# A Modica-Mortola approximation for branched transport and applications

Edouard Oudet,[*] Filippo Santambrogio [†]

September 28, 2010

## Abstract

The $M^\alpha$ energy which is usually minimized in branched transport problems among singular 1-dimensional rectifiable vector measures is approximated by means of a sequence of elliptic energies defined on more regular vector fields. The procedure recalls the one of Modica-Mortola related to the approximation of the perimeter. In our context, the double-well potential is replaced by a concave term. The paper contains a proof of $\Gamma-$convergence and numerical simulations of optimal networks based on that previous result.

## 1 Introduction

The name "branched transport" is often used to address transport problems where the cost of transportation of a moving mass $m$ along a distance $l$ is proportional to $l$ and sub-additive with respect to $m$ (like for instance $m^\alpha$ with $0 < \alpha < 1$). The mass distributions of sources and destinations are given and one looks for the paths followed by the particles that are optimal for the total cost $\sum_i l_i m_i^\alpha$. The adjective "branched" stands for one of the main features of optimal solutions: they gather mass together, masses tend to move jointly as long as possible and then branch towards different destinations. This behavior gives rise to tree-shaped structures.

In the case of finite graphs this kind of problems dates back to the '60 in the community of operational research (the first paper on the subject is [17]). More recently, several different approaches and results concerning the generalization to continuous frameworks have been introduced by the community of optimal transport. The first paper in this direction is the one by Q. Xia ([27]), which will be our main reference. Other approaches have been proposed by Bernot, Caselles, Maddalena, Morel and Solimini (see [19, 6]). The equivalences between

---

[*]LAMA, Université de Savoie, Campus scientifique, F- 73376 Le Bourget du Lac Cedex, `edouard.oudet@univ-savoie.fr`

[†]Département de Mathématiques, Bât. 425, Faculté des Sciences, Université Paris-Sud XI, F-91405 Orsay cedex, `filippo.santambrogio@math.u-psud.fr`

the different models as well as a survey of the whole theory are presented in a recent book by Bernot, Caselles and Morel, [7], who are responsible for most of the results therein.

All the different approaches that have been proposed have their own advantages. In this paper we stick to the divergence-constrained formulation proposed by Q. Xia. In this model, an optimal branched transport is obtained in minimizing the energy $M^\alpha(u)$ among vector measures $u$ satisfying $\nabla \cdot u = f$, where $f = f^+ - f^-$ is the difference of two given probability measures representing the sources and the destinations of the transport. The $M^\alpha$ energy is defined (see (3) for a more precise formulation) as $M^\alpha(u) = \int_M \theta^\alpha d\mathcal{H}^1$ for all measures which are concentrated on a rectifiable set $M$ and having a vector density of modulus $\theta$ w.r.t. the Hausdorff measure $\mathcal{H}^1$ on $M$ (and $+\infty$ on the other measures). The sources and destinations, which were finite combination of Dirac masses in the discrete case, are replaced here by generic probability measures and the rectifiable sets $M$ generalize the framework of finite graphs. See Section 2 where the model by Xia is quickly sketched.

A satisfactory numerical treatment of these issues (even in the discrete case) is far from being obtained. This fact motivates the need of reliable approximation strategies. This paper presents a precise approximation result, in terms of $\Gamma-$convergence (see [13]), and an efficient numerical method based on it. Actually, since the continuous formulation of branched transport problems passes through a divergence-constrained formulation, it is natural to approximate it by means of problems which concern more regular vector fields (i.e. which are not measures concentrated on one-dimensional graphs, but have a density that is at least weakly differentiable).

We propose an approximation based on functionals with a concave and a Dirichlet term whose coefficients increase the weight of the first term and vanish on the latter as far the approximation parameter $\varepsilon$ goes to 0. In a very imprecise way, we could say that we consider the minimization of a functional of the kind

$$\min \frac{1}{\varepsilon} \int |u|^\alpha + \varepsilon \int |\nabla u|^2 \qquad (1)$$

under constraints or penalization on $\nabla \cdot u$. We will define an energy $M_\varepsilon^\alpha$, which is similar to (1) after some suitable rescaling, and prove that $M_\varepsilon^\alpha$ converges in the sense of $\Gamma-$convergence to $M^\alpha$. We will recall the definition and properties of this type of convergence (introduced in [14]) in Section 3, and we only insist here on the fact that it is the good notion of convergence for variational problems which provides convergence of minimizers. In this framework, we think that this $\Gamma-$convergence result has at least two important points of interest, from a theoretical and from a numerical point of view. The theoretical aspects, the main results, the difficulties and open questions that arise as well as some suggestions to overcome them are also presented (with no proofs), in the brief note [25] published by the second author.

**Elliptic approximation of singular energies** First, let us spend some words on the theoretical aspects of this kind of approximation. Actually, the

2

interest of this $\Gamma-$convergence result does not only rely in its direct numerical application but also on its comparison with previous elliptic approximation of Modica and Mortola ([20]) at the very beginning of $\Gamma-$convergence time. We explain their result in more details in Section 4. Let us here recall that in their case the variable $u$ was a scalar function and the approximating functionals were defined by

$$\frac{1}{\varepsilon} \int F(u) + \varepsilon \int |\nabla u|^2,$$

$F$ being a double-well potential, enforcing at the limit $u$ to take values in $\{0, 1\}$. With this sequence of functionals the optimal energy concentrates on low-dimensional set which is the interface between the two phases $u = 0$ and $u = 1$. The same happens in our situation: the concave power $F$ plays the role of a double well at $u = 0$ and $|u| = \infty$ and the energy concentrates on a one-dimensional graph.

Being $u$ a vector field in our problem, one could also evoke Ginzburg-Landau theory with its approximation (see [8, 9] where the problem of the convergence of the minimizers of

$$\min \frac{1}{\varepsilon} \int (1 - |u|)^2 + \varepsilon \int |\nabla u|^2$$

is first addressed). Nevertheless, due to the divergence which is a bounded measure, our problem is essentially scalar since asymptotically only one direction is locally relevant. Actually a singular vector measure concentrated on a lower-dimensional object is always oriented along a tangent direction if its divergence is a measure (otherwise it would be a first-order distribution). Moreover, in dimension two it is possible to take advantage of the usual decomposition of a vector field into a gradient plus a rotated gradient so that the information on the divergence "fixes" the gradient part. In this way one would arrive to consider limits of functionals like

$$\min \int F(\nabla u) + \varepsilon \int |D^2 u|^2$$

which are of the form of those studied by Aviles and Giga (Modica-Mortola results for higher order energies, see [4, 5] where the second order term only contains the Laplacian, and lately [1] with the whole Hessian).

Probably the main goal of this paper, from the theoretical point of view, is creating a bridge between two different topics in Calculus of Variations: the approximation of free discontinuity problems on the one hand and the optimization of transport networks. The first one, much linked with elliptic PDEs has already been object of high-quality researches for decades and is studied in relation with its applications in material sciences and image segmentation. The second is, in its continuous version, more recent and linked to the theory of optimal transport by Monge and Kantorovitch with applications ranging from economics to biology and geophysics (see [23]). This topic also involves lots of questions from Elliptic PDEs, dimensional reduction and geometric measure theory. See

[3, 2, 12, 20] for the whole theory and the main examples of $\Gamma-$convergence applied to free discontinuity problems: notice that here, the discontinuities (or "jumps") are replaced by a bilateral singularity. More precisely, the measures solutions of the limit problem vanish almost everywhere and are concentrated on one-dimensional sets. They have a double jump, whose "intensity" (measured with respect to $\mathcal{H}^1$ instead of $\mathcal{L}^d$) enters the limit functional.

There is another class of variational models which is worth mentioning in connection with this study : in a paper devoted to different applications, see [10] Bouchitté et al., the authors give an approximation result which is very similar for some energies defined on atomic measures. We could say that in our paper the 1D networks play the same role in a 2D ambient space as the points played in 1D in [10]. Moreover, the same authors are currently studying, by slicing - but quite different from ours -techniques, the extension to branched transport problems in 2D (also taking advantage of the reformulation in terms of rotated gradients, so as to get to a BV problem). Even if the techniques are different than those which are used here, the difficulties are almost the same.

**Numerical applications**  From the beginning, one of our motivation was to develop some approximation results also for numerical purposes. Actually, this is not a minor point if we consider the fact (see [29]) that the exact identification of global optimal networks, in the combinatorial context, is NP hard (with respect to the number of sources and targets). In order to tackle this difficulty we introduce in this paper a continuous and relaxed framework based on our relaxation.

The main difficulties related to the numerical optimization of problem (4) are related both to the approximation of irregular functions and to the strongly non-convex cost functional.

Theorem (5.1) makes it possible to replace the problem on singular measures by a sequence of optimization problems addressed on smooth functions spaces. Moreover, we observe that for $\varepsilon >> 1$ the functional $M_\varepsilon^\alpha$ is close from being convex, since the Dirichlet part is the most important one. We base our optimization strategy on this last observation. Solving the relaxed problems where $M_\varepsilon^\alpha$ replaces $M^\alpha$, for $\varepsilon$ large, we expect to identify by a standard descent method a good approximation of the solution $u_\varepsilon$. Then, decreasing the value of $\varepsilon$ step by step, and solving the associated sequence of optimization problems starting from the previous optimal $u_\varepsilon$, we can expect to avoid a great number of local minimizers. The same kind of strategy, based on $\Gamma-$convergence and on this choice of starting from huge values of $\varepsilon$ to decrease them later, has already been used by the first author for different problems (see [22]).

Notice that our strategy does not give any warranty to identify at the end of the process a global optimum of the original problem. As a matter of fact, branching in wrong directions may occur when $\varepsilon$ tends to 0. Nevertheless, we observe in our experiments that this approach is surprisingly efficient on the tests cases that we considered.

After the success of the numerical procedures based on Theorem 5.1 which

only deals with the case $\alpha$ larger than $\frac{1}{2}$ (in $\mathbb{R}^2$), we decided to try to attack the case $\alpha = 0$ which corresponds to Steiner problem. The minimal connection problem by Steiner is very classical in computational geometry and amounts to finding the smallest connected set $S$ (in the sense of its length, i.e. we want to minimize $\mathcal{H}^1(S)$) containing a given finite set of points $(x_i)_{i=0,\dots,n}$. This problem may be seen as a particular case of branched transport, when one takes $\alpha = 0$ and takes as a source measure $f^+ = \delta_{x_0}$ and as a destination $f^- = \sum_{i=1}^{n} \frac{1}{n} \delta_{x_i}$. The need to reach $\alpha = 0$ imposed to extend the $\Gamma-$convergence result, thus obtaining Theorem 6.1. Notice that this new result concerns the exponents $\alpha > 0$, and not $\alpha = 0$. Numerically speaking, this leads us to obtain solutions if Steiner's as a limit of irrigation problems where $\alpha \to 0$.

**Structure of the paper** Section 2 recalls Xia's formulation of branched transport problems. The section 3 is a short introduction to $\Gamma-$convergence. Section 4 is related to the form of our relaxed functionals. Mainly we concentrate our attention to the choice of right exponents which are not obvious at all (in particular, the function $F$ in (1) will be of the form $F(|u|) = |u|^\beta$ with $\beta \neq \alpha$). Section 5 presents the detailed proof of our main result ($\Gamma-$convergence of the energies, in dimension two only, under no divergence conditions). Section 6 provides the modifications to be performed on the proofs of Section 5 to cover the case $\alpha \leq 1 - \frac{1}{d}$ (which is needed in order to approximate Steiner problem). In Section 7, before passing to numerics, we discuss the completeness of the theoretical results we prove and we provide some open questions. In particular, we insist on the fact that we do not know how to deal with the divergence constraint and that we are not able to prove compactness of the minimizers of the approximated problems. Anyway, for the sake of numerical simulations, we ignore them: as we previously said we are not able to prove that the numerical procedure we establish converge to a global minimizer of the limit problem, and we only use it to select a candidate "well-selected" local minimizer. For more detailed discussions about these open problems and the way to overcome their difficulties, we will refer to [25].

