

# CLASSIFICATION OF AREA-STRICT LIMITS OF PLANAR BV HOMEOMORPHISMS

DANIEL CAMPBELL, AAPO KAURANEN, AND EMANUELA RADICI

ABSTRACT. We present a classification of area-strict limits of planar  $BV$  homeomorphisms. This class of mappings allows for cavitations and fractures but fulfil a suitable generalization of the INV condition. As pointed out by J. Ball [4], these features are expected in limit configurations of elastic deformations. In [12], De Philippis and Pratelli introduced the *no-crossing* condition which characterizes the  $W^{1,p}$  closure of planar homeomorphisms. In the current paper we show that a suitable version of this concept is equivalent with a map,  $f$ , being the area-strict limit of BV homeomorphisms. This extends our results from [10], where we proved that the *no-crossing BV* condition for a BV map was equivalent with the map being the m-strict limit of homeomorphisms (i.e.  $f_k \xrightarrow{*} f$  and  $|D_1 f_k|(\Omega) + |D_2 f_k|(\Omega) \rightarrow |D_1 f|(\Omega) + |D_2 f|(\Omega)$ ). Further we show that the *no-crossing BV* condition is equivalent with a seemingly stronger version of the same condition.

## 1. INTRODUCTION

Over the past few years the classification of weak and strong limits of Sobolev diffeomorphisms has attracted a lot of interest for its relevance to variational models of nonlinear elasticity and geometric function theory. The pioneering work in the area was by Iwaniec and Onninen [23] followed by the more recent result of De Philippis and Pratelli [12]. Thanks to these results, the classification of weak and strong Sobolev limits of Sobolev homeomorphisms in the planar setting is now well understood.

In the case that  $p \geq 2$  the authors of [23] utilise the approximation techniques of [21, 22] to prove that it is exactly monotone maps which characterize the weak closure of  $W^{1,p}$  homeomorphisms.

The authors of [12] however, approach the problem using a different technique introduced in [20] for the diffeomorphic approximation of  $W^{1,1}$  homeomorphisms. They prove that the weak and strong closures of  $W^{1,p}$  homeomorphisms coincide for all  $1 \leq p < \infty$  (assuming uniform integrability in the case of  $p = 1$ ). This result was examined in the case that the mappings in question equal identity on the boundary of a square. When  $p < 2$  the closure contains mappings with discontinuities and so monotonicity is too restrictive to characterize the class. On the other hand, they, quite surprisingly, proved that the INV condition of Müller and Spector [25] is satisfied by maps that cannot be the weak Sobolev limits of homeomorphisms (see [12, Section 5.2]) and the situation is not saved even by the restriction of  $J_f > 0$ -a.e. for the limit maps (see [12, Section 5.3]).

---

2010 *Mathematics Subject Classification.* Primary 46E35; Secondary 30E10, 58E20.

*Key words and phrases.* No-crossing condition, homeomorphisms, BV mappings, Strict closure.

The first author was supported by the grant GACR 20-19018Y.

Therefore, it was necessary to introduce a new condition which the authors called the *no-crossing* condition. In essence a Sobolev mapping belongs to the (weak or strong) closure of homeomorphisms if and only if its restriction to ‘almost any’ grid of horizontal and vertical lines can be uniformly approximated by continuous injective maps. Let us note it was proved in [9] that the condition that the injective approximation of the map on a grid cannot be weakened to the condition where the map can be uniformly approximated by injective maps on a single injective Lipschitz curve. One advantage of the no-crossing condition is that it does not require the map to be defined everywhere but only up to a  $\mathcal{H}^1$ -negligible set. The necessity of this was demonstrated in [12] where they construct limits of planar Sobolev homeomorphisms presenting cavitations. Therefore, studying the closure of planar homeomorphisms in the  $BV$  setting opens itself as a natural question. As was shown in [8] there are maps in the limit class which exhibit more complicated discontinuities, like fractures.

The continuity and invertibility properties of candidates for energy minimizing deformations in elasticity theory were studied in the pioneering works of Ball [5, 6]. Thanks to the concept of non-interpenetration of matter, it is natural to minimise in classes of homeomorphisms which satisfy certain boundary conditions. In some models, however, the energy functional does not guarantee the existence of a homeomorphic solution and in these cases one is led to find a larger class which contains the limit maps. While doing so, however, one wants the class to be restrictive enough to demand its maps exhibit required behaviour of elastic deformations; for example, the non-interpenetration of the material. One such condition is the INV condition of Müller and Spector from [25]. This is a kind of monotonicity condition which allows for cavitation, a phenomenon which has been observed experimentally in deformed elastic materials (see [14, Figure 4]). Ball proposed generalizing the mathematical model so as to allow for both cavitations and fracture-type singularities. This is motivated by several experimental observations, e.g. on ductile fracture of titanium alloys [26, 27] or in [31, 15] on elastomers, where the experiments suggest that the strains at the cavity surface produced during cavitation are so large that fracture occurs at the same time.

One approach to the question of modelling deformations allowing for discontinuities was studied by Henao and Mora Corral by introducing a term in the energy functional which penalizes new surface created by the deformations. In the series of results [16, 17, 18, 19] they prove that the minimizers are one-to-one almost everywhere and can exhibit fractures.

Another approach was proposed by the authors and Hencl in [8]. Our motivation was fuelled by interesting  $BV$  relaxation results obtained by Kristensen and Rindler [24] and Rindler and Shaw [30] for Dirichlet-type boundary conditions, and by Baía, Krömer and Kružík [3] for Neumann-type conditions. Further motivation came from recent results in [28] and [29] by the third author and Pratelli on the *strict* and *area-strict* approximation of  $BV$  homeomorphisms by diffeomorphisms. We remind the reader that a sequence  $f_k : \Omega \rightarrow \mathbb{R}^n$  of  $BV$  functions converges *strictly* to  $f \in BV(\Omega, \mathbb{R}^n)$  if  $f_k \rightarrow f$  in  $L^1(\Omega, \mathbb{R}^n)$  and  $|Df_k|(\Omega) \rightarrow |Df|(\Omega)$ . The sequence converges *area-strictly* if it converges strictly and it is possible to decompose  $Df_k$  as the sum of two measures  $\mu_k + \nu_k$  such that  $|\mu_k - D^a f|(\Omega) \rightarrow 0$  and  $|\nu_k|(\Omega) \rightarrow |D^s f|(\Omega)$ , where  $D^a f$  and  $D^s f$  denote the absolutely and the singular part of  $Df$  respectively.

In this paper we study classes of area-strict limits and the so-dubbed ‘m-strict’ limits of  $BV$  homeomorphisms. We say that a sequence of  $BV$  maps  $f_k$  converges m-strictly to  $f$  on  $\Omega$  if  $f_k \rightarrow f$  in  $L^1(\Omega)$  and

$$(1.1) \quad |D_1 f_k|(\Omega) + |D_2 f_k|(\Omega) \rightarrow |D_1 f|(\Omega) + |D_2 f|(\Omega).$$

It is obvious that area-strict convergence implies strict convergence which in turn implies m-strict convergence. It is not hard to construct examples which show that each convergence is sharply weaker than the previous.

In [8] it was shown that strict limits of planar  $BV$  homeomorphisms can exhibit cavities and fractures but still preserve a sort of monotonicity property. Although there is not enough topological information to meaningfully generalize the INV condition, a type of topological image can be defined and it was shown in [8] that the intersection of this topological image of two disjoint sets has zero measure. As remarked in [10], the proof of this fact in [8] used only the m-strict convergence (not strict convergence).

It was proved in [10] that the m-strict limits of  $BV$  homeomorphisms satisfy the  $NCBV$  condition (see Definition 3.8 for the precise definition and the following paragraph for an intuition, it is a generalization of the NC condition of [12]) and any map satisfying the  $NCBV$  condition can be approximated m-strictly by  $BV$  homeomorphisms. Given these facts, there were grounds to claim in [10] that the class of m-strict limits of homeomorphisms is an appropriate class within which to conduct  $BV$  relaxations which are physically relevant in elasticity. On the other hand, as in the Sobolev case strong and weak closure of diffeomorphisms coincide, also in  $BV$  the counterpart of strong and weak closure should behave the same way. The strong closure of diffeomorphisms does not say anything meaningful for elasticity, hence the strong counterpart has to be understood in this weaker sense of area-strict convergence. Roughly speaking, area-strict convergence is equivalent to strong convergence in those portions of the domain where the singular part of the derivative is small, while is equivalent to strict convergence where the singular part is concentrated. These classes are also an attractive option since they are much more widely used and studied. We prove that in fact, somewhat surprisingly the classes coincide (although we do not claim that an m-strict converging sequence must also converge strictly). An extra advantage of this result is that being an m-strict limit may be an easier condition to check since one can restrict oneself to behaviour on almost all lines reducing the dimension of the problem.

We shall now endeavour to give the reader a rough idea of what the  $NCBV$  condition is. We consider a  $BV$  map  $f$  defined on the unit square. On almost every horizontal and vertical line the restriction of a  $BV$  map is a  $BV$  map from the line. It is not hard to prove that for  $\mathcal{L}^2$ -almost every point  $X$  of the domain the restriction of the  $BV$  map to the horizontal and vertical line intersecting  $X$  is continuous at  $X$ . Call  $\Gamma$  the union of a finite number of horizontal and vertical lines on each of which the restriction of  $f$  is  $BV$  and  $f_{|\Gamma}$  is continuous at the intersection point of any two lines. There is a countable number of points outside of which  $f_{|\Gamma}$  is continuous and each of these points is a jump and lies on exactly one horizontal or vertical segment. We declare the topological image of each point  $X \in \Gamma$  to be the segment connecting the one-sided limits of  $f_{|\Gamma}$  at  $X$ . Then the topological image of each of the horizontal and vertical lines is now a Lipschitz curve. We parametrize this Lipschitz curve by some

Lipschitz mapping from the grid  $\Gamma$  and call this map the geometric representative of  $f$ . We say that  $f$  satisfies the *NCBV* condition if for every such  $\Gamma$  the geometric representative of  $f|_{\Gamma}$  can be uniformly approximated by a continuous injective map. See section 3.1 for our concept of restricting a *BV* map onto a grid and the precise definition of the *NCBV* condition. The *NCBV*<sup>+</sup> condition is essentially the same as the *NCBV* condition but the lines do not have to be only vertical and horizontal, instead,  $\Gamma$  is the union of piecewise linear injective paths that intersect each other at most once and the intersection point of any pair is distinct from another pair.

In the following we denote  $Q(c, r)$  the square centered at  $c$  with side  $2r$ , in particular  $Q(0, 1)$  denotes  $(-1, 1)^2$ . Our main result is the following:

**Theorem 1.1.** *Let  $f \in BV(Q(0, 1); Q(0, 1))$  and  $f(x, y) = (x, y)$  for every  $(x, y) \in \partial Q(0, 1)$ . Then the following conditions are equivalent*

- (1)  *$f$  satisfies the *NCBV* condition,*
- (2)  *$f$  satisfies the *NCBV*<sup>+</sup> condition,*
- (3) *there exists a sequence  $f_k \in BV(Q(0, 1), Q(0, 1))$  of diffeomorphisms with  $f_k = \text{id}$  on  $\partial Q(0, 1)$  converging to  $f$  area-strictly,*
- (4) *there exists a sequence  $f_k \in BV(Q(0, 1), Q(0, 1))$  of diffeomorphisms with  $f_k = \text{id}$  on  $\partial Q(0, 1)$  converging to  $f$  weakly in *BV* and  $|D_1 f_k|(Q(0, 1)) + |D_2 f_k|(Q(0, 1)) \rightarrow |D_1 f|(Q(0, 1)) + |D_2 f|(Q(0, 1))$ .*

**1.1. Overview of the proof.** Since (3) implies (4) is obvious and since (4) implies (1) has been proved in [10], there are two implications to prove, i.e. (1) implies (2) and (2) implies (3).

To prove (1) implies (2) we assume that we have a good non-straight grid for  $f$  called  $\Gamma$ . Our aim is to create a system of horizontal and vertical segments on which  $f$  remains close to  $f$  on  $\Gamma$ . The non-straight grid  $\Gamma$  is the union of  $\gamma_i([0, 1])$ , where  $\gamma_i$  are piecewise linear, injective and continuous. We call  $X_{i,j}$  the unique element of  $\gamma_i([0, 1]) \cap \gamma_j([0, 1])$  whenever the intersection is non-empty.

Our first step is to eliminate the intersection points  $X_{i,j}$  of  $\Gamma$ . These points are chosen so that  $\lim_{r \rightarrow 0} r^{-1} |Df|(Q(X_{i,j}, r)) = 0$ . We choose a small  $r > 0$  and 4 disjoint spirals (see Figure 1) with which we replace part of each segment of  $\Gamma$  ending at  $X_{i,j}$ . By choosing the lines of the spiral carefully and by choosing the  $r$  small enough we guarantee that the oscillation of  $f|_{\Gamma \cap Q(X_{i,j}, r)}$  and the oscillation of  $f$  on the spiral are both much smaller than some  $\sigma$ . Then any injective approximation of  $f$  on the spiral with error  $\sigma/2$  is also an injective approximation of  $f$  on  $\Gamma \cap Q(X_{i,j}, r)$  with error bounded by  $\sigma$ . This step is Proposition 3.17.

For all  $X \in \Gamma$  except for a finite number of points where there are ‘large’ jumps of  $f|_{\Gamma}$  we have  $\limsup_{r \rightarrow 0} r^{-1} |Df|(Q(X, r)) \ll \sigma$ . We can then find a rectangle  $R_X$  such that the oscillation of  $f$  on  $\Gamma \cap R_X$  and the oscillation of  $f$  on  $\partial R_X$  are both much smaller than  $\sigma$ . Then any injective approximation of  $f$  on an appropriate part of  $\partial R_X$  with error  $\sigma/2$  is also an injective approximation of  $f$  on  $\Gamma \cap R_X$  with error bounded by  $\sigma$ . This step is conducted in the proof of Theorem 3.18 by applying Lemma 3.16.

It then remains to deal with the finite number of ‘large’ (say larger than  $\frac{\sigma}{40}$ ) jumps of  $f|_{\Gamma}$ . But these jump points are Lebesgue points of the polar decomposition of  $D^j f$  and so it is not hard to find a point very close by which has a very similar jump and construct a piece-wise horizontal and vertical path with endpoints on  $\Gamma$  and doing

basically just the same jump as  $f$ . See Figure 2 for a depiction of how we do this. This step is conducted in the proof of Theorem 3.18. After having done the above 3 steps we have an admissible set of horizontal and vertical segments from which we generate a good straight grid  $\tilde{\Gamma}$ .

At this point it is easy to generate an injective approximation of the geometric representative of  $f$  with error  $\sigma$  using the *NCBV* condition with error sufficiently smaller than  $\sigma$  and finding an appropriate correspondence of a subset of  $\tilde{\Gamma}$  and  $\Gamma$ .

The proof that (2) implies (3) is quite involved. Some techniques developed in [10] are also implemented here. In comparison with the approximation in [10] we have to approximate  $D^a f$  in  $L^1$  and in order to do so we have to use the techniques developed in [20] (and following papers like [7]). Further the m-strict convergence in [10] is strictly weaker than even strict convergence and therefore it is necessary to find a better way to approximate the singular part of the derivative. The new extension result needed for this approximation has been developed in [11]. It is a rotated version of the main result in [28], not only for rectangles but also for convex polygons. Let us now give a short sketch of how we use these results for the proof.

First we find a small set supporting the vast majority of  $|D^s f|$ . By dividing this set into very small squares and using a Lebesgue point-type argument for the polar decomposition of  $D^s f$  we can apply Theorem 2.3 which is taken from [11]. Thanks to Alberti's rank-one theorem and the Lebesgue point-type estimate, for each square  $Q_i$  we have a unit vector  $v_i$  such that  $D^s f|_{Q_i}$  is very close to  $u_i \otimes v_i |D^s f|_{Q_i}|$  for some unit vector  $u_i$ . This enables us to make estimates like  $|Df|(Q_i) \leq (1 + \varepsilon) |D^s f|(Q_i) \leq (1 + 2\varepsilon) |\langle D^s f, v_i \rangle|(Q_i)$  which means that Theorem 2.3 gives an estimate of the energy of our approximation on  $Q_i$  by  $(1 + C\varepsilon) |D^s f|(Q_i)$ .

From now on we work on the rest of  $Q(0, 1)$  where the energy of  $|D^s f|$  is already very small. We separate this set into very small squares and use a Lebesgue point-type argument for the derivatives to find 4 sets, call them  $A_1, A_2, A_3, A_4$ . We have that  $|Df|(A_1)$  is very small because on every square in  $A_1$  we have that  $|Df|(Q_i) \leq \varepsilon \mathcal{L}^2(Q_i)$ . The set  $A_2$  is very small and so  $|Df|(A_2)$  is also very small. The set  $A_3$  is made of squares where  $f$  is very close to a nice affine mapping and  $A_4$  is made of squares where  $f$  is very close to an affine mapping with zero Jacobian. We slightly shift the vertexes of the squares so that the behaviour of  $f$  on the boundary of the resulting convex quadrilaterals is similar to the behaviour of  $f$  on the quadrilateral (by the BV on lines characterization). This allows us to use the extension theorems from [20] on quadrilaterals in  $A_1$  and  $A_2$  (Theorem 2.1), and  $A_3$  (Theorem 2.2) while in  $A_4$  it is enough to do a straight forward triangularisation. There is then no problem in proving that the derivative of our approximation is close to  $D^a f$  in  $L^1$ .

## 2. PRELIMINARIES

**2.1. Extension Theorems.** The following is from [20, Theorem 2.1].

**Theorem 2.1.** *Let  $\varphi : \partial Q \rightarrow \mathbb{R}^2$  be a piecewise linear and one-to-one function. There is a finitely piecewise affine homeomorphism  $g : Q \rightarrow \mathbb{R}^2$  such that  $g = \varphi$  on  $\partial Q$ , and*

$$\int_Q |Dg| \leq C \operatorname{diam} Q \int_{\partial Q} |D_\tau \varphi|.$$

The following theorem is [7, Theorem 3.7]. The question is how to approximate a map which is close to a degenerate linear map  $\Phi$  (up to a rotation in the pre-image we may assume that  $\Phi = \begin{pmatrix} d & 0 \\ 0 & 0 \end{pmatrix}$ ).

**Theorem 2.2.** *Let  $d > \delta > 0$ , let  $r_0 \in (0, 1)$  and let  $\mathcal{Q}$  be a convex set and the image of  $[0, r_0]^2$  in a 2-bi-Lipschitz mapping which is equal to an affine mapping on  $\text{co}\{(0, 0), (0, r_0), (r_0, 0)\}$  and  $\text{co}\{(r_0, r_0), (0, r_0), (r_0, 0)\}$ . Then for every  $\varphi : \partial\mathcal{Q} \rightarrow \mathbb{R}^2$  finitely piece-wise linear and one-to-one mapping with*

$$(2.1) \quad \int_{\partial\mathcal{Q}} \left| D_\tau \varphi(t) - \begin{pmatrix} d & 0 \\ 0 & 0 \end{pmatrix} \tau \right| d\mathcal{H}^1(t) < \delta r_0,$$

and  $\|D_\tau \varphi\|_{L^\infty(\partial\mathcal{Q})} \leq d + 2\delta$ , there exists a finitely piece-wise affine homeomorphism  $g : \mathcal{Q} \rightarrow \mathbb{R}^2$  such that  $g = \varphi$  on  $\partial\mathcal{Q}$  and

$$(2.2) \quad \left\| Dg(x) - \begin{pmatrix} d & 0 \\ 0 & 0 \end{pmatrix} \right\|_{L^1(\mathcal{Q})} < C\delta r_0^2.$$

Let  $\mathcal{Q}_i$  be a convex quadrilateral. For a unit vector  $v_i$  we denote  $\pi_{v_i}(x) = v_i^\perp \langle x, v_i^\perp \rangle = x - v_i \langle x, v_i \rangle$  and  $\pi_{v_i^\perp}(x) = v_i \langle x, v_i \rangle = x - v_i^\perp \langle x, v_i^\perp \rangle$ . For each  $X \in \pi_{v_i}(\text{int } \mathcal{Q}_i)$  there are exactly two distinct points  $X_*, X^*$  in  $\partial\mathcal{Q}_i$  such that  $\pi_{v_i}(X_*) = \pi_{v_i}(X^*) = X$  and similarly  $\pi_{v_i^\perp}(Z_*) = \pi_{v_i^\perp}(Z^*) = Z \in \pi_{v_i^\perp}(\mathcal{Q}_i)$ . To be specific we denote  $X_*, X^*$  so that  $\langle X_*, v_i \rangle < \langle X^*, v_i \rangle$  and  $Z_*, Z^*$  so that  $\langle Z_*, v_i^\perp \rangle < \langle Z^*, v_i^\perp \rangle$ . Given  $\varphi : \partial\mathcal{Q} \rightarrow \mathbb{R}^2$  a continuous injective map, we call  $\mathcal{P}$  the bounded connected component of  $\mathbb{R}^2 \setminus \varphi(\partial\mathcal{Q})$  identified by the Jordan curve  $\varphi(\partial\mathcal{Q})$ . For any pair of points  $\mathbf{A}, \mathbf{B} \in \mathcal{P}$  we denote by

$$(2.3) \quad d_{\mathcal{P}}(\mathbf{A}, \mathbf{B}) \text{ the geodesic distance between } \mathbf{A} \text{ and } \mathbf{B} \text{ inside } \mathcal{P}.$$

The following is [11, Theorem 1.2].

**Theorem 2.3.** *Let  $\theta \in [0, 2\pi)$  be fixed and let  $v_\theta = (\cos \theta, \sin \theta)$ ,  $\mathcal{Q} \subset \mathbb{R}^2$  be a convex polygon and  $\varphi : \partial\mathcal{Q} \rightarrow \mathbb{R}^2$  be a continuous piecewise linear injective map. Then for every  $\varepsilon > 0$  there exists a finitely piecewise affine homeomorphism  $g : \mathcal{Q} \rightarrow \mathbb{R}^2$  extending  $\varphi$ , such that*

$$(2.4) \quad \begin{aligned} |\langle Dg, v_\theta \rangle|(\mathcal{Q}) &\leq \int_{\pi_{v_i}(\mathcal{Q})} d_{\mathcal{P}}(\varphi(X^*), \varphi(X_*)) d\mathcal{H}^1(X) + \varepsilon \text{ and} \\ |\langle Dg, v_\theta^\perp \rangle|(\mathcal{Q}) &\leq \int_{\pi_{v_i^\perp}(\mathcal{Q})} d_{\mathcal{P}}(\varphi(Z^*), \varphi(Z_*)) d\mathcal{H}^1(Z) + \varepsilon. \end{aligned}$$

**2.2. Properties of BV maps.** In the following we repeat some necessary results from the structure theory of BV mappings.

The following definition is [2, Definition 2.40]

**Definition 2.4** (Tangent measures). We define the set of tangent measures to a  $\mathbb{R}^m$  valued Radon measure  $\mu$  on  $\mathbb{R}^n$  at  $x \in \mathbb{R}^n$  as the set of all finite Radon measures on  $B(0, 1)$ , which are weak\* limits of

$$\frac{\mu(B(x, r))}{|\mu(B(x, r))|}$$

as  $r \rightarrow 0$ .

The following definition is [2, Definition 2.79].

**Definition 2.5** (Approximate tangent spaces). Let  $x \in G \subset \mathbb{R}^n$  be open and let  $\mu \in \mathcal{M}(G, \mathbb{R}^m)$ . We say that  $V$ , a  $k$ -dimensional vector subspace of  $\mathbb{R}^n$ , is the approximate tangent space of  $\mu$  with multiplicity  $\lambda \in \mathbb{R}$  if

$$\frac{\mu(B(x, r))}{r^k} \xrightarrow{*} \lambda \mathcal{H}_{|V}^k$$

as  $r \rightarrow 0$  and we denote this as  $\text{Tan}^k(\mu, x) := \lambda \mathcal{H}_{|V}^k$ .

