



**HAL**  
open science

# A fractal shape optimization problem in branched transport

Paul Pegon, Filippo Santambrogio, Qinglan Xia

► **To cite this version:**

Paul Pegon, Filippo Santambrogio, Qinglan Xia. A fractal shape optimization problem in branched transport. *Journal de Mathématiques Pures et Appliquées*, 2019, 123, pp.244-269. 10.1016/j.matpur.2018.06.007 . hal-01581675v2

**HAL Id: hal-01581675**

**<https://hal.science/hal-01581675v2>**

Submitted on 28 Dec 2017

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

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

# A FRACTAL SHAPE OPTIMIZATION PROBLEM IN BRANCHED TRANSPORT

PAUL PEGON<sup>†</sup>, FILIPPO SANTAMBROGIO<sup>†</sup>, AND QINGLAN XIA<sup>‡</sup>

ABSTRACT. We investigate the following question: what is the set of unit volume which can be best irrigated starting from a single source at the origin, in the sense of branched transport? We may formulate this question as a shape optimization problem and prove existence of solutions, which can be considered as a sort of “unit ball” for branched transport. We establish some elementary properties of optimizers and describe these optimal sets  $A$  as sublevel sets of a so-called landscape function which is now classical in branched transport. We prove  $\beta$ -Hölder regularity of the landscape function, allowing us to get an upper bound on the Minkowski dimension of the boundary:  $\overline{\dim}_M \partial A \leq d - \beta$  (where  $\beta := d(\alpha - (1 - 1/d)) \in (0, 1)$  is a relevant exponent in branched transport, associated with the exponent  $\alpha > 1 - 1/d$  appearing in the cost). We are not able to prove the lower bound, but we conjecture that  $\partial A$  is of non-integer dimension  $d - \beta$ . Finally, we make an attempt to compute numerically an optimal shape, using an adaptation of the phase-field approximation of branched transport introduced some years ago by Oudet and the second author.

## CONTENTS

Introduction	1
1. Preliminaries	3
1.1. The irrigation problem	3
1.2. The shape optimization problem	5
2. Existence and first properties	6
3. Hölder continuity of the landscape function	8
3.1. Main lemmas	9
3.2. Hölder regularity	10
4. On the dimension of the boundary	13
5. Numerical simulations	15
5.1. Optimization methods	16
5.2. Computing the projection	17
5.3. Algorithms and numerical experiments	19
References	23

## INTRODUCTION

Given two probability measures  $\mu, \nu$  on  $\mathbb{R}^d$ , a classical optimization problem amounts to finding a connection between the two measures which has minimal cost. In branched

---

(<sup>†</sup>) LABORATOIRE DE MATHÉMATIQUES D’ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY CEDEX, FRANCE

(<sup>‡</sup>) DEPARTMENT OF MATHEMATICS, UC DAVIS, ONE SHIELDS AVE, DAVIS, CA 95616, UNITED STATES  
*E-mail addresses:* paul.pegon@math.u-psud.fr, filippo.santambrogio@math.u-psud.fr, qlxia@math.ucdavis.edu.

2010 *Mathematics Subject Classification.* 49Q10 ; 49N60 ; 65K10 ; 28A80.

*Key words and phrases.* branched transport ; landscape function ; fractal dimension ; Morrey-Campanato spaces ; phase-field approximation, non-smooth optimization.

transport, such a connection will be performed along a 1-dimensional structure such that the cost for moving a mass  $m$  at distance  $\ell$  is proportional to  $m^\alpha \times \ell$  where  $\alpha$  is some concave exponent  $\alpha \in [0, 1]$ . The map  $t \mapsto t^\alpha$  being subadditive (even strictly subadditive for  $\alpha < 1$ ), that is to say  $(a + b)^\alpha \leq a^\alpha + b^\alpha$ , it is cheaper for masses to travel together as much as possible. Consequently, the optimal connections exhibit branching structures: for instance, if one wishes to transport one Dirac mass to two Dirac masses of mass  $1/2$ , the optimal graph will be  $Y$ -shaped.

A early model has been proposed by Gilbert in [Gil67] as an extension of Steiner problem (see [GP68]) in a discrete setting, where the connection between two atomic measures is made through weighted oriented graphs. There are two main extensions of this model to a continuous setting, i.e. with arbitrary probability measures. The first one was introduced in 2003 by the third author in [Xia03] and can be viewed as a Eulerian model. It is based on vector measures and roughly reads as:

$$\min \left\{ \int \left| \frac{dv}{d\mathcal{H}^1}(x) \right|^\alpha d\mathcal{H}^1(x) : \nabla \cdot v = \mu - \nu \right\},$$

minimizing among vector measures which have an  $\mathcal{H}^1$ -density. A Lagrangian model was introduced essentially at the same time by and Maddalena, Solimini, Morel [MSM03], and then intensively studied by Bernot, Caselles, Morel [BCM05]. It is based on measures on a set of curves, but the description of this model, which is a little more involved, is given in Section 1. An almost up-to-date reference on branched transport resides in the book by the same authors [BCM09].

Looking at the optimal branching structures computed numerically in [OS11] (in some non-atomic cases), or at natural drainage networks and their irrigations basins, one is tempted to describe them as fractal (see [RIR01]). Actually, even though the underlying network has infinitely many branching points, it is still a 1-rectifiable set, hence it is not clear in what sense fractality appears. Fractality is a notion which usually relates either to self-similarity properties of non-smooth objects, or to non-integer dimension of sets. A first rigorous result which would fall in the first category is proven by Brancolini and Solimini in [BS14]: for sufficiently diffuse measures (for example the Lebesgue measure restricted to a Lipschitz open set), the number of branches of length  $\sim \varepsilon$  stemming from a branch of length  $\ell$  is of order  $\ell/\varepsilon$ . This may read as a self-similarity property since in a way the total length is preserved when looking at subbranches at all scales.

The present paper leans towards the other notion of fractality, that is towards “fractal” dimension. Some sets in branched transport have already been proposed as candidates to exhibit non-integer dimension, for instance the boundary of adjacent irrigation basins (an open conjecture by J.-M. Morel). Here we are interested in another candidate which is related to branched transport: the boundary of what we call unit balls for branched transport. With the results of the present paper, we can only prove an upper bound on the dimension, which is non-integer, and conjecture that this upper bound is actually sharp.

The article is divided into five parts. In a preliminary section we define properly the Lagrangian framework of branched transport and its basic features, and we formulate our question as a shape optimization problem involving the irrigation distance. Section 2 is devoted to the proof of existence of minimizers and to elementary properties of minimizers. In Section 3 we prove the  $\beta$ -Hölder regularity of the landscape function, which appears in the description of optimizers, and use it to derive an upper bound on the Minkowski dimension of the boundary of the optimizers in Section 4. The final section is an attempt at computing optimizers numerically, which is particularly useful due to the fact that we are not fully able to answer theoretically the question of the fractal behavior of the boundary. This is done by adapting the Modica-Mortola approach introduced by [OS11], and allows to provide some convincing computer visualizations.

## 1. PRELIMINARIES

As preliminaries, we quickly set the Lagrangian framework of branched transport and its main features. For more details, we refer to the book [BCM09] or to [Peg17, Sections 1–2] for a simpler exposition.

**1.1. The irrigation problem.** We denote by  $\Gamma(\mathbb{R}^d)$  the set of 1-Lipschitz curves in  $\mathbb{R}^d$  parameterized on  $[0, \infty]$ , endowed with the topology of uniform convergence on compact sets.

**Irrigation plans.** We call *irrigation plan* any probability measure  $\eta \in \text{Prob}(\Gamma)$  satisfying the following finite-length condition

$$(1.1) \quad \mathbf{L}(\eta) := \int_{\Gamma} L(\gamma) \, d\eta(\gamma) < +\infty,$$

where  $L(\gamma) = \int_0^{\infty} |\dot{\gamma}(t)| \, dt$ . Notice that any irrigation plan is concentrated on  $\Gamma^1(\mathbb{R}^d) := \{\gamma : L(\gamma) < \infty\}$ . We denote by  $\text{IP}(\mathbb{R}^d)$  the set of all irrigation plans  $\eta \in \text{Prob}(\Gamma)$ . If  $\mu$  and  $\nu$  are two probability measures on  $\mathbb{R}^d$ , one says that  $\eta \in \text{IP}(\mathbb{R}^d)$  irrigates  $\nu$  from  $\mu$  if one recovers the measures  $\mu$  and  $\nu$  by sending the mass of each curve respectively to its initial point and to its final point, which means that

$$(\pi_0)_{\#}\eta = \mu \text{ and } (\pi_{\infty})_{\#}\eta = \nu,$$

where  $\pi_0(\gamma) = \gamma(0)$ ,  $\pi_{\infty}(\gamma) = \gamma(\infty) := \lim_{t \rightarrow +\infty} \gamma(t)$  and  $f_{\#}\eta$  denotes the push-forward of  $\eta$  by  $f$  whenever  $f$  is a Borel map<sup>1</sup>. We denote by  $\text{IP}(\mu, \nu)$  the set of irrigation plans irrigating  $\nu$  from  $\mu$ :

$$\text{IP}(\mu, \nu) = \{\eta \in \text{IP}(\mathbb{R}^d) : (\pi_0)_{\#}\eta = \mu, (\pi_{\infty})_{\#}\eta = \nu\}.$$

If  $\eta$  is a given irrigation plan, we define the multiplicity at  $x$ , that is the total mass passing by  $x$ , as

$$\theta_{\eta}(x) = \eta(\{\gamma \in \Gamma : x \in \gamma\}),$$

where  $x \in \gamma$  means that  $x$  belongs to the image of the curve  $\gamma$ . Finally, for any nonnegative function  $f$ , we denote by  $\int_{\gamma} f(x) \, dx$  the line integral of  $f$  along  $\gamma \in \Gamma$ :

$$\int_{\gamma} f(x) \, dx := \int_0^{+\infty} f(\gamma(t)) |\dot{\gamma}(t)| \, dt.$$

**Irrigation costs.** For  $\alpha \in [0, 1]$  we consider the *irrigation cost*  $\mathbf{I}_{\alpha} : \text{IP}(\mathbb{R}^d) \rightarrow [0, \infty]$  defined by

$$\mathbf{I}_{\alpha}(\eta) := \int_{\Gamma} \int_{\gamma} \theta_{\eta}(x)^{\alpha-1} \, dx \, d\eta(\gamma),$$

with the conventions  $0^{\alpha-1} = \infty$  if  $\alpha < 1$ ,  $0^{\alpha-1} = 1$  otherwise, and  $\infty \times 0 = 0$ . If  $\mu, \nu$  are two probability measures on  $\mathbb{R}^d$ , the irrigation (or branched transport) problem consists in minimizing the cost  $\mathbf{I}_{\alpha}$  on the set of irrigation plans which send  $\mu$  to  $\nu$ , which reads

$$(LL_{\alpha}) \quad \min_{\eta \in \text{IP}(\mu, \nu)} \int_{\Gamma} \int_{\gamma} \theta_{\eta}(x)^{\alpha-1} \, dx \, d\eta(\gamma).$$

We set  $Z_{\eta}(\gamma) = \int_{\gamma} \theta_{\eta}(x)^{\alpha-1} \, dx$  so that the cost may be expressed as

$$\mathbf{I}_{\alpha}(\eta) = \int_{\Gamma} Z_{\eta}(\gamma) \, d\eta(\gamma).$$

The following results are extracted from [BCM05; Peg17; Xia03].

<sup>1</sup>Notice that  $\lim_{t \rightarrow \infty} \gamma(t)$  exists if  $\gamma \in \Gamma^1(K)$ , and this is all we need since any irrigation plan is concentrated on  $\Gamma^1(K)$ .