The numerical issues for branched transport and Steiner problems are discussed in Sections 8-9. Section 8 describes briefly the dicretization process of vector measure on staggered grid. Then algorithms deduced to the projection on the divergence constraints and on our optimization strategy are given. Finally, four numerical experiments are carried out in section 9 to illustrate the performance of the method.

# 2 Branched transport via divergence-constrained optimization

We present here the framework of the optimization problem proposed by Xia in [27, 28] which received the attention of many authors (see for instance [7] for a whole presentation of the theory).

Let $\Omega \subset \mathbb{R}^d$ be an open set with compact closure $\overline{\Omega}$ and $\mathcal{M}(\Omega)$ the set of finite vector measures on $\overline{\Omega}$ with values in $\mathbb{R}^d$ and such that their divergence is a finite scalar measure, i.e. such that

$$\sup\left\{\int \nabla\phi \cdot du \; : \; \phi \in C^1(\overline{\Omega}), \; ||\phi||_{L^\infty} \leq 1\right\} < +\infty \tag{2}$$

(as you can see, we do not ask for $\phi$ to vanish at the boundary i.e. we take into account possible parts of $\nabla \cdot u$ on $\partial\Omega$ as well). The value of the supremum in (2) will be denoted by $|\nabla\cdot u|(\overline{\Omega})$ i.e. the "total variation of the divergence" of $u$. On this space we consider the convergence $u_\varepsilon \to u \Leftrightarrow u_\varepsilon \rightharpoonup u$ and $\nabla\cdot u_\varepsilon \rightharpoonup \nabla\cdot u$ as measures. When a function is considered as an element of this space, or a functional space as a subset of it, we always think it as the absolutely continuous measures (with respect to the Lebesgue measure on $\Omega$) which density is that function.

When we take $u \in \mathcal{M}(\Omega)$ and we write $u = U(M, \theta, \xi)$ we mean that $u$ is a rectifiable vector measure (it is the translation in the language of vector measures of the concept of rectifiable currents) $u = \theta\xi \cdot \mathcal{H}^1_{|M}$ whose density with respect to the $\mathcal{H}^1-$Hausdorff measure on $M$ is given by the real multiplicity $\theta : M \to \mathbb{R}^+$ times the orientation $\xi : M \to \mathbb{R}^d$, $\xi$ being a measurable vector field of unit vectors belonging to the (approximate) tangent space to $M$ at $\mathcal{H}^1-$almost any point.

For $0 < \alpha < 1$, we consider the energy

$$M^\alpha(u) = \begin{cases} \int_M \theta^\alpha d\mathcal{H}^1 & \text{if } u = U(M, \theta, \xi), \\ +\infty & \text{otherwise.} \end{cases} \tag{3}$$

The problem of branched transport amounts to minimizing $M^\alpha$ under a divergence constraint:

$$\min\left\{M^\alpha(u) \; : \; \nabla\cdot u = f := f^+ - f^-\right\}. \tag{4}$$

The divergence constraint is given in weak form and means

$$\int \nabla\phi \cdot du = \int \phi \, d(f^- - f^+) \; \text{ for all } \phi \in C^0(\overline{\Omega}),$$

which actually corresponds to Neumann boundary conditions

$$\nabla\cdot u = f \; \text{ in } \Omega \text{ and } \; u \cdot n = 0 \; \text{ on } \partial\Omega.$$

From now on, we will always think of Neumann boundary conditions when speaking about divergences, so that $\nabla \cdot u$ is the linear functional associating to every $\phi \in C^1(\Omega)$ (independently of the values on $\partial\Omega$) the value $\int \nabla\phi \cdot du$: if $u$ is a regular function this corresponds to a measure which is absolutely continuous inside $\Omega$ with density given by the true divergence, and which has a $\mathcal{H}^1-$part on the boundary with density given by $u \cdot n$.

*Remark* 1. Notice that this is not the original definition by Xia of the Energy $M^\alpha$: Xia proposed it in [27] as a relaxation of the case of finite graphs, but (3) can be seen as a representation formula for the relaxed energy

$$M^\alpha(u) = \inf \left\{ \liminf_n E^\alpha(G_n) \ : \ G_n \text{ finite graph, } u_{G_n} \to u \right\},$$

where

$$E^\alpha(G) := \sum_h w_h^\alpha \mathcal{H}^1(e_h), \tag{5}$$

for a weighted oriented graph $G = (e_h, \hat{e}_h, w_h)_h$ (where $e_h$ are the edges, $\hat{e}_h$ their orientations, $w_h$ the weights), and $u_G$ is the associated vector measure given by

$$u_G := \sum_h w_h \hat{e}_h \mathcal{H}^1_{|e_h},$$

(and the convergence is in the sense of $\mathcal{M}(\Omega)$). For the proof of the equivalences between the two definition, look at [28] or at Chapter 9 in [7].

Notice that in general Problem (4) admits a solution with finite energy for any pair of probability measures $(f^+, f^-)$ (or more generally, for any pair of equal mass of finite positive measures), provided $\alpha > 1 - \frac{1}{d}$ (this is proven in [27] by means of an explicit construction).

# 3 Variational approximation, preliminaries

The main theoretical result of the paper will be a $\Gamma-$convergence result for a sequence of energies approximating $M^\alpha$. We will see in Section 6 that for a complete approximation of the problem, one would need to insert the "boundary conditions" given by the divergence constraints and prove compactness for a suitable sequence of minimizers $u_\varepsilon$ of the approximating problems.

For precising what we mean by "approximating the energy" and how to use the result, let us sketch briefly the main outlines of $\Gamma-$convergence's theory, as introduced by De Giorgi (see [14] and [13]).

**Definition** 3.1. *On a metric space $X$ let $F_n : X \to \mathbb{R} \cup \{+\infty\}$ be a sequence of functions. We define the two lower-semicontinuous functions $F^-$ and $F^+$ (called $\Gamma - \liminf$ and $\Gamma - \limsup F^+$ of this sequence, respectively) by*

$$F^-(x) := \inf\{\liminf_{n \to \infty} F_n(x_n) \ : \ x_n \to x\},$$
$$F^+(x) := \inf\{\limsup_{n \to \infty} F_n(x_n) \ : \ x_n \to x\}.$$

*Should $F^-$ and $F^+$ coincide, then we say that $F_n$ actually $\Gamma-$converges to the common value $F = F^- = F^+$.*

This means that, when one wants to prove $\Gamma-$convergence of $F_n$ towards a given functional $F$, one has actually to prove two distinct facts: first we

need $F^- \geq F$ (i.e. we need to prove $\liminf_n F_n(x_n) \geq F(x)$ for any approximating sequence $x_n \to x$; not only, it is sufficient to prove it when $F_n(x_n)$ is bounded) and then $F^+ \leq F$ (i.e. we need to find a sequence $x_n \to x$ such that $\limsup_n F_n(x_n) \leq F(x)$). The definition of $\Gamma-$convergence for a continuous parameter $\varepsilon \to 0$ obviously passes through the convergence to the same limit for any subsequence $\varepsilon_n \to 0$.

Among the properties of $\Gamma-$convergence we have the following:

- if there exists a compact set $K \subset X$ such that $\inf_X F_n = \inf_K F_n$ for any $n$, then $F$ attains its minimum and $\inf F_n \to \min F$,

- if $(x_n)_n$ is a sequence of minimizers for $F_n$ admitting a subsequence converging to $x$, then $x$ minimizes $F$,

- if $F_n$ is a sequence $\Gamma-$converging to $F$, then $F_n + G$ will $\Gamma-$converge to $F + G$ for any continuous function $G : X \to \mathbb{R} \cup \{+\infty\}$.

## 4  Elliptic approximation, heuristics

As we partially mentioned, the result we will present in Section 5 is somehow inspired by, or at least recalls most of the results in the elliptic approximation of free discontinuity problems (Modica-Mortola, Ginzburg-Landau or Aviles-Giga). We will only mention the following (see [20] and [12]) because of its simplicity, even if it is probably not the closest one in this two-dimensional setting where Aviles-Giga seems closer.

**Theorem 4.1.** *Define the functional $F_\varepsilon$ on $L^1(\Omega)$ through*

$$F_\varepsilon(u) = \begin{cases} \frac{1}{\varepsilon} \int W(u(x))dx + \varepsilon \int |\nabla u(x)|^2 dx & \text{if } u \in H^1(\Omega); \\ +\infty & \text{otherwise.} \end{cases}$$

*Then, if $W(0) = W(1) = 0$ and $W(t) > 0$ for any $t \neq 0, 1$, the functionals $F_\varepsilon$ $\Gamma-$converge towards the functional $F$ given by*

$$F(u) = \begin{cases} cPer(S) & \text{if } u = 1 \text{ on } S, \ u = 0 \text{ on } S^c \text{ and } S \text{ is a finite-perimeter set;} \\ +\infty & \text{otherwise,} \end{cases}$$

*where the constant $c$ is given by $c = 2 \int_0^1 \sqrt{W(t)}dt$.*

We precised the value of the constant so that the reader will notice that similar constants are involved in our case as well. For the same reason (the analogy with the present paper) we precise also the key-ingredient in the proof of the above Theorem which is the inequality

$$\frac{1}{\varepsilon}W(u(x)) + \varepsilon|\nabla u(x)|^2 \geq 2\sqrt{W(u(x))}|\nabla u(x)| = 2|\nabla(H \circ u)|,$$

where $H$ is the primitive of $\sqrt{W}$ (so that one has $F_\varepsilon(u) \geq 2TV(H \circ u)$ where $TV$ denotes the total variation).

In our study we will consider functionals of the form

$$E_\varepsilon(u) = \varepsilon^{\gamma_1} \int_\Omega |u(x)|^\beta dx + \varepsilon^{\gamma_2} \int_\Omega |\nabla u(x)|^2 dx, \tag{6}$$

defined on $u \in H^1(\Omega; \mathbb{R}^2)$ and set to $+\infty$ outside $H^1 \subset \mathcal{M}(\Omega)$, for well-chosen exponents $\gamma_1 < 0 < \gamma_2$ and $0 < \beta < 1$.

As one can see our functional recalls Modica-Mortola's functional to recover the perimeter as a limit, where the double-well potential is replaced with a concave power. Notice that when minimizing concave powers of a function $u$ of fixed average value (which is in some sense the meaning of weak convergence, i.e. the convergence we use on $\mathcal{M}(\Omega)$), it is suitable to have either $u = 0$ or $|u|$ as large as possible. In that sense the concave potential plays the role of double well on zero and infinity. We give below an heuristic argument to determine the right exponents $\beta$, $\gamma_1$ and $\gamma_2$.

Suppose you want to approximate a measure $u$ concentrated on a segment $S$, with multiplicity $m$, and directed towards one of the direction of the segment, via a measure $u_A$ whose density is smooth and concentrated on a strip of width $A$ around $S$. The values of $u_A$ will hence be of the order of $m/A^{d-1}$ and the values of its gradient of the order of $m/A^d$. This gives a functional value of the order of

$$E_\varepsilon \approx \varepsilon^{\gamma_1} A^{d-1} \left( \frac{m}{A^{d-1}} \right)^\beta + \varepsilon^{\gamma_2} A^{d-1} \left( \frac{m}{A^d} \right)^2.$$

Minimizing the previous cost with respect to the width $A$, we get

$$A \approx \varepsilon^{\frac{\gamma_2 - \gamma_1}{2d - \beta(d-1)}} m^{\frac{2-\beta}{2d - \beta(d-1)}}; \quad E_\varepsilon \approx \varepsilon^{\gamma_2 - (\gamma_2 - \gamma_1)\frac{d+1}{2d - \beta(d-1)}} m^{2 - (2-\beta)\frac{d+1}{2d - \beta(d-1)}}.$$

The correct choice for a possible convergence result towards the energy (3) which is proportional to $m^\alpha$ is obtained by imposing

$$\beta = \frac{2 - 2d + 2\alpha d}{3 - d + \alpha(d-1)}; \quad \frac{\gamma_1}{\gamma_2} = \frac{(d-1)(\alpha - 1)}{3 - d + \alpha(d-1)}.$$

Notice that $\gamma_1$ and $\gamma_2$ may not both be determined since one can always replace $\varepsilon$ with a power of $\varepsilon$ which change single exponents but not their ratio. Notice also that the exponent $\beta$ is positive and less than 1 as soon as $\alpha \in ]1 - \frac{1}{d}, 1[$, which is the usual condition for having a well posed problem.

