ASYMPTOTIC ANALYSIS OF A CLASS OF OPTIMAL LOCATION PROBLEMS

GUY BOUCHITTÉ, CHLOÉ JIMENEZ, RAJESH MAHADEVAN

ABSTRACT. Given a density function f on an compact subset of $\mathbb{R}^d$, we look at the problem of finding the best approximation of f by discrete measures $\nu = \sum c_i \delta_{x_i}$ in the sense of the p-Wasserstein distance, subject to size constraints of the form $\sum h(c_i) \leq \alpha$ where h is a given weight function. This is an important problem with applications in economic planning of locations, in information theory and in shape optimization problems. The efficiency of the approximation can be measured by studying the rate at which the minimal distance tends to zero as α tends to infinity. In this paper, we introduce a rescaled distance which depends on a small parameter and establish a representation formula for its limit as a function of the local statistics for the distribution of the c_i's. The asymptotic problem for large α can be then treated in the case of quite general entropy functions h.

2000MSC :

Keywords : Monge-Kantorovich transportation, Wasserstein distance, entropy, optimal partitions, rate of approximation, asymptotics, Γ-convergence, functionals on measures, integral representation.

1. Introduction

Let Ω be a compact subset of $\mathbb{R}^d$ and let f be a given bounded non-negative density function on Ω. We are interested in the problem of approximation of the diffuse measure $f dx$ by discrete positive measures $\nu := \sum_{i \in \mathbb{N}} c_i \delta_{x_i}$ subject to some restrictions on the size of the approximating measures. The error in approximation can be measured using various distances. In this paper, we will use the celebrated *p-Wasserstein distance*, for $p \geq 1$. This distance which can be seen as the p^{th} moment of the measure $f\,dx$ around ν is given by:

$$W_p^p(f\,dx, \nu) = \inf \int_\Omega |x - Tx|^p f(x)\,dx \tag{1.1}$$

where the infimum is taken over all measurable functions $T : \Omega \mapsto \overline{\Omega}$ such that the image of the measure $f\,dx$ under T is the measure ν (the distance is $+\infty$ if the total mass of f is different from that of ν). Some of the common *size constraints* used (for instance when penalizing large storage of information), are of the form

$$\sharp(\operatorname{spt} \nu) \leq \alpha \quad , \quad \sum_i c_i^q \leq \alpha \quad (-\infty < q < 1) \quad , \quad \sum_i -c_i \ln c_i \leq \alpha \tag{1.2}$$

where α is the maximal permitted size. Similarly one can consider constraints of the form:

$$\sum_i c_i^q \geq \delta \tag{1.3}$$

with $q > 1$ and $\delta > 0$: such restrictions ensure that ν must include enough Dirac masses whose size is not too small. We will incorporate all such constraints in an abstract framework by considering an appropriate weight function $h : \mathbb{R}^+ \to (-\infty, +\infty]$ and the associated *entropy* functional H defined by:

$$H(\nu) := \begin{cases} \sum_i h(c_i) & \text{if } \nu \text{ is discrete,} \\ +\infty & \text{otherwise.} \end{cases} \tag{1.4}$$

We recover the size constraints in (1.2) by taking $h(t) = t^q$ for $q < 1$ or $h(t) = -t \ln t$, whereas for $q > 1$ (1.3) is obtained by choosing $h(t) = -t^q$ and $\alpha = -\delta$. The interesting limit case $q \to -\infty$ leads to consider alternatively

$$H(\nu) := \begin{cases} \sup_i c_i^{-1} & \text{if } \nu \text{ is discrete,} \\ +\infty & \text{otherwise.} \end{cases} \tag{1.5}$$

A common feature of all these examples is that the optimization problem reads

$$\mathcal{E}_H^{p,d}(f, \Omega, \alpha) := \inf\{W_p^p(f, \nu) : H(\nu) \leq \alpha\} , \tag{1.6}$$

or alternatively in the Lagrange multiplier formulation as follows

$$\inf\{W_p^p(f, \nu) + \varepsilon \, H(\nu)\} \tag{1.7}$$

where the parameter $\varepsilon > 0$ is suitably chosen in terms of α.

The existence of an optimal solution ν for (1.6) is straightforward if H is lower semicontinuous with respect to the weak-$*$ topology on measures. A systematic characterization of the lower semicontinuity property (or of the relaxation) of such functionals was obtained in [5, 6, 7]: H is lower semicontinuous iff h is lower semicontinuous and subadditive and satisfies $\lim_{t \to 0^+} \frac{h(t)}{t} = +\infty$. This holds for constraints like in (1.2). However, in the case of (1.3) where in contrast $\lim_{t \to 0^+} \frac{h(t)}{t} = 0$, the function H should be substituted with its relaxation $\overline{H}$ which by [7] reads $\overline{H}(\nu) := H(\tilde{\nu})$ being $\tilde{\nu}$ the atomic part of ν. Notice that all the functionals on measures H we are dealing with are non convex and therefore the global optimization problems under consideration are quite involved.

Let us now give some motivation for studying optimization problems of the kind (1.6). They arise in fact in various contexts of which we mention a few hereafter.

Economic planning : For example, in the problem of planning the location of a certain number of schools to meet the demands of the student population in a district, the variables which arise in the optimization problem (1.6) assume the following significance: f denotes the density of the student population; the measure ν represents the location $\{x_i\}$ and the capacity $\{c_i\}$ of the various schools; the quantity $W_p^p(f, \nu)$ measures the cost of transporting students to their schools. The government would like to plan the schools in such a way that the cost of transportation is minimized while respecting budget constraints for the construction of schools. The function H calculates the construction costs for the distribution ν of the schools.

Information theory: From the point of view of signal processing or image compression, the function f will denote a given distribution of information which needs to be quantized. The distance $W_p^p(f, \nu)$ will model the distortion associated to the quantization of f by the measure ν. This needs to be minimized under admissible limits on the size of the quantization imposed by constraints on the memory space available for storing information.

Granular media: The minimization problem (1.6) can also be reformulated as a problem of *optimal partition*. Indeed, given ν, the distance $W_p^p(f,\nu)$ can be written as:

$$W_p^p(f,\nu) = \inf \left\{ \sum_i \int_{A_i} |x - x_i|^p f(x)dx : \int_{A_i} f(x)\, dx = c_i \right\} \tag{1.8}$$

where the infimum is taken over all partitions $\{A_i\}_i$ such that $\displaystyle\int_{A_i} f(x)\, dx = c_i$ and $x_i \in \Omega$ for all indices i. Let us introduce the *shape function* $\psi_{p,f} : \mathcal{B}(\Omega) \to \mathbb{R}^+$ on the set of Borel subsets of Ω by:

$$\psi_{p,f}(A) := \inf_{a \in \Omega} \int_A |x - a|^p f(x)\, dx. \tag{1.9}$$

Then, the problem (1.6) reduces to searching for a minimal partition $\{A_i\}_i$ of Ω with respect to the following criterium

$$\inf \left\{ \sum_i \psi_{p,f}(A_i) : \sum_i h\left(\int_{A_i} f(x)\, dx \right) \leq \alpha \right\}. \tag{1.10}$$

When α increases, we expect nice (possibly periodic) distributions of small subsets which could be used to simulate granular media.

Optimal design: Recently G. Buttazzo et al. [10] considered the problem of placing a Dirichlet region made up of n balls of given radius in order to minimize the compliance of the configuration with respect to a given source term. They observed that when n tends to infinity, the asymptotic distribution of the centers is directly linked to the limit of problem (1.6) as $\alpha \to \infty$ when taking $H(\nu) = \sharp(\mathrm{spt}\nu)$. In this case the optimal partition problem is equivalent the problem of finding the optimal location of points $x_1, x_2, \dots x_n$ minimizing $\sum_{i=1}^n \psi_{p,f}(A_i)$, being $\{A_i\}$ the *Voronoï* partition induced by $\{x_i\}$:

$$A_i = \{x \in \Omega \ : \ |x - x_i| \leq |x - x_j| \ , \ \forall j \neq i\} \ .$$

An important issue in all the problems listed above is the asymptotic of the infimum value $\mathcal{E}_H^{p,d}(f,\Omega,\alpha)$ as the parameter α tends to infinity (or to some critical value). Before describing our method and main contributions, let us stress that, to the best of our knowledge, the only contributions to this asymptotic problem concern the case just mentioned before where $H(\nu) = \sharp(\mathrm{spt}\nu)$ (i.e. $q = 0$) and $|\cdot|$ is the Euclidean norm: in 1982 Bucklew and Wise [9] proved that

$$\mathcal{E}_H^{p,d}(f,\Omega,N) \sim N^{-\frac{p}{d}}\, C_{p,d} \left(\int_\Omega f^{\frac{d}{d+p}} \right)^{\frac{d+p}{d}} \tag{1.11}$$

where the universal constant $C_{p,d}$ (which depends only on p and d) represents the normalized asymptotic rate of approximation for the Lebesgue measure on a unit cell by n masses, as n goes to infinity. Such a formula motivated by applications in information theory can be useful in practice for determining the size of memory $N(\eta)$ required in order that the error $\mathcal{E}_H^{d,p}(f,\Omega,N(\eta))$ is below a reasonable upperbound η.

The exact value of the constant $C_{p,d}$ is known only in the case $d = 2$ for $p \in \{1, 2\}$: it has been proved by D.J. Newman [19] that

$$C_{2,2} = \int_P |x|^2\ dx = \frac{5}{18\sqrt{3}}\ ,$$

where P is a centered regular hexagon of Lebesgue measure one. Later on, in 2002, R. Bolton et F. Morgan [4] motivated by applications in urban economy considered the same problem for $p = 1$ (Monge distance) and proved a very similar result based on the hexagonal tiling of the plane. They found

$$C_{1,2} = \int_P |x|\, dx = \frac{\mathrm{Log}3 + \frac{4}{3}}{4\left(3\sqrt{\frac{3}{2}}\right)^{1/2}} .$$

In this paper we are going to develop a method which will enable us to derive asymptotic formulas of the kind (1.11) when constraints of the type (1.2) or of the type (1.3) are considered, as well as for many other choices of the weight function h. Having in mind that the size of the sets of the optimal partitions will shrink to zero in the limit process, we introduce a small scaling parameter ε which represents the side of a hypercube whose volume ε^d should be of the same order as the volume of our unknown small sets. Accordingly, we substitute the original weighted *entropy* functional H in (1.4) by the following ε-rescaled version:

$$H_\varepsilon(\nu) \;:=\; \sum_i c_i\, g(\frac{c_i}{\varepsilon^d}) \quad \text{where} \quad g(t) \;:=\; \frac{h(t)}{t} \quad (t > 0) \; . \tag{1.12}$$

We observe in particular that if $h(t) = t^q$ for $0 \leq q < 1$, the constraint in (1.2) can be rewritten as $H_\varepsilon(\nu) \leq 1$ being $\varepsilon = \alpha^{\frac{1}{d(q-1)}}$ and, in view of (1.11), we already know that, at least for $q = 0$, the value of the infimum $\inf\{W_p(f,\nu) \;:\; H_\varepsilon(\nu) \leq 1\}$ is of order ε as ε tends to zero. Expecting that this will be true in a large class of weighted functions h, we then define the ε- *normalized error* as follows

$$\inf \left\{ \frac{W_p(f,\nu)}{\varepsilon} \;:\; \sum_i c_i\, g(\frac{c_i}{\varepsilon^d}) \leq 1 \right\}. \tag{1.13}$$

Our goal is to characterize the limit behavior of the previous optimization problem. A first observation is that any sequence of competitors ν_ε does converge weakly-star to the prescribed density f but this information is not sufficient in order to pass to the limit in (1.13), in particular in the entropy constraint. To overcome this difficulty we look at $H_\varepsilon(\nu)$ defined in (1.12) as a linear form with respect to g (this idea goes back to the theory of Young measures). More precisely, for every discrete measure $\nu = \sum_i c_i\, \delta_{x_i}$ on $\overline{\Omega}$, we write

$$H_\varepsilon(\nu) \;=\; \int_{\overline{\Omega}\times\mathbb{R}^+} g(t)\, \lambda_\varepsilon(\nu)(dxdt) \quad \text{where} \quad \lambda_\varepsilon(\nu) := \sum_i c_i \delta_{(x_i, \frac{c_i}{\varepsilon^d})} \; .$$

The new measure $\lambda_\varepsilon(\nu)$ supported on the product $\overline{\Omega}\times\mathbb{R}^+$ has ν as first marginal whereas the dependence in t accounts for the amplitudes of the Dirac masses. Clearly, if (ν_ε) converges to the density f while $\lambda_\varepsilon(\nu_\varepsilon)$ converges tightly to some limit λ, then the first marginal of λ coincides with f whereas, for every continuous (compactly supported) function g, we will have

$$\lim_{\varepsilon\to 0} H_\varepsilon(\nu_\varepsilon) \;=\; \lim_{\varepsilon\to 0} \int_{\overline{\Omega}\times\mathbb{R}^+} g(t)\, \lambda_\varepsilon(\nu)(dxdt) \;=\; \int_{\overline{\Omega}\times\mathbb{R}^+} g(t)\, \lambda(dxdt) \; . \tag{1.14}$$

At this point the strategy becomes very clear as we are going to pass to the limit in a sequence of variational problems on the space of finite Borel measures on $\overline{\Omega}\times\mathbb{R}^+$ endowed with the topology of the tight convergence. In view of (1.14), the expected limit problem will have the form

$$\min \left\{ E(f,\lambda) \;:\; \int_{\Omega\times\mathbb{R}^+} g(t)\, \lambda(dxdt) \;\leq\; c \right\} \; , \tag{1.15}$$

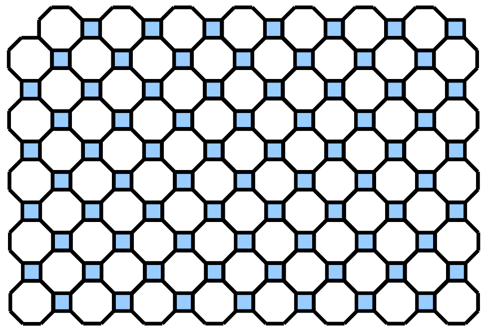

FIGURE 1. A tesselation in $\mathbb{R}^2$ with small squares and large octagons

where the Γ-limit $E(f,\lambda)$ satisfies

$$E(f,\lambda) \ := \ \inf\left\{\liminf_{\varepsilon} \frac{W_p^p(f,\nu_\varepsilon)}{\varepsilon^p} \ , \ \lambda_\varepsilon(\nu_\varepsilon) \to \lambda \ \text{ tightly}\right\} \ .$$

The main part of the paper is devoted to the identification of this Γ-limit $E_\varepsilon(f,\lambda)$ and to give a justification of the fact that the limit problem writes as (1.15). Among applications we discovered the remarkable fact that, for particular entropies h, minimizing sequences may generate microscopic partitions which are not simply uniform tessellations but consist of patterns with different sizes like for instance the one appearing in Figure 1.

The paper is organized as follows: In Section 2, we introduce some basic notations and definitions. In Section 3, we state the main results and their applications to a general class of optimum location problems. Some open problems and conjectures highlighting the role of constants are also given. The proof of Proposition 3.2 related to the case where $f = 1$ and Ω is the unit cube is given in subsequent Section 4. In Section 5, we establish the Γ-convergence of $E_\varepsilon(f,\lambda)$ (Theorem 3.1). Finally in Section 6, we justify the optimality conditions for the limit problem (1.15) (Proposition 3.6) and prove the convergence of infima (Theorem 3.5).

2. Notations and basic definitions

• **Measure theory:** The Lebesgue measure $\mathcal{L}^d(A)$ of a set $A \subset \mathbb{R}^d$ will be denoted shortly $|A|$. If $f \in L^1(\mathbb{R}^d)$, the measure of density f with respect to $\mathcal{L}^d$ will still be denoted f. We will deal with non-negative measures defined on the Borel σ-algebra $\mathcal{B}$ of a locally compact Hausdorff space X. Accordingly we introduce the following notations:

$\mathcal{M}^+(X)$	: the class of non-negative finite Borel measures on X
$\mathcal{M}_0^+(X)$	: the subclass of discrete measures in $\mathcal{M}^+(X)$
$\mathcal{P}(X)$	: the class of probability measures on X
$C_0(X)$	: the space of continuous functions vanishing at the infinity.

The trace of an element $\rho \in \mathcal{M}^+(X)$ on a Borel subset A will be denoted $\rho \llcorner A$. Two measures ρ_1, ρ_2 in $\mathcal{M}^+(X)$ are said to be *orthogonal* if there exists a Borel subset A such that $\rho_1 = \rho_1 \llcorner A$ and $\rho_2 \llcorner A = 0$. We write it $\rho_1 \perp \rho_2$. We will need the following

Definition 2.1. (locality) *A functional $G : \mathcal{P}(X) \to [0, +\infty]$ will be said to be local if whenever $\rho_1 \perp \rho_2$, one has $G(\theta\rho_1 + (1-\theta)\rho_2) = \theta\, G(\rho_1) + (1-\theta)\, G(\rho_2) \quad \forall\theta \in [0,1]$ (or equivalently for every ρ, the map $B \mapsto \rho(B) G\left(\frac{\rho \llcorner B}{\rho(B)}\right)$ is additive on disjoint sets).*

Eventually $\mathcal{M}(X) := \mathcal{M}^+(X) - \mathcal{M}^+(X)$ denotes the space of signed measures on X. Recall that it can be identified with the dual space of $C_0(X)$ and that the associated dual norm coincides with the notion of total variation: $\|\mu\| = |\mu|(X)(= \mu^+(X) + \mu^-(X))$. We will rather use the topology of weak-star (resp. tight) convergence of measures. Recall

Definition 2.2. *A sequence of measures $\mu_n \in \mathcal{M}(X)$ weak-$*$ converges to a measure μ (this is denoted $\mu_n \overset{*}{\rightharpoonup} \mu$) if and only if*

$$\langle \mu_n\,,\,\varphi\rangle \to \langle \mu\,,\,\varphi\rangle \qquad \forall \varphi \in C_0(X). \tag{2.1}$$

A sequence of measures $\mu_n \in \mathcal{M}^+(X)$ such that $\mu_n \overset{*}{\rightharpoonup} \mu$ is said to converge tightly *if in addition $\mu_n(X) \to \mu(X)$. In this case, the convergence in (2.1) can be extended to all φ in $C_b(X)$ the space of continuous and bounded functions.*

The topology of tight convergence is metrizable and Prokhorov's citerium allows to characterize compact subsets of $\mathcal{M}^+(X)$. In this paper, it will be useful to choose a special distance on $\mathcal{P}(\mathbb{R}^+)$: denoting by $\mathrm{Lip}(\varphi)$ the Lipschitz constant of an element $\varphi \in C_b(\mathbb{R}^+)$ and by $|\varphi|_\infty$ its uniform norm, we set for every $\mu \in mathcal P(\mathbb{R}^+)$

$$N(\mu) := \sup\left\{\langle \mu\,,\,\varphi\rangle\ ,\ \|\varphi\|_{BL} \le 1\right\} \quad \text{where} \quad \|\varphi\|_{BL} := |\varphi|_\infty + \mathrm{Lip}(\varphi) \ \ . \tag{2.2}$$

Then $d(\rho_1, \rho_2) := N(\rho_1 - \rho_2)$ is a distance on $\mathcal{P}(\mathbb{R}^+)$ and it is well known that the convergence in this metric is equivalent to the tight convergence of probabilities measures on $\mathbb{R}^+$ (see for instance [16], Theorem 11.3.3).

Given two locally compact spaces X_1, X_2, any measure $\lambda \in \mathcal{M}^+(X_1 \times X_2)$ can be sliced as follows (see for instance [13]).

Proposition 2.3. *(Desintegration of measures) Let $\mu \in \mathcal{M}^+(X_1)$ be the first marginal of λ. Then to μ-almost every $x \in X_1$, we can associate $\rho^x \in \mathcal{P}(X_2)$ such that:*

i) *For all Borel set $B \subset X_2$, the application $x \to \rho^x(B)$ is Borel,*

ii) $\forall \varphi \in \mathcal{C}_0(X_1 \times X_2), \quad \displaystyle\int_{X_1\times X_2} \varphi(x_1, x_2)\; d\lambda(x_1, x_2) = \int_{X_1}\left(\int_{X_2} \varphi(x_1, x_2) d\rho^{x_1}(x_2)\right) d\mu(x_1).$

We will use the notation $\lambda = \mu \otimes \rho^x$ (where ρ^x is an abuse of notation for the application $x \mapsto \rho^x$) and call ρ^x the desintegration of λ with respect to μ.

• **Mass transport, Wasserstein distance:** Recall that, given a Borel map $T : X_1 \mapsto X_2$ and $\mu \in \mathcal{M}^+(X_1)$, the *push forward* of μ under T is the measure $T^\sharp\mu$ defined on X_2 by setting: $T^\sharp\mu(B) := \mu(T^{-1}(B))$ for all Borel subset $B \subset X_2$. Accordingly, if π_i $(i = 1, 2)$ denote the projections from a product space $X_1 \times X_2$ on each X_i, then the marginals of an element $\lambda \in \mathcal{M}^+(X_1 \times X_2)$ are nothing else but the $\pi_i^\sharp\lambda$.

In the following, we will consider $X_1 = X_2 = \mathbb{R}^d$ endowed with a norm denoted by $|\cdot|$. Let us emphasize that, unless specified, this norm *will not be a priori the Euclidean norm.*

Definition 2.4. *Given $p \geq 1$ and a norm $|\cdot|$ on $\mathbb{R}^d$, the p-Wasserstein distance between two measures ν_1, ν_2 in $\mathcal{M}^+(\mathbb{R}^d)$ is given by:*

$$W_p(\nu_1,\nu_2) := \inf_{\gamma\in\mathcal{M}^+(\mathbb{R}^d\times\mathbb{R}^d)} \left\{ \left(\int |x-y|^p d\gamma(x,y)\right)^{1/p} : \pi_1^\sharp\gamma = \nu_1, \pi_2^\sharp\mu = \nu_2 \right\}. \tag{2.3}$$

It is well known that if K is a compact subset of $\mathbb{R}^d$, the topology induced by the Wasserstein distance on $\mathcal{P}(K)$ coincides with the topology of tight convergence (see [20], Remark 7.13 (ii)). Moreover, if ν_1 is non-atomic, it turns out (see for instance Ambrosio [1]) that the above definition is equivalent to the classical Monge transport formulation:

$$W_p(\nu_1,\nu_2) := \inf \left\{ \left(\int |x-Tx|^p d\nu_1(x)\right)^{1/p} \; : \; T^\sharp\nu_1 = \nu_2 \right\}. \tag{2.4}$$

Furthermore, if ν_1 is absolutely continuous with respect to the Lebesgue measure, then there always exists an optimal transport map T for $p \geq 1$ which is is unique for $p > 1$. This very difficult result is known since the end of the 90' when $|\cdot|$ is the Euclidian norm (see [20, 1] and the references therein), whereas it has been shown only very recently for a general norm (see [14], [11]). However, when ν_2 is an atomic measure, the existence of an optimal transport map can be easily proved and uniqueness holds even when $p = 1$.
Let us finally notice that if $L : \mathbb{R}^d \mapsto \mathbb{R}^d$ is an isometry with respect to the given norm $|\cdot|$ (i.e. an affine transform such that $|L(x)| = |x|$), then

$$W_p(L^\sharp\nu_1, L^\sharp\nu_2) \;=\; W_p(\nu_1,\nu_2) \;. \tag{2.5}$$

We will often use the following subadditivity property: if $\{A_i\}$ is a finite or countable family of disjoint Borel subsets $\nu_1(\mathbb{R}^d \setminus \cup_i A_i) = \nu_2(\mathbb{R}^d \setminus \cup_i A_i) = 0$, then

$$W_p^p(\nu_1,\nu_2) \quad \leq \quad \sum_i W_p^p(\nu_1 \llcorner A_i, \nu_2 \llcorner A_i) \;, \tag{2.6}$$

with the convention that $W_p(\nu_1 \llcorner A_i, \nu_2 \llcorner A_i) = +\infty$ if $\nu_1(A_i) \neq \nu_2(A_i)$.
Besides, if L_t is the homothety $x \mapsto tx$ (with $t > 0$), then

$$W_p(L_t^\sharp\nu_1, L_t^\sharp\nu_2) \;=\; t\, W_p(\nu_1,\nu_2) \;. \tag{2.7}$$

In order to lighten the notations, the dilated measure $L_t^\sharp\nu$ will sometimes be denoted $\frac{1}{t^d}\nu(\frac{\cdot}{t})$.
Eventually we will use the following classical upperbound for the Wasserstein distance between compactly supported measures (see [20], Proposition 7.10, p. 211):

$$W_p^p(\nu_1,\nu_2) \;\leq\; \frac{(\text{diam}K)^p}{2} \int |\nu_1 - \nu_2| \;, \quad \text{if } \nu_1,\nu_2 \text{ are supported in } K. \tag{2.8}$$

• **Shape function :** We introduce the following shape function (counterpart of (1.9) for $f = 1$):

$$\psi_p^\Omega(A) = \inf_{y\in\Omega} \left\{ \int_A |x-y|^p \, dx \right\} \quad , \quad A \subset \Omega \subset \mathbb{R}^d.$$

The superscript Ω will be omitted if $\Omega = \mathbb{R}^d$. Obviously $\psi_p^\Omega(A) \leq \psi_p^\Omega(B)$ if $A \subset B$. Moreover $\psi_p^\Omega(A)$ does not depend on Ω whenever Ω is convex, that is

$$\psi_p^\Omega(A) \;=\; \psi_p^{\mathbb{R}^d}(A) := \; \psi_p(A). \tag{2.9}$$

The quantity $\psi_p(A)$ represents the p^{th} moment of A with respect to its p-barycenter and it enjoys the following invariance relations, for every $x \in \mathbb{R}^d$ and $t > 0$:

$$\Psi_p(A + x) \ = \ \Psi_p(A) \quad , \quad \Psi_p(tA) \ = \ t^{p+d}\Psi_p(A) \tag{2.10}$$

In view of finding a lowerbound for the minimal partition problem (1.10), it is very useful to introduce the constant

$$\gamma_{p,d} \ := \ \inf\Big\{\psi_p(A) : A \subset \mathbb{R}^d \text{ is measurable and } |A| = 1\Big\} . \tag{2.11}$$

It is clear that the infimum is attained for a ball (with respect to the given norm). Indeed:

Lemma 2.5. *Let B denote the ball centered at the origin with unit Lebesgue measure. Then, for any measurable subset A of $\mathbb{R}^d$, we have:*

$$\psi_p(A) \ \geq \ \gamma_{p,d} \ |A|^{1+\frac{p}{d}} \quad , \quad \gamma_{p,d} = \int_B |x|^p \ dx \ . \tag{2.12}$$

Proof. By (2.10), we are reduced to prove that for every measurable set A with $|A| = 1$, we have $\int_A |x|^p \ dx \geq \ \int_B |x|^p \ dx$. We observe that $|A \cap B^c| = 1 - |A \cap B| = |B \cap A^c|$. Therefore there exists a Lebesgue measure preserving map $T : B \cap A^c \mapsto A \cap B^c$ which obviously satisfies $|Tx| \geq 1 \geq |x|$ a.e. The thesis follows since

$$\int_{A\cap B^c} |x|^p \ dx \ = \ \int_{B\cap A^c} |Tx|^p \ dx \ \geq \ \int_{B\cap A^c} |x|^p \ dx \ .$$

□

• **Tilings :** Unfortunately the optimal ball B found in Lemma 1.10 is not suitable for constructing good competiting partitions for (1.9) since in general $\mathbb{R}^d$ (or the reference domain Ω) cannot be tiled with isometric copies of B. More precisely, having in mind the invariance property (2.5) of the Wasserstein distance, let us associate with any compact subset P of $\mathbb{R}^d$ its orbit $\tilde{P}$ through the group of isometries (with respect to $|\cdot|$) and set:

$$\theta_k(P) \ := \ \sup\left\{\frac{1}{k^d}\sum_{i=1}^{\infty}|A_i| \ : \ A_i \in \tilde{P} \ , A_i \subset \left[-\frac{k}{2}, \frac{k}{2}\right]^d \ , |A_i \cap A_j| = 0 \text{ if } i \neq j\right\} \ . \tag{2.13}$$

Then we may define an *asymptotic tiling ratio of P* as follows:

$$\theta(P) \ := \ \sup\big\{\theta_k(P) \ : \ k \in \mathbb{N}\big\} \ = \ \lim_{k\to\infty} \theta_k(P) \ , \tag{2.14}$$

where the existence of the limit in the right hand side can be deduced by applying Lemma 2.8 to the subadditive set function $S(A) := \inf\Big\{|A \setminus \cup_i A_i| \ : \ A_i \in \tilde{P} \ , A_i \subset A \ , |A_i \cap A_j| = 0 \text{ if } i \neq j\Big\}$.

