White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/95025/

Version: Published Version

Article:
Adiceam, Faustin, Beresnevich, Victor, Levesley, Jason et al. (2016) Diophantine Approximation and applications in Interference Alignment. Advances in Mathematics. pp. 231-279. ISSN: 0001-8708

https://doi.org/10.1016/j.aim.2016.07.002

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Advances in Mathematics 302 (2016) 231–279

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Diophantine approximation and applications in interference alignment ☆

F. Adiceam, V. Beresnevich, J. Levesley, S. Velani *, E. Zorin

Department of Mathematics, University of York, York, YO10 5DD, UK

ARTICLE INFO

Article history:
Received 8 June 2015
Received in revised form 26 June 2016
Accepted 4 July 2016
Available online xxxx
Communicated by Kenneth Falconer

Dedicated to Maurice Dodson

MSC:
primary 11J83
secondary 11J13, 11K60, 94A12

Keywords:
Metric Diophantine approximation
Khintchine–Groshev Theorem
Non-degenerate manifolds

ABSTRACT

This paper is motivated by recent applications of Diophantine approximation in electronics, in particular, in the rapidly developing area of Interference Alignment. Some remarkable advances in this area give substantial credit to the fundamental Khintchine–Groshev Theorem and, in particular, to its far reaching generalisation for submanifolds of a Euclidean space. With a view towards the aforementioned applications, here we introduce and prove quantitative explicit generalisations of the Khintchine–Groshev Theorem for non-degenerate submanifolds of $\mathbb{R}^n$. The importance of such quantitative statements is explicitly discussed in Jafar's monograph [12, §4.7.1].
© 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction 232
2. The theory for independent variables 236

☆ FA research is supported by EPSRC Programme Grant: EP/J018260/1. VB and SV research is supported in part by EPSRC Programme Grant: EP/M021858/1.
* Corresponding author.
E-mail addresses: Faustin.Adiceam@york.ac.uk (F. Adiceam), victor.beresnevich@york.ac.uk (V. Beresnevich), jason.levesley@york.ac.uk (J. Levesley), sanju.velani@york.ac.uk (S. Velani), evgeniy.zorin@york.ac.uk (E. Zorin).

http://dx.doi.org/10.1016/j.aim.2016.07.002
0001-8708/© 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

3. Diophantine approximation on manifolds . . . 243
4. Preliminaries for Theorem 3 . . . 244
5. A strengthening and proof of Theorem 3 . . . 249
6. Proof of Theorem 5 . . . 252
7. Verifying conditions (i) & (ii) of Theorem 7 . . . 254
8. Proof of Proposition 2 . . . 268
Acknowledgments . . . 275
Appendix A. Diophantine approximation on manifolds and wireless technology . . . 275
References . . . 279

1. Introduction

The present paper is motivated by a recent series of publications, including [11,12, 14–17,21–23], which utilise the theory of metric Diophantine approximation to develop new approaches in interference alignment, a concept within the field of wireless communication networks. This new link is both surprising and striking. The key ingredient from the number theoretic side is the fundamental Khintchine–Groshev Theorem and its variations. In this paper we seek to address certain problems in Diophantine approximation which crop up, or impinge upon, the applications to interference alignment. The results obtained represent quantitative refinements of the Khintchine–Groshev Theorem that are relevant to the applications mentioned above. Indeed, the desirability of such quantitative statements is explicitly eluded to in Jafar's monograph [12, §4.7.1]. While the main content of the paper is purely number theoretic, in Appendix A we attempt to illustrate at a basic level the manner in which Diophantine Approximation plays a natural role in Interference Alignment. The appendix is very much intended for the reader whose background is not in electronics but is nevertheless interested in applications.

Although the main emphasis will be on the Khintchine–Groshev Theorem for submanifolds of $\mathbb{R}^n$ [2,5,7,8,13,20], we begin by considering the classical theory for systems of linear forms of independent variables. This approach has two benefits. Firstly, we are able to introduce the key ideas without too much technical machinery obscuring the picture. Secondly, the refinements of the classical theory produce effective results with much better constants.

In order to recall Khintchine's theorem we first define the set $\mathcal{W}(\psi)$ of *ψ-well approximable numbers*. To this end, denote by $\mathbb{R}_+$ the set of non-negative real numbers. Given a real positive function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\psi(r) \to 0$ as $r \to \infty$, let then

$$\mathcal{W}(\psi) := \{x \in \mathbb{R} : |qx - p| < \psi(q) \text{ for i.m. } (q,p) \in \mathbb{N} \times \mathbb{Z}\},$$

where 'i.m.' reads 'infinitely many'. For obvious reasons the function ψ is often referred to as an *approximating function*. The points x in $\mathcal{W}(\psi)$ are characterised by the property that they admit approximation by rational points p/q with the error at most $\psi(q)/q$.

A simple 'volume' argument together with the Borel–Cantelli Lemma from probability theory implies that

$$|\mathcal{W}(\psi)| = 0 \quad \text{if} \quad \sum_{q=1}^{\infty} \psi(q) \ < \infty \, ,$$

where $|X|$ stands for the Lebesgue measure of $X \subset \mathbb{R}$. The above convergence statement represents the easier part of the following beautiful result due to Khintchine which gives a criterion for the size of the set $\mathcal{W}(\psi)$ in terms of Lebesgue measure. In what follows, we say that $X \subset \mathbb{R}$ is *full* in $\mathbb{R}$ and write $|X| = \textsc{Full}$ if $|\mathbb{R} \setminus X| = 0$; that is, the complement of X in $\mathbb{R}$ is of Lebesgue measure zero. The following is a slightly more general version of Khintchine, see [4].

Theorem A *(Khintchine, 1924). Let ψ be an approximating function. Then*

$$|\mathcal{W}(\psi)| \ = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) < \infty \, , \\ \textsc{Full} & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) = \infty \quad \textit{and } \psi \textit{ is monotonic.} \end{cases}$$

Thus, given any monotonic approximating function ψ, for almost all[1] $x \in \mathbb{R}$ the inequality $|x - p/q| < \psi(q)/q$ holds for infinitely many rational numbers p/q if and only if the sum $\sum_{q=1}^{\infty} \psi(q)$ diverges.

There are various generalisations of Khintchine's theorem to higher dimensions — see [3] for an overview. Here we shall consider the case of systems of linear forms which originates from a paper by Groshev in 1938. In what follows, m and n will denote positive integers and $M_{m,n}$ will stand for the set of $m \times n$ matrices over $\mathbb{R}$. Given a function $\Psi : \mathbb{Z}^n \to \mathbb{R}_+$, let

$$\mathcal{W}_{m,n}(\Psi) := \{\mathbf{X} = (x_{i,j}) \in M_{m,n} : \|\mathbf{X}\mathbf{a}\| < \Psi(\mathbf{a}) \text{ for i.m. } \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\} \, ,$$

where $\mathbf{a} = (a_1, \ldots, a_n)$,

$$\|\mathbf{X}\mathbf{a}\| := \max_{1 \le i \le m} \|x_{i,1}a_1 + \ldots + x_{i,n}a_n\|$$

and $\|x\| := \min\{|x - k| : k \in \mathbb{Z}\}$ is the distance of $x \in \mathbb{R}$ from the nearest integer. Given a subset X in $M_{m,n}$, we will write $|X|_{mn}$ for its ambient (i.e. mn-dimensional) Lebesgue measure. It is easily seen that $\mathcal{W}_{1,1}(\Psi)$ coincides with $\mathcal{W}(\psi)$ when $\Psi(q) = \psi(|q|)$. Therefore the following result is the natural extension of Theorem A to higher dimensions. Notice that there is no monotonicity assumption on the approximating function.

[1] 'For almost all' means for all except from a set of Lebesgue measure zero.

Theorem B. *Let $m, n \in \mathbb{N}$ with $nm > 1$, $\psi : \mathbb{N} \to \mathbb{R}_+$ be an approximating function and*

$$\Sigma_\psi := \sum_{q=1}^{\infty} q^{n-1}\psi(q)^m \,. \tag{1}$$

Let $\Psi : \mathbb{Z}^n \to \mathbb{R}_+$ be given by $\Psi(\mathbf{a}) := \psi(|\mathbf{a}|)$ for $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, where $|\mathbf{a}| = \max_{1\le i\le n} |a_i|$. Then

$$|\mathcal{W}_{m,n}(\Psi)|_{mn} \;=\; \begin{cases} 0 & \text{if} \quad \Sigma_\psi < \infty \,, \\ \text{Full} & \text{if} \quad \Sigma_\psi = \infty \,. \end{cases}$$

Theorem B was first obtained by Groshev under the assumption that $q^n\psi(q)^m$ is monotonic in the case of divergence. The redundancy of the monotonicity condition for $n \ge 3$ follows from Schmidt's paper [18, Theorem 2] and for $n = 1$ from Gallagher's paper [10]. Theorem B as stated was eventually proved in [6] where the remaining case of $n = 2$ was addressed. The convergence case of Theorem B is a relatively simple application of the Borel–Cantelli Lemma and it holds for arbitrary functions Ψ. Thus together with Theorem A, we have the following extremely general statement in the case of convergence.

Theorem C. *Let $m, n \in \mathbb{N}$ and $\Psi : \mathbb{Z}^n \to \mathbb{R}_+$ be any function such that the sum*

$$\Sigma_\Psi := \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{a})^m \tag{2}$$

converges. Then

$$|\mathcal{W}_{m,n}(\Psi)|_{mn} = 0 \,.$$

An immediate consequence of Theorem C is the following statement.

Corollary 1. *Let Ψ be as in Theorem C. Then, for almost every $\mathbf{X} \in M_{m,n}$ there exists a constant $\kappa(\mathbf{X}) > 0$ such that*

$$\|\mathbf{X}\mathbf{a}\| \;>\; \kappa(\mathbf{X})\,\Psi(\mathbf{a}) \qquad \forall \;\; \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \,. \tag{3}$$

In recent years estimates of this kind have become an important ingredient in the study of the achievable number of degrees of freedom in various schemes on Interference Alignment from electronics communication — see, e.g., [16]. The applications typically require that $\kappa(\mathbf{X})$ is independent of $\mathbf{X}$. Unfortunately, this is impossible to guarantee with probability 1, that is on a set of full Lebesgue measure. To demonstrate this claim, let us define the following set:

$$\mathcal{B}_{m,n}(\Psi,\kappa) := \left\{\mathbf{X} \in M_{m,n} : \|\mathbf{Xa}\| > \kappa\Psi(\mathbf{a}) \quad \forall\, \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\right\}. \tag{4}$$

Then, for any κ and Ψ, the set $\mathcal{B}_{1,n}(\Psi,\kappa)$ will not contain

$$[-\kappa\Psi(\mathbf{a}), \kappa\Psi(\mathbf{a})] \times \mathbb{R}^{n-1}$$

with $\mathbf{a} = (1, 0, \dots, 0)$. This set is of positive probability. In the light of this example it becomes highly desirable to address the following problem:

Problem. Investigate the dependence between κ and the probability of $\mathcal{B}_{m,n}(\Psi,\kappa)$.

As the first step to understanding this problem we obtain the following straightforward consequence of Theorem C.

Theorem 1. *Let $m, n \in \mathbb{N}$ and μ be a probability measure on $M_{m,n}$ that is absolutely continuous with respect to Lebesgue measure on $M_{m,n}$. Let $\Psi : \mathbb{Z}^n \to \mathbb{R}_+$ be any function such that (2) converges. Then for any $\delta \in (0,1)$ there is a constant $\kappa > 0$ depending only on μ, Ψ and δ such that*

$$\mu\left(\mathcal{B}_{m,n}(\Psi,\kappa)\right) \geq 1 - \delta. \tag{5}$$

Prior to giving a proof of this theorem recall that a measure μ on $M_{m,n}$ is *absolutely continuous with respect to Lebesgue measure* if there exists a Lebesgue integrable function $f : M_{m,n} \to \mathbb{R}_+$ such that for every Lebesgue measurable subset A of $M_{m,n}$, one has that

$$\mu(A) = \int_A f, \tag{6}$$

where $\int_A f$ is the Lebesgue integral of f over A. The function f is often referred to as the *distribution (or density) of μ*.

Proof. Since μ is absolutely continuous with respect to Lebesgue measure, Theorem C implies that $\mu(\mathcal{W}_{m,n}(\Psi)) = 0$. Hence $\mu(M_{m,n} \setminus \mathcal{W}_{m,n}(\Psi)) = \mu(M_{m,n}) = 1$. Note that

$$\bigcup_{\kappa>0} \mathcal{B}_{m,n}(\Psi,\kappa) = M_{m,n} \setminus \mathcal{W}_{m,n}(\Psi)\,.$$

Theorem 1 now follows on using the continuity of measures. □

In view of our previous discussion we have that $\kappa \to 0$ as $\delta \to 0$. Then, the above problem specialises to the explicit understanding of the dependence of κ on δ. This will be the main content of the next section. Subsequent sections will be devoted to obtaining similar effective version of the convergence Khintchine–Groshev Theorem for non-degenerate

submanifolds of $\mathbb{R}^n$. This constitutes the main substance of the paper. The results are obtained by exploiting the techniques of Bernik, Kleinbock and Margulis [8] originating from the seminal work of Kleinbock and Margulis [13] on the Baker–Sprindžuk conjecture.

2. The theory for independent variables

To begin with we give an alternative proof of Theorem 1 which introduces an explicit construction that will be utilised for quantifying the dependence of κ on δ. Indeed, in the case that μ is a uniform distribution on a unit cube the proof already identifies the required dependence.

2.1. Theorem 1 revisited

By a unit cube in $M_{m,n}$ we will mean a subset of $M_{m,n}$ given by

$$\big\{(x_{i,j}) \in M_{m,n} : \alpha_{i,j} \leq x_{i,j} < \alpha_{i,j} + 1\big\}$$

for some fixed matrix $(\alpha_{i,j}) \in M_{m,n}$. Given $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and $\varepsilon > 0$, let $\mathcal{W}(\mathbf{a}, \varepsilon)$ denote the set of $\mathbf{X} \in M_{m,n}$ such that

$$\|\mathbf{X}\mathbf{a}\| \ \leq \varepsilon\,. \tag{7}$$

It is easily seen that $\mathcal{W}(\mathbf{a}, \varepsilon)$ is invariant under additive translations by an integer matrix; that is,

$$\mathcal{W}(\mathbf{a}, \varepsilon) + \mathbf{B} = \mathcal{W}(\mathbf{a}, \varepsilon)$$

for any $\mathbf{B} \in M_{m,n}(\mathbb{Z})$, where $M_{m,n}(\mathbb{Z})$ denotes the set of $m \times n$ matrices with integer entries. Furthermore, we have that

$$|\mathcal{W}(\mathbf{a}, \varepsilon) \cap P|_{mn} = (2\varepsilon)^m \tag{8}$$

for any $0 \leq \varepsilon \leq \frac{1}{2}$ and any unit cube P in $M_{m,n}$. This follows, for example, from [19, Chapter 1, Lemma 8]. Then, since

$$\Sigma_\Psi := \sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} \Psi(\mathbf{a})^m < \infty \tag{9}$$

we must have that

$$M_\Psi := \sup\{\Psi(\mathbf{a}) : \mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}\} < \infty. \tag{10}$$

In what follows we will assume that

$$2\kappa M_\Psi \leq 1\,. \tag{11}$$

This condition ensures that we can apply (8) with $\varepsilon = \kappa\Psi(\mathbf{a})$.

Fix a unit cube $P_{\mathbf{0}}$ in $M_{m,n}$ and for each $\Delta \in M_{m,n}(\mathbb{Z})$, let

$$P_\Delta := P_{\mathbf{0}} + \Delta$$

denote the additive translation of $P_{\mathbf{0}}$ by Δ. Clearly, P_Δ itself is a unit cube. Furthermore,

$$M_{m,n} = \overset{\circ}{\bigcup_{\Delta \in M_{m,n}(\mathbb{Z})}} P_\Delta\,. \tag{12}$$

Note that the union is disjoint. Using (8) and the fact that

$$M_{m,n} \setminus \mathcal{B}_{m,n}(\Psi,\kappa) = \bigcup_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \mathcal{W}(\mathbf{a},\kappa\Psi(\mathbf{a}))\,,$$

we obtain that for each $\Delta \in M_{m,n}(\mathbb{Z})$,

$$\begin{aligned} |P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)|_{mn} &\leq \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} |\mathcal{W}(\mathbf{a},\kappa\Psi(\mathbf{a})) \cap P_\Delta|_{mn} \\ &= \sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} (2\kappa\Psi(\mathbf{a}))^m \;=\; (2\kappa)^m \Sigma_\Psi\,. \end{aligned} \tag{13}$$

Since μ is a probability measure, it follows from (12) that there exists a finite subset $A \subset M_{m,n}(\mathbb{Z})$ such that

$$\mu\left(\bigcup_{\Delta\in A} P_\Delta\right) > 1 - \delta/2\,. \tag{14}$$

Let $N = \#A$ be the number of elements in A. Since μ is absolutely continuous with respect to Lebesgue measure, for every $\Delta \in A$ and any $\varepsilon_1 > 0$, there exists ε_2 such that for any measurable subset X of P_Δ,

$$|X|_{mn} < \varepsilon_2 \quad \Rightarrow \quad \mu(X) < \varepsilon_1\,. \tag{15}$$

In view of (13), applying (15) to $X = P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)$ and $\varepsilon_1 = \delta/(2N)$ implies the existence of

$$\varepsilon_2 = \varepsilon_2(\Delta,\delta,N) > 0$$

such that

$$\mu(P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) < \delta/(2N) \qquad \text{if} \qquad (2\kappa)^m \Sigma_\Psi \le \varepsilon_2(\Delta,\delta,N)\,. \tag{16}$$

In particular, the second inequality in (16) holds if

$$\kappa \le \kappa_\Delta := \frac{1}{2}\left(\frac{\varepsilon_2(\Delta,\delta,N)}{\Sigma_\Psi}\right)^{1/m}.$$

Since A is finite, there exists κ satisfying (11) and

$$0 < \kappa \le \min_{\Delta\in A} \kappa_\Delta\,.$$

Clearly, for such a choice of κ the first inequality in (16) holds for any $\Delta \in A$. Hence, by (14) and the additivity of μ we obtain that

$$\begin{aligned}\mu(M_{m,n} \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) &\le \frac{\delta}{2} + \sum_{\Delta\in A} \mu(P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) \\ &\le \frac{\delta}{2} + \sum_{\Delta\in A} \frac{\delta}{2N} \;=\; \frac{\delta}{2} + N\frac{\delta}{2N} \;=\; \delta\,.\end{aligned}$$

The upshot of this is that

$$\mu(\mathcal{B}_{m,n}(\Psi,\kappa)) = 1 - \mu(M_{m,n} \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) \ge 1-\delta\,, \tag{17}$$

which completes the proof of Theorem 1.