**Proposition 1.1** (First variation inequality for  $\mathbf{I}_\alpha$ ). *If  $\eta$  is an irrigation plan with  $\mathbf{I}_\alpha(\eta)$  finite, then for all irrigation plan  $\tilde{\eta}$  the following holds:*

$$(1.2) \quad \mathbf{I}_\alpha(\tilde{\eta}) \leq \mathbf{I}_\alpha(\eta) + \alpha \int Z_\eta(\gamma) d(\tilde{\eta} - \eta).$$

*Notice that the integral  $\int Z_\eta d(\tilde{\eta} - \eta)$  is well-defined since  $\int Z_\eta d\eta < \infty$  and  $Z_\eta$  is nonnegative, though it may be infinite.*

**Theorem 1.2** (Existence of minimizers). *For any pair of probability measures  $\mu, \nu \in \text{Prob}(\mathbb{R}^d)$  which have compact support, the problem  $(\mathbf{LI}_\alpha)$  admits a minimizer.*

**Theorem 1.3** (Irrigability). *If  $1 - \frac{1}{d} < \alpha < 1$ , for any  $\mu, \nu \in \text{Prob}(\mathbb{R}^d)$  with compact support there exists some  $\eta \in \text{IP}(\mu, \nu)$  such that  $\mathbf{I}_\alpha(\eta)$  is finite.*

From now on we assume that  $\alpha \in ]1 - \frac{1}{d}, 1[$ .

**Irrigation distance.** Let us set

$$d_\alpha(\mu, \nu) = \min\{\mathbf{I}_\alpha(\eta) : \eta \in \text{IP}(\mu, \nu)\}$$

for any pair  $\mu, \nu$  of probability measures on  $\mathbb{R}^d$ . For any compact  $K \subseteq \mathbb{R}^d$ , it induces a distance on  $\text{Prob}(K)$  which metrizes the weak- $\star$  convergence of measures in the duality with  $\mathcal{C}(K)$ . On non-compact subsets of  $\mathbb{R}^d$ , the distance  $d_\alpha$  is lower semicontinuous w.r.t. the weak- $\star$  convergence of measures in the duality with bounded and continuous functions (narrow convergence)<sup>2</sup>.

**Proposition 1.4** (Scaling law). *For any compactly supported measures  $\mu, \nu$  with equal mass, there is an upper bound on the irrigation distance depending on the mass and the diameter. We set  $\mu' = \mu - \mu \wedge \nu, \nu' = \nu - \mu \wedge \nu$  the disjoint parts of the measures and  $m = |\mu'| = |\nu'|$  their common mass. Then:*

$$d_\alpha(\mu, \nu) \leq C m^\alpha \text{diam}(\text{supp } \mu' \cup \text{supp } \nu').$$

**Landscape function.** The landscape function was introduced by the second author in [San07], in the single-source case. It has been then studied by Brancolini, Solimini in [BS11] and by the third author in [Xia14]. It will be a central tool in the study of the shape optimization problem we are going to introduce. We recall here the basic definitions and properties. Given an optimal irrigation plan  $\eta \in \text{IP}(\delta_0, \nu)$ , we say that a curve  $\gamma$  is  $\eta$ -good if

- the quantity  $Z_\eta(\gamma) = \int_\gamma \theta_\eta(x)^{\alpha-1} dx$  is finite,
- for all  $t < T(\gamma)$ ,

$$\theta_\eta(\gamma(t)) = \eta(\{\tilde{\gamma} \in \Gamma(\mathbb{R}^d) : \gamma = \tilde{\gamma} \text{ on } [0, t]\}),$$

where  $T(\gamma) = \inf\{t \in [0, \infty] : \gamma(s) = \gamma(\infty) \text{ for all } s \in [t, \infty]\}$  is the stopping time of  $\gamma$ .

One may prove by optimality that  $\eta$  is concentrated on the set of  $\eta$ -good curves. Moreover it is proven in [San07] that for all  $\eta$ -good curves  $\gamma$ , the quantity  $Z_\eta(\gamma)$  depends only on the final point  $\gamma(\infty)$  of the curve, thus we may define the landscape function  $z_\eta$  as follows:

$$z_\eta(x) = \begin{cases} Z_\eta(\gamma) & \text{if } \gamma \text{ is an } \eta\text{-good curve s.t. } x = \gamma(\infty), \\ +\infty & \text{otherwise.} \end{cases}$$

<sup>2</sup>Proving this is just an adaptation of the proof on compact sets. If  $\mu$  is fixed (for example) and  $\nu_n \rightarrow \nu$  with  $\eta_n \in \text{IP}(\mu, \nu_n)$  optimal and parameterized by arc length, assuming that the cost is bounded, the irrigation plans  $\eta_n$  are tight and one may extract a subsequence converging to some  $\eta$  which irrigates  $\nu$  and whose cost is less than  $\liminf d_\alpha(\mu, \nu_n)$  by lower semicontinuity of  $\mathbf{I}_\alpha$ .

Notice that for an optimal  $\eta$  the cost may be expressed in terms of  $z_\eta$ :

$$\mathbf{I}_\alpha(\eta) = \int_\Gamma Z_\eta(\gamma) d\eta(\gamma) = \int_{\mathbb{R}^d} z_\eta(x) d\nu(x).$$

Finally, one may show that  $z_\eta$  is lower semicontinuous and that the inequality  $z_\eta(x) \geq |x|$  holds.

**1.2. The shape optimization problem.** We ask ourselves the following question: what is the set of unit volume which is closest to the origin in the sense of irrigation? To give this a precise meaning, we embed everything in the space of probability measures; hence we want to minimize the  $d_\alpha$  distance between the unit Dirac mass at  $0 \in \mathbb{R}^d$  and sets  $E$  of unit volume, seen as the uniform measure on  $E$ . This problem reads

$$(S_\alpha) \quad \min \{d_\alpha(\delta_0, \mathbf{1}_E \mathcal{L}) : |E| = 1\},$$

where  $\mathcal{L}$  denotes the Lebesgue measure on  $\mathbb{R}^d$ . We relax this problem by minimizing on a larger set, which is the set of probability measures with Lebesgue density bounded by 1, thus getting:

$$(R_\alpha) \quad \min \{\mathbf{X}_\alpha(\nu) : \nu \leq 1, \nu \in \text{Prob}(\mathbb{R}^d)\},$$

where  $\mathbf{X}_\alpha(\nu) = d_\alpha(\delta_0, \nu)$ .

In the following, we will sometimes encounter positive measures which do not have unit mass, thus we extend the functional by setting  $\mathbf{X}_\alpha(\nu) := d_\alpha(|\nu|\delta_0, \nu)$  for any finite measure  $\nu$ .

A key tool in the analysis of this problem lies in the following proposition, proved in [San07] under slightly more restrictive hypotheses.

**Proposition 1.5** (First variation inequality for  $\mathbf{X}_\alpha$ ). *Suppose that  $\nu \in \text{Prob}(\mathbb{R}^d)$  with  $\mathbf{X}_\alpha(\nu) < \infty$ . Suppose also that  $\eta$  is an optimal irrigation plan between  $|\nu|\delta_0$  and  $\nu$ , with landscape function  $z_\eta$ . The following holds:*

$$\mathbf{X}_\alpha(\tilde{\nu}) \leq \mathbf{X}_\alpha(\nu) + \alpha \int z_\eta d(\tilde{\nu} - \nu)$$

for any  $\tilde{\nu} \in \text{Prob}(\mathbb{R}^d)$ .

Notice also that the integral  $\int z_\eta d(\tilde{\nu} - \nu)$  is well-defined since  $\int z_\eta d\nu = \mathbf{I}_\alpha(\eta) = \mathbf{X}_\alpha(\nu) < \infty$  and  $z_\eta$  is non-negative, though it may be infinite.

*Proof.* If  $\int z_\eta d\tilde{\nu} = \infty$  then there is nothing to prove. Otherwise for  $\nu$ -a.e.  $x$ ,  $z_\eta(x)$  is finite hence there are  $\eta$ -good curves reaching  $x$  and one can find a measurable<sup>3</sup> map  $g : \mathbb{R}^d \rightarrow \Gamma$  which associates with every  $x$  an  $\eta$ -good curve reaching  $x$ . Let us build an irrigation plan  $\tilde{\eta} \in \text{IP}(|\tilde{\nu}|\delta_0, \tilde{\nu})$  which is concentrated on  $\eta$ -good curves, by setting  $\tilde{\eta} = g\#\nu$ , so that

$$\int_\Gamma Z_\eta d\tilde{\eta} = \int_\Gamma z_\eta(\gamma(\infty)) d\tilde{\eta}(\gamma) = \int_{\mathbb{R}^d} z_\eta(x) d\tilde{\nu}.$$

Then, by the first variation inequality for  $\mathbf{I}_\alpha$ , we get:

$$\mathbf{X}_\alpha(\tilde{\nu}) \doteq d_\alpha(|\tilde{\nu}|\delta_0, \tilde{\nu}) \leq \mathbf{I}_\alpha(\tilde{\eta}) \leq \mathbf{I}_\alpha(\eta) + \alpha \int_\Gamma Z_\eta d(\tilde{\eta} - \eta) = \mathbf{X}_\alpha(\nu) + \alpha \int_{\mathbb{R}^d} z_\eta d(\tilde{\nu} - \nu). \quad \square$$

<sup>3</sup>One can characterize  $\eta$ -good curves as those  $\gamma$  such that  $\tilde{Z}_\eta(\gamma) < \infty$  where  $\tilde{Z}_\eta(\gamma) := \int_0^\infty |\dot{\gamma}|_{t,\eta} dt$  is a slight variation of  $Z_\eta$  defined in [San07] which is also lower semicontinuous. Hence the multifunction associating to every  $x$  the set of  $\eta$ -good curves reaching  $x$  can be written as  $\bigcup_{\ell \in \mathbb{Q}} \{\gamma \in \Gamma : \tilde{Z}_\eta(\gamma) \leq \ell, \gamma(\infty) = x\}$ , i.e. as a countable union of multifunctions with closed graph. This means that this multifunction is measurable and admits a measurable selection (see e.g. [CV77]).

## 2. EXISTENCE AND FIRST PROPERTIES

We will often denote by  $C = C(\alpha, d)$  or  $c = c(d)$  different positive constants which depend only on  $\alpha, d$  or  $d$  respectively.

**Theorem 2.1.** *The relaxed shape optimization problem  $(\mathbf{R}_\alpha)$  admits at least a minimizer.*

*Proof.* The existence of a minimizer follows from the lower semicontinuity and tightness. Indeed, any minimizing sequence  $\nu_n$  must have bounded first moment since

$$\int |x| d\nu(x) \leq \int z_\eta(x) d\nu(x) = d_\alpha(\delta_0, \nu).$$

A bound on the first moment implies tightness of the sequence and, up to extracting a subsequence, one has  $\nu_n \rightarrow \nu$ . The condition  $\nu_n \leq 1$  implies  $\nu \leq 1$  and the lower semicontinuity of  $d_\alpha$  provides the optimality of  $\nu$ .  $\square$

For  $1 > \alpha > 1 - \frac{1}{d}$ , we will denote the optimal value for the relaxed shape optimization problem  $(\mathbf{R}_\alpha)$  by:

$$e_\alpha := \min\{d_\alpha(\delta_0, \nu) : \nu \leq 1 \text{ and } \nu \in \text{Prob}(\mathbb{R}^d)\}.$$

**Lemma 2.2** (Scaling lemma). *For any finite measure  $\nu$  we have*

$$\mathbf{X}_\alpha(\nu) \geq e_\alpha |\nu|^{\alpha + \frac{1}{d}}.$$

*Proof.* For  $\lambda = |\nu|^{-1/d}$ , let  $\tilde{\nu} = \lambda^d \varphi_\#(\nu)$  be a scaling of  $\nu$  under the map  $\varphi(x) = \lambda x$  in  $\mathbb{R}^d$ . Then,  $\int_{\mathbb{R}^d} d\tilde{\nu} = \lambda^d \int_{\mathbb{R}^d} d\nu = \lambda^d |\nu| = 1$  and  $\nu \leq 1$ . Thus,

$$e_\alpha \leq d_\alpha(\tilde{\nu}, \delta_0) = \lambda^{\alpha d + 1} d_\alpha(\nu, |\nu| \delta_0) = |\nu|^{-(\alpha + \frac{1}{d})} \mathbf{X}_\alpha(\nu). \quad \square$$

For any  $\nu$ , we say that  $z$  is a landscape function of  $\nu$  if it is the landscape function  $z_\eta$  associated with some optimal irrigation plan  $\eta \in \text{IP}(\delta_0, \nu)$ .

