Balian-Low phenomena, exponential bases and neural network approximation

A Dissertation

Submitted in Partial Fulfillment of the Requirements
for the Degree of
Doctor in Mathematics
(Dr. rer. nat.)

KATHOLISCHE UNIVERSITÄT
EICHSTÄTT-INGOLSTADT

Submitted by Andrei Caragea
Department of Mathematics, MGF
Proposal for reviewer: Prof. Dr. Götz Pfander
Proposal for co-reviewer: Prof. Dr. David Walnut

Eichstätt April 2023

Contents

1 Acknowledgements

The work of the author of the present thesis has been supported over the last 3 years by the DFG project PF 450/11-1.

I would like to firstly thank my advisor, professor Götz Pfander. I was fortunate enough to meet G. Pfander during my first semester of undergraduate studies, some 15 years ago. Throughout this long span of time, my advisor has been the immutable bedrock of my academic pursuits. In all honesty, without his unwavering patience, support and encouragement, I would not be writing this thesis. Moreover, it was his trust in me that gave the courage to return to academia after a 4 year 'experiment' in the standard job market, as a programmer. I am convinced that without my advisor, the only doctor in my family would still be my dad. For all that G. Pfander has done to help, motivate and advise me, I will be forever grateful.

I have been extremely fortunate to have amazing colleagues in G. Pfander's workgroup. Much of my day-to-day progress as a mathematician is due to being surrounded by people from whom there is always something to learn. I would like to thank Prof. Felix Voigtlaender (who I view as my second mentor as well as one of the most talented mathematicians I've personally met), Dr. Dae Gwan Lee, Dr. Friedrich Philipp, Prof. Dominik Stöger, Prof. Johannes Maly and Dr. Thomas Heindl, who, despite being in a different workgroup, is still 'one of us'. I am happy to think of them as my friends.

I would also like to thank my two math teachers from my school days, Elena Gălii and Gheorghe Pădurariu. They both played significant roles in my formative years and in developing my interest for mathematics.

In terms of non-mathematicians that have played a large role in inspiring my career and have been a role-model in terms of a scientist from my generation, I would like to thank Dr. Cristiana Elena Lungu. She is probably the most interesting person I have ever met and will forever be a part of who I am and how I think.

Last but definitely not least, I want to thank my mom and dad. The trust they have always shown me is something I cannot put into words. The sacrifices they have made in order for me to find a purpose in life and their sheer conviction that such an endeavor will be successful are, on the one hand, mystifying to me, and on the other the purest expression of sincere affection I have ever seen.

Thank you!

2 A foreword

The present thesis is a cumulative dissertation. It covers research done in three rather distinct areas of mathematics and has at its core six published papers. This thesis also includes some individual work done by the author, inspired by the papers and usually quite close in topic to them. Due to the very niche and rather minor nature of some of these individual results, the author has chosen not to pursue separate journal publication for them.

Section 3 discusses results concerning Balian-Low subspace phenomena, Section 4 deals with quantitative approximation with complex valued or curse of dimensionality avoiding neural networks and Section 5 mainly explores some applications of a novel construction technique for Riesz exponential bases developed in the mid 2010s by Kozma and Nitzan. Section 6 contains the full arXiv text of the six papers. References within the thesis to specific results from the 6 papers are based on the published versions.

The ordering of Sections 3, 4, 5 and the papers discussed within is chronological from the perspective of the author's involvement with them. Broadly speaking, the structure of these sections is the same, with an introduction to the topic and a discussion of the related state of the art, followed by an overview of the results of our relevant papers and, lastly, any additional work done by the author of the present thesis inspired by or closely related to the papers.

The main goal of the exposition style chosen for this thesis was to keep the technical formalities already fully detailed in the papers to a minimum in the main body of the thesis, as long as this did not hamper legibility and the flow of ideas. The otherwise unpublished material, especially in the subsections with additional work done by the author, is presented in full formality. The emphasis in writing the parts paraphrasing content contained in the six papers was to streamline the presentation of the intuitive flow of ideas and to illustrate the author's perception of the respective topics.

3 Balian-Low phenomena for subspaces

This section is based on three papers:

- 'A Balian-Low theorem for subspaces' (co-authored with D.G. Lee, G. Pfander and F. Philipp), published in The Journal of Fourier Analysis and Applications in 2019, henceforth referred to as **Paper 1**, [Car+19];
- 'A quantitative subspace Balian-Low theorem' (co-authored with D.G. Lee, F. Philipp and F. Voigtlaender), published in Applied and Computational Harmonic Analysis in 2021, henceforth referred to as **Paper 2**, [Car+21];
- 'A Balian-Low type theorem for Gabor Riesz sequences of arbitrary density' (co-authored with D.G. Lee, F. Philipp and F. Voigtlaender), published in Mathematische Zeitschrift in 2023, henceforth referred to as **Paper 3**, [Car+23].

The section is organized as follows. Subsection 1 briefly discusses some general harmonic analysis concepts and how the Balian-Low theorem has impacted Gabor analysis. Furthermore, we address the context that inspired our three papers. Subsection 2 broadly discusses the main ideas of each paper, with a focus on the development and contributions to each paper and the connections between them. Finally, subsection 3 highlights some peripheral results and ideas of the author that were meaningful in the development of the papers, but were replaced with simpler or more general versions in the published manuscripts.

3.1 Harmonic analysis and the Balian-Low theorem

The bedrock tool of harmonic analysis is the Fourier transform, densely defined on the space $L^2(\mathbb{R})$ of square integrable functions on the real line as

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-2\pi i \xi x} f(x) dx.$$

The Fourier transform viewed as an operator is an isometry

$$\mathcal{F} : L^2(\mathbb{R}) \to L^2(\widehat{\mathbb{R}})$$

where one usually regards $\mathbb{R}$ as the time axis and $\widehat{\mathbb{R}}$ as the frequency axis. Conceptually, the Fourier transform establishes an equivalence between the time and the frequency content of a signal.

Perhaps the most fundamental illustration of this correspondence is Plancherel's formula concerning the Fourier series of a function with compact support.

Theorem 1 (Plancherel's formula)**.** *Let $\{c_n\}_{n\in\mathbb{Z}}$ be an arbitrary $\ell^2(\mathbb{Z})$ sequence (i.e., $\sum_{n\in\mathbb{Z}} |c_n|^2 < \infty$). Then there exists a unique signal $f \in L^2([0,1])$ satisfying $\hat{f}(n) = c_n$ for all $n \in \mathbb{Z}$. f can be written as an L^2 function as $f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$, with convergence in the L^2 sense.*

Informally put, this implies that a signal supported on $[0,1]$ (in time) is fully determined by the magnitude of its integer frequencies. From a different perspective, Plancherel's formula implies that the pure frequency signals $x \mapsto e^{2\pi i n x}$ span (in fact form an orthonormal basis for) the set of of all square integrable signals with time domain $[0,1]$. This latter viewpoint motivates a more detailed look at various notions of bases in a Hilbert space.

Definition 1. *Let H be a Hilbert space with respect to the inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$. A collection of vectors $\{h_n\}_{n \in \mathbb{Z}}$ is called*

- *a **Frame** if there exist constants $A, B > 0$ such that for any element $h \in H$ the following double-sided inequality holds*

$$A\|h\|_H^2 \leq \sum_{n \in \mathbb{Z}} |\langle h, h_n \rangle|^2 \leq B\|h\|_H^2.$$

- *a **Riesz basis** if the h_n vectors span H and there exist constants $A, B > 0$ such that for any coefficient sequence $\{c_n\}_{n \in \mathbb{Z}} \in \ell^2$ the following double-sided inequality holds*

$$A\|\{c_n\}\|_2^2 \leq \|\sum_{n \in \mathbb{Z}} c_n h_n\|_H^2 \leq B\|\{c_n\}\|_2^2.$$

  Alternatively, a Riesz basis is a frame with no redundancy.

- *an **Orthonormal basis** if it is a Riesz basis with the additional property that*

$$\langle h_n, h_m \rangle = \delta_{n,m}.$$

Remark 1. *There are other equivalent definitions for all of these notions. A detailed discussion can be found in Chapter 5.1 of [Grö01]. One often extends these definitions to frame (or Riesz or orthonormal) sequences if the conditions are satisfied only for a proper subspace (namely the subspace spanned by the elements of the frame or Riesz or orthonormal sequence) of the ambient Hilbert space H.*

The crucial aspect of these different notions of a basis for a Hilbert space is that they give rise to reconstruction formulas (non-unique in the case of a frame and unique for Riesz and orthonormal bases). Given a frame $\{h_n\}_n$, one can define the **frame operator**

$$S : H \to H, \quad S(h) = \sum_n \langle h, h_n \rangle h_n,$$

which is a boundedly invertible self-adjoint operator on H (Chapter 5.1 in [Grö01]). This allows the construction of the **canonical dual frame** $\{\widetilde{h}_n\}_n$ given by $\widetilde{h}_n := S^{-1} h_n$. A direct computation using the linearity and self-adjointness of S^{-1} gives the reconstruction formulas

$$h = \sum_n \langle h, h_n \rangle \widetilde{h}_n = \sum_n \langle h, \widetilde{h}_n \rangle h_n$$

which hold for all $h \in H$, implying that any signal in H can be encoded without loss as a discrete linear combination of the frame signals with coefficients determined by inner products with the dual frame (or vice-versa).

Returning to the classical Fourier context, setting $H = L^2([0,1])$ and $f_n(x) = e^{2\pi i n x}$ for all $n \in \mathbb{Z}$, the Fourier series can be rewritten as $\hat{f}(n) = \langle f, f_{-n} \rangle$ (the minus arising from the conjugation in the second term of the complex inner product) and then Plancherel's formula becomes

$$f = \sum_{n \in \mathbb{Z}} \langle f, f_{-n} \rangle f_{-n} = \sum_{n \in \mathbb{Z}} \langle f, f_n \rangle f_n.$$

Viewing the Fourier series as a frame for $L^2([0,1])$ (it is in fact an ONB) with frame operator being the identity, it is self-dual, meaning that $\widetilde{f_n} = f_n$ and thus Plancherel's formula is simply the generic frame reconstruction formula for the Fourier transform.

Gabor analysis focuses on studying basis properties of discrete (or continuous) collections of functions that are time and frequency translations of each other.

Definition 2. *Given constants $u \in \mathbb{R}$ and $\eta \in \hat{\mathbb{R}}$, the **time shift** or **translation** and **frequency shift** or **modulation** operators are defined, respectively, as*

- $T_u : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad T_u f(x) = f(x-u).$
- $M_\eta : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad M_\eta f(x) = e^{2\pi i \eta x} f(x).$

*Their composition $\pi(u,\eta) = M_\eta T_u$ is called the **time-frequency shift operator**.*

*Given a fixed L^2 function g called a **Gabor window** and a discrete time-frequency set $\Gamma \subset \mathbb{R} \times \hat{\mathbb{R}}$, the associated **Gabor system** is given by*

$$(g, \Gamma) = \{\pi(u,\eta)g : (u,\eta) \in \Gamma\}.$$

*The L^2 closure of the linear span of the Gabor system is called the associated **Gabor space** and is denoted by $\mathcal{G}(g,\Gamma)$.*

The example of the Fourier ONB for $L^2([0,1])$ can be rephrased in this language as follows. Setting $g := \mathbb{1}_{[0,1]}$ and $\Gamma = \{0\} \times \hat{\mathbb{Z}}$, we have that (g,Γ) is an ONB for the Gabor space $\mathcal{G}(g,\Gamma) = L^2([0,1])$ as the pure frequency signals are just integer modulations of the characteristic function.

Definition 3. *A subset $\Lambda \subset \mathbb{R}^2$ is called a (non-degenerate) **lattice** if there exists some (invertible) matrix A such that $\Lambda = A\mathbb{Z}^2$. Its **density** is the quantity $\frac{1}{\det A}$.*

Remark 2.
- *Starting from this point we shall omit writing $\hat{\mathbb{R}}$ (and $\hat{A}$ for subsets) for the frequency axis and we will instead write $\mathbb{R}$ (or A for subsets) since the symmetry between time and frequency domains in Gabor analysis makes the overburdened notation superfluous.*

- *There are works discussing Gabor frames and bases in the general case of arbitrary discrete sets $\Gamma \subset \mathbb{R}^2$ (e.g. [FS07]), but we restrict our work to regular Gabor systems where the time-frequency set is a (non-degenerate) lattice $\Lambda \subset \mathbb{R}^2$.*

The prototypical (non-degenerate) example of a Gabor basis is the system $(\mathbb{1}_{[0,1]}, \mathbb{Z}^2)$ which forms an ONB for $L^2(\mathbb{R})$. This can be seen as a consequence of the fact that if $f \in L^2(\mathbb{R})$, then for any integer $n \in \mathbb{Z}$, the restriction $f \cdot \mathbb{1}_{[n,n+1]}$ is an element of $L^2([n, n+1])$ which admits as an ONB the Gabor system $(\mathbb{1}_{[0,1]}, \{-n\} \times \mathbb{Z}) = (\mathbb{1}_{[n,n+1]}, \{0\} \times \mathbb{Z})$.

The main practical issue with the basis $(\mathbb{1}_{[0,1]}, \mathbb{Z}^2)$ is the fact that the frequency content of the characteristic function is not well-localized with respect to the 2-norm:

$$\begin{aligned}
\widehat{\mathbb{1}_{[0,1]}}(\xi) &= \int_{\mathbb{R}} e^{-2\pi i \xi x} \mathbb{1}_{[0,1]}(x) dx \\
&= \int_0^1 e^{-2\pi i \xi x} dx \\
&= \frac{1 - e^{-2\pi i \xi}}{2\pi i \xi} \\
&= \frac{e^{\pi i \xi} - e^{-\pi i \xi}}{2\pi i \xi} e^{-\pi i \xi} \\
&= \operatorname{sinc}(\xi) e^{-\pi i \xi},
\end{aligned}$$

where the (normalized) sinc function is defined as $\operatorname{sinc}(x) = \frac{\sin \pi x}{\pi x}$.

Definition 4. *A signal $f \in L^2(\mathbb{R})$ is said to be **well-localized** in the time-frequency plane if its position and momentum are both square integrable ([Bat88]), i.e., if*

$$\int_{\mathbb{R}} |x f(x)|^2 dx \cdot \int_{\mathbb{R}} |\xi \hat{f}(\xi)|^2 d\xi < \infty.$$

Having in mind the fact that the momentum of the signal $\xi \hat{f}(\xi)$ is the Fourier transform of the derivative of f, we define the mixed Sobolev space

$$\mathbb{H}^1 := \{f \in L^2(\mathbb{R}) : x \mapsto x \cdot f(x) \in L^2(\mathbb{R}) \text{ and } \xi \mapsto \xi \hat{f}(\xi) \in L^2(\mathbb{R})\}$$

as the set of well-localized square integrable signals.

With this definition, one can see that the characteristic function $\mathbb{1}_{[0,1]}$ is well-localized in time, but not in frequency as $x \mapsto x \cdot \operatorname{sinc}(x)$ is not square integrable. The system $(\mathbb{1}_{[0,1]}, \mathbb{Z}^2)$ forms a basis for $L^2(\mathbb{R})$ at the cost of the localization of the Gabor window. The fact that this phenomenon is general, i.e., that it occurs for all windows, is the Balian-Low theorem.

Theorem 2 (Classical Balian-Low,[Bal81; Fra85; Bat88; Dau90; BHW95])**.** *Let $g \in \mathbb{H}^1$. Then the Gabor system (g, Λ) cannot form a Riesz basis for $L^2(\mathbb{R})$.*

Remark 3.

- *The theorem was stated for ONBs by Balian and Low independently, but their suggested proofs had some technical issues. The first formal proof for the ONB setting that the author is aware of belongs to Battle and has a strong physics flavor ([Bat88]). Daubechies provided a 'contemporary' proof that extended the result to Riesz bases as well ([Dau90]).*
- *In order for any window to have the chance to form a Riesz basis with time-frequency shifts from a lattice, the density of that lattice must necessarily be* 1 *(Corollary 7.5.2 in [Grö01]). Restricting to* Λ *with density* 1*, the system* (g, Λ) *cannot even be a frame if* g *is well-localized (due to Ron-Shen duality, Theorem 7.4.3 in [Grö01], see for instance Corollary 8.4.3 in [Grö01]).*

Fundamentally, the Balian-Low theorem states that 'too much regularity' of the window prevents the Gabor system from forming a basis for the space of square integrable signals. There is a second well-established meaning of 'regularity' in this context in the literature.

Theorem 3 (Amalgam Balian-Low, [BHW95])**.** *Let* $W(C_0, \ell^1)$ *be the Wiener amalgam space*

$$W(C_0, \ell^1) := \{f \in L^2(\mathbb{R}) : \|f\|_{\infty,1} = \sum_{n\in\mathbb{Z}} \|f \cdot \mathbb{1}_{[n,n+1]}\|_\infty < \infty, f \text{ continuous}\}$$

If the Gabor system $(g, a\mathbb{Z} \times b\mathbb{Z})$ *forms a Riesz basis for* $L^2(\mathbb{R})$*, then* $g \notin W(C_0, \ell^1)$ *and* $\hat{g} \notin W(C_0, \ell^1)$*.*

Remark 4. *Due to the ubiquitous nature of the Feichtinger algebra* $\mathcal{S}_0$ *in harmonic analysis ([Fei81], p.246 of [Grö01]),*

$$\mathcal{S}_0 = \{f \in L^2(\mathbb{R}) : V_f f(u, \eta) = \langle f, \pi(u, \eta) f\rangle \in L^1(\mathbb{R}^2)\},$$

the amalgam Balian-Low theorem is usually stated in the following, slightly weaker way: if $g \in \mathcal{S}_0$*, then* (g, Λ) *cannot be a Riesz basis for* $L^2(\mathbb{R})$*. Note that the Feichtinger algebra is invariant under the Fourier transform, i.e.,* $f \in \mathcal{S}_0$ *if and only if* $\hat{f} \in \mathcal{S}_0$*. Additionally,* $\mathcal{S}_0 \subset W(C_0, \ell^1)$*.*

A crucial observation made by Cabrelli, Molter and Pfander was that amalgam regularity of the Gabor window not only prevents the Gabor system from being a basis for $L^2(\mathbb{R})$, but also limits the 'size' of the associated Gabor space for which the system is basic. More precisely, the Gabor space cannot contain any time-frequency shifts of the window that are not already in the lattice giving rise to the Gabor system.

Theorem 4 ([CMP16])**.** *Let* $\Lambda \subset \mathbb{R}^2$ *be a rational density lattice and let* $g \in \mathcal{S}_0$*. If* (g, Λ) *forms a Riesz basis for the proper subspace* $\mathcal{G}(g, \Lambda) \subset L^2(\mathbb{R})$*, then* $\pi(u, \eta)g \notin \mathcal{G}(g, \Lambda)$ *whenever* $(u, \eta) \notin \Lambda$*.*

This result raised two clear questions.

Problem 1. *Does classical regularity (i.e., $g \in \mathbb{H}^1$) also prevent the Gabor space from containing additional time-frequency shifts of the window?*

Problem 2. *Once one looks at Gabor systems (g, Λ) forming Riesz sequences (i.e., the associated Gabor spaces are proper closed subspaces of $L^2(\mathbb{R})$), the density of the lattice Λ can in principle take any value smaller than 1. The use of a winding number argument on the Zak transform of the window in the proof of Theorem 4 is contingent on the density of Λ being rational. Is this just a proof artifact or does Theorem 4 fail if the density of Λ is allowed to be irrational? Can one drop the rational density condition in the classical regularity setting?*

These two questions are the motivation for the three papers presented in this section.

3.2 A brief overview of the three papers

Broadly speaking, the main results contained in the three papers are as follows:

- **Paper 1** solves Problem 1 in the affirmative.
- **Paper 2** extends and uses the results in **Paper 1** to show that if $g \in \mathbb{H}^1$ and (g, Λ) is a Riesz sequence, then the distance between $\pi(u, \eta)g$ and $\mathcal{G}(g, \Lambda)$ (measured in $L^2(\mathbb{R})$) is proportional to the Euclidean distance between the time-frequency pair (u, η) and the lattice Λ.
- **Paper 3** settles the first part of Problem 2. We show that the rational density condition on Λ is indeed superfluous if $g \in \mathcal{S}_0$.

Before discussing each paper in more detail, we provide some more general remarks and context about the work done.

The author of the present thesis was first introduced to the topic and the paper [CMP16] and the two arising questions by his advisor, G. Pfander, during the author's Master's degree. The main result of the author's Master thesis was to show that a relaxation of the amalgam regularity condition (namely that the Zak transform, formally introduced in Definition 5 below, of the window is assumed to be continuous on almost every section) does not necessarily lead to subspace Balian-Low phenomena. Explicitly, examples were constructed of Gabor windows with almost every section continuous Zak transform that formed Riesz sequences with lattices of the form $\mathbb{Z} \times P\mathbb{Z}$, for some natural P, such that the associated Gabor spaces contained additional time-frequency shifts. A full characterization of the permissible time and frequency coordinates of the additional shits was also included.

Therefore, the continuation of this project into the PhD period of the author's studies was natural and constituted the main focus of the initial phase of the PhD. The development of the last two papers was partially covered by the author's involvement in the DFG project PF 450/11-1. The team working on the project grew organically by the author discussing the problem and partial

solutions and ideas with the other members of his advisor's workgroup. Virtually all of the main ideas present in the final form of the three papers were obtained during group discussions over the blackboard (a notable exception that springs to mind is the use of the Walnut representation in Step 1 of the proof of Theorem 8 in **Paper 3**, an idea that came from F. Voigtlaender and D.G. Lee; another exception would be one of the equivalent conditions of Theorem 11 in **Paper 3** that was the idea of F. Philipp).

The three papers are presented in chronological order of their development. **Paper 1** and **Paper 2** were relatively close chronologically speaking. Some time elapsed between the formulation of Theorem 11 in **Paper 3** and the final version of that manuscript. The essential idea of looking at irrational rotation algebras to make use of Theorem 11 was suggested to us by K. Gröchenig at the SampTA conference in 2019.

3.2.1 Paper 1

As mentioned before, the key tool for the amalgam subspace Balian-Low theorem in [CMP16] was the Zak transform.

Definition 5. *Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. The **Zak transform** (Chapter 8.1 in [Grö01]) is almost everywhere defined on $\mathbb{R}^2$ as*

$$Zf(x,\omega) = \sum_{k\in\mathbb{Z}} f(x+k)e^{-2\pi ik\omega},$$

with L^2 convergence.

The definition extends to a unitary operator $Z : L^2(\mathbb{R}) \to L^2([0,1]^2)$ due to the quasi-periodicity of the Zak transform,

$$Zf(x+m,\omega+n) = e^{2\pi im\omega} Zf(x,\omega) \quad \text{for all } (x,\omega) \in \mathbb{R}^2 \text{ and all } m,n \in \mathbb{Z}.$$

The main properties of the Zak transform are collected in Lemma 2.2 (of **Paper 1**) and an exploration of the symmetries of the Zak transform in the context of a Gabor system can be found in Section 2.

On a rational density lattice, the condition of the Gabor space containing a time-frequency shift of the Gabor window can be equivalently rewritten as a functional equation involving the Zak transform and an additional bivariate function h (see Step 3 in the proof of the main theorem of [CMP16] or the proof of Proposition 4.1 in **Paper 1**). Crucially, due to the nature of this functional equation, the helper function h inherits the regularity of the Zak transform.

In the case of the window g being in the Feichtinger algebra $\mathcal{S}_0$, the Zak transform Zg is easily seen to be continuous. This implies that the helper function h is also continuous and allowed the authors of [CMP16] to use a winding number argument (Proposition 3 in [CMP16]) to derive divisibility restrictions on the time-frequency parameters of any shift $\pi(u,\eta)g$ contained in the Gabor space $\mathcal{G}(g,\Lambda)$. These restrictions ruled out any time-frequency pairs not already contained in the lattice Λ.

In the classical Balian-Low setting of $g \in \mathbb{H}^1$, the major difficulty to overcome was the fact that the Zak transform Zg is not necessarily continuous.

The key find to address this issue was the work of Gautam ([Gau08]). To our knowledge, he was the first to realize that the correct regularity condition on the Zak transform of a well-localized window g was local vanishing mean oscillation, i.e., that $g \in \mathbb{H}^1$ implies $Zg \in VMO_{loc}(\mathbb{R}^2)$ (a detailed discussion of VMO functions can be found in Section 3 of **Paper 1**). In fact Gautam proved the following version of the Balian-Low theorem.

Theorem 5 ([Gau08])**.** *Let $g \in L^2(\mathbb{R})$. If the Gabor system $(g, \mathbb{Z}^2)$ forms a Riesz basis for $L^2(\mathbb{R})$, then $Zg \notin VMO_{loc}(\mathbb{R}^2)$. Moreover, $Zg \notin VMO_{loc}(\mathbb{R}^2)$ implies that for any $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ we have*

$$\int_{\mathbb{R}} |x|^p |g(x)|^2 dx \cdot \int_{\mathbb{R}} |\xi|^q |\hat{g}(\xi)|^2 d\xi = \infty.$$

The main result of **Paper 1** is a generalization of this and solves in slightly more generality Problem 1.

Theorem 6 (Theorem 1.4 in **Paper 1**)**.** *Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a rational density lattice. If the system (g, Λ) is a Riesz sequence and if $\pi(u, \eta)g \in \mathcal{G}(g, \Lambda)$ for some $(u, \eta) \notin \Lambda$, then for any $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ we have*

$$\int_{\mathbb{R}} |x|^p |g(x)|^2 dx \cdot \int_{\mathbb{R}} |\xi|^q |\hat{g}(\xi)|^2 d\xi = \infty.$$

Conceptually, the proof of Theorem 6 follows along the steps of the proof of the main theorem of [CMP16], by first reducing to the case of $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$ with P, Q coprime integers using symplectic operators and then writing a characterization of an additional time-frequency shift in the Gabor space as a functional equation involving Zg and a helper function h. The winding number argument for $VMO_{loc}(\mathbb{R}^2)$ functions (like h) is given in Proposition 3.6 and relies on approximating both the Zak transform and h with continuous counterparts.

An interesting additional result obtained is the following.

Theorem 7 (Theorem 1.5 in **Paper 1**)**.** *Let $g \in L^2(\mathbb{R})$ and let $\Lambda = A\mathbb{Z}^2$ for some $A \in GL(2, \mathbb{Q})$. If (g, Λ) is a Riesz sequence and $\pi(u, \eta)g \in \mathcal{G}(g, \Lambda)$ for some pair $(u, \eta) \notin \Lambda$, then $Zg \notin VMO_{loc}(\mathbb{R}^2)$.*

Considering the fact that either $g \in \mathcal{S}_0$ or $g \in \mathbb{H}^1$ imply $Zg \in VMO_{loc}(\mathbb{R}^2)$, this result provides a unified framework for Balian-Low subspace phenomena (both classical and amalgam regularity assumptions) under the condition that the lattice Λ not only has rational density, but consists only of rational vectors. This additional restriction (concerning the lattice) stems from the fact that, as far as we were able to prove, the space of vanishing oscillation functions is not fully invariant under symplectic operators (a more detailed discussion of this can be found in Proposition 3.11 of **Paper 1**).

3.2.2 Paper 2

This work can be viewed as a quantitative extension of **Paper 1** in the case of classical regularity. The motivation is loosely based on an OFDM communication viewpoint.

Assume a signal $f \in \mathcal{G}(g, \Lambda)$ is transmitted through a linear channel modeled by a simple time-frequency shift operator $\pi(u, \eta)$. The task of the receiver is to recover the original signal f from the distorted signal $\pi(u, \eta)f$ using the canonical dual frame reconstruction formula. Due to the Balian-Low theorem, if g is well-localized in time and frequency (which is often desirable), it is possible that the distorted signal $\pi(u, \eta)f$ will fall outside the Gabor space $\mathcal{G}(g, \Lambda)$ and therefore the receiver will only have access to $P_{\mathcal{G}}\pi(u, \eta)f$ (the projection of the distorted signal onto the Gabor space) instead of the actual distorted signal. The question in this toy model therefore is, assuming the original signal was the window itself, can the off-band energy loss $\|(Id - P_{\mathcal{G}})\pi(u, \eta)g\|_{L^2(\mathbb{R})}$ be quantified?

This is answered in the affirmative by the main result of **Paper 2**.

Theorem 8 (Theorem 1.3 in **Paper 2**). *Let $g \in \mathbb{H}^1$ and let $\Lambda \subset \mathbb{R}^2$ be a rational density lattice such that the system (g, Λ) forms a Riesz sequence. Then there exist constants $C_1, C_2 > 0$ such that for all $(u, \eta) \in \mathbb{R}^2$*

$$C_1 \cdot dist((u, \eta), \Lambda) \leq dist(\pi(u, \eta)g, \mathcal{G}(g, \Lambda)) \leq C_2 \cdot dist((u, \eta), \Lambda),$$

where the outer distances are Euclidean and the inner distance is measured by the L^2 norm.

The proof of this result relies on the observation that the time-frequency map $S_g : \mathbb{R}^2 \to L^2(\mathbb{R})$ given by $S_g(u, \eta) = \pi(u, \eta)g$ is (Fréchet) differentiable if $g \in \mathbb{H}^1$ with the derivative at $(0, 0)$ given by $(u, \eta) \mapsto -ug' + 2\pi i\eta Xg$, where X is the position operator mentioned before $(Xf(x) = xf(x))$. This is discussed in Lemma 3.2 (of **Paper 2**).

Section 2 (of **Paper 2**) recasts the distances involved in Theorem 8 in terms of matrix operators using the regularity of the Zak transform of the window and Section 4 shows that the function $-ug' + 2\pi i\eta Xg$ cannot lie in the Gabor space when $(u, \eta) \neq (0, 0)$ under the given assumptions, based on Theorem 6.

This establishes the estimate of Theorem 8 in a small neighborhood of the lattice Λ and then a compactness argument together with Theorem 6 is used to extend this to the rest of the time-frequency plane.

3.2.3 Paper 3

As previously pointed out, the fundamental tool for the main results of both [CMP16] and **Paper 1** was rewriting the condition of $\pi(u, \eta)g \in \mathcal{G}(g, \Lambda)$ in terms of a functional equation involving the Zak transform Zg and a helper function h. Exploiting symmetries of the Zak transform and the lattice, one could cancel out the Zak transform and be left with a functional equation involving just h. Through this equation, the regularity of h was sufficient to derive constraints on

the time-frequency parameters (u, η) that ruled out time-frequency shifts not already contained in the lattice Λ.

Crucially, the step of canceling out the Zak transform relies on the lattice Λ being symplectically equivalent to a lattice with rational vectors. Since symplectic operators preserve the density of the lattice, it is clear that this technique cannot be extended to general lattices Λ (i.e., lattices with irrational density).

Initially, we were only able to prove a different characterization of $\pi(u, \eta)g \in \mathcal{G}(g, \Lambda)$ independent of Zak transform methods, which is contained in Theorem 11 (of **Paper 3**). The main component of this characterization (condition (iv) in Theorem 11) was a new biorthogonal relation involving the window, the dual window and the adjoint lattice.

Interestingly, this was sufficient for proving that Problem 2 is true in the special case of a Gaussian window.

The generalization to arbitrary $g \in \mathcal{S}_0$ (Theorem 1 of **Paper 3**) was made possible by the suggestion of K. Gröchenig to look at the traces of projections on the irrational rotation algebra generated by the translation and modulation operators of a time-frequency shift contained in the Gabor space (a precise statement is given in Theorem 14 and discussed in the Appendix of **Paper 3**).

3.3 Additional work done by the author

The author of this thesis would like to bring forth two results he worked on that were rendered superfluous by the eventual technical development of the papers discussed here. The author considers these results too minor to warrant an independent publication.

The first result is a technical characterization of VMO functions that was motivated by what can only be described as a case of 'group miscomputation'. During the early phases of the development of **Paper 1**, we struggled with showing that the space of locally vanishing mean oscillation functions was closed under multiplication and under taking arithmetic inverses (assuming the original function is bounded away from 0). The characterization gave rise to a proof of these facts, although in a unnecessarily technically involved way (as can be seen from Lemma 3.3 and Corollary 3.4 of **Paper 1**).

The second result was the simplified toy-model case that inspired the general interplay between the differential equations on the Zak transform and the distance bounds obtained in **Paper 2**. Due to the restricted nature of this simple case, the proof is much more straightforward.

3.3.1 A characterization of VMO functions

As stated before, the inspiration for finding a technical characterization of VMO functions was the need to prove that the function space VMO_{loc} was closed under multiplication and taking (arithmetic) inverses of functions bounded away from 0. We include the relevant definitions here. Theorem 9 characterizes vanishing mean oscillation functions in terms of the 'non-separatness' of their ranges and Propositions 1 and 2 prove the desired closedness results.

The much more natural Lemma 3.3 in **Paper 1** ended up replacing Theorem 9 as the simpler tool to establish closedness results. Corollary 3.4 in **Paper 1** made the initial purpose of this side note obsolete (i.e. Propositions 1 and 2), but Theorem 9 might still present some technical interest on its own.

Definition 6. *An L^1_{loc} (i.e., locally integrable) function f is said to have vanishing mean oscillation if*

$$\lim_{a\downarrow 0} \sup_{|Q|<a} \tfrac{1}{|Q|} \int_Q |f(x) - f_Q| dx = 0.$$

Here Q is allowed to be any finite length open interval and $|Q|$ is the Lebesgue measure of Q. Additionally, f_Q is the average value of the function f over such an interval Q, i.e.

$$f_Q = \tfrac{1}{|Q|} \int_Q f(x) dx.$$

Then the space VMO is defined as

$$VMO = \{f \in L^1_{loc} : f\, has\, vanishing\, mean\, oscillation\}.$$

Definition 7. *We say that an $f \in L^\infty$ has locally separable range if there exist $r, c > 0$ and there exists a sequence of triplets of sets $\{(Q_n, U_n, V_n)\}_{n\in\mathbb{N}}$ satisfying Q_n is an interval with $|Q_n| < \frac{1}{n}$ and $U_n, V_n \subset Q_n$ such that the following two conditions hold simultaneously for all n:*

1. $\inf_{x\in U_n, y\in V_n} |f(x) - f(y)| \geq c$,
2. $min(\frac{|U_n|}{|Q_n|}, \frac{|V_n|}{|Q_n|}) \geq r$.

Remark 5. *It is important to point out that the second condition makes this definition not trivial i.e., not trivially satisfied by pretty much any function. Indeed, without the second condition, one could simply take any $r, c > 0$ and any sequence of intervals $\{Q_n\}_n$ and simply set $U_n, V_n = \varnothing$ for all n. Equivalently, since we are dealing with equivalence classes of functions defined up to zero sets, one could also take U_n, V_n to be arbitrary null subsets of Q_n.*

Some typical examples of functions that have locally separable range would be step functions or functions defined as the characteristic functions of fat Cantor sets.

Theorem 9. *Let $f \in L^\infty$. Then $f \notin VMO$ if and only if f has locally separable range.*

Proof. $\Longleftarrow$ Since f has locally separable range we know there exist $r, c > 0$ and a sequence $\{(Q_n, U_n, V_n)\}_{n\in\mathbb{N}}$ as in the definition.

Pick $Q := Q_n$ arbitrarily from the family of intervals and let f_Q be the average of f on Q. For simplicity of notation we shall also set $U := U_n$ and $V := V_n$. By triangle inequality it follows trivially that

$$\max(dist(f_Q, f(U)), dist(f_Q, f(V))) > \tfrac{c}{3}$$

since otherwise we would have that $dist(f(U), f(V)) < c$. Without loss of generality, let us assume that $\inf_{x\in U} |f(x) - f_Q| = dist(f_Q, f(U)) > \frac{c}{3}$.

Then we have that

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(x) - f_Q|dx &\geq \frac{1}{|Q|}\int_U |f(x) - f_Q|dx \\ &> \frac{1}{|Q|}\int_U \frac{c}{3}dx \\ &= \frac{c}{3}\frac{|U|}{|Q|} \\ &\geq \frac{cr}{3}.\end{aligned}$$

Since $n \in \mathbb{N}$ was chosen arbitrarily, it is obvious that f cannot have vanishing mean oscillation.

$\implies$ Let $\|f\| := \|f\|_\infty < \infty$ be the essential supremum of f and assume that f does not have vanishing mean oscillation. Then there exists a constant $c > 0$ and a sequence of intervals $\{Q_n\}_{n\in\mathbb{N}}$ such that, for any n, $\frac{1}{|Q_n|}\int_{Q_n} |f(x) - f_{Q_n}|dx \geq c$ and $|Q_n| < \frac{1}{n}$. This of course in particular implies that $\|f\| > 0$.

We partition each of the intervals Q_n into $S_n := \{x \in Q_n : |f(x) - f_{Q_n}| \geq \frac{c}{2}\}$ and $T_n := \{x \in Q_n : |f(x) - f_{Q_n}| < \frac{c}{2}\}$.

Therefore we have that

$$\begin{aligned}c &\leq \frac{1}{|Q_n|}\int_{Q_n} |f(x) - f_{Q_n}|dx \\ &= \frac{1}{|Q_n|}\int_{S_n} |f(x) - f_{Q_n}|dx + \frac{1}{|Q_n|}\int_{T_n} |f(x) - f_{Q_n}|dx \\ &\leq \frac{1}{|Q_n|}\int_{S_n} 2\|f\|dx + \frac{1}{|Q_n|}\int_{T_n} \frac{c}{2}dx \\ &\leq 2\|f\|\frac{|S_n|}{|Q_n|} + \frac{c}{2},\end{aligned}$$

and then that $\frac{|S_n|}{|Q_n|} > \frac{c}{4\|f\|}$ for all n.

Now let $W_n := \{x \in Q_n : |f(x) - f_{Q_n}| < \frac{c}{4}\}$ and suppose that $\lim_{n\to\infty} \frac{|W_n|}{|Q_n|} \neq 0$. Then, by possibly passing to a subsequence of intervals, we can find a uniform bound $t > 0$ such that $\frac{|W_n|}{|Q_n|} > t$ for all n. Now we can simply set $U_n = S_n, V_n = W_n$ and the function f has locally separable range with constants $r = min(\frac{c}{4\|f\|}, t)$ and $\frac{c}{4}$.

So we may assume without loss of generality that $\lim_{n\to\infty} \frac{|W_n|}{|Q_n|} = 0$. Now let A_n be the annulus in the complex plane centered at f_{Q_n} with inner radius $\frac{c}{4}$ and outer radius $2\|f\|$. The previous limit being 0 implies that for n large enough we have that $\frac{|f^{-1}(A_n)|}{|Q_n|} \geq \frac{24}{25}$ and that $\frac{|W_n|}{|Q_n|} \leq \frac{1}{25}$.

Now let us partition the annulus A_n into 6 congruent slices of angular width $\frac{\pi}{3}$ and consider the preimages under f of these slices. At least one of them, which we shall denote by U_n, satisfies, by the pigeonhole principle, the inequality $\frac{|U_n|}{|f^{-1}(A_n)|} \geq \frac{1}{6}$ and therefore $\frac{|U_n|}{|Q_n|} \geq \frac{4}{25}$. Without loss of generality we can rotate

the coordinates to make sure that the image $f(U_n)$ corresponds to the slice of the annulus spanning between the angles $-\frac{\pi}{6}$ and $\frac{\pi}{6}$.

Let V_n be the preimage under f of the half of the annulus opposite to $f(U_n)$ i.e., the left half spanning between the angles $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ or the half obtained by removing $f(U_n)$ and its two immediately adjacent slices.

Let us denote by Q_n^+ the preimage of the right half plane centered at f_{Q_n} and by Q_n^- that of the left half plane. Then in view of the fact that f_{Q_n} is the average of the function we have that $\int_{Q_n^+} f(x) - f_{Q_n} dx = -\int_{Q_n^-} f(x) - f_{Q_n} dx$ and therefore that $|\int_{Q_n^+} \Re(f(x) - f_{Q_n})dx| = |\int_{Q_n^-} \Re(f(x) - f_{Q_n})dx|$, where $\Re(z)$ is the real part of the complex number z.

Now clearly $|\int_{Q_n^+} \Re(f(x) - f_{Q_n})dx| \geq |\int_{U_n} \Re(f(x) - f_{Q_n})dx| \geq \frac{c}{4}\cos\frac{\pi}{6}|U_n| = \frac{\sqrt{3}c}{8}|U_n|$.

On the other hand, we also have that

$$\begin{aligned}
|\int_{Q_n^-} \Re(f(x) - f_{Q_n})| &\leq |\int_{V_n} \Re(f(x) - f_{Q_n})dx| + |\int_{W_n} \Re(f(x) - f_{Q_n})dx| \\
&\leq \int_{V_n} 2\|f\|dx + \int_{W_n} \tfrac{c}{4}dx \\
&= 2\|f\||V_n| + \tfrac{c}{4}|W_n|.
\end{aligned}$$

Combining the two inequalities and dividing by $|Q_n|$ we obtain that

$$\begin{aligned}
\tfrac{\sqrt{3}c}{50} &\leq \tfrac{\sqrt{3}c}{8}\tfrac{|U_n|}{|Q_n|} \\
&\leq 2\|f\|\tfrac{|V_n|}{|Q_n|} + \tfrac{c}{4}\tfrac{|W_n|}{|Q_n|} \\
&\leq 2\|f\|\tfrac{|V_n|}{|Q_n|} + \tfrac{c}{100},
\end{aligned}$$

so $\frac{|V_n|}{|Q_n|} \geq \frac{(2\sqrt{3}-1)c}{200\|f\|} > \frac{c}{100\|f\|}$.

Moreover, it is easy to see that $dist(f(U_n), f(V_n)) \geq \frac{c}{4}$ and therefore we have that f has locally separable range with constants $r = min(\frac{4}{25}, \frac{c}{100\|f\|})$ and $\frac{c}{4}$ thus completing the proof. □

Proposition 1. *The space $L^\infty \cap VMO$ is closed under multiplication.*

Proof. Let $f, g \in L^\infty \cap VMO$ and assume that $fg \notin VMO$. Then fg must have locally separable range so there exist constants $r, c > 0$ and a sequence $\{(Q_n, U_n, V_n)\}$ as in the definition. In particular, for any n we have that $|f(x)g(x) - f(y)g(y)| \geq c$ for all $x \in U_n, y \in V_n$.

From triangle inequality it follows that $c \leq |f(x)||g(x) - g(y)| + |g(y)||f(x) - f(y)| \leq \|f\||g(x) - g(y)| + \|g\||f(x) - f(y)|$. Therefore we must have that for all $x \in U_n, y \in V_n$ either $|f(x) - f(y)| \geq \frac{c}{2\|g\|}$ or that $|g(x) - g(y)| \geq \frac{c}{2\|f\|}$.

Let $t := \frac{c}{2max(\|f\|,\|g\|)} > 0$. Then for any $x \in U_n, y \in V_n$ we have that

$$\begin{aligned}
t &\leq |f(x) - f(y)| + |g(x) - g(y)| \\
&\leq |f(x) - f_{Q_n}| + |f(y) - f_{Q_n}| + |g(x) - g_{Q_n}| + |g(y) - g_{Q_n}|.
\end{aligned}$$

Now we can integrate over U_n and V_n (all functions are essentially bounded and the domains of integration have finite measure) to get

$$\begin{aligned}
t|U_n||V_n| &\leq \int_{U_n}\int_{V_n}(|f(x)-f_{Q_n}|+|f(y)-f_{Q_n}|+|g(x)-g_{Q_n}|+|g(y)-g_{Q_n}|)dydx \\
&= |V_n|\int_{U_n}|f(x)-f_{Q_n}|dx+|U_n|\int_{V_n}|f(y)-f_{Q_n}|dy \\
&+ |V_n|\int_{U_n}|g(x)-g_{Q_n}|dx+|U_n|\int_{V_n}|g(y)-g_{Q_n}|dy \\
&\leq max(|U_n|,|V_n|)(\int_{U_n\cup V_n}|f(x)-f_{Q_n}|dx+\int_{U_n\cup V_n}|g(x)-g_{Q_n}|dx) \\
&\leq max(|U_n|,|V_n|)(\int_{Q_n}|f(x)-f_{Q_n}|dx+\int_{Q_n}|g(x)-g_{Q_n}|dx).
\end{aligned}$$

Therefore we must have that $\frac{1}{|Q_n|}\int_{Q_n}|f(x)-f_{Q_n}|dx+\frac{1}{|Q_n|}\int_{Q_n}|g(x)-g_{Q_n}|dx \geq t\frac{min(|U_n|,|V_n|)}{|Q_n|} \geq tr$ for any n. But this immediately implies that f and g cannot simultaneously have vanishing mean oscillation, a contradiction. □

Proposition 2. *Let $f \in L^\infty \cap VMO$ be essentially bounded away from 0. Then also $\frac{1}{f} \in VMO$.*

Proof. Assume $\frac{1}{f}$ has locally separable range. Then we can find constants $r, c > 0$ and a sequence $\{(Q_n, U_n, V_n)\}_{n\in\mathbb{N}}$ as in the definition.

Then for any n we will have that $|\frac{1}{f(x)}-\frac{1}{f(y)}| \geq c$ for any $x \in U_n, y \in V_n$. But this immediately implies that $|f(x)-f(y)| \geq c\|\frac{1}{f}\|^{-2}$, so f has locally separable range with constants r and $c\|\frac{1}{f}\|^{-2}$, a contradiction. □

3.3.2 A toy-model case: the derivative of the time-frequency map does not lie in the Gabor space

Proposition 3. *Consider the lattice $\Lambda = \mathbb{Z}\times 2\mathbb{Z}$ and let $g \in \mathbb{H}^1$ such that (g,Λ) is a Riesz sequence. Then $g' \notin \mathcal{G}(g,\Lambda)$.*

Sketch of proof. First, note that $g' = -(-1)g' + 2\pi i 0Xg$ is the derivative of the time-frequency map evaluated at $(-1, 0)$.

Step 1 The Riesz sequence condition implies that the inequality

$$2A \leq |Zg(x,\omega)|^2 + |Zg(x+\tfrac{1}{2},\omega)|^2 \leq 2B \tag{1}$$

holds almost everywhere with A, B being the Riesz sequence constants (see for example Lemma 2.3 of **Paper 1** or the discussion on page 6 of [CMP16]).

Step 2 Assume that $g' \in \mathcal{G}(g,\Lambda)$. Then there exists some locally square integrable bivariate function h that is $\frac{1}{2}$ periodic in the first coordinate and 1 periodic in the second coordinate such that

$$\partial_1 Zg(x,\omega) = h(x,\omega)Zg(x,\omega),$$

where ∂_1 is the derivative with respect to the time coordinate (see for example Lemma 2.4 of **Paper 2**).

Step 3 Solving the elementary PDE leads to

$$Zg(x,\omega) = Zg(0,\omega)e^{\int_0^x h(s,\omega)ds}.$$

Plugging this solution into equation (1) gives that

$$0 < 2A \leq |Zg(0,\omega)|^2\Big(e^{2\Re\int_0^x h(s,\omega)ds} + e^{2\Re\int_0^{x+1/2} h(s,\omega)ds}\Big),$$

which implies that $Zg(0,\omega) \neq 0$ for a.e. x. Moreover, since the complex exponential is never 0, this also gives us that $Zg(x,\omega) \neq 0$ for a.e. x.

With this in mind, we can compute the absolute value of $|Zg(x+1,\omega)|$ in two ways.

On the one hand, using the quasiperiodicity of the Zak transform with respect to the time coordinate,

$$|Zg(x+1,\omega)| = |e^{2\pi i\omega}Zg(x,\omega)| = |Zg(x,\omega)|. \tag{2}$$

On the other hand, making use of the solution to the PDE and the $\frac{1}{2}$ periodicity of h with respect to the time coordinate,

$$|Zg(x+1,\omega)| = |Zg(0,\omega)e^{\int_0^{x+1} h(s,\omega)ds}| = |Zg(x,\omega)| \cdot |e^{\int_0^{1/2} h(s,\omega)ds}|^2. \tag{3}$$

Combining equations (2) and(3), and since $|Zg(x,\omega)| \neq 0$ for a.e. x, gives

$$|e^{\int_0^{1/2} h(s,\omega)ds}| = 1.$$

Now we can write, using the $\frac{1}{2}$ time-periodicity of h once again,

$$|Zg(x+\tfrac{1}{2},\omega)| = |Zg(0,\omega)e^{\int_0^x h(s,\omega)ds}| \cdot |e^{\int_0^{1/2} h(s,\omega)ds}| = |Zg(x,\omega)|, \tag{4}$$

which holds almost everywhere.

Step 4 Plugging equation (4) into equation (1) implies that the Zak transform is essentially bounded above and below, i.e., that $A \leq |Zg(x,\omega)|^2 \leq B$ holds almost everywhere.

But this implies that the Gabor system $(g, \mathbb{Z}\times\mathbb{Z})$ is a frame for $L^2(R)$ (see Corollary 8.3.2 in [Grö01]), which contradicts the classical Balian-Low theorem as $g \in \mathbb{H}^1$. □

4 Neural network approximation

This section is based on the following two papers:

- 'Neural network approximation and estimation of classifiers with classification boundary in a Barron class' (co-authored with P. Petersen and F. Voigtlaender), recently accepted for publication in Annals of Applied Probability, henceforth referred to as **Paper 4**, [CPV20];
- 'Quantitative Approximation Results for Complex-Valued Neural Networks' (co-authored with D.G. Lee, J. Maly, G. Pfander and F. Voigtlaender), published in SIAM Journal on Mathematics of Data Science in 2022, henceforth referred to as **Paper 5**, [Car+22].

The section is organized as follows. Subsection 1 shortly addresses the current role of neural networks in the framework of machine learning and discusses the state of the art and some of the open questions in the field. Subsections 2 and 3 frame **Paper 4** and **Paper 5** in the context of the main motivation behind them, i.e., avoiding the curse of dimensionality and complex variable networks, respectively, and provide a summary of the main results and techniques of the two papers as well as details about the author's contribution. Subsection 4 is a brief note on some other neural network papers and collaborations the author was involved in in a more minor capacity.

4.1 Machine learning and neural networks

On a fundamental level, *statistical learning* is the field concerned with constructing and fine-tuning mathematical models to fit given numerical data or empirical observations and measurements with the goal of maximizing the predictive power of the obtained models. Classical examples range from empirical derivations of physical laws and constants from experimental data via regression or interpolation to classification via separating surfaces using support vector machines with non-linear kernel (feature) maps (see for example Chapters 5 and 6 of [MRT18]).

In this context, *machine learning* is the application of computation resources to aid in the development and testing of the models of statistical learning. The impact the development of computing technologies has had on statistical learning cannot be overstated. As a historical anecdote, it bears pointing out that the orbital calculations performed by a small army of highly trained engineers at NASA during the Apollo program less than 60 years ago can now be done virtually instantaneously and with no errors on an inexpensive commercially available laptop. However, the massive impact and adoption of machine learning in more and more areas of everyday life cannot be explained purely by the speedup of processing data. The development of computation technology has opened the door to modeling data that is intractable to human analysis, either due to the sheer amount of data involved or due to the complexity of the models needed to make predictions based on the data.

The current machine learning revolution, while technically predicated on the evolution of computing solutions, is inextricably correlated to the development of the theory of neural networks (and in particular deep neural networks). The widespread adoption of neural networks in concrete practical implementations in a variety of areas, ranging from classification and pattern recognition to numerically solving differential equations [EY18; HSN20; Pfa+20] is a complex phenomenon with multiple likely reasons:

- Neural networks of different architectures offer a significant amount of flexibility and adaptability to concrete modeling scenarios. Historically, convolutional nets have been employed in image classification problems with state of the art defining success [KSH17; He+15]. Recurrent networks form the backbone of many language processing, speech recognition and automated translation applications (e.g. [SVL14]). Graph neural networks have recently been used to efficiently solve typical big data graph problems in areas ranging from medicine and biology [Sto+20; Duv+15] to social media [Mon+19].

- From a theoretical perspective, the simplest architecture of networks, *feed forward neural networks*, has provided a rich research field with numerous applications in classical function analytic scenarios, like approximation spaces [Gri+22; GV21] and finding numerical solutions for differential equations [GH21; Elb+22].

- Hypothesis classes of fixed architecture neural networks usually exhibit a large degree of expressivity. This is a classical field of study of neural networks with quite a few universality theorems showing that specific architectures yield dense subspaces of common function spaces [Les+93; Yar22; HS17]. This permits the approximation of complicated models with comparatively simpler neural networks.

- Expressivity typically does not come at the cost of unquantifiable or 'bad' information theoretic complexity of the chosen hypothesis class of neural networks. There are numerous works bounding the complexity of neural network hypothesis classes in terms of classic measurement tools like VC dimension [KM97; Bar+19], Radamacher complexity (e.g. [GRS18]) or entropy numbers (e.g. [GV21]). The existence of these bounds allows for guaranteed loss bounds for typical learning algorithms (e.g. empirical loss minimization).

Despite the many positives, the field of study of neural networks is still relatively young and there are still numerous open questions and poorly understood phenomena. To name a few:

- There still exists a large gap between the practical results of neural networks and most theoretical guarantees. Moreover, the extent to which neural networks can be used is intuitively atypical, considering the fact

that, historically, optimal algorithmic solutions to fundamentally different problems usually implied very different hypothesis classes of models. A more fundamental explanation of the ubiquity of neural networks as optimal algorithmic solvers is one of the holy grails of the field.

- Function spaces of realizations of given network architectures tend to exhibit exotic properties (e.g. 'very high' non-convexity, [PRV21]) and the weights responsible for achieving a specific realization are highly non-unique, implying that a neural network realization viewed as a function from the space of weights to its outputs is typically very badly conditioned.

- There is no clear overarching theory for the existence of adversarial examples in classification problems. It is well documented in practice [Bro+17; Eyk+18; TVG19] that correctly classified images that are very slightly perturbed (even imperceptibly so to most observers) can lead to erroneous classification (such perturbations are what are referred to as adversarial examples [Sze+13; GSS14]). A major objective of practical implementations of neural networks where safety is a concern is the development of training algorithms that guarantee robustness to adversarial examples.

- Commonly, in image classification problems, although a specific architecture and training algorithm can be chosen to provide state of the art generalization error bounds, the computational complexity of the model explodes as the dimension of the input grows. This is known as the *curse of dimensionality*. Obviously, for practical applications, hypothesis classes that avoid the curse of dimensionality while still being expressive enough to tackle a wide range of problems are in high demand. This topic will be further discussed in the next subsection.

- The most common iterative training algorithm for neural networks is, by a very large margin, (stochastic) gradient descent (see for instance [ALS19]). Some non-intuitive phenomena related to the use of gradient descent in network training are currently under investigation in the literature, e.g. implicit bias ([Gun+17; LMZ18]), double descent fitness curves ([Bel+19; MM22]).

In the next two subsections, we shall exclusively focus on feed forward neural networks. We conclude this subsection with a formal definition of such networks and some of the associated notions.

Definition 8. *Let $\mathbb{F}$ be the field of real or complex numbers. Let $L \in \mathbb{N}$ and $d_{in} := d_0, d_1, \ldots, d_{out} := d_L$ be natural numbers. Let $\phi : \mathbb{F} \to \mathbb{F}$ be a non-linear function. A **feed forward neural network** is a tuple of pairs*

$$\Phi := \big((A_1, b_1), (A_2, b_2), \ldots, (A_L, b_L)\big)$$

such that each A_j is a matrix with d_{j-1} columns and d_j rows with entries in $\mathbb{F}$ and each b_j is a vector in $\mathbb{F}^{d_j}$.

*d_{in} and d_{out} are called the **input** and, respectively, **output dimension**. L is called the **number of layers** of the network. If $L = 2$ (i.e., if the network just has the input layer followed by a single internal or hidden layer), the network is called **shallow**. If $L > 2$, the network is called **deep**.*

$$w(\Phi) := \max(d_0, d_1, \ldots, d_L)$$

*is called the **width** of the network and*

$$N(\Phi) := d_0 + d_1 + \cdots + d_L$$

*is the **number of neurons** of the network. The entries of the matrices A_j and vectors b_j are called the **weights** of the network and*

$$W(\Phi) := \sum_{j=1}^{L} (\|A_j\|_0 + \|b_j\|_0)$$

*is the **number of (non-zero) weights**, where $\|X\|_0$ is the number of non-zero entries of a matrix or vector X. Additionally,*

$$\|\Phi\| := \max(\|A_1\|_\infty, \ldots, \|A_L\|_\infty, \|b_1\|_\infty, \ldots, \|b_L\|_\infty)$$

*is called the **norm** of the network, i.e., its largest weight in absolute value.*

*The **realization** of the network is the function $R_\phi\Phi : \mathbb{F}^{d_{in}} \to \mathbb{F}^{d_{out}}$ given by the composition*

$$R_\phi\Phi = T_L \circ (\phi \circ T_{L-1}) \circ \cdots \circ (\phi \circ T_1),$$

*where $T_j : \mathbb{F}^{d_{j-1}} \to \mathbb{F}^{d_j}$ is the affine operator given by $T_j x = A_j x + b_j$. Crucially the application of the **activation function** ϕ is component-wise,i.e., $\phi(x) := (\phi(x_1), \phi(x_2), \ldots, \phi(x_m))^T$ for any $x := (x_1, x_2, \ldots, x_m)^T \in \mathbb{F}^m$.*

*Finally, the **architecture** of the network Φ is the configuration of neurons per layer, i.e., the tuple $(d_0, d_1, \ldots, d_L)$ together with the positions of non-zero weights on each layer.*

4.2 The curse of dimensionality and Barron functions

As previously discussed, computational suitability of a given machine learning model for a concrete task is sometimes not sufficient, in the sense that some additional flexibility for related tasks is desired while keeping in check the computational complexity.

As a toy example, consider a model that needs to determine the coordinates of a particle in a unit sized box with precision no worse than $\frac{1}{2}$ (for each coordinate). Let $x = (x_1, \ldots, x_d)$ be the position of the particle and let $I = [0, 1]^d$ be the unit box. In one spatial dimension, $d = 1$, it will suffice to do just two measurements, namely $|x - \frac{1}{4}|$ and $|x - \frac{3}{4}|$ and compare the two and declare the position to be either $\frac{1}{4}$ or $\frac{3}{4}$ (based on the smallest distance). In dimension $d = 2$, we would need to do 4 measuremens, namely $|x_i - \frac{1}{4}|$ and $|x_i - \frac{3}{4}|$ for

each $i \in \{1, 2\}$. More generally, in arbitrary dimension d, 2^d measurements would be necessary, so the computational complexity of the algorithm would be exponential in the dimension. This phenomenon of exponential dependency on the input dimension of the overall complexity of the model is referred to as the *curse of dimensionality*.

Note that at the expense of this 'bad' dependency on the input dimension, all the (binary) states of the particle in the box are differentiated. It is not difficult to avoid the curse of dimensionality if compromises can be made in terms of the expressivity of the hypothesis class of models. For example, if we only cared about the (binary) position of the particle in the first $\log_2 d$ coordinates, the overall complexity would be linear in d, but we would not distinguish between states that differ in the larger index coordinates.

A major goal of machine learning is to construct models with sufficient expressive power to cover most practical necessities but with a mild dependency on input dimension (i.e., avoiding the curse of dimensionality and having at most polynomial dependency of the complexity on the input dimension). This allows successful models to be relatively cheaply adapted to a wider range of concrete problems. Arguably the most common practical application of machine learning where the dependency on input dimension is a crucial aspect is image classification. This is because images are typically interpreted as binary or RGB matrices which are vectorized and then fed as inputs. For even the most basic example, the hand drawn digit classification problem based on the MNIST database (`http://yann.lecun.com/exdb/mnist/`), the pictures are 28×28 pixels large resulting in an input dimension of 784. Current image classification problems have input dimension in the order of 10^5 or more (for instance, for ImageNet [Den+09], it is quite standard to rescale all the images to 256×256 pixels).

Paper 4 is inspired by the work of A. Barron (in particular [Bar93]), who constructed a class of Lipschitz continuous functions that can be approximated by shallow networks with N neurons with accuracy $N^{-1/2}$ independent of the input dimension (therefore avoiding the curse of dimensionality). Crucially, Barron also showed that these functions (commonly referred to as *Barron class* functions in the literature) are sufficiently rich to cover a wide range of models (a detailed discussion with examples of Barron functions can be found in Chapter IX of [Bar93]).

The main goal of **Paper 4** is to adapt the works of Barron to classification problems. We are therefore interested in classifiers under the assumption that the boundary between different classes is of Barron class. We consider 'tube-compatible' measures μ (a fairly general class of measures that includes Lebesgue measures and product measures, see Section 1.1, Definition 3.5 and Corollary 3.6 in **Paper 4**) and show that the characteristic function $\mathbb{1}_\Omega$, where Ω is a set with Barron class boundary, can be approximated with respect to the measure μ to arbitrary precision by networks without the curse of dimensionality.

Theorem 10 (Theorem 1.1 in **Paper 4**)**.** *Let μ be a tube-compatible measure with parameter $\alpha \in (0, 1]$ and let $\Omega \subset \mathbb{R}^d$ be a set with Barron class boundary.*

Then for any $N \in \mathbb{N}$, there exists a ReLU network Φ with 3 hidden layers (i.e., $L = 4$) and $\mathcal{O}(d + N)$ neurons satisfying $\|\Phi\| = \mathcal{O}(d + N^{1/2})$ such that

$$\|\mathbb{1}_\Omega - R_\varrho \Phi\|_{L^p(\mu)} = \mathcal{O}(d^{3/(2p)} N^{-\alpha/(2p)}).$$

Remark 6.

- *Crucially, the approximation error depends polynomially (subquadratically) on the input dimension.*
- *We are using the rectified linear unit (ReLU, $\varrho(x) = \max(0, x)$) activation function, which is the most commonly used activation function in the current literature. The adaptation of the classic Barron result for ReLU networks (as opposed to the original setup with Heaviside activation in [Bar93] is done in Proposition 2.2 in **Paper 4**).*

The precise construction of the approximating networks is given in Theorem 3.7 of **Paper 4** where we show that once we choose a desired approximation error, we can fix an architecture that will work for any set Ω (with Barron class boundary). Moreover an upper bound on the total number of weights $W(\Phi)$ is obtained and a quadratic dependence on the input dimension d for W is shown to be sufficient.

The remainder of **Paper 4** is structured as follows.

Section 4 uses entropy-type bounds to show that the approximation rate obtained in Theorem 3.7 is asymptotically optimal. Theorem 4.3 shows that networks with at most W weights cannot achieve approximation rates better than $\mathcal{O}(W^{-1/2-1/(d-1)})$, whereas the result of Theorem 3.7 shows that a rate of $\mathcal{O}(W^{-1/2})$ is achieved with our construction.

Section 5 provides error bounds for the performance of the use of empirical risk minimization in determining the weights of Φ in Theorem 5.1. The remainder of this section is a first attempt at a discussion of the optimality of this rate, however this has mostly been superseded by the more recent work [PV21].

Section 6 motivates the restriction of the main result to tube-compatible measures by using a version of the no-free-lunch theorem to show that for sets with infinite VC-dimension and general measures, no non-trivial minimax approximation results can be achieved. Then we show in Lemma 6.4 that the class of sets with Barron class boundary does indeed have infinite VC-dimension.

Finally, Section 7 addresses the different versions of so-called Barron spaces seen in the literature and appropriate embeddings between them are shown.

The main contribution of the author of the present thesis was in adapting the ideas of the coauthors [PV18] for network approximation of classifiers with smooth boundaries to the Barron boundary setting, the construction of the precise architecture in Theorem 3.7 as well as the size bounds found therein and the formalization of the suitable tube-compatible measure class. Most of the results in the paper were obtained in group discussions. Section 7 was, however, entirely done by F. Voigtlaender and P. Petersen.

4.3 Quantitative universality of neural networks over the complex numbers

Despite the overwhelming success of and research interest in neural networks, the vast majority of formal literature on the topic is focused on real ReLU neural networks, i.e., networks that have vectors of real numbers as inputs and outputs and that use the ReLU activation function $\varrho(x) = \max(0, x)$. However, recent applied works (for example, [WY18; VSL17]) have shown that in some concrete instances (for example magnetic resonance imaging) where the problem to be modeled is natively expressed with complex numbers, it is beneficial to use complex valued neural networks. Empirically, it was shown in [VSL17] that using complex networks instead of splitting the real and imaginary parts of the inputs and feeding them through appropriately sized real networks has clear advantages.

It is interesting to point out that complex valued networks share a non-trivial relation to their real valued counterparts under the identification $\mathbb{C}^d \cong \mathbb{R}^{2d}$. Namely, complex valued activation functions are structurally richer than real valued ones, but the condition of linearity over $\mathbb{C}^d$ is more restrictive than the equivalent notion over $\mathbb{R}^{2d}$. It is likely that the case-by-case advantages of one class of networks over the other is dependent on the interplay between these two phenomena in the context of the concrete problem at hand.

Motivated by a universal approximation theorem for complex networks obtained in [Voi20], the main goal of **Paper 5** is to provide a systematic quantitative approximation template for the study of complex valued networks and to apply it to one of the most natural choices of a complex activation function, the so-called *modReLU* activation function

$$\sigma(z) = \begin{cases} 0 & \text{if } |z| \leq 1 \\ z - \frac{z}{|z|} & \text{else.} \end{cases}$$

We are interested in approximating functions $g : \mathbb{C}^d \to \mathbb{C}$ with Sobolev-type smoothness in the sense of real variables, i.e., the space $\mathcal{F}_{n,d}$, see Section 1.2 in **Paper 5** for details.

The main result is

Theorem 11 (Theorem 1 in **Paper 5**). *For any smoothness parameter $n \in \mathbb{N}$ and input dimension $d \in \mathbb{N}$, there is a constant $C = C(n, d) > 0$ such that for any $\varepsilon \in (0, 1)$ there exists a fixed modReLU-network architecture with at most $C \cdot \ln(2/\varepsilon)$ layers, at most $C \cdot \varepsilon^{-2d/n} \cdot \ln^2(2/\varepsilon)$ weights, all bounded by $C \cdot \varepsilon^{-44d}$ such that for any $g \in \mathcal{F}_{n,d}$ there exists a network Φ with the prescribed architecture satisfying*

$$|g(z) - R_\sigma \Phi(z)| \leq \varepsilon$$

for all z in the unit cube of $\mathbb{C}^d$.

Remark 7. *It is important to point out that these rates are exactly what one would expect via the identification $\mathbb{C}^d \cong \mathbb{R}^{2d}$ (for instance, [Yar17]), so applications where the complex structure is relevant will incur no added cost to*

approximation rates. Additionally, we stress that the architecture only depends on the approximation accuracy, smoothness parameter and ambient dimension. Only the weights of Φ depend on the individual function g to be approximated.

The majority of **Paper 5** is dedicated to the proof of Theorem 1.

Section 1.4 discusses the limited literature on complex valued networks with an emphasis on the role of the complex activation function and the rather surprising result in [Voi20] that, in the complex valued case, the class of universal activation functions is different for shallow networks and for deep networks. Also a short review of other complex activation functions is given.

Section 2 shows that one can construct shallow modReLU-networks to approximate the real and imaginary part of the input and an architecture is given (Proposition 3).

Section 3 contains a proof that the square of the real (or imaginary) part can be approximated and an architecture is given in Proposition 8. This is heavily inspired by the idea of Yarotski [Yar17] to show that $x \mapsto x^2$ can be approximated using ReLU networks.

Section 4 extends this to an approximation of the product of two complex numbers via a simple polarization argument (Proposition 9 and Corollary 10).

Section 5 constructs a partition of unity based on the modReLU activation function.

Section 6 combines the previous steps to give a proof of Theorem 1 by looking at the Taylor approximation of g relative to the constructed partition of unity.

Finally, Section 7 uses entropy-type bounds to show that the approximation rate obtained in Theorem 1 is essentially optimal (modulo log factors). Specifically, Theorem 12 shows that if one can achieve approximation rates of ε via networks with at most $\mathcal{O}(\varepsilon^{-\gamma})$ weights bounded polynomially in ε^{-1}, then necessarily $\gamma \leq 2d/n$.

The main contribution of the author of the present thesis regarding **Paper 5** had to do with the specific architecture constructions and size bounds, especially in Section 6. The vast majority of the results were achieved in group discussions, with the notable exception of Section 7 that was written by F. Voigtlaender.

4.4 Additional work done by the author

The author of the present thesis has a relatively long-standing interest in the work of F. Voigtlaender concerning neural networks and is often the first reader for a lot of his recent papers. The author also took part in the design of F. Voigtlaender's grant proposals concerning machine learning and is currently collaborating with him and T. Römer on the topic of adversarially robust networks.

The author has had minor contributions in terms of editing and technical suggestions for the following papers:

- The universal approximation theorem for complex-valued neural networks, F. Voigtlaender, `https://arxiv.org/pdf/2012.03351`, author appears in acknowledgements.

- L^p sampling numbers for the Fourier-analytic Barron space, F. Voigtlaender, `https://arxiv.org/pdf/2208.07605`, author appears in acknowledgements.
- Optimal learning of high-dimensional classification problems using deep neural networks, P. Petersen and F. Voigtlaender, `https://arxiv.org/abs/2112.12555`.

5 Structured Exponential Riesz bases

This section is based on the paper 'A note on exponential Riesz bases' (co-authored with D.G. Lee), recently accepted for publication in Sampling Theory, Signal Processing, and Data Analysis, henceforth referred to as **Paper 6**, [CL22].

The section is organized as follows. Subsection 1 discusses the background of exponential bases. Subsection 2 introduces the specific structured exponential bases that **Paper 6** discusses and the literature context and provides an overview of the results in **Paper 6**. Finally, Subsection 3 contains some additional work done by the author of this thesis on the topic of exponential bases.

The work done in this section provides partial answers to WP1 from the DFG project PF 450/11-1.

5.1 Exponential bases

As the pure frequencies $\{x \mapsto e^{2\pi inx} : n \in \mathbb{Z}\}$ form the Fourier basis (an orthonormal basis) for the space $L^2([0,1])$ of square integrable function on the unit interval, a natural question to ask is if there exists an similar Fourier basis for the space of square integrable functions supported on some arbitrary bounded measurable (with positive measure) set S.

Definition 9. *Let $S \subset \mathbb{R}^d$ be a bounded set with positive measure and let $\Lambda \subset \mathbb{R}^d$ be a uniformly discrete (in the sense that $\inf_{\lambda_1,\lambda_2\in\Lambda} |\lambda_1 - \lambda_2| > 0$) set of frequencies. Define the collection of functions*

$$\mathcal{E}(\Lambda) := \{e^{2\pi i\langle\lambda,\cdot\rangle}\mathbb{1}_S(\cdot) : \lambda \in \Lambda\}.$$

*The set $\mathcal{E}(\Lambda)$ is said to be an **orthogonal (Riesz) basis of exponentials** for $L^2(S)$ if the collection $\mathcal{E}(\Lambda)$ forms an orthogonal (or, respectively, Riesz) basis for $L^2(S)$. To simplify notation, we shall say that Λ is **orthogonal (Riesz)** spectral for S.*

Example 1. *In this notation, $\mathbb{Z}$ is orthogonally spectral for $[0,1]$. This is just the restatement of the fact that Fourier basis is an orthonormal basis for $L^2([0,1])$ using the fact that the Lebesgue measure of the unit interval is 1.*

Remark 8.

- *There is essentially no distinction between discussing orthogonal and orthonormal exponential bases. Indeed, if Λ is orthogonally spectral for S, then by definition $\langle e^{2\pi i\langle\lambda_1,\cdot\rangle}\mathbb{1}(S)(\cdot), e^{2\pi i\langle\lambda_2,\cdot\rangle}\mathbb{1}(S)(\cdot)\rangle_{L^2(S)} = 0$ whenever $\lambda_1 \neq \lambda_2$. Moreover, $\|e^{2\pi i\lambda\cdot}\mathbb{1}(S)(\cdot)\|^2_{L^2(S)} = \mu(S)$ for all $\lambda \in \Lambda$. So by rescaling all the pure frequencies by the same factor $\frac{1}{\sqrt{\mu(S)}}$, we get a orthonormal basis for $L^2(S)$. We will therefore stick to the not-normalized orthogonal bases to simplify notation.*

- *The uniform discreteness of the frequency set is a necessary condition. Indeed, if Λ is not uniformly discrete, then it cannot be Riesz spectral for any bounded set S (see for example Remark 1 in [KN15]).*

- *We will hereafter omit the multiplication by the characteristic function of S in our notation for the set $\mathcal{E}(S)$. However, it is fundamental to understand that if we say that Λ is spectral for a set S, then the pure frequency functions defined by Λ are understood as functions defined on S. This is crucial when discussing bases for unions of sets. As we will see below from Lemma 1, $\Lambda_1 = 2\mathbb{Z}$ is orthogonally spectral for $[0, \frac{1}{2}]$ and $\Lambda_2 = 2\mathbb{Z} + 1$ is orthogonally spectral for $[\frac{1}{2}, 1]$. Combining the frequency sets we get $\mathbb{Z} = \Lambda_1 \cup \Lambda_2$ and this is orthogonally spectral for $[0, 1]$ if we extend the domain of both even and odd pure frequency functions to be the entire unit interval. For example, $\mathbb{1}_{[0,1]}$ is not in the span of $\{e^{2\pi i 2k \cdot} \mathbb{1}_{[0,\frac{1}{2}]} : k \in \mathbb{Z}\} \cup \{e^{2\pi i(2k+1)\cdot} \mathbb{1}_{[\frac{1}{2},1]} : k \in \mathbb{Z}\}$.*

Lemma 1. *Let $S = [a, b]$ be some non-degenerate bounded interval, i.e., $-\infty < a < b < \infty$. Then for any $c \in \mathbb{R}$, the set $\frac{1}{b-a}\mathbb{Z} + c$ is orthogonally spectral for S.*

Sketch of proof. First of all, note that $\frac{1}{b-a}\mathbb{Z} + c$ is orthogonally spectral if and only if $\frac{1}{b-a}\mathbb{Z}$ is. This is because $\mathcal{E}(\frac{1}{b-a}\mathbb{Z} + c) = e^{2\pi i c \cdot} \mathbb{1}_{[a,b]}(\cdot)\mathcal{E}(\frac{1}{b-a}\mathbb{Z})$ and $\|e^{2\pi i c \cdot} \mathbb{1}_{[a,b]}(\cdot)\|_2^2 = b - a$.

Next note that $\frac{1}{b-a}\mathbb{Z}$ is orthogonally spectral for $[a, b]$ if and only if it is orthogonally spectral for $[0, b - a]$. This follows from the fact that for any $n \in \mathbb{Z}$, the function $x \mapsto e^{2\pi i \frac{n}{b-a} x}$ is $b - a$ periodic and so for any sequence $\{c_n\}_n \in \ell^2$ we have that

$$\| \sum_{n\in\mathbb{Z}} c_n e^{2\pi i \frac{n}{b-a} \cdot} \|_{L^2([0,b-a])} = \| \sum_{n\in\mathbb{Z}} c_n e^{2\pi i \frac{n}{b-a} \cdot} \|_{L^2([a,b])}.$$

Finally, consider the dilation operator $D_{b-a} : L^2([0, 1]) \to L^2([0, b - a])$ given by $D_{b-a}(f)(x) = \frac{1}{\sqrt{b-a}} f((b - a)x)$. Since this is an isometry and since the Fourier basis is orthonormal, it follows that $\frac{1}{b-a}\mathbb{Z}$ is orthogonally spectral for $[0, b - a]$. □

Lemma 1 can be seen as a definition for the Fourier basis on an arbitrary bounded interval. As mentioned before, a natural question is whether any bounded positive measure set admits a Fourier basis (in this context we mean an orthogonally spectral set of frequencies).

The answer to this question is *no*, in general. Possibly the most natural example of this is that the unit ball in any dimension $d \geq 2$ admits no orthogonal basis of exponentials [IKP99]. The most famous attempt of a characterization of subsets of $\mathbb{R}^d$ that admit an orthogonal basis of exponentials is due to Fuglede.

Conjecture 1 (Fuglede's conjecture, [Fug74])**.** *A set $S \subset \mathbb{R}^d$ admits an orthogonal basis of exponentials if and only if S tiles $\mathbb{R}^d$, i.e., if there exists some discrete translation set $\Gamma \subset \mathbb{R}^d$ such that $\mathbb{R}^d = \cup_{\gamma\in\Gamma}\{S + \gamma\}$ with the union disjoint up to zero sets.*

Fuglede proved the conjecture under the additional assumption that either the candidate frequency set Λ or the tiling set Γ are restricted to be lattices in $\mathbb{R}^d$

in [Fug74]. However, without one of these additional restrictions, the conjecture is in general false. Tao showed that both directions of the conjecture can fail in dimension $d \geq 5$ ([Tao04]) and this was later extended to dimensions $d = 3$ and $d = 4$ (see for example [Mat05; KM06]). The validity of the conjecture in dimensions $d = 1$ and $d = 2$ remains open to the author's knowledge.

Since the orthogonality of the pure frequency functions seems to prevent sets without particular geometric features from having orthogonal spectral sets and since it is often sufficient for applications to consider Riesz bases instead of orthonormal bases, the obvious question is whether all sets admit Riesz bases of exponentials or if the sets that do not can be characterized.

Recently, it was shown in [KNO21] that even in dimension $d = 1$ there exist sets with no Riesz bases of exponentials. The set proposed by Kozma, Nitzan and Olevskii in [KNO21] is rather exotic (the construction of the set relying on self-similarity). For 'tamer' sets S, specifically finite unions of intervals in $d = 1$, there exist positive and partially constructive methods for obtaining Riesz spectral sets. Additionally, under appropriate dilations, the frequencies can all be chosen to be integers.

Theorem 12 (Avdonin [Avd91], Seip [Sei95])**.** *Let $S \subset [0, 1]$ be a non-degenerate interval. Then there exists a frequency set $\Lambda \subset \mathbb{Z}$ that is Riesz spectral for S.*

Theorem 13 (Combining Riesz bases for intervals,[KN15])**.** *Let $S \subset [0, 1]$ be a finite union of non-degenerate intervals. Then there exists a frequency set $\Lambda \subset \mathbb{Z}$ that is Riesz spectral for S.*

The landmark paper [KN15], of which Theorem 13 is the main result, is remarkable not only for proving the existence of Riesz spectral sets for finite unions of intervals, but for introducing a novel 'constructor method' for generating Riesz spectral sets for general sets S under some conditions.

Prior to [KN15], to the author's knowledge, there existed only three classical tools of constructing Riesz spectral sets and they all referred specifically to sets S that are intervals:

1. Lemma 1 that yields translations of lattices for the frequency set.
2. **Kadec's $\frac{1}{4}$-theorem** ([Kad64] or [You01] for a modern reference): If $\Lambda = \{\lambda_n : n \in \mathbb{Z}\}$ satisfies the condition $\sup_n |\lambda_n - n| < \frac{1}{4}$, then Λ is Riesz spectral for $[0, 1]$.
3. **Avdonin's theorem** ([Avd74] or [AI95] for a modern reference), a powerful generalization of Kadec's theorem where the size of the perturbation of the Fourier basis has to be smaller than $\frac{1}{4}$ in the average: Let $\Lambda = \{\lambda_n\}_{n\in\mathbb{Z}}$ be a uniformly discrete sequence. If there exists a natural number N and a constant $0 \leq \delta < \frac{1}{4}$ such that

$$\Big| \sum_{n=mN+1}^{(m+1)N} (\lambda_n - n) \Big| \leq \delta N$$

for all $m \in \mathbb{Z}$, then Λ is Riesz spectral for $[0, 1]$.

It bears pointing out that using Avdonin's theorem alongside dilation and translation techniques similar to those of Lemma 1, one can explicitly construct an integer frequency Riesz spectral set for any interval $S \subset [0,1]$, recovering Theorem 12.

Indeed, similar to Lemma 1, one can assume without loss of generality that $S = [0, \alpha]$ for some $0 < \alpha < 1$. Then the scaled version of Avdonin's lemma implies that if $\Lambda = \{\lambda_n : n \in \mathbb{Z}\}$ satisfies

$$\Big| \sum_{n=mN+1}^{(m+1)N} (\lambda_n - \tfrac{n}{\alpha}) \Big| \leq \tfrac{\delta}{\alpha} N$$

for all integers m and for some fixed natural N and some fixed constant $0 \leq \delta < \frac{1}{4}$, then Λ is Riesz spectral for S.

Now there are two cases. If $\alpha = \frac{p}{q} \in \mathbb{Q}$ for some coprime naturals $p < q$, then the reference lattice will be $\frac{q}{p}\mathbb{Z}$. We can take $\Lambda = \{\lfloor \frac{qn}{p} \rfloor : n \in \mathbb{Z}\}$, $N = p$ and $\delta = 0$ and the conditions of the scaled Avdonin theorem will be satisfied.

If $\alpha \notin \mathbb{Q}$, then one can again take $\Lambda = \{\lfloor \frac{n}{\alpha} \rfloor : n \in \mathbb{Z}\}$. For any $\delta > 0$, the existence of some large natural number N that satisfies the scaled Avdonin theorem is then guaranteed by Weyl's equidistribution theorem (a stronger version of which is Proposition 7 discussed in **Paper 6**).

The new method for constructing Riesz spectral sets which is the core of [KN15] and the main tool behind the present section is the following technical lemma.

Lemma 2 (Lemma 2 in [KN15])**.** *Let $S \subset [0,1]$ be an arbitrary set with positive measure. Fix some natural number N and consider the 'folding' sets*

$$A_{\geq n} := \{x \in [0, \tfrac{1}{N}] : x + \frac{k}{N} \in S \text{ for at least } n \text{ distinct integers } k\}.$$

If, for each $1 \leq n \leq N$, there exists a frequency set $\Lambda_n \subset N\mathbb{Z}$ that is Riesz spectral for $A_{\geq n}$, then the set $\Lambda := \cup_{n=1}^{N} \big(\Lambda_n + n\big)$ is Riesz spectral for S.

As a toy example of an application of Lemma 2 (although this can also be proved more directly using the non-singularity of the corresponding minors of the Fourier matrix, the key ingredient of the proof of Lemma 2), consider the set $S \subset [0,1]$ to be a finite union of intervals with rational endpoints. Choose N to be the smallest common multiple of all the denominators of the individual interval endpoints. It is now easy to check that

$$A_{\geq n} = \begin{cases} [0, \frac{1}{N}] & \text{if } n \leq N\mu(S) \\ \varnothing & \text{if } n > N\mu(S), \end{cases}$$

so, using Lemma 1, we can take

$$\Lambda_n = \begin{cases} N\mathbb{Z} & \text{if } n \leq N\mu(S) \\ \varnothing & \text{if } n > N\mu(S), \end{cases}$$

implying that $\cup_{n=1}^{N\mu(S)} \big(N\mathbb{Z} + n\big)$ is Riesz spectral for S.

5.2 Complementability of the Fourier basis and hierarchical Riesz bases of exponentials

Paper 6 is motivated by two questions from WP 1 of the DFG project PF 450/11-1.

Problem 3 (Complementability of Riesz spectral sets)**.** *Let $S \subset \mathbb{R}$ be a bounded set with positive measure and let $\mathcal{C}$ be a class of bounded sets each of which are disjoint from S. Can one find some frequency set Λ that is Riesz spectral for S such that for each $S' \in \mathcal{C}$ there exists some Λ' Riesz spectral for S' satisfying the additional condition that $\Lambda \cup \Lambda'$ is Riesz spectral for $S \cup S'$? If the answer is yes, Λ is said to be complementable with respect to the class $\mathcal{C}$. What conditions on S and $\mathcal{C}$ are sufficient for complementability?*

Problem 4 (Hierarchical Riesz exponential bases)**.** *Under what conditions on the bounded measurable sets $S_1, S_2, \dots, S_L$ can we find frequency sets $\Lambda_1, \dots, \Lambda_L$ such that for each index subset $J \subset \{1, 2, \dots, L\}$, the union $\cup_{j\in J}\Lambda_j$ is Riesz spectral for $\cup_{j\in J} S_j$?*

The first problem has its origins in discussions with Pfander and Walnut and the investigation of the second is due to an attempt to generalize the following consequence of the main result of Pfander, Walnut and Revay:

Theorem 14 ('Consecutive hierarchy'[PRW21])**.** *Let $[0,1] = \cup_{k=1}^{L}[a_k, b_k)$ be a partition of the unit interval into subintervals. Then one can find an explicit partition of the integers $\mathbb{Z} = \cup_{k=1}^{L}\Lambda_k$ such that for each consecutive index subset $J \subset \{1, 2, \dots, L\}$, the frequency set $\cup_{j\in J}\Lambda_j$ is Riesz spectral for $\cup_{j\in J}[a_j, b_j)$.*

The Avdonin based approach of Theorem 14 can be used to treat a few more special cases when the index subset contains some specific gaps, but there seems to be no clear way to obtain a fully hierarchical partition of the integers with this technique.

The two main results of **Paper 6** are based on Lemma 2 by contrast.

Theorem 15 (Theorem 2 in **Paper 6**)**.** *The Fourier basis of the unit interval is complementable with respect to any finite union of bounded intervals. Given real numbers $1 \le a_1 < b_1 < a_2 < b_2 < \cdots < a_L < b_L \le N$ for some $L, N \in \mathbb{N}$, there exists some frequency set $\Lambda' \subset \left(\frac{1}{N}\mathbb{Z}\right) \setminus \mathbb{Z}$ such that Λ' is Riesz spectral for $\cup_{k=1}^{L}[a_k, b_k)$ and $\mathbb{Z} \cup \Lambda'$ is Riesz spectral for $[0,1] \cup \left(\cup_{k=1}^{L}[a_k, b_k)\right)$.*

Theorem 16 (Theorem 1 in **Paper 6**)**.** *Let $0 < a_1 < b_1 < a_2 < b_2 < \cdots < a_L < b_L < 1$ with $L \in \mathbb{N}$ and consider the set $S = \cup_{k=1}^{L}[a_k, b_k]$. If the numbers $1, a_1, a_2, \dots, a_L, b_1, b_2, \dots, b_L$ are linearly independent over $\mathbb{Q}$, then S admits a hierarchical Riesz basis of exponentials with integer frequencies.*

We conjecture that the condition of the endpoints being rationally independent in Theorem 16 can be lifted, but so far we have not been able to show this.

As stated before, the main idea behind these results is Lemma 2. Specifically for Theorem 16 two more main ingredients are used:

- Chebotarëv's theorem on roots of unity (see for example [SL96]) which states that all minors of a prime sized Fourier matrix are invertible.
- A Kronecker-Weyl equidistribution along the primes result (Proposition 7 in **Paper 6**).

Paper 6 is structured as follows:

- The first section contains the statements of the two main results, Theorems 1 and 2.
- The second section contains the technical lemmas used including the generalization of Lemma 2 contained in Lemma 6.
- The third section contains a proof of the hierarchy result.
- The fourth section contains a proof of the complementability result.
- The appendix discusses the proof of Lemma 6.

In terms of the work contribution on **Paper 6**, both Dae Gwan Lee and the author of the present thesis wrote down essentially identical (in terms of the techniques used) manuscripts based on group discussions. The published version is a synthesis of the two individual manuscripts.

We conclude this subsection with a short overview of the proofs of the two main results.

A discussion of the proof of Theorem 15. Dilating by $\frac{1}{N}$, we can set $S = [0, \frac{1}{N}]$ and $S' = \cup_{k=1}^{L} [\frac{a_k}{N}, \frac{b_k}{N}]$ such that $S \cup S' \subset [0, 1]$.

Denoting by $\{x\}$ the fractional part of a real number x, we see that depending on the ordering of the fractional parts $\{Na_1\}, \ldots, \{Na_L\}, \{Nb_1\}, \ldots, \{Nb_L\}$, the non-empty sets $A_{\geq n}$ will either be $[0, \frac{1}{N}]$ or finite unions of intervals with endpoints from the set $\{0, \frac{1}{N}, \frac{\{a_k\}}{N}, \frac{\{b_\ell\}}{N} : 1 \leq k, \ell \leq N\}$.

The conclusion now follows by applying Theorem 13 to each of the sets $A_{\geq n}$ and then Lemma 2. It is worth pointing out that this can be derived in a simpler way in the case $L = 1$, by invoking Theorem 12 and possibly Proposition 3 from **Paper 6**. □

A discussion of the proof of Theorem 16. Conceptually, the main difficulty that arises from attempting to prove the existence of a hierarchical basis as opposed to just a basis for the full union using the techniques from [KN15] is the appearance of non-consecutive shift factors for the Riesz exponential bases of the sets $A_{\geq n}$ (see Figures 1 and 2).

Lemma 2 relies on the fact that the frequency set $\cup_{k=1}^{L} \big(N\mathbb{Z} + j_k\big)$ is Riesz spectral for the union of intervals $\cup_{k=1}^{L} [\frac{\ell_k}{N}, \frac{\ell_k+1}{N}]$ (with $0 \leq \ell_k \leq N-1$) if the minor of the Fourier matrix $\big[e^{-2\pi i j_k \ell_k / N}\big]_{1 \leq k \leq L}$ corresponding to the rows j_k and columns ℓ_k is invertible. This condition is trivially satisfied if the shifting coefficients j_k are consecutive, since in that case the minor of the Fourier matrix is a Vandermonde matrix and hence invertible.

Chebotarëv's theorem provides a solution to this problem at the cost of restricting the choice of N to prime numbers (since then all minors of the Fourier matrix are invertible). This allows us to generalize Lemma 2 (in the case of prime N) to Lemma 6 in **Paper 6**.

The next issue is ensuring that excising any interval $[a_k, b_k)$ from the set S has a controllable effect on the folding sets $A_{\geq n}$. The solution we found was to enforce a particular order of the endpoints of the intervals when folded onto $[0, \frac{1}{N}]$, explicitly that

$$0 < \{Na_1\} < \{Na_2\} < \cdots < \{Na_L\} < \{Nb_L\} < \cdots < \{Nb_2\} < \{Nb_1\} < 1.$$

This ordering informally means that, apart from full copies of $[0, \frac{1}{N}]$, the contributions of the each interval $[a_k, b_k)$ to the folding sequence $A_{\geq n}(S)$ are nested. More explicitly,

$$A_{\geq n} = \begin{cases} [0, \frac{1}{N}] & \text{for } n \leq m \\ [0, \frac{\{Nb_k\}}{N}) \cup [\frac{\{Na_k\}}{N}, \frac{1}{N}) & \text{for } n = m + k, 1 \leq k \leq L \\ \varnothing & \text{for } n > m + L, \end{cases}$$

where m is some natural number $m < N$.

Again, informally speaking, this means that dropping any interval (or subunion of intervals) from the full union S will translate to 'deleting' the appropriate levels from the folding sets $A_{\geq n}$ and permuting the remaining sets (with no further interaction) (see Figure 2). The permutation introduced in this way is taken into account by the use of Lemma 6 from **Paper 6** instead of Lemma 2, as discussed above.

In order to guarantee the existence of a prime N that satisfies the desired ordering, we make use of a Kronecker-Weyl equidistribution result along the primes (Proposition 7 in **Paper 6**), an idea that dates back (to our knowledge) to Vinogradov.

Unfortunately, this approach imposes the seemingly artificial restriction of the endpoints of the intervals needing to be rationally independent.

□

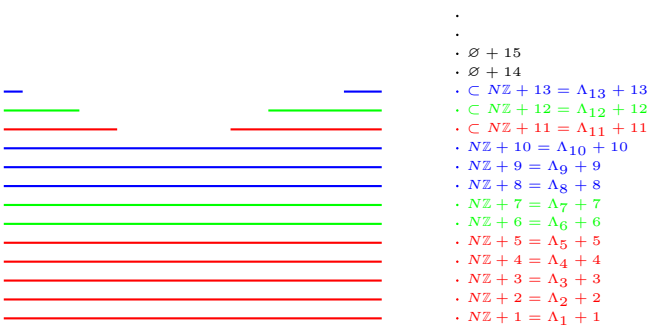

Figure 1: A toy example of S being a union of 3 intervals. N is chosen to 'nest' the contributions of each of the intervals in the sets $A_{\geq n}$. The shifts of the Riesz spectral sets can be chosen to be consecutive.

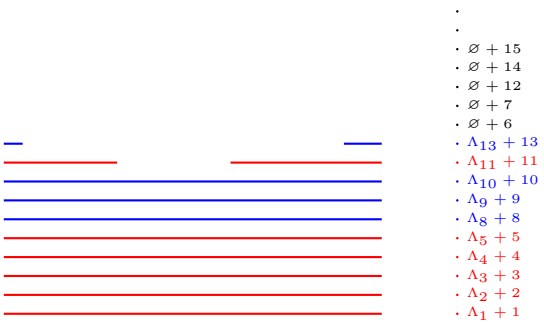

Figure 2: S' is the partial union of 2 intervals. Due to the 'nesting' of contributions of the intervals in S, the 'tower of folding sets' $A_{\geq n}(S')$ is obtained from the 'tower' $A_{\geq n}(S)$ just by eliminating the levels corresponding to the missing interval and shifting. Note the non-consecutive nature of the shifts of the Riesz spectral frequency sets which is circumvented by the use of Chebotarëv's theorem. Also note that the basis corresponding to the 'red' (or 'blue') interval have not changed from Figure 1.

5.3 Additional work done by the author

This subsection describes some additional investigations of the author of the present thesis concerning potential applications of Lemma 2 to the study of structured exponential bases for unions of intervals and some restrictions related to Kadec bases. Two threads of discussion are presented here, the first concerned with attempts to lift the restriction on the endpoints of the intervals being rationally independent in order to obtain a hierarchical basis, and the second arguing that simple Kadec bases cannot have more than one intersection point without compromising the spanning properties of their union.

5.3.1 On the topic of hierarchical bases for arbitrary finite unions of intervals

We recall that we are interested in finding hierarchical Riesz bases of exponentials for $L^2(S)$, where $S \subset [0,1]$ is a finite union of intervals of the form $S = \cup_{k=1}^{L}[a_k, b_k]$. I.e., we want to find some frequency set of integers $\Lambda \subset \mathbb{Z}$ that can be partitioned as $\Lambda = \cup_{k=1}^{L}\Lambda_k$ such that for any index subset J of $\{1, 2, \ldots, L\}$ we have that $\cup_{k\in J}\Lambda_k$ is Riesz spectral for $\cup_{k\in J}[a_k, b_k]$.

Theorem 16 solves this question in the case when $1, a_1, \ldots, a_L, b_1, \ldots, b_L$ are linearly independent over $\mathbb{Q}$, however, as pointed out before, we believe that this restriction is a proof artifact induced by our use of an equidistribution-type result and we conjecture that the rational independence of the interval endpoints condition can be dropped, although we have been unable to prove this up to this point.

A natural path of investigation towards proving this more general result is the following problem.

Problem 5. *Let $[0,1] = \cup_{k=1}^{L}[a_k, b_k]$ be a partition of the unit interval into a finite union of intervals with rational endpoints, i.e., $a_k, b_k \in \mathbb{Q}$. Then there exists a partition of the integers $\mathbb{Z} = \cup_{k=1}^{L}\Lambda_k$ that forms a hierarchical Riesz basis for the unit interval.*

Somewhat surprisingly, the 'opposite' scenario of having rational endpoints presented in Problem 5 seems more difficult to tackle than the situation of rationally independent points illustrated in Theorem 16. The author's main focus concerning this topic after the resolution of Theorem 16 has been towards Problem 5.

Currently, the best approach to tackle Problem 5, based on discussions with G. Pfander, D. Walnut and D.G. Lee, seems to be the following conjecture.

Conjecture 2. *Let N be a natural number and let $\mathcal{F} := [e^{-2\pi ik\ell/N}]_{1\leq k,\ell\leq N}$ be the associated Fourier matrix. Then there exists some permutation $\sigma : \{1, 2, \ldots, N\} \to \{1, 2, \ldots, N\}$ such that, for any subset $K \subset \{1, 2, \ldots, N\}$, the minor*

$$\mathcal{F}_K = [e^{-2\pi ik\sigma(\ell)/N}]_{k,\ell\in K}$$

of the Fourier matrix is invertible.

Remark 9.

- *This conjecture can be seen as an extension of Chebotarëv's theorem to the case when N is not a prime. Indeed, when N is prime, the conjecture is satisfied for σ being the identity permutation (or any other permutation).*
- *Numerical testing has been done in MatLab to verify the validity of the conjecture for small values of N. For $N \leq 20$, an algorithm based on generating random permutations has always found σ that satisfy the conjecture and for $N \leq 15$ an exhaustive search based on the Steinhaus-Johnson-Trotter algorithm for iteratively generating permutations without storing*

them has been implemented to find all σ that satisfy the conjecture. Unfortunately, no clear asymptotic conclusion concerning the frequency of 'good' permutations could be drawn from the runs.

This conjecture would indeed imply a positive answer to Problem 5.

Sketch of a positive solution to Problem 5 assuming Conjecture 2. Let N be the least common multiple of the denominators of the endpoints of the intervals, i.e., of all the a_k and b_k. Clearly, in general, N will not be prime and therefore the approach using equidistribution and Chebotarëv's theorem from Theorem 16 will not work in this scenario.

Note that the length of each interval will, in this case, be a integer multiple of N, let us say that $b_k - a_k = \frac{\rho_k}{N}$ with $\rho_k \in \mathbb{N}$ for all k. The ρ_k will clearly satisfy $\sum_{k=1}^{L} \rho_k = N$.

Let σ be a 'good' permutation prescribed by the assumption that Conjecture 2 holds.

The folding set chain associated to the full interval $[0,1]$ will be $A_{\geq n} = [0, \frac{1}{N}]$ for all $1 \leq n \leq N$ and we can take the frequency set $\Gamma_n = N\mathbb{Z} + \sigma(n)$ as Riesz spectral for $A_{\geq n}$.

Furthermore, for each index $1 \leq k \leq L$ we can define a frequency set

$$\Lambda_k := \cup_{m=1}^{\rho_k} \Gamma_{m+\sum_{j=1}^{k-1} \rho_j},$$

under the convention that $\sum_{j=1}^{0} \rho_j = 0$.

With this notation, given an arbitrary index subset $J \subset \{1, 2, \ldots, L\}$ and setting $S_J := \cup_{k \in J} [a_k, b_k]$, we see that

$$A_{\geq n}(S_J) = \begin{cases} [0, \frac{1}{N}] & \text{if } n \leq \sum_{k \in J} \rho_k \\ \varnothing & \text{otherwise.} \end{cases}$$

Explicitly writing the index set $J = \{k_1, k_2, \ldots, k_r\}$, where $k_1 < k_2 < \cdots < k_r$, we can now take, for example, $\Gamma_{1+\sum_{j=1}^{k_1-1} \rho_j}$ as Riesz spectral for $A_{\geq 1}(S_J)$, $\Gamma_{2+\sum_{j=1}^{k_1-1} \rho_j}$ as Riesz spectral for $A_{\geq 2}(S_J)$, $\Gamma_{1+\sum_{j=1}^{k_2-1} \rho_j}$ as Riesz spectral for $A_{\geq \rho_{k_1}+1}(S_J)$ and so on.

With these choices, it follows that $\cup_{k \in J} \Lambda_k$ will be Riesz spectral for S_J as desired. □

As an informal note on the above sketch, firstly, the choice of N guarantees that the folding sets $A_{\geq n}$ are either full intervals or empty and thus the contributions of each interval can be separated. In contrast, in the proof of Theorem 16, this was achieved by making sure that these contributions were nested. Secondly, the permutation σ from Conjecture 2 guarantees the invertibility of the involved Fourier minors when one cannot ensure that they are Vandermonde matrices or that N is prime.

A different approach to exploring Problem 5 undertaken by the author of the present thesis was to examine the possibility of achieving a nested structure

of contributions of the individual intervals to the folding sets $A_{\geq n}$ without employing any equidistribution argument. This would clearly lift the restriction of the endpoint of the intervals being rationally independent, a condition clearly violated under the assumptions of Problem 5. This gives rise to the following number theoretic question.

Problem 6. *Let $S = \cup_{k=1}^{L}[a_k, b_k] \subset [0,1]$ such that the endpoints a_k, b_k are all rational. Can one reorder the endpoints a_k and b_k arbitrarily modulo $\frac{1}{N}$ for some prime N? In other words, given an arbitrary bijective map $\sigma : \{c_1, c_2, \ldots, c_{2L}\} \to \{a_1, \ldots, a_L\} \cup \{b_1, \ldots, b_L\}$, does there exist some prime N such that*

$$\{N_{\sigma(c_1)}\} < \{N_{\sigma(c_2)}\} < \cdots < \{N_{\sigma(c_{2L})}\}?$$

Remark 10.

- *As seen from the discussion of the proof of Theorem 16, the orderings we are actually interested in are of the form $\{Na_1\} < \{Na_2\} < \cdots < \{Na_L\} < \{Nb_L\} < \cdots < \{Nb_1\}$.*
- *If a positive solution for Problem 6 is found, especially with large enough prime N, then Problem 5 can be immediately solved by following along the steps in the proof of Theorem 16, by replacing the equidistribution argument and invoking Chebotarëv's theorem once again.*

Perhaps not surprisingly, a positive solution to Problem 6 is not possible in full generality. What is more interesting is that one can provide an explicit and fairly general technical condition on the endpoints under which all reorderings are possible for infinitely many primes.

To conclude this section, we will first discuss a simple counterexample to the desired result of Problem 6 holding in general and secondly present a technical number theory lemma for sufficient additional conditions for the result to hold.

A concrete counterexample

Consider the set $S = [\frac{1}{6}, \frac{1}{4}] \cup [\frac{1}{3}, \frac{1}{2}] \cup [\frac{3}{4}, \frac{5}{6}]$, i.e., we have the endpoints $a_1 = \frac{1}{6}$, $a_2 = \frac{1}{3}$, $a_3 = \frac{3}{4}$, $b_1 = \frac{1}{4}$, $b_2 = \frac{1}{2}$ and $b_3 = \frac{5}{6}$.

We are interested in the existence of some permutation $\sigma : \{1,2,3\} \to \{1,2,3\}$ such that given some relabeling of a, b as c, d (meaning that if $c = a$ then $d = b$ or vice-versa), we can achieve a reordering of the form

$$\{Nc_{\sigma(1)}\} \leq \{Nc_{\sigma(2)}\} \leq \{Nc_{\sigma(3)}\} \leq \{Nd_{\sigma(3)}\} \leq \{Nd_{\sigma(2)}\} \leq \{Nd_{\sigma(1)}\}, \quad (5)$$

where N is a natural number (ideally a prime).

Since the least common multiple of the denominators of the endpoints is clearly 12, we only need to investigate $1 \leq N \leq 11$ in view of the Chinese remainder theorem. We present the reordering in the following table, where the numerators are listed modulo 12 (with the distinction that the class 0 will be written as 0 if it comes from a left interval endpoint, i.e., an a, and as 12 if it comes from a right interval endpoint, i.e., a b). We will also abuse notation and write $x \leq y$ instead of $\{Nx\} \leq \{Ny\}$ to improve the readability of the table.

N	a_1	b_1	a_2	b_2	a_3	b_3	Order
1	2	3	4	6	9	10	$a_1 < b_1 < a_2 < b_2 < a_3 < b_3$
2	4	6	8	12	6	8	$a_1 < a_3 = b_1 < a_2 = b_3 < b_2$
3	6	9	0	6	3	6	$a_2 < a_3 < a_1 = b_2 = b_3 < b_1$
4	8	12	4	12	0	4	$a_3 < a_2 = b_3 < a_1 < b_1 = b_2$
5	10	3	8	6	9	2	$b_3 < b_1 < b_2 < a_2 < a_3 < a_1$
6	0	6	0	12	6	12	$a_1 = a_2 < a_3 = b_1 < b_2 = b_3$
7	2	9	4	6	3	10	$a_1 < a_3 < a_2 < b_2 < b_1 < b_3$
8	4	12	8	12	0	8	$a_3 < a_1 < a_2 = b_3 < b_1 = b_2$
9	6	3	0	6	9	6	$a_2 < b_1 < a_1 = b_2 = b_3 < a_3$
10	8	6	4	12	6	4	$a_2 = b_3 < a_3 = b_1 < a_1 < b_2$
11	10	9	8	6	3	2	$b_3 < a_3 < b_2 < a_2 < b_1 < a_1$

It is not difficult to see that for $N \notin \{6, 10\}$, there is no way to start the orderings with a_j and end with b_j or vice-versa. If we fix this issue for $N = 6$, we are left with $a_1 < a_3 = b_1 < b_3$. Similarly, for $N = 10$, we are left with $b_3 < a_3 = b_1 < a_1$. None of these configurations satisfy the desired ordering prototype.

It is also interesting to point out that even if we consider shifts of S of the form $S_\varepsilon = [\frac{1}{6} + \varepsilon, \frac{1}{4} + \varepsilon] \cup [\frac{1}{3} + \varepsilon, \frac{1}{2} + \varepsilon] \cup [\frac{3}{4} + \varepsilon, \frac{5}{6} + \varepsilon]$ with small $\varepsilon \geq 0$ (which will have no impact of the existence of hierarchical Riesz bases, i.e., S admits a hierarchical Riesz basis if and only if all the S_ε do as long as $S_\varepsilon \subset [0, 1]$), the desired orderings still cannot be achieved.

Indeed, consider rational numbers $0 < \frac{p_1}{q} < \frac{p_2}{q} < \cdots < \frac{p_d}{q} < 1$ and some $\varepsilon \geq 0$ and let N be an arbitrary natural number.

If $\varepsilon = 0$, as mentioned before, the Chinese remainder theorem implies that the ordering of the numbers $\{N(\frac{p_j}{q} + \varepsilon)\}$, $1 \leq j \leq d$, is uniquely determined by the remainder of N modulo q.

If $\varepsilon > 0$, we can write $N = kq + r$ where k is some non-negative integer and $0 \leq r \leq q - 1$ and then we have that

$$\big\{N\big(\tfrac{p_j}{q} + \varepsilon\big)\big\} = \big\{(kq + r)\tfrac{p_j}{q} + N\varepsilon\big\} = \big\{r\tfrac{p_j}{q} + N\varepsilon\big\}.$$

Crucially, the $N\varepsilon$ quantity is independent of the index j, so if we fix a remainder of N modulo q, the tuple of numbers of the form $\{N(\frac{p_j}{q} + \varepsilon)\}$ corresponds to a cyclic shift of the tuple with numbers $\{N\frac{p_j}{q}\}$. The size of $\{N\varepsilon\}$ just dictates the order of the cyclic shift. Therefore, if one chooses ε appropriately, the new orderings achievable are just all the cyclic shifts of the ordering obtained when $\varepsilon = 0$.

It is not difficult to see that if we also allow cyclic shifts of the ordering in the table above, we still find no match with the desired ordering prototype specified in (5).

A technical sufficient condition for Problem 6

Before stating the sufficient condition, we discuss the Chinese remainder theorem and a simple number theoretic lemma.

Given pairwise coprime numbers $n_1, \ldots, n_d$, the Chinese remainder theorem states that there exists a ring isomorphism

$$\mathbb{Z}/(n_1 n_2 \ldots n_d)\mathbb{Z} \cong (\mathbb{Z}/n_1\mathbb{Z}) \times (\mathbb{Z}/n_2\mathbb{Z}) \times \cdots \times (\mathbb{Z}/n_d\mathbb{Z}),$$

or, equivalently, that for any integers $c_1, \ldots, c_d$, the system of equations $x \equiv c_i \bmod n_i$, $1 \leq i \leq d$, has a unique solution modulo $n_1 n_2 \ldots n_d$.

We are interested in the case where the moduli n_i are not necessarily coprime. First of all note that if x simultaneously satisfies

$$x \equiv c_i \bmod n_i$$
$$x \equiv c_j \bmod n_j,$$

then any common divisor of n_i and n_j must divide the difference $c_i - c_j$. This implies that the condition $c_i \equiv c_j \bmod gcd(n_i, n_j)$ is necessary for the existence of x. It turns out that this is also sufficient.

Theorem 17 (gCRT)**.** *For arbitrary naturals $n_1, \ldots, n_d$ and $c_1, \ldots, c_d$, the system of congruences*

$$x \equiv c_i \ mod \ n_i, \quad 1 \leq i \leq d$$

has a unique solution modulo $lcm(n_1, \ldots, n_d)$ if and only if for all $1 \leq i, j \leq d$ we have that $c_i \equiv c_j \ mod \ gcd(n_i, n_j)$.

Proof. As discussed before, we only need to show sufficiency.

Note that the congruence $x \equiv c \bmod mn$ is equivalent to $x \equiv c \bmod m$ and $x \equiv c \bmod n$ if $gcd(m, n) = 1$. One direction is immediate. For the other direction, there exist naturals k, ℓ such that $x = km + c = \ell n + c$ implying $km = \ell n$. Since m and n are coprime this means $k = k'n$ and $\ell = \ell' m$. Therefore, $x = c + k'mn$ so $x \equiv c \bmod mn$.

It follows that each congruence $x \equiv c_i \bmod n_i$ can be replaced with the system of congruences $x \equiv c_i \bmod \pi_\ell^{a_\ell^i}$ for $1 \leq \ell \leq k_i$ where $n_i = \prod_{\ell=1}^{k_i} \pi_\ell^{a_\ell^i}$ is the prime factorization of n_i.

Let $lcm(n_1, \ldots, n_d) = \prod_{\ell=1}^{k} \pi_\ell^{a_\ell}$ be the prime factorization of the least common multiple of all the moduli.

The condition $c_i \equiv c_j \bmod gcd(n_i, n_j)$ for all $1 \leq i, j \leq d$ along with the observation that $x \equiv c \bmod \pi_\ell^{a_\ell}$ implies $x \equiv c \bmod \pi_\ell^{a}$ for any $a \leq a_\ell$ means that the original system of congruences is equivalent to the system

$$x \equiv c_{\phi(\ell)} \bmod \pi_\ell^{a_\ell}, \quad 1 \leq \ell \leq k,$$

where $\phi : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, d\}$ is a function that associates to each ℓ the index of one of the n_i's that is divisible by $\pi_\ell^{a_\ell}$.

The conclusion follows immediately by applying the classical CRT to this new system of congruences. □

Lemma 3. *Let a, b, c be pairwise coprime numbers. Let x, z be the inverses of a in $\mathbb{Z}/b\mathbb{Z}$ and $\mathbb{Z}/(bc)\mathbb{Z}$ respectively. Then there exists some integer d such that $z = db + x$.*

Proof. Let y be the inverse of a in $\mathbb{Z}/c\mathbb{Z}$. We know there exist some integers k, ℓ such that $ax = kb + 1$ and $ay = \ell c + 1$. Set $z := x + (c-k)by$. Then

$$\begin{aligned} az &= ax + a(c-k)by = kb + 1 + abcy - kb(ay) = \\ &= kb + 1 + abcy - kb(\ell c + 1) = kb + 1 + abcy - kb\ell c - kb = \\ &= 1 + bc(ay - k\ell). \end{aligned}$$

We see that z is indeed the inverse modulo bc, so we can take $d := (c-k)y$. □

We are now ready to state the sufficient condition result.

Theorem 18. *Let $\frac{p_1}{q_1}, \ldots, \frac{p_d}{q_d}$ be fractions inside $(0,1)$ in reduced form, i.e., $gcd(p_i, q_i) = 1$.*

Let $\kappa := \lceil \max_{i,j} \frac{q_i}{q_j} \rceil$ and let $P := \{\pi_1 < \cdots < \pi_m\}$ be the set of primes that appear in the factorization of $lcm(q_1, \ldots, q_d) = \prod_{j=1}^{m} \pi_j^{\alpha_j}$. Then we can write each $q_i = \prod_{j=1}^{m} \pi_j^{\alpha_j^i}$ for some sequence satisfying $0 \leq \alpha_j^i \leq \alpha_j$ for all $1 \leq i \leq m$.

Assume there exists a cutoff index $1 \leq m_0 < m$ such that $\pi_{m_0} > \kappa^d \prod_{j=1}^{m_0-1} \pi_j^{\alpha_j}$ and such that for each $j \geq m_0$ there exists an index i_j satisfying $\alpha_j^{i_j} = \alpha_j$ and $\alpha_j^i = 0$ for $i \neq i_j$. Finally assume that for each i, $\sum_{j=m_0}^{m} \alpha_j^i > 0$.

Equivalently, we can write each $q_i = u_i v_i$, where $u_i = \prod_{j=1}^{m_0-1} \pi_j^{\alpha_j^i}$ and $v_i = \prod_{j=m_0}^{m} \pi_j^{\alpha_j^i}$ and let $U := lcm(u_1, \ldots, u_d)$. The assumption is equivalent to saying that the v_i are pairwise coprime, each is coprime with U, each is strictly bigger than 1 and any prime divisor of any of the v_i is bigger than $\kappa^d U$.

Then given an arbitrary permutation $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$ there exist infinitely many primes N such that $\{\frac{Np_{\sigma(1)}}{q_{\sigma(d)}}\} < \cdots < \{\frac{Np_{\sigma(d)}}{q_{\sigma(d)}}\}$.

Proof. First, note that $\frac{Np}{q} = k + \frac{s}{q}$ for some positive k and $s \in \{1, \ldots, q-1\}$ if p and q are coprime. Therefore the required ordering boils down to finding a sequence $s = (s_1, \ldots, s_d)$ such that $\frac{s_1}{q_{\sigma(1)}} < \cdots < \frac{s_d}{q_{\sigma(d)}} < 1$.

First let us denote for $1 \leq i \leq d$ the reductions of p_i modulo q_i by p_i', i.e., let $p_i' \equiv p_i \bmod q_i$ such that $p_i' < u_i$ (since p_i and q_i are coprime). Now we can define the sequence s. Let $s_1 := p'_{\sigma(1)}$ and $s_i := p'_{\sigma(i)} + U \sum_{j=1}^{i-1} \kappa^j$ for $2 \leq i \leq d$. Note that since the v_i were pairwise coprime and each was bigger than 1 by assumption, $\kappa \geq 2$. Also note that

$$\begin{aligned} s_i = p'_{\sigma(i)} + U \sum_{j=1}^{i-1} \kappa^j \leq u_{\sigma(i)} + U \sum_{j=1}^{d-1} \kappa^j \leq \\ \leq U \sum_{j=0}^{d-1} \kappa^j = U \tfrac{\kappa^d - 1}{\kappa - 1} < U\kappa^d \leq \\ \leq \min_i q_i \end{aligned}$$

since, by assumption, each q_i is divisible by a prime bigger or equal to $U\kappa^d$.

