Uniform estimates for a Modica–Mortola type approximation of branched transportation

Antonin Monteil

ESAIM: COCV 23 (2017) 309–335
DOI: 10.1051/cocv/2015049

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

UNIFORM ESTIMATES FOR A MODICA–MORTOLA TYPE APPROXIMATION OF BRANCHED TRANSPORTATION

Antonin Monteil[1]

Abstract. Models for branched networks are often expressed as the minimization of an energy M^α over vector measures concentrated on 1-dimensional rectifiable sets with a divergence constraint. We study a Modica–Mortola type approximation M^α_ε, introduced by Edouard Oudet and Filippo Santambrogio, which is defined over H^1 vector measures. These energies induce some pseudo-distances between L^2 functions obtained through the minimization problem $\min\{M^\alpha_\varepsilon(u)\colon \nabla\cdot u = f^+ - f^-\}$. We prove some uniform estimates on these pseudo-distances which allow us to establish a Γ-convergence result for these energies with a divergence constraint.

Mathematics Subject Classification. 49J45, 90B06, 90B18.

Received March 11, 2015. Accepted September 30, 2015.

1. Introduction

Branched transportation is a classical problem in optimization: it is a variant of the Monge–Kantorovich optimal transportation theory in which the transport cost for a mass m per unit of length is not linear anymore but sub-additive. More precisely, the cost to transport a mass m on a length l is considered to be proportional to $m^\alpha l$ for some $\alpha \in]0,1[$. As a result, it is more efficient to transport two masses m_1 and m_2 together instead of transporting them separately. For this reason, an optimal pattern for this problem has a "graph structure" with branching points. Contrary to what happens in the Monge–Kantorovich model, in the setting of branched transportation, an optimal structure cannot be described only using a transport plan, giving the correspondence between origins and destinations, but we need a model which encodes all the trajectories of mass particles.

Branched transportation theory is motivated by many structures that can be found in the nature: vessels, trees, river basins... Similarly, as a consequence of the economy of scale, large roads are proportionally cheaper than large ones and it follows that the road and train networks also present this structure. Surprisingly the theory has also had theoretical applications: recently, it has been used by Bethuel in [4] so as to study the density of smooth maps in Sobolev spaces between manifolds.

Branched transportation theory was first introduced in the discrete framework by Gilbert in [14] as a generalization of the Steiner problem. In this case an admissible structure is a weighted graph composed of oriented edges of length l_i on which some mass m_i is flowing. The cost associated to it is then $\sum_i l_i m_i^\alpha$ and it has to be minimized over all graphs which transport some given atomic measure to another one. More recently, the branched transportation problem was generalized to the continuous framework by Xia in [23] by means

Keywords and phrases. Branched transportation networks, Γ-convergence, phase field models.

[1] Laboratoire de Mathématiques d'Orsay, Université Paris-Sud 11, Bât. 425, 91405 Orsay, France. antonin.monteil@gmail.com

of a relaxation of the discrete energy (see also [24]). Then, many other models and generalizations have been introduced (see [15] for a Lagrangian formulation, see also [1–3] for different generalizations and regularity properties.). In this paper, we will concentrate on the model with a divergence constraint, due to Xia. However, this is not restrictive since all these models have been proved to be equivalent (see [3,19]).

In this model, a transport path is represented as a vector measure u on some open set $\Omega \subset \mathbb{R}^d$ such that $\nabla \cdot u = \mu^+ - \mu^-$ for two probability measures μ^+ and μ^-. Then the energy of u is defined as $M^\alpha(u) = \int_M \theta^\alpha d\mathcal{H}^1$ if u is a vector measure concentrated on a rectifiable 1-dimensional set M on which u has multiplicity θ w.r.t. the Hausdorff measure (see [3] for more details). In this framework, u must be considered as the momentum (the mass θ times the velocity) of a particle at some point. Then $(\nabla \cdot u)(x)$ represents the difference between incoming and outcoming mass at each point x.

In this paper, we are interested in some approximation of branched transportation proposed by Oudet and Santambrogio few years ago in [18] and which has interesting numerical applications. This model was inspired by the well known scalar phase transition model proposed by Modica and Mortola in [16]. Given $u \in H^1(\Omega, \mathbb{R}^d)$ for some bounded open subset $\Omega \subset \mathbb{R}^d$, Oudet and Santambrogio introduced the following energy:

$$M_\varepsilon^\alpha(u) = \varepsilon^{-\gamma_1} \int_\Omega |u|^\beta + \varepsilon^{\gamma_2} \int_\Omega |\nabla u|^2,$$

where $\beta \in (0,1)$ and $\gamma_1, \gamma_2 > 0$ are some exponents depending on α (see (2.4)). If u does not belong to the set $H^1(\Omega)$, the value of M_ε^α is taken as $+\infty$.

We recall the heuristic which shows why M_ε^α is an approximation of M^α (see [18]): assume that μ^- (resp. μ^+) is a point source at S_1 (resp. S_2) with mass m. Then, it is clear that the optimal path for M^α between these two measures is the oriented edge $S = (S_1, S_2)$ of length l with a mass m flowing on it. We would like to approximate this structure, seen as a vector measure u concentrated on S, by some H^1 vector fields v which are more or less optimal for M_ε^α. What we expect is that v looks like a convolution of u with a kernel ρ depending on ε and m: $v = u * \rho_R$ for some $R = R(\varepsilon, m)$, where

$$\rho_R(x) = R^{-d} \rho\left(R^{-1} x\right) \tag{1.1}$$

for some fixed smooth and compactly supported radial kernel $\rho \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Then the support of v is like a strip of width R around S so that $|v|$ is of the order of m/R^{d-1} and $|\nabla v|$ is of the order of m/R^d. This gives an estimate of $M_\varepsilon^\alpha(v)$ like

$$M_\varepsilon^\alpha(v) \simeq \varepsilon^{-\gamma_1} R^{d-1} \left(m/R^{d-1}\right)^\beta l + \varepsilon^{\gamma_2} R^{d-1} \left(m/R^d\right)^2 l. \tag{1.2}$$

With our choice for the exponents γ_1, γ_2 and β, the optimal choice for R is

$$R = \varepsilon^\gamma m^{\frac{1-\gamma}{d-1}}, \tag{1.3}$$

where

$$\gamma = \frac{2}{2d - \beta(d-1)} = \frac{\gamma_2}{d+1}. \tag{1.4}$$

This finally leads to $M_\varepsilon^\alpha(v) \simeq m^\alpha$ as expected.

It was proved in [18] that, at least in two dimensions, the energy sequence $(M_\varepsilon^\alpha)_{\varepsilon>0}$ Γ-converges to the branched transportation functional $c_0 M^\alpha$ for some constant c_0 and for some suitable topology (see Thm. 2.1, p. 313). This result has been interestingly applied to produce a numerical method. However, rather than a Γ-convergence result on M_ε^α we would need to deal with the functionals $\overline{M}_\varepsilon^\alpha$, obtained by adding a divergence constraint: it should be shown that $\overline{M}_\varepsilon^\alpha(u) := M_\varepsilon^\alpha(u) + I_{\nabla \cdot u = f_\varepsilon}$ Γ-converges to $c_0 \overline{M}^\alpha(u) := c_0 M_\varepsilon^\alpha(u) + I_{\nabla \cdot u = \mu^+ - \mu^-}$, where $f_\varepsilon \in L^2$ is some suitable approximation of $\mu^+ - \mu^-$ and $I_A(u)$ is the indicator function in the sense of convex analysis that is 0 whenever the condition is satisfied and $+\infty$ otherwise. Even if this property was not

proved in [18], the effectiveness of the numerical simulations made the authors think that it actually holds true. Note that an alternative using a penalization term was proposed in [20] to overcome this difficulty.

In Section 2 we recall Xia's formulation of branched transportation and its approximation M_ε^α introduced by Oudet and Santambrogio. The longest part of this paper, Section 3, is devoted to a local estimate which gives a bound on the minimum value $d_\varepsilon^\alpha(f^+, f^-) := \min\{M_\varepsilon^\alpha(u) \ : \ \nabla\cdot u = f\}$ depending on $\|f\|_{L^1}$, $\|f\|_{L^2}$ and $\mathrm{diam}(\Omega)$ (see Thm. 3.2, p. 314). In Section 4, we deduce a comparison between d_ε^α and the Wasserstein distance with an "error term" involving the L^2 norm of $f^+ - f^-$. As an application of this inequality, in the last section, we will prove the following Γ-convergence result which was lacking in [18]

Theorem 1.1. *Let $(f_\varepsilon)_{\varepsilon>0} \subset L^2(\Omega)$ be a sequence weakly converging to μ as measures when $\varepsilon \to 0$. Assume that the sequence $(f_\varepsilon)_{\varepsilon>0}$ satisfies*

$$\int_\Omega f_\varepsilon(x)\,\mathrm{d}x = 0 \quad \textit{and} \quad \varepsilon^{\gamma_2}\|f_\varepsilon\|_{L^2}^2 \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

There exists a constant $c_0 > 0$ such that the functional sequence $(\overline{M}_\varepsilon^\alpha)_{\varepsilon>0}$ Γ-converges to $c_0\overline{M}^\alpha$ as $\varepsilon \to 0$. Moreover c_0 is the minimum value for the minimizing problem (5.2).

This answers the open question 1 in [18,20] and validates their numerical method.

2. Mathematical setting

2.1. The branched transportation energy

In all what follows, we will use the model proposed by Xia (see [23,24]):

Let $d \geq 1$ be an integer and Ω be some open and bounded subset of $\mathbb{R}^d$. Let us denote by $\mathcal{M}_{div}(\Omega)$ the set of finite vector measures on $\overline{\Omega}$ such that their divergence is also a finite measure:

$$\mathcal{M}_{div}(\Omega) := \left\{u \text{ measure on } \overline{\Omega} \text{ valued in } \mathbb{R}^d \ : \ \|u\|_{\mathcal{M}_{div}(\Omega)} < +\infty\right\},$$

where $\|u\|_{\mathcal{M}_{div}(\Omega)} := |u|(\overline{\Omega}) + |\nabla\cdot u|(\overline{\Omega})$ with

$$|u|(\overline{\Omega}) := \sup\left\{\int_{\overline{\Omega}} \psi\cdot \mathrm{d}u \ : \ \psi \in \mathcal{C}(\overline{\Omega}, \mathbb{R}^d),\ \|\psi\|_{L^\infty} \leq 1\right\}$$

and, similarly,

$$|\nabla\cdot u|(\overline{\Omega}) := \sup\left\{\int_{\overline{\Omega}} \nabla\varphi\cdot \mathrm{d}u \ : \ \varphi \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}),\ \|\varphi\|_{L^\infty} \leq 1\right\}.$$

In all what follows, $\nabla\cdot u$ has to be thought in the weak sense, *i.e.* $\int \varphi\nabla\cdot u = -\int \nabla\varphi\cdot \mathrm{d}u$ for all $\varphi \in \mathcal{C}^1(\overline{\Omega})$. Since we do not ask φ to vanish at the boundary, $\nabla\cdot u$ may contain possible parts on $\partial\Omega$ which are equal to $u\cdot n$ when u is smooth, where n is the external unit normal vector to $\partial\Omega$. In other words, $\nabla\cdot u$ is the weak divergence of $u\mathbf{1}_\Omega$ in $\mathbb{R}^d$, where $\mathbf{1}_\Omega$ is the classical indicator function of Ω, equal to 1 on Ω and 0 elsewhere. From now on, the notation $\mathbf{1}_X$ for the classical indicator function of a set X and I_X for the indicator function in the sense of convex analysis (equal to 1 inside and $+\infty$ outside) will be used. $\mathcal{M}_{div}(\Omega)$ is endowed with the topology of weak convergence on u and on its divergence: *i.e.* $u_n \overset{\mathcal{M}_{div}(\Omega)}{\longrightarrow} u$ if $u_n \rightharpoonup u$ and $\nabla\cdot u_n \rightharpoonup \nabla\cdot u$ weakly as measures.

Given $0 < \alpha < 1$, the energy of branched transportation can be represented as follows for measures $u \in \mathcal{M}_{div}(\Omega)$:

$$M^\alpha(u) = \begin{cases} \int_M \theta^\alpha\,\mathrm{d}\mathcal{H}^1 & \text{if } u \text{ can be written as } u = U(M,\theta,\xi), \\ +\infty & \text{otherwise,} \end{cases} \tag{2.1}$$

where $U(M,\theta,\xi)$ is the rectifiable vector measure $u=\theta\xi\cdot\mathcal{H}^1_{|M}$ with density $\theta\xi$ with respect to the $\mathcal{H}^1$-Hausdorff measure on the rectifiable set M. The real multiplicity is a measurable function $\theta : M\to\mathbb{R}^+$ and the orientation $\xi : M\to S^{d-1}\subset\mathbb{R}^d$ is such that $\xi(x)$ is tangential to M for $\mathcal{H}^1$-a.e. $x\in M$.

Given two probability measures μ^+ and μ^- on $\overline{\Omega}$, the problem of branched transportation consists in minimizing M^α under the constraint $\nabla\cdot u=\mu^+-\mu^-$:

$$\inf\left\{M^\alpha(u)\ :\ u\in\mathcal{M}_{div}(\overline{\Omega})\quad\text{and}\quad\nabla\cdot u=\mu^+-\mu^-\right\}. \tag{2.2}$$

Note that, if $\mu^\pm(\partial\Omega)=0$, the divergence constraint implies a Neumann condition on u: $u\cdot n=0$ on $\partial\Omega$.

2.2. Functionals M^α_ε

For the minimum value in (2.2) to be finite whatever μ^+ and μ^- in the set of probability measures, we will require α to be sufficiently close to 1. More precisely, we make the following assumption:

$$1-\frac{1}{d}<\alpha<1. \tag{2.3}$$

Xia has shown in [23] that, under this assumption, there exists at least one vector measure $u\in\mathcal{M}_{div}(\Omega)$ such that $M^\alpha(u)<+\infty$.

We are interested in the following approximation of M^α which was introduced in [18]: for all $u\in\mathcal{M}_{div}(\Omega)$ and for all open subset $\omega\subset\Omega$,

$$M^\alpha_\varepsilon(u,\omega):=\begin{cases}\varepsilon^{-\gamma_1}\displaystyle\int_\omega|u(x)|^\beta\,\mathrm{d}x+\varepsilon^{\gamma_2}\int_\omega|\nabla u(x)|^2\,\mathrm{d}x & \text{if } u\in H^1(\omega)\\ +\infty & \text{otherwise,}\end{cases} \tag{2.4}$$

where β, γ_1 and γ_2 are three exponents depending on α and d through:

$$\beta=\frac{2-2d+2\alpha d}{3-d+\alpha(d-1)}$$

and

$$\gamma_1=(d-1)(1-\alpha)\quad\text{and}\quad\gamma_2=3-d+\alpha(d-1).$$

Note that inequality $1-1/d<\alpha<1$ implies that $0<\beta<1$. When $\omega=\Omega$, we simply write

$$M^\alpha_\varepsilon(u,\Omega)=:M^\alpha_\varepsilon(u).$$

We point out the 2-dimensional case where M^α_ε rewrites as

$$M^\alpha_\varepsilon(u)=\varepsilon^{\alpha-1}\int_\Omega|u(x)|^\beta\,\mathrm{d}x+\varepsilon^{\alpha+1}\int_\Omega|\nabla u(x)|^2\,\mathrm{d}x, \tag{2.5}$$

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

Given two densities $f^+,f^-\in L^2_+(\Omega):=\{f\in L^2(\Omega)\ :\ f\geq 0\}$ such that $\int f^+=\int f^-$, we are interested in minimizing $M^\alpha_\varepsilon(u)$ under the constraint $\nabla\cdot u=f^+-f^-$:

$$\inf\left\{M^\alpha_\varepsilon(u)\ :\ u\in H^1(\Omega)\quad\text{and}\quad\nabla\cdot u=f^+-f^-\right\}. \tag{2.6}$$

The classical theory of calculus of variation shows that this infimum is actually a minimum. A natural question that arises is then to understand the limit behavior for minimizers of these problems when ε goes to 0. A classical tool to study this kind of problems is the theory of Γ-convergence which was introduced by De Giorgi in [12]. For the definition and main properties of Γ-convergence, we refer to [8,11]. In particular, if M^α_ε Γ-converges to some energy functional M^α_0 and if (u_ε) is a sequence of minimizers for M^α_ε admitting a subsequence converging to u, then, u is a minimizer for M^α_0. By construction of M^α_ε, we expect that, up to a subsequence, M^α_ε Γ-converges to c_0M^α. In the two dimensional case, we have the following Γ-convergence theorem proved in [18].

