BEYOND THE SCOPE OF DOUBLING MEASURES: WEIGHTED MARTINGALE TRANSFORM VIA OUTER MEASURE SPACES

CHRISTOPH THIELE, SERGEI TREIL, AND ALEXANDER VOLBERG

Abstract. A new approach to Wittwer's inequality is given, resulting in a new version with arbitrary and in particular non-doubling reference measure. The approach combines embeddings into outer measure spaces with a core concavity argument. Specifically, we prove that finiteness of the quantity $Q = \sup_J \langle w\rangle_J \langle w^{-1}\rangle_J$, defined through averages $\langle . \rangle_J$ relative to the reference measure ν, implies that each martingale transform relative to ν is bounded in $L^2(w\, d\nu)$.

1. Notations and results

Following [12], by Wittwer's inequality we mean an inequality of the form

$$\sum_{I\in\mathscr{D}} |(f,h_I)(g,h_I)| \le A(\sup_{I\in\mathscr{D}}[w]_I[w^{-1}]_I)\|f\|_{w^{-1}}\|g\|_w \ .$$

Here the interval I runs through the set $\mathscr{D}$ of all dyadic intervals of some dyadic lattice on $\mathbb{R}$, the function h_I is the $L^2(\mathbb{R})$ normalized Haar function on I, and the brackets $(,)$ denote the inner product in $L^2(\mathbb{R})$. On the right-hand-side of the inequality appears a weight w, a measurable function such that w and w^{-1} are locally integrable. The average of w over an interval I is denoted by $[w]_I$. We let $d\nu$ be Lebesgue measure and identify w with the measure $wd\nu$. We denote by $\|f\|_w$ the norm in the weighted Hilbert space $L^2(w)$ defined by $\|f\|_w^2 = \int |f(x)|^2 w(x)\, d\nu(x)$. The constant A is absolute. By the duality relation

$$\|h\|_w = \sup_{\|f\|_{w^{-1}}=1} |(f,h)| \ ,$$

Wittwer's inequality is the dual form of an $L^2(w)$ norm bound of an arbitrary Haar multiplier operator with bounded coefficients, also called zero-shift operator or martingale transform operator for the dyadic martingale. Hence Wittwer's inequality states that the norm of such an operator is at most proportional to the A_2 constant $\sup_{I\in\mathscr{D}}[w]_I[w^{-1}]_I$ of the weight. The present paper originates from an attempt to understand and possibly simplify the well known proof of Wittwer's inequality using the language of L^p theory for outer measure spaces delevoped in [1]. As a result of these efforts, we found a generalization of Wittwer's inequality relative to arbitrary and in particular in general non-doubling reference measures ν, see Theorem 1.1 below. This level of generality of Wittwer's inequality came as a surprise to us. We describe the general setup in detail. Let ν be a locally finite positive measure on $\mathbb{R}$. We shall make the qualitative assumption $\nu(I) > 0$ for each dyadic interval I, so that all our

2010 *Mathematics Subject Classification.* 42B20, 42B35, 47A30.

Key words and phrases. Weighted outer measure spaces, Wittwer's inequality, Bellman function, martingale transform, bilinear embeddings.

CT is supported by the NSF grant DMS-1001535.

ST is partially supported by the NSF grant DMS-1301579.

AV is partially supported by the NSF grant DMS-1265549 and by the Hausdorff Institute for Mathematics, Bonn, Germany.

expressions are well defined. Starting from arbitrary ν one might consider the measure $\nu+\epsilon$ for a small constant ϵ to obtain strict positivity of the new measure on all I. Since none of our estimates will depend on ϵ, one may consider ϵ tending to zero whenever one applies our main theorem in a setting that is well defined for arbitrary ν. Hence our qualitative assumption is not very restrictive. We consider the Hilbert space $L^2(d\nu)$, the Hilbert space whose norm can be defined by $\|f\|_\nu^2 = \int |f|^2\, d\nu$. We write $(f,g)_\nu$ for the inner product in this Hilbert space. We introduce a new bracket for averages with respect to the measure ν:

$$\langle w\rangle_I = \frac{1}{\nu(I)}\int_I w(x)\, d\nu(x)\,.$$

The Haar system with respect to ν is the orthonormal basis in $L^2(d\nu)$ consisting of functions h_I^ν parameterized by the dyadic intervals I, supported on this interval and constant on each of the two dyadic children of this interval. We will no longer be explicity concerned with the special case of ν being the Lebesgue measure, hence we will identify w with the measure $wd\nu$ and w^{-1} with the measure $w^{-1}d\nu$ for the general measure ν.

