

# Discrete stochastic approximations of the Mumford–Shah functional

Matthias Ruf

*Département Mathématique, Université libre de Bruxelles (ULB), CP 214, boulevard du Triomphe, 1050 Brussels, Belgium*

## 1. Introduction

The Mumford–Shah functional has its origin in image segmentation problems [50]. Given a rectangle (or more generally a bounded domain  $D \subset \mathbb{R}^2$ ) and a function  $g : D \rightarrow \mathbb{R}$  representing the gray level of an image, one aims at minimizing the functional

$$MS(u, K) = \int_{D \setminus K} |\nabla u|^2 \, dx + \beta \mathcal{H}^1(K) + \gamma \int_D |u - g|^2 \, dx.$$

---

*E-mail address:* [matthias.ruf@ulb.ac.be](mailto:matthias.ruf@ulb.ac.be).

Here  $K$  is the union of a finite number of points and a finite set of smooth arcs joining these points with no other intersections. The function  $u$  is supposed to be differentiable on  $D \setminus K$ , but may have discontinuities on  $K$ . Then the pair  $(u, K)$  is an approximation of the image.  $K$  represents the sharp edges in the image, while the smooth part  $u$  yields a cartoon-like total image since it rules out fine textures away from  $K$ . Existence and regularity of minimizing pairs  $(u, K)$  is far from being trivial. In [38] it was shown that there exists a minimizing pair  $(u, K)$  among all closed sets  $K$  and  $u \in C^1(D \setminus K)$  provided  $g \in L^\infty(D)$ . As commonly done in variational problems, one first has to enlarge the set of competitors in order to obtain compactness of minimizing sequences. This leads to the nowadays well-known formulation of the Mumford–Shah functional for  $SBV$ -functions, which was first introduced in [10]: given  $u \in SBV(D)$ , the Mumford–Shah functional takes the form

$$MS(u) = \int_D |\nabla u|^2 dx + \beta \mathcal{H}^1(S_u) + \gamma \int_D |u - g|^2 dx. \quad (1)$$

Here  $S_u$  denotes the discontinuity set of  $u$ . The closed set  $K$  then can be recovered setting  $K = \overline{S_u}$  since minimizers have an essentially closed discontinuity set (see [38, Lemma 5.2]). However it is still unknown if  $K$  can be taken as a finite union of regular arcs. We refer the interested reader to the recent survey articles [39,49] for known regularity results for minimizers.

Besides the regularity of minimizers, there is the natural question how to minimize the Mumford–Shah functional (1) in practice. A very popular approach is given by the Ambrosio–Tortorelli approximation [12,13], where the surface term is replaced by a Modica–Mortola-type approximation with an additional variable. More precisely, given a small parameter  $\varepsilon > 0$  and  $0 < \eta_\varepsilon \ll \varepsilon$  one defines an elliptic approximation  $AT_\varepsilon : W^{1,2}(D) \times W^{1,2}(D) \rightarrow [0, +\infty]$  by

$$AT_\varepsilon(u, v) = \int_D (\eta_\varepsilon + v^2) |\nabla u|^2 dx + \frac{\beta}{2} \int_D \varepsilon |\nabla v|^2 + \frac{1}{\varepsilon} (v - 1)^2 dx + \gamma \int_D |u - g|^2 dx.$$

In [13] it is shown that the family  $AT_\varepsilon$  approximates the Mumford–Shah functional (1) in the sense of  $\Gamma$ -convergence (we refer to the monographs [21,35] for details on this type of convergence). In particular, up to subsequences, the  $u$ -component of any global minimizer  $(u_\varepsilon, v_\varepsilon)$  of  $AT_\varepsilon$  converges to a global minimizer of  $MS$ . This approach was recently extended to second order penalizations, that means to replace the term  $\varepsilon |\nabla v|^2$  either by  $\varepsilon^3 |\nabla^2 v|^2$  or  $\varepsilon^3 (\Delta v)^2$ , where in the second case one puts additional boundary conditions on  $v$  (see [27] for more details or [14] for an anisotropic version).

Instead of introducing a second variable, Braides and Dal Maso constructed non-local approximations. In [23] they showed that the sequence of functionals  $NL_\varepsilon : W^{1,2}(D) \rightarrow [0, +\infty]$  defined by

$$NL_\varepsilon(u) = \frac{1}{\varepsilon} \int_D f\left(\varepsilon \int_{B_\varepsilon(x) \cap D} |\nabla u(y)|^2 dy\right) dx + \gamma \int_D |u - g|^2 dx$$

$\Gamma$ -converges to  $MS$  provided  $f$  is continuous, increasing and satisfies

$$\lim_{t \rightarrow 0} \frac{f(t)}{t} = 1, \quad \lim_{t \rightarrow +\infty} f(t) = f_\infty < \infty. \quad (2)$$

In this case it turns out that  $\beta = 2f_\infty$ .

Note that both approximations are defined on more regular, but still infinite-dimensional spaces. Hence one has to discretize these spaces to numerically solve the minimization problems for the approximating functionals. On the one hand,  $\varepsilon$  should be taken very small in order to obtain almost sharp interfaces. On the other hand, to guarantee that finite elements/differences yield the same asymptotic behavior as the continuum approximations, it is proposed to take the mesh-size to be infinitesimal with respect to  $\varepsilon$  (see [16,19]). Indeed, for the Modica–Mortola approximation of the perimeter, such a choice is known to be necessary to preserve isotropy [26]. Very recently, in [15] this result was extended to finite difference discretizations of the Ambrosio–Tortorelli functional. Thus continuum approximations require in general a very fine mesh, which increases the computational effort. However, in dimension one there exists a direct approximation based on finite differences. In this case the small parameter  $\varepsilon$  represents the mesh-size of a one-dimensional grid  $\varepsilon\mathbb{Z}$ . Given a function  $u : \varepsilon\mathbb{Z} \cap (0, 1) \rightarrow \mathbb{R}$ , we define the functional  $F_\varepsilon$  by

$$F_\varepsilon(u) = \sum_{i=1}^{\lceil 1/\varepsilon \rceil - 2} \min \left\{ \varepsilon \left| \frac{u(\varepsilon(i+1)) - u(\varepsilon i)}{\varepsilon} \right|^2, \beta \right\} + \gamma \sum_{i=1}^{\lceil 1/\varepsilon \rceil - 1} \varepsilon |u(\varepsilon i) - g_\varepsilon(\varepsilon i)|^2,$$

where  $g_\varepsilon$  is a suitable discretized version of  $g \in L^\infty$ . The proof of convergence to the one-dimensional version of the Mumford–Shah functional can be found for example in [21, Chapter 8.3]. The functional  $F_\varepsilon$  above has a natural extension to higher dimensions. Indeed, given  $u : \varepsilon \mathbb{Z}^d \cap D \rightarrow \mathbb{R}$ , one sets

$$F_\varepsilon(u) = \frac{1}{2} \sum_{\substack{\varepsilon i, \varepsilon j \in \varepsilon \mathbb{Z}^d \cap D \\ |i-j|=1}} \varepsilon^{d-1} \min \left\{ \varepsilon \left| \frac{u(\varepsilon i) - u(\varepsilon j)}{\varepsilon} \right|^2, \beta \right\} + \gamma \sum_{\varepsilon i \in \varepsilon \mathbb{Z}^d \cap D} \varepsilon^d |u(\varepsilon i) - g_\varepsilon(\varepsilon i)|^2. \quad (3)$$

However, in higher dimensions the anisotropy of the lattice  $\mathbb{Z}^d$  leads to anisotropic surface integrals. For  $d = 2$  Chambolle proved in [29] that the functionals  $F_\varepsilon$   $\Gamma$ -converge to an anisotropic version of the Mumford–Shah functional given by

$$F(u) = \int_D |\nabla u|^2 dx + \beta \int_{S_u} |v_u|_1 d\mathcal{H}^1 + \gamma \int_D |u - g|^2 dx,$$

where  $|\cdot|_1$  denotes the  $l_1$ -norm of the normal vector  $v_u$  at  $x \in S_u$  (we remark that with the results obtained in this paper the  $\Gamma$ -convergence above can be extended to any dimension).

In order to avoid the anisotropy, there have been found two approaches: on the one hand, inspired by the nonlocal approximation of the Mumford–Shah functional studied in [45], one can consider long-range interactions in (3) instead of only nearest neighbors. This has been analyzed in [30] and indeed anisotropy can be reduced but the functional to be minimized gets more complex as the number of interactions grows. On the other hand, Chambolle and Dal Maso considered functionals defined on piecewise affine functions with respect to a whole class of two-dimensional triangulations. More precisely, let  $\mathcal{T}_\varepsilon(D, \theta)$  be the set of all finite triangulations containing  $D$  such that for each triangle the inner angles are at least  $\theta$  and the side lengths are between  $\varepsilon$  and  $w(\varepsilon)$ , where  $w(\varepsilon) \geq 6\varepsilon$  satisfies  $\lim_{\varepsilon \rightarrow 0} w(\varepsilon) = 0$ . Denote by  $V_\varepsilon(D, \theta)$  the set of continuous functions that are piecewise affine with respect to some  $\mathcal{T} \in \mathcal{T}_\varepsilon(D, \theta)$ . In [31] it is shown that there exists  $0 < \theta_0 < 60^\circ$  such that for all  $0 < \theta < \theta_0$  the functionals  $F_{\varepsilon, \theta}$  defined on  $V_\varepsilon(D, \theta)$  by

$$F_{\varepsilon, \theta}(u) = \frac{1}{\varepsilon} \int_D f(\varepsilon |\nabla u|^2) dx + \gamma \int_D |u - g|^2 dx$$

$\Gamma$ -converge to the Mumford–Shah functional when  $f$  satisfies (2). In this case it holds that  $\beta = f_\infty \sin(\theta)$ . We remark that the triangulation is not fixed, so it is part of the minimization problem to find the optimal one (see [20] for details on numerical minimization for slightly modified functionals). Moreover, in contrast to the approximations mentioned before, this result is restricted to dimension two.

There are also different approaches to minimize the Mumford–Shah functional, for instance level-set methods, graph cut algorithms or convex relaxation techniques. We do not go into details but refer the reader to [52] and references therein.

The motivation for this work relies on the more recent paper [53], in which Cremers and Strekalovski propose another discrete functional based on finite differences along with a very fast algorithm to compute its minimizers in real-time. Although convergence of the algorithm has not been proven so far, it is demonstrated that it works well in practice. The discrete functional has a similar form to (3), but takes into account non-pairwise interactions and suitably scaled it reads as

$$\tilde{F}_{\varepsilon, g}(u) = \sum_{\varepsilon x \in \varepsilon \mathbb{Z}^d \cap D} \varepsilon^{d-1} \min \left\{ \alpha \varepsilon \sum_{i=1}^d \left| \frac{u(\varepsilon x + \varepsilon e_i) - u(\varepsilon x)}{\varepsilon} \right|^2, 1 \right\} + \gamma \sum_{\varepsilon x \in \varepsilon \mathbb{Z}^d \cap D} \varepsilon^d |u(\varepsilon x) - g_\varepsilon(\varepsilon x)|^2. \quad (4)$$

The authors of [53] conjecture that  $\tilde{F}_{\varepsilon, g}$  approximates the Mumford–Shah functional.

In this paper we analyze the asymptotic behavior of a more general family of discrete functionals. Inspired by the structure of  $\tilde{F}_{\varepsilon, g}$  above, we allow the functionals to depend not only on pairwise but general finite differences.

Furthermore, we replace the periodic lattice by so-called stochastic lattices. As a first consequence of our analysis, which is described more in detail below, we can identify the  $\Gamma$ -limit of the functionals  $\tilde{F}_{\varepsilon,g}$  for  $d = 2$ . In particular, we show that it differs from the Mumford–Shah functional due to an anisotropic surface integral (the last point can be verified in any dimension; see Remark 4.2). Motivated by the fast algorithm for discrete approximations presented in [53], we then construct a random family of discrete functionals for that we can prove  $\Gamma$ -convergence to the Mumford–Shah functional almost surely (a.s.). The basic idea is quite simple: since the anisotropy in the  $\Gamma$ -limit of the family of functionals in (3) stems mostly from the lattice, we replace  $\mathbb{Z}^d$  by a more isotropic point set. Since there exist no isotropic, countable sets, we need to go beyond deterministic models and consider realizations of random point sets. Those have the flexibility to be isotropic at least in distribution.

### The main approximation result

We consider random, countable point sets  $\mathcal{L}(\omega) \subset \mathbb{R}^d$  that satisfy the following geometric constraints:

- (i) There exists  $R > 0$  such that  $\text{dist}(x, \mathcal{L}(\omega)) < R$  for all  $x \in \mathbb{R}^d$ ;
- (ii) There exists  $r > 0$  such that  $\text{dist}(x, \mathcal{L}(\omega) \setminus \{x\}) \geq r$  for all  $x \in \mathcal{L}(\omega)$ .

Given a small parameter  $\varepsilon > 0$  (again representing a kind of mesh-size) and  $q > 1$ , one possible approximation is the family of random functionals  $F_{\varepsilon,g}(\omega)$  defined on functions  $u : \varepsilon\mathcal{L}(\omega) \cap D \rightarrow \mathbb{R}^m$  by

$$F_{\varepsilon,g}(\omega)(u) = \sum_{\substack{(x,y) \in \mathcal{N}(\omega) \\ \varepsilon x, \varepsilon y \in D}} \varepsilon^{d-1} f\left(\varepsilon \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^2\right) + \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap D} \varepsilon^d |u(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q, \quad (5)$$

where  $\mathcal{N}(\omega)$  denotes the set of Voronoi neighbors (see Definition 2.6),  $f$  is a function satisfying (2) and  $g_\varepsilon(\omega)$  is a suitable discretization of some given  $g \in L^q(D, \mathbb{R}^m)$ . We require that the random point set  $\mathcal{L}$  is stationary and isotropic, that means  $\mathcal{L}$  and  $R\mathcal{L} + z$  have the same statistics for all  $z \in \mathbb{Z}^d$  and all  $R \in SO(d)$ . If the shift operation is realized by an ergodic group action and  $g_\varepsilon(\omega) \rightarrow g$  in  $L^q(D, \mathbb{R}^m)$ , our main result, which is stated in full generality in Theorem 4.3, can be summarized as follows:

**Theorem.** *Under the above assumptions, there exist three positive constants  $c_1, c_2, c_3$  such that with probability 1 the functionals  $F_{\varepsilon,g}(\omega)$   $\Gamma$ -converge with respect to the  $L^1(D, \mathbb{R}^m)$ -topology to the deterministic functional  $F_g$  defined by*

$$F_g(u) = \begin{cases} c_1 \int_D |\nabla u|^2 dx + c_2 \mathcal{H}^{d-1}(S_u) + c_3 \int_D |u - g|^q dx & \text{if } u \in L^q(D, \mathbb{R}^m) \cap GSBV^2(D, \mathbb{R}^m), \\ +\infty & \text{otherwise.} \end{cases}$$

Given the probabilistic assumptions above, the result is quite robust. For example, the same limit (with different constant  $c_2$ ) can be proven for the random version of (4) given by

$$\tilde{F}_{\varepsilon,g}(\omega) = \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap D} \varepsilon^{d-1} f\left(\sum_{\substack{(x,y) \in \mathcal{N}(\omega) \\ \varepsilon y \in D}} \varepsilon \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^2\right) + \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap D} \varepsilon^d |u(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q.$$

Some remarks are in order:

- (i) A point process that satisfies all our assumptions is given by the random parking process [44,51].
- (ii) The coefficients  $c_i$  are not explicit but are derived from three abstract homogenization formulas. However, for fixed  $\mathcal{L}$  one can still tune them since  $c_1$  and  $c_2$  are proportional to  $f'(0)$  and  $f_\infty$  respectively, while for  $c_3$  we can multiply the second term in (5) by some factor.
- (iii) We will prove the convergence for more general finite differences (see Section 2). Those require more technical notation that we want to avoid in this introduction. Hence we restrict the description of our analysis below to pairwise interactions via Voronoi neighbors.

- (iv) In the proof we will identify  $u$  with a piecewise constant function on Voronoi cells. However this is not needed for minimizing the functional  $F_{\varepsilon,g}(\omega)$ . In particular one only has to determine the Voronoi neighbors, but no volume of cells or piecewise affine interpolations on Delaunay triangulations. One can also avoid the determination of the Voronoi neighbors using a  $k$ -NN algorithm, but  $k$  should not be too small (see also Remark 2.7 (ii) & (iii)).
- (v) We prove that global minimizers of  $F_{\varepsilon,g}$  converge to minimizers of  $F_g$  in  $L^q(D, \mathbb{R}^m)$ . Note that this is not the natural compactness to be expected from finite energy sequences.
- (vi) The discrete functionals are still non-convex which cannot be avoided since the  $\Gamma$ -limit of any sequence of convex functionals remains convex.
- (vii) In our setting the case of vector-valued  $u$  corresponds to color images. Our arguments cannot be generalized straightforward to models for linearized elasticity.
- (viii) A different randomization of the functional (3) has been considered in another context in [25], where a random choice between the potential  $f(s) = \min\{s, 1\}$  and  $\tilde{f}(s) = s$  is analyzed. However, isotropy of the limit functional remained an open problem.
- (ix) Another approach to construct discrete approximations could be based on random point clouds similar to [42], where the authors prove an approximation result for total variation-type functionals. Point clouds have the advantage that they can be generated very fast. However, for point clouds one usually needs interactions with range  $\sim (\log(n)/n)^{\frac{1}{d}}$  for  $n$  points compared to  $\sim (1/n)^{\frac{1}{d}}$  for stochastic lattices. Otherwise the corresponding graph will not be connected. Besides the nonlocal structure there are also many redundant points due to clustering.

### Plan of the paper

We now give a short overview of the paper and explain briefly the steps to prove our main approximation theorem. Section 2 is divided into three preliminary parts. First we recall the necessary function spaces that we need for our analysis. In the second part we introduce in a rigorous way the stochastic point sets that we use to define our approximating functionals  $F_{\varepsilon,g}(\omega)$ . In the last part we introduce the class of functionals under consideration (we omit the fidelity term in most parts of the paper). For this introduction we assume the functionals to be of the form

$$F_{\varepsilon}(\omega)(u) = \sum_{\substack{(x,y) \in \mathcal{N}(\omega) \\ \varepsilon x, \varepsilon y \in D}} \varepsilon^{d-1} f\left(\varepsilon \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^p\right),$$

where for the sake of generality we take a general exponent  $p > 1$ . The localized versions of these functionals will be the main objects to be studied in the subsequent sections.

In Section 3 we present three general results that we prove on the way towards the main approximation result and which might be of independent interest. Assuming only the geometric properties of a single realization  $\mathcal{L}(\omega)$ , we prove in Theorem 3.3 that (up to subsequences) the  $\Gamma$ -limit of the family  $F_{\varepsilon}(\omega)$  always has the form of a free discontinuity functional, that means it is finite only on  $GSBV^p(D, \mathbb{R}^m)$ , where it can be written as

$$F(\omega)(u) = \int_D h(x, \nabla u) dx + \int_{S_u} \varphi(x, \nu_u) d\mathcal{H}^{d-1}. \quad (6)$$

Moreover, the density  $h$  of the  $\Gamma$ -limit coincides with the density of the  $\Gamma$ -limit of the convex functionals

$$E_{\varepsilon}(\omega)(u) = f'(0) \sum_{\substack{(x,y) \in \mathcal{N}(\omega) \\ \varepsilon x, \varepsilon y \in D}} \varepsilon^d \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^p,$$

while the surface density  $\varphi$  of the functional  $F(\omega)$  in (6) is given by surface density of the  $\Gamma$ -limit of the Ising-type energies defined on functions  $u : \varepsilon \mathcal{L}(\omega) \rightarrow \{\pm e_1\}$  by

$$I_{\varepsilon}(\omega)(u) = \frac{f_{\infty}}{2} \sum_{\substack{(x,y) \in \mathcal{N}(\omega) \\ \varepsilon x, \varepsilon y \in D}} \varepsilon^{d-1} |u(\varepsilon x) - u(\varepsilon y)|.$$

In Theorem 3.5 we state a general stochastic homogenization result for the functionals  $F_\varepsilon(\omega)$  in the case of stationary, ergodic stochastic lattices. In particular, the  $\Gamma$ -limit of  $F_\varepsilon(\omega)$  exists a.s., is deterministic and on its domain it is of the form

$$F(u) = \int_D h(\nabla u) \, dx + \int_{S_u} \varphi(v_u) \, d\mathcal{H}^{d-1},$$

that means in contrast to (6) the densities  $h$  and  $\varphi$  do not depend on  $x$  and are deterministic. Theorem 3.8 contains the  $\Gamma$ -convergence including the convergence of minimizers when we add the discrete fidelity term in the stationary, ergodic setting.

In Section 4 we apply the results of Section 3. On the one hand, we identify the  $\Gamma$ -limit of the functionals in (4). On the other hand, with Theorem 4.3 we obtain our main approximation result about the Mumford–Shah functional when we assume additionally that the stochastic lattice is isotropic in distribution.

Section 5 contains the proof of Theorem 3.3. While the integral form (6) of any  $\Gamma$ -limit is obtained by standard techniques combining the abstract methods of  $\Gamma$ -convergence with an integral representation theorem, the characterizations of the integrands  $h$  and  $\varphi$  by the  $\Gamma$ -limits of the sequences  $E_\varepsilon(\omega)$  and  $I_\varepsilon(\omega)$  is the most delicate step in this paper. Although similar results have been obtained in a continuum setting (see [24, 28, 43]), we cannot use interpolation and copy the argument. This has several reasons: on the one hand, in dimensions larger than two, piecewise affine interpolations on Delaunay tessellations might be degenerate due to very flat tetrahedrons. On the other hand, even in a planar setting Voronoi cells can have very short boundary sides, so that the discrete functional overestimates the length of interfaces. Moreover, fine constructions based on geometric measure theory can be incompatible with the prescribed lattice structure. Thus our arguments, which are nevertheless inspired by the continuum case, need to use the discrete environment as long as possible. The complete strategy is explained more in detail at the beginning of Section 5.

In Section 6 we prove the remaining results of Section 3. With the characterizations proven in the previous section, Theorem 3.5 is a straightforward consequence of the results on discrete-to-continuum stochastic homogenization for elastic and Ising-type energies obtained in the two papers [5, 6], respectively. Also adding the fidelity term is quite straightforward.

The appendix contains a technical argument how to choose  $\Gamma$ -converging diagonal sequences in our special setting when the functionals are not equicoercive.

## 2. Setting of the problem and preliminaries

We first introduce some notation that will be used in this paper. Given a measurable set  $B \subset \mathbb{R}^d$  we denote by  $|B|$  its  $d$ -dimensional Lebesgue measure, while more generally  $\mathcal{H}^k(A)$  stands for the  $k$ -dimensional Hausdorff measure. We denote by  $\mathbb{1}_B$  the characteristic function of  $B$ . If  $B$  is finite,  $\#B$  means its cardinality. Given an open set  $O \subset \mathbb{R}^d$ , we denote by  $\mathcal{A}(O)$  the family of all bounded, open subsets of  $O$ , while  $\mathcal{A}^R(O)$  means the bounded, open subsets with Lipschitz boundary. For  $x \in \mathbb{R}^d$  or  $y \in \mathbb{R}^m$  we denote by  $|x|$  and  $|y|$  the Euclidean norm. Given a matrix  $\xi \in \mathbb{R}^{m \times d}$ , we let  $|\xi|$  be its Frobenius norm. As usual  $B_\varrho(x_0)$  denotes the open ball with radius  $\varrho$  centered at  $x_0 \in \mathbb{R}^d$ . We simply write  $B_\varrho$  when  $x_0 = 0$ . Given  $v \in S^{d-1}$ , we let  $v_1 = v, v_2, \dots, v_d$  be an orthonormal basis of  $\mathbb{R}^d$  and we define the cube  $Q_v$  as

$$Q_v = \left\{ z \in \mathbb{R}^d : |\langle z, v_i \rangle| < 1/2 \right\},$$

where the brackets  $\langle \cdot \rangle$  denote the scalar product. Given  $x_0 \in \mathbb{R}^d$  and  $\varrho > 0$ , we set  $Q_v(x_0, \varrho) = x_0 + \varrho Q_v$ . For  $x_0 \in \mathbb{R}^d$ ,  $v \in S^1$  and  $a, b \in \mathbb{R}^m$  we define the function  $u_{x_0, v}^{a, b}$  by the formula

$$u_{x_0, v}^{a, b}(x) := \begin{cases} a & \text{if } \langle x - x_0, v \rangle > 0, \\ b & \text{otherwise.} \end{cases} \quad (7)$$

The notation  $\text{co}(x_1, \dots, x_d)$  means the convex hull of finitely many points in  $\mathbb{R}^d$ . We will use  $\|u\|_{L^p(A)}$  for the  $L^p(A, \mathbb{R}^m)$ -norm. There should be no confusion about the dimension  $m$ . The symbol  $\otimes$  stands for the outer product of vectors, that is, for any  $a \in \mathbb{R}^m, b \in \mathbb{R}^d$  we have  $a \otimes b \in \mathbb{R}^{m \times d}$  with  $(a \otimes b)_{ij} := a_i b_j$ . Finally, the letter  $C$  stands for a generic positive constant that may change every time it appears.

### 2.1. Generalized special functions of bounded variation

We briefly recall the function spaces we are going to use in this paper. We refer to [7, 11, 36] for more details.

Let  $O \subset \mathbb{R}^d$  be an open set. We denote by  $BV(O, \mathbb{R}^m) \subset L^1(O, \mathbb{R}^m)$  the space of vector-valued functions of bounded variation. We write  $Du$  for the matrix-valued distributional derivative of  $u$ , which can be decomposed as

$$Du(B) = \int_B \nabla u \, dx + \int_{S_u \cap B} (u^+(x) - u^-(x)) \otimes \nu_u(x) \, d\mathcal{H}^{d-1} + D^c u(B).$$

In the above formula  $\nabla u$  denotes the absolutely continuous part of  $Du$  with respect to the Lebesgue measure,  $S_u$  is the so-called jump set of  $u$  with (measure-theoretic) normal vector  $\nu_u \in S^{d-1}$  and  $u^-$ ,  $u^+$  denote the one-sided traces of  $u$  at  $\mathcal{H}^{d-1}$ -a.e.  $x \in S_u$ . The remainder  $D^c u$  is called the Cantor-part, but it will play no role in this paper.

Indeed, the space of *special functions of bounded variation* is defined as the set of those  $u \in BV(O, \mathbb{R}^m)$  such that  $D^c u = 0$ . We write  $u \in SBV(O, \mathbb{R}^m)$ . Given  $p \in (1, +\infty)$ , we define  $SBV^p(O, \mathbb{R}^m) \subset SBV(O, \mathbb{R}^m)$  as the set of those functions such that  $\nabla u \in L^p(O, \mathbb{R}^{m \times d})$  and  $\mathcal{H}^{d-1}(S_u) < +\infty$ . Due to a lack of compactness in many free discontinuity problems, we have to enlarge this class. We say that a Borel-function  $u : O \rightarrow \mathbb{R}^m$  is a *generalized special function of bounded variation*, if  $\Phi \circ u \in SBV_{\text{loc}}(O, \mathbb{R}^m)$  for every function  $\Phi \in C^1(\mathbb{R}^m, \mathbb{R}^m)$  such that  $\nabla \Phi$  has compact support and write  $u \in GSBV(O, \mathbb{R}^m)$ . In this case, the approximate differential  $\nabla u(x)$  still exists a.e. and there is a well-defined jump set  $S_u$ , which is countably  $(d-1)$ -rectifiable. Finally, we set  $GSBV^p(O, \mathbb{R}^m)$  as those functions  $u \in GSBV(O, \mathbb{R}^m)$  such that  $\nabla u \in L^p(O, \mathbb{R}^{m \times d})$  and  $\mathcal{H}^{d-1}(S_u) < +\infty$ . As shown in [36, Section 2], the set  $GSBV^p(O, \mathbb{R}^m)$  is a vector space and, if  $\Phi \in C^1(\mathbb{R}^m, \mathbb{R}^m)$  is such that  $\nabla \Phi$  has compact support, then  $\Phi \circ u \in SBV^p(O, \mathbb{R}^m)$ . Moreover, given  $u \in GSBV^p(O, \mathbb{R}^m)$ , one can define a Borel-function  $\nu_u : S_u \rightarrow S^{d-1}$  and two Borel-functions  $u^+, u^- : S_u \rightarrow \mathbb{R}^m$  still satisfying a weak trace condition for  $\mathcal{H}^{d-1}$ -a.e.  $x \in S_u$ .

For our analysis we make use of a special family of smooth truncations as in [28] which essentially allows to reduce many proofs to the space  $SBV^p$ . Consider a function  $\phi \in C_c^\infty(\mathbb{R})$  such that  $\phi(t) = t$  for all  $|t| \leq 1$ ,  $\phi(t) = 0$  for  $t \geq 3$  and  $\|\phi'\|_\infty \leq 1$ . We define the function  $\Phi \in C_c^\infty(\mathbb{R}^m, \mathbb{R}^m)$  by

$$\Phi(u) = \begin{cases} \phi(|u|) \frac{u}{|u|} & \text{if } u \neq 0, \\ 0 & \text{if } u = 0. \end{cases}$$

As shown at the beginning of [28, Section 4] the function  $\Phi$  is 1-Lipschitz. Given  $k > 0$  we further set  $\Phi_k(u) = k\Phi(\frac{u}{k})$ , which is still 1-Lipschitz. Then we have the following approximation result, which is a consequence of dominated convergence and [7, Propositions 1.1-1.3 & Theorem 3.7].

**Lemma 2.1.** *Let  $u \in GSBV^p(D, \mathbb{R}^m) \cap L^1(D, \mathbb{R}^m)$  and let  $k > 0$ . Defining the truncation  $T_k u = \Phi_k(u)$ , the function  $T_k u$  belongs to  $SBV^p(D, \mathbb{R}^m) \cap L^\infty(D, \mathbb{R}^m)$  and*

- (i)  $\lim_{k \rightarrow +\infty} T_k u = u$  a.e. and in  $L^1(D, \mathbb{R}^m)$ ;
- (ii)  $\nabla T_k u(x) = \nabla \Phi_k(u(x)) \nabla u(x)$  for a.e.  $x \in D$ ;
- (iii)  $S_{T_k u} \subset S_u$ ,  $\lim_{k \rightarrow +\infty} \mathcal{H}^{d-1}(S_{T_k u}) = \mathcal{H}^{d-1}(S_u)$  and  $\nu_u = \pm \nu_{T_k u}$   $\mathcal{H}^{d-1}$ -a.e. on  $S_{T_k u}$ .

### 2.2. Stochastic lattices

Next we introduce the random point sets that we use for the discrete approximations. Throughout this paper we let  $\Omega$  be a probability space with a complete  $\sigma$ -algebra  $\mathcal{F}$  and probability measure  $\mathbb{P}$ . We call a random variable  $\mathcal{L} : \Omega \rightarrow (\mathbb{R}^d)^\mathbb{N}$  a stochastic lattice. The following definition, which has been introduced in [17] in the context of quantum models (and is known in a deterministic setting as Delone sets), essentially forbids clustering of points as well as arbitrarily big empty regions in space.

**Definition 2.2 (Admissible lattices).** Let  $\mathcal{L}$  be a stochastic lattice.  $\mathcal{L}$  is called admissible if there exist  $R > r > 0$  such that the following two conditions hold a.s.:



- (i)  $\text{dist}(x, \mathcal{L}(\omega)) < R$  for all  $x \in \mathbb{R}^d$ ;
- (ii)  $\text{dist}(x, \mathcal{L}(\omega) \setminus \{x\}) \geq r$  for all  $x \in \mathcal{L}(\omega)$ .

**Remark 2.3.** We also make use of the associated Voronoi tessellation  $\mathcal{V}(\omega) = \{\mathcal{C}(x)\}_{x \in \mathcal{L}(\omega)}$ , where the (random) Voronoi cells with nuclei  $x \in \mathcal{L}(\omega)$  are defined as

$$\mathcal{C}(x) = \{z \in \mathbb{R}^d : |z - x| \leq |z - y| \text{ for all } y \in \mathcal{L}(\omega)\}.$$

If  $\mathcal{L}(\omega)$  is admissible, then [6, Lemma 2.3] yields the inclusions  $B_{\frac{r}{2}}(x) \subset \mathcal{C}(x) \subset B_R(x)$ .

Next we introduce some notions from ergodic theory that build the basis for stochastic homogenization.

**Definition 2.4.** We say that a family of measurable functions  $\{\tau_z\}_{z \in \mathbb{Z}^d}$ ,  $\tau_z : \Omega \rightarrow \Omega$ , is an additive group action on  $\Omega$  if

$$\tau_0 = \text{id} \quad \text{and} \quad \tau_{z_1+z_2} = \tau_{z_2} \circ \tau_{z_1} \quad \text{for all } z_1, z_2 \in \mathbb{Z}^d.$$

An additive group action is called measure preserving if

$$\mathbb{P}(\tau_z^{-1}B) = \mathbb{P}(B) \quad \text{for all } B \in \mathcal{F}, z \in \mathbb{Z}^d.$$

Moreover,  $\{\tau_z\}_{z \in \mathbb{Z}^d}$  is called ergodic if, in addition, for all  $B \in \mathcal{F}$  we have the implication

$$(\tau_z(B) = B \quad \forall z \in \mathbb{Z}^d) \Rightarrow \mathbb{P}(B) \in \{0, 1\}.$$

In terms of a stochastic lattice the probabilistic properties read as follows:

**Definition 2.5.** A stochastic lattice  $\mathcal{L}$  is said to be stationary if there exists an additive, measure preserving group action  $\{\tau_z\}_{z \in \mathbb{Z}^d}$  on  $\Omega$  such that for all  $z \in \mathbb{Z}^d$

$$\mathcal{L} \circ \tau_z = \mathcal{L} + z.$$

If in addition  $\{\tau_z\}_{z \in \mathbb{Z}^d}$  is ergodic, then  $\mathcal{L}$  is called ergodic, too.