2.2. Quantifying the dependence of κ on δ

We now turn our attention to quantifying the dependence of κ on δ within the context of Theorem 1. To this end, we will make use of the L_p norm. Given a Lebesgue measurable function $f : M_{m,n} \to \mathbb{R}_+$, a measurable subset X of $M_{m,n}$ and $p \ge 1$, we write $f \in L_p(X)$ if the Lebesgue integral

$$\int_X |f|^p := \int_{M_{m,n}} |f|^p \chi_X$$

exists and is finite. Here χ_X is the characteristic function of X. For $f \in L_p(X)$, the L_p norm of f on X is defined by

$$\|f\|_{p,X} := \left(\int_X |f|^p\right)^{1/p}.$$

In the case $p = \infty$, the L_∞-norm on X is defined as the essential maximum of $|f|$ on X; that is,

$$\|f\|_{\infty,X} := \inf\{c \in \mathbb{R} : |f(x)| \leq c \text{ for almost all } x \in X\} .$$

If $\|f\|_{\infty,X} < \infty$, then we write $f \in L_\infty(X)$. For example, if f is continuous and X is a non-empty open subset of $M_{m,n}$, then $\|f\|_{\infty,X}$ is simply the supremum of f on X. The following lemma gathers together two well know facts regarding the L_p norm.

Lemma 1.

(1) *For any $p \geq 1$ and any measurable subsets $X \subset Y$,*

$$\|f\|_{p,X} \leq \|f\|_{p,Y} .$$

(2) (*Hölder's inequality*) *For any $1 \leq p, q \leq \infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$,*

$$\left| \int_X fg \right| \leq \|f\|_{p,X} \|g\|_{q,X} .$$

The next lemma is a corollary of Lemma 1.

Lemma 2. *Let $p > 1$ and μ be a probability measure on $M_{m,n}$ with density f. Let X be a Lebesgue measurable subset of $M_{m,n}$. If $f \in L_p(X)$, then*

$$\mu(X) \leq \|f\|_{p,X} |X|_{mn}^{1-1/p} .$$

Proof. By definition, we have that

$$\mu(X) = \int_X f .$$

Define q by the equation $\frac{1}{p} + \frac{1}{q} = 1$. Then by Hölder's inequality, we have that

$$\mu(X) = \int_X f \times 1 \ \leq \ \|f\|_{p,X} \|1\|_{q,X} \ = \ \|f\|_{p,X} \left(\int_X 1^q \right)^{1/q} \ \leq \ \|f\|_{p,X} |X|_{mn}^{1-1/p}$$

as required. □

We are now in the position to provide an effective version of Theorem 1. Let $P_{\mathbf{0}}$ and A be the same as in §2.1. In particular, assume that (14) holds. Furthermore, assume

that there exists some $p > 1$ such that for every $\Delta \in A$, the density f of μ has finite L_p norm on P_Δ.
Let κ be such that (11) is satisfied. In this case, (13) holds for every P_Δ with $\Delta \in A$. By Lemmas 1 and 2,

$$\mu(P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) \leq \|f\|_{p,P_\Delta} \cdot |P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)|_{mn}^{1-1/p}.$$

Using (13), we obtain that

$$\mu(P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) \leq \|f\|_{p,P_\Delta} \cdot \Big((2\kappa)^m \Sigma_\Psi\Big)^{1-1/p} \tag{18}$$

where Σ_Ψ is given by (9). It follows that

$$\begin{aligned}\mu(M_{m,n} \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) &\leq \frac{\delta}{2} + \sum_{\Delta\in A} \mu(P_\Delta \setminus \mathcal{B}_{m,n}(\Psi,\kappa)) \\ &\leq \frac{\delta}{2} + \Big((2\kappa)^m \Sigma_\Psi\Big)^{1-1/p} \Sigma_f \;\leq\; \delta\end{aligned}$$

if

$$\kappa \leq \frac{1}{2}\left(\Sigma_\Psi^{-1}\left(\frac{\delta}{2\Sigma_f}\right)^{\frac{p}{p-1}}\right)^{1/m},$$

where

$$\Sigma_f := \sum_{\Delta\in A} \|f\|_{p,P_\Delta}. \tag{19}$$

Since A is finite, the quantity Σ_f is also finite. The upshot of the above discussion is the following statement.

Theorem 2 *(Effective version of Theorem 1). Let $m, n \in \mathbb{N}$, μ, Ψ be as in Theorem 1, let M_Ψ be given by (10) and let f denote the density of μ. Furthermore, let $P_{\mathbf{0}}$ be any unit cube in $M_{m,n}$ and A be any finite subset of $M_{m,n}(\mathbb{Z})$ satisfying (14). Assume there exists $p > 1$ such that $f \in L_p(P_\Delta)$ for any $\Delta \in A$ and also assume that the quantity Σ_f is given by (19). Then, for any $\delta \in (0,1)$, inequality (5) holds with*

$$\kappa := \frac{1}{2}\min\left\{\frac{1}{M_\Psi},\ \left(\Sigma_\Psi^{-1}\left(\frac{\delta}{2\Sigma_f}\right)^{\frac{p}{p-1}}\right)^{1/m}\right\}. \tag{20}$$

In this formula, the quotient $p/(p-1)$ should be taken as equal to 1 when $p = \infty$.

Remark 1. In the case when Ψ is even, that is $\Psi(-\mathbf{a}) = \Psi(\mathbf{a})$ for all $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, one can improve formula (20) for κ by replacing Σ_Ψ with $\frac{1}{2}\Sigma_\Psi$. This is an obvious consequence of the fact that in this case the sets $\mathcal{W}(\mathbf{a}, \kappa\Psi(\mathbf{a}))$ and $\mathcal{W}(-\mathbf{a}, \kappa\Psi(-\mathbf{a}))$ coincide and therefore do not have to be counted twice within the proof.

There are various simplifications and specialisations of Theorem 2 when we have extra information regarding the measure μ. The following is a natural corollary which is particularly relevant for probability measures μ with bounded distribution f and mean value about the origin.

Corollary 2. *Let $m, n \in \mathbb{N}$, μ, Ψ, M_Ψ be as in Theorem 1 and let the density f of μ be bounded above by a constant $K > 0$. Furthermore, let T be the smallest positive integer such that*

$$\mu\left([-T, T)^{mn}\right) \geq 1 - \delta/2. \tag{21}$$

Then, for any $\delta \in (0, 1)$, inequality (5) holds with

$$\kappa := \frac{1}{2} \min\left\{ \frac{1}{M_\Psi},\ \left(\frac{\delta}{2K(2T)^{mn}\Sigma_\Psi} \right)^{1/m} \right\}. \tag{22}$$

Proof. With respect to Theorem 2, let $p = \infty$ and A be the collection of cubes P_Δ that exactly tiles $[-T, T)^{mn}$. Then, $\#A = (2T)^{mn}$ and thus $\Sigma_f \leq K(2T)^{mn}$. Now, (22) trivially follows from (20). □

2.3. Numerical examples

In what follows, we will use the standard Gaussian error function

$$\operatorname{erf}(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2} \mathrm{d}t\,.$$

It is readily verified that the function erf is continuous, strictly increasing and that

$$\lim_{x \to -\infty} \operatorname{erf}(x) = 0 \qquad \text{and} \qquad \lim_{x \to +\infty} \operatorname{erf}(x) = 1.$$

As usual, for $0 < y < 1$, define $\operatorname{erf}^{-1}(y)$ to be the unique value $x \in \mathbb{R}$ such that $\operatorname{erf}(x) = y$. Furthermore, define formally $\operatorname{erf}^{-1}(0) := -\infty$ and $\operatorname{erf}^{-1}(1) := +\infty$.

Consider now Corollary 2 in the case when $m = n = 1$ and when μ follows the standard Gaussian distribution $\mathcal{N}(0, 1)$. It can then be verified that Corollary 2 implies that inequality (5) holds with

$$\kappa = \frac{\delta\sqrt{2\pi}}{8 \cdot N \cdot \Sigma_\Psi}\,, \tag{23}$$

where $N := \lceil \mathrm{erf}^{-1}(1-\delta/4) \rceil$. Here $\lceil x \rceil$ is the "ceiling" of x, that is the smallest integer that is bigger than or equal to $x \in \mathbb{R}$. We now consider explicit approximating functions. First, let Ψ be the function given by $\Psi(q) = 0$ if $q \leq 0$,

$$\Psi(q) := \frac{1}{2q \cdot \log^2 q} \quad \text{if} \quad q \geq 2 \quad \text{and} \quad \Psi(1) := 1/2\,.$$

Then $\Sigma_\Psi < 1.555$ and on making use of (23) we obtain the following table for values of N and κ:

δ	0.5	0.25	0.1	0.01	10^{-3}	10^{-5}
N	2	2	3	4	4	5
κ	0.05	0.025	0.0067	$5 \cdot 10^{-4}$	$5 \cdot 10^{-5}$	$4 \cdot 10^{-7}$

It follows for instance from this set of data that for 99% of the values of the random variable x with normal distribution $\mathcal{N}(0,1)$, one has that

$$\|qx\| \;>\; \frac{1}{2000} \cdot \Psi(q) \qquad \text{for all} \quad q \in \mathbb{N}.$$

In the next example, we fix a $Q \in \mathbb{N}$ and consider the approximating function Ψ given by

$$\Psi(q) := \begin{cases} \frac{1}{Q} & \text{if } 1 \leq q \leq Q, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\Sigma_\Psi = 1$ and one can readily verify that

(i) for at least 75% of the values of the random variable x with normal distribution $\mathcal{N}(0,1)$, one has that

$$\|qx\| > \frac{1}{13Q} \qquad \text{for all} \quad q \in [-Q, Q],\ q \neq 0,$$

(ii) for at least 90% of the values of the random variable x with normal distribution $\mathcal{N}(0,1)$, one has that

$$\|qx\| > \frac{1}{50Q} \qquad \text{for all} \quad q \in [-Q, Q],\ q \neq 0.$$

3. Diophantine approximation on manifolds

The aim is to establish an analogue of Theorem 2 for submanifolds $\mathcal{M}$ of $\mathbb{R}^n$. More precisely, we consider the set $\mathcal{B}_n(\Psi,\kappa)\cap\mathcal{M}$, where

$$\mathcal{B}_n(\Psi,\kappa) := \mathcal{B}_{1,n}(\Psi,\kappa)\,.$$

The fact that the points of interest are of dependent variables, reflecting the fact that they lie on $\mathcal{M}$, introduces major difficulties in attempting to describe the measure theoretic structure of $\mathcal{B}_n(\Psi,\kappa)\cap\mathcal{M}$.

Non-degenerate manifolds. In order to make any reasonable progress with the above problems it is not unreasonable to assume that the manifolds $\mathcal{M}$ under consideration are *non-degenerate*. Essentially, these are smooth submanifolds of $\mathbb{R}^n$ which are sufficiently curved so as to deviate from any hyperplane. Formally, a manifold $\mathcal{M}$ of dimension d embedded in $\mathbb{R}^n$ is said to be *non-degenerate* if it arises from a non-degenerate map $\mathbf{f}:\mathbf{U}\to\mathbb{R}^n$ where $\mathbf{U}$ is an open subset of $\mathbb{R}^d$ and $\mathcal{M}:=\mathbf{f}(\mathbf{U})$. The map $\mathbf{f}:\mathbf{U}\to\mathbb{R}^n, \mathbf{x}\mapsto\mathbf{f}(\mathbf{x})=(f_1(\mathbf{x}),\dots,f_n(\mathbf{x}))$ is said to be *l-non-degenerate at* $\mathbf{x}\in\mathbf{U}$, where $l\in\mathbb{N}$, if $\mathbf{f}$ is l times continuously differentiable on some sufficiently small ball centred at $\mathbf{x}$ and the partial derivatives of $\mathbf{f}$ at $\mathbf{x}$ of orders up to l span $\mathbb{R}^n$. The map $\mathbf{f}$ is *non-degenerate* at $\mathbf{x}$ if it is l-non-degenerate at $\mathbf{x}$ for some $l\in\mathbb{N}$. As is well known, any real connected analytic manifold not contained in any hyperplane of $\mathbb{R}^n$ is non-degenerate at every point [13].

Observe that if the dimension of the manifold $\mathcal{M}$ is strictly less than n then we have that $|\mathcal{B}_n(\Psi,\kappa)\cap\mathcal{M}|_n=0$ irrespective of the approximating function Ψ and κ. Thus, when referring to the Lebesgue measure of the set $\mathcal{B}_n(\Psi,\kappa)\cap\mathcal{M}$ it is always with reference to the induced Lebesgue measure on $\mathcal{M}$. More generally, given a subset S of $\mathcal{M}$ we shall write $|S|_{\mathcal{M}}$ for the measure of S with respect to the induced Lebesgue measure on $\mathcal{M}$. Without loss of generality, we will assume that $|\mathcal{M}|_{\mathcal{M}}=1$ as otherwise the measure can be re-normalised accordingly.

The following statement is a straightforward consequence of the main result of Bernik, Kleinbock and Margulis in [8].

Theorem BKM. *Let $\mathcal{M}$ be a non-degenerate submanifold of $\mathbb{R}^n$. Let $\Psi:\mathbb{Z}^n\to\mathbb{R}_+$ be monotonically decreasing in each variable and such that*

$$\Sigma_\Psi := \sum_{\mathbf{q}\in\mathbb{Z}^n\setminus\{0\}}\Psi(\mathbf{q})<\infty\,. \tag{24}$$

Then, for any $\delta\in(0,1)$, there is a constant $\kappa>0$ depending on $\mathcal{M}$, Σ_Ψ and δ only such that

$$|\mathcal{B}_n(\Psi,\kappa)\cap\mathcal{M}|_{\mathcal{M}}\geq 1-\delta. \tag{25}$$

Remark 2. Theorem BKM holds for arbitrary probability measures supported on $\mathcal{M}$ that are absolutely continuous with respect to the induced Lebesgue measure on $\mathcal{M}$, thus giving an analogue of Theorem 2 for manifolds. As in the case of Theorem 1, the more general result follows from the Lebesgue statement.

It is worth pointing out that the main result in [8] actually implies that the union $\bigcup_{\kappa>0} \mathcal{B}_n(\Psi,\kappa) \cap \mathcal{M}$ has full measure on $\mathcal{M}$. Theorem BKM as stated above follows from [8, Theorem 1.1][2] on using the continuity of measures. Our main goal is to quantify the dependence of κ on δ. Theorem 6 of §5 below explicitly quantifies this dependence. However, the statement is rather technical and we prefer to state for now a cleaner result that shows that the dependency between κ and δ is polynomial.

Theorem 3. *Let $l \in \mathbb{N}$ and let $\mathcal{M}$ be a compact d-dimensional C^{l+1} submanifold of $\mathbb{R}^n$ that is l-non-degenerate at every point. Let μ be a probability measure supported on $\mathcal{M}$ absolutely continuous with respect to $|\,.\,|_{\mathcal{M}}$. Let $\Psi : \mathbb{Z}^n \to \mathbb{R}_+$ be a monotonically decreasing function in each variable satisfying (24). Then there exist positive constants κ_0, C_1 depending on Ψ and $\mathcal{M}$ only and C_0 depending on the dimension of $\mathcal{M}$ only such that for any $0 < \delta < 1$, the inequality*

$$\mu(\mathcal{B}_n(\Psi,\kappa) \cap \mathcal{M}) \geq 1 - \delta \tag{26}$$

holds with

$$\kappa := \min\left\{\kappa_0,\ C_0\Sigma_\Psi^{-1}\delta,\ C_1\delta^{d(n+1)(2l-1)}\right\}. \tag{27}$$

4. Preliminaries for Theorem 3

4.1. Localisation and parameterisation

Since $\mathcal{M}$ is non-degenerate everywhere, we can restrict ourself to considering a sufficiently small neighbourhood of an arbitrary point on $\mathcal{M}$. By compactness, $\mathcal{M}$ then can be covered with a finite subcollection of such neighbourhoods. Therefore, in view of the finiteness of the cover, the existence of κ_0, C_0 and C_1 satisfying Theorem 3 globally will follow from the existence of these parameters for every neighbourhood in the finite cover: κ_0, C_0 and C_1 should be taken to be the minimum of their local values.

Now as we can work with $\mathcal{M}$ locally, we can parameterise it with some map $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ defined on a ball $\mathbf{U}$ in $\mathbb{R}^d$, where $d = \dim \mathcal{M}$. Note that $\mathbf{f}$ must be at least C^2 in order to ensure that $\mathcal{M}$ is non-degenerate. Without loss of generality we assume that

$$\mathcal{M} = \{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in \mathbf{U}\}.$$

[2] Throughout, results and page numbers within [8] are with reference to the arXiv version: math/0210298v1.

Furthermore, using the Implicit Function Theorem if necessary, we can make $\mathbf{f}$ to be a Monge parametrisation, that is $\mathbf{f}(\mathbf{x}) = (x_1, \dots, x_d, f_{d+1}(\mathbf{x}), \dots, f_n(\mathbf{x}))$, where $\mathbf{x} = (x_1, \dots, x_d)$. Note that $\mathbf{f}$ can be assumed to be bi-Lipschitz on $\mathbf{U}$. This readily follows from the fact that $\mathbf{f}$ is C^1 but possibly requires a further shrinking of $\mathbf{U}$.

