Large Data Limit for a Phase Transition Model with the p-Laplacian on Point Clouds

Riccardo Cristoferi[1] and Matthew Thorpe[2]

[1]Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh, PA 15213, USA

[2]Department of Applied Mathematics and Theoretical Physics,
University of Cambridge,
Cambridge, CB3 0WA, UK

February 2018

Abstract

The consistency of a nonlocal anisotropic Ginzburg-Landau type functional for data classification and clustering is studied. The Ginzburg-Landau objective functional combines a double well potential, that favours indicator valued function, and the p-Laplacian, that enforces regularity. Under appropriate scaling between the two terms minimisers exhibit a phase transition on the order of $\varepsilon = \varepsilon_n$ where n is the number of data points. We study the large data asymptotics, i.e. as $n \to \infty$, in the regime where $\varepsilon_n \to 0$. The mathematical tool used to address this question is Γ-convergence. In particular, it is proved that the discrete model converges to a weighted anisotropic perimeter.

1 Introduction

The analysis of big data is one of the most important challenges we currently face. A typical problem concerns partitioning the data based on some notion of similarity. When the method makes use of a (usually small) subset of the data for which there are labels then this is known as a *classification problem*. When the method only uses geometric features, i.e. there are are no a-priori known labels, then this is known as a *clustering problem*. We refer to both problems as labelling problems.

A popular method to to represent the geometry of a given data set is construct a graph embedded in an ambient space $\mathbb{R}^d$, Typically the labelling task is fulfilled via a minimization procedure. In the machine learning community, successfully implemented approaches considered minimizing graph cuts and total variation (see, for instance, [4, 10, 13–16, 37, 45, 46, 48, 49]).

Of capital importance for evaluating a labelling method is whether it is *consistent* of not; namely it is desirable that the minimization procedure approaches some limit minimization procedure when the number of elements of the data set goes to infinity. Indeed, practitioners want to know if the labelling obtained by applying a specific minimization algorithm is an approximation of a limit (minimizing) object. For a consistent methodology properties of the large data limit will be evident when a large, but finite, number of data points is being considered. In particular, this can also be used to justify, a posteriori, the use of a certain procedure in order to obtain some desired features of the classification. Furthermore, understanding the large data limits can open up new algorithms.

This paper is part of an ongoing project aimed at justifying analytically the consistency of several models for soft labelling used by practitioners. Here we consider a generalization of the approach introduced by Bertozzi and Flenner in [7] (see also [17], for an introduction on this topic see [54]), where a Ginzburg-Landau (or Modica-Mortola, see [38, 39]) type functional is used as the underlining energy to minimize in the context of the soft classification problem. Our goal is to prove the consistency of the model.

In future works we will consider other kind of approaches to the two phases soft labelling problems, for instance the Ginzburg-Landau functional we consider is based on the p-Laplacian, one can also consider the normalised p-Laplacian or the random walk Laplacian (see [40, 45]). One can also cosider models for multi-phase labelling (see [29]) and convergence of the associated gradient flows.

The paper is organized as follows: in the following subsection we define the discrete model, and in Subsection 1.2 we define the continuum limiting problem. The main results are given in Section 1.3 with the proofs presented in Sections 4 and 5. Section 2 contains some prelimimary material we include for the convenience of the reader. Finally, Section 3 is devoted to the proofs of some technical results that are of interest in their own right, and are later used in the proofs in Section 3.

1.1 Finite Data Model

In the graph representation of a data set, vertexes are points $X_n := \{x_i\}_{i=1}^n \subset \mathbb{R}^d$ connected with weighted edges $\{W_{ij}\}_{i,j=1}^n$, where each $W_{ij} \geq 0$ is meant to represent particular similarities between the vertexes x_i and x_j, and in some sense represent the geometry. The larger W_{ij} is the more similar the points x_i and x_j are and "the closer they are on the graph".

Let us consider the problem of partitioning a set of data in two classes. A partition of the set of points X_n is a map $u : X_n \to \{-1, 1\}$, where -1 and 1 represent the two classes. This is referred to as *hard* labelling, since u can only assume a finite number of values. From the computational point of view it is preferable to work with functions whose values range in the whole interval $[-1, 1]$, *i.e.*, labellings $u : X_n \to [-1, 1]$, thus allowing for a *soft* labelling. Labels that are close to 1, or to -1, are supposed to be in the same class. The model used to obtain the binary classification should then force the labelling to be either 1 or -1 when the number of data points is large.

In order to scale the weights on the edges of the graph we define W_{ij} through a kernel $\eta : \mathbb{R}^d \to \mathbb{R}$. More precisely, we define the graph weights by $W_{ij} = \eta_\varepsilon(x_i - x_j) = \frac{1}{\varepsilon^d}\eta((x_i - x_j)/\varepsilon)$ where ε controls the scale of interactions on the graph; in particular choosing ε large implies the graph is dense, and choosing ε small implies the graph is disconnected. Later assumptions, see Remark 1.4, imply that we scale $\varepsilon = \varepsilon_n$ such that the graph is eventually connected (with probability one).

We now introduce the discrete functional we are going to study.

Definition 1.1. *For $p \geq 1$ and $n \in \mathbb{N}$ define the functional $\mathcal{G}_n^{(p)} : L^1(X_n) \to [0, \infty)$ by*

$$\mathcal{G}_n^{(p)}(u) := \frac{1}{\varepsilon_n n^2} \sum_{i,j=1}^n W_{ij} |u(x_i) - u(x_j)|^p + \frac{1}{\varepsilon_n n} \sum_{i=1}^n V(u(x_i)) \,,$$

where

$$W_{ij} := \eta_{\varepsilon_n}(x_i - x_j) := \frac{1}{\varepsilon_n^d} \eta\left(\frac{x_i - x_j}{\varepsilon_n}\right) \,. \tag{1}$$

The first term in $\mathcal{G}_n$ plays the role of penalising oscillations, intuitively one wants a labelling solution such that if x_i and x_j are close on the graph then the labels are also close. The first term, when $p = 2$, can also be written as $\frac{1}{\varepsilon_n n}\langle u, Lu\rangle_{\mu_n}$ where L is the graph Laplacian. The second term penalises soft labellings. In particular we assume that $V(t) = 0$ if and only if $t \in \{\pm 1\}$ and $V(t) > 0$ for all $t \neq \pm 1$. Hence any soft labelling is given a penalty of $\frac{1}{\varepsilon_n n}\sum_{i=1}^n V(u(x_i))$, as $\varepsilon_n \to 0$ this penality blows up unless u takes the values ± 1 almost everywhere.

The function η plays the role of a mollifier, and that explains the definition of η_{ε_n}. Moreover, to justify the scaling $\frac{1}{\varepsilon_n}$ we reason as follows: since η has support contained in a ball, we get

$$|u(x_i) - u(x_j)|^p \sim \varepsilon_n^p |\nabla u|^p \,.$$

So that, dividing by ε_n will give us the typical form of the singular perturbation used in the gradient theory of phase transitions (see [38]), namely

$$\int_X \frac{1}{\varepsilon_n} V(u) + \varepsilon_n^{p-1} |\nabla u|^p \,.$$

The consistency of the model is studied by using Γ-convergence (see Section 2.4), a very important tool introduced by De Giorgi in the 70's to understand the limiting behavior of a sequence of functionals (see [21]). This kind of variational convergence gives, almost immediately, convergence of minimizers.

1.2 Infinite Data Model

In order to define the limiting functional, we first introduce some notation.

Definition 1.2. *Let $\nu \in \mathbb{R}^d$. Define $\nu^\perp := \{z \in \mathbb{R}^d \,:\, z \cdot \nu = 0\}$. Moreover, for $x \in \mathbb{R}^d$, set*

$$\mathcal{C}(x,\nu) := \Big\{ C \subset \nu^\perp \,:\, C \textit{ is a } (d-1)\textit{-dimensional cube centred at } x \Big\} \,.$$

*For $C \in \mathcal{C}(x,\nu)$, we denote by $v_1, \dots, v_{d-1}$ its principal directions (where each v_i is a unit vector normal to the i^{th} face of C), and we say that a function $u : \mathbb{R}^d \to \mathbb{R}$ is C-*periodic *if $u(y + r v_i) = u(y)$ for all $y \in \mathbb{R}^d$, all $r \in \mathbb{N}$ and all $i = 1, \dots, d-1$.*

Finally, we consider the following space of functions:

$$\mathcal{U}(C,\nu) := \Big\{ u : \mathbb{R}^d \to [-1,1] \,:\, u \textit{ is } C\textit{-periodic}, \lim_{y\cdot\nu \to \infty} u(y) = 1, \textit{ and } \lim_{y\cdot\nu\to-\infty} u(y) = -1 \Big\} \,.$$

We now define the limiting (continuum) model.

Definition 1.3. *Let $p \geq 1$. Define the functional $\mathcal{G}_\infty^{(p)} : L^1(X) \to [0,\infty]$ by*

$$\mathcal{G}_\infty^{(p)}(u) := \begin{cases} \displaystyle\int_{\partial^*\{u=1\}} \sigma^{(p)}(x, \nu_u(x))\rho(x)\,\mathrm{d}\mathcal{H}^{d-1}(x) & \textit{if } u \in BV(X;\{\pm 1\}) \,, \\[2ex] +\infty & \textit{else} \,, \end{cases}$$

where

$$\sigma^{(p)}(x,\nu) := \inf\left\{ \frac{1}{\mathcal{H}^{d-1}(C)} G^{(p)}(u,\rho(x),T_C) \,:\, C \in \mathcal{C}(x,\nu)\,, u \in \mathcal{U}(C,\nu) \right\}\,,$$

and, for $C \in \mathcal{C}(x,\nu)$, we set $T_C := \{z + t\nu \,:\, z \in C, t \in \mathbb{R}\}$. Finally, for $\lambda \in \mathbb{R}$ and $A \subset \mathbb{R}^d$ define

$$G^{(p)}(u,\lambda,A) := \lambda \int_A \int_{\mathbb{R}^d} \eta(h)|u(z+h) - u(z)|^p \,\mathrm{d}h\,\mathrm{d}z + \int_A V(u(z))\,\mathrm{d}z\,.$$

Notice that, while the discrete functional $\mathcal{G}_n^{(p)}$ is nonlocal, the functional $\mathcal{G}_\infty^{(p)}$ is a local one. The minimization problem defining $\sigma^{(p)}$ is called the *cell problem* and it is common in phase transitions problems (see related works in Subsection 1.4). Although not explicit, we have at least information on the form of the limiting functional: an anisotropic weighted perimeter. This shows that minimizers of $\mathcal{G}_\infty^{(p)}$ are sets $E \subset X$ whose boundary ∂E (or, to be precise, reduced boundary $\partial^* E$) will likely be in the region where ρ is small and orthogonal to directions ν for which $\sigma^{(p)}(\nu)$ is as small as possible.

Finally, we want to point out that one of the main issues we have to deal with is that, for each $n \in \mathbb{N}$, the data set X_n is a discrete set, while in the limit the data is given by a probability measure μ on the set X, hence why we call $\mathcal{G}_\infty$ the continuum model. Thus, we will need to compare functions (the labeling) defined on different sets. To do so we will implement the strategy introduced by García-Trillos and Slepčev in [33], that consists in extending a function $u : X_n \to \mathbb{R}$ to a function $v : X \to \mathbb{R}$ in an optimal piecewise constant way. Optimal here is meant in the sense of optimal transportation. In particular, a sequence of maps $\{u_n\}_{n=1}^\infty$ with $u_n \in L^1(X_n)$, is said to converge in the TL^1 topology to a map $u \in L^1(X)$ if there exists a sequence $\{T_n\}_{n=1}^\infty \subset L^1(X;X_n)$ converging to the identity map in $L^1(X)$ and with

$$\mu(T_n^{-1}(B)) = \frac{1}{n}\#\{x_i \in B \,:\, i = 1,2,\ldots,n\}$$

for every Borel set $B \subset X$, such that $u_n \circ T_n \to u$ in $L^1(X)$. We review the TL^1 topology in more detail in Section 2.2.

1.3 Main Results

This section is devoted to the precise statements of the main results of this paper.

Let $X \subset \mathbb{R}^d$ be a bounded, connected and open set with Lipschitz boundary. Fix $\mu \in \mathcal{P}(X)$ and assume the following.

(A1) $\mu \ll \mathcal{L}^d$, has a continuous density $\rho : X \to [c_1, c_2]$ for some $0 < c_1 \le c_2 < \infty$.

We extend ρ to a function defined in the whole space $\mathbb{R}^d$ by setting $\rho(x) := 0$ for $x \in \mathbb{R}^d \setminus X$. For all $n \in \mathbb{N}$, consider a point cloud $X_n = \{x_i\}_{i=1}^n \subset X$ and let μ_n be the associated empirical measure (see Definition 2.1). Let $\{\varepsilon_n\}_{n=1}^\infty$ be a positive sequence converging to zero and such that the following rate of convergence holds:

(A2) $\dfrac{\mathrm{dist}_\infty(\mu_n,\mu)}{\varepsilon_n} \to 0$, where $\mathrm{dist}_\infty(\mu_n,\mu)$ is the ∞-Wasserstein distance between the measures μ_n and μ, see Definition 2.3.

Remark 1.4. When $x_i \overset{\text{iid}}{\sim} \mu$ then (with probability one), hypothesis (A2) is implied by $\varepsilon_n \gg \delta_n$, where δ_n is defined in Theorem 2.10. Notice that for $d \geq 3$ this lower bound on ε_n that ensures the graph with vertices x_n and edges weighted by W_{ij} (see (1)) is eventually connected (see [42, Theorem 13.2]). The lower bound can potentially be improved when x_i are not independent. For example if $\{x_i\}_{i=1}^n$ form a regular graph then μ_n converges to the uniform measure and the lower bound is given by $\varepsilon_n \gg n^{-\frac{1}{d}}$.

The double well potential $V : \mathbb{R}^d \to \mathbb{R}$ satisfies the following.

(B1) V is continuous.

(B2) $V^{-1}\{0\} = \{\pm 1\}$ and $V \geq 0$.

(B3) There exists $\tau > 0, R_V > 1$ such that for all $|s| \geq R_V$ that $V(s) \geq \tau|s|$.

The assumptions on V imply that in the limit there are only two phases ± 1. Assumption (B3) is used to establish compactness, in particular it is used to show that minimisers can be bounded in L^∞ by 1.

Recall that the graph weights are defined by $W_{ij} = \eta_{\varepsilon_n}(x_i - x_j)$. We assume that $\eta : \mathbb{R}^d \to [0, \infty)$ is a measurable functions satisfying the following.

(C1) $\eta \geq 0, \eta(0) > 0$ and η is continuous at $x = 0$.

(C2) η is an even function, i.e. $\eta(-x) = \eta(x)$.

(C3) η has support in $B(0, R_\eta)$, for some $R_\eta > 0$.

(C4) For all $\delta > 0$ there exists c_δ, α_δ such that if $|x - z| \leq \delta$ then $\eta(x) \geq c_\delta \eta(\alpha_\delta z)$, furthermore $c_\delta \to 1, \alpha_\delta \to 1$ as $\delta \to 0$.

Remark 1.5. Note that (C3) and (C4) imply that $\|\eta\|_{L^\infty} < \infty$ and, in particular, $\int_{\mathbb{R}^d} \eta(x)|x| \, \mathrm{d}x < \infty$. Indeed, given $\delta > 0$, it is possible to cover $B(0, R_\eta)$ with a finite family $\widetilde{B}_\delta(x_1), \ldots, \widetilde{B}_\delta(x_r)$ of sets of the form

$$\widetilde{B}_\delta(x_i) := \{ \alpha_\delta z \,:\, |z - x_i| < \delta \} .$$

Hypothesis (C2) is justified by the fact that η plays the role of an interaction potential. Finally, hypothesis (C4) is a version of continuity of η we need in order to perform our technical computations. We note that (C4) is general enough to include $\eta(x) = \chi_A$ where $A \subset \mathbb{R}^d$ is open, bounded, convex and $0 \in A$, see [52, Proposition 2.2].

The main result of the paper is the following theorem.

Theorem 1.6. *Let $p \geq 1$ and assume (A1-2), (B1-3) and (C1-4) are in force. Then, the following holds:*

- *(compactness) let $u_n \in L^1(\mu_n)$ satisfy $\sup_{n\in\mathbb{N}} \mathcal{G}_n^{(p)}(u_n) < \infty$, then u_n is relatively compact in TL^1 and each cluster point u has $\mathcal{G}_\infty^{(p)}(u) < \infty$;*
- *(Γ-convergence) $\Gamma\text{-}\lim_{n\to\infty}(TL^1)\mathcal{G}_n^{(p)} = \mathcal{G}_\infty^{(p)}$.*

Since the proof of Theorem 1.6 is quite long, we briefly sketch here the main idea behind the Γ-convergence result. We approximately follow the method of [33] where the authors considered the continuum limit of total variation on point clouds. We will show the convergence of the discrete nonlocal functional $\mathcal{G}_n^{(p)}$ to the continuum local one $\mathcal{G}_\infty^{(p)}$ via an intermediate nonlocal continuum functional $\mathcal{F}_{\varepsilon_n}^{(p)}$ (defined in (3)). In particular, we will prove that:

(i) the functionals $\mathcal{F}_{\varepsilon_n}^{(p)}$ Γ-converge in $L^1(X)$ to $\mathcal{G}_\infty^{(p)}$, see Section 3, where we implement a strategy similar to the one of [1], where the authors considered the functional $\mathcal{F}_{\varepsilon_n}^{(p)}$ with $\rho \equiv 1$ and $p = 2$,

(ii) it is possible to bound from below $\mathcal{G}_n^{(p)}$ with $\mathcal{F}_{\varepsilon_n'}^{(p)}$ (see (46)), where $\lim_{n\to\infty} \frac{\varepsilon_n'}{\varepsilon_n} = 1$, from which the liminf inequality follows,

(iii) if $u \in BV(X;\pm 1)$ and we set $u_n := u \llcorner X_n$, we get that $u_n \to u$ in $TL^1(X)$ and, up to an error term (negligible in the limit), we can get an upper bound of $\mathcal{G}_n^{(p)}(u_n)$ with $\mathcal{F}_{\varepsilon_n'}^{(p)}(u)$, where $\lim_{n\to\infty} \frac{\varepsilon_n'}{\varepsilon_n} = 1$. This will give us the limsup inequality.

Similarly, the compactness property follows by comparing $\mathcal{G}_n^{(p)}$ with the intermediary functional $\mathcal{F}_{\varepsilon_n}^{(p)}$.

As an application of the Theorem 1.6, we consider the functional $\mathcal{G}_n^{(p)}$ with a data fidelity term.

Definition 1.7. *Let $k_n : X_n \times \mathbb{R} \to \mathbb{R}$ and $k_\infty : X \times \mathbb{R} \to \mathbb{R}$. Define the functionals $\mathcal{K}_n : L^1(X_n) \to \mathbb{R}$ and $\mathcal{K}_\infty : L^1(X) \to \mathbb{R}$ by*

$$\mathcal{K}_n(u) := \frac{1}{n} \sum_{i=1}^{n} k_n(x_i, u(x_i)) \,,$$

and

$$\mathcal{K}_\infty(u) := \int_X k_\infty(x, u(x)) \rho(x) \, \mathrm{d}x \,,$$

respectively.