Let us emphasize that through out the paper we use the standard notation  $\mathcal{H}^k$  to denote the  $k$ -dimensional Hausdorff measure. The following is [2, Theorem 2.81, (b)]

**Theorem 2.6** (Strict convergence of approximate tangent spaces). *Let  $\mu$  be a  $\mathbb{R}^m$  valued Radon measure on  $\mathbb{R}^n$  and assume that  $\limsup_{r \rightarrow 0^+} r^{-k} |\mu|(B(x, r)) < \infty$ . Call  $f$  the function of the polar decomposition of  $\mu$  i.e.  $\mu = f|\mu|$  and assume further that  $x$  is a Lebesgue point of  $f$  with respect to  $|\mu|$ . Then*

$$\nu = \text{Tan}^k(\mu, x) \text{ if and only if } |\nu| = \text{Tan}^k(|\mu|, x).$$

The following definition is [2, Definition 3.67]

**Definition 2.7** (Approximate jump points). Let  $f \in L_{\text{loc}}^1(G, \mathbb{R}^m)$  and let  $x \in G \subset \mathbb{R}^n$  be open. We say that  $x$  is an approximate jump point of  $f$  if there exist  $a \neq b \in \mathbb{R}^m$  and a  $v \in \mathbb{R}^n$ ,  $|v| = 1$  such that

$$\lim_{r \rightarrow 0} \int_{B(x, r, v, +)} |f(y) - a| d\mathcal{L}^n(y) = 0 \quad \text{and} \quad \lim_{r \rightarrow 0} \int_{B(x, r, v, -)} |f(y) - b| d\mathcal{L}^n(y) = 0$$

where  $B(x, r, v, +) := \{y \in B(x, r); \langle y - x, v \rangle > 0\}$  and  $B(x, r, v, -) := \{y \in B(x, r); \langle y - x, v \rangle < 0\}$ .

We denote  $f^+(x) := a$  and  $f^-(x) := b$  and up to fixing the orientation of the vector  $v$  the notation is unique. We denote the set of all jump points of  $f$  as  $J_f$ .

The following is the Federer-Vol'pert theorem as in [2, Theorem 3.78]

**Theorem 2.8** (Federer-Vol'pert). *For any  $f \in \text{BV}(\Omega, \mathbb{R}^m)$  the discontinuity set  $S_f$  is countably  $\mathcal{H}^{n-1}$ -rectifiable and  $\mathcal{H}^{n-1}(S_f \setminus J_f) = 0$ . Moreover  $Df|_{J_f} = (f^+ - f^-) \otimes v \mathcal{H}_{|J_f}^{n-1}$  and*

$$\begin{aligned} \text{Tan}^{n-1}(J_f, x) &= (v(x))^\perp \\ \text{Tan}^{n-1}(|Df|_{|J_f}, x) &= |f^+(x) - f^-(x)| \mathcal{H}_{|(v(x))^\perp}^{n-1} \end{aligned}$$

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

The following is Alberti's Rank one theorem from [1] as in [2, Theorem 3.94]

**Theorem 2.9** (Rank one). *Let  $f \in \text{BV}(\Omega, \mathbb{R}^m)$  and  $g$  is the function such that  $Df = g|Df|$  then there exist unit vectors  $u, v$  such that  $g(x) = u(x) \otimes v(x)$  for  $|D^s f|$ -almost every  $x \in \Omega$ .*

**Lemma 2.10.** *Let  $f \in \text{BV}(\Omega)$  and let  $A \subset \Omega$  be open then  $|\langle Df, v \rangle|(A) \leq |Df|(A)$  for any  $|v| = 1$ .*

*Proof.* By [13, Theorem 5.2] we have a sequence of smooth functions  $f_k$  strictly converging to  $f$  in  $BV(\Omega)$ . For every unit vector  $v$  it is obvious that  $|\langle Df_k(x), v \rangle| \leq |Df_k(x)|$  everywhere in  $\Omega$ . Now integrating over  $A$  and using the lower semi continuity of the variation we conclude.  $\square$

### 3. THE NCBV PROPERTY FOR $BV$ MAPS

**3.1.  $BV$  on grids.** In [12] the authors introduced a property called the  $NC$  condition that characterises the limits of  $W^{1,p}$  homeomorphisms from  $Q(0, 1)$  onto  $Q(0, 1)$  equalling the identity on the boundary. In [10] we generalized this condition for  $BV$  maps, which can even fail to be continuous on a 1-rectifiable set. The essence of these conditions is that a certain map can be approximated uniformly by a continuous injective map with error arbitrarily small. Our approach requires choosing a finite number of points where  $f$  is continuous which is in the following proposition.

**Proposition 3.1.** *Let  $f \in BV(Q(0, 1), \mathbb{R}^2)$  and let  $v_1, v_2, v_3, v_4 \in \mathbb{R}^2$  with  $|v_i| = 1$ . Then for almost every  $(x, y) \in Q(0, 1)$  it holds that*

$$(3.1) \quad f|_{(x,y)+\cup_i v_i \mathbb{R}} \text{ is continuous at } (x, y)$$

and

$$(3.2) \quad \lim_{r \rightarrow 0} r^{-1} |Df|(Q((x, y), r)) = 0.$$

*Proof.* The claim (3.1) follows from [2, Theorem 3.107] and (3.2) follows from [2, Proposition 3.92].  $\square$

**Corollary 3.2.** *Let  $f \in BV(Q(0, 1), \mathbb{R}^2)$ , There exists a set  $N_1, N_2 \subset [-1, 1]$ ,  $\mathcal{L}^1(N_1) = \mathcal{L}^1(N_2) = 0$  such that*

- (1)  $f|_{\{x\} \times [-1, 1]}$  is  $BV$  on  $\{x\} \times [-1, 1]$ ,
- (2)  $f|_{[-1, 1] \times \{y\}}$  is  $BV$  on  $[-1, 1] \times \{y\}$  and
- (3)  $f|_{\{x\} \times [-1, 1] \cup [-1, 1] \times \{y\}}$  is continuous at  $(x, y)$

for  $\mathcal{L}^1$ -almost every  $y \in [-1, 1]^2$  if  $x \in [-1, 1] \setminus N_1$  and for  $\mathcal{L}^1$ -almost every  $x \in [-1, 1]^2$  if  $y \in [-1, 1] \setminus N_2$ .

*Proof.* The claim follows from Proposition 3.1 and the Fubini theorem.  $\square$

Let

$$(3.3) \quad (f^+(\cdot) - f^-(\cdot)) \otimes v(\cdot) \mathcal{H}^1_{J_f} = D^j f$$

be the standard decomposition of  $D^j f$  mentioned in Theorem 2.8. Also by Theorem 2.8 for  $\mathcal{H}^1$  almost every  $x \in J_f$  it holds that

$$(3.4) \quad \frac{1}{r} \int_{B(x,r) \cap J_f} |(f^+(y) - f^-(y)) \otimes v(y) - (f^+(x) - f^-(x)) \otimes v(x)| d\mathcal{H}^1_{J_f}(y) \rightarrow 0$$

and we call these points Lebesgue points of  $(f^+(\cdot) - f^-(\cdot)) \otimes v(\cdot)$ .

**Definition 3.3** (Admissible curves for  $f$ ). Let  $f \in BV(Q(0, 1), Q(0, 1))$ . Let  $\gamma_i : [0, 1] \rightarrow Q(0, 1) \subset \mathbb{R}^2$ ,  $i = 1, \dots, K$  be finitely piecewise linear mappings with  $\gamma_i([0, 1]) \cap \gamma_j([0, 1])$  contains at most one point  $\{X_{i,j}\}$  for all  $i \neq j$ . Let  $\Gamma = \bigcup_{i=1}^K \gamma_i([0, 1]) \subset Q(0, 1)$ . We call  $\Gamma$  admissible for  $f$  if

- (1)  $f \circ \gamma_i \in BV((0, 1), \mathbb{R}^2)$  for all  $i$ ,



- (2)  $f|_\Gamma$  is continuous at  $\gamma_i(s_{i,k})$  at every point  $s_{i,k}$ , where  $\{s_{i,k}\}$  is the finite set of endpoints of intervals on which  $\gamma_i$  is linear,
- (3)  $f|_\Gamma$  is continuous at each  $X_{i,j}$  for every  $i, j$  such that  $\gamma_i([0, 1]) \cap \gamma_j([0, 1]) \neq \emptyset$ ,
- (4)  $\lim_{r \rightarrow 0} r^{-1} |Df|(Q(\gamma_i(s_{i,k}), r)) = 0$  for each  $s_{i,k}$ ,
- (5)  $\lim_{r \rightarrow 0} r^{-1} |Df|(Q(X_{i,j}, r)) = 0$  for each  $X_{i,j}$ .

Note that by Proposition 3.1 ‘almost every’ polyline is admissible.

**Definition 3.4** (Good straight grid). Let  $f \in BV(Q(0, 1), Q(0, 1))$  and let  $\Gamma = (\bigcup_{i=1}^K \{x_i\} \times [-1, 1]) \cup (\bigcup_{j=1}^K [-1, 1] \times \{y_j\})$  be admissible for  $f$  in the sense of Definition 3.3. We call  $\Gamma$  a good straight grid for  $f$  if

- (1) every point of  $\Gamma \cap J_f$  is a Lebesgue point of  $(f^+(\cdot) - f^-(\cdot)) \otimes v(\cdot)$  in the sense of (3.4),
- (2)  $\langle (1, 0), v(x, y_j) \rangle \neq 0$ ,  $\langle (0, 1), v(x_i, y) \rangle \neq 0$  for all  $x, y \in [-1, 1]$  such that  $(x_i, y), (x, y_j) \in \Gamma \cap J_f$ , where  $v(x, y)$  is the vector from Theorem 2.9.

**Definition 3.5.** Let  $f \in BV(Q(0, 1), Q(0, 1))$  and let  $\Gamma \subset Q(0, 1)$  be the union of a finite number of horizontal and vertical segments. Let  $\tilde{\Gamma}$  be the smallest straight grid such that  $\tilde{\Gamma} \supset \Gamma$ . Then we say that  $\tilde{\Gamma}$  is the straight grid generated by  $\Gamma$ .

Now we define the object  $\Gamma$  which we refer to as a ‘non-straight grid’.

**Definition 3.6** (Good non-straight grid). Let  $f \in BV(Q(0, 1), Q(0, 1))$ . Let  $\gamma_i : [0, 1] \rightarrow Q(0, 1) \subset \mathbb{R}^2$ ,  $i = 1, \dots, K$  be finitely piecewise linear mappings with  $\gamma_i([0, 1]) \cap \gamma_j([0, 1])$  contains at most one point  $\{X_{i,j}\}$  if  $i \neq j$ . Let  $\Gamma = \bigcup_{i=1}^K \gamma_i([0, 1])$  be admissible for  $f$ . We call  $\Gamma$  a good non-straight grid for  $f$  if

- (1) every point of  $\Gamma \cap J_f$  is a Lebesgue point of  $(f^+(\cdot) - f^-(\cdot)) \otimes v(\cdot)$  with respect to  $\mathcal{H}_{J_f}^1$ ,
- (2) the derivative  $\gamma_i'(t)$  exists and  $\langle \gamma_i'(t), v(\gamma_i(t)) \rangle \neq 0$  whenever  $\gamma_i(t) \in J_f$  where  $v(x, y)$  is the vector from Theorem 2.9.

**Remark 3.7.** A simple affine change of variables together with Corollary 3.2 proves the following. For every pair of distinct directions  $v_1, v_2 \in \mathbb{R}^2$ ,  $|v_j| = 1$  almost every line  $L_1$  parallel to  $v_1$  yields an admissible  $\Gamma = L_1 \cup L_2$  for almost every line  $L_2$  parallel to  $v_2$ . Informally we can say that ‘almost every’ non-straight grid is admissible for  $f$ . In Theorem 4.3 we show that ‘almost every’ non-straight grid is good for  $f$ .

In this paper we often restrict a planar  $BV$  map onto lines or good grids (both straight and non-straight). Since, for almost every line, the restriction of a  $BV$  map is one-dimensional  $BV$  on the line, it stands to reason, for a good choice of grid  $\Gamma$ , that  $f|_\Gamma$  is also in  $BV$ . On the other hand, the space  $BV(\Gamma)$  is not very standard so we now explain what we mean by this.

Let  $f : Q(0, 1) \rightarrow \mathbb{R}^2$  and let  $\Gamma$  be a good non-straight grid. We say that  $f|_\Gamma$  is  $BV$  on  $\Gamma$  if

- (1)  $f \circ \gamma_i \in BV([0, 1], \mathbb{R}^2)$
- (2)  $f|_\Gamma$  is continuous at each  $X_{i,j}$ .

By  $D_\tau f$  we denote the measure on  $\Gamma$  given by  $\sum_{i=1}^K D_{\frac{\gamma_i'}{|\gamma_i'|}} f|_{\gamma_i([0,1])}$ . The condition (2) above ensures that  $\{X_{i,j} : 1 \leq i < j \leq K\}$  is a negligible set in  $|D_\tau f|_\Gamma$ . This means

that we can interpret  $D_\tau$  as the distributional derivative tangential to  $\Gamma$  outside the set  $\{X_{i,j} : 1 \leq i < j \leq K\} \cup \{\gamma_i(s_{i,m}) : 1 \leq i \leq K\}$ , where the points  $s_{i,m}$  are those defined in Definition 3.3.

If  $f \in BV(\Gamma)$  and  $|D_\tau f|_\Gamma$  is absolutely continuous with respect to  $\mathcal{H}^1$  then we say that  $f|_\Gamma$  is in  $W^{1,1}(\Gamma)$ .

As a matter of convention, when we write  $\partial_\tau f|_\Gamma(X)$  we refer to the classical partial derivative of the mapping  $f$  in the direction tangential to  $\Gamma$  at  $X$ . We avoid writing this at points  $X_{i,j}$  to avoid ambiguity. It is known for  $f|_\Gamma$  in  $BV$  on  $\Gamma$  (in the sense described above) that  $\partial_\tau f|_\Gamma$  exists almost everywhere and the absolutely continuous part of  $D_\tau f|_\Gamma$  can be represented by  $\partial_\tau f|_\Gamma \mathcal{H}^1|_\Gamma$ .

Throughout the paper  $\nabla f(x)$  denotes the approximative derivative of  $f$  at  $x$  which for  $f \in BV(Q(0,1), \mathbb{R}^2)$  exists  $\mathcal{L}^2$  almost everywhere for the correct representative.

Let  $\gamma : [0,1] \rightarrow Q(0,1)$  be a bi-Lipschitz parametrisation of a curve such that  $f \circ \gamma \in BV(Q(0,1), \mathbb{R}^2)$ . We define  $h : \gamma([0,1]) \rightarrow \mathbb{R}^2$ , the geometric representative of  $f$  on  $\gamma([0,1])$  as explained below. For each  $t$  we define

$$\begin{aligned} Y(t) &= \lim_{s \rightarrow t^-} f \circ \gamma(s), & Z(t) &= \lim_{s \rightarrow t^+} f \circ \gamma(s), \\ l(t) &= |Df \circ \gamma|([0,t]), & L(t) &= |Df \circ \gamma|([0,t]). \end{aligned}$$

We define the a curve  $\tilde{h} : [0, 1 + L(1)] \rightarrow \mathbb{R}^2$  as the constant speed parametrization of the segment  $[Y(t)Z(t)]$  from  $[l(t) + t, L(t) + t]$ . Finally, we define

$$(3.5) \quad h(x) = \tilde{h}([1 + L(1)]\gamma^{-1}(x))$$

on  $\gamma([0,1])$ .

Let  $\Gamma$  be a good non-straight grid for  $f$  then we define  $h$ , the geometric representative of  $f$  on  $\Gamma$ , as follows. For each  $1 \leq i \leq K$  we find finitely many intervals  $[s_{i,m}, s_{i,m+1}]$  covering  $[0,1]$  such that  $\gamma_i$  is linear on each  $[s_{i,m}, s_{i,m+1}]$ ,  $f|_\Gamma$  is continuous at each  $\gamma_i(s_{i,m})$  and for each  $X_{i,j}$  there is an  $m$  such that  $\gamma(s_{i,m}) = X_{i,j}$ . We define the map  $h$  on each segment  $[\gamma_i(s_{i,m})\gamma_i(s_{i,m+1})]$  as in (3.5). This gives us a map  $h$  defined on the whole of  $\Gamma$ . When we refer to a segment of  $\Gamma$  we mean one of the segments  $[\gamma_i(s_{i,m})\gamma_i(s_{i,m+1})]$ .

**Definition 3.8** (No-Crossing  $BV$  condition). Let  $f \in BV(Q(0,1), \mathbb{R}^2)$  with  $f(x) = x$  on  $\partial Q(0,1)$ . We say that  $f$  satisfies the  $NCBV$  condition if for every  $\sigma > 0$  and for every good straight grid  $\Gamma$  for  $f$ , there exists a continuous injective map  $H_\sigma : \Gamma \rightarrow \mathbb{R}^2$  such that  $|h(x) - H_\sigma(x)| < \sigma$  for all  $x \in \Gamma$ , where  $h$  is the geometric representative<sup>1</sup> of  $f$  on  $\Gamma$ .

**Definition 3.9** (No-Crossing  $BV+$  condition). Let  $f \in BV(Q(0,1), \mathbb{R}^2)$  with  $f(x) = x$  on  $\partial Q(0,1)$ . We say that  $f$  satisfies the  $NCBV^+$  condition if for every  $\sigma > 0$  and for every good non-straight grid  $\Gamma$  for  $f$ , there exists a continuous injective map  $H_\sigma : \Gamma \rightarrow \mathbb{R}^2$  such that  $|h(x) - H_\sigma(x)| < \sigma$  for all  $x \in \Gamma$ , where  $h$  is the geometric representative<sup>2</sup> of  $f$  on  $\Gamma$ .

The following lemma was published in [10]. We include its proof here for the convenience of the reader.

<sup>1</sup>See the previous paragraphs for the definition.

<sup>2</sup>See the previous paragraphs for the definition.

**Lemma 3.10.** *Let  $X, Y \in \mathbb{R}^2$ ,  $\varepsilon, \delta \in [0, 1]$  and  $L := |X - Y| > 0$ . Let  $C \in B(X, \delta L)$  and  $D \in B(Y, \delta L)$ . Let  $\eta : [0, 1] \rightarrow \mathbb{R}^2$  be a path (with constant speed parametrisation) joining points  $C$  and  $D$  with arc length  $l(\eta) \leq (1 + \varepsilon)L$ . Let  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  be the constant speed parametrisation of the line segment joining  $X$  and  $Y$ . Then for every  $t \in [0, 1]$   $|\eta(t) - \gamma(t)| \leq \sqrt{3\varepsilon + 12\delta}L$ .*

*Proof.* We assume without loss of generality that  $X$  is the origin and  $Y = (L, 0)$ . Fix a point  $t \in [0, 1]$ . We use the following notation  $l_1 = l(\eta|_{[0,t]})$  and  $l_2 = l(\eta|_{[t,1]})$  and thus  $l_1 + l_2 = l(\eta) \leq (1 + \varepsilon)L$ . At least one of the angles  $\sphericalangle(\eta(t), (tL, 0), (L, 0))$  and  $\sphericalangle((0, 0), (tL, 0), \eta(t))$  is clearly in the range  $[\pi/2, \pi]$ . We give the proof in the case where  $\sphericalangle(\eta(t), (tL, 0), (L, 0)) \in [\pi/2, \pi]$ . The proof in the other case is essentially the same.

We denote the distance between  $\eta(t)$  and  $(tL, 0)$  by  $E$  and the distance between  $(0, 0)$  and  $\eta(t)$  by  $D$ . By the assumption  $\sphericalangle(\eta(t), (tL, 0), (L, 0)) \in [\pi/2, \pi]$  we obtain from law of cosines that

$$(3.6) \quad l_1 + \delta \geq D \geq \sqrt{L^2 t^2 + E^2}.$$

On the other hand, we have  $l_1 = l(\eta)t \leq (1 + \varepsilon)Lt + \delta L$ . Combining this with (3.6) and simplifying results in the estimate

$$E^2 \leq (3\varepsilon + 12\delta)L^2$$

□

**Definition 3.11** (Good arrival grids). Let the mapping  $f \in BV(Q(0, 1); Q(0, 1))$ , let  $\Gamma$  be a good starting grid for  $f$  and let  $\gamma$  be the geometrical representative of  $f$  on  $\Gamma$ . Let  $\kappa > 0$ , let the numbers  $-1 = w_0 < w_1 < w_2 \cdots < w_{N+1} = 1$  and  $-1 = z_0 < z_1 < z_2 \cdots < z_{M+1} = 1$  satisfy  $w_{n+1} - w_n < \kappa$  and  $z_{m+1} - z_m < \kappa$  for every  $0 \leq n \leq N$  and  $0 \leq m \leq M$ . We say that

$$(3.7) \quad \mathcal{G} = \bigcup_{n=0}^{N+1} \{w_n\} \times [-1, 1] \cup \bigcup_{m=0}^{M+1} [-1, 1] \times \{z_m\} \subseteq Q(0, 1)$$

is a *good arrival grid for  $f$  associated with  $\Gamma$  and with side-length  $\kappa$*  if  $P := \gamma^{-1}(\mathcal{G}) \cap \Gamma$  is a finite set and for every  $p \in P$  it holds that

- $p$  is not a cross of the grid  $\Gamma$  (i.e. a point  $(x_i, y_j)$ ),
- $\gamma(p)$  is not a cross of the grid  $\mathcal{G}$  (i.e. a point  $(w_n, z_m)$ ),
- $p$  is a point where the derivative  $\partial_\tau \gamma(p)$  exists and  $\partial_\tau \gamma(p) \neq 0$ ,
- $\partial_\tau \gamma(p)$  is not parallel to the side of  $\mathcal{G}$  containing  $\gamma(p)$ .

An important fact is that good arrival grids always exist. More precisely, we have the following property, whose proof is a simple variant of the proof of [12, Lemma 3.6] and can be found in [9, Lemma 4.4].

**Lemma 3.12.** *Let  $f \in BV(Q(0, 1); Q(0, 1))$  and let  $\Gamma$  be a good starting grid for  $f$  in  $Q(0, 1)$ . Then the geometrical representative  $\gamma$  of  $f$  on  $\Gamma$  is in  $W^{1,1}(\Gamma, Q(0, 1))$ . Moreover, there exists  $\bar{\kappa} = \bar{\kappa}(L) > 0$  such that for any  $0 < \kappa < \bar{\kappa}$  and any  $\Sigma \subset \Gamma$   $\mathcal{H}^1$ -negligible set, there exists a good arrival grid  $\mathcal{G}$  for  $f$  associated with  $\Gamma$ , with side-length  $\kappa$ , and such that  $\gamma^{-1}(\mathcal{G}) \cap \Sigma = \emptyset$ .*

We define the concept of the generalized segment, already introduced in [12], which will be useful throughout the proof of Theorem 1.1.

**Definition 3.13** (generalized segments). Let  $\mathcal{G} \subset Q(0, 1) \subset \mathbb{R}^2$  be a grid (the finite union of horizontal and vertical lines). Let  $R$  be a rectangle of the grid  $\Gamma$  (the closure of a component of  $Q(0, 1) \setminus \Gamma$ ). Let  $X \neq Y$  and  $X, Y \in \partial R \subset \mathcal{G}$ . Given  $\xi > 0$  a small parameter, the *generalized segment*  $[XY]$  between  $X$  and  $Y$  in  $R$  is defined as the standard segment  $[XY]$  if the two points are not in the same side of  $\partial R$ ; otherwise,  $[XY]$  is the union of two segments of the form  $[XM]$  and  $[MB]$  where  $M$  is the point inside  $R$  whose distance from the side containing  $X$  and  $Y$  is  $\xi|X - Y|/2$  and the projection of  $M$  on the segment  $[XY]$  is the mid-point of  $[XY]$ .

The following claim about generalized segments holds.