Let $\mathcal{B}_n(\Psi, \kappa, \mathcal{M})$ denote the orthogonal projection of $\mathcal{B}_n(\Psi, \kappa) \cap \mathcal{M}$ onto the set of parameters $\mathbf{U}$. Thus,

$$\mathcal{B}_n(\Psi, \kappa, \mathcal{M}) := \{\mathbf{x} \in \mathbf{U} : \|\mathbf{a}.\mathbf{f}(\mathbf{x})\| > \kappa\Psi(\mathbf{a}) \ \text{ for all } \ \mathbf{a} \in \mathbb{Z}^n, \ \mathbf{a} \neq \mathbf{0}\}. \tag{28}$$

The set $\mathcal{B}_n(\Psi, \kappa) \cap \mathcal{M}$ and its projection $\mathcal{B}_n(\Psi, \kappa, \mathcal{M})$ are related by the bi-Lipschitz map $\mathbf{f}$. Since bi-Lipschitz maps only affect the Lebesgue measure of a set by a multiplicative constant (in this case the constant will depend on $\mathbf{f}$ only), it suffices to prove Theorem 3 for the project set. More precisely, Theorem 3 is equivalent to showing that there exist positive constants κ_0, C_0 and $C_1 > 0$ depending on Ψ and $\mathbf{f}$ only such that for any $0 < \delta < 1$,

$$|\mathcal{B}_n(\Psi, \kappa, \mathcal{M})|_d \geq (1 - \delta)|\mathbf{U}|_d \tag{29}$$

holds with κ given by (27).

4.2. Auxiliary statements

We will denote the standard L_1 (resp. Euclidean, infinity) norm on $\mathbb{R}^d$ by $\|\,.\,\|_1$ (resp. $\|\,.\,\|_2$, $\|\,.\,\|_\infty$). Also as before, given an $x \in \mathbb{R}$, $\|x\|$ will denote the distance of x from the nearest integer. The notation $B(\mathbf{x}, r)$ will refer to the Euclidean open ball of radius $r > 0$ centred at $\mathbf{x}$ and $\mathbb{S}^{d-1}$ will denote the unit sphere in dimension $d \geq 1$ (with respect to the Euclidean norm). Furthermore, throughout

$$V_d := \frac{\pi^{d/2}}{\Gamma\left(1 + d/2\right)}$$

is the volume of the d-dimensional unit ball and N_d denotes the Besicovitch covering constant.

Remark 3. For further details on the Besicovitch covering constant, cf. [9]. We will only need in what follows the inequality $N_d \leq 5^d$ satisfied by this constant.

The proof of Theorem 3 involves two separate cases that take into consideration the relative size of the gradient of $\mathbf{f}(\mathbf{x}) \cdot \mathbf{q}$, where $\mathbf{q} = (q_1, \dots, q_n) \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and $\mathbf{f}(\mathbf{x}) \cdot \mathbf{q} = f_1(\mathbf{x})q_1 + \cdots + f_n(\mathbf{x})q_n$ is the standard inner product of $\mathbf{f}(\mathbf{x})$ and $\mathbf{q}$. The first case of 'big gradient' is considered within the next result and is an adaptation of [8, Theorem 1.3].

In what follows, ∂_β will denote partial derivation with respect to a multi-index $\beta = (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$, where $\mathbb{N}_0$ will stand for the set of non-negative integers, that is $\mathbb{N}_0 := \{0, 1, 2, \dots\}$. Furthermore, $|\beta|$ we will mean the order of derivation, that is $|\beta| := \beta_1 + \cdots + \beta_d$. Also, ∂_i^k will denote the differential operator corresponding to the kth derivative with respect to the ith variable, that is, $\partial_i^k := \partial^k / \partial x_i^k$.

Theorem 4. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball of radius r and $\mathbf{f} \in C^2(2\mathbf{U})$, where $2\mathbf{U}$ is the ball with the same centre as $\mathbf{U}$ and radius $2r$. Let*

$$L^* := \sup_{|\beta|=2,\ \mathbf{x}\in 2\mathbf{U}} \|\partial_\beta \mathbf{f}(\mathbf{x})\|_\infty \qquad \text{and} \qquad L := \max\left(L^*, \frac{1}{4r^2}\right). \tag{30}$$

Then, for every $\delta' > 0$ and every $\mathbf{q} \in \mathbb{Z}^n \setminus \{0\}$, the set of $\mathbf{x} \in \mathbf{U}$ such that

$$\|\mathbf{f}(\mathbf{x}) \cdot \mathbf{q}\| < \delta'$$

and

$$\|\nabla \mathbf{f}(\mathbf{x})\mathbf{q}\|_\infty \geq \sqrt{ndL\|\mathbf{q}\|_\infty} \tag{31}$$

has measure at most $K_d \delta' |\mathbf{U}|_d$, where $\nabla \mathbf{f}(\mathbf{x})\mathbf{q}$ is the gradient of $\mathbf{f}(\mathbf{x}) \cdot \mathbf{q}$ and

$$K_d := \frac{4^{2d+1} d^{d/2} N_d}{V_d} \tag{32}$$

is a constant depending on d only.

Proof. The proof of Theorem 4 follows on appropriately applying [8, Lemma 2.2]. For convenience we refer to this lemma as L2.2. We take M in L2.2 to be equal to the quantity ndL, where L is defined by (30). We set δ in L2.2 to be equal to δ' appearing in Theorem 4. Then, in view of (30) and the fact that $n, d, \|\mathbf{q}\|_\infty \geq 1$, it follows that the hypotheses of L2.2 (namely [8, Eq. (2.1a) & Eq. (2.1b)]) are satisfied. Thus, the conclusion of L2.2 implies Theorem 4 with the constant C_d in L2.2 equal to K_d appearing in Theorem 4. The explicit value of K_d is calculated by 'tracking' the values of the auxiliary constants C_d', C_d'' and C_d''' appearing in [8]. Namely,[3]

$$C_d' = \frac{V_d}{2^{2d} d^{d/2}}, \qquad C_d'' = 2^{d+2}, \qquad C_d''' = \frac{C_d''}{C_d'},$$

[3] There are two typos in the proof of L2.2 that one should be aware of when verifying the values of the constants given here. On page 6 line −2, the inclusion regarding $U(x)$ is the wrong way round, it should read $U(x) \supset B(x, \frac{\rho}{4\sqrt{d}})$. Next, on page 7 line 11, in the rightmost term of the displayed set of inequalities the quantity δ is missing, it should read $C_d''' \delta |U(x)|_d$. These typos do not affect the validity of the proof given in [8].

and then

$$K_d = 2^d C_d''' N_d = \frac{2^d C_d'' N_d}{C_d'} = \frac{2^{4d+2} d^{d/2} N_d}{V_d}. \quad \square$$

Next in Theorem 5 below we consider the case of 'small gradient'. This is an explicit version of [8, Theorem 1.4]. First we introduce auxiliary constants.

Given a C^l map $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$ defined on a ball $\mathbf{U}$ in $\mathbb{R}^d$, the supremum of $s \in \mathbb{R}$ such that for any $\mathbf{x} \in \mathbf{U}$ and any $\mathbf{v} \in \mathbb{S}^{n-1}$ there exists an integer k, $0 < k \leq l$, and a unit vector $\mathbf{u} \in \mathbb{S}^{d-1}$ satisfying

$$\left| \frac{\partial^k (\mathbf{f} \cdot \mathbf{v})}{\partial \mathbf{u}^k}(\mathbf{x}) \right| \geq s \tag{33}$$

will be called *the measure of l-non-degeneracy of* $(\mathbf{f}, \mathbf{U})$ and will be denoted by $s(l; \mathbf{f}, \mathbf{U})$. Here and elsewhere for a unit vector $\mathbf{u} \in \mathbb{S}^{d-1}$, $\partial^k / \partial \mathbf{u}^k$ will denote the derivative in direction $\mathbf{u}$ of order k.

As in Theorem 4, the radius of the ball $\mathbf{U}$ will be denoted by r. Throughout, we let $\mathbf{x}_0$ denote the centre of $\mathbf{U}$. Also, given a real number $\lambda > 0$, we let $\lambda \mathbf{U}$ denote the scaled ball of radius λr and with the same centre $\mathbf{x}_0$ as that of $\mathbf{U}$. With this in mind, consider the balls

$$\begin{aligned} \mathbf{U}^+ &:= 3^{d+1}\mathbf{U}, \\ \tilde{\mathbf{U}} &:= 3^{n+1}\mathbf{U}, \\ \tilde{\mathbf{U}}^+ &:= 3^{n+d+2}\mathbf{U}. \end{aligned} \tag{34}$$

For technical reasons, that will soon become apparent, in order to deal with the 'small gradient' case we make the following assumption on the map $\mathbf{f} : \mathbf{U} \to \mathbb{R}^n$.

Assumption 1. *The map $\mathbf{f} = (f_1, \ldots, f_n)$ is an n-tuple of C^{l+1} functions defined on the closure of $\tilde{\mathbf{U}}^+$ which is l-non-degenerate everywhere on the closure of $\tilde{\mathbf{U}}^+$.*

Remark. In view of the discussion of §4.1, there is no loss of generality in imposing Assumption 1 within the context of Theorem 3.

We denote by s_0 the measure of non-degeneracy of $\mathbf{f}$ on $\tilde{\mathbf{U}}^+$. Note that Assumption 1 ensures that

$$s_0 := s(l; \mathbf{f}, \tilde{\mathbf{U}}^+) > 0. \tag{35}$$

Also, notice that it ensures the existence of a constant $M \geq 1$ such that for all $k \leq l+1$ and all $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{S}^{d-1}$,

$$\sup_{\mathbf{x}\in\tilde{\mathbf{U}}^+}\left\|\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}_1\dots\partial\mathbf{u}_k}(\mathbf{x})\right\|_2\leq M\,, \tag{36}$$

where $\partial\mathbf{u}_i$ means differentiation in direction $\mathbf{u}_i$. Note that the left-hand side of (33) is the length of the projection of $\partial^k\mathbf{f}(\mathbf{x})/\partial\mathbf{u}^k$ on the line passing through $\mathbf{v}$ and hence it is no bigger than M. This implies that

$$s_0\ \leq\ M\,.$$

Without loss of generality, we will assume that the radius r of the ball $\mathbf{U}$ satisfies

$$r\leq\min\left\{\frac{s_0\cdot\sigma(l,d)}{2\cdot 3^{n+d+2}\sqrt{d}M},\frac{\eta s_0}{4\cdot 10^7 3^{n+d+2}\,dMl^{l+2}(l+1)!}\right\}, \tag{37}$$

where

$$\eta:=\min\left\{\frac{1}{16},\left(\frac{V_d}{2^{d+2}dl(l+1)^{1/l}5^d}\right)^{d(2l-1)(2l-2)}\right\} \tag{38}$$

and where

$$\sigma(l,d):=\frac{1}{2^{3l(d-1)/2}}\cdot\phi\left(\left(\sqrt{2}\cdot(2l)^{2+l(l-1)/2}\cdot(l+1)!\right)^{-1},2,l\right)^{d-1} \tag{39}$$

with the quantity $\phi(\omega,B,k)$ defined as

$$\phi(\omega,B,k):=\frac{\omega^{k(k-1)/2}}{2\sqrt{2}\cdot B^k\cdot(k+1)!} \tag{40}$$

for any integer $k\geq 1$ and any real numbers $\omega,B>0$.

Furthermore, define the following constants determined by $\mathbf{f}$ and $\mathbf{U}$:

$$\rho_1:=\frac{s_0}{4l^l(l+1)!\sqrt{d}}(2r)^l\,, \tag{41}$$

$$\tau:=\frac{r^l s_0}{4l^l(l+1)!},$$

and

$$\rho_2:=\frac{s_0}{4l^l(l+1)!}\left(\frac{\tau}{M}\right)^{l-1}\frac{\left(\tau\left(1-1/\sqrt{2}\right)\right)^2}{\sqrt{\left(\frac{s_0}{4l^l(l+1)!}\left(\frac{\tau}{M}\right)^l\right)^2+\left(\tau\left(1+\frac{1}{\sqrt{2}}\right)\right)^2}}\cdot \tag{42}$$

Finally, let

$$\rho:=\frac{\rho_1\rho_2}{\sqrt{\rho_1^2+(\rho_2+2M^2)^2}}\cdot \tag{43}$$

Theorem 5. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball and $\mathbf{f} = (f_1, \dots, f_n)$ be an n-tuple of C^{l+1} functions satisfying Assumption 1. Then, for any $0 < \delta' \leq 1$, any n-tuple $\mathbf{T} = (T_1, \dots, T_n)$ of real numbers ≥ 1 and any $K > 0$ satisfying*

$$\frac{\delta' K T_1 \cdot \dots \cdot T_n}{\max_{i=1,\dots,n} T_i} \leq 1, \tag{44}$$

define the set $A(\delta', K, \mathbf{T})$ to be

$$A(\delta', K, \mathbf{T}) := \left\{ \mathbf{x} \in \mathbf{U} : \exists\ \mathbf{q} \in \mathbb{Z}^n \setminus \{0\} \text{ such that } \left\{ \begin{array}{l} \|\mathbf{f}(\mathbf{x}) \cdot \mathbf{q}\| < \delta' \\ \|\nabla \mathbf{f}(\mathbf{x}) \mathbf{q}\|_\infty < K \\ |q_i| < T_i, i = 1, \dots, n \end{array} \right\} \right. . \tag{45}$$

Then

$$|A(\delta', K, \mathbf{T})|_d \leq E \left(\sqrt{n+d+1} \cdot \varepsilon_1 \right)^{1/d(2l-1)} |\mathbf{U}|_d \,,$$

where

$$\varepsilon_1 := \max \left(\delta', \left(\frac{\delta' K T_1 \cdot \dots \cdot T_n}{\max\limits_{1 \leq i \leq n} T_i} \right)^{\frac{1}{n+1}} \right) \tag{46}$$

and

$$E := C(n+1)(3^d N_d)^{n+1} \rho^{-1/d(2l-1)} \,, \tag{47}$$

in which ρ is given by (43) and C is the constant explicitly given by (62) below.

At first glance the statement of Theorem 5 looks very similar to [8, Theorem 1.4]. We stress that the key difference is that in our statement the constants are made fully explicit. The proof of Theorem 5 is rather involved and will be the subject of §6.

5. A strengthening and proof of Theorem 3

In view of the discussion of §4.1, Theorem 3 will follow immediately on establishing a stronger result (Theorem 6 below), which explicitly characterises the dependence on Ψ and $\mathcal{M}$ of the constants κ_0, C_0 and C_1 appearing within the statement of Theorem 3. In the case that the function $\mathbf{f}$ defining the manifold under consideration is explicitly given, the values of these constants may be improved by following the methodology of the proof of Theorem 6 as many computations will then be made simpler.

Let

$$C_\Psi := \sup_{\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{q})\prod_{i=1}^{n} q_i^+, \qquad \text{where } q_i^+ := \max\{1,|q_i|\}\,. \tag{48}$$

It is a well known fact that, under the assumption that Ψ is monotonically decreasing in each variable, relation (24) implies that $0 < C_\Psi < \infty$. Also define the constant

$$S_n := \sum_{\mathbf{t}\in\mathbb{Z}^n} 2^{-\frac{\|\mathbf{t}\|_\infty}{2d(2l-1)(n+1)}},$$

which is clearly finite and positive as the sum converges.

Theorem 6. *Let $\mathbf{U}\subset\mathbb{R}^d$ be a ball whose radius satisfies (37) and let $\mathbf{f}=(f_1,\dots,f_n)$ be an n-tuple of C^{l+1} functions satisfying Assumption 1. Let Σ_Ψ, L, K_d and E be given by (24), (30), (32) and (47) respectively and let*

$$K_0 = E\left(\sqrt{n+d+1}\cdot\left(C_\Psi 2^{n-1/2}\sqrt{ndL}\right)^{\frac{1}{n+1}}\right)^{1/d(2l-1)}.$$

Given any $\delta>0$, let

$$\kappa := \min\left\{\frac{1}{C_\Psi 2^{n-1/2}\sqrt{ndL}},\ \frac{\delta}{2K_d\Sigma_\Psi},\ \left(\frac{\delta}{2K_0S_n}\right)^{d(n+1)(2l-1)}\right\}.$$

Then

$$|\mathcal{B}_n(\Psi,\kappa,\mathcal{M})|_d \geq (1-\delta)|\mathbf{U}|_d. \tag{49}$$

Clearly the above is an explicit version of Theorem 3 in the case when μ is Lebesgue measure. The arguments given in the proof of Theorem 2 are easily adapted to deal with the general situation.