We make the following assumptions on k_n, k_∞:

(D1) $k_n \geq 0$, $k_\infty \geq 0$.

(D2) There exist $\beta > 0$ and $q \geq 1$ such that $k_n(x,u) \leq \beta(1 + |u|^q)$, for all $n \in \mathbb{N}$ and almost all $x \in X_n$.

(D3) For almost every $x \in X$ the following holds: let $u_n \to u$ be a converging real valued sequence and $x_n \to x$, then
$$\lim_{n\to\infty} k_n(x_n, u_n) = k_\infty(x, u) \,.$$

Remark 1.8. For example, we can use this form of $\mathcal{K}_n, \mathcal{K}_\infty$ to include a data fidelity term in a specific subset of X. Let $B \subset X$ be an open set with $\mathrm{Vol}(B) > 0$ and $\mathrm{Vol}(\partial B) = 0$. Let $\lambda_n \geq 0$ with $\lambda_n \to \lambda$ as $n \to \infty$. Let $y_n \in L^1(X_n)$ and $y_\infty \in L^1(X)$ with $\sup_{n\in\mathbb{N}} \|y_n\|_{L^\infty} < \infty$ and such that $y_n(x_{i_n}) \to y_\infty(x)$ for almost every $x \in X$ and any sequence $x_{i_n} \to x$. Define

$$k_n(x,u) := \begin{cases} \lambda_n |y_n(x) - u|^q & \text{in } B \cap X_n \,, \\ 0 & \text{on } X_n \setminus B \,, \end{cases}$$

$$k_\infty(x,u) := \begin{cases} \lambda_n |y_\infty(x) - u|^q & \text{in } B\,, \\ 0 & \text{on } X \setminus B\,. \end{cases}$$

Then k_n and k_∞ satisfy hypothesis (D1-3). Indeed, (D1) follows directly from the definition of the fidelity terms, while (D3) holds thanks to the continuity and the fact that $\mathrm{Vol}(\partial B) = 0$. Finally, in order to prove (D2) we simply notice that

$$k_n(x,u) = \lambda_n |y_n(x) - u|^q \leq \sup_{n\in\mathbb{N}} \lambda_n 2^{q-1} \left(\|y_n\|^q_{L^\infty} + |u|^q \right) \leq \beta(1 + |u|^q)\,.$$

for some $\beta > 0$.

We now consider the minimisation problem

$$\text{minimise} \quad \mathcal{G}_n^{(p)}(u) + \mathcal{K}_n(u) \quad \text{over } u \in L^1(X_n)\,.$$

Corollary 1.9. *In addition to Assumptions (A1-3), (B1-2), (C1-4), (D1-3), assume that for the same $q \geq 1$ as in Assumption (D2) there exists $\tau, R_V > 0$ such that for all $|s| \geq R_V$ that $V(s) \geq \tau |s|^q$. Then any sequence of almost minimizers of $\mathcal{G}_n^{(p)} + \mathcal{K}_n$ is compact in TL^1. And furthermore, any cluster point of almost minimizers is a minimizer of $\mathcal{G}_\infty^{(p)} + \mathcal{K}_\infty$ in $L^1(X)$.*

We prove the corollary in Section 5.

Finally, we would like to comment on the hypothesis $\rho \geq c_1 > 0$. If we drop it, we can still get the following result:

Corollary 1.10. *Let $p \geq 1$ and assume (A2), (B1-3) and (C1-4) are in force and that $\rho \in [0, c_2]$, for some $c_2 < \infty$. Set $X_+ := \{x \in X \,:\, \rho(x) > 0\}$ and define the functional $\widetilde{\mathcal{G}}_\infty^{(p)} : L^1(X) \to [0, +\infty]$ as*

$$\widetilde{\mathcal{G}}_\infty^{(p)}(u) := \begin{cases} \displaystyle\int_{\partial^*\{u=1\}} \sigma^{(p)}(x, \nu_u(x))\rho(x)\,\mathrm{d}\mathcal{H}^{d-1}(x) & \text{if } u \in BV_{loc}(X_+; \{\pm 1\})\,, \\ \\ +\infty & \text{else}\,, \end{cases}$$

where $BV_{loc}(X_+;\{\pm 1\})$ denotes the space of functions $u \in L^1(X;\pm 1)$ such that $u \in BV(K;\pm 1)$ for any compact set $K \subset X_+$. Then, the following holds:

- *(compactness) for any compact set $K \subset X_0$ we have that any sequence $\{u_n\}_{n=1}^\infty \subset L^1(K;\mu_n)$ satisfying $\sup_{n\in\mathbb{N}} \mathcal{G}_n^{(p)}(u_n) < \infty$ is relatively compact in TL^1 and each cluster point u has $\widetilde{\mathcal{G}}_\infty^{(p)}(u) < \infty$;*
- *(Γ-convergence) $\Gamma\text{-}\lim_{n\to\infty}(TL^1)\mathcal{G}_n^{(p)} = \widetilde{\mathcal{G}}_\infty^{(p)}$.*

1.4 Related Works

The functional $\mathcal{G}_n^{(p)}$ in the case $p = 1$ has been considered by the second author and Theil in [52], where a similar Γ-convergence result has been proved. The difference is that, in the case $p = 1$, the limit energy density function $\sigma^{(1)}$ can be given explicitly, via an integral. In [53] van Gennip and Bertozzi studied the Ginzburg-Landau functional on 4-regular graphs for $p = 2$ proving limits for $\varepsilon \to 0$ and $n \to \infty$ (both simultaneously and independently).

The TL^p topology, as introduced by García Trillos and Slepčev [33], provides a notion of convergence upon which the Γ-convergence framework can be applied. This method has now been applied in many works, see, for instance, [20, 23, 30, 31, 33–35, 46, 52]. Further studies on this topology can be found in [33, 34, 50, 51].

The literature on phase transitions problems is quite extensive. Here we just recall some of the main results, starting from the pioneering work [39] of Modica and Mortola and of Mortola [38], (see also Sternberg [47]) where the scalar isotropic case has been studied. The vectorial case has been considered by Kohn and Sternberg in [36], Fonseca and Tartar in [27] and Baldo [5]. A study of the anisotropic case has been carried out by Bouchitté [8] and Owen [41] in the scalar case, and by Barroso and Fonseca [6] and Fonseca and Popovici [26] in the vectorial case.

Nonlocal approximations of local functionals of the perimeter type go back to the work [1] of Alberti and Bellettini (see also [2]). Several variants and extensions have been considered since then (see, for instance, Savin and Valdinoci [44] and Esedoḡlu and Otto [24]). In particular, nonlocal functionals have been used by Brezis, Bourgain and Mironescu in [9] to characterized Sobolev spaces (see also the work [43] of Ponce)

Approximations of (anisotropic) perimeter functionals via energies defined in the discrete setting have been carried out by Braides and Yip in [12] and by Chambolle, Giacomini and Lussardi in [18].

2 Background

2.1 Notation

In the following χ_E will denote the characteristic function of a set $E \subset \mathbb{R}^d$, while $\mathrm{Vol}(E) = \mathcal{L}^d(E)$ will denote its d-dimensional Lebesgue measure and $\mathcal{H}^{d-1}(E)$ its $(d-1)$-Hausdorff measure. Moreover, with $B(x, r)$ we will denote the ball centered at $x \in \mathbb{R}^d$ with radius $r > 0$ and we set $\mathbb{S}^{d-1} := \partial B(0, 1)$. The identity map will be denoted by Id.

Given an open set $X \subset \mathbb{R}^d$, we define the space

$$\mathcal{P}(X) := \{ \text{ Radon measures } \mu \text{ on } X \text{with } \mu(X) = 1 \} .$$

Given a set of data points $\{x_i\}_{i=1}^n$ we define the empirical measure as follows.

Definition 2.1. *For all $n \in \mathbb{N}$ let $X_n := \{x_i\}_{i=1}^n$ be a set of n random variables. We define the* empirical measure μ_n *as*

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{x_i} ,$$

where δ_x denotes the Dirac delta centered at x.

We state our results in terms of a general sequence of empirical measures μ_n that converge weak* to some $\mu \in \mathcal{P}(X)$. An important special case is when x_i are independent and identically distributed (which we abbreviate to iid) from μ.

Remark 2.2. When $x_i \overset{\text{iid}}{\sim} \mu$ then μ_n converges, with probability one, to μ weakly* in the sense of measures, see for example [22, Theorem 11.4.1], (and we write $\mu_n \overset{w^*}{\rightharpoonup} \mu$), *i.e.*

$$\int_X \varphi \, \mathrm{d}\mu_n \to \int_X \varphi \, \mathrm{d}\mu$$

as $n \to \infty$, for all $\varphi \in C_c(X)$.

We write $L^p(X,\mu;Y)$ for he space of L^p integrable, with respect to μ, functions from X to Y. We will often suppress the Y dependence and just write $L^p(X,\mu)$. Moreover, if $\mu = \mathcal{L}^d$ then we will often write $L^p(X) = L^p(X,\mu)$. If $\mu = \mu_n$ is the empirical measure we also write $L^p(X_n) = L^p(X,\mu_n)$.

2.2 Transportation theory

In this section we collect the fundamental material needed in order to explain how to compare functions defined in different spaces, namely a function $w \in L^1(X,\mu)$ and a function $u \in L^1(X_n,\mu_n)$, where $X \subset \mathbb{R}^d$ is an open set and $X_n \subset X$ is a finite set of points. This is fundamental in stating our Γ-convergence result (Theorem 1.6). The TL^p space was introduced in [33] and consists of comparing w and a piecewise constant extension of the function u in L^p. In particular, we take a map $T : X \to X_n$ and we consider the function $v : X \to \mathbb{R}$ defined as $v := u \circ T$. In order that this defines a metric one needs to impose conditions on T, the natural conditions are that T "matches the measure μ with μ_n" and is *optimal* in the sense that matching moves as little mass as possible (see Theorem 2.10). This will be done by using the optimal transport distance that we recall now (see also [55] for background on optimal transport and [33, 51] for a further description of the TL^p space).

Definition 2.3. *Let $X \subset \mathbb{R}^d$ be an open set and let $\mu, \lambda \in \mathcal{P}(X)$. We define the set of couplings $\Gamma(\mu,\lambda)$ between μ and λ as*

$$\Gamma(\mu,\lambda) := \{ \pi \in \mathcal{P}(X \times X) \, : \, \pi(A \times X) = \mu(A), \, \pi(X \times A) = \lambda(A), \, \textit{for all } A \subset X \} \, .$$

For $p \in [1,+\infty]$, we define the p-Wasserstein distance between μ and λ as follows:

- *when $1 \leq p < \infty$,*

$$\mathrm{dist}_p(\mu,\lambda) := \inf \left\{ \left(\int_{X \times X} |x-y|^p \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}} \, : \, \pi \in \Gamma(\mu,\lambda) \right\} ,$$

- *when $p = \infty$,*

$$\mathrm{dist}_\infty(\mu,\lambda) := \inf \{ \mathrm{esssup}_\pi \{ |x-y| \, : \, (x,y) \in X \times X \} \, : \, \pi \in \Gamma(\mu,\lambda) \} ,$$

  where esssup_π *denotes the essential supremum with respect to the measure π.*

Remark 2.4. The infimum problems in the above definition are known as the Kantorovich optimal transport problem and the distance is commonly called the p^{th} Wasserstein distance or sometimes the earth movers distance. It is possible to see (see [55]) that the infimum is actually achieved. Moreover, the metric dist_p is equivalent to the weak* convergence of probability measures $\mathcal{P}(X)$ (plus convergence of p^{th} moments).

We now consider the case we are interested in: take $\mu \in \mathcal{P}(X)$ with $\mu = \rho\mathcal{L}^d$ (where $\mathcal{L}^d$ is the d-dimensional Lebesgue measure on $\mathbb{R}^d$) and assume the density ρ is such that $0 < c_1 \leq \rho \leq c_2 < \infty$. Then it is possible to see that the Kantorovich minimization problem is equivalent to the Monge optimal transport problem (see [28]). In particular, for $p \in [1,+\infty)$ it holds that

$$\mathrm{dist}_p(\mu,\lambda) = \min \left\{ \|\mathrm{Id} - T\|_{L^p(X,\mu)} \, : \, T : X \to X \text{ Borel}, \, T_\#\mu = \lambda \right\} ,$$

where

$$\|\mathrm{Id} - T\|^p_{L^p(X,\mu)} := \int_X |x - T(x)|^p \rho(x) \,\mathrm{d}x$$

and we define the *push forward* measure $T_{\#}\mu \in \mathcal{P}(X)$ as $T_{\#}\mu(A) := \mu\left(T^{-1}(A)\right)$ for all $A \subset X$. In the case $p = +\infty$ we get

$$\mathrm{dist}_\infty(\mu, \lambda) = \inf\left\{ \|\mathrm{Id} - T\|_{L^\infty(X,\mu)} \,:\, T : X \to X \text{ Borel},\ T_{\#}\mu = \lambda \right\}.$$

A map T is called a *transport map* between μ and λ if $T_{\#}\mu = \lambda$.

Throughout the paper we will assume the empirical measures μ_n converges weakly* to μ (see Remark 2.2 for iid samples) so by Remark 2.4 there exists a sequence of Borel maps $\{T_n\}_{n=1}^\infty$ with $T_n : X \to X_n$ and $(T_n)_{\#}\mu = \mu_n$ such that

$$\lim_{n\to\infty} \|\mathrm{Id} - T_n\|^p_{L^p(X,\mu)} = 0\,.$$

Such a sequence of functions $\{T_n\}_{n=1}^\infty$ will be called *stagnating*. We are now in position to define the notion of convergence for sequences $u_n \in L^p(X_n)$ to a continuum limit $u \in L^p(X, \mu)$.

Definition 2.5. *Let $u_n \in L^p(X_n)$, $w \in L^p(X, \mu)$ where $X_n = \{x_i\}_{i=1}^n$ and assume that the empirical measure μ_n converges weak* to μ. We say that $u_n \to w$ in $TL^p(X)$, and we write $u_n \xrightarrow{TL^p} w$, if there exists a sequence of stagnating transport maps $\{T_n\}_{n=1}^\infty$ between μ and μ_n such that*

$$\|v_n - w\|_{L^p(X,\mu)} \to 0\,, \tag{2}$$

as $n \to \infty$, where $v_n := u_n \circ T_n$.

Remark 2.6. It is easy to see that if (2) holds for one sequence of stagnating maps, then it holds for all sequences of stagnating maps [33, Proposition 3.12]. Moreover, since ρ is bounded above and below it holds

$$\|v_n - \mathrm{Id}\|_{L^p(X,\mu)} \to 0 \quad \Leftrightarrow \quad \|v_n - \mathrm{Id}\|_{L^p(X)} \to 0\,.$$

We have introduced TL^p convergence $u_n \xrightarrow{TL^p} u$ by defining transport maps $T_n : X \to X_n$ which "optimally partition" the space X after which we define a piecewise constant extension of u_n to the whole of X. This constructionist approach is how we use TL^p convergence in our proofs. However, this description hides the metric properties of TL^p. We briefly mention here the metric structure which characterises the convergence given in Definition 2.5. We define the $TL^p(X)$ space as the space of couplings (u, μ) where $\mu \in \mathcal{P}(X)$ has finite p^{th} moment and $u \in L^p(\mu)$. We define the distance $d_{TL^p} : TL^p(X) \times TL^p(X) \to [0, +\infty)$ for $p \in [1, +\infty)$ by

$$\begin{aligned} d_{TL^p}((u,\mu),(v,\lambda)) &:= \min_{\pi\in\Gamma(\mu,\nu)} \left(\int_{X^2} |x-y|^p + |u(x) - v(y)|^p \,\mathrm{d}\pi(x,y) \right)^{\frac{1}{p}} \\ &= \inf_{T_{\#}\mu=\lambda} \left(\int_X |x - T(x)|^p + |u(x) - v(T(x))|^p \,\mathrm{d}\mu(x) \right)^{\frac{1}{p}}, \end{aligned}$$

or for $p = +\infty$ by

$$d_{TL^\infty}((u,\mu),(v,\lambda)) := \inf_{\pi\in\Gamma(\mu,\nu)} \left(\operatorname*{ess\,inf}_{\pi} \left\{ |x-y| + |u(x) - v(y)| \,:\, (x,y) \in X \times X \right\} \right)$$

$$= \inf_{T_{\#}\mu=\lambda}\left(\operatorname*{ess\,inf}_{\mu}\{|x-T(x)|+|u(x)-v(T(x))|\,:\,x\in X\}\right).$$

Proposition 2.7. *The distance d_{TL^p} is a metric and furthermore, $d_{TL^p}((u_n,\mu_n),(u,\mu))\to 0$ if and only if $\mu_n \overset{w^*}{\rightharpoonup} \mu$ and there exists a sequence of stagnating transport maps $\{T_n\}_{n=1}^\infty$ between μ and μ_n such that $\|u_n\circ T_n - u\|_{L^p(X,\mu)}\to 0$.*

The proof is given in [33, Remark 3.4 and Proposition 3.12]. Note that Definition 2.5 characterises TL^p convergence.

In order to be able to write the discrete functional we will need the following result.

Lemma 2.8. *Let $\lambda\in\mathcal{P}(X)$ and let $T:X\to X$ be a Borel map. Then, for any $u\in L^1(X,\lambda)$ it holds*

$$\int_X u\,\mathrm{d}T_{\#}\lambda = \int_X u\circ T\,\mathrm{d}\lambda\,.$$

Proof. Let $s:X\to\mathbb{R}$ be a simple function. Write

$$s=\sum_{i=1}^k a_i\chi_{U_i}\,.$$

Then

$$\int_X s\,\mathrm{d}T_{\#}\lambda=\sum_{i=1}^k a_i T_{\#}\lambda(U_i)=\sum_{i=1}^k a_i\lambda\left(T^{-1}(U_i)\right)=\int_X s\circ T\,\mathrm{d}\lambda\,.$$

The result then follows directly from the definition of integral. □

Remark 2.9. Applying the above result to the empirical measures $\mu_n=\frac{1}{n}\sum_{i=1}^n\delta_{x_i}$ and $u\in L^1(X_n)$ we get

$$\frac{1}{n}\sum_{i=1}^n u(x_i)=\int_X v_n(x)\,\mathrm{d}\mu(x)\,,$$

where $v_n:=u\circ T_n$ for any T_n such that $(T_n)_{\#}\mu=\mu_n$.

In [32] the authors, García Trillos and Slepčev, obtain the following rate of convergence for a sequence of stagnating maps. This is of crucial importance for applying the results of this paper to the iid setting.

