Calc. Var. (2022) 61:183
https://doi.org/10.1007/s00526-022-02293-6

Calculus of Variations

Regularity of gradient vector fields giving rise to finite Caccioppoli partitions

Roger Moser[1]

Received: 19 January 2022 / Accepted: 5 July 2022 / Published online: 15 July 2022

Abstract
For a finite set $A \subseteq \mathbb{R}^n$, consider a function $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ such that $\nabla u \in A$ almost everywhere. If A is convex independent, then it follows that u is piecewise affine away from a closed, countably $\mathcal{H}^{n-1}$-rectifiable set. If A is affinely independent, then u is piecewise affine away from a closed $\mathcal{H}^{n-1}$-null set.

Mathematics Subject Classification 49Q20 · 46E35

1 Introduction

For $n \in \mathbb{N}$, consider a finite set $A \subseteq \mathbb{R}^n$. We study continuous functions $u\colon \mathbb{R}^n \to \mathbb{R}$ such that the weak gradient ∇u satisfies $\nabla u \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^n)$ and $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. This means that whenever $\Omega \subseteq \mathbb{R}^n$ is open and bounded, the sets $\{x \in \Omega\colon \nabla u(x) = a\}$, for $a \in A$, form a Caccioppoli partition of Ω as discussed, e.g., by Ambrosio et al. [1, Sect. 4.4]. The theory of Caccioppoli partitions therefore applies and gives some information on the structure of ∇u and of u. The fact that we are dealing with a gradient, however, gives rise to a better theory, especially under additional assumptions on the geometry of A. We work with the following notions in this paper.

Definition 1 A set $A \subset \mathbb{R}^n$ is called *convex independent* if any $a \in A$ does not belong to the convex hull of $A \setminus \{a\}$. It is called *affinely independent* if any $a \in A$ does not belong to the affine span of $A \setminus \{a\}$.

If either of these conditions is satisfied, then we can prove statements on the regularity of u that finite Caccioppoli partitions do not share in general. In fact, we will see that u is locally piecewise affine away from a closed, countably $\mathcal{H}^{n-1}$-rectifiable set (if A is convex independent) or away from a closed $\mathcal{H}^{n-1}$-null set (if A is affinely independent).

In order to make this more precise, we introduce some notation. Given $r > 0$ and $x \in \mathbb{R}^n$, we write $B_r(x)$ for the open ball of radius r centred at x. Given $a \in \mathbb{R}^n$, the function

Communicated by X. Ros-Oton.

✉ Roger Moser
r.moser@bath.ac.uk

[1] Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK

$\lambda_a : \mathbb{R}^n \to \mathbb{R}$ is defined by $\lambda_a(x) = a \cdot x$ for $x \in \mathbb{R}^n$. Given two functions $v, w : \mathbb{R}^n \to \mathbb{R}$, we write $v \wedge w$ and $v \vee w$, respectively, for the functions with $(v \wedge w)(x) = \min\{v(x), w(x)\}$ and $(v \vee w)(x) = \max\{v(x), w(x)\}$ for $x \in \mathbb{R}^n$.

Definition 2 Given a function $u : \mathbb{R}^n \to \mathbb{R}$, the *regular set* of u, denoted by $\mathcal{R}(u)$, consists of all $x \in \mathbb{R}^n$ such that there exist $a, b \in \mathbb{R}^n$, $c \in \mathbb{R}$, and $r > 0$ with $u(\cdot - x) = \lambda_a \wedge \lambda_b + c$ in $B_r(0)$ or $u(\cdot - x) = \lambda_a \vee \lambda_b + c$ in $B_r(0)$. The *singular set* of u is its complement $\mathcal{S}(u) = \mathbb{R}^n \setminus \mathcal{R}(u)$.

The condition for $\mathcal{R}(u)$ allows the possibility that $a = b$, in which case u is affine near x. If $a \neq b$, then it is still piecewise affine near x. Obviously $\mathcal{R}(u)$ is an open set and $\mathcal{S}(u)$ is closed.

It would be reasonable to include functions consisting of more than two affine pieces in the definition of $\mathcal{R}(u)$, for example $(\lambda_{a_1} \wedge \lambda_{a_2}) \vee \lambda_{a_3} + c$ for $a_1, a_2, a_3 \in \mathbb{R}^n$ and $c \in \mathbb{R}$. For the results of this paper, however, this would make no difference, therefore we choose the simpler definition.

For $s \geq 0$, we denote the s-dimensional Hausdorff measure in $\mathbb{R}^n$ by $\mathcal{H}^s$. The notation $\mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ is used for the space of functions with weak gradient in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^n)$. Thus the hypotheses of the following theorems are identical to the assumptions at the beginning of the introduction.

Theorem 3 *Suppose that A is a finite, affinely independent set. Let $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ with $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. Then $\mathcal{H}^{n-1}(\mathcal{S}(u)) = 0$.*

Theorem 4 *Suppose that A is a finite, convex independent set. Let $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ with $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. Then $\mathcal{S}(u)$ is countably $\mathcal{H}^{n-1}$-rectifiable.*

For $n = 2$, Theorem 3 was proved in a previous paper [10]. For higher dimensions, the result is new. Theorem 4 is new even for $n = 2$. For $n = 1$, both statements are easy to prove.

The results are optimal in terms of the Hausdorff measures involved. Furthermore, the assumption of convex/affine independence is necessary. Indeed, there are examples of finite sets $A \subseteq \mathbb{R}^2$ and functions $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^2)$ with $\nabla u(x) \in A$ almost everywhere such that

- $\mathcal{H}^2(\mathcal{S}(u)) > 0$; or
- $\mathcal{H}^1(\mathcal{S}(u)) > 0$ and A is convex independent; or
- $\mathcal{H}^s(\mathcal{S}(u)) = \infty$ for any $s < 1$ and A is affinely independent.

All of these can be found in the author's previous paper [10].

Our results are easily adapted to functions defined on an open set $\Omega \subseteq \mathbb{R}^n$ with the corresponding properties. For the sake of simplicity, we do not discuss this case in detail.

Apart from being of obvious geometric interest, functions as described above appear in problems from materials science. They naturally arise as limits in Γ-convergence theories in the spirit of Modica and Mortola [8, 9] for quantities such as

$$\int_\Omega \left(\epsilon |\nabla^2 u|^2 + \frac{W(\nabla u)}{\epsilon} \right) dx, \tag{1}$$

where $\Omega \subseteq \mathbb{R}^n$ is an open set and $W : \mathbb{R}^n \to [0, \infty)$ is a function with $A = W^{-1}(\{0\})$. Functionals of this sort appear in certain models for the surface energy of nanocrystals [7, 13, 14]. For $\Omega \subseteq \mathbb{R}^2$, functions $u \in \mathrm{BV}^2(\Omega)$ with $\nabla u \in \{(\pm 1, 0), (0, \pm 1)\}$ have also been used by Cicalese et al. [3] for a different sort of Γ-limit arising from a model for frustrated spin systems.

Functionals similar to (1), but for maps $u\colon \Omega \to \mathbb{R}^n$, also appear in certain models for phase transitions in elastic materials (see, e.g., the seminal paper of Ball and James [2] or the introduction into the theory by Müller [11]). In this context, due to the frame indifference of the underlying models, the set $W^{-1}(\{0\})$ is typically not finite. Sometimes, however, the frame indifference is disregarded (as in the paper by Conti et al. [4]), or the theory gives a limit with $\nabla u \in \mathrm{BV}(\Omega; A)$ for a finite set $A \subseteq \mathbb{R}^{n\times n}$ anyway (such as in recent results of Davoli and Friedrich [5, 6]). In such a case, Theorem 3 and Theorem 4 are potentially useful, as they apply to the components (or other one-dimensional projections) of u.

In the proof of Theorem 3, we use some of the tools from the author's previous paper [10]. In particular, we will analyse the intersections of the graph of u with certain hyperplanes in $\mathbb{R}^{n+1}$. We will see that these intersections correspond to the graphs of functions with $(n-1)$-dimensional domains and with properties similar to u. The key ideas from the previous paper, however, are specific to $\mathbb{R}^2$, so we eventually use different arguments. In this paper, we use the theory of $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^n)$ to a much greater extent. The central argument will consider approximate jump points of ∇u. Near such a point, we know that u is close to a piecewise affine function in a measure theoretic sense by definition. We then use an induction argument (with induction over n) to show that u is in fact piecewise affine near $\mathcal{H}^{n-1}$-almost every approximate jump point.

We also need to analyse points where u has an approximate limit, and they are of interest for the proofs of both Theorems 3 and 4. This part of the analysis is significantly simpler and relies on the fact that for any $a \in A$, the function $v(x) = u(x) - a\cdot x$ has some monotonicity properties.

In the rest of the paper, we study a fixed function $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ with $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. Then u is automatically Lipschitz continuous. Since we are interested only in the local properties of u, we may assume that it is also bounded. (Otherwise we can modify it outside of a bounded set with the construction described in [10, Sect. 6].) We define the function $\boldsymbol{U}\colon \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$\boldsymbol{U}(x) = \begin{pmatrix} x \\ u(x) \end{pmatrix}, \quad x \in \mathbb{R}^n.$$

We use the notation $\mathrm{graph}(u) = \boldsymbol{U}(\mathbb{R}^n)$ for the graph of u.

As we sometimes work with points in $\mathbb{R}^{n+1}$ (especially points on $\mathrm{graph}(u)$) and their projections onto $\mathbb{R}^n$ simultaneously, we use the following notation. A generic point in $\mathbb{R}^{n+1}$ is denoted by $\boldsymbol{x} = (x_1, \dots, x_{n+1})^T$, and then we write $x = (x_1, \dots, x_n)^T$. Thus $\boldsymbol{x} = \left(\begin{smallmatrix} x \\ x_{n+1} \end{smallmatrix}\right)$. We think of elements of $\mathbb{R}^n$ and of $\mathbb{R}^{n+1}$ as column vectors, and this is sometimes important, as we use them as columns in certain matrices.

As our function satisfies in particular the condition $\nabla u \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^n)$, the theory of this space will of course be helpful. In this context, we mostly follow the notation and terminology of Ambrosio et al. [1]. We also use several of the results found in this book.

2 Approximate faces and edges of the graph

In this section, we decompose $\mathbb{R}^n$ into three sets $\mathcal{F}$, $\mathcal{E}$, and $\mathcal{N}$. These are defined such that we expect regularity in $\mathcal{F}$ under the assumptions of either of the main theorems, and also in $\mathcal{E}$ under the assumptions of Theorem 3. The third set, $\mathcal{N}$, will be an $\mathcal{H}^{n-1}$-null set. The sets $\mathcal{F}$ and $\mathcal{E}$ are characterised, up to $\mathcal{H}^{n-1}$-null sets, by the condition that ∇u has an approximate limit or an approximate jump, respectively. Since much of our analysis examines $\mathrm{graph}(u)$, it is also convenient to think of $\mathcal{F}$ as the set of points where the graph behaves approximately

like the (n-dimensional) faces of a polyhedral surface, whereas $\mathcal{E}$ corresponds to approximate (($n-1$)-dimensional) edges.

First, however, we define three related sets $\mathcal{F}'$, $\mathcal{E}'$, and $\mathcal{N}'$. The sets $\mathcal{F}$, $\mathcal{E}$, and $\mathcal{N}$ will be derived from these later.

Consider the set $\mathcal{F}' \subseteq \mathbb{R}^n$, comprising all points $x \in \mathbb{R}^n$ such that there exists $a \in \mathbb{R}^n$ satisfying

$$\lim_{r \searrow 0} \fint_{B_r(x)} |\nabla u - a| \, d\mathcal{H}^n = 0.$$

In other words, this is the set of all points where ∇u has an approximate limit a. It is then clear that $a \in A$. The complement $\mathbb{R}^n \setminus \mathcal{F}'$ is called the approximate discontinuity set of ∇u.

Furthermore, let $\mathcal{E}'$ be the set of all $x \in \mathbb{R}^n$ such that there exist $a_-, a_+ \in \mathbb{R}^n$ with $a_- \neq a_+$ and there exists $\eta \in S^{n-1}$ such that

$$\lim_{r \searrow 0} \fint_{\{\tilde{x} \in B_r(x) \colon (\tilde{x}-x)\cdot\eta > 0\}} |\nabla u - a_+| \, d\mathcal{H}^n = 0 \tag{2}$$

and

$$\lim_{r \searrow 0} \fint_{\{\tilde{x} \in B_r(x) \colon (\tilde{x}-x)\cdot\eta < 0\}} |\nabla u - a_-| \, d\mathcal{H}^n = 0. \tag{3}$$

This is the approximate jump set of ∇u. Again, the points a_-, a_+ belong to A.

