RATIONAL POINTS NEAR MANIFOLDS AND KHINTCHINE THEOREM

VICTOR BERESNEVICH AND SHREYASI DATTA

ABSTRACT. In this paper we complete the long standing challenge to establish a Khintchine type theorem for arbitrary nondegenerate manifolds in \mathbb{R}^n . In particular, our main result finally removes the analyticity assumption from the Khintchine type theorem proved in [Ann. of Math. 175 (2012), 187-235]. Furthermore, we obtain a Jarnìk type refinement of our main result, which uses Hausdorff measures. The results are also obtained in the inhomogeneous setting. The proofs are underpinned by a sharper version of a 'quantitative nondivergence' result of Bernik–Kleinbock–Margulis and a duality argument which we use to study rational points near manifolds.

1. Introduction

Throughout $\boldsymbol{\theta} \in \mathbb{R}^n$, $\psi : (0, \infty) \to (0, 1)$ is a non-increasing function such that $\psi(q) \to 0$ as $q \to \infty$ and \mathcal{L}_n denotes Lebesgue measure on \mathbb{R}^n . In this paper we will be interested in the following set that naturally arises in a variety of problems in the theory of inhomogeneous Diophantine approximation:

$$\mathcal{S}_n^{\theta}(\psi) := \left\{ \mathbf{x} \in \mathbb{R}^n : \left\| \mathbf{x} - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| \le \frac{\psi(q)}{q} \text{ for i.m. } (\mathbf{p}, q) \in \mathbb{Z}^n \times \mathbb{N} \right\}, \tag{1.1}$$

where 'i.m.' reads 'infinitely many' and $\|\cdot\|$ denotes the sup norm. The points \mathbf{x} lying in $\mathcal{S}_n^{\boldsymbol{\theta}}(\psi)$ are often referred to as $(\psi, \boldsymbol{\theta})$ -approximable. When $\boldsymbol{\theta} = \mathbf{0}$, we deal with homogeneous approximations to points \mathbf{x} in \mathbb{R}^n by rational points \mathbf{p}/q , in which case $\mathbf{x} \in \mathcal{S}_n^{\mathbf{0}}(\psi)$ is called ψ -approximable. Since the set $\mathcal{S}_n^{\boldsymbol{\theta}}(\psi)$ is invariant under translations by integer vectors, it is often restricted to the unit hyper-cube $[0,1]^n$.

By the inhimogeneous version of Khintchine's classical results, see [1, §1.1], we fully understand the probability of a random point $\mathbf{x} \in [0, 1]^n$ being $(\psi, \boldsymbol{\theta})$ -approximable:

The Inhomogeneous Khintchine theorem: For any monotonic ψ and $\theta \in \mathbb{R}^n$

$$\mathcal{L}_n(\mathcal{S}_n^{\boldsymbol{\theta}}(\psi) \cap [0,1]^n) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} \psi(q)^n < \infty, \\ 1 & \text{if } \sum_{q=1}^{\infty} \psi(q)^n = \infty. \end{cases}$$

$$(1.2)$$

A longstanding notorious problem in the theory of Diophantine approximation is to obtain Khintchine type results when \mathbf{x} is restricted to a subset of \mathbb{R}^n , such as a manifold or a fractal, which supports a natural (probability) measure. In particular, Khintchine

type results have long been search for in the case of so-called nondegenerate manifolds. In this paper we complete this quest by establishing the following key result.

Theorem 1.1. Let $n \geq 2$, ψ be non-increasing, $\boldsymbol{\theta} \in \mathbb{R}^n$ and \mathfrak{M} be any nondegenerate submanifold of \mathbb{R}^n . Then, almost no/almost every point $\mathbf{x} \in \mathfrak{M}$ is $(\psi, \boldsymbol{\theta})$ -approximable provided that the sum

$$\sum_{q=1}^{\infty} \psi(q)^n \tag{1.3}$$

converges/diverges.

Remark 1.1. Nondegenerate manifolds satisfy a natural and mild geometric condition. While leaving the definition of nondegeneracy until §1.1, we now describe our key new ideas and how Theorem 1.1 builds upon previous findings. The convergence case of this theorem was previously established in [13] in the homogeneous setting, that is for $\theta = 0$, and thus the convergence case of Theorem 1.1 can be regarded as an extension of [13] to the inhomogeneous setting. However, the divergence case of Theorem 1.1 is new, even for $\theta = 0$ – previous divergence results did not reach the full class of nondegenerate manifolds even in the homogeneous case. Indeed, for manifolds of dimension d > 2 the divergence case has previously been known only for analytic nondegenerate manifolds – this is the main result of [5]. Beyond analytic manifolds, the divergence case of Theorem 1.1 was only ever established for nondegenerate curves $(d=1, n \geq 2)$, which is the main result of [11]. We note that both [5] and [11] adopt a ubiquity approach relying on an application of quantitative non-divergence estimates first introduced by Kleinbock and Margulis in their landmark work [31]. It is the application of the quantitative non-divergence estimates that represents a challenge resulting in additional assumptions on nondegenerate manifolds in [5] and [11]. In this paper we present a new approach using a version of the quantitative non-divergence estimate of Bernik-Kleinbock-Margulis [18] and an adaptation of a duality argument from [13]. In §3, we also establish a strengthening of the quantitative non-divergence estimate of [18] which we use to obtain the Hausdorff measure version of Theorem 1.2, namely Theorem 2.1 stated below.

Remark 1.2. Apart from papers [5, 11, 13] that we mentioned in the previous remark, over the years the search for a Khintchine type theorem for manifolds has yielded a number of other milestones. In this context, we ought to mention the case of planar curves (n = 2, d = 1), where a homogeneous Khintchine type theorem was first established for arbitrary C^3 nondegenerate curves in the case of convergence [34] and divergence [8]. Subsequently, these were taken to the inhomogeneous setting in [16], while the C^3 assumption was removed, in fact relaxed to weak nondegeneracy - a condition that is locally closed in the uniform convergence topology, in [14] and [26]. Previous results in higher dimensions, have been established for various subclasses of nondegenerate manifolds with various geometric restrictions which were excluding curves, see [5, 10, 27, 28, 32, 33] and references within.

Remark 1.3. It should be noted that there is the dual to (1.1) setting, e.g. see [6, 19], which even in the case of manifolds has been well understood for quite some time

now. Indeed, the story begins with the pioneering work of Kleinbock and Margulis [31] addressing approximating functions of the form $\psi(t) = t^{-w}$. Subsequently, a full dual analogue of Khintchine's theorem for nondegenerate manifolds was obtained in [4, 15, 18] for homogeneous approximations and in [2, 12] for inhomogeneous approximations. We also refer the reader to [9, 23, 24, 29, 30] for relevant results for affine subspaces which represent the degenerate case.

1.1. Nondegeneracy of maps and manifolds. In view of the measure theoretic nature of the problems we investigate, it suffices to consider manifolds locally. Hence, without loss of generality we will assume that $\mathcal{M} = \mathbf{f}(\mathbf{U})$, where **U** is an open subset of \mathbb{R}^d , $1 \leq d \leq n$ and $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ is a nonsingular C^1 map, that is the Jacobian matrix of \mathbf{f} has rank d. Here and elsewhere d stands for the dimension of \mathcal{M} . Following [31], we will say that a map $\mathbf{f}: \mathbf{U} \to \mathbb{R}^n$ is l-nondegenerate at $\mathbf{x} \in \mathbf{U}$ if \mathbf{f} is C^l on a neighbourhood of \mathbf{x} and there are n linearly independent partial derivatives of \mathbf{f} of orders $\leq l$. We will say that **f** is nondegenerate at **x** if it is *l*-nondegenerate at **x** for some l. We say that f is nondegenerate in U if f is nondegenerate at (Lebesgue) almost every $\mathbf{x} \in \mathbf{U}$. Respectively, manifold M is called nondegenerate if a neighborhood of (almost) every point can be parametrised by a nondegenerate map. If we are given a (local) parameterisation $\mathbf{f}: \mathbf{U} \to \mathbb{R}^n$ of \mathcal{M} so that $\mathcal{M} = \mathbf{f}(\mathbf{U})$, where \mathbf{U} is an open subset of \mathbb{R}^d , $d = \dim \mathcal{M}$, then we will conveniently use the pushforward $\mathbf{f}_{\star}\mathcal{L}_d$ of d-dimensional Lebesgue measure. This is defined for $A \subset \mathcal{M}$ as follows: $\mathbf{f}_{\star}\mathcal{L}_d(A) := \mathcal{L}_d(\mathbf{f}^{-1}(A))$. In view of the above remarks, Theorem 1.1 can now be equivalently reformulated as follows:

Theorem 1.2. Let $n \geq 2$, $\mathbf{U} \subset \mathbb{R}^d$ be an open subset, $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be nonsingular and nondegenerate in \mathbf{U} . Let $\boldsymbol{\theta} \in \mathbb{R}^n$ and ψ be non-increasing. Then

$$\mathcal{L}_d(\mathbf{f}^{-1}(\mathbf{S}_n^{\boldsymbol{\theta}}(\psi))) = \begin{cases} 0 & \text{if } \sum_{q} \psi(q)^n < \infty, \\ \mathcal{L}_d(\mathbf{U}) & \text{if } \sum_{q} \psi(q)^n = \infty. \end{cases}$$
 (1.4)

1.2. Rational points near manifolds. The key to the divergence result is the construction of an appropriate ubiquitous system of rational points lying near a given manifold. A by-product of this is a lower bound on the number of rational points lying close to a given manifold, namely a lower bound on

$$\#\left\{ (\mathbf{p}, q) \in \mathbb{Z}^{n+1} : \inf_{\mathbf{y} \in \mathcal{M}} \left\| \mathbf{y} - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| \le \frac{\psi(e^t)}{e^t}, \ 0 < q < e^t \right\}, \tag{1.5}$$

where \mathcal{M} is a given manifold, or rather a generic open subset of a given manifold.

The convergence case generally requires an optimal (up to a constant factor) upper bound for (1.5). We note that as a consequence of the convergence of (1.3), the monotonic function ψ must satisfy the inequality

$$\psi(t) < t^{-1/n} \tag{1.6}$$

for all sufficiently large t.

Regarding (1.5), in the groundbreaking work [34] an optimal upper bound was established for any nonplanar C^3 curve in \mathbb{R}^3 for the best possible range of approximating functions ψ . But in higher dimensions obtaining the expected (heuristic) upper bound for (1.5) turns out to be a difficult problem. Indeed, until recently all the results in higher dimensions required extra conditions on the nondegenerate manifolds; see [10, 27]. In [13] for any nondegenerate manifolds a sharp upper bound was established only near a part of the manifold, that excludes a small Lebesgue measure set. More is known about the lower bound; see [5, 11, 32].

To state our new results we now introduce further notation. For any $\Delta \subset \mathbf{U}$, define

$$N_{\boldsymbol{\theta}}^{\star}(\Delta; \varepsilon, t) := \# \left\{ (\mathbf{p}, q) \in \mathbb{Z}^{n+1} \mid e^{t-1} \leq q \leq e^{t}, \inf_{\mathbf{x} \in \Delta} \left\| \mathbf{f}(\mathbf{x}) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| \leq \frac{\varepsilon}{e^{t}} \right\}$$

and

$$N_{\boldsymbol{\theta}}(\Delta; \varepsilon, t) := \# \left\{ (\mathbf{p}, q) \in \mathbb{Z}^{n+1} \mid 1 \le q \le e^t, \inf_{\mathbf{x} \in \Delta} \left\| \mathbf{f}(\mathbf{x}) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| \le \frac{\varepsilon}{e^t} \right\}.$$

Let

$$\eta = \eta(n, d) := \begin{cases} \frac{1}{n-1} & \text{if } d > 1, \\ \frac{3}{2n-1} & \text{if } d = 1. \end{cases}$$
(1.7)

By $x \ll_{\delta} y$, we mean $x \leq cy$, where c > 0 is a constant that depends only on δ , a parameter. Now we state our main theorems on counting rational points near manifolds.

Theorem 1.3 (Lower bound). Let $n \geq 2$, $1 \leq d < n$, $\boldsymbol{\theta} \in \mathbb{R}^n$, $\mathbf{U} \subset \mathbb{R}^d$ be an open subset, $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be nonsingular and nondegenerate at $\mathbf{x}_0 \in \mathbf{U}$. Let η be as in (1.7). Then there exists an open neighborhood B_0 of \mathbf{x}_0 such that for any ball $B \subset B_0$, there exists $t_0 = t_0(B, n, \mathbf{f})$ such that for every $t \geq t_0$,

$$N_{\theta}^{\star}(B;\varepsilon,t) \gg_{\mathbf{x}_0,d,\mathbf{f}} \varepsilon^m e^{(d+1)t} \mathcal{L}_d(B)$$
 (1.8)

for any $\varepsilon \in (0,1)$ satisfying

$$\varepsilon \gg_{n,\mathbf{f},\mathbf{x}_0} e^{-t\eta}$$
. (1.9)

Compared to previous results, Theorem 1.3 removes any extra conditions on the manifolds beyond the nondegeneracy. Indeed, the previous results; see [5, 11, 32] and the references therein, have either dimension restriction or smoothness/analyticity assumptions on the manifolds, although their range of ε may be bigger than that in Theorem 1.3. Thus even for $\theta = 0$ Theorem 1.3 is new.

