Uniform estimates in one- and two-dimensional time-frequency analysis

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

MICHAŁ WARCHALSKI

aus
Katowice, Polen

Bonn 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Christoph Thiele
2. Gutachter: Prof. Dr. Herbert Koch

Tag der Promotion: 30.10.2018
Erscheinungsjahr: 2019

Abstract

This thesis is concerned with two special cases of the singular Brascamp-Lieb inequality, namely, the trilinear forms corresponding to the one- and two-dimensional bilinear Hilbert transform. In this work we study the uniform estimates in the parameter space of these two objects. The questions of the uniform bounds in one dimension arose from investigating Calderón's commutator, implying an alternative proof of its boundedness. Another reason for studying this problem is that, as the parameters degenerate, one can recover the bounds for the classical Hilbert transform, which is a well understood operator. Analogously, it is natural to investigate the two dimensional form, whose parameter space turns out to be considerably more involved and offering many more questions concerning the uniform bounds.

The thesis consists of four chapters.

In Chapter 1 we investigate the parameter space of the bilinear Hilbert transform. We complete the classification of the two dimensional form that was first given by Demeter and Thiele. We also describe the parameter space, reducing its dimensionality, and discuss the related geometry, which raises many open questions concerning the uniform bounds in two dimensions.

In Chapter 2 we prove the uniform bounds for the bilinear Hilbert transform in the local L^1 range, which extends the previously known range of exponents for this problem. This a joint work with Gennady Uraltsev.

In Chapter 3, which is an elaboration on Chapter 2, we prove the uniform bounds for the Walsh model of the bilinear Hilbert transform in the local L^1 range in the framework of the iterated outer L^p spaces. This theorem was already proven by Oberlin and Thiele, however, in their work they did not use the outer measure structure.

Finally, Chapter 4 is dedicated to proving the uniform bounds for the Walsh model of the two dimensional bilinear Hilbert transform, in a two parameter setting in the vicinity of the triple that corresponds to the two dimensional singular integral.

Acknowledgements

First and foremost, I would like to thank my advisor, Professor Christoph Thiele, for introducing me to time-frequency analysis and guiding me through most of my mathematical path. I am indebted to him, for his generosity in sharing his wide knowledge, constant optimism, great patience and tremendous amount of time he dedicated to me during my studies. They are all invaluable to me.

I would like to thank my older academic brother and collaborator, Gennady Uraltsev, for the plethora of discussions we had, during which I learned incredibly much mathematics from him. Without his strong intuition, always positive attitude and passion for science, we definitely would not have been able to complete our joint work.

I am indebted to Mariusz Mirek, for his great support in the beginning of my mathematical journey, without whom my PhD studies would have probably not happened. I am also grateful to Diogo Oliveira e Silva, Olli Saari and Pavel Zorin-Kranich for all the discussions we had, which made me understand harmonic analysis better.

I would like to thank the rest of my academic siblings at the University of Bonn, for making my PhD studies much more joyful and for being always very helpful, not only with mathematics: Polona Durcik, Marco Fraccaroli, Shaoming Guo, João Pedro Ramos, Johanna Richter, Joris Roos.

I am grateful to the other members of the Analysis and PDE group, for having a possibility to do exciting mathematics in this extremely friendly environment. I would also like to acknowledge the support I received from the Hausdorff Center for Mathematics and the Bonn Graduate School of Mathematics as well as thank them for organizing many stimulating scientific events.

I would like to thank Professor Paweł Głowacki for giving me a beautiful introduction to mathematical analysis and his support in the beginning of my studies.

I am grateful to all my friends, fellow students at and outside of the University of Bonn, for making all these years much more enjoyable and exciting. Unfortunately, there would not be enough space on this page to mention all of them.

Last but not least, I am thankful to my family and Ana, for the unlimited support they give me every day. I would like to dedicate this thesis to them.

Contents

Introduction

In this thesis we are concerned with a singular variant of the Brascamp-Lieb inequality, whose classical version is defined as

$$\int_{\mathbb{R}^m} \prod_{j=1}^{n} F_j(\Pi_j x)\,dx \le C \prod_{j=1}^{n} \|F_j\|_{L^{p_j}(\mathbb{R}^{k_j})}, \tag{0.1}$$

where $F_j \in \mathbb{R}^m \to \mathbb{C}$ are measurable functions and $\Pi_j\colon \mathbb{R}^m \to \mathbb{R}^{k_j}$ and $\Pi\colon \mathbb{R}^m \to \mathbb{R}^k$ are surjective linear maps. Bennett, Carbery, Christ and Tao in [Ben+08] gave a complete description of (0.1), proving that the above inequality holds if and only if for every subspace V of $\mathbb{R}^m$ it holds that $\dim(V) \le \sum_{j=1}^{n} \frac{1}{p_j} \dim(\Pi_j V)$ together with the equality for $V = \mathbb{R}^m$. When one integrates the product against a Calderón-Zygmund kernel in (0.1), then it becomes a so-called singular Brascamp-Lieb inequality. It is generally of the form

$$\int_{\mathbb{R}^m} \prod_{j=1}^{n} F_j(\Pi_j x) K(\Pi x)\,dx \le C \prod_{j=1}^{n} \|F_j\|_{L^{p_j}(\mathbb{R}^{k_j})}, \tag{0.2}$$

where $\Pi\colon \mathbb{R}^m \to \mathbb{R}^k$ is a surjective linear map and K is a Calderón-Zygmund kernel on $\mathbb{R}^k$. Multilinear inequalities of the form (0.2) form a very vast family of problems and cover a large portion of questions considered in harmonic analysis including, among others, the classical Hilbert transform, paraproducts, the bilinear Hilbert transform and the simplex Hilbert transform. Various examples of singular Brascamp-Lieb integrals were thoroughly discussed in the work of Durcik [Dur17], where she proved multilinear L^p estimates for a so-called entangled form, which falls into this general class.

In this dissertation we are interested in two special cases of the multilinear form appearing in (0.2):

- The trilinear form associated with the bilinear Hilbert transform

$$\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3) := \int_{\mathbb{R}^2} \prod_{j=1}^{3} f_j(x - \beta_j t)\,dx\,\frac{dt}{t}, \tag{0.3}$$

  where f_1, f_2, f_3 are Schwartz functions on $\mathbb{R}$ and $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$ with $\sum_{j=1}^{3} \beta_j = 0$. Note that the above trilinear form is obtained from (0.2), assuming $m = 2$, $n = 3$, $\Pi_j(x,t) = x - \beta_j t$ for $j = 1, 2, 3$ and $K(t) = 1/t$, $\Pi(x,t) = t$.

- the trilinear form associated with the two dimensional bilinear Hilbert transform

$$\mathrm{BHF}^K_{\vec{B}}(g_1, g_2, g_3) := \int_{\mathbb{R}^4} \prod_{j=1}^{3} g_j((x,y) + B_j(s,t)) K(s,t)\,dx\,dy\,ds\,dt, \tag{0.4}$$

where g_1, g_2, g_3 are Schwartz functions on $\mathbb{R}^2$, $\vec{B} = (B_1, B_2, B_3) \in (\mathbb{R}^{2\times 2})^3$ is a triple of 2×2 real matrices with $\sum_{j=1}^3 B_j = 0$ and $K\colon \mathbb{R}^2 \setminus \{0,0\} \to \mathbb{R}$ is a two dimensional Calderón-Zygmund kernel, i.e. satisfying

$$|\partial^\alpha \hat{K}(\xi,\eta)| \leq |(\xi,\eta)|^{-|\alpha|},$$

for all $\alpha \in \mathbb{Z}_+^2$ up to a high order and $(\xi,\eta) \neq (0,0)$. Note that the above trilinear form is obtained from (0.2), assuming $m = 4$, $n = 3$, $\Pi_j(x,y,s,t) = (x,y) - B_j(s,t)$ for $j = 1,2,3$ and $\Pi(x,y,s,t) = (s,t)$.

One is interested in proving the singular Brascamp-Lieb inequalities in those two special cases

$$|\operatorname{BHF}_{\vec{\beta}}(f_1,f_2,f_3)| \leq C_{p_1,p_2,p_3,\vec{\beta}} \prod_{j=1}^3 \|f_j\|_{L^{p_j}(\mathbb{R})}, \tag{0.5}$$

$$|\operatorname{BHF}^K_{\vec{B}}(g_1,g_2,g_3)| \leq C_{p_1,p_2,p_3,\vec{B}} \prod_{j=1}^3 \|g_j\|_{L^{p_j}(\mathbb{R}^2)}. \tag{0.6}$$

By scaling, the exponents in (0.5), (0.6) should satisfy $1/p_1 + 1/p_2 + 1/p_3 = 1$.

In Chapter 1 we study the geometry of the parameter space of the bilinear Hilbert transform. While it is well understood in one dimension, it is a more involved object in two dimensions. One attempt to classify various cases in two dimensions was given in [DT10] by Demeter and Thiele, however they did not how the parameters degenerate. The main purpose of this chapter is classification of $\vec{B}$ up to symmetries that do not affect the defining constants of the kernel K, hence giving a good description of the related geometry. This makes it a good starting point for proving (0.6) with a constant independent of $\vec{B}$, which is a completely open problem. Below we discuss some background in one and two dimensions, and the content of this chapter.

Observe that up to a symmetry, there are essentially 3 different cases of (0.3). If one assumes that all β_j are equal, then using the translation symmetry $\operatorname{BHF}_{\vec{\beta}}$ equals zero and the inequality (0.5) is clearly satisfied. If two of the components of $\vec{\beta}$ are the same, then $\operatorname{BHF}_{\vec{\beta}}$ up to a symmetry it equals

$$\int_{\mathbb{R}} Hf_1(x)\, f_2(x)\, f_3(x)\, dx,$$

which implies (0.5) for $1 < p_1, p_2, p_3 < \infty$ using the boundedness of the Hilbert transform. The third possibility is when β_j's are pairwise distinct. The first proof in this case was given by Lacey and Thiele in [LT97], where they proved (0.5) in the range $2 < p_1, p_2, p_3 < \infty$.

The dependence of the constant in (0.5) is not explicitly stated in terms of $\vec{\beta}$ in [LT97], however, one can show it that it behaves linearly in $\min_{i\neq j} |\beta_i - \beta_j|^{-1}$. The authors of [LT99] asked, whether there exists a constant $C_{p_1,p_2,p_3} < \infty$ independent of $\vec{\beta}$, such that

$$|\operatorname{BHF}_{\vec{\beta}}(f_1,f_2,f_3)| \leq C_{p_1,p_2,p_3} \prod_{j=1}^3 \|f_j\|_{L^{p_j}(\mathbb{R})} \tag{0.7}$$

holds for triples of Schwartz functions and, moreover, what is the range of exponents in which the above inequality holds. This question has already been extensively studied by Thiele [Thi02a], Grafakos and Li [GL04] and Li [Li06]. Since the following chapter is concerned with extending the range of exponents for (0.7) we discuss the background of this problem in detail later on. Concerning the geometry, applying the translation and the dilation symmetry of the form it is not difficult to show that the parameter space in one dimension can be identified with $S^1 \cup \{0\}$, where $\{0\}$ corresponds to the aforementioned trivial 0 form, a finite set of points on S^1 is identified with the Hilbert transform and the rest of the circle corresponds to the nondegenerate case.

As opposed to the one dimensional form, (0.4) has in total 10 different cases. First, if one assumes that that all B_1, B_2, B_3 are singular, then (0.4) degenerates to a one dimensional operator or a strongly singular two dimensional operator. In the first case its boundedness follows from the one dimensional time-frequency analysis and paraproduct theory. Otherwise, as shown in Chapter 1, it is an operator whose boundedness is strongly related to the boundedness of the triangular Hilbert transform. The latter is known to be a difficult open problem and in fact this is the only case in which (0.6) is not known. If one assumes that $\vec{B}$ is such that at least one of B_1, B_2, B_3 is nonsingular, then there are several possibilities: it is a fully two dimensional form, a so-called one and half dimensional form or a so-called twisted paraproduct. In [DT10] Demeter and Thiele gave the first proof in the first two cases, for exponents satisfying $2 < p_1, p_2, p_3 < \infty$. Their methods consisted of using two dimensional as well as one and half dimensional time-frequency analysis. The latter case was later resolved by Vjekoslav Kovač in [Kov12], which initiated the so-called twisted technology. The authors of [DT10] provided also a classification of the cases, assuming that one of the matrices is nonsingular. In the first main result of Chapter 1, Theorem 1.8, we complete the classification given in [DT10], including the aforementioned cases when all B_1, B_2, B_3 are singular.

Similarly as in one dimension, it is natural to ask whether L^p bounds hold with a constant independent of $\vec{B}$, which is not provided by the methods in [DT10]. This brings us to the conjecture.

Conjecture 1. *Let K be a Calderón-Zygmund kernel satisfying* (1.2) *and assume that* $2 < p_1, p_2, p_3 < \infty$ *with* $\sum_{j=1}^{3} 1/p_j = 1$. *There exists a constant* $0 < C_{p_1,p_2,p_3} < \infty$, *such that for all* $g_1, g_2, g_3 \in \mathcal{S}(\mathbb{R}^2)$

$$|\operatorname{BHF}_{\vec{B}}^{K}(g_1, g_2, g_3)| \leq C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|g_j\|_{L^{p_j}(\mathbb{R}^2)} \tag{0.8}$$

holds uniformly in $\vec{B} \in (\mathbb{R}^{2\times 2})^3$.

As described above, the parameter space of (0.4) is much richer than of its one dimensional counterpart and there is a number possibilities in which $\vec{B}$ can approach various degenerate cases. The conjecture is completely open, except for the cases which correspond to the one dimensional bilinear Hilbert transform and the bounds follows from the one dimensional theory of the uniform estimates. Since (0.8) implies the boundedness of the triangular Hilbert transform, which is a difficult open problem, the full version of Conjecture 1 seems to be out of reach for the current state of the art. However, there are several different degenerations for which the L^p bounds are known to hold. The main goal of Chapter 1 is to study the geometry of triples $\vec{B}$, which possibly makes it a good starting point for studying Conjecture 1 further. The authors of [DT10] were not concerned with the uniform estimates and did not consider how applying the symmetries of (1.1) affects the kernel K. In the main theorem of Chapter 1, Theorem 1.13 we describe the manifold of parameters in two dimensions, up to only these symmetries that do not change the defining constants of K, essentially identifying it with $(S^1)^3 \cup (S^1)^2 \cup \{0\}$. This is motivated by the aforementioned parametrization in one dimension by $S^1 \cup \{0\}$. The parametrization in one dimension is significantly easier, because all matrices in one dimension commute. Since this is not the case in two dimensions, it requires more care to carry out a similar process.

Chapter 2 is concerned with extending the range of exponents for the one dimensional inequality (0.5). The content of Chapter 2 is a joint work with Gennady Uraltsev. Below we discuss the background and its content.

First estimates of the type (0.5) were given by Lacey and Thiele in [LT97], in the range $2 < p_1, p_2, p_3 < \infty$, corresponding to the open triangle c in Figure 1. They subsequently extended the range of exponents for the inequality in [LT99] the open triangles a_1, a_2, a_3 in Figure 1. The works [LT97], [LT99], inspired by the works of Carleson [Car66] and Fefferman [Fef73], initiated the modern time-frequency analysis.

In [LT97] Lacey and Thiele proved that (0.5) holds in the range of exponents $2 < p_1, p_2, p_3 < \infty$, with a constant dependent only on p_1, p_2, p_3 and $\vec{\beta}$, which corresponds to the open triangle c in Figure 1. The range was extended in [LT99] to the one corresponding to the convex hull of the open triangles a_1, a_2, a_3 in Figure 1. The works [LT97], [LT99], inspired by the works of Carleson [Car66] and Fefferman [Fef73], initiated the modern time-frequency analysis.

Since the form (1.4) is symmetric under permutations of the coordinates of $\vec{\beta}$, let us assume from now on that $\vec{\beta}$ is in the neighbourhood of the degenerate case $\beta_2 = \beta_3$. In this case the trilinear form becomes (2.3) and the Hilbert transform is not bounded in L^∞, thus one cannot expect the uniform bounds to hold for $\alpha_1 \leq 0$. This region corresponds in Figure 1 to the one below the line spanned by $(0, 0, 1)$, $(0, 1, 0)$. Moreover, the region spanned by a_1, a_2, a_3 in the picture is the maximal range for which parameter dependent bounds for the bilinear Hilbert transform are known. Taking the intersection of these two regions we obtain the convex hull of the open triangles b_3, b_2, a_3 and a_2.

The uniform estimate (0.7) was investigated in several papers. The first time inequality (0.7) was proven with a constant independent of $\vec{\beta}$ by Thiele in [Thi02a], where he showed a weak type inequality at the two upper corners of the triangle c in Figure 1. Next (0.7) was proven by Grafakos and Li in [GL04] in the open triangle c and Li [Li06] extended the bounds to the range corresponding to the open triangles a_1, a_2. Interpolating these results, one obtains (0.7) for the exponents corresponding the convex hull of the open triangles a_2, a_3 and c, see Figure 1. However, up to date, it was not known whether the uniform bounds hold in the neighbourhood of points $(1/p_1, 1/p_2, 1/p_3) = (0, 0, 1)$, $(0, 1, 0)$. The following main result of Chapter 2 resolves this issue.

Theorem 0.1. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $\vec{\beta}$ and all triples of Schwartz functions f_1, f_2, f_3 the inequality* (0.7) *holds.*

The range of exponents in the above theorem corresponds to the convex hull of the open triangles b_1, b_2, b_3. This extends the uniform inequality (0.7) to the exponents corresponding to the convex hull of the open triangles a_2, a_3, b_2 and b_3 in Figure 1, after interpolating with the theorem of Li [Li06].

In order to prove Theorem 0.1, we refine the outer measure approach progressively developed in the papers [DT15], [DPO15], [Ura16]. This approach was initiated in the paper [DT15], where Do and Thiele reformulated the problem of boundedness of the bilinear Hilbert transform into proving an outer Hölder inequality on the upper half space $\mathbb{R}^3_+ := \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$, which can be identified with the symmetries of (0.3), and an embedding theorem for exponents in the range $2 < p < \infty$. In [DPO15] Di Plinio and Ou extended it to the range $1 < p < \infty$, which was afterwards reformulated by Uraltsev in [Ura16] as an iterated embedding theorem. The approach of [DT15] using the refinements of [DPO15] and [Ura16] can be very roughly outlined as follows. One embeds any Schwartz function on $\mathbb{R}$, f via

$$F^\varphi(f)(y, \eta, t) := f * \varphi_{\eta,t}(y)$$

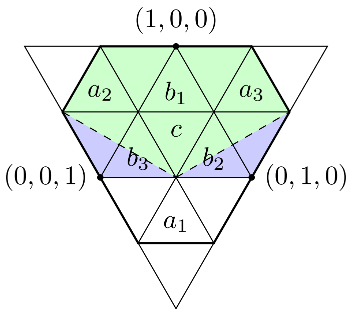

Figure 1: Range of exponents $(\alpha_1, \alpha_2, \alpha_3) = (1/p_1, 1/p_2, 1/p_3)$ with $\sum_{j=1}^{3} \alpha_j = 1$. The uniform bounds were previously known to hold in the convex hull of the open triangles a_2, a_3 and c. The result of Chapter 2 implies the uniform bounds in the convex of the open triangles a_2, a_3, b_2 and b_3.

where φ is a Schwartz function with sufficiently small support. Performing the wave packet decomposition one essentially rewrites

$$\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3) \approx \int_{\mathbb{R}^3_+} \prod_{j=1}^{3} F^{\varphi}(f_j)(y, \alpha_j \eta + \delta\beta_j t^{-1}, |\alpha_j|^{-1} t)\, dt\, d\eta\, dy,$$

where $\vec{\alpha} \in \mathbb{R}^3$ is the unit vector perpendicular to both $(1, 1, 1)$ and $\vec{\beta}$, and $\delta := \min(|\alpha_1|, |\alpha_2|, |\alpha_3|)$. Applying the outer Hölder inequality [DT15] and using the embedding theorem from [DPO15] for each f_j separately in the framework of [Ura16], the right hand side of the previous display is bounded by

$$\prod_{j=1}^{3} \|F(f_j)\|_{L^{p_j}\mathcal{L}^{q_j}(S)} \lesssim \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R})}. \tag{0.9}$$

On the left hand side are the outer L^p norms that we precisely introduce in Chapter 2. We follow the above approach and the main difficulty in our case is to prove a trilinear inequality for the wave packet decomposition of $\mathrm{BHF}_{\vec{\beta}}$, with a constant uniform in the parameter $\vec{\beta}$. We then complete the proof combining that trilinear inequality with (0.9).

Chapter 3 and Chapter 4 are dedicated to proving Walsh analogues of (0.7) and (0.8) respectively. The so-called Walsh models of multilinear forms are often studied by time-frequency analysts along with their continuous analogues, as many technical issues disappear due to perfect time-frequency localization of the Walsh wave packets. On the other hand, they are still similar enough to the original problem, so that they are a well established way for understanding and presenting the gist of the problem. Walsh models appeared in the context of the bilinear Hilbert transform in a number of articles, for example, [Thi95], [Thi02b], [OT11], [DDP13]. Below, we first discuss the content Chapter 3 and then we discuss the content of Chapter 4.

Oberlin and Thiele in [OT11] proved the uniform inequality (0.7) for a Walsh model of the bilinear Hilbert transform in the range that corresponds to the convex hull of the open triangles a_2, a_3, b_2 and b_3 in Figure 1. In Chapter 3, we reprove the result of [OT11] in the local L^1 range in the framework of the outer L^p spaces. This can be thought as a demonstration of the techniques that are used in Chapter 2 in the context of the continuous form.

In order to define the Walsh model we introduce the set of tiles, where the wave packets are time-frequency localized. We call a tile the Cartesian product $I \times \omega$, where $I, \omega \subset \mathbb{R}_+$ are dyadic intervals and denote the set of tiles with $\mathbb{X}$. The L^2 normalized wave packets associated with tiles are defined recursively via the following identities

$$\varphi_{I\times[0,|I|^{-1})} = |I|^{-1/2}\mathbb{1}_I(x), \qquad \varphi_{J^-\times\omega} + \varphi_{J^+\times\omega} = \varphi_{J\times\omega^-} + \varphi_{J\times\omega^+},$$

for any dyadic intervals $I, J, \omega \subset \mathbb{R}_+$ with $|J||\omega| = 2$, where J^- and J^+ are dyadic children of J. Similarly as in the continuous case, given a function $f \in \mathcal{S}(\mathbb{R})$ we associate it with the embedded function via

$$F(f)(P) = \int f(x)\varphi_P(x)\,dx,$$

where φ_P is the Walsh wave packet associated with a tile P. Set $F_j = F(f_j)$ for $j = 1,2,3$. We indicate the dyadic sibling of a dyadic interval I as $I^{\circlearrowleft}$ and by $P^{\circlearrowleft}$ the tile $I_P \times \omega_p^{\circlearrowleft}$. The trilinear form on the embedded functions associated to the Walsh bilinear Hilbert transform is given for $L \in \mathbb{N}$ by

$$\Lambda_L(F_1, F_2, F_3) := \sum_{P\in\mathbb{X}} |I_P|^{-1/2} F_1(P^{\circlearrowleft}) \sum_{Q\in P^L} F_2(Q)F_3(Q)h_{I_P}(c(I_Q)),$$

where $P^L = \{Q \in \mathbb{X} \colon I_Q \subset I_P,\ |I_Q| = 2^{-L}|I_P|,\ \omega_Q = 2^L\omega_P\}$. In the above expression we used the Haar function h_{I_p} and the center of the interval I_Q, $c(I_Q)$.

The main result of Chapter 3 is the following theorem.

Theorem 0.2. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$ and $1/q_1 + 1/q_2 + 1/q_3 > 1$ with $2 < q_1, q_2, q_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $L \geq 2$ and all triples of Schwartz functions f_1, f_2, f_3*

$$|\Lambda_L(F(f_1), F(f_2), F(f_3))| \leq C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|F(f_j)\|_{L^{p_j}\mathcal{L}^{q_j}(S)}. \tag{0.10}$$

On the right hand side of (0.10) are iterated outer L^p norms developed in [Ura16] that we define precisely in Section 3. Each of them separately can be controlled using the Walsh iterated embedding theorem, proved by Uraltsev in [Ura17], so that the right hand side of (0.10) is bounded by $\prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R})}$. The results of Chapter 3 and Chapter 2 are a continuation of studies in [War15], where the uniform bounds on Λ_L were proven in the local L^2 range.

In Chapter 4 we study a Walsh model of (0.8) for diagonal triples $\vec{B}$ that approach the trilinear form associated with the dimensional singular integral. This can be seen as the simplest setting for two parameter uniform bounds and thus, it is a natural question to investigate first. Below we discuss the content of this chapter.

We call a multitile the Cartesian product $R \times \Omega$, where $R := I_1 \times I_2, \Omega := \omega_1 \times \omega_2 \subset \mathbb{R}_+$ are dyadic rectangles and $|I_j||\omega_j| = 1$ for $j = 1,2$. Here we denote the set of multitiles with $\mathbb{X}$. The L^2 normalized wave packet associated with a multitile P is defined as

$$\varphi_P(x,y) := \varphi_{P_1}(x)\varphi_{P_2}(y),$$

where for $j = 1,2$, $P_j = I_j \times \omega_j$ and φ_{P_j} is the one dimensional Walsh wave packet.

Given a Schwartz function f on $\mathbb{R}^2$ we associate it to the embedded function via

$$F(f)(P) = \langle f, \varphi_P \rangle.$$

Let $f_1,\ f_2,\ f_3$ be a triple of Schwartz functions on $\mathbb{R}^2$. Set $F_j = F(f_j)$ for $j = 1, 2, 3$. For a multitile $P = R \times \Omega$, where $\Omega = \omega_1 \times \omega_2$ we denote

$$\Omega^{\circlearrowleft} = \omega_1^{\circlearrowleft} \times \omega_2, \qquad\qquad P^{\circlearrowleft} = R \times \Omega^{\circlearrowleft},$$

where $\omega^{\circlearrowleft}$ is the dyadic sibling of a dyadic interval ω. For a $K \in \mathbb{Z}$ we denote with $\mathcal{R}^K$ the set of all dyadic rectangles $I \times J$ with $|I| = 2^K |J|$ and denote with $\mathbb{X}^K$ the set of all multitiles $P = R \times \Omega$ with $R \in \mathcal{R}^K$. Given $K, L \in \mathbb{N}$, we define the trilinear form on the embedded functions associated with the two dimensional Walsh bilinear Hilbert transform by

$$\Lambda_{K,L}(F_1, F_2, F_3) := \sum_{P \in \mathbb{X}} |R_P|^{-1/2} F_1(P^{\circlearrowleft}) \sum_{Q \in P^{K,L}} F_2(Q) F_3(Q) h_{R_P}(c(R_Q)),$$

where for $P \in \mathbb{X}$, $P^{K,L} = \{Q \in \mathbb{X}^K \colon R_Q \subset R_P,\ \Omega_Q = \Omega_P^{K,L}\}$, $c(R_Q)$ is the center of R_Q and $\Omega_P^{K,L} := 2^L \omega_1 \times 2^{L+K} \omega_2$ for $\Omega_P = \omega_1 \times \omega_2$. Moreover, $h_{R_P}(x, y) = \varphi_P(x, y) \varphi_{P^{\circlearrowleft}}(x, y)$.

The goal of Chapter 4 is to prove the uniform bounds for the Walsh model of the two dimensional bilinear Hilbert transform modularizing it as an iterated outer L^p estimate for $\Lambda_{K,L}$ uniform in K, L and the Walsh iterated embedding theorem. Here is the main theorem of this chapter.

Theorem 0.3. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$ and $1/q_1 + 1/q_2 + 1/q_3 > 1$ with $2 < q_1, q_2, q_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $K, L \geq 2$, all triples of Schwartz functions $f_1,\ f_2,\ f_3$*

$$|\Lambda_{K,L}(F(f_1), F(f_2), F(f_3))| \leq C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|F(f_j)\|_{L^{p_j} \text{Ł}^{q_j}(S)}. \tag{0.11}$$

On the right hand side of (0.11) are the two dimensional counterparts of the iterated outer L^p norms developed in [Ura16] that we define precisely in Chapter 4. The two dimensional Walsh iterated embedding theorem, which we prove in Section 5 of Chapter 4, implies that for $j = 1, 2, 3$

$$\|F(f_j)\|_{L^{p_j} \text{Ł}^{q_j}(S)} \leq C_{p_j} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}.$$

We record that the uniform bounds for a Walsh model of the two dimensional bilinear Hilbert transform were already studied in [War15], where they were proven in the one-parameter case $K \geq 2$, $L = \infty$.

Notation

We write $A \lesssim B$, if there exists a positive and finite constant such that $A \leq CB$ and its value in the argument is either absolute or irrelevant. We also write $A \simeq B$ if $A \lesssim B$ and $B \lesssim A$. We write $A \lesssim_p B$ if $C = C_p$ depends on a parameter p. We also usually discard factors involving π, coming from the Fourier transform or its inverse.

Chapter 1

Parameter space of the bilinear Hilbert transform

1.1 Introduction

It is well known that the trilinear form associated with the one dimensional bilinear Hilbert transform can be parametrized by $S^1 \cup \{0\}$, where the trilinear forms corresponding to the Hilbert transform are associated with a finite subset on the circle and the origin corresponds to the trivial 0 form. In this chapter we are mostly concerned with the parameter space of the trilinear form associated with the two dimensional Hilbert transform, defined as

$$\mathrm{BHF}_{\vec{B}}^{K}(f_1, f_2, f_3) := \int_{\mathbb{R}^4} \prod_{j=1}^{3} f_j((x,y) + B_j(s,t))K(s,t)\, dx\, dy\, ds\, dt, \tag{1.1}$$

where f_j are Schwartz functions on $\mathbb{R}^2$, $\vec{B} = (B_1, B_2, B_3) \in (\mathbb{R}^{2\times 2})^3$ is a triple of 2×2 real matrices and $K\colon \mathbb{R}^2 \setminus \{0,0\} \to \mathbb{R}$ is a two dimensional Calderón-Zygmund kernel, i.e. satisfying

$$|\partial^\alpha \hat{K}(\xi,\eta)| \le |(\xi,\eta)|^{-|\alpha|}, \tag{1.2}$$

for all $\alpha \in \mathbb{Z}_+^2$ up to a high order and $(\xi,\eta) \neq (0,0)$. One is interested in proving the inequality for all triples of Schwartz functions on $\mathbb{R}^2$

$$|\,\mathrm{BHF}_{\vec{B}}^{K}(f_1, f_2, f_3)| \le C_{p_1,p_2,p_3,\vec{B}} \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}, \tag{1.3}$$

for exponents satisfying $\sum_{j=1}^{3} \frac{1}{p_j} = 1$, which is dictated by scaling.

The goal of this chapter is to describe the parameter space $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ by exploiting its symmetries. Such parametrization is more challenging than in the one dimensional case, since the 2×2 matrices do not commute in general. In Theorem 1.8 we complete the classification of cases for the two dimensional bilinear Hilbert transform that appeared already in the paper by Demeter and Thiele [DT10], where we include some more degenerate forms. In Theorem 1.14 describe the parameter manifold in two dimensions, essentially as $(S^1)^3 \cup (S^1)^2 \cup \{0\}$. The point of this parametrization is that we use only these symmetries that do not affect the constant in (1.3). Therefore, it is a good starting point for studying the inequality (1.3) uniformly in $\vec{B}$.

In Section 1.2 we recall the parametrization of the one dimensional bilinear Hilbert transform. After that we introduce and state the main results of this chapter in Section 1.3, and make connections with known results and open problems in two dimensional time-frequency analysis. Section 1.4 contains the proofs of our main results. Finally, in the last section we make some further remarks about the uniform bounds in two dimensions.

1.2 Prelude - parametrization in one dimension

In the following we quickly recall the degenerate cases and the parametrization of the one dimensional bilinear Hilbert transform. For convenience of the reader, we recall that it is given for a triple of Schwartz functions on $\mathbb{R}$ by

$$\mathrm{BHF}^{1D}_{\vec{\beta}}(f_1, f_2, f_3) = \int_{\mathbb{R}^2} \prod_{j=1}^{3} f_j(x - \beta_j t)\, dx\, \frac{dt}{t}. \tag{1.4}$$

One is interested in the estimate for triples of Schwartz functions

$$\mathrm{BHF}^{1D}_{\vec{\beta}}(f_1, f_2, f_3) \le C_{p_1,p_2,p_3,\vec{\beta}} \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R})}. \tag{1.5}$$

with $\sum_{j=1}^{3} 1/p_j = 1$ dictated by scaling. Next, we define a function that differentiates between degenerate and nondegenerate cases for (1.4).

Definition 1.1. *Let $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$. Define*

$$h^{1D}(\vec{\beta}) = (r(\beta_2 - \beta_3), r(\beta_3 - \beta_1), r(\beta_1 - \beta_2)),$$

where $r(A)$ denotes the rank rank of a matrix (in this case, either 0 or 1). We call $\vec{\beta}$ degenerate if $h^{1D}(\vec{\beta}) \neq (1,1,1)$ and nondegenerate otherwise.

The one dimensional bilinear Hilbert transform is called degenerate if one of the ranks above equals zero. More precisely, here are all the possibilities.

Proposition 1.2. *Let $\vec{\beta} \in \mathbb{R}^3$. Up to a permutation of β_1, β_2, β_3 it satisfies one and only of the following conditions*

$$h^{1D}(\vec{\beta}) = (1,1,1), \tag{1.6}$$

$$h^{1D}(\vec{\beta}) = (1,1,0), \tag{1.7}$$

$$h^{1D}(\vec{\beta}) = (0,0,0). \tag{1.8}$$

Remark 1.3. *Note that $h^{1D}(\vec{\beta}) = (1,0,0)$ is not possible.*

In order to reduce dimensionality of the parameter space one exploits the symmetries of the trilinear form. By simple change of variables we have the following.

Proposition 1.4. *Let f_1, f_2, f_3 be three Schwartz functions on $\mathbb{R}$. Assume that $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$. Moreover, let $a \in \mathbb{R}$. Then*

- *Translation invariance: we have*

$$\mathrm{BHF}^{1D}_{\vec{\beta}}(f_1, f_2, f_3) = \mathrm{BHF}^{1D}_{\vec{\beta}-(a,a,a)}(f_1, f_2, f_3).$$

- *Multiplication invariance: if $a \neq 0$, then we have*

$$\mathrm{BHF}^{1D}_{\vec{\beta}}(f_1, f_2, f_3) = \mathrm{BHF}^{1D}_{a\vec{\beta}}(f_1, f_2, f_3).$$

Remark 1.5. *Observe that the above invariances do not change the constant in* (1.5)*.*

For all $\vec{\beta}$ satisfying (1.6) the proof of (1.5) is essentially the same and requires time-frequency analysis [LT97], [LT99]. Assuming that $\vec{\beta}$ satisfies (1.7), boundedness of $\mathrm{BHF}^{1D}_{\vec{\beta}}$ is equivalent to boundedness of the Hilbert transform. If $\vec{\beta}$ satisfies (1.8), then by the translation symmetry of the form it is easy to verify that $\mathrm{BHF}^{1D}_{\vec{\beta}}$ equals zero. However, when one is trying to prove bounds with $C = C_{\vec{\beta}}$ independent of $\vec{\beta}$, it is useful to reduce the dimensionality of the parameter space. Using the translation symmetry we may assume that

$$\beta_1 + \beta_2 + \beta_3 = 0. \tag{1.9}$$

Let $\vec{\beta}_\gamma = (\gamma_1, \gamma_2, -\gamma_1 - \gamma_2)$, where $\gamma = (\gamma_1, \gamma_2)$. By invariance of the measure dt/t under rescaling $\lambda t \mapsto t$, one may assume that $\gamma_1^2 + \gamma_2^2 \in \{1, 0\}$, which gives the following.

Proposition 1.6. *Let $\vec{\beta} \in \mathbb{R}^3$ satisfy* (1.9)*. There exists a nonzero $a \in \mathbb{R}$ such that up to a permutation $\vec{\beta}$ satisfies*

- *$a\vec{\beta} = \vec{\beta}_\gamma$, with $\gamma \in S^1$ such that no two coordinates of $\vec{\beta}_\gamma$ are equal, if and only if $\vec{\beta}$ corresponds to* (1.6)*,*

- *$a\vec{\beta} = \vec{\beta}_\gamma$, with $\gamma \in S^1$ such that exactly two coordinates of $\vec{\beta}_\gamma$ are equal, if and only if $\vec{\beta}$ corresponds to* (1.7)*,*

- *$a\vec{\beta} = \vec{\beta}_{(0,0)}$, if and only if $\vec{\beta}$ corresponds to* (1.8)*.*

Hence, the space of parameters can be identified with $S^1 \cup \{0\}$. This way the degenerate $\vec{\beta}$'s become a finite set on the circle, which corresponds to the Hilbert transform, and the origin, which corresponds to the trivial 0 form, while all the other points on the circle correspond to the nondegenerate case. Note that all transformations that we performed on $\vec{\beta}$ do not affect the constant $C_{\vec{\beta}}$ in (1.5) and hence it is a correct way of case classification for the uniform bounds.

1.3 Main results

In this section we introduce and state the main results of this chapter. We start off along the lines of the previous section with a classification in terms of ranks of $\vec{B}$ and its linear combinations, as well as study the symmetries of the trilinear form. Subsequently, we present the two main theorems, concerning classification and geometry of $\vec{B} \in (\mathbb{R}^{2\times 2})^3$.

1.3.1 Classification in terms of ranks

First, we shall define what we call a degenerate case in two dimensions.

Definition 1.7. *Let $\vec{B} = (B_1, B_2, B_3) \in (\mathbb{R}^{2\times 2})^3$. Set $\vec{B}^T = (B_1^T, B_2^T, B_3^T)$. Define the function*

$$h(\vec{B}) = (r(\vec{B}), r(\vec{B}^T), r(B_2 - B_3), r(B_3 - B_1), r(B_1 - B_2)),$$

where we treat $\vec{B}, \vec{B}^T$ as 6×2 matrices and $r(A)$ denotes the rank of a matrix A.

We call $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ a degenerate triple if $h(\vec{B}) \neq (2, 2, 2, 2, 2)$ and nondegenerate otherwise.

In the following theorem we classify $\vec{B}$ according to the value of $h(\vec{B})$.

Theorem 1.8. *Let $\vec{B} \in (\mathbb{R}^{2\times 2})^3$. Up to a permutation of B_1, B_2, B_3 it satisfies one and only one of the following conditions*

(I) $h(\vec{B}) = (2, 2, 2, 2, 2)$,

(II) $h(\vec{B}) = (2, 2, 2, 2, 1)$,

(III) $h(\vec{B}) = (2, 2, 2, 2, 0)$,

(IV) $h(\vec{B}) = (2, 2, 2, 1, 1)$,

(V) $h(\vec{B}) = (2, 1, 1, 1, 1)$,

(VI) $h(\vec{B}) = (1, 2, 1, 1, 1)$,

(VII) $h(\vec{B}) = (1, 1, 1, 1, 1)$

(VIII) $h(\vec{B}) = (1, 1, 1, 1, 0)$,

(IX) $h(\vec{B}) = (0, 0, 0, 0, 0)$.

The estimate (1.3) is known to hold for $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ in all of the cases above, except for Case (V). In Proposition 1.15 below we show that this case is very closely related to the well known and difficult open problem of boundedness of the triangular Hilbert transform. For Cases (VI) - (VIII), (1.5) follows from one dimensional paraproduct theory, see [CM75], [Mus+] and time-frequency analysis, see [LT97], [LT99], while for Case (III) it follows from the standard two dimensional singular integral theory. In Case (IX), it is easy to verify that $\mathrm{BHF}^K_{\vec{B}}$ equals zero. Concerning the remaining cases, in [DT10] Demeter and Thiele proved that (1.3) holds for $\vec{B}$ corresponding to Case (I) and Case (II). The boundedness for Case (IV) was proven by Vjekoslav Kovač in [Kov12].

1.3.2 Symmetries of the form

Theorem 1.8 gives an overview of triples $\vec{B}$, however, in what follows we wish to reduce the dimensionality of this (12 parameter) space as much as possible, similarly as in one dimension one reduces the initially 3 dimensional parameter space of vectors $\vec{\beta}$ to a one dimensional space. In Proposition 1.10 we study translation and multiplication invariance of the form, which are crucial for further classification. For a function $f\colon \mathbb{R}^2 \to \mathbb{C}$ and a 2×2 matrix A set

$$f^A(x, y) := f(A(x, y)).$$

We define for $\vec{B} = (B_1, B_2, B_2) \in (\mathbb{R}^{2\times 2})^3$ matrix A the left and the right multiplication operations as follows

$$A\vec{B} = (AB_1, AB_2, AB_3), \qquad \vec{B}A = (B_1A, B_2A, B_3A).$$

Remark 1.9. *If we treat $\vec{B}$ as a 2×6 matrix, then the left multiplication is simply the matrix multiplication of $\vec{B}$ from the left by A and the right multiplication is the multiplication of $\vec{B}$ from the right by 6×6 matrix $Id_3 \otimes A$, where Id_3 is the identity 3×3 matrix and $\otimes$ is the tensor product.*

Proposition 1.10. *Let f_1, f_2, f_3 be three Schwartz functions on $\mathbb{R}^2$ and $0 < p_1, p_2, p_3 < \infty$ with $\sum_{j=1}^{3} 1/p_j = 1$. Assume that B_j is a 2×2 real matrix for $j = 1, 2, 3$. Moreover, let A be a 2×2 real matrix. Then*

- *Translation invariance: we have*
$$\mathrm{BHF}^K_{\vec{B}}(f_1, f_2, f_3) = \mathrm{BHF}^K_{\vec{B}-(A,A,A)}(f_1, f_2, f_3). \tag{1.10}$$
- *Left multiplication invariance: if A is nonsingular, then we have*
$$\mathrm{BHF}^K_{\vec{B}}(f_1, f_2, f_3) = |\det A^{-1}|\, \mathrm{BHF}^K_{A\vec{B}}(f_1^{A^{-1}}, f_2^{A^{-1}}, f_3^{A^{-1}}).$$
- *Right multiplication invariance: if A is nonsingular, then we have*
$$\mathrm{BHF}^K_{\vec{B}}(f_1, f_2, f_3) = |\det A|\, \mathrm{BHF}^{K\circ A}_{\vec{B}A}(f_1, f_2, f_3). \tag{1.11}$$

Remark 1.11. *By a change of variables and Proposition 1.10, the translation and the left multiplication of a triple $\vec{B}$ do not change the constant with which* (1.3) *holds. Observe that the right multiplication, when applied with a non-orthogonal matrix, changes both the kernel and its constants in* (1.3)*, hence there is no straightforward invariance of* (1.3) *this case.*

We also have the following invariance of the function h under left and right multiplication.

Proposition 1.12. *Let $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ and let $C, D \in \mathbb{R}^{2\times 2}$ be nonsingular. Then*

$$h(\vec{B}) = h(C\vec{B}D).$$

1.3.3 Classification of the parameter space modulo the symmetries

In the next theorem we give every case in Theorem 1.8 a canonical form. This completes the classification given in [DT10] as well as will simplify the discussion later on. In view of the translation symmetry (1.10), from now on we consider triples of matrices $\vec{B} = (B_1, B_2, B_3)$ satisfying

$$B_1 + B_2 + B_3 = 0. \tag{1.12}$$

Theorem 1.13. *Let $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ satisfy* (1.12)*. There exist two nonsingular $C, D \in \mathbb{R}^{2\times 2}$ such that up to a permutation, $\vec{B}$ satisfies exactly one of the following with some $\lambda, \mu \in \mathbb{R}$*

(1) (a)

$$C\vec{B}D = (\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \begin{pmatrix} -1-\lambda & 0 \\ 0 & -1-\mu \end{pmatrix}),$$

with $\lambda, \mu \neq -2, -1/2, 1$,

(b)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}\lambda & \mu\\ -\mu & \lambda\end{pmatrix}, \begin{pmatrix}-1-\lambda & -\mu\\ \mu & -1-\lambda\end{pmatrix}),$$

with $\mu \neq 0$,

(c)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}, \begin{pmatrix}-1-\lambda & -1\\ 0 & -1-\lambda\end{pmatrix}),$$

with $\lambda \neq -2, -1/2, 1$,

(2) *(a)*

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & \lambda\end{pmatrix}, \begin{pmatrix}-2 & 0\\ 0 & -1-\lambda\end{pmatrix}),$$

with $\lambda \neq -2, -1/2, 1$,

(b)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-2 & -1\\ 0 & -2\end{pmatrix}),$$

(3)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-2 & 0\\ 0 & -2\end{pmatrix}),$$

(4)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}, \begin{pmatrix}-2 & 0\\ 0 & 1\end{pmatrix}),$$

(5)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}-1 & 0\\ -1 & 0\end{pmatrix})$$

(6)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}-2 & 0\\ 0 & 0\end{pmatrix}),$$

(7)

$$C\vec{B}D = (\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}\lambda & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}-1-\lambda & 0\\ 0 & 0\end{pmatrix}),$$

with $\lambda \neq -2, -1/2, 1$,

(8)

$$C\vec{B}D = (\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}),$$

(9)

$$C\vec{B}D = (\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}).$$

We call a triple $\vec{B}$ canonical for Case(n) if it satisfies the condition for Case(n) with $C = D = Id$. Moreover, if $\vec{B}$ corresponds to Case(n), then $h(\vec{B})$ corresponds to Case(R(n)) in Theorem 1.8, where R(n) is the Roman representation of n.

Note that Theorem 1.13 together with Proposition 1.12 implies Theorem 1.8. Case (1) and (2) above have several subcases, all corresponding to Case (I) and Case (II), respectively. In what follows we are not going to differentiate between (1a), (1b), (1c), since the proofs of boundedness of $\mathrm{BHF}_{\vec{B}}$ in these cases [DT10] are identical, i.e. for our problem they are essentially the same. We also remark that the proofs of (2a) and (2b) in [DT10] are similar. It is thus arguable that they could be considered as a single case, but for historical reasons [DT10] we decided to treat them as two subcases.

1.3.4 Geometry of the parameter space

The classification given in Theorem 1.13 effectively distinguishes different cases, however it does not describe how the parameters degenerate. Namely, it requires multiplying the matrices from the right by all nonsingular matrices and, in view of (1.11), it affects the defining constants of the Calderón-Zygmund kernel K. In Theorem 1.14 below we put emphasis on uniformity and classify $\vec{B}$ up to multiplication from the right by orthogonal matrices, which does not affect the constant in (1.2). As we are going to see below, the parameter space has essentially three connected components. The first one corresponds to the forms that act in both coordinates and is homeomorphic to the three dimensional manifold $S^1 \times S^1 \times S^1$. The forms acting in one variable only form the two dimensional manifold homeomorphic to $S^1 \times S^1$ with a submanifold homeomorphic to S^1 corresponding to the bilinear Hilbert transform in one dimension. The trivial 0 form corresponds to $\{0\}$.

From now on we denote by $D_{\alpha,\beta}$ the diagonal matrix with eigenvalues α, β and by R_θ the rotation by θ. We define the parameter space as follows. Let

$$\Omega := S^1 \times S^1 \times [0, 2\pi) \subset \mathbb{R}^5,$$

where we identify the endpoints of the interval, hence treat it as S^1; however, in the following it will be handy to keep the explicitly parametrization in terms of angle. For a $(\beta, \gamma, \theta) \in \Omega$ we represent the triple that corresponds to a point $(\beta, \gamma, \theta) \in \Omega$

$$\vec{B}_{\beta,\gamma,\theta} = (D_{\beta_1,\gamma_1}, D_{\beta_2,\gamma_2} R_\theta, -D_{\beta_1,\gamma_1} - D_{\beta_2,\gamma_2} R_\theta).$$

Let $U \subset \Omega$ be defined as

$$U = \{(\beta, \gamma, \theta) \in \Omega \colon \beta, \gamma \neq (0, \pm 1)\}$$

Note that the closure of U equals Ω. Define the mapping $F\colon U \to \mathbb{R}^2$ given by

$$F(\beta,\gamma,\theta) := \Big(\frac{\beta_2\gamma_2}{\beta_1\gamma_1}, \Big(\frac{\beta_2}{\beta_1} + \frac{\gamma_2}{\gamma_1}\Big)\cos\theta\Big).$$

The role of function F is to encode the eigenvalues of the matrices of the triples $\vec{B}$, which lets us distinguish between different cases appearing in Theorem 1.13. Having defined the set of parameters we can finally state the main theorem of this chapter.

Theorem 1.14. *Let $\vec{B} \in (\mathbb{R}^{2\times 2})^3$ satisfy* (1.12)*. There exist a nonsingular $C \in \mathbb{R}^{2\times 2}$ and an orthogonal $Q \in \mathbb{R}^{2\times 2}$ such that up to a permutation $\vec{B}$ satisfies*

(A)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in \Omega$ such that none of the conditions below is satisfied, if and only if $\vec{B}$ corresponds to Case (1)*.*

(B) (a)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in U$, $F(\beta,\gamma,\theta) = (\lambda, \lambda+1)$, $\lambda \neq -2, -1/2, 1$, if and only if $\vec{B}$ corresponds to Case (2a)*.*

(b)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in U$, $F(\beta,\gamma,\theta) = (1,2)$ and $\theta \neq 0, \pi$, if and only if $\vec{B}$ corresponds to Case (2b)*.*

(C)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in U$, $F(\beta,\gamma,\theta) = (1,2)$ and $\theta = 0, \pi$, if and only if $\vec{B}$ corresponds to Case (3)*.*

(D)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in U$, $F(\beta,\gamma,\theta) = (-2,-1)$, if and only if $\vec{B}$ corresponds to Case (4)*,*

(E)

$$C\vec{B}Q = \vec{B}_{\beta,\gamma,\theta},$$

with $(\beta,\gamma,\theta) \in \Omega$, $\beta = (1,0)$ and $\gamma = (0,1)$ and $\theta = \pi/2, 3\pi/2$, if and only if $\vec{B}$ corresponds to Case (5)*,*

(F)

$$C\vec{B}Q = \vec{B}_{\beta,(0,0),0},$$

with $\beta \in S^1$*,* $\beta_1 = \beta_2$*, if and only if* $\vec{B}$ *corresponds to Case* (6)*.*

(G)

$$C\vec{B}Q = \vec{B}_{\beta,(0,0),0},$$

with $\beta \in S^1$*,* $\beta_3 = 0$ *and* $\beta_1 \neq -2\beta_2, -1/2\beta_2, \beta_2$*, if and only if* $\vec{B}$ *corresponds to Case* (7)*.*

(H)

$$C\vec{B}Q = \vec{B}_{\beta,(0,0),\theta},$$

with $\beta \in S^1$ *and* $\theta \neq 0, \pi$*, if and only if* $\vec{B}$ *corresponds to Case* (8)*.*

(I)

$$C\vec{B}Q = \vec{B}_{(0,0),(0,0),0},$$

if and only if $\vec{B}$ *corresponds to Case* (9)*.*

1.3.5 Uniform bounds conjecture

In view of the classification given in Theorem 1.14 we state a conjecture that implies the uniform bounds in all $\vec{B} \in (\mathbb{R}^{2\times 2})^3$.

Conjecture. *Let* $\mathcal{K}$ *be a family of Calderón-Zygmund kernels* K*, such that* (1.2) *holds. There exists a constant* $C_{p_1,p_2,p_3} < \infty$*, such that for all* $f_1, f_2, f_3 \in \mathcal{S}(\mathbb{R}^2)$ *and* $2 < p_1, p_2, p_3 < \infty$ *with* $\sum_{j=1}^{3} 1/p_j = 1$

$$\mathrm{BHF}^K_{\vec{B}_{\beta,\gamma,\theta}}(f_1, f_2, f_3) \leq C_{p_1,p_2,p_3} \prod_{i=1}^{3} \|f_i\|_{L^{p_i}(\mathbb{R}^2)},$$

uniformly in $(\beta, \gamma, \theta) \in \Omega$ *and* $K \in \mathcal{K}$*.*

In view of Proposition 1.19 below, in order to obtain uniform bounds for all $(\beta, \gamma, \theta) \in \Omega$, it suffices to prove the uniform bounds for any dense subset of Ω. Correspondingly, in order prove (1.3) for some $B_{\beta,\gamma,\theta}$, it is enough to prove the uniform bounds for $\{B_{\beta_n,\gamma_n,\theta_n}\}$ for a sequence $(\beta_n, \gamma_n, \theta_n) \to (\beta, \gamma, \theta)$. We discuss a number of uniform questions related to Conjecture 1.3.5 in the last section of this chapter.

1.3.6 Connection to the triangular Hilbert transform

In this subsection we shall see how Case (5) relates to the triangular Hilbert transform. Let f_1, f_2, f_3 be three Schwartz function on $\mathbb{R}^2$. Let $\vec{B}$ be the canonical triple in Case (5). The triangular Hilbert transform is defined as

$$\Lambda_\triangle(f_1, f_2, f_3) := \int_{\mathbb{R}^3} \prod_{j=1}^{3} f_j((x, y) + B_j(s, 0)) \, \frac{ds}{s} \, dx \, dy$$

In [KTZK15] it was shown that, if one assumes that $\Lambda_\triangle$ is L^p bounded, then (1.3) holds for odd and homogeneous kernels K of degree 2 uniformly in $\vec{B}$. Moreover, in the same paper, the inequality (1.3) was proven for a dyadic model of $\Lambda_\triangle$, under an additional assumption that one of the three functions is of a special form.

In the following we show that Case (5) above corresponds to the triangular Hilbert transform, in the sense that the triangular Hilbert transform can be recovered choosing an appropriate kernel. Specifically, we have the following proposition.

Proposition 1.15. *Let $\vec{B}$ the canonical triple for Case (5). There exists a Calderón-Zygmund kernel K such that for all triples f_1, f_2, f_3 of Schwartz functions on $\mathbb{R}^2$ we have*

$$\mathrm{BHF}^K_{\vec{B}}(f_1, f_2, f_3) = \Lambda_\triangle(f_1, f_2.f_3).$$

1.4 Proofs

Note that Theorem 1.8 follows directly from Proposition 1.12 and Theorem 1.13, which we prove later on in this section.

Proof of Proposition 1.10. Translation invariance:

$$\mathrm{BHF}^K_{\vec{B}-(A,A,A)} = \mathrm{BHF}^K_{\vec{B}}$$

follows from a simple change of variables in (x, y). Thus the inequality does not change.

Left multiplication invariance: rewrite $\mathrm{BHF}^K_{\vec{B}}(f_1, f_2, f_3)$ as follows

$$\int_{\mathbb{R}^4} \prod_{j=1}^{3} f_j^{A^{-1}}(A(x,y) + AB_j(s,t))K(s,t)\, dx\, dy\, ds\, dt.$$

Changing variables $A(x, y) \mapsto (x, y)$ this is equal to

$$\begin{aligned} &|\det A^{-1}| \int_{\mathbb{R}^4} \prod_{j=1}^{3} f_j^{A^{-1}}((x,y) + AB_j(s,t))K(s,t)\, dx\, dy\, ds\, dt \\ &= |\det A^{-1}|\, \mathrm{BHF}^K_{A\vec{B}}(f_1^{A^{-1}}, f_2^{A^{-1}}, f_3^{A^{-1}}). \end{aligned}$$

Right multiplication invariance: this follows by the change of variables $(s, t) \mapsto A(s, t)$. Note that $K(A(s, t))$ remains a Calderón-Zygmund kernel (possibly with different constants). □

Proof of Proposition 1.12. Clearly multiplying any of $B_2 - B_3$, $B_3 - B_1$, $B_1 - B_2$ from the left and from the right by a nonsingular matrix does not change their ranks. Thus, we only have to prove $r(C\vec{B}D) = r(\vec{B})$ and $r(D^T\vec{B}^TC^T) = r(\vec{B}^T)$, for nonsingular $C, D \in \mathbb{R}^{2\times2}$. By symmetry, it suffices to show that $r(C\vec{B}) = r(\vec{B})$ and $r(\vec{B}D) = r(\vec{B})$. Using Remark 1.9, the first identity follows, because C is a 2×2 matrix of rank 2 and the second identity follows, because $Id_3 \otimes D$ is a 6×6 matrix of rank 6. □

Proof of Theorem 1.13.

