A Modica-Mortola approximation for branched transport

Filippo Santambrogio

A Modica-Mortola approximation for branched transport

Filippo Santambrogio *

September 16, 2009

Abstract

The M^α energy which is usually minimized in branched transport problems among singular 1-dimensional rectifiable vector measures with prescribed divergence is approximated (and convergence is proved) by means of a sequence of elliptic energies, defined on more regular vector fields. The procedure recalls the Modica-Mortola one for approximating the perimeter, and the double-well potential is replaced by a concave power.

1 Introduction

The name "branched transport" is now often used for addressing all the transport problems where the cost for a mass m moving on a distance l is proportional to l but not to m but sub-additive and typically proportional to a power m^α ($0 < \alpha < 1$). The distributions of sources and destinations are given and one looks for the path followed by each particle, sums up the mass which moves together on each part of the path, and associates to every configuration its total cost $\sum_i l_i m_i^\alpha$. The adjective "branched" in the name stands for one of the main features of the optimal solutions: they gather mass together, masses tend to move jointly as long as possible, and then they branch towards different destinations, thus giving rise to a tree-shaped structure.

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 [16]). 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 ([24]), which will be our main reference. Other approaches have been proposed by Bernot, Caselles, Maddalena, Morel and Solimini (see [17, 6]). The equivalences between the different models as well as a survey on the whole theory are presented in a

*CEREMADE, Université Paris-Dauphine, Place de Lattre de Tassigny, F-75775 Paris Cedex 16, filippo@ceremade.dauphine.fr

recent book by Bernot, Caselles and Morel, [7], who are responsible for most of the results therein.

A satisfactory numerical treatment of these issues is far from being obtained and this is one of the reasons for the need for approximation results. This paper presents a precise approximation result, in terms of $\Gamma-$convergence (see [12]), which can be the origin of possible numerical methods based on PDEs, after an idea by Jean-Michel Morel. Actually, the continuous formulation of branched transport problems passes through a divergence-constrained formulation (see next session, where the model by Xia is quickly sketched) and 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 term and a Dirichlet term, with coefficients which increase the weight of the first and vanish on the latter as far the approximation parameter ε goes to 0. In this way one would need to solve a problem of the kind

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

under constraints or penalization on $\nabla \cdot u$. The optimality conditions for this problem read as an elliptic vector PDE (a system) with an unknown pressure coming from the divergence constraint. In dimension two, the decomposition of any vector field into a gradient plus an orthogonal gradient allows to translate this into a system of scalar fourth-order elliptic equations. Local minima may be found by a gradient descent method passing through a parabolic evolution equations, while global ones are more difficult to detect due to the concavity of F. Numerical methods for this kind of equations are much more developed and efficient rather than the combinatory-based ones which characterize the study of finite networks. A partial use of genetics algorithm could thereafter help in avoiding local minima.

Besides possibles numerical applications, the interest of this $\Gamma-$convergence result also comes from its comparison with the elliptic approximation of the perimeter functional, proposed by Modica and Mortola ([18]) at the beginning of $\Gamma-$convergence time. In their case u was scalar and F was a double-well potential, enforcing at the limit u (after a suitable rescaling of the functional, so that the F-part has a coefficient going to infinity and the Dirichlet part a negligible one) to take values in $\{0, 1\}$. And the energy was, in the limit as $\varepsilon \to 0$, concentrated on a lower-dimensional structure, i.e. the interface between the two phases $u = 0$ and $u = 1$. The same will happen here: the concave power F will play the role of a double well at $u = 0$ and $|u| = \infty$ and the energy will concentrate on a one-dimensional graph.

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

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

is first addressed). Yet, it is easy to notice, due to the divergence which is a bounded measure, that the problem is essentially scalar, as at the limit there will be locally one direction only which will be relevant. Actually a singular vector measure, concentrated on a lower-dimensional object, must be oriented along a tangent direction if we want its divergence to be a measure (otherwise it is 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 information on the divergence "fix" the gradient part and all the functional may be expressed through the other gradient. In this way one would arrive to consider the limits of something like

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

which are functionals 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 the paper is creating a 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 [21]). Elliptic PDEs, dimensional reduction and geometric measure theory are very much involved as well. See [2, 3, 11, 18] 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, in the sense that the rectifiable graphs in the limit problem vanish almost everywhere and are concentrated on one-dimensional sets, thus having a double jump, whose "intensity" (measured with respect to $\mathcal{H}^1$ instead of $\mathcal{L}^d$) enters the limit functional.