**Theorem 2.3.** *Let  $\nu$  be a minimizer of  $(\mathbf{R}_\alpha)$  and  $z$  a landscape function of  $\nu$ . Then  $\nu$  is the indicator of a set  $A$  which is a sublevel set of  $z$ :*

$$(2.1) \quad A = \{x : z(x) \leq z^*\}, \text{ with } z^* = \frac{e_\alpha}{\alpha} \left( \alpha + \frac{1}{d} \right).$$

*In particular,  $A$  is a solution to problem  $(\mathbf{S}_\alpha)$  and it is a compact and path-connected set.*

*Proof.* We show that  $\nu$  also minimizes the first variation of  $\mathbf{X}_\alpha$ , that is  $\mu \mapsto \int z d\mu$ . Take  $\tilde{\nu}$  a competitor for  $(\mathbf{R}_\alpha)$ . By Proposition 1.5, one has:

$$\mathbf{X}_\alpha(\tilde{\nu}) \leq \mathbf{X}_\alpha(\nu) + \alpha \int z d(\tilde{\nu} - \nu),$$

but  $\mathbf{X}_\alpha(\nu) \leq \mathbf{X}_\alpha(\tilde{\nu})$ , thus

$$\int z d\nu \leq \int z d\tilde{\nu}$$

for any  $\tilde{\nu}$ . So as to minimize this quantity,  $\nu$  must concentrate its mass on the points where  $z$  takes its lowest values. More precisely, there is a value  $z^* \in [0, \infty]$  such that

$$\nu(x) \begin{cases} = 1 & \text{if } z(x) < z^*, \\ \in [0, 1] & \text{if } z(x) = z^*, \\ = 0 & \text{if } z(x) > z^*. \end{cases}$$

Indeed, we just take  $z^* = \sup\{t \in \mathbb{R} : |\{z(x) \leq t\}| < 1\}$ . Since  $\int z d\nu = e_\alpha > 0$ , necessarily  $z^* > 0$ . This kind of arguments is typical in optimization problems under an upper density constraint, as it was for instance done for crowd motion applications in [MRCS14].

**Step 1:**  $z^* \leq \frac{e_\alpha}{\alpha} \left( \alpha + \frac{1}{d} \right)$ .

For  $0 \leq k < z^*$ , we consider the competitor  $\tilde{\nu} = \mathbf{1}_{\{z \leq k\}}$  and set  $|\tilde{\nu}| = 1 - m$ , noting that  $m > 0$  by definition of  $z^*$ . Using Lemma 2.2 and Proposition 1.5, one gets

$$e_\alpha(1 - m)^{\alpha + \frac{1}{d}} \leq \mathbf{X}_\alpha(\tilde{\nu}) \leq \mathbf{X}_\alpha(\nu) + \alpha \int z d(\tilde{\nu} - \nu) = e_\alpha - \alpha \int_{\{z > k\}} z d\nu.$$

Since  $\nu(\{z > k\}) = 1 - |\{z \leq k\}| = m$ , it follows that

$$(2.2) \quad e_\alpha(1 - m)^{\alpha + \frac{1}{d}} \leq e_\alpha - \alpha km.$$

As  $\alpha + \frac{1}{d} > 1$ , the map  $t \mapsto t^{\alpha + \frac{1}{d}}$  is (strictly) convex, thus

$$e_\alpha \left( 1 - \left( \alpha + \frac{1}{d} \right) m \right) \leq e_\alpha(1 - m)^{\alpha + \frac{1}{d}} \leq e_\alpha - \alpha km,$$

hence forgetting the middle term, subtracting  $e_\alpha$  and dividing by  $m$ :

$$\alpha k \leq e_\alpha \left( \alpha + \frac{1}{d} \right).$$

Taking the limit  $k \rightarrow z^*$  yields:

$$(2.3) \quad z^* \leq \frac{e_\alpha}{\alpha} \left( \alpha + \frac{1}{d} \right).$$

**Step 2:**  $\nu = \mathbf{1}_A$  where  $A = \{z \leq z^*\}$ .

Take the competitor  $\tilde{\nu} = \mathbf{1}_{\{z \leq z^*\}}$  and set  $|\tilde{\nu}| = 1 + m$ ,  $m \geq 0$ . Using again the scaling lemma and the first variation of  $\mathbf{X}_\alpha$  one gets:

$$e_\alpha(1 + m)^{\alpha + \frac{1}{d}} \leq e_\alpha + \alpha \int_{z=z^*} z d(\tilde{\nu} - \nu) = e_\alpha + \alpha z^* m.$$

Now by strict convexity of  $t \mapsto t^{\alpha + \frac{1}{d}}$ , if  $m > 0$  then one has

$$e_\alpha(1 + m)^{\alpha + \frac{1}{d}} > e_\alpha \left( 1 + \left( \alpha + \frac{1}{d} \right) m \right),$$

thus

$$e_\alpha \left( \alpha + \frac{1}{d} \right) m < \alpha z^* m,$$

which contradicts (2.3). Consequently  $m = 0$ , hence  $\nu = \tilde{\nu} = \mathbf{1}_{\{z \leq z^*\}}$ .

**Step 3:** Compactness and connectedness.

$A$  is closed since  $z$  is lower semicontinuous and bounded since  $z(x) \geq |x|$  for all  $x \in \mathbb{R}^d$ . It is path-connected since any point  $x$  with  $z(x) \leq z^*$  is the endpoint of an  $\eta$ -good curve  $\gamma$  starting from 0 and  $\gamma \subseteq A$  because  $z$  is increasing along this curve.

**Step 4:**  $z^* \geq \frac{e_\alpha}{\alpha} \left( \alpha + \frac{1}{d} \right)$ .

Take  $x_0 \in A$  with maximal Euclidean norm. Then the half ball  $H_r(x_0) := B_r(x_0) \cap \{x : \langle x - x_0, x_0 \rangle > 0\}$  is included in  $A^c$ . We consider the competitor  $\tilde{\nu} = \mathbf{1}_{A \sqcup H_r(x_0)}$ , with mass  $|\tilde{\nu}| = 1 + m$ , where  $m = |H_r(x_0)| = cr^d$  for some constant  $c = c(d)$ . To irrigate  $\tilde{\nu}$ , we pay at most the cost of irrigation of  $\nu$ , plus the price for moving an extra mass  $m$  from 0 to  $x_0$  along the irrigation plan, plus the cost for moving this mass to  $B_r(x_0) \setminus A$ , which we can bound by  $Cm^\alpha r$  thanks to Proposition 1.4, as follows:

$$\begin{aligned} \mathbf{X}_\alpha(\tilde{\nu}) &= d_\alpha((1 + m)\delta_0, \tilde{\nu}) \leq d_\alpha((1 + m)\delta_0, \nu + m\delta_{x_0}) + d_\alpha(\nu + m\delta_{x_0}, \nu + \mathbf{1}_{H_r(x_0)}) \\ &= \mathbf{X}_\alpha(\nu + m\delta_{x_0}) + d_\alpha(m\delta_{x_0}, \mathbf{1}_{H_r(x_0)}) \\ &\leq e_\alpha + \alpha m z(x_0) + Crm^\alpha, \end{aligned}$$



where  $C = C(\alpha, d)$  is some positive constant. Combining this inequality with the following convexity inequality

$$\mathbf{X}_\alpha(\tilde{\nu}) \geq e_\alpha(1+m)^{\alpha+\frac{1}{d}} \geq e_\alpha \left( 1 + \left( \alpha + \frac{1}{d} \right) m \right),$$

and dividing by  $m > 0$ , one gets:

$$e_\alpha \left( \alpha + \frac{1}{d} \right) \leq \alpha z(x_0) + Cr^{1+d\alpha-d}.$$

Passing to the limit  $r \rightarrow 0$ , we obtain

$$z^* \geq z(x_0) \geq \frac{e_\alpha}{\alpha} \left( \alpha + \frac{1}{d} \right). \quad \square$$

### 3. HÖLDER CONTINUITY OF THE LANDSCAPE FUNCTION

The Hölder regularity of the landscape function has been proved in [San07] under some regularity assumptions on  $\nu$  using Campanato spaces (these spaces were introduced in [Cam63], see [Giu03, Section 2.3] for a modern exposition). Namely, if  $\nu$  is of the form  $\nu = f \mathcal{L}_{\perp E}$  where the density  $f(x)$  and the fraction of mass  $\Theta_E(x, r) := \frac{|E \cap B_r(x)|}{|B_r(x)|}$  lying in  $E$  are bounded from below by some constant  $c > 0$  for all  $x \in E$  and all  $r \leq \text{diam } E$ , then  $z$  is  $\beta$ -Hölder continuous where

$$\beta := d \left( \alpha - \left( 1 - \frac{1}{d} \right) \right) = 1 + d\alpha - d,$$

is a number which is strictly between 0 and 1 as  $1 > \alpha > 1 - \frac{1}{d}$ . Another proof for more general regularity assumptions on  $\nu$  has been given in [BS11]. In our case, we do not know a priori that  $A$  is regular (on the contrary we suspect it has a fractal boundary), hence the Hölder regularity of  $z$  does not follow from previous works. Exploiting the fact that  $A$  is optimal, we are going to show that  $z$  is  $\beta$ -Hölder continuous adapting classical computations to pass from Campanato to Hölder spaces. More precisely, setting  $A_r(x) := A \cap B_r(x)$  and  $z_r(x)$  the mean of  $z$  on  $A_r(x)$ , we are going to prove the following sequence of inequalities, for arbitrary  $r \leq \text{diam } A$ :

$$(3.1) \quad \int_{A_r(x)} |z - z_r(x)| \leq Cr^\beta,$$

$$(3.2) \quad |z_r(x) - z_{r/2}(x)| \leq Cr^\beta,$$

$$(3.3) \quad z_r(x) - z(x) \leq Cr^\beta,$$

$$(3.4) \quad |z(x) - z_r(x)| \leq Cr^\beta,$$

$$(3.5) \quad |z_{|y-x|}(x) - z_{|y-x|}(y)| \leq C|y-x|^\beta.$$

Notice that the two last inequalities imply that  $z$  is indeed  $\beta$ -Hölder continuous:

$$\begin{aligned} |z(y) - z(x)| &\leq |z(y) - z_{|y-x|}(y)| + |z_{|y-x|}(x) - z_{|y-x|}(y)| + |z(x) - z_{|y-x|}(x)| \\ &\leq C|y-x|^\beta. \end{aligned}$$

The main difficulty we will encounter is that we will quite easily obtain estimates of the form

$$\dots \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}},$$

and will need to get rid of the term  $\Theta_A(x, r)^{1-\alpha}$ , i.e. treat the case when it becomes small.