Let us now check that this choice of the sequence s satisfies the desired order, i.e., that

$$\frac{p_{\sigma(1)}}{q_{\sigma(1)}} < \frac{p_{\sigma(2)}+\kappa U}{q_{\sigma(2)}} < \frac{p_{\sigma(3)}+(\kappa^2+\kappa)U}{q_{\sigma(3)}} < \cdots < \frac{p_{\sigma(d)}+(\kappa^{d-1}+\kappa^{d-2}+\cdots+1)U}{q_{\sigma(d)}}.$$

Using the definition of κ (and the fact that the v_i are pairwise coprime) and the fact that $0 \leq p'_i \leq u_i \leq U$, observe that $\frac{s_{i+1}}{s_i} \geq \frac{U\sum_{j=1}^{i-1}\kappa^j}{U+U\sum_{j=1}^{i-2}\kappa^j} = \kappa > \frac{q_{\sigma(i+1)}}{q_{\sigma(i)}}$ for $i \geq 2$ and trivially $\frac{s_2}{s_i} \geq k > \frac{q_{\sigma(2)}}{q_{\sigma(1)}}$. So the sequence s gives the needed order.

We now claim that we can find infinitely many primes N that satisfy the congruence system

$$Np_{\sigma(i)} \equiv s_i \bmod q_{\sigma(i)} \tag{6}$$

for $1 \leq i \leq d$.

For $1 \leq i \leq d$, let r_i be the inverse of p_i in $\mathbb{Z}/q_i\mathbb{Z}$, which exists by assumption that the fractions are in reduced form and let t_i be the inverse of p_i (and therefore of p'_i) in $\mathbb{Z}/u_i\mathbb{Z}$. Applying Lemma 3, there exist integers k_i such that $r_i = t_i + k_i u_i$.

We can rewrite the system of congruences that we need to solve as

$$N \equiv s_i r_{\sigma(i)} \bmod q_{\sigma(i)}. \tag{7}$$

Using the above notation, we have that

$$s_i r_{\sigma(i)} = (p'_{\sigma(i)} + U\sum_{j=1}^{i-1}\kappa^j)(t_{\sigma(i)} + k_{\sigma(i)}u_{\sigma(i)}) = 1 + \ell_i u_{\sigma(i)}$$

for some integer ℓ_i, $1 \leq i \leq d$.

Now, $s_i r_{\sigma(i)} - s_j r_{\sigma(j)} = \ell_i u_{\sigma(i)} - \ell_j u_{\sigma(j)}$, so it is divisible by $gcd(u_{\sigma(i)}, u_{\sigma(j)})$ which is equal by construction to $gcd(q_{\sigma(i)}, q_{\sigma(j)})$, $1 \leq i, j \leq d$. Therefore, we can apply gCRT to conclude that there exists some $1 \leq N_0 \leq lcm(q_1, \ldots, q_d) - 1$ such that for any $k \geq 0$, the number $N = N_0 + k \cdot lcm(q_1, \ldots, q_d)$ satisfies the system of congruences (7) and further on (6).

Next we claim that N_0 is coprime with $Q := lcm(q_1, \ldots, q_d)$.

Recall that the list of prime divisors of Q was $P = \{\pi_1 < \cdots < \pi_m\}$. So assume there exists some index $j \leq m$ such that π_j divides N_0. Let $1 \leq i \leq d$ be any index such that π_j divides $q_{\sigma(i)} = u_{\sigma(i)}v_{\sigma(i)}$. Since N_0 satisfies the system of congruences (7), it follows that π_j also divides $s_i r_{\sigma(i)}$.

Assume π_j divides $u_{\sigma(i)}$. Since $s_i r_{\sigma(i)} \equiv 1 \bmod u_{\sigma(i)}$, this implies π_j divides 1, a contradiction. It follows that π_j does not divide any u_i, or, equivalently, that the index j satisfies $j \geq m_0$. In particular, by construction, we have that $\pi_j \geq U\kappa^d$.

Now, since π_j is a prime, it must divide either s_i or $r_{\sigma(i)}$. However, as we have shown before, each s_i is strictly smaller than $U\kappa^d$, which implies that there exists some integer k such that $k\pi_j = r_{\sigma(i)}$. This means that $k\pi_j$ is the inverse of $p_{\sigma(i)}$ in $\mathbb{Z}/q_{\sigma(i)}\mathbb{Z}$ which is impossible since π_j divides $q_{\sigma(i)}$.

We have a contradiction, therefore N_0 must be coprime with Q.

The proof is completed by applying Dirichlet's theorem on arithmetic progressions (a 'folklore' result, a proof of which can be found in [Ser73]), which guarantees that any arithmetic progression $a+kb$ contains infinitely many primes whenever a and b are coprime. □

Remark 11. *The requirements of Theorem 18 are neither easy to state in a non-convoluted but still precise way, nor are they too mild.*

Informally speaking, what is required for the machinery of Theorem 18 to work is that one can partition the list of prime divisors of the least common multiple of the denominators $q_1, \dots, q_d$ into two subsets of primes, U and V. Any prime divisor common to multiple denominators must come from the set U and each q_j must have at least one divisor from the set V. Finally, the set V must be sufficiently separated from the set U, meaning that the smallest prime in V must be sufficiently large compared to any prime in U, and the notion of 'sufficiently large' depends on the largest ratio between denominators and how many different rational numbers we are trying to reorder.

For instance, cases when at least two denominators are equal are not covered by Theorem 18. However, in a scenario like

$$S = [\tfrac{7}{2018}, \tfrac{45}{2206}] \cup [\tfrac{49}{1237}, \tfrac{101}{2186}] \cup [\tfrac{200}{1823}, \tfrac{200}{1399}] \cup [\tfrac{1000}{1669}, \tfrac{1500}{1847}],$$

Theorem 18 does apply and leads to the existence of a hierarchical Riesz exponential basis for S.

5.3.2 A brief discussion on Kadec bases with non-trivial intersection

We begin this discussion with a definition of Kadec bases based on the Kadec $\frac{1}{4}$ theorem.

Definition 10. *Let $a > 0$ be a real number. In view of the Kadec $\frac{1}{4}$ theorem, a separated set $\Lambda = \{an + \alpha + c_n : n \in \mathbb{Z} \text{ and } \alpha \in (-\frac{a}{2}, \frac{a}{2}]\}$ is called an a-**Kadec basis** (for any interval of length $\frac{1}{a}$) if there exists some $0 \leq \delta < \frac{1}{4}$ such that $|c_n| \leq a\delta$ for all $n \in \mathbb{Z}$.*

In the particular case $a = 1$, we can uniquely specify a 1-Kadec basis through a function $f : \mathbb{Z} \to \mathbb{R}$, $n \mapsto n + \beta + c_n$ for any choice of $\beta \in (-\frac{1}{2}, \frac{1}{2}]$ and any sequence $\|\{c_n\}\|_\infty < \frac{1}{4}$.

Remark 12. *The condition $\alpha \in (-\frac{a}{2}, \frac{a}{2}]$ is not an actual restriction, since any choice of a global shift factor α can be reduced modulo a at the cost of possibly reindexing the sequence $\{c_n\}$.*

The motivation behind the present chapter is a surprising construction due to D. G. Lee that shows that two 2-Kadec bases with non-trivial intersection can have a union that is a 1-Kadec basis.

Before discussing this example, let us note for the sake of completion that the typical or 'non-pathologic' situation is that the union of two 2-Kadec bases 'should' constitute a 1-Kadec basis only if they do not intersect.

The prototypical example would be the 2-Kadec bases $\Lambda_1 = 2\mathbb{Z}$ and $\Lambda_2 = 2\mathbb{Z}+1$, corresponding, respectively, to $\alpha = 0$ and $c_n = 0$, and $\alpha = 1$ and $c_n = 0$. Clearly, their union corresponds to the standard Fourier basis frequencies.

An example of non-trivial intersection due to D. G. Lee

Fix an arbitrary $0 < \varepsilon < \frac{1}{4}$ and define the frequency sets

$$\Lambda_1 = \{\dots, -4, -2, 0, 1+\varepsilon, 3+\varepsilon, \dots\} = \{2n - \tfrac{1-\varepsilon}{2} + c_n : n \in \mathbb{Z}\}$$

and

$$\Lambda_2 = -\Lambda_1,$$

where

$$c_n = \begin{cases} \frac{1-\varepsilon}{2} & \text{if } n \leq 0 \\ -\frac{1-\varepsilon}{2} & \text{if } n > 0. \end{cases}$$

It follows that $|c_n| = |\frac{1-\varepsilon}{2}| < \frac{1}{2}$, so Λ_1 and therefore Λ_2 are both 2-Kadec bases. Additionally, $\Lambda_1 \cap \Lambda_2 = \{0\}$ and

$$\Lambda := \Lambda_1 \cup \Lambda_2 = \{\dots, -4, -3-\varepsilon, -2, -1-\varepsilon, 0, 1+\varepsilon, 2, 3+\varepsilon, 4, \dots\}.$$

Since $\varepsilon < \frac{1}{4}$, Λ is clearly a 1-Kadec basis.

It is clear that any a-Kadec basis will satisfy an appropriately dilated version of Avdonin's theorem and that perturbing finitely many point of the a-Kadec basis will again result is a sequence that again satisfies Avdonin's theorem (on a possibly larger averaging window) as long as the perturbation does not 'collapse' different points to a single point.

Using this idea, we can shift the intersection point of the two bases arbitrarily. For example, if we set $\Lambda_1' = \{\dots, -4, -2, \varepsilon, 2, 3+\varepsilon, 5+\varepsilon, \dots\}$ and $\Lambda_2' = \Lambda_2$, we would have that $\Lambda_1' \cup \Lambda_2'$ is still Riesz spectral for $[0, 1]$, but $\Lambda_1' \cap \Lambda_2' = \{2\}$.

However perturbations that interfere with the number of intersection points will not preserve the basis property.

For example, taking $\Lambda_1' = \{\dots, -4, -2, \varepsilon, 1+\varepsilon, 3+\varepsilon, \dots\}$ and $\Lambda_2' = \Lambda_2$ would result in $\Lambda_1' \cap \Lambda_2' = \varnothing$ and $\Lambda_1' \cup \Lambda_2' = \Lambda_1 \cup \Lambda_2 \cup \{\varepsilon\}$ which can therefore not be a basis for $[0, 1]$ as it will have redundancy.

Alternatively, taking $\Lambda_1' = \{\dots, -4, -2, 0, 2, 3+\varepsilon, 5+\varepsilon, \dots\}$ and $\Lambda_2' = \Lambda_2$ will give $\Lambda_1' \cap \Lambda_2' = \{0, 2\}$ and $\Lambda_1' \cup \Lambda_2' = \big(\Lambda_1 \cup \Lambda_2\big) \setminus \{1+\varepsilon\}$ which cannot be a basis for $[0, 1]$ due to the minimality of any basis.

A limitation on the size of the intersection of 2-Kadec bases

We will show that if two 2-Kadec bases Λ_1 and Λ_2 are such that their union $\Lambda := \Lambda_1 \cup \Lambda_2$ is a 1-Kadec basis, then Λ_1 and Λ_2 cannot have more than one point in common. Throughout the remainder of this discussion, we shall assume that $\Lambda_1 = \{2n + \alpha_1 + c_{1,n} : n \in \mathbb{Z}\}$ and $\Lambda_2 = \{2n + \alpha_2 + c_{2,n} : n \in \mathbb{Z}\}$ are 2-Kadec bases such that $\Lambda_1 \cup \Lambda_2 =: \Lambda = f(\mathbb{Z})$, where $f(n) = n + \beta + c_n$ is a 1-Kadec basis. Recall that this implies that $\alpha_1, \alpha_2 \in (-1, 1]$, that $\beta \in (-\frac{1}{2}, \frac{1}{2}]$, that $\|\{c_{1,n}\}\|_\infty, \|\{c_{2,n}\}\|_\infty < \frac{1}{2}$ and that $\|\{c_n\}\|_\infty < \frac{1}{4}$.

Due to this assumption, it follows that for $n \in \mathbb{Z}$ we must have that either $f(n) \in \Lambda_1$ or that $f(n) \in \Lambda_2$. As we pointed out in the examples before, the most 'natural' setup that allows this is one where the elements of Λ_1 interlace those of Λ_2, i.e., if $f(n) \in \Lambda_1$ then $f(n+1) \in \Lambda_2$ and vice-versa. We first show that breaking this interlacing structure comes with specific restrictions.

Lemma 4. *It is impossible for 3 consecutive integers to be represented by the same 2-Kadec basis. In other words, there can exist no $k \in \mathbb{Z}$ such that $f(k), f(k+1), f(k+2)$ are elements, w.l.o.g., of Λ_1. Additionally if $f(k), f(k+1)$ are both elements of, w.l.o.g., Λ_1, then they are represented by elements with consecutive indices.*

Proof. Let's assume this is possible. Then there exist integers n, m, p such that for some $\alpha_1, \beta, \{c_{1,n}\}_n, \{c_n\}_n$ as described before, the following system of equations is satisfied:

$$\begin{cases} 2n + \alpha_1 + c_{1,n} & = k + \beta + c_k \\ 2(n+m) + \alpha_1 + c_{1,n+m} & = k + 1 + \beta + c_{k+1} \\ 2(n+m+p) + \alpha_1 + c_{1,n+m+p} & = k + 2 + \beta + c_{k+2}. \end{cases}$$

Subtracting the first equation from the second and rearranging we get that

$$2m - 1 = (c_{k+1} - c_k) - (c_{1,n+m} - c_{1,n}),$$

which implies that $2m - 1 \in (-\frac{3}{2}, \frac{3}{2})$ which is only possible if $m = 1$. This takes care of the additional statement of the lemma. We can do the same with the last two equations to see that $p = 1$. We can rewrite the system of equations as

$$\begin{cases} 2n + \alpha_1 + c_{1,n} & = k + \beta + c_k \\ 2(n+1) + \alpha_1 + c_{1,n+1} & = k + 1 + \beta + c_{k+1} \\ 2(n+2) + \alpha_1 + c_{1,n+2} & = k + 2 + \beta + c_{k+2}. \end{cases}$$

Subtracting the first equation from the third and rearranging we get

$$2 = (c_{k+2} - c_k) - (c_{1,n+2} - c_{1,n}),$$

which is impossible as $2 \notin (-\frac{3}{2}, \frac{3}{2})$. □

Corollary 1. *If $f(k), f(k+1) \in \Lambda_1$, then necessarily $f(k-1), f(k+2) \in \Lambda_2$ or vice-versa (with respect to swapping Λ_1 with Λ_2). Therefore if $f(k) \in \Lambda_1 \cap \Lambda_2$, then necessarily $f(k-1) \in \Lambda_1$ and $f(k+1) \in \Lambda_2$ or vice-versa. Put differently, this means that $f(k-1), f(k) \in \Lambda_1$ and $f(k), f(k+1) \in \Lambda_2$ or vice-versa. In other words, a point in the intersection $\Lambda_1 \cap \Lambda_2$ induces the appearance of consecutive elements from Λ in both Λ_1 and Λ_2.*

Lemma 5. *Assume that for some integers k, ℓ we have that $f(k), f(k+1), f(k+\ell), f(k+\ell+1)$ are elements, w.l.o.g., of Λ_1. Then necessarily ℓ is even.*

Proof. Using Lemma 4, this implies that for some integers n, m, the following system is satisfied

$$\begin{cases} 2n + \alpha_1 + c_{1,n} & = k + \beta + c_k \\ 2(n+1) + \alpha_1 + c_{1,n+1} & = k + 1 + \beta + c_{k+1} \\ 2(n+m) + \alpha_1 + c_{1,n+m} & = k + \ell + \beta + c_{k+\ell} \\ 2(n+m+1) + \alpha_1 + c_{1,n+m+1} & = k + \ell + 1 + \beta + c_{k+\ell+1}. \end{cases}$$

Subtracting the second equation from the third, the first from the third and the first from the fourth, respectively, and rearranging we obtain

$$\begin{cases} 2m - \ell - 1 & = (c_{k+\ell} - c_{k+1}) - (c_{1,n+m} - c_{1,n+1}) \\ 2m - \ell & = (c_{k+\ell} - c_k) - (c_{1,n+m} - c_{1,n}) \\ 2m - \ell + 1 & = (c_{k+\ell+1} - c_k) - (c_{1,n+m+1} - c_{1,n}), \end{cases}$$

which implies that $2m - \ell - 1, 2m - \ell$ and $2m - \ell + 1$ are different integers simultaneously contained in the interval $(-\frac{3}{2}, \frac{3}{2})$. This necessarily means that $2m - \ell = 0$, i.e., that $\ell = 2m$, so ℓ is even. □

Theorem 19. *Let Λ_1 and Λ_2 be 2-Kadec bases such that $\Lambda := \Lambda_1 \cup \Lambda_2$ is a 1-Kadec basis. Then $|\Lambda_1 \cap \Lambda_2| \leq 1$.*

Proof. Let us define the "defect" index set for Λ split between Λ_1 and Λ_2. Let $D_s := \{k \in \mathbb{Z} : f(k), f(k+1) \in \Lambda_s\}$ for $s \in \{1, 2\}$. With this notation, we can summarize the conclusions of Lemmas 4 and 5 and of Corollary 1 as follows (for $s \in \{1, 2\}$):

- if $k \in D_s$, then $k + 1 \notin D_s$;
- if $k, \ell \in D_s$, then $k \equiv \ell \mod 2$;
- $|D_s| \geq |\Lambda_1 \cap \Lambda_2|$. In fact for each $k \in \Lambda_1 \cap \Lambda_2$, $k \in D_1$ and $k - 1 \in D_2$ or vice-versa.

Now assume the contrary. This implies that D_1 and D_2 have each at least 2 elements. By relabeling if necessary (swapping between Λ_1 and Λ_2) we can assume in view of these properties that D_1 contains only even indices and D_2 only odd indices. By assumption, there exists some $k \in \mathbb{Z}$ such that $f(k) \in \Lambda_1 \cap \Lambda_2$ and for some $k' > k + 1$ (note that as pointed out before we cannot have $f(k)$ and $f(k+1)$ simultaneously in the intersection) we also have that $f(k') \in \Lambda_1 \cap \Lambda_2$. W.l.o.g., we assume that k is even (the odd case can be treated the same way). This will imply that $k \in D_1$ (otherwise $k \in D_2$).

By assumption, there exists a smallest integer $\ell \geq 1$ such that $k, k + 2\ell \in D_1$ and $\{k + 1, \ldots, k + 2\ell - 1\} \cap D_1 = \varnothing$. We now distinguish between two cases.

Case 1. Assume $\{k + 1, \ldots, k + 2\ell - 1\} \cap D_2 = \varnothing$. This implies that Λ_1 and Λ_2 interlace on the set $\{f(k+1), f(k+2), \ldots, f(k+2\ell-1)\}$. In other words, $f(k+1) \in \Lambda_1$, $f(k+2) \in \Lambda_2$, $f(k+3) \in \Lambda_1$ and so on. This implies that

$f(odd) \in \Lambda_1$ for all $k < odd < k+2\ell$. In particular $f(k+2\ell-1) \in \Lambda_1$. But since $f(k+2\ell) \in \Lambda_1$ by assumption, this means that $k+2\ell-1 \in D_1$, contradiction.

Case 2. Assume $\{k+1, \ldots, k+2\ell-1\} \cap D_2 \neq \varnothing$ and let $k+r$ be the smallest element in the intersection. Note that this implies that $r = 2s+1$ for some natural s, since r must be odd. Once again we must have that Λ_1 and Λ_2 interlace on the set $\{f(k+1), f(k+2), \ldots, f(k+2s+1)\}$. Once again $f(odd) \in \Lambda_1$ for all $k < odd \leq k+2s+1$. In particular $f(k+2s+1) \in \Lambda_1$ and since $k+2s+1 \in D_2$, it follows that $f(k+2s+1) \in \Lambda_2$ as well. This implies that $s > 0$, since otherwise we would have that both $f(k)$ and $f(k+1)$ are simultaneously in the intersection $\Lambda_1 \cap \Lambda_2$, which would imply that either $k-1, k \in D_1$ or $k-1, k \in D_2$, a contradiction. So we are left with $f(k+2s+1) \in \Lambda_1 \cap \Lambda_2$, which implies, taking parity into account, that $k+2s \in D_1$. However, since $s > 0$, $k+2s \in \{k+1, k+2, \ldots, k+2\ell-1\} \cap D_1$, again a contradiction. $\square$

References

[AI95] Sergei A. Avdonin and Sergei A. Ivanov. *Families of exponentials. The method of moments in controllability problems for distributed parameter systems*, Translated from the Russian and revised by the authors. Cambridge University Press, Cambridge, 1995, pp. xvi+302. ISBN: 0-521-45243-0.

[ALS19] Z. Allen-Zhu, Y. Li, and Z. Song. "A Convergence Theory for Deep Learning via Over-Parameterization". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, Sept. 2019, pp. 242–252. URL: https://proceedings.mlr.press/v97/allen-zhu19a.html.

[Avd74] S. A. Avdonin. "On the question of Riesz bases of exponential functions in L^2". In: *Vestnik Leningrad. Univ. No. 13 Mat. Meh. Astronom.* Vyp. 3 (1974), pp. 5–12, 154.

[Avd91] S. A. Avdonin. "The existence of basis subfamilies of a Riesz basis from exponentials". In: *Vestnik Leningrad. Univ. Mat. Mekh. Astronom.* vyp. 3 (1991), pp. 115–116, 134. ISSN: 0024-0850.

[Bal81] R. Balian. "Un principe d'incertitude fort en théorie du signal ou en mécanique quantique". In: *C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre* 292.20 (1981), pp. 1357–1362. ISSN: 0249-6305.

[Bar+19] P. L. Bartlett, N. Harvey, C. Liaw, and A. Mehrabian. "Nearly-tight VC-dimension and pseudodimension bounds for piecewise linear neural networks". In: *J. Mach. Learn. Res.* 20 (2019), Paper No. 63, 17. ISSN: 1532-4435.

[Bar93] A. R. Barron. "Universal approximation bounds for superpositions of a sigmoidal function". In: *IEEE Trans. Inform. Theory* 39.3 (1993), pp. 930–945. ISSN: 0018-9448. DOI: 10.1109/18.256500. URL: https://doi.org/10.1109/18.256500.

[Bat88] G. Battle. "Heisenberg proof of the Balian-Low theorem". In: *Lett. Math. Phys.* 15.2 (1988), pp. 175–177. ISSN: 0377-9017. DOI: 10.1007/BF00397840. URL: https://doi.org/10.1007/BF00397840.

[Bel+19] M. Belkin, D. Hsu, S. Ma, and S. Mandal. "Reconciling modern machine-learning practice and the classical bias-variance trade-off". In: *Proc. Natl. Acad. Sci. USA* 116.32 (2019), pp. 15849–15854. ISSN: 0027-8424. DOI: 10.1073/pnas.1903070116. URL: https://doi.org/10.1073/pnas.1903070116.

[BHW95] J. J. Benedetto, C. Heil, and D. Walnut. "Differentiation and the Balian-Low theorem". In: *J. Fourier Anal. Appl.* 1.4 (1995), pp. 355–402. ISSN: 1069-5869. DOI: 10.1007/s00041-001-4016-5. URL: https://doi.org/10.1007/s00041-001-4016-5.

[Bro+17] T. B. Brown, D. Mané, A. Roy, M. Abadi, and J. Gilmer. "Adversarial Patch". In: *arXiv:1712.09665* (2017).

[Car+19] A. Caragea, D. G. Lee, G. Pfander, and F. Philipp. "A Balian-low theorem for subspaces". In: *J. Fourier Anal. Appl.* 25.4 (2019), pp. 1673–1694. ISSN: 1069-5869. DOI: 10.1007/s00041-018-9634-2. URL: https://doi.org/10.1007/s00041-018-9634-2.

[Car+21] A. Caragea, D. G. Lee, F. Philipp, and F. Voigtlaender. "A quantitative subspace Balian-Low theorem". In: *Appl. Comput. Harmon. Anal.* 55 (2021), pp. 368–404. ISSN: 1063-5203. DOI: 10.1016/j.acha.2021.06.005. URL: https://doi.org/10.1016/j.acha.2021.06.005.

[Car+22] A. Caragea, D. G. Lee, J. Maly, G. Pfander, and F. Voigtlaender. "Quantitative approximation results for complex-valued neural networks". In: *SIAM J. Math. Data Sci.* 4.2 (2022), pp. 553–580. DOI: 10.1137/21M1429540. URL: https://doi.org/10.1137/21M1429540.

[Car+23] Andrei Caragea, Dae Gwan Lee, Friedrich Philipp, and Felix Voigtlaender. "A Balian-Low type theorem for Gabor Riesz sequences of arbitrary density". In: *Math. Z.* 303.2 (2023), Paper No. 48, 22. ISSN: 0025-5874. DOI: 10.1007/s00209-022-03182-6. URL: https://doi.org/10.1007/s00209-022-03182-6.

[CL22] A. Caragea and D. G. Lee. "A note on exponential Riesz bases". In: *Sampl. Theory Signal Process. Data Anal.* 20.2 (2022), Paper No. 13, 14. ISSN: 2730-5716. DOI: 10.1007/s43670-022-00031-9. URL: https://doi.org/10.1007/s43670-022-00031-9.

[CMP16] C. Cabrelli, U. Molter, and G. Pfander. "Time-frequency shift invariance and the amalgam Balian-Low theorem". In: *Appl. Comput. Harmon. Anal.* 41.3 (2016), pp. 677–691. ISSN: 1063-5203. DOI: 10.1016/j.acha.2015.04.003. URL: https://doi.org/10.1016/j.acha.2015.04.003.

[CPV20] A. Caragea, P. Petersen, and F. Voigtlaender. "Neural network approximation and estimation of classifiers with classification boundary in a Barron class". In: *arXiv preprint arXiv:2011.09363* (2020).

[Dau90] I. Daubechies. "The wavelet transform, time-frequency localization and signal analysis". In: *IEEE Trans. Inform. Theory* 36.5 (1990), pp. 961–1005. ISSN: 0018-9448. DOI: 10.1109/18.57199. URL: http://dx.doi.org/10.1109/18.57199.

[Den+09] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li F.-F. "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

[Duv+15] D. K. Duvenaud, J. Maclaurin D.and Iparraguirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams. "Convolutional networks on graphs for learning molecular fingerprints". In: *Advances in neural information processing systems* 28 (2015).

[Elb+22] D. Elbrächter, P. Grohs, A. Jentzen, and C. Schwab. "DNN expression rate analysis of high-dimensional PDEs: application to option pricing". In: *Constr. Approx.* 55.1 (2022), pp. 3–71. ISSN: 0176-4276. DOI: 10.1007/s00365-021-09541-6. URL: https://doi.org/10.1007/s00365-021-09541-6.

[EY18] W. E and B. Yu. "The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems". In: *Commun. Math. Stat.* 6.1 (2018), pp. 1–12. ISSN: 2194-6701. DOI: 10.1007/s40304-018-0127-z. URL: https://doi.org/10.1007/s40304-018-0127-z.

[Eyk+18] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, and D. Song. "Robust Physical-World Attacks on Deep Learning Visual Classification". In: *Proceedings of CVPR*. June 2018.

[Fei81] H. G. Feichtinger. "On a new Segal algebra". In: *Monatsh. Math.* 92.4 (1981), pp. 269–289. ISSN: 0026-9255. DOI: 10.1007/BF01320058. URL: https://doi.org/10.1007/BF01320058.

[Fra85] E. Francis. *Low, Complete sets of wave packets, A passion for physics-essays in honor of Geoffrey Chew (C. DeTar et al., ed.)* 1985.

[FS07] H. G. Feichtinger and W. Sun. "Sufficient conditions for irregular Gabor frames". In: *Adv. Comput. Math.* 26.4 (2007), pp. 403–430. ISSN: 1019-7168. DOI: 10.1007/s10444-004-7210-6. URL: https://doi.org/10.1007/s10444-004-7210-6.

[Fug74] B. Fuglede. "Commuting self-adjoint partial differential operators and a group theoretic problem". In: *J. Functional Analysis* 16 (1974), pp. 101–121. DOI: 10.1016/0022-1236(74)90072-x. URL: https://doi.org/10.1016/0022-1236(74)90072-x.

[Gau08] S. Z. Gautam. "A critical-exponent Balian-Low theorem". In: *Math. Res. Lett.* 15.3 (2008), pp. 471–483. ISSN: 1073-2780. DOI: 10.4310/MRL.2008.v15.n3.a7. URL: https://doi.org/10.4310/MRL.2008.v15.n3.a7.

[GH21] P. Grohs and L. Herrmann. "Deep neural network approximation for high-dimensional parabolic Hamilton-Jacobi-Bellman equations". In: *arXiv preprint arXiv:2103.05744* (2021).

[Gri+22] R. Gribonval, G. Kutyniok, M. Nielsen, and F. Voigtlaender. "Approximation spaces of deep neural networks". In: *Constr. Approx.* 55.1 (2022), pp. 259–367. ISSN: 0176-4276. DOI: 10.1007/s00365-021-09543-4. URL: https://doi.org/10.1007/s00365-021-09543-4.

[Grö01] K. Gröchenig. *Foundations of time-frequency analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2001, pp. xvi+359. ISBN: 0-8176-4022-3. DOI: 10.1007/978-1-4612-0003-1. URL: https://doi.org/10.1007/978-1-4612-0003-1.

[GRS18] N. Golowich, A. Rakhlin, and O. Shamir. "Size-Independent Sample Complexity of Neural Networks". In: *Proceedings of the 31st Conference On Learning Theory*. Ed. by Sébastien Bubeck, Vianney Perchet, and Philippe Rigollet. Vol. 75. Proceedings of Machine Learning Research. PMLR, June 2018, pp. 297–299. URL: https://proceedings.mlr.press/v75/golowich18a.html.

[GSS14] I. J. Goodfellow, J. Shlens, and C. Szegedy. "Explaining and harnessing adversarial examples". In: *arXiv preprint arXiv:1412.6572* (2014).

[Gun+17] S. Gunasekar, B. E. Woodworth, S. Bhojanapalli, B. Neyshabur, and N. Srebro. "Implicit regularization in matrix factorization". In: *Advances in Neural Information Processing Systems* 30 (2017).

[GV21] P. Grohs and F. Voigtlaender. "Proof of the theory-to-practice gap in deep learning via sampling complexity bounds for neural network approximation spaces". In: *arXiv preprint arXiv:2104.02746* (2021).

[He+15] K. He, X. Zhang, S. Ren, and J. Sun. "Delving deep into rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *Proceedings of ICCV*. 2015, pp. 1026–1034.

[HS17] B. Hanin and M. Sellke. "Approximating continuous functions by relu nets of minimal width". In: *arXiv preprint arXiv:1710.11278* (2017).

[HSN20] J. Hermann, Z. Schätzle, and F. Noé. "Deep-neural-network solution of the electronic Schrödinger equation". In: *Nature Chemistry* 12.10 (2020), pp. 891–897.

[IKP99] A. Iosevich, N. Katz, and S. Pedersen. "Fourier bases and a distance problem of Erdős". In: *Math. Res. Lett.* 6.2 (1999), pp. 251–255. ISSN: 1073-2780. DOI: 10.4310/MRL.1999.v6.n2.a13. URL: https://doi.org/10.4310/MRL.1999.v6.n2.a13.

[Kad64] M. I. Kadets. "The exact value of the Paley–Wiener constant". In: *Doklady Akademii Nauk*. Vol. 155. 6. Russian Academy of Sciences. 1964, pp. 1253–1254.

[KM06] M. N. Kolountzakis and M. Matolcsi. "Complex Hadamard matrices and the spectral set conjecture". In: *Collect. Math.* Vol. Extra (2006), pp. 281–291. ISSN: 0010-0757.

[KM97] M. Karpinski and A. Macintyre. "Polynomial bounds for VC dimension of sigmoidal and general Pfaffian neural networks". In: vol. 54. 1, part 2. 1st Annual Dagstuhl Seminar on Neural Computing (1994). 1997, pp. 169–176. DOI: 10.1006/jcss.1997.1477. URL: https://doi.org/10.1006/jcss.1997.1477.

[KN15] G. Kozma and S. Nitzan. "Combining Riesz bases". In: *Invent. Math.* 199.1 (2015), pp. 267–285. ISSN: 0020-9910. DOI: 10.1007/s00222-014-0522-3. URL: https://doi.org/10.1007/s00222-014-0522-3.

[KNO21] G. Kozma, S. Nitzan, and A. Olevskii. "A set with no Riesz basis of exponentials". In: *arXiv preprint arXiv:2110.02090* (2021).

[KSH17] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "Imagenet classification with deep convolutional neural networks". In: *Communications of the ACM* 60.6 (2017), pp. 84–90.

[Les+93] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken. "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function". In: *Neural networks* 6.6 (1993), pp. 861–867.

[LMZ18] Y. Li, T. Ma, and H. Zhang. "Algorithmic regularization in over-parameterized matrix sensing and neural networks with quadratic activations". In: *Conference On Learning Theory*. PMLR. 2018, pp. 2–47.

[Mat05] M. Matolcsi. "Fuglede's conjecture fails in dimension 4". In: *Proc. Amer. Math. Soc.* 133.10 (2005), pp. 3021–3026. ISSN: 0002-9939. DOI: 10.1090/S0002-9939-05-07874-3. URL: https://doi.org/10.1090/S0002-9939-05-07874-3.

[MM22] S. Mei and A. Montanari. "The generalization error of random features regression: precise asymptotics and the double descent curve". In: *Comm. Pure Appl. Math.* 75.4 (2022), pp. 667–766. ISSN: 0010-3640. DOI: 10.1002/cpa.22008. URL: https://doi.org/10.1002/cpa.22008.

[Mon+19] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein. "Fake news detection on social media using geometric deep learning". In: *arXiv preprint arXiv:1902.06673* (2019).

[MRT18] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of machine learning*. Adaptive Computation and Machine Learning. Second edition of [MR3057769]. MIT Press, Cambridge, MA, 2018, pp. xv+486. ISBN: 978-0-262-03940-6.

[Pfa+20] D. Pfau, J. S. Spencer, A. GDG Matthews, and W M. C Foulkes. "Ab initio solution of the many-electron Schrödinger equation with deep neural networks". In: *Physical Review Research* 2.3 (2020), p. 033429.

[PRV21] P. Petersen, M. Raslan, and F. Voigtlaender. "Topological properties of the set of functions generated by neural networks of fixed size". In: *Found. Comput. Math.* 21.2 (2021), pp. 375–444. ISSN: 1615-3375. DOI: 10.1007/s10208-020-09461-0. URL: https://doi.org/10.1007/s10208-020-09461-0.

[PRW21] G. Pfander, S. Revay, and D. Walnut. "Exponential bases for partitions of intervals". In: *arXiv preprint arXiv:2109.04441* (2021).

[PV18] P. Petersen and F. Voigtlaender. "Optimal approximation of piecewise smooth functions using deep ReLU neural networks". In: *Neural Networks* 108 (2018), pp. 296–330.

[PV21] P. Petersen and F. Voigtlaender. "Optimal learning of high-dimensional classification problems using deep neural networks". In: *arXiv preprint arXiv:2112.12555* (2021).

[Sei95] K. Seip. "A simple construction of exponential bases in L^2 of the union of several intervals". In: *Proc. Edinburgh Math. Soc. (2)* 38.1 (1995), pp. 171–177. ISSN: 0013-0915. DOI: 10.1017/S0013091500006295. URL: https://doi.org/10.1017/S0013091500006295.

[Ser73] J.-P. Serre. *A course in arithmetic*. Graduate Texts in Mathematics, No. 7. Translated from the French. Springer-Verlag, New York-Heidelberg, 1973, pp. viii+115.

[SL96] P. Stevenhagen and H. W. Lenstra Jr. "Chebotarëv and his density theorem". In: *Math. Intelligencer* 18.2 (1996), pp. 26–37. ISSN: 0343-6993. DOI: 10.1007/BF03027290. URL: https://doi.org/10.1007/BF03027290.

[Sto+20] J. M. Stokes et al. "A deep learning approach to antibiotic discovery". In: *Cell* 180.4 (2020), pp. 688–702.

[SVL14] I. Sutskever, O. Vinyals, and Q. V. Le. "Sequence to sequence learning with neural networks". In: *Advances in neural information processing systems* 27 (2014).

[Sze+13] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus. "Intriguing properties of neural networks". In: *arXiv preprint arXiv:1312.6199* (2013).

[Tao04] T. Tao. "Fuglede's conjecture is false in 5 and higher dimensions". In: *Math. Res. Lett.* 11.2-3 (2004), pp. 251–258. ISSN: 1073-2780. DOI: 10.4310/MRL.2004.v11.n2.a8. URL: https://doi.org/10.4310/MRL.2004.v11.n2.a8.

[TVG19] S. Thys, W. Van Ranst, and T. Goedemé. "Fooling automated surveillance cameras: adversarial patches to attack person detection". In: *Proceedings of CVPR Workshops*. 2019.

[Voi20] F. Voigtlaender. "The universal approximation theorem for complex-valued neural networks". In: *arXiv preprint arXiv:2012.03351* (2020).

[VSL17] P. Virtue, X Yu S., and M. Lustig. "Better than real: Complex-valued neural nets for MRI fingerprinting". In: *2017 IEEE international conference on image processing (ICIP)*. IEEE. 2017, pp. 3953–3957.

[WY18] M. Wolter and A. Yao. "Complex gated recurrent neural networks". In: *Advances in neural information processing systems* 31 (2018).

[Yar17] D. Yarotsky. "Error bounds for approximations with deep ReLU networks". In: *Neural Networks* 94 (2017), pp. 103–114.

[Yar22] D. Yarotsky. "Universal approximations of invariant maps by neural networks". In: *Constr. Approx.* 55.1 (2022), pp. 407–474. ISSN: 0176-4276. DOI: 10.1007/s00365-021-09546-1. URL: https://doi.org/10.1007/s00365-021-09546-1.

[You01] R. M. Young. *An introduction to nonharmonic Fourier series*. first. Academic Press, Inc., San Diego, CA, 2001, pp. xiv+234. ISBN: 0-12-772955-0.

6 Full texts of the six papers

6.1 Paper 1

A BALIAN-LOW THEOREM FOR SUBSPACES

ANDREI CARAGEA, DAE GWAN LEE, GÖTZ E. PFANDER, AND FRIEDRICH PHILIPP

Abstract. We extend the Balian-Low theorem to Gabor subspaces of $L^2(\mathbb{R})$ by involving the concept of additional time-frequency shift invariance. We prove that if a Gabor system on a lattice of rational density is a Riesz sequence generating a subspace which is invariant under an additional time-frequency shift, then its generator cannot decay fast simultaneously in time and frequency.

1. Introduction

The Balian-Low theorem is an uncertainty principle in time-frequency analysis which in its original form states that a generator of a Gabor orthonormal basis of the space of square integrable functions on the real line cannot be well-localized simultaneously in time and frequency.

Theorem 1.1 ([2, 12])**.** *If the functions $e^{2\pi inx}g(x-m)$, $(m,n)\in\mathbb{Z}\times\mathbb{Z}$, form an orthonormal basis of $L^2(\mathbb{R})$, then*

$$\Big(\int |x-\alpha|^2|g(x)|^2\,dx\Big)\cdot\Big(\int |\omega-\beta|^2|\widehat{g}(\omega)|^2\,d\omega\Big)=\infty,\quad \alpha,\beta\in\mathbb{R}. \tag{1.1}$$

The result generalizes from $\mathbb{Z}\times\mathbb{Z}$ to separable lattices of the form $a\mathbb{Z}\times b\mathbb{Z}$, where $ab=1$; the latter being in fact necessary for $e^{2\pi ibnx}g(x-am)$, $(am,bn)\in a\mathbb{Z}\times b\mathbb{Z}$, to form an orthonormal basis of $L^2(\mathbb{R})$. The results in this paper though achieve generalizations in the case $ab>1$ by involving an additional invariance by time-frequency shifts. Before discussing the Balian Low theorem, its extensions, and our results in more depth, we state our main result in its simplest form for illustration:

Theorem 1.2. *If $ab\geq 1$ is rational, the functions $e^{2\pi ibnx}g(x-am)$, $(am,bn)\in a\mathbb{Z}\times b\mathbb{Z}$, form an orthonormal system and its closed linear span contains $e^{2\pi i\eta x}g(x-u)$ for some $(u,\eta)\notin a\mathbb{Z}\times b\mathbb{Z}$, then*

$$\Big(\int |x-\alpha|^2|g(x)|^2\,dx\Big)\cdot\Big(\int |\omega-\beta|^2|\widehat{g}(\omega)|^2\,d\omega\Big)=\infty,\quad \alpha,\beta\in\mathbb{R}.$$

In the last two decades, the Balian-Low theorem has inspired significant research in time-frequency analysis and has itself been generalized in various ways (see, e.g., [1, 7, 10, 14]). Gautam [7] recognized that g having a finite uncertainty product (1.1)

2010 *Mathematics Subject Classification.* 42C15, 42C30, 30H35.

Key words and phrases. Balian-Low theorem, Amalgam Balian-Low theorem, Additional shift invariance, Gabor frames, Time-frequency analysis, VMO (vanishing mean oscillation) functions.

implies that its Zak transform Zg has locally vanishing mean oscillation and that the latter actually prevents the system $\{e^{2\pi inx}g(x-m) : m,n \in \mathbb{Z}\}$ to be a Riesz basis of $L^2(\mathbb{R})$. We will introduce the reader in Sections 2 and 3 to both the Zak transform and the concept of vanishing mean oscillation (VMO). In fact, Gautam proved the following theorem.

Theorem 1.3 ([7]). *If $g \in L^2(\mathbb{R})$ such that the Gabor system $\{e^{2\pi inx}g(x-m) : m,n \in \mathbb{Z}\}$ is a Riesz basis of $L^2(\mathbb{R})$, then $Zg \notin \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$. Moreover, if $Zg \notin \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$, then for any $p,q \in (1,\infty)$ with $\frac{1}{p}+\frac{1}{q}=1$ we have*

$$\Big(\int |x-\alpha|^q |g(x)|^2\,dx\Big)\cdot\Big(\int |\omega-\beta|^p |\widehat{g}(\omega)|^2\,d\omega\Big)=\infty, \quad \alpha,\beta\in\mathbb{R}. \tag{1.2}$$

In particular, if $\{e^{2\pi inx}g(x-m) : m,n \in \mathbb{Z}\}$ constitutes a Riesz basis of $L^2(\mathbb{R})$, then (1.2) holds for any p and q as above.

In this paper, we generalize Theorem 1.3 in two ways. First, as the attentive reader might have noticed, Theorem 1.3 is only proved and formulated for the most simple lattice $\mathbb{Z}\times\mathbb{Z}$. In our results we consider general rational lattices and lattices of rational density. Secondly, we work with Gabor systems that constitute a Riesz basis of their closed linear span instead of $L^2(\mathbb{R})$, as indicated in Theorem 1.2. Our first main result reads as follows.

Theorem 1.4. *Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density such that the Gabor system $\{e^{2\pi ibx}g(x-a) : (a,b)\in\Lambda\}$ is a Riesz basis of its closed linear span $\mathcal{G}(g,\Lambda)$. If $e^{2\pi i\eta x}g(x-u) \in \mathcal{G}(g,\Lambda)$ for some $(u,\eta)\notin\Lambda$, then (1.2) holds for all $p,q\in(1,\infty)$ with $\frac{1}{p}+\frac{1}{q}=1$.*

The conclusion of Theorem 1.4 can be strengthened significantly if we restrict ourselves to rational lattices, i.e., lattices that only consist of rational points. Recall that, given a field $\mathbb{F}$, by $\mathrm{GL}(n,\mathbb{F})$ one usually denotes the group of invertible matrices in $\mathbb{F}^{n\times n}$ and $\mathrm{SL}(n,\mathbb{F})$ stands for the subgroup of $\mathrm{GL}(n,\mathbb{F})$ consisting of the matrices with determinant 1.

Theorem 1.5. *Let $g\in L^2(\mathbb{R})$ and let $\Lambda = A\mathbb{Z}^2$ with $A\in\mathrm{GL}(2,\mathbb{Q})$, such that the Gabor system $\{e^{2\pi ibx}g(x-a) : (a,b)\in\Lambda\}$ is a Riesz basis of its closed linear span $\mathcal{G}(g,\Lambda)$. If $e^{2\pi i\eta x}g(x-u)\in\mathcal{G}(g,\Lambda)$ for some $(u,\eta)\notin\Lambda$, then $Zg\notin\mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.*

Note that neither of the above two theorems implies the other, since rational lattices are of rational density but the condition $Zg\notin\mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ is stronger than (1.2) as seen in Theorem 1.3 (cf. Problem 4.3).

In [8], Gebardo and Han already generalized the Balian-Low theorem to Gabor frames for subspaces of $L^2(\mathbb{R})$. One of their main results states that *if $ab>1$, $g\in L^2(\mathbb{R})$, and $\{e^{2\pi ibmx}g(x-an) : m,n\in\mathbb{Z}\}$ forms an overcomplete frame for its closed linear span, then (1.1) holds.* As it is mentioned in [8], the word "overcomplete" cannot be dropped in the statement, as is revealed by choosing the Gaussian for g for which $\{e^{2\pi ibmx}g(x-an) : m,n\in\mathbb{Z}\}$ always is a Riesz sequence. Our theorems therefore complement the result from [8] inasmuch as we replace the term "overcomplete frame for its closed linear span" by "Riesz basis of its closed linear span which has an additional time-frequency shift invariance". In particular, we obtain the following corollary.

Corollary 1.6. *Let $g(x) = e^{-x^2}$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density < 1. Then $e^{2\pi i\eta x} g(x-u) \notin \mathcal{G}(g,\Lambda)$ for all $(u,\eta) \in \mathbb{R}^2 \backslash \Lambda$.*

An important variant of the Balian-Low theorem is the so-called amalgam version, known as the Amalgam Balian-Low theorem, which replaces the condition (1.1) by $g \notin \mathcal{S}_0(\mathbb{R})$, where $\mathcal{S}_0(\mathbb{R})$ denotes the Feichtinger algebra, given by

$$\mathcal{S}_0(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}) : \int f(x)\, e^{-(x-t)^2}\ e^{2\pi i x\nu}\, dx\ \in L^1(t,\nu) \right\}.$$

Recently [4], the Amalgam Balian-Low theorem has been generalized to Gabor subspaces of $L^2(\mathbb{R})$ in a similar fashion as Theorem 1.2 generalizes the Balian-Low theorem. Specifically, the main theorem in [4] reads as Theorem 1.4 with "(1.2) holds for ..." replaced by "$g \notin \mathcal{S}_0(\mathbb{R})$". In fact, the question whether $g \notin \mathcal{S}_0(\mathbb{R})$ can be replaced by (1.1) was posed as an open problem in [4]. Hence, Theorem 1.4 gives a positive answer to this question and goes beyond.

As is well known, the techniques used in proving the Balian-Low theorem are much more involved than those used in the proof of the Amalgam Balian-Low theorem. Therefore, and as we want to point out, the problem of replacing $g \notin \mathcal{S}_0(\mathbb{R})$ by (1.1) or (1.2) is by far not a matter of a straight-forward procedure.

The Balian-Low theorem and its amalgam version are not equivalent. In fact, as pointed out in [3], none of these two classical theorems implies the other. Therefore, it seems desirable to find a space $V \subset L^2(\mathbb{R})$ which contains both $\mathcal{S}_0(\mathbb{R})$ and the set of functions with a finite uncertainty product as in (1.1) such that the functions in V fail to be generators of Gabor Riesz bases of $L^2(\mathbb{R})$. In fact, Theorem 1.3 provides such a space, namely the space of functions whose Zak transform is locally VMO. Hence, the following easy consequence of Theorem 1.5 is a unification of the two classical theorems for rational lattices of the critical density 1.

Theorem 1.7. *Let $g \in L^2(\mathbb{R})$ and let $\Lambda = A\mathbb{Z}^2$ with $A \in \mathrm{GL}(2,\mathbb{Q})$, $\det A = 1$, such that the Gabor system $\{e^{2\pi i ax} g(x-b) : (a,b) \in \Lambda\}$ is a Riesz basis of $L^2(\mathbb{R})$. Then $Zg \notin \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.*

The paper is organized as follows. In Section 2 we introduce the reader to the notions and notations used throughout the paper. Section 3 introduces and discusses the functions that are locally of vanishing mean oscillation (VMO). We prove several statements on invariance properties of $\mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$ which we make use of in the proofs of our main results Theorem 1.4 and Theorem 1.5 in Section 4, but which also seem to be new and are interesting in their own right.

2. Preliminaries

In this section, we collect basic notions and tools in time-frequency analysis that are necessary for formulating and proving our main results. Recall that a lattice in $\mathbb{R}^2$ is a set of the form $A\mathbb{Z}^2$ with some $A \in \mathrm{GL}(2,\mathbb{R})$ and its density is given by $|\det A|^{-1}$. We define the time-frequency shift operator by $(u,\eta) \in \mathbb{R}^2$ as

$$\pi(u,\eta) : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad \pi(u,\eta) f(x) = e^{2\pi i\eta x} f(x-u).$$

Using this notation, the Gabor system generated by $g \in L^2(\mathbb{R})$ and a lattice $\Lambda \subset \mathbb{R}^2$ is simply written as $(g, \Lambda) := \{\pi(u,\eta)g : (u,\eta) \in \Lambda\}$. By $\mathcal{G}(g,\Lambda)$ we denote its closed linear span in $L^2(\mathbb{R})$, i.e., $\mathcal{G}(g,\Lambda) = \overline{\text{span}}\,\{\pi(u,\eta)g : (u,\eta) \in \Lambda\}$. For the convenience of the reader, we state some easily verifiable properties of the time-frequency shift operator in the following lemma.

Lemma 2.1. *The following statements hold.*

(a) *For $a, b, c, d \in \mathbb{R}$, we have*

$$\pi(a,b)\pi(c,d) = e^{-2\pi i a d}\pi(a+c, b+d) = e^{-2\pi i(ad-bc)}\pi(c,d)\pi(a,b).$$

(b) *For fixed $f \in L^2(\mathbb{R})$, the mapping $(a,b) \mapsto \pi(a,b)f$ is continuous from $\mathbb{R}^2$ to $L^2(\mathbb{R})$.*

The *Fourier transform* is defined on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ by

$$\mathcal{F}f(\omega) = \widehat{f}(\omega) := \int_{\mathbb{R}} f(x)\, e^{-2\pi i x\omega}\, dx, \quad \omega \in \mathbb{R}.$$

It is well known that the operator $\mathcal{F}$ extends to a unitary operator from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. It can be used to define the Sobolev space $H^s(\mathbb{R})$, $s > 0$, as follows:

$$H^s(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} (1+|\omega|^2)^s |\widehat{f}(\omega)|^2\, d\omega < \infty \right\}.$$

The *Zak transform* of $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is defined (a.e.) by

$$Zf(x,\omega) = \sum_{k\in\mathbb{Z}} f(x+k)\, e^{-2\pi i k\omega}, \quad (x,\omega) \in \mathbb{R}^2.$$

As is easily seen, the function Zf is quasi-periodic, i.e., for $m, n \in \mathbb{Z}$ we have

$$Zf(x+m, \omega+n) = e^{2\pi i m\omega}\, Zf(x,\omega) \quad \text{for a.e. } (x,\omega) \in \mathbb{R}^2. \tag{2.1}$$

The mapping $f \mapsto Zf|_{[0,1]^2}$ extends continuously to a unitary map from $L^2(\mathbb{R})$ onto $L^2([0,1]^2)$. Here, if $f \in L^2(\mathbb{R})$, by Zf we mean the quasi-periodic extension of $Zf|_{[0,1]^2}$ to $\mathbb{R}^2$, which is an element of $L^2_{\text{loc}}(\mathbb{R}^2)$. We summarize some useful properties of the Zak transform in the following lemma.

Lemma 2.2. *Let $f \in L^2(\mathbb{R})$. Then the following relations hold for a.e. $(x,\omega) \in \mathbb{R}^2$.*

(a) $Zf(x+m, \omega+n) = e^{2\pi i m\omega}\, Zf(x,\omega)$ *for all* $m, n \in \mathbb{Z}$.
(b) $Z\pi(u,\eta)f(x,\omega) = e^{2\pi i \eta x}\, Zf(x-u, \omega-\eta)$ *for all* $(u,\eta) \in \mathbb{R}^2$.
(c) $(Z\pi(m,n)f)(x,\omega) = e^{2\pi i(nx - m\omega)}\, Zf(x,\omega)$ *for all* $m, n \in \mathbb{Z}$.
(d) $Z\widehat{f}(x,\omega) = e^{2\pi i x\omega}\, Zf(-\omega, x)$.
(e) $f(x) = \int_0^1 Zf(x,\omega)\, d\omega$.

The following technical lemma will be used to prove our main results. A similar statement can be found in [8, Proposition 3.8] (see also [4, Lemma 5]). However, since the present setting is slightly different as in [8] and [4], we give a full proof of the statement.

Lemma 2.3. *Let $P, Q \in \mathbb{N}$, $g \in L^2(\mathbb{R})$, and assume that the Gabor system $(g, \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z})$ is a Riesz sequence in $L^2(\mathbb{R})$ with Riesz bounds A and B. Then the matrix function*

$$\mathbf{A}(x,\omega) := \Big(Zg(x - \tfrac{k}{P} - \tfrac{\ell}{Q}, \omega)\Big)_{k,\ell=0}^{P-1,Q-1} \in \mathbb{C}^{P\times Q}, \qquad (x,\omega) \in \mathbb{R}^2.$$

is essentially bounded from above and from below. More precisely, for a.e. $(x,\omega) \in \mathbb{R}^2$ we have that

$$PA\|\xi\|^2 \le \|\mathbf{A}(x,\omega)\xi\|^2 \le PB\|\xi\|^2 \qquad \text{for all } \xi \in \mathbb{C}^Q. \tag{2.2}$$

In particular, $Zg \in L^\infty(\mathbb{R}^2)$, which also implies $g \in L^\infty(\mathbb{R})$.

Proof. Let $F \in L^\infty(R_P, \mathbb{C}^Q)$ be arbitrary, where $R_P := (0, \frac{1}{P}) \times (0,1)$. Then there exists $(c_{m,n})_{m,n\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$ such that $F_\ell = \sum_{s,n\in\mathbb{Z}} c_{sQ+\ell,n} e^{2\pi i(nPx - s\omega)}$, $\ell = 0, \dots, Q-1$, where F_ℓ denotes the ℓ-th coordinate of F. Using the properties of the Zak transform, we have (extending F to $\mathbb{R}^2$ periodically)

$$\begin{aligned}
\Big\| \sum_{m,n\in\mathbb{Z}} c_{m,n} \pi\left(\tfrac{m}{Q}, nP\right) g \Big\|_{L^2(\mathbb{R})}^2 &= \Big\| \sum_{\ell=0}^{Q-1} \sum_{s,n\in\mathbb{Z}} c_{sQ+\ell,n} e^{2\pi i(nPx - s\omega)} (Zg)\left(x - \tfrac{\ell}{Q}, \omega\right) \Big\|_{L^2([0,1]^2)}^2 \\
&= \Big\| \sum_{\ell=0}^{Q-1} F_\ell(x,\omega) (Zg)\left(x - \tfrac{\ell}{Q}, \omega\right) \Big\|_{L^2([0,1]^2)}^2 \\
&= \int_0^{\frac{1}{P}} \int_0^1 \sum_{k=0}^{P-1} \Big| \sum_{\ell=0}^{Q-1} (Zg)\left(x - \tfrac{k}{P} - \tfrac{\ell}{Q}, \omega\right) F_\ell(x,\omega) \Big|^2 \, d\omega \, dx \\
&= \int_0^{\frac{1}{P}} \int_0^1 \|\mathbf{A}(x,\omega) F(x,\omega)\|_2^2 \, d\omega \, dx.
\end{aligned}$$

Hence, for every $F \in L^\infty(R_P, \mathbb{C}^Q)$ we obtain

$$PA\|F\|_{L^2(R_P,\mathbb{C}^Q)}^2 \le \int_0^{\frac{1}{P}} \int_0^1 \|\mathbf{A}(x,\omega)F(x,\omega)\|_2^2 \, d\omega \, dx \le PB\|F\|_{L^2(R_P,\mathbb{C}^Q)}^2.$$

Let $\mathcal{D}$ be a countable dense set in $\mathbb{C}^Q$ (e.g., $\mathcal{D} = (\mathbb{Q} + i\mathbb{Q})^Q$). For $\xi \in \mathcal{D}$ let $\mathcal{L}(\xi)$ denote the set consisting of all Lebesgue points in R_P of the map $(x,\omega) \mapsto \|\mathbf{A}(x,\omega)\xi\|_2^2$ and put $\mathcal{L} := \bigcap_{\xi\in\mathcal{D}} \mathcal{L}(\xi)$. Then $R_P \setminus \mathcal{L}$ has zero measure. For $(x_0,\omega_0) \in \mathcal{L}$, $\xi \in \mathcal{D}$, and $\varepsilon > 0$ define $F = F_{\xi,\varepsilon,x_0,\omega_0} := \frac{1}{\sqrt{\pi}\varepsilon} \chi_{B_\varepsilon(x_0,\omega_0)} \xi$, where $B_\varepsilon(x_0,\omega_0)$ denotes the euclidian ball with center (x_0,ω_0) and radius ε. Then, for ε small enough, $\|F\|_{L^2(R_P,\mathbb{C}^Q)} = \|\xi\|$ and

$$\int_0^{\frac{1}{P}} \int_0^1 \|\mathbf{A}(x,\omega)F(x,\omega)\|_2^2 \, d\omega \, dx = \tfrac{1}{\pi\varepsilon^2} \int_{B_\varepsilon(x_0,\omega_0)} \|\mathbf{A}(x,\omega)\xi\|_2^2 \, d(x,\omega).$$

Letting $\varepsilon \to 0$ yields

$$PA\|\xi\|^2 \le \|\mathbf{A}(x_0,\omega_0)\xi\|_2^2 \le PB\|\xi\|^2.$$

By a density argument this holds for all $\xi \in \mathbb{C}^Q$, which establishes (2.2) for all $(x,\omega) \in \mathcal{L}$ and thus (due to the quasi-periodicity of $\mathbf{A}$) for a.e. $(x,\omega) \in \mathbb{R}^2$. Now, choosing the first

standard basis vector of $\mathbb{C}^Q$ for ξ, we obtain that $|Zg(x,\omega)|^2 \leq PB$ for a.e. $(x,\omega) \in \mathbb{R}^2$. Hence, $Zg \in L^\infty(\mathbb{R}^2)$. Also, Lemma 2.2(e) yields $g \in L^\infty(\mathbb{R})$. □

3. Functions of Vanishing Mean Oscillation (VMO)

A cube in $\mathbb{R}^n$ of side length $\delta > 0$ is a set of the form $I_1 \times \cdots \times I_n$ where each $I_i \subset \mathbb{R}$ is a closed interval of length δ. For a function $F \in L^1_{\rm loc}(\mathbb{R}^n)$ and a bounded measurable set $\Delta \subset \mathbb{R}^n$ with $|\Delta| > 0$, we define

$$F_\Delta := \frac{1}{|\Delta|}\int_\Delta F\,dx \qquad \text{and} \qquad M_\Delta(F) := (|F - F_\Delta|)_\Delta.$$

Also, for a bounded open set $U \subset \mathbb{R}^n$ and $\varepsilon > 0$, let

$$S_{\varepsilon,U}(F) := \sup\{M_Q(F) : Q \subset U \text{ cube, } |Q| < \varepsilon\}.$$

Definition 3.1. Let U be a bounded open subset of $\mathbb{R}^n$.

(a) A function $F \in L^1_{\rm loc}(\mathbb{R}^n)$ is said to be of *bounded mean oscillation* (BMO) on U if $\sup_Q M_Q(F) < \infty$, where the supremum is taken over all bounded cubes Q contained in U. The space of all such functions is denoted by $\mathrm{BMO}(U)$. We write $F \in \mathrm{BMO}_{\rm loc}(\mathbb{R}^n)$ if $F \in \mathrm{BMO}(U)$ for every bounded open set $U \subset \mathbb{R}^n$.

(b) A function $F \in L^1_{\rm loc}(\mathbb{R}^n)$ is said to be of *vanishing mean oscillation* (VMO) on U if $F \in \mathrm{BMO}(U)$ and $\lim_{\varepsilon\to 0} S_{\varepsilon,U}(F) = 0$. The space of all such functions is denoted by $\mathrm{VMO}(U)$. Likewise, we write $F \in \mathrm{VMO}_{\rm loc}(\mathbb{R}^n)$ if $F \in \mathrm{VMO}(U)$ for every bounded open set $U \subset \mathbb{R}^n$.

Remark 3.2. (a) It is easily seen that for any $F, G \in L^1_{\rm loc}(\mathbb{R}^n)$ and a cube $Q \subset U$, $(F+G)_Q = F_Q + G_Q$ and $M_Q(F+G) \leq M_Q(F) + M_Q(G)$, which leads to $S_{\varepsilon,U}(F+G) \leq S_{\varepsilon,U}(F) + S_{\varepsilon,U}(G)$. This shows that the sets $\mathrm{BMO}(U)$ and $\mathrm{VMO}(U)$ are linear spaces. Also, $\|\cdot\|_{\mathrm{BMO}(U)} := \sup_{Q\subset U} M_Q(F)$ induces a semi-norm on $\mathrm{BMO}(U)$.

(b) It is straightforward that $L^\infty(\mathbb{R}^n) \subset \mathrm{BMO}(\mathbb{R}^n)$. Also, every bounded uniformly continuous function on $\mathbb{R}^n$ belongs to $\mathrm{VMO}(\mathbb{R}^n)$ [15].

In the sequel, we will use the notation

$$\mathrm{VMO}^\infty_{\rm loc}(\mathbb{R}^n) := \mathrm{VMO}_{\rm loc}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n).$$

The following lemma shows in particular that $\mathrm{VMO}^\infty_{\rm loc}(\mathbb{R}^n)$ is closed under multiplication and is therefore an algebra.

Lemma 3.3. *The following statements hold.*

(i) *If $F, G \in L^1_{\rm loc}(\mathbb{R}^n)$, then for any cube $Q \subset \mathbb{R}^n$ we have*

$$|F_Q G_Q - (FG)_Q| \leq \tfrac{1}{2}\max\{\|F\|_\infty, \|G\|_\infty\}(M_Q(F) + M_Q(G)). \tag{3.1}$$

Also, whenever $U \subset \mathbb{R}^n$ is a bounded open set and $\varepsilon > 0$, then

$$S_{\varepsilon,U}(FG) \;\leq\; \tfrac{3}{2}\max\{\|F\|_\infty, \|G\|_\infty\}\big(S_{\varepsilon,U}(F) + S_{\varepsilon,U}(G)\big). \tag{3.2}$$

Consequently, $F, G \in \mathrm{VMO}^\infty_{\rm loc}(\mathbb{R}^n)$ implies $FG \in \mathrm{VMO}^\infty_{\rm loc}(\mathbb{R}^n)$.

(ii) *If* $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$ *and* $\operatorname{ess\,inf}|F| > 0$, *then for any bounded open set* $U \subset \mathbb{R}^n$ *there exists* $\varepsilon = \varepsilon_U > 0$ *such that*

$$|F_Q| \geq (\operatorname{ess\,inf}|F|)/2 \quad \textit{for all cubes } Q \subset U \textit{ with } |Q| < \varepsilon, \tag{3.3}$$

and

$$S_{\varepsilon,U}(1/F) \;\leq\; \frac{4}{(\operatorname{ess\,inf}|F|)^2}\, S_{\varepsilon,U}(F). \tag{3.4}$$

Consequently, $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$ *and* $\operatorname{ess\,inf}|F| > 0$ *imply* $1/F \in \mathrm{VMO}^{\infty}_{\mathrm{loc}}(\mathbb{R}^n)$.

Proof. (i) Let $F, G \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let $Q \subset U$ be a cube. Then

$$\frac{1}{|Q|}\int_Q |FG - (FG)_Q|\,dx \leq \frac{1}{|Q|}\int_Q |FG - F_Q G_Q|\,dx + |F_Q G_Q - (FG)_Q|.$$

We estimate the first term on the right hand side as

$$\begin{aligned}\frac{1}{|Q|}\int_Q |FG - F_Q G_Q|\,dx &\leq \frac{1}{|Q|}\int_Q \left(|F||G - G_Q| + |F - F_Q||G_Q|\right)\,dx \\ &\leq \max\{\|F\|_\infty, \|G\|_\infty\}(M_Q(F) + M_Q(G)).\end{aligned}$$

For the second term, we observe that

$$|F_Q G_Q - (FG)_Q| = \frac{1}{|Q|}\left|\int_Q F\,(G_Q - G)\,dx\right| \leq \frac{1}{|Q|}\int_Q |F|\,|G - G_Q|\,dx \;\leq\; \|F\|_\infty M_Q(G)$$

and

$$|F_Q G_Q - (FG)_Q| = \frac{1}{|Q|}\left|\int_Q (F_Q - F)\,G\,dx\right| \leq \frac{1}{|Q|}\int_Q |F - F_Q|\,|G|\,dx \;\leq\; \|G\|_\infty M_Q(F)$$

so that

$$|F_Q G_Q - (FG)_Q| \leq \tfrac{1}{2}\max\{\|F\|_\infty, \|G\|_\infty\}(M_Q(F) + M_Q(G)).$$

Therefore, $M_Q(FG) \leq \frac{3}{2}\max\{\|F\|_\infty, \|G\|_\infty\}(M_Q(F)+M_Q(G))$ from which (3.2) follows.

(ii) Assume that $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$ and $C := \operatorname{ess\,inf}|F| > 0$, and let $U \subset \mathbb{R}^n$ be an open set. Since $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$, we have $S_{\varepsilon,U}(F) \leq C/2$ for some $\varepsilon = \varepsilon_U > 0$. Let $Q \subset U$ be any cube with $|Q| < \varepsilon$. Then $|F(x) - F_Q| + |F_Q| \geq |F(x)| \geq C$ a.e. so that $M_Q(F) + |F_Q| \geq C$. Using the fact that $M_Q(F) \leq S_{\varepsilon,U}(F) \leq C/2$, we obtain $|F_Q| \geq C/2$. Now, observe that

$$M_Q(1/F) = \frac{1}{|Q|}\int_Q \left|\frac{1}{F} - \left(\frac{1}{F}\right)_Q\right| dx \;\leq\; \frac{1}{|Q|}\int_Q \left|\frac{1}{F} - \frac{1}{F_Q}\right| dx + \left|\frac{1}{F_Q} - \left(\frac{1}{F}\right)_Q\right|.$$

The first term can be estimated by

$$\frac{1}{|Q|}\int_Q \left|\frac{1}{F} - \frac{1}{F_Q}\right| dx = \frac{1}{|Q|}\int_Q \frac{|F_Q - F|}{|F_Q F|}\,dx \;\leq\; \frac{2M_Q(F)}{C^2}$$

and the second term by

$$\left|\frac{1}{F_Q} - \left(\frac{1}{F}\right)_Q\right| \leq \frac{2}{C}\left|1 - F_Q\left(\frac{1}{F}\right)_Q\right| = \frac{2}{C|Q|}\left|\int_Q \left(1 - \frac{F_Q}{F}\right) dx\right|$$

$$\leq \frac{2}{C^2|Q|}\int_Q |F - F_Q|\, dx = \frac{2M_Q(F)}{C^2}.$$

Thus, we have $M_Q(1/F) \leq 4M_Q(F)/C^2$, which yields (3.4) . □

A successive application of (3.1) in Lemma 3.3(i) gives the following corollary.

Corollary 3.4. *If $F, F_1, \ldots, F_n \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$, then $\prod_{i=1}^n F_i \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$. Moreover, there exists a constant $C > 0$ which only depends on $\|F_1\|_\infty, \ldots, \|F_n\|_\infty$ such that for any bounded open set $U \subset \mathbb{R}^n$, $\varepsilon > 0$ and a cube $Q \subset U$ with $|Q| < \varepsilon$,*

$$\left| \left(\prod_{i=1}^n F_i \right)_Q - \prod_{i=1}^n (F_i)_Q \right| \leq C \sum_{i=1}^n S_{\varepsilon,U}(F_i).$$

Corollary 3.5. *Let $\mathbf{B}(x) = [B_{j,k}(x)]_{j,k=1}^N$, $N \in \mathbb{N}$, be such that each $B_{j,k} : \mathbb{R}^n \to \mathbb{C}$ belongs to $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ and $\overline{B_{k,j}}(x) = B_{j,k}(x)$ for all j, k, i.e., $\mathbf{B}(x) = \mathbf{B}(x)^*$. If there exist constants $\alpha, \beta > 0$ such that $\alpha I_N \leq \mathbf{B}(x) \leq \beta I_N$ a.e., then each entry of $\mathbf{B}(x)^{-1}$ belongs to $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$.*

Proof. Let $\mathbf{C}(x) = \operatorname{adj} \mathbf{B}(x)$ be the adjugate matrix of $\mathbf{B}(x)$. By Lemma 3.3(i), each entry of $\mathbf{C}(x)$ is in $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ and so is $\det \mathbf{B}(x)$. Since $\det \mathbf{B}(x) \geq \alpha^N$ a.e., it follows from Lemma 3.3(ii) that $(\det \mathbf{B}(x))^{-1} \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$. Again by Lemma 3.3(i), we conclude that each entry of $\mathbf{B}(x)^{-1} = (\det \mathbf{B}(x))^{-1}\mathbf{C}(x)$ belongs to $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^n)$. □

The next proposition will play a key role in the proofs of our main theorems. It was proved for a continuous function H in [4, Proposition 3]. Here, we relax the condition to $H \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$ which is much weaker than H being continuous.

Proposition 3.6. *Let $P_1, P_2, N \in \mathbb{N}$, $M_1, M_2 \in \mathbb{Z}$, and $(u, \eta) \in \mathbb{Q}^2$, $(u, \eta) \neq (0, 0)$, such that $Nu, N\eta \in \mathbb{Z}$. If $H \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$ is $\frac{1}{P_1}$-periodic in x, $\frac{1}{P_2}$-periodic in ω and*

$$\prod_{n=0}^{N-1} H(x + nu, \omega + n\eta) = e^{2\pi i(M_1 x + M_2 \omega)} \quad \text{for a.e. } (x, \omega) \in \mathbb{R}^2, \tag{3.5}$$

then NP_1 divides M_1 and NP_2 divides M_2.

Proof. First, we note that $H \in \mathrm{VMO}(\mathbb{R}^2)$ since H is periodic. For $r > 0$ and $F \in L^\infty(\mathbb{R}^2)$, we define the mean function

$$F_{[r]}(x, \omega) := \frac{1}{|Q_r(x, \omega)|}\int_{Q_r(x,\omega)} F(z)\, dz, \qquad (x, \omega) \in \mathbb{R}^2,$$

which takes the average of F over the cube $Q_r(x, \omega)$ of side length r centered at (x, ω),

$$Q_r(x, \omega) = [x - \tfrac{r}{2}, x + \tfrac{r}{2}] \times [\omega - \tfrac{r}{2}, \omega + \tfrac{r}{2}].$$

It is easily seen that $F_{[r]}$ is continuous (even Lipschitz continuous); moreover, if F is periodic, then $F_{[r]}$ inherits the periodicity of F. Setting $H_n(x, \omega) := H(x + nu, \omega + n\eta)$ for $n = 0, \ldots, N - 1$, Corollary 3.4 implies that

$$\left| \prod_{n=0}^{N-1} (H_n)_{[r]} - \left(\prod_{n=0}^{N-1} H_n \right)_{[r]} \right| \leq C \sum_{n=0}^{N-1} S_{r^2,\mathbb{R}^2}(H_n) = CNS_{r^2,\mathbb{R}^2}(H),$$

where $C > 0$ is a constant which depends only on $\|H\|_\infty$. Using

$$\left(\prod_{n=0}^{N-1} H_n\right)_{[r]}(x,\omega) = \left(e^{2\pi i(M_1 x + M_2 \omega)}\right)_{[r]} = \operatorname{sinc}(M_1 r)\operatorname{sinc}(M_2 r)\, e^{2\pi i(M_1 x + M_2 \omega)},$$

we obtain

$$\left|\prod_{n=0}^{N-1} (H_n)_{[r]}(x,\omega) - e^{2\pi i(M_1 x + M_2 \omega)}\right| \leq CNS_{r^2,\mathbb{R}^2}(H) + |\operatorname{sinc}(M_1 r)\operatorname{sinc}(M_2 r) - 1|$$

for all $(x,\omega) \in \mathbb{R}^2$. Note that the right hand side does not depend on (x,ω) and tends to zero as $r \to 0$. Since $\prod_{n=0}^{N-1}(H_n)_{[r]}(x,\omega)$ and $e^{2\pi i(M_1 x + M_2\omega)}$ are continuous functions in (x,ω), there exist continuous functions $\rho_r : \mathbb{R}^2 \to \mathbb{C}$, $r > 0$, such that

$$\prod_{n=0}^{N-1} (H_n)_{[r]}(x,\omega) = \rho_r(x,\omega)\, e^{2\pi i(M_1 x + M_2 \omega)}.$$

It is easily seen that $\rho_r(x,\omega)$ converges uniformly to 1 on $\mathbb{R}^2$ as $r \to 0$. Noting that $(H_n)_{[r]}(x,\omega) = H_{[r]}(x + nu, \omega + n\eta)$ for $n = 0, \ldots, N-1$, the equation above can be written as

$$\prod_{n=0}^{N-1} H_{[r]}(x + nu, \omega + n\eta) = \rho_r(x,\omega)\, e^{2\pi i(M_1 x + M_2 \omega)}. \tag{3.6}$$

Here, the mean function $H_{[r]}$ inherits the periodicity of H, and is therefore $\frac{1}{P_1}$-periodic in x and $\frac{1}{P_2}$-periodic in ω. Note that the periodicity of H together with (3.5) yields $M_1/P_1, M_2/P_2 \in \mathbb{Z}$. This shows that $e^{2\pi i(M_1 x + M_2\omega)}$ is $\frac{1}{P_1}$-periodic in x and $\frac{1}{P_2}$-periodic in ω, and therefore by (3.6), so is $\rho_r(x,\omega)$. On the other hand, by replacing x and ω respectively with $x + u$ and $\omega + \eta$ in (3.5), taking into account $Nu, N\eta \in \mathbb{Z}$, and using the periodicity of H, we find that $M_1 u + M_2 \eta \in \mathbb{Z}$. Applying the same trick to (3.6) then gives

$$\rho_r(x + u, \omega + \eta) = \rho_r(x,\omega)$$

for all $r > 0$ and $(x,\omega) \in \mathbb{R}^2$. As $\rho_r \to 1$ uniformly, there exists a branch of $\sqrt[N]{\cdot}$ such that $\sqrt[N]{\rho_r}$ is continuous for r small enough, say, $r \leq r_0$, $r_0 > 0$. Now, setting $G_r(x,\omega) := H_{[r]}(x,\omega)/\sqrt[N]{\rho_r(x,\omega)}$ for $r \leq r_0$ and combining all these facts yields

$$\prod_{n=0}^{N-1} G_r(x + nu, \omega + n\eta) = \prod_{n=0}^{N-1} \frac{H_{[r]}(x + nu, \omega + n\eta)}{\sqrt[N]{\rho_r(x + nu, \omega + n\eta)}} = e^{2\pi i(M_1 x + M_2 \omega)}.$$

Note that G_r is continuous and $\frac{1}{P_1}$-periodic in x, $\frac{1}{P_2}$-periodic in ω. The fact that NP_1 divides M_1 and NP_2 divides M_2 now follows from [4, Proposition 3]. □

In the remainder of this section, we consider functions in $\mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$ that are not necessarily in $L^\infty(\mathbb{R}^n)$.

Lemma 3.7. *Let $\Delta_1, \Delta_2 \subset \mathbb{R}^n$ be bounded measurable sets with $\Delta_1 \subset \Delta_2$ and $|\Delta_1| > 0$. Then for any $F \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ we have*

$$M_{\Delta_1}(F) \,\leq\, 2\frac{|\Delta_2|}{|\Delta_1|}\, M_{\Delta_2}(F).$$

Proof. Note that for any $F \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$,

$$\begin{aligned} M_{\Delta_2}(F) &= \frac{1}{|\Delta_2|}\int_{\Delta_2} |F - F_{\Delta_2}|\, dx \,\geq\, \frac{1}{|\Delta_2|}\int_{\Delta_1} |(F - F_{\Delta_1}) - (F_{\Delta_2} - F_{\Delta_1})|\, dx \\ &\geq \frac{|\Delta_1|}{|\Delta_2|}\big(M_{\Delta_1}(F) - |F_{\Delta_2} - F_{\Delta_1}|\big), \end{aligned}$$

which is equivalent to

$$M_{\Delta_1}(F) \leq \frac{|\Delta_2|}{|\Delta_1|} M_{\Delta_2}(F) + |F_{\Delta_2} - F_{\Delta_1}|.$$

Estimating the last term by

$$|F_{\Delta_2} - F_{\Delta_1}| = \left|\frac{1}{|\Delta_1|}\int_{\Delta_1} (F - F_{\Delta_2})\, dx\right| \,\leq\, \frac{1}{|\Delta_1|}\int_{\Delta_1} |F - F_{\Delta_2}|\, dx \leq \frac{|\Delta_2|}{|\Delta_1|} M_{\Delta_2}(F),$$

we obtain the desired inequality. □

Lemma 3.8. *Let $A \in \mathrm{GL}(n, \mathbb{R})$ and $b \in \mathbb{R}^n$ and define an affine mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ by $\Phi(x) = Ax + b$. Then for any $F \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and any cube $Q \subset \mathbb{R}^n$ with center $c \in \mathbb{R}^n$ and side length $\delta > 0$ we have*

$$M_Q(F \circ \Phi) \,\leq\, \frac{2n^{n/2}\|A\|_{\mathrm{op}}^n}{|\det A|}\, M_{\widetilde{Q}}(F),$$

where $\widetilde{Q}$ is the cube with center $\Phi(c)$ and side length $\sqrt{n}\|A\|_{\mathrm{op}}\delta$. Consequently, if $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$, then $F \circ \Phi \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$.

Proof. Note that the set $\Phi(Q)$ is a parallelepiped in $\mathbb{R}^n$ with volume $|AQ| = |\det A||Q|$. For any $G \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we have

$$(G \circ \Phi)_Q = \frac{1}{|\det A||Q|}\int_{\Phi^{-1}(\Phi(Q))} G(\Phi(x))|\det A|\, dx = \frac{1}{|\Phi(Q)|}\int_{\Phi(Q)} G\, dx = G_{\Phi(Q)}$$

so that

$$M_Q(F\circ\Phi) = (|F\circ\Phi - (F\circ\Phi)_Q|)_Q = (|F - F_{\Phi(Q)}|\circ\Phi)_Q = (|F - F_{\Phi(Q)}|)_{\Phi(Q)} = M_{\Phi(Q)}(F).$$

It is easy to see that the cube $\widetilde{Q}$ contains $\Phi(Q)$. Hence, Lemma 3.7 implies that

$$M_Q(F \circ A) \,\leq\, 2\frac{|\widetilde{Q}|}{|\Phi(Q)|} M_{\widetilde{Q}}(F) = \frac{2n^{n/2}\|A\|_{\mathrm{op}}^n}{|\det A|}\, M_{\widetilde{Q}}(F).$$

This proves the lemma. □

Proposition 3.9. *For $F \in L^n_{\mathrm{loc}}(\mathbb{R}^n)$, $\phi \in C^1(\mathbb{R}^n)$ and a cube $Q \subset \mathbb{R}^n$ of side length $\delta > 0$, we have*

$$M_Q(\phi F) \le \|\phi\|_{L^\infty(Q)} M_Q(F) + \|\phi'\|_Q \|F\|_{L^n(Q)},$$

where $\|\phi'\|_Q = \sup_{x\in Q} \|\nabla\phi(x)\|_{\ell_1}$. Consequently, if $F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n) \cap L^n_{\mathrm{loc}}(\mathbb{R}^n)$ and $\phi \in C^1(\mathbb{R}^n)$, then $\phi F \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^n)$.

Proof. Let $F \in L^n_{\mathrm{loc}}(\mathbb{R}^n)$, $\phi \in C^1(\mathbb{R}^n)$ and $Q \subset \mathbb{R}^n$ a cube of side length $\delta > 0$. Then for any $x, y \in Q$,

$$|\phi(x) - \phi(y)| \le \|\phi'\|_Q \cdot \|x - y\|_\infty \le \delta \|\phi'\|_Q$$

so that

$$\begin{aligned} M_Q(\phi F) &\le \frac{1}{|Q|}\int_Q |\phi(F - F_Q)|\,dx + \frac{1}{|Q|}\int_Q |\phi F_Q - (\phi F)_Q|\,dx \\ &\le \|\phi\|_{L^\infty(Q)} M_Q(F) + \frac{1}{|Q|}\int_Q \left|\frac{1}{|Q|}\int_Q [\phi(x) - \phi(y)]\,F(y)\,dy\right| dx \\ &\le \|\phi\|_{L^\infty(Q)} M_Q(F) + \frac{\delta\|\phi'\|_Q}{|Q|}\int_Q |F(y)|\,dy \\ &\le \|\phi\|_{L^\infty(Q)} M_Q(F) + \frac{\delta\|\phi'\|_Q}{|Q|}\|F\|_{L^n(Q)}|Q|^{1-1/n}, \end{aligned}$$

which gives the desired estimate. □

As for any $f \in L^2(\mathbb{R})$ the Zak transform Zf is locally square-integrable, we deduce the following corollary.

Corollary 3.10. *If $f \in L^2(\mathbb{R})$ satisfies $Zf \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$, then for any $\phi \in C^1(\mathbb{R}^2)$ we have that $\phi Zf \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.*

Proposition 3.11. *For $\alpha, \beta \in \mathbb{R}\backslash\{0\}$, define the operators*

$$D_\alpha : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad D_\alpha f(x) = \sqrt{|\alpha|} f(\alpha x) \quad \text{(dilation by } \alpha\text{)}$$
$$C_\beta : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad C_\beta f(x) = e^{2\pi i\beta x^2} f(x) \quad \text{(multiplication by a chirp } e^{2\pi i\beta x^2}\text{)}.$$

If $g \in L^2(\mathbb{R})$ satisfies $Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$, then $Z\hat{g}$, $Z(D_\alpha g)$, $Z(C_\beta g) \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ for all $\alpha, \beta \in \mathbb{Q}\backslash\{0\}$.

Proof. By Lemma 2.2, we have

$$Z\hat{g}(x,\omega) = e^{2\pi i x\omega} Zg(-\omega, x) = e^{2\pi i x\omega} Zg\left(\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)(x,\omega)^T\right).$$

Then Lemma 3.8 and Corollary 3.10 immediately yield that $Z\hat{g} \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$. Next, let us write $\alpha = p/q \in \mathbb{Q}\backslash\{0\}$, where $p, q \in \mathbb{Z} \setminus \{0\}$ are coprime, and $A = \mathrm{diag}(\alpha, \alpha^{-1})$. It is known [11] that

$$Z(D_\alpha g)(x,\omega) = \frac{1}{\sqrt{pq}} \sum_{\ell=0}^{q-1}\sum_{r=0}^{p-1} e^{2\pi i\ell\omega} Zg\left(A(x,\omega)^T + (\alpha\ell, r/p)^T\right) \quad \text{if } \alpha > 0$$

and

$$Z(D_\alpha g)(x,\omega) = \frac{1}{\sqrt{-pq}} \sum_{\ell=0}^{q-1} \sum_{r=0}^{p-1} e^{-2\pi i \ell \omega} Zg\left(A(x,\omega)^T - (\alpha\ell, r/p)^T\right) \quad \text{if } \alpha < 0.$$

Likewise, Lemma 3.8 and Corollary 3.10 imply that $Z(D_\alpha g) \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$. Finally, observe that

$$C_\beta = D_{\gamma^{-1}} C_{\beta\gamma^2} D_\gamma \tag{3.7}$$

for any $\beta, \gamma \in \mathbb{R}\backslash\{0\}$. If $\beta \in \mathbb{Q}\backslash\{0\}$, then there exists $\gamma \in \mathbb{N}$ such that $\beta\gamma^2 \in \mathbb{Z}$. Therefore, it suffices to show that $Z(C_m g) \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ for $m \in \mathbb{Z}$. For this, note that

$$\begin{aligned} Z(C_m g)(x,\omega) &= \sum_{k\in\mathbb{Z}} e^{2\pi i m(x+k)^2} g(x+k) e^{-2\pi i k\omega} = e^{2\pi i m x^2} \sum_{k\in\mathbb{Z}} g(x+k) e^{-2\pi i k(\omega - 2mx)} \\ &= e^{2\pi i m x^2} Zg(x, \omega - 2mx) = e^{2\pi i m x^2} Zg\left(\left(\begin{smallmatrix} 1 & 0 \\ -2m & 1 \end{smallmatrix}\right)(x,\omega)^T\right). \end{aligned}$$

Again, the claim now follows from Lemma 3.8 and Corollary 3.10. □

Remark 3.12. In [7], it is claimed, referring to the quasi-periodicity of Zak transform, that $Zf \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$ implies $Zf \in \mathrm{VMO}(\mathbb{R}^2)$. However, this is not true in general. For example, the function $f(x) = \mathbf{1}_{[0,1)}(x)\sin(\pi x)$ satisfies $Zf \in \mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)\backslash\,\mathrm{VMO}(\mathbb{R}^2)$. To see this, given any $\delta \in (0,1)$ let $k \in \mathbb{N}$ be such that $|\operatorname{sinc}(k\delta)| \le \frac{1}{2}$ and

$$Q = [k + \tfrac{1-\delta}{2}, k + \tfrac{1+\delta}{2}] \times [-\tfrac{\delta}{2}, \tfrac{\delta}{2}].$$

Then

$$(Zf)_Q = \left(\frac{1}{\delta}\int_{k+(1-\delta)/2}^{k+(1+\delta)/2} \sin(\pi(x-k))\,dx\right)\left(\frac{1}{\delta}\int_{-\delta/2}^{\delta/2} e^{2\pi i k\omega}\,d\omega\right) = \operatorname{sinc}(\delta/2)\operatorname{sinc}(k\delta)$$

and thus

$$\begin{aligned} M_Q(Zf) &= \frac{1}{\delta^2}\int_Q \left|\sin(\pi(x-k))e^{2\pi i k\omega} - \operatorname{sinc}(\delta/2)\operatorname{sinc}(k\delta)\right| d(x,\omega) \\ &\ge \frac{1}{\delta^2}\int_Q \left[\sin(\pi(x-k)) - \operatorname{sinc}(\delta/2)|\operatorname{sinc}(k\delta)|\right] d(x,\omega) \\ &= \operatorname{sinc}(\delta/2)(1 - |\operatorname{sinc}(k\delta)|) \ge \operatorname{sinc}(1/2)(1 - |\operatorname{sinc}(k\delta)|) \ge \frac{1}{\pi}, \end{aligned}$$

which shows that $Zf \notin \mathrm{VMO}(\mathbb{R}^2)$.

4. Proof of Theorems 1.4 and 1.5

The core of the proof of both theorems is the following proposition, which is simply Theorem 1.5 for lattices of the form $\frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$. In the proofs of the two theorems below we shall extend Proposition 4.1 to more general lattices by means of so-called metaplectic operators.

Proposition 4.1. *Let $g \in L^2(\mathbb{R})$ and let $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$ with $P, Q \in \mathbb{N}$ coprime, such that the Gabor system $\{e^{2\pi i b x} g(x-a) : (a,b) \in \Lambda\}$ is a Riesz basis of its closed linear span $\mathcal{G}(g,\Lambda)$. If $e^{2\pi i \eta x} g(x-u) \in \mathcal{G}(g,\Lambda)$ for some $(u,\eta) \in \mathbb{R}^2 \backslash \Lambda$, then $Zg \notin \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.*

Before we prove Proposition 4.1, we state a lemma that allows us to replace $(u,\eta) \in \mathbb{R}^2\backslash\Lambda$ with a rational pair $(u,\eta) \in \mathbb{Q}^2\backslash\Lambda$.

Lemma 4.2. *For $g \in L^2(\mathbb{R})$ and $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$ with $P, Q \in \mathbb{N}$, define the set*

$$M = \big\{(\alpha,\beta) \in \mathbb{R}^2 : \pi(\alpha,\beta)g \in \mathcal{G}(g,\Lambda)\big\} \quad (\supset \Lambda).$$

If $M\backslash\Lambda \neq \varnothing$, then there exists $(u,\eta) \in (M \cap \mathbb{Q}^2)\backslash\Lambda$.

Proof. Note that, by Lemma 2.1, M is a closed set containing the lattice $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$. Assume that $(u,\eta) \in M\backslash\Lambda$ and define the set

$$\Lambda_1 := \Big\{(\tfrac{m}{Q}, nP) + k(u,\eta) : m,n,k \in \mathbb{Z}\Big\} \subset M.$$

Assume that $u \notin \mathbb{Q}$, $\eta \in \mathbb{Q}$, $\eta = \frac{c}{d}$, $c, d \in \mathbb{Z}$, $d \neq 0$, and put $v := dPu \notin \mathbb{Q}$. Let $x := \frac{1}{2Q}$. Since $\{\frac{m}{Q} + \ell v : m, \ell \in \mathbb{Z}\}$ is dense in $\mathbb{R}$, we can choose $m, \ell \in \mathbb{Z}$ such that $x' := u + \frac{m}{Q} + \ell v$ is arbitrary close to x. Then

$$(x',\eta) = \Big(\tfrac{m}{Q}, -\ell c P\Big) + (1 + \ell d P)(u,\eta) \in \Lambda_1.$$

Hence, $(\frac{1}{2Q},\eta) \in \overline{\Lambda_1}\backslash\Lambda \subset M\backslash\Lambda$.

Assume now that $\eta \notin \mathbb{Q}$. Since $\{nP + k\eta : n,k \in \mathbb{Z}\}$ is dense in $\mathbb{R}$, there exist sequences $(n_j)_{j\in\mathbb{N}}$ and $(k_j)_{j\in\mathbb{N}}$ in $\mathbb{Z}$ such that $n_j P + k_j \eta \to \frac{1}{2}$ as $j \to \infty$. For each $j \in \mathbb{N}$ pick $m_j \in \mathbb{Z}$ such that $\frac{m_j}{Q} + k_j u \in [0, \frac{1}{Q}]$. Then this sequence is bounded and thus has a convergent subsequence. By x denote its limit. Then $(x, \frac{1}{2}) \in M\backslash\Lambda$. If $x \in \mathbb{Q}$, we have reached our aim. Otherwise the above reasoning applies again. □

Proof of Proposition 4.1. In the sequel we aim to deduce a contradiction from the three following assumptions (where $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$ with $P, Q \in \mathbb{N}$):

(i) $Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.
(ii) The Gabor system (g,Λ) is a Riesz basis of $\mathcal{G}(g,\Lambda)$.
(iii) $\pi(u,\eta)g \in \mathcal{G}(g,\Lambda)$ for some $(u,\eta) \in \mathbb{R}^2\backslash\Lambda$.

Due to Lemma 4.2 we may assume in (iii) that $(u,\eta) \in \mathbb{Q}^2\backslash\Lambda$. Our strategy is as follows: First, from (ii) and (iii) we obtain an equation of the form

$$\mathbf{A}(x,\omega) = e^{-2\pi i \eta(x+u)} D_P^{-1} \mathbf{A}(x+u, \omega+\eta)\mathbf{M}(x+u,\omega+\eta),$$

which holds for a.e. $(x,\omega) \in \mathbb{R}^2$. Here, $\mathbf{A}$ is the matrix function from Lemma 2.3, D_P is a constant diagonal scaling matrix and $\mathbf{M}$ is a matrix function satisfying certain periodicity properties. We then iterate this equation by successively replacing (x,ω) by $(x+u, \omega+\eta)$. As $\mathbf{A}$ is quasi-periodic and $u, \eta \in \mathbb{Q}$, this process ends at a certain point with $\mathbf{A}(x,\omega)$ on both sides and we end up with an equation

$$\prod_{n=1}^{N} \mathbf{M}(x+nu, \omega+n\eta) = e^{2\pi i(M_1 x + M_2 \omega)} I_Q.$$

By applying the determinant on both sides, we are in the situation of Proposition 3.6, which finally implies that $(u,\eta)\in\Lambda$.

So, let us assume that (i)–(iii) are satisfied. Since the system $(g,\Lambda)=\{\pi(\frac{m}{Q},nP)g : m,n\in\mathbb{Z}\}$ is a Riesz basis of $\mathcal{G}(g,\Lambda)$ by (ii) and $\pi(u,\eta)g\in\mathcal{G}(g,\Lambda)$, there exists $(c_{m,n})_{m,n\in\mathbb{Z}}\in\ell^2(\mathbb{Z}^2)$ such that

$$\pi(u,\eta)g=\sum_{m,n\in\mathbb{Z}}c_{m,n}\pi(\tfrac{m}{Q},nP)g=\sum_{\ell=0}^{Q-1}\sum_{s,n\in\mathbb{Z}}c_{sQ+\ell,n}\,\pi(s+\tfrac{\ell}{Q},nP)g,$$

which converges in $L^2(\mathbb{R})$. Denoting $G:=Zg$, an application of the Zak transform gives (see Lemma 2.2)

$$\begin{aligned}e^{2\pi i\eta x}\,G(x-u,\omega-\eta)&=\sum_{\ell=0}^{Q-1}\sum_{s,n\in\mathbb{Z}}c_{sQ+\ell,n}e^{2\pi i(nPx-s\omega)}G(x-\tfrac{\ell}{Q},\omega)\\&=\sum_{\ell=0}^{Q-1}F_\ell(x,\omega)\,G(x-\tfrac{\ell}{Q},\omega)\qquad\text{for a.e. }(x,\omega)\in\mathbb{R}^2,\end{aligned}\tag{4.1}$$

where $F_\ell(x,\omega):=\sum_{s,n\in\mathbb{Z}}c_{sQ+\ell,n}e^{2\pi i(nPx-s\omega)}$. By definition, each F_ℓ is 1-periodic in ω (i.e., $F_\ell(x,\omega+1)=F_\ell(x,\omega)$ for a.e. $(x,\omega)\in\mathbb{R}^2$) and $\frac{1}{P}$-periodic in x. Replacing x by $x-\frac{k}{P}$ in (4.1), $k=0,\dots,P-1$, yields

$$e^{2\pi i\eta(x-\frac{k}{P})}G(x-u-\tfrac{k}{P},\omega-\eta)=(\mathbf{A}(x,\omega)F(x,\omega))_k,$$

where $\mathbf{A}$ is the matrix function from Lemma 2.3 and $F:=(F_0,\dots,F_{Q-1})^T$. Thus, we have

$$\mathbf{A}(x,\omega)F(x,\omega)=e^{2\pi i\eta x}D_P\mathbf{A}(x-u,\omega-\eta)e_0\qquad\text{for a.e. }(x,\omega)\in\mathbb{R}^2,\tag{4.2}$$

where $D_P=\operatorname{diag}(\exp(-2\pi i\eta\frac{k}{P}))_{k=0}^{P-1}$ and e_j is the $(j+1)$-th standard basis vector of $\mathbb{C}^Q$, $j=0,\dots,Q-1$. Note that each entry of $\mathbf{A}$ is a function in $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$ by (i), Lemma 3.8, and Lemma 2.3. The identity in (4.2) implies (cf. Lemma 2.3)

$$F(x,\omega)=e^{2\pi i\eta x}\left(\mathbf{A}(x,\omega)^*\mathbf{A}(x,\omega)\right)^{-1}\mathbf{A}(x,\omega)^*D_P\mathbf{A}(x-u,\omega-\eta)e_0.$$

Hence, from the periodicity of F_ℓ and Corollary 3.5 we infer that $F_\ell\in\mathrm{VMO}(\mathbb{R}^2)\cap L^\infty(\mathbb{R}^2)$ for $\ell=0,\dots,Q-1$. From (4.2) we also obtain

$$\mathbf{A}(x-u,\omega-\eta)e_0=e^{-2\pi i\eta x}D_P^{-1}\mathbf{A}(x,\omega)F(x,\omega).$$

Note that $\mathbf{A}(x,\omega)e_\ell=\mathbf{A}(x-\frac{\ell}{Q},\omega)e_0$ and $\mathbf{A}(x-\frac{\ell}{Q},\omega)=\mathbf{A}(x,\omega)\mathbf{R}(\omega)^\ell$, where $\mathbf{R}(\omega)\in\mathbb{C}^{Q\times Q}$ is the matrix

$$\begin{pmatrix}0&&&e_w\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{pmatrix},$$

with $e_\omega := e^{-2\pi i\omega}$. Therefore,

$$\begin{aligned}\mathbf{A}(x-u,\omega-\eta)e_\ell &= e^{-2\pi i\eta(x-\frac{\ell}{Q})}D_P^{-1}\mathbf{A}(x-\tfrac{\ell}{Q},\omega)F(x-\tfrac{\ell}{Q},\omega)\\ &= e^{-2\pi i\eta(x-\frac{\ell}{Q})}D_P^{-1}\mathbf{A}(x,\omega)\mathbf{R}(\omega)^\ell F(x-\tfrac{\ell}{Q},\omega).\end{aligned}$$

Hence, if $\mathbf{M}(x,\omega)\in\mathbb{C}^{Q\times Q}$ denotes the matrix with columns $e^{2\pi i\eta\frac{\ell}{Q}}\mathbf{R}(\omega)^\ell F(x-\frac{\ell}{Q},\omega)$, $\ell=0,\dots,Q-1$, we obtain

$$\mathbf{A}(x-u,\omega-\eta)=e^{-2\pi i\eta x}D_P^{-1}\mathbf{A}(x,\omega)\mathbf{M}(x,\omega). \tag{4.3}$$

Note that $\mathbf{M}$ has the same periodicity in x and ω as F. Moreover, an easy calculation leads to

$$\mathbf{M}(x-\tfrac{1}{Q},\omega)=e^{-\frac{2\pi i}{Q}}\mathbf{R}(\omega)^{-1}\mathbf{M}(x,\omega)\mathbf{R}(\omega). \tag{4.4}$$

Now, let us iterate the relation (4.3):

$$\begin{aligned}\mathbf{A}(x,\omega) &= e^{-2\pi i\eta(x+u)}D_P^{-1}\mathbf{A}(x+u,\omega+\eta)\mathbf{M}(x+u,\omega+\eta)\\ &= e^{-2\pi i\eta(2x+3u)}D_P^{-2}\mathbf{A}(x+2u,\omega+2\eta)\mathbf{M}(x+2u,\omega+2\eta)\mathbf{M}(x+u,\omega+\eta)\\ &= \dots\\ &= e^{2\pi i\eta\left(Nx+\frac{N(N-1)}{2}u\right)}D_P^{-N}\mathbf{A}(x+Nu,\omega+N\eta)\prod_{n=1}^{N}\mathbf{M}(x+nu,\omega+n\eta),\end{aligned}$$

where the matrix product is to be read in terms of left multiplication. As $(u,\eta)\in\mathbb{Q}^2$, we may choose N such that $M_1:=-N\eta\in\mathbb{Z}$, $M_2:=-Nu\in\mathbb{Z}$, $Nu\eta\in 2\mathbb{Z}$, and $M_1/P=-N\eta/P\in\mathbb{Z}$. Then $D_P^{-N}=I_P$ and Lemma 2.2 yields $\mathbf{A}(x+Nu,\omega+N\eta)=\mathbf{A}(x-M_2,\omega-M_1)=e^{-2\pi iM_2\omega}\mathbf{A}(x,\omega)$, hence

$$\mathbf{A}(x,\omega)=e^{2\pi i(-M_1x-M_2\omega)}\mathbf{A}(x,\omega)\prod_{n=1}^{N}\mathbf{M}(x+nu,\omega+n\eta).$$

Since $\mathbf{A}(x,\omega)$ has a left inverse for a.e. $(x,\omega)\in\mathbb{R}^2$ by Lemma 2.3, we get

$$\prod_{n=1}^{N}\mathbf{M}(x+nu,\omega+n\eta)=e^{2\pi i(M_1x+M_2\omega)}I_Q\qquad\text{for a.e. }(x,\omega)\in\mathbb{R}^2.$$

Finally, we define the function $H:=\det\mathbf{M}$. Since each entry of $\mathbf{M}$ is contained in $\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$, we have $H\in\mathrm{VMO}^\infty_{\mathrm{loc}}(\mathbb{R}^2)$. In addition, H is $\frac{1}{P}$-periodic in x and 1-periodic in ω. But by (4.4), H is also $\frac{1}{Q}$-periodic in x. Since it satisfies

$$\prod_{n=0}^{N-1}H(x+nu,\omega+n\eta)=e^{2\pi i(QM_1x+QM_2\omega)}\qquad\text{for a.e. }(x,\omega)\in\mathbb{R}^2,$$

Proposition 3.6 implies that both NP and NQ divide $QM_1=-NQ\eta$, and N divides $QM_2=-QNu$. The last relation gives $u\in\frac{1}{Q}\mathbb{Z}$. From the first two relations it follows that P divides $Q\eta$ and that $\eta\in\mathbb{Z}$. But as P and Q are coprime, P divides η, i.e., $\eta\in P\mathbb{Z}$. This is the desired contradiction. □

In order to extend Proposition 4.1 to arbitrary rational lattices and lattices of rational density we make use of the so-called *metaplectic operators*. To describe this class of operators we first mention that any matrix in $\mathrm{SL}(2,\mathbb{R})$ can be expressed as a finite product of matrices of the form

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}, \qquad \alpha, \beta \in \mathbb{R}\backslash\{0\}. \tag{4.5}$$

Indeed, if $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc = 1$, then

$$S = \begin{pmatrix} 1 & 0 \\ ca^{-1} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -ab & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -a & 0 \\ 0 & -a^{-1} \end{pmatrix} \quad \text{if } a \neq 0$$

and

$$S = \begin{pmatrix} 1 & 0 \\ -cd & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} b^{-1} & 0 \\ 0 & b \end{pmatrix} \quad \text{if } a = 0.$$

This in particular shows that if $S \in \mathrm{SL}(2,\mathbb{Q})$, then the parameters α, β can be chosen to be rational.

It is known [9] that to each matrix $S \in \mathrm{SL}(2,\mathbb{R})$ there corresponds a (so-called metaplectic) unitary operator $U_S : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ such that

$$U_S \pi(x,\omega) U_S^* = \pi(S(x,\omega)^T) \quad \text{for all } (x,\omega) \in \mathbb{R}^2.$$

The operator U_S is unique up to scalar multiplication with unimodular constants. If $S, T \in \mathrm{SL}(2,\mathbb{R})$, the operator $U_S U_T$ is obviously a metaplectic operator corresponding to ST. That is, we have $U_{ST} = U_S U_T$. As is easily seen, the three types of matrices in (4.5), which generate $\mathrm{SL}(2,\mathbb{R})$, correspond to the metaplectic operators $\mathcal{F}$ (Fourier transform), D_α, and C_β (defined in Proposition 3.11), respectively.

Proof of Theorem 1.5. As in the proof of Proposition 4.1, in addition to the assumptions of Theorem 1.5, we shall assume that $Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ and derive a contradiction. We have $\Lambda = A\mathbb{Z}^2$, where $A \in \mathrm{GL}(2,\mathbb{Q})$. We may write $\det A = \frac{P}{Q}$, where $P, Q \in \mathbb{Z} \setminus \{0\}$ are coprime numbers. Then $B := \mathrm{diag}(\frac{1}{Q}, P)A^{-1} \in \mathrm{SL}(2,\mathbb{Q})$ can be expressed as a finite product of matrices (4.5) with $\alpha, \beta \in \mathbb{Q}\backslash\{0\}$. Hence, the metaplectic operator U_B can be written as a finite product of operators of the type $\mathcal{F}$, D_α, and C_β. Now, set

$$(u_1, \eta_1)^T := B(u,\eta)^T, \qquad g_1 := U_B g \in L^2(\mathbb{R}), \qquad \text{and} \qquad \Lambda_1 = B\Lambda = \tfrac{1}{Q}\mathbb{Z} \times P\mathbb{Z}. \tag{4.6}$$

Proposition 3.11 implies that $Zg_1 \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$. Also, $\pi(B\lambda)g_1 = U_B \pi(\lambda) U_B^{-1} g_1 = U_B \pi(\lambda) g$ for every $\lambda \in \Lambda$, which implies that (g_1, Λ_1) is a Riesz basis for its closed linear span $\mathcal{G}(g_1, \Lambda_1) = U_B \mathcal{G}(g, \Lambda)$. Moreover, the condition $\pi(u,\eta)g \in \mathcal{G}(g,\Lambda)$ immediately translates to

$$\pi(u_1,\eta_1)g_1 = \pi(B(u,\eta)^T) U_B g = U_B \pi(u,\eta) g \in U_B \mathcal{G}(g,\Lambda) = \mathcal{G}(g_1,\Lambda_1).$$

But as $(u_1, \eta_1) \notin \Lambda_1$ this is not possible by Proposition 4.1. □

Proof of Theorem 1.4. Let us first discuss the condition (1.2). Since $g \in L^2(\mathbb{R})$, the condition holds for some $\alpha, \beta \in \mathbb{R}$ if and only if it holds for all $\alpha, \beta \in \mathbb{R}$. At the same time, (1.2) exactly means that $g \notin H^{p/2}(\mathbb{R})$ or $g \notin H^{q/2}(\mathbb{R})$. Towards a contradiction,

in addition to the assumptions of Theorem 1.4, we assume that there exist some $p, q \in (1,\infty)$ with $\frac{1}{p}+\frac{1}{q}=1$ such that the product on the left hand side of (1.2) (with, e.g., $\alpha=\beta=0$) is finite. In other words, we assume that $g \in H^{p/2}(\mathbb{R})$ and $\hat{g} \in H^{q/2}(\mathbb{R})$.

Let Λ be an arbitrary lattice in $\mathbb{R}^2$ with rational density P/Q, where P and Q are coprime integers, and let $g \in L^2(\mathbb{R})$ be as in Theorem 1.4. Also, let $(u,\eta) \in \mathbb{R}^2 \backslash \Lambda$ such that $\pi(u,\eta)g \in \mathcal{G}(g,\Lambda)$. As in the proof of Theorem 1.5, we choose a matrix $B \in \mathrm{SL}(2,\mathbb{R})$ (here, B is allowed to have irrational entries) such that $B\Lambda = \Lambda_1 := \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$. Define (u_1,η_1), g_1, and Λ_1 as in (4.6). Then $g_1 \in L^2(\mathbb{R})$, (g_1,Λ_1) is a Riesz basis for its closed linear span, and $\pi(u_1,\eta_1)g_1 \in \mathcal{G}(g_1,\Lambda_1)$. Hence, by Proposition 4.1 it suffices to show that $Zg_1 \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$.

To simplify notations, for $p \in (1,\infty)$ let $\mathbb{H}^p(\mathbb{R})$ be the space of functions $f \in H^{p/2}(\mathbb{R})$ whose Fourier transform $\hat{f}$ is contained in $H^{q/2}(\mathbb{R})$, where $\frac{1}{p}+\frac{1}{q}=1$. Also, let $\mathbb{H} := \bigcup_{p\in(1,\infty)} \mathbb{H}^p(\mathbb{R})$. In what follows, we prove that $U\mathbb{H} \subset \mathbb{H}$ for any metaplectic operator U. It then follows that $g_1 = U_B g \in \mathbb{H}$. And since, by [7], we have $Z\mathbb{H} \subset \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$, we obtain $Zg_1 \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$, which was our aim.

Since every metaplectic operator is a finite product of the Fourier transform $\mathcal{F}$, dilations, and chirp muliplication, we only have to prove that $\mathcal{F}\mathbb{H} \subset \mathbb{H}$, $D_\alpha \mathbb{H} \subset \mathbb{H}$ for each $\alpha \in \mathbb{R}\setminus\{0\}$, and $C_1\mathbb{H} \subset \mathbb{H}$ (cf. (3.7)). Using the representation $H^s(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \omega^{2s}\widehat{f}^{\,2} \in L^1(\mathbb{R})\}$, the first two claims are almost immediate. We will now prove that $C_1\mathbb{H}^p(\mathbb{R}) \subset \mathbb{H}^p(\mathbb{R})$ for $p \in (1,\infty)$. So, let $g \in \mathbb{H}^p(\mathbb{R})$. Then $\hat{g} \in H^{q/2}(\mathbb{R})$ implies $x^q g^2 \in L^1(\mathbb{R})$ and hence $\widehat{C_1 g} \in H^{q/2}(\mathbb{R})$. In order to show that $C_1 g \in H^{p/2}(\mathbb{R})$, we make use of the following representation of fractional Sobolev spaces (see, e.g., [13]):

$$H^s(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}) : \frac{f(x)-f(y)}{(x-y)^{\frac{1}{2}+s}} \in L^2(\mathbb{R}^2) \right\}$$

for $s \in (0,1)$ and

$$H^s(\mathbb{R}) = \left\{ f \in H^m(\mathbb{R}) : f^{(m)} \in H^\sigma(\mathbb{R}) \right\}$$

for $s = m+\sigma$ with $m \in \mathbb{N}$ and $\sigma \in [0,1)$. Let $c(x) := e^{2\pi i x^2}$, $x \in \mathbb{R}$. Then $C_1 g(x) = c(x)g(x)$ and it is clear that $C_1 g \in H^{p/2}(\mathbb{R})$ if $p/2 \in \mathbb{N}$. We only prove the claim here for $1<p<2$. The rest is then straightforward. We have (setting $s=p/2$)

$$\frac{C_1 g(x) - C_1 g(y)}{(x-y)^{\frac{p+1}{2}}} = c(x)\frac{g(x)-g(y)}{(x-y)^{\frac{p+1}{2}}} + \frac{c(x)-c(y)}{(x-y)^{\frac{p+1}{2}}} g(y).$$

Hence, as $g \in H^{p/2}(\mathbb{R})$, it is left to show that the second summand is in $L^2(\mathbb{R}^2)$. To this end, fix $y \in \mathbb{R}$ and observe that

$$\int_{|x-y|>1} \frac{|c(x)-c(y)|^2}{|x-y|^{p+1}}\,dx \;\le\; c_1$$

and

$$\int_{|x-y|\le 1} \frac{|c(x)-c(y)|^2}{|x-y|^{p+1}}\,dx \;\le\; c_2(1+y^2),$$

where $c_1, c_2 > 0$ only depend on p. Hence,

$$\int_{\mathbb{R}} \frac{|c(x) - c(y)|^2}{|x-y|^{p+1}}\, dx \;\leq\; c(1+y^2)$$

for each $y \in \mathbb{R}$. Now, as $\hat{g} \in H^{q/2}(\mathbb{R})$, we have $(1+y^2)^{q/2}|g(y)|^2 \in L^1(\mathbb{R})$, and $p < 2$ implies $q/2 > 1$. Thus, also $(1+y^2)|g(y)|^2 \in L^1(\mathbb{R})$, and the proof is complete. □

Problem 4.3. As already mentioned, we have restricted ourselves to rational lattices in Theorem 1.5 while Theorem 1.4 considers a broader class of lattices, namely the lattices of rational density. The main reason for this is that we could not prove whether the set $\{g \in L^2(\mathbb{R}) : Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)\}$ is invariant under irrational dilations. If this were true, Theorem 1.5 would hold not only for rational lattices but for arbitrary lattices of rational density. We leave the following as open problems:

(1) Is it true that $Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ for $g \in L^2(\mathbb{R})$ implies $Z(D_\alpha g) \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$ for every $\alpha \in \mathbb{R}\backslash\mathbb{Q}$?
(2) Is there a good description of the space consisting of the functions $g \in L^2(\mathbb{R})$ that satisfy $Zg \in \mathrm{VMO}_{\mathrm{loc}}(\mathbb{R}^2)$?

Acknowledgements. D.G. Lee and G.E. Pfander acknowledge support by the DFG Grants PF 450/6-1 and PF 450/9-1. The authors would like to thank S. Zubin Gautam, Jeffrey A. Hogan, and David F. Walnut for valuable discussions. They also thank the referees for having read the manuscript carefully and their various valuable comments.

References

[1] A. Aldroubi, Q. Sun, and H. Wang, Uncertainty principles and Balian-Low type theorems in principal shift-invariant spaces, Appl. Comput. Harmon. Anal. 30 (2011), 337–347.
[2] R. Balian, Un principe d'incertitude fort en théorie du signal ou en mécanique quantique, C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre 292 (1981), 1357–1362.
[3] J.J. Benedetto, C. Heil, and D.F. Walnut, Differentiation and the Balian-Low theorem, J. Fourier Anal. Appl. 1 (1995), 355–402.
[4] C. Cabrelli, U. Molter, and G.E. Pfander, Time-frequency shift invariance and the Amalgam Balian-Low theorem, Appl. Comput. Harmon. Anal. 41 (2016), 677–691.
[5] O. Christensen, An introduction to frames and Riesz bases, Birkhäuser, Boston, Basel, Berlin 2003.
[6] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inf. Theory 36 (1990), 961–1005.
[7] S.Z. Gautam, A critical-exponent Balian-Low theorem, Math. Res. Lett. 15 (2008), 471–483.
[8] J. Gebardo and D. Han, Balian-Low phenomenon for subspace Gabor frames, J. Math. Phys. 45 (2004), 3362–3378.
[9] K. Gröchenig, Foundations of time-frequency analysis, Birkhäuser, Boston, Basel, Berlin, 2001.
[10] C. Heil and R. Tinaztepe, Modulation spaces, BMO, and the Balian-Low theorem, Sampl. Theory Signal Image Process. 11 (2012), 25–41.
[11] A.J.E.M. Janssen, The Zak transform: A signal transform for sampled time-continuous signals, Philips J. Res. 43 (1988), 23–69.
[12] F. Low, Complete sets of wave packets, A Passion for Physics—Essays in Honor of Geoffrey Chew (edited by C. DeTar et al.), Singapore, World Scientific, 17–22, 1985.
[13] E. Di Nezza, G. Palatucci, and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces, Bull. Sci. math. 136 (2012), 521–573.