The paper will start with two short sections on preliminaries, Section 2 on Xia's formulation of branched transport problems, Section 3 on $\Gamma-$convergence. Then, we will discuss briefly in Section 4 why to choose the form of the approximation we will choose, mainly concentrating on the choice of the exponents, since they are not obvious (in particular, the function F in (1) will be of the form $F(|u|) = |u|^\beta$ with $\beta \neq \alpha$). Section 5 will present the detailed proof of the main result ($\Gamma-$convergence of the energies, in dimension two only, under no divergence conditions). Section 6 will suggest how to apply the result for producing interesting approximated variational problems, underlining what can be done with the tools we have so far and what could be worthwhile to prove in possible future investigations.

2 Branched transport via divergence-constrained optimization

We present here the framework of the optimization problem proposed by Xia in [24, 25] and then studied by 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 ϕ vanishing 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 mass of the measure "total variation of the divergence" of u. On this space we consider the convergence $u_\varepsilon\rightharpoonup u \Leftrightarrow u_\varepsilon\rightharpoonup u$ 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 of absolutely continuous measures (with respect to the Lebesgue measure on Ω) and the functions represent their densities.

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) $u=\theta\xi\cdot\mathcal{H}^1_{\lfloor M}$ 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 $\xi: M\to\mathbb{R}^d$, ξ 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^α 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 absolutely continuous inside Ω 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^α: Xia proposed it in [24] as a relaxation from the case of finite graphs, but formula (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

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

(and the convergence is in the sense of $\mathcal{M}(\Omega)$). For the proof of the equivalences between the two definition, look at [25] 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 finite positive measures), provided $\alpha > 1 - 1/d$ (this is proven in [24] by means of an explicit construction).

3 Variational approximation, preliminaries

The main result of the paper will be a $\Gamma-$convergence result for a sequence of energies approximating M^α. 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 to prove compactness for a suitable sequence of minimizers u_ε 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 [13] and [12]).

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 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, intuition and 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 [18] and [11]) 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_ε 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_ε $\Gamma-$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 that one of the key-ingredient in the proof of the above Theorem 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 $F_\varepsilon(u) \geq 2TV(H \circ u)$).

In our study we will instead 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 two exponents $\gamma_1 < 0 < \gamma_2$.

As one can see the 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 concave powers, in their minimization, if the average value for u is fixed in a region (which is in some sense the meaning of weak convergence, i.e. the convergence we use on $\mathcal{M}(\Omega)$), prefer either $u = 0$ or $|u|$ being as large as possible, i.e. there is sort of a double well on zero and infinity.

We give here a heuristics for determining the exponents β, γ_1 and γ_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 (for instance by convolution).

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 of the order of

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

In the minimization one will choose the optimal width A, depending on m and ε, and this gives

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

The correct choice for a possible convergence result towards the energy (3) which is proportional to m^α is obtained by imposing that the exponent of m is α and the exponent of ε is zero, i.e.

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

Notice that γ_1 and γ_2 may not both be determined since one can always replace ε with a power of ε, thus changing the single exponents but not their ratio. Notice also that the exponent β is positive and less than 1 as soon as $\alpha \in]1-1/d, 1[$, which is the usual condition.

Finally, it is worthwhile to remark that for this choice of exponents the dependence of A with respect to ε 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 a strip of width A without changing its mass) to u. It proves nothing, but it is coherent with the convergence result we want to prove.

We conclude by underlining the case of dimension 2, since it will be the only one we will solve: in this case one has $\beta = \frac{4\alpha-2}{\alpha+1}$ and $\gamma_1/\gamma_2 = (\alpha-1)/(\alpha+1)$.

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 $\alpha > 1 - 1/D$. This is quite astonishing if one thinks that, also for $\alpha \leq 1 - 1/d$, there are measures $f^{\pm}$ which admit possible solution with finite energy (in particular when both the measures are supported on a same lower-dimensional set). Yet, the problem lies in the kind of approximation we require, which uses measures u_ε which are more regular and in particular are absolutely continuous with respect to $\mathcal{L}^d$ and may not be concentrated on lower-dimensional sets.