1. First we show that for any triple $\vec{B}$ there exist nonsingular C, D such that $C\vec{B}D$ corresponds to one of the cases.

Assume that one of B_1, B_2, B_3 is nonsingular. Without loss of generality we may assume that B_1 is nonsingular. Multiplying from the left by B_1^{-1} we may assume that $\vec{B} = (Id, B_2, B_3)$.

Moreover, we may multiply from the left by A and right by A^{-1}, for an appropriate, so that AB_2A^{-1} is of canonical Jordan form. Hence there exist two nonsingular matrices C, D such that up to a permutation $C\vec{B}D$ belongs to at least one of Cases (1) - (4).

Now, assume that all B_1, B_2, B_3 are singular. If they are all zero, then $\vec{B}$ corresponds to Case (9). Otherwise, without loss of generality B_1 and B_2 have rank 1 (by (1.12) it is not possible that only one of the coordinates of $\vec{B}$ is nonzero). Let (v, w) denote the 2×2 matrix with vectors v, w as columns. It must hold that $B_1 = (\lambda_1 v_1, \mu_1 v_1)$, $B_2 = (\lambda_2 v_2, \mu_2 v_2)$, where $\lambda_j, \mu_j \in \mathbb{R}$ for $j = 1, 2$ and at least one number of the pair λ_1, λ_2 and of the pair μ_1, μ_2 is nonzero. Now, there are two possibilities, either v_1, v_2 are linearly independent or not. If they are, then multiplying from the left we may assume that $B_1 = (\lambda_1 e_1, \mu_1 e_1)$, $B_2 = (\lambda_2 e_2, \mu_2 e_2)$, where e_1, e_2 are the standard basis vectors. Moreover, multiplying from the right by an appropriate matrix we may assume that the first row of B_1 is $(1, 0)$, second row is zero and the first row of B_2 is zero and the second row is $(1, 0)$ (it cannot be $(0, 1)$ because $r(B_3) \leq 1$). This correspond to Case (5). If v_1, v_2 are linearly dependent, then using similar arguments one can show that $\vec{B}$ corresponds to one of Case (6)-(8).

2. Now we prove that there is exactly one case that $\vec{B}$ corresponds to. First, note that it follows from Proposition 1.12 that $h(\vec{B})$ is invariant under multiplying $\vec{B}$ from the left and right by a nonsingular matrix. Thus h differentiates between all of the cases except maybe (1a), (1b) (1c) and the pair (2a), (2b), whose values of h coincide. In order prove the statement for these assume that we have two triples $\vec{A}$, $\vec{B}$, each of which contains at least one nonsingular matrix and there exist two nonsingular matrices C, D such that

$$\{CA_1D, CA_2D, CA_3D\} = \{B_1, B_2, B_3\}.$$

Without loss of generality, we may assume that $\vec{A}$, $\vec{B}$ are both canonical triples and $A_1 = B_1 = Id$. If $C = D^{-1}$, then $\{A_2, A_3\}$ and $\{B_2, B_3\}$ must have the same Jordan form, which implies that $\{Id, A_2, A_3\}$ must correspond to the same Case as $\{Id, B_2, B_3\}$. If $C \neq D^{-1}$, then, without loss of generality assume that $CD = B_2$ and $CA_2D = Id$. This implies that $D^{-1}A_2D = B_2^{-1}$. In other words the Jordan canonical form of A_2 is B_2^{-1}. It is easy to verify that that is not possible if $\vec{A}$, $\vec{B}$ belong to different subcases of Case (1), or similarly, if $\vec{A}$ belongs Case (2a) and $\vec{B}$ belongs to Case (2b). □

It follows from the previous theorem that the "more degenerate" cases (5) - (8) have some structural properties that can be easily explained in terms of column vectors. The following corollary will be helpful in the proof of 1.14.

Corollary 1.16. *Let (v, w) denote the 2×2 matrix with vectors v, w as columns.*

- $B_1 = (\lambda v_1, \mu v_1)$, $B_2 = (\lambda v_2, \mu v_2)$ *for two linearly independent vectors v_1, v_2 and a nonzero vector (λ, μ) if and only if $\vec{B}$ corresponds to Case* (5).
- $B_1 = (v, v)$, $B_2 = \lambda(v, v)$ *for a nonzero vector v and $\lambda = -2, -1/2, 1$ if and only if $\vec{B}$ corresponds to Case* (6).
- $B_1 = (v, v)$, $B_2 = \lambda(v, v)$ *for a nonzero vector v and $\lambda \neq -2, -1/2, 1$ if and only if $\vec{B}$ corresponds to Case* (7).
- $B_1 = (\lambda_1 v, \mu_1 v)$, $B_2 = (\lambda_2 v, \mu_2 v)$ *for a nonzero vector v and two linearly independent vectors (λ_1, μ_1), (λ_2, μ_2) if and only if $\vec{B}$ corresponds to Case* (8).

Proof. Follows from the classification given in Theorem 1.13. □

Proof of Theorem 1.14.

($\Longleftarrow$):

1. First assume that $\vec{B}$ belongs to one of Cases (1) - (4) in Theorem 1.13, i.e. without loss of generality we may assume that B_1 is nonsingular.

First step: replace B_1, B_2 by $B_1^{-1}B_1$, $B_1^{-1}B_2$. This way we may reduce to $\vec{B} = (Id, B, -Id - B)$.

Second step: recall that there exists the polar decomposition of any real matrix B, meaning that

$$B = P\widetilde{Q}$$

where $\widetilde{Q}$ is a real orthogonal matrix and P is symmetric, positive semi-definite. We can diagonalize P conjugating by an orthogonal real matrix V

$$D = VPV^T.$$

Thus, conjugating $B = P\widetilde{Q}$ by V

$$VBV^T = VPV^TV\widetilde{Q}V^T = DQ,$$

where $Q = V\widetilde{Q}V^T$ is an orthogonal matrix. Hence conjugating by orthogonal real matrices we may reduce $\vec{B}$ to the form $(Id, DQ, -Id - DQ)$.

Third step: let λ and μ be the (real) eigenvalues of D. Multiplying from the left by

$$\begin{pmatrix} \frac{1}{\sqrt{1+\lambda^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+\mu^2}} \end{pmatrix}$$

we reduce to $(D_{\beta_1,\gamma_1}, D_{\beta_2,\gamma_2}Q, -D_{\beta_1,\gamma_1} - D_{\beta_2,\gamma_2}Q)$, with $\beta, \gamma \in S^1$. By the characterization of 2×2 orthogonal matrices Q is either a rotation or a reflection. Note that $D_{-1,1}$ times a reflection is a rotation. Hence, we may always replace a reflection with a rotation and obtain the desired form of the triple $\vec{B}$.

2. Assume that $\vec{B}$ belongs to Case (5) in Theorem 1.13. By Corollary 1.16 a triple that belongs to this case is of the form

$$B_1 = (\lambda v_1, \mu v_1), \qquad B_2 = (\lambda v_2, \mu v_2), \qquad B_3 = (-\lambda(v_1 + v_2), -\mu(v_1 + v_2)),$$

for two linearly independent vectors v_1, v_2 and a nonzero vector (λ, μ). Multiplying from the left by the matrix that maps $v_1 \mapsto (1, 0)$ and $v_2 \mapsto (0, 1)$ we reduce this triple to

$$\begin{pmatrix} \lambda & \mu \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 \\ \lambda & \mu \end{pmatrix}, \qquad \begin{pmatrix} -\lambda & -\mu \\ -\lambda & -\mu \end{pmatrix}$$

Multiplying once more from the left we may normalize $\|(\lambda, \mu)\| = 1$ and multiplying from the left by the transpose of the rotation that maps $(\lambda, \mu) \mapsto (1, 0)$ we further transform the triple to

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}.$$

3. Assume that $\vec{B}$ belongs to one of Cases (6)-(8) in Theorem 1.13. First, suppose that $\vec{B}$ corresponds to Case (8). Then by Lemma 1.16 $B_1 = (\lambda_1 v, \mu_1 v)$ and $B_2 = (\lambda_2 v, \mu_2 v)$ for a

nonzero vector v and two linearly independent vectors (λ_1, μ_1), (λ_2, μ_2). Multiplying from the left by the matrix that maps $v \mapsto (1,0)$ we may assume that

$$B_1 = \begin{pmatrix} \lambda_1 & \mu_1 \\ 0 & 0 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \lambda_2 & \mu_2 \\ 0 & 0 \end{pmatrix},$$

Multiplying $\vec{B}$ from the right by the transpose of the rotation that maps $(\lambda_1, \mu_1) \mapsto (\alpha, 0)$ (for some $\alpha \geq 0$ depending on the length of (λ_1, μ_1)) we may reduce $\vec{B}$ to $(D_{\alpha,0}, D_{\beta,0}Q, -D_{\alpha,0} - D_{\beta,0}Q)$, for some $\alpha, \beta \geq 0$. Finally, multiplying from the left we can normalize so that the triple has the desired form. Using similar arguments one can prove the desired form of triples $\vec{B}$ corresponding to Cases (6), (7).

$(\implies)$:

1. Consider $\vec{B}_{\beta,\gamma,\theta}$ that satisfies one of the conditions (A) - (D). Without loss of generality we may assume that B_1 is nonsingular, i.e. $\beta_1, \gamma_1 \neq 0$. Multiplying $\vec{B}_{\beta,\gamma,\theta}$ from the left by $D_{\beta_1^{-1},\gamma_1^{-1}}$ it becomes $(Id, B, -Id - B)$, where

$$B = \begin{pmatrix} \frac{\beta_2}{\beta_1}\cos\theta & -\frac{\beta_2}{\beta_1}\sin\theta \\ \frac{\gamma_2}{\gamma_1}\sin\theta & \frac{\gamma_2}{\gamma_1}\cos\theta \end{pmatrix}.$$

Observe that the eigenvalues λ_1, λ_2 of B satisfy

$$F(\beta, \gamma, \theta) = (\lambda_1\lambda_2, \lambda_1 + \lambda_2).$$

We shall need the following lemma.

Lemma 1.17. *Assume that $\vec{B} = (Id, B, -Id - B)$. $\vec{B}$ corresponds to*

- *Case* (2a) *if and only if B has two eigenvalues and exactly one of them in* $\{-2, -\frac{1}{2}, 1\}$
- *Case* (2b) *if and only if B is similar to a Jordan block with an eigenvalue in* $\{-2, -\frac{1}{2}, 1\}$
- *Case* (3) *if and only if B is diagonalizable with two equal eigenvalues in* $\{-2, -\frac{1}{2}, 1\}$
- *Case* (4) *if and only if B has two different eigenvalues in* $\{-2, -\frac{1}{2}, 1\}$

Otherwise, $\vec{B}$ corresponds to Case (1).

Proof. Follows from changing the basis so that B has the Jordan canonical form and Theorem 1.13. □

Then, the proof follows from analysis of the product and the sum of possible pairs of eigenvalues for B_2 given the value of $F(\beta, \gamma, \theta)$ and using Lemma 1.17, assuming that $B_1 = Id$.

2. Assume that $\vec{B}_{\beta,\gamma,\theta}$ corresponds to Case (E). Note that in view of Corollary 1.16 this triple must correspond to Case (5).

3. Assume that $\vec{B}_{\beta,\gamma,\theta}$ corresponds to Case (H), i.e. $\beta \in S^1$, $\gamma = (0,0)$ and $\theta \neq 0, \pi$, then by Corollary 1.16 it corresponds to (8). Similarly, the desired implication follows for triples $\vec{B}_{\beta,\gamma,\theta}$ corresponding to Case (F), (G).

This finishes the proof of the theorem. □

Proof of Proposition 1.15. Let φ, ψ be smooth, compactly support functions on $\mathbb{R}$. Additionally assume that φ has mean zero, ψ has mean one. Define for $k \in \mathbb{Z}$, $\varphi_k(u) := \frac{1}{2^k}\varphi(\frac{u}{2^k})$, $\psi_k(u) := \frac{1}{2^k}\psi(\frac{u}{2^k})$ and assume that for $s \neq 0$

$$\frac{1}{s} = \sum_{k\in\mathbb{Z}} \varphi_k(s).$$

Define K as follows

$$K(s,t) = \sum_{k\in\mathbb{Z}} \varphi_k(s)\psi_k(t).$$

One can easily check that K is a valid Calderón-Zygmund kernel. Moreover, we have

$$\begin{aligned}
&\mathrm{BHF}_{\vec{B}}(f_1, f_2, f_3)\\
&= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^3} \int_{\mathbb{R}} \prod_{j=1}^{3} f_j((x,y) + B_j(s,0))\varphi_k(s)\psi_k(t)\, dt\, ds\, dx\, dy\\
&= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^3} \int_{\mathbb{R}} \prod_{j=1}^{3} f_j((x,y) + B_j(s,0))\varphi_k(s)\, ds\, dx\, dy\\
&= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^3} \int_{\mathbb{R}} \prod_{j=1}^{3} f_j((x,y) + B_j(s,0))\, \frac{ds}{s}\, dx\, dy = \Lambda_{\triangle}(f_1, f_2, f_3).
\end{aligned}$$

□

1.5 Closing remarks

In Theorem 1.14, we described the parameter space of (1.1) as having three connected components, homeomorphic to $(S^1)^3$, $(S^1)^2$ (with a submanifold homemorphic to S^1, which is corresponding to the one dimensional bilinear Hilbert transform) and a single point. Moreover, we distinguished several subsets of $(S^1)^3$, given by preimages of certain values of the function F, corresponding to different operators in harmonic analysis. In this section, we give a summary of the uniform questions on the manifold $(S^1)^3$. If there exists a sequence of points corresponding to Case (X), convergent to a point corresponding to Case (Y), we say that Case (Y) can be approached by Case (X).

Proposition 1.18. *We have that:*

- *Case* (Ba) *can be approached by Case* (A)
- *Case* (Bb) *can be approached by Case* (Ba) *and Case* (A),
- *Case* (C) *can be approached by Case* (Bb), *Case* (Ba) *and Case* (A).
- *Case* (D) *can be approached by Case* (Ba) *and Case* (A),
- *Case* (E) *can be approached by Case* (D), *Case* (Ba), *Case* (Bb) *and Case* (A),

Proof. Case (A): it can approach all other cases simply by density: the triples corresponding (A) are dense in $S^1 \times S^1 \times [0, 2\pi)$.

Case (Ba): it can approach all cases except for (A), which it cannot approach because of the value of function F. For the other cases, except for Case (E), it can be seen choosing a convergent sequence of parameters $(\beta_n, \gamma_n, \theta_n) \in U$ such that

$$\lim_{n\to\infty} F(\beta_n, \gamma_n, \theta_n) = (\lambda_n, \lambda_n + 1),$$

where $\lambda_n \notin \{-2, -1/2, 1\}$ with the limit in $\{-2, -1/2, 1\}$. Note, however, that since the function F is not defined for Case (E), it requires a different argument. In this situation it is enough to notice that there exists a sequence of parameters $(\beta_n, \gamma_n, \theta_n) \in U$ with

$$\begin{aligned}&(\beta_n, \gamma_n, \theta_n) \to ((1,0),(0,1),\pi/2),\\ &F(\beta_n, \gamma_n, \theta_n) = (\lambda, \lambda+1), \qquad \lambda \notin \{-2,-1/2,1\}.\end{aligned}$$

Case (Bb): it can be seen that it can approach Case (C) by choosing a convergent sequence $(\beta_n, \gamma_n, \theta_n) \in U$ with $F(\beta_n, \gamma_n, \theta_n) = (1,2)$ and $\theta_n \to 0$. Moreover, it can approach Case (E) arguing like in the previous paragraph.

Cases (C) and (E) correspond to a finite set in $S^1 \times S^1 \in [0, 2\pi)$ and hence it cannot approach any other case on the manifold.

Case (D): it can approach (E) using similar argument as before. One can see that it cannot approach any other case on $S^1 \times S^1 \times [0, 2\pi)$ by investigating the values of the function F it corresponds to. □

At the end of this chapter we prove a continuity result for the form BHF with respect to triples $\vec{B}$. Precisely, we have the following.

Proposition 1.19. *Let $0 < p_1, p_2, p_3 < \infty$ with $\sum_{j=1}^{3} 1/p_j = 1$. Let BHF^ε denote the truncation of the integral defining BHF to $\varepsilon \le |(t,s)| \le 1/\varepsilon$. Suppose that $\vec{B}_n \to \vec{B}$ and there exists a constant $C > 0$ such that for any $\varepsilon > 0$, $n \in \mathbb{N}$ and any triple of Schwartz functions f_1, f_2, f_3 on $\mathbb{R}^2$*

$$\mathrm{BHF}^\varepsilon_{\vec{B}_n}(f_1, f_2, f_3) \le C \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}.$$

Then for any triple of Schwartz function f_1, f_2, f_3 and any $\varepsilon > 0$

$$\mathrm{BHF}^\varepsilon_{\vec{B}}(f_1, f_2, f_3) \le C \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}.$$

Note that in view of Proposition 1.19, for Conjecture 1.3.5 it suffices to prove boundedness in a dense set of parameters.

Proof of Proposition 1.19. Let us fix a triple of Schwartz functions f_1, f_2, f_3. Since $\varepsilon > 0$, for any $\delta > 0$ and $n \ge N_\delta$ large enough we have

$$\begin{aligned}&|\mathrm{BHF}^\varepsilon_{\vec{B}}(f_1, f_2, f_3)|\\ &\le |\mathrm{BHF}^\varepsilon_{\vec{B}}(f_1, f_2, f_3) - \mathrm{BHF}^\varepsilon_{\vec{B}_n}(f_1, f_2, f_3)| + |\mathrm{BHF}^\varepsilon_{\vec{B}_n}(f_1, f_2, f_3)|\\ &\le \delta + |\mathrm{BHF}^\varepsilon_{\vec{B}_n}(f_1, f_2, f_3)|\\ &\le \delta + C \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}.\end{aligned}$$

This finishes the proof. □

Chapter 2

Uniform bounds for the bilinear Hilbert transform in local L^1

2.1 Introduction

In this chapter we present a joint work with Gennady Uraltsev which will be a part of a publication. Thus, we start with giving a self-contained introduction to the problem, which may be somewhat repetitive when compared with the introduction of this thesis.

The trilinear form associated through duality to the bilinear Hilbert transform is given by

$$\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3) := \int_{\mathbb{R}} \int_{\mathbb{R}} \prod_{j=1}^{3} f_j(x - \beta_j t)\, dx\, \frac{dt}{t}, \tag{2.1}$$

where f_1, f_2, f_3 are Schwartz functions on the real line and $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$ is a unit vector with pairwise distinct coordinates perpendicular to $\vec{1} := (1, 1, 1)$. One is interested in proving the a priori L^p bounds for this form

$$|\,\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3)| \le C_{p_1,p_2,p_3,\beta} \|f_1\|_{L^{p_1}(\mathbb{R})} \|f_2\|_{L^{p_2}(\mathbb{R})} \|f_3\|_{L^{p_3}(\mathbb{R})}. \tag{2.2}$$

By scaling, the exponents in (2.2) should satisfy $1/p_1 + 1/p_2 + 1/p_3 = 1$, which we will assume throughout.

In [LT97] Lacey and Thiele proved first estimates of the type (2.2). They showed that (2.2) holds in the range $2 < p_1, p_2, p_3 < \infty$, with a constant dependent only on p_1, p_2, p_3 and $\vec{\beta}$. This corresponds to the open triangle c in Figure 2.1. The range of exponents for the inequality (2.2) was extended in [LT99] to the range that coincides with the convex hull of the open triangles a_1, a_2, a_3 in Figure 2.1. The bounds outside of the range $1 < p_1, p_2, p_3 < \infty$ are in the sense of restricted weak type, we refer to [Thi06] for details of restricted weak type interpolation. Inspired by the works of Carleson [Car66] and Fefferman [Fef73], the main tool that was used by the authors of [LT97], [LT99] was time-frequency analysis, i.e. techniques based on localizing functions f_1, f_2, f_3 both in space and frequency. As noted in [Dem+08], the time-frequency approach shares some similarities with Bourgain's argument in [Bou88] in the context of convergence of bilinear ergodic averages.

When two of the components of $\vec{\beta}$ are equal, the trilinear form $\mathrm{BHF}_{\vec{\beta}}$ becomes a composition of the Hilbert transform and the pointwise product. More precisely, up to a symmetry it equals

$$\int_{\mathbb{R}} H f_1(x)\, f_2(x)\, f_3(x)\, dx, \tag{2.3}$$

which immediately implies boundedness for $1 < p_1, p_2, p_3 < \infty$ by Hölder's inequality and boundedness of the Hilbert transform. While in [LT97], [LT99] the dependence of the constant in (2.2) is not explicitly stated in terms of $\vec{\beta}$, one can show it grows linearly in $\min_{i \neq j} |\beta_i - \beta_j|^{-1}$. This raised the question asked in [LT99]: can one prove that

$$|\operatorname{BHF}_{\vec{\beta}}(f_1, f_2, f_3)| \leq C_{p_1,p_2,p_3} \|f_1\|_{L^{p_1}(\mathbb{R})} \|f_2\|_{L^{p_2}(\mathbb{R})} \|f_3\|_{L^{p_3}(\mathbb{R})} \tag{2.4}$$

holds with a constant C_{p_1,p_2,p_3} independent of $\vec{\beta}$ and if so, in what range of exponents? The form is symmetric under permutations of $(\beta_1, \beta_2, \beta_3)$, hence from now on let us assume that $\vec{\beta}$ is in the vicinity of the degenerate case $\beta_2 = \beta_3$. Since in the degenerate case the trilinear form equals (2.3) and the classical Hilbert transform is not L^∞ bounded, uniform bounds cannot hold for $\alpha_1 \leq 0$. This corresponds in Figure 2.1 to the region below the line spanned by $(0,0,1)$, $(0,1,0)$. Moreover, the maximal range for which the parameter dependent bounds (2.2) are known, is the convex hull of the open triangles a_1, a_2, a_3. The intersection of the two regions is the convex hull of the open triangles b_3, b_2, a_3 and a_2.

A lot of progress has been made in the direction of the uniform bounds. The inequality (2.4) was proven with a constant independent of $\vec{\beta}$ in several papers: Thiele [Thi02a] proved a weak type inequality at the two upper corners of the triangle c in Figure 2.1, Grafakos and Li [GL04] showed the inequality in the open triangle c, and Li [Li06] proved the bounds in the open triangles a_1, a_2. By interpolation one obtains (2.4) in the range corresponding the convex hull of the open triangles a_2, a_3 and c, see Figure 2.1. What however was not known up to date, is whether the uniform bounds hold in the vicinity of $(1/p_1, 1/p_2, 1/p_3) = (0,0,1)$, $(0,1,0)$. The purpose of this article is to resolve precisely this issue. Here is our main result.

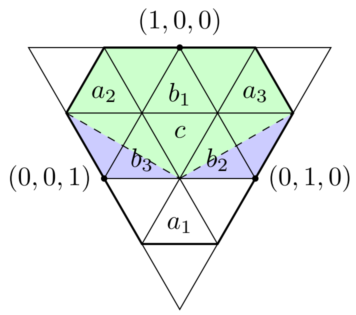

Figure 2.1: Range of exponents $(\alpha_1, \alpha_2, \alpha_3) = (1/p_1, 1/p_2, 1/p_3)$ with $\sum_{j=1}^3 \alpha_j = 1$. The uniform bounds were previously known to hold in the convex hull of the open triangles a_2, a_3 and c. Theorem 2.1 implies the uniform bounds in the convex of the open triangles a_2, a_3, b_2 and b_3.

Theorem 2.1. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $\vec{\beta}$ and all triples of Schwartz functions f_1, f_2, f_3 the inequality* (2.4) *holds.*

Interpolated with the result of Li [Li06] this extends the uniform inequality (2.4) to the exponents corresponding to the convex hull of the open triangles a_2, a_3, b_2 and b_3, see Figure 2.1. We remark that Oberlin and Thiele [OT11] proved a counterpart of the uniform inequality (2.4) for a Walsh model of the bilinear Hilbert transform in the same range.

It is stated in [LT97], [Thi02a] that Calderón considered the bilinear Hilbert transform in the 1960's in the context of the Calderón first commutator. This operator is given by

$$C_1(f)(x) = \int \frac{A(x)-A(y)}{(x-y)^2} f(y)\,dy,$$

where A is a Lipschitz function. It is a well known result of Calderón [Cal65] that C_1 is L^p bounded for $1 < p < \infty$. As said in [Thi02a], one of the initially unsuccessful approaches, which motivated the study of the bilinear Hilbert transform, was to rewrite it formally using the mean value theorem as

$$C_1(f)(x) = \int_0^1 \int f(y) A'(y+\alpha(x-y)) \frac{1}{x-y}\,dy\,d\alpha.$$

By duality, in order to prove L^p boundedness of C_1, it suffices to show that the form $\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, A')$ is bounded for $p_3 = \infty$ and $1 < p_1, p_2 < \infty$ with $1/p_1 + 1/p_2 = 1$, and a constant independent of $\vec{\beta}$. Therefore Theorem 2.1 together with [Li06] gives an alternative proof of Calderón's result. We record that that yet another proof of this theorem was given by Muscalu [Mus14]. Let us also remark that recently in [Gre+16], the uniform bounds found an application in the context of a trilinear form acting on functions on $\mathbb{R}^2$, which possesses the full $GL_2(\mathbb{R})$ dilation symmetry. The boundedness of this form is reduced to a fiber-wise application of the result from [GL04], see [Gre+16] for details.

On the technical side, we refine the outer measure approach gradually developed in the sequence of papers [DT15], [DPO15], [Ura16]. In the paper [DT15], Do and Thiele reformulated the problem of boundedness of the bilinear Hilbert transform into proving an outer Hölder inequality on the upper half space $\mathbb{R}^3_+ := \mathbb{R}\times\mathbb{R}\times\mathbb{R}_+$ and an embedding theorem. Their methods work in the range $2 < p < \infty$. The embedding was later extended to the range $1 < p < \infty$ by Di Plinio and Ou in [DPO15] and reformulated in [Ura16] as an iterated embedding theorem. We shall follow the latter approach. In key Theorem 2.2 below we prove an inequality that can be viewed as a trilinear outer L^p estimate for the wave packet decomposition of $\mathrm{BHF}_{\vec{\beta}}$ uniform in the parameter $\vec{\beta}$. We record that while in [DT15], [DPO15], [Ura16] the main difficulty are embedding theorems, in this chapter we are concerned with the multilinear inequality. Having it proven, we can apply off the shelf, though difficult, embedding theorem shown in [DPO15].

It is well known that the trilinear form $\mathrm{BHF}_{\vec{\beta}}$ is symmetric under translations, modulations and dilations. Following [DT15], we parametrize these actions by (y, η, t) in the upper half space $\mathbb{R}^3_+$. Let Φ be the class of Schwartz functions whose Fourier transform is supported in $(-1,1)$ and such that for a fixed large natural number N and a constant $A > 0$ satisfy

$$\sup_{n,m\le N} \sup_{x\in\mathbb{R}} (1+|x|)^n |\varphi^{(m)}(x)| \le A < \infty$$

Moreover, let $\Phi^* \subset \Phi$ be the class of Schwartz functions whose Fourier transform is supported in $(-2^{-8}b, 2^{-8}b)$ for some $0 < b < 2^{-8}$, which is fixed throughout this chapter. For $\varphi \in \Phi$ set $\varphi_{\eta,t}(x) := \frac{1}{t} e^{i\eta x} \varphi(\frac{x}{t})$ and

$$F^{\varphi}(f)(y,\eta,t) := f * \varphi_{\eta,t}(y), \tag{2.5}$$

$$F(f)(y,\eta,t) := \sup_{\varphi\in\Phi} |F^{\varphi}(f)(y,\eta,t)|, \tag{2.6}$$

$$F^*(f)(y,\eta,t) := \sup_{\varphi\in\Phi^*} |F^{\varphi}(f)(y,\eta,t)|,$$

$$\boldsymbol{F}(f)(y,\eta,t) := (F(f)(y,\eta,t), F^*(f)(y,\eta,t)) \tag{2.7}$$

where $(y,\eta,t) \in \mathbb{R}^3_+$. In the vein of [DT15] we rewrite the problem of boundedness of the bilinear Hilbert transform as a problem for a trilinear integral over $\mathbb{R}^3_+$

$$\Lambda_{\vec{\beta}}(F^\varphi(f_1), F^\varphi(f_2), F^\varphi(f_3)) := \int_{\mathbb{R}^3_+} \prod_{j=1}^{3} F^\varphi(f_j)(y, \alpha_j\eta + \delta\beta_j t^{-1}, |\alpha_j|^{-1}t)\, dt\, d\eta\, dy, \tag{2.8}$$

where $\vec{\alpha} := (\alpha_1,\alpha_2,\alpha_3) \in \mathbb{R}^3$ is the unit vector perpendicular to $\vec{\beta}$ and $\vec{1}$, and $\delta := \min(|\alpha_1|,|\alpha_2|,|\alpha_3|)$. It was shown in [Ura16] that the results of [DT15] imply that for $\varphi \in \Phi^*$ the following inequality holds

$$|\Lambda_{\vec{\beta}}(F^\varphi(f_1), F^\varphi(f_2), F^\varphi(f_3))| \le C_{p_1,p_2,p_3,\vec{\beta}} \prod_{j=1}^{3} \|F^\varphi(f_j)\|_{L^{p_j}\L^{q_j}(S)} \tag{2.9}$$

for $\sum_{j=1}^3 1/p_j = 1$ with $1 < p_j < \infty$ and $\sum_{j=1}^3 1/q_j = 1$ with $2 < q_j < \infty$. On the right hand side of (2.9) are iterated outer L^p norms developed in [Ura16] that we define precisely in Section 2.3. Following [Ura16] we write the embedding theorem of [DPO15] as

$$\|F^\varphi(f)\|_{L^p\L^q(S)} \le C_p\|f\|_{L^p(\mathbb{R})} \qquad \text{for } p > 1 \text{ and } q > \max(p',2) \tag{2.10}$$

Coupled with (2.9) it in particular implies L^p boundedness of the bilinear Hilbert transform (2.1) in the local L^1. In this chapter we prove a counterpart of (2.9) with a constant uniform in the parameter $\vec{\beta}$. Here is our result.

Theorem 2.2. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1,p_2,p_3 < \infty$ and $1/q_1 + 1/q_2 + 1/q_3 > 1$ with $2 < q_1,q_2,q_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $\vec{\beta}$ and all triples of Schwartz functions f_1, f_2, f_3*

$$\sup_{\varphi\in\Phi^*} |\Lambda_{\vec{\beta}}(F^\varphi(f_1), F^\varphi(f_2), F^\varphi(f_3))| \le C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|\boldsymbol{F}(f_j)\|_{L^{p_j}\L^{q_j}(S^\infty,S)}. \tag{2.11}$$

Again, we postpone the precise definitions of iterated L^p norms to Section 2.3. There are several differences between our result (2.11) and (2.9). First of all, given the nature of the problem, we have to prove the estimate with a constant independent of $\vec{\beta}$. Moreover, as opposed to [DT15] we do not prove a Hölder inequality, but prove the inequality using the Marcinkiewicz multilinear interpolation for outer L^p spaces. This is caused by the fact that we keep the absolute values outside of the form, since one needs to decompose the functions in question further, using so-called telescoping. Another difference is the appearance of the supremum embedding (2.7) instead of (2.5) on the right hand side. The supremum is required by our methods. Observe that we get the supremum on the left hand side "for free", simply because the inequality holds for any φ in the given class. We shall need a counterpart of the embedding theorem (2.10) for (2.6). Let $p > 1$ and $q > \max(p',2)$. Then

$$\|\boldsymbol{F}(f_j)\|_{L^p\L^q(S^\infty,S)} \le C_{p,q}\|f_j\|_{L^p(\mathbb{R})} \qquad \text{for } j = 1,2,3 \tag{2.12}$$

The proof of (2.12) is an simple modification of the arguments in [DPO15]. We record that the supremum embedding (2.6) was already considered by Muscalu, Tao and Thiele in [MTT02], where they proved the uniform bounds for a n-linear counterpart of the bilinear Hilbert transform in the local L^2 range. One of the ingredients in their proof is essentially equivalent to the above embedding theorem for $2 < p < \infty$ in a discretized setting.

Coupled with the embedding theorem (2.12), Theorem 2.2 implies boundedness of the bilinear Hilbert transform uniformly in $\vec{\beta}$ in the local L^1 Banach triangle, Theorem 2.1. This chapter is a continuation of studies in [War15], where the uniform estimate (2.4) was reproved in the local L^2 range using the outer measure approach. In that case no iterated outer L^p theory was needed. Instead, it was shown that noniterated counterparts of the trilinear outer L^p inequality and the embedding theorem hold with a constant independent of the parameter $\vec{\beta}$ for $p_j > 2$.

2.1.1 Structure of the chapter

The rest of this chapter is organised as follows.

In Section 2.2 we obtain a wave packet decomposition for the bilinear Hilbert transform Proposition 2.3. Having the decomposition in hand, we give a proof of Theorem 2.1 assuming Theorem 2.2.

In Section 2.3 we recall the abstract outer L^p spaces. We prove multilinear interpolation for outer L^p spaces with a general trilinear form Λ, Proposition 2.10. Then, we review the outer measure structure on $\mathbb{R}^3_+$ and adapt it for our purpose. In particular we define sizes (i.e. seminorms) of functions on $\mathbb{R}^3_+$ and the generated outer L^p norms. Due to the nature of our problem we need sizes that dependent on the parameter $\vec{\beta}$.

In Section 2.4 we prove several auxiliary inequalities for outer L^p spaces on $\mathbb{R}^3_+$, including the fact that $\vec{\beta}$ dependent outer L^p norms are dominated by the L^p norms which are independent of the degeneration, Proposition 2.31. The main advantage of this fact is that we can use the iterated embedding (2.12), which is independent of $\vec{\beta}$.

In Section 2.5 we prove the trilinear inequality for the iterated L^p spaces, Theorem 2.2. The proof requires two localized estimates for $\Lambda_{\vec{\beta}}$ uniform in $\vec{\beta}$, corresponding to the two iterations of the outer measure structure. The first one is a time-scale localized estimate Proposition 2.52 and the second one is a frequency-scale localized estimate, Proposition 2.53.

2.2 Wave packet decomposition

From now on we fix $\vec{\beta}$ and all constants in our statements are going to be independent of $\vec{\beta}$. In this section we obtain a wave packet decomposition (2.8) for (2.1), and give a proof of Theorem 2.1 assuming Theorem 2.2. At the end we introduce a slightly less symmetric equivalent trilinear form, which is, however, easier to deal with.

2.2.1 Wave packet decomposition in $\mathbb{R}^3_+$ and proof of Theorem 2.1

We follow the wave packet decomposition in [DT15], however since here we are concerned with the uniform bounds, it is important to keep explicit dependence on $\vec{\beta}$ as it degenerates. A similar, but discretized, wave packet decomposition for the uniform bounds appears for example in [Thi02a], [MTT02]. Roughly, the difference is that in [Thi02a], [MTT02] the phase plane projections on the enlarged time-frequency rectangles of area δ^{-1} are considered, while our decomposition in the discrete setting splits the enlarged rectangle into δ^{-1} rectangles of area 1 with a common frequency interval, see also Chapter 3 for such discretized decomposition.

From now on, assume that $|\beta_2 - \beta_3| \ll 1$, hence $|\alpha_1| \ll 1$, $|\alpha_2|, |\alpha_3| \simeq 1$, $\alpha_2 \approx -\alpha_3$.

Proposition 2.3. *There exist $\varphi \in \Phi^*$ and constants $c_1, c_2 \neq 0$ independent of the parameter $\vec{\beta}$ such that*

$$\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3) = c_1 \Lambda_{\vec{\beta}}(F_1^{\varphi}, F_2^{\varphi}, F_3^{\varphi}) + c_2 \int_{\mathbb{R}} f_1(x) f_2(x) f_3(x)\, dx,$$

where $F_j^\varphi := F^\varphi(f_j)$ *for* $j = 1, 2, 3$.

Using the wave packet decomposition and assuming the iterated Hölder inequality Theorem 2.2 we can now give a proof of Theorem 2.1.

Proof of Theorem 2.1. We have that there exists $\varphi \in \Phi$ and constants c_1, c_2 independent of $\vec{\beta}$ such that

$$\mathrm{BHF}_{\vec{\beta}}(f_1, f_2, f_3) = c_1 \Lambda_{\vec{\beta}}(F_1^\varphi, F_2^\varphi, F_3^\varphi) + c_2 \int_{\mathbb{R}} f_1(x) f_2(x) f_3(x) dx.$$

holds. The second integral on the right hand side is clearly bounded in the local L^1 by the Hölder inequality. Applying Theorem 2.2 we obtain that

$$|\Lambda_{\vec{\beta}}(F_1^\varphi, F_2^\varphi, F_3^\varphi)| \leq C_{p_1,p_2,p_3} \|\boldsymbol{F}_1\|_{L^{p_1}\mathcal{L}^{q_1}(S)} \|\boldsymbol{F}_2\|_{L^{p_2}\mathcal{L}^{q_2}(S)} \|\boldsymbol{F}_3\|_{L^{p_3}\mathcal{L}^{q_3}(S)}$$

for $1 < p_1, p_2, p_3 < \infty$ with $1/p_1 + 1/p_2 + 1/p_3 = 1$ and any $2 < q_1, q_2, q_3 < \infty$ with $1/q_1 + 1/q_2 + 1/q_3 > 1$. Choosing such q_j's with $q_j > \max(p_j', 2)$ and applying (2.12) (for more details, see Proposition 2.25 below) the last display is bounded by

$$C_{p_1,p_2,p_3} \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}.$$

This finishes the proof of Theorem 2.1. □

Now we give the proof of Proposition 2.3.

Proof of Proposition 2.3. Choose $\vec{\alpha} \in \mathbb{R}^3$, so that together $\vec{1} := (1, 1, 1)$, $\vec{\beta}$ they form an orthonormal basis of $\mathbb{R}^3$. Moreover, since we assumed that $|\beta_2 - \beta_3|$ is small, we have $|\alpha_1| = \min(|\alpha_1|, |\alpha_2|, |\alpha_3|)$. The wave packet decomposition shall be obtained in terms of the embedding (2.6). Let $\vec{f}$ be the tensor product of f_1, f_2 and f_3. Let us rewrite

$$\begin{aligned} &\mathrm{BHF}_\beta(f_1, f_2, f_3) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f_1(x - \beta_1 t) f_2(x - \beta_2 t) f_3(x - \beta_3 t) dx \frac{dt}{t} \\ &= \int_{\mathbb{R}^3} \vec{f}(x \cdot \vec{1} + \rho \cdot \vec{\alpha} - t \cdot \vec{\beta})\, dx\, \delta_0(\rho)\, d\rho \,\frac{d\sigma}{\sigma} \end{aligned}$$

The right hand side is equal to

$$-i \int_{\mathbb{R}^3} \widehat{\vec{f}}(\hat{x} \cdot \vec{1} + \hat{\rho} \cdot \vec{\alpha} - \hat{\sigma} \cdot \vec{\beta})\, \delta_0(\hat{x}) d\hat{x}\, d\hat{\rho}\, sgn(\hat{\sigma})\, d\hat{\sigma}$$

Adding and subtracting a multiple of $\int_{\mathbb{R}} f_1(x) f_2(x) f_3(x) dx$ we may concentrate on the half-line $\hat{\sigma} \in (0, \infty)$. Thus in the following we shall perform a wave packet decomposition of

$$\int_{\mathbb{R}^3} \widehat{\vec{f}}(\hat{x} \cdot \vec{1} + \hat{\rho} \cdot \vec{\alpha} - \hat{\sigma} \cdot \vec{\beta})\, \delta_0(\hat{x}) d\hat{x}\, d\hat{\rho}\, \mathbb{1}_{(0,\infty)}(\hat{\sigma})\, d\hat{\sigma}. \tag{2.13}$$

The time-frequency decomposition depends on the fact, how do we decompose $\mathbb{1}_{(0,\infty)}(\hat{t})$ inside the integral. Let $\widehat{\vec{\varphi}}$ be the tensor product of $\hat{\varphi}_1$, $\hat{\varphi}_2$, $\hat{\varphi}_3$. We set [1]

$$\hat{\varphi}_j := D^\infty_{|\alpha_j|} \hat{\varphi}, \qquad \text{for } j = 1, 2, 3, \tag{2.14}$$

[1] Note that here we could also dilate φ with any $|\widetilde{\alpha_j}|$ comparable with $|\alpha_j|$. We make use of this observation in the next subsection.

where $\varphi \in \Phi^*$. Moreover, let $\vec{\tau} = |\alpha_1|\vec{\beta}$. In order to perform a time-frequency decomposition for the uniform problem we shall prove that

$$\int_0^\infty \int_{\mathbb{R}} \widehat{\vec{\varphi}}(t\hat{\sigma}\vec{\beta} + t\hat{x}\vec{1} + t\hat{\rho}\vec{\alpha} - \eta\vec{\alpha} - \vec{\tau})\, d\eta\, \frac{dt}{t} = c\mathbb{1}_{(0,\infty)}(\hat{\sigma}), \tag{2.15}$$

holds for a constant c uniform in $\vec{\beta}$ for $\hat{x} = 0$ and for any $\hat{\rho} \in \mathbb{R}$. Changing the variables and inserting $\hat{x} = 0$, we shall prove that

$$\int_0^\infty \int_{\mathbb{R}} \widehat{\vec{\varphi}}(t\hat{\sigma}\vec{\beta} - \eta\vec{\alpha} - \vec{\tau})\, d\eta\, \frac{dt}{t} = c\mathbb{1}_{(0,\infty)}(\hat{\sigma}).$$

First of all, note that by the change of variables $t\hat{\sigma} \to t$, the left hand side is constant for $\hat{\sigma} > 0$ and equal zero for $\hat{\sigma} < 0$. Assume that $\hat{\sigma} > 0$. Observe that the left hand side of the previous display is comparable with

$$\int_0^\infty \widehat{\vec{\varphi}}(t\vec{\beta} - \vec{\tau})\, \frac{dt}{t}$$

with the constant equal to the measure of $|\mathrm{proj}_{\langle\vec{\alpha}\rangle}(\mathrm{supp}(\widehat{\vec{\varphi}}) \cap \langle\vec{1}\rangle^\perp)| \simeq 1$, where $\mathrm{proj}_{\langle v\rangle}(A)$ is the projection of a set A onto the line spanned by a vector v. The last display is further comparable with

$$\|\widehat{\vec{\varphi}}\|_{L^\infty} \cdot |\mathrm{proj}_{\langle\vec{\beta}\rangle}(\mathrm{supp}(\widehat{\vec{\varphi}}) \cap \langle\vec{1}\rangle^\perp)| \cdot |\vec{\tau}|^{-1} \simeq 1 \cdot |\alpha_1| \cdot |\alpha_1|^{-1} = 1,$$

where $|\mathrm{proj}_{\langle\vec{\beta}\rangle}(\mathrm{supp}(\widehat{\vec{\varphi}}) \cap \langle\vec{1}\rangle^\perp)| \simeq |\alpha_1|$, because since we assumed that $|\beta_2 - \beta_3|$ is small, we have $\vec{\beta} \approx \pm\frac{1}{\sqrt{3}}\vec{1} \pm \frac{\sqrt{2}+1}{\sqrt{3}}(-1,0,0)$ and clearly $|\mathrm{proj}_{\langle(1,0,0)\rangle}(\mathrm{supp}(\widehat{\vec{\varphi}}) \cap \langle\vec{1}\rangle^\perp)| \simeq |\alpha_1|$.

This proves (2.15) and implies that the absolute value of (2.13) is comparable with the absolute value of

$$\int_{\mathbb{R}^3} \widehat{\vec{f}}(\hat{x} \cdot \vec{1} + \hat{\rho} \cdot \vec{\alpha} - \hat{\sigma} \cdot \vec{\beta}) \times \int_0^\infty \int_{\mathbb{R}} \widehat{\vec{\varphi}}(t\hat{\sigma}\vec{\beta} + t\hat{x}\vec{1} + t\hat{\rho}\vec{\alpha} - \eta\vec{\alpha} - \vec{\tau})\, d\eta\, \frac{dt}{t}\, \delta_0(\hat{x}) d\hat{x}\, d\hat{\rho}\, d\hat{\sigma}$$

By the choice of $\hat{\varphi}_i$'s, after an application of the inverse Fourier transform this is equal to

$$\int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}} \prod_{j=1}^3 f_j * M_{\alpha_j\eta t^{-1} + \tau_j t^{-1}} D^1_{|\alpha_j|^{-1}t}\varphi(x)\, dx\, d\eta\, \frac{dt}{t}$$

Changing variables $\eta \mapsto t\eta$ this equals

$$\int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}} \prod_{j=1}^3 F^\varphi(f_j)(x, \alpha_j\eta + \tau_j t^{-1}, |\alpha_j|^{-1}t)\, dx\, d\eta\, dt.$$

Since $\vec{\tau} = |\alpha_1|\vec{\beta} = \delta\vec{\beta}$,

$$\int_{\mathbb{R}^3_+} \prod_{j=1}^3 F^\varphi(f_j)(x, \alpha_j\eta + \delta\beta_j t^{-1}, |\alpha_j|^{-1}t)\, dx\, d\eta\, dt.$$

□

2.2.2 Simplifying assumptions on $\Lambda_{\vec{\beta}}$

For notational convenience later on, we replace the symmetric trilinear form from Proposition 2.3 by a less symmetric trilinear form, which is comparable.

Note that if $|\alpha_1| \ll 1$, then in (2.14) we may take φ_j's, such that $\hat{\varphi}_1 := D^{\infty}_{|\alpha_1|}\hat{\varphi}$ as earlier, and $\hat{\varphi}_2 = \hat{\varphi}_3 = D^{\infty}_{|\alpha_2|}\hat{\varphi}$. We obtain that the trilinear form in the previous display is comparable, with a uniform constant, with the following form

$$\int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}} F^{\varphi}(f_1)(x,\alpha_1\eta + |\alpha_1|\beta_1 t^{-1}, |\alpha_1|^{-1}t) \times \prod_{j=2}^{3} F^{\varphi}(f_j)(x,\alpha_j\eta + |\alpha_1|\beta_j t^{-1}, |\alpha_2|^{-1}t)\, dx\, d\eta\, dt.$$

From now on we set

$$\delta := |\alpha_1|/|\alpha_2| \tag{2.16}$$

Note that $\delta \ll 1$. Changing variables $t \mapsto |\alpha_1|t$, $\eta \mapsto \alpha_1^{-1}\eta$ the above is equal to ± 1 times

$$\int_{\mathbb{R}^3_+} F^{\varphi}(f_1)(x,\eta + \beta_1 t^{-1}, t) \prod_{j=2}^{3} F^{\varphi}(f_j)(x,\alpha_1^{-1}\alpha_j\eta + \delta\beta_j(\delta t)^{-1}, \delta t)\, dt\, d\eta\, dx. \tag{2.17}$$

In the rest of the chapter we are going to work with the above form and denote it with $\Lambda_{\vec{\beta}}(\cdot,\cdot,\cdot)$.

2.3 Outer L^p spaces

In this section we describe the outer L^p space setting on $\mathbb{R}^3_+ = \mathbb{R}\times\mathbb{R}\times\mathbb{R}_+$ that we will be using in this chapter. The outer measure L^p space framework follows the one introduced in [DT15] and further developed in [Ura16].

2.3.1 Generalities

We recollect the theory of outer L^p spaces introduced in [DT15]. Let $\mathbb{X}$ be a locally compact metric space; denote by $\mathcal{P}(\mathbb{X})$ the collection of all its subsets, by $\mathcal{B}(\mathbb{X})$ the set of Borel functions on $\mathbb{X}$ and by $C(\mathbb{X})$ the set of continuous functions on $\mathbb{X}$. Let us fix a collection of *generating sets* $\mathbb{T} \subset \mathcal{P}(\mathbb{X})$ that are locally compact Borel measurable subsets.

Definition 2.4 (Outer measure)**.** *An* outer measure *on $\mathbb{X}$ is a set functional $\mu\colon \mathcal{P}(\mathbb{X}) \to [0,\infty]$ that is σ-subadditive*

$$\mu\Big(\bigcup_{n\in\mathbb{N}} E_n\Big) \le \sum_{n\in\mathbb{N}} \mu(E_n), \qquad \text{for } E_n \subset \mathbb{X},$$

monotone

$$\mu(E) \le \mu(E'), \qquad \text{for } E \subset E' \subset \mathbb{X}$$

and $\mu(\emptyset) = 0$.

We refer to a function $\overline{\mu}\colon \mathbb{T} \to [0,\infty]$ as a premeasure that generates the outer measure μ via

$$\mu(E) := \inf\Big\{\sum_{n\in\mathbb{N}} \overline{\mu}(T_n)\colon E \subset \bigcup_{n\in\mathbb{N}} T_n,\ \text{where } T_n \in \mathbb{T}\Big\}. \tag{2.18}$$

Let $\mathbb{T}^\cup$ be the set of countable unions of sets in $\mathbb{T}$ i.e.

$$E \in \mathbb{T}^\cup \text{ if } E = \bigcup_{n\in\mathbb{N}} T_n, \qquad \text{where } T_n \in \mathbb{T}.$$

Definition 2.5 (Size). *A* size *on $\mathbb{X}$ is a functional $\|\cdot\|_{\mathcal{S}}\colon \mathcal{B}(\mathbb{X}) \to [0,\infty]^{\mathbb{T}}$ that satisfies*

1. *if $F, G \in \mathcal{B}(\mathbb{X})$ with $|F| \le |G|$ and $T \in \mathbb{T}$, then $\|F\|_{\mathcal{S}(T)} \lesssim \|G\|_{\mathcal{S}(T)}$,*
2. *there exists $C > 0$ such that for every $T \in \mathbb{T}$ $\|F+G\|_{\mathcal{S}(T)} \le C(\|F\|_{\mathcal{S}(T)} + \|G\|_{\mathcal{S}(T)})$ for all $F, G \in \mathcal{B}(\mathbb{X})$,*
3. *for every $T \in \mathbb{T}$, $\lambda \in \mathbb{C}$ and for every $F \in \mathcal{B}(\mathbb{X})$, $\|\lambda F\|_{\mathcal{S}(T)} = |\lambda| \|F\|_{\mathcal{S}(T)}$.*

Definition 2.6 (Outer L^p norms). *Given an outer measure μ, a size $\|\cdot\|_{\mathcal{S}}$, and a generating collection $\mathbb{T} \subset \mathcal{P}(\mathbb{X})$, the outer L^p quasi-norm of a function $F \in \mathcal{B}(\mathbb{X})$ with $p \in (0,\infty)$ is given by*

$$\|F\|_{L^p_\mu(\mathcal{S})} := \Big(p \int_0^\infty \lambda^p \mu(\|F\|_{\mathcal{S}} > \lambda) \frac{d\lambda}{\lambda}\Big)^{1/p},$$

where

$$\mu(\|F\|_{\mathcal{S}} > \lambda) := \inf\{\mu(E_\lambda) : E_\lambda \in \mathbb{T}^\cup,\ \|\mathbb{1}_{\mathbb{X}\setminus E_\lambda} F\|_{L^\infty(\mathcal{S})} \le \lambda\}$$

where the outer L^∞ quasi-norm is given by

$$\|F\|_{L^\infty(\mathcal{S})} := \sup_{T\in\mathbb{T}} \|F\|_{\mathcal{S}(T)}.$$

The outer weak L^p norms are given by

$$\|F\|_{L^{p,\infty}_\mu(\mathcal{S})} := \Big(\sup_{\lambda>0} p\lambda^p \mu(\|F\|_{\mathcal{S}} > \lambda)\Big)^{1/p}.$$

For example, the standard Lebesgue $L^p(\mathbb{R}^n)$ space can be constructed as an outer L^p space. Taking the generating collection $\mathbb{T}$ to be set of dyadic cubes in $\mathbb{R}^n$ and $\overline{\mu}$ the standard measure of a cube, μ becomes the Lebesgue measure. Setting $\|\cdot\|_{\mathcal{S}(T)}$ to be the average of $|F|$ over a cube T, $\|\cdot\|_{L^\infty(\mathcal{S})}$ is comparable with $\|\cdot\|_{L^\infty(\mathbb{R}^n)}$ and consequently $\|\cdot\|_{L^p_\mu(\mathcal{S})}$ is comparable with $\|\cdot\|_{L^p(\mathbb{R}^n)}$. Let us also remark that defining L^0 appropriately, it is possible to view the outer L^p spaces as interpolation spaces between L^0 and L^∞. See [War15] for details of such approach to the outer L^p spaces.

Given a size family $\|\cdot\|_{\mathcal{S}(T)}$ and $c > 0$ set

$$\|F\|_{c\mathcal{S}(T)} := c\|F\|_{\mathcal{S}(T)}.$$

Given two size families $\|\cdot\|_{\mathcal{S}_1(T)}$ and $\|\cdot\|_{\mathcal{S}_2(T)}$ we define the size family

$$\|F\|_{(\mathcal{S}_1+\mathcal{S}_2)(T)} := \|F\|_{\mathcal{S}_1(T)} + \|F\|_{\mathcal{S}_2(T)}.$$

We also define the size family on $\boldsymbol{F} = (F^1, F^2) \in \mathcal{B}(\mathbb{X}) \times \mathcal{B}(\mathbb{X})$ via

$$\|\boldsymbol{F}\|_{(\mathcal{S}_1,\mathcal{S}_2)(T)} := \|F^1\|_{\mathcal{S}_1(T)} + \|F^2\|_{\mathcal{S}_2(T)}.$$

Convex set decomposition

We introduce convex sets, which are counterparts of the convex trees of tiles in the general setting. They come in handy when one wants to represent the outer L^p norm as an ℓ^p sum over pairwise disjoint sets.

Definition 2.7 (Convex sets)**.**

- *We call a subset $\Delta T \subset \mathbb{X}$ a convex tree if $\Delta T = T \setminus K$ for some $T \in \mathbb{T}$ and $K \in \mathbb{T}^{\cup}$.*
- *We call a set $E \subset \mathbb{X}$ convex if it is of the form $E = K_1 \setminus K_2$, with $K_1, K_2 \in \mathbb{T}^{\cup}$.*

In the next lemma we control an ℓ^p sum coming from appropriately selected pairwise disjoint convex sets. We shall need this result, when proving outer L^p comparison inequalities in Section 2.4. The proof is a standard decomposition of the space into level sets and then making them pairwise disjoint.

Lemma 2.8 (Decomposition into convex trees)**.** *Suppose we are given an outer measure space $(\mathbb{X}, \mu, \|\cdot\|_{\mathcal{S}})$. Fix $F\colon \mathbb{X} \to \mathbb{R}$ and let $0 < p < \infty$. There exists a decomposition into pairwise disjoint convex trees $\mathbb{X} = \bigcup_{k\in\mathbb{Z}} \bigcup_{\Delta T \in \Phi_k} \Delta T$ such that*

$$\left(\sum_{k\in\mathbb{Z}} \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F \mathbb{1}_{\Delta T}\|_{L^\infty(\mathcal{S})}^p \right)^{1/p} \lesssim \|F\|_{L^p_\mu(\mathcal{S})},$$

with a constant independent of F.