[14] S. Nitzan and J.-F. Olsen, A quantitative Balian-Low theorem, J. Fourier Anal. Appl. 19 (2013), 1078–1092.
[15] D. Sarason, Functions of vanishing mean oscillation, Trans. Am. Math. Soc. 207 (1975), 391–405.

Author affiliations

A. Caragea: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
E-mail address: andrei.caragea@gmail.com
URL: http://www.ku.de/?acaragea

D.G. Lee: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
E-mail address: daegwans@gmail.com

G.E. Pfander: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
E-mail address: pfander@ku.de
URL: http://www.ku.de/?pfander

F. Philipp: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
E-mail address: fmphilipp@gmail.com
URL: http://www.ku.de/?fmphilipp

6.2 Paper 2

A QUANTITATIVE SUBSPACE BALIAN-LOW THEOREM

ANDREI CARAGEA, DAE GWAN LEE, FRIEDRICH PHILIPP, AND FELIX VOIGTLAENDER

Abstract. Let $\mathcal{G} \subset L^2(\mathbb{R})$ be the subspace spanned by a Gabor Riesz sequence (g, Λ) with $g \in L^2(\mathbb{R})$ and a lattice $\Lambda \subset \mathbb{R}^2$ of rational density. It was shown recently that if g is well-localized both in time and frequency, then $\mathcal{G}$ cannot contain any time-frequency shift $\pi(z)g$ of g with $z \in \mathbb{R}^2 \setminus \Lambda$. In this paper, we improve the result to the quantitative statement that the L^2-distance of $\pi(z)g$ to the space $\mathcal{G}$ is equivalent to the Euclidean distance of z to the lattice Λ, in the sense that the ratio between those two distances is uniformly bounded above and below by positive constants. On the way, we prove several results of independent interest, one of them being closely related to the so-called weak Balian-Low theorem for subspaces.

1. Introduction

The Balian-Low theorem is a well known and fundamental result in time-frequency analysis, which asserts that a Gabor system cannot be a Riesz basis for $L^2(\mathbb{R})$ if its generating window is well localized both in time and frequency. More precisely, it states the following:

Theorem 1.1 (Balian-Low Theorem)**.** *Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that the Gabor system $\{e^{2\pi ibx} g(x-a) : (a,b) \in \Lambda\}$ is a Riesz basis for $L^2(\mathbb{R})$ (and therefore Λ is of density 1). Then*

$$\left(\int x^2 \, |g(x)|^2 \, dx \right) \left(\int \omega^2 \, |\widehat{g}(\omega)|^2 \, d\omega \right) = \infty. \tag{1.1}$$

Recently, the following generalization of the Balian-Low theorem was proved in [4] (see also [6] for a similar generalization of the amalgam Balian-Low theorem).

Theorem 1.2 ([4])**.** *Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density such that the Gabor system $\{e^{2\pi ibx} g(x-a) : (a,b) \in \Lambda\}$ is a Riesz basis for its closed linear span $\mathcal{G}(g, \Lambda)$. If there exists a time-frequency shift $e^{2\pi i \eta x} g(x-u)$, $(u, \eta) \in \mathbb{R}^2 \setminus \Lambda$, of g which is contained in $\mathcal{G}(g, \Lambda)$, then* (1.1) *holds.*

Note that condition (1.1) is equivalent to having $g \notin H^1(\mathbb{R})$ or $\widehat{g} \notin H^1(\mathbb{R})$, where $H^1(\mathbb{R})$ denotes the usual Sobolev space in $L^2(\mathbb{R})$ of regularity order 1. Therefore, Theorem 1.2 can be rephrased as follows: if $g, \widehat{g} \in H^1(\mathbb{R})$, then the time-frequency shift $e^{2\pi i \eta x} g(x-u)$ has a positive L^2-distance to the space $\mathcal{G}(g, \Lambda)$ whenever $(u, \eta) \in \mathbb{R}^2$ has a positive Euclidean distance to the lattice Λ. As our main result, we are going to prove the following quantitative version of Theorem 1.2 which relates the two mentioned distances. In the sequel, we denote by $\mathbb{H}^1(\mathbb{R})$ the set of all $g \in H^1(\mathbb{R})$ satisfying $\widehat{g} \in H^1(\mathbb{R})$.

2010 *Mathematics Subject Classification.* Primary: 42C15. Secondary: 42C30, 42C40.

Key words and phrases. Balian-Low Theorem; Weak subspace Balian-Low Theorem; Gabor systems; Time frequency shift invariance; Zak transform.

Theorem 1.3. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density such that $\{e^{2\pi ibx} g(x-a) : (a,b) \in \Lambda\}$ is a Riesz basis for its closed linear span $\mathcal{G}(g,\Lambda)$. Then there exist constants $C_1, C_2 > 0$ such that for all $(u,\eta) \in \mathbb{R}^2$ we have*

$$C_1 \cdot \operatorname{dist}\big((u,\eta),\Lambda\big) \leq \operatorname{dist}\big(e^{2\pi i\eta x} g(x-u), \mathcal{G}(g,\Lambda)\big) \leq C_2 \cdot \operatorname{dist}\big((u,\eta),\Lambda\big). \tag{1.2}$$

The *upper* bound in (1.2) in fact holds for any $g \in \mathbb{H}^1(\mathbb{R})$ and any lattice $\Lambda \subset \mathbb{R}^2$, regardless of $\{e^{2\pi ibx} g(x-a) : (a,b) \in \Lambda\}$ being a Riesz sequence or the lattice Λ having rational density; besides, an explicit constant C_2 can be found easily; see Proposition 4.1 below. On the other hand, finding an explicit constant C_1 is more elusive. Even in the case where (g,Λ) forms an orthonormal system, we were only able to derive a constant C_1 such that (1.2) holds for (u,η) close to the lattice Λ; see Theorem 5.4. We expect such a constant to depend on the Riesz bounds of $\{e^{2\pi ibx} g(x-a) : (a,b) \in \Lambda\}$ and on the norms $\|g\|_{L^2}$, $\|g\|_{H^1}$, and $\|\widehat{g}\|_{H^1}$.

Quantitative Balian-Low estimates for general elements in the Gabor space. Writing $\pi(u,\eta) f(x) = e^{2\pi i\eta x} f(x-u)$, one might wonder whether the estimate

$$\operatorname{dist}\big(\pi(u,\eta) f, \mathcal{G}(g,\Lambda)\big) \asymp \operatorname{dist}\big((u,\eta),\Lambda\big) \cdot \|f\|_{L^2}$$

holds for general $f \in \mathcal{G}(g,\Lambda)$ and not just for $f = g$. In general this is *not* the case. Indeed, if g is a Gaussian, then $(g, 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z})$ is a Riesz basis for its closed linear span $\mathcal{G}(g, 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z})$, but there exists a function $0 \neq f \in \mathcal{G}(g, 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z})$ satisfying $f(\cdot - 1) \in \mathcal{G}(g, 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z})$ (see Example A.17 for details). Therefore, we see that the distance $\operatorname{dist}\big(\pi(1,0) f, \mathcal{G}(g, 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z})\big)$ vanishes, even though $\operatorname{dist}((1,0), 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z}) \cdot \|f\|_{L^2} \neq 0$.

Implications regarding the OFDM communication scheme. One motivation for analyzing the distance of the time-frequency shift $\pi(u,\eta)g$ to the Gabor space $\mathcal{G}(g,\Lambda)$ stems from the communication scheme called *orthogonal frequency division multiplexing* (OFDM). In OFDM, the sender wants to transmit the coefficients $c = (c_{k,\ell})_{k,\ell\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$ to the receiver. This is done by selecting a fixed Gabor Riesz sequence $\big(\pi(k\alpha,\ell\beta)g\big)_{k,\ell\in\mathbb{Z}}$ to form the *transmission signal* $Fc = \sum_{k,\ell\in\mathbb{Z}} c_{k,\ell}\, \pi(k\alpha,\ell\beta)g$, which is then sent to the receiver through a *communication channel*. Mathematically, the effect of the channel is modeled as a linear operator $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$; that is, the signal that arrives at the receiver is TFc instead of Fc.

The first step of the reconstruction procedure in OFDM is to apply the *reconstruction operator* R given by $Rf = \big(\langle f, \pi(k\alpha,\ell\beta)g^\circ\rangle\big)_{k,\ell\in\mathbb{Z}}$ to the signal TFc, thereby obtaining the sequence $\widetilde{c} = RTFc \in \ell^2(\mathbb{Z}^2)$. Here, g° is the dual window for the Riesz sequence $\big(\pi(k\alpha,\ell\beta)g\big)_{k,\ell\in\mathbb{Z}}$; i.e., $g^\circ \in \mathcal{G} := \mathcal{G}(g, \alpha\mathbb{Z}\times\beta\mathbb{Z})$ satisfies the biorthogonal property $\langle \pi(k\alpha,\ell\beta)g, \pi(k'\alpha,\ell'\beta)g^\circ\rangle = \delta_{k,k'}\delta_{\ell,\ell'}$ for $k,k',\ell,\ell' \in \mathbb{Z}$. At least for the ideal communication channel $T = \mathrm{Id}_{L^2(\mathbb{R})}$, this guarantees perfect reconstruction, meaning that $\widetilde{c} = c$. For more general channels, this is not the case, but one might hope to reconstruct c by applying a suitable (linear) *post-processing operator* P to $\widetilde{c}$.

In fact, there exists such a bounded post-processing operator P satisfying $PRTFc = c$ for all $c \in \ell^2(\mathbb{Z}^2)$ if and only if the operator RT is bounded below on the Gabor space $\mathcal{G}$, meaning that $\|RTf\|_{\ell^2} \gtrsim \|f\|_{L^2}$ for all $f \in \mathcal{G}$. It is not hard to see that $\|Rh\|_{\ell^2} \asymp \|\mathbb{P}h\|_{L^2}$ for $h \in L^2(\mathbb{R})$, where we denote by $\mathbb{P}$ the orthogonal projection onto the Gabor space $\mathcal{G}$. For the important special case that T is a pure time-frequency shift $\pi(u,\eta)$, reconstruction is thus possible if and only if $\|\mathbb{P}\pi(u,\eta)f\|_{L^2} \gtrsim \|f\|_{L^2}$ for all $f \in \mathcal{G}$, which is equivalent to the existence of a constant $c < 1$ satisfying

$$\operatorname{dist}\big(\pi(u,\eta)f, \mathcal{G}\big) = \big\|(I - \mathbb{P})\pi(u,\eta)f\big\|_{L^2} \leq c\,\|f\|_{L^2}, \qquad \forall\, f \in \mathcal{G}. \tag{1.3}$$

The term $\operatorname{dist}(\pi(u,\eta)f, \mathcal{G})$ measures the *off-band energy loss* caused by the time-frequency shift $\pi(u,\eta)$, that is, the proportion of the signal energy that gets "pushed out of the Gabor space" by applying the time-frequency shift $\pi(u,\eta)$. Even in the case where the off-band energy loss is small enough so that (1.3) holds, it is interesting to know more precise upper and lower bounds for this quantity, since it influences the stability of the reconstruction. Theorem 1.3 shows that in the case $f = g$, the off-band energy loss is of the order $\operatorname{dist}((u,\eta), \alpha\mathbb{Z}\times\beta\mathbb{Z})$.

Structure of the proof. The outline of the proof of Theorem 1.3 is as follows. First of all, we note that it suffices to establish the inequality (1.2) for (u,η) in a neighborhood of the origin $(0,0)$; the inequality then holds for all $(u,\eta) \in \mathbb{R}^2$ (with possibly different constants C_1 and C_2) by Theorem 1.2 and a compactness argument. In order to analyze the behavior of the quantity $\operatorname{dist}(\pi(u,\eta)g, \mathcal{G}(g,\Lambda))$ for (u,η) close to the origin, we first show that the *time-frequency map* $S_g : (a,b) \mapsto \pi(a,b)g$ is differentiable at $(0,0)$ with (Fréchet) derivative $(a,b) \mapsto -ag' + 2\pi ibXg$, where X is the position operator defined formally by $Xf(x) = xf(x)$; see Lemma 3.2. We then prove in Proposition 4.4 that $-ag' + 2\pi ibXg$ is not contained in $\mathcal{G}(g,\Lambda)$ unless $a = b = 0$. Denoting by $\mathbb{P}$ the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda)$, this implies that there exists a constant $\gamma > 0$ with $\|(I-\mathbb{P})(-ag' + 2\pi ibXg)\|_{L^2} \geq \gamma \|(a,b)\|_2$ for all $(a,b) \in \mathbb{R}^2$. The claim then follows immediately because $(a,b) \mapsto (I-\mathbb{P})(-ag' + 2\pi ibXg)$ linearizes the map $(a,b) \mapsto (I-\mathbb{P})(e^{2\pi ibx}g(x-a))$ in a neighborhood of $(0,0)$.

The main ingredients of the proof are thus the differentiability of the time-frequency map (see Section 3) and the fact that none of its directional derivatives $-ag' + 2\pi ibXg$ with $(a,b) \in \mathbb{R}^2\backslash\{(0,0)\}$ are contained in $\mathcal{G}(g,\Lambda)$ (see Proposition 4.4). While the former is probably folklore (although we could not find a reference), the latter seems to be a new result and should be interesting in its own right. We also point out a close relationship between Proposition 4.4 and the weak Balian-Low theorem for subspaces from [13]; see Remark 4.5 for a detailed discussion.

As mentioned above, we were unable to derive a closed-form formula for the constant C_1 in Equation (1.2). However, if we assume the Gabor system $\{e^{2\pi ibx}g(x-a) : (a,b)\in\Lambda\}$ to be *orthonormal*, then we can find an *explicit* constant $C_1 > 0$ such that (1.2) holds for all (u,η) in a neighborhood of the lattice Λ; see Theorem 5.4. This result then leads to a statement similar to Theorem 1.2 but without assuming the rational density of Λ; see Corollary 5.5.

The paper is organized as follows: In Section 2, we show how the main properties that we are interested in (the regularity of g, the property of (g,Λ) being a Riesz sequence, and the distance $\operatorname{dist}(\pi(\mu), \mathcal{G}(g,\Lambda)))$ can be described via the Zak transform and certain associated matrix multiplication operators. Section 3 contains the aforementioned differentiability result for the time-frequency map of $\mathbb{H}^1(\mathbb{R})$ functions. The proof of Theorem 1.3 is given in Section 4. Finally, in Section 5 we provide an explicit local lower bound C_1 in the case where the Gabor system is orthonormal.

Several results that are technical or only tangentially related to the core arguments are deferred to A. Although most of them should be well-known or be considered folklore, we either give detailed references or include their proofs for the sake of completeness.

2. Preparations

Notation. Let us begin with collecting some notation which will be used throughout the paper. We set $\mathbb{N} := \{1, 2, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The closure of a subset $M \subset X$ of a metric space X will be denoted by $\overline{X}$. The *Lebesgue measure* of a Borel set $E \subset \mathbb{R}^n$ is denoted by $\boldsymbol{\lambda}(E)$. If $g : \mathbb{R} \to \mathbb{C}$ is measurable, we write Xg for the function $x \mapsto x\, g(x)$, that is, $(Xg)(x) = x\, g(x)$, $x \in \mathbb{R}$.

Let $\mathcal{H}$ be a Hilbert space, and $\Phi = (\varphi_i)_{i \in I}$ be a family of vectors in $\mathcal{H}$. This family is called a *frame* for $\mathcal{H}$ if $A \|f\|_{\mathcal{H}}^2 \leq \sum_{i \in I} |\langle f, \varphi_i \rangle|^2 \leq B \|f\|_{\mathcal{H}}^2$ for all $f \in \mathcal{H}$ and certain constants $A, B \in (0, \infty)$. If Φ is a frame for its closed linear span $\overline{\operatorname{span}}\{\varphi_i : i \in I\}$, then we say that Φ is a *frame sequence*. We say that Φ is a *Riesz sequence* if there are $A, B \in (0, \infty)$ such that $A \|c\|_{\ell^2} \leq \left\| \sum_{i \in I} c_i \varphi_i \right\| \leq B \|c\|_{\ell^2}$ for all finitely supported sequences $c = (c_i)_{i \in I} \in \ell^2(I)$. If Φ is a Riesz sequence and $\operatorname{span}\{\varphi_i : i \in I\}$ is dense in $\mathcal{H}$, we say that Φ is a *Riesz basis* for $\mathcal{H}$. Each Riesz basis is a frame.

Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator on a (complex) Hilbert space $\mathcal{H}$. The *spectrum* of T will be denoted by $\sigma(T)$; that is,

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not boundedly invertible}\}.$$

We denote by $\varrho(T)$ the complement set of $\sigma(T)$ in $\mathbb{C}$ which is called the *resolvent set* of T. For a bounded linear operator $A : \mathcal{H} \to \mathcal{K}$ between two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, we define

$$\sigma_0(A) := \sqrt{\min \sigma(A^* A)} \in [0, \infty) \quad \text{and} \quad \sigma_1(A) := \sqrt{\inf[\sigma(A^* A) \setminus \{0\}]} \in [0, \infty]. \tag{2.1}$$

Note that $\sigma(A^* A) = \{0\}$ if and only if $A = 0$, in which case we have $\sigma_1(A) = \infty$; on the other hand, we have $\sigma_1(A) < \infty$ for $A \neq 0$. If A is a matrix, then $\sigma_0(A)$ is the smallest singular value of A, while $\sigma_1(A)$ is the smallest *positive* singular value of A.

We occasionally consider the vector-valued L^2 space $L^2(\Omega; \mathbb{C}^k)$, which we equip with the inner product $\langle f, g \rangle = \int \langle f(\omega), g(\omega) \rangle_{\mathbb{C}^k} d\mu(\omega)$, where $\langle \cdot, \cdot \rangle_{\mathbb{C}^k}$ denotes the standard inner product on $\mathbb{C}^k$.

The *Fourier transform* $\widehat{g}$ of $g \in L^2(\mathbb{R})$ is defined by

$$\widehat{g}(\omega) := \lim_{R \to \infty} \int_{-R}^{R} g(x)\, e^{-2\pi i x \omega}\, dx\,,$$

where the limit is taken in $L^2(\mathbb{R})$. For $a, b \in \mathbb{R}$ we also define the time-frequency shift operator

$$[\pi(a, b) g](x) := e^{2\pi i b x} g(x - a), \qquad x \in \mathbb{R}\,,$$

which can be expressed as $\pi(a, b) = M_b T_a$ where T_a and M_b denote the operators of *translation* by $a \in \mathbb{R}$ and *modulation* by $b \in \mathbb{R}$, respectively. For $k \in \mathbb{N}$, we set $\mathbb{H}^k(\mathbb{R}) := \{f \in H^k(\mathbb{R}) : \widehat{f} \in H^k(\mathbb{R})\}$, with the usual (complex-valued) L^2-Sobolev space $H^k(\mathbb{R}) = W^{k,2}(\mathbb{R})$.

A *lattice* in $\mathbb{R}^2$ is a set $\Lambda = A\mathbb{Z}^2$ with $A \in \mathrm{GL}(2, \mathbb{R})$. Its *density* is defined as $|\det A|^{-1}$. If Λ is a lattice in $\mathbb{R}^2$ and $g \in L^2(\mathbb{R})$, we denote by (g, Λ) the *Gabor system* generated by g and Λ, that is,

$$(g, \Lambda) := \{\pi(\lambda) g : \lambda \in \Lambda\}.$$

The *Gabor space* generated by g and Λ is defined as $\mathcal{G}(g, \Lambda) := \overline{\operatorname{span}}\,(g, \Lambda)$, with the closure taken in $L^2(\mathbb{R})$.

The *Zak transform* of $g \in L^2(\mathbb{R})$ is defined as

$$Zg(x,\omega) = \lim_{N\to\infty} \sum_{k=-N}^{N} e^{2\pi i k\omega} g(x-k), \qquad (x,\omega) \in (0,1)^2, \tag{2.2}$$

where the limit is taken in $L^2((0,1)^2)$. The Zak transform $g \mapsto Zg$ is a unitary operator from $L^2(\mathbb{R})$ to $L^2((0,1)^2)$. In the following, we will consider the Zak transform Zg of $g \in L^2(\mathbb{R})$ as an (a.e. defined) function on $\mathbb{R}^2$, by using Equation (2.2) on all of $\mathbb{R}^2$, where the limit is taken in $L^2_{\mathrm{loc}}(\mathbb{R}^2)$. This extended Zak transform has the following properties (all of which hold for a.e. $(x,\omega) \in \mathbb{R}^2$):

(a) $Zg(x+m,\omega+n) = e^{2\pi i m\omega}\, Zg(x,\omega)$ for all $m,n \in \mathbb{Z}$.
(b) $Z[\pi(u,\eta)g](x,\omega) = e^{2\pi i \eta x}\, Zg(x-u,\omega-\eta)$ for all $(u,\eta) \in \mathbb{R}^2$.
(c) $(Z[\pi(m,n)g])(x,\omega) = e^{2\pi i(nx-m\omega)}\, Zg(x,\omega)$ for all $m,n \in \mathbb{Z}$.
(d) $Z\widehat{g}(x,\omega) = e^{2\pi i x\omega}\, Zg(-\omega,x)$.
(e) $g(x) = \int_0^1 Zg(x,\omega)\, d\omega$ and $\widehat{g}(\omega) = \int_0^1 e^{-2\pi i x\omega} Zg(x,\omega)\, dx$.

For all these properties, we refer to [12, Chapter 8]. The property (a) of Zg is called *quasi-periodicity*.

2.1. Reduction to matrix multiplication operators

In this subsection, we show that the properties and quantities that we are interested in—the distance $\mathrm{dist}(\pi(\mu)g, \mathcal{G}(g,\Lambda))$ and whether (g,Λ) is a Riesz sequence—can be conveniently reformulated using certain matrix multiplication operators

$$M_A : L^2(\Omega;\mathbb{C}^k) \to L^2(\Omega;\mathbb{C}^\ell), f \mapsto A(\cdot)f(\cdot).$$

Here, the matrix function $A : \Omega \to \mathbb{C}^{\ell\times k}$ will be defined using the Zak transform. More details regarding these matrix multiplication operators can be found in A.1.

We start by considering the Gabor system (g,Λ) associated to the lattice $\Lambda = \frac{1}{Q}\mathbb{Z}\times P\mathbb{Z}$ (where $P,Q \in \mathbb{N}$) and connect the spectral properties of the *frame operator*

$$\mathbf{S}: \quad L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad f \mapsto \sum_{\lambda\in\Lambda} \langle f, \pi(\lambda)g\rangle\, \pi(\lambda)g$$

and the *Gram operator* $\mathbf{G} : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ defined by

$$\mathbf{G}(c_{n,k})_{n,k\in\mathbb{Z}} = \Big(\big\langle \sum_{n,k\in\mathbb{Z}} c_{n,k}\pi(Q^{-1}n, Pk)g,\ \pi(Q^{-1}m, P\ell)g \big\rangle \Big)_{m,\ell\in\mathbb{Z}}$$

to matrix multiplication operators on the domain $R_P := (0,\frac{1}{P})\times(0,1)$. This relies on using the unitary operators $\mathcal{V} : L^2((0,1)^2) \to L^2(R_P,\mathbb{C}^P)$ and $\mathcal{U} : \ell^2(\mathbb{Z}^2) \to L^2(R_P,\mathbb{C}^Q)$, defined by

$$(\mathcal{V}f)(x,\omega) := \big(f(x+\tfrac{k}{P},\omega)\big)_{k=0}^{P-1} \quad \text{and} \quad \mathcal{U}c = \Big(\sum_{s,n\in\mathbb{Z}} c_{sQ+\ell,n}\, e_{s,n} \Big)_{\ell=0}^{Q-1}, \tag{2.3}$$

where $f \in L^2((0,1)^2)$ and $c = (c_{n,m})_{n,m\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$, and where we use the function $e_{s,n}(x,\omega) := P^{1/2}\cdot e^{2\pi i(nPx - s\omega)}$ defined for $(x,\omega) \in R_P$. Furthermore, we denote by $S_n \in \mathbb{C}^{n\times n}$ the cyclic shift operator satisfying $S_n e_i = e_{i-1}$ for $i \in \{1,\dots,n-1\}$ and $S_n e_0 = e_{n-1}$ for the standard basis $\{e_0,\dots,e_{n-1}\}$ of $\mathbb{C}^n$. Finally, for $\omega \in \mathbb{R}$ we define the matrices

$$L_\omega := S_P\, \mathrm{diag}(e^{2\pi i\omega}, 1, \dots, 1) \in \mathbb{C}^{P\times P}$$

and

$$R_\omega = \operatorname{diag}(e^{-2\pi i\omega}, 1, \dots, 1)\, S_Q^{-1} \in \mathbb{C}^{Q\times Q}\,.$$

Lemma 2.1. *For $P, Q \in \mathbb{N}$ and $g \in L^2(\mathbb{R})$, $g \neq 0$, let us define the matrix function $A_g : \mathbb{R}^2 \to \mathbb{C}^{P\times Q}$ by*

$$A_g(x,\omega) := \tfrac{1}{\sqrt{P}}\big(Zg(x + \tfrac{k}{P} - \tfrac{\ell}{Q}, \omega)\big)_{k,\ell=0}^{P-1,Q-1}.$$

Then for a.e. $(x,\omega) \in \mathbb{R}^2$ we have

$$A_g(x + \tfrac{1}{P}, \omega) = L_\omega A_g(x,\omega) \qquad \text{and} \qquad A_g(x - \tfrac{1}{Q}, \omega) = A_g(x,\omega) R_\omega\,. \tag{2.4}$$

In particular, $A_g^* A_g$ is $(\frac{1}{P}, 1)$-periodic and $A_g A_g^*$ is $(\frac{1}{Q}, 1)$-periodic.

If $\Lambda = \frac{1}{Q}\mathbb{Z}\times P\mathbb{Z}$, then (g,Λ) is a Bessel sequence if and only if $Zg \in L^\infty(\mathbb{R}^2)$. In this case, the synthesis operator

$$\mathbf{T}: \quad \ell^2(\mathbb{Z}^2) \to L^2(\mathbb{R}), \quad (c_{n,m})_{n,m\in\mathbb{Z}} \mapsto \sum_{n,m\in\mathbb{Z}} c_{n,m}\, \pi(Q^{-1}n, Pm) g,$$

the frame operator $\mathbf{S}$, and the Gram operator $\mathbf{G}$ of (g,Λ) satisfy

$$\mathbf{T} = (\mathcal{V}Z)^* M_{A_g} \mathcal{U}, \quad \mathbf{S} = (\mathcal{V}Z)^* M_{A_g A_g^*}(\mathcal{V}Z), \quad \text{and} \quad \mathbf{G} = \mathcal{U}^* M_{A_g^* A_g} \mathcal{U}, \tag{2.5}$$

respectively, where $M_{A_g A_g^*}$ (respectively $M_{A_g^* A_g}$ or M_{A_g}) is the matrix multiplication operator (cf. A.1) with respect to $A_g A_g^*$ (resp. $A_g^* A_g$ or A_g) acting on $L^2(R_P; \mathbb{C}^P)$ (resp. $L^2(R_P; \mathbb{C}^Q)$).

If $Zg \in L^\infty(\mathbb{R}^2)$, the following statements hold:

(a) *(g,Λ) is a Riesz sequence if and only if $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A_g(z)) > 0$.*
(b) *(g,Λ) is a frame sequence if and only if $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_1(A_g(z)) > 0$.*
(c) *(g,Λ) is a frame for $L^2(\mathbb{R})$ if and only if $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A_g(z)^*) > 0$.*

Proof. Let $A := A_g$. We have $A(x + \frac{1}{P}, \omega) = P^{-\frac{1}{2}} \cdot \big(Zg(x + \frac{k+1}{P} - \frac{\ell}{Q}, \omega)\big)_{k,\ell=0}^{P-1,Q-1}$, where—due to the quasi-periodicity of Zg—we see that

$$\begin{aligned} &Zg\big(x + \tfrac{k+1}{P} - \tfrac{\ell}{Q}, \omega\big) \\ &= \begin{cases} \sqrt{P} \cdot \big(A(x,\omega)\big)_{k+1,\ell} & \text{if } k < P-1\,, \\ e^{2\pi i\omega} Zg(x - \frac{\ell}{Q}, \omega) = \sqrt{P} \cdot e^{2\pi i\omega} \cdot \big(A(x,\omega)\big)_{0,\ell} & \text{if } k = P-1\,. \end{cases} \end{aligned}$$

In matrix notation, this means precisely that A satisfies the first relation in (2.4), and the $(\frac{1}{P}, 1)$-periodicity of A^*A follows from $L_\omega^* L_\omega = \mathrm{Id}_{\mathbb{C}^P}$ and from $A(x,\omega+1) = A(x,\omega)$. The second relation in (2.4) can be proved similarly and shows that AA^* is $(\frac{1}{Q}, 1)$-periodic.

Let $\mathbf{T}_0$ denote the pre-synthesis operator of (g,Λ), that is,

$$\mathbf{T}_0: \quad \ell_0(\mathbb{Z}^2) \to L^2(\mathbb{R}), \quad \mathbf{T}_0(c_{m,n})_{m,n\in\mathbb{Z}} := \sum_{m,n\in\mathbb{Z}} c_{m,n}\, \pi(\tfrac{m}{Q}, nP) g\,,$$

where $\ell_0(\mathbb{Z}^2)$ is the space of all elements of $\ell^2(\mathbb{Z}^2)$ with only finitely many non-zero entries. For $c \in \ell_0(\mathbb{Z}^2)$, the properties of the Zak transform listed after Equation (2.2) show that

$$(Z\,\mathbf{T}_0\, c)(x,\omega) = \sum_{m,n\in\mathbb{Z}} c_{m,n}\, Z[\pi(\tfrac{m}{Q}, nP)\, g](x,\omega)$$

$$= P^{-1/2} \cdot \sum_{\ell=0}^{Q-1} h_\ell(x,\omega)\, Zg(x - \tfrac{\ell}{Q}, \omega) = \langle A(x,\omega)h(x,\omega), e_0\rangle_{\mathbb{C}^P},$$

where $h_\ell(x,\omega) := P^{1/2} \sum_{s,n\in\mathbb{Z}} c_{sQ+\ell,n}\, e^{2\pi i(nPx - s\omega)}$ and $h := (h_\ell)_{\ell=0}^{Q-1} = \mathcal{U}c$ with $\mathcal{U}$ defined in Equation (2.3). Since h is $(\frac{1}{P}, 1)$-periodic, we obtain for $k \in \{0, \dots, P-1\}$ that

$$(Z\,\mathbf{T}_0\, c)(x + \tfrac{k}{P}, \omega) = \langle A(x + \tfrac{k}{P}, \omega)h(x,\omega), e_0\rangle_{\mathbb{C}^P} = \langle A(x,\omega)h(x,\omega), e_k\rangle_{\mathbb{C}^P}\,.$$

Here, we used the identity $A(x + \frac{1}{P}, \omega) = L_\omega A(x,\omega)$ from the beginning of the proof to get

$$\langle A(x + \tfrac{k}{P}, \omega)h(x,\omega), e_0\rangle = \langle L_\omega^k A(x,\omega)h(x,\omega), e_0\rangle = \langle A(x,\omega)h(x,\omega), (L_\omega^*)^k e_0\rangle\,,$$

where a straightforward induction shows $(L_\omega^*)^k e_0 = \big(\operatorname{diag}(e^{-2\pi i\omega}, 1, \dots, 1)\, S_P^*\big)^k e_0 = e_k$ for $k = 0, \dots, P-1$. With the operator $\mathcal{V}$ defined in Equation (2.3), we have thus shown

$$(\mathcal{V}Z\mathbf{T}_0\, c)(x,\omega) = A(x,\omega)h(x,\omega); \quad \text{that is,} \quad \mathcal{V}Z\mathbf{T}_0 = M_A\,\mathcal{U}|_{\ell_0(\mathbb{Z}^2)}\,.$$

Since the operators $\mathcal{V}, Z, \mathcal{U}$ are unitary, this shows that $\mathbf{T}_0$ is bounded if and only if M_A is bounded, that is, if and only if each entry of A is essentially bounded (on R_P), which—by quasi-periodicity—exactly means that $Zg \in L^\infty(\mathbb{R}^2)$. In particular, this shows that (g,Λ) is a Bessel sequence if and only if $\mathbf{T}_0$ is bounded, if and only if $Zg \in L^\infty(\mathbb{R}^2)$.

Let us assume for the rest of this proof that $Zg \in L^\infty(\mathbb{R}^2)$. Then $\mathcal{V}Z\mathbf{T} = M_A\mathcal{U}$, where $\mathbf{T} = \overline{\mathbf{T}_0} = (\mathcal{V}Z)^* M_A\mathcal{U}$ is the synthesis operator of (g,Λ). Clearly, $M_A^* = M_{A^*}$ is the (bounded) multiplication operator with A^*; thus $M_A^* M_A = M_{A^*A}$ and $M_A M_A^* = M_{AA^*}$. Since $\mathbf{S} = \mathbf{T}\mathbf{T}^*$ and $\mathbf{G} = \mathbf{T}^*\mathbf{T}$, this proves (2.5).

By definition, (g,Λ) is a Riesz sequence if and only if the synthesis operator $\mathbf{T}$ is bounded below. Lemma A.3 shows that this holds if and only if $\mathbf{G} = \mathbf{T}^*\mathbf{T}$ is boundedly invertible, that is, if and only if $0 \in \varrho(\mathbf{G})$. Similarly, (g,Λ) is a frame for $L^2(\mathbb{R})$ if and only if $0 \in \varrho(\mathbf{S})$. Likewise, (g,Λ) is a frame sequence if and only if $(0,\varepsilon_0] \subset \varrho(\mathbf{G})$ for some $\varepsilon_0 > 0$ (see Lemmas A.2 and A.4). Hence, (g,Λ) is a Riesz sequence if and only if $0 \in \varrho(M_{A^*A})$, a frame sequence if and only if $(0,\varepsilon_0] \subset \varrho(M_{A^*A})$ for some $\varepsilon_0 > 0$, and a frame for $L^2(\mathbb{R})$ if and only if $0 \in \varrho(M_{AA^*})$.

The statements (a)–(c) now follow from Lemma A.1 (ii) and (iii). Here, it is used for properties (a) and (b) that $\sigma_i(A_g(z))$ only depends on $A_g^*(z)A_g(z)$, which is $(P^{-1}, 1)$-periodic, so that $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_i(A_g(z)) = \operatorname{essinf}_{z\in R_p} \sigma_i(A_g(z))$ for $i \in \{1,2\}$. Finally, for property (c), it is used that if (g,Λ) is a frame for $L^2(\mathbb{R})$, then $P/Q \le 1$ (see [12, Corollary 7.5.1]), so that $R_Q \subset R_P$. This implies $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A_g^*(z)) = \operatorname{essinf}_{z\in R_P} \sigma_0(A_g^*(z))$, since $z \mapsto A_g(z)A_g^*(z)$ is $(Q^{-1}, 1)$-periodic. Conversely, if $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A_g^*(z)) > 0$, then we also have $\operatorname{essinf}_{z\in R_P} \sigma_0(A_g^*(z)) > 0$, so that $0 \in \varrho(M_{AA^*})$ by Lemma A.1. □

In the next lemma, we derive a formula for the matrix function $A_{\widetilde{g}}$ associated to the *dual window* $\widetilde{g}$ of the Riesz sequence (g,Λ). This means that—considered on the Gabor space $\mathcal{G}(g,\Lambda)$—the Gabor system $(\widetilde{g},\Lambda)$ is the canonical dual frame to (g,Λ). In the proof of the lemma, we will use that $\widetilde{g} = \mathbf{S}^\dagger g$ satisfies this property, where $\mathbf{S}^\dagger$ is the *pseudo-inverse* (see A.2) of the (pre)-frame operator $\mathbf{S}$ of (g,Λ), which is given by $\mathbf{S}f = \sum_{\lambda\in\Lambda} \langle f, \pi(\lambda)g\rangle\, \pi(\lambda)g$. For completeness, we sketch a proof of this fact. Let $\mathbf{S}_0 := \mathbf{S}|_{\mathcal{G}(g,\Lambda)} : \mathcal{G}(g,\Lambda) \to \mathcal{G}(g,\Lambda)$ denote the restriction of $\mathbf{S}$ to $\mathcal{G}(g,\Lambda)$. Note that $\mathbf{S}_0$ is invertible since (g,Λ) is a frame for $\mathcal{G}(g,\Lambda)$, and that $\mathcal{G}(g,\Lambda) = \operatorname{ran}\mathbf{S} = (\ker\mathbf{S})^\perp$ since $\mathbf{S}$ is self-adjoint. Therefore, the pseudo-inverse of $\mathbf{S}$ is given by $\mathbf{S}^\dagger = \mathbf{S}_0^{-1}\mathbb{P}$, where $\mathbb{P}$ denotes

the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda)$. Hence, $g_0 := \mathbf{S}^\dagger g = \mathbf{S}_0^{-1} g$. Finally, a straightforward but tedious computation shows that $\pi(\lambda)\mathbf{S} = \mathbf{S}\pi(\lambda)$ for all $\lambda \in \Lambda$, which also implies that $\pi(\lambda)\mathcal{G}(g,\Lambda) = \mathcal{G}(g,\Lambda)$. Therefore, $\mathbf{S}_0[\pi(\lambda)g_0] = \pi(\lambda)\mathbf{S}_0 g_0 = \pi(\lambda)g$, showing that $\big(\pi(\lambda)g_0\big)_{\lambda\in\Lambda} = \big(\mathbf{S}_0^{-1}(\pi(\lambda)g)\big)_{\lambda\in\Lambda}$ is indeed the canonical dual frame of (g,Λ).

With this preparation, we can now prove the announced lemma.

Lemma 2.2. *Let $g \in L^2(\mathbb{R})$, $P, Q \in \mathbb{N}$, $\Lambda = \frac{1}{Q}\mathbb{Z}\times P\mathbb{Z}$, and assume that (g,Λ) is a Riesz sequence. Let $\widetilde{g}$ be the dual window of (g,Λ) and $G := Zg$, $\widetilde{G} := Z\widetilde{g}$, $A := A_g$, and $\widetilde{A} := A_{\widetilde{g}}$, with A_g and $A_{\widetilde{g}}$ as in Lemma 2.1. Then*

$$\widetilde{A} = A(A^*A)^{-1} \quad \text{almost everywhere on } \mathbb{R}^2. \tag{2.6}$$

Moreover, for arbitrary $\mu = (u,\eta) \in \mathbb{R}^2$ we have

$$\operatorname{dist}^2\big(\pi(\mu)g, \mathcal{G}(g,\Lambda)\big) = \|g\|_{L^2}^2 - \int_0^1 \int_0^{1/P} \|H_\mu(x,\omega)\, e_0\|_{\mathbb{C}^P}^2 \, dx\, d\omega\,,$$

where $e_0 = (1,0,\dots,0)^T \in \mathbb{C}^Q$,

$$H_\mu(z) = P^{1/2} \cdot A(z) \cdot \big(A(z)^*A(z)\big)^{-1} \cdot A(z)^* \cdot e^{2\pi i \eta D_P} \cdot A(z-\mu) \in \mathbb{C}^{P\times Q}\,,$$

and $D_P = \operatorname{diag}(k/P)_{k=0}^{P-1}$, with the notation $e^{2\pi i\eta D_P} := \operatorname{diag}\big((e^{2\pi i\eta k/P})_{k=0}^{P-1}\big)$.

Proof. In this proof we shall make use of the notion of the *pseudo-inverse* $T^\dagger$ of an operator T with closed range. For the definition of this notion and a review of some of its properties, we refer to A.2.

Let $\mathbf{S}$ be the frame operator of (g,Λ). Then the range of $\mathbf{S}$ is $\operatorname{ran}\mathbf{S} = \mathcal{G}(g,\Lambda)$, which is closed in $L^2(\mathbb{R})$. Hence, by Lemma A.6 we have $\mathbf{S}^\dagger = \varphi(\mathbf{S})$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is defined by $\varphi(0) = 0$ and $\varphi(t) = 1/t$ for $t \neq 0$. As seen before the statement of the lemma, $\widetilde{g} = \mathbf{S}^\dagger g = \varphi(\mathbf{S})g$. Furthermore, Lemma A.6 shows $\varphi(A(z)A(z)^*) = (A(z)A(z)^*)^\dagger$ for every $z \in R_P$. Hence, an application of Equation (2.5) and of Lemma A.1 (iv) shows that

$$\widetilde{G} = Z\widetilde{g} = Z[\varphi(\mathbf{S})g] = Z(\mathcal{V}Z)^*\varphi(M_{AA^*})(\mathcal{V}Z)g = \mathcal{V}^* M_{\varphi(AA^*)}\mathcal{V}G,$$

with $\mathcal{V}$ as defined in Equation (2.3). Therefore,

$$\mathcal{V}\widetilde{G} = (AA^*)^\dagger(\mathcal{V}G) \qquad \text{a.e. on } R_P. \tag{2.7}$$

In order to extend this relation to $\mathbb{R}^2$, define $(\widetilde{\mathcal{V}}f)(x,\omega) := \big(f(x+\frac{k}{P},\omega)\big)_{k=0}^{P-1}$ for $f : \mathbb{R}^2 \to \mathbb{C}$ and $(x,\omega) \in \mathbb{R}^2$. Let $z = (x,\omega) \in R_P$ be arbitrary, and set $z_{n,k} = (x+\frac{n+k}{P},\omega)$ for $k \in \{0,\dots,P-1\}$ and $n \in \mathbb{Z}$. Using Equation (2.4), we see that

$$\begin{aligned}\widetilde{G}(z_{n,k}) = (\,Z\,\widetilde{g}\,)\big(x+\tfrac{n+k}{P},\omega\big) &= \sqrt{P}\cdot\Big(A_{\widetilde{g}}\big(x+\tfrac{n}{P},\omega\big)\Big)_{k,0} \\ &= \sqrt{P}\cdot\Big(L_\omega^n\, A_{\widetilde{g}}(x,\omega)\Big)_{k,0} = \Big(L_\omega^n\,\big(\widetilde{G}(z_{0,\ell})\big)_{\ell=0}^{P-1}\Big)_k\,.\end{aligned}$$

Similarly, Equation (2.4) shows that $\big(G(z_{n,k})\big)_{k=0}^{P-1} = L_\omega^n\big(G(z_{0,k})\big)_{k=0}^{P-1}$. Thus, we get for $(x,\omega) \in R_P$ and $n \in \mathbb{Z}$ that

$$\begin{aligned}(\widetilde{\mathcal{V}}\widetilde{G})(x+\tfrac{n}{P},\omega) &= \Big(\widetilde{G}(z_{n,k})\Big)_{k=0}^{P-1} = L_\omega^n\Big[\,\Big(\widetilde{G}(z_{0,k})\Big)_{k=0}^{P-1}\,\Big] = L_\omega^n\big([\mathcal{V}\widetilde{G}](x,\omega)\big) \\ &\overset{\text{(Eq. (2.7))}}{=} L_\omega^n\big[A(x,\omega)A(x,\omega)^*\big]^\dagger\big(G(z_{0,k})\big)_{k=0}^{P-1} \\ &\overset{\text{(Eq. (2.4))}}{=} L_\omega^n\,\big[L_\omega^{-n}A(x+\tfrac{n}{P},\omega)A(x+\tfrac{n}{P},\omega)^*L_\omega^n\big]^\dagger\, L_\omega^{-n}\big(G(z_{n,k})\big)_{k=0}^{P-1}\end{aligned}$$

$$\begin{aligned}(\text{Corollary } A.7) &= \big[A(x+\tfrac{n}{P},\omega)A(x+\tfrac{n}{P},\omega)^*\big]^\dagger \big(G(z_{n,k})\big)_{k=0}^{P-1}\\ &= \big[A(x+\tfrac{n}{P},\omega)A(x+\tfrac{n}{P},\omega)^*\big]^\dagger (\widetilde{\mathcal{V}}G)(x+\tfrac{n}{P},\omega)\,.\end{aligned}$$

In combination with the 1-periodicity in the second variable of all involved functions, this implies

$$\widetilde{\mathcal{V}}\widetilde{G} = (AA^*)^\dagger(\widetilde{\mathcal{V}}G) \qquad \text{a.e. on } \mathbb{R}^2.$$

Since $(\widetilde{\mathcal{V}}G)(x-\frac{\ell}{Q},\omega)$ is the ℓ-th column of the matrix $\sqrt{P}\cdot A(x,\omega)$, since $(\widetilde{\mathcal{V}}\widetilde{G})(x-\frac{\ell}{Q},\omega)$ is the ℓ-th column of $\sqrt{P}\,\widetilde{A}(x,\omega)$, and because AA^* is $(\frac{1}{Q},1)$-periodic, we obtain the identity $\widetilde{A} = (AA^*)^\dagger A = A(A^*A)^\dagger = A(A^*A)^{-1}$, see Lemma A.5 (iv). Here, we used that A^*A is invertible almost everywhere by Lemma 2.1 (a). We have thus proved Equation (2.6).

Now, denote the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda) = \operatorname{ran}\mathbf{S}$ by $\mathbb{P}$. Then, for any $\mu=(u,\eta)\in\mathbb{R}^2$ we have

$$\operatorname{dist}^2\big(\pi(\mu)g,\mathcal{G}(g,\Lambda)\big) = \|(I-\mathbb{P})\pi(\mu)g\|_{L^2}^2 = \|g\|_{L^2}^2 - \|\mathbb{P}\,\pi(\mu)g\|_{L^2}^2\,.$$

Next, Lemmas A.6 and A.1 show $M^\dagger_{A_gA_g^*} = \varphi(M_{A_gA_g^*}) = M_{\varphi(A_gA_g^*)} = M_{(A_gA_g^*)^\dagger}$, which implies $\mathbf{S}^\dagger = (\mathcal{V}Z)^* M^\dagger_{A_gA_g^*}\mathcal{V}Z = (\mathcal{V}Z)^* M_{(A_gA_g^*)^\dagger}\mathcal{V}Z$ thanks to Equation (2.5) and Corollary A.7. Now, Lemma A.5 shows $\mathbb{P} = \mathbf{S}\mathbf{S}^\dagger$. Hence, Equations (2.5) and (A.3) show

$$\mathbb{P} = (\mathcal{V}Z)^* M_{AA^*}M_{(AA^*)^\dagger}(\mathcal{V}Z) = (\mathcal{V}Z)^* M_{(AA^*)(AA^*)^\dagger}(\mathcal{V}Z) = (\mathcal{V}Z)^* M_{P_{\operatorname{ran}A}}(\mathcal{V}Z)$$

and $P_{\operatorname{ran}A} = P_{\operatorname{ran}(AA^*)} = (AA^*)\,(AA^*)^\dagger = A(A^*A)^\dagger A^* = A(A^*A)^{-1}A^*$. For arbitrary $f\in L^2(\mathbb{R})$, we thus see that

$$\|\mathbb{P}f\|_{L^2}^2 = \|M_{P_{\operatorname{ran}A}}\mathcal{V}Zf\|^2_{L^2(R_P,\mathbb{C}^P)} = \int_0^1\int_0^{1/P} \|A(A^*A)^{-1}A^*\mathcal{V}Zf\|^2_{\mathbb{C}^P}\,dx\,d\omega\,.$$

Finally, since $Z[\pi(\mu)g](x,\omega) = e^{2\pi i\eta x}Zg(x-u,\omega-\eta)$ for $\mu=(u,\eta)$, we see

$$\begin{aligned}(\mathcal{V}Z[\pi(\mu)g])(x,\omega) &= \big(e^{2\pi i\eta(x+\frac{k}{P})}Zg(x+\tfrac{k}{P}-u,\omega-\eta)\big)_{k=0}^{P-1}\\ &= e^{2\pi i\eta x}e^{2\pi i\eta D_P}\big(Zg(x+\tfrac{k}{P}-u,\omega-\eta)\big)_{k=0}^{P-1}\\ &= e^{2\pi i\eta x}\,P^{1/2}\,e^{2\pi i\eta D_P}A(x-u,\omega-\eta)\,e_0\,.\end{aligned}$$

Now, the claim follows from $|e^{2\pi i\eta x}|=1$. □

In proving the next result, we crucially use that if $\lambda=(\alpha,\beta)\in\mathbb{R}^2$ and $\mu=(a,b)\in\mathbb{R}^2$, then $\pi(\lambda)\pi(\mu)f = e^{-2\pi i\alpha b}\pi(\lambda+\mu)f$, as can be verified by a direct calculation. In particular, this implies $\|T\pi(\lambda)\pi(\mu)f\|_{L^2} = \|T\pi(\lambda+\mu)f\|_{L^2}$ for any linear operator $T: L^2(\mathbb{R})\to L^2(\mathbb{R})$.

Lemma 2.3. *Let $g\in L^2(\mathbb{R})$ and let $\Lambda\subset\mathbb{R}^2$ be a lattice. If $\mathbb{P}$ denote the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda)$. Then $\mathbb{P}$ commutes with the operators $\pi(\lambda)$, $\lambda\in\Lambda$. In particular*

$$\operatorname{dist}\big(\pi(\mu+\lambda)f,\mathcal{G}(g,\Lambda)\big) = \operatorname{dist}\big(\pi(\mu)f,\mathcal{G}(g,\Lambda)\big) \qquad \forall\,\mu\in\mathbb{R}^2, f\in L^2()\mathbb{R}, \lambda\in\Lambda.$$

Proof. Let $\mathcal{G} := \mathcal{G}(g,\Lambda)$. Lemma A.16 shows $\pi(-\lambda)\mathcal{G}\subset\mathcal{G}$ for all $\lambda\in\Lambda$. This implies $\pi(\lambda)\mathcal{G}^\perp\subset\mathcal{G}^\perp$: Indeed, $\langle\pi(\lambda)f,h\rangle = e^{2\pi i\lambda_1\lambda_2}\langle f,\pi(-\lambda)h\rangle = 0$ for $f\in\mathcal{G}^\perp$ and $h\in\mathcal{G}$. Now, if $f\in L^2(\mathbb{R})$, then we can write $f=f_1+f_2$ with $f_1\in\mathcal{G}$ and $f_2\in\mathcal{G}^\perp$; hence, $\pi(\lambda)f = \pi(\lambda)f_1+\pi(\lambda)f_2$ with $\pi(\lambda)f_1\in\mathcal{G}$ and $\pi(\lambda)f_2\in\mathcal{G}^\perp$, which implies $\mathbb{P}[\pi(\lambda)f] = \pi(\lambda)f_1 = \pi(\lambda)[\mathbb{P}f]$.

As to the "in particular"-part, we observe for $\mu \in \mathbb{R}^2$ and $\lambda \in \Lambda$ that

$$\begin{aligned}\|(I-\mathbb{P})\pi(\mu+\lambda)f\|_{L^2} &= \|(I-\mathbb{P})\pi(\lambda)\pi(\mu)f\|_{L^2} \\ &= \|\pi(\lambda)(I-\mathbb{P})\pi(\mu)f\|_{L^2} = \|(I-\mathbb{P})\pi(\mu)f\|_{L^2}\,.\end{aligned}$$

The claim now follows by noting $\operatorname{dist}(f,\mathcal{G}) = \|(I-\mathbb{P})f\|_{L^2}$ for $f \in L^2(\mathbb{R})$. □

2.2. Describing the regularity of g via the Zak transform

The following lemma is probably folklore. However, since we could not find any reference for it (one direction is proved in [8, Proof of Thm. 2.3]), we give a full proof here. Recall that $\mathbb{H}^1(\mathbb{R}) = \{f \in H^1(\mathbb{R})\colon \widehat{f} \in H^1(\mathbb{R})\}$.

Lemma 2.4. *Let $g \in L^2(\mathbb{R})$. Then $g \in \mathbb{H}^1(\mathbb{R})$ if and only if $Zg \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$. In this case, the weak derivatives of Zg are given by*

$$\partial_1 Zg = Z(g') \quad \text{and} \quad \partial_2 Zg(x,\omega) = -2\pi i\big([Z(Xg)](x,\omega) - x\cdot Zg(x,\omega)\big)\,. \tag{2.8}$$

Proof. "⇒:" Assume that $g \in \mathbb{H}^1(\mathbb{R})$ and let $V \subset \mathbb{R}^2$ be nonempty, open, and bounded. Let us first assume that $g \in C_c^\infty(\mathbb{R})$ (such a function of course is in $\mathbb{H}^1(\mathbb{R})$). Recalling the definition (2.2) of the Zak transform, we see that on V, Zg is defined by a *finite* sum (hence $Zg \in C^\infty(V)$), and the first relation in (2.8) is easily verified. For the second relation, we note

$$\begin{aligned}\partial_2[Zg(x,\omega)] &= \sum_{k=-\infty}^{\infty} \partial_\omega[e^{2\pi i k\omega}\, g(x-k)] = \sum_{k=-\infty}^{\infty} 2\pi i k\, e^{2\pi i k\omega}\, g(x-k) \\ &= \sum_{k=-\infty}^{\infty} e^{2\pi i k\omega}[2\pi i x\, g(x-k) - 2\pi i (Xg)(x-k)] \\ &= 2\pi i\cdot\big[x\cdot Zg(x,\omega) - Z[Xg](x,\omega)\big]\,,\end{aligned}$$

as claimed in (2.8).

Now, let $g \in \mathbb{H}^1(\mathbb{R})$ be arbitrary. Since $C_c^\infty(\mathbb{R})$ is dense in $H^1(\mathbb{R})$ (see for instance [2, Section E10.8]), we find a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R})$ which converges to g in $H^1(\mathbb{R})$, that is, $\varphi_n \to g$ and $\varphi_n' \to g'$ in $L^2(\mathbb{R})$. For $\phi \in C_c^\infty(V)$ we then have (with $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_{L^2(V)}$)

$$\begin{aligned}&|\langle Zg, \partial_1\phi\rangle + \langle Zg', \phi\rangle| \\ &\qquad \le |\langle Z(g-\varphi_n), \partial_1\phi\rangle| + |\langle Z\varphi_n, \partial_1\phi\rangle + \langle Z\varphi_n', \phi\rangle| + |\langle Z(g'-\varphi_n'), \phi\rangle|\,.\end{aligned}$$

The middle term vanishes by partial integration and since $\partial_1(Z\varphi_n) = Z(\varphi_n')$; the other two terms tend to zero as $n \to \infty$. Hence, $\langle Zg, \partial_1\phi\rangle = -\langle Zg', \phi\rangle$.

The relation $\langle Zg, \partial_2\phi\rangle = 2\pi i\langle Z(Xg) - XZg, \phi\rangle$ is proven similarly, by noting that since $(1+|X|)\, g \in L^2(\mathbb{R})$, one can find[1] a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R})$ satisfying $\|(1+|X|)\,(g-\varphi_n)\|_{L^2} \to 0$, whence $\varphi_n \to g$ in L^2 and $X\varphi_n \to Xg$ in L^2, and therefore $Z(X\varphi_n) \to Z(Xg)$ and $X\,Z\varphi_n \to X\,Zg$ with convergence in $L^2_{\mathrm{loc}}(\mathbb{R}^2)$.

Because of $Zg' \in L^2(V)$ and $Z(Xg) - XZg \in L^2(V)$, this proves that $Zg \in H^1(V)$ and that (2.8) holds on V. Since $V \subset \mathbb{R}^2$ was an arbitrary non-empty, open, bounded set, we have proved one implication.

[1] Indeed, given $\varepsilon > 0$, there is a compactly supported h such that $\|(1+|X|)\,(g-h)\|_{L^2} < \varepsilon$, say $\operatorname{supp} h \subset [-N,N]$. Pick $\varphi \in C_c^\infty(\mathbb{R})$ satisfying $\|\varphi - h\|_{L^2} < \varepsilon/(1+2N)$ and $\operatorname{supp}\varphi \subset [-2N, 2N]$, whence $\|(1+|X|)\,(\varphi-h)\|_{L^2} \le (1+2N)\,\|\varphi-h\|_{L^2} < \varepsilon$, so that finally $\|(1+|X|)\,(g-\varphi)\|_{L^2} < 2\varepsilon$.

"$\Leftarrow$:" Assume that $G := Zg \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$. Lemma A.14 shows that, after changing G on a null-set, we can assume that $G(x,\cdot)$ is locally absolutely continuous on $\mathbb{R}$ with derivative $(\partial_2 G)(x,\cdot) \in L^2_{\mathrm{loc}}(\mathbb{R})$ for almost every $x \in \mathbb{R}$ and *simultaneously* that $G(\cdot,\omega)$ is locally absolutely continuous on $\mathbb{R}$ with derivative $(\partial_1 G)(\cdot,\omega) \in L^2_{\mathrm{loc}}(\mathbb{R})$ for almost every $\omega \in \mathbb{R}$.

According to the properties of the Zak transform, $g(x) = \int_0^1 G(x,\omega)\,d\omega$ for almost all $x \in \mathbb{R}$; see the list of properties below Equation (2.2). Let us fix one $x_0 \in \mathbb{R}$ for which this is true. Hence, for almost all $x \in \mathbb{R}$ we have

$$\begin{aligned} g(x) &= \int_0^1 G(x,\omega)\,d\omega = \int_0^1 \left(G(x_0,\omega) + \int_{x_0}^x \partial_1 G(t,\omega)\,dt \right) d\omega \\ &= g(x_0) + \int_{x_0}^x \left(\int_0^1 \partial_1 G(t,\omega)\,d\omega \right) dt = g(x_0) + \int_{x_0}^x \phi(t)\,dt\,, \end{aligned}$$

where $\phi(t) := \int_0^1 \partial_1 G(t,\omega)\,d\omega$. Note that $\phi \in L^1_{\mathrm{loc}}(\mathbb{R})$ since $\partial_1 G \in L^2_{\mathrm{loc}}(\mathbb{R}^2)$. Hence, possibly after redefining g on a set of measure zero, g is locally absolutely continuous on $\mathbb{R}$. To see that actually $\phi \in L^2(\mathbb{R})$ (and hence $g \in H^1(\mathbb{R})$), recall from the properties of the Zak transform that $G(t+n,\omega) = e^{2\pi i n\omega} G(t,\omega)$ for almost all $(t,\omega) \in \mathbb{R}^2$. Hence,

$$\|\phi\|_{L^2}^2 = \sum_{n\in\mathbb{Z}} \int_0^1 \left| \int_0^1 \partial_1 G(t+n,\omega)\,d\omega \right|^2 dt = \int_0^1 \sum_{n\in\mathbb{Z}} \left| \int_0^1 e^{2\pi i n\omega} \partial_1 G(t,\omega)\,d\omega \right|^2 dt.$$

Now, set $g_t(\omega) := \partial_1 G(t,\omega)$ (which is in $L^2((0,1))$ for a.e. $t \in \mathbb{R}$). Then

$$\int_{\mathbb{R}} |\phi(t)|^2\,dt = \int_0^1 \sum_{n\in\mathbb{Z}} |\widehat{g_t}(n)|^2\,dt = \int_0^1 \int_0^1 |g_t(\omega)|^2\,d\omega\,dt = \|\partial_1 G\|^2_{L^2([0,1]^2)}.$$

Hence, $g \in H^1(\mathbb{R})$ with $g'(x) = \int_0^1 \partial_1 G(x,\omega)\,d\omega$.

To see that also $\widehat{g} \in H^1(\mathbb{R})$, define $F : \mathbb{R}^2 \to \mathbb{C}, (x,\omega) \mapsto e^{-2\pi i x\omega} G(x,\omega)$. Since $G_x := G(x,\cdot)$ is locally absolutely continuous for almost all $x \in \mathbb{R}$, the product rule for Sobolev functions (see for instance [2, Section 4.25]) shows that also $F_x := F(x,\cdot)$ satisfies this property. Moreover, the product rule also shows for almost all $x \in \mathbb{R}$ that we have

$$\begin{aligned} F_x'(\omega) &= e^{-2\pi i x\omega}(-2\pi i x G_x(\omega) + G_x'(\omega)) \\ &= e^{-2\pi i x\omega}\big(-2\pi i x\, G(x,\omega) + (\partial_2 G)(x,\omega)\big) =: H(x,\omega) \end{aligned}$$

for almost all $\omega \in \mathbb{R}$. Note that $H \in L^2_{\mathrm{loc}}(\mathbb{R}^2)$, since $G \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$. This easily implies that the function $\psi : \mathbb{R} \to \mathbb{C}, \omega \mapsto \int_0^1 H(x,\omega)\,dx$, is almost everywhere well-defined and satisfies $\psi \in L^1_{\mathrm{loc}}(\mathbb{R})$.

Next, recall the inversion formula of the Zak transform (see the list of properties below Equation (2.2)), stating $\widehat{g}(\omega) = \int_0^1 e^{-2\pi i x\omega} G(x,\omega)\,dx = \int_0^1 F(x,\omega)\,dx$ for almost all $\omega \in \mathbb{R}$. Fix some $\omega_0 \in \mathbb{R}$ for which this holds, and note for almost all $\omega \in \mathbb{R}$ that

$$\widehat{g}(\omega) = \int_0^1 F_x(\omega)\,dx = \int_0^1 \Big(F_x(\omega_0) + \int_{\omega_0}^{\omega} F_x'(\gamma)\,d\gamma \Big)\,dx = \widehat{g}(\omega_0) + \int_{\omega_0}^{\omega} \psi(\gamma)\,d\gamma\,.$$

Hence—possibly after changing $\widehat{g}$ on a null-set—we see that $\widehat{g}$ is locally absolutely continuous, with $\widehat{g}'(\omega) = \psi(\omega)$, so that it remains to show $\psi \in L^2(\mathbb{R})$.

To see this, note for arbitrary $n \in \mathbb{Z}$ that $G_x(\omega+n) = Zg(x,\omega+n) = Zg(x,\omega) = G_x(\omega)$, and hence also $G_x'(\omega+n) = G_x'(\omega)$, which finally implies for almost all $x \in \mathbb{R}$ that

$H(x,\omega+n) = e^{-2\pi i n x} H(x,\omega)$ for almost all $\omega \in \mathbb{R}$. Therefore, we see for any $n \in \mathbb{Z}$ that $\psi(\omega+n) = \int_0^1 e^{-2\pi i n x} H(x,\omega)\,dx = \widehat{H_\omega}(n)$, where H_ω is defined by $H_\omega(x) := H(x,\omega)$ for $x \in [0,1]$, so that $H_\omega \in L^2([0,1])$ for almost all $\omega \in \mathbb{R}$. Thus, we finally arrive at

$$\begin{aligned}\int_{\mathbb{R}} |\psi(\omega)|^2\,d\omega &= \sum_{n\in\mathbb{Z}} \int_0^1 |\psi(\omega+n)|^2\,d\omega = \int_0^1 \sum_{n\in\mathbb{Z}} |\widehat{H_\omega}(n)|^2\,d\omega \\ &= \int_0^1 \|H_\omega\|_{L^2}^2\,d\omega = \int_0^1 \int_0^1 |H(x,\omega)|^2\,dx\,d\omega < \infty\,. \qquad \square\end{aligned}$$

2.3. Symplectic operators and the regularity of the dual window

In this subsection, we show that if (g,Λ) is a Riesz sequence with $g \in \mathbb{H}^1(\mathbb{R})$, then the canonical dual window $\widetilde{g}$ belongs to $\mathbb{H}^1(\mathbb{R})$ as well.

For proving this—and also several other results—we shall make use of so-called *symplectic operators* to generalize statements involving lattices of the form $Q^{-1}\mathbb{Z}\times P\mathbb{Z}$, $P,Q \in \mathbb{N}$, to general lattices of rational density. To explain this, let $\Lambda \subset \mathbb{R}^2$ be such a general lattice of rational density. Then there exists a matrix $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ such that $B\Lambda = Q^{-1}\mathbb{Z}\times P\mathbb{Z}$ with $P,Q \in \mathbb{N}$ co-prime. Indeed, we have $\Lambda = A\mathbb{Z}^2$ for some $A \in \mathbb{R}^{2\times 2}$ with $\det A \in \mathbb{Q}\backslash\{0\}$, that is, $|\det A| = P/Q$ for some co-prime $P,Q \in \mathbb{N}$. Now define $B_0 := |\det A|^{1/2} \cdot A^{-1}$ if $\det A > 0$, and if instead $\det A < 0$, then let $B_0 := |\det A|^{1/2} \cdot \operatorname{diag}(-1,1) \cdot A^{-1}$. It is not hard to check that $\det B_0 = 1$, and that $B_0\Lambda = |\det A|^{1/2}\,\mathbb{Z}^2$. Thus, the matrix $B := \operatorname{diag}\left((PQ)^{-1/2}, (PQ)^{1/2}\right) B_0$ satisfies $\det B = 1$ and $B\Lambda = Q^{-1}\mathbb{Z}\times P\mathbb{Z}$.

Next, since $\det B = 1$, we see from [12, Lemma 9.4.1 and Equation (9.39)] that there is a unitary operator $U_B : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ satisfying

$$U_B\,\rho(z) = \rho(Bz)\,U_B \qquad \text{for all } z \in \mathbb{R}^2\,, \tag{2.9}$$

where

$$\rho(a,b)f := e^{-\pi i a b}\pi(a,b)f, \quad \text{for } f \in L^2(\mathbb{R}) \text{ and } a,b \in \mathbb{R}. \tag{2.10}$$

Such an operator U_B is called *symplectic*. As a consequence of Schur's Lemma (see [12, Lemma 9.3.2]), the operator U_B is unique up to multiplication with unimodular constants; thus, we see for $B, B_1, B_2 \in \mathrm{SL}(2,\mathbb{R})$ that

$$U_{B_1B_2} = c_{B_1,B_2}\,U_{B_1}U_{B_2} \qquad \text{and} \qquad U_B^* = c_B \cdot U_{B^{-1}} \tag{2.11}$$

for certain constants $c_{B_1,B_2}, c_B \in \mathbb{C}$ with $|c_{B_1,B_2}| = 1 = |c_B|$.

For us, an important property of symplectic operators is that they leave $\mathbb{H}^1(\mathbb{R})$ invariant. To see this, recall from [4, discussion around Equation (4.5)] that each matrix $B \in \mathrm{SL}(2,\mathbb{R})$ can be written as a product of matrices of the form

$$B_0 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B_\alpha^{(1)} := \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix}, \quad \text{and} \quad B_\beta^{(2)} := \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}$$

with $\alpha,\beta \in \mathbb{R}\backslash\{0\}$. Furthermore, if we define operators $D_\alpha : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ and $C_\beta : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by $D_\alpha f(x) := |\alpha|^{1/2} \cdot f(\alpha x)$ and $C_\beta f(x) = e^{\pi i \beta x^2} \cdot f(x)$, then a direct computation shows that the choices $U_{B_\alpha^{(1)}} := D_\alpha$ and $U_{B_\beta^{(2)}} := C_\beta$ make (2.9) valid. Likewise, if we let $U_{B_0} := \mathcal{F}$ be the Fourier transform, then (2.9) is satisfied as well.

Thus, in view of (2.11), it suffices to show that $\mathbb{H}^1(\mathbb{R})$ is invariant under the operators $\mathcal{F}$, D_α, and C_β. For $\mathcal{F}$ and D_α, this is trivial. Finally, for C_β recall that $f \in L^2(\mathbb{R})$

is in $\mathbb{H}^1(\mathbb{R})$ if and only if $Xf \in L^2(\mathbb{R})$ and if f is locally absolutely continuous with $f' \in L^2(\mathbb{R})$. As a consequence of the product rule for Sobolev functions (see for instance [2, Section 4.25]), it follows that if $g \in \mathbb{H}^1(\mathbb{R})$, then $C_\beta g$ is locally absolutely continuous, with

$$(C_\beta\, g)'(x) = 2\pi i\beta x \cdot e^{\pi i\beta x^2} \cdot g(x) + e^{\pi i\beta x^2} \cdot g'(x) \in L^2(\mathbb{R})\,.$$

Since $XC_\beta\, g \in L^2(\mathbb{R})$ holds trivially, we have $C_\beta\, g \in \mathbb{H}^1(\mathbb{R})$, as desired.

To see an application of symplectic operators, note that if Λ is a lattice of rational density with $B\Lambda = Q^{-1}\mathbb{Z}\times P\mathbb{Z}$ for some $B \in \mathrm{SL}(2,\mathbb{R})$, and if $g \in L^2(\mathbb{R})$ is such that (g,Λ) is a Riesz sequence, one may define

$$g_1 := U_B\, g \qquad \text{and} \qquad \Lambda_1 := B\Lambda = \tfrac{1}{Q}\mathbb{Z}\times P\mathbb{Z}\,.$$

Then (2.9) implies $\pi(B\lambda)g_1 = c_\lambda U_B\, \pi(\lambda)g$, $\lambda \in \Lambda$, where $c_\lambda = c_\lambda(B)$ is a unimodular constant. Hence, (g_1,Λ_1) is a Riesz basis for its closed linear span $\mathcal{G}(g_1,\Lambda_1) = U_B\, \mathcal{G}(g,\Lambda)$. This reduction to the separable lattice Λ_1 will be crucial in the proof of the following proposition.

Proposition 2.5. *Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density such that (g,Λ) is a Riesz sequence. Let $\widetilde{g}$ be the dual window of (g,Λ). Then $g \in \mathbb{H}^1(\mathbb{R})$ if and only if $\widetilde{g} \in \mathbb{H}^1(\mathbb{R})$.*

Proof. Let us first prove the claim for $\Lambda = Q^{-1}\mathbb{Z}\times P\mathbb{Z}$, where $P, Q \in \mathbb{N}$. Assume that $g \in \mathbb{H}^1(\mathbb{R})$. By Lemma 2.4, $Zg \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$. Let us denote by A_g and $A_{\widetilde{g}}$ the matrix functions introduced in Lemma 2.1. Using that lemma, we conclude that each entry of A_g is contained in $L^\infty(\mathbb{R}^2)$ and that there exists $c > 0$ such that $\sigma_0(A_g(z)) \geq c$ for a.e. $z \in \mathbb{R}^2$. Therefore, a combination of Equation (2.6) and Lemmas A.5 and A.10 shows that each entry of $A_{\widetilde{g}} = A_g(A_g^* A_g)^{-1} = (A_g^\dagger)^*$ is contained in $H^1_{\mathrm{loc}}(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$. In view of the definition of $A_{\widetilde{g}}$, this shows that $Z\widetilde{g} \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$, whence Lemma 2.4 implies $\widetilde{g} \in \mathbb{H}^1(\mathbb{R})$. Since $(\widetilde{g},\Lambda)$ is also a Riesz basis for $\mathcal{G}(g,\Lambda) = \mathcal{G}(\widetilde{g},\Lambda)$ with (g,Λ) being the dual Riesz basis, interchanging the roles of g and $\widetilde{g}$ in the above arguments shows that $\widetilde{g} \in \mathbb{H}^1(\mathbb{R})$ implies $g \in \mathbb{H}^1(\mathbb{R})$.

Now, let $\Lambda \subset \mathbb{R}^2$ be an arbitrary lattice of rational density. As seen before Equation (2.9), there is a matrix $B \in \mathrm{SL}(2,\mathbb{R})$ such that $\Lambda_1 := B\Lambda = Q^{-1}\mathbb{Z}\times P\mathbb{Z}$ for certain $P, Q \in \mathbb{N}$. Let $g_1 := U_B\, g$. Then (g_1,Λ_1) is a Riesz basis for $\mathcal{G}(g_1,\Lambda_1) = U_B\, \mathcal{G}(g,\Lambda)$. Furthermore, since $\pi(B\lambda)g_1 = c_\lambda U_B \pi(\lambda)g$ for $\lambda \in \Lambda$, where $|c_\lambda| = 1$, it is not hard to see that the frame operator $\widetilde{\mathbf{S}}$ for (g_1,Λ_1) is given by $\widetilde{\mathbf{S}} = U_B \mathbf{S} U_B^*$, where $\mathbf{S}$ is the frame operator of (g,Λ). Hence, as discussed before Lemma 2.2, the dual window of (g_1,Λ_1) is given by via the pseudo-inverse as $\widetilde{g_1} = \widetilde{\mathbf{S}}^\dagger g_1 = U_B \mathbf{S}^\dagger U_B^* U_B g = U_B \widetilde{g}$, where used that $\widetilde{\mathbf{S}}^\dagger = U_B \mathbf{S}^\dagger U_B^*$ due to Corollary A.7.

Now, suppose that $g \in \mathbb{H}^1(\mathbb{R})$. As seen in the discussion before this proposition, symplectic operators leave $\mathbb{H}^1(\mathbb{R})$ invariant; thus, $g_1 = U_B\, g \in \mathbb{H}^1(\mathbb{R})$. Hence, by what we showed above, we see that $\widetilde{g}_1 \in \mathbb{H}^1(\mathbb{R})$, which implies $\widetilde{g} = U_B^*\, \widetilde{g}_1 = c_B\, U_{B^{-1}}\, \widetilde{g}_1 \in \mathbb{H}^1(\mathbb{R})$. Finally, by interchanging the roles of g and $\widetilde{g}$ we see that $\widetilde{g} \in \mathbb{H}^1(\mathbb{R})$ implies $g \in \mathbb{H}^1(\mathbb{R})$. □

3. Differentiability of the Time-Frequency Map

In this section, we show that for $g \in \mathbb{H}^1(\mathbb{R})$ the map $(a,b) \mapsto e^{2\pi i b x} g(x-a)$ is differentiable at the origin, with the derivative given by $(a,b) \mapsto -ag' + 2\pi i b X g$. In the

proof, we will make use of the following simple estimate. Recall that the sinc function is defined by $\operatorname{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$ for $x \in \mathbb{R}\backslash\{0\}$ and $\operatorname{sinc}(0) := 1$.

Lemma 3.1. *We have* $\left|\frac{\sin(x)}{x} - e^{-ix}\right| \leq |x|$ *for all* $x \in \mathbb{R}\backslash\{0\}$. *Consequently,*

$$|\operatorname{sinc}(x) - e^{-i\pi x}| \leq \min\{2, \pi|x|\} \quad \text{for all } x \in \mathbb{R}\,. \tag{3.1}$$

Proof. The first inequality is equivalent to $|\sin(x) - xe^{-ix}| \leq x^2$ and thus to

$$f(x) := (\sin(x) - x\cos(x))^2 + x^2\sin^2(x) - x^4 \leq 0\,.$$

Since f is even, it suffices to prove $f(x) \leq 0$ for $x > 0$. We have

$$\begin{aligned} f'(x) &= 2x\sin(x)\big(\sin(x) - x\cos(x)\big) + 2x\sin^2(x) + 2x^2\sin(x)\cos(x) - 4x^3 \\ &= 4x\sin^2(x) - 4x^3 = 4x(\sin(x) - x)(\sin(x) + x)\,. \end{aligned}$$

As $\sin(x) < x$ and $\sin(x) + x > 0$ for $x > 0$, we have that $f'(x) < 0$ for $x > 0$. Since $f(0) = 0$, this proves the claim. Equation (3.1) is a direct consequence of the first estimate combined with $|\operatorname{sinc}(x)| \leq 1$ and $|e^{-i\pi x}| \leq 1$. □

For $g \in L^2(\mathbb{R})$ define the map $S_g : \mathbb{R}^2 \to L^2(\mathbb{R})$ by

$$S_g(a,b) := \pi(a,b)g = e^{2\pi ib\,(\cdot)}g(\cdot - a), \quad a, b \in \mathbb{R}.$$

It is well known (see e.g. [7, Lemma 2.9.2]) that S_g is continuous for every $g \in L^2(\mathbb{R})$. Here, we will show that S_g is differentiable if $g \in \mathbb{H}^1(\mathbb{R})$. We will first prove the differentiability of S_g at the origin and then use it to prove the differentiability of S_g at arbitrary points $(u,\eta) \in \mathbb{R}^2$. For the convenience of the reader, we recall the notion of (Fréchet)-differentiability: The map $S_g : \mathbb{R}^2 \to L^2(\mathbb{R})$ is *(Fréchet)-differentiable at* $(u,\eta) \in \mathbb{R}^2$ if there is a bounded linear map $T : \mathbb{R}^2 \to L^2(\mathbb{R})$ satisfying $\frac{|S_g(u+a,\eta+b) - S_g(u,\eta) - T\binom{a}{b}|}{|(a,b)|} \to 0$ as $(a,b) \to 0$. Such a map T is referred to as the derivative of S_g at (u,η), denoted by $S_g'(u,\eta)$; see e.g. [17, Chapter XIII, §2] for more details.

Lemma 3.2. *For any* $g \in \mathbb{H}^1(\mathbb{R})$, *the map* S_g *is (Fréchet)-differentiable at* $(0,0)$ *with*

$$S_g'(0,0)\binom{a}{b} = -ag' + 2\pi ibXg,$$

where X *is the position operator defined formally by* $Xf(x) = xf(x)$.

If $g \in \mathbb{H}^2(\mathbb{R})$ *(that is,* $g \in H^2(\mathbb{R})$ *and* $\widehat{g} \in H^2(\mathbb{R})$*), then*

$$\|S_g(a,b) - g - (-ag' + 2\pi ibXg)\|_{L^2} \;\leq\; C_g \cdot \|(a,b)\|_2^2 \quad \forall\,(a,b) \in \mathbb{R}^2\,, \tag{3.2}$$

where

$$C_g := 3\pi^2 \max\left\{\|X^2 g\|_{L^2}, \|\omega^2\widehat{g}\|_{L^2}, \|Xg'\|_{L^2}\right\}. \tag{3.3}$$

Remark. As shown in Lemma A.11, we indeed have $Xg' \in L^2(\mathbb{R})$ if $g \in \mathbb{H}^2(\mathbb{R})$.

Proof. Let $\Psi_g : \mathbb{R}^2 \to L^2(\mathbb{R})$ be defined by $\Psi_g\binom{a}{b} := -ag' + 2\pi ibXg$; in particular, Ψ_g is linear. We have to prove that

$$\lim_{(a,b)\to(0,0)} \frac{\|S_g(a,b) - g - \Psi_g\binom{a}{b}\|_{L^2}}{\sqrt{a^2+b^2}} = 0. \tag{3.4}$$

To see this, we write

$$\begin{aligned} &[S_g(a,b) - g - \Psi_g\binom{a}{b}](x) \\ &= e^{2\pi ibx}g(x-a) - g(x) + ag'(x) - 2\pi ibxg(x) \\ &= e^{2\pi ibx}\big(g(x-a) - g(x)\big) + (e^{2\pi ibx} - 1 - 2\pi ibx)g(x) + ag'(x) \end{aligned}$$

$$= e^{2\pi ibx}\big(T_a g - g + ag'\big)(x) + \big[e^{2\pi ibx} - 1 - 2\pi ibx\big]g(x) + a\big[1 - e^{2\pi ibx}\big]g'(x)\,. \tag{3.5}$$

To estimate the middle term in (3.5), recall that $\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x} = \frac{e^{\pi ix} - e^{-\pi ix}}{2\pi ix}$ and hence

$$e^{2\pi ibx} - 1 - 2\pi ibx = 2\pi ibx e^{\pi ibx}\big[\operatorname{sinc}(bx) - e^{-\pi ibx}\big]. \tag{3.6}$$

Therefore,

$$\int \left|e^{2\pi ibx} - 1 - 2\pi ibx\right|^2 |g(x)|^2 dx = 4\pi^2 b^2 \int x^2 \left|\operatorname{sinc}(bx) - e^{-\pi ibx}\right|^2 |g(x)|^2 dx. \tag{3.7}$$

Using the estimate (3.1), we find that this expression is not larger than

$$4\pi^4 |b|^3 \int_{|x|<1/\sqrt{|b|}} x^2 |g(x)|^2\, dx \;+\; 16\pi^2 b^2 \int_{|x|\geq 1/\sqrt{|b|}} x^2 |g(x)|^2\, dx\,.$$

Hence, we obtain

$$\begin{aligned} &\left(\frac{\|(e^{2\pi ibx} - 1 - 2\pi ibx)\cdot g\|_{L^2}}{\sqrt{a^2+b^2}}\right)^2 \\ &\leq \frac{1}{b^2}\int \left|e^{2\pi ibx} - 1 - 2\pi ibx\right|^2 |g(x)|^2\, dx \\ &\leq 4\pi^4 |b|\, \|Xg\|_{L^2}^2 + 16\pi^2 \int_{|x|\geq |b|^{-1/2}} x^2\, |g(x)|^2\, dx\,, \end{aligned} \tag{3.8}$$

which tends to zero as $b \to 0$ as a consequence of $Xg \in L^2$ and the dominated convergence theorem.

For the first term in (3.5), observe that Plancherel's theorem yields

$$\|T_a g - g + ag'\|_{L^2}^2 = \int |e^{2\pi i(-a)\omega} - 1 - 2\pi i(-a)\omega|^2 \cdot |\widehat{g}(\omega)|^2\, d\omega\,. \tag{3.9}$$

Thus, using that $\big(\omega \mapsto \omega \cdot \widehat{g}(\omega)\big) \in L^2$, we can conclude from our calculations in (3.8) that

$$0 \leq \left(\frac{\|T_a g - g + a\, g'\|_{L^2}}{\sqrt{a^2+b^2}}\right)^2 \leq \frac{\|T_a g - g + ag'\|_{L^2}^2}{a^2} \xrightarrow[a\to 0]{} 0\,.$$

Finally, using the estimates $|e^{2\pi ibx} - 1| \leq 2\pi|bx|$ and $|e^{2\pi ibx} - 1| \leq 2$, we can treat the last summand in (3.5) as follows:

$$\big\|a \cdot [1 - e^{2\pi ibx}] \cdot g'(x)\big\|_{L^2}^2 \leq 4\pi^2 a^2 |b| \int_{|x|\leq |b|^{-\frac{1}{2}}} |g'(x)|^2\, dx + 4a^2 \int_{|x|\geq |b|^{-\frac{1}{2}}} |g'(x)|^2\, dx.$$

Hence,

$$\left(\frac{\|a \cdot g'(x) \cdot (1 - e^{2\pi ibx})\|_{L^2}}{\sqrt{a^2+b^2}}\right)^2 \;\leq\; 4\pi^2 |b| \cdot \|g'\|_{L^2}^2 + 4 \int_{|x|\geq 1/\sqrt{|b|}} |g'(x)|^2\, dx\,,$$

which tends to zero as $(a,b) \to (0,0)$, again as a consequence of the dominated convergence theorem and $g' \in L^2$. By recalling (3.5), we thus see that (3.4) holds.

Assume now that $g \in \mathbb{H}^2(\mathbb{R})$. In order to prove (3.2), we recall Equations (3.7) and (3.1) to see that

$$\begin{aligned} \int \left|e^{2\pi ibx} - 1 - 2\pi ibx\right|^2 |g(x)|^2\, dx &= 4\pi^2 b^2 \int x^2 \left|\operatorname{sinc}(bx) - e^{-\pi ibx}\right|^2 |g(x)|^2\, dx \\ &\leq 4\pi^4 b^4 \int x^4 |g(x)|^2\, dx = 4\pi^4 b^4 \|X^2 g\|_{L^2}^2\,. \end{aligned}$$

Likewise, we use Equations (3.9), (3.1), and (3.6) to obtain

$$\|T_a g - g + ag'\|_{L^2}^2 \;\leq\; 4\pi^4 a^4 \int \omega^4 \, |\widehat{g}(\omega)|^2 \, d\omega = 4\pi^4 a^4 \|\omega^2 \widehat{g}(\omega)\|_{L^2}^2 \,.$$

Furthermore,

$$a^2 \int \left| e^{2\pi i b x} - 1 \right|^2 |g'(x)|^2 \, dx \;\leq\; 4\pi^2 a^2 b^2 \int x^2 |g'(x)|^2 \, dx = 4\pi^2 a^2 b^2 \|X g'\|_{L^2}^2 \,.$$

Thus, Equation (3.5), combined with the elementary estimate $|ab| \leq \frac{1}{2}(a^2 + b^2)$, shows that

$$\|S_g(a,b) - g - (-ag' + 2\pi i b X g)\|_{L^2} \;\leq\; \tfrac{2}{3} C_g \cdot (a^2 + b^2 + |a||b|) \;\leq\; C_g \cdot \|(a,b)\|_2^2 \,,$$

and the lemma is proved. □

Corollary 3.3. *For any $g \in \mathbb{H}^1(\mathbb{R})$, the map S_g is continuously (Fréchet)-differentiable with*

$$S_g'(\mu) \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = -a\pi(\mu) g' + 2\pi i b X \pi(\mu) g, \qquad \mu \in \mathbb{R}^2.$$

Proof. Let $\mu, \lambda \in \mathbb{R}^2$, $\mu = (u, \eta)$, $\lambda = (a, b)$. Then

$$\pi(\mu + \lambda) g - \pi(\mu) g = M_{\eta+b} T_{u+a} \, g - M_\eta T_u \, g = M_\eta \, (\pi(a,b) - I) \, T_u \, g \,.$$

Now, since $T_u \, g \in \mathbb{H}^1(\mathbb{R})$ with $(T_u \, g)' = T_u \, g'$, Lemma 3.2 shows that

$$(\pi(a,b) - I) \, T_u \, g = -a T_u \, g' + 2\pi i b X T_u \, g + \varepsilon(a,b) \,,$$

where $\varepsilon(a,b) = \varepsilon_u(a,b) \in L^2(\mathbb{R})$ satisfies $\lim_{(a,b)\to(0,0)} \frac{\|\varepsilon(a,b)\|_{L^2}}{\|(a,b)\|_2} = 0$. Thus,

$$\begin{aligned} \pi(\mu + \lambda) g - \pi(\mu) g &= -a M_\eta T_u g' + 2\pi i b M_\eta (X T_u \, g) + M_\eta \, \varepsilon(a,b) \\ &= -a\pi(\mu) g' + 2\pi i b X \pi(\mu) g + \widetilde{\varepsilon}(a,b) \,, \end{aligned}$$

where $\widetilde{\varepsilon} := M_\eta \, \varepsilon$. As $\|\widetilde{\varepsilon}(a,b)\|_{L^2} = \|\varepsilon(a,b)\|_{L^2}$, the claim is proved. □

4. Proof of Theorem 1.3

As mentioned in the introduction, an upper bound in (1.2) is not difficult to achieve. It even holds without the additional assumptions of Λ having rational density or (g, Λ) forming a Riesz sequence.

Proposition 4.1. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let Λ be a lattice in $\mathbb{R}^2$. Then*

$$\operatorname{dist}\big(\pi(\mu) g, \mathcal{G}(g, \Lambda)\big) \;\leq\; \sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2} \cdot \operatorname{dist}(\mu, \Lambda) \quad \text{for all } \mu \in \mathbb{R}^2.$$

Proof. Let $\lambda \in \Lambda$ be a closest point (in Euclidean distance) in Λ to μ. Then $(0,0)$ is a closest point in Λ to $z := \mu - \lambda$, and thus $\operatorname{dist}(\mu, \Lambda) = \operatorname{dist}(z, \Lambda) = \|z\|_2$. By Lemma 2.3 we have

$$\operatorname{dist}\big(\pi(\mu) g, \mathcal{G}(g, \Lambda)\big) = \operatorname{dist}\big(\pi(z) g, \mathcal{G}(g, \Lambda)\big) \leq \|\pi(z) g - g\|_{L^2} \,.$$

Now, if $z = (u, \eta)$, then Plancherel's theorem shows that

$$\begin{aligned} \|\pi(z) g - g\|_{L^2} &\leq \|(\pi(u,\eta) - \pi(0,\eta)) g\|_{L^2} + \|\pi(0,\eta) g - g\|_{L^2} \\ &= \|M_\eta (T_u - I) g\|_{L^2} + \|(M_\eta - I) g\|_{L^2} \\ &= \|(M_{-u} - I) \widehat{g}\|_{L^2} + \|(M_\eta - I) g\|_{L^2} \,. \end{aligned}$$

Next, recall that $|e^{2ix} - 1| = |e^{ix} - e^{-ix}| = 2|\sin(x)| \leq 2|x|$ for $x \in \mathbb{R}$. Using this estimate, we observe for $f \in L^2(\mathbb{R})$ with $\widehat{f} \in H^1(\mathbb{R})$ and $\alpha \in \mathbb{R}$ that

$$\|(M_\alpha - I)f\|_{L^2}^2 = \int \left|e^{2\pi i\alpha x} - 1\right|^2 |f(x)|^2 \, dx \leq 4\pi^2\alpha^2 \int x^2 \, |f(x)|^2 \, dx \,,$$

that is, $\|(M_\alpha - I)f\|_{L^2} \leq 2\pi|\alpha| \cdot \|Xf\|_{L^2} = |\alpha| \cdot \|2\pi i X f\|_{L^2}$. Hence, if we define $\Omega\widehat{f}(\omega) := \omega\,\widehat{f}(\omega)$ for $\omega \in \mathbb{R}$ and $f \in \mathbb{H}^1(\mathbb{R})$, we find that

$$\operatorname{dist}\big(\pi(\mu)g, \mathcal{G}(g,\Lambda)\big) \;\leq\; |u| \cdot \|2\pi i\,\Omega\,\widehat{g}\|_{L^2} + |\eta| \cdot \|2\pi i X g\|_{L^2} \,.$$

Since $2\pi i\,\Omega\,\widehat{g} = \mathcal{F}[g']$, Plancherel's theorem and the Cauchy-Schwarz inequality yield the claim. □

Remark 4.2. If $\Lambda = A\mathbb{Z}^2$ with $A \in \mathrm{GL}(2,\mathbb{R})$, then the maximal distance of a point $\mu \in \mathbb{R}^2$ to the lattice Λ is bounded above by $2^{-1/2}\|A\|_{\mathrm{op}}$. Therefore, for each time-frequency shift $\pi(\mu)g$ of g we have that

$$\operatorname{dist}\big(\pi(\mu)g, \mathcal{G}(g,\Lambda)\big) \leq \sqrt{\frac{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}{2}}\,\|A\|_{\mathrm{op}} \,.$$

In other words, the better g is localized in both time and frequency, the closer the time-frequency shifts of g scatter around $\mathcal{G}(g,\Lambda)$. However, due to the uncertainty principle (see e.g., [12, Theorem 2.2.1]), the constant in the above inequality is easily seen to satisfy $\sqrt{\frac{\|g'\|_{L^2}^2+\|2\pi i X g\|_{L^2}^2}{2}} \geq \sqrt{\pi}$.

In the proof of the next proposition we consider matrix-valued ordinary differential equations (ODEs) of the form

$$X'(t) = X(t)M(t), \tag{4.1}$$

where $X : \mathbb{R} \to \mathbb{C}^{m\times n}$ and where $M : \mathbb{R} \to \mathbb{C}^{n\times n}$ has locally integrable entries. A *solution* of this ODE is a matrix function $X : \mathbb{R} \to \mathbb{C}^{m\times n}$ with (locally) absolutely continuous entries for which $X'(t) = X(t)M(t)$ holds for a.e. $t \in \mathbb{R}$.

Lemma 4.3. *If X_1 and X_2 are two solutions to the ODE* (4.1) *such that $X_1(0) = X_2(0)$, then $X_1(t) = X_2(t)$ for* all $t \in \mathbb{R}$.

Proof. Since the classical ODE theory deals with continuously differentiable solutions to equations with coefficient functions fulfilling a Lipschitz condition, we cannot quite apply that theory. As we will see, however, the same proof idea still works.

Indeed, since $X := X_1 - X_2$ is a solution to the ODE $X' = X \cdot M$ with $X(0) = 0$, it suffices to show that any such function satisfies $X \equiv 0$. Since X is continuous, the set $\Gamma := \{t \in \mathbb{R} \colon X(t) = 0\}$ is closed. Since $\mathbb{R}$ is connected and since $0 \in \Gamma \neq \varnothing$, it is therefore enough to show that Γ is also open.

Thus, let $x_0 \in \Gamma$ be fixed but arbitrary. Since M is locally integrable, there is some $\varepsilon > 0$ such that $\int_{x_0-\varepsilon}^{x_0+\varepsilon} \|M(t)\|_{\mathrm{op}}\,dt \leq \frac{1}{2}$. Now, set $I := [x_0 - \varepsilon, x_0 + \varepsilon]$, and denote by $\mathbf{X} := C(I;\mathbb{C}^{m\times n})$ the space of all continuous functions $f : I \to \mathbb{C}^{m\times n}$, equipped with the norm $\|f\|_{\mathbf{X}} := \sup_{t\in I}\|f(t)\|_{\mathrm{op}}$. It is not hard to see that $\mathbf{X}$ is a Banach space. Furthermore, define the linear operator

$$T : \mathbf{X} \to \mathbf{X}, f \mapsto Tf \quad \text{where} \quad (Tf)(t) := \int_{x_0}^{t} f(s)\,M(s)\,ds \quad \text{for } t \in I \,.$$

Note that indeed $Tf \in \mathbf{X}$ if $f \in \mathbf{X}$, since M is locally integrable, so that $f \cdot M$ is integrable on I. Next, observe

$$\|Tf(t)\|_{\mathrm{op}} \leq \Big| \int_{x_0}^{t} \|f(s)\|_{\mathrm{op}} \cdot \|M(s)\|_{\mathrm{op}} \, ds \Big| \leq \|f\|_{\mathbf{X}} \cdot \int_I \|M(s)\|_{\mathrm{op}} \, ds \leq \frac{1}{2} \, \|f\|_{\mathbf{X}} \,,$$

and hence $\|T\|_{\mathbf{X}\to\mathbf{X}} \leq \frac{1}{2} < 1$. From this, it follows using a Neumann series argument that $\mathrm{id} - T : \mathbf{X} \to \mathbf{X}$ is invertible.

Finally, since $X(x_0) = 0$ and $X'(t) = X(t)M(t)$, we have

$$X(t) = X(t) - X(x_0) = \int_{x_0}^{t} X'(s) \, ds = \int_{x_0}^{t} X(s)M(s) \, ds = \big(T[X|_I]\big)(t)$$

for all $t \in I$, which means that $f := X|_I$ satisfies $(\mathrm{id} - T)f = 0$. Hence $f = 0$, which means that $X \equiv 0$ on $(x_0 - \varepsilon, x_0 + \varepsilon)$. Thus, $(x_0 - \varepsilon, x_0 + \varepsilon) \subset \Gamma$, so that Γ is open. $\square$

The following proposition can be seen as a weak Balian-Low-type theorem for subspaces. For a comparison with related results, see Remark 4.5 below.

Proposition 4.4. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice of rational density such that (g, Λ) is a Riesz basis for its closed linear span $\mathcal{G}(g, \Lambda)$. Then*

$$-ag' + 2\pi i b X g \notin \mathcal{G}(g, \Lambda) \quad \text{for all } (a, b) \in \mathbb{R}^2 \backslash \{(0, 0)\}.$$

Proof. Let us assume towards a contradiction that $\gamma := -ag' + 2\pi i b X g \in \mathcal{G}(g, \Lambda)$ for some $(a, b) \in \mathbb{R}^2 \backslash \{(0, 0)\}$. We divide the proof into five steps.

Step 1: In the first four steps of the proof, we only consider separable lattices of the form $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$ for certain $P, Q \in \mathbb{N}$.

Let $G := Zg \in L^2_{\mathrm{loc}}(\mathbb{R}^2)$ denote the Zak transform of g, and recall from Lemma 2.1 the definition of the function $A_g \in L^2_{\mathrm{loc}}(\mathbb{R}^2; \mathbb{C}^{P \times Q})$ given by

$$A_g(x, \omega) = P^{-1/2} \cdot \big(G(x + \tfrac{k}{P} - \tfrac{\ell}{Q}, \omega)\big)_{k,\ell=0}^{P-1, Q-1} .$$

Since $g \in \mathbb{H}^1(\mathbb{R})$, Lemma 2.4 shows that $G \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$, so that all component functions of A_g are in $H^1_{\mathrm{loc}}(\mathbb{R}^2)$ as well. In this step, we show that A_g satisfies a certain differential equation; see Equation (4.2) below.

Since $\gamma \in \mathcal{G}(g, \Lambda)$ and $\Lambda = \frac{1}{Q}\mathbb{Z} \times P\mathbb{Z}$, Lemma A.16 shows $\pi(\frac{L}{Q}, 0)\gamma \in \mathcal{G}(g, \Lambda)$ for each $L \in \{0, \dots, Q-1\}$. This means that for each $L \in \{0, \dots, Q-1\}$ there is a sequence $(c^{(L)}_{m,n})_{m,n\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$ such that

$$\pi(\tfrac{L}{Q}, 0)\gamma = \sum_{m,n\in\mathbb{Z}} c^{(L)}_{m,n} \, \pi(\tfrac{n}{Q}, Pm) g = \sum_{\ell=0}^{Q-1} \sum_{m,s\in\mathbb{Z}} c^{(L)}_{m,sQ+\ell} \, \pi(s + \tfrac{\ell}{Q}, Pm) g \,.$$

By using the properties (a)–(c) of the Zak transform listed below Equation (2.2), this implies for each $L \in \{0, \dots, Q-1\}$ that

$$\begin{aligned}(Z\gamma)\big(x - \tfrac{L}{Q}, \omega\big) &= Z\big[\pi(\tfrac{L}{Q}, 0)\gamma\big](x, \omega) = \sum_{\ell=0}^{Q-1} \sum_{m,s\in\mathbb{Z}} c^{(L)}_{m,sQ+\ell} \, Z\big[\pi(s + \tfrac{\ell}{Q}, Pm) g\big](x, \omega) \\ &= - \sum_{\ell=0}^{Q-1} f^{(L)}_{\ell}(x, \omega) \cdot G\big(x - \tfrac{\ell}{Q}, \omega\big)\end{aligned}$$

where $f_\ell^{(L)}(x,\omega) := -\sum_{m,s\in\mathbb{Z}} c_{m,sQ+\ell}^{(L)} e^{2\pi i(Pmx-s\omega)}$. Note that each $f_\ell^{(L)}$ is locally square-integrable on $\mathbb{R}^2$ and $(\frac{1}{P},1)$-periodic.

Now, recall from Lemma 2.4 that $(\partial_2 G)(x,\omega) = 2\pi i\big(x\,G(x,\omega) - Z(Xg)(x,\omega)\big)$ and $\partial_1 G = Zg'$. Therefore,

$$\begin{aligned}(Z\gamma)(x,\omega) &= Z[-ag' + 2\pi ibXg](x,\omega)\\ &= -a\cdot\partial_1 G(x,\omega) + 2\pi ib\cdot x\,G(x,\omega) - b\cdot\partial_2 G(x,\omega).\end{aligned}$$

Thus, we arrive at

$$\begin{aligned}&a\,\partial_1 G(x-\tfrac{L}{Q},\omega) + b\,\partial_2 G(x-\tfrac{L}{Q},\omega)\\ &= 2\pi ib\,\big(x-\tfrac{L}{Q}\big)\,G(x-\tfrac{L}{Q},\omega) + \sum_{\ell=0}^{Q-1} f_\ell^{(L)}(x,\omega)\,G(x-\tfrac{\ell}{Q},\omega).\end{aligned}$$

Denoting by $e_0,\dots,e_{Q-1}$ the standard basis vectors of $\mathbb{C}^Q$, plugging $x+\frac{k}{P}$ instead of x into the preceding displayed equation, and recalling that $f_\ell^{(L)}$ is $(\frac{1}{P},1)$-periodic, we obtain for each $L\in\{0,\dots,Q-1\}$ that

$$\begin{aligned}&a\,\partial_1 A_g(x,\omega)\,e_L + b\,\partial_2 A_g(x,\omega)\,e_L\\ &= 2\pi ib\big[\big(x-\tfrac{L}{Q}\big)A_g(x,\omega)e_L + D_P A_g(x,\omega)\,e_L\big] + A_g(x,\omega)\,f^{(L)}(x,\omega)\,,\end{aligned}$$

where $f^{(L)} := (f_\ell^{(L)})_{\ell=0}^{Q-1}$ and $D_P := \operatorname{diag}(k/P)_{k=0}^{P-1}$. This leads to

$$\begin{aligned}a\,\partial_1 A_g(x,\omega) + b\,\partial_2 A_g(x,\omega) &= 2\pi ib[(xA_g(x,\omega) + D_P A_g(x,\omega)]\\ &\quad + A_g(x,\omega)(F(x,\omega) - 2\pi ibD_Q)\,,\end{aligned}$$

where $D_Q := \operatorname{diag}(L/Q)_{L=0}^{Q-1}$ and $F := [f^{(0)} \mid \dots \mid f^{(Q-1)}] \in L^2_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{C}^{Q\times Q})$. As a consequence of Fubini's theorem (and since $(a,b)\neq(0,0)$), there is a null-set $N_0\subset\mathbb{R}^2$ such that $(t\mapsto F(x+ta,\omega+tb))\in L^2_{\mathrm{loc}}(\mathbb{R};\mathbb{C}^{Q\times Q})$ for all $(x,\omega)\in\mathbb{R}^2\backslash N_0$.

Note that the preceding displayed equation holds for almost all $(x,\omega)\in\mathbb{R}^2$. Therefore, if we let $v_t := v + t\,(a,b)$ for $v\in\mathbb{R}^2$ and $t\in\mathbb{R}$, then Lemma A.15 yields a null-set $N_1\subset\mathbb{R}^2$ such that if $v=(x,\omega)\in\mathbb{R}^2\backslash N_1$, then

$$\begin{aligned}&a\,(\partial_1 A_g)(v_t) + b\,(\partial_2 A_g)(v_t)\\ &= 2\pi ib\big[(x+ta)\,A_g(v_t) + D_P A_g(v_t)\big] + A_g(v_t)\cdot\big(F(v_t) - 2\pi ib\,D_Q\big)\\ &= 2\pi ib\,D_P A_g(v_t) + A_g(v_t)[2\pi ib(x+ta) + F(v_t) - 2\pi ib\,D_Q]\\ &= 2\pi ib\,D_P A_g(v_t) + A_g(v_t)W_v(t)\end{aligned}\tag{4.2}$$

for almost all $t\in\mathbb{R}$. In the last step we introduced the matrix

$$W_v(t) := 2\pi ib(x+ta)I_Q + F(v_t) - 2\pi ib\,D_Q \in \mathbb{C}^{Q\times Q},\quad t\in\mathbb{R},$$

where I_Q denotes the Q-dimensional identity matrix. Note $W_v\in L^2_{\mathrm{loc}}(\mathbb{R};\mathbb{C}^{Q\times Q})$ for all $v\in\mathbb{R}^2\backslash N_0$.

Step 2: In this step, we construct a particularly nice representative of $G = Zg$.

Recall from Step 1 that $G\in H^1_{\mathrm{loc}}(\mathbb{R})$. Next, define $\varrho := (a,b)\in\mathbb{R}^2\backslash\{0\}$, and choose $\theta\in\mathbb{R}^2$ with $\|\theta\|_2 = 1$ and $\theta\perp\varrho$. Define $T:\mathbb{R}^2\to\mathbb{R}^2, (t,s)\mapsto t\varrho + s\theta$, and note that T is linear and bijective, so that the same holds also for T^{-1}. In particular, T and T^{-1} are Lipschitz continuous, and thus map null-sets to null-sets. Furthermore, since T and

T^{-1} are Lipschitz continuous, the change-of-variables formula for Sobolev functions (see for instance [21, Theorem 2.2.2]) shows that $\widetilde{G} := G \circ T \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$, and that

$$\begin{aligned} D\widetilde{G}(t,s) &= DG(T(t,s)) \cdot DT(t,s) = \big([\partial_1 G](t\varrho + s\theta), [\partial_2 G](t\varrho + s\theta)\big) \cdot \begin{pmatrix} a & \theta_1 \\ b & \theta_2 \end{pmatrix} \\ &= \big(a\,[\partial_1 G](t\varrho + s\theta) + b\,[\partial_2 G](t\varrho + s\theta), *\big) \end{aligned} \tag{4.3}$$

for almost all $(t,s) \in \mathbb{R}^2$. By Lemma A.12, there is a null-set $N_2 \subset \mathbb{R}$ such that for all $s \in \mathbb{R}\backslash N_2$, Equation (4.3) holds for almost all $t \in \mathbb{R}$.

Lemma A.14 yields a null-set $N_3 \subset \mathbb{R}$, and a (pointwise defined) Borel function $\widetilde{G}_0 : \mathbb{R}^2 \to \mathbb{C}$ such that $\widetilde{G}_0 = \widetilde{G}$ almost everywhere, and such that for all $s \in \mathbb{R}\backslash N_3$, the function $t \mapsto \widetilde{G}_0(t,s)$ is continuous and in $H^1_{\mathrm{loc}}(\mathbb{R})$ with $\frac{d}{dt}\widetilde{G}_0(t,s) = (\partial_1 \widetilde{G})(t,s)$ almost everywhere. In view of Equation (4.3), we thus see for all $s \in \mathbb{R}\backslash(N_2 \cup N_3)$ that

$$\frac{d}{dt}\widetilde{G}_0(t,s) = a\,[\partial_1 G](t\varrho + s\theta) + b\,[\partial_2 G](t\varrho + s\theta) \quad \text{for almost all } t \in \mathbb{R}.$$

Note that since $\widetilde{G}_0 = \widetilde{G} = G \circ T$ almost everywhere and since T and T^{-1} map null-sets to null-sets, we have $G = \widetilde{G}_0 \circ T^{-1} =: G_0$ almost everywhere. By Lemma A.15, there is thus a null-set $N_4 \subset \mathbb{R}^2$ such that

$$\forall\,(x,\omega) \in \mathbb{R}^2\backslash N_4\colon \quad G_0(x + ta, \omega + tb) = G(x + ta, \omega + tb) \text{ for a.e. } t \in \mathbb{R}\,. \tag{4.4}$$

Since T is Lipschitz continuous, the set $N_5 := T(\mathbb{R}\times(N_2 \cup N_3)) \subset \mathbb{R}^2$ is a null-set. For any $(x,\omega) \in \mathbb{R}^2\backslash N_5$, we have $(x,\omega) = T(t_0,s_0) = t_0\varrho + s_0\theta$ for certain $(t_0,s_0) \in \mathbb{R}\times(\mathbb{R}\backslash(N_2 \cup N_3))$. By the properties from above, this means that the map

$$\begin{aligned} \mathbb{R} \to \mathbb{C}, \quad t \mapsto G_0(x + ta, \omega + tb) &= G_0\big((x,\omega) + t\varrho\big) \\ &= G_0\big((t + t_0)\varrho + s_0\theta\big) = \widetilde{G}_0(t + t_0, s_0) \end{aligned}$$

is continuous and in $H^1_{\mathrm{loc}}(\mathbb{R})$ with derivative

$$\begin{aligned} \frac{d}{dt}G_0(x + ta, \omega + tb) &= \frac{d}{dt}\widetilde{G}_0(t + t_0, s_0) \\ &= a\,[\partial_1 G]\big((t + t_0)\varrho + s_0\theta\big) + b\,[\partial_2 G]\big((t + t_0)\varrho + s_0\theta\big) \\ &= a\,[\partial_1 G](x + ta, \omega + tb) + b\,[\partial_2 G](x + ta, \omega + tb) \end{aligned} \tag{4.5}$$

for almost all $t \in \mathbb{R}$, for each fixed $(x,\omega) \in \mathbb{R}^2\backslash N_5$.

Finally, let $N_6 := \bigcup_{k,\ell\in\mathbb{Z}} \big((N_4 \cup N_5) + (\frac{\ell}{Q} - \frac{k}{P}, 0)\big) \subset \mathbb{R}^2$, which is a null-set. If $(x,\omega) \in \mathbb{R}^2\backslash N_6$, then $\big(x + \frac{k}{P} - \frac{\ell}{Q}, \omega\big) \in \mathbb{R}^2\backslash(N_4 \cup N_5)$ for all $k,\ell \in \mathbb{Z}$.

Step 3: In this step, we use the "nice" representative G_0 of G to construct for almost all $v = (x,\omega) \in \mathbb{R}^2$ two locally absolutely continuous functions $R_v : \mathbb{R} \to \mathbb{C}^{P\times Q}$ and $L_v : \mathbb{R} \to \mathbb{C}^{P\times Q}$ which satisfy the differential equations $R_v'(t) = R_v(t)W_v(t)$ and $L_v'(t) = L_v(t)W_v(t)$ for almost all $t \in \mathbb{R}$, for the matrix function $W_v \in L^2_{\mathrm{loc}}(\mathbb{R}; \mathbb{C}^{Q\times Q})$ defined in Step 1. We then use this differential equation to deduce $R_v = L_v$. In Step 4 we will finally employ this identity to complete the proof for the case $\Lambda = \frac{1}{Q}\mathbb{Z}\times P\mathbb{Z}$.

First, define

$$A\colon \quad \mathbb{R}^2 \to \mathbb{C}^{P\times Q}, \quad (x,\omega) \mapsto P^{-1/2} \cdot \big(G_0(x + \tfrac{k}{P} - \tfrac{\ell}{Q}, \omega)\big)_{k,\ell=0}^{P-1,Q-1},$$

noting $A = A_g$ almost everywhere. Next, note for $v = (x,\omega) \in \mathbb{R}^2\backslash(N_0 \cup N_1 \cup N_6)$ that $(x + \frac{k}{P} - \frac{\ell}{Q}, \omega) \in \mathbb{R}^2\backslash(N_4 \cup N_5)$ for all $k,\ell \in \mathbb{Z}$, so that Equations (4.5), (4.2), and

(4.4) show that the function $E_v : \mathbb{R} \to \mathbb{C}^{P\times Q}, t \mapsto A(v_t) = A(x + ta, \omega + tb)$ is locally absolutely continuous and satisfies

$$\begin{aligned}
E_v'(t) &= P^{-1/2}\Big(a\,(\partial_1 G)\big(x + \tfrac{k}{P} - \tfrac{\ell}{Q} + ta, \omega + tb\big) \\
&\qquad\qquad + b\,(\partial_2 G)\big(x + \tfrac{k}{P} - \tfrac{\ell}{Q} + ta, \omega + tb\big)\Big)_{k,\ell=0}^{P-1,Q-1} \\
&= a\,(\partial_1 A_g)(x + ta, \omega + tb) + b\,(\partial_2 A_g)(x + ta, \omega + tb) \\
&= a\,[\partial_1 A_g](v_t) + b\,[\partial_2 A_g](v_t) \\
&= 2\pi i b\, D_P A_g(v_t) + A_g(v_t) W_v(t) \\
&= 2\pi i b\, D_P A(v_t) + A(v_t) W_v(t) = 2\pi i b\, D_P\, E_v(t) + E_v(t) W_v(t) \qquad (4.6)
\end{aligned}$$

for almost all $t \in \mathbb{R}$.

Next, Lemma 2.1 shows that $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A_g(z)) > 0$, since (g, Λ) is a Riesz sequence. Hence, we also have $\operatorname{essinf}_{z\in\mathbb{R}^2} \sigma_0(A(z)) > 0$, which means that $(A^*A)(x, \omega)$ is invertible for almost all $(x, \omega) \in \mathbb{R}^2$, say for all $(x, \omega) \in \mathbb{R}^2 \backslash N_7$.

For $v = (x, \omega) \in \mathbb{R}^2 \backslash (N_0 \cup N_1 \cup N_6 \cup N_7)$, set $C_v := A(v)\big(A^*(v)A(v)\big)^{-1} A^*(v)$ (so that $C_v \in \mathbb{C}^{P\times P}$) and furthermore

$$\begin{aligned}
R_v : &\quad \mathbb{R} \to \mathbb{C}^{P\times Q}, \quad t \mapsto e^{-2\pi i t b\, D_P}\; A(x + ta, \omega + tb) = e^{-2\pi i t b\, D_P}\; E_v(t), \\
L_v : &\quad \mathbb{R} \to \mathbb{C}^{P\times Q}, \quad t \mapsto C_v \cdot e^{-2\pi i t b\, D_P} \cdot A(x + ta, \omega + tb) = C_v \cdot R_v(t),
\end{aligned}$$

where as before $D_P = \operatorname{diag}(k/P)_{k=0,\dots,P-1} \in \mathbb{R}^{P\times P}$.

Since $v = (x, \omega) \in \mathbb{R}^2 \backslash (N_0 \cup N_1 \cup N_6)$, we see as a consequence of the product rule for Sobolev functions (see for instance [2, Section 4.25]) and of Equation (4.6) that R_v is locally absolutely continuous, with

$$R_v'(t) = -2\pi i b D_P R_v(t) + e^{-2\pi i t b\, D_P} E_v'(t) = R_v(t)\, W_v(t) \quad \text{for almost all } t \in \mathbb{R}.$$

where the last equality follows from Equation (4.6) combined with the elementary identity $e^{-2\pi i t b\, D_P} D_P = D_P\, e^{-2\pi i t b\, D_P}$. This easily implies that L_v is locally absolutely continuous as well, with $L_v'(t) = C_v\, R_v'(t) = C_v\, R_v(t)\, W_v(t) = L_v(t)\, W_v(t)$ for almost all $t \in \mathbb{R}$. Finally, note that

$$L_v(0) = C_v \cdot E_v(0) = A(v)\big(A^*(v)A(v)\big)^{-1} A^*(v)A(v) = A(v) = R_v(0)\,.$$

Therefore, Lemma 4.3 shows $L_v(t) = R_v(t)$ for all $v \in \mathbb{R}^2 \backslash (N_0 \cup N_1 \cup N_6 \cup N_7)$ and all $t \in \mathbb{R}$.