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^α_ε to cM^α, 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 $M^\alpha_\varepsilon(u_\varepsilon)$ is bounded. This implies at first that all the u_ε are H^1 functions, and that

$$\int |u_\varepsilon|^\beta \leq C\varepsilon^{1-\alpha} \quad \text{and} \quad \int |\nabla u_\varepsilon|^2 \leq C\varepsilon^{-1-\alpha}. \tag{7}$$

Step 1 *Lower bounds on M^α_ε*

Consider a rectangle $R \subset \Omega$. Suppose for simplicity that it is oriented according to the x and y axes, i.e. $R = [a, b] \times [c, d]$, and set $v_\varepsilon := [(u_\varepsilon)_x]_+$ (the positive part of the $x-$component of u_ε) and $v'_\varepsilon := \partial v_\varepsilon/\partial y$. Consider also a function $\phi_1 : R \to [0, 1]$, depending on y only, such that $\phi_1 = 1$ on $[a, b] \times [c + \delta, d - \delta]$ and $\phi_1 = 0$ for $y \in \{c, d\}$. For every x, set $R_x := \{x\} \times [c, d]$ and $R'_x := \{x\} \times [c + \delta, d - \delta]$.

Fix a value of $x \in [a, b]$ and let A_ε be the maximal value of v_ε on R'_x (which is well-defined for a.e. x) and $L_\varepsilon = A^{\beta-1}_\varepsilon$, $f_\varepsilon(t) = \sqrt{(t^\beta - L_\varepsilon t)_+}$ and $F_\varepsilon(t) = \int_0^t f_\varepsilon(s)ds$. One can write

$$v^\beta_\varepsilon = f^2_\varepsilon(v_\varepsilon) + L_\varepsilon v_\varepsilon - (v^\beta_\varepsilon - L_\varepsilon v_\varepsilon)_- \geq f^2_\varepsilon(v_\varepsilon) + L_\varepsilon v_\varepsilon - L_\varepsilon v_\varepsilon I_{R_x \setminus R'_x} = f^2_\varepsilon(v_\varepsilon) + L_\varepsilon v_\varepsilon I_{R'_x}$$

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

$$\begin{aligned}\varepsilon^{\alpha-1} v^\beta_\varepsilon + \varepsilon^{\alpha+1}(v'_\varepsilon)^2 &\geq \varepsilon^{\alpha-1} L_\varepsilon v_\varepsilon I_{R'_x} + \varepsilon^{\alpha-1} f^2_\varepsilon(v_\varepsilon) + \varepsilon^{\alpha+1}(v'_\varepsilon)^2 \\ &\geq \varepsilon^{\alpha-1} L_\varepsilon v_\varepsilon I_{R'_x} + 2\varepsilon^\alpha f_\varepsilon(v_\varepsilon)v'_\varepsilon = \varepsilon^{\alpha-1} L_\varepsilon v_\varepsilon I_{R'_x} + 2\varepsilon^\alpha (F_\varepsilon(v_\varepsilon))'.\end{aligned}$$

By multiplying times ϕ_1 and integrating on R_x with respect to y, one has

$$\int_{R_x} \left(\varepsilon^{\alpha-1} v_\varepsilon^\beta + \varepsilon^{\alpha+1}(v_\varepsilon')^2\right)\phi_1(y)dy$$
$$\geq \varepsilon^{\alpha-1} L_\varepsilon \int_{R_x'} v_\varepsilon dy + 2\varepsilon^\alpha TV(F_\varepsilon(v_\varepsilon)\phi_1) - 2\varepsilon^\alpha \int_{R_x\setminus R_x'} F_\varepsilon(v_\varepsilon)|\phi_1'|dy. \quad (8)$$

Since $v_\varepsilon\phi_1$ vanishes at both boundaries and reaches the value $F_\varepsilon(A_\varepsilon)$ inside 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 (8) will tend to zero as $\varepsilon \to 0$, after integration with respect to x, thanks to Lemma 5.2 below, since ϕ_1 is a fixed regular function and hence ϕ_1' is bounded, $F_\varepsilon(t) \leq \int_0^t s^{\beta/2} ds = ct^{1+\beta/2}$ and $|v_\varepsilon| \leq |u_\varepsilon|$.

Lemma 5.2. *For any bounded energy sequence u_ε we have*

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

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

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

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

Now, for any pair of conjugate exponents p and q, we have

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

Hence we have

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

where the exponent γ_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. □

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_ε) 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_ε is seen as a measure on $\overline{\Omega}$ which belongs to $L^2(\Omega) + L^2(\partial\Omega)$.

Hence, we will ignore the last term and, for every x, we look at the quantity

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

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

$$\min_{A\in]0,+\infty[} \varepsilon^{\alpha-1} A^{\beta-1} m + 4c_0 \varepsilon^\alpha A^{1+\beta/2} = c_2 m^\alpha, \tag{9}$$

the minimum being realized by

$$A = \left(\frac{m(1-\beta)}{\varepsilon 2c_0(2+\beta)} \right)^{2/(4-\beta)}. \tag{10}$$

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 (9) we need to use the relations between β and α.

This is the first point where we start seeing an expression recalling somehow M^α.