**Proposition 3.14.** *Let  $R \subset \mathbb{R}^2$  be a rectangle and let  $a, b \in \partial R$ . Let  $S$  be a generalized segment from  $a$  to  $b$  in  $R$  with parameter  $\xi > 0$  and let  $\tilde{S} \subset S$  be a closed and connected subset of  $S$ . Then*

$$\mathcal{H}^1(\tilde{S}) \leq (1 + \xi) \text{diam}(\tilde{S})$$

*Proof.* If  $S$  is a segment the claim is immediate. In fact,  $\mathcal{H}^1(\tilde{S}) \leq \text{diam}(\tilde{S})$ . If  $S$  is the union of 2 segments, then after rotation and translation we can interpret  $S$  as the graph of the function  $\xi|x|$ . We have  $\text{diam}(\tilde{S}) \geq \text{diam}(\pi(\tilde{S}))$ , where  $\pi$  is the projection onto the horizontal axis. Using the area formula

$$\mathcal{H}^1(\tilde{S}) = \sqrt{1 + \xi^2} \text{diam}(\pi(\tilde{S})) \leq (1 + \xi) \text{diam}(\tilde{S}).$$

□

### 3.2. Equivalence of $NCBV$ and $NCBV^+$ conditions.

**Proposition 3.15.** *Let  $f \in BV(Q(0, 1))$  be an  $NCBV^+$  map, then  $f$  is an  $NCBV$  map.*

*Proof.* It suffices to see that every good straight grid for  $f$  is also a good non-straight grid for  $f$ , which is evident. □

In the following lemma we refer by  $N_1, N_2 \subset [-1, 1]$  to the null sets from Corollary 3.2. Further, for all  $(y_1, y_2, x_1, x_2) \in [-1, 1]^4$  we define the (boundary of a) rectangle  $R_{(y_1, y_2, x_1, x_2)} = [(x_1, y_1)(x_2, y_1)] \cup [(x_2, y_1)(x_2, y_2)] \cup [(x_2, y_2)(x_1, y_2)] \cup [(x_1, y_2)(x_1, y_1)]$ .

**Lemma 3.16.** *Let  $f \in BV(Q(0, 1))$  be an  $NCBV$  map and let  $\sigma > 0$ . Then for every point  $(x_0, y_0) \in Q(0, 1)$  where*

$$(3.8) \quad \liminf_{r \rightarrow 0} r^{-1} |Df|(Q((x_0, y_0), r)) \leq \frac{1}{34} \sigma$$

*there exists a sequence of positive numbers  $r_n \rightarrow 0$  and a sequence of sets  $A_n \subset ((y_0 - r_n, y_0 - \frac{1}{2}r_n) \setminus N_1) \times ((y_0 + \frac{1}{2}r_n, y_0 + r_n) \setminus N_1) \times ((x_0 - r_n, x_0 - \frac{1}{2}r_n) \setminus N_2) \times ((x_0 + \frac{1}{2}r_n, x_0 + r_n) \setminus N_2)$  with  $\mathcal{L}^4(A_n) > 0$  and for any  $(y_1, y_2, x_1, x_2) \in A_n$  it holds that the straight grid generated by  $R_{(y_1, y_2, x_1, x_2)}$  is good for  $f$  and*

$$(3.9) \quad |D_\tau f|_{R_{(y_1, y_2, x_1, x_2)}}(R_{(y_1, y_2, x_1, x_2)}) < \frac{1}{4} \sigma.$$

*Proof.* By (3.8), we find a sequence  $r_n \rightarrow 0$  so that  $|Df|(Q((x_0, y_0), r_n)) < \frac{1}{33}\sigma r_n$ . We denote  $C_{y,n} = [x_0 - r_n, x_0 + r_n] \times \{y\}$  for any  $y \in [y_0 - r_n, y_0 + r_n]$ . By a standard disintegration argument (see [2, Theorem 2.28] and [2, Theorem 3.107]) and Lemma 2.10 we have

$$\begin{aligned} \frac{r_n \sigma}{33} &> |Df|(Q((x_0, y_0), r_n)) \\ &\geq \langle Df, (1, 0) \rangle(Q((x_0, y_0), r_n)) \\ &= \int_{y_0 - r_n}^{y_0 + r_n} |D_\tau f|_{C_{y,n}}|(C_{y,n}) dy. \end{aligned}$$

We use the argument

$$\int_A |D_\tau f|_{C_{y,n}}|(C_{y,n}) dy \geq \lambda \mathcal{L}^1(\{y \in A; |D_\tau f|_{C_{y,n}}|(C_{y,n}) \geq \lambda\})$$

for both  $A = (y_0 - r_n, y_0 - \frac{1}{2}r_n)$  and for  $A = (y_0 + \frac{1}{2}r_n, y_0 + r_n)$  to get

$$\begin{aligned} \mathcal{L}^1(\{t \in A; |D_\tau f|_{C_{y,n}}|(C_{y,n}) < \lambda\}) &= \frac{1}{2}r_n - \mathcal{L}^1(\{t \in A; |D_\tau f|_{C_{y,n}}|(C_{y,n}) \geq \lambda\}) \\ &\geq \frac{1}{2}r_n - \frac{1}{\lambda} \int_A |D_\tau f|_{C_{y,n}}|(C_{y,n}) dy \\ &> \frac{1}{2}r_n - \frac{r_n \sigma}{33\lambda}. \end{aligned}$$

Choose  $\lambda = \frac{1}{16}\sigma$ , then

$$(3.10) \quad \mathcal{L}^1(\{t \in A; |D_\tau f|_{C_{y,n}}|(C_{y,n}) < \frac{1}{16}\sigma\}) > 0.$$

The same estimates hold also for vertical lines.

Call  $A_n$  the subset of  $((y_0 - r_n, y_0 - \frac{1}{2}r_n) \setminus N_1) \times ((y_0 + \frac{1}{2}r_n, y_0 + r_n) \setminus N_1) \times ((x_0 - r_n, x_0 - \frac{1}{2}r_n) \setminus N_2) \times ((x_0 + \frac{1}{2}r_n, x_0 + r_n) \setminus N_2)$  whose each component satisfies (3.10). Then  $A_n$  has positive  $\mathcal{L}^4$  measure. Without loss of generality we may assume that each point of  $A_n$  is a Lebesgue point of the set (with respect to  $\mathcal{L}^4$ ). By Proposition 3.2 have that (almost every)  $[-1, 1] \times \{x_1\} \cup [-1, 1] \times \{x_2\} \cup \{y_1\} \times [-1, 1] \cup \{y_2\} \times [-1, 1]$  is admissible for  $f$ .  $\square$

**Proposition 3.17.** *Let  $f \in BV(Q(0, 1), \mathbb{R}^2)$  be an NCBV map, let  $\sigma > 0$  and let  $(x_0, y_0) \in Q(0, 1)$  satisfy*

$$(3.11) \quad \lim_{r \rightarrow 0} r^{-1} |Df|(Q((x_0, y_0), r)) = 0.$$

Let  $v_1, \dots, v_4 \in \mathbb{R}^2$  be distinct vectors with  $|v_i| = 1$  and call the rays  $\gamma_i = (x_0, y_0) + v_i[0, \infty)$ . Then there exists a constant  $\delta > 0$  depending on the vectors  $v_i$  and an  $r_0 = r_0(\sigma) > 0$  such that for any  $0 < r \leq r_0$  there exists a set  $A_r \subset Q((x_0, y_0), \delta r)$  with  $\mathcal{L}^2(A_r) > 0$  and for every  $(\tilde{x}, \tilde{y}) \in A_r$  there exist four polylines  $\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3, \tilde{\gamma}_4 \subset Q((x_0, y_0), r)$  with the following properties

- (1) the point  $(\tilde{x}, \tilde{y})$  is an endpoint of all  $\tilde{\gamma}_i$ ,  $i = 1, \dots, 4$  and the second endpoint of  $\tilde{\gamma}_i$  (we call it  $(x_i, y_i)$ ) lies on  $\gamma_i \setminus \{(\tilde{x}, \tilde{y})\}$ ,
- (2) the point  $(\tilde{x}, \tilde{y})$  is the only common point of the curves  $\tilde{\gamma}_i \cup (x_i, y_i) + (0, \infty)v_i$  i.e.  $[\tilde{\gamma}_i \cup ((x_i, y_i) + (0, \infty)v_i)] \cap [\tilde{\gamma}_j \cup ((x_j, y_j) + (0, \infty)v_j)] = \{(\tilde{x}, \tilde{y})\}$  for  $1 \leq i < j \leq 4$ ,
- (3) the collection  $\{\tilde{\gamma}_i \cup (x_i, y_i) + (0, \infty)v_i\}_{i=1}^4$  is admissible for  $f$ ,

- (4) all segments of  $\tilde{\gamma}_i$  are parallel to either  $(1, 0)$  or  $(0, 1)$ ,  
(5)  $\text{osc}(f, \tilde{\gamma}_i) \leq \frac{1}{4}\sigma$ .

*Proof.* We separate the plane into 4 quadrants  $P_1 = \{(x, y) \in \mathbb{R}^2; x > x_0, y \geq y_0\}$ ,  $P_2 = \{(x, y) \in \mathbb{R}^2; x \leq x_0, y > y_0\}$ ,  $P_3 = \{(x, y) \in \mathbb{R}^2; x < x_0, y \leq y_0\}$  and  $P_4 = \{(x, y) \in \mathbb{R}^2; x \geq x_0, y < y_0\}$ . Notice that each  $\gamma_i$  is contained in exactly one of the quadrants. Further observe that there exists a  $0 < \delta < \frac{1}{8}$  such that  $(\gamma_i - u) \setminus Q((x_0, y_0), \frac{1}{2})$  is contained in exactly one quadrant for any  $|u| < \delta$  (in the case of  $v_i = (\pm 1, 0)$  or  $(0, \pm 1)$  this may not be the same as the original quadrant). In every case having chosen  $|u| < \delta$  we have a uniquely determined quadrant  $P_i + u$  containing  $\gamma_i$ .

We define the polyline ‘spiral’

$$\begin{aligned} \gamma_1^* = & [(0, 0), (1, 0)] \cup [(1, 0), (1, 2)] \cup [(1, 2), (-3, 2)] \cup [(-3, 2), (-3, -4)] \\ & \cup [(-3, -4), (5, -4)] \cup [(5, -4), (5, 6)] \cup [(5, 6), (-7, 6)] \end{aligned}$$

and  $\gamma_2^*, \gamma_3^*, \gamma_4^*$  are rotations of  $\gamma_1^*$  by 90, 180 and 270 degrees clockwise. See Figure 1 for an illustration. We use these polylines as a starting point. In the following we construct  $\tilde{\gamma}_i$  from  $\gamma_i^*$  by translating, rescaling and truncating them.

We call

$$\gamma_{i,u,r}^* = (x_0, y_0) + u + r\gamma_i^* \quad \text{and} \quad \Gamma(u, r) = \bigcup_{i=1}^4 \gamma_{i,u,r}^*$$

for each  $u \in \{(x, y) \in \mathbb{R}^2; |x| + |y| \leq \delta r\} =: \tilde{Q}_r$  and  $r > 0$ . Then  $\Gamma(u, r) \subset Q((x_0, y_0), 8r)$  for each  $u \in \tilde{Q}_r$ . For every  $u \in \tilde{Q}_r$  we decompose  $u = s(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) + t(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$  for  $s, t \in [-\frac{1}{\sqrt{2}}\delta r, \frac{1}{\sqrt{2}}\delta r]$ . There exists a number  $N$  such that for any fixed  $s$  there is at most  $N$  points of the set  $\Gamma(0, r)$  on the corresponding line  $s(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) + \mathbb{R}(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ .

As explained above, by  $|D_\tau f|_{\Gamma(u,r)}$  we denote the measure on  $\Gamma(u, r)$  given by the one-dimensional variation of  $f$  on the segments of  $\Gamma(u, r)$ . Let us calculate

$$\begin{aligned} (3.12) \quad & \int_{\tilde{Q}_r} |D_\tau f|_{\Gamma(u,r)}(\Gamma(u, r)) d\mathcal{L}^2(u) \\ & = \int_{-\frac{\delta r}{\sqrt{2}}}^{\frac{\delta r}{\sqrt{2}}} \int_{-\frac{\delta r}{\sqrt{2}}}^{\frac{\delta r}{\sqrt{2}}} |D_\tau f|_{\Gamma(s(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) + t(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), r)}(\Gamma(s(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) + t(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), r)) dt ds \\ & \leq \int_{-\frac{\delta r}{\sqrt{2}}}^{\frac{\delta r}{\sqrt{2}}} N \left( |\langle Df, (0, 1) \rangle|(Q((x_0, y_0), 2r)) + |\langle Df, (1, 0) \rangle|(Q((x_0, y_0), 2r)) \right) ds \\ & \leq CN\delta r |Df|(Q(x_0, y_0), 8r). \end{aligned}$$

Using (3.11) we get that

$$\int_{\tilde{Q}_r} |Df|_{\Gamma(u,r)}(\Gamma(u, r)) d\mathcal{L}^2(u) \leq C\delta^{-1}r^{-1} |Df|(Q(x_0, y_0), 7r) \rightarrow 0$$

as  $r \rightarrow 0$ . Using the standard average values argument (as in the proof of (3.10)) we find an  $r_0 > 0$  such that for any  $0 < r < r_0$  there exists an  $\tilde{A}_r \subset \tilde{Q}_r$  of positive  $\mathcal{L}^2$

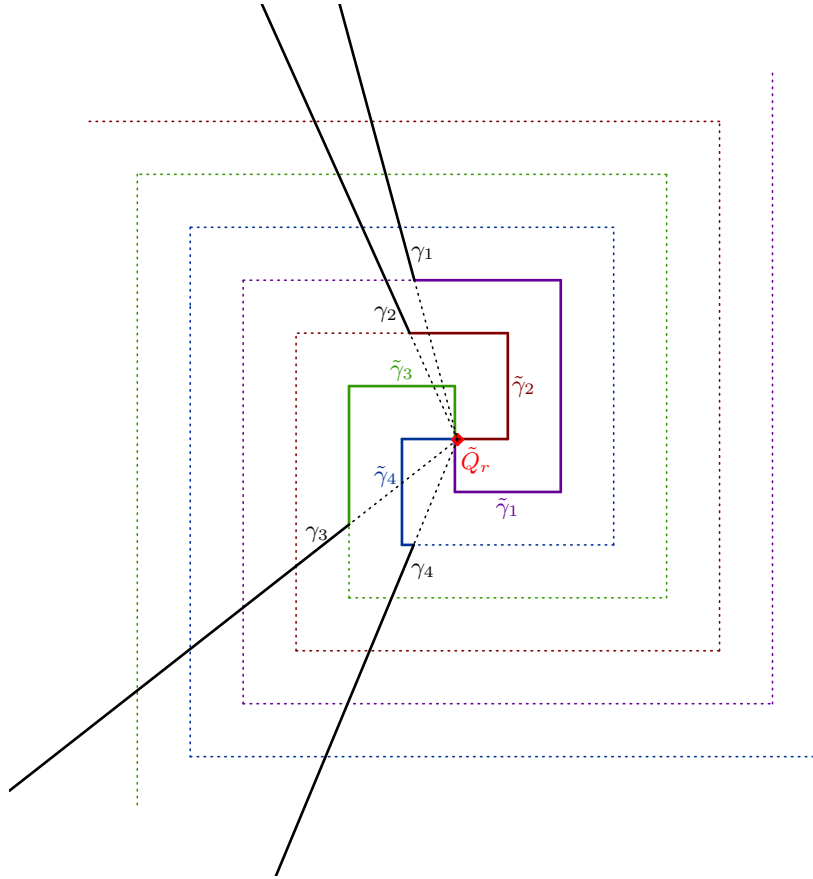


FIGURE 1. The approach to replacing four segments with a common endpoint by four polylines “spirals” parallel to coordinate axes. Here the paths have already been renumbered so that the first quadrant is the successor of the second of the two quadrants containing two segments which we have also renumbered.

measure such that

$$(3.13) \quad |D_\tau f|_{\Gamma(u,r)}(\Gamma(u,r)) \leq \frac{1}{4}\sigma \text{ for any } u \in \tilde{A}_r.$$

We may assume that all points of  $\tilde{A}_r$  are Lebesgue points for  $f$ .

Now we truncate the curves  $\tilde{\gamma}_{i,u,r}^*$  for  $u \in \tilde{A}_r$  so that they satisfy (1) and (2). The set  $\Gamma(u,r)$  is the union of four polylines,  $\gamma_{i,u,r}^*$ , each of which is a ‘square spiral’-type curve anti-clockwise around  $(x_0, y_0) + u$ . Recall that for each  $u \in A_r$  and for each  $\gamma_i$  we have a uniquely determined quadrant  $P_j$  such that  $\gamma_i \setminus Q((x_0, y_0), \frac{1}{2}r) \subset P_j + u$ . We call  $P_{j+1}$  the successor of  $P_j$  and  $P_{j-1}$  its predecessor calculating the indexes mod 4.

Assume that we have  $u \in A_r$  fixed. We start working with the quadrant whose predecessor contains the greatest number of segments  $\gamma_i$ . In the case where each quadrant contains exactly one segment we may start with any quadrant. In the case when there are two quadrants each containing two segments but neither is the successor of the other we start with the successor of either quadrant containing segments. If there are two successive quadrants both containing two segments each, we start with the

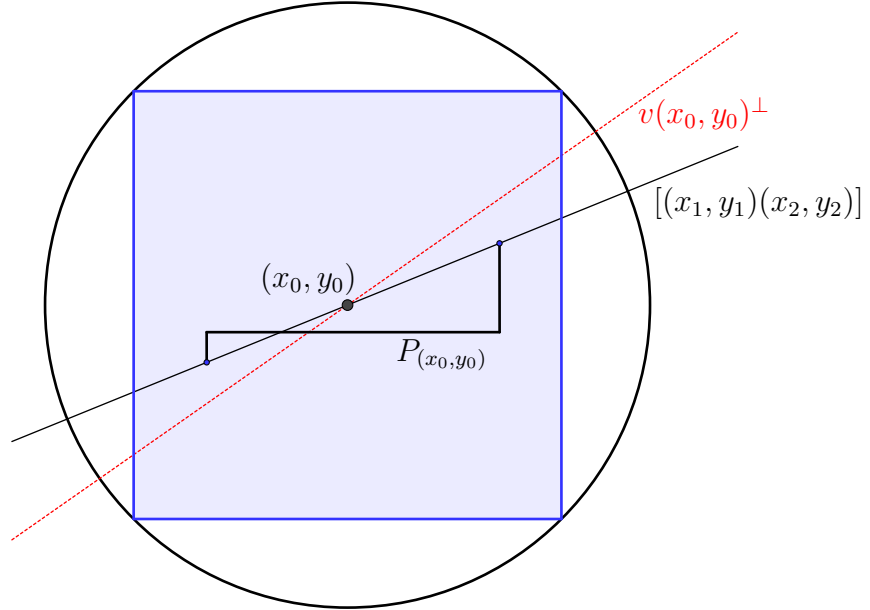


FIGURE 2. The situation close to the points of  $J_f \cap \Gamma$ . The jump set ‘approaches’ the approximate tangent space. Because  $(x_0, y_0)$  is a ‘Lebesgue’ jump point there are many horizontal segments with jumps very close to the jump at  $(x_0, y_0)$ . Such a horizontal segment can easily be connected to  $[(x_1, y_1)(x_2, y_2)]$  by segments parallel to coordinate axes.

quadrant succeeding the second of these two quadrants. Thus we have determined a starting quadrant (and without loss of generality) we assume it is  $P_1$ .

We define  $\tilde{\gamma}_1$  as the part of  $\gamma_{1,u,r}^*$  which goes from  $(\tilde{x}, \tilde{y}) = (x_0, y_0) + u$  to its first intersection with the first segment it meets, (if necessary re-number it to be)  $\gamma_1$ . Then we define each curve  $\tilde{\gamma}_i$  similarly;  $\tilde{\gamma}_i$  is the part of  $\gamma_{i,u,r}^*$  which goes from  $(\tilde{x}, \tilde{y}) = (x_0, y_0) + u$  to its first intersection with the segment (assumed after re-numbering to be)  $\gamma_i$  i.e. the first segment which has not yet been taken by a previous curve. Since there is a finite number of scenarios it is not difficult to check that the algorithm ensures points (1) and (2) hold by checking on a case by case basis. An illustration of this process is in Figure 1.

Point (4) is obvious. Point (5) is an obvious result of (3.13). Then, by Proposition 3.1  $\{\tilde{\gamma}_i \cup (x_i, y_i) + (0, \infty)v_i\}$  is admissible for  $f$  for almost every  $u \in \tilde{A}_r$ , which is point (3) where  $A_r = (x_0, y_0) + \tilde{A}_r$ .  $\square$

**Theorem 3.18.** *Let  $f \in BV(Q(0, 1), \mathbb{R}^2)$  and  $f = \text{id}$  on  $\partial Q(0, 1)$ . It holds that  $f$  is an  $NCBV^+$  map if and only if it is an  $NCBV$  map.*

*Proof.* Thanks to Proposition 3.15 it suffices to show that if  $f$  satisfies  $NCBV$  then  $f$  satisfies  $NCBV^+$ . Therefore we take a non-straight grid  $\Gamma$  good for  $f$  and, for any fixed  $\sigma$ , we construct  $\tilde{\Gamma}$  a straight grid good for  $f$  and an injective continuous map  $g$  from  $\hat{\Gamma} \subset \tilde{\Gamma}$  onto  $\Gamma$  with the property that given any injective approximation  $h$  of  $f$  on  $\tilde{\Gamma}$ , with  $\|h - f\|_{\infty, \tilde{\Gamma}} < \sigma/4$  it holds that  $\|h \circ g^{-1} - f\|_{\infty, \Gamma} < \sigma$ .



We call  $F_\sigma$  the set of  $(x_0, y_0) \in \Gamma \cap J_f$  Lebesgue points for  $(f^+(\cdot) - f^-(\cdot)) \otimes v(\cdot)$  in the sense of (3.4) such that  $|f^+(\cdot) - f^-(\cdot)| \geq \frac{1}{40}\sigma$ . Then  $F_\sigma$  is a finite set and for each  $(x_0, y_0) \in F_\sigma$  there exists exactly one segment  $[(x_1, y_1)(x_2, y_2)]$  of  $\Gamma$  containing  $(x_0, y_0)$ . Without loss of generality we assume that  $[(x_1, y_1)(x_2, y_2)] \cap F_\sigma = \{(x_0, y_0)\}$  and that  $|\langle (1, 0), v(x_0, y_0) \rangle| \geq \frac{1}{2}$ .

By the Federer-Vol'pert theorem, Theorem 2.8 (and Theorem 2.6) we have the strict convergence of  $r^{-1}|Df|(r[\cdot + (x_0, y_0)])$  to  $|f^+(x_0, y_0) - f^-(x_0, y_0)|\mathcal{H}^1_{|v(x_0, y_0)^\perp}$  on  $B(0, 1)$  as  $r \rightarrow 0$  for each point of  $(x_0, y_0) \in \Gamma \cap F_\sigma$ . Therefore we get

$$(3.14) \quad r^{-1}|\langle Df, (1, 0) \rangle|((x_0 - r, x_0 + r) \times (y_0 - \frac{1}{2}r, y_0 + \frac{1}{2}r)) \rightarrow |f^+(x_0, y_0) - f^-(x_0, y_0)|.$$

Further we have that  $[(x_1, y_1)(x_2, y_2)]$  is not parallel to the tangent space  $v(x_0, y_0)^\perp$ . Therefore we observe the existence of a polyline  $P_{(x_0, y_0)} \subset Q((x_0, y_0), r)$  with segments parallel to coordinate axes with one endpoint  $(\tilde{x}_{(x_0, y_0)}, \tilde{y}_{(x_0, y_0)}) \in [(x_1, y_1)(x_0, y_0)] \setminus \{(x_0, y_0)\}$  and the other on  $(\hat{x}_{(x_0, y_0)}, \hat{y}_{(x_0, y_0)}) \in [(x_0, y_0)(x_2, y_2)] \setminus \{(x_0, y_0)\}$  such that the straight grid generated by  $P_{(x_0, y_0)}$  (call it  $\tilde{\Gamma}_1$ ) is a good straight grid for  $f$  and

$$(3.15) \quad \left| |D_\tau f|_{\tilde{\Gamma}_1}(P_{(x_0, y_0)}) - |f^+(x_0, y_0) - f^-(x_0, y_0)| \right| \leq \left(\frac{\sigma}{100}\right)^2.$$

To ease our notation we denote  $S_{(x_0, y_0), \sigma} = [(\tilde{x}_{(x_0, y_0)}, \tilde{y}_{(x_0, y_0)})(\hat{x}_{(x_0, y_0)}, \hat{y}_{(x_0, y_0)})]$ . Further, thanks to (3.14), we assume that  $r > 0$  is chosen so small that

$$(3.16) \quad |D_\tau f|_{\Gamma}(S_{(x_0, y_0), \sigma}) \leq |f^+(x_0, y_0) - f^-(x_0, y_0)| + \left(\frac{\sigma}{100}\right)^2.$$

For a depiction of the process see Figure 2. Since we are in position to choose the  $x$ -coordinate of vertical segments and the  $y$ -coordinate for horizontal segments for the polylines  $P_{(x_0, y_0)}$  from sets of positive measure. Therefore we may assume that the horizontal segments do not have  $y$ -coordinate in the set  $N_1$  from Corollary 3.2 and the vertical segments do not have  $x$ -coordinate in the set  $N_2$  from Corollary 3.2 and that the straight grid generated by  $\bigcup_{(x_0, y_0) \in \Gamma \cap F_\sigma} P_{(x_0, y_0)}$  is good for  $f$ . Further we may assume that we choose  $r$  so small that the squares  $Q((x_0, y_0), r)$  for  $(x_0, y_0) \in F_\sigma \cap \Gamma$  are pairwise disjoint and do not contain any of the intersection points  $X_{i,j}$  of  $\Gamma$ .