We call  $\mathcal{L}$  isotropic, if for every  $R \in SO(d)$  there exists a measure preserving function  $\tau'_R : \Omega \rightarrow \Omega$  such that

$$\mathcal{L} \circ \tau'_R = R\mathcal{L}.$$

In order to define gradient-like structures, we equip a stochastic lattice with a set of directed edges. We summarize the necessary properties in a separate definition:

**Definition 2.6 (Admissible edges).** Let  $\mathcal{L}$  be an admissible stochastic lattice and  $\mathcal{E} \subset \mathcal{L}^2$ . We call  $\mathcal{E}$  admissible edges if for all  $i, j \in \mathbb{N}$  the set  $\{\omega \in \Omega : (\mathcal{L}(\omega)_i, \mathcal{L}(\omega)_j) \in \mathcal{E}(\omega)\}$  is  $\mathcal{F}$ -measurable and

- (i) there exists  $M > R$  such that a.s.

$$\sup\{|x - y| : (x, y) \in \mathcal{E}(\omega)\} < M; \tag{8}$$

- (ii) the Voronoi neighbors defined by  $\mathcal{N}(\omega) := \{(x, y) \in \mathcal{L}(\omega)^2 : \mathcal{H}^{d-1}(\mathcal{C}(x) \cap \mathcal{C}(y)) \in (0, +\infty)\}$  are contained in  $\mathcal{E}(\omega)$  up to symmetrizing, that means

$$\mathcal{N}(\omega) \subset \mathcal{E}(\omega) \cup \{(y, x) \in \mathcal{L}(\omega)^2 : (x, y) \in \mathcal{E}(\omega)\}. \tag{9}$$

If  $\mathcal{L}$  is stationary or isotropic, we say that the edges  $\mathcal{E}$  are stationary or isotropic if  $\mathcal{E} \circ \tau_z = \mathcal{E} + (z, z)$  for all  $z \in \mathbb{Z}^d$  or  $\mathcal{E} \circ \tau'_R = R\mathcal{E}$  for all  $R \in SO(d)$ .

Up to enlarging  $M$ , by Remark 2.3 we may assume in addition that

$$\sup_{x \in \mathcal{L}(\omega)} \#\{y \in \mathcal{L}(\omega) : (x, y) \in \mathcal{E}(\omega) \text{ or } (y, x) \in \mathcal{E}(\omega)\} \leq M. \tag{10}$$



**Remark 2.7.**

- (i) In the proof of [6, Lemma A.2] it is shown, that the choice  $\mathcal{E}(\omega) = \mathcal{N}(\omega)$  satisfies the measurability assumption. Then we can add for example non-Voronoi neighbors by selecting them based on a maximal distance.
- (ii) In order to avoid the computation of the Voronoi neighbors (or the Delaunay triangulation), one can also use a  $k$ -NN algorithm (taking all points in case of a tie to avoid anisotropy). By a naive volume bound based on Remark 2.3 it suffices to take

$$k = \left\lceil \left(4Rr^{-1} + 1\right)^d - \left(2Rr^{-1} - 1\right)^d \right\rceil - 2.$$

Indeed, if  $(x, y) \in \mathcal{N}(\omega)$  is a pair of Voronoi neighbors, then for any point  $z \in \mathcal{C}(x) \cap \mathcal{C}(y)$  it holds that  $|z - x| < R$  and the ball  $B_{|x-z|}(z)$  is contained in  $B_{2|x-z|}(x)$  and contains no other point of  $\mathcal{L}(\omega)$ . For all other points  $x' \in \mathcal{L}(\omega)$  with  $|x' - x| \leq |y - x|$  (including  $x$ ) the balls  $B_{\frac{r}{2}}(x')$  are pairwise disjoint and contained in the set  $B_{2|x-z|+\frac{r}{2}}(x) \setminus B_{|x-z|-\frac{r}{2}}(z)$ . The volume of this set is monotone in  $|x - z| \geq r/2$ , which implies the claimed bound. For the random parking model we have the optimal ratio  $R/r = 2$  at the packing limit, which yields  $k = 70$  for  $d = 2$ . This bound is far from being optimal. Nevertheless, a more detailed treatment of numerical issues is beyond the scope of this paper.

- (iii) We use (9) only for proving Lemma 5.1. Hence in (9) the set  $\mathcal{N}(\omega)$  can be replaced by any set of edges such that this lemma remains valid.

Having introduced the random framework, we will need it again only in Section 6. To simplify the notation, we will drop the dependence on  $\omega$  for some quantities. If so, we tacitly assume that we have a realization satisfying properties (i) and (ii) of both Definitions 2.2 and 2.6.

### 2.3. A generalized weak-membrane energy

In order to discretize vectorial Mumford–Shah-type functionals we basically follow the approach used on periodic lattices for the scalar case. However we go beyond pairwise interactions. Due to the possibly non-ordered edges  $\mathcal{E}(\omega)$  this requires some notation. For  $M \in \mathbb{N}$  satisfying (10) we denote by

$$\mathcal{P}_+(M) = \{\mathbf{p} : [0, +\infty) \rightarrow \mathbb{N}_0 : \#\mathbf{p}^{-1}(\mathbb{N}) < +\infty, \sum_{v \in \mathbf{p}^{-1}(\mathbb{N})} \mathbf{p}(v) \leq M\}$$

the set of all multisets over  $[0, +\infty)$  with at most  $M$  elements. Note that if  $v_1, \dots, v_k \in [0, +\infty)$  with  $k \leq M$ , then up to permutation we can identify these points with a unique  $\mathbf{p} \in \mathcal{P}_+(M)$  setting  $\mathbf{p}(v_i) = \#\{j : v_j = v_i\}$  for all  $1 \leq i \leq k$  and zero elsewhere. In this sense we sometimes use the more common notation  $\mathbf{p} = \{v_i\}_b$ , where the  $b$  indicates a badge in which elements can occur several times in contrast to ordinary sets. For  $\mathbf{p} \in \mathcal{P}_+(M)$  we set  $\|\mathbf{p}\|_1 = \sum_{v \in \mathbf{p}^{-1}(\mathbb{N})} \mathbf{p}(v)v$ . In this paper we fix a bounded function  $f : \mathcal{P}_+(M) \rightarrow [0, +\infty)$  satisfying the following two structural assumptions: there exists  $0 < \alpha < +\infty$  such that

$$\lim_{\|\mathbf{p}\|_1 \rightarrow 0} \frac{f(\mathbf{p})}{\|\mathbf{p}\|_1} = \alpha \tag{11}$$

and  $f$  is monotone increasing in the following sense: for all  $v, v' \in [0, +\infty)^M$  with  $v_i \leq v'_i$  for all  $1 \leq i \leq M$  and for any  $1 \leq k \leq M$  we have

$$f(\{v_1, \dots, v_k\}_b) \leq f(\{v'_1, \dots, v'_k\}_b). \tag{12}$$

We also assume that  $f$  is lower semicontinuous in the sense that for all sequences  $(v^n)_n \subset [0, +\infty)^M$  converging to some  $v \in [0, +\infty)^M$  and for all  $1 \leq k \leq M$  it holds that

$$f(\{v_1, \dots, v_k\}_b) \leq \liminf_{n \rightarrow +\infty} f(\{v_1^n, \dots, v_k^n\}_b). \tag{13}$$

**Remark 2.8.** Note that from the boundedness of  $f$ , the monotonicity assumption and the property (11) it follows that there exist constants  $C_f > c_f > 0$  such that

$$c_f \min\{\|p\|_1, 1\} \leq f(p) \leq C_f \min\{\|p\|_1, 1\}. \quad (14)$$

Moreover, again by boundedness and monotonicity, for every  $1 \leq l \leq k \leq M$  there exists the limit

$$\beta(l, k) = \lim_{N \rightarrow +\infty} f(l\mathbb{1}_{\{N\}} + (k-l)\mathbb{1}_{\{0\}}) > 0. \quad (15)$$

While we will frequently use the property (12) for our analysis, the lower semicontinuity (13) will just guarantee the existence of minimizers for the discrete approximations.

In order to define the discrete approximation of a continuous functional we scale a stochastic lattice by a small parameter  $\varepsilon > 0$ . Let us fix a reference set  $D \subset \mathbb{R}^d$ , which we assume to be a bounded Lipschitz domain, and a growth exponent  $p \in (1, +\infty)$ . Given  $u : \mathcal{E}\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$ , an open set  $A \in \mathcal{A}(D)$  and  $\eta > 0$ , we define the function  $\eta|\nabla_{\omega,\varepsilon}(u, A)|^p : \mathcal{E}\mathcal{L}(\omega) \rightarrow \mathcal{P}_+(M)$  by

$$\eta|\nabla_{\omega,\varepsilon}(u, A)|^p(\varepsilon x) = \left\{ \eta \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^p : (x, y) \in \mathcal{E}(\omega), \varepsilon x, \varepsilon y \in A \right\}_b.$$

Then we define the localized discrete approximations (which we also call energies) as

$$F_\varepsilon(\omega)(u, A) = \sum_{\varepsilon x \in \mathcal{E}\mathcal{L}(\omega) \cap A} \varepsilon^{d-1} f(\varepsilon|\nabla_{\omega,\varepsilon}(u, A)|^p(\varepsilon x)).$$

We simply write  $F_\varepsilon(\omega)$  for  $F_\varepsilon(\omega)(\cdot, D)$ , which will be the functional of interest in this paper.

**Remark 2.9.** We chose the abstract framework above for two reasons.

- (i) We take directed edges to define  $\eta|\nabla_{\omega,\varepsilon}(u, A)|^p(\varepsilon x)$  in order to include the functional (4);
- (ii) we define the function  $f$  on multisets to handle pairwise and non-pairwise gradients simultaneously. For pairwise interactions, we set  $f(\mathbf{p}) = \sum_{v \in \mathbf{p}^{-1}(\mathbb{N})} \mathbf{p}(v) f_0(v)$  with  $f_0$  satisfying (2). The other example of the introduction is given by  $f(\mathbf{p}) = f_0(\|\mathbf{p}\|_1)$ .

As we aim at using the abstract theory of  $\Gamma$ -convergence, we will identify a discrete variable with its piecewise constant interpolation on the Voronoi cells, that means with functions of the class

$$\mathcal{PC}_\varepsilon^\omega = \{u : \mathbb{R}^d \rightarrow \mathbb{R}^m : u|_{\mathcal{C}(x)} \text{ is constant for all } x \in \mathcal{L}(\omega)\}.$$

With a slight abuse of notation we extend the functional to  $F_\varepsilon(\omega) : L^1(D, \mathbb{R}^m) \times \mathcal{A}(D) \rightarrow [0, +\infty]$  by

$$F_\varepsilon(\omega)(u, A) = \begin{cases} F_\varepsilon(\omega)(u, A) & \text{if } u \in \mathcal{PC}_\varepsilon^\omega, \\ +\infty & \text{otherwise.} \end{cases} \quad (16)$$

**Remark 2.10.** We work in the space  $L^1$  due to the applications we have in mind. A priori there is no equicoercivity in this space and therefore it seems more natural to define the functionals for example on measurable functions. However, we will show in Lemma 5.6 that this can be circumvented using a fidelity term that is part of the Mumford–Shah functional anyway.

### 3. Presentation of the general results

Having introduced the necessary notation, we now give an overview of the theoretical results that we prove on the way to obtain the discretization of the Mumford–Shah functional. Readers who are interested specifically in the Mumford–Shah functional can go to Section 4.2.

### 3.1. Integral representation and separation of bulk and surface effects

Our first result establishes a general  $\Gamma$ -convergence result (up to subsequences), which holds without any probabilistic assumptions on the lattice, but pointwise for admissible lattices and edges. We show that any  $\Gamma$ -limit has an integral representation which is characterized by the separate  $\Gamma$ -limits of an associated bulk energy and a surface energy. More precisely, a formal linearization of  $F_\varepsilon(\omega)$  at  $u = 0$  yields the bulk energy  $E_\varepsilon(\omega)(\cdot, A) : \mathcal{PC}_\varepsilon^\omega \rightarrow [0, +\infty]$  defined by

$$E_\varepsilon(\omega)(u, A) = \alpha \sum_{\substack{(x,y) \in \mathcal{E}(\omega) \\ \varepsilon x, \varepsilon y \in A}} \varepsilon^d \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^p, \quad (17)$$

where  $\alpha$  is given by (11). At least when  $u$  is the discretization of a Lipschitz function, the functional  $E_\varepsilon(\omega)(u, A)$  should be a good approximation of  $F_\varepsilon(\omega)(u, A)$  since  $\|\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p\|_1$  vanishes when  $\varepsilon \rightarrow 0$ . On the contrary, when  $u$  is the discretization of a (macroscopically) piecewise constant function, the elements in  $\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p$  either equal zero or blow up. Hence one expects that  $F_\varepsilon(\omega)(u, A)$  should be well-approximated by the functional  $I_\varepsilon(\omega) : \{v : \varepsilon \mathcal{L}(\omega) \rightarrow \{\pm e_1\}\} \rightarrow [0, +\infty]$  given by the formula

$$I_\varepsilon(\omega)(v, A) = \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap A} \varepsilon^{d-1} \beta \left( \sum_{\substack{\varepsilon y \in \varepsilon \mathcal{L}(\omega) \cap A \\ (x,y) \in \mathcal{E}(\omega)}} \frac{1}{2} |v(\varepsilon x) - v(\varepsilon y)|, \#\{\varepsilon y \in \varepsilon \mathcal{L}(\omega) \cap A : (x, y) \in \mathcal{E}(\omega)\} \right), \quad (18)$$

where the function  $\beta$  is given by (15) and  $\beta(0, k) := 0$  for all  $1 \leq k \leq M$ . Indeed, the function  $\beta$  takes into account how many components of the discrete gradient blow up while all others remain zero.

In order to formulate the theorem we recall the  $\Gamma$ -convergence results obtained for the energies  $E_\varepsilon(\omega)$  and  $I_\varepsilon(\omega)$  (both extended to  $L^p(D)$  via  $+\infty$ ). The following result was proven in [5, Theorem 3]<sup>1</sup>:

**Theorem 3.1** ([5]). *For every sequence  $\varepsilon \rightarrow 0$  there exists a subsequence  $\varepsilon_n$  such that for every  $A \in \mathcal{A}^R(D)$  the functionals  $E_{\varepsilon_n}(\omega)(\cdot, A) \Gamma(L^p(D, \mathbb{R}^m))$ -converge to a functional  $E(\omega) : L^p(D, \mathbb{R}^m) \times \mathcal{A}^R(D) \rightarrow [0, +\infty]$  that is finite only on  $W^{1,p}(A, \mathbb{R}^m)$ , where it takes the form*

$$E(\omega)(u, A) = \int_A q(x, \nabla u) \, dx$$

for some non-negative Carathéodory-function  $q$  that is quasiconvex in the second variable for a.e.  $x \in D$  and satisfies the growth conditions

$$\frac{1}{C} |\xi|^p - C \leq q(x, \xi) \leq C(|\xi|^p + 1).$$

The  $\Gamma$ -convergence of functionals of the type (18) was treated in [6, Theorem 3.2]<sup>2</sup>:

**Theorem 3.2** ([6]). *For every sequence  $\varepsilon \rightarrow 0$  there exists a subsequence  $\varepsilon_n$  such that for every  $A \in \mathcal{A}^R(D)$  the functionals  $I_{\varepsilon_n}(\omega)(\cdot, A) \Gamma(L^p(D, \mathbb{R}^m))$ -converge to a functional  $I(\omega) : L^p(D, \mathbb{R}^m) \times \mathcal{A}^R(D) \rightarrow [0, +\infty]$  that is finite only on  $BV(A, \{\pm e_1\})$ , where it takes the form*

$$I(\omega)(u, A) = \int_A s(x, v_u) \, d\mathcal{H}^{d-1}$$

for some measurable function  $s$  that satisfies the growth conditions

<sup>1</sup> Interactions via a random edge set  $\mathcal{E}(\omega)$  are not covered by the results in [5]. However, the proof works for general finite range interactions with  $p$ -growth and coercive Voronoi neighbor interactions. Due to (9) those assumptions are fulfilled.

<sup>2</sup> In [6] the proofs are given only for pairwise interactions, but as already mentioned in [6, Section 6.3], up to minor modifications they cover also multi-body interactions. Hence we decided not to repeat the arguments.

$$\frac{1}{C} \leq s(x, v) \leq C.$$

Our first main result states that the  $\Gamma$ -limit of  $F_\varepsilon(\omega)$  can be characterized by the  $\Gamma$ -limits of  $E_\varepsilon(\omega)$  and  $I_\varepsilon(\omega)$  in the following sense:

**Theorem 3.3.** *Let  $\varepsilon_n \rightarrow 0$ . The  $\Gamma(L^1(D))$ -limit of  $F_{\varepsilon_n}(\omega)$  exists if and only if the  $\Gamma(L^p(D))$ -limits of  $E_{\varepsilon_n}(\omega)(\cdot, D)$  and  $I_{\varepsilon_n}(\omega)(\cdot, D)$  exist. In this case the  $\Gamma$ -limit  $F(\omega)$  is finite only  $GSBV^p(D, \mathbb{R}^m)$ , where it is given by*

$$F(\omega)(u) = \int_D q(x, \nabla u(x)) \, dx + \int_{S_u} s(x, v_u) \, d\mathcal{H}^{d-1}$$

with the integrands given by Theorems 3.1 and 3.2.

**Remark 3.4.** Both Theorems 3.1 and 3.2 provide asymptotic formulas for the integrands  $q$  and  $s$ , respectively. To state them, we need some notation. Given a set  $A \in \mathcal{A}^R(D)$  and  $\delta > 0$ , we set

$$\partial_\delta A := \{x \in \mathbb{R}^d : \text{dist}(x, \partial A) \leq \delta\}. \quad (19)$$

For a pointwise well-defined function  $\bar{u} \in L^\infty_{\text{loc}}(\mathbb{R}^d, \mathbb{R}^m)$ , we define the set of discrete functions taking the boundary value  $\bar{u}$  on  $\partial_\delta A$  as

$$\mathcal{PC}_{\varepsilon, \delta}^\omega(\bar{u}, A) = \{u \in \mathcal{PC}_\varepsilon^\omega : u(\varepsilon x) = \bar{u}(\varepsilon x) \text{ for all } \varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap \partial_\delta A\}. \quad (20)$$

Then, as shown in Step 1 of the proof of [5, Theorem 2], for a.e.  $x_0 \in D$  and every  $\xi \in \mathbb{R}^{m \times d}$  it holds that

$$q(x_0, \xi) = \lim_{\varrho \rightarrow 0} \varrho^{-d} \lim_{n \rightarrow +\infty} \left( \inf \left\{ E_{\varepsilon_n}(\omega)(v, Q_{e_1}(x_0, \varrho)) : v \in \mathcal{PC}_{\varepsilon_n, M\varepsilon_n}^\omega(\xi(\cdot - x_0), Q_{e_1}(x_0, \varrho)) \right\} \right),$$

where  $M$  denotes the maximal range of interactions given by Definition 2.6.

The formula we use for  $s$  can be found at the beginning of the proof of [22, Theorem 5.8]. It states that for all  $x_0 \in D$  and  $v \in S^1$  we have

$$s(x_0, v) = \limsup_{\varrho \rightarrow 0} \frac{1}{\varrho^{d-1}} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow +\infty} \left( \inf \left\{ I_{\varepsilon_n}(\omega)(v, Q_v(x_0, \varrho)) : v \in \mathcal{PC}_{\varepsilon_n, \delta}^\omega(u_{x_0, v}^{-e_1, e_1}, Q_{e_1}(x_0, \varrho)) \right\} \right),$$

where  $u_{x_0, v}^{-e_1, e_1}$  is defined in (7). Note that the minimization problem defining  $s$  is automatically restricted to functions with values in  $\{\pm e_1\}$ .

### 3.2. Stochastic homogenization

The second main result uses the statistical properties of the lattice and the edges (more precisely only stationarity and ergodicity) in order to avoid passing to subsequences for the  $\Gamma$ -convergence result. It is a straightforward consequence of Theorem 3.3 and the homogenization results proven in [5, 6].

**Theorem 3.5.** *Assume that  $\mathcal{L}$  is a stationary and ergodic stochastic lattice with admissible stationary edges in the sense of Definitions 2.2 and 2.6. Then  $\mathbb{P}$ -a.s. the functionals  $F_\varepsilon(\omega)$   $\Gamma$ -converge in the  $L^1(D, \mathbb{R}^m)$ -topology to a deterministic functional  $F : L^1(D, \mathbb{R}^m) \rightarrow [0, +\infty]$  with domain  $L^1(D, \mathbb{R}^m) \cap GSBV^p(D, \mathbb{R}^m)$ , where it is given by*

$$F(u) = \int_D h(\nabla u) \, dx + \int_{S_u} \varphi(v_u) \, d\mathcal{H}^{d-1}$$

for some convex,  $p$ -homogeneous function  $h$  and some convex, one-homogeneous function  $\varphi$ .

**Remark 3.6.**

- (i) By Theorem 3.3 the functions  $h$  and  $\varphi$  are given by the formulas in Remark 3.4 for  $q(0, \cdot)$  and  $s(0, \cdot)$ , respectively. However, we can take  $\varrho = 1$  in both formulas and in the formula for  $s(0, \cdot)$  one can avoid the additional limit in  $\delta$  by setting  $\delta = 2M\varepsilon$ . Moreover, every remaining lim sup-expression can be replaced by a limit.
- (ii) The above  $\Gamma$ -convergence result also holds locally on each  $A \in \mathcal{A}^R(D)$  for the same set of  $\omega$ .
- (iii) Theorem 3.5 still holds if we drop the ergodicity assumption. Then the integrands are  $\tau$ -invariant, but possibly random.

### 3.3. Convergence of minimizers in the stationary, ergodic setting

Now we add a discrete fidelity term to our approximating functional  $F_\varepsilon(\omega)$ , which will approximate the continuum fidelity term that penalizes the distance to the measured image.

In order to define the discrete approximation, we consider a discrete measurement of a given continuum function. More precisely, throughout this section we fix an exponent  $q > 1$  and consider a sequence  $g_\varepsilon(\omega) : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$ , for which we assume that there exists  $g \in L^q(D, \mathbb{R}^m)$  such that  $\mathbb{P}$ -a.s.

$$g_\varepsilon(\omega) \rightarrow g \quad \text{in } L^q(D, \mathbb{R}^m). \quad (21)$$

**Remark 3.7.** For every given  $g \in L^q(D, \mathbb{R}^m)$ , we find a sequence with this approximation property by first extending  $g = 0$  on  $\mathbb{R}^d \setminus D$  and then setting

$$g_\varepsilon(\omega)(\varepsilon x) = \frac{1}{|B_\varepsilon(\varepsilon x)|} \int_{B_\varepsilon(\varepsilon x)} g(z) \, dz.$$

To see this, we may assume that  $m = 1$ . It is a consequence of the (generalized) Lebesgue differentiation theorem (see [40, Remark 1.160]) that  $g_\varepsilon(\omega) \rightarrow g$  a.e. on  $D$ . Since  $g_\varepsilon$  is bounded in modulus by the maximal function of  $g$  (which belongs itself to  $L^q(\mathbb{R}^d)$ ), we deduce (21) from dominated convergence.

Given a sequence  $g_\varepsilon(\omega)$  satisfying (21), we introduce  $F_{\varepsilon,g}(\omega) : L^1(D, \mathbb{R}^m) \rightarrow [0, +\infty]$  defined by

$$F_{\varepsilon,g}(\omega)(u) = \begin{cases} F_\varepsilon(\omega)(u) + \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap D} \varepsilon^d |u(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q & \text{if } u \in \mathcal{PC}_\varepsilon^\omega, \\ +\infty & \text{otherwise,} \end{cases} \quad (22)$$

where  $F_\varepsilon(\omega)$  is defined in (16). Note that we chose a discrete fidelity term not depending on the measure of the Voronoi cells. The identification of the  $\Gamma$ -limit of  $F_{\varepsilon,g}(\omega)$  is contained in the following theorem.

**Theorem 3.8.** Let  $q \in (1, +\infty)$  and  $g_\varepsilon(\omega)$  satisfy (21). Under the assumptions of Theorem 3.5, there exists a constant  $\gamma > 0$  such that  $\mathbb{P}$ -a.s. the functionals  $F_{\varepsilon,g}(\omega)$  defined in (22)  $\Gamma$ -converge with respect to the  $L^1(D, \mathbb{R}^m)$ -topology to the functional  $F_g : L^1(D, \mathbb{R}^m) \rightarrow [0, +\infty]$  with domain  $L^q(D, \mathbb{R}^m) \cap GSBVP(D, \mathbb{R}^m)$ , where it is defined by

$$F_g(u) = \int_D h(\nabla u) \, dx + \int_{S_u} \varphi(v_u) \, d\mathcal{H}^{d-1} + \gamma \int_D |u - g|^q \, dx$$

with the functions  $h$  and  $\varphi$  given by Theorem 3.5.

Moreover, for each  $\varepsilon > 0$  there exists a global minimizer  $\hat{u}_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  of the functional  $F_{\varepsilon,g}(\omega)$  and if  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  is any sequence such that

$$\lim_{\varepsilon \rightarrow 0} \left( F_{\varepsilon, g}(\omega)(u_\varepsilon) - \min_{u \in \mathcal{PC}_\varepsilon^\omega} F_{\varepsilon, g}(\omega)(u) \right) = 0,$$

then a.s. it is compact in  $L^q(D, \mathbb{R}^m)$  and each cluster point as  $\varepsilon \rightarrow 0$  is a global minimizer of  $F_g$ .

**Remark 3.9.** Up to a further subsequence, the statement of Theorem 3.8 remains valid in the non-stationary setting of Theorem 3.3. However, the  $\Gamma$ -limit contains a heterogeneous fidelity term of the form  $\int_D \gamma(x) |u - g|^q dx$  for some positive function  $\gamma \in L^\infty(D)$  such that also  $1/\gamma \in L^\infty(D)$ . In the case of a stationary, but non-ergodic group action the function  $\gamma$  does not depend on  $x$ , but might be random.

#### 4. Applications and the main result

We postpone the technical proofs of the results of Section 3 to the last two sections. First we use them to derive the  $\Gamma$ -limit of the functionals in (4) and prove our main result, that is, the approximation of the Mumford–Shah functional announced in the introduction.

##### 4.1. The $\Gamma$ -limit for forward differences on $\mathbb{Z}^2$

Our aim is to analyze the asymptotic behavior of the discrete functional proposed in [53]. It is based on the square lattice  $\mathcal{L} = \mathbb{Z}^2$  with edges that yield standard forward differences, that is,  $\mathcal{E}_{FD} = \{(x, x + e_i) : x \in \mathbb{Z}^2, i = 1, 2\}$ , where  $\{e_i\}_{i=1,2}$  denotes the standard basis of  $\mathbb{R}^2$ . The discrete approximation is then defined for functions  $u : \varepsilon \mathbb{Z}^2 \rightarrow \mathbb{R}^m$  and after rescaling and dropping the fidelity term it reads as

$$\tilde{F}_\varepsilon(u, A) = \sum_{\varepsilon x \in \varepsilon \mathbb{Z}^2 \cap A} \varepsilon \min\{\alpha \varepsilon^{-1} (|u(\varepsilon x + \varepsilon e_1) - u(\varepsilon x)|^2 + |u(\varepsilon x + \varepsilon e_2) - u(\varepsilon x)|^2), 1\}. \quad (23)$$

The set of edges satisfies (9) as the Voronoi neighbors are given by all  $x, y \in \mathbb{Z}^2$  such that  $|x - y| = 1$ . Moreover  $\tilde{F}_\varepsilon$  has the required structure to apply our results which can be seen by setting  $f(p) = \min\{\alpha \|p\|_1, 1\}$ . In [53] Cremers and Strelakovski conjectured that  $\tilde{F}_\varepsilon$  approximates the Mumford–Shah functional. With the results of Section 3 we can identify the  $\Gamma$ -limit of  $\tilde{F}_\varepsilon$ , which differs from the Mumford–Shah functional due to an anisotropic surface integral.

**Corollary 4.1.** *The functionals  $\tilde{F}_\varepsilon$  defined in (23)  $\Gamma$ -converge with respect to the  $L^1(D, \mathbb{R}^m)$ -topology to the functional  $F_0 : L^1(D, \mathbb{R}^m)$  with domain  $L^1(D, \mathbb{R}^m) \cap GSBV^2(D, \mathbb{R}^m)$ , where it is given by*

$$F_0(u) = \alpha \int_D |\nabla u|^2 + \int_{S_u} \varphi_0(v) d\mathcal{H}^1,$$

with the function  $\varphi_0 : \mathbb{R}^2 \rightarrow [0, +\infty)$  defined by

$$\varphi_0(v) = \begin{cases} |v_1| + |v_2| & \text{if } v_1 \cdot v_2 < 0, \\ \max\{|v_1|, |v_2|\} & \text{if } v_1 \cdot v_2 \geq 0. \end{cases}$$

**Proof.** As outlined above, we can apply Theorem 3.5 to the sequence  $\tilde{F}_\varepsilon$ . Moreover, from Theorem 3.3 we deduce that the function  $h$  in Theorem 3.5 is given by the density of the  $\Gamma$ -limit of the sequence  $E_\varepsilon$  defined in (17). With our choice of  $f$  and  $\mathcal{E}_{FD}$  the functional  $E_\varepsilon$  takes the form

$$E_\varepsilon(u) = \alpha \sum_{\substack{\varepsilon x, \varepsilon y \in \varepsilon \mathbb{Z}^2 \cap D \\ |x - y| = 1}} \frac{\varepsilon^2}{2} \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|^2.$$

In this case [4, Remark 5.3] yields that

$$\Gamma(L^p(D))\text{-}\lim_{\varepsilon \rightarrow 0} E_\varepsilon(u) = \alpha \int_D |\nabla u|^2 dx$$

for all  $u \in W^{1,2}(D, \mathbb{R}^m)$ , so that  $h(\xi) = \alpha|\xi|^2$  and it remains to identify the surface integrand  $\varphi$  in Theorem 3.5 as  $\varphi_0$ . By convexity, the function  $\varphi$  is continuous, so it suffices to treat the case  $v_1 \cdot v_2 \neq 0$ . Note that for forward differences the functional  $I_\varepsilon$  defined in (18) is given by

$$I_\varepsilon(u, A) = \frac{1}{2} \sum_{\varepsilon x \in \varepsilon \mathbb{Z}^2 \cap A} \varepsilon \max\{|u(\varepsilon x + \varepsilon e_i) - u(\varepsilon x)| : i \in \{1, 2\} \text{ with } \varepsilon x + \varepsilon e_i \in A\},$$

where  $u : \varepsilon \mathbb{Z}^2 \rightarrow \{\pm e_1\}$ . From Remarks 3.4 and 3.6 we know that  $\varphi$  is given by the formula

$$\varphi(v) = \lim_{\varepsilon \rightarrow 0} \left( \inf \left\{ I_\varepsilon(v, Q_v(0, 1)) : v \in \mathcal{PC}_{\varepsilon, 4\varepsilon}(u_{0,v}^{-e_1, e_1}, Q_v(0, 1)) \right\} \right). \quad (24)$$

Consider for fixed  $\varepsilon \ll 1$  any function  $v \in \mathcal{PC}_{\varepsilon, 4\varepsilon}(u_{0,v}^{-e_1, e_1}, Q_v(0, 1))$  with values in  $\{\pm e_1\}$ . We locally construct a function  $\tilde{v} \in BV(Q_v(0, 1), \{\pm e_1\})$  as follows: on  $Q_{e_1}(\varepsilon x, \varepsilon)$  with  $x \in \mathbb{Z}^2$  such that  $Q_{e_1}(\varepsilon x, \varepsilon) \subset Q_v(0, t)$  we set

$$\tilde{v}(y) = \begin{cases} v(\varepsilon x) & \text{if } \prod_{i=1}^2 |v(\varepsilon x + \varepsilon e_i) - v(\varepsilon x)| = 0, \\ v(\varepsilon x) & \text{if } \prod_{i=1}^2 |v(\varepsilon x + \varepsilon e_i) - v(\varepsilon x)| \neq 0 \text{ and } \langle y - \varepsilon x, e_1 + e_2 \rangle \leq 0, \\ v(\varepsilon x + \varepsilon e_1) & \text{if } \prod_{i=1}^2 |v(\varepsilon x + \varepsilon e_i) - v(\varepsilon x)| \neq 0 \text{ and } \langle y - \varepsilon x, e_1 + e_2 \rangle > 0, \end{cases}$$

while we define  $\tilde{v} = u_{0,v}^{-e_1, e_1}$  on all cubes  $Q_{e_1}(\varepsilon x, \varepsilon)$  with  $Q_{e_1}(\varepsilon x, \varepsilon) \cap \partial Q_v(0, 1) \neq \emptyset$ . The latter implies  $\tilde{v} = u_{0,v}^{-e_1, e_1}$  on  $\partial Q_v(0, 1)$  in the sense of traces. Moreover we modified the jump set away from the boundary in such a way that it contains diagonal lines of length  $\sqrt{2}$  instead of corners formed by the upper and the right hand side of a cube. Setting  $Q_\varepsilon := \{y \in Q_v(0, 1) : \text{dist}(y, \partial Q_v(0, 1)) > 2\varepsilon\}$ , the above construction and the boundary conditions for  $v$  imply

$$I_\varepsilon(v, Q_v(0, 1)) \geq \int_{S_{\tilde{v}} \cap Q_\varepsilon} \varphi_0(v_{\tilde{v}}) d\mathcal{H}^1 \geq \int_{S_{\tilde{v}} \cap Q_v(0, 1)} \varphi_0(v_{\tilde{v}}) d\mathcal{H}^1 - C\varepsilon. \quad (25)$$

Observe that  $\varphi_0$  is convex. Hence the functional on the right hand side is  $BV$ -elliptic in the sense that

$$\int_{S_{\tilde{v}} \cap Q_v(0, 1)} \varphi_0(v_{\tilde{v}}) d\mathcal{H}^1 \geq \varphi_0(v)$$

for all  $\tilde{v} \in BV(Q_v(0, 1), \{\pm e_1\})$  such that  $\tilde{v} = u_{0,v}^{-e_1, e_1}$  on  $\partial Q_v(0, 1)$  in the sense of traces (see [9] for more details). Since  $v \in \mathcal{PC}_{\varepsilon, 4\varepsilon}(u_{0,v}^{-e_1, e_1}, Q_v(0, 1))$  was arbitrary, we conclude from (24) and (25) that  $\varphi(v) \geq \varphi_0(v)$ .