According to a result by Federer and Vol'pert (which can be found in the book by Ambrosio et al. [1, Theorem 3.78]), there exists an $\mathcal{H}^{n-1}$-null set $\mathcal{N}' \subseteq \mathbb{R}^n$ such that

$$\mathbb{R}^n = \mathcal{F}' \cup \mathcal{E}' \cup \mathcal{N}'.$$

Furthermore, the set $\mathcal{E}'$ is countably $\mathcal{H}^{n-1}$-rectifiable.

Given $x \in \mathbb{R}^n$ and $\rho > 0$, we define the function $u_{x,\rho} \colon \mathbb{R}^n \to \mathbb{R}$ with

$$u_{x,\rho}(\tilde{x}) = \frac{1}{\rho} \left(u(x + \rho\tilde{x}) - u(x) \right)$$

for $\tilde{x} \in \mathbb{R}^n$. For x fixed, the family of functions $(u_{x,\rho})_{\rho>0}$ is clearly bounded in $C^{0,1}(K)$ for any compact set $K \subseteq \mathbb{R}^n$. Therefore, the theorem of Arzelà–Ascoli implies that there exists a sequence $\rho_k \searrow 0$ such that u_{x,ρ_k} converges locally uniformly. If we have in fact a limit for $\rho \searrow 0$, then we write

$$T_x u = \lim_{\rho \searrow 0} u_{x,\rho}$$

and call this limit the *tangent function* of u at x.

If $x \in \mathcal{F}'$ and $a \in A$ is the approximate limit of ∇u at x, then for any sequence $\rho_k \searrow 0$, the limit of u_{x,ρ_k} can only be λ_a. Hence in this case, there exists a tangent function $T_x u$, which is exactly this function. Similarly, if $x \in \mathcal{E}'$, then $T_x u$ exists and

$$T_x u(\tilde{x}) = \begin{cases} \lambda_{a_-}(\tilde{x}) & \text{if } \tilde{x} \cdot \eta < 0, \\ \lambda_{a_+}(\tilde{x}) & \text{if } \tilde{x} \cdot \eta \geq 0. \end{cases}$$

Because $T_x u$ is a continuous function, this means that

$$\eta = \pm \frac{a_+ - a_-}{|a_+ - a_-|}. \tag{4}$$

Then we conclude that $T_x u = \lambda_{a_-} \wedge \lambda_{a_+}$ or $T_x u = \lambda_{a_-} \vee \lambda_{a_+}$, depending on the sign.

If we consider the functions $a_-, a_+ \colon \mathcal{E}' \to A$ and $\eta \colon \mathcal{E}' \to S^{n-1}$ such that (2) and (3) are satisfied on $\mathcal{E}'$, then the previously used result [1, Theorem 3.78] also implies that

$$D\nabla u \llcorner \mathcal{E}' = (a_+ - a_-) \otimes \eta \, \mathcal{H}^{n-1} \llcorner \mathcal{E}'. \tag{5}$$

Let $\gamma = \min\{|a-b| \colon a, b \in A\}$. Then for any Borel set $\Omega \subseteq \mathbb{R}^n$, we conclude that

$$|D\nabla u|(\Omega) \geq \gamma \mathcal{H}^{n-1}(\mathcal{E}' \cap \Omega).$$

Hence $\mathcal{E}'$ has locally finite $\mathcal{H}^{n-1}$-measure.

Now define

$$\mathcal{F} = \left\{ x \in \mathcal{F}' \colon \lim_{\rho \searrow 0} \rho^{1-n} |D\nabla u|(B_r(x)) = 0 \right\}.$$

Then standard results [1, Theorem 2.56 and Lemma 3.76] imply that $\mathcal{H}^{n-1}(\mathcal{F}' \setminus \mathcal{F}) = 0$.

Recall the map $\boldsymbol{U} \colon \mathbb{R}^n \to \mathbb{R}^{n+1}$ defined in the introduction. Set $\mathcal{F}^* = \boldsymbol{U}(\mathcal{F})$ and $\mathcal{E}^\dagger = \boldsymbol{U}(\mathcal{E}')$. Then $\mathcal{E}^\dagger$ is a countably $\mathcal{H}^{n-1}$-rectifiable subset of $\mathbb{R}^{n+1}$. Hence at $\mathcal{H}^{n-1}$-almost every $\boldsymbol{x} \in \mathcal{E}^\dagger$, the measure $\mathcal{H}^{n-1} \llcorner \mathcal{E}^\dagger$ has a tangent measure [1, Theorem 2.83] of the form $\mathcal{H}^{n-1} \llcorner T_{\boldsymbol{x}} \mathcal{E}^\dagger$, where $T_{\boldsymbol{x}} \mathcal{E}^\dagger$ is an $(n-1)$-dimensional linear subspace of $\mathbb{R}^{n+1}$ (the approximate tangent space of $\mathcal{E}^\dagger$ at $\boldsymbol{x}$). Let $\mathcal{E}^*$ be the set of all $\boldsymbol{x} \in \mathcal{E}^\dagger$ where this is the case. Furthermore, let $\mathcal{E} = \boldsymbol{U}^{-1}(\mathcal{E}^*)$. Then $\mathcal{E}' \setminus \mathcal{E}$ is an $\mathcal{H}^{n-1}$-null set.

Thus if we define $\mathcal{N} = \mathbb{R}^n \setminus (\mathcal{F} \cup \mathcal{E})$, then $\mathcal{N}$ is an $\mathcal{H}^{n-1}$-null set and we have the disjoint decomposition

$$\mathbb{R}^n = \mathcal{F} \cup \mathcal{E} \cup \mathcal{N}.$$

For later use, we also define $\mathcal{N}^* = \boldsymbol{U}(\mathcal{N})$.

3 Proof of Theorem 4

In this section we prove our second main result, Theorem 4. The proof is based on the following proposition, which will also be useful for the Proof of Theorem 3 later on.

Proposition 5 *Suppose that $A \subseteq \mathbb{R}^n$ is finite and convex independent. Let $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ be a function with $\nabla u(x) \in A$ for almost all $x \in \mathbb{R}^n$. Then there exist $r > 0$ and $\epsilon > 0$ with the following property. Suppose that there exists $a \in A$ such that*

$$\mathcal{H}^n(\{x \in B_1(0) \colon \nabla u(x) \neq a\}) \leq \epsilon \tag{6}$$

and

$$|D\nabla u|(B_1(0)) \leq \epsilon.$$

Then $\nabla u(x) = a$ for almost every $x \in B_r(0)$.

Proof Because A is convex independent, there exists $\omega \in S^{n-1}$ such that

$$a \cdot \omega < \min_{b \in A \setminus \{a\}} b \cdot \omega.$$

(For example, we may choose a point b_0 in the convex hull of $A \setminus \{a\}$ that minimises the distance to a. Then $\omega = (b_0 - a)/|b_0 - a|$ has this property.) As A is finite, there also exists $\delta \in (0, 1)$ such that the inequality $a \cdot \xi \leq \min_{b \in A \setminus \{a\}}(b \cdot \xi)$ holds even for ξ in the cone

$$C = \left\{ \xi \in \mathbb{R}^n \colon \xi \cdot \omega \geq \delta |\xi| \right\}.$$

Consider the function $v\colon \mathbb{R}^n \to \mathbb{R}$ with $v(x) = u(x) - a \cdot x$ for $x \in \mathbb{R}^n$. Then for any $\xi \in C$,

$$\xi \cdot \nabla v(x) = \xi \cdot \nabla u(x) - a \cdot \xi \geq 0$$

almost everywhere. Thus v is monotone along lines parallel to ξ. (This is true for every such line by the continuity of v.) Furthermore, for almost every $x \in \mathbb{R}^n$, we find that either $\nabla u(x) = a$ or $\omega \cdot \nabla v(x) > 0$.

Suppose that $\nabla u = a$ does *not* hold almost everywhere in $B_r(0)$. Then there exist $x_-, x_+ \in B_r(0)$ with $v(x_-) < v(x_+)$. Define

$$C_- = (x_- - C) \cap B_1(0) \quad \text{and} \quad C_+ = (x_+ + C) \cap B_1(0).$$

Then for any $x' \in C_-$ and $x'' \in C_+$, we conclude that

$$v(x') \leq v(x_-) < v(x_+) \leq v(x'').$$

We now foliate a part of $B_1(0)$ by line segments parallel to ω. For $R \in (0, 1]$, let $Z_R = \{x \in B_R(0)\colon \omega \cdot x = 0\}$. For every $z \in Z_R$, consider the line segment

$$L_z = \left\{ z + t\omega\colon\ -\frac{1}{2} \leq t \leq \frac{1}{2} \right\}.$$

Provided that r is chosen sufficiently small, we can find $R \in (0, 1]$ such that

$$\left\{ z - \frac{\omega}{2}\colon z \in Z_R \right\} \subseteq C_- \quad \text{and} \quad \left\{ z + \frac{\omega}{2}\colon z \in Z_R \right\} \subseteq C_+.$$

Hence for any $z \in Z_R$,

$$v\left(z + \frac{\omega}{2}\right) - v\left(z - \frac{\omega}{2}\right) \geq v(x_+) - v(x_-) > 0.$$

In particular, the restriction of v to the line segment L_z is not constant. For $z \in Z_R$, define $\Lambda_z = \{x \in L_z\colon \nabla u(x) = a\}$. Then it follows that $\mathcal{H}^1(\Lambda_z) < 1$ for $\mathcal{H}^{n-1}$-almost all $z \in Z_R$.

On the other hand, because of (6), we also know that

$$\mathcal{H}^{n-1}\left(\left\{z \in Z_R\colon \mathcal{H}^1(\Lambda_z) = 0\right\}\right) \leq \epsilon.$$

Thus if we define $Z' = \left\{z \in Z_R\colon 0 < \mathcal{H}^1(\Lambda_z) < 1\right\}$, then

$$\mathcal{H}^{n-1}(Z') \geq \mathcal{H}^{n-1}(Z_R) - \epsilon.$$

Set $c = \min_{b \in A} |a - b|$. For $\mathcal{H}^{n-1}$-almost any $z \in Z'$, the function $t \mapsto \nabla u(z + t\omega)$ belongs to $\mathrm{BV}\big((-\frac{1}{2}, \frac{1}{2}); \mathbb{R}^n\big)$ and its total variation is at least c. Hence [1, Theorem 3.103] implies that

$$|D\nabla u|(B_1(0)) \geq c\mathcal{H}^{n-1}(Z') \geq c\big(\mathcal{H}^{n-1}(Z_R) - \epsilon\big).$$

If ϵ is sufficiently small, then this means in particular that $|D\nabla u|(B_1(0)) > \epsilon$. Thus we have proved the contrapositive of Proposition 5. □

Proof of Theorem 4 We show that $\mathcal{F} \subseteq \mathcal{R}(u)$. To this end, fix $x \in \mathcal{F}$ and consider the rescaled functions $u_{x,\rho}$ for $\rho > 0$. Since $x \in \mathcal{F}$, we know that $\nabla u_{x,\rho} \to a$ in $L^1(B_1(0))$ as $\rho \searrow 0$ for some $a \in A$. Furthermore, since

$$|D\nabla u_{x,\rho}|(B_1(0)) = \rho^{1-n}|D\nabla u|(B_\rho(x)) \to 0$$

as $\rho \searrow 0$, the function $u_{x,\rho}$ satisfies the inequalities of Proposition 5 for ρ sufficiently small. Hence $\nabla u_{x,\rho}(\tilde{x}) = a$ for almost every $\tilde{x} \in B_r(0)$, which implies that

$$u(\tilde{x}) = u(x) + a \cdot (\tilde{x} - x)$$

for all $\tilde{x} \in B_{\rho r}(x)$. Hence $x \in \mathcal{R}(u)$.

We conclude that $\mathcal{S}(u) \subseteq \mathcal{E} \cup \mathcal{N}$. As seen in Sect. 2, the set $\mathcal{E}$ is countably $\mathcal{H}^{n-1}$-rectifiable and $\mathcal{N}$ is an $\mathcal{H}^{n-1}$-null set. Thus Theorem 4 follows. □

4 Specialising to a regular *n*-simplex

The rest of the paper is devoted to the Proof of Theorem 3. This proof relies on the same general strategy as the Proof of Theorem 4 to some extent: we use some monotonicity properties of u again, together with the theory of BV-functions, to show that if u is close to $\lambda_{a_1} \wedge \lambda_{a_2}$ (for two points $a_1, a_2 \in A$) in a cube centred at 0, and if $|D\nabla u|(Q)$ is not much larger than the corresponding quantity for $\lambda_{a_1} \wedge \lambda_{a_2}$, then u actually coincides with $\lambda_{a_1} \wedge \lambda_{a_2}$ up to a translation and addition of a constant in a smaller set. (The same applies to $\lambda_{a_1} \vee \lambda_{a_2}$.) The details, however, are much more involved than in the Proof of Theorem 4.