5.1. Proof of *Theorem 6* modulo *Theorem 5*

For $\kappa>0$ and any $\mathbf{q}\in\mathbb{Z}^n$, define

$$A(\kappa;\mathbf{q}) := \{\mathbf{x}\in\mathbf{U} : \|\mathbf{f}(\mathbf{x})\cdot\mathbf{q}\| < \kappa\Psi(\mathbf{q}) \ \&\ (31) \text{ holds}\}$$

and

$$A^c(\kappa;\mathbf{q}) := \{\mathbf{x}\in\mathbf{U} : \|\mathbf{f}(\mathbf{x})\cdot\mathbf{q}\| < \kappa\Psi(\mathbf{q}) \ \&\ (31) \text{ does not hold}\}\,.$$

Clearly it suffices to prove that

$$\left|\bigcup_{\mathbf{q}\in\mathbb{Z}^n\setminus\{0\}} A(\kappa;\mathbf{q})\right|_d \leq \frac{\delta}{2}|\mathbf{U}|_d \qquad \text{and} \qquad \left|\bigcup_{\mathbf{q}\in\mathbb{Z}^n\setminus\{0\}} A^c(\kappa;\mathbf{q})\right|_d \leq \frac{\delta}{2}|\mathbf{U}|_d\,. \tag{50}$$

By Theorem 4 with $\delta' = \kappa\Psi(\mathbf{q})$, we immediately have that $|A(\kappa;\mathbf{q})|_d \leq K_d\kappa\Psi(\mathbf{q})|\mathbf{U}|_d$. Then, summing over all $\mathbf{q}\in\mathbb{Z}^n\setminus\{0\}$ gives

$$\left|\bigcup_{\mathbf{q}\in\mathbb{Z}^n\setminus\{0\}} A(\kappa;\mathbf{q})\right|_d \leq K_d\Sigma_\Psi\kappa|\mathbf{U}|_d \leq \frac{\delta}{2}\,|\mathbf{U}|_d. \tag{51}$$

Now to establish the second inequality in (50), given an n-tuple $\mathbf{t} = (t_1,\dots,t_n)\in\mathbb{N}_0^n$, define the set

$$A^c_{\mathbf{t}} := \bigcup_{\substack{\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{Z}^n\setminus\{\mathbf{0}\}\\ 2^{t_i}\leq q_i^+<2^{t_i+1}}} A^c(\kappa;\mathbf{q})\,, \tag{52}$$

where $q_i^+ = \max\{1,|q_i|\}$. Observe that

$$\bigcup_{\mathbf{q}\in\mathbb{Z}^n} A^c(\kappa;\mathbf{q}) = \bigcup_{\mathbf{t}\in\mathbb{N}_0^n} A^c_{\mathbf{t}}\,. \tag{53}$$

By (48) and the monotonicity of Ψ in each variable, for every $\mathbf{q} = (q_1,\dots,q_n)\in\mathbb{Z}^n\setminus\{\mathbf{0}\}$ satisfying the inequalities $2^{t_i}\leq q_i^+ < 2^{t_i+1}$, we have that

$$\Psi(\mathbf{q}) \leq C_\Psi\left(\prod_{i=1}^n q_i^+\right)^{-1} \leq C_\Psi\prod_{i=1}^n 2^{-t_i} = C_\Psi 2^{-\sum_{i=1}^n t_i}$$

and

$$\|\mathbf{q}\|_\infty \leq 2^{\max_i t_i+1}\,.$$

Now let

$$\delta' = \kappa C_\Psi 2^{-\sum_{i=1}^n t_i}, \quad K = \sqrt{ndL2^{\max_i t_i+1}} \quad \text{and} \quad T_i = 2^{t_i+1}\ (1\leq i\leq n)\,.$$

Then, $A^c_{\mathbf{t}}$ is easily seen to be contained in the set $A(\delta',K,\mathbf{T})$ defined within Theorem 5. Clearly $T_1,\dots,T_n\geq 1$ and $K>0$. Since $\kappa < C_\Psi^{-1}$, we have that $0<\delta'<1$. Finally, (44) is satisfied, since

$$\frac{\delta' K T_1\cdot\dots\cdot T_n}{\max_{i=1,\dots,n} T_i} = \frac{\kappa C_\Psi 2^{-\sum_{i=1}^n t_i}\sqrt{ndL2^{\max_i t_i+1}}\prod_i 2^{t_i+1}}{2^{\max_i t_i+1}}$$

$$= \frac{\kappa C_\Psi 2^n \sqrt{ndL}}{2^{(\max_i t_i+1)/2}} = \frac{\kappa C_\Psi 2^{n-1/2}\sqrt{ndL}}{2^{\|\mathbf{t}\|_\infty/2}} < 1\,, \tag{54}$$

where the last inequality follows from the definition of κ. Therefore, Theorem 5 is applicable and it follows that

$$|A^c_{\mathbf{t}}|_d \le |A(\delta', K, \mathbf{T})|_d \le E\left(\sqrt{n+d+1}\cdot\varepsilon_1\right)^{1/d(2l-1)} |\mathbf{U}|_d\,,$$

where E is given by (47) and where, from (54), the definition of δ' and the fact that $\kappa C_\Psi < 1$,

$$\varepsilon_1 = \max\left(\kappa C_\Psi 2^{-\sum_{i=1}^n t_i}, \left(\frac{\kappa C_\Psi 2^{n-1}\sqrt{ndL}}{2^{\|\mathbf{t}\|_\infty/2}}\right)^{\frac{1}{n+1}}\right) = \left(\frac{\kappa C_\Psi 2^{n-1}\sqrt{ndL}}{2^{\|\mathbf{t}\|_\infty/2}}\right)^{\frac{1}{n+1}}.$$

Then, using (53) and summing over all $\mathbf{t}\in\mathbb{N}_0^n$, we find that

$$\left|\bigcup_{\mathbf{q}\in\mathbb{Z}^n} A^c(\kappa;\mathbf{q})\right|_d \le \sum_{\mathbf{t}\in\mathbb{N}_0^n} E\left(\sqrt{n+d+1}\cdot\left(\frac{\kappa C_\Psi 2^{n-1}\sqrt{ndL}}{2^{\|\mathbf{t}\|_\infty/2}}\right)^{\frac{1}{n+1}}\right)^{1/d(2l-1)} |\mathbf{U}|_d$$

$$= K_0 S_n \kappa^{1/d(n+1)(2l-1)} |\mathbf{U}|_d \le \frac{\delta}{2}|\mathbf{U}|_d\,,$$

where the latter inequality follows from the definition of κ. This establishes the second inequality in (50) and thus completes the proof of Theorem 6 modulo Theorem 5.

6. Proof of Theorem 5

To establish Theorem 5, we will follow the basic strategy set out in the proof of [8, Theorem 1.4]. We stress that non-trivial modifications and additions are required to make the constants explicit. To begin with, we state a simplified form of [8, Theorem 6.2] and, to this end, various notions are now introduced.

Given a finite dimensional real vector space W, ν will denote a submultiplicative function on the exterior algebra $\bigwedge W$; that is, ν is a continuous function from $\bigwedge W$ to $\mathbb{R}_+$ such that

$$\nu(t\mathbf{w}) = |t|\,\nu(\mathbf{w}) \quad \text{and} \quad \nu(\mathbf{u}\wedge\mathbf{w}) \le \nu(\mathbf{u})\nu(\mathbf{w}) \tag{55}$$

for any $t\in\mathbb{R}$ and $\mathbf{u},\mathbf{w}\in\bigwedge W$. Given a discrete subgroup Λ of W of rank $k\ge 1$, let $\nu(\Lambda) := \nu(\mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_k)$, where $\mathbf{v}_1,\dots,\mathbf{v}_k$ is a basis of Λ (this definition makes sense from the first equation in (55)). Also, $\mathcal{L}(\Lambda)$ will denote the set of all non-zero primitive subgroups of Λ. Furthermore, given $C,\alpha>0$ and $V\subset\mathbb{R}^d$, a function $f:\mathbf{x}\in V\mapsto f(\mathbf{x})\in\mathbb{R}$ is said to be (C,α)-good on V if for any open ball $B\subset V$ and any $\epsilon>0$,

$$|\{\mathbf{x}\in B : |f(\mathbf{x})| < \epsilon \sup_{\mathbf{x}\in B}|f(\mathbf{x})|\}|_d \le C\epsilon^\alpha |B|_d.$$

Theorem 7. *([8, Theorem 6.2]) Let W be a $d+n+1$ dimensional real vector space and Λ be a discrete subgroup of W of rank k. Let a ball $B = B(\mathbf{x_0}, r) \subset \mathbb{R}^d$ and a map $H : \hat{B} \to \mathrm{GL}(W)$ be given, where $\hat{B} := 3^k B$. Take $C, \alpha > 0$, $0 < \tilde{\rho} < 1$ and let ν be a submultiplicative function on $\bigwedge W$. Assume that for any $\Gamma \in \mathcal{L}(\Lambda)$,*

(i) *the function $\mathbf{x} \to \nu(H(\mathbf{x})\Gamma)$ is (C, α)-good on $\hat{B}$,*
(ii) *there exists $\mathbf{x} \in B$ such that $\nu(H(\mathbf{x})\Gamma) \geq \tilde{\rho}$, and*
(iii) *for all $\mathbf{x} \in \hat{B}$, $\#\{\Gamma \in \mathcal{L}(\Lambda) \mid \nu(H(\mathbf{x})\Gamma) < \tilde{\rho}\} < \infty$.*

Then, for any positive $\varepsilon \leq \rho$, one has

$$|\{\mathbf{x} \in B : \exists\ \mathbf{v} \in \Lambda \setminus \{0\} \text{ such that } \nu(H(\mathbf{x})\mathbf{v}) < \varepsilon\}|_d \ \leq\ k(3^d N_d)^k C \left(\frac{\varepsilon}{\rho}\right)^{\alpha} |B|_d. \quad (56)$$

Theorem 5 is now deduced from Theorem 7 in the following manner. With respect to the parameters appearing in Theorem 7, we let

$$W = \mathbb{R}^{n+d+1}$$

and

$$\nu_* \text{ be the submultiplicative function introduced in [8, §7].}$$

There is nothing to gain in formally recalling the definition of ν_*. All we need to know is that ν_* as given in [8] has the property that

$$\nu_*(\mathbf{w}) \leq \|\mathbf{w}\|_2 \qquad \forall\ \ \mathbf{w} \in \bigwedge W \quad (57)$$

and that its restriction to W coincides with the Euclidean norm. Next, the discrete subgroup Λ appearing in Theorem 7 is defined as

$$\Lambda := \left\{ \begin{pmatrix} p \\ \mathbf{0} \\ \mathbf{q} \end{pmatrix} \in \mathbb{R}^{n+d+1} : p \in \mathbb{Z}, \mathbf{q} \in \mathbb{Z}^n \right\}. \quad (58)$$

Note that it has rank $k = n+1$, therefore the ball $\hat{B}$ appearing in the statement of Theorem 7 coincides with the ball $\tilde{\mathbf{U}}$ defined by (34). Finally, we let the map H send $\mathbf{x} \in \tilde{\mathbf{U}}$ to the product of matrices

$$H(\mathbf{x}) := DU_{\mathbf{x}}, \quad (59)$$

where

$$U_{\mathbf{x}} := \begin{pmatrix} 1 & 0 & \mathbf{f}(\mathbf{x}) \\ 0 & I_d & \nabla \mathbf{f}(\mathbf{x}) \\ 0 & 0 & I_n \end{pmatrix} \in \mathrm{SL}_{n+d+1}(\mathbb{R}) \quad (60)$$

and D is the diagonal matrix

$$D := \operatorname{diag}\Big(\frac{\varepsilon_1}{\delta'}, \underbrace{\frac{\varepsilon_1}{K}, \dots, \frac{\varepsilon_1}{K}}_{d \text{ times}}, \frac{\varepsilon_1}{T_1}, \dots, \frac{\varepsilon_1}{T_n}\Big) \tag{61}$$

defined via the constants δ', K, $T_1, \dots, T_n$, and ε_1 appearing in Theorem 5.

With the above choice of parameters, on using (57), it is easily verified that the set $A(\delta', K, \mathbf{T})$ defined by (45) within the context of Theorem 5 is contained in the set on the left-hand side of (56) with $\varepsilon := \varepsilon_1\sqrt{n+d+1}$. The upshot is that Theorem 5 follows from Theorem 7 on verifying conditions (i), (ii) and (iii) therein with appropriate constants C, α and ρ. With this in mind, we note that condition (iii) is already established in [8, §7] for any $\tilde{\rho} \le 1$. In §7 below, we will verify the remaining conditions (i) & (ii) with the following explicit constants:

$$C := \left(\frac{d(n+2)(d(n+1)+2)}{2}\right)^{\alpha/2} \max\left(C_1^*, 2C_{d,l}\right), \tag{62}$$

where

$$C_1^* := \max\left(\frac{2M}{s_0 \cdot \sigma(l,d)}, \frac{2^{d+2}}{V_d} \cdot dl(l+1) \cdot \frac{M}{s_0} \cdot \left(\frac{2l^l+1}{\sigma(l,d)}\right)^{1/l}\right) \tag{63}$$

(here, $\sigma(l,d)$ is the quantity defined in (39)),

$$C_{d,l} := \frac{2^{d+1} dl(l+1)^{1/l}}{V_d}, \tag{64}$$

$$\alpha := \frac{1}{d(2l-1)} \tag{65}$$

and $\tilde{\rho} = \rho$ as defined by (43) (note that $\rho < 1$). This will establish Theorem 5.

7. Verifying conditions (i) & (ii) of Theorem 7

Unless stated otherwise, throughout this section, Λ will be the discrete subgroup of $\mathbb{R}^{n+d+1}$ given by (58) and $\Gamma \in \mathcal{L}(\Lambda)$ will be a primitive subgroup of Λ. Verifying condition (i) of Theorem 7 is based on two separate cases: one when the rank of Γ is one and the other case of rank ≥ 2.

7.1. Rank one case of condition (i)

The key to verifying condition (i) in the case that Γ is of rank one is the following explicit version of [8, Proposition 3.4]. Notice that it and its corollary are themselves independent of rank and indeed Γ.

Proposition 1. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball, $\mathcal{F} \subset \mathcal{C}^{l+1}\left(\tilde{\mathbf{U}}^+\right)$ be a family of real valued functions and λ and γ be positive real numbers such that:*

(1) *the set $\{\nabla f : f \in \mathcal{F}\}$ is compact in $\mathcal{C}^{l-1}\left(\tilde{\mathbf{U}}^+\right)$,*
(2) $\sup\limits_{f\in\mathcal{F}} \sup\limits_{\mathbf{x}\in\tilde{\mathbf{U}}^+} |\partial_\beta f(\mathbf{x})| \leq \lambda$ *for any multi-index $\beta \in \mathbb{N}_0^d$ with $1 \leq |\beta| \leq l+1$,*
(3) $\inf\limits_{f\in\mathcal{F}} \sup\limits_{\mathbf{u}\in\mathbb{S}^{d-1}} \sup\limits_{1\leq k\leq l} \left|\frac{\partial^k f}{\partial \mathbf{u}^k}(\mathbf{x}_0)\right| \geq \gamma$, *where $\mathbf{x}_0$ is the centre of $\mathbf{U}$,*
(4) $\dfrac{\gamma\cdot\sigma(l,d)}{2\cdot 3^{n+d+2}\sqrt{d}\lambda} \geq r$, *where r is the radius of $\mathbf{U}$ as defined in (37) and $\sigma(l,d)$ is defined in (39).*

Then, for any $f \in \mathcal{F}$, we have that

(a) *f is $\left(C_1, \frac{1}{dl}\right)$-good on $\tilde{\mathbf{U}}$,*
(b) *$\|\nabla f\|_\infty$ is $\left(C_1, \frac{1}{d(l-1)}\right)$-good on $\tilde{\mathbf{U}}$,*

where

$$C_1 = C_1(\gamma,\lambda) := \max\left(\frac{2\lambda}{\gamma\cdot\sigma(l,d)}, \frac{2^{d+2}}{V_d}\, dl(l+1)\frac{\lambda}{\gamma}\left(\frac{2l^l+1}{\sigma(l,d)}\right)^{1/l}\right). \tag{66}$$

Remark. Hypothesis (2) is additional to those made in [8, Proposition 3.4]. In short, it is this "extra" hypothesis that yields an explicit formula for the constant C_1. Note that by the definition of C_1^* as given by (63), we have that

$$C_1^* = C_1(s_0, M)\,.$$

Using the explicit constant C_1 appearing in Proposition 1, it is possible to adapt the proof of [8, Corollary 3.5] to give the following statement.

Corollary 3. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball and $\mathbf{f} = (f_1, \ldots, f_n)$ be an n-tuple of C^{l+1} functions satisfying Assumption 1. With reference to Proposition 1, let*

$$\gamma := s_0 \qquad \text{and} \qquad \lambda := M.$$

Then, for any linear combination $f = c_0 + \sum_{i=1}^n c_i f_i$ with $c_0, \ldots, c_n \in \mathbb{R}$, we have that

(a) *f is $\left(C_1^*, \frac{1}{dl}\right)$-good on $\tilde{\mathbf{U}}$,*
(b) *$\|\nabla f\|_\infty$ is $\left(C_1^*, \frac{1}{d(l-1)}\right)$-good on $\tilde{\mathbf{U}}$.*

Corollary 3 allows us to verify condition (i) of Theorem 7 in the case that Γ is a primitive subgroup of Λ of rank 1. Indeed, in view of (59) and of the discussions following

equations (55) and (57), $\nu_*(H(\mathbf{x})\Gamma)$ is the Euclidean norm of $H(\mathbf{x})\mathbf{w} = DU_{\mathbf{x}}\mathbf{w}$, where $\mathbf{w}$ is a basis vector of Γ. It is readily seen that the coordinate functions of $H(\mathbf{x})\mathbf{w}$ are either constants, or $f(\mathbf{x})$, or $\partial f(\mathbf{x})/\partial x_i$ for some $f = c_0 + \sum_{i=1}^n c_i f_i$ with $c_0, \dots, c_n \in \mathbb{R}$. Hence, by Corollary 3 and [8, Lemma 3.1 (b,d)] we obtain that the function $\|H(\cdot)\Gamma\|_\infty$ is (C_1^*, α)-good on $\tilde{\mathbf{U}}$, where α is given by (65). In turn, on using [8, Lemma 3.1(c)] and the fact that

$$\frac{1}{\sqrt{n+d+1}} \leq \frac{\|H(\mathbf{x})\mathbf{w}\|_\infty}{\|H(\mathbf{x})\mathbf{w}\|_2} \leq 1,$$

we have that $\nu_*(H(\,.\,)\Gamma)$ is $\left(C_1^*(n+d+1)^{\alpha/2}, \alpha\right)$-good on $\tilde{\mathbf{U}}$. It then follows from [8, Lemma 3.1(d)] that

$$\nu_*(H(\,.\,)\Gamma) \text{ is } (C, \alpha)\text{-good on } \tilde{\mathbf{U}}.$$

Proof of Corollary 3. In view of [8, Lemma 3.1.a], it suffices to prove the corollary under the assumption that $\|(c_1, \dots, c_n)\|_2 = 1$. Thus, with reference to Proposition 1, define

$$\mathcal{F} := \left\{ c_0 + \sum_{i=1}^n c_i f_i \,:\, \|(c_1, \dots, c_n)\|_2 = 1 \right\}.$$

The corollary will follow on verifying the four hypotheses of Proposition 1. Thus, hypothesis (1) is easily seen to be satisfied. Hypothesis (2) is a consequence of (36) and of the Cauchy–Schwarz inequality while hypothesis (3) follows straightforwardly from the definition of the measure of non-degeneracy s_0 in (33) and (35). Finally, hypothesis (4) is guaranteed by (37) and the choices of γ and λ. □

7.2. Proof of Proposition 1

The proof of Proposition 1 relies on the following lemma:

Lemma 3. *Let f be a real-valued function of class C^k $(k \geq 1)$ defined in a neighbourhood of $\mathbf{x} \in \mathbb{R}^d$ $(d \geq 1)$. Assume that there exists an index $1 \leq i_0 \leq d$ and a real number $\mu > 0$ such that*

$$\left| \frac{\partial^k f}{\partial x_{i_0}^k}(\mathbf{x}) \right| \geq \mu.$$

Then there exists a rotation $S : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\left| \frac{\partial^k (f \circ S)}{\partial x_i^k}(\mathbf{x}) \right| \geq \mu \cdot \sigma(k, d)$$

for all indices $1 \leq i \leq d$, where the quantity $\sigma(k, d)$ is defined in (39).