**3.1. Main lemmas.** The following lemma will be key to prove the regularity of the landscape function.

**Lemma 3.1** (Maximum deviation). *There is a constant  $C = C(d, \alpha) > 0$  such that the following holds:*

$$(3.6) \quad \forall y \in A_r(x), \quad z^* - z(y) \leq C \frac{r^\beta}{\Theta_{A^c}(x, r)^{1-\alpha}}.$$

*Proof.* We consider the competitor  $\tilde{\nu} = \mathbf{1}_{A \cup B_r(x)}$ , with mass  $|\tilde{\nu}| = 1 + m$  where  $m = |B_r(x) \setminus A|$ . For any  $y \in A_r(x)$ , let us irrigate  $\tilde{\nu}$  from 0 by irrigating  $\nu$  from 0, moving an extra mass  $m$  from 0 to  $y$  along the irrigation plan, then irrigating  $\mathbf{1}_{B_r(x) \setminus A}$  from this mass at  $y$ . Using Lemma 2.2 and Proposition 1.4, we have

$$e_\alpha(1+m)^{\alpha+\frac{1}{d}} \leq \mathbf{X}_\alpha(\tilde{\nu}) \leq \mathbf{X}_\alpha(\nu) + \alpha \int z \, d(m\delta_y) + Crm^\alpha.$$

By convexity,

$$e_\alpha \left( 1 + \left( \alpha + \frac{1}{d} \right) m \right) \leq (1+m)^{\alpha+\frac{1}{d}} e_\alpha \leq e_\alpha + \alpha m z(y) + Crm^\alpha,$$

thus, knowing that  $e_\alpha \left( \alpha + \frac{1}{d} \right) = \alpha z^*$  by (2.1):

$$\alpha m z^* \leq \alpha m z(y) + Crm^\alpha.$$

By definition,  $m = \omega_d r^d \Theta_{A^c}(x, r)$  where  $\omega_d$  is the volume on the unit  $d$ -dimensional ball, hence

$$z^* - z(y) \leq Cr(r^d \Theta_{A^c}(x, r))^{\alpha-1} = C \frac{r^\beta}{\Theta_{A^c}(x, r)^{1-\alpha}}. \quad \square$$

*Remark 3.1.* One can see that if  $\Theta_A(x, r)$  becomes small, then  $\Theta_{A^c}(x, r)$  is large (close to 1), and actually all values of  $z$  in  $A_r(x)$  become close to the same value  $z^*$  up to  $Cr^\beta$ .

**Lemma 3.2** (Mean deviation). *There is some constant  $C = C(d, \alpha) > 0$  such that*

$$\int_{A_r(x)} |z(y) - z_r(x)| \, dy \leq Cr^\beta$$

for all  $r > 0$  and all  $x \in A$ .

*Proof.* We will first show that

$$\int_{A_r(x)} |z(y) - z_r(x)| \, dy \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Denoting by  $\bar{z}_r(x)$  the central median of  $z$  on the set  $A_r(x)$ , there is a disjoint union  $A_r(x) = A^- \sqcup A^+$  such that  $|A^-| = |A^+| = \frac{|A_r(x)|}{2}$  and  $z \leq \bar{z}_r(x)$  on  $A^-$ ,  $z \geq \bar{z}_r(x)$  on  $A^+$ . Let us consider the competitor  $\tilde{\nu} = \mathbf{1}_A - \mathbf{1}_{A^+} + \mathbf{1}_{A^-}$ . By the first variation lemma:

$$\mathbf{X}_\alpha(\tilde{\nu}) \leq \mathbf{X}_\alpha(\nu) + \alpha \int z \, d(\tilde{\nu} - \nu).$$

Recall that  $\mathbf{X}_\alpha(\rho) = d_\alpha(\delta_0, \rho)$  when  $\rho$  is a probability measure, which is the case for  $\nu$  and  $\tilde{\nu}$ , and that  $d_\alpha$  is a distance. Thus by the triangle inequality:

$$\alpha \int z \, d(\nu - \tilde{\nu}) \leq d_\alpha(\delta_0, \nu) - d_\alpha(\delta_0, \tilde{\nu}) \leq d_\alpha(\nu, \tilde{\nu}).$$

We know that  $d_\alpha(\nu, \tilde{\nu}) \leq C|\tilde{\nu} - \nu|^\alpha \text{diam}(\text{supp}(\tilde{\nu} - \nu)) \leq C|A_r(x)|^\alpha r$  for some  $C = C(\alpha, d) > 0$ . Moreover notice that

$$\begin{aligned} \int z \, d(\nu - \tilde{\nu}) &= \int_{A^+} z(y) \, dy - \int_{A^-} z(y) \, dy \\ &= \int_{A^+} (z(y) - \bar{z}_r(x)) \, dy + \int_{A^-} (\bar{z}_r(x) - z(y)) \, dy \\ &= \int_{A_r(x)} |z(y) - \bar{z}_r(x)| \, dy. \end{aligned}$$

Consequently:

$$\int_{A_r(x)} |z(y) - \bar{z}_r(x)| \, dy \leq C|A_r(x)|^{\alpha-1} r \leq C \frac{|A_r(x)|^{\alpha-1}}{|B_r(x)|^{\alpha-1}} r^{1+d(\alpha-1)} = C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Moreover, one has

$$|z_r(x) - \bar{z}_r(x)| = \left| \int_{A_r(x)} z(y) - \bar{z}_r(x) \, dy \right| \leq \int_{A_r(x)} |z(y) - \bar{z}_r(x)| \, dy \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}$$

which leads to

$$\int_{A_r(x)} |z(y) - z_r(x)| \, dy = \int_{A_r(x)} |z(y) - \bar{z}_r(x)| \, dy + |z_r(x) - \bar{z}_r(x)| \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Now we get rid of  $\Theta_A(x, r)^{1-\alpha}$ . If  $\Theta_A(x, r) \geq 1/2$ , we get the desired inequality. On the other hand, if  $\Theta_{A^c}(x, r) \geq 1/2$ , by Lemma 3.1, we have

$$0 \leq z^* - z(y) \leq Cr^\beta, \quad \forall y \in A_r(x),$$

which also implies that

$$0 \leq z^* - z_r(x) \leq Cr^\beta.$$

By these two inequalities, we have

$$|z(y) - z_r(x)| \leq Cr^\beta, \quad \forall y \in A_r(x).$$

Now, taking the mean over  $A_r(x) \ni y$  leads to the wanted inequality as well:

$$\int_{A_r(x)} |z(y) - z_r(x)| \, dy \leq Cr^\beta. \quad \square$$

*Remark 3.2.* Notice that the estimate

$$\int_{A_r(x)} |z(y) - z_r(x)| \, dy \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}$$

is valid in general: we only use the fact that  $\nu$  is an indicator function (a density bounded from below would suffice). The optimality of  $\nu$  comes into play to get rid of  $\Theta_A(x, r)$ .

### 3.2. Hölder regularity.

**Proposition 3.3** (Small-scale difference). *For all  $x \in A$  and all  $r > 0$  one has*

$$|z_r(x) - z_{r/2}(x)| \leq Cr^\beta.$$

*Proof.* First we show that

$$|z_r(x) - z_{r/2}(x)| \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Indeed, by Lemma 3.2,

$$\begin{aligned} |z_r(x) - z_{r/2}(x)| &\leq \frac{\int_{A_{r/2}(x)} |z(y) - z_r(x)| dy}{|A_{r/2}(x)|} \leq \frac{|A_r(x)|}{|A_{r/2}(x)|} \int_{A_r(x)} |z(y) - z_r(x)| dy \\ &\leq 2^d \frac{\Theta_A(x, r)}{\Theta_A(x, r/2)} Cr^\beta \leq C \frac{r^\beta}{\Theta_A(x, r/2)}. \end{aligned}$$

As before, if  $\Theta_A(x, r/2) \geq 1/2$  we get the desired estimate. Otherwise  $\Theta_{A^c}(x, r/2) \geq 1/2$  and  $\Theta_{A^c}(x, r) \geq 2^{-d}\Theta_{A^c}(x, r/2) \geq 2^{-1-d}$ . Now, by Lemma 3.1,

$$0 \leq z^* - z(y) \leq C \frac{r^\beta}{\Theta_{A^c}(x, r)} \leq Cr^\beta, \quad \forall y \in A_r(x).$$

Consequently  $0 \leq z^* - z_{r/2}(x) \leq Cr^\beta$  and  $0 \leq z^* - z_r(x) \leq Cr^\beta$  which implies that

$$|z_r(x) - z_{r/2}(x)| \leq Cr^\beta. \quad \square$$

**Lemma 3.4** (Lower deviation to the mean). *There is a constant  $C = C(d, \alpha) > 0$  such that for all  $x \in A$  and all  $r > 0$  one has:*

$$(3.7) \quad \forall y \in A_r(x), \quad z_r(x) - z(y) \leq Cr^\beta.$$

*Proof.* First we show that

$$z_r(x) - z(y) \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Remove the mass  $m = |A_r(x)|$  going to  $A_r(x)$  from the irrigation plan, make it travel along the plan to any fixed  $y \in A_r(x)$  and then send it to  $A_r(x)$ : this should cost more. This implies

$$\alpha m z(y) - \alpha \int_{A_r(x)} z + Cm^\alpha r \geq 0,$$

which may be rewritten as

$$z_r(x) - z(y) \leq Cm^{\alpha-1} r \leq C \frac{r^\beta}{\Theta_A(x, r)^{1-\alpha}}.$$

Now if  $\Theta_A(x, r) \geq 1/2$  one gets the desired result. Otherwise  $\Theta_{A^c}(x, r) \geq 1/2$  and Lemma 3.1 yields:

$$0 \leq z^* - z(y) \leq Cr^\beta, \quad \forall y \in A_r(x).$$

Thus  $0 \leq z^* - z_r(x) \leq Cr^\beta$  and for any fixed  $y \in A_r(x)$ ,

$$|z_r(x) - z(y)| \leq |z_r(x) - z^*| + |z^* - z(y)| \leq Cr^\beta,$$

from which we also get the wanted inequality.  $\square$

**Lemma 3.5** (Deviation to the mean). *For all  $x \in A$  and all  $r > 0$ , one has*

$$|z(x) - z_r(x)| \leq Cr^\beta.$$

*Proof.* By Proposition 3.3, one has

$$|z(x) - z_r(x)| \leq |z(x) - z_{r/2}(x)| + |z_{r/2}(x) - z_r(x)| \leq |z(x) - z_{r/2}(x)| + Cr^\beta,$$

which means by setting  $f(r) = |z(x) - z_r(x)|$  for  $r > 0$  that:

$$f(r) \leq f(r/2) + Cr^\beta.$$

Consequently for all  $k \in \mathbb{N}$

$$f(r) \leq f(r \cdot 2^{-(k+1)}) + Cr^\beta \sum_{i=0}^k 2^{-i\beta}$$

thus

$$f(r) \leq \limsup_{\varepsilon \rightarrow 0} f(\varepsilon) + Cr^\beta \sum_{i=0}^{\infty} 2^{-i\beta} \leq \limsup_{\varepsilon \rightarrow 0} f(\varepsilon) + Cr^\beta.$$

Now let us prove that  $f(\varepsilon) \rightarrow 0$  when  $\varepsilon \rightarrow 0$ , i.e.  $z_\varepsilon(x) \xrightarrow{\varepsilon \rightarrow 0} z(x)$ . We already know that  $z$  is lower semi-continuous hence  $z(x) \leq \liminf_{\varepsilon \rightarrow 0} z_\varepsilon(x)$ . Moreover using (3.7), we have

$$\limsup_{\varepsilon \rightarrow 0} z_\varepsilon(x) \leq \limsup_{\varepsilon \rightarrow 0} (z(x) + C\varepsilon^\beta) = z(x),$$

which implies that  $z_\varepsilon(x) \rightarrow z(x)$  when  $\varepsilon \rightarrow 0$ . Therefore the inequality  $f(r) \leq Cr^\beta$  holds, that is to say:

$$|z(x) - z_r(x)| \leq Cr^\beta. \quad \square$$

**Lemma 3.6** (Large scale difference). *For any  $x, y \in A$ , one has:*

$$|z_{|y-x|}(x) - z_{|y-x|}(y)| \leq C|y-x|^\beta.$$

*Proof.* Set  $r = |y-x|$ , and  $\Delta_r = B_r(x) \cap B_r(y)$ . Notice that,  $\Delta_r$  being a fixed fraction of  $B_r(x)$  (independant of  $r$ ),  $|\Delta_r| = c|B_r|$  for some  $c = c(d) \in (0, 1)$ .