Instead of considering *any* affinely independent set A, it is convenient in this analysis to assume that $a_0, \ldots, a_n \in \mathbb{R}^n$ are the corners of a regular n-simplex of side length $\sqrt{2n+2}$ centred at 0, and that $A = \{a_0, \ldots, a_n\}$. This means in particular that

$$\sum_{i=0}^{n} a_i = 0. \tag{7}$$

We further assume that the matrix with columns $a_0 - a_1, \ldots, a_0 - a_n$ has a positive determinant. Theorem 3 can then be reduced to this situation by composing u with an affine transformation. The details are given on page 26 below.

As it is sometimes convenient to permute $a_0, \ldots, a_n$ cyclically, we regard $0, \ldots, n$ as members of $\mathbb{Z}_{n+1} = \mathbb{Z}/(n+1)\mathbb{Z}$ in this context. Thus $a_{i+n+1} = a_i$.

The condition that our simplex has side length $\sqrt{2n+2}$ means that $|a_i| = \sqrt{n}$ for every $i \in \mathbb{Z}_{n+1}$. Indeed, by the calculations of Parks and Wills [12], the dihedral angle of the regular n-simplex is $\arccos \frac{1}{n}$. As each a_i is orthogonal to one of the faces, this means that $a_i \cdot a_j = -\frac{1}{n}|a_i||a_j|$ for $i \neq j$, and therefore $2n+2 = |a_i - a_j|^2 = \frac{2n+2}{n}|a_i||a_j|$. From this we conclude that

$$|a_i| = \sqrt{n}$$

for $i \in \mathbb{Z}_{n+1}$ and

$$a_i \cdot a_j = -1$$

for $i \neq j$.

In the following arguments, we mostly study u in terms of $\mathrm{graph}(u)$. With A chosen as above, we have a positively oriented, orthonormal basis $(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_{n+1})$ of $\mathbb{R}^{n+1}$ (defined shortly) consisting of vectors normal to the expected faces of $\mathrm{graph}(u)$. (This is the reason why we choose this specific set A.) To facilitate our analysis, we will often represent $\mathrm{graph}(u)$ with respect to this basis, or equivalently, apply the corresponding linear transformation in $\mathrm{SO}(n+1)$.

The vectors $\boldsymbol{v}_i$ are defined as follows. For $i \in \mathbb{Z}_{n+1}$, let

$$\boldsymbol{v}_i = \frac{1}{\sqrt{n+1}} \begin{pmatrix} -a_i \\ 1 \end{pmatrix}.$$

Then

$$|\boldsymbol{v}_i|^2 = \frac{|a_i|^2 + 1}{n+1} = 1,$$

whereas for $i \neq j$,

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \frac{a_i \cdot a_j + 1}{n+1} = 0.$$

Hence $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n+1})$ is indeed an orthonormal basis of $\mathbb{R}^{n+1}$. Furthermore,

$$\begin{aligned} \det \begin{pmatrix} -a_1 & \cdots & -a_{n+1} \\ 1 & \cdots & 1 \end{pmatrix} &= \det \begin{pmatrix} a_0 - a_1 & \cdots & a_0 - a_n & -a_0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} a_0 - a_1 & \cdots & a_0 - a_n \end{pmatrix}. \end{aligned}$$

(In the first step, we have used the fact that $a_{n+1} = a_0$ and subtracted the last column from each of the other columns of the matrix.) Hence the above assumption guarantees that the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n+1})$ gives the standard orientation of $\mathbb{R}^{n+1}$.

We now use the notation $\lambda_i = \lambda_{a_i}$, recalling that this is the linear function with $\lambda_i(x) = a_i \cdot x$ for $x \in \mathbb{R}^n$. For $i \in \mathbb{Z}_{n+1}$, we set

$$\mathcal{F}_i = \{x \in \mathcal{F} \colon T_x u = \lambda_i\}.$$

Thus we have the disjoint decomposition

$$\mathcal{F} = \bigcup_{i \in \mathbb{Z}_{n+1}} \mathcal{F}_i.$$

Furthermore, we define $\mathcal{F}_i^* = \boldsymbol{U}(\mathcal{F}_i)$.

Of course $\boldsymbol{U} \colon \mathbb{R}^n \to \operatorname{graph}(u)$ is a bi-Lipschitz map. Thus in order to understand $\mathcal{F}$, $\mathcal{E}$, or $\mathcal{F}_i$, it suffices to study $\mathcal{F}^*$, $\mathcal{E}^*$, or $\mathcal{F}_i^*$ and how $\boldsymbol{U}^{-1}$ transforms them. In particular, the following is true.

Lemma 6 *For any Borel set $\Omega \subseteq \mathbb{R}^n$,*

$$\mathcal{H}^{n-1}(\mathcal{E}^* \cap (\Omega \times \mathbb{R})) = \sqrt{\frac{n+1}{2}} \mathcal{H}^{n-1}(\mathcal{E} \cap \Omega) = \frac{1}{2} |D \nabla u|(\Omega).$$

Proof We use the area formula [1, Theorem 2.91]. Hence we need to calculate the Jacobian of $\boldsymbol{U}$ restricted to the approximate tangent spaces of $\mathcal{E}$.

More precisely, since $\mathcal{E}$ is countably $\mathcal{H}^{n-1}$-rectifiable, there exists an approximate tangent space $T_x\mathcal{E}$ at $\mathcal{H}^{n-1}$-almost every $x \in \mathcal{E}$. Because $\boldsymbol{U}$ is Lipschitz continuous, the tangential derivative $d^{\mathcal{E}}\boldsymbol{U}(x)$ (i.e., the derivative of the restriction of $\boldsymbol{U}$ to $T_x\mathcal{E}$) exists at $\mathcal{H}^{n-1}$-almost every $x \in \mathcal{E}$ [1, Theorem 2.90]. We write L^* for the adjoint of a linear operator L. Then

$$J_{\mathcal{E}}\boldsymbol{U}(x) = \sqrt{\det\big((d^{\mathcal{E}}\boldsymbol{U}(x))^* \circ d^{\mathcal{E}}\boldsymbol{U}(x)\big)}$$

is the Jacobian of $\boldsymbol{U}$ at x with respect to $T_x\mathcal{E}$. The area formula implies that

$$\mathcal{H}^{n-1}(\boldsymbol{U}(\mathcal{E} \cap \Omega)) = \int_{\mathcal{E} \cap \Omega} J_{\mathcal{E}}\boldsymbol{U}(x)\, d\mathcal{H}^{n-1}.$$

Thus in order to prove the first identity, it suffices to show that

$$J_{\mathcal{E}}\boldsymbol{U}(x) = \sqrt{\frac{n+1}{2}}$$

for $\mathcal{H}^{n-1}$-almost every $x \in \mathcal{E}$.

To this end, consider $x \in \mathcal{E}$. Because of (4), we know that $T_x\mathcal{E} = (a_i - a_j)^{\perp}$ for some $i, j \in \mathbb{Z}_{n+1}$ with $i \neq j$ at $\mathcal{H}^{n-1}$-almost every such point. For $\xi \in (a_i - a_j)^{\perp}$, we know that

$$\frac{1}{\rho}(u(x+\rho\xi) - u(x)) = u_{x,\rho}(\xi) \to T_x u(\xi)$$

as $\rho \searrow 0$. The convergence is in fact uniform on compact subsets of $(a_i - a_j)^{\perp}$. Moreover, since $T_x u = \lambda_i \wedge \lambda_j$ or $T_x u = \lambda_j \vee \lambda_j$, its restriction to $(a_i - a_j)^{\perp}$ is linear with $T_x u(\xi) = a_i \cdot \xi$. Hence $d^{\mathcal{E}}u(x)$ exists, and so does $d^{\mathcal{E}}\boldsymbol{U}(x)$. We calculate

$$d^{\mathcal{E}}\boldsymbol{U}(x)\xi = \begin{pmatrix} \xi \\ a_i \cdot \xi \end{pmatrix}.$$

For simplicity, we assume that $i = n-1$ and $j = n$. The space $(a_i - a_j)^{\perp}$ is spanned by the vectors $a_0, \ldots, a_{n-2}$. Suppose that we choose an orthonormal basis $(\epsilon_0, \ldots, \epsilon_{n-2})$ of $T_x\mathcal{E}$. Let $L\colon T_x\mathcal{E} \to T_x\mathcal{E}$ denote the linear operator that maps ϵ_i to a_i for $i = 0, \ldots, n-2$. Then $d^{\mathcal{E}}\boldsymbol{U}(x) \circ L$ is represented by the matrix

$$M_1 = \begin{pmatrix} a_0 & \cdots & a_{n-2} \\ a_0 \cdot a_{n-1} & \cdots & a_{n-2} \cdot a_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & \cdots & a_{n-2} \\ -1 & \cdots & -1 \end{pmatrix}$$

with respect to the above basis. Hence

$$J_{\mathcal{E}}\boldsymbol{U}(x) = \sqrt{\frac{\det(M_1^T M_1)}{\det(L^* \circ L)}}.$$

We write I_k for the identity $k \times k$-matrix. Then

$$M_1^T M_1 = \begin{pmatrix} a_0 \cdot a_0 + 1 & \cdots & a_0 \cdot a_{n-2} + 1 \\ \vdots & \ddots & \vdots \\ a_{n-2} \cdot a_0 + 1 & \cdots & a_{n-2} \cdot a_{n-2} + 1 \end{pmatrix} = (n+1) I_{n-1}$$

and $\det(M_1^T M_1) = (n+1)^{n-1}$.

As L maps an $(n-1)$-cube of side length 1 to the parallelepiped spanned by $a_0, \ldots, a_{n-2}$, we know that $\sqrt{\det(L^* \circ L)}$ is the $(n-1)$-volume of the latter. Thus if M_2 is the $n \times (n-1)$-matrix with columns $a_0, \ldots, a_{n-2}$, then

$$\det(L^* \circ L) = \det(M_2^T M_2).$$

We further compute

$$M_2^T M_2 = \begin{pmatrix} a_0 \cdot a_0 & \cdots & a_0 \cdot a_{n-2} \\ \vdots & \ddots & \vdots \\ a_{n-2} \cdot a_0 & \cdots & a_{n-2} \cdot a_{n-2} \end{pmatrix} = \begin{pmatrix} n & -1 & \cdots & -1 \\ -1 & n & & \vdots \\ \vdots & & \ddots & -1 \\ -1 & \cdots & -1 & n \end{pmatrix}.$$

In order to calculate the determinant, we first subtract the first row of this matrix from each of the other rows. We obtain

$$\det(M_2^T M_2) = \det \begin{pmatrix} n & -1 & \cdots & \cdots & -1 \\ -(n+1) & n+1 & 0 & \cdots & 0 \\ -(n+1) & 0 & n+1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -(n+1) & 0 & \cdots & 0 & n+1 \end{pmatrix}$$
$$= (n+1)^{n-2} \det \begin{pmatrix} n & -1 & \cdots & \cdots & -1 \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

In the last matrix, we now add to the first row the sum of all the other rows. Thus

$$\det(M_2^T M_2) = (n+1)^{n-2} \det \begin{pmatrix} 2 & 0 & \cdots & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix} = 2(n+1)^{n-2}.$$

Hence

$$J_{\mathcal{E}} \boldsymbol{U}(x) = \sqrt{\frac{\det(M_1^T M_1)}{\det(M_2^T M_2)}} = \sqrt{\frac{n+1}{2}}.$$

In order to prove the second identity, we recall that $|a_i - a_j| = \sqrt{2n+2}$ for $i \neq j$. Hence $|D\nabla u|(\Omega) = \sqrt{2n+2}\mathcal{H}^{n-1}(\mathcal{E} \cap \Omega) = 2\mathcal{H}^{n-1}(\mathcal{E}^* \cap (\Omega \times \mathbb{R}))$ according to (5). □

5 Slicing the graph

We still assume that A consists of the corners of the regular n-simplex from Sect. 4 and we assume that $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n)$ is bounded and satisfies $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. In this section, we analyse the graph of u. In particular, we examine intersections of graph(u) with hyperplanes perpendicular to one of the vectors $\boldsymbol{v}_i$. We will see that almost all such intersections can be represented as the graphs of functions in $\mathrm{BV}^2_{\mathrm{loc}}(P)$, where

$$P = \left\{ y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0 \right\},$$

and with gradient taking one of n different values almost everywhere. That is, we have a function with properties similar to u, but with an $(n-1)$-dimensional domain. This observation will eventually make it possible to prove Theorem 3 with the help of an induction argument.

