closed convex hull of $X$.

**Proposition 3.5.** *The function $\tilde{u}$ has the following properties.*

(i) *For all $x \in \mathbb{R}^2 \setminus \overline{\text{co}}(X)$, there exists $x_0 \in \partial\overline{\text{co}}(X)$ such that $\tilde{u}$ is affine on the half-line $\{x_0 + t(x - x_0) ; t \geq 0\}$. Moreover, $\nabla\tilde{u}(x) \in \text{argmax}_{y \in \overline{Y}}\langle y, x - x_0 \rangle \subseteq \partial Y$.*

(ii) *For all $x \in \overline{\text{co}}(X) \setminus \overline{X}$, there exists $x_0 \in \partial X$ such that $\tilde{u}$ is affine on the line segment $[x, x_0]$.*

*Remark* 3.6. The condition $\nabla\tilde{u}(x) \in \text{argmax}_{y \in \overline{Y}}\langle y, x - x_0 \rangle$ actually means that $x - x_0$ is in $\mathcal{N}_{\overline{Y}}(y)$, the normal cone to $\overline{Y}$ at $y = \nabla\tilde{u}(x)$. In fact, all the points $x'$ such that $x' - x_0 \in \mathcal{N}_{\overline{Y}}(y)$ are mapped to the same gradient value $y = \nabla\tilde{u}(x)$, and $\tilde{u}$ is affine on that set.

Figure 1 provides an illustration of Proposition 3.5 and Remark 3.6. The connected blue regions in the source space represent points which are mapped to the same gradient value. Such regions are invariant by translation by $\mathcal{N}_{\overline{Y}}(y)$. In $\overline{\text{co}}(X) \setminus \overline{X}$ (here $X$ is not convex) points are mapped into the interior of $Y$, onto a line segment (black dashed line) which is a discontinuity set for the gradient of the inverse optimal transport map. Gradients are also constant on some convex sets corresponding to the subgradients of the inverse optimal map on this line segment. These regions are naturally connected to $X$. The behavior of optimal transport maps in the case of non-convex supports is analysed in depth in [Fig09].

*Proof.* In view of the uniqueness (up to a constant) stated in Proposition 3.1, we may assume without loss of generality that the function $u_0$ used in the construction of $\tilde{u}$ (see (10)) is convex lower semi-continuous and satisfies

$$\forall x \in \mathbb{R}^2 \setminus \overline{\text{co}} X, \quad u_0(x) = +\infty, \tag{12}$$

as this does not change its being a Brenier solution.

Let $y_0 \in \overline{Y}$ be a slope. It is optimal for $x \in \mathbb{R}^2$ in (10) iff

$$x \in \partial\left(u_0^* + \chi_{\overline{Y}}\right)(y_0) = \partial u_0^*(y_0) + \mathcal{N}_{\overline{Y}}(y_0) = (\partial u_0)^{-1}(y_0) + \mathcal{N}_{\overline{Y}}(y_0), \tag{13}$$

as there is a point $y' \in Y$ where $\chi_{\overline{Y}}$ is continuous and $u_0^*$ is finite. In the above equation, $\chi_{\overline{Y}}$ and $\mathcal{N}_{\overline{Y}}$ respectively stand for the characteristic function and the normal cone of $\overline{Y}$ at $y_0$,

$$\chi_{\overline{Y}}(y) = \begin{cases} 0 & \text{if } y \in \overline{Y}, \\ +\infty & \text{otherwise,} \end{cases}, \qquad \mathcal{N}_{\overline{Y}}(y_0) = \left\{x' \in \mathbb{R}^2 ; \forall y \in \overline{Y}, \langle x', y - y_0 \rangle \leq 0\right\}. \tag{14}$$

Equation (13) is equivalent to the existence of some $x_0 \in (\partial u_0)^{-1}(y_0) \subseteq \overline{\text{co}} X$ such that $\langle x - x_0, y - y_0 \rangle \leq 0$ for all $y \in \overline{Y}$ (or equivalently $y_0 \in \text{argmax}_{y \in \overline{Y}}\langle y, x - x_0 \rangle$). Clearly, if $y_0$ is optimal for $x$, it is also optimal for $x + t(x - x_0)$ for $t > -1$.
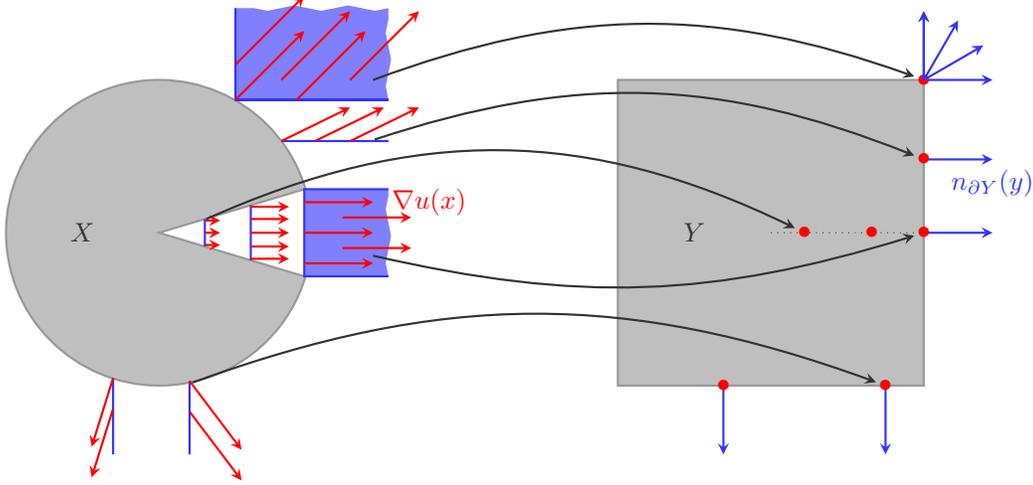
Figure 1: Illustration of Proposition 3.5 showing the correspondence between gradients (red) and their support (blue) in source space and target space. The connected blue regions in the source space represent points which are mapped to the same gradient value (in red).

Moreover, sets of the form (13) cover the whole space $\mathbb{R}^2$, since by compactness and semi-continuity, there always exists an optimal $y_0$ for (13). Incidentally that slope is in fact $\nabla \tilde{u}(x)$ since, provided $y_0$ is optimal for $x$,

$$\forall e \in \mathbb{R}^2, \forall t > 0, \quad \tilde{u}(x + te) - \tilde{u}(x) \geq \langle x + te, y_0 \rangle - u_0^*(y_0) - (\langle x, y_0 \rangle - u_0^*(y_0)),$$
$$\text{hence} \quad t\langle e, \nabla \tilde{u}(x) \rangle + o(t) \geq t\langle e, y_0 \rangle.$$

Dividing by $t \to 0^+$ yields $y_0 = \nabla \tilde{u}(x)$.

To summarize, we have proved $(i)$ since $x - x_0 \neq 0$ for all $x \in \mathbb{R}^2 \backslash \overline{\text{co}}(X)$, hence $\{x + t(x - x_0) \; ; \; t > 0\}$ does indeed define a half line. As for $(ii)$, $\nabla \tilde{u}(x) \in \overline{Y}$, hence there exists $x_0 \in (\partial u_0)^{-1}(\nabla \tilde{u}(x)) \cap \overline{X}$, so that $\nabla \tilde{u}(x_0) = \nabla \tilde{u}(x)$ (and since $\{x' \; ; \; \nabla \tilde{u}(x') = \nabla \tilde{u}(x)\}$ is convex, it is not restrictive to assume $x_0 \in \partial X$). $\qquad \square$

*Remark* 3.7. Another point of view, using standard tools of convex analysis (see e.g. [Roc83, Th. 16.4]) is to interpret $\tilde{u}$ as an infimal convolution

$$\forall x \in \mathbb{R}^2, \quad \tilde{u}(x) = \inf \left\{ u_0(x') + \sigma_{\overline{Y}}(x - x') \; ; \; x' \in \text{co}\, \overline{X} \right\}. \tag{15}$$

where $\sigma_{\overline{Y}} : x \mapsto \sup_{y \in \overline{Y}} \langle y, x \rangle$ is the support function of $\overline{Y}$.

## 3.3   Discussion

From Proposition 3.5, we see that the function $\tilde{u}$ defined in proposition 3.1 formally satisfies the equations

$$\det(D^2 u)\, g(\nabla u) = f \text{ on } X, \tag{16}$$
$$\det(D^2 u) = 0 \text{ on } \mathbb{R}^2 \setminus X, \tag{17}$$
$$\min_{e \in \mathbb{S}^1} \langle (D^2 u)e, e \rangle = 0 \text{ on } \mathbb{R}^2 \setminus X, \tag{18}$$
$$\sup_{e \in \mathbb{S}^1} \{\langle \nabla u, e \rangle - \sigma_{\overline{Y}}(e)\} = 0 \text{ on } \mathbb{R}^2 \setminus \overline{\text{co}}(X). \tag{19}$$

8

Indeed, (16) is the Monge-Ampère equation (see Theorem 2.1), (17) and (18) follow from the affine property in Proposition 3.5. If $\tilde{u}$ is smooth and the minimum eigenvalue of $D^2 u$ is null, this also enforces convexity (see [Obe08] for the connection with the convexe enveloppe problem). As for (19), since $Y$ is convex, the inequality $\sup_{e \in \mathbb{S}^1}\{\langle \nabla u, e \rangle - \sigma_{\overline{Y}}(e)\} \leq 0$ is an equivalent formulation of the BV2 boundary condition $\nabla \tilde{u}(\mathbb{R}^2) \subseteq \overline{Y}$. The equality actually means that $\nabla \tilde{u}(x) \in \partial Y$ for $x \in \mathbb{R}^2 \setminus \overline{\mathrm{co}}(X)$.

The discretization strategy is presented in Section 4 and then the convergence proof in Section 5. The convergence will hold in the Aleksandrov/Brenier setting but the limit solution regularity itself will depend on the the regularity of $f$ and $g$.

*Remark* 3.8 (Uniqueness). It is not difficult to show that $\tilde{u}$ is a viscosity solutions of equations (16-19) but it is is much harder to prove uniqueness for this class of equations, see [CIL92] and its references to oblique boundary conditions . However, $\tilde{u}$ coincide with the unique Brenier solution on $X$ and the $\mathbb{R}^2$ extension is also unique. More precisely (see remark 3.2) $\tilde{u}$ is unique up to a constant.

# 4   Finite Difference Discretization

This section explains how our scheme is built from the set of the equations of Section 3.3 and discuss the properties of the resulting discrete system.

We will consider a sequence of discretization steps $(h_n)_{n \in \mathbb{N}}$, $h_n > 0$, $h_n \searrow 0^+$, and we define an infinite lattice of points $\mathcal{G}_n \overset{\text{def.}}{=} h_n \mathbb{Z}^2$. We work in a compact square domain $D \subseteq \mathbb{R}^2$ (say $D = [-1, 1]^2$) which contains $\overline{X}$ in its interior. We assume without loss of generality that $0 \in X$. A discrete solution $U_n \in \mathbb{R}^{\mathrm{card}(\mathcal{G}_n \cap D)}$ is defined on that grid : if $u$ is a continuous solution of our problem, its discrete interpolant on the grid is $U_n[x] = u(x)$ for all $x \in \mathcal{G}_n \cap D$.

We will use the following finite differences formulae in each grid direction $e$,

$$\delta_e^{h_n} U_n[x] \overset{\text{def.}}{=} U_n[x + h_n\, e] - U_n[x]$$

and

$$\Delta_e^{h_n} U_n[x] \overset{\text{def.}}{=} \delta_e^{h_n} U_n[x] + \delta_{-e}^{h_n} U_n[x].$$

## 4.1   Discretization of the target $Y$

As it is defined on a finite grid, our discrete scheme is only able to estimate the directional gradient in a finite number of directions. Hence, the constraint we can impose in practice when discretizing (19) is that the gradient belongs to a polygonal approximation $Y_n$ to $\overline{Y}$. More precisely given a finite family of irreducible vectors $\mathcal{V}_n \subseteq \mathbb{Z}^2 \setminus \{0\}$, we consider the corresponding set

$$Y_n \overset{\text{def.}}{=} \left\{ y \in \mathbb{R}^2 \; ; \; \forall e \in \mathcal{V}_n, \; \langle y, e \rangle \leq \sigma_{\overline{Y}}(e) \right\} \tag{20}$$

which is nonempty, closed and convex. We assume that

1. $\mathcal{V}_n$ contains $\{(0, 1), (1, 0), (0, -1), (-1, 0)\}$ (so that $Y_n$ is compact)

2. $\mathcal{V}_n \subseteq \mathcal{V}_{n'}$ for $n' \leq n$ (so that $\overline{Y} \subseteq Y_{n'} \subseteq Y_n$),

3. $\bigcap_{n \in \mathbb{N}} Y_n = \overline{Y}$.

4. The following inclusion holds for $n$ large enough

$$\overline{X} + h_n \mathcal{V}_n \subseteq D. \tag{21}$$

The third point holds as soon as $\overline{Y}$ is defined by a finite number of inequalities of the form (20) (in which case the constraint $\nabla u(x) \in \overline{Y}$ is exactly imposed) or provided that $\{e/|e| \; ; \; e \in \bigcup_{n \in \mathbb{N}} \mathcal{V}_n\}$ is dense in $\mathbb{S}^1$. The fourth point will be useful in Proposition 4.3 to ensure that the variations of $U_n$ inside $\overline{X}$ are bounded by $\sigma_{Y_n}$.