In order to prove the reverse inequality, first note that

$$I_\varepsilon(u, A) \leq \frac{1}{4} \sum_{\substack{\varepsilon x, \varepsilon y \in \varepsilon \mathbb{Z}^2 \cap A \\ |x-y|=1}} \varepsilon |u(\varepsilon x) - u(\varepsilon y)|. \quad (26)$$

The  $\Gamma$ -limit of the right hand side of (26) is well-known. It is finite only on  $BV(D, \{\pm e_1\})$  and given by  $\int_{S_u} |v|_1 d\mathcal{H}^1$  (see [3]). By comparison we obtain  $\varphi(v) \leq |v_1| + |v_2|$ . This finishes the proof in the case  $v_1 \cdot v_2 < 0$ . If  $v_1 \cdot v_2 > 0$ , denote by  $i_0$  the index such that  $|v_{i_0}| = \max\{|v_1|, |v_2|\}$  and set  $i_1 = \{1, 2\} \setminus \{i_0\}$ . We define a candidate for the minimum problem in (24) setting  $u_v(\varepsilon x) = u_{0,v}^{-e_1, e_1}(\varepsilon x)$  for all  $x \in \mathbb{Z}^2$ . By definition it satisfies the correct boundary conditions. A straightforward analysis shows that for any  $x \in \mathbb{Z}^2$  with  $u_v(\varepsilon x) \neq u_v(\varepsilon x + \varepsilon e_{i_1})$  we have  $u_v(\varepsilon x) \neq u_v(\varepsilon x + \varepsilon e_{i_0})$ . Thus it suffices to count just the interactions along the direction  $e_{i_0}$ . Those can be bounded by  $\varepsilon^{-1}|v_{i_0}| + C$ , so that

$$I_\varepsilon(u_v, Q_v(0, 1)) \leq |v_{i_0}| + C\varepsilon = \max\{|v_1|, |v_2|\} + C\varepsilon.$$

From (24) we conclude that  $\varphi(v) \leq \varphi_0(v)$  which finishes the proof.  $\square$

**Remark 4.2.** For the  $d$ -dimensional version of  $\tilde{F}_\varepsilon$  (defined in the introduction), one still has the existence of the  $\Gamma$ -limit with an anisotropic surface integrand. To see the latter, one first shows that  $\varphi(e_1) = 1$ . Then for the vector  $v_0 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$  the discretization of  $u_{0,v_0}^{-e_1, e_1}$  yields an upper bound  $\varphi(v_0) \leq \frac{1}{\sqrt{2}}$  (actually equality holds). The precise  $\Gamma$ -limit in higher dimensions is beyond the scope of this paper.



#### 4.2. Approximations of the Mumford–Shah functional

Now we use our general results to provide a discretization of the vector-valued Mumford–Shah functional. To this end, we need to take our parameters  $p = q = 2$ . However, it might be of interest to obtain also other exponents for the fidelity term and therefore we consider the general case  $q > 1$  and just fix  $p = 2$  to focus on the isotropy issue. We suggest to take as stochastic lattice the so-called *random parking process*. For the precise geometric construction of this process by suitably choosing projected points of a homogeneous Poisson point process in dimension  $d + 1$ , we refer the reader to the two papers [51,44]. Here we recall that the random parking process defines a stochastic lattice  $\mathcal{L}_{RP}$  that is admissible, stationary, ergodic and, most important for our applications, isotropic in the sense of Definition 2.5. Moreover, we can choose for instance  $\mathcal{E}(\omega) = \mathcal{N}(\omega)$  to obtain stationary and isotropic edges (see also Remark 2.7 for other possible choices). We prove our result for general stochastic lattices satisfying these assumptions. Note that the following theorem covers in particular the two functionals presented in the introduction.

**Theorem 4.3.** *Fix  $p = 2$  and let  $q \in (1, +\infty)$  and  $g_\varepsilon(\omega)$  satisfy (21). Assume that  $\mathcal{L}$  is an admissible stochastic lattice that is stationary, ergodic and isotropic with admissible stationary and isotropic edges. Then there exist constants  $c_1, c_2, c_3 > 0$  such that  $\mathbb{P}$ -a.s. the functionals  $F_{\varepsilon,g}(\omega)$  defined in (22)  $\Gamma$ -converge with respect to the  $L^1(D, \mathbb{R}^m)$ -topology to the functional  $F_g : L^1(D, \mathbb{R}^m) \rightarrow [0, +\infty]$  with domain  $L^q(D, \mathbb{R}^m) \cap GSBV^2(D, \mathbb{R}^m)$ , where it is defined by*

$$F_g(u) = c_1 \int_D |\nabla u|^2 dx + c_2 \mathcal{H}^{d-1}(S_u) + c_3 \int_D |u - g|^q dx.$$

**Remark 4.4.** In the scalar case  $m = 1$ , the statement of Theorem 4.3 is valid for every  $p > 1$ . Indeed, we already know that the function  $h$  has to be  $p$ -homogeneous. Following the proof below, stochastic isotropy implies that it is constant on  $S^{d-1}$ . Hence  $h(\xi) = c_1 |\xi|^p$  for some  $c_1 > 0$ . The formula for the surface term does not depend on  $p$ .

**Remark 4.5.** Theorem 4.3 and the convergence of minimizers in Theorem 3.8 yield the full discretization of the Mumford–Shah functional. In practice it is of course impossible to create the stochastic lattice on the whole space but one has to take a finite particle approximation. Moreover the minimization of the discrete functionals  $F_\varepsilon(\omega)$  is still nontrivial due to non-convexity. However, first numerical tests have shown promising results and we plan to further investigate our approach in the future.

**Proof of Theorem 4.3.** Due Theorem 3.8 it only remains to show that  $h(\xi) = c_1 |\xi|^2$  and  $\varphi(v) = c_2$  for some constants  $c_1, c_2 > 0$ . By Theorem 3.3 the function  $h$  is also the density of the  $\Gamma$ -limit of the functionals  $E_\varepsilon(\omega)$  defined in (17). For  $p = 2$  these are non-negative quadratic forms, so we deduce from [35, Theorem 11.1] that  $h$  is a non-negative quadratic form, too. We write it explicitly as

$$h(\xi) = \sum_{i,k=1}^m \sum_{j,l=1}^d h_{ijkl} \xi_{ij} \xi_{kl},$$

where the coefficients satisfy the symmetry condition  $h_{ijkl} = h_{klij}$ . Since the discrete functional is invariant under orthogonal transformations  $u \mapsto Qu$  it holds that  $h(Q\xi) = h(\xi)$  for all  $\xi \in \mathbb{R}^{m \times d}$  and  $Q \in O(m)$ . Moreover, reasoning exactly as for the case  $m = d$  treated in [5, Theorem 9] one can further show that ergodicity and isotropy imply  $h(\xi R) = h(\xi)$  for all  $\xi \in \mathbb{R}^{m \times d}$  and all  $R \in SO(d)$ . We argue that  $h$  depends only on the singular values. To this end, we fix a matrix  $\xi \in \mathbb{R}^{m \times d}$  and consider any singular value decomposition  $\xi = Q\Sigma V^T$  with orthogonal matrices  $Q \in O(m)$  and  $V \in O(d)$  and a diagonal matrix  $\Sigma \in \mathbb{R}^{m \times d}$ . If  $V \in SO(d)$  then  $h(\xi) = h(\Sigma)$ . Otherwise we replace  $V$  by a rotation observing that

$$A = QP_m^{1,2}P_m^{1,2}\Sigma P_d^{1,2}P_d^{1,2}V^T,$$

where  $P_n^{1,2}$  denotes the  $n \times n$ -matrix which differs from the identity by exchanging the first and the second column. In this case  $h(\xi) = h(P_m^{1,2}\Sigma P_d^{1,2})$  since the matrix  $P_d^{1,2}V^T$  belongs to  $SO(d)$ . Set  $l = \min\{d, m\}$  and write the singular

values as  $\lambda(\xi) \in \mathbb{R}^l$  with non-negative coefficients. We conclude that there exists a permutation  $P(\xi) \in \{I, P_l^{1,2}\}$  such that

$$h(\xi) = \sum_{i,k=1}^l h_{iikk} (P(\xi)\lambda(\xi))_i (P(\xi)\lambda(\xi))_k.$$

Thus it is enough to characterize the coefficients  $h_{iikk}$ . We will test several diagonal matrices  $\xi$ . To simplify notation, given  $v \in \mathbb{R}^l$  we denote by  $\text{diag}(v) \in \mathbb{R}^{m \times d}$  the diagonal matrix with elements  $v$ . As a first step, note that we can find  $Q \in O(m)$  and  $R \in SO(d)$  such that  $\text{diag}(e_i) = Q \text{diag}(-e_j)R$ . Thus by invariance  $h_{iiii} = h_{jjjj}$  for all  $i, j = 1, \dots, l$ . Now consider  $i \neq k$ . We argue that  $h_{iikk} = 0$ . To this end, we use that there exists a matrix  $Q \in O(m)$  such that  $\text{diag}(e_i + e_k) = Q \text{diag}(e_i - e_k)$ , which yields again by invariance that

$$h_{iiii} + h_{kkkk} + h_{iikk} + h_{kkii} = h_{iiii} + h_{kkkk} - h_{iikk} - h_{kkii}.$$

Due the symmetry condition on the coefficients of  $h$  we obtain  $h_{iikk} = 0$ . Setting  $c_1 = h_{1111}$  we have shown that

$$h(\xi) = c_1 \sum_{i=1}^l \lambda_i(\xi)^2 = c_1 |\xi|^2.$$

It follows from Theorem 3.1 that  $c_1 > 0$ .

We now turn to the surface integrand  $\varphi$  and prove that  $\varphi(Rv) = \varphi(v)$  for all  $R \in SO(d)$ . Since  $\varphi$  is deterministic by ergodicity, we can take expectations in the asymptotic formula given by Remark 3.4 and simplified via Remark 3.6. Since  $\tau'_R$  is measure preserving, dominated convergence and a change of variables yield

$$\begin{aligned} \varphi(Rv) &= \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \inf\{I_{\varepsilon}(\omega)(v, Q_{Rv}(0, 1)) : v \in \mathcal{PC}_{\varepsilon, 2M\varepsilon}^{\omega}(u_{0, Rv}^{-e_1, e_1}, Q_{Rv}(0, 1))\} d\mathbb{P}(\omega) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \inf\{I_{\varepsilon}(\tau'_{R^T}\omega)(v, Q_v(0, 1)) : v \in \mathcal{PC}_{\varepsilon, 2M\varepsilon}^{\tau'_{R^T}\omega}(u_{0, v}^{-e_1, e_1}, Q_v(0, 1))\} d\mathbb{P}(\omega) \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \inf\{I_{\varepsilon}(\omega')(v, Q_v(0, 1)) : v \in \mathcal{PC}_{\varepsilon, 2M\varepsilon}^{\omega'}(u_{0, v}^{-e_1, e_1}, Q_v(0, 1))\} d\mathbb{P}(\omega') = \varphi(v), \end{aligned}$$

where we used from the first to the second line that by isotropy of  $\mathcal{L}$  and  $\mathcal{E}$  the discrete functional in (18) satisfies  $I_{\varepsilon}(\tau'_{R^T}\omega)(u, Q_{Rv}(0, 1)) = I_{\varepsilon}(\omega)(u \circ R, Q_v(0, 1))$  for every  $R \in SO(d)$ . We finish the proof setting  $c_2 = \varphi(e_1)$  since Theorem 3.2 implies  $c_2 > 0$ .  $\square$

### 5. Separation of scales: proof of Theorem 3.3

This section is devoted to the proof of Theorem 3.3, which will constitute the most involved part of the paper. In a first part we use [18, Theorem 1] to represent (up to subsequences) the  $\Gamma$ -limit on  $SBV^p$  as an integral functional. In a second and third part we study asymptotic formulas for the integrands of that representation which allow to conclude the proof. Several times we will need the following property of Voronoi neighbors in a stochastic lattice.

**Lemma 5.1.** *Let  $\mathcal{L}(\omega)$  be an admissible set of points with constants  $R > r > 0$  as in Definition 2.2. Then for all  $x, y \in \mathcal{L}(\omega)$  there exists a path  $P(x, y) = \{x_0 = x, x_1, \dots, x_n = y\}$  such that  $(x_{i-1}, x_i) \in \mathcal{N}(\omega)$  for all  $1 \leq i \leq n$  and*

$$P(x, y) \subset \text{co}(x, y) + B_{2R}(0).$$

*In particular, there exists a constant  $C_{r,R} < +\infty$  such that  $\#P(x, y) \leq C_{r,R}|x - y|$ .*

**Proof.** For  $0 < \delta \ll 1$  consider the collection of segments

$$G_{\delta}(x, y) = \{z + \lambda(y - x) : z \in B_{\delta}(x), 0 \leq \lambda \leq 1\}.$$

We argue that there exists a segment  $s^* = \{z_0 + \lambda(y - x) : 0 \leq \lambda \leq 1\} \subset G_\delta(x, y)$  that does not intersect any Voronoi facet of the tessellation  $\mathcal{V}(\omega)$  of dimension less than  $d - 1$ . Indeed, assume by contradiction that the claim is false for all  $z_0 \in B_\delta(x)$ . By Remark 2.3 we find finitely many Voronoi facets of dimension less than  $d - 1$  whose projections onto the hyperplane containing  $x$  and orthogonal to  $y - x$  cover a  $d - 1$ -dimensional set. Since projections onto hyperplanes are Lipschitz continuous, we obtain a contradiction. Hence, for  $\delta$  small enough, we can construct a path of Voronoi neighbors connecting  $x, y$  by suitably numbering the set

$$P(x, y) = \{z \in \mathcal{L}(\omega) : \mathcal{C}(z) \cap s_* \neq \emptyset\},$$

which satisfies  $P(x, y) \subset \text{co}(x, y) + B_{R+\delta}(0)$  as claimed. Combining Remark 2.3 with a covering argument we obtain in addition

$$\#P(x, y) \leq |B_{\frac{r}{2}}(0)|^{-1}(|x - y| + 2R)(2R)^{d-1} \leq |B_{\frac{r}{2}}(0)|^{-1} \frac{2(2R)^d}{r} |x - y| =: C_{r,R} |x - y|. \quad \square$$

### 5.1. Integral representation on $SBV^p$

We are going to prove the following intermediate result:

**Proposition 5.2.** *Given any sequence  $\varepsilon \rightarrow 0$  there exists a subsequence  $\varepsilon_n$  such that for all  $A \in \mathcal{A}^R(D)$  the functionals  $F_{\varepsilon_n}(\omega)(\cdot, A)$   $\Gamma$ -converge in the  $L^1(D, \mathbb{R}^m)$ -topology to a functional  $F(\omega)(\cdot, A) : L^1(D, \mathbb{R}^m) \rightarrow [0, +\infty]$ . If  $u \in SBV^p(A, \mathbb{R}^m)$  then  $F(\omega)(u, A)$  can be written as*

$$F(\omega)(u, A) = \int_A h(x, \nabla u) \, dx + \int_{S_u \cap A} \varphi(x, u^+ - u^-, \nu) \, d\mathcal{H}^{d-1},$$

where, for  $x_0 \in D$ ,  $\nu \in S^1$ ,  $a \in \mathbb{R}^m$  and  $\xi \in \mathbb{R}^{d \times m}$ , the integrands are given by

$$h(x_0, \xi) = \limsup_{\varrho \rightarrow 0} \frac{1}{\varrho^d} m(\omega)(\xi(\cdot - x_0), Q_\nu(x_0, \varrho)),$$

$$\varphi(x_0, a, \nu) = \limsup_{\varrho \rightarrow 0} \frac{1}{\varrho^{d-1}} m(\omega)(u_{x_0, \nu}^{a, 0}, Q_\nu(x_0, \varrho))$$

with the function  $u_{x_0, \nu}^{a, 0}$  defined in (7) and the function  $m(\omega)(\bar{u}, A)$  defined for any  $\bar{u} \in SBV^p(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}^R(D)$  by

$$m(\omega)(\bar{u}, A) = \inf\{F(\omega)(v, A) : v \in SBV^p(A, \mathbb{R}^m), v = \bar{u} \text{ in a neighborhood of } \partial A\}.$$

In order to prove this result we will analyze the localized  $\Gamma$ -lim inf and  $\Gamma$ -lim sup  $F'(\omega), F''(\omega) : L^1(D, \mathbb{R}^m) \times \mathcal{A}(D) \rightarrow [0, +\infty]$  of the functionals  $F_\varepsilon(\omega)$ , which are defined by

$$F'(\omega)(u, A) = \inf_{\varepsilon \rightarrow 0} \{\liminf F_\varepsilon(\omega)(u_\varepsilon, A) : u_\varepsilon \rightarrow u \text{ in } L^1(D, \mathbb{R}^m)\},$$

$$F''(\omega)(u, A) = \inf_{\varepsilon \rightarrow 0} \{\limsup F_\varepsilon(\omega)(u_\varepsilon, A) : u_\varepsilon \rightarrow u \text{ in } L^1(D, \mathbb{R}^m)\}.$$

**Remark 5.3.** It is well-known that both functionals are  $L^1(D, \mathbb{R}^m)$ -lower semicontinuous. Moreover, note that for any  $u \in L^1(D, \mathbb{R}^m)$  there exists indeed a sequence  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  such that  $u_\varepsilon \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$ . For  $u \in C_c(D, \mathbb{R}^m)$  this follows from Remark 2.3. In the general case one can use a density argument and construct suitable diagonal sequences.

Our aim is to apply the integral representation of [18, Theorem 1] to a slightly modified functional. To this end, below we establish several properties of  $F'(\omega)$  and  $F''(\omega)$  which are necessary in the context of integral representation. However, at first we prove a truncation lemma that allows to reduce many arguments to the case of bounded functions.

**Lemma 5.4.** Let  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$ . For any  $k > 0$  set  $T_k u_\varepsilon$  as in Lemma 2.1. Then, for any  $A \in \mathcal{A}(D)$ , it holds that  $F_\varepsilon(\omega)(T_k u_\varepsilon, A) \leq F_\varepsilon(\omega)(u_\varepsilon, A)$ . In particular, whenever  $u \in L^\infty(D, \mathbb{R}^m)$  we can compute  $F'(\omega)(u, A)$  and  $F''(\omega)(u, A)$  considering sequences  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  such that  $|u_\varepsilon(\varepsilon x)| \leq 3\|u\|_\infty$  for all  $x \in \mathcal{L}(\omega)$ . Moreover, for all  $u \in L^1(D, \mathbb{R}^m)$  we have

$$\begin{aligned} \lim_{k \rightarrow +\infty} F'(\omega)(T_k u, A) &= F'(\omega)(u, A), \\ \lim_{k \rightarrow +\infty} F''(\omega)(T_k u, A) &= F''(\omega)(u, A). \end{aligned}$$

**Proof.** For the estimate at the discrete level, it suffices to combine the fact that  $|T_k u_\varepsilon(\varepsilon x) - T_k u_\varepsilon(\varepsilon y)| \leq |u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)|$  for all  $x, y \in \mathcal{L}(\omega)$  with the monotonicity assumption (12). In order to restrict the class of approximating sequences, we use the first estimate and the fact that any truncated sequence  $T_k u_\varepsilon$  with  $k = \|u\|_\infty$  still converges to  $u$  in  $L^1(D, \mathbb{R}^m)$ . The continuity property at the limit follows from  $L^1(D, \mathbb{R}^m)$ -lower semicontinuity of both functionals and the fact that the discrete upper bound is conserved in the limit.  $\square$

We next show that  $F''(\omega)$  is local.

**Lemma 5.5** (Locality of  $F''$ ). Let  $A \in \mathcal{A}^R(D)$ . If  $u, v \in L^1(D, \mathbb{R}^m)$  and  $u = v$  a.e. on  $A$ , then  $F''(\omega)(u, A) = F''(\omega)(v, A)$ .

**Proof.** Due to Remark 5.3 there exist sequences  $u_\varepsilon, v_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  converging to  $u$  and  $v$  in  $L^1(D, \mathbb{R}^m)$ , respectively, and such that

$$\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A) = F''(\omega)(u, A), \quad \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(v_\varepsilon, A) = F''(\omega)(v, A).$$

Define  $\tilde{u}_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  by

$$\tilde{u}_\varepsilon(\varepsilon x) = \mathbb{1}_A(\varepsilon x) u_\varepsilon(\varepsilon x) + (1 - \mathbb{1}_A(\varepsilon x)) v_\varepsilon(\varepsilon x).$$

Since  $|\partial A| = 0$  and  $u_\varepsilon$  and  $v_\varepsilon$  are equiintegrable, it follows that  $\tilde{u}_\varepsilon \rightarrow v$  in  $L^1(D, \mathbb{R}^m)$ . Then by definition

$$F''(\omega)(v, A) \leq \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(\tilde{u}_\varepsilon, A) = \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A) = F''(\omega)(u, A).$$

Exchanging the roles of  $u$  and  $v$  concludes the proof.  $\square$

The following two lemmata provide a lower bound for  $F'$  and an upper bound for  $F''$ . Together with the lower bound we obtain an equicoercivity property under an additional equiintegrability assumption.

**Lemma 5.6** (Compactness and lower bound). Assume that  $A \in \mathcal{A}^R(D)$  and  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  are such that

$$\sup_{\varepsilon > 0} F_\varepsilon(\omega)(u_\varepsilon, A) < +\infty.$$

If  $u_\varepsilon$  is equiintegrable on  $A$ , then there exists a subsequence (not relabeled) such that  $u_\varepsilon \rightarrow u$  in  $L^1(A, \mathbb{R}^m)$  for some  $u \in L^1(A, \mathbb{R}^m) \cap GSBV^p(A, \mathbb{R}^m)$ . Moreover we have the estimate

$$\frac{1}{c} \left( \int_A |\nabla u|^p dx + \mathcal{H}^{d-1}(S_u \cap A) \right) \leq F'(\omega)(u, A)$$

for some constant  $c > 0$  independent of  $\omega, A$  and  $u$ .

**Proof.** We first construct a suitable function  $v_\varepsilon \in SBV^p(A, \mathbb{R}^m)$  that is asymptotically close to  $u_\varepsilon$ . Given a triangulation  $\mathcal{T}^d$  of the cube  $[0, 1]^d$  we construct a periodic triangulation of  $\mathbb{R}^d$  via

$$\mathcal{T} = \{T = z + T_d : z \in \mathbb{Z}^d, T_d \in \mathcal{T}^d\}.$$

We may assume that  $\text{diam}(T) < R$  for all  $T \in \mathcal{T}$ . Define  $u_\varepsilon^{\text{aff}}$  as a continuous piecewise affine interpolation of  $u_\varepsilon$  on  $\varepsilon\mathcal{T}$  as follows: for each  $z \in \mathbb{Z}^d$  we choose a point  $x(z) \in \mathcal{L}(\omega)$  such that  $z \in \mathcal{C}(x)$  and set

$$u_\varepsilon^{\text{aff}}(\varepsilon z) = u_\varepsilon(\varepsilon x(z)).$$

Next we decompose the scaled lattice as  $\varepsilon\mathcal{L}(\omega) = L_{0,\varepsilon} \cup L_{1,\varepsilon}$ , where

$$L_{0,\varepsilon} = \{\varepsilon x \in \varepsilon\mathcal{L}(\omega) : \|\varepsilon|\nabla_{\omega,\varepsilon}(u, A)|^p(\varepsilon x)\|_1 \leq 1\}.$$

Let us also group the simplices overlapping with  $A$  according to

$$\begin{aligned} \mathcal{T}_1 &= \{T \in \mathcal{T} : T \cap A \neq \emptyset, \inf_{z \in \varepsilon T} \text{dist}(z, \partial A) \leq 8R\varepsilon\}, \\ \mathcal{T}_2 &= \{T \in \mathcal{T} : T \cap A \neq \emptyset, \inf_{z \in \varepsilon T} \text{dist}(z, \partial A) > 8R\varepsilon \text{ and } \inf_{z \in \varepsilon T} \text{dist}(z, L_{1,\varepsilon}) \leq 6R\varepsilon\}, \\ \mathcal{T}_3 &= \{T \in \mathcal{T} : T \cap A \neq \emptyset, \inf_{z \in \varepsilon T} \text{dist}(z, \partial A) > 8R\varepsilon \text{ and } \inf_{z \in \varepsilon T} \text{dist}(z, L_{1,\varepsilon}) > 6R\varepsilon\}. \end{aligned}$$

Given  $u_\varepsilon^{\text{aff}}$  we define  $v_\varepsilon$  on the interior of each simplex  $\varepsilon T \in \varepsilon\mathcal{T}$  setting

$$v_{\varepsilon|\varepsilon T} = \begin{cases} u_{\varepsilon|\varepsilon T}^{\text{aff}} & \text{if } T \in \mathcal{T}_3, \\ 0 & \text{otherwise.} \end{cases}$$

Note that  $v_\varepsilon \in SBVP(A, \mathbb{R}^m)$ . We start estimating the difference of  $u_\varepsilon$  and  $v_\varepsilon$  on  $A$ . Consider any simplex  $\varepsilon T$  with  $T \in \mathcal{T}_1$ . Then  $\varepsilon T \subset \partial A + B_{9R\varepsilon}(0)$ . Since  $\partial A$  is a Lipschitz boundary it admits a  $(d-1)$ -dimensional Minkowski content. Hence there exists a constant  $C = C_R > 0$  such that for  $\varepsilon$  small enough we have

$$\left| \bigcup_{T \in \mathcal{T}_1} \varepsilon T \cap A \right| \leq |\{z \in A : \text{dist}(z, \partial A) \leq 9R\varepsilon\}| \leq C\mathcal{H}^{d-1}(\partial A)\varepsilon. \quad (27)$$

Next, if  $T \in \mathcal{T}_2$ , then there exists  $\varepsilon x \in L_{1,\varepsilon} \cap A$  such that  $\varepsilon T \subset B_{7R\varepsilon}(\varepsilon x)$ . From (14) and the definition of  $L_{1,\varepsilon}$  we deduce

$$\left| \bigcup_{T \in \mathcal{T}_2} \varepsilon T \cap A \right| \leq C\varepsilon^d \#\{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap A : \|\varepsilon|\nabla_{\omega,\varepsilon}(u, A)|^p(\varepsilon x)\|_1 > 1\} \leq C\varepsilon F_\varepsilon(\omega)(u_\varepsilon, A) \leq C\varepsilon. \quad (28)$$

Finally, if  $z \in \varepsilon T \cap A$  for some  $T \in \mathcal{T}_3$ , then by definition  $\text{dist}(z, L_{1,\varepsilon}) > 6R\varepsilon$  and  $\text{dist}(z, \partial A) > 8R\varepsilon$ . Let us write  $T = \text{co}(z^0, \dots, z^d) \in \mathcal{T}$  and  $z = \sum_i \lambda_i^z z^i$ . Choosing  $x \in \mathcal{L}(\omega)$  such that  $z \in \varepsilon(\mathcal{C}(x) \cap T)$ , Remark 2.3 yields

$$|v_\varepsilon(z) - u_\varepsilon(z)| \leq \sum_{i=0}^d \lambda_i^z |u_\varepsilon^{\text{aff}}(\varepsilon z^i) - u_\varepsilon(\varepsilon x)| \leq \sum_{y \in \mathcal{L}(\omega) \cap B_{3R}(x)} |u_\varepsilon(\varepsilon y) - u_\varepsilon(\varepsilon x)| \quad (29)$$

except for a null set where  $u_\varepsilon$  is not well-defined. Given  $y \in \mathcal{L}(\omega) \cap B_{3R}(x)$ , we let  $P(y, x) = \{y = x_0, x_1, \dots, x_n = x\}$  be the path of Voronoi neighbors given by Lemma 5.1. Since  $|x - y| \leq 3R$ , we have

$$|x_i - \varepsilon^{-1}z| \leq \text{dist}(x_i, \text{co}(x, y)) + |y - x| + |x - \varepsilon^{-1}z| < 2R + 3R + R = 6R.$$

In particular we conclude that  $\varepsilon x_i \in L_{0,\varepsilon} \cap A$  for all  $0 \leq i \leq n$ . Using (9), the definition of the set  $L_{0,\varepsilon}$  implies that  $|u_\varepsilon(\varepsilon x_i) - u_\varepsilon(\varepsilon x_{i+1})| \leq \varepsilon^{1-\frac{1}{p}}$  for all  $0 \leq i \leq n-1$ . By Remark 2.3 the number of Voronoi neighbors in  $B_{6R}(\varepsilon^{-1}z)$  is uniformly bounded, so that from (29) we infer the bound

$$|v_\varepsilon(z) - u_\varepsilon(z)| \leq C\varepsilon^{1-\frac{1}{p}}. \quad (30)$$

Since we have set  $v_\varepsilon = 0$  on all  $\varepsilon T \in \varepsilon(\mathcal{T}_1 \cup \mathcal{T}_2)$  and  $u_\varepsilon$  is equiintegrable by assumption, we conclude from (27), (28) and (30) that

$$\lim_{\varepsilon \rightarrow 0} \|v_\varepsilon - u_\varepsilon\|_{L^1(A)} = 0. \quad (31)$$

Moreover, the sequence  $v_\varepsilon$  is still equiintegrable on  $A$ . We will show the convergence for the sequence  $v_\varepsilon$ . We start estimating the size of its jump set. We have at most two contributions. The first one comes from discontinuities along edges of simplices in  $\mathcal{T}_1$ . For those we have, again for  $\varepsilon$  small enough, the estimate

$$\mathcal{H}^{d-1}\left(\bigcup_{T \in \mathcal{T}_1} \partial \varepsilon T \cap A\right) \leq C \varepsilon^{-1} |\{z \in \mathbb{R}^d : \text{dist}(z, \partial A) \leq 9R\varepsilon\}| \leq C \mathcal{H}^{d-1}(\partial A),$$

where we used the Lipschitz regularity of  $\partial A$  and a reverse isoperimetric inequality for the finitely many simplices in  $\mathcal{T}^d$ . The other contribution is given by

$$\mathcal{H}^{d-1}\left(\bigcup_{T \in \mathcal{T}_2} \partial \varepsilon T \cap A\right) \leq \sum_{T \in \mathcal{T}_2} C \varepsilon^{d-1} \leq C F_\varepsilon(\omega)(u_\varepsilon, A) \leq C,$$

where the second inequality follows by the same reasoning used in the lines preceding (28). The last two estimates imply that

$$\mathcal{H}^{d-1}(S_{v_\varepsilon} \cap A) \leq \mathcal{H}^{d-1}\left(\bigcup_{T \in \mathcal{T}_1 \cup \mathcal{T}_2} \partial \varepsilon T \cap A\right) \leq C. \quad (32)$$

To estimate the gradient it suffices to consider simplices  $T = \text{co}(z^0, \dots, z^d) \in \mathcal{T}_3$ . Write any basis vector  $e_k$  as  $e_k = \sum_{i=1}^d \lambda_i^k (z^i - z^0)$ . Due to the periodicity of the triangulation  $\mathcal{T}$  the coefficients  $\lambda_i^k$  are equibounded with respect to the simplices. Take  $x \in \mathcal{L}(\omega)$  such that  $u_\varepsilon^{\text{aff}}(\varepsilon z^0) = u_\varepsilon(\varepsilon x)$ . Since  $v_\varepsilon$  is affine on  $\varepsilon T$  we have

$$\begin{aligned} |\partial_k v_{\varepsilon|_{\varepsilon T}}| &= \left| \sum_{i=1}^d \lambda_i^k \frac{v_\varepsilon(\varepsilon z^i) - v_\varepsilon(\varepsilon z^0)}{\varepsilon} \right| \leq C \sum_{i=0}^d \left| \frac{u_\varepsilon^{\text{aff}}(\varepsilon z^i) - u_\varepsilon^{\text{aff}}(\varepsilon z^0)}{\varepsilon} \right| \\ &\leq C \sum_{y \in \mathcal{L}(\omega) \cap B_{3R}(x)} \varepsilon^{-1} |u_\varepsilon(\varepsilon y) - u_\varepsilon(\varepsilon x)| \leq C \sum_{\substack{x_i, x_j \in B_{5R}(x) \\ (x_i, x_j) \in \mathcal{N}(\omega)}} \varepsilon^{-1} |u_\varepsilon(\varepsilon x_i) - u_\varepsilon(\varepsilon x_j)|, \end{aligned}$$

where the last inequality follows by the same reasoning we used for proving (30). Now observe that if  $T \in \mathcal{T}_3$  and  $|x_i - x| < 5R$ , then  $\varepsilon x_i \in L_{0,\varepsilon} \cap A$ . Thus taking the  $p$ -th power of the above estimate and using (9) and (14) we obtain

$$\begin{aligned} |\nabla v_{\varepsilon|_{\varepsilon T}}|^p &\leq C \sum_{\substack{x_i, x_j \in B_{5R}(x) \\ (x_i, x_j) \in \mathcal{N}(\omega)}} \varepsilon^{-p} |u_\varepsilon(\varepsilon x_i) - u_\varepsilon(\varepsilon x_j)|^p \leq C \varepsilon^{-1} \sum_{x_i \in \mathcal{L}(\omega) \cap B_{5R}(x)} \|\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p(\varepsilon x_i)\|_1 \\ &\leq C \varepsilon^{-1} \sum_{x_i \in \mathcal{L}(\omega) \cap B_{5R}(x)} f(\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p(\varepsilon x_i)). \end{aligned}$$

We sum the last estimate over all  $T \in \mathcal{T}_3$ . Since  $T \subset B_{7R}(x_i)$  we count each lattice point  $x_i$  at most  $C$  times and thus conclude

$$\int_A |\nabla v_\varepsilon(z)|^p dz \leq C \sum_{\varepsilon x \in \mathcal{L}(\omega) \cap A} \varepsilon^{d-1} f(\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p(\varepsilon x)) = C F_\varepsilon(\omega)(u_\varepsilon, A) \leq C. \quad (33)$$

By (32) and (33), the compactness theorem for  $GSBV(A, \mathbb{R}^m)$ -functions [7, Theorem 2.2] implies that, up to subsequences,  $v_\varepsilon \rightarrow u \in GSBV(A, \mathbb{R}^m)$  in measure and by equiintegrability also in  $L^1(A, \mathbb{R}^m)$ . Moreover  $\nabla v_\varepsilon \rightharpoonup \nabla u$  in  $L^p(A, \mathbb{R}^{m \times d})$  and from lower semicontinuity [7, Theorem 3.7] we deduce

$$\int_A |\nabla u|^p dz + \mathcal{H}^{d-1}(S_u \cap A) \leq C \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A) + C \mathcal{H}^{d-1}(\partial A) \leq C.$$

Thus by definition  $u \in GSBV^p(A)$  which finishes the proof of compactness. In order to prove the lower bound, note that the argument above shows that for any open set  $A' \subset \subset A$  it holds that

$$\int_{A'} |\nabla u|^p dz + \mathcal{H}^{d-1}(S_u \cap A') \leq C \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A),$$

provided the right hand side is finite and  $u_\varepsilon \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$ . By the definition of  $F'(\omega)$  and the arbitrariness of  $A'$  we obtain the desired estimate.  $\square$

As a next step we estimate  $F''(\omega)$  from above.