Step 2 *Comparison with* M^α

One can call μ_ε 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 μ_ε stay bounded in the set of positive Radon measures on Ω. Hence we can suppose $\mu_\varepsilon \rightharpoonup \mu$.

As we did previously, take a rectangle R and keep the same notations. Keeping the positive $x-$part only and the derivative with respect to y only one has $\mu_\varepsilon \geq \left(\varepsilon^{\alpha-1} v_\varepsilon^\beta + \varepsilon^{\alpha+1}|v'_\varepsilon|^2\right) \cdot \mathcal{L}^2$, hence we get

$$\int_R \phi_1(y) d\mu_\varepsilon \geq \int_a^b [m'_\varepsilon(x)]^\alpha dx - r_\varepsilon, \tag{11}$$

r_ε being a negligible rest corresponding to the last term in (8).

We would like this estimate to pass to the limit as $\varepsilon \to 0$, using $u_\varepsilon \to u$. This is not yet possible: to get it, take again a function $\phi_2 : R \to [0,1]$, similar to ϕ_1, depending on y only, but such that $\phi_2 = 1$ on $[a,b] \times [c+2\delta, d-2\delta]$ and $\phi_2(y) = 0$ for $y \in [c, c+2\delta] \cup [d-2\delta, d]$. For every x, set

$$m''_{\varepsilon,+}(x) := \int_{R_x} v_\varepsilon \phi_2 dy,\ m''_{\varepsilon,-}(x) := \int_{R_x} [(u_\varepsilon)_x]_- \phi_2 dy,\ m_\varepsilon(x) := \int_{R_x} (u_\varepsilon)_x \phi_2 dy.$$

Notice that estimate (11) implies $\int_R \phi_1(y) d\mu_\varepsilon \geq \int_a^b [m''_{\varepsilon,+}(x)]^\alpha dx$ (up to the negligible term r_ε). Moreover, performing for every x the same proof as before with

the negative part instead of the positive one, and using $\max\{m''_{\varepsilon,+}(x), m''_{\varepsilon,-}(x)\} \geq |m_\varepsilon(x)|$, one gets

$$\int_R \phi_1(y) d\mu_\varepsilon \geq \int_a^b |m_\varepsilon(x)|^\alpha dx. \tag{12}$$

The measures $m_\varepsilon(x)dx$ 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 the right hand side of (12). 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 $(\pi_x)_\#(\phi_2(u_\varepsilon)_x)$).

What we may prove is that the functions $x \mapsto m_\varepsilon(x)$ are uniformly BV in x and this will allow for L^1 and pointwise convergence. We will use the fact that $u_\varepsilon \to u$ in $\mathcal{M}(\Omega)$ implies a bound on $\nabla \cdot u_\varepsilon$. Take a function $\psi : [a,b] \to \mathbb{R}$ with $\psi(a) = \psi(b) = 0$. Consider

$$\begin{aligned}
\int_a^b m_\varepsilon(x)\psi'(x)dx &= \int_R (u_\varepsilon)_x(x,y)\phi_2(y)\psi'(x)\, dydx \\
&= \int_R u_\varepsilon(x,y) \cdot \nabla(\phi_2(y)\psi(x))dydx - \int_R (u_\varepsilon)_y(x,y)\phi_2'(y)\psi(x)\, dydx \\
&\leq |\nabla \cdot u_\varepsilon|(\overline{\Omega})||\psi||_{L^\infty(a,b)} + ||u_\varepsilon||_{L^1}||\phi_2'||_{L^\infty(a,b)}||\psi||_{L^\infty(a,b)} \leq C||\psi||_{L^\infty(a,b)}.
\end{aligned}$$

This proves the BV bound we needed and implies that $(\pi_x)_\#(\phi u_x)$ is a measure on $[a,b]$ which "belongs to BV" (i.e. is absolutely continuous and has a BV density). Let us call $m(x)$ its density. Moreover, one has $m_\varepsilon(x) \to m(x)$ for almost any x. Passing to the limit in (12) as $\varepsilon \to 0$ one gets, by Fatou's Lemma,

$$\int_R \phi_2(y) d\mu \geq \int_a^b |m(x)|^\alpha dx.$$

It is quite straightforward that one can let δ go to 0 and get rid of the functions ϕ_1 and ϕ_2 (which actually depend on δ). This gives

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

where $\tilde{m}(x)$ is defined as the density at x of the projection π_x of the measure u_x restrained at $[a,b]\times]c,d[$. Notice that this measure has a density as one can realize by writing it as a series of measures with densities with finite series of L^1 norms (i.e. we take a sequence of δ_n going to zero and see that the corresponding m_n are all L^1 functions, and the series $\sum_n ||m_{n+1} - m_n||_{L^1}$ converges since $\sum_n \int |(\phi_2)_{n+1} - (\phi_2)_n|d|u_x| < +\infty$).