Finally, it is worthwhile to remark that for this choice of exponents the dependence of $A$ with respect to $\varepsilon$ is of the form $A \approx \varepsilon^{\gamma_2/(d+1)}$ (and $\gamma_2 > 0$). This implies $\lim_{\varepsilon \to 0} A = 0$, which gives weak convergence of the approximation we chose (enlarging $u$ on a strip of width $A$ without changing its mass) to $u$. It doesn't prove anything but it is coherent with the our expected convergence result.

We conclude by underlining the case of the dimension 2 since it is the only one we will be able to solve. In that case one has $\beta = \frac{4\alpha - 2}{\alpha + 1}$ and $\gamma_1/\gamma_2 =$

$(\alpha - 1)/(\alpha + 1)$. This allows us to precise our choice for the approximating energies $M_\varepsilon^\alpha$:

$$M_\varepsilon^\alpha(u) = \begin{cases} \varepsilon^{\alpha-1} \int_\Omega |u(x)|^\beta dx + \varepsilon^{\alpha+1} \int_\Omega |\nabla u(x)|^2 dx & \text{if } u \in H^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases} \qquad (7)$$

with $\beta = \frac{4\alpha-2}{\alpha+1}$.

*Remark* 2. From this heuristics and from the proof that we will present in the following, the reader may see that the construction only works for $\alpha > 1 - \frac{1}{d}$. This is quite astonishing if one thinks that, also for $\alpha \leq 1 - \frac{1}{d}$, there are measures $f^\pm$ which admit possible solution with finite energy (in particular when both measures are supported on a same lower-dimensional set). Yet, the problem lies in the kind of approximation we require, which uses measures $u_\varepsilon$ which are more regular and in particular which are absolutely continuous with respect to $\mathcal{L}^d$ and may not be concentrated on lower-dimensional sets. Section 6 explains how to deal with the case where $\alpha$ is smaller than the threshold $1 - \frac{1}{d}$ and how to adapt the arguments we will give in next section to the case $\alpha \in ]0, \frac{1}{2}]$ for $d = 2$.

# 5 Our main $\Gamma-$convergence result

This Section will be devoted to the proof of the following theorem:

**Theorem 5.1.** *Suppose $d = 2$ and $\alpha \in ]1/2, 1[$: then we have $\Gamma-$convergence of the functionals $M_\varepsilon^\alpha$ to $cM^\alpha$, with respect to the convergence of $\mathcal{M}(\Omega)$, as $\varepsilon \to 0$, where $c$ is a finite and positive constant (the value of $c$ is actually $c = \alpha^{-1} \left(4c_0\alpha/(1-\alpha)\right)^{1-\alpha}$, being $c_0 = \int_0^1 \sqrt{t^\beta - t}\, dt$).*

As usual in several $\Gamma-$convergence proofs, we will work separately on the two inequalities we need.

## 5.1 $\Gamma - \liminf$ inequality

In this part we will consider an arbitrary sequence $u_\varepsilon \to u$ and we will suppose that $M_\varepsilon^\alpha(u_\varepsilon)$ is bounded. This implies at first that all the $u_\varepsilon$ are $H^1$ functions, and that

$$\int |u_\varepsilon|^\beta \leq C\varepsilon^{1-\alpha} \quad \text{and} \quad \int |\nabla u_\varepsilon|^2 \leq C\varepsilon^{-1-\alpha}. \qquad (8)$$

Not only they are $H^1$ functions, but we can actually assume, for this part of the proof, that they are more regular, say $C^1$, since we could otherwise replace them with other more regular functions, still denoted by $u_\varepsilon$, with the same limit in $\mathcal{M}(\Omega)$, and the same energies $M_\varepsilon^\alpha$ up to negligible remainder (this is possible thanks to the fact that $C^1$ is dense in $H^1$).

**Step 1** *Sliced lower bounds on $M_\varepsilon^\alpha$ :* We are going to establish below for each

10

$x$ a lower bound for the integral $\int \varepsilon^{\alpha-1}|u_\varepsilon(x,y)|^\beta + \varepsilon^{\alpha+1}|\nabla u_\varepsilon(x,y)|^2\,dy$ which involves a term of the form $|\int u_\varepsilon(x,y)\,dy|^\alpha$ up to some technical details (cut off functions, negligible remainders, ...)

Consider a rectangle $R \subset \Omega$. Suppose for simplicity that it is oriented according to the $x$ and $y$ axes, i.e. $R = [a,b] \times [c,d]$ set $v_\varepsilon := [(u_\varepsilon)_x]_+$ (the positive part of the $x-$component of $u_\varepsilon$) and $v'_\varepsilon := \partial v_\varepsilon/\partial y$. Consider also two $C^1$ functions $\phi_1, \phi_2 : R \to [0,1]$ such that: $\partial\phi_1/\partial x = \partial\phi_2/\partial x = 0$, $\phi_1 = 1$ on $[a,b] \times [c+\delta, d-\delta]$, $\phi_1(x,y) = 0$ for $y \in \{c,d\}$ and $\phi_2 = 1$ on $[a,b] \times [c+2\delta, d-2\delta]$ and $\phi_2(y) = 0$ for $y \in [c, c+\delta] \cup [d-\delta, d]$. For every $x$, set $R_x := \{x\} \times [c,d]$ and $\hat{R}_x := \{x\} \times [c+\delta, d-\delta]$. Finally we define $m_{\varepsilon,\delta}(x) = \int_{R_x} (u_\varepsilon)_x \phi_2\,dy$. The precise estimate that is proved in this step is

$$\int_{R_x} \left(\varepsilon^{\alpha-1}|u_\varepsilon(x,y)|^\beta + \varepsilon^{\alpha+1}|\nabla u_\varepsilon(x,y)|^2\right)\phi_1(y)\,dy \geq c|m_{\varepsilon,\delta}(x)|^\alpha - r_{\varepsilon,\delta}(x),$$

for some constant $c$ and a function $r_{\varepsilon,\delta}$ which will be proved to be negligible in step 2. Fix $x \in [a,b]$ and let $A_\varepsilon$ be the maximal value of $v_\varepsilon$ on $\hat{R}_x$ and $L_\varepsilon = A_\varepsilon^{\beta-1}$, $f_\varepsilon(t) = \sqrt{(t^\beta - L_\varepsilon t)_+}$ and $F_\varepsilon(t) = \int_0^t f_\varepsilon(s)ds$. One can write

$$v_\varepsilon^\beta = f_\varepsilon^2(v_\varepsilon) + L_\varepsilon v_\varepsilon - (v_\varepsilon^\beta - L_\varepsilon v_\varepsilon)_- \geq f_\varepsilon^2(v_\varepsilon) + L_\varepsilon v_\varepsilon - L_\varepsilon v_\varepsilon I_{R_x\setminus\hat{R}_x} = f_\varepsilon^2(v_\varepsilon) + L_\varepsilon v_\varepsilon I_{\hat{R}_x}$$

where the second inequality comes from the fact that $v_\varepsilon \leq A_\varepsilon$ implies $v_\varepsilon^\beta - L_\varepsilon v_\varepsilon \geq 0$ and $A_\varepsilon = \max_{\hat{R}_x} v_\varepsilon$. Considering the other term as well one has

$$\varepsilon^{\alpha-1}v_\varepsilon^\beta + \varepsilon^{\alpha+1}(v'_\varepsilon)^2 \geq \varepsilon^{\alpha-1}L_\varepsilon v_\varepsilon I_{\hat{R}_x} + \varepsilon^{\alpha-1}f_\varepsilon^2(v_\varepsilon) + \varepsilon^{\alpha+1}(v'_\varepsilon)^2$$
$$\geq \varepsilon^{\alpha-1}L_\varepsilon v_\varepsilon I_{\hat{R}_x} + 2\varepsilon^\alpha f_\varepsilon(v_\varepsilon)|v'_\varepsilon| = \varepsilon^{\alpha-1}L_\varepsilon v_\varepsilon I_{\hat{R}_x} + 2\varepsilon^\alpha|F_\varepsilon(v_\varepsilon)'|.$$

By multiplying times $\phi_1$ and integrating on $R_x$ with respect to $y$, one has

$$\int_{R_x}\left(\varepsilon^{\alpha-1}v_\varepsilon^\beta + \varepsilon^{\alpha+1}(v'_\varepsilon)^2\right)\phi_1(y)dy$$
$$\geq \varepsilon^{\alpha-1}L_\varepsilon\int_{\hat{R}_x}v_\varepsilon dy + 2\varepsilon^\alpha TV(F_\varepsilon(v_\varepsilon)\phi_1) - 2\varepsilon^\alpha\int_{R_x\setminus\hat{R}_x}F_\varepsilon(v_\varepsilon)|\phi'_1|dy, \quad (9)$$

where $TV$ stands again for the total variation. Since $v_\varepsilon\phi_1$ vanishes at both boundaries and reaches the value $F_\varepsilon(A_\varepsilon)$ inside $\hat{R}_x$, the total variation $TV(F_\varepsilon(v_\varepsilon)\phi_1)$ is at least $2F_\varepsilon(A_\varepsilon) = 2c_0 A_\varepsilon^{1+\beta/2}$. This value may be computed by a change of variable ($t = A_\varepsilon s$):

$$F_\varepsilon(A_\varepsilon) = \int_0^{A_\varepsilon}\sqrt{(t^\beta - L_\varepsilon t)_+}dt = \int_0^1\sqrt{A_\varepsilon^\beta s^\beta - L_\varepsilon A_\varepsilon s}\,A_\varepsilon ds = A_\varepsilon^{1+\beta/2}c_0$$

(we used $L_\varepsilon = A_\varepsilon^{\beta-1}$ and $t^\beta \geq L_\varepsilon t$ for $t \leq A_\varepsilon$).

Notice that the last term in (9) is bounded by $r_{\varepsilon,\delta}(x) := c_\delta\varepsilon^\alpha\int_{R_x}|u_\varepsilon|^{1+\beta/2}$ since $\phi'_1$ is bounded by a constant depending of $\delta$, $F_\varepsilon(t) \leq \int_0^t s^{\beta/2}ds = ct^{1+\beta/2}$ and $|v_\varepsilon| \leq |u_\varepsilon|$.

Since we will prove in step 2 that $r_{\varepsilon,x}(x)$ tends to zero as $\varepsilon \to 0$, after integration with respect to $x$ (see Lemma 5.2 below), we will ignore this term. For every $x$, we look at the quantity

$$\varepsilon^{\alpha-1} A_\varepsilon^{\beta-1} \hat{m}_\varepsilon(x) + 4c_0 \varepsilon^\alpha A_\varepsilon^{1+\beta/2},$$

where $\hat{m}_\varepsilon(x) := \int_{\hat{R}_x} v_\varepsilon dy$. To estimate it from below, we will minimize over possible values of $A_\varepsilon$. We have

$$\min_{A \in ]0,+\infty[} \varepsilon^{\alpha-1} A^{\beta-1} m + 4c_0 \varepsilon^\alpha A^{1+\beta/2} = c_2 m^\alpha, \tag{10}$$

the minimum being realized by

$$A = \left( \frac{m(1-\beta)}{\varepsilon 2 c_0 (2+\beta)} \right)^{2/(4-\beta)}. \tag{11}$$

We do not precise here the coefficient $c_2$ appearing in the minimal value but the correct computation is the one in the statement of the theorem. Notice that the exponent $2/(4-\beta)$ equals $(\alpha+1)/3$. For computing the last equality in (10) we need to use the relations between $\beta$ and $\alpha$. By collecting all this estimates and assuming $m_{\varepsilon,\delta}(x) \geq 0$, we obtain

$$\int_{R_x} \left( \varepsilon^{\alpha-1} |u_\varepsilon(x,y)|^\beta + \varepsilon^{\alpha+1} |\nabla u_\varepsilon(x,y)|^2 \right) \phi_1(y) \, dy$$

$$\geq \int_{R_x} \left( \varepsilon^{\alpha-1} |v_\varepsilon(x,y)|^\beta + \varepsilon^{\alpha+1} |v_\varepsilon'(x,y)|^2 \right) \phi_1(y) \, dy$$

$$\geq c_2 |\hat{m}_\varepsilon(x)|^\alpha - r_{\varepsilon,\delta}(x)$$

$$\geq c_2 | \int_{R_x} v_\varepsilon \phi_2 dy|^\alpha - r_{\varepsilon,\delta}(x) \geq c_2 |m_{\varepsilon,\delta}(x)|^\alpha - r_{\varepsilon,\delta}(x)$$

If $m_{\varepsilon,\delta}(x) < 0$, we can reproduce the same strategy considering the negative parts instead of the positive parts to obtain the same estimate.