Theorem 2.1. *Assume that $d = 2$ and $\alpha \in (1/2, 1)$. Then, there exists a constant $c > 0$ such that $(M_\varepsilon^\alpha)_{\varepsilon>0}$ Γ-converges to cM^α in $\mathcal{M}_{div}(\Omega)$ when ε goes to 0.*

Nevertheless, this does not imply the Γ-convergence of $M_\varepsilon^\alpha(u) + I_{\nabla\cdot u = f^+ - f^-}$ to $M_\varepsilon^\alpha(u) + I_{\nabla\cdot u = f^+ - f^-}$. Indeed, the Γ-convergence is stable under the addition of continuous functionals but not l.s.c. functionals. Consequently, we cannot deduce, from this theorem, the behavior of minimizers for (2.6). For instance, it is not clear that there exists a recovery sequence (u_ε), *i.e.* u_ε converges to u in $\mathcal{M}_{div}(\Omega)$ and $M_\varepsilon^\alpha(u_\varepsilon)$ converges to $M^\alpha(u)$ as $\varepsilon \to 0$, with prescribed divergence $\nabla \cdot u_\varepsilon = f^+ - f^-$. To this aim, we require some estimates on these energies and this is the purpose of this paper.

2.3. Distance of branched transportation

We remind our hypothesis $1 - 1/d < \alpha < 1$. In [23], Xia has remarked that, as in optimal transportation theory, M^α induces a distance d^α on the space $\mathcal{P}(\overline{\Omega})$ of probability measures on $\overline{\Omega}$:

$$d^\alpha(\mu^+, \mu^-) = \inf\left\{M^\alpha(u) \ : \ u \in \mathcal{M}_{div}(\Omega) \quad \text{such that} \quad \nabla \cdot u = \mu^+ - \mu^-\right\},$$

for all $\mu^+, \mu^- \in \mathcal{P}(\overline{\Omega})$. Thanks to our assumption $\alpha > 1 - 1/d$, d^α is finite for all $\mu^\pm \in \mathcal{P}(\overline{\Omega})$ and it induces a distance on the set $\mathcal{P}(\overline{\Omega})$ which metrizes the topology of weak convergence of measures. Actually, d^α has a stronger property which is a comparison with the Wasserstein distance.

Proposition 2.2. *Let μ^+ and μ^- be two probability measures on $\overline{\Omega}$. We denote by W_p the Wasserstein distance associated to the cost $(x, y) \to |x - y|^p$ for $p \geq 1$. Then, one has*

$$W_{1/\alpha}(\mu^+, \mu^-) \leq d^\alpha(\mu^+, \mu^-) \leq C\, W_1(\mu^+, \mu^-)^{1-d(1-\alpha)},$$

for a constant $C > 0$ only depending on d, α and the diameter of Ω.

We refer to [17] for a proof of this property (see also [3,9] for an alternative proof) and [21,22] for the definition and main properties of the Wasserstein distance. In the same way, we define d_ε^α as follows:

$$d_\varepsilon^\alpha(f^+, f^-) = \inf\left\{M_\varepsilon^\alpha(u) \ : \ u \in H^1(\mathbb{R}^d) \quad \text{such that} \quad \nabla \cdot u = f^+ - f^-\right\}, \tag{2.7}$$

where $f^+, f^- \in L^2_+(\Omega)$ satisfy $\int_\Omega f^+ = \int_\Omega f^-$. Although d^α is a distance, it is not the case for d_ε^α which does not satisfy the triangular inequality. Actually, because of the second term involving $|\nabla u|^2$, M_ε^α is not subadditive. However, for $u_1, \ldots, u_n$ in $\mathcal{M}_{div}(\Omega)$, the inequality $|\nabla u_1 + \ldots + \nabla u_n|^2 \leq n\{|\nabla u_1|^2 + \ldots + |\nabla u_n|^2\}$ implies

$$M_\varepsilon^\alpha\left(\sum_{i=1}^n u_i\right) \leq n \sum_{i=1}^n M_\varepsilon^\alpha(u_i).$$

In particular, d_ε^α is a pseudo-distance in the sense that the three properties in the following proposition are satisfied.

Proposition 2.3. *Let f^+, f^- and $f_1, \ldots, f_n$ be L^2 densities,* i.e. *L^2 non negative functions whose integral is equal to 1. Then one has*

1. *$d_\varepsilon^\alpha(f^+, f^-) = 0$ implies $f^+ = f^-$,*
2. *$d_\varepsilon^\alpha(f^+, f^-) = d_\varepsilon^\alpha(f^-, f^+)$,*
3. *$d_\varepsilon^\alpha(f_0, f_n) \leq n\left[d_\varepsilon^\alpha(f_0, f_1) + d_\varepsilon^\alpha(f_1, f_2) + \cdots + d_\varepsilon^\alpha(f_{n-1}, f_n)\right]$.*

3. Local estimate

We remind our assumption (2.3) which insures that $d^\alpha(\mu^+,\mu^-)$ is always finite. Our goal is to prove that d^α_ε enjoys a property similar to the following one.

Proposition 3.1. *Let $Q_0=(0,L)^d\subset\mathbb{R}^d$ be a cube of side length $L>0$. There exists some constant $C>0$ only depending on d and α such that for all non negative Borel finite measure μ of total mass $\theta>0$,*

$$d^\alpha(\mu,\theta\delta_0)\leq C\ \theta^\alpha L,$$

where δ_0 is the Dirac measure at the point c_{Q_0}, the center of Q_0.

Since $d^\alpha_\varepsilon(f^+,f^-)$ is only defined on L^2 functions $f^\pm$, to do so, we first have to replace $\theta\delta_0$ by some kernel which concentrates at the origin when ε goes to 0. Let $\rho\in\mathcal{C}^1_c(B,\mathbb{R}^+)$ be a radial non negative function such that $\int_{\mathbb{R}^d}\rho=1$, where $B\subset\mathbb{R}^d$ is the unit ball centered at the origin, and define $\rho_{\theta,\varepsilon}:=\rho_R$ as in (1.1), where

$$R=:R_{\theta,\varepsilon}=\varepsilon^\gamma\theta^{\frac{1-\gamma}{d-1}}.$$

Here, we recall that R and $\gamma=\frac{\gamma_2}{d+1}$ were introduced in (1.4). Let Q be a cube in $\mathbb{R}^d$ centered at some point $c_Q\in\mathbb{R}^d$ and $f\in L^2_+(Q)$ be a density such that $\int_Q f=:\theta_Q$. Then, we will denote by ρ_Q the kernel $\theta\rho_{\theta,\varepsilon}$ refocused at c_Q with a small abuse of notation (indeed, ρ_Q also depends on f):

$$\rho_Q(x)=\theta_Q\rho_{\theta_Q,\varepsilon}(x-c_Q).$$

The main result of this section is the following theorem

Theorem 3.2 (Local estimate)**.** *Let us set $Q_0=(0,L)^d$ for some $L>0$. There exists $C>0$ only depending on α, ρ and d such that for all $f\in L^2_+(Q_0)$ with $\int_{Q_0}f=\theta$, we have*

- *If $\operatorname{supp}\rho_{Q_0}\subset Q_0$ then, there exists $u\in H^1_0(Q_0)$ such that $\nabla\cdot u=f-\rho_{Q_0}$ and*

$$d^\alpha_\varepsilon(f,\rho_{Q_0})\leq M^\alpha_\varepsilon(u)\leq C\big\{\theta^\alpha L+\varepsilon^{\gamma_2}\|f\|^2_{L^2}\big\}\quad\text{and}\quad\|u\|_{L^1}\leq C\,L\,\theta.$$

- *Otherwise, there exists $u\in H^1_0(\widetilde{Q}_0)$ such that*

$$d^\alpha_\varepsilon(f,\rho_{Q_0})\leq M^\alpha_\varepsilon(u)\leq C\varepsilon^{\gamma_2}\|f\|^2_{L^2}\quad\text{and}\quad\|u\|_{L^1}\leq C\,L\,\theta,$$

where $\widetilde{Q}_0=2\operatorname{supp}\rho_{Q_0}:=B(c_{Q_0},2R_{\theta,\varepsilon})$.

Remark 3.3. The Dirichlet term, $\varepsilon^{\gamma_2}\|f\|^2_{L^2}$, in the estimates above is easily understandable. Indeed, if ε is very large so that one can get rid of the first term in the energy M^α_ε, then, one can use a classical Dirichlet type estimate, that is Theorem 3.4 below. On the contrary, for ε very small, the Γ-limit result on M^α_ε tells us that these energies are close to M^α so that it is natural to hope a similar estimate as the one for M^α: that is to say an estimate from above by $\theta^\alpha L$ (see [3]).

The main difficulty to prove Theorem 3.2 is the non subadditivity of the pseudo-distances $\lambda^\alpha_\varepsilon$. Indeed, our proof is based on a dyadic construction used by Xia in [23] to prove Proposition 3.1 (see also [3]). This gives a singular vector measure u which is concentrated on a graph. Since M^α_ε contains a term involving the L^2 norm of ∇u, we have to regularize u by taking a convolution with the kernel $\rho_{\theta,\varepsilon}$ on each branch of the graph (θ being the mass traveling on it). Unfortunately in this way, two different branches are no longer disjoints.

It is useful to see that we have a first candidate for the minimization problem (2.7). This candidate is of the form $u=\nabla\phi$, where ϕ is the solution of the Dirichlet problem

$$\begin{cases}\Delta\phi=f^+-f^-\ \text{in}\quad Q,\\ \phi\quad=0\qquad\quad\ \text{on}\quad\partial Q.\end{cases}\tag{3.1}$$

Then $u = \nabla\phi$ satisfies $\nabla \cdot u = f^+ - f^-$ in Q and $u(x) \in \mathbb{R}n$ a.e. on ∂Q where n stands for the external unit normal vector to ∂Q. Alternatively, one could consider Neumann homogeneous boundary conditions for ϕ rather than Dirichlet boundary conditions. Then, one would obtain $u(x) \cdot n = 0$ a.e. on ∂Q. Theorem 3.4 below gives a better result in the sense that the candidate u vanishes at the boundary:

Theorem 3.4. *Let $Q_0 = (0, L)^d$ be a cube of side length $L > 0$. There exists $C > 0$ only depending on d such that for all $f \in L^2_0(Q_0)$, there exists $u \in H^1_0(Q_0, \mathbb{R}^2)$ solving $\nabla \cdot u = f$ and satisfying $\|u\|_{L^1(Q_0)} \leq CL\,\|f\|_{L^1(Q_0)}$ together with*

$$\|u\|_{H^1_0(Q_0)} := \left(\int_{Q_0} |\nabla u|^2\right)^{1/2} \leq C\,\|f\|_{L^2(Q_0)},$$

where $L^2_0(Q_0) = \left\{ f \in L^2(Q_0) \ : \ \int_{Q_0} f(x)\,\mathrm{d}x = 0 \right\}$.

For a proof of this theorem, see, for instance, Theorem 2 in [7]: the only difference with Theorem 3.4 is that we add the estimate $\|u\|_{L^1(Q_0)} \leq CL\,\|f\|_{L^1(Q_0)}$ which can be easily obtained following the proof of Bourgain and Brezis. The corresponding property formulated on a Lipschitz bounded connected domain Ω is also true (see Thm. 2' in [7]) except that the constant C could depend on Ω in this case.

Of course, this candidate is usually not optimal for (2.7) and this does not allow for a good estimate because of the first term in the definition of M^α_ε. For this reason, we have to use the dyadic construction of Xia up to a certain level ("diffusion level") from which we simply use Theorem 3.4.

3.1. Dyadic decomposition of Q_0 and "diffusion level" associated to f

Let us call "dyadic descent" of $Q_0 = (0, L)^d$ the set $\mathcal{Q} = \bigcup_{j\geq 0} \mathcal{Q}_j$, where $\mathcal{Q}_j$ is the jth "dyadic generation":

$$\mathcal{Q}_j = \left\{ (x_1, \ldots, x_d) + 2^{-j}Q_0 \ : \ x_i \in \{k2^{-j}L \ : \ 0 \leq k \leq 2^j - 1\} \quad \text{for} \quad i = 1, \ldots, d \right\}.$$

Note that $\mathrm{Card}(\mathcal{Q}_j) = 2^{jd}$. For each $Q \in \mathcal{Q}$, let us define

- $\mathcal{D}(Q)$: the descent of Q, the family of all dyadic cubes contained in Q.
- $\mathcal{A}(Q)$: the ancestry of Q, the family of all dyadic cubes containing Q.
- $\mathcal{C}(Q)$: the family of children of Q composed of the 2^d biggest dyadic cubes strictly included in Q.
- $F(Q)$: the father of Q, the smallest dyadic cube strictly containing Q.

We now remind the dyadic construction described in [23] which irrigates f from a point source. We first introduce some notations: fix a function $f \in L^2_+(Q_0)$ with integral θ and let $Q \in \mathcal{Q}$ be a dyadic cube centered at $c_Q \in \mathbb{R}^d$. Then we introduce θ_Q the mass associated to the cube Q as

$$\theta_Q = \int_Q f.$$

If $\theta_Q \neq 0$, we also define the kernel associated to Q through

$$\overline{\rho}_Q(x) = \rho_R(x), \tag{3.2}$$

where ρ_R is defined in (1.1) for

$$R = R_Q := \varepsilon^\gamma \theta_Q^{\frac{1-\gamma}{d-1}}, \quad \gamma = \frac{\gamma_2}{d+1}.$$

Here γ was defined in Define also the weighted recentered kernel by

$$\rho_Q(x) = \theta_Q \overline{\rho}_Q(x - c_Q) \tag{3.3}$$

if $\theta_Q \neq 0$ and $\rho_Q(x) = 0$ otherwise. Lastly, we introduce the point source associated to the cube Q as

$$\mathcal{S}_Q := \theta_Q \times \text{Dirac measure at point } c_Q.$$

We are now able to construct a vector measure X such that $M^\alpha(X) < +\infty$. First define the measures X_Q as below:

$$X_Q = \sum_{Q' \in \mathcal{C}(Q)} \theta_{Q'}\, n_{Q'}\, \mathcal{H}^1_{\lfloor [c_Q, c_{Q'}] \rfloor}, \tag{3.4}$$

where $n_{Q'} = \dfrac{c_{Q'} - c_Q}{\|c_{Q'} - c_Q\|}$. Then, we have

$$\nabla \cdot X_Q = \sum_{Q' \in \mathcal{C}(Q)} \mathcal{S}_{Q'} - \mathcal{S}_Q$$

and the energy estimate

$$M^\alpha(X_Q) \leq 2^{d-2} \theta_Q^\alpha \operatorname{diam}(Q),$$

where $\operatorname{diam}(Q)$ stands for the diameter of Q. Finally, the measure $X = \sum_{Q \in \mathcal{Q}} X_Q$ solves $\nabla \cdot X = f - \mathcal{S}_{Q_0}$ and satisfies

$$M^\alpha(X) \leq C \theta^\alpha \operatorname{diam}(Q_0).$$

Indeed, it is enough to apply the following lemma with $\lambda = \alpha$:

Lemma 3.5. *Let $Q \in \mathcal{Q}$ and $\lambda \in]1 - 1/d, 1]$. There exists a constant $C = C(\lambda, d)$ such that*

$$\sum_{Q' \in \mathcal{D}(Q)} \theta_{Q'}^\lambda \operatorname{diam}(Q') \leq C \theta_Q^\lambda \operatorname{diam}(Q).$$

Proof. Let $j_0 \geq 0$ be such that $Q \in \mathcal{Q}_{j_0}$. The definition of $\mathcal{D}(Q)$, the Jensen inequality and the fact that $d - 1 - \lambda d < 0$ give

$$\begin{aligned}
\sum_{Q' \in \mathcal{D}(Q)} \theta_{Q'}^\lambda \operatorname{diam}(Q') &= \sum_{j \geq 0} 2^{-j} \operatorname{diam}(Q) \sum_{Q' \in \mathcal{D}(Q) \cap \mathcal{Q}_{j_0 + j}} \theta_{Q'}^\lambda \\
&\leq \operatorname{diam}(Q) \sum_{j \geq 0} 2^{-j} 2^{jd} \left(2^{-jd} \sum_{Q' \in \mathcal{D}(Q) \cap \mathcal{Q}_{j_0 + j}} \theta_{Q'} \right)^\lambda \\
&\leq \theta_Q^\lambda \operatorname{diam}(Q) \sum_{j \geq 0} 2^{j(d - 1 - \lambda d)} \\
&\leq C \theta_Q^\lambda \operatorname{diam}(Q).
\end{aligned}$$

□

Now, the idea is to replace each term in (3.4) by its convolution with the kernel $\overline{\rho}_{Q'}$. Unfortunately, this will make appear extra divergence terms around each node. We have to modify X so as to make this extra divergence vanish using, for instance, Theorem 3.4. Furthermore, we cannot follow the construction for all generations $j \geq 1$, otherwise the "enlarged edges" (convolution of a dyadic edge and the kernel $\rho_{\theta,\varepsilon}$) may overlap. This is the reason why we introduce the following definition:

Definition 3.6 ("Diffusion level")**.** For a cube Q_0 and $f \in L^2_+(Q_0)$ we define the set $\mathcal{D}(Q_0, f)$ or $\mathcal{D}(f) \subset \mathcal{Q}$ as the maximal element for the inclusion in the set

$$\Lambda = \{ D \subset \mathcal{Q} \,:\, \forall Q \in D,\ \mathcal{A}(Q) \cup \mathcal{C}(F(Q)) \subset D \quad \text{and} \quad \operatorname{supp} \rho_Q \subset Q \}.$$

If $\Lambda = \emptyset$, that is $\operatorname{supp}\rho_{Q_0} \not\subseteq Q_0$, we take the convention $\mathcal{D}(f) = \emptyset$. For all $x \in Q_0$, define also the "generation index" of x associated to f as

$$j(f,x) = \max\{j \ : \ \exists Q \in \mathcal{D}(f) \cap \mathcal{Q}_j,\ x \in Q\} \in \mathbb{N} \cup \{\pm\infty\},$$

where the convention $\max(\emptyset) = -\infty$ has been used.