Theorem 1.1. *We have for every weight function w such that w and w^{-1} are measurable and absolutely continuous with respect to ν and locally integrable with respect to ν the inequality*

$$\sum_{I\in\mathscr{D}} |(f,h_I^\nu)_\nu(g,h_I^\nu)_\nu| \le A(\sup_{I\in\mathscr{D}}\langle w\rangle_I\langle w^{-1}\rangle_I)\|f\|_{w^{-1}}\|g\|_w \tag{1.1}$$

for some absolute constant A. Moreover, we have for each dyadic interval I the local inequality

$$|\langle g\rangle_I\langle f\rangle_I|\,\nu(I) + \sum_{J\in\mathscr{D}(I)} |(f,h_J^\nu)_\nu(g,h_J^\nu)_\nu| \le A(\sup_{J\in\mathscr{D}(I)}\langle w\rangle_J\langle w^{-1}\rangle_J)\|f1_I\|_{w^{-1}}\|g1_I\|_w\ , \tag{1.2}$$

where $\mathscr{D}(I)$ denotes the collection of dyadic intervals contained in I.

The number $\sup_{I\in\mathscr{D}}\langle w\rangle_I\langle w^{-1}\rangle_I$ may be called the relative A_2 constant of w with respect to the reference measure ν. Note that the global inequality follows from the local inequality by a limiting argument, approximating f and g by functions of bounded support. Our proof of the theorem will then only be concerned with the local inequality. Note further that the first term in the local inequality, the martingale average term, is easily estimated by Cauchy–Schwarz:

$$|\langle g\rangle_I\langle f\rangle_I|\,\nu(I) \le A((\langle w\rangle_I\langle w^{-1}\rangle_I)^{1/2}\|f1_I\|_w\|g1_I\|_{w^{-1}}\ .$$

Since the A_2 constant is always bounded below by 1 by another Cauchy–Schwarz,

$$1 = \nu(J)^{-1}\int_J w^{1/2}w^{-1/2}d\nu \le (\langle w\rangle_J\langle w^{-1}\rangle_J)^{1/2}\ ,$$

the square root on the right-hand-side of the previous display can be ignored and one obtains the desired bound for the martingale average term. It will then suffice to estimate the martingale difference term on the left hand side of (1.2). Note further that the theorem can be interpreted on any martingale which is equivalent to the real line or a dyadic interval. Consider the family of martingale transform operators on an interval I:

$$g \to \varepsilon\langle g\rangle_I 1_I + \sum_{J\in\mathscr{D}(I)} \varepsilon_J(g,h_J^\nu)_\nu h_J^\nu,\ \ \varepsilon,\varepsilon_J = \pm 1\ .$$

The local Wittwer inequality (1.2) means precisely that the martingale transform operators are bounded in $L^2(w)$ uniformly over the choices of $\varepsilon,\varepsilon_J$, with norm at most the relative A_2 constant. Conversely, if for fixed w the martingale operators are uniformly bounded in $L^2(w)$ by some constant C, then we can easily conclude that the A_2 constant of w relative to ν on the interval I is finite and bounded by C^2. Namely, for any subinterval $J\subset I$ there is a martingale transform operator T which averages functions relative to ν on subintervals of length $|J|$, so that

$$T(g1_J) = \langle g\rangle_J 1_J \ .$$

Then we have

$$\langle w\rangle_J\langle w^{-1}\rangle_J = (T(w1_J), w^{-1})_\nu \nu(J)^{-1}$$
$$\leq C\|w1_J\|_{w^{-1}}\|w^{-1}1_J\|_w \nu(J)^{-1} = C(\langle w\rangle_J\langle w^{-1}\rangle_J)^{1/2} \ .$$

Dividing the inequality by the square root term appearing on the right-hand-side and squaring the resulting inequality proves the claim.

Corollary 1.2. *The condition that the relative A_2 constant is finite is equivalent to uniform boundedness in $L^2(w)$ of all martingale transform operators on I relative to ν.*

While our simple necessity argument leaves a discrepancy in power dependence on the A_2 constant between the sufficient an necessary condition, it is well known that in the case of ν being Lebesgue measure the first power in the A_2 constant in Wittwer's inequality is sharp. This can be seen by testing similar functions as above but supported on sibling dyadic intervals.