We use some tools from the author's previous paper [10] in this section. Given $i \in \mathbb{Z}_{n+1}$, let $\Phi_i : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ be the linear map with

$$\Phi_i(\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{v}_{i+1} \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{v}_{i+n+1} \cdot \boldsymbol{x} \end{pmatrix},$$

so that $\Phi_i(\boldsymbol{v}_{i+k})$ is the k-th standard basis vector in $\mathbb{R}^{n+1}$. For $t \in \mathbb{R}$, let

$$\Gamma_i(t) = \left\{ y \in \mathbb{R}^n : \begin{pmatrix} y \\ t \end{pmatrix} \in \Phi_i(\operatorname{graph}(u)) \right\}.$$

This corresponds to the intersection of $\operatorname{graph}(u)$ with a hyperplane orthogonal to $\boldsymbol{v}_i$ after rotation by Φ_i, or in other words, a slice of $\operatorname{graph}(u)$.

We further define the functions

$$\underline{g}_i(y) = \sup \left\{ t \in \mathbb{R} : u(tv_i + y_1 v_{i+1} + \cdots + y_n v_{i+n}) > \frac{t + y_1 + \cdots + y_n}{\sqrt{n+1}} \right\}$$

and

$$\overline{g}_i(y) = \inf \left\{ t \in \mathbb{R} : u(tv_i + y_1 v_{i+1} + \cdots + y_n v_{i+n}) < \frac{t + y_1 + \cdots + y_n}{\sqrt{n+1}} \right\}.$$

Note that for a fixed $y \in \mathbb{R}^n$, the set

$$\left\{ t \in \mathbb{R} : u(tv_i + y_1 v_{i+1} + \cdots + y_n v_{i+n}) = \frac{t + y_1 + \cdots + y_n}{\sqrt{n+1}} \right\}$$

corresponds to the intersection of $\operatorname{graph}(u)$ with a line parallel to v_i, so the functions $\underline{g}_i$ and $\overline{g}_i$ tell us something about the geometry of $\operatorname{graph}(u)$ as well.

The following properties of $\underline{g}_i$ and $\overline{g}_i$ have been proved elsewhere for $n = 2$ [10, Lemma 16]. The proof carries over to higher dimensions as well. We therefore do not repeat it here.

Lemma 7 *For any $i \in \mathbb{Z}_{n+1}$, the following statements hold true.*

(i) *The function $\underline{g}_i$ is lower semicontinuous and $\overline{g}_i$ is upper semicontinuous.*
(ii) *The identity $\underline{g}_i = \overline{g}_i$ holds almost everywhere in $\mathbb{R}^n$.*
(iii) *For any $y \in \mathbb{R}^n$, the inequality $\underline{g}_i(y) \le \overline{g}_i(y)$ holds true and*

$$\{y\} \times [\underline{g}_i(y), \overline{g}_i(y)] \subseteq \Phi_i(\operatorname{graph}(u)).$$

(iv) *Let $t \in \mathbb{R}$ and $y \in \mathbb{R}^n$. Then $y \in \Gamma_i(t)$ if, and only if, $\underline{g}_i(y) \le t \le \overline{g}_i(y)$.*
(v) *For all $y \in \mathbb{R}^n$ and all $\zeta \in (0, \infty)^n$, the inequality $\overline{g}_i(y + \zeta) \le \underline{g}_i(y)$ is satisfied; and if equality holds, then*

$$\underline{g}_i(y) = \underline{g}_i(y + s\zeta) = \overline{g}_i(y + s\zeta) = \overline{g}_i(y + \zeta)$$

for all $s \in (0, 1)$.
(vi) *For all $y \in \mathbb{R}^n$ and all $\zeta \in [0, \infty)^n$, the inequalities $\underline{g}_i(y) \ge \underline{g}_i(y + \zeta)$ and $\overline{g}_i(y) \ge \overline{g}_i(y + \zeta)$ are satisfied.*

Now consider the hyperplane $P \subseteq \mathbb{R}^n$ given by

$$P = \left\{ y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0 \right\}$$

and its unit normal vector

$$\sigma = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n.$$

Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$ and define

$$b_i = \sigma - \sqrt{n} e_i$$

for $i = 1, \dots, n$. Then

$$|b_i|^2 = n - 1$$

and

$$b_i \cdot b_j = -1$$

for $i \neq j$. Hence $b_1, \dots, b_n$ are the corners of a regular $(n-1)$-simplex in P centred at 0 with side length $\sqrt{2n}$. (Indeed the construction is similar to the standard $(n-1)$-simplex.) Thus they are the $(n-1)$-dimensional counterparts to $a_0, \dots, a_n$.

Given a function $f\colon P \times \mathbb{R} \to \mathbb{R}$, we write $\tilde{\nabla} f$ for its gradient with respect to the variable $p \in P$. We want to show the following.

Proposition 8 *Let $i \in \mathbb{Z}_{n+1}$. Then there exists a function $f_i\colon P \times \mathbb{R} \to \mathbb{R}$ such that for almost every $t \in \mathbb{R}$,*

- *The function $p \mapsto f_i(p,t)$ belongs to $\mathrm{BV}^2_{\mathrm{loc}}(P)$ and $\tilde{\nabla} f_i(p,t) \in \{b_1, \dots, b_n\}$ for $\mathcal{H}^{n-1}$-almost every $p \in P$; and*
- *Its graph is $\Gamma_i(t)$, that is, $\Gamma_i(t) = \{p + f_i(p,t)\sigma\colon p \in P\}$.*

Before we can prove this result, we need a few lemmas.

Lemma 9 *Let $i \in \mathbb{Z}_{n+1}$. Suppose that $t \in \mathbb{R}$ and $y_-, y_+ \in \Gamma_i(t)$. Then*

$$\big(y_- + [0,\infty)^n\big) \cap \big(y_+ - [0,\infty)^n\big) \subseteq \Gamma_i(t).$$

Proof We first prove that

$$\big(y_- + (0,\infty)^n\big) \cap \big(y_+ - (0,\infty)^n\big) \subseteq \Gamma_i(t).$$

Let

$$y \in \big(y_- + (0,\infty)^n\big) \cap \big(y_+ - (0,\infty)^n\big).$$

Define $\zeta_- = y - y_-$ and $\zeta_+ = y_+ - y$. Then $\zeta_-, \zeta_+ \in (0,\infty)^n$. According to Lemma 7, this means that

$$t \geq \underline{g}_i(y_-) \geq \overline{g}_i(y_- + \zeta_-) = \overline{g}_i(y) \geq \underline{g}_i(y) = \underline{g}_i(y_+ - \zeta_+) \geq \overline{g}_i(y_+) \geq t.$$

Hence $y \in \Gamma_i(t)$. By the semicontinuity of $\underline{g}_i$ and $\overline{g}_i$, we also conclude that

$$\underline{g}_i(y) \leq t \leq \overline{g}_i(y)$$

for all $y \in \big(y_- + [0,\infty)^n\big) \cap \big(y_+ - [0,\infty)^n\big)$. □

Lemma 10 *Let $i \in \mathbb{Z}_{n+1}$. Let $t \in \mathbb{R}$ and $p \in P$. Suppose that*

$$\{s \in \mathbb{R} \colon p + s\sigma \in \Gamma_i(t)\} = [s_-, s_+].$$

Then

$$\Gamma_i(t) \cap \left(p + s_-\sigma - (0,\infty)^n\right) = \emptyset$$

and

$$\Gamma_i(t) \cap \left(p + s_+\sigma + (0,\infty)^n\right) = \emptyset.$$

Proof Let $y \in p + s_-\sigma - (0,\infty)^n$. Choose $s < s_-$ such that $y \in p + s\sigma - (0,\infty)^n$ as well. Then Lemma 7 implies that

$$\underline{g}_i(y) \geq \overline{g}_i(p + s\sigma) \geq \underline{g}_i(p + s\sigma) \geq \overline{g}_i(p + s_-\sigma) \geq t. \tag{8}$$

Moreover, since $p + s\sigma \notin \Gamma_i(t)$, we know that $\underline{g}_i(p + s\sigma) \neq t$. Therefore, we do not have equality everywhere in (8). Hence $y \notin \Gamma_i(t)$. The proof of the second statement is similar. □

Lemma 11 *There exists a constant C such that the following holds true. Suppose that $v \colon \mathbb{R}^n \to \mathbb{R}$ is smooth and bounded with $a_j \cdot \nabla v > -1$ for all $j \in \mathbb{Z}_{n+1}$ and $\sup_{\mathbb{R}^n} |v| \leq M$. Let $i \in \mathbb{Z}_{n+1}$. Let $\phi \colon P \times \mathbb{R} \to \mathbb{R}$ be the unique function such that*

$$\begin{pmatrix} p + \phi(p,t)\sigma \\ t \end{pmatrix} \in \Phi_i(\mathrm{graph}(v))$$

for $p \in P$ and $t \in \mathbb{R}$. Then

$$|\tilde{\nabla}\phi(p,t)| \leq \sqrt{n} \tag{9}$$

for all $p \in P$ and $t \in \mathbb{R}$. Moreover, for any $R > 0$,

$$\int_{-R}^{R} \int_{P \cap B_R(0)} |\tilde{\nabla}^2 \phi| \, d\mathcal{H}^{n-1} \, dt \leq C \int_{B_{C(M+R)}(0)} |\nabla^2 v| \, dx.$$

Since the proof of this statement is lengthy, we postpone it to the next section. We now prove Proposition 8.

Proof of Proposition 8 Let $t \in \mathbb{R}$ and $p \in P$. Since u is bounded, the line

$$\left\{ t\boldsymbol{v}_i + \sum_{k=1}^{n} (p_k + s\sigma_k)\boldsymbol{v}_{i+k} \colon s \in \mathbb{R} \right\}$$

(which is not horizontal) must intersect $\mathrm{graph}(u)$. Hence there exists $s \in \mathbb{R}$ with $p + s\sigma \in \Gamma_i(t)$.

If there are $s_-, s_+ \in \mathbb{R}$ with $s_- < s_+$ such that $p + s_-\sigma \in \Gamma_i(t)$ and $p + s_+\sigma \in \Gamma_i(t)$, then Lemma 9 implies that $\Gamma_i(t)$ has non-empty interior, denoted by $\mathring{\Gamma}_i(t)$. Because of Lemma 7.(v), we know that $\underline{g}_i(y) = \overline{g}_i(y) = t$ for every $y \in \mathring{\Gamma}_i(t)$. Hence for $t_1 \neq t_2$, it follows that $\mathring{\Gamma}_i(t_1) \cap \mathring{\Gamma}_i(t_2) = \emptyset$. Therefore, there can only be countably many $t \in \mathbb{R}$ such that $\mathring{\Gamma}_i(t) \neq \emptyset$. For all other values, we see that $\Gamma_i(t)$ is a graph of a function over P. We denote this function by $f_i(\cdot, t)$.

We extend f_i arbitrarily to the remaining values of t.

If t is such that $\mathring{\Gamma}_i(t) = \emptyset$, then Lemma 10 shows that for every $y \in \Gamma_i(t)$, the set $\Gamma_i(t)$ is between the cones $y + (0,\infty)^n$ and $y - (0,\infty)^n$. Clearly there exists $L > 0$ such that

$$\{p + s\sigma \colon p \in P, s \in \mathbb{R} \text{ with } |s| > L|p|\} \subseteq (0,\infty)^n \cup (-(0,\infty)^n).$$

It follows that $f_i(\cdot, t)$ is Lipschitz continuous with Lipschitz constant at most L.

Next we employ an approximation argument in conjunction with Lemma 11. Using a standard mollifier, we can find a sequence of smooth, uniformly bounded functions $v_k \colon \mathbb{R}^n \to \mathbb{R}$ such that $v_k \to u$ locally uniformly as $k \to \infty$ and [1, Proposition 3.7]

$$|D\nabla u|(\Omega) = \lim_{k\to\infty} \int_\Omega |\nabla^2 v_k| \, dx$$

whenever $\Omega \subseteq \mathbb{R}^n$ is an open, bounded set with $|D\nabla u|(\partial\Omega) = 0$. It is then easy to modify v_k (e.g., replacing it with $(1 - 1/k)v_k$) such that in addition, it satisfies $a_j \cdot \nabla v_k > -1$ in $\mathbb{R}^n$ for every $j \in \mathbb{Z}_{n+1}$. Hence Lemma 11 applies to v_k.