**Step 2** *Lower bounds pass to the limit :* In this step we integrate the previous lower bound with respect to $x$ and consider the limit when $\varepsilon$ tends to zero. After getting rid off the term $r_{\varepsilon,\delta}$, we establish a strong compactness result which is necessary to pass to the limit in the none-convex term $\int |m_{\varepsilon,\delta}(x)|^\alpha \, dx$. Secondly we let $\delta$ go to zero and obtain an estimate without cutoff function.

One can call $\mu_\varepsilon$ the positive measure $\left( \varepsilon^{\alpha-1} |u_\varepsilon|^\beta + \varepsilon^{\alpha+1} |\nabla u_\varepsilon|^2 \right) \cdot \mathcal{L}^2$. Since $\int_\Omega d\mu_\varepsilon = M_\varepsilon^\alpha(u_\varepsilon)$, the measures $\mu_\varepsilon$ stay bounded in the set of positive Radon measures on $\Omega$. Hence we can suppose $\mu_\varepsilon \rightharpoonup \mu$.

As we did previously, take a rectangle $R$ and keep the same notations. We have got from step 1:

$$\int_R \phi_1(y) d\mu_\varepsilon \geq c_2 \int_a^b [m_{\varepsilon,\delta}(x)]^\alpha dx - \int_a^b r_{\varepsilon,\delta}(x) \, dx. \tag{12}$$

The following lemma allows to handle the last term in (12) proving that it is negligible as $\varepsilon$ tends to zero.

**Lemma 5.2.** *For any bounded energy sequence $u_\varepsilon$ we have*

$$\lim_{\varepsilon \to 0} \varepsilon^\alpha \int_R |u_\varepsilon|^{1+\beta/2} = 0.$$

*Proof.* First of all write $|u_\varepsilon| \le 1 + w_\varepsilon$, where $w_\varepsilon = (|u_\varepsilon| - 1)_+$ and then

$$\int_R |u_\varepsilon|^{1+\beta/2} \le C + C \int_R w_\varepsilon^{1+\beta/2}.$$

Notice than that $|\{|u_\varepsilon| > 1\}| \le \int_R |u_\varepsilon|^\beta \le C\varepsilon^{1-\alpha} \to 0$, thus $w_\varepsilon$ vanishes on a large part of $R$. This allows to apply standard Sobolev-Poincaré inequalities $||w_\varepsilon||_{L^r} \le C||w_\varepsilon||_{H^1}$ (in dimension two any exponent $r < +\infty$ is admitted). Remember $w_\varepsilon \le |u_\varepsilon|$ and $|\nabla w_\varepsilon| \le |\nabla u_\varepsilon|$.

Now, for any pair of conjugate exponents $p$ and $q$ greater than 1, we have

$$\int_R w_\varepsilon^{1+\beta/2} \le \left( \int_\Omega |w_\varepsilon|^\beta \right)^{\frac{1}{p}} \left( \int_\Omega |w_\varepsilon|^{(1+\beta/2-\beta/p)q} \right)^{\frac{1}{q}}$$

$$\le C\varepsilon^{(1-\alpha)/p} ||u_\varepsilon||_{L^{(1+\beta/2-\beta/p)q}}^{1+\beta/2-\beta/p} \le C\varepsilon^{(1-\alpha)/p} ||u_\varepsilon||_{H^1}^{1+\beta/2-\beta/p}$$

$$\le C\varepsilon^{(1-\alpha)/p-(1+\beta/2-\beta/p)(\alpha+1)/2}.$$

Hence we have

$$\varepsilon^\alpha \int_R |u_\varepsilon|^{1+\beta/2} \le C\varepsilon^\alpha + C\varepsilon^{\gamma_p},$$

where the exponent $\gamma_p$, from the previous computations, is given by

$$\gamma_p = \alpha + \frac{1-\alpha}{p} - \frac{(2p + \beta(p-2))(\alpha+1)}{4p} = \frac{\alpha(2-p)}{2p}.$$

It goes to zero provided $\gamma_p > 0$, and it is sufficient to choose $p < 2$ in order to get the result. $\square$

*Remark* 3. Notice that the proof would have been easier if one supposed that $u_\varepsilon \cdot n = 0$ (since one could have directly applied Sobolev-Poincaré to $u_\varepsilon$) on $\partial\Omega$, which is quite natural. Yet, for the sake of generality, we admitted possible divergences concentrated on the boundary, i.e.non vanishing values of the normal component. Actually, in this $H^1$ setting, the divergence of $u_\varepsilon$ is seen as a measure on $\overline{\Omega}$ with an $L^2$ part in $\Omega$ and a boundary part corresponding to $u_\varepsilon \cdot n$.

We know consider the limit of the term $\int_a^b [m_{\varepsilon,\delta}(x)]^\alpha dx$ when $\varepsilon$ tends to zero. The measures $m_{\varepsilon,\delta}(x)dx$ are the projections on $[a,b]$ of $\phi_2 \cdot (u_\varepsilon)_x$. They obviously weakly converge to the projection of $\phi_2 u_x$. Yet, this weak convergence is not sufficient for getting the convergence (nor for lower semicontinuity) of this term. This is due to the non-convex behavior of the function $m \mapsto |m|^\alpha$. To prove this convergence we need more compactness (and hence a stronger convergence of $(\pi_x)_\#(\phi_2(u_\varepsilon)_x)$).

13

What we may prove is that the functions $x \mapsto m_{\varepsilon,\delta}(x)$ are uniformly $BV$ in $x$ and this will allow for $L^1$ and pointwise convergence. We recall that $u_\varepsilon \to u$ in $\mathcal{M}(\Omega)$ which means by definition that $u_\varepsilon$ and its divergence weakly converge as measures. In particular both $|\nabla \cdot u_\varepsilon|(\overline{\Omega})$ and $||u_\varepsilon||_{L^1}$ are bounded. Take a function $\psi : [a,b] \to \mathbb{R}$ with $\psi(a) = \psi(b) = 0$. Consider

$$\int_a^b m_{\varepsilon,\delta}(x)\psi'(x)dx = \int_R (u_\varepsilon)_x(x,y)\phi_2(y)\psi'(x)\,dydx$$

$$= \int_R u_\varepsilon(x,y) \cdot \nabla(\phi_2(y)\psi(x))dydx - \int_R (u_\varepsilon)_y(x,y)\phi_2'(y)\psi(x)\,dydx$$

$$\leq |\nabla \cdot u_\varepsilon|(\overline{\Omega})||\psi||_{L^\infty(a,b)} + ||u_\varepsilon||_{L^1}||\phi_2'||_{L^\infty(a,b)}||\psi||_{L^\infty(a,b)} \leq C||\psi||_{L^\infty(a,b)}.$$

This proves the BV bound we needed and implies that $(\pi_x)_\#(\phi_2 u_x)$ is a measure on $[a,b]$ "belonging to BV" (i.e. is absolutely continuous and has a BV density). Let us call $m_\delta(x)$ its density. Moreover, one has $m_{\varepsilon,\delta}(x) \to m_\delta(x)$ for almost any $x$. Passing to the limit one gets, by Fatou's Lemma,

$$\int_R \phi_1(y)d\mu \geq \int_a^b |m_\delta(x)|^\alpha dx.$$

It is quite straightforward that one can let $\delta$ go to 0 and get rid of the functions $\phi_1$ and $\phi_2$ (which actually depend on $\delta$). In fact, in the limit as $\delta$ tends to zero the measures $\phi_2 u_x$ converges strongly to the measure $1_{[a,b]\times]c,d[} \cdot u_x$. This implies strong convergence for $m_\delta$ to $(\pi_x)_\#(1_{[a,b]\times]c,d[} \cdot u_x)$. Notice that this last measure is also absolutely continuous since all the $m_\delta$ are in $L^1$. Hence if we denote by $m(x)$ its density, it satisfies

$$\mu([a,b]\times]c,d[) \geq \int_a^b |m(x)|^\alpha dx. \tag{13}$$

**Step 3** *Rectifiability and comparison with $M^\alpha$* : In this step we prove that the estimate of step 2 implies that $u$, simply because of its attainability as a limit of $u_\varepsilon$ with bounded $M_\varepsilon^\alpha$ energies, is a rectifiable measure. Finally, this point is used to obtain $\mu(\overline{\Omega}) \geq M^\alpha(u)$.

The proof of the rectifiability is contained in the two following Lemmas 5.3, 5.4.

**Lemma 5.3.** *Suppose $u_\varepsilon \to u$ in $\mathcal{M}(\Omega)$ and $M_\varepsilon^\alpha(u_\varepsilon) \leq C$. Then $u$ is a one-dimensional rectifiable vector measure, i.e. it is of the form $u = U(M, \theta, \xi)$.*

*Proof.* We will use the rectifiability theorem proved through different techniques by Federer and White (see [15, 26]) and already used to prove rectifiability of measures with finite $M^\alpha-$mass (see also [28]). This theorem roughly states that $u$ is rectifiable if and only if almost all its $(d-1)-$dimensional slices parallel to the coordinate axes are countable collections of Dirac masses. Take the coordinate axes and disintegrate $u_x$ with respect to the projection on the variable

$x$ of $|u_x|$ (that we will call for simplicity $m(x)dx$ on the variable $x$, even if before this denoted the projection of a signed measure). In this way we get some signed measures $\nu_x$. For every $n$, divide $\Omega$ into $2^n$ horizontal strips of equal width and call $f_{i,n}(x)$ the integral of $\nu_x$ on the $i$-th interval strip (without its boundary). The estimate (13) that we already established easily gives

$$\mu(\Omega) \geq \int m(x)^\alpha \sum_i f_{i,n}(x)^\alpha dx.$$

Up to choosing the levels where to put the boundary of the strips, we can ensure that for almost any $x$ no mass is given to the boundaries by $|u|$ (if this is not the case, simply translate a little bit the strips, and this will happen again for a countable set of choices only). Hence, when we pass from $n$ to $n+1$, the mass of the previous strip is exactly the sum of the mass of the two new strips. Notice that, due to $|a + b|^\alpha \leq |a|^\alpha + |b|^\alpha$, the sequence $n \mapsto \sum_i f_{i,n}(x)^\alpha$ is increasing. Call $G(x)$ its limit: we have $\int m(x)^\alpha G(x)dx < +\infty$. This implies that $G$ is finite almost everywhere. Thanks to Lemma 5.4, almost every $\nu_x$ is purely atomic, and the same may be performed on the direction $y$. This allows to apply the White's criterion and proves that $u$ is rectifiable. □

**Lemma 5.4.** *For $\alpha < 1$ and a measure $\nu$ on a interval (say $[0, 1[$), set*

$$G_\alpha^{(sup)}(\nu) = \sup \left\{ \sum_{i=0}^{2^n-1} |\nu(I_{i,n})|^\alpha; \quad n \in \mathbb{N} \quad I_{i,n} = \left[ \frac{i}{2^n}, \frac{i+1}{2^n} \right[ \right\}$$

*and*

$$G_\alpha(\nu) = \begin{cases} \sum_{k \in \mathbb{N}} (a_k)^\alpha & \text{if } \nu = \sum_{k \in \mathbb{N}} a_k \delta_{x_k} \\ +\infty & \text{otherwise} \end{cases}.$$

*Then we have $G_\alpha^{(sup)}(\nu) = G_\alpha(\nu)$, and in particular if $G_\alpha^{(sup)}(\nu) < +\infty$ then $\nu$ is purely atomic.*

*Proof.* Let us start from proving $G_\alpha^{(sup)}(\nu) \geq G_\alpha(\nu)$: for any $n$, build a measure $\nu_n$ which is purely atomic, with one atom at each point $i2^{-n}$ and choose the mass of such an atom equal to that of $\nu$ on $I_{i,n}$. These measures converge weakly to $\nu$, hence we have

$$G_\alpha(\nu) \leq \liminf G_\alpha(\nu_n) \leq G_\alpha^{(sup)}(\nu).$$

Then, we prove the opposite inequality. We can suppose $G_\alpha(\nu) < +\infty$. Hence $\nu = \sum_{k \in \mathbb{N}} a_k \delta_{x_k}$ is purely atomic and for every pair $(i, n)$ we have, thanks to subadditivity,

$$|\nu(I_{i,n})|^\alpha \leq \sum_{k \,:\, x_k \in I_{i,n}} a_k^\alpha.$$

Summing up, we get

$$\sum_{i=0}^{2^n-1} |\nu(I_{i,n})|^\alpha \leq G_\alpha(\nu)$$

15

and the proof is obtained by taking the sup over $n$.　　　　　　　□

Once we know about $u$ being rectifiable, one can choose rectangles $R$ shrinking around a tangent segment to the set $M$ at a point $x_0$ (this works for $\mathcal{H}^1-$almost any point $x_0$) and get $\mu \geq \theta^\alpha \cdot \mathcal{H}^1_{|M}$, which implies the thesis.