**Lemma 5.7** (Upper bound). *Let  $u \in L^1(D, \mathbb{R}^m)$ . There exists a constant  $c > 0$  independent of  $\omega$  and  $u$  such that for all  $A \in \mathcal{A}^R(D)$  with  $u \in GSBV^p(A, \mathbb{R}^m)$  it holds that*

$$F''(\omega)(u, A) \leq c \left( \int_A |\nabla u|^p dx + \mathcal{H}^{d-1}(S_u \cap A) \right).$$

**Proof.** Take any ball  $B_L$  such that  $D \subset\subset B_L$ . For the moment let us assume that  $u \in SBV^p(B_L, \mathbb{R}^m)$  is such that

- (i)  $\mathcal{H}^{d-1}(\overline{S_u} \setminus S_u \cap B_L) = 0$ ,
- (ii)  $\overline{S_u}$  is the intersection of  $B_L$  with a finite number of pairwise disjoint  $(d-1)$ -simplices,
- (iii)  $u \in W^{k,\infty}(B_L \setminus \overline{S_u}, \mathbb{R}^m)$  for all  $k \in \mathbb{N}$ .

We define an admissible sequence to bound  $F''(\omega)(u|_D, A)$  setting

$$u_\varepsilon(\varepsilon x) = \begin{cases} u(\varepsilon x) & \text{if } \varepsilon x \in B_L \setminus \overline{S_u}, \\ 0 & \text{otherwise.} \end{cases}$$

Using the properties (ii) and (iii) from above it follows by Remark 2.3 that  $u_\varepsilon \rightarrow u|_D$  in  $L^1(D, \mathbb{R}^m)$ . To bound the energy, consider first the case that  $\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap A$  is such that  $\text{dist}(\varepsilon x, \overline{S_u}) \geq 3M\varepsilon$ . Then for all  $y \in \mathcal{L}(\omega)$  with  $(x, y) \in \mathcal{E}(\omega)$  we have by Jensen's inequality and the regularity of  $u$  that

$$\left| \frac{u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)}{\varepsilon} \right|^p = \left| \int_0^1 \nabla u(\varepsilon x + s\varepsilon(y-x))(y-x) ds \right|^p \leq |y-x|^p \int_0^1 |\nabla u(\varepsilon x + s\varepsilon(y-x))|^p ds.$$

Integrating both sides over  $\varepsilon \mathcal{C}(x)$  we infer from Fubini's theorem and Remark 2.3 the bound

$$\begin{aligned} \varepsilon^d \left| \frac{u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)}{\varepsilon} \right|^p &\leq C \int_{\varepsilon \mathcal{C}(x)} \int_0^1 |\nabla u(\varepsilon x + s\varepsilon(y-x))|^p ds dz \\ &\leq C \left( \int_{\varepsilon \mathcal{C}(x)} \int_0^1 |\nabla u(z + s\varepsilon(y-x))|^p ds dz + c_u^p \int_{\varepsilon \mathcal{C}(x)} |z - \varepsilon x|^p dz \right) \\ &\leq C \left( \int_0^1 \int_{\varepsilon(\mathcal{C}(x) + s(y-x))} |\nabla u(z)|^p dz ds + c_u^p \varepsilon^{d+p} \right), \end{aligned} \quad (34)$$

where  $c_u$  denotes the  $L^\infty$ -norm of  $D^2u$  on  $B_L \setminus \overline{S_u}$ . Here we used that by Remark 2.3 we have

$$t(\varepsilon x + s\varepsilon(y-x)) + (1-t)(z + s\varepsilon(y-x)) \in B_{2M\varepsilon}(\varepsilon x) \subset B_L \setminus \overline{S_u}, \quad (35)$$

for all  $z \in \varepsilon \mathcal{C}(x)$  and  $s, t \in [0, 1]$ , so that  $c_u$  indeed provides Lipschitz estimates for  $\nabla u$ . Due to (14) we have  $f(p) \leq C_f \|p\|_1$ , so that (10), (34) and (35) imply



$$\varepsilon^{d-1} f(\varepsilon |\nabla_{\omega, \varepsilon}(u, A)|^p(\varepsilon x)) \leq C \left( \int_{B_{2M\varepsilon}(\varepsilon x)} |\nabla u(z)|^p dz + c_u^p \varepsilon^{d+p} \right). \quad (36)$$

In order to control the contribution from the remaining lattice points, fix a set  $A' \in \mathcal{A}^R(\mathbb{R}^d)$  such that  $A \subset \subset A'$ . Then, for  $\varepsilon$  small enough, Remark 2.3 yields the estimate

$$\varepsilon^{d-1} \# \{ \varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap A : \text{dist}(\varepsilon x, \overline{S_u}) < 3M\varepsilon \} \leq \frac{|(\overline{S_u} \cap A') + B_{4M\varepsilon}(0)|}{\varepsilon |B_{\frac{r}{2}}(0)|}.$$

Recall that  $\overline{S_u}$  is the intersection of  $B_L$  with a finite union of pairwise disjoint  $(d-1)$ -simplices, so that  $\overline{S_u} \cap \overline{A'}$  admits a  $(d-1)$ -dimensional Minkowski content. Hence, letting  $\varepsilon \rightarrow 0$ , it holds that

$$\limsup_{\varepsilon \rightarrow 0} \frac{|(\overline{S_u} \cap A') + B_{4M\varepsilon}(0)|}{\varepsilon |B_{\frac{r}{2}}(0)|} \leq C \mathcal{H}^{d-1}(\overline{S_u} \cap \overline{A'}) = \mathcal{H}^{d-1}(S_u \cap \overline{A'}),$$

where we used assumption (i) in the second identity. Since  $f$  is bounded, from (36) and the bound above we conclude that

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A) &\leq \limsup_{\varepsilon \rightarrow 0} \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap A} C \left( \int_{B_{2M\varepsilon}(\varepsilon x)} |\nabla u(z)|^p dz + c_u^p \varepsilon^{d+p} \right) + C \mathcal{H}^{d-1}(S_u \cap \overline{A'}) \\ &\leq C \int_{A'} |\nabla u(z)|^p dz + C \mathcal{H}^{d-1}(S_u \cap \overline{A'}). \end{aligned}$$

Letting  $A' \downarrow \overline{A}$  in this estimate yields by definition of  $F''(\omega)$  that

$$F''(\omega)(u|_D, A) \leq C \int_A |\nabla u(z)|^p dz + C \mathcal{H}^{d-1}(S_u \cap \overline{A}). \quad (37)$$

From this estimate we can now prove the claim by density. First we assume that  $u \in SBV^p(A, \mathbb{R}^m) \cap L^\infty(A, \mathbb{R}^m)$ . Due to the Lipschitz regularity of  $\partial A$  we can use a local reflection argument to extend  $u$  to a function  $\tilde{u} \in SBV^p(B_L, \mathbb{R}^m) \cap L^\infty(B_L, \mathbb{R}^m)$  such that  $\mathcal{H}^{d-1}(S_{\tilde{u}} \cap \partial A) = 0$ . By [33, Theorem 3.1] applied to the large set  $B_L$  we find a sequence  $u_n \in SBV^p(B_L, \mathbb{R}^m)$  fulfilling assumptions (i)-(iii) of the first part such that  $u_n \rightarrow \tilde{u}$  in  $L^1(B_L, \mathbb{R}^m)$ ,  $\nabla u_n \rightarrow \nabla \tilde{u}$  in  $L^p(B_L, \mathbb{R}^{m \times d})$  and  $\limsup_n \mathcal{H}^{d-1}(S_{u_n} \cap \overline{A}) \leq \mathcal{H}^{d-1}(S_{\tilde{u}} \cap \overline{A}) = \mathcal{H}^{d-1}(S_u \cap \overline{A})$ . From locality and lower semicontinuity of  $F''(\omega)(\cdot, A)$  and (37) we deduce

$$F''(u, A) = F''(\tilde{u}|_D, A) \leq \liminf_n F''(\omega)(u_n|_D, A) \leq C \int_A |\nabla u(z)|^p dz + C \mathcal{H}^{d-1}(S_u \cap \overline{A}).$$

It remains to remove the  $L^\infty$ -bound. To this end, given any  $u \in GSBV^p(A, \mathbb{R}^m) \cap L^1(D, \mathbb{R}^m)$ , we consider the truncated sequence  $T_k u \in SBV^p(A, \mathbb{R}^m) \cap L^\infty(A, \mathbb{R}^m)$ . Then  $u_k \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$  and, as in the previous reasoning, the claim follows by lower semicontinuity of  $F''(\omega)(\cdot, A)$ , Lemma 2.1 and the estimate established for bounded functions.  $\square$

The following technical lemma establishes an almost subadditivity of the set function  $A \mapsto F''(\omega)(u, A)$ .

**Proposition 5.8** (Almost subadditivity). *Let  $A, B \in \mathcal{A}^R(D)$ . Moreover let  $A' \in \mathcal{A}^R(D)$  be such that  $A' \subset \subset A$ . Then, for all  $u \in L^1(D, \mathbb{R}^m)$ ,*

$$F''(\omega)(u, A' \cup B) \leq F''(\omega)(u, A) + F''(\omega)(u, B).$$

**Proof.** We can assume that  $F''(u, A)$  and  $F''(u, B)$  are both finite. Since  $A' \cup B \in \mathcal{A}(D)$ , Lemma 5.4 allows us to reduce the proof to the case  $u \in L^\infty(D, \mathbb{R}^m)$ . Let  $u_\varepsilon, v_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  both converge to  $u$  in  $L^1(D, \mathbb{R}^m)$  such that

$$\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon, A) = F''(\omega)(u, A), \quad \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(v_\varepsilon, B) = F''(\omega)(u, B). \quad (38)$$

By Lemma 5.4 we may assume that  $\|u_\varepsilon\|_\infty, \|v_\varepsilon\|_\infty \leq 3\|u\|_\infty$ , so that both sequences actually converge to  $u$  in  $L^p(D, \mathbb{R}^m)$ . Fix  $h \leq \text{dist}(A', A^c)$  and  $N \in \mathbb{N}$ . For  $i = 1, \dots, N$  we define the sets

$$A_i := \left\{ x \in A : \text{dist}(x, A') < i \frac{h}{2N} \right\}.$$

Let  $0 \leq \Theta_i \leq 1$  be a cut-off function between the sets  $A_i$  and  $A_{i+1}$ , that means  $\Theta_i = 1$  on  $A_i$  and  $\Theta_i = 0$  on  $\mathbb{R}^d \setminus A_{i+1}$ . We may assume that  $\|\nabla \Theta_i\|_\infty \leq \frac{4N}{h}$ . Then define  $w_\varepsilon^i \in \mathcal{PC}_\varepsilon^\omega$  by

$$w_\varepsilon^i(\varepsilon x) = \Theta_i(\varepsilon x)u_\varepsilon(\varepsilon x) + (1 - \Theta_i(\varepsilon x))v_\varepsilon(\varepsilon x).$$

Note that for fixed  $i \in \{1, \dots, N\}$  it holds that  $w_\varepsilon^i \rightarrow u$  in  $L^p(D, \mathbb{R}^m)$ . We define the layer-like set

$$S_\varepsilon^i := \{x \in A' \cup B : \text{dist}(x, A_{i+1} \setminus A_{i-1}) < 3M\varepsilon\}.$$

Then by the definition of the localized functionals we can decompose  $F_\varepsilon(\omega)(w_\varepsilon^i, A' \cup B)$  via

$$\begin{aligned} F_\varepsilon(\omega)(w_\varepsilon^i, A' \cup B) &\leq F_\varepsilon(\omega)(u_\varepsilon, A_i) + F_\varepsilon(\omega)(v_\varepsilon, B \setminus \overline{A_{i+1}}) + F_\varepsilon(\omega)(w_\varepsilon^i, S_\varepsilon^i) \\ &\leq F_\varepsilon(\omega)(u_\varepsilon, A) + F_\varepsilon(\omega)(v_\varepsilon, B) + F_\varepsilon(\omega)(w_\varepsilon^i, S_\varepsilon^i). \end{aligned} \quad (39)$$

We show that the last term is negligible. This will be done by averaging. Observe that

$$\begin{aligned} w_\varepsilon^i(\varepsilon y) - w_\varepsilon^i(\varepsilon x) &= \Theta_i(\varepsilon y)(u_\varepsilon(\varepsilon y) - u_\varepsilon(\varepsilon x)) + (1 - \Theta_i(\varepsilon y))(v_\varepsilon(\varepsilon y) - v_\varepsilon(\varepsilon x)) \\ &\quad + (\Theta_i(\varepsilon y) - \Theta_i(\varepsilon x))(u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)) \end{aligned}$$

for all  $x, y \in \mathcal{L}(\omega)$ . Applying the convexity inequality  $(a + b + c)^p \leq 3^{p-1}(a^p + b^p + c^p)$  and the mean value theorem for  $\Theta_i$ , we obtain for all  $(x, y) \in \mathcal{E}(\omega)$  the bound

$$\begin{aligned} \varepsilon \left| \frac{w_\varepsilon^i(\varepsilon y) - w_\varepsilon^i(\varepsilon x)}{\varepsilon} \right|^p &\leq 3^{p-1} \varepsilon \left| \frac{u_\varepsilon(\varepsilon y) - u_\varepsilon(\varepsilon x)}{\varepsilon} \right|^p + 3^{p-1} \varepsilon \left| \frac{v_\varepsilon(\varepsilon y) - v_\varepsilon(\varepsilon x)}{\varepsilon} \right|^p + \frac{(12MN)^p}{3h^p} \varepsilon |u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)|^p. \end{aligned}$$

Summing this estimate over all  $\varepsilon y \in \varepsilon \mathcal{L}(\omega) \cap S_\varepsilon^i$  with  $(x, y) \in \mathcal{E}(\omega)$  we infer

$$\begin{aligned} \|\varepsilon |\nabla_{\omega, \varepsilon}(w_\varepsilon^i, S_\varepsilon^i)|^p(\varepsilon x)\|_1 &\leq 3^{p-1} \|\varepsilon |\nabla_{\omega, \varepsilon}(u_\varepsilon, S_\varepsilon^i)|^p(\varepsilon x)\|_1 + 3^{p-1} \|\varepsilon |\nabla_{\omega, \varepsilon}(v_\varepsilon, S_\varepsilon^i)|^p(\varepsilon x)\|_1 \\ &\quad + CN^p \varepsilon |u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)|^p. \end{aligned} \quad (40)$$

Note that for all  $\lambda \geq 1$  and  $x, y \in \mathbb{R}$  it holds that

$$\min\{\lambda x + y, 1\} \leq \lambda \min\{x, 1\} + \min\{y, 1\}.$$

We combine this estimate with the bound (14) and the monotonicity of  $x \mapsto \min\{x, 1\}$  to deduce from (40) that

$$\begin{aligned} \varepsilon^{d-1} f(\varepsilon |\nabla_{\omega, \varepsilon}(w_\varepsilon^i, S_\varepsilon^i)|^p(\varepsilon x)) &\leq C \varepsilon^{d-1} f(\varepsilon |\nabla_{\omega, \varepsilon}(u_\varepsilon, S_\varepsilon^i)|^p(\varepsilon x)) + C \varepsilon^{d-1} f(\varepsilon |\nabla_{\omega, \varepsilon}(v_\varepsilon, S_\varepsilon^i)|^p(\varepsilon x)) \\ &\quad + CN^p \varepsilon^d |u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)|^p. \end{aligned}$$

Summing this inequality yields

$$F_\varepsilon(\omega)(w_\varepsilon^i, S_\varepsilon^i) \leq C(F_\varepsilon(\omega)(u_\varepsilon, S_\varepsilon^i) + F_\varepsilon(\omega)(v_\varepsilon, S_\varepsilon^i)) + CN^p \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap S_\varepsilon^i} \varepsilon^d |u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)|^p.$$

For  $\varepsilon$  small enough we have  $S_\varepsilon^i \cap S_\varepsilon^j = \emptyset$  for  $|i - j| \geq 3$ . Moreover,  $S_\varepsilon^i \subset A \cap B$  for  $i \geq 2$  as well as  $S_\varepsilon^i \subset \subset A$  with a uniform distance to  $\partial A$ . Thus averaging the last inequality and applying (38) gives

$$\begin{aligned} \frac{1}{N-1} \sum_{i=2}^N F_\varepsilon(\omega)(w_\varepsilon^i, S_\varepsilon^i) &\leq \frac{C}{N} (F_\varepsilon(\omega)(u_\varepsilon, A) + F_\varepsilon(\omega)(v_\varepsilon, B)) + CN^{p-1} \sum_{\varepsilon \mathcal{C}(x) \subset A} \varepsilon^d |u_\varepsilon(\varepsilon x) - v_\varepsilon(\varepsilon x)|^p \\ &\leq \frac{C}{N} + CN^{p-1} \|u_\varepsilon - v_\varepsilon\|_{L^p(D)}^p. \end{aligned}$$

Since we have  $u_\varepsilon - v_\varepsilon \rightarrow 0$  also in  $L^p(D, \mathbb{R}^m)$ , the last term vanishes when  $\varepsilon \rightarrow 0$ . For every  $\varepsilon > 0$  let  $i_\varepsilon \in \{2, \dots, N\}$  be such that

$$F_\varepsilon(\omega)(w_\varepsilon^{i_\varepsilon}, S_\varepsilon^{i_\varepsilon}) \leq \frac{1}{N-1} \sum_{i=2}^N F_\varepsilon(\omega)(w_\varepsilon^i, S_\varepsilon^i) \leq \frac{C}{N} + CN^{p-1} \|u_\varepsilon - v_\varepsilon\|_{L^p(D)}^2 \quad (41)$$

and set  $w_\varepsilon := w_\varepsilon^{i_\varepsilon}$ . Note that  $w_\varepsilon$  still converges to  $u$  strongly in  $L^p(D, \mathbb{R}^m)$ . Hence, using (38), (39) and (41), we conclude that

$$F''(\omega)(u, A' \cup B) \leq \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(w_\varepsilon, A' \cup B) \leq F''(\omega)(u, A) + F''(\omega)(u, B) + \frac{C}{N}.$$

The claim follows by letting  $N \rightarrow +\infty$ .  $\square$

In the lemma below we state the last property that we need to prove Proposition 5.2.

**Lemma 5.9** (Inner regularity). *Let  $u \in L^1(D)$ . Then for any  $A \in \mathcal{A}^R(D)$  it holds that*

$$F''(\omega)(u, A) = \sup_{A' \subset \subset A} F''(\omega)(u, A').$$

**Proof.** It suffices to prove one inequality since  $A \mapsto F''(\omega)(u, A)$  is monotone with respect to set inclusion. For  $k \in \mathbb{N}$  define the set  $A_k = \{x \in A : \text{dist}(x, \partial A) > 2^{-k}\}$ . Then for  $k$  large enough we have that  $A_k, A \setminus A_k \in \mathcal{A}^R(D)$  (see [44, Lemma 2.2]). We first treat the case  $u \notin GSBVP(A, \mathbb{R}^m)$  and prove that  $\limsup_k F''(\omega)(u, A_k) = +\infty$ . Assume by contradiction that this sequence is bounded. Then, for each  $k$  we have  $u \in GSBVP(A_k, \mathbb{R}^m)$  by Lemma 5.6 and thus  $u \in GSBV(A, \mathbb{R}^m)$ . Since the measure of the jump set and the  $L^p$ -norm of the gradient in  $A_k$  are equibounded with respect to  $k$ , we reach the contradiction  $u \in GSBVP(A, \mathbb{R}^m)$ . Now assume that  $u \in GSBVP(A, \mathbb{R}^m)$ . Note that  $A = A_{k+1} \cup A \setminus \overline{A_k}$ . Hence Lemma 5.7 and Proposition 5.8 imply

$$\begin{aligned} F''(\omega)(u, A) &\leq F''(\omega)(u, A_{k+2}) + C \left( \|\nabla u\|_{L^p(A \setminus \overline{A_k})}^p + \mathcal{H}^{d-1}(S_u \cap (A \setminus \overline{A_k})) \right) \\ &\leq \sup_{A' \subset \subset A} F''(\omega)(u, A') + C \left( \|\nabla u\|_{L^p(A \setminus \overline{A_k})}^p + \mathcal{H}^{d-1}(S_u \cap (A \setminus \overline{A_k})) \right). \end{aligned}$$

Letting  $k \rightarrow +\infty$  proves the claim since  $u \in GSBVP(A, \mathbb{R}^m)$ .  $\square$

**Proof of Proposition 5.2.** Given a sequence  $\varepsilon \rightarrow 0^+$ , by the compactness property of  $\Gamma$ -convergence on separable metric spaces (see [21, Proposition 1.42]) we find a subsequence  $\varepsilon_n$  such that

$$\Gamma\text{-}\lim_n F_{\varepsilon_n}(\omega)(u, R) =: \tilde{F}(\omega)(u, R)$$

exists for every  $u \in L^1(D, \mathbb{R}^m)$  and all sets  $R \in \mathcal{R}$ , where  $\mathcal{R}$  denotes the class of all subsets of  $D$  that are finite unions of rectangles with rational vertices. Due to Lemma 5.9 and monotonicity, for every  $u \in L^1(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}^R(D)$  we conclude that

$$\Gamma\text{-}\limsup_n F_{\varepsilon_n}(\omega)(u, A) \leq \sup_{R \in \mathcal{R}} \tilde{F}(\omega)(u, R) \leq \Gamma\text{-}\liminf_n F_{\varepsilon_n}(\omega)(u, A),$$

so that we can define  $\tilde{F}(\omega)(u, A) := \Gamma\text{-}\lim_n F_{\varepsilon_n}(\omega)(u, A)$  also for all  $u \in L^1(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}^R(D)$ . We extend  $\tilde{F}(\omega)(u, \cdot)$  to  $\mathcal{A}(D)$  via its inner regular envelope

$$F(\omega)(u, A) := \sup \{ \tilde{F}(\omega)(u, A') : A' \subset \subset A, A' \in \mathcal{A}^R(D) \}.$$

By Lemma 5.9 this functional indeed extends  $\tilde{F}(\omega)(u, \cdot)$ . Next we need to slightly perturb the functional. Given  $\eta > 0$ , for any  $u \in SBV^p(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}(D)$  we define the auxiliary functional  $\mathcal{F}_\eta : SBV^p(D, \mathbb{R}^m) \times \mathcal{A}(D) \rightarrow [0, +\infty)$  as

$$\mathcal{F}_\eta(u, A) = F(\omega)(u, A) + \eta \int_{S_u \cap A} |u^+ - u^-| d\mathcal{H}^{d-1}.$$

We argue that  $\mathcal{F}_\eta$  satisfies the assumptions of [18, Theorem 1], that means,

- (i)  $\mathcal{F}_\eta(u, \cdot)$  is the restriction to  $\mathcal{A}(D)$  of a Radon measure;
- (ii)  $\mathcal{F}_\eta(u, A) = \mathcal{F}_\eta(v, A)$  whenever  $u = v$  a.e. on  $A \in \mathcal{A}(D)$ ;
- (iii)  $\mathcal{F}_\eta(\cdot, A)$  is  $L^1(D, \mathbb{R}^m)$ -lower semicontinuous;
- (iv) there exists  $c > 0$  such that

$$\begin{aligned} & \frac{1}{c} \left( \int_A |\nabla u|^p dx + \int_{S_u \cap A} (1 + |u^+ - u^-|) d\mathcal{H}^{d-1} \right) \leq \mathcal{F}_\eta(u, A) \\ & \leq c \left( \int_A (1 + |\nabla u|^p) + \int_{S_u \cap A} (1 + |u^+ - u^-|) d\mathcal{H}^{d-1} \right). \end{aligned}$$

(i): We verify the De Giorgi–Letta criterion (see [40, Theorem 1.62]). Clearly  $A \mapsto \mathcal{F}_\eta(u, A)$  is a non-negative, increasing and inner regular set function with  $\mathcal{F}_\eta(u, \emptyset) = 0$ . Moreover, discrete superadditivity on disjoint sets is transferred from  $F_\varepsilon(\omega)(u, \cdot)$  to  $F'(\omega)(u, \cdot)$ , which implies that  $\mathcal{F}_\eta(u, A \cup B) \geq \mathcal{F}_\eta(u, A) + \mathcal{F}_\eta(u, B)$  whenever  $A \cap B = \emptyset$ . In order to prove subadditivity, let  $A, B \in \mathcal{A}(D)$  and consider  $S \subset \subset A \cup B$  such that  $S \in \mathcal{A}^R(D)$ . Let us define the set  $A_k = \{x \in A : \text{dist}(x, \partial A) > 2^{-k}\}$  and similarly we define  $B_k$ . Then the family  $\{A_k \cup B_k\}_k$  forms an open cover of  $\bar{S}$ . By compactness we find an index  $k_0$  such that  $\bar{S} \subset A_{k_0} \cup B_{k_0}$ . Next we regularize the sets  $A_{2k_0}, B_{2k_0}$  and  $A_{4k_0}, B_{4k_0}$  by standard methods to find further sets  $A_0, A_1, B_0, B_1 \in \mathcal{A}^R(D)$  such that  $\bar{S} \subset A_0 \cup B_0$  and  $A_0 \subset \subset A_1 \subset \subset A$  and  $B_0 \subset \subset B_1 \subset \subset B$ . Then by Proposition 5.8

$$\tilde{F}(\omega)(u, S) \leq \Gamma\text{-}\limsup_n F_{\varepsilon_n}(\omega)(u, A_0 \cup B_0) \leq \tilde{F}(\omega)(u, A_1) + \tilde{F}(\omega)(u, B_1) \leq F(\omega)(u, A) + F(\omega)(u, B).$$

Taking the supremum over such  $S$  yields the subadditivity of  $A \mapsto F(\omega)(u, A)$ . The corresponding property for the perturbation term is straightforward. Thus the De Giorgi–Letta criterion applies and we infer that  $\mathcal{F}_\eta(u, \cdot)$  is the trace of a Borel measure. By Lemma 5.7 this measure is finite, so it is a Radon measure.

(ii)+(iii): The locality property follows from Lemma 5.5 and the definition of  $F(\omega)$  by inner approximation as well as locality of the perturbation term. By the properties of  $\Gamma$ -limits we know that  $\tilde{F}(\omega)(\cdot, A)$  is  $L^1(D, \mathbb{R}^m)$ -lower semicontinuous and so is  $F(\omega)(\cdot, A)$  as the supremum of lower semicontinuous functionals.  $L^1(D, \mathbb{R}^m)$ -lower semicontinuity of the perturbation term along sequences such that  $F(\omega)(u_n, A)$  remains bounded follows from the bounds established in Lemma 5.6 that still hold for  $F(\omega)(u, A)$ . Indeed those bounds yield stronger compactness so that we can combine [7, Theorems 2.2 and 3.7] to conclude lower semicontinuity. Hence  $\mathcal{F}(\cdot, A)$  is lower semicontinuous as the sum of (finite) lower semicontinuous functionals.

(iv): The bounds follow from Lemmata 5.6 and 5.7, which still hold for  $F(\omega)$  in place of  $\tilde{F}(\omega)$ , and the definition of the perturbation term.

From [18, Theorem 1] and the fact that  $F_\varepsilon(\omega)(u + z, A) = F_\varepsilon(\omega)(u, A)$  for all  $z \in \mathbb{R}^m$  we deduce that  $\mathcal{F}_\eta$  has the representation

$$\mathcal{F}_\eta(u, A) = \int_A h_\eta(x, \nabla u) dx + \int_{S_u \cap A} \varphi_\eta(x, u^+ - u^-, \nu_u) d\mathcal{H}^{d-1} \quad (42)$$

for all  $u \in SBV^p(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}(D)$  with the integrands given by the asymptotic formulas

$$h_\eta(x_0, \xi) = \limsup_{\varrho \rightarrow 0} \frac{1}{\varrho^d} \inf\{\mathcal{F}_\eta(v, Q_v(x_0, \varrho)) : v = \xi(\cdot - x_0) \text{ in a neighborhood of } \partial Q_v(x_0, \varrho)\},$$

$$\varphi_\eta(x_0, a, v) = \limsup_{\varrho \rightarrow 0} \frac{1}{\varrho^{d-1}} \inf\{\mathcal{F}_\eta(v, Q_v(x_0, \varrho)) : v = u_{x_0, v}^{a, 0} \text{ in a neighborhood of } \partial Q_v(x_0, \varrho)\}.$$

Hence, due to locality, for every  $u \in L^1(D, \mathbb{R}^m)$  and  $A \in \mathcal{A}^R(D)$  such that  $u \in SBVP(A, \mathbb{R}^m)$  we have

$$\Gamma\text{-}\lim_n F_{\varepsilon_n}(\omega)(u, A) = F(\omega)(u, A) = \int_A h_\eta(x, \nabla u) \, dx + \int_{S_u \cap A} (\varphi_\eta(x, u^+ - u^-, v_u) - \eta|u^+ - u^-|) \, d\mathcal{H}^{d-1}.$$

It remains to prove the formulas for the integrands stated in Proposition 5.2. Note that the mapping  $\eta \mapsto \varphi_\eta(x_0, a, v)$  is increasing and non-negative on  $(0, +\infty)$ . Hence there exists the limit  $\varphi(x_0, a, v) = \lim_{\eta \rightarrow 0} \varphi_\eta(x_0, a, v)$ . By the same reasoning there exists  $h(x_0, \xi) = \lim_{\eta \rightarrow 0} h_\eta(x_0, \xi)$ , so that monotone convergence implies that

$$F(\omega)(u, A) = \int_A h(x, \nabla u) \, dx + \int_{S_u \cap A} \varphi(x, u^+ - u^-, v_u) \, d\mathcal{H}^{d-1}$$

for every  $A \in \mathcal{A}^R(D)$  and every  $u \in L^1(D, \mathbb{R}^m)$  such that  $u \in SBVP(A, \mathbb{R}^m)$ . Since  $F(\omega)(u, A) \leq \mathcal{F}_\eta(u, A)$ , we further know by the definition of  $m(\omega)(v, A)$  (cf. the statement of Proposition 5.2) that

$$\limsup_{\varrho \rightarrow 0} \varrho^{1-d} m(\omega)(u_{x_0, v}^{a, 0}, Q_v(x_0, \varrho)) \leq \lim_{\eta \rightarrow 0} \varphi_\eta(x_0, a, v) = \varphi(x_0, a, v).$$

In order to show the reverse inequality, we note that Lemma 5.4 implies a very weak maximum principle for  $F(\omega)$ : there exists  $u_\varrho \in SBVP(Q_v(x_0, \varrho), \mathbb{R}^m)$  admissible for the definition of  $m(\omega)(u_{x_0, v}^{a, 0}, Q_v(x_0, \varrho))$  and satisfying  $\|u\|_\infty \leq 3|a|$  such that

$$\varrho^{1-d} F(\omega)(u_\varrho, Q_v(x_0, \varrho)) \leq m(\omega)(u_{x_0, v}^{a, 0}, Q_v(x_0, \varrho)) + \varrho.$$

Then clearly  $|u_\varrho^+ - u_\varrho^-| \leq 6|a|$  for  $\mathcal{H}^{d-1}$ -a.e.  $x \in S_{u_\varrho}$ . With the lower bound of Lemma 5.6 we obtain

$$\begin{aligned} \varphi_\eta(x_0, a, v) &\leq \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \left( F(\omega)(u_\varrho, Q_v(x_0, \varrho)) + 6|a|\eta \mathcal{H}^{d-1}(S_{u_\varrho} \cap Q_v(x_0, \varrho)) \right) \\ &\leq \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \left( m(\omega)(u_{x_0, v}^{a, 0}, Q_v(x_0, \varrho)) + C|a|\eta F(\omega)(u_\varrho, Q_v(x_0, \varrho)) \right) \\ &\leq (1 + C|a|\eta) \limsup_{\varrho \rightarrow 0} \varrho^{1-d} m(\omega)(u_{x_0, v}^{a, 0}, Q_v(x_0, \varrho)). \end{aligned}$$

Since the term on the right hand side is finite, we conclude by taking the limit as  $\eta \rightarrow 0$ . The proof of the formula for  $h$  is the same except that we can choose  $u_\varrho$  even such that  $\|u_\varrho\|_\infty \leq C|\xi|\varrho$ , so that there is no need to let  $\eta \rightarrow 0$  at the end. This finishes the proof of Proposition 5.2.  $\square$

## 5.2. Characterization of the bulk density

In this section we show that the function  $h$  given by Proposition 5.2 agrees with the density of the  $\Gamma$ -limit of the discrete functionals defined in (17).

Let us briefly explain the strategy: for both functionals  $F_\varepsilon(\omega)$  and  $E_\varepsilon(\omega)$  we consider recovery sequences for affine functions instead of minimization problems with affine boundary conditions since this approach shortens the proof. Our aim is to modify those sequences on a small set such that  $\|\varepsilon|\nabla_{\omega, \varepsilon}(u_\varepsilon, A)|^p\|_1$  is equiintegrable because at that level  $F_\varepsilon(\omega)$  and  $E_\varepsilon(\omega)$  become comparable. It is well-known that if a sequence is bounded in  $W^{1, p}$ , then up to a subsequence one can modify it on a set whose measure goes to zero in such a way that the gradients become  $p$ -equiintegrable (see [41, Lemma 1.2]). This is usually achieved via an abstract Lipschitz-extension on the set where the maximal function of the norm of the gradient is very large. Indeed, this classical strategy goes back at least to [1]. To control the size of the set where the function is modified, one estimates the maximal function by the gradient itself, which is possible only for  $p > 1$ . In our case no a priori  $L^p$ -bounds on the gradient are available. However, in [48] the

modification procedure was extended to  $BV$ -sequences with vanishing singular part. In our setting the basic idea is the following observation: on a ball of radius  $\varrho$  the energy of a recovery sequence can be trivially bounded by  $\sim \varrho^d$ . Morally speaking, this shows that the set where no  $L^p$ -bounds are available is relatively small with a quantitative rate in  $\varrho$ . In suitable diagonal regimes of  $\varrho$  and  $\varepsilon$  this can be exploited to still perform a modification on small sets which makes the gradients equiintegrable. However, we emphasize that we have to transfer these ideas to the discrete environment.

The analysis splits into three different parts. As a first step, we argue that affine functions indeed fully characterize the function  $h$  and establish the framework for the diagonal argument on small balls  $B_\varrho(x_0)$ . Then we introduce the notion of discrete maximal functions on stochastic lattices and prove the doubling property of the counting measure as well as a Poincaré inequality. Those properties allow to perform modifications on small sets by abstract Lipschitz extensions. Finally, with these tools at hand, we can carefully modify recovery sequences on smaller and smaller balls to conclude by a blow-up argument.

Although it is just a technical detail, we first prove that  $h$  is a Carathéodory function, which is necessary to conclude its quasiconvexity.