Let  $r > 0$  be chosen so small that  $Q(X_{i,j}, r) \cap \Gamma$  is exactly the union of four segments,  $Q(X_{i,j}, r)$  are pairwise disjoint and are also disjoint with each  $Q((x_0, y_0), r)$  for each  $(x_0, y_0) \in F_\sigma \cap \Gamma$ . For each point  $X_{i,j}$  intersection point of  $\gamma_i([0, 1]) \neq \gamma_j([0, 1]) \subset \Gamma$  we find polylines (contained in  $Q(X_{i,j}, r)$ ) from Proposition 3.17, where the four segments from the claim of Proposition 3.17 (there called  $\gamma_1, \dots, \gamma_4$ ) refer to the four segments of  $\Gamma$ , which intersect at  $X_{i,j}$ . We will refer to these polylines as  $\tilde{\gamma}_k^{X_{i,j}}$ ,  $k = 1, 2, 3, 4$ . By the fact that  $A_r$  has positive measure we can guarantee that the straight grid generated by  $\{\tilde{\gamma}_k^{X_{i,j}} : i, j, k\}$  and by the polylines  $\{P_{(x_0, y_0)} : (x_0, y_0) \in F_\sigma\}$  is a good straight grid for  $f$ . We call this grid  $\tilde{\Gamma}_2$ .

Recall the notation; for every point  $(x_0, y_0) \in F_\sigma$  we have a segment (above denoted by  $S_{x_0, y_0, \sigma}$ ). Using an obvious adaptation of the notation of Proposition 3.17, with  $[(x_k(X_{i,j}), y_k(X_{i,j}))X_{i,j}]$  we denote the segment of endpoints  $(x_k(X_{i,j}), y_k(X_{i,j}))$  and  $X_{i,j}$  and we define the set

$$\kappa = \Gamma \setminus \left( \bigcup_{(x_0, y_0) \in F_\sigma} P_{(x_0, y_0)} \cup \bigcup_{X_{i,j}} \bigcup_{k=1}^4 [(x_k(X_{i,j}), y_k(X_{i,j}))X_{i,j}] \right).$$

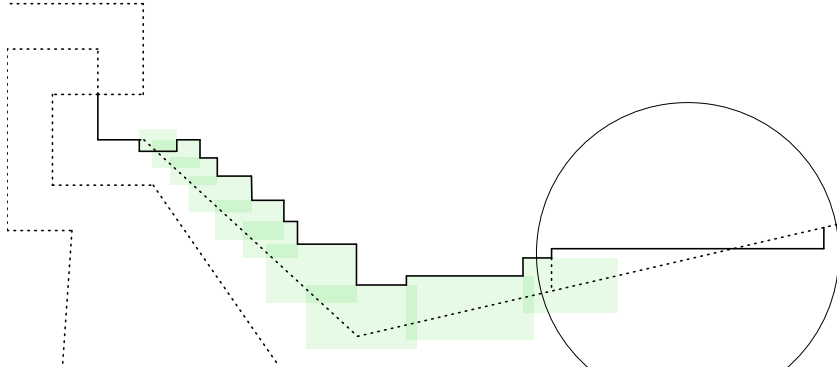


FIGURE 3. The above figure depicts how we choose a polyline parallel to coordinate axes along the boundary of chosen rectangles (shown in green) between (in this case) an intersection point and a jump point

Obviously  $\kappa$  is a compact set. All points of  $\kappa$  satisfy (3.8). Thus by Lemma 3.16 we have a fine covering of  $\kappa$  with rectangles each of which satisfies (3.9) and each rectangle is associated to a particular point on  $\kappa$ . We choose a finite covering of  $\kappa$  of the rectangles from Lemma 3.16 with the rectangles chosen so small that whenever a pair of rectangles intersect then either

- $R_1, R_2$  are associated with points on the same segment of  $\Gamma$  and the associated points are not disconnected by some  $(x_0, y_0) \in F_\sigma$
- $R_1, R_2$  are associated with points on neighbouring segments of  $\Gamma$  whose common endpoint is some  $\gamma_i(s_{i,k})$  (where  $s_i, k$  has been defined in Definition 3.3), not a point  $X_{i,j}$ .

Each rectangle can be made slightly bigger allowing us to choose horizontal lines with  $y$ -coordinate in  $[-1, 1] \setminus N_1$  (similarly for  $x$ -coordinate) since the set  $A_n$  has positive measure. Therefore we may assume that the grid generated by the sides of all the rectangles and  $\tilde{\Gamma}_2$  is a good straight grid for  $f$ . We may also assume that the covering we have chosen is minimal in the sense that it does not contain a strict sub-covering of  $\kappa$ . This grid is exactly  $\tilde{\Gamma}$  mentioned at the start of the proof.

Now we want to choose the set  $\hat{\Gamma} \subset \tilde{\Gamma}$ . Each of the polylines  $P_{(x_0, y_0)}$  intersect exactly two of the rectangles  $R_1, R_2$  chosen in the previous paragraph using Lemma 3.16 (each containing exactly one endpoint of  $P_{(x_0, y_0)}$ ). The part of  $P_{(x_0, y_0)}$  included in  $\hat{\Gamma}$  is  $\overline{P_{(x_0, y_0)} \setminus (R_1 \cup R_2)}$ .

For each  $X_{i,j}$ , we defined four polylines  $\tilde{\gamma}_1^{X_{i,j}}, \dots, \tilde{\gamma}_4^{X_{i,j}}$ , each of which intersects exactly one (and distinct) rectangle  $R_1, \dots, R_4$  chosen using Lemma 3.16. We include each  $\overline{\tilde{\gamma}_k^{X_{i,j}} \setminus R_k}$  in  $\hat{\Gamma}$  for  $k = 1, 2, 3, 4$  and do this at each  $X_{i,j}$ .

Let  $R$  be a rectangle chosen using Lemma 3.16 such that we already have exactly one point in  $\hat{\Gamma} \cap \partial R$ . Then  $R$  has a neighbouring  $R'$  (also chosen using Lemma 3.16) and there exists a path on  $\partial R$  from the previously chosen point in  $\hat{\Gamma} \cap \partial R$  to a point in  $\partial R'$ . We include this path in  $\hat{\Gamma}$ .

If we have a rectangle  $R$  chosen using Lemma 3.16 such that  $\hat{\Gamma} \cap \partial R$  contains exactly two points. Then there exists a path on  $\partial R$  between the two points of  $\hat{\Gamma} \cap \partial R$ . We include this path in  $\hat{\Gamma}$ . After dealing with the finite number of rectangles we have defined  $\hat{\Gamma}$ .

Now we describe how to define  $g$  on  $\hat{\Gamma}$ . We define  $g$  on  $\hat{\Gamma} \cap P_{(x_0, y_0)}$  as the constant speed map from  $\hat{\Gamma} \cap P_{(x_0, y_0)}$  onto the segment  $S_{x_0, y_0, \sigma}$ . Then  $g$  is injective and continuous on  $\hat{\Gamma} \cap P_{(x_0, y_0)}$ . On  $\hat{\Gamma} \cap \tilde{\gamma}_k^{X_{i,j}}$  we define  $g$  as the constant speed map onto the segment  $[(x_k(X_{i,j}), y_k(X_{i,j})) X_{i,j}]$ . For any pair of rectangles  $R_1, R_2$  chosen using Lemma 3.16 we choose a point  $Z_{R_1, R_2} \in \Gamma \cap R_1 \cap R_2$ . Then for each rectangle  $R$  chosen using Lemma 3.16 we define  $g$  as the constant speed map from  $\hat{\Gamma} \cap \partial R$  onto the part of  $\Gamma$  between the two points  $Z_{R, R'}$  and  $Z_{R, R''}$ , where  $R'$  and  $R''$  are the two neighbours of  $R$ . Alternatively if  $R$  neighbours  $R'$  and  $P_{(x_0, y_0)}$  we define  $g$  as constant speed onto the part of  $\Gamma$  between  $Z_{R, R'}$  and the corresponding endpoint of the polyline  $P_{(x_0, y_0)}$ . Finally, if  $R$  neighbours  $R'$  and  $\tilde{\gamma}_k^{X_{i,j}}$  we define  $g$  as constant speed onto the part of  $\Gamma$  between  $Z_{R, R'}$  and the corresponding endpoint of the polyline  $\tilde{\gamma}_k^{X_{i,j}}$ .

We separate  $\hat{\Gamma}$  into three pieces; parts of  $\partial R$ , the polylines  $\tilde{\gamma}_k^{X_{i,j}} \cap \hat{\Gamma}$  and the polylines  $P_{(x_0, y_0)} \cap \hat{\Gamma}$ . In each case we refer to the corresponding polyline as  $E$ . In the first two cases we have that  $|Df_{\hat{\Gamma}}|(E) \leq \frac{1}{4}\sigma$  by Lemma 3.16 and Proposition 3.17. Then an injective  $H_\sigma$  with  $\|H_\sigma - f\|_{\infty, \hat{\Gamma}} < \sigma/4$  has oscillation bounded by  $\frac{1}{2}\sigma$  on  $E$ . The same oscillation estimates holds on the segment  $\Gamma \cap R$  and  $f$  is continuous at the intersection points of  $\Gamma \cap \partial R$ . Therefore it holds that  $\|H_\sigma \circ g^{-1} - f\|_{\infty, E} < \sigma$ . The only case that needs closer consideration is the case of large jumps, i.e.  $E = P_{(x_0, y_0)} \cap \hat{\Gamma}$ . But in this case we have chosen the endpoints of  $P_{(x_0, y_0)} \cap \hat{\Gamma}$  so that their image is very close to the corresponding one sided limits at the jump point  $(x_0, y_0)$  (by (3.15) and (3.16) the error is bounded by  $4(\frac{1}{100}\sigma)^2$ ). Then by Lemma 3.10 we have that  $\|H_\sigma \circ g^{-1} - f\|_{\infty, E} < \frac{1}{4}\sigma$  and so we get our required result.  $\square$

#### 4. CONSTRUCTION OF A SEQUENCE CONVERGING AREA-STRICTLY TO $f$ IF $f$ SATISFIES THE $NCBV^+$ CONDITION

To approximate  $f$  area-strictly we need to isolate the majority of the singular part of  $Df$  from the majority of the absolutely continuous part of  $Df$ . That is the goal of the following lemma.

**Lemma 4.1.** *Let  $f \in BV(Q(0, 1))$  and let  $\varepsilon > 0$  then there exists a finite number of squares  $\{Q_i\}_{i=1}^{N_0}$  (whose union we denote as  $\tilde{F}_\varepsilon = \bigcup_{i=1}^{N_0} Q_i$ ) such that*

$$|D^a f|(\tilde{F}_\varepsilon) \leq \varepsilon |D^a f|(Q(0, 1))$$

and

$$|D^s f|(Q(0, 1) \setminus \tilde{F}_\varepsilon) \leq \varepsilon |D^s f|(Q(0, 1))$$

Further, if  $|D^a f|(Q(0, 1)) > 0$ , then

$$|D^s f|(Q(0, 1) \setminus \tilde{F}_\varepsilon) \leq \varepsilon^2 |D^a f|(Q(0, 1)).$$

*Proof.* We assume that  $|D^a f|(Q(0, 1)) > 0$  because the opposite case easily follows from the proof for this case. We may assume that  $|D^s f|(Q(0, 1)) > 0$  because otherwise the claim is nothing but the absolute continuity of the integral. Let  $\delta > 0$  be a number chosen small enough that

$$|D^a f|(A) \leq \varepsilon \min \{1, |D^a f|(Q(0, 1)), |D^s f|(Q(0, 1))\}$$

for any  $A \subset Q(0, 1)$  such that  $\mathcal{L}^2(A) < \delta$ . We call the set

$$S = \left\{ x \in Q(0, 1); \lim_{r \rightarrow 0} \frac{|Df|(Q(x, r))}{r^2} = \infty \right\}$$

and similarly we call

$$A_M^R = \left\{ x \in Q(0, 1): \frac{|Df|(Q(x, r))}{r^2} > 4M \text{ for every } 0 < r < R \right\}.$$

Recall that by [2, Proposition 3.92]

$$(4.1) \quad |D^s f|(Q(0, 1) \setminus S) = 0.$$

By the Vitali-Besicovitch theorem [2, Theorem 2.19] we have a countable collection of pairwise disjoint squares  $Q(x_i, r_i)$  with  $x_i \in A_M^R$  and  $0 < r_i < R$  covering  $A_M^R$  (up to a set of  $\mathcal{L}^n + |Df|$ -measure 0). Using  $\frac{|Df|(Q(x_i, r_i))}{r_i^2} > 4M$  we get

$$\mathcal{L}^2(A_M^R) \leq \sum_{i=1}^{\infty} 4r_i^2 < \frac{|Df|(Q(0, 1))}{M}.$$

and by choosing  $M = \delta^{-1} 2|Df|(Q(0, 1))$  we have

$$(4.2) \quad \mathcal{L}^2(A_M^R) < \frac{\delta}{2}.$$

Obviously for all  $M > 0$  and all  $R_0 > 0$  we have

$$S \subset \bigcup_{0 < R < R_0} A_M^R,$$

especially the inclusion  $A_M^{R_1} \subset A_M^{R_2}$  holds for any  $R_1 > R_2 > 0$ . This shows that  $\mathcal{L}^2(S) = 0$  but also, using (4.1), that  $|D^s f|(S \setminus A_M^R) \rightarrow 0$  as  $R \rightarrow 0^+$  (the limit makes sense thanks to the previous inclusion). Find then an  $R_0$  such that

$$(4.3) \quad \begin{aligned} |D^s f|(S \setminus A_M^{R_0}) &< \frac{\varepsilon}{2} |D^s f|(Q(0, 1)), & |D^s f|(S \setminus A_M^{R_0}) &< \frac{\varepsilon^2}{2} |D^a f|(Q(0, 1)), \\ \mathcal{L}^2(A_M^{R_0} + B_{R_0}(0)) &\leq \delta, \end{aligned}$$

where with  $A_M^{R_0} + B_{R_0}(0)$  we mean the set  $\{x \in \mathbb{R}^2 : \text{dist}(x, A_M^{R_0}) < R_0\}$ .

The Vitali-Besicovitch theorem gives a countable number of pairwise disjoint squares  $Q_i = Q(x_i, R_i)$ ,  $x_i \in A_M^{R_0}$  with  $R_i \leq R_0$ . Since  $\lim_{N \rightarrow \infty} |D^s f|(A_M^{R_0} \setminus \bigcup_{i=1}^N Q_i) = 0$  we can find a  $N_0$  such that

$$|D^s f|\left(A_M^{R_0} \setminus \bigcup_{i=1}^{N_0} Q_i\right) < \frac{\varepsilon}{2} |D^s f|(Q(0, 1))$$

and similarly

$$|D^s f|\left(A_M^{R_0} \setminus \bigcup_{i=1}^{N_0} Q_i\right) < \frac{\varepsilon^2}{2} |D^a f|(Q(0, 1)).$$

Combining the last two estimates with (4.1) and (4.3) we have

$$|D^s f|\left(Q(0, 1) \setminus \bigcup_{i=1}^{N_0} Q_i\right) < \varepsilon |D^s f|(Q(0, 1)) \text{ and}$$

$$|D^s f|\left(Q(0, 1) \setminus \bigcup_{i=1}^{N_0} Q_i\right) < \varepsilon^2 |D^a f|(Q(0, 1)).$$

On the other hand, by the third of (4.3), (4.2) and the choice of  $\delta$ , we have

$$|D^a f|\left(\bigcup_{i=1}^{N_0} Q_i\right) \leq \varepsilon \min\{1, |D^a f|(Q(0, 1)), |D^s f|(Q(0, 1))\}.$$

□

The following theorem is used to divide  $Q(0, 1)$  up into small squares (side length is  $2^{1-K}$ ) which have different properties. In the squares of type  $E_{\varepsilon, K}$  the majority of the behaviour comes from the singular part of the derivative and the direction map of the polar decomposition of the derivative satisfies a Lebesgue-point-type estimate based on the choice of  $\varepsilon$ . Further we categorise the other squares, which are used to approximate  $D^a f$  in  $L^1$ . The categories are  $G_{\varepsilon, \alpha, K}$  (where  $f$  is very close to a nice affine map),  $T_{\varepsilon, \alpha, K}$  (where  $f$  is very close to a non-constant affine map with zero Jacobian) and  $W_{\varepsilon, \alpha, K}$  (where  $f$  is nearly constant or does not behave ‘ $\varepsilon$ -similarly’ to any affine map). In the following theorem we refer to the set  $\tilde{F}_\varepsilon$  defined in Lemma 4.1.

**Theorem 4.2.** *Let  $f \in BV(Q(0, 1))$  be an NCBV map and let  $\varepsilon > 0$ . There exists an  $0 < \alpha_0 < \varepsilon$  such that for any  $0 < \alpha \leq \alpha_0$  the following holds. There exists a  $K = K(\varepsilon, \alpha) \in \mathbb{N}$  such that the division of  $Q(0, 1)$  into  $K$ -dyadic squares  $\{Q_i\}_{i=1}^{2^{2K}}$  has the following properties*

- (1) *calling  $F_{\varepsilon, K}$  the set of  $K$ -dyadic squares which themselves intersect  $\tilde{F}_\varepsilon$  or have a neighbouring  $K$ -dyadic square that intersects  $\tilde{F}_\varepsilon$  and calling  $\tilde{F}_{\varepsilon, K} = \bigcup_{Q_i \in F_{\varepsilon, K}} Q_i$  it holds that*

$$|D^a f|(\tilde{F}_{\varepsilon, K}) \leq 2\varepsilon |D^a f|(Q(0, 1)),$$

- (2) *it holds that*

$$|D^s f|(Q(0, 1) \setminus \tilde{F}_{\varepsilon, K}) \leq \varepsilon |D^s f|(Q(0, 1))$$

*and if  $|D^a f|(Q(0, 1)) > 0$  then*

$$|D^s f|(Q(0, 1) \setminus \tilde{F}_{\varepsilon, K}) \leq \varepsilon^2 |D^a f|(Q(0, 1)),$$

- (3) *there exists a subselection  $E_{\varepsilon, K}$  of squares of  $F_{\varepsilon, K}$  (whose union we denote as  $\tilde{E}_{\varepsilon, K}$ ) such that*

$$|D^s f|(Q(0, 1) \setminus \tilde{E}_{\varepsilon, K}) \leq 2\varepsilon |D^s f|(Q(0, 1))$$

and for any  $Q_i \in E_{\varepsilon, K}$  there exists a  $w_i \in Q_i \cap S$  such that

$$(4.4) \quad \int_{S \cap 2Q_i} |g(z) - g(w_i)| d|D^s f|(z) \leq \varepsilon |D^s f|(4Q_i \cap S),$$

where  $g \in L^1(|D^s f|, \mathbb{R}^2)$ ,  $g|D^s f| = D^s f$  is the polar decomposition of  $D^s f$  and  $g(w_i) = u_i \otimes v_i$  for an appropriate  $|u_i| = |v_i| = 1$ .

- (4) Further, all squares  $Q_i$  with  $Q_i \cap \tilde{F}_\varepsilon = \emptyset$ , are separated into three disjoint categories  $G_{\varepsilon, \alpha, K}$ ,  $T_{\varepsilon, \alpha, K}$  and  $W_{\varepsilon, \alpha, K}$  (their unions denoted by  $\tilde{G}_{\varepsilon, \alpha, K}$ ,  $\tilde{T}_{\varepsilon, \alpha, K}$  and  $\tilde{W}_{\varepsilon, \alpha, K}$ ) such that

$$(4.5) \quad |D^a f|(\tilde{W}_{\varepsilon, \alpha, K}) \leq 8\varepsilon [|D^a f|(Q(0, 1)) + 1]$$

and, for any  $Q_i = Q(c_i, 2^{-K}) \in G_{\varepsilon, \alpha, K} \cup T_{\varepsilon, \alpha, K}$  there exists an  $w_i \in Q(c_i, 2^{-K-2})$  and a set  $Z_{i, \alpha} \subset Q(c_i, 2^{1-K})$  with  $\mathcal{L}^2(Q(c_i, 2^{1-K}) \setminus Z_{i, \alpha}) \leq 2^{-2K-9}$  such that

$$(4.6) \quad \begin{aligned} & \int_{Q(c_i, 2^{2-K})} |\nabla f(y) - \nabla f(w_i)| d\mathcal{L}^2(y) \leq \varepsilon \alpha^2 2^{-2K}, \\ & \|f(\cdot) - f(w_i) - \nabla f(w_i)(\cdot - w_i)\|_{L^\infty(Z_{i, \alpha})} < \alpha^4 2^{-K}, \\ & \alpha_0 \leq |\nabla f(w_i)| \leq \alpha_0^{-1}, \end{aligned}$$

and

$$|D^s f|(Q_i) \leq \varepsilon |D^a f|(Q_i).$$

In the case that  $Q_i \in G_{\varepsilon, \alpha, K}$  it holds that  $\alpha_0 < \det \nabla f(w_i)$  and in the case  $Q_i \in T_{\varepsilon, \alpha, K}$  it holds that  $\det \nabla f(w_i) = 0$ .

*Proof.*

**Step 1. Prove (1) and (2) by applying Lemma 4.1.**

The set  $\tilde{F}_\varepsilon$  is the union of a finite number of disjoint squares. Then as  $K \rightarrow \infty$  we clearly have  $\mathcal{L}^2(\tilde{F}_{\varepsilon, K} \setminus \tilde{F}_\varepsilon) \rightarrow 0$ . Therefore we find a  $K_0$  such that for any  $K \geq K_0$  we have  $\mathcal{L}^2(\tilde{F}_{\varepsilon, K} \setminus \tilde{F}_\varepsilon) < \delta$ , where  $\delta$  is so small that  $|D^a f|(A) \leq \varepsilon |D^a f|(Q(0, 1))$  as soon as  $\mathcal{L}^2(A) < \delta$ . Since  $\tilde{F}_\varepsilon$  was chosen so that  $|D^a f|(\tilde{F}_\varepsilon) \leq \varepsilon |D^a f|(Q(0, 1))$ , we satisfy point (1) of the claim. The set  $\tilde{F}_{\varepsilon, K} \supset \tilde{F}_\varepsilon$  and so point (2) of our claim is immediate from Lemma 4.1.