Proof. First of all note that using standard argument we can replace the integral in the definition of the outer L^p norm so that

$$\sum_{k\in Z} 2^{kp} \mu(\|F\|_{\mathcal{S}} > 2^k) \lesssim \|F\|_{L^p_\mu(\mathcal{S})}^p.$$

Now for each $k \in \mathbb{Z}$, by definition, choose a collection of $\Psi_k \subset \mathbb{T}$ such that

$$\sum_{T \in \widetilde{\mathcal{T}}_k} \mu(T) \lesssim \mu(\|F\|_{\mathcal{S}} > 2^k), \qquad \|F \mathbb{1}_{\mathbb{X} \setminus \bigcup \widetilde{\mathcal{T}}_k}\|_{L^\infty(\mathcal{S})} \le 2^k.$$

We just need to make all the selected sets pairwise disjoint what will lead to convex trees. Fix k and let $T \in \widetilde{\mathcal{T}}_k$. Define

$$\Delta T := T \setminus \bigcup_{\widetilde{\mathcal{T}}_k \ni T' \neq T} T' \setminus \bigcup_{n>k} \bigcup \widetilde{\mathcal{T}}_n$$

Note that each ΔT is a convex tree. Let $\mathcal{T}_k$ denote the collection of such ΔT for each $k \in \mathbb{Z}$. Moreover, $\mathbb{X} = \bigcup_{k\in\mathbb{Z}} \bigcup_{\Delta T \in \mathcal{T}_k} \Delta T$, ΔT's are pairwise disjoint and

$$\|F \mathbb{1}_{\Delta T}\|_{L^\infty(\mathcal{S})} \le \|F \mathbb{1}_{\mathbb{X} \setminus \widetilde{\mathcal{T}}_{k+1}}\| \lesssim 2^k,$$
$$\sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \le \sum_{T \in \widetilde{\mathcal{T}}_k} \mu(T) \lesssim \mu(\|F\|_{\mathcal{S}} > 2^k).$$

This gives

$$\begin{aligned} \sum_{k\in\mathbb{Z}} \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F \mathbb{1}_{\Delta T}\|_{L^\infty(\mathcal{S})}^p &\le \sum_{k\in\mathbb{Z}} 2^{kp} \sum_{T \in \widetilde{\mathcal{T}}_k} \mu(T) \\ &\le \sum_{k\in\mathbb{Z}} 2^{kp} \mu(\|F\|_{\mathcal{S}} > 2^k) \lesssim \|F\|_{L^p_\mu(\mathcal{S})}^p. \end{aligned}$$

□

Marcinkiewicz interpolation

We state the Marcinkiewicz interpolation for outer L^p spaces, Proposition 1.7 in [Ura17]. It is analogous to the classical interpolation theorem, the only difference is that replace L^p norms with their outer counterparts.

Proposition 2.9 (Marcinkiewicz interpolation)**.** *Let* $(\mathbb{X}_1, \mu_1, \|\cdot\|_{S_1})$, $(\mathbb{X}_2, \mu_2, \|\cdot\|_{S_2})$ *be two outer measure spaces with sizes. Assume that* $p_1, p_2, q_1, q_2 \in (0, \infty]$ *and let* T *be an operator that satisfies*

- $|T(\lambda F)| = |\lambda T(F)|$ *for all* $F \in L^{p_1}(\mathbb{X}_1) + L^{q_1}(\mathbb{X}_1)$ *and* $\lambda \in \mathbb{C}$,
- $|T(F+G)| \leq C(|T(F)| + |T(G)|)$ *for all* $F, G \in L^{p_1}(\mathbb{X}_1) + L^{q_1}(\mathbb{X}_1)$,
- *for all* $F \in L^{p_1}(S_1)$
$$\|T(F)\|_{L^{p_2,\infty}_{\mu_2}(S_2)} \leq C_0 \|F\|_{L^{p_1}_{\mu_1}(S_1)}$$
- *for all* $F \in L^{q_1}(S_1)$
$$\|T(F)\|_{L^{q_2,\infty}_{\mu_2}(S_2)} \leq C_1 \|F\|_{L^{q_1}_{\mu_1}(S_1)}.$$

Then for any $\theta \in (0,1)$, $\frac{1}{r_{1,\theta}} = \frac{1-\theta}{p_1} + \frac{\theta}{q_1}$ *and* $\frac{1}{r_{2,\theta}} = \frac{1-\theta}{p_2} + \frac{\theta}{q_2}$ *it holds that*

$$\|T(F)\|_{L^{r_{2,\theta}}_{\mu_2}(S_2)} \lesssim_{\theta, p_1, p_2, q_1, q_2} C_0^{1-\theta} C_1^{\theta} \|F\|_{L^{r_{1,\theta}}_{\mu_1}(S_1)}.$$

Multilinear Marcinkiewicz interpolation

We prove multilinear Marcinkiewicz interpolation for outer L^p spaces, Proposition 2.10. We record that it is quite reminiscent of the restricted type interpolation that appears for example in [Thi06]. In this chapter we shall need the next proposition only in the case $n = 3$.

Proposition 2.10 (Multilinear Marcinkiewicz interpolation)**.** *Suppose we are given a collection of outer measure spaces with sizes* $(\mathbb{X}, \mu_j, \|\cdot\|_{\mathcal{S}_j})$ *for* $j = 1, 2, .., n$. *Let* Λ *be an* n*-linear form defined for* n*-tuples of functions* $\in \mathcal{B}(\mathbb{X})$. *Suppose that* $F_j^0, F_j \in \mathcal{B}(\mathbb{X})$ *for* $j = 1, 2, .., n$, $1/\vec{p} := (1/p_1, 1/p_2, ..., 1/p_n)$ *with* $\sum_i 1/p_j = 1$ *and* $1 < p_j < \infty$ *are such that for all* $V_j, W_j \in \mathbb{T}^{\cup}$, $j = 1, 2, .., n$

$$\Lambda(\mathbb{1}_{V_1 \setminus W_1} F_1^0, \mathbb{1}_{V_2 \setminus W_3} F_2^0, .., \mathbb{1}_{V_n \setminus W_n} F_n^0) \lesssim \prod_{j=1}^n \mu_j(V_j)^{\alpha_j} \|\mathbb{1}_{V_j \setminus W_j} F_j\|_{L^\infty(\mathcal{S}_j)}$$

holds for $\sum_i \alpha_i = 1$, *with* $\vec{\alpha} := (\alpha_1, \alpha_2, .., \alpha_n)$ *in the neighbourhood of* $1/\vec{p}$. *Then*

$$|\Lambda(F_1^0, F_2^0, .., F_n^0)| \lesssim \prod_{j=1}^n \|F_j\|_{L^{p_j}_{\mu}(\mathcal{S}_j)}. \tag{2.19}$$

We could not use the outer Hölder inequality of [DT15] for our purpose, since it requires a stronger assumption than we were able to obtain, for details see the later part of this chapter, in particular (2.41).

Proof. In the proof we set $\|\cdot\|_{L^{p_j}(\mathcal{S}_j)} := \|\cdot\|_{L^{p_j}_{\mu_j}(\mathcal{S}_j)}$. Assume by scaling that $\|F_j\|_{L^{p_j}(\mathcal{S}_j)} = 1$ for $j = 1, 2, .., n$.

First we split each F_j according the level sets of $\|\cdot\|_{\mathcal{S}_j}$ at level $2^{k_j/p_j}$

$$\Lambda(F_1^0, F_2^0, .., F_n^0) \\ = \sum_{k_1,k_2,..,k_n} \Lambda(F_1^0 \mathbb{1}_{W_{k_1} \setminus W_{k_1-1}}, F_2^0 \mathbb{1}_{W_{k_2} \setminus W_{k_2-1}}, .., F_n^0 \mathbb{1}_{W_{k_n} \setminus W_{k_n-1}})$$

so that we have the properties

- $\mu_j(W_{k_j}) \lesssim A_{j,k_j} := \mu_j(\|F_j\|_{\mathcal{S}_j} > 2^{-k_j/p_j}) \lesssim 2^{k_j}\|F_j\|^{p_j}_{L^{p_j}(\mathcal{S}_j)}$
- $\|F_j \mathbb{1}_{W^c_{k_j-1}}\|_{L^\infty(\mathcal{S}_j)} \lesssim 2^{k_j/p_j}$

for $j = 1, 2, 3$. Using the restricted type and the properties listed above one obtains that this is bounded by

$$\sum_{k_1,k_2,..,k_n \in \mathbb{Z}} \prod_{j=1}^{n} \mu_j(W_{k_j})^{\alpha_j - 1/p_j} \mu_j(W_{k_j})^{1/p_j} \|F_j \mathbb{1}_{W^c_{k_j-1}}\|_{L^\infty(\mathcal{S}_j)} \\ \lesssim \sum_{k_1,k_2,..,k_n \in \mathbb{Z}} 2^{\sum_{j=1}^n (\alpha_j - 1/p_j)k_j} \prod_{j=1}^{n} \|F_j\|^{(\alpha_j - 1/p_j)p_j}_{L^{p_j}(S_j)} 2^{k_j/p_j} A^{1/p_j}_{j,k_j} \\ \lesssim \sum_{k_1,k_2,..,k_n \in \mathbb{Z}} 2^{\sum_{j=1}^n (\alpha_j - 1/p_j)k_j} \prod_{j=1}^{n} 2^{k_j/p_j} A^{1/p_j}_{j,k_j},$$

Let k be the average of k_1, k_2, k_3. Choosing $\vec{\alpha}$ in the neighbourhood of $1/\vec{p}$ appropriately we can assume that

$$2^{\sum_{j=1}^3 (\alpha_j - 1/p_j)k_j} = 2^{\sum_{j=1}^3 (\alpha_j - 1/p_j)(k_j - k)} \lesssim 2^{-\varepsilon \max_{j=1,2,..,n} |k_j - k|}.$$

We can then bound the previous display as follows changing the summation parameters to k, $\overline{k}_j := k_j - k \in \frac{1}{n}\mathbb{Z}$ for $j = 1, 2, .., n-1$

$$\sum_{\overline{k}_i, k \in \frac{1}{n}\mathbb{Z}} 2^{-\varepsilon \max_{0<i<n} |\overline{k}_i|} \prod_{j=1}^{n-1} 2^{(k+\overline{k}_j)/p_j} A^{1/p_j}_{j,k+\overline{k}_j} 2^{(k - \sum_{i=1}^{n-1} \overline{k}_i)/p_n} A^{1/p_n}_{n,k-\sum_{i=1}^{n-1}\overline{k}_i}$$

Applying Hölder's inequality in the sum over k with the exponents p_1, p_2, p_3 we obtain that it is bounded by (since $\overline{k}_j$ are fixed translations now) $\prod_{j=1}^n \|F_j\|_{L^{p_j}(\mathcal{S}_j)} = 1$. This means that we are just left with the series

$$\sum_{\overline{k}_1, \overline{k}_2, .., \overline{k}_{n-1} \in \frac{1}{n}\mathbb{Z}} 2^{-\varepsilon \max_{0<j<n} |\overline{k}_j|}$$

which is summable. □

The following short lemma lets us dominate a vector valued outer L^p norms by a sum of outer L^p norms taken coordinatewise.

Lemma 2.11. *Suppose we are given an outer measure space $(\mathbb{X}, \mu)$ together with two families of sizes $\|\cdot\|_{\mathcal{S}_1}$, $\|\cdot\|_{\mathcal{S}_2}$ indexed by $T \in \mathbb{T}$. For any $0 < p \leq \infty$ and $\boldsymbol{F} = (F^1, F^2) \in \mathcal{B}(\mathbb{X}) \times \mathcal{B}(\mathbb{X})$ it holds that*

$$\|F^1\|_{L^p_\mu(\mathcal{S}_1)} + \|F^2\|_{L^p_\mu(\mathcal{S}_2)} \lesssim \|\boldsymbol{F}\|_{L^p_\mu((S_1,S_2))} \lesssim \|F^1\|_{L^p_\mu(\mathcal{S}_1)} + \|F^2\|_{L^p_\mu(\mathcal{S}_2)}.$$

Proof. For $0 < p < \infty$ the lemma follows directly from the following standard fact and integrating over $\lambda \in \mathbb{R}_+$

$$\frac{1}{2}\sum_{j=1}^{2} \mu(\|F^j\|_{\mathcal{S}_j} > \lambda) \leq \mu(\|\boldsymbol{F}\|_{\mathcal{S}} > \lambda) \leq \sum_{j=1}^{2} \mu(\|F^j\|_{\mathcal{S}_j} > \lambda/2).$$

For $p = \infty$ it follows from definition. □

2.3.2 Outer measures and sizes on $\mathbb{R}^3_+$

Trees and outer measures in time-frequency-scale space

We consider $\mathbb{R}^3_+ = \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$ endowed with the euclidean metric. Let $\boldsymbol{\Theta} = (\Theta, \Theta^{(in)})$ where $\Theta \supset \Theta^{(in)} \ni 0$ be two open intervals and let $\delta \in (0, 1]$. The outer L^p structure we introduce on $\mathbb{R}^3_+$ depends on these parameters.

Definition 2.12 (Trees)**.** *For $(x, \xi, s) \in \mathbb{R}^3_+$ we define*

$$\begin{aligned}
T_{\boldsymbol{\Theta},\delta}(x,\xi,s) :=& \big\{(y,\eta,t) \in \mathbb{R}^3_+ \colon t < \min(s - |y-x|, \delta s),\ t(\eta-\xi) \in \Theta\big\} \\
=& T^{(in)}_{\boldsymbol{\Theta},\delta}(x,\xi,s) \cup T^{(out)}_{\boldsymbol{\Theta},\delta}(x,\xi,s), \\
T^{(in)}_{\boldsymbol{\Theta},\delta}(x,\xi,s) :=& \big\{(y,\eta,s) \in T_{\boldsymbol{\Theta},\delta}(x,\xi,s) \colon t(\eta-\xi) \in \Theta^{(in)}\big\}, \\
T^{(out)}_{\boldsymbol{\Theta},\delta}(x,\xi,s) :=& \big\{(y,\eta,t) \in T_{\boldsymbol{\Theta},\delta}(x,\xi,t) \colon t(\eta-\xi) \in \Theta^{(out)}\big\} \quad \Theta^{(out)} := \Theta \setminus \Theta^{(in)}.
\end{aligned}$$

The set of all trees with parameters $(\boldsymbol{\Theta}, \delta)$ is denoted by $\mathbb{T}_{\boldsymbol{\Theta},\delta}$.

For each $T_{\boldsymbol{\Theta},\delta}(x,\xi,s) \in \mathbb{T}_{\boldsymbol{\Theta},\delta}$ we define the pre-measure

$$\overline{\mu_{\boldsymbol{\Theta},\delta}}(T_{\boldsymbol{\Theta},\delta}(x,\xi,s)) = s.$$

that generates an outer measure $\mu_{\boldsymbol{\Theta},\delta}$ as in (2.18).

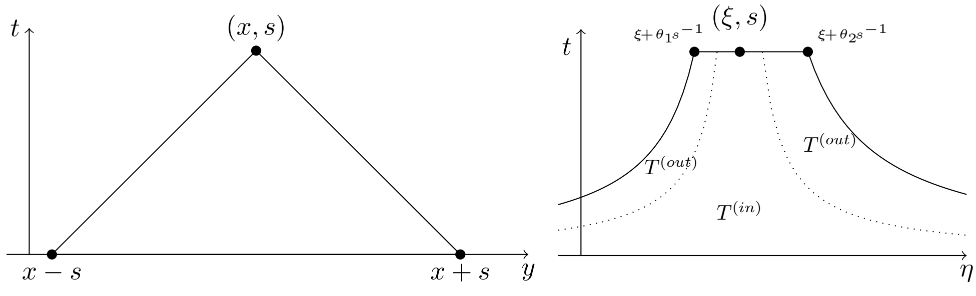

Figure 2.2: The tree $T(x,\xi,s) \in \mathbb{T}_{\boldsymbol{\Theta},1}$, where $\Theta = (\theta_1, \theta_2)$.

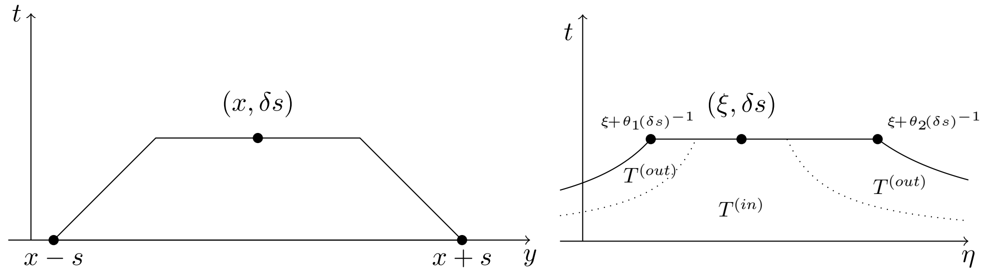

Figure 2.3: The tree $T(x, \xi, s) \in \mathbb{T}_{\Theta,\delta}$, where $\Theta = (\theta_1, \theta_2)$.

Generalized trees

The following is the definition of generalized trees. We use such sets for their nice geometric properties, as they are closed under finite intersections, however note that we do not define sizes for them. Moreover, one can recover trees and strips (Definition 2.24) as special cases of generalized trees.

Definition 2.13 (Generalized trees)**.** *For $x \in \mathbb{R}$, $\xi_1 \in [-\infty, \infty)$, $\xi_2 \in (-\infty, \infty]$, $s \in \mathbb{R}_+$ and $0 < \delta \le 1$ we define the generalized tree $\mathcal{T}_{\Theta}(x, \xi_1, \xi_2, s, \delta)$ as*

$$\{(y, \eta, t) \in \mathbb{R}^3_+ \colon 0 < t < \min(\delta s, s - |y - x|),\ \xi_1 + \theta_1 t^{-1} \le \eta \le \xi_2 + \theta_2 t^{-1}\},$$

where $\Theta = (\theta_1, \theta_2)$. We denote the set of all generalized trees $\mathcal{T}_{\Theta}$ with $\mathscr{T}_{\Theta}$.

We recover trees and strips from generalized trees, since

$$\mathcal{T}_{\Theta}(x, \xi, \xi, s, \delta) = T_{\Theta,\delta}(x, \xi, s) \in \mathbb{T}_{\Theta,\delta}.$$

and

$$\mathcal{T}_{\Theta}(x, -\infty, \infty, s, 1) = D(x, s) \in \mathbb{D},$$

where $\mathbb{D}$ is the set of strips, Definition 2.24.

Boundaries

We shall define boundaries of $A \cap T$ for sets A having nice geometric properties. Such objects come up naturally later on, when we differentiate functions of the form $F\mathbb{1}_A$ and a derivative falls on $\mathbb{1}_A$, where F is an embedding. They are controlled via the boundary sizes, defined below.

Definition 2.14. *Let $A \subset \mathbb{R}^+_3$ be measurable. For every tree $T = T_{\Theta,\delta}(x, \xi, s) \in \mathbb{T}_{\Theta,\delta}$, $y \in B_s(x)$ and $\theta \in \Theta$ we define*

$$A^{\theta}_T(y) := \{t \in \mathbb{R}_+ \colon (y, \xi + \theta t^{-1}, t) \in A \cap T\}.$$

Definition 2.15. *We call a measurable $A \subset \mathbb{R}^3_+$ boundary admissible if for every $T(x, \xi, s) \in \mathbb{T}_{\Theta,\delta}$, $A^{\theta}_T(y) = (t^{\theta,-}_{A,T}(y), t^{\theta,+}_{A,T}(y))$ is such that $t^{\theta,\pm}_{A,T} \colon B_s(x) \to \mathbb{R}_+$ is a Lipschitz function. Moreover, for such A, a tree $T = T(x, \xi, s) \in \mathbb{T}_{\Theta,\delta}$ and $\theta \in \Theta$ we set*

$$\partial^{\theta}_{A,T} = \{(y, \xi + \theta(t^{\pm,\theta}_{A,T}(y))^{-1}, t^{\pm,\theta}_{A,T}(y)) \colon y \in B_{s(1-\delta)}(x)\}.$$

We also set $\partial^{\theta}_T = \{(y, \theta(\delta s)^{-1}, \delta s) \colon y \in B_{s(1-\delta)}(x)\}$.

Remark 2.16. *The reason for introducing boundary admissible sets is that we want the quantities in following Definition 2.19 to make sense. In practice, the only boundary admissible sets that we will be dealing with are of the form $K \setminus L$, where $K, L \in \mathscr{T}_{\mathbf{\Theta}}^{\cup}$, i.e. they are both countable unions of generalized trees.*

Sizes in time-frequency-scale space

The sizes we introduce depend on the parameters $(\mathbf{\Theta}, \delta)$. $\|\cdot\|_{S^2}$ and $\|\cdot\|_{S^\infty}$, which are very similar to the standard sizes introduced in [DT15], additionally parametrized by δ.

Definition 2.17 (The size S^2)**.** *For $F \in \mathcal{B}(\mathbb{R}^3_+)$ we introduce the size family $\|\cdot\|_{S^2_{\mathbf{\Theta},\delta}(T)}$ indexed by $T = T_{\mathbf{\Theta},\delta}(x,\xi,u) \in \mathbb{T}_{\mathbf{\Theta},\delta}$ as*

$$\|F\|_{S^2_{\mathbf{\Theta},\delta}(T)} := \Big(\int_{T^{(out)} \cap \{(y,\eta,t) \in \mathbb{R}^3_+ \,:\, y \in B_{s(1-\delta)}(x)\}} |F(y,\eta,t)|^2 \, dy \, d\eta \, dt\Big)^{1/2}$$

i.e.

$$\|F\|_{S^2_{\mathbf{\Theta},\delta}(T)} = \Big(\int_{\Theta^{(out)}} \int_{B_{s(1-\delta)}(x)} \int_0^{\min(\delta s, s-|y-x|)} |F(y, \xi + \theta t^{-1}, \delta t)|^2 \, \frac{dt}{t} \, dy \, d\theta\Big)^{1/2}.$$

Definition 2.18 (The size S^∞)**.** *For $F \in \mathcal{B}(\mathbb{R}^3_+)$ we introduce the size family $\|\cdot\|_{S^2_{\mathbf{\Theta},\delta}(T)}$ indexed by $T = T_{\mathbf{\Theta},\delta}(x,\xi,u) \in \mathbb{T}_{\mathbf{\Theta},\delta}$ as*

$$\|F\|_{S^\infty_{\mathbf{\Theta},\delta}(T)} := \sup_{(y,\eta,t) \in T} |F(y,\eta,t)|$$

We introduce the boundary size R_A which is the supremum over L^2 averages over the boundaries $\partial^\theta_{A,T}$. They are used to control contribution from boundaries appearing as byproduct of integrating functions of the form $F\mathbb{1}_A$ by parts.

Definition 2.19 (The size R_A)**.** *Given a tree $T = T_{\mathbf{\Theta},\delta}(x,\xi,s) \in \mathbb{T}_{\mathbf{\Theta},\delta}$, a boundary admissible set $A \subset \mathbb{R}^3_+$ we introduce the size family for $F \in C(\mathbb{R}^3_+)$ (continuous functions)*

$$\|F\|_{\widetilde{R}_{\mathbf{\Theta},\delta,A}(T)} = \sup_{\theta \in \Theta^{(in)}} \|F\|_{R^\theta_{\mathbf{\Theta},\delta,A}(T)},$$

where

$$\|F\|_{R^\theta_{\mathbf{\Theta},\delta,A}(T)} = \|F\|_{R^{\theta,-}_{\mathbf{\Theta},\delta,A}(T)} + \|F\|_{R^{\theta,+}_{\mathbf{\Theta},\delta,A}(T)},$$

with

$$\begin{aligned} \|F\|_{R^{\theta,\pm}_{\mathbf{\Theta},\delta,A}(T)} &:= \lim_{\varepsilon \to 0} \Big(\frac{1}{s} \int_{\partial^{\theta,\pm,\varepsilon}_{A,T}} |F\mathbb{1}_A|^2\Big)^{1/2} \\ &:= \lim_{\varepsilon \to 0} \Big(\frac{1}{s} \int_{B_{s(1-\delta)}(x)} |F\mathbb{1}_A(y, \xi + \theta t^{\theta,\pm}_{A,T}(y)^{-1}, t^{\theta,\pm}_{A,T}(y) \mp \varepsilon)|^2 \, dy\Big)^{1/2}. \end{aligned}$$

indexed by $T = T_{\mathbf{\Theta},\delta}(x,\xi,s) \in \mathbb{T}_{\mathbf{\Theta},\delta}$. Since $t^{\pm}_{A,T}$ are Lipschitz functions, the above integral makes sense. Finally, given a tree $T = T_{\mathbf{\Theta},\delta}(x,\xi,s) \in \mathbb{T}_{\mathbf{\Theta},\delta}$ we define

$$\|F\|_{R_{\mathbf{\Theta},\delta,A}(T)} := \sup_{\delta/4 < \widetilde{\delta} \leq \delta} \|F\|_{\widetilde{R}_{\mathbf{\Theta},\widetilde{\delta},A}(T_{\mathbf{\Theta},\widetilde{\delta}}(x,\xi,s))}, \tag{2.20}$$

where each $T_{\mathbf{\Theta},\widetilde{\delta}}(x,\xi,s) \in \mathbb{T}_{\mathbf{\Theta},\widetilde{\delta}}$.

Remark 2.20. *The reason for introducing the boundary size as a limit as $\varepsilon \to 0$ is only a technicality, caused by the fact that we do not necessarily have $\partial^{\theta}_{A,T} \subset A$. However, for $F \in C(\mathbb{R}^3_+)$ we actually have the equality (note that we do not restrict to $\mathbb{1}_A$)*

$$\begin{aligned}\|F\|_{R^{\theta,\pm}_{\mathbf{\Theta},\delta,A}(T)} &= (\frac{1}{s}\int_{\partial^{\theta,\pm}_{A,T}} |F|^2)^{1/2} \\ &= (\frac{1}{s}\int_{B_{s(1-\delta)}(x)} |F(y,\xi+\theta t^{\theta,\pm}_{A,T}(y)^{-1}, t^{\theta,\pm}_{A,T}(y))|^2\, dy)^{1/2}.\end{aligned}$$

Remark 2.21. *The functions F that we will be dealing with in this chapter come from embeddings, i.e. it holds that $F = F(f)$. It is not difficult to see that $\{F^{\varphi}(f) \colon \varphi \in \Phi\}$ satisfies the following: for any $\overline{x} \in \mathbb{R}^3_+$ and $\varepsilon > 0$ there exists an open neighbourhood $U_{\overline{x}}$, such that for any $\varphi \in \Phi$, $\overline{x}_1, \overline{x}_2 \in U_{\overline{x}}$ it holds that $|F^{\varphi}(f)(\overline{x}_1) - F^{\varphi}(f)(\overline{x}_2)| < \varepsilon$. Hence $F(f)$, which is the supremum over $\varphi \in \Phi$, belongs to $C(\mathbb{R}^3_+)$.*

Remark 2.22. *We shall be using only a discretized collection of trees and the supremum in* (2.20) *is needed, so that that there are no "gaps" between scales. This is particularly important in the proof of Lemma 2.46, however all other estimates in this chapter, where we use $R_{\mathbf{\Theta},\delta,A}$, hold with $\widetilde{R}_{\mathbf{\Theta},\delta,A}$ as well.*

We additionally set for a Lebesgue measurable $A \subset \mathbb{R}^3_+$

$$\|F\|_{S^2_{\mathbf{\Theta},\delta,A}} := \|F\mathbb{1}_A\|_{S^2_{\mathbf{\Theta},\delta}}, \qquad \|F\|_{S^{\infty}_{\mathbf{\Theta},\delta,A}} := \|F\mathbb{1}_A\|_{S^{\infty}_{\mathbf{\Theta},\delta}}.$$

Let $\boldsymbol{F} = (F^1, F^2)$ where $F^1, F^2 \in \mathcal{B}(\mathbb{R}^3_+)$. Using the above we introduce the size families for $0 < \delta \leq 1$

$$\begin{aligned}\|\boldsymbol{F}\|_{S_{\mathbf{\Theta},\delta,A}(T)} &:= \|F^1\|_{R_{\mathbf{\Theta},\delta,A}(T)} + \|F^1\mathbb{1}_T\|_{L^{\infty}(\delta^{1/2}S^{\infty}_{\mathbf{\Theta},1,A})} + \|F^2\mathbb{1}_T\|_{L^{\infty}(\delta^{1/2}S^2_{\mathbf{\Theta},1,A})} + \|F^2\|_{S^2_{\mathbf{\Theta},\delta,A}(T)}, \\ \|\boldsymbol{F}\|_{S^{\gamma}_{\mathbf{\Theta},\delta,A}(T)} &:= \|\boldsymbol{F}\|^{1-\gamma}_{S_{\mathbf{\Theta},\delta,A}(T)}\|\boldsymbol{F}\|^{\gamma}_{L^{\infty}(S_{\mathbf{\Theta},\delta,A})}.\end{aligned}$$

Remark 2.23. *Observe that for $\delta = 1$, $\|\boldsymbol{F}\|_{S_{\mathbf{\Theta},1,A}(T)}$ is comparable with $\|F^1\mathbb{1}_T\|_{L^{\infty}(S^{\infty}_{\mathbf{\Theta},1,A})} + \|F^2\mathbb{1}_T\|_{L^{\infty}(S^2_{\mathbf{\Theta},1,A})}$.*

Strips, outer measures and iterated sizes in time-scale space

We briefly recollect the iterated L^p spaces from [Ura16], which was introduced as a framework to deal with outer L^p embeddings below local L^2. Roughly, one may think of the idea of restricting functions $F(f)$ to strips below, as being related to Calderón-Zygmund decomposition of the underlying function f. In fact, the iterated embedding (see [DPO15] and [Ura16] for its iterated version) consists of applying a refined version of it, which is known as the multi-frequency Calderón-Zygmund decomposition [NOT09].

Definition 2.24 (Strips)**.** *A* strip *at a point $(x, s) \in \mathbb{R}^2_+ = \mathbb{R} \times \mathbb{R}_+$ is given by*

$$D(x, s) := \{(y, \eta, s) \in \mathbb{R}^3_+ \colon |y - x| < s - t\}.$$

We denote the family of all strips by $\mathbb{D}$.

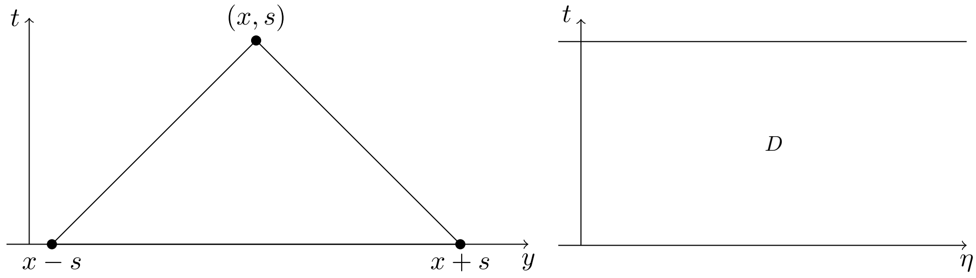

Figure 2.4: The strip $D(x,s) \in \mathbb{D}$.

The outer measure ν is generated by strips via the pre-measure

$$\overline{\nu}\big(D(x,s)\big) := s.$$

Let $\|\cdot\|_{S(T)}$ be a size family indexed by $T \in \mathbb{T}_{\mathbf{\Theta},\delta}$ and μ an outer measure generated by $\mathbb{T}_{\mathbf{\Theta},\delta}$. The iterated Ł^q size family indexed by $D \in \mathbb{D}$ is given for $F \in \mathcal{B}(\mathbb{R}^3_+)$ by

$$\|F\|_{\text{Ł}^q_\mu(S)(D)} := \nu(D)^{-1/q}\|F\mathbb{1}_D\|_{L^q_\mu(S)}.$$

Remark on the embedding theorem

(2.12) is implied by the following.

Proposition 2.25. *Let f be a Schwartz function on $\mathbb{R}$ and $\mathbf{\Theta} = (\Theta, \Theta^{(in)})$. For $1 < p \le \infty$ and $q > \max(p', 2)$ it holds that*

$$\|F(f)\|_{L^p_\nu \text{Ł}^q_{\mu_{\mathbf{\Theta},1}}(S^\infty_{\mathbf{\Theta},1})} \le C_{p,q}\|f\|_{L^p(\mathbb{R})},$$

and moreover

$$\|F^*(f)\|_{L^p_\nu \text{Ł}^q_{\mu_{\mathbf{\Theta},1}}(S_{\mathbf{\Theta},1})} \le C_{p,q}\|f\|_{L^p(\mathbb{R})}.$$

The proof of the above proposition follows along the lines of [DPO15] in the framework of [Ura16], choosing for each $(y,\eta,t) \in \mathbb{R}^3_+$ a wave packet $\varphi \in \Phi$ ($\varphi \in \Phi^*$) that almost attains the supremum in the definition of $F(f)$ ($F^*(f)$, respectively).

Choice of parameters and notation

Here we introduce most of the notation and fix the parameters that we are going to use throughout this chapter.

Recall that throughout this chapter we fix $0 < b < 2^{-8}$. Moreover, we assume that $|\alpha_1| \ll 1$, hence $|\alpha_2|, |\alpha_3| \simeq 1$, $|\beta_1| \simeq 1$. In the following we shall fix three pairs of intervals $\mathbf{\Theta}_i = (\Theta_j, \Theta_j^{(in)})$, $j \in \{1,2,3\}$ to define three collections of $(\mathbf{\Theta},\delta)$-trees and associated outer measure structures on $\mathbb{R}^3_+$ that are compatible with the trilinear form $\Lambda_{\vec{\beta}}$.

Let $\Theta = B_1(0) \subset \mathbb{R}$, δ be as in (2.16) and

$$a_j = \delta\alpha_1^{-1}\alpha_j. \tag{2.21}$$

Note that since we assume $|\alpha_1| \ll 1$, we have $|a_j| \simeq 1$ set

$$\Theta_1 := \Theta + \beta_1, \qquad \Theta_2 = a_2\Theta + \delta\beta_2, \qquad \Theta_3 = a_3\Theta + \delta\beta_3 \tag{2.22}$$

For $j = 1, 2, 3$ fix $\Theta_j^{(in)}$ such that $\Theta_i \supset \Theta_i^{(in)} \supset B_b(0)$ and so that for $j = 2, 3$

$$a_j^{-1}(\Theta_j^{(in)} - \delta\beta_j) + \beta_1 \subset \Theta_1^{(out)} = \Theta_1 \setminus \Theta_1^{(in)}.$$

This can be done uniformly in $\vec{\beta}$ using the assumption $|\beta_1|, |\beta_2|, |\beta_3|, |\alpha_2|, |\alpha_3| \approx 1$. It is sufficient to have $b < 2^{-8} \ll |\beta_1|$, and to set $\Theta_j^{(in)} = B_b(0)$ for $j = 1, 2, 3$.

In this chapter we restrict ourselves to a discretized collection of trees and strips in $\mathbb{R}^3_+ = \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$ and $\mathbb{R}^2_+ = \mathbb{R} \times \mathbb{R}_+$. Let

$$\Delta\mathbb{R}^3_+ := \left\{(4^k n, \xi, 4^k) \in \mathbb{R}^3_+ \colon n \in \mathbb{Z},\ \xi \in \mathbb{R},\ k \in \mathbb{Z}\right\}$$

and $\Delta\mathbb{T}_{\boldsymbol{\Theta},\delta}$ be the set of trees $T_{\boldsymbol{\Theta},\delta}(x, \xi, s) \in \mathbb{T}_{\boldsymbol{\Theta},\delta}$ with $(x, \xi, s) \in \Delta\mathbb{R}^3$. From now on we denote

$$\mathbb{T} = \Delta\mathbb{T}_{\boldsymbol{\Theta},1}, \qquad \mathbb{T}_\delta = \Delta\mathbb{T}_{\boldsymbol{\Theta},\delta},$$

$$\mathbb{T}^j = \begin{cases} \Delta\mathbb{T}_{\boldsymbol{\Theta}_1,1} & j = 1 \\ \Delta\mathbb{T}_{\boldsymbol{\Theta}_j,\delta} & j \in \{2, 3\}. \end{cases}$$

Moreover let

$$\Delta\mathbb{R}^2_+ := \{(4^k n, 4^k) \in \mathbb{R}^2_+ \colon k, n \in \mathbb{Z}\}$$

and let $\Delta\mathbb{D}$ be the set of strips $D(x, s)$ with $(x, s) \in \Delta\mathbb{R}^2_+$. From now on we overload the notation and set

$$\mathbb{D} := \Delta\mathbb{D}.$$

For $\varphi \in \Phi$ and $j = 1, 2, 3$ we set

$$F_j^\varphi = F^\varphi(f_j), \qquad F_j = F(f_j), \qquad F_j^* = F^*(f_j), \qquad \boldsymbol{F}_j = (F_j, F_j^*).$$

Let us denote the measures and sizes restricted to the discretized collection with the preceding Δ. In order to ease the notation, we set for $j = 1, 2, 3$

$$S_j := \Delta S_{\boldsymbol{\Theta}_j,1}, \qquad \mu_j = \Delta\mu_{\boldsymbol{\Theta}_j,1},$$

and for $j = 2, 3$ we set

$$S_{j,\delta,A} := \Delta S_{\boldsymbol{\Theta}_j,\delta,A}, \qquad \mu_{j,\delta} := \Delta\mu_{\Theta_j,\delta}.$$

Moreover, for $j = 1, 2, 3$ we set

$$\|\boldsymbol{F}_j\|_{L^p Ł^q(S)} := \|F_j\|_{L^p_\nu Ł^q_{\mu_{\boldsymbol{\Theta}_j,1}}(S^\infty_{\boldsymbol{\Theta}_j,1})} + \|F_j^*\|_{L^p_\nu Ł^q_{\mu_{\boldsymbol{\Theta}_j,1}}(S^2_{\boldsymbol{\Theta}_j,1})}.$$

For technical reasons, in order to work on a compact set in the time-scale direction, we define ε-dependent sets for $\varepsilon > 0$. We assume throughout that the ε is a very small number such that $\varepsilon^{-1} \in \mathbb{N}$, so that $D_\varepsilon \in \mathbb{D}$ and $W_\varepsilon \in \mathbb{D}^\cup$.

$$A_\varepsilon := D_\varepsilon \setminus W_\varepsilon, \tag{2.23}$$

where

$$D_\varepsilon = D(0, 4^{\varepsilon^{-1}}), \qquad W_\varepsilon = \bigcup_{j=-4^{2\varepsilon^{-1}}}^{4^{2\varepsilon^{-1}}} D(j4^{-\varepsilon^{-1}}, 4^{-\varepsilon^{-1}}).$$

We shall restrict the functions on $\mathbb{R}^+_3$ to A_ε; all estimates in this chapter will be independent of the ε and thus by standard limiting argument as $\varepsilon \to 0$, we recover the full estimates.

Geometry of discretized trees

In this part we study pairwise disjoint decompositions of the discretized collection of trees and strips.

In order to ease the notation in the following proofs, we define the three families of dyadic intervals

$$\mathbb{I}_j := \{[4^k n - 4^k, 4^k n + 4^k) : k, n \in \mathbb{Z},\ n \equiv j \ (\mathrm{mod}\ 3)\}. \tag{2.24}$$

Next four lemmata are simple observations, which imply that up to a finite overlap, one can think of trees and strips as coming from the standard dyadic grid.

Lemma 2.26. *For any $j \in \{0,1,2\}$ and any two $I, I' \in \mathbb{I}_j$, I and I' are either disjoint or one is contained in another. Moreover, if $T(x,\xi,s) \in \mathbb{T}$, then $I_T \in \mathbb{I}_j$ for exactly one $j \in \{0,1,2\}$.*

Proof. Follows easily from the definition of $\mathbb{I}_j$ in (2.24), using the fact that for $k \in N$, $4^k \equiv 1$ modulo 3. □

Lemma 2.27. *There exists a decomposition $\mathbb{T} = \mathbb{T}^0_\Delta \cup \mathbb{T}^1_\Delta \cup \mathbb{T}^2_\Delta$, such that for $j = 0,1,2$ the intervals $\{I_T : T \in \mathbb{T}^j_\Delta\}$ have the following property: for any two of them they are either pairwise disjoint or one is contained in the other.*

Proof. For $j = 0,1,2$, let $\mathbb{T}^j_\Delta \subset \mathbb{T}$ be the set of trees with $I_T \in \mathbb{I}_j$. Together with Lemma 2.26 this gives the desired decomposition. □

Lemma 2.28. *There exists a decomposition $\mathbb{D} = \mathbb{D}^0_\Delta \cup \mathbb{D}^1_\Delta \cup \mathbb{D}^2_\Delta$, such that for $j = 0,1,2$ the intervals $\{I_D : D \in \mathbb{D}^j_\Delta\}$ have the following property: for any two of them they are either pairwise disjoint or one is contained in the other.*

Proof. Exactly the same as the proof of the previous lemma. □

Lemma 2.29. *Let $|\cdot|$ denote the standard Lebesgue measure and let $V \in \mathbb{D}^\cup$. Then there exist $D_m \in \mathbb{D}$ for $m = 1,2,..$, such that $V = \bigcup_{m=1}^\infty D_m$ and for any measurable $A \subset \mathbb{R}$*

$$\sum_{m=1}^\infty |I_{D_m} \cap A| \lesssim |I_V \cap A|.$$

Proof. Let $V = \bigcup \mathcal{V}$, where $\mathcal{V} \subset \mathbb{D}$. Let $\mathcal{V}_j$ be the family of maximal strips in $\mathcal{V}$ which are elements of $\mathbb{D}^j_\Delta$ in Lemma 2.28 for $j = 0,1,2$. By maximality we have $V = \bigcup_{j=0}^2 \bigcup \mathcal{V}_j$. Moreover $D \in \mathcal{V}_j$ are pairwise disjoint for $j = 0,1,2$ by Lemma 2.29, hence

$$\sum_{D \in \mathcal{V}_j} |I_D \cap A| \le |I_V \cap A|.$$

Putting these three collections together we obtain the desired result. □

Geometry of the trilinear form

In this part we investigate the interplay of the discretized trees with the trilinear form Λ. We introduce the maps that capture the geometry of the trilinear form, which is expressed in the form of the so-called transfer properties below. Let

$$\pi_1(y,\eta,t) = (y, \eta + \beta_1 t^{-1}, t), \qquad \pi_j(y,\eta,t) = (y, a_j\eta + \delta\beta_j t^{-1}, t), \quad \text{for } j = 2,3. \tag{2.25}$$

Moreover for $T(x,\xi,s) \in \mathbb{T}$ we define

$$\rho(T(x,\xi,s)) = T_\delta(x,\delta^{-1}\xi,s) \in \mathbb{T}_\delta. \tag{2.26}$$

With the above two definitions we have for any $T(x,\xi,s) \in \mathbb{T}$

$$\pi_1(T(x,\xi,s)) = T^1(x,\xi,s) \in \mathbb{T}^1, \qquad \pi_j(\rho(T(x,\xi,s)) = T^j(x,\delta^{-1}a_j\xi,s) \in \mathbb{T}^j.$$

We extend π_1 canonically to sets of the form $K \cap M \setminus L$ for $K,L,M \in \mathbb{T}^\cup$ and π_j canonically to sets of the form $K \cap M \setminus L$ for $K,L,M \in \mathbb{T}^\cup_\delta$ and $j=2,3$. By definition of Λ, (2.17) and (2.25), we easily obtain the following fact, which we call the transfer properties. It lets us move the characteristic functions of subsets of $\mathbb{R}^3_+$ between the functions in the trilinear form, under appropriate assumptions.

Lemma 2.30 (Transfer properties)**.**

1. *Let $E = K \setminus L$ with $K,L \in \mathbb{T}^\cup_\delta$ and let $A = V \setminus W$ for $V,W \in \mathbb{D}^\cup$. We have*
$$\Lambda(G_1, G_2 \mathbb{1}_{\pi_2(E)} \mathbb{1}_A, G_3) = \Lambda(G_1, G_2, G_3 \mathbb{1}_{\pi_3(E)} \mathbb{1}_A), \tag{2.27}$$
2. *Let $K \in \mathbb{T}^\cup$, where $K = \bigcup_i T_i$. We have*
$$\Lambda(G_1 \mathbb{1}_{\pi_1(\bigcup_i T_i)}, G_2, G_3) = \Lambda(G_1 \mathbb{1}_{\pi_1(\bigcup_i T_i)}, G_2 \mathbb{1}_{\pi_2(\bigcup_i \rho(T_i))}, G_3 \mathbb{1}_{\pi_3(\bigcup_i \rho(T_i))}). \tag{2.28}$$
3. *Let $V \in \mathbb{D}^\cup$. We have*
$$\Lambda(G_1 \mathbb{1}_V, G_2, G_3) = \Lambda(G_1 \mathbb{1}_V, G_2 \mathbb{1}_V, G_3 \mathbb{1}_V). \tag{2.29}$$
4. *Moreover if $K \in \mathbb{T}^\cup_\delta$, then there exists $\widetilde{K} \in \mathbb{T}^\cup$ with $\mu(\widetilde{K}) \lesssim \mu_\delta(K)$, such that*
$$\Lambda(G_1, G_2 \mathbb{1}_{\pi_2(K)}, G_3) = \Lambda(G_1 \mathbb{1}_{\pi_1(\widetilde{K})}, G_2 \mathbb{1}_{\pi_2(K)}, G_3). \tag{2.30}$$

Proof. (2.27): follows from the definition (2.25).

Set $h_j(y,\eta,t) = (y, \delta^{-1}(a_j\eta + \delta\beta_j t^{-1}), \delta t)$ for $j=2,3$. Note that by the definition (2.17) we have for any tree $T \in \mathbb{T}$

$$\Lambda(G_1 \mathbb{1}_{\pi_1(T)}, G_2, G_3) = \Lambda(G_1, G_2 \mathbb{1}_{h_2(T)}, G_3) = \Lambda(G_1, G_2, G_3 \mathbb{1}_{h_3(T)}). \tag{2.31}$$

Now, let us prove (2.28): Observe that for $T \in \mathbb{T}$, we have $\pi_2(\rho(T)) \supset h(T)$. Together with (2.31) and (2.27) it finishes the proof for $K = T$. We extend it canonically to unions of trees.

(2.29): Observe that for any $V \in \mathbb{D}^\cup$, there exists a countable collection of trees $\{T_i\}$, such that $V = \pi_1(\bigcup_i T_i)$. Then, the property follows from (2.28), since for $j=2,3$, $\pi_j(\bigcup_i \rho(T_i)) \subset \pi_j(\bigcup_i T_i) = V$.

(2.30): Observe that for any $T(x,\xi,s) \in \mathbb{T}$ there exists $n \in \mathbb{Z}$, such that setting $\widetilde{x} = 4sn$ and $x_j = \widetilde{x} + j4s$ for $j = -2,..,2$ and $\xi_k = \xi + ks^{-1}$ for $k = -4,..,4$, we have $T(x_j, \xi_k, 4s) \in \mathbb{T}$ for $j = 1,2,3$ and moreover it holds that

$$\pi_2(\rho(T)) \subset \bigcup_{j=-2}^{2} \bigcup_{k=-4}^{4} h(T(x_j,\xi_k,4s)).$$

Applying (2.31) we have

$$\begin{aligned}
&\Lambda(G_1 \mathbb{1}_{\pi_1(\bigcup_{j=-2}^2 \bigcup_{k=-4}^4 T(x_j,\xi_k,4s))}, G_2 \mathbb{1}_{\pi_2(\rho(T))}, G_3) \\
&= \Lambda(G_1, G_2 \mathbb{1}_{\bigcup_{j=-2}^2 \bigcup_{k=-4}^4 h(T(x_j,\xi_k,4s))} \mathbb{1}_{\pi_2(\rho(T))}, G_3) \\
&= \Lambda(G_1, G_2 \mathbb{1}_{\pi_2(\rho(T))}, G_3).
\end{aligned}$$

We extend the result for unions $K \in \mathbb{T}^\cup$ analogously. ∎

2.4 Inequalities for outer L^p spaces on $\mathbb{R}_3^+$

In this section we prove several outer L^p inequalities for $F \in \mathcal{B}(\mathbb{R}^3_+)$, which are shown via mostly geometric arguments. We shall exploit them in Section 2.5. Without loss of generality, let $\Theta = (-1,1)$, $\Theta^{(in)} = (-b,b)$ and $\boldsymbol{\Theta} = (\Theta, \Theta^{(in)})$. The constants in general may depend on $\boldsymbol{\Theta}$, however we shall need the result of this section for a finite set of $\boldsymbol{\Theta}$'s, so it does not cause any problem. Let us fix the parameter $0 < \delta \leq 1$. In this subsection we set $\mathbb{T} := \mathbb{T}_{\boldsymbol{\Theta},1}$, $\mu := \mu_{\boldsymbol{\Theta},1}$ (recall Section 2.3.2 for notation) and

$$\|\cdot\|_{S^2} := \|\cdot\|_{S^2_{\boldsymbol{\Theta},1}}, \qquad \|\cdot\|_{S^\infty} := \|\cdot\|_{S^\infty_{\boldsymbol{\Theta},1}}, \qquad \|\cdot\|_S := \|\cdot\|_{S^2} + \|\cdot\|_{S^\infty}.$$

Moreover, we set $\mathbb{T}_\delta := \Delta\mathbb{T}_{\boldsymbol{\Theta},\delta}$, $\mu_\delta := \Delta\mu_{\boldsymbol{\Theta},\delta}$ (Δ stands for discretized) and

$$\|\cdot\|_{S^2_{\delta,A}} := \|\cdot\|_{S^2_{\boldsymbol{\Theta},\delta,A}}, \qquad \|\cdot\|_{R_{\delta,A}} := \|\cdot\|_{R_{\boldsymbol{\Theta},\delta,A}},$$
$$\|\cdot\|_{\delta^{1/2}S^\infty_{1,A}} := \|\cdot\|_{\delta^{1/2}S^\infty_{\boldsymbol{\Theta},\delta,A}}, \qquad \|\cdot\|_{\delta^{1/2}S^2_{1,A}} := \|\cdot\|_{\delta^{1/2}S^2_{\boldsymbol{\Theta},1,A}}.$$

2.4.1 Outer L^p domination on $\mathbb{R}^3_+$

The main result of this subsection are the following three propositions that let us control the contribution of δ dependent sizes $\|\boldsymbol{F}\|_{S_{\delta,A}}$ in terms of $\|\boldsymbol{F}\|_{S_A}$. The main advantage of this fact is that we can use the iterated embedding theorem for the δ independent sizes, Proposition 2.25.

$$A = A_\varepsilon \cap \bigcap_{j=1}^m V_j \setminus W_j, \tag{2.32}$$

where $V_j, W_j \in \mathbb{D}^\cup$ and A_ε is as in (2.23). Let $\mathcal{B}_0(\mathbb{R}^3_+)$ be the set of functions F, such that for any $t^-, t^+ \in \mathbb{R}_+$

$$\lim_{\xi\to\infty} \sup_{(y,\eta,t)\in\mathbb{R}\times(-\xi,\xi)^c\times(t^-,t^+)} |F(y,\eta,t)| = 0. \tag{2.33}$$

This technical assumption as well as restricting to A_ε are needed only for the definition of the selection algorithms below. Moreover, observe that (2.33) is satisfied for embedded functions $F = F(f)$, where f is a Schwartz function, to which we are going to apply the results of this section.

All the constants in this section will be independent of δ and A, unless explicitly stated.

Proposition 2.31. *The following inequality holds for $0 < \delta \leq 1$, $2 < p \leq \infty$ and any $\boldsymbol{F} = (F^1, F^2)$, where $F^1, F^2 \in C(\mathbb{R}^3_+) \cap \mathcal{B}_0(\mathbb{R}^3_+)$*

$$\|\boldsymbol{F}\|_{L^p_{\mu_\delta}(S_{\delta,A})} \lesssim_p \|\boldsymbol{F}\|_{L^p_\mu(S_{1,A})}.$$

By an application of Lemma 2.11, Proposition 2.32 follows from the following three. The first proposition dominates the outer L^p norm of the boundary size $R_{\delta,A}$ by the outer L^p generated by S. Note that in the statement we require $F \in C(\mathbb{R}^3_+)$, which is satisfied for functions we apply this proposition to, see Remark 2.21.

Proposition 2.32. *The following inequality holds for $0 < \delta \leq 1$, $2 < p \leq \infty$ and any $F \in C(\mathbb{R}^3_+)$*

$$\|F\|_{L^p_{\mu_\delta}(R^2_{\delta,A})} \lesssim_p \|F\mathbb{1}_A\|_{L^p_\mu(S^\infty)},$$

together with the weak bound

$$\|F\|_{L^{2,\infty}_{\mu}(R^2_{\delta,A})} \lesssim \|F\mathbb{1}_A\|_{L^2_\mu(S^\infty)},$$

with constants independent of A and δ.

The second proposition dominates the S^2 portion $\|F\|_{L^p_{\mu_\delta}(S^2_{\delta,A})}$ by $\|F\|_{L^2_\mu(S)}$.

Proposition 2.33. *The following inequality holds for $0<\delta\le 1$, $2<p\le\infty$ and any $F\in\mathcal{B}_0(\mathbb{R}^3_+)$*

$$\|F\|_{L^p_{\mu_\delta}(S^2_{\delta,A})} \lesssim_p \|F\mathbb{1}_A\|_{L^p_\mu(S)},$$

together with the weak bound

$$\|F\|_{L^{2,\infty}_{\mu}(S^2_{\delta,A})} \lesssim \|F\mathbb{1}_A\|_{L^2_\mu(S)},$$

with constants independent of A and δ.

The third proposition dominates the S^∞ or S^2 portion of the size multiplied by the small factor $\delta^{1/2}$, which compensates for the measure μ_δ.

Proposition 2.34. *Let $U=S^2$ or $U=S^\infty$. For $0<\delta\le 1$, $2<p\le\infty$ and any $F\in\mathcal{B}(\mathbb{R}^3_+)$*

$$\|F\|_{L^p_{\mu_\delta}(\delta^{1/2}U)} \lesssim_p \|F\|_{L^p_\mu(U)}.$$

Remark 2.35. *Applying the above lemma to $F\mathbb{1}_A$ one immediately obtains*

$$\|F\|_{L^p_{\mu_{\delta,A}}(\delta^{1/2}U_{1,A})} \lesssim_p \|F\mathbb{1}_A\|_{L^p_\mu(U)}.$$

First we prove Proposition 2.32 and Proposition 2.33, and then we prove Proposition 2.34 at the end of this subsection. By standard limiting procedure we may also assume that

$$\|F\mathbb{1}_A\|_{L^2(S)} < \infty. \tag{2.34}$$

The key ingredient of the proof of Proposition 2.32 and Proposition 2.33 is a Bessel type inequality which can be thought of as $\|F\|_{L^2(E)} \le C\|F\mathbb{1}_E\|_{L^2(S)}$ with a constant uniform in the degeneration parameter δ and A, where E is typically a set selected during a selection algorithm in the upper half space; below we introduce two such algorithms that will be useful in this context. We record that such sets E are sometimes called strongly disjoint. Moreover, note that the left hand side of the inequality is the L^2 norm over a possibly singular set E, that has to be understood properly.

Our arguments depend mostly on the geometry of $\mathbb{R}^3_+$ and $\mathbb{T}$. Here are the two key lemmata. The first one estimates the contribution from the boundary under an appropriate geometric assumption.

Lemma 2.36. *Let $\{T_j\}_{j=1}^m$ be a collection of trees with $T_j\in\mathbb{T}_{\delta_j}$ for $\delta/4<\delta_j\le\delta$ and $E\subset\mathbb{R}^3_+$ be such that for any convex tree ΔT, there are at most C indices in $\{1,...,m\}$ such that $\partial^{\theta_j}_{A,j} := \partial^{\theta_j}_{A,T_j}\cap E$ intersects $\Delta T\cap(\{y\}\times\mathbb{R}\times\mathbb{R}_+)$. Then*

$$\sum_{j=1}^{M} \mu_{\delta_j}(T_j)\|F\mathbb{1}_E\|^2_{R^{\theta_j}_{\delta_j,A}(T_j)} \lesssim \|F\mathbb{1}_A\|^2_{L^2_\mu(S^\infty)}.$$

Proof. Using Lemma 2.8 we decompose $\bigcup_k \bigcup_{\Delta T \in \mathcal{T}_k} \Delta T$ for $F\mathbb{1}_A$, where each ΔT is a convex tree. Now, we have

$$\begin{aligned}
&\sum_{j=1}^{M} \mu_{\delta_j}(T_j) \|F\mathbb{1}_E\|^2_{R^{\theta_j}_{\delta_j,A}(T_j)} \\
&\leq \sum_k \sum_{\Delta T \in \mathcal{T}_k} \sum_{\substack{1 \leq j \leq M: \\ \partial^{\theta_j}_{A,j} \cap \Delta T \neq \emptyset}} \mu_{\delta_j}(T_j) \|F\mathbb{1}_E \mathbb{1}_{\Delta T}\|^2_{R^{\theta_j}_{\delta_j,A}(T_j)} \\
&\leq C \sum_k \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F\mathbb{1}_{\Delta T} \mathbb{1}_E \mathbb{1}_A\|^2_{L^\infty(S^\infty)} \\
&\lesssim C \|F\mathbb{1}_E \mathbb{1}_A\|^2_{L^2_\mu(S^\infty)} \leq C \|F\mathbb{1}_A\|^2_{L^2_\mu(S^\infty)},
\end{aligned}$$

where we used that the integral $\mu_{\delta_j}(T_j)\|F\mathbb{1}_{\Delta T}\|^2_{R_{\delta_j,A}(T_j)}$ is over a subset of $\mathbb{R}$ whose measure does not exceed $\mu(\Delta T)$. □

The next lemma controls the contribution from the S^2 portion.