If both  $\Theta_{A^c}(x, r) \geq \frac{c}{2}$  and  $\Theta_{A^c}(y, r) \geq \frac{c}{2}$ , then by Lemma 3.1 one has:

$$0 \leq z^* - z_r(x) \leq C \frac{r^\beta}{\Theta_{A^c}(x, r)^{1-\alpha}} \leq Cr^\beta, \text{ and } 0 \leq z^* - z_r(y) \leq C \frac{r^\beta}{\Theta_{A^c}(y, r)^{1-\alpha}} \leq Cr^\beta,$$

which implies the desired inequality

$$(3.8) \quad |z_r(x) - z_r(y)| \leq Cr^\beta.$$

On the other hand, if either  $\Theta_{A^c}(x, r)$  or  $\Theta_{A^c}(y, r)$  is less than  $c/2$ , say  $\Theta_{A^c}(x, r) \leq \frac{c}{2}$ , we claim the desired inequality (3.8) still holds. Indeed, for all  $u \in A_r(x) \cap A_r(y)$  one has

$$|z_r(x) - z_r(y)| \leq |z_r(x) - z(u)| + |z_r(y) - z(u)|$$

thus integrating over  $A_r(x) \cap A_r(y)$  in  $u$  one gets:

$$\begin{aligned} |z_r(x) - z_r(y)| &\leq \frac{1}{|A_r(x) \cap A_r(y)|} \left[ \int_{A_r(x)} |z(u) - z_r(x)| du + \int_{A_r(y)} |z(u) - z_r(y)| du \right] \\ &\leq Cr^\beta \frac{|A_r(x)| + |A_r(y)|}{|A_r(x) \cap A_r(y)|}, \end{aligned}$$

the last inequality resulting from Lemma 3.2. Note that

$$|A_r(x) \cap A_r(y)| = |\Delta_r \cap A| \geq |\Delta_r| - |B_r(x) \setminus A| = c|B_r(x)| - |B_r(x) \setminus A|$$

which implies that

$$\frac{|A_r(x)| + |A_r(y)|}{|A_r(x) \cap A_r(y)|} \leq \frac{2|B_r(x)|}{c|B_r(x)| - |B_r(x) \setminus A|} = \frac{2}{c - \Theta_{A^c}(x, r)} \leq \frac{4}{c}.$$

Thus, in this case, we still have

$$|z_r(x) - z_r(y)| \leq Cr^\beta \frac{|A_r(x)| + |A_r(y)|}{|A_r(x) \cap A_r(y)|} \leq Cr^\beta. \quad \square$$

**Theorem 3.7** (Hölder continuity). *The function  $z$  is  $\beta$ -Hölder continuous on  $A$ . More precisely:*

$$\forall x, y \in A, \quad |z(y) - z(x)| \leq C|y-x|^\beta,$$

for some constant  $C = C(\alpha, d)$ .

*Proof.* By Lemma 3.5 and Lemma 3.6,

$$|z(y) - z(x)| \leq |z(y) - z_{|y-x|}(y)| + |z_{|y-x|}(y) - z_{|y-x|}(x)| + |z(y) - z_{|y-x|}(y)| \leq 3C|y-x|^\beta. \quad \square$$

As a consequence of this result we may quantify the minimal size of a ball one can put inside  $A$  around  $x$  in terms of  $z^* - z(x)$  and prove that  $A$  has non-empty interior.

**Proposition 3.8** (Interior points). *For some constant  $C = C(\alpha, d)$  the following holds:*

$$(3.9) \quad \forall x \in A, \quad B_{r(x)}(x) \subseteq A,$$

where  $r(x) := C(z^* - z(x))^{1/\beta} \geq 0$ . In particular

$$\{x \in A : z(x) < z^*\} \subseteq \overset{\circ}{A} \text{ and } \partial A \subseteq \{x \in A : z(x) = z^*\}.$$

*Proof.* It suffices to prove (3.9) for  $x_0 \in A$  satisfying  $z(x_0) < z^*$ . Consider a point  $x \in A^c$ . Take a point  $y \in A$  which is closest to  $x$ : it is possible since  $A$  is compact. By Lemma 3.1, we know that for small  $r$

$$\Theta_{A^c}(y, r)^{1-\alpha}(z^* - z(y)) \leq Cr^\beta.$$

But by construction  $y$  is such that  $\liminf_{r \rightarrow 0} \Theta_{A^c}(y, r) \geq 1/2$ , and since the right-hand side tends to 0, necessarily  $z(y) = z^*$ . By the Hölder continuity of  $z$  stated in Theorem 3.7,

$$z^* - z(x_0) = |z(y) - z(x_0)| \leq C|y - x_0|^\beta \leq C|x - x_0|^\beta,$$

where the last inequality follows from the fact that  $|y - x_0| \leq |y - x| + |x - x_0| \leq 2|x - x_0|$  because  $y$  minimizes the distance from  $x$ . Hence, for all  $x \in A^c$ ,  $|x - x_0| \geq C(z^* - z(x_0))^{1/\beta} = r(x_0)$ , which implies the desired result.  $\square$

#### 4. ON THE DIMENSION OF THE BOUNDARY

We are interested in the dimension of the boundary  $\partial A$ , our guess being that it should be non-integer, and lie between  $d - 1$  and  $d$ . Here we look at the Minkowski dimension (also called box-counting dimension). Given a set  $X$ , we denote by  $N_\varepsilon(X)$  the maximum amount of disjoint balls of radius  $\varepsilon$  centered at points of  $X$ .

**Definition 4.1** (Minkowski dimension). We define the upper Minkowski dimension of  $X$  by

$$\overline{\dim}_M(X) = \limsup_{\varepsilon \rightarrow 0} \frac{\log(N_\varepsilon(X))}{-\log \varepsilon},$$

and the lower Minkowski dimension by

$$\underline{\dim}_M(X) = \liminf_{\varepsilon \rightarrow 0} \frac{\log(N_\varepsilon(X))}{-\log \varepsilon}.$$

When these coincide we just call it the Minkowski dimension and denote it by  $\dim_M(X)$ .

We shall get an upper bound on the upper Minkowski dimension. We say that  $X$  is of dimension smaller than  $\delta$  if  $\overline{\dim}_M X \leq \delta$ .

**Lemma 4.1.** *There is a constant  $C = C(\alpha, d)$  such that for all  $k \leq z^*$ ,*

$$|\{x \in A : k < z(x) \leq z^*\}| \leq C(z^* - k).$$

*Proof.* Consider the competitor  $\tilde{\nu} = \mathbf{1}_{\{z \leq k\}}$  with total mass  $|\tilde{\nu}| = 1 - m$ , where  $m = |\{x \in A : k < z(x) \leq z^*\}|$ . As in (2.2), one has

$$e_\alpha(1 - m)^{\alpha+1/d} \leq e_\alpha - \alpha km$$

hence knowing that  $\alpha z^* = (\alpha + 1/d)e_\alpha$  and developing the term on the left-hand side at order 2, we obtain:

$$-\alpha m z^* + \frac{e_\alpha}{2} \left(\alpha + \frac{1}{d}\right) \left(\alpha + \frac{1}{d} - 1\right) m^2 \leq -\alpha km$$

Thus

$$m \leq C(z^* - k)$$

with  $1/C = e_\alpha(\alpha + 1/d)(\alpha + 1/d - 1)/(2\alpha)$ .  $\square$

**Theorem 4.2.** *The set  $\partial A$  is of dimension less than  $d - \beta$ .*

*Proof.* For  $\varepsilon > 0$  fixed, take disjoint balls  $(B_i)_{i \in I}$  of radius  $\varepsilon$ , where  $N := |I| = N_\varepsilon(\partial A)$ . We set  $B_i^+ = B_i \setminus A$ ,  $B_i^- = B_i \cap A$ . We split the set of balls into two parts: those which have a larger intersection with  $A$  rather than  $A^c$ , and vice-versa. Namely, we set

$$\begin{aligned} I^+ &= \{i \in I : |B_i^+| \geq |B_i|/2\}, & N^+ &= |I^+|, \\ I^- &= \{i \in I : |B_i^-| \geq |B_i|/2\}, & N^- &= |I^-|, \end{aligned}$$

so that  $I = I^+ \cup I^-$  and  $N \leq N^+ + N^-$ . We are going to bound  $N^+$  and  $N^-$  by some power of  $\varepsilon$ .

**Step 1:** Bound on  $N^-$ .

Since  $z$  is  $\beta$ -Hölder continuous on  $A$ , one has for each  $B_i = B_\varepsilon(x_i)$ :

$$\forall x \in B_i \cap A, \quad |z(x) - z^*| < C\varepsilon^\beta,$$

since the center  $x_i$  lies in  $\partial A \subseteq \{z = z^*\}$  according to Proposition 3.8. Consequently

$$(\partial A)_\varepsilon \cap A \subseteq \{z^* - C\varepsilon^\beta < z \leq z^*\},$$

thus because of Lemma 4.1:

$$|(\partial A)_\varepsilon \cap A| \leq |\{z^* - C\varepsilon^\beta < z \leq z^*\}| \leq C\varepsilon^\beta.$$

Using the previous inequality and the fact that  $|B_i^-| \geq |B_i|/2 \geq C\varepsilon^d$  for  $i \in I^-$ , one has:

$$CN^-\varepsilon^d \leq \sum_{i \in I^-} |B_i^-| \leq |(\partial A)_\varepsilon \cap A| \leq C\varepsilon^\beta,$$

which implies:

$$(4.1) \quad N^- \leq C\varepsilon^{-(d-\beta)}.$$

**Step 2:** Bound on  $N^+$ .

We consider the competitor  $\tilde{\nu} = \mathbf{1}_{\tilde{A}}$  where  $\tilde{A} = A \cup \bigcup_{i \in I^+} B_i^+$ . It has a mass  $|\tilde{\nu}| = 1 + m$  where  $m = \sum_{i \in I^+} |B_i^+|$ . To irrigate  $\tilde{\nu}$ , we send an extra mass  $|B_i^+|$  to each center  $x_i$  along the irrigation plan, which costs  $\alpha|B_i^+|z^*$ , then we send this mass towards  $B_i^+$ , which costs at most  $C|B_i^+|^\alpha\varepsilon$ . But one should get a cost no less than  $e_\alpha(1+m)^{\alpha+1/d}$  by the scaling lemma. Moreover, with a development of order 2 one has:

$$\begin{aligned} (1+m)^{\alpha+1/d} &\geq 1 + (\alpha+1/d)m + 1/2 \cdot (\alpha+1/d)(\alpha+1/d-1)(1+m)^{\alpha+1/d-2}m^2 \\ &\geq 1 + (\alpha+1/d)m + Cm^2 \end{aligned}$$

because for  $\varepsilon$  small,  $1+m$  is less than 2 for example. Consequently one may say:

$$e_\alpha(1 + (\alpha+1/d)m + Cm^2) \leq e_\alpha + \alpha mz^* + \sum_{i \in I^+} C\varepsilon|B_i^+|^\alpha.$$

Recall that  $\alpha z^* = e_\alpha(\alpha+1/d)$ , thus after simplifying one gets for some  $C > 0$ :

$$(4.2) \quad m^2 \leq C \sum_{i \in I^+} |B_i^+|^\alpha \varepsilon \leq CN^+\varepsilon^{1+\alpha d}.$$

Notice that for  $i \in I^+$ ,  $|B_i^+| \geq |B_i|/2 \geq C\varepsilon^d$ , so that

$$m = \sum_{i \in I^+} |B_i^+| \geq CN^+\varepsilon^d.$$

Injecting this into (4.2), one gets:

$$(N^+\varepsilon^d)^2 \leq CN^+\varepsilon^{1+\alpha d},$$

thus

$$(4.3) \quad N^+ \leq C\varepsilon^{1+\alpha d-2d} = C\varepsilon^{-(d-\beta)}.$$

Putting (4.1) and (4.3) together yields:

$$N_\varepsilon(\partial A) = N \leq N^+ + N^- \leq \frac{C}{\varepsilon^{d-\beta}},$$

and

$$\overline{\dim}_M(\partial A) = \limsup_{\varepsilon \rightarrow 0} \frac{\log(N_\varepsilon(\partial A))}{-\log(\varepsilon)} \leq d - \beta,$$

which means that  $\partial A$  is of dimension smaller than  $d - \beta$ .  $\square$

This result pushes us to propose the following conjecture:

**Conjecture 4.3.** *The boundary  $\partial A$  is of dimension  $d - \beta$ , in the sense that:*

$$\dim_H(\partial A) = \dim_M(\partial A) = d - \beta.$$

Proving this requires to establish the inequality  $\dim_H(\partial A) \geq d - \beta$ , for which we do not have a working strategy yet.