We outline the proof of the generalized Wittwer inequality. We consider the Haar system h_I^w relative to $L^2(w)$ and expand the Haar function h_I^ν as

$$h_I^\nu = \alpha_I^w h_I^w + \rho_I^w \chi_I^w \ ,$$

where χ_I^w is a multiple of the indicator function of I, normalized in $L^2(w)$. Similarly we proceed with w^{-1} in place of w. The martingale difference term in Wittwer's inequality involves a tensor product of two Haar functions, which we expand as

$$h_I^\nu \otimes h_I^\nu = I + II + III + IV =$$
$$\alpha_I^w \alpha_I^{w^{-1}} h_I^w \otimes h_I^{w^{-1}} + \alpha_I^w \rho_I^{w^{-1}} h_I^w \otimes \chi_I^{w^{-1}} + \rho_I^w \alpha_I^{w^{-1}} \chi_I^w \otimes h_I^{w^{-1}} + \rho_I^w \rho_I^{w^{-1}} \chi_I^w \otimes \chi_I^{w^{-1}} \ .$$

By the triangle inequality we reduce Wittwer's inequality to four inequalities along the above decomposition. The first inequality is easiest and estimated by an $l^\infty \times l^2 \times l^2$ Hölder inequality

$$\sum_{J\in\mathscr{D}(I)} |\alpha_J^w \alpha_J^{w^{-1}} (f, h_J^w)_\nu (g, h_J^{w^{-1}})_\nu|$$

$$\leq \sup_{J\in\mathscr{D}(I)} |\alpha_J^w \alpha_J^{w^{-1}}| \left(\sum_{J\in\mathscr{D}(I)} |(f,h_J^w)_\nu|^2\right)^{1/2} \left(\sum_{J\in\mathscr{D}(I)} |(g,h_J^{w^{-1}})_\nu|^2\right)^{1/2} \ .$$

Using the trivial but elucidating identification $(f, h_J^w)_\nu = (fw^{-1}, h_J^w)_w$ and analoguously for g, and using orthonormality of the respective Haar systems, we estimate term I by

$$\leq (\sup_{J\in\mathscr{D}(I)} |\alpha_I^w \alpha_I^{w^{-1}}|)\|fw^{-1}1_I\|_w \|gw1_I\|_{w^{-1}} = (\sup_{J\in\mathscr{D}(I)} |\alpha_I^w \alpha_I^{w^{-1}}|)\|f1_I\|_{w^{-1}} \|g1_I\|_w \ .$$

The explicit calculation of the α-s in (2.6) below reveals that

$$\sup_{J\in\mathscr{D}(I)} |\alpha_I^w \alpha_I^{w^{-1}}| \leq \sup_{J\in\mathscr{D}(I)} \langle w\rangle_J \langle w^{-1}\rangle_J$$

and concludes the proof for term I. The remaining terms, of which by symmetry we only need to estimate III and IV, will be done in analogous but more involved steps. We can write the above Hölder inequality as

$$\sum_{J\in\mathscr{D}(I)} |\alpha_J^w \alpha_J^{w^{-1}} (f, h_J^w)_\nu (g, h_J^{w^{-1}})_\nu| \leq \|\alpha_J^w \rho_J^{w^{-1}}\|_{l^\infty} \times \|(f,h_J^w)_\nu\|_{l^2} \times \|(g,h_J^{w^{-1}})_\nu\|_{l^2} \ .$$

We will replace the norms on the right-hand-side by more elaborate outer L^p norms as introduced in [1]. The estimation by such norms will use a successive combination of Cauchy–Schwarz and outer measure space Hölder inequalities, and has the effect of separating the contributions from the function f, the function g, and the factor involving α-s and ρ-s. We will review the important definitions and facts about outer L^p theory in the present paper. The use of orthogonality of the Haar systems above can be summarized as

$$\|(fw^{-1}, h_J^w)_w\|_{l^2} \leq C\|fw^{-1}\|_w \ ,$$

an inequality whose more involved pendants in outer L^p theory we will call embedding theorems. Somewhat at the heart of the matter will then be the control of the terms involving the coefficients α and ρ, these terms depends only on the weight w and the measure ν, but not on the functions f or g. In the above case I, the relevant embedding theorem turns out quite trivial and reads as

$$\|\alpha_J^w \alpha_J^{w^{-1}}\|_{l^\infty} \leq C \sup_{J\subset I} \langle w\rangle_J \langle w^{-1}\rangle_J \ .$$

