PIECEWISE AFFINE APPROXIMATIONS FOR FUNCTIONS OF BOUNDED VARIATION

JAN KRISTENSEN AND FILIP RINDLER

ABSTRACT. BV functions cannot be approximated well by piecewise constant functions, but this work will show that a good approximation is still possible with (countably) piecewise affine functions. In particular, this approximation is area-strictly close to the original function and the L^1-difference between the traces of the original and approximating functions on a substantial part of the mesh can be made arbitrarily small. Necessarily, the mesh needs to be adapted to the singularities of the BV function to be approximated, and consequently, the proof is based on a blow-up argument together with explicit constructions of the mesh. In the case of $W^{1,1}$-Sobolev functions we establish an optimal $W^{1,1}$-error estimate for approximation by piecewise affine functions on uniform regular triangulations. The piecewise affine functions are standard quasi-interpolants obtained by mollification and Lagrange interpolation on the nodes of triangulations, and the main new contribution here compared to for instance Clément (RAIRO Analyse Numérique 9 (1975), no. R-2, 77–84) and Verfürth (M2AN Math. Model. Numer. Anal. 33 (1999), no. 4, 695–713) is that our error estimates are in the $W^{1,1}$-norm rather than merely the L^1-norm.

1. INTRODUCTION

Functions of bounded variation have important applications in many branches of mathematical physics, among them optimization [4], free-discontinuity problems [3] (this and the previous reference also contain good introductions to the theory of BV functions), and hyperbolic systems of conservation laws [13]. However, apart from applications to image segmentation and related models (see [8] and the references quoted there), the theory of numerical approximations of such functions is not very well developed and indeed, because of the nature of singularities of BV functions, a more thorough analysis is required. In this work we consider the basic question of whether a BV function can be approximated well by (countably) piecewise affine functions. Before we start our investigation, though, we remark that any good approximation by piecewise-constant BV functions must fail, this will be shown in Proposition 4 below.

For $d > 1$ an integer, let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$. The density of piecewise affine functions in $W^{1,p}(\Omega; \mathbb{R}^m)$ for exponents $1 \le p < \infty$ is fundamental for finite element methods, and for example also useful in proving lower semicontinuity of variational functionals, see for example [17, 16, 12]. Such an approximation result is also easily seen to be true in the space BV of functions of bounded variation, if we switch from norm convergence to the so-called (area-)strict convergence (see below).

For some applications, however, one needs a better approximation result, in which also the trace differences over the boundaries of the mesh are controlled. This need becomes immediately obvious for example in the analysis of energy functionals in free-discontinuity problems, which typically involve a term measuring jumps over the discontinuity surface

(see for example [9]), and where, to produce good approximants, one needs to control also the approximation of jumps.

The purpose of this note is to prove the following approximation theorem, related to a weaker result in [21] and the numerical and analytical studies [5, 31, 25, 2, 6, 32].

Theorem 1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, λ a finite Borel measure on Ω and $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$. Then, for every $\varepsilon > 0$ there exist a countable family $\mathcal{R}$ of (rotated) rectangles and simplices $R \subset \Omega$ and $v \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ with the following properties:*

(i) *The sets in $\mathcal{R}$ are non-overlapping and $(\mathcal{L}^d + |Du|)(\Omega \setminus \bigcup \mathcal{R}) = 0$.*
(ii) *For each $R \in \mathcal{R}$ the restriction $v|_R$ is affine.*
(iii) $\|u - v\|_{\mathrm{L}^1(\Omega; \mathbb{R}^m)} + |\langle Du \rangle(\Omega) - \langle Dv \rangle(\Omega)| < \varepsilon$.
(iv) $\mathcal{R} = \mathcal{R}_g \cup \mathcal{R}_b$ *with* $\lambda\left(\bigcup \mathcal{R}_b\right) < \varepsilon$ *and* $\displaystyle\sum_{R \in \mathcal{R}_g} \int_{\partial R} |u - v| \, \mathrm{d}\mathcal{H}^{d-1} < \varepsilon$.
(v) $v|_{\partial \Omega} = u|_{\partial \Omega}$.

We refer to the family $\mathcal{R}$ as a mesh for Ω. Observe that $\mathcal{R}$ need not be locally finite in Ω and therefore the $\mathrm{W}^{1,1}$-function v in the theorem, called an $\mathcal{R}$-**piecewise affine function**, need not be continuous. The property (v) should be understood in the sense of trace in BV (see [3]).

Let us also briefly comment on the notation for measures. If $\mu = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d} \mathcal{L}^d + \mu^s$ is the Lebesgue–Radon–Nikodým decomposition of the measure μ, we define the following measure (related to the area functional)

$$\langle \mu \rangle(A) := |(\mu, \mathcal{L}^d)|(A) = \int_A \sqrt{1 + \left|\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}\right|^2} \, \mathrm{d}x + |\mu^s|(A)$$

for every Borel set $A \subset \Omega$, where $|(\mu, \mathcal{L}^d)|$ and $|\mu^s|$ denote the total variation measures of the vector measures $(\mu, \mathcal{L}^d)$ and μ^s, respectively. We always define the total variation measures with respect to the natural euclidean norms (see [3]).

Strong convergence in BV is not very useful here and instead we shall mainly be concerned with three weaker notions of convergence in BV: A sequence $(u_j) \subset \mathrm{BV}(\Omega; \mathbb{R}^m)$ is said to converge **weakly*** to $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ if $\|u_j - u\|_{\mathrm{L}^1} \to 0$ and $Du_j \overset{*}{\rightharpoonup} Du$ in $\mathrm{C}_0(\Omega, \mathbb{R}^{m \times d})^*$, it is said to converge **strictly** if $\|u_j - u\|_{\mathrm{L}^1} \to 0$ and $|Du_j|(\Omega) \to |Du|(\Omega)$ as $j \to \infty$; and it is said to converge $\langle \cdot \rangle$-**strictly** (or **area-strictly**) if $\|u_j - u\|_{\mathrm{L}^1} \to 0$ and $\langle Du_j \rangle(\Omega) \to \langle Du \rangle(\Omega)$. The latter mode of convergence is the natural replacement for the strong convergence in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ since one can show that smooth functions are $\langle \cdot \rangle$-strictly but not strongly dense in $\mathrm{BV}(\Omega; \mathbb{R}^m)$. In fact, we even have the following stronger result on $\langle \cdot \rangle$-strict density of C^∞ smooth functions under trace constraints (see Lemma 1 of [19] for a slightly more general statement or Lemma B.1 in [7]):

Lemma 2. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$. There exists a sequence $(v_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ such that*

$$v_j \to u \quad \langle \cdot \rangle\text{-strictly and } v_j = u \text{ in the sense of trace on } \partial \Omega \text{ for all } j \in \mathbb{N}.$$

If $u \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$, then we can in addition arrange that $v_j \to u$ strongly in $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$.

See also [28] for an interesting generalization. Many reasonable integral functionals with linear growth are continuous with respect to the $\langle \cdot \rangle$-strict convergence by Reshetnyak's continuity theorem [26] and extensions, see for example Theorem 4 and the appendix of [20]; lower semicontinuity in this situation is discussed in [14].