It is useful to see that this estimate implies that u, simply because of its attainability as a limit of u_ε with bounded M_ε^α energies, is a rectifiable measure. This is proven thanks to the 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 [14, 23]) and already used to prove rectifiability of measures with finite $M^\alpha-$mass (see also [25]). This theorem roughly states that u is rectifiable if and only if almost all its $(d-1)-$dimensional slices parallel to the coordinate axis are countable collections of Dirac masses. Take the coordinate axis and disintegrate u_x with respect to its projection $m(x)dx$ on the variable x, thus getting signed measures ν_x. For every n, divide Ω into 2^n horizontal strips of equal width and call $f_{i,n}(x)$ the integral of ν_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 ν_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 ν 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 ν is purely atomic.

Proof. Let us start from proving $G_\alpha^{(sup)}(\nu) \geq G_\alpha(\nu)$: for any n, build a measure ν_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 ν on $I_{i,n}$. This measures converge weakly to ν, 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

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

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 $\mu \geq \theta^\alpha \cdot \mathcal{H}^1_{\lfloor M}$, which implies the thesis.

This kind of proofs follow a standard scheme, i.e. considering measures μ_ε whose mass gives the value of the approximating energy, and providing estimates on the limit measure μ. 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 μ (here it is the density w.r.t. $\mathcal{H}^1$). Similar proofs (typically comparing μ 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 [10] and [20]

5.2 $\Gamma - \limsup$ inequality

Step 3 *The case of a single segment*

Consider the case $u = \theta \cdot \mathcal{H}^1_{\lfloor S}$, being S a segment (for simplicity, $S = [0,1] \times \{0\}$). For producing a "recovery sequence" one can inspire himself at the lower bound proof. We will look for a profile u_ε with the following properties:

- the $x-$component only of u_ε must be present and must have the same sign, so that $u_\varepsilon = v_\varepsilon e_1$;
- $\nabla v_\varepsilon = v_\varepsilon' e_2$ (i.e. v_ε only depends on y);
- the Cauchy-Schwartz inequality used in (8) 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_ε 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_ε must vanish at the boundary of a certain rectangle, so that one gan avoid using the function ϕ_1;
- the maximum A_ε must be optimal in (9);

- as $\varepsilon \to 0$ weak convergence to the measure $u = \theta \cdot \mathcal{H}^1_{|S}$ 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 $\int_{\mathbb{R}} z(t)dt = \theta$ 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_ε will be A_ε, realized at zero, and the 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 (10), the ratio $A_\varepsilon^{\beta/2}/(\varepsilon A_\varepsilon^2)$ equals $(1-\beta)\theta/(2c_0(2+\beta))$. This is very good since there is no more ε.

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). This means that there are several solutions, and $\int_0^\infty z(t)dt$ may be any possible value larger than $\int_0^\infty z_0(t)dt$ (z_0 being the only solution with no flat part $z = 1$ around $t = 0$).

We want now to check that θ is larger than the lower bound for the integrals, and this gives the final property that z needed to satisfy.

We need

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

(the integral has been computed by change of variables $z = z_0(t)$, which is possible for the solution z_0 only, since it is injective). 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 $C_0 \leq c_0(2+\beta)/(1-\beta)$.

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 is enough to get the desired inequality.

Obviously, one needs after that to perform a correction of u_ε 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_ε 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. In this case we need to estimate. Denote for simplicity by $f_\varepsilon(r)$ the radial profile one performs and by B_ε the half ball

where it is non-null. We need to estimate

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

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_ε which is of the order of $\varepsilon^{-2/(4-\beta)}$): both terms in the energy will be of the order of $\varepsilon^{(\alpha+1)/3}$, 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_ε 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 4 *Conclusion by density*

As we pointed out in Section 2, it is well known from the works by Q. Xia on (see [24]) that the energy M^α 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" in the space $\mathcal{M}(\Omega)$ and $\Gamma-$convergence theory guarantees that it is enough to build recovery sequences for such a class (see [12]). 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_ε coming from different segments will happen only on regions whose size is of the order of A_ε^{-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_ε 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_ε 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_ε^{-1} around the points. The mass of u_ε 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_ε 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, then there exists a vector field in $C^1(\overline{B_R})$ such that $\nabla\cdot v=g$, $v=0$ on ∂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 $r \in [0, R/3]$, $|\chi| \leq 1$, $|\chi'| \leq C/R$, $|\chi''| \leq C/R^2$ 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 ∇ and the tangential component of ∇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 [15]), since one has

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

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 al the ingredients give the desired estimate (we pass to the L^∞ norms for the sake of simplicity). □

With this lemma in mind, we can notice that at every node, the divergence of the function u_ε 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 $||g||_{L^\infty} \leq CR^{-2}$ and $||\nabla g||_{L^\infty} \leq CR^{-3}$.