Theorem 1.4 (Upper bound). Let $n \geq 2$, $1 \leq d < n$, $\theta \in \mathbb{R}^n$, $\mathbf{U} \subset \mathbb{R}^d$ be an open subset, $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be nonsingular and l-nondegenerate at $\mathbf{x}_0 \in \mathbf{U}$. Then there exists an open neighborhood B_0 of \mathbf{x}_0 such that for some $t_0 = t_0(B_0, n, \mathbf{f})$, for every $t \geq t_0$,

$$N_{\boldsymbol{\theta}}^{\star}(B_0; \varepsilon, t) \ll_{n, \mathbf{f}, B_0} \varepsilon^m e^{(d+1)t} + e^{(d+1)t} \left(\varepsilon^{n-1/2} e^{3t/2}\right)^{-\alpha}, \tag{1.10}$$

where $\alpha = \frac{1}{d(2l-1)(n+1)}$. In particular, if $\varepsilon > e^{\frac{-3t}{2md(2l-1)(n+1)+2n-1}}$ then the first term in (1.10) dominates and the bound is optimal up to a constant factor.

Even in the case of $\boldsymbol{\theta} = \mathbf{0}$, the above theorem extends [32, Theorem 1.5], where it was shown that if \mathbf{f} , as in Theorem 1.4, is also $\lceil \frac{n+1}{\delta} \rceil$ times continuously differentiable for some $0 < \delta \leq \frac{1}{2n+10}$, then $N_{\mathbf{0}}^{\star}(B_0; \varepsilon, t) \ll_{n,\mathbf{f},B_0} \varepsilon^m e^{(d+1)t}$ when $\varepsilon > e^{\frac{-3t}{2md(2l-1)(n+1)+2n-1}} e^{\frac{(2n+1)\delta t}{2md(2l-1)(n+1)+2n-1}}$. This is a strictly smaller range of ε than the range of ε in Theorem 1.4.

2. Hausdorff dimension and measure results

In this section we provide refinements of the results stated in §1 using Hausdorff measures and dimension. To begin with, we recall the following Hausdorff measure generalisation of Khintchine's theorem.

The Inhomogeneous Khintchine-Jarník theorem: For any non-increasing ψ , any $\theta \in \mathbb{R}^n$ and any $0 < s \le n$

$$\mathcal{H}^{s}(\mathcal{S}_{n}^{\theta}(\psi) \cap [0,1]^{n}) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q^{n-s} \psi(q)^{s} < \infty, \\ \mathcal{H}^{s}([0,1]^{n}) & \text{if } \sum_{q=1}^{\infty} q^{n-s} \psi(q)^{s} = \infty. \end{cases}$$
(2.1)

Here and elsewhere \mathcal{H}^s denotes s-dimensional Hausdorff measure. We refer the reader to [20], or [7, Section 12.1] or [1, §1.1 & 1.3] for a detailed historical account and further generalisations. The case of s = n in (2.1) corresponds to the Inhomogeneous Khintchine theorem stated in §1. The following two results of this paper are Hausdorff measure generalisations of Theorem 1.2.

Theorem 2.1 (Divergence case). Let $n \geq 2$, $\boldsymbol{\theta} \in \mathbb{R}^n$, ψ be a non-increasing function such that $\psi(q)^{1/\eta}q \to \infty$ as $q \to \infty$, where η is given by (1.7). Let d, \mathbf{U} and \mathbf{f} be as in Theorem 1.2 and d+m=n. Then for any $\frac{(n+1)}{1+\eta}-m < s \leq d$, we have that

$$\mathcal{H}^{s}\left(\mathbf{f}^{-1}(\mathcal{S}_{n}^{\boldsymbol{\theta}}(\psi))\right) = \mathcal{H}^{s}(\mathbf{U}) \text{ if } \sum q^{n} \left(\frac{\psi(q)}{q}\right)^{s+m} = \infty.$$
 (2.2)

In particular, if $\tau(\psi) := \liminf_{q \to \infty} \frac{-\log \psi(q)}{\log q}$ satisfies $\frac{1}{n} < \tau(\psi) < \eta$, then

$$\dim_H \left(\mathbf{f}^{-1}(\mathcal{S}_n^{\boldsymbol{\theta}}(\psi)) \right) \ge \frac{n+1}{\tau(\psi)+1} - m. \tag{2.3}$$

Note that the divergence case of Theorem 1.2 is the s=d case of Theorem 2.1.

Remark 2.1. For d=1, Theorem 2.1 was previously established in [11, Theorem 1.3], however our proof gives an alternate simpler argument. The main new substance of Theorem 2.1 is the case d>1 which extends [5, Theorem 2.5] from analytic to any nondegenerate manifolds, although, when d>1, the range of s (resp. the range of $\tau(\psi)$) we have in Theorem 2.1 is smaller than that in [5, Theorem 2.5].

Theorem 2.2 (Convergence case). Let $n \geq 2$, $\theta \in \mathbb{R}^n$, ψ be a non-increasing function, d, U and f be as in Theorem 1.2 and d + m = n. Suppose that s > 0,

$$\mathcal{H}^s\{\mathbf{x} \in \mathbf{U} : \mathbf{f} \text{ is not } l\text{-nondegenerate at } \mathbf{x}\} = 0,$$
 (2.4)

$$\sum_{q>1} q^n \left(\frac{\psi(q)}{q}\right)^{s+m} < \infty \tag{2.5}$$

and

$$\sum_{t>1} \left(\frac{\psi(e^t)}{e^t} \right)^{(s-d)/2} \left(\psi(e^t)^{n-1/2} e^{3t/2} \right)^{-\alpha} < \infty , \qquad (2.6)$$

where

$$\alpha = \frac{1}{d(2l-1)(n+1)}. (2.7)$$

Then

$$\mathcal{H}^s\left(\mathbf{f}^{-1}(\mathcal{S}_n^{\boldsymbol{\theta}}(\psi))\right) = 0. \tag{2.8}$$

Remark 2.2. It should be noted that even in the case of $\theta = 0$, the above theorem is new. It removes the smoothness condition in [32, Theorem 1.11] and also improves the summation condition in [32, Equation 1.15].

Next, we have the following corollary that follows on combining Theorem 2.2 and Theorem 2.1. It improves the range of τ in [13, Corollary 2.9], removes the smoothness assumption from [32, Corollary 1.13], and extends these results to any $\theta \in \mathbb{R}^n$. We write $S_n^{\theta}(\tau)$ for $S_n^{\theta}(\psi)$ with $\psi(q) = q^{-\tau}$.

Corollary 2.1 (Exact Hausdorff dimension). Let $n \geq 2$, m, d, \mathbf{f} and α be as before, and

$$\dim_{H}(\{\mathbf{x} \in \mathbf{U} \mid \mathbf{f} \text{ is not } l\text{-nondegenerate at } \mathbf{x}\}) \le \frac{n+1}{\tau+1} - m, \tag{2.9}$$

where

$$\frac{1}{n} < \tau < \frac{3\alpha + 1}{\alpha(2n - 1) + n}$$
 (2.10)

Then

$$\dim_{H}(\mathbf{f}^{-1}(\mathcal{S}_{n}^{\boldsymbol{\theta}}(\tau))) = \frac{n+1}{\tau+1} - m. \tag{2.11}$$

2.1. **Spectrum of inhomogeneous exponents.** We discuss the implications of our results for the inhomogeneous version of a problem by Bugeaud and Laurent regarding the Diophantine exponents of points on the Veronese curves $(x, x^2, ..., x^n)$. Given an $x \in \mathbb{R}$ and $\theta \in \mathbb{R}^n$, first recall the definition of inhomogeneous exponents from [22]:

$$\lambda_n^{\boldsymbol{\theta}}(x) := \sup\{\tau \mid (x, x^2, \dots, x^n) \in \mathcal{S}_n^{\boldsymbol{\theta}}(\tau)\}. \tag{2.12}$$

Define

$$\operatorname{Spec}^{\theta}(\lambda_n) = \left\{ \lambda \in [0, \infty] \mid \lambda_n^{\theta}(x) = \lambda \text{ for some } x \in \mathbb{R} \right\}.$$

Note that when $\boldsymbol{\theta} = \mathbf{0}$, $\lambda_n^{\mathbf{0}}(\frac{p}{q}) = \infty$ for every $\frac{p}{q} \in \mathbb{Q}$ and by Dirichlet's theorem $\lambda_n^{\mathbf{0}}(x) \geq 1/n$ for every $x \in \mathbb{R}$. Thus the definition in [21] and [13] matches with the above one when $\boldsymbol{\theta} = \mathbf{0}$. By Theorem 1.2 we get that for almost every $x \in \mathbb{R}$, $\lambda_n^{\boldsymbol{\theta}}(x) = \frac{1}{n}$.

Thus we know $\frac{1}{n} \in \operatorname{Spec}^{\theta}(\lambda_n)$. The following is a special case of Corollary 2.1 for $n \geq 2$, and m = n - 1, d = 1. Then for $\mathbf{f}(x) = (x, \dots, x^n)$, and τ satisfying

$$\frac{1}{n} \le \tau < \frac{2n^2 + n + 2}{(n^2 + n + 1)(2n - 1)},\tag{2.13}$$

we have

$$\dim_{H}(\{\mathbf{x} \in \mathbf{U} : \mathbf{f}(\mathbf{x}) \in \mathcal{S}_{n}^{\boldsymbol{\theta}}(\tau)\}) = \frac{n+1}{\tau+1} - n + 1.$$
 (2.14)

This in particular gives the following corollary.

Corollary 2.2. For every n > 3,

$$\left[\frac{1}{n}, \frac{1}{n} + \frac{n+1}{n(2n-1)(n^2+n+1)}\right] \subset \operatorname{Spec}^{\theta}(\lambda_n). \tag{2.15}$$

3. Sharp Quantitative Nondivergence

The proofs of the main result will rely on an effective upper bound for the measure of the set $\mathfrak{S}_{\mathbf{f}}(\delta, K, \mathbf{T})$ given by

$$\left\{ \mathbf{x} \in \mathbf{U} : \exists (a_0, \mathbf{a}) \in \mathbb{Z} \times \mathbb{Z}_{\neq \mathbf{0}}^n \text{ such that } \|\nabla \mathbf{f}(\mathbf{x}) \mathbf{a}^\top\| < K \\ |a_i| < T_i (1 \le i \le n) \right\}, \tag{3.1}$$

where $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{T} = (T_1, \dots, T_n)$, $\delta, K, T_1, \dots, T_n$ are positive real parameters, $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ is a map defined on an open set $\mathbf{U} \subset \mathbb{R}^d$, ∇ stands for the gradient of a real-valued function, $\|\cdot\|$ denotes the sup norm and \mathbf{a}^{\top} denotes the transpose of \mathbf{a} . An upper bound on the measure of this set was already established in [18]; here we provide an improved estimates which is crucial in order to get a larger range of exponents $\tau(\psi)$ in Theorem 2.1 than we could get otherwise. We prove the following theorem which is of independent interest.

Theorem 3.1. Let $\mathbf{U} \subset \mathbb{R}^d$ be open, $\mathbf{x}_0 \in \mathbf{U}$ and $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be l-nondegenerate at \mathbf{x}_0 . Write n = m + d and suppose that

$$\mathbf{f} = (x_1, \dots, x_d, f_1(\mathbf{x}), \dots, f_m(\mathbf{x})), \quad \text{where } \mathbf{x} = (x_1, \dots, x_d). \tag{3.2}$$

Then there exists a ball $B_0 \subset \mathbf{U}$ centred at \mathbf{x}_0 and a constant E > 0 depending on B_0 and \mathbf{f} such that for any choice of $\delta, K, T_1, \dots, T_n$ satisfying

$$0 < \delta \le 1,$$
 $T_1, \dots, T_n \ge 1,$ $K > 0$ and $\delta^n < K \frac{T_1 \cdots T_n}{\max_i T_i}$ (3.3)

and any ball $B \subset B_0$ of radius $0 < r \le 1$ we have that

$$\mathcal{L}_d(\mathfrak{S}_{\mathbf{f}}(\delta, K, \mathbf{T}) \cap B) \ll$$

$$\ll \delta \prod_{i=1}^{d} \min\{K, T_i\} \cdot (T_{d+1} \cdots T_n) \mathcal{L}_d(B) + E\left(\delta \min\{K, r^{-1}\} \frac{T_1 \cdots T_n}{\max_i T_i}\right)^{\alpha}, \quad (3.4)$$

where α is given by (2.7) and the implied constant depends on l, m and d only.

Theorem 3.1 can be regarded as a refinement of the following simplified version of Theorem 1.4 from [18], which is also used in [13] in this form.

Theorem 3.2 ([18, Theorem 1.4]). Let $\mathbf{U} \subset \mathbb{R}^d$ be open, $\mathbf{x}_0 \in \mathbf{U}$ and $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be l-nondegenerate at \mathbf{x}_0 . Then there exists a ball $B_0 \subset \mathbf{U}$ centred at \mathbf{x}_0 and a constant $E \geq 1$ such that for any choice of $\delta, K, T_1, \dots, T_n$ satisfying (3.3) we have that

$$\mathcal{L}_d(\mathfrak{S}_{\mathbf{f}}(\delta, K, \mathbf{T}) \cap B_0) \leq E\left(\delta K \frac{T_1 \cdots T_n}{\max_i T_i}\right)^{\alpha} \mathcal{L}_d(B_0), \qquad (3.5)$$

where α is given by (2.7).