A general strategy to ensure the above four assumptions is to choose

$$\mathcal{V}_n = \left\{ e \in \mathbb{Z}^2 \; ; \; e \text{ irreducible and } \|e\|_\infty \leq \rho_n \right\} \tag{22}$$

where $\rho_n \to +\infty$ and $\rho_n/h_n \to 0$ as $n \to +\infty$ (for instance $\rho_n = \sqrt{h_n}$). The result of this approximation strategy is illustrated in Figure 2 for three values $\rho_n \in \{1, 2, 3\}$. It appears that $\rho_n = 3$ already yields a sharp polygonal approximation of the considered ellipse.

In any case, the first three points above ensure that $Y_n$ converges towards $\overline{Y}$ in the sense of the Hausdorff topology [Sch93, Lemma 1.8.1], that is

$$\lim_{n \to +\infty} \max \left( \sup_{y \in \overline{Y}} d(y, Y_n), \sup_{y \in Y_n} d(y, \overline{Y}) \right) = 0. \tag{23}$$
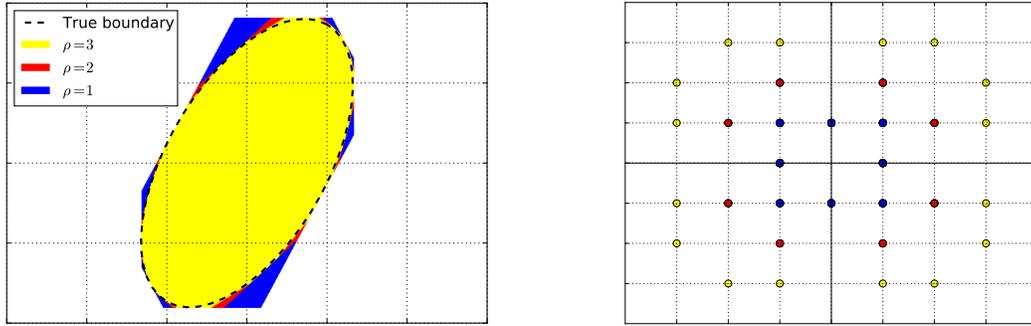


Figure 2: Left: Approximation of an ellipse using (20) for the values $\rho_n \in \{1, 2, 3\}$ (the true boundary is shown in dashed red line). The ellipse is obtained by the rotation with angle $\pi/3$ of $\left\{(y_1, y_2) \in \mathbb{R}^2 \; ; \; \frac{1}{a^2} y_1^2 + \frac{1}{b^2} y_2^2 = 1\right\}$ with $a = 2$, $b = 1$. Right: the corresponding irreducible vectors used in $\mathcal{V}_n$.

## 4.2 Discretization of the Monge-Ampère operator in $X$

We will use the (MA-LBR) scheme for the Monge-Ampère operator $\det(D^2 u)$ discretization [BCM16]. It relies on the notion of superbase :

**Definition 4.1.** A basis of $\mathbb{Z}^2$ is a pair $(e', e'') \in (\mathbb{Z}^2)^2$ such that $|\det(e', e'')| = 1$.
A superbase of $\mathbb{Z}^2$ is a triplet $(e, e', e'') \in (\mathbb{Z}^2)^3$ such that $e + e' + e'' = 0$, and $(e', e'')$ is a basis of $\mathbb{Z}^2$.

The finite difference MA-LBR operator, is a consistent and monotone approximation of the Monge-Ampère operator given by

$$\mathrm{MA}_n(U_n)[x] \stackrel{\text{def.}}{=} \frac{1}{h_n^4} \min_{\substack{(e, e', e'') \in (\mathbb{Z}^2)^3 \\ \text{superbase}}} \mathbf{h}(\Delta_e^{h_n^+} U_n[x], \Delta_{e'}^{h_n^+} U_n[x], \Delta_{e''}^{h_n^+} U_n[x]) \tag{24}$$

10

where $\Delta_e^{h_n+}U_n[x] = \max(\Delta_e^{h_n}U_n[x], 0)$ and for $a, b, c \in \mathbb{R}_+$ we define

$$\mathbf{h}(a, b, c) := \begin{cases} bc \text{ if } a \geq b+c, \text{ and likewise permuting } a, b, c, \\ \frac{1}{2}(ab + bc + ca) - \frac{1}{4}(a^2 + b^2 + c^2) \text{ otherwise.} \end{cases} \qquad (25)$$

Note that the minimum in (24) is in fact restricted to superbases $(e, e', e'')$ such that

$$\{x, x \pm h_n e, x \pm h_n e', x \pm h_n e''\} \subseteq D \qquad (26)$$

(that is, both $x + h_n e$ and $x - h_n e$ belong to $D$ and similarly for $e'$ and $e''$). The number $N$ of such superbases is *a priori* very large, but the adaptive algorithm proposed in [BCM16] yields a dramatic speed-up as it is asymptotically sufficient to test $\log(N)$ superbases out of $N$. We use this refinement in our simulations and we refer to [BCM16] for the detail of the adaptive algorithm. Also, it is shown in [BCM16] that the largest width of the vectors in the optimal superbase grows with the condition number of the Hessian of $u$. In practice, it is therefore possible to limit the minimization to a stencil defined by its width, i.e. vectors on the grid $\mathbb{Z} \times \mathbb{Z}$ with a given maximum norm.

The scheme consistency is remarkable. Given a quadratic form $u(x) = \frac{1}{2}\langle Mx, x \rangle$, where $M$ a strictly positive definite matrix with condition number $\kappa$, its grid interpolation $U_n$ satisfies

$$\mathrm{MA}_n(U_n)[x] = \det(M)$$

provided that $\{x, x \pm h_n e, x \pm h_n e', x \pm h_n e''\} \subseteq D$ for all irreducible vectors $e \in \mathbb{Z}^2$ (and the corresponding $e'$ and $e''$) such that $\|e\|_2 \leq \kappa\sqrt{2}$ (see [BCM16]).

The MA-LBR operator also provides interesting "discrete" convexity properties (see [BCM16] and proposition A.3 in [Mir16a]. That notion is used in Appendix A to study the finite difference approximation of the gradient.

Finally the following property states that the MA-LBR operator overestimates the subgradient of the convex envelope of $U_n$ at grid points. The proof was communicated to us by J.-M Mirebeau, and we reproduce it in Appendix B with his kind permission. This result will be useful in the convergence proof.

**Lemma 4.2.** *Let* $\underline{u} : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ *be the largest convex lower semi-continuous function which minorizes* $U_n[x]$ *at all points* $x \in \mathcal{G}_n \cap D$. *Then,*

$$\forall x \in \mathcal{G}_n \cap X, \quad |\partial \underline{u}(x)| \leq \mathrm{MA}_n(U_n)[x]\omega_n, \qquad (27)$$

*where* $\omega_n = h_n^2$ *is the area of one cell of the grid.*

## 4.3 Discretization in of the Monge-Ampère operator and the BV2 conditions in $D \setminus X$

The MA-LBR scheme is only suitable to discretize strictly convex functions. When at least one eigenvalue decreases to 0, the stencil width (see above) becomes infinite. To capture the flat behavior of solutions in $\mathbb{R}^2 \setminus X$, we need to add more and more directions to the minimization.

Instead, we apply the Wide-Stencil (WS) formulation proposed by Oberman in [Obe08] to discretize (18), that is, the minimum eigenvalue of the Hessian should be 0. This simply yields the scheme

$$\mathrm{MA}_n^0(U_n)[x] \stackrel{\text{def.}}{=} \min_{e \in V(x)} \Delta_e^{h_n} U_n[x], \qquad (28)$$

where $V(x)$ denotes the set of irreducible vectors $e \in \mathbb{Z}^2 \setminus \{0\}$ such that $\{x - e, x, x + e\} \subseteq D$

11

Additionally, we take advantage of the points in $D \setminus \overline{X}$ to impose the (BV2) boundary condition, by modifying the scheme (28) as follows :

$$\widetilde{\mathrm{MA}_n^0}(U_n)[x] \stackrel{\text{def.}}{=} \min_{e \in V(x) \cup \mathcal{V}_n} \tilde{\Delta}_e^{h_n} \tag{29}$$

where for all $x \in \mathcal{G}_n \cap D$,

$$\tilde{\Delta}_e^{h_n} U_n[x] \stackrel{\text{def.}}{=} \tilde{\delta}_e^{h_n} U_n[x] + \tilde{\delta}_{-e}^{h_n} U_n[x],$$

$$\text{and } \tilde{\delta}_e^{h_n} U_n[x] \stackrel{\text{def.}}{=} \begin{cases} U_n[x + h_n e] - U_n[x] & \text{if } x + he \in D, \\ \sigma_{\overline{Y}}(h_n e) & \text{otherwise.} \end{cases}$$

The rationale of that scheme comes from (18) and (19): for fixed $e \in \mathbb{R}^2$, imposing $\tilde{\Delta}_e^{h_n} U_n[x] = 0$ is consistent with

$$\langle D^2 u(x)e, e \rangle = 0 \text{ if } x \in \mathrm{int}(D) \setminus X,$$

$$\langle \nabla u(x), e \rangle = \sigma_{\overline{Y}}(e) \text{ if } x \in \partial D \text{ and } e \text{ points outwards } D.$$

The formal consistency with with (18) and (19) is straightforward. In pratice, the same stencil can be used for $V(x)$, i.e. discretization of the degenerate Monge-Ampère operator and $\mathcal{V}_n$, i.e. the discretization of the target geometry.

## 4.4  Gradient Approximation

Except when $g$ is the constant density on $Y$, one needs to discretize the gradient $\nabla u$ in order to discretize the Monge-Ampère equation (16).

In Section 5, we prove the convergence of the scheme as $n \to +\infty$. The main assumption to obtain this convergence is that the discrete gradient $\mathcal{D}_n^1 U_n$ satisfies the following uniform convergence

$$\lim_{n \to +\infty} \sup_{x \in \overline{X} \cap \mathcal{G}_n} \sup_{y \in \partial \tilde{u}_n(x)} \left| \mathcal{D}_n^1 U_n[x] - y \right| = 0, \tag{30}$$

where $\tilde{u}_n$ is the continuous interpolation of $U_n$ defined in the next Section (Eq. (41)). Provided a solution $U_n$ to the scheme exists, Theorem 5.5 then ensures that its interpolation $\tilde{u}_n$ converges towards the minimal Brenier solution $\tilde{u}$ such that $\tilde{u}(0) = 0$.

In particular, we prove in Appendix A that this property holds for

- the centered finite difference on the cartesian grid,

$$\mathcal{D}_n^1 U_n[x] \stackrel{\text{def.}}{=} \frac{1}{2h_n} \begin{pmatrix} \delta_{(1,0)}^{h_n} U_n[x] - \delta_{(-1,0)}^{h_n} U_n[x] \\ \delta_{(0,1)}^{h_n} U_n[x] - \delta_{(0,-1)}^{h_n} U_n[x] \end{pmatrix}, \tag{31}$$

- the forward and backward finite differences on the cartesian grid,

$$\mathcal{D}_n^1 U_n[x] \stackrel{\text{def.}}{=} \frac{1}{h_n} \begin{pmatrix} \delta_{(1,0)}^{h_n} U_n[x] \\ \delta_{(0,1)}^{h_n} U_n[x] \end{pmatrix}, \quad \mathcal{D}_n^1 U_n[x] \stackrel{\text{def.}}{=} \frac{1}{h_n} \begin{pmatrix} -\delta_{(-1,0)}^{h_n} U_n[x] \\ -\delta_{(0,-1)}^{h_n} U_n[x] \end{pmatrix}. \tag{32}$$

12

## 4.5 Summary of the scheme and property of the discrete system

Finally, for fixed $h \in (0, h_0]$, we plan on computing $U_n$ solution of

$$\forall x \in \mathcal{G}_n \cap D, \quad \begin{cases} \mathrm{MA}_n(U_n)[x] - \frac{\bar{f}_n[x]}{g((\mathcal{D}_n^1 U_n)[x])} = 0 & \text{if } x \in X \\ \widetilde{\mathrm{MA}_n^0}(U_n)[x] = 0 & \text{otherwise} \\ U_n[0] = 0 \end{cases} \tag{33}$$

where $\bar{f}_n[x] \stackrel{\text{def.}}{=} \frac{1}{h_n^2} \int_{[-h_n/2, h_n/2]^2} f(x+t) \mathrm{d}t$ is a local average of the density $f$. The added scalar equation $U_n[0] = 0$ fixes the constant.

Now, we list several properties of the scheme which will be useful for the proof of convergence.