The main result of this paper establishes density of (countably) piecewise affine functions in BV in the area-strict sense and such that additionally conditions (iv), (v) are satisfied. We remark that it is easy to satisfy conditions (i)–(iii) of Theorem 1 by simply mollifying u,

which gives a strictly close smooth function (by Lemma 2) and then strongly approximating this smooth function by a piecewise affine function. Unfortunately, on this route one loses control of the sum of integrals in (iv), and it is precisely this condition, which is important for instance in applications to free-discontinuity problems.

In Theorem 3 below we establish substantially stronger approximation results for $W^{1,1}$-Sobolev functions. This is done by showing that suitable quasi-interpolants on any given regular and uniform triangulation yield an approximation with error bounds that only depend on the L^1-modulus of continuity of the weak gradient of the approximated $W^{1,1}$-function and regularity/uniformity constants of the triangulation. We refer to Section 2 for notation and terminology. Hence, by virtue of the Kolmogorov-Riesz characterization of compactness in L^1, we have in particular a uniform rate of approximation by quasi-interpolants on all norm-compact subsets of $W^{1,1}$. Of course, similar results remain true in $W^{1,p}$ for $p > 1$, but we shall refrain from stating these results explicitly here and instead focus on the case $p = 1$.

Theorem 3. *Assume that $\mathcal{T}$ is a regular triangulation of $\mathbb{R}^d$ which is also uniform in the sense that there exists a constant $\gamma \in (0, 1)$ such that*

$$\gamma k \leq \operatorname{diam} \tau \leq k \qquad \text{for all } \tau \in \mathcal{T},$$

where

$$k := \sup_{\tau \in \mathcal{T}} \operatorname{diam} \tau.$$

Then, there exists a constant C, only depending on the dimensions, γ, and the parameter of regularity of the triangulation $\mathcal{T}$, such that for any bounded Lipschitz domain Ω in $\mathbb{R}^d$ and any Sobolev function $u \in W^{1,1}(\Omega; \mathbb{R}^m)$ there exists a $\mathcal{T}$-piecewise affine function $a \in W^{1,\infty}(\Omega, \mathbb{R}^m)$ with

$$\int_\Omega \big(|u - a| + |\nabla u - \nabla a| \big) \, \mathrm{d}x \leq C\omega(3k),$$

where ω is a suitable L^1-modulus of continuity for ∇u; a modulus of continuity is suitable if it is a L^1-modulus of continuity for ∇U where U is any $W^{1,1}$-extension of u from Ω to $\mathbb{R}^d$, see (2.3) for the precise definition.

We end this introduction by returning to the issue of approximating with pure jump functions. As mentioned above, if in Theorem 1 we try to approximate with pure jump functions, that is, (countably) piecewise constant BV functions, instead of piecewise affine functions, then the mode of approximation cannot be area-strict. This also gives another reason why condition (iv) in Theorem 1 is of interest:

Proposition 4. *Let $m \in \mathbb{N}$ be a natural number and assume that $u \in C_c^\infty(B(0, 1), \mathbb{R}^m)$ is not identically zero. Then there can be no sequence of piecewise constant functions $u_j \in BV(B(0, 1), \mathbb{R}^m)$ such that $u_j \to u \langle \cdot \rangle$-strictly in BV.*

Proof. Let $F \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ be any continuous integrand satisfying $F \equiv 0$ on $\nabla u(B(0, 1))$, which is a bounded set by assumption, and such that $F(z) = |z|$ for large values of $|z|$. Then clearly its recession function equals $|z|$:

$$F^\infty(z) := \lim_{\substack{t \to \infty \\ z' \to z}} \frac{F(tz')}{t} = |z|.$$

Now if u_j were piecewise constant BV functions such that $u_j \to u \langle \cdot \rangle$-strictly in BV, then by Reshetnyak's continuity theorem [26] (also see [29] and the appendix of [20]), we would have

$$|Du_j| \overset{*}{\rightharpoonup} |Du|$$

and, since $Du_j = D^s u_j$,

$$|Du_j| = F^\infty \left(\frac{\mathrm{d} D^s u_j}{\mathrm{d} |D^s u_j|} \right) |D^s u_j| \overset{*}{\rightharpoonup} F \left(\frac{\mathrm{d} Du}{\mathrm{d} \mathcal{L}^d} \right) \mathcal{L}^d = 0,$$

where both convergences are weak* in $\mathrm{C}_0(B(0,1))^*$. But this is clearly impossible because $Du \neq 0$. $\square$

For another example for the usual strict convergence in the vectorial case, consider the continuous and 1-homogeneous integrand

$$F \colon \mathbb{R}^{2\times 2} \to \mathbb{R}, \qquad F(z) := |\det z|^{1/2}.$$

Then, for any $u_j \to u$ strictly in $\mathrm{BV}(\mathbb{R}^2; \mathbb{R}^2)$, Reshetnyak's continuity theorem implies

$$F \left(\frac{\mathrm{d} Du_j}{\mathrm{d} |Du_j|} \right) |Du_j| \overset{*}{\rightharpoonup} F \left(\frac{\mathrm{d} Du}{\mathrm{d} |Du|} \right) |Du| =: F(Du).$$

However, if the u_j's are pure jump functions, then rank $\frac{\mathrm{d} Du_j}{\mathrm{d} |Du_j|} \leq 1$ $|Du_j|$-almost everywhere, but F is zero on the rank-one cone, and hence $F(Du) = 0$ as a measure would follow. Thus it is not possible to approximate in the strict sense any BV function u with $F(Du) \neq 0$ by piecewise constant BV functions.

This observation should be contrasted with the case of real-valued BV-functions (corresponding to $m = 1$) and the usual strict convergence, where a discretisation of the coarea formula allows strict approximation by pure jump functions, see [9] and [32, Theorem 3].

2. Approximation on a triangulation and proof of Theorem 3

We use standard notation and terminology for simplices and triangulations, and the reader can find all necessary background in the monograph [23]. However, it is convenient to recall the key concepts and definitions. Let τ be a d-**simplex** in $\mathbb{R}^d$, that is, τ is the convex hull of $d+1$ affinely independent points $v_0, \ldots, v_d$ of $\mathbb{R}^d$, the **vertices** of τ. A face of τ is any convex combination of a (non-empty) subset of its vertices, and as such it is a lower dimensional simplex.