We can say that the constant E appearing in (3.4) is related but may be different to the constant E appearing in (3.5); see the details in the proof of Theorem 3.1. We note that Theorem 3.2 together with the following statement appearing as Theorem 1.3 in [18] is already enough to prove a weaker version of Theorem 3.1 in which one would have $\min\{K, T^{1/2}\}$ with $T = \max_{1 \le i \le n} T_i$ in place of the term $\min\{K, r^{-1}\}$ in (3.4).

Theorem 3.3 ([18, Theorem 1.3]). Let $\mathbf{U} \subset \mathbb{R}^d$ be open, $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be a C^2 map. Let B be a ball of radius r > 0 such that $2B \subset \mathbf{U}$, and let

$$L := \sup_{\substack{\mathbf{x} \in 2B \\ |\beta| = 2}} \|\partial_{\beta} \mathbf{f}(\mathbf{x})\| < \infty.$$
 (3.6)

Then, for any $\mathbf{a} \in \mathbb{R}^n$ such that

$$\|\mathbf{a}\| \ge \frac{1}{4nLr^2} \tag{3.7}$$

the set

$$\mathfrak{S}'_{\mathbf{f}}(\delta, \mathbf{a}) := \left\{ \mathbf{x} \in B : \exists \ a_0 \in \mathbb{Z} \ such \ that \quad \begin{aligned} |a_0 + \mathbf{f}(\mathbf{x}) \mathbf{a}^\top| &< \delta \\ \|\nabla \mathbf{f}(\mathbf{x}) \mathbf{a}^\top\| &\ge \sqrt{ndL \|\mathbf{a}\|} \end{aligned} \right\}$$
(3.8)

has measure at most $C_d \delta \mathcal{L}_d(B)$, where $C_d > 0$ is a constant depending on d only.

We note that Theorem 1.3 in [18] is stated for $\mathbf{a} \in \mathbb{Z}^n$, however nowhere is the proof does it use the fact that \mathbf{a} is integer, and assuming that $\mathbf{a} \in \mathbb{R}^n$ is enough. The proof of Theorem 3.1 will use the following strengthening of Theorem 3.3.

Theorem 3.4. Let $\mathbf{U} \subset \mathbb{R}^d$ be open, $\mathbf{x}_0 \in \mathbf{U}$ and $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be l-nondegenerate at \mathbf{x}_0 . Then there exists a ball B_0 centred at \mathbf{x}_0 and satisfying $2B_0 \subset \mathbf{U}$, and a constant $C_1 > 0$ depending on \mathbf{x}_0 such that for any ball $B \subset B_0$ of radius r > 0 and any $\mathbf{a} \in \mathbb{R}^n$ satisfying

$$\|\mathbf{a}\| \ge \max\left\{\frac{1}{4nLr^2}, \frac{ndL}{C_1^2}\right\},$$
 (3.9)

where L is given by (3.6), the set

$$\mathfrak{S}_{\mathbf{f}}''(\delta, \mathbf{a}) := \left\{ \mathbf{x} \in \mathbf{U} : \exists \ a_0 \in \mathbb{Z} \ such \ that \quad \begin{aligned} |a_0 + \mathbf{f}(\mathbf{x})\mathbf{a}^\top| < \delta \\ |\nabla \mathbf{f}(\mathbf{x})\mathbf{a}^\top| \ge r^{-1} \end{aligned} \right\}$$
(3.10)

satisfies

$$\mathcal{L}_d(\mathfrak{S}''_{\mathbf{f}}(\delta, \mathbf{a}) \cap B) \le C_{l,d} \delta \mathcal{L}_d(B)$$
,

where $C_{l,d} > 0$ is a constant depending on l and d only.

We remark that since \mathbf{f} is non-degenerate at \mathbf{x}_0 , we necessarily have that $\mathbf{f} \in C^2$ on a neighbourhood of \mathbf{x}_0 . This means that the constant L, given by (3.6), is well defined. Furthermore, we can assume that L is non-zero.

The proof of Theorem 3.4 will use the following proposition which extends [3, Proposition 1].

Proposition 3.1. Let $0 < \delta < 1$, $k \ge 2$, $\Theta \ge 1/2$, $I \subset \mathbb{R}$ be an interval and $F: I \to \mathbb{R}$ be a C^k function such that $\inf_{x \in I} |F^{(k)}(x)| > 0$. Let $\mathfrak{S}^{(1)}(F, \delta, \Theta, I)$ be the set of $x \in I$ such that

$$|F(x) + a| < \delta, \qquad |F'(x)| > \Theta^{-1} \mathcal{L}_1(I)^{-1}$$
 (3.11)

for some $a \in \mathbb{Z}$. Then

$$\mathcal{L}_1(\mathfrak{S}^{(1)}(F,\delta,\Theta,I)) \le 8k\Theta\delta\mathcal{L}_1(I). \tag{3.12}$$

Proof. Without loss of generality, we will assume that $\delta < \frac{1}{4}$ as otherwise (3.12) is trivially true. Let $\sigma(a)$ be the set of $x \in I$ satisfying (3.11) for a given a. Clearly, $\sigma(a)$ can be written as a union of non-empty disjoint intervals $\sigma(a) = \bigcup_i \sigma_i(a)$. In what follows we will assume without loss of generality that this union is minimal, that is $\sigma_i(a) \cup \sigma_j(a)$ is not an interval if $i \neq j$. Fix any $\varepsilon > 0$ and for each choice of a and i, fix any $\alpha_{a,i} \in \sigma_i(a)$ such that

$$\inf_{x \in \sigma_i(a)} |F'(x)| \ge \frac{|F'(\alpha_{a,i})|}{1+\varepsilon} \,. \tag{3.13}$$

By the Mean Value Theorem, (3.11) and (3.13), $\mathcal{L}_1(\sigma_i(a)) \leq 2(1+\varepsilon)\delta|F'(\alpha_{a,i})|^{-1}$. Therefore,

$$\mathcal{L}_1(\mathfrak{S}^{(1)}(F,\delta,I)) \le \sum_{a} \sum_{i} \mathcal{L}_1(\sigma_i(a)) \le 2(1+\varepsilon)\delta \sum_{a} \sum_{i} |F'(\alpha_{a,i})|^{-1}.$$
 (3.14)

Since $\inf_{x\in I} |F^{(k)}(x)| > 0$, F' and F'' have $s \leq (k-1) + (k-2) = 2k-3$ zeros inside I. Thus I can be written as a disjoint union of subintervals I_0, \ldots, I_s such that the above zeros are one of the endpoints of these subintervals. For each j, let A_j be the set of all $\alpha_{a,i} \in I_j$. Thus $\bigcup_a \bigcup_i \{\alpha_{a,i}\} = \bigcup_j A_j$. Note that on each interval I_j , F' is monotonic and does not change sign. For simplicity we will assume that F' is increasing (the decreasing case is considered exactly the same way) and positive. Therefore F is increasing. Then, we claim that if $\alpha_{a_1,i_1} \leq \alpha_{a_2,i_2}$ are two consecutive elements in I_j then we must have that $a_1 \neq a_2$. To prove this, assume for the moment that the claim is not true and let $\alpha_{a,i_1} \leq \alpha_{a,i_2}$ be two consecutive elements in I_j . Let c_1 be the right most end point of $\sigma_{i_1}(a)$, c_2 be the left most end point of $\sigma_{i_2}(a)$. By the minimality of the union $\sigma(a) = \bigcup_i \sigma_i(a)$, we must have that $c_1 < c_2$ and $|F(c_i) + a| = \delta$ for i = 1, 2. This implies that F' must vanish between c_1 and c_2 , contrary to the fact that F' is

not zero inside I_j , thus proving the claim. Next, using the inequalities $\delta < \frac{1}{4}$ and $|F(\alpha_{a_\ell,i_\ell}) + a_\ell| < \delta$ for $\ell = 1, 2$, we obtain that

$$\frac{1}{2} \le |a_1 - a_2| - 2\delta \le |F(\alpha_{a_1, i_1}) - F(\alpha_{a_2, i_2})|.$$

Then, using the Mean Value Theorem and the assumption that F' is increasing, we get that

$$\frac{1}{2} \le |F(\alpha_{a_1,i_1}) - F(\alpha_{a_2,i_2})| \le |F'(\alpha_{a_2,i_2})|(\alpha_{a_2,i_2} - \alpha_{a_1,i_1}),$$

whence

$$|F'(\alpha_{a_2,i_2})|^{-1} \le 2(\alpha_{a_2,i_2} - \alpha_{a_1,i_1}).$$

On summing the latter over $\alpha_{a_2,i_2} \in I_j$ and using the fact that the sum of right hand side is $\leq 2\mathcal{L}_1(I_j)$, we get that

$$\sum_{\alpha_{a,i} \in A_j} |F'(\alpha_{a,i})|^{-1} \le 2\mathcal{L}_1(I_j) + \max_{\alpha_{a,i} \in A_j} |F'(\alpha_{a,i})|^{-1} \stackrel{\text{(3.11)}}{\le} 2\mathcal{L}_1(I_j) + \Theta \mathcal{L}_1(I).$$

Finally, since the number of different intervals I_j is $\leq 2k-2$ and $\sum_j \mathcal{L}_1(I_j) = \mathcal{L}_1(I)$, we get that

$$\sum_{a} \sum_{i} |F'(\alpha_{a,i})|^{-1} \le 2\mathcal{L}_1(I) + (2k-2)\Theta \mathcal{L}_1(I) \le (2k+2)\Theta \mathcal{L}_1(I),$$

since $\Theta > \frac{1}{2}$. Combining this with (3.14) and letting $\varepsilon \to 0$ gives the required bound.

Proof of Theorem 3.4. Since **f** is *l*-non-degenerate at \mathbf{x}_0 , there is a constant $C_0 > 0$ such that for any $\mathbf{a} \in \mathbb{R}^n$ with $\|\mathbf{a}\| \geq 1$ there exists a multi-index β with $1 \leq |\beta| = k \leq l$ such that

$$|\partial_{\beta} \mathbf{f}(\mathbf{x}_0) \mathbf{a}^{\top}| \geq C_0 ||\mathbf{a}||.$$

This can be explained as follows. First, let us consider

$$\mathcal{F} := \{g = \mathbf{f}(\cdot)\mathbf{a}^\top \mid \mathbf{a} \in \mathbb{R}^n, \|\mathbf{a}\| = 1\}.$$

By the compactness of \mathcal{F} and the nondegeneracy of \mathbf{f} at \mathbf{x}_0 , we have that

$$\inf_{g \in \mathcal{F}} \max_{|\beta| \le l} |\partial_{\beta} g(\mathbf{x}_0)|$$

is attained at some g and is nonzero. Thus we can take $C_0 = \inf_{g \in \mathcal{F}} \max_{|\beta| \leq l} |\partial_{\beta} g(\mathbf{x}_0)|$. The subsequent part of the proof will use the following

Claim: There exists a rotation of the coordinate systems in \mathbb{R}^d around \mathbf{x}_0 , which may depend on \mathbf{a} , such that $|\tilde{\partial}_i^k \mathbf{f}(\mathbf{x}_0) \mathbf{a}^\top| \geq 2C_1 \|\mathbf{a}\|$ for $i = 1, \ldots, d$ for some $C_1 > 0$ independent of \mathbf{a} , where $\tilde{\partial}_i^k$ refers to derivation in the rotated coordinate system.

For completeness we will provide a proof of this claim in §3.1. For now we continue with the proof of Theorem 3.4. By continuity, there is an open neighbourhood $U_0 \subset U$ of \mathbf{x}_0 such that for every $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ there is an appropriate rotation of the coordinate systems around \mathbf{x}_0 such that

$$|\tilde{\partial}_i^k \mathbf{f}(\mathbf{x}) \mathbf{a}^\top| \ge C_1 \|\mathbf{a}\| \quad \text{for } i = 1, \dots, d, \text{ and for all } \mathbf{x} \in \mathbf{U}_0.$$
 (3.15)

Now let B_0 be a (Euclidean) ball centred at \mathbf{x}_0 such that whenever a ball B lies in B_0 , any cube \hat{B} circumscribed around B is contained in \mathbf{U}_0 .

Fix any ball $B \subset B_0$ and let r denote the radius of B. Fix any $\mathbf{a} \in \mathbb{R}^n$ with $\|\mathbf{a}\| \geq 1$ and let \hat{B} be the cube circumscribed around B which is aligned with the rotated coordinate system for which (3.15) holds. By the choice of B_0 we have that $\hat{B} \subset \mathbf{U}_0$. Clearly,

$$\mathcal{L}_d(\mathfrak{S}''_{\mathbf{f}}(\delta, \mathbf{a}) \cap B) \le \mathcal{L}_d(\mathfrak{S}''_{\mathbf{f}}(\delta, \mathbf{a}) \cap \hat{B}). \tag{3.16}$$

Since the rotation of the coordinate system rotates the gradient, the second inequality in (3.10) implies that $\|\tilde{\nabla}\mathbf{f}(\mathbf{x})\mathbf{a}^{\top}\| \geq C_2 r^{-1}$ in the new coordinate system, where $0 < C_2 \leq 1$ depends on d only. The condition on the gradient means that for some $1 \leq i \leq d$ we have that $|\tilde{\partial}_i \mathbf{f}(\mathbf{x})\mathbf{a}^t| \geq C_2 r^{-1}$.