Lemma 2.37. *Let $\{T_j\}_{j=1}^m$ be a collection of trees and $E \subset \mathbb{R}^3_+$ such that there exists a constant $C > 0$ for any $y \in \mathbb{R}$ and any convex tree ΔT, the area of $E \cap \Delta T^{(in)} \cap (\{y\} \times \mathbb{R} \times \mathbb{R}_+)$ is bounded by C. Then*

$$\sum_{j=1}^{M} \mu_{\delta,D}(T_j) \|F\mathbb{1}_E\|^2_{S^2_{\delta,A,D}(T_j)} \lesssim \|F\mathbb{1}_A\|^2_{L^2_\mu(S)}.$$

Proof. Using Lemma 2.8 we decompose $\bigcup_k \bigcup_{\Delta T \in \mathcal{T}_k} \Delta T$ for $F\mathbb{1}_A$ and split $E = E^{(out)} \cup E^{(in)}$, where

$$E^{(out)} = E \cap \bigcup_k \bigcup_{\Delta T \in \mathcal{T}_k} \Delta T^{(out)}, \qquad E^{(in)} = E \cap \bigcup_k \bigcup_{\Delta T \in \mathcal{T}_k} \Delta T^{(in)}$$

First of all, using only pairwise disjointness of ΔT

$$\begin{aligned}
&\sum_{j=1}^{M} \mu_{\delta,D}(T_j) \|F\mathbb{1}_{E^{(out)}}\|^2_{S^2_{\delta,A,D}(T_j)} \\
&\leq \sum_k \sum_{\Delta T \in \mathcal{T}_k} \sum_{j=1}^{M} \mu_{\delta,D}(T_j) \|F\mathbb{1}_{E^{(out)}} \mathbb{1}_{\Delta T}\|^2_{S^2_{\delta,A,D}(T_j)} \\
&\lesssim \sum_k \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F\mathbb{1}_{\Delta T} \mathbb{1}_E \mathbb{1}_A\|^2_{L^\infty(S^2)} \\
&\lesssim \|F\mathbb{1}_E \mathbb{1}_A\|^2_{L^2_\mu(S^2)} \leq \|F\mathbb{1}_A\|^2_{L^2_\mu(S^2)}.
\end{aligned}$$

We are left with estimating the part restricted to $E^{(in)}$. We have

$$\begin{aligned}
&\sum_{j=1}^{M} \mu_{\delta,D}(T_j) \|F\mathbb{1}_{E^{(in)}}\|^2_{S^2_{\delta,A,D}(T_j)} \\
&\lesssim \sum_k \sum_{\Delta T \in \mathcal{T}_k} \sum_{\substack{0 \leq j \leq M \\ T_j^{(out)} \cap \Delta T^{(in)} \neq \emptyset}} \mu_{\delta,D}(T_j) \|F\mathbb{1}_E \mathbb{1}_{\Delta T^{(in)}}\|^2_{S^2_{\delta,A,D}(T_j)}
\end{aligned}$$

$$\lesssim \sum_k \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F\mathbb{1}_A\|^2_{L^\infty(S^\infty)} \sup_{y\in\mathbb{R}} |E \cap \Delta T^{(in)} \cap (\{y\} \times \mathbb{R} \times \mathbb{R}_+)|$$
$$\lesssim \sum_k \sum_{\Delta T \in \mathcal{T}_k} \mu(\Delta T) \|F\mathbb{1}_A\|^2_{L^\infty(S^\infty)} \lesssim \|F\mathbb{1}_A\|^2_{L^2_\mu(S^\infty)}.$$

□

In the proofs of Proposition 2.32 and Proposition 2.33 we introduce auxiliary selection algorithms, one for $R_{\delta,A}$ and one for $S^2_{\delta,A}$, respectively, which we use to carefully decompose the level sets of the corresponding size. Similar procedures are usually used in the context of proving embedding theorems, see for example [LT99], [DT15]. The generated collections of forests have intrinsically very nice geometric properties, so that they can be controlled by the previous two lemmata.

R-selection algorithm and Proof of Proposition 2.32

First we introduce the selection algorithm for $\|\cdot\|_{R_{\delta,A}}$ and bound the contribution of the selected set.

Definition 2.38 (R-selection algorithm)**.** *Let initially $E_0 = \emptyset$ and $\mathbb{X}_0 = \mathbb{R}^3_+$. In the n-th step of the procedure for $n \geq 0$ we proceed as follows: if there exists $T = T(x, \xi, s) \in \mathbb{T}_\delta$ with*

$$\|F\mathbb{1}_{\mathbb{X}_n}\|_{R_{\delta,A}(T)} > \lambda, \tag{2.35}$$

then we choose a $T(x, \xi, s) \in \mathbb{T}_D$ with maximal s. This is possible, since, by Lemma 2.36 applied with $E = \mathbb{R}^3_+$ and a single tree and (2.34), there exists an upper bound for s for trees satisfying (2.35) dependent on λ; moreover all possible s come from a discrete set and thus we can choose $T(x, \xi, s)$ with maximal s.

We then set $E_{n+1} := E_n \cup \partial^{\theta_n}_{A,n}$, where $\partial^{\theta_n}_{A,n} := \partial^{\theta_n}_{A,\widetilde{T}} \cap \mathbb{X}_n$ with $\widetilde{T} = \widetilde{T}(x, \xi, s) \in \mathbb{T}_{\widetilde{\delta}}$, $\delta/4 < \widetilde{\delta} \leq \delta$ and θ_n all chosen such that

$$\|F\mathbb{1}_{\mathbb{X}_n}\|_{R^{\theta_n}_{\delta,A}(\widetilde{T})} > \lambda$$

Next, we set $T_n := T$, $\widetilde{T}_n := \widetilde{T}$, $\delta_n := \widetilde{\delta}$, $\mathbb{X}_{n+1} := \mathbb{X}_n \setminus T_n$ and iterate the procedure. It will terminate, since we have a lower bound for all possible s (because we are restricted to A_ε) and because we have an upper bound for the sum of measures of the selected trees (by Lemma 2.40). Let M be the number of the last iteration. Then, we clearly have $\|F\mathbb{1}_{\mathbb{X}_M}\|_{L^\infty(R_{\delta,A})} \leq \lambda$.

Remark 2.39. *Analogously, one may define the selection algorithm for $\|\cdot\|_{S^\infty}$, which is essentially equivalent to the above $\delta = 1$. Setting $\|\cdot\|_{R_{1,A}} := \|\cdot\|_{S^\infty}$ all the related bounds below continue to hold.*

We have the following bound for the L^2 norm over the picked boundaries.

Lemma 2.40. *Let $\{T_j, \widetilde{T}_j\}_{j=1}^M$ be the trees selected during the algorithm given in Definition 2.38 together with selected $\{\theta_j\}_{j=1}^M$ and $E := E_M$. We have*

$$\sum_{j=1}^M \mu_{\delta_j}(\widetilde{T}_j) \|F\mathbb{1}_E\|^2_{R^{\theta_j}_{\delta_j,A}(\widetilde{T}_j)} \lesssim \|F\mathbb{1}_A\|^2_{L^2_\mu(S^\infty)}.$$

We now give a proof of Proposition 2.32 assuming Lemma 2.40.

Proof of Proposition 2.32. We prove the strong L^∞ bound, weak L^2 bound and interpolate. The L^∞ bound follows since for any tree $T \in \mathbb{T}_\delta$ have $\|F\|_{R_{\delta,A}(T)} \leq \|F\|_{L^\infty(S^\infty)}$. In order to prove the weak L^2 bound we choose any $\lambda > 0$, run the selection algorithm in Definition 2.38, obtaining a collection of trees $\{T_j\}_{j=1}^M$ and E_λ. Then observe that applying Proposition 2.40 in the second inequality we obtain

$$\lambda^2 \mu_\delta(\|F\|_{R_{\delta,A}} > \lambda) \lesssim \sum_{j=1}^{M} \mu_\delta(T_j) \|F \mathbb{1}_{E_\lambda}\|^2_{R^{\theta_j}_{\delta,A}(T_j)} \lesssim \|F \mathbb{1}_A\|^2_{L^2_\mu(S^\infty)}.$$

□

Proof of Lemma 2.40. The proof follows from the next lemma together with Lemma 2.36.

Lemma 2.41. *For $y \in \mathbb{R}$ and for a given convex tree ΔT, the selection algorithm from Definition 2.38 yields at most 25 indices $1 \leq j \leq M$, such that $\partial^{\theta_j}_{A,j}$ intersects $\Delta T \cap (\{y\} \times \mathbb{R} \times \mathbb{R}_+)$.*

Proof. Suppose that $\Delta T = \Delta T(0, 0, 1)$ and fix $y \in \mathbb{R}$. Let us reenumerate the given trees so that $T_j = T(x_j, \xi_j, s_j)$ for $j = 1, 2, ..., N$ (in this order) are such that $\partial^{\theta_j}_{A,j} \cap \Delta T \neq \emptyset$. Let t_j be the largest scale in $\partial^{\theta_j}_{A,j} \cap \Delta T$, i.e. such that

$$\partial^{\theta_j,+}_{A,j} \cap \Delta T = \{(y, \xi_j + \theta_j t_j^{-1}, t_j)\}.$$

Note that we have $t_j \geq t_{j+1}$ for $1 \leq j \leq N-1$. Let for $j = 1, 2, ..., N$, $\xi_j^\pm = \xi_j \pm b t_N^{-1}$. Observe that it holds that

$$|\xi_i^+ - \xi_j^+| + |\xi_i^- - \xi_j^-| \geq \frac{1}{2} t_N^{-1}.$$

for $i \neq j$, $1 \leq i, j \leq N$, otherwise due to the way we select trees, it would not be possible that both points $(y, \xi_i + \theta_i t_i^{-1}, t_i)$, $(y, \xi_j + \theta_j t_j^{-1}, t_j)$ belong to E. On the other hand, since $\partial^{\theta_j}_{A,j} \cap \Delta T \neq \emptyset$, we know that for each $1 \leq j \leq N$, so we have $\xi_j \in (-2t_j^{-1}, 2t_j^{-1})$ (because $(y, \xi_j + \theta_j t_j^{-1}, t_j) \in \Delta T$) and hence $\xi_j^\pm \in (-3t_N^{-1}, 3t_N^{-1})$. Let us further reenumerate ξ_j, so that they are in increasing order. The above properties of ξ_j's imply that

$$(N-1) t_N^{-1} \frac{1}{2} \leq \sum_{j=1}^{N-1} |\xi_j^- - \xi_{j+1}^-| + |\xi_j^+ - \xi_{j+1}^+| \leq 12 t_N^{-1} \implies N \leq 25.$$

□

Using the above lemma and Lemma 2.36 we obtain the desired bound. □

S^2-selection algorithm and Proof of Proposition 2.33

Before we introduce the selection algorithm, we shall refine the discretized collection of trees slightly, which will come in handy. Similarly as in [DT15], we consider

$$\mathbb{R}^3_{+,D} := \{(4^k n, 4^{-k} bl, 4^k) \colon k, l, n \in \mathbb{Z}\} \subset \mathbb{R}^3_+.$$

and set $\mathbb{T}_{\delta,D} = \{T(x, \xi, s) \in \mathbb{T}_\delta \colon (x, \xi, s) \in \mathbb{R}^3_{+,D}\}$, moreover let $\mu_{\delta,D}$ be μ_δ restricted to $\mathbb{T}_{\delta,D}$ and let $\|\cdot\|_{S^2_{\delta,A,D}}$ be the size $\|\cdot\|_{S^2_{\delta,A}}$ restricted to $\mathbb{T}_{\delta,D}$. It is correct to restrict ourselves to the discretized collection of tree $\mathbb{T}_{\delta,D}$, since we have the following lemma.

Lemma 2.42. *For $0 < q \leq \infty$ and for any $F \in \mathcal{B}(\mathbb{R}^3_+)$*

$$\|F\|_{L^q_{\mu_\delta}(S^2_{\delta,A})} \lesssim_q \|F\|_{L^q_{\mu_{\delta,D}}(S^2_{\delta,A,D})}.$$

Proof. The proof bases on the following straightforward fact: for any $T(x,\xi,s) \in \mathbb{T}_\delta$, there exist $T_j(x,\xi_j,s) \in \mathbb{T}_{\delta,D}$ for $j = 1,2,3,4$, such that $T^{(out)} \subset \bigcup_{j=1}^4 T_j^{(out)}$. As a direct consequence, for any $\mathbb{T}_\delta$ we have

$$\|F\|^2_{S^2_{\delta,A}(T)} \leq \sum_{j=1}^4 \|F\mathbb{1}_{T_j^{(out)}}\|^2_{S^2_{\delta,A}(T)} \leq \sum_{j=1}^4 \|F\|^2_{S^2_{\delta,A,D}(T_j)} \leq 4\|F\|_{L^\infty(S^2_{\delta,A,D})},$$

and it follows that $\|F\|_{L^\infty(S_{\delta,A})} \leq 4\|F\|_{L^\infty(S_{\delta,A,D})}$, which gives the endpoint $q = \infty$ in the statement of the lemma. Moreover, any $T \in \mathbb{T}_{\delta,D}$ is clearly an element of $\mathbb{T}_\delta$, which implies, together with the endpoint that we just proved, the following inequality for the level sets

$$\mu_\delta(\|F\|_{S^2_{\delta,A}} > 4\lambda) \leq \mu_{\delta,D}(\|F\|_{S^2_{\delta,A,D}} > \lambda).$$

Multiplying the above by λ^{q-1} and integrating over $\lambda \in \mathbb{R}_+$ we obtain the statement in the range $0 < q < \infty$. □

Now, we shall discuss the S^2 selection algorithm and bound the contribution of the trees selected within that procedure. For a tree $T(x,\xi,s) \in \mathbb{T}_D$ let

$$T_+(x,\xi,s) = \{(y,\eta,t) \in T(x,\xi,s) \colon \eta \geq \xi,\ y \in B_{(1-\delta)s}(x)\}\},$$

$$T_-(x,\xi,s) = \{(y,\eta,t) \in T(x,\xi,s) \colon \eta \leq \xi,\ y \in B_{(1-\delta)s}(x)\}.$$

Definition 2.43 (S^2-selection algorithm)**.** *Initially $E_0 = \emptyset$ and $\mathbb{X}_0 := \mathbb{R}^3_+$. In the n-th step of the algorithm for $n \geq 0$, we proceed as follows: if there exists a tree $T(x,\xi,s) \in \mathbb{T}_{\delta,D}$ with*

$$\|F\mathbb{1}_{T_+}\mathbb{1}_{\mathbb{X}_0}\|_{S^2_{\delta,A,D}(T)} > \lambda, \tag{2.36}$$

which maximizes s for the maximal possible value of ξ. This is possible because of the following observations: by (2.33) and restriction to A_ε given in (2.32) there exists and upper bound for admissible ξ's; moreover ξ's come from a discrete set, so we may choose maximal ξ; moreover, by Lemma 2.37 applied with $E = \mathbb{R}^3_+$ and a single tree, there exists an upper bound for s of a tree $T(x,\xi,s)$ which satisfies (2.36); since all s come from a discrete set, we may choose maximal s.

We set $E_n = E_{n-1} \cup (T_+^{(out)} \cap \mathbb{X}_0)$ and set $\mathbb{X}_n := \mathbb{X}_{n-1} \setminus T(x,\xi,s)$. We iterate the procedure until there are no more trees satisfying (2.36). It will terminate, since we have a lower bound for all possible s (because we are restricted to A_ε in (2.32)) and because we have an upper bound for the sum of measures of the selected trees (by Lemma 2.40). Let M be the number of the last iteration. Then we clearly have $\|F\mathbb{1}_{\mathbb{X}_M}\|_{L^\infty(S_{\delta,A,D})} \leq \lambda$.

Analogously we define the selection algorithm for T_-, with the only difference that at every step we select a tree $T(x,\xi,s)$ with minimal ξ.

We bound the contribution of the selected trees in the next lemma.

Lemma 2.44. *Let $\{T_j\}_{j=1}^M$ be the trees selected during the algorithm given in Definition 2.43 with $E := E_M$. We have*

$$\sum_{j=1}^M \mu_{\delta,D}(T_j)\|F\mathbb{1}_E\|^2_{S^2_{\delta,A,D}(T_j)} \lesssim \|F\mathbb{1}_A\|^2_{L^2_\mu(S)}.$$

Proof. Without loss of generality we prove the statement for the selection algorithm for T_+. The proof follows from the following lemma together with Lemma 2.37.

Lemma 2.45. *Fix $y \in \mathbb{R}$ and let ΔT be a convex tree. The area of $E \cap \Delta T^{(in)} \cap (\{y\} \times \mathbb{R} \times \mathbb{R}_+)$ is bounded uniformly in A, ΔT and y.*

Proof. In the proof we abuse the notation and write T_j, S_j instead of T_j^+, S_j^+.

Without loss of generality assume that $\Delta T = T(0,0,s) \setminus K$, where $K \in \mathbb{T}^{\cup}$. Let us set for any selected tree $T_j \in \mathbb{T}_\delta$, selected at j-th step

$$F_{T_j} := \mathbb{X}_j \cap T_j^{(out)} \cap \Delta T^{(in)} \cap (\{y\} \times \mathbb{R} \times \mathbb{R}_+)$$

Let us reenumerate the trees so that $T_j = T(x_j, \xi_j, s_j)$ for $j = 0, 1, ..., M$ are the selected trees (in this order), such that $F_{T_j} \neq \emptyset$.

Observe that for any index $j \in \mathbb{N}$, we have $|F_{T_j}| \lesssim 1$. In order to see this: let t_0 be the top scale such that there exists $\eta \in \mathbb{R}$ with $(y, \eta, t_0) \in F_{T_j}$. Observe that we have:

$$F_{T_j} \subset \{(y, \eta, t) \in \Delta T^{(in)} \colon bt^{-1} - t_0^{-1} < \eta < bt^{-1},\ t \leq t_0\}.$$

Hence

$$|F_{T_j}| \leq |F| \leq \int_0^{t_0} \int_{b - tt_0^{-1} < \theta < b} d\theta \, \frac{dt}{t} \leq \int_0^{t_0} tt_0^{-1} \frac{dt}{t} = 1. \tag{2.37}$$

Let k be the first index (if it exists, otherwise set $k := M$), such that there is no point $(y, \eta, t) \in F_{T_k}$, with $t = s_k$. Note that then all $j > k$, satisfy $F_{T_j} = \emptyset$. Moreover, observe that by Definition 2.43 and since the selected trees are discretized in frequency, the number k is bounded with a constant dependent only on b. Thus, the observation (2.37) concludes the proof of the lemma.

□

□

The proof that Lemma 2.44 implies Proposition 2.33 is analogous to the proof of Proposition 2.32 after using Lemma 2.42, which reduces the matters to the discrete collection of trees $\mathbb{T}_{\delta,D}$.

Proof of Proposition 2.34

In the proof we use the following two observations: first of all, any tree $T \in \mathbb{T}$ satisfies $\mu_\delta(T) = \delta^{-1}\mu(T)$. On the other hand, the factor $\delta^{1/2}$ compensates for this fact, because $\|F\|_{L^\infty(S)}$ is independent of δ. Thus, we obtain the inequality with a constant $\lesssim 1$.

Proof of Proposition 2.34. Let U be one of S^2, S^∞. We prove the statement for $q = \infty$ and for the weak endpoint $q = 2, \infty$, and interpolate. Note that the former follows from $\delta \leq 1$. Concerning the other endpoint we have to argue differently. Observe that for any $T \in \mathbb{T}$ there exists $T_\delta \in \mathbb{T}_\delta$ with $T_\delta \supset T$ and $\mu_\delta(T_\delta) = \delta^{-1}\mu(T)$. Hence, we have for any $\lambda > 0$

$$\mu_\delta(\|F\|_U > \lambda) \lesssim \delta^{-1}\mu(\|F\|_U > \lambda)$$

Applying the above inequality and Chebyshev's inequality for outer L^p spaces we obtain

$$\mu_\delta(\delta^{1/2}\|F\|_U > \lambda) \lesssim \lambda^{-2}(\delta^{1/2})^2\delta^{-1}\|F\|_{L^2_\mu(U)} = \lambda^{-2}\|F\|_{L^2_\mu(U)}.$$

We finish the proof by an application of Proposition 2.9. □

2.4.2 Boundary lemma

We will need the following lemma in the next section, when we control the boundary given by the intersection of unions of trees $\in \mathbb{T}_{\mathbf{\Theta},\delta}$ and unions of strips $\in \mathbb{D}$. The contribution from the trees is estimated by means of $R_{\delta,A}$ and $\delta^{1/2}S^{\infty}_{\delta,A}$ size, where A comes only from strips, which is acceptable by Proposition 2.32 and Proposition 2.34. In the proof we decompose the boundary coming from a union of trees into a finitely overlapping collection of boundaries coming from single trees, where we use that the trees are discretized in space.

Lemma 2.46. *Let $F \in C(\mathbb{R}^3_+)$, $T = T(0,0,s) \in \mathbb{T}_{\mathbf{\Theta},\delta}$ and $C = P \cap A$, where*

$$P = \bigcap_{j=1}^{n} K_j \setminus L, \qquad\qquad A = \bigcap_{j=1}^{m} V_j \setminus W$$

where $K_j, L \in \mathbb{T}^{\cup}_{\mathbf{\Theta},\delta}$ and $V_j, W \in \mathbb{D}^{\cup}$. Let $\theta \in \Theta^{(in)}$ and let $t^{\theta;\pm}_{C,T}\colon [-s,s] \to \mathbb{R}_+$ be the parametrization of C^{θ}_{T}. Then

$$\int_{-s}^{s} |F(y, \theta t^{\theta;\pm}_{C,T}(y)^{-1}, t^{\theta;\pm}_{C,T}(y))|^2\, dy \lesssim_n s(\|F\mathbb{1}_C\|^2_{L^\infty(R_{\delta,A})} + \|F\mathbb{1}_C\|^2_{L^\infty(\delta^{1/2}S^{\infty}_{\delta,A})})$$

Proof. We prove the statement for $t^{\theta;-}_{C,T}$, since the argument for $t^{\theta;+}_{C,T}$ is the same, hence from now on we discard writing $\pm$ in the superscripts. Notice that the integral over $\{|y| \geq (1-\delta)s\}$ is bounded by $s\|F\mathbb{1}_C\|^2_{L^\infty(\delta^{1/2}S^\infty)}$. Set

$$\partial^{\theta}_{C,T} = \{(y, \theta t^{\theta}_{C,T}(y)^{-1}, \theta t^{\theta}_{C,T}(y)) : |y| < (1-\delta)s\}.$$

Consider the decomposition of $\partial^{\theta}_{C,T}$

$$\partial^{\theta}_{C,T} = (\partial^{\theta}_{C,T} \cap \partial^{\theta}_{A,T}) \cup (\partial^{\theta}_{C,T} \cap \partial^{\theta}_{P,T}) = O_1 \cup O_2$$

For the first part we have

$$\int_{O_1} |F|^2 \leq s\|F\mathbb{1}_C\|_{R^{\theta}_{\delta,A}(T)} \leq s\|F\mathbb{1}_C\|_{L^\infty(R_{\delta,A})}.$$

Now we shall bound the integral over O_2. Consider the splitting

$$O_2 = (\bigcup_{j=1}^{n} O_2 \cap \partial^{\theta}_{K_j,T}) \cup (O_2 \cap \partial^{\theta}_{L,T}) = \bigcup_{j=1}^{n} \widetilde{O}_j \cup \overline{O}.$$

we consider only $\widetilde{O}_1$, but the proof for different $\widetilde{O}_j$'s and $\overline{O}$ is the same. Set $K := K_1$, $\partial^{\theta}_{K} := \partial^{\theta}_{K,T}$ and note that without loss of generality (up to increasing the constant by factor 3) we may assume that $K \subset \mathbb{T}^0_{\Delta}$ (see Lemma 2.27). Let t^{θ}_{K} be the parametrization of ∂^{θ}_{K}. Let $\partial^{\theta}_{K} = S_1 \cup S_2$, where S_1 is the set of points $(y,\eta,t) \in \partial^{\theta}_{K}$, such that y is a local maximum of the function t^{θ}_{K} and let $S_2 = \partial^{\theta}_{K} \setminus S_1$. Observe that S_1 is a union of pairwise disjoint of "horizontal" segments $\mathcal{I}$ that are parallel to $\{(y,0,0)\colon y \in \mathbb{R}\}$ and S_2 is a union of pairwise disjoint curves $\mathcal{C}$ whose projection onto $\{(y,0,t)\colon y \in \mathbb{R}, t \in \mathbb{R}_+\}$ are segments which form the angle $\pi/4$ or $3\pi/4$ with $\{(y,0,0)\colon y \in \mathbb{R}\}$. For every segment $I \in \mathcal{I}$, we select all minimal (in terms of measure) trees $\widetilde{T} = T_\delta(x,\xi,u) \in \mathbb{T}_\delta$ such that there exists $\widetilde{\delta}$ satisfying: $\delta/4 < \widetilde{\delta} \leq \delta$ and $\overline{T} = T_{\widetilde{\delta}}(x,\xi,u) \in \mathbb{T}_{\widetilde{\delta}}$ satisfies $I \cap \partial^{\theta}_{\overline{T}} \neq \emptyset$. For a fixed I, we denote the collection of selected $\widetilde{T}$ with $\widetilde{\mathcal{T}}_I$ and set $\widetilde{\mathcal{T}} := \bigcup_{I\in\mathcal{I}} \widetilde{\mathcal{T}}_I$. This way we cover all points that belong to S_1 with $\{\partial^{\theta}_{\widetilde{T}}\colon \widetilde{T} \in \widetilde{\mathcal{T}}\}$. Moreover, observe that this collection of boundaries also covers all points belonging to S_2, and therefore the whole set ∂^{θ}_{K}. We have the following.

Lemma 2.47. *Let $I \in \mathcal{I}$. We have*

$$\sum_{\widetilde{T} \in \widetilde{\mathcal{T}}_I} |I_{\widetilde{T}}| \lesssim |I|.$$

Proof. Note that I is of the form $\overline{I} \times \{\eta\} \times \{t\}$ for some $\eta \in \mathbb{R}$, $t \in \mathbb{R}_+$ and an interval $\overline{I} \subset \mathbb{R}$. Since $\mathbb{T}$ is the discretized collection of trees and by the way we defined $I \in \mathcal{I}$, we have that

$$|I| \gtrsim \delta^{-1} t. \tag{2.38}$$

Using Lemma 2.27, it suffices to prove the statement for $\widetilde{\mathcal{T}}_I \subset \mathbb{T}^0_\Delta$, i.e. for pairwise disjoint $\partial^\theta_{\widetilde{T}}$. By definition, each $\widetilde{T}(x, \xi, u) \in \widetilde{\mathcal{T}}_I$ satisfies $\partial^\theta_{\widetilde{T}} = \widetilde{I} \times \{\eta\} \times \{\delta u\}$ for some interval $\widetilde{I} \subset \mathbb{R}$ and $u \in \mathbb{R}_+$ such that $\delta u \simeq t$. Thus,

$$|\partial^\theta_{\widetilde{T}}| \simeq |I_{\widetilde{T}}| = \delta^{-1}(\delta u) \simeq \delta^{-1} t$$

Moreover, $\partial^\theta_{\overline{T}}$ (see the definition of $\overline{T}$) are pairwise disjoint intervals of the same length comparable with $\delta^{-1}t$ like in the previous display, which have nonempty intersection with I. Together with (2.38), that implies the desired inequality. □

Since $I \in \mathcal{I}$ are pairwise disjoint intervals, hence applying the above lemma we obtain that $\sum_{\widetilde{T} \in \widetilde{\mathcal{T}}} |I_{\widetilde{T}}| \lesssim s$. Using this bound, we may further estimate

$$\begin{aligned}\int_{\widetilde{O}} |F|^2 &\le \sum_{\widetilde{T} \in \widetilde{\mathcal{T}}} |I_{\widetilde{T}}| (\|F \mathbb{1}_C\|^2_{R_{\delta,A}(\widetilde{T})} + \|F \mathbb{1}_C\|^2_{\delta^{1/2} S^\infty_{\delta,A}(\widetilde{T})}) \\ &\lesssim s(\|F \mathbb{1}_C\|^2_{L^\infty(R_{\delta,A})} + \|F \mathbb{1}_C\|^2_{L^\infty(\delta^{1/2} S^\infty_{\delta,A})}).\end{aligned}$$

□

2.4.3 Quasi-monotonicity of iterated outer L^p norms on $\mathbb{R}^3_+$

In this subsection we shall state and prove several auxiliary lemmata about the iterated L^p norms, which we will need in the next section. First, we show that $\|\cdot\|_{\L^p}$ sizes are decreasing in p.

Lemma 2.48 (Monotonicity of iterated sizes)**.** *Let and $0 < p \le q \le \infty$. We have for $F \in \mathcal{B}(\mathbb{R}^3_+)$*

$$\|F\|_{\L^q_\mu(S)} \lesssim_{p,q} \|F\|_{\L^p_\mu(S)}.$$

Proof. Recall that $\|F\|_{\L^\infty(S)} = \|F\|_{L^\infty(S)}$. First, assume that for any $p < \infty$

$$\|F\|_{L^\infty(S)} \lesssim_p \|F\|_{\L^p(S)}. \tag{2.39}$$

Using (2.39) we have for $0 < p < q < \infty$ and $\|F\|_{\L^p(S)} < \infty$ (otherwise there is nothing to show), and for any $D \in \mathbb{D}$, by the definition of the outer L^p norm

$$\begin{aligned}&\nu(D)^{-1/q} \|F \mathbb{1}_D\|_{L^q(S)} \lesssim_{p,q} \|F\|^{1-p/q}_{L^\infty(S)} \nu(D)^{-1/q} \|F \mathbb{1}_D\|^{p/q}_{L^p(S)} \\ &\lesssim \|F\|^{1-p/q}_{L^\infty(S)} \|F \mathbb{1}_D\|^{p/q}_{\L^p(S)} \lesssim \|F \mathbb{1}_D\|_{\L^p(S)}.\end{aligned}$$

We still have to show (2.39). First assume that $\|F\|_{L^\infty(S)} < \infty$. Let $T \in \mathbb{T}$ be a tree such that $\|F\mathbb{1}_T\|_{L^\infty(S)} > \|F\|_{L^\infty(S)}/2$. Let $D \in \mathbb{D}$ be such that $I_D = I_T$. We have

$$\begin{aligned}\|F\|_{Ł^p(S)} &\geq \nu(D)^{-1/p}\|F\mathbb{1}_D\|_{L^p(S)}\\ &\gtrsim \nu(D)^{-1/p}\nu(D)^{1/p}\|F\|_{L^\infty(S)} = \|F\|_{L^\infty(S)}.\end{aligned}$$

If $\|F\|_{L^\infty} = \infty$, choosing a tree such that $\|F\mathbb{1}_T\|_{L^\infty} > n$ we similarly obtain $\|F\|_{Ł^p(S)} \gtrsim n$. Since this holds for any $n \in \mathbb{N}$, it implies $\|F\|_{Ł^p(S)} = \infty$. □

The following fact lets us relate the outer L^q norm to the averaged $Ł^q$.

Lemma 2.49. *Let $V \in \mathbb{D}^\cup$. Then for $F \in \mathcal{B}(\mathbb{R}^3_+)$ it holds that*

$$\|F\mathbb{1}_V\|_{L^q_\mu(S)} \lesssim_q \nu(V)^{1/q}\|F\mathbb{1}_V\|_{Ł^q_\mu(S)} \qquad q \in (0,\infty]$$

Proof. Using Lemma 2.29, decompose $V = \bigcup_{m=1}^\infty D_m$, where $D_m \in \mathbb{D}$ for $m = 1, 2, ...$

For $q = \infty$ there is nothing to show. Assume that $0 < q < \infty$. Let $\mathbb{T}_m \subset \mathbb{T}$ for $m = 1, 2, ..$ be such that

$$\sum_{T\in\mathbb{T}_m} \mu(T) \lesssim \mu(\|F\mathbb{1}_{D_m}\|_S > \lambda), \qquad \|F\mathbb{1}_{D_m}\mathbb{1}_{\mathbb{X}\setminus\bigcup\mathbb{T}_m}\|_{L^\infty(S)} \leq \lambda$$

Set $\mathbb{T}' = \bigcup_m \mathbb{T}_m$. Let $T \in \mathbb{T}$ be any tree. For every m, let T_m be the tree, whose top is the maximal tile $P \in D_m \cap T$ (it can be empty). Observe that it holds that

$$\begin{aligned}\|F\mathbb{1}_V\mathbb{1}_{\mathbb{X}\setminus\bigcup\mathbb{T}'}\|^2_{S(T)} &= \frac{1}{\mu(T)}\sum_m \mu(T_m)\|F\mathbb{1}_{D_m}\mathbb{1}_{\mathbb{X}\setminus\bigcup\mathbb{T}'}\|^2_{S(T_m)}\\ &\leq \lambda^2\frac{1}{\mu(T)}\sum_m \mu(T_m) \lesssim \lambda^2,\end{aligned}$$

where we used that (applying Lemma 2.29) $\sum_m \mu(T_m) \leq \mu(T)$. Thus we have

$$\mu(\|F\mathbb{1}_V\|_S > \lambda) \leq \sum_{T\in\mathbb{T}'}\mu(T) \lesssim \sum_m \mu(\|F\mathbb{1}_{D_m}\| > \lambda).$$

In consequence

$$\begin{aligned}\lambda^q\mu(\|F\mathbb{1}_V\|_S > \lambda) &\lesssim \sum_m \lambda^q\mu(\|F\mathbb{1}_{D_m}\|_S > \lambda)\\ &\lesssim \sum_m \nu(D_m)\|F\mathbb{1}_{D_m}\|^q_{Ł^q_\mu(S)} \lesssim \nu(V)\|F\mathbb{1}_V\|^q_{Ł^q_\mu(S)},\end{aligned}$$

where we used that (applying Lemma 2.29) $\sum_m \nu(D_m) \leq \nu(V)$ holds. The full result follows Proposition 2.9. □

The next lemma reverts the inequality in Lemma 2.48 if F is appropriately localized, losing a factor coming from the localization.

Lemma 2.50. *Let $A \subset \mathbb{R}^3_+$ be a Borel set and let $V \in \mathbb{D}^\cup$. Then for any $0 < t \leq q \leq \infty$ and $F \in \mathcal{B}(\mathbb{R}^3_+)$ one has*

$$\|F\mathbb{1}_A\mathbb{1}_V\|_{L^t_\mu(S)} \lesssim_{q,t} \mu(V\cap A)^{1/t-1/q}\nu(V)^{1/q}\|F\mathbb{1}_A\mathbb{1}_V\|_{Ł^q_\mu(S)}. \tag{2.40}$$

Proof. Note that we can assume that $\|F\|_{L^\infty(S)} < \infty$, otherwise the right hand side of (2.40) is infinite and there is nothing to show. Moreover, by scaling invariance we assume that $\|F\mathbb{1}_A\mathbb{1}_V\|_{L^\infty(S)} \leq 1$. Note that for all $\lambda > 0$

$$\mu(\|F\mathbb{1}_A\mathbb{1}_V\|_S > \lambda) \leq \mu(V \cap A).$$

holds. Note that, if $\mu(V \cap A) = 0$ or $\|F\mathbb{1}_A\mathbb{1}_V\|_{L^q_\mu(S)}$, then there is nothing to show. Let $C = \mu(V \cap A)^{-1/q}\|F\mathbb{1}_A\mathbb{1}_V\|_{L^q_\mu(S)}$. We have

$$\begin{aligned}
&\|F\mathbb{1}_A\mathbb{1}_V\|_{L^t_\mu(S)} \\
&\lesssim (\int_0^C \lambda^t \mu(\|F\mathbb{1}_A\mathbb{1}_V\|_S > \lambda)\,\frac{d\lambda}{\lambda})^{1/t} + (\int_C^\infty \lambda^t \mu(\|F\mathbb{1}_A\mathbb{1}_V\|_S > \lambda)\,\frac{d\lambda}{\lambda})^{1/t} \\
&\lesssim C\mu(V \cap A)^{1/t} + C^{1-q/t}\|F\|^{q/t}_{L^q_\mu(S)} \lesssim \mu(V \cap A)^{1/t-1/q}\|F\|_{L^q_\mu(S)}.
\end{aligned}$$

The conclusion follows now from Lemma 2.49. □

The following lemma controls the counting function of a forest coming from the selection algorithm in terms of $Ł^q$ norm.

Lemma 2.51 (Counting function estimates)**.** *Let $V \in \mathbb{D}^\cup$ and let $\mathcal{T}_\lambda$ be the collection of trees selected according to Definition 2.38(Definition 2.43) at a certain level $\lambda > 0$ for function $F\mathbb{1}_V$. Let $N_{\mathcal{T}_\lambda}$ be its counting function. Then for any $1 \leq p < \infty$, $2 \leq q < \infty$ and $F \in \mathcal{B}_c(\mathbb{R}^3_+)$ the following bounds hold*

$$\|N_{\mathcal{T}_\lambda}\|_{L^p} \lesssim_p \nu(V)^{1/p}\lambda^{-q}\|F\mathbb{1}_V\|^q_{Ł^q_\mu(S)}$$

together with the BMO endpoint

$$\|N_{\mathcal{T}_\lambda}\|_{BMO} \lesssim \lambda^{-q}\|F\mathbb{1}_V\|^q_{Ł^q_\mu(S)}.$$

Proof. Let E_λ be the strongly disjoint set selected at level λ for function $F\mathbb{1}_V$.

We show the result for endpoints $p = 1$ and BMO; the full statement follows then from the inequality $\|f\|_{L^p} \lesssim_p \|f\|^{1/p}_{L^1}\|f\|^{1-1/p}_{BMO}$ (see, [CZ05]). First of all, applying Lemma 2.40(Lemma 2.44) we have

$$\|N_{\mathcal{T}_\lambda}\|_{L^1} \lesssim \lambda^{-2}\|F\mathbb{1}_V\|^2_{L^2_\mu(S)}.$$

Applying Lemma 2.50 to the right hand side we obtain

$$\|N_{\mathcal{T}_\lambda}\|_{L^1} \lesssim \lambda^{-2}\|N_{\mathcal{T}_\lambda}\|^{1-2/q}_{L^1}\nu(V)^{2/q}\|F\mathbb{1}_V\|^2_{Ł^q},$$

which gives

$$\|N_{\mathcal{T}_\lambda}\|_{L^1} \lesssim \nu(V)\lambda^{-q}\|F\mathbb{1}_V\|^q_{Ł^q}.$$

We still need to prove the BMO bound. Fix an interval I and let $\mathcal{T}^I_\lambda \subset \mathcal{T}_\lambda$ be the subcollection of selected trees with their top intervals contained in I; moreover let $E^I_\lambda = E_\lambda \cap \bigcup \mathcal{T}^I_\lambda$. Notice that $\mathcal{T}^I_\lambda \subset D_{3I} \cap \mathcal{T}_\lambda$, where $D_{3I} = D(c(I), 3|I|)$ is the strip generated by $3I$. Apply Lemma 2.40(Lemma 2.44) to E^I_λ. We obtain

$$\|N_{\mathcal{T}^I_\lambda}\|_{L^1} \lesssim \lambda^{-2}\|F\mathbb{1}_V\mathbb{1}_{D_{3I}}\|^2_{L^2_\mu(S)}$$

Applying Lemma 2.50, this time using D_I (instead of V) as the time-scale localization, we obtain

$$\|N_{\mathcal{T}_\lambda^I}\|_{L^1} \lesssim |I|\lambda^{-q}\|F\mathbb{1}_V\|^q_{\mathcal{L}^q},$$

where we used that $\nu(D_{3I}) = 3|I|$. Dividing by $|I|$ and taking the supremum on the left hand we get that

$$\|N_{\mathcal{T}_\lambda}\|_{BMO} \lesssim \lambda^{-q}\|F\mathbb{1}_V\|^q_{\mathcal{L}^q}.$$

□

2.5 Trilinear iterated L^p estimate

From now on we fix $\vec{\beta} \in \mathbb{R}^3$ and set $\Lambda := \Lambda_{\vec{\beta}}$. The main result of this section is the following proposition.

Proposition 2.52. *Let* $1 < p_1, p_2, p_3 < \infty$ *with* $\sum_{j=1}^3 1/p_j = 1$ *and* $2 < q_1, q_2, q_3 < \infty$ *with* $\sum_{j=1}^3 1/q_j > 1$. *Let for* $j = 1, 2, 3$, f_j *be a Schwartz function and let* $F_j^\varphi := F^\varphi(f_j)$, $F_j = \sup_{\varphi\in\Phi} F_j^\varphi$, $F_j^* = \sup_{\varphi\in\Phi^*} F_j^\varphi$ *and* $\boldsymbol{F}_j = (F_j, F_j^*)$. *Moreover, assume that* $V_j, W_j \in \mathbb{D}^\cup$ *for* $j = 1, 2, 3$. *Then*

$$|\Lambda(F_1^\varphi\mathbb{1}_{V_1\setminus W_1}, F_2^\varphi\mathbb{1}_{V_2\setminus W_2}, F_3^\varphi\mathbb{1}_{V_3\setminus W_3})| \lesssim \prod_{j=1}^3 \nu(V_j)^{1/p_j}\|\boldsymbol{F}_j\mathbb{1}_{V_j\setminus W_j}\|_{L^\infty\mathcal{L}^{q_j}(S)}. \tag{2.41}$$

Note that in conjunction with Proposition 2.10 the above inequality implies Theorem 2.2. We could not use the outer Hölder inequality from [DT15] for our purpose, since it requires a stronger assumption than we were able to obtain. Namely the outer Hölder inequality would require $\min_j \mu(V_j)$ instead of $\prod_j \mu(V_j)^{1/p_j}$ on the right hand side of (2.41). The other reason is that our multilinear form is nonpositive and, as opposed to [DT15], we do not view it as L^1 norm. Although one could try to deal with the nonpositivity introducing nonpositive sizes to view the left hand side of (2.19) as an outer L^1 norm, it does not seem likely that one can obtain much better gain than $\prod_j \mu(V_j)^{1/p_j}$ in (2.41), since V_1 scales differently than V_2 and V_3.

Before we prove Proposition 2.52 we show a localized estimate at the level of trees.

Proposition 2.53. *Let* $1 \le p_1, p_2, p_3 \le \infty$ *with* $\sum_{j=1}^3 1/p_j = 1$. *Let for* $j = 1, 2, 3$, f_j *be a Schwartz function and let* $F_j^\varphi := F^\varphi(f_j)$. *Assume that* $K_j, L_j, M_j, N_j \in (\mathbb{T}^j)^\cup$, $V_j, W_j \in \mathbb{D}^\cup$ *and* μ_j *is for* $j = 1, 2, 3$. *Moreover, set* $A = A_\varepsilon \cap V_1 \cap \bigcap_{j=2}^3 (V_j \setminus W_j)$, *where* A_ε *is as in (2.23),* $G_j^\varphi := F_j^\varphi\mathbb{1}_{V_j\setminus W_j}\mathbb{1}_{M_j\setminus N_j}$, $G_j = \sup_{\varphi\in\Phi} G_j^\varphi$, $G_j^* = \sup_{\varphi\in\Phi^*} G_j^\varphi$ *and* $\boldsymbol{G}_j = (G_j, G_j^*)$. *Moreover, let* $S_j := S_{j,\delta,A}^\gamma$, $\mu_j = \mu_{j,\delta}$ *for* $j = 2, 3$ *(see, Section 2.3.2). Then*

$$|\Lambda(G_1^\varphi\mathbb{1}_{K_1\setminus L_1}, G_2^\varphi\mathbb{1}_{K_2\setminus L_2}, G_3^\varphi\mathbb{1}_{K_3\setminus L_3})| \lesssim \prod_{j=1}^3 \mu_j(K_j)^{1/p_j}\|\boldsymbol{G}_j\mathbb{1}_{K_j\setminus L_j}\|_{L^\infty(S_j)}.$$

Remark. *Observe that optimizing in* p_j *and* $\mu_j(K_j)$, *we can make* $\prod_{j=1}^3 \mu_j(K_j)^{1/p_j}$ *to be equal* $\min_{j=1,2,3} \mu_j(K_j)$.

Remark. *Observe that* G_j *are additionally restricted to* $M_j \setminus N_j \in (\mathbb{T}^j)^\cup$. *That is because we shall need such additional localization in the proof of Proposition 2.52.*

Note that applying Proposition 2.10, the previous proposition immediately implies

$$|\Lambda(G_1^{\varphi}, G_2^{\varphi}, G_3^{\varphi})| \lesssim \prod_{j=1}^{3} \|G_j\|_{L^{p_j}(S_j)}.$$

In the first subsection we make a couple of remarks concerning boundary integrals, which we shall be using throughout this section. In the second subsection we prove several technical facts, which let us streamline the exposition of the proof of Proposition 2.53. In the second subsection we prove Proposition 2.53. In the third subsection we finally proceed with the proof of Proposition 2.52.

2.5.1 Remark on boundary integrals

In this section we will be using the fundamental theorem of calculus and Green's theorem multiple times, while integrating by parts functions restricted to sets with boundary. In the following two propositions we state the versions of the these facts that we are going to apply. Note that these facts require restriction to a compact set with piecewise linear boundary, which is the technical reason why we restrict functions in $\mathbb{R}^3_+$ to the set A_ε, which we introduced in (2.23).

Proposition 2.54 (Fundamental theorem of calculus)**.** *Let $y \in \mathbb{R}^2_+$ and $a, b, c \in \mathbb{R}_+$ with $0 < a < b$. We have for any $G \in C^1(\mathbb{R}^2_+)$ and $(y,t) \in \mathbb{R}^2_+$*

$$\begin{aligned} G(y,t)\mathbb{1}_{(a,b)}(c) &= \int_0^c \partial_u (G(y,u)\mathbb{1}_{[a,b]}(u))\, du \\ &= \int_0^c \partial_u G(y,u)\mathbb{1}_{(a,b)}(u)\, du + \int_0^c G(y,u)\partial_u \mathbb{1}_{[a,b]}(u)\, du, \end{aligned}$$

with $\partial_u \mathbb{1}_{[a,b]}(u)\, du$ being the measure $\delta_a(u) - \delta_b(u)$.

Let $C \subset \mathbb{R}^2_+$ be a compact region of the form

$$C = \{(y,t) \in \mathbb{R}^2_+ : t^-(y) \le t \le t^+(y)\},$$

where $t^\pm : [a,b] \to \mathbb{R}_+$, where $a < b$, are piecewise linear functions whose Lipschitz constants are bounded by 1 and let us set $C(y) = (t^-(y), t^+(y))$.

Proposition 2.55 (Green's theorem)**.** *Let D be a region bounded by a positively oriented, piecewise linear simple closed curve in $\mathbb{R}^2$. For any $G, H \in C^1(\mathbb{R}^2_+)$ we have*

$$\int_D \partial_y G(y,t)H(y,t)\, dy\, dt = \oint_{\partial D} G(y,t)H(y,t)\, dt - \int_D G(y,t)\partial_y H(y,t)\, dy\, dt.$$

In particular

$$\begin{aligned} &\int\int \partial_y G(y,t)H(y,t)\mathbb{1}_{C(y)}(t)\, dy\, dt \\ &= -\int\int G(y,t)\partial_y H(y,t)\mathbb{1}_{C(y)}(t)\, dy\, dt + \int\int G(y,t)H(y,t)\partial_y \mathbb{1}_{C(y)}(t)\, dy\, dt, \end{aligned}$$

with $\int \partial_y \mathbb{1}_{C(y)}(t)\, dy\, dt$ being $\oint_{\partial C} dt$ as in the statement of classical Green's theorem.

Note that we have the following estimates to control the boundary terms coming from differentiating in space and in scale (in the first and in the second variable), we have

$$|\int_a^b \int_0^s G(y,t)\partial_t \mathbb{1}_{C(y)}(t)\,dt\,dy| \leq \int_a^b |G(y,t^-(y))|\,dy + \int_a^b |G(y,t^+(y))|\,dy, \tag{2.42}$$

and

$$|\int_a^b \int_0^s G(y,t)\partial_y \mathbb{1}_{C(y)}(t)\,dt\,dy| \leq \int_a^b |G(y,t^-(y))|\,dy + \int_a^b |G(y,t^+(y))|\,dy. \tag{2.43}$$

2.5.2 Preliminaries: properties of embeddings in $\mathbb{R}^3_+$

This part is dedicated to proving several technical facts about sizes which will shorten the exposition of the proof of the single tree estimate in the subsequent subsection.

Throughout this subsection, for simplicity we fix $\boldsymbol{\Theta} = (\Theta, \Theta^{(in)})$ with $\Theta = (-1,1)$ and $\Theta^{(in)} = (-b,b)$. The proof for different Θ is analogous. The constants may be dependent on $\boldsymbol{\Theta}$, however in the end we apply the results only to a finite set of parameters introduced in (2.22). We also fix an arbitrary Schwartz function f and a number $s > 0$. Moreover, we set $F^\varphi = F^\varphi(f)$ for $\varphi \in \Phi$, $F = F(f)$, $F^* = F^*(f)$ and $\boldsymbol{F} = (F, F^*)$, see (2.6). All the constants in this subsection are independent of δ, f, unless explicitly stated.

Definition 2.56. *Let $G \in B(\mathbb{R}^2_+)$. Define*

$$\|G\|_V := \|G\|_{V^2} + \|G\|_{V^\infty}, \tag{2.44}$$

where

$$\|G\|_{V^\infty} := \sup_{(y,t)\in\mathbb{R}\times\mathbb{R}_+} |G(y,t)|$$

$$\|G\|_{V^2} := \sup_{T(x,0,w)\in\mathbb{T}} (\frac{1}{w}\int_{x-w}^{x+w}\int_0^w |G(y,t)|^2 \frac{dt}{t}\,dy)^{1/2}.$$

Moreover, define

$$\|G\|_{\boldsymbol{V}^2} := (\frac{1}{s}\int_{-s}^{s}\int_0^{\min(\delta s, s-|y|)} |G(y,t)|^2 \frac{dt}{t}\,dy)^{1/2}. \tag{2.45}$$

Definition 2.57. *Let $G \in C(\mathbb{R}^2_+)$ be a measurable function and*

$$B(y) = (b^-(y), b^+(y)) \subset (0, \min(\delta s, s-|y|)),$$

where $b^\pm\colon [-s,s] \to \mathbb{R}_+$ are piecewise linear, Lipschitz functions. Define

$$\|G\|_{R,\boldsymbol{V}_B} := \sum_{j\in\{-,+\}}^{2} (\int_{-s}^{s} |G(y,b^j(y))|^2\,dy)^{1/2} \tag{2.46}$$

The first lemma of this subsection lets us dominate the S^2 portion of the size over a single θ-dependent hyperbola $\{\theta t^{-1}\}$. Moreover, observe that we convolve f with a function that is only of mean zero and does not necessarily have the Fourier support away from zero. By virtue of this fact, we are able to take the supremum over θ in the proof of the key “overlapping tree estimate”, Lemma 2.77, in the next subsection.

Lemma 2.58.

1. *Let* $T = T(0,0,s) \in \mathbb{T}_{\mathbf{\Theta},\delta}$ *and* $C = (K\setminus L)\cap A_\varepsilon$*, where* $K, L \in \mathscr{T}_{\mathbf{\Theta}}^{\cup}$ *(recall Definition 2.13 and (2.23)). Moreover set* $C^\theta := C_T^\theta$ *with* $\theta \in \Theta^{(in)}$*. Let* $G^\psi(y,t) := f * \psi_t(y) := F^{\overline{\psi}}(y, \theta t^{-1}, t)$ *and additionally assume that* $\hat{\overline{\psi}}_{\theta,1}(\xi) = \xi\hat{\overline{\varphi}}_{\theta,1}(\xi)$*, where* $\overline{\varphi} \in \Phi^*$*. We have*

$$\|G^\psi \mathbb{1}_{C^\theta}\|_{\boldsymbol{V}^2} \lesssim \|\boldsymbol{F}\|_{L^\infty(S_{\delta,C})}. \tag{2.47}$$

$$\|G^\psi \mathbb{1}_{C^\theta}\|_{V} \lesssim \|\boldsymbol{F}\|_{L^\infty(S_{1,C})}. \tag{2.48}$$

2. *Let* $T = T(0,0,s) \in \mathbb{T}_{\mathbf{\Theta},1}$ *and* $B = (K \setminus L) \cap A_\varepsilon$*, where* $K, L \in \mathscr{T}_{\mathbf{\Theta}}^{\cup}$ *(recall Definition 2.13 and (2.23)). Moreover set* $B^\theta := B_T^\theta$ *with* $\theta \in \Theta^{(in)}$*. Let* $G^\varphi(y,t) := f * \varphi_t(y) := F^{\overline{\varphi}}(y, (\theta+\beta_1)t^{-1}, t)$ *and let* $\overline{\varphi} \in \Phi^*$*, i.e.* $\hat{\overline{\varphi}}$ *is supported on* $(-a,a)$*, where* $a \le 2^{-8}b$*, where* $b \le 2^{-8}$*. We have*

$$\|G^\varphi \mathbb{1}_{B^\theta}\|_{V} \lesssim \|\boldsymbol{F}\|_{L^\infty(S_{1,B})}.$$

Remark 2.59. *The restriction of* C *to* A_ε *is used to ensure that the considered scales are bounded from below, so that* $c^\pm$ *are away from zero, where* $(c^-(y), c^+(y)) = C_T^\theta$.

Proof. (1). First we prove (2.47). We set $C^\theta(y) = (c^-(y), c^+(y))$. Note that by the assumptions we have $\psi = \overline{\psi}_{\theta,1}$. Moreover, set $\varphi = \overline{\varphi}_{\theta,1}$ and observe $\hat{\psi}$ is supported on $(-2b, 2b)$ and $\hat{\psi}(\xi) = \xi\hat{\varphi}(\xi)$. Let $a < 2^{-8}b$ be fixed throughout the proof.