This inequality initially may seem as a bilinear embedding theorem in the functions w and w^{-1}. However, the right-hand-side is not merely a product of two norms, one measuring w and the other measuring w^{-1}, but it is an intermingled expression which at each interval J separately multiplies quantities involving w and w^{-1}. Ultimately one needs to incorporate the nonlinear dependence of w and w^{-1} at every interval J. Thus the latter embedding theorem is of more nonlinear type. In cases III and IV we will use concavity arguments, that is arguments of Bellmann function type, to prove these nonlinear embedding theorems. A specialty in case IV occurs, where we do not separate the functions f and g from each other, but prove directly a bilinear embedding theorem for these two functions. The reason is that one has to again use the interplay between w and w^{-1} at every scale.

We comment a bit on the history of the subject. Wittwer's inequality, especially linearity in the A_2 constant, proved in [12] for $\nu = dx$, was a quite non-trivial fit. The proof was heavily based on a rather difficult two weight estimate of [6], prompting us to seek a simpler proof. The sharp dependence on the A_2 constant was later used very essentially by Petermichl–Volberg [8] to find a sharp weighted estimate for the Ahlfors–Beurling operator. In turn their result on the Ahlfors–Beurling operator applies to give sharp regularity estimates for the Beltrami euqtion and solve the Astala–Iwaniec–Saksman problem. For the Hilbert transform the linear estimate in terms of the A_2 characteristic was done later by Petermichl in [9]. The full generality of the A_2 conjecture, that is the linear–in–A_2 bound for all Calderón–Zygmund operators, was done by Hytönen in [2]. More recently there were different proofs, e. g. in [3] and in Lerner's work [4]. All this has been done with respect to underlying Lebesgue measure and was later adopted to the case of doubling underlying measure. Recently Treil–Volberg [10] proved a result about weighted singular shifts in the non-homogeneous situation. In this paper the authors asked specifically about the conditions on the boundedness of the martingale transform for arbitrary ν. We answer this question here.

For the unweighted case of the boundedness of martingale transform in $L^p(\nu)$ for arbitrary ν one can consult the celebrated series of papers by Burkholder. A very interesting approach to other dyadic singular operators and their boundedness in unweighted $L^p(\nu)$ one can find in [**?**].

The authros are grateful to Carlos Pérez for organizing an inspiring Summer School in Santander, Spain, where the research on this paper was initiated.

2. Boundedness of term III

This section is devoted to proving the desired bound for term III, that is

$$\sum_{J\in\mathscr{D}(I)} |\rho_J^w \alpha_J^{w^{-1}} (f,\chi_J^w)_\nu (g,h_J^{w^{-1}})_\nu| \le A(\sup_{J\in\mathscr{D}(I)} \langle w\rangle_J \langle w^{-1}\rangle_J) \|f1_I\|_{w^{-1}} \|g1_I\|_w \ .$$

We denote

$$Q := \sup_{J\in\mathscr{D}(I)} \langle w\rangle_J \langle w^{-1}\rangle_J \ . \tag{2.1}$$

We begin with applying the Cauchy–Schwarz inequality, estimating the left-hand-side of the desired inequality by

$$\left(\sum_{J\in\mathscr{D}(I)} |\rho_J^w \alpha_J^{w^{-1}} (f,\chi_J^w)_\nu|^2\right)^{1/2} \left(\sum_{J\in\mathscr{D}(I)} |(g,h_J^{w^{-1}})_\nu|^2\right)^{1/2} \ .$$

The second factor is estimated as in the considerations for term I by $\|g1_I\|_w$. It then remains to prove

$$\left(\sum_{J\in\mathscr{D}(I)} |\rho_J^w \alpha_J^{w^{-1}} (f,\chi_J^w)_\nu|^2\right)^{1/2} \le AQ\|f1_I\|_{w^{-1}d\nu} \ . \tag{2.2}$$