Notice that $\theta(P)$ is invariant under dilatation. We will say that *a norm on $\mathbb{R}^d$ has the tiling property* if the unit ball B satisfies $\theta(B) = 1$. For instance, the l^1 or the l^∞ norm on $\mathbb{R}^d$ have the tiling property whereas the Euclidean ball satisfies $\theta(B) = \frac{\pi}{4} < 1$ for $d = 2$ and more generally $\theta(B) = \frac{\omega_d}{2^d} < 1$ for $d > 1$. When $\theta(B) < 1$, another constant will be important in the upperbound estimates, namely

$$\Gamma_{p,d} \ := \ \inf\Big\{\psi_p(P) \ : \ P \text{ compact} \ , \ \theta(P) = 1 \ , \ |P| = 1\Big\} \ . \tag{2.15}$$

It is clear that $\gamma_{p,d} \leq \Gamma_{p,d} \leq \psi_p(]0, 1[^d) < +\infty$. On the other hand $\gamma_{p,d} = \Gamma_{p,d}$ if the norm enjoys the tiling property. The problem of the existence of an optimal compact P in (2.15) is open. It seems reasonable to conjecture that in the general case the infimum is reached for a convex compact subset, which in the case of a finite group of isometries, should be polyhedral. In [4], it is proved that for $p = 1$ and the Euclidean norm on $\mathbb{R}^2$, the infimum among convex polytopes is reached when P is a regular hexagon.

• **Scaling and local statistics :** Given $\nu = \sum_i c_i \delta_{x_i}$ in $\mathcal{M}_0^+(A)$ with total mass $m = \sum_i c_i$, we associate the element of $\mathcal{M}_0^+(\mathbb{R}^+)$ defined by

$$\rho(\nu) := \sum_i \frac{c_i}{m} \delta_{c_i} \ . \tag{2.16}$$

This probability measure $\rho(\nu)$ represents the statistical distribution of all masses carried by ν.

Now in view of the rescaling argument described in Section 1, for every value of the small parameter ε, we associate with ν the measure $\lambda_\varepsilon(\nu) \in \mathcal{M}_0^+(A \times \mathbb{R}^+)$ defined by

$$\lambda_\varepsilon(\nu) \ := \ \sum_i c_i \, \delta_{(x_i, \frac{c_i}{\varepsilon^d})} \ . \tag{2.17}$$

The second marginal of $\lambda_\varepsilon(\nu)$ produces after normalization the ε-rescaled counterpart of (2.16):

$$\rho_\varepsilon(\nu) \ := \ \frac{1}{\varepsilon^d} \rho(\nu)(\frac{\cdot}{\varepsilon^d}) \ = \ \sum_i \frac{c_i}{m} \, \delta_{\frac{c_i}{\varepsilon^d}} \ . \tag{2.18}$$

• **Γ- convergence :** For the convenience of the reader, we recall the definition and the main property of the Γ- convergence. Further details can be found for instance in [2, 15].

Definition 2.6. *(sequential Γ-convergence) A family of extended real valued functions F_ε defined on a topological space X is said to be sequentially Γ-convergent to a functional F if the two following statements hold:*
*(i) (*lower bound*) for every sequence $\{\lambda_\varepsilon\}$ converging to $\lambda \in X$, we have:*

$$\liminf_{\varepsilon \to 0} F_\varepsilon(\lambda_\varepsilon) \geq F(\lambda). \tag{2.19}$$

*(ii) (*recovering sequence*) for every $\lambda \in X$ there exists a sequence $\{\lambda_\varepsilon\}$ converging to λ such that*

$$\limsup_{\varepsilon \to 0} F_\varepsilon(\lambda_\varepsilon) \leq F(\lambda). \tag{2.20}$$

When properties i) and ii) are satisfied, we write $F = \Gamma - \lim\limits_{\varepsilon \to 0} F_\varepsilon$.

Proposition 2.7. *Let $F_\varepsilon : X \to]-\infty, +\infty]$ be a sequence of functionals such that:*
(i) $F = \Gamma - \lim\limits_{\varepsilon \to 0} F_\varepsilon$,
(ii) $\sup_\varepsilon F_\varepsilon(\lambda_\varepsilon) < +\infty \ \Rightarrow \ \{\lambda_\varepsilon\}$ is sequentially relatively compact.
Then we have the convergence $\inf F_\varepsilon \to \inf F$ *and every cluster point of a minimizing sequence $\{\lambda_\varepsilon\}$ (i.e. such that $F_\varepsilon(\lambda_\varepsilon) - \inf F_\varepsilon \to 0$) achieves the minimum of F.*

• **Ergodicity :** We will need the following result related to subadditive processes which is often used in ergodic theory (cf. [18] , [3]).

Lemma 2.8. *Let $S : \mathcal{B} \to \mathbb{R}^+$ be a function on Borel subsets of $\mathbb{R}^d$ such that:*
*- (*subadditivity*) $S(A \cup B) \leq S(A) + S(B)$ for all A, B in $\mathcal{B}$ such that $A \cap B = \emptyset$*
*- (*translation invariance*) $S(x + A) = S(A)$ for all A in $\mathcal{B}$ and x in $\mathbb{R}^d$.*

Then, setting $Q_k :=]-\frac{k}{2}, \frac{k}{2}[^d$, we have

$$\liminf_{k \to \infty} \frac{S(Q_k)}{k^d} \ = \ \limsup_{k \to \infty} \frac{S(Q_k)}{k^d} \ = \ \inf_{k > 0} \frac{S(Q_k)}{k^d} \ .$$

3. Main results and applications.

From now on we will assume for simplicity that the reference domain Ω is a *compact* subset of $\mathbb{R}^d$. It is endowed with a norm $|\cdot|$ (which is not necessarily the Euclidian norm). We fix a positive density f in $L^1(\Omega)$ and we assume from now on that

$$f \text{ is lower semicontinuous on } \Omega \text{ and } f \geq \alpha \text{ for a suitable constant } \alpha > 0. \tag{3.1}$$

For every value of the parameter $p \geq 1$, we look at the asymptotic limit of the ε-normalized error given by the expression (1.13). To that aim we are going to construct a variational problem involving the limit in the sense of $\Gamma-$convergence of the sequence of functionals $\{E_\varepsilon(f,\cdot)\}$ defined on $\mathcal{M}^+(\Omega\times\mathbb{R}^+)$ by:

$$E_\varepsilon(f,\lambda) = \begin{cases} \dfrac{1}{\varepsilon^p} W_p^p(f,\nu) & \text{if } \lambda = \lambda_\varepsilon(\nu)\ ,\ \nu \in \mathcal{M}_0^+(\Omega) \\ +\infty \text{ otherwise.} & \end{cases} \tag{3.2}$$

being $\lambda_\varepsilon(\nu)$ given by (2.17). It is important to notice that if $\{\lambda_\varepsilon\}$ is a sequence of $\mathcal{M}^+(\Omega\times\mathbb{R}^+)$ such that

$$\lambda_\varepsilon = \lambda_\varepsilon(\nu_\varepsilon), \quad \nu_\varepsilon \in \mathcal{M}_0^+(\Omega), \quad \sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon) < +\infty,$$

then $W_p^p(f,\nu_\varepsilon) \to 0$ yielding that $\{\nu_\varepsilon\}$ converges weakly-$*$ to f as $\varepsilon \to 0$. Moreover, as will be shown in assertion i) of Theorem 3.1, $\{\lambda_\varepsilon\}$ is tight and so every weak-$*$ cluster point λ of $\{\lambda_\varepsilon\}$ admits f as a first marginal. Therefore such a λ will be of the form $\lambda = f\otimes\rho^x$ being $\{\rho^x\}$ the parametrized family of probabilities defined in Proposition 2.3. It turns out that this family $\{\rho^x\}$ determines completely the Γ-limit $E(f,\lambda)$ of $\{E_\varepsilon(f,\cdot)\}$ with respect to the tight convergence and will play the role of the unknown in our final optimization problem.

3.1. Identification of the Γ- limit.

We begin by constructing a functional $G:\mathcal{P}(\mathbb{R}^+)\mapsto\mathbb{R}$ which will be shown to coincide with the Γ- limit of $E_\varepsilon(f,\lambda)$ in the particular case where $\lambda = 1\!\!1_Q\otimes\rho$ that is $f = 1\!\!1_Q$ with $Q = [-1/2,1/2]^d$ and $\rho^x = \rho$ does not depend on x.

Given any small $\delta > 0$, using the notation in (2.16), we introduce the set function

$$S_\delta(\rho,A) := \inf\left\{ W_p^p(1\!\!1_A,\nu) + \frac{|A|}{\delta^p} N(\rho-\rho(\nu)) : \nu\in\mathcal{M}_0^+(A)\right\}, \tag{3.3}$$

where A is a Borel subset of $\mathbb{R}^d$. Then denoting by Q_k the cube $[-k/2,k/2]^d$, we define a functional on $\mathcal{P}(\mathbb{R}^+)$ by setting

$$G_\delta(\rho) := \inf_{k>0} \frac{S_\delta(\rho,Q_k)}{|Q_k|}. \tag{3.4}$$

After noticing that G_δ is monotone with respect to δ, we set

$$G(\rho) \ := \ \sup_{\delta>0} G_\delta(\rho) \ = \lim_{\delta\to 0} G_\delta(\rho). \tag{3.5}$$

We are now able to state our main asymptotic result. The proof appears in Section 4.

Theorem 3.1. *Let f satisfy (3.1) and let us denote by τ the topology of the tight convergence on $\mathcal{M}^+(\Omega\times\mathbb{R}^+)$. Then:*
i) Any sequence $\{\lambda_\varepsilon\}$ such that $\sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon) < +\infty$ is τ-relatively compact,
ii) The τ-sequential Γ-limit of the sequence of functionals $\{E_\varepsilon(f,\cdot)\}$ exists as ε tend to 0 and we have:

$$\Gamma - \lim_{\varepsilon\to 0} E_\varepsilon(f,\lambda) = E(f,\lambda)$$

where:

$$E(f,\lambda) = \begin{cases} \int_\Omega f(x)^{1-\frac{p}{d}}\, G(\rho^x)\ dx & \text{if } (\pi_1)^\sharp \lambda = f\,dx \quad (\lambda = f\otimes \rho^x)\ , \\ +\infty \ \text{otherwise.} \end{cases} \tag{3.6}$$

The main properties of the functional $G : \mathcal{P}(\mathbb{R}^+) \to [0,+\infty]$ are summarized in the following proposition whose proof is given in Section 4. The constants $\gamma_{p,d}$ and $\Gamma_{p,d}$ appeared respectively in (2.11) and (2.15). The space $\mathcal{P}(\mathbb{R}^+)$ is embedded with the topology of tight convergence and can be metrized by using the distance $d(\rho,\rho') = N(\rho-\rho')$.

Proposition 3.2. *i) The functional G is convex and lower semicontinuous.*

*ii) (*Scale invariance*) For every $a>0$ and $\rho\in\mathcal{P}(\mathbb{R}^+)$, there holds:*

$$G\left(L_a^\sharp(\rho)\right) \ = \ a^{\frac{p}{d}}\, G(\rho)\ . \tag{3.7}$$

iii) Let P be a compact subset of $\mathbb{R}^d$ such that $|P|=1$. Then we have

$$G(\theta\delta_a + (1-\theta)\delta_0) \ \le\ \theta\,\psi_p(P)\,a^{p/d} \quad \text{for every } 0\le\theta\le\theta(P)\ . \tag{3.8}$$

Furthermore, if P is a ball, then $\psi_p(P)=\gamma_{p,d}$, we obtain the equality

$$G(\theta\delta_a + (1-\theta)\delta_0) \ =\ \gamma_{p,d}\ \theta\, a^{p/d} \quad \text{for every } 0\le\theta\le\theta(B)\ . \tag{3.9}$$

iv) Set $\omega_{p,d} := G(\delta_1)$. Then $\gamma_{p,d}\le\omega_{p,d}\le\Gamma_{p,d}$, and for every $\rho\in\mathcal{P}(\mathbb{R}^+)$, there holds:

$$\gamma_{p,d}\int t^{p/d}\,\rho(dt) \ \le\ G(\rho) \ \le\ \omega_{p,d}\int t^{p/d}\,\rho(dt)\ . \tag{3.10}$$

*v) (*Non locality*) The locality property (Definition 2.1) holds for G if and only if $\gamma_{p,d}=\omega_{p,d}$. In particular it is the case when $\theta(B)=1$.*

Remark 3.3. a) By Theorem 3.1, it turns out that $G(\rho)$ coincides with $E(\mathbb{1}_Q, \mathbb{1}_Q\otimes\rho)$. Thus the definition of G given in (3.5) does not depend on the choice of the norm N which was used merely to define the δ-regularized functional G_δ.
b) Unfortunately we are not able to give an explicit representation of the non local functional G unless the norm satisfies the tiling property. This is a quite challenging issue and several related questions seem to be deep and difficult open problems. Some of them are mentioned in Section 3.6.
c) It is an easy consequence of the first inequality in(3.10) that the level sets of G are compact for the tight convergence. On the other hand, by the second inequality and assertion v), we have the explicit form $G(\rho)=\gamma_{p,d}\int t^{p/d}\rho(dt)$ whenever G is local.

3.2. **Convergence of the constrained optimization problems.** The main application of Theorem 3.1 consists in passing to the limit as $\varepsilon\to 0$ in constrained problems of the kind (1.13). Following the notations introduced in Section 2, namely (2.17), these problems can be written in the following form

$$(\mathcal{P}_\varepsilon^c) \qquad \inf\left\{E_\varepsilon(f,\lambda) : \int g(t)\ d\lambda(x,t)\le c\right\}$$

and accordingly the expected limit problem reads

$$(\mathcal{P}^c) \qquad \inf\left\{E(f,\lambda) : \int g(t)\ d\lambda(x,t)\le c\right\}.$$

In order to establish the convergence of $(\mathcal{P}^c_\varepsilon)$ to $(\mathcal{P}^c)$, we need to add some technical assumptions on the size function $g\,:]0,+\infty[\to\mathbb{R}$. However these conditions have to be compatible with the simple cases considered in the introduction (namely when g blows up at 0). We will assume that g can be decomposed as

$$g(s) := \ \beta(s) \ + \ V(s) \ , \tag{3.11}$$

where V is a possibly unbounded potential $V\,:]0,+\infty]\to[0,+\infty]$ such that

$$V \textit{ is lower semicontinuous, proper and non increasing} \ , \tag{3.12}$$

and β is a *continuous function* on $\mathbb{R}^+$ satisfying one of the conditions a) or b) below

$$\begin{array}{llll} a) & \displaystyle\lim_{t\to+\infty} \frac{|\beta(t)|}{t^{p/d}} \ = \ 0 & \textit{if } f\in L^\infty(\Omega), \\ b) & \sup|\beta| \ < \ +\infty & \textit{if } f\notin L^\infty(\Omega). \end{array} \tag{3.13}$$

We set

$$\bar{c} \ := \ \left(\int_\Omega f\,dx\right)\ \inf_{\mathbb{R}^+} g \ , \tag{3.14}$$

and we define the infimum value function

$$m(c) \ := \ \inf(\mathcal{P}^c) \ .$$

Under the assumptions above, we have

Lemma 3.4. *The function $m(c)$ is convex and non -increasing increasing. It is finite and continuous on $]\bar{c},\infty[$ and for every $c>\bar{c}$ the problem $(\mathcal{P}^c)$ admits at least one solution.*

Proof. : The function $m(c)$ is clearly non-increasing and inherits its convexity property from that of $E(f,\cdot)$. On the other hand, if $c>\bar{c}$, there exits by (3.14) a real s_0 such that $g(s_0)\int_\Omega f\,dx\le c$. Then $\lambda_0 := f\otimes\delta_{s_0}$ satisfies the constraint and has finite energy $E(f,\lambda_0)=\omega_{p,d}\,s_0^{p/d}\,\int_\Omega f^{1-p/d}\,dx$, since $f\ge\alpha>0$. Therefore $m(c)$ is finite and continuous on $]\bar{c},+\infty[$.

Let us show that the infimum is achieved if $m(c)$ is finite. Let $\{\lambda_n\}$ be a sequence such that $E(f,\lambda_n)<m(c)+1/n$ and $\int g(t)\,d\lambda_n\le c$. Assume first that $\{\lambda_n\}$ is tight so that, up to a subsequence, it converges to some λ. Then, by the lower semicontinuity of $E(f,\cdot)$, we infer that $m(c)\ge\liminf_n E(f,\lambda_n)\ge E(f,\lambda)$. In addition, thanks to the assumption (3.13) on g, we have $\int g(t)\,d\lambda\le c$ (see the complete argument in step 1 of the proof of Theorem 3.5). Therefore $E(f,\lambda)\ge m(c)$ and we have found an admissible λ such that $E(f,\lambda)=m(c)$. To show the tightness property, we use (3.10) from which we deduce that for every $M>0$, $E(f,\lambda_n)\ \ge\ M^{-\frac{p}{d}}\,\gamma_{p,d}\,\int_{\Omega^M\times\mathbb{R}^+} t^{p/d}\,d\lambda_n$ being $\Omega^M := \{x\in\Omega\ :\ f(x)\le M\}$. Then, for every $k>0$, there holds

$$\lambda_n(\Omega\times[k,+\infty[)\ \le\ \frac{1}{k^{\frac{p}{d}}}\int_{\Omega^M\times\mathbb{R}^+} t^{p/d}\,d\lambda_n+\int_{\Omega\setminus\Omega^M} f\,dx\ \le\ \frac{M^{\frac{p}{d}}}{\gamma_{p,d}\,k^{\frac{p}{d}}}\left(m(c)+\frac{1}{n}\right)+\int_{\Omega\setminus\Omega^M} f\,dx.$$

Thus $\displaystyle\limsup_{k\to+\infty}\sup_n\{\lambda_n(\Omega\times[k,+\infty[)\}\le\int_{\Omega\setminus\Omega^M} f\,dx$ which vanishes as $M\to\infty$. □

Theorem 3.5. *Let f satisfy (3.1) and let g be of the form (3.11) satisfy (3.12)(3.13). Then for every $c>\bar{c}$, there holds*

$$\lim_{\varepsilon\to0}\ \inf\mathcal{P}^c_\varepsilon\ =\ \inf\mathcal{P}^c\ =\ \min\mathcal{P}^c.$$

Futhermore every minimizing sequence $\lambda_\varepsilon(=\lambda_\varepsilon(\nu_\varepsilon))$ for $(\mathcal{P}^c_\varepsilon)$ converges tightly, possibly after extracting a subsequence, to a limit of the form $\lambda=f\otimes\rho^x$ where λ is a minimizer for $(\mathcal{P}^c)$.

Let us emphasize that this result is not a straightforward consequence of Theorem 3.1. Indeed in many cases, the size function $g(t)$ we have in mind is *unbounded* (in particular near 0) and it is very tricky to adapt the recovering sequences $\{\lambda_\varepsilon\}$ so that they satisfy $\int g(t)\ d\lambda_\varepsilon(x,t) \le c$. In fact it seems difficult to weaken the assumption (3.13). In particular in the case b), the boundedness assumption on f cannot be omitted as it is shown in the example of Remark 3.13. Let us also notice that in the statement of the theorem, we do not need the existence of a minimizer for $(\mathcal{P}^c_\varepsilon)$.

3.3. **Reformulation of the limit problem and optimality conditions.** Let $\lambda = f \otimes \rho^x$ a solution of $(\mathcal{P}^c)$. The related optimal limit configuration is described by a local statistic ρ^x we want to characterize. To that aim we associate with g the infimum value function

$$\Phi_g : s \in \mathbb{R} \ \mapsto\ \min\left\{ G(\rho) :\ \rho \in \mathcal{P}(\mathbb{R}^+)\ ;\ \int g(t)\ d\rho(t)\ (= \langle \rho, g\rangle)\ \le s \right\}. \tag{3.15}$$

Noticing that the constraint in $(\mathcal{P}^c)$ can be simply recast as

$$\int_\Omega u(x)\ dx \le c \quad \text{where} \quad u(x) := f(x)\,\langle \rho^x, g\rangle\ ,$$

we are going to reduce our limit problem $(\mathcal{P}^c)$ to the simpler one

$$(\mathcal{Q}^c) \qquad \inf_{u \in \mathrm{L}^1(\Omega)} \left\{ \int_\Omega \Phi_g\left(\frac{u(x)}{f(x)}\right) f(x)^{1-\frac{p}{d}}\, dx \ :\ \int_\Omega u\, dx \le c \right\}\ .$$

Proposition 3.6. *Under the assumptions of Theorem 3.5, the problem $(\mathcal{Q}^c)$ admits solutions and for every $c > \bar{c}$, there holds $m(c) = \min \mathcal{Q}^c = \min \mathcal{P}^c$.*
Futhermore $\lambda = f \otimes \rho^x$ is minimal with respect to $(\mathcal{P}^c)$ if and only if there exists u optimal for $(\mathcal{Q}^c)$ such that :

$$G(\rho^x)\ =\ \Phi_g\left(\frac{u(x)}{f(x)}\right) \quad , \quad \langle \rho^x, g\rangle\ = \frac{u(x)}{f(x)} \qquad a.e.\ x \in \Omega\ . \tag{3.16}$$

This result will be proved in Section 6. We will need before some properties of the integrand Φ_g defined in (3.15) and to that aim introduce the comparison function:

$$\varphi_g(s)\ :=\ \inf\left\{ t^{p/d}\ :\ t > 0\ ,\ g(t)\ \le\ s \right\} \tag{3.17}$$

(with the convention that $\varphi_g(s) = +\infty$ if $g(t) > s$ for all t).
The following result is proved in Section 4.

Lemma 3.7. *The function $\Phi_g(s)$ is convex right-continuous monotone non increasing and the infimum in (3.15) is achieved whenever it is finite (in particular for $s > \inf g$). Furthermore, for every s, one has*

$$\gamma_{p,d}\ \varphi_g^{**}(s)\ \le\ \Phi_g(s)\ \le\ \omega_{p,d}\ \varphi_g^{**}(s)\ , \tag{3.18}$$

where φ_g^{**} denotes the convex envelope of φ_g. In particular, there holds $\lim\limits_{s\to\infty} \Phi_g(s) = 0$ if g is finite on $]0, +\infty[$.

Remark 3.8. The computation of the minimum value $m(c)$ is very easy once the convex integrand Φ_g has been computed and this will be illustrated in Subsection 3.5 . Since Φ_g is non-increasing it is clear that the minimum is reached for u such that $\int_\Omega u dx = c$. On the other hand, if one wishes to treat this integral constraint by means of a Lagrange multiplier, the following explicit formula can be derived by a straightforward computation (Fenchel conjugate of integral functionals in $L^1(\Omega)$).