**Proposition 4.3.** *If $U_n$ is a solution to (33), then*

1. *For all $x \in \mathcal{G}_n \cap X$, and all $e$ irreducible such that $\{x+h_n e, x, x-h_n e\} \subseteq D$, $\Delta_e^{h_n} U_n[x] > 0$.*

2. *For all $x \in \mathcal{G}_n \cap (D \setminus X)$, and all $e \in \mathcal{V}_h$ such that $\{x+h_n e, x, x-h_n e\} \subseteq D$, $\Delta_e^{h_n} U_n[x] \geq 0$.*

3. *For all $x \in \mathcal{G}_n \cap D$, $e \in \mathcal{V}_n$ and $(k, \ell) \in \mathbb{N}^2$, such that $k \leq \ell$,*

$$-\sigma_{\overline{Y}}(-h_n e) \leq U_n[x + (k+1)h_n e] - U_n[x + kh_n e] \\ \leq U_n[x + (\ell+1)h_n e] - U_n[x + \ell h_n e] \leq \sigma_{\overline{Y}}(h_n e) \tag{34}$$

   *whenever $x + i h_n e \in D$ for $i \in \{k, k+1, \ell, \ell+1\}$.*

4. *If $x \in D$ and $e \in \mathcal{V}_n$ irreducible are such that $U_n[x + h_n e] - U_n[x] = \sigma_{\overline{Y}}(h_n e)$, then $U_n[x + k h_n e] = U_n[x] + k \sigma_{\overline{Y}}(h_n e)$ for all integer $k \geq 0$ such that $x + k h_n e \in D$.*

5. *There exists $C > 0$ such that for all $x, x' \in D$,*

$$\left| U_n[x] - U_n[x'] \right| \leq C \left\| x - x' \right\|_1. \tag{35}$$

*Proof.* The first point follows from $\mathrm{MA}_n(U_n)[x] = \frac{\bar{f}_n(x)}{g((\mathcal{D}_n^1 U_n)[x])} > 0$ and the inequality $h(a, b, c) \leq \min\{ab, bc, ca\}$ (see [BCM16]). As for the second, point it follows immediately from the definition of $\tilde{\Delta}_e^{h_n} U_n$ and the scheme $D \setminus X$.

To prove the third point, let us first assume that $\ell \in \mathbb{Z}^2$ is such that $\{x + \ell h_n e, x + (\ell+1)h_n e\} \subseteq D$ but $x + (\ell+2)h_n e \notin D$. By (21) we deduce that $x + (\ell+1)h_n e \in D \setminus \overline{X}$, so that

$$0 \leq \tilde{\Delta}_e^{h_n} U_n[x + (\ell+1)h_n e] = U_n[x + \ell h_n e] - U_n[x + (\ell+1)h_n e] + \sigma_{\overline{Y}}(h_n e),$$

which yields $U_n[x + (\ell+1)h_n e] - U_n[x + \ell h_n e] \leq \sigma_{\overline{Y}}(h_n e)$. Similarly $-\sigma_{\overline{Y}}(-h_n e) \leq U_n[x + (k+1)h_n e] - U_n[x + kh_n e]$ if $x + (k-1)h_n e \notin D$ but $\{x + kh_n e, x + (k+1)h_n e\} \subseteq D$. The other intermediate inequalities follow from $\Delta_e^{h_n} U_n[x + i h_n e] \geq 0$.

The fourth point is a consequence of (34).

Now, we deal with the last point. Write $x - x' = khe_1 + \ell h e_2$ where $e_1 \stackrel{\text{def.}}{=} (1, 0) \in \mathcal{V}_n$ and $e_2 \stackrel{\text{def.}}{=} (0, 1) \in \mathcal{V}_n$. Applying (34), we get

$$U_n[x] - U_n[x'] \leq h \left( k^+ \sigma_{\overline{Y}}(e_1) + (-k)^+ \sigma_{\overline{Y}}(-e_1) + \ell^+ \sigma_{\overline{Y}}(e_2) + (-\ell)^+ \sigma_{\overline{Y}}(-e_2) \right),$$
$$\text{and} \quad U_n[x'] - U_n[x] \leq h \left( k^+ \sigma_{\overline{Y}}(-e_1) + (-k)^+ \sigma_{\overline{Y}}(e_1) + \ell^+ \sigma_{\overline{Y}}(-e_2) + (-\ell)^+ \sigma_{\overline{Y}}(e_2) \right),$$

where, as before, $k^+ \stackrel{\text{def.}}{=} \max(k, 0)$. Hence (35) holds with $C = \max\{\sigma_{\overline{Y}}(e_1), \sigma_{\overline{Y}}(-e_1), \sigma_{\overline{Y}}(e_2), \sigma_{\overline{Y}}(-e_2)\}$. $\qquad \square$

**Degenerate Ellipticity.** In [Obe06], a theory of convergence is built for non-linear schemes based on viscosity theory using a modified version of monotonicity called Degenerate Ellipticity (DE). To the best of our knowledge, there are only two instances of DE scheme for the Monge-Ampère operator : the Wide-Stencil and MA-LBR schemes.

We discuss here the potential application of this theory to our scheme.
An abstract scheme is represented at each point $x \in \mathcal{G}_n \cap D$ by an equation of the form

$$S(x, U_n[x], \{\delta_{y-x} U_n[x]\}_{y \in \mathcal{G}_n \cap D \setminus \{x\}}) = 0. \tag{36}$$

**Definition 4.4.** The scheme $S$ is Degenerate Elliptic (DE) if for all $x \in \mathcal{G}_n \cap D$, $S(x, \cdot, \cdot)$ is nonincreasing in its first variable, and nondecreasing in any other variable.

A DE Lipschitz and proper scheme is well defined and has a unique solution.

These properties can be checked on (33) when $g$ is the uniform density. Otherwise and even when $g$ is uniformly Lipschitz, the gradient approximation we propose is not degenerate elliptic. A fix proposed by Froese and Oberman [FO13] is to remark that a uniform positive bound on the Monge-Ampère operator is sufficient to dominate the DE defect potentially induced by the gradient discretization. Unfortunately this bound is not straightforward to establish. In practice, Froese and Oberman impose it by truncating their scheme. We found that our scheme works just as is but we have no further theoretical results.

**Newton solver.** As in [BCM16, BFO14, Mir15, LR05, Mér11, Lév15b] we use a damped Newton algorithm to solve the system (33). For the description of the algorithm and proof of its convergence we refer to the paper above. The crucial ingredient is to prove global convergence of Newton iterates is the invertibility of the Jacobian matrix of the system. See for example [Mir15] where the case of Dirichlet boundary conditions is treated.

In our case there the invertibilty of the Jacobian remains an open problem, regarding :

- The proposed discretization of the BV2 condition.

- The non constant $g$ density case

In practice and in the implementation, the system (33) is unchanged but for the residual inversion in the Newton procedure we use an inexact Jacobian wich forces positive definiteness as follows.
For each line $x$ of the Jacobian : set $dF_1[x] = \partial_{q_1} F(x, D_C^h U_n[x])$ and $dF_2[x] = \partial_{q_2} F(x, D_C^h U_n[x])$ and add coefficients of an upwind type discretization on the classical five point stencil :

$$
\begin{aligned}
\text{At column } x + (1,0)\, h: \quad & Gw = -\max(0, dF_1[x]) \\
\text{At column } x + (-1,0)\, h: \quad & Ge = \min(0, dF_1[x]) \\
\text{At column } x + (0,1)\, h: \quad & Gn = -\max(0, dF_2[x]) \\
\text{At column } x + (0,-1)\, h: \quad & Gs = \min(0, dF_2[x]) \\
\text{On the diagonal}: \quad & Gn + Gs + Gw + Ge
\end{aligned}
\tag{37}
$$

Finally, $g$ needs to be defined globally in case the iterate generate gradients outside of the target $Y$.

# 5 Convergence

We now assume that for all discretization steps $(h_n)_{n \in \mathbb{N}}$, $h_n > 0$, $h_n \searrow 0^+$, we have a unique well-defined solution $(U_n[x])_{x \in \mathcal{G}_n \cap D}$ and proceed to show that it converges to the Brenier solution of the problem.

## 5.1 Convex extension as an interpolation and its properties

Let us recall that we assume that $g$ is continuous on $\overline{Y}$, whereas $f$ is only Lebesgue integrable. We will assume that there exists $(\alpha_f, \beta_f, \alpha_g, \beta_g) \in (0, +\infty)^4$, such that

$$\forall x \in \overline{X}, \ \alpha_f \leq f(x) \leq \beta_f, \tag{38}$$

$$\forall y \in \overline{Y}, \ \alpha_g \leq g(y) \leq \beta_g. \tag{39}$$

In fact, possibly changing $g$ in $\mathbb{R}^2 \setminus \overline{Y}$, it is not restrictive to assume that $g$ is continuous on $Y_0 \supseteq \overline{Y}$ and that (39) holds for $y \in Y_0$.

From the values $(U_n[x])_{x \in \mathcal{G}_n \cap D}$ of the discrete problem (see Section 4.5), we build the following function $\tilde{u}_n : \mathbb{R}^2 \to \mathbb{R}$, with

$$\tilde{u}_n(x) \overset{\text{def.}}{=} \sup \left\{ L(x) \ ; \ L : \mathbb{R}^2 \to \mathbb{R} \text{ is affine, } \nabla L \in Y_n, \text{ and } \forall x' \in \mathcal{G}_n \cap D, L(x') \leq U_n[x'] \right\}. \tag{40}$$

Equivalently,

$$\forall x \in \mathbb{R}^2, \quad \tilde{u}_n(x) \overset{\text{def.}}{=} \sup_{y \in Y_n} \left( \langle x, y \rangle - U_n^\star(y) \right) \tag{41}$$

$$\text{where} \quad U_n^\star(y) \overset{\text{def.}}{=} \sup_{x \in \mathcal{G}_n \cap D} \left( \langle y, x \rangle - U_n[x] \right). \tag{42}$$

The following proposition gathers some properties which typically hold with such constructions (see also [CMOB15]).

**Proposition 5.1.** *The following properties hold.*

1. *$\tilde{u}_n$ is convex, finite-valued, and by construction $\tilde{u}_n(x) \leq U_n[x]$ for all $x \in \mathcal{G}_n \cap D$.*

2. *$\partial \tilde{u}_n(\mathbb{R}^2) = \partial \tilde{u}_n(\mathcal{G}_n \cap D) = Y_n$.*

3. *Let $x \in \mathcal{G}_n \cap D$. If $\tilde{u}_n(x) < U_n[x]$, then $|\partial \tilde{u}_n(x)| = 0$.*

*Proof.* The first point is left to the reader. To prove the second one, consider any slope $y \in Y_n$. There is some $x \in \mathcal{G}_n \cap D$ which minimizes $x' \mapsto U_n[x'] - \langle y, x' \rangle$ over $\mathcal{G}_n \cap D$ (let $\alpha$ denote the corresponding minimum). Then the affine function $L : x' \mapsto \langle y, x' \rangle + \alpha$ satisfies $L(x') \leq U_n[x']$ for all $x' \in \mathcal{G}_n \cap D$, and $L(x) = U_n(x)$. Hence, by construction of $\tilde{u}_n$, $L(x') \leq \tilde{u}_n(x')$ for all $x' \in \mathbb{R}^2$, and $L(x) = \tilde{u}_n(x)$, which means that $y \in \partial \tilde{u}_n(x)$. As a result $Y_n \subseteq \partial \tilde{u}_n(\mathcal{G}_n \cap D)$. The fact that $\partial \tilde{u}_n(\mathbb{R}^2) \subseteq Y_n$ follows from $\partial \tilde{u}_n(\mathbb{R}^2) \subseteq \text{dom}(\tilde{u}_n^*) = \text{dom}(U_n^\star + \chi_{Y_n})$.

Now, let us prove the third point. Assume by contradiction that $|\partial \tilde{u}_n(x)| > 0$. Since $\partial \tilde{u}_n(x)$ is convex, it must have nonempty interior, hence there is some $y_0 \in Y_n$, $r > 0$ such that $B(y_0, r) \subseteq \partial \tilde{u}_n(x)$. For all $x' \in \mathbb{R}^2$, all $p \in B(y_0, r)$, $\tilde{u}_n(x') \geq \tilde{u}_n(x) + \langle p, x' - x \rangle$ hence

$$\tilde{u}_n(x') \geq \tilde{u}_n(x) + \langle y_0, x' - x \rangle + r \, \|x' - x\| .$$

For $x' \in \mathcal{G}_n \cap D \setminus \{x\}$, $U_n[x'] \geq \tilde{u}_n(x')$ and $\|x' - x\| \geq h$. Hence, for $\alpha \in (0, 1]$ small enough, the affine function $L : x' \mapsto \tilde{u}_n(x) + \langle y_0, x' - x \rangle + \alpha r h$ satisfies $L(x') \leq U_n[x']$ for all $x' \in \mathcal{G}_n \cap D$, and $\tilde{u}_n(x) < L(x)$ which contradicts the definition of $\tilde{u}_n$. $\square$

15

The next Lemma describes more specific properties of $\tilde{u}_n$ which follow from the construction of $U_n$.

**Lemma 5.2.** *The family $\{\tilde{u}_n\}_{n\in\mathbb{N}} \subseteq \mathscr{C}(\mathbb{R}^2)$ is relatively compact for the topology of uniform convergence on compact sets. Moreover,*

$$\forall x \in \mathcal{G}_n \cap X, \quad |\partial\tilde{u}_n(x)| \le \mathrm{MA}_n(U_n)[x], \tag{43}$$

$$\forall x \in \mathcal{G}_n \cap (D \setminus X), \quad |\partial\tilde{u}_n(x)| = 0. \tag{44}$$

*Proof.* We observe that $(\tilde{u}_n)_{n\in\mathbb{N}}$ is uniformly equicontinuous by Proposition 5.1 and the fact that $Y_n \subseteq Y_0$. Moreover, from Proposition 4.3 and $U_n(0) = 0$, we deduce that

$$\forall n \in \mathbb{N}, \forall x \in \mathcal{G}_n \cap D, \quad |U_n[x]| \le C \sup_{x' \in D} \|x'\|_1.$$

As a result, $\tilde{u}_n(0) \in [-C\sup_{x'\in D}\|x'\|_1, 0]$. We deduce the claimed compactness by applying the Ascoli-Arzelà theorem.