**Lemma 5.10.** *Let  $h : D \times \mathbb{R}^{m \times d} \rightarrow [0, +\infty)$  be given by Proposition 5.2. Then, for every  $\xi \in \mathbb{R}^{m \times d}$  the map  $x \mapsto h(x, \xi)$  is measurable and, for every  $x \in D$ , the map  $\xi \mapsto h(x, \xi)$  is continuous.*

**Proof.** Denoting by  $u_{x_0, \xi}$  the affine function  $u_{x_0, \xi}(x) = \xi(x - x_0)$ , the function  $x \mapsto h_\eta(x, \xi)$  defined in the proof of Proposition 5.2 is the Radon-Nikodym derivative of the measure  $\mathcal{F}_\eta(u_{x_0, \xi}, \cdot)$  with respect to the Lebesgue-measure. Hence it is measurable and so is  $x \mapsto h(x, \xi)$  as the pointwise limit. In order to prove continuity in  $\xi$ , let us write  $Q_\varrho = Q_v(x_0, \varrho)$  and fix  $\xi_1, \xi_2 \in \mathbb{R}^{m \times d}$  and  $\eta > 0$ . Consider a smooth function  $0 \leq \Theta \leq 1$  such that  $\Theta = 1$  on  $Q_\varrho$  and  $\text{supp}(\Theta) \subset Q_{(1+\eta)\varrho}$  satisfying in addition  $\|\nabla \Theta\|_\infty \leq \frac{2}{\varrho\eta}$ . Given  $u_1 \in SBV^p(Q_\varrho, \mathbb{R}^m)$  such that  $u_1 = u_{x_0, \xi_1}$  in a neighborhood of  $\partial Q_\varrho$  (extended to the whole space) we define  $u_2 \in SBV^p(Q_{(1+2\eta)\varrho}, \mathbb{R}^m)$  by

$$u_2 = \Theta u_1 + (1 - \Theta)u_{x_0, \xi_2}.$$

Then  $\mathcal{H}^{d-1}(S_{u_2} \cap (Q_{(1+2\eta)\varrho} \setminus Q_\varrho)) = 0$  and  $u_2 = u_{x_0, \xi_2}$  in a neighborhood of  $\partial Q_{(1+2\eta)\varrho}$ . Hence by Lemma 5.7

$$\begin{aligned} \frac{m(\omega)(u_{x_0, \xi_2}, Q_{(1+2\eta)\varrho})}{((1+2\eta)\varrho)^d} &\leq \frac{F(\omega)(u_2, Q_{(1+2\eta)\varrho})}{\varrho^d} \leq \frac{F(\omega)(u_1, Q_\varrho)}{\varrho^d} + C\varrho^{-d} \int_{Q_{(1+2\eta)\varrho} \setminus \overline{Q_\varrho}} |\nabla u_2|^p dz \\ &\leq \frac{F(\omega)(u_1, Q_\varrho)}{\varrho^d} + C((1+2\eta)^d - 1) \left( (|\xi_1|^p + |\xi_2|^p) + \frac{\varrho^p}{(\varrho\eta)^p} |\xi_1 - \xi_2|^p \right). \end{aligned}$$

Since  $u_1$  was arbitrary, by the definition of  $h$  we obtain for fixed  $0 < \eta < \frac{1}{2}$  the inequality

$$h(x_0, \xi_2) = \limsup_{\varrho \rightarrow 0} \frac{m(\omega)(u_{x_0, \xi_2}, Q_{(1+2\eta)\varrho})}{((1+2\eta)\varrho)^d} \leq h(x_0, \xi_1) + C\eta \left( (|\xi_1|^p + |\xi_2|^p) + \frac{1}{\eta^p} |\xi_1 - \xi_2|^p \right).$$

Exchanging the roles of  $\xi_1$  and  $\xi_2$ , for any  $0 < \eta < \frac{1}{2}$  we infer

$$|h(x_0, \xi_1) - h(x_0, \xi_2)| \leq C\eta \left( (|\xi_1|^p + |\xi_2|^p) + \frac{1}{\eta^p} |\xi_1 - \xi_2|^p \right).$$

This estimate implies continuity since the right hand side can be made arbitrarily small on bounded sets by first adjusting  $\eta$  and then the difference  $|\xi_1 - \xi_2|$ .  $\square$

In the next lemma we prove that  $h$  is determined by the behavior of  $F(\omega)(u_{x_0, \xi}, B_\varrho(x_0))$  for  $\varrho \rightarrow 0$ .

**Lemma 5.11.** *Let  $\varepsilon_n$  and  $F(\omega)$  be as in Proposition 5.2. Then there exists a null set  $N \subset D$  such that for every  $x_0 \in D \setminus N$  and all  $\xi \in \mathbb{R}^{m \times d}$  it holds that*

$$|B_1| h(x_0, \xi) = \lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)),$$

where  $u_{x_0, \xi}$  denotes the affine function  $x \mapsto \xi(x - x_0)$ .

**Proof.** It follows from the definition of  $F(\omega)$  that

$$\varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) = |B_1| \int_{B_\varrho(x_0)} h(z, \xi) dz.$$

Due to Lebesgue's differentiation theorem there exists a null set  $N_\xi \subset D$  such that for any  $x_0 \in D \setminus N_\xi$  we have

$$|B_1| h(x_0, \xi) = \lim_{\delta \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)). \quad (43)$$

It remains to show that the null sets can be chosen independent of  $\xi$ . To this end, observe that the restriction of  $F(\omega)$  to  $W^{1,p}(D, \mathbb{R}^m)$  is  $L^1(D, \mathbb{R}^m)$ -lower semicontinuous. Using the defining formula provided by Proposition 5.2 and testing an affine function, by Lemma 5.7 we find that  $0 \leq h(x_0, \xi) \leq C|\xi|^p$  and Lemma 5.10 yields that  $h$  is a Carathéodory-function. Hence by standard results (see [34, Theorem 1.13]) there exists a null set  $N' \subset D$  such that for every  $x \in D \setminus N'$  the function  $\xi \mapsto h(x, \xi)$  is quasiconvex on  $\mathbb{R}^{m \times d}$  and therefore locally Lipschitz continuous. Moreover the Lipschitz constant is uniformly bounded on compact sets. Let us define the null set  $N = N' \cup \bigcup_{\xi \in \mathbb{Q}^{m \times d}} N_\xi$ . Then for any  $x_0 \in D \setminus N$  it holds that  $h(x_0, \xi) = \lim_n h(x_0, \xi_n)$ , where  $\mathbb{Q}^{m \times d} \ni \xi_n \rightarrow \xi$ . On the other hand, whenever  $\xi_n \rightarrow \xi$ , then without loss of generality  $\sup_n |\xi_n| \leq (|\xi| + 1)$  and by the uniform local Lipschitz continuity we obtain

$$\int_{B_\varrho(x_0)} |h(x, \xi_n) - h(x, \xi)| dx \leq C_\xi |\xi - \xi_n|.$$

Thus, for  $x_0 \in D \setminus N$  we can pass to the limit in  $n$  on both sides in (43) finishing the proof.  $\square$

Motivated by the previous lemma we analyze the limit functional  $F(\omega)(u_{x_0, \xi}, B_\varrho(x_0))$  by studying recovery sequences for the affine function  $u_{x_0, \xi}$ . As explained at the beginning of this subsection, we have to localize the analysis to small balls  $B_\varrho(x_0)$  and consider diagonal sequences  $(\varepsilon_\varrho, \varrho)$ . Therefore we need the auxiliary result below, which is an immediate consequence of a change of variables. We omit its proof.

**Lemma 5.12.** *Let  $\varepsilon_n$  and  $E(\omega)$  be as in Theorem 3.1. For  $\varrho > 0$  and  $x_0 \in D$  such that  $B_\varrho(x_0) \subset D$ , define the functional  $G_{\varepsilon_n, \varrho}(x_0, \omega) : L^p(B_1, \mathbb{R}^m) \rightarrow [0, +\infty]$  to be finite only for  $u : \frac{\varepsilon_n}{\varrho}(\mathcal{L}(\omega) - \frac{x_0}{\varepsilon_n}) \rightarrow \mathbb{R}^m$  with value*

$$G_{\varepsilon_n, \varrho}(x_0, \omega)(u) = \alpha \sum_{\substack{(x, y) \in \mathcal{E}(\omega) - \frac{x_0}{\varepsilon_n} \\ \frac{\varepsilon_n}{\varrho}x, \frac{\varepsilon_n}{\varrho}y \in B_1}} \left( \frac{\varepsilon_n}{\varrho} \right)^d \left| \frac{u(\frac{\varepsilon_n}{\varrho}x) - u(\frac{\varepsilon_n}{\varrho}y)}{\varepsilon_n \varrho^{-1}} \right|^p.$$

*Then  $G_{\varepsilon_n, \varrho}(x_0, \omega) \Gamma(L^p(B_1, \mathbb{R}^m))$ -converges to the functional  $E_\varrho(x_0, \omega) : L^p(B_1, \mathbb{R}^m) \rightarrow [0, +\infty]$  with domain  $W^{1,p}(B_1, \mathbb{R}^m)$ , where it is given by*

$$E_\varrho(x_0, \omega)(u) = \int_{B_1} q(x_0 + \varrho y, \nabla u(y)) dy.$$

In order to use diagonal sequences of the functionals  $G_{\varepsilon, \varrho}(x_0, \omega)$ , we first identify the  $\Gamma$ -limit of the functionals  $E_\varrho(x_0, \omega)(u, A)$  when  $\varrho \rightarrow 0$ . This is already contained in [43, Lemma 2.1] for scalar problems. Here we provide a short proof in the vectorial case. We essentially follow the lines of [35, Theorem 5.14] up to some necessary modifications.

**Lemma 5.13** (Blow-up by  $\Gamma$ -convergence). *Let  $E_\varrho(x_0, \omega)$  be a functional as in Lemma 5.12. Then there exists a null set  $N \subset D$  such that for all  $x_0 \in D \setminus N$  it holds that*

$$\Gamma(L^p(B_1, \mathbb{R}^m))\text{-}\lim_{\varrho \rightarrow 0} E_\varrho(x_0, \omega)(u) = \begin{cases} \int_{B_1} q(x_0, \nabla u(y)) dy & \text{if } u \in W^{1,p}(B_1, \mathbb{R}^m), \\ +\infty & \text{otherwise.} \end{cases}$$



**Proof.** By coercivity, the  $\Gamma$ -liminf can be finite only on  $W^{1,p}(B_1, \mathbb{R}^m)$ . Let  $N' \subset D$  be the null set where the function  $\xi \mapsto q(x, \xi)$  is not quasiconvex or does not fulfill the bound  $\frac{1}{C}|\xi|^p - C \leq q(x, \xi) \leq C(|\xi|^p + 1)$ . We redefine  $q(x, \xi) = |\xi|^p$  for all  $x \in N'$  not changing the functional  $E_\varrho(x_0, \omega)$ . Then, according to [34, Proposition 2.32], there exists a constant  $C_q$  such that

$$|q(x, \xi) - q(x, \zeta)| \leq C_q(1 + |\xi|^{p-1} + |\zeta|^{p-1})|\xi - \zeta| \quad (44)$$

for all  $x \in D$  and all  $\xi, \zeta \in \mathbb{R}^{m \times d}$ . Moreover, by Lebesgue's differentiation theorem, for every  $\xi \in \mathbb{R}^{m \times d}$  there exists a null set  $N_\xi$  such that for every  $x_0 \in D \setminus N_\xi$  we have

$$\lim_{\varrho \rightarrow 0} \int_{B_1} |q(x_0 + \varrho y, \xi) - q(x_0, \xi)| dy = \lim_{\varrho \rightarrow 0} \int_{B_\varrho(x_0)} |q(z, \xi) - q(x_0, \xi)| dz = 0. \quad (45)$$

Set  $N = N' \cup \bigcup_{\xi \in \mathbb{Q}^{m \times d}} N_\xi$ , fix  $x_0 \in D \setminus N$  and consider a sequence  $\varrho_j \rightarrow 0$ . We first prove that, along a suitable subsequence, for each  $\xi \in \mathbb{R}^{m \times d}$  the functions  $y \mapsto q_j(y, \xi) = q(x_0 + \varrho_j y, \xi)$  converge a.e. on  $B_1$  to  $q(x_0, \xi)$ . Indeed, due to (45) the convergence holds true in  $L^1(B_1)$  for every  $\xi \in \mathbb{Q}^{m \times d}$ . Thus we find a ( $\xi$ -dependent) subsequence such that  $q_{j_k}(\cdot, \xi) \rightarrow q(x_0, \xi)$  a.e. on  $B_1$ . Enumerating the  $\xi \in \mathbb{Q}^{m \times d}$  we can ensure that the subsequences are nested. Then, by a diagonal argument, we find a common subsequence  $j_k$  such that  $q_{j_k}(\cdot, \xi) \rightarrow q(x_0, \xi)$  a.e. on  $B_1$  for every  $\xi \in \mathbb{Q}^{m \times d}$ . Using (44) we can extend the convergence to all  $\xi \in \mathbb{R}^{m \times d}$ .

We now prove the  $\Gamma$ -convergence along this subsequence. By dominated convergence the existence of a recovery sequence is provided through the pointwise limit. In order to prove the lower bound, we can assume that  $u_j \rightarrow u$  in  $L^p(B_1, \mathbb{R}^m)$  and that  $u_j$  is bounded in  $W^{1,p}(B_1, \mathbb{R}^m)$ . By [41, Lemma 1.2] there exists a subsequence (not relabeled) and another sequence  $z_j \in W^{1,p}(B_1, \mathbb{R}^m)$  such that  $|\nabla z_j|^p$  is equiintegrable and

$$\lim_j |\{z_j \neq u_j \text{ or } \nabla z_j \neq \nabla u_j\}| = 0. \quad (46)$$

Note that the above property implies that  $z_j$  is also bounded in  $W^{1,p}(B_1, \mathbb{R}^m)$ . Otherwise, along a subsequence,  $w_j = \frac{z_j}{\|z_j\|_{L^p}}$  converges in  $W^{1,p}(B_1, \mathbb{R}^m)$  to a non-zero constant, but the sequence  $v_j = \frac{u_j}{\|z_j\|_{L^p}}$  converges to 0 strongly in  $W^{1,p}(B_1, \mathbb{R}^m)$ . This contradicts (46) which remains valid for  $w_j$  and  $v_j$ . Thus we deduce from (46) that  $z_j \rightarrow u$  in  $W^{1,p}(B_1, \mathbb{R}^m)$ . Moreover, equiintegrability of  $|\nabla z_j|^p$ , the bound  $0 \leq q(x, \xi) \leq C(|\xi|^p + 1)$  and (46) imply that

$$\liminf_j E_{\varrho_j}(x_0, \omega)(u_j) \geq \liminf_j E_{\varrho_j}(x_0, \omega)(z_j). \quad (47)$$

Next, for any  $\delta > 0$  we find again by equiintegrability and the upper bound on  $q$  a number  $\eta_\delta > 0$  such that for any measurable set  $G \subset B_1$  with  $|G| \leq \eta_\delta$  it holds that

$$\sup_j \int_G q(x_0, \nabla z_j(y)) dy < \delta. \quad (48)$$

Hence let us choose  $t_\delta > 0$  such that  $\sup_j |\{|\nabla z_j| > t_\delta\}| \leq \eta_\delta$  and set  $K_\delta = C_q(1 + 2t_\delta^{p-1})$  with  $C_q$  given by (44). By a compactness argument we find  $\xi_1, \dots, \xi_N \in \mathbb{R}^{m \times d}$  such that  $|\xi_i| < t_\delta$  and

$$\{\xi \in \mathbb{R}^{m \times d} : |\xi| \leq t_\delta\} \subset \bigcup_{i=1}^N \{\xi \in \mathbb{R}^{m \times d} : |\xi - \xi_i| < \frac{\delta}{K_\delta}\}.$$

Due to Egorov's theorem there exists a set  $G \subset B_1$  with  $|G| \leq \eta_\delta$  such that the sequences  $q_j(y, \xi_i)$  converge uniformly to  $q(x_0, \xi_i)$  on  $B_1 \setminus G$ . Hence there exists  $j_\delta \in \mathbb{N}$  such that for all  $j \geq j_\delta$ , all  $i$  and all  $y \in B_1 \setminus G$  we have  $|q_j(y, \xi_i) - q(x_0, \xi_i)| < \delta$ . Using (44) and the triangle inequality we obtain

$$|q_j(y, \xi) - q(x_0, \xi)| \leq 2C_q(1 + 2t_\delta^{p-1}) \min_i |\xi_i - \xi| + \delta \leq 3\delta$$

for all  $y \in B_1 \setminus G$ , all  $\xi \in \mathbb{R}^{m \times d}$  with  $|\xi| \leq t_\delta$  and all  $j \geq j_\delta$ . Since  $q(x, \xi) \geq 0$ , we infer that for those  $j$

$$\begin{aligned}
E_{q_j}(x_0, \omega)(z_j) &\geq \int_{\{|\nabla z_j| \leq t_\delta\} \setminus G} q_j(y, \nabla z_j(y)) \, dy \geq \int_{\{|\nabla z_j| \leq t_\delta\} \setminus G} q(x_0, \nabla z_j(y)) \, dy - 3|B_1|\delta \\
&\geq \int_{B_1} q(x_0, \nabla z_j(y)) \, dy - (3|B_1| + 2)\delta,
\end{aligned}$$

where we used (48) and the definition of  $t_\delta$ . By quasiconvexity and the growth conditions on  $q$  the last term is lower semicontinuous with respect to weak convergence in  $W^{1,p}(B_1, \mathbb{R}^m)$ . Hence (47) implies

$$\liminf_j E_{q_j}(x_0, \omega)(u_j) \geq \int_{B_1} q(x_0, \nabla u(y)) \, dy - C\delta.$$

By the arbitrariness of  $\delta$  we obtain the lower bound. Since the limit functional is independent of any subsequence, we established the full  $\Gamma$ -convergence result.  $\square$

**Remark 5.14.** Taking the same null set  $N \subset D$  as in Lemma 5.13, the convergence (45) holds for all  $x_0 \in D \setminus N$  and all  $\xi \in \mathbb{R}^d$  again by (44).

Now we are almost in a position to use a diagonal sequence to recover the function  $q(x_0, \xi)$ . However, in general there exists no metric characterizing the  $\Gamma$ -convergence when equicoercivity fails, so that diagonal arguments are not always available. Therefore we provide an explicit construction similar to [37] in the appendix. With this metric at hand we can derive the following result.

**Lemma 5.15.** *Under the assumptions of Lemma 5.12, let  $x_0 \in D \setminus N$  where  $N$  is given by Lemma 5.13. For every  $\varrho$  there exists  $\varepsilon(\varrho) > 0$  such that whenever we chose  $\varepsilon_{n(\varrho)} \leq \varepsilon(\varrho)$  it holds that*

$$\Gamma(L^p(B_1, \mathbb{R}^m))\text{-}\lim_{\varrho \rightarrow 0} G_{\varepsilon_{n(\varrho)}, \varrho}(x_0, \omega)(u) = \begin{cases} \int_{B_1} q(x_0, \nabla u(y)) \, dy & \text{if } u \in W^{1,p}(B_1, \mathbb{R}^m), \\ +\infty & \text{otherwise.} \end{cases}$$

**Proof.** This result is an immediate consequence of Lemma A.1 and Remark A.2, which allow to combine Lemma 5.13 and a diagonal argument with respect to the metric  $\mathfrak{d}$  constructed in the appendix.  $\square$

We now come to the second part and derive Lipschitz-estimates from bounds on the gradient's discrete maximal function. For the remainder of this subsection it will be convenient to view a stochastic lattice also as an undirected graph  $G = (\mathcal{L}(\omega), \mathbb{B}(\omega))$  with edges  $\mathbb{B}(\omega) = \mathcal{E}(\omega) \cup \{(y, x) \in \mathcal{L}(\omega)^2 : (x, y) \in \mathcal{E}(\omega)\}$ . We say that  $P$  is a path of length  $n$  if  $P = \{x_0, \dots, x_n\}$  with  $(x_{i-1}, x_i) \in \mathbb{B}(\omega)$  for all  $i = 1, \dots, n$ . Given  $x, y \in \mathcal{L}(\omega)$ , we define the graph distance as

$$d_G(x, y) = \inf\{\text{length of a path } P \text{ such that } x, y \in P\}.$$

We denote by  $B_G(x, \eta) = \{y \in \mathcal{L}(\omega) : d_G(x, y) \leq \eta\}$  the closed ball with radius  $\eta$  with respect to the graph metric. In the next lemma we establish a doubling property of the counting measure and a weak Poincaré inequality that allow us to relate Lipschitz continuity to discrete maximal functions of gradients. Given  $\varepsilon > 0$  and  $u : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  we define the length of its edge gradient  $|\nabla_{e, \varepsilon} u| : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}$  by

$$|\nabla_{e, \varepsilon} u|(\varepsilon x) = \sum_{(x, y) \in \mathbb{B}(\omega)} \left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right|. \quad (49)$$

Note that in the above definition we take into account the undirected edges  $e \in \mathbb{B}(\omega)$  (hence the subscript).

**Lemma 5.16.** *Let  $G = (\mathcal{L}(\omega), \mathbb{B}(\omega))$  be a graph associated to an admissible stochastic lattice. Then there exists a constant  $C = C(r, R, M) > 0$  such that for all  $x \in \mathcal{L}(\omega)$ ,  $\eta > 0$  and  $u : \mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  it holds that*

$$(i) \#B_G(x, 2\eta) \leq C\#B_G(x, \eta),$$

(ii)

$$\sum_{y \in B_G(x, \eta)} |u(y) - u_{B_G(x, \eta)}| \leq C\eta \sum_{y \in B_G(x, C\eta)} |\nabla_{e,1} u|(y),$$

where the average  $u_{B_G(x, \eta)}$  is defined as

$$u_{B_G(x, \eta)} = \frac{1}{\#B_G(x, \eta)} \sum_{y \in B_G(x, \eta)} u(y).$$

**Remark 5.17.** Given  $\eta > 0$  and a scalar function  $v : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}$  we define the maximal function  $\mathcal{M}_\eta^\varepsilon v : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}$  as

$$\mathcal{M}_\eta^\varepsilon v(\varepsilon x) = \sup_{0 < s < \eta} \left( \frac{1}{\#B_G(x, \frac{s}{\varepsilon})} \sum_{y \in B_G(x, \frac{s}{\varepsilon})} |v(\varepsilon y)| \right). \quad (50)$$

Then, assuming the doubling property, the Poincaré inequality is equivalent to the estimate

$$\left| \frac{u(\varepsilon x) - u(\varepsilon y)}{\varepsilon} \right| \leq \bar{C} d_G(x, y) \left( \mathcal{M}_{\bar{C}\varepsilon d_G(x, y)}^\varepsilon |\nabla_{e, \varepsilon} u|(\varepsilon x) + \mathcal{M}_{\bar{C}\varepsilon d_G(x, y)}^\varepsilon |\nabla_{e, \varepsilon} u|(\varepsilon y) \right), \quad (51)$$

where  $\bar{C}$  is a constant independent of  $u : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  and  $x, y \in \mathcal{L}(\omega)$ . For  $\varepsilon, m = 1$  this fact can be found in a much more general context in [46, Lemma 5.15]. For  $\varepsilon > 0$  and  $m = 1$  it follows by applying the inequality to  $v : \mathcal{L}(\omega) \rightarrow \mathbb{R}$  defined as  $v(x) = \varepsilon^{-1}u(\varepsilon x)$  upon noticing that  $\mathcal{M}_\eta^1 |\nabla_{e,1} v|(x) = \mathcal{M}_{\varepsilon\eta}^\varepsilon |\nabla_{e, \varepsilon} u|(\varepsilon x)$ . In particular the constant  $\bar{C}$  is independent of  $\varepsilon$ . For  $m \geq 2$  the inequality remains true arguing for each component and increasing the constant  $\bar{C}$  by at most a factor of  $m$ .

**Proof of Lemma 5.16.** (i): We may assume that  $\eta \geq \frac{1}{2}$ . Our aim is to compare the graph-metric with the Euclidean one. Given  $x, y \in \mathcal{L}(\omega)$  and an optimal path  $P = \{x_0 = x, x_1, \dots, x_n = y\}$ , by (8) we have

$$|x - y| \leq \sum_{i=1}^{d_G(x, y)} |x_{i-1} - x_i| < M d_G(x, y). \quad (52)$$

On the other hand, it follows from (9) and Lemma 5.1 that

$$d_G(x, y) \leq C_{r,R} |x - y|. \quad (53)$$

Using again Remark 2.3 and (52), for  $\eta \geq \frac{1}{2}$  we deduce that

$$\#B_G(x, 2\eta) \leq \#(\mathcal{L}(\omega) \cap B_{2M\eta}(x)) \leq |B_{\frac{r}{2}}|^{-1} |B_{4M\eta}(x)|,$$

while due to (53) for any  $\varrho > 0$  it holds that

$$|B_\varrho(x)| \leq |B_R(0)| \#(\mathcal{L}(\omega) \cap B_{2\varrho}(x)) \leq |B_R(0)| \#B_G(x, 2C_{r,R}\varrho). \quad (54)$$

The claim now follows from the scaling properties of the Lebesgue measure by choosing  $\varrho = \frac{\eta}{2C_{r,R}}$ .

(ii): We can assume that  $\eta \geq 1$ . Due to the triangle inequality we have

$$\sum_{y \in B_G(x, \eta)} |u(y) - u_{B_G(x, \eta)}| \leq \frac{1}{\#B_G(x, \eta)} \sum_{y, z \in B_G(x, \eta)} |u(y) - u(z)|. \quad (55)$$

Fix  $y, z \in B_G(x, \eta)$  and consider the path  $P(y, z) = \{x_0 = y, x_1, \dots, x_n = z\}$  given by Lemma 5.1. Then

$$|u(y) - u(z)| \leq \sum_{i=1}^n |u(x_{i-1}) - u(x_i)|. \quad (56)$$

The triangle inequality, Lemma 5.1 and (52) imply that for all  $x_i \in P(y, z)$

$$\begin{aligned} d_G(x_i, x) &\leq d_G(x_i, y) + d_G(y, x) \leq \#P(y, z) + \eta \leq C_{r,R}|y - z| + \eta \\ &\leq C_{r,R}(|y - x| + |x - z|) + \eta \leq (2MC_{r,R} + 1)\eta. \end{aligned}$$

Setting  $C = 2MC_{r,R} + 1$ , we deduce that  $P(y, z) \subset B_G(x, C\eta)$ . Conversely, given an edge  $(x', y') \in \mathbb{B}(\omega)$  with  $x', y' \in B_G(x, C\eta)$ , we denote by

$$N(x', y') = \#\{(y, z) \in B_G(x, \eta) \times B_G(x, \eta) : x', y' \subset P(y, z)\}$$

the number of pairs  $(y, z)$  such that this edge is contained in the path given by Lemma 5.1. As a consequence of (54), (55), (56) and (9) we have

$$\begin{aligned} \sum_{y \in B_G(x, \eta)} |u(y) - u_{B_G(x, \eta)}| &\leq \frac{1}{\#B_G(x, \eta)} \sum_{\substack{(x', y') \in \mathbb{B}(\omega) \\ x', y' \in B_G(x, C\eta)}} N(x', y') |u(x') - u(y')| \\ &\leq C \sup_{\substack{(x', y') \in \mathbb{B}(\omega) \\ x', y' \in B_G(x, C\eta)}} \frac{N(x', y')}{\eta^d} \sum_{y \in B_G(x, C\eta)} |\nabla_{e,1} u|(y). \end{aligned} \quad (57)$$

It remains to prove a suitable upper bound for  $N(x', y')$ . Since  $x', y' \in B_G(x, C\eta)$ , it follows by (52) that  $x', y' \in B_{MC\eta}(x)$  and therefore

$$N(x', y') \leq \#\{(y, z) \in B_{2MC\eta}(x') \times B_{2MC\eta}(x') : x', y' \subset P(y, z)\},$$

where we used that  $C \geq 1$ . Consider then the boundary-like set  $\Gamma = \{b \in \mathcal{L}(\omega) : \mathcal{C}(b) \cap \partial B_{2MC\eta}(x') \neq \emptyset\}$  and for each  $b \in \Gamma$  let us define the cylinder-type set

$$Z(b, x') = \{a + \lambda(x' - b) : \lambda \in [0, 2], a \in B_{6R}(b)\}.$$

For the moment fix any  $y, z \in B_{2MC\eta}(x')$  such that  $x', y' \in P(y, z)$ . By Lemma 5.1 there exists a point  $x_* \in \text{co}(y, z)$  such that  $|x_* - x'| \leq 2R$ . Without loss of generality we assume that  $|y - x_*| \leq |z - x_*|$  and denote by  $p$  the unique point  $p \in \{x' + \lambda(z - x') : \lambda \geq 0\} \cap \partial B_{2MC\eta}(x')$ . Then we find  $b \in \Gamma$  with  $|p - b| \leq R$ . We argue that  $y, z \in Z(b, x')$ . For  $z$  this follows upon writing  $z = p + \lambda_z(x' - p)$  with  $\lambda_z \in [0, 1]$  and choosing  $a = \lambda_z b + (1 - \lambda_z)p$  and  $\lambda = \lambda_z$  in the definition of  $Z(b, x')$ . Regarding  $y$ , recall that we assume  $|y - x_*| \leq |z - x_*|$ , so that for some  $\lambda_y \in [1, 2]$  we can write

$$y = z + \lambda_y(x_* - z) = p + \lambda_z(x' - p) + \lambda_y(x_* - p - \lambda_z(x' - p)).$$

Setting  $\lambda = \lambda_z + \lambda_y - \lambda_z \lambda_y \in [0, 2]$  and using the ansatz  $a = b + \xi$  in the definition of  $Z(b, x')$ , we find

$$\xi = (p - b)(1 - \lambda_y) + \lambda_y(x_* - x'),$$

so that  $|\xi| \leq 5R$  and therefore  $y \in Z(b, x')$  as well. Thus we have proven that

$$N(x', y') \leq 2(\#\Gamma) \left( \sup_{b \in \Gamma} \#(\mathcal{L}(\omega) \cap Z(b, x')) \right)^2. \quad (58)$$

Now we use again Remark 2.3 combined with a covering argument and the fact that  $\eta \geq 1$  to find

$$\#\Gamma \leq |B_{\frac{r}{2}}(0)|^{-1} \left( |B_{2MC\eta+R}(x')| - |B_{2MC\eta-R}(x')| \right) \leq C \left( (2MC\eta + R)^d - (2MC\eta - R)^d \right) \leq C\eta^{d-1}.$$

With the same technique we obtain the bound

$$\#(\mathcal{L}(\omega) \cap Z(b, x')) \leq 2|B_{\frac{r}{2}}(0)|^{-1} (14R)^{d-1} (|x' - b| + 7R) \leq C\eta.$$

Combining the last two estimates with (57) and (58) we conclude the proof.  $\square$

The next lemma is a discrete analogue of [41, Lemma 1.2].