Denote by $\Phi_g^* : \mathbb{R} \to]-\infty, +\infty]$ the Fenchel conjugate of Φ_g. Then, the Fenchel conjugate of $m(c)$ is finite only on $\mathbb{R}^-$ and is given at $-\eta$ for every $\eta > 0$ by the following formula:

$$m^*(-\eta) \;=\; -\inf_{u\in \mathrm{L}^1(\Omega)} \left\{ \int_\Omega \Phi_g\left(\frac{u}{f}\right) f^{1-\frac{p}{d}} \;+\; \eta \int_\Omega u\,dx \right\} \;=\; \int_\Omega \Phi_g^*\Big(\frac{-\eta}{f^{-\frac{p}{d}}}\Big)\, f^{1-\frac{p}{d}}\,dx \;. \tag{3.19}$$

3.4. **Optimal partitions may exhibit patterns with different sizes.** In view of Proposition 3.6, the function Φ_g introduced in (3.15) plays a central role in the characterization of the optimal asymptotic statistics represented by $\{\rho^x\}$. Finding for instance ρ^x to be a Dirac mass a.e. on a subset $A \subset \Omega$ means that locally on A the minimizing partitions exhibit patterns with only a single size. This will be typically the case when the disposition of the small subsets of the partition becomes almost periodic (or quasiperiodic). From Lemma 3.7 it is easy to deduce the following criterium

Lemma 3.9. *Let $s > \inf g$. Then the minimal value $\Phi_g(s)$ in (3.15) is achieved for a single Dirac mass if and only if $\Phi_g(s) = \omega_{p,d}\, \varphi_g(s)$. In this case, there holds $\varphi_g^{**}(s) = \varphi_g(s)$.*

Proof. It is clear that if the infimum is reached for a Dirac mass, then $\Phi_g(s) = \omega_{p,d}\,\varphi_g(s)$ and the second inequality in (3.18) leads to $\varphi_g(s) = \varphi_g^{**}(s)$. Conversely if $\Phi_g(s) = \omega_{p,d}\,\varphi_g(s)$, it is optimal to use the Dirac mass at $\bar{t} = \min\{t \geq 0 : g(t) \leq s\}$. □

A direct consequence of Lemma 3.9 is that in general the optimal ρ^x is not a Dirac mass. Indeed assume that $p \geq d$ and consider the non-increasing function g defined by

$$g(t) = 2 \quad \text{if } t < 1 \quad , \quad g(t) = 1 \quad \text{if } 1 \leq t < 2 \quad , \quad g(t) = 0 \quad \text{if } t \geq 2 \;.$$

It satisfies (3.11) and (3.12) with $\beta = 0, V = g$. Then φ_g and its convexification (extended by the value $+\infty$ for $t < 0$) are given by

$$\varphi_g(s) \;=\; \begin{cases} 2^{p/d} & \text{if } 0 \leq s < 1 \\ 1 & \text{if } 1 \leq s < 2 \\ 0 & \text{if } s \geq 2 \end{cases} \quad , \quad \varphi_g^{**}(s) \;=\; \begin{cases} 2^{p/d}(1-s) + s & \text{if } 0 \leq s \leq 1 \\ 2 - s & \text{if } 1 \leq s < 2 \\ 0 & \text{if } s > 2 \end{cases}$$

Assume for simplicity that $f = 1$ on Ω. Then the constant function $u(x) = s_0$ where $s_0 = \frac{c}{|\Omega|}$ solves problem $(\mathcal{Q}^c)$. Clearly if c is given so that s_0 ranges in $]0, 1[$ where $\varphi_g^{**} < \varphi_g$, the minimal value $\Phi_g(s_0) = G(\rho^x)$ can never be reached for ρ^x a Dirac mass. We can be even more precise in the case where $\omega_{p,d}$ agrees with $\gamma_{p,d}$ (for instance if the norm has the tiling property): in this case by (3.18) and since $s_0 \in]0, 1[$, $\Phi_g(s_0) = \gamma_{p,d}(2^{p/d}(1 - s_0) + s_0)$ and the optimal ρ^x is unique and given by $\rho^x = s_0\delta_1 + (1-s_0)\delta_2$. This serves as evidence to conclude that, asymptotically as ε becomes small, the optimal partitions associated with problem $(\mathcal{P}_\varepsilon^c)$ should exhibit patterns made of two kind of sets like in Figure 1.

3.5. **Applications to some particular entropies.** We may now apply Theorem 3.5 in order to find the asymptotic behaviour as $n \to \infty$ of the expression

$$\mathcal{E}_H^{p,d}(f, \Omega, n) \;=\; \inf\{W_p^p(f, \nu) : H(\nu) \leq n\}$$

defined in the introduction, for the classical entropies $H(\sum_i c_i\delta_{x_i}) = \sum_i h(c_i)$ associated with the size constraints (1.2) or (1.3). The associated functions g are given by $g(t) = t^{q-1}$ for $q < 1$, $g(t) = -\ln t$ for $q = 1$ and $g(t) = -t^{q-1}$ for $q > 1$. Accordingly we introduce, the following universal constants (recall that d is the dimension of the ambient space and $p \geq 1$):

$$C_{p,d}(q) \ := \ \begin{cases} \inf\left\{G(\rho) \ : \ \ \rho \in \mathcal{P}(\mathbb{R}^+) \ , \ \displaystyle\int_{\mathbb{R}^+} t^{q-1}\, d\rho \leq 1\right\} & \text{if } q < 1, \\ \inf\left\{G(\rho) \ : \ \ \rho \in \mathcal{P}(\mathbb{R}^+) \ , \ \displaystyle\int_{\mathbb{R}^+} t^{q-1}\, d\rho \geq 1\right\} & \text{if } q > 1. \end{cases} \tag{3.20}$$

Rewriting the constraint (3.20) in the condensed form $\left(\int_{\mathbb{R}^+} t^{q-1}\, d\rho\right)^{\frac{1}{q-1}} \geq 1$, we find a monotone dependence with respect to q and by passing to the limit as $q \to 1$ or as $q \to -\infty$ (for fixed ρ), we are led to extend the definition to the case $q = 1$ and $q = -\infty$ as follows

$$C_{p,d}(1) := \inf\left\{G(\rho) : \int_{\mathbb{R}^+} \ln t \ d\rho \geq 0\right\} \quad , \quad C_{p,d}(-\infty) := \inf\Big\{G(\rho) : \mathrm{spt}\rho \subset [1,\infty)\Big\}. \tag{3.21}$$

Lemma 3.10. *The extended function $q \in [-\infty, +\infty] \mapsto C_{p,d}(q)$ is monotone non-increasing, continuous on $[-\infty, 1 + \frac{p}{d}[$. The infima is attained in (3.21) and in (3.20) for every $q \leq 1 + p/d$. Furthermore there holds*

$$C_{p,d}(-\infty) \leq \omega_{p,d} \quad , \quad C_{p,d}(1 + \frac{p}{d}) = \gamma_{p,d} \quad , \quad C_{p,d}(q) = 0 \ \textit{whenever}\ q > 1 + \frac{p}{d} \ . \tag{3.22}$$

Proposition 3.11. *Let for every $q \neq 1$, $r = r(p,d,q) := \frac{p}{d(1-q)}$. Then*

(i) For $H(\nu) := \sum_x (\nu(\{x\})^q$ and $-\infty < q < 1$, we have

$$\lim_{n\to\infty} n^{\frac{p}{d(1-q)}}\, \mathcal{E}_H^{p,d}(f,\Omega,n) = \ C_{p,d}(q) \left(\int_\Omega f^{\frac{1+qr}{1+r}}\, dx\right)^{1+r} \tag{3.23}$$

(ii) For $H(\nu) := \sup_x \left\{(\nu(\{x\})^{-1}\right\}$, we have

$$\lim_{n\to\infty} n^{p/d}\, \mathcal{E}_H^{p,d}(f,\Omega,n) = \ C_{p,d}(-\infty) \int_\Omega f^{1-\frac{p}{d}}\, dx \tag{3.24}$$

(iii) For $H(\nu) := -\sum_x (\nu(\{x\})^q$ with $q \in (1, 1+p/d)$ and assuming that $f \in L^\infty(\Omega)$, we have

$$\lim_{n\to\infty} n^{\frac{p}{d(q-1)}}\, \mathcal{E}_H^{p,d}(f,\Omega,-\frac{1}{n}) \ = C_{p,d}(q) \left(\int_\Omega f^{\frac{1+qr}{1+r}}\, dx\right)^{1+r} \tag{3.25}$$

(iv) For $H(\nu) := \sum_x (-\nu(\{x\}) \ln(\nu(\{x\})))$, $f \in L^\infty(\Omega)$ and setting $I(f) := \int_\Omega f(x)\, dx$, we have

$$\lim_{n\to\infty} exp\left(\frac{p\, n}{d\ I(f)}\right) \mathcal{E}_H^{p,d}(f,\Omega,n) \ = \ C_{p,d}(1)\ I(f)\ exp\left(-\frac{p}{d}\frac{\int_\Omega (f \ln f)\, dx}{I(f)}\right) \tag{3.26}$$

Remark 3.12. By using (1.10) and the shape function (2.9), we can give for $f = \mathbb{1}_Q$ an equivalent statement covering both cases $q > 1$ and $q < 1$:

$$C_{p,d}(q) = \lim_{\delta\to 0} \delta^{-\frac{p}{d}} \ \min_{\mathcal{P}} \left\{\sum_i \psi_p(A_i) : \ \left(\sum_i |A_i|^q\right)^{\frac{1}{q-1}} \geq \delta\right\}, \quad \text{for } q \neq 1 \tag{3.27}$$

For $q=1$ and $q=-\infty$, we obtain similarly

$$\begin{aligned} C_{p,d}(1) &:= \lim_{\delta\to 0} \delta^{-\frac{p}{d}} \min_{\mathcal{P}} \left\{ \sum_i \psi_p(A_i) : \ -\sum_i |A_i| \log(|A_i|) \geq \log\delta \right\} \\ C_{p,d}(-\infty) &:= \lim_{\delta\to 0} \delta^{-\frac{p}{d}} \min_{\mathcal{P}} \left\{ \sum_i \psi_p(A_i) : \ \inf_i |A_i| \geq \delta \right\} . \end{aligned}$$

In previous formulae, the minima are taken over all partitions $\mathcal{P}=(A_i)_i$ of the unit cube Q and the infinitesimal δ is given respectively by $\delta = n^{-\frac{1}{|q-1|}}$ for $q\neq 1$, $\delta=\exp(-n)$ for $q=1$ and $\delta=n^{-1}$ for $q=-\infty$.

Remark 3.13. i) The location problem where a prescribed number of points is imposed can be deduced by taking $q=0$. We then recover the asymptotic behavior obtained in [4] in the case $p=1$ and in [8, 9] in the case $p\geq 1$.
ii) The explicit computation of he constant $C_{p,d}(q)$ is a very difficult task except in the case where the norm considered in $\mathbb{R}^d$ has the tiling property (then $C_{p,d}(q)=\gamma_{p,d}=\Gamma_{p,d}$ for all $q\leq 1+p/d$). However, in the case of the Euclidean norm, it is known that if $p=1,2$ then $C_{p,2}(0)=\psi_p(H_6)$ where H_6 is a regular hexagon of unit Lebesgue measure (cf. Bolton and Morgan [4] and Newman [19]). As a consequence, by noticing that $\omega_{p,2}\leq\Gamma_{p,2}\leq\psi_p(H_6)$ (see the assertion iv) of Proposition 3.2) and exploiting (3.22), we find that the non-increasing function $q\to C_{p,2}(q)$ is in fact constant on $]-\infty,0]$; more precisely, for all $q\leq 0$, there holds : $C_{p,2}(q)=\omega_{p,2}=\Gamma_{p,2}=\psi_p(H_6)$.
iii) By adapting the proof of (3.22), it is possible to show that for every $f\in L^1(\Omega)$ the limit of the left hand member of (3.25) vanishes as $n\to\ \infty$ whenever $q>1+p/d$ (which is is consistent with the fact that $C_{p,d}(q)=0$ in (3.22)). On the other hand, for $1<q<1+p/d$, it may happen that the equality in (3.25) becomes a strict inequality if one removes the boundedness assumption on f. This can be seen in the example below, where the limit in the left hand member of (3.25) vanishes.

Example: Take $d=1$, $\Omega=(0,1)$ and $f=(1-\alpha)x^{-\alpha}$ where $0\leq\alpha<1$. As in the assertion iii) of Proposition 3.11, we consider the entropy H associated with $h(t)=t^q$ where $q\in]1,1+p/d[$. For a given $\delta>0$ and $q>1$, we set $\theta_\delta=\delta^{1/q}$ and $l_\delta=\theta_\delta^{1/(1-\alpha)}$ so that $\int_0^{l_\delta} f\,dx=\theta_\delta$. Noticing that $\nu_\delta=\theta_\delta\delta_0+\tilde\nu$ satisfies the size constraint $H(\nu_\delta)\leq-\delta$ ($h(t)=-t^q$), for any measure $\tilde\nu$ on $(l_\delta,1)$ with total mass $1-\theta_\delta$ and optimizing with respect to such a $\tilde\nu$, it is easy to check that $\mathcal{E}_H^{1,p}(f,\Omega,-\delta)\leq\theta_\delta l_\delta^p$. Then $\lim_{\delta\to 0}\delta^{-p/(q-1)}\,\mathcal{E}_H^{1,p}(f,\Omega,-\delta)=0$ whenever $q>q_\alpha$ where $q_\alpha:=\frac{p+1-\alpha}{p\alpha+1-\alpha}$ (observe that $1<q_\alpha<1+p$ for $\alpha\in(0,1)$). Thus applying (3.25) would lead to $C_{p,1}(q)=0$ which is clearly false for $q\leq 1+p$ where $C_{p,1}(q)=\gamma_{p,1}=\frac{1}{(p+1)2^p}$. Therefore the boundedness assumption for f in Proposition 3.11 iii) is crucial.

3.6. **Open problems and conjectures.** i) *About the constants $\omega_{p,d}$ and $\Gamma_{p,d}$.* Recall that the constant $\omega_{p,d}$ introduced in Proposition 3.2 satisfies $\omega_{p,d}\leq\Gamma_{p,d}$ (being $\Gamma_{p,d}$ defined in (2.15)). It seems reasonable to conjecture that this inequality is actually an equality. In this case, the condition $\omega_{p,d}=\gamma_{p,d}$ characterizing the locality property of G would be equivalent to $\theta(B)=1$ (that is the norm on $\mathbb{R}^d$ enjoys the tiling property).

A first step in the direction of the conjecture could be to prove the existence of an optimal compact set with respect to the infimum in (2.15). Notice that, for $d=2$, $p=1,2$ and if $|\cdot|$ is the Euclidean norm (thus $\theta(B)<1$), we know from [4, 19] that $\omega_{p,2}=\Gamma_{p,2}=\Psi_p(P)$ where P is a regular hexagon.

ii) *About the constants $C_{p,d}(q)$.* We observed that the value of $C_{p,d}(q)$ is monotone non-increasing with respect to q and that the maximum value $C_{p,d}(-\infty)$ is below $\omega_{p,d}$. On the other hand $C_{p,d}(0)$ is nothing else but the constant appearing in the optimal location problem. In the case $d = 2, p = 1$ and $|\cdot|$ the Euclidean norm, owing to the results in [4], we have that $C_{1,2}(0) = \omega_{1,2} = \Gamma_{1,2}$. Thus the function $C_{1,1}(q)$ is constant on $(-\infty, 0]$. We conjecture that this constancy is also true for $C_{p,d}(q)$ in the general case and even that, for $q \in]-\infty, 1[$, there holds $C_{p,d}(q) = \omega_{p,d}$. This conjecture could be proved by showing that the infimum appearing in the definition (3.20) of $C_{p,d}(q)$ is reached for ρ being the Dirac mass δ_1.

4. The functional $G(\rho)$: properties and constrained minimization.

This section concerns the approximation of the uniform density on the unit cube. The related functional G has been introduced in (3.5). We will establish some fundamental properties of this functional stated in Proposition 3.2 and then we will study the minimization of G under entropy constraints (as stated in Lemma 3.7).

First, we observe that, for every $\delta > 0$, S_δ given in (3.3) is both subadditive and translation invariant and by applying Lemma 2.8 we may write

$$G_\delta(\rho) \; = \; \lim_{k\to\infty} \frac{S_\delta(\rho, Q_k)}{|Q_k|}. \tag{4.1}$$

On the other hand from the definition of S_δ, one deduces easily the following Lipschitz estimate

$$|S_\delta(\rho_1, A) - S_\delta(\rho_2, A)| \le \frac{|A|}{\delta^p} N(\rho_1 - \rho_2) \quad \forall \rho_1, \rho_2 \in \mathcal{P}(\mathbb{R}^+) \ , \tag{4.2}$$

for every measurable set A, so that by (4.1) and (4.2) we also have

$$|G_\delta(\rho_1) - G_\delta(\rho_2)| \le \frac{1}{\delta^p} N(\rho_1 - \rho_2). \tag{4.3}$$

4.1. **Some preliminary results.**

We shall write $G_\delta(\rho)$ in a form useful for future calculations. For this we first note the change of variables formulae for arbitrary $\varepsilon > 0$ and $\nu \in \mathcal{M}_0^+(A)$:

$$W_p^p\left(1\!\!1_{\varepsilon A}\, ,\, \varepsilon^d\, L_\varepsilon^\sharp(\nu)\right) \; = \; \varepsilon^{p+d}\, W_p^p(1\!\!1_A, \nu) \quad , \quad \rho_\varepsilon\left(\varepsilon^d\, L_\varepsilon^\sharp(\nu)\right) = \rho\left(L_\varepsilon^\sharp(\nu)\right) = \rho(\nu) \ . \tag{4.4}$$

Lemma 4.1. *For every $r > 0$ and $\rho \in \mathcal{P}(\mathbb{R}^+)$, there holds*

$$\lim_{\varepsilon\to 0} \left\{ \inf \left\{ \frac{1}{\varepsilon^p\, r^d} W_p^p(1\!\!1_{Q_r}, \nu) + \frac{1}{\delta^p} N(\rho - \rho_\varepsilon(\nu)) : \nu \in \mathcal{M}_0^+(Q_r) \right\} \right\} \; = \; G_\delta(\rho) \ . \tag{4.5}$$

Proof. We start with (3.4) taking $k = \frac{r}{\varepsilon}$ and apply (4.4) with $A = Q_{\frac{r}{\varepsilon}}$:

$$\begin{aligned} G_\delta(\rho) \; &= \; \lim_{\varepsilon\to 0} \frac{S_\delta(\rho, Q_{\frac{r}{\varepsilon}})}{\left(\frac{r}{\varepsilon}\right)^d} \\ &= \; \lim_{\varepsilon\to 0} \left(\frac{\varepsilon}{r}\right)^d \inf \left\{ W_p^p(1\!\!1_{Q_{\frac{r}{\varepsilon}}}, \nu) + \frac{|Q_{\frac{r}{\varepsilon}}|}{\delta^p} N(\rho - \rho(\nu)) : \; \nu \in \mathcal{M}_0^+(Q_{\frac{r}{\varepsilon}}) \right\} \\ &= \; \lim_{\varepsilon\to 0} \left\{ \inf \left\{ \frac{1}{\varepsilon^p\, r^d} W_p^p\left(1\!\!1_{Q_r}, \nu(\frac{\cdot}{\varepsilon})\right) + \frac{1}{\delta^p} N\left(\rho - \rho_\varepsilon(\nu(\frac{\cdot}{\varepsilon}))\right) : \nu \in \mathcal{M}_0^+(Q_{\frac{r}{\varepsilon}}) \right\} \right\}. \end{aligned}$$

The last infimum can be taken equivalently with respect to $\tilde{\nu} = \nu(\frac{\cdot}{\varepsilon}))$ running over $\mathcal{M}_0^+(Q_r)$. $\square$

Lemma 4.2. *Let $r > 0$ and $a > 0$. Then, for every sequence $\{\nu_\varepsilon\}$ in $\mathcal{M}_0^+(Q_r)$, we have*

$$N(\rho_\varepsilon(\nu_\varepsilon) - \rho) \to 0 \quad \Rightarrow \quad \liminf_{\varepsilon\to 0} \frac{W_p^p(a\,\mathbb{1}_{Q_r}, \nu_\varepsilon)}{\varepsilon^p\, r^d} \;\geq\; a^{1-\frac{p}{d}}\, G(\rho)\ . \tag{4.6}$$

Futhermore, for every $\rho \in \mathcal{P}(\mathbb{R}^+)$ such that $G(\rho) < +\infty$, there exists a sequence $\{\nu_\varepsilon\}$ in $\mathcal{M}_0^+(Q_r)$ such that

$$N(\rho_\varepsilon(\nu_\varepsilon) - \rho) \to 0 \quad , \quad \lim_{\varepsilon\to 0} \frac{W_p^p(a\,\mathbb{1}_{Q_r}, \nu_\varepsilon)}{\varepsilon^p\, r^d} \;=\; a^{1-\frac{p}{d}}\, G(\rho)\ . \tag{4.7}$$

Proof. Recalling that G_δ converges increasingly to G and using a classical diagonalization argument (see [2]), it is easy to check that (4.5) implies the statements (4.6) and (4.7) when $a = 1$. The extension to the general case is a consequence of the following observation: given any sequence $\{\nu_\varepsilon\}$ in $\mathcal{M}_0^+(Q_r)$ such that $\nu_\varepsilon \to a\mathbb{1}_{Q_r}$ and $\rho_\varepsilon(\nu_\varepsilon) \to \rho$, the new sequence $\{\nu'_\varepsilon\}$ obtained by setting $\nu'_\varepsilon := L^\sharp_{a^{\frac{1}{d}}}(\nu_\varepsilon)$ satisfies for $r' = r\, a^{\frac{1}{d}}$:

$$\nu'_\varepsilon \;\to\; \mathbb{1}_{Q_{r'}} \quad , \quad \rho_\varepsilon(\nu'_\varepsilon) \;\to\; \rho \quad , \quad W_p^p(\mathbb{1}_{Q_{r'}}, \nu'_\varepsilon) \;=\; a^{\frac{p}{d}}\; W_p^p(a\,\mathbb{1}_{Q_r}, \nu_\varepsilon)\ .$$

(here we used the fact that $\rho_\varepsilon(\nu'_\varepsilon) = \rho_\varepsilon(\nu_\varepsilon)$ and, in order to derive the last equality, that $L^\sharp_{a^{\frac{1}{d}}}(a\,\mathbb{1}_{Q_r}) = \mathbb{1}_{Q_{r'}}$ combined with relation (2.7)). □

Lemma 4.3. *Let f be a measurable function such that $\alpha \leq f \leq M$ where α, M are positive constants. Then for every $\nu \in \mathcal{M}_0^+(\Omega)$ we have:*

$$\frac{W_p^p(f,\nu)}{\varepsilon^p} \geq \frac{\alpha}{M^{1+p/d}}\,\gamma_{p,d} \int_{\Omega\times\mathbb{R}^+} t^{p/d}\; \lambda_\varepsilon(\nu)(dxdt)\ . \tag{4.8}$$

Proof. Let $\nu = \sum_i c_i \delta_{x_i}$ in $\mathcal{M}_0^+(\Omega)$ be such that $W_p^p(f,\nu) < +\infty$ and let $\{A_i\}_{i=1}^\infty$ be a partition of Ω satisfying

$$\int_{A_i} f(x)\,dx = c_i \quad\text{and}\quad W_p^p(f,\nu) = \sum_i \int_{A_i} |x - x_i|^p f(x)\,dx.$$

Then, by using (2.12) and $\alpha \leq f \leq M$, we have

$$\int_{A_i} |x - x_i|^p f(x)\; dx \;\geq\; \alpha\gamma_{p,d}\,|A_i|^{1+p/d} \;\geq\; \frac{\alpha\,\gamma_{p,d}}{M^{1+p/d}}\, c_i^{1+p/d}\ ,$$

yielding that

$$\frac{W_p^p(f,\nu)}{\varepsilon^p} \;=\; \frac{1}{\varepsilon^p}\sum_i \int_{A_i} |x{-}x_i|^p\, f(x)\,dx \geq \frac{1}{\varepsilon^p}\frac{\alpha\,\gamma_{p,d}}{M^{1+p/d}}\sum_i c_i^{1+p/d} \;=\; \frac{\alpha\,\gamma_{p,d}}{M^{1+p/d}}\int_{\Omega\times\mathbb{R}^+} t^{p/d}\,\lambda_\varepsilon(\nu)(dxdt).$$

□

Lemma 4.4. *For every $\delta > 0$ the function G_δ is convex.*

Proof. G_δ being a continuous function for very $\delta > 0$, it is enough to show that

$$G_\delta\left(\frac{\rho^1 + \rho^2}{2}\right) \;\leq\; \frac{1}{2}G_\delta(\rho^1) + \frac{1}{2}G_\delta(\rho^2) \qquad \forall \rho^1, \rho^2 \in \mathcal{P}(\mathbb{R}^+).$$

According to the property (4.5), we can choose two sequences of measures $\{\nu_\varepsilon^s\}$, $s \in \{1,2\}$ supported in Q and so that:

$$G_\delta(\rho^s) = \lim_{\varepsilon\to 0}\left(\frac{W_p^p(\mathbb{1}_Q, \nu_\varepsilon^s)}{\varepsilon^p} + \frac{1}{\delta^p} N(\rho^s - \rho_\varepsilon(\nu_\varepsilon^s))\right). \tag{4.9}$$

We split the 2^d vertices $\{a_j\}$ of the cube Q into two subfamilies $\{a_j\,;\, j\in J_s\}$, $s\in\{1,2\}$ each of them having 2^{d-1} elements. Setting $A_j := a_j + Q$, we obtain a covering of $Q_2 =]-1,1[^d$ (a.e.) by considering $\{A_j\,,\, j\in J_1\cup J_2\}$. For each $j\in J_1\cup J_2$, the push forward by the a_j-translation of ν^s_ε provides a measure $\nu^s_\varepsilon(x-a_j)$ on A_j and we obtain a measure on Q_2 of total mass 2^d by setting:

$$\nu_\varepsilon = \sum_{j\in J_1}\nu^1_\varepsilon(x-a_j) \;+\; \sum_{j\in J_2}\nu^2_\varepsilon(x-a_j)\ .$$

By the sub-additivity property (2.6) and (2.7), one has

$$W_p^p(\mathbb{1}_{Q_2},\nu_\varepsilon) \;\leq\; 2^{d-1}\left(W_p^p(\mathbb{1}_Q,\nu^1_\varepsilon)+W_p^p(\mathbb{1}_Q,\nu^2_\varepsilon)\right)\ , \tag{4.10}$$

whereas it is easy to verify that $\rho_\varepsilon(\nu_\varepsilon)=\frac{1}{2}\rho_\varepsilon(\nu^1_\varepsilon)+\frac{1}{2}\rho_\varepsilon(\nu^2_\varepsilon)$ so that

$$N\big(\frac{1}{2}\rho^1+\frac{1}{2}\rho^2-\rho_\varepsilon(\nu_\varepsilon)\big) \;\leq\; \frac{1}{2}N(\rho_\varepsilon(\nu^1_\varepsilon)-\rho^1) \;+\; \frac{1}{2}N(\rho_\varepsilon(\nu^2_\varepsilon)-\rho^2)\ . \tag{4.11}$$

By applying (4.5) with $\rho=\frac{1}{2}(\rho^1+\rho^2)$ and $r=2$ and exploiting (4.9)(4.10)(4.11), we are led to

$$\begin{aligned}
G_\delta\left(\frac{\rho^1+\rho^2}{2}\right) \;&\leq\; \liminf_{\varepsilon\to 0}\left(\frac{W_p^p(\mathbb{1}_{Q_2},\nu_{\varepsilon,\delta})}{\varepsilon^p\,2^d}+\frac{1}{\delta^p}N\big(\frac{1}{2}\rho^1+\frac{1}{2}\rho_2-\rho_\varepsilon(\nu_\varepsilon)\big)\right)\\
&\leq \limsup_{\varepsilon\to 0}\frac{1}{2}\left(\frac{W_p^p(\mathbb{1}_Q,\nu^1_\varepsilon)}{\varepsilon^p}+\frac{N(\rho^1-\rho_\varepsilon(\nu^1_\varepsilon))}{\delta^p}\right)+\frac{1}{2}\left(\frac{W_p^p(\mathbb{1}_Q,\nu^2_\varepsilon)}{\varepsilon^p}+\frac{N(\rho^2-\rho_\varepsilon(\nu^2_\varepsilon))}{\delta^p}\right)\\
&\leq\; \frac{1}{2}G_\delta(\rho^1)+\frac{1}{2}G_\delta(\rho^2).
\end{aligned}$$

□

We finish this subsection with a technical result related to the N-metric on $\mathcal{P}(\mathbb{R}^+)$.