**Step 2. Find a  $K_1$  that allows us to prove (3).**

Let us call  $g$  the function of the so-called polar decomposition of  $D^s f$ . Then  $|g| = 1$   $|D^s f|$ -almost everywhere and  $D^s f = g|D^s f|$ . The function  $g \in L^1(Q(0, 1), |D^s f|, \mathbb{R}^2)$  and so  $|D^s f|$  almost every point of  $S$  is a Lebesgue point of  $g$  with respect to  $|D^s f|$ . Recall that the singular part of the derivative of  $f$  is supported on  $S$ , i.e.  $D^s f = D^s f|_S$  (see [2, Proposition 3.92]). As a result of this (see [13, Theorem 1.33]) for  $|D^s f|$ -almost every  $w \in S$  it holds that

$$\frac{1}{|D^s f|(Q(w, r) \cap S)} \int_{Q(w, r) \cap S} |g(z) - g(w)| d|D^s f|(z) \xrightarrow{r \rightarrow 0^+} 0.$$

Thus, for any given  $\varepsilon$  the  $|D^s f|$  measure of the set of points  $w \in S$  such that

$$\int_{Q(w, r) \cap S} |g(z) - g(w)| d|D^s f|(z) > \varepsilon |D^s f|(Q(w, r) \cap S)$$

for some  $0 < r < 2^{-K}$  tends to zero as  $K \rightarrow \infty$ . Thus we find a  $K_1$  so that the  $|D^s f|$  measure of this set is bounded by  $\varepsilon |D^s f|(Q(0, 1))$ . Call  $\tilde{X}_{\varepsilon, K}$  the set of  $w \in S$  such that

$$\int_{Q(w, r) \cap S} |g(z) - g(w)| d|D^s f|(z) \leq \varepsilon |D^s f|(Q(w, r) \cap S) \text{ for all } 0 < r < 2^{-K_1}.$$

For any  $K \geq K_1 + 1$  we have for any  $Q_i = Q(c_i, 2^{-K}) \in F_{\varepsilon, K}$  such that  $Q_i \cap \tilde{X}_{\varepsilon, K} \neq \emptyset$  and for any choice of  $w_i \in Q_i \cap \tilde{X}_{\varepsilon, K}$  that

$$\begin{aligned} \int_{Q(c_i, 2^{-K}) \cap S} |g(z) - g(w_i)| d|D^s f|(z) &\leq \int_{Q(w_i, 2^{1-K}) \cap S} |g(z) - g(w_i)| d|D^s f|(z) \\ &\leq \varepsilon |D^s f|(Q(w_i, 2^{1-K})) \\ &\leq \varepsilon |D^s f|(Q(c_i, 2^{2-K})). \end{aligned}$$

We define the collection  $E_{\varepsilon, K}$  (for  $K \geq \max\{K_0, K_1 + 1\}$ ) as those squares  $Q_i \in F_{\varepsilon, K}$  such that  $Q_i \cap \tilde{X}_{\varepsilon, K} \neq \emptyset$ . By the choice of  $K_1$ , we have that

$$|D^s f|\left(\bigcup_{Q_i \in F_{\varepsilon, K} \setminus E_{\varepsilon, K}} Q_i\right) = |D^s f|(S \setminus \tilde{X}_{\varepsilon, K}) < \varepsilon |D^s f|(Q(0, 1)).$$

These two estimates together with (2) are point (3) of our claim.

**Step 3. Choose an appropriate  $\alpha_0 > 0$  and for every  $0 < \alpha < \alpha_0$  find an appropriate  $K(\varepsilon, \alpha)$ .**

For every  $K \geq \max\{K_0, K_1 + 1\}$  we call  $W'_{\varepsilon, K}$  the collection of  $Q_i$ , the  $K$ -dyadic squares  $Q_i$  such that

$$Q_i \notin F_{\varepsilon, K} \text{ and } |D^s f|(2Q_i) > \varepsilon |D^a f|(Q_i).$$

Denote  $\tilde{W}'_{\varepsilon, K} = \bigcup_{Q_i \in W'_{\varepsilon, K}} Q_i$ . Notice that for any  $Q_i \notin F_{\varepsilon, K}$  we have  $2Q_i \cap \tilde{F}_\varepsilon = \emptyset$  by the definition of  $F_{\varepsilon, K}$ . Then, by the second estimate of point (2), for every  $K$  we have

$$(4.7) \quad |D^a f|(\tilde{W}'_{\varepsilon, K}) \leq \varepsilon^{-1} \sum_{Q_i \in W'_{\varepsilon, K}} |D^s f|(2Q_i) \leq 4\varepsilon |D^a f|(Q(0, 1)).$$

The constant 4 is the overlap multiplicity bound for  $\{2Q_i\}$ , i.e. a bound for  $\sum_i \chi_{2Q_i}$ . The remaining squares  $Q(0, 1) \supset Q_i \notin F_{\varepsilon, K} \cup W'_{\varepsilon, K}$  satisfy the estimate  $|D^s f|(2Q_i) \leq \varepsilon |D^a f|(Q_i)$ .

Call  $\tilde{P}_{\alpha_0}$  the set where

$$\begin{aligned} \tilde{P}_{\alpha_0} = &\{w \in Q(0, 1) \setminus S; |\nabla f(w)| > \alpha_0^{-1}\} \cup \{w \in Q(0, 1); 0 < |\nabla f(w)| < \alpha_0\} \\ &\cup \{w \in Q(0, 1); 0 < \det \nabla f(w) < \alpha_0\} \end{aligned}$$

Our first observation is that as  $\alpha_0 \rightarrow 0$  we have

$$\begin{aligned} \mathcal{L}^2\left(\{w \in Q(0, 1) \setminus S; |\nabla f(w)| > \alpha_0^{-1}\}\right) &\rightarrow 0 \\ \mathcal{L}^2\left(\{w \in Q(0, 1); 0 < |\nabla f(w)| < \alpha_0\}\right) &\rightarrow 0 \\ \mathcal{L}^2\left(\{w \in Q(0, 1); 0 < \det \nabla f(w) < \alpha_0\}\right) &\rightarrow 0, \end{aligned}$$

because as we send  $\alpha_0 \rightarrow 0$  the sets (which are nested) tend to the empty set. Recall the choice of the parameter  $\delta$ , chosen such that  $|D^a f|(A) \leq \varepsilon |D^a f|(Q(0, 1))$  for any  $A$  such that  $\mathcal{L}^2(A) < \delta$ . We find an  $0 < \alpha_0 < \varepsilon$  such that  $\mathcal{L}^2(\tilde{P}_{\alpha_0}) < \frac{\delta}{16}$ .

From [2, Theorem 3.83] we have that  $\mathcal{L}^2$ -almost every point of  $Q(0, 1)$  is a point of approximate differentiability of  $f$ . We define  $\tilde{Y}_{\varepsilon, \alpha, K}$  as the set of points  $w \in Q(0, 1)$  such that

$$(4.8) \quad \frac{1}{\mathcal{L}^2(Q(w, 8r))} \int_{Q(w, 8r)} |\nabla f(z) - \nabla f(w)| d\mathcal{L}^2(y) > 2^{-8} \varepsilon \alpha^2 \quad \text{or} \\ \mathcal{L}^2\left(\left\{y \in Q(w, 4r) : |f(z) - f(w) - \nabla f(w)(z - w)| > \alpha^4 r\right\}\right) \geq 2^{-9} r^2$$

for some  $0 < r < 2^{-K}$ . By [2, Theorem 3.83] the  $\mathcal{L}^2$  measure of  $\tilde{Y}_{\varepsilon, \alpha, K}$  tends to zero as  $K \rightarrow \infty$ . Therefore we find a  $K_2(\alpha)$  sufficiently large such that for any  $0 < \alpha < \alpha_0$  and any  $K \geq K_2(\alpha)$  we have  $\mathcal{L}^2(\tilde{Y}_{\varepsilon, \alpha, K}) < \delta/16$ . From now on we require that  $K \geq K_2$ . Also we require  $2^{-K_2} < \varepsilon \alpha_0$ .

**Step 4. Designate the squares  $G_{\varepsilon, \alpha, K}, T_{\varepsilon, \alpha, K}, W_{\varepsilon, \alpha, K}$  and prove (4).**

Notice that the bound on the measure of  $\tilde{P}_{\alpha_0}$  and  $\tilde{Y}_{\varepsilon, \alpha, K}$  implies that the union of all  $K$ -dyadic squares  $Q_i = Q(c_i, 2^{-K})$  such that  $Q(c_i, 2^{-K-2}) \subset \tilde{P}_{\alpha_0} \cup \tilde{Y}_{\varepsilon, \alpha, K}$  has measure at most  $2\delta$ . Therefore

$$(4.9) \quad |D^a f|(\tilde{P}_{\alpha_0} \cup \tilde{Y}_{\varepsilon, \alpha, K}) \leq 2\varepsilon |D^a f|(Q(0, 1))$$

because of the choice of  $\delta$ . We call  $W_{\varepsilon, \alpha, K}$  the collection of those squares either

- (i)  $Q_i \in W'_{\varepsilon, K}$  or
- (ii)  $Q(c_i, 2^{-K-2}) \subset \tilde{P}_{\alpha_0} \cup \tilde{Y}_{\varepsilon, \alpha, K}$  or
- (iii)  $|D^a f|(Q_i) < 2\alpha_0 2^{-2K+2} = 2\alpha_0 \mathcal{L}^2(Q_i)$

In case (iii) we use the fact that  $\mathcal{L}^2(Q(0, 1)) = 4$  the choice of  $\alpha_0 < \varepsilon$  to get

$$|D^a f|\left(\bigcup_{\{i: |D^a f|(Q_i) < 2\alpha_0 2^{-2K+2}\}} Q_i\right) \leq \sum_{\{i: |D^a f|(Q_i) < 2\alpha_0 2^{-2K+2}\}} |D^a f|(Q_i) \leq 2\alpha_0 4 < 8\varepsilon.$$

This in combination with (4.7) (for case (i)) and (4.9) (for case (ii)) prove the estimate (4.5).

All the other squares  $Q_i \notin F_{\varepsilon, K} \cup W_{\varepsilon, \alpha, K}$  have a point  $w_i \in Q(c_i, 2^{-K-1}) \setminus (P_{\alpha_0} \cup Y_{\varepsilon, \alpha, K})$  and therefore they satisfy the estimates

$$|\nabla f(w_i)| \leq \alpha_0^{-1}$$

and

$$\text{either } \det \nabla f(w_i) \geq \alpha_0 \text{ or } \det \nabla f(w_i) = 0$$

and

$$\int_{Q(c_i, 2^{2-K})} |\nabla f(z) - \nabla f(w_i)| d\mathcal{L}^2(z) \leq \int_{Q(w_i, 2^{3-K})} |\nabla f(z) - \nabla f(w_i)| d\mathcal{L}^2(z) \\ \leq \varepsilon \alpha^2 2^{-2K}.$$

Further the fact that  $|D^a f|(Q_i) \geq 2\alpha_0 \mathcal{L}^2(Q_i)$  implies that  $|\nabla f(w_i)| \geq 2\alpha_0 - \varepsilon \alpha \alpha_0$  and since  $\alpha < \varepsilon < 1$  we have

$$\alpha_0 \leq |\nabla f(w_i)| \leq \alpha_0^{-1}.$$



Moreover, using  $Q(c_i, 2^{1-K}) \subset Q(w_i, 2^{2-K})$  and  $r = 2^{-K}$  in (4.8) we have that

$$\mathcal{L}^2\left(\{z \in Q(c_i, 2^{1-K}) : |f(z) - f(w_i) - \nabla f(w_i)(z - w_i)| > \alpha^4 2^{-K}\}\right) < 2^{-2K-9}.$$

Then,  $Z_{i,\alpha} \subset Q(c_i, 2^{1-K})$  being the set where  $|f(\cdot) - f(w_i) - \nabla f(w_i)(\cdot - w_i)| \leq \alpha^4 2^{-K}$  satisfies  $\mathcal{L}^2(Q_i \setminus Z_{i,\alpha}) < 2^{-2K-9}$ . Thus we have proved point (4).  $\square$

In [2, Proposition 3.92] the authors introduced the set

$$\Theta_f = \{w \in Q(0, 1); \liminf_{r \rightarrow 0} r^{-1} |Df|(B(w, r)) > 0\}.$$

We adapt slightly this notion and in the following theorem we use the sets

$$\Theta_f^\beta = \{w \in Q(0, 1); \liminf_{r \rightarrow 0} r^{-1} |Df|(B(w, r)) > \frac{1}{10}\beta\}$$

for  $\beta > 0$ .

In the Theorem 4.3, for each pair of neighbouring vertexes  $V, \tilde{V}$  of some square  $Q_i$  chosen in Theorem 4.2, we find sets  $H_{V, \tilde{V}}$  such that when we create a quadrilateral by shifting  $V$  and  $\tilde{V}$  to a pair of points in  $H_{V, \tilde{V}}$  then the behaviour of  $f|_{\llbracket V, \tilde{V} \rrbracket}$  corresponds to the behaviour of  $f$  inside  $Q_i$ . In fact we can create a good non-straight grid for  $f$  by joining neighbouring shifted vertices with segments. Further we get the useful estimates (4.10) and (4.11).

**Theorem 4.3.** *Let  $f \in BV(Q(0, 1))$  be an NCBV map, let  $\varepsilon, \beta > 0$ , and let  $\alpha_0 < \varepsilon$  be the number given by Theorem 4.2 and let  $0 < \alpha < \min\{\alpha_0, 2^{-9}\}$ . Let  $K = K(\varepsilon, \alpha) \in \mathbb{N}$  and  $\{Q_i\}_{i=1}^{2^{2K}}$  be the  $K$ -dyadic squares chosen in Theorem 4.2. There exists a constant  $C > 0$  such that the following holds. Let  $V$  be a vertex of a square  $Q_i$  and let  $\tilde{V}$  be one of its neighbouring vertices. There exists a set  $H_{V, \tilde{V}} \subset Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2})$  with  $\mathcal{L}^4(H_{V, \tilde{V}}) \geq \frac{4}{5} \mathcal{L}^4(Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2}))$  with the following properties:*

- (1) *for every pair  $(X, \tilde{X}) \in H_{V, \tilde{V}}$  both  $X$  and  $\tilde{X}$  are Lebesgue points of  $f$  and we assume that*

$$f(X) = \lim_{r \rightarrow 0} \int_{B(X, r)} f(z) dz \quad \text{and} \quad f(\tilde{X}) = \lim_{r \rightarrow 0} \int_{B(\tilde{X}, r)} f(z) dz,$$

- (2) *for any pair  $(X, \tilde{X}) \in H_{V, \tilde{V}}$  the map  $f|_{L_{X, \tilde{X}}}$  is continuous at  $X$  and  $\tilde{X}$ , where  $L_{X, \tilde{X}}$  is the line passing through  $X$  and  $\tilde{X}$ ,*  
 (3) *for any pair  $(X, \tilde{X}) \in H_{V, \tilde{V}}$  the segment  $[X\tilde{X}]$  intersects the set  $\Theta_f^\beta$  only at Lebesgue points of the function  $(f^+(x) - f^-(x)) \otimes v(x)$  with respect to  $\mathcal{H}_{\Theta_f^\beta}^1$  and further  $\langle X - \tilde{X}, v(x) \rangle \neq 0$  at every such point of intersection,*  
 (4) *for any pair  $(X, \tilde{X}) \in H_{V, \tilde{V}}$  it holds that  $|Df|([X\tilde{X}]) = 0$ ,*  
 (5) *for any pair  $(X, \tilde{X}) \in H_{V, \tilde{V}}$  the estimate holds*

$$(4.10) \quad |D_\tau f|_{\llbracket X\tilde{X} \rrbracket}([X\tilde{X}]) \leq C 2^K |Df|(2Q_i).$$

*Further, if  $V$  and  $\tilde{V}$  are both vertices of some  $Q_j \in G_{\varepsilon, \alpha, K} \cup T_{\varepsilon, \alpha, K}$  (the collection of squares defined in Theorem 4.2) then*

$$(4.11) \quad |D_\tau [f(\cdot) - \nabla f(w_j)(\cdot)]|_{\llbracket X\tilde{X} \rrbracket}([X\tilde{X}]) \leq C \varepsilon |D^a f|(Q(c_j, 2^{1-K})) 2^K$$

where  $w_j$  is the point chosen in Theorem 4.2. Moreover

$$(4.12) \quad X, \tilde{X} \in Z_{j,\alpha},$$

where  $Z_{j,\alpha}$  is the set from Theorem 4.2.

*Proof.* To satisfy point (1) it suffices to consider a ‘good’ representative of  $f$  and eliminate the set of non-Lebesgue points from  $Q(V, 2^{-K-2})$  and  $Q(\tilde{V}, 2^{-K-2})$ , which is a set of zero measure. Call  $N_1$  the set of all  $(X, \tilde{X}) \in Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2})$  such that either  $X$  is not a Lebesgue point of  $f$  or  $\tilde{X}$  is not a Lebesgue point of  $f$ . Clearly  $\mathcal{L}^4(N_1) = 0$ .

During the course of this proof we use the mapping

$$\Psi : (X, \tilde{X}) \rightarrow \left( \frac{X - \tilde{X}}{|X - \tilde{X}|}, \left\langle X, \frac{X - \tilde{X}}{|X - \tilde{X}|} \right\rangle, \left\langle \tilde{X}, \frac{X - \tilde{X}}{|X - \tilde{X}|} \right\rangle, \left\langle X, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{X - \tilde{X}}{|X - \tilde{X}|} \right\rangle \right).$$

In the image of  $\Psi$  we have the measure

$$\mu := \mathcal{H}^1_{\{|X|=1\}} \times \mathcal{L}^3.$$

The first space,  $\mathcal{H}^1_{\{|X|=1\}}$ , is bi-Lipschitz equivalent with  $\mathcal{L}^1_{(0,2\pi)}$  and so the measure  $\mu$  is locally bi-Lipschitz equivalent with  $\mathcal{L}^4$  on  $\mathbb{R}^4$  which is  $\mathcal{H}^4$  on  $\mathbb{R}^4$ . In this sense  $\mu$  is equivalent with  $\mathcal{H}^4$  on  $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1^2 + x_2^2 = 1\}$ . We have the validity of the area formula (for example see [13, Theorem 3.8]) for  $\Psi$ , i.e.

$$\int_E \mathbf{J}_4 \Psi(X, \tilde{X}) \, d\mathcal{L}^4(X, \tilde{X}) = \int_{\mathbb{R}^5} \mathcal{H}^0(\Psi^{-1}(Z)) \, d\mathcal{H}^4(Z),$$

where  $\mathbf{J}_4 \Psi(X, \tilde{X}) = \sqrt{\sum_{\lambda \in \Lambda(5,4)} [\det \nabla \Psi_\lambda(X, \tilde{X})]^2}$  is calculated by the well-known Cauchy-Binet formula. In fact it is easy to see that  $\Psi$  is injective and

$$\int_E \mathbf{J}_4 \Psi(X, \tilde{X}) \, d\mathcal{L}^4(X, \tilde{X}) = \mathcal{H}^4(\Psi(E)) \approx \mu(\Psi(E)).$$

When the arguments satisfy  $|X - \tilde{X}| \approx 2^{-K}$  then  $\Psi$  is in fact  $2^K$  bi-Lipschitz. It is not difficult to calculate under these circumstances that  $\mathbf{J}_4 \Psi \approx 2^K$ . In particular

$$(4.13) \quad \mu(\Psi(N)) = 0 \text{ exactly when } \mathcal{L}^4(N) = 0$$

and in the following we use this property repeatedly.

From [2, Theorem 3.107] we have that for any choice of  $|u| = 1$ ,  $\mathcal{L}^2$  almost every choice of  $X \in Q(V, 2^{-K-2})$ , and  $\mathcal{H}^1$  almost every choice of  $\tilde{X} \in Q(\tilde{V}, 2^{-K-2})$  with  $\tilde{X} \in X + \mathbb{R}u$  there exists a partial derivative in the direction  $u$  at  $X$  and at  $\tilde{X}$  and therefore  $f_{|[X, \tilde{X}]}$  is continuous at  $X$  (and  $\tilde{X}$ ). Call  $N_2$  the set of all  $(X, \tilde{X}) \in Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2})$  such that either there is no partial derivative in the direction  $\frac{X - \tilde{X}}{|X - \tilde{X}|}$  at  $X$  or there is no partial derivative in the direction  $\frac{X - \tilde{X}}{|X - \tilde{X}|}$  at  $\tilde{X}$ . We have from the above that  $\mu(\Psi(N_2)) = 0$ . Then (4.13) implies that  $\mathcal{L}^4(N_2) = 0$ .

As a step towards proving (3) we show that the pairs of  $X, \tilde{X}$  whose corresponding segments intersect  $\Theta_f^\beta$  at non-Lebesgue points of the map in (3) has zero  $\mathcal{L}^4$  measure. Although this is a standard result of structure theory we give some details here. We denote the so-called jump set of  $f$  as  $J_f$ . The set  $\Theta_f^0 \supset \Theta_f^\beta$  is a superset of  $J_f$  (see

[2, Proposition 3.92]) and  $\mathcal{H}^1(\Theta_f^0 \setminus J_f) = 0$  implying that  $D^j f = (f^+ - f^-) \otimes u \mathcal{H}_{|\Theta_f^0}^1$  (see [2, Lemma 3.76, Theorem 3.77]). Then, since  $|f^+(x) - f^-(x)| > \frac{1}{10}\beta$  for  $\mathcal{H}^1$  almost every  $x \in \Theta_f^\beta$ , we have that  $\Theta_f^\beta$  is both 1-rectifiable and  $\mathcal{H}^1(\Theta_f^\beta) < \infty$  (see Theorem 2.8). Also we have that  $f^+ - f^- \in L^1(\Theta_f^\beta, \mathcal{H}^1)$ . Therefore  $\mathcal{H}^1$  almost every point of  $\Theta_f^\beta$  is a Lebesgue point of  $f^+ - f^-$  with respect to  $\mathcal{H}_{|\Theta_f^0}^1$ . That is, calling  $T_1$  the set of non-Lebesgue points of  $f^+ - f^-$  with respect to  $\mathcal{H}_{|\Theta_f^0}^1$ , we have  $\mathcal{H}^1(T_1) = 0$ . This in turn implies that for any direction  $|u| = 1$  and its corresponding projection  $\pi_u(\cdot) := \cdot - u\langle \cdot, u \rangle$  we have that  $\mathcal{H}^1(\pi_u(T_1)) = 0$ . Let us call  $N_3$  the set of  $(X, \tilde{X})$  such that  $[X\tilde{X}]$  intersects  $T_1$ . Since  $\mathcal{H}^1\left(\pi_{\frac{X-\tilde{X}}{|X-\tilde{X}|}}(T_1)\right) = 0$  for every possible value of  $\frac{X-\tilde{X}}{|X-\tilde{X}|}$ , the Fubini theorem gives that  $\mu(\Psi(N_3)) = 0$  and, by (4.13),  $\mathcal{L}^4(N_3) = 0$ .

Now we show that  $\mathcal{L}^4$  almost every choice of  $X$  and  $\tilde{X}$  does not meet  $\Theta_f^\beta$  tangentially, which is a claim of point (3). The set of directions  $|u| = 1$  such that

$$\mathcal{H}^1(\{w \in \Theta_f^\beta : v(w) \perp u\}) > 0$$

is at most countable and so has zero measure since  $\mathcal{H}^1(\Theta_f^\beta) < \infty$ . We call this set of directions  $T_2$ . Therefore, for almost every direction  $|u| = 1$  we have that

$$\mathcal{H}^1\left(\pi_u(Q(V, 2^{-K-2})) \cap \pi_u(\{w \in \Theta_f^\beta : v(w) \perp u\})\right) = 0$$

and the same estimate holds after replacing  $V$  with  $\tilde{V}$ . Choose any direction  $|u| = 1$  with  $u \notin T_2$  such that  $\mathcal{H}^1(\pi_u(Q(V, 2^{-K-2})) \cap \pi_u(Q(\tilde{V}, 2^{-K-2}))) > 0$ . We have that the set of points  $\hat{X}$  such that  $\hat{X} \in \pi_u(Q(V, 2^{-K-2})) \cap \pi_u(Q(\tilde{V}, 2^{-K-2}))$  and  $\hat{X} \in \pi_u(\{w \in \Theta_f^\beta : v(w) \perp u\})$  has  $\mathcal{H}^1$  measure equal zero. This holds for any vector  $|u| = 1$ ,  $u \notin T_2$ . Call  $N_4$  the set of pairs  $(X, \tilde{X})$  such that  $\frac{X-\tilde{X}}{|X-\tilde{X}|} \in T_2$  or

$$\pi_{\frac{X-\tilde{X}}{|X-\tilde{X}|}}(X) \in \pi_{\frac{X-\tilde{X}}{|X-\tilde{X}|}}\left(\{x \in \Theta_f^\beta : v(x) \perp \frac{X-\tilde{X}}{|X-\tilde{X}|}\}\right).$$

The Fubini theorem and (4.13) guarantee that  $\mathcal{L}^4(N_4) = 0$ .