The inequality (43) follows from Lemma 4.2. Indeed, let $\underline{u} : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ be the largest convex l.s.c. function which minorizes $U_n[x]$ at all points $x \in \mathcal{G}_n \cap D$. If $\tilde{u}_n(x) < U_n[x]$, then $|\partial\tilde{u}_n(x)| = 0$ and there is nothing to prove. Otherwise, $\tilde{u}_n(x) = \underline{u}(x) = U_n[x]$ and $\tilde{u}_n \le \underline{u}$ imply that $\partial\tilde{u}_n(x) \subseteq \partial\underline{u}(x)$. As a result,

$$\forall x \in \mathcal{G}_n \cap X, \quad |\partial\tilde{u}_n(x)| \le |\partial\underline{u}(x)| \le \mathrm{MA}_n(U_n)[x].$$

Now, we prove (44). Let $x \in \mathcal{G}_n \cap (D \setminus X)$. Again, if $\tilde{u}_n(x) < U_n[x]$, then $|\partial\tilde{u}_n(x)| = 0$, so we assume that $\tilde{u}_n(x) = U_n[x]$. Assuming by contradiction that $|\partial\tilde{u}_n(x)| > 0$ there must exist again some $y_0 \in Y_n$, $r > 0$ such that $B(y_0, r) \subseteq Y$, and $\tilde{u}_n(x') \ge \tilde{u}_n(x) + \langle y_0, x' - x \rangle + r\|x' - x\|$.

Since $\min_{e\in\mathcal{V}_n} \tilde{\Delta}_e U_n[x] = 0$, there exists $e \in \mathbb{Z}^2 \setminus \{0\}$ such that $\tilde{\Delta}_e U_n[x] = 0$. If $\{x - h_n e, x, x + h_n e\} \subseteq D$, then

$$0 = (U_n[x + h_n e] - U_n[x]) + (U_n[x - h_n e] - U_n[x]) \ge (\tilde{u}_n(x + h_n e) - \tilde{u}_n(x)) + (\tilde{u}_n(x - h_n e) - \tilde{u}_n(x))$$
$$\ge 2rh_n\|e\| > 0,$$

which is impossible. On the other hand, if $x + h_n e \notin D$ (the case $x - h_n e \notin D$ is similar), then $x - h_n e \in D$ and $U_n[x - h_n e] - U_n[x] = -h_n\sigma_{\overline{Y}}(e)$. The slope of $\tilde{u}_n$ is monotone in the direction $e$, and it cannot exceed $\sigma_{\overline{Y}}(e/|e|)$, hence $\tilde{u}_n(x + te) = \tilde{u}_n(x) + t\sigma_{\overline{Y}}(e)$ for $t \ge -h_n$. But this contradicts the inequality

$$\tilde{u}_n(x + te) \ge \tilde{u}_n(x) + t\langle y_0, e \rangle + r|t|\|e\|.$$

We conclude that $|\partial\tilde{u}_n(x)| = 0$. $\qquad\square$

## 5.2 A semi-discrete optimal transport problem

Let us define the following measures

$$\mu_n \stackrel{\text{def.}}{=} \sum_{x\in\mathcal{G}_n\cap D} F_h[x]\delta_x, \quad \nu_n \stackrel{\text{def.}}{=} \sum_{x\in\mathcal{G}_n\cap D} g(\mathcal{D}_n^1 U_n[x])\mathbb{1}_{\partial\tilde{u}_n(x)}\mathcal{L}^2, \tag{45}$$

where $F_h[x]$ is defined for all $x \in \mathcal{G}_n \cap D$ as

$$F_h[x] \stackrel{\text{def.}}{=} g(\mathcal{D}_n^1 U_n[x])|\partial\tilde{u}_n(x)| \le g(\mathcal{D}_n^1 U_n[x])\mathrm{MA}_n(U_n)[x]\omega_n = \bar{f}_n(x)\omega_n, \tag{46}$$

16

and $\omega_n = h_n^2$ is the area of one cell of the grid.

Incidentally, let us note that from Proposition 5.1 and Lemma 5.2, the Monge-Ampère measure (see [Gut01]) associated with $\tilde{u}_n$ satisfies

$$\det(D^2\tilde{u}_n) \stackrel{\text{def.}}{=} |\partial\tilde{u}_n| = \sum_{x\in\mathcal{G}_n\cap X} \frac{F_h[x]}{g(\mathcal{D}_n^1 U_n[x])} \delta_x. \tag{47}$$

**Lemma 5.3.** *There exists $v \in \mathscr{C}^1(\mathbb{R}^2)$, and non-negative measures $\hat{\mu} \in \mathcal{M}(\mathbb{R}^2)$, $\hat{\nu} \in \mathcal{M}(\mathbb{R}^2)$, such that*

- *up to the extraction of a subsequence, the measures $\mu_n$ and $\nu_n$ respectively weakly converge to $\hat{\mu}$ and $\hat{\nu}$ as $h_n \to 0^+$,*

- *$\hat{\mu}(\mathbb{R}^2) = \hat{\nu}(\mathbb{R}^2)$ and $\nabla v$ is the optimal transport of $\hat{\mu}$ to $\hat{\nu}$.*

*Proof.* Let us consider the transport plan $\gamma_n \in \mathcal{M}(\mathbb{R}^2 \times \mathbb{R}^2)$

$$\gamma_n \stackrel{\text{def.}}{=} \sum_{x\in\mathcal{G}_n\cap D} g(\mathcal{D}_n^1 U_n[x])\delta_x \otimes \mathbb{1}_{\partial\tilde{u}_n(x)}\mathcal{L}^2$$

We note that $\gamma_n$ is an optimal transport plan between $\mu_n$ and $\nu_n$ since its support is contained in the graph of $\partial\tilde{u}_n$ [San15].

Moreover, we observe that

$$\mu_n \le \omega_n \sum_{x\in X\cap\mathcal{G}_n} \bar{f}_n(x)\delta_x = \sum_{x\in X\cap\mathcal{G}_n} \left(\int_{[x-h_n/2,x+h_n/2]^2} f(x+t)\mathrm{d}t\right)\delta_x, \tag{48}$$

where the right hand-side converges towards $\mu$ in the weak-* topology (*i.e.* induced by compactly supported continuous test functions). Additionally, since

$$\nu_n \le \beta_g \mathcal{L}^2 \llcorner Y_n, \tag{49}$$

we deduce that the supports of $\mu_n$, $\nu_n$, $\gamma_n$ are respectively contained in the compact sets $\overline{X}$, $Y_n$, $\overline{X} \times Y_n$, and their masses are uniformly bounded. Hence, there exists Radon measures $\hat{\mu}$, $\hat{\nu} \in \mathcal{M}(\mathbb{R}^2)$, $\hat{\gamma} \in \mathcal{M}(\mathbb{R}^2 \times \mathbb{R}^2)$ such that, up to the extraction of a (not relabeled) subsequence, $\mu_n$, $\nu_n$ and $\gamma_n$ respectively converge to $\hat{\mu}$, $\hat{\nu}$, $\hat{\gamma}$ in the weak-* topology. As for $\tilde{u}_n$, we already know from Lemma 5.2 that we may extract an additional subsequence so that $\tilde{u}_n$ converges uniformly on compact sets to some (convex function) $v \in \mathscr{C}(\mathbb{R}^2)$. We also note that $\hat{\gamma}$ has respective marginals $\hat{\mu}$ and $\hat{\nu}$, and $\hat{\mu}$, $\hat{\nu}$ have densities with respect to the Lebesgue measure $\hat{f}$, $\hat{g}$ which satisfy

$$\hat{f} \le f \quad \text{and} \quad \hat{g} \le \beta_g \mathbb{1}_{\overline{Y}}.$$

Let us note that $g(\mathcal{D}_n^1 U_n[x]) \ge \alpha_g$, hence

$$\nu_n \ge \alpha_g \sum_{x\in\mathcal{G}_n\cap D} (\mathbb{1}_{\partial\tilde{u}_n(x)}\mathcal{L}^2) = \alpha_g\mathcal{L}^2 \llcorner Y_n,$$

thus $\mu_n(\mathbb{R}^2) = \nu_n(\mathbb{R}^2) \ge \alpha_g |Y|$, and in the limit, $\hat{\mu}(\mathbb{R}^2) = \hat{\nu}(\mathbb{R}^2) \ge \alpha_g |Y| > 0$.

Moreover, any element of $\operatorname{supp}\hat{\gamma}$ is the limit of (a subsequence of) elements $(x_n, y_n)$ of $\operatorname{supp}\gamma_n \subseteq \partial\tilde{u}_n$. Passing to the limit in the subdifferential inequality

$$\forall x' \in \mathbb{R}^2, \quad \tilde{u}_n(x') \ge \tilde{u}_n(x_n) + \langle y_n, x' - x_n \rangle,$$

17

we obtain
$$\operatorname{supp}\hat{\gamma} \subseteq \operatorname{graph}\partial v.$$

As a result, and since $\hat{\mu}$ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^2$, $\hat{\gamma} = (I \times \nabla v)_{\sharp}\hat{\mu}$ and $\nabla v$ is the optimal transport map between $\hat{\mu}$ and $\hat{\nu}$.

Additionally, passing to the limit in (47) with $\frac{F_h[x]}{g(\mathcal{D}_n^1 U_n[x])} \leq \frac{\beta_f}{\alpha_g}\omega_n$, we get

$$\det(D^2 v) \leq \frac{\beta_f}{\alpha_g}\mathbb{1}_{\overline{X}}, \quad \text{and} \quad \partial v(\mathbb{R}^2) = \overline{Y}.$$

and from [FL09, Theorem 2.1], we deduce that $v \in \mathscr{C}^1(\mathbb{R}^2)$. $\qquad\square$

**Lemma 5.4.** *Assume that* (30) *holds. Then, with the notations of Lemma 5.3, $\hat{\mu} = \mu$ and $\hat{\nu} = \nu$.*

*Proof.* For all $y \in \overline{Y} \subseteq Y_n$, there exists $x \in \mathcal{G}_n \cap D$ such that $y \in \partial\tilde{u}_n(x)$. Except for a set of $y$ with zero Lebesgue measure, that $x$ is unique [Gut01, Lemma 1.1.12], and $x \in \overline{X}$ (see Lemma 5.2). Then, denoting by $\omega_g$ the modulus of continuity of $g$ over $Y_0$, we get

$$\left|g(\mathcal{D}_n^1 U_n[x]) - g(y)\right| \leq \omega_g\left(\max_{x' \in \mathcal{G}_n \cap \overline{X}} \max_{y' \in \partial\tilde{u}_n(x'_n)} \left|\mathcal{D}_n^1 U_n[x'] - y'\right|\right) \longrightarrow 0$$

as $n \to +\infty$, where the convergence of the right-hand side follows from (30).

As a result the following convergence holds

$$\lim_{n \to +\infty} \left\|\sum_{x \in \mathcal{G}_n \cap D} g(\mathcal{D}_n^1 U_n[x])\mathbb{1}_{\partial\tilde{u}_n(x)} - g\right\|_{L^\infty(\overline{Y})} = 0. \tag{50}$$

We deduce that $\hat{g} \equiv g$ on $\overline{Y}$ and from (50) and (23) we obtain

$$\begin{aligned}
\int_{\mathbb{R}^2} \hat{f}(x)\mathrm{d}x &= \lim_{n \to +\infty} \mu_n(\mathbb{R}^2) \\
&= \lim_{n \to +\infty} \nu_n(\mathbb{R}^2) \\
&= \lim_{n \to +\infty} \int_{Y_n} \left(\sum_{x \in \mathcal{G}_n \cap D} g(\mathcal{D}_n^1 U_n[x])\mathbb{1}_{\partial\tilde{u}_n(x)}(y)\right) \mathrm{d}y \\
&= \int_{\overline{Y}} g(y)\mathrm{d}y = \int_{\mathbb{R}^2} f(x)\mathrm{d}x.
\end{aligned}$$

Since, $\hat{f} \leq f$, we get $f = \hat{f}$. $\qquad\square$

To sum up, we are now in position to prove the following result.

**Theorem 5.5.** *As $n \to +\infty$, the function $\tilde{u}_n$ converges uniformly on compact sets towards the unique minimal Brenier solution $\tilde{u}$ which satisfies $\tilde{u}(0) = 0$.*

*Proof.* Up to the extraction of a sequence, we obtain from Lemma 5.3 and 5.4 that $\tilde{u}_n$ converges uniformly on compact sets towards some convex function $v \in \mathscr{C}^1(\mathbb{R}^2)$ which satisfies in the Brenier sense

$$\det(D^2 v) = f(x)/g(\nabla v(x)), \quad \text{and} \quad \partial v(\mathbb{R}^2) = \overline{Y}.$$

However, the function $v$ may depend on the choice of the subsequence. By Proposition 3.1 which states the uniqueness of the minimal Brenier solution up to an additive constant, we can prove that $v = \tilde{u}$ by proving that $v(0) = 0$. Then we obtain that the full family $(\tilde{u}_n)_{n\in\mathbb{N}}$ converges towards $\tilde{u}$ by uniqueness of the cluster point.