This kind of proofs follow a standard scheme, i.e. considering measures $\mu_\varepsilon$ whose mass gives the value of the approximating energy, and providing estimates on the limit measure $\mu$. This estimates are obtained through local inequality on $\int_R d\mu$, so that shrinking $R$ around a point one gets information on the density of $\mu$ (here it is the density w.r.t. $\mathcal{H}^1$). Similar proofs (typically comparing $\mu$ to $\mathcal{L}^d$ instead of $\mathcal{H}^1$), are quite used in $\Gamma-$convergence problems in the setting of transport and location: see for instance [11] and [21]

## 5.2　$\Gamma - \limsup$ **inequality**

**Step 4** *The case of a single segment with constant multiplicity :* In this step we produce a recovery sequence for the case of $u = \theta_0 \cdot \mathcal{H}^1_{|S}$, being $S$ a segment and $\theta_0$ a constant positive multiplicity on it (for simplicity, $S = [0,1] \times \{0\}$).

For this construction we will consider the inequalities obtained in the lower band proof and try to obtain equalities instead. We look for a profile $u_\varepsilon$ having the following properties:

- the $x-$component only of $u_\varepsilon$ must be present and must have everywhere the same sign, so that $u_\varepsilon = v_\varepsilon e_1$;

- $\nabla v_\varepsilon = v'_\varepsilon e_2$ (i.e. $v_\varepsilon$ only depends on $y$);

- the Cauchy-Schwartz inequality used in (9) must be an equality, i.e. one needs
$$v'_\varepsilon = \pm \frac{1}{\varepsilon} \sqrt{v_\varepsilon^\beta - L_\varepsilon v_\varepsilon};$$

- the total variation must actually be given by twice the maximum (i.e. $v_\varepsilon$ must be monotone on the two separate intervals before and after reaching the maximum: we will realize it by taking a maximal value at $y = 0$ and symmetric monotone profiles around 0);

- $v_\varepsilon$ must vanish at the boundary of a certain rectangle, so that one can avoid using the function $\phi_1$;

- the maximum value $A_\varepsilon$ must be optimal in (10);

- as $\varepsilon \to 0$ weak convergence to the measure $u = \theta_0 \cdot \mathcal{H}^1_{|S}$ is needed: we will realize it by taking different rescaling of the same profile $z$ (say, $v_\varepsilon(y) = A_\varepsilon z(A_\varepsilon y)$), so that $A_\varepsilon \to \infty$ and $\int_{\mathbb{R}} z(t)dt = \theta_0$ will be sufficient.

Look at the conditions that the profile $z$ must satisfy: we need $z(0) = 1$ and $z'(t)t \leq 0$, so that the maximum of $v_\varepsilon$ will be $A_\varepsilon$, realized at zero, and the

16

monotonicity conditions as well will be satisfied. Moreover, on $t \geq 0$ (for $t \leq 0$ just symmetrize), $z$ must satisfy

$$z'(t) = -\frac{A_\varepsilon^{\beta/2}}{\varepsilon A_\varepsilon^2}\sqrt{z^\beta - z}.$$

Thanks to (11), the ratio $A_\varepsilon^{\beta/2}/(\varepsilon A_\varepsilon^2)$ equals $2c_0(2+\beta)/((1-\beta)\theta_0)$. This is very good since there is no more $\varepsilon$.

Notice that this equation has a strong non-uniqueness, since the function $z \mapsto \sqrt{z^\beta - z}$ is non-Lipschitz. For instance the constant one is a solution, but a solution going from 1 to 0 exists as well (just get it by starting from a different starting point and see that it has to reach both 0 and 1 in finite time). We will cal $z_0$ the only solution with no flat part $z = 1$ around $t = 0$. This solution is also the only one to be injective on each of the two segments where it is positive and monotone. We claim that if we take the profile $z_0$ the last property we needed (i.e. weak convergence) is satisfied. This only requires that $\theta_0$ equals the integral of $z_0$, i.e. we need

$$\frac{\theta_0}{2} = \int_0^\infty z_0(t)dt = \int_0^1 \frac{z\,(1-\beta)\,\theta_0}{\sqrt{z^\beta - z}\,(2+\beta)\,2c_0}dz$$

(the integral has been computed by change of variables $z = z_0(t)$, which is possible for the solution since $z_0$ is injective on $[0, +\infty[\cap\{z_0 > 0\})$. This means that we must compare $C_0 := \int_0^1 \frac{z}{\sqrt{z^\beta - z}}dz$ and $c_0 = \int_0^1 \sqrt{z^\beta - z}dz$ and prove $C_0 = c_0(2+\beta)/(1-\beta)$.

Compute $c_0$ by integrating by part:

$$c_0 = \int_0^1 1 \cdot \sqrt{z^\beta - z}dz = -\int_0^1 z\frac{\beta z^{\beta-1} - 1}{2\sqrt{z^\beta - z}}dz = \frac{(1-\beta)}{2}C_0 - \frac{\beta}{2}c_0,$$

which implies $c_0 = C_0\frac{1-\beta}{2+\beta}$ and gives the desired equality.

Obviously, one needs after that to perform a correction of $u_\varepsilon$ near $x = 0$ and $x = 1$ so that the function actually belongs to $H^1$ (in order to avoid discontinuities at the two ends of the segment), and to control the extra energy one pays, as well as the divergence.

One possibility for the case of the single segment is the following: the profile we got for $v_\varepsilon$ is of the form $A_\varepsilon z(A_\varepsilon t)$ and one can simply replicate it radially on a half disk, so as to ensure regularity. Denote for simplicity by $f_\varepsilon(r)$ the radial profile one performs and by $B_\varepsilon$ the half ball where it is non-null. We need to estimate

$$\varepsilon^{\alpha-1}\int_{B_\varepsilon} |f_\varepsilon|^\beta + \varepsilon^{\alpha+1}\int_{B_\varepsilon} |f_\varepsilon'|^2, \quad \int_{B_\varepsilon} |f_\varepsilon'|, \quad \text{and} \quad \int_{B_\varepsilon} |f_\varepsilon|$$

Just use $|f_\varepsilon| \leq A_\varepsilon$, $|f_\varepsilon'| \leq CA_\varepsilon^2$ and $|B_\varepsilon| \leq CA_\varepsilon^{-2}$ (and the value of $A_\varepsilon$ which is of the order of $\varepsilon^{-2/(4-\beta)}$): both terms in the energy will be of the order of

17

$\varepsilon^{(\alpha+1)/3}$, the second term to be estimated (the divergence) will be bounded and the last will be of the order of $\varepsilon^{(\alpha+1)/3}$ as well.

In this way we have produced a sequence $u_\varepsilon$ that converges to $u$ (weak convergence is ensured by construction, weak convergence of $\nabla \cdot u_\varepsilon$ comes from the bound we just proved).

**Step 5** *Conclusion by density* Thanks to general properties of $\Gamma-$convergence, it is sufficient to build a recovery sequence for a class of limit measures $u$ which are "dense in energy". In this step we use the results of the previous step established for single segments to show that we can build these sequences for every $u$ concentrated on a finite graph, i.e. $u = u_G$, and then conclude.

As we pointed out in Section 2, it is well known from the works by Q. Xia on (see [27]) that the energy $M^\alpha$ is obtained as a relaxation of the same energy defined on finite graphs. This implies that the class of finite graphs is dense in energy (w.r.t. the energy $M^\alpha$) in the space $\mathcal{M}(\Omega)$ and $\Gamma-$convergence theory guarantees that it is enough to build recovery sequences for such a class (see [13]). One can also impose the condition $G \cap \partial\Omega = \emptyset$, so to avoid problems at the boundary.

For dealing with $u = u_G$, where $G$ is a finite graph, one can simply consider separately the segments composing $G$ and apply the previous construction of the previous step. Possible superpositions of the part of $u_\varepsilon$ coming from different segments will happen only on regions whose size is of the order of $A_\varepsilon^{-1}$ and hence negligible in the limit (and on such a region, one can use $|u_1+u_2|^\beta \leq |u_2|^\beta + |u_2|^\beta$ and $|\nabla(u_1 + u_2)|^2 \leq 2|\nabla u_1|^2 + 2|\nabla u_1|^2$). Not only, the number of nodes will be finite and hence the bound on the divergence will stay valid.

**Improvement** *Better connections at the junctions*

The construction for the $\Gamma-\limsup$ that we just detailed provides a sequence $u_\varepsilon$ converging to $u$ in $\mathcal{M}(\Omega)$ which works very well in the case of a single segment but which could be improved in general. Actually, in the case of a single segment the divergence of $u_\varepsilon$ has the same mass as that of $u$, since it replaces two Dirac masses at the two extremal points with two diffuse masses, concentrated on half-balls of radius $cA_\varepsilon^{-1}$ around the points. The mass of $u_\varepsilon$ itself is only slightly larger of the mass of $u$, due to the part we added at the extremities.

This changes a lot when one considers more than one segment, since the divergence of $u_\varepsilon$ at the nodes will be given by the superposition of different densities, each corresponding to a segment: the integral will vanish due to compensations, but $||\nabla \cdot u_\varepsilon||_{L^1}$ could have increased a lot.

This may be corrected thanks to the following lemma.

**Lemma 5.5.** *If $g \in C^1(B_R)$ is a function with zero mean on a ball of radius $R$ in two dimensions, then there exists a vector field in $C^1(\overline{B_R})$ such that $\nabla \cdot v = g$, $v = 0$ on $\partial B_R$ and*

$$||v||_{L^\infty} \leq C \left( R||g||_{L^\infty} + R^2||\nabla g||_{L^\infty} \right); \quad ||\nabla v||_{L^2}^2 \leq C \left( R^2||g||_{L^\infty}^2 + R^4||\nabla g||_{L^\infty}^2 \right),$$

*where $C$ is a universal constant.*

*Proof.* Take $w$ the solution of the elliptic problem

$$\begin{cases} \Delta w = g & \text{in } B_R, \\ w = 0 & \text{on } \partial B_R. \end{cases}$$

Thanks to $\int g = 0$ one has $\int_{\partial B_R} \partial w / \partial n = 0$. This allows to define a function $\phi : \partial B_R \to \mathbb{R}$ with $\phi' = \partial u / \partial n$. Take then a cut-off function $\chi(r)$ such that $\chi(r) = 1$ if $r \in [2R/3, R]$ and $\chi(r) = 0$ if $r \in [0, R/3]$, $|\chi| \le 1$, $|\chi'| \le C/R$, $|\chi''| \le C/R^2$ and define $\psi(x) = \chi(|x|)\phi(Rx/|x|)$.

Now take $v = \nabla w + Rot(\nabla \psi)$, where $Rot$ denotes a $90°$ clockwise rotation. In this way $\nabla \cdot v = \nabla \cdot \nabla w = g$, since the rotated gradient part is divergence-free, and both the normal and the tangential component of $v$ on the boundary vanish (since the tangential component of $\nabla \psi$ compensates the normal one of $\nabla w$ and the tangential component of $\nabla w$ and the normal of $\nabla \psi$ are zero).