Lemma 2.60. *For any* $c > 1$ *we may decompose* ψ *as follows*

$$\psi(x) = \int_1^c \psi_u^{u,-}(x) + \psi_u^{u,+}(x)\, \frac{du}{u^2} + c^{-1}\varphi_c^c(x) \tag{2.49}$$

where for each $u \in [1, +\infty]$ *such that*

$$\psi_u^{u,+}(x) := u^{-1}\psi^{u,+}\big(\frac{x}{u}\big) \quad \psi_u^{u,-}(x) := u^{-1}\psi^{u,-}\big(\frac{x}{u}\big) \quad \varphi_c^c(x) := c^{-1}\varphi^c\big(\frac{x}{c}\big)$$

with

$$\begin{aligned} &supp(\hat{\psi}^{u,-}(\xi+3b)), supp(\hat{\psi}^{u,+}(\xi-3b))) \subset (-a/2, a/2) \\ &e^{i3bx}\psi^{u,-}(x) \in \Phi^*, \quad e^{-i3bx}\psi^{u,+}(x) \in \Phi^*, \quad \varphi^c \in \Phi. \end{aligned} \tag{2.50}$$

In particular for any $0 < t < c$ *it holds that*

$$\psi_t(x) = t\int_t^c \psi_u^{u/t,-}(x) + \psi_u^{u/t,+}(x)\, \frac{du}{u^2} + c^{-1}t\varphi_c^{c/t}(x). \tag{2.51}$$

Proof. Let $\eta \ge 0$ be an even nonzero Schwartz function supported on $(-3b - a/2, -3b + a/2) \cup (3b - a/2, 3b + a/2)$. It follows from the change of variables $u\xi \to u$ that for $\xi \in (-2b, 2b) \setminus \{0\}$

$$1 = \int_1^\infty \overline{\eta}(u\xi)\, \frac{du}{u},$$

where $\overline{\eta}(u) = 2(\int_{-\infty}^\infty \eta(t)t^{-1}dt)^{-1}\eta(u)$. Note that (we discard $2\pi i$, which is irrelevant here)

$$\psi(x) = \int_1^\infty \varphi * (\eta')_u(x)\frac{du}{u^2}.$$

The above holds, since on the Fourier side we have for $\xi \in \mathbb{R} \setminus \{0\}$

$$\hat{\psi}(\xi) = \int_1^\infty \hat{\psi}(\xi)\overline{\eta}(u\xi)\,\frac{du}{u} = \int_1^\infty \hat{\varphi}(\xi)\,u\xi\overline{\eta}(u\xi)\,\frac{du}{u^2},$$

and both functions in questions are mean zero. $\mathcal{F}^{-1}(\hat{\varphi}(\xi)u\xi\overline{\eta}(u\xi)) = \varphi * (\overline{\eta}')_u = \psi_u^u$, where $\psi^u = \psi^{u,-} + \psi^{u,+}$, with some u-dependent $\psi^{u,-}, \psi^{u,+}$ satisfying (2.50). Hence, the only thing left to show is that there exists $\varphi^c \in \Phi$ such that

$$\int_c^\infty \psi_u^u(x)\,\frac{du}{u^2} = c^{-1}\varphi_c^c(x),$$

Changing variables on the left hand side we obtain

$$\int_c^\infty \psi_u^u(x)\,\frac{du}{u^2} = c^{-1}\int_1^\infty \psi_{uc}^{uc}(x)\,\frac{du}{u^2} = c^{-1}\varphi_c^c(x),$$

where $\varphi^c(x) := \int_1^\infty \psi_u^{uc}(x)\,\frac{du}{u^2}$. Note that for every $u > 1$, $\psi^{uc} \in \Phi^*$, meaning that it has uniformly bounded derivatives up to a high order and its support sufficiently small around zero. It is not difficult to check that this implies that $\varphi^c \in \Phi$. This finishes the proof of (2.49). (2.51) follows from applying (2.49) with $\widetilde{c} = ct^{-1}$ and a change of variables. This finishes the proof of the lemma. $\square$

Applying (2.51) with $c = c^+(y)$ we bound $s\|G^\psi \mathbb{1}_{C^\theta}\|_{\boldsymbol{V}^2}$ by the sum of

$$\int_{-s}^{s}\int_{c^-(y)}^{c^+(y)} c^+(y)^{-2}t^2|f * \varphi_{c^+(y)}^{c^+(y)/t}|^2\,\frac{dt}{t}\,dy \tag{2.52}$$

$$\int_{-s}^{s}\int_{c^-(y)}^{c^+(y)} \Big|\int_t^{c^+(y)} f * \psi_u^{u/t}(y)\,\frac{t}{u}\,\frac{du}{u}\Big|^2\,\frac{dt}{t}\,dy \tag{2.53}$$

Note that (2.52) is bounded by

$$\int_{-s}^{s} \sup_{t\in(c^-(y),c^+(y))} |f * \varphi_{c^+(y)}^{c^+(y)/t}|^2\,dy \lesssim s\|F\|^2_{R^\theta_{\delta,C}(T)}.$$

Concerning (2.53), we consider $\varphi_u^{u/t,+}$ since the argument for $\varphi_u^{u/t,-}$ is analogous. Also, let $\psi_u^{y,u} := \psi_u^{u/t_{y,u},+}$ with $t_0 > 0$ be such that $|f * \psi_u^{y,u}(y)| \geq \frac{1}{2}\sup_{t<u}|f * \psi_u^{u/t,+}(y)|$. Now, observe that that there exists a wave packet $\overline{\psi}^{y,u} \in \Phi^*$ supported on $(-a/2, a/2)$, such that for some $\overline{\theta} \in \Theta^{(in)} + 2b$

$$f * \psi_u^{y,u} = F^{\overline{\psi}^{y,u}}(y, \overline{\theta}u^{-1}, u)$$

From now on, for simplicity, we write $F^{y,u} := F^{\overline{\psi}^{y,u}}$. The consideration above imply that (2.53) is bounded by

$$\int_{-s}^{s}\int_{c^-(y)}^{c^+(y)} \Big|\int_t^{c^+(y)} |F^{y,u}(y, \overline{\theta}u^{-1}, u)|\,\frac{t}{u}\,\frac{du}{u}\Big|^2\,\frac{dt}{t}\,dy$$

Let $g(t) = t\mathbb{1}_{(0,1)}(t)$. Note that the integral over t is the L^2 norm of the convolution

$$(F^{y,\cdot}(y, \overline{\theta}\cdot^{-1}, \cdot)\mathbb{1}_{C^\theta(y)}(\cdot)) * g$$

in the multiplicative group $(\mathbb{R}_+, \frac{dt}{t})$. Using Young's inequality, it is then estimated by

$$\int_{-s}^{s}\int_0^{\delta s} |F^{y,u}(y,\overline{\theta}u^{-1},u)|^2 \mathbb{1}_{C^{\theta}(y)}(u)\,\frac{du}{u}\,dy.$$

The previous display is estimated further by

$$\int_{-s}^{s}\int_0^{\delta s} |F^{y,u}(y,\overline{\theta}u^{-1},u)|^2 \mathbb{1}_{C^{\overline{\theta}}(y)\cap C^{\theta}(y)}(u)\,\frac{du}{u}\,dy$$
$$+\int_{-s}^{s}\int_0^{\delta s} |F^{y,u}(y,\overline{\theta}u^{-1},u)|^2 \mathbb{1}_{C^{\theta}(y)\setminus C^{\overline{\theta}}(y)}(u)\,\frac{du}{u}\,dy.$$

First we bound the second summand above and then the first. In order to bound the second summand it is enough to show the following.

Lemma 2.61. *Let $\overline{\theta} \in \Theta^{(in)} + 2b$. Then*

$$C^{\theta}(y)\setminus C^{\overline{\theta}}(y) \subset (c^-(y), 2c^-(y)) \cup (\frac{1}{2}c^+(y), c^+(y)),$$

Proof. 1. First we show that

$$t \in C^{\theta}(y), t \notin C^{\overline{\theta}}(y) \implies \text{there exists } s \in C^{\overline{\theta}}(y) \text{ with } \frac{1}{2} \le \frac{s}{t} \le 2. \tag{2.54}$$

The left hand side of the implication means that there exists a tree separating the points $(y, \theta t^{-1}, t)$ and $(y, \overline{\theta}t^{-1}, t)$, i.e. there exists a ξ such that

$$\theta t^{-1} \le \xi + t^{-1} \le \overline{\theta}t^{-1} \tag{2.55}$$

or

$$\theta t^{-1} \le \xi - t^{-1} \le \overline{\theta}t^{-1}. \tag{2.56}$$

Let us consider only the first case, since the other one is very similar. Since $|\theta|, |\overline{\theta}| \ll 1$, the hyperbola $\overline{\theta}u^{-1}$ is "steeper" than $\xi + u^{-1}$ and there exists $s < t$ such that

$$\overline{\theta}s^{-1} = \xi + s^{-1}.$$

Subtracting this equality from the left inequality in (2.55) one obtains

$$\theta t^{-1} - \overline{\theta}s^{-1} \le t^{-1} - s^{-1} \implies (1-\overline{\theta})s^{-1} \le (1-\theta)t^{-1}$$
$$\implies \frac{s}{t} \ge \frac{1-\overline{\theta}}{1-\theta} \ge \frac{1}{2},$$

where the last inequality follows from $|\theta|, |\overline{\theta}| \ll 1$. Similarly, one can show that if (2.56) is satisfied, then there exists $\frac{s}{t} \le 2$ with $s \in C^{\overline{\theta}}(y)$. This ends the proof of (2.54).

2. Note that for any $\widetilde{\theta} \in \Theta$, $C^{\widetilde{\theta}}(y)$ is a connected set, hence if $c^-(y), c^+(y) \in C^{\overline{\theta}}(y)$, then $C^{\theta}(y) \subset C^{\overline{\theta}}(y)$ and there is nothing to prove. Also, note that if $c^-(y)$, $c^+(y)$ are at most factor 2 away from each other, then there is also nothing to prove.

Assume that the above is not the case and $c^+(y) \notin C^{\overline{\theta}}(y)$. Note that (2.54) implies that $\frac{c^+(y)}{2} \in C^{\overline{\theta}}(y)$. Similarly, if $c^-(y) \notin C^{\overline{\theta}}(y)$, then $2c^-(y) \in C^{\overline{\theta}}(y)$. □

Using the above lemma we bound

$$\begin{aligned}&\int_{-s}^{s}\int_{0}^{\delta s}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\mathbb{1}_{C^\theta(y)\setminus C^{\overline{\theta}}(y)}(u)\,\frac{du}{u}\,dy\\&\lesssim\int_{-s}^{s}\sup_{u\in(c^-(y),2c^-(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\,dy\\&+\int_{-s}^{s}\sup_{u\in(\frac{1}{2}c^+(y),c^+(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\,dy.\end{aligned}$$

Note that $\sup_{u\in(c^-(y),2c^-(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|\leq|F(y,\theta c^-(y)^{-1},c^-(y)|$ and similar holds for the term involving $c^+(y)$. Hence, the right hand side of the last display is bounded by $2s\|F\|^2_{L^\infty(R_{\delta,C})}+2s\|F\mathbb{1}_C\|^2_{L^\infty(\delta^{1/2}S^\infty)}$. We still have to deal with the L^2 integral restricted to $C^{\overline{\theta}}$.

Lemma 2.62. *Let $\overline{\theta}\in\Theta^{(in)}+2b$. Then*

$$\int_{-s}^{s}\int_{0}^{\delta s}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\mathbb{1}_{C^{\overline{\theta}}(y)\cap C^\theta(y)}(u)\,\frac{du}{u}\,dy\lesssim s\|\boldsymbol{F}\|^2_{L^\infty(S_{\delta,C})}.$$

Remark 2.63. *The point of the lemma is essentially to replace the left hand side by its average over θ on an interval of size ~ 1. In order to that we split the integral into a part that is close to the boundary and the other part, where we have "enough space" to perform integration over θ.*

Proof. Observe that

$$C^{\overline{\theta}}(y)\cap C^\theta(y):=D(y)\cup\widetilde{C}^{\overline{\theta}}(y),$$

where

$$D(y):=(c^-(y),2c^-(y))\cup(\frac{1}{2}c^+(y),c^+(y)),\qquad \widetilde{C}^{\overline{\theta}}(y):=C^{\overline{\theta}}(y)\cap C^\theta(y)\setminus D^{\overline{\theta}}(y).$$

Note that

$$\begin{aligned}&\int_{-s}^{s}\int_{D^{\overline{\theta}}(y)}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\mathbb{1}_{C^{\overline{\theta}}(y)}(u)\,\frac{du}{u}\,dy\\&\lesssim\int_{-s}^{s}\sup_{u\in(c^-(y),2c^-(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\,dy\\&\qquad+\int_{-s}^{s}\sup_{u\in(\frac{1}{2}c^+(y),c^+(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\,dy\\&\lesssim s(\|F\|^2_{L^\infty(R_{\delta,C})}+\|F\mathbb{1}_C\|^2_{L^\infty(\delta^{1/2}S^\infty)}),\end{aligned}$$

since $\sup_{u\in(c^-(y),2c^-(y))}|F^{y,u}(y,\overline{\theta}u^{-1},u)|^2\leq|F(y,\theta c^-(y)^{-1},c^-(y))|$, where $\theta=\overline{\theta}-2b\in\Theta^{(in)}$ and similarly for the term involving $c^+(y)$.

Now we bound the contribution of $\widetilde{C}^\theta(y)$. We have the following observation.

Lemma 2.64. *For any $\widetilde{\theta}\in\Theta^{(in)}+2b$ and $u\in\widetilde{C}^{\overline{\theta}}(y)$, we have $(y,\widetilde{\theta}u^{-1},u)\in C$.*

Proof. This follows directly from Lemma 2.61. □

Let $c < b$. Applying the previous lemma and making c small enough we may ensure that for any $\widetilde{\theta} \in [\overline{\theta}, \overline{\theta} + c)$ and $u \in \widetilde{C^{\theta}}(y)$ we have $|F^{y,u}(y, \overline{\theta} u^{-1}, u)| \leq |F^*(y, \widetilde{\theta} u^{-1}, u)|$. This implies

$$\begin{aligned}
&\int_{-s}^{s} \int_{0}^{\delta s} |F^{y,u}(y, \overline{\theta} u^{-1}, u)|^2 \mathbb{1}_{\widetilde{C^{\theta}}(y)}(u) \, \frac{du}{u} \, dy \\
&\lesssim_c \int_{\overline{\theta}}^{\overline{\theta}+c} \int_{-s}^{s} \int_{\delta s} |F(y, \widetilde{\theta} u^{-1}, u)|^2 \mathbb{1}_C(y, \widetilde{\theta} u^{-1}, u) \, \frac{du}{u} \, dy \, d\widetilde{\theta} \\
&\lesssim s(\|F^*\|^2_{L^\infty(S^2_{\delta,C})} + \|F^* \mathbb{1}_C\|^2_{L^\infty(\delta^{1/2} S^2)}),
\end{aligned}$$

where the last inequality follows from $[\overline{\theta}, \overline{\theta} + c) \subset \Theta^{(out)}$ (since $\overline{\theta} \in \Theta^{(in)} + 2b$). □

(2.48) is proven analogously to (2.47); the only difference is that we do not use the boundary size $R_{\delta,C}$, but S^∞.

2. Notice that $G^\varphi(y,t) = G^\psi(y,t)$, where $\psi = \varphi_{\theta+\beta_1}$. Note that the first part of the proof applies to such ψ and follows similarly. □

The following two simple lemmata below will be extensively used when we integrate by parts the trilinear form.

Lemma 2.65 (Differentiating wave packets in space)**.**

$$\partial_y F^\varphi(y, \xi, u) = u^{-1} F^{\varphi'}(y, \xi, u). \tag{2.57}$$

Proof. Straightforward. □

Lemma 2.66 (Differentiating wave packets in scale)**.**

$$\partial_u F^\varphi(y, \xi, u) = u^{-1} F^{\vartheta}(y, \xi, u), \tag{2.58}$$

where $\vartheta(x) = (x\varphi(x))'$.

Proof. It follows from the identity

$$\partial_u u^{-1} \varphi(\frac{x}{u}) = -u^{-2} \varphi(\frac{x}{u}) - x u^{-3} \varphi'(\frac{x}{u}) = u^{-1} u^{-1} \psi'(\frac{x}{u}),$$

where $\psi(x) = x\varphi(x)$. Hence, the right hand side equals $u^{-1}\vartheta_u$, with $\vartheta = \psi'$. □

Lemma 2.67. *Let $\varphi \in \Phi^*$ and $0 < \varepsilon < 2^{-10}b$. Then there is a decomposition such that $\varphi = \overline{\varphi} + \psi$, where $\overline{\varphi}, \psi \in \Phi^*$, $\hat{\overline{\varphi}}$ is constant on $(-\varepsilon, \varepsilon)$ and there exists $\vartheta \in \Phi^*$ such that $\psi = \vartheta'$.*

Proof. Let $\eta \in \Phi^*$ be such that $\hat{\eta}$ is supported on $(-2\varepsilon, 2\varepsilon)$ and constant 1 on $(-\varepsilon, \varepsilon)$. Define ψ, $\overline{\varphi}$ as follows

$$\hat{\psi}(\xi) = \hat{\varphi}(\xi) - \hat{\eta}(\xi)\hat{\varphi}(0), \qquad \hat{\overline{\varphi}}(\xi) = \hat{\eta}(\xi)\hat{\varphi}(0).$$

Clearly $\varphi = \psi + \overline{\varphi}$ and $\hat{\overline{\varphi}}$ is constant on $(-\varepsilon, \varepsilon)$. Let

$$\vartheta(x) := \int_{-\infty}^{x} \psi(y) \, dy = \int_{-\infty}^{x} \varphi(y) - \hat{\varphi}(0)\eta(y) \, dy.$$

$\vartheta \in \Phi$, because $\varphi, \eta \in \Phi^*$. Moreover, $\vartheta' = \psi$. The last identity implies also that $\hat{\vartheta}$ is supported on $(-2^{-9}b, 2^{-9}b)$, because both $\hat{\varphi}$ and $\hat{\eta}$ are supported on that interval. Hence $\vartheta \in \Phi^*$. □

Till the end of this subsection, unless otherwise stated we set

$$B(y) := (b^-(y), b^+(y)) \subset (0, s - |y|),$$

where $b^\pm \colon B_s(0) \to \mathbb{R}_+$ are piecewise linear, Lipschitz functions and

$$C(y) := (c^-(y), c^+(y)) \subset (0, \min(\delta s, s - |y|)).$$

where $c^\pm \colon B_s(0) \to \mathbb{R}_+$ are piecewise linear, Lipschitz functions.

We use the following definition of BMO

$$\|f\|_{BMO} := \sup_I \fint_I |f - \fint_I f|.$$

In the following lemma we prove boundedness of a version of the Hilbert transform, operator (2.59), that involves restriction to scales $t \in B(y)$, under appropriate Lipschitzity condition. However, note that the bounds (2.60), (2.61) are in terms V.

Lemma 2.68. *Let*

$$LH(y) = \int_0^s H(y,t) \mathbb{1}_{B(y)}(t_1) \frac{dt}{t} \mathbb{1}_{[-s,s]}(y). \tag{2.59}$$

Let $\overline{\varphi}' = \varphi$ for $\overline{\varphi} \in \Phi^*$. We have

$$\|LG^\varphi\|_{L^2} \lesssim s^{1/2} \sup_{\zeta \in \{\overline{\varphi}, \varphi, \varphi'\}} \|G^\zeta \mathbb{1}_B\|_V, \tag{2.60}$$

$$\|LG^\varphi\|_{BMO} \lesssim \sup_{\zeta \in \{\overline{\varphi}, \varphi, \varphi'\}} \|G^\zeta \mathbb{1}_B\|_V, \tag{2.61}$$

and for $2 < p < \infty$

$$\|LG^\varphi\|_{L^p} \lesssim_p s^{1/p} \sup_{\zeta \in \{\overline{\varphi}, \varphi, \varphi'\}} \|G^\zeta \mathbb{1}_B\|_V. \tag{2.62}$$

Proof. In the proof we omit writing the complex conjugate. By assumption, $\overline{\varphi}' = \varphi$, moreover we set $\widetilde{\varphi} := \varphi'$.

(2.60): Let $C := \|LG^\varphi\|_{L^2}$ and $D := s^{1/2} \sup_{\zeta \in \{\overline{\varphi}, \varphi, \varphi'\}} \|G^\zeta \mathbb{1}_B\|_V$, we want to prove that $C \lesssim D$. Note that C^2 equals

$$\int_{-s}^s \int_0^s \int_0^s G^\varphi(y, t_1) \mathbb{1}_{B(y)}(t_1) G^\varphi(y, t_2) \mathbb{1}_{A(y)}(t_2) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \, dy$$

by symmetry we bound only

$$\int_{-s}^s \int_0^s \int_{0<t_1<t_2} G^\varphi(y, t_1) \mathbb{1}_{B(y)}(t_1) G^\varphi(y, t_2) \mathbb{1}_{B(y)}(t_2) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \, dy$$

We shall now integrate by parts moving derivative from $G^\varphi(y, t_1)$ to $G^\varphi(y, t_2)$.

Hence, applying Proposition 2.55 it is enough to estimate the following four integrals

$$\int_{-s}^s \int_0^s \int_{t_1<t_2<s} G^{\widetilde{\varphi}}(y, t_2) \mathbb{1}_{B(y)}(t_2) \frac{t_1}{t_2} \frac{dt_2}{t_2} G^{\overline{\varphi}}(y, t_1) \mathbb{1}_{B(y)}(t_1) \frac{dt_1}{t_1} \, dy \tag{2.63}$$

$$\int_{-s}^{s}\int_{0}^{s}\int_{0<t_1<t_2} G^{\overline{\varphi}}(y,t_1)\partial_y \mathbb{1}_{B(y)}(t_1)G^{\varphi}(y,t_2)\mathbb{1}_{B(y)}(t_2)\,dt_1\,\frac{dt_2}{t_2}\,dy \tag{2.64}$$

$$\int_{-s}^{s}\int_{0}^{s}\int_{0<t_1<t_2} G^{\overline{\varphi}}(y,t_1)\mathbb{1}_{B(y)}(t_1)G^{\varphi}(y,t_2)\partial_y\mathbb{1}_{B(y)}(t_2)\,dt_1\,\frac{dt_2}{t_2}\,dy \tag{2.65}$$

Concerning (2.63) we apply Cauchy-Schwarz in t_1 with measure $\frac{dt_1}{t_1}$ first and then we use Young's convolution inequality for the multiplicative group $(\mathbb{R}_+,\frac{dt}{t})$:

$$\lesssim \int_{-s}^{s} \|G^{\overline{\varphi}}\mathbb{1}_{B(y)}(t) * t\mathbb{1}_{(0,1)}(t)\|_{L^2(\frac{dt}{t})}\,\|G^{\widetilde{\varphi}}\mathbb{1}_{B(y)}(t)\|_{L^2(\frac{dt}{t})}\,dy \lesssim D^2$$

(2.64): applying (2.43) and the Cauchy-Schwarz inequality in y, it is bounded by

$$s^{1/2}\|G^{\overline{\varphi}}\|_V\Big(\int_{-s}^{s}|\int_0^s G^{\varphi}(y,t_2)\mathbb{1}_{B(y)}(t_2)\,\frac{dt_2}{t_2}|^2\,dy\Big)^{1/2}\lesssim DC$$

(2.65): applying (2.43) and using Cauchy-Schwarz twice, in t_1 and then in y, it is estimated by

$$\begin{aligned}
&\int_{-s}^{s}\int_{0}^{s}\int_{0<t_1<t_2}|G^{\overline{\varphi}}(y,t_1)|\mathbb{1}_{B(y)}(t_1)\frac{t_1^{1/2}}{t_1^{1/2}}\,dt_1|G^{\varphi}(y,t_2)\partial_y\mathbb{1}_{B(y)}(t_2)|\frac{dt_2}{t_2}\,dy\\
&\lesssim \int_{-s}^{s}\int_{0}^{s}(\int_{0<t_1<t_2}|G^{\overline{\varphi}}(y,t_1)|^2\mathbb{1}_{B(y)}(t_1)\,\frac{dt_1}{t_1})^{1/2}|G^{\varphi}(y,t_2)\partial_y\mathbb{1}_{B(y)}(t_2)|dt_2\,dy\\
&\lesssim (\int_{-s}^{s}\int_{0}^{s}|G^{\overline{\varphi}}(y,t_1)|^2\mathbb{1}_{B(y)}(t_1)\frac{dt_1}{t_1}\,dy)^{1/2}\|G^{\varphi}\mathbb{1}_B\|_V\lesssim D^2.
\end{aligned}$$

We have just proven that $C^2\lesssim D^2+CD$. Hence, either $C^2\lesssim D^2$ or $C^2\lesssim CD$. This implies $C\lesssim D$.

(2.61): Fix an interval I. Without loss of generality we may assume that either $I\subset[-s,s]$ or $[-s,s]\subset I$. In the latter case the bound follows from the L^2 estimate. Also, without loss of generality, we may assume that $I=[-w,w]$. Hence we consider the first case and using Jensen's inequality it is enough to bound

$$\Big(\fint_{-w}^{w}|\int_0^s G^{\varphi}(y,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}-\fint_{-w}^{w}\int_0^s G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|^2dy\Big)^{1/2}$$

Note that by the already shown L^2 estimate

$$\fint_{-w}^{w}|\int_0^w G^{\varphi}(y,t)\mathbb{1}_{B(y)}(t)\frac{dt}{t}|^2\,dy\lesssim \sup_{\zeta\in\{\overline{\varphi},\varphi,\varphi'\}}\|G^{\zeta}\mathbb{1}_B\|_V^2$$

and by Jensen's inequality

$$\begin{aligned}
&\fint_{-w}^{w}|\fint_{-w}^{w}\int_0^w G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|^2\,dy\\
&\lesssim \fint_{-w}^{w}\fint_{-w}^{w}|\int_0^w G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}|^2\,dz\,dy\lesssim \sup_{\zeta\in\{\overline{\varphi},\varphi,\varphi'\}}\|G^{\zeta}\mathbb{1}_B\|_V^2
\end{aligned}$$

Hence, we are just left with bounding

$$\fint_{-w}^{w}|\int_w^s G^{\varphi}(y,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}-\fint_{-w}^{w}\int_w^s G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|^2dy$$

Using Lemma 2.69 the difference inside the absolute value is bounded by

$$|\fint_{-w}^{w}\int_{w}^{s} G^{\varphi}(y,t)\mathbb{1}_{B(y)}(t) - G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz| \lesssim \sup_{\zeta\in\{\varphi,\varphi'\}} \|G^{\zeta}\mathbb{1}_B\|_V.$$

Using $L^2 - BMO$ interpolation (see, [CZ05] for example), we conclude the proof. □

Lemma 2.69. *Let $w < s$ and $y \in (-w, w)$. We have*

$$|\fint_{-w}^{w}\int_{w}^{s} G^{\varphi}(y,t)\mathbb{1}_{B(y)}(t) - G^{\varphi}(z,t)\mathbb{1}_{B(z)(t)}\,\frac{dt}{t}\,dz| \lesssim \sup_{\zeta\in\{\varphi,\varphi'\}} \|G^{\zeta}\mathbb{1}_B\|_V$$

Proof. Using the fundamental theorem of calculus, we have two integrals to bound. The first one is estimated using (2.43)

$$\begin{aligned}
&\fint_{-w}^{w} |\int_{z}^{y}\int_{w}^{s} G^{\varphi}(\zeta,t)\partial_{\zeta}\mathbb{1}_{B(\zeta)}(t)\,\frac{dt}{t}\,d\zeta|\,dz\\
&= \fint_{-w}^{w}\frac{1}{w}\int_{z}^{y}\sum_{j=1}^{2}|G^{\varphi}(\zeta,b_j(\zeta))|\,d\zeta\,dz \lesssim \|G^{\varphi}\mathbb{1}_B\|_V,
\end{aligned}$$

and the second one is estimated by

$$\begin{aligned}
&\fint_{-w}^{w}\int_{z}^{y}\int_{w}^{s} |G^{\varphi'}(\zeta,t)|\mathbb{1}_{B(\zeta)}(t)\,\frac{dt}{t^2}\,d\zeta\,dz\\
&\lesssim \|G^{\varphi'}\mathbb{1}_B\|_V \fint_{-w}^{w}\frac{1}{w}\int_{z}^{y} d\zeta\,dz \lesssim \|G^{\varphi'}\mathbb{1}_B\|_V.
\end{aligned}$$

□

The subsequent lemma is essentially boundedness of the maximal truncation of the Hilbert transform in terms of the size. Once again we adapt the inequality to the varying scale restriction $t \in B(y)$ and bound the operator in terms of the size S.

Lemma 2.70. *Define*

$$LH(y) = \sup_{w>0} |\int_{w}^{s} H(y,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}|\mathbb{1}_{[-s,s]}(y).$$

Let $\overline{\varphi}' = \varphi$ for $\overline{\varphi} \in \Phi^*$. We have for $2 \le p < \infty$

$$\|LG^{\varphi}\|_{L^p} \lesssim s^{1/p} \sup_{\zeta\in\{\overline{\varphi},\varphi,\varphi'\}} \|G^{\zeta}\mathbb{1}_B\|_V.$$

Proof. Fix $0 < w < s$. We shall obtain a pointwise bound for $|\int_w^s G^{\varphi}(y,t)\,\frac{dt}{t}|$. Let $\widetilde{I}_w$ be the interval of length w centered at y and let $I_w = \widetilde{I}_w \cap [-s,s]$; note that $|\widetilde{I}_w|$ and $|I_w|$ are comparable. Subtracting an averaged term it is enough to bound the following three expressions

$$|\int_{w}^{s} G^{\varphi}\mathbb{1}_{B(y)}(t) - \fint_{I_w}\int_{w}^{s} G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|, \tag{2.66}$$

$$|\fint_{I_w}\int_{0}^{w} G^{\varphi}(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|, \tag{2.67}$$

$$|\fint_{I_w}\int_0^s G^\varphi(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz|. \tag{2.68}$$

Using Lemma 2.69, (2.66) is bounded by

$$|\fint_{I_w}\int_w^s G^\varphi(y,t)\mathbb{1}_{B(y)}(t) - G^\varphi(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}\,dz| \lesssim \sup_{\zeta\in\{\varphi,\varphi'\}}\|G^\zeta\mathbb{1}_B\|_V.$$

By an application of the Cauchy-Schwarz inequality and using (2.60), (2.67) is estimated by

$$\begin{aligned}&\fint_{I_w}|\int_0^w G^\varphi(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}|\,dz\\ &\lesssim \Big(\fint_{I_w}|\int_0^w G^\varphi(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}|^2\,dz\Big)^{1/2} \lesssim \sup_{\zeta\in\{\overline{\varphi},\varphi,\varphi'\}}\|G^\zeta\mathbb{1}_B\|_V.\end{aligned}$$

(2.68) is estimated by

$$\fint_{I_w}|\int_0^s G^\varphi(z,t)\mathbb{1}_{B(z)}(t)\,\frac{dt}{t}|\,dz \lesssim M(\int_0^s G^\varphi(\cdot,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}\mathbb{1}_{[-s,s]})(y),$$

where M is the maximal function. Hence

$$\begin{aligned}\|LG^\varphi\|_{L^p} &\lesssim \|\|G^\varphi\mathbb{1}_B\|_V\mathbb{1}_{[-s,s]}\|_{L^p} + \|M(\int_0^s G^\varphi(\cdot,t)\,\frac{dt}{t}\mathbb{1}_{[-s,s]})\|_{L^p}\\ &\lesssim s^{1/p}\sup_{\zeta\in\{\overline{\varphi},\varphi,\varphi'\}}\|G^\zeta\mathbb{1}_B\|_V,\end{aligned}$$

where we used boundedness of the maximal function and (2.62). □

Applying Cauchy-Schwarz together with the previous lemma we obtain the following.

Lemma 2.71 (Bilinear estimate)**.** *Let $G\in C(\mathbb{R}^2_+)$. We have for $p\geq 1$*

$$(\int_{-s}^s |G_1(y,c^\pm(y))G_2(y,c^\pm(y))|^p\,dy)^{1/p} \lesssim s^{1/p}\prod_{j=1}^2\|G_j\|_{R,\boldsymbol{V}_C}^{1/p}\|G_j\mathbb{1}_C\|_V^{1-1/p}. \tag{2.69}$$

Proof. This follows from Cauchy-Schwarz and pulling out the L^∞ norm with the appropriate power. □

We shall need the following time-scale localized square-function estimate.

Lemma 2.72. *Let $\varphi\in\Phi^*$ and let $\theta\in\Theta$. Define*

$$LH(y) = (\int_0^{\delta s}|H(y,t)|^2\mathbb{1}_{C(y)}(t)\,\frac{dt}{t})^{1/2}\mathbb{1}_{[-s,s]}(y).$$

We have

$$\|(LG^\varphi)^2\|_{L^1} \lesssim s\|G^\varphi\mathbb{1}_C\|_{\boldsymbol{V}^2}^2, \tag{2.70}$$

$$\|(LG^\varphi)^2\|_{BMO} \lesssim \sup_{\zeta\in\{\varphi,\varphi'\}}\|G^\zeta\mathbb{1}_C\|_V^2, \tag{2.71}$$

and for $2<p<\infty$

$$\|LG^\varphi\|_{L^p} \lesssim_p s^{1/p}\|G^\varphi\mathbb{1}_C\|_{\boldsymbol{V}^2}^{2/p}(\sup_{\zeta\in\{\varphi,\varphi'\}}\|G^\zeta\mathbb{1}_C\|_V)^{1-2/p}. \tag{2.72}$$

Proof. (2.70): Follows by definition.

(2.71): It is enough to consider $I \subset [-s, s]$ and $[-s, s] \subset I$. The latter case follows from (2.70). Let us consider the first, without loss of generality $I = [-w, w]$. We shall bound

$$\fint_{-w}^{w} |\int_0^{\delta s} |G^\varphi(y,t)|^2 \mathbb{1}_{C(y)}(t) \frac{dt}{t} - \fint_{-w}^{w} \int_0^{\delta s} |G^\varphi(z,t)|^2 \mathbb{1}_{C(z)}(t) \frac{dt}{t} \, dz| \, dy$$

Notice that

$$\fint_{-w}^{w} \int_0^{w} |G^\varphi(y,t)|^2 \mathbb{1}_{C(y)}(t) \frac{dt}{t} \, dy \lesssim \|G^\varphi \mathbb{1}_C\|_V^2,$$

and

$$\fint_{-w}^{w} \fint_{-w}^{w} \int_0^{w} |G^\varphi(z,t)|^2 \mathbb{1}_{C(y)}(t) \frac{dt}{t} \, dz \, dy \lesssim \|G^\varphi \mathbb{1}_C\|_V^2.$$

Hence we are just left with estimating

$$\fint_{-w}^{w} |\int_w^{\delta s} |G^\varphi(y,t)|^2 \mathbb{1}_{C(y)}(t) \frac{dt}{t} - \fint_{-w}^{w} \int_w^{\delta s} |G^\varphi(z,t)|^2 \mathbb{1}_{C(z)}(t) \frac{dt}{t} \, dz| \, dy$$

We can rewrite the difference inside the absolute value of the previous display as

$$\fint_{-w}^{w} \int_w^{\delta s} |G^\varphi(y,t)|^2 \mathbb{1}_{C(y)}(t) - |G^\varphi(z,t)|^2 \mathbb{1}_{C(y)}(t) \frac{dt}{t} \, dz$$

Applying the fundamental theorem of calculus to the above innermost difference, similarly Lemma 2.69, it is enough to control the following two integrals

$$\begin{aligned} &\fint_{-w}^{w} \int_w^{\delta s} \int_z^y |G^{\varphi'}(\zeta,t)||G^\varphi(\zeta,t)| \mathbb{1}_{C(\zeta)}(t) \, d\zeta \frac{dt}{t^2} \, dz \\ &\lesssim w \fint_{-w}^{w} \int_w^\infty \frac{1}{t^2} \, dt \, dz \sup_{\zeta \in \{\varphi, \varphi'\}} \|G^\zeta \mathbb{1}_C\|_V^2 \lesssim \sup_{\zeta \in \{\varphi, \varphi'\}} \|G^\zeta \mathbb{1}_C\|_V^2, \end{aligned}$$

and when the derivative falls on $\mathbb{1}_C$ we apply (2.43)

$$\begin{aligned} &\fint_{-w}^{w} \int_z^y \int_w^{\delta s} |G^\varphi(\zeta,t)|^2 \partial_\zeta \mathbb{1}_{C(\zeta)}(t) \frac{dt}{t} \, d\zeta \, dz \\ &\lesssim \fint_{-w}^{w} \frac{1}{w} \int_z^y \sum_{j=1}^{2} |G^\varphi(\zeta, b_j(\zeta))|^2 \, d\zeta \, dz \lesssim \|G^\varphi \mathbb{1}_C\|_V^2. \end{aligned}$$

(2.72) follows from the $L^1 - BMO$ interpolation ([CZ05]). □

Corollary 2.73. *Let $\varphi_1, \varphi_2 \in \Phi^*$. Let $1 < p < \infty$. Then*

$$\begin{aligned} &(\int_{-s}^{s} |\int_0^{\delta s} |G_1^{\varphi_1}(y,t) G_2^{\varphi_2}(y,t)| \mathbb{1}_{C(y)}(t) \frac{dt}{t}|^p \, dy)^{1/p} \\ &\qquad \lesssim s^{1/p} \prod_{j=1}^{3} \|G_j^{\varphi_j} \mathbb{1}_C\|_{\mathbf{V}^2}^{1/p} \|G_j^{\varphi_j} \mathbb{1}_C\|_V^{1-1/p} \end{aligned}$$

Proof. Applying the Cauchy-Schwarz inequality in u and y the left hand side is bounded by

$$\begin{aligned}
&\Big(\int_{-s}^{s}\prod_{j=1}^{2}\Big|\int_0^{\delta s}|G_j^{\varphi}(y,t)|^2\mathbb{1}_{C(y)}(t)\,\frac{dt}{t}\Big|^{p/2}dy\Big)^{1/p}\\
&\lesssim \prod_{j=1}^{2}\Big(\int_{-s}^{s}\Big|\int_0^{\delta s}|G_j^{\varphi}(y,t)|^2\mathbb{1}_{C(y)}(t)\,\frac{dt}{t}\Big|^{p}dy\Big)^{1/2p}
\end{aligned}$$

The statement follows from (2.72) applied with the exponent $2p$. □

2.5.3 Proof of Proposition 2.53

In this subsection we prove Proposition 2.53. Let $V_j, W_j \in \mathbb{D}^{\cup}$, for $j = 1, 2, 3$ and let

$$A_j := A_\varepsilon \cap V_j \setminus W_j, \qquad \text{for } j = 1, 2, 3,$$

where A_ε is as in (2.23), be fixed throughout this section. Moreover, set

$$A := V_1 \cap A_2 \cap A_3,$$

All the implicit constants will be independent of δ, A. The main ingredient of the proof of Proposition 2.53 is the following lemma. The point of it is that we reduce the full estimate from Proposition 2.53 to the single tree estimate.

Lemma 2.74. *Let $T \in \mathbb{T}$ and let*

$$P_1 = \bigcap_{j=1}^{m} M_j \setminus L_1, \qquad \textit{where } M_j, L_1 \in \mathbb{T}^{\cup} \textit{ for } j = 1, 2, .., m,$$

$$P = \bigcap_{j=1}^{n} N_j \setminus L, \qquad \textit{where } N_j, L \in \mathbb{T}_\delta^{\cup} \textit{ for } j = 1, 2, .., n.$$

Then for $\varphi \in \Phi^*$ and $0 < \gamma < 1$ it holds that

$$\begin{aligned}
&\Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(T)}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}, F_2^{\varphi}\mathbb{1}_{A}\mathbb{1}_{\pi_2(P)}, F_3\mathbb{1}_{A}\mathbb{1}_{\pi_3(P)})\\
&\lesssim_{m,n} \mu(T)\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}\|_{L^\infty(S_1)}\prod_{j=2}^{3}\|\boldsymbol{F}_j\mathbb{1}_{A}\mathbb{1}_{\pi_j(P)}\|_{L^\infty(S^{\gamma}_{j,\delta,A})}.
\end{aligned}$$

Proof of Proposition 2.53 assuming Lemma 2.74. We gradually reduce the estimate in three steps.

Step 1: First of all note that by (2.25), it is enough to prove that for all $K_1, L_1, M_1 \in \mathbb{T}^{\cup}$ and $K_j, L_j, M_j \in \mathbb{T}_\delta^{\cup}$ for $j = 2, 3$ it holds that

$$\begin{aligned}
&\Lambda(F_1^{\varphi}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(K_1\cap M_1\setminus L_1)}, F_2^{\varphi}\mathbb{1}_{A_2}\mathbb{1}_{\pi_2(K_2\cap M_2\setminus L_2)}, F_3\mathbb{1}_{A_3}\mathbb{1}_{\pi_3(K_3\cap M_3\setminus L_3)})\\
&\lesssim (\min_{j=1,2,3}\mu_j(\pi_j(K_j)))\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(K_1\cap M_1\setminus L_1)}\|_{L^\infty(S_1)}\prod_{j=2}^{3}\|\boldsymbol{F}_j\mathbb{1}_{A_j}\mathbb{1}_{\pi_j(K_j\cap M_j\setminus L_j)}\|_{L^\infty(S^{\gamma}_{j,\delta,A})}.
\end{aligned}$$

Step 2: Let

$$P_1 = K_1 \cap M_1 \setminus L_1, \qquad P = K_2 \cap M_2 \cap K_3 \cap M_3 \setminus (L_2 \cup L_3).$$

It follows from (2.27) and (2.29), that in order to prove the estimate from the previous step it is enough to show the following for all $K_1, L_1, M_2 \in \mathbb{T}^{\cup}$, $K_j, L_j, M_j \in \mathbb{T}^{\cup}_{\delta}$, $j = 2, 3$

$$\begin{aligned}
&\Lambda(F_1^{\varphi}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}, F_2^{\varphi}\mathbb{1}_{A}\mathbb{1}_{\pi_2(P)}, F_3\mathbb{1}_{A}\mathbb{1}_{\pi_3(P)}) \\
&\lesssim \min(\mu(M_j), \min_j \mu_\delta(N_j))\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}\|_{L^\infty(S_1)} \prod_{j=2}^{3} \|\boldsymbol{F}_j\mathbb{1}_{A}\mathbb{1}_{\pi_j(P)}\|_{L^\infty(S^{\gamma}_{j,\delta,A})}.
\end{aligned}$$

Step 3: in this step we show that in order prove the estimate from the previous step it is enough to show the following for all $T \in \mathbb{T}$, $P_1 = M_1 \cap M_2 \setminus L_1$, $P = \bigcap_{j=1}^{4} N_j \setminus L$:

$$\begin{aligned}
&\Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(T)}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}, F_2^{\varphi}\mathbb{1}_{A_3}\mathbb{1}_{\pi_2(P)}, F_3^{\varphi}\mathbb{1}_{A_2}\mathbb{1}_{\pi_3(P)}) \\
&\lesssim \mu(T)\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}\|_{L^\infty(S_1)} \prod_{j=2}^{3} \|\boldsymbol{F}_j\mathbb{1}_{A_j}\mathbb{1}_{\pi_j(P)}\|_{L^\infty(S^{\gamma}_{j,\delta,A})}.
\end{aligned}$$

Let

$$\overline{P} := \bigcap_{j=1}^{2} M_j \cap \bigcap_{j=1}^{4} \widetilde{N}_j,$$

where $\widetilde{N}_j$ are like in (2.30). Let $\bigcup_{i=1}^{J} T_i \supset \overline{P}$ be such that

$$\sum_{i=1}^{J} \mu(T_i) \lesssim \min(\min_j \mu(M_j), \min_j \mu(\widetilde{N}_j)).$$

The above condition is possible to satisfy, by covering one of the sets M_j, $\widetilde{N}_j$, which has smallest measure. We can make T_j's pairwise disjoint obtaining convex trees $\bigcup_{i=1}^{J} T_i \setminus \overline{T}_i$. Using (2.30) and assuming the estimate in the previous display we obtain

$$\begin{aligned}
&\Lambda(F_1^{\varphi}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}, F_2^{\varphi}\mathbb{1}_{A_2}\mathbb{1}_{\pi_2(P)}, F_3^{\varphi}\mathbb{1}_{A_3}\mathbb{1}_{\pi_3(P)}) \\
&\quad = \Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(\bigcup_{i=1}^{J} T_i \setminus \overline{T}_i)}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}, F_2^{\varphi}\mathbb{1}_{A_2}\mathbb{1}_{\pi_2(P)}, F_3^{\varphi}\mathbb{1}_{A_3}\mathbb{1}_{\pi_3(P)}) \\
&\quad = \sum_{i=1}^{J} \Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(T_i)}\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1\setminus\overline{T}_i)}, F_2^{\varphi}\mathbb{1}_{A_2}\mathbb{1}_{\pi_2(P)}, F_3^{\varphi}\mathbb{1}_{A_3}\mathbb{1}_{\pi_3(P)}) \\
&\quad \lesssim \sum_{i=1}^{J} \mu(T_i)\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}\|_{L^\infty(S_1)} \prod_{j=2}^{3} \|\boldsymbol{F}_j\mathbb{1}_{A_j}\mathbb{1}_{\pi_j(P)}\|_{L^\infty(S^{\gamma}_{j,\delta A})} \\
&\quad \lesssim \min(\min_j \mu(M_j), \min_j \mu(\widetilde{N}_j))\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(P_1)}\|_{L^\infty(S_1)} \prod_{j=2}^{3} \|\boldsymbol{F}_j\mathbb{1}_{A_j}\mathbb{1}_{\pi_j(P)}\|_{L^\infty(S^{\gamma}_{j,\delta,A})}.
\end{aligned}$$

It finishes the proof of this reduction since $\min_j \mu(\widetilde{N}_j) \lesssim \min_j \mu_\delta(N_j)$. □

Since the estimate in Lemma 2.74 is translation and modulation invariant, it is implied by the following.

Lemma 2.75. *Let*

$$P_1 = \bigcap_{j=1}^{m} M_j \setminus L_1, \qquad P = \bigcap_{j=1}^{n} N_j \setminus L,$$

where $M_j, L_1 \in \mathbb{T}^\cup$, $N_j, L \in \mathbb{T}^\cup_\delta$,

$$B = A_1 \cap \pi_1(P_1), \qquad C = A \cap \pi_2(P).$$

Let a_j *be like in* (2.21), $T = T(0,0,s) \in \mathbb{T}$ *and set*

$$B^\theta(y) := B^{\theta+\beta_1}_{\pi_1(T)}(y), \qquad C^\theta(y) := C^{a_2\theta+\delta\beta_2}_{\pi_2(\rho(T))}(y).$$

Then

$$\int_\Theta \int_{|y|<s} \int_0^{s-|y|} F_1^\varphi(y, \frac{\theta+\beta_1}{t}, t) \mathbb{1}_{B^\theta(y)}(t) \prod_{j=2}^3 F_j^\varphi(y, \frac{a_j\theta+\delta\beta_j}{\delta t}, \delta t) \mathbb{1}_{C^\theta(y)}(\delta t) \frac{dt}{t}\, dy\, d\theta$$
$$\lesssim s \|\boldsymbol{F}_1 \mathbb{1}_B\|_{L^\infty(S_1)} \prod_{j=2}^3 \|\boldsymbol{F}_j \mathbb{1}_C\|_{L^\infty(S^\gamma_{j,\delta,A})}.$$

Remark 2.76. *Note that if* $\widetilde{C} = A \cap \pi_3(P \setminus L)$, *then* $C^\theta = \widetilde{C}^{a_3\theta+\delta\beta_3}_{\pi_3(\rho(T))}$.

Proof. First, restrict the outermost integral to $\theta \in \Theta^{(out)}$. Applying Hölder's inequality with exponents $\infty, 2, 2$ in (y, θ, t) and interpolating to obtain the γ factor in $\|\cdot\|_{S^{2,\gamma}_{j,\delta,C}}$ for $j = 2, 3$, the left hand side is bounded by the desired quantity.

In order to complete the proof of the lemma and in view of Lemma 2.46 it is enough to show that for every $\theta \in \Theta^{(in)}$ the double integral in (y,t) is bounded by

$$s \|\boldsymbol{F}_1 \mathbb{1}_B\|_{L^\infty(S_1)} \prod_{j=2}^3 \|\boldsymbol{F}_j\|_{L^\infty(S^\gamma_{j,\delta,C})},$$

This estimate follows from Lemma 2.77 combined with Lemma 2.58. □

For $\zeta \in \Phi^*$ and $j = 1,2,3$ we set $G_j^\zeta(y,t) := f_j * \zeta_t(y)$. Let ϑ, $\widetilde{\varphi}$ be such that

$$G_1^\vartheta(y,t) = F_1^\varphi(y, (\theta+\beta_1)t^{-1}, t),$$
$$G_j^{\widetilde{\varphi}}(y,t) = F_j^\varphi(y, (\alpha_j\theta+\delta\beta_j)t^{-1}, t) \quad \text{for } j = 2,3.$$

Let $\|\cdot\|_V, \|\cdot\|_{V^\infty} \|\cdot\|_{\boldsymbol{V}^2}$ and $\|\cdot\|_{RV_B}$ be defined as in (2.44), (2.45) and (2.46). Let $\widetilde{\varphi} = \overline{\varphi} + \psi$ be the decomposition given by Lemma 2.67. Set

$$G_1(y,t) := \sup_{\zeta \in \{\vartheta, \vartheta'\}} |G_1^\zeta(y,t)|,$$
$$G_j(y,t) := \sup_{\zeta \in \{\cdot\overline{\varphi}(\cdot), \overline{\varphi}', \psi, \psi'\}} |G_j^\zeta(y,t)|, \quad \text{for } j = 2,3.$$

The next lemma is the crucial "overlapping tree estimate". A priori we do not have enough cancellation to apply Hölder's inequality as we did in case $\theta \in \Theta^{(out)}$, so we first telescope restricted functions $F_2\mathbb{1}_C$, $F_3\mathbb{1}_C$ and then integrate by parts, which yields the boundary terms, that are ultimately controlled by the boundary sizes $\|\cdot\|_{R_{j,\delta,A}}$. We have the following.

Lemma 2.77. *Assume that* $\theta \in \Theta^{(in)}$. *It holds that*

$$\int_{|y|<s} \int_0^{s-|y|} G_1^\vartheta(y,t) \mathbb{1}_{B^\theta(y)}(t) G_2^{\widetilde{\varphi}}(y,\delta t) G_3^{\widetilde{\varphi}}(y,\delta t) \mathbb{1}_{C^\theta(y)}(\delta t) \frac{dt}{t}\, dy$$
$$\lesssim s \|G_1 \mathbb{1}_B\|_V \prod_{j=2}^3 (\|G_j\|^{1-\gamma}_{RV_{C^\theta}} \|G_j \mathbb{1}_{C^\theta}\|^\gamma_V + \delta^{1/2} \|G_j \mathbb{1}_{C^\theta}\|_{V^\infty} + \|G_j \mathbb{1}_{C^\theta}\|^{1-\gamma}_{\boldsymbol{V}^2} \|G_j \mathbb{1}_{C^\theta}\|^\gamma_V).$$

Proof of Lemma 2.77. To ease the notation we set $B(y) := B^\theta(y)$, $C(y) := C^\theta(y)$. Using Lemma 2.67 we decompose $\widetilde{\varphi} = \overline{\varphi} + \psi$, which implies that for $j = 2, 3$

$$G_j^{\widetilde{\varphi}} = G_j^{\overline{\varphi}} + G_j^{\psi}.$$

This way we obtain four integrals to bounds.

$$\int_{|y|<s} \int_0^{s-|y|} G_1^\vartheta(y,t) \mathbb{1}_{B(y)}(t) G_2^\psi(y,\delta t) G_3^\psi(y,\delta t) \mathbb{1}_{C(y)}(\delta t) \, \frac{dt}{t} \, dy \tag{2.73}$$

$$\int_{|y|<s} \int_0^{s-|y|} G_1^\vartheta(y,t) \mathbb{1}_{B(y)}(t) G_2^{\overline{\varphi}}(y,\delta t) G_3^{\overline{\varphi}}(y,\delta t) \mathbb{1}_{C(y)}(\delta t) \, \frac{dt}{t} \, dy \tag{2.74}$$

$$\int_{|y|<s} \int_0^{s-|y|} G_1^\vartheta(y,t) \mathbb{1}_{B(y)}(t) G_2^{\overline{\varphi}}(y,\delta t) G_3^\psi(y,\delta t) \mathbb{1}_{C(y)}(\delta t) \, \frac{dt}{t} \, dy \tag{2.75}$$

$$\int_{|y|<s} \int_0^{s-|y|} G_1^\vartheta(y,t) \mathbb{1}_{B(y)}(t) G_2^\psi(y,\delta t) G_3^{\overline{\varphi}}(y,\delta t) \mathbb{1}_{C(y)}(\delta t) \, \frac{dt}{t} \, dy \tag{2.76}$$

Note that the bound for (2.73) follows simply by applying Cauchy-Schwarz with exponents $\infty, 2, 2$ and interpolating to obtain the desired right hand side. We still have the other three terms to estimate.

We start with (2.74). In order to ease the notation, let us set $H_1 := G_1^\vartheta$, $\widetilde{H}_1 := G_1^{\varphi'}$; $H_2 := G_j^{\overline{\varphi}}$, $\overline{H}_2 := G_2^{\zeta'}$, where $\zeta(x) = x\overline{\varphi}(x)$ and $\widetilde{H}_2 := G_2^\zeta$; $H_3 := G_3^{\overline{\varphi}}$ and $\widetilde{H}_3 := G_3^{\overline{\varphi}'}$. Using Proposition 2.54 we rewrite $H_2 H_3 \mathbb{1}_C$ as

$$H_2(y,\delta t) H_3(y,\delta t) \mathbb{1}_{C(y)}(\delta t) = \int_0^{\delta t} \partial_u (H_2(y,u) H_3(y,u) \mathbb{1}_{C(y)}(u)) \, du,$$

so that using (2.58) up to a symmetry we have to control the two integrals

$$\int_{|y|<s} \int_0^{s-|y|} H_1(y,t) \mathbb{1}_{B(y)}(t) \int_0^{\delta t} H_2(y,u) H_3(y,u) \partial_u \mathbb{1}_{C(y)}(u) \, du \, \frac{dt}{t} \, dy \tag{2.77}$$

$$\int_{|y|<s} \int_0^{s-|y|} H_1(y,t) \mathbb{1}_{B(y)}(t) \int_0^{\delta t} \overline{H}_2(y,u) H_3(y,u) \mathbb{1}_{C(y)}(u) \, \frac{du}{u} \, \frac{dt}{t} \, dy \tag{2.78}$$

Concerning (2.77), applying (2.42) and changing the order of integration we have

$$\begin{aligned} &\int_{|y|<s} \int_0^{\delta(s-|y|)} \prod_{j=2}^3 H_j(y,u) \partial_u \mathbb{1}_{C(y)}(u) \int_{\delta^{-1}u}^{s-|y|} H_1(y,t) \mathbb{1}_{B(y)}(t) \, \frac{dt}{t} \, du \, dy \\ \leq &\int_{|y|<s} \sup_M |\int_M^s H_1(y,t) \mathbb{1}_{B(y)}(t) \, \frac{dt}{t}| \int_0^{\delta s} |\prod_{j=2}^3 H_j(y, c^\pm(y))| \, dy \end{aligned}$$

Applying Hölder's inequality with exponents $1/\gamma, 1/(1-\gamma)$ in y, applying Lemma 2.70 to H_1 and (2.69) to the product $H_2 H_3$ we obtain the desired inequality for (2.77). Now we proceed with (2.78). Using Proposition 2.55 and (2.57) we move the derivative from $\overline{H}_2$ to one of the other four factors obtaining another four integrals to bound

$$\int_{|y|<s}\int_0^{s-|y|}\widetilde{H}_1(y,t)\mathbb{1}_{B(y)}(t)\int_0^{\delta t}\widetilde{H}_2(y,u)H_3(y,u)\mathbb{1}_{C(y)}(u)\frac{u}{t}\frac{du}{u}\frac{dt}{t}\,dy \tag{2.79}$$

$$\int_{|y|<s}\int_0^{s-|y|}H_1(y,t)\partial_y\mathbb{1}_{B(y)}(t)\int_0^{\delta t}\widetilde{H}_2(y,u)H_3(y,u)\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \tag{2.80}$$

$$\int_{|y|<s}\int_0^{s-|y|}H_1(y,t)\mathbb{1}_{B(y)}(t)\int_0^{\delta t}\widetilde{H}_2(y,u)\widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)\,\frac{du}{u}\frac{dt}{t}\,dy \tag{2.81}$$

$$\int_{|y|<s}\int_0^{s-|y|}H_1(y,t)\mathbb{1}_{B(y)}(t)\int_0^{\delta t}\widetilde{H}_2(y,u)H_3(y,u)\partial_y\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \tag{2.82}$$

We shall now bound each of the above four integrals.