In this way, each cube in $\mathcal{D}(f)$ contains the support of its associated kernel. Moreover, if Q is an element of $\mathcal{D}(f)$, then all its ancestry and its brothers (*i.e.* elements of the set $\mathcal{C}(F(Q))$) are elements of $\mathcal{D}(f)$. $\mathcal{D}(f)$ can be constructed by induction as follows: first take $j = 0$ and $\mathcal{D}(f) = \emptyset$. If $\operatorname{supp}\rho_{Q_0} \subset Q_0$ then add Q_0 to the set $\mathcal{D}(f)$ and j is replaced by $j+1$. For all cubes Q in $\Lambda \cap \mathcal{Q}_{j-1}$: if all cubes $Q' \in \mathcal{C}(Q) \subset \mathcal{Q}_j$ are such that their associated kernels are supported on Q' then $\mathcal{D}(f)$ is replaced by $\mathcal{D}(f) \cup \mathcal{C}(Q)$. If $\mathcal{D}(f)$ has been changed at this stage j is replaced by $j+1$ and the preceding step is reiterated. This process is repeated for $j \geq 1$ until it fails.

Let $\mathcal{D}_{\min}(f)$ be the set of all cubes in $\mathcal{D}(f)$ which are minimal for the inclusion. If $\mathcal{D}_{\min}(f) \neq \emptyset$, we also define

$$D(f) = \bigcup_{Q \in \mathcal{D}_{\min}(f)} Q.$$

Note that this is actually a disjoint union: two distinct cubes in $\mathcal{D}_{\min}(f)$ are disjoint. Indeed, for Q, $Q' \in \mathcal{D}_{\min}(f) \subset \mathcal{Q}$, either $Q \cap Q' = \emptyset$ or Q and Q' are comparable: $Q \subset Q'$ or $Q' \subset Q$. In the last case, since Q and Q' are minimal, we deduce that $Q = Q'$.

Moreover, it is not difficult to see that, if $\mathcal{D}_{\min}(f) \neq \emptyset$, then $D(f) = \{x \in Q_0 \ : \ j(f,x) \text{ is finite}\}$ and also that $\overline{f}(x) = 0$ whenever $j(f,x) = +\infty$, where $\overline{f}$ is the precise representative of f (*i.e.* the limit of the mean values of f on small cubes). Indeed, assume that $Q \in \mathcal{D}(f)$ is a cube of side length L_Q. Then, by definition, $\operatorname{supp}\rho_Q \subset Q$ and for some constant C depending on ρ and for $\nu = \frac{1-\gamma}{d-1}$, one has $\varepsilon^\gamma \theta_Q^\nu \leq C L_Q$ and so

$$⨍_Q f := L_Q^{-d}\theta_Q \leq \varepsilon^{-\gamma/\nu} L_Q^{1/\nu - d}.$$

Since $1/\nu - d = \frac{(d-1)(\alpha d - d + 1)}{d+1}$ is positive, we deduce that L_Q cannot be arbitrarily small if there exists $x \in Q$ such that $\overline{f}(x) > 0$. Moreover, if $f(x) \geq \eta$ a.e. for some $\eta > 0$, then there exists some constant $C_\eta > 0$ depending on η, ε, d and α such that

$$\forall Q \in \mathcal{D}(f),\ L_Q \geq C_\eta. \tag{3.5}$$

In particular, one can deduce that $\mathcal{D}_{\min}(f) = \emptyset$ if and only if $\mathcal{D}(f) = \emptyset$ or $f(x) = 0$ a.e. Indeed, if $\mathcal{D}(f) = \emptyset$, then it is clear that $\mathcal{D}_{\min}(f) = \emptyset$. Conversely, assume that $\mathcal{D}_{\min}(f) \neq \emptyset$ (*i.e.* $Q_0 \in \mathcal{D}(f)$) and that there exists $x \in Q_0$ such that $\overline{f}(x) > 0$. Since $\bigcup_{Q \in \mathcal{D}(f)} \partial Q$ is negligible for the Lebesgue measure, one can assume that $x \in \bigcup_{Q \in \mathcal{D}(f)} Q$. Then $0 \leq j(f,x) < +\infty$ and so there exists a minimal cube $Q \in \mathcal{D}(f)$ containing x. Then $Q \in \mathcal{D}_{\min}(f)$. Indeed, if $Q' \in \mathcal{D}(f)$ and $Q' \subsetneq Q$, then $\mathcal{A}(Q) \subset \mathcal{D}(f)$ and there exists $Q'' \in \mathcal{A}(Q)$ such that $Q'' \subsetneq Q$ and $x \in Q''$ which is a contradiction.

We are now able to define two approximations of f which are useful for our problem. The first is a dyadic approximation of f by an atomic measure,

$$\Lambda_\varepsilon f = \begin{cases} \sum_{Q \in \mathcal{D}_{\min}(f)} \mathcal{S}_Q & \text{if } \mathcal{D}_{\min}(f) \neq \emptyset, \\ \mathcal{S}_{Q_0} & \text{otherwise,} \end{cases}$$

where we recall the definition of $\mathcal{S}_Q := \theta_Q \delta_{c_Q}$. We also define an approximation in $H^1(Q_0)$,

$$\lambda_\varepsilon f = \begin{cases} \displaystyle\sum_{Q \in \mathcal{D}_{\min}(f)} \rho_Q & \text{if } \mathcal{D}_{\min}(f) \neq \emptyset, \\ \rho_{Q_0} & \text{otherwise,} \end{cases}$$

where ρ_Q is defined in (3.3). The following result shows in which sense $\lambda_\varepsilon f$ is an approximation of f and justifies the term "diffusion level". Indeed, this proposition indicates that we get a good estimate by using a local diffusion from $\lambda_\varepsilon f$ to f, *i.e.* minimizing $\int_Q |\nabla u|^2$ over the constraint $\nabla \cdot u = \lambda_\varepsilon f - f$ for all $Q \in \mathcal{D}_{\min}(f)$ (see Thm. 3.4).

Proposition 3.7. *There exists a constant $C > 0$ depending on d and ρ such that for all $f \in L^2_+(Q_0)$,*

$$d^\alpha_\varepsilon(\lambda_\varepsilon f, f) + d^\alpha(\Lambda_\varepsilon f, f) \leq C\, \varepsilon^{\gamma_2} \|f\|^2_{L^2(Q_0)}.$$

More precisely, if $\operatorname{supp} \rho_{Q_0} \subset Q_0$*, there exists $u \in H^1_0(Q_0)$ such that $\nabla \cdot u = f - \lambda_\varepsilon f$ as well as*

$$M^\alpha_\varepsilon(u) \leq C\, \varepsilon^{\gamma_2} \|f\|^2_{L^2} \quad \text{and} \quad \|u\|_{L^1} \leq C\, \operatorname{diam}(Q_0) \|f\|_{L^1}.$$

If $\operatorname{supp} \rho_{Q_0} \not\subseteq Q_0$ *the same estimates hold but the condition $u \in H^1_0(Q_0)$ has to be replaced by $u \in H^1_0(\widetilde{Q}_0)$, where $\widetilde{Q}_0$ is a cube containing Q_0 and* $\operatorname{supp} \rho_{Q_0}$*.*

Proof. First assume that $\operatorname{supp} \rho_{Q_0} \subset Q_0$ *i.e.* $Q_0 \in \mathcal{D}(f)$. If $\mathcal{D}_{\min}(f) = \emptyset$, then $f(x) = 0$ a.e. and the proposition is trivial. Hence, one can assume that $\mathcal{D}_{\min}(f) \neq \emptyset$. Then f is supported on $D(f)$ and $\mathcal{D}_{\min}(f) =: \{Q_i\}_{i \in I}$ is a finite or countable partition of $D(f)$. Denote for simplicity $D_i := \operatorname{diam}(Q_i)$, $f_i := f \mathbf{1}_{Q_i}$ (restriction of f to Q_i), $\theta_i := \theta_{Q_i}$ and $\rho_i := \rho_{Q_i} = \theta_i\, \rho_{R_i}$ for $i \in I$, where

$$R_i := R_{Q_i} = \varepsilon^\gamma \theta_i^{\frac{1-\gamma}{d-1}}.$$

Since Q_i is minimal in $\mathcal{D}(f)$, we deduce that, for some constants $C, C' > 0$,

$$C' R_i \leq D_i \leq C R_i. \tag{3.6}$$

Indeed, the first inequality follows from the fact that $\operatorname{supp} \rho_i \subset Q_i$ and $\operatorname{diam}(\operatorname{supp} \rho_i) = cR_i$ for some constant c depending on ρ. For the second inequality observe that, since Q_i is minimal, there exists $Q \in \mathcal{C}(Q_i)$ such that $\operatorname{supp} \rho_Q \not\subseteq Q$ and hence $R_Q \geq c' \operatorname{diam}(Q) = c'/2 D_i$ for some constant $c' > 0$ depending on ρ. Since $\theta_Q \leq \theta_{Q_i} = \theta_i$, one has $R_Q \leq R_i$ and the second inequality follows.

Now, Theorem 3.4 allows us to find $u_i \in H^1_0(Q_i)$ such that $\nabla \cdot u_i = g_i$, $\|u_i\|_{H^1(Q_i)} \leq C\, \|g_i\|_{L^2(Q_i)}$ and $\|u_i\|_{L^1(Q_i)} \leq C\, D_i \|g_i\|_{L^1(Q_i)}$, where $g_i := f_i - \rho_i$. Since u_i vanishes at ∂Q_i, one can extend u_i by 0 out of Q_i to get a function in $H^1(\mathbb{R}^d)$: for the sake of simplicity, this function is still denoted by u_i. Consequently, $u = \sum_i u_i$ belongs to $H^1_0(Q_0)$ and $\nabla \cdot u = f - \lambda_\varepsilon f$. It remains to estimate $M^\alpha_\varepsilon(u)$ and $\|u\|_{L^1(Q_0)}$. First of all,

$$\|u\|_{L^1(Q_0)} \leq \sum_i \|u_i\|_{L^1(Q_i)} \leq C\, \operatorname{diam}(Q_0) \sum_i \|g_i\|_{L^1(Q_i)}$$

and the inequality $\|g_i\|_{L^1(Q_i)} \leq 2\theta_i$ leads to $\|u\|_{L^1} \leq 2C\, \operatorname{diam}(Q_0) \|f\|_{L^1}$ as required.

Let us compute the L^2-norm of ρ_i:

$$\|\rho_i\|^2_{L^2(Q_i)} = \theta_i^2\, \|\rho_{R_i}\|^2_{L^2(Q_i)} = \theta_i^2 R_i^{-d} \|\rho\|^2_{L^2(Q_i)} = C \theta_i^2 R_i^{-d}.$$

By a Cauchy–Schwarz inequality,

$$\theta_i^2 = \left(\int_{Q_i} f_i \right)^2 \leq |Q_i| \|f_i\|^2_{L^2(Q_i)} = D_i^d \|f_i\|^2_{L^2(Q_i)} \tag{3.7}$$

which, together with (3.6), gives

$$\|\rho_i\|^2_{L^2(Q_i)} \leq C\, R_i^d \|f_i\|^2_{L^2(Q_i)} R_i^{-d} = C\, \|f_i\|^2_{L^2(Q_i)}.$$

Since $\|u_i\|_{H^1(Q_i)} \le C\|f_i - \rho_i\|_{L^2(Q_i)}$, we get $\|u_i\|_{H^1(Q_i)} \le C\|f_i\|_{L^2(Q_i)}$. Now, because the energy M_ε^α is local and since each u_i is supported on Q_i, one has

$$M_\varepsilon^\alpha(u) = \sum_{i=1}^n M_\varepsilon^\alpha(u_i) = \sum_{i=1}^n \left(\varepsilon^{-\gamma_1} \int_{Q_i} |u_i|^\beta + \varepsilon^{\gamma_2} \int_{Q_i} |\nabla u_i|^2 \right)$$

By construction of u_i, one has

$$\int_{Q_i} |\nabla u_i|^2 \le \|u_i\|^2_{H^1(Q_i)} \le C\|g_i\|^2_{L^2(Q_i)} \le 2C \left(\|\rho_i\|^2_{L^2(Q_i)} + \|f_i\|^2_{L^2(Q_i)} \right) \le C' \|f_i\|^2_{L^2(Q_i)}.$$

It remains to estimate the first term. First of all, we use the Hölder and Poincaré inequalities as follows:

$$\int_{Q_i} |u_i|^\beta \le |Q_i|^{1-\beta/2} \left(\int_{Q_i} |u_i|^2 \right)^{\beta/2} \le D_i^{d-d\beta/2} \left(D_i^2 \int_{Q_i} |\nabla u_i|^2 \right)^{\beta/2} \le D_i^\nu \|f_i\|^\beta_{L^2(Q_i)},$$

where $\nu = \beta + d - \frac{d\beta}{2}$. In view of (3.7) and (3.6), we have

$$D_i \le CR_i = C\varepsilon^\gamma \theta_i^{\frac{1-\gamma}{d-1}} \le C\varepsilon^\gamma (D_i^{\frac{d}{2}} \|f_i\|_{L^2(Q_i)})^{\frac{1-\gamma}{d-1}}$$

and, introducing $\delta := 1 - \frac{d(1-\gamma)}{2(d-1)}$,

$$D_i^\delta \le C\varepsilon^\gamma \|f_i\|_{L^2}^{\frac{1-\gamma}{d-1}}. \tag{3.8}$$

Finally, since $-\gamma_1 + \frac{\gamma\nu}{\delta} = \gamma_2$ and $\beta + \frac{\nu(1-\gamma)}{\delta(d-1)} = 2$, we get

$$\varepsilon^{-\gamma_1} \int_{Q_i} |u_i|^\beta \le C\varepsilon^{-\gamma_1 + \frac{\gamma\nu}{\delta}} \|f_i\|^{\beta + \frac{\nu(1-\gamma)}{\delta(d-1)}}_{L^2(Q_i)} = C\varepsilon^{\gamma_2} \|f_i\|^2_{L^2(Q_i)}.$$

The proof of the second inequality is quite similar but easier:

$$d^\alpha(\Lambda_\varepsilon f, f) \le \sum_{i=1}^n d^\alpha(\mathcal{S}_{Q_i}, f_i) \le \sum_{i=1}^n \theta_i^\alpha D_i.$$

Once again, applying (3.7) and then (3.8), we get

$$d^\alpha(\Lambda_\varepsilon f, f) \le C\,\varepsilon^{\gamma_2} \|f\|^2_{L^2}.$$

In the case where $\operatorname{supp} \rho_{Q_0} \not\subseteq Q_0$, *i.e.* $R_{Q_0} := \varepsilon^\gamma \theta_Q^{\frac{1-\gamma}{d-1}} \ge CL$ (L being the side length of Q_0 and C a constant depending on ρ), the proof is the same. Indeed, we just apply Theorem 3.4 to $g = f - \rho_{Q_0}$ and the same computations as above lead to the same result. □

3.2. Proof of Theorem 3.2

Let $Q_0 = (0, L)^d$, $L > 0$ and $f \in L^2_+(Q_0)$ with $\int_{Q_0} f = \theta$. In the case where $\operatorname{supp} \rho_{Q_0} \not\subseteq Q_0$, Theorem 3.2 is a particular case of Proposition 3.7. Consequently, one can assume that $\operatorname{supp} \rho_{Q_0} \subset Q_0$ *i.e.* $Q_0 \in \mathcal{D}(f)$. In the case where $\mathcal{D}(f) = \{Q_0\}$, one has $\lambda_\varepsilon f = \rho_{Q_0}$ and Theorem 3.2 is a consequence of Proposition 3.7 as well. For this reason, one can assume that $\mathcal{C}(Q_0) \subset \mathcal{D}(f)$. Moreover, up to replacing f by $f + \eta$ for some small constant $\eta > 0$ and passing to the limit when $\eta \to 0$, one can assume that $\mathcal{D}(f)$ is finite. Indeed, in view of (3.5), $\mathcal{D}(f + \eta)$ is finite since for all $Q \in \mathcal{D}(f + \eta)$, $\operatorname{diam}(Q) \ge C_\eta > 0$.