Theorem 2.10. *Let $X\subset\mathbb{R}^d$ be a bounded, connected and open set with Lipschitz boundary. Let $\mu\in\mathcal{P}(X)$ of the form $\mu=\rho\mathcal{L}^d$ with $0<c_1\le\rho\le c_2<\infty$. Let $\{x_i\}_{i=1}^\infty$ be a sequence of independent and identically distributed random variables distributed on X according to the measure μ, and let μ_n be the associated empirical measure. Then, there exists a constant $C>0$ such that, with probability one, there exists a sequence $\{T_n\}_{n=1}^\infty$ of maps $T_n:X\to X$ with $(T_n)_{\#}\mu=\mu_n$ and*

$$\limsup_{n\to\infty}\frac{\|T_n-\mathrm{Id}\|_{L^\infty(X)}}{\delta_n}\le C\,,$$

where

$$\delta_n:=\begin{cases}\sqrt{\frac{\log\log n}{n}} & \text{if } d=1\,,\\ \frac{(\log n)^{\frac{3}{4}}}{\sqrt{n}} & \text{if } d=2\,,\\ \left(\frac{\log n}{n}\right)^{\frac{1}{d}} & \text{if } d\ge 3\,.\end{cases}$$

Remark 2.11. The proof for $d = 1$ is simpler and follows from the law of iterated logarithms. Notice that the connectedness of X is essential in order to get the above result.

By the above theorem our main result, Theorem 1.6, holds with probability one when $x_i \overset{\text{iid}}{\sim} \mu$ and the graph weights are scaled by ε_n with $\varepsilon_n \gg \delta_n$.

2.3 Sets of finite perimeter

In this section we recall the definition and the basic facts about sets of finite perimeter. We refer the reader to [3] for more details.

Definition 2.12. *Let $E \subset \mathbb{R}^d$ with $\mathrm{Vol}(E) < \infty$ and let $X \subset \mathbb{R}^d$ be an open set. We say that E has* finite perimeter *in X if*

$$|D\chi_E|(X) := \sup \left\{ \int_E \mathrm{div}\varphi \, \mathrm{d}x \, : \, \varphi \in C^1_c(X;\mathbb{R}^d) \, , \, \|\varphi\|_{L^\infty} \leq 1 \right\} < \infty \, .$$

Remark 2.13. If $E \subset \mathbb{R}^d$ is a set of finite perimeter in X it is possible to define a finite vector valued Radon measure $D\chi_E$ on A such that

$$\int_{\mathbb{R}^d} \varphi \, \mathrm{d}D\chi_E = \int_E \mathrm{div}\varphi \, \mathrm{d}x$$

for all $\varphi \in C^1_c(X;\mathbb{R}^d)$.

Definition 2.14. *Let $X \subset \mathbb{R}^d$ be an open set and let $u \in L^1(X;\pm 1)$ with $\|u\|_{L^1(X)} < \infty$. We say that u is of* bounded variation *in X, and we write $u \in BV(X;\pm 1)$, if $\{u = 1\} := \{x \in X \, : \, u(x) = 1\}$ has finite perimeter in X.*

Definition 2.15. *Let $E \subset \mathbb{R}^d$ be a set of finite perimeter in the open set $X \subset \mathbb{R}^d$. We define $\partial^* E$, the* reduced boundary *of E, as the set of points $x \in \mathbb{R}^d$ for which the limit*

$$\nu_E(x) := -\lim_{r \to 0} \frac{D\chi_E(x + rQ)}{|D\chi_E|(x + rQ)}$$

exists and is such that $|\nu_E(x)| = 1$. Here Q denotes the unit cube of $\mathbb{R}^d$ centered at the origin with sides parallel to the coordinate axes. The vector $\nu_E(x)$ is called the measure theoretic exterior normal *to E at x.*

We now recall the structure theorem for sets of finite perimeter due to De Giorgi, see [3, Theorem 3.59] for a proof of the following theorem.

Theorem 2.16. *Let $E \subset \mathbb{R}^d$ be a set with finite perimeter in the open set $X \subset \mathbb{R}^d$. Then*

(i) for all $x \in \partial^* E$ the set $E_r := \frac{E - x}{r}$ converges locally in $L^1(\mathbb{R}^d)$ as $r \to 0$ to the halfspace orthogonal to $\nu_E(x)$ and not containing $\nu_E(x)$,

(ii) $D\chi_E = \nu_E \, \mathcal{H}^{d-1} \llcorner \partial^* E$,

(iii) the reduced boundary $\partial^* E$ *is* $\mathcal{H}^{d-1}$*-rectifiable,* i.e., *there exist Lipschitz functions* $f_i : \mathbb{R}^{d-1} \to \mathbb{R}^d$ *such that*

$$\partial^* E = \bigcup_{i=1}^{\infty} f_i(K_i)\,,$$

where each $K_i \subset \mathbb{R}^{d-1}$ *is a compact set.*

Remark 2.17. Using the above result it is possible to prove that (see [25])

$$\nu_E(x) = -\lim_{r\to 0} \frac{D\chi_E(x + rQ)}{r^{d-1}}$$

for all $x \in \partial^* E$, where Q is a unit cube centred at 0 with sides parallel to the co-ordinate axis.

The construction of the recovery sequences in Section 3.4 and Section 4.3 will be done for a special class of functions, that we introduce now.

Definition 2.18. *We say that a function* $u \in L^1(X; \pm 1)$ *is* polyhedral *if* $u = \chi_E - \chi_{X\setminus E}$*, where* $E \subset X$ *is a set whose boundary is a Lipschitz manifold contained in the union of finitely many affine hyperplanes. In particular,* $u \in BV(X, \pm 1)$*.*

Using the result [3, Theorem 3.42] and the fact that it is possible to approximate every smooth surface with polyhedral sets, it is possible to obtain the following density result.

Theorem 2.19. *Let* $u \in BV(X; \{\pm 1\})$*. Then there exists a sequence* $\{u_n\}_{n=1}^{\infty} \subset BV(X; \{\pm 1\})$ *of polyhedral functions such that* $u_n \to u$ *in* $L^1(X)$ *and* $|Du_n|(X) \to |Du|(X)$*. In particular* $Du_n \overset{w^*}{\rightharpoonup} Du$*.*

Finally, we recall a result due to Reshetnvyak in the form we will need in this paper (for a proof of the general case see, for instance, [3, Theorem 2.38]).

Theorem 2.20. *Let* $\{E_n\}_{n=1}^{\infty}$ *be a sequence of sets of finite perimeter in the open set* $X \subset \mathbb{R}^d$ *such that* $D\chi_{E_n} \overset{w^*}{\rightharpoonup} D\chi_E$ *and* $|D\chi_{E_n}|(X) \to |D\chi_E|(X)$*, where* E *is a set of finite perimeter in* X*. Let* $f : X \times \mathbb{S}^{d-1} \to [0, \infty)$ *be an upper semi-continuous function. Then*

$$\limsup_{n\to\infty} \int_{\partial^* E_n \cap X} f\left(x, \nu_{E_n}(x)\right) \, \mathrm{d}\mathcal{H}^{d-1}(x) \le \int_{\partial^* E \cap X} f\left(x, \nu_E(x)\right) \, \mathrm{d}\mathcal{H}^{d-1}(x)\,.$$

2.4 Γ-convergence

We recall the basic notions and properties of Γ-convergence (in metric spaces) we will use in the paper (for a reference, see [11, 19]).

Definition 2.21. *Let* (A, d) *be a metric space. We say that* $F_n : A \to [-\infty, +\infty]$ Γ*-converges to* $F : A \to [-\infty, +\infty]$*, and we write* $F_n \overset{\Gamma\text{-(d)}}{\longrightarrow} F$ *or* $F = \Gamma\text{-}\lim(d)_{n\to\infty} F_n$*, if the following hold true:*

(i) for every $x \in A$ *and every* $x_n \to x$ *we have*

$$F(x) \le \liminf_{n\to\infty} F_n(x_n)\,;$$

(ii) for every $x \in A$ there exists $\{x_n\}_{n=1}^\infty \subset A$ (the so called recovery sequence*) with $x \to x$ such that*

$$\limsup_{n\to\infty} F_n(x_n) \leq F(x)\,.$$

The notion of Γ-convergence has been designed in order for the following convergence of minimisers and minima result to hold.

Theorem 2.22. *Let (A, d) be a metric space and let $F_n \overset{\Gamma-(\mathrm{d})}{\longrightarrow} F$, where F_n and F are as in the above definition. Let $\{\varepsilon_n\}_{n=1}^\infty$ with $\varepsilon_n \to 0^+$ as $n \to \infty$ and let $x_n \in A$ be a ε_n-*minimizers *for F_n, that is*

$$F_n(x_n) \leq \max\left\{\inf_A F_n + \frac{1}{\varepsilon_n}\,,\, -\frac{1}{\varepsilon_n}\right\}\,.$$

Then every cluster point of $\{x_n\}_{n=1}^\infty$ is a minimizer of F.

Remark 2.23. The condition defining an ε-minimizer takes into account the fact that the infimum of the functional can be $-\infty$.

In the context of this paper we apply Theorem 2.22 in order to prove Corollary 1.9. In particular, we show that $\mathcal{G}_n^{(p)} + \mathcal{K}_n$ Γ-converges to $\mathcal{G}_\infty + \mathcal{K}_\infty$ and satisfies a compactness property. We note that in general

$$\Gamma\text{-}\lim_{n\to\infty}(d)(F_n + G_n) \neq \Gamma\text{-}\lim_{n\to\infty}(d)F_n + \Gamma\text{-}\lim_{n\to\infty}(d)G_n.$$

However, with a suitable strong notion of convergence of $G_n \to G$ we can infer the additivity of Γ-limits.

Proposition 2.24. *Let (A, d) be a metric space and let $F_n \overset{\Gamma\text{-}(\mathrm{d})}{\longrightarrow} F$. Assume $G_n(u_n) \to G(u)$ and $G(u) > -\infty$ for any sequence $u_n \to u$ with $\sup_{n\in\mathbb{N}} F_n(u_n) < +\infty$ and $F(u) < +\infty$ then*

$$\Gamma\text{-}\lim_{n\to\infty}(d)(F_n + G_n) = \Gamma\text{-}\lim_{n\to\infty}(d)F_n + \Gamma\text{-}\lim_{n\to\infty}(d)G_n.$$

The assumption in the above proposition is similar to the notion of continuous convergence, see [19, Definition 4.7 and Proposition 6.20]. In our context continuous convergence is not quite the right concept, indeed we define the fidelity term $\mathcal{K}_n$ on TL^p by

$$\mathcal{K}_n(u, \nu) = \begin{cases} \frac{1}{n}\sum_{i=1}^n \mathcal{K}_n(x_i, u(x_i)) & \text{if } \nu = \mu_n \\ +\infty & \text{else.} \end{cases}\,.$$

Hence $\mathcal{K}_n$ does not continuous converge to $\mathcal{K}_\infty$ (defined on TL^p analogously), however we show, in Section 5, that $\mathcal{K}_n$ does satisfy the assumptions in Proposition 2.24.

3 Convergence of the Non-Local Continuum Model

We first introduce the intermediary functional $\mathcal{F}_\varepsilon^{(p)}$ that is a non-local continuum approximation of the discrete functional $\mathcal{G}_n^{(p)}$.

Definition 3.1. *Let $p \geq 1$, $\varepsilon > 0$, $s_\varepsilon > 0$, and let $A \subset X$ be an open and bounded set. Define the functional $\mathcal{F}_\varepsilon^{(p)}(\cdot, A) : L^1(X) \to [0, \infty]$ by*

$$\mathcal{F}_\varepsilon^{(p)}(u, A) = \frac{s_\varepsilon}{\varepsilon} \int_{A\times A} \eta_\varepsilon(x-z)|u(x)-u(z)|^p \rho(x)\rho(z) \,\mathrm{d}x \,\mathrm{d}z + \frac{1}{\varepsilon}\int_A V(u(x))\rho(x)\,\mathrm{d}x\,. \tag{3}$$

When $A = X$, we will simply write $\mathcal{F}_\varepsilon^{(p)}(u)$.

This section is devoted to proving the following result.

Theorem 3.2. *Let $p \geq 1$, and assume $s_\varepsilon \to 1$ as $\varepsilon \to 0$. Under conditions (A1), (B1-3) and (C1-3) the following holds:*

- *(Compactness) Let $\varepsilon_n \to 0^+$ and $u_n \in L^1(X, \mu)$ satisfy $\sup_{n\in\mathbb{N}} \mathcal{F}_{\varepsilon_n}^{(p)}(u_n) < \infty$, then u_n is relatively compact in $L^1(X, \mu)$ and each cluster point u has $\mathcal{G}_\infty^{(p)}(u) < \infty$;*
- *(Γ-convergence) $\Gamma\text{-}\lim_{\varepsilon\to 0}(L^1)\mathcal{F}_\varepsilon^{(p)} = \mathcal{G}_\infty^{(p)}$ and furthermore, if $u \in L^1(X, \mu)$ is a polyhedral function then $\limsup_{\varepsilon\to 0} \mathcal{F}_\varepsilon^{(p)}(u) = \mathcal{G}_\infty^{(p)}(u)$.*

The result of Theorem 3.2 is a generalization of a result by Alberti and Bellettini (see [1]) that we are going to recall for the reader's convenience. First, we introduce notation for the special case when $p = 2$ and $\rho = 1$ on X.

Definition 3.3. *Let $\varepsilon > 0$, and define the functionals $\mathcal{E}_\varepsilon, \mathcal{E}_0 : L^1(X) \to [0, \infty]$ by*

$$\mathcal{E}_\varepsilon(u) := \frac{1}{\varepsilon}\int_X \int_X \eta_\varepsilon(x-z)|u(x)-u(z)|^2 \,\mathrm{d}x\,\mathrm{d}z + \frac{1}{\varepsilon}\int_X V(u(x))\,\mathrm{d}x$$

$$\mathcal{E}_0(u) := \begin{cases} \displaystyle\int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) & \textit{if } u \in BV(X, \{\pm 1\})\,, \\ +\infty & \textit{else}\,, \end{cases}$$

respectively, where

$$\hat{\sigma}(\nu) := \min\left\{ E(f;\nu) \,|\, f : \mathbb{R} \to \mathbb{R}, \lim_{t\to\infty} f(t) = 1, \lim_{t\to -\infty} f(t) = -1 \right\}$$
$$E(f;\nu) := \int_{-\infty}^{\infty}\int_{\mathbb{R}^d} \eta(h)|f(t+h_\nu) - f(t)|^2 \,\mathrm{d}h\,\mathrm{d}t + \int_{-\infty}^{\infty} V(f(t))\,\mathrm{d}t$$

and $h_\nu := h \cdot \nu$.

Remark 3.4. We make the following observations.

1. The fact that there exists a minimizer of $\hat{\sigma}(\nu)$ follows from [2, Theorem 2.4].
2. The minimum in $\hat{\sigma}(\nu)$ can be taken over non-decreasing functions with $tf(t) \geq 0$ for all $t \in \mathbb{R}$.

3. If η is isotropic, *i.e.*, $\eta(h) = \eta(|h|)$, then $E(f;\nu)$ is independent of ν and therefore $\sigma(\nu)$ is a constant and the functional $\mathcal{E}_0$ is a multiple of the perimeter $\mathcal{H}^{d-1}(\partial^*\{u=1\})$.

The following theorem is a combination of two results, [1, Theorem 1.4] and [2, Theorem 3.3].

Theorem 3.5. *Assume that $X \subset \mathbb{R}^d$ is open, V satisfies conditions (B1-3), and η satisfies conditions (C2-3). Then the following holds:*

- *(Compactness) Any sequence $\varepsilon_n \to 0^+$ and $\{u_n\} \subset L^1(X)$ with $\sup_{n\in\mathbb{N}} \mathcal{E}_{\varepsilon_n}(u_n) < \infty$ is relatively compact in $L^1(X)$, furthermore any cluster point u satisfies $\mathcal{E}_0(u) < \infty$;*
- *(Γ-convergence) $\Gamma\text{-}\lim_{\varepsilon\to 0}(L^1)\mathcal{E}_\varepsilon = \mathcal{E}_0$.*

The proof of Theorem 3.2 is based on the proof of [1, Theorem 1.4], where we have to deal with the fact that, in our case, we are considering a generic exponent $p \geq 1$, and that we have a density ρ. The compactness proof is in Section 3.1 and the Γ-convergence in Sections 3.3-3.4.

3.1 Compactness

The aim of this section is to show that any sequence $\{u_n\}_{n=1}^\infty \subset L^1(X,\mu)$ with $\sup_{n\in\mathbb{N}} \mathcal{F}^{(p)}_{\varepsilon_n}(u_n) < \infty$ is relatively compact in $L^1(X,\mu)$ and that $\mathcal{G}^{(p)}_\infty(u) < \infty$ for any cluster point $u \in L^1(X,\mu)$. This will prove the first part of Theorem 3.2.

The strategy of the proof is to apply the Alberti and Bellettini compactness result in Theorem 3.5. When $p = 2$ this follows from the upper and lower bounds on ρ that imply an 'equivalence' between $\mathcal{E}_{\varepsilon_n}$ and $\mathcal{F}^{(2)}_{\varepsilon_n}$. When $p \neq 2$ we approximate u_n with a sequence v_n satisfying $v_n(x) \in \{\pm 1\}$ then since $|v_n(x) - v_n(z)|^2 = 2^{2-p}|v_n(x) - v_n(y)|^p$ we can easily find an equivalence between $\mathcal{E}_{\varepsilon_n}$ and $\mathcal{F}^{(p)}_{\varepsilon_n}$. We start with the preliminary result that shows $\mathcal{E}_\infty(u) < \infty \implies \mathcal{G}^{(p)}_\infty(u) < \infty$.

Proposition 3.6. *Let $X \subset \mathbb{R}^d$ be open and bounded, and let $u \in L^1(X;\{\pm 1\})$. Under assumptions (A1), (B1-2), (C1,3) we have*

$$\begin{aligned}\int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty \quad &\Leftrightarrow \quad \mathcal{H}^{d-1}(\partial^*\{u=1\}) < +\infty \\ &\Rightarrow \quad \int_{\partial^*\{u=1\}} \sigma^{(p)}(x,\nu_u(x))\rho(x)\,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty.\end{aligned}$$

Remark 3.7. The above proposition implies that

$$\mathcal{E}_0(u) < +\infty \quad \Leftrightarrow \quad u \in BV(X;\{\pm 1\}) \quad \Leftrightarrow \quad \mathcal{G}^{(p)}_\infty(u) < +\infty$$

since by definition of $\mathcal{G}^{(p)}_\infty$ if $\mathcal{G}^{(p)}_\infty(u) < +\infty$ then $u \in BV(X;\{\pm 1\})$.

Remark 3.8. The missing implication,

$$\int_{\partial^*\{u=1\}} \sigma^{(p)}(x,\nu_u(x))\rho(x)\,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty \implies \mathcal{H}^{d-1}(\partial^*\{u=1\}) < +\infty\,,$$

is also true. However the natural way to prove this is to first show that the minimisation problem in $\sigma^{(p)}(x,\nu)$ (for any $x \in X$ and $\nu \in \mathbb{S}^{d-1}$) can be be reduced to a minimisation problem over functions $f : \mathbb{R} \to \mathbb{R}$ similar to $\hat{\sigma}(\nu)$ (but with an additional x dependence). This is non-trivial and would take considerable space. Since we only use the proposition to imply that if $\mathcal{E}_0(u) < +\infty$ then $\mathcal{G}_\infty^{(p)}(u) < +\infty$ then extending the result is not needed. We refer to [2, Section 3] where the authors carry out analogous computations for $\rho \equiv 1$ and $p = 2$.