(2.79): we have for $u<\delta t$

$$H_3(y,u)\mathbb{1}_{C(y)}(u)\frac{u^{1/2}}{t^{1/2}}\lesssim\delta^{1/2}\|G_3\mathbb{1}_C\|_{V^\infty}$$

Moreover

$$\int_0^{\delta t}|\widetilde{H}_2(y,u)|\mathbb{1}_{C(y)}(u)\frac{u^{1/2}}{t^{1/2}}\,\frac{du}{u}$$

is dominated the convolution of $|\widetilde{H}_2(y,\cdot)|\mathbb{1}_{(0,\delta s)}$ with $\mathbb{1}_{(0,1)}t^{1/2}$ in the multiplicative group $\mathbb{R}_+$ endowed with measure $\frac{dt}{t}$. Applying Cauchy-Schwarz and Young's convolution inequality for $(\mathbb{R}_+,\frac{dt}{t})$ we therefore obtain that (2.79) is bounded by $\delta^{1/2}\|G_3\mathbb{1}_C\|_{V^\infty}$ times

$$\begin{aligned}&\int_{|y|<s}\int_0^{s-|y|}\widetilde{H}_1(y,t)\mathbb{1}_{B(y)}(t)\int_0^{\delta t}\widetilde{H}_2(y,u)\mathbb{1}_{C(y)}(u)\frac{u^{1/2}}{t^{1/2}}\,\frac{du}{u}\,\frac{dt}{t}\,dy\\ \lesssim&\int_{|y|<s}(\int_0^{s-|y|}|\widetilde{H}_1(y,t)\mathbb{1}_{B(y)}(t)|^2\frac{dt}{t})^{1/2}(\int_0^{\delta s}|\widetilde{H}_2(y,t)\mathbb{1}_{C(y)}(t)|^2\frac{dt}{t})^{1/2}\,dy\end{aligned}$$

Applying Cauchy-Schwarz in y and interpolating we obtain the desired inequality.

(2.80): it is dominated by

$$\int_0^s\int_{|y|<s-t}|H_1(y,t)\partial_y\mathbb{1}_{B(y)}(t)|\sup_{u\in(0,\delta t)}\delta|\widetilde{H}_2(y,u)H_3(y,u)\mathbb{1}_{C(y)}(u)|\,dy\,dt$$

We have

$$\delta^{1/2}|\widetilde{H}_2(y,u)\mathbb{1}_{C(y)}(u)|\lesssim\delta^{1/2}\|G_3\mathbb{1}_C\|_{V^\infty}$$

and similarly for H_3. The last thing to notice is that by (2.43)

$$\int_0^s\int_{|y|<s-t}|H_1(y,t)\partial_y\mathbb{1}_{B(y)}(t)|dy\,dt\lesssim s\|G_1\mathbb{1}_{A_1}\|_V.$$

(2.81): changing the order of integration, it is bounded by

$$\int_{|y|<s} \sup_M |\int_M^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}| \times \int_0^{\delta s} |\widetilde{H}_2(y,u)\widetilde{H}_3(y,u)|\mathbb{1}_{C(y)}(u)\,\frac{du}{u}\,dy.$$

Applying Hölder's inequality in y with exponents $1/\gamma, 1/(1-\gamma)$ and Lemma 2.70 to H_1, Lemma 2.73 to the product H_2H_3 it is bounded by the desired quantity.

(2.82): applying (2.43) and changing the order of integration it is bounded by

$$\int_{|y|<s} \sup_M |\int_M^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)\,\frac{dt}{t}||\widetilde{H}_2(y,c^{\pm}(y))H_3(y,c^{\pm}(y))|\,dy$$

Applying Hölder's inequality with exponents $1/\gamma, 1/(1-\gamma)$ in y and applying Lemma 2.70 to H_1 and (2.69) to the product $H_2\widetilde{H}_3$ we obtain the desired inequality.

Now we bound (2.76). Notice that (2.75) can be treated exactly the same way. Let $H_2 := G_2^{\psi}$, $\widetilde{H}_2 := G_2^{\psi'}$ $H_3 := G_3^{\overline{\varphi}}$, $\overline{H}_3 := H_3^{\vartheta'}$, where $\vartheta(x) = x\overline{\varphi}(x)$ and $\widetilde{H}_3 := H_3^{\vartheta}$. In this case we telescope only H_3, we apply to it Proposition 2.54 in the following way

$$H_3(y,\delta t)\mathbb{1}_{C(y)}(\delta t) = \int_0^{\delta t} \partial_u(H_3(y,u)\mathbb{1}_{C(y)}(u))\,du.$$

Additionally using (2.58) we have to control the two integrals

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t} H_3(y,u)\partial_u\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \quad (2.83)$$

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t} \overline{H}_3(y,u)\mathbb{1}_{C(y)}(u)\,\frac{du}{u}\,\frac{dt}{t}\,dy \quad (2.84)$$

Concerning (2.83), we first bound it applying (2.42) and Cauchy-Schwarz in t by

$$\int_{|y|<s}(\int_0^s |H_1(y,t)|^2\mathbb{1}_{B(y)}(t)\,\frac{dt}{t})^{1/2}(\int_0^s |H_2(y,\delta t)|^2\mathbb{1}_{C(y)}(\delta t)\,\frac{dt}{t})^{1/2}|H_3(y,c^{\pm}(y))|\,dy$$

Applying Hölder's inequality in with exponents $(1/\gamma, 2/(1-\gamma), 2/(1-\gamma))$ in y and applying (2.72) to G_1 and H_2, and using simple interpolation to control the term involving H_3 we bound (2.83) by the desired quantity. Now we proceed with (2.84). Applying Proposition 2.55 and using (2.57) we move derivative from $\overline{H}_3$ to one of the other five factors obtaining another five integrals to bound

$$\int_{|y|<s}\int_0^{s-|y|} \widetilde{H}_1(y,t)\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t} \widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)\frac{u}{t}\,\frac{du}{u}\,\frac{dt}{t}\,dy \quad (2.85)$$

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\partial_y\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t} \widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \quad (2.86)$$

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)\widetilde{H}_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t}\widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)\frac{u}{\delta t}\,\frac{du}{u}\,\frac{dt}{t}\,dy \quad (2.87)$$

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\partial_y\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t}\widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \quad (2.88)$$

$$\int_{|y|<s}\int_0^{s-|y|} H_1(y,t)\mathbb{1}_{B(y)}(t)H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\int_0^{\delta t}\widetilde{H}_3(y,u)\partial_y\mathbb{1}_{C(y)}(u)\,du\,\frac{dt}{t}\,dy \quad (2.89)$$

We shall now bound each of the above five integrals.

(2.85): We have for $u<\delta t$

$$H_2(y,\delta t)\mathbb{1}_{C(y)}(\delta t)\frac{u^{1/2}}{t^{1/2}}\lesssim \delta^{1/2}\|G_2\mathbb{1}_C\|_{V^\infty}$$

Note that

$$\int_0^{\delta t}|\widetilde{H}_3(y,u)|\mathbb{1}_{C(y)}(u)\frac{u^{1/2}}{t^{1/2}}\,\frac{du}{u}$$

is dominated the convolution of $|\widetilde{H}_3(y,\cdot)|\mathbb{1}_{(0,\delta(s-|y|))}$ with $\mathbb{1}_{(0,1)}t^{1/2}$ in the multiplicative group $\mathbb{R}_+$ endowed with measure $\frac{dt}{t}$. Applying Cauchy-Schwarz and Young's convolution inequality for $(\mathbb{R}_+,\frac{dt}{t})$ we therefore obtain that (2.85) is bounded by $\delta^{1/2}\|G_2\mathbb{1}_C\|_{V^\infty}$ times

$$\begin{aligned}&\int_{|y|<s}\int_0^{s}|\widetilde{H}_1(y,t)|\mathbb{1}_{B(y)}(t)\int_0^{\delta t}|\widetilde{H}_3(y,u)|\mathbb{1}_{C(y)}(u)\frac{u^{1/2}}{t^{1/2}}\,\frac{du}{u}\,\frac{dt}{t}\,dy\\ \lesssim&\int_{|y|<s}(\int_0^{s}|\widetilde{H}_1(y,t)\mathbb{1}_{B(y)}(t)|^2\frac{dt}{t})^{1/2}(\int_0^{\delta s}|\widetilde{H}_3(y,t)\mathbb{1}_{C(y)}(t)|^2\frac{dt}{t})^{1/2}\,dy\end{aligned}$$

Applying Cauchy-Schwarz in y and interpolating we obtain the desired inequality.

(2.86): using (2.43) it is dominated by

$$\int_{|y|<s}|H_1(y,b^{\pm}(y))|\sup_{u\in(0,\delta t)}\delta^{1/2}|H_2(y,u)\mathbb{1}_{C(y)}(u)|\sup_{u\in(0,\delta t)}\delta^{1/2}|\widetilde{H}_3(y,u)\mathbb{1}_{C(y)}(u)|\,dy$$

We have

$$\delta^{1/2}|\widetilde{H}_2(y,u)\mathbb{1}_{C(y)}(u)|\lesssim\delta^{1/2}\|G_2\mathbb{1}_{C(y)}\|_{V^\infty},$$

and similarly for H_3. The last thing to notice is that

$$\int_{|y|<s}|H_1(y,b^{\pm}(y))|dy\lesssim s\|G_1\mathbb{1}_B\|_V.$$

(2.87) changing variables $\delta t\mapsto t$ is bounded by

$$\|G_1\mathbb{1}_B\|_V\int_{|y|<s}\int_0^{\delta s}|\widetilde{H}_2(y,t)\mathbb{1}_{C(y)}(t)|\int_0^{t}|\widetilde{H}_3(y,u)|\mathbb{1}_{C(y)}(u)\frac{u}{t}\,\frac{du}{u}\,\frac{dt}{t}\,dy.$$

Applying Cauchy-Schwarz in t, observing that the integral involving $\widetilde{H}_3$ is bounded by the convolution of $H_3(y,t)\mathbb{1}_{(0,\delta s)}(t)$ with $t\mathbb{1}_{(0,1)}$ in the multiplicative group $(\mathbb{R}_+, \frac{dt}{t})$ and applying Young's convolution inequality, and finally Cauchy-Schwarz in y, we bound the above by the desired quantity using interpolation.

(2.88): since $1_{C(y)}(\delta t) = \mathbb{1}_{\delta^{-1}C(y)}(t)$, it is bounded by $\|G_1\mathbb{1}_B\|_V$ times

$$\delta^{-1}\int_{|y|<s} |H_2(y,c^{\pm}(y))| \int_0^{c^{\pm}(y)} |\widetilde{H}_3(y,u)|\mathbb{1}_{C(y)}(u)\,du\,\frac{\delta}{c^{\pm}(y)}\,dy$$

Applying Cauchy-Schwarz in u and then dividing and multiplying du by $c^{\pm}(y)$, it is bounded by

$$\int_{|y|<s} |H_2(y,c^{\pm}(y))|c^{\pm}(y)(\int_0^{c^{\pm}(y)} |\widetilde{H}_3(y,u)|^2\mathbb{1}_{C(y)}(u)\,\frac{du}{c^{\pm}(y)})^{1/2}\,\frac{1}{c^{\pm}(y)}\,dy$$

Applying Cauchy-Schwarz in y, the above integral is further estimated by

$$(\int_{|y|<s} |H_2(y,c^{\pm}(y))|^2\,dy)^{1/2}(\int_{|y|<s}\int_0^{\delta s} |\widetilde{H}_3(y,u)|^2\mathbb{1}_{C(y)}(u)\,\frac{du}{u}\,dy)^{1/2},$$

which together with interpolation, gives the desired bound.

(2.89): it is bounded exactly the same way as (2.83). This finishes the proof. □

2.5.4 Proof of Proposition 2.52

In this subsection we put all the previous results of this chapter together and prove Proposition 2.52. The main difficulty is to show it in the case when $(1/p_1, 1/p_2, 1/p_3)$ is in the neighbourhood of $(0,1,0)$, or symmetrically, in the neighbourhood of $(0,0,1)$. We remark that the proof can be considerably simplified if $(1/p_1, 1/p_2, 1/p_3)$ is in the neighbourhood of $(1,0,0)$, however here we present the argument that works for all cases. We record that similarly to [OT11], in the proof we decompose Λ according to the level sets of F_1 and then prove that the summands decay exponentially, what yields the desired inequality.

Proof of Lemma 2.52. In the proof we use the notation introduced in Section 2.3.2. By homogeneity we may assume that for $j \in \{1,2,3\}$

$$\nu(V_j)^{1/p_j}\|\boldsymbol{F}_j\mathbb{1}_{V_j\setminus W_j}\|_{L^\infty \L^{q_j}(S)} \le 1. \tag{2.90}$$

Let A_ε be as in (2.23). We set $A_j = A_\varepsilon \cap V_j \setminus W_j$ for $j = 1,2,3$, and $A = V_1 \cap A_2 \cap A_3$. All the implicit constants will be independent of δ, A and may depend on γ. Let $V \in \mathbb{D}^{\cup}$ be a covering of $V_1 \cap V_2 \cap V_3$ such that $\nu(V) \le \min(\nu(V_1), \nu(V_2), \nu(V_3))$. This requirement can be satisfied, since each of V_1 V_2, V_3 is a covering of $V_1 \cap V_2 \cap V_3$. Finally, we set $\overline{A} = V \setminus W_2 \setminus W_3$.

Hence, using (2.90) and by standard limiting argument as $\varepsilon \to 0$, it suffices to show

$$|\Lambda(F_1^{\varphi}\mathbb{1}_{A_1}, F_2^{\varphi}\mathbb{1}_{A_2}, F_3^{\varphi}\mathbb{1}_{A_3})| \lesssim 1.$$

Note that (2.90) and Lemma 2.48 imply that $\|\boldsymbol{F}_1\mathbb{1}_{A_1}\|_{L^\infty_{\mu_1}(S_1)} \le \nu(V_1)^{-1/p_1}$. Let us run the selection algorithm from Definition 2.43 and Remark 2.39 for $F_1\mathbb{1}_{A_1}$, over $k \in \mathbb{N}$, such that $E_k := \bigcup \mathcal{T}_k$ corresponds to the level $2^{-k}\nu(V_1)^{-1/p_1}$. Additionally defining $E_{-1} = \emptyset$ we have

$$\|\boldsymbol{F}_1\mathbb{1}_{A_1}\mathbb{1}_{\pi_1(E_k\setminus E_{k-1})}\|_{L^\infty_{\mu_1}(S_1)} \lesssim 2^{-k}\nu(V_1)^{-1/p_1}.$$

and $E_k \setminus E_{k-1}$ are pairwise disjoint. Set $N_{\mathcal{T}_k} = \sum_{\Delta T \in \mathcal{T}_k} \mathbb{1}_{I_T}$ to be the counting function of the forest $\mathcal{T}_k$. Using (2.90) and Lemma 2.51 we obtain (since $\mu(E_k) \le \|N_{\mathcal{T}_k}\|_{L^1}$)

$$\mu(E_k) \lesssim 2^{q_1 k}\nu(V_1)\nu(V_1)^{q_1/p_1}\|\boldsymbol{F}_1 \mathbb{1}_{A_1}\|^{q_1}_{Ł^{q_1}(S)} \lesssim 2^{q_1 k}\nu(V_1). \tag{2.91}$$

Denote $\rho(E_k) = \bigcup_{T \in \mathcal{T}_k} \rho(T)$ (see the definition, (2.25), (2.26)). Note that $E_k \setminus E_{k-1}$ are pairwise disjoint so we may split the trilinear form using (2.28) into

$$\begin{aligned}&|\Lambda(F_1^{\varphi}\mathbb{1}_{A_1}, F_2^{\varphi}\mathbb{1}_{A_2}, F_3^{\varphi}\mathbb{1}_{A_3})|\\ &\le \sum_{k \ge 0} |\Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(E_k \setminus E_{k-1})}\mathbb{1}_{A_1}, F_2^{\varphi}\mathbb{1}_{\pi_2(\rho(E_k))}\mathbb{1}_{A_2}, F_3^{\varphi}\mathbb{1}_{\pi_3(\rho(E_k))}\mathbb{1}_{A_3})|\end{aligned}$$

Fix $k \in \mathbb{Z}_+$. Applying Proposition 2.53 together with Proposition 2.10 and changing variables in the definition of the outer L^p norm we obtain

$$\begin{aligned}&|\Lambda(F_1^{\varphi}\mathbb{1}_{\pi_1(E_k \setminus E_{k-1})}\mathbb{1}_{A_1}, F_2^{\varphi}\mathbb{1}_{\pi_2(\rho(E_k))}\mathbb{1}_{A_2}, F_3^{\varphi}\mathbb{1}_{\pi_3(\rho(E_k))}\mathbb{1}_{A_3})|\\ &\lesssim \|\boldsymbol{F}_1\mathbb{1}_{\pi_1(E_k \setminus E_{k-1})}\mathbb{1}_{A_1}\|_{L^{t_1}_{\mu_1}(S_1)} \prod_{j=2}^{3} \|\boldsymbol{F}_j \mathbb{1}_{\pi_j(\rho(E_k))}\mathbb{1}_{\overline{A}}\|^{1-\gamma}_{L^{t_j(1-\gamma)}_{\mu_{j,\delta,A}}(S_{j,\delta,A})} \|\boldsymbol{F}_j \mathbb{1}_{\overline{A}}\|^{\gamma}_{L^{\infty}_{\mu_j}(S_j)}\end{aligned} \tag{2.92}$$

for any $t_i \in [1,\infty]$ such that $\sum_{i=1}^{3} t_i^{-1} = 1$. Using Lemma 2.50 and (2.91), it follows that

$$\|\boldsymbol{F}_1\mathbb{1}_{\pi_1(E_k \setminus E_{k-1})}\mathbb{1}_{A_1}\|_{L^{t_1}_{\mu_1}(S_1)} \lesssim 2^{q_1 k(1/t_1 - 1/q_1)}\nu(V_1)^{1/t_1 - 1/p_1}. \tag{2.93}$$

For the terms involving F_2, as long as $t_2 \in (2, q_2]$ we use Proposition 2.31, Lemma 2.50 and Lemma 2.48 to obtain that

$$\begin{aligned}&\|\boldsymbol{F}_2 \mathbb{1}_{\pi_2(\rho(E_k))}\mathbb{1}_{\overline{A}}\|^{1-\gamma}_{L^{t_2(1-\gamma)}_{\mu_2}(S_{2,\delta,A})} \|\boldsymbol{F}_2 \mathbb{1}_{\overline{A}}\|^{\gamma}_{L^{\infty}_{\mu_2}(S_2)}\\ &\lesssim \|\boldsymbol{F}_2 \mathbb{1}_{\pi_2(\rho(E_k))}\mathbb{1}_{\overline{A}}\|^{1-\gamma}_{L^{t_2(1-\gamma)}_{\mu_2}(S_2)} \|\boldsymbol{F}_2 \mathbb{1}_{\overline{A}}\|^{\gamma}_{L^{\infty}_{\mu_2}(S_2)}\\ &\lesssim \mu(V \cap \pi_2(\rho(E_k)))^{1/t_2 - (1-\gamma)/q_2}\nu(V)^{(1-\gamma)/q_2}\|\boldsymbol{F}_2 \mathbb{1}_{\pi_2(\rho(E_k))}\mathbb{1}_{\overline{A}}\|^{1-\gamma}_{Ł^{q_2}_{\mu_2}(S)} \|\boldsymbol{F}_2 \mathbb{1}_{\overline{A}}\|^{\gamma}_{L^{\infty}_{\mu_2}(S_2)}\\ &\lesssim \mu(V \cap \pi_2(\rho(E_k)))^{1/t_2 - (1-\gamma)/q_2}\nu(V)^{(1-\gamma)/q_2}\|\boldsymbol{F}_2 \mathbb{1}_{\overline{A}}\|_{Ł^{q_2}_{\mu_2}(S_2)}\\ &\lesssim \mu(V \cap \pi_2(\rho(E_k)))^{1/t_2 - (1-\gamma)/q_2}\nu(V)^{(1-\gamma)/q_2 - 1/p_2}.\end{aligned}$$

Let $V = \bigcup_{m=1}^{\infty} D_m$, where D_m's are given by Lemma 2.29. Then, using Lemma 2.29

$$\begin{aligned}\mu(V \cap \pi_2(\rho(E_k))) &\lesssim \sum_{T \in \Phi_k} \mu(V \cap \pi_2(\rho(T))) \le \sum_{T \in \Phi_k}\sum_{m=1}^{\infty} \mu(D_m \cap \pi_2(\rho(T)))\\ &\lesssim \sum_{T \in \Phi_k}\sum_{m=1}^{\infty} |I_{D_m} \cap I_{\pi_2(\rho(T))}| \lesssim \sum_{T \in \Phi_k} |I_V \cap I_{\pi_2(\rho(T))}| = \|N_{\pi_2(\rho(\Phi_k))}\mathbb{1}_{I_V}\|_{L^1},\end{aligned}$$

where $I_V = \bigcup_{m=1}^{\infty} I_{D_m}$. Using $N_{\Phi_k} = N_{\pi_2(\rho(\Phi_k))}$ and applying Hölder's inequality for $1 \le p \le \infty$ we obtain that the last display is bounded by

$$\nu(V)^{1-1/p}\|N_{\Phi_k}\|_p = \nu(V)\frac{\|N_{\Phi_k}\|_p}{\nu(V)^{1/p}}.$$

By Lemma 2.51 and (2.90) we have for $1 \le p < \infty$

$$\|N_{\Phi_k}\|_p \lesssim_p \nu(V_1)^{1/p}2^{q_1 k}\nu(V_1)^{q_1/p_1}\|\boldsymbol{F}_1\mathbb{1}_{A_1}\|^{q_1}_{Ł^{q_1}_{\mu_1}(S_1)} \le 2^{q_1 k}\nu(V_1)^{1/p}.$$

Thus, for any $p \in [1,\infty)$, we have

$$\begin{aligned}
&\|\boldsymbol{F}_2 \mathbb{1}_{\boldsymbol{E}_k} \mathbb{1}_{\overline{A}}\|^{1-\gamma}_{L^{t_2(1-\gamma)}_{\mu_2}(\boldsymbol{S}_{2,\delta,A})} \|\boldsymbol{F}_2 \mathbb{1}_{\overline{A}}\|^{\gamma}_{L^{\infty}_{\mu_2}(S_2)} \\
&\lesssim 2^{q_1 k(1/t_2-(1-\gamma)/q_2)} \left(\frac{\nu(V)}{\nu(V_1)}\right)^{-\left(1/t_2-(1-\gamma)/q_2\right)/p} \nu(V)^{1/t_2-1/p_2}.
\end{aligned} \tag{2.94}$$

The same result holds for the term with F_3. Putting the bounds (2.94) and (2.93) into (2.92) we obtain using $\sum_{j=1}^{3} 1/t_j = \sum_{j=1}^{3} 1/p_j = 1$

$$\begin{aligned}
&|\Lambda(F_1^{\varphi} \mathbb{1}_{A_1} \mathbb{1}_{\pi_1(E_k \setminus E_{k-1})}, F_2^{\varphi} \mathbb{1}_{A_2} \mathbb{1}_{\pi_2(\rho(E_k))}, F_3^{\varphi} \mathbb{1}_{A_3} \mathbb{1}_{\pi_3(\rho(E_k))})| \\
&\lesssim 2^{q_1 k\left(1-1/q_1-1/q_2-1/q_3+\gamma(1/q_2+1/q_3)\right)} \\
&\qquad \times \left(\frac{\nu(V)}{\nu(V_1)}\right)^{(1/p_1-1/t_1)} \left(\frac{\nu(V)}{\nu(V_1)}\right)^{-\left(1/t_2+1/t_3-(1-\gamma)\left(1/q_2+1/q_3\right)\right)/p}.
\end{aligned}$$

As long as $\gamma \ll 1$ and since $\sum_j 1/q_j > 1$, we have that $q_1(1-1/q_1-1/q_2-1/q_3+\gamma(1/q_2+1/q_3)) < 0$ that makes the above expression summable over $k \in \mathbb{N}$. If $t_1 > p_1$ and $p \in (1,\infty]$ is chosen large enough then the exponent of $\frac{\nu(V)}{\nu(V_1)}$ is positive. Since $\nu(V) \leq \nu(V_1)$, this concludes the proof. □

Chapter 3

Uniform bounds for Walsh bilinear Hilbert transform in local L^1

3.1 Introduction

The goal of this chapter is to prove the uniform bounds for a Walsh model of the bilinear Hilbert transform modularizing it as a multilinear iterated outer L^p estimate uniform in the degeneration parameter and the Walsh iterated embedding. This is elaboration on Chapter 2. We record that considering Walsh models is a well established way of studying multilinear forms as well as explaining the key ideas of technically more involved statements in time-frequency analysis.

Uniform bounds for the Walsh model were already proven by Oberlin and Thiele in [OT11]. Our argument, although motivated by their approach, involves a number of refinements. For convenience of the reader, below we introduce the Walsh model that we will be dealing with and state the main result, which we already did in the introduction of the thesis.

We call a tile the Cartesian product $I \times \omega$, where $I, \omega \subset \mathbb{R}_+$ are dyadic intervals. The Walsh phase plane is $\mathbb{R}_+^2$ together with the set of tiles $\mathbb{X}$. The L^2 normalized wave packets associated with tiles are defined recursively via the following identities

$$\varphi_{I\times[0,|I|^{-1})} = |I|^{-1/2}\mathbb{1}_I(x), \qquad \varphi_{J^-\times\omega} + \varphi_{J^+\times\omega} = \varphi_{J\times\omega^-} + \varphi_{J\times\omega^+},$$

for any dyadic intervals $I, J, \omega \subset \mathbb{R}_+$ with $|J||\omega| = 2$, where J^- and J^+ are dyadic children of J.

Given a function $f \in L^p(\mathbb{R})$ we associate it with the embedded function via

$$F(f)(P) := \langle f, \varphi_P \rangle := \int_{\mathbb{R}} f(x)\varphi_P(x)\,dx.$$

Let $F_j \colon \mathbb{X} \to \mathbb{R}$, $j = 1, 2, 3$. We indicate the dyadic sibling of a dyadic interval I by $I^{\circlearrowleft}$ and by $P^{\circlearrowleft}$ the tile $I_P \times \omega_p^{\circlearrowleft}$. The trilinear form on the embedded functions associated to the Walsh bilinear Hilbert transform is formally given for $L \in \mathbb{Z}_+$ by

$$\Lambda_L(F_1, F_2, F_3) := \sum_{P\in\mathbb{X}} F_1(P^{\circlearrowleft}) \sum_{Q\in P^L} F_2(Q)F_3(Q)h_{I_P}(c(I_Q)),$$

where h_{I_P} is the L^2 normalized Haar function, $c(I_Q)$ is the center of of the interval I_Q and

$$P^L = \{Q \in \mathbb{X}\colon I_Q \subset I_P,\ |I_Q| = 2^{-L}|I_P|,\ \omega_Q = 2^L \omega_P\},$$

where $2^L\omega = [2^L a, 2^L b)$ for an interval $\omega = [a, b)$. Here is the main result of the chapter.

Theorem 3.1. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$ and $1/q_1 + 1/q_2 + 1/q_3 > 1$ with $2 < q_1, q_2, q_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $L \in \mathbb{Z}_+$ and all triples of Schwartz functions on $\mathbb{R}$, f_1, f_2, f_3*

$$|\Lambda_L(F(f_1), F(f_2), F(f_3))| \le C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|F(f_j)\|_{L^{p_j} Ł^{q_j}(S)}. \tag{3.1}$$

On the right hand side of (3.1) are iterated outer L^p norms developed in [Ura16] that we define precisely in Section 3. For embedded functions, they can be controlled using the following Walsh iterated embedding theorem, proved by Uraltsev in [Ura17].

Theorem 3.2. *Let $1 < p \le \infty$, $\max(p', 2) < q \le \infty$. Then*

$$\|F(f)\|_{L^p Ł^q(S)} \le C_{p,q} \|f\|_{L^p(\mathbb{R})}.$$

We prove Theorem 3.1 in the framework of outer L^p spaces using a counterpart of multilinear Marcinkiewicz interpolation for outer L^p spaces, Proposition 2.10, as we were not able to verify the assumption of the outer Hölder inequality developed in [DT15].

The trilinear form Λ_L is strongly related with the Walsh bilinear Hilbert transform considered by Oberlin and Thiele in [OT11]. In the following we introduce a model very similar to the one that they investigated. For a $P \in \mathbb{X}$ we define the phase plane projections

$$\Pi_P f(x) := \langle f, \varphi_P \rangle \varphi_P(x), \qquad \Pi_{P^L} f(x) := \sum_{Q \in P^L} \Pi_Q(x).$$

The Walsh model of the bilinear Hilbert transform is defined as

$$\mathrm{BHF}_L(f_1, f_2, f_3) := \int \sum_{P \in \mathbb{X}} \widetilde{\varphi}_P(x) \Pi_{P^\circ} f_1(x) \Pi_{P^L} f_2(x) \Pi_{P^L} f_3(x) dx.$$

In [OT11], Oberlin and Thiele proved the following.

Theorem 3.3. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $L \in \mathbb{Z}_+$ and all triples of Schwartz functions on $\mathbb{R}$, f_1, f_2, f_3 the inequality*

$$|\mathrm{BHF}_L(f_1, f_2, f_3)| \le C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R})}$$

holds.

Their theorem works in a wider range of exponents, however, as it was highlighted in their work, the most difficult case is when the exponents are in the neighbourhood of $(1/p_1, 1/p_2, 1/p_3) = (0, 1, 0)$ or $(1/p_1, 1/p_2, 1/p_3) = (0, 0, 1)$. Theorem 3.1 coupled with Theorem 3.2 implies Theorem 3.3, since

$$\mathrm{BHF}_L(f_1, f_2, f_3) = \Lambda_L(F(f_1), F(f_2), F(f_3)).$$

In Section 3.2 we introduce the outer L^p spaces on $\mathbb{X}$. In Section 3.3 we prove auxiliary inequalities, including domination of L-dependent outer L^p norms that L-independent norms as well as quasi-monotonicity of the iterated L^p norms. In Section 3.4 we prove Theorem 3.1. In the appendix we recall the properties of the Walsh wave packets.

3.2 Outer L^p spaces in time-frequency-scale space

In Chapter 3, Section 2.3 we introduced abstract outer measures and sizes, while in this section we introduce our particular choice of these objects that we will be working with till the end of this chapter. From now on, we assume that $\mathbb{X}$ is the set of tiles, $\mathcal{B}(X)$ is the set of (Borel) functions on $\mathbb{X}$.

3.2.1 Outer measures in time-frequency-scale space

First, we define the generating collection on $\mathbb{X}$, which we call trees. They can be seen as time-frequency localized subsets of the time-frequency-scale space $\mathbb{X}$. We shall need their variants parametrized by the parameter L, similarly as in [OT11]. They can be seen as Walsh analogues of the trees introduced in Chapter 2.

Definition 3.4 (Trees and measures)**.** *For $I \times \omega \in \mathbb{X}$ and $L \in \mathbb{N}$ we define an L-tree as follows:*

$$T(I, \omega) := \{P \in \mathbb{X} \colon I_P \subset I,\, \omega_P \cup \omega_P^{\circlearrowleft} \supset \omega\} \cap \{P \in \mathbb{X} \colon |I_P| \leq 2^{-L}|I|\}.$$

Moreover we set

$$T^{(ov)} := \{P \in T \colon \omega_P \supset \omega\}, \qquad\qquad T^{(lac)} = T \setminus T^{(ov)}.$$

We set $I_T = I$ and $\omega_T = \omega$ if $T = T(I, \omega)$ and we denote the family of L-trees with $\mathbb{T}_L$ and with $\mathbb{T}_L^{\cup}$ the family of countable unions of $T \in \mathbb{T}_L$. The generating pre-measure is given by

$$\overline{\mu}^L(T) := |I_T|$$

and it generates μ^L. Additionally, we set $\mathbb{T} := \mathbb{T}_0$ and $\mu := \mu^0$.

Analogously as in Chapter 2 we introduce the generating collection of strips that we shall use as the generating collection for the iterated outer L^p spaces, which were developed in [Ura16].

Definition 3.5 (Strips and measures)**.** *We define time-scale strips as subsets of $\mathbb{X}$ and the associated pre-measure by setting*

$$D(I) = \{P \in \mathbb{X} \colon I_P \subset I\} \qquad \overline{\nu}\big(D(I)\big) := |I|.$$

We denote $I_D = I$ if $D = D(I)$, with $\mathbb{D}$ the set of strips and with $\mathbb{D}^{\cup}$ the family of countable unions of $D \in \mathbb{D}$.

Remark 3.6. *Note that for any $D \in \mathbb{D}$, $D \in \mathbb{T}^{\cup}$.*

Note that strips can be identified with time localized subsets of the time-scale space, as for a tile $P = I \times \omega$ and a strip D, the condition $P \in D$ is independent of the frequency component ω.

3.2.2 Sizes in time-frequency-scale space

In this subsection we introduce sizes, which we shall be working with. Before we do that, we define the top of a tree, which is essentially the boundary of a tree intersected with a set of the form $V \setminus W$, where $V, W \in \mathbb{T}^{\cup}$.

Definition 3.7 (Order on $\mathbb{X}$)**.** *We write $P \leq Q$ for $P, Q \in \mathbb{X}$ with $I_P \subset I_Q$ and $\omega_P \supset \omega_Q$.*

Definition 3.8 (Tops)**.** *Let $L \in \mathbb{N}$, $T \in \mathbb{T}_L$, $A = K \setminus M$, where $K, M \in \mathbb{T}^{\cup}$. Let $\mathcal{P}^1_{A,T}$ be the set of maximal tiles contained in $A \cap T^{(ov)}$ and $\mathcal{P}^2_{A,T}$ be the set of minimal tiles contained in $A \cap T^{(ov)}$. We define the top of T with respect to A as*

$$\mathrm{Top}_A(T) := \mathcal{P}^1_{A,T} \cup \mathcal{P}^2_{A,T}.$$

We think of tops coming from $\mathbb{D}^{\cup}$ as being "rough", as they behave essentially like 0-trees (they are unions of 0-trees), as opposed to L-trees for $L > 0$, which are "smooth".

We introduce the sizes for functions $F \in \mathcal{B}(\mathbb{X})$. Except for more standard S^∞ and S^2 sizes, which are additionally parametrized by the parameter L, the key difference comparing to previous works on the Walsh uniform bounds [OT11], [War15] is that we use what we call top sizes, which control the contribution from the rough boundary introduced by strips.

Definition 3.9 (L-sizes in time-frequency-scale space)**.** *Let $F \in \mathcal{B}(\mathbb{X})$ and $T \in \mathbb{T}_L$. Define*

$$\|F\|_{S^{2,L}(T)} := \left(\frac{1}{|I_T|} \sum_{P \in T^{(lac)}} |F(P)|^2 \right)^{1/2},$$

$$\|F\|_{S^{\infty,L}(T)} := |I_{P_T}|^{-1/2} \, |F(P_T)|.$$

For $L = 0$, we set $S^2 := S^{2,0}$, $S^\infty := S^{\infty,0}$ and

$$\|F\|_{S(T)} := \|F\|_{S^2(T)} + \|F\|_{S^\infty(T)},$$

For $L \in \mathbb{Z}_+$ and $A = K \setminus L$, $K, L \in \mathbb{T}^{\cup}$ we define

$$\|F\|_{S_A^{\mathrm{Top},L}(T)} := \left(\frac{1}{|I_T|} \sum_{P \in \mathrm{Top}_A(T)} |F(P)|^2 \right)^{1/2},$$

$$\|F\|_{S_A^L(T)} := \|F\|_{S_A^{\mathrm{Top},L}(T)} + \|F \mathbb{1}_A\|_{S^{2,L}(T)}.$$

Next, we recall the definition of the Walsh iterated sizes, which were originally introduced in [Ura16]. For a given function on $\mathbb{X}$, they are essentially the supremum of its outer L^q averages over strips.

Definition 3.10 (Iterated sizes in time-scale space)**.** *Given a function $F \colon \mathbb{X} \to \mathbb{R}$, $0 < q < \infty$ and outer measure μ on $\mathbb{X}$ and size $\| \cdot \|_S$ we define the iterated size as*

$$\|F\|_{Ł^q_\mu(S)} := \sup_{D \in \mathbb{D}} \nu(D)^{-1/q} \|F \mathbb{1}_D\|_{L^q_\mu(S)}.$$

Remark 3.11. *If $V \in \mathbb{D}^{\cup}$, we may assume without loss of generality that $V = \bigcup_{m=1}^\infty D_m$, where D_m's are pairwise disjoint and $\in \mathbb{D}$. This is because $D \in \mathbb{D}$ are based on dyadic intervals, meaning that for any two $D, D' \in \mathbb{D}$ we either have $D \cap D' = \emptyset$ or one is contained in the other.*

3.2.3 Notation

From now on we fix $L \in \mathbb{Z}_+$ and set $\boldsymbol{S}_A := S_A^L$ and $\boldsymbol{\mu} := \mu^L$. In order to ease the notation we also set

$$\|F_1\|_{L^p(S)} := \|F_1\|_{L^p_\mu(S)}, \qquad \|F_1\|_{L^p(Ł^q(S))} := \|F_1\|_{L^p_\nu(Ł^q_\mu(S))}$$

and for $j = 2, 3$

$$\|F_j\|_{L^p(\boldsymbol{S}_A)} := \|F_j\|_{L^p_{\mu^L}(\boldsymbol{S}_A)}, \qquad \|F_j\|_{L^p(Ł^q(S))} := \|F_j\|_{L^p_\nu(Ł^q_\mu(S))}.$$

3.3 Inequalities for outer L^p spaces on $\mathbb{X}$

3.3.1 Outer L^p comparison

Throughout this subsection we assume that

$$\|F\|_{L^2_\mu(S)}, \|F\|_{L^\infty_\mu(S)} < \infty \tag{3.5}$$

and assume that $\mathbb{X}$ is the set of tiles with

$$\omega_P \subset [0, N) \tag{3.6}$$

for a large $N \in \mathbb{Z}_+$. Since all bounds in this section are be independent of N, we may pass to the limit to obtain the result of this section for the whole set of tiles.

Let $A = V \setminus W$ for $V, W \in \mathbb{D}^\cup$ be fixed throughout this subsection. Our main result of this section is the following lemma that lets us dominate $\|F\|_{L^p(\boldsymbol{S}_A)} \lesssim \|F\|_{L^p(S)}$. Later it lets us apply the iterated embedding theorem for S from [Ura17]. The proof is essentially a combination of the Bessel inequality for the outer L^2 space together with dominating the measure of the level set by an ℓ^2 sum of coefficients coming from pairwise disjoint tiles.

Lemma 3.12 (Comparison of outer L^p spaces)**.** *The following inequalities hold for all functions $F \in \mathcal{B}(\mathbb{X})$:*

$$\|F\|_{L^q(\boldsymbol{S}_A)} \lesssim_q \|F \mathbb{1}_A\|_{L^q(S)} \qquad \forall q \in (2, \infty]$$
$$\|F\|_{L^{2,\infty}(\boldsymbol{S}_A)} \lesssim \|F \mathbb{1}_A\|_{L^2(S)}$$

with a constant that is dependent on q but independent of $L \in \mathbb{Z}_+$, F.

Remark 3.13. *It is easy to check that $\boldsymbol{\mu} \geq \mu$ while $\|\cdot\|_{L^\infty(\boldsymbol{S})} \leq \|\cdot\|_{L^\infty(S)}$. From these two conditions no non-trivial relation between $\|\cdot\|_{L^q(\boldsymbol{S}_A)} := \|\cdot\|_{L^q_{\boldsymbol{\mu}}(\boldsymbol{S}_A)}$ and $\|\cdot\|_{L^q(S)} := \|\cdot\|_{L^q_\mu(S)}$ can be deduced.*

The main advantage of Lemma 3.12 is that we do not have to prove any uniform iterated embedding and can apply Theorem 3.2 as a black box.

Definition 3.14 (Strongly disjoint)**.** *We call a subset $\mathcal{P}$ of $\mathbb{X}$ strongly disjoint if for any two distinct $P, P' \in \mathcal{P}$, $P \cap P' = \emptyset$.*

The following lemma will be crucial for us in this subsection. It can be thought of as the Bessel inequality for the outer L^2 space.

Lemma 3.15 (Outer L^2 sizes of strongly disjoint sets)**.** *Let $\mathcal{P} \subset \mathbb{X}$ be a strongly disjoint set of tiles. Then*

$$\|\mathbb{1}_\mathcal{P} F\|_{l^2(\mathbb{X})} \lesssim \|\mathbb{1}_\mathcal{P} F\|_{L^2(S)}$$

Proof. Using Lemma 2.8 decompose $\mathcal{P}$ into $\bigcup_k \bigcup_{\Delta T \in \Phi_k} \Delta T$ for $F\mathbb{1}_\mathcal{P}$, where ΔT are pairwise disjoint convex trees of the collection $\mathbb{T}$ and Φ_k corresponds to the 2^k-level set of $F\mathbb{1}_\mathcal{P}$. Define

$$\mathcal{P}^{(lac)} := \mathcal{P} \cap \bigcup_k \bigcup_{\Delta T \in \Phi_k} \Delta T^{(lac)}, \qquad \mathcal{P}^{(ov)} := \mathcal{P} \cap \bigcup_k \bigcup_{\Delta T \in \Phi_k} \Delta T^{(ov)}$$

First of all, using only pairwise disjointness of ΔT

$$\|\mathbb{1}_{\mathcal{P}^{(lac)}} F\|^2_{l^2(\mathbb{X})} \leq \sum_k \sum_{\Delta T \in \Psi_k} |I_T| \|\mathbb{1}_{\Delta T} \mathbb{1}_\mathcal{P} F\|^2_{S(T)}$$

while for the latter we have that

$$\begin{aligned}\|\mathbb{1}_{\mathcal{P}^{(ov)}}F\|^2_{l^2(\mathbb{X})} &\lesssim \sum_k \sum_{\Delta T\in\Psi_k} \sum_{\substack{P\in\mathcal{P}^{(ov)}\\ P\in\Delta T^{(ov)}}} |I_P|\frac{|F(P)|^2}{|I_P|}\\ &\leq \sum_k \sum_{\Delta T\in\Psi_k} |I_T|\|\mathbb{1}_{\Delta T}\mathbb{1}_{\mathcal{P}}F\|^2_{S(T)} \lesssim \|\mathbb{1}_{\mathcal{P}}F\|^2_{L^2(S)}.\end{aligned}$$

This concludes the proof. □

We introduce two auxiliary selection algorithms, one for $\boldsymbol{S}_A^{\mathrm{Top}}$ and the other one for $\boldsymbol{S}_A^2$. Similar procedures are usually used in the context of proving embedding theorems, see for example [LT97], [DT15]. The generated collection of forests has intrinsically very nice disjointness properties that we shall exploit in what follows.

Definition 3.16 (Top-selection algorithm)**.** *Initially $\mathcal{P}=\emptyset$ and $\mathbb{X}_0\subset\mathbb{X}$. In the n-th step of the procedure we choose a tree $T_n\in\mathbb{T}_L$ such that*

$$\|F\mathbb{1}_{\mathbb{X}_n}\|_{\boldsymbol{S}_A^{\mathrm{Top}}(T_n)}>\lambda \tag{3.7}$$

and s is maximal. This is possible due to Lemma 3.15, since the tiles in $\mathrm{Top}_A(T)$ can be split into two pairwise disjoint collection of tiles. As a result we obtain an upper bound for $\boldsymbol{\mu}(T)$

$$\boldsymbol{\mu}(T)\leq\lambda^{-2}\|F\mathbb{1}_{\mathrm{Top}_A(T)\cap\mathbb{X}_n}\|^2_{\ell^2(\mathbb{X})}\lesssim\lambda^{-2}\|F\|^2_{L^2(S)}<\infty.$$

We then add $S_n=\mathrm{Top}_A(T_n)\cap\mathbb{X}_n$ to the set $\mathcal{P}$ and set $\mathbb{X}_{n+1}:=\mathbb{X}_n\setminus T_n$ and iterate the procedure. It will terminate, since by the assumption (3.6)*, there is a lower bound on $\boldsymbol{\mu}(T)$; together with* (3.9) *and* (3.5) *there can be only finitely many $T\in\mathbb{T}_L$ satisfying* (3.7)*.*

In the following definition we denote we with ω^-,ω^+, the left and the right sibling of a dyadic interval $\omega\subset\mathbb{R}_+$, respectively. If such sibling does not exists, we set $\omega^\pm=\emptyset$.

Definition 3.17 (2-selection algorithm)**.** *Initially $\mathcal{P}=\emptyset$ and $\mathbb{X}_0\subset\mathbb{X}$. For a tree $T(I,\omega)\in\mathbb{T}_L$ let*

$$T^+(I,\omega)=\{P\in T\colon \omega_T\subset\omega_P^-\}.$$

In the n-th step of the procedure we proceed as follows: if there exists a tree $T\in\mathbb{T}_L$ with

$$\|F\mathbb{1}_{T^+}\mathbb{1}_{\mathbb{X}_n}\|_{\boldsymbol{S}_A^2(T)}>\lambda \tag{3.8}$$

We select a tree $T_n(I,\omega)$ which maximizes $\boldsymbol{\mu}(T)$ for the maximal possible value of ξ, where ξ is the middle of ω. This requirement can be satisfied, since there exists an upper bound for $\boldsymbol{\mu}(T)$: the tiles in $T^{(lac)}$ are pairwise disjoint, thanks to Lemma 3.15 we obtain

$$\boldsymbol{\mu}(T)\leq\lambda^{-2}\|F\mathbb{1}_{T^{(lac)}\cap\mathbb{X}_n}\|^2_{\ell^2(\mathbb{X})}\lesssim\lambda^{-2}\|F\|^2_{L^2(S)}<\infty.$$

This in turn implies that $|\omega|$'s in question are bounded from below, hence all possible ξ's are in a discrete set. Together with the assumption (3.6) *there must exist ω for which ξ attains its maximum. We add $S_n=T_n^+\cap\mathbb{X}_n$ to $\mathcal{P}$ and set $\mathbb{X}_{n+1}:=\mathbb{X}_n\setminus T_n$. We iterate the procedure until there are no more trees satisfying* (3.8)*. It will terminate, since by the assumption* (3.6)*, there is a lower bound on $\boldsymbol{\mu}(T)$; together with* (3.9) *and* (3.5) *there can be only finitely many $T\in\mathbb{T}_L$ satisfying* (3.8)*.*

Analogously we define the selection algorithm for

$$T^-(I,\xi) = \{(P \in T \colon \omega_T \subset \omega_P^+\},$$

with the only difference that we select trees $T(I,\omega)$ that maximize $\boldsymbol{\mu}(T)$ for the minimal possible value of ξ.

Since the selection algorithms for T^+ and T^- are entirely symmetric in the proof we will be working with only one of them.

Definition 3.18 (Selection algorithm)**.** *The selection algorithm at level λ consists of running the* Top*-selection algorithm at level λ and then 2-selection algorithm at level λ starting with $\mathbb{X}_0 = \mathbb{X}_M$, where is M is the the number of the last iteration of the* Top*-selection algorithm.*

The described algorithm yields a collection of trees $\Phi_\lambda = \Phi_\lambda^\infty \cup \Phi_\lambda^2$, where Φ_λ^∞, Φ_λ^2 are trees selected while running Top*-selection and 2-selection algorithm respectively. It also yields a collection of selected tiles $\mathcal{P}_\lambda$.*

Lemma 3.19. *If $\mathcal{P} \subset \mathbb{X}$ comes from the selection algorithm given in Definition 3.18 at level $\lambda > 0$, then $\mathcal{P}$ is strongly disjoint.*

Remark 3.20. *In particular an application of Lemma 3.15 implies that, if T_n are the selected trees, $n = 0, 1, 2, ..,$ then they satisfy for any $M \in \mathbb{N}$*

$$\sum_{n=0}^{M} \mu(T_n) \lesssim \lambda^{-2} \|F \mathbb{1}_A\|^2_{L^2(S)}. \tag{3.9}$$

Proof. We prove this lemma by contradiction.

First assume that $\mathcal{P}$ comes from the Top-selection algorithm Definition 3.16. Suppose that there exist $P, P' \in \mathcal{P}$, such that $P \in \mathrm{Top}_A(T)$, $P' \in \mathrm{Top}_A(T')$ and $P \cap P' \neq \emptyset$, where T, T' are two distinct selected trees. Without loss of generality suppose that T was selected earlier than T'. Then it is not possible $I_{P'} \subset I_P$, because this would mean that $P' \in T$, which is a contradiction. If this is not the case, then we necessarily have $|I_P| < |I_{P'}|$. But since $I_P, I_{P'}$ are members of the partition of I_T, I_T' generated by V, respectively, this would imply that $|I_T| < |I_{T'}|$. This would imply that T' was selected earlier. Contradiction.

Now assume that $\mathcal{P}$ comes from the 2-selection algorithm. Without loss of generality assume that $P \in T$, $P' \in T'$, $P \cap P' \neq \emptyset$ and T was selected earlier than T'. It is not possible that $I_{P'} \subset I_P$, since this would mean $P' \in T$ was being selected, which is a contradiction. Otherwise $\omega_{P'} \subset \omega_P$ and the inclusion is strict. However, this would imply that $\xi_{T'} > \xi_T$, where ξ_T is the midpoint of ω for $T = T(I,\omega)$, analogously $\xi_{T'}$. Then, T' should have been selected before T was selected. Contradiction. □

Proof of Lemma 3.12. By interpolation it suffices to show the statement for $q = \infty$ and at the weak endpoint $q = 2$. The case $q = \infty$ is trivial since we have that

$$\|F\|_{L^\infty(\boldsymbol{S}_A)} \lesssim \|F \mathbb{1}_A\|_{L^\infty(S)}$$

and the measure does not play a role.

For the endpoint $q = 2$ it is enough to show that

$$\lambda^2 \boldsymbol{\mu}(\|F\|_{\boldsymbol{S}_A} > \lambda) \lesssim \|F \mathbb{1}_A\|_{L^2(S)}$$

Let Ψ be the collection of trees and $\mathcal{P}$ the selected set of tiles during the selection algorithm. Using Lemma 3.19 we know that $\mathcal{P}$ is strongly disjoint. (3.9) implies

$$\lambda^2 \boldsymbol{\mu}(\|F\|_{\boldsymbol{S}_A} > \rho) \lesssim \sum_{T\in\Psi} \lambda^2 \mu(T) \lesssim \|F\mathbb{1}_A\|_{L^2(S)}.$$

The full statement follows now Proposition 2.9. □

3.3.2 Quasi monotonicity of outer L^p spaces on $\mathbb{X}$

In this subsection we show several inequalities concerning monotonicity and reverse quasi-monotonicity of the iterated L^p norms, which we will use in the proof of Proposition 3.25. Most of the proofs follow along the lines of Section 2.4, mostly with only minor changes, which we point out. First lemmas states that $Ł^p$ sizes are decreasing in p.

Lemma 3.21 (Monotonicity of iterated sizes)**.** *Let $0 < p \le q \le \infty$ and let $F \in \mathcal{B}(\mathbb{X})$. Then*

$$\|F\|_{Ł^q_\mu(S)} \lesssim_{p,q} \|F\|_{Ł^p_\mu(S)}.$$

Proof. Identical with the proof of Lemma 2.48. □

The following fact lets us relate the outer L^q norm restricted to $V \in \mathbb{D}^\cup$, with the averaged $Ł^q$.

Lemma 3.22. *Let $V \in \mathbb{D}^\cup$ and let $F \in \mathcal{B}(\mathbb{X})$. Then for $0 < q \le \infty$*

$$\|F\mathbb{1}_V\|_{L^q(S)} \lesssim_q \nu(V)^{1/q} \|F\mathbb{1}_V\|_{Ł^q(S)}.$$

Proof. By Remark 3.11 we may write $V = \bigcup_{m=1}^\infty D_m$, where $D_m \in \mathbb{D}$ for $m = 1, 2, ...$ and D_m's are pairwise disjoint. Then the argument follows along the lines the proof of Lemma 2.49. □

The next lemma reverts the inequality in Lemma 3.21 if F appropriately localized, losing a factor coming from the localization.

Lemma 3.23. *Let $A \subset \mathbb{X}$, $V \in \mathbb{D}^\cup$ and let $F \in \mathcal{B}(\mathbb{X})$. Then for any $0 < t \le q \le \infty$*

$$\|F\mathbb{1}_A\mathbb{1}_V\|_{L^t(S)} \lesssim_{q,t} \mu(V \cap A)^{1/t-1/q} \nu(V)^{1/q} \|F\mathbb{1}_A\mathbb{1}_V\|_{Ł^q(S)}.$$

Proof. Along the lines of the proof of Lemma 2.50. □

The following lemma controls the counting function of a forest coming from the selection algorithm in terms of $Ł^q$ norm. We will use it in the proof of Proposition 3.25 together with a careful decomposition of the trilinear form according to the level sets.

Lemma 3.24 (Counting function estimates)**.** *Let $V \in \mathbb{D}^\cup$ and let $F \in \mathcal{B}(\mathbb{X})$. Assume that $\Phi_\lambda \subset \mathbb{T}$ is the forest selected according to Definition 3.18 at a certain level $\lambda > 0$ for function $F\mathbb{1}_V$. Let $N_{\Phi_\lambda} = \sum_{T\in\Phi_\lambda} \mathbb{1}_{I_T}$ be its counting function. Then for any $1 \le p < \infty$ and $2 \le q < \infty$ the following bounds hold*

$$\|N_{\Phi_\lambda}\|_{L^p} \lesssim_p \nu(V)^{1/p} \lambda^{-q} \|F\mathbb{1}_V\|^q_{Ł^q(S)}$$

together with the BMO endpoint

$$\|N_{\Phi_\lambda}\|_{BMO} \lesssim \lambda^{-q} \|F\mathbb{1}_V\|^q_{Ł^q(S)}.$$

Proof. Identical with the proof of Lemma 2.51, after setting $E_\lambda := \mathcal{P}_\lambda$, $\mathcal{T}_\lambda := \Phi_\lambda$ in that argument. □

3.4 Iterated L^p bounds

From now on we fix $L \in \mathbb{Z}_+$ and set $\Lambda := \Lambda_L$. For notation, refer to Subsection 3.2.3. The main result of this section is the following proposition.

Proposition 3.25. *Let $1 < p_1, p_2, p_3 < \infty$ with $\sum_{j=1}^3 1/p_j = 1$ and $2 < q_1, q_2, q_3 < \infty$ with $\sum_{j=1}^3 1/q_j > 1$. Let for $j = 1, 2, 3$, f_j be a Schwartz function on $\mathbb{R}$ and let $F_j := F(f_j)$. Assume that $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$. Then*

$$|\Lambda(F_1 \mathbb{1}_{V_1 \setminus W_1}, F_2 \mathbb{1}_{V_2 \setminus W_2}, F_3 \mathbb{1}_{V_3 \setminus W_3})| \lesssim \prod_{j=1}^3 \nu(V_j)^{1/p_j} \|F_j \mathbb{1}_{V_j \setminus W_j}\|_{L^\infty \L^{q_j}(S)}. \tag{3.10}$$

Note that in conjunction with Proposition 2.10 the above inequality implies Theorem 3.1. Similarly as we noted in Chapter 2, we could not use the outer Hölder inequality from [DT15], which would require $\min_j \mu(V_j)$ instead of $\prod_j \mu(V_j)^{1/p_j}$ on the right hand side of (3.10) and moreover, our trilinear form is nonpositive. Just like in Chapter 2, it does not seem feasible that one can obtain much better gain than $\prod_j \mu(V_j)^{1/p_j}$ in (3.10), since V_1 scales differently than V_2 and V_3.

Before we prove Proposition 3.25 we show the localized estimate at the level of trees.