Our aim is to prove that there exists $C > 0$ only depending on α, d and ρ such that

$$d^\alpha_\varepsilon(f, \rho_{Q_0}) \le C\big\{\theta^\alpha L + \varepsilon^{\gamma_2} \|f\|^2_{L^2(Q_0)}\big\}.$$

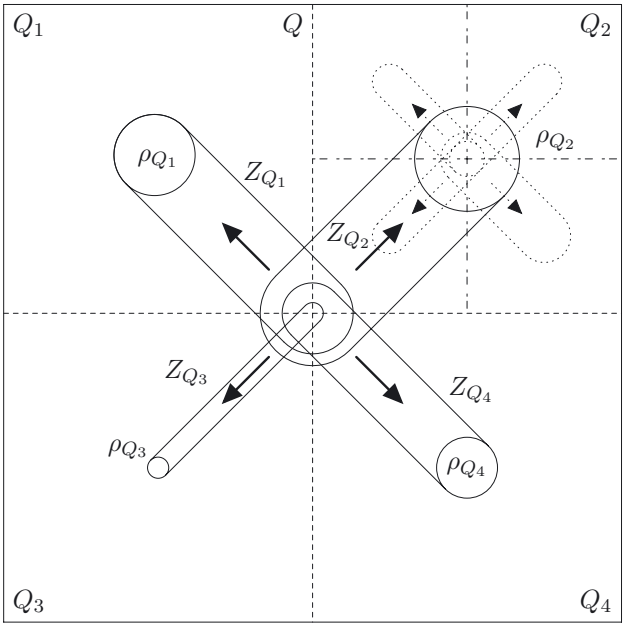

FIGURE 1. A square Q and its 4 dyadic children Q_i with the associated vector field Z_Q.

The idea of the proof is to approximate the vector field $X = \sum X_Q$ of the previous section (see (3.4)) by a vector field in H^1 using the kernel ρ. In this part, we will use the notations of the previous section: in particular, the definition of $\mathcal{D}(f)$ in Definition 3.6, the measures X_Q in (3.4) and $X = \sum_{Q\in\mathcal{D}(f)} X_Q$.

Since $\mathcal{C}(Q_0) \subset \mathcal{D}(f)$, we can construct the regularized vector field Y by the formula

$$Y = \sum_{\substack{Q\in\mathcal{D}(f)\\ Q\neq Q_0}} Z_Q,$$

where, for all $Q \in \mathcal{D}(f)$ such that $Q \neq Q_0$ (see Fig. 1),

$$Z_Q := \theta_Q \; n_Q \; \overline{\rho}_Q * \mathcal{H}^1_{\lfloor [c_{F(Q)}, c_Q]}, \tag{3.9}$$

n_Q being the normalized vector $n_Q = \frac{c_Q - c_{F(Q)}}{\|c_Q - c_{F(Q)}\|}$ and $\overline{\rho}_Q$ being defined in (3.2).

By definition of the kernel $\overline{\rho}_Q$, one has

$$M^\alpha_\varepsilon(Z_Q) \leq C\theta_Q^\alpha \; \mathrm{diam}(Q). \tag{3.10}$$

This a consequence of the choice of R_Q as a minimizer in (1.2). Indeed, for the sake of simplicity, let us assume that $\mathrm{supp}\,\rho$ is the unit ball centered at the origin. Then Z_Q is concentrated on a strip of width R_Q around the segment $S = [c_{F(Q)}, c_Q]$, *i.e.*

$$\mathrm{supp}\, Z_Q \subset \tilde{S} := \{x \in \mathbb{R}^d \; : \; \mathrm{dist}(x, S) \leq R_Q\} \tag{3.11}$$

and Z_Q satisfies the two estimates

$$\|Z_Q\|_{L^\infty} \leq C\theta_Q R_Q^{1-d} \quad \text{and} \quad \|\nabla Z_Q\|_{L^\infty} \leq C\theta_Q R_Q^{-d}. \tag{3.12}$$

Then, the same computations as in (1.2) and the fact that $R_Q \leq \operatorname{diam}(Q)$ give (3.10).

Let us estimate the L^1 norm of Y which has to be controlled by θ as stated in Theorem 3.2:

$$\|Y\|_{L^1(Q_0)} \leq \sum_{j\geq 1} \sum_{Q\in\mathcal{D}(f)\cap\mathcal{Q}_j} \|Z_Q\|_{L^1(\tilde{S})} \leq \sum_{j\geq 1} \sum_{Q\in\mathcal{D}(f)\cap\mathcal{Q}_j} \theta_Q\, L\, 2^{-j} \;= L\,\theta.$$

Note that

$$\nabla\cdot Y = \rho_{Q_0} - h - \lambda_\varepsilon f,$$

where h stands for the extra divergence. h can be written as

$$h = \sum_{Q\in\mathcal{D}_{fr.}} \left\{ \rho_Q - \sum_{Q'\in\mathcal{C}(Q)} \rho_{Q',Q} \right\},$$

where $\rho_{Q',Q}$ represents the kernel $\rho_{Q'}$ translated at c_Q, center of Q, and, for the sake of simplicity, the set of all cubes Q such that $\mathcal{C}(Q)\subset\mathcal{D}(f)$ has been denoted by $\mathcal{D}_{fr.}$:

$$\mathcal{D}_{fr.} := \{Q\in\mathcal{D}(f) \;:\; \mathcal{C}(Q)\subset\mathcal{D}(f)\}.$$

Since $\nabla\cdot Y = \rho_{Q_0} - f + (f-\lambda_\varepsilon f) - h \neq \rho_{Q_0} - f$, we have to slightly modify the vector field Y. This will be done replacing Y by

$$V = Y + V_1 + V_2,$$

where V_1, $V_2 \in H^1(Q_0,\mathbb{R}^d)$ are constructed so that $\nabla\cdot V_1 = h$ and $\nabla\cdot V_2 = \lambda_\varepsilon f - f$. The construction of V_1 and the estimate of $M^\alpha_\varepsilon(V_1)$, $\|V_1\|_{L^1}$ will be the object of the first step. In the second step we prove that $M^\alpha_\varepsilon(Y)\leq C\theta^\alpha L$. Then, Proposition 3.7 allows us to construct $V_2\in H^1$ such that $\nabla\cdot V_2 = \lambda_\varepsilon f - f$ with an estimate on $M^\alpha_\varepsilon(V_2)$ and $\|V_2\|_{L^1}$.

3.2.1. First step: Correction at the nodes, construction of V_1

For all $Q\in\mathcal{D}_{fr.}$, let B_Q be the support of ρ_Q. Since $\operatorname{supp}\rho$ has been supposed to be the unit ball centered at the origin and $\rho_Q(x) = \theta_Q\rho_{R_Q}(x-c_Q)$, we have $B_Q = B(c_Q,R_Q)\subset Q$. Let us define the extra divergence corresponding to this node,

$$h_Q = \rho_Q - \sum_{Q'\in\mathcal{C}(Q)} \rho_{Q',Q}.$$

For each $Q\in\mathcal{D}_{fr.}$, thanks to Theorem 3.4, we can find $V_Q\in H^1_0(B_Q)$ such that $\nabla\cdot V_Q = h_Q$ and $\|V_Q\|_{H^1(B_Q)}\leq C\,\|h_Q\|_{L^2(B_Q)}$. But in this case, because h_Q is radial up to a translation, we essentialy use the proposition in dimension 1 which is quite easy and gives better estimates. Let us give more details on this point:

Lemma 3.8. *Let $d\geq 1$ and $B = B(0,R)\subset\mathbb{R}^d$ be a ball centered at the origin. There exists a constant $C>0$ only depending on d such that the following holds.*

Let $F\in L^\infty(B)$ be a radial function: i.e. *for a.e. $x\in B$, $F(x) = f(|x|)$ for some $f\in L^\infty(0,R)$. Assume that $\int_B F = 0$. Then, there exists a radial function $V\in W^{1,\infty}_0(B,\mathbb{R}^d)$ such that $\nabla\cdot V = F$ and*

$$\|\nabla V\|_{L^\infty(Q_0)} \leq C\,\|F\|_{L^\infty(Q_0)}.$$

Proof. First of all, by a scaling argument, one can assume that $R=1$. The vector field $V: B\to\mathbb{R}^d$ defined by $V(x) = v(|x|)x$ for some Lipschitz function $v:\mathbb{R}^+\to\mathbb{R}$ satisfies

$$\nabla\cdot V(x) = r^{1-d}[r^d v(r)]'$$

in the distributional sense. Thus, if v is chosen as

$$v(r) = r^{-d}\int_0^r f(s)s^{d-1}\,\mathrm{d}s,$$

then V solves the following problem:

$$\begin{cases} \nabla \cdot V(x) = F(x) & \text{on } B, \\ V(x) = 0 & \text{on } \partial B. \end{cases}$$

Moreover, for a.e. $x \in B$, we have $\nabla V(x) = v'(|x|)x \otimes \frac{x}{|x|} + v(|x|)\,\mathrm{Id}$, where Id is the matrix identity. In particular, we get $\|\nabla V\|_{L^\infty} \leq C(\|rv'(r)\|_{L^\infty} + \|v\|_{L^\infty})$. The second term in the RHS on the preceding equation is estimated by $\|v\|_{L^\infty} \leq r^{1-d}\|f\|_{L^\infty} r^{d-1} = \|f\|_{L^\infty}$. For the first term, one has $v'(r) = -dr^{-d-1}\int_0^r f(s)s^{d-1}\,\mathrm{d}s + r^{-1}f(r)$ and so $\|rv'(r)\|_{L^\infty} \leq C\|f\|_{L^\infty}$. Thus, $\|\nabla V\|_{L^\infty} \leq C\|F\|_{L^\infty}$. □

Applying Lemma 3.8 to $F = h_Q$ and $R = R_Q$ gives $V_Q \in H^1_0(B_Q)$ such that $\nabla \cdot V_Q = h_Q$ and

$$\|\nabla V_Q\|_{L^\infty(B_Q)} \leq C\theta_Q R_Q^{-d} \quad \|\nabla V_Q\|_{L^1(B_Q)} \leq |B_Q|\|\nabla V_Q\|_{L^\infty(B_Q)} \leq C\theta_Q. \tag{3.13}$$

Moreover, since V_Q is supported on $B_Q = B(c_Q, R_Q)$, we deduce that $\|V_Q\|_{L^\infty(B_Q)} \leq R_Q\|\nabla V_Q\|_{L^\infty(B_Q)} \leq C\theta_Q R_Q^{1-d}$ so that V_Q satisfies the same estimate as (3.12). In particular, we get $M^\alpha_\varepsilon(V_Q) \leq C\theta_Q^\alpha \operatorname{diam}(Q)$. Now define

$$V_1 = \sum_{Q \in \mathcal{D}_{fr.}} V_Q.$$

Since $\|V_Q\|_{L^1(B_Q)} \leq CR_Q\|\nabla V_Q\|_{L^1(B_Q)} \leq C\operatorname{diam}(Q)\theta_Q$, Lemma 3.5 implies

$$\|V_1\|_{L^1(Q_0)} \leq C\operatorname{diam}(Q_0)\theta_{Q_0} \leq C'L\|f\|_{L^1(Q_0)}$$

as required. Then, using the definition of M^α_ε in (2.4) and the subadditivity of $x \to |x|^\beta$, one gets

$$M^\alpha_\varepsilon(V_1) \leq \varepsilon^{-\gamma_1} \sum_{Q \in \mathcal{D}_{fr.}} \int |V_Q|^\beta + 2\,\varepsilon^{\gamma_2} \int \sum_{Q,Q' \in \mathcal{D}_{fr.}\,:\,Q' \subset Q} |\nabla V_{Q'} : \nabla V_Q|\,, \tag{3.14}$$

where $A : B$ stands for the euclidian product of two matrices $A = (A_{ij})_{1\leq i,j\leq d}$, $B = (B_{ij})_{1\leq i,j\leq d}$ of size $d \times d$: $A : B := \sum_{ij} A_{ij}B_{ij}$. For the estimate of $|\nabla V_1|^2$, we have used the identity $|\nabla V_1|^2 = \nabla V_1 : \nabla V_1 = \sum_{Q,Q' \in \mathcal{D}_{fath}} \nabla V_Q : \nabla V_{Q'}$. Since V_Q is supported on Q, $\nabla V_Q : \nabla V_{Q'}$ vanishes except when $Q \cap Q' \neq \emptyset$, *i.e.* $Q \subset Q'$ or $Q' \subset Q$, thus justifying the factor 2 and the inclusion $Q' \subset Q$ in (3.14).

We need to estimate the two terms in (3.14). Since $M^\alpha_\varepsilon(V_Q) \leq C\theta_Q^\alpha \operatorname{diam}(Q)$, thanks to Lemma 3.5, this term is less or equal than $C\theta^\alpha L$ as required. Using the inequality $\|fg\|_{L^1} \leq \|f\|_{L^\infty}\|g\|_{L^1}$, one can estimate the second term of (3.14) by

$$2\,\varepsilon^{\gamma_2} \sum_{Q,Q' \in \mathcal{D}_{fr.}\,:\,Q' \subset Q} \|\nabla V_Q\|_{L^\infty(B_Q)}\|\nabla V_{B_{Q'}}\|_{L^1(B_{Q'})}.$$

Note that it would be more natural to use a Cauchy–Schwarz inequality (L^2-L^2) at this step but, using it, we were not able to deduce the estimate by $\theta^\alpha L$. Once again, since $R_{Q'} \leq \operatorname{diam}(Q')$, we have

$$\|\nabla V_{Q'}\|_{L^1(B_{Q'})} \leq C\theta_{Q'} \leq \operatorname{diam}(Q')R_{Q'}^{-1}\theta_{Q'} = C\operatorname{diam}(Q')\varepsilon^{-\gamma}\theta_{Q'}^{1-\frac{1-\gamma}{d-1}}. \tag{3.15}$$

Since $1 - \frac{1}{d} < 1 - \frac{1-\gamma}{d-1} < 1$, Lemma 3.5 gives

$$\sum_{Q' \in \mathcal{D}_{fr.}\,:\,Q' \subset Q} \|\nabla V_{Q'}\|_{L^1(B_{Q'})} \leq C\varepsilon^{-\gamma}\operatorname{diam}(Q)\theta_Q^{1-\frac{1-\gamma}{d-1}}.$$

Now, elementary computations on exponents α, γ_2, γ and Lemma 3.5 give successively $\gamma_2 = (d+1)\gamma$, $\alpha = 2 - (d+1)\frac{1-\gamma}{d-1}$ and

$$C\varepsilon^{\gamma_2} \sum_{Q \in \mathcal{D}_{fr.}} \operatorname{diam}(Q)\theta_Q R_Q^{-d}\varepsilon^{-\gamma}\theta_Q^{1-\frac{1-\gamma}{d-1}} = C \sum_{Q \in \mathcal{D}_{fr.}} \operatorname{diam}(Q)\theta_Q^\alpha \leq C\theta^\alpha L.$$

Finally, we have obtained the desired inequality: $M^\alpha_\varepsilon(V_1) \leq C\,\theta^\alpha L$.

3.2.2. Second step: Estimate of the energy of Y on the node set

In order to get estimates on Y, it is convenient to divide Q_0 into 2 domains: the node set N and its complementary N^c, where

$$N := \bigcup_{Q\in\mathcal{D}(f)} B(c_Q, cR_Q)$$

and $c > 0$ is a constant which will be chosen later. By analogy with V_1, one can write $Y_{\lfloor N}$ as a sum of vector fields Y_Q, where

$$Y_Q = \begin{cases} \mathbf{1}_{B(c_Q,cR_Q)}\left(Z_Q - \sum_{Q'\in\mathcal{C}(Q)} Z_{Q'}\right) & \text{if } Q \in \mathcal{D}_{fr.} \text{ (see (3.9))},\\ \mathbf{1}_{B(c_Q,cR_Q)} Z_Q & \text{otherwise.}\end{cases}$$

Now, from (3.12), we deduce the estimates (3.13) satisfied by V_Q are also true for Y_Q and consequently, we obtain $M^\alpha_\varepsilon(Y, N) \le C\,\theta^\alpha L$ as well (see (2.4) for the definition of $M^\alpha_\varepsilon(Y, N)$).