As the proof of Lemma 3 is lengthy, before giving it, we show how to deduce Proposition 1 from it.

Deduction of Proposition 1 from Lemma 3. Let $\mathbf{x}_0 = (v_1, v_2, \dots, v_d)$ denote the centre of $\mathbf{U}$. Hypothesis (3) of Proposition 1 implies that for any $f \in \mathcal{F}$, there exists a unit vector $\mathbf{u} \in \mathbb{S}^{d-1}$ and an index $1 \leq k \leq l$ such that

$$\left| \frac{\partial^k f}{\partial \mathbf{u}^k}(\mathbf{x}_0) \right| \geq \gamma.$$

Even if it means applying a first rotation to the coordinate system that brings the x_1 axis onto the line spanned by the vector $\mathbf{u}$, it may be assumed, without loss of generality, that the above inequality reads as

$$\left| \partial_1^k f(\mathbf{x}_0) \right| \geq \gamma.$$

From Lemma 3, up to another rotation of the coordinate system, one can guarantee that

$$\left| \partial_i^k f(\mathbf{x}_0) \right| \geq \gamma \cdot \sigma(k,d) \ := \ C_2$$

for all indices $1 \leq i \leq d$. Now, for a fixed index i, it follows from a Taylor expansion at $\mathbf{x}_0$ that, for any $\mathbf{x} = (x_1, \dots, x_d) \in \tilde{\mathbf{U}}^+$,

$$\partial_i^k f\left(\mathbf{x}\right) = \partial_i^k f\left(\mathbf{x}_0\right) + \sum_{j=1}^{d} R_j(\mathbf{x}; \mathbf{x}_0) \left(x_j - v_j\right),$$

where, by hypothesis (2), $R_j(\mathbf{x}; \mathbf{x}_0)$ satisfies the inequality

$$|R_j(\mathbf{x}; \mathbf{x}_0)| \ \leq \ \sup_{\mathbf{x} \in \tilde{\mathbf{U}}^+} \left| \left(\partial_j \circ \partial_i^k\right) f(\mathbf{x}) \right| \leq \lambda.$$

In view of hypothesis (4), we have furthermore that

$$\|\mathbf{x} - \mathbf{x}_0\|_2 \leq 3^{n+d+2} r \leq \frac{\gamma \cdot \sigma(l,d)}{2\sqrt{d}\lambda} \leq \frac{C_2}{2\sqrt{d}\lambda}.$$

Thus, for all indices $1 \leq i \leq d$,

$$\begin{aligned} \left| \partial_i^k f\left(\mathbf{x}\right) \right| \ &\geq \ C_2 - \sum_{i=1}^{d} |x_j - u_j| \, \lambda \ = \ C_2 - \lambda \, \|\mathbf{x} - \mathbf{x}_0\|_1 \\ &\geq \ C_2 - \lambda \sqrt{d} \, \|\mathbf{x} - \mathbf{x}_0\|_2 \ \geq \ \frac{C_2}{2}. \end{aligned} \tag{67}$$

Next, observe that any cube circumscribed about $\tilde{\mathbf{U}}$ lies inside of $\tilde{\mathbf{U}}^+$. It then follows on applying [8, Lemma 3.3] with $A_1 = \lambda$ and $A_2 = C_2/2$ that the function f is $\left(C', \frac{1}{dk}\right)$-good on $\tilde{\mathbf{U}}$, where

$$C' := \frac{2^d}{V_d} dk(k+1) \left(\frac{2\lambda}{\gamma \cdot \sigma(k,d)} (k+1) \left(2k^k + 1\right) \right)^{1/k}$$

$$\leq \frac{2^{d+2}}{V_d} dk(k+1) \frac{\lambda}{\gamma} \left(\frac{2k^k + 1}{\sigma(k,d)} \right)^{1/k}.$$

A computation then shows that

$$C' \leq \frac{2^{d+2}}{V_d} dl(l+1) \frac{\lambda}{\gamma} \left(\frac{2l^l + 1}{\sigma(l,d)} \right)^{1/l}.$$

Part (a) of Proposition 1 is now a consequence of [8, Lemma 3.1.d]. Regarding part (b), the proof is essentially the same as that of [8, Proposition 3.4.b] with the constant C replaced with the explicit constant C_1 given by (66). □

We now proceed with the proof of Lemma 3 which requires several intermediate results. The first one is rather intuitive.

Lemma 4. *Let $C > 0$ be a real number and $p \geq 1$ be an integer. Then every section of the cube $[0, C]^p$ with a $(p-1)$-dimensional subspace of $\mathbb{R}^p$ has a volume at most $\sqrt{2}C^{p-1}$.*

Proof. See [1, Theorem 4]. □

Lemma 5. *Let $k \geq 1$ denote an integer and let $\mathbf{w} := (w_0, \ldots, w_k) \in \mathbb{R}^{k+1}$. Let $\omega, B > 0$ be real numbers. Furthermore, assume that the $k+1$ real numbers $0 < t_0 < \cdots < t_k$ satisfy the following two assumptions:*

(1) $\min\limits_{0 \leq i \neq j \leq k} |t_i - t_j| \geq \omega$,
(2) $\max\limits_{0 \leq i \leq k} |t_i|^k \leq B$.

Then, there exists an index $0 \leq j \leq k$ such that

$$\left| \sum_{i=0}^{k} w_i t_j^i \right| \geq \phi(\omega, B; k) \cdot \|\mathbf{w}\|_2,$$

where $\phi(\omega, B; k)$ is the quantity defined in (40).

The following notation will be used in the course of the proof of Lemma 5: given a point $\mathbf{x} \in \mathbb{R}^n$ and a set $A \subset \mathbb{R}^n$, $\operatorname{dist}(\mathbf{x}, A)$ will denote the quantity

$$\operatorname{dist}(\mathbf{x}, A) := \inf\{\|\mathbf{x} - \mathbf{a}\|_2 : \mathbf{a} \in A\}. \tag{68}$$

Proof. Let $\mathbf{X} := (\mathbf{x}_1, \cdots, \mathbf{x}_{k+1}) \in M_{k+1,k+1}$ denote the matrix defined by the following $k+1$ column vectors in $\mathbb{R}^{k+1}$:

$$\begin{aligned} \mathbf{x}_1 &:= (1, t_0, \ldots, t_0^k)^T, \\ &\vdots \\ \mathbf{x}_{k+1} &:= (1, t_k, \ldots, t_k^k)^T. \end{aligned}$$

Together with the origin, these points form a simplex $\mathcal{S}(\mathbf{X})$ in $\mathbb{R}^{k+1}$ whose volume $|\mathcal{S}(\mathbf{X})|_{k+1}$ satisfies the well-known equation

$$|\mathcal{S}(\mathbf{X})|_{k+1} = \frac{1}{(k+1)!} \det \begin{pmatrix} \mathbf{x}_1 & \ldots & \mathbf{x}_{k+1} & \mathbf{0} \\ 1 & \ldots & 1 & 1 \end{pmatrix}.$$

The formula for the determinant of a Vandermonde matrix together with hypothesis (1) then yields the inequality

$$|\mathcal{S}(\mathbf{X})|_{k+1} \geq \frac{\omega^{k(k-1)/2}}{(k+1)!}.$$

Note that hypothesis (2) implies that all the vectors $\mathbf{x}_1, \cdots, \mathbf{x}_{k+1}$ lie in the hypercube $\mathcal{B} := [0, B]^{k+1}$. As a consequence, the volume of the section of the simplex $\mathcal{S}(\mathbf{X})$ with any hyperplane does not exceed the volume of the section of $\mathcal{S}(\mathbf{X})$ with $\mathcal{B}$ which, from Lemma 4, is at most $\sqrt{2}B^k$. Also, given a hyperplane $\mathcal{P}$, it should be clear that

$$|\mathcal{S}(\mathbf{X})|_{k+1} \leq 2 \cdot \max_{1 \leq j \leq k+1} \operatorname{dist}(\mathbf{x}_j, \mathcal{P}) \cdot |\mathcal{P} \cap \mathcal{S}(\mathbf{X})|_k .$$

The upshot of this discussion is that the following inequality holds:

$$\max_{1 \leq j \leq k} \operatorname{dist}(\mathbf{x}_j, \mathcal{P}) \geq \frac{\omega^{k(k-1)/2}}{2\sqrt{2}B^k(k+1)!} := \phi(\omega, B, k). \tag{69}$$

Consider now the hyperplane $\mathcal{P} = \mathbf{w}^{\perp}$ and let j be one of the indices realising the maximum in (69). The conclusion of the lemma is then a direct consequence of the equation

$$\operatorname{dist}(\mathbf{x}_j, \mathbf{w}^{\perp}) = \frac{\left|\sum_{i=0}^k w_i t_j^i\right|}{\|\mathbf{w}\|_2}. \quad \square$$

The next result contains the main substance of the proof of Lemma 3.

Lemma 6. *Let f be a real valued function of class C^k $(k \geq 1)$ defined in a neighbourhood of $(x_0, y_0) \in \mathbb{R}^2$. Let $c > 0$ be a real number such that*

$$\left| \frac{\partial^k f}{\partial x^k}(x_0, y_0) \right| \geq c.$$

Then, there exist two orthonormal vectors $\mathbf{u}, \mathbf{v} \in \mathbb{S}^1$ such that

$$\min\left\{ \left| \frac{\partial^k f}{\partial \mathbf{u}^k}(x_0, y_0) \right|, \left| \frac{\partial^k f}{\partial \mathbf{v}^k}(x_0, y_0) \right| \right\} \geq \frac{c}{2^{3k/2}} \cdot \phi\left(\left(\sqrt{2}(2k)^{2+k(k-1)/2}(k+1)! \right)^{-1}, 2, k \right)$$
$$= c \cdot \sigma(k, 2).$$

Proof. Set

$$\mathbf{w} = (w_0, \ldots, w_k) := \left(\binom{k}{j} \frac{\partial^k f}{\partial x^{k-j} \partial y^j}(x_0, y_0) \right)_{0 \leq j \leq k} \in \mathbb{R}^{k+1}.$$

It readily follows from the assumptions of the lemma that

$$\|\mathbf{w}\|_2 \geq c.$$

Let $\lambda > 0$ be a real number such that, for all indices $0 \leq j \leq k$,

$$\left| \frac{\partial^k f}{\partial x^{k-j} \partial y^j}(x_0, y_0) \right| \leq \lambda.$$

We thus have the inequality

$$\|\mathbf{w}\|_2 \leq 2^k (k+1)\lambda. \tag{70}$$

Define now $k+1$ real numbers $t_0, \ldots, t_k$ as follows:

$$t_i := \frac{1}{2} + \frac{i}{2k},$$

where $i = 0, \ldots, k$.

With the choices of the parameters $\omega := 1/(2k)$ and $B = 1$, Lemma 5 applied to the vector $\mathbf{w}$ and to the system of points $(t_i)_{0 \leq i \leq k}$ yields the existence of a point $t_j \in [1/2, 1]$ such that

$$\left| \sum_{i=0}^{k} w_i t_j^i \right| \geq \|\mathbf{w}\|_2 \cdot \phi\left(\frac{1}{2k}, 1, k \right). \tag{71}$$

Let

$$\epsilon := \frac{1}{2k} \cdot \phi\left(\frac{1}{2k}, 1, k\right) \leq \frac{1}{2k} \tag{72}$$

denote a constant and

$$g : t \in [0,1] \mapsto \sum_{i=0}^{k} w_i t^i$$

a function. Note that for all $t \in [0,1]$,

$$|g'(t)| = \left|\sum_{i=1}^{k} w_i i t^{i-1}\right| \leq 2^k \lambda \cdot \frac{k(k+1)}{2} = 2^{k-1}\lambda k(k+1).$$

This implies that for all $t \in [t_j - \epsilon, t_j + \epsilon]$, where t_j is the constant appearing in (71), the following inequalities hold:

$$\begin{aligned}
|g(t)| = \left|\sum_{i=0}^{k} w_i t^i\right| &\geq |g(t_j)| - |g(t) - g(t_j)| \\
&\geq |g(t_j)| - \epsilon \cdot 2^{k-1}\lambda k(k+1) \\
&\underset{(71)\&(72)}{\geq} \phi\left(\frac{1}{2k}, 1, k\right) \cdot \left(\|\mathbf{w}\|_2 - 2^{k-2}\lambda(k+1)\right) \\
&\underset{(70)}{\geq} \frac{\|\mathbf{w}\|_2}{2} \cdot \phi\left(\frac{1}{2k}, 1, k\right).
\end{aligned}$$

Consider now the image $[a,b] \subset [1,2]$ of the interval $[t_j - \epsilon, t_j + \epsilon] \cap [1/2, 1]$ under the map $t \mapsto 1/t$. It is then readily verified that

$$|b - a| \geq \epsilon.$$

With the choices of the parameters

$$\omega := \frac{\epsilon}{k} = \frac{1}{2k^2} \cdot \phi\left(\frac{1}{2k}, 1, k\right)$$

and $B = 2$, apply once more Lemma 5, this time to the vector $\left((-1)^i w_i\right)_{0 \leq i \leq k}$ and to the set of points

$$\tilde{t}_i = a + \frac{b-a}{k} \cdot i, \qquad 0 \leq i \leq k.$$

This yields the existence of $\tilde{t}_j \in [a,b]$ such that

$$\left|\sum_{i=0}^{k}(-1)^i w_i \tilde{t}_j^i\right| \geq \|\mathbf{w}\|_2 \cdot \phi\left(\frac{\epsilon}{k}, 2, k\right).$$

The upshot of this is that, when considering the point $s := 1/\tilde{t}_j$, the following two inequalities hold simultaneously:

$$\left|\sum_{i=0}^{k} w_i s^i\right| \geq \frac{\|\mathbf{w}\|_2}{2} \cdot \phi\left(\frac{1}{2k}, 1, k\right)$$

and

$$\left|\sum_{i=0}^{k}(-1)^i w_i s^{-i}\right| \geq \|\mathbf{w}\|_2 \cdot \phi\left(\frac{1}{2k^2} \cdot \phi\left(\frac{1}{2k}, 1, k\right), 2, k\right).$$

Since $s \in [1/2, 1]$, it is easily seen that one can find a unit vector $(u_1, u_2) \in \mathbb{S}^1$ such that $s = u_2/u_1$ and $u_1, u_2 \in [1/(2\sqrt{2}), 1]$. Let $\mathbf{u} \in \mathbb{S}^1$ and $\mathbf{v} \in \mathbb{S}^1$ denote the two orthonormal vectors defined as $\mathbf{u} := (u_1, u_2)$ and $\mathbf{v} := (u_2, -u_1)$.

Note then that

$$\begin{aligned}\left|\sum_{i=0}^{k} w_i u_1^{k-i} u_2^i\right| = u_1^k \left|\sum_{i=0}^{k} w_i s^i\right| &\geq \frac{1}{2^{1+3k/2}} \cdot \|\mathbf{w}\|_2 \cdot \phi\left(\frac{1}{2k}, 1, k\right) \\ &\geq \frac{c}{2^{1+3k/2}} \cdot \phi\left(\frac{1}{2k}, 1, k\right)\end{aligned}$$

and, similarly,

$$\left|\sum_{i=0}^{k}(-1)^i w_i u_2^{k-i} u_1^i\right| = u_2^k \left|\sum_{i=0}^{k}(-1)^i w_i s^{-i}\right| \geq \frac{c}{2^{3k/2}} \cdot \phi\left(\frac{1}{2k^2} \cdot \phi\left(\frac{1}{2k}, 1, k\right), 2, k\right).$$

Since, from the definition of the vector $\mathbf{w}$,

$$\frac{\partial^k f}{\partial \mathbf{u}^k}(x_0, y_0) = \sum_{i=0}^{k} w_i u_1^{k-i} u_2^i$$

and

$$\frac{\partial^k f}{\partial \mathbf{v}^k}(x_0, y_0) = \sum_{i=0}^{k}(-1)^i w_i u_2^{k-i} u_1^i,$$

this completes the proof of the lemma from the definition of ϕ in (40). □

We now have all the ingredients at our disposal to prove Lemma 3.