Lemma 4.5. *Let $\beta\geq 1$. Then for every ρ,ρ' in $\mathcal{P}(\mathbb{R}^+)$ and all $r,s\in[0,\beta]$*

$$N(L^\sharp_s(\rho)-L^\sharp_s(\rho')) \quad\leq\quad \beta\,N(\rho-\rho') \tag{4.12}$$

$$N(L^\sharp_r(\rho)-L^\sharp_s(\rho')) \quad\leq\quad \beta\,|r-s| \;+\; 2\,\rho([\beta,+\infty[) \;+\; \beta\,N(\rho-\rho')\ . \tag{4.13}$$

Proof. Let $\varphi:\mathbb{R}^+\mapsto\mathbb{R}$ such that $\|\varphi\|_{\rm BL}\leq 1$. Then, as $s\leq\beta$, there holds $\|\varphi(s\,\cdot)\|_{\rm BL}\leq\beta\,{\rm Lip}(\varphi)+|\varphi|_\infty\leq\beta\|\varphi\|_{\rm BL}$ so that relation (4.12) holds. On the other hand, we have $|\varphi(rt)-\varphi(st)|\;\leq\;\beta\,|r-s|$ if $t\leq\beta$ whereas $|\varphi(rt)-\varphi(st)|\;\leq\;2$ is always true. Thus

$$N(L^\sharp_r(\rho)-L^\sharp_s(\rho)) \;=\; \sup_{\|\varphi\|_{\rm BL}\leq 1}|\langle L^\sharp_r(\rho)-L^\sharp_s(\rho),\varphi\rangle| \;\leq\; \beta|r-s|+2\rho([\beta,+\infty[)\ .$$

The relation (4.13) follows by (4.12) and by using triangle inequality.

□

4.2. **Proof of Proposition 3.2:** *Proof of i).* It is a consequence of the fact that G is the supremum of the family of functions $\{G_\delta\}$ which are convex by Lemma 4.4 and continuous by (4.3).

Proof of ii). Applying (4.7) for $r=1$ at $L^\sharp_a(\rho)$, we can choose a sequence a sequence $\{\nu_\varepsilon\}$ in $\mathcal{M}^+_0(Q)$ such that

$$\rho_\varepsilon(\nu_\varepsilon)\to L^\sharp_a(\rho)\quad,\quad \lim_{\varepsilon\to 0}\frac{W_p^p(\mathbb{1}_Q,\nu_\varepsilon)}{\varepsilon^p} \;=\; G(L^\sharp_a(\rho))\ .$$

Consider the new sequence defined by $\nu'_\varepsilon := \frac{1}{a}\,\nu_\varepsilon$. Then as $\rho_\varepsilon(\nu'_\varepsilon)=L^\sharp_{a^{-1}}(\rho_\varepsilon(\nu_\varepsilon))$, we obtain that $\rho_\varepsilon(\nu'_\varepsilon)\to\rho$. Therefore, applying (4.6), it follows that

$$G(L^\sharp_a(\rho)) \;=\; \lim_{\varepsilon\to 0}\frac{W_p^p(\mathbb{1}_Q,\nu_\varepsilon)}{\varepsilon^p} \;=\; \lim_{\varepsilon\to 0}\,a\frac{W_p^p(a^{-1}\mathbb{1}_Q,\nu'_\varepsilon)}{\varepsilon^p} \;\geq\; a\,(a^{-1})^{1-\frac{p}{d}}\,G(\rho) \;=\; a^{\frac{p}{d}}\,G(\rho)\ .$$

The converse inequality comes out by switching a to a^{-1} while replacing ρ by $L_a^\sharp \rho$.

Proof of iii). By the scale invariance (3.7), it is enough to consider the case $a = 1$. Given $\theta \in]0, \theta(P)[$, there exists in Q a family of small disjoint subsets $P_{i,\varepsilon} = x_i^\varepsilon + \varepsilon P$ with $i \in I_\varepsilon$ such that the total measure of $A_\varepsilon = \cup_{i\in I_\varepsilon} P_{i,\varepsilon}$ converges to θ as $\varepsilon \to 0$. All subsets $P_{i,\varepsilon}$ share the same measure ε^d so that $\varepsilon^d \sharp(I_\varepsilon) \to \theta$. We may cover the complement $B_\varepsilon = Q \setminus A_\varepsilon$ with disjoint hypercubes of smaller size $Q_{j,\varepsilon} = y_{j,\varepsilon} + r_{j,\varepsilon} Q$ where $j \in J_\varepsilon$ and $r_{j,\varepsilon} \leq \varepsilon^2$ (but $r_{j,\varepsilon}$ may depend on j) so that $\sum_{j\in J_\varepsilon} r_{j,\varepsilon}^d \to 1 - \theta$. Then we set

$$\nu_\varepsilon = \sum_{i\in I_\varepsilon} \varepsilon^d \delta_{x_{i,\varepsilon}} + \sum_{j\in J_\varepsilon} r_{j,\varepsilon}^d \delta_{y_{j,\varepsilon}}.$$

Then we can majorize the optimal transport of ν_ε to $\mathbb{1}_Q\, dx$ as follows

$$\limsup_{\varepsilon\to 0} \frac{W_p^p(\nu_\varepsilon, \mathbb{1}_Q)}{\varepsilon^p} \;\leq\; \limsup_{\varepsilon\to 0} \left(\sharp(I_\varepsilon)\, \varepsilon^d\, \Psi_p(P) + \sum_{j\in J_\varepsilon} \frac{r_{j,\varepsilon}^{p+d}}{\varepsilon^p}\, \Psi_p(Q) \right) \;\leq\; \theta \Psi_p(P).$$

On the other hand, it is easy to check that $N(\rho_\varepsilon(\nu_\varepsilon) - (\theta\delta_1 + (1-\theta)\delta_0)) \to 0$. The inequality (3.8) follows by applying (4.6) in Lemma 4.2 (with $r = a = 1$). If P is the unit ball, then $\Psi_p(P) = \gamma_{p,d}$ and $G(\theta\delta_1 + (1-\theta)\delta_0) \leq \gamma_{p,d}\, \theta$ for every $\theta \in [0, \theta(B)]$. By (3.10) (whose proof is given below), the converse equality holds also.

Proof of iv). First we check that $\omega_{p,d} \leq \Gamma_{p,d} < +\infty$. Clearly $P = [0,1|^d$ satisfies $\theta(P) = 1$ and therefore $\Gamma_{p,d} \leq \Psi_p([0,1|^d) < +\infty$. Now if P is another compact subset such that $|P| = 1$ and $\theta(P) = 1$, by applying (3.8) with $\theta = a = 1$, we obtain $\omega_{p,d} := G(\delta_1) \leq \Psi_p(P)$ thus $\omega_{p,d} \leq \Gamma_{p,d}$ by minimizing over such P.

Let us prove the first inequality in (3.10). We assume without loss of generality that $G(\rho) < +\infty$. Then by applying (4.7) with $a = r = 1$ and the estimate (4.8) taking $f = \mathbb{1}_Q$, we obtain the desired inequality after passing to the limit $\varepsilon \to 0$. We prove now the upper bound for G in (3.10).

From the scaling property and as $\omega_{p,d} = G(\delta_1)$, we deduce that, for every $a \in \mathbb{R}^+$, there holds $G(\delta_a) = \omega_{p,d}\, a^{p/d}$. Now by the the convexity of G already established, it follows immediately that

$$G(\rho) \leq \omega_{p,d} \int t^{p/d} \rho(dt) \qquad , \quad \forall \rho \in \mathcal{M}_0^+(\mathbb{R}^+) \ .$$

Thanks to the lower semicontinuity of G, this inequality can be extended to probality measures ρ with compact support by using a sequence of discrete probability measures supported in a fixed compact set and converging weakly to ρ (so that the right hand side converges). Eventually, if $\rho \in \mathcal{P}(\mathbb{R}^+)$ is a general measure such that $\int t^{p/d}\, d\rho < +\infty$, we reduce to the previous case by considering $\rho_n := \frac{1}{\rho([0,n])} \rho \llcorner [0,n]$ and sending $n \to \infty$:

$$G(\rho) \;\leq\; \liminf_n G(\rho_n) \;\leq\; \liminf_n \frac{\omega_{p,d}}{\rho([0,n])} \int_{[0,n]} t^{\frac{p}{d}} \rho(dt) \;=\; \omega_{p,d} \int_{\mathbb{R}^+} t^{\frac{p}{d}} \rho(dt) \ .$$

Proof of v). It is clear that if $\omega_{p,d} = \gamma_{p,d}$, the inequality (3.10) implies that $G(\rho) = \gamma_{p,d} \int t^{p/d} \rho(dt)$ so that G has a very simple explicit form and is local. Converserly, if G is local (in the sense of Definition 2.1), for every $\theta \in [0,1]$ one has

$$G(\theta\, \delta_1 + (1-\theta)\, \delta_0) \;=\; \theta\, G(\delta_1) + (1-\theta)\, G(\delta_0) \;=\; \theta\, \omega_{p,d}.$$

Then applying (3.9) yields the equality $\theta\,\omega_{p,d} = \theta\,\gamma_{p,d}$ for every $\theta \in [0, \theta(B)]$, hence $\omega_{p,d} = \gamma_{p,d}$. This concludes the proof of the proposition. □

4.3. **Proof of Lemma 3.7.** The first statement can be derived in the same way as in the proof of Lemma 3.4. Let us prove (3.18). From the second inequality in (3.10) we immediately obtain the majorization $\Phi_g(s) \leq \omega_{p,d}\varphi_g(s)$ and therefore, the convexity of Φ_g leads directly to the second inequality in (3.18). In order to show the first inequality in (3.18), we introduce

$$\beta(s) := \inf\left\{\int t^{p/d}\rho(dt) : \int g(t)\rho(dt) \leq s\right\} . \tag{4.14}$$

It can be seen, similarly as for the function Φ_g, that β is convex, non-increasing and lower semicontinuous. Also, from the first inequality in (3.10), we have $\gamma_{p,d}\,\beta(s) \ \leq\ \Phi_g(s)$. Thus we are reduced to check that $\beta(s) = \varphi_g^{**}(s)$. The Fenchel conjugate β^* is finite only if $s^* < 0$ and is given by:

$$\begin{aligned}\beta^*(s^*) &:= \sup_{s,\rho}\{s^* s - \int t^{p/d}\rho(dt) : \int g(t)\rho(dt) \leq s\}\\ &= \sup_{\rho}\{s^* \int g(t)\rho(dt) - \int t^{p/d}\rho(dt)\} = \sup_{t}\{s^* g(t) - t^{p/d}\} = \varphi_g^*(s^*) \ .\end{aligned}$$

Passing again to the Fenchel conjugate, we infer that $\beta(s) = \beta^{**}(s) = \varphi_g^{**}(s)$. □

4.4. **Weak-strong lower semicontinuity of** $E(f, \lambda)$**.** We are going to prove that the functional $E(f, \lambda)$, defined in (3.6), is lower semicontinuous with respect to strong convergence in the first variable $f \in L^1(\Omega)$ and the tight convergence of the second variable $\lambda \in \mathcal{M}(\Omega \times \mathbb{R}^+)$.

Lemma 4.6. *Let $\{f_n, \lambda_n\}_n$ be a sequence in $L^1(\Omega; [\alpha, +\infty)) \times \mathcal{M}(\Omega \times \mathbb{R}^+)$ such that $\int_\Omega |f_n - f|\, dx \ \to 0$ and $\lambda_n \overset{*}{\rightharpoonup} \lambda$ tightly as $n \to \infty$. Then there holds: $\liminf_{n\to\infty} E(f_n, \lambda_n) \ \geq\ E(f, \lambda)$.*

Proof. Writing $\lambda_n = f_n \otimes \rho_n^x$ and $\lambda = f \otimes \rho^x$ for suitable families of probability ρ_n^x, ρ^x. It is enough to prove that for every $\delta > 0$

$$\liminf_{n\to\infty} \int_\Omega f_n(x)^{1-\frac{p}{d}} G_\delta(\rho_n^x)\, dx \geq \int_\Omega f(x)^{1-\frac{p}{d}} G_\delta(\rho^x)\, dx. \tag{4.15}$$

Indeed G_δ goes increasing to G as $\delta \to 0$ and we can pass to the limit in the right hand side by Beppo Levi's theorem. Now we exploit the Lipschitz continuity of G_δ with respect to the norm $N(\cdot)$ which entails that the sequence $\{G_\delta(\rho_n^x), n \in \mathbb{N}\}$ is bounded in $L^\infty(\Omega)$ and then, possibly after extracting a subsequence, does converge weakly-star to some function $m(x) \in L^\infty(\Omega)$. On the other hand, as $f_n \geq \alpha$, we have that $f_n^{1-\frac{p}{d}} \leq \alpha^{-\frac{p}{d}} f_n$ and thus by dominated convergence $f_n^{1-\frac{p}{d}} \to f^{1-\frac{p}{d}}$ strongly in $L^1(\Omega)$. It follows that the left hand side of (4.15) converges to $\int_\Omega m(x)\, f^{1-\frac{p}{d}}\, dx$ and we are reduced to show that

$$m(x) \ \geq\ G_\delta(\rho^x) \quad \text{a.e. } x \in \Omega\ . \tag{4.16}$$

Next we claim that

$$< \rho_n^x, \varphi > \overset{*}{\rightharpoonup} < \rho^x, \varphi > \quad \text{weakly-star in } L^\infty(\Omega) \quad \forall \varphi \in \mathcal{C}_0(\mathbb{R}^+). \tag{4.17}$$

Indeed $\{\rho_n^x\,,\, n \in \mathbb{N}\}$ is a bounded sequence in $L^\infty(\Omega, \mathcal{M}^+(\mathbb{R}^+))$ and so we may assume, without loss of generality, that there exists a function $\tilde{\rho}^x \in L^\infty(\Omega, \mathcal{M}^+(\mathbb{R}^+))$ such that

$$\int_\Omega \langle \rho_n^x, \varphi\rangle\ \psi(x)\ dx \to \int_\Omega \langle \tilde{\rho}^x, \varphi\rangle \psi(x)\ dx \quad \forall \psi \in \mathrm{L}^1(\Omega),\ \forall \varphi \in \mathcal{C}_0(\mathbb{R}^+).$$

Then, by the convergence $\lambda_n \overset{*}{\rightharpoonup} \lambda$, one checks easily that $\tilde{\rho}^x$ agrees with ρ^x (thus the claim (4.17)):

$$\begin{aligned}\int_\Omega \langle \tilde{\rho}^x, \varphi\rangle f(x)\,\psi(x)\,dx &= \lim_{n\to\infty} \int_\Omega \langle \rho_n^x, \varphi\rangle f_n(x)\psi(x)\,dx \\ &= \lim_{n\to\infty} \langle \lambda_n, \psi(x)\varphi(t)\rangle = \langle \lambda, \psi(x)\,\varphi(t)\rangle = \int_\Omega \langle \rho^x, \varphi\rangle f(x)\psi(x)\,dx\,.\end{aligned}$$

Given a Borel subset $A \subset \Omega$, by the convexity of G_δ and Jensen's inequality, one has for every n

$$G_\delta\left(\frac{1}{|A|}\int_A \rho_n^x\,dx\right) \;\leq\; \frac{1}{|A|}\int_A G_\delta(\rho_n^x)\,dx\;.$$

By (4.17), we can pass to the limit in the left hand side using the continuity of G_δ and in the right hand by exploiting the weak star convergence $G_\delta(\rho_n^x) \overset{*}{\rightharpoonup} m(x)$:

$$G_\delta\left(\frac{1}{|A|}\int_A \rho^x\,dx\right) \;\leq\; \frac{1}{|A|}\int_A m(x)\,dx\;.$$

By taking $A = x_0 + \varepsilon B$ where $\varepsilon \to 0$, we conclude that (4.16) holds at every Lebesgue point x_0 of the functions $m(x)$ and ρ^x. As $x \mapsto \rho^x$ ranges in the dual of a separable metric space, the inequality in (4.16) holds for a.e. $x_0 \in \Omega$. The proof of Lemma 4.6 is complete. □

5. Proof of the Γ-convergence Theorem

We begin by some premilinary lemmas

Lemma 5.1. *Let be given $\{\lambda_\varepsilon\}_{\varepsilon>0}$ a sequence of measures such that $\sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon) < \infty$. Then*

i) For every constant $k>0$, we can find a modified sequence λ_ε^k such that

$$\mathrm{supp}(\lambda_\varepsilon^k) \subset \Omega\times[0,k] \quad , \quad E_\varepsilon(f,\lambda_\varepsilon) \geq E_\varepsilon(f,\lambda_\varepsilon^k) \quad , \quad \int |\lambda_\varepsilon^k - \lambda_\varepsilon| \leq 2\,\lambda_\varepsilon\,(\Omega\times[k,+\infty[)\,.$$

ii) For any measurable $\tilde{f}$ such that $0 \leq \tilde{f} \leq f$ a.e., we can find a new sequence $\tilde{\lambda}_\varepsilon$ such that

$$E_\varepsilon(f,\lambda_\varepsilon) \geq E_\varepsilon(\tilde{f},\tilde{\lambda}_\varepsilon) \quad , \quad \int |\tilde{\lambda}_\varepsilon - \lambda_\varepsilon| \;=\; \int_\Omega |f-\tilde{f}|\,dx\;.$$

Proof. The finiteness of $E_\varepsilon(f,\lambda_\varepsilon)$ implies that, for each $\varepsilon > 0$, there exists a discrete measure $\nu_\varepsilon = \sum_i c_{i,\varepsilon}\delta_{x_{i,\varepsilon}}$ such that $\lambda_\varepsilon = \lambda_\varepsilon(\nu_\varepsilon)$ and $E_\varepsilon(f,\lambda_\varepsilon) = \varepsilon^{-p}\,W_p^p(f,\nu_\varepsilon)$. Let $\{A_{i,\varepsilon}\ :\ i \in I_\varepsilon\}$ be a partition of Ω such that

$$\int_{A_{i,\varepsilon}} f(x)\,dx = c_{i,\varepsilon} \quad \text{and} \quad W_p^p(f,\nu_\varepsilon) = \sum_i \int_{A_{i,\varepsilon}} |x - x_{i,\varepsilon}|^p\,f(x)\,dx. \tag{5.1}$$

For the assertion i), we set $I_\varepsilon^k = \{i \in I_\varepsilon\ :\ c_{i,\varepsilon} > k\,\varepsilon^d\}$ and, for every $i \in I_\varepsilon^k$, we partition again $A_{i,\varepsilon}$ into a family $\{A_{i,\varepsilon}^j\}_j$ such that:

$$\int_{A_{i,\varepsilon}^j} f(x)\;dx \leq k\,\varepsilon^d. \tag{5.2}$$

We choose a family of points $\{y_{i,\varepsilon}^j\}_j$ in Ω, all distinct, such that

$$\int_{A_{i,\varepsilon}^j} |x - y_{i,\varepsilon}^j|^p f(x)\;dx \leq \int_{A_{i,\varepsilon}^j} |x - x_{i,\varepsilon}|^p f(x)\;dx \tag{5.3}$$

and define the new measure

$$\nu_\varepsilon^k := \sum_{i\notin I_\varepsilon^k} c_{i,\varepsilon}\delta_{x_{i,\varepsilon}} + \sum_{i\in I_\varepsilon^k}\sum_j c_{i,\varepsilon}^j \delta_{y_{i,\varepsilon}^j} \quad , \quad c_{i,\varepsilon}^j = \int_{A_{i,\varepsilon}^j} f(x)\,dx \ . \tag{5.4}$$

Subsequently, we define $\lambda_\varepsilon^k := \lambda_\varepsilon(\nu_\varepsilon^k)$ which by construction is supported in $\Omega\times[0,k]$. We have

$$\int_{\Omega\times\mathbb{R}^+} |\lambda_\varepsilon - \lambda_\varepsilon^k| \ = \ \int_\Omega |\nu_\varepsilon - \nu_\varepsilon^k| \ \le \ 2\sum_{i\in I_\varepsilon^k} c_{i,\varepsilon} \ = \ 2\int_{\Omega\times\{t>k\}} d\lambda_\varepsilon(x,t)$$

and the inequality $E_\varepsilon(f,\lambda_\varepsilon) \ge E_\varepsilon(f,\lambda_\varepsilon^k)$ follows from (5.1) (5.3) since

$$W_p^p(f,\nu_\varepsilon) \ \ge \ \sum_{i\notin I_{i,\varepsilon}} \int_{A_{i,\varepsilon}} |x-x_{i,\varepsilon}|^p\, f(x)\,dx + \sum_{i\in I_{i,\varepsilon}}\sum_j \int_{A_{i,\varepsilon}^j} |x-y_{i,\varepsilon}^j|^p\, f(x)\,dx \ \ge \ W_p^p(f,\nu_\varepsilon^k) \ .$$

In order to prove ii), it is enough to consider $\tilde\nu_\varepsilon := \sum_i \tilde c_{i,\varepsilon}\,\delta_{x_{i,\varepsilon}}$, with $\tilde c_{i,\varepsilon} := \int_{A_{i,\varepsilon}} \tilde f\,dx$. Let T_ε be an optimal transport map for $W_p(f,\nu_\varepsilon)$. Then clearly $\tilde\nu_\varepsilon = T_\varepsilon^\sharp(\tilde f)$ and therefore $E_\varepsilon(\tilde f,\tilde\lambda_\varepsilon) \le E_\varepsilon(f,\lambda_\varepsilon)$. Moreover $\int_{\Omega\times\mathbb{R}^+} |\tilde\lambda_\varepsilon - \lambda_\varepsilon| \ = \ \sum_i |\tilde c_{i,\varepsilon} - c_{i,\varepsilon}| = \int_\Omega |f-\tilde f|\,dx$. □

Lemma 5.2. *Let $\rho_\varepsilon^x, \rho^x : \Omega \mapsto \mathcal{P}(\mathbb{R}^+)$ be Lebesgue measurable functions and let $f \in L^1(\Omega,\mathbb{R}^+)$. Then the following assertion are equivalent:*

$$i) \qquad \forall\varphi\in \mathrm{BL}(\mathbb{R}^+)\ ,\ \lim_{\varepsilon\to 0}\int_\Omega |\langle \rho_\varepsilon^x - \rho^x\ ,\ \varphi\rangle|\ f(x)\,dx = 0$$

$$ii) \qquad \lim_{\varepsilon\to 0}\int_\Omega N(\rho_\varepsilon^x - \rho^x)\, f(x)\,dx = 0\ .$$

Proof. The only non trivial part is $i) \Rightarrow ii)$. Let $\{\varphi_n\} \subset \mathrm{BL}(\mathbb{R}^+)$ a dense subset in the unit ball of $C_0(\mathbb{R}^+)$. Then $d(\rho,\rho') := \sum_{n=0}^\infty 2^{-n}|\langle \rho-\rho',\varphi_n\rangle|$ is a distance on $\mathcal{P}(\mathbb{R}^+)$ which (like $N(\rho-\rho')$) induces the topology of tight convergence . From i), it is easy to check that $\lim_{\varepsilon\to 0}\int_\Omega d(\rho_\varepsilon^x,\rho^x)\,f(x)\,dx = 0$. Then we chose a subsequence ε' of ε such that: $\rho_{\varepsilon'}^x \to \rho^x$ tightly a.e. on $\{f>0\}$ and $\lim_{\varepsilon'\to 0}\int_\Omega N(\rho_{\varepsilon'}^x - \rho^x)\,f(x)\,dx \ = \ \limsup_{\varepsilon\to 0}\int_\Omega N(\rho_\varepsilon^x-\rho^x)\,f(x)\,dx$. The conclusion follows thanks to dominated convergence Theorem.