Step 4: We complete the proof for the case $\Lambda = Q^{-1}\mathbb{Z} \times P\mathbb{Z}$. To this end, let $t \in \mathbb{R}$ be arbitrary, and note that the matrix function $H_{(-ta,-tb)}$ defined in Lemma 2.2 satisfies for *almost* all $v = (x, \omega) \in \mathbb{R}^2$ that

$$\begin{aligned}
&H_{(-ta,-tb)}(x, \omega) \\
&= P^{1/2} \cdot A_g(v) \cdot \big(A_g^*(v) A_g(v)\big)^{-1} A_g^*(v) \cdot e^{-2\pi i t b\, D_P} \cdot A_g(x + ta, \omega + tb) \\
&= P^{1/2} \cdot C_v \cdot e^{-2\pi i t b\, D_P} \cdot A(x + ta, \omega + tb) \\
&= P^{1/2} \cdot L_v(t) = P^{1/2} \cdot R_v(t).
\end{aligned}$$

Hence,

$$\|H_{(-ta,-tb)}(x, \omega) e_0\|_{\mathbb{C}^P}^2 = P\, \|R_v(t) e_0\|_{\mathbb{C}^P}^2 = P\, \|A(v_t) e_0\|_{\mathbb{C}^P}^2 = P\, \|A_g(v_t) e_0\|_{\mathbb{C}^P}^2$$

$$= \sum_{k=0}^{P-1} \left|G\big(x + ta + \tfrac{k}{P}, \omega + tb\big)\right|^2 \quad \text{for a.e. } (x,\omega) \in \mathbb{R}^2.$$

By Lemma 2.2 and by the quasi-periodicity of $G = Zg$ (which implies that $|G|$ is $(1,1)$-periodic), this implies that

$$\begin{aligned}
\operatorname{dist}^2\big(\pi(-ta,-tb)g, \mathcal{G}(g,\Lambda)\big) &= \|g\|_{L^2}^2 - \int_0^1 \int_0^{\frac{1}{P}} \sum_{k=0}^{P-1} \left|G\big(x+ta+\tfrac{k}{P}, \omega+tb\big)\right|^2 dx\, d\omega \\
&= \|g\|_{L^2}^2 - \int_0^1 \int_0^1 |G(x+ta, \omega+tb)|^2 \, dx\, d\omega \\
&= \|g\|_{L^2}^2 - \int_0^1 \int_0^1 |G(x,\omega)|^2 \, dx\, d\omega = 0\,.
\end{aligned}$$

That is, $\pi(-ta,-tb)g \in \mathcal{G}(g,\Lambda)$ for each $t \in \mathbb{R}$. By Theorem 1.2, this means that $(-ta,-tb) \in \Lambda$ for every $t \in \mathbb{R}$. Because of $(a,b) \neq (0,0)$ and since $\Lambda \subset \mathbb{R}^2$ is discrete, this yields the desired contradiction.

Step 5: Let $\Lambda \subset \mathbb{R}^2$ be an arbitrary lattice of rational density, and assume again that $-ag' + 2\pi i b X g \in \mathcal{G}(g,\Lambda)$ for some $a, b \in \mathbb{R}$. Then there exists a matrix $B \in \mathrm{GL}(2,\mathbb{R})$ with $\det B = 1$ and certain $P, Q \in \mathbb{N}$ such that $\Lambda_1 := B\Lambda = Q^{-1}\mathbb{Z} \times P\mathbb{Z}$. With the symplectic operator U_B (see (2.9)), set $g_1 := U_B\, g$. Then (g_1, Λ_1) is a Riesz basis for $\mathcal{G}(g_1,\Lambda_1) = U_B\, \mathcal{G}(g,\Lambda)$ and, as $\mathbb{H}^1(\mathbb{R})$ is invariant under symplectic operators (see the discussion after Equation (2.11)), we have $g_1 \in \mathbb{H}^1(\mathbb{R})$. For $f \in \mathbb{H}^1(\mathbb{R})$, let us set $T_f(x,\omega) := \rho(x,\omega)f$, $x, \omega \in \mathbb{R}$, cf. (2.10). Using Corollary 3.3 we find that

$$\begin{aligned}
&T_f'(x,\omega)\, \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \\
&= e^{-\pi i x\omega} \left[(-\pi i \omega S_f(x,\omega) + \partial_1 S_f(x,\omega))\, a + (-\pi i x S_f(x,\omega) + \partial_2 S_f(x,\omega))\, b\right] \\
&= e^{-\pi i x\omega} \left[-\pi i (a\omega + bx) S_f(x,\omega) - a\pi(x,\omega) f' + 2\pi i b X \pi(x,\omega) f\right].
\end{aligned}$$

In particular,

$$T_f'(0,0)\, \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = -af' + 2\pi i b X f. \tag{4.7}$$

We have (see (2.9))

$$U_B\, T_g(x,\omega) = \rho(B\, \left(\begin{smallmatrix} x \\ \omega \end{smallmatrix}\right)) g_1 = T_{g_1}(B\, \left(\begin{smallmatrix} x \\ \omega \end{smallmatrix}\right))\,.$$

Differentiating this with respect to (x,ω) gives $U_B\, T_g'(x,\omega) = T_{g_1}'(B\, \left(\begin{smallmatrix} x \\ \omega \end{smallmatrix}\right)) \circ B$. Hence, by Equation (4.7), we see that

$$U_B\, (-ag' + 2\pi i b X g) = U_B\, T_g'(0,0)\, \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = T_{g_1}'(0,0)\big(B\, \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)\big) = -\alpha g_1' + 2\pi i \beta X g_1\,,$$

where $\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right) = B\, \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$. That is, $-\alpha g_1' + 2\pi i \beta X g_1 \in U_B \mathcal{G}(g,\Lambda) = \mathcal{G}(g_1,\Lambda_1)$, which, by the first part of this proof, implies that $\alpha = \beta = 0$ and thus $a = b = 0$. □

Remark 4.5. Proposition 4.4 is closely related to the so-called weak subspace Balian-Low Theorem (cf. [13, Thm. 8]) which states that if $g \in L^2(\mathbb{R})$ and $\Lambda \subset \mathbb{R}^2$ is a lattice such that (g,Λ) is a Riesz basis for its closed linear span $\mathcal{G}$, then at least one of the distributions $g', Xg, \widetilde{g}', X\widetilde{g}$ is not contained in $\mathcal{G}$, where $\widetilde{g}$ denotes the dual window of (g,Λ). More precisely, Proposition 4.4 implies that if $g', Xg \in L^2(\mathbb{R})$ and $\Lambda \subset \mathbb{R}^2$ is a lattice of rational density such that (g,Λ) is a Riesz sequence (and hence also $\widetilde{g}', X\widetilde{g} \in L^2(\mathbb{R})$ by Proposition 2.5), then *none* of $g', Xg, \widetilde{g}', X\widetilde{g}$ is contained in $\mathcal{G}$. In fact, it even asserts that none of the real linear combinations of ig' and Xg except 0 can

belong to $\mathcal{G}$. Similarly, none of the real linear combinations of $i\widetilde{g}'$ and $X\widetilde{g}$ except 0 can belong to $\mathcal{G}$.

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. Let us denote by $\mathbb{P}$ the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G} := \mathcal{G}(g,\Lambda)$. Proposition 4.4 implies that the $\mathbb{R}$-linear mapping

$$\mathbb{R}^2 \to L^2(\mathbb{R}),\ \ (a,b) \mapsto (\mathrm{Id}-\mathbb{P})(-ag' + 2\pi i b X g)\,,$$

with $L^2(\mathbb{R})$ considered as an $\mathbb{R}$-linear space, is injective. Since $\mathbb{R}^2$ is finite-dimensional, this implies $\|(\mathrm{Id}-\mathbb{P})(-ag'+2\pi i b X g)\|_{L^2} \geq 2\gamma\|(a,b)\|_2$ for some $\gamma > 0$ and all $(a,b) \in \mathbb{R}^2$. On the other hand, Lemma 3.2 gives a family of functions $\{\varepsilon(a,b)\}_{(a,b)\in\mathbb{R}^2} \subset L^2(\mathbb{R})$ such that

$$\pi(a,b)g - g = -ag' + 2\pi i b X g + \varepsilon(a,b) \quad \text{and} \quad \lim_{(a,b)\to(0,0)} \frac{\|\varepsilon(a,b)\|_{L^2}}{\|(a,b)\|_2} = 0\,.$$

In particular, there exists some $\delta > 0$ such that $\|\varepsilon(a,b)\|_{L^2} \leq \gamma\|(a,b)\|_2$ for $\|(a,b)\|_2 < \delta$. Combining these observations and the fact that $(\mathrm{Id}-\mathbb{P})g = 0$, we see for $\|(a,b)\|_2 < \delta$ that

$$\begin{aligned} 2\gamma\|(a,b)\|_2 &\leq \|(\mathrm{Id}-\mathbb{P})(-ag' + 2\pi i b X g)\|_{L^2} = \big\|\big(\mathrm{Id}-\mathbb{P}\big)\big(\pi(a,b)g - \varepsilon(a,b)\big)\big\|_{L^2} \\ &\leq \|(\mathrm{Id}-\mathbb{P})\pi(a,b)g\|_{L^2} + \|\varepsilon(a,b)\|_{L^2} \leq \mathrm{dist}(\pi(a,b)g,\mathcal{G}) + \gamma\,\|(a,b)\|_2\,, \end{aligned}$$

that is, $\mathrm{dist}(\pi(a,b)g,\mathcal{G}) \geq \gamma\|(a,b)\|_2$ for $\|(a,b)\|_2 < \delta$.

Now, consider the compact set $R := \{\mu \in \mathbb{R}^2 : \|\mu\|_2 = \mathrm{dist}(\mu,\Lambda)\}$ and denote by $B = B_\delta(0,0) \subset \mathbb{R}^2$ the open ball of radius $\delta > 0$ centered at $(0,0)$. By possibly shrinking δ, we may assume that $B \subset R$; in fact, since Λ is discrete, there is some $\delta_0 > 0$ such that $\|\lambda\|_2 \geq 2\delta_0$ for all $\lambda \in \Lambda\backslash\{0\}$. We then have $B \subset R$ as soon as $0 < \delta \leq \delta_0$.

We will show that $\|(\mathrm{Id}-\mathbb{P})\pi(a,b)g\|_{L^2} \geq \gamma'\|(a,b)\|_2$ for a suitable $\gamma' > 0$ and all $(a,b) \in R\backslash B$. Towards a contradiction, suppose that there is no such $\gamma' > 0$. Then there exists a sequence $(\mu_n)_{n\in\mathbb{N}} \subset R\backslash B$ such that $(\mathrm{Id}-\mathbb{P})\pi(\mu_n)g \to 0$ as $n \to \infty$. As $R\backslash B$ is compact, we may assume that $\mu_n \to \mu_0$ as $n \to \infty$ for some $\mu_0 \in R\backslash B$. But then, since $\mu \mapsto \pi(\mu)g$ is continuous, it follows that $(\mathrm{Id}-\mathbb{P})\pi(\mu_0)g = 0$, that is, $\pi(\mu_0)g \in \mathcal{G}$, which by Theorem 1.2 is only possible if $\mu_0 \in \Lambda$; but this implies $\|\mu_0\|_2 = \mathrm{dist}(\mu_0,\Lambda) = 0$, in contradiction to $\mu_0 \in R\backslash B$.

Hence, $\mathrm{dist}(\pi(a,b)g,\mathcal{G}) = \|(\mathrm{Id}-\mathbb{P})\pi(a,b)g\|_{L^2} \geq \gamma'\|(a,b)\|_2$ for some $\gamma' > 0$ and all $(a,b) \in R\backslash B$. As a consequence, we have with $C_1 := \min\{\gamma,\gamma'\} > 0$,

$$\mathrm{dist}(\pi(\mu)g,\mathcal{G}) \geq C_1 \cdot \|\mu\|_2 = C_1 \cdot \mathrm{dist}(\mu,\Lambda) \quad \text{for all } \ \mu \in R.$$

Finally, we note that for each $\mu \in \mathbb{R}^2$ there exist $\lambda \in \Lambda$ and $\nu \in R$ with $\mu = \lambda + \nu$; indeed, there exists $\lambda \in \Lambda$ with $\|\mu - \lambda\|_2 = \mathrm{dist}(\mu,\Lambda)$, and then $\nu := \mu - \lambda$ satisfies $\|\nu\|_2 = \mathrm{dist}(\mu,\Lambda) = \mathrm{dist}(\nu,\Lambda)$. Thus, we obtain (see Lemma 2.3)

$$\mathrm{dist}(\pi(\mu)g,\mathcal{G}) = \mathrm{dist}(\pi(\nu)g,\mathcal{G}) \ \geq\ C_1 \cdot \mathrm{dist}(\nu,\Lambda) = C_1 \cdot \mathrm{dist}(\mu,\Lambda).$$

In view of Proposition 4.1, this completes the proof. □

5. An Explicit Local Bound

As mentioned in the introduction, we were unable to derive an *explicit* constant C_1 for (1.2). Nevertheless, we can find a constant C_1 that is valid for (u,η) close to the lattice Λ. For this, however, we have to assume that (g,Λ) is an orthonormal sequence.

The following result makes a first step towards finding such a constant C_1; it improves Proposition 4.4 under the additional assumption of orthonormality.

Proposition 5.1. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g, Λ) is an orthonormal basis of its closed linear span $\mathcal{G}(g, \Lambda)$. Then for any $(a, b) \in \mathbb{R}^2$,*

$$\operatorname{dist}\left(-ag' + 2\pi i b X g,\, \mathcal{G}(g,\Lambda)\right) \geq \frac{\pi}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}} \, \|(a,b)\|_2 \,.$$

Remark 5.2. The classical uncertainty principle (see e.g., [12, Theorem 2.2.1]), combined with elementary computations, implies because of $\|g\|_{L^2} = 1$ that the lower bound appearing in Proposition 5.1 is bounded by

$$\frac{\pi}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}} \leq \sqrt{\pi/2}.$$

The proof of Proposition 5.1 hinges crucially on the following lemma which describes a general property of Hilbert spaces.

Lemma 5.3. *Let $\mathcal{H}$ be a Hilbert space, and let $f, g \in \mathcal{H}$ with $f \neq 0$ or $g \neq 0$. Then*

$$\|af + bg\|^2 \geq \frac{\|f\|^2 \cdot \|g\|^2 - (\operatorname{Re}\langle f, g\rangle)^2}{\|f\|^2 + \|g\|^2} \cdot \|(a,b)\|_2^2 \geq \frac{(\operatorname{Im}\langle f, g\rangle)^2}{\|f\|^2 + \|g\|^2} \cdot \|(a,b)\|_2^2$$

for all $a, b \in \mathbb{R}$.

Proof. Let $\alpha := \|f\|_{\mathcal{H}}^2$, $\gamma := \|g\|_{\mathcal{H}}^2$, and $\beta := \operatorname{Re}\langle f, g\rangle$. Moreover, set $A := \alpha + \gamma$ and $B := \alpha\gamma - \beta^2$. Because of $f \neq 0$ or $g \neq 0$, we have $A > 0$. Besides, the Cauchy-Schwarz inequality shows $\beta \leq |\beta| \leq \sqrt{\alpha\gamma}$, and thus $B \geq 0$. Finally, a direct computation shows $A^2 - 4B = (\alpha + \gamma)^2 - 4(\alpha\gamma - \beta^2) = (\alpha - \gamma)^2 + 4\beta^2 \geq 0$.

Given these notations, another direct computation shows for $a, b \in \mathbb{R}$ that

$$\|af + bg\|_{\mathcal{H}}^2 = \langle af + bg, af + bg\rangle_{\mathcal{H}} = \left\langle \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right), M \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \right\rangle_{\mathbb{R}^2} \quad \text{where} \quad M := \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}. \tag{5.1}$$

Note that the matrix M is real-symmetric, with characteristic polynomial

$$\chi_M(\lambda) = \det \begin{pmatrix} \lambda - \alpha & -\beta \\ -\beta & \lambda - \gamma \end{pmatrix} = \lambda^2 - A\,\lambda + B \,,$$

which has the roots

$$\lambda_{1/2} = \frac{A}{2} \pm \sqrt{\frac{A^2}{4} - B} = \frac{A \pm \sqrt{A^2 - 4B}}{2} \,.$$

Therefore, and because of $\sqrt{A^2 - 4B} \leq \sqrt{A^2} = A$, the smallest eigenvalue of M satisfies

$$\lambda_{\min} = \frac{A - \sqrt{A^2 - 4B}}{2} = \frac{1}{2} \frac{A^2 - (A^2 - 4B)}{A + \sqrt{A^2 - 4B}} = \frac{2B}{A + \sqrt{A^2 - 4B}} \geq \frac{B}{A} \geq 0 \,.$$

Since M is real symmetric, this implies $\langle x, Mx\rangle_{\mathbb{R}^2} \geq \frac{B}{A}\|x\|_2^2$ for all $x \in \mathbb{R}^2$.

Now, Equation (5.1) shows that $\|af + bg\|_{\mathcal{H}}^2 = \langle \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right), M \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)\rangle_{\mathbb{R}^2} \geq \frac{B}{A} \cdot \|(a,b)\|_2^2$ for all $a, b \in \mathbb{R}$, which establishes the first part of the claim. For the second part, note that the Cauchy-Schwarz inequality implies

$$\begin{aligned} B = \alpha\gamma - \beta^2 &= \|f\|_{\mathcal{H}}^2 \, \|g\|_{\mathcal{H}}^2 - (\operatorname{Re}\langle f, g\rangle_{\mathcal{H}})^2 \\ &\geq |\langle f, g\rangle_{\mathcal{H}}|^2 - (\operatorname{Re}\langle f, g\rangle_{\mathcal{H}})^2 = (\operatorname{Im}\langle f, g\rangle_{\mathcal{H}})^2 \,. \end{aligned}$$

□

Proof of Proposition 5.1. Denote by $\mathbb{P}$ the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda)$ in $L^2(\mathbb{R})$. Since (g,Λ) is an orthonormal sequence,

$$\mathbb{P}f = \sum_{\lambda\in\Lambda} \langle f, \pi(\lambda)g\rangle \pi(\lambda)g\,, \quad \text{whence} \quad \langle \mathbb{P}g', iXg\rangle = \sum_{\lambda\in\Lambda} \langle g', \pi(\lambda)g\rangle\langle \pi(\lambda)g, iXg\rangle\,.$$

Let $a,b\in\mathbb{R}$. By integration by parts and translation, and by using the elementary identity $(\pi(a,b))^* = e^{-2\pi iab}\pi(-a,-b)$, we see that

$$\begin{aligned}\langle g', \pi(a,b)g\rangle &= -\Big\langle g,\ \frac{d}{dx}\big(e^{2\pi ibx}\, g(x-a)\big)\Big\rangle \\ &= -\big\langle g, 2\pi ib\cdot \pi(a,b)g\big\rangle - \big\langle g, \pi(a,b)g'\big\rangle \\ &= 2\pi ib\cdot\big\langle g,\ \pi(a,b)g\big\rangle - e^{-2\pi iab}\cdot\big\langle \pi(-a,-b)g, g'\big\rangle\,,\end{aligned} \tag{5.2}$$

as well as

$$\begin{aligned}\langle \pi(a,b)g, iXg\rangle &= \big\langle g, e^{-2\pi iab}\,\pi(-a,-b)[iXg]\big\rangle \\ &= e^{2\pi iab}\cdot\big\langle g,\ M_{-b}\big[\, i\ ((\cdot)+a)\ g(\cdot+a)\,\big]\big\rangle \\ &= e^{2\pi iab}\ \big(-ia\big\langle g,\ M_{-b}[g(\cdot+a)]\big\rangle + \big\langle -iXg,\ M_{-b}[g(\cdot+a)]\big\rangle\big) \\ &= e^{2\pi iab}\ \big(-ia\big\langle g,\ \pi(-a,-b)g\big\rangle - \big\langle iXg,\ \pi(-a,-b)g\big\rangle\big)\,.\end{aligned} \tag{5.3}$$

From Equations (5.2) and (5.3), we see by orthonormality of (g,Λ) for arbitrary $(a,b)\in\Lambda$ that

$$\begin{aligned}\langle g', \pi(a,b)g\rangle &= \begin{cases} -\langle g, g'\rangle\,, & \text{if } (a,b)=0\,,\\ -e^{-2\pi iab}\,\langle \pi(-a,-b)g, g'\rangle\,, & \text{otherwise}\end{cases} \\ &= -e^{-2\pi iab}\,\langle \pi(-a,-b)g, g'\rangle\end{aligned}$$

and

$$\begin{aligned}\langle \pi(a,b)g, iXg\rangle &= \begin{cases} -\langle iXg, g\rangle\,, & \text{if } (a,b)=0\,,\\ -e^{2\pi iab}\,\langle iXg, \pi(-a,-b)g\rangle\,, & \text{otherwise}\end{cases} \\ &= -e^{2\pi iab}\,\langle iXg, \pi(-a,-b)g\rangle\,.\end{aligned}$$

Combining these identities, we arrive at

$$\langle g', \pi(a,b)g\rangle\langle \pi(a,b)g, iXg\rangle = \langle \pi(-a,-b)g, g'\rangle\langle iXg, \pi(-a,-b)g\rangle\,,$$

for all $(a,b)\in\Lambda$. Therefore, with $\mu=-\lambda$, we see that

$$\begin{aligned}\langle \mathbb{P}g', iXg\rangle &= \sum_{\lambda\in\Lambda}\langle g', \pi(\lambda)g\rangle\langle \pi(\lambda)g, iXg\rangle \\ &= \sum_{\mu\in\Lambda}\langle \pi(\mu)g, g'\rangle\langle iXg, \pi(\mu)g\rangle = \langle iXg, \mathbb{P}g'\rangle\,,\end{aligned}$$

which shows that $\mathrm{Im}\langle \mathbb{P}g', iXg\rangle = 0$.

We now intend to use partial integration to get $\langle g', Xg\rangle = -\|g\|_{L^2}^2 - \langle Xg, g'\rangle$; however, since $Xg\notin H^1(\mathbb{R})$, we cannot *directly* apply such a partial integration. Instead, pick $\varphi\in C_c^\infty(\mathbb{R})$ with $0\le\varphi\le 1$, $\operatorname{supp}\varphi\subset(-2,2)$, and $\varphi\equiv 1$ on $(-1,1)$, and set $\varphi_n:\mathbb{R}\to[0,1], x\mapsto\varphi(x/n)$. We then have $\varphi_n\to 1$ pointwise, so that the dominated convergence theorem implies $\langle f, \varphi_n\cdot h\rangle\to\langle f,h\rangle$ for all $f,h\in L^2(\mathbb{R})$. Likewise, we have

$\varphi_n'(x) = n^{-1} \cdot \varphi'(x/n)$ and hence $\varphi_n' \to 0$ uniformly, which implies $\langle f, \varphi_n' \cdot h\rangle \to 0$ for $f, h \in L^2(\mathbb{R})$. Overall, since $\varphi_n X g \in H^1(\mathbb{R})$, we thus see

$$\begin{aligned}\langle g', Xg\rangle &= \lim_{n\to\infty} \langle g', \varphi_n X g\rangle = \lim_{n\to\infty} -\langle g, (\varphi_n X g)'\rangle \\ &= -\lim_{n\to\infty} \big[\langle g, \varphi_n' \cdot Xg\rangle + \langle g, \varphi_n g\rangle + \langle g, \varphi_n X g'\rangle\big] = -\|g\|_{L^2}^2 - \langle Xg, g'\rangle\,.\end{aligned}$$

Here, we used in the last step that $\langle g, \varphi_n X g'\rangle = \langle Xg, \varphi_n g'\rangle$ with $Xg, g' \in L^2(\mathbb{R})$.

In view of the last displayed equation, we get $2\operatorname{Re}\langle g', Xg\rangle = -\|g\|_{L^2}^2 = -1$, and hence $\operatorname{Im}\langle g', 2\pi i X g\rangle = -2\pi \operatorname{Re}\langle g', Xg\rangle = \pi$. Therefore,

$$\operatorname{Im}\langle (I-\mathbb{P})g', 2\pi i X g\rangle = \operatorname{Im}\langle g', 2\pi i X g\rangle - 2\pi \operatorname{Im}\langle \mathbb{P} g', iXg\rangle = \pi.$$

Setting $f := (I-\mathbb{P})[-g']$ and $h := (I-\mathbb{P})[2\pi i X g]$, we have shown up to now that $\operatorname{Im}\langle f, h\rangle = -\pi \neq 0$, which in particular implies that $f \neq 0$ and $h \neq 0$. Thus, an application of Lemma 5.3 shows for arbitrary $(a,b) \in \mathbb{R}^2$ that

$$\begin{aligned}&\operatorname{dist}^2\big(-ag' + 2\pi i b X g,\, \mathcal{G}(g,\Lambda)\big) \\ &= \|(I-\mathbb{P})(-ag' + 2\pi i b X g)\|_{L^2}^2 = \|a\cdot f + b\cdot h\|_{L^2}^2 \\ &\geq \frac{(\operatorname{Im}\langle f, h\rangle)^2}{\|f\|_{L^2}^2 + \|h\|_{L^2}^2} \cdot \|(a,b)\|_2^2 \geq \frac{\pi^2 \cdot \|(a,b)\|_2^2}{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}\,.\end{aligned}$$

This concludes the proof of the proposition. □

Theorem 5.4. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g,Λ) is an orthonormal basis of its closed linear span $\mathcal{G}(g,\Lambda)$. Then there exists $\varepsilon > 0$ such that*

$$\operatorname{dist}\big(\pi(\mu)g,\, \mathcal{G}(g,\Lambda)\big) \;\geq\; \frac{\pi/2}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}}\, \operatorname{dist}(\mu, \Lambda) \quad \forall\, \mu \in \Lambda + B_\varepsilon(0).$$

If $g \in \mathbb{H}^2(\mathbb{R})$, then ε can be chosen as $\varepsilon := \pi \Big/ \Big(2C_g \sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}\Big)$ with C_g as in Equation (3.3).

Proof. For $(a,b) \in \mathbb{R}^2$ let $\gamma(a,b) := \pi(a,b)g - g - (-ag' + 2\pi i b X g)$. Denote by $\mathbb{P}$ the orthogonal projection from $L^2(\mathbb{R})$ onto $\mathcal{G}(g,\Lambda)$. Due to Proposition 5.1 we have

$$\begin{aligned}\frac{\pi}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}} \|(a,b)\|_2 &\leq \|(I-\mathbb{P})(-ag' + 2\pi i b X g)\|_{L^2} \\ &= \big\|(I-\mathbb{P})\big(\pi(a,b)g - g - \gamma(a,b)\big)\big\|_{L^2} \\ &\leq \|(I-\mathbb{P})\pi(a,b)g\|_{L^2} + \|\gamma(a,b)\|_{L^2}\,.\end{aligned}$$

In the last inequality we used that $(I-\mathbb{P})g = 0$ and $\|I-\mathbb{P}\| = 1$. By Lemma 3.2 there exists $\varepsilon > 0$ such that

$$\|\gamma(a,b)\|_{L^2} \leq \frac{\pi/2}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}}\, \|(a,b)\|_2 \qquad \text{for } \|(a,b)\|_2 < \varepsilon\,.$$

Moreover, this is satisfied in the case $g \in \mathbb{H}^2(\mathbb{R})$ if ε is as given in the theorem (see Lemma 3.2). Hence, if $(\alpha,\beta) \in \Lambda + B_\varepsilon(0)$, say $(\alpha,\beta) = \lambda + (a,b)$ with $\lambda \in \Lambda$ and $(a,b) \in B_\varepsilon(0)$, then (see Lemma 2.3)

$$\|(I-\mathbb{P})\pi(\alpha,\beta)g\|_{L^2} = \|(I-\mathbb{P})\pi(a,b)g\|_{L^2} \geq \frac{\pi/2}{\sqrt{\|g'\|_{L^2}^2 + \|2\pi i X g\|_{L^2}^2}} \cdot \|(a,b)\|_2\,.$$

This proves the theorem. □

Remark. In the case $g \in \mathbb{H}^1(\mathbb{R})$, the value of ε in Theorem 5.4 depends on the convergence to zero of the following quantities (see the proof of Lemma 3.2):

$$\int_{|x|>b} x^2|g(x)|^2\,dx, \quad \int_{|x|>b} |g'(x)|^2\,dx \quad \text{and} \quad \int_{|\omega|>b} \omega^2|\widehat{g}(\omega)|^2\,d\omega \quad \text{as } b \to \infty.$$

Note that the lattice Λ in Theorem 5.4 is not necessarily of rational density. The following corollary suggests that the rational density condition of Λ in Theorems 1.2 and 1.3 might be redundant.

Corollary 5.5. *Let $g \in \mathbb{H}^1(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g, Λ) is an orthonormal basis of its closed linear span $\mathcal{G}(g, \Lambda)$. Then there exists an $N \in \mathbb{N}$ such that $\pi(\mu)g \notin \mathcal{G}(g, \Lambda)$ for all $\mu \in \mathbb{R}^2 \backslash \frac{1}{N}\Lambda$; that is, $\mathcal{G}(g, \Lambda)$ is invariant only under time-frequency shifts with parameters in a subset of $\frac{1}{N}\Lambda$.*

Proof. This follows by combining Theorem 5.4 with [5, Lemma 3.1]. □

A. Auxiliary results

A.1. Matrix multiplication operators

Let (Ω, Σ, μ) be a measure space. To avoid trivialities, we assume that there exists $S \in \Sigma$ with $0 < \mu(S) < \infty$. Now, let $B : \Omega \to \mathbb{C}^{n\times m}$ be a measurable matrix-valued function. Then the multiplication operator

$$M_B : \quad \mathrm{dom}(M_B) \subset L^2(\Omega, \mathbb{C}^m) \to L^2(\Omega, \mathbb{C}^n)$$

is defined by

$$(M_B f)(\omega) := B(\omega)f(\omega), \qquad \omega \in \Omega,\ f \in \mathrm{dom}(M_B),$$

where

$$\mathrm{dom}(M_B) := \left\{ f \in L^2(\Omega; \mathbb{C}^m) \colon \int \|B(\omega)f(\omega)\|^2_{\mathbb{C}^n}\,d\mu(\omega) < \infty \right\}.$$

It is easy to see that the operator M_B is bounded if and only if each entry of $B(\cdot)$ is essentially bounded as a function on Ω, if and only if $\mathrm{dom}(M_B) = L^2(\Omega; \mathbb{C}^m)$.

Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded self-adjoint operator in a Hilbert space $\mathcal{H}$. Then for any continuous, real-valued function $\varphi \in C(\sigma(A); \mathbb{R})$, the operator $\varphi(A)$ is defined by $\varphi(A) := \lim_{n\to\infty} p_n(A)$, where $(p_n)_{n\in\mathbb{N}}$ is a sequence of real-valued polynomials converging uniformly to φ on $\sigma(A) \subset \mathbb{R}$ and the limit is taken with respect to the operator norm. Since $\|p(A)\| = \|p\|_{C(\sigma(A))}$ for polynomials p, this definition is meaningful. One then has $\|\varphi(A)\| = \|\varphi\|_{C(\sigma(A))}$ and $\sigma(\varphi(A)) = \{\varphi(\lambda) : \lambda \in \sigma(A)\}$. Furthermore, $\varphi(A)$ is self-adjoint for all $\varphi \in C(\sigma(A); \mathbb{R})$, since this is easily seen to hold for all polynomials p_n. For more details on this *continuous functional calculus*, see [20, Section VII.1].

For the case $n = m$, the next lemma connects the spectral properties of the multiplication operator M_B to those of the matrices $B(\omega)$, $\omega \in \Omega$.

Lemma A.1. Let $B : \Omega \to \mathbb{C}^{n\times n}$ be a measurable, essentially bounded matrix-valued function satisfying $B(\omega) = B(\omega)^*$ for a.e. $\omega \in \Omega$. Then the following statements hold:

(i) The operator M_B is bounded and self-adjoint.
(ii) For a.e. $\omega \in \Omega$ we have

$$\sigma(B(\omega)) \subset \sigma(M_B).$$

(iii) For every set $N \subset \Omega$ of zero measure,

$$\sigma(M_B) \subset \overline{\bigcup_{\omega \in \Omega \setminus N} \sigma(B(\omega))}\,.$$

(iv) For every function $\varphi \in C(\sigma(M_B); \mathbb{R})$ we have

$$\varphi(M_B) = M_{\varphi(B)},$$

where $(\varphi(B))(\omega) := \varphi(B(\omega))$ is well-defined for almost all $\omega \in \Omega$.

Proof. Part (i) follows easily from $\langle B(\omega) f(\omega), g(\omega)\rangle_{\mathbb{C}^n} = \langle f(\omega), B(\omega) g(\omega)\rangle_{\mathbb{C}^n}$, which holds for almost all $\omega \in \Omega$.

For the proofs of both (ii) and (iii), we will use [16, Proposition 1], which shows for any $\lambda \in \mathbb{C}$ that

$$\lambda \in \varrho(M_B) \iff \exists\, \varepsilon > 0 : \mu(\{\omega \in \Omega \colon \sigma(B(\omega)) \cap B_\varepsilon(\lambda) \neq \varnothing\}) = 0\,. \tag{A.1}$$

To prove (ii), let us assume towards a contradiction that the claim is false; that is, the set

$$\Omega_0 := \big\{\omega \in \Omega \colon \sigma(B(\omega)) \cap \varrho(M_B) \neq \varnothing\big\}$$

is *not* a null-set. Since $\varrho(M_B)$ is an open set, we have $\varrho(M_B) = \bigcup_{k\in\mathbb{N}} I_k$ for certain compact sets $I_k \subset \varrho(M_B) \subset \mathbb{C}$. Setting $\Omega_k := \{\omega \in \Omega \colon \sigma(B(\omega)) \cap I_k \neq \varnothing\}$, we then have $\Omega_0 = \bigcup_{k\in\mathbb{N}} \Omega_k$, so that there is some $k \in \mathbb{N}$ for which Ω_k is not a null-set. Let us choose a dense subset $\{\lambda_n \colon n \in \mathbb{N}\}$ of I_k, and define

$$\Omega_{m,n} := \big\{\omega \in \Omega \colon \sigma(B(\omega)) \cap B_{1/m}(\lambda_n) \neq \varnothing\big\} \qquad \text{for } m, n \in \mathbb{N}\,.$$

By density, we have $I_k \subset \bigcup_{n\in\mathbb{N}} B_{1/m}(\lambda_n)$ for every $m \in \mathbb{N}$, and hence $\Omega_k \subset \bigcup_{n\in\mathbb{N}} \Omega_{m,n}$. Thus, for each $m \in \mathbb{N}$, there is some $n_m \in \mathbb{N}$ such that Ω_{m,n_m} is not a null-set.

Since $(\lambda_{n_m})_{m\in\mathbb{N}}$ is a sequence in the compact set I_k, there is a subsequence $(\lambda_{n_{m_\ell}})_{\ell\in\mathbb{N}}$ such that $\lambda_{n_{m_\ell}} \to \lambda \in I_k \subset \varrho(M_B)$ as $\ell \to \infty$. By (A.1), there is some $\varepsilon > 0$ such that $\Theta := \{\omega \in \Omega \colon \sigma(B(\omega)) \cap B_\varepsilon(\lambda) \neq \varnothing\}$ is a null-set. But for $\ell \in \mathbb{N}$ large enough, we have $\frac{1}{m_\ell} + |\lambda_{n_{m_\ell}} - \lambda| < \varepsilon$, and hence $B_{1/m_\ell}(\lambda_{n_{m_\ell}}) \subset B_\varepsilon(\lambda)$, which shows that $\Omega_{m_\ell, n_{m_\ell}} \subset \Theta$ is a null-set. This is the desired contradiction.

To prove (iii) let $\lambda \in \sigma(M_B)$ and let $N \subset \Omega$ be of zero measure. If $k \in \mathbb{N}$ is arbitrary, then by (A.1), the set $\{\omega \in \Omega : \sigma(B(\omega)) \cap B_{1/k}(\lambda) \neq \varnothing\}$ does not have measure zero, and thus has non-empty intersection with $\Omega\setminus N$. Hence, we can pick $\omega_k \in \Omega\setminus N$ and $\lambda_k \in \sigma(B(\omega_k))$ such that $|\lambda_k - \lambda| < 1/k$. This proves the inclusion in (iii).

Part (iv) is obvious for polynomials φ. Given general $\varphi \in C(\sigma(M_B); \mathbb{R})$, we can approximate φ uniformly on $\sigma(M_B) \subset \mathbb{R}$ by polynomials p_n. Then $\varphi(M_B) - p_n(M_B)$ converges to zero in operator norm, and $p_n(M_B) = M_{p_n(B)}$. Hence, we see for every $f \in L^2(\Omega; \mathbb{C}^n)$ that $\varphi(M_B) f = \lim_{n\to\infty} M_{p_n(B)} f$. But by (ii), we have $\sigma(B(\omega)) \subset \sigma(M_B)$ for almost every $\omega \in \Omega$. For each such $\omega \in \Omega$,

$$\|p_n(B(\omega)) - \varphi(B(\omega))\| = \|p_n - \varphi\|_{C(\sigma(B(\omega)))} \leq \|p_n - \varphi\|_{C(\sigma(M_B))}\,.$$

Thus, we see $[M_{p_n(B)} f](\omega) = p_n(B(\omega)) f(\omega) \to \varphi(B(\omega)) f(\omega) = [M_{\varphi(B)} f](\omega)$ for almost every ω, for every $f \in L^2(\Omega; \mathbb{C}^n)$. Since also have $M_{p_n(B)} f \to \varphi(M_B) f$ with convergence in $L^2(\Omega, \mathbb{C}^n)$, this implies $\varphi(M_B) f = M_{\varphi(B)} f$, as claimed. □

A.2. Operators with closed range and their pseudo-inverse

In this subsection, we review the notion of the pseudo-inverse of an operator with closed range and some of its elementary properties. All of these properties are well-known in general; yet, as some readers might not be familiar with them we decided to include the essentials. Throughout this subsection $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{L}$ denote Hilbert spaces.

Lemma A.2. Let $A : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator. Then

$$(\ker A)^\perp = \overline{\operatorname{ran} A^*}. \tag{A.2}$$

Moreover, the following statements are equivalent:

(a) $\operatorname{ran} A$ is closed in $\mathcal{K}$.
(b) $\operatorname{ran}(AA^*)$ is closed in $\mathcal{K}$.
(c) $\operatorname{ran}(A^*A)$ is closed in $\mathcal{H}$.
(d) $\operatorname{ran} A^*$ is closed in $\mathcal{H}$.
(e) $\sigma_1(A) > 0$.
(f) $\sigma_1(A^*) > 0$.

In one of these properties holds, then the following identities hold:

$$\operatorname{ran}(AA^*) = \operatorname{ran} A, \quad \operatorname{ran}(A^*A) = \operatorname{ran} A^*, \quad \text{and} \quad \sigma_1(A) = \sigma_1(A^*). \tag{A.3}$$

Proof. The identity (A.2) is a simple exercise (see [14, Theorem 58.2]).

For the equivalence of (a)–(f), we refer to [19, Theorem 2].

Next, if (a)–(f) hold, then Equation (A.2) shows $(\ker A)^\perp = \operatorname{ran}(A^*)$. This implies $\operatorname{ran} A = \operatorname{ran}(A|_{(\ker A)^\perp}) = \operatorname{ran}(A|_{\operatorname{ran} A^*}) = \operatorname{ran}(AA^*)$, which proves the first part of Equation (A.3). The second part follows by applying the first part to A^* instead of A.

The last identity in (A.3) follows directly from the definition of σ_1 and the well-known Jacobson lemma which states that for arbitrary bounded linear operators $S : \mathcal{H} \to \mathcal{K}$ and $T : \mathcal{K} \to \mathcal{H}$ we have $\sigma(ST)\backslash\{0\} = \sigma(TS)\backslash\{0\}$. It can indeed be easily seen that $\lambda \in \varrho(TS)\backslash\{0\}$ implies $\lambda \in \varrho(ST)$, by virtue of the identity

$$(ST - \lambda I_{\mathcal{K}})^{-1} = \frac{1}{\lambda}\left[S(TS - \lambda I_{\mathcal{H}})^{-1}T - I_{\mathcal{K}}\right].$$

By symmetry, this implies $\varrho(TS)\backslash\{0\} = \varrho(ST)\backslash\{0\}$. □

Lemma A.3. A bounded operator $A : \mathcal{K} \to \mathcal{H}$ is bounded below (meaning that there is $c > 0$ with $\|Ax\|_{\mathcal{H}} \geq c\,\|x\|_{\mathcal{K}}$ for all $x \in \mathcal{K}$) if and only if $A^*A : \mathcal{K} \to \mathcal{K}$ is bounded below.

Furthermore, a bounded *self-adjoint* operator $T : \mathcal{H} \to \mathcal{H}$ is bounded below if and only if T is boundedly invertible.

Proof. Using the bounded inverse theorem, it is easy to see that a bounded operator T between two Hilbert spaces is bounded below if and only if $\ker T = \{0\}$ and if $\operatorname{ran} T$ is closed. Lemma A.2 shows that $\operatorname{ran} A$ is closed if and only if $\operatorname{ran}(A^*A)$ is closed. Since furthermore $\ker A = \ker(A^*A)$, we obtain the first claim.

For the second part of the claim, let $T : \mathcal{H} \to \mathcal{H}$ be bounded, self-adjoint, and bounded below. As seen above, this implies that $\operatorname{ran} T$ is closed and that $\ker T = \{0\}$. Therefore, Equation (A.2) shows $\mathcal{H} = (\ker T)^\perp = \overline{\operatorname{ran} T^*} = \operatorname{ran} T$. Hence, $T : \mathcal{H} \to \mathcal{H}$ is bijective, so that the bounded inverse theorem shows that T is boundedly invertible. It is clear that if T is boundedly invertible, then T is bounded below. □

The next lemma follows directly from [7, Cor. 5.5.2 and Cor. 5.5.3].

Lemma A.4. A Bessel sequence $(\varphi_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ is a *frame sequence* if and only if its *analysis operator* $\mathbf{A} : \mathcal{H} \to \ell^2(I), f \mapsto \big(\langle f, \varphi_i\rangle\big)_{i\in I}$ has closed range.

Let $A : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator with closed range. Then the operator

$$A_0 : (\ker A)^\perp \to \operatorname{ran} A, \qquad x \mapsto A\,x, \tag{A.4}$$

is boundedly invertible by the bounded inverse theorem. Hence, the *pseudo-inverse*

$$A^\dagger := \iota_{(\ker A)^\perp} \circ A_0^{-1} \circ P_{\operatorname{ran} A}$$

of A defines a bounded linear operator from $\mathcal{K}$ to $\mathcal{H}$. Here, $\iota_{(\ker A)^\perp}$ is the inclusion map $(\ker A)^\perp \to \mathcal{H}$, $x \mapsto x$.

In the following lemma we list some of the properties of the pseudo-inverse.

Lemma A.5. Let $A : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator with closed range. Then the following hold:

(i) $A^\dagger A = P_{(\ker A)^\perp}$.
(ii) $AA^\dagger = P_{\operatorname{ran} A}$.
(iii) $(A^\dagger)^* = (A^*)^\dagger$.
(iv) $(A^*A)^\dagger A^* = A^\dagger = A^*(AA^*)^\dagger$.

Proof. Properties (i)–(iii) can be found in [7, Lemma 2.5.2].

For the first identity in (iv), we refer to [9, Theorem 1]. The remaining identity follows from the first one and (iii) by applying the first part of (iv) on the right-hand side of the identity $A^\dagger = \big((A^*)^\dagger\big)^*$. □

Lemma A.6. Let $A : \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator with closed range and set $c := \sigma_1(A)$. Then $\sigma(A) \subset \{0\} \cup (\mathbb{R}\backslash(-c, c))$ and $A^\dagger = \varphi(A)$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is defined by $\varphi(t) = \frac{1}{t}$ for $t \neq 0$ and $\varphi(0) = 0$.

Remark. Since 0 is an isolated point of $\sigma(A) \subset \{0\} \cup \big(\mathbb{R}\backslash(-c, c)\big)$, $\varphi|_{\sigma(A)}$ is continuous.

Proof. Lemma A.2 shows $c = \sigma_1(A) > 0$. By definition of $\sigma_1(A)$ (see Equation (2.1)), we thus see that $A^2 = A^*A$ satisfies $\sigma(A^2) \subset \{0\} \cup [c^2, \infty)$. As $\sigma(A^2) = \{\lambda^2 : \lambda \in \sigma(A)\}$ and since $\sigma(A) \subset \mathbb{R}$ because of $A^* = A$, it follows that $\sigma(A) \subset \{0\} \cup (\mathbb{R}\backslash(-c, c))$. In particular, this entails that $\varphi|_{\sigma(A)}$ is continuous.

To prove $A^\dagger = \varphi(A)$, define $\psi := \mathbb{1}_{\{0\}}$ and note $\psi \in C(\sigma(A); \mathbb{R})$ since 0 is an isolated point of $\sigma(A)$ (or even $0 \notin \sigma(A)$). Since $\psi^2 = \psi$, we see that $P := \psi(A)$ satisfies $P^2 = P = P^*$, so that $P = P_V$ is the orthogonal projection onto a closed subspace $V \subset \mathcal{H}$. For $x \in \ker A$ we have $Ax = 0x$, so that [20, Theorem VII.1(d)] shows $Px = \psi(A)x = \psi(0)x = x$; hence, $\ker A \subset V$. Conversely, we have $\mathrm{id}_{\sigma(A)} \cdot \psi \equiv 0$ and hence $0 = (\mathrm{id}_{\sigma(A)} \cdot \psi)(A) = AP$, which shows $V = \operatorname{ran} P \subset \ker A$ and hence $V = \ker A$.

Next, observe that $\varphi \cdot \mathrm{id}_{\sigma(A)} = 1 - \psi$, whence $\varphi(A)A = \mathrm{id}_{\mathcal{H}} - P = P_{V^\perp} = A^\dagger A$, where the last step used Lemma A.5(a). Hence, $\varphi(A) = A^\dagger$ on $\operatorname{ran} A$. Finally, we have $\varphi(A)P_V = (\varphi \cdot \psi)(A) = 0$, meaning $\varphi(A) = 0 = A^\dagger$ on $V = \ker A = (\operatorname{ran} A)^\perp$. Overall, this shows $\varphi(A) = A^\dagger$, as claimed. □

Corollary A.7. Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded, self-adjoint operator with closed range, and let $U : \mathcal{K} \to \mathcal{H}$ be unitary. Then $U^*AU : \mathcal{K} \to \mathcal{K}$ is also bounded and self-adjoint with closed range, and we have $(U^*AU)^\dagger = U^*A^\dagger U$.

Proof. It is clear that U^*AU is bounded and self-adjoint with closed range. Furthermore, a direct calculation shows $p(U^*AU) = U^*p(A)U$ for every polynomial $p \in \mathbb{R}[x]$. By definition of the continuous spectral calculus, we thus get $\varphi(U^*AU) = U^*\varphi(A)U$ for all $\varphi \in C(\sigma(A); \mathbb{R})$, where we note $\sigma(A) = \sigma(U^*AU)$. Now, the claim follows from Lemma A.6. □

A.3. Some properties of Sobolev functions

A.3.1. Essentially bounded (matrix-valued) Sobolev functions

Our main objective in this subsection is to prove that the space of matrix-valued functions with all entries in $H^1(\Omega) \cap L^\infty(\Omega)$ is stable under matrix multiplication and inversion. For this, the following lemma will be crucial.

Lemma A.8 ([3, Cor. 2.7])**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $\gamma : \mathbb{C} \to \mathbb{C}$ be a Lipschitz continuous map. In case of $\boldsymbol{\lambda}(\Omega) = \infty$, assume additionally that $\gamma(0) = 0$. If $f \in H^1(\Omega)$, then $\gamma \circ f \in H^1(\Omega)$.

Lemma A.9. Let $\Omega \subset \mathbb{R}^n$ be open and let $f, g \in H^1(\Omega) \cap L^\infty(\Omega)$. Then:

(a) $f \cdot g \in H^1(\Omega) \cap L^\infty(\Omega)$.
(b) If $\operatorname{essinf} |f| > 0$, then also $1/f \in H^1(\Omega) \cap L^\infty(\Omega)$.

Proof. (a) Clearly, $fg \in L^2(\Omega) \cap L^\infty(\Omega)$. Further, [2, Section 4.25] shows that the weak derivatives of fg exist and satisfy

$$\partial_j(fg) = (\partial_j f) \cdot g + f \cdot (\partial_j g) \quad \text{for } j \in \{1, \dots, n\}.$$

As $\partial_j f, \partial_j g \in L^2(\Omega)$ and $f, g \in L^\infty(\Omega)$ it follows that $\partial_j(fg) \in L^2(\Omega)$.

(b) Let $r := \operatorname{essinf} |f| > 0$. We trivially have $1/f \in L^\infty(\Omega)$. Note that $\boldsymbol{\lambda}(\Omega) < \infty$ since $f \in H^1(\Omega) \subset L^2(\Omega)$ and $|f(x)| \geq r > 0$ almost everywhere. Let $B := \{z \in \mathbb{C} \colon |z| < r\}$ and define $\gamma_0 : \mathbb{C} \setminus B \to \mathbb{C}, z \mapsto z^{-1}$. Then γ_0 is well-defined and Lipschitz continuous, since $|z^{-1} - w^{-1}| = \left|\frac{w-z}{zw}\right| \leq r^{-2}|w - z|$ for $z, w \in \mathbb{C} \setminus B$. Now, [10, Theorem 1 in Section 3.1.1] shows that there exists a Lipschitz continuous extension $\gamma : \mathbb{C} \to \mathbb{C}$ of γ_0. Since $|f(x)| \geq r$ almost everywhere, we have $\gamma \circ f = \gamma_0 \circ f = 1/f$ almost everywhere, and Lemma A.8 shows $1/f = \gamma \circ f \in H^1(\Omega)$. □

In the following we denote by $H^1(\Omega; \mathbb{C}^{k\times\ell})$ the space of all matrix-valued functions $A : \Omega \to \mathbb{C}^{k\times\ell}$ for which each component function is in $H^1(\Omega)$. We similarly define $L^p(\Omega; \mathbb{C}^{k\times\ell})$ for $p \in [1, \infty]$.

Lemma A.10. Let $\Omega \subset \mathbb{R}^n$ be open and let $A \in H^1(\Omega; \mathbb{C}^{k\times\ell}) \cap L^\infty(\Omega; \mathbb{C}^{k\times\ell})$ and $B \in H^1(\Omega; \mathbb{C}^{\ell\times m}) \cap L^\infty(\Omega; \mathbb{C}^{\ell\times m})$. Then the following statements hold:

(a) $AB \in H^1(\Omega; \mathbb{C}^{k\times m}) \cap L^\infty(\Omega; \mathbb{C}^{k\times m})$.
(b) If $k = \ell$ and $\operatorname*{essinf}_{x\in\Omega} \sigma_0(A(x)) > 0$, then $A^{-1} \in H^1(\Omega; \mathbb{C}^{k\times k}) \cap L^\infty(\Omega; \mathbb{C}^{k\times k})$.
(c) If $\operatorname*{essinf}_{x\in\Omega} \sigma_0(B(x)) > 0$, then $B^\dagger \in H^1(\Omega; \mathbb{C}^{m\times\ell}) \cap L^\infty(\Omega; \mathbb{C}^{m\times\ell})$.

Proof. Statement (a) follows from Lemma A.9 (a), since $(AB)_{j,n} = \sum_t A_{j,t} B_{t,n}$. For (b) we first observe that Leibniz's formula

$$\det A(x) = \sum_{\sigma \in S_k} \Big[\operatorname{sign}(\sigma) \prod_{j=1}^{k} A_{\sigma(j),j}(x) \Big]$$

and Lemma A.9 (a) yield $\det A \in H^1(\Omega) \cap L^\infty(\Omega)$. Now, the condition on A implies that $A(x)$ is invertible for a.e. $x \in \Omega$ so that $A(x)^{-1}$ indeed exists for a.e. $x \in \Omega$. Moreover, for a.e. $x \in \Omega$, for the smallest eigenvalue $\lambda(x)$ of $A(x)^* A(x)$ we have that $\lambda(x) \geq c := \operatorname{essinf}(\sigma_0(A))^2 > 0$. Therefore, we conclude that

$$|\det A(x)|^2 = \det(A(x)^* A(x)) \geq \lambda(x)^k \geq c^k$$

for a.e. $x \in \Omega$. Hence, Lemma A.9 (b) shows that $(\det A)^{-1} \in H^1(\Omega) \cap L^\infty(\Omega)$. Also,

$$\|A(x)^{-1}\|^2 = \|[A(x)^* A(x)]^{-1}\| \leq \frac{1}{c}$$

for a.e. $x \in \Omega$ implies that $A^{-1} \in L^\infty(\Omega; \mathbb{C}^{k\times k})$. Finally, $A^{-1} \in H^1(\Omega; \mathbb{C}^{k\times k})$ follows from Lemma A.9 (a), combined with the so-called *cofactor formula* for the inverse of a matrix (see for instance [15, Equations (5-22) and (5-23)]). It states for $A \in \mathbb{C}^{k\times k}$ with $k > 1$ and $\det A \neq 0$ that

$$A^{-1} = \frac{\operatorname{adj} A}{\det A} \qquad \text{with} \quad (\operatorname{adj} A)_{i,j} = (-1)^{i+j} \cdot \det A^{(j,i)}\,,$$

where $A^{(j,i)}$ is obtained from A by deleting its j-th row and its i-th column. In the remaining case $k = 1$, we have $A^{-1} = (\det A)^{-1} \in H^1(\Omega) \cap L^\infty(\Omega)$ as well.

The condition on B in (c) implies that $B(x)^* B(x)$ is invertible for a.e. $x \in \Omega$ with $\operatorname{essinf}_{x\in\Omega} \sigma_0(B(x)^* B(x)) = \operatorname{essinf}_{x\in\Omega} \sigma_0(B(x))^2 > 0$. The claim now follows from (a), (b), and the identity $B(x)^\dagger = (B(x)^* B(x))^{-1} B(x)^*$ (see Part (iv) of Lemma A.5). □

A.3.2. A certain property of the space $\mathbb{H}^2(\mathbb{R})$

Lemma A.11. If $g \in \mathbb{H}^2(\mathbb{R})$, then $Xg' \in L^2(\mathbb{R})$ with the estimate

$$\begin{aligned}\|X\, g'\|_{L^2} &\leq 45 \cdot \left(\|g''\|_{L^2}^2 + \|X^2\, g\|_{L^2}^2 + \|g'\|_{L^2}^2\right)^{1/2} \\ &\leq 45 \cdot \left((1+4\pi^2)\, \|g''\|_{L^2}^2 + \|X^2\, g\|_{L^2}^2 + 4\pi^2\, \|g\|_{L^2}^2\right)^{1/2}.\end{aligned}$$

Proof. It follows from [1, Lemma 5.4] that for any $\eta > 0$ and $f \in C^2([0,\eta])$,

$$|f'(0)|^2 \leq \frac{C}{\eta} \cdot \left(\eta^2 \cdot \int_0^\eta |f''(t)|^2\, dt + \eta^{-2} \cdot \int_0^\eta |f(t)|^2\, dt\right),$$

where $C := 2 \cdot 9^2$. One can see that this remains true for $f \in H^2\big((0,\eta)\big)$, by a density argument since $H^2\big((0,\eta)\big) \hookrightarrow C^1([0,\eta])$ (see for instance [1, Thm. 4.12, Part II]).

Given $g \in \mathbb{H}^2(\mathbb{R})$ and $x \in [1,\infty)$, we can apply the above estimate to the function $t \mapsto g(x+t)$ to obtain

$$|g'(x)|^2 \leq C \cdot \fint_0^\eta \left(\eta^2 \cdot |g''(x+t)|^2 + \eta^{-2} \cdot |g(x+t)|^2\right) dt\,,$$

where we denote by $\fint_\Omega f(x)\, dx = \frac{1}{\boldsymbol{\lambda}(\Omega)} \int_\Omega f(x)\, dx$ the *average* of f over Ω, with $\boldsymbol{\lambda}(\Omega)$ denoting the Lebesgue measure of Ω.

Now, fix $n \in \mathbb{N}_0$ for the moment, and let $x \in [2^n, 2^{n+1})$. If we set $\eta = 2^{-n}$, then $\frac{2^n}{\eta} = 2^{2n} \leq x^2 \leq (x+t)^2$ for all $t \in [0,\eta]$. Therefore,

$$|x \cdot g'(x)|^2 \leq 4 \cdot 2^{2n} \cdot |g'(x)|^2 \leq 4C \fint_0^{2^{-n}} \left(|g''(x+t)|^2 + |(x+t)^2 \cdot g(x+t)|^2\right) dt\,. \quad \text{(A.5)}$$

For brevity, set $F(y) := |g''(y)|^2 + |y^2 \cdot g(y)|^2$. Then, for any $t \in [0, 2^{-n}] \subset [0,1]$, we have $2^{n+1} + t \leq 2^{n+2}$ and hence $\int_{2^n}^{2^{n+1}} F(x+t)\, dx \leq \int_{2^n}^{2^{n+2}} F(y)\, dy$. By combining this observation with the trivial estimate $\fint_\Omega G(t)\, dt \leq \|G\|_{L^\infty(\Omega)}$, and by integrating Equation (A.5) over $x \in [2^n, 2^{n+1})$, we arrive at

$$\int_{2^n}^{2^{n+1}} |x \cdot g'(x)|^2\, dx \leq 4C \cdot \fint_0^{2^{-n}} \int_{2^n}^{2^{n+1}} F(x+t)\, dx\, dt \leq 4C \cdot \int_{2^n}^{2^{n+2}} F(y)\, dy\,.$$

Summing over $n \in \mathbb{N}_0$, we conclude that

$$\begin{aligned}\int_1^\infty |x \cdot g'(x)|^2 \, dx &= \sum_{n=0}^\infty \int_{2^n}^{2^{n+1}} |x \cdot g'(x)|^2 \, dx \\ &\leq 4C \cdot \int_1^\infty F(y) \cdot \sum_{n=0}^\infty \mathbb{1}_{(2^n, 2^{n+2})}(y) \, dy \\ &\leq 12C \cdot \left(\|g''\|_{L^2((1,\infty))}^2 + \|X^2 \, g\|_{L^2((1,\infty))}^2 \right).\end{aligned} \tag{A.6}$$

Here the last step used that $\sum_{n=0}^\infty \mathbb{1}_{(2^n, 2^{n+2})}(y) \leq 3$; indeed, if $2^n < y < 2^{n+2}$, then each $k \in \mathbb{Z}$ for which also $2^k < y < 2^{k+2}$ satisfies $2^n < 2^{k+2}$ and $2^k < 2^{n+2}$, so that $k \in \{n-1, n, n+1\}$.

By applying estimate (A.6) to $h : \mathbb{R} \to \mathbb{C}, x \mapsto g(-x)$ instead of g, we easily get $\int_{-\infty}^{-1} |x \cdot g'(x)|^2 \, dx \leq 12C \cdot \left(\|g''\|_{L^2((-\infty,-1))}^2 + \|X^2 \, g\|_{L^2((-\infty,-1))}^2 \right)$. Adding this to (A.6) and using the trivial estimate $\int_{-1}^1 |x \cdot g'(x)|^2 \, dx \leq \|g'\|_{L^2}^2$, we finally arrive at

$$\int_{\mathbb{R}} |x \cdot g'(x)|^2 \, dx \leq \|g'\|_{L^2}^2 + 12C \cdot \left(\|g''\|_{L^2}^2 + \|X^2 \, g\|_{L^2}^2 \right).$$

This easily implies the first part of the stated estimate.

For the last part, recall that $\mathcal{F}[g'](\xi) = 2\pi i \xi \, \widehat{g}(\xi)$ and $\mathcal{F}[g''](\xi) = (2\pi i \xi)^2 \, \widehat{g}(\xi)$. Thanks to Plancherel's theorem and the elementary estimate $|\xi|^2 \leq 1 + |\xi|^4$, we thus see

$$\begin{aligned}\|g'\|_{L^2}^2 &= \int_{\mathbb{R}} |2\pi\xi \cdot \widehat{g}(\xi)|^2 \, d\xi \leq (2\pi)^2 \cdot \int_{\mathbb{R}} |\widehat{g}(\xi)|^2 + |(2\pi i \xi)^2 \, \widehat{g}(\xi)|^2 \, d\xi \\ &= (2\pi)^2 \cdot \left(\|g\|_{L^2}^2 + \|g''\|_{L^2}^2 \right).\end{aligned}$$

Together with the first part of the lemma, this implies the second part. □

A.3.3. Sobolev functions on slices and the AC-property

Let $A \subset \mathbb{R}^n$ be Borel measurable, where $n > 1$. For $i \in \{1, \dots, n\}$ and $x \in \mathbb{R}^{n-1}$ we define the following Borel measurable subset of $\mathbb{R}$:

$$A_{i,x} = \{t \in \mathbb{R} : (x_1, \dots, x_{i-1}, t, x_i, \dots, x_{n-1}) \in A\}.$$

Note that $A_{i,x}$ is open if A is so. The following lemma is an easy consequence of Fubini's theorem.

Lemma A.12. A Borel set $N \subset \mathbb{R}^n$ has measure zero if and only if for some (and then all) $i \in \{1, \dots, n\}$ and a.e. $x \in \mathbb{R}^{n-1}$ the set $N_{i,x}$ has measure zero in $\mathbb{R}$.

We say that a function $h : U \to \mathbb{C}$, where $U \subset \mathbb{R}$ is open, is *locally absolutely continuous (LAC)* on U if it is LAC on each connected component of U; this is equivalent to h being LAC on each open subinterval of U. Here, a function $f : I \to \mathbb{C}$ with an open interval $I \subset \mathbb{R}$ is called locally absolutely continuous if there is a function $g \in L^1_{\mathrm{loc}}(I)$ such that $f(x) - f(y) = \int_y^x g(t) \, dt$ for all $x, y \in I$. In particular, each LAC function is continuous.

Definition A.13. Let $\Omega \subset \mathbb{R}^n$ be an open set. A (pointwise defined) function $f : \Omega \to \mathbb{C}$ is said to have the *AC-property* (on Ω), if for each $i \in \{1, \dots, n\}$ and almost all $x \in \mathbb{R}^{n-1}$ the function

$$f_{i,x} : \Omega_{i,x} \to \mathbb{C}, \quad t \mapsto f(x_1, \dots, x_{i-1}, t, x_i, \dots, x_{n-1})$$

is LAC on $\Omega_{i,x}$.

Note that the classical partial derivatives $\partial_i f$ of a function $f : \Omega \to \mathbb{C}$ having the AC-property exist a.e. on Ω by [11, Theorem 3.35] and Lemma A.12.

Lemma A.14. Let $\Omega \subset \mathbb{R}^n$ be open and let $f \in W^{1,1}_{\mathrm{loc}}(\Omega)$. Then there is a representative $g : \Omega \to \mathbb{C}$ of f which has the AC-property on Ω. In particular, we have $\partial_i g = D_i f$ a.e. on Ω, $i = 1, \dots, n$. Here, $D_i f$ denotes the weak derivative of f.

Proof. Let $\Omega^{(0)} := \varnothing$ and $\Omega^{(k)} := \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > 1/k\} \cap (-k, k)^n$, $k \in \mathbb{N}$. Then each $\Omega^{(k)}$ is open in $\mathbb{R}^n$, $\overline{\Omega^{(k)}} \subset \Omega$ is compact, $\Omega^{(k)} \subset \Omega^{(k+1)}$, and $\bigcup_k \Omega^{(k)} = \Omega$. By [18, Thm. 1.41], for each $k \in \mathbb{N}$ there exists a representative $f^{(k)}$ of f which has the AC-property on $\Omega^{(k)}$. It is clear that the function $g : \Omega \to \mathbb{C}$,

$$g := \sum_{k=1}^{\infty} \mathbb{1}_{\Omega^{(k)} \setminus \Omega^{(k-1)}} \cdot f^{(k)}$$

is a representative of f. Let us show that it has the AC-property on Ω.

First of all, for each $k \in \mathbb{N}$ and $i \in \{1, \dots, n\}$ there exists a set $L_i^{(k)} \subset \mathbb{R}^{n-1}$ of measure zero such that $f_{i,x}^{(k)}$ is LAC on $\Omega_{i,x}^{(k)}$ for every $x \in \mathbb{R}^{n-1} \setminus L_i^{(k)}$. Let $L := \bigcup_{i,k} L_i^{(k)}$.

Fix $k \in \mathbb{N}$. Then $f^{(k+1)} = f^{(k)}$ a.e. on $\Omega^{(k)}$. In particular, by Lemma A.12, for each $i \in \{1, \dots, n\}$ there exists a set $M_i^{(k)} \subset \mathbb{R}^{n-1}$ of measure zero such that for all $x \in \mathbb{R}^{n-1} \setminus M_i^{(k)}$ we have that $f_{i,x}^{(k+1)} = f_{i,x}^{(k)}$ a.e. on $\Omega_{i,x}^{(k)}$. Let $M := \bigcup_{i,k} M_i^{(k)}$.

Let $N := L \cup M \subset \mathbb{R}^{n-1}$. Then N is a null-set, and for each $i \in \{1, \dots, n\}$, each $k \in \mathbb{N}$, and each $x \in \mathbb{R}^{n-1} \setminus N$ we have that $f_{i,x}^{(k)}$ is LAC on $\Omega_{i,x}^{(k)}$ and $f_{i,x}^{(k+1)} = f_{i,x}^{(k)}$ on $\Omega_{i,x}^{(k)}$; indeed, since $f_{i,x}^{(k+1)} = f_{i,x}^{(k)}$ almost everywhere on $\Omega_{i,x}^{(k)}$, and since both functions are continuous on the open set $\Omega_{i,x}^{(k)}$, they agree everywhere on $\Omega_{i,x}^{(k)}$. Now, if $i \in \{1, \dots, n\}$, $x \in \mathbb{R}^{n-1} \setminus N$, and if $K \subset \Omega_{i,x}$ is compact, then $\bigcup_k \Omega_{i,x}^{(k)}$ is an open cover of K. Thus, there is some $k = k(i, x, K) \in \mathbb{N}$ such that $K \subset \Omega_{i,x}^{(k)}$ and $g_{i,x} = f_{i,x}^{(k)}$ on K. Therefore, $g_{i,x}$ is LAC on $\Omega_{i,x}$.

For the "in particular"-part, it suffices to prove that $\partial_i g = D_i f$ almost everywhere on every open rectangular cell $R = \prod_{j=1}^n (a_j, b_j)$ satisfying $\overline{R} \subset \Omega$. To see this, set $R_i := \prod_{j \neq i} (a_j, b_j) \subset \mathbb{R}^{n-1}$, and observe that for any $\varphi \in C_c^\infty(R)$ we have

$$\begin{aligned} &\int_R (\partial_i g) \cdot \varphi \, dy + \int_R g \cdot (\partial_i \varphi) \, dx \\ &= \int_{R_i} \int_{a_i}^{b_i} g'_{i,x}(t) \varphi_{i,x}(t) \, dt \, dx + \int_{R_i} \int_{a_i}^{b_i} g_{i,x}(t) \varphi'_{i,x}(t) \, dt \, dx = 0 \,. \end{aligned}$$

Hence, $\int_R (\partial_i g - D_i f) \varphi \, dx = \int_R (f - g) \partial_i \varphi \, dx = 0$ for every $\varphi \in C_c^\infty(R)$. The claim thus follows from the fundamental lemma of the calculus of variations (see for instance [2, Section 4.22]). □

We close with this subsection with a result that generalizes Lemma A.12 in the case $n = 2$ to sections of $\mathbb{R}^2$ that are not necessarily parallel to the coordinate axes.

Lemma A.15. Let $N \subset \mathbb{R}^2$ be a null-set, and let $(a, b) \in \mathbb{R}^2 \setminus \{0\}$. Then there is a null-set $N_0 \subset \mathbb{R}^2$ such that for all $(x, \omega) \in \mathbb{R}^2 \setminus N_0$, we have

$$(x + ta, \omega + tb) \in \mathbb{R}^2 \setminus N \quad \text{for almost all } t \in \mathbb{R}.$$

Remark. The set of $t \in \mathbb{R}$ for which $(x + ta, \omega + tb) \in \mathbb{R}^2 \setminus N$ depends on (x, ω).

Proof. Set $\theta := (a,b) \in \mathbb{R}^2\backslash\{0\}$, and choose $\varrho \in \mathbb{R}^2\backslash\{0\}$ with $\varrho \perp \theta$. Let us define $T : \mathbb{R}^2 \to \mathbb{R}^2, (t,s) \mapsto t\theta + s\varrho$. Note that T is linear and bijective, so that the same holds of T^{-1}. In particular, T and T^{-1} are Lipschitz continuous, and thus map null-sets to null-sets.

Let $\widetilde{N} := T^{-1}N \subset \mathbb{R}^2$. By Lemma A.12, there is a null-set $\widetilde{N}_1 \subset \mathbb{R}$ such that for all $s \in \mathbb{R}\backslash\widetilde{N}_1$, the set $\widetilde{N}_{1,s} = \{t \in \mathbb{R} \colon (t,s) \in \widetilde{N}\}$ is a null-set. Let $N_0 := T(\mathbb{R}\times\widetilde{N}_1)$, and note that $N_0 \subset \mathbb{R}^2$ is indeed a null-set.

We claim that if $(x,\omega) \in \mathbb{R}^2\backslash N_0$, then $(x+ta, \omega+tb) \in \mathbb{R}^2\backslash N$ for almost all $t \in \mathbb{R}$. To see this, let $(x,\omega) \in \mathbb{R}^2\backslash N_0$. This implies $(x,\omega) = T(t_0,s_0)$ for certain $(t_0,s_0) \in \mathbb{R}\times(\mathbb{R}\backslash\widetilde{N}_1)$, so that $\widetilde{N}_{1,s_0}$ is a null-set. Finally, if $t \in \mathbb{R}\backslash(\widetilde{N}_{1,s_0} - t_0)$ (which holds for almost all $t \in \mathbb{R}$), then $t+t_0 \notin \widetilde{N}_{1,s_0}$, which means that $(t+t_0, s_0) \notin \widetilde{N} = T^{-1}N$, and hence $(x+ta, \omega+tb) = T(t+t_0,s_0) \in \mathbb{R}^2\backslash N$, as claimed. □

A.4. Invariance properties of Gabor spaces

Lemma A.16. Let $g \in L^2(\mathbb{R})$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice. Define $\mathcal{G} := \mathcal{G}(g,\Lambda)$ Then $\pi(\lambda)\mathcal{G} \subset \mathcal{G}$ for all $\lambda \in \Lambda$.

Proof. For $\lambda, \lambda' \in \Lambda$, there exists a unimodular constant $c = c(\lambda,\lambda') \in \mathbb{C}$ satisfying $\pi(\lambda)\pi(\lambda') = c\,\pi(\lambda+\lambda')$. Hence, $\pi(\lambda)[\pi(\lambda')g] \in \mathcal{G}$. Since $\mathcal{G}$ is spanned by the elements $\pi(\lambda')g$, $\lambda' \in \Lambda$, this shows $\pi(\lambda) \subset \mathcal{G}$ for all $\lambda \in \Lambda$. □

A.5. Failure of the main result for general elements of $\mathcal{G}(g,\Lambda)$

We close this paper with an example showing that the relation

$$\operatorname{dist}\big(\pi(u,\eta)f, \mathcal{G}(g,\Lambda)\big) \asymp \operatorname{dist}((u,\eta),\Lambda)\cdot\|f\|_{L^2}, \tag{A.7}$$

which holds for $f = g$, does *not* extend to general $f \in \mathcal{G}(g,\Lambda)$. The example is constructed based on a footnote in [6].

Example A.17. Let $\varphi : \mathbb{R} \to \mathbb{R}, x \mapsto e^{-\pi x^2}$ denote the Gaussian. We will repeatedly make use of the following two facts: First, [12, Theorem 7.5.3] shows that if $\alpha,\beta > 0$, then $(\varphi, \alpha\mathbb{Z}\times\beta\mathbb{Z})$ is a frame for $L^2(\mathbb{R})$ if and only if $\alpha\beta < 1$. By Ron-Shen duality (see [12, Theorem 7.4.3]), this implies that $(\varphi, \alpha\mathbb{Z}\times\beta\mathbb{Z})$ is a Riesz sequence (a Riesz basis for its closed linear span) if and only if $\alpha\beta > 1$.

Set $\Lambda := 2\mathbb{Z}\times\frac{2}{3}\mathbb{Z}$ and $\Lambda_0 := \Lambda \cup ((1,0)+\Lambda) = \mathbb{Z}\times\frac{2}{3}\mathbb{Z}$. Then (φ,Λ_0) is a frame for $L^2(\mathbb{R})$ but *not* a Riesz sequence. Thus, the synthesis operator

$$T : \quad \ell^2(\mathbb{Z}^2) \to L^2(\mathbb{R}), \quad (c_{k,\ell})_{k,\ell\in\mathbb{Z}} \mapsto \sum_{k,\ell\in\mathbb{Z}} c_{k,\ell}\ \pi(k, \tfrac{2}{3}\ell)\varphi$$

is surjective, but *not* injective, since otherwise the bounded inverse theorem would imply that T is boundedly invertible, meaning that (φ,Λ_0) is a Riesz basis for $L^2(\mathbb{R})$. In other words, there exist ℓ^2 sequences $c = (c_{m,n})_{m,n\in\mathbb{Z}}$ and $d = (d_{m,n})_{m,n\in\mathbb{Z}}$ with $(c,d) \neq 0$ and

$$\begin{aligned}\sum_{m,n\in\mathbb{Z}} c_{m,n}\,\pi(2m, \tfrac{2}{3}n)\varphi &= \sum_{m,n\in\mathbb{Z}} d_{m,n}\,\pi(2m+1, \tfrac{2}{3}n)\varphi \\ &= \pi(1,0)\Big[\sum_{m,n\in\mathbb{Z}} \widetilde{d}_{m,n}\,\pi(2m, \tfrac{2}{3}n)\varphi\Big],\end{aligned}$$

where $\widetilde{d}_{m,n} := e^{\frac{4}{3}\pi i n} d_{m,n}$ for $m, n \in \mathbb{Z}$. Then $f := \sum_{m,n\in\mathbb{Z}} \widetilde{d}_{m,n}\, \pi(2m, \frac{2}{3}n)\varphi$ satisfies $f \in \mathcal{G}(\varphi, \Lambda)$ and $\pi(1,0)f \in \mathcal{G}(\varphi, \Lambda)$. Now, once we show that $f \neq 0$, we will have disproved (A.7).

To see that $f \neq 0$, we note that (φ, Λ) is a Riesz sequence. If $f = 0$, we would have $\widetilde{d} = 0$ and therefore $d = 0$. In turn, the above identity gives $0 = \sum_{m,n\in\mathbb{Z}} c_{m,n}\pi(2m, \frac{2}{3}n)\varphi$, whence $c = 0$, again since (φ, Λ) is a Riesz sequence. Therefore, $f = 0$ implies $(c, d) = 0$ which is a contradiction.

Acknowledgments

D.G. Lee acknowledges support by the DFG Grants PF 450/6-1 and PF 450/9-1. A. Caragea acknowledges support by the DFG Grant PF 450/11-1. The authors would like to thank Götz E. Pfander and Peter Jung for fruitful discussions. The authors thank the editor and anonymous reviewers for their valuable comments and suggestions which improved the paper.

References

[1] R.A. Adams, J.J.F. Fournier, *Sobolev spaces*, second edition, Elsevier/Academic Press, Amsterdam, 2003.
[2] H.W. Alt, *Linear functional analysis*, Springer-Verlag London, Ltd., London, 2016.
[3] W. Arendt, M. Kreuter, *Mapping theorems for Sobolev spaces of vector-valued functions*, Studia Math. 240 (2018), 275–299.
[4] A. Caragea, D.G. Lee, G.E. Pfander, F. Philipp, *A Balian-Low theorem for subspaces*, J. Fourier Anal. Appl. 25 (2019), 1673–1694.
[5] A. Caragea, D. G. Lee, F. Philipp, F. Voigtlaender, *Time-Frequency Shift Invariance of Gabor Spaces with an S_0-Generator*, arXiv preprint, arXiv:1904.12345.
[6] C. Cabrelli, U. Molter, G.E. Pfander, *Time-frequency shift invariance and the Amalgam Balian-Low theorem*, Appl. Comput. Harmon. Anal. 41 (2016), 677–691.
[7] O. Christensen, *An introduction to frames and Riesz bases*, second edition, Birkhauser/Springer, 2016.
[8] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inf. Theory 36 (1990), 961–1005.
[9] C. A. Desoer, B. H. Whalen, *A note on pseudoinverses*, J. Soc. Indust. Appl. Math. 11 (1963), 442–447.
[10] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992.
[11] G.B. Folland, *Real Analysis*, second edition, John Wiley & Sons, Inc., New York, 1999.
[12] K. Gröchenig, *Foundations of time-frequency analysis*, Birkhäuser, Boston, Basel, Berlin, 2001.
[13] K. Gröchenig, D. Han, C. Heil, G. Kutyniok, *The Balian–Low theorem for symplectic lattices in higher dimensions*, Appl. Comput. Harmon. Anal. 13 (2002), 169–176.
[14] H. Heuser, *Funktionalanalysis*, fourth edition, B. G. Teubner, Stuttgart, 2006.
[15] K. Hoffman, R. Kunze, *Linear algebra*, second edition, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.
[16] A. Holderrieth, *Matrix multiplication operators generating one parameter semigroups*, Semigroup Forum 42 (1991), 155–166.
[17] S. Lang, *Real and Functional Analysis*, third edition, Springer-Verlag, New York, 1993.
[18] J. Malý, W.P. Ziemer, *Fine regularity of solutions of elliptic partial differential equations*, Mathematical Surveys and Monographs, Vol. 51, Americal Mathematical Society, 1997.
[19] M. Ould-Ali, B. Messirdi, *On closed range operators in Hilbert space*, Int. J. Algebra 4 (2010), 953–958.
[20] M. Reed, B. Simon, *Modern methods of mathematical physics*, Vol. 1, Academic Press, 1980.
[21] W.P. Ziemer, *Weakly differentiable functions*, Springer-Verlag, New York, 1989.

A. Caragea: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
Email address: andrei.caragea@gmail.com
URL: http://www.ku.de/?acaragea

D.G. Lee: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
Email address: daegwans@gmail.com
URL: http://www.ku.de/?lee

F. Philipp: Technische Universität Ilmenau, Institute for Mathematics, Weimarer Strasse 25, D-98693 Ilmenau, Germany
Email address: friedrich.philipp@tu-ilmenau.de

F. Voigtlaender: KU Eichstätt-Ingolstadt, Mathematisch-Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany
Email address: felix@voigtlaender.xyz
URL: http://www.ku.de/?voigtlaender
URL: http://voigtlaender.xyz

6.3 Paper 3

A BALIAN–LOW TYPE THEOREM FOR GABOR RIESZ SEQUENCES OF ARBITRARY DENSITY

ANDREI CARAGEA, DAE GWAN LEE, FRIEDRICH PHILIPP, AND FELIX VOIGTLAENDER

Abstract. We consider Gabor Riesz sequences generated by a lattice $\Lambda \subset \mathbb{R}^2$ and a window function $g \in L^2(\mathbb{R})$ which is well localized in both time and frequency. When g belongs to the Feichtinger algebra, we prove that only those time-frequency shifts with parameters from the lattice Λ leave the corresponding Gabor space invariant. This improves on earlier results where only lattices of rational density were considered. A slightly weaker result is proved—again for lattices of general density—under the regularity assumptions of the classical Balian-Low theorem, where both g and its Fourier transform belong to the Sobolev space $H^1(\mathbb{R})$. The proof relies on a combination of methods from time-frequency analysis and the theory of C^*-algebras, specifically the so-called irrational rotation algebra.

1. Introduction

When working with Gabor frames, the window function g should have a good time-frequency localization, so that the frame coefficients faithfully reflect the time-frequency behavior of the analyzed function. The *Feichtinger algebra* $S_0(\mathbb{R}^d)$ [12, 19] is a particularly popular window class. Among other advantages, choosing a window from S_0 ensures that the canonical dual window also belongs to the Feichtinger algebra [15], so that for example the membership of a function $f \in L^2(\mathbb{R}^d)$ in the modulation space $M^{p,q}(\mathbb{R}^d)$ can be characterized in terms of the decay properties of its frame coefficients. One crucial obstruction, however, is that a Gabor system with window belonging to $S_0(\mathbb{R}^d)$ can not form an orthonormal basis—in fact not even a Riesz basis—for $L^2(\mathbb{R}^d)$. We call this phenomenon the S_0 *Balian-Low theorem*; it is a consequence of the Amalgam Balian-Low theorem [2, Theorem 3.2]. The same no-go type result holds for the case where g belongs to the space $\mathbb{H}^1(\mathbb{R}^d)$ consisting of functions in the L^2-Sobolev space $H^1(\mathbb{R}^d)$ whose Fourier transform also belongs to $H^1(\mathbb{R}^d)$. This is the classical Balian-Low theorem; see [14, Theorem 8.4.5] for the case of orthonormal bases, and [10, Theorem 2.3] for the general case.

Yet, even though a Gabor system with $g \in S_0(\mathbb{R}^d)$ cannot form a Riesz basis for all of $L^2(\mathbb{R}^d)$, it might still be a *Riesz sequence*, that is, a Riesz basis for its closed linear span $\mathcal{G}(g, \Lambda)$, at least if $\mathcal{G}(g, \Lambda)$ is a *proper* subspace of $L^2(\mathbb{R}^d)$. In this case, one might wonder about further properties—in addition to being a proper subspace of $L^2(\mathbb{R}^d)$—that the Gabor space $\mathcal{G}(g, \Lambda)$ has to have. One important property in time-frequency analysis is the invariance of $\mathcal{G}(g, \Lambda)$ under time-frequency shifts $T_a M_b$. For lattices Λ *of rational density* and for dimension $d = 1$, it was observed in [4] that if (g, Λ) is a Riesz sequence and if $g \in S_0(\mathbb{R})$, then the set of parameters $(a, b) \in \mathbb{R}^2$ such that $\mathcal{G}(g, \Lambda)$ is invariant under the time-frequency shift $T_a M_b$ is exactly equal to Λ. A multi-dimensional variant of this was derived in [5].

2010 *Mathematics Subject Classification.* Primary: 42C15. Secondary: 42C30, 42C40.

Key words and phrases. Gabor systems; Riesz sequences; Time-frequency shifts; Balian-Low theorem; Ron-Shen duality; Completeness of Gabor systems; Feichtinger algebra; Irrational rotation algebra.

These results generalize the S_0 Balian-Low theorem to subspaces of $L^2(\mathbb{R})$. Indeed, to derive the S_0 Balian-Low theorem from the above result, note that if $\mathcal{G}(g,\Lambda) = L^2(\mathbb{R})$ then $\mathcal{G}(g,\Lambda)$ is invariant under *all* time-frequency shifts, even under those with $(a,b) \notin \Lambda$; hence, g cannot belong to S_0. A corresponding generalization of the *classical* Balian-Low theorem was proved in [6]; a quantitative version can be found in [7].

We emphasize that in all articles [4, 5, 6, 7] it is assumed that the generating lattice Λ has rational density. This restriction is needed in order to utilize the Zak transform which is used extensively in [4, 5, 6, 7]. It is thus natural to ask whether the results in [4] and [6] still hold for lattices with *irrational* density.

In a sense, this question has analogies with the research concerning the regularity of the canonical dual window of a Gabor frame. In 1997 it was shown (see [13, Theorem 3.4]) that if $g \in S_0(\mathbb{R}^d)$ generates a Gabor frame for $L^2(\mathbb{R}^d)$ over a lattice *of rational density*, then the canonical dual window also belongs to $S_0(\mathbb{R}^d)$. It was conjectured in the same article that this property continues to hold for general lattices. Six years later, this conjecture was confirmed by Gröchenig and Leinert [15] by using C^*-algebra methods.

Here, we likewise extend the result in [4] to arbitrary lattices:

Theorem 1.1. *If $g \in S_0(\mathbb{R})$ and $\Lambda \subset \mathbb{R}^2$ is a lattice such that the Gabor system (g,Λ) is a Riesz basis for its closed linear span $\mathcal{G}(g,\Lambda)$, then the time-frequency shifts T_aM_b that leave $\mathcal{G}(g,\Lambda)$ invariant satisfy $(a,b) \in \Lambda$.*

As indicated above, the Zak transform is a powerful tool for analyzing Gabor systems generated by lattices with rational density; yet, it is not of much use in the case of irrational density lattices. Consequently, the methods used in the present paper differ substantially from those in [4, 5, 6, 7]: Instead of applying the Zak transform and thus dealing with functions on $\mathbb{R}^2$, we work directly with the given objects and exploit the rich theory of time-frequency analysis. Along the way, we obtain several new statements related to time-frequency shift invariance that are interesting in their own right.

The proof of Theorem 1.1 consists of several steps. First, for $g \in S_0(\mathbb{R})$ and only assuming that (g,Λ) is a *frame sequence*—that is, a frame for its closed linear span—we prove the following dichotomy:

> *Either (g,Λ) spans all of $L^2(\mathbb{R})$, or the set of $(a,b) \in \mathbb{R}^2$ for which T_aM_b leaves $\mathcal{G}(g,\Lambda)$ invariant is a lattice containing Λ as a sublattice;* (D)

see Theorem 3.4. This result significantly reduces the range of parameters (a,b) that we need to consider. Next, we give a characterization for the invariance of $\mathcal{G}(g,\Lambda)$ under a time-frequency shift T_aM_b with $(a,b) \notin \Lambda$ in terms of the adjoint system of (g,Λ); see Theorem 4.2. This characterization holds for general $g \in L^2(\mathbb{R})$, not only for $g \in S_0(\mathbb{R})$. Combining this characterization with a deep existing result about traces of projections in the so-called *irrational rotation algebra* (see [24, 25]), we arrive at the conclusion of Theorem 1.1.

With Theorem 1.1 established for g in the Feichtinger algebra, it is natural to ask whether the same statement holds in the setting of the *classical* Balian-Low theorem, that is, when g has finite uncertainty product $\big(\int x^2|g(x)|^2\,dx\big)\cdot\big(\int \omega^2|\widehat{g}(\omega)|^2\,d\omega\big) < \infty$, a condition which we simply write as $g \in \mathbb{H}^1$. Unfortunately, we were not able to prove a full-fledged version of Theorem 1.1 for $g \in \mathbb{H}^1$; the best we could do is to show that the dichotomy (D) described above for $g \in S_0$ still holds for $g \in \mathbb{H}^1$.

The outline of the paper is as follows: After recalling the necessary background on Janssen's representation, time-frequency shift invariance, symplectic operators, and the

two spaces $S_0(\mathbb{R})$ and $\mathbb{H}^1$ in Section 2, the paper proper starts in Section 3, where we prove the dichotomy (D) described above, for $g \in S_0(\mathbb{R})+\mathbb{H}^1$. Next, in Section 4 we show that one can reduce to the case of a separable lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$, with an additional time-frequency shift of the form $T_{\alpha/\nu}$ for some $\nu \in \mathbb{N}_{\geq 2}$. For this setting, we then derive a characterization in terms of the adjoint Gabor system. Throughout Section 4, the generating function g is only assumed to be in $L^2(\mathbb{R})$. The paper culminates in Section 5, where we prove Theorem 1.1. Finally, Appendix A contains a short treatise on the irrational rotation algebra and a corresponding result that is crucial for our proof of Theorem 1.1.

2. Preliminaries

For $a, b \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$ we define the operators of translation by a and modulation by b as

$$T_a f(x) := f(x-a) \quad \text{and} \quad M_b f(x) := e^{2\pi i b x} f(x),$$

respectively. Both T_a and M_b are unitary operators on $L^2(\mathbb{R})$ and hence so is the *time-frequency shift*

$$\pi(a,b) := T_a M_b = e^{-2\pi i a b}\, M_b T_a.$$

A *lattice* $\Lambda \subset \mathbb{R}^2$ is any set of the form $\Lambda = A\mathbb{Z}^2$ with an invertible matrix $A \in \mathbb{R}^{2\times 2}$. The *density* of Λ is defined by $d(\Lambda) = |\det A|^{-1}$. Note that $A\mathbb{Z}^2 = \mathbb{Z}^2$ if and only if $A \in \mathbb{Z}^{2\times 2}$ and $\det A = \pm 1$. This will be used heavily in the proof of Proposition 3.1 below.

A lattice Λ is called *separable* if A can be chosen to be diagonal, i.e., $\Lambda = \alpha\mathbb{Z}\times\beta\mathbb{Z}$ with $\alpha, \beta > 0$. The next lemma shows that every lattice can be transformed into a separable one by means of a symplectic matrix; this will be used frequently.

Lemma 2.1. *Let $A \in \mathbb{R}^{2\times 2}$ be a non-singular matrix. Then there exists $C \in \mathbb{R}^{2\times 2}$ with $\det C = 1$ such that CA is diagonal, i.e., $CA\mathbb{Z}^2$ is separable.*

Proof. Write $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and note $\Delta := ad - bc \neq 0$. If $a \neq 0$, choose $C = \left(\begin{smallmatrix} 1+bc/\Delta & -ab/\Delta \\ -c/a & 1 \end{smallmatrix}\right)$. Then a simple calculation yields $\det C = 1$ and $CA = \operatorname{diag}(a, \Delta/a)$. In the case $a = 0$ we have $b \neq 0 \neq c$ as A is non-singular. Then $C := \left(\begin{smallmatrix} -d/b & 1 \\ -1 & 0 \end{smallmatrix}\right)$ satisfies $\det C = 1$ and $CA = \operatorname{diag}(c, -b)$. □

For a subset $M \subset L^2(\mathbb{R})$, we denote its closure by $\overline{M}$. Then, for $g \in L^2(\mathbb{R})$ and a lattice $\Lambda \subset \mathbb{R}^2$ we set

$$(g,\Lambda) := \big\{\pi(\lambda)g : \lambda \in \Lambda\big\} \qquad \text{and} \qquad \mathcal{G}(g,\Lambda) := \overline{\operatorname{span}}\,(g,\Lambda) \subset L^2(\mathbb{R}).$$

For the Fourier transform, we use the normalization $\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx$ for $f \in L^1(\mathbb{R})$. It is well-known that $\mathcal{F}$ extends to a unitary map $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$.

2.1. Bessel vectors and Janssen's representation

Let $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ be a separable lattice with $\alpha, \beta > 0$. The *adjoint lattice* of Λ is defined as $\Lambda^\circ = \frac{1}{\beta}\mathbb{Z} \times \frac{1}{\alpha}\mathbb{Z}$. We say that $g \in L^2(\mathbb{R})$ is a *Bessel vector* for Λ if the system (g,Λ) is a Bessel system in $L^2(\mathbb{R})$, meaning that the *analysis operator* $C_{\Lambda,g}$ corresponding to (g,Λ) is bounded as an operator from $L^2(\mathbb{R})$ to $\ell^2(\mathbb{Z}^2)$. It is defined by

$$C_{\Lambda,g} f = \big(\langle f, T_{m\alpha} M_{n\beta} g\rangle\big)_{m,n\in\mathbb{Z}}, \qquad f \in L^2(\mathbb{R}).$$

We denote the set of Bessel vectors for Λ by $\mathcal{B}_\Lambda$. This is a linear subspace of $L^2(\mathbb{R})$ which is dense because it contains the Schwartz space $\mathcal{S}(\mathbb{R})$; see [14, Corollary 6.2.3]. It is well known that $\mathcal{B}_\Lambda = \mathcal{B}_{\Lambda^\circ}$ (see [26, Theorem 2.2(a)]) and that

$$\sum_{m,n\in\mathbb{Z}} \langle f, T_{m\alpha}M_{n\beta}g\rangle\langle T_{m\alpha}M_{n\beta}h, u\rangle = \frac{1}{\alpha\beta}\sum_{k,\ell\in\mathbb{Z}} \langle h, T_{\frac{k}{\beta}}M_{\frac{\ell}{\alpha}}g\rangle\langle T_{\frac{k}{\beta}}M_{\frac{\ell}{\alpha}}f, u\rangle \tag{2.1}$$

whenever at least three of $f, g, h, u \in L^2(\mathbb{R})$ are Bessel vectors for Λ; this follows from [20, Proposition 2.4]. Formula (2.1) yields a useful representation (the so-called *Janssen representation*) of the cross frame operator $S_{\Lambda,g,h} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ associated to Bessel vectors $g, h \in \mathcal{B}_\Lambda$. This operator is defined by

$$S_{\Lambda,g,h}f := \sum_{m,n\in\mathbb{Z}} \langle f, T_{m\alpha}M_{n\beta}g\rangle \cdot T_{m\alpha}M_{n\beta}h, \qquad f \in L^2(\mathbb{R}). \tag{2.2}$$

Equation (2.1) implies that

$$S_{\Lambda,g,h}f = \frac{1}{\alpha\beta}\sum_{k,\ell\in\mathbb{Z}} \big\langle h, T_{\frac{k}{\beta}}M_{\frac{\ell}{\alpha}}g\big\rangle \cdot T_{\frac{k}{\beta}}M_{\frac{\ell}{\alpha}}f \qquad \text{if} \quad f, g, h \in \mathcal{B}_\Lambda. \tag{2.3}$$

The series in Equations (2.2) and (2.3) both converge unconditionally in $L^2(\mathbb{R})$.

2.2. Time-frequency shift invariance

For a closed linear subspace $\mathcal{G} \subset L^2(\mathbb{R})$, we denote by $\mathfrak{I}(\mathcal{G})$ the set of all pairs $(a, b) \in \mathbb{R}^2$ such that $\mathcal{G}$ is invariant under the time-frequency shift $\pi(a, b)$; that is,

$$\mathfrak{I}(\mathcal{G}) := \big\{z \in \mathbb{R}^2 : \pi(z)\mathcal{G} \subset \mathcal{G}\big\}.$$

If $\mathcal{G} = \mathcal{G}(g, \Lambda)$ for some $g \in L^2(\mathbb{R})$ and a lattice $\Lambda \subset \mathbb{R}^2$, then clearly $\Lambda \subset \mathfrak{I}(\mathcal{G})$. Any time-frequency shift $\pi(z)$ with $z \in \mathfrak{I}(\mathcal{G})\backslash\Lambda$ will be called an *additional time-frequency shift* for $\mathcal{G}(g, \Lambda)$. For Gabor spaces $\mathcal{G} = \mathcal{G}(g, \Lambda)$, the set $\mathfrak{I}(\mathcal{G})$ has some additional structure:

Lemma 2.2 ([3, Proposition A.1])**.** *Let $g \in L^2(\mathbb{R})$, let $\Lambda \subset \mathbb{R}^2$ be a lattice, and define $\mathcal{G} := \mathcal{G}(g, \Lambda)$. If $z \in \mathbb{R}^2$, then $z \in \mathfrak{I}(\mathcal{G})$ if and only if $\pi(z)g \in \mathcal{G}$. Moreover, $\mathfrak{I}(\mathcal{G})$ is a closed additive subgroup of $\mathbb{R}^2$.*

Lemma 2.2 shows that $z \in \mathfrak{I}(\mathcal{G})$ implies $-z \in \mathfrak{I}(\mathcal{G})$, i.e., $\pi(z)\mathcal{G} \subset \mathcal{G}$ and $\pi(z)^{-1}\mathcal{G} \subset \mathcal{G}$. Hence, we have $\pi(z)\mathcal{G} = \mathcal{G}$ whenever $z \in \mathfrak{I}(\mathcal{G})$.

The next lemma characterizes the case when $\mathcal{G}$ is invariant under *all* time-frequency shifts.

Lemma 2.3. *For a closed linear subspace $\mathcal{G} \subset L^2(\mathbb{R})$, $\mathcal{G} \neq \{0\}$, we have $\mathfrak{I}(\mathcal{G}) = \mathbb{R}^2$ if and only if $\mathcal{G} = L^2(\mathbb{R})$.*

Proof. Clearly, if $\mathcal{G} = L^2(\mathbb{R})$, then $\mathfrak{I}(\mathcal{G}) = \mathbb{R}^2$. Conversely, assume that $\mathfrak{I}(\mathcal{G}) = \mathbb{R}^2$ and let $f \in \mathcal{G}^\perp$ and $g \in \mathcal{G}\backslash\{0\}$. Then $\langle f, \pi(z)g\rangle = 0$ for all $z \in \mathbb{R}^2$, so that the short-time Fourier transform $V_g f$ of f with window g satisfies $V_g f \equiv 0$. By [14, Corollary 3.2.2] and since $g \neq 0$, this implies $f = 0$. We have thus shown $\mathcal{G}^\perp = \{0\}$, whence $\mathcal{G} = L^2(\mathbb{R})$, since $\mathcal{G}$ is a closed subspace of $L^2(\mathbb{R})$. □

2.3. Symplectic operators

It is often useful to reduce a statement involving a non-separable lattice to one that involves a separable lattice, since separable lattices are usually easier to handle. For this reduction, we will use so-called *symplectic operators* (see [14, Section 9.4]). Since we are working in dimension $d = 1$, a matrix $B \in \mathbb{R}^{2\times 2}$ is *symplectic* if and only if $\det B = 1$; see [14, Lemma 9.4.1]. For any such matrix B, it is shown in [14, Equation (9.39)] that there exists a unitary operator $U_B : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ such that

$$U_B \rho(z) = \rho(Bz) U_B, \quad z \in \mathbb{R}^2, \tag{2.4}$$

where (as in [14, Page 185 and Equation (9.25)])

$$\rho(a,b) := e^{\pi i ab} \cdot \pi(a,b).$$

In the sequel, we fix for each $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ one choice of the operator U_B, and for functions $g \in L^2(\mathbb{R})$, closed subspaces $\mathcal{G} \subset L^2(\mathbb{R})$, and sets $\Lambda \subset \mathbb{R}^2$ we write

$$g_B := U_B\, g, \qquad \mathcal{G}_B := U_B\, \mathcal{G}, \qquad \text{and} \qquad \Lambda_B := B\Lambda.$$

As shown in [14, Page 197], given $B, C \in \mathbb{R}^{2\times 2}$ with $\det B = \det C = 1$, we have $U_B U_C = \theta_{B,C}\, U_{BC}$ for some $\theta_{B,C} \in \mathbb{C}$ with $|\theta_{B,C}| = 1$.

Note that (2.4) implies

$$\pi(z) g \in \mathcal{G} \iff \pi(Bz) g_B \in \mathcal{G}_B, \qquad z \in \mathbb{R}^2. \tag{2.5}$$

Therefore, (g, Λ) is a frame (Riesz basis, resp.) for its closed linear span $\mathcal{G}$ if and only if (g_B, Λ_B) is a frame (Riesz basis, resp.) for its closed linear span $\mathcal{G}_B$. Thanks to Lemma 2.2, the equivalence (2.5) also implies that

$$\mathfrak{I}(\mathcal{G}_B) = B\, \mathfrak{I}(\mathcal{G}). \tag{2.6}$$

2.4. The Feichtinger algebra

We denote by $S_0(\mathbb{R})$ the *Feichtinger algebra*, which is the space of functions $f \in L^2(\mathbb{R})$ such that $\langle f, \pi(\cdot)\varphi\rangle \in L^1(\mathbb{R}^2)$ for some (and hence every; see [14, Proposition 12.1.2]) Schwartz function $\varphi \neq 0$.

Recall that $S_0(\mathbb{R})$ is invariant under each operator U_B (cf. [14, Proposition 12.1.3]), so that $g \in S_0(\mathbb{R})$ always implies $g_B \in S_0(\mathbb{R})$ for $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$. Also, each $g \in S_0(\mathbb{R})$ is a Bessel vector for any (separable) lattice (see e.g. [14, Propositions 6.2.2 and 12.1.4]). Since for $g, h \in S_0(\mathbb{R})$ and any $\alpha, \beta > 0$ the sequence $\big(\langle h, T_{m\alpha} M_{n\beta} g\rangle\big)_{m,n\in\mathbb{Z}}$ belongs to $\ell^1(\mathbb{Z}^2)$ (see [14, Corollary 12.1.12]), it follows from (2.3) and from the density of $\mathcal{B}_\Lambda$ in $L^2(\mathbb{R})$ that

$$S_{\Lambda,g,h} = \frac{1}{\alpha\beta} \sum_{k,\ell\in\mathbb{Z}} \big\langle h, T_{\frac{k}{\beta}} M_{\frac{\ell}{\alpha}} g\big\rangle \cdot T_{\frac{k}{\beta}} M_{\frac{\ell}{\alpha}} \quad \text{with} \quad \Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}, \tag{2.7}$$

where the series converges absolutely in operator norm.

2.5. The space $\mathbb{H}^1$

Let $H^1(\mathbb{R})$ denote the space of all functions f in $L^2(\mathbb{R})$ for which the weak derivative f' exists and belongs to $L^2(\mathbb{R})$. In other words, $H^1(\mathbb{R}) = W^{1,2}(\mathbb{R})$ is an L^2-Sobolev-space. It is well known (see [21, Theorem 7.16]) that each $f \in H^1(\mathbb{R})$ has a representative that is absolutely continuous on $\mathbb{R}$ and whose classical derivative exists and coincides with the weak derivative f' almost everywhere.

By $\mathbb{H}^1$ we denote the space of all functions $f \in H^1(\mathbb{R})$ whose Fourier transform $\widehat{f}$ also belongs to $H^1(\mathbb{R})$. Equivalently, a function $f \in L^2(\mathbb{R})$ is in $\mathbb{H}^1$ if and only if $f', Xf \in L^2(\mathbb{R})$, where Xf represents the function $\mathbb{R} \to \mathbb{C}, x \mapsto xf(x)$. The space $\mathbb{H}^1$ also coincides with the modulation space $M^2_m(\mathbb{R})$ with the weight $m(x,\omega) = 1+\sqrt{x^2+\omega^2}$; see [16, Corollary 2.3].

As shown in [6, Proof of Theorem 1.4], the space $\mathbb{H}^1$ is invariant under symplectic operators, meaning that $U_B g \in \mathbb{H}^1$ if $g \in \mathbb{H}^1$ and $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$.

3. Time-frequency shift invariance: A closer look

In this section, we first establish a certain trichotomy concerning the set of invariant time-frequency shifts. We then show that one of the three cases of the trichotomy is excluded if the generator function g is "sufficiently nice".

The next theorem establishes the trichotomy: the invariance set $\mathfrak{I}(\mathcal{G})$ either fills the whole space $\mathbb{R}^2$, or it consists of equispaced lines that are aligned with the lattice, or it is a refinement of Λ (and in particular a lattice itself). Note that this holds regardless of the regularity of the generator g or the (frame) properties of the Gabor system (g,Λ).

Proposition 3.1. *Let H be a closed additive subgroup of $\mathbb{R}^2$ and suppose that $H \supset \Lambda$ for a non-degenerate lattice $\Lambda \subset \mathbb{R}^2$. Then there exist $\lambda_1, \lambda_2 \in \Lambda$ satisfying $\Lambda = \mathbb{Z}\cdot\lambda_1 + \mathbb{Z}\cdot\lambda_2$ and $m, n \in \mathbb{N}_{\geq 1}$ such that exactly one of the following conditions holds:*

(1) $H = \mathbb{R}^2$.
(2) $H = \mathbb{R}\cdot\lambda_1 + \mathbb{Z}\cdot\frac{\lambda_2}{n}$.
(3) $H = \mathbb{Z}\cdot\frac{\lambda_1}{m} + \mathbb{Z}\cdot\frac{\lambda_2}{n}$.

In particular, if $\Lambda \subset \mathbb{R}^2$ is a lattice and $g \in L^2(\mathbb{R})$, then one of the above cases holds for $H = \mathfrak{I}(\mathcal{G}(g,\Lambda))$.

Proof. By [17, Theorem 9.11], there are $\alpha, \beta \in \mathbb{N}_0$ and linearly independent vectors $x_1, \dots, x_\alpha, y_1, \dots, y_\beta \in \mathbb{R}^2$ (hence, $\alpha+\beta \leq 2$) such that

$$H = \mathbb{R}\,x_1 + \cdots + \mathbb{R}\,x_\alpha + \mathbb{Z}\,y_1 + \cdots + \mathbb{Z}\,y_\beta.$$

Since H contains the non-degenerate lattice Λ (and thus two linearly independent vectors), we must have $\alpha+\beta = 2$. Hence, there are three cases:

(i) $(\alpha,\beta) = (2,0)$ and hence $H = \mathbb{R}^2$,
(ii) $(\alpha,\beta) = (1,1)$, so that $H = \mathbb{R}v + \mathbb{Z}w$ with linearly independent $v, w \in \mathbb{R}^2$,
(iii) $(\alpha,\beta) = (0,2)$, so that H is a (non-degenerate) lattice.

Clearly, in Case (i), Condition (1) of the statement of the theorem holds. Let us discuss the case (ii): $H = \mathbb{R}\cdot v + \mathbb{Z}\cdot w$. Let $\Lambda = \mathbb{Z}\cdot\mu + \mathbb{Z}\cdot\lambda$ be an arbitrary representation of Λ. Since $\Lambda \subset H$, there exist $m, n \in \mathbb{Z}$ and $s, t \in \mathbb{R}$ such that

$$[\mu,\lambda] = [sv+mw,\, tv+nw] = [v,w]\begin{bmatrix} s & t \\ m & n\end{bmatrix}.$$

Note that μ, λ are linearly independent, and hence $sn - tm \neq 0$, so that $d = (sn-tm)^{-1}$ is well-defined. Furthermore, we see

$$[v,w] = d\cdot[\mu,\lambda]\begin{bmatrix} n & -t \\ -m & s\end{bmatrix},$$

which shows that $v = d(n\mu - m\lambda)$ and thus $\mathbb{R}\cdot v = \mathbb{R}\cdot\lambda_1$ with some $\lambda_1 \in \Lambda$. By rescaling λ_1, we can ensure that $\frac{1}{k}\lambda_1 \notin \Lambda$ for each $k \in \mathbb{Z}\backslash\{-1,0,1\}$. Note because of $\mathbb{R}\cdot v = \mathbb{R}\cdot\lambda_1$ that $H = \mathbb{R}\cdot\lambda_1 + \mathbb{Z}\cdot w$.

Now, there exists $\lambda_2 \in \Lambda$ such that $\Lambda = \mathbb{Z}\cdot\lambda_1 + \mathbb{Z}\cdot\lambda_2$. Indeed, writing $\Lambda = A\mathbb{Z}^2$ with $A = [a_1, a_2] \in \mathbb{R}^{2\times 2}$ invertible, there exist $i, j \in \mathbb{Z}$ such that $\lambda_1 = ia_1 + ja_2$. The numbers i, j are necessarily coprime, since $\frac{1}{k}\lambda_1 \notin \Lambda$ for $k \in \mathbb{Z}\backslash\{-1,0,1\}$. Hence, by Bézout's lemma there exist $k, \ell \in \mathbb{Z}$ such that $i\ell - jk = 1$. Set $\lambda_2 = ka_1 + \ell a_2$. Then $[\lambda_1, \lambda_2]\mathbb{Z}^2 = A\left[\begin{smallmatrix} i & k \\ j & \ell \end{smallmatrix}\right]\mathbb{Z}^2 = A\mathbb{Z}^2 = \Lambda$. Since $\lambda_2 \in \Lambda \subset H$, there exist $\sigma \in \mathbb{R}$ and $\nu \in \mathbb{Z}$ such that $\lambda_2 = \sigma\lambda_1 + \nu w$. Then $\nu \neq 0$ and so $H = \mathbb{R}\cdot\lambda_1 + \mathbb{Z}\cdot(\frac{\lambda_2}{\nu} - \frac{\sigma}{\nu}\lambda_1) = \mathbb{R}\cdot\lambda_1 + \mathbb{Z}\cdot\frac{\lambda_2}{|\nu|}$. Hence, Condition (2) of the statement of the theorem holds.

Assume now that Case (iii) holds: H is a lattice, i.e., $H = \mathbb{Z}\cdot v + \mathbb{Z}\cdot w$ with linearly independent vectors $v, w \in \mathbb{R}^2$. Write $\Lambda = \mathbb{Z}\cdot\mu + \mathbb{Z}\cdot\lambda$ with linearly independent $\lambda, \mu \in \mathbb{R}^2$. Then, because of $\Lambda \subset H$, there exist $a, b, c, d \in \mathbb{Z}$ such that

$$[\mu, \lambda] = [av + cw, bv + dw] = [v, w]\begin{bmatrix} a & b \\ c & d \end{bmatrix} =: [v, w]\cdot A, \tag{3.1}$$

with $A \in \mathbb{Z}^{2\times 2}$. Let $A = MDN^{-1}$ be the Smith canonical form of A (see, for instance [22, Theorem 26.2] or [18, Theorem 3.8]), where $M, N, D \in \mathbb{Z}^{2\times 2}$ with $\det M = \det N = 1$ and D is a diagonal matrix. Note that A (and hence D) is invertible; this follows from (3.1) since μ and λ are linearly independent. Moreover, note that $[\mu, \lambda]ND^{-1} = [v, w]M$.

Define $\lambda_1, \lambda_2 \in \Lambda$ via $[\lambda_1, \lambda_2] := [\mu, \lambda]N$. Then

$$\Lambda = [\mu, \lambda]\mathbb{Z}^2 = [\mu, \lambda]N\mathbb{Z}^2 = [\lambda_1, \lambda_2]\mathbb{Z}^2 = \mathbb{Z}\cdot\lambda_1 + \mathbb{Z}\cdot\lambda_2.$$

Further, writing $D = \operatorname{diag}(m, n)$ with $m, n \in \mathbb{Z}\backslash\{0\}$, we see

$$H = [v, w]\mathbb{Z}^2 = [v, w]M\mathbb{Z}^2 = [\mu, \lambda]ND^{-1}\mathbb{Z}^2 = [\lambda_1, \lambda_2]D^{-1}\mathbb{Z}^2 = \mathbb{Z}\cdot\tfrac{\lambda_1}{|m|} + \mathbb{Z}\cdot\tfrac{\lambda_2}{|n|}.$$

This completes the proof of the theorem, since the conditions (1)–(3) are clearly mutually exclusive. □

The example below shows that Case (2) in Proposition 3.1 can occur for $H = \mathfrak{I}(\mathcal{G}(g, \Lambda))$ for every lattice Λ with density smaller than one—even if (g, Λ) is a Riesz sequence.

Example 3.2. Due to Equation (2.6) and Lemma 2.1 it suffices to construct an example for a separable lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ with $\alpha, \beta > 0$, $\alpha\beta > 1$. For $m \in \mathbb{Z}$, define $E_m := m\alpha + [0, \frac{1}{\beta}]$, and let $g := \sqrt{\beta}\cdot\mathbf{1}_{E_0}$. Then

$$M_{n\beta}T_{m\alpha}g(x) = \sqrt{\beta}\cdot e^{2\pi i n\beta x}\cdot\mathbf{1}_{[0,\frac{1}{\beta}]}(x - m\alpha) = \sqrt{\beta}\cdot e^{2\pi i n\beta x}\cdot\mathbf{1}_{E_m}(x).$$

Hence, for any $m \in \mathbb{Z}$ the system $(T_{m\alpha}M_{n\beta}g)_{n\in\mathbb{Z}}$ is an orthonormal basis for the subspace $L^2(E_m)$ of $L^2(\mathbb{R})$. Note that, since $\frac{1}{\beta} < \alpha$, we have $[0, \frac{1}{\beta}] \subsetneq [0, \alpha]$. The system (g, Λ) is thus an orthonormal basis for $\mathcal{G} := L^2(E) \subset L^2(\mathbb{R})$, where $E = \bigcup_{m\in\mathbb{Z}} E_m$. Note that $\mathbb{R}\backslash E$ has positive (even infinite) measure, so that $\mathcal{G} \subsetneq L^2(\mathbb{R})$. Moreover, for any $\omega \in \mathbb{R}$ we have $M_\omega g = \sqrt{\beta}\cdot e^{2\pi i\omega\cdot}\cdot\mathbf{1}_{E_0} \in L^2(E_0) \subset \mathcal{G}$. Therefore, Lemma 2.2 shows $\{0\} \times \mathbb{R} \subset \mathfrak{I}(\mathcal{G})$, which can only occur in Case (2) of Proposition 3.1, since we would have $\mathcal{G} = L^2(\mathbb{R})$ in Case (1), see Lemma 2.3.

Note that the function g in Example 3.2 is not well localized in frequency. In the remainder of this section, we show that Case (2) in Proposition 3.1 *cannot* occur if (g, Λ) is a frame sequence with a sufficiently nice window g. In this case, the trichotomy

from Proposition 3.1 becomes a dichotomy. By g being "sufficiently nice" we mean that $g \in \mathbb{W}(C, \ell^2)$, where

$$\mathbb{W}(C, \ell^2) := \big\{ f \in L^2(\mathbb{R}) : U_B f \in W(C, \ell^2) \text{ for all } B \in \mathbb{R}^{2\times 2} \text{ with } \det B = 1 \big\}.$$

Here, $W(C, \ell^2)$ is the so-called *Wiener Amalgam space* consisting of all *continuous* functions $f : \mathbb{R} \to \mathbb{C}$ satisfying

$$\|f\|_{W(C,\ell^2)} := \left(\sum_{k\in\mathbb{Z}} \sup_{x\in[k-1,k+1]} |f(x)|^2 \right)^{1/2} < \infty.$$

Recall from Section 2.3 that if $g \in \mathbb{W}(C, \ell^2)$ and if $B, C \in \mathbb{R}^{2\times 2}$ satisfy $\det B = \det C = 1$, then there is $\theta_{B,C} \in \mathbb{C}$ satisfying $U_C U_B g = \theta_{B,C}\, U_{CB} g \in W(C, \ell^2)$. This shows that $U_B g \in \mathbb{W}(C, \ell^2)$ whenever $g \in \mathbb{W}(C, \ell^2)$ and $\det B = 1$.