Proof of Lemma 3. Denote the coordinates of the vector $\mathbf{x} \in \mathbb{R}^d$ as $\mathbf{x} = (x_1, \dots, x_d)$. Even if it means relabelling the axes, assume furthermore without loss of generality that $i_0 = 1$ in the statement of the lemma. The proof then goes by induction on $d \geq 1$, the conclusion being trivial when $d = 1$. When $d = 2$, Lemma 3 reduces to Lemma 6. Assume therefore that $d \geq 3$. It then readily follows from the induction hypothesis applied to the function $(x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} \mapsto f(x_1, \dots, x_{d-1}, x_d)$ that there exists a rotation $S_1 : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\left| \frac{\partial^k (f \circ S_1)}{\partial x_i^k}(\mathbf{x}) \right| \geq \mu \cdot \sigma(k, d-1) \qquad (1 \leq i \leq d-1).$$

Consider now the function $(x_1, x_d) \in \mathbb{R}^2 \mapsto f(x_1, \dots, x_{d-1}, x_d)$. Applying Lemma 6 to this function with $c = \mu \cdot \sigma(k, d-1)$ therein provides the existence of a rotation $S_2 : \mathbb{R}^d \mapsto \mathbb{R}^d$ acting on the plane (x_1, x_d) and leaving its orthogonal unchanged such that

$$\min \left\{ \frac{\partial (f \circ S_1 \circ S_2)}{\partial x_1^k}(\mathbf{x}), \frac{\partial (f \circ S_1 \circ S_2)}{\partial x_d^k}(\mathbf{x}) \right\} \geq \mu \cdot \sigma(k, d-1) \cdot \sigma(k, 2) = \mu \cdot \sigma(k, d).$$

The lemma follows upon setting $S = S_1 \circ S_2$. □

7.3. Higher rank case of condition (i)

The key to verifying condition (i) of Theorem 7 in the case when Γ is of rank greater than one is Proposition 2 below. In short, it is an explicit version of [8, Proposition 4.1] in the particular case when the set $\mathcal{G}$ appearing therein is given by

$$\mathcal{G} := \left\{ (\mathbf{u}_1 \cdot \mathbf{f}, \mathbf{u}_2 \cdot \mathbf{f} + u_0) \ : \ u_0 \in \mathbb{R}, \, \mathbf{u}_1, \mathbf{u}_2 \in \mathbb{S}^{n-1}, \, \mathbf{u}_1 \perp \mathbf{u}_2 \right\}. \tag{73}$$

The statement is concerned with the *skew gradient* of a map as defined in [8, §4]. We recall the definition. Let $\mathbf{g} = (g_1, g_2) : \tilde{\mathbf{U}}^+ \to \mathbb{R}^2$ be a differentiable function. The skew gradient $\widetilde{\nabla}\mathbf{g} : \tilde{\mathbf{U}}^+ \to \mathbb{R}^2$ is defined by

$$\widetilde{\nabla}\mathbf{g}(\mathbf{x}) := g_1(\mathbf{x}) \nabla g_2(\mathbf{x}) - g_2(\mathbf{x}) \nabla g_1(\mathbf{x}).$$

If we write $\mathbf{g}(\mathbf{x})$ in terms of polar coordinates; i.e. via the usual functions $\rho(\mathbf{x})$ and $\theta(\mathbf{x})$, it is then readily verified that

$$\widetilde{\nabla}\mathbf{g}(\mathbf{x}) = \rho^2(\mathbf{x}) \nabla \theta(\mathbf{x}). \tag{74}$$

Essentially, the skew gradient measures how different the pair of functions g_1 and g_2 are from being proportional to each other.

Proposition 2. *Let* $\mathbf{U} \subset \mathbb{R}^d$ *be a ball and* $\mathbf{f} = (f_1, \dots, f_n)$ *be an* n*-tuple of* C^{l+1} *functions satisfying Assumption 1. Let* ρ_2, $C_{d,l}$ *and* $\mathcal{G}$ *be given by* (42), (64) *and* (73) *respectively. Then,*

(a) *for all* $\mathbf{g} \in \mathcal{G}$,

$$\|\tilde{\nabla}\mathbf{g}\|_2 \quad \textit{is} \quad \left(2C_{d,l}, \frac{1}{d(2l-1)}\right) \textit{-good} \quad \textit{on } \tilde{\mathbf{U}}$$

(b) *for all* $\mathbf{g} \in \mathcal{G}$,

$$\sup_{\mathbf{x}\in\mathbf{U}} \|\tilde{\nabla}\mathbf{g}(\mathbf{x})\|_2 \geq \rho_2 \,. \tag{75}$$

This proposition together with Corollary 3 and the basic properties of (C, α)-good functions given in [8, Lemma 3.1] enables us to deduce the following statement, which establishes condition (i) in the higher rank case.

Corollary 4. *Let* $\mathbf{U} \subset \mathbb{R}^d$ *be a ball and* $\mathbf{f} = (f_1, \dots, f_n)$ *be an* n*-tuple of* C^{l+1} *functions satisfying Assumption 1. Let* Λ *be the discrete subgroup given by* (58) *and* $\Gamma \in \mathcal{L}(\Lambda)$ *be a primitive subgroup of* Λ*. Furthermore, let* H *be the map given by* (59)*. Then, the function*

$$\mathbf{x} \mapsto \nu_* \left(H(\mathbf{x})\Gamma\right) \tag{76}$$

is (C, α)*-good on the ball* $\tilde{\mathbf{U}}$ *with constants* C *and* α *given by* (62) *and* (65) *respectively.*

Proof. Let k denote the rank of Γ. The case $k = 1$ has already been established as a consequence of Corollary 3 in §7.1. Assume therefore that $k \geq 2$. It is shown in [8, §7, Eq. (7.3)] that there exist real numbers $a, b, \mu \in \mathbb{R}$ such that, for all $\mathbf{x} \in \tilde{\mathbf{U}}$, $\nu_* \left(H(\mathbf{x})\Gamma\right)$ given by (76) can be expressed as the Euclidean norm of a vector $\mathbf{w}(\mathbf{x})$. Furthermore, there exists an orthonormal system of vectors of the form $\mathcal{S} = \{\mathbf{e}_0, \mathbf{e}_1^*, \dots, \mathbf{e}_d^*, \mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ when $k \leq n$ or of the form $\mathcal{S} = \{\mathbf{e}_0, \mathbf{e}_1^*, \dots, \mathbf{e}_d^*, \mathbf{v}_0, \dots, \mathbf{v}_{k-1}\}$ when $k = n + 1$ such that $\mathbf{w}(\mathbf{x})$ is a linear combination of

$$L_d(k) := \frac{(k+1)(dk+2)}{2}$$

skew products of elements of $\mathcal{S}$ whose coefficients are of any of the following form:

$$a + b\mathbf{f} \cdot \mathbf{v}_0 \tag{77}$$

$$b \tag{78}$$

$$b\,\mathbf{f} \cdot \mathbf{v}_i \qquad (1 \leq i \leq k-1) \tag{79}$$

$$b\,\mu\,\partial_s(\mathbf{f} \cdot \mathbf{v}_i) \qquad (1 \leq i \leq k,\ 1 \leq s \leq d) \tag{80}$$

$$\mu\, X(i,s) \qquad (1 \le i \le k-1,\ 1 \le s \le d) \tag{81}$$

$$b\,\mu\, Y(i,j,s) \qquad (1 \le i < j \le k-1,\ 1 \le s \le d), \tag{82}$$

where

$$X(i,s) := (\mathbf{f}\cdot\mathbf{v}_i)\,\partial_s\,(a + b\mathbf{f}\cdot\mathbf{v}_0) - (a + b\mathbf{f}\cdot\mathbf{v}_0)\,\partial_s\,(\mathbf{f}\cdot\mathbf{v}_i)$$

and

$$Y(i,j,s) := (\mathbf{f}\cdot\mathbf{v}_i)\,\partial_s\,(\mathbf{f}\cdot\mathbf{v}_j) - (\mathbf{f}\cdot\mathbf{v}_j)\,\partial_s\,(\mathbf{f}\cdot\mathbf{v}_i)\,.$$

- It follows from part (a) of Corollary 3 and [8, Lemma 3.1(a,d)] that the coordinate functions given by (77), (78) and (79) are (C',α)-good, where

$$C' := \frac{C}{(L_d(n+1)\cdot d)^{\alpha/2}}.$$

- It follows from part (b) of Corollary 3 and [8, Lemma 3.1(a,d)] that, when the index i is fixed, the maximum over s of the coordinate functions given by (80), that is, the quantity $\|b\,\mu\,\nabla(\mathbf{f}\cdot\mathbf{v}_i)\|_\infty$, is (C',α)-good.
- It follows from Proposition 2 and [8, Lemma 3.1(a,d)] that, for fixed indices i and j, the Euclidean norm over s of the coordinate functions given by (81) and (82), that is, the quantities $\|\mu\,\tilde{\nabla}(\mathbf{f}\cdot\mathbf{v}_i, a + b\mathbf{f}\cdot\mathbf{v}_0)\|_2$ and $\|b\,\mu\,\tilde{\nabla}(\mathbf{f}\cdot\mathbf{v}_i, \mathbf{f}\cdot\mathbf{v}_j)\|_2$ respectively, are (C',α)-good. On using the relation

$$\frac{1}{\sqrt{d}} \le \frac{\|\cdot\|_\infty}{\|\cdot\|_2} \le 1$$

valid in $\mathbb{R}^d$ and [8, Lemma 3.1(c)], it follows that $\|\mu\,\tilde{\nabla}(\mathbf{f}\cdot\mathbf{v}_i, a + b\mathbf{f}\cdot\mathbf{v}_0)\|_\infty$ and $\|b\,\mu\,\tilde{\nabla}(\mathbf{f}\cdot\mathbf{v}_i, \mathbf{f}\cdot\mathbf{v}_j)\|_\infty$ are $(d^{\alpha/2}C',\alpha)$-good.

The upshot of the above together with [8, Lemma 3.1(b)] is that the maximum of the coordinate functions (77)–(82) is $(d^{\alpha/2}C',\alpha)$-good. In turn, on using the relation

$$\frac{1}{\sqrt{L_d(k)}} \le \frac{\|\cdot\|_\infty}{\|\cdot\|_2} \le 1$$

valid in $\mathbb{R}^{L_d(k)}$ and [8, Lemma 3.1(c)], we have that

$$\nu_*\,(H(\,.\,)\Gamma) \quad \text{is} \quad \Big(C'(d\cdot L_d(k))^{\alpha/2},\alpha\Big)\text{-good.}$$

As $k \le n+1$, the desired statement follows. □

Modulo the proof of Proposition 2, we have completed the task of verifying condition (i) of Theorem 7. The proof of the proposition is rather lengthy and therefore is postponed till after we have verified condition (ii) of Theorem 7.

7.4. Verifying condition (ii) of Theorem 7 modulo Proposition 2

The following lemma, which although not explicitly stated, is essentially proved in [8, §7], see [8, Eq. (7.5)] and onwards. The key difference is that we make use of Proposition 2 in place of [8, Proposition 4.1] and so are able to give explicit values of ρ_1 and ρ.

Lemma 7. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball and $\mathbf{f} = (f_1, \dots, f_n)$ be an n-tuple of C^{l+1} functions satisfying Assumption 1. Let $\rho_1, \rho > 0$ be given by (41) and (43) respectively and assume that for any $\mathbf{v} \in \mathbb{S}^{n-1}$ and $p \in \mathbb{R}$ we have that*

$$\sup_{\mathbf{x}\in\mathbf{U}} |\mathbf{f}(\mathbf{x})\cdot\mathbf{v} + p| \geq \rho_1 \quad \text{and} \quad \sup_{\mathbf{x}\in\mathbf{U}} \|\nabla\left(\mathbf{f}(\mathbf{x})\cdot\mathbf{v}\right)\|_\infty \geq \rho_1. \tag{83}$$

Furthermore, let Λ be the discrete subgroup given by (58), $\Gamma \in \mathcal{L}(\Lambda)$ be a primitive subgroup of Λ and H be the map given by (59). Then

$$\sup_{\mathbf{x}\in\mathbf{U}} \nu_*(H(\mathbf{x})\Gamma) \geq \rho.$$

The following statement immediately verifies condition (ii) of Theorem 7. It is the above lemma without the assumptions made in (83).

Corollary 5. *Let $\mathbf{U} \subset \mathbb{R}^d$ be a ball and $\mathbf{f} = (f_1, \dots, f_n)$ be an n-tuple of C^{l+1} functions satisfying Assumption 1. Let Λ be the discrete subgroup given by (58) and $\Gamma \in \mathcal{L}(\Lambda)$ be a primitive subgroup of Λ. Furthermore, let H be the map given by (59). Then*

$$\sup_{\mathbf{x}\in\mathbf{U}} \nu_*(H(\mathbf{x})\Gamma) \geq \rho, \tag{84}$$

where ρ is given by (43).

Proof of Corollary 5. The desired statement follows directly from Lemma 7 on verifying the inequalities associated with (83). Let $\mathbf{v} \in \mathbb{S}^{n-1}$. By the definition of $s_0 := s(l; \mathbf{f}, \tilde{\mathbf{U}}^+)$, there exists a $\mathbf{u} \in \mathbb{S}^{d-1}$ and $1 \leq k \leq l$ such that

$$\left|\frac{\partial^k(\mathbf{f}\cdot\mathbf{v})}{\partial\mathbf{u}^k}(\mathbf{x}_0)\right| \geq s_0. \tag{85}$$

Recall, that $\mathbf{x}_0$ is the centre of $\mathbf{U}$. It follows that for any $\mathbf{x} \in \mathbf{U}$, we have that:

$$\left|\mathbf{v}\cdot\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x})\right| = \left|\mathbf{v}\cdot\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x}_0)\right| - \left|\mathbf{v}\cdot\left(\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x}_0) - \frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x})\right)\right|$$

$$\geq s_0 - \left\|\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x}_0) - \frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x})\right\|_2. \tag{86}$$

Let $\mathbf{s}'$ denote the unit vector

$$\mathbf{s}' := \frac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|_2}.$$

By Lagrange's Theorem, there exists $\mathbf{x}'$ between $\mathbf{x}_0$ and $\mathbf{x}$ such that

$$\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x}_0) = \frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x}) + \|\mathbf{x}-\mathbf{x}_0\|_2\,\frac{\partial}{\partial\mathbf{s}'}\left(\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}\right)(\mathbf{x}').$$

It then follows from (86) and the definition of M in (36) that

$$\left|\mathbf{v}\cdot\frac{\partial^k\mathbf{f}}{\partial\mathbf{u}^k}(\mathbf{x})\right| \geq s_0 - M\,\|\mathbf{x}-\mathbf{x}_0\|_2\,.$$

This together with the fact that $r < s_0/2M$ — a direct consequence of (37) —, implies that

$$\left|\frac{\partial^k(\mathbf{f}\cdot\mathbf{v})}{\partial\mathbf{u}^k}(\mathbf{x})\right| \geq s_0/2 \qquad \forall\ \mathbf{x}\in\mathbf{U}\,. \tag{87}$$

The upshot is that the hypotheses of [8, Lemma 3.6] are satisfied. A straightforward application of [8, Lemma 3.6] together with (41) implies that

$$\sup_{x\in\mathbf{U}}|\mathbf{f}(\mathbf{x})\cdot\mathbf{v}+p| \geq \frac{s_0/2}{2k^k(k+1)!}(2r)^k \geq \frac{s_0}{4l^l(l+1)!}(2r)^l = \rho_1\sqrt{d} \geq \rho_1$$

for any $p\in\mathbb{R}$. Thus the first inequality appearing in (83) is established.

It remains to prove the second inequality in (83); that is, that for any $\mathbf{v}\in\mathbb{S}^{n-1}$,

$$\sup_{\mathbf{x}\in\mathbf{U}}\|\nabla\,(\mathbf{f}(\mathbf{x})\cdot\mathbf{v})\|_\infty \geq \rho_1 = \frac{s_0}{4l^l(l+1)!\sqrt{d}}(2r)^l. \tag{88}$$

Recall from above that for any $\mathbf{v}\in\mathbb{S}^{n-1}$ we can find a vector $\mathbf{u}\in\mathbb{S}^{d-1}$ such that (85) and (87) hold. Furthermore, observe that

$$\frac{\partial(\mathbf{f}\cdot\mathbf{v})}{\partial\mathbf{u}}(\mathbf{x}) = \mathbf{u}^t\cdot\nabla\,(\mathbf{f}(\mathbf{x})\cdot\mathbf{v})\,. \tag{89}$$

We proceed by considering two cases, depending on whether or not $k=1$ in (85).

- Suppose $k = 1$ in (85). Then it follows from (87) and (89) that

$$\left|\mathbf{u}^t \cdot \nabla\left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}\right)\right| \geq \frac{s_0}{2} \qquad \forall\, \mathbf{x} \in \mathbf{U}\,.$$

  On applying the Cauchy–Schwartz inequality, we obtain that

$$\left\|\nabla\left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}\right)\right\|_2 \geq \frac{s_0}{2} \qquad \forall\, \mathbf{x} \in \mathbf{U}\,.$$

  This together with the fact that $\|\,.\,\|_2 \leq \sqrt{d}\,\|\,.\,\|_\infty$ implies the second inequality appearing in (83).

- Suppose $k \geq 2$ in (85). Consider the function $g(\mathbf{x}) := \frac{\partial(\mathbf{f}\cdot\mathbf{v})}{\partial \mathbf{u}}(\mathbf{x})$ defined on $\mathbf{U}$. Then by (87), we have that

$$\left|\frac{\partial^{k-1} g}{\partial \mathbf{u}^{k-1}}(\mathbf{x})\right| \geq \frac{s_0}{2} \qquad \forall\, \mathbf{x} \in \mathbf{U}\,.$$

  Thus, the hypotheses of [8, Lemma 3.6] are satisfied for the function $g(\mathbf{x})$ and a straightforward application of that lemma together with (41) implies that

$$\sup_{\mathbf{x}\in\mathbf{U}} |g(\mathbf{x})| \geq \frac{s_0}{4(l-1)^{l-1} l!}(2r)^{l-1} > \rho_1 \sqrt{d}. \tag{90}$$

  Now the Cauchy–Schwartz inequality and (89) imply that

$$\left\|\nabla\left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}\right)\right\|_2 \geq \left|\mathbf{u}^t \cdot \nabla\left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}\right)\right| = |g(\mathbf{x})| \qquad \forall\, \mathbf{x} \in \mathbf{U}\,.$$

  This together with (90) and the fact that $\|\,.\,\|_2 \leq \sqrt{d}\,\|\,.\,\|_\infty$ imply the desired statement; namely that

$$\sup_{\mathbf{x}\in\mathbf{U}} \left\|\nabla\left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}\right)\right\|_\infty \geq \rho_1\,. \quad \square$$

The upshot of §7 is that we have verified conditions (i) & (ii) of Theorem 7 as desired modulo Proposition 2.