□

5.1. **Compactness.** We prove the part i) of Theorem 3.1. Let $\{\lambda_\varepsilon\}$ any sequence such that $C := \sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon) < +\infty$. Thanks to Prokhorov's Theorem (see [16], chapter 11), the following condition is sufficient to get the τ-relative compacity of $\{\lambda_\varepsilon\}$:

$$\lim_{k\to+\infty}\limsup_{\varepsilon\to 0}\lambda_\varepsilon(\Omega\times[k,+\infty[) = 0. \tag{5.5}$$

We estimate $\lambda_\varepsilon(\Omega\times[k,+\infty[)$ for any fixed $k>0$ and $\varepsilon>0$ as follows. Let $\nu_\varepsilon = \sum_{i\in I_\varepsilon} c_{i,\varepsilon}\delta_{x_{i,\varepsilon}}$ be associated with λ_ε and a partition $\{A_{i,\varepsilon},\ i\in I_\varepsilon\}$ such that

$$W_p^p(f,\pi_1^\sharp\lambda_\varepsilon) = \sum_{i\in I_\varepsilon}\int_{A_{i,\varepsilon}} |x-x_{i,\varepsilon}|^p\ f(x)dx\ ,$$

For any fixed $M>0$, we set:

$$I_\varepsilon^M := \left\{ i \in I_\varepsilon : \ \fint_{A_{i,\varepsilon}} f(x)\ dx \le M \right\} \ , \ \lambda_\varepsilon^M := \sum_{i\in I_\varepsilon^M} c_{i,\varepsilon}\, \delta_{\left(x_{i,\varepsilon}, \frac{c_{i,\varepsilon}}{\varepsilon^d}\right)} \ , \ \Omega_\varepsilon^M := \bigcup_{i\in I_\varepsilon^M} A_{i,\varepsilon} \ .$$

We notice that:

$$|\Omega\backslash\Omega_\varepsilon^M| < \frac{1}{M}\int_\Omega f(x)\ dx \quad , \quad c_{i,\varepsilon} \le M|A_{i,\varepsilon}| \quad \forall i \in I_\varepsilon^M. \tag{5.6}$$

The following inequalities hold:

$$\begin{aligned} \lambda_\varepsilon(\Omega\times[k,+\infty[) &\le \lambda_\varepsilon^M(\Omega\times[k,+\infty[) + \lambda_\varepsilon(\Omega\backslash\Omega_\varepsilon^M\times\mathbb{R}^+) \\ &= \int_{\Omega\times[k,+\infty[} d\lambda_\varepsilon^M(x,t) + \int_{\Omega\backslash\Omega_\varepsilon^M} f(x)\ dx \\ &\le \frac{1}{k^{p/d}}\int_{\Omega\times\mathbb{R}^+} t^{p/d}\ d\lambda_\varepsilon^M(x,t) + \eta(M)\ , \end{aligned} \tag{5.7}$$

where $\eta(M)$ vanishes as $M\to+\infty$ since, by (5.6), the measure of $\Omega\backslash\Omega_\varepsilon^M$ goes uniformly to 0. In order to majorize the integral in the right hand member of (5.7), we use (2.12) and (5.6). Recalling that $f\ge\alpha$ and exploiting (5.6), we have for every $i\in I_\varepsilon^M$:

$$\int_{A_{i,\varepsilon}} \frac{|x-x_{i,\varepsilon}|^p}{\varepsilon^p}\ f(x)dx \quad \ge \quad \alpha\,\gamma_{p,d}\,\frac{|A_{i,\varepsilon}|^{1+\frac{p}{d}}}{\varepsilon^p} \ \ge\ \frac{\alpha\,\gamma_{p,d}}{M^{1+\frac{p}{d}}}\,\frac{c_{i,\varepsilon}^{1+\frac{p}{d}}}{\varepsilon^p}\ ,$$

so that, summing with respect to $i\in I_\varepsilon^M$:

$$\begin{aligned} E_\varepsilon(f,\lambda_\varepsilon) \ \ge \sum_{i\in I_\varepsilon^M}\int_{A_{i,\varepsilon}} \frac{|x-x_{i,\varepsilon}|^p}{\varepsilon^p}\ f(x)dx \quad &\ge\quad \frac{\alpha\,\gamma_{p,d}}{M^{1+\frac{p}{d}}}\sum_{i\in I_\varepsilon^M}\left(\frac{c_{i,\varepsilon}}{\varepsilon^d}\right)^{p/d} \\ &\ge\quad \frac{\alpha\,\gamma_{p,d}}{M^{1+\frac{p}{d}}}\int_{\Omega\times\mathbb{R}^+} t^{p/d}\ d\lambda_\varepsilon^M(x,t)\ . \end{aligned} \tag{5.8}$$

Recording $C:=\sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon)$, by (5.7) and (5.8), we get $\limsup_{\varepsilon\to0}\lambda_\varepsilon(\Omega\times[k,+\infty[) \le \dfrac{C'}{k^{\frac{p}{d}}} + \eta(M)$ with $C' = \dfrac{C}{\alpha\,\gamma_{p,d}}\,M^{1+\frac{p}{d}}$. The claim (5.5) follows by sending k and M to infinity. □

5.2. **Lower bound inequality.** We will now prove the Γ-*liminf inequality*, namely, for every sequence $\lambda_\varepsilon\in\mathcal{M}(\Omega\times\mathbb{R}^+)$ converging tightly to a measure λ in $\mathcal{M}(\Omega\times\mathbb{R}^+)$ we will show that

$$\liminf_{\varepsilon\to0} E_\varepsilon(f\ ,\ \lambda_\varepsilon) \ge E(f,\lambda). \tag{5.9}$$

Without loss of generality we will always assume that

$$\sup_{\varepsilon>0} E_\varepsilon(f,\lambda_\varepsilon) \le C < \infty. \tag{5.10}$$

The proof will be accomplished in two steps.

Step 1: We will prove (5.9) under the assumption that f has a continuous representative and that

$$\exists M>0 \quad : \quad \mathrm{supp}(\lambda_\varepsilon)\subset\Omega\times[0,M] \quad \forall\varepsilon>0. \tag{5.11}$$

From (5.10), it follows that $\lambda_\varepsilon = \lambda_\varepsilon(\nu_\varepsilon)$ for some $\nu_\varepsilon := \sum_i c_{i,\varepsilon}\delta_{x_{i,\varepsilon}}$ with support in Ω and that $W_p^p(f\,dx\,,\,\nu_\varepsilon)\ \le\ C\,\varepsilon^p$. In particular, $\nu_\varepsilon \stackrel{*}{\rightharpoonup} f dx$ and the first marginal of λ coincides with f.

Therefore we may write λ as $\lambda = f \otimes \rho^x$ for a suitable family $\{\rho^x\}$ of probability measures in $\mathcal{P}(\mathbb{R}^+)$ (see Proposition 2.3). Let $T_\varepsilon = \sum_i x_{i,\varepsilon} 1\!\!1_{A_{i,\varepsilon}}$ an optimal transport so that

$$W_p^p(f\,dx\,,\,\nu_\varepsilon) = \sum_i \int_{A_{i,\varepsilon}} |x - x_{i,\varepsilon}|^p f(x)\,dx \quad \text{and} \quad \int_{A_{i,\varepsilon}} f(x)\,dx = c_{i,\varepsilon}\ .$$

We introduce the following sequence of measures:

$$m_\varepsilon := \frac{|T_\varepsilon x - x|^p}{\varepsilon^p} f(x)\,dx. \tag{5.12}$$

In view of (5.10) it follows that m_ε is a bounded sequence of measures on Ω and, without loss of generality, we may assume that m_ε weakly-$*$ converges to a measure $m := m_a dx + m_s$. Here, we have denoted by m_a the density of the absolutely continuous part of m with respect to the Lebesgue measure and by m_s its singular part. So,

$$\liminf_{\varepsilon\to 0} E_\varepsilon(f, \lambda_\varepsilon) \geq \int_\Omega m_a(x)\,dx.\ ,$$

We are reduced to show that

$$m_a(x) \geq f(x)^{1-p/d}\, G(\rho^x) \quad \text{a.e. } x \in \Omega. \tag{5.13}$$

Let $x_0 \in \Omega$. Here and in the following $Q_r(x_0) = x_0 + rQ_1$ is the cube of size r centered at x_0. Sometimes we will omit the subscript x_0. Fix x_0 to be a Lebesgue point of the density m_a, i.e. $m_a(x_0) = \lim_{r\to 0} \dfrac{m(Q_r(x_0))}{r^d}$. Let $r > 0$ be such that $m_\varepsilon(Q_r(x_0)) \to m(Q_r(x_0))$ as $\varepsilon \to 0$ (this is true for all but countably many $r > 0$). By the continuity of f there exists a sequence $\gamma_r \in]0, 1[$ such that $\gamma_r \nearrow 1$ as $r \to 0$ and

$$\gamma_r\, k_0 \ \leq f(x) \ \leq \ \frac{1}{\gamma_r}\, k_0 \quad \forall x \in Q_r(x_0) \quad \text{where} \quad k_0 := f(x_0)\ . \tag{5.14}$$

By (5.12) and (5.14), we have

$$\frac{m_\varepsilon(Q_r(x_0))}{r^d} \ \geq \ \frac{\gamma_r}{r^d \varepsilon^p} \left(\sum_i \int_{Q_r(x_0)\cap A_{i,\varepsilon}} |x - x_{i,\varepsilon}|^p \ k_0\,dx \right). \tag{5.15}$$

We may now rewrite the right hand side of (5.15) in terms of the cost of transporting the density $k_0 1\!\!1_{Q_r}$ onto a new measure $\nu_\varepsilon^r(x_0)$ (we will often omit x_0). Denoting $x_{i,\varepsilon}^r$ the projection of $x_{i,\varepsilon}$ on $Q_r(x_0)$, we set

$$\nu_\varepsilon^r(x_0) := \sum_i \nu_{i,\varepsilon}^r \delta_{x_{i,\varepsilon}^r} \quad \text{where} \quad \nu_{i,\varepsilon}^r := k_0\, |A_{i,\varepsilon} \cap Q_r|\ . \tag{5.16}$$

Thus, we get

$$\frac{m_\varepsilon(Q_r(x_0))}{r^d} \ \geq \ \frac{\gamma_r}{r^d \varepsilon^p}\ W_p^p(k_0\, 1\!\!1_{Q_r}, \nu_\varepsilon^r). \tag{5.17}$$

As by (5.10) the left hand side of (5.17) remains bounded, we find that, for every r, the sequence $\{\lambda_\varepsilon(\nu_\varepsilon^r)\}_\varepsilon$ is tight as well as its second marginal $\{\rho_\varepsilon(\nu_\varepsilon^r)\}_\varepsilon$. Let $\bar\rho^r$ be a tight limit point of the latter sequence. By passing to the limit in (5.17) with the help of estimate (4.6), it follows that that

$$\frac{m(Q_r(x_0))}{r^d} \ \geq \ \liminf_{\varepsilon\to 0} \gamma_r \frac{1}{r^d \varepsilon^p} W_p^p(k_0 1\!\!1_{Q_r}, \nu_\varepsilon^r) \geq \ k_0^{1-\frac{p}{d}}\, G(\bar\rho^r)\ . \tag{5.18}$$

As the left hand side of (5.18) converges as $r \to 0$ to a finite limit $m_a(x_0)$, we infer from the lower bound in (3.10) that the sequence $\{\bar\rho^r\}$ is tight and converges tightly to some $\bar\rho$. Thus by the lower semicontinuity of G and (5.18)

$$m_a(x_0) \;=\; \lim_{r\to 0} \frac{m(Q_r(x_0))}{r^d} \;\geq\; k_0^{1-\frac{p}{d}}\, G(\bar\rho)\ . \tag{5.19}$$

The claim (5.13) follows from (5.19) once we can prove that the unique limit point is $\bar\rho = \rho^{x_0}$. This is the goal of the following lemma which achieves our first step.

Lemma 5.3. *Assume that x_0 is a Lebesgue point for f and for the map $x \mapsto f(x)\rho^x$. Then the sequence $\{\rho_\varepsilon(\nu_\varepsilon^r)(x_0)\}$ satisfies*

$$\lim_{r\to 0}\left(\lim_{\varepsilon\to 0}\ N\Big(\rho_\varepsilon(\nu_\varepsilon^r)(x_0) - \rho^{x_0}\Big)\right) \;=\; 0\ .$$

Proof. : Let us introduce the local statistic of ν_ε on Q_r

$$\rho_\varepsilon^r(x_0) \;:=\; \frac{1}{\nu_\varepsilon(Q_r(x_0))} \sum_{x_{i,\varepsilon}\in Q_r(x_0)} c_{i,\varepsilon}\, \delta_{\frac{c_{i,\varepsilon}}{\varepsilon^d}}\,. \tag{5.20}$$

Notice that by (5.11), $\rho_\varepsilon^r(x_0)$ remains with compact support in $[0,M]$. As $\lambda_\varepsilon(\nu_\varepsilon)$ converges tightly to $\lambda = f\otimes\rho^x$ which has no mass on $\partial Q_r(x_0)\times\mathbb{R}^+$, we infer that for every $r>0$, $\rho_\varepsilon^r(x_0)$ does converge tightly to the the averaged $\rho^r(x_0)$ given by

$$<\rho^r(x_0),\varphi> \;=\; \frac{\int_{Q_r(x_0)} <\rho^x,\varphi>\ f(x)dx}{\int_{Q_r(x_0)} f(x)\,dx}\ .$$

By hypotheses x_0 is a Lebesgue point for $\rho^x f(x)$ and f, so $\rho^r(x_0)$ converges itself to ρ^{x_0} as $r\to 0$.

Let us focus now on the sequence $\{\rho_\varepsilon(\nu_\varepsilon^r)\}$. As noticed before it is tight and without loss of generality we may suppose that $\rho_\varepsilon(\nu_\varepsilon^r)$ converges tightly to some probability $\bar\rho^r$ as $\varepsilon\to 0$, and then that $\bar\rho^r$ converges to $\bar\rho$ as $r\to 0$. In order to prove the lemma, we need only to check that the probability measures $\bar\rho$ and ρ^{x_0} coincide and for that it is enough to check that

$$\rho^{x_0}(]a,b[)\ \leq \bar\rho([a,b]) \qquad \text{whenever } 0\leq a<b. \tag{5.21}$$

Given a,b as above, we consider arbitrary $a'<a$ and $0<r'<r$. For every such a',r',r we will establish the inequality:

$$\rho^{r'}(]a,b[) \quad\leq\quad \frac{1}{\gamma_r^2}\,(\frac{r}{r'})^d \quad \bar\rho^r([a'\gamma_r, \frac{b}{\gamma_r}]). \tag{5.22}$$

The inequality (5.21) will follow by letting $a'\to a$, $r'\to r$ and $r\to 0$ successively. Let us define (for simplicity of notation, we skip the dependence with respect to a,a',b,r,r'):

$$\begin{aligned}
I_\varepsilon &= \{i\ :\ x_{i,\varepsilon}\in Q_{r'}\ :\ a < \frac{c_{i,\varepsilon}}{\varepsilon^d} < b\}\ ,\\
J_\varepsilon &= \{i\in I_\varepsilon\ :\ a' \leq \frac{c_{i,\varepsilon}^r}{\varepsilon^d} < b\}\ ,\quad \text{where } c_{i,\varepsilon}^r := \int_{A_{i,\varepsilon}\cap Q_r} f(x)\,dx\ ,\\
K_\varepsilon &= \{i\in I_\varepsilon\ :\ \frac{c_{i,\varepsilon}-c_{i,\varepsilon}^r}{\varepsilon^d}\ \geq\ a-a'\}\ .
\end{aligned}$$

Noticing that $|x-x_{i,\varepsilon}|\geq \frac{r-r'}{2}$ whenever $x_{i,\varepsilon}\in Q_{r'}$ and $x\in A_{i,\varepsilon}\setminus Q_r$, the following inequality holds

$$\forall i\in I_\varepsilon \quad c_{i,\varepsilon}\ \leq\ c_{i,\varepsilon}^r\ +\ \frac{2^p\, m_{i,\varepsilon}}{|r-r'|^p}\ , \qquad \text{where}\quad m_{i,\varepsilon} := \int_{A_{i,\varepsilon}} |x-x_{i,\varepsilon}|^p\, f(x)\,dx\ . \tag{5.23}$$

In particular there holds $\sum_i m_{i,\varepsilon} \geq\ 2^{-p}\ (a-a')\,\varepsilon^d\,|r-r'|^p\,\sharp(K_\varepsilon)$. By the assumptions (5.10) and (5.11), we have

$$\sum_i m_{i,\varepsilon}\ \leq\ C\,\varepsilon^p \quad , \quad c_{i,\varepsilon}\ \leq\ M\,\varepsilon^d\ .$$

Observe that, for $i \in K_\varepsilon$, we have $c_{i,\varepsilon} \leq \frac{M 2^p m_{i,\varepsilon}}{|r-r'|^p(a-a')}$. Therefore summing subsequently over J_ε and K_ε, we obtain:

$$\sum_{i\in J_\varepsilon} c_{i,\varepsilon}\ \leq\ \sum_{i\in J_\varepsilon} c^r_{i,\varepsilon}\ +\ \frac{C\,2^p\,\varepsilon^p}{|r-r'|^p} \quad , \quad \sum_{i\in K_\varepsilon} c_{i,\varepsilon}\ \leq\ \frac{M\,C\,2^p\,\varepsilon^p}{(a-a')\,|r-r'|^p}\ . \tag{5.24}$$

On the other hand, by (5.14), we have

$$\gamma_r\,c^r_{i,\varepsilon}\ \leq\ \nu^r_{i,\varepsilon}\ \leq\ \ \gamma_r^{-1}\,c^r_{i,\varepsilon} \quad (\text{thus for all } i\in J_\varepsilon\ ,\ a'\,\gamma_r\ \leq\ \frac{\nu^r_{i,\varepsilon}}{\varepsilon^d}\ \leq\ \frac{b}{\gamma_r})\ . \tag{5.25}$$

Eventually we observe that $I_\varepsilon \subset J_\varepsilon \cup K_\varepsilon$ and therefore by (5.24)(5.25):

$$\sum_{i\in I_\varepsilon} c_{i,\varepsilon}\ \leq\ \gamma_r^{-1}\sum\left\{\nu^r_{i,\varepsilon}\ :\ a'\gamma_r \leq \frac{\nu^r_{i,\varepsilon}}{\varepsilon^d} \leq \frac{b}{\gamma_r}\right\}\ +\ C'\,\varepsilon^p\ ,$$

where C' is a suitable constant. Recalling the definitions (5.16) and (5.20), we may rewrite the latter inequality as

$$\rho^{r'}_\varepsilon(]a,b[)\ \nu_\varepsilon(Q_{r'}) \quad \leq \quad \gamma_r^{-1}\,k_0\ r^d\ \rho_\varepsilon(\nu^r_\varepsilon)\left([a'\gamma_r,\frac{b}{\gamma_r}]\right)\ +\ C'\,\varepsilon^p. \tag{5.26}$$

Taking into account that $\lim_{\varepsilon\to 0}\nu_\varepsilon(Q_{r'}) = \int_{Q_{r'}} f\,dx \geq \gamma_r k_0\,{r'}^d$, we deduce the inequality (5.22) by passing to the limit in (5.26) as $\varepsilon \to 0$. The proof of Lemma 5.3 is achieved. $\square$

Step 2: We now relax the continuity condition on f and the hypothese (5.11) of Step 1. We choose a sequence of continuous functions $\{f_n\}$ such that $f_n \leq f$ for all n and $\int_\Omega |f-f_n|dx \to 0$ as $n\to\infty$. By the second assertion of Lemma 5.1, the exists a new sequence $\{\lambda^n_\varepsilon\}$ such that:

$$E_\varepsilon(f,\lambda_\varepsilon)\ \geq\ E_\varepsilon(f_n,\lambda^n_\varepsilon) \quad , \quad \int|\lambda^n_\varepsilon-\lambda_\varepsilon| = \int_\Omega |f_n-f|\,dx\ . \tag{5.27}$$

Now by the first assertion of Lemma 5.1, for every $k>0$, we can modify the latter sequence in a sequence $\{\lambda^{n,k}_\varepsilon\}$ so that

$$\mathrm{supp}\lambda^{n,k}_\varepsilon \subset \Omega\times[0,k] \quad , \quad E_\varepsilon(f_n,\lambda^n_\varepsilon)\ \geq\ E_\varepsilon(f_n,\lambda^{n,k}_\varepsilon) \quad , \quad \int|\lambda^n_\varepsilon-\lambda^{n,k}_\varepsilon| \leq 2\int_{\Omega\times[k,+\infty[} d\lambda^n_\varepsilon\ . \tag{5.28}$$

From (5.27) and (5.28), we infer that $C \geq\ E_\varepsilon(f,\lambda_\varepsilon)\ \geq\ E_\varepsilon(f_n,\lambda^n_\varepsilon) \geq\ E_\varepsilon(f_n,\lambda^{n,k}_\varepsilon)$. Thus, by the assertion i) of Theorem 3.1, for all n and k, the sequences $\{\lambda^n_\varepsilon\}$ are tight and, up to a subsequence, converge to some $\lambda^{n,k}$ and λ, respectively. In addition we have:

$$\int|\lambda^{n,k}-\lambda| \leq \liminf_\varepsilon \int|\lambda^{n,k}_\varepsilon-\lambda_\varepsilon|\ \leq\ 2\int_{\Omega\times[k,+\infty[} d\lambda^n + \int_\Omega |f_n-f|\,dx\ .$$

We are allowed to apply Step 1 to the sequence $\{(f_n,\lambda^{n,k}_\varepsilon)\}$ as it satisfies the continuity assumption and the condition (5.11). We obtain that for every n,k:

$$\liminf_{\varepsilon\to 0}\ E_\varepsilon(f,\lambda_\varepsilon)\ \geq\ \liminf_{\varepsilon\to 0}\ E_\varepsilon(f_n,\lambda^{n,k}_\varepsilon)\ \geq E(f_n,\lambda^{n,k})\ .$$

Next we apply the strong-weak lower semicontinuity result of Lemma 4.6 to the pair $(f_n, \lambda^{n,k})$ which converges to (f, λ) as $n, k \to \infty$. It follows that

$$\liminf_{\varepsilon \to 0} E_\varepsilon(f, \lambda_\varepsilon) \ \geq \ \liminf_{n,k \to \infty} E(f_n, \lambda^{n,k}) \ \geq \ E(f, \lambda) \ .$$

This concludes the proof of the lower bound inequalty. $\square$

5.3. **Proof of the upper bound.** In this subsection, we prove the Γ-*limsup inequality*, namely that for every $\lambda \in \mathcal{M}(\Omega \times \mathbb{R}^+)$ there exists a sequence of measures $\{\lambda_\varepsilon\}$ such that

$$\lambda_\varepsilon \to \lambda \quad \text{tightly in } \mathcal{M}(\Omega \times \mathbb{R}^+) \quad \text{and} \quad \limsup_{\varepsilon \to 0} \ E_\varepsilon(f \, , \, \lambda_\varepsilon) \leq E(f, \lambda). \tag{5.29}$$

Without loss of generality, we may assume that $E(f, \lambda)$ is finite and consequently, the first marginal of λ is f and λ can be decomposed as $\lambda = f \otimes \rho^x$. Recalling the scale invariance property (3.7), it is convenient to rewrite the expected limit energy given in (3.6) as

$$E(f, \lambda) \ = \ \int_\Omega G(\hat{\rho}^x) \, f(x) \, dx \qquad \text{where} \qquad \hat{\rho}^x \ := \ L^\sharp_{f^{-1}(x)} \, \rho^x \ . \tag{5.30}$$

Here, as later in the paper, L_t denotes for every t the dilatation of factor t in $\mathbb{R}^+$.