3.2.3. Third step: Estimate of the energy of Y out of the node set

Reminding that

$$Y = \sum_{\substack{Q\in\mathcal{D}(f)\\ Q\neq Q_0}} Z_Q,$$

considering that M^α_ε is not subadditive (due to the term $|\nabla Y|^2$), the first thing to do is to understand to which extent the supports of Z_Q can intersect. To this aim, let us note that if the constant $c > 0$ in (3.2.2) is chosen equal to $\sqrt{d}$ or more, due to (3.11), then each Z_Q restricted to N^c is supported on Q (see Fig. 1): $\operatorname{supp} Z_Q \cap N^c \subset Q$. In particular, this implies that

$$\operatorname{supp} Z_Q \cap \operatorname{supp} Z_{Q'} \cap N^c \neq \emptyset \implies Q \cap Q' \neq \emptyset \implies Q \subset Q' \quad \text{or} \quad Q' \subset Q.$$

For this reason, $M^\alpha_\varepsilon(Y, N^c)$ can be estimated exactly in the same way as we did for the estimate of $M^\alpha_\varepsilon(V_1)$ in (3.14). Moreover, the Young inequality, $\|f * \mu\|_{L^1} \le \|f\|_{L^1}|\mu|(\mathbb{R}^d)$, valid for all $f \in L^1(\mathbb{R}^d)$, $\mu \in \mathcal{M}(\mathbb{R}^d)$, and the definition of Z_Q in (3.9), easily give

$$\|\nabla Z_{Q'}\|_{L^1(Q')} \le C\theta_{Q'} R_{Q'}^{-1} \operatorname{diam}(Q').$$

Since this estimate (which is the same as (3.15)) and (3.10) are the only ones we have used in the first step for the estimate of $M^\alpha_\varepsilon(V_1)$, we get $M^\alpha_\varepsilon(Y, N^c) \le C\theta^\alpha L$ as well.

3.2.4. End of the proof of Theorem 3.2

Finally, the vector field $V = Y + V_1 + V_2$, where V_2 is given by Proposition 3.7, satisfies $\nabla \cdot V = \rho_{Q_0} - f$,

$$M^\alpha_\varepsilon(V) \le 3\{M^\alpha_\varepsilon(Y) + M^\alpha_\varepsilon(V_1) + M^\alpha_\varepsilon(V_2)\} \le C\{\theta^\alpha L + \varepsilon^{\gamma_2}\|f\|^2_{L^2}\}$$

and

$$\|V\|_{L^1} \le \|Y\|_{L^1} + \|V_1\|_{L^1} + \|V_2\|_{L^1} \le CL\|f\|_{L^1}.$$

4. Estimate between d^α_ε and the Wasserstein distance

Our aim is to prove an estimate on the pseudo-distances d^α_ε similar to Proposition 2.2. Because of the Dirichlet term in the definition of M^α_ε, d^α_ε cannot be estimated only by the Wasserstein distance W_1 but one has to add a term involving $\|f^+ - f^-\|_{L^2}$. Using Theorem 3.2, we are going to prove the following theorem.

Theorem 4.1. *Let $Q = (0, L)^d$ be a a cube of side length $L > 0$ in $\mathbb{R}^d$ and $\varepsilon \in (0, 1)$. There exists $C > 0$ only depending on α, d and L such that for all $f^+, f^- \in L^2_+(Q)$ with $\int_Q f^+ = \int_Q f^- = 1$, there exists $u \in H^1(\mathbb{R}^d)$ compactly supported on the set $Q_\varepsilon := \{x \in \mathbb{R}^d \ : \ \mathrm{dist}(x, Q) \leq C\varepsilon^\gamma\}$ satisfying $\nabla \cdot u = f := f^+ - f^-$ as well as*

$$d^\alpha_\varepsilon(f^+, f^-) \leq M^\alpha_\varepsilon(u) \leq C\, H\big(W_1^{1-d(1-\alpha)}(f^+, f^-) + \varepsilon^{\gamma_2}\|f\|^2_{L^2}\big) \quad \text{and} \quad \|u\|_{L^1} \leq C, \tag{4.1}$$

where $H : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is the scalar function defined by $H(x) = x + x^\lambda$ for some $\lambda \in (0, 1)$ depending on α, and W_1 stands for the Wasserstein distance associated to the Monge cost $(x, y) \to |x - y|$.

Remark 4.2. One can replace the condition $\int f^\pm = 1$ by $\int f^\pm = \theta \geq 0$. Then, the constant C will also depend on θ: $C = C(\theta, \alpha, d, L)$. However, we can easily check that C is locally bounded with respect to θ, *i.e.* it is uniform for bounded values of θ.

Remark 4.3. It is tempting to think that estimate (4.1) also holds when $H(x) = x$ which would be the natural choice. Indeed, if ε is taken very small, since M^α_ε Γ-converge to M^α and because of Proposition 2.2, one can expect that $d^\alpha_\varepsilon(f^+, f^-) \simeq d^\alpha(f^+, f^-) \leq CW_1(f^+, f^-)^{1-d(1-\alpha)}$. On the contrary, when ε is very large, because of Theorem 3.4, one can expect that $d^\alpha_\varepsilon(f^+, f^-) \simeq \varepsilon^{\gamma_2}\|f\|^2_{L^2}$. However, for technical reasons, due to the lack of subadditivity of the second term (Dirichlet energy) in the definition of M^α_ε, we were not able to reach the case $H(x) = x$.

Proof. Our method to prove this proposition is an adaptation of that of Morel and Santambrogio in [17] (see also Prop. 6.16, p. 64 in [3]).

Up to replacing (f^+, f^-) by $(f^+ - f^+ \wedge f^-, f^- - f^+ \wedge f^-)$, one can assume that $f^+ \wedge f^- = 0$, where for all $x \in Q$, $(f^- \wedge f^+)(x) = \inf(f^-(x), f^+(x))$. Indeed, it is sufficient to note that, if $\mu^\pm$ are two measures with the same mass and ν is a positive measure on Q then we have $W_1(\mu^+ + \nu, \mu^- + \nu) = W_1(\mu^+, \mu^-)$.

For the sake of simplicity, in all the proof, $C > 0$ will denote some constant only depending on α, d and L and big enough so that all the inequalities below are satisfied.

Let $f^+, f^- \in L^2_+(Q)$ be two densities on the cube $Q = (0, L)^d$ such that $\int_Q f^\pm = 1$. Chose an optimal transport plan Π between f^+ and f^- for the Monge–Kantorovich problem associated to the cost $c(x, y) = |x-y|$. Hence Π satisfies the constraint $P^\pm_\# \Pi = f^\pm(x)\,\mathrm{d}x$ where P^+ (resp. P^-) is the projection on the first variable x (resp. the second variable y) and $\mathrm{d}x$ is the Lebesgue measure. Moreover we have

$$\int_Q |x - y|\,\mathrm{d}\Pi(x, y) = W_1(f^+, f^-) =: W. \tag{4.2}$$

So as to use the local estimate of the previous part, let us classify the set of ordered pairs (x, y) with respect to the distance $|x - y|$. More precisely, for $j \geq 0$, set

$$X_j = \{(x, y) \in Q^2 \ : \ d_j \leq |x - y| < d_{j+1}\},$$

where $d_j = (2^j - 1)\, w$ and $w \in (0, 1)$ will be chosen later. In particular, $d_0 = 0$ and X_j is empty if $d_j > \mathrm{diam}(Q)$, *i.e.* $j > J := \left\lfloor \ln_2\left(\frac{\mathrm{diam}(Q)}{w} + 1\right)\right\rfloor$. For this reason, one can restrict to integers $j \leq J \leq C(1 + |\ln w|)$: we will assume that $d_j \leq \mathrm{diam}(Q)$. Moreover, (4.2) immediately gives the estimate

$$\sum_j d_j\theta_j \leq W\,, \quad \text{where} \quad \theta_j = \Pi(X_j). \tag{4.3}$$

Next, for each integer $j \in [1, J]$, consider a uniform partition of Q into cubes Q_{jk}, $k = 1, \ldots, K_j$, with side length d_{j+1}. It is easy to estimate K_j by

$$K_j \leq Cd^{-d}_{j+1}. \tag{4.4}$$

For $j \geq 0$, set

$$\Pi_j = \Pi_{|X_j} \; ; \; \theta_j = \Pi(X_j); \; f_j^{\pm} = P_{\#}^{\pm} \Pi_j \quad \text{and} \quad f_j = f_j^+ - f_j^-,$$

Clearly, one has

$$\Pi = \sum_j \Pi_j \quad \text{and} \quad f^{\pm} = \sum_j f_j^{\pm}.$$

In the same way, for $j \geq 0$ and $1 \leq k \leq K_j$, set

$$\Pi_{jk} = \Pi_{|X_j \cap (Q_{jk} \times Q)} \; ; \; \theta_{jk} = \Pi_{jk}(Q^2) \quad \text{and} \quad f_{jk}^{\pm} = P_{\#}^{\pm} \, \Pi_{jk}$$

so that

$$\Pi_j = \sum_k \Pi_{jk}; \; \theta_j = \sum_k \theta_{jk} \quad \text{and} \quad f_j^{\pm} = \sum_k f_{jk}^{\pm}.$$

Π_{jk} represents the part of the transport plan Π corresponding to points in Q_{jk} which are sent at a distance comparable to d_{j+1}. In particular, f_{jk}^+ is supported on Q_{jk} and f_{jk}^- is supported on the cube $\widetilde{Q}_{jk}$ with the same center but twice the side length of Q_{jk}. As we did in (3.3), let us define ρ_{jk} the kernel associated to Q_{jk} by

$$\rho_{jk}(x) = \begin{cases} (R_{jk})^{-d} \rho(R_{jk}(x - c_{jk})) & \text{if } \theta_{jk} \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\rho \in \mathcal{C}_c^1(\mathbb{R}^d, \mathbb{R}^+)$, $R_{jk} = \varepsilon^{\gamma} \theta_{jk}^{\frac{1-\gamma}{d-1}}$ and c_{jk} is the center of Q_{jk}. For the sake of simplicity, let us assume that $\operatorname{supp} \rho$ is the unit ball centered at the origin. Let $B_{jk} := B(c_{jk}, r_{jk})$ be the smallest ball containing $\widetilde{Q}_{jk}$ and $\operatorname{supp} \rho_{jk} = B(c_{jk}, R_{jk})$: *i.e.* $r_{jk} = \max\{R_{jk}, \operatorname{diam}(Q_{jk})\}$. Thanks to Theorem 3.2, it is possible to find a vector field $u_{jk} \in H_0^1(B_{jk})$ satisfying $\nabla \cdot u_{jk} = f_{jk} := f_{jk}^+ - f_{jk}^-$, $\|u_{jk}\|_{L^1(B_{jk})} \leq C \theta_{jk}$ and

$$M_{jk} := M_{\varepsilon}^{\alpha}(u_{jk}) \leq C \, \{\theta_{jk}^{\alpha} d_{j+1} + \varepsilon^{\gamma_2} \|f_{jk}\|_{L^2(B_{jk})}^2\}. \tag{4.5}$$

Moreover, if $R_{jk} \geq d_{j+1}/2$, the first term in the right-hand side of (4.5) can be omitted since one has

$$\theta_{jk}^{\alpha} d_{j+1} \leq C \varepsilon^{\gamma_2} \|f_{jk}\|_{L^2}^2. \tag{4.6}$$

Indeed, in this case, writing $\theta := \theta_{jk}$ and $R := R_{jk}$, one has $\theta^{\alpha} d_{j+1} \leq 2 \theta^{\alpha} R$ and, using $2 - \alpha = \frac{(1-\gamma)(d+1)}{d-1}$, we get $\theta^{\alpha} R = [\theta^{\alpha-2} R^{1+d}][\theta^2 R^{-d}] = \varepsilon^{\gamma_2} R^{-d} \theta^2$. Then, (4.6) follows from the fact that, by the Cauchy–Schwarz inequality, we have

$$R^{-d} \theta^2 \leq R^{-d} |B_{jk}| \int_{B_{jk}} (f_{jk})^2 \leq C \int_{B_{jk}} (f_{jk})^2.$$

Now, let us define the vector field $u = \sum_{j,k} u_{jk}$, which satisfies

$$\nabla \cdot u = \sum_{j,k} \nabla \cdot u_{jk} = \sum_{j,k} f_{jk} = f := f^+ - f^-.$$

First note that

$$\|u\|_{L^1(Q)} \leq C \sum \|u_{jk}\|_{L^1(B_{jk})} \leq 2C \sum \theta_{jk} = 2C.$$

In order estimate the energy of u, a similar development of $|\sum \nabla u_{jk}|^2$ as in (3.14) and the Cauchy–Schwarz inequality give

$$M_{\varepsilon}^{\alpha}(u) \leq J \sum_{j=1}^{J} M_{\varepsilon}^{\alpha} \left(\sum_{k=1}^{K_j} u_{jk} \right) \leq C \, J \sum_j \left\{ \sum_k M_{jk} + \sum_{(k,l) \in I_j} \sqrt{M_{jk}} \sqrt{M_{jl}} \right\}, \tag{4.7}$$

where I_j stands for the set of pairs (k, l) satisfying $k \neq l$, $\theta_{jk} \geq \theta_{jl}$ and $B_{jk} \cap B_{jl} \neq \emptyset$. We have to estimate the two terms in the right-hand side of (4.7).

4.1. Estimate of the first term in (4.7)

We recall that $M_{jk} \leq \theta_{jk}^{\alpha} d_{j+1} + \varepsilon^{\gamma_2} \|f_{jk}\|^2_{L^2(B_{jk})}$. For the second term, note that

$$\sum_{j,k} \|f_{jk}\|^2_{L^2} \leq \|f\|^2_{L^2}. \tag{4.8}$$

Indeed, since $f^+ \wedge f^- = 0$, for all j,k, one has $f^+_{jk} \wedge f^-_{jk} = 0$ as well. In particular, $\|f_{jk}\|^2_{L^2(B_{jk})} = \|f^+_{jk}\|^2_{L^2(B_{jk})} + \|f^-_{jk}\|^2_{L^2(B_{jk})}$, $\|f\|^2_{L^2(Q)} = \|f^+\|^2_{L^2(Q)} + \|f^-\|^2_{L^2(Q)}$ and (4.8) follows from the super-additivity of the power function $x \to |x|^p$ for $p \geq 1$: $|x+y|^p \geq |x|^p + |y|^p$ for $x, y \in \mathbb{R}$ whenever $xy \geq 0$.

For the first term, applying successively the Jensen inequality with power $\alpha \in (0,1)$, the Hölder inequality, (4.3) and the fact that $K_j d_{j+1} = C d_{j+1}^{1-d}$ (see (4.4)), one gets

$$\begin{aligned}
\sum_{j,k} \theta_{jk}^{\alpha} d_{j+1} &\leq \sum_j d_{j+1} K_j [\theta_j / K_j]^{\alpha} = \sum_j [d_{j+1}\theta_j]^{\alpha} [d_{j+1} K_j]^{1-\alpha} \\
&\leq \left(\sum_j \theta_j \, d_{j+1} \right)^{\alpha} \left(\sum_j d_{j+1} K_j \right)^{1-\alpha} \\
&\leq C(w+W)^{\alpha} \left(\sum_j [w(2^{j+1}-1)]^{1-d} \right)^{1-\alpha} \\
&\leq C'(w^{\alpha} + W^{\alpha}) w^{(1-d)(1-\alpha)}
\end{aligned}$$

since $\theta_0 d_1 \leq d_1 = w$ (we cannot estimate this term by W because $d_0 = 0$) and, because of (4.3), $\sum_{j\geq 1} \theta_j \, d_{j+1} \leq 3 \sum_{j \geq 1} \theta_j \, d_j \leq 3W$. Finally, we get

$$\sum_{j,k} M_{jk} \leq C \left\{ w^{1-d(1-\alpha)} + W^{\alpha} w^{-(d-1)(1-\alpha)} + \varepsilon^{\gamma_2} \|f\|^2_{L^2} \right\}. \tag{4.9}$$

4.2. Estimate of the second term in (4.7)

Before following these computations, we need to understand what the condition "$B_{jk} \cap B_{jl} \neq \emptyset$" is meaning. Assume that $(k,l) \in I_j$. From $Q_{jk} \cap Q_{jl} = \emptyset$, we see that either $\operatorname{supp} \rho_{jk}$ or $\operatorname{supp} \rho_l^j$ is not included in Q_{jk} (resp. Q_{jl}). Since, by definition of I_j, we have $\theta_{jk} \geq \theta_{jl}$, this implies that $R_{jk} \geq d_{j+1}/2$. Therefore, as we noticed after formula (4.5),

$$M_{jk} \leq \varepsilon^{\gamma_2} \|f_{jk}\|^2_{L^2(B_{jk})}$$

and (4.6) also implies that

$$\theta_{jl}^{\alpha} d_{j+1} \leq \theta_{jk}^{\alpha} d_{j+1} \leq C \varepsilon^{\gamma_2} \|f_{jk}\|^2_{L^2(B_{jk})}.$$