Let us fix a direction  $|u| = 1$  then almost every line  $L$  parallel to  $u$  has  $|Df|(L \cap Q(0, 1)) = 0$ . Therefore using the bi-Lipschitz quality of  $\Psi$  we have that the set of pairs  $(X, \tilde{X})$  (call it  $N_5$ ) such that  $|Df|([X\tilde{X}]) > 0$  satisfies  $\mathcal{L}^4(N_5) = 0$ .

Let  $|u| = 1$  be any vector such that  $P_{u, V, \tilde{V}, K} := \pi_u(Q(V, 2^{-K-2})) \cap \pi_u(Q(\tilde{V}, 2^{-K-2})) \neq \emptyset$ . For any  $p \in P_{u, V, \tilde{V}, K}$  let us also denote  $M_{u, V, K, p} := \pi_u^{-1}(p) \cap Q(V, 2^{-K-2})$  and  $M_{u, \tilde{V}, K, p} := \pi_u^{-1}(p) \cap Q(\tilde{V}, 2^{-K-2})$ . Then, (because each  $[X, \tilde{X}] \subset 2Q_i$ ), we have by the Fubini theorem, [2, Theorem 3.107] and Lemma 2.10

$$\begin{aligned} \int_{P_{u, V, \tilde{V}, K}} \int_{M_{u, V, K, U} \times M_{u, \tilde{V}, K, U}} |D_\tau f|_{|[X\tilde{X}]|}([X\tilde{X}]) d\mathcal{H}^1 \times \mathcal{H}^1(X, \tilde{X}) d\mathcal{H}^1(U) \\ \leq C2^{-2K} |\langle Df, u \rangle|(2Q_i) \\ \leq C2^{-2K} |Df|(2Q_i). \end{aligned}$$

Integrating this with respect to  $u$  and then using the change of variables formula with  $\Psi$  (note that  $\mathbf{J}_4\Psi \approx 2^K$ ) we get

$$\int_{Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2})} |D_\tau f|_{[X\tilde{X}]} |([X\tilde{X}])| d\mathcal{L}^4(X, \tilde{X}) \leq C 2^{-3K} |Df|(2Q_i).$$

Call  $N := N_1 \cup N_2 \cup N_3 \cup N_4 \cup N_5$ , then  $\mathcal{L}^4(N) = 0$ . Using the Chebyshev inequality we find a constant  $\lambda > 0$  (dependent only on  $f$ ) and a set  $H_{V, \tilde{V}} \subset Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2}) \setminus N$  such that  $\mathcal{L}^4(H_{V, \tilde{V}}) \geq \frac{4}{5} \mathcal{L}^4(Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2}))$  and

$$|D_\tau f|_{[X\tilde{X}]} |([X\tilde{X}])| \leq C \lambda 2^K |Df|(2Q_i)$$

for any  $(X, \tilde{X}) \in H_{V, \tilde{V}}$ , this is (4.10).

Now let us prove (4.11) and (4.12). We assume that  $Q_j = Q(c_j, 2^{-K}) \in G_{\varepsilon, \alpha, K} \cup T_{\varepsilon, \alpha, K}$ . Specifically by Theorem 4.2 there exists  $w_j \in Q(c_j, 2^{-K-2})$

$$(4.14) \quad \int_{Q(c_j, 2^{-K})} |\nabla f(z) - \nabla f(w_j)| d\mathcal{L}^2(z) \leq \varepsilon \alpha^2 2^{-2K},$$

and

$$(4.15) \quad |D^s f|(2Q_j) \leq \varepsilon |D^a f|(Q_j)$$

and finally that there exists a set  $Z_{j, \alpha} \subset Q(c_j, 2^{1-K})$  with  $\mathcal{L}^2(Q(c_j, 2^{1-K}) \setminus Z_{j, \alpha}) \leq 2^{-2K-9}$  such that

$$\|f(\cdot) - f(w_j) - \nabla f(w_j)(\cdot - w_j)\|_{L^\infty(Z_{j, \alpha})} < \alpha^4 2^{-K}$$

and

$$\alpha_0 \leq |\nabla f(w_j)| \leq \alpha_0^{-1}, \quad \alpha_0 < \det \nabla f(w_j).$$

We integrate over all lines parallel to  $|u| = 1$  a vector such that  $P_{u, V, \tilde{V}, K} := \pi_u(Q(V, 2^{-K-2})) \cap \pi_u(Q(\tilde{V}, 2^{-K-2})) \neq \emptyset$  and over  $M_{u, V, K, U} = \pi_u^{-1}(U) \cap Q(V, 2^{-K-2})$  and  $M_{u, \tilde{V}, K, U} = \pi_u^{-1}(U) \cap Q(\tilde{V}, 2^{-K-2})$ . First we decompose into the singular and absolutely continuous part

$$\begin{aligned} & \int_{P_{u, V, \tilde{V}, K}} \int_{M_{u, V, K, U}} \int_{M_{u, \tilde{V}, K, U}} |D_\tau [f(\cdot) - \nabla f(w_j)(\cdot)]|_{[X\tilde{X}]} |([X\tilde{X}])| d\tilde{X} dX d\mathcal{H}^1(U) \\ & \leq \int_{P_{u, V, \tilde{V}, K}} \int_{M_{u, V, K, U}} \int_{M_{u, \tilde{V}, K, U}} \int_{[X\tilde{X}]} |\nabla f(z) - \nabla f(w_j)| d\mathcal{H}^1(z) d\tilde{X} dX d\mathcal{H}^1(U) \\ & \quad + \int_{P_{u, V, \tilde{V}, K}} \int_{M_{u, V, K, U}} \int_{M_{u, \tilde{V}, K, U}} \langle |D^s f|, u \rangle ([X\tilde{X}]) d\tilde{X} dX d\mathcal{H}^1(U). \end{aligned}$$

Now we use Fubini and the fact that  $\{X \in Q(V, 2^{-K-2}) : \pi_u(X) \in P_{u, V, \tilde{V}, K}\}, \{\tilde{X} \in Q(\tilde{V}, 2^{-K-2}) : \pi_u(\tilde{X}) \in P_{u, V, \tilde{V}, K}\} \subset 2Q_i$ . Also we use the estimate that every slice of

$Q(V, 2^{-K-2})$  has diameter bounded by  $2^{-K}$  to get

$$\begin{aligned} & \int_{P_{u,V,\tilde{V},K}} \int_{M_{u,V,K,U}} \int_{M_{u,\tilde{V},K,U}} \int_{[X\tilde{X}]} |\nabla f(z) - \nabla f(w_j)| d\mathcal{H}^1(z) d\tilde{X} dX d\mathcal{H}^1(U) \\ & + \int_{P_{u,V,\tilde{V},K}} \int_{M_{u,V,K,U}} \int_{M_{u,\tilde{V},K,U}} \langle |D^s f|, u \rangle ([X\tilde{X}]) d\tilde{X} dX d\mathcal{H}^1(U) \\ & \leq 2^{-2K} \int_{2Q_i} |\nabla f - \nabla f(w_i)| + 2^{-2K} |D^s f|(2Q_i) \end{aligned}$$

We use (4.14) and (4.15) and then  $\alpha < \alpha_0 \leq |\nabla f(w_i)|$  to get

$$\begin{aligned} 2^{-2K} \int_{2Q_i} |\nabla f - \nabla f(w_i)| + 2^{-2K} |D^s f|(2Q_i) \\ \leq \varepsilon \alpha^2 2^{-4K} + C 2^{-2K} \varepsilon |D^a f|(Q_i) \\ \leq C \varepsilon \alpha^2 2^{-4K} + C 2^{-4K} \varepsilon |\nabla f(w_i)| + C \varepsilon \alpha^2 2^{-4K} \\ \leq C 2^{-4K} \varepsilon |\nabla f(w_i)|. \end{aligned}$$

Integrating the above equations over all  $|u| = 1$  and using the change of variables formula with  $\Psi$  (recall that  $\mathbf{J}_4 \Psi \approx 2^K$ ) and denoting  $A = (Q(V, 2^{-K-2}) \cap Z_{j,\alpha}) \times (Q(\tilde{V}, 2^{-K-2}) \cap Z_{j,\alpha})$  we get

$$\int_A |D_\tau[f(\cdot) - \nabla f(w_j)(\cdot)]|_{[X\tilde{X}]}([X\tilde{X}]) d\mathcal{L}^4(X, \tilde{X}) \leq C \varepsilon 2^{-3K} |D^a f|(2Q_i).$$

Because  $\mathcal{L}^2(Q(c_j, 2^{1-K}) \setminus Z_{j,\alpha}) \leq 2^{-2K-9}$ , we have that

$$\begin{aligned} \mathcal{L}^2(Q(V, 2^{-K-2}) \cap Z_{j,\alpha}) & \geq \mathcal{L}^2(Q(V, 2^{-K-2})) - \mathcal{L}^2(Q(c_j, 2^{1-K}) \setminus Z_{j,\alpha}) \\ & \geq 2^{-2K-4} - 2^{-2K-9} = \frac{31}{32} 2^{-2K-4} \\ & = \frac{31}{32} \mathcal{L}^2(Q(V, 2^{-K-2})). \end{aligned}$$

This implies that

$$\mathcal{L}^4\left((Q(V, 2^{-K-2}) \cap Z_{j,\alpha}) \times (Q(\tilde{V}, 2^{-K-2}) \cap Z_{j,\alpha})\right) \geq \frac{9}{10} \mathcal{L}^4(Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2})).$$

Again we have, up to increasing the value of  $\lambda$ , that

$$|D_\tau[f(\cdot) - \nabla f(w_j)(\cdot)]|_{[X\tilde{X}]}([X\tilde{X}]) \leq C \lambda \varepsilon 2^K |D^a f|(2Q_i)$$

for all  $(X, \tilde{X}) \in H_{V,\tilde{V}} \subset (Z_{j,\alpha} \times Z_{j,\alpha})$  while simultaneously  $\mathcal{L}^4(H_{V,\tilde{V}}) \geq \frac{4}{5} \mathcal{L}^4(Q(V, 2^{-K-2}) \times Q(\tilde{V}, 2^{-K-2}))$ , thus proving (4.11) and concluding our proof.  $\square$

In the following proposition we use the following notation. Let  $\mathcal{P} \subset \mathbb{R}^2$  be an injective continuous image of a circle with  $\mathcal{H}^1(\mathcal{P}) < \infty$ . Denote the closure of the bounded component of  $\mathbb{R}^2 \setminus \mathcal{P}$  by  $\tilde{\mathcal{P}}$ . Let  $p_1, p_2 \in \tilde{\mathcal{P}}$ , we define

$$(4.16) \quad d_{\mathcal{P}}(p_1, p_2) = \inf \left\{ l(\gamma) : \gamma \text{ is a path joining } p_1, p_2; \gamma \subset \tilde{\mathcal{P}} \right\},$$

where by a path joining  $p_1$  and  $p_2$  in  $\tilde{\mathcal{P}}$  we mean a continuous curve  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  such that  $\gamma(0) = p_1$  and  $\gamma(1) = p_2$ .

The following proposition is the utilization of the  $NCBV^+$  condition on a good non-straight grid  $\Gamma$  chosen using the previous theorem. It gives us a map  $\varphi$  defined

on  $\Gamma$ . The utility of  $\varphi$  is that we are able to find a homeomorphic extension of  $\varphi$  with estimates that allow us to prove area-strict convergence.

**Proposition 4.4.** *For every  $\varepsilon > 0$  let  $\alpha_0 > 0$  as in Theorem 4.2. For every  $0 < \alpha < \alpha_0$  there exists a good non-straight grid for  $f$  called  $\Gamma$  and a function  $\varphi$  defined on  $\Gamma$  such that*

- (1)  $\Gamma$  is admissible for  $f$ ,
- (2) every component of  $Q(0, 1) \setminus \Gamma$  is a convex quadrilateral and contains exactly one point  $c_i$ , where  $c_i$  is the centre of a square  $Q_i$  the set  $\{Q_i\}_{i=1}^{2^{2K}}$  of  $K$ -dyadic squares in Theorem 4.2 and Theorem 4.3 (thanks to this we call the quadrilaterals  $\mathcal{Q}_i$  the components of  $Q(0, 1) \setminus \Gamma$  containing  $c_i$ ),
- (3) if  $Q_i \in E_{\varepsilon, K}$  (the set from Theorem 4.2) and finding  $u_i, v_i$  such that that the function  $g$  of Theorem 4.2 satisfies  $g(w_i) = u_i \otimes v_i$  we have

$$(4.17) \quad \int_{\pi_{v_i}(\mathcal{Q}_i)} d_{\varphi(\partial\mathcal{Q}_i)}(\varphi(X_*), \varphi(X^*)) d\mathcal{H}^1 \leq (1 + \varepsilon)|Df|(\mathcal{Q}_i) + C\varepsilon 2^{-2K}$$

and

$$(4.18) \quad \int_{\pi_{v_i^\perp}(\mathcal{Q}_i)} d_{\varphi(\partial\mathcal{Q}_i)}(\varphi(Z^*), \varphi(Z_*)) d\mathcal{H}^1(Z) \leq C\varepsilon|D^s f|(4Q_i \cap S) + C2^{-2K}\varepsilon.$$

where  $\pi_{v_i}(x) = v_i^\perp \langle x, v_i^\perp \rangle$  and  $\pi_{v_i^\perp}(x) = v_i \langle x, v_i \rangle$  and where for each  $X \in \pi_{v_i}(\mathcal{Q}_i)$  the points  $X_*, X^*$  are the two distinct points in  $\partial\mathcal{Q}_i$  such that  $\pi_{v_i}(X_*) = \pi_{v_i}(X^*) = X$  and similarly  $\pi_{v_i^\perp}(Z_*) = \pi_{v_i^\perp}(Z^*) = Z \in \pi_{v_i^\perp}(\mathcal{Q}_i)$ ,

- (4) for all  $i = 1, \dots, 2^{2K}$  it holds that

$$|D_\tau \varphi|(\partial\mathcal{Q}_i) \leq C2^K |Df|(2Q_i),$$

- (5) if  $Q_i \in G_{\varepsilon, \alpha, K}$  then  $\varphi(x, y) = f(x, y)$  at each  $(x, y)$  vertex of  $\mathcal{Q}_i$  and  $\varphi$  is linear on each side of  $\partial\mathcal{Q}_i$
- (6) if  $Q_i \in T_{\varepsilon, \alpha, K}$  then

$$(4.19) \quad \int_{\partial\mathcal{Q}_i} |\partial_\tau \varphi - \nabla f(w_i)\tau| d\mathcal{H}^1 \leq C\varepsilon 2^K |Df|(2Q_i).$$

*Proof.*

### Step 1. Choice of $\Gamma$ .

Each vertex  $V$  of each square  $Q_i$  has at most four neighbouring vertexes, call them  $\tilde{V}_1, \dots, \tilde{V}_4$ . By the Fubini theorem we have the existence of an  $X \in Q(V, 2^{-K-2})$  such that  $\mathcal{L}^2(\{\tilde{X} : (X, \tilde{X}) \in H_{V, \tilde{V}_j}\}) \geq \frac{4}{5}\mathcal{L}^2(Q(\tilde{V}_j, 2^{-K-2}))$  holds for all  $j = 1, 2, 3, 4$  simultaneously. It follows that it is possible to choose  $X_V \in Q(V, 2^{-K-2})$  for each  $V$  such that for every pair of neighbours  $V$  and  $\tilde{V}$  we have that  $(X_V, X_{\tilde{V}}) \in H_{V, \tilde{V}}$ . The squares  $Q_i$  can be described as the convex hull of their vertexes  $V_1, \dots, V_4$  and the corresponding quadrilateral  $\mathcal{Q}_i$  is the convex hull of  $X_{V_1}, \dots, X_{V_4}$ . By definition it is not hard to check that  $\{\mathcal{Q}_i\}$  are pairwise disjoint outside their mutual boundaries, they are convex quadrilaterals and  $X_V$  lies in the boundary of  $\mathcal{Q}_i$  exactly when  $V$  is a vertex of  $Q_i$ . This is point (2). Theorem 4.3 and  $(X_V, X_{\tilde{V}}) \in H_{V, \tilde{V}}$  guarantees point (1) of our claim. For every unit vector  $v$  the set  $\{(X, \tilde{X}) : \frac{X - \tilde{X}}{|X - \tilde{X}|} = v\}$  has Hausdorff

dimension 3 and therefore has  $\mathcal{L}^4$  measure 0. Therefore it is not restrictive to assume that the sides of  $\mathcal{Q}_i$  are not parallel to  $v_i$  or  $v_i^\perp$ .

In order to prove points (3)-(6) we need to define the map  $\varphi$ . Before we start the construction of  $\varphi$  itself we add extra lines to the grid  $\Gamma$  to get an augmented grid  $\tilde{\Gamma}$ . Although we do not need to use  $f$  on  $\tilde{\Gamma} \setminus \Gamma$  we use the extra lines added to get a good parametrization of the geometric representative of  $f$  on  $\Gamma$ . This is equivalent to the concept of guidelines from [10].

**Step 2. Construction of  $\tilde{\Gamma}$  by the choice of guidelines.**

For each  $Q_i \in E_{\alpha,K}$ , by (4.4), we have  $|u_i| = |v_i| = 1$  such that

$$\int_{S \cap 2Q_i} |g(z) - u_i \otimes v_i| d|D^s f|(z) \leq \varepsilon |D^s f|(4Q_i \cap S).$$

For almost every  $X \in \pi_{v_i}(\mathcal{Q}_i)$  we have  $f|_{[X + \mathbb{R}v_i] \cap Q(0,1)}$  in  $BV$  on  $[X + \mathbb{R}v_i] \cap Q(0,1)$ . The corresponding claim holds for almost every  $Z \in \pi_{v_i^\perp}(\mathcal{Q}_i)$ . It follows from the  $BV$  on lines characterization and Lemma 2.10 that

$$|\langle Df, v_i \rangle|(\mathcal{Q}_i) = \int_{\pi_{v_i}(\mathcal{Q}_i)} |D_\tau f|_{[X + \mathbb{R}v_i] \cap \mathcal{Q}_i}|([X + \mathbb{R}v_i] \cap Q(0,1)) d\mathcal{H}^1(X)$$

and, by (4.4),

$$|\langle Df, v_i^\perp \rangle|(\mathcal{Q}_i) = \int_{\pi_{v_i^\perp}(\mathcal{Q}_i)} |D_\tau f|_{[Z + \mathbb{R}v_i^\perp] \cap \mathcal{Q}_i}|([Z + \mathbb{R}v_i^\perp] \cap Q(0,1)) d\mathcal{H}^1(Z) \leq \varepsilon |D^s f|(4Q_i).$$

Since  $\Gamma$  is admissible for  $f$  we have  $f|_\Gamma$  is  $BV$  on  $\Gamma$ . Therefore for each pair,  $V_1, V_2$  of neighbouring vertices of  $\mathcal{Q}_i$

$$(4.20) \quad \text{there exists a finite set } J_{V_1, V_2} = \{X \in [V_1 V_2] : |D_\tau f|_{[V_1 V_2]}(\{X\}) > \varepsilon 2^{-2K}\}.$$

We call the cardinality of this finite set  $\mathfrak{K}$ .

We now chop  $\mathcal{Q}_i$  into slices parallel to  $v_i$ . Let us have a locally finite decomposition of  $\mathbb{R}$  into pairwise disjoint intervals indexed by  $m \in \mathbb{Z}$  called  $I_m$ . Then we define  $S_{i,m} = \mathcal{Q}_i \cap \{(x, y) \in \mathbb{R}^2 : \langle (x, y), v_i^\perp \rangle \in I_m\}$ . Because the sides of  $\mathcal{Q}_i$  are not parallel to  $v_i$  it holds that for every  $X \in [\pi_{v_i}(\mathcal{Q}_i)]^\circ$  there are exactly two distinct points  $X_*, X^* \in \partial\mathcal{Q}_i$  such that  $\pi_{v_i}(X_*) = \pi_{v_i}(X^*) = X$ . For simplicity denote the points  $X_*$  and  $X^*$  such that  $\langle X_*, v_i \rangle < \langle X^*, v_i \rangle$ . In fact, the sides of  $\partial\mathcal{Q}$  are separated into two categories, either the points of the side are all  $X^*$ -type points (an upper side) or all the points of the side are  $X_*$ -type points (a lower side). Then we denote  $S_{i,m}^+$  the set of points  $X^* \in \partial\mathcal{Q}_i \cap S_{i,m}$  that lie on an ‘‘upper’’ side (with respect to  $v_i$ ). Similarly we call  $S_{i,m}^-$  the set of points  $X^* \in \partial\mathcal{Q}_i \cap S_{i,m}$  that lie on a ‘‘lower’’ side (with respect to  $v_i$ ). By choosing the intervals  $I_m$  carefully we can achieve that either

$$(4.21) \quad |D_\tau f|_\Gamma|(\partial\mathcal{Q}_i \cap S_{i,m}) < \varepsilon 2^{-2K} \quad \text{or} \quad \mathcal{H}^1(S_{i,m}^\pm) < \frac{\varepsilon}{2^{2K} |D_\tau f|_\Gamma|(\Gamma) \mathfrak{K}}$$

for every  $i, m$ . We may assume also that  $S_{i,m}^\pm$  is always a segment.

The argument from the previous paragraph can be repeated in the perpendicular direction  $v_i^\perp$  (and if necessary increase the number  $\mathfrak{K}$ ). This gives us a finite number

of segments  $T_{i,m}^\pm$  covering  $\partial\mathcal{Q}_i$  and either

$$|D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap T_{i,m}) < \varepsilon 2^{-2K} \quad \text{or} \quad \mathcal{H}^1(T_{i,m}^\pm) < \frac{\varepsilon}{2^{2K}|D_\tau f|_\Gamma(\Gamma)\mathfrak{R}}.$$

As before we separate  $\partial\mathcal{Q}_i$  into upper and lower sides (with respect to  $v_i^\perp$ ) and points on lower sides we denote by  $Z_*$  and points on upper sides we denote by  $Z^*$ . Note that even if  $S_{i,m}^- = T_{i,m}^-$  then  $S_{i,m}^+ \neq T_{i,m}^+$ .

By Remark 3.7, we have for  $\mathcal{H}^1$ -almost every  $X \in \pi_{v_i}(\mathcal{Q}_i)$  that

$$(4.22) \quad \Gamma \cup ([X + v_i\mathbb{R}] \cap \mathcal{Q}_i) \cup ([Z + v_i^\perp\mathbb{R}] \cap \mathcal{Q}_i) \text{ is admissible for } f, \text{ for } \mathcal{H}^1\text{-almost every } Z.$$

Then for every  $S_{i,m}$  such that  $|D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap S_{i,m}) < \varepsilon 2^{-2K}$  we find an  $X_{i,m}$  satisfying (4.22) and such that

$$(4.23) \quad \mathcal{H}^1(\pi_{v_i}(S_{i,m}))|D_\tau f|_{[X_{i,m} + \mathbb{R}v_i] \cap \mathcal{Q}_i}|([X_{i,m} + \mathbb{R}v_i] \cap \mathcal{Q}_i) \leq |D_{v_i} f|(S_{i,m} \cap \mathcal{Q}_i).$$

By  $X_{i,m}^*$  we denote the point in  $[X_{i,m} + \mathbb{R}v_i] \cap S_{i,m}^+$  and by  $X_{i,m,*}$  we denote the point in  $[X_{i,m} + \mathbb{R}v_i] \cap S_{i,m}^-$  (for clarification see Figure 4).