Step 1 *(Construction of the approximating sequence).* We consider a partition of $\mathbb{R}^d$ made of cubes $\{Q_{j,\varepsilon} : j \in \mathbb{N}\}$ of size ε'. In the limit process as $\varepsilon \to 0$, it is important that ε' is of same oder as ε (in order to control $\varepsilon^{-1} W_p(f, \widetilde{f}_\varepsilon)$ for $\widetilde{f}_\varepsilon$ being a piecewise constant approximation of f). Therefore we choose $\varepsilon' = k\varepsilon$, where k is an arbitrary large integer designed to tend to infinity afterwards. Then we set

$$I_\varepsilon := \{j \in \mathbb{N} : |\Omega \cap Q_{j,\varepsilon}| > 0\} \ , \ J_\varepsilon := \{j \in I_\varepsilon : |Q_{j,\varepsilon} \setminus \Omega| = 0 \ \} \ , \ K_\varepsilon := I_\varepsilon \setminus J_\varepsilon \ .$$

Clearly the Lebesgue measure of ${\Omega'}_\varepsilon = \Omega \setminus \Omega_\varepsilon$ where $\Omega_\varepsilon := \cup_{j \in J_\varepsilon} Q_{j,\varepsilon}$ tends to 0 as $\varepsilon \to 0$ We introduce the piecewise constant approximants of f and λ:

$$\widetilde{f}_\varepsilon \ = \ \sum_{I_\varepsilon} \ f_{j,\varepsilon} \, 1\!\!1_{Q_{j,\varepsilon} \cap \Omega} \quad , \quad \widetilde{\lambda}_\varepsilon \ = \ \sum_{I_\varepsilon} \ f_{j,\varepsilon} \, 1\!\!1_{Q_{j,\varepsilon} \cap \Omega} \otimes \rho_{j,\varepsilon} \ , \tag{5.31}$$

where we have set

$$f_{j,\varepsilon} := \frac{1}{|Q_{j,\varepsilon} \cap \Omega|} \int_{Q_{j,\varepsilon} \cap \Omega} \ f(x) \, dx \quad , \quad \rho_{j,\varepsilon} := \frac{1}{\int_{Q_{j,\varepsilon} \cap \Omega} f(x) \, dx} \int_{Q_{j,\varepsilon} \cap \Omega} \ \rho^x \, f(x) \, dx \ .$$

We search an approximating sequence $\{\nu_\varepsilon\}$ of the form $\nu_\varepsilon \ = \ \sum_{j \in I_\varepsilon} \nu_{j,\varepsilon}$ where each $\nu_{j,\varepsilon}$ belongs to $\mathcal{M}_0^+(Q_{j,\varepsilon} \cap \Omega)$ and satisfies $\nu_{j,\varepsilon}(Q_{j,\varepsilon}) \ = \ \int_{Q_{j,\varepsilon} \cap \Omega} f \, dx$. The choice of $\nu_{j,\varepsilon}$ for $j \in K_\varepsilon$ will not be relevant in the final estimate since the measure of ${\Omega'}_\varepsilon$ vanishes as $\varepsilon \to 0$. For $j \in J_\varepsilon$, we consider

$$\nu_{j,\varepsilon} \ = \ f_{j,\varepsilon} \, \varepsilon^d \ R^\sharp_{j,\varepsilon}(\mu_{j,\varepsilon}) \quad , \quad \mu_{j,\varepsilon} \in \mathcal{M}_0^+(Q_k) \ , \ \int \mu_{j,\varepsilon} = k^d \ , \tag{5.32}$$

where $R^\sharp_{j,\varepsilon}$ is the affine function mapping Q_k onto $Q_{j,\varepsilon}$. The discrete measure $\mu_{j,\varepsilon}$ has to be selected in order that $\lambda_\varepsilon(\nu_\varepsilon)$ is close to the approximation $\widetilde{\lambda}_\varepsilon$ of λ while keeping the Wasserstein distance between ν_ε and $\widetilde{f}_\varepsilon$ as small as possible. To make this idea precise, let us consider a generic smooth test function on $\Omega \times \mathbb{R}^+$ of the kind $\psi(x) \, \varphi(t)$. Up to a small error vanishing as $\varepsilon \to 0$ and in the same way as for $\widetilde{f}_\varepsilon$, ψ can be substituted with a piecewise function $\widetilde{\psi}_\varepsilon$ taking value $\psi_{j,\varepsilon}$ on each $Q_{j,\varepsilon}$. Recalling (2.16), a simple computation shows that, for ν_ε and $\widetilde{\lambda}_\varepsilon$ by (5.31) (5.32), we have:

$$\begin{aligned}\left\langle \lambda\,,\,\widetilde{\psi}_\varepsilon(x)\,\varphi(t)\right\rangle &= \left\langle \widetilde{\lambda}_\varepsilon, \widetilde{\psi}_\varepsilon(x)\,\varphi(t)\right\rangle \;=\; \sum_{j\in I_\varepsilon} \left|Q_{j,\varepsilon}\cap\Omega\right| f_{j,\varepsilon}\,\psi_{j,\varepsilon}\,\langle\rho_{j,\varepsilon},\varphi\rangle \qquad (5.33)\\ \left\langle \lambda_\varepsilon(\nu_\varepsilon)\,,\,\widetilde{\psi}_\varepsilon(x)\,\varphi(t)\right\rangle &= \sum_{j\in I_\varepsilon} \left|Q_{j,\varepsilon}\cap\Omega\right| f_{j,\varepsilon}\,\psi_{j,\varepsilon}\left\langle L^\sharp_{f_{j,\varepsilon}}(\rho(\mu_{j,\varepsilon})),\varphi\right\rangle \qquad (5.34)\end{aligned}$$

Owing to (5.33)(5.34), in order that $\lambda_\varepsilon(\nu_\varepsilon)$ is close to λ, we need to choose $\rho(\mu_{j,\varepsilon})$ close to

$$\widehat{\rho}_{j,\varepsilon} := L^\sharp_{f^{-1}_{j,\varepsilon}}(\rho_{j,\varepsilon})\ . \qquad (5.35)$$

To that aim, for given $\delta>0$ and by using the definition (3.3) of the set function S_δ, we choose $\mu_{j,\varepsilon}\in\mathcal{M}^+_0(Q_k)$ for $j\in J_\varepsilon$ so that (this depends on k,δ):

$$W_p^p(\mathbb{1}_{Q_k},\mu_{j,\varepsilon}) + \frac{k^d}{\delta^p} N(\widehat{\rho}_{j,\varepsilon}-\rho(\mu_{j,\varepsilon})) \;\le\; S_\delta(\widehat{\rho}_{j,\varepsilon},Q_k) \;+\; \varepsilon\ . \qquad (5.36)$$

Summarizing, we have constructed a triple indexed sequence of admissible measures by setting

$$\lambda_{\varepsilon,k,\delta} := \lambda_\varepsilon(\nu_\varepsilon) \quad , \quad \nu_\varepsilon = \sum_{j\in I_\varepsilon}\nu_{j,\varepsilon} \quad \text{with } \nu_{j,\varepsilon} \text{ given by (5.32)(5.35)(5.36)} \qquad (5.37)$$

We will need the following

Lemma 5.4. *Let $\widehat{\rho}^x_\varepsilon:\Omega\mapsto\mathcal{P}(\mathbb{R}^+)$ be the piecewise constant function defined by $\widehat{\rho}^x_\varepsilon=\sum_{j\in I_\varepsilon}\widehat{\rho}_{j,\varepsilon}\,\mathbb{1}_{Q_{j,\varepsilon}}$. Then we have*

$$\lim_{\varepsilon\to0}\int_\Omega N(\widehat{\rho}^x_\varepsilon-\widehat{\rho}^x)\,f(x)\,dx \;=\; 0$$

Proof. As $f, f_{j,\varepsilon}\ge\alpha$, we apply the inequality (4.13) for every $\beta\ge\max\{1,\alpha^{-1}\}$: for every $x\in Q_{j,\varepsilon}$, one has

$$\begin{aligned}N(\widehat{\rho}^x_\varepsilon-\widehat{\rho}^x) &= N(L^\sharp_{f^{-1}_{j,\varepsilon}}(\rho_{j,\varepsilon})-L^\sharp_{f(x)^{-1}}(\rho^x))\\ &\le \beta\,|f^{-1}_{j,\varepsilon}-f(x)^{-1}| \;+\; 2\,\rho^x([\beta,+\infty[) \;+\; \beta\,N(\rho_{j,\varepsilon}-\rho^x)\ .\end{aligned}$$

Let us multiply by $f(x)$ and integrate over Ω. Setting $\rho^x_\varepsilon:=\sum_{j\in I_\varepsilon}\rho_{j,\varepsilon}\,\mathbb{1}_{Q_{j,\varepsilon}}$, we derive

$$\int_\Omega N(\widehat{\rho}^x_\varepsilon-\widehat{\rho}^x)\,f(x)\,dx \;\le\; \frac{\beta}{\alpha^2}\int_\Omega|f-\widetilde{f}_\varepsilon|\,dx+\frac{2}{\alpha}\lambda(\Omega\times[\beta,+\infty[)+\beta\int_\Omega N(\rho^x_\varepsilon-\rho^x)\,f(x)\,dx. \quad (5.38)$$

We are now reduced to show that

$$\lim_{\varepsilon\to0}\int_\Omega N(\rho^x_\varepsilon-\rho^x)\,f(x)\,dx=0. \qquad (5.39)$$

Indeed, passing to the limit as $\varepsilon\to0$ in (5.38), we obtain $\limsup_{\varepsilon\to0}\int_\Omega N(\widehat{\rho}^x_\varepsilon-\widehat{\rho}^x)\,f(x)\,dx \le 2\lambda(\Omega\times[\beta,+\infty[)$ and the conclusion follows by sending $\beta\to\infty$. To prove claim (5.39), we check the condition i) Lemma 5.2. Let $\varphi\in\mathrm{BL}(\mathbb{R}^+)$ and set $g(x):=\langle\rho^x,\varphi\rangle$. Then $\tilde{g}_\varepsilon(x):=\langle\rho^x_\varepsilon,\varphi\rangle$ coincides on each $Q_{j,\varepsilon}$ with the average of g on $Q_{j,\varepsilon}$ with respect to the weighted density $\mu=f(x)\,dx$. Therefore as well known, we have $0=\lim_{\varepsilon\to0}\int_\Omega|\tilde{g}_\varepsilon-g|\,\mu(dx)=\lim_{\varepsilon\to0}\int_\Omega|\langle\rho^x-\rho^x_\varepsilon,\varphi\rangle|\,f(x)\,dx$. □

Step 2 *(Upper bound estimate).* We show that

$$\limsup_{\delta\to0}\ \limsup_{k\to\infty}\ \limsup_{\varepsilon\to0}\ E_\varepsilon(f,\lambda_{\varepsilon,k,\delta}) \;\le\; E(f,\lambda)\ . \qquad (5.40)$$

Let $\gamma > 1$ and R $(= R_{\gamma,p})$ a constant such that $(a+b)^p \le \gamma\, a^p + R\, b^p$ holds for all $a, b \in \mathbb{R}^+$. By applying the triangle inequality to the distance W_p combined with the previous inequality and the subadditivity property (2.6), we get

$$\begin{aligned} E_\varepsilon(f, \lambda_{\varepsilon,k,\delta}) &\le \frac{W_p^p(f \llcorner \Omega_\varepsilon, \nu_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} + \frac{W_p^p(f \llcorner \Omega'_\varepsilon, \nu_\varepsilon \llcorner \Omega'_\varepsilon)}{\varepsilon^p} \\ &\le \gamma\, \frac{W_p^p(\tilde f_\varepsilon \llcorner \Omega_\varepsilon, \nu_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} + R\, \frac{W_p^p(f \llcorner \Omega_\varepsilon, \tilde f_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} + \frac{W_p^p(f \llcorner \Omega'_\varepsilon, \nu_\varepsilon \llcorner \Omega'_\varepsilon)}{\varepsilon^p} \ . \end{aligned} \tag{5.41}$$

The last two terms in the right hand side of (5.41) vanish in the limit as $\varepsilon \to 0$. Indeed by (2.8), we have, for every $j \in I_\varepsilon$, $W_p^p(f \llcorner Q_{j,\varepsilon}, \tilde f_\varepsilon \llcorner Q_{j,\varepsilon}) \le (\sqrt{d}k)^p \varepsilon^p \int_{Q_{j,\varepsilon}} |f - \tilde f_\varepsilon|$. Thus, by using (2.6) and (2.8), we obtain

$$\begin{aligned} \frac{W_p^p(f \llcorner \Omega_\varepsilon, \tilde f_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} &\le \sum_{j \in J_\varepsilon} \frac{W_p^p(f \llcorner Q_{j,\varepsilon}, \tilde f_\varepsilon \llcorner Q_{j,\varepsilon})}{\varepsilon^p} \le (\sqrt{d}k)^p \int_{\Omega_\varepsilon} |f - \tilde f_\varepsilon| \\ \frac{W_p^p(f \llcorner \Omega'_\varepsilon, \nu_\varepsilon \llcorner \Omega'_\varepsilon)}{\varepsilon^p} &\le \sum_{j \in K_\varepsilon} \frac{1}{\varepsilon^p} (\operatorname{diam} Q_{j,\varepsilon})^p \int_{\Omega \cap Q_{j,\varepsilon}} f dx \ \le\ (\sqrt{d}k)^p \int_{\Omega \setminus \Omega'_\varepsilon} f dx \end{aligned}$$

Therefore, from (5.41), by letting $\gamma \to 1$ after $\varepsilon \to 0$, we deduce that for every $k, \delta > 0$

$$\limsup_{\varepsilon \to 0} E_\varepsilon(f, \lambda_{\varepsilon,k,\delta}) \quad \le \quad \limsup_{\varepsilon \to 0} \frac{W_p^p(\tilde f_\varepsilon \llcorner \Omega_\varepsilon, \nu_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} \ . \tag{5.42}$$

Now we exploit our choice (5.36). By construction, we have

$$\begin{aligned} \frac{W_p^p(\tilde f_\varepsilon \llcorner \Omega_\varepsilon, \nu_\varepsilon \llcorner \Omega_\varepsilon)}{\varepsilon^p} &\le \frac{1}{\varepsilon^p} \sum_{j \in J_\varepsilon} W_p^p(f_{j,\varepsilon} \mathbb{1}_{Q_{j,\varepsilon}}, \nu_\varepsilon) \\ &= \frac{1}{\varepsilon^p} \sum_{j \in J_\varepsilon} \varepsilon^{p+d} f_{j,\varepsilon}\, W_p^p(\mathbb{1}_{Q_k}, \mu_{j,\varepsilon}) \\ &\le \sum_{j \in J_\varepsilon} \int_{Q_{j,\varepsilon}} \frac{S_\delta(\widehat\rho_{\varepsilon,j}, Q_k) + \varepsilon}{k^d}\, f(x)\, dx \ . \end{aligned}$$

Eventually in the last term, on each $Q_{j,\varepsilon}$ we substitute $\widehat\rho_{\varepsilon,j}$ with the function $\hat\rho^x$ defined in (5.30) majorizing the error thanks to Lipschitz estimate (4.2). The inequality (5.42) becomes

$$\limsup_{\varepsilon \to 0} E_\varepsilon(f, \lambda_{\varepsilon,k,\delta}) \ \le\ \int_\Omega \frac{S_\delta(\hat\rho^x, Q_k)}{k^d}\, f(x)\, dx \ +\ \frac{1}{\delta^p} \limsup_{\varepsilon \to 0} R_\varepsilon \ , \tag{5.43}$$

where, by Lemma 5.4, the remainder R_ε given by

$$R_\varepsilon \ (= R_{\varepsilon,k,\delta}) \ =\ \sum_{j \in J_\varepsilon} \int_{Q_{j,\varepsilon}} N(\hat\rho^x - \hat\rho_{j,\varepsilon})\, f(x)\, dx \ , \tag{5.44}$$

converges to zero as $\varepsilon \to 0$ (for k, δ being fixed). Then passing to the limit in (5.43) first as $k \to \infty$ using the dominated convergence theorem and eventually as $\delta \to 0$ using (3.5) and monotone convergence, we are led to

$$\limsup_{\delta \to 0}\ \limsup_{k \to \infty}\ \limsup_{\varepsilon \to 0}\ E_\varepsilon(f, \lambda_{\varepsilon,k,\delta}) \ \le\ \limsup_{\delta \to 0} \int_\Omega G_\delta(\hat\rho^x)\, f(x)\, dx \ \le\ \int_\Omega G(\widehat\rho^x)\, f(x)\, dx \ ,$$

which by (5.30) is nothing else but (5.40).

Step 3. *(tight convergence)* We show now that, for every $\Psi \in C_b(\Omega \times \mathbb{R}^+)$, there holds

$$\lim_{\delta\to 0}\lim_{k\to\infty}\lim_{\varepsilon\to 0}\ \big\langle \lambda_{\varepsilon,k,\delta}\,,\, \Psi(x,t)\big\rangle \ =\ \big\langle \lambda\,,\, \Psi(x,t)\big\rangle \tag{5.45}$$

By (5.40) and the assertion i) of Theorem 3.1 (already proved in section 5.1), the family $\{\lambda_{\varepsilon,k,\delta}\}$ is relatively compact for the tight topology. Thus , by using a density argument, it is not restrictive to assume that $\Psi(x,t) = \psi(x)\,\varphi(t)$ where ψ and φ are Lipschitz continuous. We may futher assume that $|\psi| \le 1$ and that $|\varphi|_\infty + \mathrm{Lip}(\varphi) \le 1$. We use the piecewise approximation $\widetilde{\psi}_\varepsilon$ of ψ and the relations (5.33)(5.34) in Step 3 to derive

$$\begin{aligned}\big|\langle \lambda_{\varepsilon,k,\delta} - \lambda\,,\, \psi(x)\,\varphi(t)\rangle\big| \ &\le\ 2\|f\|_{L^1(\Omega)}\,|\widetilde{\psi}_\varepsilon - \psi|_\infty + \sum_{j\in I_\varepsilon} \big|Q_{j,\varepsilon}\cap\Omega\big|\, f_{j,\varepsilon}\,\psi_{j,\varepsilon}\,\Big|\langle L^\sharp_{f_{j,\varepsilon}}(\rho(\mu_{j,\varepsilon})) - \rho_{j,\varepsilon}, \varphi\rangle\Big| \\ &\le\ C\,k\varepsilon\ +\ \sum_{j\in I_\varepsilon}\big|Q_{j,\varepsilon}\cap\Omega\big|\, f_{j,\varepsilon}\ N\left(L^\sharp_{f_{j,\varepsilon}}(\rho(\mu_{j,\varepsilon})) - \rho_{j,\varepsilon}\right) \end{aligned}\tag{5.46}$$

Here C denotes the Lipschitz constant of ψ and we have used the fact that $\int|\lambda_{\varepsilon,k,\delta} - \lambda| \le 2\|f\|_{L^1(\Omega)}$, $|\psi|$, $|\varphi|$ and $|\psi_{j,\varepsilon}|$ are smaller than 1 and the definition (2.2). Let us fix $\eta \ge 1$ and set

$$J^\eta_\varepsilon := \{j \in J_\varepsilon\ :\ f_{j,\varepsilon} \le \eta\} \quad , \quad \Omega^\eta_\varepsilon := \bigcup_{j\in J^\eta_\varepsilon} Q_{j,\varepsilon}\ .$$

By (5.35) and applying (4.12) with $\beta = \eta$ and $s = f_{j,\varepsilon}$, we are led to the following upper bound

$$N\left(L^\sharp_{f_{j,\varepsilon}}(\rho(\mu_{j,\varepsilon})) - \rho_{j,\varepsilon}\right) \le \begin{cases} \eta\ N\left(\rho(\mu_{j,\varepsilon}) - \widehat{\rho}_{j,\varepsilon}\right) & \text{if } j\in J^\eta_\varepsilon \\ 2 & \text{otherwise}\end{cases} \quad ,$$

so that (5.46) yields

$$\big|\langle \lambda_{\varepsilon,k,\delta} - \lambda\,,\, \psi(x)\,\varphi(t)\rangle\big|\ \le\ C\,k\varepsilon\ +\ 2\int_{\Omega\setminus\Omega^\eta_\varepsilon} f\,dx\ +\ \eta\sum_{j\in J^\eta_\varepsilon} k^d\varepsilon^d\, f_{j,\varepsilon}\ N\left(\rho(\mu_{j,\varepsilon}) - \widehat{\rho}_{j,\varepsilon}\right) \tag{5.47}$$

Now by (5.36) and (4.2), for all $j \in J_\varepsilon$ and $x \in Q_{j,\varepsilon}$, one has:

$$N(\widehat{\rho}_{j,\varepsilon} - \rho(\mu_{j,\varepsilon}))\ \le\ \frac{\delta^p}{k^d}(S_\delta(\widehat{\rho}_{j,\varepsilon}, Q_k)\ +\ \varepsilon)\ \le\ \delta^p\,\frac{S_\delta(\widehat{\rho}^x, Q_k)}{k^d}\ +\ N(\widehat{\rho}_{j,\varepsilon} - \widehat{\rho}^x) + \varepsilon\frac{\delta^p}{k^d},$$

so that recalling (5.44), and after multiplying by $f(x)$ and integrating on each $Q_{j,\varepsilon}$:

$$\sum_{j\in J^\eta_\varepsilon} k^d\varepsilon^d\, f_{j,\varepsilon}\ N\left(\rho(\mu_{j,\varepsilon}) - \widehat{\rho}_{j,\varepsilon}\right)\ \le\ \delta^p\int_\Omega \frac{S_\delta(\widehat{\rho}^x, Q_k)}{k^d}\,f(x)\,dx\ +\ R_{\varepsilon,k,\delta} + \varepsilon\frac{\delta^p}{k^d}\int_\Omega f\,dx \tag{5.48}$$

Passing first to the limit as $\varepsilon \to 0$, taking in to account (5.44), (5.47) and (5.48) , we get

$$\limsup_{\varepsilon\to 0}\big|\langle \lambda_{\varepsilon,k,\delta} - \lambda\,,\, \psi(x)\,\varphi(t)\rangle\big|\ \le\ \ \eta\,\delta^p\int_\Omega \frac{S_\delta(\widehat{\rho}^x, Q_k)}{k^d}\,f(x)\,dx\ +\ \limsup_{\varepsilon\to 0}\ 2\int_{\Omega\setminus\Omega^\eta_\varepsilon} f(x)\,dx \tag{5.49}$$

Now, by (4.2) and noticing that the function $S_\delta(\rho)$ vanishes for ρ being the Dirac mass at 0, , it is easy to check that the ratio $S_\delta(\widehat{\rho}^x, Q_k)\ k^{-d}$ remain bounded by $\frac{2}{\delta^p}$. Moreover it pointwise converges to $G_\delta(\widehat{\rho}^x)$ as $k\to\infty$ and since , by (5.30), $\int_\Omega G_\delta(\widehat{\rho}^x)\,f(x)\,dx \le \int_\Omega G(\widehat{\rho}^x)\,f(x)\,dx < +\infty$, by dominated convergence, we deduce from (5.49) that

$$\limsup_{\delta\to 0}\ \limsup_{k\to\infty}\ \limsup_{\varepsilon\to 0}\big|\langle \lambda_{\varepsilon,k,\delta} - \lambda\,,\, \psi(x)\,\varphi(t)\rangle\big|\ \le\ \limsup_{k\to\infty}\ \limsup_{\varepsilon\to 0}\ 2\int_{\Omega\setminus\Omega^\eta_\varepsilon} f(x)\,dx\ .$$

This being true whatever is $\eta > 1$, we obtain the claim (5.45) by noticing that $|\Omega \setminus \Omega_\varepsilon^\eta| \leq |\Omega'_\varepsilon| + \frac{1}{\eta}\int_\Omega f(x)\,dx$ which vanishes as $\varepsilon \to 0$ and $\eta \to \infty$.

Last step. Consider $\{\Psi_n\}$ a dense sequence in the unit ball of $C_b(\Omega \times \mathbb{R}^+)$ and set

$$a_{\delta,k,\varepsilon} := [E_\varepsilon(f, \lambda_{\varepsilon,k,\delta}) - E(f,\lambda)]^+ \ + \ \sum_{n=0}^{\infty} 2^{-n} \left|\langle \lambda_{\varepsilon,k,\delta} - \lambda \, , \, \Psi_n\rangle\right| .$$

By (5.40) and (5.45), we have $\limsup_{\delta\to 0} \limsup_{k\to\infty} \limsup_{\varepsilon \to 0} \ a_{\delta,k,\varepsilon} \ \leq \ 0$. Therefore by a classical diagonalization argument (see for instance [2], Corollary 1.16), we can choose sequences $\delta(\varepsilon), k(\varepsilon)$ such that $\delta(\varepsilon) \to 0, \ k(\varepsilon) \to +\infty$ and $a_{\delta(\varepsilon),k(\varepsilon),\varepsilon} \to 0$ as $\varepsilon \to 0$. We are led to

$$\limsup_{\varepsilon\to 0} E_\varepsilon\big(f, \lambda_{\delta(\varepsilon),k(\varepsilon),\varepsilon}\big) \ \leq \ E(f,\lambda) \quad , \quad \lim_{\varepsilon \to 0} \langle \lambda_{\delta(\varepsilon),k(\varepsilon),\varepsilon} - \lambda, \Psi_n\rangle = 0 \quad ,\forall n \, .$$

Then as $E(f,\lambda) < +\infty$, the sequence $\lambda_\varepsilon := \lambda_{\delta(\varepsilon),k(\varepsilon),\varepsilon}$ is τ relatively compact (by the assertion i) of Theorem3.1) and clearly admits λ as unique cluster point. We have therefore constructed a sequence $\{\lambda_\varepsilon\}$ which fulfills all requirements in (5.29). This concludes the proof of the upper bound inequality. The proof of Theorem 3.1 is now complete. □

6. Convergence of the infimum problem and optimality conditions.

In this section, we prove Theorem 3.5, Proposition 3.6, Lemma 3.10 and Proposition 3.11.