**Lemma 5.18.** *Let  $G = (\mathcal{L}(\omega), \mathbb{B}(\omega))$  be admissible, let  $x_0 \in \mathbb{R}^d$ ,  $\lambda > 0$  and set  $\bar{k} = 3 + 6\bar{C}C_{r,R}M$  with  $\bar{C}$  and  $C_{r,R}$  given by (51) and Lemma 5.1, respectively. If  $\varepsilon \downarrow 0$  and  $u_\varepsilon : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  is a sequence such that*

$$\sup_{\varepsilon > 0} \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap B_{\bar{k}\lambda}(x_0)} \varepsilon^d |\nabla_{e,\varepsilon} u_\varepsilon|^p(\varepsilon x) \leq C,$$

*then there exists a subsequence  $\varepsilon_j$  and a sequence  $w_j : \varepsilon_j\mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  such that  $|\nabla_{e,\varepsilon_j} w_j|^p$  is equiintegrable on  $B_{2\lambda}(x_0)$  and moreover*

$$\lim_j \varepsilon_j^d \# \{ \varepsilon_j x \in \varepsilon_j\mathcal{L}(\omega) \cap B_{2\lambda}(x_0) : u_{\varepsilon_j} \not\equiv w_j \text{ on } \varepsilon_j B_G(x, 1) \} = 0.$$

**Proof.** Define a function  $V_\varepsilon : \varepsilon\mathcal{L}(\omega) \rightarrow [0, +\infty)$  by

$$V_\varepsilon(\varepsilon x) = \begin{cases} |\nabla_{e,\varepsilon} u_\varepsilon|(\varepsilon x) & \text{if } \varepsilon x \in B_{\bar{k}\lambda}(x_0), \\ 0 & \text{otherwise.} \end{cases}$$

By piecewise constant interpolation on the Voronoi cells  $\{\mathcal{E}\mathcal{C}(x)\}$  of  $\varepsilon\mathcal{L}(\omega)$  we can view  $V_\varepsilon$  as an element of  $L^p(\mathbb{R}^d)$ . From our assumption we deduce that

$$\|V_\varepsilon\|_{L^p(\mathbb{R}^d)}^p \leq C \sum_{\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap B_{\bar{k}\lambda}(x_0)} \varepsilon^d |\nabla_{e,\varepsilon} u_\varepsilon|^p(\varepsilon x) \leq C.$$

Hence the sequence  $V_\varepsilon$  is bounded in  $L^p(\mathbb{R}^d)$ . We claim that  $\mathcal{M}_\infty^\varepsilon V_\varepsilon$  is also bounded in  $L^p(\mathbb{R}^d)$ , where  $\mathcal{M}_\infty^\varepsilon V_\varepsilon$  is the discrete maximal function of  $V_\varepsilon$  defined in (50). To this end, we show that it can be pointwise controlled by the standard Hardy–Littlewood maximal function. Indeed, by (52), (54) and Remark 2.3, for any function  $v : \varepsilon\mathcal{L}(\omega) \rightarrow \mathbb{R}$  we can estimate

$$\begin{aligned} |\mathcal{M}_\infty^\varepsilon v(\varepsilon x)| &\leq \sup_{\eta > \frac{\varepsilon}{2}} \left( \frac{1}{\varepsilon^d \# B_G(x, \frac{\eta}{\varepsilon})} \sum_{y \in B_G(x, \frac{\eta}{\varepsilon})} \varepsilon^d |v(\varepsilon y)| \right) \leq \sup_{\eta > \frac{\varepsilon}{2}} \left( \frac{C}{|B_\eta(x)|} \sum_{y \in B_{M\frac{\eta}{\varepsilon}}(x)} \varepsilon^d |v(\varepsilon y)| \right) \\ &\leq \sup_{\eta > \frac{\varepsilon}{2}} \left( \frac{C}{|B_\eta(\varepsilon x)|} \int_{B_{3M\eta}(\varepsilon x)} |v(z)| \, dz \right) \leq C \sup_{\eta > 0} \left( \frac{1}{|B_\eta(\varepsilon x)|} \int_{B_\eta(\varepsilon x)} |v(z)| \, dz \right). \end{aligned}$$

Thus boundedness of the Hardy–Littlewood maximal function operator on  $L^p(\mathbb{R}^d)$  (see [40, Theorem 2.91]) implies that

$$\|\mathcal{M}_\infty^\varepsilon V_\varepsilon\|_{L^p(\mathbb{R}^d)}^p \leq C \|V_\varepsilon\|_{L^p(\mathbb{R}^d)}^p \leq C. \quad (59)$$

For  $l > 0$  we introduce the scalar truncation operator  $\delta_l : \mathbb{R} \rightarrow \mathbb{R}$  given by  $\delta_l(x) = \min\{l, |x|\} \frac{x}{|x|}$  (with  $\delta_l(0) = 0$ ). Applying [40, Lemma 2.31], we know that there exists a subsequence  $\varepsilon_j$  and an increasing sequence of positive integers  $l_j^p \rightarrow +\infty$  such that the sequence  $\delta_{l_j^p} \circ (|\mathcal{M}_\infty^{\varepsilon_j} V_{\varepsilon_j}|^p) = (\delta_{l_j} \circ |\mathcal{M}_\infty^{\varepsilon_j} V_{\varepsilon_j}|)^p$  is equiintegrable on  $\mathbb{R}^d$ . Define the sets

$$\begin{aligned} R'_j &:= \{x \in \mathcal{L}(\omega) : \mathcal{M}_\infty^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j x) \leq l_j\}, \\ R_j &:= \bigcup_{x \in R'_j} B_G(x, 1). \end{aligned}$$

Note that  $R'_j \subset R_j$ . By Remark 2.3 and (59) we have

$$\varepsilon_j^d \# (\mathcal{L}(\omega) \setminus R_j) \leq \varepsilon_j^d \# (\mathcal{L}(\omega) \setminus R'_j) \leq \frac{C}{l_j^p} \int_{\mathbb{R}^d} |\mathcal{M}_\infty^{\varepsilon_j} V_{\varepsilon_j}(z)|^p \, dz \leq \frac{C}{l_j^p}. \quad (60)$$

Next we bound  $\mathcal{M}_\infty^{\varepsilon_j} V_{\varepsilon_j}$  on the set  $R_j$ . Given  $y \in B_G(x, 1)$  with  $x \in R'_j$ , for  $\eta \geq 2\varepsilon_j$  it holds that

$$B_G\left(x, \frac{\eta}{2\varepsilon_j}\right) \subset B_G\left(x, \frac{\eta}{\varepsilon_j} - 1\right) \subset B_G\left(y, \frac{\eta}{\varepsilon_j}\right) \subset B_G\left(x, \frac{\eta}{\varepsilon_j} + 1\right) \subset B_G\left(x, \frac{2\eta}{\varepsilon_j}\right).$$

Hence applying twice the doubling property proven in Lemma 5.16 (i) we conclude that

$$\sup_{\eta \geq 2\varepsilon_j} \left( \frac{1}{\#B_G(y, \frac{\eta}{\varepsilon_j})} \sum_{z \in B_G(y, \frac{\eta}{\varepsilon_j})} |V_{\varepsilon_j}(\varepsilon_j z)| \right) \leq C \left( \mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j x) \right) \leq Cl_j.$$

If  $\eta < 2\varepsilon_j$ , then by Remark 2.3 and (8) we still have the inequality and inclusion

$$\frac{1}{C} \#B_G(x, 2) \leq \#B_G\left(y, \frac{\eta}{\varepsilon_j}\right) \quad \text{and} \quad B_G\left(y, \frac{\eta}{\varepsilon_j}\right) \subset B_G(x, 2),$$

so that for  $y \in R_j$  we obtain the estimate

$$\mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j y) \leq Cl_j. \quad (61)$$

Next we want to use (51) and Kirsztbraun's extension theorem on  $\varepsilon_j \mathcal{L}(\omega) \cap B_{3\lambda}(x_0)$ . To this end, note that by (52) and (53), for all  $x, y \in R_j \cap \varepsilon_j^{-1} B_{3\lambda}(x_0)$  and  $z \in B_G(y, \overline{C}d_G(x, y))$ , it holds that

$$|\varepsilon_j z - x_0| < 3\lambda + \varepsilon_j |z - y| \leq 3\lambda + \varepsilon_j M \overline{C} d_G(x, y) \leq 3\lambda + 6\lambda M \overline{C} C_{r,R} = \bar{k}\lambda.$$

Since  $V_{\varepsilon_j}$  agrees with  $|\nabla_{e,\varepsilon_j} u_{\varepsilon_j}|$  on  $\varepsilon_j \mathcal{L}(\omega) \cap B_{\bar{k}\lambda}(x_0)$ , it follows from (61) that for  $\varepsilon_j y \in \varepsilon_j R_j \cap B_{3\lambda}(x_0)$  we have the estimate

$$\mathcal{M}_{\overline{C}\varepsilon_j d_G(x,y)}^{\varepsilon_j} |\nabla_{e,\varepsilon_j} u_{\varepsilon_j}|(\varepsilon_j y) = \mathcal{M}_{\overline{C}\varepsilon_j d_G(x,y)}^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j y) \leq \mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j y) \leq Cl_j.$$

Combining (51) and Kirsztbraun's extension theorem we find Lipschitz-functions  $w_j : \varepsilon_j \mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  that agree with  $u_{\varepsilon_j}$  on the set  $\varepsilon_j R_j \cap B_{3\lambda}(x_0)$  and with Lipschitz constant bounded by  $Cl_j$  (up to enlarging  $C$  due to (53)). We claim that  $|\nabla_{e,\varepsilon_j} w_j|^p$  is equiintegrable on  $B_{2\lambda}(x_0)$ . To verify this assertion, we observe that for  $j = j(\lambda)$  large enough, by the definition of  $R_j$  we have that  $|\nabla_{e,\varepsilon_j} w_j| = |\nabla_{e,\varepsilon_j} u_{\varepsilon_j}|$  on  $\varepsilon_j R'_j \cap B_{5\lambda/2}(x_0)$ . Hence for  $\varepsilon_j x \in \varepsilon_j R'_j \cap B_{5\lambda/2}(x_0)$  we deduce

$$|\nabla_{e,\varepsilon_j} w_j|^p(\varepsilon_j x) = |\nabla_{e,\varepsilon_j} u_{\varepsilon_j}|^p(\varepsilon_j x) = |V_{\varepsilon_j}(\varepsilon_j x)|^p \leq |\mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}(\varepsilon_j x)|^p = (\delta_{l_j} \circ |\mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}|)^p(\varepsilon_j x),$$

while on  $\varepsilon_j \left( \mathcal{L}(\omega) \setminus R'_j \right)$  the bound on the Lipschitz constant implies

$$|\nabla_{e,\varepsilon_j} w_j|^p(\varepsilon_j x) \leq Cl_j^p = C(\delta_{l_j} \circ |\mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}|)^p(\varepsilon_j x).$$

Hence equiintegrability on  $B_{2\lambda}(x_0)$  transfers from  $(\delta_{l_j} \circ |\mathcal{M}_{\infty}^{\varepsilon_j} V_{\varepsilon_j}|)^p$  to  $|\nabla_{e,\varepsilon_j} w_j|^p$ . Finally, note that

$$\{\varepsilon_j x \in \varepsilon_j \mathcal{L}(\omega) \cap B_{2\lambda}(x_0) : u_{\varepsilon_j} \neq w_j \text{ on } \varepsilon_j B_G(x, 1)\} \subset \varepsilon_j \left( \mathcal{L}(\omega) \setminus R'_j \right),$$

so that the second claim of the lemma follows from (60).  $\square$

Now we are finally in a position to compare the two discrete functionals  $F_{\varepsilon}(\omega)$  and  $E_{\varepsilon}(\omega)$ .

**Proposition 5.19** (Separation of bulk effects). *Let  $\varepsilon_n$  and  $F(\omega)$  be as in Proposition 5.2. Then for a.e.  $x_0 \in D$  and every  $\xi \in \mathbb{R}^{m \times d}$  it holds that*

$$|B_1|h(x_0, \xi) = \lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_{\varrho}(x_0)) = |B_1|q(x_0, \xi),$$

where  $q$  is an (equivalent) integrand given by the  $\Gamma$ -limit of  $E_{\varepsilon_n}(\omega)(\cdot, D)$ , which in particular exists.

**Proof.** The first equality follows from Lemma 5.11, so we turn to the proof of the second one. We apply Theorem 3.1, so that, passing to a further subsequence (not relabeled), we may assume that  $E_{\varepsilon_n}(\omega)$   $\Gamma$ -converges to some integral functional  $E(\omega)$  with density  $q(x, \xi)$ . Let us fix  $x_0 \in D$  satisfying the first equality and such that Lemmata 5.13 and 5.15 hold. Choose  $0 < \varrho_0 < 1$  such that  $B_{\varrho_0}(x_0) \subset D$ . Lemma 5.4 yields a sequence  $u_{\varepsilon_n} \in \mathcal{PC}_{\varepsilon_n}^{\omega}$  that is equibounded in  $L^{\infty}(\mathbb{R}^d, \mathbb{R}^m)$ ,  $u_{\varepsilon_n} \rightarrow u_{x_0, \xi}$  in  $L^p(D, \mathbb{R}^m)$  and

$$\lim_{\varepsilon_n \rightarrow 0} F_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho_0}(x_0)) = F(\omega)(u_{x_0, \xi}, B_{\varrho_0}(x_0)).$$

Now consider  $0 < \varrho < \varrho_0$ . By discrete superadditivity and  $\Gamma$ -convergence on  $\mathcal{A}^R(D)$  we find that

$$\begin{aligned} \limsup_{\varepsilon_n \rightarrow 0} F_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho}(x_0)) &\leq \lim_{\varepsilon_n \rightarrow 0} F_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho_0}(x_0)) - \liminf_{\varepsilon_n \rightarrow 0} F_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho_0}(x_0) \setminus \overline{B_{\varrho}(x_0)}) \\ &\leq F(\omega)(u_{x_0, \xi}, B_{\varrho_0}(x_0)) - F(\omega)(u_{x_0, \xi}, B_{\varrho_0}(x_0) \setminus \overline{B_{\varrho}(x_0)}) \\ &= F(\omega)(u_{x_0, \xi}, B_{\varrho}(x_0)), \end{aligned} \quad (62)$$

where in the last equality we used that the limit energy of  $u_{x_0, \xi}$  does not concentrate on  $\partial B_{\varrho}(x_0)$ . (62) shows that  $u_{\varepsilon_n}$  is also a recovery sequence on each  $B_{\varrho}(x_0)$  for  $0 < \varrho < \varrho_0$ . Next we introduce a constant whose value will become clear later in the proof (cf. the constant  $k$  in Lemma 5.18). Choose  $k$  satisfying

$$3 + 6\overline{C} C_{r,R} M + |\xi| \leq k,$$

where  $\overline{C}$  and  $C_{r,R}$  are given by (51) and Lemma 5.1, respectively. Since  $|u_{x_0, \xi}| \leq |\xi| \varrho$  on  $B_{\varrho}(x_0)$ , Lemma 5.4 implies that the truncated functions  $T_{k\varrho} u_{\varepsilon_n}$  also yield a recovery sequence on  $B_{\varrho}(x_0)$ . Now consider a sequence  $\varrho_j \rightarrow 0$ . For any  $\varrho = \varrho_j \in (0, (3Mk^2)^{-1}\varrho_0)$  we choose  $\varepsilon_{\varrho} = \varepsilon_{n(\varrho)} \leq \min\{\varrho^{\frac{p}{p-1}}, \varrho^2\}$  non-decreasing in  $\varrho$ , satisfying Lemma 5.15 and such that

$$\begin{aligned} F_{\varepsilon_{\varrho}}(\omega)(T_{k\varrho} u_{\varepsilon_{\varrho}}, B_{3Mk^2\varrho}(x_0)) &\leq C|\xi|^p \varrho^d, \\ \int_{B_{\varrho}(x_0)} |T_{k\varrho} u_{\varepsilon_{\varrho}} - u_{x_0, \xi}|^p dx &\leq \varrho^{p+1}. \end{aligned} \quad (63)$$

Note that the first estimate is realizable due to Lemma 5.4 and the fact that  $u_{\varepsilon_n}$  is a recovery sequence also on  $B_{3Mk^2\varrho}(x_0)$ . Finally, we can also require that

$$\lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_{\varrho}(x_0)) = \lim_{\varrho \rightarrow 0} \varrho^{-d} F_{\varepsilon_{\varrho}}(\omega)(T_{k\varrho} u_{\varepsilon_{\varrho}}, B_{\varrho}(x_0)). \quad (64)$$

Our analysis relies on several modifications of the sequence  $T_{k\varrho} u_{\varepsilon_{\varrho}}$  and a rescaling to  $B_1$ .

### Step 1 Construction of Lipschitz competitors

The following argument is well-known for continuum functionals and we adapt it carefully to the discrete setting. Let us set  $v_{\varrho} : \mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  as  $v_{\varrho}(x) = \varepsilon_{\varrho}^{-1} T_{k\varrho} u_{\varepsilon_{\varrho}}(\varepsilon_{\varrho} x)$ . Then, for given  $\Lambda > 0$ , we define the sets

$$R_{\varrho}^{\Lambda} := \{x \in \mathcal{L}(\omega) \cap \varepsilon_{\varrho}^{-1} B_{k\varrho}(x_0) : \mathcal{M}_{k^2\varrho\varepsilon_{\varrho}^{-1}}^1 |\nabla_{e,1} v_{\varrho}|(x) \leq \Lambda\},$$

$$S_{\varrho}^{\Lambda} := \{x \in \mathcal{L}(\omega) : |\nabla_{e,1} v_{\varrho}|(x) \geq \frac{\Lambda}{2}\},$$

where  $\mathcal{M}_{\eta}^1$  denotes the discrete maximal function operator defined in (50) and the norm of the discrete gradient is given by the formula (49). First we estimate the cardinality of  $(\mathcal{L}(\omega) \cap \varepsilon_{\varrho}^{-1} B_{k\varrho}(x_0)) \setminus R_{\varrho}^{\Lambda}$ . To this end, note that for every  $x \in (\mathcal{L}(\omega) \cap \varepsilon_{\varrho}^{-1} B_{k\varrho}(x_0)) \setminus R_{\varrho}^{\Lambda}$  there exists a number  $0 < \eta_x \leq k^2 \varrho \varepsilon_{\varrho}^{-1}$  such that

$$\Lambda \# B_G(x, \eta_x) < \sum_{y \in B_G(x, \eta_x)} |\nabla_{e,1} v_{\varrho}|(y) =: |\nabla_1 v_{\varrho}|(B_G(x, \eta_x)).$$

Applying Vitali's covering lemma on separable metric spaces we find a (finite) collection of disjoint balls  $B_G(x_i, \eta_i)$  with  $x_i \in (\mathcal{L}(\omega) \cap \varepsilon_{\varrho}^{-1} B_{k\varrho}(x_0)) \setminus R_{\varrho}^{\Lambda}$  satisfying the above inequality and

$$(\mathcal{L}(\omega) \cap \varepsilon_{\varrho}^{-1} B_{k\varrho}(x_0)) \setminus R_{\varrho}^{\Lambda} \subset \bigcup_i B_G(x_i, 5\eta_i).$$

Since the balls are disjoint we conclude that

$$\Lambda \# \left( \bigcup_i B_G(x_i, \eta_i) \right) < |\nabla_1 v_{\varrho}| \left( \bigcup_i B_G(x_i, \eta_i) \right) \leq |\nabla_1 v_{\varrho}| \left( \bigcup_i B_G(x_i, \eta_i) \cap S_{\varrho}^{\Lambda} \right) + \frac{\Lambda}{2} \# \left( \bigcup_i B_G(x_i, \eta_i) \right),$$



where we used the definition of  $S_Q^\Lambda$  in the last estimate. Rearranging terms we obtain

$$\# \left( \bigcup_i B_G(x_i, \eta_i) \right) < \frac{2}{\Lambda} |\nabla_1 v_Q| \left( \bigcup_i B_G(x_i, \eta_i) \cap S_Q^\Lambda \right). \quad (65)$$

To reduce notation, for  $x \in \mathcal{L}(\omega)$  we set  $N_x = \{y \in \mathcal{L}(\omega) : (x, y) \in \mathbb{B}(\omega)\}$ . Moreover define

$$\mathcal{J}_Q = \{x \in \mathcal{L}(\omega) : \varepsilon_Q |\nabla_{e,1} v_Q|^p(x) \geq 1\}.$$

Note that for any  $x \in \mathcal{J}_Q$  there exists a point  $y \in N_x$  with  $\varepsilon_Q |v_Q(x) - v_Q(y)|^p \geq \frac{1}{M^p}$ . Hence the growth condition (14) implies the inequality

$$1 \leq C \left( f(\varepsilon_Q |\nabla_{\omega,1}(v_Q, \mathbb{R}^d)|^p(x)) + f(\varepsilon_Q |\nabla_{\omega,1}(v_Q, \mathbb{R}^d)|^p(y)) \right). \quad (66)$$

In order to control the location of such  $y$ , observe that by (52) we have  $B_G(x_i, \eta_i) \subset B_{M\eta_i}(x_i)$ , which in turn implies (for  $M \geq 1$ ) that

$$\bigcup_i \bigcup_{x \in B_G(x_i, \eta_i)} N_x \subset \varepsilon_Q^{-1} B_{3Mk^2_Q}(x_0). \quad (67)$$

Here we also used (8) and that  $Mk^2_Q \geq M\varepsilon_Q$ . Next we sum the estimate (66) over  $x$ . Note that due to (10) every term can appear at most  $M+1$  times. By the definition of  $v_Q$  we obtain

$$\varepsilon_Q^{d-1} \# \left( \bigcup_i B_G(x_i, \eta_i) \cap \mathcal{J}_Q \right) \leq C F_{\varepsilon_Q}(\omega)(T_{k_Q} u_{\varepsilon_Q}, B_{3Mk^2_Q}(x_0)) \leq C |\xi|^p \varrho^d,$$

where we applied the first bound in (63). By truncation we further know that  $|\nabla_{e,1} v_Q|(x) \leq C\varepsilon_Q^{-1} k_Q$  for all  $x \in \mathcal{L}(\omega)$ , so that

$$|\nabla_1 v_Q| \left( \bigcup_i B_G(x_i, \eta_i) \cap \mathcal{J}_Q \right) \leq C |\xi|^p \left( \frac{\varrho}{\varepsilon_Q} \right)^d \varrho. \quad (68)$$

In order to estimate the remaining contributions in the right hand side of (65) we use Hölder's inequality in the form

$$|\nabla_1 v_Q| \left( \bigcup_i B_G(x_i, \eta_i) \cap S_Q^\Lambda \setminus \mathcal{J}_Q \right) \leq \# \left( \bigcup_i B_G(x_i, \eta_i) \cap S_Q^\Lambda \setminus \mathcal{J}_Q \right)^{\frac{p-1}{p}} \left( \sum_{x \in \bigcup_i B_G(x_i, \eta_i) \setminus \mathcal{J}_Q} |\nabla_{e,1} v_Q|^p(x) \right)^{\frac{1}{p}}. \quad (69)$$

In the last term we have to pass from the undirected gradient to the directed version: for  $x \notin \mathcal{J}_Q$  and  $y \in N_x$  it holds that  $\varepsilon_Q |v_Q(x) - v_Q(y)|^p \leq 1$ . Hence we infer from the bound (14) that for  $x \in \bigcup_i B_G(x_i, \eta_i) \setminus \mathcal{J}_Q$

$$\begin{aligned} |\nabla_{e,1} v_Q(x)|^p(x) &\leq C \sum_{y \in N_x} |v_Q(x) - v_Q(y)|^p = \frac{C}{\varepsilon_Q} \sum_{y \in N_x} \min\{\varepsilon_Q |v_Q(x) - v_Q(y)|^p, 1\} \\ &\leq \frac{C}{\varepsilon_Q} \sum_{y \in N_x} f(\varepsilon_Q |\nabla_{\omega,1}(v_Q, \varepsilon_Q^{-1} B_{3Mk^2_Q}(x_0))|^p(x)) + f(\varepsilon_Q |\nabla_{\omega,1}(v_Q, \varepsilon_Q^{-1} B_{3Mk^2_Q}(x_0))|^p(y)), \end{aligned}$$

where we used again (67). We sum this estimate and by (10) each term is counted at most  $2M$  times. Thus in combination with the first estimate in (63) we have

$$\left( \sum_{x \in \bigcup_i B_G(x_i, \eta_i) \setminus \mathcal{J}_Q} |\nabla_{e,1} v_Q|^p(x) \right)^{\frac{1}{p}} \leq C \varepsilon_Q^{-\frac{d}{p}} \left( F_{\varepsilon_Q}(\omega)(T_{k_Q} u_{\varepsilon_Q}, B_{3Mk^2_Q}(x_0)) \right)^{\frac{1}{p}} \leq \left( \frac{\varrho}{\varepsilon_Q} \right)^{\frac{d}{p}} C |\xi|. \quad (70)$$

Combining this estimate with (69) leads to

$$|\nabla_1 v_Q| \left( \bigcup_i B_G(x_i, \eta_i) \cap S_Q^\Lambda \setminus \mathcal{J}_Q \right) \leq C \# \left( \bigcup_i B_G(x_i, \eta_i) \cap S_Q^\Lambda \setminus \mathcal{J}_Q \right)^{\frac{p-1}{p}} \left( \frac{\varrho}{\varepsilon_Q} \right)^{\frac{d}{p}} |\xi|.$$

In order to bound the cardinality term, note that by the definition of  $S_Q^\Lambda$  and (70) it holds that

$$\# \left( \bigcup_i B_G(x_i, \eta_i) \cap S_\varrho^\Lambda \setminus \mathcal{J}_\varrho \right) \left( \frac{\Lambda}{2} \right)^p \leq \sum_{x \in \bigcup_i B_G(x_i, \eta_i) \setminus \mathcal{J}_\varrho} |\nabla_{e,1} v_\varrho|^p(x) \leq C |\xi|^p \left( \frac{\varrho}{\varepsilon_\varrho} \right)^d.$$

Plugging this estimate into the previous one yields

$$|\nabla_1 v_\varrho| \left( \bigcup_i B_G(x_i, \eta_i) \cap S_\varrho^\Lambda \setminus \mathcal{J}_\varrho \right) \leq C \Lambda^{1-p} |\xi|^p \left( \frac{\varrho}{\varepsilon_\varrho} \right)^d. \quad (71)$$

Applying twice the doubling property of Lemma 5.16 and combining (65), (68) and (71) we infer that

$$\#((\mathcal{L}(\omega) \cap \varepsilon_\varrho^{-1} B_{k\varrho}(x_0)) \setminus R_\varrho^\Lambda) \leq \# \left( \bigcup_i B_G(x_i, 5\eta_i) \right) \leq C \# \left( \bigcup_i B_G(x_i, \eta_i) \right) \leq C |\xi|^p \left( \frac{\varrho}{\varepsilon_\varrho} \right)^d \left( \varrho \Lambda^{-1} + \Lambda^{-p} \right).$$

We choose  $\Lambda = \Lambda_\varrho$  as  $\Lambda^{p-1} = \varrho^{-1}$ , so that the last inequality can be written as

$$\#((\mathcal{L}(\omega) \cap \varepsilon_\varrho^{-1} B_{k\varrho}(x_0)) \setminus R_\varrho^{\Lambda_\varrho}) \leq C |\xi|^p \left( \frac{\varrho}{\varepsilon_\varrho} \right)^d \Lambda_\varrho^{-p}. \quad (72)$$

With this choice of  $\Lambda_\varrho$ , we now construct the Lipschitz competitor. First observe that for any  $x, y \in \mathcal{L}(\omega) \cap \varepsilon_\varrho^{-1} B_{k\varrho}(x_0)$  the definition of  $k$  yields

$$\overline{C} d_G(x, y) \leq \overline{C} C_{r,R} |x - y| \leq 2 \overline{C} C_{r,R} k \varrho \varepsilon_\varrho^{-1} \leq k^2 \varrho \varepsilon_\varrho^{-1},$$

so that (51) and (53) imply for any  $x, y \in R_\varrho^{\Lambda_\varrho}$  the Lipschitz estimate

$$|v_\varrho(x) - v_\varrho(y)| \leq 2 \overline{C} C_{r,R} \Lambda_\varrho |x - y| \leq k \Lambda_\varrho |x - y|.$$

Using Kirszbraun's extension theorem we find a Lipschitz function  $\tilde{v}_\varrho : \mathcal{L}(\omega) \rightarrow \mathbb{R}^m$  with Lipschitz constant  $k \Lambda_\varrho$  that agrees with  $v_\varrho$  on  $R_\varrho^{\Lambda_\varrho}$ . Moreover, by truncation via the operator  $T_{3\varepsilon_\varrho^{-1}k\varrho}$  we can additionally assume that  $\|\tilde{v}_\varrho\|_\infty \leq 9\varepsilon_\varrho^{-1}k\varrho$ .

**Step 2** From Lipschitz continuity to equiintegrability of discrete gradients

It will be convenient to rescale the function  $\tilde{v}_\varrho$  constructed in the first step onto  $B_1$ . First we introduce some notation. We set  $\sigma_\varrho = \frac{\varepsilon_\varrho}{\varrho}$  and  $\mathcal{L}(\omega)_\varrho := \mathcal{L}(\omega) - \frac{x_0}{\varepsilon_\varrho}$ . For any  $x \in \mathcal{L}(\omega)_\varrho$  we further denote by  $N_{x,\varrho} = N_{x+\frac{x_0}{\varepsilon_\varrho}} - \frac{x_0}{\varepsilon_\varrho}$  the set of adjacent points in the undirected shifted graph. In the notation for the discrete gradients we will replace  $e$  by  $e_0$  and  $\omega$  by  $\omega_0$ , respectively. Define  $u^\varrho : \sigma_\varrho \mathcal{L}(\omega)_\varrho \rightarrow \mathbb{R}^m$  via

$$u^\varrho(\sigma_\varrho x) = \sigma_\varrho \tilde{v}_\varrho \left( x + \frac{x_0}{\varepsilon_\varrho} \right).$$

By the properties of  $\tilde{v}_\varrho$  established in the first step, the function  $u^\varrho$  satisfies

- (i)  $\|u^\varrho\|_\infty \leq 9k$ ;
- (ii)  $|u^\varrho(\sigma_\varrho x) - u^\varrho(\sigma_\varrho y)| \leq k \Lambda_\varrho \sigma_\varrho |x - y|$  for all  $x, y \in \mathcal{L}(\omega)_\varrho$ ;
- (iii)  $u^\varrho(\sigma_\varrho x) = \sigma_\varrho v_\varrho(x + \frac{x_0}{\varepsilon_\varrho}) = \varrho^{-1} T_{k\varrho} u_{\varepsilon_\varrho}(\varepsilon_\varrho x + x_0)$  for all  $x \in R_\varrho^{\Lambda_\varrho} - \frac{x_0}{\varepsilon_\varrho}$ .

We aim at applying Lemma 5.18 with  $x_0 = 0$ ,  $\lambda = 1$  and the vanishing sequence  $\sigma_\varrho$  (note that the shift of the graph preserves admissibility). Due to (ii) we have the bound  $|\nabla_{e_0, \sigma_\varrho} u^\varrho|^p(\sigma_\varrho x) \leq C \Lambda_\varrho^p$ . In combination with (iii) and a change of variables we derive the bound

$$\begin{aligned} & \sum_{\substack{\sigma_\varrho x \in \sigma_\varrho \mathcal{L}(\omega)_\varrho \cap B_k \\ N_{x,\varrho} \setminus (R_\varrho^{\Lambda_\varrho} - \frac{x_0}{\varepsilon_\varrho}) \neq \emptyset}} \sigma_\varrho^d |\nabla_{e_0, \sigma_\varrho} u^\varrho|^p(\sigma_\varrho x) + \sum_{\substack{\sigma_\varrho x \in \sigma_\varrho \mathcal{L}(\omega)_\varrho \cap B_k \\ N_{x,\varrho} \subset (R_\varrho^{\Lambda_\varrho} - \frac{x_0}{\varepsilon_\varrho})}} \sigma_\varrho^d |\nabla_{e_0, \sigma_\varrho} u^\varrho|^p(\sigma_\varrho x) \\ & \leq C \Lambda_\varrho^p \sigma_\varrho^d \# \left( (\mathcal{L}(\omega) \cap \varepsilon_\varrho^{-1} B_{k\varrho}(x_0)) \setminus R_\varrho^{\Lambda_\varrho} \right) + \sum_{x \in R_{\varepsilon_\varrho}^{\Lambda_\varrho}} \sigma_\varrho^d |\nabla_{e,1} v_\varrho|^p(x). \end{aligned}$$

By definition of the maximal function operator it holds that  $\varepsilon_\varrho |\nabla_{e,1} v_\varrho|^p(x) \leq \varepsilon_\varrho \Lambda_\varrho^p \leq 1$  for  $x \in R_\varrho^{\Lambda_\varrho}$ , so that we can use (72) and the same reasoning as for (70) to obtain the estimate

$$\sum_{\sigma_\varrho x \in \sigma_\varrho \mathcal{L}(\omega)_\varrho \cap B_k} \sigma_\varrho^d |\nabla_{e,1} v_\varrho|^p(\sigma_\varrho x) \leq C |\xi|^p + C \varrho^{-d} F_\varepsilon(\omega)(T_{k\varrho} u_{\varepsilon_\varrho}, B_{3Mk^2\varrho}(x_0)) \leq C |\xi|^p.$$

Hence our choice of  $k$  allows to apply Lemma 5.18 and we obtain a subsequence  $\varrho_j$  and a sequence  $w_j : \sigma_{\varrho_j} \mathcal{L}(\omega)_{\varrho_j} \rightarrow \mathbb{R}^m$  such that, setting  $\sigma_j = \sigma_{\varrho_j}$ , the sequence  $|\nabla_{e,1} w_j|^p$  is equiintegrable on  $B_2$  and

$$\lim_j \sigma_j^d \# \{ \sigma_j x \in \sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_2 : u^{\varrho_j} \not\equiv w_j \text{ on } \sigma_j B_{G_{\varrho_j}}(x, 1) \} = 0, \quad (73)$$

where  $G_{\varrho_j}$  denotes the undirected shifted graph. Finally, by a truncation argument based on the operator  $T_{9k}$  we can assume that  $\|w_j\|_\infty \leq 27k$ .

**Step 3** Proof of  $\lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) \geq |B_1| q(x_0, \xi)$

Let  $w_j$  be the sequence constructed in Step 2. First we estimate the  $L^p(B_1)$ -norm of the sequence  $w_j - u_{0, \xi}$ . Define the set

$$U_j = \left\{ \sigma_j x \in \sigma_j \mathcal{L}(\omega)_{\varrho_j} : B_{G_{\varrho_j}}(x, 1) \subset R_{\varrho_j}^{\Lambda_{\varrho_j}} - x_0 \varepsilon_{\varrho_j}^{-1} \right\} \setminus R'_j,$$

where  $R'_j$  denotes the set in (73). Then by construction  $w_j(\sigma_j x) = \varrho_j^{-1} T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}(x_0 + \varepsilon_j x)$  for all  $\sigma_j x \in B_2 \cap U_j$ . Moreover, by (10) and (72) we can bound the cardinality of the complement via

$$\sigma_j^d \# (\sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_2 \setminus U_j) \leq \sigma_j^d \# R'_j + C \sigma_j^d \# \left( \mathcal{L}(\omega) \cap \varepsilon_{\varrho_j}^{-1} B_{k\varrho_j}(x_0) \setminus R_{\varrho_j}^{\Lambda_{\varrho_j}} \right) \leq \sigma_j^d \# R'_j + C |\xi|^p \Lambda_{\varrho_j}^p,$$

so that by (73) and the choice of  $\Lambda_{\varrho_j}$  we have

$$\lim_{j \rightarrow +\infty} \sigma_j^d \# (\sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_2 \setminus U_j) = 0. \quad (74)$$

Due to the  $L^\infty$ -bound on  $w_j$ , for  $j$  large enough a change of variables yields

$$\|w_j - u_{0, \xi}\|_{L^p(B_1)}^p \leq \frac{C}{\varrho_j^p} \int_{B_{\varrho_j}(x_0)} |T_{k\varrho_j} u_{\varepsilon_{\varrho_j}} - u_{x_0, \xi}|^p dz + C \sigma_j^d \# (\sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_2 \setminus U_j).$$

Hence (63) and (74) imply that  $w_j \rightarrow u_{0, \xi}$  in  $L^p(B_1, \mathbb{R}^m)$ . Now we turn to the energy estimates. Fix  $\eta > 0$ . Since  $|\nabla_{e,1} w_j|^p$  is piecewise constant with respect to the Voronoi tessellation of  $\sigma_j \mathcal{L}(\omega)_{\varrho_j}$ , Remark 2.3, (74) and the equiintegrability of  $|\nabla_{e,1} w_j|^p$  on  $B_2$  imply that there exists  $j_\eta$  such that for all  $j \geq j_\eta$

$$\sum_{\sigma_j x \in \sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_1 \setminus U_j} \sigma_j^d |\nabla_{e,1} w_j|^p(\sigma_j x) \leq \eta. \quad (75)$$

For  $t > 0$  let us further introduce the sets

$$S_j(t) = \{ \sigma_j x \in \sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_1 : |\nabla_{e,1} w_j|^p(\sigma_j x) > t \}.$$

Again due to the equiintegrability of  $|\nabla_{e,1} w_j|^p$  on  $B_2$  we find  $t_\eta > 0$  such that for  $j \geq j_\eta$  we have

$$\sum_{\sigma_j x \in S_j(t_\eta)} \sigma_j^d |\nabla_{e,1} w_j|^p(\sigma_j x) \leq C \int_{B_2 \cap \{ |\nabla_{e,1} w_j|^p > t_\eta \}} |\nabla_{e,1} w_j|^p(z) dz \leq \eta. \quad (76)$$

Moreover, if  $\varepsilon_{\varrho_j} x \in (x_0 + \varrho_j(B_1 \cap U_j \setminus S_j(t_\eta)))$ , then

$$\|\varepsilon_{\varrho_j} |\nabla_{\omega, \varepsilon_{\varrho_j}}(T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}, B_{\varrho_j}(x_0))|^p(\varepsilon_j x) - \varrho_j^{-1} x_0\|_1 \leq \varepsilon_{\varrho_j} |\nabla_{e,1} w_j|^p(\sigma_j x - \varrho_j^{-1} x_0) \leq \varepsilon_{\varrho_j} t_\eta. \quad (77)$$

The right hand side converges to zero. Thus, after enlarging  $j_\eta$ , assumption (11) yields

$$f(\varepsilon_{\varrho_j} |\nabla_{\omega, \varepsilon_{\varrho_j}} (T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}, B_{\varrho_j}(x_0))|^p (\varepsilon_j x)) \geq (1 - \eta) \alpha \|\varepsilon_{\varrho_j} |\nabla_{\omega, \varepsilon_{\varrho_j}} (T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}, B_{\varrho_j}(x_0))|^p (\varepsilon_{\varrho_j} x)\|_1$$

for all  $j \geq j_\eta$  and all  $\varepsilon_{\varrho_j} x \in (x_0 + \varrho_j(B_1 \cap U_j \setminus S_j(t_\eta)))$ . For the remaining lattice points we can use (75) and (76), so that from a change of variables we deduce the lower bound

$$\begin{aligned} \varrho_j^{-d} F_{\varepsilon_{\varrho_j}}(\omega)(T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}, B_{\varrho_j}(x_0)) &\geq (1 - \eta) \alpha \sum_{\sigma_j x \in \sigma_j \mathcal{L}(\omega)_{\varrho_j} \cap B_1} \sigma_j^{d-1} \|\sigma_j |\nabla_{\omega_0, \sigma_j}(w_j, B_1)|^p (\sigma_j x)\|_1 - 2\eta \\ &= (1 - \eta) G_{\varepsilon_{\varrho_j}, \varrho_j}(x_0, \omega)(w_j) - 2\eta \end{aligned}$$

with the functional  $G_{\varepsilon, \varrho}(x_0, \omega)$  defined in Lemma 5.12. Since we have chosen  $x_0$  and  $\varepsilon_\varrho$  such that Lemma 5.15 holds, we deduce from (64) and the convergence  $w_j \rightarrow u_{0, \xi}$  in  $L^p(B_1, \mathbb{R}^m)$  that

$$\begin{aligned} \lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) &= \lim_j \varrho_j^{-d} F_{\varepsilon_{\varrho_j}}(\omega)(T_{k\varrho_j} u_{\varepsilon_{\varrho_j}}, B_{\varrho_j}(x_0)) \geq \liminf_j (1 - \eta) G_{\varepsilon_{\varrho_j}, \varrho_j}(x_0, \omega)(w_j) - 2\eta \\ &\geq (1 - \eta) \int_{B_1} q(x_0, \xi) dz - 2\eta = (1 - \eta) |B_1| q(x_0, \xi) - 2\eta. \end{aligned}$$

Since  $\eta > 0$  was arbitrary, we conclude that for a.e.  $x_0 \in D$  and all  $\xi \in \mathbb{R}^{m \times d}$

$$\lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) \geq |B_1| q(x_0, \xi). \quad (78)$$

Note that the exceptional set may depend on the subsequence chosen at the beginning.