We know that $v$ is a solution to the Monge-Ampère equation, hence (in our setting) also in the Aleksandrov sense. As a result, for all $r > 0$, $\int_{B(0,r)} f \mathrm{d}\mathcal{L}^2 = \int_{\partial v(B(0,r))} g \mathrm{d}\mathcal{L}^2$, and

$$0 < \alpha_f \, |B(0,r)| \le \beta_g \, |\partial v(B(0,r))| \,.$$

But by weak-convergence of the Monge-Ampère measures [Gut01, Lemma 1.2.2],

$$|\partial v(B(0,r))| \le \liminf_{n\to+\infty} |\partial \tilde{u}_n \left( B(0,r) \right)| = \liminf_{n\to+\infty} |\partial \tilde{u}_n \left( B(0,r) \cap \mathcal{G}_n \right)| \tag{51}$$

As a result, for all $n$ large enough, there exists $x_{n,r} \in B(0,r) \cap \mathcal{G}_n$ such that $|\partial \tilde{u}_n(x_{n,r})| > 0$, and thus, by Proposition 5.1, $\tilde{u}_n(x_{n,r}) = U_n[x_{n,r}]$. By a diagonal argument we construct a sequence $x_n^*$ such that $x_n^* \to 0$ and $\tilde{u}_n(x_n^*) = U_n[x_n^*]$. By Proposition 4.3

$$|\tilde{u}_n(x_n^*)| = |U_n[x_n^*]| \le C \, \|x_n^*\|_1 \,. \tag{52}$$

Passing to the limit $n \to +\infty$, we get $v(0) = 0$. As a result $v = \tilde{u}$ and the full family converges towards $\tilde{u}$. $\qquad\square$

# 6 Numerical study

The numerical method, as we implemented it, depends on two main parameters :

- $h$ the step size of the cartesian grid discretizing the square $D$ wich contains the support of $f$.

- The stencil width, i.e. the maximun norm of the vectors on the grid $\mathbb{Z}^2$ which will be used in the variational schemes. More precisely we use superbases with vector with maximum norm given by the stencil width in (24), we use the same set of vectors for $V(x)$ in (24) and finally we also use the same vectors to define $\mathcal{V}_n$ in (20) for the discretization of the target.

There are also parameters linked to the precision of the damped Newton algorithm, but they have a very limited impact on the efficiency of the method.

We run the Newton method until the residual is stationary in $\infty$ norm which usually takes a few dozen interations. The value of the residual depends on the discretization parameter but in all of our experiments was always between 1e-10 and 1e-15.

The target density $g$ is defined in all $\mathbb{R}^2$ using a constant extension. In practice we initialized with a potential $u$ which gradient maps inside the target $Y$ but it cannot be excluded to hit out along the iterations.

## 6.1 Experiment with $h$

A standard test is to optimally map a uniform density square to a uniform density circle. We use here the a stencil width of 5 (we recall that means that we will use vectors on the grid $\mathbb{Z}^2$ of
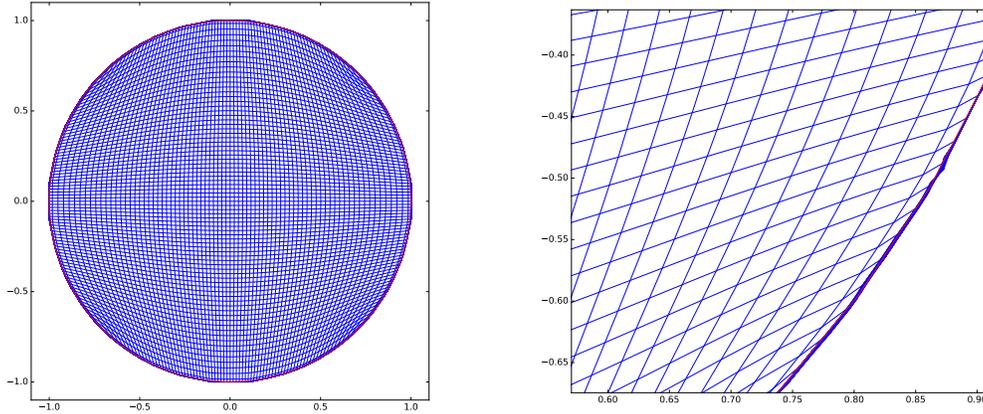
Figure 3: Left: Map deformation of the cartesian grid for square to ball test case. Right: Zoom of the Map deformation of the cartesian grid (as expected, $D \setminus X$ is mapped to $\partial Y$).

maximum $\ell^\infty$ norm 5).

$X$ is a square $[0.2, 0.8]^2$ immersed in a bigger square $D = [0, 1]^2$.

Table 1 shows, for different values of $h$, the number of iterations the norms of the reached residuals and the value of the $u(x_0)$, which is not exactly 0 as the geometry of the $Y$ is not exactly preserved by the discretization.

Figure 3 shows the deformation of the computational grid under the gradient map. Notice that the grid volume is well preserved and that all grid points in $D \setminus X$ are collapsed onto the boundary of $Y$ which is approximated as a polygon (see next section).

| $N = \frac{1}{h}$ | # It. | $\|Res.\|_\infty$ | $\|Res.\|_2$ | $u(x_0)$ |
|---|---|---|---|---|
| 128 | 32 | 4e-12 | 7e-13 | 1.4e-03 |
| 256 | 72 | 3e-12 | 3e-11 | 4.5e-02 |
| 512 | 103 | 1e-10 | 1e-11 | 1.7e-02 |
| 1024 | 209 | 3e-10 | 5e-11 | 1.1e-02 |

Table 1: Number of iterations / Norms of the reached residuals / Value of forced constant

On a standard Laptop a non optimized Julia implementation needs less than 5 minutes to solve the 512 case and 1.5 hours for the $1024 \times 1024$ case.

## 6.2 Experiments with stencil width

In all this section $N = 128$, and we vary the stencil width, meaning the number of vector on the grid used in scheme. The target is an heptagone whose normals directions are not necessarily vectors in the stencil.

Table 2 shows, for increasing values of the stencil width , the number of iterations the norms of the reached residuals and the value of the $u(x_0)$, which decreases as the accuracy of the discretization of $Y$ improves.

Figure 4 to 6 shows the deformation of the computational grid under the gradient map and zooms. They show how the geometry of the computed target improves with the domain discretization which also depends on the stencil width. Notice again that grid volume is well preserved and that all grid points in $D \setminus X$ are collapsed onto the boudary of $Y$.

| S. width | # It. | $\|Res.\|_\infty$ | $\|Res.\|_2$ | $u(x_0)$ |
|---|---|---|---|---|
| 2 | 32 | 9e-12 | 1.5e-12 | 1.5e-01 |
| 4 | 42 | 9e-12 | 6e-12 | 6.4e-02 |
| 8 | 39 | 8e-12 | 4e-12 | 3.9e-02 |
| 16 | 42 | 4e-12 | 7e-13 | 2.2e-02 |

Table 2: Stationary residual reached in $\ell^\infty$ and $\ell^2$ norms / Number of iterations to reach stationary minimal residual / Value of forced constant
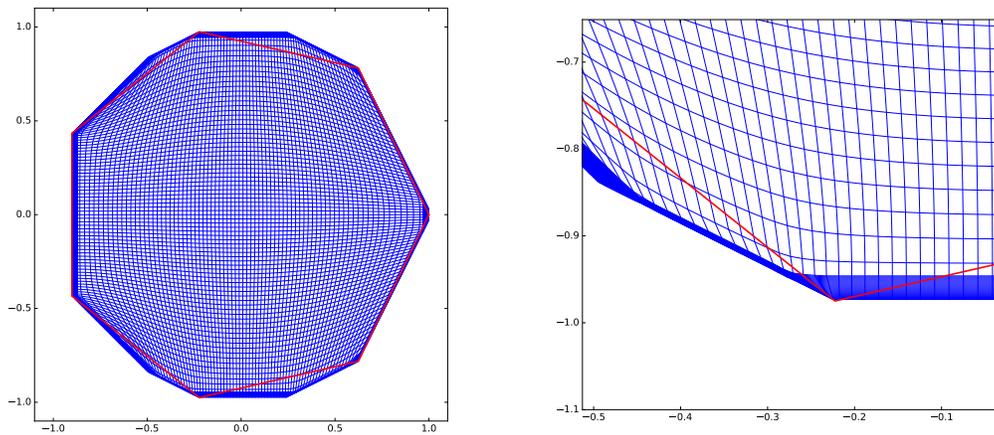


Figure 4: Left: Mapping a square to heptagon, Stencil width = 2. Right: zoom.
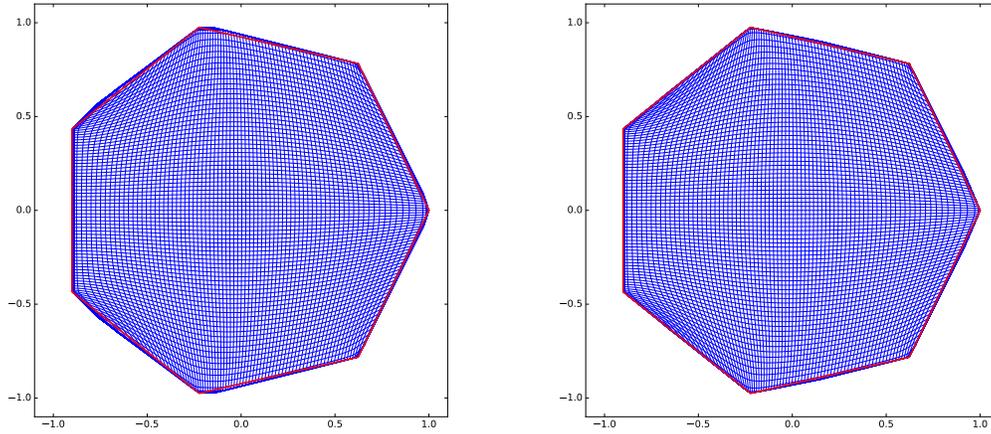
21

Figure 5: Mapping a square to heptagon. Left: stencil width = 4. Right: stencil width = 8.

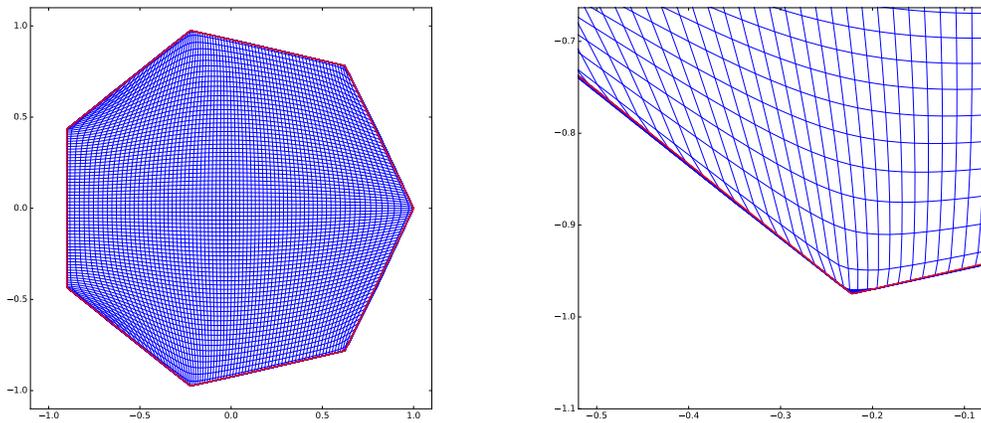

Figure 6: Left: Mapping a square to heptagon, Stencil width = 16. Right: zoom.

## 6.3 Experiments with inhomogeneous source density $f$

In this section we show results obtained with heterogeneous densities $f$ mapped to a constant density ball $Y$, $N = 128$ and stencil width is 5 . If the source density is given analytically and not as grid values, the convergence proof holds assuming that the grid value discretization is computed as average of $f$ on grid cells.

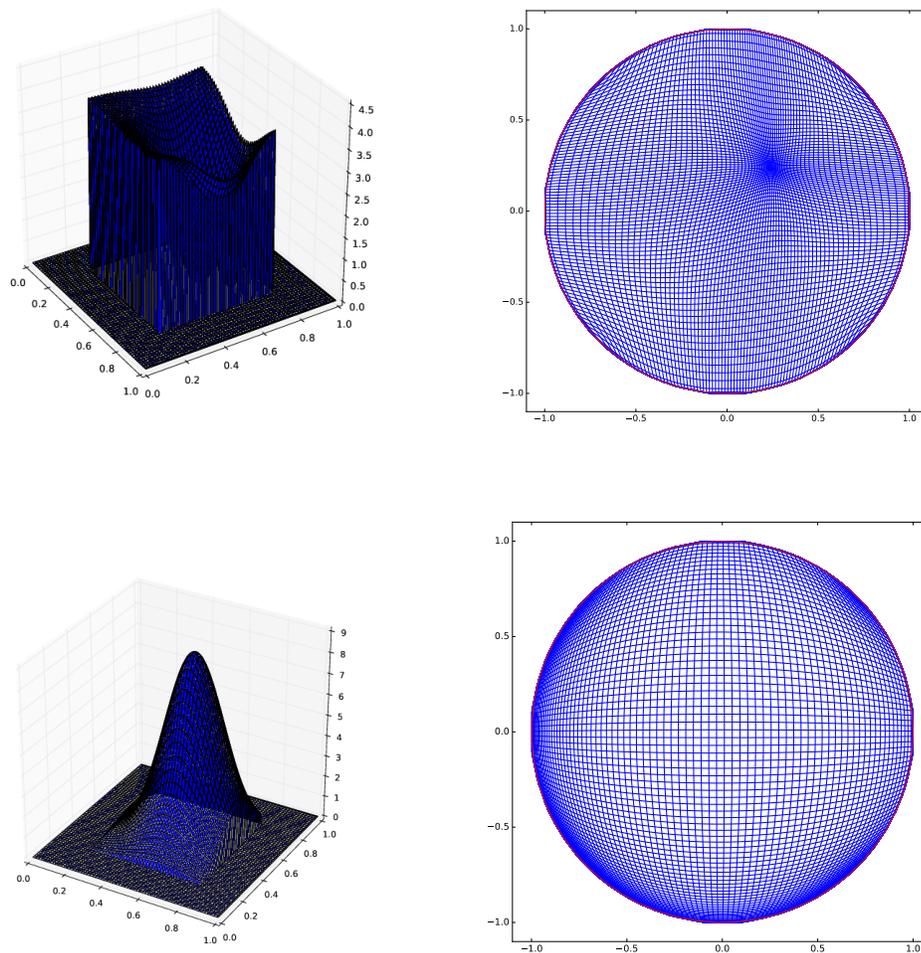Figures 7 and 8 show heterogeneous sources and the corresponding deformation of the cartesian grid.