Now, with the view to using Proposition 3.1 we perform a reduction to dimension one. Indeed, by Fubini's theorem, the right hand side of (3.16) is bounded by

$$d(2r)^{d-1} \sup_{\mathbf{x}'} \mathcal{L}_1(\mathfrak{S}^{(1)}(F_{\mathbf{x}'}, \delta, \Theta, I)), \qquad (3.17)$$

where \mathbf{x}' is the projection of $\mathbf{x} = (x_1, \dots, x_d)$ obtained by removing x_i for some i and $F_{\mathbf{x}'}(x_i) = \mathbf{f}(x_1, x_2, \dots, x_d)\mathbf{a}^{\top}$ is defined on an interval I of length $\mathcal{L}_1(I) = 2r$ and $\Theta = C_2^{-1}$. Suppose that $k \geq 2$. Then, by Proposition 3.1 and the fact that $k \leq l$, we get that (3.17) is bounded by $8dlC_2^{-1}\delta(2r)^d$, which implies the required result.

It remains to consider the case k = 1. First, note that by (3.15), we have $\|\tilde{\nabla}\mathbf{f}(\mathbf{x})\mathbf{a}^{\top}\| \ge C_1\|\mathbf{a}\|^{1/2}\|\mathbf{a}\|^{1/2} \ge C_1\frac{\sqrt{ndL\|\mathbf{a}\|}}{C_1} \ge r^{-1}$ for all $\mathbf{x} \in \hat{B}$. Thus the second inequality in the definition of $\mathfrak{S}''_{\mathbf{f}}(\delta, \mathbf{a}) \cap \hat{B}$ is redundant, and moreover using (3.9) the aforementioned set is the same as $\mathfrak{S}'_{\mathbf{f}}(\delta, \mathbf{a}) \cap \hat{B}$. Thus, we apply Theorem 3.3 to complete the proof. \square

Proof of Theorem 3.1. If $K < r^{-1}$ then $\min\{K, r^{-1}\} = K$ and (3.4) becomes an immediate consequence of Theorem 3.2. Thus for the rest of the proof we will assume that $K \ge r^{-1}$. In this case, since $\delta^n \le 1 \le r^{-1} \frac{T_1 \cdots T_n}{\max_i T_i}$, by Theorem 3.2 again we have that

$$\mathcal{L}_d(\mathfrak{S}_{\mathbf{f}}(\delta, r^{-1}, \mathbf{T})) \ll E\left(\delta \min\{K, r^{-1}\} \frac{T_1 \cdots T_n}{\max_i T_i}\right)^{\alpha}$$
(3.18)

while

$$\mathfrak{S}_{\mathbf{f}}(\delta, K, \mathbf{T}) \setminus \mathfrak{S}_{\mathbf{f}}(\delta, r^{-1}, \mathbf{T}) \subset \bigcup_{\mathbf{a}} \mathfrak{S}_{\mathbf{f}}''(\delta, \mathbf{a}),$$
 (3.19)

where the union is taken over all $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n$ such that for some $\mathbf{x} \in B_0$

$$\|\nabla \mathbf{f}(\mathbf{x})\mathbf{a}^{\top}\| < K \quad \text{and} \quad |a_i| \le T_i \quad (1 \le i \le n).$$
 (3.20)

By the right hand side of (3.20), there are at most $\ll T_{d+1} \cdots T_n$ choices for a_{d+1}, \ldots, a_n . In turn, in view of (4.3), by the left and right hand sides of (3.20), for each fixed choice of a_{d+1}, \ldots, a_n , there are at most $\ll \prod_{i=1}^d \min\{K, T_i\}$ choices for a_1, \ldots, a_d . Thus, the union in (3.19) contains $\ll \prod_{i=1}^d \min\{K, T_i\} T_{d+1} \cdots T_n$ sets. In the case (3.9) is satisfied, Theorem 3.4 is applicable and the corresponding contribution from the union

in (3.19) coincides with the first term of the estimate in (3.4). Finally, the union in the right of (3.19) taken over **a** such that

$$\|\mathbf{a}\| \le \max\left\{\frac{1}{4nLr^2}, \frac{ndL}{C_1^2}\right\} \tag{3.21}$$

is obviously contained in $\mathfrak{S}_{\mathbf{f}}(\delta, T_0, T_0, \cdots, T_0)$ with

$$T_0 = C_0 \max\left\{1, \frac{1}{4nLr^2}, \frac{ndL}{C_1^2}\right\} \ge 1$$

for a suitably chosen constant C_0 depending on \mathbf{f} only. In this last case we again apply Theorem 3.2, and obtain an upper bound $\ll_{\mathbf{f},B} E\delta^{\alpha} \leq E\left(\delta \min\{K,r^{-1}\}\frac{T_1\cdots T_n}{\max_i T_i}\right)^{\alpha}$ since $K \geq r^{-1} \geq 1$.

3.1. **Proof of the Claim.** Here we give a proof of the Claim stated within the proof of Theorem 3.4. This will be done in two steps: first we show the existence of a suitable constant for a given **a**, then we prove that the constant can be chosen to be independent of **a**.

Step 1: Let us fix $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. First, we show that there exists a rotation of coordinate system (depending on \mathbf{a}) around \mathbf{x}_0 such that $|\tilde{\partial}_i^k \mathbf{f}(\mathbf{x}_0) \mathbf{a}^\top| \neq 0$ for every $i = 1, \ldots, d$. Here $\tilde{\partial}$ denotes differentiation in the new coordinate system. To this end, given $A = (a_{ij}) \in \mathrm{GL}_d(\mathbb{R})$, differentiating the map $g := \mathbf{f}(\cdot)\mathbf{a}^\top$ we have, by the chain rule, that

$$\begin{array}{rcl} \partial_{1}^{k}g \circ A(A^{-1}\mathbf{x}_{0}) & = \sum_{\alpha=(i_{1},...,i_{d}),|\alpha|=k} a_{11}^{i_{1}} \cdots a_{d1}^{i_{d}} \; \partial_{\alpha}g(\mathbf{x}_{0}) \; , \\ & \vdots \\ \partial_{d}^{k}g \circ A(A^{-1}\mathbf{x}_{0}) & = \sum_{\alpha=(i_{1},...,i_{d}),|\alpha|=k} a_{1d}^{i_{1}} \cdots a_{dd}^{i_{d}} \; \partial_{\alpha}g(\mathbf{x}_{0}) \; . \end{array}$$

We wish to find $A \in O_n(\mathbb{R})$ such that $x_i' \neq 0$ for all i = 1, ..., d where

$$x'_1 = P(a_{11}, \dots, a_{d1}),$$

 \vdots
 $x'_d = P(a_{1d}, \dots, a_{dd}),$

and

$$P(x_1, \dots, x_d) = \sum_{i_1 + \dots + i_d = k} x_{(i_1, \dots, i_d)}^0 x_1^{i_1} \cdots x_d^{i_d},$$

and $x_{(i_1,\ldots,i_d)}^0 = \partial_{(i_1,\ldots,i_d)}g(\mathbf{x}_0)$. Note that P is a homogeneous polynomial of degree k. Since $\partial_{\beta}g(\mathbf{x}_0) \neq 0$, i.e., $x_{\beta}^0 \neq 0$, P is a nonzero polynomial. Also, since P is homogeneous, if $P(\mathbf{y}) \neq 0$ for some $\mathbf{y} = (y_1,\ldots,y_d) \neq \mathbf{0}$, we also have $P(\mathbf{y}') \neq 0$ for $\mathbf{y}' = \mathbf{y}/\|\mathbf{y}\|$.

We prove by induction that for any $s \leq d$ we can construct nonzero orthonormal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ such that $\prod_{i=1}^s P(\mathbf{v}_i) \neq 0$. When s=1, the statement follows since P is a nonzero and homogeneous. Suppose $\mathbf{v}_1, \cdots, \mathbf{v}_s, s < d$ are constructed such that $\prod_{i=1}^s P(\mathbf{v}_i) \neq 0$. Since P is nonzero at each of these vectors, there exists $W_i, i=1,\ldots,s$ open balls centered at \mathbf{v}_i of radius r_i such that P is nonzero at each point in W_i for

 $i=1,\ldots,s$. Let \mathcal{L} be the subspace that is perpendicular to each \mathbf{v}_i . Let us take any unit vector (length 1) $\mathbf{v} \in \mathcal{L}$, and $V:=\{cO\mathbf{v} \mid \|O-\mathrm{I}\| < r, O \in \mathrm{O}(n), c>0\}$ an open neighborhood of \mathbf{v} , where $r<\min r_i$. Since P is nonzero polynomial, there exists a unit vector $\mathbf{w} \in V$ such that $P(\mathbf{w}) \neq 0$. Thus $\mathbf{w} = O\mathbf{v}$ for some $O \in \mathrm{O}(n)$ with $\|O-\mathrm{I}\| < r$. Then for $i=1,\ldots,s, \|O\mathbf{v}_i-\mathbf{v}_i\| < r < r_i$, hence $\mathbf{v}_i'=O\mathbf{v}_i \in W_i$. By the construction of $W_i, P(\mathbf{v}_i') \neq 0$ for $i=1,\ldots,s$. Since $\mathbf{v}_1,\ldots,\mathbf{v}_s,\mathbf{v}$ are orthonormal, $\mathbf{v}_1',\ldots,\mathbf{v}_s',\mathbf{w}$ are also orthonormal. Let us denote $\mathbf{w} = \mathbf{v}_{s+1}'$. Thus we constructed orthonormal vectors $\mathbf{v}_i', i=1,\ldots,s+1$ such that $\prod_{i=1}^{s+1} P(\mathbf{v}_i') \neq 0$.

To summarize the above, we have that for every $\mathbf{a} \in \mathbb{R}^n \setminus \{0\}$, there exists $O_{\mathbf{a}} \in O_n(\mathbb{R})$ such that $\tilde{\partial}_i^k \mathbf{f}(\mathbf{x}_0) \mathbf{a}^\top := \partial_i^k \mathbf{f} \circ O_{\mathbf{a}}(O_{\mathbf{a}}^{-1}\mathbf{x}_0) \mathbf{a}^\top \neq 0$, for all $i = 1, \ldots, d$. This implies

$$\min_{i=1}^{d} |\partial_i^k \mathbf{f} \circ O_{\mathbf{a}}(O_{\mathbf{a}}^{-1} \mathbf{x}_0) \mathbf{a}^{\top}| > 0.$$
 (3.22)

This concludes the first part of the claim.

Step 2: Finally, we show the existence of $C_1 > 0$ as in the claim. Let

$$C_1 := \inf_{\|\mathbf{a}\|=1} \sup_{O \in O_n(\mathbb{R})} \min_{i=1}^d |\partial_i^k \mathbf{f} \circ O(O^{-1}\mathbf{x}_0) \mathbf{a}^\top|.$$

The goal is to show that $C_1 > 0$. We prove this by contradiction. If $C_1 = 0$, then there exists a sequence \mathbf{a}_m with $\|\mathbf{a}_m\| = 1$ such that

$$\sup_{O \in O_n(\mathbb{R})} \min_{i=1}^d |\partial_i^k \mathbf{f} \circ O(O^{-1}\mathbf{x}_0) \mathbf{a}_m^\top| < \frac{1}{m}.$$

By the compactness of the unit sphere $\{\mathbf{a} \in \mathbb{R}^n \mid ||\mathbf{a}|| = 1\}$, there exists a convergent subsequence $\mathbf{a}_{m_l} \to \mathbf{a}$, where $||\mathbf{a}|| = 1$, such that

$$\sup_{O \in O_n(\mathbb{R})} \min_{i=1}^d |\partial_i^k \mathbf{f} \circ O(O^{-1} \mathbf{x}_0) \mathbf{a}^\top| = 0.$$

This contradicts (3.22) and completes the proof of the claim.

4. Auxiliary statements and definitions

4.1. Some geometry of numbers. Before moving on to the proofs of the main results, we recall some auxiliary facts about lattices. In what follows, given a lattice Λ in \mathbb{R}^{n+1} ,

$$\delta_1(\Lambda), \ldots, \delta_{n+1}(\Lambda)$$

will denote Minkowski's successive minima of the closed unit Euclidean ball centered at $\mathbf{0}$ on the lattice Λ . For $g \in \mathrm{GL}_{n+1}(\mathbb{R})$, by Minkowski's second theorem, we have that

$$\delta_1(g\mathbb{Z}^{n+1})\cdots\delta_{n+1}(g\mathbb{Z}^{n+1}) \asymp_n |\det g|$$
.

In particular,

$$\frac{|\det g|}{\delta_1(g\mathbb{Z}^{n+1})} \ll_n \delta_{n+1}(g\mathbb{Z}^{n+1})^n. \tag{4.1}$$

For any $g \in GL_{n+1}(\mathbb{R})$, we define the dual matrix $g^* = \sigma^{-1}(g^{\top})^{-1}\sigma$. Here and elsewhere, A^{\top} denotes the transpose of the matrix/row/column A, $\sigma = \sigma_{n+1}$ and

$$\sigma_k = \begin{bmatrix} 0 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 0 \end{bmatrix}$$

is the long Weyl element of $GL_k(\mathbb{R})$, where k is the size of the matrix.

We state the following lemma that appears in [13, Lemma 3.3]. We note that it was stated for $g \in SL_{n+1}(\mathbb{R})$ in [13] but it is true for $g \in GL_{n+1}(\mathbb{R})$ for the same reasons given in [13], in particular by Theorems 21.5 and 23.2 from [25].