## 5. NUMERICAL SIMULATIONS

Our goal now is to compute solutions to our shape optimization problem numerically. To perform numerical simulations, we use the Eulerian framework of branched transport, first defined by the third author in [Xia03]. This framework is based on vector measures with a measure divergence, i.e. measures  $v \in \mathcal{M}^d(\mathbb{R}^d)$  such that  $\nabla \cdot v \in \mathcal{M}(\mathbb{R}^d)$ , the set of such measures being denoted by  $\mathcal{M}_{div}(\mathbb{R}^d)$ . The cost is the so-called  $\alpha$ -mass:

$$M_\alpha(v) = \begin{cases} \int \left| \frac{dv}{d\mathcal{H}^1}(x) \right|^\alpha d\mathcal{H}^1(x) & \text{if } v \text{ is 1-rectifiable,} \\ +\infty & \text{otherwise.} \end{cases}$$

An elliptic approximation of this functional was introduced by Oudet and the second author in [OS11] (see also [San10]), in the spirit of Modica and Mortola [MM77]. The approximate functional is defined for  $\varepsilon > 0$  by:

$$M_\varepsilon^\alpha(v) = \varepsilon^{-\sigma_1} \int |v(x)|^\sigma dx + \varepsilon^{\sigma_2} \int \frac{|v(x)|^2}{2} dx$$

for suitably chosen  $\sigma, \sigma_1, \sigma_2$ . It is proven in [OS11] that  $M_\varepsilon^\alpha$   $\Gamma$ -converges to  $M^\alpha$  as  $\varepsilon$  goes to 0, for a suitable topology on  $\mathcal{M}_{div}(\mathbb{R}^d)$ . Moreover, the  $\Gamma$ -convergence result also holds imposing an equality constraint on the divergence  $\nabla \cdot v = f_\varepsilon$ , for a suitable sequence  $f_\varepsilon \rightarrow f$ , as proven in [Mon17]. The results of [OS11] are proven in dimension  $d = 2$ , but in [Mon15] there is a proof of how to extend to higher dimension, in the case  $\alpha > 1 - 1/d$  (in dimension  $d = 2$  there is also a version of the  $\Gamma$ -convergence result for  $\alpha \leq 1/2$ ). Also note that, recently, other phase-field approximations for branched transport or other network problems have been studied, see for instance [BOO16; FCM16; BLS15].

Here we adapt the approach of [OS11] to our shape optimization problem by adding this time an inequality constraint on the divergence.

Recall that the Lagrangian and Eulerian frameworks are equivalent [Peg17], so that the irrigation distance may be computed in the following way:

$$d_\alpha(\mu, \nu) = \inf_v \{M^\alpha(v) : \nabla \cdot v = \mu - \nu\}.$$

Consequently the shape optimization problem  $(R_\alpha)$  rewrites, in relaxed form, as:

$$(ES) \quad \min_v \{M^\alpha(v) : \mu - 1 \leq \nabla \cdot v \leq \mu\} \quad \text{where } \mu = \delta_0.$$

Setting  $a = \mu - 1, b = \mu$  and some mollified versions  $a_\varepsilon = \mu_\varepsilon - 1, b_\varepsilon = \mu_\varepsilon$ , for example a convolution of  $\mu$  with the standard mollifier of suitable size  $r_\varepsilon$  (e.g.  $\varepsilon^{\sigma_2} r_\varepsilon^{-d} = o(1)$ ) as in [Mon17]), we define the following approximate problem, for  $\varepsilon > 0$ :

$$(AS) \quad \min_v \{M_\varepsilon^\alpha(v) : a_\varepsilon \leq \nabla \cdot v \leq b_\varepsilon\}.$$



Let us remark that the above-mentioned  $\Gamma$ -convergence results do not allow us to say that this problem approximates (ES), as the inequality constraint on the divergence is not directly in these works. We leave this question for further investigation, as our aim is for now to make a first attempt to compute numerically an optimal shape for the original problem (S $_{\alpha}$ ).

**5.1. Optimization methods.** We tackle problem (AS) by descent methods. Two difficulties arise: first of all, the functional  $M_{\varepsilon}^{\alpha}$  is not convex hence there is no guarantee that the methods converge, and if they do, they may converge to a local minimizer which is not necessarily a global minimizer ; secondly, this is a constrained problem, hence we will need to compute projections or resort to proximal methods to handle the constraint. The simplest approach is to use a first-order method, for instance to perform a projected gradient descent on the functional  $M_{\varepsilon}^{\alpha}$  for  $\varepsilon$  fixed (but small):

**The projected gradient method.**

$$\begin{cases} v_0 \in C \\ v_{n+1} = p_C(v_n - \tau_n \nabla M_{\varepsilon}^{\alpha}(v_n)), \end{cases}$$

where

$$C = \{v : a_{\varepsilon} \leq \nabla \cdot v \leq b_{\varepsilon}\}$$

is the convex set of admissible vector fields for (AS).

Computing the projection  $p_C$  is not an easy task, even more so as we want fast computations since this projection should be done at each step of the algorithm. Actually, this projection step will be quite costly (at least in our approach), hence we need to pass to a higher order method to get to an approximate minimizer in a reasonable number of iterations.

Recall that the projected gradient method is a particular case of the proximal gradient method, which we describe briefly. Consider a problem of the form

$$\min_v f(v) + g(v)$$

where  $f$  is smooth and  $g$  “proximable”, in the sense that one may easily compute its proximal operator

$$\text{prox}_g^{\tau}(v) = \underset{v'}{\text{argmin}} g(v') + \frac{1}{2\tau} |v' - v|^2.$$

The proximal gradient method consists in doing at each step an explicit descent for  $f$  and an implicit descent for  $g$ :

**The proximal gradient method.**

$$\begin{cases} v_0 \text{ given} \\ v_{n+1} = \text{prox}_g^{\tau_n}(v_n - \tau_n \nabla f(v_n)). \end{cases}$$

The projected gradient method corresponds to the case

$$g(v) = \begin{cases} 0 & \text{if } v \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

If there was no function  $g$ , we recover the classical gradient descent method. Notice that there is an implicit choice in this method, since we compute gradients which depend on the scalar product. There is no reason that the canonical scalar product is well adapted to the function we want to minimize. Following the work of Lee, Sun and Saunders [LSS14] on Newton-type proximal methods, one may “twist” the scalar product, leading to the more general method:

**A “twisted” proximal gradient method.**

$$(5.1) \quad \begin{cases} v_0 \text{ given} \\ v_{n+1} = \text{prox}_g^{\tau_n, H_n}(v_n - \tau_n \nabla_{H_n} f(v_n)), \end{cases}$$

where  $\nabla_H f(x)$  is the gradient of  $f$  with respect to the scalar product  $\langle x, y \rangle_H = \langle Hx, y \rangle$  for  $H$  an invertible self-adjoint operator, and

$$\text{prox}_g^{\tau, H}(v) = \underset{v'}{\text{argmin}} g(v') + \frac{1}{2\tau} \|v' - v\|_H^2.$$

The best quadratic model of  $f$  around a point  $x_0$  is

$$Qf_{x_0}(x) = f(x_0) + \langle \nabla_H f(x_0), x \rangle_H + 1/2 \langle x, x \rangle_H,$$

with  $H = H_f(x_0)$  being the Hessian of  $f$  at  $x$ , thus it is natural to consider (5.1) with  $H_n = H_f(x_n)$ . Notice indeed that if  $g$  is zero, the proximal operator is the identity and that  $\nabla_H f(v) = H^{-1} \nabla f$ , so that one recovers Newton’s method:

$$v_{n+1} = v_n - \tau_n H_n^{-1} \nabla f(v_n),$$

which is known to converge quadratically for smooth enough  $f$ . This is why this method is called *proximal Newton method*. However, for large-scale problems, computing and storing the Hessian is very costly, thus an alternative is to set  $H_n$  to be an approximation of the Hessian of  $f$  at  $v_n$ , thus leading to *proximal quasi-Newton methods*. These methods were introduced in [LSS14], which we refer to for further detail and theoretical results of convergence.

A very popular choice for  $H_n$  is given by the L-BFGS method (see [LN89]), which is a quasi-Newton method building in some sense the “best” approximation of the Hessian at  $v_n$  using only the information of the points  $v_k$  and the gradients  $\nabla f(v_k)$  for a fixed number of previous steps  $k = n, n-1, \dots, n-L+1$ . The interest is that no matrix is stored, and there is a very efficient way to compute the matrix-vector product  $H_n^{-1} \cdot v$  using simple algebra. Therefore, we decided to implement a proximal L-BFGS method, which in our case reads:

**The projected L-BFGS method.**

$$(5.2) \quad \begin{cases} v_0 \text{ given} \\ v_{n+1} = p_C^{\tilde{H}_n}(v_n - \tau_n \tilde{H}_n^{-1} \nabla f(v_n)), \end{cases}$$

where  $\tilde{H}_n$  is the approximate Hessian computed with the L-BFGS method with  $L$  steps and  $p_C^{\tilde{H}_n}$  is the projection on  $C$  with respect to the norm  $\|\cdot\|_{\tilde{H}_n}$ .

The algorithm to compute the matrix-vector product  $\tilde{H}_n^{-1} \cdot x$  is given in Section 5.3.

**5.2. Computing the projection.** The difficulty lies in the computation of the projection, that is on the proximal operator. A box constraint on the variable is very easy to deal with, but here we are faced with box constraints on  $\nabla \cdot v$ , that is on a linear operator applied to  $v$ . Moreover, we want to compute a projection with respect to a twisted scalar product  $\langle \cdot, \cdot \rangle_H$ , which adds some extra difficulty. For simplicity of notations, we rename  $a_\varepsilon, b_\varepsilon$  as  $a, b$ . By definition, finding the projection  $p_C^H(v_0)$  of  $v_0$  amounts to solving the optimization problem:

$$(P) \quad \min \left\{ \frac{\|v - v_0\|_H^2}{2} : a \leq \nabla \cdot v \leq b, v // \partial\Omega \right\}.$$

Note that, when one considers the divergence operator as an operator acting on vector fields defined on the whole  $\mathbb{R}^d$  (extended to 0 outside  $\Omega$ ), the Neumann boundary condition above exactly corresponds to the fact that the divergence has no mass on  $\partial\Omega$ , which can be considered as included in the inequality constraints.