**Step 4** Proof of  $\lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) \leq |B_1| q(x_0, \xi)$

To prove the reverse inequality in (78) we take a sequence  $u_{\varepsilon_n} \in \mathcal{PC}_{\varepsilon_n}^\omega$  converging to  $u_{x_0, \xi}$  in  $L^p(D, \mathbb{R}^m)$  and such that

$$\lim_{\varepsilon_n \rightarrow 0} E_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho_0}(x_0)) = E(\omega)(u_{x_0, \xi}, B_{\varrho_0}(x_0)).$$

The arguments are very similar to Steps 2 and 3, so we just sketch them. As in (62) one can show that the truncated functions  $T_{k\varrho} u_{\varepsilon_n}$  form a recovery sequence on all balls  $B_\varrho(x_0)$  with  $0 < \varrho < \varrho_0$ . This time we apply Lemma 5.18 to  $T_{k\varrho} u_{\varepsilon_n}$  with the chosen  $x_0$  and  $\lambda = \varrho$ . Note that the assumptions are satisfied since for  $\varrho \leq \varrho_0/(2k)$  we have that

$$\sum_{\varepsilon_n x \in \varepsilon_n \mathcal{L}(\omega) \cap B_{k\varrho}(x_0)} \varepsilon_n^d |\nabla_{e, \varepsilon_n} T_{k\varrho} u_{\varepsilon_n}|^p (\varepsilon_n x) \leq C E_{\varepsilon_n}(\omega)(T_{k\varrho} u_{\varepsilon_n}, B_{2k\varrho}(x_0)) \leq C E_{\varepsilon_n}(\omega)(u_{\varepsilon_n}, B_{\varrho_0}(x_0)).$$

Hence we find a subsequence  $\varepsilon_{n_j}$  and a sequence  $v_j \in \mathcal{PC}_{\varepsilon_{n_j}}^\omega$  (both depending on  $\varrho$ ) such that, setting  $\varepsilon_j = \varepsilon_{n_j}$ , the sequence  $|\nabla_{e, \varepsilon_j} v_j|^p$  is equiintegrable on  $B_{2\varrho}(x_0)$  and

$$\lim_j \varepsilon_j^d \# \{ \varepsilon_j x \in \varepsilon_j \mathcal{L}(\omega) \cap B_{2\varrho}(x_0) : T_{k\varrho} u_{\varepsilon_j} \not\equiv v_j \text{ on } \varepsilon_j B_G(x, 1) \} = 0. \quad (79)$$

By truncation we may further assume that  $\|v_j\|_\infty \leq 9k\varrho$ . Fix  $\eta > 0$  and for  $t > 0$  define the sets  $S_j(t)$  by

$$S_j(t) := \{ \varepsilon_j x \in \varepsilon_j \mathcal{L}(\omega) \cap B_\varrho(x_0) : |\nabla_{e, \varepsilon_j} v_j|^p (\varepsilon_j x) > t \}.$$

We choose  $t_\eta > 0$  (possibly depending on  $\varrho$ ) such that, keeping in mind the inequality  $f(p) \leq C_f \|p\|_1$ , for  $j$  large enough it holds that

$$\varrho^{-d} \sum_{\varepsilon_j x \in S_j(t_\eta)} \varepsilon_j^{d-1} f(\varepsilon_j |\nabla_{\omega, \varepsilon_j}(v_j, B_\varrho(x_0))|^p (\varepsilon_j x)) \leq C \varrho^{-d} \int_{B_{2\varrho}(x_0) \cap \{ |\nabla_{e, \varepsilon_j} v_j|^p > t_\eta \}} |\nabla_{e, \varepsilon_j} v_j|^p(z) dz \leq \eta, \quad (80)$$

which is possible due to the equiintegrability of  $|\nabla_{e, \varepsilon_j} v_j|^p$  on  $B_{2\varrho}(x_0)$ . Denoting by  $W_j$  the set in (79), the same arguments yield  $j_\eta \in \mathbb{N}$  such that for all  $j \geq j_\eta$  we have

$$\varrho^{-d} \sum_{\varepsilon_j x \in W_j \cap B_\varrho(x_0)} \varepsilon_j^{d-1} f(\varepsilon_j |\nabla_{\omega, \varepsilon_j}(v_j, B_\varrho(x_0))|^p (\varepsilon_j x)) \leq \eta. \quad (81)$$

Similar to (77), for  $\varepsilon_j x \in \varepsilon_j \mathcal{L}(\omega) \cap B_\varrho(x_0) \setminus (W_j \cup S_j(t_\eta))$  we know that  $\|\varepsilon_j |\nabla_{w, \varepsilon_j}(v_j, B_\varrho(x_0))|^p\|_1 \leq \varepsilon_j t_\eta$ . Hence assumption (11) and the bounds (80) and (81) imply for large enough  $j$  the estimate

$$\varrho^{-d} E_{\varepsilon_j}(\omega)(T_{k\varrho} u_{\varepsilon_j}, B_\varrho(x_0)) \geq \varrho^{-d} (1 - \eta) F_{\varepsilon_j}(\omega)(v_j, B_\varrho(x_0)) - 2\eta.$$

From (79) and the uniform boundedness of  $v_j$  we infer that  $v_j \rightarrow u_{x_0, \xi}$  in  $L^1(B_\varrho(x_0), \mathbb{R}^m)$ . By a modification on  $\varepsilon_j \mathcal{L}(\omega) \setminus B_\varrho(x_0)$  not affecting the energy, this convergence also holds in  $L^1(D, \mathbb{R}^m)$  and therefore we deduce from  $\Gamma$ -convergence along the subsequence  $\varepsilon_j$  and the previous inequality that

$$\varrho^{-d} E(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) = \lim_j \varrho^{-d} E_{\varepsilon_j}(\omega)(T_{k\varrho} u_{\varepsilon_j}, B_\varrho(x_0)) \geq \varrho^{-d} (1 - \eta) F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) - 2\eta.$$

In view of Remark 5.14 and the arbitrariness of  $\eta$  we conclude that

$$|B_1|q(x_0, \xi) = \lim_{\varrho \rightarrow 0} \varrho^{-d} E(\omega)(u_{x_0, \xi}, B_\varrho(x_0)) \geq \lim_{\varrho \rightarrow 0} \varrho^{-d} F(\omega)(u_{x_0, \xi}, B_\varrho(x_0)).$$

Combined with (78), this estimate yields the claim along the chosen subsequence. In the general case, we obtain that along any subsequence of  $\varepsilon_n$  the  $\Gamma$ -limit of  $E_\varepsilon(\omega)$  is uniquely defined by the integrand  $h(x, \xi)$ , so that the  $\Gamma$ -limit along the sequence  $\varepsilon_n$  exists by the Urysohn-property of  $\Gamma$ -convergence, although the integrand might differ on a negligible set depending on the subsequence.  $\square$

### 5.3. Characterization of the surface density and conclusion

Having identified the bulk term, we now show that the computation of the surface integrand  $\varphi(x, a, v)$  can be performed with the discrete functional  $F_\varepsilon(\omega)$  restricted to functions taking only the two values  $a$  and  $0$ . Then we prove that the surface density  $\varphi$  agrees with the function  $s$  defined in Remark 3.4 via the energy  $I_\varepsilon(\omega)$ .

We study the asymptotic minimization problems given by Proposition 5.2 and their connection to boundary value problems for the discrete functionals  $F_\varepsilon(\omega)$ . More precisely, as a first step we compare the two quantities

$$\begin{aligned} m_\varepsilon^\delta(\omega)(\bar{u}, A) &= \inf\{F_\varepsilon(\omega)(v, A) : v \in \mathcal{PC}_{\varepsilon, \delta}^\omega(\bar{u}, A)\}, \\ m(\omega)(\bar{u}, A) &= \inf\{F(\omega)(v, A) : v \in SBVP(A, \mathbb{R}^m), v = \bar{u} \text{ in a neighborhood of } \partial A\}, \end{aligned}$$

where the limit functional  $F(\omega)$  is given (up to subsequences) by Proposition 5.2 and the set  $\mathcal{PC}_{\varepsilon, \delta}^\omega(\bar{u}, A)$ , which takes into account discrete boundary conditions, is defined in (20). We restrict the class of boundary conditions to pointwise well-defined functions  $\bar{u} \in SBVP(D, \mathbb{R}^m) \cap L^\infty(D, \mathbb{R}^m)$  such that, setting  $\bar{u}_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  as  $\bar{u}_\varepsilon(\varepsilon x) = \bar{u}(\varepsilon x)$ , it holds that

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\bar{u}_\varepsilon, B) &\leq C \int_B |\nabla \bar{u}|^p dx + C \mathcal{H}^{d-1}(S_{\bar{u}} \cap \bar{B}), \\ \bar{u}_\varepsilon &\rightarrow \bar{u} \text{ in } L^1(D, \mathbb{R}^m), \quad \mathcal{H}^{d-1}(S_{\bar{u}} \cap \partial A) = 0 \end{aligned} \tag{82}$$

for some  $C > 0$  uniformly for  $B \in \mathcal{A}^R(D)$ . In particular, as seen in the proof of Lemma 5.7, we allow for piecewise smooth functions with polyhedral jump set that has no mass on  $\partial A$ . We have the following convergence result.

**Lemma 5.20** (Approximation of minimum values). *Let  $\varepsilon_n$  and  $F(\omega)$  be as in Proposition 5.2. Then, for any  $A \in \mathcal{A}^R(D)$  and  $\bar{u}$  as in (82), it holds that*

$$\lim_{\delta \rightarrow 0} \liminf_n m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A) = \lim_{\delta \rightarrow 0} \limsup_n m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A) = m(\omega)(\bar{u}, A).$$

**Proof.** First note that by monotonicity the limits with respect to  $\delta$  exist. Moreover, from the first assumption in (82) it follows that  $m_\varepsilon^\delta(\omega)(\bar{u}, A)$  is equibounded. For  $n \in \mathbb{N}$  let  $u_n \in \mathcal{PC}_{\varepsilon_n, \delta}^\omega(\bar{u}, A)$  be such that  $m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A) = F_{\varepsilon_n}(\omega)(u_n, A)$ . Since  $\bar{u} \in L^\infty$  we can apply Lemma 5.4 and assume without loss of generality that  $|u_n(\varepsilon_n x)| \leq 3\|\bar{u}\|_\infty$  for all  $x \in \mathcal{L}(\omega)$ . By Lemma 5.6 we know that, up to a subsequence (not relabeled),  $u_n \rightarrow u$  in  $L^1(A, \mathbb{R}^m)$  for some  $u \in L^1(A, \mathbb{R}^m) \cap GSBVP(A, \mathbb{R}^m)$ . Using Remark 2.3 and again (82) we infer that  $u = \bar{u}$  on  $\partial_\delta A$ . Note that  $u \in L^\infty(A, \mathbb{R}^m)$ , which implies  $u \in SBVP(A, \mathbb{R}^m)$ . Up to extension we can assume that  $u$  is admissible in the infimum problem defining  $m(\omega)(\bar{u}, A)$  and Proposition 5.2 yields

$$m(\omega)(\bar{u}, A) \leq F(\omega)(u, A) \leq \liminf_n F_{\varepsilon_n}(\omega)(u_n, A) \leq \liminf_n m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A).$$

As  $\delta$  was arbitrary, we conclude that  $m(\omega)(\bar{u}, A) \leq \lim_{\delta \rightarrow 0} \liminf_n m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A)$ .

In order to prove the remaining inequality, for given  $\theta > 0$  we let  $u \in SBV^p(A, \mathbb{R}^m)$  be such that  $u = \bar{u}$  in a neighborhood of  $\partial A$  and  $F(\omega)(u, A) \leq m(\omega)(\bar{u}, A) + \theta$ . Take  $u_n \in \mathcal{PC}_{\varepsilon_n}^\omega$  converging to  $u$  in  $L^1(D, \mathbb{R}^m)$  and satisfying

$$\lim_n F_{\varepsilon_n}(\omega)(u_n, A) = F(\omega)(u, A). \quad (83)$$

We will modify  $u_n$  such that it fulfills the discrete boundary conditions. The argument follows the proof of Proposition 5.8. Since  $u = \bar{u}$  in a neighborhood of  $\partial A$ , there exist sets  $A' \subset \subset A'' \subset \subset A$  such that  $A', A'' \in \mathcal{A}^R(D)$  and

$$u = \bar{u} \quad \text{on } A \setminus A'. \quad (84)$$

Fix  $N \in \mathbb{N}$ . For  $h \leq \text{dist}(A', \partial A'')$  and  $i \in \{1, \dots, N\}$  we define the sets

$$A_i = \left\{ x \in A : \text{dist}(x, A') < i \frac{h}{2N} \right\}.$$

Let  $\Theta_i$  be a cut-off function between the sets  $A_i$  and  $A_{i+1}$  with  $\|\nabla \Theta_i\|_\infty \leq \frac{4N}{h}$  and define  $u_n^i \in \mathcal{PC}_{\varepsilon_n}^\omega$  by

$$u_n^i(\varepsilon_n x) = \Theta_i(\varepsilon_n x) u_n(\varepsilon_n x) + (1 - \Theta_i(\varepsilon_n x)) \bar{u}(\varepsilon_n x).$$

Up to extending  $u$  on  $D \setminus A$  via  $u|_{D \setminus A} = \bar{u}$ , by (82) and (84) we can assume that for each  $i \in \{1, \dots, N\}$  it holds that  $u_n^i \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$ . Setting  $S_n^i := \{x \in A : \text{dist}(x, A_{i+1} \setminus A_{i-1}) < 3M\varepsilon_n\}$ , we have

$$F_{\varepsilon_n}(\omega)(u_n^i, A) \leq F_{\varepsilon_n}(\omega)(u_n, A) + F_{\varepsilon_n}(\omega)(\bar{u}_{\varepsilon_n}, A \setminus \overline{A'}) + F_{\varepsilon_n}(\omega)(u_n^i, S_n^i). \quad (85)$$

Using the same arguments as in the proof of Proposition 5.8 we infer that

$$F_\varepsilon(\omega)(u_n^i, S_n^i) \leq C \left( F_{\varepsilon_n}(\omega)(u_n, S_n^i) + F_{\varepsilon_n}(\omega)(\bar{u}_{\varepsilon_n}, S_n^i) \right) + CN^p h^{-p} \sum_{\varepsilon_n x \in \varepsilon_n \mathcal{L}(\omega) \cap S_n^i} \varepsilon_n^d |u_n(\varepsilon_n x) - \bar{u}_{\varepsilon_n}(\varepsilon_n x)|^p.$$

By construction  $S_n^i \cap S_n^j = \emptyset$  for  $|i - j| \geq 3$  and  $S_n^i \subset \subset A \setminus A'$  for  $i \geq 2$ . Averaging the previous inequality and using (82) and (83) yields

$$\frac{1}{N-1} \sum_{i=2}^N F_\varepsilon(\omega)(u_n^i, S_n^i) \leq \frac{C}{N} + CN^{p-1} h^{-p} \|u_n - \bar{u}_{\varepsilon_n}\|_{L^p(A \setminus A')}^p.$$

By equiboundedness, properties (82) and (84) imply that  $u_n - \bar{u}_{\varepsilon_n} \rightarrow 0$  in  $L^p(A \setminus A', \mathbb{R}^m)$ . For every  $n$  we choose  $i_n \in \{2, \dots, N\}$  such that

$$F_{\varepsilon_n}(\omega)(u_n^{i_n}, S_n^{i_n}) \leq \frac{C}{N} + CN^{p-1} h^{-p} \|u_n - \bar{u}_{\varepsilon_n}\|_{L^p(A \setminus A')}^p. \quad (86)$$

Note that  $u_n^{i_n}$  still converges to  $u$  in  $L^1(D, \mathbb{R}^m)$ . Moreover,  $u_n^{i_n}(\varepsilon_n x) = \bar{u}(\varepsilon_n x)$  for all  $\varepsilon_n x \in \varepsilon_n \mathcal{L}(\omega) \cap A \setminus A''$ . Hence  $u_n^{i_n} \in \mathcal{PC}_{\varepsilon_n, \delta}^\omega(\bar{u}, A)$  for all  $\delta > 0$  small enough. From (83), (85) and (86) we obtain

$$\begin{aligned} \limsup_n m_{\varepsilon_n}^\delta(\omega)(\bar{u}, A) &\leq \limsup_n F_{\varepsilon_n}(\omega)(u_n^{i_n}, A) \leq F(\omega)(u, A) + \limsup_n F_{\varepsilon_n}(\omega)(\bar{u}_{\varepsilon_n}, A \setminus \overline{A'}) + \frac{C}{N} \\ &\leq m(\omega)(\bar{u}, A) + \theta + C \int_{A \setminus \overline{A'}} |\nabla \bar{u}|^p dx + C \mathcal{H}^{d-1}(S_{\bar{u}} \cap A \setminus A') + \frac{C}{N}, \end{aligned}$$

where we used (82) with  $B = A \setminus \overline{A'}$ . As  $\theta > 0$  was arbitrary, the claim follows letting first  $\delta \rightarrow 0$ , then  $N \rightarrow +\infty$  and finally  $A' \uparrow A$ .  $\square$

In view of Proposition 5.2 and Lemma 5.20 we can further characterize the surface densities of possible  $\Gamma$ -limits of the family  $F_\varepsilon(\omega)$  by analyzing the quantities  $m_\varepsilon^\delta(\omega)(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho))$ . To this end, we define the class of interfaces

$$\mathcal{S}_{\varepsilon,\delta}^\omega(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho)) = \{u \in \mathcal{PC}_{\varepsilon,\delta}^\omega(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho)) : u(\varepsilon x) \in \{a, 0\} \text{ for all } x \in \mathcal{L}(\omega)\}.$$

We have the following important result, which also implies that the minimization defining the surface energy density can be performed on characteristic functions of sets of finite perimeter instead of general Caccioppoli partitions.

**Proposition 5.21** (*Separation of surface effects*). *Let  $\varepsilon_n \rightarrow 0$ . Then, for all  $x_0 \in D$ , all  $a \in \mathbb{R}^m$  and all  $v \in S^{d-1}$  it holds that*

$$\begin{aligned} \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \lim_{\delta \rightarrow 0} \limsup_n \left( \inf \left\{ F_{\varepsilon_n}(\omega)(u, Q_v(x_0, \varrho)) : u \in \mathcal{S}_{\varepsilon_n,\delta}^\omega(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho)) \right\} \right) \\ = \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \lim_{\delta \rightarrow 0} \limsup_n m_{\varepsilon_n}^\delta(\omega)(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho)). \end{aligned}$$

**Proof.** Note that it suffices to bound the first term by the second one. To reduce notation, we set  $Q_\varrho := Q_v(x_0, \varrho)$  and write  $\varepsilon$  instead of  $\varepsilon_n$ . If  $a = 0$  then both sides are zero. Thus assume that  $a \neq 0$ . Fix  $u_\varepsilon \in \mathcal{PC}_{\varepsilon,\delta}^\omega(u_{x_0,v}^{a,0}, Q_\varrho)$  such that

$$F_\varepsilon(\omega)(u_\varepsilon, Q_\varrho) \leq F_\varepsilon(\omega)(u_{x_0,v}^{a,0}, Q_\varrho) \leq C\varrho^{d-1}, \quad (87)$$

where the last inequality is a consequence of Remark 2.3 and the boundedness of the discrete density  $f$  (provided  $\varepsilon$  is small enough). In what follows we construct a sequence  $v_\varepsilon \in \mathcal{S}_{\varepsilon,\delta}^\omega(u_{x_0,v}^{a,0}, Q_\varrho)$  that has almost the same energy. Given  $\theta > 0$ , due to the monotonicity (12) one can choose  $L_\theta \geq 1$  such that

$$|\beta(k, k) - f(p)| < \theta f(p) \quad \forall p \in \mathcal{P}_+(M) \text{ with } p^{-1}(\mathbb{N}) \subset [L_\theta, +\infty) \text{ and } \sum_{v \in p^{-1}(\mathbb{N})} p(v) = k \quad (88)$$

for all  $1 \leq k \leq M$ , where  $\beta(k, k)$  is given by (15). We fix  $L_\theta$  from now on and consider the set of edges

$$\mathcal{J}_{u_\varepsilon} = \{(x, y) \in \mathcal{E}(\omega) : \varepsilon x, \varepsilon y \in Q_\varrho \text{ and } \varepsilon^{1-p} |u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)|^p \geq L_\theta\}.$$

Denote by  $u_\varepsilon^i$  the  $i^{\text{th}}$  component of  $u_\varepsilon$ . If  $a_i = 0$  we set  $v_\varepsilon^i(\varepsilon x) = 0$  for all  $x \in \mathcal{L}(\omega)$ . Otherwise, we assume that  $a_i > 0$ . The remaining case requires only minor modifications. For  $t \in \mathbb{R}$  we define

$$S_\varepsilon^i(t) := \{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap Q_\varrho : u_\varepsilon^i(\varepsilon x) > t\}.$$

To reduce notation, we also introduce the set

$$\mathcal{R}_\varepsilon^i(t) = \{(x, y) \in \mathcal{E}(\omega) : \varepsilon x \in Q_\varrho \cap S_\varepsilon^i(t), \varepsilon y \in Q_\varrho \setminus S_\varepsilon^i(t) \text{ or vice versa}\}.$$

Observe that for  $(x, y) \in \mathcal{E}(\omega)$  with  $\varepsilon x, \varepsilon y \in Q_\varrho$  we have  $(x, y) \in \mathcal{R}_\varepsilon^i(t)$  if and only if  $t \in [u_\varepsilon^i(\varepsilon x), u_\varepsilon^i(\varepsilon y))$  or  $t \in [u_\varepsilon^i(\varepsilon y), u_\varepsilon^i(\varepsilon x))$ . Hence for such  $x, y$  the following coarea-type estimate holds true:

$$\int_0^{a_i} \mathbb{1}_{\{(x,y) \in \mathcal{R}_\varepsilon^i(t)\}} dt \leq |u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)|.$$

Summing this estimate, we infer from (10) and Hölder's inequality that

$$\begin{aligned} \int_0^{a_i} \varepsilon^{d-1} \#(\mathcal{R}_\varepsilon^i(t) \setminus \mathcal{J}_{u_\varepsilon}) dt &\leq \sum_{\substack{(x,y) \in \mathcal{E}(\omega) \setminus \mathcal{J}_{u_\varepsilon} \\ \varepsilon x, \varepsilon y \in Q_\varrho}} \varepsilon^{d-1} |u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)| \\ &\leq C \varepsilon^{\frac{dp-d}{p}} (\#(\varepsilon \mathcal{L}(\omega) \cap Q_\varrho))^{\frac{p-1}{p}} \left( \sum_{\substack{(x,y) \in \mathcal{E}(\omega) \setminus \mathcal{J}_{u_\varepsilon} \\ \varepsilon x, \varepsilon y \in Q_\varrho}} \varepsilon^d \left| \frac{u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)}{\varepsilon} \right|^p \right)^{\frac{1}{p}}. \end{aligned}$$

In order to estimate the last sum, recall that  $L_\theta \geq 1$ . Thus the definition of the set  $\mathcal{J}_{u_\varepsilon}$ , (10) and assumption (14) imply for  $\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap Q_\varrho$  the uniform bound

$$\begin{aligned} \sum_{\substack{\varepsilon y \in \varepsilon\mathcal{L}(\omega) \cap Q_\varrho \\ (x,y) \in \mathcal{E}(\omega) \setminus \mathcal{J}_{u_\varepsilon}}} \varepsilon \left| \frac{u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)}{\varepsilon} \right|^p &\leq CL_\theta \min \left\{ \varepsilon \sum_{\substack{\varepsilon y \in \varepsilon\mathcal{L}(\omega) \cap Q_\varrho \\ (x,y) \in \mathcal{E}(\omega) \setminus \mathcal{J}_{u_\varepsilon}}} \left| \frac{u_\varepsilon(\varepsilon x) - u_\varepsilon(\varepsilon y)}{\varepsilon} \right|^p, 1 \right\} \\ &\leq CL_\theta \min \{ \|\varepsilon |\nabla_{\omega,\varepsilon}(u_\varepsilon, Q_\varrho)|^p(\varepsilon x)\|_1, 1 \} \leq CL_\theta f(\varepsilon |\nabla_{\omega,\varepsilon}(u_\varepsilon, Q_\varrho)|^p(\varepsilon x)). \end{aligned}$$

Moreover, for  $\varepsilon = \varepsilon(\varrho)$  small enough the cardinality term can be bounded by  $\#(\varepsilon\mathcal{L}(\omega) \cap Q_\varrho) \leq C(\varrho\varepsilon^{-1})^d$ , so that

$$\int_0^{a_i} \varepsilon^{d-1} \#(\mathcal{R}_\varepsilon^i(t) \setminus \mathcal{J}_{u_\varepsilon}) dt \leq C\varrho^{\frac{dp-d}{p}} (L_\theta F_\varepsilon(\omega)(u_\varepsilon, Q_\varrho))^{\frac{1}{p}} \leq CL_\theta \varrho^{\frac{dp-1}{p}},$$

where we applied (87) in the second inequality. Hence there exists  $t_\varepsilon^i \in (0, a_i)$  such that

$$\varepsilon^{d-1} \#(\mathcal{R}_\varepsilon^i(t_\varepsilon^i) \setminus \mathcal{J}_{u_\varepsilon}) \leq C|a_i|^{-1} L_\theta \varrho^{\frac{dp-1}{p}}. \quad (89)$$

Define  $v_\varepsilon^i$  by its values on  $\varepsilon\mathcal{L}(\omega)$  setting

$$v_\varepsilon^i(\varepsilon x) = \begin{cases} 0 & \text{if } u_\varepsilon(\varepsilon x) \leq t_\varepsilon^i, \\ a_i & \text{if } u_\varepsilon(\varepsilon x) > t_\varepsilon^i. \end{cases}$$

As  $t_\varepsilon^i \in (0, a_i)$ , the boundary conditions imposed on  $u_\varepsilon$  imply that the function  $v_\varepsilon$  satisfies  $v_\varepsilon(\varepsilon x) = u_{x_0,v}^{a,0}(\varepsilon x)$  for all  $\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap \partial_\delta Q_\varrho$ , so that  $v_\varepsilon \in \mathcal{S}_{\varepsilon,\delta}^\omega(u_{x_0,v}^{a,0}, Q_\varrho)$ . In order to estimate the energy difference, let  $\varepsilon x \in \varepsilon\mathcal{L}(\omega) \cap Q_\varrho$  be such that  $\|\varepsilon |\nabla_{\omega,\varepsilon}(v_\varepsilon, Q_\varrho)|^p(\varepsilon x)\|_1 \neq 0$ . We distinguish two exhaustive cases: either  $(x, y) \in \mathcal{J}_{u_\varepsilon}$  for all  $\varepsilon y \in \varepsilon\mathcal{L}(\omega) \cap Q_\varrho$  with  $(x, y) \in \mathcal{E}(\omega)$ , so that (12) and (88) yield

$$f(\varepsilon |\nabla_{\omega,\varepsilon}(v_\varepsilon, Q_\varrho)|^p(\varepsilon x)) \leq (1 + \theta) f(\varepsilon |\nabla_{\omega,\varepsilon}(u_\varepsilon, Q_\varrho)|^p(\varepsilon x)), \quad (90)$$

or there exists  $y \in \mathcal{L}(\omega)$  with  $(x, y) \in \mathcal{R}_\varepsilon^i(t_\varepsilon^i) \setminus \mathcal{J}_{u_\varepsilon}$  for some  $i$ . In this case we can use the estimate (89) to bound the number of such  $x$ . Since  $f$  is bounded by assumption, we deduce from (90) that

$$\varrho^{1-d} F_\varepsilon(\omega)(v_\varepsilon, Q_\varrho) \leq (1 + \theta) \varrho^{1-d} F_\varepsilon(\omega)(u_\varepsilon, Q_\varrho) + C \sum_{i:a_i \neq 0} |a_i|^{-1} L_\theta \varrho^{\frac{p-1}{p}}.$$

Taking the appropriate infimum on each side, then letting  $\varepsilon \rightarrow 0$  before  $\delta \rightarrow 0$  and  $\varrho \rightarrow 0$ , we conclude the proof as  $\theta > 0$  was arbitrary.  $\square$

Now we can relate the surface density  $\varphi$  to the  $\Gamma$ -limit of the functionals  $I_\varepsilon(\omega)$  defined in (18).

**Proposition 5.22.** *Let  $\varepsilon_n$  and  $F(\omega)$  be as in Proposition 5.2. Then for every  $x_0 \in D$ ,  $a \in \mathbb{R}^m \setminus \{0\}$  and  $v \in S^1$  it holds that*

$$\varphi(x_0, a, v) = s(x_0, v),$$

where  $s$  is the surface tension of the  $\Gamma$ -limit of  $I_{\varepsilon_n}(\omega)(\cdot, D)$ , which in particular exists.

**Proof.** Choosing any subsequence of  $\varepsilon_n$  (not relabeled), the  $\Gamma$ -limit of  $F_{\varepsilon_n}(\omega)$  remains the same. Hence, combining Propositions 5.2 and 5.21 with Lemma 5.20 yields the formula

$$\varphi(x_0, a, v) = \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \lim_{\delta \rightarrow 0} \limsup_n \left( \inf \left\{ F_{\varepsilon_n}(\omega)(u, Q_v(x_0, \varrho)) : u \in \mathcal{S}_{\varepsilon_n,\delta}^\omega(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho)) \right\} \right). \quad (91)$$

For any  $u \in \mathcal{S}_{\varepsilon_n,\delta}^\omega(u_{x_0,v}^{a,0}, Q_v(x_0, \varrho))$  we define the function  $v \in \mathcal{S}_{\varepsilon_n,\delta}^\omega(u_{x_0,v}^{-e_1,e_1}, Q_v(x_0, \varrho))$  by



$$v(\varepsilon x) = \begin{cases} -e_1 & \text{if } u(\varepsilon x) = a, \\ e_1 & \text{otherwise.} \end{cases}$$

Let  $I_\varepsilon(\omega)$  be the Ising-type energy defined in (18). Then by the monotonicity (12) it holds that

$$\begin{aligned} I_{\varepsilon_n}(\omega)(v, Q_v(x_0, \varrho)) &\geq F_{\varepsilon_n}(\omega)(u, Q_v(x_0, \varrho)) \\ &\geq \min_{1 \leq l \leq k \leq M} \left\{ \frac{f(p)}{\beta(l, k)} : p = l \mathbb{1}_{\{|a|^{p_{\varepsilon_n}^1 - p}\}} + (k - l) \mathbb{1}_{\{0\}} \right\} I_{\varepsilon_n}(\omega)(v, Q_v(x_0, \varrho)). \end{aligned}$$

Since we have chosen a subsequence at the beginning of the proof, we may assume that  $I_{\varepsilon_n}(\omega)$   $\Gamma$ -converges to some surface integral functional  $I(\omega)$  as in Theorem 3.2 with density  $s(x, v)$ . Since  $|a| \neq 0$ , the definition of  $\beta(l, k)$  in (15) and Remark 3.4 imply

$$\varphi(x_0, a, v) = \limsup_{\varrho \rightarrow 0} \varrho^{1-d} \lim_{\delta \rightarrow 0} \limsup_n \inf \{ I_{\varepsilon_n}(\omega)(v, Q_v(x_0, \varrho)) : v \in \mathcal{S}_{\varepsilon_n, \delta}^\omega(u_{x_0, v}^{-e_1, e_1}, Q_v(x_0, \varrho)) \} = s(x_0, v).$$

Since the subsequence was arbitrary, the Urysohn-property of  $\Gamma$ -convergence yields that  $I_{\varepsilon_n}(\omega)$  indeed  $\Gamma$ -converges and the surface densities of the limits  $F(\omega)$  and  $I(\omega)$  agree.  $\square$

Eventually, we can prove our first main result.