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 something of the order of

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

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

6 Applications 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, as we mentioned in the introduction.

We want to replace the problem of minimizing M^α under divergence constraints with a simpler problem, i.e. minimizing M^α_ε.

The idea would be solving

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

being f_ε a suitable approximation of $f = f^+ - f^-$., and proving that the minimizers of (14) 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_ε. We will come back to this question, which is still open, later.

What we will consider here are penalization methods, which are quite natural as far as numerics is concerned. One could decide to replace the (quite severe) condition $\nabla\cdot u = f$ at the limit with a weaker one, concerning a distance between $\nabla\cdot u$ and f, possibly with a very high penalization coefficient. Considering problems of the form

$$\min\quad M^\alpha_\varepsilon(u) + G(\nabla\cdot u; f), \tag{15}$$

where G is a functional defined on pairs of finite measures, which is continuous w.r.t. weak convergence, and such that $G((\mu,\nu),(f^+,f^-)) = 0$ if and only if $\mu = \nu$, would play the game, since the $\Gamma-$convergence result could be easily extended, thanks to the last property of this kind of convergence that we presented in Section 3.

Yet, this would converge to

$$\min\quad M^\alpha(u) + G(\nabla\cdot u; f), \tag{16}$$

which in general is not exactly the same as imposing $\nabla\cdot u = f$. If on the contrary one wants a penalization of the kind $\varepsilon^{-1}G$, so that at the limit one needs $G = 0$, i.e. $\nabla\cdot u = f$, then the whole $\Gamma-$convergence result would have to be re-established (and would present some difficulties), since we are no more in the same framework of adding a same continuous functional.

Hence, we will stick to the case of a single penalization and we will start from the following consideration which works quite well if the divergence is given by the difference of two fixed mass positive measures.

As explained in [24], the quantity $d_\alpha(\mu,\nu) := \min\{M^\alpha(u)\ :\ \nabla\cdot u = \mu-\nu\}$ is a distance on the set of positive measures with the same finite mass that metrizes weak convergence.

As a consequence of d_α being a distance, due to triangular inequality, solving

$$\min_{\mu^+,\mu^-\in\mathcal{P}(\overline{\Omega})} 2d_\alpha(f^+,\mu^+) + \min\{M^\alpha(u)\ :\ \nabla\cdot u = \mu^+-\mu^-\} + 2d_\alpha(f^-,\mu^-)$$

amount to choosing $\mu^\pm = f^\pm$ and solving $\min\{M^\alpha(u) \ : \ \nabla \cdot u = f^+ - f^-\}$.

This means that choosing $2d_\alpha$ as a penalization is a clever strategy (even if it only allows for considering probabilities or, in general, fixed mass measures).

To describe a precise problem where we can insert this penalization we need to introduce a slightly different space. Consider the space $Y(\Omega) \subset \mathcal{M}(\Omega) \times \mathcal{P}(\overline{\Omega}) \times \mathcal{P}(\overline{\Omega})$ defined by

$$Y(\Omega) := \{(u, \mu, \nu) \in \mathcal{M}(\Omega) \times \mathcal{P}(\overline{\Omega}) \times \mathcal{P}(\overline{\Omega}) \ : \ \nabla \cdot u = \mu - \nu\},$$

endowed with the obvious topology of componentwise weak convergence.

On this space consider the functionals

$$(u, \mu, \nu) \mapsto M^\alpha_\varepsilon(u) + 2d_\alpha(f^+, \mu) + 2d_\alpha(f^-, \mu). \tag{17}$$

It is an easy consequence of our previous $\Gamma-$convergence result and of the fact that the quantity $G((\mu, \nu), (f^+, f^-)) := 2d_\alpha(f^+, \mu) + 2d_\alpha(f^-, \mu)$ is continuous on this space that we still have $\Gamma-$convergence to the functional $M^\alpha(u) + G((\mu, \nu), (f^+, f^-))$. The minimization of this last functional being equivalent to the minimization of $M^\alpha(u)$ under the divergence constraint $\nabla \cdot u = \mu - \nu$ one has obtained a useful approximation of any branched transport problem.