Each point $x \in \tau$ admits a unique **barycentric representation** as a convex combination of the vertices of τ:

$$x = \sum_{j=0}^d \lambda_j v_j,$$

where $\lambda_j = \lambda_j(x) \in [0,1]$, $j = 1, \ldots, d$, and $\sum_{j=0}^d \lambda_j \equiv 1$. The functions $\tau \ni x \mapsto \lambda_j(x) \in [0,1]$ are easily seen to be restrictions of affine functions (again denoted) $\lambda_j \colon \mathbb{R}^d \to \mathbb{R}$ satisfying $\lambda_j(\tau) = [0,1]$ and $\sum_{j=0}^d \lambda_j \equiv 1$ on $\mathbb{R}^d$. If $f \colon \tau \to \mathbb{R}^m$ is a function, then the Lagrange interpolation of f on τ is the (unique) affine function $a \colon \mathbb{R}^d \to \mathbb{R}^m$ that agrees with f at the vertices of τ. In terms of the λ_j's, it can be written as

$$a(x) = \sum_{j=0}^\infty \lambda_j(x) f(v_j).$$

Let us also record that the standard d-simplex

$$\sigma = \sigma^d := \left\{ x \in \mathbb{R}^d \ : \ x_j \geq 0 \text{ for all } j, \sum_{j=1}^d x_j \leq 1 \right\}$$

can be used as reference in the sense that given any d-simplex τ we can find a vector $c \in \mathbb{R}^d$ and a linear automorphism $M \in \mathrm{GL}(d)$ such that

$$\tau = c + M\sigma. \tag{2.1}$$

This remark is useful when we define regularity of a triangulation below.

On a d-simplex we have the following elementary bound, where we note that the right-hand side is the local Riesz potential of the second derivative evaluated at the vertices of the simplex. This is the natural counterpart in the context of Lagrange interpolation of the standard potential estimates obtained by expressing functions in terms of their derivatives and the Newtonian potential.

Lemma 5. *Let $\tau = c + M\sigma$, where $c \in \mathbb{R}^d$ and $M \in \mathrm{GL}(d)$ and assume that $1/\alpha \leq \det M \leq \alpha$ for some constant $\alpha > 1$. There exists a constant $C = C(d, m, \alpha)$, only depending on the dimensions and the regularity constant α, with the following property: Assume that $f : \tau \to \mathbb{R}^m$ is a C^2 function with Lagrange interpolant a on the d-dimensional simplex τ. Then,*

$$\int_\tau |\nabla f - \nabla a| \, \mathrm{d}x \leq C(\operatorname{diam} \tau)^d \sum_{j=0}^d \int_\tau \frac{|\nabla^2 f(x)|}{|x - v_j|^{d-1}} \, \mathrm{d}x.$$

Proof. We start from the formula

$$\sum_{j=0}^d (x - v_j) \otimes \nabla \lambda_j(x) = -I \quad \in \mathbb{R}^{d \times d},$$

which is a consequence of $\sum_j \nabla \lambda_j \equiv 0$. Thus,

$$\nabla f(x) - \nabla a(x) = -\nabla f(x) \left(\sum_{j=0}^d (x - v_j) \otimes \nabla \lambda_j \right) + \sum_{j=0}^d \big(f(x) - f(v_j) \big) \otimes \nabla \lambda_j$$

$$= \sum_{j=0}^d \left(\int_0^1 \big(\nabla f(v_j + t(x - v_j)) - \nabla f(x) \big) \cdot (x - v_j) \, \mathrm{d}t \right) \otimes \nabla \lambda_j$$

$$= \sum_{j=0}^d \left(\int_0^1 \int_1^t (x - v_j)^T \, \nabla^2 f(v_j + s(x - v_j)) \, (x - v_j) \, \mathrm{d}s \, \mathrm{d}t \right) \otimes \nabla \lambda_j.$$

Consequently, we find

$$\int_\tau |\nabla f(x) - \nabla a(x)| \, \mathrm{d}x$$

$$\leq \sum_{j=0}^d \int_\tau |\nabla \lambda_j| \int_0^1 |\nabla^2 f(v_j + s(x - v_j))| \cdot |x - v_j|^2 \, \mathrm{d}s \, \mathrm{d}x$$

$$= \sum_{j=0}^d \int_0^1 \int_0^1 \int_{\{\lambda_j = h\} \cap \tau} |\nabla^2 f(v_j + s(x - v_j))| \cdot |x - v_j|^2 \, \mathrm{d}\mathcal{H}^{d-1}(x) \, \mathrm{d}s \, \mathrm{d}h, \tag{2.2}$$

where we used (an elementary version of) the coarea formula for the last equality. Now, for each fixed $h \in (0, 1)$ we change variables via $y = \Phi(x, s) := v_j + s(x - v_j)$, where $(x, s) \in (\tau \cap \lambda_j^{-1}\{h\}) \times (0, 1)$ and $y \in \tau_h := \tau \cap \{\lambda_j \geq h\}$, for which we can estimate

$$\mathrm{d}\mathcal{H}^{d-1}(x) \, \mathrm{d}s = \frac{\mathrm{d}y}{|\det D\Phi|} \sim \frac{\mathrm{d}y}{s^{d-1} h \cdot \operatorname{diam} \tau} \sim \frac{(\operatorname{diam} \tau)^{d-2}}{|y - v_j|^{d-1} \cdot h} \, \mathrm{d}y,$$

since $(\operatorname{diam}\tau)s \sim |y - v_j|$, and where the implied constants depend on α and the dimensions. Furthermore, $|x - v_j| \sim (\operatorname{diam}\tau)h$ and hence

$$\int_0^1 \int_{\{\lambda_j = h\} \cap \tau} |\nabla^2 f(v_j + s(x - v_j))| \cdot |x - v_j|^2 \, d\mathcal{H}^{d-1}(x) \, ds$$

$$\leq C(\operatorname{diam}\tau)^2 \int_0^1 \int_{\{\lambda_j = h\} \cap \tau} |\nabla^2 f(v_j + s(x - v_j))| \cdot h^2 \, d\mathcal{H}^{d-1}(x) \, ds$$

$$\leq C(\operatorname{diam}\tau)^d \int_{\tau_h} \frac{|\nabla^2 f(y)|}{|y - v_j|^{d-1}} \cdot h \, dy.$$

Plugging this into (2.2), we conclude

$$\int_\tau |\nabla f(x) - \nabla a(x)| \, dx \leq C(\operatorname{diam}\tau)^d \sum_{j=0}^d \int_0^1 \int_{\tau_h} \frac{|\nabla^2 f(y)|}{|y - v_j|^{d-1}} \cdot h \, dy \, dh$$

$$\leq C(\operatorname{diam}\tau)^d \sum_{j=0}^d \int_\tau \frac{|\nabla^2 f(y)|}{|y - v_j|^{d-1}} \, dy.$$

This completes the proof. $\qquad\square$

Next we consider triangulations (where, unless stated otherwise, triangulations could possibly be infinite). We start with some standard notation and terminology:

Definition 6. *A **triangulation** of an open subset $\Omega \subset \mathbb{R}^d$ is a family $\mathcal{T}$ of d-simplices $\tau \subset \Omega$ satisfying the following three conditions:*
  (i) *$\overline{\Omega} = \overline{\bigcup \mathcal{T}}$;*
  (ii) *if τ', $\tau'' \in \mathcal{T}$ and $\tau' \cap \tau'' \neq \emptyset$, then $\tau' \cap \tau''$ is a common face;*
  (iii) *every $\tau \in \mathcal{T}$ lies in an open set V which intersects only a finite number of simplices from $\mathcal{T}$.*

*The triangulation $\mathcal{T}$ is called **regular** if there exists a constant $\alpha \geq 1$ such that whenever $\tau \in \mathcal{T}$ is represented as in (2.1) then*

$$\alpha^{-1} \leq |\det M| \leq \alpha.$$

A function $f \colon \Omega \to \mathbb{R}^m$ is said to be piecewise affine with respect to $\mathcal{T}$ (or $\mathcal{T}$-piecewise affine) if for each $\tau \in \mathcal{T}$ the restriction $f|_\tau$ is an affine function (that is, equals the restriction $a|_\tau$ of an affine function $a \colon \mathbb{R}^d \to \mathbb{R}^m$).