From the above convergence, it follows that for any sequence of points $\boldsymbol{x}_k \in \operatorname{graph}(v_k)$, if $\boldsymbol{x}_k \to \boldsymbol{x}$ as $k \to \infty$, then $\boldsymbol{x} \in \operatorname{graph}(u)$. If we define ϕ_k as in Lemma 11, then for any fixed $t \in \mathbb{R}$, the functions $\phi_k(\cdot, t)$ are uniformly bounded in $C^{0,1}(P \cap B_R(0))$ for any $R > 0$. Hence there is a subsequence that converges locally uniformly. If t is such that $\Gamma_i(t)$ is the graph of $f_i(\cdot, t)$, then it is clear that the limit of any such subsequence must coincide with $f_i(\cdot, t)$. Hence in this case, we have the locally uniform convergence $\phi_k(\cdot, t) \to f_i(\cdot, t)$ as $k \to \infty$. The second inequality in Lemma 11 implies that

$$\limsup_{k\to\infty} \int_{-R}^{R} \int_{P\cap B_R(0)} |\tilde{\nabla}^2 \phi_k| d\mathcal{H}^{n-1} \, dt < \infty$$

for any $R > 0$. By Fatou's lemma,

$$\int_{-R}^{R} \liminf_{k\to\infty} \int_{P\cap B_R(0)} |\tilde{\nabla}^2 \phi_k| d\mathcal{H}^{n-1} \, dt < \infty.$$

Therefore, for almost every $t \in (-R, R)$, there exists a subsequence $(\phi_{k_\ell}(\cdot, t))_{\ell\in\mathbb{N}}$ converging to $f_i(\cdot, t)$ locally uniformly and such that

$$\limsup_{\ell\to\infty} \int_{P\cap B_R(0)} |\tilde{\nabla}^2 \phi_{k_\ell}| d\mathcal{H}^{n-1} < \infty.$$

We conclude that $f_i(\cdot, t) \in \mathrm{BV}^2(P \cap B_R(0))$ for almost all $t \in (-R, R)$. Applying this argument for a sequence $R_\ell \to \infty$, we then see that $f_i(\cdot, t) \in \mathrm{BV}^2_{\mathrm{loc}}(P)$ for almost all $t \in \mathbb{R}$.

We finally need to show that $\tilde{\nabla} f_i(p, t) \in \{b_1, \dots, b_n\}$ for almost every $t \in \mathbb{R}$ and $\mathcal{H}^{n-1}$-almost every $p \in P$.

Consider the function $w_i \colon \mathbb{R}^n \to \mathbb{R}$ with

$$w_i(x) = \frac{u(x) - a_i \cdot x}{\sqrt{n+1}}, \quad x \in \mathbb{R}^n.$$

Then for every $t \in \mathbb{R}$,

$$\begin{aligned}\Gamma_i(t) \times \{t\} &= \Phi_i\big(\{\boldsymbol{x} \in \operatorname{graph}(u) \colon \boldsymbol{x} \cdot \boldsymbol{\nu}_i = t\}\big) \\ &= \Phi_i\left(\left\{\begin{pmatrix} x \\ u(x)\end{pmatrix} : x \in \mathbb{R}^n \text{ with } w_i(x) = t\right\}\right).\end{aligned}$$

Note further that $\mathcal{F}_i$ coincides up to an $\mathcal{H}^n$-null set with $\{x \in \mathbb{R}^n \colon \nabla w_i(x) = 0\}$. Let $\mathcal{Z} \subset \mathbb{R}^n$ denote the set of all points where u is *not* differentiable. By Rademacher's theorem, this is an $\mathcal{H}^n$-null set. Hence the coarea formula gives

$$0 = \int_{\mathcal{F}_i \cup \mathcal{Z}} |\nabla w_i| \, dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}\big(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup \mathcal{Z})\big) \, dt.$$

In particular, for almost all $t \in \mathbb{R}$,

$$\mathcal{H}^{n-1}\big(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup \mathcal{Z})\big) = 0.$$

As the map $\boldsymbol{U}$ (defined in the introduction) is Lipschitz continuous, we conclude that $\boldsymbol{U}\left(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup \mathcal{Z})\right)$ is an $\mathcal{H}^{n-1}$-null set, too. Therefore, for $\mathcal{H}^{n-1}$-almost all $y \in \Gamma_i(t)$, the unique point $x \in \mathbb{R}^n$ with

$$\Phi_i(\boldsymbol{U}(x)) = \begin{pmatrix} y \\ t \end{pmatrix}$$

belongs to $\mathbb{R}^n \setminus \mathcal{Z}$ and satisfies $\nabla u(x) \in A \setminus \{a_i\}$.

To put it differently, for almost every $t \in \mathbb{R}$, the following holds true: for $\mathcal{H}^{n-1}$-almost every $p \in P$ the derivative of u exists at the point

$$\Theta(p,t) = t v_i + \sum_{k=1}^{n} (p_k + f_i(p,t)\sigma_k) v_{i+k}$$

and belongs to $A \setminus \{a_i\}$. Furthermore, we know that $f_i(\cdot, t)$ is differentiable at $\mathcal{H}^{n-1}$-almost every p by Rademacher's theorem. At a point $p \in P$ where both statements hold true, we can differentiate the equation

$$u(\Theta(p,t)) = \frac{t + \sqrt{n} f_i(p,t)}{\sqrt{n+1}}.$$

(The right-hand side is the $(n+1)$-st component of

$$t\boldsymbol{v}_i + \sum_{k=1}^{n} (p_k + f_i(p,t)\sigma_k)\boldsymbol{v}_{i+k} = \Phi_i^{-1}\begin{pmatrix} p + f_i(p,t)\sigma \\ t \end{pmatrix}$$

because $p \in P$ and by the definition of σ.) For any $\varpi \in P$, we thus obtain

$$-\left(\sum_{k=1}^{n} \varpi_k a_{i+k} + \varpi \cdot \tilde{\nabla} f_i(p,t) \sum_{k=1}^{n} \sigma_k a_{i+k}\right) \cdot \nabla u(\Theta(p,t)) = \sqrt{n}\, \varpi \cdot \tilde{\nabla} f_i(p,t).$$

If $\nabla u(\Theta(p,t)) = a_j$ for some $j \neq i$, then this simplifies to

$$-(n+1)\varpi_{j-i} - \frac{1}{\sqrt{n}} \varpi \cdot \tilde{\nabla} f_i(p,t) = \sqrt{n} \varpi \cdot \tilde{\nabla} f_i(p,t).$$

Hence

$$\varpi \cdot \tilde{\nabla} f_i(p,t) = -\sqrt{n}\, \varpi_{j-i} = b_{j-i} \cdot \varpi.$$

We therefore conclude that $\tilde{\nabla} f_i(p,t) = b_{j-i}$ at such a point. □

6 Proof of Lemma 11

In this section we give the postponed Proof of Lemma 11. To this end, we first need another lemma.

Lemma 12 *Let Λ denote the $(n \times n)$-matrix with columns*

$$\sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik} a_i, \quad k = 1, \dots, n.$$

Then

$$\det(\Lambda) = (-1)^n (n+1)^{\frac{n-1}{2}} \det \begin{pmatrix} \gamma_{01} & \cdots & \gamma_{0n} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \end{pmatrix}.$$

Proof Let M denote the $((n+1) \times (n+1))$-matrix with columns

$$\sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik} \boldsymbol{v}_i, \quad k = 1, \dots, n, \quad \text{and} \quad \sum_{i \in \mathbb{Z}_{n+1}} \boldsymbol{v}_i.$$

Then, since $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n+1})$ is a positively oriented orthonormal basis of $\mathbb{R}^{n+1}$, we conclude that

$$\det(M) = \det \begin{pmatrix} \gamma_{01} & \cdots & \gamma_{0n} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \end{pmatrix}.$$

On the other hand,

$$M = \frac{1}{\sqrt{n+1}} \begin{pmatrix} & & 0 \\ & -\Lambda & \vdots \\ & & 0 \\ m_1 & \cdots & m_n \; n+1 \end{pmatrix},$$

where $m_k = \sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik}$. Here we use the properties of $a_0, \dots, a_n$, including Eq. (7). Hence

$$\det(M) = (-1)^n (n+1)^{\frac{1-n}{2}} \det(\Lambda).$$

The claim follows immediately. □

Proof of Lemma 11 Suppose that $\boldsymbol{x} = \binom{x}{v(x)} \in \operatorname{graph}(v)$. Then

$$N = \frac{1}{\sqrt{1 + |\nabla v|^2}} \begin{pmatrix} -\nabla v(x) \\ 1 \end{pmatrix}$$

is a normal vector to $\operatorname{graph}(v)$ at $\boldsymbol{x}$. We compute

$$\Phi_i(N) = \frac{1}{\sqrt{(n+1)(1 + |\nabla v|^2)}} \begin{pmatrix} 1 + a_{i+1} \cdot \nabla v(x) \\ \vdots \\ 1 + a_{i+n+1} \cdot \nabla v(x) \end{pmatrix}.$$

Under the assumptions of Lemma 11, all the components of this vector are positive. Hence it is not parallel to $(0, \dots, 0, 1)^T$ and not perpendicular to $(1, \dots, 1, 0)^T$. The implicit function theorem therefore implies that

$$\Phi_i(\operatorname{graph}(v)) \cap \left\{ \boldsymbol{y} \in \mathbb{R}^{n+1} : y_{n+1} = t \right\}$$

is a smooth $(n-1)$-dimensional manifold for every $t \in \mathbb{R}$ and that the function ϕ from the statement of the lemma is smooth.

If we define $\Xi\colon P\times\mathbb{R}^2\to\mathbb{R}^{n+1}$ such that

$$\Xi(p,s,t)=t\boldsymbol{v}_i+\sum_{k=1}^{n}(p_k+s\sigma_k)\boldsymbol{v}_{i+k}$$

for $p\in P$ and $s,t\in\mathbb{R}$, then ϕ is characterised by the condition that

$$\Xi(p,\phi(p,t),t)\in\operatorname{graph}(v)$$

for all $t\in\mathbb{R}$ and $p\in P$. Hence

$$v\big(\Xi(p,\phi(p,t),t)\big)=\Xi_{n+1}(p,\phi(p,t),t). \tag{10}$$

We now differentiate this equation.

We compute

$$\frac{\partial\Xi}{\partial t}=v_i=-\frac{a_i}{\sqrt{n+1}},\quad \frac{\partial\Xi_{n+1}}{\partial t}=\frac{1}{\sqrt{n+1}}.$$

For $\varpi\in P$,

$$\varpi\cdot\tilde{\nabla}\Xi=-\frac{1}{\sqrt{n+1}}\sum_{k=1}^{n}\varpi_k a_{i+k},\quad \varpi\cdot\tilde{\nabla}\Xi_{n+1}=\frac{1}{\sqrt{n+1}}\sum_{k=1}^{n}\varpi_k=0.$$

Finally,

$$\frac{\partial\Xi}{\partial s}=\sum_{k=1}^{n}\sigma_k v_{i+k}=-\frac{1}{\sqrt{n^2+n}}\sum_{k=1}^{n}a_{i+k}=\frac{a_i}{\sqrt{n^2+n}},\quad \frac{\partial\Xi_{n+1}}{\partial s}=\sqrt{\frac{n}{n+1}}.$$

We define $\Theta(p,t)=\Xi(p,\phi(p,t),t)$. Differentiating (10), we now conclude that

$$\left(\frac{1}{\sqrt{n}}\frac{\partial\phi}{\partial t}(p,t)-1\right)a_i\cdot\nabla v(\Theta(p,t))=\sqrt{n}\frac{\partial\phi}{\partial t}(p,t)+1$$

and

$$\left(\frac{1}{\sqrt{n}}\varpi\cdot\tilde{\nabla}\phi(p,t)a_i-\sum_{k=1}^{n}\varpi_k a_{i+k}\right)\cdot\nabla v(\Theta(p,t))=\sqrt{n}\varpi\cdot\tilde{\nabla}\phi(p,t). \tag{11}$$

Hence

$$\frac{\partial\phi}{\partial t}(p,t)=\sqrt{n}\frac{a_i\cdot\nabla v(\Theta(p,t))+1}{a_i\cdot\nabla v(\,\Theta(p,t))-n} \tag{12}$$

and

$$\varpi\cdot\tilde{\nabla}\phi(p,t)=\sqrt{n}\frac{\sum_{k=1}^{n}\varpi_k a_{i+k}\cdot\nabla v(\Theta(p,t))}{a_i\cdot\nabla v(\Theta(p,t))-n}.$$

Fix $t\in\mathbb{R}$ and $p\in P$. Since $\nabla v(\Theta(p,t))$ is in the interior of the convex hull of the set $\{a_j\colon j\in\mathbb{Z}_{n+1}\}$, there exist $\tau_j\in(0,1)$ for $j\in\mathbb{Z}_{n+1}$ such that

$$\sum_{j\in\mathbb{Z}_{n+1}}\tau_j=1$$

and

$$\nabla v(\Theta(p,t))=\sum_{j\in\mathbb{Z}_{n+1}}\tau_j a_j.$$

Then

$$a_i \cdot \nabla v(\Theta(p,t)) - n = n\tau_i - \sum_{j\neq i} \tau_j - n = (n+1)(\tau_i - 1),$$

while

$$\sum_{k=1}^{n} \varpi_k a_{i+k} \cdot \nabla v(\Theta(p,t)) = \sum_{k=1}^{n} \varpi_k \left(n\tau_{i+k} - \sum_{j\neq i+k} \tau_j \right) = (n+1) \sum_{k=1}^{n} \varpi_k \tau_{i+k}.$$

We further note that

$$\tau_{i+1}^2 + \cdots + \tau_{i+n}^2 \leq (\tau_{i+1} + \cdots + \tau_{i+n})^2 = (1-\tau_i)^2.$$

The Cauchy–Schwarz inequality therefore implies that

$$\left| \sum_{k=1}^{n} \varpi_k a_{i+k} \cdot \nabla v(\Theta(p,t)) \right| \leq (n+1)(1-\tau_i)|\varpi|.$$

It follows that

$$|\varpi \cdot \tilde{\nabla}\phi(p,t)| \leq \sqrt{n}|\varpi|,$$

and inequality (9) is proved.