Proof of Proposition 3.6. Step 1. We show

$$\mathcal{H}^{d-1}(\partial^*\{u=1\}) < +\infty \qquad \Longrightarrow \qquad \int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty.$$

Choose $\hat{f}(t) = +1$ if $t \geq 0$ and $f(t) = -1$ if $t < 0$. Then,

$$E(\hat{f};\nu) = \int_{-\infty}^{\infty}\int_{\mathbb{R}^d} \eta(h)|\hat{f}(t+h_\nu) - \hat{f}(t)|^2\,\mathrm{d}h\,\mathrm{d}t = \int_{\mathbb{R}^d} \eta(h)|h_\nu|\,\mathrm{d}h \leq R_\eta \int_{\mathbb{R}^d} \eta(h) =: C$$

where $C \in (0,\infty)$, $h_\nu = h \cdot \nu$ and R_η is given by assumption (C3). Note that C is independent of ν. So,

$$\int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) \leq \int_{\partial^*\{u=1\}} E(\hat{f};\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) \leq C\mathcal{H}^{d-1}(\partial^*\{u=1\}).$$

Step 2. We show

$$\int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x))\,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty \qquad \Longrightarrow \qquad \mathcal{H}^{d-1}(\partial^*\{u=1\}) < +\infty.$$

Using assumption (C1) it is possible to find $a > 0$ and $c > 0$ satisfying $\eta(h) \geq c$ for all $|h| \leq a$. Let $\hat{f} : \mathbb{R} \to \mathbb{R}$ such that

$$\hat{f} \text{ is non-decreasing}, \quad \lim_{t\to\infty} \hat{f}(t) = -\lim_{t\to-\infty} \hat{f}(t) = 1, \quad \hat{f}(t)t \geq 0 \text{ for } t \in \mathbb{R}. \tag{4}$$

If $\hat{f}(\frac{a}{2}) \leq \frac{1}{2}$ then $\hat{f}(t) \in [0,\frac{1}{2})$ for $t \in [0,\frac{a}{2}]$ and

$$E(\hat{f};\nu) \geq \frac{a}{2} \inf_{t\in[0,\frac{1}{2}]} V(t) =: \tilde{a}_1 > 0. \tag{5}$$

Otherwise, if $\hat{f}\left(\frac{a}{2}\right) > \frac{1}{2}$ then for $h \geq \frac{a}{2}$

$$\int_{\mathbb{R}} |\hat{f}(t+h) - \hat{f}(t)|^2\,\mathrm{d}t \geq \int_{\frac{a}{2}-h}^{0} |\hat{f}(t+h) - \hat{f}(t)|^2\,\mathrm{d}t \geq \frac{h-\frac{a}{2}}{4}.$$

Similarly, if $h \leq -\frac{a}{2}$ we have

$$\int_{\mathbb{R}} |\hat{f}(t+h) - \hat{f}(t)|^2\,\mathrm{d}t \geq \frac{\frac{a}{2}-h}{4}.$$

So, for all $|h| \geq \frac{a}{2}$ it holds

$$\int_{\mathbb{R}} |\hat{f}(t+h) - \hat{f}(t)|^2\,\mathrm{d}t \geq \frac{|h-\frac{a}{2}|}{4}.$$

Therefore

$$
\begin{aligned}
E(\hat{f};\nu) &\geq \int_{\mathbb{R}^d} \eta(h) \int_{\mathbb{R}} |\hat{f}(t+h_\nu) - \hat{f}(t)|^2 \,\mathrm{d}t \,\mathrm{d}h \\
&\geq \frac{1}{4} \int_{\{h\in\mathbb{R}^d \,:\, |h_\nu| \geq \frac{3a}{4}\}} \eta(h) \left| h_\nu - \frac{a}{2} \right| \,\mathrm{d}h \\
&\geq \frac{ac}{16} \int_{\{h\in\mathbb{R}^d \,:\, |h_\nu| \geq \frac{3a}{4}\}} \chi_{B(0,a)} \,\mathrm{d}h \\
&=: \tilde{a}_2 > 0\,. \qquad (6)
\end{aligned}
$$

Set $\tilde{a} := \min\{\tilde{a}_1, \tilde{a}_2\} > 0$. Using (5) and (6) we have, for any $f : \mathbb{R} \to \mathbb{R}$ satisfying (4), that $E(\hat{f};\nu) \geq \tilde{a}$. We get

$$
\tilde{a}\mathcal{H}^{d-1}(\partial^*\{u=1\}) \leq \int_{\partial^*\{u=1\}} \hat{\sigma}(\nu_u(x)) \,\mathrm{d}\mathcal{H}^{d-1}(x) = \mathcal{E}_0(u) < \infty\,.
$$

Step 3. We show

$$
\mathcal{H}^{d-1}(\partial^*\{u=1\}) < +\infty \quad \Longrightarrow \quad \int_{\partial^*\{u=1\}} \sigma^{(p)}(x,\nu_u(x))\rho(x) \,\mathrm{d}\mathcal{H}^{d-1}(x) < +\infty.
$$

Following the same argument as in the first part of the proof we let $\hat{f}(t) = 1$ for $t \geq 0$ and $\hat{f}(t) = -1$ for $t < 0$. Fix $\nu \in \mathbb{S}^{d-1}$ and let $f_\nu(x) = \hat{f}(x\cdot\nu)$. For any $\bar{x} \in X$ and $C \in \mathcal{C}(\bar{x},\nu)$ we clearly have $f_\nu \in \mathcal{U}(C,\nu)$. So,

$$
\begin{aligned}
\frac{1}{\mathcal{H}^{d-1}(C)} G^{(p)}(f_\nu, \rho(\bar{x}), T_C) &= \rho(\bar{x}) \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} \eta(h) |\hat{f}(t+h_\nu) - \hat{f}(t)|^p \,\mathrm{d}h \,\mathrm{d}t \\
&\leq 2^p c_2 \int_{\mathbb{R}^d} \eta(h) |h_\nu| \,\mathrm{d}h \\
&\leq \tilde{c}
\end{aligned}
$$

where $\tilde{c} \in (0,\infty)$ can be chosen to be independent of ν and in the last step we used assumption (C3). Then,

$$
\begin{aligned}
\int_{\partial^*\{u=1\}} \sigma^{(p)}(x,\nu_u(x))\rho(x) \,\mathrm{d}\mathcal{H}^{d-1}(x) &\leq \int_{\partial^*\{u=1\}} \frac{1}{\mathcal{H}^{d-1}(C)} G^{(p)}(f_\nu, \rho(\bar{x}), T_C)\rho(\bar{x}) \,\mathrm{d}\mathcal{H}^{d-1}(\bar{x}) \\
&\leq c_2 \tilde{c} \mathcal{H}^{d-1}(\partial^*\{u=1\}).
\end{aligned}
$$

This concludes the proof. □

By the above proposition it is enough to show that any sequence $\{u_n\}_{n=1}^\infty \subset L^1(X,\mu)$ satisfying $\sup_{n\in\mathbb{N}} \mathcal{F}_{\varepsilon_n}(u_n) < +\infty$ is relatively compact and any cluster point u satisfies $\mathcal{E}_0(u) < \infty$. We do this by a direct comparison with $\mathcal{E}_{\varepsilon_n}$.

Proof of Theorem 3.2 (Compactness). Assume $\varepsilon_n \to 0^+$, $s_n := s_{\varepsilon_n} \to 1$ and $\sup_{n\in\mathbb{N}} \mathcal{F}_{\varepsilon_n}(u_n) < +\infty$. Let

$$
v_n(x) := \mathrm{sign}(u_n) := \begin{cases} +1 & \text{if } u_n(x) \geq 0 \\ -1 & \text{if } u_n(x) < 0. \end{cases}
$$

We claim that

(1) $\|u_n - v_n\|_{L^1} \to 0$,

(2) $\sup_{n\in\mathbb{N}} \mathcal{F}^{(p)}_{\varepsilon_n}(v_n) < +\infty$.

Step 1. Let us first prove (1). Fix $\delta > 0$ and let

$$\begin{aligned} K_n^{(\delta)} &= \{x \in X \,:\, |u_n(x)| \geq 1+\delta\} \\ L_n^{(\delta)} &= \{x \in X \,:\, |u_n(x)| \leq 1-\delta\}\,. \end{aligned}$$

Note that for $x \in X \setminus (K_n^{(\delta)} \cup L_n^{(\delta)})$ we have $|v_n(x) - u_n(x)| \leq \delta$. Now,

$$\begin{aligned} \int_X |u_n(x) - v_n(x)|\,\mathrm{d}x &\leq \delta \mathrm{Vol}(X) + \int_{K_n^{(\delta)}} |u_n(x) - v_n(x)|\,\mathrm{d}x + \int_{L_n^{(\delta)}} |u_n(x) - v_n(x)|\,\mathrm{d}x \\ &\leq \delta \mathrm{Vol}(X) + \int_{K_n^{(\delta)}} |u_n(x)|\,\mathrm{d}x + \mathrm{Vol}(K_n^{(\delta)}) + 2\mathrm{Vol}(L_n^{(\delta)}). \end{aligned}$$

Since V is continuous and zero only at ± 1 then there exists $\gamma_\delta > 0$ such that $V(t) \geq \gamma_\delta$ for all $t \in (-\infty, -1-\delta) \cup (-1+\delta, 1-\delta) \cup (1+\delta, +\infty)$. Hence $V(u_n(x)) \geq \gamma_\delta$ for all $x \in K_n^{(\delta)} \cup L_n^{(\delta)}$. This implies,

$$\mathrm{Vol}(K_n^{(\delta)}) \leq \frac{1}{\gamma_\delta} \int_{K_n^{(\delta)}} V(u_n(x))\,\mathrm{d}x \leq \frac{\varepsilon_n}{c_1 \gamma_\delta} \mathcal{F}^{(p)}_{\varepsilon_n}(u_n).$$

By the same calculation $\mathrm{Vol}(L_n^{(\delta)}) \leq \frac{\varepsilon_n}{c_1\gamma_\delta} \mathcal{F}^{(p)}_{\varepsilon_n}(u_n)$. Furthermore,

$$\begin{aligned} \int_{K_n^{(\delta)}} |u_n(x)|\,\mathrm{d}x &= \int_{K_n^{(\delta)} \cap \{|u_n(x)| \leq R_V\}} |u_n(x)|\,\mathrm{d}x + \int_{\{|u_n(x)| > R_V\}} |u_n(x)|\,\mathrm{d}x \\ &\leq R_V \mathrm{Vol}(K_n^{(\delta)}) + \frac{1}{\tau} \int_{\{|u_n(x)| > R_V\}} V(u_n(x))\,\mathrm{d}x \\ &\leq \left(\frac{R_V}{\gamma_\delta} + \frac{1}{\tau} \right) \frac{\varepsilon_n \mathcal{F}^{(p)}_{\varepsilon_n}(u_n)}{c_1}. \end{aligned}$$

So, $\lim_{n\to\infty} \|u_n - v_n\|_{L^1} \leq \delta \mathrm{Vol}(X)$. Since this is true for all $\delta > 0$ then we have $\lim_{n\to\infty} \|u_n - v_n\|_{L^1} = 0$ which proves claim (1).

Step 2. In order to prove (2) we reason as follows. If $|u_n(x)| \geq \frac{1}{2}$ then $\mathrm{sign}(u_n(x)) \neq \mathrm{sign}(u_n(y))$ implies $|u_n(x) - u_n(y)| \geq \frac{1}{2}$. Now since,

$$|v_n(x) - v_n(y)| = \begin{cases} 0 & \text{if } \mathrm{sign}(u_n(x)) = \mathrm{sign}(u_n(y)) \\ 2 & \text{if } \mathrm{sign}(u_n(x)) \neq \mathrm{sign}(u_n(y)) \end{cases}$$

then $|v_n(x) - v_n(y)| \leq 4|u_n(x) - u_n(y)|$ when $|u_n(x)| \geq \frac{1}{2}$.

Let $M_n = \{x \in X \,:\, |u_n(x)| \leq \frac{1}{2}\}$. We have,

$$\begin{aligned} \mathcal{F}^{(p)}_{\varepsilon_n}(v_n) = \frac{s_n}{\varepsilon_n} &\int_{M_n \times X} \eta_{\varepsilon_n}(x-z) |v_n(x) - v_n(z)|^p \rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\ &+ \frac{s_n}{\varepsilon_n} \int_{M_n^c \times X} \eta_{\varepsilon_n}(x-z) |v_n(x) - v_n(y)|^p \rho(x)\rho(y)\,\mathrm{d}x\,\mathrm{d}z \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{2^p c_2^2 s_n}{\varepsilon_n}\int_{M_n\times X}\eta_{\varepsilon_n}(x-z)\,\mathrm{d}x\,\mathrm{d}z \\
&\qquad\qquad + \frac{4^p s_n}{\varepsilon_n}\int_{M_n^c\times X}\eta_{\varepsilon_n}(x-z)|u_n(x)-u_n(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\leq \frac{2^p c_2^2 s_n}{\varepsilon_n}\int_{\mathbb{R}^d}\eta(w)\,\mathrm{d}w\mathrm{Vol}(M_n) + 4^p\mathcal{F}^{(p)}_{\varepsilon_n}(u_n)
\end{aligned}
$$

where $V(t)\geq\gamma>0$ for all $t\in[\frac{1}{2},\frac{1}{2}]$. Now,

$$
\frac{1}{\varepsilon_n}\mathrm{Vol}(M_n)\leq\frac{1}{\gamma\varepsilon_n}\int_{\{|u_n(x)|\leq\frac{1}{2}\}}V(u_n(x))\,\mathrm{d}x\leq\frac{1}{\gamma c_1}\mathcal{F}^{(p)}_{\varepsilon_n}(u_n).
$$

Hence $\sup_{n\in\mathbb{N}}\mathcal{F}^{(p)}_{\varepsilon_n}(v_n)<+\infty$.

Step 3. We conclude the proof by noticing that, since $v_n\in L^1(X;\{\pm1\})$ we have $\mathcal{F}^{(p)}_{\varepsilon_n}(v_n)=2^{p-2}s_n\mathcal{E}_{\varepsilon_n}(v_n)$. So v_n is relatively compact in L^1 by Theorem 3.5 and by Proposition 3.6 $\mathcal{G}^{(p)}_\infty(u)<+\infty$ for any cluster point u of $\{v_n\}_{n=1}^\infty$. □

3.2 Preliminary results

Here we prove some technical results needed in the proof of the Γ-convergence result stated in Theorem 3.2. We start by proving some continuity properties of the function $\sigma^{(p)}$.

Lemma 3.9. *Under assumptions (A1) and (C3) the followings hold:*

(i) the function $(x,\nu)\mapsto\sigma^{(p)}(x,\nu)$ is upper semi-continuous on $X\times\mathbb{S}^{d-1}$,

(ii) for every $\nu\in\mathbb{S}^{d-1}$, the function $x\mapsto\sigma^{(p)}(x,\nu)$ is continuous on X.

Proof. Proof of (i). Fix $\bar{x}\in X$ and $\nu\in\mathbb{S}^{d-1}$ and let $\{x_n\}_{n=1}^\infty\subset X$ and $\{\nu_n\}_{n=1}^\infty\subset\mathbb{S}^{d-1}$ such that $x_n\to\bar{x}$ and $\nu_n\to\nu$ as $n\to\infty$. Let R_n be a rotation such that $R_n\nu=\nu_n$. Fix $t>0$ and let $D\in\mathcal{C}(\bar{x},\nu)$ and $w\in\mathcal{U}(D,\nu)$ be such that

$$
\frac{1}{\mathcal{H}^{d-1}(D)}G^{(p)}\left(w,\rho(\bar{x}),T_D\right)\leq\sigma^{(p)}(\bar{x},\nu)+t\,.
$$

Notice that without loss of generality we can assume $|w|\leq 1$. For $n\in\mathbb{N}$ define $C_n\in\mathcal{C}(x_n,\nu_n)$ and $u_n\in\mathcal{U}(C_n,\nu_n)$ by

$$
C_n:=R_n(D-\bar{x})+x_n\,,
$$

$$
u_n(x):=w(R_n^{-1}(x-x_n)+\bar{x})\,,
$$

respectively. Then

$$
\begin{aligned}
\sigma^{(p)}(x_n,\nu_n)&\leq\frac{1}{\mathcal{H}^{d-1}(C_n)}G^{(p)}(u_n,\rho(x_n),T_{C_n}) \\
&\leq\frac{1}{\mathcal{H}^{d-1}(D)}G^{(p)}(w,\rho(\bar{x}),T_D)+\delta_n \\
&\leq\sigma^{(p)}(\bar{x},\nu)+t+\delta_n\,,
\end{aligned}
$$

where

$$\delta_n := \frac{1}{\mathcal{H}^{d-1}(D)} \left| G^{(p)}(u_n, \rho(x_n), T_{C_n}) - G^{(p)}(w, \rho(\bar{x}), T_D) \right| .$$

We claim that $\delta_n \to 0$ as $n \to \infty$. Since $t > 0$ is arbitrary, this will prove the upper semi-continuity. Notice that, for every $h \in B_{R_\eta}(0)$, the following inequality holds

$$\int_{T_D} |w(z+h) - w(z)|^p \, \mathrm{d}z \leq 2^p \mathcal{H}^{d-1}(D) R_\eta^p . \tag{7}$$

Indeed, setting

$$v(z) := \begin{cases} +1 & \text{if } z \cdot \nu \geq 0 \, , \\ -1 & \text{if } z \cdot \nu < 0 \, , \end{cases}$$

we get

$$\begin{aligned} \int_{T_D} |w(z+h) - w(z)|^p \, \mathrm{d}z &\leq \int_{T_{D_n}} |v(z+h) - v(z)|^p \, \mathrm{d}z \\ &= 2^p \mathcal{H}^{d-1}(D) |h \cdot \nu|^p \\ &\leq 2^p \mathcal{H}^{d-1}(D) R_\eta^p . \end{aligned}$$

where in the first inequality we used the convexity of the function $s \mapsto |s|^p$. Fix $\varepsilon > 0$. Using the continuity of ρ, for n sufficiently large we have that $|\rho(x_n) - \rho(\bar{x})| < \varepsilon$. Thus, using (7), we get

$$\begin{aligned} \delta_n &= \frac{1}{\mathcal{H}^{d-1}(D)} \Bigg| \rho(x_n) \int_{T_D} \int_{\mathbb{R}^d} \eta(R_n h) \, |w(z+h) - w(z)|^p \, \mathrm{d}h \, \mathrm{d}z \\ &\qquad\qquad - \rho(\bar{x}) \int_{T_D} \int_{\mathbb{R}^d} \eta(h) \, |w(z+h) - w(z)|^p \, \mathrm{d}h \, \mathrm{d}z \Bigg| \\ &\leq \frac{|\rho(x_n) - \rho(\bar{x})|}{\mathcal{H}^{d-1}(D)} \int_{\mathbb{R}^d} \eta(h) \int_{T_D} |w(z+h) - w(z)|^p \, \mathrm{d}z \, \mathrm{d}h \\ &\qquad\qquad + \frac{\rho(x_n)}{\mathcal{H}^{d-1}(D)} \int_{\mathbb{R}^d} |\eta(R_n h) - \eta(h)| \int_{T_D} |w(z+h) - w(z)|^p \, \mathrm{d}z \, \mathrm{d}h \\ &\leq 2^p R_\eta^p \left(\varepsilon R_\eta^d \omega_d \|\eta\|_{L^\infty} + c_2 \int_{\mathbb{R}^d} |\eta(R_n h) - \eta(h)| \, \mathrm{d}h \right) \end{aligned}$$

where ω_d denotes the volume of the unit ball in $\mathbb{R}^d$. In order to show that the second term in the parenthesis vanishes, we use an argument similar to the one for proving that translations are continuous in L^p. For every $s > 0$ let $\widetilde{\eta}_s : \mathbb{R} \to [0, \infty)$ be a continuous function with support in B_{2R_η} such that $\|\eta - \widetilde{\eta}_s\|_{L^1(\mathbb{R}^d)} < s$. Then, for every $r > 0$ there exists $\bar{n} \in \mathbb{N}$ such that $|\widetilde{\eta}_s(R_n h) - \widetilde{\eta}_s(h)| < r$ for all $h \in \mathbb{R}^d$ and all $n \geq \bar{n}$. So that, for $n \geq \bar{n}$

$$\int_{\mathbb{R}^d} |\eta(R_n h) - \eta(h)| \, \mathrm{d}h \leq \|\eta - \widetilde{\eta}_s\|_{L^1(\mathbb{R}^d)} + \int_{\mathbb{R}^d} |\widetilde{\eta}_s(R_n h) - \widetilde{\eta}_s(h)| \, \mathrm{d}h \leq s + r \mathrm{Vol}(B_{2R_\eta}) .$$

Since r and s are arbitrary, we conclude that $\int_{\mathbb{R}^d} |\eta(R_n h) - \eta(h)| \, \mathrm{d}h \to 0$ as $n \to \infty$ and, in turn, that $\delta_n \to 0$ as $n \to \infty$.