Before we prove the announced theorem let us show that the function classes considered in this paper (namely, $S_0(\mathbb{R})$ and $\mathbb{H}^1$) are contained in $\mathbb{W}(C, \ell^2)$.

Lemma 3.3. *We have $S_0(\mathbb{R}) \subset \mathbb{W}(C, \ell^2)$ and $\mathbb{H}^1 \subset \mathbb{W}(C, \ell^2)$.*

Proof. If $g \in S_0(\mathbb{R})$, then [14, Proposition 12.1.3] shows that $U_B g \in S_0(\mathbb{R})$ for each $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$. Similarly, if $g \in \mathbb{H}^1$, then [6, Proof of Theorem 1.4] shows that $U_B g \in \mathbb{H}^1$ for each $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$. Therefore, it suffices to show that $S_0(\mathbb{R}) \subset W(C, \ell^2)$ and $\mathbb{H}^1 \subset W(C, \ell^2)$.

First, if $f \in S_0(\mathbb{R})$, then [14, Proposition 12.1.4] shows that $\widehat{f} \in L^1(\mathbb{R})$. By Fourier inversion, this implies that f has a continuous representative. Furthermore, by [14, Proposition 12.1.4] we have $f \in W(L^\infty, \ell^1)$. Combined with the embedding $\ell^1(\mathbb{Z}) \hookrightarrow \ell^2(\mathbb{Z})$, this easily implies $f \in W(L^\infty, \ell^2)$ and thus $f \in W(C, \ell^2)$.

Next, if $f \in \mathbb{H}^1 \subset H^1 = W^{1,2}(\mathbb{R})$, then [21, Theorem 7.16] shows (after changing f on a null-set) that f is absolutely continuous, and hence continuous, and satisfies $f(x) - f(y) = \int_y^x f'(t)\, dt$ for all $y < x$, where $f' \in L^2(\mathbb{R})$ is the weak derivative of f. Now, note that if $n \in \mathbb{Z}$ and $x, y \in [n-1, n+1]$, then

$$\begin{aligned} |f(x)| &\leq |f(y)| + \int_{\min\{x,y\}}^{\max\{x,y\}} |f'(t)|\, dt \leq |f(y)| + \int_{n-1}^{n+1} |f'(t)|\, dt \\ &\leq |f(y)| + \sqrt{2} \left(\int_{n-1}^{n+1} |f'(t)|^2\, dt \right)^{1/2}, \end{aligned}$$

and hence $|f(x)|^2 \leq 2|f(y)|^2 + 4 \int_{n-1}^{n+1} |f'(t)|^2\, dt$. Integrating this over $y \in [n-1, n+1]$ gives

$$2|f(x)|^2 \leq 2 \int_{n-1}^{n+1} |f(y)|^2\, dy + 8 \int_{n-1}^{n+1} |f'(t)|^2\, dt,$$

for all $x \in [n-1, n+1]$, which finally implies

$$\|f\|_{W(C,\ell^2)}^2 \leq \sum_{n\in\mathbb{Z}} \left(\int_{n-1}^{n+1} |f(y)|^2\, dy + 4 \int_{n-1}^{n+1} |f'(t)|^2\, dt \right) \lesssim \|f\|_{L^2}^2 + \|f'\|_{L^2}^2 < \infty,$$

and hence $f \in W(C, \ell^2)$. □

Our next result shows that Case (2) in Proposition 3.1 cannot occur if (g, Λ) is a frame sequence with generator $g \in \mathbb{W}(C, \ell^2) \backslash \{0\}$.

Theorem 3.4. *Let $g \in \mathbb{W}(C, \ell^2)\backslash\{0\}$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g, Λ) is a frame for $\mathcal{G} = \mathcal{G}(g, \Lambda)$. Then either $\mathfrak{I}(\mathcal{G}) = \mathbb{R}^2$ or there exist $\lambda_1, \lambda_2 \in \Lambda$ and $m, n \in \mathbb{N}_{\geq 1}$ such that*

$$\Lambda = \mathbb{Z} \cdot \lambda_1 + \mathbb{Z} \cdot \lambda_2 \qquad \text{and} \qquad \mathfrak{I}(\mathcal{G}) = \mathbb{Z} \cdot \tfrac{\lambda_1}{m} + \mathbb{Z} \cdot \tfrac{\lambda_2}{n}. \tag{3.2}$$

Proof. Let us assume $\mathfrak{I}(\mathcal{G}) \subsetneq \mathbb{R}^2$. Writing $H = \mathfrak{I}(\mathcal{G})$, the two possibilities in (3.2) represent the cases (1) and (3) from the trichotomy in Proposition 3.1. It is thus enough to show that Case (2) from that theorem cannot occur. Therefore, we assume towards a contradiction that Case (2) holds, i.e., there are $\lambda_1, \lambda_2 \in \Lambda$ and $n \in \mathbb{N}_{\geq 1}$ such that $\Lambda = \mathbb{Z} \cdot \lambda_1 + \mathbb{Z} \cdot \lambda_2$ and $\mathfrak{I}(\mathcal{G}) = \mathbb{Z} \cdot \frac{\lambda_1}{n} + \mathbb{R} \cdot \lambda_2$.

Step 1. We first derive a contradiction for the case $\lambda_1 = (\alpha, 0)^\top$ and $\lambda_2 = (0, \beta)^\top$ with some $\alpha, \beta > 0$. Then $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$, and $\{0\} \times \mathbb{R} \subset \mathfrak{I}(\mathcal{G})$. For $f \in \mathcal{G}$ we thus have $M_\omega f \in \mathcal{G}$ for all $\omega \in \mathbb{R}$. By [27, Theorem 9.17] (applied to the translation invariant space $\mathcal{F}^{-1}\mathcal{G}$, with $\mathcal{F}$ denoting the Fourier transform), there exists a Borel measurable set $E \subset \mathbb{R}$ such that $\mathcal{G} = L^2(E)$, where we consider $L^2(E)$ as a closed subspace of $L^2(\mathbb{R})$, in the sense that $L^2(E) = \{f \in L^2(\mathbb{R}) \colon f = 0 \text{ a.e. on } \mathbb{R}\backslash E\}$.

Our goal is to show that $E = \mathbb{R}$, up to null-sets. This will imply $\mathcal{G} = L^2(E) = L^2(\mathbb{R})$ and hence $\mathfrak{I}(\mathcal{G}) = \mathbb{R}^2$, providing the desired contradiction. Towards proving $E = \mathbb{R}$, let us consider for given $f \in L^2(\mathbb{R})$ the continuous function $\Gamma_f : \mathbb{R} \to \mathbb{R}$ defined by

$$\Gamma_f(\omega) := \langle S M_\omega f, M_\omega f \rangle, \quad \omega \in \mathbb{R},$$

where $S : L^2(\mathbb{R}) \to \mathcal{G}$ denotes the frame operator of (g, Λ). By [14, Proposition 7.1.1], the operator S has the *Walnut representation*

$$\langle Sf, h \rangle = \beta^{-1} \sum_{n \in \mathbb{Z}} \langle G_n \cdot T_{\frac{n}{\beta}} f, h \rangle \qquad \forall\, f, h \in L^\infty(\mathbb{R}) \text{ with compact support},$$

where only finitely many terms of the sum do not vanish, and where

$$G_n(x) := \sum_{m \in \mathbb{Z}} g(x - m\alpha) \cdot \overline{g(x - \tfrac{n}{\beta} - m\alpha)}, \qquad x \in \mathbb{R}, n \in \mathbb{Z}.$$

The fact that $g \in W(C, \ell^2)$ easily implies that the series defining G_n converges locally uniformly, and that the G_n are continuous functions. Since G_n is also α-periodic, this means that each G_n is bounded.

Now, since multiplication with G_n commutes with the modulation M_ω, using the identity $T_{n/\beta} M_\omega = e^{-2\pi i \frac{n}{\beta}\omega} M_\omega T_{n/\beta}$, we get

$$\Gamma_f(\omega) = \beta^{-1} \sum_{n \in \mathbb{Z}} e^{-2\pi i \frac{n}{\beta}\omega} \langle G_n \cdot T_{\frac{n}{\beta}} f, f \rangle \qquad \forall\, f \in L^\infty(\mathbb{R}) \text{ with compact support}, \tag{3.3}$$

where there are only finitely many $n \in \mathbb{Z}$ (depending only on f, but not on the choice of ω) for which $\langle G_n \cdot T_{\frac{n}{\beta}} f, f \rangle \neq 0$.

As (g, Λ) is a frame for $\mathcal{G}$ and $M_\omega f \in \mathcal{G}$ for all $\omega \in \mathbb{R}$ and $f \in \mathcal{G}$, there exists $A > 0$ such that $\Gamma_f(\omega) = \langle S M_\omega f, M_\omega f \rangle \geq A \|f\|_{L^2}^2$ for all $f \in \mathcal{G}$. Let us write $L_c^\infty(E)$ for the set of all compactly supported $f \in L^\infty(\mathbb{R})$ which satisfy $f = 0$ on $\mathbb{R}\backslash E$, and note that $L_c^\infty(E) \subset L^2(E) = \mathcal{G}$. For $f \in L_c^\infty(E)$, integrate the estimate $\Gamma_f(\omega) \geq A \|f\|_{L^2}^2$ over $[0, \beta]$ and apply Equation (3.3) to see

$$\beta\, A\, \|f\|_{L^2}^2 \;\leq\; \beta^{-1} \sum_{n \in \mathbb{Z}} \langle G_n \cdot T_{\frac{n}{\beta}} f, f \rangle \int_0^\beta e^{-2\pi i \frac{n}{\beta}\omega}\, d\omega = \langle G_0 f, f \rangle = \langle hf, f \rangle,$$

where $h := G_0 = \sum_{m\in\mathbb{Z}} |T_{m\alpha} g|^2$. We have thus shown

$$\int_E \big(h(x) - \beta A\big) \cdot |f(x)|^2 \, dx \;\geq\; 0 \qquad \forall\, f \in L^\infty_c(E).$$

Using standard arguments, this implies that $h(x) \geq \beta A$ for almost all $x \in E$.

Since $T_{m\alpha}\, g \in \mathcal{G} = L^2(E)$ and thus $T_{m\alpha}\, g(x) = 0$ for almost all $x \in \mathbb{R}\backslash E$ and arbitrary $m \in \mathbb{Z}$, it follows that $h(x) = 0$ for almost all $x \in \mathbb{R}\backslash E$. Recall from above that $h(x) \geq \beta A$ for almost all $x \in E$; thus, $h(x) \in \{0\} \cup [\beta A, \infty)$ almost everywhere. Also recall from above that $h = G_0$ is continuous. Hence, the *open* set $h^{-1}((0, \beta A))$ has measure zero and is thus empty; that is, $h(x) \in \{0\} \cup [\beta A, \infty)$ for *all* $x \in \mathbb{R}$. By the intermediate value theorem, this implies that $h(x) \geq \beta A$ for all $x \in \mathbb{R}$ (since $h \geq |g|^2$ and $g \not\equiv 0$) and thus, indeed, $E = \mathbb{R}$ (up to null-sets), since $h(x) = 0$ a.e. on $\mathbb{R}\backslash E$.

Step 2. Let Λ be a general lattice. Recall that $\Lambda = \mathbb{Z}\lambda_1 + \mathbb{Z}\lambda_2$ and $\mathfrak{I}(\mathcal{G}) = \mathbb{Z}\frac{\lambda_1}{n} + \mathbb{R}\lambda_2$. By Lemma 2.1 there exists $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ such that $B[\lambda_1, \lambda_2] = \operatorname{diag}(\alpha, \beta)$ for certain $\alpha, \beta \in \mathbb{R}\backslash\{0\}$. We thus obtain $\Lambda_B = B\Lambda = B[\lambda_1, \lambda_2]\mathbb{Z}^2 = |\alpha|\mathbb{Z} \times |\beta|\mathbb{Z}$ and

$$\mathfrak{I}(\mathcal{G}_B) = B\mathfrak{I}(\mathcal{G}) = B[\lambda_1, \lambda_2] \operatorname{diag}(\tfrac{1}{n}, 1)(\mathbb{Z} \times \mathbb{R}) = \operatorname{diag}(\tfrac{|\alpha|}{n}, |\beta|)(\mathbb{Z} \times \mathbb{R}) = \tfrac{|\alpha|}{n}\mathbb{Z} \times \mathbb{R};$$

see (2.6). In particular, $\{0\} \times \mathbb{R} \subset \mathfrak{I}(\mathcal{G}_B)$. Hence, since $g_B = U_B g \in \mathbb{W}(C, \ell^2)$ and (g_B, Λ_B) is a frame for $\mathcal{G}_B = U_B\mathcal{G} \subsetneq L^2(\mathbb{R})$ (cf. Subsection 2.3), we are in the situation of Step 1, which we proved to be impossible. □

By combining Theorem 3.4 and Lemma 3.3, we obtain the following corollary.

Corollary 3.5. *Let $g \in S_0(\mathbb{R})\backslash\{0\}$ or $g \in \mathbb{H}^1\backslash\{0\}$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g, Λ) is a Riesz basis for $\mathcal{G} := \mathcal{G}(g, \Lambda)$. Then $\mathfrak{I}(\mathcal{G})$ is a refinement of Λ as in* (3.2).

Proof. By the Balian-Low theorem [10, Theorem 2.3] and the Amalgam Balian-Low theorem [2, Theorem 3.2], it is not possible that $\mathcal{G} = L^2(\mathbb{R})$. Therefore, Lemma 2.3 implies $\mathfrak{I}(\mathcal{G}) \neq \mathbb{R}^2$. The rest follows from Lemma 3.3 and Theorem 3.4. □

4. Time-frequency shift invariance: Duality

Let us consider a Gabor Riesz sequence (g, Λ) with $g \in \mathbb{W}(C, \ell^2)$ as in the previous section, and assume that $\mathcal{G} := \mathcal{G}(g, \Lambda) \subsetneq L^2(\mathbb{R})$, but that there exists an additional time-frequency shift, meaning $\mathfrak{I}(\mathcal{G}) \neq \Lambda$. In view of Theorem 1.1 it is our goal to show that this is impossible, at least if $g \in S_0$. To make the situation more accessible, we first reduce to the case where $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ is separable, and where the additional time-frequency shift is of the form $(\frac{\alpha}{\nu}, 0)^\top$ for some $\nu \in \mathbb{N}_{\geq 2}$, meaning that $T_{\alpha/\nu}\, g \in \mathcal{G}$. After that, we provide a characterization of this simplified condition in terms of the adjoint Gabor system. It is this characterization that we will use to prove our main result, Theorem 1.1, in the next section.

Lemma 4.1. *Let $g \in \mathbb{W}(C, \ell^2)\backslash\{0\}$ and let $\Lambda \subset \mathbb{R}^2$ be a lattice such that (g, Λ) is a frame for $\mathcal{G} := \mathcal{G}(g, \Lambda)$. If $\mathcal{G} \neq L^2(\mathbb{R})$ and $\mathfrak{I}(\mathcal{G}) \neq \Lambda$, there exist a matrix $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ and $\alpha, \beta > 0$ such that $\Lambda_B = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ and $(\frac{\alpha}{\nu}, 0)^\top \in \mathfrak{I}(\mathcal{G}_B)$ for some $\nu \in \mathbb{N}$, $\nu \geq 2$ (i.e., $T_{\frac{\alpha}{\nu}} g_B \in \mathcal{G}_B$).*

Proof. Due to Theorem 3.4 and Lemma 2.3, we have $\Lambda = [\lambda_1, \lambda_2]\,\mathbb{Z}^2$ and $\mathfrak{I}(\mathcal{G}) = [\frac{\lambda_1}{m}, \frac{\lambda_2}{n}]\,\mathbb{Z}^2$ for suitable vectors $\lambda_1, \lambda_2 \in \mathbb{R}^2$ and $m, n \in \mathbb{N}\backslash\{0\}$. We may safely assume that $m \neq 1$.

Indeed, since $\mathfrak{I}(\mathcal{G}) \neq \Lambda$, we have $(m,n) \neq (1,1)$. If $m = 1$, then with $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ also $\Lambda = [\lambda_1, \lambda_2] J \mathbb{Z}^2 = [-\lambda_2, \lambda_1]\mathbb{Z}^2$ and $\mathfrak{I}(\mathcal{G}) = [\frac{-\lambda_2}{n}, \frac{\lambda_1}{m}]\mathbb{Z}^2$.

Now, by Lemma 2.1 there exists a matrix $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ such that $B[\lambda_1, \lambda_2] = \operatorname{diag}(\alpha, \beta)$, where $\alpha, \beta \in \mathbb{R}\backslash\{0\}$. Hence, $\Lambda_B = B\Lambda = |\alpha|\mathbb{Z} \times |\beta|\mathbb{Z}$ and

$$\mathfrak{I}(\mathcal{G}_B) = B\mathfrak{I}(\mathcal{G}) = B[\lambda_1, \lambda_2] \operatorname{diag}(\tfrac{1}{m}, \tfrac{1}{n})\mathbb{Z}^2 = \tfrac{|\alpha|}{m}\mathbb{Z} \times \tfrac{|\beta|}{n}\mathbb{Z};$$

see (2.6). In particular, $(\frac{|\alpha|}{m}, 0)^\top \in \mathfrak{I}(\mathcal{G}_B)$ and $m \geq 2$. □

In what follows, fix $g \in L^2(\mathbb{R})$, $\alpha, \beta > 0$, $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$, and $\nu \in \mathbb{N}_{\geq 2}$, and assume that (g, Λ) is a frame for $\mathcal{G} = \mathcal{G}(g, \Lambda)$. The adjoint system $\mathcal{F} := \{T_{k/\beta} M_{\ell/\alpha}\, g : k, \ell \in \mathbb{Z}\}$ is then a frame for its closed linear span $\mathcal{K}$ by [26, Theorem 2.2 (c)]. Note that $\mathcal{K} = L^2(\mathbb{R})$ if and only if (g, Λ) is a Riesz sequence (cf. [26, Thm. 2.2 (e)] or [14, Theorem 7.4.3]).

It is a natural question to ask what the existence of an additional time-frequency shift of the form $T_{\frac{\alpha}{\nu}} g \in \mathcal{G}$ means for the adjoint system $\mathcal{F}$. To describe this, we set

$$\mathcal{F}_s := \big\{T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} M_{\frac{s}{\alpha}} g \colon k, \ell \in \mathbb{Z}\big\}, \qquad s = 0, \dots, \nu - 1.$$

Again by [26, Theorem 2.2 (c)], $\mathcal{F}_0$ is a frame sequence if and only if the system $(g, \frac{\alpha}{\nu}\mathbb{Z} \times \beta\mathbb{Z})$ is a frame sequence. In this case, each $\mathcal{F}_s$ is a frame sequence because $M_{s/\alpha}\mathcal{F}_0$ is, and multiplying the vectors of a frame sequence by unimodular constants results in a frame sequence. We set $\mathcal{L}_s := \overline{\operatorname{span}}\, \mathcal{F}_s$ for $s \in \{0, \dots, \nu - 1\}$. Note that

$$\mathcal{K} = \mathcal{L}_0 + \cdots + \mathcal{L}_{\nu-1}.$$

Indeed, the inclusion "$\supset$" is trivial. Conversely, since $\mathcal{F}$ is a frame sequence, each $f \in \mathcal{K}$ satisfies $f = \sum_{k,\ell\in\mathbb{Z}} c_{k,\ell}\, T_{k/\beta} M_{\ell/\alpha}\, g$ with a suitable sequence $c = (c_{k,\ell})_{k,\ell\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$. Since $\mathcal{F}$ is a Bessel sequence, $f_s := \sum_{k,\ell\in\mathbb{Z}} c_{k,\ell\nu+s} T_{\frac{k}{\beta}} M_{\frac{\ell\nu+s}{\alpha}} g \in \mathcal{L}_s$ is well-defined for $s \in \{0, \dots, \nu - 1\}$, and $f = f_0 + \cdots + f_{\nu-1} \in \mathcal{L}_0 + \cdots + \mathcal{L}_{\nu-1}$. Finally, it is clear that $\mathcal{L}_s = M_{\frac{s}{\alpha}} \mathcal{L}_0$.

In the sequel, the symbol $\boxplus$ denotes the *direct* (not necessarily orthogonal) sum of subspaces, whereas $\oplus$ is used to denote an orthogonal sum. The next theorem characterizes the existence of an additional time-frequency shift for $\mathcal{G}$ in terms of properties of the adjoint system $\mathcal{F}$.

Theorem 4.2. *Let $g \in L^2(\mathbb{R})$ and $\alpha, \beta > 0$, and assume that $(g, \alpha\mathbb{Z} \times \beta\mathbb{Z})$ is a frame sequence with canonical dual window $\gamma \in \mathcal{G}$, where $\mathcal{G} = \mathcal{G}(g, \alpha\mathbb{Z} \times \beta\mathbb{Z})$. Let $\nu \in \mathbb{N}_{\geq 2}$, and define the systems $\mathcal{F}_s$ and the spaces $\mathcal{K}, \mathcal{L}_s$ as above, and set $S_{\gamma,g} := S_{\frac{1}{\beta}\mathbb{Z}\times\frac{\nu}{\alpha}\mathbb{Z},\gamma,g}$, with notation as in Equation (2.2). Then the following statements are equivalent:*

(i) $T_{\frac{\alpha}{\nu}} g \in \mathcal{G}$.

(ii) $(\alpha\beta)^{-1} S_{\gamma,g} M_{\frac{s}{\alpha}} g = \delta_{s,0} \cdot g$ *for* $s = 0, \dots, \nu - 1$.

(iii) $\mathcal{K} = \mathcal{L}_0 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}$.

(iv) $\big\langle T_{\frac{k}{\beta}} M_{\frac{\ell}{\alpha}} \gamma, g \big\rangle = 0$ *for all* $k \in \mathbb{Z}$ *and all* $\ell \in \mathbb{Z}\backslash\nu\mathbb{Z}$.

If one of (i)–(iv) *holds, then for each $s = 0, \dots, \nu - 1$ the system $\mathcal{F}_s$ is a frame for $\mathcal{L}_s$ and the operator $P_s := (\alpha\beta)^{-1} M_{s/\alpha} S_{\gamma,g} M_{-s/\alpha}$ is the (possibly non-orthogonal) projection onto $\mathcal{L}_s$ with respect to the decomposition $L^2(\mathbb{R}) = (\mathcal{L}_0 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}) \oplus \mathcal{K}^\perp$.*

Proof. First, note by Ron-Shen duality (see [26, Theorem 2.2(c)]) that $\mathcal{F}$ is a frame sequence. We will frequently use the following fact (see [26, Theorem 2.3]):

$$(\alpha\beta)^{-1}\gamma \text{ is the canonical dual window of } \mathcal{F} = \left\{T_{\frac{k}{\beta}} M_{\frac{\ell}{\alpha}} g : k, \ell \in \mathbb{Z}\right\} ; \tag{4.1}$$

in particular, $\gamma \in \overline{\operatorname{span}}\, \mathcal{F} = \mathcal{K}$.

For the rest of the proof we set $P := (\alpha\beta)^{-1} S_{\gamma,g} = (\alpha\beta)^{-1} S_{\frac{1}{\beta}\mathbb{Z}\times\frac{\nu}{\alpha}\mathbb{Z},\gamma,g}$. It is well known (see for instance [14, Equation (5.25)]) that

$$P T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} = T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} P \qquad \text{for all } k, \ell \in \mathbb{Z}. \tag{4.2}$$

Moreover, Equation (2.3) applied to the lattice $\frac{1}{\beta}\mathbb{Z} \times \frac{\nu}{\alpha}\mathbb{Z}$ shows for $f \in \mathcal{B}_{\frac{1}{\beta}\mathbb{Z}\times\frac{\nu}{\alpha}\mathbb{Z}}$ that

$$Pf = \sum_{m,n\in\mathbb{Z}} c_{m,n} \cdot T_{\frac{m\alpha}{\nu}} M_{n\beta} f \qquad \text{with} \qquad c_{m,n} = \tfrac{1}{\nu} \langle g, T_{\frac{m\alpha}{\nu}} M_{n\beta} \gamma \rangle. \tag{4.3}$$

Let us denote the orthogonal projection onto the subspace $\mathcal{K} = \overline{\operatorname{span}}\, \mathcal{F}$ by $P_{\mathcal{K}}$. Note that Equation (4.1) implies $S_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z},(\alpha\beta)^{-1}\gamma,g}|_{\mathcal{K}} = \mathrm{id}_{\mathcal{K}}$ and $S_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z},(\alpha\beta)^{-1}\gamma,g}|_{\mathcal{K}^\perp} \equiv 0$, so that $P_{\mathcal{K}} = S_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z},(\alpha\beta)^{-1}\gamma,g}$. Similarly, the orthogonal projection $P_{\mathcal{G}}$ onto $\mathcal{G}$ satisfies $P_{\mathcal{G}} = S_{\alpha\mathbb{Z}\times\beta\mathbb{Z},\gamma,g}$. Next, using (4.3) and the elementary identity $\sum_{s=0}^{\nu-1} e^{2\pi i \frac{ms}{\nu}} = \nu \cdot \mathbb{1}_{\nu\mathbb{Z}}(m)$, we obtain

$$\begin{aligned}
\sum_{s=0}^{\nu-1} M_{\frac{s}{\alpha}} P M_{-\frac{s}{\alpha}} f &= \sum_{s=0}^{\nu-1} \sum_{m,n\in\mathbb{Z}} c_{m,n} \cdot M_{\frac{s}{\alpha}} T_{\frac{m\alpha}{\nu}} M_{n\beta} M_{-\frac{s}{\alpha}} f \\
&= \sum_{m,n\in\mathbb{Z}} c_{m,n} \Big(\sum_{s=0}^{\nu-1} e^{2\pi i \frac{ms}{\nu}} \Big) T_{\frac{m\alpha}{\nu}} M_{n\beta} f \\
&= \nu \sum_{m,n\in\mathbb{Z}} c_{\nu m,n} T_{m\alpha} M_{n\beta} f = \sum_{m,n\in\mathbb{Z}} \langle g, T_{m\alpha} M_{n\beta} \gamma \rangle \cdot T_{m\alpha} M_{n\beta} f \\
(\text{Equation (2.3)}) &= (\alpha\beta)^{-1} S_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z},\gamma,g} f = S_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z},(\alpha\beta)^{-1}\gamma,g} f = P_{\mathcal{K}} f,
\end{aligned} \tag{4.4}$$

for all $f \in \mathcal{B}_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z}}$ and hence for all $f \in L^2(\mathbb{R})$ by density. Here, we used that if $f \in \mathcal{B}_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z}}$, then $M_{-\frac{s}{\alpha}} f \in \mathcal{B}_{\frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z}} \subset \mathcal{B}_{\frac{1}{\beta}\mathbb{Z}\times\frac{\nu}{\alpha}\mathbb{Z}}$. Next, for $s = 0, \dots, \nu-1$, we see by another application of Equation (4.3) that

$$\begin{aligned}
M_{-\frac{s}{\alpha}} P M_{\frac{s}{\alpha}} g &= \frac{1}{\nu} \sum_{m,n\in\mathbb{Z}} \langle g, T_{\frac{m\alpha}{\nu}} M_{n\beta} \gamma \rangle \cdot M_{-\frac{s}{\alpha}} T_{\frac{m\alpha}{\nu}} M_{n\beta} M_{\frac{s}{\alpha}} g \\
&= \frac{1}{\nu} \sum_{m,n\in\mathbb{Z}} \sum_{r=0}^{\nu-1} \big\langle g, T_{\frac{\nu m - r}{\nu}\alpha} M_{n\beta} \gamma \big\rangle \cdot M_{-\frac{s}{\alpha}} T_{\frac{\nu m - r}{\nu}\alpha} M_{n\beta} M_{\frac{s}{\alpha}} g \\
&= \frac{1}{\nu} \sum_{r=0}^{\nu-1} e^{2\pi i \frac{sr}{\nu}} \cdot T_{-\frac{r\alpha}{\nu}} \sum_{m,n\in\mathbb{Z}} \big\langle T_{\frac{r\alpha}{\nu}} g, T_{m\alpha} M_{n\beta} \gamma \big\rangle \cdot T_{m\alpha} M_{n\beta}\, g \\
&= \frac{1}{\nu} \sum_{r=0}^{\nu-1} e^{2\pi i \frac{sr}{\nu}} \cdot T_{-\frac{r\alpha}{\nu}} P_{\mathcal{G}} T_{\frac{r\alpha}{\nu}} g,
\end{aligned} \tag{4.5}$$

where $P_{\mathcal{G}}$ is the orthogonal projection onto $\mathcal{G}$. Equation (4.5) shows that the vectors

$$v = \big(M_{-\frac{s}{\alpha}} P M_{\frac{s}{\alpha}} g\big)_{s=0}^{\nu-1} \qquad \text{and} \qquad u = \big(T_{-\frac{r\alpha}{\nu}} P_{\mathcal{G}} T_{\frac{r\alpha}{\nu}} g\big)_{r=0}^{\nu-1}$$

in $\left(L^2(\mathbb{R})\right)^\nu$ satisfy $F_\omega u = \sqrt{\nu}\cdot v$, where F_ω is the DFT-matrix $F_\omega = \nu^{-1/2}\left(\omega^{sr}\right)_{s,r=0}^{\nu-1}$ with $\omega = e^{2\pi i/\nu}$.

With this preparation, we now prove the equivalence of the statements (i)–(iv).

(i)⇔(ii): If $T_{\frac{\alpha}{\nu}} g \in \mathcal{G}$, then Lemma 2.2 shows that $T_{\frac{r\alpha}{\nu}} g \in \mathcal{G}$ for all $r \in \mathbb{Z}$, so that $T_{-\frac{r\alpha}{\nu}} P_{\mathcal{G}} T_{\frac{r\alpha}{\nu}} g = g$ for all $r \in \mathbb{Z}$. Since $\frac{1}{\nu}\sum_{r=0}^{\nu-1} e^{2\pi i \frac{sr}{\nu}} = \delta_{s,0}$ for $s \in \{0,\dots,\nu-1\}$, Property (ii) then follows from (4.5). Conversely, if (ii) holds, then $v = (g,0,\dots,0)^\top$, which implies that $u = \sqrt{\nu}\cdot F_\omega^* v = (g,g,\dots,g)^\top$. In particular, $T_{-\alpha/\nu} P_{\mathcal{G}} T_{\alpha/\nu} g = g$, i.e., $T_{\alpha/\nu}\, g \in \mathcal{G}$.

(ii)⇒(iii): Since $Pg = g$, it is a consequence of (4.2) that $P|_{\mathcal{L}_0} = I|_{\mathcal{L}_0}$. Furthermore, for $s \in \{1,\dots,\nu-1\}$ and $k,\ell \in \mathbb{Z}$, Equation (4.2) implies

$$P T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} M_{\frac{s}{\alpha}} g = T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} P M_{\frac{s}{\alpha}} g = 0,$$

which shows $P|_{\mathcal{L}_s} = 0$. By using these observations and by noting $\mathcal{L}_r = M_{r/\alpha}\mathcal{L}_0$, we see for $r,s \in \{0,\dots,\nu-1\}$ that $P_r|_{\mathcal{L}_r} = M_{r/\alpha} P M_{-r/\alpha}|_{\mathcal{L}_r} = I|_{\mathcal{L}_r}$ and furthermore $P_r|_{\mathcal{L}_s} = M_{r/\alpha} P M_{-r/\alpha}|_{\mathcal{L}_s} = 0$ for $s \neq r$. Hence, the sum $\mathcal{K} = \mathcal{L}_0 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}$ is direct, and $P_s|_{\mathcal{K}} = M_{s/\alpha} P M_{-s/\alpha}|_{\mathcal{K}}$ is the projection onto $\mathcal{L}_s$ with respect to this decomposition. Finally, since $\gamma \in \mathcal{K}$ and since $\mathcal{K}$ is invariant under $T_{k/\beta} M_{\ell/\alpha}$, it follows by definition of $P_s = (\alpha\beta)^{-1} M_{s/\alpha}\, S_{\frac{1}{\beta}\mathbb{Z}\times\frac{\nu}{\alpha}\mathbb{Z},\gamma,g}\, M_{-s/\alpha}$ that $P_s|_{\mathcal{K}^\perp} = 0$. Therefore, P_s is the projection onto $\mathcal{L}_s$ with respect to the decomposition $L^2(\mathbb{R}) = (\mathcal{L}_0 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}) \oplus \mathcal{K}^\perp$.

Finally, we show that $\mathcal{F}_s$ is a frame for $\mathcal{L}_s$, where it clearly suffices to show this for $s = 0$. Since $\mathcal{F}_0$ is a Bessel sequence, [8, Corollary 5.5.2] shows that we only need to prove that the synthesis operator

$$D: \quad \ell^2(\mathbb{Z}^2) \to L^2(\mathbb{R}), \qquad (c_{k,\ell})_{k,\ell\in\mathbb{Z}} \mapsto \sum_{k,\ell\in\mathbb{Z}} c_{k,\ell}\, T_{\frac{k}{\beta}}\, M_{\frac{\ell\nu}{\alpha}}\, g$$

has closed range $\operatorname{ran} D = \mathcal{L}_0$. By definition of $\mathcal{L}_0 = \overline{\operatorname{span}}\,\mathcal{F}_0$, we see $\operatorname{ran} D \subset \mathcal{L}_0$. Conversely, if $f \in \mathcal{L}_0$, then $f = Pf = (\alpha\beta)^{-1} S_{\gamma,g} f = (\alpha\beta)^{-1} D c \in \operatorname{ran} D$ for the sequence $c = (c_{k,\ell})_{k,\ell\in\mathbb{Z}} \in \ell^2(\mathbb{Z}^2)$ given by $c_{k,\ell} = \langle f, T_{k/\beta} M_{\ell\nu/\alpha}\gamma\rangle$.

(iii)⇒(ii): Since $P = (\alpha\beta)^{-1} S_{\gamma,g}$, we see by definition of $S_{\gamma,g}$ that $\operatorname{ran} P \subset \mathcal{L}_0$. Hence, $Pg - g \in \mathcal{L}_0$. On the other hand, again as a consequence of $\operatorname{ran} P \subset \mathcal{L}_0$ we see that $M_{s/\alpha} P M_{-s/\alpha}\, g \in \mathcal{L}_s$, so that Equation (4.4) implies

$$\mathcal{L}_0 \ni P\, g - g = P\, g - P_{\mathcal{K}}\, g = -\sum_{s=1}^{\nu-1} M_{s/\alpha}\, P\, M_{-s/\alpha}\, g \in \mathcal{L}_1 + \cdots + \mathcal{L}_{\nu-1}, \tag{4.6}$$

and thus $Pg = g$ since the sum $\mathcal{L}_0 + \cdots + \mathcal{L}_{\nu-1}$ is direct. Similarly, for any $s \in \{1,\dots,\nu-1\}$ we get because of $\operatorname{ran} P \subset \mathcal{L}_0$ that $M_{s/\alpha} P M_{-s/\alpha}\, g \in \mathcal{L}_s$; but this implies as in Equation (4.6) that

$$\mathcal{L}_s \ni M_{\frac{s}{\alpha}} P M_{-\frac{s}{\alpha}}\, g = P_{\mathcal{K}} g - \sum_{r\neq s} M_{\frac{r}{\alpha}} P M_{-\frac{r}{\alpha}}\, g \in \mathcal{L}_0 + \operatorname{span}\{\mathcal{L}_r : r \neq s\} = \operatorname{span}\{\mathcal{L}_r : r \neq s\}.$$

Again, since $\mathcal{L}_0 + \cdots + \mathcal{L}_{\nu-1}$ is a direct sum, this implies $P M_{-\frac{s}{\alpha}}\, g = 0$ for $s = 1,\dots,\nu-1$. Since P commutes with $M_{\pm\nu/\alpha}$ (see (4.2)), we have $P M_{(\nu-s)/\alpha}\, g = 0$ and therefore $P M_{s/\alpha}\, g = 0$ for $s = 1,\dots,\nu-1$.

(i)⇒(iv): Note that $(\gamma, \alpha\mathbb{Z}\times\beta\mathbb{Z})$ is a frame sequence and that $\mathcal{G} = \mathcal{G}(\gamma, \alpha\mathbb{Z}\times\beta\mathbb{Z})$. Further, Lemma 2.2 shows that $T_{\alpha/\nu} g \in \mathcal{G}$ if and only if $\mathcal{G}$ is invariant under $T_{\alpha/\nu}$, if and

only if $T_{\alpha/\nu}\gamma \in \mathcal{G}$. Let us consider the setting above with g and γ interchanged: Define

$$\mathcal{F}_s^* := \big\{T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} M_{\frac{s}{\alpha}} \gamma \colon k, \ell \in \mathbb{Z}\big\}, \qquad s = 0, \dots, \nu - 1.$$

Then, by using the implication "(i)⇒(iii)" in this setting, we get $\mathcal{K}^* = \mathcal{L}_0^* \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}^*$, where $\mathcal{L}_s^* := \overline{\operatorname{span}}\, \mathcal{F}_s^*$ and $\mathcal{K}^* = \overline{\operatorname{span}}\, \mathcal{F}^*$ with $\mathcal{F}^* := \big\{T_{k/\beta} M_{\ell/\alpha} \gamma \colon k, \ell \in \mathbb{Z}\big\}$. Note that $\mathcal{K} = \mathcal{K}^*$ by Equation (4.1).

We have $S_{g,\gamma} = S_{\gamma,g}^*$. Hence, $M_{s/\alpha} P^* M_{-s/\alpha}$ is the projection onto $\mathcal{L}_s^*$ with respect to the decomposition $L^2(\mathbb{R}) = (\mathcal{L}_0^* \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}^*) \oplus \mathcal{K}^\perp$. In particular, using the general formula $(\ker T)^\perp = \overline{\operatorname{ran} T^*}$ for a bounded operator $T : \mathcal{H} \to \mathcal{H}$ (see [9, Remarks after Theorem II.2.19]) and the elementary identity $(A + B)^\perp = A^\perp \cap B^\perp$ for subspaces $A, B \subset \mathcal{H}$, we get

$$\mathcal{L}_0^* = \overline{\operatorname{ran} P^*} = (\ker P)^\perp = \big(\mathcal{L}_1 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}\big)^\perp \cap \mathcal{K}.$$

For $k, \ell \in \mathbb{Z}$ and $s \in \{1, \dots, \nu - 1\}$ this implies $\big\langle T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} \gamma, M_{\frac{s}{\alpha}} g \big\rangle = 0$, which is equivalent to (iv).

(iv)⇒(ii): For $s \in \{1, \dots, \nu - 1\}$, we have

$$P M_{\frac{s}{\alpha}} g = \frac{1}{\alpha\beta} \sum_{k,\ell \in \mathbb{Z}} \big\langle M_{\frac{s}{\alpha}} g, T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} \gamma \big\rangle T_{\frac{k}{\beta}} M_{\frac{\ell\nu}{\alpha}} g = 0.$$

Thanks to Equation (4.4), this implies $g = P_{\mathcal{K}} g = P g$. Overall, we have thus shown $P M_{s/\alpha}\, g = \delta_{s,0}\, g$ for all $s \in \{0, \dots, \nu - 1\}$. □

Note that with $P := (\alpha\beta)^{-1} S_{\gamma,g}$, the condition $Pf = f$ for $f \in \mathcal{L}_0$ means that $(\alpha\beta)^{-1}\gamma$ is a dual window for the frame sequence $\mathcal{F}_0 = \big(g, \frac{1}{\beta}\mathbb{Z} \times \frac{\nu}{\alpha}\mathbb{Z}\big)$. However, it is possible that $\gamma \notin \mathcal{L}_0 = \overline{\operatorname{span}}\, \mathcal{F}_0$.

5. Proof of the main theorem

In this section, we prove our main result, Theorem 1.1, which we state here once more for the convenience of the reader.

Theorem 1.1. *If $g \in S_0(\mathbb{R})$ and $\Lambda \subset \mathbb{R}^2$ is a lattice such that the Gabor system (g, Λ) is a Riesz basis for its closed linear span $\mathcal{G}(g, \Lambda)$, then the time-frequency shifts $T_a M_b$ that leave $\mathcal{G}(g, \Lambda)$ invariant satisfy $(a, b) \in \Lambda$.*

Proof. The claim is true if Λ has rational density; see [4, Theorem 1]. Thus, assume that Λ has irrational density $d(\Lambda) \in \mathbb{R}\backslash\mathbb{Q}$. Write $\Lambda = A\mathbb{Z}^2$ with an invertible matrix $A \in \mathbb{R}^{2\times 2}$.

Due to the Amalgam Balian-Low theorem [2, Theorem 3.2], it is not possible that $\mathcal{G} := \mathcal{G}(g, \Lambda) = L^2(\mathbb{R})$. Hence, $\mathcal{G} \neq L^2(\mathbb{R})$. Suppose towards a contradiction that $\mathfrak{I}(\mathcal{G}) \supsetneq \Lambda$. According to Lemma 4.1 there exist $B \in \mathbb{R}^{2\times 2}$ with $\det B = 1$ and $\alpha, \beta > 0$ such that $\Lambda_B = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ and $T_{\frac{\alpha}{\nu}} g_B \in \mathcal{G}_B$ for some $\nu \in \mathbb{N}_{\geq 2}$. Set $h := g_B$ and $\mathcal{G}_h := \mathcal{G}_B = \mathcal{G}(h, \alpha\mathbb{Z} \times \beta\mathbb{Z})$. Then $h \in S_0(\mathbb{R})$ by [14, Proposition 12.1.3] and $T_{\frac{\alpha}{\nu}} h \in \mathcal{G}_h$. Furthermore, note that with (g, Λ), also $(h, \alpha\mathbb{Z} \times \beta\mathbb{Z}) = (g_B, \Lambda_B)$ is a Riesz sequence (cf. Subsection 2.3).

Let γ be the canonical dual window for $(h, \alpha\mathbb{Z} \times \beta\mathbb{Z})$. By Ron-Shen duality (see [14, Theorem 7.4.3]), the adjoint system $(h, \frac{1}{\beta}\mathbb{Z} \times \frac{1}{\alpha}\mathbb{Z})$ is a frame for $L^2(\mathbb{R})$. Let $\gamma^\natural$ denote

the canonical dual window of $(h, \frac{1}{\beta}\mathbb{Z} \times \frac{1}{\alpha}\mathbb{Z})$, and note by Wexler-Raz orthogonality (see [14, Theorem 7.3.1]) that $\langle h, \gamma^\natural \rangle = (\alpha\beta)^{-1}$. Next, note that [26, Theorem 2.3] shows $\gamma^\natural = (\alpha\beta)^{-1}\gamma$ and hence $\langle h, \gamma \rangle = \alpha\beta \cdot \langle h, \gamma^\natural \rangle = 1$, which will be used below.

Since $T_{\alpha/\nu} h \in \mathcal{G}_h$, Theorem 4.2 implies that $P_0 := (\alpha\beta)^{-1} S_{\frac{1}{\beta}\mathbb{Z} \times \frac{\nu}{\alpha}\mathbb{Z}, \gamma, h}$ is an idempotent (i.e., $P_0^2 = P_0$). We now wish to apply Theorem A.1 to derive a contradiction. To this end, first note that $\alpha\beta = \big(d(\Lambda_B)\big)^{-1} = |\det BA| = |\det A| = \big(d(\Lambda)\big)^{-1} \in \mathbb{R} \backslash \mathbb{Q}$. Next, set $U := M_\beta$ and $V := T_{\frac{\alpha}{\nu}}$. A direct calculation shows that

$$UV = e^{2\pi i \theta} VU, \qquad \text{where} \qquad \theta := \tfrac{\alpha\beta}{\nu} \in \mathbb{R} \backslash \mathbb{Q}.$$

Note that also $\gamma \in S_0(\mathbb{R})$; see [1, Theorem 7]. Hence, we may use Equation (2.7) and obtain

$$P_0 = \frac{1}{\nu} \sum_{m,n \in \mathbb{Z}} \big\langle h, T_{\frac{m\alpha}{\nu}} M_{n\beta} \gamma \big\rangle T_{\frac{m\alpha}{\nu}} M_{n\beta} = \sum_{m,n \in \mathbb{Z}} \frac{1}{\nu} \big\langle h, V^m U^n \gamma \big\rangle \, V^m U^n, \tag{5.1}$$

with coefficient sequence $a = (a_{m,n})_{m,n \in \mathbb{Z}} := \big(\frac{1}{\nu} \langle h, V^m U^n \gamma \rangle\big)_{m,n \in \mathbb{Z}} \in \ell^1(\mathbb{Z}^2)$. Therefore, Theorem A.1 shows that $\frac{1}{\nu} = a_{0,0} \in \mathbb{Z} + \theta\mathbb{Z}$, say $\frac{1}{\nu} = m + n\theta$ for some $m, n \in \mathbb{Z}$. We must have $n \neq 0$, since otherwise $\frac{1}{\nu} = m \in \mathbb{Z}$, in contradiction to $\nu \geq 2$. Thus, $\theta = \frac{1}{n\nu} - \frac{m}{n} \in \mathbb{Q}$, which is the desired contradiction, since $\theta = \frac{\alpha\beta}{\nu}$ is irrational. □

Remark 5.1. On a first look, it might appear as if the proof of Theorem 1.1 would also apply in case of $g \in \mathbb{H}^1$: First, the classical Balian-Low theorem implies that $\mathcal{G} := \mathcal{G}(g, \Lambda) \subsetneq L^2(\mathbb{R})$, so that Lemma 4.1 allows the reduction to a Gabor Riesz sequence (h, Λ) with $h \in \mathbb{H}^1$, a separable lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$, and an additional time-frequency shift of the form $T_{\alpha/\nu}\, h \in \mathcal{G}$. One can then apply Theorem 4.2 to see that that $L^2(\mathbb{R}) = \mathcal{L}_0 \boxplus \cdots \boxplus \mathcal{L}_{\nu-1}$. In the S_0-case, we then employed Janssen's representation (5.1) for the projection $P_0 = (\alpha\beta)^{-1} S_{\frac{1}{\beta}\mathbb{Z} \times \frac{\nu}{\alpha}\mathbb{Z}, \gamma, h}$, which then led to success in the proof of Theorem 1.1, thanks to existing results concerning the structure of the irrational rotation algebra. However, in the case $h \in \mathbb{H}^1$ the series in (5.1) might not converge *in operator norm*, so that one does not know whether P_0 belongs to the irrational rotation algebra. Thus, the proof breaks down at this point.

Acknowledgments

A. Caragea acknowledges support by the DFG Grant PF 450/11-1. D.G. Lee acknowledges support by the DFG Grants PF 450/6-1 and PF 450/9-1. F. Voigtlaender acknowledges support by the DFG in the context of the Emmy Noether junior research group VO 2594/1-1.

The authors take pleasure in thanking Karlheinz Gröchenig for suggesting the idea of considering the trace on the irrational rotation algebra. In fact, we had established the characterization in Theorem 4.2, but were unable to prove that a projection as in (5.1) cannot exist, until Karlheinz Gröchenig suggested to us at SampTA 2019 that we should try to use the trace on the irrational rotation algebra. Without his hint, we would probably not have managed to prove our main result, Theorem 1.1.

In addition, the authors would like to thank Radu Balan, Ilya Krishtal, Götz E. Pfander, and Jordy van Velthoven for fruitful discussions and hints.

A. Appendix

In this section, we make use of a deep result concerning the structure of the *irrational rotation algebra* $\mathcal{A}_\theta$ (see [11, 24, 25]) to prove the following auxiliary statement, which is a crucial ingredient for the proof of our main result, Theorem 1.1. As usual, we denote the set of all bounded linear operators from a Hilbert space $\mathcal{H}$ into itself by $\mathcal{B}(\mathcal{H})$.

Theorem A.1. *Let $\mathcal{H} \neq \{0\}$ be a Hilbert space and let $U, V \in \mathcal{B}(\mathcal{H})$ be unitary and such that $UV = e^{2\pi i\theta} VU$, for some $\theta \in \mathbb{R}\backslash\mathbb{Q}$. If $a = (a_{k,\ell})_{k,\ell\in\mathbb{Z}} \in \ell^1(\mathbb{Z}^2)$ is such that the operator $P_a := \sum_{k,\ell\in\mathbb{Z}} a_{k,\ell} V^k U^\ell$ satisfies $P_a^2 = P_a$, then $a_{0,0} \in \mathbb{Z} + \theta\mathbb{Z}$.*

The proof will make use of some parts of the theory of C^*-algebras, which we recall here for the convenience of the reader, based on [23]. Readers familiar with C^*-algebras will probably want to skip this part—except possibly Lemma A.2.

A C^*-algebra is a (complex) Banach algebra $(A, \|\cdot\|)$, additionally equipped with a map $A \to A, x \mapsto x^*$ (called the *involution* on A), satisfying the following properties:

- $(x+y)^* = x^* + y^*$, $(\lambda\, x)^* = \overline{\lambda}\, x^*$, and $(x\, y)^* = y^*\, x^*$ and $(x^*)^* = x$ for all $x, y \in A$ and $\lambda \in \mathbb{C}$;
- $\|x^*\| = \|x\|$ and $\|x^* x\| = \|x\|^2$ for all $x \in A$.

An element $p \in A$ is called an *idempotent* if $p^2 = p$. An idempotent p is called a *projection* if additionally $p = p^*$ holds. A C^*-algebra A is called *unital* if it contains a (necessarily unique) element $1 \in A$ satisfying $1 \neq 0$ and $x\,1 = 1\,x = x$ for all $x \in A$. In a unital C^*-algebra A, an element $x \in A$ is called *unitary* if $x^* x = 1 = x x^*$. If A is a unital C^*-algebra and $a \in A$, then $\sigma(a^* a) \subset [0, \infty)$; see [23, Theorem 2.2.4]. Here, $\sigma(b) = \{\lambda \in \mathbb{C} \colon b - \lambda 1 \text{ not invertible in } A\}$.

Lemma A.2. *Any idempotent e in a unital C^*-algebra A is similar to a projection $p \in A$. That is, there exist a projection $p \in A$ and an invertible element $a \in A$ such that $e = a^{-1} p a$.*

Proof. We set $b := e^* - e$ and $z := 1 + b^* b$. Note that z is invertible since $\sigma(b^* b) \subset [0, \infty)$. We have

$$ez = e + (e - ee^*)(e^* - e) = ee^* e = e + (e - e^*)(e^* e - e) = ze.$$

Consequently, $ez^{-1} = z^{-1} e$ and, as $z = z^*$, also $e^* z^{-1} = z^{-1} e^*$. Now, define the element $p := ez^{-1} e^*$. We have $p^* = p$. Furthermore, since we just saw that z^{-1} commutes with e and e^* and that $ee^* e = ze$, we also see that $p^2 = z^{-2}(ee^* e)e^* = z^{-1} ee^* = p$. Hence, p is a projection. We further observe that $ep = p$ and $pe = ez^{-1} e^* e = z^{-1} ee^* e = z^{-1} ze = e$. Set $a := 1 - p + e$. Then we see because of

$$(1 \mp p \pm e)(1 \pm p \mp e) = 1 \pm p \mp e \mp p - p + e \pm e + p - e = 1$$

that a is invertible with $a^{-1} = 1 + p - e$. Hence, from $ae = e - pe + e = e$ we obtain

$$aea^{-1} = e(1 + p - e) = e + ep - e = ep = p,$$

which proves the lemma. □

A closed subspace B of a C^*-algebra A is called a *C^*-subalgebra* of A if it is closed under both multiplication and involution. It is clear that B is then itself a C^*-algebra. As usual, given a subset $S \subset A$, there is a smallest (with respect to inclusion) C^*-subalgebra of A containing S. We call it the *C^*-algebra generated by S*, and denote it by $C^*(S)$.

A map $\varphi : A \to B$ between two C^*-algebras A and B is called a $*$-homomorphism if it is linear and satisfies $\varphi(x\,y) = \varphi(x)\,\varphi(y)$ as well as $\varphi(x^*) = [\varphi(x)]^*$ for all $x, y \in A$. A bijective $*$-homomorphism is called a $*$-isomorphism. Any $*$-homomorphism $\varphi : A \to B$ necessarily satisfies $\|\varphi(x)\|_B \leq \|x\|_A$ for all $x \in A$, and is hence continuous; see [23, Theorem 2.1.7].

Proof of Theorem A.1. We will make use of the so-called *irrational rotation algebra* $\mathcal{A}_\theta$, as introduced for instance in [11, Chapter VI]. The actual definition of this algebra is not relevant for us; we will only need to know that it satisfies the following properties:

- $\mathcal{A}_\theta$ is a unital C^*-algebra;
- The algebra $\mathcal{A}_\theta$ is *universal* among all unital C^*-algebras generated by unitary elements U, V satisfying $UV = e^{2\pi i\theta}VU$. Thus, defining $\mathcal{A} := C^*(U, V)$ as a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$ with U, V as in the statement of Theorem A.1, there is a $*$-isomorphism $\varphi : \mathcal{A} \to \mathcal{A}_\theta$; this follows from [11, Theorem VI.1.4].
- As shown in [11, Corollary VI.1.2 and Proposition VI.1.3], there is a unique (*unital*) *trace* $\tau : \mathcal{A}_\theta \to \mathbb{C}$. By definition of a trace, this means in particular that τ is linear and continuous, satisfying $\tau(1) = 1$ and $\tau(xy) = \tau(yx)$ for all $x, y \in \mathcal{A}_\theta$.
- For any projection $p \in \mathcal{A}_\theta$, we have $\tau(p) \in \mathbb{Z} + \theta\mathbb{Z}$; see [25, Theorem 1.2]. We remark that this result was originally proven in [24].

Let us define $\tau^\natural := \tau \circ \varphi$, and note that $\tau^\natural : \mathcal{A} \to \mathbb{C}$ is continuous. It is easy to see that $\tau^\natural$ is linear with $\tau^\natural(\mathrm{id}_\mathcal{H}) = 1$ and $\tau^\natural(AB) = \tau^\natural(BA)$ for all $A, B \in \mathcal{A}$; this is called the *cyclicity* of the trace. Next, from the relation $UV = e^{2\pi i\theta}VU$, we immediately get for $k, \ell \in \mathbb{Z}$ that

$$V^k U^\ell = e^{-2\pi i\ell\theta} V^{k-1} U^\ell V = e^{-2\pi ik\theta} U V^k U^{\ell-1}.$$

Thus, noting that $V^k U^\ell \in \mathcal{A}$, we obtain $\tau^\natural(V^k U^\ell) = e^{-2\pi i\ell\theta}\tau^\natural(V^k U^\ell) = e^{-2\pi ik\theta}\tau^\natural(V^k U^\ell)$ by cyclicity. As θ is irrational, this implies $\tau^\natural(V^k U^\ell) = \delta_{\ell,0}\delta_{k,0}$. Next, since we have $\|V^k U^\ell\| = 1$ for all $k, \ell \in \mathbb{Z}$ and since $a \in \ell^1(\mathbb{Z}^2)$, we see that $P_a = \sum_{k,\ell\in\mathbb{Z}} a_{k,\ell} V^k U^\ell \in \mathcal{A}$, with unconditional convergence of the defining series. Hence,

$$\tau^\natural(P_a) = \sum_{k,\ell\in\mathbb{Z}} a_{k,\ell} \cdot \tau^\natural(V^k U^\ell) = a_{0,0}.$$

Since $P_a^2 = P_a$ and since $\varphi : \mathcal{A} \to \mathcal{A}_\theta$ is a $*$-homomorphism, we see that $e := \varphi(P_a) \in \mathcal{A}_\theta$ is an idempotent. By Lemma A.2 there exist $b, p \in \mathcal{A}_\theta$ such that b is invertible, p is a projection, and $e = b^{-1}pb$. Thanks to the cyclicity of the trace, we thus see that $a_{0,0} = \tau^\natural(P_a) = \tau(e) = \tau(b^{-1}pb) = \tau(p) \in \mathbb{Z} + \theta\mathbb{Z}$, as claimed. □

References

[1] R. Balan, P.G. Casazza, C. Heil, and Z. Landau, *Density, overcompleteness, and localization of frames. II. Gabor systems*, J. Fourier Anal. Appl. 12(3) (2006), 309–344.

[2] J.J. Benedetto, C. Heil, and D.F. Walnut, *Differentiation and the Balian-Low theorem*, J. Fourier Anal. Appl. 1 (1995), 355–402.

[3] C. Cabrelli, D.G. Lee, U. Molter, and G.E. Pfander, *Time-frequency shift invariance of Gabor spaces generated by integer lattices*, J. Math. Anal. Appl. 474 (2019), 1289–1305.

[4] C. Cabrelli, U. Molter, and G.E. Pfander, *Time-frequency shift invariance and the Amalgam Balian-Low theorem*, Appl. Comput. Harmon. Anal. 41 (2016), 677–691.

[5] C. Cabrelli, U. Molter, and G.E. Pfander, *An Amalgam Balian-Low Theorem for symplectic lattices of rational density*, Proceedings International Conference on Sampling Theory and Applications, Washington DC, 2015.
[6] A. Caragea, D.G. Lee, G.E. Pfander, and F. Philipp, *A Balian-Low theorem for subspaces*, J. Fourier Anal. Appl. 25 (2019), 1673–1694.
[7] A. Caragea, D.G. Lee, F. Philipp, and F. Voigtlaender, *A quantitative subspace Balian-Low theorem*, Appl. Comput. Harmon. Anal. 55 (2021) 368–404.
[8] O. Christensen, *An introduction to frames and Riesz bases*, Birkhäuser/Springer, [Cham], 2016.
[9] J. B. Conway, *A course in functional analysis*, second edition, Springer-Verlag, New York, 1990.
[10] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inform. Theory 36(5) (1990), 961–1005.
[11] K. R. Davidson, *C^*-algebras by example*, American Mathematical Society, Providence, RI, 1996.
[12] H.G. Feichtinger, *On a new Segal algebra*, Monatsh. Math. 92 (1981), 269–289.
[13] H.G. Feichtinger and K. Gröchenig, *Gabor frames and time-frequency analysis of distributions*, J. Funct. Anal. 146 (1997), 464–495.
[14] K. Gröchenig, *Foundations of time-frequency analysis*, Birkhäuser, Boston, Basel, Berlin, 2001.
[15] K. Gröchenig and M. Leinert, *Wiener's lemma for twisted convolution and Gabor frames*, J. Am. Math. Soc. 17 (2003), 1–18.
[16] C. Heil and R. Tinaztepe, *Modulation spaces, BMO, and the Balian-Low theorem*, Sampl. Theory Signal Image Process. 11 (2012), 25–41.
[17] E. Hewitt and K.A. Ross, *Abstract harmonic analysis, Volume I: Structure of topological groups, Integration Theory, Group Representations*, second edition, Springer-Verlag, 1963.
[18] N. Jacobson, *Basic Algebra I*, second edition, W.H. Freeman and Company, New York, 1985.
[19] M.S. Jakobsen, *On a (no longer) new Segal algebra: A review of the Feichtinger algebra*, J. Fourier Anal. Appl. 24 (2018), 1579–1660.
[20] A.J.E.M. Janssen, *Duality and biorthogonality for Weyl-Heisenberg frames*, J. Fourier Anal. Appl. 1 (1995), 403–436.
[21] G. Leoni, *A first course in Sobolev spaces*, second edition, American Mathematical Society, 2017.
[22] C.C. MacDuffee, *The theory of matrices*, Chelsea Publishing Company, New York, 1946.
[23] G. J. Murphy, *C^*-algebras and operator theory*, Academic Press, Inc., Boston, MA, 1990.
[24] M. Pimsner and D. Voiculescu, *Imbedding the irrational rotation C^*-algebra into an AF-algebra*, J. Operator Theory 4(2) (1980), 201–210.
[25] M. A. Rieffel, *C^*-algebras associated with irrational rotations*, Pacific J. Math. 93(2) (1981), 415–429.
[26] A. Ron and Z. Shen, *Weyl-Heisenberg frames and Riesz bases in $L_2(\mathbb{R}^d)$*, Duke Math. J. 89 (1997), 237–282.
[27] W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill, Inc., 1987.

Author Affiliations

A. Caragea: KU Eichstätt–Ingolstadt, Mathematisch–Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany

Email address: andrei.caragea@gmail.com

D.G. Lee: KU Eichstätt–Ingolstadt, Mathematisch–Geographische Fakultät, Ostenstrasse 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany

Email address: daegwans@gmail.com

F. Philipp: Technische Universität Ilmenau, Institute for Mathematics, Weimarer Strasse 25, D-98693 Ilmenau, Germany

Email address: friedrich.philipp@tu-ilmenau.de

F. Voigtlaender: KU Eichstätt-Ingolstadt, Lehrstuhl Reliable Machine Learning, Ostenstrasse 26, 85072 Eichstätt, Germany

Email address: felix.voigtlaender@ku.de

6.4 Paper 4

Neural network approximation and estimation of classifiers with classification boundary in a Barron class

Andrei Caragea*, Philipp Petersen†, Felix Voigtlaender‡

March 11, 2022

Abstract

We prove bounds for the approximation and estimation of certain binary classification functions using ReLU neural networks. Our estimation bounds provide a priori performance guarantees for empirical risk minimization using networks of a suitable size, depending on the number of training samples available. The obtained approximation and estimation *rates* are independent of the dimension of the input, showing that the curse of dimensionality can be overcome in this setting; in fact, the input dimension only enters in the form of a polynomial factor. Regarding the regularity of the target classification function, we assume the interfaces between the different classes to be locally of Barron-type. We complement our results by studying the relations between various Barron-type spaces that have been proposed in the literature. These spaces differ substantially more from each other than the current literature suggests.

1 Introduction

This article concerns the approximation and statistical estimation of *high-dimensional, discontinuous functions by neural networks*. More precisely, we study a certain class of target functions for classification problems, such as those encountered when automatically labeling images. For such problems, deep learning methods—based on the training of deep neural networks with gradient-based methods—achieve state of the art performance [35, 33]. The underlying functional relationship of such an (image) classification task is typically extremely high-dimensional. For example, the most widely used image data-bases used to benchmark classification algorithms are MNIST [36] with 28×28 pixels per image, CIFAR-10/CIFAR-100 [32] with 32×32 pixels per image and ImageNet [15, 33] which contains high-resolution images that are typically down-sampled to 256×256 pixels. Compared to practical applications, these benchmark datasets are relatively low-dimensional. Yet, already for MNIST, the simplest of those databases, the input dimension for the classification function is $d = 784$.

2020 *Mathematics Subject Classification.* 68T07, 41A25, 41A46, 42B35, 46E15

Key words and phrases. ReLU neural networks, Deep neural networks, Approximation, Empirical Risk minimization, Classification, Barron spaces.

Funding. AC thankfully acknowledges support by the German Research Foundation (DFG), project number PF 450/11–1. FV thankfully acknowledges support by the German Research Foundation (DFG) in the context of the Emmy Noether junior research group VO 2594/1–1.

*KU Eichstätt–Ingolstadt, Mathematisch–Geographische Fakultät, Ostenstraße 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany

†Faculty of Mathematics and Research Platform Data Science @ Uni Vienna, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria, e-mail: philipp.petersen@univie.ac.at

‡Faculty of Mathematics, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria, e-mail: felix.voigtlaender@univie.ac.at

It is well known in classical approximation theory that high-dimensional approximation problems typically suffer from the so-called *curse of dimensionality* [11, 43]. This term describes the fact that the problems of approximation or estimation typically become exponentially more complex for increasing input dimension. Yet, given the overwhelming success of deep learning methods in practice, high-dimensional input does not seem to be a prohibitive factor.

One of the first theoretical results in neural network approximation offering a partial explanation for this ostensible clash of theory and practical observations was found in [4]. There it was demonstrated that for a certain class of functions with variation bounded in a suitable sense (these functions are, in particular, Lipschitz continuous), neural networks with one hidden layer of N neurons achieve an approximation accuracy of the order of $N^{-1/2}$ in the $L^2(\mu)$-norm for a probability measure μ on a d-dimensional ball. Notably, this approximation rate is *independent* of the ambient dimension d. Neural networks can thus overcome the curse of dimensionality for this class of functions. This is particularly significant, since the considered class of functions (nowadays so-called a *Barron class*) is so large that every *linear* method of approximation for it is subject to the curse of dimensionality; see [4, Theorem 6]. The result of [4] has since been extended and generalized in various ways; we refer to Subsection 1.2 for an overview.

In contrast to the (Lipschitz) continuous functions considered in [4], our interest lies in the approximation of *discontinuous classification functions*. Such functions are of the form $\sum_{k=1}^K q_k \mathbb{1}_{\Omega_k}$, where the sets $\Omega_k \subset \mathbb{R}^d$ are disjoint and describe $K+1 \in \mathbb{N}$ classes (we also consider $(\bigcup_{k=1}^K \Omega_k)^c$ as a class). Here $\mathbb{1}_{\Omega_k}$ denotes the indicator function of Ω_k; that is, $\mathbb{1}_{\Omega_k}(x) = 1$ if $x \in \Omega_k$ and 0 otherwise. Moreover, $(q_k)_{k=1}^K$ correspond to the labels of the classes and could for example be unit vectors, as in $q_k = e_k \in \mathbb{R}^K$ for $k = 1, \dots K$, in the case of one-hot-encoding or $(q_k)_{k=1}^K \subset \mathbb{N}$ for integer labels. These functions were discussed previously in [45] and [28, 29], where it was shown that the regularity of the boundary determines the approximation rate. However, the results of [45, 28, 29] are based on classical notions of smoothness regarding the boundary and suffer from the curse of dimensionality. In this article, we assume the class interfaces to be locally of bounded variation in the sense used in [4]. The following subsection gives an overview of our results and the employed proof methods.

1.1 Our results

We present upper and lower bounds for the approximation and estimation of classification functions using deep neural networks with the ReLU activation function as hypothesis space. The classification functions that we consider are of the form $\sum_{k=1}^K q_k \mathbb{1}_{\Omega_k}$, where each $\Omega_k \subset \mathbb{R}^d$ is an open set such that $\partial\Omega_k$ is locally a $d-1$-dimensional Barron function. In the sequel, we only consider the case of two complementary classes, that is, $K = 1$; the generalization to more summands is straightforward.

Measure of approximation accuracy: In contrast to ReLU neural networks, the indicator functions $\mathbb{1}_\Omega$ are discontinuous. Uniformly approximating $\mathbb{1}_\Omega$ using ReLU neural networks is thus impossible. Therefore, we measure the approximation error in terms of the measure of the set on which the true function and the approximation differ; since both functions are bounded in absolute value by 1, this easily implies corresponding error estimates in $L^p(\mu)$ for arbitrary exponents $p \in [1, \infty)$. Here, we consider those measures μ that are *tube compatible* with an exponent $\alpha \in (0, 1]$, meaning that the measure around any ε tube of the graph of a function decays like ε^α as $\varepsilon \downarrow 0$. This notion is broad enough to include a large class of product measures on $\mathbb{R}^d$, as well as all measures of the form $d\mu = f d\nu$, where f is a bounded density and ν a tube compatible measure. We also show in Section 6 that for general (not

tube compatible) measures, no nontrivial approximation rates can be derived.

Regularity assumptions on the class interfaces: Similar to the notion of C^k-domains or Lipschitz domains, we assume the boundary $\partial\Omega \subset \mathbb{R}^d$ to be locally parametrized by Barron-regular functions. Here, inspired by [4], we say that a function $f : U \subset \mathbb{R}^k \to \mathbb{R}$ is of Barron-type, if it can be represented as

$$f(x) = c + \int_{\mathbb{R}^k} \big(e^{i\langle x,\xi\rangle} - 1\big) F(\xi)\, d\xi \quad \text{for } x \in U, \qquad \text{where} \qquad \int_{\mathbb{R}^k} |\xi| \cdot |F(\xi)|\, d\xi < \infty. \quad (1.1)$$

For more formal discussion of our assumptions, we refer to Definitions 2.1 and 3.3. We also remark that recently other notions of Barron-type functions have been proposed in the literature; these are discussed briefly below and in full detail in Section 7.

Upper bounds on the approximation rate: A simplified but honest version of our main approximation result reads as follows:

Theorem 1.1. *Let μ be a finite measure, tube compatible with exponent $\alpha \in (0, 1]$. Let $\Omega \subset \mathbb{R}^d$ be such that $\partial\Omega$ can locally be parametrized by functions of Barron-type. Then, for every $N \in \mathbb{N}$ the function $\mathbb{1}_\Omega$ can be approximated using ReLU neural networks with three hidden layers and a total of $\mathcal{O}(d+N)$ neurons to accuracy $\mathcal{O}\big(d^{3/(2p)} N^{-\alpha/(2p)}\big)$ in the $L^p(\mu)$ norm. Moreover, the magnitude of the weights in the approximating neural networks can be chosen to be $\mathcal{O}(d+N^{1/2})$.*

For example, if μ is the Lebesgue measure, then $\alpha = 1$. We note that the accuracy of our approximation *does* depend on the dimension, but the dimension enters only as a multiplicative factor which is polynomial in d.

The proof of Theorem 1.1 is structured as follows:

1. We use a classical result of Barron [3] that yields uniform approximation of functions with a bounded Fourier moment. Because of a minor inaccuracy in the original result, we reprove this theorem in Proposition 2.2.

2. Approximation of *horizon functions*. We show that we can efficiently approximate *horizon functions*, meaning functions of the form $\mathbb{1}_{x_1 \le f(x_2,\dots,x_d)}$ where f is a $d-1$ dimensional function of Barron-type. For the proof, we use a) that ReLU neural networks efficiently approximate the Heaviside function, b) the compositional structure of NNs, and c) the approximation result from Step 1.

3. The classification function $\mathbb{1}_\Omega$ is only *locally* represented by horizon functions as in Step 2. Using a *ReLU-based partition of unity*, we show that the result from Step 2 can be improved to an approximation of the full classification function $\mathbb{1}_\Omega$.

The details of the above argument are presented in the proof of Theorem 3.7.

Lower bounds on the approximation rate: We show that the established upper bounds on the approximation rates can, in general, not be significantly improved. More precisely, for the Lebesgue measure $d\mu = \mathbb{1}_{[-1,1]^d} d\lambda$, we show that for the set of classification functions considered above, approximation with $L^1(\mu)$ error decaying asymptotically faster than $N^{-1/2-1/(d-1)}$ for $N \to \infty$ is not possible. For large input dimensions d, this almost matches the upper bound $N^{-1/2}$ from Theorem 1.1.

We prove two forms of this result. First, in Theorem 4.3, we consider neural networks for which the individual weights are suitably quantized and grow at most polynomially with the total number $W \in \mathbb{N}$ of neural network parameters. We show that no sequence of such

neural networks achieves an asymptotic approximation rate faster than $W^{-1/2-1/(d-1)}$. This result follows by showing that efficient approximation of horizon functions implies efficient approximation of the associated interface functions, a technique previously applied in [45]. Then, known entropy bounds for certain Besov spaces contained in the classical Barron spaces can be used; this is inspired by ideas from [4].

For "quantized" networks, we can allow arbitrary network architectures. As our second result, we show in Theorem 4.4 that the assumption of weight quantization can be dropped, provided that the depths of the approximating neural networks are assumed to be uniformly bounded. It is still required, however, that the magnitude of the individual weights only grows polynomially with the total network size. The proof of this second result is based on a previously established "quantization lemma"; see [13, Lemma 3.7] and [21, Lemma VI.8].

Upper bounds on learning: Based on our approximation results, we study the problem of estimating classifier functions of the form described above from a given set of training samples. Precisely, we analyze the performance of the standard empirical risk minimization procedure, where we use the 0-1 loss as the loss function and a suitable class of ReLU neural networks as the hypothesis space.

To describe the result in more detail, let us denote by Φ_S the empirical risk minimizer based on a training sample $S = \big((X_1, Y_1), \dots, (X_m, Y_m)\big)$ with $(X_1, \dots, X_m) \overset{i.i.d.}{\sim} \mathbb{P}$ and $Y_i = \mathbb{1}_\Omega(X_i)$. Assuming that the boundary $\partial\Omega$ is locally parametrized by functions of Barron class and that $\mathbb{P}$ is tube compatible with exponent $\alpha \in (0, 1]$, we derive bounds on the risk of Φ_S, that is, on $\mathbb{P}\big(\Phi_S(X) \neq \mathbb{1}_\Omega(X)\big)$ where $X \sim \mathbb{P}$.

In Theorem 5.1, we show that, if the hypothesis class is a certain set of ReLU neural networks with three hidden layers and $N \sim (dm/\ln(dm))^{1/(1+\alpha)}$ neurons, then—with probability at least $1 - \delta$ with respect to the choice of the training sample S—the risk of any empirical risk minimizer Φ_S is at most

$$\mathcal{O}\bigg(d^{3/2} \cdot \Big(\frac{\ln(dm)}{dm}\Big)^{\alpha/(2+2\alpha)} + \Big(\frac{\ln(1/\delta)}{m}\Big)^{1/2}\bigg).$$

In particular, if $\alpha = 1$, which is the case for the uniform probability measure, then the risk is at most $\mathcal{O}\big(d^{3/4} \ln(dm) \cdot m^{-1/4} + \sqrt{\ln(1/\delta)/m}\,\big)$. This is similar to the estimation bounds established in [5] for Barron regular functions.

Different notions of Barron spaces: In this article we mainly use the Fourier-analytic notion of Barron-type functions as introduced in [4]; see Equation (1.1). We will refer to this space as the *classical Barron space*, or the *Fourier-analytic Barron space*. In recent years, other types of function spaces have been studied under the name "Barron-type spaces" as well; see for instance [16, 17, 38, 55]. In contrast to the Fourier-analytic definition of [4], these more recent articles consider Barron spaces that essentially consist of all "infinitely wide" neural networks with a certain control over the network parameters. More formally, given an activation function ϕ (which is either the ReLU or a Heaviside function), the elements of the associated Barron space are all functions that can be written as

$$f(x) = \int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} a \cdot \phi\big(\langle w, x\rangle + c\big)\, d\mu(a, w, c)$$

for a probability measure μ satisfying

$$\int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} |a| \cdot \phi(|w| + |c|)\, d\mu(a, w, c) < \infty\,.$$

We will refer to these spaces as the *infinite-width Barron spaces*. We emphasize that in contrast to the Fourier-analytic Barron spaces, these infinite-width Barron spaces do depend on the choice of the activation function ϕ; they thus do *not* contain *all conceivable* "infinite-width" networks.

The relationship between the infinite-width and Fourier-analytic Barron spaces is not immediately obvious. Already in [3] it was shown that the Fourier-analytic Barron space is contained in the infinite-width Barron space associated to the Heaviside function. It is not clear, however, whether this also holds for the ReLU activation function. In Section 7, we will review approaches in the literature that address this embedding problem and prove that the classical Barron space is *not* contained in the infinite-width Barron space associated to the ReLU. In fact, we show in Proposition 7.4 the stronger result that if we consider a generalized Fourier-analytic Barron space that consists of all functions $f : \mathbb{R}^d \to \mathbb{R}$ such that their Fourier transform $\widehat{f}$ exists and satisfies $\left\| \xi \mapsto (1 + |\xi|)^\alpha \widehat{f}(\xi) \right\|_{L^1(\mathbb{R}^d)} < \infty$, then this space is contained in the infinite-width Barron space for the ReLU function only if $\alpha \geq 2$.

1.2 Previous work

In this section, we discuss previous research concerning the performance of neural networks for approximating and estimating classification functions, as well as existing results concerning dimension-independence in approximation and estimation problems. We distinguish between results of Barron-type, i.e., approaches following the ideas of [4], and other approaches. We first discuss extensions of [4] for shallow neural networks (i.e., networks with one hidden layer). Here, we in particular discuss the article [17], which is the only other work that we are aware of that studies *classification problems* (as opposed to regression problems) in the context of Barron-type functions. Secondly, we discuss extensions to deep neural networks and then review other related approaches not involving Barron-type spaces. Finally, we explain how our work complements the existing literature.

1.2.1 Previous work considering shallow neural networks

In [4], it was shown that shallow neural networks can break the curse of dimensionality for approximating functions f that have one finite Fourier moment; more precisely, one can achieve $\|f - \Phi_N\|_{L^2(\mu)} \lesssim N^{-1/2}$, where Φ_N is a shallow neural network with N neurons and μ is a probability measure on a ball in $\mathbb{R}^d$. The main insight in [4] is that functions with one finite Fourier moment belong to the closed convex hull of the set of half planes; that is, they admit an integral representation

$$f(x) = \int_{\mathbb{R}^d \times \mathbb{R}} \alpha(w, c) H(c + w^T x) \, d\nu(w, c) \tag{1.2}$$

where ν is a probability measure satisfying $\int_{\mathbb{R}^d \times \mathbb{R}} |\alpha(w, c)| \, d\nu(w, c) < \infty$ and $H = \mathbb{1}_{[0,\infty)}$ is the Heaviside function. The approximation rate of $N^{-1/2}$ is then a consequence of an approximate and probabilistic version of Caratheodory's theorem; see for instance [54, Theorem 0.0.2]. The paper [3] generalized these results from approximation in $L^2(\mu)$ to uniform approximation. Furthermore, in [5] these results are extended to obtain estimation bounds for the class of functions with one bounded Fourier moment. Essentially, using $n \sim N^2$ i.i.d. samples, a neural network with N neurons can be found that approximates f up to an L^2-error of the order of $N^{-1/2} \sim n^{-1/4}$.

Recently, several extensions of these original results by Barron to different spaces have been proposed. The Barron-type spaces introduced in [38, 18, 16, 55], are motivated by the integral representation (1.2). Specifically, given an activation function $\phi : \mathbb{R} \to \mathbb{R}$ and an

exponent $p \in [1, \infty]$, the *p-infinite-width Barron space* consists of all functions of the form

$$f(x) = \int_{\mathbb{S}^{d-1} \times [-1,1]} a(w,b)\, \phi(b + \langle x, w \rangle)\, d\pi(w,b), \tag{1.3}$$

for $x \in \mathbb{R}^d$, where π is a probability measure on $\mathbb{S}^{d-1} \times [-1,1]$ and $a \in L^p(\pi)$. It is shown that for certain values of p, the functions in the *p-infinite-width Barron space* can be efficiently estimated and approximated by neural networks with activation function ϕ, without dependency on the dimension. Let us add here that it was shown in [44] that functions of the form (1.3) arise naturally as the solutions of appropriately regularized learning problems.

We also mention the result [39], in which a slightly improved approximation rate is obtained for networks with the Heaviside function, albeit under a slightly stronger assumption on the functions to be approximated. Essentially, it is assumed in [39] that Equation (1.2) holds with a *bounded* function a instead of an integrable one. A further related result has been obtained in [51], where the above results were extended to more general activation functions and to approximation with respect to L^2-Sobolev norms. In addition, lower bounds on the approximation of Barron functions by shallow neural networks have recently been studied in [52].

The work in the present paper complements these results by clarifying the relation between the spaces of functions that can be represented as in Equations (1.2) and (1.3) and those that have one finite Fourier moment, as considered in the original papers by Barron; see Section 7 for more details.

1.2.2 Shallow neural networks for classification problems

The article [17] studies the problem of learning a classification function associated to two disjoint classes C_+, C_-. Instead of describing the accuracy of approximation and estimation with respect to the typical square loss, the paper focuses on the hinge loss and certain cross-entropy type losses. In this framework, a classification problem is considered solvable with respect to a hypothesis class if there exist elements in that hypothesis class that assume different signs on the two classes. It is shown in [17] that for general C_+, C_- such a problem is solvable by Barron regular functions if and only if the sets C_+, C_- have positive distance. Since for these functions the approximation and estimation behavior using shallow neural networks is well studied, as reviewed in the previous subsection, this observation yields approximation and estimation bounds by shallow neural networks for the classification problem.

In contrast to the setting considered in [17], in the present paper we analyze classification problems for which the different classes are *not* required to have a positive distance. Instead, we impose a regularity condition on the class boundaries and assume that the underlying probability measure is *tube-compatible*, meaning that it should not be too strongly concentrated at the class boundary.

1.2.3 Deep neural networks and the curse of dimensionality

It is natural to wonder whether deeper networks can improve on shallow neural networks regarding approximation and estimation problems. The fundamental property enabling "dimension-free" approximation by *shallow* neural networks is that the function to be approximated should belong to the closed convex hull of the set of simple neurons. The corresponding property for deep networks has been identified to be a certain summability property of the weights of approximating neural networks. In [6, Equation 1], this summability property is called the *variation* of the neural network. To avoid ambiguities with the (total) variation of a measure or function, we will refer to this notion as the *weight variation*. More precisely, the weight variation is the ℓ^1 norm of the entries of the product of the weight matrices of the

neural networks[1]. In [6, Theorem 1] it is shown that if arbitrarily large neural networks are of bounded weight variation, then these neural networks can be well approximated by smaller neural networks. Here the size of the neural networks is measured via the encoding complexity of the weights. Moreover, the reduction in size is independent of the dimension. The weight variation also serves as a motivation for the so-called *path norm* that is fundamental to the definition of generalized Barron spaces associated to compositional function representation in [55]. This path norm can be understood as the continuous counterpart of the weight variation. Correspondingly, the elements of the generalized Barron spaces in [55] are those functions that can be obtained as limits of deep neural networks with bounded variation, for increasing width. We also mention [37], which studies approximation of functions that are compositions of n classical Barron functions and shows that these can then be efficiently approximated by neural networks with $n+1$ layers.

In a somewhat similar vein, we show that if one is interested not in approximating Barron-regular functions themselves, but rather classification functions for which the class boundaries are Barron-regular, then this can be done efficiently with (somewhat) deep ReLU neural networks, namely using networks with 3 hidden layers; see Theorem 3.7.

1.2.4 Non-Barron-type results on curse of dimensionality

Functions of Barron-type are not the only functions that can be approximated by deep neural networks without the curse of dimensionality. Other function classes that allow for approximation with only minor (in particular sub-exponential) dependencies on the dimension include the following: functions that have a graph-like structure and are compositions of low dimensional functions, [46], [45, Section 5], [49, 14, 42]; bandlimited functions [41]; and also solutions of some classes of high-dimensional PDEs [25, 27, 30, 9, 12, 34, 24, 20, 48] and SDEs [10, 47], under the assumption that the right-hand side of the equation is itself well-approximated (i.e., without suffering from the curse of dimensionality) by neural networks.

In the present paper, we show that the classification functions with Barron regular decision boundaries also belong to this list of well-approximable functions.

1.2.5 Deep neural networks for classification problems

The approximation and estimation of classification functions of the form $\sum_{k=1}^{K} f_k \mathbb{1}_{\Omega_k}$, where each $\Omega_k \subset \mathbb{R}^d$ is an open set such that $\partial\Omega_k$ is piecewise smooth and $f_k : \mathbb{R}^d \to \mathbb{R}$ are smooth, is studied in [28, 29, 45]. In these works, it is shown that the achievable approximation and estimation rates are primarily determined by the smoothness of the boundaries $\partial\Omega_k$, in the sense that, given sufficient regularity of the f_k, smoother class boundaries yield better approximation and estimation rates. The general strategy of the approximation theoretical aspects of these works is closely related to the approach taken in this article. Indeed, the approximation of classification functions is reduced to that of horizon functions $\mathbb{1}_{x_1 \leq f(x_2,\dots,x_d)}$ where f is a $d-1$-dimensional smooth function. In addition, the articles [28, 29] establish estimation bounds by invoking classical bounds on the covering numbers of the involved neural network spaces to bound the generalization error of empirical risk minimization.

1.2.6 Delineation of our work

In the present article, we discuss a concrete set of practically relevant functions, namely those arising in classification tasks where the interfaces between classes are sufficiently regular, which formally means that they are locally described by Barron-type functions. As indicated earlier, these results are based on a combination of two ideas: First, a classical result of Barron

[1]Here, it should be noted that all weights are assumed to be non-negative in [6], which is accomplished there without loss of generality by a slight modification of the activation function.

showing uniform and dimension-independent approximation of Barron-type functions [3] and, second, a strategy to emulate functions with regular jump curves by neural networks, originally introduced in [45].

The results are neither a direct consequence of the study of (generalized) Barron spaces nor can they be derived directly from the results of [45]. Indeed, the functions that we discuss (classification functions with Barron-regular boundary) do not have a representation by neural networks with bounded weights or bounded variation of the weights. In fact, it can be shown (see [55, Theorem 2.7]) that functions in the (generalized) Barron spaces are always Lipschitz continuous, which is not satisfied for the classification functions that we consider. The key difference between our approach and alternative studies of Barron spaces is that in those works the boundedness of the (sum of the) network weights or a related property such as a bounded weight variation plays a central role. In contrast, we allow a moderate weight growth that is essentially inversely proportional to the approximation error. Besides, in contrast to [17] we study classification problems for which the different classes do *not* have positive distance to each other. Furthermore, the required regularity of the class boundaries for our results is explicitly stated, e.g. in terms of a finite Fourier moment; this is in contrast to the more implicit integral representation property required for the infinite-width Barron spaces considered in [17].

Finally, in contrast to [28, 29, 45], the results in the present paper do not suffer from the curse of dimensionality.

1.3 Structure of the paper

After introducing general and neural network related notation in Subsections 1.4 and 1.5, we start in Section 2 by formally defining the Fourier-analytic Barron class, and proving that such functions can be uniformly approximated with error $\mathcal{O}(N^{-1/2})$ using shallow ReLU networks with $\mathcal{O}(N)$ neurons and controlled weights. We reprove this result since the argument in [3] for handling general sigmoidal activation functions contains a technical inaccuracy.

In Section 3, we give the precise definition of sets with boundary in the Barron class, and we show that indicator functions of such sets can be well approximated by ReLU neural networks. The complementing lower bounds and estimation bounds are derived in Sections 4 and 5. For the approximation and estimation results, we always assume that the measure under consideration is tube compatible; Section 6 shows that this is unavoidable. Finally, in Section 7, we discuss the relation between the Fourier-analytic Barron space that we consider and the alternative Barron spaces considered in the literature.

Several mainly technical results are deferred to the appendices.

1.4 General notation

We will use the following notation: For $n \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$, we write $\underline{n} := \{1, 2, \dots, n\}$; in particular, $\underline{0} = \varnothing$. For an arbitrary set M, we write $|M| = \#M \in \mathbb{N}_0 \cup \{\infty\}$ for the number of elements of M.

Given $a \in \mathbb{R}^d$, we denote the entries of a by $a_1, \dots, a_d \in \mathbb{R}$. For $a, b \in \mathbb{R}^d$ we write $a \leq b$ if and only if $a_i \leq b_i$ for all $i \in \underline{d}$. In this case, we define $[a, b] := \prod_{i=1}^d [a_i, b_i]$. For $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ with $d > 1$ and $i \in \underline{d}$, we set $x^{(i)} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \in \mathbb{R}^{d-1}$.

The standard scalar product of $x, y \in \mathbb{R}^d$ will be denoted by $\langle x, y \rangle = \sum_{i=1}^d x_i \, y_i$, and the Euclidean norm of x is written as $|x| := \sqrt{\langle x, x \rangle}$. For a continuous function f defined on a set $Q \subset \mathbb{R}^d$, we define $\|f\|_{\sup} := \sup_{x \in Q} |f(x)|$. For a set X and two functions $f, g \colon X \to \mathbb{R}^+$, we write $f(x) \lesssim g(x)$ if $f(x) \leq C g(x)$ for a constant $C > 0$ and all $x \in X$. This constant is referred to as the implicit constant of the estimate.

Finally, given a class $\mathcal{F}$ of $\{0, 1\}$-valued (or $\{\pm 1\}$-valued) functions, we denote the VC-

dimension of $\mathcal{F}$ by $\mathrm{VC}(\mathcal{F}) \in \mathbb{N}_0 \cup \{\infty\}$. We refer to [50, Chapter 6] for the definition of the VC dimension.

1.5 Neural network notation

In this subsection, we briefly introduce our notation regarding neural networks. To avoid ambiguities, we define neural networks in a way that allows a precise counting of the number of neurons and layers. This is done by differentiating between a neural network as a set of weights and the associated realization which represents the function that is described through these weights.

Definition 1.2. Let $d, L \in \mathbb{N}$. A *neural network (NN)* Φ *with input dimension* d *and* L *layers* is a sequence of matrix-vector tuples

$$\Phi = \big((A_1, b_1), (A_2, b_2), \dots, (A_L, b_L)\big),$$

where, for $N_0 = d$ and certain $N_1, \dots, N_L \in \mathbb{N}$, each A_ℓ is an $N_\ell \times N_{\ell-1}$ matrix, and $b_\ell \in \mathbb{R}^{N_\ell}$.

For a NN Φ and an activation function $\phi : \mathbb{R} \to \mathbb{R}$, we define the associated *realization of the NN* Φ as

$$R_\phi \Phi : \quad \mathbb{R}^d \to \mathbb{R}^{N_L}, \quad x \mapsto x_L = R_\phi \Phi(x),$$

where the output $x_L \in \mathbb{R}^{N_L}$ results from the scheme

$$\begin{aligned} x_0 &:= x \in \mathbb{R}^d = \mathbb{R}^{N_0}, \\ x_\ell &:= \phi\left(A_\ell \, x_{\ell-1} + b_\ell\right) \in \mathbb{R}^{N_\ell} \qquad \text{for } \ell = 1, \dots, L-1, \\ x_L &:= A_L \, x_{L-1} + b_L \in \mathbb{R}^{N_L}. \end{aligned}$$

Here ϕ is understood to act component-wise. We call $N(\Phi) := d + \sum_{j=1}^{L} N_j$ the *number of neurons* of the NN Φ, $L = L(\Phi)$ the *number of layers*, and $W(\Phi) := \sum_{j=1}^{L} (\|A_j\|_0 + \|b_j\|_0)$ is called the *number of weights* of Φ. Here, $\|A\|_0$ and $\|b\|_0$ denote the number of non-zero entries of the matrix A or the vector b. Moreover, we refer to N_L as the *output dimension* of Φ. The activation function $\varrho : \mathbb{R} \to \mathbb{R}, x \mapsto \max\{0, x\}$ is called the *ReLU*. We call $R_\varrho \Phi$ a *ReLU neural network*. Finally, the vector $(d, N_1, N_2, \dots, N_L) \in \mathbb{N}^{L+1}$ is called the *architecture* of Φ.

Remark 1.3. With notation as above, the number of *hidden* layers of Φ is $L - 1$. A special type of neural networks are those with one hidden layer, i.e., $L = 2$; these are called *shallow neural networks*. Realizations of such networks have the form

$$\mathbb{R}^d \ni x \mapsto e + \sum_{i=1}^{N} a_i \, \phi(\langle c_i, x \rangle + b_i),$$

where $N \in \mathbb{N}$, $a_i, b_i, e \in \mathbb{R}$ and $c_i \in \mathbb{R}^d$ for $i = 1, \dots, N$.

One important property of neural networks is that one can construct complicated neural networks by combining simpler ones. The following remark collects several standard operations that were analyzed in [45].

Remark 1.4. Let Φ_1, Φ_2 be two neural networks with input dimensions $d_1, d_2 \in \mathbb{N}$, L_1, L_2 layers and architectures $(d_1, N_1, N_2, \dots, N_{L_1}) \in \mathbb{N}^{L_1+1}$ and $(d_2, M_1, M_2, \dots, M_{L_2}) \in \mathbb{N}^{L_2+1}$, respectively. Furthermore, let $\phi : \mathbb{R} \to \mathbb{R}$.

- If $d_2 = N_{L_1}$, then there exists a neural network Φ_3 such that $R_\phi\Phi_3 = R_\phi\Phi_2 \circ R_\phi\Phi_1$. Moreover, Φ_3 can be chosen to have architecture

$$(d_1,\ N_1,\ N_2,\ \ldots,\ N_{L_1-1},\ M_1,\ M_2,\ \ldots,\ M_{L_2}) \in \mathbb{N}^{L_1+L_2}$$

  and to satisfy $L(\Phi_3) = L_1 + L_2 - 1$ and $W(\Phi_3) \leq W(\Phi_1) + W(\Phi_2) + N_{L_1-1}M_1$.

- If $L_1 = L_2$, $d_1 = d_2$ and $N_{L_1} = M_{L_1}$, then, given arbitrary $a, b \in \mathbb{R}$ there exists a neural network Φ_4 such that $R_\phi\Phi_4 = aR_\phi\Phi_1 + bR_\phi\Phi_2$. Moreover, Φ_4 can be chosen to have architecture

$$(d_1,\ N_1 + M_1,\ N_2 + M_2,\ \ldots,\ N_{L_1-1} + M_{L_1-1},\ N_{L_1})$$

  and to satisfy $L(\Phi_4) = L_1$ and $W(\Phi_4) \leq W(\Phi_1) + W(\Phi_2)$.

2 Uniform approximation of Barron-type functions using ReLU networks

In this section, we formalize the notion of the (Fourier-analytic) Barron space that we will use in the sequel. We then prove that functions in the Barron class can be approximated up to error $\mathcal{O}(N^{-1/2})$ using shallow ReLU neural networks with N neurons. For neural networks with the Heaviside activation function, this result is due to Barron [3]. Furthermore, it is claimed in [3] that the result extends to neural networks with sigmoidal activation functions, which would then also imply the same property for the ReLU activation function ϱ, since $\phi(x) = \varrho(x) - \varrho(x-1)$ is sigmoidal. However, regarding the extension to sigmoidal activation functions there seems to be a gap in the proof presented in [3]. Namely, it is argued in the bottom left column on Page 3 of [3] that if f is uniformly continuous and $\|f - f_T\|_{\sup} \lesssim T^{-1/2}$ where f_T is of the form $f_T(x) = c_0 + \sum_{k=1}^T c_k \mathbb{1}_{(0,\infty)}(\langle a_k, x\rangle + b_k)$ with $c_k, b_k \in \mathbb{R}$ and $a_k \in \mathbb{R}^d$, then one can also achieve $\|f - g_T\|_{\sup} \lesssim T^{-1/2}$ for $g_T(x) = C_0 + \sum_{k=1}^T C_k\, \phi(\langle A_k, x\rangle + B_k)$, where ϕ is measurable and *sigmoidal*, meaning ϕ is bounded with $\lim_{x\to\infty}\phi(x) = 1$ and $\lim_{x\to-\infty}\phi(x) = 0$. As we could not verify this claim, we provide an alternative proof for the case of the ReLU activation function, based on the main ideas in [3]. In addition, our more careful proof shows that one can choose the weights of the neural network to be uniformly bounded, independent of the desired approximation accuracy.

We first formalize the notion of *Barron class functions*, essentially as introduced in [4, 3].

Definition 2.1. Let $\varnothing \neq X \subset \mathbb{R}^d$ be bounded. A function $f : X \to \mathbb{R}$ is said to be of *Barron class with constant* $C > 0$, if there are $x_0 \in X$, $c \in [-C, C]$, and a measurable function $F : \mathbb{R}^d \to \mathbb{C}$ satisfying

$$\int_{\mathbb{R}^d} |\xi|_{X,x_0} \cdot |F(\xi)|\, d\xi \leq C \quad \text{and} \quad f(x) = c + \int_{\mathbb{R}^d} \left(e^{i\langle x,\xi\rangle} - e^{i\langle x_0,\xi\rangle}\right) \cdot F(\xi)\, d\xi \qquad \forall\, x \in X, \tag{2.1}$$

where we used the notation $|\xi|_{X,x_0} := \sup_{x\in X} |\langle \xi, x - x_0\rangle|$. We write $\mathcal{B}_C(X, x_0)$ for the class of all such functions.

Remark. The precise choice of the "base point" $x_0 \in X$ is immaterial, in the sense that it at most changes the resulting norm by a factor of 2. Indeed, let $x_0, x_1 \in X$ and assume that f satisfies (2.1) with $|c| \leq C$. Then we see for arbitrary $\xi \in \mathbb{R}^d$ and $x \in X$ that

$$|\langle \xi, x - x_1\rangle| \leq |\langle \xi, x - x_0\rangle| + |\langle \xi, x_0 - x_1\rangle| = |\langle \xi, x - x_0\rangle| + |\langle \xi, x_1 - x_0\rangle| \leq 2\,|\xi|_{X,x_0},$$

meaning $|\xi|_{X,x_1} \leq 2\,|\xi|_{X,x_0}$ and hence $\int_{\mathbb{R}^d} |\xi|_{X,x_1} \cdot |F(\xi)|\, d\xi \leq 2C$. Furthermore, setting $c' := c + \int_{\mathbb{R}^d} \left(e^{i\langle x_1,\xi\rangle} - e^{i\langle x_0,\xi\rangle}\right) F(\xi)\, d\xi$, we have $f(x) = c' + \int_{\mathbb{R}^d} \left(e^{i\langle x,\xi\rangle} - e^{i\langle x_1,\xi\rangle}\right) F(\xi)\, d\xi$ and

$|e^{i\langle x_1,\xi\rangle} - e^{i\langle x_0,\xi\rangle}| \leq |\langle x_1 - x_0, \xi\rangle| \leq |\xi|_{X,x_0}$, which implies $|c'| \leq C + \int_{\mathbb{R}^d} |\xi|_{X,x_0} \, |F(\xi)| \, d\xi \leq 2C$. Overall, this shows that $f \in \mathcal{B}_{2C}(X, x_1)$ and hence $\mathcal{B}_C(X, x_0) \subset \mathcal{B}_{2C}(X, x_1)$.

Based on this, it is straightforward to see

$$\forall\, \varnothing \neq Y \subset X \text{ and } x_0 \in X, y_0 \in Y : \quad \mathcal{B}_C(X, x_0) \subset \mathcal{B}_{2C}(Y, y_0).$$

For the sake of clarity, note that if $\{x_0\} \subset Y \subset X$ and $f \in \mathcal{B}_C(X, x_0)$, then clearly $f|_Y \in \mathcal{B}_C(Y, x_0)$ as the conditions in (2.1) are already satisfied. Therefore the inclusion $\mathcal{B}_C(X, x_0) \subset \mathcal{B}_{2C}(Y, y_0)$ from above is to be understood, by slight abuse of notation, in the sense of function restrictions.

The following result shows that functions from the Barron class can be uniformly approximated with error $\mathcal{O}(N^{-1/2})$ using shallow ReLU neural networks with $\mathcal{O}(N)$ neurons. It also shows that the weights of the approximating network can be chosen to be bounded in a suitable way. We emphasize that the result is *not* covered by [16, Theorem 12], since the Fourier-analytic Barron space that we use here is *not* contained in the Barron space considered in [16]; see Proposition 7.4.

Proposition 2.2. *There is a universal constant $\kappa > 0$ with the following property: For any bounded set $X \subset \mathbb{R}^d$ with nonempty interior, for any $C > 0$, $x_0 \in X$ and $f \in \mathcal{B}_C(X, x_0)$, and any $N \in \mathbb{N}$, there is a shallow neural network Φ with $8N$ neurons in the hidden layer such that*

$$\|f - R_\varrho \Phi\|_{\sup} \leq \kappa \sqrt{d} \cdot C \cdot N^{-1/2}.$$

Furthermore, one can choose all weights and biases of Φ to be bounded by

$$\big(5 + \vartheta(X, x_0)\big) \cdot \big(1 + \|x_0\|_{\ell^1}\big) \cdot \sqrt{C}, \quad \text{where} \quad \vartheta(X, x_0) := \sup_{\xi \in \mathbb{R}^d \setminus \{0\}} \Big(\|\xi\|_{\ell^\infty} / |\xi|_{X,x_0} \Big).$$

Remark 2.3. The quantity $\vartheta(X, x_0)$ roughly speaking measures how big of a rectangle the set X contains. More precisely, assume that $X \supset [a, b]$ where $b_i - a_i \geq \varepsilon > 0$ for all $i \in \underline{d}$. Then we see with the standard basis $(e_1, \dots, e_d)$ of $\mathbb{R}^d$ that

$$\varepsilon \, |\xi_i| = \big|\langle \xi, a + \varepsilon \, e_i - x_0\rangle - \langle \xi, a - x_0\rangle\big| \leq |\langle \xi, a + \varepsilon e_i - x_0\rangle| + |\langle \xi, a - x_0\rangle| \leq 2 \sup_{x \in X} |\langle \xi, x - x_0\rangle|.$$

Since this holds for all $i \in \underline{d}$, we see $|\xi|_{X,x_0} \geq \frac{\varepsilon}{2} \|\xi\|_{\ell^\infty}$ and hence $\vartheta(X, x_0) \leq \frac{2}{\varepsilon}$.

Note that since X has nonempty interior, we can always find a sufficiently small non-degenerate rectangle in X; therefore, $|\xi|_{X,x_0} \gtrsim \|\xi\|_{\ell^\infty}$ for all $\xi \in \mathbb{R}^d$.

Proof. It is enough to prove the claim for the case $C = 1$. Indeed, for $f \in \mathcal{B}_C(X, x_0)$, we have $\widetilde{f} := f/C \in \mathcal{B}_1(X, x_0)$. Applying the claim to $\widetilde{f}$, we thus get $\|\widetilde{f} - \widetilde{g}\|_{\sup} \leq \kappa\sqrt{d} \cdot N^{-1/2}$, where $\widetilde{g}(x) = \sum_{i=1}^{8N} a_i \, \varrho(b_i + \langle w_i, x\rangle)$ with $\|w_i\|_{\ell^\infty}, |a_i|, |b_i| \leq (5 + \vartheta(X, x_0)) \cdot (1 + \|x_0\|_{\ell^1})$. Hence, defining $g(x) = \sum_{i=1}^{8N} \sqrt{C} a_i \, \varrho(\sqrt{C} b_i + \langle \sqrt{C} w_i, x\rangle)$, we have $g(x) = C \cdot \widetilde{g}(x)$, which easily yields the claim for f. We will thus assume $C = 1$ in what follows. The actual proof is divided into three steps.

Step 1 *(Writing f as an expectation of indicators of half-spaces):* Let $c \in [-C, C]$ and $F : \mathbb{R}^d \to \mathbb{C}$ such that Equation (2.1) is satisfied. The case where $F = 0$ almost everywhere is easy to handle; we thus assume that $F \neq 0$ on a set of positive measure.

Set $X_0 := \{x - x_0 \colon x \in X\}$, and define $f_0 : X_0 \to \mathbb{R}$ by $f_0(x) := f(x + x_0) - c$ and $F_0 : \mathbb{R}^d \to \mathbb{C}, \xi \mapsto e^{i\langle x_0,\xi\rangle} \, F(\xi)$. With this notation, we have $f_0(x) = \int_{\mathbb{R}^d} (e^{i\langle x,\xi\rangle} - 1) \cdot F_0(\xi) \, d\xi$

and $\int_{\mathbb{R}^d} |\xi|_{X_0} \cdot |F_0(\xi)|\, d\xi \leq C$, where $|\xi|_{X_0} := \sup_{x \in X_0} |\langle x, \xi\rangle| = |\xi|_{X, x_0}$. Thus, (the proof of) [3, Theorem 2] shows for all $x \in X_0$ that

$$f_0(x) = v \cdot \int_{\mathbb{R}^d} \int_0^1 \Big(\mathbb{1}_{(0,\infty)}\big(-\langle \xi/|\xi|_{X_0},\ x\rangle - t\big) - \mathbb{1}_{(0,\infty)}\big(\langle \xi/|\xi|_{X_0},\ x\rangle - t\big)\Big) \cdot s(\xi,t) \cdot p(\xi,t)\, dt\, d\xi,$$

where, using the polar decomposition $F_0(\xi) = |F_0(\xi)| \cdot e^{i\,\theta_\xi}$, the function $s : \mathbb{R}^d \times [0,1] \to \{\pm 1\}$ is given by $s(\xi,t) = \operatorname{sign}\big(\sin(t\,|\xi|_{X_0} + \theta_\xi)\big)$, while $p : \mathbb{R}^d \times [0,1] \to [0,\infty)$ is defined as $p(\xi,t) = \frac{1}{v} \cdot |\xi|_{X_0} \cdot \big|\sin(t\,|\xi|_{X_0} + \theta_\xi)\big| \cdot |F_0(\xi)|$. Finally,

$$v = \int_{\mathbb{R}^d} \int_0^1 |\xi|_{X_0} \cdot |\sin(t|\xi|_{X_0} + \theta_\xi)| \cdot |F_0(\xi)|\, dt\, d\xi \leq C$$

is chosen such that p is a probability density function. It is easy to see $v > 0$ since $F_0 \neq 0$ on a set of positive measure.

For brevity, define $\Omega := (\mathbb{R}^d \setminus \{0\}) \times [0,1]$. Furthermore, set $\xi^* := \xi/|\xi|_{X_0}$ for $\xi \in \mathbb{R}^d \setminus \{0\}$ (where we note that $|\xi|_{X_0} > 0$ since X_0 has nonempty interior), and for $x \in X_0$ define

$$\Gamma_x : \quad \Omega \to [-1,1], \quad (\xi,t) \mapsto \mathbb{1}_{(0,\infty)}(-\langle \xi^*, x\rangle - t) - \mathbb{1}_{(0,\infty)}(\langle \xi^*, x\rangle - t).$$

Finally, let us set $V_\pm := \int_{\mathbb{R}^d} \int_0^1 \mathbb{1}_{s(\xi,t)=\pm 1} \cdot p(\xi,t)\, dt\, d\xi$, and define probability measures $\mu_\pm$ on Ω via

$$d\mu_\pm := \frac{1}{V_\pm} \cdot \mathbb{1}_{s(\xi,t)=\pm 1} \cdot p(\xi,t)\, dt\, d\xi.$$

Note that $V_+, V_- \geq 0$ and $V_+ + V_- = 1$. Also note that strictly speaking $\mu_\pm$ is only well-defined in case of $V_\pm > 0$. In case of $V_\pm = 0$, one can simply drop the respective term in what follows; we leave the straightforward modifications to the reader.

Given all these notations, we see that $f_0 = v \cdot (V_+ \cdot f_+ - V_- \cdot f_-)$, where

$$f_\pm : X_0 \to \mathbb{R} \quad \text{is defined by} \quad f_\pm(x) := \int_\Omega \Gamma_x(\xi,t)\, d\mu_\pm(\xi,t).$$

It is enough to show $\|f_\pm - R_\varrho \Phi_\pm\|_{\sup} \leq N^{-1/2} \cdot \big(\frac{C}{v\,V_\pm} + \kappa_0 \sqrt{d}\big)$ for a shallow neural network $\Phi_\pm$ with $4\,N$ neurons in the hidden layer and with all weights and biases bounded by $4 + \vartheta(X, x_0)$. Indeed, once this is shown, it is easy to see that there exists a shallow network Φ with $8N$ neurons in the hidden layer satisfying

$$R_\varrho \Phi(x) = c + v\,V_+ \cdot R_\varrho \Phi_+(x - x_0) - v\,V_- \cdot R_\varrho \Phi_-(x - x_0).$$

Because of $f(x) = c + f_0(x - x_0) = c + v\,V_+ \cdot f_+(x - x_0) - v\,V_- \cdot f_-(x - x_0)$ and $0 < v \leq C$, this yields

$$\begin{aligned} \|f - R_\varrho \Phi\|_{\sup} &\leq N^{-1/2} \cdot \Big(v\,V_+ \cdot \big(\tfrac{C}{v\,V_+} + \kappa_0 \sqrt{d}\big) + v\,V_- \cdot \big(\tfrac{C}{v\,V_-} + \kappa_0 \sqrt{d}\big)\Big) \\ &= N^{-1/2} \cdot \big(2C + v\kappa\,\sqrt{d}\big) \leq \big(2 + \kappa_0\,\sqrt{d}\big) \cdot C \cdot N^{-1/2} \leq \kappa\,\sqrt{d} \cdot C \cdot N^{-1/2} \end{aligned}$$

for a suitable absolute constant $\kappa > 0$. Again, since $0 < v \leq C$ and $c \in [-C, C]$ as well as $0 \leq V_\pm \leq 1$, and since we assume $C = 1$, it is easy to see that Φ can be chosen in such a way that all weights of Φ are bounded by $\big(4 + \vartheta(X, x_0)\big) \cdot (1 + \|x_0\|_{\ell^1})$. Here, we use that if $\|w\|_{\ell^\infty}, |b| \leq 4 + \vartheta(X, x_0)$, then $\varrho(\langle w, x - x_0\rangle + b) = \varrho(\langle w, x\rangle + b - \langle w, x_0\rangle)$, where

$$\big|b - \langle w, x_0\rangle\big| \leq |b| + \big(4 + \vartheta(X, x_0)\big)\, \|x_0\|_{\ell^1} \leq \big(4 + \vartheta(X, x_0)\big) \cdot \big(1 + \|x_0\|_{\ell^1}\big).$$

Step 2 *(Approximating f by an expectation of ReLU networks):* For $\varepsilon > 0$, define

$$H_\varepsilon : \quad \mathbb{R} \to [0,1], \quad x \mapsto \frac{1}{\varepsilon}\big(\varrho(x) - \varrho(x-\varepsilon)\big),$$

noting that $H_\varepsilon(x) = \mathbb{1}_{(0,\infty)}(x)$ for all $x \in \mathbb{R} \setminus (0,\varepsilon)$. Next, for $\varepsilon > 0$ and $x \in X_0$, set

$$N_{\varepsilon,x} : \quad \Omega \to [-1,1], \quad (\xi,t) \mapsto H_\varepsilon(-\langle \xi^*, x\rangle - t) - H_\varepsilon(\langle \xi^*, x\rangle - t).$$

Setting $J_{\xi,x}^{(\varepsilon)} := [-\langle \xi^*, x\rangle - \varepsilon,\ -\langle \xi^*, x\rangle] \cup [\langle \xi^*, x\rangle - \varepsilon,\ \langle \xi^*, x\rangle]$, we have $\Gamma_x(\xi,t) = N_{\varepsilon,x}(\xi,t)$ for all $(\xi,t) \in \Omega$ with $t \notin J_{\xi,x}^{(\varepsilon)}$. Thus, using the bound $0 \leq p(\xi,t) \leq \frac{1}{v}|\xi|_{X_0}\,|F(\xi)|$ and the definitions of $f_\pm$ and $\mu_\pm$, we see for all $x \in X_0$ that

$$\begin{aligned}\Big| f_\pm(x) - \int_\Omega N_{\varepsilon,x}(\xi,t)\,d\mu_\pm(\xi,t)\Big| &\leq \int_{\mathbb{R}^d\setminus\{0\}} \int_0^1 2 \cdot \mathbb{1}_{J_{\xi,x}^{(\varepsilon)}}(t) \cdot \frac{1}{V_\pm}\, p(\xi,t)\,dt\,d\xi \\ &\leq \frac{4\varepsilon}{v\,V_\pm} \int_{\mathbb{R}^d} |\xi|_{X_0} \cdot |F(\xi)|\,d\xi \leq \frac{4\varepsilon C}{v\,V_\pm}.\end{aligned}$$

We now choose $\varepsilon := \frac{1}{4}N^{-1/2}$ and define $f_{\pm,\varepsilon} : X_0 \to \mathbb{R}, x \mapsto \int_\Omega N_{\varepsilon,x}(\xi,t)\,d\mu_\pm(\xi,t)$. Then the preceding estimate shows that $\|f_\pm - f_{\pm,\varepsilon}\|_{\sup} \leq N^{-1/2} \cdot \frac{C}{v\,V_\pm}$.

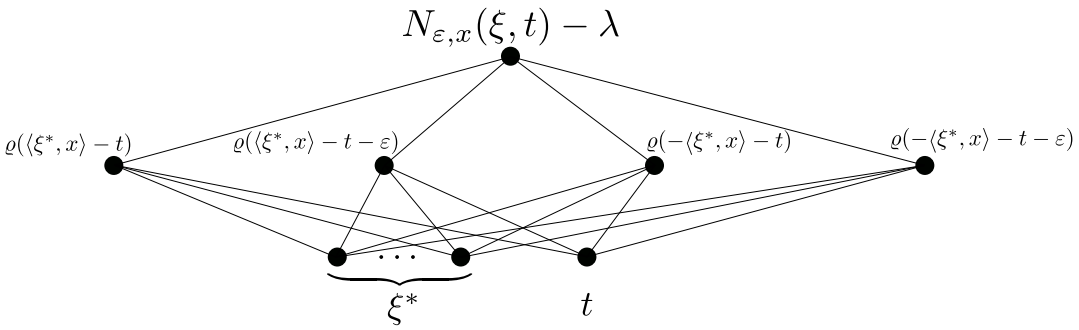

Figure 1: Representation of the function $(\xi,t) \mapsto N_{\varepsilon,x}(\xi,t) - \lambda$ as a ReLU network with $L = 2$ layers, $W = 4d + 16$ weights, and $U = 5$ computation units (using the notation of [7]).

Step 3 *(Using bounds for empirical processes to complete the proof):* Denote by $\mathcal{G}_0$ the set of all functions $g : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ that are implemented by ReLU neural networks with the architecture shown in Figure 1 (that is, fully connected with one hidden layer containing four neurons). Then the VC dimension bound for neural networks shown in [7, Theorem 6] implies that there is an absolute constant $\kappa_1 \in \mathbb{N}$ such that

$$\mathrm{VC}(\{\mathbb{1}_{g>0} \colon g \in \mathcal{G}_0\}) \leq \kappa_1\, d.$$

Moreover, using the map $\Theta : \Omega \to \mathbb{R}^d \times [0,1], (\xi,t) \mapsto (\xi^*, t) = (\xi/|\xi|_{X_0}, t)$, the construction in Figure 1 shows for arbitrary $\lambda \in \mathbb{R}$ that

$$\{\mathbb{1}_{N_{\varepsilon,x}>\lambda} \colon x \in X_0\} \subset \{\mathbb{1}_{g\circ\Theta>0} \colon g \in \mathcal{G}_0\}.$$

Directly from the definition of the VC dimension, we see that composing a class of functions with a fixed map (in this case, Θ) can not increase the VC dimension, so that we get $\mathrm{VC}(\{\mathbb{1}_{N_{\varepsilon,x}>\lambda} \colon x \in X_0\}) \leq \kappa_1\, d$ for all $\varepsilon > 0$ and $\lambda \in \mathbb{R}$.