8. Proof of Proposition 2

In order to prove Proposition 2, we first establish an explicit version of [8, Lemma 4.3]. Throughout this section, the notation introduced in (68) will be used.

Lemma 8. *Let $B \subset \mathbb{R}^d$ be a ball of radius 1 and let B_∞ denote the hypercube circumscribed around B with edges parallel to the coordinate axes. Assume further that $\mathbf{p} = (p_1, p_2) : B \mapsto \mathbb{R}^2$ is a polynomial map of degree at most $l \geq 1$ such that*

$$\sup_{\mathbf{x},\mathbf{y}\in B_\infty} \|\mathbf{p}(\mathbf{x})-\mathbf{p}(\mathbf{y})\|_2 \leq 2 \tag{91}$$

and

$$\sup_{\mathbf{x}\in B} \operatorname{dist}(\mathcal{L},\mathbf{p}(\mathbf{x})) \geq \frac{1}{8} \tag{92}$$

for any straight line $\mathcal{L}\subset\mathbb{R}^2$*. Then,*

$$\sup_{\mathbf{x}\in B} \|\widetilde{\nabla}\mathbf{p}(\mathbf{x})\|_2 \geq \frac{1}{86\,016\sqrt{10}}\left(1+\sup_{\mathbf{x}\in B}\|\mathbf{p}(\mathbf{x})\|_2\right) \tag{93}$$

and

$$\sup_{\mathbf{x}\in B,\, i=1,2} \|\nabla p_i(\mathbf{x})\|_2 \leq 2l^2\sqrt{d}. \tag{94}$$

Proof. Regarding (93), if we assume that $\sup_{\mathbf{x}\in B}\|\mathbf{p}(\mathbf{x})\|_2 > 6$, the argument used to prove [8, Lemma 4.3] gives the stronger inequality in which the constant factor $1/(86\,016\sqrt{10})$ is replaced by $1/64$. Thus, without loss of generality, assume that

$$\sup_{\mathbf{x}\in B}\|\mathbf{p}(\mathbf{x})\|_2 \leq 6\,. \tag{95}$$

It is easily inferred from (92) that there exists $\mathbf{x}_1\in\overline{B}$, the closure of B, such that $\|\mathbf{p}(\mathbf{x}_1)\|_2\geq 1/8$. Working in polar coordinates and choosing the straight line $\mathcal{L}_1$ joining the origin to $\mathbf{p}(\mathbf{x}_1)$ to be the polar axis, let $(\rho(\mathbf{x}),\theta(\mathbf{x}))$ denote the polar coordinates of a vector $\mathbf{x}\in\mathbb{R}^2$. Thus, $\rho(\mathbf{p}(\mathbf{x}_1))\geq 1/8$. Furthermore, from (92), there exists $\mathbf{x}_2\in\overline{B}$ such that $\operatorname{dist}(\mathcal{L}_1,\mathbf{p}(\mathbf{x}_2))\geq 1/8$ and therefore, together with (95), we have that

$$\begin{aligned}|\theta(\mathbf{p}(\mathbf{x}_2))| &\geq |\sin\theta(\mathbf{p}(\mathbf{x}_2))| = \frac{\operatorname{dist}(\mathcal{L}_1,\mathbf{p}(\mathbf{x}_2))}{\rho(\mathbf{p}(\mathbf{x}_2))}\\ &\geq \frac{1/8}{6} = \frac{1}{48}.\end{aligned} \tag{96}$$

Now let Δ be the straight line joining $\mathbf{p}(\mathbf{x}_1)$ and $\mathbf{p}(\mathbf{x}_2)$. Furthermore, let $\mathcal{L}_2$ denote the x-coordinate axis, (x_1,y_1) the Cartesian coordinates of $\mathbf{p}(\mathbf{x}_2)$ and $(\rho(\mathbf{p}(\mathbf{x}_1)),0)$ the Cartesian coordinates of $\mathbf{p}(\mathbf{x}_1)$. Then the Cartesian equation of Δ is

$$\Delta \;:\; y_1x-(x_1-\rho(\mathbf{p}(\mathbf{x}_1)))y-\rho(\mathbf{p}(\mathbf{x}_1))y_1=0.$$

It follows from the choice of the points $\mathbf{x}_1$ and $\mathbf{x}_2$ together with (91), (92) and (95) that

$$\frac{1}{8}\leq|y_1|\leq 6,\quad \rho(\mathbf{p}(\mathbf{x}_1))\geq\frac{1}{8}\quad\text{and}\quad |x_1-\rho(\mathbf{p}(\mathbf{x}_1))|\leq 2.$$

Therefore, the distance from the origin O to Δ satisfies the inequality

$$\operatorname{dist}(\Delta, O) = \frac{|(\rho(\mathbf{p}(\mathbf{x}_1))y_1|}{\sqrt{y_1^2 + (x_1 - \rho(\mathbf{p}(\mathbf{x}_1))^2}} \geq \frac{(1/8)^2}{\sqrt{6^2+2^2}} = \frac{1}{128\sqrt{10}}. \tag{97}$$

Let J denote the straight line segment $[\mathbf{x}_1, \mathbf{x}_2]$ and let $\mathbf{u}$ be the unit vector

$$\mathbf{u} := \frac{\mathbf{x}_2 - \mathbf{x}_1}{\|\mathbf{x}_2 - \mathbf{x}_1\|_2}.$$

Restricting $\mathbf{p}$ to J, Lagrange's Theorem guarantees the existence of $\mathbf{y} \in (\mathbf{x}_1, \mathbf{x}_2)$ such that

$$\theta(\mathbf{p}(\mathbf{x}_2)) = \frac{\partial \theta}{\partial \mathbf{u}}(\mathbf{y}) |J|.$$

It then follows via (74), (96) and (97) that

$$\begin{aligned} \|\tilde{\nabla}\mathbf{p}(\mathbf{y})\|_2 &\geq |\mathbf{u} \cdot \tilde{\nabla}\mathbf{p}(\mathbf{y})| = \rho^2(\mathbf{y}) \left| \frac{\partial \theta}{\partial \mathbf{u}}(\mathbf{y}) \right| \\ &\geq \operatorname{dist}(\Delta, O) \frac{|\theta(\mathbf{p}(\mathbf{x}_2))|}{|J|} = \operatorname{dist}(\Delta, O) \frac{|\theta(\mathbf{p}(\mathbf{x}_2))|}{\|\mathbf{x}_2 - \mathbf{x}_1\|_2} \\ &\geq \frac{1}{128\sqrt{10} \times 48 \times 2} = \frac{1}{12\,288\sqrt{10}}. \end{aligned}$$

Thus,

$$\sup_{\mathbf{x} \in B} \|\tilde{\nabla}\mathbf{p}(\mathbf{x})\|_2 \geq \frac{1}{12\,288\sqrt{10}} = \frac{7}{86\,016\sqrt{10}} \geq \frac{1}{86\,016\sqrt{10}} \left(1 + \sup_{\mathbf{x} \in B} \|\mathbf{p}(\mathbf{x})\|_2\right).$$

This completes the proof of (93). We now turn out attention to (94).

Let $i \in \{1, 2\}$. It may be assumed without loss of generality that $p_i(0, \ldots, 0) = 0$ and that the ball B is centred at the origin. Then, for given $x_2, \ldots, x_d$ in $\mathbb{R}$, consider the polynomial in one variable $p(x) := p_i(x, x_2, \ldots, x_d)$, which is of degree at most l. It follows from (91) that

$$\sup_{|x| \leq 1} |p(x)| \leq 2.$$

Hence by Markov's inequality for polynomials, we have that

$$\sup_{|x| \leq 1} \left| \frac{\mathrm{d}p}{\mathrm{d}x}(x) \right| = \sup_{|x_1| \leq 1} \left| \frac{\partial p_i}{\partial x_1}(x_1, x_2, \ldots, x_d) \right| \leq 2l^2.$$

This together with the fact that $\|\,.\,\|_2 \leq \sqrt{d}\,\|\,.\,\|_\infty$ implies that

$$\max\left\{\sup_{\mathbf{x}\in B}\|\nabla p_1(\mathbf{x})\|_2\,,\sup_{\mathbf{x}\in B}\|\nabla p_2(\mathbf{x})\|_2\right\} \leq 2l^2\sqrt{d}$$

and therefore completes the proof of the lemma. □

We now have all the ingredients in place to prove Proposition 2.

Proof of Proposition 2. The proposition is an explicit version of [8, Proposition 4.1]. Within our setup in which $\mathcal{G}$ is given by (73), the starting point for the proof of part (a) of [8, Proposition 4.1] corresponds to the existence of positive constants δ, c, and α with

$$0<\delta<1/8 \qquad \text{and} \qquad 2C_{d,l}N_d\delta^{1/(d(2l-1)(2l-2))}\leq 1 \tag{98}$$

such that for every $\mathbf{g}\in\mathcal{G}$ one has

$$\forall\,\mathbf{v}\in\mathbb{S}^1 \quad \exists\,\mathbf{u}\in\mathbb{S}^{d-1} \quad \exists\,k\leq l \quad : \quad \inf_{\mathbf{x}\in\tilde{\mathbf{U}}}\left|\mathbf{v}\cdot\frac{\partial^k\mathbf{g}}{\partial\mathbf{u}^k}(\mathbf{x})\right|\geq c \tag{99}$$

and

$$\sup_{\mathbf{x},\mathbf{y}\in\tilde{\mathbf{U}}}\|\partial_\beta\mathbf{g}(\mathbf{x})-\partial_\beta\mathbf{g}(\mathbf{y})\|\leq\frac{\delta c\alpha}{8\xi l^l(l+1)!}=\frac{\delta c\alpha}{16l^{l+2}(l+1)!\sqrt{d}} \tag{100}$$

for all multi-indices β with $|\beta|=l$. Here, $\xi=2l^2\sqrt{d}$ is the quantity in right-hand side of (94) and the real number α is required to be less than the constant appearing in the right-hand side of (93), that is,

$$\alpha\leq\frac{1}{86\,016\sqrt{10}}. \tag{101}$$

The statements (99) and (100) correspond exactly to [8, Eq. (4.5a) & Eq. (4.5b)] with V replaced by $\tilde{\mathbf{U}}$.

The proof of part (a) of Proposition 2 follows from the existence of the constants δ, c, and α as established in the proof of part (a) of [8, Proposition 4.1]. It remains for us to show that, given the definition of r in (37), it is indeed possible to choose such constants in such a way that the relations (98)–(101) hold.

With this in mind, set

$$\delta:=\eta,$$

where η is defined by (38). It follows from the definition of η and the well known bound $N_d\leq 5^d$ for the Besicovitch constant (cf. Remark 3 p.245) that (98) is satisfied with $\delta=\eta$. We proceed with verifying (99) and (100).

Regarding (99), let $\mathbf{g} = (\mathbf{u}_1 \cdot \mathbf{f}, \mathbf{u}_2 \cdot \mathbf{f} + u_0) \in \mathcal{G}$. Also, let $\mathbf{u} := (\mathbf{u}_1, \mathbf{u}_2)$ with $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{S}^{n-1}$ and let $\mathbf{v} := (v_1, v_2) \in \mathbb{S}^1$. Furthermore, let $\mathbf{w}$ denote the vector $\mathbf{w} := v_1\mathbf{u}_1 + v_2\mathbf{u}_2$. Since by the definition of $\mathcal{G}$, the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal, it follows that $\mathbf{w} \in \mathbb{S}^{n-1}$. Now observe that for any multi-index β such that $|\beta| \leq l$,

$$\mathbf{v} \cdot \partial_\beta \mathbf{g} = \mathbf{w} \cdot \partial_\beta \mathbf{f}.$$

By the definition of $s_0 = s(l; \mathbf{f}, \tilde{\mathbf{U}}^+)$, there exists $\mathbf{s} \in \mathbb{S}^{d-1}$ and $1 \leq k \leq l$ such that

$$\left| \mathbf{v} \cdot \frac{\partial^k \mathbf{g}}{\partial \mathbf{s}^k}(\mathbf{x}_0) \right| = \left| \mathbf{w} \cdot \frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}_0) \right| \geq s_0 \, .$$

As per usual, $\mathbf{x}_0$ denotes here the centre of $\mathbf{U}$. It follows that for any $\mathbf{x} \in \tilde{\mathbf{U}}^+$, we have that

$$\begin{aligned} \left| \mathbf{v} \cdot \frac{\partial^k \mathbf{g}}{\partial \mathbf{s}^k}(\mathbf{x}) \right| &\geq \left| \mathbf{w} \cdot \frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}_0) \right| - \left| \mathbf{w} \cdot \left(\frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}_0) - \frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}) \right) \right| \\ &\geq s_0 - \left\| \frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}_0) - \frac{\partial^k \mathbf{f}}{\partial \mathbf{s}^k}(\mathbf{x}) \right\|_2 . \end{aligned} \tag{102}$$

The same arguments as those used to prove (87) can be employed to show that

$$\left| \mathbf{v} \cdot \frac{\partial^k \mathbf{g}}{\partial \mathbf{s}^k}(\mathbf{x}) \right| \geq \frac{s_0}{2} \qquad \forall \, \mathbf{x} \in \tilde{\mathbf{U}}^+ .$$

This proves (99) with

$$c := \frac{s_0}{2}.$$

We now turn our attention to (100). With $\mathbf{g}$ and $\mathbf{u}$ as above, first note that for any $\mathbf{x} \in \tilde{\mathbf{U}}^+$ and for any multi-index β such that $|\beta| = l$, we have that

$$\|\partial_\beta \mathbf{g}(\mathbf{x}) - \partial_\beta \mathbf{g}(\mathbf{x}_0)\|_2 = \|(\mathbf{u}_1 \cdot (\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)), \mathbf{u}_2 \cdot (\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)))\|_2 \, . \tag{103}$$

Next, note that from the Cauchy–Schwarz inequality, we have that, for $i = 1, 2$,

$$(\mathbf{u}_i \cdot (\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)))^2 \leq \|\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)\|_2^2. \tag{104}$$

On combining (103) and (104), we find that

$$\|\partial_\beta \mathbf{g}(\mathbf{x}) - \partial_\beta \mathbf{g}(\mathbf{x}_0)\|_2 \leq \sqrt{2} \|\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)\|_2.$$

Now, since $\mathbf{f}$ satisfies Assumption 1, in view of (36), we obtain that

$$\|\partial_\beta \mathbf{f}(\mathbf{x}) - \partial_\beta \mathbf{f}(\mathbf{x}_0)\|_2 \leq M\sqrt{2}\,\|\mathbf{x}-\mathbf{x}_0\|_2\,.$$

Hence, for any $\mathbf{x}, \mathbf{y} \in \tilde{\mathbf{U}}^+$ we have that

$$\|\partial_\beta \mathbf{g}\,(\mathbf{x}) - \partial_\beta \mathbf{g}\,(\mathbf{y})\|_2 \;\leq\; 2M\,(\|\mathbf{x}-\mathbf{x}_0\|_2 + \|\mathbf{y}-\mathbf{x}_0\|_2)\,.$$

In view of (37), we also have that

$$\begin{aligned}\|\mathbf{x}-\mathbf{x}_0\|_2 + \|\mathbf{y}-\mathbf{x}_0\|_2 &\leq 2\cdot 3^{n+d+2}\cdot r\\ &\leq 2\cdot 3^{n+d+2}\cdot \frac{\eta s_0}{4\cdot 10^7 3^{n+d+2}\,dMl^{l+2}(l+1)!}\\ &= \frac{\eta s_0}{2\cdot 10^7\,dMl^{l+2}(l+1)!}.\end{aligned}$$

The upshot is that

$$\sup_{\mathbf{x},\mathbf{y}\in\tilde{\mathbf{U}}^+} \|\partial_\beta \mathbf{g}\,(\mathbf{x}) - \partial_\beta \mathbf{g}\,(\mathbf{y})\|_2 \;\leq\; \frac{\eta s_0}{10^7\,dl^{l+2}(l+1)!} \tag{105}$$

for any multi-index β with $|\beta| = l$. This proves (100) with

$$\alpha := \frac{32}{10^7\sqrt{d}}\,,$$

which clearly satisfies (101).