Proof of (ii). Fix $\nu \in \mathbb{S}^{d-1}$, $\bar{x} \in X$ and let $x_n \to \bar{x}$.

Step 1. We claim that $\sigma^{(p)}(\bar{x},\nu) \leq \liminf_{n\to\infty} \sigma^{(p)}(x_n,\nu)$. Without loss of generality, let us assume that

$$\liminf_{n\to\infty} \sigma^{(p)}(x_n,\nu) = \lim_{n\to\infty} \sigma^{(p)}(x_n,\nu) < \infty\,. \tag{8}$$

For every $n \in \mathbb{N}$ let $C_n \in \mathcal{C}(x_n,\nu)$ and $u_n \in \mathcal{U}(C_n,\nu)$ be such that

$$\frac{1}{\mathcal{H}^{d-1}(C_n)} G^{(p)}(u_n, \rho(x_n), T_{C_n}) \leq \sigma^{(p)}(x_n,\nu) + \frac{1}{n}. \tag{9}$$

Set $\lambda_n := \bar{x} - x_n$ and define

$$\widetilde{C}_n := C_n + \lambda_n\,, \qquad \widetilde{u}_n(x) := u_n(x-\lambda_n).$$

So, $\widetilde{C}_n \in \mathcal{C}(\bar{x},\nu)$ and $\widetilde{u}_n \in \mathcal{U}(\widetilde{C}_n,\nu)$. Using (9), we get

$$\begin{aligned}
\sigma^{(p)}(\bar{x},\nu) &\leq \frac{1}{\mathcal{H}^{d-1}(\widetilde{C}_n)} G^{(p)}(\widetilde{u}_n, \rho(\bar{x}), T_{\widetilde{C}_n}) \\
&\leq \frac{G^{(p)}(u_n,\rho(x_n),T_{C_n})}{\mathcal{H}^{d-1}(C_n)} + \frac{1}{\mathcal{H}^{d-1}(C_n)} \left| G^{(p)}(\widetilde{u}_n, \rho(\bar{x}), T_{\widetilde{C}_n}) - G^{(p)}(u_n,\rho(x_n),T_{C_n}) \right| \\
&\leq \sigma^{(p)}(x_n,\nu) + \frac{1}{n} + \frac{1}{\mathcal{H}^{d-1}(C_n)} \left| G^{(p)}(\widetilde{u}_n, \rho(\bar{x}), T_{\widetilde{C}_n}) - G^{(p)}(u_n,\rho(x_n),T_{C_n}) \right| \qquad (10)
\end{aligned}$$

To estimate the last term, we reason as follows. First of all, we notice that

$$\int_{T_{C_n}} V\left(u_n(z)\right) \mathrm{d}z = \int_{T_{\widetilde{C}_n}} V\left(\widetilde{u}_n(z)\right) \mathrm{d}z\,. \tag{11}$$

Fix $\varepsilon > 0$. Using the continuity of ρ, there exists $\bar{n} \in \mathbb{N}$, such that $|\rho(x_n) - \rho(\bar{x})| < \varepsilon$ for all $n \geq \bar{n}$. From (11) we get that

$$\begin{aligned}
&\frac{1}{\mathcal{H}^{d-1}(C_n)} \left| G^{(p)}(\widetilde{u}_n, \rho(\bar{x}), T_{\widetilde{C}_n}) - G^{(p)}(u_n,\rho(x_n),T_{C_n}) \right| \\
&\qquad \leq \frac{1}{\mathcal{H}^{d-1}(C_n)} |\rho(x_n) - \rho(\bar{x})| \int_{T_{C_n}} \int_{\mathbb{R}^d} \eta(h) \left|u_n(z+h) - u_n(z)\right|^p \,\mathrm{d}h\,\mathrm{d}z \\
&\qquad \leq \frac{\varepsilon}{\mathcal{H}^{d-1}(C_n)} \int_{T_{C_n}} \int_{\mathbb{R}^d} \eta(h) \left|u_n(z+h) - u_n(z)\right|^p \,\mathrm{d}h\,\mathrm{d}z \\
&\qquad \leq \varepsilon 2^p R_\eta^{d+p} \omega_d \|\eta\|_{L^\infty}\,,
\end{aligned}$$

where in the last step we used (7). Using the arbitrariness of ε, together with (8) and (9), we conclude that

$$\frac{1}{\mathcal{H}^{d-1}(C_n)} \left| G^{(p)}(\widetilde{u}_n, \rho(\bar{x}), T_{\widetilde{C}_n}) - G^{(p)}(u_n,\rho(x_n),T_{C_n}) \right| \to 0\,,$$

as $n \to \infty$. Thus, by taking the liminf in (10), we conclude that

$$\sigma^{(p)}(\bar{x},\nu) \leq \liminf_{n\to\infty} \sigma^{(p)}(x_n,\nu)\,.$$

Step 2. With a similar argument, it is possible to prove that $\sigma^{(p)}(\bar{x},\nu) \geq \limsup_{n\to\infty} \sigma^{(p)}(x_n,\nu)$. This concludes the proof of the continuity of the map $x \mapsto \sigma^{(p)}(x,\nu)$. □

Remark 3.10. Notice that the above result did not require the existence of a solution for the infimum problem defining $\sigma^{(p)}$.

We notice that the main feature of the Γ-convergence of $\mathcal{F}^{(p)}_{\varepsilon_n}$ to $\mathcal{G}^{(p)}_{\infty}$ is that we recover, in the limit, a local functional starting from nonlocal ones. To be more precise, let $A, B \subset X$ be disjoint sets. Then, it holds that

$$\mathcal{F}^{(p)}_{\varepsilon_n}(u, A \cup B) = \mathcal{F}^{(p)}_{\varepsilon_n}(u, A) + \mathcal{F}^{(p)}_{\varepsilon_n}(u, B) + 2\widetilde{\Lambda}_{\varepsilon_n}(u, A, B) \tag{12}$$

where we define the nonlocal deficit

$$\widetilde{\Lambda}_{\varepsilon}(u, A, B) := \frac{s_\varepsilon}{\varepsilon} \int_A \int_B \eta_\varepsilon(x - z)|u(x) - u(z)|^p \rho(x)\rho(z)\, \mathrm{d}x\, \mathrm{d}z\,. \tag{13}$$

On the other hand, for the limiting functional we have

$$\mathcal{G}^{(p)}_{\infty}(u, A \cup B) = \mathcal{G}^{(p)}_{\infty}(u, A) + \mathcal{G}^{(p)}_{\infty}(u, B)\,, \tag{14}$$

where, for $u \in BV(X; \pm 1)$, we set

$$\mathcal{G}^{(p)}_{\infty}(u, A) := \int_{\partial^*\{u=1\}\cap A} \sigma^{(p)}(x, \nu_u(x))\rho(x)\, \mathrm{d}\mathcal{H}^{d-1}(x)\,.$$

Identity (12) states that the functionals $\mathcal{F}^{(p)}_{\varepsilon_n}$ are nonlocal, while (14) is the locality property of the limiting functional $\mathcal{G}^{(p)}_{\infty}$. Thus, we expect the nonlocal deficit to disappear in the limit, *i.e.*, that if $u_{\varepsilon_n} \to u$ in $L^1(X)$, then

$$\widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n}, A, B) \to 0\,, \tag{15}$$

as $n \to \infty$. For technical reasons we also need the nonlocal deficit's without weighting by ρ or s_ε:

$$\Lambda_{\varepsilon}(u, A, B) := \frac{1}{\varepsilon} \int_A \int_B \eta_\varepsilon(x - z)|u(x) - u(z)|^p\, \mathrm{d}x\, \mathrm{d}z\,.$$

By continuity of ρ if A and B are sets in X that are close to $\bar{x}$ then $\widetilde{\Lambda}_\varepsilon(u, A, B) \approx s_\varepsilon \rho^2(\bar{x})\Lambda_\varepsilon(u, A, B)$. In [1] the authors prove that the limit of the nonlocal deficit is determined by the behavior of u_{ε_n} close to the boundaries of A and B and, in turn, that (15) holds in certain cases of interest. Here we only state the main technical result of [1] in a version we need in the paper, addressing the interested reader to the paper by Alberti and Bellettini for the details.

Proposition 3.11. *Let $v_n \to v$ in $L^1(X)$ with $|v_n| \leq 1$. Then, for all $\bar{x} \in \mathbb{R}^d$ and for all $\nu \in \mathbb{S}^{d-1}$ the following holds: given $C \in \mathcal{C}(\bar{x}, \nu)$ consider the strip T_C and any cube $Q \subset \mathbb{R}^d$ whose intersection with $\nu^\perp$ is C. Then, for a.e. $t > 0$:*

- *(i)* $\Lambda_{\varepsilon_n}(v_n, tT_C, \mathbb{R}^d \setminus tT_C) \to 0$ *as* $n \to \infty$,
- *(ii)* $\Lambda_{\varepsilon_n}(v_n, tQ, tT_C \setminus tQ) \to 0$ *as* $n \to \infty$.

Remark 3.12. The boundness assumption on the sequence $\{v_n\}_{n=1}^\infty$ allows one to obtain the proof of the above result directly from [1, Proposition 2.5 and Theorem 2.8]. In particular in [1] the authors prove the result for $p = 1$ and $\rho \equiv 1$, using the L^∞ bound on v_n one can easily bound the more general case considered here by the L^1 case. With similar computations it is also possible to obtain the same result without the L^∞ bound.

Finally, notice that when $A, B \in \mathbb{R}^d$ are disjoint sets with $\mathrm{d}(A,B) > 0$, using the fact that the function η has support in the ball $B(0, R_\eta)$ (see (C3)), it is easy to prove that there exists $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$ it holds

$$\Lambda_{\varepsilon_n}(v_n, A, B) = 0\,.$$

For technical reasons we need to introduce a scaled version of the functional $G^{(p)}$.

Definition 3.13. *For $\varepsilon > 0$, $p \geq 1$, $u : \mathbb{R}^d \to \mathbb{R}$, $\lambda \in \mathbb{R}$, and $A \subset \mathbb{R}^d$, we define*

$$G^{(p)}_\varepsilon(u, \lambda, A) := \frac{\lambda}{\varepsilon} \int_A \int_{\mathbb{R}^d} \eta_\varepsilon(h) |u(z+h) - u(z)|^p \,\mathrm{d}h \,\mathrm{d}z + \frac{1}{\varepsilon} \int_A V(u(z)) \,\mathrm{d}z\,.$$

Let $r > 0$ and $x \in X$. For a set $A \subset \mathbb{R}^d$, we define $x + rA := \{x + ry \,:\, y \in A\}$. Moreover, for a function $u : \mathbb{R}^d \to \mathbb{R}$, we set

$$R_{x,r}u(y) := u(x + ry)\,. \tag{16}$$

Using a change of variable, it is easy to see that the following scaling property holds true:

$$G^{(p)}_\varepsilon(u, \lambda, x + rA) = r^{d-1} G^{(p)}_{\varepsilon/r}(R_{x,r}u, \lambda, A)\,. \tag{17}$$

3.3 The Liminf Inequality

This section is devoted at proving the following: let $u_{\varepsilon_n} \to u$ in L^1, then

$$\mathcal{G}^{(p)}_\infty(u) \leq \liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n})\,. \tag{18}$$

We will follow the proof of [1, Theorem 1.4], with some modifications due to the presence of the density ρ.

Proof of Theorem 3.2 (Liminf). Let $\varepsilon_n \to 0^+$ and $u_{\varepsilon_n} \to u$ in $L^1(X, \mu)$. Assume without loss of generality that

$$\liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}) = \lim_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}) < \infty\,. \tag{19}$$

Step 1. By compactness (see Section 3.1) it holds $u = \chi_A$ for some set $A \subset X$ of finite perimeter in X. In order to prove (18) we use the strategy introduced by Fonseca and Müller in [25]. Write

$$\mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}) = \int_X g_{\varepsilon_n}(x) \mathrm{d}x \tag{20}$$

and set $\mathrm{d}\lambda_{\varepsilon_n} := g_{\varepsilon_n} \mathrm{d}\mathcal{L}^d \llcorner X$, so that

$$|\lambda_{\varepsilon_n}|(X) = \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n})\,. \tag{21}$$

Using (19), (20), (21) , up to a subsequence (not relabeled) it holds $\lambda_{\varepsilon_n} \stackrel{*}{\rightharpoonup} \lambda$ for some finite Radon measure λ on X. Then

$$|\lambda|(X) \leq \liminf_{n\to\infty} |\lambda_{\varepsilon_n}|(X)\,. \tag{22}$$

In view of (21) and (22), the liminf inequality (18) is implied by the following claim: for $\mathcal{H}^{d-1}$-a.e. $\bar{x} \in \partial^*\{u = 1\}$ it holds

$$\sigma^{(p)}(\bar{x}, \nu(\bar{x}))\rho(\bar{x}) \leq \frac{\mathrm{d}\lambda}{\mathrm{d}\theta}(\bar{x}) ,$$

where $\theta := \mathcal{H}^{d-1} \llcorner \partial^*\{u = 1\}$. In order to prove the claim we reason as follows. For $\mathcal{H}^{d-1}$-a.e. $\bar{x} \in \partial^*\{u = 1\}$ it is possible to find the density of λ with respect to θ via (recall Remark 2.17)

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\theta}(\bar{x}) = \lim_{r \to 0} \frac{\lambda(\bar{x} + rQ)}{r^{d-1}} , \tag{23}$$

where Q is a unit cube centered at the origin and having $\nu(\bar{x})$, the measure theoretic exterior normal to A at $\bar{x}$, as one of its axes. Let $\bar{x} \in \partial^*\{u = 1\}$. Theorem 2.16 implies that

$$R_{\bar{x},r} u \to v_{\bar{x}} \tag{24}$$

in $L^1_{loc}(\mathbb{R}^N)$ as $r \to 0$, where

$$v_{\bar{x}}(x) := \begin{cases} -1 & x \cdot \nu(\bar{x}) \geq 0 , \\ 1 & x \cdot \nu(\bar{x}) < 0 . \end{cases}$$

Let $\bar{x} \in \partial^*\{u = 1\}$ be a point for which (23) and (24) hold. Without loss of generality, we can assume that

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\theta}(\bar{x}) < \infty . \tag{25}$$

Since $u_{\varepsilon_n} \to u$ in L^1 and $\lambda_{\varepsilon_n} \overset{*}{\rightharpoonup} \lambda$, it is possible to find a (not relabeled) subsequence $\{\varepsilon_n\}_{n=1}^\infty$ and a sequence $\{r_n\}_{n=1}^\infty$ with $r_n \to 0^+$ and $\frac{\varepsilon_n}{r_n} \to 0^+$, such that

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\theta}(\bar{x}) = \lim_{n \to \infty} \frac{\lambda_{\varepsilon_n}(\bar{x} + r_n Q)}{r_n^{d-1}} \tag{26}$$

and

$$R_{\bar{x},r_n} u_{\varepsilon_n} \to v_{\bar{x}} .$$

Using the fact that X is open, we can assume that $\bar{x} + r_n Q \subset X$ for all $n \in \mathbb{N}$. Thus

$$\frac{\lambda_{\varepsilon_n}(\bar{x} + r_n Q)}{r_n^{d-1}} \geq \frac{\mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, \bar{x} + r_n Q)}{r_n^{d-1}} . \tag{27}$$

Step 2. We claim that

$$\delta_n := \frac{|\mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, \bar{x} + r_n Q) - \rho(\bar{x})\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, \rho(\bar{x}), \bar{x} + r_n Q)|}{r_n^{d-1}} \to 0 , \tag{28}$$

as $n \to \infty$, where

$$\widetilde{\mathcal{F}}^{(p)}_{\varepsilon}(u, \xi, A) := \frac{\xi s_\varepsilon}{\varepsilon} \int_A \int_A \eta_\varepsilon(x - z)|u(x) - u(z)|^p \, \mathrm{d}x \, \mathrm{d}z + \frac{1}{\varepsilon} \int_A V(u(x)) \, \mathrm{d}x ,$$

for $\varepsilon > 0$, $A \subset X$, $u : A \to \mathbb{R}$ and $\xi \in \mathbb{R}$. Indeed, fix $t > 0$. Thanks to hypothesis (A1) the function ρ is continuous in X. Then, it is possible to find $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$ and all $y \in \bar{x} + r_n Q$ it holds

$$|\rho(y) - \rho(\bar{x})| < t .$$

Thus,

$$
\begin{aligned}
\delta_n &\le \frac{t}{r_n^{d-1}\varepsilon_n}\int_{\bar{x}+r_nQ} V(u_{\varepsilon_n}(x))\,\mathrm{d}x + \frac{ts_{\varepsilon_n}\rho(\bar{x})}{r_n^{d-1}\varepsilon_n}\int_{\bar{x}+r_nQ}\int_{\bar{x}+r_nQ}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\,\mathrm{d}x\,\mathrm{d}z \\
&\quad + \frac{ts_{\varepsilon_n}}{r_n^{d-1}\varepsilon_n}\int_{\bar{x}+r_nQ}\int_{\bar{x}+r_nQ}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\le \frac{t}{r_n^{d-1}\varepsilon_n c_1}\int_{\bar{x}+r_nQ} V(u_{\varepsilon_n}(x))\rho(x)\,\mathrm{d}x \\
&\quad + \frac{ts_{\varepsilon_n}(c_1+c_2)}{r_n^{d-1}\varepsilon_n c_1^2}\int_{\bar{x}+r_nQ}\int_{\bar{x}+r_nQ}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\rho(z)\rho(x)\,\mathrm{d}x\,\mathrm{d}z \\
&\le \frac{t(c_1+c_2)}{c_1^2}\frac{\lambda_{\varepsilon_n}(\bar{x}+r_nQ)}{r_n^{d-1}},
\end{aligned}
$$

where in the last step we used (27). By (25) and (26) $\lim_{n\to\infty}\delta_n \le Ct$ for some constant $C<\infty$. Since $t>0$ is arbitrary, this proves the claim.