In order to prove the second statement of Lemma 11, we need to differentiate (11) again with respect to p. We write $\Lambda : M$ for the Frobenius inner product between two matrices Λ and M. We also drop the arguments (p,t) in the derivatives of ϕ and in Θ. Then for all $\varpi, \xi \in P$,

$$\begin{aligned} &\sqrt{\frac{n+1}{n}} (\xi \otimes \varpi) : \tilde{\nabla}^2 \phi \\ &= \frac{\left(\frac{\xi \cdot \tilde{\nabla}\phi}{\sqrt{n}} a_i - \sum_{k=1}^{n} \xi_k a_{i+k} \right) \otimes \left(\frac{\varpi \cdot \tilde{\nabla}\phi}{\sqrt{n}} a_i - \sum_{k=1}^{n} \varpi_k a_{i+k} \right)}{n - a_i \cdot \nabla v(\Theta)} : \nabla^2 v(\Theta). \end{aligned}$$

As we have already seen that $|\tilde{\nabla}\phi| \leq \sqrt{n}$, it follows that there is a constant $C_1 = C_1(n)$ such that

$$|\tilde{\nabla}^2 \phi| \leq \frac{C_1 |\nabla^2 v(\Theta)|}{n - a_i \cdot \nabla v(\Theta)}.$$

Choose an orthonormal basis $(\eta_1, \ldots, \eta_{n-1})$ of P. Next we examine the derivative $d\Theta$, and more specifically, its determinant.

Define $\gamma_{1k}, \ldots, \gamma_{nk}$ such that $\eta_k = (\gamma_{1k}, \ldots, \gamma_{nk})^T$. For $t \in \mathbb{R}$ and $p \in P$, we also define

$$\gamma_{n+1,k}(p,t) = -\frac{1}{\sqrt{n}} \eta_k \cdot \tilde{\nabla}\phi(p,t), \quad k = 1, \ldots, n-1,$$

and

$$\gamma_{n+1,n}(p,t) = 1 - \frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial t}(p,t).$$

Finally, we set $\gamma_{\ell n} = 0$ for $\ell = 1, \ldots, n$. We compute

$$\eta_k \cdot \tilde{\nabla}\Theta(p,t) = \frac{1}{\sqrt{n+1}} \left(\frac{1}{\sqrt{n}} \eta_k \cdot \tilde{\nabla}\phi(p,t) a_i - \sum_{\ell=1}^{n} \gamma_{\ell k} a_{i+\ell} \right)$$

and

$$\frac{\partial \Theta}{\partial t}(p,t) = \left(\frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial t}(p,t) - 1 \right) \frac{a_i}{\sqrt{n+1}}.$$

Hence we can represent $d\Theta$ by the matrix with columns

$$-\sum_{\ell=1}^{n+1} \frac{\gamma_{\ell k} a_{i+\ell}}{\sqrt{n+1}}, \quad k = 1, \dots, n,$$

with respect to the basis of $P \times \mathbb{R}$ induced by $\eta_1 \dots, \eta_{n-1}$. Lemma 12 now tells us that

$$\begin{aligned}
\det(d\Theta) &= \pm \frac{1}{\sqrt{n+1}} \det \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n+1,1} & \cdots & \gamma_{n+1,n} & 1 \end{pmatrix} \\
&= \pm \frac{1}{\sqrt{n+1}} \det \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1,n-1} & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{n,n-1} & 0 & 1 \\ \gamma_{n+1,1} & \cdots & \gamma_{n+1,n-1} & \gamma_{n+1,n} & 1 \end{pmatrix} \\
&= \mp \sqrt{\frac{n}{n+1}} \gamma_{n+1,n} \det \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1,n-1} & \sigma_1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{n,n-1} & \sigma_n \end{pmatrix}.
\end{aligned}$$

As $(\eta_1, \dots, \eta_{n-1}, \sigma)$ form an orthonormal basis of $\mathbb{R}^n$, we find that

$$|\det(d\Theta)| = \sqrt{\frac{n}{n+1}} |\gamma_{n+1,n}| = \frac{1}{\sqrt{n+1}} \left| \sqrt{n} - \frac{\partial \phi}{\partial t} \right|.$$

Recalling (12), we now obtain

$$|\det(d\Theta)| = \frac{\sqrt{n^2+n}}{n - a_i \cdot \nabla v(\Theta)}.$$

We also note that the map Θ is injective. Given $R > 0$, we therefore compute

$$\begin{aligned}
&\int_{-R}^{R} \int_{P \cap B_R(0)} |\tilde{\nabla}^2 \phi| \, d\mathcal{H}^{n-1} \, dt \\
&\leq C_1 \int_{-R}^{R} \int_{P \cap B_R(0)} \frac{|\nabla^2 v(\Theta)|}{n - a_i \cdot \nabla v(\Theta)} \, d\mathcal{H}^{n-1} \, dt \\
&= \frac{C_1}{\sqrt{n^2+n}} \int_{-R}^{R} \int_{P \cap B_R(0)} |\nabla^2 v(\Theta)| |\det(d\Theta)| \, d\mathcal{H}^{n-1} \, dt \\
&= \frac{C_1}{\sqrt{n^2+n}} \int_{\Theta((P \cap B_R(0)) \times (-R,R))} |\nabla^2 v| \, dx.
\end{aligned} \tag{13}$$

It remains to examine the set $\Theta((P \cap B_R(0)) \times (-R, R))$. Recall that we have the assumption $\sup_{\mathbb{R}^n} |v| \leq M$ in Lemma 11. Thus (10) implies that

$$|\Xi_{n+1}(p, \phi(p,t), t)| \leq M.$$

Since

$$\Xi_{n+1}(p, \phi(p,t), t) = \frac{t + \sqrt{n}\phi(p,t)}{\sqrt{n+1}},$$

this means that

$$|\phi(p,t)| \le M\sqrt{\frac{n+1}{n}} + \frac{R}{\sqrt{n}}$$

when $t \in (-R, R)$. Hence there exists a constant $C_2 = C_2(n)$ such that

$$|\Theta(p,t)| \le C_2(M+R)$$

for all $p \in P \cap B_R(0)$ and all $t \in (-R, R)$. Thus (13) implies the second inequality of Lemma 11. □

7 Proof of Theorem 3

In this section we combine the previous results to prove the second main theorem. We first consider a function $u \in \mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ such that graph(u) is close to the graph of $\lambda_i \wedge \lambda_j$ or $\lambda_i \vee \lambda_j$ in a cube in $\mathbb{R}^{n+1}$ with edges parallel to $\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_{n+1}$. We will give a condition which implies that such a function actually coincides with $\lambda_i \wedge \lambda_j$ or $\lambda_i \vee \lambda_j$ up to a constant in part of the domain.

For $i, j \in \mathbb{Z}_{n+1}$ with $i \ne j$ and for $r, R > 0$, we define

$$Q_{ij}(r,R) = \left\{ \sum_{k \in \mathbb{Z}_{n+1}} c_k \boldsymbol{\nu}_k : c_i, c_j \in (-r, r) \text{ and } c_k \in (-R, R) \text{ for } k \notin \{i, j\} \right\}.$$

Again we consider the map $\boldsymbol{U} : \mathbb{R}^n \to \mathbb{R}^{n+1}$ with $\boldsymbol{U}(x) = \binom{x}{u(x)}$ for $x \in \mathbb{R}^n$. Recall that $\mathcal{E}^* = \boldsymbol{U}(\mathcal{E})$. The following is the key statement for the Proof of Theorem 3.

Proposition 13 *Let $n \in \mathbb{N}$. For any $\delta > 0$ there exist $\epsilon > 0$ with the following properties. Let $i, j \in \mathbb{Z}_{n+1}$ with $i \ne j$. Suppose that $|u(0)| \le \epsilon$ and either*

$$|u - \lambda_i \wedge \lambda_j| \le \epsilon \quad \text{in } \boldsymbol{U}^{-1}\big(Q_{ij}(1,1)\big) \tag{14}$$

or

$$|u - \lambda_i \vee \lambda_j| \le \epsilon \quad \text{in } \boldsymbol{U}^{-1}\big(Q_{ij}(1,1)\big). \tag{15}$$

Then

$$\mathcal{H}^{n-1}\big(\mathcal{E}^* \cap Q_{ij}(\tfrac{1}{4}, 1)\big) \ge 2^{n-1}(1-\delta). \tag{16}$$

If, in addition,

$$\mathcal{H}^{n-1}\big(\mathcal{E}^* \cap Q_{ij}(1,1)\big) \le 2^{n-1}(1+\epsilon), \tag{17}$$

then there exist $\alpha, \beta \in \mathbb{R}$ such that

$$u = (\lambda_i + \alpha) \wedge (\lambda_j + \beta) \quad \text{in } \boldsymbol{U}^{-1}\big(Q_{ij}(\tfrac{1}{2}, \tfrac{1}{2})\big) \tag{18}$$

or

$$u = (\lambda_i + \alpha) \vee (\lambda_j + \beta) \quad \text{in } \boldsymbol{U}^{-1}\big(Q_{ij}(\tfrac{1}{2}, \tfrac{1}{2})\big). \tag{19}$$

Before we can prove Proposition 13, we need a few more lemmas. First we need some more information on the functions f_i from Proposition 8. Recall that $f_i(\cdot, t) \in \mathrm{BV}^2_{\mathrm{loc}}(P)$ for almost all $t \in \mathbb{R}$.

Given $i \in \mathbb{Z}_{n+1}$ and given $t \in \mathbb{R}$ such that $f_i(\cdot, t) \in \mathrm{BV}^2_{\mathrm{loc}}(P)$, let $\mathcal{D}'_i(t) \subseteq P$ denote the approximate jump set of $\tilde{\nabla} f_i(\cdot, t)$. Thus this set is defined analogously to $\mathcal{E}'$, but for the function $f_i(\cdot, t)$ instead of u. Furthermore, we set

$$\mathcal{D}^\dagger_i(t) = \left\{ p + f_i(p,t)\sigma : p \in \mathcal{D}'_i(t) \right\},$$

in analogy to $\mathcal{E}^\dagger$.