We intend to further separate f from the product of α and ρ. This will be done by an outer Hölder inequality. We first review the relevant notions from [1]. We consider the collection $\mathscr{D}(I)$ of dyadic intervals as an outer measure space. The *outer measure* of an arbitrary subset $\mathcal{K}$ of $\mathscr{D}(I)$ is defined by coverings of $\mathcal{K}$ by collections $\mathscr{D}(J)$. Precisely, define the outer measure

$$\mu_\nu(\mathcal{K}) := \inf_{\mathcal{K}\subset\bigcup_{J\in\mathcal{J}}\mathscr{D}(J)} \left(\sum_{J\in\mathcal{J}} \nu(J)\right) \ .$$

We also define an outer measure relative to w:

$$\mu_w(\mathcal{K}) = \inf_{\mathcal{K}\subset\bigcup_{J\in\mathcal{J}}\mathscr{D}(J)} \left(\sum_{J\in\mathcal{J}} w(J)\right) \ .$$

and similarly for w^{-1}. Given a function $H : \mathscr{D}(I) \to \mathbb{C}$ we define the following *sizes* of this function

$$S_\nu^p(H,\mathscr{D}(J)) := \left(\frac{1}{\nu(J)} \sum_{K\in\mathscr{D}(J)} |H(K)|^p \nu(K)\right)^{1/p} \ ,$$

$$S_w^p(H,\mathscr{D}(J)) := \left(\frac{1}{w(J)} \sum_{K\in\mathscr{D}(J)} |H(K)|^p w(K)\right)^{1/p} \ ,$$

and

$$S^\infty(H,\mathscr{D}(J)) := \sup_{K\in\mathscr{D}(J)} |H(K)| \ ,$$

where the last definition is independent of the underlying measure ν or w. We define analogous sizes for w^{-1}. For S any of the above sizes, we define the *outer L^∞ norm*

$$\|H\|_{L^\infty(\mathscr{D}(I),\nu,S)} := \sup_{J\in\mathscr{D}(I)} S(H,\mathscr{D}(J)) \ .$$

Moreover, define for $\lambda > 0$ and any of the sizes S the *super level measure*

$$\mu_\nu(S(H) > \lambda) := \inf_{\mathcal{K}\subset\mathscr{D}(I): \|H1_{\mathscr{D}(I)\setminus\mathcal{K}}\|_{L^\infty(\mathscr{D}(I),\nu,S)} < \lambda} \mu_\nu(\mathcal{K}) \ .$$

This is a generalization of the usual outer measure of the set of all J with $H(J) > \lambda$, but in general it is an abstract version of this notion which only coincides with the classical case if S is the S^∞ size. Finally we define the *outer L^p norm* with respect to the size S for $1 < p < \infty$ by the Choquet integral

$$\|H\|_{L^p(\mathscr{D}(I),\nu,S)} = \left(\int_0^\infty \lambda^{p-1}\mu_\nu(S(H) > \lambda)\, d\lambda\right)^{1/p} \ .$$

Again we define similar notions with the measure ν replaced by w or w^{-1} throughout. By the Radon–Nikodym lemma of [1], which we will sketch in the appendix, we can estimate the left-hand-side of (2.2) as

$$\left(\sum_{J\in\mathscr{D}(I)} |\rho_J^w \alpha_J^{w^{-1}}(f,\chi_J^w)_\nu w(J)^{-1/2}|^2 w(J)\right)^{1/2}$$
$$\le \|\rho_J^w \alpha_J^{w^{-1}}(f,\chi_J^w)_\nu w(J)^{-1/2}\|_{L^2(\mathscr{D}(I),w,S_w^2)} \ .$$

This is a generalization of the fact that a square sum is - or at least is controlled by - an L^2 norm. By the outer Hölder inequality of [1], which we will also sketch in the appendix, the right-hand-side of the last display can be estimated by

$$\|\rho_J^w \alpha_J^{w^{-1}}\|_{L^\infty(\mathscr{D}(I),w,S_w^2)} \times \|(f,\chi_J^w)_\nu w(J)^{-1/2}\|_{L^2(\mathscr{D}(I),w,S^\infty)} \ .$$

For the reader unfamiliar with the notion of outer measure spaces, we highlight that such an outer Hölder inequality has two levels, one Hölder type inequality at the level of sizes, here with exponents $2\times\infty \to 2$, and one at the level of outer L^p norms, here with exponents $\infty \times 2 \to 2$. The term involving f is estimated by the following lemma, which essentially encodes boundedness of the w-Hardy Littlewood maximal operator in the space $L^2(w)$.