Step 3. Observe that for any $\lambda \ge 0$, $\varepsilon > 0$, $r > 0$ and $v \in L^1$ we have

$$
\min\left\{1, \frac{s_\varepsilon}{s_{\frac{\varepsilon}{r}}}\right\}\tilde{\mathcal{F}}^{(p)}_{\frac{\varepsilon}{r}}(R_{\bar{x},r}v,\lambda,Q) \le \frac{1}{r^{d-1}}\tilde{\mathcal{F}}^{(p)}_{\varepsilon}(v,\lambda,\bar{x}+rQ) \le \max\left\{1, \frac{s_\varepsilon}{s_{\frac{\varepsilon}{r}}}\right\}\tilde{\mathcal{F}}^{(p)}_{\frac{\varepsilon}{r}}(R_{\bar{x},r}v,\lambda,Q).
$$

Let $C = Q\cap\nu(\bar{x})^\perp \in \mathcal{C}(\bar{x},\nu(\bar{x}))$. Define the function $w_n:\mathbb{R}^d\to\mathbb{R}$ as the periodic extension of the function that is $R_{\bar{x},r_n}u_{\varepsilon_n}$ in Q and $v_{\bar{x}}$ in $T_C\setminus Q$. Set $\varepsilon_n' := \frac{\varepsilon_n}{r_n}$ and $s_n' = \min\left\{1, \frac{s_{\varepsilon_n}}{s_{\varepsilon_n'}}\right\}$. Using (27) and (28) together with the scaling identity (17) we get

$$
\begin{aligned}
\frac{\lambda_{\varepsilon_n}(\bar{x}+r_nQ)}{r_n^{d-1}} &\ge \rho(\bar{x})\frac{\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n},\rho(\bar{x}),\bar{x}+r_nQ)}{r_n^{d-1}} - \delta_n \\
&\ge s_n'\rho(\bar{x})\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(R_{\bar{x},r_n}u_{\varepsilon_n},\rho(\bar{x}),Q) - \delta_n \\
&= s_n'\rho(\bar{x})\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),Q) - \delta_n \\
&\ge s_n'\rho(\bar{x})G^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),T_C) - \delta_n \\
&\qquad - s_n'\rho(\bar{x})\left|\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),Q) - G^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),T_C)\right| \\
&= s_n'\rho(\bar{x})\left(\frac{\varepsilon_n}{r_n}\right)^{d-1} G^{(p)}\left(R_{0,\varepsilon_n'}w_n,\rho(\bar{x}),\frac{r_n}{\varepsilon_n}T_C\right) - \delta_n \\
&\qquad - s_n'\rho(\bar{x})\left|\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),Q) - G^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),T_C)\right|.
\end{aligned}
\tag{29}
$$

We would like to say that

$$
\left|\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),Q) - G^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),T_C)\right| \to 0
$$

as $n\to\infty$. Unfortunately, this might not be true. In order to overcome this difficulty, take $t\in(0,1)$. Notice that we can bound

$$
\left|\widetilde{\mathcal{F}}^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),tQ) - G^{(p)}_{\varepsilon_n'}(w_n,\rho(\bar{x}),tT_C)\right| \le 2\rho(\bar{x})\Lambda_{\varepsilon_n'}(w_n,tQ,tT_C\setminus tQ)
$$

$$
\begin{aligned}
&+ \rho(\bar{x})\Lambda_{\varepsilon_n'}(w_n, tT_C, \mathbb{R}^d \setminus tT_C) + \rho(\bar{x})\Lambda_{\varepsilon_n'}(w_n, tT_C \setminus tQ, tT_C \setminus tQ) \\
&+ \frac{r_n \rho(\bar{x})}{\varepsilon_n} |1 - s_{\varepsilon_n'}| \int_{tQ} \int_{tQ} \eta_{\frac{\varepsilon_n}{r_n}}(y - z) |w_n(y) - w_n(z)|^p \, \mathrm{d}y \, \mathrm{d}z \,. \qquad (30)
\end{aligned}
$$

Now,

$$
\begin{aligned}
\frac{r_n \rho(\bar{x})}{\varepsilon_n} \int_{tQ} \int_{tQ} \eta_{\varepsilon_n'}(y - z) |w_n(y) - w_n(z)|^p \, \mathrm{d}y \, \mathrm{d}z &\leq G^{(p)}_{\varepsilon_n'}(w_n, \rho(\bar{x}), T_C) \qquad (31) \\
&= \left(\frac{\varepsilon_n}{r_n}\right)^{d-1} G^{(p)}\left(R_{0,\varepsilon_n'} w_n, \rho(\bar{x}), \frac{r_n}{\varepsilon_n} T_C\right) .
\end{aligned}
$$

Moreover, using Proposition 3.11 we get that for a.e. $t \in (0, 1)$ it holds

$$
\Lambda_{\varepsilon_n'}(w_n, tQ, tT_C \setminus tQ) \to 0 \,, \qquad \Lambda_{\varepsilon_n'}(w_n, tT_C, \mathbb{R}^d \setminus tT_C) \to 0 \qquad (32)
$$

as $n \to \infty$. Finally, using Remark 3.12 and the fact that w_n is constant on $T_C \setminus Q$ it is easy to see that

$$
\lim_{t \to 1} \lim_{n \to \infty} \Lambda_{\varepsilon_n'}(w_n, tT_C \setminus tQ, tT_C \setminus tQ) = 0 \,. \qquad (33)
$$

Hence, from (28), (29), (30), (31), (32) and (33) and recalling that $s_n' \to 1$ we get

$$
\lim_{n \to \infty} \frac{\lambda_{\varepsilon_n}(\bar{x} + r_n Q)}{r_n^{d-1}} \geq \rho(\bar{x}) \sigma^{(p)}(\bar{x}, \nu(\bar{x}))
$$

as required $\square$

3.4 The Limsup Inequality

This section is devoted at proving the following: let $u \in BV(X, \{\pm 1\})$, then it is possible to find $\{u_{\varepsilon_n}\}_{n=1}^\infty \subset L^1(X)$ with $u_{\varepsilon_n} \to u$ in $L^1(X, \mu)$ such that

$$
\limsup_{n \to \infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}) \leq \mathcal{G}^{(p)}_{\infty}(u) \,. \qquad (34)
$$

Without loss of generality, we can assume $\mathcal{G}^{(p)}_{\infty}(u) < \infty$, namely $u \in BV(X; \pm 1)$. The proof will follow the lines of the argument used to prove [1, Theorem 5.2].

Proof of Theorem 3.2 (limsup). We first prove the result for polyhedral functions then, via a diagonalisation argument, generalise to arbitrary functions in $BV(X; \pm 1)$. We fix the sequence $\varepsilon_n \to 0^+$ now.

Step 1. Polyhedral functions. Assume $u \in BV(X, \{\pm 1\})$ is a polyhedral function (see Definition 2.18). Then we claim that there exists a sequence $\{u_{\varepsilon_n}\}_{n=1}^\infty$ with $|u_{\varepsilon_n}| \leq 1$, converging uniformly to u on every compact set $K \subset X \setminus \partial^*\{u = 1\}$, and in particular $u_{\varepsilon_n} \to u$ in $L^1(X, \mu)$, such that (34) holds.

Let us denote by E the polyhedral set $\{u = 1\}$ and by $E_1, \ldots, E_k$ its faces. It is possible to cover $\partial E \cap X$ with a finite family of sets $\bar{A}_1, \ldots, \bar{A}_k$, where each A_i is an open set satisfying the following properties:

(i) ∂A_i can be written as the union of two Lipschitz graphs over the face E_i,

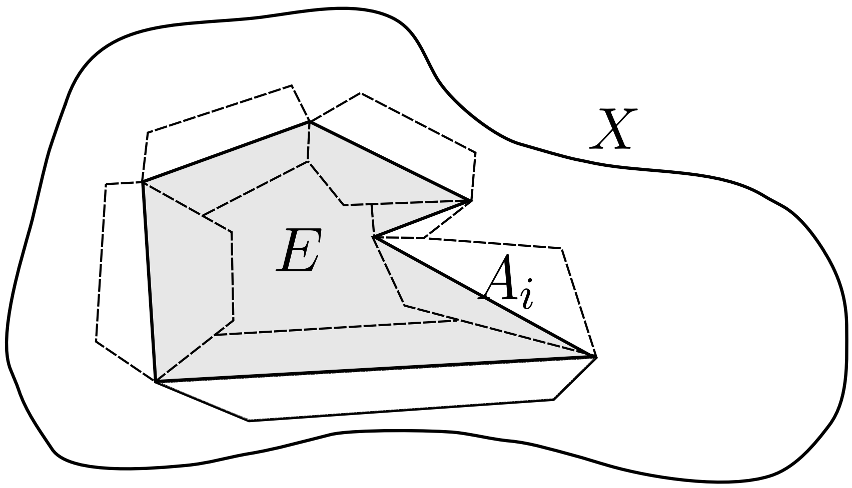

Figure 1: The polyhedral set E (shaded) and the sets A_i (dotted lines).

(ii) every point in the relative interior of E_i belongs to A_i,

(iii) $\mathcal{H}^{d-1}\left(\bar{A}_i \cap \bigcup_{j\neq i} E_j\right) = 0$,

(iv) $A_i \cap A_j = \emptyset$ if $i \neq j$.

Set

$$A_0 := \{u = 1\} \setminus \bigcup_{i=1}^{k} \bar{A}_i \,, \qquad A_{k+1} := X \setminus \bigcup_{i=0}^{k} \bar{A}_i \,.$$

We then define u_{ε_n} in each A_i separately. Set $u_{\varepsilon_n}(x) := 1$ for $x \in A_0$ and $u_{\varepsilon_n}(x) := -1$ for $x \in A_{k+1}$. Now fix $i \in \{1, \dots, k\}$ and $n \in \mathbb{N}$. In order to define u_{ε_n} in A_i, we reason as follows. Denote by ν the normal of the hyperplane containing E_i. Without loss of generality, we can assume $\nu = e_d$ and $E_i \subset \{x_d = 0\}$. A point $x \in \mathbb{R}^d$ will be denoted as

$$x = (x', x_d) \,, \qquad x' \in \mathbb{R}^{d-1} \,, \quad x_d \in \mathbb{R} \,.$$

Fix $\xi > 0$. Using the continuity of σ (see Lemma 3.9) and of ρ in X (see (A1)) it is possible to find a finite family of $d-1$ dimensional disjoint cubes $\{Q_j\}_{j=1}^{M_\xi}$, for some $M_\xi \in \mathbb{N}$, of side $r_\xi > 0$ lying in the hyperplane containing E_i, having $E_i \subseteq \cup_{j=1}^{M_\xi} Q_j$, $E_i \cap Q_j \neq \emptyset$, and satisfying the following properties: denoting by $\{x_j\}_{j=1}^{M_\xi}$ their centers (or, in the case the center of a cube Q_j is not contained in E_i, a point of $Q_j \cap E_i$) we have

$$\left| \int_{E_i} \sigma^{(p)}(x, \nu)\rho(x) \, \mathrm{d}\mathcal{H}^{d-1}(x) - r_\xi^{d-1} \sum_{j=1}^{M_\xi} \sigma^{(p)}(x_j, \nu)\rho(x_j) \right| < \eta \,. \tag{35}$$

It is possible to find, for every $j = 1, \dots M_\xi$, $C_j \in \mathcal{C}(x_j, \nu)$ and $w_j \in \mathcal{U}(C_j, \nu)$ such that

$$\frac{1}{\mathcal{H}^{d-1}(C_j)} G^{(p)}(w_j, \rho(x_j), T_{C_j}) < \sigma^{(p)}(x_j, \nu) + \frac{\xi}{\rho(x_j) M_\xi r_\xi^{d-1}} \,. \tag{36}$$

We can assume $|w_j| \leq 1$. For every $j = 1, \dots, M_\xi$, let $L_j \in \mathbb{N}$ be such that

$$\frac{1}{\mathcal{H}^{d-1}(C_j)} \left| G^{(p)}(w_j, \rho(x_j), T_{C_j}) - G^{(p)}(\widetilde{w}_j, \rho(x_j), T_{C_j}) \right| < \frac{\xi}{\rho(x_j) M_\xi r_\xi^{d-1}} , \tag{37}$$

where

$$\widetilde{w}_j(x) := \begin{cases} w_j(x) & \text{if } |x_d| < L_j , \\ +1 & \text{if } x_d > L_j , \\ -1 & \text{if } x_d < -L_j . \end{cases} \tag{38}$$

For every $j = 1, \dots, M_\xi$ cover Q_j with copies of $\varepsilon_n C_j$. Denote them by $\{Q^{(n)}_{j,s}\}_{s=1}^{k_j^{(n)}}$ and by $\{y^{(n)}_{j,s}\}_{s=1}^{k_j^{(n)}}$ their centers. Notice that it might be necessary to consider the intersection of some cubes with Q_j, and that

$$k_j^{(n)} = \frac{1}{\varepsilon_n^{d-1}} r_\xi^{d-1} \left(\mathcal{H}^{d-1}(C_j) \right)^{-1} . \tag{39}$$

Define the function $v_j^{(n)} : \mathbb{R}^d \to [-1, +1]$ as the periodic extension of

$$v_j^{(n)}(x) := \sum_{s=1}^{k_j^{(n)}} \widetilde{w}_j^n \left(\frac{x - (x_j + y_j^{(n)})}{\varepsilon_n} \right) \chi_{Q^{(n)}_{j,s}}(x') . \tag{40}$$

We are now in position to define the function u_{ε_n} in A_i: for $x \in A_i$ define

$$u_{\varepsilon_n}(x) := \sum_{j=1}^{M_\xi} \chi_{Q_j^{(n)}}(x') v_j^{(n)}(x) .$$

Using (38) and (40) we have that $u_{\varepsilon_n} \to u$ as $n \to \infty$ uniformly on compact sets $K \subset X \setminus \partial^* \{u = 1\}$. We now prove the validity of inequality (34). We claim that:

(i) $\widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n}, A_i, A_j) \to 0$ as $n \to \infty$ for all $i \neq j$, where $\widetilde{\Lambda}$ is defined by (13);

(ii) $\limsup_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, A_i) \leq \mathcal{G}^{(p)}_\infty(u, A_i)$ for all $i = 0, \dots, k+1$.

If the above claims hold true, then we can conclude as follows: we have

$$\begin{aligned} \limsup_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}) &\leq \sum_{i=0}^{k+1} \limsup_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, A_i) + 2 \sum_{i<j=0}^{k+1} \limsup_{n\to\infty} \widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n}, A_i, A_j) \\ &\leq \sum_{i=0}^{k+1} \mathcal{G}^{(p)}_\infty(u, A_i) \\ &= \mathcal{G}^{(p)}_\infty(u) . \end{aligned}$$

We start by proving claim (ii). It is easy to see that it holds true for $i = 0, k+1$. Fix $i \in \{1, \dots, k\}$. Noticing that

$$A_i \subset \bigcup_{j=1}^{M} \bigcup_{s=1}^{k_j^{(n)}} T_{Q^{(n)}_{j,s}} ,$$

we get

$$
\begin{aligned}
\mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n}, A_i) &\leq \mathcal{F}^{(p)}_{\varepsilon_n}\left(u_{\varepsilon_n}, \bigcup_{j=1}^{M_\xi}\bigcup_{s=1}^{k_j^{(n)}} T_{Q_{j,s}^{(n)}}\right) \\
&\leq \sum_{j=1}^{M_\xi}\sum_{s=1}^{k_j^{(n)}}\Bigg[\frac{s_{\varepsilon_n}}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}\int_{\mathbb{R}^d}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\qquad\qquad + \frac{1}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}} V(u_{\varepsilon_n}(x))\rho(x)\,\mathrm{d}x\Bigg]. \qquad (41)
\end{aligned}
$$

Here, for every j and s we are using the whole cube $Q_{j,s}^{(n)}$, not only its intersection with Q_j.