Similarly, for each  $T_{i,m}$  we find an  $Z_{i,m} \in T_{i,m}$  such that

$$(4.24) \quad \Gamma \cup \bigcup_{m,i} ([X_{i,m} + v_i\mathbb{R}] \cap \mathcal{Q}_i) \cup ([Z_{i,m} + v_i^\perp\mathbb{R}] \cap \mathcal{Q}_i)$$

is admissible for  $f$  and

$$(4.25) \quad \mathcal{H}^1(\pi_{v_i^\perp}(T_{i,m}))|D_\tau f|_{[Z_{i,m} + \mathbb{R}v_i^\perp] \cap \mathcal{Q}_i}|([Z_{i,m} + \mathbb{R}v_i^\perp] \cap \mathcal{Q}_i) \leq |D_{v_i^\perp} f|(T_{i,m} \cap \mathcal{Q}_i)$$

for every  $T_{i,m}$  such that  $|D_\tau f|_\Gamma(T_{i,m}) < \varepsilon 2^{-2K}$ . The choice of  $X_{i,m}, Z_{i,m}$  will be exactly what we need to get the estimate (4.17) and (4.18). To fix a temporary notation, we call  $\hat{\Gamma}$  the good non-straight grid for  $f$  which is constructed by repeating the process in (4.24) for every  $\mathcal{Q}_i$  such that  $\mathcal{Q}_i \in E_{\varepsilon,K}$ .

The Figure 4 can help orient the reader in the following construction. For  $S_{i,m}$  such that  $|D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap S_{i,m}) \geq \varepsilon 2^{-2K}$  we find pairs of points  $X_{i,m}^+, X_{i,m}^- \in \pi_{v_i}(S_{i,m})$  satisfying (4.22) such that the sets

$$S_{i,m}^l = \{(x, y) \in S_{i,m} : \langle (x, y), v_i^\perp \rangle \leq \langle X_{i,m}^-, v_i^\perp \rangle\}$$

and

$$S_{i,m}^r = \{(x, y) \in S_{i,m} : \langle (x, y), v_i^\perp \rangle \geq \langle X_{i,m}^+, v_i^\perp \rangle\}$$

satisfy

$$(4.26) \quad |D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap \pi_{v_i}^{-1}(S_{i,m}^l)) < \varepsilon 2^{-2K} \quad \text{and} \quad |D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap \pi_{v_i}^{-1}(S_{i,m}^r)) < \varepsilon 2^{-2K}.$$

Similarly, for  $T_{i,m}$  such that  $|D_\tau f|_\Gamma(\partial\mathcal{Q}_i \cap T_{i,m}) \geq \varepsilon 2^{-2K}$  we find pairs of points  $Z_{i,m}^+, Z_{i,m}^- \in \pi_{v_i}(T_{i,m})$  such that

$$\tilde{\Gamma} := \hat{\Gamma} \cup \bigcup_m ([X_{i,m}^\pm + v_i\mathbb{R}] \cap \mathcal{Q}_i) \cup ([Z_{i,m}^\pm + v_i^\perp\mathbb{R}] \cap \mathcal{Q}_i)$$

is admissible for  $f$  and such that the sets

$$T_{i,m}^l = \{(x, y) \in \mathbb{R}^2 : \langle (x, y), v_i \rangle \leq \langle Z_{i,m}^-, v_i \rangle\}$$

and

$$T_{i,m}^r = \{(x, y) \in \mathbb{R}^2 : \langle (x, y), v_i \rangle \geq \langle Z_{i,m}^+, v_i \rangle\}$$



satisfy

$$(4.27) \quad |D_\tau f|_\Gamma|(\partial Q_i \cap \pi_{v_i^\pm}^{-1}(T_{i,m}^l)) < \varepsilon 2^{-2K} \quad \text{and} \quad |D_\tau f|_\Gamma|(\partial Q_i \cap \pi_{v_i^\pm}^{-1}(T_{i,m}^r)) < \varepsilon 2^{-2K}.$$

As argued above we have that almost every choice of  $X_{i,m}^\pm$  permits almost any choice of  $Z_{i,m}^\pm$  and so there is no obstacle in choosing  $X_{i,m}^\pm$  and  $Z_{i,m}^\pm$  for every  $m$  and for each  $Q_i$  so that the resulting set  $\tilde{\Gamma}$  is a good non-straight grid for  $f$ .

### Step 3. Injective approximations of the geometric representative of $f$ .

By  $\hat{\varphi}$  we denote the geometric representative of  $f$  on  $\tilde{\Gamma}$  as defined in (3.5) and its following paragraphs. Then Lemma 3.12 provides a good arrival grid,  $\mathcal{G}$ , associated with  $\tilde{\Gamma}$  and  $\hat{\varphi}$  with side length  $\kappa = \frac{\varepsilon}{2^{K+1}}$  in the sense of Definition 3.11.

By the definition of the good arrival grid  $\mathcal{G}$ , the set  $P := \hat{\varphi}^{-1}(\mathcal{G}) \cap \tilde{\Gamma}$  is finite and does not contain any vertices of  $\tilde{\Gamma}$ . Nor does  $\hat{\varphi}(\tilde{\Gamma})$  intersect any vertex of  $\mathcal{G}$  (the points denoted as  $(w_n, z_m)$  in Definition 3.11). Further, for every point  $(x, y) \in P$ , it holds that the derivative of  $\hat{\varphi}$  at  $(x, y)$  tangential to  $\tilde{\Gamma}$  (we denote it as  $\partial_\tau \hat{\varphi}(x, y)$ ) has non-zero component perpendicular to the side of  $\mathcal{G}$  containing  $\hat{\varphi}(x, y)$  (the existence of  $\partial_\tau \hat{\varphi}$  at all points of  $P$  is a requirement of the good arrival grid, see Definition 3.11). Therefore there exists a smallest perpendicular component whose size is  $v > 0$ .

For each point  $a \in P$  we have some  $d_a > 0$  such that when  $(x, y) \in \tilde{\Gamma}$  and  $|(x, y) - a| < d_a$  then

$$\hat{\varphi}(x, y) - \hat{\varphi}(a) - \partial_\tau \hat{\varphi}(a)[(x, y) - a] < \frac{v}{3}|(x, y) - a|.$$

Since  $P$  is finite we define  $d := \min_{a \in P} d_a > 0$ . As a consequence, the images through  $\hat{\varphi}$  of the endpoints of the segments  $B(a, d) \cap \tilde{\Gamma}$  have distance at least  $\frac{vd}{2}$  from  $\hat{\varphi}(a)$ . By making  $d$  smaller if necessary we can assume that each  $B(a, d) \cap \tilde{\Gamma}$  is a segment and each pair of these (finitely many) segments is disjoint neither does any of the segments contain any vertex of  $\tilde{\Gamma}$ . The choice of the number  $d$  has been made so that the following holds; let  $c$  be an endpoint of the segment of  $B(a, d) \cap \tilde{\Gamma}$  then  $|\hat{\varphi}(c) - \hat{\varphi}(a)| \geq \frac{vd}{2}$  and this holds for all  $a \in P$ .

On the other hand, being  $\hat{\varphi}(\tilde{\Gamma} \setminus \bigcup_{a \in P} B(a, d))$  a closed set, it follows that there exists a  $\sigma_0 > 0$  such that

$$(4.28) \quad \text{dist} \left( \hat{\varphi} \left( \tilde{\Gamma} \setminus \bigcup_{a \in P} B(a, d) \right), \mathcal{G} \right) \geq 3\sigma_0.$$

We have that  $\tilde{\Gamma}$  is a good non-straight grid and so the  $NCBV^+$  condition enjoyed by  $f$  guarantees the existence of injective uniform approximations of  $\hat{\varphi}$ . Equation (4.28) immediately implies that for any  $0 < \sigma \leq \sigma_0$  and any  $\tilde{\varphi}_\sigma$ , continuous injective approximation of  $\hat{\varphi}$  with  $\|\tilde{\varphi}_\sigma - \hat{\varphi}\|_{\infty, \tilde{\Gamma}} \leq \sigma$ , it holds that  $\tilde{\varphi}_\sigma^{-1}(\mathcal{G}) \subset \bigcup_{a \in P} B(a, d) \cap \tilde{\Gamma}$ .

Let

$$\rho' = \min\{|\mathbf{a} - \mathbf{b}| : \mathbf{a} \in \hat{\varphi}(P), \mathbf{b} \text{ vertex of } \mathcal{G}\}$$

and

$$\rho'' = \min\{|\mathbf{a} - \mathbf{b}| : \mathbf{a} \neq \mathbf{b}, \mathbf{a}, \mathbf{b} \in \hat{\varphi}(P)\}$$

Finally, we set

$$\rho = \min\{\rho', \rho'', \sigma_0, \frac{1}{100}\}.$$

Notice that  $\rho$  is positive due to the properties of good arrival grid. Let

$$(4.29) \quad 0 < \sigma \leq \frac{\varepsilon^2 \rho}{12(2^K + 1)}.$$

Then by applying the  $NCBV^+$  condition to  $\hat{\varphi}$  we get a continuous injective  $\tilde{\varphi}_\sigma$  with  $\|\tilde{\varphi}_\sigma - \hat{\varphi}\|_{\infty, \Gamma} < \sigma$ .

We adjust the map  $\tilde{\varphi}_\sigma$  as follows. For each  $a \in P = \hat{\varphi}^{-1}(\mathcal{G}) \cap \tilde{\Gamma}$  we find the first and last point (i.e. the points furthest away from  $a$ ) on the segment  $B(a, d) \cap \tilde{\Gamma}$  (call them  $a^-$  and  $a^+$  respectively) such that  $\tilde{\varphi}_\sigma(a^\pm) \in B(\hat{\varphi}(a), 2\sigma)$ . Notice that (4.28) and the choice of  $\sigma$  imply that

$$\hat{\varphi}\left(\tilde{\Gamma} \setminus \bigcup_{a \in P} B(a, d)\right) \cap \left(\bigcup_{a \in P} B(\hat{\varphi}(a), 2\sigma)\right) = \emptyset,$$

hence we have that  $\tilde{\varphi}_\sigma$  intersects  $B(\hat{\varphi}(a), 2\sigma)$  only on the segments  $B(a, d) \cap \tilde{\Gamma}$ ,  $a \in P$ . We define

$$\tilde{\tilde{\varphi}}_\sigma(t) = \frac{|t - a^+|}{|a^+ - a^-|} \tilde{\varphi}_\sigma(a^-) + \frac{|t - a^-|}{|a^+ - a^-|} \tilde{\varphi}_\sigma(a^+) \quad \text{for all } t \in [a^- a^+] \text{ and } a \in P$$

then we set

$$\tilde{\tilde{\varphi}}_\sigma(t) = \tilde{\varphi}_\sigma(t) \quad \text{for } t \in \tilde{\Gamma} \setminus \bigcup_{a \in P} [a^- a^+].$$

By construction, it follows that  $\tilde{\tilde{\varphi}}_\sigma(t)$  is again continuous, injective and  $\|\tilde{\tilde{\varphi}}_\sigma - \hat{\varphi}\|_{L^\infty(\tilde{\Gamma})} \leq 7\sigma$ . Since  $\tilde{\tilde{\varphi}}_\sigma(a^-)$  and  $\tilde{\tilde{\varphi}}_\sigma(a^+)$  must be separated by  $\mathcal{G}$  there is exactly one point  $\tilde{a}$  in each  $[a^- a^+]$  which is mapped onto  $\mathcal{G} \cap B(\hat{\varphi}(a), 2\sigma)$ .

At this stage we use  $\tilde{\tilde{\varphi}}_\sigma$  and the arrival grid  $\mathcal{G}$  to define a piecewise linear map from  $\tilde{\Gamma}$  to  $\mathbb{R}^2$ . We will call this map  $\varphi$ . We start by specifying the image,  $\varphi(\tilde{\Gamma})$ . For each segment of  $\tilde{\Gamma}$  we have a finite number of points  $\tilde{a}$  such that  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}) \in \mathcal{G}$ . Whenever we have a pair of adjacent points  $\tilde{a}_1, \tilde{a}_2$  lying on a common segment of  $\tilde{\Gamma}$  such that  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_1)$  and  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_2)$  lie on two distinct sides of a rectangle in  $\mathcal{G}$  we define the segment  $S_{a_1, a_2} = [\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_1) \tilde{\tilde{\varphi}}_\sigma(\tilde{a}_2)]$  where  $a_1$  and  $a_2$  are the unique points in  $P$  for which  $\tilde{a}_i \in B(a_i, d)$ .

Let us now consider a pair of adjacent  $\tilde{a}_1$  and  $\tilde{a}_2$  for which  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_1), \tilde{\tilde{\varphi}}_\sigma(\tilde{a}_2)$  lie on the same side of a rectangle in  $\mathcal{G}$ . Firstly notice that for any such pair  $\tilde{a}_1$  and  $\tilde{a}_2$  there exists an  $0 < \xi_{a_1, a_2}$  so small that the generalized segments (see Definition 3.13) with  $\xi = \xi_{a_1, a_2}$  intersect only those previously defined straight segments  $S_{a_3, a_4}$  for which  $\tilde{\tilde{\varphi}}_\sigma([\tilde{a}_1 \tilde{a}_2])$  was already intersecting  $\tilde{\tilde{\varphi}}_\sigma([\tilde{a}_3 \tilde{a}_4])$ . We define

$$\xi = \frac{1}{2} \min \{ \xi_{a_1, a_2} : a_1, a_2 \in P \text{ adjacent and } f(a_1), f(a_2) \text{ lie on a common side of } \mathcal{G} \}.$$

We assume that  $\xi < \kappa < \varepsilon$  and define  $S_{a_1, a_2}$  as the generalised segment from  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_1)$  to  $\tilde{\tilde{\varphi}}_\sigma(\tilde{a}_2)$  with the chosen  $\xi$ .

It is very easy to check that any pair  $S_{a_1, a_2}$  and  $S_{a_3, a_4}$  as defined above, where  $a_1, a_2$  and  $a_3, a_4$  are pairs of adjacent points of  $P$  lying on a common segment of  $\tilde{\Gamma}$ , intersect each other if and only if  $\tilde{\tilde{\varphi}}_\sigma([\tilde{a}_1 \tilde{a}_2])$  intersects  $\tilde{\tilde{\varphi}}_\sigma([\tilde{a}_3 \tilde{a}_4])$ . Further, any two (distinct) paths can have at most one intersection.

Now we are in a position to define the map  $\varphi$  on  $\tilde{\Gamma}$ . We define  $\varphi(a) = \tilde{\tilde{\varphi}}_\sigma(\tilde{a})$  for all  $a \in P$  and for all the corresponding  $\tilde{a}$ . Further, for every  $\{X_{i,j}\} = \gamma_i([0, 1]) \cap \gamma_j([0, 1])$ ,

intersection point of the grid  $\tilde{\Gamma}$ , there exists exactly two pairs of adjacent  $a_1, a_2 \in P$  (both lying on  $\gamma_i$ ) and  $a_3, a_4 \in P$  (both lying on  $\gamma_j$ ) closest to  $X_{i,j}$  on  $\gamma_i$  and  $\gamma_j$  respectively. That is there exists a  $t_1, t_2, t_3, t_4 \in [0, 1]$  such that  $\gamma_i(t_1) = a_1, \gamma_i(t_2) = a_2, \gamma_j(t_3) = a_3, \gamma_j(t_4) = a_4$  further  $\gamma_i((t_1, t_2)) \cap P = \emptyset$  and  $\gamma_j((t_3, t_4)) \cap P = \emptyset$ .

Then, by construction, there exists exactly one point of intersection call it  $\mathbf{X}_{i,j}$  in the set  $S_{a_1, a_2} \cap S_{a_3, a_4}$  and we define  $\varphi(X_{i,j}) = \mathbf{X}_{i,j}$ . Thus we have separated the grid  $\tilde{\Gamma}$  into simple segments lying between adjacent intersections with  $\mathcal{G}$ , intersecting segments of  $\tilde{\Gamma}$  or a combination of the two. In each case there is a clear correspondence between the endpoints of remaining segments in  $\tilde{\Gamma}$  and (parts of the possibly generalized) segments defined in the previous paragraph. We define  $\varphi$  by parametrizing these segments (or possibly paths consisting of 2 segments) at constant speed from the corresponding segments in  $\tilde{\Gamma}$ .

Thus we obtain a continuous injective piecewise linear mapping  $\varphi : \tilde{\Gamma} \rightarrow Q(0, 1)$  satisfying  $\|\hat{\varphi} - \varphi\|_{L^\infty(\tilde{\Gamma})} \leq 4\kappa \leq \varepsilon 4(2^K + 1)^{-1}$ . The last step we make in order to define  $\varphi$  is to redefine it as linear on each side of  $\partial Q_i \in G_{\varepsilon, \alpha, K}$  keeping the same values at its vertices. By the definition of  $G_{\varepsilon, \alpha, K}$  (especially (4.6)) it is not hard to see that this modification keeps  $\varphi$  continuous, injective and piecewise linear. We refer to [7], Proof of Theorem 4.1, Step 3, the detailed argument. Thus we have achieved point (5) of the claim.

#### Step 4. Estimates.

We claim that for any pair of points  $a, b \in [ab] \subset \tilde{\Gamma}$  we have that

$$(4.30) \quad |D_\tau \varphi|([ab]) \leq (1 + \xi)(1 + \varepsilon) (|D_\tau \hat{\varphi}|([ab]) + 4\kappa).$$

Indeed it holds that  $|\hat{\varphi}(a) - \varphi(a)| \leq 3\sigma \leq \frac{1}{2}\varepsilon\rho$  for each  $a \in P$ . Immediately from the definition of  $\rho$  we obtain that for any pair  $a, a' \in P$  adjacent on a segment of  $\tilde{\Gamma}$  we have that either  $\hat{\varphi}(a) = \hat{\varphi}(a')$  or  $|\hat{\varphi}(a) - \hat{\varphi}(a')| \geq \rho$ . In the first case we know that the length of the curve given by  $\hat{\varphi}$  on  $[aa']$  is at least  $100\sigma$  by (4.28),  $\varepsilon < \frac{1}{100}$ , and  $\sigma \leq \varepsilon^2\rho \leq \varepsilon\sigma_0$ . On the other hand we have  $|\varphi(a) - \varphi(a')| < 6\sigma$  and so the length of  $S_{aa'}$  is at most  $(1 + \xi)6\sigma \leq 12\sigma < 100\sigma$ . Therefore, on such segments we in fact have that  $|D_\tau \varphi|([aa']) < |D_\tau \hat{\varphi}|([aa'])$ .

In the second case we have that  $|\hat{\varphi}(a) - \hat{\varphi}(a')| \geq \rho$  and, by (4.29), that

$$|\varphi(a) - \hat{\varphi}(a)| < 3\sigma \leq \varepsilon\rho$$

and the same holds also for  $a'$ . We can estimate by the triangle inequality and (4.29) that

$$\begin{aligned} |\varphi(a) - \varphi(a')| &\leq |\varphi(a) - \hat{\varphi}(a)| + |\hat{\varphi}(a) - \hat{\varphi}(a')| + |\hat{\varphi}(a') - \varphi(a')| \\ &\leq |\hat{\varphi}(a) - \hat{\varphi}(a')| + 6\sigma \\ &\leq (1 + \varepsilon)|\hat{\varphi}(a) - \hat{\varphi}(a')|. \end{aligned}$$

Now, because the length of the generalized segment between  $\mathbf{a}$  and  $\mathbf{b}$  with parameter  $\xi$  has length bounded by  $(1 + \xi)|\mathbf{a} - \mathbf{b}|$  (see Proposition 3.14), we get that

$$(4.31) \quad |D_\tau \varphi|([aa']) \leq (1 + \xi)(1 + \varepsilon)|D_\tau \hat{\varphi}|([aa'])$$

for any  $a, a' \in P$  adjacent on a segment in  $\tilde{\Gamma}$ . Summing over subsegments we see immediately that the same holds for any  $a, a' \in P$  lying on a segment of  $\tilde{\Gamma}$  but not necessarily adjacent.

The argument for a general pair  $a, b$  both lying on a single segment of  $\tilde{\Gamma}$  is as follows. We use the estimate (4.31) on the maximal segment  $[a_1 a_2] \subset [a, b]$  for  $a_1, a_2 \in P$ . Now it remains to estimate the length of the image of  $[aa_1]$  and  $[a_2 b]$ ; or if  $[ab] \cap P = \emptyset$  then we have to estimate the length of the image of  $[ab]$  knowing that  $[ab] \cap P = \emptyset$ . Let us deal with the former case first; the latter case readily follows from the first. We can deal with each segment separately so let us estimate the length of  $\varphi([aa_1])$ . The image of  $[aa_1]$  is a generalized segment contained in some rectangle, i.e.  $\{(x, y) \in \mathbb{R}^2 : w_n < x < w_{n+1}, z_m < y < z_{m+1}\}$  of  $\mathcal{G}$ . The diameter of each rectangle of  $\mathcal{G}$  is bounded by  $2\frac{\varepsilon}{2^{K+1}}$ . Then also the diameter of  $\varphi([aa_1])$  is also bounded by  $2\frac{\varepsilon}{2^{K+1}}$ . On the other hand by Proposition 3.14 the length of  $\varphi([aa_1])$  is bounded by  $(1 + \xi)$  times its diameter. Therefore, the length of  $\varphi([aa_1])$  is bounded by  $2(1 + \xi)\frac{\varepsilon}{2^{K+1}}$ . The same estimate holds for the length of  $\varphi([a_2 b])$  and indeed for  $\varphi([ab])$  if  $[ab] \cap P = \emptyset$ . Assuming then that  $[ab] \cap P \neq \emptyset$  we combine (4.31) with

$$\mathcal{H}^1(\varphi([aa_1])), \mathcal{H}^1(\varphi([a_2 b])) \leq (1 + \xi)2\frac{\varepsilon}{2^{K+1}}$$

we get (4.30), since  $\kappa = \frac{\varepsilon}{2^{K+1}}$ . In the case that  $[ab] \cap P = \emptyset$  then  $\mathcal{H}^1(\varphi([ab])) \leq (1 + \xi)2\frac{\varepsilon}{2^{K+1}}$  and so (4.30) holds.

The estimate in (4) follows immediately from  $(X_V, X_{\tilde{V}}) \in H_{V, \tilde{V}}$  for each adjacent  $V, \tilde{V}$  and (4.10). Similarly we get the estimate (6) immediately from (4.11).

Let us now prove the estimate (4.17). As in (4.21), each  $S_{i,m}$  has either

$$|D_\tau f|_{\partial \mathcal{Q}_i}(S_{i,m}^+), |D_\tau f|_{\partial \mathcal{Q}_i}(S_{i,m}^-) < \varepsilon 2^{-2K}$$

or

$$\max\{|D_\tau f|_{\partial \mathcal{Q}_i}(S_{i,m}^+), |D_\tau f|_{\partial \mathcal{Q}_i}(S_{i,m}^-)\} \geq \varepsilon 2^{-2K}, \text{ but } \mathcal{H}^1(\pi_{v_i}(S_{i,m})) < \frac{\varepsilon}{2^{2K}|D_\tau f|_\Gamma(\Gamma)}.$$

We start with the first case. Figure 4 should help identify the following sets. We define the set

$$A_{i,m}^+ := [S_{i,m-1}^+ \cup S_{i,m}^+ \cup S_{i,m+1}^+] \cap \{X \in \mathbb{R}^2 : \langle Y_{i,m-1}, v_i^\perp \rangle \leq \langle X, v_i^\perp \rangle \leq \langle Y_{i,m+1}, v_i^\perp \rangle\},$$

$$A_{i,m}^- := [S_{i,m-1}^- \cup S_{i,m}^- \cup S_{i,m+1}^-] \cap \{X \in \mathbb{R}^2 : \langle Y_{i,m-1}, v_i^\perp \rangle \leq \langle X, v_i^\perp \rangle \leq \langle Y_{i,m+1}, v_i^\perp \rangle\},$$

where  $Y_{i,m-1} = X_{i,m-1}$  if  $S_{i,m-1}$  is in the first category of (4.21) and  $Y_{i,m-1} = X_{i,m-1}^+$  if  $S_{i,m-1}$  is in the second category of (4.21) and  $Y_{i,m+1} = X_{i,m+1}$  if  $S_{i,m+1}$  is in the first category of (4.21) and  $Y_{i,m+1} = X_{i,m+1}^-$  if  $S_{i,m+1}$  is in the second category of (4.21). In all of the above cases we have

$$|D_\tau f|_{\partial \mathcal{Q}_i}(A_{i,m}^+ \cap S_{i,m+1}^+), |D_\tau f|_{\partial \mathcal{Q}_i}(A_{i,m}^+ \cap S_{i,m-1}^+) \leq \varepsilon 2^{-2K}.$$

By the choice of  $S_{i,m}$  being first category in (4.21) we have  $|D_\tau f|_{\partial \mathcal{Q}_i}(A_{i,m}^+) \leq 3\varepsilon 2^{-2K}$ . For each  $X^* \in S_{i,m}^+$  we can estimate (using (4.30))

$$|D_\tau \varphi|([X^* X_{i,m}^*]) \leq |D_\tau f|_{\partial \mathcal{Q}_i}(A_{i,m}^+) + 4\kappa \leq 3\varepsilon 2^{-2K} + 4\varepsilon 2^{-K}.$$

The above estimate is done on the bottom side in the same way, specifically

$$|D_\tau \varphi|([X_* X_{i,m,*}]) \leq |D_\tau f|_{\partial \mathcal{Q}_i}(A_{i,m}^-) + 4\kappa \leq 3\varepsilon 2^{-2K} + 4\varepsilon 2^{-K}.$$

for all  $X_* \in S_{i,m}^-$ .