Note that triangulations $\mathcal{T}$ must be locally finite in Ω, but that they can become infinite towards the boundary of Ω. Therefore a $\mathcal{T}$-piecewise affine $W^{1,1}$-function will in general only be locally Lipschitz in Ω.

We are now ready for the main purpose of this section.

Proof of Theorem 3. Using a standard extension theorem (see [22]) we may assume that $u \in W^{1,1}(\mathbb{R}^d; \mathbb{R}^m)$ and let for each $h \geq 0$,

$$\omega(h) := \sup_{|y| \leq h} \int_{\mathbb{R}^d} |\nabla u(x + y) - \nabla u(x)| \, dx \tag{2.3}$$

be an L^1-modulus of continuity for ∇u. We only show the integral bound for derivatives, the corresponding bound for $u - v$ is similar, but easier.

Apply Lemma 5 to the regularized mapping $f = \varphi_\varepsilon \star u$, where $\varphi \in C_c^\infty(\mathbb{R}^d)$ is non-negative, has compact support in $B(0, 1)$, integral equal to 1, and $\varphi_\varepsilon(x) := \varepsilon^{-d} \varphi(x/\varepsilon)$ for $\varepsilon > 0$ to be chosen later. For integers $1 \leq k, l \leq d$ we have

$$\partial_{kl} f = \frac{1}{\varepsilon} (\partial_l \varphi)_\varepsilon \star (\partial_k u - Z^k) \qquad \text{for any } Z = (Z^1, \ldots, Z^d) \in \mathbb{R}^{m \times d}.$$

Let τ be a fixed simplex in the triangulation $\mathcal{T}$. Consequently, we may estimate using Lemma 5,

$$\int_\tau |\nabla f(x) - \nabla a(x)| \, \mathrm{d}x$$

$$\leq \frac{C}{\varepsilon} \sum_{j=0}^d (\operatorname{diam}\tau)^d \int_\tau \int_{B(0,\varepsilon)} \frac{|(\nabla\varphi)_\varepsilon(y)| \cdot |\nabla u(x-y) - Z|}{|x - v_j|^{d-1}} \, \mathrm{d}y \, \mathrm{d}x$$

$$\leq \frac{C}{\varepsilon} \sum_{j=0}^d (\operatorname{diam}\tau)^d \varepsilon^{-d} \|\nabla\varphi\|_{L^\infty} \int_\tau \int_{B(0,\varepsilon)} \frac{|\nabla u(x-y) - Z|}{|x - v_j|^{d-1}} \, \mathrm{d}y \, \mathrm{d}x$$

$$\leq \frac{C}{\varepsilon^{d+1}} \|\nabla\varphi\|_{L^\infty} (\operatorname{diam}\tau)^d \sum_{j=0}^d \int_{\tau+B(0,\varepsilon)} |\nabla u(w) - Z| \int_{B(0,\varepsilon)} \frac{\mathrm{d}y}{|w+y-v_j|^{d-1}} \, \mathrm{d}w.$$

Here we remark that all constants can be chosen independently of τ, because the triangulation is regular. Taking

$$\varepsilon = k := \sup_{\tau\in\mathcal{T}} \operatorname{diam}\tau,$$

we note

$$\int_{B(0,\varepsilon)} \frac{\mathrm{d}y}{|w+y-v_j|^{d-1}} \leq \int_{B(v_j-w,3\varepsilon)} \frac{\mathrm{d}y}{|w+y-v_j|^{d-1}} = \int_{B(0,3\varepsilon)} \frac{\mathrm{d}x}{|x|^{d-1}} = 3\omega_d\varepsilon,$$

and so, for all $Z \in \mathbb{R}^{m\times d}$,

$$\int_\tau |\nabla f - \nabla a| \, \mathrm{d}x \leq c \int_{\tau+B(0,k)} |\nabla u - Z| \, \mathrm{d}x, \tag{2.4}$$

where $c = 3(d+1)\omega_d C\|\varphi\|_{L^\infty}$. Let us denote $\tau^+ := \tau + B(0,k)$ and plug $Z := (\nabla u)_{\tau^+} := \fint_{\tau^+} \nabla u \, \mathrm{d}x$ into (2.4). Then we estimate

$$\int_\tau |\nabla f - \nabla a| \, \mathrm{d}x \leq c \int_{\tau^+} |\nabla u - (\nabla u)_{\tau^+}| \, \mathrm{d}x$$

$$\leq c \fint_{\tau^+} \int_{\tau^+} |\nabla u(x) - \nabla u(y)| \, \mathrm{d}y \, \mathrm{d}x$$

$$\leq c \fint_{\tau^+} \int_{B(0,\operatorname{diam}\tau^+)} |\nabla u(x) - \nabla u(x+w)| \, \mathrm{d}w \, \mathrm{d}x$$

$$\leq c \fint_{B(0,\operatorname{diam}\tau^+)} \int_{\tau^+} |\nabla u(x) - \nabla u(x+w)| \, \mathrm{d}x \, \mathrm{d}w.$$

$$\leq \frac{c}{k^d} \int_{B(0,3k)} \int_{\tau^+} |\nabla u(x) - \nabla u(x+w)| \, \mathrm{d}x \, \mathrm{d}w.$$

Next we sum these integrals over all simplices $\tau \in \mathcal{T}$ with $\tau \cap \Omega \neq \emptyset$ and use the bounded overlap property of $\{\tau^+\}_{\tau\in\mathcal{T}}$, which follows from the uniformity of the triangulation, to get

$$\int_\Omega |\nabla f - \nabla a| \, \mathrm{d}x \leq \frac{c_1}{k^d} \int_{B(0,3k)} \int_{\mathbb{R}^d} |\nabla u(x) - \nabla u(x+w)| \, \mathrm{d}x \, \mathrm{d}w$$

$$\leq c_1 3^d \omega(3k).$$

Here the constant c_1 depends on γ, α and the dimensions. Using finally the standard fact that also $\|\nabla u - \nabla f\|_{L^1(\Omega;\mathbb{R}^m)} \leq \omega(k)$ we finish the proof. $\square$

Remark that by virtue of the Kolmogorov-Riesz characterization of compactness in L^1, see for example [18], and Theorem 3 we therefore obtain a uniform rate of convergence by quasi-interpolants for all functions u in any fixed compact subset of $W^{1,1}(\Omega;\mathbb{R}^m)$. The relation between compactness and regularity is also discussed in [15].