Notice that in this framework the imprvement that we proposed at the end of Section 5 is crucial: without this kind of correction on the divergence, one would got approximating measures u_ε (in the $\Gamma - \limsup$ construction) whose divergence is no more the difference of two probabilities, since some extra mass at every node is added both to the positive and to the negative part.

Nevertheless, the approximation is not satisfactory yet, especially for numerical purposes: we are trying to suggest methods to approximate the minimization of M^α and we propose to use the distance d_α itself!! For computing it one should probably solve a problem of the same kind and no progress would have been done.

A possible escape we suggest is replacing d_α with other quantities, which are larger but still vanish when the two measures coincide. This is possible for instance thanks to [19], where the inequality $d_\alpha \leq CW_1^{2\alpha-1}$ is proven, W_1 being the usual Wasserstein distance on $\mathcal{P}(\overline{\Omega})$. We will call G_1 the functional $G_1((\mu, \nu), (f^+, f^-)) := CW_1^{2\alpha-1}(\mu, f^+) + CW_1^{2\alpha-1}(\nu, f^-)$. The distance W_1 has the advantages of being independent of the branched transport problem and of admitting more numerical methods for computing it (it is, by the way, a dual Lipschitz norm). Notice that other possible attachment terms, easier to compute, such as the L^2 norm, are not adequate because they are not continuous with respect to the convergence we have (and because in general we do not want to require $f \in L^2$).

It is anyway important to stress that, the approach on $\mathcal{M}(\Omega)$ without penalization stays 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 [22], for instance for urban planning or biological shape optimization.

Yet, even if there are, as we showed, was to overcome the problem, we find anyway 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 $u \mapsto M^\alpha_\varepsilon(u) + I_{\nabla\cdot u = f_\varepsilon}$ 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_ε 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_ε 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_ε. 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 $M^\alpha_\varepsilon(u_\varepsilon) \leq C$ will not be sufficient for 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 "bounded energy configuration have not necessarily bounded mass, but optimal configuration do have". This is due to the fact that, if f^+ and f^- are probabilities, then $m \leq 1$ on optimal configurations (and no cycle are possible, by the way). Notice that this statement does not depend on $m \mapsto m^\alpha$ being concave, but simply increasing in m. The same kind of behavior is likely to be true on the approximating problems, but after months of work a proof of this fact did not appear.

Hence we conclude with two points: we ask a question and we give a final suggestion for getting a useful Γ-convergence result.

The suggestion is quite *naïve* : just take a sufficiently large number K so that every minimizer u of the limit problem satisfies $|u|(\overline{\Omega}) \leq K$ and then restrict the analysis to the compact subset $Y_K(\Omega) := \{(u,\mu,\nu) \in Y(\Omega) \ : \ |u|(\overline{\Omega}) \leq K\}$. Unfortunately, there exists no continuous functional Ψ on $Y(\Omega)$ which is coercive on the mass of u. If it was the case, one could have added a term f the form $F(u) = (\Psi(u) - K)^2_+$ to the limit and approximating functionals. It is on the contrary a reasonable question, simpler than the one we are going to detail below, whether there exists a continuous functional Ψ such that, for the minimizers u_ε, the condition $\Psi(u_\varepsilon) \leq C$ implies a bound on the mass (taking advantage of possible extra estimates on u_ε).

Thus, the problems

$$\min\left\{M^\alpha_\varepsilon(u) + G_1((\mu,\nu)(f^+,f^-))\ (u,\mu,\nu) \in Y_K(\Omega)\right\}$$

approximate (with $\Gamma-$convergence) the limit problem given by $\min_{Y_K(\Omega)}\ M^\alpha(u) + G_1((\mu,\nu),(f^+,f^-))$, which is equivalent to $\min\ M^\alpha(u) \ : \ \nabla\cdot u = f$.

Here as well a proper use of the estimate given in the last improvement in Section 5 are important. Actually, one needs to produce a recovery sequence u_ε for the $\Gamma - \limsup$ construction which does not violate the mass constraint given by K. This may be obtained in the following way: instead of considering

as a dense subset the set of all vector measures u corresponding to finite graphs, take the set of all such measures u satisfying $|u| < K$. This set is still dense in energy. Afterwards, one notices that the construction which was provided satisfied $|u_\varepsilon|(\overline{\Omega}) \leq |u|(\overline{\Omega}) + C\varepsilon^{\alpha+1}$ (where the constant C depends actually on u, but not on ε, since it depends on the number of nodes and of branches at every node). This means that, if u satisfies the strict inequality $|u| < K$, this will be true for u_ε as well for ε sufficiently small.