We only need to check the bounds on the norms. These bounds come from standard elliptic regularity theory (see for instance [16]), since one has

$$\begin{aligned}
||\nabla w||_{L^\infty} &\le C\left(R||g||_{L^\infty} + R^2||\nabla g||_{L^\infty}\right), \\
||\phi||_{L^\infty} &\le CR||\nabla\phi||_{L^\infty} \le CR||\nabla w||_{L^\infty}, \\
||D^2 w||_{L^2}^2 &\le C||\Delta w||_{L^2}^2 = C||g||_{L^2}^2, \\
||D^2\phi||_{L^2}^2 &\le CR||D^2 w||_{L^2(\partial B_R)}^2 \le CR^2||D^3 w||_{L^2}^2 + C||D^2 w||_{L^2}^2 \\
&\le CR^2||\nabla g||_{L^2}^2 + C||g||_{L^2}^2.
\end{aligned}$$

The last line of inequalities come from the combination of a trace inequality in Sobolev spaces applied to $D^2 w$ (where the two coefficients of the gradient and the function part have different scaling with respect to $R$) with a regularity estimate for Dirichlet problems (estimating the $H^{k+2}$ norm of the solution with the $H^k$ norm of the datum). Combining all the ingredients give the desired estimate (we pass to the $L^\infty$ norms for the sake of simplicity). $\qquad\square$

With this lemma in mind, we can notice that at every node, the divergence of the function $u_\varepsilon$ we gave before is composed by the zero-mean sum of some functions $g_i$ of the form $f'(r)(x \cdot e_i)/r$ (this is the divergence of the vector field directed as $e_i$ with radial intensity we used above), where $e_i$ is the direction of the corresponding segment. Each function is supported on a half ball whose simmetry axes follows $e_i$ and the radius is of the order of $A_\varepsilon^{-1} \approx \varepsilon^{2/(4-\beta)}$. Setting for each node $g = \sum_i g_i$ one has $\int g = 0$, the support of $g$ is included in $B_R$ with $R \approx \varepsilon^{2/(4-\beta)}$ and $||g||_{L^\infty} \le CR^{-2}$ and $||\nabla g||_{L^\infty} \le CR^{-3}$.

Hence, one can add at every node a vector field $v$ as in Lemma 5.5 so as to erase the extra divergence. This would cost no more than a quantity of the order of

$$\varepsilon^{\alpha-1}R^2||v||_{L^\infty}^\beta + \varepsilon^{\alpha+1}||\nabla v||_{L^2}^2 \le C\varepsilon^{\alpha-1}R^{2-\beta} + C\varepsilon^{\alpha+1}\frac{1}{R^2} = C\varepsilon^{(\alpha+1)/3}.$$

Also the mass of the vector field is not that changed, since we only added $R^2||v||_{L^\infty} \le C\varepsilon^{(\alpha+1)/3}$.

# 6   The case $0 < \alpha \le 1/2$ in dimension 2

In Section 4 we provided an heuristic computation for the exponent $\beta$ to choose in order to prove the $\Gamma-$convergence result we look for, and, even in the simplest case $d = 2$ we found a value for $\beta$ which is negative for $\alpha \le \frac{1}{2}$. This prevented us to go on with such a case since we do not want to minimize terms involving $\int |u|^\beta$ for $\beta < 0$. Yet, it is possible to overcome this difficulty, if we want to approximate the enregy $M^\alpha$ and we accept to define the energies $M_\varepsilon^\alpha$ in a less homogeneous way. Even if this paper does not follow the strategy of [10], the fact that [10] proves the result for costs which are not exactly power functions, but asymptotically equivalent to a power, suggests that is possible to prove the following:

**Theorem 6.1.** *Suppose $d = 2$ and $\alpha \in ]0, 1/2[$; let $B : [0, +\infty[ \to [0, +\infty[$ a continuous function such that*

$$B(0) = 0, \ B > 0 \ on \ ]0, +\infty[ \ \lim_{t \to \infty} \frac{B(t)}{t^\beta} = 1, \ B'(0) > 0, \ \beta = \frac{2\alpha - 1}{\alpha + 1} \in ]-2, 0[.$$

*Define $M_\varepsilon^B$ through*

$$M_\varepsilon^B(u) = \varepsilon^{\alpha-1} \int_\Omega B(|u(x)|)dx + \varepsilon^{\alpha+1} \int_\Omega |\nabla u(x)|^2 dx,$$

*if $u \in H_0^1(\Omega; \mathbb{R}^2)$. Set $M_\varepsilon^B(u) = +\infty$ if $u \notin H_0^1 \subset \mathcal{M}(\Omega)$. Then we have $\Gamma-$convergence of the functionals $M_\varepsilon^B$ to $cM^\alpha$, with respect to the convergence of $\mathcal{M}(\Omega)$, as $\varepsilon \to 0$, where $c$ is a finite and positive constant (the value of $c$ being again $c = \alpha^{-1}(4c_0\alpha/(1 - \alpha))^{1-\alpha}$, being $c_0 = \int_0^1 \sqrt{t^\beta - t}dt$, where $c_0 \in \mathbb{R}$ is a consequence of $\beta \in ]-2, 0[$).*

*Remark* 4. Notice that the assumption $u = 0$ on $\partial\Omega$ was not present in the case $\alpha > \frac{1}{2}$. It has been added for technical reasons, to get a bound on $|u|$ which was guaranteed, in the case $\alpha > \frac{1}{2}$, by the fact that $\beta$ was positive and large values of $|u|$ were already penalized. Here this fact is replaced by an homogeneous boundary condition. Anyway, this is not a problem, since this only means that the flow must vanish and no divergence is permitted on $\partial\Omega$ (i.e. $f^+$ and $f^-$ may not give mass to the boundary). This is not restrictive as soon as one considers measures $f^\pm$ whose support is far from $\partial\Omega$: since this section is mainly written for applications purposes, we consider that this is always possible, up to enlarging the domain $\Omega$.

The proof may be obtained from the arguments in Section 5 under the some modifications.

$\Gamma - \liminf$ **part:**

**Modifications to step 1:**   Define $f_\varepsilon$ through $f_\varepsilon^2(t) = (B(t) - L_\varepsilon t)_+$, where the constant $L_\varepsilon$ is defined from $A_\varepsilon := \max_{\hat{R}_x} v_\varepsilon$ in the follwing way: $L_\varepsilon :=$

$\min\{B(t)/t : t \in ]0, A_\varepsilon]\}$, so that $B(v_\varepsilon) - L_\varepsilon v_\varepsilon \geq 0$ in $\hat{R}_x$. From the behaviour of the function $B$, we see that for large $A_\varepsilon$ the minimum above is realized at $t = A_\varepsilon$ and $L_\varepsilon A_\varepsilon = B(A_\varepsilon)$ (but anyway we have $L_\varepsilon A_\varepsilon \leq B(A_\varepsilon)$). The function $F_\varepsilon$ is defined as before as the primitive of $f_\varepsilon$ and it satisfies $F_\varepsilon(t) \leq Ct^{1+\beta/2}$. On the other hand, when computing $F_\varepsilon(A_\varepsilon)$ we get

$$
\begin{aligned}
F_\varepsilon(A_\varepsilon) &= \int_0^{A_\varepsilon} \sqrt{B(t) - L_\varepsilon t}\, dt = A_\varepsilon^{1+\beta/2} \int_0^1 \sqrt{\frac{B(A_\varepsilon t)}{A_\varepsilon^\beta} - \frac{L_\varepsilon A_\varepsilon}{A_\varepsilon^\beta} t}\, dt \\
&\geq A_\varepsilon^{1+\beta/2} \int_0^1 \sqrt{\frac{B(A_\varepsilon t)}{A_\varepsilon^\beta} - \frac{B(A_\varepsilon)}{A_\varepsilon^\beta} t}\, dt := c_0(A_\varepsilon) A_\varepsilon^{1+\beta/2},
\end{aligned}
$$

where $c_0(A_\varepsilon)$ converges to $c_0$ if $A_\varepsilon \to +\infty$.

**Modifications to Lemma 5.2:** We define $w_\varepsilon$ from $u_\varepsilon$ but, since $\beta < 0$, we can no more guarantee that $w_\varepsilon$ vanishes on a large zone of $\Omega$. Hence we need boundary conditions to guarantee the Sobolev-Poincaré inequality $||w_\varepsilon||_{L^r} \leq ||\nabla w_\varepsilon||_{L^2}$. Then, we use Hölder inequality to estimate the $L^{1+\beta/2}$ norm of $u_\varepsilon$:

$$
\int_R u_\varepsilon^{1+\beta/2} 1_{\{|u_\varepsilon| \geq 1\}} \leq \left( \int_\Omega |u_\varepsilon|^\beta 1_{\{|u_\varepsilon| \geq 1\}} \right)^{\frac{1}{p}} \left( \int_\Omega (1 + |w_\varepsilon|)^{(1+\beta/2-\beta/p)q} \right)^{\frac{1}{q}}.
$$

Then we go on by saying that, for $t \geq 1$, we have $t^\beta \leq CB(t)$, which allows to estimate the first factor by $\varepsilon^{(1-\alpha)/p}$ and using Sobolev-Poincaré on $w_\varepsilon$. The thesis is obtained in the same way as in Lemma 5.2, for $p < 2$

**Modifications to Equation** (10)**:** The minimization that we need to consider now is
$$
\min_{A \in ]0, +\infty[} \varepsilon^{\alpha-1} A^{\beta-1} m + 4c_0(A)\varepsilon^\alpha A^{1+\beta/2}.
$$

Here we can prove that the minimal value is given by a quantity $c_2(m/\varepsilon)m^\alpha$, for an increasing function $c_2(t) \leq c_2$, satisfying $\lim_{t \to \infty} c_2(t) = c_2$. This is possible if one notices that, either we have $A^{1-\beta} \leq (m/\varepsilon)^{1-\alpha}/c_2$, and in this case we get $\varepsilon^{\alpha-1} A^{\beta-1} m \geq c_2 m^\alpha \geq c_2(m/\varepsilon)m^\alpha$, or we have $A^{1-\beta} > (m/\varepsilon)^{1-\alpha}/c_2$. In this last case one can replace the value $c_0(A)$ in the left hand side with $c(m/\varepsilon) := c_0\left(((m/\varepsilon)^{1-\alpha}/c_2)^{1/(1-\beta)}\right)$, which is a quantity depending on $m/\varepsilon$ only, and no more on $A$. The same computation as in Section 5 gives the result. The constant in the optimal value is obtained from $c(m/\varepsilon)$ exactly as it was obtained before from $c_0$, and since $\lim_t c(t) = c_0$ our aim is achieved.

**Other modifications to step 2:** No special attention is needed, since the only difference is that we get, in (12), the term $\int c_2(m_{\varepsilon,\delta}(x)/\varepsilon)|m_{\varepsilon,\delta}(x)|^\alpha dx$. This term is not exactly the integral of the $\alpha$-th power, but it converges to, and Fatou's Lemma may be applied as well. From that point on, including lemma 5.3, the proof stays unchanged.

$\Gamma - \limsup$ **part:**

To produce a a recovery sequence $u_\varepsilon$ the strategy is the same as in the case $\beta > 0$: first define it in the case of a single segment, than glue the sequence on finite unions of segments and go on by density. In order to guarantee the additional constraint $u = 0$ on $\partial\Omega$, one just needs to restrict his attention to segments which are far from the boundary, but graphs composed of these segments are anyway a dense class. The sequence that will be chosen will be *exactly* the one described in the previous section, i.e. we will use profiles $z$ solving $z' = \pm\sqrt{z^\beta - z}$ (and not $z' = \pm\sqrt{B(z) - z}$). By solving this differential equation (which has non-uniqueness and, again, we take the solution with no flat part around $z = 1$) we get a profile $z_0$ which arrives to $z = 0$ with infinite slope. The same profile will be extended to zero. It is not evident that these functions belong to the space $H^1$ and the main estimate to prove it is

$$\int |z_0'(t)|^2 dt = \int |z_0'(t)|\sqrt{z_0(t)^\beta - z_0(t)}dt = 2\int_0^1 \sqrt{z^\beta - z}\,dz < +\infty,$$

where the same change of variable $z = z_0(t)$ has been used and the integral is finite since $\beta > -2$.