Now, using the bound in Proposition A.1 and recalling that $\mathbb{E}_{(\xi,t)\sim\mu_\pm}[N_{\varepsilon,x}(\xi,t)] = f_{\pm,\varepsilon}(x)$, we see that if we choose $(\xi_1,t_1),\dots,(\xi_N,t_N) \overset{\text{i.i.d.}}{\sim} \mu_\pm$, then there is a universal constant $\kappa_2 > 0$ satisfying for all $N \in \mathbb{N}$ that[2]

$$\mathbb{E}\Big[\sup_{x\in X_0}\Big|f_{\pm,\varepsilon}(x) - \frac{1}{N}\sum_{i=1}^{N} N_{\varepsilon,x}(\xi_i,t_i)\Big|\Big] \leq \kappa_2 \cdot \sqrt{\frac{\kappa_1 d}{N}}. \tag{2.2}$$

In particular, there is one specific realization $\big((\xi_1,t_1),\dots,(\xi_N,t_N)\big) \in \Omega^N$ such that

$$\sup_{x\in X_0}\Big|f_{\pm,\varepsilon}(x) - \frac{1}{N}\sum_{i=1}^{N} N_{\varepsilon,x}(\xi_i,t_i)\Big| \leq \kappa\,\sqrt{d}\,N^{-1/2}.$$

Clearly, $g_{\pm,\varepsilon} : \mathbb{R}^d \to \mathbb{R}, x \mapsto \frac{1}{N}\sum_{i=1}^{N} N_{\varepsilon,x}(\xi_i,t_i)$ is implemented by a shallow ReLU network with $4N$ neurons in the hidden layer, as follows from

$$\frac{1}{N}N_{\varepsilon,x}(\xi_i,t_i) = \frac{\varepsilon^{-1}}{N}\cdot\Big(\varrho\big(-\langle\xi_i^*,x\rangle - t_i\big) - \varrho\big(-\langle\xi_i^*,x\rangle - t_i - \varepsilon\big) - \varrho\big(\langle\xi_i^*,x\rangle - t_i\big) + \varrho\big(\langle\xi_i^*,x\rangle - t_i - \varepsilon\big)\Big).$$

Now, note by definition of $\vartheta(X,x_0)$ and $\xi^* = \xi/|\xi|_{X_0}$ that $\|\xi_i^*\|_{\ell^\infty} \leq \vartheta(X,x_0)$. Furthermore, $|t_i| \leq 1$. Finally, by choice of $\varepsilon = \frac{1}{4}N^{-1/2}$, we see $\varepsilon^{-1}/N = 4N^{-1/2} \leq 4$. Overall, we thus see that $g_{\pm,\varepsilon} = R_\varrho\Phi_\pm$ where the shallow neural network $\Phi_\pm$ has $4N$ neurons in the hidden layer and all weights and biases bounded by $4 + \vartheta(X,x_0)$. □

3 Approximation of sets with Barron class boundary

In this section, we show that indicator functions of sets with Barron class boundary are well approximated by ReLU neural networks. Essentially the only property of Barron class functions that we will need is that they can be uniformly approximated up to error $\mathcal{O}(N^{-1/2})$ by shallow ReLU networks with N neurons and suitably bounded weight. Thus, to allow for a slightly more general result, we introduce a "Barron approximation space" containing all such functions.

Definition 3.1. Let $d \in \mathbb{N}$ and let $X \subset \mathbb{R}^d$ be bounded with nonempty interior. For $C > 0$, we define the *Barron approximation set* $\mathcal{BA}_C(X)$ as the set of all functions $f : X \to \mathbb{R}$ such that for every $N \in \mathbb{N}$ there is a shallow neural network Φ with N neurons in the hidden layer such that

$$\|f - R_\varrho\Phi\|_{\sup} \leq \sqrt{d}\cdot C \cdot N^{-1/2}$$

and such that all weights (and biases) of Φ are bounded in absolute value by

$$\sqrt{C}\cdot\Big(5 + \inf_{x_0\in X}\big[\,\|x_0\|_{\ell^1} + \vartheta(X,x_0)\,\big]\Big), \quad \text{where} \quad \vartheta(X,x_0) := \sup_{\xi\in\mathbb{R}^d\setminus\{0\}}\big(\|\xi\|_{\ell^\infty}/|\xi|_{X,x_0}\big).$$

The set $\mathcal{BA}(X) = \bigcup_{C>0}\mathcal{BA}_C(X)$ is called the *Barron approximation space*.

[2] Strictly speaking, Proposition A.1 yields a bound for

$$\sup_{X_{00}\subset X_0 \text{ finite}} \mathbb{E}\Big[\sup_{x\in X_{00}}\big|f_{\pm,\varepsilon}(x) - N^{-1}\textstyle\sum_{i=1}^{N} N_{\varepsilon,x}(\xi_i,t_i)\big|\Big].$$

But since $x \mapsto f_{\pm,\varepsilon}(x)$ and $x \mapsto N_{\varepsilon,x}(\xi_i,t_i)$ are continuous, this coincides with the expression in Equation (2.2).

Remark 3.2. a) Using Proposition 7.4, it is not hard to see $\mathcal{B}_C(X, x_0) \subset \mathcal{BA}_{\kappa_0 C}(X)$ for every $C > 0$, with a constant $\kappa_0 > 0$ that is *absolute*, (i.e., independent of all other quantities and objects).

b) For the infinite-width Barron space $\mathcal{B}_\varrho(X)$ associated to the ReLU function (which will be formally introduced in Section 7), it follows from [16, Theorem 12] that

$$\mathcal{B}_{\varrho,C}(X) := \big\{ f \in \mathcal{B}_\varrho(X) \,:\, \|f\|_{\mathcal{B}_\varrho(X)} \leq C \big\} \subset \mathcal{BA}_{\sigma C}(X),$$

where the constant $\sigma > 0$ scales polynomially with d and linearly with $\sup_{x \in X} \|x\|_{\ell^\infty}$.

c) If $Y \subset X$ has nonempty interior, we have $\vartheta(X, y_0) \leq \vartheta(Y, y_0)$ for all $y_0 \in Y$ and hence $\inf_{x_0 \in X} \big[\|x_0\|_{\ell^1} + \vartheta(X, x_0)\big] \leq \inf_{y_0 \in Y} \big[\|y_0\|_{\ell^1} + \vartheta(X, y_0)\big] \leq \inf_{y_0 \in Y} \big[\|y_0\|_{\ell^1} + \vartheta(Y, y_0)\big]$. Based on this, it is straightforward to see

$$f|_Y \in \mathcal{BA}_C(Y) \quad \text{if } f \in \mathcal{BA}(X) \text{ and } Y \subset X \text{ has nonempty interior.} \tag{3.1}$$

Using the notion of Barron approximation spaces, we can now formally define sets with Barron class boundary.

Definition 3.3. Let $d \in \mathbb{N}_{\geq 2}$ and $B > 0$ and let $Q = [a, b] \subset \mathbb{R}^d$ be a rectangle. A function $F : Q \to \mathbb{R}$ is called a *Barron horizon function with constant* B, if there are $i \in \underline{d}$ and $f \in \mathcal{BA}_B\big([a^{(i)}, b^{(i)}]\big)$ as well as $\theta \in \{\pm 1\}$ such that

$$F(x) = \mathbb{1}_{\theta x_i \leq f(x^{(i)})} \qquad \forall\, x \in Q.$$

We write $\mathcal{BH}_B(Q)$ for the set of all such functions.

Finally, given $M \in \mathbb{N}$ and $B > 0$, a compact set $\Omega \subset \mathbb{R}^d$ is said to have a *Barron class boundary with constant* B if there exist rectangles $Q_1, \dots, Q_M \subset \mathbb{R}^d$ such that $\Omega \subset \bigcup_{i=1}^M Q_i$ where the rectangles have disjoint interiors (i.e., $Q_i^\circ \cap Q_j^\circ = \varnothing$ for $i \neq j$) and such that $\mathbb{1}_{Q_i \cap \Omega} \in \mathcal{BH}_B(Q_i)$ for each $i \in \underline{M}$. We write $\mathcal{BB}_{B,M}(\mathbb{R}^d)$ for the class of all such sets. Also, a family $(Q_j)_{j=1}^M$ of rectangles as above is called an *associated cover* of Ω.

Remark 3.4. By Remark 3.2, the set of functions with Barron class boundary contains all characteristic functions of sets whose boundary is locally described by functions in the Fourier-analytic Barron space or the infinite-width Barron space associated to the ReLU.

The following example illustrates the above definition.

Example. (1) Every set Ω of the form $\Omega = \{x \in Q \colon x_1 \leq f(x_2, \dots, x_d)\}$ for a rectangle $Q = [a^{(1)}, b^{(1)}] \times Q' \subset \mathbb{R}^d$ and a function $f : Q' \to \mathbb{R}$ from the Fourier-analytic Barron class $\mathcal{B}_B(Q', x_0)$ (for arbitrary $x_0 \in Q'$) belongs to $\mathcal{BB}_{\kappa_0 B, 1}(\mathbb{R}^d)$, for the absolute constant $\kappa_0 > 0$ from Remark 3.2.

Examples for such functions f are discussed in great length in [4, Section IX]; here, we just mention three special cases. First, for the Gaussian $f(x) = e^{-|x|^2/2}$, it holds that $f \in \mathcal{B}_{\sqrt{d}}(Q', x_0)$ for any rectangle $Q' \subset \mathbb{R}^{d-1}$ and any $x_0 \in Q'$; thus, one only has a polynomial dependence on the dimension. Second, if $Q' = [0,1]^{d-1}$ and $f(x) = \sum_{k \in \mathbb{Z}^{d-1}} c_k e^{2\pi i \langle k, x \rangle}$, then $f \in \mathcal{B}_C(Q', 0)$ for $C = |c_0| + \sum_{k \in \mathbb{Z}^{d-1}} |k|\,|c_k|$; this essentially follows as in [4, Section IX, Point (16)]. Finally, if $f \in C^k(Q')$ for $k \geq 2 + \lfloor (d-1)/2 \rfloor$, then [4, Section IX, Point (15)] shows that f belongs to $\mathcal{B}_C(Q', x_0)$, for a suitable $C = C(f, Q') > 0$. This last observation, however, is more of qualitative than of quantitative use, since the resulting constant C is often quite large if d is large.

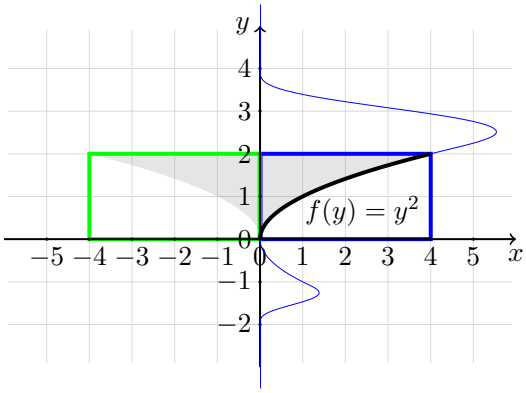

Figure 2: An illustration of the cusp domain Ω (shown in gray) discussed in Part (2) of the example below. The blue and green boxes show the rectangles $Q_1, Q_2 \subset \mathbb{R}^2$ satisfying $\Omega \subset Q_1 \cup Q_2$ and such that $\mathbb{1}_{\Omega \cap Q_i}$ is a Barron horizon function. For more details see the example below.

(2) The class of sets with Barron class boundary also contains sets that are not necessarily Lipschitz domains. An example of such a domain is the cusp domain

$$\Omega = \big\{(x,y) \in [-4,4] \times [0,2] \quad : \quad y \geq \sqrt{|x|}\big\}$$

shown in Figure 2. Indeed, we claim for the rectangles $Q_1 = [0,4]\times[0,2]$ and $Q_2 = [-4,0]\times[0,2]$ that $\mathbb{1}_{\Omega\cap Q_i}$ is a Barron horizon function. We only verify this for Q_1. To see this, note that the function $f : [0,2] \to \mathbb{R}, y \mapsto y^2$ can be extended to a function $f \in C_c^3(\mathbb{R})$; one such extension is shown in Figure 2. As seen above, this implies that $f \in \mathcal{B}_C([0,2],0)$ for a certain $C > 0$. Because of $\mathbb{1}_{\Omega\cap Q_1}(x,y) = \mathbb{1}_{x\leq f(y)}$ for $(x,y) \in Q_1$, this implies that $\mathbb{1}_{\Omega\cap Q_1}$ is a Barron horizon function.

We will show in Section 6 that it is impossible to derive nontrivial minimax bounds for the class of sets with Barron boundary for the case of *general* probability measures. For this reason, we will restrict to the following class of measures.

Definition 3.5. Let μ be a finite Borel measure on $\mathbb{R}^d$. We say that μ is *tube compatible* with parameters $\alpha \in (0,1]$ and $C > 0$ if for each measurable function $f : \mathbb{R}^{d-1} \to \mathbb{R}$, each $i \in \underline{d}$ and each $\varepsilon \in (0,1]$, we have

$$\mu\big(T_{f,\varepsilon}^{(i)}\big) \leq C \cdot \varepsilon^{\alpha} \quad \text{where} \quad T_{f,\varepsilon}^{(i)} := \big\{x \in \mathbb{R}^d \colon |x_i - f(x^{(i)})| \leq \varepsilon\big\}.$$

The set $T_{f,\varepsilon}^{(i)}$ is called a *tube of width ε (associated to f).*

Remark 3.6. The definition might appear technical, but it is satisfied for a wide class of product measures. For instance, if $\mu_1, \ldots, \mu_d$ are Borel probability measures on $\mathbb{R}^d$ such that each distribution function $F_i(x) = \mu_i((-\infty, x])$ is α-Hölder continuous with constant C, then the product measure $\mu = \mu_1 \otimes \cdots \otimes \mu_d$ is tube compatible with parameters α and $2^{\alpha}C$, since Fubini's theorem shows for $\mu^{(i)} := \mu_1 \otimes \cdots \mu_{i-1} \otimes \mu_{i+1} \otimes \cdots \otimes \mu_d$ that

$$\mu(T_{f,\varepsilon}^{(i)}) = \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} \mathbb{1}_{|y-f(x)|\leq\varepsilon} \, d\mu_i(y) \, d\mu^{(i)}(x),$$

where

$$\int_{\mathbb{R}} \mathbb{1}_{|y-f(x)|\leq\varepsilon}\, d\mu_i(y) = \mu_i([f(x)-\varepsilon, f(x)+\varepsilon]) = F_i\big(f(x)+\varepsilon\big) - F_i\big(f(x)-\varepsilon\big) \leq C\cdot(2\varepsilon)^\alpha = 2^\alpha C\cdot\varepsilon^\alpha,$$

from which we easily get $\mu\big(T_{f,\varepsilon}^{(i)}\big) \leq 2^\alpha C \cdot \varepsilon^\alpha$, as claimed.

Measures that do not have a product structure can be tube compatible as well. For example, if μ is tube compatible with parameters $\alpha \in (0,1]$ and $C > 0$, then any measure ν of the form $d\nu = f\, d\mu$ with a bounded density function f will be tube compatible, with parameters α and $C \cdot \sup_x f(x)$.

Next, we give our main approximation result for functions $\mathbb{1}_\Omega$, where Ω is a set with Barron class boundary.

Theorem 3.7. *Let $d \in \mathbb{N}_{\geq 2}$, $M, N \in \mathbb{N}$, $B, C > 0$, and $\alpha \in (0,1]$, and let $\Omega \in \mathcal{BB}_{B,M}(\mathbb{R}^d)$.*

There exists a neural network I_N with 3 hidden layers such that for each tube compatible measure μ with parameters α, C, we have

$$\mu(\{x \in \mathbb{R}^d : \mathbb{1}_\Omega(x) \neq R_\varrho I_N(x)\}) \leq 6CMB^\alpha\, d^{3/2}\, N^{-\alpha/2}.$$

Moreover, $0 \leq R_\varrho I_N(x) \leq 1$ for all $x \in \mathbb{R}^d$ and the architecture of I_N is given by

$$\mathcal{A} = \big(d,\ M(N+2d+2),\ M(4d+2),\ M,\ 1\big).$$

Thus, I_N has at most $7M(N+d)$ neurons and at most $54d^2 M\, N$ non-zero weights. The weights (and biases) of I_N are bounded in magnitude by $d(4+R)(1+B) + \sqrt{N}\cdot\big(B^{-1}+B^{-1/2}\big)$, where $R = \sup_{x\in\Omega}\|x\|_{\ell^\infty}$.

Proof. The proof will proceed in three parts. First we construct a neural network that satisfies a certain approximation accuracy, without going into much detail regarding the architecture of this network. Afterwards, we analyze the network architecture, and bound the network weights.

Network construction and approximation bound:

Step 1. (Construction of neural networks locally approximating boundaries) Let $(Q_j)_{j=1}^M$ be an associated cover of Ω. Fix $m \in \underline{M}$ and write $Q_m := [a,b]$. By the assumption $\Omega \in \mathcal{BB}_{B,M}(\mathbb{R}^d)$, there exist $i = i(m) \in \underline{d}$ and $\theta_m \in \{\pm 1\}$ as well as a function $f_m \in \mathcal{BA}_B(Q_m^i)$ such that $\mathbb{1}_\Omega(x) = \mathbb{1}_{\theta_m x_i \leq f_m(x^{(i)})}$ for all $x \in Q_m$. Here, we used the notation $Q_m^i := \prod_{j\neq i}[a_j, b_j]$. With $R = \sup_{x\in\Omega}\|x\|_{\ell^\infty}$ as in the theorem statement, note that if we replace each Q_j by $\widetilde{Q_j} := Q_j \cap [-R,R]^d$, then the family $(\widetilde{Q_j})_{j=1}^M$ is still a cover of Ω consisting of rectangles. Furthermore, Equation (3.1) shows that $f_m \in \mathcal{BA}_C(\widetilde{Q_m^i})$, and we clearly have $\mathbb{1}_\Omega(x) = \mathbb{1}_{\theta_m x_i \leq f_m(x^{(i)})}$ for all $x \in \widetilde{Q_m}$. Therefore, we can assume in the following that $Q_m \subset [-R,R]^d$ for all $m \in \underline{M}$.

Now, by Definitions 3.3 and 3.1, there exists a shallow neural network I_N^m with N neurons in the hidden layer such that $\|f_m - R_\varrho I_N^m\|_{\sup} \leq \gamma N^{-1/2}$ where $\gamma := B\sqrt{d-1}$. Furthermore, all weights and biases of I_N^m are bounded by $\sqrt{B}\cdot\big(6 + \vartheta(Q_m^i, q_m) + \|q_m\|_{\ell^1}\big)$ for some $q_m \in Q_m^i \subset [-R,R]^{d-1}$.

Step 2. (Construction of neural networks approximating horizon functions) Set

$$S_m := \big\{x \in Q_m : f_m(x^{(i)}) \geq \theta_m\, x_i\big\} \qquad \text{and} \qquad T_m := \big\{x \in Q_m : R_\varrho I_N^m(x^{(i)}) \geq \theta_m\, x_i\big\},$$

where I_N^m is the network obtained in the previous step. Recalling $\|f_m - R_\varrho I_N^m\|_{\sup} \leq \gamma\, N^{-1/2}$ and using the notation $S_m \triangle T_m = (S_m \setminus T_m) \cup (T_m \setminus S_m)$, we then see

$$\begin{aligned}
&S_m \triangle T_m \\
&= \big\{x \in Q_m : f_m(x^{(i)}) < \theta_m x_i \leq R_\varrho I_N^m(x^{(i)})\big\} \cup \big\{x \in Q_m : R_\varrho I_N^m(x^{(i)}) < \theta_m x_i \leq f_m(x^{(i)})\big\} \\
&\subset \big\{x \in Q_m : -\gamma N^{-1/2} \leq f_m(x^{(i)}) - \theta_m x_i < 0\big\} \cup \big\{x \in Q_m : 0 \leq f_m(x^{(i)}) - \theta_m x_i < \gamma N^{-1/2}\big\} \\
&\subset \big\{x \in Q_m : |f_m(x^{(i)}) - \theta_m x_i| \leq \gamma N^{-1/2}\big\}.
\end{aligned}$$

Since μ is α, C tube compatible and since $\mathbb{1}_\Omega(x) = \mathbb{1}_{S_m}(x)$ for $x \in Q_m$, it follows that

$$\begin{aligned}
\mu(\{x \in Q_m : \mathbb{1}_\Omega(x) \neq \mathbb{1}_{T_m}(x)\}) &= \mu(\{x \in Q_m : \mathbb{1}_{S_m}(x) \neq \mathbb{1}_{T_m}(x)\}) \\
&= \mu(S_m \triangle T_m) \leq C\gamma^\alpha N^{-\alpha/2}.
\end{aligned}$$

Next, we define the approximate Heaviside function $H_\delta : \mathbb{R} \to [0,1]$ by

$$H_\delta(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{x}{\delta} & \text{if } 0 \leq x \leq \delta \\ 1 & \text{if } x \geq 1. \end{cases}$$

Since H_δ can be realized by a ReLU neural network (via $H_\delta(x) = \frac{1}{\delta}(\varrho(x) - \varrho(x-\delta))$), we next approximate the characteristic function of T_m by an appropriate approximate Heaviside function applied to $R_\varrho I_N^m(x^{(i)}) - \theta_m x_i$.

To this end, note for $\delta > 0$ and an arbitrary measurable function $\phi : \mathbb{R}^{d-1} \to \mathbb{R}$ that

$$\begin{aligned}
\big\{(t,u) \in \mathbb{R}^{d-1} \times \mathbb{R} : \mathbb{1}_{\phi(t) \geq u} \neq H_\delta(\phi(t) - u)\big\} &= \{(t,u) : 0 < H_\delta(\phi(t) - u) < 1\} \\
&\subset \{(t,u) : 0 \leq \phi(t) - u \leq \delta\} \\
&\subset \{(t,u) : |\phi(t) - u| \leq \delta\}.
\end{aligned}$$

Therefore, by picking $\delta = \gamma N^{-1/2}$ and using the tube compatibility of the measure we see that $\mu(\{x \in Q_m : \mathbb{1}_{T_m}(x) \neq R_\varrho J_N^m(x)\}) \leq C\gamma^\alpha N^{-\alpha/2}$, where J_N^m is chosen such that $R_\varrho J_N^m(x) = H_{\gamma N^{-1/2}}\big(R_\varrho I_N^m(x^{(i)}) - \theta_m x_i\big)$. Note that $0 \leq R_\varrho J_N^m \leq 1$.

Step 3. (Localization to patches) Next, we want to truncate each realization $R_\varrho J_N^m$ such that it is supported on Q_m and we want to realize these truncations as ReLU neural networks. This is based on a simplified version of the argument in [45, Lemma A.6] For the sake of completeness, we recall the construction from [45, Lemma A.6].

Let $[a,b] = \prod_{i=1}^d [a_i, b_i]$ be a rectangle in $\mathbb{R}^d$, let $0 < \varepsilon \leq \frac{1}{2}\min_{i \in \underline{d}}(b_i - a_i)$ and define $[a+\varepsilon, b-\varepsilon] := \prod_{i=1}^d [a_i + \varepsilon, b_i - \varepsilon]$. Furthermore, define the functions $t_i : \mathbb{R} \to \mathbb{R}$, for $i \in \underline{d}$, by

$$t_i(u) := \begin{cases} 0 & \text{if } u \in \mathbb{R} \setminus [a_i, b_i] \\ 1 & \text{if } u \in [a_i + \varepsilon, b_i - \varepsilon] \\ \frac{u - a_i}{\varepsilon} & \text{if } u \in [a_i, a_i + \varepsilon] \\ \frac{b_i - u}{\varepsilon} & \text{if } u \in [b_i - \varepsilon, b_i], \end{cases}$$

and $\eta_\varepsilon : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ by $\eta_\varepsilon(x,y) = \varrho\big(\sum_{i=1}^d t_i(x_i) + \varrho(y) - d\big)$. Note that for $y \in [0,1]$, if $x \in [a+\varepsilon, b-\varepsilon]$, we have $\eta_\varepsilon(x,y) = \varrho(\varrho(y)) = y$; furthermore, if $x \in \mathbb{R}^d \setminus [a,b]$, we have $0 \leq \eta_\varepsilon(x,y) \leq \varrho(d - 1 + \varrho(y) - d) = \varrho(y-1) = 0$. This implies for any function $g : \mathbb{R}^d \to [0,1]$ that $\{x \in \mathbb{R}^d : \eta_\varepsilon(x, g(x)) \neq \mathbb{1}_{[a,b]}(x) \cdot g(x)\} \subset [a,b] \setminus [a+\varepsilon, b-\varepsilon]$. Note additionally that

the function η_ε can be implemented by a ReLU neural network and that $0 \leq t_i \leq 1$, so that $0 \leq \eta_\varepsilon(x,y) \leq \varrho(\varrho(y)) \leq 1$ for all $y \in [0,1]$, by monotonicity of the ReLU.

Returning now to the neural networks constructed in the previous step we distinguish two cases: First, if the rectangle Q_m has width along some coordinate direction i less than $2\gamma N^{-1/2}$ (Q_m is a "small rectangle"), then we see for a suitable (constant) function $g_m : \mathbb{R}^{d-1} \to \mathbb{R}$ that $Q_m \subset T^{(i)}_{g_m, 2\gamma N^{-1/2}}$ and hence $\mu(Q_m) \leq 2^\alpha C\gamma^\alpha N^{-\alpha/2} \leq 2dC\gamma^\alpha N^{-\alpha/2}$, since $\alpha \leq 1$. We thus choose L_N^m to be a trivial neural network with input dimension $d+1$, meaning $R_\varrho L_N^m(x,y) = 0$ for all $x \in \mathbb{R}^d$ and $y \in \mathbb{R}$. We then have

$$\mu\big(\big\{x \in \mathbb{R}^d : \mathbb{1}_{Q_m}(x) R_\varrho J_N^m(x) \neq R_\varrho L_N^m(x, R_\varrho J_N^m(x))\big\}\big) \leq \mu(Q_m) \leq 2d\, C\, \gamma^\alpha\, N^{-\alpha/2}.$$

Otherwise (if Q_m is a "large rectangle"), writing $Q_m = [a,b]$, we have $\frac{\gamma}{\sqrt{N}} \leq \frac{1}{2}\min_{i \in \underline{d}}(b_i - a_i)$, and it is not hard to see that $[a,b] \setminus [a+\gamma N^{-1/2}, b-\gamma N^{-1/2}]$ is contained in the union of $2d$ tubes of width $\gamma N^{-1/2}$. Therefore, choosing L_N^m such that $R_\varrho L_N^m = \eta_{\gamma N^{-1/2}}$, we obtain

$$\mu\big(\big\{x \in \mathbb{R}^d : \mathbb{1}_{Q_m}(x) R_\varrho J_N^m(x) \neq R_\varrho L_N^m(x, R_\varrho J_N^m(x))\big\}\big) \leq 2dC\gamma^\alpha N^{-\alpha/2}.$$

In both cases, the function $x \mapsto R_\varrho L_N^m(x, R_\varrho J_N^m(x))$ is supported on Q_m and vanishes on the boundary of Q_m (due to continuity).

Step 4. (Finishing the construction and error estimate) To summarize, on each rectangle Q_m we have

$$\begin{aligned}
&\mu(\{x \in \mathbb{R}^d : \mathbb{1}_{\Omega \cap Q_m}(x) \neq R_\varrho L_N^m(x, R_\varrho J_N^m(x))\}) \\
&\leq \mu(\{x \in \mathbb{R}^d : R_\varrho L_N^m(x, R_\varrho J_N^m(x)) \neq \mathbb{1}_{Q_m}(x) R_\varrho J_N^m(x)\}) \\
&\quad + \mu(\{x \in Q_m : R_\varrho J_N^m(x) \neq \mathbb{1}_{T_m}(x)\}) + \mu(\{x \in Q_m : \mathbb{1}_{T_m}(x) \neq \mathbb{1}_\Omega(x)\}) \\
&\leq 2dC\gamma^\alpha N^{-\alpha/2} + C\gamma^\alpha N^{-\alpha/2} + C\gamma^\alpha N^{-\alpha/2} \\
&= 2(d+1)C\gamma^\alpha N^{-\alpha/2}.
\end{aligned}$$

Now, defining the neural network I_N such that $R_\varrho I_N(x) := \sum_{m=1}^M R_\varrho L_N^m(x, R_\varrho J_N^m(x))$, we obtain because of $\mathbb{1}_\Omega = \sum_{m=1}^M \mathbb{1}_{\Omega \cap Q_m}$ (almost everywhere) that

$$\mu\big(\big\{x \in \mathbb{R}^d : \mathbb{1}_\Omega(x) \neq R_\varrho I_N(x)\big\}\big) \leq 2M(d+1)C\gamma^\alpha N^{-\alpha/2} = 2(d+1)(d-1)^{\alpha/2} CMB^\alpha N^{-\alpha/2}.$$

To simplify the estimate, using that $\alpha \leq 1$, we see $(d+1)(d-1)^{\alpha/2} \leq (d+1)^{3/2} \leq (2d)^{3/2}$, since $d \geq 2$. Finally, note that $2^{1+3/2} = 2^{5/2} < 6$. Combining these estimates we see that $2(d+1)(d-1)^{\alpha/2} CMB^\alpha N^{-\alpha/2} \leq 6CMB^\alpha\, d^{3/2} N^{-\alpha/2}$.

Additionally, recall from above that $0 \leq R_\varrho J_N^m \leq 1$ for every $m \in \underline{M}$. As seen in Step 3, this implies that $\zeta_m(x) := R_\varrho L_N^m(x, R_\varrho J_N^m(x))$ satisfies $0 \leq \zeta_m(x) \leq 1$ for all $x \in \mathbb{R}^d$. Since each ζ_m is supported on Q_m and vanishes on the boundary of Q_m, and since the rectangles Q_m have disjoint interiors, this implies that $0 \leq R_\varrho I_N \leq 1$ as well.

The architecture:

Now let us examine the architecture of each L_N^m in more detail. For each rectangle Q_m, the flowchart of computations performed by each L_N^m can be visually represented as in Figure 3.

In the following, we explicitly describe each of the layers of the network computing L_N^m; we then describe how these networks are combined to obtain I_N.

Inputs. The input layer with d neurons corresponding to the d coordinates of an input $x \in \mathbb{R}^d$.

Layer 1. This layer will contain $N + 2d + 2$ neurons split into 3 categories:

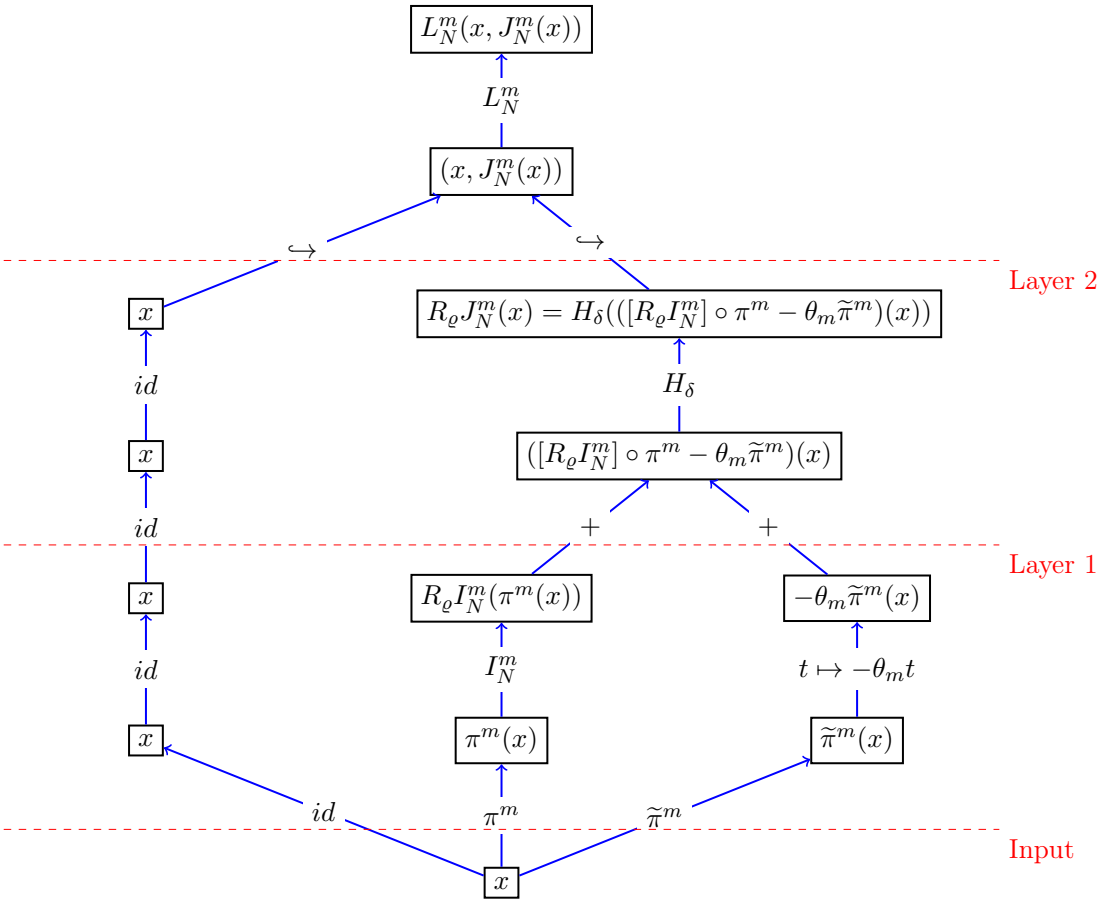

Figure 3: Visualization of the neural network L_N^m for the case of a "large" rectangle Q_m.

- $2d$ neurons computing $\varrho(x_i)$ and $\varrho(-x_i)$ respectively for each $i \in \underline{d}$.
- N neurons corresponding to the neurons in the hidden layer of the networks I_N^m. Explicitly, writing $R_\varrho I_N^m(x) = D + \sum_{k=1}^N C_k\, \varrho(B_k + \langle A_k, x\rangle)$ with $D, B_k, C_k \in \mathbb{R}$ and $A_k \in \mathbb{R}^{d-1}$ for $k \in \underline{N}$, the k-th of these neurons will compute $\phi_k(x) = \varrho\big(B_k + \langle (\pi^m)^T A_k, x\rangle\big)$, where π^m is the projection that sends x to $x^{(i)}$ (with $i = i(m)$), viewed as a $(d-1) \times d$ matrix.
- 2 neurons computing $\varrho(\pm\theta_m \widetilde{\pi}^m(x))$, respectively, where $\widetilde{\pi}^m$ is the projection that sends x to x_i (where $i = i(m)$), viewed as a $1 \times d$ matrix.

Layer 2. This layer will contain $4d + 2$ neurons split into 2 categories:

- 4 neurons for each coordinate $i \in \underline{d}$ computing the building blocks for the t_i functions in Step 3: $t_i^1(u_i) = \varrho(u_i - a_i)$, $t_i^2(u_i) = \varrho(u_i - a_i - \varepsilon)$, $t_i^3(u_i) = \varrho(u_i - b_i + \varepsilon)$ and $t_i^4(u_i) = \varrho(u_i - b_i)$, where $u_i := \varrho(x_i) - \varrho(-x_i) = x_i$. Note that $t_i(u_i) = \frac{t_i^1 - t_i^2 - t_i^3 + t_i^4}{\varepsilon}(u_i)$. Furthermore, recall that we chose $\varepsilon = \gamma N^{-1/2}$.
- 2 neurons computing the parts of the approximate Heaviside function H_δ, computing,

respectively,

$$\psi_1(x) := \varrho\Big(D + \sum_{k=1}^{N} C_k\phi_k(x) - \varrho\big(\theta_m\widetilde{\pi}^m(x)\big) + \varrho\big(-\theta_m\widetilde{\pi}^m(x)\big)\Big)$$

and

$$\psi_2(x) := \varrho\Big(D + \sum_{k=1}^{N} C_k\phi_k(x) - \varrho\big(\theta_m\widetilde{\pi}^m(x)\big) + \varrho\big(-\theta_m\widetilde{\pi}^m(x)\big) - \delta\Big),$$

where we recall from above that $R_\varrho I_N^m(\pi^m(x)) = D + \sum_{k=1}^{N} C_k\phi_k(x)$ and $\delta = \gamma N^{-1/2}$. Therefore, $\frac{1}{\delta}(\psi_1(x) - \psi_2(x)) = H_\delta\big(R_\varrho I_N^m(\pi^m(x)) - \theta_m\widetilde{\pi}^m(x)\big) = R_\varrho J_N^m(x)$; in particular, $\psi_1(x) - \psi_2(x) \geq 0$.

Layer 3. This layer will have a single neuron, either computing the zero function (in the case of a "small rectangle" Q_m), or (in the case of a "large rectangle") computing

$$\eta_\varepsilon(x, R_\varrho J_N^m(x)) = R_\varrho L_N^m(x, R_\varrho J_N^m(x)) = \varrho\Big(\tfrac{1}{\varepsilon}\sum_{i=1}^{d}\big(t_i^1 - t_i^2 - t_i^3 + t_i^4\big)(x_i) + \tfrac{1}{\delta}\big(\psi_1(x) - \psi_2(x)\big) - d\Big).$$

We used here that $(\psi_1 - \psi_2)(x) \geq 0$, so the difference is invariant under ϱ.

Now, the full network I_N can be realized with one more layer (the output layer), so that $R_\varrho I_N(x) = \sum_{m=1}^{M} R_\varrho L_N^m(x, R_\varrho J_N^m(x))$.

Thus, I_N can be realized by a ReLU neural network with 3 hidden layers, architecture $\mathcal{A} = \big(d,\, M(N+2d+2),\, M(4d+2),\, M,\, 1\big)$, and $d+1+M(N+6d+5) \leq 7M(N+d)$ neurons.

Now let us estimate the number of non-zero weights of I_N which we will denote by $W(I_N)$. An immediate bound can be found by taking the product of the number of neurons on every pair of consecutive layers in the L_N^m networks, summing up over the layers, multiplying by M, adding M to account for the weights of the final output layer, and finally adding the total number of non-input neurons to account for the biases. We thus see

$$W(I_N) \leq M \cdot \big(d(N+2d+2) + (N+2d+2)(4d+2) + (4d+2)\cdot 1\big) + M + MN + 6Md + 5M + 1,$$

so that a rough estimate shows $W(I_N) \leq 54Md^2\,N$.

Bounding the magnitude of the weights and biases:

Let us now acquire an upper bound for the absolute value of the weights and biases of I_N. Note first of all that for the networks I_N^m we have two cases depending on the size of the corresponding rectangle $Q_m = \prod_{i=1}^{d}[a_i, b_i]$:

- If $\min_i(b_i - a_i) < 2\gamma N^{-1/2}$, we can set all weights of the "subnetwork" corresponding to the rectangle Q_m to be zero.
- If $\min_i(b_i - a_i) \geq 2\gamma N^{-1/2}$, then by Remark 2.3, we have $\vartheta(Q_m, q_m) \leq \gamma^{-1}N^{1/2}$. Since furthermore $\|q_m\|_{\ell^1} \leq (d-1)R$, our choice of I_N^m in Step 1 ensures that the weights and biases of I_N^m are bounded by $\sqrt{B}\cdot(6 + \vartheta(Q_m, q_m) + \|q_m\|_{\ell^1}) \leq \sqrt{B}\cdot(6 + \gamma^{-1}N^{1/2} + dR)$.

In either case, we see that the weights and biases on the first layer are bounded in absolute value by $1 + \sqrt{B}\cdot(6 + \gamma^{-1}N^{1/2} + dR)$.

For the second layer, the weights corresponding to the first $4d$ neurons are bounded by $1 + \varepsilon + R$ and for the last 2 neurons again by $1 + \sqrt{B}\cdot(6 + \gamma^{-1}N^{1/2} + dR)$. Finally for the third layer, the weights and biases are bounded by $\max(\frac{1}{\varepsilon}, \frac{1}{\delta}, d) \leq d + \gamma^{-1}N^{1/2}$.

In conclusion, the weights of I_N will have magnitudes bounded by

$$\begin{aligned}&\max\left\{1+6\sqrt{B}+\sqrt{B}\gamma^{-1}N^{1/2}+\sqrt{B}dR,\ 1+\varepsilon+R,\ d+\gamma^{-1}N^{1/2}\right\}\\&\leq d(4+R)(1+B)+\sqrt{N}\cdot\left(B^{-1}+B^{-1/2}\right).\end{aligned}$$

Here, we used that $d \geq 2$, combined with several elementary estimates including the bound $\sqrt{B} \leq 1 + B$. □

4 Lower bounds for approximating sets with Barron class boundary

In this section, we present a lower bound on the achievable minimax rate for approximating functions $f \in \mathcal{BB}_{B,M}(\mathbb{R}^d)$ in L^1 with respect to the Lebesgue measure on $[-1,1]^d$. In fact, we show that the approximation rate provided by Theorem 3.7 is *almost optimal* even when only horizon functions with boundary from the Fourier-analytic Barron space are considered. More precisely, we will see in Theorem 4.3 below that neural networks with W weights cannot obtain a better approximation error than $\mathcal{O}(W^{-\frac{1}{2}-\frac{1}{d-1}})$ over the class of horizon functions with boundary from the Fourier-analytic Barron space; in contrast, our upper bound from Theorem 3.7 guarantees that an approximation error of $\mathcal{O}(W^{-1/2})$ is achievable. Thus, even though the two rates of approximation do not precisely agree, the difference between them vanishes for increasing input dimension $d \to \infty$; therefore, we speak of *almost optimality*.

Since the arguments in this section are heavily based on the Fourier transform, we start by fixing its normalization. Concretely, for $f \in L^1(\mathbb{R}^d)$, we define

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-i\langle x,\xi\rangle}\, dx,$$

so that the inverse Fourier transform is given by $\mathcal{F}^{-1}f(x) = \widehat{f}(-x)$; see e.g. [22, Section 4.3.1].

Our first step towards lower bounds is to relate the covering numbers of certain sets of horizon functions to covering numbers of certain subsets of Besov spaces. To fix the terminology, recall that if Θ is a subset of a normed vector space X, then a set $\varnothing \neq M \subset X$ is called an ε-net for Θ (in X), if $\sup_{x\in\Theta} \inf_{m\in M} \|x - m\|_X \leq \varepsilon$.

Proposition 4.1. *Let $d \in \mathbb{N}_{\geq 2}$ and $\Omega := (-1,1)^{d-1}$. Given a function $f : \Omega \to \mathbb{R}$, define the associated* horizon function *as*

$$H_f : \quad (-1,1)^{d-1} \times (-1,1) \to \{0,1\}, \quad (x,y) \mapsto \mathbb{1}_{y\leq f(x)}.$$

For each $s > \frac{d+1}{2}$ and $C > 0$, there is a constant $\lambda = \lambda(d,s,C) > 0$ with the following property: If $\varepsilon > 0$ and if $M \subset L^1\big((-1,1)^d\big)$ is an ε-net (in $L^1((-1,1)^d)$) for

$$\mathrm{HF}(\mathcal{B}_C) := \big\{H_f \,:\, f \in \mathcal{B}_C([-1,1]^{d-1})\big\},$$

then there exists a set $M' \subset B_0^{1,\infty}(\Omega)$ satisfying $|M'| \leq |M|$ which is a $\lambda\varepsilon$-net (in $B_0^{1,\infty}(\Omega)$) for $\mathcal{G} := \big\{f \in B_s^{2,2}(\Omega) \colon \|f\|_{B_s^{2,2}} \leq 1\big\}$.

Remark 4.2. Here, for an open set $\Omega \subset \mathbb{R}^{d-1}$, we use the definition of the Besov spaces $B_s^{p,q}(\Omega)$ as in [19, Section 2.5.1], that is, $B_s^{p,q}(\Omega) = \big\{f|_\Omega \,:\, f \in B_s^{p,q}(\mathbb{R}^{d-1})\big\}$ with norm

$$\|f\|_{B_s^{p,q}(\Omega)} := \inf\big\{\|g\|_{B_s^{p,q}(\mathbb{R}^{d-1})} \colon g \in B_s^{p,q}(\mathbb{R}^{d-1}) \text{ and } f = g|_\Omega\big\}.$$

Proof. We divide the proof into four steps.

Step 1: For completeness, we prove the well-known embedding $L^1(\Omega) \hookrightarrow B_0^{1,\infty}(\Omega)$. Clearly, it is enough to prove $L^1(\mathbb{R}^{d-1}) \hookrightarrow B_0^{1,\infty}(\mathbb{R}^{d-1})$. To this end, recall from [19, Section 2.2.1] that the norm on $B_0^{1,\infty}(\mathbb{R}^{d-1})$ is given by

$$\|f\|_{B_0^{1,\infty}} = \sup_{j\in\mathbb{N}_0} \left\| \mathcal{F}^{-1}\big[\varphi_j \, \widehat{f}\,\big] \right\|_{L^1},$$

where $\varphi_0, \varphi \in \mathcal{S}(\mathbb{R}^{d-1})$ are suitably chosen and $\varphi_k(\xi) = \varphi_1(2^{-k+1}\xi)$ for $k \in \mathbb{N}$. Note that $\|\mathcal{F}^{-1}\varphi_k\|_{L^1} = \big\|2^{(k-1)(d-1)}(\mathcal{F}^{-1}\varphi_1)(2^{k-1}\cdot)\big\|_{L^1} = \|\mathcal{F}^{-1}\varphi_1\|_{L^1}$, whence $\|\mathcal{F}^{-1}\varphi_k\|_{L^1} \leq C_1 < \infty$ for all $k \in \mathbb{N}_0$. By Young's inequality for convolutions, this implies

$$\left\| \mathcal{F}^{-1}\big[\varphi_j \, \widehat{f}\,\big] \right\|_{L^1} = C_0 \left\| (\mathcal{F}^{-1}\varphi_j) * f \right\|_{L^1} \leq C_1' \cdot \|f\|_{L^1},$$

so that $\|f\|_{B_0^{1,\infty}} \leq C_1' \cdot \|f\|_{L^1}$ for $f \in L^1(\mathbb{R}^{d-1})$, with $C_1' = C_1'(d) > 0$.

Step 2: We show existence of $c = c(s,d,C) > 0$ such that every $f \in \mathcal{G}$ satisfies $\|cf\|_{\sup} \leq \frac{1}{4}$ and $cf \in \mathcal{B}_C([-1,1]^{d-1})$. We remark that this inclusion was (up to minute differences) already observed in [4, Example 15 on Page 941] and [18, Theorem 3.1]. We provide the proof here for the sake of completeness.

To this end, we first prove $\int_{\mathbb{R}^{d-1}} (1+|\xi|)\,|\widehat{f}(\xi)|\,d\xi \leq C_2 \cdot \|f\|_{B_s^{2,2}}$ for all $f \in B_s^{2,2}(\mathbb{R}^{d-1})$, for a suitable constant $C_2 = C_2(s,d) > 0$. First, recall from [53, Sections 2.2.2, 2.3.2, and 2.5.6] the well-known identity $B_s^{2,2}(\mathbb{R}^{d-1}) = F_s^{2,2}(\mathbb{R}^{d-1}) = H^{s,2}(\mathbb{R}^{d-1})$, where the norm on the Sobolev space $H^{s,2}(\mathbb{R}^{d-1})$ is given by $\|f\|_{H^{s,2}}^2 = \int_{\mathbb{R}^{d-1}} (1+|\xi|^2)^s\,|\widehat{f}(\xi)|^2\,d\xi$. Using the Cauchy-Schwarz inequality, we therefore see for $f \in B_s^{2,2}(\mathbb{R}^{d-1})$ that

$$\begin{aligned}
\int_{\mathbb{R}^{d-1}} \big(1+|\xi|\big)\,|\widehat{f}(\xi)|\,d\xi &\lesssim \int_{\mathbb{R}^{d-1}} \big(1+|\xi|^2\big)^{\frac{1-s}{2}} \big(1+|\xi|^2\big)^{s/2}\,|\widehat{f}(\xi)|\,d\xi \\
&\leq \Big(\int_{\mathbb{R}^{d-1}} \big(1+|\xi|^2\big)^{1-s}\,d\xi \Big)^{1/2} \Big(\int_{\mathbb{R}^{d-1}} \big(1+|\xi|^2\big)^{s}\,|\widehat{f}(\xi)|^2\,d\xi \Big)^{1/2} \\
&\lesssim \|f\|_{H^{s,2}} \lesssim \|f\|_{B_s^{2,2}}.
\end{aligned}$$

Here, we used that $2\cdot(1-s) < -(d-1)$, so that $\int_{\mathbb{R}^{d-1}} (1+|\xi|^2)^{1-s}\,d\xi < \infty$.

Now, by definition of $\mathcal{G}$ and of $B_s^{2,2}(\Omega)$, for each $f \in \mathcal{G}$, there is $F \in B_s^{2,2}(\mathbb{R}^{d-1})$ with $f = F|_\Omega$ and $\|F\|_{B_s^{2,2}} \leq 2$. As seen above, this entails $\int_{\mathbb{R}^{d-1}} (1+|\xi|)\,|\widehat{F}(\xi)|\,d\xi \leq 2C_2$. On the one hand, this implies by Fourier inversion for all $x \in \Omega$ that

$$|f(x)| = |F(x)| = |(\mathcal{F}^{-1}\widehat{F})(x)| \leq (2\pi)^{-(d-1)/2} \int_{\mathbb{R}^{d-1}} |\widehat{F}(\xi)|\,d\xi \leq (2\pi)^{-(d-1)/2} 2C_2 \leq 2C_2.$$

On the other hand, $f(x) = F(x) = \mathcal{F}^{-1}\widehat{F}(x) = \int_{\mathbb{R}^{d-1}} e^{i\langle x,\xi\rangle} (2\pi)^{-\frac{d-1}{2}}\,\widehat{F}(\xi)\,d\xi$ for $x \in [-1,1]^{d-1}$ and, in the notation of Definition 2.1, $|\xi|_{[-1,1]^{d-1},0} = \sup_{x\in[-1,1]^{d-1}} |\langle \xi, x\rangle| = \|\xi\|_{\ell^1} \lesssim |\xi|$, meaning $\int_{\mathbb{R}^{d-1}} |\xi|_{[-1,1]^{d-1},0} \cdot (2\pi)^{-(d-1)/2} |\widehat{F}(\xi)|\,d\xi \lesssim C_2 \lesssim 1$. By combining these observations, it is easy to see $\|cf\|_{\sup} \leq \frac{1}{4}$ and $cf \in \mathcal{B}_C([-1,1]^{d-1})$, for $c = c(s,d,C) > 0$ small enough.

Step 3: We show that

$$\|H_f - H_g\|_{L^1((-1,1)^d)} \geq \|f-g\|_{L^1(\Omega)} \qquad \forall\, f,g : \Omega \to \big[-\tfrac{1}{2},\tfrac{1}{2}\big] \text{ measurable.}$$

To see this, first note by Fubini's theorem that

$$\begin{aligned}\|H_f - H_g\|_{L^1((-1,1)^d)} &= \int_{(-1,1)^{d-1}} \int_{-1}^{1} |\mathbb{1}_{y\leq f(x)} - \mathbb{1}_{y\leq g(x)}| \, dy \, dx \\ &\overset{(*)}{=} \int_{(-1,1)^{d-1}} |f(x) - g(x)| \, dx = \|f - g\|_{L^1(\Omega)}.\end{aligned}$$

Here, the step marked with $(*)$ used that $\int_{-1}^{1} |\mathbb{1}_{y\leq f(x)} - \mathbb{1}_{y\leq g(x)}| \, dy = |f(x) - g(x)|$, which is trivial if $f(x) = g(x)$. Otherwise, if $f(x) > g(x)$, then

$$\big|\mathbb{1}_{y\leq f(x)} - \mathbb{1}_{y\leq g(x)}\big| = \mathbb{1}_{(g(x),f(x))}(y),$$

which implies the claimed estimate. Here, we implicitly used that $(g(x), f(x)) \subset [-1, 1]$, since $f(x), g(x) \in [-\frac{1}{2}, \frac{1}{2}]$. For $f(x) < g(x)$, one can argue similarly.

Step 4: We complete the proof. To this end, write $N := |M|$ and $M = \{G_1, \dots, G_N\}$. With $c > 0$ as in Step 2, for each $i \in \underline{N}$, choose $f_i \in \mathcal{G}$ with

$$\|H_{cf_i} - G_i\|_{L^1} \leq \varepsilon + \inf_{f\in\mathcal{G}} \|H_{cf} - G_i\|_{L^1}. \tag{4.1}$$

We claim that $M' := \{f_1, \dots, f_N\} \subset L^1(\Omega) \subset B_0^{1,\infty}(\Omega)$ is a $\lambda\varepsilon$-net for $\mathcal{G}$ (in $B_0^{1,\infty}(\Omega)$), for a suitable choice of $\lambda = \lambda(d, s, C) > 0$.

To see this, let $f \in \mathcal{G}$ be arbitrary. By Step 2, we have $cf \in \mathcal{B}_C([-1, 1]^{d-1})$ and hence $H_{cf} \in \mathrm{HF}(\mathcal{B}_C)$. Since M is an ε-net for $\mathrm{HF}(\mathcal{B}_C)$ (in $L^1((-1, 1)^d)$), this implies that there exists $i \in \underline{N}$ with $\|H_{cf} - G_i\|_{L^1} \leq 2\varepsilon$. Since $f, f_i \in \mathcal{G}$ and hence $\|c\, f\|_{\sup}, \|c\, f_i\|_{\sup} \leq \frac{1}{2}$ by Step 2, the estimates from Steps 1 and 3 show

$$\begin{aligned}\|f - f_i\|_{B_0^{1,\infty}(\Omega)} &\leq C_1' \, \|f - f_i\|_{L^1(\Omega)} = \frac{C_1'}{c} \, \big\|c\, f - c\, f_i\big\|_{L^1(\Omega)} \leq \frac{C_1'}{c} \, \big\|H_{cf} - H_{cf_i}\big\|_{L^1((-1,1)^d)} \\ &\leq \frac{C_1'}{c}\big(\|H_{cf} - G_i\|_{L^1} + \|G_i - H_{cf_i}\|_{L^1}\big) \\ &\overset{(**)}{\leq} \frac{C_1'}{c}\big(\|H_{cf} - G_i\|_{L^1} + \varepsilon + \|H_{cf} - G_i\|_{L^1}\big) \leq \frac{5C_1'}{c}\varepsilon.\end{aligned}$$

Here, the step marked with $(**)$ is justified by Equation (4.1). □

Based on Proposition 4.1, we can now prove our first lower bound for the approximation of Barron-class horizon functions. This result uses the notion of (τ, ε)-*quantized networks* introduced in [45, Definition 2.9]. Precisely, given $\tau \in \mathbb{N}$ and $\varepsilon \in (0, \frac{1}{2})$, we say that a network Φ is (τ, ε)-quantized, if all the weights and biases of Φ belong to the set $[-\varepsilon^{-\tau}, \varepsilon^{-\tau}] \cap 2^{-\tau\lceil \log_2(1/\varepsilon)\rceil}\mathbb{Z}$. Similar notions of quantized networks have been employed in [13, 21] in the context of lower bounds on approximation rates.

Theorem 4.3. *Let $d \in \mathbb{N}_{\geq 2}$, $\tau \in \mathbb{N}$, and $C, \sigma > 0$. With notation as in Proposition 4.1, assume that there are $C_1, C_2 > 0$ and a null-sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0, \infty)$ such that for every $f \in \mathcal{B}_C([-1, 1]^{d-1})$ and $n \in \mathbb{N}$, there is a network Φ with d-dimensional input and 1-dimensional output, with (τ, ε_n)-quantized weights, and such that*

$$\|H_f - R_\varrho\Phi\|_{L^1((-1,1)^d)} \leq C_1\, \varepsilon_n \qquad \text{and} \qquad W(\Phi) \leq C_2 \cdot \varepsilon_n^{-\sigma}.$$

Then $\frac{1}{\sigma} \leq \frac{1}{2} + \frac{1}{d-1}$.

Proof. Let $\Omega_0 := B_{1/2}(0) = \{x \in \mathbb{R}^{d-1} : |x| < 1/2\}$, noting that this is a bounded C^∞-domain in the sense of [53, Section 3.2.1], and that $\Omega_0 \subset \Omega = (-1,1)^{d-1}$. Let us fix $s > \frac{d+1}{2}$ for the moment, and define $A := B^{2,2}_s(\Omega_0)$ and $B := B^{1,\infty}_0(\Omega_0)$. The proof is based on existing *entropy bounds* for the embedding $A \hookrightarrow B$. More precisely, writing $U_A := \{x \in A : \|x\|_A \leq 1\}$ (and similarly for U_B), the k-th entropy number of this embedding is defined as

$$e_k := \inf\left\{\varepsilon > 0 \quad : \quad \exists f_1, \dots, f_{2^{k-1}} \in B \text{ such that } U_A \subset \bigcup_{i=1}^{2^{k-1}} (f_i + \varepsilon\, U_B)\right\};$$

see [19, Definition 1 in Section 1.3.1]. Furthermore, [19, Theorem 1 in Section 3.3.3] shows that there is a constant $c = c(d,s) > 0$ satisfying

$$e_k \geq c \cdot k^{-s/(d-1)} \qquad \forall\, k \in \mathbb{N}.$$

Given a neural network Φ, let us write $d_{\text{in}}(\Phi)$ and $d_{\text{out}}(\Phi)$ for the input- and output-dimension of Φ, respectively. Fix $n \in \mathbb{N}$ with $\varepsilon := \varepsilon_n < 1/2$, and define

$$M_n := \left\{R_\varrho \Phi \,:\, \Phi \text{ is } (\tau,\varepsilon)\text{-quantized NN with } d_{\text{in}}(\Phi) = d, d_{\text{out}}(\Phi) = 1, \text{ and } W(\Phi) \leq C_2 \cdot \varepsilon^{-\sigma}\right\}.$$

Note that $\lceil \log_2(1/\varepsilon)\rceil \leq 1 + \log_2(1/\varepsilon) \leq 2\log_2(1/\varepsilon)$, whence $2^{\tau\lceil \log_2(1/\varepsilon)\rceil} \leq 2^{2\tau\log_2(1/\varepsilon)} = \varepsilon^{-2\tau}$. Furthermore, note for arbitrary $a, b > 0$ that

$$\left|[-a,a] \cap b\mathbb{Z}\right| = \left|[-b^{-1}a, b^{-1}a] \cap \mathbb{Z}\right| \leq 1 + 2b^{-1}a\ ,$$

which shows

$$\left|[-\varepsilon^{-\tau}, \varepsilon^{-\tau}] \cap 2^{-\tau\lceil \log_2(1/\varepsilon)\rceil}\mathbb{Z}\right| \leq 1 + 2\,\varepsilon^{-\tau}\, 2^{\tau\lceil \log_2(1/\varepsilon)\rceil} \leq 1 + 2\varepsilon^{-3\tau} \leq \varepsilon^{-5\tau},$$

and hence $\left|[-\varepsilon^{-\tau}, \varepsilon^{-\tau}] \cap 2^{-\tau\lceil \log_2(1/\varepsilon)\rceil}\mathbb{Z}\right| \leq 2^K$ for $K := \lceil \log_2(\varepsilon^{-5\tau})\rceil \leq 6\tau\log_2(1/\varepsilon)$. Therefore, an application of [45, Lemma B.4] shows that there is a constant $C_3 = C_3(d) \in \mathbb{N}$ satisfying

$$|M_n| \leq 2^{C_3 C_2 \varepsilon^{-\sigma}\cdot(\lceil \log_2(C_2\,\varepsilon^{-\sigma})\rceil + 6\tau\log_2(1/\varepsilon))} \leq 2^{C_4 \varepsilon^{-\sigma}\log_2(1/\varepsilon)},$$

with $C_4 = C_4(C_2, d, \tau, \sigma) > 0$.

By assumption of the theorem to be proven and because of our choice $\varepsilon = \varepsilon_n$, we see with notation as in Proposition 4.1 that M_n is a $C_1\varepsilon$-net (in $L^1((-1,1)^d)$) for $\mathrm{HF}(\mathcal{B}_C)$. Therefore, with $\lambda = \lambda(d,s,C) > 0$ as in Proposition 4.1, there is a $\lambda C_1 \varepsilon$-net $M'_n \subset B^{1,\infty}_0(\Omega)$ for $\mathcal{G} := \{f \in B^{2,2}_s(\Omega) : \|f\|_{B^{2,2}_s} \leq 1\}$ satisfying $|M'_n| \leq |M_n| \leq 2^{k-1}$ for $k := 1 + \lceil C_4 \varepsilon^{-\sigma}\log_2(1/\varepsilon)\rceil$. Defining $M''_n := \{f|_{\Omega_0} : f \in M'_n\} \subset B^{1,\infty}_0(\Omega_0)$, we thus see that M''_n is a $\lambda C_1 \varepsilon$-net for U_A.

Overall, we thus see because of $k \leq C_5 \varepsilon^{-\sigma}\log_2(1/\varepsilon)$ that

$$\lambda\, C_1\, \varepsilon \geq e_k \geq c \cdot k^{-s/(d-1)} \geq c \cdot C_5^{-s/(d-1)} \cdot \varepsilon^{s\sigma/(d-1)} \cdot \big(\log_2(1/\varepsilon)\big)^{-s/(d-1)}.$$

Note that this holds for all $\varepsilon = \varepsilon_n \to 0$ as $n \to \infty$. This is only possible if $s\sigma/(d-1) \geq 1$, meaning $\frac{1}{\sigma} \leq \frac{s}{d-1}$. Since $s > \frac{d+1}{2}$ can be chosen arbitrarily, this implies as claimed that $\frac{1}{\sigma} \leq \frac{1}{2}\frac{d+1}{d-1} = \frac{1}{2} + \frac{1}{d-1}$. □

The strength of the lower bound in Theorem 4.3 is that it applies to networks of *arbitrary depth*; but it requires the neural networks to be quantized. Our final lower bound shows that for neural networks of a fixed maximal depth, one can replace the quantization assumption by a suitable growth condition on the magnitude of the weights.

Theorem 4.4. *Let $d \in \mathbb{N}_{\geq 2}$, $L, N \in \mathbb{N}$, and $\gamma, C, C_1, C_2, C_3 > 0$. Suppose that there is an infinite set $\mathcal{W} \subset \mathbb{N}$ such that for each $W \in \mathcal{W}$ and each $f \in \mathcal{B}_C([-1,1]^{d-1})$ there is a neural network Φ with d-dimensional input and 1-dimensional output and with all weights bounded in absolute value by $C_1 W^N$ such that*

$$\|H_f - R_\varrho \Phi\|_{L^1((-1,1)^d)} \leq C_2 \cdot W^{-\gamma}, \qquad W(\Phi) \leq C_3 \cdot W, \quad \text{and} \quad L(\Phi) \leq L.$$

Then $\gamma \leq \frac{1}{2} + \frac{1}{d-1}$.

Proof. Let $k := \lceil \max\{\gamma^{-1} N + C_1,\ \gamma^{-1} + C_3\} \rceil$ and $m := 3kL$. For $W \in \mathcal{W}$ large enough, we have $\varepsilon := \varepsilon_W := W^{-\gamma} \leq \frac{1}{2}$. For this choice of W and given $f \in \mathcal{B}_C([-1,1]^{d-1})$, let Φ as in the assumption of the theorem. Note that $x \leq 2^x \leq \varepsilon^{-x}$ for all $x \geq 0$, and hence $W(\Phi) \leq C_3 \cdot W = C_3 \cdot \varepsilon^{-1/\gamma} \leq \varepsilon^{-(\gamma^{-1} + C_3)} \leq \varepsilon^{-k}$. Likewise, all weights of Φ are bounded in absolute value by $C_1 W^N = C_1 \varepsilon^{-\frac{N}{\gamma}} \leq \varepsilon^{-(\frac{N}{\gamma} + C_1)} \leq \varepsilon^{-k}$.

Overall, the "quantization lemma" [21, Lemma VI.8] shows that there exists an (m, ε)-quantized network Ψ with d-dimensional input and 1-dimensional output and such that

$$W(\Psi) \leq W(\Phi) \leq C_3 \cdot W = C_3 \cdot \varepsilon^{-1/\gamma} \qquad \text{and} \qquad \|R_\varrho \Phi - R_\varrho \Psi\|_{\sup} \leq \varepsilon,$$

where the $\|\cdot\|_{\sup}$ norm is taken over $(-1,1)^d$. Hence, $\|H_f - R_\varrho \Psi\|_{L^1((-1,1)^d)} \leq (2^d + C_2)\,\varepsilon$. Since $\varepsilon_W = W^{-\gamma} \to 0$ as $W \to \infty$ with $W \in \mathcal{W}$, Theorem 4.3 shows that $\gamma = \frac{1}{1/\gamma} \leq \frac{1}{2} + \frac{1}{d-1}$, as claimed. □

5 Estimation bounds

In this section, we provide error bounds for the performance of empirical risk minimization for learning the indicator function of a set with boundary of Barron class. We also briefly discuss the optimality of these results. More precisely, we show that the best one can hope for is to (roughly) double the "estimation-error rate" that we obtain. We conjecture that neither the lower bound nor the derived rate are optimal, but we were unable to prove this.

In the following theorem, given a subset $\Omega \subset \mathbb{R}^d$, we use the notation

$$\chi_\Omega : \quad \mathbb{R}^d \to \{\pm 1\}, \quad x \mapsto \begin{cases} 1, & \text{if } x \in \Omega, \\ -1, & \text{otherwise.} \end{cases}$$

Moreover, for $\mathcal{A} = (d, N_1, \dots, N_L) \in \mathbb{N}^{L+1}$, we denote by $\mathcal{NN}(\mathcal{A})$ the set of neural networks Φ with input dimension d, L layers, and N_ℓ neurons in the ℓth layer for all $\ell \in \{1, \dots, L\}$. Finally, we define $\operatorname{sign} : \mathbb{R} \to \{\pm 1\}$ by $\operatorname{sign}(x) = 1$ for $x \geq 0$ while $\operatorname{sign}(x) = -1$ if $x < 0$.

Theorem 5.1. *Let $B, C \geq 1$, $M \in \mathbb{N}$, $d \in \mathbb{N}_{\geq 2}$, $\alpha \in (0,1]$, and $m \in \mathbb{N}$. Define*

$$N := \left\lceil \Big((BC)^2 d\, m / \ln(BCMd\, m) \Big)^{1/(1+\alpha)} \right\rceil \in \mathbb{N}$$

and $\mathcal{A} := \big(d,\ M(N + 2d + 2),\ M(4d+2),\ M,\ 1\big)$. Let $\mathbb{P}$ be a tube compatible probability measure on $\mathbb{R}^d$ with parameters C, α, and let $\Omega \in \mathcal{BB}_{B,M}(\mathbb{R}^d)$. Let $S_X = (X_1, \dots, X_m) \overset{\text{iid}}{\sim} \mathbb{P}$ and define $Y_i := \chi_\Omega(X_i)$ for $i \in \underline{m}$.

Then, given $\delta \in (0,1)$, with probability at least $1 - \delta$ regarding the choice of S_X, any

$$\Phi^* \in \operatorname*{argmin}_{\Phi \in \mathcal{NN}(\mathcal{A})} \sum_{i=1}^m \mathbb{1}_{\operatorname{sign}(R_\varrho \Phi(X_i)) \neq Y_i} \tag{5.1}$$

satisfies

$$\mathbb{P}\Big(\operatorname{sign}\big(R_\varrho\Phi^*(X)\big) \neq \chi_\Omega(X)\Big) \leq C_0 \cdot \left(BCM\, d^{3/2} \cdot \Big(\frac{\ln(BCMdm)}{(BC)^2 dm}\Big)^{\gamma/2} + \Big(\frac{\ln(1/\delta)}{m}\Big)^{1/2} \right), \tag{5.2}$$

where $X \sim \mathbb{P}$. Here, $C_0 \geq 1$ is an absolute constant and $\gamma = \frac{\alpha}{1+\alpha}$.

Remark. 1) The set $\big\{\big(\operatorname{sign}(f(X_1)), \dots, \operatorname{sign}(f(X_m))\big) \colon f : \mathbb{R}^d \to \mathbb{R}\big\} \subset \{\pm 1\}^m$ is finite, which implies that a minimizer as in Equation (5.1) always exists.

2) In the common case where $\alpha = 1$ (for instance, if $d\mathbb{P}(x) = \mathbb{1}_{[0,1]^d}(x)\, dx$), we have $\gamma = 1/2$, so that one gets

$$\mathbb{P}\Big(\operatorname{sign}\big(R_\varrho\Phi^*(X)\big) \neq \chi_\Omega(X)\Big) \lesssim \Big(\frac{\ln m}{m}\Big)^{1/4} + \Big(\frac{\ln(1/\delta)}{m}\Big)^{1/2}.$$

Proof. All "implied constants" appearing in this proof are understood to be *absolute* constants.

Define $\Lambda := (BC)^2 dm / \ln(BCMdm)$, so that $N = \lceil \Lambda^{1/(1+\alpha)} \rceil$. If $\Lambda \leq 1$, then the right-hand side of Equation (5.2) is at least 1, so that the estimate is trivial. We can thus assume without loss of generality that $\Lambda > 1$, so that $N \geq 2$ and $N \leq 1 + \Lambda^{1/(1+\alpha)} \leq 2\,\Lambda^{1/(1+\alpha)}$.

Let $\mathcal{H} := \{\operatorname{sign} \circ R_\varrho\Phi \colon \Phi \in \mathcal{NN}(\mathcal{A})\}$. Note that since at most every neuron in layer ℓ can be connected to every neuron in layer $\ell + 1$, the number $W(\mathcal{A})$ of weights of a network with architecture $\mathcal{A}$ satisfies $W(\mathcal{A}) \lesssim M^2 d^2 N$. Therefore, [8, Theorem 2.1] shows that there are absolute constants $C_1, C_2 > 0$ such that

$$\operatorname{VC}(\mathcal{H}) \leq C_1 \cdot M^2 d^2 N \cdot \ln(M^2 d^2 N) \leq C_2 \cdot M^2 d^2 N \cdot \ln(dMN).$$

Next, recall that $\Lambda \geq 1$ and hence $N \leq 2\Lambda^{1/(1+\alpha)} \leq 2\Lambda \lesssim (BC)^2 dm$. Therefore, $\ln(dMN) \lesssim 1 + \ln((BCd)^2 Mm) \lesssim \ln(BCMd\,m)$, which easily implies that

$$\begin{aligned}
\sqrt{\frac{\operatorname{VC}(\mathcal{H})}{m}} &\lesssim m^{-1/2} Md\sqrt{N}\sqrt{\ln(dMN)} \\
&\lesssim \big(\ln(BCdMm)\big)^{\frac{1}{2}(1-\frac{1}{1+\alpha})} \cdot M \cdot d^{\frac{1}{2}(2+\frac{1}{1+\alpha})} \cdot (BC)^{1/(1+\alpha)} \cdot m^{\frac{1}{2}(\frac{1}{1+\alpha}-1)} \\
&= (BC)^{1-\gamma} \cdot M \cdot d^{\frac{3}{2}-\frac{\gamma}{2}} \cdot \big(\ln(BCMd\,m)\big)^{\gamma/2} \cdot m^{-\gamma/2} =: (*).
\end{aligned} \tag{5.3}$$

To make use of this estimate, note that the *Fundamental theorem of statistical learning theory* (see [50, Theorem 6.8 and Definitions 4.1 and 4.3]) shows for arbitrary $\varepsilon, \delta \in (0,1)$ that if we set

$$L_{\mathbb{P}}(h) := \mathbb{P}\big(h(X) \neq \chi_\Omega(X)\big) \quad \text{and} \quad L_S(h) := \frac{1}{m}\sum_{i=1}^m \mathbb{1}_{h(X_i) \neq \chi_\Omega(X_i)},$$

then, with probability at least $1 - \delta$ with respect to the choice of $S = (X_1, \dots, X_m) \overset{iid}{\sim} \mathbb{P}$, we have

$$\forall\, h \in \mathcal{H} : \quad |L_{\mathbb{P}}(h) - L_S(h)| \leq \varepsilon, \tag{5.4}$$

provided that $m \geq C_3 \frac{\operatorname{VC}(\mathcal{H}) + \ln(1/\delta)}{\varepsilon^2}$. Using the estimate $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$, it is easy to see that the condition on m is satisfied if $\varepsilon \geq \sqrt{C_3} \cdot \big(\sqrt{\operatorname{VC}(\mathcal{H})/m} + \sqrt{\ln(1/\delta)/m}\big)$. Finally, thanks to Equation (5.3), we see that there is an absolute constant $C_0 > 0$ (which we can without loss of generality take to satisfy $C_0 \geq 24$) such that this condition holds as soon as

$$\varepsilon \geq \varepsilon_0 := \frac{C_0}{4} \cdot \Big[(*) + \Big(\frac{\ln(1/\delta)}{m}\Big)^{1/2}\Big].$$

This is satisfied if we take ε as one fourth of the right-hand side of Equation (5.2); for this, note that in case of $\varepsilon \geq 1$, Estimate (5.4) is trivially satisfied.

Now, choosing ε to be one fourth of the right-hand side of Equation (5.2), we know that with probability at least $1 - \delta$ with respect to the choice of S, Equation (5.4) holds. Let us assume that $S = (X_1, \dots, X_m)$ is chosen such that this holds. Now, Theorem 3.7 shows that there is $\Phi_0 \in \mathcal{NN}(\mathcal{A})$ such that

$$\mathbb{P}(\{x \in \mathbb{R}^d \colon \mathbb{1}_\Omega(x) \neq R_\varrho \Phi_0(x)\}) \leq 6\, BCM\, d^{3/2}\, N^{-\alpha/2} \leq \frac{C_0}{4} BCM\, d^{3/2}\, \Lambda^{-\gamma/2} \leq \varepsilon.$$

It is not hard to see that there exists $\Phi_1 \in \mathcal{NN}(\mathcal{A})$ satisfying $R_\varrho \Phi_1 = -1 + 2R_\varrho \Phi_0$ and that if $\mathbb{1}_\Omega(x) = R_\varrho \Phi_0(x)$, then $h_1(x) = R_\varrho \Phi_1(x) = \chi_\Omega(x)$ for $h_1 := \operatorname{sign} \circ (R_\varrho \Phi_1) \in \mathcal{H}$. Therefore, $L_{\mathbb{P}}(h_1) = \mathbb{P}(h_1(X) \neq \chi_\Omega(X)) \leq \mathbb{P}(\mathbb{1}_\Omega(X) \neq R_\varrho \Phi_0(X)) \leq \varepsilon$. Overall, if $\Phi^* \in \mathcal{NN}(\mathcal{A})$ satisfies Equation (5.1), and if we set $h^* := \operatorname{sign} \circ R_\varrho \Phi^*$, then Equation (5.4) shows

$$L_{\mathbb{P}}(h^*) \leq L_S(h^*) + \varepsilon \leq L_S(h_1) + \varepsilon \leq L_{\mathbb{P}}(h_1) + 2\varepsilon \leq 3\varepsilon \leq 4\varepsilon = \text{RHS}\,(5.2)\ ,$$

which proves Equation (5.2). □

Remark 5.2 (Quantifying the non-optimality of the learning bound). By taking $\delta \sim m^{-\gamma/2}$, it is not hard to see that the bound in Theorem 5.1 implies that the learning algorithm

$$\big((X_1, \chi_\Omega(X_1)), \dots, (X_m, \chi_\Omega(X_m))\big) \mapsto A_S := \operatorname{sign} \circ R_\varrho \Phi_S^*$$

with Φ_S^* a solution to Equation (5.1) satisfies

$$\mathbb{E}_S\big[\|A_S - \chi_\Omega\|_{L^1(\mathbb{P})}\big] \lesssim \big[\ln(m)/m\big]^{\gamma/2};$$

here, we used that $|A_S - \chi_\Omega| \leq 2 \cdot \mathbb{1}_{A_S \neq \chi_\Omega}$. For the uniform measure $d\mathbb{P} = 2^{-d} \mathbb{1}_{[-1,1]^d}\, dx$, we have $\gamma = 1/2$, and therefore $\mathbb{E}_S\big[\|A_S - \chi_\Omega\|_{L^1([-1,1]^d)}\big] \lesssim \big[\ln(m)/m\big]^{1/4}$. In the remainder of this remark, we sketch an argument showing that no learning algorithm $S \mapsto A_S$ can satisfy

$$\mathbb{E}_S\big[\|A_S - \chi_\Omega\|_{L^1([-1,1]^d)}\big] \lesssim m^{-\theta} \quad \text{with} \quad \theta > \theta^* := \frac{1}{2} \frac{d + 2 + \mathbb{1}_{2\mathbb{Z}+1}(d)}{d - 1}. \tag{5.5}$$

Note that $\theta^* \to \frac{1}{2}$ as $d \to \infty$, which still leaves a gap between this lower bound and the estimation-error rate $m^{-1/4}$ that we obtain.

We expect the lower bound of (5.5) to be suboptimal. One reason why we assume so is that, for a general estimation problem, where the error of estimating a density from m measurements is measured with respect to the Kullback-Leibler divergence, [56, Theorem 1] yields a general lower bound in terms of the metric entropy of the class of densities. As we have seen in the proof of Theorem 4.3, the metric entropy of the set of horizon functions can be lower bounded by using fact that a ball in $B_s^{2,2}$ for $s > (d+1)/2$ embedds into the Fourier-analytic Barron space. By this observation it can be seen using [56, Theorem 1] that a lower bound on the expected error of estimating χ_Ω from m measurements as in Theorem 5.1 measured with respect to the Kullback-Leibler divergence is given by $\mathcal{O}(m^{-(\frac{d+1}{4d} + \delta)})$ for any $\delta > 0$. Note that this rate almost matches the upper bound given in Theorem 5.1 for the L^1 estimation error. The argument in [56] yields bounds for L^2 distances under additional assumptions, see [56, Theorems 4,5,6]. However, none of these assumptions are satisfied in our case.

To prove (5.5), assume by way of contradiction that some learning algorithm $S \mapsto A_S$ satisfies Equation (5.5), uniformly for all $\Omega \in \mathcal{BB}_{1,1}(\mathbb{R}^d)$. Let $Q := (-1,1)^{d-1}$ and $s := 1 + \lfloor \frac{d+1}{2} \rfloor$, as well as $\mathcal{G} := \{f \in W^{s,2}(Q) \colon \|f\|_{W^{s,2}} \leq 1\}$ with the usual Sobolev space $W^{s,2}(Q)$. Since $s > \frac{d+1}{2}$, we see as in the proof of Proposition 4.1 that there is $c > 0$ such that

$$\forall\, f \in \mathcal{G} : \quad \|c\, f\|_{\sup} \leq 1 \quad \text{and} \quad \Omega_f := \big\{(x,t) \in [-1,1]^{d-1} \times [-1,1] \,:\, t \leq c\, f(x)\big\} \in \mathcal{BB}_{1,1}(\mathbb{R}^d).$$

Let $W = (W_1, \dots, W_m) \overset{\text{iid}}{\sim} U([-1,1]^d)$, and write $W_i = (X_i, X_i')$ with $X_i \in [-1,1]^{d-1}$ and $X_i' \in [-1,1]$. Given $f \in \mathcal{G}$, let

$$Y_i := -1 + 2 \cdot \mathbb{1}_{X_i' \leq c f(X_i)} = \chi_{\Omega_f}(W_i),$$

and set $S_f := \big((W_1, Y_1), \dots, (W_m, Y_m)\big)$. By Equation (5.5), there is $C > 0$ independent of m such that

$$\mathbb{E}_W\big[\|A_{S_f} - \chi_{\Omega_f}\|_{L^1([-1,1]^d)}\big] \leq C \cdot m^{-\theta}.$$

Note that S_f is uniquely determined by fixing W and f, and that S_f does not depend fully on f, but only on m point samples of f. Define

$$B_W : \quad [-1,1]^{d-1} \to \mathbb{R}, \quad x \mapsto \frac{1}{c} \cdot \Big(-1 + \int_{-1}^{1} \frac{1 + A_{S_f}(x,t)}{2}\, dt \Big).$$

Note that $B : (W, f) \mapsto B_W$ is a Monte-Carlo algorithm in the sense of [26, Section 2], and for each (random) choice of W, B computes its output based on m point samples of f. To motivate the definition of B_W, note because of $\|c\, f\|_{\sup} \leq 1$ that

$$\int_{-1}^{1} \frac{1 + \chi_{\Omega_f}(x,t)}{2}\, dt = \int_{-1}^{1} \mathbb{1}_{\Omega_f}(x,t)\, dt = \int_{-1}^{1} \mathbb{1}_{t \leq c\, f(x)}\, dt = \int_{-1}^{c\, f(x)} dt = c\, f(x) + 1,$$

and hence $f(x) = \frac{1}{c}\Big(-1 + \int_{-1}^{1} \frac{1 + \chi_{\Omega_f}(x,t)}{2}\, dt \Big)$. This implies

$$\|B_W - f\|_{L^1([-1,1]^{d-1})} \leq \frac{1}{2c} \int_{[-1,1]^{d-1}} \Big| \int_{-1}^{1} A_{S_f}(x,t) - \chi_{\Omega_f}(x,t)\, dt \Big|\, dx \leq \frac{1}{2c} \|A_{S_f} - \chi_{\Omega_f}\|_{L^1},$$

and hence

$$\mathbb{E}_W \|B_W - f\|_{L^1} \leq \frac{1}{2c} \mathbb{E}_W \|A_{S_f} - \chi_{\Omega_f}\|_{L^1} \leq \frac{C}{2c} \cdot m^{-\theta}.$$

Note that this holds for every $f \in \mathcal{G}$ and recall from above that $B : (W, f) \mapsto B_W$ is a Monte-Carlo algorithm that depends on f only through m point samples. However, it is known from information-based complexity (see for instance [26, Theorem 6.1]) that such an error bound for a Monte-Carlo algorithm can only hold if $\theta \leq \frac{s}{d-1} = \frac{1}{2} \frac{d+2+\mathbb{1}_{2\mathbb{Z}+1}(d)}{d-1} = \theta^*$.

6 The case against general measures

In this section, we show that for general probability measures, one cannot derive any non-trivial minimax bound regarding the approximation of sets with Barron class boundary using ReLU neural networks.

The following general result shows that for sets of infinite VC-dimension and general probability measures, no non-trivial minimax approximation results using neural networks can be derived. To conveniently formulate the result, we use the notation

$$\mathcal{NN}_{N,L} := \big\{\Phi \,:\, \Phi \text{ NN with input dimension } d, \text{ with } L(\Phi) \leq L \text{ and } N(\Phi) \leq N\big\}.$$

Furthermore, we continue to write ϱ for the ReLU. The proof of the following lemma is based on (the proof of) the no-free-lunch theorem as presented in [50, Theorem 5.1].

Proposition 6.1. *Let $\Omega \subset \mathbb{R}^d$ be Borel measurable and let $\mathcal{F} \subset \{F : \Omega \to \{0,1\} \colon F \text{ measurable}\}$ such that $\mathrm{VC}(\mathcal{F}) = \infty$.*

Then for arbitrary $N, L \in \mathbb{N}$ we have

$$\sup_{\mu \text{ Borel prob. measure on } \Omega} \sup_{F \in \mathcal{F}} \inf_{\Phi \in \mathcal{NN}_{N,L}} \big\| F - \mathbb{1}_{(0,\infty)} \circ R_\varrho \Phi \big\|_{L^1(\mu)} \geq \frac{1}{16}.$$

Remark 6.2. Even without composing the ReLU neural network $R_\varrho \Phi$ with $\mathbb{1}_{(0,\infty)}$, the above result implies that

$$\sup_{\mu \text{ Borel prob. measure on } \Omega} \sup_{F \in \mathcal{F}} \inf_{\Phi \in \mathcal{NN}_{N,L}} \big\| F - R_\varrho \Phi \big\|_{L^1(\mu)} \geq \frac{1}{32}.$$

This follows by first noting that $\{R_\varrho \Phi \colon \Phi \in \mathcal{NN}_{N,L}\}$ is closed under addition of constant functions and secondly by noting that

$$|y - \mathbb{1}_{(0,\infty)}(z - \tfrac{1}{2})| \leq 2\,|y - z| \qquad \forall\, y \in \{0,1\} \text{ and } z \in \mathbb{R}. \tag{6.1}$$

This estimate is trivial in case of $y = \mathbb{1}_{(0,\infty)}(z - \frac{1}{2})$; thus, let us assume that $y \neq \mathbb{1}_{(0,\infty)}(z - \frac{1}{2})$. Then there are two cases: First, if $z \leq \frac{1}{2}$, then $\mathbb{1}_{(0,\infty)}(z - \frac{1}{2}) = 0$ and $y = 1$, which implies that $2\,|y - z| \geq 2(y - z) \geq 1 = |y - \mathbb{1}_{(0,\infty)}(z - \frac{1}{2})|$. If otherwise $z > \frac{1}{2}$, then $\mathbb{1}_{(0,\infty)}(z - \frac{1}{2}) = 1$ and $y = 0$, so that $2\,|y - z| = 2|z| \geq 1 = |y - \mathbb{1}_{(0,\infty)}(z - \frac{1}{2})|$. This proves (6.1).

Proof of Proposition 6.1. Let $N, L \in \mathbb{N}$ be arbitrary. As shown for instance in [2, Theorem 8.7], if we consider the function class $\mathcal{N} := \{\mathbb{1}_{(0,\infty)} \circ R_\varrho \Phi \colon \Phi \in \mathcal{NN}_{N,L}\}$, then $\mathrm{VC}(\mathcal{N}) < \infty$. By the fundamental theorem of statistical learning theory (see for instance [50, Theorem 6.7]), this means that $\mathcal{N}$ has the *uniform convergence property*, which implies (see [50, Definition 4.3]) that there is some $n \in \mathbb{N}$ such that for each measurable $F : \Omega \to \{0,1\}$ and each Borel probability measure μ on Ω, if we choose $S_X = (X_1, \ldots, X_n) \overset{\text{i.i.d.}}{\sim} \mu$, then with probability at least $1 - \frac{1}{10}$ with respect to the choice of S_X, we have

$$\sup_{\phi \in \mathcal{N}} \big| R_{\mu,F}(\phi) - R_{S_X,F}(\phi) \big| \leq \frac{1}{32}, \tag{6.2}$$

where

$$R_{\mu,F}(\phi) = \mu\big(\{x \in \Omega \colon \phi(x) \neq F(x)\}\big) \qquad \text{and} \qquad R_{S_X,F}(\phi) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\phi(X_i) \neq F(X_i)}.$$

Note $|F - \phi| \in \{0,1\}$, whence $R_{\mu,F}(\phi) = \|F - \phi\|_{L^1(\mu)}$ and $R_{S_X,F}(\phi) = \frac{1}{n} \sum_{i=1}^{n} |\phi(X_i) - F(X_i)|$.

Since $\mathrm{VC}(\mathcal{F}) = \infty$, there is a set $\Omega_0 \subset \Omega$ of cardinality $|\Omega_0| = 2n$ such that Ω_0 is shattered by $\mathcal{F}$, meaning that if we set $\mathcal{G} := \big\{g : \Omega_0 \to \{0,1\}\big\}$, then $\mathcal{G} = \{f|_{\Omega_0} \colon f \in \mathcal{F}\}$. Let $\mu := U(\Omega_0)$ denote the uniform distribution on Ω_0, meaning $\mu(\{x\}) = 1/|\Omega_0|$ for all $x \in \Omega_0$, and assume towards a contradiction that

$$\sup_{F \in \mathcal{F}} \inf_{\phi \in \mathcal{N}} \|F - \phi\|_{L^1(\mu)} < \frac{1}{16}. \tag{6.3}$$

Now, given any $S = \big((X_i, Y_i)\big)_{i=1,\ldots,n} \in \big(\Omega_0 \times \{0,1\}\big)^n$, choose $\phi_S \in \mathcal{N}$ satisfying

$$\phi_S \in \operatorname*{argmin}_{\phi \in \mathcal{N}} \sum_{i=1}^{n} |\phi(X_i) - Y_i|. \tag{6.4}$$

Such a function ϕ_S exists, since the expression $\sum_{i=1}^N |\phi(X_i) - Y_i|$ only depends on $\phi|_{\Omega_0}$, while $\{\phi|_{\Omega_0} \colon \phi \in \mathcal{N}\} \subset \{0,1\}^{\Omega_0}$ is a finite set. Here, $\{0,1\}^{\Omega_0} = \{\psi : \Omega_0 \to \{0,1\}\}$ is the set of all functions from Ω_0 to $\{0,1\}$.

For $S_X = (X_1, \dots, X_n) \in \Omega_0^n$ and $g \in \mathcal{G}$, let us define $S_X(g) := \big((X_i, g(X_i))\big)_{i=1,\dots,n}$. Now, given an arbitrary $g \in \mathcal{G}$, recall from above that $g = F|_{\Omega_0}$ for some $F \in \mathcal{F}$. Thanks to (6.3), there is thus some $\phi^* \in \mathcal{N}$ satisfying $\|g - \phi^*\|_{L^1(\mu)} = \|F - \phi^*\|_{L^1(\mu)} < \frac{1}{16}$. Overall, we thus see that with probability at least $1 - \frac{1}{10}$ with respect to the choice of $S_X = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} \mu$, we have

$$\begin{aligned}
\|g - \phi_{S_X(g)}\|_{L^1(\mu)} = \|F - \phi_{S_X(F)}\|_{L^1(\mu)} &= R_{\mu,F}\big(\phi_{S_X(F)}\big) \\
&\overset{(6.2)}{\leq} \frac{1}{32} + R_{S_X,F}\big(\phi_{S_X(F)}\big) \\
&\overset{(6.4)}{\leq} \frac{1}{32} + R_{S_X,F}\big(\phi^*\big) \\
&\overset{(6.2)}{\leq} \frac{1}{16} + R_{\mu,F}\big(\phi^*\big) = \frac{1}{16} + \|F - \phi^*\|_{L^1(\mu)} \\
&\overset{(6.3)}{<} \frac{1}{8}.
\end{aligned}$$

Since $|g - \phi_{S_X(g)}| \leq 1$, we thus see for every $g \in \mathcal{G}$ that

$$\mathbb{E}_{S_X}\, \|g - \phi_{S_X(g)}\| \leq \frac{1}{10} + \frac{1}{8} < \frac{1}{4} \quad \text{and hence} \quad \mathbb{E}_{S_X}\Big[\frac{1}{|\mathcal{G}|}\sum_{g \in \mathcal{G}} \|g - \phi_{S_X(g)}\|_{L^1(\mu)}\Big] < \frac{1}{4}.$$

In the last part of the proof, we will show that this is impossible, by showing for *every* $S_X = (X_1, \dots, X_n) \in \Omega_0^n$ that $\frac{1}{|\mathcal{G}|}\sum_{g \in \mathcal{G}} \|g - \phi_{S_X(g)}\|_{L^1(\mu)} \geq \frac{1}{4}$.

Thus, let $S_X = (X_1, \dots, X_n) \in \Omega_0^n$ be fixed, and set $\Omega_1 := \{X_1, \dots, X_n\}$, noting that $|\Omega_0 \setminus \Omega_1| \geq n$. Given $g \in \mathcal{G}$ and $x \in \Omega_0$, define

$$g^{(x)} \colon \quad \Omega_0 \to \{0,1\}, \quad y \mapsto \begin{cases} g(y), & \text{if } y \neq x, \\ 1 - g(x), & \text{otherwise.} \end{cases}$$

It is easy to see that $\mathcal{G} \to \mathcal{G}, g \mapsto g^{(x)}$ is bijective, since $(g^{(x)})^{(x)} = g$. Furthermore, given any $x \in \Omega_0 \setminus \Omega_1$, note that $S_X(g) = S_X(g^{(x)})$, so that

$$\big|g(x) - \phi_{S_X(g)}(x)\big| + \big|g^{(x)}(x) - \phi_{S_X(g^{(x)})}(x)\big| = \big|g(x) - \phi_{S_X(g)}(x)\big| + \big|g^{(x)}(x) - \phi_{S_X(g)}(x)\big| = 1.$$

Overall, we thus see

$$\begin{aligned}
\frac{1}{|\mathcal{G}|}\sum_{g \in \mathcal{G}} \|g - \phi_{S_X(g)}\|_{L^1(\mu)} &\geq \frac{1}{2n}\frac{1}{|\mathcal{G}|} \sum_{x \in \Omega_0 \setminus \Omega_1} \sum_{g \in \mathcal{G}} |g(x) - \phi_{S_X(g)}(x)| \\
&\geq \frac{1}{2n}\frac{1}{2|\mathcal{G}|} \sum_{x \in \Omega_0 \setminus \Omega_1} \sum_{g \in \mathcal{G}} \big[|g(x) - \phi_{S_X(g)}(x)| + |g^{(x)}(x) - \phi_{S_X(g^{(x)})}(x)|\big] \\
&= \frac{|\Omega_0 \setminus \Omega_1|}{2n} \cdot \frac{|\mathcal{G}|}{2|\mathcal{G}|} \geq \frac{1}{4},
\end{aligned}$$

as claimed. This completes the proof. □

In Proposition 6.1, the measure μ might depend on the choice of $N, L \in \mathbb{N}$. The next result shows that even if one restricts to a *fixed* measure μ for all $N, L \in \mathbb{N}$, the approximation rate can get arbitrarily bad.

Proposition 6.3. *Let $\Omega \subset \mathbb{R}^d$ be Borel measurable and let $\mathcal{F} \subset \{F : \Omega \to \{0,1\} : F \text{ measurable}\}$ such that $\mathrm{VC}(\mathcal{F}) = \infty$.*

Then for each null-sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ and arbitrary sequences $(N_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ and $(L_n)_{n\in\mathbb{N}} \subset \mathbb{N}$, there is a Borel probability measure μ on Ω and some $n_0 \in \mathbb{N}$ such that

$$\sup_{F\in\mathcal{F}} \inf_{\Phi\in\mathcal{NN}_{N_n,L_n}} \|F - \mathbb{1}_{(0,\infty)} \circ R_\varrho\Phi\|_{L^1(\mu)} \geq \varepsilon_n \qquad \forall\, n \in \mathbb{N}_{\geq n_0}.$$

Proof. Define $\tau_n := \sup_{k\geq n} \varepsilon_k$, as well as $N'_n := \max\{N_1, \dots, N_n\}$ and $L'_n := \max\{L_1, \dots, L_n\}$ for $n \in \mathbb{N}$. Note that $(\tau_n)_{n\in\mathbb{N}}$ is a non-increasing null-sequence; in particular, $\tau_n \geq 0$ for all $n \in \mathbb{N}$. Choose a strictly increasing sequence $(n_\ell)_{\ell\in\mathbb{N}} \subset \mathbb{N}$ satisfying $\tau_{n_\ell} \leq 2^{-5-\ell}$, so that $\kappa := \sum_{\ell=1}^\infty \tau_{n_\ell} \leq 2^{-5} \sum_{\ell=1}^\infty 2^{-\ell} = \frac{1}{32}$. Now, Proposition 6.1 yields for each $\ell \in \mathbb{N}$ a Borel probability measure μ_ℓ and some $F_\ell \in \mathcal{F}$ satisfying $\inf_{\Phi\in\mathcal{NN}_{N'_{n_\ell},L'_{n_\ell}}} \|F_\ell - \mathbb{1}_{(0,\infty)} \circ R_\varrho\Phi\|_{L^1(\mu_\ell)} \geq \frac{1}{32}$. Fix some $\omega_0 \in \Omega$ and define $\mu := 32 \sum_{\ell=1}^\infty \tau_{n_\ell}\, \mu_{\ell+1} + (1 - 32\kappa)\delta_{\omega_0}$, so that μ is a Borel probability measure on Ω.

Now, given any $n \in \mathbb{N}_{\geq n_1}$, let $\ell \in \mathbb{N}$ with $n_\ell \leq n < n_{\ell+1}$, so that $\tau_{n_\ell} = \sup_{k\geq n_\ell} \varepsilon_k \geq \varepsilon_n$ and $N_n \leq N'_n \leq N'_{n_{\ell+1}}$ as well as $L_n \leq L'_n \leq L'_{n_{\ell+1}}$. Therefore,

$$\begin{aligned}
&\sup_{F\in\mathcal{F}} \inf_{\Phi\in\mathcal{NN}_{N_n,L_n}} \big\|F - \mathbb{1}_{(0,\infty)} \circ R_\varrho\Phi\big\|_{L^1(\mu)} \\
&\geq 32\,\tau_{n_\ell} \cdot \inf_{\Phi\in\mathcal{NN}_{N'_{n_{\ell+1}},L'_{n_{\ell+1}}}} \|F_{\ell+1} - \mathbb{1}_{(0,\infty)} \circ R_\varrho\Phi\|_{L^1(\mu_{\ell+1})} \geq \tau_{n_\ell} \geq \varepsilon_n.
\end{aligned}$$

Since $n \in \mathbb{N}_{\geq n_1}$ was arbitrary, we are done. □

Finally, we show that the class of Barron horizon functions (and thus also the class of sets with boundary of Barron class) has infinite VC dimension, so that the previous results apply in this setting.

Lemma 6.4. *Let $d \geq 2$ and $Q = [-1,1]^d$, as well as $C > 0$ and $M \in \mathbb{N}$. Then*

$$\mathrm{VC}\big(\mathcal{BB}_{C,M}(\mathbb{R}^d)\big) \geq \mathrm{VC}\big(\mathcal{BH}_C(Q)\big) = \infty.$$

Proof. Let $n \in \mathbb{N}$ be arbitrary. For each $k \in \underline{n}$, choose $\varphi_n^{(k)} \in C_c^\infty\big((\frac{k-1}{n}, \frac{k}{n}) \times (-1,1)^{d-2}\big)$ satisfying $\varphi_n^{(k)} \geq 0$ and $\varphi_n^{(k)}(\frac{k-1}{n} + \frac{1}{2n}, 0, \dots, 0) = 1$. Define $X := [-1,1]^{d-1}$ and use Remark 3.2 to select $C' > 0$ satisfying $\mathcal{B}_{C'}(X,0) \subset \mathcal{BA}_C(X)$. It is easy to see that there is some $\tau_n > 0$ satisfying $\tau_n\, \varphi_n^{(k)} \in \mathcal{B}_{C'/n}(X)$. Now, given $\theta = (\theta_1, \dots, \theta_n) \in \{0,1\}^n$, define

$$f_n^{(\theta)} := \tau_n \sum_{k=1}^n (2\theta_k - 1)\varphi_n^{(k)} \in \mathcal{B}_{C'}(X,0) \subset \mathcal{BA}_C(X).$$

This implies $H_n^{(\theta)} \in \mathcal{BH}_C(Q)$, where $H_n^{(\theta)}(x) := \mathbb{1}_{f_n^{(\theta)}(x_1,\dots,x_{d-1}) \geq x_d}$. Furthermore, in view of $f_n^{(\theta)}(\frac{k-1}{n} + \frac{1}{2n}, 0, \dots, 0) = (2\theta_k - 1)\tau_n$, we see that

$$H_n^{(\theta)}\big(\tfrac{k-1}{n} + \tfrac{1}{2n}, 0, \dots, 0\big) = \mathbb{1}_{(2\theta_k-1)\tau_n \geq 0} = \mathbb{1}_{2\theta_k \geq 1} = \theta_k.$$

Therefore, $\mathcal{BH}_C(Q)$ shatters the set $\big\{\big(\frac{k-1}{n} + \frac{1}{2n}, 0 \dots, 0\big) : k \in \underline{n}\big\} \subset Q$, which shows that $\mathrm{VC}\big(\mathcal{BH}_C(Q)\big) \geq n$. Since this holds for every $n \in \mathbb{N}$, we are done. □

7 Three kinds of Barron spaces

In the literature (see for instance [16, 18, 38]), there are at least three different function spaces that are referred to as *Barron spaces*. In the terminology that we used in the introduction, these are the *Fourier-analytic Barron space* and the *infinite-width Barron spaces,* either using the ReLU or the Heaviside activation function. In the current literature, the relationship between these spaces has only been understood partially. Therefore, we clarify this issue in this section.

To fix the terminology, let us write $\mathcal{P}_d$ for the set of all Borel probability measures on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$. Given a (measurable) function $\phi : \mathbb{R} \to \mathbb{R}$ and $\mu \in \mathcal{P}_d$, we write

$$\mu_\phi(x) := \int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} a \cdot \phi\big(\langle w, x\rangle + c\big)\, d\mu(a, w, c) \quad \text{for } x \in \mathbb{R}^d,$$

whenever the integral exists. Let us denote the Heaviside function by $H := \mathbb{1}_{[0,\infty)}$ and the ReLU by $\varrho : \mathbb{R} \to \mathbb{R}, x \mapsto \max\{0, x\}$. Then, given a set $\varnothing \neq U \subset \mathbb{R}^d$ and $s \geq 0$, we define

$$\begin{aligned}
\mathcal{B}_H(U) &:= \Big\{ f : U \to \mathbb{R} \,:\, \exists\, \mu \in \mathcal{P}_d :\, \|\mu\|_H < \infty \text{ and } \forall\, x \in U : f(x) = \mu_H(x) \Big\}, \\
\mathcal{B}_\varrho(U) &:= \Big\{ f : U \to \mathbb{R} \,:\, \exists\, \mu \in \mathcal{P}_d :\, \|\mu\|_\varrho < \infty \text{ and } \forall\, x \in U : f(x) = \mu_\varrho(x) \Big\}, \\
\mathcal{B}_{\mathcal{F},s}(U) &:= \Big\{ f : U \to \mathbb{R} \,:\, \exists\, F : \mathbb{R}^d \to \mathbb{C} : \|F\|_{\mathcal{F},s} < \infty \text{ and } \forall\, x \in U \colon f(x) = \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} F(\xi)\, d\xi \Big\},
\end{aligned}$$

where

$$\|\mu\|_H := \int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} |a|\, d\mu(a, w, c) \quad \text{and} \quad \|\mu\|_\varrho := \int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} |a| \cdot (|w| + |c|)\, d\mu(a, w, c),$$

while $\|F\|_{\mathcal{F},s} := \int_{\mathbb{R}^d} (1 + |\xi|)^s\, |F(\xi)|\, d\xi$. Finally, the norms on these spaces are given by

$$\|f\|_{\mathcal{B}_H} := \inf\big\{ \|\mu\|_H \quad : \quad \mu \in \mathcal{P}_d \text{ and } f = \mu_H|_U \big\}$$

and similarly for $\|f\|_{\mathcal{B}_\varrho}$, while

$$\|f\|_{\mathcal{B}_{\mathcal{F},s}} := \inf\Big\{ \|F\|_{\mathcal{F},s} \colon F : \mathbb{R}^d \to \mathbb{C} \text{ measurable and } f(x) = \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} F(\xi)\, d\xi \text{ for all } x \in U \Big\}.$$

From the literature, the following properties of these spaces are known.

Lemma 7.1. *Let $\varnothing \neq U \subset \mathbb{R}^d$ be bounded. Then the following hold:*

1) $\mathcal{B}_\varrho(U) \hookrightarrow \mathcal{B}_H(U)$. If U has nonempty interior, then the inclusion is strict.

2) $\mathcal{B}_{\mathcal{F},1}(U) \hookrightarrow \mathcal{B}_H(U)$.

3) $\mathcal{B}_{\mathcal{F},2}(U) \hookrightarrow \mathcal{B}_\varrho(U)$.

Remark. Regarding part *1)*, an easy modification of the proof shows that it would in fact suffice for U to satisfy $\{tx + (1-t)y \colon t \in [0,1]\} \subset U$ for certain $x \neq y$, even if U is not open.

Proof. **Ad 1):** Every function in $\mathcal{B}_\varrho(U)$ is Lipschitz continuous; see [18, Theorem 3.3]. On the other hand, choosing μ to be a Dirac measure, we see $H_{w,c} = \big(x \mapsto H(\langle w, x\rangle + c)\big) \in \mathcal{B}_H(U)$ for arbitrary $w \in \mathbb{R}^d$ and $c \in \mathbb{R}$. If U has nonempty interior, one can choose w, c in such a way that $H_{w,c}$ is discontinuous on U, and therefore cannot belong to $\mathcal{B}_\varrho(U)$. This shows that the inclusion has to be strict if U has nonempty interior.

The inclusion $\mathcal{B}_\varrho(U) \subset \mathcal{B}_H(U)$ is probably folklore; since we could not locate a reference, however, we provide the proof. Since U is bounded, we have $U \subset \overline{B_R}(0)$ for a suitable $R > 0$. Set $C := 1 + R$ and note that

$$\varrho(y) = \int_0^C H(y-t)\,dt \qquad \forall\, y \in \mathbb{R} \text{ with } |y| \leq R.$$

Now, given $w \in \mathbb{R}^d$ and $c \in \mathbb{R}$, define $\theta_{w,c} := |w| + |c|$ and note note $|\langle w, x\rangle + c| \leq C \cdot \theta_{w,c}$ for all $x \in U$. Recall that $\varrho(\gamma x) = \gamma\,\varrho(x)$ for $\gamma \geq 0$ and $x \in \mathbb{R}$. Therefore, given a measure $\mu \in \mathcal{P}_d$, and setting $\Omega := \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$, we see for all $x \in U$ that

$$\begin{aligned}
\mu_\varrho(x) &= \int_\Omega a \cdot \varrho(\langle w, x\rangle + c)\,d\mu(a,w,c) = \int_\Omega a\,\theta_{w,c} \cdot \varrho(\langle \tfrac{w}{\theta_{w,c}}, x\rangle + \tfrac{c}{\theta_{w,c}})\,d\mu(a,w,c) \\
&= \int_\Omega \int_0^C a\,\theta_{w,c} \cdot H(\langle \tfrac{w}{\theta_{w,c}}, x\rangle + \tfrac{c}{\theta_{w,c}} - t)\,dt\,d\mu(a,w,c) \\
&= \int_\Omega \alpha \cdot H(\langle \omega, x\rangle + s)\,d\nu(\alpha,\omega,s) = \nu_H(x),
\end{aligned}$$

where $\nu := \Theta^{-1}(\mu \otimes \lambda)$ is the pushforward of the product measure $\mu \otimes \lambda$ (with λ denoting the Lebesgue measure on $[0,C]$) under the map

$$\Theta: \quad \Omega \times [0,C] \to \Omega, \quad \big((a,w,c),t\big) \mapsto \big(a \cdot \theta_{w,c},\ \tfrac{w}{\theta_{w,c}},\ \tfrac{c}{\theta_{w,c}} - t\big).$$

Finally, note that

$$\|\nu\|_H = \int_\Omega |\alpha|\,d\nu(\alpha,\omega,s) = \int_\Omega \int_0^C |a\cdot\theta_{w,c}|\,dt\,d\mu(a,w,c) \leq C \int_\Omega |a|\,(|w|+|c|)\,d\mu(a,w,c) = C\,\|\mu\|_\varrho.$$

This easily shows that $\|f\|_{\mathcal{B}_H(U)} \leq C \cdot \|f\|_{\mathcal{B}_\varrho(U)} < \infty$ for all $f \in \mathcal{B}_\varrho(U)$.

Ad 2): This follows from [3, Theorem 2].

Ad 3): This essentially follows from [16, Theorem 9], which is itself a consequence of (the proof of) [31, Theorem 6].

More precisely, since $U \subset \mathbb{R}^d$ is bounded, we can choose $x_0 \in \mathbb{R}^d$ and $R \geq 1$ such that $U \subset x_0 + [0,R]^d$. Let $f \in \mathcal{B}_{\mathcal{F},2}(U)$ with $\|f\|_{\mathcal{B}_{\mathcal{F},2}} \leq 1$. This implies $f(x) = \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} F(\xi)\,d\xi$ for $x \in U$, where $\|F\|_{\mathcal{F},2} \leq 2$. Define $G, H : \mathbb{R}^d \to \mathbb{C}$ by $G(\xi) = \frac{1}{2}\big(F(\xi) + \overline{F(-\xi)}\big)$ and $H(\xi) = R^{-d} \cdot e^{i\langle \frac{x_0}{R}, \xi\rangle} \cdot G(\xi/R)$. A direct calculation shows $\|G\|_{\mathcal{F},2} \leq 2$ and $\|H\|_{\mathcal{F},2} \leq 2R^2$. Next, define $g, h : \mathbb{R}^d \to \mathbb{R}$ by $g(x) := \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} G(\xi)\,d\xi$ and $h(x) := \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} H(\xi)\,d\xi$. It is straightforward to verify $h(\frac{x-x_0}{R}) = g(x) = f(x)$ for $x \in U$.

By elementary properties of the Fourier transform, we see $\int_{\mathbb{R}^d} |\xi|^2\,|\widehat{h}(\xi)|\,d\xi \leq C$ and $h \in C^1$ with $\|h\|_{\sup}, \|\nabla h\|_{\sup} \leq C$ where $C = C(d,R)$. Thanks to [16, Theorem 9], this implies $\|h\|_{\mathcal{B}_\varrho([0,1]^d)} \leq C' < \infty$. Therefore, $h(y) = \int_\Omega a\,\varrho(c + \langle w, x\rangle)\,d\mu(a,w,c)$ for all $y \in [0,1]^d$, where $\mu \in \mathcal{P}_d$ satisfies $\|\mu\|_\varrho \leq 2C'$. Because of $y = \frac{x-x_0}{R} \in [0,1]^d$ for $x \in U$, this implies $f(x) = h(\frac{x-x_0}{R}) = \int_\Omega a\,\varrho\big(\langle \frac{w}{R}, x\rangle + c - \frac{\langle w,x_0\rangle}{R}\big)\,d\mu(a,w,c) = \nu_\varrho(x)$, where $\nu = \Psi^*\mu$ is the pushforward of μ under the map $\Psi : \Omega \to \Omega, (a,w,c) \mapsto \big(a, \frac{w}{R}, c - \frac{\langle w,x_0\rangle}{R}\big)$. A direct calculation shows $\|\nu\|_\varrho \leq (1+|x_0|)\|\mu\|_\varrho \leq C''$ for $C'' = C''(d,R,x_0)$. Hence, $f \in \mathcal{B}_\varrho(U)$ with $\|f\|_{\mathcal{B}_\varrho(U)} \leq C''$. □

The previous lemma collected several relations between the different Barron-type spaces from the literature. The question of how the spaces $\mathcal{B}_\varrho$ and $\mathcal{B}_{\mathcal{F},1}$ are related, however, has,

to the best of our knowledge, not been answered until now. While it is claimed in [18, Theorem 3.1] that $\mathcal{B}_{\mathcal{F},1}$ embeds continuously into $\mathcal{B}_\varrho$, citing [4] as a reference, we believe that this mischaracterizes the results of [4]. In fact, in [4] (or rather [3]), it is merely shown that $\mathcal{B}_{\mathcal{F},1}$ embeds into $\mathcal{B}_H$, not $\mathcal{B}_\varrho$. As we will see in Proposition 7.4 below, we actually have $\mathcal{B}_{\mathcal{F},1} \not\subseteq \mathcal{B}_\varrho$. The proof will be based on the following lemma, which shows that the partial derivatives of functions in $\mathcal{B}_\varrho$ are "uniformly of bounded variation along the coordinate axes". This lemma is similar in spirit to [18, Example 4.1], which essentially corresponds to the one-dimensional case of the result given here. In the following lemma, we use for a Lipschitz continuous function $g : \mathbb{R}^d \to \mathbb{R}$, $i, j \in \underline{d}$, and $x \in \mathbb{R}^d$, the following functions

$$g_{j,i,x} : \quad \mathbb{R} \to \mathbb{R}, \quad t \mapsto (\partial_j g)(x + te_i), \tag{7.1}$$

where $(e_1, \dots, e_d)$ denotes the standard basis of $\mathbb{R}^d$.

Lemma 7.2. *Let $\varnothing \neq U \subset \mathbb{R}^d$ be bounded. For every $f \in \mathcal{B}_\varrho(U)$, there exists a Lipschitz continuous function $g : \mathbb{R}^d \to \mathbb{R}$ satisfying $f = g|_U$ and*

$$\sup_{i,j \in \underline{d}, x \in \mathbb{R}^d} \|g_{j,i,x}\|_{BV} \leq 4 \, \|f\|_{\mathcal{B}_\varrho}, \tag{7.2}$$

where we write $\|h\|_{BV} := \|h\|_{\sup} + \mathrm{TV}(h)$ for $h : \mathbb{R} \to \mathbb{R}$, with $\mathrm{TV}(h)$ denoting the total variation of h; see for instance [23, Chapter 3.5] for the definition.

Remark. The partial derivative $\partial_j g$ appearing in Equation (7.1) above is the weak derivative of g, and thus a priori only uniquely defined up to changes on a null-set. What is meant is that there is a version of this derivative such that $g_{j,i,x}$ is of bounded variation for all $i, j \in \underline{d}$ and $x \in \mathbb{R}^d$, and such that Equation (7.2) holds.

Proof. The claim is clear in case of $\|f\|_{\mathcal{B}_\varrho} = 0$; thus, let us assume that $\|f\|_{\mathcal{B}_\varrho} > 0$. By definition of $\mathcal{B}_\varrho(U)$ there is a probability measure $\mu \in \mathcal{P}_d$ satisfying $\|\mu\|_\varrho \leq \frac{5}{4} \|f\|_{\mathcal{B}_\varrho}$ and $f = \mu_\varrho|_U$. Define $\Omega := \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ and $g := \mu_\varrho$, noting that $g : \mathbb{R}^d \to \mathbb{R}$ is well-defined, since

$$\int_\Omega |a \, \varrho(\langle w, x \rangle + c)| \, d\mu(a, w, c) \leq (1 + |x|) \int_\Omega |a| \cdot (|w| + |c|) \, d\mu(a, w, c) < \infty$$

for each $x \in \mathbb{R}^d$. Furthermore, since ϱ is 1-Lipschitz, we see for all $x, y \in \mathbb{R}^d$ that

$$|g(x) - g(y)| \leq \int_\Omega |a| \cdot |\langle w, x - y \rangle| \, d\mu(a, w, c) \leq |x - y| \int_\Omega |a| \cdot (|w| + |c|) \, d\mu(a, w, c) \leq |x - y| \cdot \|\mu\|_\varrho,$$

meaning that g is Lipschitz continuous.