And, as obvious, the question is

Open question 2. *Prove a bound on the L^1 norm of the minimizers u_ε (or on suitable minimizers u_ε, if needed). If possible, prove it for minimizers u_ε which minimize M_ε^α under a divergence constraint $\nabla \cdot u_\varepsilon = f_\varepsilon$ (so that it will be true even if we add penalizations on the divergence). Is is possible to do it simply by writing and exploiting optimality conditions under the form of elliptic (second or fourth order) PDEs? do we need to pass through more geometrical tools such has some surrogate of the no-cycle condition? Is this bound equivalent to a bound on some m^α (being m the flux thrugh some segments or curves, as in what we proved in Step 1, where $m(x)$ is the flux through R_x) coupled with some bound on the fluxes m?*

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 very one-dimensional).

Acknowledgements

Jean-Michel Morel is warmly thanked for the impulse he gave to this subject at his very starting point in 2006 and for the constant help and attention towards both the topic and the author since then.

Giovanni Alberti and Giuseppe Buttazzo contributed to the current version of the paper by means of the interesting remarks and clarifying comments they proposed when this work was first presented in a seminar in Pisa.

The author also took advantage of the support of the ANR project OTARIE (Optimal transport : Theory and Applications to cosmological Reconstruction and Image processing) and of the French-Italian Galilée Project "Modèles de réseaux de transport: phénomènes de congestion et de branchement"

References

[1] L. Ambrosio, C. De Lellis, C. Mantegazza, Line energies for gradient vector fields in the plane, *Calc. Var. Partial Differential Equations* (9) no. 4, 1999, 327–255.

[2] L. Ambrosio and V. M. Tortorelli, Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.* (43), no. 8, 1990, 999–1036.

[3] L. Ambrosio and V. M. Tortorelli, On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital. B (7)*, 6, no. 1, 1992, 105–123.

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

[5] P. Aviles and Y. Giga, Singularities and rank one properties of Hessian measures. *Duke Math. J.* 58 (1989), no. 2, 441–467.

[6] M. Bernot, V. Caselles and J.-M. Morel, Traffic plans, *Publicacions Matemàtiques* (49), no. 2, 417–451, 2005.

[7] M. Bernot, V. Caselles and J.-M. Morel, *Optimal transportation networks, Models and Theory*, Lecture Notes in Mathematics, Springer, Vol. 1955 (2008).

[8] F. Bethuel, H. Brezis and F. Hélein, Asymptotics for the minimization of a Ginzburg-Landau functional, *Calc. Var. and Par. Diff. Eq.*, (1), no. 2, 1993, 123–148

[9] F. Bethuel, H. Brezis and F. Hélein, *Ginzburg-Landau Vortices*, Birkhäuser, 1994.

[10] G. Bouchitté, C. Jimenez, M. Rajesh: Asymptotique d'un problème de positionnement optimal, *C. R. Acad. Sci. Paris* Ser. I, (335), 2002, 1–6.

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

[12] G. Dal Maso: *An Introduction to* $\Gamma-$*convergence*. Birkhauser, Basel, 1992.

[13] E. De Giorgi and T. Franzoni, Su un tipo di convergenza variazionale, *Atti Acc. Naz. Lincei Rend.*, (58) no. 8, 1975, pp. 842-850.

[14] H. Federer: *Geometric measure theory*. Springer Verlag, New York, 1969.

[15] D. Gilbarg and N. Trudinger *Elliptic Partial Differential Equations of Second Order*, Springer Verlag, Berlin, 1977.

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

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

[18] L. Modica and S. Mortola, Un esempio di Γ-convergenza. (Italian) *Boll. Un. Mat. Ital. B* (14), no. 1, 285–299, 1977.

[19] J.-M. Morel and F. Santambrogio, Comparison of distances between measures, *Appl. Math. Lett.*, (20), no. 4, 2007, 427–432.

[20] S. Mosconi and P. Tilli: Γ−Convergence for the Irrigation Problem, *J. of Conv. Anal.* (12), no.1, 2005, 145–158.

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

[22] F. Santambrogio, Optimal Channel Networks, Landscape Function and Branched Transport, *Interfaces and Free Boundaries*, (9), 2007, 149–169.

[23] B. White, Rectifiability of flat chains, *Ann. of Math.* (150), no. 1, 165–184.

[24] Q. Xia, Optimal Paths related to Transport Problems. *Comm. Cont. Math.* (5), no. 2, 251–279, 2003.

[25] Q. Xia, Interior regularity of optimal transport paths. *Calc. Var. and and Par. Diff. Eq.* (20), no. 3, 2004, 283–299.