As a convex optimization, such a problem admits a dual problem, which we are going to use. We set

$$\psi(w) = \begin{cases} 0 & \text{if } a \leq w \leq b, \\ +\infty & \text{if not,} \end{cases}.$$

whose Legendre transform is

$$g(u) = \psi^*(u) = \int bu_+ - \int au_-,$$

so that  $\psi = \psi^{**} = g^*$ . Let us derive formally the dual problem by an inf – sup exchange:

$$\begin{aligned} \inf_{v // \partial\Omega} \left\{ \frac{\|v - v_0\|_H^2}{2} : a \leq \nabla \cdot v \leq b \right\} &= \inf_v \frac{1}{2} \|v - v_0\|_H^2 + \psi(\nabla \cdot v) \\ &= \inf_v \frac{1}{2} \|v - v_0\|_H^2 + \sup_u -\langle \nabla u, v \rangle - g(u) \\ &= \inf_v \sup_u \frac{1}{2} \|v - v_0\|_H^2 - \langle \nabla_H u, v \rangle_H - g(u) \\ &\geq \sup_u -g(u) + \inf_v \frac{1}{2} \|v - v_0\|_H^2 - \langle \nabla_H u, v \rangle_H \\ &= -\inf_u g(u) + \sup_v \langle \nabla_H u, v \rangle_H - \frac{1}{2} \|v - v_0\|_H^2 \\ &= -\inf_u g(u) + \frac{\|\nabla_H u\|_H^2}{2} + \langle \nabla_H u, v_0 \rangle_H \\ &= -\inf_u g(u) + \frac{1}{2} \int H^{-1} \nabla u \cdot \nabla u - \int u(\nabla \cdot v_0). \end{aligned}$$

Hence the dual problem reads:

$$(D) \quad \min_u \underbrace{\frac{1}{2} \int H^{-1} \nabla u \cdot \nabla u - \int u(\nabla \cdot v_0)}_{f(u)} + \underbrace{\int bu_+ - \int au_-}_{g(u)}.$$

The inf – sup interversion can be justified with equality via Fenchel's duality [Bre11, Chapter 1] in a well-chosen Banach space. Hence there is no duality gap:

$$\min (P) + \min (D) = 0.$$

As a consequence solving the dual problem provides a solution to the primal one. Indeed if  $u$  is optimal for (D) then  $v = v_0 + \nabla_H u$  is optimal for (P). Now let us justify why it was interesting to pass by the resolution of a dual problem. Such a problem is of the form

$$(5.3) \quad \min_u f(u) + g(u),$$

where  $f$  is smooth, with gradient  $\nabla f(u) = -\nabla \cdot (H^{-1} \nabla u) - \nabla \cdot v_0$ , and  $g$  is proximable:

$$\text{prox}_g^\tau(u)(x) = \begin{cases} u(x) - \tau a & \text{if } u(x) < \tau a, \\ 0 & \text{if } \tau a \leq u(x) \leq \tau b, \\ u(x) - \tau b & \text{if } u(x) > \tau b. \end{cases}$$

We know how to compute the proximal operator and the gradient of  $f$ , since L-BFGS provides a simple method to compute the product  $H^{-1}x$ . Problems of the form (5.3) with  $f$  smooth (and computable gradient) and  $g$  proximable can be tackled with first-order methods such as the proximal gradient method described in the previous section (also called ISTA) or a fast proximal gradient method called FISTA, introduced in [BT09]. We opted for the

latter, which is a slight modification of the proximal gradient method using an intermediary point:

$$(FISTA) \quad \begin{cases} u_0 \in H^1(\mathbb{R}^d), \\ \tilde{u}_n = u_n + \lambda_n(u_n - u_{n-1}), \\ u_{n+1} = \text{prox}_g^\tau(\tilde{u}_n - \tau \nabla f(\tilde{u}_n)), \end{cases}$$

where  $\lambda_n$  is given by some recursive formula (we refer to [BT09] for the details). It enjoys a theoretical and practical rate of convergence which is higher than ISTA and which is that of the classical gradient method:

$$f(u_n) - f_{opt} \leq \frac{2L_f |u_0 - u_{opt}|^2}{(n+1)^2}.$$

**5.3. Algorithms and numerical experiments.** Following the work of [OS11], we discretize our problem on a staggered grid : we divide the cube  $Q = [-1, 1]^2$  into  $M^2$  subcubes of side  $2/M$ , the functions  $U$  are defined at the center of the small cubes, while the  $x$  component  $V^x$  of a vector fields  $V$  is defined on the vertical edges of the grid and the  $y$  component  $V^y$  on the horizontal edges of the grid. This is quite convenient to compute the discrete divergence of a vector field and the discrete gradient of a function.

- Unknowns:  $(V_{i,j}^x)_{\substack{1 \leq i \leq M \\ 1 \leq j \leq M+1}}, (V_{i,j}^y)_{\substack{1 \leq i \leq M+1 \\ 1 \leq j \leq M}}$ , with

$$V_{1,j}^x = V_{M+1,j}^x = V_{i,1}^y = V_{i,M+1}^y = 0,$$

which means that  $V$  is parallel to the boundary.

- Objective function:

$$F(V) = \varepsilon^{-\sigma_1} h^2 \sum_{i,j} N(\hat{V}_{i,j})^\sigma + \varepsilon^{\sigma_2} h^2 / 2 \left( \sum_{i,j} |\nabla_{i,j} V^x|^2 + \sum_{i,j} |\nabla_{i,j} V^y|^2 \right).$$

There are several definitions to give to make sense of  $F$ . First of all  $N$  is a smooth approximation of the norm, of the form

$$N(x) = (|x|^2 + \varepsilon_s^2)^{1/2} \quad \text{for } \varepsilon_s \text{ small.}$$

The discrete vector field  $\hat{V}_{i,j} = (\hat{V}_{i,j}^x, \hat{V}_{i,j}^y)$  is an interpolation of  $(V^x, V^y)$  defined at the centers of the cubes:

$$\hat{V}_{i,j}^x = \frac{V_{i,j}^x + V_{i+1,j}^x}{2}, \quad \hat{V}_{i,j}^y = \frac{V_{i,j}^y + V_{i,j+1}^y}{2}, \quad 1 \leq i, j \leq M.$$

Finally the discrete gradient is defined as usual by

$$\begin{aligned} \nabla_{i,j} V^x &= ((V_{i,j+1}^x - V_{i,j}^x)/h, (V_{i+1,j}^x - V_{i,j}^x)/h), & 1 \leq i \leq M-1, 1 \leq j \leq M, \\ \nabla_{i,j} V^y &= ((V_{i,j+1}^y - V_{i,j}^y)/h, (V_{i+1,j}^y - V_{i,j}^y)/h), & 1 \leq j \leq M-1, 1 \leq i \leq M. \end{aligned}$$

We may now give the main algorithm and its sub-methods.

**Algorithm 1** Proximal L-BFGS for  $F$ 


---

Data: tolerance  $tol$ , initial vector field  $V_0$ , step  $\tau_0$ , source  $\delta$   
 $V \leftarrow V_0, U \leftarrow U_0$   
**compute**  $error$   
**while**  $error > tol$  **do**  
     $\tau \leftarrow \tau_0$   
    **repeat**  
         $G \leftarrow \text{MULTIPLYBFGS}(\nabla F(V))$   
         $V, U \leftarrow \text{PROJECT}(V - \tau G, U, \delta, \tau)$   
         $\tau \leftarrow \tau/2$   
    **until**  $F(V)$  has decreased  
    **update** L-BFGS data  
    **compute**  $error$   
**end while**

---

The update step for L-BFGS data consists in storing in  $Y, Z, r$  the points and gradients of the  $L$  previous steps, so that at step  $n$ :

$$Y_{L-k} = \nabla F(V_{n-k}) - \nabla F(V_{n-k-1}), \quad Z_{L-k} = V_{n-k} - V_{n-k-1}$$

for all  $k = 0, \dots, L-1$ , and  $r_k = 1/(Y_k \cdot Z_k)$  for all  $k = 0, \dots, L-1$ . Notice here that we do a simple backtracking line search by reducing the stepsize  $\tau$  until the energy has decreased, for example until it has sufficiently decreased and satisfies the Armijo rule. Also, notice that the potential  $U$  computed at step  $n$  is used at the next step as initial data ; this trick extensively speeds up the computation of the projection. Finally, we took as error measurement some relative difference between two consecutive steps.

Now, as stated in Section 5.2, the projection on  $C$  with respect to  $\|\cdot\|_H$  is computed via the FISTA method, as follows:

**Algorithm 2** Project  $V_0$  on  $C$  with respect to  $\|\cdot\|_H$ 


---

Data: tolerance  $tol_p$ , step  $\tau_p$   
**function**  $\text{PROJECT}(V_0, U_0, \delta, \tau)$   
     $D_0 \leftarrow \nabla \cdot V_0$   
     $U \leftarrow U_0$   
    **while**  $error > tol_p$  **do**  
         $t_p \leftarrow t; t \leftarrow (1 + \sqrt{1 + 4t_p^2})/2; s \leftarrow (t_p - 1)/t$   
         $G \leftarrow \text{MULTIPLYBFGS}(\nabla U)$   
         $U_i \leftarrow U + s(U - U_{old})$   
         $U_{old} \leftarrow U$   
         $U \leftarrow \text{PROX}(U_i - \tau_p(\nabla \cdot G - D_0), \delta, \tau)$   
        **compute**  $error$   
    **end while**  
     $V \leftarrow V_0 + \text{MULTIPLYBFGS}(\nabla U)$   
    **return**  $V, U$   
**end function**

---

The PROX function is just the proximal operator associated with the discrete counterpart of  $g : u \mapsto \int bu_+ - \int au_-$  where  $a = \delta - 1, b = \delta$ . Thus  $P = \text{PROX}(U, \delta, \tau)$  is defined by:

$$P_i = \begin{cases} U_i - \tau(\delta_i - 1) & \text{if } U_i < \tau(\delta_i - 1), \\ 0 & \text{if } \tau(\delta_i - 1) \leq U_i \leq \tau\delta_i, \\ U_i - \tau\delta_i & \text{if } U_i > \tau\delta_i. \end{cases}$$

For the sake of completeness, we give a simple method to compute the L-BFGS multiplication  $H^{-1}X$  (see [LN89; Noc80] for details).

---

**Algorithm 3** L-BFGS multiplication  $H^{-1}X$

---

```

function MULTIPLYBFGS( $X$ )
   $G \leftarrow X$ 
  for  $i = L, \dots, 1$  do
     $s_i \leftarrow r_i Z_i \cdot G$ 
     $G \leftarrow G - s_i Y_i$ 
  end for
   $G \leftarrow (Z_L \cdot Y) / (Y_L \cdot Y_L) G$ 
  for  $i = 1, \dots, L$  do
     $t \leftarrow r_i Y_i \cdot G$ 
     $G \leftarrow G + (s_i - t) Z_i$ 
  end for
  return  $G$ 
end function

```

---

We present some numerical results obtained with  $\varepsilon_s = 10^{-4}$ , on a  $M \times M$  grid with  $M = 201$  and  $\varepsilon = 3h$  where  $h = 2/M$ , the code being written in Julia. We have started with random initial values for  $V$  and a smooth approximation  $\delta$  of the Dirac  $\delta_0$ . After some days of computation on a standard laptop, one gets the following shapes and underlying networks.

With no surprise, the shape for  $\alpha = 0.85$  is rounder than those obtained for  $\alpha = 0.55$  and  $\alpha = 0.65$ . These two are quite similar, but a simple zoom shows that the one with the smallest value of  $\alpha$  is a slightly more irregular than the other. The corresponding irrigation networks are also coherent with the expected results: the branches have larger multiplicity (close to the origin) for smaller  $\alpha$ .