To prove part (b) of Proposition 2, we closely follow the proof of part (b) of [8, Proposition 4.1]. The new ingredient in our proof is the calculation of explicit constants at appropriate places. With this in mind, let $\mathbf{g} = (g_1, g_2) \in \mathcal{G}$ and take B appearing at the start of the proof of [8, Proposition 4.1(b)] to be $\mathbf{U}$, so that $\hat{B} = \frac{1}{2}\mathbf{U}$. We claim that there exists a point $\mathbf{y} \in \hat{B}$ such that

$$\|\mathbf{g}(\mathbf{y})\|_2 \;\geq\; \tau := \frac{r^l s_0}{4l^l(l+1)!}. \tag{106}$$

To see that this is so, take $\mathbf{v} := (1,0) \in \mathbb{S}^1$. In view of (99), there exists a vector $\mathbf{u} \in \mathbb{S}^{d-1}$ and $1 \leq k \leq l$ such that

$$\left|\mathbf{v}\cdot\frac{\partial^k \mathbf{g}}{\partial \mathbf{u}^k}\,(\mathbf{x})\right| \;=\; \left|\frac{\partial^k g_1}{\partial \mathbf{u}^k}\,(\mathbf{x})\right| \;\geq\; c := \frac{s_0}{2} \quad \forall\, \mathbf{x}\in\mathbf{U}\,.$$

Thus, on applying [8, Lemma 3.6] to the function g_1 and the ball $\hat{B}$, we obtain that

$$\sup_{\mathbf{x},\mathbf{y}\in\hat{B}} |g_1(\mathbf{x}) - g_1(\mathbf{y})| \geq \frac{r^l c}{l^l(l+1)!} = \frac{r^l s_0}{2l^l(l+1)!}.$$

This implies the existence of a point $\mathbf{y} \in \hat{B}$ such that

$$\|\mathbf{g}(\mathbf{y})\|_2 \geq |g_1(\mathbf{y})| \geq \tau$$

as claimed. Next, observe that for any $\mathbf{w} \in \mathbb{S}^{d-1}$ and any $\mathbf{x} \in \mathbf{U}$,

$$\frac{\partial \mathbf{g}}{\partial \mathbf{w}}(\mathbf{x}) = \begin{pmatrix} \mathbf{u}_1 \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{w}}(\mathbf{x}) \\ \mathbf{u}_2 \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{w}}(\mathbf{x}) \end{pmatrix}.$$

Therefore, on using the Cauchy–Schwarz inequality, we obtain via (36) that

$$\left\| \frac{\partial \mathbf{g}}{\partial \mathbf{w}}(\mathbf{x}) \right\|_2 \leq \sqrt{2} \left\| \frac{\partial \mathbf{f}}{\partial \mathbf{w}}(\mathbf{x}) \right\|_2 \leq \sqrt{2}M\,. \tag{107}$$

Now, observe that, in view of (37), of the definition of τ and of the fact that $s_0 \leq M$, we have that

$$\tau \leq r\,M.$$

Consider the ball $B' \subset B = \mathbf{U}$ with radius $\tau/(2M) \leq r/2$ centred at $\mathbf{y}$, where $\mathbf{y}$ satisfies (106). Take a vector $\mathbf{v} \in \mathbb{S}^1$ orthogonal to $\mathbf{g}(\mathbf{y})$. In view of (99), there exists a vector $\mathbf{u} \in \mathbb{S}^{d-1}$ and $1 \leq k \leq l$ such that

$$\left| \mathbf{v} \cdot \frac{\partial^k \mathbf{g}}{\partial \mathbf{u}^k}(\mathbf{x}) \right| \geq c = \frac{s_0}{2} \quad \forall\, \mathbf{x} \in \mathbf{U}\,.$$

Thus, on applying [8, Lemma 3.6] to the function $\mathbf{x} \to \mathbf{v}\cdot\mathbf{g}(\mathbf{x})$ and the ball B', we obtain that

$$\sup_{\mathbf{x}\in B'} |\mathbf{v}\cdot\mathbf{g}(\mathbf{x})| \geq \frac{s_0}{4l^l(l+1)!} \left(\frac{\tau}{M}\right)^l. \tag{108}$$

On the other hand, the upper bound (107) implies that

$$\sup_{\mathbf{x}\in B'} \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\|_2 \leq \frac{\tau}{2M}\sqrt{2}M = \frac{\tau}{\sqrt{2}}. \tag{109}$$

The upshot of (108) and (109) is that we are able to apply [8, Lemma 4.2] to the map $\mathbf{g} : B' \to \mathbb{R}^2$ to yield (75) and thereby complete the proof of part (b) of Proposition 2. For ease of comparison, we point out that the quantities a, δ and w appearing in the statement of [8, Lemma 4.2] correspond to τ, $\tau/\sqrt{2}$ and the right-hand side of (108) respectively. □

Acknowledgments

The main catalysts for this paper were the 'York-BT Mathematics of MIMO' meeting at BT, Adastral Park on 14 June 2013 and the subsequent international 'Workshop on interactions between number theory and wireless communication' at the University of York between 9–23 May 2014. The aim of these meetings was to explore potential applications to wireless technology of current research on Diophantine approximation. We would like to take this opportunity to thank the 'electronic' participants for putting up with our naive questions and often nonsensical ramblings. In particular, we would like to thank to Alistar Burr (Dept. of Electronics, York) and Keith Briggs (BT Labs) for their enthusiastic support in making sure that the meetings actually happened — they were our main link to the electronics world!

All five authors of this paper are number theorists and we have all benefited hugely from the presence of Maurice Dodson in our lives at some stage in our careers in some form or another. Not only does he have a wonderful mind and personality but his intellectual curiosity knows no boundaries. We would like to think that some of his curiosity has rubbed off onto us and for this we are forever in his debt. There is so much to gain in engaging in dialogue with researchers in other disciplines and attempting to overcome 'language' barriers. In short, thank you Maurice for being a genuine intellectual.

SV would like to thank his teenage daughters Iona and Ayesha for still being a joyous addition in his life but would strongly urge them to improve their taste in music. Also a massive thanks to Bridget Bennett for sticking around even at fifty — congratulations! EZ is enormously grateful to Aljona who not only recently gave birth to their daughter Alyssa but looked after her during his obsession with this project. Infinite thanks to Alyssa herself, just for her simple existence and for all the joy and happiness she has already brought to his life.

Appendix A. Diophantine approximation on manifolds and wireless technology

In short, interference alignment is a linear precoding technique that attempts to align signals in time, frequency, or space. The following exposition is an attempt to illustrate at a basic level the role of Diophantine approximation in implementing this technique. We stress that this section is not meant for the "electronics" experts. We consider two examples. The first basic example brings into play the theory of Diophantine approximation while the second slightly more complicated example also brings into play the manifold theory.

Example 1. There are two people (*users*) S_1 and S_2 who wish to send (*transmit*) a message (*signal*) $u \in \{0,1\}$ and $v \in \{0,1\}$ respectively along a single communication channel (could be a cable or radio channel) to a person (*receiver*) R. Suppose there is a certain degree of fading (*channel coefficients*) associated with the messages during transmission along the channel. This for instance could be dependent on the distance

of the users to the receiver and in the case of a radio channel, the reflection caused by obstacles such as buildings in the path of the signal. It is worth stressing that this aspect of "fading" associated with a signal should not be confused with the more familiar aspect of a signal being corrupted by "noise" that will be discussed a little later. Let h_1 and h_2 denote the fading factors associated with the messages being sent by S_1 and S_2 respectively. These are strictly positive numbers and assume their sum is one. Also, assume that the channel is additive. That is to say that R receives the message:

$$y = h_1x_1 + h_2x_2 \qquad \text{where} \quad x_1 = u \quad \text{and} \quad x_2 = v\,. \tag{110}$$

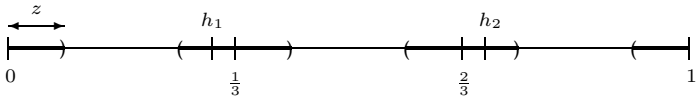
Specifically, the outcomes of y are

$$y = \begin{cases} 0 & \text{if} \quad u = v = 0 \\ h_1 & \text{if} \quad u = 0 \text{ and } v = 1 \\ h_2 & \text{if} \quad u = 1 \text{ and } v = 0 \\ 1 = h_1 + h_2 & \text{if} \quad u = v = 1\,, \end{cases} \tag{111}$$

and if $h_1 \neq h_2$, the receiver is obviously able to recover the messages u and v. Moreover, the greater the mutual separation of the above four outcomes in the unit interval $I = [0, 1]$, the better the tolerance for error (*noise*) during the transmission of the signal. The noise can be a combinations of various factors but often the largest contributing factor is the interference caused by other communication channels. If z denotes the noise, then instead of (110), in practice R receives the message:

$$y = h_1x_1 + h_2x_2 + z \qquad \text{where} \quad x_1 = u \quad \text{and} \quad x_2 = v\,. \tag{112}$$

Now let d denote the minimum distance between the four outcomes of $y \in I$ which are explicitly given by (111). Then as long as the absolute value $|z|$ of the noise is strictly less than $d/2$, the receiver is able to recover the messages u and v. This is simply due to the fact that intervals of radius $d/2$ centred at the four outcomes of y are disjoint. In this basic example, it is easy to see that the maximum separation between the four outcomes is attained when $h_1 = 1/3$ and $h_2 = 2/3$. In this case $d = 1/3$, and we are able to recover the messages u and v as long as $|z| < 1/6$. The upshot is that the closer the real numbers h_1 and h_2 are to $1/3$ and $2/3$ the better the tolerance for noise. Hence, at the most fundamental level we are interested in the simultaneous approximation property of real numbers by rational numbers. In practice, it is the probabilistic aspect of the approximation property that is important — knowing that the numbers h_1 and h_2 lie within a 'desirable' neighbourhood of the points $1/3$ and $2/3$ with reasonably high probability is key. This naturally brings into play the theory of metric Diophantine approximation.



Note that from a probabilistic point of view, the chances that $h_1 = h_2$ is zero and is therefore insignificant. Furthermore, within the context of this basic example, by weighting (*precoding*) the messages u and v appropriately before the transmission stage it is possible to ensure optimal separation ($d = 1/3$) at the receiver regardless of the values of h_1 and h_2. Indeed, suppose $x_1 = \frac{1}{3}h_1^{-1}u$ and $y_2 = \frac{2}{3}h_2^{-1}v$ are transmitted instead of u and v. Then, without taking noise into consideration, R receives the message

$$y = h_1x_1 + h_2x_2 = \tfrac{1}{3}u + \tfrac{2}{3}v \tag{113}$$

and so the specifics outcomes are

$$y = \begin{cases} 0 & \text{if} \quad u = v = 0 \\ 1/3 & \text{if} \quad u = 0 \text{ and } v = 1 \\ 2/3 & \text{if} \quad u = 1 \text{ and } v = 0 \\ 1 & \text{if} \quad u = v = 1\,. \end{cases} \tag{114}$$

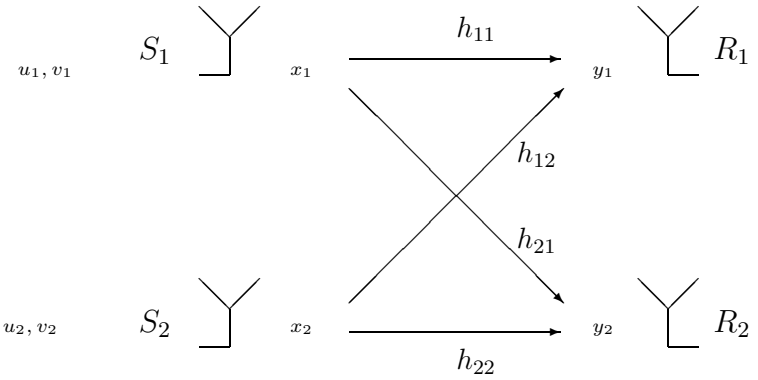
Example 2. There are two users S_1 and S_2 as before but this time there are also two receivers R_1 and R_2. Suppose S_1 wishes to simultaneously transmit independent signals u_1 and v_1 as a single signal, say $x_1 = u_1 + v_1$ where u_1 is intended for R_1 and v_1 for R_2. Similarly, suppose S_2 wishes to simultaneously transmit independent signals u_2 and v_2 as a single signal, say $x_2 = u_2 + v_2$ where u_2 is intended for R_1 and v_2 for R_2. As in the first example, for the sake of simplicity, we can assume that the signals $u_1, u_2, v_1, v_2 \in \{0, 1\}$. Now let h_{11} and h_{21} denote the channel coefficients associated with signals being sent by S_1 to R_1 and R_2 respectively. Similarly, let h_{12} and h_{22} denote the channel coefficients associated with signals being sent by S_2 to R_1 and R_2. Assume that the channel is additive and let y_1 (respectively y_2) denote the signal at receiver R_1 (respectively R_2). Thus,

$$y_1 = h_{11}x_1 + h_{12}x_2 \tag{115}$$

$$y_2 = h_{21}x_1 + h_{22}x_2 \tag{116}$$

where

$$x_1 = u_1 + v_1 \quad \text{and} \quad x_2 = u_2 + v_2\,. \tag{117}$$



Recall, that R_1 (respectively R_2) only cares about recovering the signals u_1 and u_2 (respectively v_1 and v_2) from y_1 (respectively y_2). For the moment, let us just concentrate on the signal received by R_1; namely

$$y_1 = h_{11}u_1 + h_{12}u_2 + h_{11}v_1 + h_{12}v_2\,.$$

It is easily seen that this corresponds to a received signal in Example 1 modified to incorporate four users and one receiver. This time there are potentially 16 different outcomes. In short, the more users, the more outcomes and therefore the smaller the mutual separation between them and in turn the smaller the tolerance for noise. Now there is one aspect of the setup in this example that we have not yet exploited. The receiver R_1 is not interested in the signals v_1 and v_2. So if they could be deliberately aligned via precoding into a single component $v_1 + v_2$, then y_1 would look like a received signal associated with just 3 users rather than 4. With this in mind, suppose instead of transmitting x_1 and x_2 given by (117), S_1 and S_2 transmit the signals

$$x_1 = h_{22}u_1 + h_{12}v_1 \quad \text{and} \quad x_2 = h_{21}u_2 + h_{11}v_2 \tag{118}$$

respectively. Then, it can be verified that the received signals given by (115) and (116) can be written as

$$y_1 = (h_{11}h_{22})u_1 + (h_{21}h_{12})u_2 + (h_{11}h_{12})(v_1 + v_2)$$

$$y_2 = (h_{21}h_{12})v_1 + (h_{11}h_{22})v_2 + (h_{21}h_{22})(u_1 + u_2)\,.$$

In other words, the unwanted, interfering signals at either receiver are aligned to a one dimensional subspace of four dimensional space. Notice that in the above equations the six coefficients are only of four variables, namely $h_{i,j}$, $i, j = 1, 2$, and thus represent dependent quantities. This, together with our findings from Example 1, naturally brings into play the manifold theory of metric Diophantine approximation.

Example 2 is a simplified version of Example 3 appearing in [16, §III]. For a deeper and more practical understanding of the link between interference alignment and metric

Diophantine approximation on manifolds the reader is urged to look at [16] and [12, §4.7].

References

[1] K. Ball, Cube slicing in $\mathbb{R}^n$, Proc. Amer. Math. Soc. 97 (1986) 465–473.
[2] V. Beresnevich, A Groshev type theorem for convergence on manifolds, Acta Math. Hungar. 94 (1–2) (2002) 99–130.
[3] V. Beresnevich, V. Bernik, M. Dodson, S. Velani, Classical metric Diophantine approximation revisited, in: Analytic Number Theory, Cambridge Univ. Press, Cambridge, 2009, pp. 38–61.
[4] V. Beresnevich, D. Dickinson, S. Velani, Measure theoretic laws for limsup sets, Mem. Amer. Math. Soc. 179 (846) (2006), x+91.
[5] V. Beresnevich, D. Dickinson, S. Velani, Diophantine approximation on planar curves and the distribution of rational points, Ann. of Math. (2) 166 (2) (2007) 367–426.
[6] V. Beresnevich, S. Velani, Classical metric Diophantine approximation revisited: the Khintchine—Groshev Theorem, Int. Math. Res. Not. IMRN (1) (2010) 69–86.
[7] V. Beresnevich, E. Zorin, Explicit bounds for rational points near planar curves and metric Diophantine approximation, Adv. Math. (2010) 3064–3087.
[8] V.I. Bernik, D. Kleinbock, G.A. Margulis, Khintchine—type theorems on manifolds: the convergence case for standard and multiplicative versions, Int. Math. Res. Not. IMRN (2001) 453–486.
[9] Z. Füredi, P.A. Loeb, On the best constant for the Besicovitch covering theorem, Proc. Amer. Math. Soc. 121 (4) (1994) 1063–1073.
[10] P.X. Gallagher, Metric simultaneous Diophantine approximation II, Mathematika 12 (1965) 123–127.
[11] A. Ghasemi, A.S. Motahari, A.K. Khandani, Interference alignment for the K user MIMO interference channel, in: 2010 IEEE International Symposium on Information Theory Proceedings, ISIT, 2010.
[12] S.A. Jafar, Interference alignment – a new look at signal dimensions in a communication network, Found. Trends Commun. Inf. Theory 7 (1) (2010).
[13] D.Y. Kleinbock, G.A. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds, Ann. of Math. (2) 148 (1998) 339–360.
[14] S.H. Mahboubi, A.S. Motahari, A.K. Khandani, Layered Interference alignment: achieving the total DOF of MIMO X-channels, in: ISIT 2010, Austin, Texas U.S.A., 2010.
[15] A.S. Motahari, S. Oveis-Gharan, M.A. Maddah-Ali, A.K. Khandani, Real Interference Alignment, 2010.
[16] A.S. Motahari, S. Oveis-Gharan, M.A. Maddah-Ali, A.K. Khandani, Real interference alignment: exploiting the potential of single antenna systems, IEEE Trans. Inform. Theory 60 (8) (2014) 4799–4810.
[17] U. Niesen, P. Whiting, The degrees of freedom of compute-and-forward, IEEE Trans. Inform. Theory 58 (8) (2012).
[18] W.M. Schmidt, A metrical theorem in diophantine approximation, Canad. J. Math. 12 (1960) 619–631.
[19] V. Sprindžuk, Metric Theory of Diophantine Approximation, John Wiley & Sons, New York–Toronto–London, 1979 (English transl.).
[20] R. Vaughan, S. Velani, Diophantine approximation on planar curves: the convergence theory, Invent. Math. 166 (1) (2006) 103–124.
[21] Y. Wu, S. Shamai, S. Verdú, Information dimension and the degrees of freedom of the interference channel, IEEE Trans. Inform. Theory 61 (1) (2015) 256–279.
[22] J. Xie, S. Ulukus, Real interference alignment for the K-user Gaussian interference compound wiretap channel, in: Communication, Control, and Computing, Allerton, 2010.
[23] M. Zamanighomi, Z. Wang, Multiple-antenna interference channels with real interference alignment and receive antenna joint processing based on simultaneous Diophantine approximation, IEEE Trans. Inform. Theory 60 (8) (2014) 4757–4769.