Proposition 3.26. *Let $1 \le p_1, p_2, p_3 \le \infty$ with $\sum_{j=1}^3 1/p_j = 1$. Let for $j = 1, 2, 3$, f_j be a Schwartz function on $\mathbb{R}$ and let $F_j := F(f_j)$. Assume that $K_1, M_1 \in \mathbb{T}^\cup$, $K_j, M_j \in \mathbb{T}_L^\cup$ for $j = 2, 3$ and $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$. Moreover, set $A = \bigcap_{j=2}^3 (V_j \setminus W_j)$, $G_j := F_j \mathbb{1}_{V_j \setminus W_j}$, $S_1 := S$, $\mu_1 := \mu$ and $S_j := \boldsymbol{S}_A$, $\mu_j := \boldsymbol{\mu}$ for $j = 2, 3$. Then*

$$|\Lambda(G_1 \mathbb{1}_{K_1 \setminus M_1}, G_2 \mathbb{1}_{K_2 \setminus M_2}, G_3 \mathbb{1}_{K_3 \setminus M_3})| \lesssim \prod_{j=1}^3 \mu_j(K_j)^{1/p_j} \|G_j \mathbb{1}_{K_j \setminus M_j}\|_{L^\infty(S_j)}.$$

Remark. *Observe that optimizing in p_j and $\mu_j(K_j)$, we can make $\prod_{j=1}^3 \mu_j(K_j)^{1/p_j}$ to be equal $\min_{j=1,2,3} \mu_j(K_j)$.*

Note that applying Proposition 2.10, the previous Proposition immediately implies $|\Lambda(G_1, G_2, G_3)| \lesssim \prod_{j=1}^3 \|G_j\|_{L^{p_j}(S_j)}$. However, we shall need an improvement of this inequality with additional localization.

Corollary 3.27. *Let $1 < p_1, p_2, p_3 < \infty$ with $1/p_1 + 1/p_2 + 1/p_3 = 1$. Let for $j = 1, 2, 3$, f_j be a Schwartz function on $\mathbb{R}$ and let $F_j := F(f_j)$. Assume that $K_1, M_1 \in \mathbb{T}^\cup$, $K_j, M_j \in \mathbb{T}_L^\cup$ for $j = 2, 3$ and $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$. Moreover, set $A = \bigcap_{j=2}^3 (V_j \setminus W_j)$, $H_j := F_j \mathbb{1}_{V_j \setminus W_j} \mathbb{1}_{K_j \setminus M_j}$, $S_1 := S$, $\mu_1 := \mu$ and $S_j := \boldsymbol{S}_A$, $\mu_j := \mu$ for $j = 2, 3$. Then*

$$|\Lambda(H_1, H_2, H_3)| \lesssim \prod_{j=1}^3 \|H_j\|_{L^{p_j}(S_j)}.$$

Proof. Let $\widetilde{K}_j, \widetilde{M}_j \in \mathbb{T}^\cup$ for $j = 1, 2, 3$. Note that the intersection of two trees is a union of trees. Thus the intersection of two unions of trees is a union of trees and we may apply Proposition 3.26 with $\overline{K}_j = K_j \cap \widetilde{K}_j$, $\overline{M}_j = M_j \cup \widetilde{M}_j$ and with $G_j = F_j \mathbb{1}_{V_j \setminus W_j}$ for $j = 1, 2, 3$. We obtain

$$\begin{aligned} &|\Lambda(H_1 \mathbb{1}_{\widetilde{K}_1 \setminus \widetilde{M}_1}, H_2 \mathbb{1}_{\widetilde{K}_2 \setminus \widetilde{M}_2}, H_3 \mathbb{1}_{\widetilde{K}_3 \setminus \widetilde{M}_3})| \\ &= |\Lambda(G_1 \mathbb{1}_{\overline{K}_1 \setminus \overline{M}_1}, G_2 \mathbb{1}_{\overline{K}_2 \setminus \overline{M}_2}, G_3 \mathbb{1}_{\overline{K}_3 \setminus \overline{M}_3})| \end{aligned}$$

$$\lesssim \prod_{j=1}^{3} \mu_j(\overline{K}_j)^{1/p_j} \|G_j \mathbb{1}_{\overline{K}_j \setminus \overline{M}_j}\|_{L^\infty(S_j)} \lesssim \prod_{j=1}^{3} \mu_j(\widetilde{K}_j)^{1/p_j} \|H_j \mathbb{1}_{\widetilde{K}_j \setminus \widetilde{M}_j}\|_{L^\infty(S_j)}$$

We conclude the proof applying Proposition 2.10. □

In the first subsection we prove Proposition 3.26 and in the second subsection we show Proposition 3.25.

3.4.1 Proof of Proposition 3.26

In this subsection we prove Proposition 3.26. Before we start the actual argument we discuss the geometry of the trilinear form and make some convenient reductions.

We set a bit of notation for this subsection. Let $A_j = V_j \setminus W_j$, where $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$ be fixed throughout this subsection. Define for $T = T(I, \omega) \in \mathbb{T}$, $\boldsymbol{T} := T(I, 2^L\omega) \in \mathbb{T}_L$ and for $K = \bigcup_k T_k \in \mathbb{T}^\cup$, define $\boldsymbol{K} := \bigcup_k \boldsymbol{T}_k$. By the definition of the trilinear form, we have the following fact.

Lemma 3.28 (Transfer property)**.** *Let $F_1, F_2, F_3 \in \mathcal{B}(\mathbb{X})$ and let $K, M \in \mathbb{T}^\cup$, $V, W \in \mathbb{D}^\cup$. Then*

$$\Lambda(F_1 \mathbb{1}_{K \setminus M}, F_2, F_3) = \Lambda(F_1, F_2 \mathbb{1}_{\boldsymbol{K} \setminus \boldsymbol{M}}, F_3) = \Lambda(F_1, F_2, F_3 \mathbb{1}_{\boldsymbol{K} \setminus \boldsymbol{M}}), \tag{3.11}$$

$$\Lambda(F_1, F_2 \mathbb{1}_{V \setminus W}, F_3) = \Lambda(F_1, F_2, F_3 \mathbb{1}_{V \setminus W}).$$

Using the previous lemma, we shall make an appropriate reduction, in order to streamline the exposition of the proof. Proposition 3.26 follows from the next lemma.

Lemma 3.29. *Let $T \in \mathbb{T}$ and $K \in \mathbb{T}^\cup$ and let $A = V \setminus W$ where $V, W \in \mathbb{D}^\cup$. Let for $j = 1, 2, 3$, f_j be a Schwartz function on $\mathbb{R}$ and let $F_j := F(f_j)$. Then*

$$|\Lambda(F_1 \mathbb{1}_{T \setminus K}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A)| \lesssim \mu(T) \|F_1 \mathbb{1}_{T \setminus K}\|_{L^\infty(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_A \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{K}}\|_{L^\infty(\boldsymbol{S}_A)}.$$

Proof of Proposition 3.26 assuming Lemma 3.29. We prove that the above inequality suffices in order to show Proposition 3.26 in two steps.

Step 1: in this step we prove that the desired inequality from Proposition 3.26 follows, if we assume that for any $T \in \mathbb{T}$, $K \in \mathbb{T}^\cup$, $A = V \setminus W$ with $V, W \in \mathbb{D}^\cup$ it holds that

$$\Lambda(F_1 \mathbb{1}_{T \setminus K} \mathbb{1}_{A_2}, F_2 \mathbb{1}_{A_2}, F_3 \mathbb{1}_{A_3}) \lesssim \mu(T) \|F_1 \mathbb{1}_{A_1} \mathbb{1}_{T \setminus K}\|_{L^\infty(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_{A_j} \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{K}}\|_{L^\infty(\boldsymbol{S}_A)}$$

Let $K = K_1 \cap K_2 \cap K_3$ and $M = M_1 \cup M_2 \cup M_3$, and let $K = \bigcup_{i=1}^n T_i$ be such that $\sum_{i=1}^n \mu(T_i) \lesssim \mu(K)$. We can make T_i's pairwise disjoint obtaining convex trees $\bigcup_{i=1}^n T_i \setminus N_i = K$. Using Lemma 3.28 in the first inequality and the assumed inequality

$$\begin{aligned}
&\Lambda(F_1 \mathbb{1}_{A_1} \mathbb{1}_{K_1 \setminus M_1}, F_2 \mathbb{1}_{A_2} \mathbb{1}_{\boldsymbol{K}_2 \setminus \boldsymbol{M}_2}, F_3 \mathbb{1}_{A_3} \mathbb{1}_{\boldsymbol{K}_3 \setminus \boldsymbol{M}_3}) \\
&= \sum_{i=1}^{n} \Lambda(F_1 \mathbb{1}_{T_i \setminus N_i \setminus M} \mathbb{1}_{A_1}, F_2 \mathbb{1}_{A_2}, F_3 \mathbb{1}_{A_3}) \\
&\lesssim \sum_{i=1}^{n} \mu(T_i) \|F_1 \mathbb{1}_{A_1} \mathbb{1}_{T_i \setminus N_i \setminus M}\|_{L^\infty(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_{A_j} \mathbb{1}_{\boldsymbol{T}_i \setminus \boldsymbol{N}_i \setminus \boldsymbol{M}}\|_{L^\infty(\boldsymbol{S}_A)}
\end{aligned}$$

$$\lesssim \mu(K)\|F_1 \mathbb{1}_{A_1} \mathbb{1}_{K_1\setminus M_1}\|_{L^\infty(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_{A_j} \mathbb{1}_{\boldsymbol{K}_j\setminus \boldsymbol{M}_j}\|_{L^\infty(\boldsymbol{S}_A)}.$$

Step 2: in this step we prove that the inequality assumed in the previous step follows from the inequality in the statement of Lemma 2.74. Note that using Remark 3.11, for all $T \in \mathbb{T}$, $K \in \mathbb{T}^\cup$ and $V, W \in \mathbb{D}^\cup$ we have

$$(T \setminus K) \cap (V \setminus W) = \bigcup_{k=1}^{n} (T_k \setminus \widetilde{K}),$$

where $T_k \in \mathbb{T}$ are pairwise disjoint in space and $\sum_{k=1}^{n} \mu(T_k) \le \mu(T)$, and $\widetilde{K} = K \cup W \in \mathbb{T}^\cup$. Moreover inside Λ we can freely transfer the characteristic functions between F_2 and F_3, using Lemma 3.28. Let $A = A_2 \cap A_3$, note that $A \in \mathbb{D}^\cup$. Using the assumed inequality we have

$$\begin{aligned}
&\Lambda(F_1 \mathbb{1}_{T\setminus K} \mathbb{1}_{A_1}, F_2 \mathbb{1}_{A_2}, F_3 \mathbb{1}_{A_3}) \\
&= \sum_{k=1}^{n} \Lambda(F_1 \mathbb{1}_{T_k\setminus \widetilde{K}}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A) \\
&\lesssim \sum_{k=1}^{n} \mu(T_k)\|F_1 \mathbb{1}_{T_k\setminus \widetilde{K}}\|_{L^\infty(S)} \|F_2 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T}_k\setminus \widetilde{\boldsymbol{K}}}\|_{L^\infty(\boldsymbol{S}_A)} \|F_3 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T}_k\setminus \widetilde{\boldsymbol{K}}}\|_{L^\infty(\boldsymbol{S}_A)} \\
&\lesssim \mu(T)\|F_1 \mathbb{1}_{A_1} \mathbb{1}_{T\setminus K}\|_{L^\infty(S)} \|F_2 \mathbb{1}_{A_2} \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}}\|_{L^\infty(\boldsymbol{S}_A)} \|F_3 \mathbb{1}_{A_3} \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}}\|_{L^\infty(\boldsymbol{S}_A)}.
\end{aligned}$$

□

Finally we give a proof of Lemma 3.29. We bound the trilinear form over the lacunary and the overlapping tree. The former is significantly easier and follows from a single application of classical Hölder's inequality. The latter is more involved and requires discrete integration parts combined with a geometric argument and Hölder's inequality. We record that such partial integration is often called telescoping in the context of paraproducts. It is the Walsh counterpart of the argument from Chapter 2, where we used integration by parts and Green's theorem. Some parts of the argument are motivated by [OT11], however here the setting is somewhat different, since we are restricted to sets of the type $V \in \mathbb{D}^\cup$ and have to take care of the boundary terms.

Proof of Lemma 3.29. Let $A = V \setminus W$, where $V, W \in \mathbb{D}^\cup$. Note that

$$\begin{aligned}
&|\Lambda(F_1 \mathbb{1}_{T^{(lac)}\setminus K}, F_2 \mathbb{1}_{V\setminus W}, F_3 \mathbb{1}_{V\setminus W})| \\
&\lesssim \mu(T)\|F_1 \mathbb{1}_{T\setminus K}\|_{S(T)} \|F_2 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}}\|_{\boldsymbol{S}^2(\boldsymbol{T})} \|F_3 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}})\|_{\boldsymbol{S}^2(\boldsymbol{T})},
\end{aligned}$$

just by an application of $(\infty, 2, 2)$-Hölder's inequality, since both F_2 and F_3 are restricted to the lacunary tree. We still have to show

$$\begin{aligned}
&|\Lambda(F_1 \mathbb{1}_{T^{(ov)}\setminus K}, F_2 \mathbb{1}_{V\setminus W}, F_3 \mathbb{1}_{V\setminus W})| \\
&\lesssim \mu(T)\|F_1 \mathbb{1}_{T\setminus K}\|_{L^\infty(S)} \|F_2 \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}} \mathbb{1}_A\|_{L^\infty(\boldsymbol{S}_A)} \|F_3 \mathbb{1}_{\boldsymbol{T}\setminus \boldsymbol{K}} \mathbb{1}_A\|_{L^\infty(\boldsymbol{S}_A)}.
\end{aligned}$$

Note that we have

$$\Lambda(F_1 \mathbb{1}_{T^{(ov)}\setminus K}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A) = \Lambda(F_1 \mathbb{1}_{T^{(ov)}} \mathbb{1}_B, F_2 \mathbb{1}_C, F_3 \mathbb{1}_C),$$

where

$$B = T \setminus K, \qquad C = A \cap (\boldsymbol{T} \setminus \boldsymbol{K}).$$

The right hand side can be rewritten as

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(Q)h_{I_p}(x) \sum_{Q\in P^L} F_2(Q)F_3(Q)\mathbb{1}_C(Q)\frac{\mathbb{1}_{I_Q}(x)}{|I_Q|}\,dx, \tag{3.12}$$

where h_{I_p} is the L^2 normalized Haar function. Fix $Q \in \boldsymbol{T}$. Let $\mathcal{P}_1$ be the set of maximal tiles in $C\cap \boldsymbol{T}^{(ov)}$, $\mathcal{P}_2$ be the set of minimal tiles in $C\cap\boldsymbol{T}^{(ov)}$; note that $\mathrm{Top}_C(\boldsymbol{T}) = \mathcal{P}_1\cup\mathcal{P}_2$. Define

$$\widetilde{\mathcal{Q}} = C\cap\{\widetilde{Q}\in\boldsymbol{T}^{(lac)}\colon \widetilde{Q}^{\circlearrowleft}\le Q,\ \widetilde{Q}^{\circlearrowleft}\notin\mathcal{P}_2\}.$$

and $\mathcal{Q} = (C\cap\{Q\})\cup\widetilde{\mathcal{Q}}$. Note that $(\mathcal{P}_1\cap\{\widetilde{Q}<Q\})\cup\mathcal{Q}$ and $\mathcal{P}_2\cap\{\widetilde{Q}\le Q\}$ are two decompositions of the same subset of $\mathbb{R}^2_+$ into pairwise disjoint tiles. Using Corollary 3.34 that gives

$$F(Q)\mathbb{1}_C(Q)\varphi_Q = -\sum_{\widetilde{Q}\in\widetilde{\mathcal{Q}}} F(\widetilde{Q})\mathbb{1}_C(\widetilde{Q})\varphi_{\widetilde{Q}} - \sum_{\substack{\widetilde{Q}\in\mathcal{P}_1\\ \widetilde{Q}<Q}} F(\widetilde{Q})\varphi_{\widetilde{Q}} + \sum_{\substack{\widetilde{Q}\in\mathcal{P}_2\\ \widetilde{Q}\le Q}} F(\widetilde{Q})\varphi_{\widetilde{Q}} \tag{3.13}$$

Observe that the above identity can be seen as a discrete integration by parts. Note that it follows from Lemma 3.32, that the following cancellation identities hold for any two different $\widetilde{Q}_1,\widetilde{Q}_2\in\widetilde{\mathcal{Q}}$ and any $\widetilde{Q}\in\widetilde{\mathcal{Q}}$, $\overline{Q}\in\mathrm{Top}_C(\boldsymbol{T})\cap\{\widetilde{Q}\le Q\}$

$$\int h_{I_P}(x)\varphi_{\widetilde{Q}_1}(x)\varphi_{\widetilde{Q}_2}(x)\,dx = 0, \qquad \int h_{I_P}(x)\varphi_{\widetilde{Q}}(x)\varphi_{\overline{Q}}(x)\,dx = 0. \tag{3.14}$$

From now on let for a tile P, $k_P\in\mathbb{Z}$ be such that $2^{k_P} = |I_P|$. Using $|I_Q|^{-1}\mathbb{1}_{I_Q} = \varphi_Q^2$, applying (3.13) to F_2 and F_3 in (3.12), and using (3.14), we are left with estimating

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x) \sum_{\substack{Q\in\boldsymbol{T}^{(lac)}\\ k_Q\le k_P-L}} F_2(Q)F_3(Q)\mathbb{1}_C(Q)\frac{\mathbb{1}_{I_Q}(x)}{|I_Q|}\,dx, \tag{3.15}$$

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x) \sum_{\substack{Q\in\mathrm{Top}_C(\boldsymbol{T})\\ k_Q\le k_P-L}} F_2(Q)F_3(Q)\frac{\mathbb{1}_{I_Q}(x)}{|I_Q|}\,dx, \tag{3.16}$$

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)\prod_{j=2}^{3} \sum_{\substack{Q\in\mathcal{P}_{j-1}\\ k_Q\le k_P-L}} F_j(Q)\varphi_Q(x)\,dx. \tag{3.17}$$

and an integral symmetric to the last one. We first bound (3.15). Changing the order of summation (3.15) equals

$$\int \sum_{Q\in\boldsymbol{T}^{(lac)}} \Big(\sum_{\substack{P\in T^{(ov)}\\ k_Q+L\le k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)\Big)F_2(Q)F_3(Q)\mathbb{1}_C(Q)\frac{\mathbb{1}_{I_Q}(x)}{|I_Q|}\,dx$$

Applying $(\infty,2,2)$-Hölder's inequality in Q and x this is bounded by

$$\mu(T)\Big\|\sup_l\Big|\sum_{\substack{P\in T^{(ov)}\\ l\le k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)\Big|\Big\|_{L^\infty}\prod_{j=2}^{3}\|F_j\mathbb{1}_C\|_{\boldsymbol{S}^2(\boldsymbol{T})}$$

The above display is bounded by the desired quantity by an application of Lemma 3.31. We shall now estimate (3.16). Changing the order of summation we rewrite (3.16) as

$$\int \sum_{Q\in \mathrm{Top}_C(\boldsymbol{T})} \Big(\sum_{\substack{P\in T^{(ov)}\\ k_Q+L\leq k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)\Big)F_2(Q)F_3(Q)\frac{\mathbb{1}_{I_Q}(x)}{|I_Q|}\,dx.$$

Applying $(\infty,2,2)$-Hölder's inequality in Q and x, we estimate the above display by

$$\mu(T)\|\sup_l|\sum_{\substack{P\in T^{(ov)}\\ l\leq k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)|\|_{L^\infty}\prod_{j=2}^{3}\|F_j\|_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})}$$

The second factor is bounded similarly as before by Lemma 3.31. Concerning the last two, we can decompose $\mathrm{Top}_C(\boldsymbol{T})$ as follows.

Lemma 3.30. *Let $C=A\cap(\boldsymbol{T}\setminus\boldsymbol{K})$ be as above. There exists $K'\in\mathbb{T}^{\cup}$ with*

$$\mathrm{Top}_C(\boldsymbol{T})\subset\Big(\bigcup_{\boldsymbol{T}'\in\boldsymbol{K}'}\mathrm{Top}_A(\boldsymbol{T}')\cup\mathrm{Top}_A(\boldsymbol{T})\Big)\cap C, \tag{3.18}$$

$$\sum_{\boldsymbol{T}'\in\boldsymbol{K}'}\boldsymbol{\mu}(\boldsymbol{T}')\lesssim\boldsymbol{\mu}(\boldsymbol{T}). \tag{3.19}$$

Proof. Let $T=T(I,\omega)$, $\boldsymbol{T}=T^L(I,2^L\omega)$ and $K=\bigcup\overline{T}$, $\boldsymbol{K}=\bigcup\overline{\boldsymbol{T}}$. Let $\mathcal{P}$ be the set of maximal tiles $P\in T^{(ov)}$ such that $P\in K$. Define the set of trees having the tiles of $\mathcal{P}$ as tops $\widetilde{K}=\{T(I_P,\omega_P)\colon P\in\mathcal{P}\}$. Note that $T\cap K=T\cap\widetilde{K}$ and $\sum_{\widetilde{T}\in\widetilde{K}}\mu(\widetilde{T})\leq\mu(T)$. That implies $\boldsymbol{T}\cap\boldsymbol{K}=\boldsymbol{T}\cap\widetilde{\boldsymbol{K}}$ and $\sum_{\widetilde{\boldsymbol{T}}\in\widetilde{\boldsymbol{K}}}\boldsymbol{\mu}(\widetilde{\boldsymbol{T}})\leq\boldsymbol{\mu}(\boldsymbol{T})$. Moreover, let $\mathcal{P}'$ be the set of dyadic parents of the tiles $\mathcal{P}$ which belong to T and let $K'=\{T(I_P,\omega_P)\colon P\in\mathcal{P}'\}$. Note that (3.19) is satisfied for such choice of $\boldsymbol{K}'$. We shall now validate the condition (3.18).

Let $P\in\mathrm{Top}_C(\boldsymbol{T})$, i.e. P is either a maximal or minimal tile in $\boldsymbol{T}^{(ov)}\cap A\setminus\widetilde{\boldsymbol{K}}$. First, suppose it is a maximal tile. If P does not have a dyadic parent in $\boldsymbol{T}$, then automatically $P\in\mathrm{Top}_A(\boldsymbol{T})$. Otherwise, let $\widetilde{P}$ be the dyadic parent of P in $\boldsymbol{T}^{(ov)}$. Then $\widetilde{P}\in\boldsymbol{T}\setminus A$ or $\boldsymbol{T}\cap\widetilde{\boldsymbol{K}}$. In the first case, by definition, we have $P\in\mathrm{Top}_A(\boldsymbol{T})$. Note that the second case is in fact not possible, since $\widetilde{P}\in\widetilde{\boldsymbol{K}}$ implies $P\in\widetilde{\boldsymbol{K}}$. Now, suppose that $P\in\mathrm{Top}_C(\boldsymbol{T})$ is a minimal tile in $\boldsymbol{T}^{(ov)}\cap A\setminus\boldsymbol{K}$. Let $\widetilde{P}$ be one of the dyadic children of P. Again, we have $\widetilde{P}\in\boldsymbol{T}\setminus A$ or $\widetilde{P}\in\boldsymbol{T}\cap\widetilde{\boldsymbol{K}}$. In the first case, $P\in\mathrm{Top}_A(\boldsymbol{T})$. In the second case, there exists $\widetilde{T}\in\widetilde{K}$ such that $\widetilde{P}\in\widetilde{\boldsymbol{T}}$. Let T' be the tree whose top tile belongs to $T^{(ov)}$ and is a dyadic parent of the top tile of $\widetilde{T}$. Then, we have $P\in\mathrm{Top}_A(\boldsymbol{T}')$. □

Using Lemma 3.30 applied for $j=2,3$

$$\begin{aligned}&\|F_j\|_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})}\\ &\lesssim\|F_j\mathbb{1}_C\|_{\boldsymbol{S}_A^{\mathrm{Top}}(\boldsymbol{T})}+\mu(T)^{-1/2}\Big(\sum_{\boldsymbol{T}'\in\boldsymbol{K}'}\mu(T')\|F_j\mathbb{1}_C\|^2_{\boldsymbol{S}_A^{\mathrm{Top}}(\boldsymbol{T}')}\Big)^{1/2}\\ &\lesssim\|F_j\mathbb{1}_C\|_{L^\infty(\boldsymbol{S}_A)}.\end{aligned}$$

We are left with estimating (3.17). Note that we can rewrite

$$\prod_{j=2}^{3}\sum_{\substack{Q\in\mathcal{P}_{j-1}\\ k_Q\leq k_P-L}}F_j(Q)\varphi_Q(x)=\sum_{\substack{Q_1\in\mathcal{P}_1\\ k_{Q_1}\leq k_P-L}}\sum_{\substack{Q_2\in\mathcal{P}_2\\ k_{Q_2}\leq k_P-L}}F_2(Q_1)\varphi_{Q_1}(x)F_3(Q_2)\varphi_{Q_2}(x),$$

Changing the order of summation, (3.17) becomes

$$\int \sum_{Q_1\in\mathcal{P}_1}\sum_{Q_2\in\mathcal{P}_2}\sum_{\substack{P\in T^{(ov)}\\ k_{Q_1}+L\le k_P\\ k_{Q_2}+L\le k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x)F_2(Q_1)\varphi_{Q_1}(x)F_3(Q_2)\varphi_{Q_2}(x)\,dx.$$

The above is bounded by

$$\int \sup_l |\sum_{\substack{P\in T^(ov)\\ l\le k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)|\prod_{j=2}^{3}\sum_{Q\in\mathcal{P}_{j-1}}|F_j(Q)||\varphi_Q(x)|\,dx$$

Applying $(\infty,2,2)$-Hölder's inequality in x, using Lemma 3.31, spatial disjointness of tiles $\mathcal{P}_1$ and spatial disjointness of tiles in $\mathcal{P}_2$ we bound the last display by

$$\mu(T)\|F_1\mathbb{1}_B\|_{L^\infty(S)}\|F_2\|_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})}\|F_3\|_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})}.$$

Another application of Lemma 3.30 finishes the proof of the proposition.

□

At the end of this subsection we prove that the Walsh counterpart of the maximal truncation of the Hilbert transform is bounded in terms of the size S. This estimate was already proven and used in [OT11]. We include the proof for convenience of the reader.

Lemma 3.31. *Let $T\in\mathbb{T}$ and $F:=F(f)$, where f is a Schwartz function on $\mathbb{R}$. Then*

$$\|\sup_l |\sum_{\substack{P\in T^{(ov)}\\ l\le k_P}} F(P^{\circlearrowleft})\mathbb{1}_B(P)h_{I_P}(x)|\|_{L^\infty(\mathbb{R})}\lesssim \|F\mathbb{1}_B\|_{L^\infty(S)}.$$

Proof. First of all, observe that the supremum over l can be dominated by the maximal function, hence it can be discarded. Hence, it is enough to bound

$$\|\sum_{P\in T^{(lac)}} F(P)\mathbb{1}_B(P)h_{I_P}(x)\|_{L^\infty(\mathbb{R})}.$$

Since $B\cap T$ is a convex set, the above display is bounded by $\|F\mathbb{1}_B\|_{L^\infty(S)}$ using Corollary 3.35. □

3.4.2 Proof of Proposition 3.25

In this subsection we prove Proposition 3.25. The main difficulty is to prove Lemma 3.25 in the case when $(1/p_1,1/p_2,1/p_3)$ is in the neighbourhood of $(0,0,1)$ or $(0,1,0)$. We remark that the proof can be simplified if $(1/p_1,1/p_2,1/p_3)$ is in the neighbourhood of $(1,0,0)$, however here we present the version of the argument that works for all cases. Similarly as in [OT11] we split the trilinear form according to the level sets of F_1. We then apply Hölder inequality on the level of trees and use the inequalities from Section 3.3 to bound the resulting outer L^p norms in terms of the iterated norms. Note that that this argument is very similar to the one at the end of Chapter 2, however here we do not have to deal with with extra factors with a small exponent γ, see Chapter 2 for details.

Proof of Lemma 3.25. Since we can freely move characteristic functions between F_2 and F_3 we may assume that $V_2 = V_3$. Moreover, since $V_1 \in \mathbb{T}^{\cup}$, after an application of (3.11) we may assume that $V_2 \subset V_1$. By homogeneity we may assume that for $j \in \{1,2,3\}$

$$\nu(V_j)^{1/p_j} \|F_j \mathbb{1}_{V_j \setminus W_j}\|_{\mathcal{L}^{q_j}(S)} \le 1. \tag{3.20}$$

Hence it suffices to show

$$|\Lambda(F_1 \mathbb{1}_{V_1\setminus W_1}, F_2 \mathbb{1}_{V_2\setminus W_2}, F_3 \mathbb{1}_{V_3\setminus W_3})| \lesssim 1.$$

In order to shorten displays, throughout the proof we set $A_j := V_j \setminus W_j$ for $j = 1,2,3$.

Note that (3.20) and Lemma 3.21 imply that $\|F_1 \mathbb{1}_{A_1}\|_{L^\infty(S)} \le \nu(V_1)^{-1/p_1}$. Let us run the selection algorithm Definition 3.18 for $k \in \mathbb{N}$, such that $E_k := \bigcup \Phi_k$ corresponds to the level $2^{-k}\nu(V_1)^{-1/p_1}$. Additionally defining $E_{-1} = \emptyset$ we have

$$\|F_1 \mathbb{1}_{A_1} \mathbb{1}_{E_k\setminus E_{k-1}}\|_{L^\infty(S)} \lesssim 2^{-k}\nu(V_1)^{-1/p_1}.$$

and $E_k \setminus E_{k-1}$ are pairwise disjoint. Using (3.20) and Lemma 3.24 we obtain

$$\mu(E_k) \lesssim 2^{q_1 k}\nu(V_1)\nu(V_1)^{q_1/p_1}\|F_1\|^{q_1}_{\mathcal{L}^{q_1}(S)} \lesssim 2^{q_1 k}\nu(V_1). \tag{3.21}$$

Set $N_{\Phi_k} = \sum_{\Delta T \in \Phi_k} \mathbb{1}_{I_T}$ to be the counting function of the forest Φ_k. Note that E_k for $k \in \mathbb{Z}$ are pairwise disjoint so we may split the the trilinear form using Lemma 3.28 into

$$\begin{aligned}&|\Lambda(\mathbb{1}_{A_1}F_1, \mathbb{1}_{A_2}F_2, \mathbb{1}_{A_3}F_3)|\\ &\le \sum_{k\ge 0} |\Lambda(F_1 \mathbb{1}_{E_k\setminus E_{k-1}}\mathbb{1}_{A_1}, F_2 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}, F_3 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_3})|\end{aligned}$$

Fix $k \in \mathbb{Z}_+$. Applying Lemma 3.29 together with Proposition 2.10 we have

$$\begin{aligned}&|\Lambda(F_1 \mathbb{1}_{E_k\setminus E_{k-1}}\mathbb{1}_{A_1}, F_2 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}, F_3 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_3})|\\ &\lesssim \|F_1 \mathbb{1}_{E_k\setminus E_{k-1}}\mathbb{1}_{A_1}\|_{L^{t_1}(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_j}\|_{L^{t_j}(\boldsymbol{S}_{A_j})}\end{aligned} \tag{3.22}$$

for any $t_i \in [1,\infty]$ such that $\sum_{i=1}^{3} t_i^{-1} = 1$. Using (3.21) and Lemma 3.23, it follows that

$$\|F_1 \mathbb{1}_{E_k}\mathbb{1}_{A_1}\|_{L^{t_1}(S)} \lesssim 2^{q_1 k(1/t_1 - 1/q_1)}\nu(V_1)^{1/t_1 - 1/p_1}. \tag{3.23}$$

In the case of terms involving F_2, as long as $t_2 \in (2, q_2]$ we use Lemma 3.12, Lemma 3.23 and Lemma 3.21 to obtain that

$$\begin{aligned}&\|F_2 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}\|_{L^{t_2}(\boldsymbol{S}_{A_2})}\\ &\lesssim \|F_2 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}\|_{L^{t_2}(S)}\\ &\lesssim \mu(V_2 \cap \boldsymbol{E}_k)^{1/t_2 - 1/q_2}\nu(V_2)^{1/q_2}\|F_2 \mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}\|_{\mathcal{L}^{q_2}(S)}\\ &\lesssim \mu(V_2 \cap \boldsymbol{E}_k)^{1/t_2 - 1/q_2}\nu(V_2)^{1/q_2}\|F_2 \mathbb{1}_{A_2}\|_{\mathcal{L}^{q_2}(S)}\\ &\lesssim \mu(V_2 \cap \boldsymbol{E}_k)^{1/t_2 - 1/q_2}\nu(V_2)^{1/q_2 - 1/p_2}.\end{aligned}$$

Using Remark 3.11, let $V_2 = \bigcup_{m=1}^{\infty} D_m$, where D_m's are pairwise disjoint strips. Then

$$\mu(V_2 \cap \boldsymbol{E}_k) \lesssim \sum_{\boldsymbol{T}\in\boldsymbol{\Phi}_k} \mu(V_2 \cap \boldsymbol{T}) \le \sum_{\boldsymbol{T}\in\boldsymbol{\Phi}_k}\sum_{m=1}^{\infty} \mu(D_m \cap \boldsymbol{T})$$

$$\lesssim \sum_{\boldsymbol{T}\in\boldsymbol{\Phi}_k}\sum_{m=1}^{\infty}|I_{D_m}\cap I_{\boldsymbol{T}}| \lesssim \sum_{\boldsymbol{T}\in\boldsymbol{\Phi}_k}|I_{V_2}\cap I_{\boldsymbol{T}}| = \|N_{\boldsymbol{\Phi}_k}\mathbb{1}_{I_{V_2}}\|_{L^1}$$

Notice that $N_{\Phi_k} = N_{\boldsymbol{\Phi}_k}$ and by Hölder's inequality with $1 \leq p \leq \infty$ the previous display is bounded by

$$\nu(V_2)^{1-1/p}\|N_{\Phi_k}\|_p = \nu(V_2)\frac{\|N_{\Phi_k}\|_p}{\nu(V_2)^{1/p}}.$$

By Lemma 3.24 and (3.20) we have for $1 \leq p < \infty$

$$\|N_{\Phi_k}\|_p \lesssim_p \nu(V_1)^{1/p}2^{q_1 k}\nu(V_1)^{q_1/p_1}\|F_1\mathbb{1}_{A_1}\|^{q_1}_{\mathcal{L}^{q_1}(S)} \leq 2^{q_1 k}\nu(V_1)^{1/p}.$$

Thus, for any $p \in [1,\infty)$, we have

$$\begin{aligned}&\|F_2\mathbb{1}_{\boldsymbol{E}_k}\mathbb{1}_{A_2}\|_{L^{t_2}(\boldsymbol{S}_{A_2})} \\ &\lesssim 2^{q_1 k(1/t_2-1/q_2)}\left(\frac{\nu(V_2)}{\nu(V_1)}\right)^{-\left(1/t_2-1/q_2\right)/p}\nu(V_2)^{1/t_2-1/p_2}.\end{aligned} \tag{3.24}$$

The same result holds for the term with F_3. Let $V_{2,3} = V_2 = V_3$. Putting the bounds (3.24) and (3.23) into (3.22), we obtain using $\sum_{j=1}^{3} 1/t_j = \sum_{j=1}^{3} 1/p_j = 1$

$$\begin{aligned}&|\Lambda(F_1\mathbb{1}_{A_1}\mathbb{1}_{E_k}, F_2\mathbb{1}_{A_2}\mathbb{1}_{\boldsymbol{E}_k}, F_3\mathbb{1}_{A_3}\mathbb{1}_{\boldsymbol{E}_k})| \\ &\lesssim 2^{q_1 k(1-1/q_1-1/q_2-1/q_3)} \\ &\times\left(\frac{\nu(V_{2,3})}{\nu(V_1)}\right)^{(1/p_1-1/t_1)}\left(\frac{\nu(V_{2,3})}{\nu(V_1)}\right)^{-(1/t_2+1/t_3-(1/q_2+1/q_3))/p}.\end{aligned}$$

By assumption we have that $1 - 1/q_1 - 1/q_2 - 1/q_3 < 0$, which makes the above expression summable in $k \in \mathbb{N}$. If $t_1 > p_1$ and $p \in (1,\infty]$ is chosen large enough then the exponent of $\frac{\nu(V_{2,3})}{\nu(V_1)}$ is positive. Since $\nu(V_{2,3}) \leq \nu(V_1)$, this concludes the proof. □

3.5 Appendix - Walsh wave packets

In this chapter we used the following facts about the Walsh wave packets.

Lemma 3.32. *If two tiles P, Q are disjoint, then φ_P and φ_Q are orthogonal, i.e.* $\langle\varphi_P,\varphi_Q\rangle = 0$.

The above lemma was proven in [Thi06].

Lemma 3.33. *Let $\mathcal{P}$ be a finite collection of tiles and assume that a tile Q is covered by the tiles in $\mathcal{P}$. Then φ_Q is in the linear span of $\{\varphi_P\colon P\in\mathcal{P}\}$.*

The above lemma was also proven in [Thi06]. As a corollary we obtain the following.

Corollary 3.34. *If $\mathcal{P}$, $\mathcal{P}'$ are two different collection of multitiles, each of which is pairwise disjoint and $\bigcup\mathcal{P} = \bigcup\mathcal{P}'$, then for any Schwartz function f on $\mathbb{R}$*

$$\sum_{P\in\mathcal{P}}F(f)(P)\varphi_P(x) = \sum_{P'\in\mathcal{P}'}F(f)(P')\varphi_{P'}(x).$$

Note that if $E \subset \mathbb{X}$ is a convex set of multitiles, then $\bigcup E$ is the union of areas covered by the maximal multitiles that do not belong to E. Moreover, observe that these multitiles are spatially pairwise disjoint. Hence, we obtain

Corollary 3.35. *Let f be a Schwartz function on $\mathbb{R}$, $E \subset \mathbb{X}$ be a convex set and let $\bigcup \mathcal{P} = \bigcup E$, where $\mathcal{P}$ is a set of pairwise disjoint multitiles. Then*

$$\| \sum_{P \in \mathcal{P}} F(f)(P) \mathbb{1}_E(P) \varphi_P(x) \|_{L^\infty(\mathbb{R})} \lesssim \sup_{P \in \mathbb{X}} |F(P) \mathbb{1}_E(P)| |I_P|^{-1/2}.$$

Chapter 4

Uniform bounds for a Walsh model of the two dimensional bilinear Hilbert transform in local L^1

4.1 Introduction

In the first chapter we viewed the parameter space of the fully two dimensional bilinear Hilbert transform as a three dimensional manifold with subsets corresponding to different well known operators in harmonic analysis. We analysed the interaction of the submanifolds, which gives a wide spectrum of questions concerning the uniform bounds. In this chapter we discuss the uniform estimates for a set of parameters that admit the full two dimensional time-frequency decomposition and degenerate to a singular integral in two dimensions. Namely, we consider $\vec{B} = (B_1, B_2, B_3)$ with

$$B_1 = \begin{pmatrix} -\beta_1 - \beta_2 & 0 \\ 0 & -\gamma_1 - \gamma_2 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \beta_1 & 0 \\ 0 & \gamma_1 \end{pmatrix}, \qquad B_3 = \begin{pmatrix} \beta_2 & 0 \\ 0 & \gamma_2 \end{pmatrix}$$

with $|(\beta_1, \beta_2)|, |(\gamma_1, \gamma_2)| = 1$, such that $\beta_1 \neq \beta_2$, $\gamma_1 \neq \gamma_2$ and prove the uniform bounds for a Walsh model with that triple of matrices as $|\beta_1 - \beta_2|, |\gamma_1 - \gamma_2| \to 0$. We remark that if one assumes $\beta_1 = \gamma_1$, $\beta_2 = \gamma_2$, then the proof follows along the lines of the argument in one dimension, which was presented in Chapter 3. Different speed of convergence of $|\beta_1 - \beta_2|$, $|\gamma_1 - \gamma_2|$ is the main difference compared to the one dimensional case. For convenience of the reader, in the following we introduce the Walsh model of $\mathrm{BHF}_{\vec{B}}$ and then state the main result, which we already did in the introduction of the thesis.

We define a multitile to be the Cartesian product $P_1 \times P_2$, where P_1 and P_2 are tiles. For a multitile $P = P_1 \times P_2$ we denote

$$R_P := I_{P_1} \times I_{P_2}, \qquad \Omega_P := \omega_{P_1} \times \omega_{P_2},$$

where for $j = 1, 2$, I_{P_j} is the spatial interval and ω_{P_j} is the frequency interval of P_j. We overload the notation and represent multitiles either as a product of two tiles $P = P_1 \times P_2$ or as a product of the spatial and the frequency component $P = R_P \times \Omega_P$. The L^2 normalized wave packet associated with a multitile $P = P_1 \times P_2$ is defined as

$$\varphi_P(x,y) := \varphi_{P_1}(x)\varphi_{P_2}(y),$$

where φ_P is the one dimensional Walsh wave packet, defined in Chapter 3.

Given a Schwartz function f on $\mathbb{R}^2$ we associate it to the embedded function via

$$F(f)(P) = \langle f, \varphi_P\rangle.$$

Let f_1, f_2, f_3 be a triple of Schwartz functions on $\mathbb{R}^2$. Set $F_j = F(f_j)$ for $j = 1,2,3$. For a multitile $P = R \times \Omega$, where $\Omega = \omega_1 \times \omega_2$ we denote

$$\Omega^{\circlearrowleft} = \omega_1^{\circlearrowleft} \times \omega_2$$
$$P^{\circlearrowleft} = R \times \Omega^{\circlearrowleft},$$

where $\omega^{\circlearrowleft}$ is the dyadic sibling of a dyadic interval ω.

For a $K \in \mathbb{Z}$ we denote with $\mathcal{R}^K$ the set of all dyadic rectangles $I \times J$ with $|I| = 2^K|J|$ and denote with $\mathbb{X}^K$ the set of all multitiles $P = R \times \Omega$ with $R \in \mathcal{R}^K$. From now on we fix $K \in \mathbb{Z}_+$. Moreover, we set $\mathcal{R} := \mathcal{R}^0$, $\mathbb{X} := \mathbb{X}^0$.

The trilinear form associated to the two dimensional Walsh bilinear Hilbert transform is given formally for $K, L \in \mathbb{Z}_+$ and triples of functions $F_1 \colon \mathbb{X} \to \mathbb{R}$ and $F_j \colon \mathbb{X}^K \to \mathbb{R}$, for $j = 2,3$, by

$$\Lambda_{K,L}(F_1,F_2,F_3) := \sum_{P\in\mathbb{X}} F_1(P^{\circlearrowleft}) \sum_{Q\in P^L} F_2(Q)F_3(Q)h_{R_P}(c(R_Q)),$$

where for $P \in \mathbb{X}$

$$P^L := \{Q \in \mathbb{X}^K : R_Q \subset R_P,\ \Omega_Q = \Omega_P^{K,L}\},$$

where for $\Omega_P = \omega_1 \times \omega_2$, we set $\Omega_P^{K,L} := 2^L\omega_1 \times 2^{L+K}\omega_2$, with $2^L\omega = [2^La, 2^Lb)$ for an interval $\omega = [a,b)$, $c(R_Q)$ is the center of R_Q and

$$h_{R_P}(x,y) = |R_P|^{1/2}\varphi_P(x,y)\varphi_{P^{\circlearrowleft}}(x,y)$$

is the L^2 normalized Haar function.

The goal of this chapter is to prove the uniform bounds for the Walsh model of the two dimensional bilinear Hilbert transform modularizing it as an iterated outer L^p estimate for $\Lambda_{K,L}$ uniform in the parameters K, L, Theorem 4.1, and the Walsh iterated embedding (4.1).

Theorem 4.1. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1,p_2,p_3 < \infty$ and $1/q_1 + 1/q_2 + 1/q_3 > 1$ with $2 < q_1,q_2,q_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $K, L \in \mathbb{Z}_+$, all triples of Schwartz functions f_1, f_2, f_3 on $\mathbb{R}^2$*

$$|\Lambda_{K,L}(F(f_1),F(f_2),F(f_3))| \leq C_{p_1,p_2,p_3}\prod_{j=1}^{3}\|F(f_j)\|_{L^{p_j}\mathcal{L}^{q_j}(S)}.$$

On the right hand side of (2.11) are iterated outer L^p norms developed in [Ura16] that we define precisely in Section 4.2. The Walsh iterated embedding theorem, which we prove in Section 4.5, implies that for $j = 1,2,3$

$$\|F(f_j)\|_{L^{p_j}\mathcal{L}^{q_j}(S)} \leq C_{p_j,q_j}\|f_j\|_{L^{p_j}(\mathbb{R}^2)}. \tag{4.1}$$

We prove Theorem 4.1 in the framework of outer L^p spaces using a counterpart of the multilinear Marcinkiewicz interpolation for outer L^p spaces, Proposition 2.10.

Let us introduce the Walsh model for the bilinear Hilbert transform. For a $P \in \mathbb{X}$ we define the phase plane projections

$$\Pi_P f(x,y) := \langle f, \varphi_P \rangle \varphi_P(x,y), \qquad \Pi_{P^L} f(x,y) := \sum_{Q \in P^L} \Pi_Q(x,y).$$

We define the Walsh model of the two dimensional bilinear Hilbert transform as

$$\mathrm{BHF}_{K,L}(f_1,f_2,f_3) := \int \sum_{P \in \mathbb{X}} \widetilde{\varphi}_P(x,y) \Pi_{P^{\circ}} f_1(x,y) \prod_{j=2}^{3} \Pi_{P^L} f_j(x,y)\, dx\, dy,$$

where $\widetilde{\varphi}_P$ is the L^∞ normalized wave packet. We have the following.

Theorem 4.2. *Let $1/p_1 + 1/p_2 + 1/p_3 = 1$ with $1 < p_1, p_2, p_3 < \infty$. There exists a constant $C_{p_1,p_2,p_3} < \infty$ such that for all $K, L \in \mathbb{Z}_+$ and all triples of Schwartz functions f_1, f_2, f_3 on $\mathbb{R}^2$ the inequality*

$$|\,\mathrm{BHF}_{K,L}(f_1,f_2,f_3)| \leq C_{p_1,p_2,p_3} \prod_{j=1}^{3} \|f_j\|_{L^{p_j}(\mathbb{R}^2)}$$

holds.

Theorem 4.1 coupled with (4.1) implies Theorem 4.2, since

$$\mathrm{BHF}_{K,L}(f_1,f_2,f_3) = \Lambda_{K,L}(F_1,F_2,F_3).$$

Note that the trilinear form $\mathrm{BHF}_{K,L}$ has cancellation in the first frequency component and no cancellation in the second one. Thus, one could analogously define a trilinear form having cancellation in the second component. The uniform estimates for the latter follow using essentially the same arguments as in this chapter and thus we decided not to present them here.

In Section 4.2 we introduce the outer measures in $(\mathbb{R}_+ \times \mathbb{R}_+)^2$. For precise definitions and notions concerning the abstract outer L^p spaces, see Section 2.3 in Chapter 2. In Section 4.3 we present an auxiliary result which lets us dominate L-dependent outer L^p norms with L-independent ones, Proposition 4.12 so that we can apply (4.1), similarly to what we did in Chapter 3. In Section 4.4 we prove Theorem 4.1. In the last section we give a proof of the Walsh iterated embedding Theorem 4.29, which implies (4.1).

We record that the material presented in this chapter is somewhat similar to Chapter 3, hence we skip a couple of proofs that follow exactly the same way. Also, for a more elaborate introduction to the outer L^p spaces, see Chapter 2 and Chapter 3.

4.2 Outer L^p spaces in time-frequency space

In this section we introduce the outer L^p structure on $\mathbb{X}$ and $\mathbb{X}^K$ that we will be working with till the end of the chapter. By $\mathcal{B}(\mathbb{X})$, $\mathcal{B}(\mathbb{X}^K)$ we denote the set of (Borel) functions on $\mathbb{X}$, $\mathbb{X}^K$, respectively. For a fixed $K \in \mathbb{Z}_+$, which corresponds to the ratio between the sides of R_P, the parameter $L \in \mathbb{Z}_+$ is treated similarly to the degeneration parameter in Chapter 3.

4.2.1 Outer measures in time-frequency-scale space

First we introduce K-trees which are the counterpart of the trees in one dimension for the set of multitiles $\mathbb{X}^K$. For a natural number K, let $\widetilde{\Omega} \in \mathcal{R}^{-K}$ be the dyadic parent of a dyadic rectangle $\Omega \in \mathcal{R}^{-K}$.

Definition 4.3 (K-trees and measures). *For $K \in \mathbb{N}$, multitile $R \times \Omega \in \mathbb{X}^K$, we define the K-tree as follows:*

$$T(R, \Omega) := \{P \in \mathbb{X}^K : R_P \subset R, \widetilde{\Omega_P} \supset \Omega\}$$

Moreover we set

$$T^{(ov)} := \{P \in T : \Omega_P \supset \Omega\}, \qquad T^{(lac)} = T \setminus T^{(ov)}.$$

We set $R_T = R$ and $\Omega_T = \Omega$ if $T = T(R, \Omega)$ and we denote the family of K-trees with $\mathbb{T}^K$ and the family of their countable unions with $(\mathbb{T}^K)^{\cup}$. The pre-measure is given by

$$\overline{\mu}^K(T) := |R_T|$$

and it generates μ^K. We also set $\mathbb{T} := \mathbb{T}^0$ and $\mu := \mu^0$.

We introduce the (K, L)-trees which are the counterpart of the L-trees for the set of multitiles $\mathbb{X}^K$. They can be thought of as dilated K-trees, where the K-dependent ratio defining the set $\mathbb{X}^K$ is fixed and the L-dependent scaling is applied to the multitiles simultaneously in both directions, so that the (K, L)-trees are still subsets of $\mathbb{X}^K$.

Definition 4.4 ((K, L)-trees and measures). *For natural numbers K, L, rectangle $R \times \Omega$ with $R \in \mathcal{R}$, $\Omega \in \mathcal{R}^{-K}$, we define the (K, L)-tree as follows:*

$$T(R, \Omega) := \{P \in \mathbb{X}^K : R_P \subset R, \widetilde{\Omega_P} \supset \Omega\} \cap \{P \in \mathbb{X} : |R_P| \leq 2^{-2L-K}|R|\}$$

Moreover we set

$$T^{(ov)} := \{P \in T : \Omega_P \supset \Omega\}, \qquad T^{(lac)} = T \setminus T^{(ov)}.$$

We set $R_T = R$ and $\Omega_T = \Omega$ if $T = T(R, \Omega)$ and we denote the family of (K, L)-trees with $\mathbb{T}^{K,L}$. The pre-measure is given by

$$\overline{\mu}^{K,L}(T) := |R_T|$$

and it generates $\mu^{K,L}$.

4.2.2 Outer measures in time-scale space

To account for space localization we introduce time-scale K-strips as subsets of $\mathbb{X}^K$ and the associated pre-measure.

Definition 4.5 (Strips and measures). *Let for $R \in \mathcal{R}^K$. Define*

$$D(R) = \{P \in \mathbb{X}^K : R_P \subset R\} \qquad \overline{\nu}\big(D(R)\big) := |R|.$$

We denote $R_D = R$ if $D = D(R)$ and the sets of K-strips with $\mathbb{D}^K$. Moreover for $K = 0$ we set $\mathbb{D} := \mathbb{D}^0$. We denote with $(\mathbb{D}^K)^{\cup}$ the family of countable unions of $D \in \mathbb{D}^K$.

Remark 4.6. *If $V \in (\mathbb{D}^K)^{\cup}$, we may assume without loss of generality that $V = \bigcup_{m=1}^{\infty} D_m$, where D_m's are pairwise disjoint and $\in \mathbb{D}^K$. This is because $D \in \mathbb{D}^K$ are based on dyadic rectangles with the same ratios between their sides, meaning that for any two $D, D' \in \mathbb{D}$ we either have $D \cap D' = \emptyset$ or one is contained in the other.*

4.2.3 Sizes in time-frequency-scale space

We define the K-sizes for a fixed $K \in \mathbb{N}$. They are defined analogously to the sizes S, S^2 in Chapter 2.

Definition 4.7 (K-sizes in time-frequency-scale space). *Let $F \in \mathcal{B}(\mathbb{X}^K)$ and $T \in \mathbb{T}^K$. Define*

$$\|F\|_{S^{2,K}(T)} := \left(\frac{1}{|R_T|} \sum_{P \in T^{(lac)}} |F(P)|^2 \right)^{1/2},$$
$$\|F\|_{S^{\infty,K}(T)} := |R_{P_T}|^{-1/2} |F(P_T)|,$$
$$\|F\|_{S^K(T)} := \|F\|_{S^{2,K}(T)} + \|F\|_{S^{\infty,K}(T)}.$$

We also set for $K = 0$, $\|\cdot\|_{S^2} := \|\cdot\|_{S^{2,0}}$, $\|\cdot\|_{S^\infty} := \|\cdot\|_{S^{\infty,0}}$ and

$$\|F\|_{S(T)} := \|F\|_{S^2(T)} + \|F\|_{S^\infty(T)}.$$

We define the top of a (K, L)-tree analogously as in Chapter 3. Before we do that, we shall introduce the order on the set of multitiles $\mathbb{X}^K$.

Definition 4.8 (Order on $\mathbb{X}$). *We write $P \le Q$ for $P, Q \in \mathbb{X}^K$ with $R_P \subset R_Q$ and $\Omega_P \supset \Omega_Q$.*

Definition 4.9 (Tops). *Let $K, L \in \mathbb{Z}_+$, $T \in \mathbb{T}^{K,L}$, $A = K \setminus M$, where $K, M \in (\mathbb{T}^K)^\cup$. Let $\mathcal{P}^1_{A,T}$ be the set of maximal multitiles contained in $A \cap T^{(ov)}$ and $\mathcal{P}^2_{A,T}$ be the set of minimal multitiles contained in $A \cap T^{(ov)}$. We define the top of T with respect to A as*

$$\mathrm{Top}_A(T) := \mathcal{P}^1_{A,T} \cup \mathcal{P}^2_{A,T}.$$

Having defined the tops, we can finally introduce the sizes for the (K, L)-trees.

Definition 4.10 ((K, L)-sizes in time-frequency-scale space). *Let $F \in \mathcal{B}(\mathbb{X}^K)$ and $T \in \mathbb{T}^{K,L}$. Define*

$$\|F\|_{S^{2,K,L}(T)} := \left(\frac{1}{|R_T|} \sum_{P \in T^{(lac)}} |F(P)|^2 \right)^{1/2},$$

$$\|F\|_{S^{\infty,K,L}(T)} := |R_{P_T}|^{-1/2} |F(P_T)|,$$

$$\|F\|_{S_A^{\mathrm{Top}}(T)} := \left(\frac{1}{|R_T|} \sum_{P \in \mathrm{Top}_A(T)} |F(P)|^2 \right)^{1/2},$$

$$\|F\|_{S_A^{K,L}(T)} = \|F\|_{S_A^{\mathrm{Top},K,L}(T)} + \|F \mathbb{1}_A\|_{S^{2,K,L}(T)}.$$

We also define the interpolated size with a parameter $0 < \gamma < 1$

$$\|F\|_{S_{A,\gamma}^{K,L}(T)} := \|F\|^{1-\gamma}_{S_A^{K,L}(T)} \|F \mathbb{1}_A\|^{\gamma}_{L^\infty(S^{\infty,K,L})}.$$

Definition 4.11 (Iterated sizes in time-scale space). *Given an outer measures space with size $(\mathbb{X}, \mu, \|\cdot\|_S)$ and a function $F \in \mathcal{B}(\mathbb{X})$ we define the iterated size as*

$$\|F\|_{\mathcal{L}^q_\mu(S)} := \sup_{D \in \mathbb{D}} |R_D|^{-1/q} \|F \mathbb{1}_D\|_{L^q_\mu(S)}.$$

4.2.4 Choice of parameters and notation

From now on we fix natural numbers $K, L \in \mathbb{Z}_+$ and set $\boldsymbol{\mu} := \mu^{K,L}$, $\|\cdot\|_{\boldsymbol{S}_A} := \|\cdot\|_{S_A^{K,L}}$, $\|\cdot\|_{\boldsymbol{S}_{A,\gamma}} := \|\cdot\|_{S_{A,\gamma}^{K,L}}$. We also use the following notation

$$\|F_1\|_{L^p(S)} := \|F_1\|_{L^p_{\mu}(S)}, \qquad \|F_1\|_{L^p(\L^q(S))} := \|F_1\|_{L^p_{\nu}(\L^q_{\mu}(S))},$$

and for $j = 2, 3$

$$\|F_j\|_{L^p(\boldsymbol{S}_{A,\gamma})} := \|F_j\|_{L^p_{\boldsymbol{\mu}}(\boldsymbol{S}_{A,\gamma})}, \qquad \|F_j\|_{L^p(\L^q(S))} := \|F_j\|_{L^p_{\nu^K}(\L^q_{\mu^K}(S^K))}.$$

When we omit the subscript A, we mean $A = \mathbb{X}$.