Now, (4.5), the subadditivity of the square root function, the preceding inequality, (4.8) and Cauchy–Schwarz inequality give in turn

$$\begin{aligned}
\sum_{(k,l)\in I_j} \sqrt{M_{jk}} \sqrt{M_{jl}} &\leq C \sum_{(k,l)\in I_j} \sqrt{\varepsilon^{\gamma_2} \|f_{jk}\|_2^2} \left(\sqrt{\varepsilon^{\gamma_2} \|f_{jl}\|_2^2} + \sqrt{\theta_{jl}^{\alpha} d_{j+1}} \right) \\
&\leq C \varepsilon^{\gamma_2} \sum_{(k,l)\in I_j} \|f_{jk}\|_2^2 + \|f_{jk}\|_2 \|f_{jl}\|_2 \\
&\leq C \varepsilon^{\gamma_2} \left\{ K_j \|f_j\|^2_{L^2(Q)} + \sqrt{\sum_{k,l} \|f_{jk}\|_2^2} \sqrt{\sum_{k,l} \|f_{jk}\|_2^2} \right\} \\
&\leq 2C \varepsilon^{\gamma_2} K_j \|f_j\|^2_{L^2(Q)}.
\end{aligned}$$

From $K_j \le d_{j+1}^{-d} \le 2^{-dj} w^{-d}$ and $\|f_j\|_{L^2(Q)}^2 \le \|f\|_{L^2(Q)}^2$, we obtain in the end that

$$\sum_j \sum_{(k,l)\in I_j} \sqrt{M_{jk}}\sqrt{M_{jl}} \le C w^{-d} \varepsilon^{\gamma_2} \|f\|_{L^2}^2. \tag{4.10}$$

4.3. End of the proof

Let $F = \varepsilon^{\gamma_2}\|f\|_{L^2}^2$. We remind the definition of $W = W_1(f^+, f^-)$. One can assume that $f^- \neq f^+$ so that $F, W > 0$. Now, (4.7), (4.9), (4.10) and the fact that $J \le C(1 + \ln w)$ yield

$$M_\varepsilon^\alpha(u) \le C(1 + |\ln w|)\left\{w^\nu + W^\alpha w^{\nu-\alpha} + w^{-d}F\right\},$$

where $\nu := 1 - d(1-\alpha) \in (0,1)$ and so $\alpha - \nu = -(d-1)(1-\alpha) < 0$. Let us fix some $\delta \in (0,1)$ small enough so that $0 < \nu \pm \delta < 1$ and $\nu - \alpha \pm \delta < 0$. For some constant c depending on δ, one has $1 + |\ln w| \le c(w^\delta + w^{-\delta})$ and so

$$M_\varepsilon^\alpha(u) \le C\left\{w^{\nu\pm\delta} + W^\alpha w^{\nu-\alpha\pm\delta} + w^{-d\pm\delta}F\right\},$$

where the sum is taken over the values of ± 1 ($+1$ or -1) in the right-hand side. Then, we make the choice $w = W + F^\lambda > 0$ for some $\lambda = \lambda(\alpha, d) > 0$ which will be fixed later. Note that all the estimates above are valid only if $w < 1$. However, if $W + F^\lambda \ge 1$ then the right-hand side of (4.1) is greater than some positive constant and (4.1) easily follows from Theorem 3.2 since $H(x) \ge x$. Thus, one can assume that $w \in (0,1)$.

Since $0 < \nu \pm \delta < 1$, we get $w^{\nu\pm\delta} \le W^{\nu\pm\delta} + F^{\lambda(\nu\pm\delta)}$ and, because $-d \pm \delta < 0$, $\nu - \alpha \pm \delta < 0$, we have $w^{\nu-\alpha\pm\delta} \le W^{\nu-\alpha\pm\delta}$ and $w^{-d\pm\delta} \le F^{\lambda(-d\pm\delta)}$ which gives

$$M_\varepsilon^\alpha(u) \le C\left\{W^{\nu\pm\delta} + F^{\lambda(\nu\pm\delta)} + W^{\nu\pm\delta} + F^{1+\lambda(-d\pm\delta)}\right\}.$$

We fix $\lambda > 0$ small enough so that $1 + \lambda(-d \pm \delta) > 0$: in this way, all the exponents in the preceding formula are positive. Finally, (4.1) follows from the fact that we have W, $F \le 1$ as a consequence of W, $F \le W^{1-d(1-\alpha)} + F$. □

Remark 4.4. Since $\min\{w^\nu + w^{-d}F \ : \ w \in (0,1)\} = cF^{\frac{1}{d+\nu}}$ and $\frac{1}{d+\nu} < 1$, one cannot obtain an estimate of the form $M_\varepsilon^\alpha(u) \le C(W + F)$ as expected. However, one could improve a bit (4.1) by a better estimate of the number of indices l such that $(k,l) \in I_j$.

5. A Γ-convergence result

Let $\Omega \subset \mathbb{R}^2$ be a bounded open set and $\mu = \mu^+ - \mu^-$ be a finite measure, where $\mu^\pm$ are two probability measures compactly supported on Ω. We recall the definition of the set

$$\mathcal{M}_{div}(\Omega) = \left\{u : \Omega \to \mathbb{R}^2 \ : \ u \text{ and } \nabla\cdot u \text{ are finite measures on } \overline{\Omega}\right\}$$

which is endowed with the topology of weak star convergence on vector measures and their divergence. As weak star topology is never metrizable in infinite dimensional Banach spaces, the space $\mathcal{M}_{div}(\Omega)$ is not metrizable. Indeed, assume that X is some infinite dimensional Banach space such that X' is metrizable. In particular X' admits a countable neighborhood basis $(V_n)_{n\ge1}$ which one can assume to be of the form

$$V_n = \{\varphi \ : \ |\langle \varphi\, ; x_i\rangle| < \varepsilon_n \text{ for } i = 1, \dots, n\}$$

for some linearly independent family of vectors $(x_i)_{i\ge1} \subset X$ and $\varepsilon_n > 0$. Then the Hahn–Banach Theorem easily provides a sequence $(\varphi_n)_{n\ge1}$ satisfying $\varphi_n(x_i) = 0$ for all $i \le n \in \mathbb{N}^*$ and $\|\varphi_n\|_{X'} = n$. In particular the sequence $(\varphi_n)_n$ weakly converges to 0 as $n \to \infty$ which is a contradiction with the fact that $(\varphi_n)_n$ is norm unbounded.

However, every bounded subsets of the dual space of a separable Banach space are metrizable for the weak star topology. In particular, for the natural norm $\|u\|_{\mathcal{M}_{div}(\Omega)} = |\nabla \cdot u|(\Omega) + |u|(\Omega)$ given by the total variation of u and its divergence, we know that all bounded subsets of $\mathcal{M}_{div}(\Omega)$ are metrizable: for all $M > 0$, there exists a metric d_M for the weak star convergence of u and $\nabla \cdot u$ on the set

$$\mathcal{M}_M(\Omega) = \{u \in \mathcal{M}_{div}(\Omega) \ : \ |u|(\Omega) + |\nabla \cdot u|(\Omega) \leq M\}.$$

In [18] the Γ-convergence of the functional sequence M_ε^α to M^α was proved. Our aim is to prove that this property remains true when adding a divergence constraint. Since, for $u \in H^1(\Omega)$, one has $\nabla \cdot u \in L^2$, one cannot prescribe $\nabla \cdot u = \mu$ if μ is not in L^2. For this reason, we first have to define a regularization of μ. Let $(f_\varepsilon)_{\varepsilon>0} \subset L^2$ be a sequence of L^2 functions weakly converging to μ as measures and satisfying

$$\int_\Omega f_\varepsilon(x)\,\mathrm{d}x = 0 \quad \text{and} \quad \varepsilon^{\gamma_2}\|f_\varepsilon\|_{L^2}^2 \underset{\varepsilon\to 0}{\longrightarrow} 0. \tag{5.1}$$

This choice is going to be useful for the proof of Theorem 1.1. For example, we can define f_ε as

$$f_\varepsilon := \rho_\varepsilon * \mu,$$

where $\rho_\varepsilon(x) = \varepsilon^{-2\gamma}\rho(\varepsilon^{-\gamma}x)$ for some compactly supported $\rho \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^+)$ such that $\int_\Omega \rho = 1$ and γ is still defined as $\gamma = \frac{\gamma_2}{d+1} = \frac{\alpha+1}{3}$. Now, let us define the functionals $\overline{M}_\varepsilon^\alpha$ (resp $\overline{M}^\alpha$) adding a divergence constraint on $u \in \mathcal{M}_{div}(\Omega)$:

$$\overline{M}^\alpha(u) = \begin{cases} M^\alpha(u) & \text{if } \nabla \cdot u = \mu, \\ +\infty & \text{otherwise,} \end{cases}$$

$$\overline{M}_\varepsilon^\alpha(u) = \begin{cases} M_\varepsilon^\alpha(u) & \text{if } \nabla \cdot u = f_\varepsilon, \\ +\infty & \text{otherwise.} \end{cases}$$

The main result of this section is Theorem 1.1:

Theorem 5.1. *There exists a constant c_0 such that the functional sequence $(\overline{M}_\varepsilon^\alpha)_{\varepsilon>0}$ Γ-converges to $c_0\overline{M}^\alpha$ as $\varepsilon \to 0$. Moreover c_0 is given by the minimum value for the minimization problem (5.2).*

We first remind how to build a recovery sequence in the case of a mass θ flowing on a single segment S, *i.e.* $u = \theta\mathcal{H}^1_{|S}$. To this aim, we need to find a structure close to u which is almost optimal for M_ε^α. We proceed by a slicing argument:

Let u be any vector measure in $\mathcal{M}_{div}(\Omega)$. Take some $\nu \in S^1 := \{x \in \mathbb{R}^2 \ : \ |x| = 1\}$ which has to be thought as the tangent vector to S in the case where $u = \theta\mathcal{H}^1_{|S}$. Let us consider $v = [(u\cdot\nu)_+]_{|\nu^\perp}$ (restriction on $\nu^\perp$ of the positive part of $u \cdot \nu$) the flux of u across the hyperplane $\nu^\perp = \{x \in \mathbb{R}^2 \ : \ x \cdot \nu = 0\}$ and assume that $\int v = \theta$. Then $M_\varepsilon^\alpha(u)$ can be controlled from below by integrals on subintervals of $\mathbb{R}\nu$ of the following Cahn–Hilliard type energy (see [10] for physical motivations):

$$F_\varepsilon^\beta(v) = \varepsilon^{-\gamma_1}\int_{\mathbb{R}} v^\beta + \varepsilon^{\gamma_2}\int_{\mathbb{R}} |\nabla v|^2.$$

This kind of models for droplets equilibrium was studied by Bouchitté *et al.* in [6] for instance (see also [5]). $F_\varepsilon^\beta(v)$ can be renormalized through the formula: $v(x) = \theta R_{\theta,\varepsilon}^{-d} w(R_{\theta,\varepsilon}^{-1}x)$, where $R_{\theta,\varepsilon} = \varepsilon^\gamma\theta^{\frac{1-\gamma}{d-1}}$. Then, the constraint $\int v = \theta$ turns into $\int w = 1$ and $F_\varepsilon^\beta(v) = \theta^\alpha F^\beta(w)$, where

$$F^\beta(w) = \left\{\int_{\mathbb{R}} w^\beta + \int_{\mathbb{R}} |\nabla w|^2\right\}.$$

Then, the existence of an optimal profile w is given by

Lemma 5.2. *There exists a profile $w \in H^1_{loc}(\mathbb{R},\mathbb{R}^+)$ solution of the minimization problem*

$$\min\left\{\int_{\mathbb{R}} w^\beta + \int_{\mathbb{R}} |w'|^2 \ :\ w \in H^1_{loc}(\mathbb{R},\mathbb{R}^+) \quad \textit{and} \quad \int_{\mathbb{R}} w = 1\right\}. \tag{5.2}$$

Moreover, w is compactly supported, Lipschitz continuous on $\mathbb{R}$ and $\mathcal{C}^\infty$ inside its support, i.e. *on the open set $\{w>0\}$. Last of all, it is possible to choose w such that it is even and non-increasing on $\mathbb{R}^+$.*

Remark 5.3. Although we restrict to the two dimensional case, some works by Bouchitté *et al.* (see [13]) suggest that Lemma 5.2 and Theorem 1.1 could be generalized in every dimension as well. However, the aim of this paper is to use the tools of Section 4 so as to establish the Γ-lim sup property for functionals $\overline{M}^\alpha_\varepsilon$ (with divergence constraint) and, from the point of view of the complexity of the proof, this is independent of the dimension. Since the Γ-lim inf property was only established in the $2D$ case in [18], we prefer to stay in this framework. Actually, the difficulty to prove a Γ-convergence result of $\overline{M}^\alpha_\varepsilon$ (resp. M^α_ε) to $\overline{M}^\alpha$ (resp. M^α) in every dimension would concern the Γ-lim inf part and this is not the purpose of this paper.

Remark 5.4. Note that the minimum value in (5.2) is related to the best constant in the one-dimensional Gagliardo-Nirenberg inequality $\int_{\mathbb{R}}|u| \le C\left(\int_{\mathbb{R}}|u'|^2\right)^{\frac{1-\beta}{2+\beta}}\left(\int_{\mathbb{R}}|u|^\beta\right)^{\frac{3}{2+\beta}}$:

$$\frac{1}{C} = \inf\left\{\left(\int_{\mathbb{R}}|u'|^2\right)^{\frac{1-\beta}{2+\beta}}\left(\int_{\mathbb{R}}u^\beta\right)^{\frac{3}{2+\beta}} \ :\ u\in H^1_{loc}(\mathbb{R},\mathbb{R}^+) \quad \text{and} \quad \int_{\mathbb{R}} u = 1\right\}.$$

Proof. First notice that there exists a finite energy configuration, *i.e.* $w \in H^1(\mathbb{R},\mathbb{R})$ such that $F^\beta(w) < +\infty$. Indeed, every compactly supported and nonnegative $\mathcal{C}^1$ function has finite energy. Let take a minimizing sequence, *i.e.* $(w_n)_n \subset H^1(\mathbb{R}) \subset \mathcal{C}^0(\mathbb{R})$ such that $F^\beta(w_n) \to c_\beta$. One can assume that w_n is even and non-increasing on $\mathbb{R}^+$. Indeed if w^* stands for the spherical rearrangement of some $w \in H^1(\mathbb{R},\mathbb{R}^+)$, then one has $\int_{\mathbb{R}}|w^*|^\beta = \int_{\mathbb{R}}|w|^\beta$ and the classical Polya–Szego Theorem states that the spherical rearrangement reduces the Dirichlet energy of w:

$$\int_{\mathbb{R}}|(w^*)'|^2 \le \int_{\mathbb{R}}|w'|^2.$$

In particular $F^\beta(w^*) \le F^\beta(w)$ as announced. Since $(w_n')_n$ is bounded in $L^2(\mathbb{R})$, one can assume that it weakly converges in $L^2(\mathbb{R})$. Moreover, as $w_n \ge 0$ a.e. and $\int_{\mathbb{R}} w_n = 1$, $(w_n)_n$ is bounded in $L^1(\mathbb{R})$. Thanks to the Poincaré-Wirtinger inequality, one deduces that $(w_n)_n$ is bounded in $H^1_{loc}(\mathbb{R})$. Up to extraction, one can assume that $(w_n)_n$ weakly converges in $H^1_{loc}(\mathbb{R})$. Let call $w \in H^1_{loc}(\mathbb{R})$ the limit. In particular, $(w_n)_n$ strongly converges to w in $L^1_{loc}(\mathbb{R})$ and so w is even, nonnegative and non-increasing on $\mathbb{R}^+$. Moreover, the Fatou lemma and the weak convergence of w_n' yields

$$F^\beta(w) \le \liminf_{n\to\infty} F^\beta(w_n) = c_\beta.$$

In order to prove that w is a global minimizer it remains to prove that w satisfies the constraint $\int w = 1$. Indeed, from the Fatou Lemma we can only deduce that $\int w \le 1$. One has to prove the strong convergence of w_n in $L^1(\mathbb{R})$. Since w_n converges in L^1_{loc}, it is enough to prove that the sequence (w_n) is tight. Let $R>0$. For all $n \ge 1$, since w_n is non increasing on $[0,R]$, one has $\|w_n\|_{L^\infty(\{x\,:\,|x|>R)} = w_n(R)$ and Markov's inequality yields $w_n(R) \le \frac{1}{2R}\int_{\mathbb{R}} w_n = \frac{1}{2R}$. Hence

$$\int_{|x|>R} w_n(x)\,\mathrm{d}x \le w_n(R)^{1-\beta}\int_{\mathbb{R}} w_n^\beta(x)\,\mathrm{d}x \le (2R)^{\beta-1}F^\beta(w_n) \le \frac{C}{R^{1-\beta}}$$

for some constant $C>0$ non depending on n which implies that $(w_n)_n$ is tight since $1-\beta>0$. Now, let check the regularity of w: Lipschitz continuous and smooth inside its support. Note that we already know that $w \in \mathcal{C}^{0,1/2}(\mathbb{R})$ thanks to the Sobolev embedding $H^1_{loc}(\mathbb{R}) \subset \mathcal{C}^{0,1/2}(\mathbb{R})$. Let check the regularity of w inside its

support. Since w is even and non-increasing on $\mathbb{R}^+$, the set $\{x \ : \ w(x) > 0\}$ is an interval $(-R, R)$ for some $R > 0$. Since w is a minimizer of the minimizing problem (5.2), w also satisfies the following Euler–Lagrange equation

$$\forall x \in (-R, R), \ -w''(x) + \beta w(x)^{\beta-1} = \lambda, \tag{5.3}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier associated to the volume constraint $\int w = 1$. Note that $\lambda = -w''(0) + \beta w(0)^{\beta-1} > 0$. Indeed, since $x = 0$ is a global maximum of w, $w'(0) = 0$ and $w''(0) \leq 0$. From (5.3), one deduces that w is smooth on $(-R, R)$. Now, multiplying (5.3) by w' and integrating it on $[0, x]$ or $[x, 0]$ yields

$$\forall x \in (-R, R), \ \frac{w'(x)^2}{2} = w(x)^\beta - \lambda w(x) + \lambda w(0) - w(0)^\beta.$$