Note that,

$$
\begin{aligned}
\mathcal{A}_{j,s}^{(n)} &:= \left\{(x,z)\in\mathbb{R}^d\times T_{Q_{j,s}^{(n)}} \,:\, \eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p \neq 0\right\} \\
&\subseteq B(x_j, \sqrt{2d}r_\xi)\times\left(T_{Q_{j,s}^{(n)}}\cap\{|z_d|\leq\varepsilon_n(L_j+R_\eta)\}\right)
\end{aligned}
$$

for ε_n sufficiently small (compared to r_ξ). Hence there exists $\delta_\xi>0$ such that for all $(x,z)\in\mathcal{A}_{j,s}^{(n)}$ we have $|\rho(x)-\rho(x_j)|\leq\delta_\xi$ and $|\rho(z)-\rho(x_j)|\leq\delta_\xi$ where $\delta_\xi\to 0$ as $\xi\to 0^+$. This implies

$$
\begin{aligned}
s_{\varepsilon_n}\rho(x)\rho(z) &\leq |s_{\varepsilon_n}-1|\rho(x)\rho(z)+\rho(x)\rho(z) \\
&\leq |s_{\varepsilon_n}-1|c_2^2+\rho(x_j)\rho(z)+\delta_\xi c_2 \\
&\leq |s_{\varepsilon_n}-1|c_2^2+\rho^2(x_j)+2\delta_\xi c_2\,.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\frac{s_{\varepsilon_n}}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}\int_{\mathbb{R}^d}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\qquad\leq\left(|s_{\varepsilon_n}-1|c_2^2+2\delta_\xi c_2+\rho^2(x_j)\right)\frac{1}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}\int_{\mathbb{R}^d}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\,\mathrm{d}x\,\mathrm{d}z\,.
\end{aligned}
$$

Similarly,

$$
\frac{1}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}V(u_{\varepsilon_n}(x))\rho(x)\,\mathrm{d}x\leq(\rho(x_j)+\delta_\xi)\,\frac{1}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}V(u_{\varepsilon_n}(x))\,\mathrm{d}x\,.
$$

Let $\gamma_{\xi,n}=1+\frac{2\delta_\xi}{c_2}+|s_{\varepsilon_n}-1|$ and notice that

$$
\lim_{\xi\to 0}\lim_{n\to\infty}\gamma_{\xi,n}=1
$$

(there is a subtle dependence that requires ε_n be small compared to r_ξ which means it is necessary to first take the limits in this order). Then

$$
\frac{s_{\varepsilon_n}}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}\int_{\mathbb{R}^d}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z+\frac{1}{\varepsilon_n}\int_{T_{Q_{j,s}^{(n)}}}V(u_{\varepsilon_n}(x))\rho(x)\,\mathrm{d}x
$$

$$\leq \gamma_{\xi,n}\rho(x_j)\left(\frac{\rho(x_j)}{\varepsilon_n}\int_{T_{Q^{(n)}_{j,s}}}\int_{\mathbb{R}^d}\eta_{\varepsilon_n}(x-z)|u_{\varepsilon_n}(x)-u_{\varepsilon_n}(z)|^p\,\mathrm{d}x\,\mathrm{d}z+\frac{1}{\varepsilon_n}\int_{T_{Q^{(n)}_{j,s}}}V(u_{\varepsilon_n}(x))\,\mathrm{d}x\right)$$
$$\leq \gamma_{\xi,n}\rho(x_j)G^{(p)}_{\varepsilon_n}(u_{\varepsilon_n},\rho(x_j),T_{Q^{(n)}_{j,s}})\,.$$

Hence,, using (35), (36), (37), (39) and (41), we get

$$\begin{aligned}
\mathcal{F}^{(p)}_{\varepsilon_n}(u_{\varepsilon_n},A_i) &\leq \gamma_{\xi,n}\sum_{j=1}^{M_\xi}\sum_{s=1}^{k^{(n)}_j}\rho(x_j)G^{(p)}_{\varepsilon_n}(u_{\varepsilon_n},\rho(x_j),T_{Q^{(n)}_{j,s}})\\
&= \gamma_{\xi,n}\sum_{j=1}^{M_\xi}\sum_{s=1}^{k^{(n)}_j}\rho(x_j)\varepsilon_n^{d-1}G^{(p)}_1(\widetilde{w}_j,\rho(x_j),T_{C_j})\\
&= \gamma_{\xi,n}\sum_{j=1}^{M_\xi}\frac{r_\xi^{d-1}\rho(x_j)}{\mathcal{H}^{d-1}(C_j)}G^{(p)}_1(\widetilde{w}_j,\rho(x_j),T_{C_j})\\
&\leq \gamma_{\xi,n}\left(r_\xi^{d-1}\sum_{j=1}^{M_\xi}\rho(x_j)\sigma^{(p)}(x_j,\nu)+2\xi\right)\\
&\leq \gamma_{\xi,n}\left(\int_{E_i}\sigma^{(p)}(x,\nu)\rho(x)\,\mathrm{d}\mathcal{H}^{d-1}(x)+3\xi\right)\,.
\end{aligned}$$

Taking limsup as $n\to\infty$ and $\xi\to 0$ implies (ii). Note that u_{ε_n} is not the recovery sequence, rather each u_{ε_n} depended on ξ (through L_j). Hence, if we make the ξ dependence explicit, we showed $\limsup_{n\to\infty}\mathcal{F}^{(p)}_{\varepsilon_n}(u^{(\xi)}_{\varepsilon_n},A_i)\leq\limsup_{n\to\infty}\gamma_{\xi,n}\mathcal{G}^{(p)}_\infty(u,A_i)$. By a diagonalisation argument we can find a sequence $\xi_n\to 0$ such that $\limsup_{n\to\infty}\mathcal{F}^{(p)}_{\varepsilon_n}(u^{(\xi_n)}_{\varepsilon_n},A_i)\leq\mathcal{G}^{(p)}_\infty(u,A_i)$.

We are thus left to prove claim (i). Analogously to Remark 3.12, $\widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n},A_i,A_j)\to 0$ as $n\to\infty$ for all i,j for which $\mathrm{d}(A_i,A_j)>0$. Let us now consider indexes $i\neq j$ for which $\mathrm{d}(A_i,A_j)=0$. Write

$$\begin{aligned}
\widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n},A_i,A_j) &= \frac{1}{\varepsilon_n}\int_{\mathbb{R}^d}\int_{A_{\varepsilon_n h}}\eta(h)|u_{\varepsilon_n}(z+\varepsilon_n h)-u_{\varepsilon_n}(z)|^p\rho(z+\varepsilon_n h)\rho(z)\,\mathrm{d}z\,\mathrm{d}h\\
&= \frac{1}{\varepsilon_n}\int_{B(0,R_\eta)}\int_{\widetilde{A}_{\varepsilon_n h}}\eta(h)|u_{\varepsilon_n}(z+\varepsilon_n h)-u_{\varepsilon_n}(z)|^p\rho(z+\varepsilon_n h)\rho(z)\,\mathrm{d}z\,\mathrm{d}h\,,
\end{aligned}\tag{42}$$

where $A_{\varepsilon_n h}:=\{z\in A_i\,:\,z+\varepsilon_n h\in A_j\}$ and,

$$\widetilde{A}_{\varepsilon_n h}:=\{z\in A_i\,:\,z+\varepsilon_n h\in A_j,|z_d|\leq\varepsilon_n\max\{\max_k L^{(i)}_k,\max L^{(j)}_k+R_\eta\}\,.\tag{43}$$

where we denote the dependence of A_i and A_j on the constants L_k from (38). Since $\mathrm{Vol}(\widetilde{A}_{\varepsilon_n h})=O(\varepsilon_n^2)$, from (42) and (43) we get

$$\widetilde{\Lambda}_{\varepsilon_n}(u_{\varepsilon_n},A_i,A_j)\to 0\,,$$

as $n\to\infty$.

Step 2. The general case. Let $u\in BV(X;\pm 1)$. Using Theorem 2.19 it is possible to find a sequence $\{v_n\}_{n=1}^\infty$ of polyhedral function such that $v_n\to u$ in L^1 (which, in turn, implies that $Dv_n\stackrel{w^*}{\rightharpoonup}Du$) and

$|Dv_n|(X) \to |Du|(X)$. Using Step 1 and a diagonalisation argument we get that there exists a sequence $\{u_n\}_{n=1}^\infty$ with $u_n \to u$ in $L^1(X)$ such that

$$\mathcal{F}_{\varepsilon_n}^{(p)}(u_n) \le \mathcal{G}_\infty^{(p)}(v_n) + \frac{1}{n}\,.$$

Then, Theorem 2.20 together with Lemma 3.9 gives us that

$$\limsup_{n\to\infty} \mathcal{G}_\infty^{(p)}(v_n) = \mathcal{G}_\infty^{(p)}(u)\,.$$

This concludes the proof. □

4 Convergence of the Graphical Model

In this Section we prove Theorem 1.6. In particular, in Section 4.1 we prove the compactness part of Theorem 1.6 and in Sections 4.2-4.3 we prove the Γ-convergence result.

4.1 Compactness

Proof of Theorem 1.6 (Compactness). Step 1. We first show that there exist $c_n, \alpha_n > 0$ with $\alpha_n, c_n \to 1$ as $n \to \infty$, such that

$$\eta\left(\frac{T_n(x) - T_n(z)}{\varepsilon_n}\right) \ge c_n \eta\left(\frac{\alpha_n(x-z)}{\varepsilon_n}\right)\,. \tag{44}$$

Let $\delta_n := \frac{2\|T_n - \mathrm{Id}\|_{L^\infty}}{\varepsilon_n}$. By Assumption (C4) we can find α_n, c_n such that, for all $a, b \in \mathbb{R}^d$ with $|a-b| \le \delta_n$, we have

$$\eta(a) \ge c_n \eta(\alpha_n b)\,. \tag{45}$$

Since by assumption (A2) we have that $\delta_n \to 0$ then α_n, c_n can be chosen such that $\alpha_n \to 1, c_n \to 1$. Now if we let $a := \frac{T_n(x)-T_n(z)}{\varepsilon_n}$ and $b := \frac{x-z}{\varepsilon_n}$ we have

$$|a-b| = \frac{|T_n(x) - T_n(z) + z - x|}{\varepsilon_n} \le \frac{2\|T_n - \mathrm{Id}\|_{L^\infty}}{\varepsilon_n} = \delta_n$$

and therefore, by (45), we get $\eta\left(\frac{T_n(x)-T_n(z)}{\varepsilon_n}\right) \ge c_n \eta\left(\frac{\alpha_n(x-z)}{\varepsilon_n}\right)$ as required.

Step 2. Let $v_n := u_n \circ T_n$. Using Lemma 2.8 and (44) we have that

$$\begin{aligned}
\mathcal{G}_n^{(p)}(u_n) &= \frac{1}{\varepsilon_n^{d+1}} \int_X \int_X \eta\left(\frac{T_n(x) - T_n(z)}{\varepsilon_n}\right) |v_n(x) - v_n(z)|^p \rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\quad + \frac{1}{\varepsilon_n} \int_X V(v_n(x))\rho(x)\,\mathrm{d}x \\
&\ge \frac{c_n}{\varepsilon_n^{d+1}} \int_X \int_X \eta\left(\frac{\alpha_n(x-z)}{\varepsilon_n}\right) |v_n(x) - v_n(z)|^p \rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z \\
&\quad + \frac{1}{\varepsilon_n} \int_X V(v_n(x))\rho(x)\,\mathrm{d}x \\
&= \frac{c_n}{\alpha_n^{d+1}(\varepsilon_n')^{d+1}} \int_X \int_X \eta\left(\frac{x-z}{\varepsilon_n'}\right) |v_n(x) - v_n(z)|^p \rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{\varepsilon_n}\int_X V(v_n(x))\rho(x)\,\mathrm{d}x \\
&= \frac{\varepsilon_n'}{\varepsilon_n}\mathcal{F}^{(p)}_{\varepsilon_n'}(v_n)
\end{aligned} \tag{46}
$$

where $\varepsilon_n' := \frac{\varepsilon_n}{\alpha_n}$ and $\mathcal{F}^{(p)}_{\varepsilon}$ is defined by (3) with $s_n := \frac{c_n\varepsilon_n}{\alpha_n^{d+1}\varepsilon_n'}$ (in (3) s depended on ε not on n, since we have fixed the sequence ε_n then clearly we could write n in terms of ε_n', however this would make the notation cumbersome).

Step 3. Since $\lim_{n\to\infty}\frac{\varepsilon_n'}{\varepsilon_n} = 1$, we can infer that $\mathcal{F}^{(p)}_{\varepsilon_n'}(v_n)$ is bounded and hence by Theorem 3.2 the sequence $\{v_n\}_{n=1}^\infty$ is relatively compact in L^1. Therefore the sequence $\{u_n\}_{n=1}^\infty$ is relatively compact in TL^1, with any limit u satisfying $\mathcal{G}^{(p)}_\infty(u) < \infty$. □

4.2 The Liminf Inequality

Proof of Theorem 1.6 (Liminf). For any $u \in L^1(\mu)$ and any $u_n \in L^1(\mu_n)$ with $u_n \to u$ in TL^1 we claim that

$$
\liminf_{n\to\infty} \mathcal{G}^{(p)}_n(u_n) \geq \mathcal{G}^{(p)}_\infty(u).
$$

Indeed, taking the liminf on both sides of (46) and using Theorem 3.2 we have

$$
\liminf_{n\to\infty} \mathcal{G}^{(p)}_n(u_n) \geq \liminf_{n\to\infty} \frac{\varepsilon_n'}{\varepsilon_n}\mathcal{F}^{(p)}_{\varepsilon_n'}(v_n) \geq \mathcal{G}^{(p)}_\infty(u)
$$

since $\lim_{n\to\infty}\frac{\varepsilon_n'}{\varepsilon_n} = 1$. □

4.3 The Limsup Inequality

Proof of Theorem 1.6 (Limsup). The aim of this section is to prove the following: given $u \in L^1(X,\mu)$ it is possible to find a sequence $\{u_n\}_{n=1}^\infty \subset L^1(X_n)$ with $u_n \to u$ in $TL^1(X)$ such that

$$
\limsup_{n\to\infty} \mathcal{G}^{(p)}_n(u_n) \leq \mathcal{G}^{(p)}_\infty(u)\,.
$$

Without loss of generality we can assume $\mathcal{G}^{(p)}_\infty(u) < +\infty$. In particular, $u \in BV(X;\pm 1)$.

We divide the proof in two cases: we first assume that u is a polyhedral function and then we extend the argument to any function u with $\mathcal{G}^{(p)}_\infty(u) < +\infty$ via a diagonalisation argument.

Case 1. Assume that u is a polyhedral function (see Definition 2.18) and consider the sequence $u_n := u \llcorner X_n$.

Step 1. We show that $u_n \to u$ in TL^1. Setting $v_n = u_n \circ T_n$, we are required to show $v_n \to u$ in L^1. Assume $u = \chi_A - \chi_{A^c}$ where A is a polyhedral set. Set

$$
A_n := \{x \in X \,:\, \mathrm{dist}(x,\partial A) \leq \|T_n - \mathrm{Id}\|_{L^\infty}\}\,. \tag{47}
$$

Then

$$
\|u_n \circ T_n - u\|_{L^1(X)} = \int_X |u_n(T_n(x)) - u(x)|\,\mathrm{d}x
$$

$$
\begin{aligned}
&= \int_{A_n} |u_n(T_n(x)) - u(x)| \,\mathrm{d}x \\
&\leq 2\mathrm{Vol}\,(A_n) \\
&= O(\|T_n - \mathrm{Id}\|_{L^\infty}).
\end{aligned}
$$

It follows that $u_n \to u$ in TL^1.

Step 2. We show that there exists $\hat{\alpha}_n, \hat{c}_n \to 1$ such that

$$
\eta\left(\frac{T_n(x) - T_n(z)}{\varepsilon_n}\right) \leq \hat{c}_n \eta\left(\frac{\hat{\alpha}_n(x-z)}{\varepsilon_n}\right). \tag{48}
$$

To show the validity of (48) we use the following subclaim: for all $\hat{\delta} > 0$ sufficiently small there exists $\hat{\alpha}_{\hat{\delta}}, \hat{c}_{\hat{\delta}} > 0$ such that $\hat{\alpha}_{\hat{\delta}} \to 1, \hat{c}_{\hat{\delta}} \to 1$, as $\hat{\delta} \to 0$, and, for any $\hat{a}, \hat{b} \in \mathbb{R}^d$, it holds

$$
|\hat{a} - \hat{b}| < \hat{\delta} \quad \Rightarrow \quad \hat{c}_{\hat{\delta}} \eta(\hat{\alpha}_{\hat{\delta}} \hat{b}) \geq \eta(\hat{a}). \tag{49}
$$

Then (48) can be obtained as follows: for any $n \in \mathbb{N}$ take

$$
\hat{a} := \frac{T_n(x) - T_n(z)}{\varepsilon_n}, \qquad \hat{b} := \frac{x-z}{\varepsilon_n}, \qquad \hat{\delta} := \frac{2\|T_n - \mathrm{Id}\|_{L^\infty}}{\varepsilon_n},
$$

and let $\hat{\alpha}_n, \hat{c}_n$ be the numbers given by the subclaim for which (49) holds. Then, since $|\hat{a} - \hat{b}| \leq \delta_n$, we infer (48).

To prove the subclaim, we let α_δ, c_δ be as in Assumption (C4) and $\hat{\delta} > 0$ and choose $\delta := \min\left\{1, \frac{\hat{\delta}}{\inf_{\delta \in (0,1]} \alpha_\delta}\right\}$, without loss of generality we assume that $\inf_{\delta \in (0,1]} \alpha_\delta \in (0,\infty)$ and trivially $\delta \to 0$ as $\hat{\delta} \to 0$. We assume that $\frac{\hat{\delta}}{\inf_{\delta \in (0,1]} \alpha_\delta} \leq 1$. Let $\hat{a}, \hat{b} \in \mathbb{R}^d$ with $|\hat{a} - \hat{b}| < \hat{\delta}$, and define $a := \frac{\hat{a}}{\alpha_\delta}$ and $b := \frac{\hat{b}}{\alpha_\delta}$. Since, $|a - b| \leq \frac{\hat{\delta}}{\alpha_\delta} \leq \frac{\hat{\delta}}{\inf_{\delta \in (0,1]} \alpha_\delta} = \delta$ then

$$
\eta(b) \geq c_\delta \eta(\alpha_\delta a) \quad \Rightarrow \quad \frac{1}{c_\delta} \eta\left(\frac{\hat{b}}{\alpha_\delta}\right) \geq \eta(\hat{a}).
$$

Let $\hat{c}_{\hat{\delta}} := \frac{1}{c_\delta}$, $\hat{\alpha}_{\hat{\delta}} := 1/\alpha_\delta$ then $\hat{\delta} \to 0$ implies $\delta \to 0$ which in turn implies $\alpha_\delta, c_\delta \to 1$ and therefore $\hat{\alpha}_{\hat{\delta}}, \hat{c}_{\hat{\delta}} \to 1$. This proves the claim.