4.3 Outer L^p comparison

In this section we prove that (K, L)-dependent outer L^p norms can be bounded by the K-dependent outer L^p norms for $2 < p \leq \infty$. By virtue of this fact, later on in the proof of Theorem 4.4 we may directly apply the iterated embedding theorem for the L-independent sizes, Theorem 4.29.

Throughout this section we assume that

$$\|F\|_{L^2_{\mu^K}(S^K)}, \|F\|_{L^\infty_{\mu^K}(S^K)} < \infty \tag{4.7}$$

are finite and consider $\mathbb{X}^K$ the set of multitiles with

$$\Omega_P \subset [0, N)^2 \tag{4.8}$$

for a large integer N. All bounds in this section will be independent of N, hence, by standard limiting procedure, we may extend them to the whole collection of multitiles. Let $A = V \setminus W$, $V, W \in \mathbb{D}^{\cup}$ be fixed throughout this section. The following result holds.

Lemma 4.12 (Comparison of uniform outer measure spaces). *The following inequality hold for all functions $F \in \mathcal{B}(\mathbb{X}^K)$:*

$$\|F\|_{L^q_{\boldsymbol{\mu}}(\boldsymbol{S}_A)} \lesssim_q \|F\|_{L^q_{\mu^K}(S^K)} \qquad \forall q \in (2, \infty],$$
$$\|F\|_{L^{2,\infty}_{\boldsymbol{\mu}}(\boldsymbol{S}_A)} \lesssim \|F\|_{L^2_{\mu^K}(S^K)},$$

with a constant that is dependent on q but independent of K, L.

The proof of Lemma 4.12 follows along the lines of Lemma 3.12 in Chapter 3, given Lemma 4.14 and Lemma 4.19 below. The only modification is that, because the frequency components of multitiles are rectangles, in the counterpart of the 2-selection algorithm below one has to consider four types of trees $T^{-,-}$, $T^{+,-}$, $T^{-,+}$, $T^{+,+}$ instead of T^-, T^+ as it was done in the one dimensional case. For convenience of the reader we provide the details below.

Definition 4.13 (Strongly disjoint). *We call a set $\mathcal{P} \subset \mathbb{X}^K$ strongly disjoint if for any two distinct $P, P' \in \mathcal{P}$, $P \cap P' = \emptyset$.*

Lemma 4.14 (Outer L^2 sizes of strongly disjoint sets). *Let $\mathcal{P} \subset \mathbb{X}^K$ be a strongly disjoint set of multitiles. Then*

$$\|\mathbb{1}_{\mathcal{P}} F\|_{l^2(\mathbb{X}^K)} \lesssim \|\mathbb{1}_{\mathcal{P}} F\|_{L^2_{\mu^K}(S^K)}$$

Proof. Exactly the same as the proof of Lemma 3.15. □

We introduce two auxiliary selection algorithms, one for $\boldsymbol{S}_A^{\mathrm{Top}}$ and the other one for $\boldsymbol{S}_A^2$.

Definition 4.15 (Top-selection algorithm)**.** *Initially $\mathcal{P} = \emptyset$ and $\mathbb{X}_0 \subset \mathbb{X}^K$. In the n-th step of the procedure we choose a tree $T_n \in \mathbb{T}^{K,L}$ such that*

$$\|F \mathbb{1}_{\mathbb{X}_n}\|_{\boldsymbol{S}_A^{\mathrm{Top}}(T_n)} > \lambda \tag{4.9}$$

and $\boldsymbol{\mu}(T_n)$ is maximal. This possible, because the multitiles in $\mathrm{Top}_A(T)$ can be split into two pairwise disjoint collections of multitiles and as a result, using Lemma 4.14 and (4.7), we obtain an upper bound for $\boldsymbol{\mu}(T)$

$$\boldsymbol{\mu}(T) \leq \lambda^{-2} \|F \mathbb{1}_{\mathrm{Top}_A(T) \cap \mathbb{X}_n}\|^2_{\ell^2(\mathbb{X}^K)} \lesssim \lambda^{-2} \|F\|^2_{L^2_{\mu^K}(S^K)} < \infty.$$

We then add $\mathrm{Top}_A(T_n) \cap \mathbb{X}_n$ to the set $\mathcal{P}$, set $\mathbb{X}_{n+1} := \mathbb{X}_n \setminus T_n$ and iterate the procedure. It will terminate because, by (4.8) there is a lower bound for $\boldsymbol{\mu}(T)$, so by (4.7) and (4.11), there can be only finitely many $T \in \mathbb{T}^{K,L}$ satisfying (4.9).

Remark 4.16. *In the special case $L = 0$ and $A = \mathbb{X}^K$ the above selection algorithm is the selection algorithm for $S^\infty = S^{\mathrm{Top}}$.*

Let Ω be a dyadic rectangle and $\widetilde{\Omega} = \widetilde{\omega}_1 \times \widetilde{\omega}_2$ be the dyadic parent of Ω. We set $\Omega^{a,b} := \widetilde{\omega}_1^a \times \widetilde{\omega}_2^b$ for $a, b \in \{-,+\}$, where ω^- is the left and ω^+ the right half of the dyadic interval ω.

Definition 4.17 (2-selection algorithm)**.** *Initially $\mathcal{P} = \emptyset$ and $\mathbb{X}_0 \subset \mathbb{X}^K$. For a tree $T(R,\Omega) \in \mathbb{T}^{K,L}$ let*

$$T^{+,+}(R,\Omega) := \{P \in T \colon \Omega_T \subset (\widetilde{\Omega_P})^{+,+} \text{ and } \Omega_T \not\subset \Omega_P\},$$

Analogously we define $T^{+,-}$, $T^{-,+}$, $T^{-,-}$. We first describe the selection procedure for $T^{+,+}$. In the n-th step of the procedure we proceed as follows: if there exists a tree $T \in \mathbb{T}^{K,L}$ with

$$\|F \mathbb{1}_{T^{+,+}} \mathbb{1}_{\mathbb{X}_n}\|_{\boldsymbol{S}_A^2(T)} > \lambda \tag{4.10}$$

we select a tree $T_n(R,\Omega)$ which maximizes $\boldsymbol{\mu}(T)$ among $\xi = (\xi_1,\xi_2)$ with the maximal value of $-\xi_1 - \xi_2$, where $\xi = (\xi_1,\xi_2)$ is the middle of Ω. This requirement can be satisfied since such $\boldsymbol{\mu}(T)$ are bounded from above: since the multitiles in $T^{+,+}$ are pairwise disjoint, thanks to Lemma 4.14 and (4.7) we obtain

$$\boldsymbol{\mu}(T) \leq \lambda^{-2} \|F \mathbb{1}_{T^{+,+} \cap \mathbb{X}_n}\|^2_{\ell^2(\mathbb{X}^K)} \lesssim \lambda^{-2} \|F\|^2_{L^2_{\mu^K}(S^K)} < \infty.$$

This in turn implies that $|\Omega|$'s are bounded from below, hence all possible ξ's are all in a discrete set. Together with (4.8) we obtain that $-\xi_1 - \xi_2$ attains its maximum for some ξ for some $T_n := T(R,\Omega)$. We add $S_n = T_n^{+,+} \cap \mathbb{X}_n$ to $\mathcal{P}$ and set $\mathbb{X}_{n+1} := \mathbb{X}_n \setminus T_n$. We iterate the procedure until there are no more trees satisfying (4.10). It will terminate since by (4.8) there is a lower bound for $\boldsymbol{\mu}(T)$, so (4.7) and (4.11) imply that there can be only finitely many $T \in \mathbb{T}^{K,L}$ satisfying (4.10). Analogously we define the selection algorithm for $T^{+,-}$, $T^{-,+}$, $T^{-,-}$ at each step selecting a tree with size larger than λ and maximizing $\boldsymbol{\mu}(T)$ for the maximal value of $-\xi_1 + \xi_2$, $\xi_1 - \xi_2$, $\xi_1 + \xi_2$, respectively.

Since the selection algorithms for $T^{+,+}$, $T^{+,-}$, $T^{-,+}$, $T^{-,-}$ are entirely symmetric and in the proof we will be working only with $T^{+,+}$.

Definition 4.18 (Selection algorithm)**.** *The selection algorithm at level λ consists of running the selection algorithm according to Definition 4.15 at level λ and then the selection algorithm according to Definition 4.17 at level λ, starting with $\mathbb{X}_0 = \mathbb{X}_M$, where is M is the the number of the last iteration of the algorithm in Definition 4.15. The described algorithm yields a collections of trees $\Phi_\lambda = \Phi_\lambda^\infty \cup \Phi_\lambda^2$, where Φ_λ^∞, Φ_λ^2 are trees selected while running* Top*-selection and* 2*-selection algorithm respectively. It also yields a collection of selected tiles $\mathcal{P}_\lambda$.*

Lemma 4.19. *If $\mathcal{P} \subset \mathbb{X}^K$ comes from the selection algorithm given in Definition 4.18, then $\mathcal{P}$ is strongly disjoint.*

Remark 4.20. *In particular an application of Lemma 4.14 implies that, if T_n are the selected trees, $n = 0, 1, 2, ..$, then they satisfy for any $M \in \mathbb{N}$*

$$\sum_{n=0}^{M} \mu(T_n) \lesssim \lambda^{-2} \|F \mathbb{1}_A\|_{L^2(S)}^2. \tag{4.11}$$

Proof. We prove this lemma by contradiction.

First assume that $\mathcal{P}$ comes from the algorithm in Definition 4.15. Suppose that there exist $P, P' \in \mathcal{P}$, such that $P \in \mathrm{Top}_A(T)$, $P' \in \mathrm{Top}_A(T')$ and $P \cap P' \neq \emptyset$, where T, T' are two distinct selected trees. Without loss of generality suppose that T was selected earlier than T'. Then it is not possible $R_{P'} \subset R_P$, because this would mean that $P' \in T$, which is a contradiction. If this is not the case, then we necessarily have $|R_P| < |R_{P'}|$. But since $R_P, R_{P'}$ are members of the partition of R_T, R_T' generated by V, respectively, this would imply that $|R_T| < |R_{T'}|$. This would imply that T' was selected earlier. Contradiction.

Now assume that $\mathcal{P}$ comes from the 2-selection algorithm (for $T^{+,+}$). Without loss of generality assume that $P \in T$, $P' \in T'$, $P \cap P' \neq \emptyset$ and T was selected earlier than T'. It is not possible that $R_{P'} \subset R_P$, since this would mean that P' could have been selected while T was being selected, which is a contradiction. Otherwise $\Omega_{P'} \subset \Omega_P$ and the inclusion is strict. However, this would imply that $-\xi_{T',1} - \xi_{T',2} > -\xi_{T,1} - \xi_{T,2}$ and T' should have been selected before T was selected. Contradiction. □

4.4 Iterated L^p bounds

We fix $K, L \in \mathbb{Z}_+$ and set $\Lambda := \Lambda_{K,L}$. The main result of this section is the following proposition, which combined with Proposition 2.10 implies Theorem 4.1.

Proposition 4.21. *Let $1 < p_1, p_2, p_3 < \infty$ with $\sum_{j=1}^3 1/p_j = 1$ and $2 < q_1, q_2, q_3 < \infty$ with $\sum_{j=1}^3 1/q_j > 1$. Let for $j = 1, 2, 3$, f_j be a Schwartz function and let $F_j := F(f_j)$. Assume that $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$. Then*

$$|\Lambda(F_1 \mathbb{1}_{V_1 \setminus W_1}, F_2 \mathbb{1}_{V_2 \setminus W_2}, F_3 \mathbb{1}_{V_3 \setminus W_3})| \lesssim \prod_{j=1}^{3} \nu(V_j)^{1/p_j} \|F_j \mathbb{1}_{V_j \setminus W_j}\|_{L^\infty \L^{q_j}(S)}. \tag{4.12}$$

Given the next proposition, the proof of (4.12) follows along the lines of the proof of Proposition 3.25 in Chapter 3, with the only difference being the factor γ, one can deal with it similarly as we did in Chapter 2, in the proof of Proposition 2.52.

Proposition 4.22. *Let $1 \leq p_1, p_2, p_3 \leq \infty$ with $\sum_{j=1}^3 1/p_j = 1$. Let for $j = 1, 2, 3$, f_j be a Schwartz function on $\mathbb{R}^2$ and let $F_j := F(f_j)$. Assume that $M_1, N_1 \in \mathbb{T}^\cup$, $M_j, N_j \in (\mathbb{T}^L)^\cup$ for*

$j = 2, 3$ *and* $V_j, W_j \in \mathbb{D}^\cup$ *for* $j = 1, 2, 3$. *Moreover, set* $A = \bigcap_{j=2}^{3}(V_j \setminus W_j)$, $G_j := F_j \mathbb{1}_{V_j \setminus W_j}$, $S_1 := S$ *and* $S_j := \boldsymbol{S}_{A,\gamma}$ *for* $j = 2, 3$. *Then*

$$|\Lambda(G_1 \mathbb{1}_{M_1 \setminus N_1}, G_2 \mathbb{1}_{M_2 \setminus N_2}, G_3 \mathbb{1}_{M_3 \setminus N_3})| \lesssim \prod_{j=1}^{3} \mu(M_j)^{1/p_j} \|G_j \mathbb{1}_{M_j \setminus N_j}\|_{L^\infty(S_j)}.$$

4.4.1 Proof of Proposition 4.22

In this section we prove Proposition 4.22. The proof is similar the proof of its one dimensional counterpart, Proposition 2.53, but we provide here the details. Let $A_j = V_j \setminus W_j$, where $V_j, W_j \in \mathbb{D}^\cup$ for $j = 1, 2, 3$ be fixed throughout this subsection. For $T = T(I, \omega) \in \mathbb{T}$, define $\boldsymbol{T}(I_k, \omega_k) := T^{K,L}(I, \Omega^{K,L}) \in \mathbb{T}^{K,L}$ and for $K = \bigcup_k T(I_k, \omega_k) \in \mathbb{T}^\cup$, define $\boldsymbol{K} := \bigcup_k \boldsymbol{T}(I_k, \omega_k)$.

Exactly the same way as in Chapter 3, Section 5, one can reduce Proposition 4.22 to the following.

Lemma 4.23. *Let* $T \in \mathbb{T}$ *and* $M \in \mathbb{T}^\cup$ *and let* $A = V \setminus W$ *where* $V, W \in \mathbb{D}^\cup$. *Let for* $j = 1, 2, 3$, f_j *be a Schwartz function on* $\mathbb{R}^2$ *and let* $F_j := F(f_j)$. *Then*

$$|\Lambda(F_1 \mathbb{1}_{T \setminus M}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A)| \lesssim \mu(T) \|F_1 \mathbb{1}_{T \setminus M}\|_{L^\infty(S)} \prod_{j=2}^{3} \|F_j \mathbb{1}_A \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{M}}\|_{L^\infty(\boldsymbol{S}_{A,\gamma})}.$$

Proof of Lemma 4.23. First, observe that

$$\begin{aligned}&|\Lambda(F_1 \mathbb{1}_{(T \setminus M)^{(lac)}}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A)| \\ &\lesssim \mu(T) \|F_1 \mathbb{1}_{T \setminus M}\|_{S(T)} \|F_2 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{M}}\|_{\boldsymbol{S}^2(\boldsymbol{T})} \|F_3 \mathbb{1}_A \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{M}}\|_{\boldsymbol{S}^2(\boldsymbol{T})},\end{aligned}$$

just by an application of $(\infty, 2, 2)$-Hölder's inequality, since F_2 and F_3 are restricted to the lacunary tree. Using Lemma 4.25 we bound the right hand side of the previous display by the desired quantity. We still have to show

$$\begin{aligned}&|\Lambda(F_1 \mathbb{1}_{(T \setminus K)^{(ov)}}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A)| \\ &\lesssim \mu(T) \|F_1 \mathbb{1}_{T \setminus K}\|_{L^\infty(S)} \|F_2 \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{K}} \mathbb{1}_A\|_{L^\infty(\boldsymbol{S}_A)} \|F_3 \mathbb{1}_{\boldsymbol{T} \setminus \boldsymbol{K}} \mathbb{1}_A\|_{L^\infty(\boldsymbol{S}_A)}.\end{aligned}$$

Note that we have

$$\Lambda(F_1 \mathbb{1}_{(T \setminus M)^{(ov)}}, F_2 \mathbb{1}_A, F_3 \mathbb{1}_A) = \Lambda(F_1 \mathbb{1}_{T^{(ov)}} \mathbb{1}_B, F_2 \mathbb{1}_C, F_3 \mathbb{1}_C),$$

where

$$B = T \setminus M, \qquad C = A \cap (\boldsymbol{T} \setminus \boldsymbol{M}). \tag{4.13}$$

The right hand side of the penultimate display can be rewritten as

$$\int \sum_{P \in T^{(ov)}} F_1(P^{\circlearrowleft}) \mathbb{1}_B(P) h_{R_P}(x, y) \sum_{Q \in P^{K,L}} F_2(Q) F_3(Q) \mathbb{1}_C(Q) \frac{\mathbb{1}_{R_Q}(x, y)}{|R_Q|} \, dx \, dy, \tag{4.14}$$

where h_{R_P} is the L^2 normalized Haar function. Fix $Q \in \boldsymbol{T}$. Let $\mathcal{P}_1$ be the set of maximal multitiles in $C \cap \boldsymbol{T}^{(ov)}$, $\mathcal{P}_2$ be the set of minimal multitiles in $C \cap \boldsymbol{T}^{(ov)}$; note that $\operatorname{Top}_C(\boldsymbol{T}) = \mathcal{P}_1 \cup \mathcal{P}_2$. Define

$$\overline{\mathcal{Q}} := C \cap \{\overline{Q} \in \boldsymbol{T}^{(lac)} : \overline{Q}^{\circlearrowleft} \leq Q, \; \overline{Q}^{\circlearrowleft} \notin \mathcal{P}_2\}.$$

and $\mathcal{Q} = (C \cap \{Q\}) \cup \overline{\mathcal{Q}}$. Note that $(\mathcal{P}_1 \cap \{\overline{Q} \colon \overline{Q} < Q\}) \cup \mathcal{Q}$ and $\mathcal{P}_2 \cap \{\overline{Q} \colon \overline{Q} \le Q\}$ are two decompositions of the same subset of $\mathbb{R}^4_+$ into pairwise disjoint multitiles. This gives

$$F(Q)\mathbb{1}_C(Q)\varphi_Q = -\sum_{\overline{Q}\in\overline{\mathcal{Q}}} F(\overline{Q})\mathbb{1}_C(\overline{Q})\varphi_{\overline{Q}} - \sum_{\substack{\overline{Q}\in\mathcal{P}_1\\ \overline{Q}<Q}} F(\overline{Q})\varphi_{\overline{Q}} + \sum_{\substack{\overline{Q}\in\mathcal{P}_2\\ \overline{Q}\le Q}} F(\overline{Q})\varphi_{\widetilde{Q}} \tag{4.15}$$

Observe that the above identity can be seen as a discrete integration by parts. Note that by Lemma 4.35 the following cancellation identities hold for any two different $\overline{Q}_1, \overline{Q}_2 \in \overline{\mathcal{Q}}$ and any $\overline{Q} \in \overline{\mathcal{Q}}$, $\overline{\overline{Q}} \in \mathrm{Top}_C(\boldsymbol{T}) \cap \{\widetilde{Q} \colon \widetilde{Q} \le Q\}$.

$$\int h_{R_P}(x,y)\varphi_{\overline{Q}_1}(x,y)\varphi_{\overline{Q}_2}(x,y)\,dx\,dy = 0, \tag{4.16}$$
$$\int h_{R_P}(x,y)\varphi_{\overline{Q}}(x,y)\varphi_{\overline{\overline{Q}}}(x,y)\,dx\,dy = 0.$$

From now on let for a multitile P, $k_P \in \mathbb{Z}$ be such that $2^{2k_P} = |R_P|$. Using $|R_Q|^{-1}\mathbb{1}_{R_Q} = \varphi_Q^2$ and applying (4.15) and (4.16) to F_2 and F_3 in (4.14), there are no cross terms coming from $\boldsymbol{T}^{(lac)}$ and $\mathrm{Top}_C(\boldsymbol{T})$, as well as no terms involving $\overline{Q}_1, \overline{Q}_2 \in \boldsymbol{T}^{(lac)}$ with $\overline{Q}_1 \ne \overline{Q}_2$. Moreover, multitiles in $\mathcal{P}_1$ are spatially pairwise disjoint and the same holds for $\mathcal{P}_2$. Thus, (4.14) equals to a linear combination of

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y) \sum_{\substack{Q\in\boldsymbol{T}^{(lac)}\\ k_Q\le k_P-L}} F_2(Q)F_3(Q)\mathbb{1}_C(Q)\frac{\mathbb{1}_{R_Q}(x,y)}{|R_Q|}\,dx\,dy, \tag{4.17}$$

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y) \sum_{\substack{Q\in\mathrm{Top}_C(\boldsymbol{T})\\ k_Q\le k_P-L}} F_2(Q)F_3(Q)\frac{\mathbb{1}_{R_Q}(x,y)}{|R_Q|}\,dx\,dy, \tag{4.18}$$

$$\int \sum_{P\in T^{(ov)}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y) \prod_{j=2}^{3} \sum_{\substack{Q\in\mathcal{P}_{j-1}\\ k_Q\le k_P-L}} F_j(Q)\varphi_Q(x,y)\,dx\,dy. \tag{4.19}$$

and an integral symmetric to the last one. We first bound (4.17). Changing the order of summation (4.17) equals

$$\int \sum_{Q\in\boldsymbol{T}^{(lac)}} \Big(\sum_{\substack{P\in T^{(ov)}\\ k_Q+L\le k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y)\Big)F_2(Q)F_3(Q)\mathbb{1}_C(Q)\frac{\mathbb{1}_{R_Q}(x,y)}{|R_Q|}\,dx\,dy.$$

Applying $(\frac{1}{\gamma}, \frac{2}{1-\gamma}, \frac{2}{1-\gamma})$-Hölder's inequality in Q and (x,y) this is bounded by

$$\|TF_1\|_{L^{1/\gamma}} \prod_{j=2}^{3} \|GF_j\|_{L^{2/(1-\gamma)}}, \tag{4.20}$$

where

$$TF(x) := \sup_l \Big| \sum_{\substack{P\in T^{(ov)}\\ l\le k_P}} F(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y)\Big| \tag{4.21}$$

and

$$GF(x) := \left(\sum_{P \in T^{(lac)}} |F(P)|^2 \mathbb{1}_C(P) |\varphi_P(x,y)|^2 \right)^{1/2}. \tag{4.22}$$

Applying Lemma 4.26 and Lemma 4.27 we bound (4.20) by the desired quantity. We shall now estimate (4.18). Changing the order of summation we rewrite (4.18) as

$$\int \sum_{Q \in \mathrm{Top}_C(\boldsymbol{T})} (\sum_{\substack{P \in T^{(ov)} \\ k_Q + L \leq k_P}} F_1(P^{\circlearrowleft}) \mathbb{1}_B(P) h_{R_P}(x,y)) F_2(Q) F_3(Q) \frac{\mathbb{1}_{R_Q}(x,y)}{|R_Q|} \, dx \, dy.$$

Using Lemma 4.24 we bound it, after an application of $(\infty, 2, 2)$-Hölder's inequality in Q and $(\frac{1}{\gamma}, \frac{2}{1-\gamma}, \frac{2}{1-\gamma})$-Hölder's inequality in (x,y), by

$$\|TF_1\|_{L^{1/\gamma}} \prod_{j=2}^{3} \|HF_j\|_{L^{2/(1-\gamma)}}, \tag{4.23}$$

where

$$HF(x) := \left(\sum_{P \in \mathrm{Top}_C(\boldsymbol{T})} |F(P)| \mathbb{1}_C(P) |\varphi_P(x,y)|^2 \right)^{1/2} \tag{4.24}$$

The first factor is bounded similarly as before by Lemma 4.26. Concerning the last two, we have the following decomposition of the top.

Lemma 4.24. *Let C be as above. There exists $\Phi' \in \mathbb{T}^{\cup}$ with*

$$\mathrm{Top}_C(\boldsymbol{T}) \subset (\bigcup_{\boldsymbol{T}' \in \boldsymbol{\Phi}'} \mathrm{Top}_A(\boldsymbol{T}') \cup \mathrm{Top}_A(\boldsymbol{T})) \cap C,$$

$$\sum_{\boldsymbol{T}' \in \boldsymbol{\Phi}'} \boldsymbol{\mu}(\boldsymbol{T}') \lesssim \boldsymbol{\mu}(\boldsymbol{T}).$$

Proof. Analogous to the proof of Lemma 3.30. □

Using Lemma 4.24 for F_j with $j = 2, 3$ we obtain

$$\begin{aligned} &\|F_j\|_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})} \\ &\lesssim \|F_j \mathbb{1}_C\|_{\boldsymbol{S}_A^{\mathrm{Top}}(\boldsymbol{T})} + \mu(T)^{-1/2} (\sum_{\boldsymbol{T}' \in \boldsymbol{\Phi}'} \mu(T') \|F_j \mathbb{1}_C\|^2_{\boldsymbol{S}_A^{\mathrm{Top}}(\boldsymbol{T}')})^{1/2} \\ &\lesssim \|F_j \mathbb{1}_C\|_{L^\infty(\boldsymbol{S}_A)}. \end{aligned}$$

The above, together with Lemma 4.28, bounds (4.23).

We are left with estimating (4.19). Note that we can rewrite

$$\prod_{j=2}^{3} \sum_{\substack{Q \in \mathcal{P}_{j-1} \\ k_Q \leq k_P - L}} F_j(Q) \varphi_Q(x,y) = \sum_{\substack{Q_1 \in \mathcal{P}_1 \\ k_{Q_1} \leq k_P - L}} \sum_{\substack{Q_2 \in \mathcal{P}_2 \\ k_{Q_2} \leq k_P - L}} F_2(Q_1) \varphi_{Q_1}(x,y) F_3(Q_2) \varphi_{Q_2}(x,y),$$

Changing the order of summation, (4.19) becomes

$$\int \sum_{Q_1\in\mathcal{P}_1}\sum_{Q_2\in\mathcal{P}_2}\sum_{\substack{P\in T^{(ov)}\\ k_{Q_1}+L\leq k_P\\ k_{Q_2}+L\leq k_P}} F_1(P^{\circlearrowleft})\mathbb{1}_B(P)h_{R_P}(x,y)F_2(Q_1)\varphi_{Q_1}(x,y)F_3(Q_2)\varphi_{Q_2}(x,y)\,dx\,dy.$$

The above is bounded by

$$\int |TF_1(x,y)|\prod_{j=2}^{3}\sum_{Q\in\mathcal{P}_{j-1}}|F_j(Q)||\varphi_Q(x,y)|\,dx\,dy$$

Applying $(\frac{1}{\gamma},\frac{2}{1-\gamma},\frac{2}{1-\gamma})$-Hölder's inequality in (x,y), using spatial disjointness of multitiles $\mathcal{P}_1$ and spatial disjointness of multitiles in $\mathcal{P}_2$ we bound the last display by

$$\|TF_1\|_{L^{1/\gamma}(S)}\prod_{j=2}^{3}\|HF_j\|_{L^{2/(1-\gamma)}}.$$

Another application of Lemma 4.26, Lemma 4.28 and Lemma 4.24 finishes the proof of the proposition. □

At the end of this section we prove the following sequence of lemmata, which we used in the above proof.

Lemma 4.25. *Let C be as in* (4.13). *For $F\in\mathcal{B}(\mathbb{X}^K)$ and $2<p\leq\infty$*

$$\|F\mathbb{1}_C\|_{\boldsymbol{S}^2(\boldsymbol{T})}\lesssim\|F\mathbb{1}_C\|^{2/p}_{\boldsymbol{S}^2(\boldsymbol{T})}\|F\mathbb{1}_C\|^{1-2/p}_{L^\infty(S^\infty)}.$$

Proof. Note that rewriting the left hand side of the inequality as an averaged L^2 norm it is enough to show

$$\mu(T)^{-1/2}\|\sum_{Q\in\boldsymbol{T}^{(lac)}}F(Q)\mathbb{1}_C(Q)\varphi_Q(x,y)\|_{L^2(x,y)}\lesssim\|F\mathbb{1}_C\|_{L^\infty(S^\infty)},$$

$$\mu(T)^{-1/2}\|\sum_{Q\in\boldsymbol{T}^{(lac)}}F(Q)\mathbb{1}_C(Q)\varphi_Q(x,y)\|_{L^2(x,y)}\lesssim\|F\mathbb{1}_C\|_{\boldsymbol{S}^2(\boldsymbol{T})},$$

and interpolate. While the first inequality, follows from Corollary 4.39 since C is a convex collection of multitiles, the second inequality follows by definition. □

Lemma 4.26. *Let T be defined as in* (4.21). *For $F\in\mathcal{B}(\mathbb{X})$ and $2<p\leq\infty$*

$$\|TF\|_{L^p}\lesssim\mu(T)^{1/p}\|F\mathbb{1}_B\|_{L^\infty(S)}.$$

Proof. First of all, observe that the supremum over l can be dominated by the maximal function and can be discarded. Hence, it is enough prove BMO and L^2 bounds: for any dyadic rectangle $R\in\mathcal{R}$

$$\|\sum_{\substack{P\in T^{(lac)}\\ R_P\subset R}}F(P)\mathbb{1}_B(P)h_{R_P}(x,y)\|_{L^\infty(x,y)}\lesssim\|F\mathbb{1}_B\|_{L^\infty(S)},\tag{4.25}$$

and

$$\| \sum_{P\in T^{(lac)}} F(P)\mathbb{1}_B(P)h_{R_P}(x,y)\|_{L^2(x,y)} \lesssim \mu(T)^{1/2}\|F\mathbb{1}_B\|_{L^\infty(S)}. \tag{4.26}$$

The boundedness of (4.25) follows from Corollary 4.39, since we can restrict the sum to the convex collection of multitiles $B \cap \{P\colon R_P \subset R\}$. Similarly one proves (4.26).

Using the log-convexity of L^p norms we obtained the desired inequality. □

Lemma 4.27. *Let G be defined as in* (4.22)*. For $F \in \mathcal{B}(\mathbb{X}^K)$ and $2 < p < \infty$*

$$\|GF\|_{L^p} \lesssim \mu(T)^{1/p}\|F\mathbb{1}_C\|^{2/p}_{\boldsymbol{S}^2(\boldsymbol{T})}\|F\mathbb{1}_C\|^{1-2/p}_{L^\infty(S^\infty)}.$$

Proof. Using interpolation, it is enough to prove the BMO and L^1 bounds

$$\|(GF)^2\|_{BMO} \lesssim \|F\mathbb{1}_C\|^2_{L^\infty(S^\infty)}, \tag{4.27}$$

$$\|(GF)^2\|_{L^1} \lesssim \mu(T)\|F\mathbb{1}_C\|^2_{\boldsymbol{S}^2(\boldsymbol{T})}. \tag{4.28}$$

In order to show (4.27) it is enough to prove that for any dyadic rectangle $R \in \mathcal{R}^K$

$$|R|^{-1}\| \sum_{\substack{Q\in \boldsymbol{T}^{(lac)} \\ R_Q\subset R}} F(Q)\mathbb{1}_C(Q)\varphi_P(x,y)\|^2_{L^2} \lesssim \|F\mathbb{1}_C\|^2_{L^\infty(S^\infty)}$$

This inequality follows from Corollary 4.39, since C is a convex set of multitiles. The bound (4.28) follows from definition. □

Lemma 4.28. *Let H be defined as in* (4.24)*. For $F \in \mathcal{B}(\mathbb{X}^K)$ and $2 < p < \infty$*

$$\|HF\|_{L^p} \lesssim \mu(T)^{1/p}\|F\mathbb{1}_C\|^{2/p}_{\boldsymbol{S}_C^{\mathrm{Top}}(\boldsymbol{T})}\|F\mathbb{1}_C\|^{1-2/p}_{L^\infty(S^\infty)}$$

Proof. Note that we have

$$\|HF\|_{L^\infty} \lesssim \|F\mathbb{1}_C\|_{L^\infty(S^\infty)},$$

since $\mathrm{Top}_C(\boldsymbol{T})$ is a union of two sets of spatially pairwise disjoint multitiles. Moreover

$$\|HF\|_{L^2} \lesssim \mu(T)^{1/2}\|F\mathbb{1}_C\|_{S_C^{\mathrm{Top}}(\boldsymbol{T})},$$

by definition. Using the log-convexity of L^p norms we obtain the desired estimate. □

4.5 Embedding theorem

The goal of this section is to prove the inequality (4.1). We set $\mathbb{X} := \mathbb{X}^K$, $\mathbb{T} := \mathbb{T}^K$, $\mu := \mu^K$, $\nu := \nu^K$ and $S^2 := S^{2,K}$, $S^\infty := S^{\infty,K}$, $S := S^K$. (4.1) is implied by the following embedding theorem.

Theorem 4.29. *Let $1 < p \leq \infty$, $\max(p', 2) < q \leq \infty$, $K \in \mathbb{N}$ and $F := F(f)$. Then for any Schwartz function f on $\mathbb{R}^2$*

$$\|F\|_{L^p Ł^q(S)} \leq C_{p,q}\|f\|_{L^p(\mathbb{R}^2)},$$

where C_p is independent of K.

This is a two dimensional counterpart of the Walsh embedding theorem that was proven in [Ura17]. It follows similarly like the proof in [Ura17], however, in order to prove uniform bounds in parameter K we use the L^p bounds for the strong maximal function [SM93] and Bessel's inequality independent of K, Corollary 4.36.

Before we prove Theorem 4.29 we show a couple of lemmata. From now on all the implied constants are independent of K.

The first lemma dominates the $\|\cdot\|_{L^p(S)}$ norm with $\|\cdot\|_{L^p(S^\infty)}$. It considerably simplifies the selection algorithm in the proof of Theorem 4.29.

Lemma 4.30. *Let $0 < p \le \infty$, f be a Schwartz function on $\mathbb{R}^2$ and $F := F(f)$. Then*

$$\|F\|_{L^p(S)} \lesssim \|F\|_{L^p(S^\infty)}.$$

Proof. Let $0 < p < \infty$. It is enough to show that for any $\lambda > 0$

$$\mu(\|F\|_{L^\infty(S)} > C\lambda) \lesssim \mu(\|F\|_{L^\infty(S^\infty)} > \lambda).$$

Let $\Psi \subset \mathbb{T}^\cup$ be such that

$$\sum_{T\in\Psi} \mu(T) \lesssim \mu(\|F\|_{L^\infty(S^\infty)} > \lambda), \qquad \|F\mathbb{1}_{E^c}\|_{L^\infty(S^\infty)} \le \lambda,$$

where $E = \bigcup \Psi$. It is enough to show that $\|F\mathbb{1}_{E^c}\|_{L^\infty(S^2)} \lesssim \lambda$. Let $T \in \mathbb{T}$. Since $T \setminus E$ is a convex set, using Corollary 4.39

$$\|F\mathbb{1}_{E^c}\|_{S^2(T)} \lesssim \|F\mathbb{1}_{E^c}\mathbb{1}_T\|_{L^\infty(S^\infty)} \le \lambda.$$

This finishes the proof of $0 < p < \infty$. Taking $\lambda = 2\|F\|_{L^\infty(S^\infty)}$ in the above argument we also cover the case $p = \infty$. □

The following lemma appeared already in [War15] and it is a two dimensional Walsh counterpart of Carleson embedding theorem proven in [DT15]. Below, we present a simplified version of the argument in [War15]. We will not need the statement explicitly, however, we shall need the selection algorithm and Lemma 4.33, which are the core of its proof.

Lemma 4.31 (Local L^2 Walsh embedding theorem)**.** *Let $2 < p \le \infty$. Then*

$$\|F\|_{L^p(S)} \lesssim_p \|f\|_{L^p(\mathbb{R}^2)},$$

together with the weak type estimate

$$\|F\|_{L^{2,\infty}(S)} \lesssim \|f\|_{L^2(\mathbb{R}^2)}.$$

Remark 4.32. *Note that applying the argument below with $F\mathbb{1}_D$, $D = D(R_D) \in \mathbb{D}$, one similarly obtains that for $2 < p \le \infty$*

$$\|F\mathbb{1}_D\|_{L^p(S)} \lesssim_p \|f\mathbb{1}_{R_D}\|_{L^p(\mathbb{R}^2)}. \tag{4.29}$$

Proof. We prove the theorem for $p = \infty$, for the weak endpoint $p = 2$ and interpolate applying Proposition 2.9. Using Lemma 4.30, it is enough to show the statement for S^∞.

$p = \infty$: notice that for any $P \in \mathbb{X}$

$$|F(P)| \lesssim \|Mf\|_{L^\infty(\mathbb{R}^2)} \lesssim \|f\|_{L^\infty(\mathbb{R}^2)},$$

where M is the strong maximal function in $\mathbb{R}^2$.

$p = 2$: Let us fix $\lambda > 0$. We run the selection algorithm as in Remark 4.16 obtaining a collection of multitiles $\mathcal{P} = \{P_n : n \in \mathbb{N}\}$ and a family of trees $\Phi = \{T_n : n \in \mathbb{N}\} \subset \mathbb{X}$ whose tops are multitiles in $\mathcal{P}$. Then, it is enough to prove that

$$\sum_{n\in\mathbb{N}} \mu(T_n) = \sum_{n\in\mathbb{N}} |R_{P_n}| \lesssim \lambda^{-2}\|f\|_{L^2(\mathbb{R}^2)}.$$

By Lemma 4.19, $\mathcal{P}$ forms a strongly disjoint subset of $\mathbb{X}$ in the sense of Definition 4.13. We have the following lemma.

Lemma 4.33. *Let $\mathcal{P}$ be a strongly disjoint set of multitiles and let g be a function such that $|F(g)(P)||R_P|^{-1/2} > \lambda$ for every $P \in \mathcal{P}$. Then*

$$\sum_{P\in\mathcal{P}} |R_P| \lesssim \lambda^{-2}\|g\|^2_{L^2(\mathbb{R}^2)}.$$

Proof. We have

$$\sum_{P\in\mathcal{P}} |R_P| \leq \sum_{P\in\mathcal{P}} \lambda^{-2}|F(g)(P)|^2 \lesssim \lambda^{-2}\|g\|_{L^2(\mathbb{R}^2)},$$

where we applied Corollary 4.36 in the second inequality. □

Applying Lemma 4.33, we finish the proof. □

Now we are ready to prove Theorem 4.29.

Proof of Theorem 4.29. First of all, observe that (4.29) implies that for any $D \in \mathbb{D}$

$$\nu(D)^{-1/q}\|F\mathbb{1}_D\|_{L^q(S)} \lesssim \|f\|_{L^\infty(\mathbb{R}^2)},$$

which implies the statement for $p = \infty$. From now on, let us fix $1 < p < \infty$. Using Lemma 4.30, it is enough to show the statement for $S = S^\infty$. Moreover, without loss of generality, assume that $\|f\|_{L^p(\mathbb{R}^2)} = 1$. By interpolation, it is enough to show

$$\|F\|_{L^{p,\infty}Ⱡ^{q,\infty}(S^\infty)} \lesssim \|f\|_{L^p(\mathbb{R}^2)},$$

for any $1 < r < p$ and $q = \max(2, r')$, together with the endpoint with $q = \infty$. Throughout the proof fix $\lambda > 0$ and set

$$\mathbb{I}_\lambda := \{\text{ maximal dyadic rectangle } R \in \mathcal{R}^K : M_r f(x) > \lambda \text{ on } R\},$$

where M_r is the L^r strong maximal function([SM93]) and let $\mathcal{K}_\lambda = \{D(R) : R \in \mathbb{I}_\lambda\} \subset \mathbb{D}$ and $K_\lambda = \bigcup \mathcal{K}_\lambda$.

1. Endpoint $q = \infty$. We have

$$\nu(K_\lambda) \lesssim \lambda^{-p}, \qquad \|F\mathbb{1}_{K^c_\lambda}\|_{L^\infty(S^\infty)} \lesssim \lambda$$

Note that the first inequality follows from L^p boundedness of the M_r strong maximal function for $r < p$. The second inequality follows simply by the definition of K_λ. Together they imply $\|F\|_{L^{p,\infty}Ⱡ^\infty(S^\infty)} \lesssim 1 = \|f\|_{L^p}$.

2. Endpoint $q = \max(2, r')$. It is enough to show that

$$\nu(K_\lambda) \lesssim \lambda^{-p}, \qquad \|F\mathbb{1}_{K^c_\lambda}\|_{Ⱡ^{q,\infty}(S^\infty)} \lesssim \lambda.$$

Note that the first condition is satisfied just like in the previous step. The condition on the right hand side can be rephrased as: for all $D \in \mathbb{D}$ and all $\tau > 0$ there exists E_τ such that

$$\mu(E_\tau) \leq \lambda^q \tau^{-q} \nu(D), \qquad \|F \mathbb{1}_{K_\lambda^c} \mathbb{1}_{E_\tau^c} \mathbb{1}_D\|_{L^\infty(S^\infty)} \leq \tau.$$

Similarly as in Section 4.3, we may assume that (4.7) and (4.8). For any $D \in \mathbb{D}$ let Φ_τ^D be the collection of trees given by the selection algorithm in Remark 4.16 at level τ for $F \mathbb{1}_{K_\lambda^c} \mathbb{1}_D$. Let

$$N_{BMO} := \sup_{D \in \mathbb{D}} \frac{1}{\nu(D)} \sum_{T \in \Phi_\tau^D} \mu(T)$$

Using the aforementioned assumptions, we may consider only $D \in \mathbb{D}$ with $\nu(D)$ uniformly bounded from below and we may assume that there are only finitely may selected trees. Hence, from now on we may assume that $N_{BMO} < \infty$. Note that it is enough to show that $N_{BMO} \lesssim \lambda^q \tau^{-q}$. In order to see that this suffices, take any $D \in \mathbb{D}$, set $E_\tau^D = \bigcup \Phi_\tau^D$ and observe that

$$\|F \mathbb{1}_{K_\lambda^c} \mathbb{1}_D \mathbb{1}_{(E_\tau^D)^c}\|_{L^\infty(S^\infty)} \leq \lambda$$

and

$$\mu(E_\tau^D) \leq \sum_{T \in \Phi_\tau^D} \mu(T) \leq \nu(D) N_{BMO} \lesssim \lambda^q \tau^{-q} \nu(D).$$

Note that since $\|F \mathbb{1}_{K_\lambda^c}\|_{L^\infty(S^\infty)} \leq \lambda$, it is enough to assume $\tau < \lambda$. In consequence we may also assume that $N_{BMO} > 1$, since otherwise there is nothing to prove. We have the following lemma.

Lemma 4.34. *Let $J \in \mathbb{D}$. There exists a function g_J such that*

$$\|g_J\|_{L^2}^2 \lesssim \lambda^2 N_{BMO}^{1-2/q} \nu(J)$$

and for all selected trees $T \in \Phi_\tau^J$ we have $|F(g_J)(P_T)| \mu(T)^{-1/2} > \tau$.

In the proof we perform the Walsh multi-frequency Calderón-Zygmund decomposition, similarly to what is done in [OT11]. This technology was originally developed by Nazarov, Oberlin and Thiele in [NOT09].

Proof. For any selected tree $T \in \Phi_\tau^J$ we have $R_T \subset R_J$. On the other hand for any $T \in \Phi_\tau^J$ and $D \in \mathcal{K}_\lambda$ we either have $R_D \subset R_T$ or $R_D \cap R_T = \emptyset$. Let $\mathcal{P}_D$ be the set of maximal multitiles $P = R_P \times \Omega_P \in D$, such that there exists a tree $T \in \Phi_\tau^J$ with $R_D \subset R_T$ and $\Omega_T \subset \Omega_P$. In this manner we define

$$g_D(x) = \sum_{P \in \mathcal{P}_D} F(P) \varphi_P(x).$$

and split f as follows

$$f(x) = \underbrace{\sum_{D \in \mathcal{K}_\lambda :\, D \subset J} g_D(x) + f \mathbb{1}_{R_J \setminus \bigcup \mathbb{I}_\lambda}(x)}_{g_J} + b(x).$$

Observe that by definition of $\mathbb{I}_\lambda$ and, since we assumed $N_{BMO} > 1$ and $q \geq 2$

$$\|f \mathbb{1}_{R_J \setminus \bigcup \mathbb{I}_\lambda}\|_{L^2}^2 \leq \lambda^2 \nu(J) \leq \lambda^2 N_{BMO}^{1-2/q} \nu(J).$$

On the other hand, observe that counting function $N^J = \sum_{T\in\Phi_\tau^J} 1_{R_T}$ is constant on R_D for each $D \in \mathcal{K}_\lambda$; let N_D^J be its constant value on R_D. In consequence, for every $D \in \mathcal{K}_\lambda$

$$\|g_D\|_{L^2}^2 = \sum_{P\in\mathcal{P}_D} |F(P)|^2 \implies \|g_D\|_{L^2}^2 \le N_D^J |R_D| (\fint_{R_D} |f|)^2$$

and by Corollary 4.36 we also have $\|g_D\|_{L^2}^2 \le |R_D| \fint_{R_D} |f|^2$. Using the Riesz-Thorin interpolation theorem we obtain

$$\|g_D\|_{L^2}^2 \lesssim (\fint |f|^r)^{2/r} (N_D^J)^{1-2/q} |R_D|.$$

Summing the above inequality over $D \in \mathcal{K}_\lambda$ and applying Hölder's inequality we obtain

$$\sum_{D\in\mathcal{K}_\lambda :\, D\subset J} \|g_D\|_{L^2}^2 \le \sum_{D\in\mathcal{K}_\lambda D:\, D\subset J} (\fint_{R_D} |f|^r)^{2/r} (N_D^J)^{1-2/q} |R_D|$$
$$\lesssim \lambda^2 \int_{R_J} (N^J)^{1-2/q} \lesssim \lambda^2 (\int_{R_J} N^J)^{1-2/q} \nu(J)^{2/q} \lesssim \lambda^2 N_{BMO}^{1-2/q} \nu(J).$$

Finally, note that for any $T \in \Phi_\tau^J$ we have $F(b)(P_T) = 0$, hence $|F(g_J)(P_T)|\mu(T)^{-1/2} > \tau$. This finishes the proof of the lemma. □

Let $J \in \mathbb{D}$ be a such that N_{BMO} is almost attained and let g_J be like in the previous lemma. Applying Lemma 4.34 and Lemma 4.33 we obtain

$$N_{BMO}\nu(J) \lesssim \sum_{T\in\Phi_\tau^J} \mu(T) \lesssim \tau^{-2} \|g_J\|_{L^2}^2 \lesssim \lambda^2 \tau^{-2} N_{BMO}^{1-2/q} \nu(J).$$

This gives

$$N_{BMO} \lesssim \lambda^q \tau^{-q},$$

which concludes the proof. □

4.6 Appendix - Walsh wave packets in two dimensions

In this chapter we used the following facts about the Walsh wave packets in two dimensions.

Lemma 4.35. *If two multitiles P, Q are disjoint, then φ_P and φ_Q are orthogonal, i.e. $\langle \varphi_P, \varphi_Q \rangle = 0$.*

The above lemma was proven in [War15]. As a corollary we obtain Bessel's inequality for the wave packets in two dimensions.

Corollary 4.36. *If $\mathcal{P}$ is a set of pairwise disjoint multitiles and let f be a Schwartz function on $\mathbb{R}^2$, then*

$$\sum_{P\in\mathcal{P}} |F(f)(P)|^2 \le \|f\|_{L^2(\mathbb{R}^2)}.$$

We shall also need the following.

Lemma 4.37. *Let $\mathcal{P}$ be a finite collection of multitiles and assume that a multitile Q is covered by the multitiles in $\mathcal{P}$. Then φ_Q is in the linear span of $\{\varphi_P : P \in \mathcal{P}\}$.*

The lemma was proven in [War15]. As a corollary we obtain the following.

Corollary 4.38. *If $\mathcal{P}$, $\mathcal{P}'$ are two different collection of multitiles, each of which is pairwise disjoint and $\bigcup \mathcal{P} = \bigcup \mathcal{P}'$, then for any Schwartz function f on $\mathbb{R}^2$*

$$\sum_{P\in\mathcal{P}} F(f)(P)\varphi_P(x) = \sum_{P'\in\mathcal{P}'} F(f)(P')\varphi_{P'}(x).$$

Note that if $E \subset \mathbb{X}^K$ is a convex set of multitiles, then $\bigcup E$ can be represented as a union of maximal multitiles that do not belong to E. Moreover, observe that these multitiles are spatially pairwise disjoint. Hence, we obtain

Corollary 4.39. *Let f be a Schwartz function on $\mathbb{R}^2$, $E \subset \mathbb{X}^K$ be a convex set of multitiles and let $\bigcup \mathcal{P} = \bigcup E$, where $\mathcal{P}$ is a set of pairwise disjoint multitiles. Then*

$$\| \sum_{P\in\mathcal{P}} F(f)(P)\mathbb{1}_E(P)\varphi_P(x,y)\|_{L^\infty(\mathbb{R}^2)} \lesssim \sup_{P\in\mathbb{X}^K} |F(P)\mathbb{1}_E(P)||R_P|^{-1/2}.$$

Curriculum vitae

2010/10 - 2013/06	B.Sc. in Mathematics University of Wrocław, Poland
2013/10 - 2015/07	M.Sc. in Mathematics, University of Bonn, Germany
2015/10 - 2018/09	Ph.D. studies in Mathematics, University of Bonn, Germany

2013/10 - 2015/09	Qualifying Hausdorff Scholarship Master's scholarship by Bonn International Graduate School of Mathematics
2015/10 - 2018/09	Hausdorff Scholarship PhD scholarship by Bonn International Graduate School of Mathematics

Bibliography

[Ben+08] Jonathan Bennett et al. „The Brascamp–Lieb inequalities: finiteness, structure and extremals“. In: *Geometric and Functional Analysis* 17.5 (2008), pp. 1343–1415.

[Bou88] Jean Bourgain. *Return times of dynamical systems.* Tech. rep. 1988.

[Cal65] Alberto P Calderón. „Commutators of singular integral operators“. In: *Proceedings of the National Academy of Sciences* 53.5 (1965), pp. 1092–1099.

[Car66] Lennart Carleson. „On convergence and growth of partial sums of Fourier series“. In: *Acta Mathematica* 116.1 (1966), pp. 135–157.

[CM75] Ronald R Coifman and Yves Meyer. „On commutators of singular integrals and bilinear singular integrals“. In: *Transactions of the American Mathematical Society* 212 (1975), pp. 315–331.

[CZ05] Jiecheng Chen and Xiangrong Zhu. „A note on BMO and its application“. In: *Journal of mathematical analysis and applications* 303.2 (2005), pp. 696–698.

[DDP13] Ciprian Demeter and Francesco Di Plinio. „Endpoint bounds for the quartile operator“. In: *Journal of Fourier Analysis and Applications* 19.4 (2013), pp. 836–856.

[Dem+08] Ciprian Demeter et al. „Breaking the duality in the return times theorem“. In: *Duke Mathematical Journal* 143.2 (2008), pp. 281–355.

[DPO15] Francesco Di Plinio and Yumeng Ou. „A modulation invariant Carleson embedding theorem outside local L^2“. In: *arXiv preprint arXiv:1510.06433* (2015).

[DT10] Ciprian Demeter and Christoph Thiele. „On the two-dimensional bilinear Hilbert transform“. In: *American journal of mathematics* 132.1 (2010), pp. 201–256.

[DT15] Yen Do and Christoph Thiele. „L^p theory for outer measures and two themes of Lennart Carleson united“. In: *Bulletin of the American Mathematical Society* 52.2 (2015), pp. 249–296.

[Dur17] Polona Durcik. „The continuous analysis of entangled multilinear forms and applications“. PhD thesis. Universität Bonn, 2017.

[Fef73] Charles Fefferman. „Pointwise convergence of Fourier series“. In: *Annals of Mathematics* (1973), pp. 551–571.

[GL04] Loukas Grafakos and Xiaochun Li. „Uniform bounds for the bilinear Hilbert transforms, I“. In: *Annals of mathematics* (2004), pp. 889–933.

[Gre+16] Philip Gressman et al. „On a trilinear singular integral form with determinantal kernel“. In: *Proceedings of the American Mathematical Society* 144.8 (2016), pp. 3465–3477.

[Kov12] Vjekoslav Kovač. „Boundedness of the twisted paraproduct“. In: *Revista Matematica Iberoamericana* 28.4 (2012), pp. 1143–1165.

[KTZK15] Vjekoslav Kovač, Christoph Thiele, and Pavel Zorin-Kranich. „Dyadic triangular Hilbert transform of two general functions and one not too general function“. In: *Forum of Mathematics, Sigma*. Vol. 3. Cambridge University Press. 2015.

[Li06] Xiaochun Li. „Uniform bounds for the bilinear Hilbert transforms, II“. In: *Revista Matematica Iberoamericana* 22.3 (2006), pp. 1069–1126.

[LT97] Michael Lacey and Christoph Thiele. „Lp Estimates on the Bilinear Hilbert Transform for“. In: *Annals of Mathematics* (1997), pp. 693–724.

[LT99] Michael Lacey and Christoph Thiele. „On Calderón's conjecture“. In: *Annals of mathematics* 149 (1999), pp. 475–496.

[MTT02] Camil Muscalu, Terence Tao, and Christoph Thiele. „Uniform estimates on multilinear operators with modulation symmetry“. In: *Journal d'Analyse Mathématique* 88.1 (2002), pp. 255–309.

[Mus+] Camil Muscalu et al. „A short proof of the coifman-meyer multilinear theorem“. In: ().

[Mus14] Camil Muscalu. „Calderón commutators and the Cauchy integral on Lipschitz curves revisited I. First commutator and generalizations“. In: *Revista Matemática Iberoamericana* 30.2 (2014), pp. 727–750.

[NOT09] Fedor Nazarov, Richard Oberlin, and Christoph Thiele. „A Calderón Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain“. In: *arXiv preprint arXiv:0912.3010* (2009).

[OT11] Richard Oberlin and Christoph Thiele. „New uniform bounds for a Walsh model of the bilinear Hilbert transform“. In: *Indiana university mathematics journal* (2011), pp. 1693–1711.

[SM93] Elias M Stein and Timothy S Murphy. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Vol. 3. Princeton University Press, 1993.

[Thi02a] Christoph Thiele. „A uniform estimate“. In: *Annals of mathematics* 156.2 (2002), pp. 519–563.

[Thi02b] Christoph Thiele. „A uniform estimate for quartile operators“. In: *Revista Matemática Iberoamericana* 18.1 (2002), pp. 115–134.

[Thi06] Christoph Thiele. *Wave packet analysis*. 105. American Mathematical Soc., 2006.

[Thi95] Christoph Martin Thiele. „Time-frequency analysis in the discrete phase plane“. PhD thesis. Yale University, 1995.

[Ura16] Gennady Uraltsev. „Variational Carleson embeddings into the upper 3-space“. In: *arXiv preprint arXiv:1610.07657* (2016).

[Ura17] Gennady Uraltsev. „Time-Frequency Analysis of the Variational Carleson Operator using outer-measure Lp spaces“. PhD thesis. Universität Bonn, 2017.

[War15] Michał Warchalski. „Uniform estimates in time-frequency analysis“. MA thesis. Universität Bonn, 2015.