FIGURE 1. Triangulation used in the proof of Lemma 7 ($\eta = \mathrm{dist}(Q^c, Q_0)$).

3. PROOF OF THEOREM 1

We first establish the following gluing lemma:

Lemma 7. *Let Q be an open (possibly rotated) cube in $\mathbb{R}^d$, let $Q_0 \subset\subset Q$ be a concentric similar open subcube, and $u \in \mathrm{BV}(Q; \mathbb{R}^m)$ with $|Du|(\partial Q_0) = 0$. We further assume:*

(i) *The closure $\overline{Q_0}$ is the disjoint union of the closure of a finite number of similar (open) rectangles $S_1, \ldots, S_m$ that are translations of one another along a single axis parallel to one of the sides of Q_0, and with the length η of the short side of the S_j, it holds that $\mathrm{dist}(\partial Q, Q_0) = \eta$ (like in Figure 1).*

(ii) *We are given a function $w \in \mathrm{W}^{1,\infty}(Q_0; \mathbb{R}^m)$ that is affine when restricted to any S_j, and satisfies*

$$\sum_{k=1}^{n} \int_{\partial S_k} |u - w| \, \mathrm{d}\mathcal{H}^{d-1} \leq \varepsilon |Du|(Q).$$

Then, on $A := Q \setminus \overline{Q_0}$ there exists a countably piecewise affine function $a \in \mathrm{W}^{1,1}(A; \mathbb{R}^m)$ satisfying $a = u$ on ∂Q, $a = w$ on ∂Q_0, and

$$\|u - a\|_{\mathrm{L}^1(A;\mathbb{R}^m)} \leq C\varepsilon \|u\|_{\mathrm{L}^1(A;\mathbb{R}^m)}, \quad |\langle Da \rangle(A) - \langle Du \rangle(A)| \leq C\varepsilon \langle Du \rangle(Q). \qquad (3.1)$$

Here, $C = C(d, m)$ is a dimensional constant.

Proof of Lemma 7. We start by selecting a smooth cut-off function $\rho \colon \overline{Q} \to [0, 1]$ with $\rho = 1$ on Q_0, $\rho = 0$ on ∂Q and $|\nabla \rho| \leq 2/\mathrm{dist}(Q_0, \partial Q)$ satisfying

$$\int_A \rho\big(|w| + |u|\big) \, \mathrm{d}x \leq \varepsilon |Du|(Q)$$

and

$$\int_A \big(\rho|\nabla w - \nabla u| + |w - u||\nabla \rho|\big)\ \mathrm{d}x + \int_A \rho\ \mathrm{d}|D^s u| \leq \varepsilon |Du|(Q).$$

We omit the routine details of this and only remark that it can be achieved for instance by suitably mollifying the indicator function for Q_0 and using our assumptions. Put $g = \rho w + (1 - \rho)u$, and apply Lemma 2 to find $f \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(A; \mathbb{R}^m)$ satisfying $f = w$ on ∂Q_0, $f = u$ on ∂Q and

$$\int_A |f - g|\ \mathrm{d}x + |\langle Dg \rangle (A) - \langle Df \rangle (A)| \leq \varepsilon |Du|(Q).$$

Next, we construct a triangulation $\mathcal{T} = \bigcup_{j \in \mathbb{N}} \tau_j$ of A as in Figure 1. This triangulation is of Whitney-type towards the outer boundary ∂Q,

$$\mathrm{diam}\,\tau \sim \mathrm{dist}(\tau, \partial Q). \tag{3.2}$$

Our $\mathcal{T}$ also has the property that the simplices match the elements $S_1, \ldots, S_n$ at the inner boundary ∂Q_0, meaning that each rectangle $S_k \cap \partial Q_0$ is a finite union of simplices $\tau \cap \partial Q_0$ where $\tau \in \mathcal{T}$ (but not necessarily only one like in the figure). We take a sufficiently fine triangulation until we reach the point where the corresponding $\mathcal{T}$-piecewise affine function $a \colon A \to \mathbb{R}^m$ obtained by Lagrange interpolating g on the nodes of $\mathcal{T}$ satisfies

$$\int_A \big(|g - a| + |\langle \nabla g \rangle - \langle \nabla a \rangle|\big)\ \mathrm{d}x \leq \varepsilon |Du|(Q).$$

Of course the implied constants in (3.2) change when the triangulation is refined, but this is unimportant. The important fact to note is that since the simplices shrink as we approach the outer boundary ∂Q we hereby achieve that $a = g$ on ∂Q (in the sense of trace), and therefore $a = u$ on ∂Q. On the inner boundary ∂Q_0 we have $g = w$, and since $w|_{\partial Q_0}$ is piecewise affine with respect to the (lower dimensional) triangulation $\{\tau \cap Q_0\}_{\tau \in \mathcal{T}}$ we also have that $a = w$ on ∂Q_0. Finally, we estimate

$$\int_A |g - u|\ \mathrm{d}x \leq \int_A \rho\big(|w| + |u|\big)\ \mathrm{d}x \leq \varepsilon |Du|(Q),$$

$$|\langle Dg \rangle (A) - \langle Du \rangle (A)| \leq \int_A \rho\big(|\nabla w - \nabla u| + |w - u||\nabla \rho|\big)\ \mathrm{d}x + \int_A \rho\ \mathrm{d}|D^s u|$$
$$\leq \varepsilon |Du|(Q).$$

We conclude by use of the triangle inequality. $\square$

   We can now show our main result.

Proof of Theorem 1. We introduce the following notation: Let $C = (-1,1)^d$ and for $x_0 \in \mathbb{R}^d$, $r > 0$ set $Q(x_0, r) := x_0 + rC$. For every unit vector $n \in \mathbb{S}^{d-1}$ select a rotation $P \in \mathrm{SO}(d)$ with $Pe_1 = n$, where $e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^d$. Let $Q_n(x_0, r) := x_0 + rP(C)$, so that $Q_n(x_0, r)$ is an open cube with center x_0, sidelength $2r$, and two faces orthogonal to n.
   Step 1. For $\mathcal{L}^d$-almost every $x_0 \in \Omega$ and $0 < r < \mathrm{dist}(x_0, \partial \Omega)/\sqrt{d}$ put

$$u_r(y) := \frac{u(x_0 + ry) - \tilde{u}(x_0)}{r}, \qquad y \in C,$$

where $\tilde{u}(x_0)$ is the value of the precise representative of u at x_0 (for this and other properties of BV-functions see [3]). We recall that for $\mathcal{L}^d$-almost all $x_0 \in \Omega$, u is approximately differentiable at x_0 and so, there is a sequence $r \downarrow 0$ (not specifically labelled) with the property that

$$u_r \to u_0 \quad \langle \cdot \rangle\text{-strictly in BV}(C; \mathbb{R}^m) \quad \text{with} \quad u_0(x) := \nabla u(x_0)x, \ x \in C, \tag{3.3}$$

with the approximate gradient $\nabla u(x_0)$ of u at x_0. The fact that one can indeed choose a sequence $r = r_n \downarrow 0$ such that $u_r \to u_0$ $\langle \cdot \rangle$-strictly follows for example from Lemma 3.1 in [27]

about the construction of strictly converging blow-ups, applied to the measure $(Du, \mathcal{L}^d)$. Since the trace operator is strictly continuous, we also have that

$$\int_{\partial C} |u_r - u_0| \, d\mathcal{H}^{d-1} \to 0 \quad \text{as} \quad r \downarrow 0; \tag{3.4}$$

see for instance [20], pp. 53–54, and the references given there. Let us denote the set of such points $x_0 \in \Omega$ by G_1.