Now, note that $x \mapsto \varrho(\langle w, x \rangle + c)$ either vanishes identically (in case of $w = c = 0$) or otherwise is differentiable on $\{x \in \mathbb{R}^d : \langle w, x \rangle + c \neq 0\}$ which is open and of full measure, with partial derivatives $\partial_j[\varrho(\langle w, x \rangle + c)] = w_j H(\langle w, x \rangle + c)$. Furthermore, $x \mapsto \varrho(\langle w, x \rangle + c)$ is Lipschitz continuous and hence weakly differentiable, and the weak derivative coincides almost everywhere with the classical derivative; see for instance [22, Theorems 4 and 5 in Section 5.8]. Therefore, for any $\varphi \in C_c^\infty(U)$ and $j \in \underline{d}$, Fubini's theorem shows that

$$\begin{aligned} \int_U f(x) \, \partial_j \varphi(x) \, dx &= \int_\Omega a \int_U \varrho(\langle w, x \rangle + c) \partial_j \varphi(x) \, dx \, d\mu(a, w, c) \\ &= - \int_\Omega \int_U a \, w_j \, H(\langle w, x \rangle + c) \varphi(x) \, dx \, d\mu(a, w, c) = - \int_U \varphi(x) g_j(x) \, dx, \end{aligned}$$

meaning that $g_j = \partial_j g$ for

$$g_j : \quad \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto \int_U a\, w_j\, H(\langle w, x\rangle + c)\, d\mu(a, w, c).$$

Now, using the convention $\operatorname{sign}(x) = 1$ for $x \geq 0$ and $\operatorname{sign}(x) = -1$ if $x < 0$, given $i, j \in \underline{d}$, define

$$M_{\alpha,\beta} := \big\{(a, w, c) \in \Omega : \ \operatorname{sign}(aw_j) = \alpha \text{ and } \operatorname{sign}(w_i) = \beta\big\} \quad \text{for} \quad \alpha, \beta \in \{\pm 1\}.$$

Since the Heaviside function H is non-decreasing, it is then straightforward to see for each $x \in \mathbb{R}^d$ that each of the functions $F_{\alpha,\beta,x} : \mathbb{R} \to \mathbb{R}, t \mapsto \int_{M_{\alpha,\beta}} a\, w_j\, H(w_i\, t + \langle w, x\rangle + c)\, d\mu(a, w, c)$ is monotonic and $g_{j,i,x} = \sum_{\alpha,\beta\in\{\pm1\}} F_{\alpha,\beta,x}$. Furthermore, each of the $F_{\alpha,\beta,x}$ is bounded; precisely,

$$|F_{\alpha,\beta,x}(t)| \leq \int_{M_{\alpha,\beta}} |a|\, |w_j|\, d\mu(a, w, c) \leq \int_{M_{\alpha,\beta}} |a| \cdot (|w| + |c|)\, d\mu(a, w, c),$$

so that $\sum_{\alpha,\beta\in\{\pm1\}} \|F_{\alpha,\beta,x}\|_{\sup} \leq \|\mu\|_{\varrho}$. It is easy to see (see [23, Section 3.5]) that every monotonic function $h : \mathbb{R} \to \mathbb{R}$ satisfies $\|h\|_{\mathrm{BV}} \leq 3\|h\|_{\sup}$. Therefore, $g_{j,i,x}$ is of bounded variation with

$$\|g_{j,i,x}\|_{\mathrm{BV}} \leq \sum_{\alpha,\beta\in\{\pm1\}} \|F_{\alpha,\beta,x}\|_{\mathrm{BV}} \leq 3\|\mu\|_{\varrho} \leq \frac{15}{4}\|f\|_{\mathcal{B}_\varrho},$$

which easily implies the claim. □

We will also need the following technical lemma. It is a well-known property of BV functions; see for instance the proof of [1, E6.10]. For the sake of completeness and for readers unfamiliar with functions of bounded variation, we provide a proof in Appendix B.

Lemma 7.3. *Let $g : \mathbb{R} \to \mathbb{R}$ be bounded and of bounded variation. Then, for arbitrary $\varphi \in C_c^\infty(\mathbb{R})$, we have $|\int_{\mathbb{R}} \varphi'(t)\, g(t) dt| \leq \|\varphi\|_{\sup}\, \mathrm{TV}(g)$.*

With these preparations, we can finally show that for most domains U, we have that $\mathcal{B}_{\mathcal{F},1}(U)$ is not contained in $\mathcal{B}_\varrho(U)$.

Proposition 7.4. *Let $U \subset \mathbb{R}^d$ have nonempty interior and let $\alpha \geq 0$. If $\mathcal{B}_{\mathcal{F},\alpha}(U) \subseteq \mathcal{B}_\varrho(U)$, then $\alpha \geq 2$. In particular, $\mathcal{B}_{\mathcal{F},1}(U) \nsubseteq \mathcal{B}_\varrho(U)$.*

Proof. The proof is divided into three steps:

Step 1 *(Setup of Banach spaces X, Y'):* We define $\beta := \max\{1, \alpha\}$ and

$$X := \big\{\mathcal{F}^{-1} f : \ f \in L^1_{(1+|\xi|)^\beta}(\mathbb{R}^d; \mathbb{C}) \text{ and } \mathcal{F}^{-1} f \text{ is real-valued}\big\}$$

with norm $\|\mathcal{F}^{-1} f\|_X := \|f\|_{L^1_{(1+|\xi|)^\beta}} = \int_{\mathbb{R}^d} (1 + |\xi|)^\beta\, |f(\xi)|\, d\xi$, which is well-defined since the Fourier transform is injective on $L^1(\mathbb{R}^d) \supset L^1_{(1+|\xi|)^\beta}(\mathbb{R}^d)$. It is straightforward to verify that X is a Banach space (with $\mathbb{R}$ as the scalar field, since we require $\mathcal{F}^{-1} f$ to be real-valued for $f \in X$), and by differentiation under the integral it is easy to see that $X \hookrightarrow C_b^1(\mathbb{R}^d)$, where $C_b^1(\mathbb{R}^d) = \{g \in C^1(\mathbb{R}^d; \mathbb{R}) \colon \|g\|_{C_b^1} < \infty\}$ and $\|g\|_{C_b^1} := \|g\|_{\sup} + \sum_{j=1}^d \|\partial_j g\|_{\sup}$.

Since U has nonempty interior, we have $U_0 := x_0 + (-3\varepsilon, 3\varepsilon)^d \subset U$ for certain $x_0 \in \mathbb{R}^d$ and $\varepsilon \in (0, 1)$. Let $Y := \big(C_c^\infty((-2\varepsilon, 2\varepsilon); \mathbb{R}), \|\cdot\|_{\sup}\big)$, and let Y' denote the dual space of Y. Note that Y' is a Banach space (see for instance [23, Proposition 5.4]), even though Y is not.

Step 2 *(Constructing a bounded operator $\Gamma : X \to Y'$):* Assume that $\mathcal{B}_{\mathcal{F},\alpha} \subset \mathcal{B}_\varrho(U)$. Then, since $|\xi|^\alpha \leq (1+|\xi|)^\beta$, we have for $f \in X$ that $f|_U \in \mathcal{B}_{\mathcal{F},\alpha}(U) \subset \mathcal{B}_\varrho(U)$, so that Lemma 7.2 shows that there is a Lipschitz continuous function $g : \mathbb{R}^d \to \mathbb{R}$ satisfying $g|_U = f|_U$ and such that for some choice of the weak derivative $\partial_1 g$ of g, if we set $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^d$ and $g_x : \mathbb{R} \to \mathbb{R}, t \mapsto (\partial_1 g)(x + t\,e_1)$, then $\sup_{x\in\mathbb{R}^d} \|g_x\|_{\mathrm{BV}} \leq 4\,\|f|_U\|_{\mathcal{B}_\varrho(U)} =: C_f$.

Since $f|_{U_0} = g|_{U_0}$ and f is continuously differentiable, we have $\partial_1 g = \partial_1 f$ almost everywhere on U_0. By Fubini's theorem, this implies for almost every $z \in (-\varepsilon,\varepsilon)^{d-1}$ that $g_{x_0+(0,z)}(t) = (\partial_1 g)(x_0 + (0,z) + t\,e_1) = (\partial_1 f)(x_0 + (0,z) + t\,e_1)$ for almost all $t \in (-2\varepsilon, 2\varepsilon)$.

For arbitrary $\varphi \in Y$ and z as above, we thus see by Lemma 7.3 that

$$\Big|\int_{\mathbb{R}} \varphi'(t)\,(\partial_1 f)\big(x_0+(0,z)+t\,e_1\big)\,dt\Big| = \Big|\int_{\mathbb{R}} \varphi'(t)\,g_{x_0+(0,z)}(t)dt\Big| \leq \big\|g_{x_0+(0,z)}\big\|_{\mathrm{BV}} \|\varphi\|_{\sup} \leq C_f\,\|\varphi\|_{\sup}.$$

Recall that this holds for almost all $z \in (-\varepsilon,\varepsilon)^{d-1}$, and thus in particular for a dense subset of $(-\varepsilon,\varepsilon)^{d-1}$. By continuity of $\partial_1 f$, we can thus take the limit $z \to 0$ to see that $|\int_{\mathbb{R}} \varphi'(t)\,(\partial_1 f)(x_0 + t\,e_1)\,dt| \leq C_f \cdot \|\varphi\|_{\sup}$ for all $\varphi \in Y$. We have thus shown that the linear map

$$\Gamma : \quad X \to Y', \quad f \mapsto \Big(\varphi \mapsto \int_{\mathbb{R}} \varphi'(t)\,(\partial_1 f)(x_0 + t\,e_1)\,dt\Big)$$

is well-defined. Note that if $f_n \xrightarrow[n\to\infty]{X} f$, then $\partial_1 f_n \to \partial_1 f$ with uniform convergence. Using this observation, it is straightforward to verify that Γ has closed graph, and is thus a bounded linear map, thanks to the closed graph theorem.

Finally, note that if $f \in X \cap C^2(\mathbb{R}^d)$, then we see by partial integration that

$$\Big|\int_{\mathbb{R}} \varphi(t)\,(\partial_1^2 f)(x_0+t\,e_1)\,dt\Big| = \Big|\int_{\mathbb{R}} \varphi'(t)\,(\partial_1 f)(x_0+t\,e_1)\,dt\Big| \leq \|\Gamma f\|_{Y'}\|\varphi\|_{\sup} \leq \|\Gamma\|\,\|f\|_X\,\|\varphi\|_{\sup}$$

for all $\varphi \in Y = C_c^\infty((-2\varepsilon,2\varepsilon);\mathbb{R})$. By the dual characterization of the L^1-norm (see for instance [1, Corollary 6.13]), this implies

$$\int_{-2\varepsilon}^{2\varepsilon} \big|(\partial_1^2 f)(x_0 + t\,e_1)\big|\,dt \leq \|\Gamma\| \cdot \|f\|_X \qquad \forall\, f \in X \cap C^2(\mathbb{R}^d). \tag{7.3}$$

Step 3 *(Completing the proof):* Pick $\gamma \in C_c^\infty(\mathbb{R}^d)$ with $0 \leq \gamma \leq 1$ and such that $\gamma \equiv 1$ on $U_0 = x_0 + (-3\varepsilon, 3\varepsilon)^d$. For $n \in \mathbb{N}$, define $f_n : \mathbb{R}^d \to \mathbb{R}, x \mapsto \gamma(x) \cdot \cos\big(\frac{n\pi}{\varepsilon}(x - x_0)\cdot e_1\big)$. Writing $T_y f(x) = f(x-y)$ and $M_\xi f(x) = e^{i\langle x,\xi\rangle}\,f(x)$ for translation and modulation, and using the identity $\cos(x) = \frac{1}{2}(e^{ix} + e^{-ix})$, it is easy to see $f_n = \frac{1}{2} T_{x_0}\big[M_{n\pi e_1/\varepsilon} T_{-x_0}\gamma + M_{-n\pi e_1/\varepsilon} T_{-x_0}\gamma\big]$, where $e_1 = (1,0,\dots,0)$. Consequently, elementary properties of the Fourier transform show that $\widehat{f_n} = \frac{1}{2} M_{-x_0}\big[T_{n\pi e_1/\varepsilon} M_{x_0}\widehat{\gamma} + T_{-n\pi e_1/\varepsilon} M_{x_0}\widehat{\gamma}\big]$, and hence

$$|\widehat{f_n}(\xi)| \leq \frac{1}{2}\big(|\widehat{\gamma}(\xi - \tfrac{n\pi}{\varepsilon}e_1)| + |\widehat{\gamma}(\xi + \tfrac{n\pi}{\varepsilon}e_1)|\big) \qquad \text{for all } \xi \in \mathbb{R}^d.$$

Since $1 + |\xi| \leq 1 + |\xi \pm \frac{n\pi}{\varepsilon}e_1| + \frac{n\pi}{\varepsilon} \leq (1 + |\xi \pm \frac{n\pi}{\varepsilon}e_1|)(1 + \frac{n\pi}{\varepsilon}) \leq \frac{2n\pi}{\varepsilon}(1 + |\xi \pm \frac{n\pi}{\varepsilon}e_1|)$, this shows that

$$\begin{aligned}
\|f_n\|_X &= \int_{\mathbb{R}^d} (1+|\xi|)^\beta \cdot |\widehat{f_n}(\xi)|\,d\xi \\
&\leq \frac{1}{2}\Big(\frac{2\pi n}{\varepsilon}\Big)^\beta \int_{\mathbb{R}^d} \sum_{\theta\in\{\pm 1\}} \big(1 + |\xi + \theta\tfrac{n\pi}{\varepsilon}e_1|\big)^\beta \big|\widehat{\gamma}(\xi + \theta\tfrac{n\pi}{\varepsilon}e_1)\big|\,d\xi \\
&\leq \Big(\frac{2\pi n}{\varepsilon}\Big)^\beta \int_{\mathbb{R}^d} (1+|\eta|)^\beta\,|\widehat{\gamma}(\eta)|\,d\eta \lesssim n^\beta.
\end{aligned} \tag{7.4}$$

On the other hand, for $t \in (-2\varepsilon, 2\varepsilon)$ we see because of $\gamma \equiv 1$ on U_0 that

$$(\partial_1^2 f_n)(x_0 + t\,e_1) = \frac{d^2}{dt^2} f_n(x_0 + t\,e_1) = \frac{d^2}{dt^2} \cos\left(\frac{n\pi}{\varepsilon} t\right) = -\left(\frac{n\pi}{\varepsilon}\right)^2 \cos\left(\frac{n\pi}{\varepsilon} t\right)$$

and hence

$$\begin{aligned}\int_{-2\varepsilon}^{2\varepsilon} |(\partial_1^2 f_n)(x_0 + t\,e_1)|\,dt &= (n\pi/\varepsilon)^2 \int_{-2\varepsilon}^{2\varepsilon} \left|\cos\left(n\pi t/\varepsilon\right)\right| dt = \frac{n\pi}{\varepsilon} \int_{2\pi n}^{2\pi n} |\cos(s)|\,ds \\ &\overset{(*)}{=} \frac{4\pi\, n^2}{\varepsilon} \int_0^{\pi} |\cos(s)|\,ds = \frac{8\pi n^2}{\varepsilon}.\end{aligned}$$

Here, we used at $(*)$ that $s \mapsto |\cos(s)|$ is π-periodic and even. Combining the last calculation with Equations (7.3) and (7.4), we arrive at $n^2 \lesssim \int_{-2\varepsilon}^{2\varepsilon} |(\partial_1^2 f_n)(x_0 + t\,e_1)|\,dt \lesssim \|f_n\|_X \lesssim n^\beta$, for all $n \in \mathbb{N}$. This is only possible if $\beta \geq 2$, and since $\beta = \max\{\alpha, 1\}$ this requires $\alpha \geq 2$. □

A A bound for empirical processes with finite pseudo-dimension

In this section, we prove a "uniform law of large numbers," similar to the pseudo-dimension based generalization bound in [40, Theorem 11.8], which is used in the third part of the proof of Proposition 2.2. The result given here is probably well-known; but since we could not locate a reference, we provide the proof. The main difference to the bound in [40] is that we estimate the expected sampling error, instead of giving a high probability bound; this allows us to omit a log factor. Furthermore, we use a complexity measure of the hypothesis class that differs slightly from the usual pseudo-dimension.

Proposition A.1. *There is a universal constant $\kappa > 0$ with the following property: If $(\Omega, \mathcal{F}, \mu)$ is a probability space, if $a, b \in \mathbb{R}$ with $a < b$, and if $\varnothing \neq \mathcal{G} \subset \{g : \Omega \to [a, b] \colon g \text{ measurable}\}$ satisfies*

$$d := \sup_{\lambda \in \mathbb{R}} \mathrm{VC}(\{I_{g,\lambda} \colon g \in \mathcal{G}\}) < \infty, \qquad \text{where} \qquad I_{g,\lambda} : \quad \Omega \to \{0,1\}, \quad \omega \mapsto \mathbb{1}_{g(\omega) > \lambda},$$

then for any $n \in \mathbb{N}$ and $S = (X_1, \dots, X_n) \overset{i.i.d.}{\sim} \mu$, we have

$$\mathbb{E}_S\left[\sup_{g \in \mathcal{G}} \left|\mathbb{E}_{X \sim \mu}[\,g(X)\,] - \frac{1}{n}\sum_{i=1}^n g(X_i)\right|\right] \leq \kappa \cdot (b - a) \cdot \sqrt{\frac{d}{n}}.$$

Remark. Here, as in most sources studying empirical processes (see e.g. [54, Section 7.2]), we interpret $\mathbb{E}[\sup_{i \in I} X_i]$ as $\sup_{I_0 \subset I \text{ finite}} \mathbb{E}[\sup_{i \in I_0} X_i]$, in order to avoid measurability issues.

Proof. Given a sample $S = (X_1, \dots, X_n) \in \Omega^n$, we write $\mu_S := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ for the associated empirical measure. We want to bound

$$\mathbb{E}\Big[\sup_{g \in \mathcal{G}} \left|\mathbb{E}_{X \sim \mu}[g(X)] - \mathbb{E}_{X \sim \mu_S}[g(X)]\right|\Big],$$

where the outer expectation is with respect to $S = (X_1, \dots, X_n) \overset{\text{i.i.d.}}{\sim} \mu$. First, by replacing $\mathcal{G}$ with $\mathcal{G}^* := \{g - a \colon g \in \mathcal{G}\}$, it is easy to see that we can assume $a = 0$ without loss of generality. Define $M := b = b - a$. Then, for any $g \in \mathcal{G}$ and any probability measure ν on Ω, the *layer cake formula* (see e.g. [23, Proposition 6.24]) shows

$$\mathbb{E}_{X \sim \nu}[g(X)] = \int_0^M \nu(\{\omega \in \Omega : g(\omega) > \lambda\})\,d\lambda = \int_0^M \mathbb{E}_{X \sim \nu}[I_{g,\lambda}(X)]\,d\lambda.$$

Therefore,

$$\begin{aligned}\Big|\mathbb{E}_{X\sim\mu}[g(X)] - \mathbb{E}_{X\sim\mu_S}[g(X)]\Big| &= \left|\int_0^M \mathbb{E}_{X\sim\mu}\big[I_{g,\lambda}(X)\big] - \mathbb{E}_{X\sim\mu_S}\big[I_{g,\lambda}(X)\big]\,d\lambda\right| \\ &\leq \int_0^M \Big|\mathbb{E}_{X\sim\mu}\big[I_{g,\lambda}(X)\big] - \mathbb{E}_{X\sim\mu_S}\big[I_{g,\lambda}(X)\big]\Big|\,d\lambda.\end{aligned}$$

In combination with the elementary estimate $\sup_{g\in\mathcal{G}}\int_0^M \Gamma_g(\lambda)\,d\lambda \leq \int_0^M \sup_{g\in\mathcal{G}}\Gamma_g(\lambda)\,d\lambda$ and Tonelli's theorem, this implies

$$\begin{aligned}\mathbb{E}\Big[\sup_{g\in\mathcal{G}}\big|\mathbb{E}_{X\sim\mu}[g(X)] - \mathbb{E}_{X\sim\mu_S}[g(X)]\big|\Big] &\leq \mathbb{E}\Big[\sup_{g\in\mathcal{G}}\int_0^M \Big|\mathbb{E}_{X\sim\mu}\big[I_{g,\lambda}(X)\big] - \mathbb{E}_{X\sim\mu_S}\big[I_{g,\lambda}(X)\big]\Big|\,d\lambda\Big] \\ &\leq \int_0^M \mathbb{E}\Big[\sup_{g\in\mathcal{G}}\Big|\mathbb{E}_{X\sim\mu}\big[I_{g,\lambda}(X)\big] - \mathbb{E}_{X\sim\mu_S}\big[I_{g,\lambda}(X)\big]\Big|\Big]\,d\lambda \\ &\overset{(*)}{\leq} \int_0^M \kappa\cdot\sqrt{\frac{\mathrm{VC}(\{I_{g,\lambda}\colon g\in\mathcal{G}\})}{n}}\,d\lambda \leq \kappa\cdot M\cdot\sqrt{\frac{d}{n}}.\end{aligned}$$

Here, the step marked with $(*)$ is an immediate consequence of the bound for the suprema of empirical processes based on the VC dimension given in [54, Theorem 8.3.23]. □

B A technical bound involving the total variation

Proof of Lemma 7.3. **Step 1:** We first show that if $h:\mathbb{R}\to\mathbb{R}$ is non-decreasing and bounded, then $|\int_{\mathbb{R}}\varphi'(t)\,h(t)\,dt| \leq \|\varphi\|_{\sup}\cdot\lim_{x\to\infty}[h(x)-h(-x)]$ for every $\varphi\in C_c^\infty(\mathbb{R})$. To see this, define $c := \lim_{x\to-\infty}h(x)$ and $\widetilde{h}:\mathbb{R}\to\mathbb{R}, x\mapsto\lim_{y\downarrow x}h(y)-c$. It is straightforward to see that $\widetilde{h}$ is non-decreasing, bounded, and right-continuous with $\lim_{x\to-\infty}\widetilde{h}(x)=0$, so that $\widetilde{h}\in\mathrm{NBV}$ in the notation of [23, Section 3.5]. Furthermore, since a monotonic function can have at most countably many discontinuities (see [23, Theorem 3.23]), we have $\widetilde{h}=h-c$ on the complement of a countable set, and hence almost everywhere. Since we also have $\int_{\mathbb{R}}\varphi'(t)\,dt=0$ thanks to the compact support of φ, if we denote by μ the unique Borel measure on $\mathbb{R}$ satisfying $\widetilde{h}(x)=\mu\big((-\infty,x]\big)$ for all $x\in\mathbb{R}$, then the partial integration formula in [23, Theorem 3.36] shows as claimed that

$$\begin{aligned}\left|\int_{\mathbb{R}}\varphi'(t)\,h(t)\,dt\right| = \left|\int_{\mathbb{R}}\varphi'(t)\,[h(t)-c]\,dt\right| &= \left|\int_{\mathbb{R}}\varphi'(t)\,\widetilde{h}(t)\,dt\right| = \left|\int_{\mathbb{R}}\widetilde{h}(t)\,d\varphi(t)\right| = \left|\int_{\mathbb{R}}\varphi(t)\,d\mu(t)\right| \\ &\leq \|\varphi\|_{\sup}\cdot\mu(\mathbb{R}) = \|\varphi\|_{\sup}\cdot\lim_{x\to\infty}\big[\,\widetilde{h}(x)-\widetilde{h}(-x)\,\big] \\ &= \|\varphi\|_{\sup}\cdot\lim_{x\to\infty}\big[h(x)-h(-x)\big].\end{aligned}$$

Step 2: Define

$$T_g:\ \mathbb{R}\to\mathbb{R},\ x\mapsto\sup\Big\{\sum_{j=1}^n|g(x_j)-g(x_{j-1})|\quad:\quad n\in\mathbb{N}\text{ and }-\infty<x_0<\cdots<x_n=x\Big\}.$$

Then T_g is non-decreasing and satisfies $\lim_{x\to-\infty}T_g(x)=0$ and $\lim_{x\to\infty}T_g(x)=\mathrm{TV}(g)$; furthermore, $g_1:=\frac{1}{2}(T_g+g)$ and $g_2:=\frac{1}{2}(T_g-g)$ are both non-decreasing and bounded with $g=g_1-g_2$; all of these properties can be found in [23, Section 3.5]. Note that $\lim_{x\to\infty}[\,g_i(x)-g_i(-x)\,]=\frac{1}{2}\mathrm{TV}(g)+\frac{(-1)^{i-1}}{2}\lim_{x\to\infty}[g(x)-g(-x)]$. In combination with the estimate from Step 1 (applied to $h=g_i$), this implies as claimed that

$$\left|\int_{\mathbb{R}}\varphi'(t)\,g(t)\,dt\right| \leq \sum_{i=1}^2\left|\int_{\mathbb{R}}\varphi'(t)\,g_i(t)\,dt\right| \leq \|\varphi\|_{\sup}\sum_{i=1}^2\lim_{x\to\infty}\big[\,g_i(x)-g_i(-x)\,\big] = \|\varphi\|_{\sup}\,\mathrm{TV}(g).\ \square$$

References

[1] H. W. Alt. *Linear functional analysis*. Universitext. Springer-Verlag London, Ltd., London, 2016. doi:10.1007/978-1-4471-7280-2.

[2] M. Anthony and P. L. Bartlett. *Neural network learning: theoretical foundations*. Cambridge University Press, Cambridge, 1999. doi:10.1017/CBO9780511624216.

[3] A. R. Barron. Neural net approximation. In *Proc. 7th Yale Workshop on Adaptive and Learning Systems*, volume 1, pages 69–72, 1992.

[4] A. R. Barron. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans. Inform. Theory*, 39(3):930–945, 1993. doi:10.1109/18.256500.

[5] A. R. Barron. Approximation and estimation bounds for artificial neural networks. *Machine learning*, 14(1):115–133, 1994. doi:10.1023/A:1022650905902.

[6] A. R. Barron and J. M. Klusowski. Approximation and estimation for high-dimensional deep learning networks. *arXiv preprint arXiv:1809.03090*, 2018.

[7] P. L. Bartlett, N. Harvey, C. Liaw, and A. Mehrabian. Nearly-tight VC-dimension and Pseudodimension Bounds for Piecewise Linear Neural Networks. *J. Mach. Learn. Res.*, 20:63–1, 2019.

[8] P. L. Bartlett, V. Maiorov, and R. Meir. Almost linear VC dimension bounds for piecewise polynomial networks. In *Advances in neural information processing systems*, pages 190–196, 1999. doi:10.1162/089976698300017016.

[9] C. Beck, L. Gonon, and A. Jentzen. Overcoming the curse of dimensionality in the numerical approximation of high-dimensional semilinear elliptic partial differential equations. *arXiv preprint arXiv:2003.00596*, 2020.

[10] S. Becker, P. Cheridito, A. Jentzen, and T. Welti. Solving high-dimensional optimal stopping problems using deep learning. *European Journal of Applied Mathematics*, 32(3):470–514, 2021. doi:10.1017/S0956792521000073.

[11] R. Bellman. On the theory of dynamic programming. *Proc. Nat. Acad. Sci. U.S.A.*, 38:716–719, 1952. doi:10.1073/pnas.38.8.716.

[12] J. Berner, P. Grohs, and A. Jentzen. Analysis of the Generalization Error: Empirical Risk Minimization over Deep Artificial Neural Networks Overcomes the Curse of Dimensionality in the Numerical Approximation of Black–Scholes Partial Differential Equations. *SIAM Journal on Mathematics of Data Science*, 2(3):631–657, 2020. doi:10.1137/19M125649X.

[13] H. Bölcskei, P. Grohs, G. Kutyniok, and P. Petersen. Optimal approximation with sparsely connected deep neural networks. *SIAM Journal on Mathematics of Data Science*, 1(1):8–45, 2019. doi:10.1137/18M118709X.

[14] A. Cloninger and T. Klock. A deep network construction that adapts to intrinsic dimensionality beyond the domain. *Neural Networks*, 141:404–419, 2021. doi:https://doi.org/10.1016/j.neunet.2021.06.004.

[15] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei. ImageNet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. IEEE, 2009. doi:10.1109/CVPR.2009.5206848.

[16] W. E, C. Ma, S. Wojtowytsch, and L. Wu. Towards a mathematical understanding of neural network-based machine learning: what we know and what we don't. *arXiv preprint arXiv:2009.10713*, 2020.

[17] W. E and S. Wojtowytsch. A priori estimates for classification problems using neural networks. *arXiv preprint arXiv:2009.13500*, 2020.

[18] W. E and S. Wojtowytsch. Representation formulas and pointwise properties for Barron functions. *Calc. Var. Partial Differ. Equ.*, 61(2):1–37, 2022. doi:10.1007/s00526-021-02156-6.

[19] D. E. Edmunds and H. Triebel. *Function spaces, entropy numbers, differential operators*, volume 120 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996. doi:10.1017/CBO9780511662201.

[20] D. Elbrächter, P. Grohs, A. Jentzen, and C. Schwab. DNN expression rate analysis of high-dimensional PDEs: Application to option pricing. *Constr. Approx.*, 55(1):3–71, 2022. doi:10.1007/s00365-021-09541-6.

[21] D. Elbrächter, D. Perekrestenko, P. Grohs, and H. Bölcskei. Deep neural network approximation theory. *IEEE Trans. Inf. Theory*, 67(5):2581–2623, 2021. doi:10.1109/TIT.2021.3062161.

[22] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010. doi:10.1090/gsm/019.

[23] G. B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999.

[24] L. Grüne. Overcoming the curse of dimensionality for approximating Lyapunov functions with deep neural networks under a small-gain condition. *IFAC-PapersOnLine*, 54(9):317–322, 2021. 24th International Symposium on Mathematical Theory of Networks and Systems MTNS 2020. doi:https://doi.org/10.1016/j.ifacol.2021.06.152.

[25] J. Han, A. Jentzen, and W. E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018. doi:10.1073/pnas.1718942115.

[26] S. Heinrich. Random Approximation in Numerical Analysis. In *Proceedings of the Conference "Functional Analysis" Essen*, pages 123–171, 1994. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.45.2552.

[27] M. Hutzenthaler, A. Jentzen, T. Kruse, T. Anh Nguyen, and P. von Wurstemberger. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. *Proceedings of the Royal Society A*, 476(2244):20190630, 2020. doi:10.1098/rspa.2019.0630.

[28] M. Imaizumi and K. Fukumizu. Deep neural networks learn non-smooth functions effectively. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 869–878. PMLR, 2019.

[29] M. Imaizumi and K. Fukumizu. Advantage of Deep Neural Networks for Estimating Functions with Singularity on Curves. *arXiv preprint arXiv:2011.02256*, 2020.

[30] A. Jentzen, D. Salimova, and T. Welti. A proof that deep artificial neural networks overcome the curse of dimensionality in the numerical approximation of Kolmogorov partial differential equations with constant diffusion and nonlinear drift coefficients. *Commun. Math. Sci.*, 19(5):1167 – 1205, 2021. doi:10.4310/CMS.2021.v19.n5.a1.

[31] J. M. Klusowski and A. R. Barron. Risk bounds for high-dimensional ridge function combinations including neural networks. *arXiv preprint arXiv:1607.01434*, 2016.

[32] A. Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009. URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012. doi:10.1145/3065386.

[34] F. Laakmann and P. Petersen. Efficient approximation of solutions of parametric linear transport equations by ReLU DNNs. *Adv. Comput. Math.*, 47(1):1–32, 2021. doi:10.1007/s10444-020-09834-7.

[35] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. doi:10.1038/nature14539.

[36] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi:10.1109/5.726791.

[37] H. Lee, R. Ge, T. Ma, A. Risteski, and S. Arora. On the ability of neural nets to express distributions. In *Conference on Learning Theory*, pages 1271–1296, 2017.

[38] C. Ma, L. Wu, and W. E. A priori estimates of the population risk for two-layer neural networks. *Communications in Mathematical Sciences*, 17(5):1407–1425, 2019. doi:10.4310/CMS.2019.v17.n5.a11.

[39] Y. Makovoz. Uniform approximation by neural networks. *J. Approx. Theory*, 95(2):215–228, 1998. URL: https://doi.org/10.1006/jath.1997.3217, doi:10.1006/jath.1997.3217.

[40] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of Machine Learning*. The MIT Press, 2nd edition, 2018.

[41] H. Montanelli, H. Yang, and Q. Du. Deep ReLU networks overcome the curse of dimensionality for generalized bandlimited functions. *J. Comput. Math.*, 39(6):801–815, 2021. doi:10.4208/jcm.2007-m2019-0239.

[42] R. Nakada and M. Imaizumi. Adaptive approximation and generalization of deep neural network with intrinsic dimensionality. *Journal of Machine Learning Research*, 21(174):1–38, 2020.

[43] E. Novak and H. Woźniakowski. Approximation of infinitely differentiable multivariate functions is intractable. *J. Complexity*, 25(4):398–404, 2009. doi:10.1016/j.jco.2008.11.002.

[44] R. Parhi and R. D. Nowak. Banach space representer theorems for neural networks and ridge splines. *J. Mach. Learn. Res.*, 22(43):1–40, 2021.

[45] P. Petersen and F. Voigtlaender. Optimal approximation of piecewise smooth functions using deep relu neural networks. *Neural Networks*, 108:296–330, 2018. doi:10.1016/j.neunet.2018.08.019.

[46] T. Poggio, H. Mhaskar, L. Rosasco, B. Miranda, and Q. Liao. Why and when can deep—but not shallow—networks avoid the curse of dimensionality: a review. *International Journal of Automation and Computing*, 14(5):503–519, 2017. doi:10.1007/s11633-017-1054-2.

[47] C. Reisinger and Y. Zhang. Rectified deep neural networks overcome the curse of dimensionality for nonsmooth value functions in zero-sum games of nonlinear stiff systems. *Analysis and Applications*, 18(06):951–999, 2020. doi:10.1142/S0219530520500116.

[48] C. Schwab and J. Zech. Deep learning in high dimension: Neural network expression rates for generalized polynomial chaos expansions in UQ. *Analysis and Applications*, 17(01):19–55, 2019. doi:10.1142/S0219530518500203.

[49] U. Shaham, A. Cloninger, and R. R. Coifman. Provable approximation properties for deep neural networks. *Applied and Computational Harmonic Analysis*, 44(3):537–557, 2018. doi:10.1016/j.acha.2016.04.003.

[50] S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

[51] J. W. Siegel and J. Xu. Approximation rates for neural networks with general activation functions. *Neural Networks*, 128:313–321, 2020. doi:10.1016/j.neunet.2020.05.019.

[52] J. W. Siegel and J. Xu. Optimal approximation rates and metric entropy of ReLU^k and cosine networks. *arXiv preprint arXiv:2101.12365*, 2021.

[53] H. Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010.

[54] R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

[55] S. Wojtowytsch and W. E. On the Banach spaces associated with multi-layer ReLU networks: Function representation, approximation theory and gradient descent dynamics. *arXiv preprint arXiv:2007.15623*, 2020.

[56] Y. Yang and A. Barron. Information-theoretic determination of minimax rates of convergence. *Ann. Statist.*, 27(5):1564–1599, 1999. URL: https://doi.org/10.1214/aos/1017939142, doi:10.1214/aos/1017939142.

6.5 Paper 5

Quantitative approximation results for complex-valued neural networks *

A. Caragea†, D.G. Lee†, J. Maly†, G.E. Pfander†, and F. Voigtlaender†‡

Abstract. Until recently, applications of neural networks in machine learning have almost exclusively relied on real-valued networks. It was recently observed, however, that complex-valued neural networks (CVNNs) exhibit superior performance in applications in which the input is naturally complex-valued, such as MRI fingerprinting. While the mathematical theory of real-valued networks has, by now, reached some level of maturity, this is far from true for complex-valued networks. In this paper, we analyze the expressivity of complex-valued networks by providing explicit quantitative error bounds for approximating C^n functions on compact subsets of $\mathbb{C}^d$ by complex-valued neural networks that employ the modReLU activation function, given by $\sigma(z) = \mathrm{ReLU}(|z|-1)\,\mathrm{sgn}(z)$, which is one of the most popular complex activation functions used in practice. We show that the derived approximation rates are optimal (up to log factors) in the class of modReLU networks with weights of moderate growth.

Key words. Deep neural networks, Complex-valued neural networks, function approximation, modReLU activation function

AMS subject classifications. 68T07, 41A25, 41A46.

1. Introduction.

Motivated by the remarkable practical success of machine learning algorithms based on deep neural networks (collectively called *deep learning* [18]) in applications like image recognition [17] and machine translation [25], the expressive power of such neural networks is the topic of an active and rich area of study [20, 22, 31, 32]. Results on the expressivity of real-valued neural networks date back to the 90s, when the main focus was on networks with *smooth* activation functions [21]. More recently, emphasis has shifted towards networks using the *rectified linear unit (ReLU)* activation function $\varrho(x) = \max\{0, x\}$, as those networks have been observed to yield similar expressive power at a greatly reduced training time cost [10, 18].

Due to the missing support for complex arithmetic in the leading deep learning software libraries [26], practical applications of deep neural networks have almost exclusively employed *real-valued* neural networks. Recently, however, there has been an increased interest in *complex-valued neural networks (CVNNs)* for problems in which the input is naturally complex-valued and in which a faithful treatment of phase information is important [26, 28]. For instance, for the problem of MRI fingerprinting, CVNNs significantly outperform their real-valued counterparts [28]. Moreover, CVNNs have demonstrated greatly improved stability and

*Version: December 6, 2021

Funding: A. Caragea acknowledges support by the DFG Grant PF 450/11-1. D.G. Lee acknowledges support by the DFG Grants PF 450/6-1 and PF 450/9-1. FV acknowledges support by the German Science Foundation (DFG) in the context of the Emmy Noether junior research group VO 2594/1–1.

†KU Eichstätt–Ingolstadt, Mathematisch–Geographische Fakultät, Ostenstraße 26, Kollegiengebäude I Bau B, 85072 Eichstätt, Germany (andrei.caragea@ku.de, daegwans@gmail.com, johannes.maly@ku.de, pfander@ku.de, felix.voigtlaender@ku.de)

‡Department of Mathematics, Technical University of Munich, 85748 Garching bei München, Germany.

convergence properties for the setting of recurrent neural networks [4, 30].

Motivated by the increased interest in complex-valued neural networks, we herein initiate the analysis of their expressive power, quantified by their ability to approximate functions of a given regularity. Specifically, we analyze how well CVNNs with the modReLU activation function (defined in Subsection 1.1) can approximate functions of Sobolev regularity $\mathcal{W}^{n,\infty}$ on compact subsets of $\mathbb{C}^d$ (see Subsection 1.2). The explicit result is given in Subsection 1.3.

1.1. Complex-valued neural networks and the modReLU function. In a complex-valued neural network (CVNN), each neuron computes a function of the form $\boldsymbol{z} \mapsto \sigma(\boldsymbol{w}^T \boldsymbol{z} + b)$ with $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{C}^N$ and $b \in \mathbb{C}$, where $\sigma : \mathbb{C} \to \mathbb{C}$ is a complex activation function.

Formally, a *complex-valued neural network (CVNN)* is a tuple $\Phi = \big((A_1, b_1), \dots, (A_L, b_L)\big)$, where $L =: L(\Phi) \in \mathbb{N}$ denotes the *depth* of the network and where $A_\ell \in \mathbb{C}^{N_\ell \times N_{\ell-1}}$ and $b_\ell \in \mathbb{C}^{N_\ell}$ for $\ell \in \{1, \dots, L\}$. Then $d_{\text{in}}(\Phi) := N_0$ and $d_{\text{out}}(\Phi) := N_L$ denote the input- and output-dimension of Φ. Given any function $\sigma : \mathbb{C} \to \mathbb{C}$, the *network function* associated to the network Φ (also called the *realization* of Φ) is the function

$$R_\sigma \Phi := T_L \circ (\sigma \circ T_{L-1}) \circ \cdots \circ (\sigma \circ T_1) : \mathbb{C}^{d_{\text{in}}(\Phi)} \to \mathbb{C}^{d_{\text{out}}(\Phi)} \quad \text{where} \quad T_\ell \, \boldsymbol{z} = A_\ell \, \boldsymbol{z} + b_\ell,$$

and where σ acts componentwise on vectors, meaning $\sigma\big((z_1, \dots, z_k)\big) = \big(\sigma(z_1), \dots, \sigma(z_k)\big)$. The functions T_L and $\sigma \circ T_\ell$ for $\ell \in \{1, \dots, L-1\}$ are the functions computed by the different *layers* of the network Φ. The network Φ is called *shallow* if $L = 2$, i.e., if Φ has only one "internal layer" (neither an input, nor an output layer), which is usually called a *hidden layer*.

The *number of neurons* $N(\Phi)$ of Φ is $N(\Phi) := \sum_{\ell=0}^{L} N_\ell$, the *width* (or *breadth*) of Φ is $B(\Phi) := \max_{0 \le \ell \le L} N_\ell$, and the *number of weights* of Φ is $W(\Phi) := \sum_{j=1}^{L} (\|A_j\|_{\ell^0} + \|b_j\|_{\ell^0})$, where $\|A\|_{\ell^0}$ denotes the number of nonzero entries of a matrix or vector A. Moreoever, writing $\|A\|_\infty := \max_{i,j} |A_{i,j}|$ for a matrix (or vector) A, we define the norm of the network Φ as $\|\Phi\| := \max_{1 \le \ell \le L} \max\{\|A_\ell\|_\infty, \|b_\ell\|_\infty\}$. We then say that *the weights of* Φ *are bounded by* $C \ge 0$ if $\|\Phi\| \le C$.

Finally, we will also use the notion of a *network architecture*[1]. Formally, this is a tuple $\mathcal{A} = \big((N_0, \dots, N_L), (I_1, \dots, I_L), (J_1, \dots, J_L)\big)$ where $(N_0, \dots, N_L)$ determines the depth L of the network and the number of neurons N_ℓ in each layer. The sets $J_\ell \subset \{1, \dots, N_\ell\}$ and $I_\ell \subset \{1, \dots, N_\ell\} \times \{1, \dots, N_{\ell-1}\}$ determine which weights of the network are permitted to be nonzero. Thus, *a network* Φ *is of architecture* $\mathcal{A}$ as above if $\Phi = \big((A_1, b_1), \dots, (A_L, b_L)\big)$ where $A_\ell \in \mathbb{C}^{N_\ell \times N_{\ell-1}}$ and $b_\ell \in \mathbb{C}^{N_\ell}$, and if furthermore $(A_\ell)_{j,k} = 0$ if $(j,k) \notin I_\ell$ and $(b_\ell)_j = 0$ if $j \notin J_\ell$. The number of weights and neurons of an architecture $\mathcal{A}$ are defined as $W(\mathcal{A}) := \sum_{\ell=1}^{L} (|J_\ell| + |I_\ell|)$ and $N(\mathcal{A}) := \sum_{\ell=0}^{L} N_\ell$, respectively.

In the present paper, we focus on neural networks using the *modReLU* activation function

$$\sigma : \quad \mathbb{C} \to \mathbb{C}, \qquad z \mapsto \varrho(|z| - 1) \operatorname{sgn}(z) = \begin{cases} 0, & \text{if } |z| \le 1, \\ z - \frac{z}{|z|}, & \text{if } |z| \ge 1 \end{cases} \tag{1.1}$$

[1] The term "network architecture" as used here does *not* refer to conceptual network architectures like feed-forward networks, recursive neural networks, and others. Instead, since we are *only* concerned with fully connected feed-forward networks, the "network architecture" only prescribes the network shape in terms of the number of layers, the number of neurons per layer, and which weights of the network may be non-zero. This terminology is widespread in the literature studying the approximation properties of neural networks; see e.g. [22, 31].

proposed in [4] as a generalization of the ReLU activation function $\varrho : \mathbb{R} \to \mathbb{R}, x \mapsto \max\{0, x\}$ to the complex domain. Note that the complex sign function is defined as $\operatorname{sgn}(z) = \frac{z}{|z|}$ for $z \neq 0$, and $\operatorname{sgn}(z) = 0$ else. We briefly discuss other activation functions in Subsection 1.4.

1.2. Smoothness assumptions. We are interested in approximating functions $f : \mathbb{C}^d \to \mathbb{C}$ that belong to the Sobolev space $\mathcal{W}^{n,\infty}$, with differentiability understood in the sense of real variables. Specifically, let

$$Q_{\mathbb{C}^d} := \big\{\boldsymbol{z} = (z_1, \dots, z_d) \in \mathbb{C}^d : \operatorname{Re} z_k, \operatorname{Im} z_k \in [0,1] \text{ for all } 1 \leq k \leq d\big\}$$

be the (real) unit cube in $\mathbb{C}^d$. As in the definition of $Q_{\mathbb{C}^d}$, we will use throughout the paper boldface characters to denote real and complex vectors.

Identifying $\boldsymbol{z} = (z_1, \dots, z_d) \in \mathbb{C}^d$ with $\boldsymbol{x} = \big(\operatorname{Re}(z_1), \dots, \operatorname{Re}(z_d), \operatorname{Im}(z_1), \dots, \operatorname{Im}(z_d)\big) \in \mathbb{R}^{2d}$, we will consider $\mathbb{C}^d \cong \mathbb{R}^{2d}$ as usual. With this in mind, a complex function $g : \mathbb{C}^d \to \mathbb{C}$ can be identified with a pair of functions $g_{\mathrm{Re}}, g_{\mathrm{Im}} : \mathbb{R}^{2d} \to \mathbb{R}$ given by $g_{\mathrm{Re}} = \operatorname{Re}(g)$ and $g_{\mathrm{Im}} = \operatorname{Im}(g)$.

Given a real function $f : [0,1]^{2d} \to \mathbb{R}$ and $n \in \mathbb{N}$, we write $f \in \mathcal{W}^{n,\infty}([0,1]^{2d}; \mathbb{R})$ if f is $n-1$ times continuously differentiable with all derivatives of order $n-1$ being Lipschitz continuous. We then define

$$\|f\|_{\mathcal{W}^{n,\infty}} := \max\Big\{\max_{|\alpha| \leq n-1} \|\partial^\alpha f\|_{L^\infty}, \max_{|\alpha| = n-1} \operatorname{Lip}(\partial^\alpha f)\Big\}.$$

Using this norm, we define the unit ball in the Sobolev space $\mathcal{W}^{n,\infty}$ as

$$F_{n,d} := \big\{f \in \mathcal{W}^{n,\infty}([0,1]^{2d}; \mathbb{R}) \quad : \quad \|f\|_{\mathcal{W}^{n,\infty}} \leq 1\big\}$$

and define the set of functions that we seek to approximate by

$$\mathcal{F}_{n,d} := \big\{g : Q_{\mathbb{C}^d} \to \mathbb{C} \quad : \quad g_{\mathrm{Re}}, g_{\mathrm{Im}} \in F_{n,d}\big\}.$$

1.3. Main result. Our main result provides explicit error bounds for approximating functions $g \in \mathcal{F}_{n,d}$ using modReLU networks. This result can be seen as a generalization to the complex domain of the approximation bounds for ReLU networks developed in [31].

Theorem 1. *For any $d, n \in \mathbb{N}$, there exists $C = C(d,n) > 0$ with the following property:*

Given any $\varepsilon \in (0,1)$ there exists a modReLU*-network architecture $\mathcal{A}$ with no more than $C \cdot \ln(2/\varepsilon)$ layers and no more than $C \cdot \varepsilon^{-2d/n} \cdot \ln^2(2/\varepsilon)$ weights such that for any $g \in \mathcal{F}_{n,d}$ there exists a network Φ of architecture $\mathcal{A}$ with all weights bounded by $C \cdot \varepsilon^{-44d}$ and such that $|g(\boldsymbol{z}) - R_\sigma \Phi(\boldsymbol{z})| \leq \varepsilon$ for all $\boldsymbol{z} \in Q_{\mathbb{C}^d}$.*

The exponent $-\frac{2d}{n}$ in place of $-\frac{d}{n}$ in the real setting is a consequence of the identification $\mathbb{C}^d \cong \mathbb{R}^{2d}$. More precisely, making the identifications $\mathbb{C}^d \cong \mathbb{R}^{2d}$ and $\mathbb{C} \cong \mathbb{R}^2$ and using (real-valued) ReLU networks (with two output channels), the results in [31] show that—up to logarithmic factors—ReLU networks achieve *the same* approximation bounds as those shown in Theorem 1 for modReLU-CVNNs. Thus, as far as the asymptotic approximation rate is concerned, modReLU-CVNNs do not *strictly improve* on the approximation capabilities of ReLU networks, but they can *match* their approximation power. This is an important theoretical finding, since even though CVNNs were found to have advantages in several applications [4,28],

up to now, no quantitative approximation results for CVNNs were known whatsoever—only universal approximation type results were available [2, 29]. Our results show that, at least for the approximation problem considered here, there is no additional "cost" in using CVNNs, compared to ReLU networks.

Since Theorem 1 only provides asymptotic rates (i.e., no explicit bound on the constant C is provided) and since the C^n assumption regarding the function to be approximated or learned cannot be verified in practical applications, the theorem is of limited use for guiding deep learning practitioners. Rather, it is intended as a first step towards mathematically understanding the expressivity of CVNNs and is furthermore expected to be informative for other theoretical works, for instance for analyzing the performance of CVNNs for approximating the solutions of PDEs, similar to the results in [11–13].

Remark 2. Note that the architecture and therefore the size of the network Φ is independent of the function g to approximate, once we fix an approximation accuracy ε and the parameters n and d. Only the choice of weights depends on g.

1.4. Comparison to existing work.

Approximation results for CVNNs. While the approximation properties of real-valued neural networks are comparatively well understood by now, the corresponding questions for complex-valued networks remain mostly open. In fact, even the property of universality—well studied for real-valued networks [9, 15, 16, 19]—was only settled for very specific activation functions [1–3, 14], until the recent paper [29] resolved the question. This universal approximation theorem for CVNNs highlights that the properties of complex-valued networks are significantly more subtle than those of their real-valued counterparts: Real-valued networks (either shallow or deep) are universal if and only if the activation function is not a polynomial [19]. In contrast, shallow *complex-valued* networks are universal if and only if the real part or the imaginary part of the activation function σ is not *polyharmonic*, while *deep* complex-valued networks (with more than one hidden layer) are universal if and only if σ is neither holomorphic, nor antiholomorphic, nor a polynomial (in z and $\overline{z}$). For instance, *deep* networks with the activation function $\sigma(z) = \overline{z} \cdot e^z$ are universal, but *shallow* networks with this activation function are not.

Aside of these purely qualitative universality results, *no quantitative approximation bounds for complex-valued networks are known whatsoever*. The present paper is thus the first to provide such bounds.

Role of the activation function. As empirically observed in [4, 28], the main advantage of complex-valued networks over their real-valued counterparts stems from the fact that the set of implementable *complex* activation functions is much richer than in the real-valued case. In fact, each real-valued activation function $\rho : \mathbb{R} \to \mathbb{R}$ can be lifted to the complex function $\sigma(z) := \rho(\operatorname{Re} z)$; then, $\sigma(\boldsymbol{w}^T \boldsymbol{z} + b) = \rho(\boldsymbol{\alpha}^T \boldsymbol{x} - \boldsymbol{\beta}^T \boldsymbol{y} + \operatorname{Re} b)$ for $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$ and $\boldsymbol{w} = \boldsymbol{\alpha} + i\boldsymbol{\beta}$. Thus, identifying $\mathbb{C}^d \cong \mathbb{R}^{2d}$, every real-valued network can be written as a complex-valued one. Therefore, one can in principle transfer every approximation result involving real-valued networks to a corresponding complex-valued result. Similar arguments apply to activation functions of the form $\sigma(z) = \rho(\operatorname{Re} z) + i\rho(\operatorname{Im} z)$.

However, using such "intrinsically real-valued" activation functions forfeits the main benefits of using complex-valued networks, namely increased expressivity and a faithful handling of phase and magnitude information. Therefore, the two most prominent complex-valued

activation functions appearing in the literature (see [4, 26, 28]) are the *modReLU* (see Equation (1.1)) and the *complex cardioid* (given by $\sigma(z) = \frac{z}{2} \cdot \big(1 + \frac{\operatorname{Re} z}{|z|}\big)$), neither of which is of the form $\rho(\operatorname{Re}(z))$ for a real activation function ρ.

In the present work, we focus on the modReLU activation function because it satisfies the natural *phase homogeneity property* $\sigma(e^{i\theta} z) = e^{i\theta}\sigma(z)$. Investigating the complex cardioid—and other complex-valued activation functions—is an interesting topic for future work.

Role of the network depth. Deep networks greatly outperform their shallow counterparts in applications [18]; therefore, much research has been devoted to rigorously quantify the influence of the network depth on the expressivity of (real-valued) neural networks. The precise findings depend on the activation function: While for *smooth* activation functions, already *shallow* networks with $\mathcal{O}(\varepsilon^{-d/n})$ weights and neurons can uniformly approximate functions $f \in C^n([0,1]^d)$ up to error ε (see [21]), this is not true for ReLU networks. To achieve the same approximation rate, ReLU networks need at least $\mathcal{O}(1 + \frac{n}{d})$ layers [22–24]. The proofs of these bounds crucially use that the ReLU is piecewise linear. Since this is not true of the modReLU, these arguments do not apply here.

Regarding sufficiency, the best known approximation result for ReLU networks [31] shows—similar to our main theorem—that ReLU networks with depth $\mathcal{O}(\ln(2/\varepsilon))$ and $\mathcal{O}(\varepsilon^{-d/n}\ln(2/\varepsilon))$ weights can approximate functions $f \in C^n([0,1]^d)$ uniformly up to error ε. For networks with bounded depth, similar results are only known for approximation in L^p [22] or for approximation in terms of the network *width* instead of the number of nonzero weights [20]. It is an interesting question whether these two results extend to modReLU networks as well.

Finally, we mention an intriguing result in [32] which shows that *extremely deep* ReLU networks (for which the number of layers is proportional to the number of weights) with *extremely complicated weights* (meaning the number of significant digits per weight grows unboundedly as $\varepsilon \downarrow 0$) can approximate functions $f \in C^n([0,1]^d)$ up to error ε using only $\mathcal{O}(\varepsilon^{-d/(2n)})$ weights (up to log factors). Due to the prohibitive complexity of the network weights this bound has limited practical significance, but is an extremely surprising and insightful mathematical result. We expect that the arguments in [32] can be extended to modReLU networks, but leave this as future work.

Optimality. For modReLU networks with polynomial growth of the individual weights and logarithmic growth of the depth (as in Theorem 1), the approximation rate of Theorem 1 is essentially optimal. We prove this in detail in Section 7, Theorem 12. Our proof relies on *entropy arguments*, which are closely related to the proof techniques based on rate distortion theory as used in [7,22]. Furthermore, for deriving suitable covering bounds for certain network sets (which then give rise to entropy bounds), we borrow several proof ideas from [6].

For ReLU networks, a similar optimality result holds for networks with logarithmic growth of the depth *even without assumptions on the magnitude of the network weights* [31]. The proof relies on sharp bounds for the VC dimension of ReLU networks [5]. For modReLU networks, a similar question is more subtle, since to the best of our knowledge no analogous VC dimension bounds are available. We thus leave it as future work to study optimality *without* assumptions on the magnitude of the network weights.

1.5. Structure of the paper. Inspired by [31], our proof of Theorem 1 proceeds by locally approximating g using Taylor polynomials, and then showing that these Taylor polynomials

and a suitable partition of unity can be well approximated by modReLU networks. To prove this, we first show in Section 2 that modReLU networks *of constant size* can approximate the functions $z \mapsto \operatorname{Re} z$ and $z \mapsto \operatorname{Im} z$ arbitrarily well—only the magnitude of the individual weights of the network grows as the approximation accuracy improves. Then, based on proof techniques in [31], we show in Section 3 that modReLU networks with $\mathcal{O}(\ln^2(2/\varepsilon))$ weights and $\mathcal{O}(\ln(2/\varepsilon))$ layers can approximate the function $z \mapsto (\operatorname{Re} z)^2$ up to error ε. By a polarization argument, this also allows to approximate the product function $(z, w) \mapsto zw$; see Section 4. After describing in Section 5 how a partition of unity can be implemented with modReLU networks, we combine all the ingredients in Section 6 to prove Theorem 1. Finally, Section 7 proves that Theorem 1 is essentially optimal.

2. Approximating real and imaginary parts. This section shows that modReLU networks *of constant size* can approximate the functions $z \mapsto \operatorname{Re} z$ and $z \mapsto \operatorname{Im} z$ arbitrarily well:

Proposition 3. *For any $R \geq 1$ and $\varepsilon \in (0,1)$, there exist functions $\operatorname{Re}_{R,\varepsilon}, \operatorname{Im}_{R,\varepsilon} : \mathbb{C} \to \mathbb{C}$ that are implemented by* shallow *σ-networks with 5 neurons and 10 weights, all bounded in absolute value by $C \cdot R^3/\varepsilon^3$ with an absolute constant $C > 0$, satisfying*

$$|\operatorname{Re}_{R,\varepsilon}(z) - \operatorname{Re}(z)| \leq \varepsilon \quad \text{and} \quad |\operatorname{Im}_{R,\varepsilon}(z) - \operatorname{Im}(z)| \leq \varepsilon \qquad \text{for all } z \in \mathbb{C} \text{ with } |z| \leq R.$$

To prove Proposition 3, we need two ingredients: First, modReLU networks can implement the identity function on bounded subsets of $\mathbb{C}$ *exactly*. To be precise, for arbitrary $R > 0$ it holds that $\operatorname{Id}_R(z) = z$ for $z \in \mathbb{C}$ with $|z| \leq R$, where

$$\operatorname{Id}_R(z) := \sigma\big(2z + 2R + 2\big) - \sigma(z + R + 1) - (R + 1). \tag{2.1}$$

Indeed, for $w \in \mathbb{C}$ with $|w| \geq 1$, we have $\sigma(2w) - \sigma(w) = 2w - \frac{2w}{|2w|} - (w - \frac{w}{|w|}) = w$. For $z \in \mathbb{C}$ with $|z| \leq R$, setting $w = z + R + 1$ so that $|w| \geq 1$ gives $\operatorname{Id}_R(z) = z$.

As the second ingredient, we use the following functions, parameterized by $h > 0$:

$$\operatorname{Im}_h(z) := \frac{-i}{h^2} \cdot \Big(\operatorname{sgn}\big(hz + \tfrac{1}{h}\big) - 1\Big) \quad \text{and} \quad \operatorname{Re}_h(z) := \frac{1}{h^2} \cdot \Big(\operatorname{sgn}\big(hz - \tfrac{i}{h}\big) + i\Big), \quad z \in \mathbb{C}.$$

The next lemma shows that these complex-valued functions well approximate the real-valued functions Re and Im. The proof of Proposition 3 will then consist of showing that Im_h and Re_h can be implemented by modReLU networks.

Lemma 4. *For $z \in \mathbb{C}$ and $0 < h \leq \frac{1}{2+2|z|}$, we have*

$$\Big|\operatorname{Im}(z) - \operatorname{Im}_h(z)\Big| \leq 2h|z| \qquad \text{and} \qquad \Big|\operatorname{Re}(z) - \operatorname{Re}_h(z)\Big| \leq 2h|z|. \tag{2.2}$$

Proof. See Appendix A.1. ∎

Proof of Proposition 3. Set $h := \frac{\varepsilon}{2+2R}$, noting that indeed $0 < h \leq \frac{1}{2+2|z|}$ and $h\,|z| \leq \frac{\varepsilon}{2}$ whenever $|z| \leq R$. Note that $w := hz - \frac{i}{h}$ satisfies $|w| \leq \frac{1}{h} + h\,|z| \leq \frac{2}{h} =: R'$ and $|w| = |hz - \frac{i}{h}| \geq \frac{1}{h} - h\,|z| \geq 2 - \frac{1}{2} \geq 1$, so that $\operatorname{sgn}(w) = w - \sigma(w) = \operatorname{Id}_{R'}(w) - \sigma(w)$, with

$\mathrm{Id}_{R'}$ as in Equation (2.1). Putting together the definitions of h, R', w and of Id_R, we see that

$$\begin{aligned}\mathrm{Re}_h(z) &= h^{-2}\cdot\big(\mathrm{Id}_{R'}(w)-\sigma(w)+i\big)\\ &= h^{-2}\cdot\Big(\sigma(2hz+\tfrac{4}{h}+2-\tfrac{2i}{h})-\sigma(hz+\tfrac{2}{h}+1-\tfrac{i}{h})-\sigma(hz-\tfrac{i}{h})+i-\tfrac{2}{h}-1\Big) \qquad (2.3)\\ &=: \mathrm{Re}_{R,\varepsilon}(z)\end{aligned}$$

is implemented by a shallow σ-network with 5 neurons and 10 weights (see Figure 1), where all the weights are bounded by $\frac{4}{h^3} \le C\,R^3/\varepsilon^3$ for an absolute constant $C>0$. Finally, Lemma 4 shows $|\mathrm{Re}(z)-\mathrm{Re}_{R,\varepsilon}(z)|\le\varepsilon$ for all $z\in\mathbb{C}$ with $|z|\le R$. The claim concerning the approximation of $\mathrm{Im}(z)$ is shown similarly. ∎

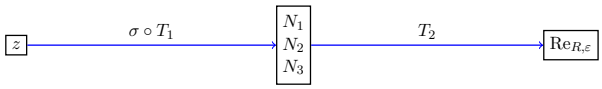

Figure 1. *Architecture of the network* $\mathrm{Re}_{R,\varepsilon}$*, where* $T_1(\cdot)=A_1(\cdot)+b_1$ *for* $A_1=\big(\frac{\varepsilon}{1+R},\frac{\varepsilon}{2(1+R)},\frac{\varepsilon}{2(1+R)}\big)^T$ *and* $b_1=\big(\frac{(8-4i)(1+R)}{\varepsilon}+2,\frac{(4-2i)(1+R)}{\varepsilon}+1,\frac{-2i(1+R)}{\varepsilon}\big)^T$*, and* $T_2(\cdot)=A_2(\cdot)+b_2$ *for* $A_2=\frac{4(1+R)^2}{\varepsilon^2}(1,-1,-1)$ *and* $b_2=\frac{4(1+R)^2}{\varepsilon^2}\cdot\big(-1-\frac{4(1+R)}{\varepsilon}+i\big)$.

3. Approximating the squared real part.

The main result of this section is Proposition 8 below, showing that the function $z\mapsto(\mathrm{Re}(z))^2$ on the set $\{z\in\mathbb{C}:|z|\le R,\ |\mathrm{Re}(z)|\le 1\}$ can be uniformly approximated up to error ε by modReLU networks with $\mathcal{O}\big(\ln(2/\varepsilon)\big)$ layers and $\mathcal{O}\big(\ln^2(2/\varepsilon)\big)$ weights of size $\mathcal{O}\big(R^6\,\varepsilon^{-7}\big)$.

As a first step towards proving Proposition 8, we show that modReLU networks can approximate functions of the form $z\mapsto\varrho(\mathrm{Re}(z)+c)$ with the usual ReLU ϱ; this will then allow us to use the approximation of the square function by ReLU networks as derived in [31].

Proposition 5. *For any choice of* $R\ge 1$, $c\in\mathbb{R}$, *and* $\varepsilon\in(0,1)$, *there exist functions* $\varrho_{R,\varepsilon}^{\mathrm{Re},c},\varrho_{R,\varepsilon}^{\mathrm{Im},c}:\mathbb{C}\to\mathbb{C}$ *that are implemented by depth-3* σ*-networks with* 6 *neurons and* 11 *weights, all bounded in absolute value by* $C\cdot R^3/\varepsilon^3+2|c|$ *(with an absolute constant* C*), satisfying* $|\varrho_{R,\varepsilon}^{\mathrm{Re},c}(z)-\varrho(\mathrm{Re}(z)+c)|\le\varepsilon$ *and* $|\varrho_{R,\varepsilon}^{\mathrm{Im},c}(z)-\varrho(\mathrm{Im}(z)+c)|\le\varepsilon$ *for all* $z\in\mathbb{C}$ *with* $|z|\le R$.

Proof. Let us first prove the statement for $\varrho_{R,\varepsilon}^{\mathrm{Re},c}$. To this end, first note that the modReLU $\sigma:\mathbb{C}\to\mathbb{C}$ is 1-Lipschitz; see Lemma 6 below.

Now, set $\delta:=1/\big(2\cdot(R+|c|)\big)$ and define

$$\varrho_{R,\varepsilon}^{\mathrm{Re},c}:\quad\mathbb{C}\to\mathbb{C},\quad z\mapsto\frac{1}{\delta}\cdot\sigma\big(1+\delta\cdot\big(\mathrm{Re}_{R,\varepsilon}(z)+c\big)\big), \qquad (3.1)$$

where $\mathrm{Re}_{R,\varepsilon}$ is as in Proposition 3. Now, a direct computation (see also Figure 2) shows that $\sigma(x+1)=\varrho(x)$ for $x\in[-2,\infty)$. Because of $|\delta\cdot(\mathrm{Re}(z)+c)|\le\delta\cdot(|z|+|c|)\le\frac12$ for $|z|\le R$, this implies $\frac{1}{\delta}\cdot\sigma\big(1+\delta\cdot(\mathrm{Re}(z)+c)\big)=\frac{1}{\delta}\cdot\varrho\big(\delta\cdot(\mathrm{Re}(z)+c)\big)=\varrho\big(\mathrm{Re}(z)+c\big)$. Combined with the 1-Lipschitz continuity of σ, we thus see

$$\begin{aligned}|\varrho(\mathrm{Re}(z)+c)-\varrho_{R,\varepsilon}^{\mathrm{Re},c}(z)| &= \big|\tfrac{1}{\delta}\cdot\sigma(1+\delta\cdot(\mathrm{Re}(z)+c))-\tfrac{1}{\delta}\cdot\sigma(1+\delta\cdot(\mathrm{Re}_{R,\varepsilon}(z)+c))\big|\\ &\le|\mathrm{Re}(z)-\mathrm{Re}_{R,\varepsilon}(z)|\le\varepsilon \quad \text{for all } z\in\mathbb{C}\text{ with }|z|\le R.\end{aligned}$$

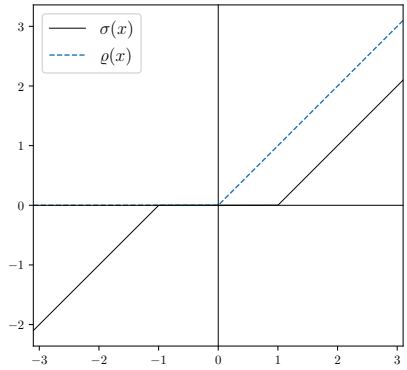

Figure 2. *A plot of the modReLU function σ on $[-3,3]$. The plot shows that $\sigma(x+1)=\varrho(x)$ for $x\in[-2,\infty)$.*

Based on the properties of $\mathrm{Re}_{R,\varepsilon}$ from Proposition 3 (see also Equation (2.3) noting that $h=\frac{\varepsilon}{2+2R}$ in that equation), it follows that $\varrho^{\mathrm{Re},c}_{R,\varepsilon}$ is implemented by a depth-3 σ-network with 6 neurons and 11 weights (see Figure 3), all bounded in absolute value by $2\,|c|+C\cdot R^3/\varepsilon^3$. The construction of $\varrho^{\mathrm{Im},c}_{R,\varepsilon}$ is similar, replacing $\mathrm{Re}_{R,\varepsilon}(z)$ with $\mathrm{Im}_{R,\varepsilon}(z)$. ∎

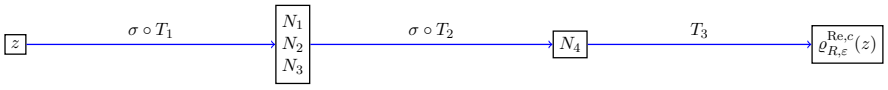

Figure 3. *Architecture of the network $\varrho^{\mathrm{Re},c}_{R,\varepsilon}$, where $T_1(\cdot)=A_1(\cdot)+b_1$ for $A_1=\left(\frac{\varepsilon}{1+R},\frac{\varepsilon}{2(1+R)},\frac{\varepsilon}{2(1+R)}\right)^T$ and $b_1=\left(\frac{(8-4i)(1+R)}{\varepsilon}+2,\frac{(4-2i)(1+R)}{\varepsilon}+1,\frac{-2i(1+R)}{\varepsilon}\right)^T$, and $T_2(\cdot)=A_2(\cdot)+b_2$ for $A_2=\frac{2(1+R)^2}{(R+|c|)\varepsilon^2}(1,-1,-1)$ and $b_2=\frac{2(1+R)^2}{(R+|c|)\varepsilon^2}\left(-1-\frac{4(1+R)}{\varepsilon}+i\right)+1+\frac{c}{2(R+|c|)}$, and finally $T_3(z)=2(R+|c|)\cdot z$.*

The next lemma shows that $\sigma:\mathbb{C}\to\mathbb{C}$ is 1-Lipschitz, which was used in the proof above.

Lemma 6. *The modReLU function $\sigma:\mathbb{C}\to\mathbb{C}$ defined in Equation (1.1) is 1-Lipschitz, i.e., $|\sigma(z)-\sigma(w)|\leq|z-w|$ for all $z,w\in\mathbb{C}$.*

Proof. Simply note that

$$\big|\sigma(z)-\sigma(w)\big|=\begin{cases}0 & \text{if } |z|,|w|\leq 1,\\ |\sigma(z)|=|z|-1\leq|z|-|w|\leq|z-w| & \text{if } |z|>1 \text{ and } |w|\leq 1,\\ |\sigma(w)|=|w|-1\leq|w|-|z|\leq|w-z| & \text{if } |z|\leq 1 \text{ and } |w|>1,\\ \big|(z-\frac{z}{|z|})-(w-\frac{w}{|w|})\big|\leq|z-w| & \text{if } |z|,|w|>1,\end{cases}$$

where we used that if $z,w\in\mathbb{C}$ with $|z|,|w|>1$, then

$$\begin{aligned}\big|(z-\tfrac{z}{|z|})-(w-\tfrac{w}{|w|})\big|^2&=\big|\tfrac{z}{|z|}\cdot(|z|-1)-\tfrac{w}{|w|}\cdot(|w|-1)\big|^2\\&=(|z|-1)^2+(|w|-1)^2-2(|z|-1)(|w|-1)\,\mathrm{Re}\left(\tfrac{z}{|z|}\,\tfrac{\overline{w}}{|w|}\right)\\&=|z|^2+|w|^2-2|z||w|\,\mathrm{Re}\left(\tfrac{z}{|z|}\,\tfrac{\overline{w}}{|w|}\right)-\big(2|z|+2|w|-2\big)\big(1-\mathrm{Re}\left(\tfrac{z}{|z|}\,\tfrac{\overline{w}}{|w|}\right)\big)\\&\leq|z|^2+|w|^2-2|z||w|\,\mathrm{Re}\left(\tfrac{z}{|z|}\,\tfrac{\overline{w}}{|w|}\right)=|z-w|^2.\end{aligned}$$

∎

Our next goal is to construct σ-networks approximating the function $z \mapsto (\operatorname{Re} z)^2$. This will be based on combining Proposition 5 with the approximation of the real function $x \mapsto x^2$ by ReLU networks, as presented in [31].

The construction in [31] is based on the following auxiliary functions, depicted in Figure 4:

$$\begin{aligned} g:&\quad \mathbb{R}\to\mathbb{R}, \quad g(x) &&:= 2\varrho(x) - 4\varrho(x-\tfrac{1}{2}) + 2\varrho(x-1),\\ g_k:&\quad \mathbb{R}\to\mathbb{R}, \quad g_k(x) &&:= \underbrace{g\circ\cdots\circ g}_{k-\text{times}}(x) \quad \text{for } k\in\mathbb{N},\end{aligned}$$

$$f_m:\quad \mathbb{R}\to\mathbb{R}, \quad f_m(x) := x - \sum_{k=1}^{m} \frac{g_k(x)}{2^{2k}} \quad \text{for } m\in\mathbb{N}\cup\{0\}.$$

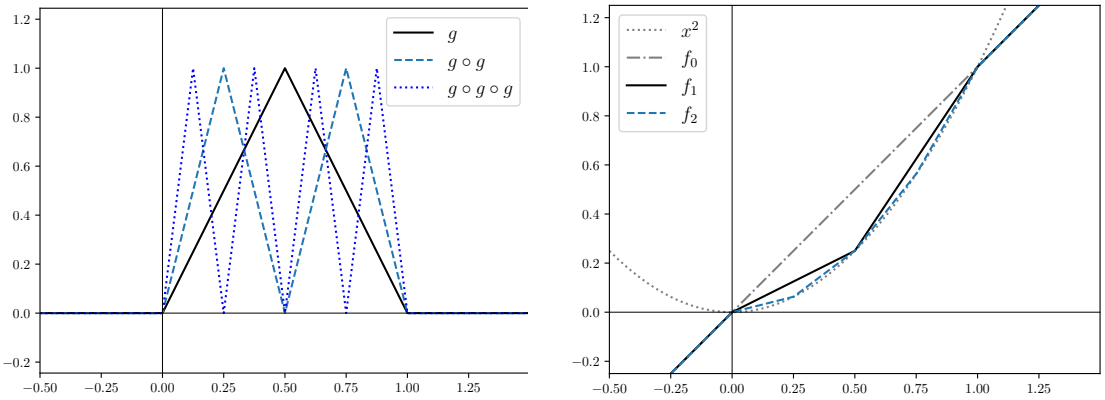

Figure 4. *A plot of the function g, its compositions $g\circ g$ and $g\circ g\circ g$, and the square approximations f_m, for $m = 0, 1, 2$. Evidently, g is 2-Lipschitz.*

One can show (cf. [31, Proof of Proposition 2]) that

$$\big|x^2 - f_m(x)\big| \le 2^{-2m-2} \quad \text{for } 0\le x\le 1. \tag{3.2}$$

Further, we define

$$\begin{aligned} g^{\mathrm{Re}}:&\quad \mathbb{C}\to\mathbb{C}, \quad g^{\mathrm{Re}}(z) &&:= g\big(\operatorname{Re}(z)\big) = 2\varrho\big(\operatorname{Re}(z)\big) - 4\varrho\big(\operatorname{Re}(z)-\tfrac{1}{2}\big) + 2\varrho\big(\operatorname{Re}(z)-1\big),\\ g^{\mathrm{Re},k}:&\quad \mathbb{C}\to\mathbb{C}, \quad g^{\mathrm{Re},k}(z) &&:= \underbrace{g^{\mathrm{Re}}\circ\cdots\circ g^{\mathrm{Re}}}_{k-\text{times}}(z) = g_k\big(\operatorname{Re}(z)\big) \quad (\text{since } g:\mathbb{R}\to\mathbb{R}).\end{aligned}$$

As is clear from Figure 4, the function $g:\mathbb{R}\to\mathbb{R}$ is 2-Lipschitz, which in turn implies that $g^{\mathrm{Re}}:\mathbb{C}\to\mathbb{C}$ is 2-Lipschitz; indeed, $\big|g(\operatorname{Re}(z)) - g(\operatorname{Re}(z'))\big| \le 2\,|\operatorname{Re}(z) - \operatorname{Re}(z')| \le 2\,|z-z'|$ for $z, z'\in\mathbb{C}$. In view of Proposition 5, we consider the approximation of g^{Re} and $g^{\mathrm{Re},k}$ respectively by the following functions:

$$\begin{aligned} g^{\mathrm{Re}}_{R,\varepsilon}:&\quad \mathbb{C}\to\mathbb{C}, \quad g^{\mathrm{Re}}_{R,\varepsilon}(z) &&:= 2\varrho^{\mathrm{Re},0}_{R,\varepsilon}(z) - 4\varrho^{\mathrm{Re},-1/2}_{R,\varepsilon}(z) + 2\varrho^{\mathrm{Re},-1}_{R,\varepsilon}(z),\\ g^{\mathrm{Re},k}_{R,\varepsilon}:&\quad \mathbb{C}\to\mathbb{C}, \quad g^{\mathrm{Re},k}_{R,\varepsilon}(z) &&:= \underbrace{g^{\mathrm{Re}}_{R,\varepsilon}\circ\cdots\circ g^{\mathrm{Re}}_{R,\varepsilon}}_{k-\text{times}}(z) \quad \text{for } k\in\mathbb{N},\end{aligned}$$

where $R \geq 1$ and $\varepsilon \in (0,1)$. As the last preparation for the proof of Proposition 8, we need the following technical lemma concerning the size of $g^{\mathrm{Re},k}_{R,\varepsilon}(z)$.

Lemma 7. *Let $R \geq 1$, $0 < \varepsilon < \min\{1, \frac{R}{8}\}$, and $k \in \mathbb{N}$. Then $\left|g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)\right| \leq R+1$ for all $z \in \mathbb{C}$ with $|z| \leq R+1$.*

Proof. Proposition 5 implies that if $z \in \mathbb{C}$ with $|z| \leq R+1$, then

$$\left|g^{\mathrm{Re}}_{R+1,\varepsilon}(z) - g^{\mathrm{Re}}(z)\right| \leq (2+4+2)\cdot\varepsilon \leq R$$

and since $|g^{\mathrm{Re}}(z)| = |g(\mathrm{Re}(z))| \leq 1$ for all $z \in \mathbb{C}$ (see Figure 4), we have

$$|g^{\mathrm{Re}}_{R+1,\varepsilon}(z)| \leq \left|g^{\mathrm{Re}}_{R+1,\varepsilon}(z) - g^{\mathrm{Re}}(z)\right| + \left|g^{\mathrm{Re}}(z)\right| \leq R+1.$$

This shows that $g^{\mathrm{Re}}_{R+1,\varepsilon}$ maps $\{z \in \mathbb{C} : |z| \leq R+1\}$ into itself. It then follows by induction that $g^{\mathrm{Re},k}_{R+1,\varepsilon}$ maps $\{z \in \mathbb{C} : |z| \leq R+1\}$ into itself, as claimed. ∎

Proposition 8. *Let $R \geq 3$ and $0 < \varepsilon < \min\{1, \frac{R}{8}\}$. There exists a function $\Phi_{R,\varepsilon} : \mathbb{C} \to \mathbb{C}$ that is implemented by a σ-network of depth $\mathcal{O}(\ln(2/\varepsilon))$ and width $\mathcal{O}(\ln(2/\varepsilon))$ and with the number of weights and neurons bounded by $\mathcal{O}(\ln^2(2/\varepsilon))$ and all weights bounded by $\mathcal{O}(R^6/\varepsilon^7)$ and such that $|(\mathrm{Re}\, z)^2 - \Phi_{R,\varepsilon}(z)| \leq \varepsilon$ for all $z \in \mathbb{C}$ with $|z| \leq R$ and $|\mathrm{Re}\, z| \leq 1$.*

Proof. First, it holds for any $z \in \mathbb{C}$ with $|z| \leq R+1$ that

$$\left|g^{\mathrm{Re}}(z) - g^{\mathrm{Re}}_{R+1,\varepsilon}(z)\right| \leq (2+4+2)\cdot\varepsilon = 8\varepsilon \tag{3.3}$$

and $\left|g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)\right| \leq R+1$ for all $k \in \mathbb{N}$, by Proposition 5 and Lemma 7 respectively. We claim that this implies

$$|g^{\mathrm{Re},k}(z) - g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)| \leq 8\varepsilon\cdot(2^k-1) \qquad \forall\, k \in \mathbb{N} \text{ and } |z| \leq R+1. \tag{3.4}$$

Indeed, for $k = 1$ we have $g^{\mathrm{Re},k}_{R+1,\varepsilon} = g^{\mathrm{Re}}_{R+1,\varepsilon}$ and $g^{\mathrm{Re},k} = g^{\mathrm{Re}}$, so that Equation (3.3) shows $|g^{\mathrm{Re},k}(z) - g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)| \leq 8\varepsilon = 8\varepsilon\cdot(2^k-1)$.

Next, suppose Equation (3.4) holds for some $k \in \mathbb{N}$. Lemma 7 shows that $w := g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)$ satisfies $|w| \leq R+1$. Further, setting $w' := g^{\mathrm{Re},k}(z)$, Equation (3.4) shows $|w-w'| \leq 8\varepsilon\cdot(2^k-1)$. Thus, using that g^{Re} is 2-Lipschitz, we see

$$\begin{aligned}
|g^{\mathrm{Re},k+1}(z) - g^{\mathrm{Re},k+1}_{R+1,\varepsilon}(z)| &= |g^{\mathrm{Re}}(w') - g^{\mathrm{Re}}_{R+1,\varepsilon}(w)| \\
&\leq |g^{\mathrm{Re}}(w') - g^{\mathrm{Re}}(w)| + |g^{\mathrm{Re}}(w) - g^{\mathrm{Re}}_{R+1,\varepsilon}(w)| \\
&\overset{(*)}{\leq} 2\cdot 8\varepsilon\cdot(2^k-1) + 8\varepsilon = 8\varepsilon\cdot(2^{k+1}-1)
\end{aligned}$$

where Equation (3.3) was used at $(*)$. Thus, Equation (3.4) holds for $k+1$ if it holds for k.

Now, using the function Id_R from Equation (2.1) (which is implemented by a 2-layer σ-network with 7 weights, all bounded by $2R+2$) and the function $\mathrm{Re}_{R,\varepsilon}$ from Proposition 3, we define for $m \in \mathbb{N}$,

$$f_{m,R,\varepsilon}(z) := \mathrm{Re}_{R,\varepsilon} \circ \mathrm{Id}_R \circ \cdots \circ \mathrm{Id}_R(z) - \sum_{k=1}^{m} \frac{g^{\mathrm{Re},k}_{R+1,\varepsilon} \circ \mathrm{Id}_R \circ \cdots \circ \mathrm{Id}_R(z)}{4^k} \quad \text{for } z \in \mathbb{C},$$

where the number of the "factors" Id_R is chosen such that all (sub)networks have the same depth and thus can be added/subtracted—see Appendices A.2 and A.3 for details on implementing composition and summation of networks. It then follows for $m \in \mathbb{N}$ and $|z| \leq R$ that

$$\begin{aligned} |f_m(\mathrm{Re}(z)) - f_{m,R,\varepsilon}(z)| &\leq |\mathrm{Re}(z) - \mathrm{Re}_{R,\varepsilon}(z)| + \sum_{k=1}^{m} \frac{|g^{\mathrm{Re},k}(z) - g^{\mathrm{Re},k}_{R+1,\varepsilon}(z)|}{4^k} \\ &\leq \varepsilon + 8\varepsilon \sum_{k=1}^{m} 2^{-k} \leq 9\varepsilon. \end{aligned} \tag{3.5}$$

Setting $m := \lceil \frac{1}{2}\ln(\frac{1}{\varepsilon})/\ln(2) \rceil \in \mathbb{N}$ (so that $2^{-2m-2} \leq \varepsilon$) and combining (3.2) and (3.5), we deduce for $|z| \leq R$ with $0 \leq \mathrm{Re}(z) \leq 1$ that

$$|\mathrm{Re}(z)^2 - f_{m,R,\varepsilon}(z)| \leq |\mathrm{Re}(z)^2 - f_m(\mathrm{Re}(z))| + |f_m(\mathrm{Re}(z)) - f_{m,R,\varepsilon}(z)| \leq 10\varepsilon. \tag{3.6}$$

We will now extend this result to $z \in \mathbb{C}$ with $|z| \leq R$ and $|\mathrm{Re}(z)| \leq 1$. Given such a z, define $w := \frac{1}{2}(z+1)$, noting that $|w| \leq R$ (since $R \geq 1$) and $0 \leq \mathrm{Re}\, w \leq 1$. Therefore, applying Equation (3.6) to w instead of z, we see $|(\mathrm{Re}\, w)^2 - f_{m,R,\varepsilon}(w)| \leq 10\varepsilon$. Note that $(\mathrm{Re}\, w)^2 = \frac{1}{4}(1 + \mathrm{Re}\, z)^2 = \frac{1}{4} + \frac{1}{2}\mathrm{Re}\, z + \frac{1}{4}(\mathrm{Re}\, z)^2$ and hence $(\mathrm{Re}\, z)^2 = 4\,(\mathrm{Re}\, w)^2 - 2\,\mathrm{Re}\, z - 1$. Thus, setting

$$h_{m,R,\varepsilon}(z) := 4\, f_{m,R,\varepsilon}\big(\tfrac{1}{2}(z+1)\big) - 2\,\mathrm{Re}_{R,\varepsilon} \circ \mathrm{Id}_R \circ \cdots \circ \mathrm{Id}_R(z) - 1,$$

where again Id_R is used to match the depth of the (sub)networks, we see

$$|(\mathrm{Re}\, z)^2 - h_{m,R,\varepsilon}(z)| \leq 4 \cdot |(\mathrm{Re}\, w)^2 - f_{m,R,\varepsilon}(w)| + 2 \cdot |\mathrm{Re}(z) - \mathrm{Re}_{R,\varepsilon}(z)| \leq 42\varepsilon.$$

It remains to bound the depth, width, and number of weights of the σ-network defining the function $\Phi_{R,\varepsilon}(z) := h_{m,R,\varepsilon}(z)$, and to estimate the size of the weights. The following estimates regarding these quantities should be fairly intuitive; the reader interested in the full details is referred to Appendices A.2 and A.3. Note that $f_{m,R,\varepsilon}$, with our choice of $m = \lceil \frac{1}{2}\ln(\frac{1}{\varepsilon})/\ln(2) \rceil = \mathcal{O}(\ln(2/\varepsilon))$, is a σ-network with depth and width $\mathcal{O}(m)$, and with $\mathcal{O}(m^2)$ neurons and weights, all of which are bounded by $\mathcal{O}(\ln(2/\varepsilon)R^6/\varepsilon^6) \subset \mathcal{O}(R^6/\varepsilon^7)$. Consequently, $\Phi_{R,\varepsilon}$ is a σ-network whose depth and width is $\mathcal{O}(\ln(2/\varepsilon))$, whose number of weights, and neurons are $\mathcal{O}(\ln^2(2/\varepsilon))$, and whose weights are bounded by $\mathcal{O}(R^6/\varepsilon^7)$. ∎

4. Approximating the product of complex numbers.

In this section, we approximate the map $\mathbb{C}^2 \to \mathbb{C}, (z,w) \mapsto zw$ using modReLU networks. To do so, we first approximate the function $\mathbb{C}^2 \to \mathbb{C}, (z,w) \mapsto \mathrm{Re}(z)\,\mathrm{Re}(w)$ based on the approximation of $(\mathrm{Re}\, z)^2$ from Proposition 8 and then use a polarization argument. This idea is motivated by [31, Proposition 3].

Proposition 9. *Given $R \geq 3$ and $\varepsilon \in (0,1)$, there is a function $\widetilde{\times}_{\mathrm{Re},R,\varepsilon} \colon \mathbb{C}^2 \to \mathbb{C}$ such that*

1. *for any inputs $z, w \in \mathbb{C}$ with $|z|, |w| \leq R$ we have $|\widetilde{\times}_{\mathrm{Re},R,\varepsilon}(z,w) - \mathrm{Re}(z)\,\mathrm{Re}(w)| \leq \varepsilon$;*
2. *the function $\widetilde{\times}_{\mathrm{Re},R,\varepsilon}$ is implemented by a σ-network of depth and width $\mathcal{O}\big(\ln(R^2\,\varepsilon^{-1})\big)$, with at most $\mathcal{O}\big(\ln^2(R^2\,\varepsilon^{-1})\big)$ weights and neurons, and all weights bounded in absolute value by $\mathcal{O}\big(R^{16}\varepsilon^{-7}\big)$.*

Proof. Define $R' := 3$ and note that $0 < \varepsilon' := \frac{\varepsilon}{6R^2} < \frac{1}{54} \leq \min\{1, \frac{R'}{8}\}$. Therefore, we can apply Proposition 8 with R', ε' instead of R, ε, which produces the function $\Phi_{R',\varepsilon'} = \Phi_{3,\frac{\varepsilon}{6R^2}}$. We then set

$$\widetilde{\times}_{\mathrm{Re},R,\varepsilon}(z,w) := 2R^2 \cdot \Big(\Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{z+w}{2R}) - \Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{z}{2R}) - \Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{w}{2R})\Big) \quad \text{for } z,w \in \mathbb{C}.$$

Comparing with the equation

$$\mathrm{Re}(z)\,\mathrm{Re}(w) = 2R^2 \cdot \Big([\mathrm{Re}(\tfrac{z+w}{2R})]^2 - [\mathrm{Re}(\tfrac{z}{2R})]^2 - [\mathrm{Re}(\tfrac{w}{2R})]^2\Big) \quad \text{for } z,w \in \mathbb{C}$$

and applying Proposition 8, we see that if $z, w \in \mathbb{C}$ with $|z|, |w| \leq R$, then

$$\begin{aligned}
&\big|\mathrm{Re}(z)\,\mathrm{Re}(w) - \widetilde{\times}_{\mathrm{Re},R,\varepsilon}(z,w)\big| \\
&\leq 2R^2 \cdot \Big(\Big|[\mathrm{Re}(\tfrac{z+w}{2R})]^2 - \Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{z+w}{2R})\Big| + \Big|[\mathrm{Re}(\tfrac{z}{2R})]^2 - \Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{z}{2R})\Big| + \Big|[\mathrm{Re}(\tfrac{w}{2R})]^2 - \Phi_{3,\frac{\varepsilon}{6R^2}}(\tfrac{w}{2R})\Big|\Big) \\
&\leq 2R^2 \cdot \big(\tfrac{\varepsilon}{6R^2} + \tfrac{\varepsilon}{6R^2} + \tfrac{\varepsilon}{6R^2}\big) = \varepsilon.
\end{aligned}$$

According to Proposition 8, the function $\Phi_{R',\varepsilon'}$ is implemented by a σ-network of depth and width $\mathcal{O}\big(\ln(\frac{2}{\varepsilon'})\big)$, with $\mathcal{O}\big(\ln^2(\frac{2}{\varepsilon'})\big)$ weights and neurons, and all weights bounded by $\mathcal{O}((R')^6(\varepsilon')^{-7}))$. Consequently, the function $\Phi_{3,\frac{\varepsilon}{6R^2}}$ is implemented by a σ-network of depth and width $\mathcal{O}\big(\ln(12R^2\varepsilon^{-1})\big)$, with $\mathcal{O}\big(\ln^2(12R^2\varepsilon^{-1})\big)$ weights and neurons, and all weights bounded by $\mathcal{O}(R^{14}\varepsilon^{-7})$. Note that $\widetilde{\times}_{\mathrm{Re},R,\varepsilon}(z,w)$ is a parallel connection of three copies of $\Phi_{3,\frac{\varepsilon}{6R^2}}$ with the adjustment that all weights in the last layer are scaled by a factor of $2R^2$, and the first layer is composed with appropriate linear transforms. Hence, the function $\widetilde{\times}_{\mathrm{Re},R,\varepsilon}(z,w)$ is again implemented by a σ-network whose depth and width are $\mathcal{O}\big(\ln(R^2\,\varepsilon^{-1})\big)$, whose number of weights and neurons are $\mathcal{O}\big(\ln^2(R^2\,\varepsilon^{-1})\big)$, and whose weights are bounded by $\mathcal{O}\big(R^{16}\varepsilon^{-7}\big)$ in absolute value. ∎

As a direct consequence of Proposition 9, we obtain an approximation for the complex product function $\mathbb{C}^2 \to \mathbb{C}, (z,w) \mapsto zw$.

Corollary 10. *Given $R \geq 3$ and $\varepsilon \in (0,1)$, there is a function $\widetilde{\times}_{R,\varepsilon} \colon \mathbb{C}^2 \to \mathbb{C}$ such that*

1. *for any inputs $z, w \in \mathbb{C}$ with $|z|, |w| \leq R$ we have $|\widetilde{\times}_{R,\varepsilon}(z,w) - zw| \leq \varepsilon$;*
2. *the function $\widetilde{\times}_{R,\varepsilon}$ is implemented by a σ-network of depth and width $\mathcal{O}\big(\ln(R^2\,\varepsilon^{-1})\big)$, with at most $\mathcal{O}\big(\ln^2(R^2\,\varepsilon^{-1})\big)$ weights and neurons, and all weights bounded in absolute value by $\mathcal{O}\big(R^{16}\varepsilon^{-7}\big)$.*

Proof. Noting that for $z, w \in \mathbb{C}$,

$$\begin{aligned}
zw &= \mathrm{Re}(z)\,\mathrm{Re}(w) - \mathrm{Im}(z)\,\mathrm{Im}(w) + i\big(\mathrm{Re}(z)\,\mathrm{Im}(w) + \mathrm{Im}(z)\,\mathrm{Re}(w)\big) \\
&= \mathrm{Re}(z)\,\mathrm{Re}(w) - \mathrm{Re}(-iz)\,\mathrm{Re}(-iw) + i\big(\mathrm{Re}(z)\,\mathrm{Re}(-iw) + \mathrm{Re}(-iz)\,\mathrm{Re}(w)\big),
\end{aligned}$$

we define

$$\widetilde{\times}_{R,\varepsilon}(z,w) := \widetilde{\times}_{\mathrm{Re},R,\frac{\varepsilon}{4}}(z,w) - \widetilde{\times}_{\mathrm{Re},R,\frac{\varepsilon}{4}}(-iz,-iw) + i\big(\widetilde{\times}_{\mathrm{Re},R,\frac{\varepsilon}{4}}(z,-iw) + \widetilde{\times}_{\mathrm{Re},R,\frac{\varepsilon}{4}}(-iz,w)\big).$$

It then follows from Proposition 9 that $|\widetilde{\times}_{R,\varepsilon}(z,w) - zw| \leq \varepsilon$ for all $z, w \in \mathbb{C}$ with $|z|, |w| \leq R$. The function $\widetilde{\times}_{R,\varepsilon}(z,w)$ is a sum of four equivalent copies of $\widetilde{\times}_{\mathrm{Re},R,\frac{\varepsilon}{4}}$ and therefore is again implemented by a σ-network whose depth and width are $\mathcal{O}\big(\ln(R^2\,\varepsilon^{-1})\big)$, whose number of weights and neurons are $\mathcal{O}\big(\ln^2(R^2\,\varepsilon^{-1})\big)$, and whose weights are bounded by $\mathcal{O}\big(R^{16}\varepsilon^{-7}\big)$ in absolute value. ∎

5. Partition of unity. Define the functions $\psi^{\mathrm{Re}}, \psi^{\mathrm{Im}} : \mathbb{C} \to \mathbb{C}$ by

$$\psi^{\mathrm{Re}}(z) := 1 - \sigma(z + \tfrac{1}{2}) + \sigma(z - \tfrac{1}{2}) \quad \text{and} \quad \psi^{\mathrm{Im}}(z) := 1 + i\,\sigma(z + \tfrac{1}{2}i) - i\,\sigma(z - \tfrac{1}{2}i). \tag{5.1}$$

Note for $x \in \mathbb{R}$ that

$$\psi^{\mathrm{Re}}(x) = \begin{cases} 1 & \text{if } |x| \leq \frac{1}{2} \\ \frac{3}{2} - |x| & \text{if } \frac{1}{2} \leq |x| \leq \frac{3}{2} \\ 0 & \text{if } |x| > \frac{3}{2} \end{cases}$$

and for $z \in \mathbb{C}$ that ψ^{I}

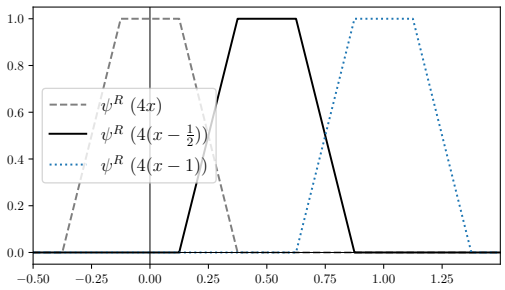

Figure 5. *A plot of the function $\psi^{\mathrm{Re}}(4\bullet)$ and its shifts, showing that they form a partition of unity.*

Let $N \geq 1$ be a natural number. For $m \in \{0, 1, \ldots, 2N\}$ define the functions $\phi^{\mathrm{Re}}_{m,N} : \mathbb{C} \to \mathbb{C}$ by $\phi^{\mathrm{Re}}_{m,N}(z) := \psi^{\mathrm{Re}}(4N(z - \frac{m}{2N}))$. It is not difficult to see that the $\phi^{\mathrm{Re}}_{m,N}$ $(m \in \{0, 1, \ldots, 2N\})$ form a partition of unity on the unit interval $[0,1] \subset \mathbb{R} \subset \mathbb{C}$, and that

$$\operatorname{supp}\big(\phi^{\mathrm{Re}}_{m,N}|_{\mathbb{R}}\big) \subset \big\{x \in \mathbb{R} \colon |x - \tfrac{m}{2N}| \leq \tfrac{3}{8}N^{-1}\big\};$$

see Figure 5. Similarly, defining $\phi^{\mathrm{Im}}_{m,N}(z) := \psi^{\mathrm{Im}}(4N(z - \frac{im}{2N}))$ for $m \in \{0, 1, \ldots, 2N\}$, we see that the $\phi^{\mathrm{Im}}_{m,N}$ form a partition of unity on the imaginary unit interval $i \cdot [0,1] \subset \mathbb{C}$.

6. Main result. In this section, we prove our main result, Theorem 1. As a preparation for the proof, we collect the following technical lemma, whose proof is deferred to Appendix A.4.

Lemma 11. *Let $\Omega \neq \varnothing$ be a set, $M \in \mathbb{N}$, $\varepsilon \in (0, \frac{1}{M+1})$, and $0 < \delta \leq \varepsilon^2$. Suppose that*

- *$\widetilde{\times} : \mathbb{C}^2 \to \mathbb{C}$ satisfies $|\widetilde{\times}(z,w) - zw| \leq \varepsilon$ for all $|z|, |w| \leq 4$;*
- *$\alpha_1, \ldots, \alpha_M : \Omega \to \mathbb{C}$ satisfy $|\alpha_j(z)| \leq 1$ for all $z \in \Omega$;*
- *$\beta_1, \ldots, \beta_M : \Omega \to \mathbb{C}$ satisfy $|\alpha_j(z) - \beta_j(z)| \leq \delta$ for all $z \in \Omega$.*

Define inductively $\gamma_1(z) := \beta_1(z)$ *and* $\gamma_{j+1}(z) := \widetilde{\times}\big(\beta_{j+1}(z), \gamma_j(z)\big)$ *for* $z \in \Omega$. *Then*

$$\Big|\gamma_M(z) - \prod_{\ell=1}^{M} \alpha_\ell(z)\Big| \leq 3M\,\varepsilon \qquad \forall\, z \in \Omega.$$

Proof of Theorem 1. As in Subsection 1.2, we identify the function $g : Q_{\mathbb{C}^d} \to \mathbb{C}$ with the pair of functions $g_{\mathrm{Re}}, g_{\mathrm{Im}} : [0,1]^{2d} \to \mathbb{R}$ and we will only explicitly show the approximation of $f := g_{\mathrm{Re}}$, since g_{Im} can be approximated in exactly the same way.

We roughly follow the structure of the proof of Theorem 1 in [31]: In the first step, we approximate f by f_*, a sum of Taylor polynomials subordinate to a partition of unity, constructed with our activation function σ in mind; see Section 5. In the second step we approximate f_* by the realization $\widetilde{f}$ of a σ-network of an appropriate architecture. An additional complication compared to the real setting considered in [31] is that we cannot access the real and imaginary parts of the inputs of f exactly with a σ network, but only approximatively; see Proposition 3.

Step 1. Employing similar notations to [31], we will denote ordered pairs (vectors) of coordinates by bold-faced characters. Given $N \in \mathbb{N}$ (specified precisely in Equation (6.1) below), let us write

$$\overline{N} := \{0, 1, \dots, 2N\}^d \times \{0, 1, \dots, 2N\}^d.$$

For $\boldsymbol{m} := (m_1, m_2, \dots, m_{2d}) \in \overline{N}$, we define on $Q_{\mathbb{C}^d} \cong [0,1]^{2d}$ the function

$$\phi_{\boldsymbol{m}}(\boldsymbol{x}) = \phi_{N,\boldsymbol{m}}(\boldsymbol{x}) = \prod_{k=1}^{d} \psi^{\mathrm{Re}}\Big(4N \cdot \big(x_k - \tfrac{m_k}{2N}\big)\Big) \prod_{\ell=d+1}^{2d} \psi^{\mathrm{Im}}\Big(4N \cdot \big(x_\ell\, i - \tfrac{i m_\ell}{2N}\big)\Big),$$

where $\boldsymbol{x} = (x_1, \dots, x_d, x_{d+1}, \dots, x_{2d})$ and where $\psi^{\mathrm{Re}}, \psi^{\mathrm{Im}}$ are given by Equation (5.1).

Based on the observations in Section 5, we see that the $\phi_{\boldsymbol{m}}$ ($\boldsymbol{m} \in \overline{N}$) form a partition of unity on $[0,1]^{2d}$ and satisfy $\operatorname{supp}(\phi_{\boldsymbol{m}}) \subset S_{\boldsymbol{m}}$ for the set

$$S_{\boldsymbol{m}} := \Big\{\boldsymbol{x} \in \mathbb{R}^{2d} : |x_k - \tfrac{m_k}{2N}| < \tfrac{1}{2N} \text{ and } |x_\ell - \tfrac{m_\ell}{2N}| < \tfrac{1}{2N} \text{ for all } 1 \leq k \leq d < \ell \leq 2d\Big\}.$$

Now for any $\boldsymbol{m} \in \overline{N}$, consider the Taylor polynomial of f at the point $\boldsymbol{x} = \frac{\boldsymbol{m}}{2N}$ of degree $n-1$, given by

$$P_{\boldsymbol{m}}(\boldsymbol{x}) = \sum_{\boldsymbol{n} \in \mathbb{N}_0^{2d}, |\boldsymbol{n}| < n} \left[\frac{\partial^{\boldsymbol{n}} f(\frac{\boldsymbol{m}}{2N})}{\boldsymbol{n}!} \cdot \left(\boldsymbol{x} - \frac{\boldsymbol{m}}{2N}\right)^{\boldsymbol{n}}\right], \qquad \text{and define} \qquad f_* := \sum_{\boldsymbol{m} \in \overline{N}} \phi_{\boldsymbol{m}} P_{\boldsymbol{m}}.$$

For any $\boldsymbol{x} \in [0,1]^{2d}$, we can bound the error by

$$\begin{aligned}
|f(\boldsymbol{x}) - f_*(\boldsymbol{x})| &= \Big|\sum_{\boldsymbol{m} \in \overline{N}} \phi_{\boldsymbol{m}}(\boldsymbol{x})\big(f(\boldsymbol{x}) - P_{\boldsymbol{m}}(\boldsymbol{x})\big)\Big| \leq \sum_{\boldsymbol{m}: \boldsymbol{x} \in S_{\boldsymbol{m}}} \big|f(\boldsymbol{x}) - P_{\boldsymbol{m}}(\boldsymbol{x})\big| \phi_{\boldsymbol{m}}(\boldsymbol{x}) \\
&\leq \max_{\boldsymbol{m}: \boldsymbol{x} \in S_{\boldsymbol{m}}} \big|f(\boldsymbol{x}) - P_{\boldsymbol{m}}(\boldsymbol{x})\big| \sum_{\boldsymbol{m} \in \overline{N}} \phi_{\boldsymbol{m}}(\boldsymbol{x}) \leq \frac{(2d)^n}{n!} \left(\frac{1}{N}\right)^n \|f\|_{\mathcal{W}^{n,\infty}} \\
&\leq \frac{(2d)^n}{n!} \left(\frac{1}{N}\right)^n,
\end{aligned}$$

where, similar to the arguments on Page 108 of [31], we used successively the fact that the $\phi_{\boldsymbol{m}}$ form a partition of unity and are supported on $S_{\boldsymbol{m}}$, a standard bound for the error of approximation by the Taylor polynomial (see e.g. the proof of [22, Lemma A.8]), and finally that f is in the unit ball of the Sobolev space, meaning $\|f\|_{\mathcal{W}^{n,\infty}} \leq 1$. Therefore by choosing

$$N := \left\lceil \left(\frac{n! \cdot \varepsilon}{2 \cdot (2d)^n} \right)^{-1/n} \right\rceil \tag{6.1}$$

(where $\lceil x \rceil$ is the smallest integer bigger or equal to x), we obtain that $\|f - f_*\|_{L^\infty} \leq \frac{\varepsilon}{2}$.

Step 2. We approximate f_* up to error $\frac{\varepsilon}{2}$ by a σ-network. To this end, note that we can rewrite f_* as

$$f_*(\boldsymbol{x}) = \sum_{\boldsymbol{m} \in \overline{N}} \sum_{\boldsymbol{n} \in \mathbb{N}_0^{2d} : |\boldsymbol{n}| < n} a_{\boldsymbol{m},\boldsymbol{n}} \cdot \phi_{\boldsymbol{m}}(\boldsymbol{x}) \cdot \left(\boldsymbol{x} - \tfrac{\boldsymbol{m}}{2N}\right)^{\boldsymbol{n}}.$$

Note that all the coefficients $a_{\boldsymbol{m},\boldsymbol{n}}$ have absolute value at most 1, since $\|f\|_{\mathcal{W}^{n,\infty}} \leq 1$. Therefore f_* is a linear combination of no more than $n^{2d}\,(2N+1)^{2d}$ terms of the form

$$f_{\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{x}) := \phi_{\boldsymbol{m}}(\boldsymbol{x}) \cdot \left(\boldsymbol{x} - \tfrac{\boldsymbol{m}}{2N}\right)^{\boldsymbol{n}}.$$

Fix $\boldsymbol{m} \in \overline{N}$ and $\boldsymbol{n} \in \mathbb{N}_0^{2d}$ with $|\boldsymbol{n}| < n$ for the moment. We want to approximate the function $f_{\boldsymbol{m},\boldsymbol{n}}$ via Lemma 11. Thus, set $S := n^{2d} \cdot |\overline{N}|$ and $M := 2d + |\boldsymbol{n}| < 2d + n$, and $\widetilde{\varepsilon} := \frac{\varepsilon}{6(2d+n) \cdot S}$, as well as $\delta := \widetilde{\varepsilon}^2$, and finally $\Omega := Q_{\mathbb{C}^d} = \left\{ \boldsymbol{z} \in \mathbb{C}^d \colon \operatorname{Re}(z_j), \operatorname{Im}(z_j) \in [0,1] \text{ for } j \in \{1, \dots, d\} \right\}$.

As a first step, we estimate $\widetilde{\varepsilon}$. Directly from Equation (6.1), we see

$$N \leq 1 + \left(\frac{2 \cdot (2d)^n}{n! \cdot \varepsilon} \right)^{1/n} \leq 1 + 4d \cdot \varepsilon^{-1} \leq 5d \cdot \varepsilon^{-1}.$$

Thus, $S = n^{2d}\,(2N+1)^{2d} \leq (12dn)^{2d} \cdot \varepsilon^{-2d}$, whence $\widetilde{\varepsilon}^{-1} \leq C_1(d,n) \cdot \varepsilon^{-2d-1} \leq C_1(d,n) \cdot \varepsilon^{-3d}$. Therefore, $\ln(2/\widetilde{\varepsilon}) \leq \ln(2C_1(d,n)) + 3d \ln(1/\varepsilon) \leq C_2(d,n) \cdot \ln(2/\varepsilon)$ for suitable constants $C_1(d,n) \geq 1$ and $C_2(d,n) \geq 1$.

Thus, Corollary 10 (applied with $\widetilde{\varepsilon}$ instead of ε) yields a function $\widetilde{\times} : \mathbb{C}^2 \to \mathbb{C}$ satisfying $|\widetilde{\times}(z,w) - z\,w| \leq \widetilde{\varepsilon}$ for all $z, w \in \mathbb{C}$ with $|z|, |w| \leq 4$, and such that $\widetilde{\times}$ is implemented by a σ-network with width and depth bounded by $C_3 \ln(2/\widetilde{\varepsilon}) \leq C_4 \cdot \ln(2/\varepsilon)$ and at most $C_3 \cdot \ln^2(2/\widetilde{\varepsilon}) \leq C_4 \cdot \ln^2(2/\varepsilon)$ weights, each bounded in absolute value by $C_3 \cdot \widetilde{\varepsilon}^{-7} \leq C_4 \cdot \varepsilon^{-21d}$. Here, $C_3 \geq 1$ is an absolute constant and $C_4 = C_4(d,n) \geq 1$.

Next, note that $\frac{8N}{\widetilde{\varepsilon}^2} \leq C_5(d,n) \cdot \varepsilon^{-7d}$. Therefore, Proposition 3 shows that there exist functions $\operatorname{Re}^\natural, \operatorname{Im}^\natural : \mathbb{C} \to \mathbb{C}$ with $|\operatorname{Re}^\natural(z) - \operatorname{Re}(z)| \leq \frac{\widetilde{\varepsilon}^2}{8N}$ and $|\operatorname{Im}^\natural(z) - \operatorname{Im}(z)| \leq \frac{\widetilde{\varepsilon}^2}{8N}$ for all $z \in \mathbb{C}$ with $|z| \leq 4$, and such that $\operatorname{Re}^\natural$ and $\operatorname{Im}^\natural$ are implemented by shallow σ-networks with 10 weights of magnitude at most $C_5\,(\frac{\widetilde{\varepsilon}^2}{8N})^{-3} \leq C_6 \cdot \varepsilon^{-21d}$, for suitable $C_6 = C_6(d,n) \geq 1$.

Finally, to apply Lemma 11, writing $\boldsymbol{n} = (n_1, \dots, n_{2d})$, we define $\alpha_k, \beta_k : \mathbb{C}^d \to \mathbb{C}$ for $1 \leq k \leq 2d + |\boldsymbol{n}| = M$ as follows:

- For $1 \leq k \leq d$, set

$$\alpha_k(\boldsymbol{z}) := \psi^{\mathrm{Re}}\big(4N \operatorname{Re}(z_k) - 2m_k\big) \quad \text{and} \quad \beta_k(\boldsymbol{z}) := \psi^{\mathrm{Re}}\big(4N \operatorname{Re}^\natural(z_k) - 2m_k\big);$$

- For $d+1 \leq k \leq 2d$, set

$$\alpha_k(\boldsymbol{z}) := \psi^{\mathrm{Im}}\big(4Ni\,\mathrm{Im}(z_{k-d}) - 2im_k\big) \quad \text{and} \quad \beta_k(\boldsymbol{z}) := \psi^{\mathrm{Im}}\big(4Ni\,\mathrm{Im}^\natural(z_{k-d}) - 2im_k\big);$$

- For $2d+n_1+\cdots+n_{\ell-1} < k \leq 2d+n_1+\cdots+n_\ell \leq 2d+n_1+\cdots+n_d$ $(1 \leq \ell \leq d)$, set

$$\alpha_k(\boldsymbol{z}) := \mathrm{Re}(z_\ell) - \frac{m_\ell}{2N} \quad \text{and} \quad \beta_k(\boldsymbol{z}) := \mathrm{Re}^\natural(z_\ell) - \frac{m_\ell}{2N};$$

- For $2d+n_1+\cdots+n_d \leq 2d+n_1+\cdots+n_{\ell-1} < k \leq 2d+n_1+\cdots+n_\ell$ $(d+1 \leq \ell \leq 2d)$, set

$$\alpha_k(\boldsymbol{z}) := \mathrm{Im}(z_{\ell-d}) - \frac{m_\ell}{2N} \quad \text{and} \quad \beta_k(\boldsymbol{z}) := \mathrm{Im}^\natural(z_{\ell-d}) - \frac{m_\ell}{2N}.$$

Since $\psi^{\mathrm{Re}}, \psi^{\mathrm{Im}}$ are 2-Lipschitz (this follows from the definition of $\psi^{\mathrm{Re}}, \psi^{\mathrm{Im}}$ and from Lemma 6), we see that $|\alpha_k(\boldsymbol{z}) - \beta_k(\boldsymbol{z})| \leq \widetilde{\varepsilon}^2 = \delta \leq 1$ for all $\boldsymbol{z} \in \Omega$ and $1 \leq k \leq M$. Furthermore, note that indeed $|\alpha_k(\boldsymbol{z})| \leq 1$ (and hence $|\beta_k(\boldsymbol{z})| \leq 2$) for all $\boldsymbol{z} \in \Omega$ and $1 \leq k \leq M$.

Overall, we can thus apply Lemma 11, which shows for

$$\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}} := \widetilde{\times}\big(\beta_1, \widetilde{\times}\big(\beta_2, \ldots, \widetilde{\times}(\beta_{2d+|\boldsymbol{n}|-1}, \beta_{2d+|\boldsymbol{n}|})\big)\big)$$

that

$$|f_{\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{z}) - \widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{z})| \leq 3M\,\widetilde{\varepsilon} \leq \frac{\varepsilon}{2S} = \frac{\varepsilon}{2}\frac{1}{n^{2d}|\overline{N}|} \qquad \forall\, \boldsymbol{z} \in \Omega = Q_{\mathbb{C}^d},$$

where we identify $\boldsymbol{z} = (z_1, \ldots, z_d) \in Q_{\mathbb{C}^d}$ with $\big(\mathrm{Re}(z_1), \ldots, \mathrm{Re}(z_d), \mathrm{Im}(z_1), \ldots, \mathrm{Im}(z_d)\big) \in [0,1]^{2d}$.