Figure 7: Left column: Source density $f$. Right column: Corresponding optimal map deformation of the grid.

In Figure 8 the source density is random on the square. This test case demonstrates the applicability of the method to Lebesgue integrable densities.
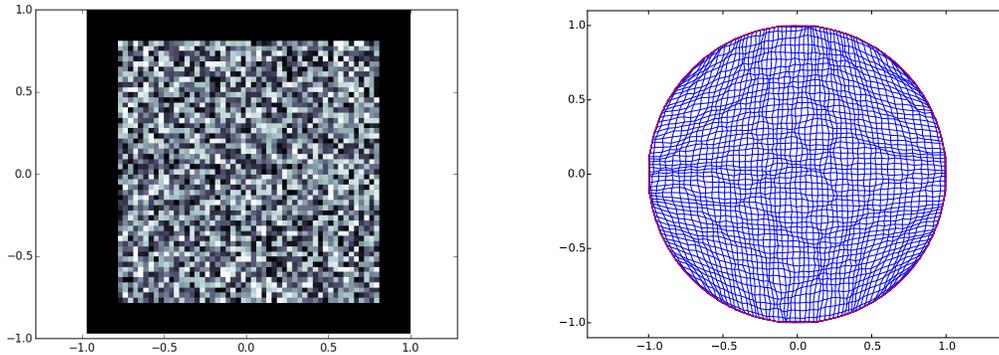
Figure 8: Left: Source density $f$. Right: Corresponding optimal map deformation of the grid.

## 6.4    Experiments with inhomogeneous target densities $g$

In this section we show the results obtained with a constant density square $X$ (not represented in the figures) mapped to heterogeneous densities $g$ defined on a ball $Y$, $N = 128$ and stencil width is 5. Figure 9 shows the case of one Gaussian and a mixture of gaussians. The boundary of the target is the red circle. The density $g$ and its gradient need to be defined in the numerical code as functions which can be evaluated anywhere in the ambient space of $Y$ as the map can hit anywhere including in $\mathbb{R}^2 \setminus Y$. If $g$ is given analytically (our case) this is easy we just extend it by a constant out of $Y$, otherwise one has to resort to interpolation.

## 6.5    Experiments with non convex and non connex sources

In this section we map constant source densities $f$ with non convex and non connex supports to a constant density ball $Y$. The solution is still $C^1$ but is is know that the inverse mapping, i.e. the Legendre-Fenchel transform of the potential, has gradient singularities. This has been analyzed by Figalli in [Fig09]. Figures 10 to 13 show different densities an the associated map deformations. The zero densities inclusions created singular structures in the target which correspond to gradient singularities of the dual map. These structures are consistent with those predicted by Figalli. We use $N = 128$ and the stencil width is 5 .

## 7    Conclusion

In this work, we have proposed an novel way to impose the BV2 boundary condition for the Monge-Ampère equation in schemes such as MA-LBR. The idea consists in slightly extending the domain of the solution so as to capture the behavior of "minimal" Brenier solutions to the optimal transport problem. Our proof of convergence as the grid stepsize goes to zero does not appeal to the theory of viscosity solutions, but rather to simple arguments combined with standard optimal transport results. The numerical experiments faithfully reproduce the typical behavior of optimal transport solutions.

Although numerically we have not encountered any particular difficulty with the resolution of the scheme, the existence of a solution to the discrete problem remains an open problem that we leave for future work.
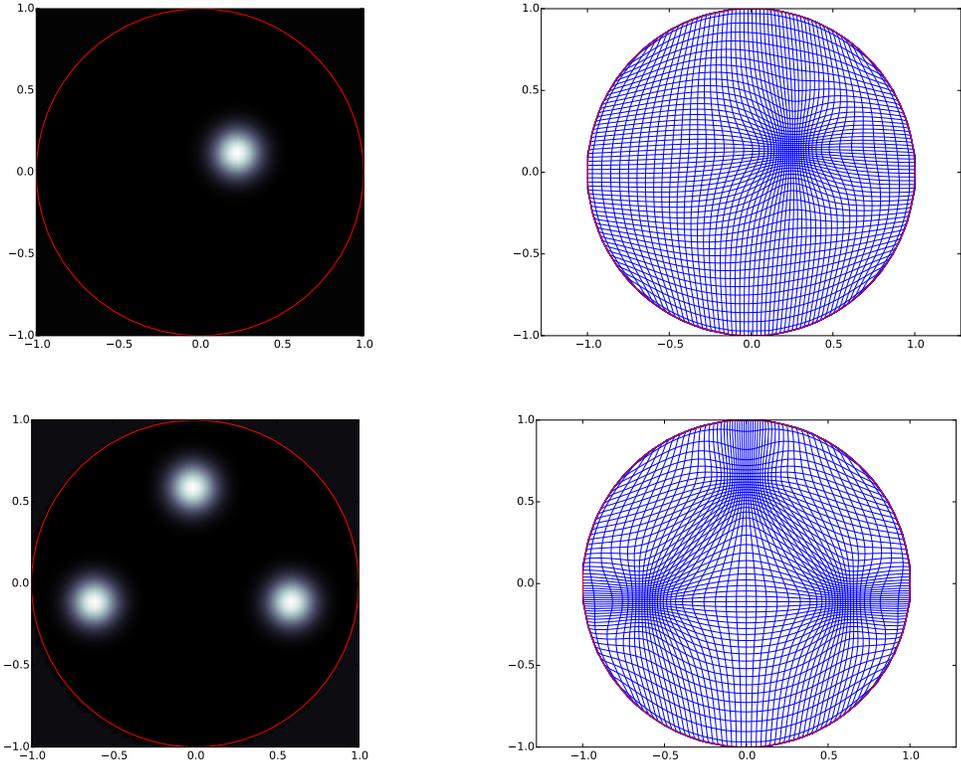
Figure 9: Left: Target density $g$. The boundary of the target is the red circle. Right: Corresponding optimal map deformation of the grid.

## Acknowledgments

## A    Convergence of the finite difference schemes

The aim of this section is to prove that (30) holds for the centered, forward and backward finite difference schemes, respectively $\mathcal{D}_n^{1,C}$ (see (31)), $\mathcal{D}_n^{1,F}$ and $\mathcal{D}_n^{1,B}$ (see (32)).

We have defined in Section 4, the values $(U_n[x])_{x \in D \cap \mathcal{G}_n}$, and in Section 5 their "interpolation" $\tilde{u}_n$ using (41) as a convex function. In Lemma 5.3, we have shown that this interpolation converges (along some sequence) towards some function $v \in \mathscr{C}^1(\mathbb{R}^2)$. We wish to prove that the discrete gradients $\mathcal{D}_n^{1,C} U_n$, $\mathcal{D}_n^{1,F} U_n$ and $\mathcal{D}_n^{1,B} U_n$ converge in some sense to the gradient $\nabla v$.

### A.1    Convergence of the gradient of convex functions

We begin with some general results on convex functions. The following lemma is standard, and mainly follows from the results in [HUL96, Section VI.6.2]. However, we provide a proof below, as this result is not explicitly stated there.

**Lemma A.1.** *Let $\Omega \subseteq \mathbb{R}^2$ be a nonempty open convex set, $C > 0$, and $\{v_n\}_{n \in \mathbb{N}}$ be a sequence of finite-valued convex functions on $\Omega$ which converges pointwise on $\Omega$ towards some function $v : \Omega \to \mathbb{R}$ as $n \to +\infty$.*
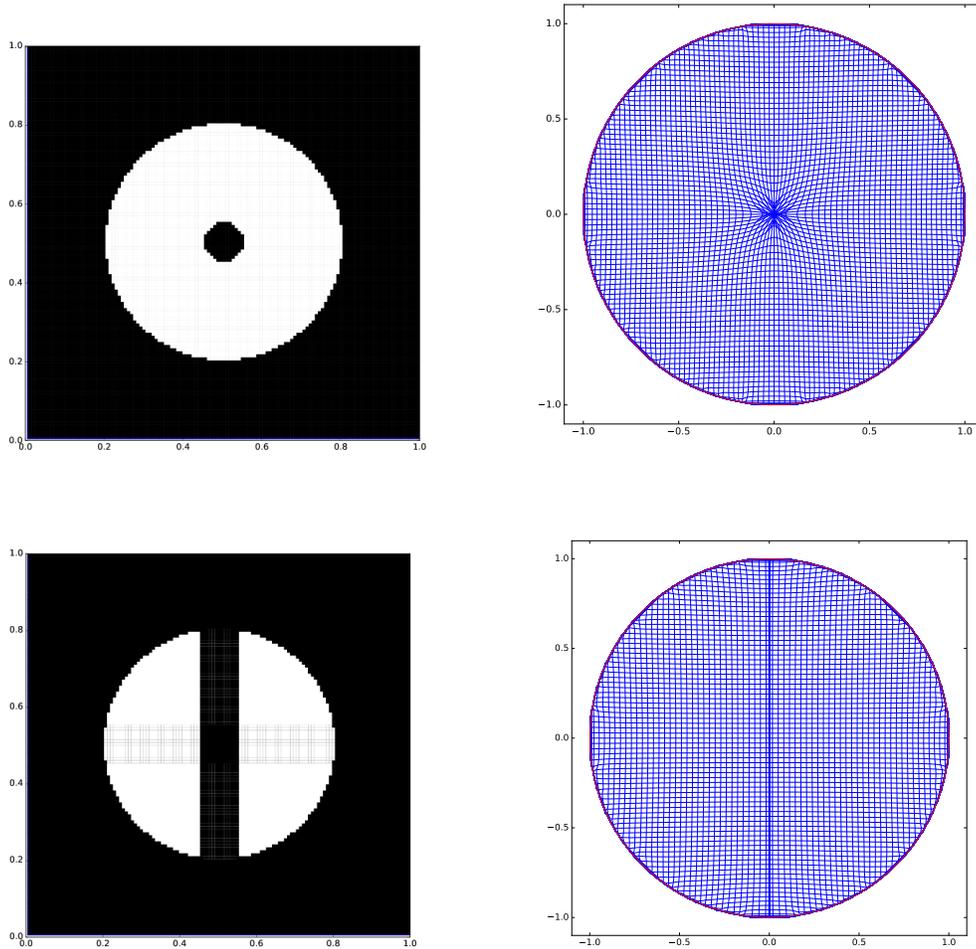
Figure 10: Left column: Source density $f$. Right column: Corresponding optimal map deformation of the grid.

If $v \in \mathscr{C}^1(\Omega)$, then for all compact set $K \subseteq \Omega$,

$$\lim_{n \to +\infty} \sup_{x \in K} \sup_{|x'-x| \leq C h_n} \sup_{y \in \partial v_n(x')} |\nabla v(x) - y|. \tag{53}$$

*Proof.* Let us recall that the convexity of $v_n$ implies that $v$ is convex, and $v_n$ converges uniformly on the compact sets of $\Omega$ towards $v$. Let $\rho > 0$ such that $\rho \leq \frac{1}{4} \operatorname{dist}(K, \mathbb{R}^2 \setminus \Omega)$, and let $K' \stackrel{\text{def.}}{=} \{x' \in \mathbb{R}^2 \; ; \; \operatorname{dist}(x', K) \leq \rho\} \subseteq \Omega$.