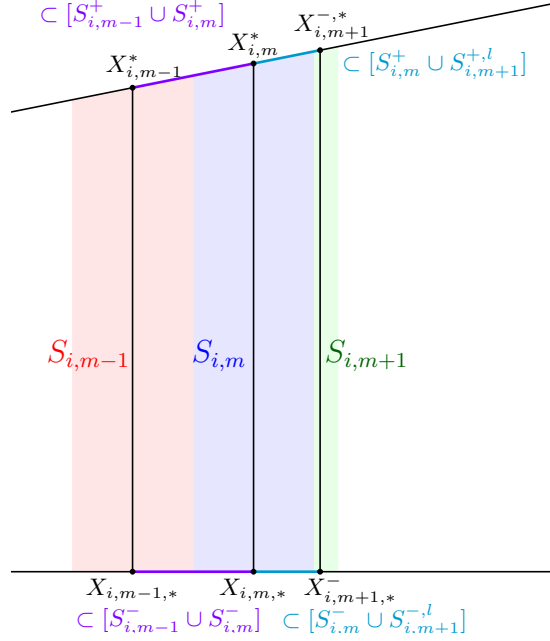


FIGURE 4. A depiction of the layout of the sets  $S_{i,m}$ ,  $S_{i,m}^{\pm}$  and the points  $X_{i,m}^*$ ,  $X_{i,m-1}^*$ ,  $X_{i,m+1}^*$ . The set  $S_{i,m-1}$  is category-1 and  $S_{i,m+1}$  is category-2. The variation along the purple and teal segments is small by the choice of  $X_{i,m-1}$ ,  $X_{i,m}$ ,  $X_{i,m+1}^-$ .

For each  $X_* \in S_{i,m}^-$  and  $X^* \in S_{i,m}^+$  with  $\pi_{v_i}(X_*) = \pi_{v_i}(X^*)$ . We define the path  $p_{X^*, X_*} \subset \tilde{\Gamma} \cap \overline{\mathcal{Q}_i}$  as

$$p(X^*, X_*) := [X^* X_{i,m}^*] \cup [X_{i,m}^* X_{i,m,*}] \cup [X_{i,m,*}, X_*].$$

We estimate using (4.30) and (4.23) that

$$\begin{aligned} |D_{\tau}\varphi|([X_{i,m}^* X_{i,m,*}]) &\leq (1 + \varepsilon)(1 + \xi) |D_{\tau}f|_{[X_{i,m} + \mathbb{R}v_i] \cap \mathcal{Q}_i} |([X_{i,m} + \mathbb{R}v_i] \cap \mathcal{Q}_i)| + 4\kappa \\ &\leq (1 + \varepsilon)(1 + \xi) \frac{|D_{v_i}f|(S_{i,m} \cap \mathcal{Q}_i)}{\mathcal{H}^1(\pi_{v_i}(S_{i,m}))} + 4\kappa. \end{aligned}$$

Combining the above estimates we get

$$|D_{\tau}\varphi|(p(X^*, X_*)) \leq (1 + \varepsilon)(1 + \xi) \frac{|D_{v_i}f|(S_{i,m} \cap \mathcal{Q}_i)}{\mathcal{H}^1(\pi_{v_i}(S_{i,m}))} + 6\varepsilon 2^{-2K} + 12\kappa.$$

This means that (see (2.3))

$$(4.32) \quad d_{\varphi(\partial\mathcal{Q}_i)}(\varphi(X^*), \varphi(X_*)) \leq (1 + \varepsilon)(1 + \xi) \frac{|D_{v_i}f|(S_{i,m} \cap \mathcal{Q}_i)}{\mathcal{H}^1(\pi_{v_i}(S_{i,m}))} + 6\varepsilon 2^{-2K} + 12\kappa$$

for every opposing pair  $X_*$ ,  $X^*$  in  $S_{i,m}$  of first category.

For opposing pairs  $X_*$ ,  $X^*$  in  $S_{i,m}$  for  $S_{i,m}$  in the second category of (4.21) we use the estimate

$$(4.33) \quad \begin{aligned} d_{\varphi(\partial\mathcal{Q}_i)}(\varphi(X^*), \varphi(X_*)) &\leq (1 + \varepsilon)(1 + \xi) |D_{\tau}f|_{\Gamma}(|\partial\mathcal{Q}_i|) + 4\kappa \\ &\leq (1 + \varepsilon)(1 + \xi) |D_{\tau}f|_{\Gamma}(|\Gamma|) + 4\kappa. \end{aligned}$$

For opposing pairs in the perpendicular direction, i.e.  $Z_* \in T_{i,m}^-, Z^* \in T_{i,m}^+$  we estimate in the same way using (4.25) in place of (4.23). For  $Z^* \in T_{i,m}^+, Z_* \in T_{i,m}^-$  we get

$$(4.34) \quad d_{\varphi(\partial Q_i)}(\varphi(Z^*), \varphi(Z_*)) \leq \begin{cases} \frac{(1+\varepsilon)(1+\xi)|D_{v_i^\perp} f|(T_{i,m} \cap Q_i)}{\mathcal{H}^1(\pi_{v_i^\perp}(T_{i,m}))} + 6\varepsilon 2^{-2K} + 12\kappa & T_{i,m}^\pm \text{ is category 1} \\ (1+\varepsilon)(1+\xi)|D_\tau f|_\Gamma(\Gamma) + 4\kappa & T_{i,m}^\pm \text{ is category 2.} \end{cases}$$

Find  $\theta_i \in [0, 2\pi)$  such that  $v_i = (\cos \theta_i, \sin \theta_i)$ . We denote the set of indexes  $m$  such that  $S_{i,m}$  is category 1 by  $C_1$  and category 2 by  $C_2$ . We denote the set of indexes  $m$  such that  $T_{i,m}$  is category 1 by  $C_3$  and category 2 by  $C_4$ . We use the admissibility of  $\tilde{\Gamma}$  (especially  $|Df|(\tilde{\Gamma}) = 0$ ) to prove (4.17)

$$\begin{aligned} & \int_{\pi_{v_i}(\mathcal{Q}_i)} d_{\varphi(\partial Q_i)}(\varphi(X^*), \varphi(X_*)) d\mathcal{H}^1(X) \\ & \leq (1+\varepsilon)(1+\xi) \left[ \sum_{m \in C_1} |D_{v_i} f|(S_{i,m} \cap Q_i) \right. \\ & \quad \left. + \sum_{m \in C_2} |D_\tau f|_\Gamma(\Gamma) \frac{\varepsilon}{2^{2K} |D_\tau f|_\Gamma(\Gamma) \mathfrak{R}} \right] \\ & \quad + [6\varepsilon 2^{-2K} + 12\kappa] \mathcal{H}^1(\pi_{v_i}(\mathcal{Q}_i)) \\ & \leq (1+\varepsilon)(1+\xi) |Df|(\mathcal{Q}_i) + C 2^{-2K} \varepsilon. \end{aligned}$$

Recall our notation  $g|D^s f| = D^s f$  and by the Theorem 2.9 we have  $g = u(X) \otimes v(X)$  for  $|D^s f|$ -almost every  $X$ . We have  $|\langle v(X), v_i^\perp \rangle| \leq C|v(X) - v_i| + C|\langle v_i, v_i^\perp \rangle|$ . Using this and (4.4) we calculate

$$\begin{aligned} & \int_{\pi_{v_i^\perp}(\mathcal{Q}_i)} d_{\varphi(\partial Q_i)}(\varphi(Z^*), \varphi(Z_*)) d\mathcal{H}^1(Z) \\ & \leq (1+\varepsilon)(1+\xi) \left[ \sum_{m \in C_3} |D_{v_i^\perp} f|(T_{i,m} \cap Q_i) + \sum_{m \in C_4} |D_\tau f|_\Gamma(\Gamma) \frac{\varepsilon}{2^{2K} |D_\tau f|_\Gamma(\Gamma) \mathfrak{R}} \right] \\ & \quad + [6\varepsilon 2^{-2K} + 12\kappa] \mathcal{H}^1(\pi_{v_i}(\mathcal{Q}_i)) \\ & \leq (1+\varepsilon)(1+\xi) |\langle Df, v_i^\perp \rangle|(\mathcal{Q}_i) + C 2^{-K} \varepsilon \\ & \leq (1+\varepsilon)(1+\xi) \varepsilon |D^s f|(4Q_i \cap S) + C 2^{-2K} \varepsilon \end{aligned}$$

Proving (4.18). □

Now it suffices to extend the mapping  $\varphi$  defined in Proposition 4.4 to get a  $BV$  homeomorphism. This is the content of the following theorem.

**Theorem 4.5.** *Let  $f \in BV(Q(0,1); Q(0,1))$ , let  $f(x,y) = (x,y)$  for  $(x,y) \in \partial Q(0,1)$  and let  $f$  satisfy the  $NCBV^+$  condition. Then there exists a sequence  $f_k \in BV(Q(0,1), Q(0,1))$ ,  $f_k = \text{id}$  on  $\partial Q(0,1)$  converging to  $f$  area-strictly.*

*Proof.* Given  $\varepsilon > 0$  we get a good non-straight grid for  $f$  called  $\Gamma$  and a mapping  $\varphi$  from Proposition 4.4.

**Step 1. Extend  $\varphi$  to get a homeomorphism.**

Let us start by extending  $\varphi$  on the ‘good’ quadrilaterals, i.e. on  $\mathcal{Q}_i$  such that  $Q_i \in G_{\varepsilon, \alpha, K}$ . We have that  $\varphi$  is linear on each side of  $\mathcal{Q}_i$ . Further by the choice of  $X, X'$ , the vertices of  $\mathcal{Q}_i$ , (specifically (4.12)) we have

$$(4.35) \quad |f(X) - f(X') - \nabla f(w_i)(X - X')| < \alpha^4 2^{-K}.$$

Recall that we have  $\alpha \leq |\nabla f(w_i)| \leq \alpha^{-1}$ ,  $\alpha < \det \nabla f(w_i)$ , which implies that  $|\nabla f(w_i)v| \geq \alpha^2$  for all  $|v| = 1$ . Therefore, since  $|X - X'| \geq 2^{-K-1}$ , we have  $|\nabla f(w_i)(X - X')| \geq 2^{-K-1}\alpha^2$ . By Proposition 4.4, step 1 we have that  $\mathcal{Q}_i$  is convex and so also is  $\nabla f(w_i)\mathcal{Q}_i$ . Therefore (4.35) guarantees that  $\varphi(\partial\mathcal{Q}_i)$  is a convex quadrilateral. Then we can choose any pair of opposing corners of  $\mathcal{Q}_i$ , call them  $X$  and  $X'$ . We define  $\varphi$  as linear continuous on the segment  $[XX']$ . Then  $\mathcal{Q}_i$  is composed of two triangles and  $\varphi$  is continuous and linear on each side of each triangle and therefore  $\varphi$  extends to an affine map on each triangle, which we call  $g$ . By the convexity of  $\varphi(\partial\mathcal{Q}_i)$  we get the injectivity of the map  $g$  on each  $\mathcal{Q}_i$ .

On the quadrilaterals  $\mathcal{Q}_i$  such that  $Q_i \in T_{\varepsilon, \alpha, K}$  we apply Theorem 2.2. For each quadrilateral  $\mathcal{Q}_i$  such that  $Q_i \in E_{\varepsilon, K}$  we find a  $\theta_i$  such that  $v_i = (\cos \theta_i, \sin \theta_i)$  and we apply Theorem 2.3 with the parameter  $\varepsilon$  from Theorem 2.3  $\varepsilon 2^{-2K}$ . On the other quadrilaterals, i.e.  $\mathcal{Q}_i$  such that  $Q_i \in W_{\varepsilon, \alpha, K} \cup [F_{\varepsilon, K} \setminus E_{\varepsilon, K}]$  we apply Theorem 2.1.

The map  $g$  is injective on  $Q(0, 1)$  because it is injective on each  $\mathcal{Q}_i$ , because  $g = \varphi$  on  $\Gamma$  and because  $\varphi$  is injective on  $\Gamma$ .

## Step 2. Convergence estimates.

It is not hard to observe the  $L^1$  convergence. Since  $\|g\|_\infty < 2$  and the set  $\mathcal{L}^2(\tilde{F}_{\varepsilon, K} \cup \tilde{P}_{\alpha_0}) < \delta < \varepsilon$  it remains to consider the set  $Q(0, 1) \setminus [\tilde{F}_{\varepsilon, K} \cup \tilde{P}_{\alpha_0}]$ . On each of the  $\mathcal{Q}_i$  we use (4.6) and the Poincaré inequality to get

$$\|f - g\|_{L^1(\mathcal{Q}_i)} \leq C 2^{-3K} \alpha_0^{-1}$$

and since  $2^{-K} < \varepsilon \alpha_0$  we get that  $\|f - g\|_1 < C\varepsilon$ .

Let us call  $\Omega = Q(0, 1) \setminus \bigcup_{Q_i \in E_{\varepsilon, K}} \mathcal{Q}_i$ . We want to estimate  $\|Dg\chi_\Omega - D^a f\|_{L^1(Q(0, 1))}$ . For  $\mathcal{L}^2$ -almost all points of  $\mathcal{Q}_i$  with  $Q_i \in G_{\varepsilon, \alpha, K}$  we have  $|Dg - \nabla f(w_i)| \leq 2\alpha^4$  by (4.35) and we chose  $\alpha < \varepsilon$ . Therefore using (4.6) we have

$$\int_{\mathcal{Q}_i} |Dg - D^a f| \leq 3\varepsilon^3 \mathcal{L}^2(\mathcal{Q}_i).$$

Summing this over  $Q_i \in G_{\varepsilon, \alpha, K}$  we bound the sum by  $\varepsilon$ .

Now we calculate on  $\mathcal{Q}_i$  for  $Q_i \in T_{\varepsilon, \alpha, K}$ . Summing the estimate from Theorem 2.2 using the estimate from (4.19) we get

$$\sum_i \int_{\mathcal{Q}_i} |Dg - D^a f| \leq C\varepsilon |Df|(Q).$$

By the definition of  $W_{\varepsilon, \alpha, K}$  and  $F_{\varepsilon, K} \setminus E_{\varepsilon, K}$  we have that

$$|Df| \left( \bigcup_{Q_i \in W_{\varepsilon, \alpha, K} \cup [F_{\varepsilon, K} \setminus E_{\varepsilon, K}]} \mathcal{Q}_i \right) \leq C\varepsilon.$$

On each  $\mathcal{Q}_i$  such that  $\mathcal{Q}_i \in W_{\varepsilon, \alpha, K} \cup [F_{\varepsilon, K} \setminus E_{\varepsilon, K}]$  we have

$$|Dg| \left( \bigcup_{\mathcal{Q}_i \in W_{\varepsilon, \alpha, K} \cup [F_{\varepsilon, K} \setminus E_{\varepsilon, K}]} \mathcal{Q}_i \right) \leq C |Df| \left( \bigcup_{\mathcal{Q}_i \in W_{\varepsilon, \alpha, K} \cup [F_{\varepsilon, K} \setminus E_{\varepsilon, K}]} \mathcal{Q}_i \right) \leq C\varepsilon.$$

Therefore, recalling that  $\|D^a f\|_{L^1(Q(0,1) \setminus \Omega)} < \varepsilon |D^a f|(Q(0,1))$  by Lemma 4.1, we conclude that

$$\|(Dg)\chi_\Omega - D^a f\|_{L^1(Q(0,1))} < C\varepsilon.$$

It remains to prove that  $\|Dg\|_{L^1(Q(0,1) \setminus \Omega)} \leq |D^s f|(Q(0,1)) + C\varepsilon$  in order to prove the area strict convergence. In the following we use the fact that

$$|Dg|(\mathcal{Q}_i) \leq |\langle Dg, v_i \rangle|(\mathcal{Q}_i) + |\langle Dg, v_i^\perp \rangle|(\mathcal{Q}_i).$$

On the set in question we defined  $g$  using Theorem 2.3 and thanks to (4.17), (4.18) and Theorem 4.2, points 2) and 3) we have

$$\begin{aligned} |Dg|(Q(0,1) \setminus \Omega) &\leq \sum_{\mathcal{Q}_i \in E_{\varepsilon, K}} |\langle Dg, v_i \rangle|(\mathcal{Q}_i) + |\langle Dg, v_i^\perp \rangle|(\mathcal{Q}_i) \\ &\leq \sum_{\mathcal{Q}_i \in E_{\varepsilon, K}} (1 + \varepsilon) |Df|(\mathcal{Q}_i) + C\varepsilon 2^{-2K} + C\varepsilon |D^s f|(4\mathcal{Q}_i \cap S) \\ &\leq (1 + C\varepsilon) |D^s f|(Q(0,1)) + C\varepsilon \end{aligned}$$

because  $\sum_i \chi_{4\mathcal{Q}_i} \leq 25$ . □

*Proof of Theorem 1.1.* The equivalence of points 1) and 2) is in Theorem 3.18. The implication 2) implies 3) is Theorem 4.5. Trivially 3) implies 4). The equivalence of 4) and 1) was proved in [10]. □

## REFERENCES

- [1] G. Alberti, Rank one property for derivatives of functions with bounded variation, *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* **123**, no. 2 (1993), 239–274.
- [2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [3] M. Baía, S. Krömer and M. Kružík, Generalized  $W^{1,1}$ -Young measures and relaxation of problems with linear growth, *SIAM: J. Math. Anal.* **50**, no. 1 (2018), 1076–1119.
- [4] J. M. Ball, *Some open problems in elasticity.*, Geometry, mechanics, and dynamics, Springer, New York, 2002.
- [5] Ball J., Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rational Mech. Anal.* **63** (1978), 337–403.
- [6] Ball J., Global invertibility of Sobolev functions and the interpenetration of matter, *Proc. Roy. Soc. Edinburgh Sect. A* **88** (1981), 315–328.
- [7] D. Campbell, Diffeomorphic approximation of Planar Sobolev Homeomorphisms in Orlicz-Sobolev spaces, *J. Funct. Anal.* **273** (2017), 125–205.
- [8] D. Campbell, S. Hencl A. Kauranen and E. Radici, Strict limits of planar  $BV$  homeomorphisms, *Nonlinear Analysis* **177** (2018), 209–237.
- [9] D. Campbell, A. Pratelli and E. Radici, Comparison between the non-crossing and the non-crossing on lines properties, *J. Math. Anal. and Appl.* **498** Issue. 1 (2021), Article 124956.
- [10] D. Campbell, A. Kauranen and E. Radici, *Classification of strict limits of planar  $BV$  homeomorphisms*, arXiv:2101.09013, 2021.
- [11] D. Campbell, A. Kauranen and E. Radici, *Minimal Extension for the  $\alpha$ -Manhattan norm*, Preprint, 2022.



- [12] G. De Philippis and A. Pratelli, The closure of planar diffeomorphisms in Sobolev spaces, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **37**, no. 1 (2020), 181–224.
- [13] L. Evans and R. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton.
- [14] A.N. Gent and P.B. Lindley, Internal rupture of bonded rubber cylinders in tension, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **249**, no. 1257 (1959), 195–205.
- [15] A. Gent and C. Wang, Fracture mechanics and cavitation in rubber-like solids, *J. Mater. Sci* **26** (1991), 3392–3395.
- [16] D. Henao and C. Mora-Corral, Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity, *Arch. Rational Mech. Anal.* **197** (2010), 619–655.
- [17] D. Henao and C. Mora-Corral, Fracture surfaces and the regularity of inverses for BV deformations, *Arch. Rational Mech. Anal.* **201** (2011), 575–629.
- [18] D. Henao and C. Mora-Corral, Lusin’s condition and the distributional determinant for deformations with finite energy, *Adv. Calc. Var.* **5** (2012), 355–409.
- [19] D. Henao and C. Mora-Corral, Regularity of inverses of Sobolev deformations with finite surface energy, *J. Funct. Anal.* **268** (2015), 2356–2378.
- [20] S. Hencl and A. Pratelli, Diffeomorphic Approximation of  $W^{1,1}$  Planar Sobolev Homeomorphisms, *J. Eur. Math. Soc* **20**, no. 3 (2018), 597–656.
- [21] T. Iwaniec, L. Kovalev and J. Onnien, Diffeomorphic approximation of Sobolev homeomorphisms, *Arch. Rational Mech. Anal* **201**, no. 3 (2011), 1047–1067.
- [22] T. Iwaniec, L. Kovalev and J. Onnien, Hopf differentials and smoothing Sobolev homeomorphisms, *International Mathematics Research Notices* **14** (2012), 3256–3277.
- [23] T. Iwaniec and J. Onninen, Limits of Sobolev homeomorphisms, *J. Eur. Math. Soc* **19**, no. 2 (2017), 473–505.
- [24] J. Kristensen and F. Rindler, Relaxation of signed integral functionals in  $BV$ , *Calc. Var. Partial Differential Equations* **37**, no. 1-2 (2010), 29–62.
- [25] S. Müller and S. J. Spector, An existence theory for nonlinear elasticity that allows for cavitation, *Arch. Rational Mech. Anal.* **131**, no. 1 (1995), 1–66.
- [26] N. Petrinic, C.R. Siviour, J.L. Curiel Sosa and B.C.F. Elliott, Improved predictive modelling of strain localisation and ductile fracture in a Ti-6Al-4V alloy subjected to impact loading, *J. Phys. IV* **134** (2006), 147–155.
- [27] N. Petrinic, C.R. Siviour, J.L. Curiel Sosa and B.C.F. Elliott, On competing volumetric and deviatoric damage mechanisms in simulation of ductile fracture in Ti64 alloy at high rates of strain, *Proceedings of Euromech EMMC-10 Conference “Multiphase and multi-component materials under dynamic loading KazimierzDolny* (2007), 467–478.
- [28] A. Pratelli and E. Radici, On the planar minimal BV extension problem, *Rendiconti Lincei: Matematica e Applicazioni* **29**, no. 3 (2018), 511–555.
- [29] A. Pratelli and E. Radici, Approximation of planar BV homeomorphisms by diffeomorphisms, *J. Func. Anal* **276** (2019), 659–686.
- [30] F. Rindler and G. Shaw, Strictly continuous extension of functionals with linear growth to the spaces  $BV$ , *Q. J. Math.* **66**, no. 3 (2015), 953–978.
- [31] M. Williams and R. Sharper, Spherical flaw instability in hydrostatic tension, *Int. J. Fract. Mech* **1** (1965), 64–71.

D. CAMPBELL: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HRADEC KRÁLOVÉ, ROKITANSKÉHO 62, 500 03 HRADEC KRÁLOVÉ, CZECH REPUBLIC

*Email address:* `daniel.campbell@uhk.cz`

A. KAURANEN: DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF JYVÄSKYLÄ, PL 35, 40014 JYVÄSKYLÄN YLIOPISTO, FINLAND

*Email address:* `aapo.p.kauranen@jyu.fi`

E. RADICI: DISIM - DEPARTMENT OF INFORMATION ENGINEERING, COMPUTER SCIENCE AND MATHEMATICS, UNIVERSITY OF L'AQUILA, VIA VETOIO 1 (COPPITO), 67100 L'AQUILA (AQ), ITALY

*Email address:* `emanuela.radici@univaq.it`