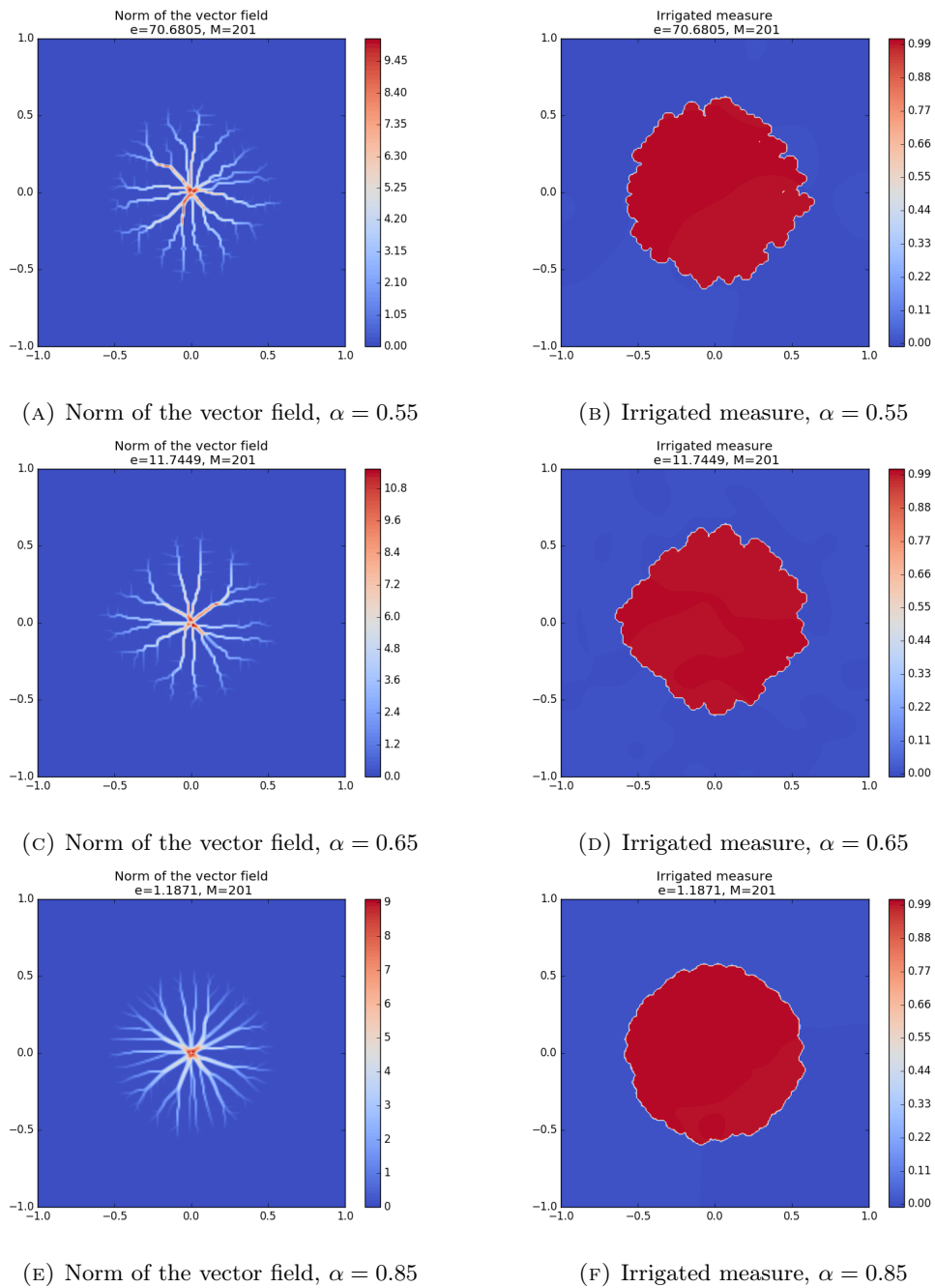


FIGURE 1. Algorithm output for different  $\alpha$ 's after  $\sim 15000$ – $25000$  iterations ( $e$  stands for the computed optimal value, which is an approximation of  $e_\alpha$ , and  $M$  for the number of discretization points on each side of the domain).

**Acknowledgments and Conflicts of Interests.** The support of the ANR project ANR-12-BS01-0014-01 GEOMETRYA and of the PGMO project MACRO, of EDF and Fondation Mathématique Jacques Hadamard, are gratefully acknowledged. The work started during a visit of the second author to UC Davis, and also profited of a visit of the third author

to Univ. Paris-Sud at Orsay, and both Mathematics Department are acknowledged for the warm hospitality.

The authors declare that no conflict of interests exists concerning this work.

## REFERENCES

- [BCM05] M. Bernot, V. Caselles, and J.-M. Morel. “Traffic plans”. In: *Publ. Mat.* 49.2 (2005), pp. 417–451. ISSN: 0214-1493. DOI: [10.5565/PUBLMAT\\_49205\\_09](https://doi.org/10.5565/PUBLMAT_49205_09).
- [BCM09] M. Bernot, V. Caselles, and J.-M. Morel. *Optimal transportation networks*. Vol. 1955. Lecture Notes in Mathematics. Models and theory. Springer-Verlag, Berlin, 2009, pp. x+200. ISBN: 978-3-540-69314-7.
- [BLS15] M. Bonnivard, A. Lemenant, and F. Santambrogio. “Approximation of length minimization problems among compact connected sets”. In: *SIAM J. Math. Anal.* 47.2 (2015), pp. 1489–1529. ISSN: 0036-1410. DOI: [10.1137/14096061X](https://doi.org/10.1137/14096061X).
- [BOO16] M. Bonafini, G. Orlandi, and E. Oudet. “Variational approximation of functionals defined on 1-dimensional connected sets: the planar case”. In: *ArXiv e-prints* (Oct. 2016). arXiv: [1610.03839](https://arxiv.org/abs/1610.03839) [[math.OC](https://arxiv.org/abs/1610.03839)].
- [Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011, pp. xiv+599. ISBN: 978-0-387-70913-0.
- [BS11] A. Brancolini and S. Solimini. “On the Hölder regularity of the landscape function”. In: *Interfaces Free Bound.* 13.2 (2011), pp. 191–222. ISSN: 1463-9963. DOI: [10.4171/IFB/254](https://doi.org/10.4171/IFB/254).
- [BS14] A. Brancolini and S. Solimini. “Fractal regularity results on optimal irrigation patterns”. In: *J. Math. Pures Appl. (9)* 102.5 (2014), pp. 854–890. ISSN: 0021-7824. DOI: [10.1016/j.matpur.2014.02.008](https://doi.org/10.1016/j.matpur.2014.02.008).
- [BT09] A. Beck and M. Teboulle. “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”. In: *SIAM J. Imaging Sci.* 2.1 (2009), pp. 183–202. ISSN: 1936-4954. DOI: [10.1137/080716542](https://doi.org/10.1137/080716542).
- [Cam63] S. Campanato. “Proprietà di hölderianità di alcune classi di funzioni”. In: *Ann. Scuola Norm. Sup. Pisa (3)* 17 (1963), pp. 175–188.
- [CV77] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin-New York, 1977, pp. vii+278.
- [FCM16] L. A. D. Ferrari, A. Chambolle, and B. Merlet. “A simple phase-field approximation of the Steiner problem in dimension two”. 24 pages, 8 figures. Sept. 2016.
- [Gil67] E. N. Gilbert. “Minimum Cost Communication Networks”. In: *Bell System Technical Journal* 46.9 (1967), pp. 2209–2227. ISSN: 1538-7305. DOI: [10.1002/j.1538-7305.1967.tb04250.x](https://doi.org/10.1002/j.1538-7305.1967.tb04250.x).
- [Giu03] E. Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003, pp. viii+403. ISBN: 981-238-043-4. DOI: [10.1142/9789812795557](https://doi.org/10.1142/9789812795557).
- [GP68] E. N. Gilbert and H. O. Pollak. “Steiner Minimal Trees”. In: *SIAM Journal on Applied Mathematics* 16.1 (1968), pp. 1–29. DOI: [10.1137/0116001](https://doi.org/10.1137/0116001). eprint: <http://dx.doi.org/10.1137/0116001>.
- [LN89] D. C. Liu and J. Nocedal. “On the limited memory BFGS method for large scale optimization”. In: *Math. Programming* 45.3, (Ser. B) (1989), pp. 503–528. ISSN: 0025-5610. DOI: [10.1007/BF01589116](https://doi.org/10.1007/BF01589116).
- [LSS14] J. D. Lee, Y. Sun, and M. A. Saunders. “Proximal Newton-type methods for minimizing composite functions”. In: *SIAM J. Optim.* 24.3 (2014), pp. 1420–1443. ISSN: 1052-6234. DOI: [10.1137/130921428](https://doi.org/10.1137/130921428).
- [MM77] L. Modica and S. Mortola. “Un esempio di  $\Gamma^-$ -convergenza”. In: *Boll. Un. Mat. Ital. B (5)* 14.1 (1977), pp. 285–299.



- [Mon15] A. Monteil. “Elliptic approximations of singular energies under divergence constraint”. PhD Thesis. Université Paris-Saclay, Dec. 2015.
- [Mon17] A. Monteil. “Uniform estimates for a Modica-Mortola type approximation of branched transportation”. In: *ESAIM Control Optim. Calc. Var.* 23.1 (2017), pp. 309–335. ISSN: 1292-8119. DOI: [10.1051/cocv/2015049](https://doi.org/10.1051/cocv/2015049).
- [MRCS14] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. “Congestion-driven dendritic growth”. In: *Discrete Contin. Dyn. Syst.* 34.4 (2014), pp. 1575–1604. ISSN: 1078-0947. DOI: [10.3934/dcds.2014.34.1575](https://doi.org/10.3934/dcds.2014.34.1575).
- [MSM03] F. Maddalena, S. Solimini, and J.-M. Morel. “A variational model of irrigation patterns”. In: *Interfaces Free Bound.* 5.4 (2003), pp. 391–415. ISSN: 1463-9963. DOI: [10.4171/IFB/85](https://doi.org/10.4171/IFB/85).
- [Noc80] J. Nocedal. “Updating quasi-Newton matrices with limited storage”. In: *Math. Comp.* 35.151 (1980), pp. 773–782. ISSN: 0025-5718. DOI: [10.2307/2006193](https://doi.org/10.2307/2006193).
- [OS11] E. Oudet and F. Santambrogio. “A Modica-Mortola approximation for branched transport and applications”. In: *Arch. Ration. Mech. Anal.* 201.1 (2011), pp. 115–142. ISSN: 0003-9527. DOI: [10.1007/s00205-011-0402-6](https://doi.org/10.1007/s00205-011-0402-6).
- [Peg17] P. Pegon. “On the Lagrangian branched transport model and the equivalence with its Eulerian formulation”. In: *Topological Optimization and Optimal Transport*. Ed. by M. Bergounioux et al. Berlin: De Gruyter, 2017, pp. 281–303.
- [RIR01] I. Rodriguez-Iturbe and A. Rinaldo. *Fractal river basins: chance and self-organization*. Cambridge University Press, 2001.
- [San07] F. Santambrogio. “Optimal channel networks, landscape function and branched transport”. In: *Interfaces Free Bound.* 9.1 (2007), pp. 149–169. ISSN: 1463-9963. DOI: [10.4171/IFB/160](https://doi.org/10.4171/IFB/160).
- [San10] F. Santambrogio. “A Modica-Mortola approximation for branched transport”. In: *Comptes Rendus Mathématique* 348.15 (2010), pp. 941–945. ISSN: 1631-073X. DOI: <http://dx.doi.org/10.1016/j.crma.2010.07.016>.
- [Xia03] Q. Xia. “Optimal paths related to transport problems”. In: *Commun. Contemp. Math.* 5.2 (2003), pp. 251–279. ISSN: 0219-1997. DOI: [10.1142/S021919970300094X](https://doi.org/10.1142/S021919970300094X).
- [Xia14] Q. Xia. “On landscape functions associated with transport paths”. In: *Discrete Contin. Dyn. Syst.* 34.4 (2014), pp. 1683–1700. ISSN: 1078-0947. DOI: [10.3934/dcds.2014.34.1683](https://doi.org/10.3934/dcds.2014.34.1683).