Observe that $\partial v_n(K')$ is bounded independently of $n$. Indeed, for any $x' \in K'$, any $y_n \in \partial v_n(x')$, if $y \neq 0$,

$$v_n \left( x' + \rho \frac{y_n}{|y_n|} \right) \geq v_n(x') + \rho |y_n|, \tag{54}$$

$$\text{hence } |y_n| \leq \frac{1}{\rho} \left( v_n \left( x' + \rho \frac{y_n}{|y_n|} \right) - v_n(x') \right), \tag{55}$$
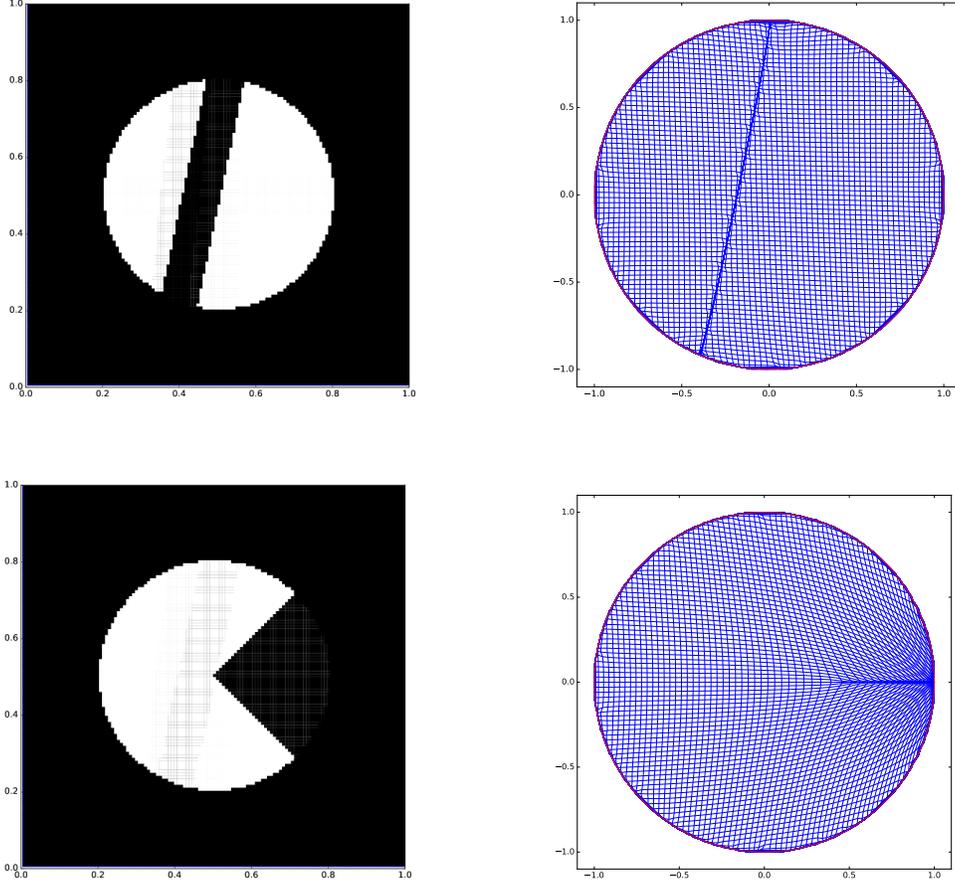
Figure 11: Left column: Source density $f$. Right column: Corresponding optimal map deformation of the grid.

and the right-hand side is uniformly bounded by uniform convergence of $v_n$ on $\{x'' \in \Omega \; ; \; \text{dist}(x, K) \leq 2\rho\}$.

Now, assume by contradiction that there is some $\varepsilon > 0$, some (not relabeled) subsequence $x_n \in K$, $x'_n \in B(x_n, Ch_n)$, and $y_n \in \partial v_n(x'_n)$ with $|y_n - \nabla v(x_n)| \geq \varepsilon$.

By compactness, there exists $y \in \mathbb{R}^2$, $x \in K$, such that (up to an additional extraction) $y_n \to y$ and $x_n \to x$. Passing to the limit in the subgradient inequality

$$\forall x'' \in \Omega, \quad v_n(x'') \geq v_n(x'_n) + \langle y_n, \, x'' - x'_n \rangle, \tag{56}$$

we obtain
$$\forall x'' \in \Omega, \quad v(x'') \geq v(x) + \langle y, \, x'' - x \rangle. \tag{57}$$

Hence $y = \nabla v(x)$ which yields a contradiction with $|y_n - \nabla v(x_n)| \geq \varepsilon$ for $n$ large enough. This yields the claimed result. $\qquad\square$

Now, we consider the finite difference scheme applied to a convex function.

**Lemma A.2.** *Let $\Omega \subseteq \mathbb{R}^2$ be a nonempty open convex set, and $\{v_n\}_{n \in \mathbb{N}}$ be a family of finite-valued convex functions on $\Omega$ which converges pointwise on $\Omega$ towards some function $v : \Omega \to \mathbb{R}$. Define $V_n[x] \overset{\text{def.}}{=} v_n(x)$ for all $x \in \Omega \cap \mathcal{G}_n$. Then, for all compact set $K \subseteq \Omega$,*

$$\lim_{n \to +\infty} \sup_{x \in K \cap \mathcal{G}_n} \left| \mathcal{D}_n^{1,C} V_n[x] - \nabla v(x) \right| \tag{58}$$
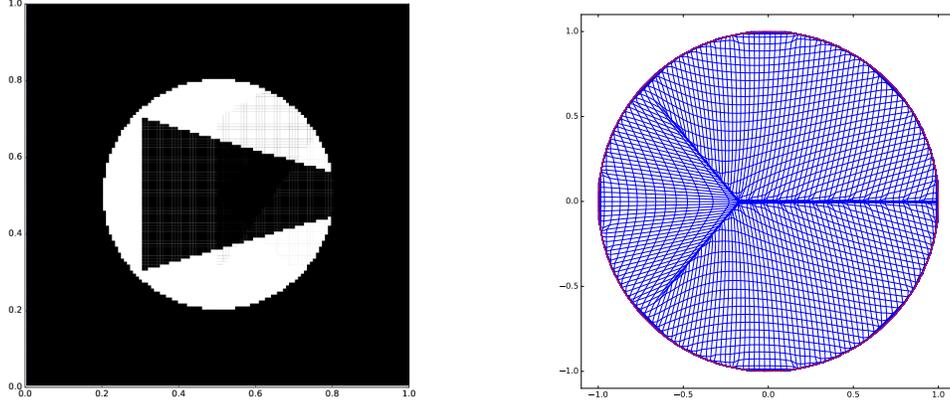
Figure 12: Left: Source density $f$. Right: Corresponding optimal map deformation of the grid.
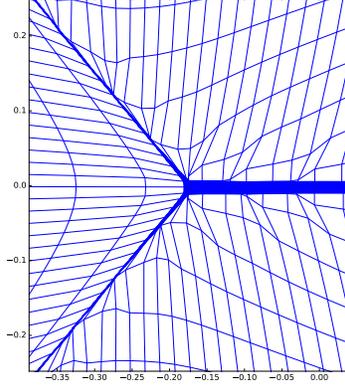


Figure 13: Zoom of Figure 12, Right.

where the centered finite difference operator $\mathcal{D}_n^{1,C}$ is defined in (31).

*Proof.* By [HUL96, Theorem VI.2.3.4], we see that

$$\delta_{(1,0)}^{h_n} V_n[x] = V_n[x + h_n(1,0)] - V_n[x] = h_n \int_0^1 \max_{y \in \partial v_n(x+sh_n(1,0))} \langle y, (1,0) \rangle \mathrm{d}s, \qquad (59)$$

so that by Lemma A.1, the above quantity converges uniformly on $K$ towards $\frac{\partial v}{\partial x_1}(x)$.

Similarly, the quantities $\delta_{(-1,0)}^{h_n} V_n[x]$, $\delta_{(0,1)}^{h_n} V_n[x]$ and $\delta_{(1,0)}^{h_n} V_n[x]$ respectively converge to $-\frac{\partial v}{\partial x_1}(x)$, $\frac{\partial v}{\partial x_2}(x)$ and $-\frac{\partial v}{\partial x_2}(x)$ uniformly. □

## A.2   Finite difference schemes for the subsampled sequences

Now, let us turn to the values $(U_n[x])_{x \in D \cap \mathcal{G}_n}$ defined in Section 4 and the function $\tilde{u}_n$ defined in (41). In the following, $v$ denotes the function constructed in Lemma 5.3, as the limit of $(\tilde{u}_n)_{n \in \mathbb{N}}$. Unfortunately, we cannot directly apply Lemma A.2 to $\tilde{u}_n$ as it is not really an interpolation of

28

$(U_n[x])_{x \in D \cap \mathcal{G}_n}$: there could be some points $x \in D \cap \mathcal{G}_n$ such that $\tilde{u}_n(x) < U_n[x]$ hence $\mathcal{D}_n^{1,C} U_n$ is not a priori the centered finite difference scheme applied to $\tilde{u}_n$.

However, the first two properties of Proposition 4.3 express the fact that $U_n$ is directionally convex, in the sense of [Mir16b, Appendix A]. In particular, defining the subsampled grids

$$\mathcal{G}_n^{(I)} \stackrel{\text{def.}}{=} h(2\mathbb{Z}) \times (2\mathbb{Z}), \quad \mathcal{G}_n^{(II)} \stackrel{\text{def.}}{=} h(2\mathbb{Z}+1) \times (2\mathbb{Z}),$$

$$\mathcal{G}_n^{(III)} \stackrel{\text{def.}}{=} h(2\mathbb{Z}) \times (2\mathbb{Z}+1), \quad \mathcal{G}_n^{(IV)} = h(2\mathbb{Z}+1) \times (2\mathbb{Z}+1),$$

there exist convex lower semi-continuous functions $v_n^{(I)}, v_n^{(II)}, v_n^{(III)}, v_n^{(IV)} : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ such that

$$\forall x \in \mathcal{G}_n^{(I)} \cap D, \quad v_n^{(I)}(x) = U_n[x], \tag{60}$$

and similarly, replacing $(I)$ with $(II)$, $(III)$, $(IV)$.

The following Lemma shows that those convex functions are not too far from our interpolation $\tilde{u}_n$, at least far from the boundary of $D$. We define $D_n \stackrel{\text{def.}}{=} \{x \in D \; ; \; x + 2h_n \mathcal{V}_n \subseteq D\}$.

**Proposition A.3.** *The following inequalities hold.*

$$\forall x \in \mathcal{G}_n^{(I)} \cap D, \quad v_n^{(I)}(x) \geq \tilde{u}_n(x), \tag{61}$$

$$\forall x \in \mathcal{G}_n^{(I)} \cap D_n, \quad \tilde{u}_n(x) \geq v_n^{(I)}(x) - Ch_n, \tag{62}$$

*where $C = 4\max\left\{\sigma_{\overline{Y}}(e) \; ; \; e \in \{(\pm 1, 0), (0, \pm 1)\}\right\}$.*

*The corresponding inequalities also hold for $v_n^{(II)}, v_n^{(III)}, v_n^{(IV)}$.*

*Proof.* The first inequality readily follows from the construction of $\tilde{u}_n$,

$$\forall x \in \mathcal{G}_n^{(I)} \cap D, \quad \tilde{u}_n(x) \leq U_n[x] = v_n^{(I)}(x).$$

Now we deal with (62). Let $x \in \mathcal{G}_n^{(I)} \cap D_n$. We first prove that $\partial v_n^{(I)}(x) \subseteq \overline{Y_n}$. For any $y \in \partial v_n^{(I)}(x)$, all $e \in \mathcal{V}_n$, and using Proposition 4.3,

$$\langle y, 2h_n e \rangle \leq v_n^{(I)}(x + 2h_n e) - v_n^{(I)}(x) = U_n[x + 2h_n e] - U_n[x] \leq \sigma_{\overline{Y}}(2h_n e)$$

Hence $\langle y, e \rangle \leq \sigma_{\overline{Y}}(e)$ for all $e \in \mathcal{V}_n$, that is $y \in \overline{Y_n}$. Moreover, by the subgradient inequality,

$$\forall x' \in \mathcal{G}_n^{(I)} \cap D, \quad U_n[x'] = v_n^{(I)}(x') \geq v_n^{(I)}(x) + \langle y, x' - x \rangle. \tag{63}$$

For all $x'' \in \mathcal{G}_n \cap D$, there exists $x' \in \mathcal{G}_n^{(I)} \cap D$ such that $\|x'' - x'\|_\infty \leq h_n$. Let $e \stackrel{\text{def.}}{=} \frac{1}{h_n}(x'' - x')$ satisfying $\|e\|_\infty \leq 1$ (hence $|\langle y, e \rangle| \leq 2\max\left\{\sigma_{\overline{Y}}(e) \; ; \; e \in \{(\pm 1, 0), (0, \pm 1)\}\right\}$). Using $U_n[x'] \leq U_n(x'') + \sigma_{\overline{Y}}(-h_n e)$ in (63) we get

$$U_n(x'') + \sigma_{\overline{Y}}(-h_n e) \geq v_n^{(I)}(x) + \langle y, x'' - x \rangle - \langle y, h_n e \rangle,$$

$$\text{hence} \quad U_n(x'') \geq \langle y, x'' - x \rangle + v_n^{(I)}(x) - Ch.$$

As a result, the affine mapping $L : x'' \mapsto \left(\langle y, x'' - x \rangle + v_n^{(I)}(x) - Ch\right)$ minorizes $U_n[x'']$ at each $x'' \in \mathcal{G}_n \cap D$, and $\nabla L = y \in \overline{Y_n}$. We deduce that $\tilde{u}_n \geq L$. In particular, $\tilde{u}_n(x) \geq L(x) = v_n^{(I)}(x) - Ch_n$. $\qquad\square$

The next proposition shows that, along sequences, $\tilde{u}_n$ and $v_n^{(I)}$ share the same limit as $h \to 0^+$.