Now, for

$$w(x) = \tilde{u}(x_0) + \nabla u(x_0)(x - x_0), \qquad x \in Q(x_0, r), \tag{3.5}$$

from (3.3), (3.4) we get by a change of variables that there exists $0 < r(x_0) < 1$ such that for $r < r(x_0)$ we have

$$\int_{Q(x_0, r)} |u - w| \, dx < \varepsilon r \mathcal{L}^d(Q(x_0, r)),$$

$$\left| \langle Du \rangle (Q(x_0, r)) - \langle Dw \rangle (Q(x_0, r)) \right| < \frac{\varepsilon}{2} \mathcal{L}^d(Q(x_0, r)),$$

$$\int_{\partial Q(x_0, r)} |u - w| \, d\mathcal{H}^{d-1} < \frac{\varepsilon}{2} \mathcal{L}^d(Q(x_0, r)). \tag{3.6}$$

Step 2. Next, for $|D^s u|$-almost all $x_0 \in \Omega$, henceforth fixed, we have $\alpha_r := \frac{|Du|(Q_n(x_0, r))}{r^d} \to \infty$ as $r \downarrow 0$, and, defining

$$u_r(y) := \frac{u(x_0 + ry) - u_{x_0, r}}{r \alpha_r}, \qquad y \in Q_n(0, 1),$$

where $u_{x_0, r} = \fint_{Q_n(x_0, r)} u \, dx$, by Lemma 3.1 of [27] we can find a sequence of r's going to 0 (not specifically labelled) such that

$$u_r \to u_0 \quad \text{strictly in BV}(Q_n(0, 2); \mathbb{R}^m) \tag{3.7}$$

and

$$u_0(y) = b\psi(y \cdot n), \qquad y \in Q_n(0, 2), \tag{3.8}$$

where $\psi \colon (-1, 1) \to \mathbb{R}$ is increasing and bounded. The fact that the blow-up limit u_0 can indeed be chosen one-directional can be proved via Alberti's Rank One Theorem [1], see Theorem 3.95 in [3]. Alternatively, we can employ the rigidity result in Lemma 3.2 of [27] (also see Remark 5.8 in *loc. cit.*); then we need to treat the additional case of an affine blow-up, but this is in fact easier than the one-directional case.

It is not hard to see that we additionally may assume that the sequence of r's is chosen such that $|Du|(\partial Q_n(x_0, r)) = 0$; we again refer to [20], pp. 54–55, for these assertions. Denote the set of such points $x_0 \in \Omega$ by G_2, and observe that $(\mathcal{L}^d + |Du|)(\Omega \setminus (G_1 \cup G_2)) = 0$.

Let N satisfy

$$N > \frac{\max\{2^{d+1}, (d-1)2^d\} |b|}{\varepsilon} (\psi(2) - \psi(-2)).$$

For the function ψ appearing in (3.8) we claim that we can require that the equidistant partition

$$-1 = t_0 < t_1 < \ldots < t_N = 1 \qquad \text{with} \qquad t_{j+1} - t_j = \frac{2}{N}.$$

consists only of continuity points of ψ. This can be achieved as follows: Select $0 < \theta < 1$ such that for the modified function $\psi_\theta(t) := \psi(t + \theta)$ all the t_j are continuity points. Since ψ has only countably many discontinuity points, such θ always exists. This corresponds to a blow-up sequence of the form

$$u_r(y) := \frac{u(x_0 + r\theta n + ry) - u_{x_0, r}}{r \alpha_r}, \qquad y \in Q_n(0, 1),$$

and in the following we need to replace $Q_n(x_0, r)$ by $Q_n(x_0 + r\theta n, r)$. We note that N still satisfies

$$N > \frac{\max\{2^{d+1}, (d-1)2^d\} |b|}{\varepsilon} (\psi(1) - \psi(-1)).$$

In the following we will however suppress such a possible shift for ease of notation.

Define φ as the piecewise affine function satisfying $\varphi(t_j) = \psi(t_j)$ for each $j = 0, \ldots, N$, and note that with (the rotation P as above),

$$S_j := P[(t_j, t_{j+1}) \times (-1, 1)^{d-1}], \qquad j = 0, \ldots, N-1,$$

we have

$$\int_{Q_n} |u_0(y) - b\varphi(y \cdot n)| \, \mathrm{d}y = 2^{d-1}|b| \int_{-1}^1 |\psi - \varphi| \, \mathrm{d}s$$

$$\leq \frac{2^d|b|}{N} \sum_{j=0}^{N-1} \psi(t_{j+1}) - \psi(t_j) < \frac{\varepsilon}{2} \tag{3.9}$$

and

$$\sum_{j=0}^{N-1} \int_{\partial S_j} |u_0(y) - b\varphi(y \cdot n)| \, \mathrm{d}\mathcal{H}^{d-1}(y)$$

$$= |b| \sum_{j=0}^{N-1} \int_{\partial S_j} |\psi(y \cdot n) - \varphi(y \cdot n)| \, \mathrm{d}\mathcal{H}^{d-1}(y)$$

$$\leq \frac{(d-1)2^{d-1}|b|}{N} \sum_{j=0}^{N-1} \psi(t_{j+1}) - \psi(t_j) < \frac{\varepsilon}{2}. \tag{3.10}$$

Next, in view of (3.7) and our choice of partition points t_j, we infer from the trace theorem that

$$\sum_{j=0}^{N-1} \int_{\partial S_j} |u_r - u_0| \, \mathrm{d}\mathcal{H}^{d-1} \to 0 \text{ as } r \downarrow 0. \tag{3.11}$$

For a point $x_0 \in G_2$, the mapping $x \mapsto b\varphi(x \cdot n)$ defined above is piecewise affine. Split the the rotated cube $Q_n(x_0, r)$ into N rectangles $S_j(x_0, r) := x_0 + rS_j$ and define the corresponding piecewise affine map

$$w(x) := u_{x_0, r} + r\alpha_r b \varphi\left(\frac{x - x_0}{r} \cdot n\right), \qquad x \in Q_n(x_0, r). \tag{3.12}$$