**Proof of Theorem 3.3.** If  $F_{\varepsilon_n}(\omega)$   $\Gamma$ -converges, then due to Propositions 5.19 and 5.22 both  $E_{\varepsilon_n}(\omega)$  and  $I_{\varepsilon_n}(\omega)$   $\Gamma$ -converge, too. Also the reverse statement follows from the same propositions, since (up to subsequences) any  $\Gamma$ -limit of  $F_{\varepsilon_n}$  is characterized by the  $\Gamma$ -limits of  $E_{\varepsilon_n}(\omega)$  and  $I_{\varepsilon_n}(\omega)$ , which do not depend on further subsequences. Taking also into account Proposition 5.2 and Lemma 5.6 the  $\Gamma$ -limit is given by

$$F(\omega)(u) = \begin{cases} \int_D q(x, \nabla u) \, dx + \int_{S_u} s(x, v_u) \, d\mathcal{H}^{d-1} & \text{if } u \in SBV^p(D, \mathbb{R}^m), \\ +\infty & \text{if } u \in L^1(D, \mathbb{R}^m) \setminus GSBV^p(D, \mathbb{R}^m). \end{cases}$$

Hence it remains to characterize the functional  $F(\omega)(u)$  for  $u \in L^1(D, \mathbb{R}^m) \cap GSBV^p(D, \mathbb{R}^m)$  such that  $u \notin SBV^p(D, \mathbb{R}^m)$ . This will be achieved via truncation. Given  $k > 0$  we have that  $T_k u \in SBV^p(D, \mathbb{R}^m)$ . Lemma 5.4 implies that  $F(\omega)(u) = \lim_{k \rightarrow +\infty} F(\omega)(T_k u)$ . In order to pass to the limit in the integral formula, we use Lemma 2.1 and the symmetry  $s(x, v) = s(x, -v)$ , which yield

$$F(\omega)(T_k, A) = \int_{D \cap \{u \leq k\}} q(x, \nabla u) \, dx + \int_{D \cap \{u > k\}} q(x, \nabla T_k u) \, dx + \int_{S_{T_k u}} s(x, v_u) \, d\mathcal{H}^{d-1}.$$

Since  $S_u = \bigcup_k S_{T_k u} \cup N$  with  $\mathcal{H}^{d-1}(N) = 0$  and the second term vanishes due to dominated convergence, we can pass to the limit and conclude the proof.  $\square$

## 6. Stochastic homogenization: proof of Theorems 3.5 and 3.8

In this section we derive the results in the random setting.

**Proof of Theorem 3.5.** By [5, Theorem 2] and [6, Theorem 5.5],<sup>3</sup> the  $\Gamma$ -limits of the two functionals  $E_\varepsilon(\omega)$  and  $I_\varepsilon(\omega)$  defined in (17) and (18) exist almost surely and have deterministic, spatially homogeneous densities. Hence the claim on the existence and form of the  $\Gamma$ -limit follows from Theorem 3.3. It remains to establish the properties of the integrands. Convexity and  $p$ -homogeneity of  $h$  follow from the fact that the  $\Gamma$ -limit of the sequence of convex and  $p$ -homogeneous functionals  $E_\varepsilon(\omega)$  is again convex and  $p$ -homogeneous [35, Theorem 11.1 and Proposition 11.6], whereas convexity of the one-homogeneous extension of  $\varphi$  follows from standard  $L^1$ -lower-semicontinuity results for functionals defined on sets of finite perimeter (see for instance [8, Theorem 3.1]).  $\square$

<sup>3</sup> As noted before, in [6] the proofs were given only for pairwise interactions. Nevertheless the same arguments apply in our setting (see also [6, Theorem 6.7]). Note that stationarity of the edges is important to apply the subadditive ergodic theorem of [2].

Finally we prove Theorem 3.8. For the convergence of minimizers we exploit the notion of biting convergence, which we recall here for reader's convenience.

**Definition 6.1** (*Biting convergence*). Let  $u_n \in L^1(D)$  be such that  $\sup_n \|u_n\|_{L^1(D)} < +\infty$ . We say that  $u_n$  converges weakly to  $u \in L^1(D)$  in the biting sense and write  $u_n \xrightarrow{b} u$ , if there exists a decreasing sequence  $S_j \subset D$  of measurable sets such that  $|S_j| \rightarrow 0$  and  $u_n \rightarrow u$  in  $L^1(D \setminus S_j)$  for all  $j \in \mathbb{N}$ .

**Remark 6.2.** Note that if  $u_n \xrightarrow{b} u$  and  $u_n \rightarrow v$  a.e., then  $u = v$ . This is a consequence of the uniqueness of the biting limit and equiintegrability of  $L^1$ -weakly convergent sequences.

**Proof of Theorem 3.8.** We first construct a candidate for the constant  $\gamma$ . Define the sequence of non-negative equi-bounded functions  $\gamma_\varepsilon(\omega) \in L^\infty(D)$  by

$$\gamma_\varepsilon(\omega)(z) = \sum_{x \in \mathcal{L}(\omega)} \frac{1}{|\mathcal{C}(x)|} \mathbb{1}_{\varepsilon \mathcal{C}(x)}(z). \quad (92)$$

We apply the ergodic theorem in order to establish weak\*-convergence of  $\gamma_\varepsilon(\omega)$ . To this end, we introduce the family of half-open boxes with integer vertices  $\mathcal{I} := \{[a, b) : a, b \in \mathbb{Z}^d, a_i < b_i \text{ for all } i\}$  and define the rescaled integral averages  $\tilde{\gamma} : \mathcal{I} \rightarrow L^1(\Omega)$  by

$$\tilde{\gamma}(I, \omega) = \int_I \sum_{x \in \mathcal{L}(\omega)} \frac{1}{|\mathcal{C}(x)|} \mathbb{1}_{\mathcal{C}(x)}(z) dz = \sum_{x \in \mathcal{L}(\omega)} \frac{|\mathcal{C}(x) \cap I|}{|\mathcal{C}(x)|}.$$

The following three properties can be verified:

- (i)  $0 \leq \tilde{\gamma}(I, \omega) \leq C|I|$  for all  $I \in \mathcal{I}$ ,
- (ii) If  $I = \bigcup_i I_i$  with finitely many, pairwise disjoint  $I_i \in \mathcal{I}$ , then  $\tilde{\gamma}(I, \omega) = \sum_i \tilde{\gamma}(I_i, \omega)$ ,
- (iii)  $\tilde{\gamma}(I, \tau_z \omega) = \tilde{\gamma}(I - z, \omega)$  for all  $z \in \mathbb{Z}^d$ .

Moreover, arguing as in [32, Lemma A.1], one can show that  $\omega \mapsto \tilde{\gamma}(I, \omega)$  is  $\mathcal{F}$ -measurable. Hence we can apply the multi-parameter additive ergodic theorem (see [47, Chapter 6, Theorem 2.8]) and conclude that  $\mathbb{P}$ -a.s. and for all  $I \in \mathcal{I}$

$$\gamma := \mathbb{E}[\tilde{\gamma}] = \lim_{n \rightarrow +\infty} \frac{\tilde{\gamma}(nI, \omega)}{|nI|},$$

where  $\mathbb{E}$  denotes the expectation. It is straightforward to extend this convergence to all sequences  $t_n \rightarrow +\infty$  and then to all cubes in  $\mathbb{R}^d$  by a continuity argument. Now we identify the weak\*-limit of  $\gamma_\varepsilon(\omega)$ . By a density argument it is enough to compute averages on cubes  $Q \subset D$ . A change of variables yields

$$\int_Q \gamma_\varepsilon(\omega)(z) dz = \varepsilon^d \tilde{\gamma}(Q/\varepsilon, \omega) \rightarrow \gamma|Q|,$$

whence  $\gamma_\varepsilon(\omega) \xrightarrow{*} \gamma$  in  $L^\infty(D)$  almost surely.

Next we prove the lower bound for the  $\Gamma$ -convergence. Passing to a subsequence, for the  $\liminf$ -inequality it suffices to consider  $u \in L^1(D, \mathbb{R}^m)$  and a sequence  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  such that  $u_\varepsilon \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$  and

$$\liminf_{\varepsilon \rightarrow 0} F_{\varepsilon, g}(\omega)(u_\varepsilon) = \lim_{\varepsilon \rightarrow 0} F_{\varepsilon, g}(\omega)(u_\varepsilon) \leq C < +\infty. \quad (93)$$

Without affecting the convergence properties or the functional we redefine  $g_\varepsilon(\omega)(\varepsilon x) = u_\varepsilon(\varepsilon x) = 0$  for all  $\varepsilon x \in \varepsilon \mathcal{L}(\omega) \setminus D$ . Then by Remark 2.3 we have

$$\|u_\varepsilon - g_\varepsilon(\omega)\|_{L^q(D)}^q \leq C \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D} \varepsilon^d |u_\varepsilon(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q \leq C,$$

which in combination with (21) implies that  $u_\varepsilon$  is bounded in  $L^q(D, \mathbb{R}^m)$ . Thus we obtain  $u \in L^q(D, \mathbb{R}^m)$ , while Theorem 3.5 and (93) yield  $u \in GSBVP(D, \mathbb{R}^m)$ . Moreover, for any  $1 \leq r < q$  we deduce the following convergence properties:

$$u_\varepsilon \rightarrow u \quad \text{in } L^r(D, \mathbb{R}^m), \quad u_\varepsilon \rightharpoonup u \quad \text{in } L^q(D, \mathbb{R}^m). \quad (94)$$

Observe that by the definition of the function  $\gamma_\varepsilon(\omega)$  it holds that

$$\sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D} \varepsilon^d |u_\varepsilon(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q \geq \int_D |\gamma_\varepsilon(\omega)(z)| |u_\varepsilon(z) - g_\varepsilon(\omega)(z)|^q dz.$$

Due to (21) and (94), the sequence  $u_\varepsilon - g_\varepsilon(\omega)$  converges to  $u - g$  in  $L^r(D, \mathbb{R}^m)$  for any  $1 \leq r < q$ . The lower semicontinuity result for pairs of weak-strong convergent sequences in [40, Theorem 7.5] and the  $\Gamma$ -convergence of Theorem 3.5 imply

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} F_{\varepsilon, g}(\omega)(u_\varepsilon) &\geq \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon) + \liminf_{\varepsilon \rightarrow 0} \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D} \varepsilon^d |u_\varepsilon(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q \\ &\geq F(u) + \gamma \int_D |u - g|^q dz, \end{aligned} \quad (95)$$

where we used that  $\gamma > 0$  to avoid the modulus. This finishes the proof of the lower bound.

For the upper bound, it suffices to consider  $u \in L^q(D, \mathbb{R}^m) \cap GSBVP(D, \mathbb{R}^m)$ . Note that for such  $u$  we can equivalently compute the  $\Gamma$ -limit of  $F_\varepsilon(\omega)$  with respect to convergence in  $L^q(D, \mathbb{R}^m)$ . Indeed, by Lemma 5.4 this is true for all truncated functions  $T_k u$  with  $k > 0$  and by lower semicontinuity with respect to  $L^q$ -convergence and again Lemma 5.4 we obtain

$$\Gamma(L^q(D))\text{-}\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u) \leq \liminf_{k \rightarrow +\infty} \left( \Gamma(L^q(D))\text{-}\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(T_k u) \right) \leq \liminf_{k \rightarrow +\infty} F(T_k u) = F(u).$$

Hence we find a sequence  $u_\varepsilon \in \mathcal{PC}_\varepsilon^\omega$  such that  $u_\varepsilon \rightarrow u$  in  $L^q(D, \mathbb{R}^m)$  and

$$\lim_{\varepsilon \rightarrow 0} F_\varepsilon(\omega)(u_\varepsilon) = F(u). \quad (96)$$

Since  $D$  has Lipschitz boundary, it satisfies an interior cone condition. Thus we find  $c_D > 0$  such that

$$|\varepsilon \mathcal{C}(x) \cap D| \geq c_D \varepsilon^d \quad \text{for all } \varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D.$$

Setting  $D_\varepsilon = \{z \in D : \text{dist}(z, \partial D) \leq 2R\varepsilon\}$ , we deduce from the above estimate that

$$\begin{aligned} \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D} \varepsilon^d |u_\varepsilon(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^q &\leq \int_D \gamma_\varepsilon(\omega)(z) |u_\varepsilon(z) - g_\varepsilon(\omega)(z)|^q dz \\ &\quad + C \int_{D_\varepsilon} |u_\varepsilon(z) - g_\varepsilon(\omega)(z)|^q dz. \end{aligned}$$

The last term vanishes when  $\varepsilon \rightarrow 0$  since the sequence  $|u_\varepsilon - g_\varepsilon|^q$  is equiintegrable on  $D$ . Moreover, by its product structure the sequence  $\gamma_\varepsilon(\omega) |u_\varepsilon - g_\varepsilon(\omega)|^q$  converges weakly in  $L^1(D, \mathbb{R}^m)$  to  $\gamma |u - g|^q$ . Therefore the last inequality implies

$$\limsup_{\varepsilon \rightarrow 0} \sum_{\varepsilon x \in \varepsilon \mathcal{L}(\omega) \cap D} \varepsilon^d |u_\varepsilon(\varepsilon x) - g_\varepsilon(\omega)(\varepsilon x)|^p \leq \gamma \int_D |u(z) - g(z)|^p dz.$$

Combined with (96) we obtain the upper bound.

Now we come to the second claim of the theorem. Existence of minimizers for fixed  $\varepsilon$  follows from  $L^\infty$ -coercivity of the fidelity term in  $F_{\varepsilon, g}(\omega)$  and the lower semicontinuity assumption (13). The last statement is true due to the fundamental property of  $\Gamma$ -convergence except that we have to prove compactness in  $L^q(D, \mathbb{R}^m)$ . As shown for the

lower bound, any sequence  $u_\varepsilon$  as in the statement is bounded in  $L^q(D, \mathbb{R}^m)$  and therefore Lemma 5.6 yields that, up to subsequences,  $u_\varepsilon \rightarrow u$  in  $L^1(D, \mathbb{R}^m)$  for some  $u \in L^q(D, \mathbb{R}^m)$ . Clearly  $u_\varepsilon$  is a recovery sequence for this  $u$ , so that

$$F_g(u) = \lim_{\varepsilon \rightarrow 0} F_{\varepsilon, g}(\omega)(u_\varepsilon).$$

Repeating the reasoning for (95) we conclude from the above limit that

$$\lim_{\varepsilon \rightarrow 0} \int_D \gamma_\varepsilon(\omega)(z) |u_\varepsilon(z) - g_\varepsilon(\omega)(z)|^q dz = \int_D \gamma |u(z) - g(z)|^q dz, \quad (97)$$

where  $\gamma_\varepsilon(\omega)$  is defined in (92). Now consider the non-negative sequence  $a_\varepsilon := |u_\varepsilon - g_\varepsilon(\omega)|^q$ . By (97) and the qualitative lower bound  $\gamma_\varepsilon(\omega)(z) \geq c$  this sequence is bounded in  $L^1(D)$ . By the biting lemma (see [40, Lemma 2.63]) and Remark 6.2 we find a subsequence (not relabeled) such that  $a_\varepsilon \xrightarrow{b} |u - g|^q$ . Taking the same sets  $S_j$  as for the biting convergence of  $a_\varepsilon$  one can prove that the product  $\gamma_\varepsilon(\omega)a_\varepsilon$  converges in the biting sense to  $\gamma|u - g|^q$ . Indeed, on  $D \setminus S_j$  the sequence  $a_\varepsilon$  is equiintegrable by the Dunford–Pettis theorem and thus strongly convergent in  $L^1(D \setminus S_j)$ . Then by the usual product rules we obtain  $\gamma_\varepsilon(\omega)a_\varepsilon \rightarrow \gamma|u - g|^q$  in  $L^1(D \setminus S_j)$ , which shows biting convergence. Now we use that  $\gamma_\varepsilon(\omega)a_\varepsilon$  is nonnegative. By (97) and [40, Proposition 2.67] this yields that  $\gamma_\varepsilon(\omega)a_\varepsilon \rightarrow \gamma|u - g|^q$  also in  $L^1(D)$ . Thus both sequences  $\gamma_\varepsilon(\omega)a_\varepsilon$  and  $a_\varepsilon$  are equiintegrable on  $D$ . By Vitali’s convergence theorem we obtain that

$$\lim_{\varepsilon \rightarrow 0} \|u_\varepsilon - g_\varepsilon\|_{L^q(D)} = \|u - g\|_{L^q(D)},$$

which, by uniform convexity of  $L^q(D, \mathbb{R}^m)$  and (21), implies that  $u_\varepsilon \rightarrow u$  in  $L^q(D, \mathbb{R}^m)$  as claimed.  $\square$

## Conflict of interest statement

The author declares that he has no conflict of interest.

## Acknowledgements

MR acknowledges financial support from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2014–2019 Grant Agreement QUANTHOM 335410). The author further appreciates the helpful suggestions of two anonymous referees.

## Appendix A

In this appendix we provide a suitable framework to use diagonal arguments along  $\Gamma$ -converging discrete energies. We need this step since the general theory [35, Chapter 10] to construct a metric for  $\Gamma$ -convergence requires an  $L^p$ -coercive lower bound for the discrete energies.

Given any function  $G : L^p(B_1, \mathbb{R}^m) \rightarrow [0, +\infty]$  not identically  $+\infty$  we define its Moreau–Yosida approximation for  $\gamma > 0$  as

$$G^\gamma(u) = \inf_{v \in L^p(B_1, \mathbb{R}^m)} \left( G(v) + \gamma \|u - v\|_{L^p(B_1)}^p \right).$$

For  $p \geq 1$  the functional  $G^\gamma$  is locally Lipschitz-continuous on  $L^p(B_1, \mathbb{R}^m)$  (see [35, Theorem 9.15]). Let  $\{w_k\}_{k \in \mathbb{N}}$  be a dense subset of  $L^p(B_1, \mathbb{R}^m)$  containing 0. Given two lower semicontinuous functions  $G, H : L^p(B_1, \mathbb{R}^m) \rightarrow [0, +\infty]$  not identically  $+\infty$  we define their distance by

$$\mathfrak{d}(G, H) = \sum_{i, k \in \mathbb{N}} \frac{1}{2^{i+k}} |\arctan(G^i(w_k)) - \arctan(H^i(w_k))|.$$

Note that on lower semicontinuous functions  $\mathfrak{d}$  is indeed a distance, since  $\mathfrak{d}(G, H) = 0$  implies by local Lipschitz continuity that  $G^i = H^i$  for all  $i \in \mathbb{N}$ . Letting  $i \rightarrow +\infty$  it follows by lower semicontinuity that  $G = H$  (see [35, Remark 9.11]). In order to state our result we need further notation: let  $h : B_1 \times \mathbb{R}^{m \times d} \rightarrow [0, +\infty)$  be a Carathéodory-function such that  $\xi \mapsto h(x, \xi)$  is quasiconvex for a.e.  $x \in B_1$  and

$$\frac{1}{C}|\xi|^p - C \leq h(x, \xi) \leq C(|\xi|^p + 1).$$

Define the functional  $\mathfrak{E}_h : L^p(B_1, \mathbb{R}^m) \rightarrow [0, +\infty]$  by

$$\mathfrak{E}_h(u) = \begin{cases} \int_{B_1} h(x, \nabla u(x)) \, dx & \text{if } u \in W^{1,p}(B_1, \mathbb{R}^m), \\ +\infty & \text{otherwise.} \end{cases}$$

Then we have the following result.

**Lemma A.1.** *Consider a sequence  $\varepsilon_j \rightarrow 0$  and let  $x_0 \in D$  and  $\varrho_j > 0$  be such that  $B_{\varrho_j}(x_0) \subset D$  and  $\varepsilon_j/\varrho_j \rightarrow 0$ . Let further  $h : B_1 \times \mathbb{R}^{m \times d} \rightarrow [0, +\infty)$  be a Carathéodory-function as above and define  $G_{\varepsilon_j, \varrho_j}(x_0, \omega)$  as in Lemma 5.12. Then the following are equivalent:*

- (i)  $\Gamma(L^p(B_1, \mathbb{R}^m))\text{-}\lim_j G_{\varepsilon_j, \varrho_j}(x_0, \omega) = \mathfrak{E}_h$ ,
- (ii)  $\lim_j \mathfrak{d}(G_{\varepsilon_j, \varrho_j}(x_0, \omega), \mathfrak{E}_h) = 0$ .

**Proof.** (ii)  $\Rightarrow$  (i): First note that both functionals are lower semicontinuous on  $L^p(B_1, \mathbb{R}^m)$  and not identically  $+\infty$ . Assumption (ii) implies that

$$\lim_j G_{\varepsilon_j, \varrho_j}^i(x_0, \omega)(w_k) = \mathfrak{E}_h^i(w_k)$$

for all  $i, k \in \mathbb{N}$ . Since  $0 \in \{w_k\}$  and  $G_{\varepsilon_j, \varrho_j}^i(x_0, \omega)(0) = 0$ , we deduce from [35, Theorem 9.15] that the sequence  $G_{\varepsilon_j, \varrho_j}^i(x_0, \omega)$  is locally equicontinuous, so that the convergence extends to  $L^p(B_1, \mathbb{R}^m)$  by density. The claim then follows from a general characterization of  $\Gamma$ -convergence (see [35, Theorem 9.5]) when we let  $i \rightarrow +\infty$ .

(i)  $\Rightarrow$  (ii): Clearly  $G_{\varepsilon_j, \varrho_j}^\gamma(x_0, \omega)(u) \leq \gamma \|u\|_{L^p(B_1)}$ , so that given a sequence  $u_j$  such that

$$G_{\varepsilon_j, \varrho_j}(x_0, \omega)(u_j) + \gamma \|u_j - u\|_{L^p(B_1)}^p \leq G_{\varepsilon_j, \varrho_j}^\gamma(x_0, \omega)(u) + \frac{1}{j}, \quad (\text{A.1})$$

$u_j$  is bounded in  $L^p(B_1, \mathbb{R}^m)$ . Since the discrete density  $f$  satisfies  $f(\mathbf{p}) \leq C_f \|\mathbf{p}\|_1$ , we have by definition

$$G_{\varepsilon_j, \varrho_j}(x_0, \omega)(u_j) = E_{\varepsilon_j/\varrho_j}(\omega)(u_j(\cdot - x_0/\varrho_j), B_1(x_0)) \geq \frac{1}{C} F_{\varepsilon_j/\varrho_j}(\omega)(u_j(\cdot - x_0/\varrho_j), B_1(x_0)).$$

From Lemma 5.6 and a change of variables we conclude that  $u_j$  is compact in  $L^1(B_1, \mathbb{R}^m)$ . Hence there exists  $v \in L^p(B_1, \mathbb{R}^m)$  such that, up to subsequences,  $u_j \rightarrow v$  in  $L^1(B_1, \mathbb{R}^m)$  and additionally  $u_j \rightharpoonup v$  in  $L^p(B_1, \mathbb{R}^m)$ . In order to use the assumption (i), we use a truncation argument. Note that  $T_k u_j \rightarrow T_k v$  in  $L^p(D, \mathbb{R}^m)$ . Then by  $\Gamma$ -convergence in  $L^p(D, \mathbb{R}^m)$  and decrease by truncation of  $G_{\varepsilon_j, \varrho_j}(x_0, \omega)$  we obtain

$$\liminf_j G_{\varepsilon_j, \varrho_j}(x_0, \omega)(u_j) \geq \liminf_j G_{\varepsilon_j, \varrho_j}(x_0, \omega)(T_k u_j) \geq \int_{B_1} h(x, \nabla T_k v) \, dx.$$

Using Lemma 2.1 we can pass to the limit in  $k$  by dominated convergence. Combined with the weak lower semicontinuity of the  $L^p$ -norm we infer from (A.1) that

$$\mathfrak{E}_h(v) + \gamma \|v - u\|_{L^p(B_1)}^p \leq \liminf_j G_{\varepsilon_j, \varrho_j}^\gamma(x_0, \omega)(u).$$

Moreover, given any  $\tilde{u} \in L^p(B_1, \mathbb{R}^m)$  let us consider a sequence  $\tilde{u}_j \rightarrow \tilde{u}$  in  $L^p(B_1, \mathbb{R}^m)$  such that  $\lim_j G_{\varepsilon_j, \varrho_j}(x_0, \omega)(\tilde{u}_j) = \mathfrak{E}_h(\tilde{u})$ . Then by the definition of the Moreau–Yosida transformation

$$\limsup_j G_{\varepsilon_j, \varrho_j}^\gamma(x_0, \omega)(u) \leq \lim_j (G_{\varepsilon_j, \varrho_j}(x_0, \omega)(\tilde{u}_j) + \gamma \|\tilde{u}_j - u\|_{L^p(B_1)}^p) = \mathfrak{E}_h(\tilde{u}) + \gamma \|\tilde{u} - u\|_{L^p(B_1)}^p.$$

Combined with the previous inequality we obtain that  $\mathfrak{E}_h^\gamma(u) = \mathfrak{E}_h(v) + \gamma \|v - u\|_{L^p(B_1)}^p$  and by setting  $\tilde{u} = v$  we showed that

$$\lim_j G_{\varepsilon_j, q_j}^\gamma(x_0, \omega)(u) = \mathfrak{E}_h^\gamma(u)$$

for all  $u \in L^p(B_1, \mathbb{R}^m)$  and all  $\gamma > 0$ . This property implies (ii) by the definition of the metric.  $\square$

**Remark A.2.** The equivalence of Lemma A.1 remains valid for two functionals  $\mathfrak{E}_{h_j}$  and  $\mathfrak{E}_h$  provided the Carathéodory functions  $h_j$  satisfy growth conditions uniformly in  $j$ . In this case the proof simplifies since the Moreau–Yosida transformations are equicoercive in  $L^p(B_1, \mathbb{R}^m)$  due to the Sobolev embedding.

## References

- [1] E. Acerbi, N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Ration. Mech. Anal.* 86 (1984) 125–145.
- [2] M.A. Akcoglu, U. Krengel, Ergodic theorems for superadditive processes, *J. Reine Angew. Math.* 323 (1981) 53–67.
- [3] R. Alicandro, A. Braides, M. Cicalese, Phase and anti-phase boundaries in binary discrete systems: a variational viewpoint, *Netw. Heterog. Media* 1 (2006) 85–107.
- [4] R. Alicandro, M. Cicalese, A general integral representation result for continuum limits of discrete energies with superlinear growth, *SIAM J. Math. Anal.* 36 (2004) 1–37.
- [5] R. Alicandro, M. Cicalese, A. Gloria, Integral representation results for energies defined on stochastic lattices and application to nonlinear elasticity, *Arch. Ration. Mech. Anal.* 200 (2011) 881–943.
- [6] R. Alicandro, M. Cicalese, M. Ruf, Domain formation in magnetic polymer composites: an approach via stochastic homogenization, *Arch. Ration. Mech. Anal.* 218 (2015) 945–984.
- [7] L. Ambrosio, Existence theory for a new class of variational problems, *Arch. Ration. Mech. Anal.* 111 (1990) 291–322.
- [8] L. Ambrosio, A. Braides, Functionals defined on partitions of sets of finite perimeter I: integral representation and  $\Gamma$ -convergence, *J. Math. Pures Appl.* 69 (1990) 285–305.
- [9] L. Ambrosio, A. Braides, Functionals defined on partitions of sets of finite perimeter II: semicontinuity, relaxation and homogenization, *J. Math. Pures Appl.* 69 (1990) 307–333.
- [10] L. Ambrosio, E. De Giorgi, New functionals in the calculus of variations (Italian), *Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat.* 82 (1988) 199–210.
- [11] L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.
- [12] L. Ambrosio, V.M. Tortorelli, Approximation of functionals depending on jumps by elliptic functionals via  $\Gamma$ -convergence, *Commun. Pure Appl. Math.* 43 (1990) 999–1036.
- [13] L. Ambrosio, V.M. Tortorelli, On the approximation of free discontinuity problems, *Boll. Unione Mat. Ital.* 6–B (1992) 105–123.
- [14] A. Bach, Anisotropic free-discontinuity functionals as the  $\Gamma$ -limit of second order elliptic functionals, *ESAIM Control Optim. Calc. Var.* (2017), <https://doi.org/10.1051/cocv/2017027>.
- [15] A. Bach, A. Braides, C.I. Zeppieri, Quantitative analysis of finite-difference approximations of free-discontinuity problems, Preprint, arXiv: 1807.05346, 2018.
- [16] G. Bellettini, A. Coscia, Discrete approximation of a free discontinuity problem, *Numer. Funct. Anal. Optim.* 15 (1994) 201–224.
- [17] X. Blanc, C. Le Bris, P.L. Lions, The energy of some microscopic stochastic lattices, *Arch. Ration. Mech. Anal.* 184 (2007) 303–339.
- [18] G. Bouchitté, I. Fonseca, G. Leoni, L. Mascarenhas, A global method for relaxation in  $W^{1,p}$  and in  $SBV_p$ , *Arch. Ration. Mech. Anal.* 165 (2002) 187–242.
- [19] B. Bourdin, Image segmentation with a finite element method, *ESAIM: M2AN* 33 (1999) 229–244.
- [20] B. Bourdin, A. Chambolle, Implementation of an adaptive finite-element approximation of the Mumford–Shah functional, *Numer. Math.* 85 (2000) 609–646.
- [21] A. Braides,  *$\Gamma$ -Convergence for Beginners*, Oxford Lecture Series in Mathematics and Its Applications, vol. 22, Oxford University Press, Oxford, 2002.
- [22] A. Braides, M. Cicalese, M. Ruf, Continuum limit and stochastic homogenization of discrete ferromagnetic thin films, *Anal. PDE* 11 (2018) 499–553.
- [23] A. Braides, G. Dal Maso, Non-local approximation of the Mumford–Shah functional, *Calc. Var. Partial Differ. Equ.* 5 (1997) 293–322.
- [24] A. Braides, A. Defranceschi, E. Vitali, Homogenization of free discontinuity problems, *Arch. Ration. Mech. Anal.* 135 (1996) 297–356.
- [25] A. Braides, A. Piatnitski, Overall properties of a discrete membrane with randomly distributed defects, *Arch. Ration. Mech. Anal.* 189 (2008) 301–323.
- [26] A. Braides, N.K. Yip, A quantitative description of mesh dependence for the discretization of singularly perturbed nonconvex problems, *SIAM J. Numer. Anal.* 50 (2012) 1883–1898.
- [27] M. Burger, T. Esposito, C.I. Zeppieri, Second-order edge-penalization in the Ambrosio–Tortorelli functional, *Multiscale Model. Simul.* 13 (2015) 1354–1389.
- [28] F. Cagnetti, G. Dal Maso, L. Scardia, C.I. Zeppieri,  $\Gamma$ -convergence of free discontinuity problems, Preprint, arXiv:1712.07093, 2017.
- [29] A. Chambolle, Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations, *SIAM J. Appl. Math.* 55 (1995) 827–863.
- [30] A. Chambolle, Finite-differences discretizations of the Mumford–Shah functional, *ESAIM: M2A* 33 (1999) 261–288.
- [31] A. Chambolle, G. Dal Maso, Discrete approximation of the Mumford–Shah functional in dimension two, *ESAIM: M2A* 33 (1999) 651–672.

- [32] M. Cicalese, M. Ruf, Discrete spin systems on random lattices at the bulk scaling, *Discrete Contin. Dyn. Syst., Ser. S* 10 (2017) 101–117.
- [33] G. Cortesani, R. Toader, A density result in  $SBV$  with respect to non-isotropic energies, *Nonlinear Anal.* 38 (1999) 585–604.
- [34] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Applied Mathematical Sciences, vol. 78, Springer, New York, 2008.
- [35] G. Dal Maso, *An Introduction to  $\Gamma$ -Convergence*, Progress in Nonlinear Differential Equations and Their Applications, vol. 8, Birkhäuser Boston Inc., Boston, MA, 1993.
- [36] G. Dal Maso, G. Francfort, R. Toader, Quasistatic crack growth in nonlinear elasticity, *Arch. Ration. Mech. Anal.* 176 (2005) 165–225.
- [37] G. Dal Maso, L. Modica, Nonlinear stochastic homogenization, *Ann. Mat. Pura Appl.* 144 (1986) 347–389.
- [38] E. De Giorgi, M. Carriero, A. Leaci, Existence theorem for a minimum problem with free discontinuity set, *Arch. Ration. Mech. Anal.* 108 (1989) 195–218.
- [39] M. Focardi, Fine regularity results for Mumford–Shah minimizers: porosity, higher integrability and the Mumford–Shah conjecture, in: N. Fusco, A. Pratelli (Eds.), *Free Discontinuity Problems*, in: PSNS, vol. 19, 2016, Pisa.
- [40] I. Fonseca, G. Leoni, *Modern Methods in the Calculus of Variations:  $L^p$  Spaces*, Springer, New York, 2010.
- [41] I. Fonseca, S. Müller, P. Pedregal, Analysis of concentration and oscillation effects generated by gradients, *SIAM J. Math. Anal.* 29 (1998) 736–756.
- [42] N. García Trillos, D. Slepčev, Continuum limit of total variation on point clouds, *Arch. Ration. Mech. Anal.* 220 (2016) 193–241.
- [43] A. Giacomini, M. Ponsiglione, A  $\Gamma$ -convergence approach to stability of unilateral minimality properties in fracture mechanics and applications, *Arch. Ration. Mech. Anal.* 180 (2006) 399–447.
- [44] A. Gloria, M.D. Penrose, Random parking, Euclidean functionals, and rubber elasticity, *Commun. Math. Phys.* 321 (2013) 1–31.
- [45] M. Gobbino, Finite difference approximation of the Mumford–Shah functional, *Commun. Pure Appl. Math.* 51 (1998) 197–228.
- [46] J. Heinonen, P. Koskela, Quasiconformal maps in metric spaces with controlled geometry, *Acta Math.* 181 (1998) 1–61.
- [47] U. Krengel, *Ergodic Theorems*, De Gruyter Studies in Mathematics, vol. 6, De Gruyter, Berlin, 1985.
- [48] C.J. Larsen, On the representation of effective energy densities, *ESAIM Control Optim. Calc. Var.* 5 (2000) 529–538.
- [49] A. Lemenant, A selective review on Mumford–Shah minimizers, *Boll. Unione Mat. Ital.* 9 (2016) 69–113.
- [50] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Commun. Pure Appl. Math.* 42 (1989) 577–685.
- [51] M.D. Penrose, Random parking, sequential adsorption, and the jamming limit, *Commun. Math. Phys.* 218 (2001) 153–176.
- [52] T. Pock, D. Cremers, H. Bischof, A. Chambolle, An algorithm for minimizing the Mumford–Shah functional, in: *IEEE 12th International Conference on Computer Vision*, 2000, pp. 1133–1140.
- [53] E. Strekalovskiy, D. Cremers, Real-time minimization of the piecewise smooth Mumford–Shah functional, *Comput. Vis. ECCV 2014* (2014) 127–141.