This sequence $u_\varepsilon$, based on the profile $z_0$, has the property that the energy $\varepsilon^{\alpha-1}\int \tilde{B}(|u_\varepsilon|) + \varepsilon^{\alpha+1}\int |\nabla u_\varepsilon|^2$ converges to $M^\alpha(u)$. Here we introduced the function $\tilde{B}$ given by $\tilde{B}(t) = t^\beta 1_{\{t>0\}}$, which is asymptotically equivalent to $B$ for $t \to \infty$ but completely different for $t \to 0$. Notice that this energy is not finite on every $H^1$ vector field (as it would be with $B$ instead of $\tilde{B}$), due to the degeneracy at $t = 0$. Yet, for this particular sequence of functions, which take the value 0 but soon leave the degeneracy zone (the profile has infinite slope close to its vanishing region), this energy turns out to be finite. Moreover, notice that $\tilde{B}(0) = B(0)$ and $\lim_{t\to\infty} B(t)/\tilde{B}(t) = 1$ and that, due to rescaling, the sequence $u_\varepsilon$ has the property that, for every fixed $M > 0$, the measure $|\{x \in \Omega : |u_\varepsilon|(x) \in ]0, M[\}|$ tends to zero. In this way one gets that the energy defined with $B$ and with $\tilde{B}$ have the same limit as $\varepsilon \to 0$ and that $u_\varepsilon$ is a recovery sequence for this energy as well.

# 7  Open questions and perspectives

Besides the interesting comparison aspects of this result with respect to the similar ones in the approximation of free discontinuity problems, one of the main goal of this study concerned possible numerical applications. To that purpose we want to replace the problem of minimizing $M^\alpha$ under divergence constraints with a simpler problem, i.e. minimizing $M_\varepsilon^\alpha$.

The idea would be to prove that the minimizers of

$$\min\left\{M_\varepsilon^\alpha(u) \ : \ \nabla \cdot u = f_\varepsilon\right\}, \tag{14}$$

$f_\varepsilon$ being a suitable approximation of $f = f^+ - f^-$, converge to the minimizers of (4).

Theorem 5.1 proves a $\Gamma-$convergence result which should give the convergence of the minimizers but the problem is that we never addressed the condition $\nabla \cdot u = f_\varepsilon$, nor we discussed the choice of $f_\varepsilon$. We will come back to this question, which is still open, later.

In [25] some suggestions to replace this divergence constraint with a penalization method are proposed. It is also possible to choose a penalization such that the penalized problem is also equivalent to the constrained one. Moreover the $\Gamma-$convergence result that we proosed in the space $\mathcal{M}(\Omega)$ without addressing the divergence constraintstays useful for a lot of problems where the divergence is not prescribed but enters the optimization (think at $\min_\mu d_\alpha(\mu, \nu) + F(\mu)$). Some of this problems are addressed in [24], for instance for urban planning or biological shape optimization.

Yet, we will not enter these issues into details here, but we find interesting to ask the following question:

**Open question 1.** *Given $f$, is it possible to find a suitable sequence $f_\varepsilon \rightharpoonup f$ so that one can prove $\Gamma-$convergence of the functionals $u \mapsto M_\varepsilon^\alpha(u) + I_{\nabla \cdot u = f_\varepsilon}$ to $u \mapsto M^\alpha(u) + I_{\nabla \cdot u = f}$ (being $I$ the indicator function in the convex analysis sense, i.e. $+\infty$ if the condition is not verified or zero if it is)? is it possible to find $f_\varepsilon$ explicitly, for instance as a convolution of $f$ with a given kernel?*

The second issue we want to address, after the one concerning divergence constraints, deals with the convergence of the minimizers. $\Gamma-$convergence is quite useless if we cannot deduce that the minimizers $u_\varepsilon$ converge, at least up to subsequences, to a minimizer $u$. Yet, this requires a little bit of compactness. The compactness we need is compactness in $\mathcal{M}(\Omega)$, i.e. we want bounds on the mass of $\nabla \cdot u_\varepsilon$ and of $u_\varepsilon$. The first bound, has been guaranteed by the fact that we decided to stick to the case of difference of probability measures. On the contrary, the bound on $|u_\varepsilon|(\Omega)$ has to be proven.

Notice that $M_\varepsilon^\alpha(u_\varepsilon) \leq C$ is not sufficient to obtain such a bound, as one can guess looking at the limit functional: think at a finite graph with a circle of length $l$ and mass $m$ on it, its energy is $m^\alpha l$ which provides no bound on $ml$ (its mass), if $m$ is allowed to be large. Actually, what happens on the limit functional is that "bounded energy configuration have not necessarily bounded mass, but optimal configuration do". This is due to the fact that, if $f^+$ and $f^-$ are probabilities, then $m \leq 1$ on optimal configurations (and no cycles are possible, by the way). Notice that this statement does not depend on $m \mapsto m^\alpha$ being concave, but simply increasing in $m$.

Again, a naive suggestion (i.e. adding a mass constraint) to overcome this difficulty is presented in [25], but the following question seems open and important :

**Open question 2.** *Prove a bound on the $L^1$ norm of the minimizers $u_\varepsilon$ (or on suitable minimizers $u_\varepsilon$, if needed). If possible, prove it for minimizers $u_\varepsilon$ which minimize $M_\varepsilon^\alpha$ under a divergence constraint $\nabla \cdot u_\varepsilon = f_\varepsilon$ (so that it will be true even if we add penalizations on the divergence).*

For finishing this questioning section, here is the last natural one:

**Open question 3.** *Prove the same results as in this paper or investigate what happens in $\mathbb{R}^d$, for $d \geq 3$.*

Notice that our proofs almost never used (up to the last improvement in Section 5) any gradient decomposition of vector fields in $\mathbb{R}^2$ such as $u = \nabla\phi + Rot\nabla\psi$. Yet, this would have been a typical trick for managing divergences in two dimensions. In our opinion the point where we used the most the fact that we are in $\mathbb{R}^2$ is when we disintegrate with respect to $x$ and we estimate the total variation in $y$ by the oscillation (which is a typical one-dimensional argument).

# 8 A numerical algorithm based on $\Gamma$-convergence regularization

Based on the ideas presented in the introduction, and on the $\Gamma-$convergence results proven in previous sections, we describe our optimization algorithm. In order to simplify the notations we restrict our description to the dimension $d = 2$ and we look for an optimal vector measure supported in $C = [0,1]^2$. It would be straightforward to adapt our method to other domains or even to the case $d = 3$ (but in dimension larger than two, no theoretical result is available for $\Gamma-$convergence). We decompose the domain $C$ into a $M^2$ cells grid with spacing $h = 1/M$. As it is standard in computational fluid dynamic, we discretized the unknown vector field on a staggered grid. More precisely, the first component $U_1$ of the vector field is placed on the midpoints of the vertical cells interfaces whereas the second components $U_2$ on the horizontal ones. Consequently the field $U_1$ is described by $(M + 1) \times M$ unknowns whereas $U_2$ is described by $M \times (M + 1)$ parameters.

We approximate the gradient at the centers of the cells of a function $U$ defined on a staggered grid by standard second order finite difference operators $\delta_x$ and $\delta_y$, defined for any discrete vector field $U$ by:

$$[\delta_x U]_{k,l} = \frac{U_{k+1,l} - U_{k,l}}{h}, \tag{15}$$

$$[\delta_y U]_{k,l} = \frac{U_{k,l+1} - U_{k,l}}{h}. \tag{16}$$

Notice that we keep the same notation for indexes $k, l$ which correspond to the midpoints of horizontal segments and the midpoints of vertical segments of the grid. If the index $(k, l)$ stands for a boundary point or to a point which is out of the grid, we assume that $U_{k,l} = 0$. The discrete cost functional $M_\varepsilon^d$ is directly deduced from (15) and (16) and a first order discretization of the integrals :

$$M_\varepsilon^d((U_i)_{k,l}) = \frac{1}{M^2} \sum_{k,l} \sum_{i=1}^{2} \varepsilon^{\gamma_1} \left( [\delta_x U_i]_{k,l}^2 + [\delta_y U_i]_{k,l}^2 \right) + \varepsilon^{\gamma_2} B(|(U_i)_{k,l}|) \tag{17}$$

where $B(r) = r^\beta$ when $\alpha > \frac{1}{2}$ and a non negative function having the required properties of theorem 6.1 if $\frac{1}{2} > \alpha > 0$. Practicality, we used a piecewise

defined $C^1$ function $B$ which is equal to $r^\beta$ for $r > r_0$ and a polynomial of degree 2 for $r \in [0, r_0]$. Additionally the divergence constraint, which is not involved in our $\Gamma$-convergence result, requires a careful numerical treatment. By the classical Helmholtz's decomposition, computing the projection on that constraint is equivalent to solve a problem of Poisson. In order to compute the solution of Poisson's problem efficiently, we implemented a Fast Fourier approach which has an almost linear complexity with respect to the number of points of the grid. Starting from an admissible vector field, we preserve the divergence constraint during the optimization process by the projection of the descent directions on the space of incompressible flows. To complete our description we give below the projection steps on the zero divergence constraint which will be apply to the gradient of the cost functional:

---

**Algorithm 1** Projection on a zero divergence constraint

---

**Require:** $(U_i)_{k,l}$
1: Define the divergence scalar field on the centers of the cells by $F_{k,l} := \nabla.(U_i)_{k,l}$ using (15) and (16).
2: Solve Poisson equation $\Delta P_{k,l} = F_{k,l}$ with Neumann's boundary conditions by Fast Fourier's Transform.
3: Update the vector field $(U_i)_{k,l} = (U_i)_{k,l} - \nabla P_{k,l}$

---

Finally, we give the successive steps of our Algorithm (we refer to [18] for technical details on the conjugated gradient algorithm and the choice of the line search methods).

---

**Algorithm 2** Numerical optimization by $\Gamma$-convergence

---

**Require:** $\varepsilon_{\text{initial}}$, $\varepsilon_{\text{final}}$, $(U_i^{\varepsilon_{\text{initial}}})$ satisfying the divergence constraint and parallel to $\partial C$, $\omega > 1$, $\delta$ (tolerance)
1: $\varepsilon := \varepsilon_{\text{initial}}$, $(U_i^\varepsilon) := (U_i^{\varepsilon_{\text{initial}}})$
2: **repeat**
3:   Compute up to a tolerance $\delta$, $(V_i^\varepsilon)$ the solution of $\min M_\varepsilon^d((V_i))$ among discrete vector fields $(V_i)$ which satisfy the boundary constraints and the divergence constraint. This step is carried out by a standard projected conjugated gradient algorithm (based on the previous projection steps) starting from $(U_i^\varepsilon)$.
4:   $(U_i^{\varepsilon/\omega}) := (V_i^\varepsilon)$, $\varepsilon := \varepsilon/\omega$
5: **until** $\varepsilon > \varepsilon_{\text{final}}$

---

# 9 Implementation and numerical results

We were able to run a series of computations on a grid of dimension $400 \times 400$ starting from randomly generated initial fields. In our experiments, we used the following parameters : $\varepsilon_{\text{initial}} = 1$, $\varepsilon_{\text{final}} = 1e - 2$, $\omega = 1.1$ and $\delta = 1e - 6$.