Lemma 4.1. Let $g \in GL_{n+1}(\mathbb{R})$ then

$$\delta_1(q\mathbb{Z}^{n+1})\delta_{n+1}(q^*\mathbb{Z}^{n+1}) \simeq_n 1. \tag{4.2}$$

4.2. Further assumptions and notation. Write n = d + m with $d, m \in \mathbb{Z}$. We will assume throughout that $1 \leq d < n$ as otherwise the manifold is essentially \mathbb{R}^n . As we already mentioned before, in view of the nonsingularity assumption on \mathbf{f} , by the Implicit Function Theorem, there will also be no loss of generality in assuming that \mathbf{f} has the following form

$$\mathbf{f} = (\mathbf{x}, f(\mathbf{x})), \text{ where } f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) \text{ and } \mathbf{x} = (x_1, \dots, x_d).$$
 (4.3)

Since d < n and \mathbf{f} is nondegenerate, we must have that \mathbf{f} is at least C^2 in the context of our theorems. Furthermore, by shrinking \mathbf{U} if necessary, we can assume without loss of generality that for some constant $M \ge 1$

$$\max_{1 \le k \le m} \max_{1 \le i, j \le d} \sup_{\mathbf{x} \in \mathbf{U}} \max \left\{ |\partial_i f_k(\mathbf{x})|, |\partial_{i,j}^2 f_k(\mathbf{x})| \right\} \le M. \tag{4.4}$$

For a vector $\mathbf{y} \in \mathbb{R}^n$ and $1 \leq i \leq n$, we denote the *i*-th coordinate of \mathbf{y} by y_i . Given $\mathbf{x} \in \mathbb{R}^k$ for some k, by $B(\mathbf{x}, r)$ (resp. $B[\mathbf{x}, r]$), we denote the open ball (resp. closed ball) in \mathbb{R}^k centred at \mathbf{x} of radius r, which is obtained using the sup norm on \mathbb{R}^k . By $\|\cdot\|$, we denote the sup norm in \mathbb{R}^k . By $[\cdot]^k$ we denote the hyper-cube (i.e. ball with sup norm) in \mathbb{R}^k of the specified radius and centred at $\mathbf{0}$. For example, $[\pi]^k$ will stand for $[-\pi, \pi]^k$.

Throughout, I_k will denote the identity matrix of $GL_k(\mathbb{R})$. Now we recall some matrices introduced in [13]. For every $\mathbf{x} \in \mathbf{U}$ define

$$J(\mathbf{x}) = \left[\partial_i f_j(\mathbf{x})\right]_{1 < i < d, 1 < j < m},$$

which is the Jacobian of the map $\mathbf{x} \to f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$. Further, following [13] for any given $\mathbf{x} \in \mathbf{U}$ define the following elements of $\mathrm{SL}_{n+1}(\mathbb{R})$:

$$z(\mathbf{x}) = \begin{bmatrix} \mathbf{I}_m & -\sigma_m^{-1} J(\mathbf{x}) \sigma_d & 0 \\ 0 & \mathbf{I}_d & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$u(\mathbf{x}) = \begin{bmatrix} \mathbf{I}_n & \sigma_n^{-1} \mathbf{f}(\mathbf{x})^\top \\ \mathbf{0} & 1 \end{bmatrix},$$

$$u_1(\mathbf{x}) = z(\mathbf{x})u(\mathbf{x})$$
.

It can be easily seen that $u_1(x)$ is the following upper triangular matrix:

$$u_{1}(\mathbf{x}) = \begin{bmatrix} 1 & 0 & \cdots & 0 & -\partial_{d} f_{m}(\mathbf{x}) & \cdots & -\partial_{1} f_{m}(\mathbf{x}) & f_{m}(\mathbf{x}) - \sum_{i=1}^{d} x_{i} \partial_{i} f_{m}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\partial_{d} f_{1}(\mathbf{x}) & \cdots & -\partial_{1} f_{1}(\mathbf{x}) & f_{1}(\mathbf{x}) - \sum_{i=1}^{d} x_{i} \partial_{i} f_{1}(\mathbf{x}) \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & x_{d} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & x_{1} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$(4.5)$$

Let us also introduce the following set closely related to the counting function $N_{\theta}^{\star}(B;\varepsilon,t)$. For every $\varepsilon > 0, B \subset \mathbf{U} \subset \mathbb{R}^d$ and $t \in \mathbb{N}$, we define

$$\mathcal{N}_{\boldsymbol{\theta}}(t, \varepsilon, B) := \left\{ (q, \mathbf{p}_d, \mathbf{p}_m) \in \mathbb{N} \times \mathbb{Z}^d \times \mathbb{Z}^m \middle| \begin{array}{l} \frac{\mathbf{p}_d + \boldsymbol{\theta}_d}{q} \in B, \ e^{t-1} \le q \le e^t \\ \left\| qf\left(\frac{\mathbf{p}_d + \boldsymbol{\theta}_d}{q}\right) - \boldsymbol{\theta}_m - \mathbf{p}_m \right\| \le \varepsilon \end{array} \right\}, \tag{4.6}$$

where f is given as in (4.3) and $\boldsymbol{\theta} = (\boldsymbol{\theta}_d, \boldsymbol{\theta}_m)$ with $\boldsymbol{\theta}_d \in \mathbb{R}^d$ and $\boldsymbol{\theta}_m \in \mathbb{R}^m$. When t > 0 is large enough and $B \subset \mathbf{U}$, one can see using Taylor's expansion that

$$\#\mathcal{N}_{\theta}(t,\varepsilon,B) \ll N_{\theta}^{\star}(B;\varepsilon,t).$$
 (4.7)

5. Lower bounds

First we recall some definitions from [11] adapted to our needs and notation. Given $\varepsilon > 0$, $\rho > 0$, $t \in \mathbb{N}$ and $B \subset \mathbf{U}$, define

$$\Delta_{\boldsymbol{\theta}}(t, \varepsilon, B, \rho) := \bigcup_{(q, \mathbf{p}_d, \mathbf{p}_m) \in \mathcal{N}_{\boldsymbol{\theta}}(t, \varepsilon, B)} B\left(\frac{\mathbf{p}_d + \boldsymbol{\theta}_d}{q}, \rho\right), \tag{5.1}$$

where $\mathcal{N}_{\theta}(t, \varepsilon, B)$ is given by (4.6). Further, given also v > 0, define

$$a_{\varepsilon,t,v} := v^{-1} \operatorname{diag}\{\underbrace{\varepsilon, \dots, \varepsilon}_{m}, \underbrace{(\varepsilon^{m}e^{t})^{-1/d}, \dots, (\varepsilon^{m}e^{t})^{-1/d}}_{d}, v^{n+1}e^{t}\}$$

and the set

$$\mathfrak{G}(v,t,\varepsilon) := \{ \mathbf{x} \in \mathbf{U} : \delta_1(a_{\varepsilon,t,v}^{-1}u_1(\mathbf{x})\mathbb{Z}^{n+1}) \ge v \}.$$
 (5.2)

The matrix $u_1(x)$ coincides with the matrix G defined in [11] up to some permutations of rows and columns and sign change in some columns. Hence, the set $\mathfrak{G}(v,t,\varepsilon)$ given above coincides with the set $\mathfrak{G}(c,Q,\psi)$ defined in [11] upon setting $c=v^{n+1}$, $Q=e^t$ and $\psi=\varepsilon$. In particular, we can now restate Corollary 2.2 from [11] as follows.

Lemma 5.1. Let \mathbf{f} , \mathbf{U} be as before. Suppose that $v \in (0,1]$, M, \tilde{t} , $\tilde{\varepsilon} > 0$ are given such that

$$\tilde{\varepsilon} \ge K_0 e^{-\frac{\tilde{t}(d+2)}{2m+d}}, \quad \text{with } K_0 \gg_{n,d} \frac{1}{v^{n+1}} \left(1 + \frac{M}{v^{n+1}}\right).$$
 (5.3)

Suppose that (4.4) holds. Let t, ε and ρ be defined from the following equations:

$$e^{t} = \frac{e^{t}}{4(n+1)}, \qquad \varepsilon = \frac{\tilde{\varepsilon}}{(1 + \frac{Md^{2}}{v^{n+1}})\frac{n+1}{v^{n+1}}},$$
 (5.4)

$$\rho = \frac{1}{2v^{n+1}} \left(\varepsilon^m e^{t(d+1)} \right)^{-1/d} = C_0 \left(\tilde{\varepsilon}^m e^{\tilde{t}(d+1)} \right)^{-1/d}, \tag{5.5}$$

where $C_0 > 0$ depends on n, m, d, M and v only. Then for any ball $B \subset \mathbf{U}$, we have that

$$\mathfrak{G}(v,t,\varepsilon) \cap B^{\rho} \subset \Delta_{\theta}(\tilde{t},\tilde{\varepsilon},B,\rho), \tag{5.6}$$

where B^{ρ} contains all $x \in B$ such that $B(x, \rho) \subset B$.

By (4.1), for any matrix $g \in \mathrm{SL}_{n+1}(\mathbb{R})$ if $\delta_1(g\mathbb{Z}^{n+1}) \leq v$ then

$$\delta_{n+1}(g\mathbb{Z}^{n+1}) \gg_n v^{-1/n}.$$

Hence

$$\mathbf{U} \setminus \mathfrak{G}(v, t, \varepsilon) = \{ \mathbf{x} \in \mathbf{U} : \delta_1(a_{\varepsilon, t, v}^{-1} u_1(\mathbf{x}) \mathbb{Z}^{n+1}) < v \}$$

$$\subset \{ \mathbf{x} \in \mathbf{U} : \delta_{n+1}(a_{\varepsilon, t, v}^{-1} u_1(\mathbf{x}) \mathbb{Z}^{n+1}) \gg_n v^{-1/n} \}$$

$$\overset{\text{Lemma } 4.1}{\subset} \left\{ \mathbf{x} \in \mathbf{U} : \delta_1((a_{\varepsilon,t,v}^{-1})^* u_1^*(\mathbf{x}) \mathbb{Z}^{n+1}) \ll_n v^{1/n} \right\}. \tag{5.7}$$

Note that

$$(a_{\varepsilon,t,v}^{-1})^* = \sigma^{-1} a_{\varepsilon,t,v} \sigma = v^{-1} \operatorname{diag} \{ v^{n+1} e^t, \underbrace{(\varepsilon^m e^t)^{-1/d}, \dots, (\varepsilon^m e^t)^{-1/d}}_{d}, \underbrace{\varepsilon, \dots, \varepsilon}_{m} \},$$

and observe that $u_1(\mathbf{x})^*$ is given by

$$u_1^{\star}(\mathbf{x}) = \begin{bmatrix} 1 & -\mathbf{x} & -f(\mathbf{x}) \\ 0 & \mathbf{I}_d & J(\mathbf{x}) \\ 0 & 0 & \mathbf{I}_m \end{bmatrix} . \tag{5.8}$$

Therefore, there is a constant $c_n > 0$ depending on n only, and $c_{n,\mathbf{f}} > 0$ depending on n and \mathbf{f} such that (5.7) is contained in $\mathfrak{S}_{\mathbf{f}}(\delta, K, \mathbf{T})$ with

$$\delta = v^{1/n} c_n v^{-n} e^{-t}, \quad K = v^{1/n} c_n v (\varepsilon^m e^t)^{1/d}, \quad T_1 = \dots = T_n = v^{1/n} c_{n,\mathbf{f}} v \varepsilon^{-1}.$$

Now we state the key proposition in this section, which follows from Theorem 3.1 following the discussion above.

Proposition 5.1. Let $\mathbf{f}: \mathbf{U} \subset \mathbb{R}^d \to \mathbb{R}^n$ be a map of the form (4.3), where \mathbf{U} is an open set in \mathbb{R}^d . Suppose that \mathbf{f} is nondegenerate at $\mathbf{x}_0 \in \mathbf{U}$. Let $0 < \kappa < 1$. Then there exists a sufficiently small neighborhood $V \subset \mathbf{U}$ of \mathbf{x}_0 , and constants 0 < v < 1 and $K_0 > 0$ such that for any ball $B \subset V$, there exists $t_0 = t(B, n, \mathbf{f}, \kappa)$ such that for all $t \geq t_0$, and $K_0 e^{-\eta t} < \varepsilon < 1$, where η is as in (1.7), we have that

$$\mathcal{L}_d(B \setminus \mathfrak{G}(v, t, \varepsilon)) < \kappa \mathcal{L}_d(B). \tag{5.9}$$

Theorem 5.1. Let $\mathbf{f}: \mathbf{U} \subset \mathbb{R}^d \to \mathbb{R}^n$ be a map of the form (4.3), where \mathbf{U} is an open set in \mathbb{R}^d , $\boldsymbol{\theta} \in \mathbb{R}^n$. Suppose \mathbf{f} is nondegenerate at $\mathbf{x}_0 \in \mathbf{U}$. Then there exists a small open ball V centered at \mathbf{x}_0 , and constants $C_0, K_0 > 0$ such that for every ball $B \subset V$, there exists $t_0 = t_0(B, n, \mathbf{f})$ such that for every $t \geq t_0$ and $\varepsilon \in (0, 1)$ satisfying

$$\varepsilon > K_0 e^{-\eta t}$$
 with η as in (1.7) (5.10)

we have that

$$\mathcal{L}_d(\Delta_{\theta}(t,\varepsilon,B,\rho)) \ge \frac{1}{2}\mathcal{L}_d(B),$$
 (5.11)

where

$$\rho = \left(\frac{C_0}{\varepsilon^m e^{(d+1)t}}\right)^{-\frac{1}{d}}.$$
(5.12)

Proof. Suppose from Proposition 5.1 we get $V \subset \mathbf{U}$ centered at \mathbf{x}_0 , and $K_0^1 > 0$, $\delta > 0$ for $\kappa = \frac{1}{3}$. Let us take $K_0 > K_0^1$ and K_0 satisfies (5.3). Using Proposition 5.1, for every ball $B \subset V$, there exists $t_0 > 0$ such that for all $t \geq t_0$,

$$\mathcal{L}_d(B \setminus \mathfrak{G}(\delta, t, \varepsilon)) \le \frac{1}{3} \mathcal{L}_d(B). \tag{5.13}$$

Now applying Lemma 5.1, for $\tilde{\varepsilon}$ and \tilde{t} as in (5.4), since $e^{-\eta t} \geq e^{-\frac{t(d+2)}{2m+d}}$ we get that

$$\mathcal{L}_d(\Delta(\tilde{t}, \tilde{\varepsilon}, B, \rho)) \ge \mathcal{L}_d(\mathfrak{G}(\delta, t, \varepsilon) \cap B^{\rho}) \ge \mathcal{L}_d(B \cap \mathfrak{G}(\delta, t, \varepsilon)) - \mathcal{L}_d(B \setminus B^{\rho}). \tag{5.14}$$

By (5.10) and (5.12), we have that $\rho \to 0$ as $t \to \infty$. Hence $\mathcal{L}_d(B \setminus B^{\rho}) \to 0$ as $t \to \infty$ and hence for all sufficiently large t we have that $\mathcal{L}_d(B \setminus B^{\rho}) \leq \frac{1}{6}\mathcal{L}_d(B)$. Combining this with (5.13) and (5.14), we get that

$$\mathcal{L}_d(\Delta(\tilde{t}, \tilde{\epsilon}, B, \rho)) \ge \frac{1}{2}\mathcal{L}_d(B)$$
.