Lemma 14 *Let $i \in \mathbb{Z}_{n+1}$. For almost any $t \in \mathbb{R}$,*

$$\mathcal{D}^\dagger_i(t) \times \{t\} \subseteq \Phi_i(\mathcal{E}^* \cup \mathcal{N}^*).$$

Hence for any $t_1, t_2 \in \mathbb{R}$ and any Borel set $\Omega \subseteq \mathbb{R}^n$,

$$\int_{t_1}^{t_2} \mathcal{H}^{n-2}(\mathcal{D}^\dagger_i(t) \cap \Omega)\, dt \le \mathcal{H}^{n-1}\big(\mathcal{E}^* \cap \Phi_i^{-1}(\Omega \times (t_1, t_2))\big).$$

Proof Let $p \in P$ and $t \in \mathbb{R}$. Set

$$\boldsymbol{x} = \Phi_i^{-1} \begin{pmatrix} p + f_i(p,t)\sigma \\ t \end{pmatrix}.$$

If $\boldsymbol{x} \in \mathcal{F}^*$, then Proposition 5 implies that $\mathrm{graph}(u)$ coincides with a hyperplane in a neighbourhood of $\boldsymbol{x}$. If that hyperplane is perpendicular to $\boldsymbol{v}_i$, then $p + f_i(p,t)\sigma \in \mathring{\Gamma}_i(t)$ and t belongs to the null set identified in Proposition 8. Otherwise, the function $f_i(\cdot, t)$ is affine near p, and hence $\Phi_i(\boldsymbol{x})$ cannot belong to $\mathcal{D}^\dagger_i(t) \times \{t\}$. This implies the first claim.

The second claim is now a consequence of the coarea formula [1, Theorem 2.93], applied to the functions $\boldsymbol{y} \mapsto y_{n+1}$ and to the countably $\mathcal{H}^{n-1}$-rectifiable set $\Phi_i(E^* \cup \mathcal{N}^*) \cap (\Omega \times [t_1, t_2])$. □

Lemma 15 *Let $k \in \{1, \dots, n\}$. Suppose that $\underline{s}, \overline{s} \in \mathbb{R}$ with $\underline{s} < \overline{s}$. For $z \in \mathbb{R}^{n-1}$, define $\ell_z(s) = (z_1, \dots, z_{k-1}, s, z_k, \dots, z_{n-1})^T$ for $s \in [\underline{s}, \overline{s}]$, and $L_z = \left\{ \ell_z(s) : \underline{s} \le s \le \overline{s} \right\}$. Fix $i \in \mathbb{Z}_{n+1}$. Then for $\mathcal{H}^{n-1}$-almost every $z \in \mathbb{R}^{n-1}$, either*

$$\underline{g}_i(y) = \overline{g}_i(y) = \underline{g}_i(y') = \overline{g}_i(y')$$

for all $y, y' \in L_z$, or there exist $\boldsymbol{y} \in L_z \times \mathbb{R}$ such that

$$\overline{g}_i(\ell_z(\overline{s})) \le y_{n+1} \le \underline{g}_i(\ell_z(\underline{s}))$$

and $\boldsymbol{y} \in \Phi_i(\mathcal{E}^*)$.

Proof Consider the projection $\Pi \colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ given by $\Pi(\boldsymbol{y}) = y$ for $\boldsymbol{y} \in \mathbb{R}^{n+1}$. Set $\Psi_i = \Pi \circ \Phi_i$. Then for $j \in \mathbb{Z}_{n+1}$ with $j \ne i$ and for $\boldsymbol{x} \in \mathcal{F}^*_j$, it is clear that $J_{\mathcal{F}^*}\Psi_i(\boldsymbol{x}) = 0$. (Indeed, the vectors $\boldsymbol{v}_{j+1}, \dots, \boldsymbol{v}_{j+n+1}$ provide an orthonormal basis of $T_{\boldsymbol{x}}\mathcal{F}^*$, and we compute $d\Psi_i(\boldsymbol{v}_i) = 0$.) Hence the area formula gives $\mathcal{H}^n(\Psi_i(\mathcal{F}^*_j)) = 0$. Using the coarea formula, we conclude that for $\mathcal{H}^{n-1}$-almost every $z \in \mathbb{R}^{n-1}$,

$$\mathcal{H}^1(L_z \cap \Psi_i(\mathcal{F}^*_j)) = 0 \tag{20}$$

for all $j \ne i$. Furthermore, since $\mathcal{E}^*$ is an $\mathcal{H}^{n-1}$-rectifiable set and $\mathcal{H}^{n-1}(\mathcal{N}^*) = 0$, we also know that for $\mathcal{H}^{n-1}$-almost every $z \in \mathbb{R}^{n-1}$,

$$\mathcal{H}^1(L_z \cap \Psi_i(\mathcal{E}^*)) = 0 \tag{21}$$

and

$$L_z \cap \Psi_i(\mathcal{N}^*) = \emptyset. \tag{22}$$

Consider a point $z \in \mathbb{R}^{n-1}$ such that (20), (21), and (22) hold true. Recall that by Lemma 7, a point $\boldsymbol{y} \in \mathbb{R}^{n+1}$ belongs to $\Phi_i(\mathrm{graph}(u))$ if, and only if, $\underline{g}_i(y) \le y_{n+1} \le \overline{g}_i(y)$. Also recall that

$$\mathrm{graph}(u) = \mathcal{E}^* \cup \mathcal{N}^* \cup \bigcup_{j \in \mathbb{Z}_{n+1}} \mathcal{F}_j^*.$$

From (20)–(22) we therefore infer that for $\mathcal{H}^1$-almost all $y \in L_z$,

$$\begin{pmatrix} y \\ t \end{pmatrix} \in \Phi_i(\mathcal{F}_i^*) \quad \text{for all } t \in [\underline{g}_i(y), \overline{g}_i(y)]. \tag{23}$$

Consider $\boldsymbol{y} \in \Phi_i(\mathcal{F}_i^*)$ with $y \in L_z$. Then, setting $\boldsymbol{x} = \Phi_i^{-1}(\boldsymbol{y})$, we have the locally uniform convergence $u_{x,\rho} \to \lambda_i$ as $\rho \searrow 0$. Hence for any compact set $K \subseteq \mathbb{R}^{n+1}$ and any $\epsilon > 0$ there exists $\rho_0 > 0$ such that

$$\frac{1}{\rho}(\mathrm{graph}(u) - \boldsymbol{x}) \cap K \subseteq \left\{\tilde{\boldsymbol{x}} \in \mathbb{R}^{n+1} \colon \operatorname{dist}(\tilde{\boldsymbol{x}}, \mathrm{graph}(\lambda_i)) < \epsilon/2\right\}$$

for all $\rho \in (0, \rho_0]$. Equivalently,

$$\Phi_i(\mathrm{graph}(u) - \boldsymbol{x}) \cap \rho \Phi_i(K) \subseteq \mathbb{R}^n \times (-\rho\epsilon/2, \rho\epsilon/2).$$

Recall that $e_1, \dots, e_n$ are the standard basis vectors in $\mathbb{R}^n$. Also recall that for any $\tilde{y} \in \mathbb{R}^n$, the points $(\tilde{y}, \underline{g}_i(\tilde{y}))$ and $(\tilde{y}, \overline{g}_i(\tilde{y}))$ belong to $\Phi_i(\mathrm{graph}(u))$. It follows that there exists $r_0 > 0$ such that for all $r \in (0, r_0]$,

$$\left|\underline{g}_i(y \pm r e_k) - \underline{g}_i(y)\right| \le r\epsilon \quad \text{and} \quad \left|\overline{g}_i(y \pm r e_k) - \overline{g}_i(y)\right| \le r\epsilon$$

and $|\underline{g}_i(y) - \overline{g}_i(y)| \le r\epsilon$. Thus

$$\frac{\partial}{\partial y_k} \underline{g}_i(y) = 0 \quad \text{and} \quad \frac{\partial}{\partial y_k} \overline{g}_i(y) = 0$$

and $\underline{g}_i(y) = \overline{g}_i(y)$. Note that this is true for $\mathcal{H}^1$-almost all $y \in L_z$. We now claim that

$$\underline{g}_i(\ell_z(\underline{s})) \ge \overline{g}_i(y) \ge \underline{g}_i(y) \ge \overline{g}_i(\ell_z(\overline{s})) \tag{24}$$

for all $y \in L_z \setminus \{\ell_z(\underline{s}), \ell_z(\overline{s})\}$. Indeed, given $y = \ell_z(s)$ with $s \in (\underline{s}, \overline{s})$, we may choose $s_- \in (\underline{s}, s)$ and $s_+ \in (s, \overline{s})$ such that $\underline{g}_i(\ell_z(s_-)) = \overline{g}_i(\ell_z(s_-))$ and $\underline{g}_i(\ell_z(s_+)) = \overline{g}_i(\ell_z(s_+))$. Then (24) follows from Lemma 7.(vi).

If (23) holds for *all* $y \in L_z$, then we immediately conclude that $\underline{g}_i$ and $\overline{g}_i$ are constant and coincide on L_z, i.e., we have the first alternative from the statement of the lemma. If there exists $y \in L_z$ such that (23) does *not* hold true, then by the above observations, we know that

$$\begin{pmatrix} y \\ t \end{pmatrix} \notin \Phi_i(\mathcal{F}_i^*)$$

holds in fact for all $t \in [\underline{g}_i(y), \overline{g}_i(y)]$. (Otherwise, we would conclude that $\underline{g}_i(y) = \overline{g}_i(y)$, which would give an immediate contradiction.) Moreover, because (23) still holds true almost everywhere on L_z, there exists a sequence $(\tilde{y}_m)_{m \in \mathbb{N}}$ in L_z such that $y = \lim_{m \to \infty} \tilde{y}_m$ and such that (23) holds for every $\tilde{y}_m$. Hence $\underline{g}_i(\tilde{y}_m) = \overline{g}_i(\tilde{y}_m)$. We set $\tilde{t}_m = \underline{g}_i(\tilde{y}_m)$. Extracting

a subsequence if necessary, we may assume that $y_{n+1} = \lim_{m\to\infty} \tilde{t}_m$ exists. Set $\boldsymbol{y} = \binom{y}{y_{n+1}}$. Then $\Phi_i^{-1}(\boldsymbol{y})$ belongs to the boundary of $\mathcal{F}_i^*$ relative to graph(u).

Proposition 5 implies that $\mathcal{F}_i^*$ is an open set relative to graph(u), and its relative boundary is contained in $\mathcal{E}^* \cup \mathcal{N}^*$. Because of (22), it follows that $\Phi_i^{-1}(\boldsymbol{y}) \in \mathcal{E}^*$. Moreover, (24) implies that

$$\overline{g}_i(\ell_z(\overline{s})) \le y_{n+1} \le \underline{g}_i(\ell_z(\underline{s})).$$

Thus $\boldsymbol{y}$ has the properties from the second alternative in the statement. □

Lemma 16 *Let $i \in \mathbb{Z}_{n+1}$. Suppose that $G \subseteq \mathbb{R}^n$ is a connected set such that $G \cap \Gamma_i(t) = \emptyset$ for all $t \in (-1, 1)$. Then either $\underline{g}_i(y) \ge 1$ for all $y \in G$ or $\overline{g}_i(y) \le -1$ for all $y \in G$.*

Proof Assume that there exists $y_0 \in G$ such that $\underline{g}_i(y_0) < 1$. Since $G \cap \Gamma_i(t) = \emptyset$ for all $t \in (-1, 1)$, this implies that

$$-1 \ge \overline{g}_i(y_0) \ge \underline{g}_i(y_0)$$

by Lemma 7.(iv).

Given $t \in (-1, 1)$, define

$$H_t = \left\{ y \in G \colon \overline{g}_i(y) \ge t \right\}.$$

Because $\overline{g}_i$ is upper semicontinuous by Lemma 7, this is a closed set relative to G. Moreover, if $y \in H_t$, it follows that

$$\overline{g}_i(y) \ge \underline{g}_i(y) \ge 1,$$

because $G \cap \Gamma_i(t') = \emptyset$ for all $t' \in (-1, 1)$. By the lower semicontinuity of $\underline{g}_i$, this means that there exists $\rho > 0$ such that $\overline{g}_i \ge \underline{g}_i \ge t$ in $B_\rho(y)$. Hence H_t is also open relative to G. Since G is connected and $y_0 \notin H_t$, it follows that $H_t = \emptyset$. This is true for all $t \in (-1, 1)$, so $\overline{g}_i(y) \le -1$ for all $y \in G$. □

We now have everything in place for the Proof of Proposition 13.

Proof of Proposition 13 We use induction over n. The statement is clear for $n = 1$. We now assume that $n \ge 2$ and the statement holds true for $n - 1$.

For simplicity, we assume that $i = 1$ and $j = 2$. We also assume that (14) holds true; the proof is similar under the assumption (15).