Hence, changing variables in (3.7), (3.11) and using (3.9), (3.10) as well as $|Du_0|(Q_n(0, 1)) = |D[b\varphi(y \cdot n)]|(Q_n(0, 1))$ and the estimate $|\sqrt{1 + |\cdot|^2} - |\cdot|| \leq 1$, it follows that there exists $r(x_0) > 0$ such that for $r < r(x_0)$,

$$\int_{Q_n(x_0, r)} |u - w| \, \mathrm{d}x < \varepsilon r |Du|(Q_n(x_0, r)),$$

$$\left| \langle Du \rangle(Q_n(x_0, r)) - \langle Dw \rangle(Q_n(x_0, r)) \right| < \frac{\varepsilon}{2} \langle Du \rangle(Q_n(x_0, r)) + 2\mathcal{L}(Q_n(x_0, r)),$$

$$\sum_{j=0}^{N-1} \int_{\partial S_j(x_0, r)} |u - w| \, \mathrm{d}\mathcal{H}^{d-1} < \frac{\varepsilon}{2} |Du|(Q_n(x_0, r)). \tag{3.13}$$

Step 3. For every $x_0 \in G := G_1 \cup G_2$ we have so far constructed a (rotated) cube $Q_0 = Q_n(x_0, r)$ with n and r depending on x_0, and for all $x_0 \in G_2$ this cube is further subdivided into rectangles $S_j(x_0, r)$ ($j = 0, \ldots, N-1$). Now, for every such Q_0 we choose

a slightly larger concentric similar cube $Q = Q(x_0, r) \supset\supset Q_0$ with the properties

$$\|u\|_{\mathrm{L}^1(Q \setminus \overline{Q_0}; \mathbb{R}^m)} \leq \varepsilon \|u\|_{\mathrm{L}^1(Q_0; \mathbb{R}^m)},$$

$$\lambda(Q \setminus \overline{Q}_0) + \langle Du \rangle(Q \setminus \overline{Q}_0) + \int_{\partial Q_0} |u - w| \, \mathrm{d}\mathcal{H}^{d-1} < \varepsilon \langle Du \rangle(Q_0). \qquad (3.14)$$

We invoke Lemma 7 with Q_0 and Q and with w as in (3.5) or (3.12) for $x_0 \in G_1$ or $x_0 \in G_2$, respectively. In particular, this yields a (countably) piecewise affine function v_Q with the properties stated in that lemma.

Step 4. We next apply the Morse Covering Theorem [24] (or see Theorem 5.51 in [3]) to cover $(\mathcal{L}^d + |Du|)$-almost all of Ω with (rotated) cubes Q from the above family. After subdividing the cubes $Q_0 = Q_0(x_0, r)$ corresponding to points $x_0 \in G_2$ into the rectangles $S_j(x_0, r)$ and the set $Q(x_0, r) \setminus Q_0(x_0, r)$ into the simplices constructed in Lemma 7, we thereby find a family $\mathcal{R}$ satisfying (i) in the statement of Theorem 1.

For the remaining properties (ii)–(v), we write

$$v = \sum_{R \in \mathcal{R}} a_R \mathbb{1}_R,$$

where a_R denotes the affine map corresponding to the rectangle or simplex $R \in \mathcal{R}$ (in particular $v_Q = \sum_{R \subset Q} a_R$ for any v_Q from the previous step). Because the rectangles/simplices are non-overlapping, the map v is well-defined and it is clear that (ii) is satisfied. For the remaining assertions we consider for $j \in \mathbb{N}$ the mapping

$$v_j = u \mathbb{1}_{H_j} + \sum_{\mathcal{L}^d(R) > \frac{1}{j}} a_R \mathbb{1}_R, \qquad \text{with} \qquad H_j = \Omega \setminus \left(\bigcup_{\mathcal{L}^d(R) > \frac{1}{j}} \overline{R} \right).$$

Since the above sum is finite, we infer that $v_j \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ and

$$\begin{aligned} Dv_j = {}& Du \llcorner H_j + u \otimes \nu_{H_j} \, \mathcal{H}^{d-1} \llcorner \partial H_j \\ &+ \sum_{\mathcal{L}^d(R) > \frac{1}{j}} \left(\nabla a_R \, \mathcal{L}^d \llcorner R + a_R \otimes \nu_R \, \mathcal{H}^{d-1} \llcorner \partial R \right) \\ = {}& Du \llcorner H_j + \sum_{\mathcal{L}^d(R) > \frac{1}{j}} \left(\nabla a_R \, \mathcal{L}^d \llcorner R + (a_R - u) \otimes \nu_R \, \mathcal{H}^{d-1} \llcorner \partial R \right). \end{aligned}$$

Here, ν_{H_j} and ν_R are the unit inner normals to H_j and R, respectively. Employing (3.6), (3.13),

$$\begin{aligned} |Dv_j|(\Omega) &= |Du|(H_j) + \sum_{\mathcal{L}^d(R) > \frac{1}{j}} \left(|\nabla a_R| \, \mathcal{L}^d(R) + \int_{\partial R} |a_R - u| \, \mathrm{d}\mathcal{H}^{d-1} \right) \\ &= |Du|(H_j) + \sum_{\mathcal{L}^d(R) > \frac{1}{j}} |\nabla a_R| \, \mathcal{L}^d(R) + \mathrm{O}(\varepsilon)(\mathcal{L}^d + |Du|)(\Omega). \end{aligned}$$

Since $|Du|(H_j) \to 0$ as $j \to \infty$, we see that $v \in \mathrm{BV}(\Omega; \mathbb{R}^m)$.

Concerning (iii), we estimate using (3.6), (3.13), (3.1), (3.14),

$$\begin{aligned} \left| \langle Du \rangle(\Omega) - \langle Dv \rangle(\Omega) \right| &\leq \sum_Q \left[\left| \langle Du \rangle(Q_0) - \langle Dv \rangle(Q_0) \right| + (\langle Du \rangle + \langle Dv \rangle)(Q \setminus Q_0) \right] \\ &\leq \sum_Q \left[\varepsilon(\mathcal{L}^d + \langle Du \rangle)(Q_0) + C(\mathcal{L}^d + \langle Du \rangle)(Q \setminus Q_0) \right] + \mathcal{L}^d(Z) \\ &\leq \varepsilon C \sum_Q (\mathcal{L}^d + \langle Du \rangle)(Q_0) + \mathcal{L}^d(Z) \\ &\leq \varepsilon C (\mathcal{L}^d + \langle Du \rangle)(\Omega) + \mathcal{L}^d(Z), \end{aligned}$$

where the summation is over all cubes used to cover Ω; inside the sum $Q_0 = Q_n(x_0, r)$ refers to the *inner* cube. The term $\mathcal{L}^d(Z)$ originates from the additional Lebesgue measure for the cubes corresponding to singular points $x \in G_2$. Because they only occur on a $\mathcal{L}^d$-negligible set, we can make this term disappear in the limit $\varepsilon \to 0$.