Since $\varepsilon \simeq \tilde{\varepsilon}$ and $e^t \simeq e^{\tilde{t}}$, the required result follows.

As a corollary of the above theorem, we get the following:

Corollary 5.1. Let $f, \theta, U, B, x_0, \varepsilon, t, C_0$ be the same as in Theorem 5.1. Then

$$\#\mathcal{N}_{\theta}(t,\varepsilon,B) \ge \frac{\varepsilon^m e^{(d+1)t}}{2V_d C_0} \mathcal{L}_d(B), \tag{5.15}$$

where $V_d := \mathcal{L}_d(B(0,1))$.

5.1. **Proof of Theorem 1.3.** Combining Corollary 5.1 and (4.7) the proof follows.

5.2. **Proof of Theorem 2.1.** The proof of Theorem 2.1 goes exactly the same as [5, pp. 196-199] and using Theorem 5.1. We should remark that the range of s is determined by the range of ε in Theorem 5.1.

5.3. Proof of divergence part in Theorem 1.2. Let us take

$$\psi_1(q) := \max \left\{ \psi(q), q^{-(\eta - \zeta)} \right\},\,$$

where $\eta - \frac{1}{n} > \zeta > 0$ and η is as in (1.7). Then $\psi_1(q)^{1/\eta}q \to \infty$ and $\sum \psi(q)^n = \infty$ implies that $\sum \psi_1(q)^n = \infty$. Since $\eta - \zeta > \frac{1}{n}$, by [17, Corollary A], $\mathcal{L}_d\left(\mathbf{f}^{-1}(\mathbb{S}_n^{\theta}(q \to q^{-\eta}))\right) = 0$. Also note that $\mathbf{f}^{-1}(\mathbb{S}_n^{\theta}(\psi_1)) = \mathbf{f}^{-1}(\mathbb{S}_n^{\theta}(\psi)) \cap \mathbf{f}^{-1}(\mathbb{S}_n^{\theta}(q \to q^{-\eta}))$. Now applying Theorem 2.1 for s = d, and ψ_1 yields the divergence part of Theorem 1.2.

6. Upper bounds

6.1. **Preliminaries.** For $\mathbf{x}_0 \in \mathbf{U}$, $1 > \varepsilon > 0$, t > 0, let us denote $\Delta_t(\mathbf{x}_0) := B\left[\mathbf{x}_0, \left(\frac{\varepsilon}{e^t}\right)^{\frac{1}{2}}\right] \subset \mathbb{R}^d$.

Also, let

$$g_{\varepsilon,t} := \phi \operatorname{diag}(\underbrace{\varepsilon^{-1}, \dots, \varepsilon^{-1}}_{n}, e^{-t}),$$

$$\phi := \left(\varepsilon^n e^t\right)^{1/n+1}, h := \frac{dt}{2(n+1)},\tag{6.1}$$

and

$$b_{t} := \operatorname{diag}\{\underbrace{e^{\frac{dt}{2(n+1)}}, \dots, e^{\frac{dt}{2(n+1)}}}_{m}, \underbrace{e^{\frac{-(m+1)t}{2(n+1)}}, \dots, e^{\frac{-(m+1)t}{2(n+1)}}}_{d}, e^{\frac{dt}{2(n+1)}}\} \in \operatorname{SL}_{n+1}(\mathbb{R}).$$
(6.2)

Let us define $d_{\varepsilon} = \operatorname{diag}(\underbrace{1,\ldots,1}_{m},\underbrace{\varepsilon^{1/2},\ldots,\varepsilon^{1/2}}_{d},1)$ for $\varepsilon > 0$. The following equation follows from matrix multiplication.

$$g_{\varepsilon,t}z(\mathbf{x})g_{\varepsilon,t}^{-1} = z(\mathbf{x}). \tag{6.3}$$

Now we make modified definitions of generic and special parts from [13]. Let $d_{\varepsilon}, g_{\varepsilon,t}, b_t, u_1(\mathbf{x})$ be as above. We define 'raw nongeneric' set to be

$$\mathfrak{M}_0^{new}(\varepsilon,t) := \{ \mathbf{x} \in \mathbf{U} \mid \delta_{n+1}(d_{\varepsilon}b_t g_{\varepsilon,t} u_1(\mathbf{x}) \mathbb{Z}^{n+1}) > \phi e^h \}.$$

Note the above definition uses a new matrix d_{ε} that was not used in [13]; see [13, Equation 4.18].

Then we define the special part to be a thickening of the above raw set, namely,

$$\mathfrak{M}^{new}(\varepsilon,t) := \bigcup_{\mathbf{x} \in \mathfrak{M}_0^{new}(\varepsilon,t)} B(\mathbf{x}, \varepsilon^{1/2} e^{-t/2}) \cap \mathbf{U}.$$
(6.4)

Note that the above thickening radius is also different than the same in [13, Equation 4.19]. Lastly, let us recall that the generic part is the complement of the special part, i.e.,

$$\mathbf{U} \setminus \mathfrak{M}^{new}(\varepsilon, t)$$
.

Similar to [13, Lemma 1.4] we have,

Lemma 6.1. If $\mathbf{f}(\mathbf{x}) \in \mathbb{S}_n^{\theta}(\psi)$ then there are infinitely many $t \in \mathbb{N}$ such that

$$\left\| \mathbf{f}(\mathbf{x}) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| < \frac{\psi(e^{t-1})}{e^{t-1}}$$
(6.5)

for some $(q, \mathbf{p}) \in \mathbb{Z}^{n+1}$ with $e^{t-1} \le q < e^t$.

The proof of the above lemma uses the fact that ψ is monotonically decreasing. Our next lemma follows from simple matrix multiplication.

Lemma 6.2. Let \mathbf{f} , \mathbf{U} be as before. Suppose for some $\mathbf{x} \in \mathbf{U}$, and $\varepsilon > 0$, and $t \in \mathbb{N}$, there exists $(q, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^n$ such that,

$$\left| qf_{j}(\mathbf{x}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}) (qx_{i} - p_{i}) - p_{d+j} \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|qx_{i} - p_{i}| \ll_{\mathbf{f},n} \varepsilon^{\frac{1}{2}} e^{t/2}, \ 1 \leq i \leq d,$$

$$|q| \ll_{\mathbf{f},n} e^{t}.$$

$$(6.6)$$

Then

$$g_{\varepsilon,t}u_1(\mathbf{x})(-\mathbf{p}\sigma_n,q) \in [c\phi]^m \times [c\phi((\varepsilon e^{-t})^{-\frac{1}{2}})]^d \times [c\phi]. \tag{6.7}$$

where ϕ, ε, t are related as (6.1) and c depends on **f** and n.

The proof is obvious from the definition of $u_1(\mathbf{x})$ in (4.5). We assume \mathbf{f} , \mathbf{U} , n, m, d, $\boldsymbol{\theta}$ be as in Theorem 1.2. By (4.3) and (4.4) the following lemma follows easily.

Lemma 6.3. Suppose for some $\mathbf{x} \in \mathbf{U}$, t > 0, $\varepsilon > 0$,

$$\|q\mathbf{f}(\mathbf{x}) - \mathbf{p} + \boldsymbol{\theta}\| < \varepsilon,$$
 (6.8)

for some $(q, \mathbf{p}) \in \mathbb{Z}^{n+1}$, with $0 < q \le e^t$. Then,

$$\left| qf_{j}(\mathbf{x}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}) (qx_{i} - p_{i} + \theta_{i}) - p_{d+j} + \theta_{d+j} \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|qx_{i} - p_{i} + \theta_{i}| \leq \varepsilon, \ 1 \leq i \leq d,$$

$$|q| \leq e^{t}.$$

$$(6.9)$$

Lemma 6.4. Let $1 \ge \varepsilon > 0, t > 0$. Suppose $\mathbf{x} \in \Delta_t(\mathbf{x}_0) := B\left[\mathbf{x}_0, \left(\frac{\varepsilon}{e^t}\right)^{\frac{1}{2}}\right] \subset \mathbb{R}^d$, and $(\mathbf{p}, q) \in \mathbb{Z}^{n+1}$, with $1 \le q \le e^t$ such that $\|\mathbf{f}(\mathbf{x}) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q}\| < \frac{\varepsilon}{e^t}$. Then

$$\left| qf_{j}(\mathbf{x}_{0}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}_{0}) (qx_{0,i} - p_{i} + \theta_{i}) - p_{d+j} + \theta_{d+j} \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|qx_{0,i} - p_{i} + \theta_{i}| \leq 2\varepsilon^{1/2} e^{t/2}, \ 1 \leq i \leq d,$$

$$|q| \leq e^{t}.$$
(6.10)

Proof. Using the hypothesis of this lemma, by Lemma 6.3, (6.9) holds. One can write $\mathbf{x} = \mathbf{x}_0 + (\varepsilon e^{-t})^{1/2} \mathbf{x}'$, where $\|\mathbf{x}'\| \le 1$. For $1 \le i \le d$,

$$|qx_{0,i} - p_i + \theta_i| < |qx_i - p_i + \theta_i| + |q||x_i - x_{0,i}| \le \varepsilon + e^t(\varepsilon e^{-t})^{\frac{1}{2}} \le 2\varepsilon^{1/2} e^{t/2}.$$

Also, by (6.9), for $1 \le j \le m$,

$$\left| qf_{j}(\mathbf{x}_{0}) - \sum_{i=1}^{d} \partial_{i}f_{j}(\mathbf{x}_{0})(qx_{0,i} - p_{i} + \theta_{i}) - p_{d+j} + \theta_{d+j} \right|$$

$$\ll_{\mathbf{f},n} \varepsilon + \sum_{i=1}^{d} |\partial_{i}f_{j}(\mathbf{x}_{0}) - \partial_{i}f_{j}(\mathbf{x})||qx_{i} - p_{i} + \theta_{i}| + |q| \left(\varepsilon e^{-t}\right)$$

$$\ll_{\mathbf{f},n} \varepsilon + (\varepsilon e^{-t})^{\frac{1}{2}}\varepsilon + \varepsilon$$

$$\ll_{\mathbf{f},n} \varepsilon.$$

To conclude some of the inequalities above, we use Taylor's expansion, $f_j(\mathbf{x}_0) = f_j(\mathbf{x}) + \sum_{i=1}^d \partial_i f_j(\mathbf{x}) (x_{0,i} - x_i) + \sum_{\beta \in \mathbb{N}^2, |\beta| = 2} \hat{f}_{j,\beta}(\mathbf{x}) (\mathbf{x}_0 - \mathbf{x})_{\beta}$, where for $\mathbf{y} \in \mathbb{R}^d$, $\mathbf{y}_{\beta} = y_k y_l$ if $\beta = (k, l) \in \mathbb{N}^2$, $|\beta| := k + l = 2$. Similarly, we also use that $\partial_i f_j(\mathbf{x}_0) = \partial_i f_j(\mathbf{x}) + \sum_{k=1}^d \hat{f}_{i,j,k}(\mathbf{x}) (x_{0,k} - x_k)$ and for every $\mathbf{x} \in \mathbf{U}$,

$$|\hat{f}_{i,j,k}(\mathbf{x})| \le \max_{\mathbf{y} \in \mathbf{U}} |\partial_{k,i}^2 f_j(\mathbf{y})| \text{ and } |\hat{f}_{i,\beta}(\mathbf{x})| \le \max_{\mathbf{y} \in \mathbf{U}} |\partial_{\beta} f_i(\mathbf{y})|.$$
 (6.11)

6.2. **Generic part.** Let us recall the definition of $\mathfrak{M}^{new}(\varepsilon,t)$ from §6.1, and we continue with the same notations as in the previous sections.