Let

$$\Lambda = \left((-\infty, 0] \times \{0\} \times \mathbb{R}^{n-2}\right) \cup \left(\{0\} \times (-\infty, 0] \times \mathbb{R}^{n-2}\right).$$

The graph of $\lambda_1 \wedge \lambda_2$ is the union of two half-hyperplanes meeting at a right angle. In fact, it is easy to see that

$$\Phi_0(\text{graph}(\lambda_1 \wedge \lambda_2)) = \Lambda \times \mathbb{R}.$$

Let

$$\epsilon' = \epsilon \sqrt{\frac{n}{n+1}}.$$

Under the assumptions of the proposition, the set $\Phi_0(\text{graph}(u)) \cap (-1, 1)^{n+1}$ is between $(\Lambda - \epsilon'\sigma) \times \mathbb{R}$ and $(\Lambda + \epsilon'\sigma) \times \mathbb{R}$, i.e.,

$$\Phi_0(\text{graph}(u)) \cap (-1, 1)^{n+1} \subseteq \bigcup_{-\epsilon' \le s \le \epsilon'} (\Lambda + s\sigma) \times \mathbb{R}.$$

Set $s_0 = \sqrt{\frac{n}{n+1}}u(0)$. Then $|s_0| \le \epsilon'$ by the assumption that $|u(0)| \le \epsilon$. Moreover, we compute

$$\Phi_0 \begin{pmatrix} 0 \\ u(0) \end{pmatrix} = \frac{u(0)}{\sqrt{n+1}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = s_0 \begin{pmatrix} \sigma \\ \frac{1}{\sqrt{n}} \end{pmatrix}.$$

Assuming that $\epsilon < \sqrt{n+1}$, we infer that $\overline{g}_0(s_0\sigma) > -1$ and $\underline{g}_0(s_0\sigma) < 1$. Using Lemmas 7.(v) and 16, we conclude that

$$\underline{g}_0(y) \ge 1 \quad \text{for } y \in (-1,1)^n \cap \bigcup_{s<-\epsilon'} (\Lambda + s\sigma)$$

and

$$\overline{g}_0(y) \le -1 \quad \text{for } y \in (-1,1)^n \cap \bigcup_{s>\epsilon'} (\Lambda + s\sigma).$$

Now consider the function $f_0 \colon P \times \mathbb{R} \to \mathbb{R}$ from Proposition 8. For almost every $t \in (-1,1)$, the graph of $f_0(\cdot,t)$, which is given by $\Gamma_0(t)$, is between $\Lambda - \epsilon'\sigma$ and $\Lambda + \epsilon'\sigma$ in the hypercube $(-1,1)^n$.

Define $\mu_1, \mu_2 \colon P \to \mathbb{R}$ by $\mu_1(p) = b_1 \cdot p$ and $\mu_2(p) = b_2 \cdot p$ for $p \in P$ (where b_1 and b_2 are the vectors defined on page 12). Let $F_t \colon P \to \mathbb{R}^n$ be the map with $F_t(p) = p + f_0(p,t)\sigma$ for $p \in P$. Then the preceding statements amount to the inequality

$$|f_0(\cdot,t) - \mu_1 \wedge \mu_2| \le \epsilon' \quad \text{in } F_t^{-1}\big((-1,1)^n\big).$$

Moreover, the condition $|f_0(0,t)| \le \epsilon'$ is clearly satisfied. Hence we may apply the induction hypothesis to the function $f_0(\cdot,t)$. We thereby obtain the inequality

$$\mathcal{H}^{n-2}\big(\mathcal{D}_0^\dagger(t) \cap \big((-\tfrac{1}{4},\tfrac{1}{4})^2 \times (-1,1)^{n-2}\big)\big) \ge 2^{n-2}(1-\delta) \tag{25}$$

for almost all $t \in (-1,1)$, provided that ϵ is sufficiently small. Using Lemma 14, we therefore obtain inequality (16). This proves the first statement of Proposition 13.

In order to prove the second statement, assume now that (17) holds true. Then

$$\int_{-1}^{1} \mathcal{H}^{n-2}\big(\mathcal{D}_0^\dagger(t) \cap (-1,1)^n\big)\, dt \le 2^{n-1}(1+\epsilon).$$

Recall that we also have inequality (25), and we may now assume that δ is arbitrarily small. Hence there exist $t_- \in (-1,-\frac{1}{2})$ and $t_+ \in (\frac{1}{2},1)$ such that

$$\mathcal{H}^{n-2}\big(\mathcal{D}_0^*(t_\pm) \cap (-1,1)^n\big) \le 2^{n-2}(1+3\delta+4\epsilon).$$

By the induction hypothesis, if δ and ϵ are sufficiently small, then

$$f_0(\cdot,t_\pm) = (\mu_1 + \alpha_\pm) \wedge (\mu_2 + \beta_\pm) \quad \text{in } F_t^{-1}\left(\left(-\tfrac{1}{2},\tfrac{1}{2}\right)^n\right)$$

for certain numbers $\alpha_-, \alpha_+, \beta_-, \beta_+ \in \mathbb{R}$. Therefore, there exist $y_-, y_+ \in \mathbb{R}^2 \times \{0\}^{n-2}$ such that

$$\Gamma_0(t_\pm) \cap \left(-\tfrac{1}{2},\tfrac{1}{2}\right)^n = (y_\pm + \Lambda) \cap \left(-\tfrac{1}{2},\tfrac{1}{2}\right)^n.$$

Clearly, by the above observations on $\Phi_0(\mathrm{graph}(u))$, this implies that $y_\pm \in B_{\epsilon'}(0)$. We assume that $\epsilon' \le \frac{1}{4}$.

If $y_- = y_+$, then

$$\Gamma_0(t) \cap \left(-\tfrac{1}{2}, \tfrac{1}{2}\right)^n = (y_+ + \Lambda) \cap \left(-\tfrac{1}{2}, \tfrac{1}{2}\right)^n$$

for every $t \in (t_-, t_+)$ as well (because any other point $y \in (-\frac{1}{2}, \frac{1}{2})^n \setminus (y_+ + \Lambda)$ satisfies either $\underline{g}_0(y) \geq t_+$ or $\overline{g}_0(y) \leq t_-$ by Lemma 7). In this case, we conclude that (18) holds true. Thus it now suffices to show that $y_- = y_+$.

We argue by contradiction here. Suppose that $y_- \neq y_+$. We assume that in fact the first components y_{1-} and y_{1+} are different. The arguments are similar if $y_{2-} \neq y_{2+}$.

If $y_{1-} \neq y_{1+}$, then for any $z \in (-\frac{1}{2}, -\frac{1}{4}) \times (-\frac{1}{2}, \frac{1}{2})^{n-2}$, it follows that

$$\underline{g}_0 \begin{pmatrix} y_{1-} \\ z \end{pmatrix} \leq t_- \leq \overline{g}_0 \begin{pmatrix} y_{1-} \\ z \end{pmatrix}$$

and

$$\underline{g}_0 \begin{pmatrix} y_{1+} \\ z \end{pmatrix} \leq t_+ \leq \overline{g}_0 \begin{pmatrix} y_{1+} \\ z \end{pmatrix}.$$

Since $t_- < t_+$, it is therefore *not* true that $\underline{g}_0$ and $\overline{g}_0$ are constant with $\underline{g}_0 = \overline{g}_0$ on $[y_{1+}, y_{1-}] \times \{z\}$. Lemma 15 now implies that for $\mathcal{H}^{n-1}$-almost every $z \in (-\frac{1}{2}, -\frac{1}{4}) \times (-\frac{1}{2}, \frac{1}{2})^{n-2}$, the set $[y_{1+}, y_{1-}] \times \{z\} \times [t_-, t_+]$ intersects $\Phi_0(\mathcal{E}^*)$. It follows that

$$\mathcal{H}^{n-1}\left(\Phi_0(\mathcal{E}^*) \cap \left((-1, 1) \times \left(-\tfrac{1}{2}, -\tfrac{1}{4}\right) \times (-1, 1)^{n-1}\right)\right) \geq \frac{1}{4}.$$

Furthermore, because of (16), we obtain the estimate

$$\mathcal{H}^{n-1}\left(\mathcal{E}^* \cap Q_{12}(1, 1)\right) \geq 2^{n-1}(1 - \delta) + \frac{1}{4}.$$

If $\delta + \epsilon < 2^{-n-1}$, then this contradicts the hypothesis. □

Finally we can prove the second main result with the help of Propositions 5 and 13.

Proof of Theorem 3 Suppose that $A \subseteq \mathbb{R}^n$ is affinely independent. Then A contains at most $n + 1$ elements. If there are fewer, then we can add additional elements to A such that it remains affinely independent. Thus we may assume without loss of generality that the size of A is exactly $n + 1$.

Now suppose that $A = \{\tilde{a}_0, \ldots, \tilde{a}_n\}$. Consider $M \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}^n$ such that $M\tilde{a}_i + c = a_i$ for $i = 0, \ldots, n$. Then the function $v \colon \mathbb{R}^n \to \mathbb{R}$ with $v(x) = u(M^T x) + c \cdot x$ has the property that $\nabla v(x) \in \{a_0, \ldots a_n\}$ for almost all $x \in \mathbb{R}^n$. Hence we may assume that A consists of the vectors $a_0, \ldots, a_n$.

Now for the sets $\mathcal{F}$, $\mathcal{E}$, and $\mathcal{N}$ as defined in Sect. 2, Proposition 5 implies that $\mathcal{F} \subseteq \mathcal{R}(u)$ with the same arguments as in the Proof of Theorem 4.

For $x \in \mathcal{E}$, the functions $u_{x,\rho}$ converge locally uniformly to $\lambda_i \wedge \lambda_j$ or to $\lambda_i \vee \lambda_j$ as $\rho \searrow 0$ for some $i, j \in \mathbb{Z}_{n+1}$ with $i \neq j$. Moreover, the approximate tangent space of $\mathcal{E}^*$ exists at the point $\boldsymbol{U}(x)$. Clearly this approximate tangent space is $\operatorname{graph}(\lambda_i) \cap \operatorname{graph}(\lambda_j)$. Hence for ρ sufficiently small, the function $u_{x,\rho}$ satisfies the hypotheses of Proposition 13, including (17). It follows that $u_{x,\rho}$ satisfies (18) or (19). In particular, it is regular near 0, and hence $x \in \mathcal{R}(u)$.

Thus $\mathcal{S}(u) \subseteq \mathcal{N}$, which is an $\mathcal{H}^{n-1}$-null set. □

Data Availibility Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs, Oxford University Press, New York (2000)
2. Ball, J.M., James, R.D.: Fine phase mixtures as minimizers of energy. Arch. Rational Mech. Anal. **100**, 13–52 (1987)
3. Cicalese, M., Forster, M., Orlando, G.: Variational analysis of a two-dimensional frustrated spin system: emergence and rigidity of chirality transitions. SIAM J. Math. Anal. **51**, 4848–4893 (2019)
4. Conti, S., Fonseca, I., Leoni, G.: A Γ-convergence result for the two-gradient theory of phase transitions. Comm. Pure Appl. Math. **55**, 857–936 (2002)
5. Davoli, E., Friedrich, M.: Two-well linearization for solid-solid phase transitions, arXiv:2005.03892 [math.AP], (2020)
6. Davoli, E., Friedrich, M.: Two-well rigidity and multidimensional sharp-interface limits for solid-solid phase transitions. Calc. Var. Partial Differ. Equ. **59**, 1–47 (2020) article no. 44
7. Liu, F., Metiu, H.: Dynamics of phase separation of crystal surfaces. Phys. Rev. B **48**, 5808 (1993)
8. Modica, L., Mortola, S.: Il limite nella Γ-convergenza di una famiglia di funzionali ellittici. Boll. Un. Mat. Ital. A (5) **14**, 526–529 (1977)
9. Modica, L., Mortola, S.: Un esempio di Γ^--convergenza. Boll. Un. Mat. Ital. B (5) **14**, 285–299 (1977)
10. Moser, R.: Structure and rigidity of functions in $\mathrm{BV}^2_{\mathrm{loc}}(\mathbb{R}^2)$ with gradients taking only three values. Proc. Lond. Math. Soc. **116**, 813–846 (2018)
11. Müller, S.: Variational models for microstructure and phase transitions, Calculus of variations and geometric evolution problems (Cetraro, 1996). Lecture Notes in Mathematics, vol. 1713, pp. 85–210. Springer, Berlin (1999)
12. Parks, H.R., Wills, D.C.: An elementary calculation of the dihedral angle of the regular n-simplex. Amer. Math. Monthly **109**, 756–758 (2002)
13. Stewart, J., Goldenfeld, N.: Spinodal decomposition of a crystal suface. Phys. Rev. A **46**, 6505–6512 (1992)
14. Watson, S.J., Norris, S.A.: Scaling theory and morphometrics for a coarsening multiscale surface, via a principle of maximal dissipation. Phys. Rev. Lett. **96**, 176103 (2006)