By a similar calculation, also (iv) holds. We let $\mathcal{R}_g$ denote the collection of rectangles and simplices outside the annuli $Q \setminus Q_0$ and $\mathcal{R}_b := \mathcal{R} \setminus \mathcal{R}_g$ and then we estimate:

$$\sum_{R \in \mathcal{R}_g} \int_{\partial R} |u - v| \, \mathrm{d}\mathcal{H}^{d-1}$$

$$= \sum_Q \left[\sum_{j=0}^{N-1} \int_{\partial S_j(x_0, r)} |u - v| \, \mathrm{d}\mathcal{H}^{d-1} + \sum_{\tau \in \mathcal{T}} \int_{\partial \tau} |u - v| \, \mathrm{d}\mathcal{H}^{d-1} \right]$$

$$\leq \sum_Q \left[\varepsilon(\mathcal{L}^d + |Du|)(Q) + C|Du|(Q_n(x_0, r) \setminus Q_0(x_0, r)) \right]$$

$$< \varepsilon(1 + C)(\mathcal{L}^d + |Du|)(\Omega).$$

For (v) we only need to observe that $v_j = u$ on $\partial \Omega$ for all $j \in \mathbb{N}$ and $v_j \to u$ strictly, hence the trace is preserved.

The only remaining part to check is whether the constructed mapping v is of class $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$. The features we shall use here are that for every cube Q as before, $v|_Q \in \mathrm{W}^{1,1}(Q; \mathbb{R}^m)$, $v = u$ on ∂Q, and that we may assume that $|Du|(\partial Q) = 0$. We have

$$Dv = \sum_Q \left(Dv \llcorner Q + v|_{\partial Q} \otimes \nu_Q \, \mathcal{H}^{d-1} \llcorner \partial Q \right).$$

By assumption $v|_{\partial Q} = u|_{\partial Q}$ and the latter coincides also with the *outer* trace of u on ∂Q since $|Du|(\partial Q) = 0$. Keeping in mind that $\sum_{Q \in \mathcal{Q}} \mathbb{1}_Q = \mathbb{1}_\Omega$ $\mathcal{L}^d$-almost everywhere, and hence in the sense of $\mathrm{L}^1(\mathbb{R}^d)$, we find

$$\sum_{Q \in \mathcal{Q}} v|_{\partial Q} \otimes \nu_Q \, \mathcal{H}^{d-1} \llcorner \partial Q = \sum_{Q \in \mathcal{Q}} u|_{\partial Q} \otimes D(\mathbb{1}_Q) = u \otimes \nu_\Omega \, \mathcal{H}^{d-1} \llcorner \partial \Omega.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[1] G. Alberti, *Rank one property for derivatives of functions with bounded variation*, Proc. Roy. Soc. Edinburgh Sect. A **123** (1993), 239–274.

[2] M. Amar and V. De Cicco, *A new approximation result for BV-functions*, C. R. Math. Acad. Sci. Paris **340** (2005), 735–738.

[3] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free-Discontinuity Problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces – Applications to PDEs and Optimization*, MPS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics, 2006.

[5] S. Bartels, *Total variation minimization with finite elements: convergence and iterative solution*, preprint, 2011.

[6] P. Bechler, R. Devore, A. Kamont, G. Petrova, and P. Wojtaszczyk, *Greedy wavelet projections are bounded on BV*, Trans. Amer. Math. Soc. **359** (2007), 619–635.

[7] M. Bildhauer, *Convex variational problems*, Lecture Notes in Mathematics, vol. 1818, Springer, 2003.

[8] B. Bourdin and A. Chambolle, *Implementation of an adaptive finite-element approximation of the Mumford-Shah functional*, Numer. Math. **85** (2000), 609–646.

[9] A. Braides and A. Coscia, *The interaction between bulk energy and surface energy in multiple integrals*, Proc. Roy. Soc. Edinburgh Sect. A **124** (1994), 737–756.

[10] P.G. Ciarlet and C. Wagschal, *Multipoint Taylor formulas and applications to the finite element method*, Numer. Math. **17** (1971), 84–100.

[11] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Analyse Numérique **9** (1975), no. R-2, 77–84.

[12] B. Dacorogna, *Direct Methods in the Calculus of Variations*, 2nd ed., Applied Mathematical Sciences, vol. 78, Springer, 2008.

[13] C. M. Dafermos, *Hyperbolic conservation laws in continuum physics*, 3rd ed., Grundlehren der mathematischen Wissenschaften, vol. 325, Springer, Berlin, 2010.

[14] S. Delladio, *Lower semicontinuity and continuity of functions of measures with respect to the strict convergence*, Proc. Roy. Soc. Edinburgh Sect. A **119** (1991), 265–278.

[15] D. Faraco and J. Kristensen, *Compactness versus regularity in the calculus of variations*, Discrete Contin. Dyn. Syst. Ser. B **17** (2012), 473–485.

[16] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1983.

[17] C. Goffman and J. Serrin, *Sublinear functions of measures and variational integrals*, Duke Math. J. **31** (1964), 159–178.

[18] H. Hanche-Olsen and H. Holden, *The Kolmogorov-Riesz compactness theorem*, Expositiones Math.

[19] J. Kristensen and F. Rindler, *Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV*, Arch. Ration. Mech. Anal. **197** (2010), 539–598.

[20] ———, *Relaxation of signed integral functionals in BV*, Calc. Var. Partial Differential Equations **37** (2010), 29–62.

[21] ———, *Erratum to "Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV"*, Arch. Ration. Mech. Anal. **203** (2012), 693–700.

[22] V. Maz'ya, *Sobolev spaces with applications to elliptic partial differential equations*, 2nd ed., Grundlehren der mathematischen Wissenschaften, vol. 342, Springer, 2011.

[23] E.E. Moise, *Geometric topology in dimensions 2 and 3*, Graduate Texts in Mathematics, vol. 47, Springer, 1977.

[24] A. P. Morse, *Perfect blankets*, Trans. Amer. Math. Soc. **61** (1947), 418–442.

[25] T. Quentin de Gromard, *Strong approximation of sets in BV(Ω)*, Proc. Roy. Soc. Edinburgh Sect. A **138** (2008), 1291–1312.

[26] Y. G. Reshetnyak, *Weak convergence of completely additive vector functions on a set*, Siberian Math. J. **9** (1968), 1039–1045.

[27] F. Rindler, *Lower semicontinuity and Young measures in BV without Alberti's Rank-One Theorem*, Adv. Calc. Var. **5** (2012), 127–159.

[28] T. Schmidt, *Strict interior approximation of sets of finite perimeter and functions of bounded variation*, Proc. Amer. Math. Soc., to appear.

[29] D. Spector, *Simple proofs of some results of Reshetnyak*, Proc. Amer. Math. Soc. **139** (2011), no. 5, 1681–1690.

[30] R. Verfürth, *Error estimates for some quasi-interpolant operators*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 4, 1766–1782.

[31] S. Waldron, *Multipoint Taylor formulae*, Numer. Math. **80** (1998), 461–494.

[32] P. Wojtaszczyk, *Projections and non-linear approximation in the space $BV(\mathbb{R}^d)$*, Proc. London Math. Soc. **87** (2003), 471–497.

[33] W. P. Ziemer, *Weakly differentiable functions*, Graduate Texts in Mathematics, vol. 120, Springer, 1989.

Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, United Kingdom.

E-mail address: `kristens@maths.ox.ac.uk`

Cambridge Centre for Analysis, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kingdom.

E-mail address: `F.Rindler@maths.cam.ac.uk`