6.1. **Preliminary estimates.** For proving Theorem 3.5, we need some technical lemmas. Let us introduce introduce the upper semicontinuous enveloppe $\widetilde{V}$ of the non increasing potential V and $\widetilde{g}$ by setting

$$\widetilde{V}(t) := \lim_{s\to t, s<t} V(s) \quad \text{for } t>0 \quad , \quad \widetilde{V}(0) = V(0+) \quad , \quad \widetilde{g}(t) := \widetilde{V}(t) + \beta(t).$$

Lemma 6.1. *Let $\{\lambda_\varepsilon\}$ be a sequence such that λ_ε converges tightly to λ and $\sup_\varepsilon E_\varepsilon(f,\lambda_\varepsilon) < +\infty$ (respectively $E(f,\lambda) < +\infty$). Then the following convergences hold*

i) $\displaystyle\limsup_{\varepsilon\to 0} \int V(t)\, d\lambda_\varepsilon(x,t) \leq \int \widetilde{V}(t)\, d\lambda(x,t)$, *if V satisfies (3.12) and $V(0+) < +\infty$.*

ii) $\displaystyle\lim_{\varepsilon\to 0} \int \beta(t)\, d\lambda_\varepsilon(x,t) = \int \beta(t)\, d\lambda(x,t)$, *if β is continuous and satisfies (3.13).*

Proof. By the compactess statement in Theorem 3.1, the sequence $\{\lambda_\varepsilon\}$ converges tightly to λ. The assumptions on V imply that $\widetilde{V}$ is non negative, bounded and upper semicontinuous. Thus it can be written as the infimum of a family of bounded continuous functions and the inequality in assertion i) follows classically. The assertion ii) is straightforward if β is bounded (assumption (3.13.b). Otherwise, under (3.13.a), we have that $f \in L^\infty(\Omega)$ and, by exploiting Lemma 4.3

$$C \ := \ \sup_\varepsilon \int t^{p/d} d\lambda_\varepsilon(x,t) \ < +\infty.$$

Assume first that $\beta \geq 0$. Then one check easily that $\beta\,\lambda_\varepsilon \overset{*}{\rightharpoonup} \beta\lambda$ whereas

$$\eta(R) \ := \ \limsup_{\varepsilon\to 0} \int_{\Omega\times[R,+\infty[} \beta(t)\, d\lambda_\varepsilon(x,t) \ \leq \ C \ \sup\left\{ \frac{\beta(s)}{s^{p/d}} \ , \ s \geq R \right\}$$

converges to zero as $R \to \infty$. Therefore the convergence of $\{\beta(t)\,\lambda_\varepsilon\}$ is tight and ii) follows. The general case follows by decomposing β with respect to its positive and negative parts. □

Lemma 6.2. *Assume (3.11) with (3.12) (3.13) and let $\widetilde{g} = \beta + \widetilde{V}$. Then, for $c > \bar{c}$*

$$\inf \mathcal{P}^c \ = \ \widetilde{m}(c) \ := \ \inf \left\{ E(f,\lambda) \ : \ \int \widetilde{g}(t)\, d\lambda(x,t) \ < \ c \right\} \ .$$

Proof. Since $\widetilde{g} \geq g$, we clearly have $\widetilde{m}(c) \geq m(c) := \inf \mathcal{P}^c$. To show the converse inequality, we take $\delta > 0$ such that $c - \delta > \bar{c}$ and fix a competitor $\lambda = f \otimes \rho^x$ such that $\int g(t)\, d\lambda(x,t) \leq c - \delta$. For every $a > 1$, we consider the measure $\lambda_a := f \otimes L_a^\sharp(\rho^x)$. We claim that

$$\int g(t)\, d\lambda_a(x,t) \ = \ \int \widetilde{g}(t)\, d\lambda_a(x,t) \qquad \text{for a.e. } a \in]1,2] \ . \tag{6.1}$$

Indeed by Fubini's formula and since the discontinuity set $\{r > 0\}$ where $r(t) := \widetilde{g}(t) - g(t)$ is at most countable, one has

$$\int_1^2 \left(\int r(t)\, d\lambda_a(x,t) \right) da \ = \ \int_\Omega f(x) \left(\int_1^2 \langle \rho^x(dt), r(at) \rangle\, da \right) dx \ = \ \int_\Omega f(x) \left\langle \rho^x, \int_1^2 r(at)\, da \right\rangle dx \ = \ 0.$$

Owing to (6.1), we can select a sequence $a_n \searrow 1$ such that $\int g(t)\, d\lambda_{a_n}(x,t) \ = \ \int \widetilde{g}(t)\, d\lambda_{a_n}(x,t)$. Besides as V is non-increasing, we have that $g(a_n t) \leq V(t) + \beta(a_n t)$ and by dominated convergence

$$\limsup_{n\to\infty} \int \widetilde{g}(t)\, d\lambda_{a_n}(x,t) \ = \ \limsup_{n\to\infty} \int g(t)\, d\lambda_{a_n}(x,t) \ \leq \ \int g(t)\, d\lambda(x,t) \ < \ c,$$

whereas by (3.6) and (3.7): $\lim_{n\to\infty} E(f, \lambda_{a_n}) = \lim_{n\to\infty} a_n^{\frac{p}{d}} E(f,\lambda) = E(f,\lambda)$. Therefore we have $\widetilde{m}(c) \leq E(f,\lambda)$ holding true for every λ admissible for $\mathcal{P}^{c-\delta}$. Thus $\widetilde{m}(c) \leq m(c-\delta)$. The conclusion follows by letting $\delta \to 0$ thanks to the continuity of the convex function m which is finite on $]\bar{c}, +\infty[$ (see Lemma 3.4). $\square$

The aim of the following lemma is to remove small Dirac masses which increase too much the size constraint when $V(0+) = +\infty$.

Lemma 6.3. *Assume that V satisfy (3.12) with $V(0+) = +\infty$ and let λ such that $E(f,\lambda) < +\infty$. Then , for every $\gamma > 1$, there exists a sequence of positive numbers $\{t_N\}$ and a double indexed sequence $\{\lambda_\varepsilon^N\}$ such that λ_ε^N is supported in $\Omega \times [t_N, +\infty[$ and*

$$i) \quad \limsup_{N\to\infty} \limsup_{\varepsilon} \int |\lambda_\varepsilon^N - \lambda| \ = \ 0 \quad , \qquad ii) \quad \limsup_{\varepsilon} \left(\int V(t)\, d\lambda_\varepsilon^N(x,t) \right) \ \leq \ \int \widetilde{V}(t)\, d\lambda(x,t) \ , \ \forall N$$

$$iii) \quad \limsup_{N\to\infty} \left(\limsup_{\varepsilon} E_\varepsilon(f, \lambda_\varepsilon^N) \right) \ \leq \ \gamma\, E(f,\lambda) \ .$$

Proof. We start with a realizing sequence $\{\lambda_\varepsilon\}$ given by Theorem 3.1, i.e. such that

$$\lambda_\varepsilon = \lambda_\varepsilon(\nu_\varepsilon) \quad , \quad \nu_\varepsilon = \sum_{i\in I_\varepsilon} c_{i,\varepsilon} \delta_{x_{i,\varepsilon}} \quad , \quad \lambda_\varepsilon \to \lambda \quad \text{tightly} \quad , \quad \limsup_{\varepsilon} E_\varepsilon(f,\lambda_\varepsilon) \leq E(f,\lambda) \ .$$

We may assume that $\int \widetilde{V}(t)\, d\lambda(x,t) < +\infty$ (since otherwise ii) would be trivially fulfilled taking $\lambda_\varepsilon^N = \lambda_\varepsilon$). Then , by the monotonicity of V, there exists a real $t^* \geq 0$ such that

$$\widetilde{V}(t) = +\infty \quad \text{if } t \in [0,t^*] \quad , \quad \widetilde{V}(t) < +\infty \quad \text{if } t > t^* \quad , \quad \lambda(\Omega \times [0,t^*]) = 0. \tag{6.2}$$

We fix a sequence $\{t_N, N \in \mathbb{N}\}$ such that

$$t_N \searrow t^* \quad , \quad \lambda(\Omega \times \{t_N\}) = 0 \ , \quad \forall N \quad , \quad \eta_N := \lambda(\Omega \times [0,t_N]) \to 0 \ . \tag{6.3}$$

Construction of the new sequence: Let T_ε be the optimal transport map corresponding to $W_p(f,\nu_\varepsilon)$ and let $A_{i,\varepsilon} = T_\varepsilon^{-1}(x_{i,\varepsilon})$ be the transport region for $x_{i,\varepsilon}$. For any large N, we set

$$I_\varepsilon^N := \left\{ i \in I_\varepsilon : \frac{c_{i,\varepsilon}}{\varepsilon^d} > t_N \right\} , \quad P_\varepsilon^N := \{ x_{i,\varepsilon} : i \in I_\varepsilon^N \} , \quad A_\varepsilon^N := \cup \{ A_{i,\varepsilon} : i \in I_\varepsilon^N \} \ , \ B_\varepsilon^N := \Omega \setminus A_\varepsilon^N \ .$$

Here, P_ε^N represents the "good destinations" and A_ε^N the associated transport subregion whereas B_ε^N represents "the costly transport subregion" where the transport has to be modified. In our construction we need to consider a covering of Ω by cubes $\cup_{j\in K_\varepsilon} Q_{j,\varepsilon}$ whose sides are of length $k\,\varepsilon$ and parallel to the axes and having non-trivial intersection with Ω. Recalling that $\alpha = \min_\Omega f$, it will be useful for proving iii) to choose the size parameter k so large that

$$\alpha\, k^d \ > \ t^* \quad , \quad 1 + \frac{t^* \sqrt{\gamma}}{\alpha\, k^d - t^*} \ < \ \gamma. \tag{6.4}$$

For every $j \in K_\varepsilon$, we denote by $\widehat{x}_{j,\varepsilon}$ the center of $Q_{j,\varepsilon}$ and select a point $p_{j,\varepsilon}$ in P_ε^N among those which are at the nearest distance to $Q_{j,\varepsilon}$. Further, we split the cubes into two sub-families according to the volume fraction of the bad transport region

$$R_\varepsilon^N := \{ j \in K_\varepsilon : b_{j,\varepsilon}^N \le t_N \} \ , \ S_\varepsilon^N := \{ j \in K_\varepsilon : b_{j,\varepsilon}^N > t_N \} \qquad \text{where} \quad b_{j,\varepsilon}^N := \frac{1}{\varepsilon^d} \int_{Q_{j,\varepsilon}\cap B_\varepsilon^N} f\ dx \ .$$

The modified transport T_ε^N is obtained by sending each portion of the bad set $Q_{j,\varepsilon} \cap B_\varepsilon^N$ to the nearest good destination $p_{j,\varepsilon}$ if the volume fraction $b_{j,\varepsilon}^N$ is smaller than t_N, to the center $\widehat{x}_{j,\varepsilon}$ otherwise while keeping T_ε unchanged on the good set. More precisely,

$$T_\varepsilon^N(x) = \begin{cases} T_\varepsilon(x) & \text{if} \quad x \in A_\varepsilon^N \ , \\ p_{j,\varepsilon} & \text{if} \quad x \in Q_{j,\varepsilon} \cap B_\varepsilon^N \text{ and } j \in R_\varepsilon^N \ , \\ \widehat{x}_{j,\varepsilon} & \text{if} \quad x \in Q_{j,\varepsilon} \cap B_\varepsilon^N \text{ and } j \in S_\varepsilon^N . \end{cases} \tag{6.5}$$

The push forward of the measure $f\,dx$ through T_ε^N is a new discrete measure ν_ε^N whose main feature is that the mass transported on each good destination point in P_ε^N has been increased while small masses have been grouped on the centers of the cells $Q_{j,\varepsilon}$:

$$\nu_\varepsilon^N := \sum_{i\in I_\varepsilon^N} (c_{i,\varepsilon} + d_{i,\varepsilon})\, \delta_{x_{i,\varepsilon}} + \sum_{j \in S_\varepsilon^N} b_{j,\varepsilon}^N \varepsilon^d \delta_{\widehat{x}_{j,\varepsilon}} \ , \quad d_{i,\varepsilon} := \sum \left\{ b_{j,\varepsilon}^N \varepsilon^d \ : \ j \in R_\varepsilon^N \ , \ p_{j,\varepsilon} = x_{i,\varepsilon} \right\}. \tag{6.6}$$

Eventually we have constructed a new sequence $\{\lambda_\varepsilon^N\}$ where $\lambda_\varepsilon^N := \lambda_\varepsilon(\nu_\varepsilon^N)$. We are going to show that it satisfies all the requirements of the Lemma.

Proof of i): Taking into account (6.6), we have

$$\begin{aligned} \int_{\Omega\times\mathbb{R}^+} |\lambda_\varepsilon^N - \lambda_\varepsilon| \ = \ \int_\Omega |\nu_\varepsilon^N - \nu_\varepsilon| \ &\le \ \sum_{i \notin I_\varepsilon^N} c_{i,\varepsilon} + \sum_{i\in I_\varepsilon^N} d_{i,\varepsilon} + \sum_{j\in S_\varepsilon^N} b_{j,\varepsilon}^N \varepsilon^d \\ &= \ \lambda_\varepsilon\left(\Omega \times [0,t_N]\right) + \sum_{j \in K_\varepsilon} b_{j,\varepsilon}^N \varepsilon^d \\ &= \ \lambda_\varepsilon\left(\Omega \times [0,t_N]\right) + \int_{B_\varepsilon^N} f\,dx \ = \ 2\,\lambda_\varepsilon\left(\Omega \times [0,t_N]\right) \ . \end{aligned}$$

The conclusion follows by (6.3) since $\limsup_\varepsilon \lambda_\varepsilon\left(\Omega\times[0,t_N]\right) \le \lambda\left(\Omega \times [0,t_N]\right) = \eta_N$.

Proof of ii): Using the fact that V is non-increasing and recalling that $V \le \widetilde{V}$, we have

$$\begin{aligned}\int V(t)\, d\lambda_\varepsilon^N &= \sum_{i\in I_\varepsilon^N} (c_{i,\varepsilon}+d_{i,\varepsilon})\, V(\frac{c_{i,\varepsilon}+d_{i,\varepsilon}}{\varepsilon^d}) + \sum_{j\in S_\varepsilon^N} b_{j,\varepsilon}^N \varepsilon^d V(b_{j,\varepsilon}^N)\\ &\le \sum_{i\in I_\varepsilon^N} c_{i,\varepsilon} V\left(\frac{c_{i,\varepsilon}}{\varepsilon^d}\right) + V(t_N)(\sum_{i\in I_\varepsilon^N} d_{i,\varepsilon} + \sum_{j\in S_\varepsilon^N} b_{j,\varepsilon}^N \varepsilon^d)\\ &\le \int_{\Omega\times[t_N,+\infty[} \widetilde{V}(t) d\lambda_\varepsilon \;+\; \widetilde{V}(t_N)\ \lambda_\varepsilon(\Omega\times[0,t_N]).\end{aligned}$$

Now, as the function $\widetilde{V}$ is non-increasing uppper semicontinuous and bounded on $[t_N,+\infty[$, we obtain ii) by passing to the limit as $\varepsilon\to 0$:

$$\limsup_{\varepsilon\to 0} \int \widetilde{V}\, d\lambda_\varepsilon^N \;\le\; \int_{\Omega\times[t_N,+\infty[} \widetilde{V}\, d\lambda \;+\; \widetilde{V}(t_N)\, \lambda(\Omega\times[0,t_N]) \;\le\; \int_{\Omega\times[0,+\infty[} \widetilde{V}\, d\lambda\ .$$

(for the last inequality we used the fact that , thanks to the choice of t_N subject to (6.3) , we have $\lambda(\Omega\times\{t_N\})=0$.)

Proof of iii) Let $j\in R_\varepsilon^N$. Then, by construction $\int_{B_\varepsilon^N\cap Q_{j,\varepsilon}} f dx = \varepsilon^d b_{j,\varepsilon}^N \le \varepsilon^d t_N$. Then

$$\int_{A_\varepsilon^N\cap Q_{j,\varepsilon}} f dx \;=\; \int_{Q_{j,\varepsilon}} f dx - \int_{B_\varepsilon^N\cap Q_{j,\varepsilon}} f dx \;\ge\; \varepsilon^d(\alpha k^d - t_N)\ ,$$

which by (6.4) is positive for large N. On the other hand, for every $x\in Q_{j,\varepsilon}$

$$|T_\varepsilon^N(x)-x| = \operatorname{dist}(x,P_\varepsilon^N) \le \operatorname{diam}(Q_{j,\varepsilon}) + \operatorname*{essinf}_{x\in Q_{j,\varepsilon}\cap A_\varepsilon^N} |x-T_\varepsilon x|\ .$$

Thus, letting C be a positive real such that that $(a+b)^p \;\le\; \sqrt{\gamma}\, a^p + C\, b^p$ holds for all $a,b\in\mathbb{R}^+$, we have

$$\int_{Q_{j,\varepsilon}\cap B_\varepsilon^N} \frac{|x-T_\varepsilon^N x|^p}{\varepsilon^p} f(x)\, dx \le \sqrt{\gamma}\, \theta_{j,\varepsilon}^N \int_{Q_{j,\varepsilon}} \frac{|x-T_\varepsilon x|^p}{\varepsilon^p} f(x)\, dx \;+\; C\,(k\sqrt{d})^p\, b_{j,\varepsilon}^N\, \varepsilon^d\ ,$$

where

$$\theta_{j,\varepsilon}^N \;:=\; \frac{\int_{Q_{j,\varepsilon}\cap B_\varepsilon^N} f\, dx}{\int_{Q_{j,\varepsilon}\cap A_\varepsilon^N} f\, dx} \;\le\; \frac{t_N}{\alpha\, k^d - t_N} \;\to\; \frac{t^*}{\alpha\, k^d - t^*}.$$

Since $|T_\varepsilon^N x - x| \le \operatorname{diam} Q_{j,\varepsilon} \le k\sqrt{d}\varepsilon$ whenever $x\in Q_{j,\varepsilon}\cap B_\varepsilon^N$ and $j\in S_\varepsilon^N$, we obtain successively:

$$\begin{aligned}E_\varepsilon(f,\lambda_\varepsilon^N) &\le \int_\Omega \frac{|x-T_\varepsilon^N x|^p}{\varepsilon^p} f(x)\, dx\\ &= \int_{A_\varepsilon^N} \frac{|x-T_\varepsilon x|^p}{\varepsilon^p} f(x)\ dx + \sum_{j\in R_\varepsilon^N\cup S_\varepsilon^N} \int_{Q_{j,\varepsilon}\cap B_\varepsilon^N} \frac{|x-T_\varepsilon^N x|^p}{\varepsilon^p} f(x)\ dx\\ &\le E_\varepsilon(f,\lambda_\varepsilon) + \sqrt{\gamma}\frac{t_N}{\alpha\, k^d - t_N} E_\varepsilon(f,\lambda_\varepsilon) + C\,(k\sqrt{d})^p \sum_{j\in R_\varepsilon^N\cup S_\varepsilon^N} b_{j,\varepsilon}^N\, \varepsilon^d\ .\end{aligned}$$

Since $\limsup_\varepsilon \sum_{j\in R_\varepsilon^N\cup S_\varepsilon^N} b_{j,\varepsilon}^N\, \varepsilon^d \le \eta_N$ which vanishes as $N\to\infty$, by(6.4), we conclude that

$$\limsup_{N\to\infty} \left(\limsup_{\varepsilon} E_\varepsilon(f,\lambda_\varepsilon^N)\right) \;\le\; \left(1+\sqrt{\gamma}\frac{t^*}{\alpha\, k^d - t^*}\right) E(f,\lambda) \;\le\; \gamma\, E(f,\lambda)\ .$$

The proof of Lemma 6.3 is achieved. ☐

We end this subsection by a direct consequence of Lemma 6.3:

Lemma 6.4. *Let $V :]0,+\infty[\to \mathbb{R}^+$ be a continuous non-increasing potential such that $V(0^+) = +\infty$ and let $\rho \in \mathcal{P}(\mathbb{R}^+)$ be such that $G(\rho) < +\infty$ and $\int V\,d\rho < +\infty$. Then there is an approximating sequence $\{\rho^N\}$ and positive reals t_N satisfying, for every N, $\operatorname{spt}(\rho^N) \subset [t_N, +\infty[$, $\int V\,d\rho^N \le \int V\,d\rho$ and such that $\int |\rho^N - \rho| \to 0$ and $G(\rho^N) \to G(\rho)$ as $N \to \infty$.*

Proof. We apply Lemma 6.3 with $\Omega = Q$, $\lambda := 1\!\!1_Q \otimes \rho$. For every $\gamma > 1$, we may find $t_N > 0$ and a sequence $\{\lambda_\varepsilon^N\}$ such that conditions i), ii) and iii) hold. In particular

$$\limsup_N \limsup_\varepsilon E_\varepsilon(1\!\!1_Q, \lambda_\varepsilon^N) \;\le\; \gamma\, E(1\!\!1_Q, \lambda) \;=\; \gamma\, G(\rho) \;<+\infty\ . \tag{6.7}$$

Therefore, by the assertion i) of Theorem 3.1, for large N, the sequence $\{\lambda_\varepsilon^N\}_\varepsilon$ is tigh and possibly after extracting subsequences converges to a measure of the kind $\lambda^N = 1\!\!1_Q \otimes \rho_x^N$, being $\{\rho_x^N\}$ a suitable family in $\mathcal{P}(\mathbb{R}^+)$ such that $\operatorname{spt}(\rho_x^N) \subset [t_N, +\infty)$. Then we have successively

$$\limsup_N E(1\!\!1_Q, \lambda^N) \le\; \limsup_N \limsup_\varepsilon E_\varepsilon(1\!\!1_Q, \lambda_\varepsilon^N) \;=\; \gamma\, G(\rho)\ , \tag{6.8}$$

$$\int_{Q\times\mathbb{R}^+} V(t)\, d\lambda^N(x,t) \;\le\; \int_{Q\times\mathbb{R}^+} V(t)\, d\lambda \;=\; \int_{\mathbb{R}^+} V\, d\rho \tag{6.9}$$

$$\limsup_N \int |\lambda^N - \lambda| \;\le\; \limsup_N \limsup_\varepsilon \int |\lambda_\varepsilon^N - \lambda| \;=\; 0\ , \tag{6.10}$$

where (6.8) is a consequence of (6.7) together with the $\Gamma-$ convergence of $E_\varepsilon(1\!\!1_Q,\cdot)$ to $E(1\!\!1_Q,\cdot)$ whereas (6.9) follows from the assertion ii) in Lemma 6.3 noticing that, by continuity assumption, $\tilde{V}$ agrees with V. Relation (6.10) is a direct consequence of the assertion i) in Lemma 6.3.

Set $\rho^N := \int_Q \rho_x^N\, dx$. Then, $\operatorname{spt}(\rho^N) \subset [t_N, +\infty[$ and we have $\int_{\mathbb{R}^+} |\rho^N - \rho| \le \int_{Q\times\mathbb{R}^+} |\lambda^N - \lambda|$ and $\int_{Q\times\mathbb{R}^+} V(t)\, d\lambda^N(x,t) = \int_{\mathbb{R}^+} V\, d\rho^N$. Eventually, by applying Jensen inequality to the convex lower semicontinuous G, we have $G(\rho^N) \le \int_Q G(\rho_x^N)\, dx = E(1\!\!1_Q, \lambda^N)$. The conclusion follows by exploiting (6.8),(6.9)(6.10) after sending N to infinity and then γ to 1. ☐

6.2. **Proof of the convergence of infima.** We may now proceed to the proof of Theorem 3.5. This is done in several steps.

Step 1. *We show that* $\liminf_{\varepsilon\to 0} \inf \mathcal{P}_\varepsilon^c \ge \inf \mathcal{P}^c$.
The inequality being trivial if $\liminf_{\varepsilon\to 0} \inf \mathcal{P}_\varepsilon^c = +\infty$, after extracting a subsequence , we may without loss of generality assume that $\sup_\varepsilon \inf \mathcal{P}_\varepsilon^c < +\infty$. Let λ_ε be an ε-approximate minimizer for $\inf \mathcal{P}_\varepsilon^c$. By Theorem 3.1, up to a subsequence, $\{\lambda_\varepsilon\}$ converges tightly to some measure λ and we have:

$$\liminf_{\varepsilon\to 0} \inf \mathcal{P}_\varepsilon^c \;=\; \liminf_{\varepsilon\to 0} E_\varepsilon(f, \lambda_\varepsilon) \;\ge\; E(f,\lambda)\ . \tag{6.11}$$

Futher as $V \ge 0$ is lower semicontinuous and by the second assertion ii) of Lemma 6.1

$$\liminf_\varepsilon \int V(t)\, d\lambda_\varepsilon(x,t) \ge \int V(t)\, d\lambda(x,t) \quad , \quad \lim_\varepsilon \int \beta(t)\, d\lambda_\varepsilon(x,t) = \int \beta(t)\, d\lambda(x,t)\ .$$

Thus $\int g(t)\, d\lambda(x,t) \le \liminf_\varepsilon \int g(t)\, d\lambda_\varepsilon(x,t) \le c$. Taking into account (6.11), we infer that

$$\liminf_{\varepsilon\to 0} \inf \mathcal{P}_\varepsilon^c \;\ge\; E(f,\lambda) \;\ge\; \inf \mathcal{P}^c \quad , \quad \text{with } \lambda \text{ admissible for } \mathcal{P}^c\ . \tag{6.12}$$

Step 2. *We assume that* $V(0_+) < +\infty$ *and prove that* $\limsup_{\varepsilon\to 0} \inf \mathcal{P}^c_\varepsilon \le \inf \mathcal{P}^c$.
Let $r > \inf \mathcal{P}^c$ (as $c > \bar{c}$, $\inf \mathcal{P}^c < +\infty$). By Lemma 6.2, there exists λ such that

$$E(f,\lambda) < r \quad , \quad \int \widetilde{g}(t)\, d\lambda(x,t) < c \ ,$$

and by Theorem 3.1, there exists a realizing sequence $\{\lambda_\varepsilon\}$ such that

$$\lambda_\varepsilon \to \lambda \quad \text{tightly} \quad , \quad \limsup_{\varepsilon\to 0} E_\varepsilon(f,\lambda_\varepsilon) \ \le \ E(f,\lambda) \ . \tag{6.13}$$

Applying Lemma 6.1, we have in addition that

$$\begin{aligned}\limsup_\varepsilon \int g(t)\, d\lambda_\varepsilon(x,t) \ &\le \ \limsup_\varepsilon \int V(t)\, d\lambda_\varepsilon(x,t) + \limsup_\varepsilon \int \beta(t)\, d\lambda_\varepsilon(x,t) \\ &\le \ \int \widetilde{V}(t)\, d\lambda(x,t) + \int \beta(t)\, d\lambda(x,t) \ \le \ \int \widetilde{g}(t)\, d\lambda(x,t) \ < c\end{aligned}$$

Thus λ_ε becomes admissible for $(\mathcal{P}^c_\varepsilon)$ while ε decreases to zero. Therefore

$$\limsup_{\varepsilon\to 0} \inf \mathcal{P}^c_\varepsilon \ \le \ \limsup_{\varepsilon\to 0} E_\varepsilon(f,\lambda_\varepsilon) \ \le \ E(f,\lambda) < r \ .$$

The conclusion of Step 2 follows by letting r tend to $\inf \mathcal{P}^c$.