Lemma 6.5. Let B be a ball in **U**. For all large enough t > 0, $0 < \varepsilon < 1$ for every $\mathbf{x}_0 \in B \cap (\mathbf{U} \setminus \mathfrak{M}^{new}(\varepsilon, t))$, we have

$$N_{\boldsymbol{\theta}}(\Delta_t(\mathbf{x}_0) \cap B; \varepsilon, t) \ll_{n, \mathbf{f}} \varepsilon^n e^t (\varepsilon e^{-t})^{\frac{-d}{2}}.$$

Proof. Without loss of generality we assume $N_{\theta}(\Delta_t(\mathbf{x}_0) \cap B; \varepsilon, t) > 1$. Then take two distinct (\mathbf{p}, q) and (\mathbf{p}', q') in $\mathbb{Z}^n \times \mathbb{N}$ such that $1 \leq q, q' \leq e^t$ and for some $\mathbf{x}, \mathbf{x}' \in \Delta_t(\mathbf{x}_0) \cap B$,

$$\left\| \mathbf{f}(x) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| \le \frac{\varepsilon}{e^t}$$
, and $\left\| \mathbf{f}(x') - \frac{\mathbf{p}' + \boldsymbol{\theta}}{q'} \right\| \le \frac{\varepsilon}{e^t}$.

By Lemma 6.4, we have

$$\left| qf_{j}(\mathbf{x}_{0}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}_{0}) (qx_{0,i} - p_{i} + \theta_{i}) - p_{d+j} + \theta_{d+j} \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|qx_{0,i} - p_{i} + \theta_{i}| \leq 2\varepsilon^{1/2} e^{t/2}, \ 1 \leq i \leq d,$$

$$|q| \leq e^{t},$$
(6.12)

and

$$\left| q' f_{j}(\mathbf{x}_{0}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}_{0}) (q' x_{0,i} - p'_{i} + \theta_{i}) - p'_{d+j} + \theta_{d+j} \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|q' x_{0,i} - p'_{i} + \theta_{i}| \leq 2\varepsilon^{1/2} e^{t/2}, \ 1 \leq i \leq d,$$

$$|q'| \leq e^{t}.$$
(6.13)

Now by subtracting the previous two equations, we get

$$\left| q'' f_{j}(\mathbf{x}_{0}) - \sum_{i=1}^{d} \partial_{i} f_{j}(\mathbf{x}_{0}) (q'' x_{0,i} - p_{i}'') - p_{d+j}'' \right| \ll_{\mathbf{f},n} \varepsilon, \ 1 \leq j \leq m,$$

$$|q'' x_{0,i} - p_{i}''| \ll 4\varepsilon^{1/2} e^{t/2}, \ 1 \leq i \leq d,$$

$$|q''| \leq 2e^{t},$$

$$(6.14)$$

where q'' = q - q' and $\mathbf{p}'' = \mathbf{p} - \mathbf{p}'$.

Now by Lemma 6.2 we have,

$$g_{\varepsilon,t}u_1(\mathbf{x}_0)(-\mathbf{p}''\sigma_n, q'') \in [c\phi]^m \times [c\phi((\varepsilon e^{-t})^{-\frac{1}{2}})]^d \times [c\phi], \tag{6.15}$$

where c depends on **f** and n. From (6.15), and by the definition of b_t as in §6.1, we have

$$b_t g_{\varepsilon,t} u_1(\mathbf{x}_0)(-\mathbf{p}''\sigma_n, q'') \in \underbrace{[c_1 \phi e^h]^m \times [c_1 \phi e^h \varepsilon^{-1/2}]^d \times [c_1 \phi e^h]}_{\Omega} \cap b_t g_{\varepsilon,t} u_1(\mathbf{x}_0) \mathbb{Z}^{n+1},$$

where $c_1 > 0$ depends on n, \mathbf{f} .

Now applying d_{ε} we get

$$d_{\varepsilon}b_{t}g_{\varepsilon,t}u_{1}(\mathbf{x}_{0})(-\mathbf{p}''\sigma_{n},q'') \in \underbrace{[c_{1}\phi e^{h}]^{m} \times [c_{1}\phi e^{h}]^{d} \times [c_{1}\phi e^{h}]}_{\Omega^{\star}} \cap d_{\varepsilon}b_{t}g_{\varepsilon,t}u_{1}(\mathbf{x}_{0})\mathbb{Z}^{n+1}.$$

Since $\mathbf{x}_0 \in \mathbf{U} \setminus \mathfrak{M}^{new}(\varepsilon, t)$, we have

$$\delta_{n+1}(d_{\varepsilon}b_tg_{\varepsilon,t}u_1(\mathbf{x}_0)\mathbb{Z}^{n+1}) \le \phi e^h.$$

Hence $c_2\Omega^*$ for some large $c_2 > 0$ will contain the fundamental domain of the lattice $d_{\varepsilon}b_tg_{\varepsilon,t}u_1(\mathbf{x}_0)\mathbb{Z}^{n+1}$. Therefore,

$$\#c_2\Omega^{\star}\cap d_{\varepsilon}b_tg_{\varepsilon,t}u_1(\mathbf{x}_0)\mathbb{Z}^{n+1}\ll_{n,\mathbf{f}}\frac{\mathcal{L}_{n+1}(\Omega^{\star})}{\det d_{\varepsilon}}=\frac{\phi^{n+1}e^{h(n+1)}}{\varepsilon^{d/2}}=\phi^{n+1}\varepsilon^{-d/2}e^{h(n+1)}=\varepsilon^ne^t(\varepsilon e^{-t})^{-d/2}.$$

Thus the number of such (\mathbf{p}'', q'') is $\ll \varepsilon^n e^t (\varepsilon e^{-t})^{-d/2}$, and note that we should get at least $N_{\theta}(\Delta_t(\mathbf{x}_0) \cap B; \varepsilon, t) - 1$ many (distinct) of them, which therefore concludes the lemma.

The proof of the following lemma is exactly the same as in [13, Lemma 5.4].

Lemma 6.6. For sufficiently large t > 0,

$$N_{\boldsymbol{\theta}}(B \setminus \mathfrak{M}^{new}(\varepsilon, t); \varepsilon, t) \leq 2(\varepsilon e^{-t})^{-d/2} \mathcal{L}_d(B) \max_{\mathbf{x}_0 \in B \cap (\mathbf{U} \setminus \mathfrak{M}^{new}(\varepsilon, t))} N_{\boldsymbol{\theta}}(\Delta_t(\mathbf{x}_0); \varepsilon, t).$$

As a consequence of Lemma 6.5, and Lemma 6.6, we have the following proposition.

Proposition 6.1. Let \mathbf{f} , \mathbf{U} be as before. Then for any $0 < \varepsilon < 1$, for any ball $B \subset \mathbf{U}$, and all sufficiently large t > 0, we have

$$N_{\theta}(B \setminus \mathfrak{M}^{new}(\varepsilon, t); \varepsilon, t) \ll_{n, \mathbf{f}} \varepsilon^m e^{(d+1)t} \mathcal{L}_d(B).$$
 (6.16)

6.3. Nongeneric part. First, for some matrices introduced in previous sections, we compute their dual matrices. For $\varepsilon > 0, t > 0$,

$$g_{\varepsilon,t}^{\star} = \phi^{-1} \operatorname{diag}(e^t, \underbrace{\varepsilon, \dots, \varepsilon}_{n}),$$
 (6.17)

$$u_1^{\star}(\mathbf{x}) = \begin{bmatrix} 1 & -\mathbf{x} & -f(\mathbf{x}) \\ 0 & \mathbf{I}_d & J(\mathbf{x}) \\ 0 & 0 & \mathbf{I}_m \end{bmatrix},$$

$$d_{\varepsilon}^{\star} = \operatorname{diag}(1, \underbrace{\varepsilon^{-1/2}, \dots, \varepsilon^{-1/2}}_{I}, \underbrace{1, \dots, 1}_{I}),$$

$$(6.18)$$

and

$$b_t^{\star} = \operatorname{diag}(e^{-h}, \underbrace{e^{\frac{(m+1)t}{2(n+1)}}, \dots, e^{\frac{(m+1)t}{2(n+1)}}}_{d}, \underbrace{e^{-h}, \dots, e^{-h}}_{m}).$$

The next lemma shows that the new special part is inside the set defined in (3.1).

Lemma 6.7. For $\varepsilon > 0, t > 0$ and **f** as before,

$$\mathfrak{M}^{new}(\varepsilon,t) \subset \mathfrak{S}_{\mathbf{f}}(ce^{-t},c\varepsilon^{-1/2}e^{-t/2},c\varepsilon^{-1})$$

where c depends on n, \mathbf{f} , and the above sets are defined in (3.1) and (6.4).

Proof. First note that, if $\mathbf{x}_0 \in \mathfrak{M}_0^{new}(\varepsilon, t)$ then

$$\delta_{n+1}(d_{\varepsilon}b_tg_{\varepsilon,t}u_1(\mathbf{x}_0)\mathbb{Z}^{n+1}) > \phi e^h$$
, where $h = \frac{dt}{2(n+1)}$.

Then by Lemma 4.1, and the fact that $(g_1g_2)^* = g_1^*g_2^*$,

$$\delta_1(d_{\varepsilon}^{\star}b_t^{\star}g_{\varepsilon}^{\star}u_1^{\star}(\mathbf{x}_0)\mathbb{Z}^{n+1}) \ll_n \phi^{-1}e^{-h}.$$

This implies there exists $(a_0, \mathbf{a}) \in \mathbb{Z}^{n+1} \setminus \mathbf{0}$ such that

$$|a_0 + \mathbf{f}(\mathbf{x}_0)\mathbf{a}^{\top}| \ll_n e^{-t}$$

$$\|\nabla \mathbf{f}(\mathbf{x}_0)\mathbf{a}^{\top}\| \ll_n \varepsilon^{-1/2} e^{-t/2}$$

$$0 < \|\mathbf{a}_m\| \ll_n \varepsilon^{-1},$$
(6.19)

where $\mathbf{a}_m = (a_{d+1}, \dots, a_n)$. Note that using the second and third inequalities in (6.19) and by (4.3) we get $\|\mathbf{a}\| \ll_{n,\mathbf{f}} \varepsilon^{-1}$.

Now for any $\mathbf{x} \in \mathfrak{M}^{new}(\varepsilon, t)$, there exists $\mathbf{x}_0 \in \mathfrak{M}_0^{new}(\varepsilon, t)$ such that $\|\mathbf{x} - \mathbf{x}_0\| \le \varepsilon^{1/2} e^{-t/2}$, we have by Taylor's expansion,

$$|a_0 + \mathbf{f}(\mathbf{x})\mathbf{a}^{\top}| \ll_{n,\mathbf{f}} e^{-t} + \varepsilon^{-1/2} e^{-t/2} \varepsilon^{1/2} e^{-t/2} + \varepsilon^{-1} (\varepsilon^{1/2} e^{-t/2})^2 \ll_{n,\mathbf{f}} e^{-t}.$$

Similarly, using Taylor's expansion we get,

$$\|\nabla \mathbf{f}(\mathbf{x})\mathbf{a}^{\top}\| \ll_{n,\mathbf{f}} \varepsilon^{-1/2} e^{-t/2} + \varepsilon^{-1} \varepsilon^{1/2} e^{-t/2} \ll_{n,\mathbf{f}} \varepsilon^{-1/2} e^{-t/2}.$$

Thus the lemma follows.

Using Lemma 6.7 and Theorem 3.2 we have the following proposition.

Proposition 6.2. Let **U** be an open set in \mathbb{R}^d , $\mathbf{f}: \mathbf{U} \to \mathbb{R}^n$ be a map satisfying (4.3). Let **f** be l-nondegenerate at $\mathbf{x}_0 \in \mathbf{U}$. Then there exists a ball B_0 centered at \mathbf{x}_0 , constants K_0, t_0 depending on n, \mathbf{f}, B_0 such that for $t \geq t_0$,

$$\mathcal{L}_d(\mathfrak{M}^{new}(\varepsilon,t)\cap B_0) \ll K_0(e^{-t}\varepsilon^{-1/2}e^{-t/2}\varepsilon^{-(n-1)})^{\alpha}\mathcal{L}_d(B_0) = (e^{3t/2}\varepsilon^{n-1/2})^{-\alpha}\mathcal{L}_d(B_0),$$

where $\alpha = \frac{1}{d(2l-1)(n+1)}$.

6.4. Completing the proof convergence. Combining Proposition 6.1 and Proposition 6.2, we have the key proposition as follows.

Proposition 6.3. Let $\mathbf{U} \subset \mathbb{R}^d$ be an open set and $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ be a map satisfying 4.3. Then for any $0 < \varepsilon < 1$ and every t > 0 there exists $\mathfrak{M}^{new}(\varepsilon, t) \subset \mathbf{U}$, which can be written as a union of balls of radius $\varepsilon^{1/2}e^{-t/2}$ of intersection multiplicity N_d . If \mathbf{f} is nondegenerate at \mathbf{x}_0 then there is a ball $B_0 \subset \mathbf{U}$ centered at \mathbf{x}_0 and constants $K_0, t_0 > 0$, which depend on n, \mathbf{f} and B_0 only, such that for $t \geq t_0$,

$$\mathcal{L}_d(\mathfrak{M}^{new}(\varepsilon,t)\cap B_0) \le K_0 \left(\varepsilon^{n-1/2} e^{3t/2}\right)^{-\alpha}, \alpha = \frac{1}{d(2l-1)(n+1)},\tag{6.20}$$

and for every ball $B \subset \mathbf{U}$ and sufficiently large t, we have

$$N_{\theta}(B \setminus \mathfrak{M}^{new}(\varepsilon, t); \varepsilon, t) \ll_{n, \mathbf{f}} \varepsilon^m e^{(d+1)t} \mathcal{L}_d(B).$$
 (6.21)

Now, we can complete the proof of Theorem 1.4, Theorem 1.2 and 2.2 using the above proposition.