Thus, setting $\widetilde{f} := \sum_{\boldsymbol{m} \in \overline{N}} \sum_{\boldsymbol{n} \in \mathbb{N}_0^{2d}, |\boldsymbol{n}|<n} a_{\boldsymbol{m},\boldsymbol{n}} \widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ and recalling that $|a_{\boldsymbol{m},\boldsymbol{n}}| \leq 1$, we see for any $\boldsymbol{z} \in \Omega$ that

$$|f_*(\boldsymbol{z}) - \widetilde{f}(\boldsymbol{z})| \leq \sum_{\boldsymbol{m} \in \overline{N}} \sum_{\boldsymbol{n} \in \mathbb{N}_0^{2d}, |\boldsymbol{n}|<n} |a_{\boldsymbol{m},\boldsymbol{n}}||f_{\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{z}) - \widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{z})| \leq |\overline{N}| \cdot n^{2d} \cdot \frac{\varepsilon}{2}\frac{1}{n^{2d}|\overline{N}|} = \frac{\varepsilon}{2},$$

and hence $|f(\boldsymbol{z}) - \widetilde{f}(\boldsymbol{z})| \leq \varepsilon$ for all $\boldsymbol{z} \in Q_{\mathbb{C}^d}$, thanks to the bound from Step 1.

Step 3 (Size of the network): Note that $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ can be expressed as a composition of the networks Id_R (with $R=2$) and $\mathrm{Re}^\natural, \mathrm{Im}^\natural, \psi^{\mathrm{Re}}, \psi^{\mathrm{Im}}$, as well as $\widetilde{\times}$ (see Figure 6) and that the number of such subnetworks that appear in $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ depends only on the dimension d and the degree of smoothness n.

Next, note that with the implied constants (potentially) depending on d and n, the following hold:

- Id_R (with $R=2$) is implemented by a σ-network with $\mathcal{O}(1)$ weights and layers, and all weights bounded by $\mathcal{O}(1)$;
- $\mathrm{Re}^\natural$ and $\mathrm{Im}^\natural$ are implemented by σ-networks with $\mathcal{O}(1)$ weights and layers and all weights bounded by $\mathcal{O}(\varepsilon^{-21d})$;
- ψ^{Re} and ψ^{Im} are implemented by σ-networks with $\mathcal{O}(1)$ weights and layers and all weights bounded by $\mathcal{O}(1)$; hence, $\psi^{\mathrm{Re}}(4N \bullet - 2m)$ and $\psi^{\mathrm{Im}}(4iN \bullet - 2im)$ (for $m \in \{0,1,\ldots,2N\}$) are implemented by σ-networks with $\mathcal{O}(1)$ weights and layers and all weights bounded by $\mathcal{O}(N) \subset \mathcal{O}(\varepsilon^{-1})$;

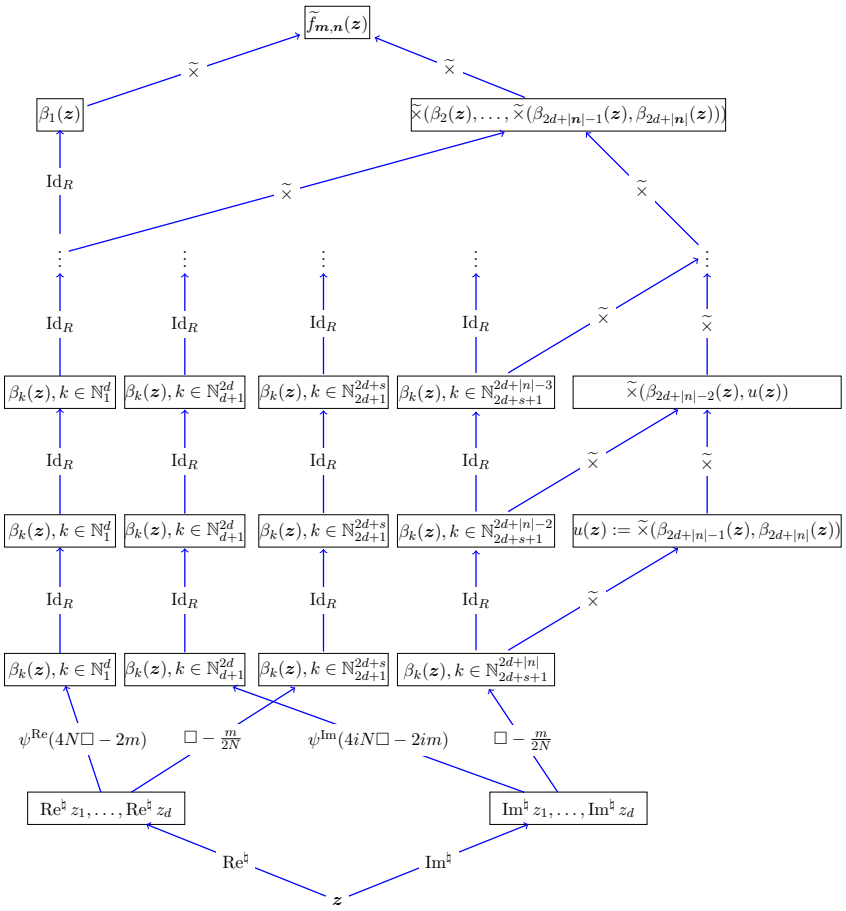

Figure 6. *Schematic of the architecture of the network implementing* $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$. *For brevity, the figure uses the notation* $\mathbb{N}_k^\ell := \{k, k+1, \dots, \ell-1, \ell\}$ *and* $s := n_1 + \cdots + n_d$ *as well as* $R := 2$.

- $\widetilde{\times}$ is implemented by a σ-network with depth and width bounded by $\mathcal{O}(\ln(2/\varepsilon))$ and $\mathcal{O}(\ln^2(2/\varepsilon))$ weights, bounded in absolute value by $\mathcal{O}(\varepsilon^{-21d})$.

This implies that $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ is implemented by a σ-network $\Phi^\varepsilon_{\boldsymbol{m},\boldsymbol{n}}$ satisfying $W(\Phi^\varepsilon_{\boldsymbol{m},\boldsymbol{n}}) \in \mathcal{O}(\ln^2(2/\varepsilon))$, $B(\Phi^\varepsilon_{\boldsymbol{m},\boldsymbol{n}}), L(\Phi^\varepsilon_{\boldsymbol{m},\boldsymbol{n}}) \in \mathcal{O}(\ln(2/\varepsilon))$, and $\|\Phi^\varepsilon_{\boldsymbol{m},\boldsymbol{n}}\| \in \mathcal{O}(\varepsilon^{-42d})$, where the implied constants (only) depend on d, n. For the full details, we refer to Appendices A.2 and A.3.

Now, since $\widetilde{f}$ is a linear combination of the no more than $n^{2d}(2N+1)^{2d}$ functions $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ with coefficients no larger in absolute value than 1 and recalling that $N \lesssim_{d,n} \varepsilon^{-1/n}$ (see Equation (6.1)), it follows that for some $C = C(d,n) > 0$ independent of ε and f, the function $\widetilde{f}$ is implemented by a σ-network with no more than $C \cdot \ln(2/\varepsilon)$ layers and no more than $C \cdot \varepsilon^{-2d/n} \cdot \ln^2(2/\varepsilon)$ weights, each bounded in absolute value by $C \cdot \varepsilon^{-44d}$.

Finally, note that $\widetilde{f} = \sum_{\boldsymbol{m} \in \overline{N}} \sum_{\boldsymbol{n} \in \mathbb{N}_0^{2d}, |\boldsymbol{n}| < n} a_{\boldsymbol{m},\boldsymbol{n}} \widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ where only the coefficients $a_{\boldsymbol{m},\boldsymbol{n}}$ depend on f, whereas the functions $\widetilde{f}_{\boldsymbol{m},\boldsymbol{n}}$ are independent of f. This easily implies that one

can choose a fixed network architecture $\mathcal{A}$ (only depending on d, n, ε but independent of f) with $L(\mathcal{A}) \leq C \cdot \ln(2/\varepsilon)$ and $W(\mathcal{A}) \leq C \cdot \varepsilon^{-2d/n} \cdot \ln^2(2/\varepsilon)$ such that $\widetilde{f}$ is implemented by a σ-network Φ_f of architecture $\mathcal{A}$ and with $\|\Phi_f\| \leq C \cdot \varepsilon^{-44d}$. ∎

7. Optimality.

In this section, we show that the approximation rate obtained in Theorem 1 cannot be improved significantly. Precisely, we prove the following result:

Theorem 12. *Let $d, n \in \mathbb{N}$, $\theta, \kappa, \gamma \geq 0$, and $C_0 \geq 1$. Assume that for every $\varepsilon \in (0,1)$ and every $g \in \mathcal{F}_{n,d}$ there exists a function g_ε implemented by a σ-network with at most $C_0 \cdot (\ln(2/\varepsilon))^\kappa$ layers and at most $C_0 \cdot \varepsilon^{-\gamma}$ weights, all bounded in absolute value by $C_0 \cdot \varepsilon^{-\theta}$ satisfying $\|g - g_\varepsilon\|_{L^\infty} \leq \frac{\varepsilon}{2}$.*

Then necessarily $\gamma \geq 2d/n$.

The proof idea consists in showing that the set of σ-networks of a given complexity satisfies certain *entropy bounds*. If the approximation rate from Theorem 1 could be strictly improved, this would then imply entropy bounds for the set $\mathcal{F}_{n,d}$ that contradict the known asymptotics of the entropy numbers of $\mathcal{F}_{n,d}$ [8].

We will derive the entropy bounds for the network sets as a consequence of the following Lipschitz bound for the realization map $\Phi \mapsto R_\sigma \Phi$. Since this bound is quite similar to the one in [6, Theorem 2.6]—although there only *real-valued* networks with the ReLU activation function are considered—we postpone the proof to Appendix A.5.

Lemma 13. *Given two networks $\Phi = \big((A_\ell, b_\ell)\big)_{\ell=1}^L$ and $\Psi = \big((B_\ell, c_\ell)\big)_{\ell=1}^L$ such that for each $\ell \in \{1, \dots, L\}$, the matrices A_ℓ, B_ℓ and vectors b_ℓ, c_ℓ have the same dimension, define $\Phi - \Psi$ to be the network $\big((A_\ell - B_\ell, b_\ell - c_\ell)\big)_{\ell=1}^L$.*

Let $R, R_0 \geq 1$ and assume $\|\Phi\|, \|\Psi\| \leq R$. Then, for every $z \in \mathbb{C}^{N_0}$ with $\|z\|_{\ell^\infty} \leq R_0$, we have

$$|R_\sigma \Phi(z) - R_\sigma \Psi(z)| \leq R_0 \cdot N_1 \cdots N_L \cdot 4^L R^{L-1} \cdot \|\Phi - \Psi\|.$$

As the final preparation for the proof of Theorem 12, we recall the notion of covering numbers and a few related facts. Given a non-empty subset $\varnothing \neq M \subset X$ of a metric space (X, d), the *covering number* $\mathrm{Cov}(M, \varepsilon) = \mathrm{Cov}_X(M, \varepsilon) \in \mathbb{N} \cup \{\infty\}$ is the minimal number $n \in \mathbb{N}$ of elements $m_1, \dots, m_n \in M$ satisfying $M \subset \bigcup_{i=1}^n \overline{B}_\varepsilon(m_i)$, where $\overline{B}_\varepsilon(m) = \{x \in X : d(x, m) \leq \varepsilon\}$ is the closed ball of radius ε around m.

It follows directly from the definitions that if $F : M \subset X \to Y$ is Lipschitz continuous with $\mathrm{Lip}(F) \leq L$ for some $L > 0$, then

$$\mathrm{Cov}_Y(F(M), \varepsilon) \leq \mathrm{Cov}_X(M, \varepsilon/L) \tag{7.1}$$

and that

$$\mathrm{Cov}\big(\textstyle\bigcup_{j=1}^n M_j, \varepsilon\big) \leq \sum_{j=1}^n \mathrm{Cov}(M_j, \varepsilon). \tag{7.2}$$

Using the identification $Q_{\mathbb{C}^d} \cong [0,1]^{2d}$, the following bound for the covering numbers of the set $\mathcal{F}_{n,d}$ (considered as a subset of $C(Q_{\mathbb{C}^d})$ with the sup-norm) is an easy consequence of [8, Theorem 3 and Theorem on Page 1086]:

Lemma 14. *For $d, n \in \mathbb{N}$ there exists a constant $C_1 = C_1(d, n) > 0$ satisfying*

$$\ln\big(\mathrm{Cov}_{C(Q_{\mathbb{C}^d})}(\mathcal{F}_{n,d}, \varepsilon)\big) \geq C_1 \cdot \varepsilon^{-2d/n} \qquad \forall\, \varepsilon \in (0, 1].$$

Furthermore, we will use the following bound for the covering numbers of subsets of $\mathbb{R}^n$, taken from [6, Lemma 2.7]:

Lemma 15. *Let $n \in \mathbb{N}$, $R \in [1,\infty)$ and $\varepsilon \in (0, e^{-1})$. Then, using the $\|\cdot\|_{\ell^\infty}$-norm on $\mathbb{R}^n$, we have*

$$\mathrm{Cov}\big([-R,R]^n, \varepsilon\big) \leq \exp\big(n \cdot \ln(\lceil R/\varepsilon \rceil)\big) \leq \exp\big(2n \cdot \ln(R/\varepsilon)\big) = (R/\varepsilon)^{2n}.$$

Using these preparations, we can finally prove Theorem 12.

Proof of Theorem 12. **Step 1:** Given $\varepsilon \in (0, e^{-1})$, set $W_\varepsilon := \lfloor C_0 \cdot \big(\frac{2}{\varepsilon}\big)^\gamma \rfloor$ and $R_\varepsilon := C_0 \cdot (\frac{2}{\varepsilon})^\theta$, as well as $L_\varepsilon := \big\lfloor C_0 \cdot \big(\ln(\frac{4}{\varepsilon})\big)^\kappa \big\rfloor$. Finally, let

$$\mathcal{NN}_\varepsilon := \big\{ R_\sigma \Phi|_{Q_{\mathbb{C}^d}} \ : \ d_{\mathrm{in}}(\Phi) = d, d_{\mathrm{out}}(\Phi) = 1, W(\Phi) \leq W_\varepsilon, L(\Phi) \leq L_\varepsilon \text{ and } \|\Phi\| \leq R_\varepsilon \big\}.$$

In this step, we show that

$$\ln\big(\mathrm{Cov}_{C(Q_{\mathbb{C}^d})}(\mathcal{NN}_\varepsilon, \varepsilon/2)\big) \leq C_1 \cdot \big(\ln(2/\varepsilon)\big)^{1+2\kappa} \cdot \varepsilon^{-\gamma} \tag{7.3}$$

for a suitable constant $C_1 = C_1(d, C_0, \kappa, \gamma, \theta) > 0$ independent of ε.

To see this, let us write $\underline{n} := \{1, \dots, n\}$ for $n \in \mathbb{N}$. Furthermore, given $L \in \underline{L_\varepsilon}$ and $\boldsymbol{N} = (N_1, \dots, N_L) \in \underline{W_\varepsilon}^L$ with $N_L = 1$ and given $\boldsymbol{J} = (J_1, \dots, J_L)$ with $J_\ell \subset \underline{N_\ell} \times \underline{N_{\ell-1}}$ and $|J_\ell| \leq W_\varepsilon$, define

$$\begin{aligned} \Lambda_{\boldsymbol{N},\boldsymbol{J}} : \ & \prod_{\ell=1}^{L} \Big([-R_\varepsilon, R_\varepsilon]^{J_\ell} \times [-R_\varepsilon, R_\varepsilon]^{N_\ell} \times [-R_\varepsilon, R_\varepsilon]^{J_\ell} \times [-R_\varepsilon, R_\varepsilon]^{N_\ell} \Big) \to C(Q_{\mathbb{C}^d}), \\ & \Big(\big(A^{(\ell)}_{j,k}\big)_{(j,k)\in J_\ell}, b^{(\ell)}, \big(B^{(\ell)}_{j,k}\big)_{(j,k)\in J_\ell}, c^{(\ell)} \Big)_{\ell=1}^{L} \mapsto R_\sigma\Big(\big(A^{(\ell)} + iB^{(\ell)}, b^{(\ell)} + ic^{(\ell)}\big)_{\ell=1}^{L} \Big), \end{aligned}$$

where $A^{(\ell)}_{j,k} = B^{(\ell)}_{j,k} = 0$ for $(j,k) \in (\underline{N_\ell} \times \underline{N_{\ell-1}}) \setminus J_\ell$.

We first claim that $\mathcal{NN}_\varepsilon \subset \bigcup_{\boldsymbol{N},\boldsymbol{J}} \mathrm{Im}(\Lambda_{\boldsymbol{N},\boldsymbol{J}})$, where the union is taken over all $\boldsymbol{N}, \boldsymbol{J}$ as above. To see this, note for $f \in \mathcal{NN}_\varepsilon$ that $f = R_\sigma \Phi$ for a network $\Phi = \big((A_\ell, b_\ell)\big)_{\ell=1,\dots,L}$ that satisfies $L \leq L_\varepsilon$, $\sum_{j=1}^{L} (\|A_j\|_{\ell^0} + \|b_j\|_{\ell^0}) = W(\Phi) \leq W_\varepsilon$, and $\|A_\ell\|_\infty, \|b_\ell\|_\infty \leq R_\varepsilon$. Since $\sigma(0) = 0$, it is easy to see simply by dropping "dead neurons" (i.e., neurons that always compute the value 0, independent of the network input) that one can assume $A_\ell \in \mathbb{C}^{N_\ell \times N_{\ell-1}}$ where $N_\ell \leq W_\varepsilon$ for $\ell \in \underline{L}$. Furthermore, the condition on the number of weights shows that for every $\ell \in \underline{L}$, one can choose a set $J_\ell \subset \underline{N_\ell} \times \underline{N_{\ell-1}}$ satisfying $|J_\ell| \leq W_\varepsilon$ and such that $(A_\ell)_{j,k} = 0$ unless $(j,k) \in J_\ell$. This easily implies $f = R_\sigma \Phi \in \mathrm{Im}(\Lambda_{\boldsymbol{N},\boldsymbol{J}})$.

Next, note for fixed $L \in \underline{L_\varepsilon}$ and $\boldsymbol{N} \in \underline{W_\varepsilon}^L$ that $N_\ell \leq d W_\varepsilon$ (even for $\ell = 0$) and hence

$$\begin{aligned} \big|\big\{ J_\ell \subset \underline{N_\ell} \times \underline{N_{\ell-1}} \colon |J_\ell| \leq W_\varepsilon \big\}\big| &= \sum_{t=0}^{\min\{W_\varepsilon, N_\ell N_{\ell-1}\}} \binom{N_\ell N_{\ell-1}}{t} \\ &\overset{(*)}{\leq} \big(e N_\ell N_{\ell-1} / \min\{W_\varepsilon, N_\ell N_{\ell-1}\}\big)^{\min\{W_\varepsilon, N_\ell N_{\ell-1}\}} \\ &\leq \big(d^2 e W_\varepsilon^2 / W_\varepsilon\big)^{\min\{W_\varepsilon, N_\ell N_{\ell-1}\}} \leq \big(d^2 e W_\varepsilon\big)^{W_\varepsilon}. \end{aligned} \tag{7.4}$$

Here, the step marked with $(*)$ used the elementary bound $\sum_{t=0}^{m} \binom{n}{t} \leq (en/m)^m$ which is valid for $1 \leq m \leq n$; see e.g. [27, Exercise 0.0.5].

As the next step, note that Lemma 13 shows that if we equip the domain of $\Lambda_{\boldsymbol{N},\boldsymbol{J}}$ with the $\|\bullet\|_{\ell^\infty}$ norm, then $\Lambda_{\boldsymbol{N},\boldsymbol{J}}$ is Lipschitz continuous with Lipschitz constant

$$\operatorname{Lip}(\Lambda_{\boldsymbol{N},\boldsymbol{J}}) \leq 2 \cdot (8 W_\varepsilon R_\varepsilon)^{L_\varepsilon} \leq \big(16 C_0^2 \cdot (2/\varepsilon)^{\gamma+\theta}\big)^{C_0 \, (\ln(4/\varepsilon))^\kappa} =: \Theta_\varepsilon.$$

Combining this with Equations (7.1), (7.2), and (7.4) and Lemma 15, we therefore see

$$\begin{aligned}
\operatorname{Cov}_{C(Q_{\mathbb{C}^d})}(\mathcal{NN}_\varepsilon, \varepsilon/2) &\leq \sum_{L,\boldsymbol{N},\boldsymbol{J}} \operatorname{Cov}_{C(Q_{\mathbb{C}^d})}\big(\operatorname{Im}(\Lambda_{\boldsymbol{N},\boldsymbol{J}}), \varepsilon/2\big) \\
&\leq \sum_{L,\boldsymbol{N},\boldsymbol{J}} \operatorname{Cov}\big([-R_\varepsilon, R_\varepsilon]^{\sum_{\ell=1}^{L}(2|J_\ell|+2W_\varepsilon)}, \tfrac{\varepsilon}{2\Theta_\varepsilon}\big) \\
&\leq L_\varepsilon \cdot W_\varepsilon^{L_\varepsilon} \cdot (d^2 e W_\varepsilon)^{W_\varepsilon} \cdot \Big(R_\varepsilon \frac{2\Theta_\varepsilon}{\varepsilon}\Big)^{8 L_\varepsilon W_\varepsilon}.
\end{aligned}$$

It is straightforward to see that $\ln(W_\varepsilon) \leq \ln(d^2 e W_\varepsilon) \lesssim \ln(2/\varepsilon)$ and $\ln(L_\varepsilon) \lesssim \ln(2/\varepsilon)$, as well as $\ln(R_\varepsilon) \lesssim \ln(2/\varepsilon)$ and $\ln(\Theta_\varepsilon) \lesssim \big(\ln(2/\varepsilon)\big)^{\kappa+1}$, where the implied constants only depend on $d, C_0, \gamma, \kappa, \theta$. In view of these estimates and because of $L_\varepsilon, W_\varepsilon \geq 1$, the preceding displayed equation shows

$$\ln\big(\operatorname{Cov}_{C(Q_{\mathbb{C}^d})}(\mathcal{NN}_\varepsilon, \varepsilon/2)\big) \lesssim L_\varepsilon \cdot W_\varepsilon \cdot \big(\ln(2/\varepsilon)\big)^{\kappa+1} \lesssim \big(\ln(2/\varepsilon)\big)^{1+2\kappa} \cdot \varepsilon^{-\gamma},$$

proving Equation (7.3).

Step 2 (Completing the proof): Let $\varepsilon \in (0, e^{-1})$. Equation (7.3) implies that there exists a constant $M_\varepsilon \in \mathbb{N}$ with $M_\varepsilon \leq \exp\big(C_1 \cdot (\ln(2/\varepsilon))^{1+2\kappa} \cdot \varepsilon^{-\gamma}\big)$ and $\mathcal{NN}_\varepsilon \subset \bigcup_{m=1}^{M_\varepsilon} \overline{B}_{\varepsilon/2}(g_m^{(\varepsilon)})$ for suitable $g_m^{(\varepsilon)} \in \mathcal{NN}_\varepsilon$, where $\overline{B}_\delta(g) := \{f \in C(Q_{\mathbb{C}^d}) \colon \|f - g\|_{L^\infty} \leq \delta\}$ is the closed ball of radius δ around g.

Now, for each $m \in \underline{M_\varepsilon}$ choose $h_m^{(\varepsilon)} \in \mathcal{F}_{n,d} \cap \overline{B}_\varepsilon(g_m^{(\varepsilon)})$ if this intersection is non-empty, and $h_m^{(\varepsilon)} = 0$ otherwise. By assumption of the theorem, for each $f \in \mathcal{F}_{n,d}$ there exists $g \in \mathcal{NN}_\varepsilon$ satisfying $\|f - g\|_{L^\infty} \leq \varepsilon/2$. Then, $\|f - g_m^{(\varepsilon)}\|_{L^\infty} \leq \frac{\varepsilon}{2} + \|g - g_m^{(\varepsilon)}\|_{L^\infty} \leq \varepsilon$ for a suitable $m \in \underline{M_\varepsilon}$, and hence $\|f - h_m^{(\varepsilon)}\|_{L^\infty} \leq \|f - g_m^{(\varepsilon)}\|_{L^\infty} + \|g_m^{(\varepsilon)} - h_m^{(\varepsilon)}\|_{L^\infty} \leq 2\varepsilon$.

Overall, this shows $\mathcal{F}_{n,d} \subset \bigcup_{m=1}^{M_\varepsilon} \overline{B}(h_m^{(\varepsilon)}, 2\varepsilon)$ and hence

$$\ln\big(\operatorname{Cov}(\mathcal{F}_{n,d}, 2\varepsilon)\big) \leq \ln(M_\varepsilon) \leq C_1 \cdot (\ln(2/\varepsilon))^{1+2\kappa} \cdot \varepsilon^{-\gamma} \qquad \forall\, \varepsilon \in (0, e^{-1}).$$

In view of Lemma 14, this is only possible if $\gamma \geq 2d/n$, which is what we wanted to show. ∎

8. Conclusion.

In the present paper, we studied the problem of approximating functions of regularity C^n defined on $\mathbb{C}^d$ using feed-forward complex-valued neural networks (CVNNs) with modReLU activation function. We showed that (ignoring logarithmic factors) a suitably constructed modReLU CVNN with $\mathcal{O}(\varepsilon^{-2d/n})$ parameters (weights) can achieve uniform approximation error ε. Moreover, we showed that this rate is near-optimal. This is as expected, since comparable real-valued neural networks obtain the same rates, cf. [31] (identifying $\mathbb{C} \simeq \mathbb{R}^2$).

Since it is known that ReLU neural networks achieve *optimal* approximation rates for C^n functions (see [7, 22, 31]), it cannot be expected that (modReLU) CVNNs *strictly improve* on ReLU networks, even in the complex setting. Rather, since CVNNs have been empirically observed to outperform real-valued neural networks in many applications involving complex-valued inputs [4,28], our goal is to initiate the study of the expressivity of CVNNs; furthermore, our goal was to rigorously prove that modReLU CVNNs can *match* the approximation capabilities of ReLU neural networks. Our results confirm that this is indeed the case.

The essential properties of the modReLU on which our proof relies are the following:

- modReLU CVNNs of a *constant size* can approximate the function $z \mapsto \operatorname{Re}(z)$ (and hence also the function $z \mapsto \operatorname{Im}(z)$) arbitrarily well; see Proposition 3;
- modReLU CVNNs of a *constant size* can approximate the "complexified" ReLU function $z \mapsto \varrho(\operatorname{Re}(z)) = \max\{0, \operatorname{Re}(z)\}$ arbitrarily well; see Proposition 5; and
- the modReLU is Lipschitz continuous.

Once these properties are known for a given activation function, the arguments used to prove our main theorem (which build upon the ideas in [31]), can be used to prove an analogous approximation result for that activation function. The Lipschitz continuity is used to control the propagation of errors among the layers of the network; it can probably be replaced by Hölder continuity and possibly even by uniform continuity. The main technical contribution of the paper is thus to verify that the above properties are satisfied for the modReLU and to show that these properties imply the main approximation result.

Appendix A. Postponed technical proofs.

A.1. Proof of Lemma 4.

Proof. Set $w := hz + \frac{1}{h}$. If $\operatorname{Im}(z) = 0$, then $w \in (0, \infty)$ and hence $\operatorname{sgn}(w) - 1 = 0 = \operatorname{Im}(z)$, so that the first part of Equation (2.2) is true. Hence, we can assume in what follows that $\operatorname{Im}(z) \neq 0$. Now, note by choice of h that $0 < h \leq \frac{1}{2}$ and $h\,|z| \leq \frac{1}{2}$, which shows that $|1 + zh^2| \geq 1 - h \cdot h\,|z| \geq \frac{3}{4}$ and therefore also $|w| = \frac{1}{h}|1 + h^2 z| \geq \frac{3}{4h} > 0$.

As a consequence, we obtain the estimate

$$\begin{aligned}\Big|\frac{1}{h^2}\frac{1}{|w|} - \frac{1}{h}\Big| &= \frac{1}{h} \cdot \Big|\frac{1}{h}\frac{1}{|h^{-1} + hz|} - 1\Big| = \frac{1}{h} \cdot \Big|\frac{1}{|1 + h^2 z|} - 1\Big| \\ &= \frac{1}{h} \cdot \frac{\big|1 - |1 + h^2 z|\big|}{|1 + h^2 z|} \leq \frac{1}{h}\frac{h^2|z|}{3/4} = \frac{4}{3}h\,|z| \leq \frac{2}{3} \leq 1.\end{aligned}$$

As $\operatorname{Im}(w) = h\,\operatorname{Im}(z) \neq 0$, this implies

$$\begin{aligned}\Big|\frac{1}{h^2}\operatorname{Im}(w/|w|) - \operatorname{Im}(z)\Big| &= h \cdot |\operatorname{Im}(z)| \cdot \Big|\frac{1}{h^2}\frac{\operatorname{Im}(w/|w|)}{\operatorname{Im}(w)} - \frac{1}{h}\Big| \\ &= h \cdot |\operatorname{Im}(z)| \cdot \Big|\frac{1}{h^2}\frac{1}{|w|} - \frac{1}{h}\Big| \leq h\,|z|.\end{aligned}$$

Next, note $\operatorname{Re}(w) = \frac{1}{h} \cdot \big(1 + h^2 \operatorname{Re}(z)\big) \geq \frac{1}{h} \cdot \big(1 - h \cdot h\,|z|\big) \geq \frac{1}{h} \cdot \big(1 - \frac{1}{2} \cdot \frac{1}{2}\big) = \frac{3}{4h} > 0$. Hence, $\operatorname{Re}(w) + |w| \geq 2\operatorname{Re}(w) \geq \frac{3}{2h}$. Since also $|\operatorname{Im}(w)| = h\,|\operatorname{Im}(z)| \leq h\,|z|$, we thus see

$$\big|\operatorname{Re}(w) - |w|\big| = \frac{\big|(\operatorname{Re}(w) - |w|)(\operatorname{Re}(w) + |w|)\big|}{\big|\operatorname{Re}(w) + |w|\big|} = \frac{\big|(\operatorname{Re}(w))^2 - |w|^2\big|}{\big|\operatorname{Re}(w) + |w|\big|} \leq \frac{|\operatorname{Im}(w)|^2}{\frac{3}{2h}}$$

and hence $\big|\operatorname{Re}(w) - |w|\big| \leq \frac{2}{3}h^3\,|z|^2$. Together with the estimate $|w| \geq \frac{3}{4h}$ from the beginning of the proof, we get

$$\frac{1}{h^2}\Big|\operatorname{Re}(w/|w|) - 1\Big| = \frac{1}{h^2}\left|\frac{\operatorname{Re}(w) - |w|}{|w|}\right| \leq \frac{1}{h^2}\frac{\frac{2}{3}h^3|z|^2}{\frac{3}{4h}} = \frac{4}{3}\frac{2}{3}h^2|z|^2 \leq h^2|z|^2.$$

Combining everything, we arrive at

$$\begin{aligned}
\Big|\operatorname{Im}(z) - \frac{-i}{h^2}\cdot\Big(\operatorname{sgn}(hz+\tfrac{1}{h}) - 1\Big)\Big| &= \Big|\frac{-i}{h^2}\cdot\Big(\frac{w}{|w|} - 1\Big) - \operatorname{Im}(z)\Big| \\
&\leq \Big|\frac{-i}{h^2}\cdot i\cdot\operatorname{Im}\Big(\frac{w}{|w|}\Big) - \operatorname{Im}(z)\Big| + \Big|\frac{-i}{h^2}\cdot\Big(\operatorname{Re}\Big(\frac{w}{|w|}\Big) - 1\Big)\Big| \\
&\leq \Big|\frac{1}{h^2}\operatorname{Im}(w/|w|) - \operatorname{Im}(z)\Big| + \frac{1}{h^2}\big|\operatorname{Re}(w/|w|) - 1\big| \\
&\leq h\,|z| + h^2|z|^2 \leq 2h\,|z|,
\end{aligned}$$

proving the first estimate in Equation (2.2). To prove the second estimate in Equation (2.2), simply note that $\operatorname{Re}(z) = \operatorname{Im}(iz)$ and $\operatorname{sgn}(iw) = i\operatorname{sgn}(w)$; hence, we get as claimed that

$$\begin{aligned}
2h\,|z| = 2h\,|iz| &\geq \Big|\operatorname{Im}(iz) - \frac{-i}{h^2}\cdot\big(\operatorname{sgn}(hiz+\tfrac{1}{h}) - 1\big)\Big| \\
&= \Big|\operatorname{Re}(z) - \frac{-i}{h^2}\cdot\big(i\operatorname{sgn}(hz-\tfrac{i}{h}) + i\,i\big)\Big| = \Big|\operatorname{Re}(z) - \frac{1}{h^2}\cdot\big(\operatorname{sgn}(hz-\tfrac{i}{h}) + i\big)\Big|.
\end{aligned}$$

■

A.2. Composition of neural networks. The composition of several neural networks is clearly again represented by a neural network. In this appendix we make this statement more precise, showing how the size of the resulting network is related to the size of the "input" networks. We note that the bounds for modReLU networks that we derive here are slightly worse than those derived for ReLU networks in [22, Section 2], owing to the fact that one can easily implement the identity function using the ReLU while this seems not to be possible (on all of $\mathbb{C}$) using the modReLU.

But first, we need some additional notation: Given a network $\Phi = \big((A_1, b_1), \dots, (A_L, b_L)\big)$, let us write $W_{\text{in}}(\Phi) := \|A_1\|_{\ell^0} + \|b_1\|_{\ell^0}$ and $W_{\text{out}}(\Phi) := \|A_L\|_{\ell^0} + \|b_L\|_{\ell^0}$ and furthermore $\|\Phi\|_{\text{in}} := \max\{\|A_1\|_\infty, \|b_1\|_\infty\}$ and $\|\Phi\|_{\text{out}} := \max\{\|A_L\|_\infty, \|b_L\|_\infty\}$. Now, assuming that $L \geq 2$ and given a further network $\Psi = \big((B_1, c_1), \dots, (B_M, c_M)\big)$ with $d_{\text{out}}(\Psi) = d_{\text{in}}(\Phi)$ and $M \geq 2$, define the composition of Φ, Ψ as

$$\Phi \bullet \Psi := \big((B_1, c_1), \dots, (B_{M-1}, c_{M-1}), (A_1 B_M, b_1 + A_1 c_M), (A_2, b_2), \dots, (A_L, b_L)\big). \tag{A.1}$$

It is straightforward to verify $R_\sigma(\Phi \bullet \Psi) = R_\sigma\Phi \circ R_\sigma\Psi$ and $L(\Phi \bullet \Psi) = L(\Phi) + L(\Psi) - 1$ and $B(\Phi \bullet \Psi) \leq \max\{B(\Phi), B(\Psi)\}$. The next lemma provides further bounds on the size of $\Phi \bullet \Psi$.

Lemma 16. *Let $\Phi^{(1)}, \dots, \Phi^{(K)}$ be neural networks of depth $L(\Phi^{(i)}) \geq 2$ for $i \in \{1, \dots, K\}$ and satisfying $d_{\text{out}}(\Phi^{(i)}) = d_{\text{in}}(\Phi^{(i+1)})$ for $i \in \{1, \dots, K-1\}$. Then the following hold:*

1. *If $W_{\text{in}}(\Phi^{(i)}), W_{\text{out}}(\Phi^{(i)}) \leq C$ for all $i \in \{1, \dots, K\}$ and some $C > 0$, then*

$$W\big(\Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}\big) \leq C^2 \cdot (K-1) + \sum_{i=1}^{K} W(\Phi^{(i)}). \tag{A.2}$$

2. *If* $\|\Phi^{(i)}\| \leq C$ *for all* $i \in \{1,\dots,K\}$ *and some* $C \geq 1$ *and* $d_{\text{out}}(\Phi^{(i)}) \leq D$ *for all* $i \in \{1,\dots,K\}$, *then*
$$\left\|\Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}\right\| \leq 2D \cdot C^2. \tag{A.3}$$
3. $L(\Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}) = \left[\sum_{i=1}^{K} L(\Phi^{(i)})\right] - (K-1)$ *and* $R_\sigma(\Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}) = R_\sigma \Phi^{(K)} \circ \cdots \circ R_\sigma \Phi^{(1)}$.

Remark 17. In particular, Equation (A.2) shows that if $B, W \geq 1$ and $W(\Phi^{(i)}) \leq W$ as well as $B(\Phi^{(i)}) \leq B$ for all $i \in \{1,\dots,K\}$, then $W(\Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}) \leq K \cdot (B^2 + W)$.

Proof. Before we prove the general case, we analyze the composition of two networks as in Equation (A.1). First, note for $A \in \mathbb{C}^{N \times K}$ and $B \in \mathbb{C}^{K \times P}$ that
$$\|AB\|_{\ell^0} = \sum_{i,j} \mathbb{1}_{(AB)_{i,j} \neq 0} \leq \sum_{i,j,\ell} \mathbb{1}_{A_{i,\ell} \neq 0} \mathbb{1}_{B_{\ell,j} \neq 0} \leq \sum_{i,\ell} \left(\mathbb{1}_{A_{i,\ell} \neq 0} \max_t \sum_j \mathbb{1}_{B_{t,j} \neq 0} \right) \leq \|A\|_{\ell^0} \|B\|_{\ell^0}.$$
A similar (but easier) calculation shows that $\|Av\|_{\ell^0} \leq \|A\|_{\ell^0}$ for $v \in \mathbb{C}^K$. Based on these estimates, we see (in the notation of Equation (A.1)) that
$$\|b_1 + A_1 c_M\|_{\ell^0} + \|A_1 B_M\|_{\ell^0} \leq \|b_1\|_{\ell^0} + \|A_1\|_{\ell^0} + \|A_1\|_{\ell^0} \|B_M\|_{\ell^0} \leq W_{\text{in}}(\Phi) \cdot (1 + W_{\text{out}}(\Psi)).$$
Directly from the definition of $\Phi \bullet \Psi$, we thus see
$$\begin{aligned} W(\Phi \bullet \Psi) &\leq W(\Psi) - W_{\text{out}}(\Psi) + W(\Phi) - W_{\text{in}}(\Phi) + W_{\text{in}}(\Phi)(1 + W_{\text{out}}(\Psi)) \\ &\leq W(\Psi) + W(\Phi) + W_{\text{in}}(\Phi) W_{\text{out}}(\Psi). \end{aligned} \tag{A.4}$$
Next, given $A \in \mathbb{C}^{N \times K}$ and $B \in \mathbb{C}^{K \times P}$ it is easy to see $|(AB)_{i,j}| \leq K \cdot \|A\|_\infty \|B\|_\infty$. Based on this, we see in the notation of Equation (A.1) that $\|A_1 B_M\|_\infty \leq d_{\text{out}}(\Psi)\, \|A_1\|_\infty \|B_M\|_\infty$ and $\|b_1 + A_1 c_M\|_\infty \leq \|b_1\|_\infty + d_{\text{out}}(\Psi) \|A_1\|_\infty \|c_M\|_\infty \leq \|\Phi\|_{\text{in}} \cdot (1 + d_{\text{out}}(\Psi) \|\Psi\|_{\text{out}})$. Thus, we see directly from the definition of $\Phi \bullet \Psi$ that
$$\|\Phi \bullet \Psi\| \leq \max\left\{\|\Phi\|,\ \|\Psi\|,\ \|\Phi\|_{\text{in}} \cdot (1 + d_{\text{out}}(\Psi)\, \|\Psi\|_{\text{out}})\right\}. \tag{A.5}$$

Now, we prove Equations (A.2) and (A.3) by induction on $K \in \mathbb{N}$. For $K = 1$ the claim is trivial. Next, assume that the claim holds for some $K \in \mathbb{N}$ and set $\Psi := \Phi^{(K)} \bullet \cdots \bullet \Phi^{(1)}$.

For proving Equation (A.2), note $W_{\text{in}}(\Phi^{(K+1)}) \leq C$ and $W_{\text{out}}(\Psi) = W_{\text{out}}(\Phi^{(K)}) \leq C$. Therefore, combining Equation (A.4) with the inductive assumption, we see
$$\begin{aligned} W(\Phi^{(K+1)} \bullet \Psi) &\leq W(\Phi^{(K+1)}) + W(\Psi) + W_{\text{in}}(\Phi^{(K+1)}) W_{\text{out}}(\Psi) \\ &\leq W(\Phi^{(K+1)}) + C^2 \cdot (K-1) + \sum_{i=1}^{K} W(\Phi^{(i)}) + C \cdot C \\ &= C^2 \cdot ((K+1) - 1) + \sum_{i=1}^{K+1} W(\Phi^{(i)}), \end{aligned}$$
completing the induction for Equation (A.2).

To prove Equation (A.3), note $\|\Psi\|_{\text{out}} = \|\Phi^{(K)}\|_{\text{out}} \leq C$ and $d_{\text{out}}(\Psi) = d_{\text{out}}(\Phi^{(K)}) \leq D$ and use Equation (A.5) and the inductive assumption to obtain
$$\|\Phi^{(K+1)} \bullet \Psi\| \leq \max\left\{\|\Phi^{(K+1)}\|,\ \|\Psi\|,\ \|\Phi^{(K+1)}\|_{\text{in}} \cdot (1 + d_{\text{out}}(\Psi)\, \|\Psi\|_{\text{out}})\right\} \leq 2D \cdot C^2,$$

completing the induction for Equation (A.3).

The last part of the lemma follows by induction after noting that $L(\Phi\bullet\Psi) = L(\Phi)+L(\Psi)-1$ and $R_\sigma(\Phi \bullet \Psi) = R_\sigma\Phi \circ R_\sigma\Psi$. ∎

A.3. Linear combinations of neural networks. In this appendix we show that the linear combinations of neural networks *of a common depth* can again be implemented as a neural network. Indeed, let $d, K \in \mathbb{N}$, and for each $j \in \{1,\dots,K\}$ let $a_j \in \mathbb{C}$ and let $\Phi^{(j)} = \big((A_1^{(j)}, b_1^{(j)}), \dots, (A_L^{(j)}, b_L^{(j)})\big)$ be a neural network with $d_{\text{in}}(\Phi^{(j)}) = d$ and $d_{\text{out}}(\Phi^{(j)}) = 1$ and of common depth $L(\Phi^{(j)}) = L$. Define $\Psi := \big((A_1, b_1), \dots, (A_L, b_L)\big)$, where

$$A_1 := \begin{pmatrix} A_1^{(1)} \\ \vdots \\ A_1^{(K)} \end{pmatrix}, \qquad A_L := \big(a_1 A_L^{(1)} \,\big|\, \cdots \,\big|\, a_K A_L^{(K)}\big), \qquad \text{and} \qquad b_\ell := \begin{pmatrix} b_\ell^{(1)} \\ \vdots \\ b_\ell^{(K)} \end{pmatrix}$$

for $\ell \in \{1,\dots,L-1\}$, as well as $A_\ell := \operatorname{diag}(A_\ell^{(1)}, \dots, A_\ell^{(K)})$ for $\ell \in \{2,\dots,L-1\}$ and $b_L := \sum_{j=1}^K a_j\, b_L^{(j)}$. It is easy to verify that

$$\begin{aligned} &R_\sigma\Psi = \sum_{j=1}^K a_j\, R_\sigma\Phi^{(j)}, \qquad L(\Psi) = L, \qquad W(\Psi) \le \sum_{j=1}^K W(\Phi^{(j)}), \\ &B(\Psi) \le \sum_{j=1}^K B(\Phi^{(j)}), \qquad \text{and} \qquad \|\Psi\| \le \sum_{j=1}^K (1+|a_j|)\big\|\Phi^{(j)}\big\|. \end{aligned} \tag{A.6}$$

Indeed, all except the first and final of these statements follow directly from the definitions. To verify the final statement, note by definition of $\|\bullet\|_\infty$ that

$$\begin{aligned} &\|A_1\|_\infty = \max_{j\in\{1,\dots,K\}} \big\|A_1^{(j)}\big\|_\infty \le \max_{j\in\{1,\dots,K\}} \big\|\Phi^{(j)}\big\| \\ \text{and} \quad &\|b_\ell\|_\infty = \max_{j\in\{1,\dots,K\}} \big\|b_\ell^{(j)}\big\|_\infty \le \max_{j\in\{1,\dots,K\}} \big\|\Phi^{(j)}\big\| \quad \text{for} \quad \ell \in \{1,\dots,L-1\}, \\ \text{as well as} \quad &\|A_L\|_\infty \le \max_{j\in\{1,\dots,K\}} |a_j| \,\big\|A_L^{(j)}\big\|_\infty \le \max_{j\in\{1,\dots,K\}} |a_j| \,\big\|\Phi^{(j)}\big\| \\ \text{and} \quad &\|b_L\|_\infty \le \sum_{j=1}^K |a_j| \,\|b_L^{(j)}\| \le \sum_{j=1}^K |a_j| \,\big\|\Phi^{(j)}\big\|, \end{aligned}$$

which implies as claimed that $\|\Psi\| \le \sum_{j=1}^K (1+|a_j|)\,\|\Phi^{(j)}\|$.

Finally, to verify the first statement, an induction with respect to $\ell \in \{1,\dots,L-1\}$ shows that if we set $T_\ell^{(j)} := A_\ell^{(j)}(\bullet)+b_\ell^{(j)}$ and $T_\ell := A_\ell(\bullet)+b_\ell$ and finally $F_\ell^{(j)} := (\sigma\circ T_\ell^{(j)})\circ\cdots\circ(\sigma\circ T_1^{(j)})$ and $F_\ell := (\sigma \circ T_\ell) \circ \cdots \circ (\sigma \circ T_1)$, then $F_\ell(z) = \big(F_\ell^{(1)}(z), \dots, F_\ell^{(K)}(z)\big)$ for $z \in \mathbb{C}^d$ and $\ell \in \{1,\dots,L-1\}$. Based on this, the first statement in Equation (A.6) follows from the definition of the realization map R_σ (see Subsection 1.1).

A.4. Proof of Lemma 11.

Proof. Define $\theta_j(z) := \prod_{\ell=1}^{j} \alpha_\ell(z)$ and $\kappa_j := \varepsilon \sum_{\ell=1}^{j} (1+\varepsilon)^\ell$. We will show inductively that $|\gamma_j(z) - \theta_j(z)| \leq \kappa_j$. This will imply the claim by taking $j = M$, since we have

$$(1+\varepsilon)^\ell \leq (1+\varepsilon)^{M+1} \leq \left(1 + \tfrac{1}{M+1}\right)^{M+1} \leq e \leq 3$$

and hence $\kappa_j \leq \kappa_M \leq 3M\,\varepsilon$.

The case $j = 1$ is trivial, since $|\gamma_1(z) - \theta_1(z)| = |\beta_1(z) - \alpha_1(z)| \leq \delta \leq \varepsilon^2 \leq \varepsilon \leq \kappa_1$. For the induction step, first note that

$$\begin{aligned}\kappa_j \leq \kappa_M = \varepsilon(1+\varepsilon) \sum_{\ell=0}^{M-1} (1+\varepsilon)^\ell = \varepsilon(1+\varepsilon)\frac{(1+\varepsilon)^M - 1}{(1+\varepsilon) - 1} \\ \leq (1+\varepsilon)^{M+1} \leq \left(1 + \tfrac{1}{M+1}\right)^{M+1} \leq e \leq 3\end{aligned}$$

and hence $|\gamma_j(z)| \leq |\theta_j(z)| + \kappa_j \leq 4$, since $|\alpha_\ell(z)| \leq 1$ for all ℓ, and thus $|\theta_j(z)| \leq 1$. Since also $|\beta_{j+1}(z)| \leq \delta + |\alpha_{j+1}(z)| \leq 1 + \delta \leq 4$, we see by the properties of $\widetilde{\times}$ for any $z \in \Omega$ that

$$\begin{aligned}\big|\gamma_{j+1}(z) - \theta_{j+1}(z)\big| &\leq \Big|\widetilde{\times}\big(\beta_{j+1}(z), \gamma_j(z)\big) - \beta_{j+1}(z)\gamma_j(z)\Big| \\ &\quad + \big|\beta_{j+1}(z)\gamma_j(z) - \beta_{j+1}(z)\theta_j(z)\big| \\ &\quad + \big|\big(\beta_{j+1}(z) - \alpha_{j+1}(z)\big)\theta_j(z)\big| \\ &\leq \varepsilon + |\beta_{j+1}(z)| \cdot \kappa_j + \delta \cdot |\theta_j(z)| \\ &\leq \varepsilon + \delta + (1+\delta)\kappa_j \leq \varepsilon(1+\varepsilon) + (1+\varepsilon)\kappa_j,\end{aligned}$$

where the last step used that $\delta \leq \varepsilon^2 \leq \varepsilon$. Finally, note by choice of κ_j that

$$\varepsilon(1+\varepsilon) + (1+\varepsilon)\kappa_j = \varepsilon(1+\varepsilon) + \varepsilon \sum_{\ell=2}^{j+1} (1+\varepsilon)^\ell = \varepsilon \sum_{\ell=1}^{j+1} (1+\varepsilon)^\ell = \kappa_{j+1}.$$

This completes the induction and thus the proof. ∎

A.5. Proof of Lemma 13.

Proof. Set $\sigma_\ell := \sigma$ for $\ell \in \{1, \ldots, L-1\}$ and $\sigma_L := \mathrm{id}_{\mathbb{C}}$. It is easy to see in each case that $\sigma_\ell(0) = 0$; furthermore, Lemma 6 implies that each σ_ℓ is 1-Lipschitz. Now, inductively define $w_0 := v_0 := z$ as well as $w_{\ell+1} := \sigma_{\ell+1}(A_{\ell+1} w_\ell + b_{\ell+1})$ and $v_{\ell+1} := \sigma_{\ell+1}(B_{\ell+1} v_\ell + c_{\ell+1})$ for $\ell \in \{0, \ldots, L-1\}$. We then have $R_\sigma \Phi(z) = w_L$ and $R_\sigma \Psi(z) = v_L$. We will show inductively that $\|v_\ell\|_{\ell^\infty} \leq R_0 \cdot (2R)^\ell \cdot N_1 \cdots N_{\ell-1}$ and $\|v_\ell - w_\ell\|_{\ell^\infty} \leq R_0 \cdot N_1 \cdots N_{\ell-1} \cdot 4^\ell R^{\ell-1} \cdot \|\Phi - \Psi\|$, which then implies the claim of the lemma.

For $\ell = 0$, we trivially have $\|v_0\|_{\ell^\infty} = \|z\|_{\ell^\infty} \leq R_0 = R_0 \cdot (2R)^\ell \cdot N_1 \cdots N_{\ell-1}$ and furthermore $\|v_0 - w_0\|_{\ell^\infty} = 0 \leq R_0 \cdot N_1 \cdots N_{\ell-1} \cdot 4^\ell R^{\ell-1} \cdot \|\Phi - \Psi\|$.

Next, if the claimed estimates hold for some $\ell \in \{0, \dots, L-1\}$, we see

$$\begin{aligned}
\big|(v_{\ell+1})_j\big| &= \big|\sigma_{\ell+1}\big((B_{\ell+1} v_\ell + c_{\ell+1})_j\big)\big| \leq \big|(B_{\ell+1} v_\ell + c_{\ell+1})_j\big| \\
&\leq |(c_{\ell+1})_j| + \sum_{m=1}^{N_\ell} |(B_{\ell+1})_{j,m}| \, |(v_\ell)_m| \leq R + N_\ell \, R \, R_0 \cdot (2R)^\ell \cdot N_1 \cdots N_{\ell-1} \\
&\leq R_0 \cdot (2R)^{\ell+1} \cdot N_1 \cdots N_\ell \cdot \big(\tfrac{1}{2} \tfrac{1}{R_0 \, (2R)^\ell \, N_1 \cdots N_\ell} + \tfrac{1}{2}\big) \leq R_0 \cdot (2R)^{\ell+1} \cdot N_1 \cdots N_\ell,
\end{aligned}$$

proving the first estimate for $\ell + 1$ instead of ℓ. In a similar way, we see

$$\begin{aligned}
\big|(w_{\ell+1})_j - (v_{\ell+1})_j\big| &= \big|\sigma_{\ell+1}\big((A_{\ell+1} w_\ell + b_{\ell+1})_j\big) - \sigma_{\ell+1}\big((B_{\ell+1} v_\ell + c_{\ell+1})_j\big)\big| \\
&\leq \big|(A_{\ell+1} \, w_\ell - B_{\ell+1} \, v_\ell)_j\big| + \|b_{\ell+1} - c_{\ell+1}\|_{\ell^\infty}.
\end{aligned} \tag{A.7}$$

Next, note that

$$\begin{aligned}
\big|(A_{\ell+1} \, w_\ell - B_{\ell+1} \, v_\ell)_j\big| &\leq \sum_{m=1}^{N_\ell} \Big[\big|(A_{\ell+1})_{j,m}\big((w_\ell)_m - (v_\ell)_m\big)\big| + \big|\big((A_{\ell+1})_{j,m} - (B_{\ell+1})_{j,m}\big)(v_\ell)_m\big|\Big] \\
&\leq N_\ell \cdot \Big(R \cdot \|w_\ell - v_\ell\|_{\ell^\infty} + \|\Phi - \Psi\| \cdot \|v_\ell\|_{\ell^\infty}\Big) \\
&\overset{(*)}{\leq} R_0 \cdot N_1 \cdots N_\ell \cdot 4^{\ell+1} R^\ell \cdot \|\Phi - \Psi\| \cdot \big(\tfrac{1}{4} + \tfrac{2^\ell}{4^{\ell+1}}\big) \\
&\leq \frac{1}{2} \cdot R_0 \cdot N_1 \cdots N_\ell \cdot 4^{\ell+1} R^\ell \cdot \|\Phi - \Psi\|,
\end{aligned}$$

where the step marked with $(*)$ used the induction hypothesis. Combining this estimate with Equation (A.7) and noting $\|b_{\ell+1} - c_{\ell+1}\|_{\ell^\infty} \leq \|\Phi - \Psi\| \leq \frac{1}{2} \cdot R_0 \cdot N_1 \cdots N_\ell \cdot 4^{\ell+1} R^\ell \cdot \|\Phi - \Psi\|$ completes the induction. ∎

REFERENCES

[1] P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia, *Neural networks in multidimensional domains: fundamentals and new trends in modelling and control*, vol. 234, Springer, 1998.

[2] P. Arena, L. Fortuna, R. Re, and M. G. Xibilia, *On the capability of neural networks with complex neurons in complex valued functions approximation*, in 1993 IEEE International Symposium on Circuits and Systems, IEEE, 1993, https://doi.org/10.1109/ISCAS.1993.394188.

[3] P. Arena, L. Fortuna, R. Re, and M. G. Xibilia, *Multilayer perceptrons to approximate complex valued functions*, International Journal of Neural Systems, 6 (1995), https://doi.org/10.1142/s0129065795000299.

[4] M. Arjovsky, A. Shah, and Y. Bengio, *Unitary evolution recurrent neural networks*, in International Conference on Machine Learning, 2016, pp. 1120–1128.

[5] P. L. Bartlett, N. Harvey, C. Liaw, and A. Mehrabian, *Nearly-tight VC-dimension and Pseudodimension Bounds for Piecewise Linear Neural Networks*, J. Mach. Learn. Res., 20 (2019), pp. 1–17.

[6] J. Berner, P. Grohs, and A. Jentzen, *Analysis of the generalization error: empirical risk minimization over deep artificial neural networks overcomes the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations*, SIAM J. Math. Data Sci., 2 (2020), pp. 631–657, https://doi.org/10.1137/19M125649X.

[7] H. Bölcskei, P. Grohs, G. Kutyniok, and P. Petersen, *Optimal approximation with sparsely connected deep neural networks*, SIAM J. Math. Data Sci., 1 (2019), pp. 8–45, https://doi.org/10.1137/18M118709X.

[8] G. F. Clements, *Entropies of several sets of real valued functions*, Pacific J. Math., 13 (1963), pp. 1085–1095, http://projecteuclid.org/euclid.pjm/1103034547.
[9] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Mathematics of control, signals and systems, 2 (1989), pp. 303–314, https://doi.org/10.1007/BF02551274.
[10] X. Glorot, A. Bordes, and Y. Bengio, *Deep sparse rectifier neural networks*, in Proceedings of the fourteenth international conference on artificial intelligence and statistics, 2011.
[11] L. Gonon, P. Grohs, A. Jentzen, D. Kofler, and D. Šiška, *Uniform error estimates for artificial neural network approximations for heat equations*, arXiv preprint arXiv:1911.09647, (2019).
[12] P. Grohs and L. Herrmann, *Deep neural network approximation for high-dimensional elliptic PDEs with boundary conditions*, IMA Journal of Numerical Analysis, (2021), https://doi.org/10.1093/imanum/drab031.
[13] P. Grohs and L. Herrmann, *Deep neural network approximation for high-dimensional parabolic Hamilton-Jacobi-Bellman equations*, arXiv preprint arXiv:2103.05744, (2021).
[14] A. Hirose, *Complex-valued neural networks: theories and applications*, vol. 5, World Scientific, 2003, https://doi.org/10.1142/5345.
[15] K. Hornik, *Approximation capabilities of multilayer feedforward networks*, Neural Networks, 4 (1991), pp. 251–257, https://doi.org/10.1016/0893-6080(91)90009-T.
[16] K. Hornik, M. Stinchcombe, and H. White, *Multilayer feedforward networks are universal approximators*, Neural Networks, 2 (1989), pp. 359–366, https://doi.org/10.1016/0893-6080(89)90020-8.
[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *ImageNet classification with deep convolutional neural networks*, Communications of the ACM, 60 (2017), https://doi.org/10.1145/3065386.
[18] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature, 521 (2015).
[19] M. Leshno, V. Lin, A. Pinkus, and S. Schocken, *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function*, Neural Networks, 6 (1993), pp. 861–867, https://doi.org/10.1016/S0893-6080(05)80131-5.
[20] J. Lu, Z. Shen, H. Yang, and S. Zhang, *Deep network approximation for smooth functions*, SIAM J. Math. Anal., 53 (2021), pp. 5465–5506, https://doi.org/10.1137/20M134695X.
[21] H. N. Mhaskar, *Neural networks for optimal approximation of smooth and analytic functions*, Neural computation, 8 (1996), https://doi.org/10.1162/neco.1996.8.1.164.
[22] P. Petersen and F. Voigtlaender, *Optimal approximation of piecewise smooth functions using deep ReLU neural networks*, Neural Netw., 108 (2018), https://doi.org/10.1016/j.neunet.2018.08.019.
[23] I. Safran and O. Shamir, *Depth-width tradeoffs in approximating natural functions with neural networks*, arXiv preprint arXiv:1610.09887, (2016).
[24] I. Safran and O. Shamir, *Depth-width tradeoffs in approximating natural functions with neural networks*, in International Conference on Machine Learning, PMLR, 2017, pp. 2979–2987.
[25] I. Sutskever, O. Vinyals, and Q. V. Le, *Sequence to sequence learning with neural networks*, in Advances in neural information processing systems, 2014.
[26] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, *Deep complex networks*, in ICLR, 2018, https://openreview.net/forum?id=H1T2hmZAb.
[27] R. Vershynin, *High-dimensional probability*, vol. 47 of Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2018, https://doi.org/10.1017/9781108231596.
[28] P. Virtue, S. X. Yu, and M. Lustig, *Better than real: Complex-valued neural nets for MRI fingerprinting*, in 2017 IEEE International Conference on Image Processing (ICIP), 2017, https://doi.org/10.1109/ICIP.2017.8297024.
[29] F. Voigtlaender, *The universal approximation theorem for complex-valued neural networks*, arXiv preprint arXiv:2012.03351, (2020).
[30] M. Wolter and A. Yao, *Complex gated recurrent neural networks*, in Advances in Neural Information Processing Systems, vol. 31, Curran Associates, Inc., 2018.
[31] D. Yarotsky, *Error bounds for approximations with deep ReLU networks*, Neural Networks, 94 (2017), pp. 103–114, https://doi.org/10.1016/j.neunet.2017.07.002.
[32] D. Yarotsky and A. Zhevnerchuk, *The phase diagram of approximation rates for deep neural networks*, Advances in Neural Information Processing Systems, 33 (2020).

6.6 Paper 6

A note on exponential Riesz bases

Andrei Caragea, Dae Gwan Lee*

Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt-Ingolstadt, 85071 Eichstätt, Germany

Abstract

We prove that if $I_\ell = [a_\ell, b_\ell)$, $\ell = 1, \dots, L$, are disjoint intervals in $[0, 1)$ with the property that the numbers $1, a_1, \dots, a_L, b_1, \dots, b_L$ are linearly independent over $\mathbb{Q}$, then there exist pairwise disjoint sets $\Lambda_\ell \subset \mathbb{Z}$, $\ell = 1, \dots, L$, such that for every $J \subset \{1, \dots, L\}$, the system $\{e^{2\pi i \lambda x} : \lambda \in \cup_{\ell \in J} \Lambda_\ell\}$ is a Riesz basis for $L^2(\cup_{\ell \in J} I_\ell)$. Also, we show that for any disjoint intervals I_ℓ, $\ell = 1, \dots, L$, contained in $[1, N)$ with $N \in \mathbb{N}$, the orthonormal basis $\{e^{2\pi i n x} : n \in \mathbb{Z}\}$ of $L^2[0, 1)$ can be complemented by a Riesz basis $\{e^{2\pi i \lambda x} : \lambda \in \Lambda\}$ for $L^2(\cup_{\ell=1}^L I_\ell)$ with some set $\Lambda \subset (\frac{1}{N}\mathbb{Z})\backslash\mathbb{Z}$, in the sense that their union $\{e^{2\pi i \lambda x} : \lambda \in \mathbb{Z} \cup \Lambda\}$ is a Riesz basis for $L^2([0, 1) \cup I_1 \cup \dots \cup I_L)$.

Keywords: exponential bases, Riesz bases, hierarchical structure, finite union of intervals, Kronecker–Weyl equidistribution along the primes
2000 MSC: 42C15

1. Introduction and Main Results

In 1995, Seip [18] showed that if S is an interval contained in $[0, 1)$, then there exists a set $\Lambda \subset \mathbb{Z}$ such that $E(\Lambda) := \{e^{2\pi i \lambda x} : \lambda \in \Lambda\}$ is a Riesz basis for $L^2(S)$. Since then, there have been various attempts towards finding/characterizing the sets S that admit a Riesz spectrum, see e.g., [1, 3, 4, 6, 7, 9]. A significant breakthrough was made by Kozma and Nitzan [10] who proved that if $[a_\ell, b_\ell)$, $\ell = 1, \dots, L$, are disjoint intervals contained in $[0, 1)$, then there exists a set $\Lambda \subset \mathbb{Z}$ such that $E(\Lambda)$ is a Riesz basis for $L^2(\cup_{\ell=1}^L [a_\ell, b_\ell))$. Recently, Pfander, Revay and Walnut [17] showed that if the intervals $[a_\ell, b_\ell)$ form a partition of $[0, 1)$, then the set of integers $\mathbb{Z}$ can be partitioned into some sets Λ_ℓ, $\ell = 1, \dots, L$, such that for each ℓ, the system $E(\Lambda_\ell)$ is a Riesz basis for $L^2[a_\ell, b_\ell)$, and moreover $E(\cup_{\ell \in J} \Lambda_\ell)$ is a Riesz basis for $L^2(\cup_{\ell \in J} S_\ell)$ whenever $J \subset \{1, \dots, L\}$ is a consecutive index set (see [17, Theorems 1 and 2]). We would like to point out that up to date, the existence of exponential Riesz bases is known only

*Corresponding author
Email addresses: andrei.caragea@gmail.com (Andrei Caragea), daegwans@gmail.com (Dae Gwan Lee)

for several classes of sets $S \subset \mathbb{R}$. Recently, Kozma, Nitzan and Olevskii [12] constructed a bounded measurable set $S \subset \mathbb{R}$ such that the space $L^2(S)$ has no exponential Riesz basis. For an overview of the known results on exponential Riesz bases, we refer to [14, Section 1].

We are interested in the following two problems:

Problem 1 (Hierarchical structured exponential Riesz bases)**.** *Given a family of disjoint sets $S_1, S_2, \ldots, S_L \subset [0,1)$ with positive measure, can we find disjoint sets $\Lambda_1, \Lambda_2, \ldots, \Lambda_L \subset \mathbb{Z}$ such that for every $J \subset \{1, \ldots, L\}$, the system $E(\cup_{\ell \in J} \Lambda_\ell)$ is a Riesz basis for $L^2(\cup_{\ell \in J} S_\ell)$?*

Problem 2 (Complementability of exponential Riesz bases)**.** *Let $\Lambda \subset \mathbb{R}$ be a discrete set and let $S \subset \mathbb{R}$ be a finite positive measure set such that $E(\Lambda)$ is a Riesz basis for $L^2(S)$. Given a finite positive measure set $S' \subset \mathbb{R} \backslash S$, can we find a discrete set $\Lambda' \subset \mathbb{R} \backslash \Lambda$ such that*

- *$E(\Lambda')$ is a Riesz basis for $L^2(S')$, and*
- *$E(\Lambda \cup \Lambda')$ is a Riesz basis for $L^2(S \cup S')$?*

The second problem is closely related to the first, as it deals with the case $L = 2$ under the assumption that the sets S_1 and Λ_1 are already fixed.

Considering the result of Kozma, Nitzan and Olevskii [12], it is necessary to restrict the sets S_ℓ, S and S' to certain classes of sets. In this paper, we will address the above problems in the case that S_ℓ, S and S' are intervals or finite unions of intervals.

Our first main result answers Problem 1 in the affirmative when $S_\ell = [a_\ell, b_\ell)$, $\ell = 1, \ldots, L$, are disjoint intervals in $[0,1)$ with the property that the numbers $1, a_1, \ldots, a_L, b_1, \ldots, b_L$ are linearly independent over $\mathbb{Q}$, which means that having $q + q_1 a_1 + \ldots + q_L a_L + q'_1 b_1 + \ldots + q'_L b_L = 0$ for some $q, q_\ell, q'_\ell \in \mathbb{Q}$ implies $q = q_\ell = q'_\ell = 0$ for all ℓ. The result is motivated by [10, p.279, Claim 2].

Theorem 1. *Let $0 < a_1 < b_1 < \cdots < a_L < b_L < 1$ with $L \in \mathbb{N}$. Assume that the numbers $1, a_1, \ldots, a_L, b_1, \ldots, b_L$ are linearly independent over $\mathbb{Q}$. There exist pairwise disjoint sets $\Lambda_\ell \subset \mathbb{Z}$, $\ell = 1, \ldots, L$, such that for every $J \subset \{1, \ldots, L\}$, the system $E(\cup_{\ell \in J} \Lambda_\ell)$ is a Riesz basis for $L^2(\cup_{\ell \in J} [a_\ell, b_\ell))$.*

Concerning Problem 2, we have the following result which builds on the fact that $E(\mathbb{Z})$ is an orthonormal basis (thus, a Riesz basis) for $L^2[0,1)$.

Theorem 2. *Let $1 \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_L < b_L \leq N$ with $L, N \in \mathbb{N}$. There exists a set $\Lambda' \subset (\frac{1}{N}\mathbb{Z}) \backslash \mathbb{Z}$ such that*

- *$E(\Lambda')$ is a Riesz basis for $L^2(\cup_{\ell=1}^{L} [a_\ell, b_\ell))$, and*
- *$E(\mathbb{Z} \cup \Lambda')$ is a Riesz basis for $L^2([0,1) \cup [a_1, b_1) \cup \cdots \cup [a_L, b_L))$.*

This theorem answers Problem 2 in the affirmative when $S = [0,1)$, $\Lambda = \mathbb{Z}$, and S' is a finite union of disjoint bounded intervals in $[1, \infty)$.

While the result of Pfander, Revay and Walnut [17] relies on Avdonin's theorem and the ergodic properties of a certain type of integer sequences, Theorem 1 is based on a refinement of the key lemma of [10] for primes, together with Chebotarëv's theorem on roots of unity and the Kronecker–Weyl equidistribution theorem along the primes.

1.1. Remarks

We state the following conjecture which improves upon Theorem 1.

Conjecture 1. *Let $[a_\ell, b_\ell)$, $\ell = 1, \dots, L$, be disjoint intervals contained in $[0,1)$, that is, $0 \le a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_L < b_L \le 1$. There exist pairwise disjoint sets $\Lambda_\ell \subset \mathbb{Z}$, $\ell = 1, \dots, L$, such that for every $J \subset \{1, \dots, L\}$, the system $E(\cup_{\ell \in J} \Lambda_\ell)$ is a Riesz basis for $L^2(\cup_{\ell \in J} [a_\ell, b_\ell))$.*

This conjecture generalizes Theorem 1 by removing the Q-linear independence of the endpoints and by allowing for contiguous intervals in $[0,1)$, i.e., $b_\ell = a_{\ell+1}$ for some $\ell \in \{1, \dots, L-1\}$. The conjecture can be easily reformulated as follows.

Conjecture 1′. *Let I_ℓ, $\ell = 1, \dots, L$, be intervals which form a partition of $[0,1)$. There exists a partition Λ_ℓ, $\ell = 1, \dots, L$, of $\mathbb{Z}$ such that for every $J \subset \{1, \dots, L\}$, the system $E(\cup_{\ell \in J} \Lambda_\ell)$ is a Riesz basis for $L^2(\cup_{\ell \in J} I_\ell)$.*

Indeed, Conjecture 1 obviously implies Conjecture 1′, and the converse is seen by considering the partition of $[0,1)$ formed using the endpoints of all $[a_\ell, b_\ell)$. Note that Conjecture 1′ generalizes the result of Pfander, Revay and Walnut [17] from consecutive index sets J to arbitrary index sets J.

Lastly, we mention that both Problems 1 and 2 remain open for more general classes of sets S_ℓ, S and S'.

2. Preliminaries

Definition. *A sequence $\{f_n\}_{n \in \mathbb{Z}}$ in a separable Hilbert space $\mathcal{H}$ is called*

- *a* frame *for $\mathcal{H}$ (with frame bounds A and B) if there are constants $0 < A \le B < \infty$ such that*

$$A \|f\|^2 \;\le\; \sum_{n \in \mathbb{Z}} |\langle f, f_n \rangle|^2 \;\le\; B \|f\|^2 \quad \text{for all } \; f \in \mathcal{H};$$

- *a* Riesz sequence *in $\mathcal{H}$ (with Riesz bounds A and B) if there are constants $0 < A \le B < \infty$ such that*

$$A \|c\|_{\ell_2}^2 \;\le\; \Big\| \sum_{n \in \mathbb{Z}} c_n f_n \Big\|^2 \;\le\; B \|c\|_{\ell_2}^2 \quad \text{for all } \; \{c_n\}_{n \in \mathbb{Z}} \in \ell_2(\mathbb{Z});$$

- *a* Riesz basis *for $\mathcal{H}$ if it is a complete Riesz sequence in $\mathcal{H}$.*

It is well-known (see e.g., [5, Proposition 3.7.3, Theorems 5.4.1 and 7.1.1] or [10, Lemma 1]) that a sequence in $\mathcal{H}$ is a Riesz basis if and only if it is both a frame and a Riesz sequence. Moreover in this case, the optimal frame bounds coincides with the optimal Riesz bounds. It is worth noting that Riesz bases are equivalent to unconditional bases that are norm-bounded above and below [5, Lemma 3.6.9]. Since every exponential function has constant norm in $L^2(S)$ with $S \subset \mathbb{R}^d$, namely $\|e^{2\pi i\lambda\cdot(\cdot)}\|_{L^2(S)} = |S|^{1/2}$ for any $\lambda \in \mathbb{R}^d$, *Riesz bases of exponentials* coincide with *unconditional bases of exponentials*.

Proposition 3 (Proposition 2.1 in [16], Proposition 5.4 in [2])**.** *Let $\{e_n\}_{n\in I}$ be an orthonormal basis of a separable Hilbert space $\mathcal{H}$, where I is a countable index set. Let $P : \mathcal{H} \to \mathcal{M}$ be the orthogonal projection from $\mathcal{H}$ onto a closed subspace $\mathcal{M}$. Let $J \subset I$, $J^c := I\backslash J$, and $0 < \alpha \leq 1$. The following are equivalent.*

(i) *$\{Pe_n\}_{n\in J} \subset \mathcal{M}$ is a frame for $\mathcal{M}$ with optimal lower bound α.*

(ii) *$\{Pe_n\}_{n\in J^c} \subset \mathcal{M}$ is a Bessel sequence with optimal bound $1-\alpha$.*

(iii) *$\{(\mathrm{Id}-P)e_n\}_{n\in J^c} \subset \mathcal{M}^\perp$ is a Riesz sequence with optimal lower bound α.*

As a direct consequence of Proposition 3, we have that for a set $\Lambda \subset \mathbb{Z}$ and a measurable set $S \subset [0,1)$, the system $E(\Lambda)$ is a frame for $L^2(S)$ if and only if $E(\mathbb{Z}\backslash\Lambda)$ is a Riesz sequence in $L^2([0,1)\backslash S)$.

Lemma 4. *Assume that $E(\Lambda)$ is a Riesz basis for $L^2(S)$ with bounds $0 < A \leq B < \infty$, where $\Lambda \subset \mathbb{R}^d$ is a discrete set and $S \subset \mathbb{R}^d$ is a measurable set. Then the following hold.*
(a) For any $a, b \in \mathbb{R}^d$, the system $E(\Lambda + a)$ is a Riesz basis for $L^2(S + b)$ with bounds A and B.
(b) For any $c > 0$, the system $E(c\Lambda)$ is a Riesz basis for $L^2(\frac{1}{c}S)$ with bounds $\frac{A}{c}$ and $\frac{B}{c}$.

Lemma 4 remains valid if all the terms "Riesz basis" are replaced by "Riesz sequence" or by "frame". A proof of Lemma 4 can be found in [14].

For any $N \in \mathbb{N}$, a measurable set $S \subset [0,1)$, and $n = 1, \ldots, N$, we define

$$A_{\geq n} = A_{\geq n}(N,S) := \Big\{ t \in [0, \tfrac{1}{N}) : t + \tfrac{k}{N} \in S \ \text{ for at least } n \text{ values of } \ k \in \{0,1,\ldots,N-1\} \Big\}. \tag{1}$$

Lemma 5 (Lemma 2 in [10])**.** *Let $N \in \mathbb{N}$ and let $S \subset [0,1)$ be a measurable set. If there exist sets $\Lambda_1, \ldots, \Lambda_N \subset N\mathbb{Z}$ such that $E(\Lambda_n)$ is a Riesz basis for $L^2(A_{\geq n})$, then $E(\cup_{n=1}^N (\Lambda_n + n))$ is a Riesz basis for $L^2(S)$.*

This lemma, which plays a central role in [10], combines Riesz bases by introducing *consecutive* shift factors n to the frequency sets Λ_n and then taking their union $\cup_{n=1}^N (\Lambda_n + n)$. For our purpose, we generalize the lemma to allow for *arbitrary* shift factors when N is prime.

Lemma 6. *Let $N \in \mathbb{N}$ be a prime and let $S \subset [0,1)$ be a measurable set. If there exist sets $\Lambda_1, \ldots, \Lambda_N \subset N\mathbb{Z}$ such that $E(\Lambda_n)$ is a Riesz basis (resp. a frame, a Riesz sequence) for $L^2(A_{\geq n})$, then for every permutation $\{j_n\}_{n=1}^N$ of $\{1, \ldots, N\}$ the system $E(\cup_{n=1}^N (\Lambda_n + j_n))$ is a Riesz basis (resp. a frame, a Riesz sequence) for $L^2(S)$.*

See Appendix A for a proof of Lemma 6.

We will use the following notation throughout the proofs. For $x \in \mathbb{R}$, we denote the fractional part of x by $\{x\}$, that is, $0 \leq \{x\} := x - \lfloor x \rfloor < 1$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Also, we adopt the convention that $[x, y) = \emptyset$ if $x = y \in \mathbb{R}$.

For the proof of Theorem 1, we will need the following version of the Kronecker–Weyl equidistribution theorem (see e.g., [8, Theorem 443] or [13, p.48, Theorem 6.3 and Example 6.1]) along the *primes*. The one-dimensional case ($d = 1$) was proved by Vinogradov [20] (see also [13, p.22]): if a is an *irrational* number, the sequence $\{2a\}, \{3a\}, \{5a\}, \{7a\}, \ldots$ is *uniformly distributed* in $[0,1)$, meaning that for every interval $I \subset [0,1)$, the ratio of the numbers $\{pa\}$ with prime $p \leq N$ that are contained in I, tends to $|I|$ as $N \to \infty$. The notion of *uniform distribution* is defined similarly in higher dimensions, see e.g., [13, p.47, Definition 6.1]. As we could not find any reference for the multi-dimensional case, we include a short proof here.

Proposition 7 (Kronecker–Weyl equidistribution along the primes)**.** *Let $d \in \mathbb{N}$ and $a_1, \ldots, a_d \in \mathbb{R}$. If the numbers $1, a_1, \ldots, a_d$ are linearly independent over $\mathbb{Q}$, which means that having $q + q_1 a_1 + \ldots + q_d\, a_d = 0$ for some $q, q_1, \ldots, q_d \in \mathbb{Q}$ implies $q = q_1 = \ldots = q_d = 0$, then the d-dimensional vectors*

$$\big(\{p\, a_1\}, \ldots, \{p\, a_d\}\big) \quad \text{for } \ p \in \mathcal{P}$$

are uniformly distributed in $[0,1)^d$, where $\mathcal{P} = \{2, 3, 5, 7, \ldots\}$ is the set of primes.

Proof. For convenience, we denote the n-th prime by p_n, that is, p_1=2, p_2=3, p_3=5, p_4=7, and so on. By Weyl's criterion (see e.g., [13, p.48, Theorems 6.2 and 6.3]), the claim is equivalent to having that for every $\boldsymbol{z} = (z_1, \ldots, z_d) \in \mathbb{Z}^d \backslash \{0\}$, the fractional part of $\langle \boldsymbol{z}, (p_n a_1, \ldots, p_n a_d) \rangle = z_1 \cdot p_n a_1 + \ldots + z_d \cdot p_n a_d = p_n \cdot (z_1 a_1 + \ldots + z_d\, a_d)$ for $n = 1, 2, \ldots$ are uniformly distributed in $[0,1)$. Note that for any fixed $\boldsymbol{z} = (z_1, \ldots, z_d) \in \mathbb{Z}^d \backslash \{0\}$, the number $\widetilde{a} := z_1 a_1 + \ldots + z_d\, a_d$ is *irrational* because $1, a_1, \ldots, a_d$ are linearly independent over $\mathbb{Q}$. Hence, the result of Vinogradov [20] implies that the numbers $\{p_n \widetilde{a}\}$, $n = 1, 2, \ldots$, are uniformly distributed in $[0,1)$, as desired. □

3. Proof of Theorem 1

Proposition 7 implies that there exist infinitely many prime numbers $N \in \mathbb{N}$ satisfying

$$\begin{aligned} 0 \ < \ \{Na_1\} \ < \ \{Na_2\} \ < \ \ldots \ < \ \{Na_{L-1}\} \ < \ \{Na_L\} \\ < \ \{Nb_L\} \ < \ \{Nb_{L-1}\} \ < \ \ldots \ < \ \{Nb_2\} \ < \ \{Nb_1\} \ < \ 1. \end{aligned} \tag{2}$$

Among such numbers, choose a large $N \in \mathbb{N}$ so that every spacing between the numbers $0 < a_1 < b_1 < a_2 < b_2 < \cdots < a_L < b_L < 1$ contains at least one of $\frac{k}{N}$, $k = 1, \ldots, N-1$, as an interior point (which clearly requires $2L+1 \le N$). This ensures that with respect to the grid $\frac{1}{N}\mathbb{Z}$ the interval $[a_\ell, b_\ell)$ is partitioned into translates of

$$\big[\tfrac{\{Na_\ell\}}{N}, \tfrac{1}{N}\big), \quad \big[0, \tfrac{\{Nb_\ell\}}{N}\big), \quad \text{and possibly some extra intervals } [0, \tfrac{1}{N}), \tag{3}$$

and that the rightmost segment of $[a_\ell, b_\ell)$, corresponding to $[0, \frac{\{Nb_\ell\}}{N})$ in (3), lies in a set $[\frac{k}{N}, \frac{k+1}{N})$ which does not intersect the next interval $[a_{\ell+1}, b_{\ell+1})$. Consequently, each of the sets $A_{\ge n} = A_{\ge n}(N, S)$, $n = 1, 2, \ldots, N$, is one of the form

$$\emptyset, \quad [0, \tfrac{1}{N}), \quad \text{and} \quad \big[\tfrac{\{Na_\ell\}}{N}, \tfrac{\{Nb_\ell\}}{N}\big) \quad \text{for some } \ell \in \{0, 1, \ldots, N-1\}.$$

Note that due to (2), the translates of $[\frac{\{Na_\ell\}}{N}, \frac{1}{N})$ and $[0, \frac{\{Nb_\ell\}}{N})$ in (3) together contribute exactly $[0, \frac{1}{N})$ and $[\frac{\{Na_\ell\}}{N}, \frac{\{Nb_\ell\}}{N})$ to the family of sets $A_{\ge n}$. The nested sets

$$A_{\ge 1} \supset A_{\ge 2} \supset \cdots \supset A_{\ge N}$$

are thus given by

$$\overbrace{[0, \tfrac{1}{N}) = \cdots = [0, \tfrac{1}{N})}^{K} \supset \overbrace{\big[\tfrac{\{Na_1\}}{N}, \tfrac{\{Nb_1\}}{N}\big) \supset \cdots \supset \big[\tfrac{\{Na_L\}}{N}, \tfrac{\{Nb_L\}}{N}\big)}^{L}$$
$$\supset \overbrace{\emptyset = \cdots = \emptyset}^{N-K-L} \qquad \text{for some integer } K \ge L.$$

Let us associate each set $A_{\ge n}$ with the interval $[a_\ell, b_\ell)$ which it originates from. The sets $[0, \frac{1}{N})$ can be associated with the intervals $[a_\ell, b_\ell)$ in various ways, but for convenience we will assume

$$\underbrace{\overbrace{[0, \tfrac{1}{N}) = \cdots = [0, \tfrac{1}{N})}^{K_1}}_{\updownarrow\, [a_1, b_1)} = \cdots = \underbrace{\overbrace{[0, \tfrac{1}{N}) = \cdots = [0, \tfrac{1}{N})}^{K_L}}_{\updownarrow\, [a_L, b_L)}$$
$$\supset \underbrace{\overbrace{\big[\tfrac{\{Na_1\}}{N}, \tfrac{\{Nb_1\}}{N}\big) \supset \cdots \supset \big[\tfrac{\{Na_L\}}{N}, \tfrac{\{Nb_L\}}{N}\big)}^{L}}_{\updownarrow\, [a_1, b_1) \qquad \updownarrow\, [a_L, b_L)} \supset \overbrace{\emptyset = \cdots = \emptyset}^{N-K-L}$$

where $K = \sum_{\ell=1}^{L} K_\ell$ with $K_\ell \in \mathbb{N}$ for all ℓ.

Step 1. *Construction of the sets $\Lambda_\ell \subset \mathbb{Z}$, $\ell = 1, \ldots, L$.*
For each $n = 1, \ldots, N$, we apply the result of Seip [18] (see the beginning of Section 1) to obtain a set $\Lambda^{(n)} \subset N\mathbb{Z}$ such that $E(\Lambda^{(n)})$ is a Riesz basis for

$L^2(A_{\geq n})$; it is easily seen that

$$\Lambda^{(n)} = \begin{cases} N\mathbb{Z} & \text{for } 1 \leq n \leq K, \\ \subsetneq N\mathbb{Z} & \text{for } K{+}1 \leq n \leq K{+}L, \\ \emptyset & \text{for } K{+}L{+}1 \leq n \leq N. \end{cases}$$

Lemma 5 implies that $E(\cup_{n=1}^{N}(\Lambda^{(n)}{+}n))$ is a Riesz basis for $L^2(S)$. For each $\ell = 1, \ldots, L$, let Λ_ℓ be the union of $\Lambda^{(n)}{+}n$ over all n such that $A_{\geq n}$ is associated with $[a_\ell, b_\ell)$, that is,

$$\begin{aligned}
\Lambda_1 &:= \bigcup_{n \in \{1,\, 2,\, \ldots,\, K_1,\, K+1\}} (\Lambda^{(n)}{+}n) = \Big(\dot{\bigcup}_{n=1}^{K_1} (N\mathbb{Z}{+}n) \Big) \,\dot{\cup}\, (\Lambda^{(K+1)}{+}K{+}1), \\
\Lambda_2 &:= \bigcup_{n \in \{K_1+1,\, K_1+2,\, \ldots,\, K_1+K_2,\, K+2\}} (\Lambda^{(n)}{+}n) \\
&= \Big(\dot{\bigcup}_{n=K_1+1}^{K_1+K_2} (N\mathbb{Z}{+}n) \Big) \,\dot{\cup}\, (\Lambda^{(K+2)}{+}K{+}2), \\
&\;\;\vdots \\
\Lambda_L &:= \bigcup_{n \in \{K_1+\cdots+K_{L-1}+1,\, \ldots,\, K,\, K+L\}} (\Lambda^{(n)}{+}n) \\
&= \Big(\dot{\bigcup}_{n=K_1+\cdots+K_{L-1}+1}^{K} (N\mathbb{Z}{+}n) \Big) \,\dot{\cup}\, (\Lambda^{(K+L)}{+}K{+}L).
\end{aligned}$$

Clearly, we have

$$\bigcup_{\ell=1}^{L} \Lambda_\ell = \bigcup_{n=1}^{N} (\Lambda^{(n)}{+}n)$$

and thus, $E(\cup_{\ell=1}^{L} \Lambda_\ell)$ is a Riesz basis for $L^2(S)$.

Step 2. For a subset $J \subset \{1, \ldots, L\}$, we set $\Lambda^J := \cup_{\ell \in J} \Lambda_\ell$ and $S^J := \cup_{\ell \in J} [a_\ell, b_\ell)$. We claim that $E(\Lambda^J)$ *is a Riesz basis for* $L^2(S^J)$.

First, note that the corresponding sets $A^J_{\geq n} := A_{\geq n}(N, S^J)$ for $n = 1, \ldots, N$ are again of the form

$$\emptyset, \quad [0, \tfrac{1}{N}), \quad \text{or} \quad \big[\tfrac{\{Na_\ell\}}{N}, \tfrac{\{Na_\ell\}}{N}\big) \quad \text{for some } \ell \in J.$$

Denoting $J = \{\ell_1, \ldots, \ell_M\}$ with $\ell_1 < \ldots < \ell_M$, we see that the nested sets

$$A^J_{\geq 1} \supset A^J_{\geq 2} \supset \cdots \supset A^J_{\geq N}$$

are given by

$$\begin{array}{c}
\overbrace{\underbrace{[0,\tfrac{1}{N})=\cdots=[0,\tfrac{1}{N})}_{\updownarrow}}^{K_{\ell_1}}=\cdots=\overbrace{\underbrace{[0,\tfrac{1}{N})=\cdots=[0,\tfrac{1}{N})}_{\updownarrow}}^{K_{\ell_M}} \\
\big[a_{\ell_1},b_{\ell_1}\big) \qquad\qquad\qquad\qquad \big[a_{\ell_M},b_{\ell_M}\big) \\
\supset \overbrace{\underbrace{\Big[\tfrac{\{Na_{\ell_1}\}}{N},\tfrac{\{Nb_{\ell_1}\}}{N}\Big)}_{\updownarrow}\ \supset\ \cdots\ \supset\ \underbrace{\Big[\tfrac{\{Na_{\ell_M}\}}{N},\tfrac{\{Nb_{\ell_M}\}}{N}\Big)}_{\updownarrow}}^{M}\ \supset\ \overbrace{\emptyset=\cdots=\emptyset}^{N-K_J-M} \\
\big[a_{\ell_1},b_{\ell_1}\big) \qquad\qquad\qquad\qquad \big[a_{\ell_M},b_{\ell_M}\big)
\end{array}$$

where $K_J := \sum_{\ell\in J} K_\ell$. Note that applying Lemma 5 directly to this setup will incur *different shift factors* in the frequency sets. Indeed, Lemma 5 implies that $E(\Lambda')$ is a Riesz basis for $L^2(S^J)$, with

$$\begin{aligned}\Lambda' := \Big(\bigcup\nolimits_{n=1}^{K_J}(N\mathbb{Z}+n)\Big)\ \bigcup\ (\Lambda^{(K+\ell_1)}+K_J+1)\ \bigcup\ (\Lambda^{(K+\ell_2)}+K_J+2)\\ \bigcup\ \cdots\ \bigcup\ (\Lambda^{(K+\ell_M)}+K_J+M),\end{aligned} \tag{4}$$

where $\Lambda^{(n)} \subset N\mathbb{Z}$ for $n = 1,\dots,N$ are the sets defined in Step 1. However, our goal is to show that $E(\cup_{m=1}^{M}\Lambda_{\ell_m})$ is a Riesz basis for $L^2(S^J)$, where

$$\begin{aligned}
\Lambda_{\ell_1} &:= \Big(\bigcup\nolimits_{n=K_1+K_2+\cdots+K_{\ell_1-1}+1}^{K_1+K_2+\cdots+K_{\ell_1}}(N\mathbb{Z}+n)\Big)\ \bigcup\ (\Lambda^{(K+\ell_1)}+K+\ell_1),\\
\Lambda_{\ell_2} &:= \Big(\bigcup\nolimits_{n=K_1+K_2+\cdots+K_{\ell_2-1}+1}^{K_1+K_2+\cdots+K_{\ell_2}}(N\mathbb{Z}+n)\Big)\ \bigcup\ (\Lambda^{(K+\ell_2)}+K+\ell_2),\\
&\ \ \vdots\\
\Lambda_{\ell_M} &:= \Big(\bigcup\nolimits_{n=K_1+K_2+\cdots+K_{\ell_M-1}+1}^{K_1+K_2+\cdots+K_{\ell_M}}(N\mathbb{Z}+n)\Big)\ \bigcup\ (\Lambda^{(K+\ell_M)}+K+\ell_M).
\end{aligned}$$

To show this, consider the $K_J{+}M$ sets

$$\begin{aligned}
\Omega_1 &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_1-1}{+}1),\\
&\vdots \qquad\quad \vdots\\
\Omega_{K_{\ell_1}} &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_1}),\\
\Omega_{K_{\ell_1}+1} &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_2-1}{+}1),\\
&\vdots \qquad\quad \vdots\\
\Omega_{K_{\ell_1}+K_{\ell_2}} &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_2}),\\
&\vdots \qquad\quad \vdots\\
\Omega_{K_{\ell_1}+K_{\ell_2}+\cdots+K_{\ell_{M-1}}} &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_M-1}{+}1),\\
&\vdots \qquad\quad \vdots\\
\Omega_{K_J} &:= N\mathbb{Z}{+}(K_1{+}K_2{+}\cdots{+}K_{\ell_M}) = N\mathbb{Z}{+}K_J,\\
\Omega_{K_J+1} &:= \Lambda^{(K+\ell_1)}{+}(K{+}\ell_1),\\
\Omega_{K_J+2} &:= \Lambda^{(K+\ell_2)}{+}(K{+}\ell_2),\\
&\vdots \qquad\quad \vdots\\
\Omega_{K_J+M} &:= \Lambda^{(K+\ell_M)}{+}(K{+}\ell_M),
\end{aligned}$$

which partitions $\cup_{m=1}^{M}\Lambda_{\ell_m}$, that is, $\uplus_{n=1}^{K_J+M}\Omega_n = \cup_{m=1}^{M}\Lambda_{\ell_m}$. Here, the sets Ω_n are exactly ordered in the way that $E(\Omega_n)$ is a Riesz basis for $L^2(A^J_{\geq n})$. Note that while the $K_J{+}M$ components of Λ' in (4) have *consecutive shift factors*, namely from 1 up to $K_J{+}M$, the shift factors associated with Ω_n are not consecutive in general. However, since $N \in \mathbb{N}$ is prime, Lemma 6 implies that $E(\cup_{m=1}^{M}\Lambda_{\ell_m}) = E(\uplus_{n=1}^{K_J+M}\Omega_n)$ is a Riesz basis for $L^2(S^J)$. This completes the proof.

4. Proof of Theorem 2

To prove Theorem 2, we will use Lemma 5 which is the key lemma of Kozma and Nitzan [10]. Note that by Lemma 4, one may replace the frequency set $\cup_{n=1}^{N}(\Lambda_n{+}n)$ in Lemma 5 by $\cup_{n=1}^{N}(\Lambda_n{+}n{-}1)$, while preserving the Riesz basis property.

We will first prove the case $L = 1$ and then extend the proof to the case $L \geq 2$.

Case $L = 1$. Given a set $V = [0,1) \cup [a,b) \subset [0,N)$ with $N \in \mathbb{N}$ and $1 \leq a < b \leq N$, let $S := \frac{1}{N}V = [0,\frac{1}{N}) \cup [\frac{a}{N},\frac{b}{N}) \subset [0,1)$. We will apply Lemma 5 directly to this set S. There are two cases, either $\{a\} \leq \{b\}$ or $\{b\} < \{a\}$.

First, assume that $0 \leq \{a\} \leq \{b\} < 1$. Then there exists a number $M \in \mathbb{N}$ such that

$$A_{\geq 1} = A_{\geq 2} = \cdots = A_{\geq M} = [0,\tfrac{1}{N}) \ \supset\ A_{\geq M+1} = \big[\tfrac{\{a\}}{N},\tfrac{\{b\}}{N}\big) \ \supset\ A_{\geq M+2} = \emptyset.$$

Clearly, we may choose the canonical frequency sets $\Lambda_1 = \cdots = \Lambda_M = N\mathbb{Z}$ for $A_{\geq 1} = A_{\geq 2} = \cdots = A_{\geq M} = [0, \frac{1}{N})$, so that for each $n = 1, \ldots, M$, the system $E(\Lambda_n)$ is a Riesz basis (in fact, an orthogonal basis) for $L^2(A_{\geq n})$. Also, there exists a set $\Lambda_{M+1} \subset N\mathbb{Z}$ such that $E(\Lambda_{M+1})$ is a Riesz basis for $L^2\big[\frac{\{a\}}{N}, \frac{\{b\}}{N}\big)$; indeed, such a set Λ_{M+1} can be obtained from the result of Seip [18] with a dilation (see Lemma 4). Then Lemma 5 with shift factors '$n-1$' in place of 'n' yields that $E((\cup_{n=1}^{M} N\mathbb{Z}+n-1) \cup (\Lambda_{M+1}+M))$ is a Riesz basis for $L^2(S) = L^2(\frac{1}{N}V)$. By a dilation, we obtain that $E((\cup_{n=1}^{M} \mathbb{Z}+\frac{n-1}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N})) = E(\mathbb{Z} \cup (\cup_{k=1}^{M-1} \mathbb{Z}+\frac{k}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N}))$ is a Riesz basis for $L^2(V) = L^2([0,1) \cup [a,b))$. Now, we claim that $E(\Lambda')$ is a Riesz basis for $L^2[a,b)$, where $\Lambda' := (\cup_{k=1}^{M-1} \mathbb{Z}+\frac{k}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N})$. To see this, we again apply Lemma 5 (the original version) to the set $S' := \frac{1}{N}V'$ with $V' = [a,b)$. One can easily check that the corresponding set $A'_{\geq n}$ is equal to the set $A_{\geq n-1}$ above, that is,

$$A'_{\geq 1} = \cdots = A'_{\geq M-1} = [0, \tfrac{1}{N}) \ \supset \ A'_{\geq M} = \big[\tfrac{\{a\}}{N}, \tfrac{\{b\}}{N}\big) \ \supset \ A'_{\geq M+1} = \emptyset. \qquad (5)$$

Then Lemma 5 implies that $E((\cup_{k=1}^{M-1} \mathbb{Z}+\frac{k}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N}))$ is a Riesz basis for $L^2[a,b)$, as claimed.

Now, assume that $0 \leq \{b\} < \{a\} < 1$. Then there exists a number $M \in \mathbb{N}$ such that

$$A_{\geq 1} = \cdots = A_{\geq M} = [0, \tfrac{1}{N}) \ \supset \ A_{\geq M+1} = \big[0, \tfrac{\{b\}}{N}\big) \cup \big[\tfrac{\{a\}}{N}, \tfrac{1}{N}\big) \ \supset \ A_{\geq M+2} = \emptyset.$$

Again, using the result of Seip [18] with a dilation, we obtain a set $\Lambda_{M+1} \subset N\mathbb{Z}$ such that $E(\Lambda_{M+1})$ is a Riesz basis for $L^2\big[\frac{\{a\}}{N}, \frac{1+\{b\}}{N}\big)$. Since all elements in $E(N\mathbb{Z})$ are $\frac{1}{N}$-periodic, it follows that $E(\Lambda_{M+1})$ is a Riesz basis for $L^2\big(\big[0, \frac{\{b\}}{N}\big) \cup \big[\frac{\{a\}}{N}, \frac{1}{N}\big)\big)$. Then, by similar arguments as in the case $\{a\} \leq \{b\}$, we deduce that $E(\mathbb{Z} \cup (\cup_{k=1}^{M-1} \mathbb{Z}+\frac{k}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N}))$ is a Riesz basis for $L^2([0,1) \cup [a,b))$, and that $E((\cup_{k=1}^{M-1} \mathbb{Z}+\frac{k}{N}) \cup (\frac{1}{N}\Lambda_{M+1}+\frac{M}{N}))$ is a Riesz basis for $L^2[a,b)$.

Case $L \geq 2$. We will use essentially the same arguments as in the case $L = 1$, but employ the main result of Kozma and Nitzan [10] instead of Seip [18]. Given a set $V = [0,1) \cup [a_1, b_1) \cup \cdots \cup [a_L, b_L) \subset [0, N)$ with $L, N \in \mathbb{N}$ and $1 \leq a_1 < b_1 < \cdots < a_L < b_L \leq N$, let $S := \frac{1}{N}V = [0, \frac{1}{N}) \cup [\frac{a_1}{N}, \frac{b_1}{N}) \cup \cdots \cup [\frac{a_L}{N}, \frac{b_L}{N}) \subset [0,1)$. As in the case $L = 1$, we will apply Lemma 5 to this set S.

Note that there are finitely many possible ordering of the values

$$0 \ \leq \ \tfrac{\{a_1\}}{N}, \tfrac{\{b_1\}}{N}, \cdots, \tfrac{\{a_L\}}{N}, \tfrac{\{b_L\}}{N} \ < \ \tfrac{1}{N},$$

where equalities are also allowed, e.g., the values are all zero if all a_ℓ and b_ℓ are integers. It is easily seen that besides 0 and $\frac{1}{N}$, these are the only possible values that can be the boundary points of $A_{\geq n}$, $n = 1, \ldots, N$. In any case, since $[0, \frac{1}{N}) \subset S$ we have

$$A_{\geq 1} = [0, \tfrac{1}{N}) \ \supset \ A_{\geq 2} \ \supset \ \cdots \ \supset \ A_{\geq N},$$

where each of the sets $A_{\geq 2}, \ldots, A_{\geq N}$ is either empty or a finite union of intervals. One can therefore use the main result of [10] with a dilation, to construct sets

$\Lambda_1{=}N\mathbb{Z}, \Lambda_2, \Lambda_3, \ldots, \Lambda_N \subset N\mathbb{Z}$ such that for each n, the system $E(\Lambda_n)$ is a Riesz basis for $L^2(A_{\geq n})$. The rest of the proof is similar to the case $L = 1$. $\square$

Acknowledgments

A. Caragea acknowledges support by the DFG (German Research Foundation) Grant PF 450/11-1. D.G. Lee acknowledges support by the DFG Grants PF 450/6-1 and PF 450/9-1. The authors would like to thank Carlos Cabrelli, Diana Carbajal, and Felix Voigtlaender for valuable comments concerning the linear independence over $\mathbb{Q}$. The names of authors are ordered alphabetically by convention, and both authors contributed equally to this work.

Appendix A. Proof of Lemma 6

To prove Lemma 6, we will follow the proof strategy of Lemma 5 [10, Lemma 2]. For any $N \in \mathbb{N}$, a measurable set $S \subset [0,1)$, and $n = 0, 1, \ldots, N$, we define

$$A_n := \Big\{ t \in [0, \tfrac{1}{N}) : t + \tfrac{k}{N} \in S \ \text{ for exactly } n \text{ values of } \ k \in \{0, 1, \ldots, N-1\} \Big\},$$
$$B_n := \Big\{ t \in S : t + \tfrac{k}{N} \in S \ \text{ for exactly } n \text{ integers } \ k \in \mathbb{Z} \Big\}.$$

Obviously, considering the set B_n modulo $\frac{1}{N}$ yields the *n-fold* of A_n, which means that each element of A_n corresponds to exactly n points of B_n that are distanced apart by multiples of $\frac{1}{N}$. Note that $\{A_n\}_{n=0}^N$ and $\{B_n\}_{n=0}^N$ form partitions of $[0, \frac{1}{N})$ and $[0,1)$, respectively, that is, $[0, \frac{1}{N}) = \cup_{n=0}^N A_n$ and $[0,1) = \cup_{n=0}^N B_n$. Also, the family $\{B_n\}_{n=1}^N$ forms a partition of S, i.e., $S = \cup_{n=1}^N B_n$. For $f \in L^2(S)$ and $n = 1, \ldots, N$, we denote by f_n the restriction of f to B_n, that is, $f_n(t) = f(t)$ for $t \in B_n$ and 0 otherwise. This yields the decomposition $L^2(S) \ni f = f_1 + \ldots + f_N$ with all f_n having disjoint support. Note that the set $A_{\geq n}$ given by (1) can be expressed as $A_{\geq n} = \cup_{\ell=n}^N A_\ell$ for $n = 1, \ldots, N$. Similarly, we define $B_{\geq n} := \cup_{\ell=n}^N B_\ell$ and $f_{\geq n} := \sum_{\ell=n}^N f_\ell$ for $n = 1, \ldots, N$. For brevity, we write $\Lambda := \cup_{\ell=1}^N (\Lambda_\ell + j_\ell)$.

Frame. Assume that $\Lambda_1, \ldots, \Lambda_N \subset N\mathbb{Z}$ are such that $E(\Lambda_n)$ is a frame for $L^2(A_{\geq n})$. To prove that $E(\Lambda)$ is a frame for $L^2(S)$, it is enough to show that there exists a constant $c > 0$ satisfying

$$\sum_{\lambda \in \Lambda} \big| \langle f, e^{2\pi i \lambda(\cdot)} \rangle_{L^2(S)} \big|^2 \ \geq \ c \, \|f_n\|_{L^2(S)}^2 - \sum_{\ell=1}^{n-1} \|f_\ell\|_{L^2(S)}^2$$
$$\text{for all } \ f \in L^2(S) \ \text{ and } \ n = 1, \ldots, N.$$

In turn, it is enough to show that there exists a constant $c > 0$ satisfying

$$\sum_{\lambda \in \Lambda} \big| \langle f_{\geq n}, e^{2\pi i \lambda(\cdot)} \rangle_{L^2(S)} \big|^2 \ \geq \ c \, \|f_n\|_{L^2(S)}^2, \quad f \in L^2(S), \ \ n = 1, \ldots, N. \quad \text{(A.1)}$$

Such reductions are essentially due to the decomposition $f = f_1 + \ldots + f_N$ with all f_n having disjoint support, and due to $S \subset [0,1)$ and $\Lambda \subset \mathbb{Z}$; see [10, Eqns. (6)–(7)] for detailed arguments.

To prove (A.1), fix any $f \in L^2(S)$ and any $n \in \{1,\ldots,N\}$. Since $f_{\geq n}$ is supported in $\cup_{\ell=n}^N B_\ell \subset S$, we have for any $\lambda \in \Lambda_\ell + j_\ell$ with $\ell \in \{1,\ldots,N\}$,

$$\begin{aligned}\big\langle f_{\geq n}, e^{2\pi i\lambda(\cdot)}\big\rangle_{L^2(S)} &= \int_0^1 f_{\geq n}(t)\, e^{-2\pi i\lambda t}\, dt \\ &= \int_0^{1/N} \sum_{k=0}^{N-1} f_{\geq n}\big(t+\tfrac{k}{N}\big)\, \exp(-2\pi i\lambda(t+\tfrac{k}{N}))\, dt \\ &= \int_0^{1/N} h_{n,\ell}(t)\, e^{-2\pi i\lambda t}\, dt = \big\langle h_{n,\ell}, e^{2\pi i\lambda(\cdot)}\big\rangle_{L^2[0,\frac{1}{N})}\,,\end{aligned}$$

where

$$h_{n,\ell}(t) := \mathbb{1}_{A_{\geq n}}(t) \cdot \sum_{k=0}^{N-1} f_{\geq n}\big(t+\tfrac{k}{N}\big)\, e^{-2\pi i j_\ell k/N}. \tag{A.2}$$

Note that for $\ell = 1,\ldots,n$, the function $h_{n,\ell}$ is supported in $A_{\geq n} \subset A_{\geq \ell}$. Since $E(\Lambda_\ell)$ is a frame for $L^2(A_{\geq \ell})$, say, with lower frame bound $\alpha_\ell > 0$, we have

$$\begin{aligned}\sum_{\lambda\in\Lambda_\ell+j_\ell} \big|\langle f_{\geq n}, e^{2\pi i\lambda(\cdot)}\rangle_{L^2(S)}\big|^2 &= \sum_{\lambda\in\Lambda_\ell+j_\ell} \big|\langle h_{n,\ell}, e^{2\pi i\lambda(\cdot)}\rangle_{L^2[0,\frac{1}{N})}\big|^2 \\ &\geq \alpha_\ell\, \|h_{n,\ell}\|^2.\end{aligned} \tag{A.3}$$

Summing up (A.3) for $\ell = 1,\ldots,n$ gives

$$\begin{aligned}\sum_{\lambda\in\Lambda} \big|\langle f_{\geq n}, e^{2\pi i\lambda(\cdot)}\rangle_{L^2(S)}\big|^2 &\geq \sum_{\ell=1}^n \sum_{\lambda\in\Lambda_\ell+j_\ell} \big|\langle f_{\geq n}, e^{2\pi i\lambda(\cdot)}\rangle_{L^2(S)}\big|^2 \\ &\geq \Big(\min_{1\leq\ell\leq n} \alpha_\ell\Big) \cdot \sum_{\ell=1}^n \|h_{n,\ell}\|^2 \\ &\geq \Big(\min_{1\leq\ell\leq n} \alpha_\ell\Big) \cdot \sum_{\ell=1}^n \|h_{n,\ell}\cdot \mathbb{1}_{A_n}\|^2.\end{aligned} \tag{A.4}$$

On the other hand, for any fixed $t \in A_n$, Equation (A.2) becomes

$$h_{n,\ell}(t) = \sum_{k=0}^{N-1} f_{\geq n}\big(t+\tfrac{k}{N}\big)\, e^{-2\pi i j_\ell k/N}$$

and collecting the equation for $\ell = 1,\ldots,n$ gives the $n\times N$ linear system

$$\big[h_{n,\ell}(t)\big]_{\ell=1}^n = \big[e^{-2\pi i j_\ell k/N}\big]_{1\leq\ell\leq n, 0\leq k\leq N-1} \big[f_{\geq n}\big(t+\tfrac{k}{N}\big)\big]_{k=0}^{N-1}.$$

Since $t \in A_n$, the vector $[f_{\geq n}\big(t+\frac{k}{N}\big)]_{0\leq k\leq N-1}$ has exactly n nonzero entries, say, at the indices $k_1 < \ldots < k_n$ from $\{0,1,\ldots,N-1\}$. This reduces the above

system to an $n\times n$ linear system

$$\big[h_{n,\ell}(t)\big]_{\ell=1}^{n} = \big[e^{-2\pi i j_\ell k_r/N}\big]_{1\le\ell\le n,1\le r\le n}\,\big[f_{\ge n}\big(t+\tfrac{k_r}{N}\big)\big]_{r=1}^{n},$$

where the associated matrix $[e^{-2\pi i j_\ell k_r/N}]_{1\le\ell\le n,1\le r\le n}$ is invertible since N is prime (by Chebotarëv's theorem on roots of unity, see e.g., [19]). As there are only finitely many possible choices of $k_1<\ldots<k_n$ from $\{0,1,\ldots,N-1\}$, there exists a constant $c'>0$ such that

$$\sum_{\ell=1}^{n}\big|h_{n,\ell}(t)\big|^2 \;\ge\; c'\sum_{k=0}^{N-1}\big|f_{\ge n}\big(t+\tfrac{k}{N}\big)\big|^2 \quad \text{for all } \; t\in A_n.$$

Integrating over $t\in A_n$ then gives

$$\sum_{\ell=1}^{n}\|h_{n,\ell}\cdot \mathbb{1}_{A_n}\|^2 \ge c'\int_{A_n}\sum_{k=0}^{N-1}\big|f_{\ge n}\big(t+\tfrac{k}{N}\big)\big|^2 = c'\,\|f_n\|^2_{L^2(B_n)} = c'\,\|f_n\|^2_{L^2(S)}.$$

Combining this inequality with (A.4) yields the desired inequality (A.1).

Riesz sequence. Assume that $\Lambda_1,\ldots,\Lambda_N\subset N\mathbb{Z}$ are such that $E(\Lambda_n)$ is a Riesz sequence in $L^2(A_{\ge n})$. We will show that $E(\Lambda)$ is a Riesz sequence in $L^2(S)$ by using the *frame* part which is proved above (a similar trick was used in [11, Lemma 7]). Let $S':=[0,1)\backslash S$ and let $A'_{\ge n}$, $n=1,\ldots,N$, be the corresponding sets of (1) for S'. It is easily seen that $A'_{\ge n}=[0,\frac{1}{N})\backslash A_{\ge N+1-n}$ for $n=1,\ldots,N$. Since $E(\Lambda_{N+1-n})$ is a Riesz sequence in $L^2(A_{\ge N+1-n})$, we deduce from Proposition 3 with a dilation that the system $E(N\mathbb{Z}\backslash\Lambda_{N+1-n})$ is a frame for $L^2([0,\frac{1}{N})\backslash A_{\ge N+1-n})=L^2(A'_{\ge n})$; see the discussion after Proposition 3. The *frame* part then implies that the system $E(\cup_{n=1}^{N}((N\mathbb{Z}\backslash\Lambda_{N+1-n})+j_{N+1-n}))$ is a frame for $L^2(S')$. Finally, again by Proposition 3, we conclude that the system $E(\Lambda)=E(\cup_{n=1}^{N}(\Lambda_n+j_n))=E(\mathbb{Z}\backslash\cup_{n=1}^{N}((N\mathbb{Z}\backslash\Lambda_{N+1-n})+j_{N+1-n}))$ is a Riesz sequence in $L^2(S)=L^2([0,1)\backslash S')$.

Riesz basis. Since a family of vectors in a separable Hilbert space is a Riesz basis if and only if it is both a frame and a Riesz sequence, this part follows immediately by combining the *frame* and *Riesz sequence* parts. □

References

[1] E. Agora, J. Antezana, C. Cabrelli, *Multi-tiling sets, Riesz bases, and sampling near the critical density in LCA groups*, Adv. Math. 285 (2015) 454–477.

[2] M. Bownik, P. Casazza, A.W. Marcus, D. Speegle, *Improved bounds in Weaver and Feichtinger conjectures*, J. Reine Angew. Math. (Crelle) 749 (2019) 267–293 (published online: Aug. 2016).

[3] C. Cabrelli, D. Carbajal, *Riesz bases of exponentials on unbounded multi-tiles*, Proc. Amer. Math. Soc. 146 (2018) 1991–2004.

[4] C. Cabrelli, K. Hare, U. Molter, *Riesz bases of exponentials and the Bohr topology*, Proc. Amer. Math. Soc. 149 (2021) 2121–2131.

[5] O. Christensen, *An introduction to frames and Riesz bases*, second ed., Birkhäuser, 2016.

[6] A. Debernardi, N. Lev, *Riesz bases of exponentials for convex polytopes with symmetric faces*, in press, J. Eur. Math. Soc., arXiv:1907.04561, 2019.

[7] S. Grepstad, N. Lev. *Multi-tiling and Riesz bases*, Adv. Math. 252 (2014) 1–6.

[8] G.H. Hardy, E.M. Wright, *An introduction to the theory of numbers*, sixth ed., Oxford Univ. Press, 2008.

[9] M.N. Kolountzakis, *Multiple lattice tiles and Riesz bases of exponentials*, Proc. Amer. Math. Soc. 143 (2015) 741–747.

[10] G. Kozma, S. Nitzan, *Combining Riesz bases*, Invent. Math. 199 (2015) 267–285.

[11] G. Kozma, S. Nitzan, *Combining Riesz bases in* $\mathbb{R}^d$, Rev. Mat. Iberoam. 32 (2016) 1393–1406.

[12] G. Kozma, S. Nitzan, A. Olevskii, *A set with no Riesz basis of exponentials*, preprint arXiv:2110.02090, 2021.

[13] L. Kuipers, H. Niederreiter, *Uniform distribution of sequences*. Wiley, New York, 1974.

[14] D.G. Lee, *On construction of bounded sets not admitting a general type of Riesz spectrum*, preprint arXiv:2108.07760, 2021.

[15] Y. Lyubarskii, K. Seip, *A splitting problem for unconditional bases of complex exponentials*, in: Y. Lyubich et al. (Eds.), Entire functions in modern analysis: Boris Levin Memorial Conference (Tel-Aviv, 1997), Israel Math. Conf. Proc. vol. 15, Bar-Ilan Univ., Ramat Gan, 2001, pp. 233–242.

[16] B. Matei, Y. Meyer, *A variant of compressed sensing*, Rev. Mat. Iberoam. 25 (2009) 669–692.

[17] G.E. Pfander, S. Revay, D. Walnut, *Exponential bases for partitions of intervals*, preprint arXiv:2109.04441, 2021.

[18] K. Seip, *A simple construction of exponential bases in* L^2 *of the union of several intervals*, Proc. Edinb. Math. Soc. 38 (1995) 171–177.

[19] P. Stevenhagen, H.W. Lenstra, *Chebotarëv and his density theorem*, Math. Intell. 18 (1996) 26–37.

[20] I.M. Vinogradov, *Representation of an odd number as a sum of three primes*, Dokl. Akad. Nauk SSSR 15 (1937) 291–294.