Since w is bounded on $\mathbb{R}$, we deduce that w is Lipschitz continuous on $\mathbb{R}$. Last of all, we prove that w is compactly supported, *i.e.* $R < \infty$. Assume by contradiction that $R = \infty$. Then $f \in \mathcal{C}^\infty(\mathbb{R})$, $f > 0$ on $\mathbb{R}$ and integrating (5.3) yields

$$\forall x \in \mathbb{R}, \ w'(x) = \int_0^x (\beta w(y)^{\beta-1} - \lambda) \, \mathrm{d}y.$$

Since $w(x) \underset{x\to\infty}{\longrightarrow} 0$ and $\beta - 1 < 0$, $w(x)^{\beta-1} \underset{x\to\infty}{\longrightarrow} +\infty$ and the right hand side of the preceding equation goes to $+\infty$ as well. Thus $w'(x) \underset{x\to\infty}{\longrightarrow} +\infty$ which is a contradiction. □

Now let $w_{\theta,\varepsilon}$ be defined by $w_{\theta,\varepsilon}(x) = \theta R_{\theta,\varepsilon}^{-d} w(R_{\theta,\varepsilon}^{-1} x)$ where w is the optimal profile of Lemma 5.2 (satisfying all claimed regularity and symmetry properties) and let us introduce the kernel $\rho_{\theta,\varepsilon}$ associated to $w_{\theta,\varepsilon}$, given by the following lemma

Lemma 5.5. *There exists a bounded and compactly supported radial kernel $\rho_{\theta,\varepsilon} \in L_c^\infty(\mathbb{R}^2, \mathbb{R}^+)$ such that $w_{\theta,\varepsilon}$ is the projection of $\rho_{\theta,\varepsilon}$ on the axis $(x_1 = 0)$:*

$$\Pi^2_\sharp \rho_{\theta,\varepsilon}(x) \, \mathrm{d}x = w_{\theta,\varepsilon}(x_2) \, \mathrm{d}x_2,$$

where Π^2 stands for the projection on the second variable, $\mathrm{d}x$ (resp. $\mathrm{d}x_2$) is the Lebesgue measure on $\mathbb{R}^2$ (resp. $\mathbb{R}$) and $\Pi_\sharp \mu$ stands for the pushforward of some measure μ by $\Pi : \mathbb{R}^2 \to \mathbb{R}$. Moreover, one can choose $\rho_{\theta,\varepsilon}$ of the form $\rho_{\theta,\varepsilon}(x) = R_{\theta,\varepsilon}^{-2} \rho(R_{\theta,\varepsilon}^{-1} x)$ for some $\rho \in L_c^\infty(\mathbb{R}, \mathbb{R}^+)$.

Proof. After the renormalization $\rho_{\theta,\varepsilon}(x) = R_{\theta,\varepsilon}^{-2} \rho(R_{\theta,\varepsilon}^{-1} x)$, it remains to find ρ satisfying $\Pi^1_\sharp \rho(x) dx = w(x_2) \, \mathrm{d}x_2$. It is not very difficult to see that a radial solution is given by the formula

$$\rho(x) = \int_{|x|}^\infty \frac{-w'(s)}{\pi \sqrt{s^2 - |x|^2}} \, \mathrm{d}s. \tag{5.4}$$

Details are left to the reader. Let justify how (5.4) implies that ρ is bounded. Since w is compactly supported, there exists $R > 0$ such that $w(x) = 0$ for $|x| > R$. In particular, ρ is compactly supported on $\overline{B}(0, R)$. Then, since $w'(0) = 0$ and since w' is bounded on $\mathbb{R}$ and smooth around 0, one has $|w'(x)| \leq C|x|$ for all $x \in \mathbb{R}$ and some $C > 0$. Hence, there exists a constant $C > 0$ such that for all $x \in \mathbb{R}^2$ such that $|x| =: r \in [0, R)$,

$$\rho(x) = \int_1^R \frac{-w'(rs) \, \mathrm{d}s}{\pi \sqrt{s^2 - 1}} \leq C \int_1^{R/r} \frac{rs \, \mathrm{d}s}{\sqrt{s^2 - 1}} \leq C \left\{ R \int_1^2 \frac{s \, \mathrm{d}s}{\sqrt{s^2 - 1}} + r \int_2^{R/r} \frac{s \, \mathrm{d}s}{\sqrt{s^2 - 1}} \right\}$$

which is bounded since $s \to \frac{s}{\sqrt{s^2-1}}$ is integrable on $[1, 2]$ and bounded on $[2, +\infty)$. □

As a consequence, in the case where $u = \theta\mathcal{H}^1_{|S}$, a recovery sequence, *i.e.* a sequence (u_ε) such that $u_\varepsilon \to u$ in $\mathcal{M}_{div}(\Omega)$ and $M^\alpha_\varepsilon(u_\varepsilon) \to M^\alpha(u)$ as $\varepsilon \to 0$, is obtained as

$$u_\varepsilon = \rho_{\theta,\varepsilon} * u.$$

In the case of a finite energy configuration, *i.e.* $u \in \mathcal{M}_{div}(\Omega)$ such that $M^\alpha(u) < \infty$, thanks to classical properties in the theory of Γ-convergence, it is enough to find a recovery sequence for u belonging to a class of measures which are dense in energy. Thanks to the work of Xia in [23] (see also [18]), we know that the class of vector measures concentrated on finite graphs is dense in energy so that one can restrict to this case. In [23], the branched transportation energy was in fact defined by relaxation of its restriction to the set of vector measures concentrated on a graph. The rectifiability of finite energy configurations and the Eulerian representation (2.1) was discovered later in [24] (see also [3]). This density property was used in [18] to prove the Γ-convergence of M^α_ε toward M^α. In the setting of functionals with divergence constraint, we need the following lemma:

Lemma 5.6. *Let $u \in \mathcal{M}_{div}(\Omega)$ be such that $M^\alpha(u) < \infty$. For all $\lambda > \gamma$, there exists a sequence $(u_\varepsilon) \subset H^1_0(\Omega)$ converging to u in $\mathcal{M}_{div}(\Omega)$ such that*

$$M^\alpha_\varepsilon(u_\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} c_0 M^\alpha(u) \quad \text{and} \quad \varepsilon^\lambda \|\nabla \cdot u_\varepsilon\|_{L^2} \quad \text{is bounded.}$$

Before proving this statement, we are going to investigate the case where u is concentrated on a finite graph. First of all, let us give some details on what "a vector measure concentrated on a finite graph G" is. Let $G = (V(G), E(G), \theta)$ be a weighted directed graph: $V(G) \subset \Omega$ is a finite set of vertices, $E(G)$ is the finite set of oriented edges $\mathbf{e} = (e, \tau_\mathbf{e})$, where $e = [a_e, b_e] \subset \Omega$ and $\tau_\mathbf{e}$ is a unit vector representing the direction of $\mathbf{e}$, and $\theta : E(G) \to (0, +\infty)$ is the weight function. Then the "vector measure associated to G" is given by:

$$u_G = \sum_{\mathbf{e}=(e,\tau_\mathbf{e})\in E(g)} \theta(\mathbf{e})\tau_\mathbf{e}\, \mathrm{d}\mathcal{H}^1_{|e}.$$

These measures u_G belong to $\mathcal{M}_{div}(\Omega)$, *i.e.* $\nabla \cdot u_G$ is a measure, and they are called "transport paths" (see Def. 2.1 in [23]). When u is a transport path, we have the following lemma:

Lemma 5.7. *Let $u = u_G \in \mathcal{M}_{div}(\Omega)$ for some weighted directed graph G. Then, there exists a sequence $(u_\varepsilon)_{\varepsilon>0}$ converging to u in $\mathcal{M}_{div}(\Omega)$ and a constant C depending on u such that, for ε small enough, $u_\varepsilon \in H^1_0(\Omega)$ and*

1. $\int_\Omega |u_\varepsilon| \le |u|(\Omega) + C\,\varepsilon^\gamma$,
2. $\int_\Omega |\nabla \cdot u_\varepsilon| \le |\nabla \cdot u|(\Omega)$,
3. $\varepsilon^\gamma \|\nabla \cdot u_\varepsilon\|_{L^2} \le C$,
4. $|M^\alpha_\varepsilon(u_\varepsilon) - c_0 M^\alpha(u)| \le C\varepsilon^\gamma$.

Proof. By definition, such a vector measure u can be written as a finite sum of measures $u_i = \theta_i\,\tau_i\,\mathcal{H}^1_{|S_i}$ concentrated on a segment $S_i \subset \Omega$ directed by τ_i with multiplicity θ_i for $i = 1, \dots, I$. We first define a regularized vector fied v_ε by $v_\varepsilon := \sum_i v_i$, where $v_i = \rho_{\theta_i,\varepsilon} * u_i$. Then, for ε small enough, v_ε is compactly supported on Ω and satisfies

$$\begin{cases} \int_\Omega |v_\varepsilon| \le |u|(\Omega), \\ |M^\alpha_\varepsilon(v_\varepsilon) - c_0 M^\alpha(u)| \le C\varepsilon^\gamma. \end{cases}$$

The first statement is a consequence of the fact that $\int \rho_{\theta_i,\varepsilon} = 1$ and the inequality $\|f * \mu\|_{L^1} \le \|f\|_{L^1}\,|\mu|(\Omega)$ for $f \in \mathcal{C}_c(\Omega)$ and for a finite measure μ on Ω. For the second statement, by definition of the kernel $\rho_{\theta,\varepsilon}$ we know that, out of the nodes set $N = \bigcup_i \mathrm{supp}(\nabla \cdot v_i)$,

$$M^\alpha_\varepsilon(v_\varepsilon, N^c) = c_0 M^\alpha(v, N^c).$$

As a result, we just have to estimate these energies on N which is a finite union of balls: the supports of $\rho_{\theta_i,\varepsilon}$ recentered at each end-point of the segment S_i. Since the radius of these balls is of the order of ε^γ, this immediately gives the fact that $M^\alpha(u,N) \leq C\varepsilon^\gamma$ for some constant $C>0$ depending on u. For the sake of simplicity, in the rest of this proof, $C>0$ will denote some constant depending on u which is large enough so that all the inequalities below are true. We are going to prove that

$$M^\alpha(u,N) + M^\alpha_\varepsilon(v_\varepsilon,N) \leq C\varepsilon^\gamma.$$

It remains to estimate $M^\alpha_\varepsilon(v_\varepsilon,N)$. Since $M^\alpha_\varepsilon(v_\varepsilon,N) \leq I\sum_i M^\alpha_\varepsilon(v_i,N)$, it is enough to estimate $M^\alpha_\varepsilon(v_i,N)$. But $\|v_i\|_{L^\infty(N)} = \|\rho_{\theta_i,\varepsilon} * u_i\|_{L^\infty(N)} \leq C\varepsilon^{-2\gamma}\|\rho\|_{L^\infty}|u_i|(N_i)$, where $N_i := N + \operatorname{supp}\rho_{\theta_i,\varepsilon} := \{x+y \ : \ x\in N,\, y\in \operatorname{supp}\rho_{\theta_i,\varepsilon}\}$. Note that $\operatorname{supp}\rho_{\theta_i,\varepsilon}$ is a ball centered at the origin with radius smaller than $C\varepsilon^\gamma$ so that N_i is a finite union of balls with radii smaller than $C\varepsilon^\gamma$ as well. In particular, using the fact that $u_i = \theta_i\,\tau_i\,\mathcal{H}^1_{|S_i}$, we get $|u_i|(N_i) \leq C\varepsilon^\gamma$ and so $\|v_i\|_{L^\infty(N)} \leq C\varepsilon^{-\gamma}$. Similarly, one has $\|\nabla v_i\|_{L^\infty(N)} = \|\nabla\rho_{\theta_i,\varepsilon} * u_i\|_{L^\infty(N)} \leq C\varepsilon^{-3\gamma}\|\nabla\rho\|_{L^\infty}|u_i|(N_i) \leq C\varepsilon^{-2\gamma}$. Now, the definition (2.5) gives

$$M^\alpha_\varepsilon(v_\varepsilon,N) = \varepsilon^{\alpha+1}\int_N |\nabla v_\varepsilon|^2 + \varepsilon^{\alpha-1}\int_N |v_\varepsilon|^\beta \leq |N|\{\varepsilon^{\alpha+1}\|\nabla v_\varepsilon\|^2_{L^\infty} + \varepsilon^{\alpha-1}\|v_\varepsilon\|^\beta_{L^\infty}\}.$$

From the inequality $|N| \leq C\varepsilon^{2\gamma}$ and the equalities $\alpha = 3\gamma-1$, $\beta\gamma = \frac{(4\alpha-2)\gamma}{\alpha+1} = 4\gamma-2$, we deduce

$$M^\alpha_\varepsilon(v_\varepsilon,N) \leq C\varepsilon^{2\gamma}\{\varepsilon^{3\gamma-4\gamma} + \varepsilon^{3\gamma-2-\beta\gamma}\} \leq 2C\varepsilon^\gamma$$

as required. In order to construct an approximating vector field with controlled divergence, we need to consider $u_\varepsilon := v_\varepsilon - w_\varepsilon$ where $w_\varepsilon \in H^1_0(N)$ is constructed as follows:

The node set N, defined above, is a finite union $N = \bigcup_{j=1}^n B_j$, where each node B_j is a ball centered at the end-point a_i of some segment $S_i = [a_i,b_i]$. Let assume that ε is small enough so that these balls are non-overlapping. Then, on each node B_j, $g_j := \nabla\cdot v_\varepsilon$ is a finite superposition of kernels like $\rho_{\theta,\varepsilon}$ recentered at c_j, the center of B_j. In particular $\|g_j\|_{L^2(B_j)} \leq C\varepsilon^{-\gamma}$ and $\int_{B_j} g_j = \int_{B_j}\nabla\cdot v_\varepsilon = (\nabla\cdot u)(B_j) =: \theta_j$. If $\theta_j = 0$, then Theorem 3.4 allows to find $w_j \in H^1_0(B_j)$ satisfying $\nabla\cdot w_j = g_j$ and $\|w_j\|_{H^1(B_j)} \leq C\,\varepsilon^{-\gamma}$. If $\theta_j \neq 0$, let say $\theta_j > 0$, we rewrite g_j as $g_j = g^+ - g^- = \lambda g^+ + (1-\lambda)g^+ - g^-$ where g^+ (resp. g^-) stands for the positive part (resp. negative part) of g and $\lambda\in(0,1]$ is chosen such that $(1-\lambda)\int_B g^+ = \int_B g^-$, *i.e.* $\theta_j = \lambda\int_{B_j} g_+$. Applying Theorem 3.4, we get $w_j \in H^1_0(B_j)$ satisfying $\nabla\cdot w_j = (1-\lambda)g^+ - g^-$ and $\|w_j\|_{H^1(N)} \leq C\,\varepsilon^{-\gamma}$. Let us define $w_\varepsilon = \sum_j w_j$ and $u_\varepsilon := v_\varepsilon - w_\varepsilon$. Since $\int_{B_j}|\nabla\cdot u_\varepsilon| = \int_{B_j}|g_j - \nabla\cdot w_j| = \lambda\int_{B_j} g^+ = \theta_j$ for all j, we have

$$\int_\Omega |\nabla\cdot u_\varepsilon| = \int_N |\nabla\cdot u_\varepsilon| \leq \sum_j \theta_j = |\nabla\cdot u|(\Omega).$$