6.5. **Proof of Theorem 1.4.** Let α , \mathbf{f} , B_0 , and \mathbf{x}_0 be as in Proposition 6.3. For $1 > \varepsilon > 0$ and all large t > 0,

$$N_{\boldsymbol{\theta}}(B_0; \varepsilon, t) \ll_{n, \mathbf{f}, B_0} \varepsilon^m e^{(d+1)t} + e^{(d+1)t} \left(\varepsilon^{n-1/2} e^{3t/2}\right)^{-\alpha}. \tag{6.22}$$

To get the second term above, we cover $\mathfrak{M}^{new}(\varepsilon,t) \cap B_0$ by $\ll \frac{\mathcal{L}_d(\mathfrak{M}(\varepsilon,t) \cap B_0)}{\varepsilon^{d/2}e^{-td/2}}$ balls of radius $\varepsilon^{1/2}e^{-t/2}$.

6.6. **Proof of convergence part in Theorem 1.2.** Since it is enough to prove this theorem for approximating function being $\max\{\psi(q), q^{-\frac{5}{4n}}\}$, without loss of generality, we assume that for all $q \in \mathbb{N}$,

$$\psi(q)^n > q^{-\frac{5}{4}}. (6.23)$$

For any $\Delta \subset \mathbb{R}^d$, any t > 0 and $0 < \varepsilon < 1$, we define

$$\mathcal{R}_{\boldsymbol{\theta}}(\Delta; \varepsilon, t) := \left\{ (\mathbf{p}, q) \in \mathbb{Z}^{n+1} \mid e^{t-1} \leq q < e^{t}, \inf_{\mathbf{x} \in \Delta} \left\| \mathbf{f}(\mathbf{x}) - \frac{\mathbf{p} + \boldsymbol{\theta}}{q} \right\| < \frac{\varepsilon}{e^{t}} \right\}.$$

By definition as in $\S1.2$,

$$N_{\boldsymbol{\theta}}^{\star}(\Delta; \varepsilon, t) = \# \mathcal{R}_{\boldsymbol{\theta}}(\Delta; \varepsilon, t).$$

Now,

$$\left\{\mathbf{x} \in B_0 \mid \mathbf{f}(\mathbf{x}) \in \mathcal{S}_n^{\boldsymbol{\theta}}(\psi)\right\} \subset \bigcup_{t \geq T} \underbrace{\left(\mathfrak{M}^{new}(e\psi(e^{t-1}), t) \cap B_0\right)}_{A_t} \bigcup_{t \geq T} \underbrace{\left(\mathbf{p}, q\right) \in \mathcal{R}_{\boldsymbol{\theta}}(B_0 \setminus \mathfrak{M}^{new}(e\psi(e^{t-1}), t); e\psi(e^{t-1}), t\right)}_{B^{\boldsymbol{\theta}}} \left\{\mathbf{x} \in B_0 \mid \left\|\mathbf{x} - \frac{\pi(\mathbf{p} + \boldsymbol{\theta})}{q}\right\| \leq \frac{\psi(e^{t-1})}{e^{t-1}}\right\}.$$

Here $\pi: \mathbb{R}^n \to \mathbb{R}^d$ is the projection of the first d many coordinates. By (6.20), we have

$$\mathcal{L}_d(A_t) \ll \left(\psi(e^t)^{n-1/2}e^{3t/2}\right)^{-\alpha} \stackrel{\text{(6.23)}}{<} \left(e^{-(1-1/2n)5t/4}e^{3t/2}\right)^{-\alpha} = \left(e^{(1+5/2n)t/4}\right)^{-\alpha},$$

where $\alpha > 0$ is as in Proposition 6.3. Next, observe that

$$\mathcal{L}_d(B_t^{\theta}) \ll \psi(e^{t-1})^m e^{(d+1)t} \left(\frac{\psi(e^{t-1})}{e^{t-1}}\right)^d \ll \psi(e^{t-1})^n e^{t-1}.$$

Since $\sum \psi(q)^n < \infty$, we have $\sum_{t \geq 1} \psi(e^t)^n e^t < \infty$. The proof is now complete by using the convergence Borel Cantelli lemma.

6.7. **Proof of Theorem 2.2.** By (2.4), it is enough to prove

$$\mathcal{H}^s(\{\mathbf{x} \in B_0 \mid \mathbf{f}(\mathbf{x}) \in \mathcal{S}_n^{\boldsymbol{\theta}}(\psi)\}) = 0,$$

where B_0 is a very small neighborhood of \mathbf{x}_0 , and \mathbf{f} is l-nondegenerate at \mathbf{x}_0 . Let us fix B_0 as in Proposition 6.3, for $\varepsilon = e\psi(e^{t-1})$, for all large t > 0, $\mathfrak{M}^{new}(\varepsilon, t)$ can be covered as union of balls of radius $\varepsilon^{1/2}e^{-t/2}$ such that number of such balls will be $\ll \mathcal{L}_d(\mathfrak{M}(\varepsilon, t)/(\varepsilon^{d/2}e^{-dt/2}))$. Hence by Proposition 6.3 we have that,

$$\mathcal{H}^s(A_t) \ll \left(\varepsilon^{1/2}e^{-t/2}\right)^{s-d} \mathcal{L}_d(\mathfrak{M}^{new}(\varepsilon,t)).$$

By Proposition 6.3, plugging $\varepsilon = e\psi(e^{t-1})$ we get that

$$\sum_{t > T} \mathcal{H}^{s}(A_{t}) \ll \sum_{t > T} \left(\frac{\psi(e^{t})}{e^{t}} \right)^{(s-d)/2} \left(\psi(e^{t})^{n-1/2} e^{3t/2} \right)^{-\alpha} \to 0,$$

by (2.6), as $T \to \infty$.

Also,

$$\sum_{t \geq T} \mathcal{H}^s(B_t^{\theta}) \ll \sum_{t \geq T} \varepsilon^m e^{(d+1)t} \left(\varepsilon e^{-t} \right)^s, \text{ where } \varepsilon = e\psi(e^{t-1})$$

and the above sum tends to zero by (2.5) as $T \to \infty$.

6.8. **Proof of Corollary 2.1.** The upper bound follows directly from Theorem 2.2 on applying $\psi(q) = q^{-\tau}$. The proof of this goes exactly as in the proof of [13, Corollary 2.9] and [32, Corollary 1.13] upon using Theorem 2.2. For the lower bound we check that τ as in (2.10) also lies in the range $(1/n, \eta)$ as in 2.1. Let us take $\frac{(n+1)\alpha}{3\alpha+1} < \delta < \frac{n+1}{n+2}$. Then it is easy to check that $\frac{1}{n-\delta} \in (1/n, \eta) \cap (1/n, \frac{3\alpha+1}{\alpha(2n-1)+n})$.

Acknowledgments. This work was supported by the EPSRC grant EP/Y016769/1. SD is grateful to Ralf Spatzier for their support, especially when it was needed the most. SD also thanks Subhajit Jana for numerous discussions around this project and for his constant encouragement.

REFERENCES

- [1] D. Allen and F. A. Ramírez. Independence inheritance and Diophantine approximation for systems of linear forms. *Int. Math. Res. Not. IMRN*, (2):1760–1794, 2023.
- [2] D. Badziahin, V. Beresnevich, and S. Velani. Inhomogeneous theory of dual Diophantine approximation on manifolds. *Adv. Math.*, 232:1–35, 2013.
- [3] V. Beresnevich. On approximation of real numbers by real algebraic numbers. *Acta Arith.*, 90(2):97–112, 1999.
- [4] V. Beresnevich. A Groshev type theorem for convergence on manifolds. *Acta Math. Hungar.*, 94(1-2):99–130, 2002.
- [5] V. Beresnevich. Rational points near manifolds and metric Diophantine approximation. Ann. of Math. (2), 175(1):187–235, 2012.
- [6] V. Beresnevich, V. Bernik, M. Dodson, and S. Velani. Classical metric Diophantine approximation revisited. In *Analytic number theory*, pages 38–61. Cambridge Univ. Press, Cambridge, 2009.
- [7] V. Beresnevich, D. Dickinson, and S. Velani. Measure theoretic laws for lim sup sets. *Mem. Amer. Math. Soc.*, 179(846):x+91, 2006.
- [8] V. Beresnevich, D. Dickinson, and S. Velani. Diophantine approximation on planar curves and the distribution of rational points. *Ann. of Math.* (2), 166(2):367–426, 2007. With an Appendix II by R. C. Vaughan.
- [9] V. Beresnevich, A. Ganguly, A. Ghosh, and S. Velani. Inhomogeneous dual Diophantine approximation on affine subspaces. *Int. Math. Res. Not. IMRN*, (12):3582–3613, 2020.
- [10] V. Beresnevich, R. C. Vaughan, S. Velani, and E. Zorin. Diophantine approximation on manifolds and the distribution of rational points: contributions to the convergence theory. *Int. Math. Res. Not. IMRN*, (10):2885–2908, 2017.

- [11] V. Beresnevich, R. C. Vaughan, S. Velani, and E. Zorin. Diophantine approximation on curves and the distribution of rational points: contributions to the divergence theory. *Adv. Math.*, 388:Paper No. 107861, 33, 2021.
- [12] V. Beresnevich and S. Velani. An inhomogeneous transference principle and Diophantine approximation. *Proc. Lond. Math. Soc.* (3), 101(3):821–851, 2010.
- [13] V. Beresnevich and L. Yang. Khintchine's theorem and diophantine approximation on manifolds. *Acta Math.*, 231:1–30, 2023.
- [14] V. Beresnevich and E. Zorin. Explicit bounds for rational points near planar curves and metric Diophantine approximation. Adv. Math., 225(6):3064–3087, 2010.
- [15] V. V. Beresnevich, V. I. Bernik, D. Y. Kleinbock, and G. A. Margulis. Metric Diophantine approximation: the Khintchine-Groshev theorem for nondegenerate manifolds. volume 2, pages 203–225. 2002. Dedicated to Yuri I. Manin on the occasion of his 65th birthday.
- [16] V. V. Beresnevich, R. C. Vaughan, and S. L. Velani. Inhomogeneous Diophantine approximation on planar curves. *Math. Ann.*, 349(4):929–942, 2011.
- [17] V. V. Beresnevich and S. L. Velani. Simultaneous inhomogeneous Diophantine approximations on manifolds. Fundam. Prikl. Mat., 16(5):3–17, 2010.
- [18] V. Bernik, D. Kleinbock, and G. A. Margulis. Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions. *Internat. Math. Res. Notices*, (9):453–486, 2001
- [19] V. I. Bernik and M. M. Dodson. Metric Diophantine approximation on manifolds, volume 137 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1999.
- [20] Y. Bugeaud. An inhomogeneous Jarník theorem. J. Anal. Math., 92:327–349, 2004.
- [21] Y. Bugeaud and M. Laurent. Exponents of Diophantine approximation and Sturmian continued fractions. *Ann. Inst. Fourier (Grenoble)*, 55(3):773–804, 2005.
- [22] Y. Bugeaud and M. Laurent. On exponents of homogeneous and inhomogeneous Diophantine approximation. *Mosc. Math. J.*, 5(4):747–766, 972, 2005.
- [23] A. Ghosh. A Khintchine-type theorem for hyperplanes. J. London Math. Soc. (2), 72(2):293–304, 2005.
- [24] A. Ghosh. Diophantine exponents and the Khintchine Groshev theorem. *Monatsh. Math.*, 163(3):281–299, 2011.
- [25] P. M. Gruber. Convex and discrete geometry, volume 336 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin, 2007.
- [26] J.-J. Huang. Rational points near planar curves and Diophantine approximation. Adv. Math., 274:490–515, 2015.
- [27] J.-J. Huang. The density of rational points near hypersurfaces. *Duke Math. J.*, 169(11):2045–2077, 2020.
- [28] J.-J. Huang. Extremal affine subspaces and Khintchine-Jarník type theorems. *Geom. Funct.* Anal., 34(1):113–163, 2024.
- [29] D. Kleinbock. Extremal subspaces and their submanifolds. Geometric and Functional Analysis (GAFA), 13(2):437–466, 2003.
- [30] D. Kleinbock. An extension of quantitative nondivergence and applications to Diophantine exponents. *Trans. Amer. Math. Soc.*, 360(12):6497–6523, 2008.
- [31] D. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. Math.*, 148:339–360, 1998.
- [32] D. Schindler, R. Srivastava, and N. Technau. Rational points near manifolds, homogeneous dynamics, and oscillatory integrals. https://arxiv.org/abs/2310.03867, 2023.
- [33] D. Schindler and S. Yamagishi. Density of rational points near/on compact manifolds with certain curvature conditions. *Adv. Math.*, 403:Paper No. 108358, 36, 2022.
- [34] R. C. Vaughan and S. Velani. Diophantine approximation on planar curves: the convergence theory. *Invent. Math.*, 166(1):103–124, 2006.

Victor Beresnevich, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF YORK, UK *Email address*: victor.beresnevich@york.ac.uk

Shreyasi Datta, Department of Mathematics, University of York, UK *Email address*: shreyasi.datta@york.ac.uk