Step 3. Since $\frac{1}{\varepsilon_n} \int V(u_n(x))\rho(x) \,\mathrm{d}x = 0$ then

$$
\mathcal{G}_n^{(p)}(u_n) = \frac{1}{\varepsilon_n n^2} \sum_{i,j=1}^n \eta_{\varepsilon_n}(x_i - x_j) |u_n(x_i) - u_n(x_j)|^p .
$$

Then, using Lemma 2.8 and (48) we get

$$
\begin{aligned}
\mathcal{G}_n^{(p)}(u_n) &= \frac{1}{\varepsilon_n} \int_X \int_X \eta_{\varepsilon_n}(T_n(x) - T_n(z)) |v_n(x) - v_n(z)|^p \rho(x)\rho(z) \,\mathrm{d}x \,\mathrm{d}z \\
&\leq \frac{\hat{c}_n}{\hat{\alpha}_n^{d+1} \hat{\varepsilon}_n'} \int_X \int_X \eta_{\hat{\varepsilon}_n'}(x - z) |v_n(x) - v_n(z)|^p \rho(x)\rho(z) \,\mathrm{d}x \,\mathrm{d}z \\
&= \frac{\hat{c}_n}{\hat{\alpha}_n^{d+1} \hat{\varepsilon}_n'} \int_X \int_X \eta_{\hat{\varepsilon}_n'}(x - z) |u(x) - u(z)|^p \rho(x)\rho(z) \,\mathrm{d}x \,\mathrm{d}z + a_n
\end{aligned}
$$

where $\varepsilon_n' := \frac{\varepsilon_n}{\hat{\alpha}_n}$ and

$$a_n := \frac{\hat{c}_n}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\int_X\int_X \eta_{\hat{\varepsilon}_n'}(x-z)\left(|v_n(x)-v_n(z)|^p - |u(x)-u(z)|^p\right)\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z\,.$$

We recall the followings inequalities: $\forall\delta>0$ there exits $C_\delta>0$ such that for any $a,b\in\mathbb{R}^d$ we have

$$|a|^p \le (1+\delta)|b|^p + C_\delta|a-b|^p\,. \tag{50}$$

Moreover, for all $p\ge 1$ and all $a,b\in\mathbb{R}$, it holds

$$|a+b|^p \le 2^{p-1}\left(|a|^p+|b|^p\right)\,. \tag{51}$$

Fix $\delta>0$. Using (50) and (51) we infer

$$a_n \le \frac{\hat{c}_n\delta}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\int_X\int_X \eta_{\hat{\varepsilon}_n'}(x-z)|u(x)-u(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z + b_n$$

where

$$b_n := \frac{2^{p-1}\hat{c}_nC_\delta}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\int_X\int_X \eta_{\hat{\varepsilon}_n'}(x-z)\Big(|v_n(x)-u(x)|^p + |v_n(z)-u(z)|^p\Big)\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z\,.$$

We show $b_n\to 0$. We have that

$$\begin{aligned} b_n &\le \frac{2^p\hat{c}_nC_\delta c_2^2}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\int_X |v_n(x)-u(x)|^p\,\mathrm{d}x\int_{\mathbb{R}^d}\eta(x)\,\mathrm{d}x \\ &\le \frac{2^{2p}\hat{c}_nC_\delta c_2^2}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\mathrm{Vol}\left(A_n\right)\int_{\mathbb{R}^d}\eta(x)\,\mathrm{d}x \\ &= O\left(\frac{\|T_n-\mathrm{Id}\|_{L^\infty}}{\hat{\varepsilon}_n'}\right)\,, \end{aligned}$$

where A_n is defined as in (47) and in the last equality, we used the fact that $\hat{c}_n\to 1$, $\hat{\alpha}_n\to 1$ as $n\to\infty$. Since $\hat{\varepsilon}_n'\asymp\varepsilon_n$, from hypothesis (A2) we get that

$$b_n\to 0\,, \tag{52}$$

as $n\to\infty$. Hence,

$$\mathcal{G}_n^{(p)}(u_n) \le \frac{\hat{c}_n(1+\delta)}{\hat{\alpha}_n^{d+1}\hat{\varepsilon}_n'}\int_X\int_X \eta_{\hat{\varepsilon}_n'}(x-z)|u(x)-u(z)|^p\rho(x)\rho(z)\,\mathrm{d}x\,\mathrm{d}z + b_n\,. \tag{53}$$

Step 4. We now conclude as follows. Using (53) we get

$$\mathcal{G}_n^{(p)}(u_n) \le (1+\delta)\mathcal{F}_{\hat{\varepsilon}_n'}(u) + b_n$$

for $\mathcal{F}_{\hat{\varepsilon}_n'}$ defined as in (3) with $s_n := \frac{\hat{c}_n}{\hat{\alpha}_n^{d+1}}$. Taking the limsup on both sides and using (52), together with Theorem 3.2 we have

$$\limsup_{n\to\infty}\mathcal{G}_n^{(p)}(u_n) \le \limsup_{n\to\infty}(1+\delta)\mathcal{F}_{\hat{\varepsilon}_n'}(u) \le (1+\delta)\mathcal{G}_\infty^{(p)}(u).$$

Taking $\delta \to 0$ completes the proof for case 1.

Case 2. To extend to arbitrary functions $u \in BV(X;\pm 1)$ we apply the following diagonalisation argument. Using Theorem 2.19 together with Lemma 3.9 and Theorem 2.20 we can find a sequence of polyhedral functions $\{u^{(m)}\}_{m=1}^\infty$ such that

$$\|u - u^{(m)}\|_{L^1} \leq \frac{1}{m}, \qquad \mathcal{G}_\infty^{(p)}(u^{(m)}) \leq \mathcal{G}_\infty^{(p)}(u) + \frac{1}{m}.$$

Using the result of Case 1, for each $m \in \mathbb{N}$ we have

$$\limsup_{n\to\infty} \mathcal{G}_n^{(p)}(u_n^{(m)}) \leq \mathcal{G}_\infty^{(p)}(u^{(m)}),$$

where $u_n^{(m)} := u^{(m)} \llcorner X_n$. For each $m \in \mathbb{N}$ let $n_m \in \mathbb{N}$ be such that

$$\mathcal{G}_n^{(p)}(u_n^{(m)}) \leq \mathcal{G}_\infty^{(p)}(u^{(m)}) + \frac{1}{m} \quad \text{and} \quad \|u_n^{(m)} \circ T_n - u^{(m)}\|_{L^1} \leq \frac{1}{m}$$

for all $n \geq n_m$. At the cost of increasing n_m we assume that $n_{m+1} > n_m$ for all m. Let $u_n := u_n^{(m)}$ for $n \in [n_m, n_{m+1})$. Then,

$$\limsup_{n\to\infty} \mathcal{G}_n^{(p)}(u_n) = \limsup_{m\to\infty} \sup_{n\in[n_m,n_{m+1})} \mathcal{G}_n^{(p)}(u_n^{(m)}) \leq \limsup_{m\to\infty} \left(\mathcal{G}_\infty^{(p)}(u^{(m)}) + \frac{1}{m} \right) \leq \mathcal{G}_\infty^{(p)}(u).$$

Similarly,

$$\lim_{n\to\infty} \|u_n \circ T_n - u\|_{L^1} \leq \limsup_{m\to\infty} \sup_{n\in[n_m,n_{m+1})} \left(\|u_n^{(m)} \circ T_n - u^{(m)}\|_{L^1} + \frac{1}{m} \right) \leq \lim_{m\to\infty} \frac{2}{m} = 0$$

therefore u_n converges to u in TL^1. Hence u_n is a recovery sequence for u. This completes the proof. □

5 Convergence of Minimizers with Data Fidelity

In this section we prove Corollary 1.9.

Proof of Corollary 1.9. In view of Theorem 2.22 and Proposition 2.24 it is enough to prove that, for any $u \in L^1(X,\mu)$ and any $\{u_n\}_{n=1}^\infty$ with $u_n \in L^1(X_n)$ such that $u_n \to u$ in $TL^1(X)$ it holds that

$$\lim_{n\to\infty} \mathcal{K}_n(u_n) = \mathcal{K}_\infty(u).$$

We can restrict ourselves to sequences $u_n \to u$ satisfying

$$\sup_{n\in\mathbb{N}} \mathcal{G}_n^{(p)}(u_n) < +\infty. \tag{54}$$

Let

$$v_n(x) := \begin{cases} u_n(x) & \text{if } x \in X_n(u_n), \\ 1 & \text{otherwise}. \end{cases}$$

where

$$X_n(u_n) := \{x \in X_n \,:\, |u_n(x)|^q \leq R_V\},$$

and R_V is as in Assumption (B3).

Step 1. We claim that

$$\lim_{n\to\infty} |\mathcal{K}_n(u_n) - \mathcal{K}_n(v_n)| \to 0\,, \tag{55}$$

as $n\to\infty$. Indeed we have that

$$\begin{aligned}\lim_{n\to\infty} |\mathcal{K}_n(u_n) - \mathcal{K}_n(v_n)| &= \lim_{n\to\infty} \left| \frac{1}{n} \sum_{x_i\notin X_n(u_n)} \left(k_n(x_i,u_n(x_i)) - k_n(x_i,1)\right) \right| \\ &= \beta \lim_{n\to\infty} \frac{1}{n} \sum_{x_i\notin X_n(u_n)} \left(3+|u_n(x_i)|^q\right).\end{aligned}$$

Now,

$$\frac{1}{n}\sum_{x_i\notin X_n(u_n)} 1 \leq \frac{1}{nR_V}\sum_{x_i\notin X_n(u_n)} |u_n(x_i)|^q \leq \frac{1}{nR_V\tau}\sum_{x_i\notin X_n(u_n)} V(u_n(x_i)) \leq \frac{\varepsilon_n \mathcal{G}^{(p)}_n(u_n)}{\tau R_V}.$$

Using (54) we conclude (55).

Step 2. We claim that $v_n\to u$ in $TL^1(X)$. By a direct computation we get

$$\begin{aligned}\|u_n - v_n\|_{L^1(\mu_n)} &= \frac{1}{n}\sum_{x_i\notin X_n(u_n)} |u_n(x_i)-1| \\ &\leq \frac{1}{n}\sum_{x_i\notin X_n(u_n)} (1+|u_n(x_i)|) \\ &\leq \frac{1}{n}\sum_{x_i\notin X_n(u_n)} (1+|u_n(x_i)|^q) \to 0,\end{aligned}$$

so $v_n\to u$ in TL^1.

Step 3. We show that $\mathcal{K}_n(v_n)\to\mathcal{K}_\infty(u)$. Consider any subsequence of v_n which we do not relabel. From Step 2 we have that $v_n\circ T_n\to u$ in L^1. Thus there exists a further subsequence (not relabeled) such that $v_n(T_n(x))\to u(x)$ for almost every $x\in X$. Using Assumption (D3) we get

$$\lim_{n\to\infty} k_n\left(T_n(x), v_n(T_n(x))\right) = k_\infty(x,u(x))\,. \tag{56}$$

Moreover it holds

$$k_n(T_n(x), v_n(T_n(x)) \leq \beta\left(1+|v_n(T_n(x))|^q\right) \leq \beta\left(1+R_V\right) \tag{57}$$

Using (56), (57) and applying the Lebesgue's dominated convergence theorem we get

$$\begin{aligned}\lim_{n\to\infty}\mathcal{K}_n(v_n) &= \lim_{n\to\infty}\int_X k_n(T_n(x), v_n(T_n(x))\,\rho(x)\,\mathrm{d}x \\ &= \int_X k_\infty(x,u(x))\rho(x)\,\mathrm{d}x \\ &= \mathcal{K}_\infty(u)\end{aligned} \tag{58}$$

Since any subsequence of v_n has a further subsequence such that $\lim_{n\to\infty} \mathcal{K}_n(v_n) = \mathcal{K}_\infty(u)$ then we can conclude the convergence is over the full sequence.

Step 4. Using (55) and (58) we conclude that

$$\lim_{n\to\infty} \mathcal{K}_n(u_n) \to \mathcal{K}_\infty(u)$$

as required $\square$

Acknowledgements

The authors thank the Center for Nonlinear Analysis at Carnegie Mellon University and the Cantab Capital Institute for the Mathematics of Information at the University of Cambridge for their support during the preparation of the manuscript. The research of R.C. was funded by National Science Foundation under Grant No. DMS-1411646. Part of the research of M.T. was funded by the National Science Foundation under Grant No. CCT-1421502.

References

[1] G. ALBERTI AND G. BELLETTINI, *A nonlocal anisotropic model for phase transitions: Asymptotic behaviour of rescaled energies*, European Journal of Applied Mathematics, 9 (1998).

[2] ——, *A nonlocal anisotropic model for phase transitions I. The optimal profile problem*, Mathematische Annalen, 310 (1998), pp. 527–560.

[3] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[4] S. ARORA, S. RAO, AND U. VAZIRANI, *Expander flows, geometric embeddings and graph partitioning*, Journal of the ACM, 56 (2009), pp. Art. 5, 37.

[5] S. BALDO, *Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids*, Annales de l'Institut Henri Poincaré. Analyse Non Linéaire, 7 (1990), pp. 67–90.

[6] A. C. BARROSO AND I. FONSECA, *Anisotropic singular perturbations the vectorial case*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 124 (1994), pp. 527–571.

[7] A. L. BERTOZZI AND A. FLENNER, *Diffuse interface models on graphs for classification of high dimensional data*, Multiscale Modeling & Simulation, 10 (2012), pp. 1090–1118.

[8] G. BOUCHITTE, *Singular perturbations of variational problems arising from a two-phase transition model*, Applied Mathematics and Optimization, 21 (1990), pp. 289–314.

[9] J. BOURGAIN, H. BREZIS, AND P. MIRONESCU, *Another look at Sobolev spaces*, in Optimal Control and Partial Differential Equations, IOS, Amsterdam, 2001, pp. 439–455.

[10] Y. BOYKOV, O. VEKSLER, AND R. ZABIH, *Fast approximate energy minimization via graph cuts*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23 (2001), pp. 1222–1239.

[11] A. BRAIDES, *Γ-Convergence for Beginners*, Oxford University Press, 2002.

[12] A. BRAIDES AND N. K. YIP, *A quantitative description of mesh dependence for the discretization of singularly perturbed nonconvex problems*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 1883–1898.

[13] X. BRESSON AND T. LAURENT, *Asymmetric Cheeger cut and application to multi-class unsupre-vised clustering*, CAM report, (2012).

[14] X. BRESSON, T. LAURENT, D. UMINSKY, AND J. H. VON BRECHT, *Convergence and energy landscape for cheeger cut clustering*, in Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS), 2012, pp. 1385–1393.

[15] ——, *Multiclass total variation clustering*, in Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS), 2013, pp. 1421–1429.

[16] X. BRESSON, X.-C. TAI, T. F. CHAN, AND A. SZLAM, *Multi-class transductive learning based on l^1 relaxations of cheeger cut and mumford-shah-potts model*, Journal of Mathematical Imaging and Vision, 49 (2014), pp. 191–201.

[17] L. CALATRONI, Y. VAN GENNIP, C.-B. SCHÖNLIEB, H. M. ROWLAND, AND A. FLENNER, *Graph clustering, variational image segmentation methods and Hough transform scale detection for object measurement in images*, Journal of Mathematical Imaging and Vision, 57 (2017), pp. 269–291.

[18] A. CHAMBOLLE, A. GIACOMINI, AND L. LUSSARDI, *Continuous limits of discrete perimeters*, M2AN Mathematical Modelling and Numerical Analysis, 44 (2010), pp. 207–230.

[19] G. DAL MASO, *An Introduction to Γ-Convergence*, Springer, 1993.

[20] E. DAVIS AND S. SETHURAMAN, *Consistency of modularity clustering on random geometric graphs*, arXiv preprint arXiv:1604.03993, (2016).

[21] E. DE GIORGI AND T. FRANZONI, *Su un tipo di convergenza variazionale*, Atti della Accademia Nazionale dei Lincei. Rendiconti. Classe di Scienze Fisiche, Matematiche e Naturali, 58 (1975), pp. 842–850.

[22] R. M. DUDLEY, *Real Analysis and Probability*, Cambridge University Press, 2002.

[23] M. M. DUNLOP, D. SLEPČEV, A. M. STUART, AND M. THORPE, *Large data and zero noise limits of graph-based semi-supervised learning algorithms*, In Preperation, (2018).

[24] S. ESEDOḠLU AND F. OTTO, *Threshold dynamics for networks with arbitrary surface tensions*, Communications on Pure and Applied Mathematics, 68 (2015), pp. 808–864.

[25] I. FONSECA AND S. MÜLLER, *Relaxation of quasiconvex functionals in* $\mathrm{BV}(\Omega, \mathbf{R}^p)$ *for integrands* $f(x, u, \nabla u)$, Archive for Rational Mechanics and Analysis, 123 (1993), pp. 1–49.

[26] I. FONSECA AND C. POPOVICI, *Coupled singular perturbations for phase transitions*, Asymptotic Analysis, 44 (2005), pp. 299–325.

[27] I. Fonseca and L. Tartar, *The gradient theory of phase transitions for systems with two potential wells*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 111 (1989), pp. 89–102.

[28] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Mathematica, 177 (1996), pp. 113–161.

[29] C. Garcia-Cardona, E. Merkurjev, A. L. Bertozzi, A. Flenner, and A. G. Percus, *Multiclass data segmentation using diffuse interface methods on graphs*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 36 (2014), pp. 1600–1613.

[30] N. García-Trillos, *Variational limits of k-nn graph based functionals on data clouds*, arXiv preprint arXiv:1607.00696, (2016).

[31] N. García Trillos and R. Murray, *A new analytical approach to consistency and overfitting in regularized empirical risk minimization*, European Journal of Applied Mathematics, 28 (2017), pp. 886–921.

[32] N. García Trillos and D. Slepčev, *On the rate of convergence of empirical measures in ∞-transportation distance*, Canadian Journal of Mathematics, 67 (2015), pp. 1358–1383.

[33] ——, *Continuum limit of total variation on point clouds*, Archive for Rational Mechanics and Analysis, 220 (2016), pp. 193–241.

[34] ——, *A variational approach to the consistency of spectral clustering*, Applied and Computational Harmonic Analysis, (2016).

[35] N. García Trillos, D. Slepčev, J. von Brecht, T. Laurent, and X. Bresson, *Consistency of cheeger and ratio graph cuts*, The Journal of Machine Learning Research, 17 (2016), pp. 6268–6313.

[36] R. Kohn and P. Sternberg, *Local minimizers and singular perturbations*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 111 (1989), pp. 69–84.

[37] E. Merkurjev, T. Kostić, and A. L. Bertozzi, *An MBO scheme on graphs for classification and image processing*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 1903–1930.

[38] L. Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Archive for Rational Mechanics and Analysis, 98 (1987).

[39] L. Modica and S. Mortola, *Un esempio di Γ-convergenza*, Bollettino dell'Unione Matematica Italiana B, 14 (1977), pp. 285–299.

[40] A. Y. Ng, M. I. Jordan, and Y. Weiss, *On spectral clustering: Analysis and an algorithm*, in Proceedings of the 15th International Conference on Neural Information Processing Systems (NIPS), 2001, pp. 849–856.

[41] N. C. Owen and P. Sternberg, *Nonconvex variational problems with anisotropic perturbations*, Nonlinear Analysis, 16 (1991), pp. 705–719.

[42] M. Penrose, *Random geometric graphs*, vol. 5 of Oxford Studies in Probability, Oxford University Press, Oxford, 2003.

[43] A. C. Ponce, *A new approach to Sobolev spaces and connections to Γ-convergence*, Calculus of Variations and Partial Differential Equations, 19 (2004), pp. 229–255.

[44] O. Savin and E. Valdinoci, *Γ-convergence for nonlocal phase transitions*, Annales de l'Institut Henri Poincaré. Analyse Non Linéaire, 29 (2012), pp. 479–500.

[45] J. Shi and J. Malik, *Normalized cuts and image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (2000), pp. 888–905.

[46] D. Slepčev and M. Thorpe, *Analysis of p-Laplacian regularization in semi-supervised learning*, arXiv preprint arXiv:1707.06213, (2017).

[47] P. Sternberg, *The effect of a singular perturbation on nonconvex variational problems*, Archive for Rational Mechanics and Analysis, 101 (1988), pp. 209–260.

[48] A. Szlam and X. Bresson, *A total variation-based graph clustering algorithm for cheeger ratio cuts*, UCLA CAM Report, (2009), pp. 1–12.

[49] ——, *Total variation and cheeger cuts*, in Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML), 2010, pp. 1039–1046.

[50] M. Thorpe, S. Park, S. Kolouri, G. K. Rohde, and D. Slepčev, *A transportation L^p distance for signal analysis*, to appear in the Journal of Mathematical Imaging and Vision, arXiv preprint arXiv:1609.08669, (2017).

[51] M. Thorpe and D. Slepčev, *Transportation L^p distances: Properties and extensions*, In preparation, (2017).

[52] M. Thorpe and F. Theil, *Asymptotic analysis of the Ginzburg-Landau functional on point clouds*, to appear in the Proceedings of the Royal Society of Edinburgh Section A: Mathematics, arXiv preprint arXiv:1604.04930, (2017).

[53] Y. van Gennip and A. L. Bertozzi, *Γ-convergence of graph Ginzburg-Landau functionals*, Advances in Differential Equations, 17 (2012), pp. 1115–1180.

[54] Y. van Gennip and C.-B. Schönlieb, *Introduction: big data and partial differential equations*, European Journal of Applied Mathematics, 28 (2017), pp. 877–885.

[55] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.