Notice that it has been possible to reach that fine tolerance criterion using an exact computation (up to round-off error) of the discrete gradient. Due to the irregularity of the vector field we want to approximate, a small error in the computation of the derivative may lead to a very unstable algorithm. Finally, let us point out the fact that the case $\frac{1}{2} > \alpha > 0$ is dramatically more difficult from a numerical point of view than the case $\alpha > \frac{1}{2}$. The difference of behavior comes from the non-monotonicity of the potential $B$ which is involved. In this situation, the number of expected local minima is larger than in the monotone cases. To tackle this difficult, we introduce at each iteration of algorithm 8 a rescaling of the potential $B$. Since asymptotically we expect that the modulus of the unknown vector field takes values close to 0 or in the range where $B(|U|)$ is equal to $|U|^\beta$ (i.e. for $|U| \geq r_0$) we force that aspect by the following procedure. At each iteration we update after step 5. our potential $B$ by setting $B(r) \leftarrow B(\frac{r_0 r}{Q_1})$ where $Q_1$ stands for the lower quartile of the values $(|U_i^\varepsilon|)$. This means that we change, through a dilation, our potential $B$ at every step. It is not difficult to convince ourselves that this corresponds to a change in the value of the parameter $\varepsilon$ : actually, the convergence being quite slow w.r.t. $\varepsilon$, we try to speed up it by selecting "well-chosen" pairs of potentials and initial conditions for the gradient descent.

We present below the results obtained with our simple approach. The following figures are the results of four different experiments with three different values of the parameter $\alpha$. On the first column of the figures, we represent two views of the graph of the given density $f_+ - f_-$. The other columns represent two views of the graph of the norm of the optimal vector field for each value of $\alpha$. As expected, "Kirchhoff's law is approximatively satisfied" by the support of the vector field which converges to a one dimensional set. Moreover we observe that two different values of $\alpha$ may lead to very different optimal structures.

We notice that "Steiner cases" which correspond to very small $\alpha$ satisfy more or less the angle condition for large multiplicity values whereas it happens that branches related to small values of the multiplicity are not completely satisfactory. This fact is due to the limitation of the method induced by the correlation between the grid size and the parameter $\varepsilon$. As a matter of fact, $\varepsilon$ can not be taken smaller than the square root of the spatial discretization to comply with the discretization error of the integrals.
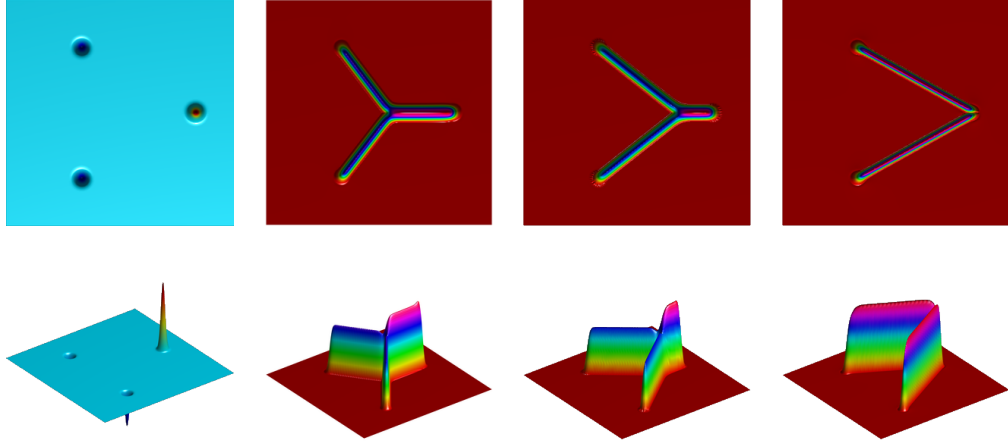
## Acknowledgements

Figure 1: Optimal irrigation of two discrete measures by one single source for $\alpha = 0.1, 0.6$ and $0.95$
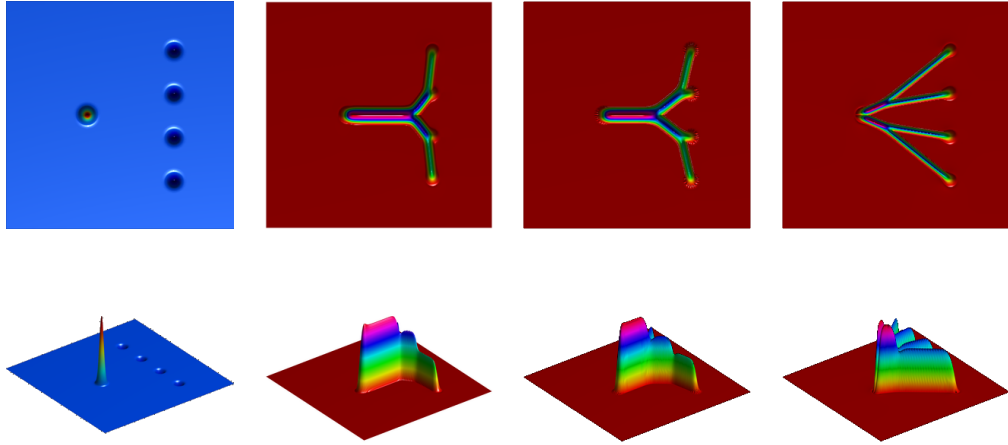


Figure 2: Optimal irrigation of four discrete measures by one single source for $\alpha = 0.1, 0.6$ and $0.95$
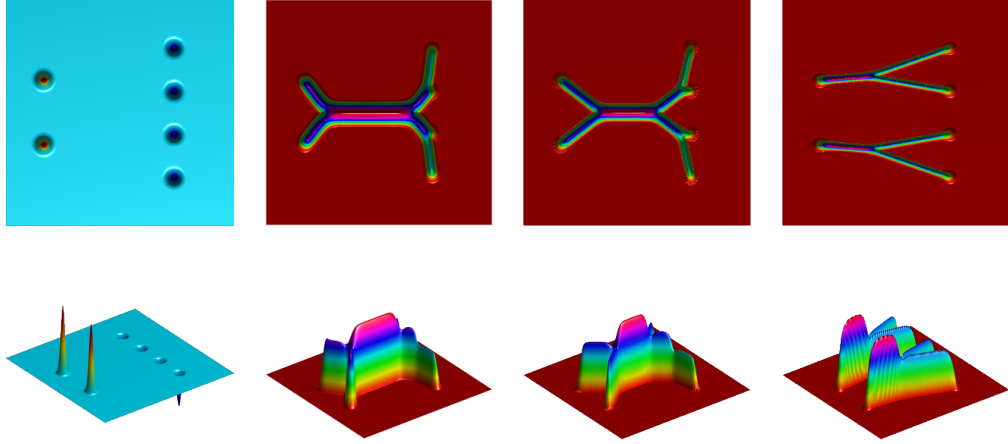
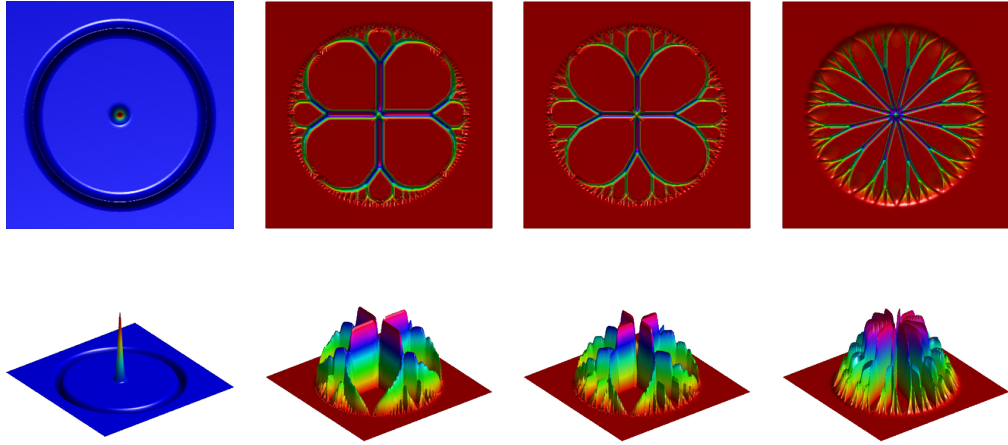Figure 3: Optimal irrigation of four discrete measures by two sources for $\alpha = 0.1, 0.6$ and $0.95$



Figure 4: Optimal irrigation of a measure supported on a circle by one single source for $\alpha = 0.6, 0.75$ and $0.95$

Italian Galilée Project "Allocation et Exploitation et Evolution Optimales des Ressources: réseaux, points et densités, modèles discrets et continus"

# References

[1] Luigi Ambrosio, Camillo De Lellis, and Carlo Mantegazza. Line energies for gradient vector fields in the plane. *Calc. Var. Partial Differential Equations*, 9(4):327–255, 1999.

[2] Luigi Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital. B (7)*, 6(1):105–123, 1992.

[3] Luigi Ambrosio and Vincenzo Maria Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.*, 43(8):999–1036, 1990.

[4] Patricio Aviles and Yoshikazu Giga. A mathematical problem related to the physical theory of liquid crystal configurations. In *Miniconference on geometry and partial differential equations, 2 (Canberra, 1986)*, volume 12 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 1–16. Austral. Nat. Univ., Canberra, 1987.

[5] Patricio Aviles and Yoshikazu Giga. Singularities and rank one properties of Hessian measures. *Duke Math. J.*, 58(2):441–467, 1989.

[6] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. Traffic plans. *Publ. Mat.*, 49(2):417–451, 2005.

[7] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. *Optimal transportation networks*, volume 1955 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Models and theory.

[8] Fabrice Bethuel, Haïm Brezis, and Frédéric Hélein. Asymptotics for the minimization of a Ginzburg-Landau functional. *Calc. Var. Partial Differential Equations*, 1(2):123–148, 1993.

[9] Fabrice Bethuel, Haïm Brezis, and Frédéric Hélein. *Ginzburg-Landau vortices*. Progress in Nonlinear Differential Equations and their Applications, 13. Birkhäuser Boston Inc., Boston, MA, 1994.

[10] Guy Bouchitté, Christophe Dubs, and Pierre Seppecher. Transitions de phases avec un potentiel dégénéré à l'infini, application à l'équilibre de petites gouttes. *C. R. Acad. Sci. Paris Sér. I Math.*, 323(9):1103–1108, 1996.

[11] Guy Bouchitté, Chloé Jimenez, and Mahadevan Rajesh. Asymptotique d'un problème de positionnement optimal. *C. R. Math. Acad. Sci. Paris*, 335(10):853–858, 2002.

[12] Andrea Braides. *Approximation of free-discontinuity problems*, volume 1694 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.

[13] Gianni Dal Maso. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[14] Ennio De Giorgi and Tullio Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 58(6):842–850, 1975.

[15] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[16] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, Berlin, 1977. Grundlehren der Mathematischen Wissenschaften, Vol. 224.

[17] E. N. Gilbert. Minimum cost communication networks. *Bell System Tech. J.*, 46:2209–2227, 1967.

[18] C. T. Kelley. *Iterative methods for optimization*, volume 18 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[19] F. Maddalena, S. Solimini, and J.-M. Morel. A variational model of irrigation patterns. *Interfaces Free Bound.*, 5(4):391–415, 2003.

[20] Luciano Modica and Stefano Mortola. Un esempio di $\Gamma^-$-convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.

[21] Sunra J. N. Mosconi and Paolo Tilli. Γ-convergence for the irrigation problem. *J. Convex Anal.*, 12(1):145–158, 2005.

[22] É. Oudet. Approximation of partitions of least perimeter by Γ-convergence : around kelvin's conjecture. `http://www.lama.univ-savoie.fr/~oudet`, 2009.

[23] I. Rodriguez-Iturbe and A. Rinaldo. *Fractal River Basins*. Cambridge University Press, 1997.

[24] Filippo Santambrogio. Optimal channel networks, landscape function and branched transport. *Interfaces Free Bound.*, 9(1):149–169, 2007.

[25] Filippo Santambrogio. A modica-mortola approximation for branched transport. *Comptes Rendus Acad. Sci.*, 348:941–945, 2010.

[26] Brian White. Rectifiability of flat chains. *Ann. of Math. (2)*, 150(1):165–184, 1999.

[27] Qinglan Xia. Optimal paths related to transport problems. *Commun. Contemp. Math.*, 5(2):251–279, 2003.

[28] Qinglan Xia. Interior regularity of optimal transport paths. *Calc. Var. Partial Differential Equations*, 20(3):283–299, 2004.

[29] Guoliang Xue, Theodore P. Lillys, and David E. Dougherty. Computing the minimum cost pipe network interconnecting one sink and many sources. *SIAM J. Optim.*, 10(1):22–42 (electronic), 1999.