**Proposition A.4.** *Let $K \subseteq \text{int}(D)$ be a compact set.*
*There exists $\hat{v} \in \mathscr{C}(K)$ such that for $n$ large enough $v_n^{(I)} \in \mathscr{C}(K)$ and up to a subsequence,*

$$\lim_{n \to +\infty} \left\| v_n^{(I)} - \hat{v} \right\|_{L^\infty(K)} = \lim_{n \to +\infty} \left\| v_n^{(II)} - \hat{v} \right\|_{L^\infty(K)} = 0, \tag{64}$$

$$\lim_{n \to +\infty} \left\| v_n^{(III)} - \hat{v} \right\|_{L^\infty(K)} = \lim_{n \to +\infty} \left\| v_n^{(IV)} - \hat{v} \right\|_{L^\infty(K)} = 0, \tag{65}$$

$$\lim_{n \to +\infty} \left\| \tilde{u}_{h_n} - \hat{v} \right\|_{L^\infty(K)} = 0. \tag{66}$$

*Proof.* The first point follows from Proposition 4.3 which ensures that $(U_n[x])_{n \in \mathbb{N}}$ is bounded uniformly in $x \in \mathcal{G}_n \cap D$. As a result, the convex function $v_n^{(I)}$ is uniformly bounded (hence uniformly Lipschitz) in a neighborhood of $K$ and Ascoli-Arzelà's theorem ensures the convergence of $v_n^{(I)}$ towards some $\hat{v}$ up to a subsequence.

By Proposition A.3, we see that $\tilde{u}_{h_n}$ must converge (pointwise) towards $\hat{v}$ on a dense subset of $K$. Since $K$ is compact and the functions are uniformly Lipschitz, the convergence is uniform on $K$.

Since Proposition A.3 also holds for $v_n^{(II)}$, $v_n^{(III)}$, $v_n^{(IV)}$, we deduce that those convex functions also converge uniformly (along the same subsequence) towards $\hat{v}$. $\qquad\square$

In Lemma 5.2, we show that $\{\tilde{u}_n\}_{n \in \mathbb{N}} \subseteq \mathscr{C}(\mathbb{R}^2)$ is precompact for the topology of uniform convergence on compact subsets of $\mathbb{R}^2$. As a consequence of Proposition A.3, if $\tilde{u}_{h_n}$ converges along some subsequence towards some function $v$, then $v_n^{(I)}, \ldots v_n^{(IV)}$ converge uniformly towards $v$ on compact subsets of $\text{int}(D)$, along the same subsequence. Now, in Lemma 5.3, we show that $v \in \mathscr{C}^1(\mathbb{R}^2)$. We may now state the main result of this section.

**Proposition A.5.** *Assume that (up to a subsequence) $\tilde{u}_{h_n}$ converges uniformly on the compacts subsets of $\mathbb{R}^2$ towards some function $v \in \mathscr{C}^1(\mathbb{R}^2)$. Then, for all $K \subseteq \text{int}(D)$ compact,*

$$\lim_{n \to +\infty} \max_{x \in K \cap \mathcal{G}_n} \left| \mathcal{D}_n^{1,C} U_n[x] - \nabla v(x) \right| = 0, \tag{67}$$

$$\lim_{n \to +\infty} \max_{x \in K \cap \mathcal{G}_n} \left| \mathcal{D}_n^{1,F} U_n[x] - \nabla v(x) \right| = 0, \tag{68}$$

$$\lim_{n \to +\infty} \max_{x \in K \cap \mathcal{G}_n} \left| \mathcal{D}_n^{1,B} U_n[x] - \nabla v(x) \right| = 0. \tag{69}$$

*Proof.* We note that from Proposition A.4, $v_n^{(I)}, \ldots v_n^{(IV)}$ also converge uniformly on compact subsets of $\text{int}(D)$ towards $v$.

We begin with the forward and backward differences defined in (32). Let $e = (1,0)$ or $e = (0,1)$. By monotonicity of the slopes (see Proposition 4.3),

$$\frac{U_n[x + 2h_n e] - U_n[x]}{2h_n} \geq \frac{U_n[x + h_n e] - U_n[x]}{h_n} \geq \frac{U_n[x] - U_n[x - h_n e]}{h_n} \geq \frac{U_n[x] - U_n[x - 2h_n e]}{2h_n} \tag{70}$$

To fix ideas, let us assume that $x \in \mathcal{G}_n^{(I)}$. Then $x \pm 2h_n e \in \mathcal{G}_n^{(I)}$ and applying Lemma (A.2) to $v_n^{(I)}$ (in some compact set $K'$, $K \subseteq K' \subseteq \text{int}(D)$), we see that the left and right-hand sides of (70) converge towards $\langle \nabla v(x), e \rangle$, uniformly in $x \in \mathcal{G}_n^{(I)} \cap K$. The same argument holds for $x$ in $\mathcal{G}_n^{(II)}$, $\mathcal{G}_n^{(III)}$ or $\mathcal{G}_n^{(IV)}$, so that we get (68) and (69).

By linear combination, we deduce (67). $\qquad\square$

# B    MA-LBR overestimates the subgradient

The following proof of Lemma 4.2 was suggested to us by J.-M. Mirebeau. We denote by $\partial_x u$ the subgradient of a convex function $u$ at a point $x$.

**Lemma B.1.** *Let $\Omega$ be a convex neighborhood of a point $x \in \mathbb{R}^d$, and let $u, v : \Omega \to \mathbb{R}^d$ be convex. If $u \leq v$, and $u(x) = v(x)$ then $\partial_x u \subseteq \partial_x v$. In particular $|\partial_x u| \leq |\partial_x v|$.*

The following Lemma is a consequence of the Brunn-Minkowski inequality (see [Sch93]).

**Lemma B.2.** *Let $\Omega$ be a convex neighborhood of a point $x \in \mathbb{R}^d$, and let $u, v : \Omega \to \mathbb{R}^d$ be convex. Let $w = (1-t)u + tv$ with $0 \leq t \leq 1$. Then $(1-t)\partial_x u + t\partial_x v \subseteq \partial_x w$, where the sign plus denotes a Minkowski sum. In particular $(1-t)|\partial_x u|^{\frac{1}{d}} + t|\partial_x v|^{\frac{1}{d}} \leq |\partial_x w|^{\frac{1}{d}}$.*

We denote by $\underline{u}$ the largest convex function bounded above by $u$.

**Proposition B.3.** *Let $X \subseteq \mathbb{R}^d$ be a finite point set, and let $x \in X$. Let $Y \subseteq X$ be symmetric w.r.t. the point $x$, i.e. $\forall y \in Y$ one has $2x - y \in Y$. Let $u : X \to \mathbb{R}$, and let $v : Y \to \mathbb{R}$ be defined by $v(y) := \frac{1}{2}(u(y) + u(2x - y))$. Then*

$$|\partial_x \underline{u}| \leq |\partial_x \underline{v}|. \tag{71}$$

*Proof.* If $u(x) > \underline{u}(x)$, then $|\partial_x \underline{u}(x)| = 0$ and there is nothing to prove. We thus assume $u(x) = \underline{u}(x)$. Introduce the restriction $u_Y := u_{|Y}$, and its symmetry $u_{-Y} := u_Y(2x - \cdot)$, and note that $\frac{1}{2}(u_Y + u_{-Y}) \leq v$.

The equality $\overline{u(x)} = \underline{u}(x)$ implies $u_Y(x) = \underline{u_Y}(x)$, hence also $u_{-Y}(x) = \underline{u_{-Y}}(x)$ and $v(x) = \underline{v}(x)$. We have

$$\underline{u} \leq \underline{u_Y} \quad \text{hence } |\partial_x \underline{u}| \leq |\partial_x \underline{u_Y}|,$$
$$\frac{1}{2}(\underline{u_Y} + \underline{u_{-Y}}) \leq \underline{v} \quad \text{hence } \frac{1}{2}|\partial_x \underline{u_Y}|^{\frac{1}{d}} + \frac{1}{2}|\partial_x \underline{u_{-Y}}|^{\frac{1}{d}} \leq |\partial_x \underline{v}|^{\frac{1}{d}}.$$

The announced result follows since $|\partial_x \underline{u_Y}| = |\partial_x \underline{u_{-Y}}|$.

$\square$

The final step to prove Lemma 4.2 is given by Remark 1.10 in [BCM16] which shows that $\mathrm{MA}_n(U_n)[x]\omega_n = |\partial_x \underline{v}(x)|$.

# References

[BB00]      Jean-David Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[BCC+15]   Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative bregman projections for regularized transportation problems. *SIAM J. Sci. Comp.*, 2015. to appear.

[BCM16]    Jean-David Benamou, Francis Collino, and Jean-Marie Mirebeau. Monotone and consistent discretization of the monge-ampere operator. *Mathematics of computation*, 85(302):2743–2775, 2016.

[BFO14]    Jean-David Benamou, Brittany D Froese, and A M Oberman. Numerical solution of the Optimal Transportation problem using the Monge–Ampère equation. *Journal of Computational Physics*, 2014.

[Bre91]     Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 1991.

[BS91]      G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[Caf92]     Luis A Caffarelli. Boundary regularity of maps with convex potentials. *Communications on pure and applied mathematics*, 45(9):1141–1151, 1992.

[Caf96]     Luis A Caffarelli. Boundary regularity of maps with convex potentials–ii. *Annals of mathematics*, 144(3):453–496, 1996.

[CIL92]     Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[CMOB15]  Guillaume Carlier, Quentin Mérigot, Edouard Oudet, and Jean-David Benamou. Discretization of functionals involving the monge-ampère operator. *Numerische mathematik*, 2015.

[CP84]      M. J. P. Cullen and R. J. Purser. An extended Lagrangian theory of semigeostrophic frontogenesis. *J. Atmospheric Sci.*, 41(9):1477–1497, 1984.

[CPSV16]   Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Scaling algorithms for unbalanced transport problems. *arXiv preprint arXiv:1607.05816*, 2016.

[Cut13]     Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In Christopher J. C. Burges, Léon Bottou, Zoubin Ghahramani, and Kilian Quentin Weinberger, editors, *Proc. NIPS*, pages 2292–2300, 2013.

[Fig09]     Alessio Figalli. Regularity properties of optimal maps between nonconvex domains in the plane, 2009.

[FL09]      Alessio Figalli and Grégoire Loeper. C1 regularity of solutions of the monge–ampère equation for optimal transport in dimension two. *Calculus of Variations and Partial Differential Equations*, 35:537–550, 08 2009.

[FO13]     Brittany D Froese and A M Oberman. Convergent Filtered Schemes for the Monge–Ampère Partial Differential Equation. *SIAM Journal on Numerical Analysis*, 2013.

[Gal16]    Alfred Galichon. Optimal transport methods in economics, 2016.

[Gut01]    Cristian E. Gutierez. *The Monge-Ampère equation*. Progress in nonlinear differential equations and their applications. Birkhäuser, Boston, Basel, 2001.

[HUL96]    J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer Verlag, Heidelberg, 1996. Two volumes - 2nd printing.

[Kan42]    L. Kantorovich. On the transfer of masses. *Dokl.Acad. Nauk. USSR*, 37:7–8, 1942.

[KMT16]    Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a newton algorithm for semi-discrete optimal transport. *arXiv preprint arXiv:1603.05579*, 2016.

[Lév15a]   Bruno Lévy. A numerical algorithm for l2 semi-discrete optimal transport in 3d. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1693–1715, 2015.

[Lév15b]   Bruno Lévy. A numerical algorithm for l2 semi-discrete optimal transport in 3d. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1693–1715, 2015.

[LR05]     Grégoire Loeper and Francesca Rapetti. Numerical solution of the Monge-Ampère equation by a Newton's algorithm. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 2005.

[Mér11]    Q. Mérigot. A multiscale approach to optimal transport. *Computer Graphics Forum*, 30(5):1583–1592, 2011.

[Mir15]    Jean-Marie Mirebeau. Discretization of the 3D Monge-Ampere operator, between wide stencils and power diagrams. *ESAIM Math. Model. Numer. Anal.*, 49(5):1511–1523, 2015.

[Mir16a]   Jean-Marie Mirebeau. Adaptive, anisotropic and hierarchical cones of discrete convex functions. *Numerische Mathematik*, 132(4):807–853, 2016.

[Mir16b]   Jean-Marie Mirebeau. Adaptive, anisotropic and hierarchical cones of discrete convex functions. *Numerische Mathematik*, 132(4):807–853, 2016.

[Obe06]    Adam 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, 2006.

[Obe08]    Adam M Oberman. Computing the convex envelope using a nonlinear partial differential equation. *Mathematical Models and Methods in Applied Sciences*, 18(5):759–780, 2008.

[OP88]     V. I. Oliker and L. D. Prussner. On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ and its discretizations. I. *Numer. Math.*, 54(3):271–293, 1988.

[Roc83]    T. Rockafellar. *Convex Analysis*, volume 224 of *Grundlehren der mathematischen Wissenschaften*. Princeton University Press, second edition, 1983.

[SAK15]  Louis-Philippe Saumier, Martial Agueh, and Boualem Khouider. An efficient numerical algorithm for the l2 optimal transport problem with periodic densities. *IMA Journal of Applied Mathematics*, 80(1):135–157, 2015.

[San15]  Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 2015.

[Sch93]  R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge Series in Environment and Behavior. Cambridge University Press, 1993.

[Urb90]  John I. E. Urbas. On the existence of nonclassical solutions for two classes of fully nonlinear elliptic equations. *Indiana Univ. Math. J.*, 39(2):355–382, 1990.

[Vil09]  C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.