Moreover, to estimate $\|\nabla\cdot u_\varepsilon\|_{L^2}$, note that $\|\nabla\cdot w_\varepsilon\|_{L^2} \leq \|w_\varepsilon\|_{H^1} \leq C\varepsilon^{-\gamma}$ and, because $\nabla\cdot v_\varepsilon$ is only composed of a finite sum of translated kernels of the form $\rho_{\theta_i,\varepsilon}$, $\|\nabla\cdot v_\varepsilon\|_{L^2} \leq C\varepsilon^{-\gamma}$ as well. In particular $\varepsilon^\gamma\|\nabla\cdot u_\varepsilon\|_{L^2}$ is bounded. Then, the Poincaré inequality yields

$$\|w_\varepsilon\|_{L^2} = \sum_j \|w_j\|_{L^2(B_j)} \leq C\sum_j \varepsilon^\gamma\|\nabla w_j\|_{L^2(B_j)} \leq C'$$

since B_j is a ball of radius $C\varepsilon^\gamma$. Consequently, by the Cauchy–Schwarz inequality, we get

$$\int_\Omega |u_\varepsilon| \leq \int_\Omega |v_\varepsilon| + \int_N |w_\varepsilon| \leq |u|(\Omega) + |N|^{1/2}\|w_\varepsilon\|_{L^2} \leq |u|(\Omega) + C\varepsilon^\gamma.$$

Similarly, by a Hölder inequality, we have

$$\int_N |w_\varepsilon|^\beta \leq |N|^{\frac{2-\beta}{2}}\|w_\varepsilon\|^\beta_{L^2} \leq (n\varepsilon^{2\gamma})^{\frac{2-\beta}{2}}\|w_\varepsilon\|^\beta_{L^2} \leq C\varepsilon^{\gamma(2-\beta)}.$$

Once again, since $\alpha = 3\gamma - 1$ and $\beta\gamma = 4\gamma - 2$, we deduce

$$M_\varepsilon^\alpha(w_\varepsilon) = \varepsilon^{\alpha+1}\int_N |\nabla w_\varepsilon|^2 + \varepsilon^{\alpha-1}\int_N |w_\varepsilon|^\beta \le C\{\varepsilon^{\alpha+1-2\gamma} + \varepsilon^{\alpha-1+\gamma(2-\beta)}\} = 2C\varepsilon^\gamma.$$

Since $M^\alpha(u,N) \le C\varepsilon^\gamma$, we get $M_\varepsilon^\alpha(u_\varepsilon,N) \le 2[M_\varepsilon^\alpha(v_\varepsilon,N) + M_\varepsilon^\alpha(w_\varepsilon,N)] \le C\varepsilon^\gamma$ which finally gives

$$|M_\varepsilon^\alpha(u_\varepsilon) - c_0M^\alpha(u)| = |M_\varepsilon^\alpha(u_\varepsilon,N) - c_0M^\alpha(u,N)| \le C\varepsilon^\gamma. \qquad \square$$

Proof of Lemma 5.6. First fix $u \in \mathcal{M}_{div}(\Omega)$ and construct a sequence $(u_n)_{n\ge1}$ converging to u such that $u_n = u_{G_n}$ is a vector measure associated to some weighted directed graph $G_n \subset \Omega$ and $M^\alpha(u_n)$ converges to $M^\alpha(u)$. Since (u_n) weakly converges in $\mathcal{M}_{div}(\Omega)$, the total variations of both measures u_n and $\nabla\cdot u_n$ are bounded by some constant $M > 0$. In the following, we use a metric d on the space $\mathcal{M}_{M+1}(\Omega)$. Extracting a subsequence if necessary, one can suppose that the two following estimates hold

$$d(u_n,u) \le 2^{-n-1} \quad \text{and} \quad |M^\alpha(u_n) - M^\alpha(u)| \le 2^{-n-1}.$$

For each $n \ge 1$, let $u_{\varepsilon,n}$ be a sequence converging to u_n as $\varepsilon \to 0$ and satisfying all properties in Lemma 5.7 for some constant $C = C_n$. Then, one can construct by induction a decreasing sequence $(\varepsilon_n)_{n\ge1} \to 0$ such that for all $n \ge 1$ and $\varepsilon \le \varepsilon_n$, $u_{\varepsilon,n} \in H_0^1(\Omega)$ and

1. $u_{\varepsilon,n} \in \mathcal{M}_{M+1}(\Omega)$,
2. $d(u_{\varepsilon,n},u_n) \le 2^{-n-1}$,
3. $|M_\varepsilon^\alpha(u_{\varepsilon,n}) - c_0M^\alpha(u_n)| \le 2^{-n-1}$,
4. $\varepsilon^{\lambda-\gamma}C_n \le 1$ so that $\varepsilon^\lambda\|\nabla\cdot u_{\varepsilon,n}\|_{L^2} \le 1$.

Indeed, assume that $\varepsilon_n > 0$ satisfies all the asked properties. Then, one can find $\varepsilon_{n+1} \in (0,\varepsilon_n)$ small enough so that

- $C_{n+1}\,\varepsilon_{n+1}^\gamma < 2^{-n-2}$ thus implying the first and third properties (see props. 1, 2 and 4 in Lem. 5.7),
- $C_{n+1}\,\varepsilon_{n+1}^{\lambda-\gamma} < 1$ which is possible since $\lambda > \gamma$
- and $d(u_{\varepsilon,n+1},u_{n+1}) \le 2^{-n-2}$ for all $\varepsilon \in (0,\varepsilon_{n+1})$ which is possible since $u_{\varepsilon,n+1}$ converges to u_{n+1} in $(\mathcal{M}_{M+1}(\Omega),d)$ as $\varepsilon \to 0$.

Now it is quite straightforward that the sequence $(u_\varepsilon)_{\varepsilon>0}$ defined by

$$u_\varepsilon = \begin{cases} u_{\varepsilon,1} & \text{if } \varepsilon > \varepsilon_1, \\ u_{\varepsilon,n} & \text{if } \varepsilon_{n+1} < \varepsilon \le \varepsilon_n \text{ for some } n \ge 1, \end{cases}$$

satisfies all the properties of Lemma 5.6. $\square$

Proof of Theorem 1.1. It is already shown in [18] that $M_\varepsilon^\alpha \xrightarrow{\Gamma} c_0M^\alpha$. We just have to prove that the $\Gamma-\limsup$ property still holds when we add the divergence constraint. In other words, it remains to prove that for all $u \in \mathcal{M}_{div}(\Omega)$ such that $\nabla\cdot u = \mu$, there exists a sequence $(v_\varepsilon)_{\varepsilon>0} \subset \mathcal{M}_{div}(\Omega)$ weakly converging to u as measures, satisfying $\nabla\cdot v_\varepsilon = f_\varepsilon$ (so that (v_ε) also converges in $\mathcal{M}_{div}(\Omega)$) and $M_\varepsilon^\alpha(v_\varepsilon) \underset{\varepsilon\to0}{\longrightarrow} c_0M^\alpha(u)$.

First of all, take a sequence $(u_\varepsilon)_{\varepsilon>0} \subset H_0^1(\Omega)$ converging to u given by Lemma 5.6 for $\lambda = \frac{5\gamma}{4}$: $M_\varepsilon^\alpha(u_\varepsilon) \to M^\alpha(u)$ with $\varepsilon^\lambda\|\nabla\cdot u_\varepsilon\|_{L^2}$ bounded. Then define $g_\varepsilon := f_\varepsilon - \nabla\cdot u_\varepsilon$ the residual divergence. Note that $\int_\Omega g_\varepsilon = 0$. Indeed $\int_\Omega f_\varepsilon = 0$ by assumption (see (5.1)) and $\int_\Omega \nabla\cdot u_\varepsilon = 0$ since $u_\varepsilon \in H_0^1(\Omega)$. Moreover, our assumptions on the sequences $(u_\varepsilon)_{\varepsilon>0}$ and $(f_\varepsilon)_{\varepsilon>0}$ yield

$$\varepsilon^{\gamma_2}\|g_\varepsilon\|_{L^2}^2 \underset{\varepsilon\to0}{\longrightarrow} 0.$$

Indeed, we know that the same estimate is satisfied by $(f_\varepsilon)_{\varepsilon>0}$ thanks to (5.1). Moreover, since $3\gamma = \gamma_2$ and $\lambda = \frac{5\gamma}{2}$, one has $\varepsilon^{\gamma_2}\|\nabla\cdot u_\varepsilon\|_{L^2} = \varepsilon^{\frac{\gamma}{2}}\varepsilon^{\frac{5\gamma}{2}}\|\nabla\cdot u_\varepsilon\|_{L^2} \le C\varepsilon^{\frac{\gamma}{2}} \to 0$. Moreover, since f_ε and $\nabla\cdot u_\varepsilon$ weakly converge to μ

as ε goes to 0, we know that g_ε weakly converges to 0. Let g_ε^+ (resp. g_ε^-) denote the positive (resp. negative) part of g_ε. In order to satisfy the divergence constraint, we may correct u_ε with a vector field w_ε, given by Theorem 4.1 (together with Rem. 4.2), such that $\nabla \cdot w_\varepsilon = g_\varepsilon$,

$$M_\varepsilon^\alpha(w_\varepsilon) \leq H\left(W_1(g_\varepsilon^+, g_\varepsilon^-)^{1-d(1-\alpha)} + \varepsilon^{\gamma_2}\|g_\varepsilon\|_{L^2}^2\right) \underset{\varepsilon\to 0}{\longrightarrow} 0 \tag{5.5}$$

and $\|w_\varepsilon\|_{\mathcal{M}_{div}(\Omega)}$ is bounded, where $H(x) = C(x + x^\delta)$ for some $C > 0$ and $\delta \in (0, 1)$. We deduce that (w_ε) is relatively compact in $\mathcal{M}_{div}(\Omega)$. From (5.5) and the $\Gamma - \liminf$ property, this implies that w_ε converges to 0 in $\mathcal{M}_{div}(\Omega)$. Now, by construction, $v_\varepsilon = u_\varepsilon + w_\varepsilon$ satisfies $\nabla \cdot v_\varepsilon = f_\varepsilon$, $v_\varepsilon \to u$ in $\mathcal{M}_{div}(\Omega)$ and $M_\varepsilon^\alpha(v_\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} c_0 M^\alpha(u)$. Indeed, this last limit is a consequence of □

Lemma 5.8. *Let Ω be some bounded open set in $\mathbb{R}^d$, $d \geq 1$. Let (u_ε), $(v_\varepsilon) \subset H^1(\Omega)$ be two sequences such that $M_\varepsilon^\alpha(u_\varepsilon - v_\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} 0$ and assume that $M_\varepsilon^\alpha(v_\varepsilon)$ is bounded. Then,*

$$|M_\varepsilon^\alpha(u_\varepsilon) - M_\varepsilon^\alpha(v_\varepsilon)| \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

Proof. Let $\nu > 0$ be some constant. For all real matrices A and B of size $d \times d$, by the Young inequality, we have

$$|A + B|^2 = |A|^2 + |B|^2 + 2A : B \leq (1 + \nu)|A|^2 + (1 + 1/\nu)|B|^2.$$

Writing $u_\varepsilon = v_\varepsilon + u_\varepsilon - v_\varepsilon$, we use the preceding inequality for $A = \nabla v_\varepsilon$, $B = \nabla(u_\varepsilon - v_\varepsilon)$ and the subadditivity of $x \to |x|^\beta$ to get

$$M_\varepsilon^\alpha(u_\varepsilon) = \varepsilon^{-\gamma_1}\int_\Omega |u_\varepsilon|^\beta + \varepsilon^{\gamma_2}\int_\Omega |\nabla u_\varepsilon|^2 \leq (1 + \nu)M_\varepsilon^\alpha(v_\varepsilon) + (1 + 1/\nu)M_\varepsilon^\alpha(u_\varepsilon - v_\varepsilon).$$

Since $M_\varepsilon^\alpha(v_\varepsilon) < C$ for some constant $C < +\infty$, we deduce that

$$M_\varepsilon^\alpha(u_\varepsilon) - M_\varepsilon^\alpha(v_\varepsilon) \leq C\nu + (1 + 1/\nu)M_\varepsilon^\alpha(u_\varepsilon - v_\varepsilon).$$

For any value of ε such that $u_\varepsilon \neq v_\varepsilon$, let take $\nu = \sqrt{M_\varepsilon^\alpha(u_\varepsilon - v_\varepsilon)} > 0$. Hence, taking the lim sup when $\varepsilon \to 0$, one gets

$$\limsup_{\varepsilon\to 0}\{M_\varepsilon^\alpha(u_\varepsilon) - M_\varepsilon^\alpha(v_\varepsilon)\} \leq C' \limsup_{\varepsilon\to 0}\sqrt{M_\varepsilon^\alpha(u_\varepsilon - v_\varepsilon)} = 0.$$

Since $M_\varepsilon^\alpha(v_\varepsilon) \leq 2[M_\varepsilon^\alpha(u_\varepsilon) + M_\varepsilon^\alpha(v_\varepsilon - u_\varepsilon)]$ and $M_\varepsilon^\alpha(v_\varepsilon)$ is bounded, we deduce that $M_\varepsilon^\alpha(u_\varepsilon)$ is bounded as well. Then we can apply all the preceding computations exchanging u_ε and v_ε to get $\limsup_{\varepsilon\to 0}\{M_\varepsilon^\alpha(v_\varepsilon) - M_\varepsilon^\alpha(u_\varepsilon)\} \leq 0$ which concludes the proof. □

Acknowledgements. I would like to thank my advisor, Filippo Santambrogio, for discussions and advices which have proved to be very useful. This work was partially supported by the PGMO research project MACRO "Modèles d'Approximation Continue de Réseaux Optimaux".

References

[1] M. Bernot, V. Caselles and J.-M. Morel, T. plans. *Publ. Mat.* **49** (2005) 417–451.
[2] M. Bernot, V. Caselles and J.-M. Morel, The structure of branched transportation networks. *Calc. Var. Partial Differ. Eq.* **32** (2008) 279–317.
[3] M. Bernot, V. Caselles and J.-M. Morel, *Optimal transportation networks: models and theory*. Vol. 1955. Springer Science & Business Media (2009).
[4] F. Bethuel, A counterexample to the weak density of smooth maps between manifolds in Sobolev spaces. Preprint arXiv:1401.1649 (2014).

[5] G. Bouchitté and P. Seppecher, Cahn and Hilliard fluid on an oscillating boundary. In *Motion by mean curvature and related topics* (*Trento, 1992*). De Gruyter, Berlin (1994) 23–42.

[6] G. Bouchitté, C. Dubs and P. 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** (1996) 1103–1108.

[7] J. Bourgain and H. Brezis, On the equation $\operatorname{div} Y = f$ and application to control of phases. *J. Amer. Math. Soc.* **16** (2003) 393–426.

[8] A. Braides, Γ-convergence for beginners. Vol. 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford (2002).

[9] L. Brasco, G. Buttazzo and F. Santambrogio, A Benamou-Brenier approach to branched transport. *SIAM J. Math. Anal.* **43** (2011) 1023–1040.

[10] J.W. Cahn and J.E. Hilliard, Free energy of a nonuniform system. i. interfacial free energy. *J. Chem. Phys.* **28** (1958) 258–267.

[11] G. Dal Maso, An introduction to Γ-convergence. Vol. 8 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA (1993).

[12] E. De Giorgi and T. Franzoni, Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **58** (1975) 842–850.

[13] C. Dubs, *Problèmes de perturbations singulières avec un potentiel dégénéré à l'infini*. Ph.D. thesis, Université de Toulon et du Var (1998).

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

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

[16] L. Modica and S. Mortola, Un esempio di γ-convergenza. *Boll. Un. Mat. Ital. B* **14** (1977) 285–299.

[17] J.-M. Morel and F. Santambrogio, Comparison of distances between measures. *Appl. Math. Lett.* **20** (2007) 427–432.

[18] E. Oudet and F. Santambrogio, A Modica–Mortola approximation for branched transport and applications. *Arch. Ration. Mech. Anal.* **201** (2011) 115–142.

[19] P. Pegon, Equivalence between branched transport models by Smirnov decomposition. To appear in *RICAM* (2017).

[20] F. Santambrogio, A Modica–Mortola approximation for branched transport. *C. R. Math. Acad. Sci. Paris* **348** (2010) 941–945.

[21] F. Santambrogio, *Optimal transport for applied mathematicians. Calculus of variations, PDEs and modeling*. Vol. 87 (2015).

[22] C. Villani, Topics in optimal transportation. Number 58 in *Graduate Studies in Mathematics*. American Mathematical Society, cop. (2003).

[23] Q. Xia, Optimal paths related to transport problems. *Commun. Contemp. Math.* **5** (2003) 251–279.

[24] Q. Xia, Interior regularity of optimal transport paths. *Calc. Var. Partial Differ. Eq.* **20** (2004) 283–299.