Step 3. *Extension of Step 2 in the unbounded case ($V(0+) = +\infty$.)*
We take r and λ as in Step 2 and consider a sequence $\{\lambda_\varepsilon\}$ satisfying (6.13). As g is unbounded near 0, we cannot prevent a priori that $\int g(t) d\lambda_\varepsilon$ blows up as $\varepsilon \to 0$. Therefore we need to modify the sequence $\{\lambda_\varepsilon\}$ removing the small Dirac masses. To that aim, for every $\gamma > 1$, we construct a double indexed sequence $\{\lambda^N_\varepsilon\}$ satisfying the conditions i), ii), iii) of Lemma 6.3. Then by the same diagonalization argument as in Subsection 5.3, we can find a sequence $N(\varepsilon)$ such that $N(\varepsilon) \to +\infty$ as $\varepsilon \to 0$ and such that $\lambda'_\varepsilon := \lambda^{N(\varepsilon)}_\varepsilon$ satisfies

$$\lambda'_\varepsilon \to \lambda \quad \text{tightly} \quad , \quad \limsup_{\varepsilon\to 0} E_\varepsilon(f,\lambda'_\varepsilon) \ \le \ \gamma\, E(f,\lambda) \quad , \quad \limsup_\varepsilon \int V(t)\, d\lambda'_\varepsilon(x,t) \ \le \ \int \widetilde{V}(t)\, d\lambda(x,t) \ .$$

Eventually by the second assertion of Lemma 6.1, $\lim_\varepsilon \int \beta(t)\, d\lambda'_\varepsilon(x,t) \ = \ \int \beta(t)\, d\lambda(x,t)$. Thus

$$\limsup_\varepsilon \int g(t)\, d\lambda'_\varepsilon(x,t) \ \le \int \widetilde{g}(t)\, d\lambda(x,t) \ < \ c \ ,$$

and the conclusion follows like in Step 2 after sending γ to 1.

Step 4. (*End of the proof*) Under the condition $c > \bar{c}$, it has been shown in Lemma 3.4 that the infimum of $\mathcal{P}^c$ is finite. Then from the previous steps, one has that $\lim_\varepsilon \inf \mathcal{P}^c_\varepsilon = \inf \mathcal{P}^c$. Let $\{\lambda_\varepsilon\}$ be a minimizing sequence for $\inf \mathcal{P}^c_\varepsilon$ and let λ be a tight limit point. Then, possibly after extracting a subsequence, the assertion (6.12) holds true yielding that λ solves $\mathcal{P}^c$. The proof of Theorem 3.5 is finished. □

6.3. **Proof of Proposition 3.6.** Let $c > \bar{c}$. By Lemma 3.4, there exists an optimal $\bar{\lambda} = f \otimes \bar{\rho}^x$ for $\mathcal{P}_c$. By the definition of Φ_g in (3.15), there holds $G(\bar{\rho}^x) \ge \ \Phi_g \left(\dfrac{\bar{u}(x)}{f(x)}\right)$ *a.e.* , where the function $\bar{u}(x) := f(x) < \bar{\rho}^x, g >$ satisfies the constraint $\int_\Omega \bar{u} \le c$. Therefore

$$m(c) = \min \mathcal{P}_c = E(f,\bar{\lambda}) \ = \ \int_\Omega G(\bar{\rho}^x)\, f(x)^{1-\frac{p}{d}}\, dx \ \ge \ \int_\Omega \Phi_g\left(\frac{\bar{u}}{f}\right) f^{1-\frac{p}{d}}\, dx \ \ge \ \inf \mathcal{Q}_c \ . \tag{6.14}$$

Let now $u \in L^1(\Omega)$ be such that $\int_\Omega u\,dx \le c$ and $\int_\Omega \Phi_g\left(\frac{u}{f}\right) f^{1-p/d}\,dx < +\infty$. By Lemma 3.7, the Borel multifunction $\Gamma : s \in \mathbb{R} \mapsto \mathrm{Argmin}\Phi_g(s)$ has non empty convex compact values in $\mathcal{P}(\mathbb{R}^+)$ for all s such that $\Phi_g(s) < +\infty$. Therefore the map $x \mapsto \Gamma\left(\frac{u(x)}{f(x)}\right)$ is a.e. non empty and Lebesgue measurable and by the classical measurable selection Theorem (see [12]), we may chose a measurable $x \mapsto \rho^x$ such that there holds a.e.

$$< \rho^x, g > f(x) \le u(x) \quad , \quad G(\rho^x) = \Phi_g\left(\frac{u(x)}{f(x)}\right) .$$

Then $\lambda = f \otimes \rho^x$ is admissible for $(\mathcal{P}_c)$ and we then deduce that $E(f,\lambda) = \int_\Omega \Phi_g\left(\frac{u}{f}\right) f^{1-p/d}\,dx \ge m(c)$. Together with (6.14), we conclude that $m(c) = \min \mathcal{Q}_c$ with $\bar{u}$ solving (Q_c). The optimality condition (3.16) for ρ^x can be recovered by writing that the inequalities in (6.14) need to be equalities. □

6.4. **Proof of Lemma 3.10.** We begin by showing (3.22). By taking $\rho = \delta_1$ which is admissible for every $q \in [-\infty, +\infty[$, we obtain that $C_{p,d}(q) \le G(\delta_1) = \omega_{p,d}$. Now, if $q > 1$, for every $\theta > 0$, we may choose $\rho = \theta\delta_a + (1-\theta)\delta_0$ which is admissible for $a = \theta^{\frac{1}{1-q}}$. Thanks to (3.9), we are led to the following upper bound holding for small values of θ:

$$\forall q > 1, \quad \forall \theta \in]0, \theta(B)] , \qquad C_{p,d}(q) \le G(\theta\delta_a + (1-\theta)\delta_0) = \gamma_{p,d}\, \theta^{1+\frac{p}{d(1-q)}} .$$

Thus, for $q = 1 + \frac{p}{d}$, we get $C_{p,d}(q) \le \gamma_{p,d}$ whereas the converse inequality and the optimality of ρ given above follows trivially from the lowerbound of G in (3.10). For $q > 1 + \frac{p}{d}$, we obtain $C_{p,d}(q) = 0$ by sending $\theta \to 0$. Notice that, in this case, the infimum in (3.20) is not attained.

Next we introduce, for every $q \in [-\infty, +\infty[$, the functional

$$S_q : \rho \in \mathcal{P}(\mathbb{R}^+) \to \begin{cases} \left(\int_{\mathbb{R}^+} t^{q-1}\,d\rho\right)^{\frac{1}{q-1}} & \text{if } q \ne 1 \\ \exp\left(\int_{\mathbb{R}^+} \ln t\,d\rho\right) & \text{if } q = 1 \\ \inf\{t \ : \ t \in \mathrm{spt}\rho\} & \text{if } q = -\infty \end{cases} ,$$

which allows us to recast (3.20) and (3.21) as $C_{p,d}(q) = \inf\{G(\rho) \ : \ S_q(\rho) \ge 1\}$. It can be easily checked that $S_{q'} \ge S_q$ whenever $q' > q$, thus the function $q \mapsto C_{p,d}(q)$ is finite non-increasing on $[-\infty, +\infty[$. To prove the continuity property on the interval $[-\infty, 1 + p/d[$ and the existence of an optimal ρ, we use the following claim:

Claim Let $q \in [-\infty, 1 + \frac{p}{d}[$. Then the following implications hold

$$q_n \to q \ , \ \rho_n \overset{*}{\rightharpoonup} \rho \ , \ \sup_n G(\rho_n) < +\infty \quad \Rightarrow \quad \limsup_{n\to\infty} S_{q_n}(\rho_n) \le S_q(\rho) \tag{6.15}$$

$$G(\rho) < +\infty \text{ and } \left\{q > 1 \text{ or } 0 \notin \mathrm{spt}(\rho)\right\} \quad \Rightarrow \quad \lim_{n\to\infty} S_{q_n}(\rho) = S_q(\rho) \tag{6.16}$$

The existence of a solution for (3.20) and (3.21) is obtained by considering a minimizing sequence $\{\rho_n\}$ such that $G(\rho_n) \to C_{p,d}(q)$. Then, by (3.10), $\{\rho_n\}$ is tight and we may assume that $\rho_n \overset{*}{\rightharpoonup} \rho$ where $\rho \in \mathcal{P}(\mathbb{R}^+)$. Then, by applying (6.15) with $q_n = q$, ρ satisfies $S_{q_n}(\rho) \ge 1$ and therefore is minimal by the lower semicontinuity of G.

Let $q_n \to q$ whith $q < 1 + p/d$. There exits ρ_n such that $G(\rho_n) = C_{p,d}(q_n)$ and $S_{q_n}(\rho_n) \ge 1$. As, by (3.22), $G(\rho_n) \le \omega_{p,d}$, the sequence $\{\rho_n\}$ is tight and we may assume that $\rho_n \overset{*}{\rightharpoonup} \rho$. Applying (6.15), we obtain $S_q(\rho) \geq \limsup_n S_q(\rho_n) \geq 1$. It follows that

$$\liminf_n C_{p,d}(q_n) \;=\; \liminf_n G(\rho_n) \;\geq\; G(\rho) \;\geq\; C_{p,d}(q)\ .$$

It remains to show that $\limsup_n C_{p,d}(q_n) \leq C_{p,d}(q)$. Let ρ such that $G(\rho) = C_{p,d}(q)$ and $S_q(\rho) \geq 1$. We notice that, for every $a > 1$, the dilated measure $L_a^\sharp(\rho)$ satisfies $S_q(L_a^\sharp(\rho)) = aS_q(\rho) > 1$. We now distinguish three cases:
If $q > 1$, we apply (6.16) with $\rho = L_a^\sharp(\rho)$ and find that $\lim_n S_{q_n}(L_a^\sharp(\rho)) = S_q(L_a^\sharp(\rho)) > 1$, thus $S_{q_n}(L_a^\sharp(\rho)) \geq 1$ holds eventually. Then recalling (3.7), we infer that $\limsup_n C_{p,d}(q_n) \leq G(L_a^\sharp(\rho)) = a^{p/d} C_{p,d}(q)$ thus the conclusion by sending $a \searrow 1$.
In the case $q < 1$, we consider an appproximating sequence $\rho^N \overset{*}{\rightharpoonup} L_a^\sharp(\rho)$ as given in Lemma 6.4 applied with $V(t) = t^{q-1}$, that is satisfying $G(\rho^N) \to G(L_a^\sharp(\rho))$ and, for every N, $\int t^{q-1}\, d\rho^N \leq \int t^{q-1}\, dL_a^\sharp(\rho)\ ,\ 0 \notin \operatorname{spt}(\rho^N)$. By (6.16), we have $\lim_n S_{q_n}(\rho^N) = S_q(\rho^N) \geq S_q(L_a^\sharp(\rho)) > 1$, and therefore $\limsup_n C_{p,d}(q_n) \leq G(\rho^N)$. The conclusion follows by sending $N \to \infty$ and then $a \searrow 1$.
For $q = 1$, we may assume that $q_n \nearrow 1$ (since $C_{p,d}(q) \leq C_{p,d}(1)$ for $q > 1$). Then we construct the sequence $\rho^N \overset{*}{\rightharpoonup} L_a^\sharp(\rho)$ by applying Lemma 6.3 with $V(t) = -\ln t$ (which is non negative after the addition of an affine function) and conclude exactly as before.

Proof of claims (6.15)(6.16): For $1 < q < 1 + p/d$ and $q_n \to q$, the condition $\sup_n G(\rho_n) < +\infty$ and the growth lower bound in (3.10) imply that the sequence $\{t^{q_n-1}\,\rho_n\}$ is tight and weakly-$*$ converges to $t^{q-1}\,\rho$. The same conclusion holds true for $q \leq 1$ if $\operatorname{spt}(\rho_n) \subset [t_0, +\infty[$ for a suitable $t_0 > 0$. Thus we have (6.15) for $q > 1$ and, by choosing $\rho_n = \rho$, we also deduce that (6.16) for all $q < 1 + p/d$.
It remains to show (6.15) when $q \leq 1$. We may assume $S_q(\rho) < +\infty$ and therefore $\rho(\{0\}) = 0$. For $q < 1$, we have to check that, for $q_n \to q$, there holds $\liminf_n \int d\mu_n \geq \int d\mu$ being $\mu_n = t^{q_n-1}\,\rho_n$ and $\mu = t^{q-1}\,\rho$. We may assume that μ_n has a uniformly bounded mass and do converge weakly-$*$ to some measure m. By using the uniform convergence of the continuous functions $t^{q_n-1} \to t^{q-1}$ on every compact subset of $]0, +\infty[$, we derive that m agrees with μ on $]0, +\infty[$ thus $\liminf_n \int d\mu_n \geq \int dm \geq \int d\mu$.
Let us show (6.15) for $q = 1$. By the monotonicity property of S_q, we may asume that $q_n \nearrow 1$ and we are reduced to prove the following inequality

$$\liminf_n \ \frac{1}{1-q_n}\ \ln\left(\int t^{q_n-1}\, d\rho_n\right) \ \geq\ -\int \ln t\ d\rho\ . \tag{6.17}$$

Without loss of generality, we may assume that the left hand member in (6.17) is finite. Then, we have $\int t^{q_n-1}\, d\rho_n \to 1$ and therefore:

$$\liminf_n \ \frac{1}{1-q_n}\ \ln\left(\int t^{q_n-1}\, d\rho_n\right) \ \geq\ \liminf_n \ \int \frac{t^{q_n-1}-1}{1-q_n}\, d\rho_n \ \geq\ \liminf_n \ \int -\ln t\, d\rho_n\ .$$

Thanks to the condition $\sup_n G(\rho_n) < +\infty$, we infer that the positive parts $(\ln t)^+\rho_n$ converges tightly to $(\ln t)^+\rho$. Then by the lower semicontinuity of $(\ln t)^-$ there holds:

$$\liminf_n \int -\ln t\, d\rho_n \geq \int (\ln t)^- d\rho - \int (\ln t)^+ d\rho \ =\ \int -\ln t\, d\rho\ ,$$

hence (6.17). The proof of the two claims and of Lemma 3.10 is finished. ☐

6.5. **Proof of Proposition 3.11.** i) Let $q < 1$ and $g(t) = t^{q-1}$. Then, setting $\varepsilon = n^{\frac{1}{d(q-1)}}$, the atomic measure $\nu = \sum_i c_i \, \delta_{x_i}$ satisfies the size constraint $H(\nu) = \sum_i c_i^q \leq n$ if and only if the rescaled measure $\lambda_\varepsilon(\nu) = \sum_i c_i \, \delta_{(x_i, \frac{c_i}{\varepsilon^d})}$ satisfies $\int g \, d\lambda_\varepsilon \leq 1$. Therefore:

$$n^{\frac{p}{d(1-q)}} \, \mathcal{E}_H^{p,d}(f, \Omega, n) \; = \; \inf \left\{ E_\varepsilon(f, \lambda) \; : \; \int g \, d\lambda \leq 1 \right\} \, .$$

By applying Theorem 3.5, the infimum in the left hand side converges to $\min \left\{ E(f, \lambda) \, : \, \int g \, d\lambda \leq 1 \right\}$ and, by Proposition 3.6, the searched limit is given by

$$l_q \; := \; \inf \left\{ \int_\Omega \Phi_q \left(\frac{u}{f} \right) \, f^{1-\frac{p}{d}} \, dx \; : \; u \in L^1(\Omega) \; \int_\Omega u \, dx \leq 1 \right\} \; , \tag{6.18}$$

where $\Phi_q(s) \; := \; \inf\{G(\rho) \; : \; \int t^{q-1} \, d\rho \; \leq \; s\}$. Noticing that, for every $a > 0$, $\int t^{q-1} \, dL_a^\sharp(\rho) = a^{q-1} \int t^{q-1} \, d\rho$ and by using the scale invariance property of G in (3.7), we derive that

$$\Phi_q(s) \; = \; C_{p,d}(q) \, s^{\frac{p}{d(q-1)}} \quad \text{if } s > 0 \quad (+\infty \text{ otherwise}) \, .$$

In particular admissible functions u for the convex optimization problem (6.18) are non negative and such an element u is optimal if and only if $\Phi_q' \left(\frac{u}{f} \right) \, f^{-\frac{p}{d}} \; = \; k$ for a suitable constant k such that, in addition, $\int u dx \; = \; 1$. A boring but straightforward computation leads to $l_q = C_{p,d}(q) \left(\int_\Omega f^{\frac{1+qr}{1+r}} \, dx \right)^{1+r}$, where $r = \frac{p}{d(1-q)}$.

ii) Let $q = -\infty$ and $g(t) = 0$ if $t \geq 1$ (and $g(t) = +\infty$ if $t < 1$). Then setting $\varepsilon = n^{-\frac{1}{d}}$, we obtain that, for every atomic measure ν, $H(\nu) \leq n$ if and only if the associated λ_ε satisfies $\int g \, d\lambda_\varepsilon \leq 0$ and therefore

$$n^{\frac{p}{d}} \, \mathcal{E}_H^{p,d}(f, \Omega, n) \; = \; \inf \left\{ E_\varepsilon(f, \lambda) \; : \; \int g \, d\lambda \leq 1 \right\} \, .$$

By applying Theorem 3.5, the infimum in the left hand side converges to $\min \left\{ E(f, \lambda) \, : \, \int g \, d\lambda \leq 0 \right\}$. Here, we have $\Phi_{-\infty}(s) := \inf\{G(\rho) : \int g \, d\rho \leq s\} = C_{p,d}(-\infty)$ if $s \geq 0$ (and $\Phi_{-\infty}(s) = +\infty$ if $s < 0$). Thus the minimum problem $\mathcal{Q}^c$ in Proposition 3.6 becomes trivial and the searched limit is given by $l_{-\infty} \; := \; C_{p,d}(-\infty) \int_\Omega f^{1-\frac{p}{d}} \, dx$, which is consistent with taking the limit as $q \to -\infty$ of the quantity l_q determined in step i) (indeed $r \to 1$ and $qr \to -\frac{p}{d}$).

iii) Let $q \in]1, 1{+}p/d]$ and $g(t) = -t^{q-1}$. Setting $\varepsilon = n^{\frac{1}{d(1-q)}}$, we see that the constraint $H(\nu) \leq -1/n$ can be recast as $\int g \, d\lambda_\varepsilon \leq -1$ and therefore

$$n^{\frac{p}{d(q-1)}} \, \mathcal{E}_H^{p,d}(f, \Omega, -\frac{1}{n}) \; = \; \inf \left\{ E_\varepsilon(f, \lambda) \; : \; \int g \, d\lambda \leq -1 \right\} \, .$$

Arguing as in i), the limit as $n \to +\infty$ is given by

$$l_q \; := \; \inf \left\{ \int_\Omega \Phi_q \left(\frac{u}{f} \right) \, f^{1-\frac{p}{d}} \, dx \; : \; u \in L^1(\Omega) \; \int_\Omega u \, dx \leq -1 \right\} \; , \tag{6.19}$$

where now $\Phi_q(s) = \; C_{p,d}(q) \, (s^-)^{\frac{p}{d(q-1)}}$, being s^- the negative part of s. It is easy to check that optimal u are non positive and by writing the Euler equation, we find for the minimal value l_q the same expression as in i) (with now an exponent r ranging into $]-1, 0[$).

iv) Let $g(t) = -\ln(t)$ and set $\varepsilon = \exp(-\frac{n}{d\,I(f)})$. Every admissible atomic measure $\nu = \sum c_i \delta_{x_i}$ is such that $\int \nu = \sum_i c_i = I(f)$. Therefore the associated λ_ε satisfies

$$\int g(t)\, d\lambda_\varepsilon = \sum_i -c_i \ln(\frac{c_i}{\varepsilon^d}) \leq n + dM\ \ln(\varepsilon) = 0$$

and we obtain the relation: $\exp\left(\frac{pn}{dM}\right) \mathcal{E}_H^{p,d}(f,\Omega,n) \ = \ \inf\left\{E_\varepsilon(f,\lambda)\ :\ \int g\, d\lambda \leq 0\right\}$. Applying Theorem 3.5 with $g = V + \beta$ where $V(t) = (-\ln t)^+$ and $\beta(t) = (\ln t)^+$, we infer from Proposition 3.6 that the limit of the left hand member is given by

$$l_1 \ := \ \inf\left\{\int_\Omega \Phi_1\left(\frac{u}{f}\right) f^{1-\frac{p}{d}}\, dx\ :\ u \in L^1(\Omega)\ \int_\Omega u\, dx \leq 0\right\}\ , \tag{6.20}$$

where $\Phi_1(s) := \inf\{G(\rho) : \int -\ln t\, d\rho \leq s\}$. Noticing that ρ is such that $\int -\ln t\, d\rho \leq s$ if and only if the dilated measure $L_a^\sharp(\rho)$ satisfies $\int -\ln t\, dL_a^\sharp(\rho) \leq 0$ for $a = \exp(-s)$, we find that $\Phi(s) = C_{p,d}(1)\exp(-\frac{ps}{d})$. The minimum of the convex optimization problem (6.20) is attained at u if and only if $\int u\, dx = 0$ and $\Phi_1'\left(\frac{u}{f}\right) f^{-\frac{p}{d}} = k$ for a suitable constant k. A straightforward computation leads to $l_1 = C_{p,d}(1)\, I(f)\, \exp\left(-\frac{p}{d\,I(f)}\int_\Omega (f \ln f)\, dx\right)$. □

References

[1] L. Ambrosio, Lecture Notes on Optimal Transport Problem, Madeira(Portugal), 2-9 July 2000, http://cvgmt.sns.it/papers/amb00a/

[2] H. Attouch, Variational Convergence of Functions and Operators, Applicable Mathematics Series, Pitman, 1984.

[3] M.A. Akcoglu and U. Krengel, Ergodic theorems for superadditive processes. J. Reine Angew. Math. 323 (1981), 53–67.

[4] R. Bolton, F. Morgan, Hexagonal Economic Regions Solve the Location Problem, *Amer. Math. Monthly* **109** (2), 2002, 165-172.

[5] G. Bouchitté and G. Buttazzo, New lower semicontinuity results for nonconvex functionals defined on measures, *Nonlinear Anal.* **15** (7), 1990, 679-692.

[6] G. Bouchitté and G. Buttazzo, Integral representations for nonconvex functionals defined on measures, *Ann. I. H. Poincaré-AN* **9** (1), 1992, 101-117.

[7] G. Bouchitté and G. Buttazzo, Relaxation for a Class of Nonconvex Functionals Defined on Measures . Annales de l'Institut H.Poincaré,Analyse Non Linéaire,10,no3,1993

[8] G. Bouchitté, C. Jimenez, M. Rajesh, Asymptotic of an optimal location problem C. R. Acad. Sci. Paris, Ser. I 335 (2002) 1-6.

[9] J. Bucklew and G. Wise, Multidimensional Asymptotic Quantization Theory with rth Power Distortion Measures, *IEEE Inform. Theory* **28** (2), 239-247, 1982.

[10] G. Buttazzo, F. Santambrogio and N. Varchon, Asymptotics of an optimal compliance-location problem *ESAIM-COCV* **12** (4), 2006, 752–769.

[11] L. Caravenna, An existence result for the Monge problem in $\mathbb{R}^n$ with norm cost function, CVGMT preprint, available at http://cvgmt.sns.it/cgi/get/papers/cara/.

[12] C. Castaing and M.Valadier, *Convex analysis and measurable multifunctions.* Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin-New York, 1977.

[13] C. Castaing, P. Raynaud de Fitte, M. Valadier, Young measures on topological spaces. With applications in control theory and probability theory. Mathematics and its Applications, 571, Kluwer.

[14] T. Champion -L. De Pascale, The Monge problem in $\mathbb{R}^d$ (2009), CVGMT preprint, available at http://cvgmt.sns.it/cgi/get/papers/chadep09/.

[15] G. Dal Maso: An introduction to Γ-convergence, Birkhaüser, Boston, (1993).

[16] R.M. Dudley, *Real analysis and probability*, Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA, (1989).

[17] C. Jimenez, Optimisation de Problèmes de Transport, Thèse, Université du Sud- Toulon Var, 2005.
[18] C. Licht, G. Michaille, Global-local subadditive ergodic theorems and application to homogenization in elasticity. *Ann. Math. Blaise Pascal* 9 (2002), no. 1, 21–62.
[19] D.J. Newman, The Hexagonal Theorem, *IEEE Inform. Theory* **28**(2), 1982, 137-138.
[20] C. Villani, Topics in Optimal transportation, Graduate studies in Mathematics, 58, AMS (2003).

UFR Sciences, Université du Sud-Toulon-Var, BP20132, 83957 La Garde Cedex, France
E-mail address: `bouchitte@univ-tln.fr`

Laboratoire de Mathematiques de Brest, 6, avenue Victor Le Gorgeu, CS 93837, F-29238 Brest Cedex 3, France
E-mail address: `chloe.jimenez@univ-brest.fr`

Departamento de Matematica, Facultad de Ciencias Fisicas y Matematicas, Universidad de Concepcion, Casilla 160-C. Concepcion. CHILE.
E-mail address: `rmahadevan@udec.cl`