Lemma 2.1. *Consider the embedding map $h \to H$ given by $H(J) = w(J)^{-1/2}(h,\chi_J^w)_w$. Then this is a bounded operator from the standard $L^2(w)$ into the outer $L^2(\mathscr{D}(I), w, S^\infty)$.*

Proof. Fix $\lambda > 0$ and consider the maximal intervals $J \in \mathscr{D}(I)$ such that

$$H(J) > \lambda.$$

Call the family of such intervals $\mathcal{H}_\lambda$. Since $H(J)$ is the average of h on the interval J with respect to the measure w, we observe with M_w^d denoting the dyadic Hardy–Littlewood operator relative to the measure w, that for any $J \in \mathcal{H}_\lambda$ and $x \in J$

$$\lambda < H(J) \le (M_w^d h)(x) \ .$$

Hence

$$w\{\cup_{J\in\mathcal{H}_\lambda} J\} \le w\{x : M_w^d h(x) > \lambda\}.$$

If we consider the new function in $\mathscr{D}(I)$, call it K, which is equal to $H(J)$ in $\mathscr{D}(I)\setminus\cup_{J\in\mathcal{H}_\lambda}\mathscr{D}(J)$ and zero in $\cup_{J\in\mathcal{H}_\lambda}\mathscr{D}(J)$, we readily see that $\|K\|_{L^\infty(\mathscr{D}(I),w,S^\infty)} \le \lambda$. Therefore, by the definition of the outer measure space $L^1(\mathscr{D}(I), w, S^\infty)$ we have

$$\mu_w(S^\infty(H) > \lambda) \le w\{\cup_{J\in\mathcal{H}_\lambda} J\} \le w\{x : M_w^d h(x) > \lambda\}.$$

Now taking the Choquet integral on both sides gives

$$\|H\|^2_{L^2(\mathscr{D}(I),w,S^\infty)} \leq \|M^d_w h\|^2_w \leq A\|h\|^2_w.$$

□

Applying this lemma with $h = fw^{-1}$ estimates

$$\|(f,\chi^w_J)_\nu w(J)^{-1/2}\|_{L^2(X,\mu_w,S^\infty)} \leq \|f1_I\|_{L^2(w^{-1})}\ .$$

It then remains to show that

$$\|\rho^w_J \alpha^{w^{-1}}_J\|_{L^\infty(X,\mu_w,S^2_w)} \leq AQ\ .$$

This inequality means that for all $J \subset I$ we have to prove

$$\left(\sum_{K\in\mathscr{D}(J)} |\rho^w_K \alpha^{w^{-1}}_K|^2 w(K)\right)^{1/2} \leq AQ(w(J))^{1/2}\ . \tag{2.3}$$

By the Radon–Nikodym lemma as above we may estimate the left-hand-side as

$$\left(\sum_{K\in\mathscr{D}(J)} |\rho^w_K \alpha^{w^{-1}}_K \langle w\rangle^{1/2}_K|^2 \nu(K)\right)^{1/2} \leq \left\|\rho^w_K \alpha^{w^{-1}}_K \langle w\rangle^{1/2}_K\right\|_{L^2(\mathscr{D}(J),\nu,S^2_\nu)}$$

$$= \left\|\rho^w_K \alpha^{w^{-1}}_K \langle w\rangle^{1/2}_K \langle w^{-1}\rangle^{1/2}_K \frac{1}{\langle w^{-1}\rangle^{1/2}_K}\right\|_{L^2(\mathscr{D}(J),\nu,S^2_\nu)}\ .$$

Using an outer Hölder' inequality we may estimate this by

$$\left\|\rho^w_K \alpha^{w^{-1}}_K \langle w\rangle^{1/2}_K \langle w^{-1}\rangle^{1/2}_K\right\|_{L^\infty(\mathscr{D}(J),\nu,S^2_\nu)} \left\|\frac{1}{\langle w^{-1}\rangle^{1/2}_K}\right\|_{L^2(\mathscr{D}(J),\nu,S^\infty_\nu)}\ .$$

Now we have the following nonlinear embedding theorem.

Lemma 2.2. *Let h be a positive function on $\mathbb{R}$, $h, h^{-1} \in L^1(\mathbb{R}, d\nu)$. Consider the function H on $\mathscr{D}(J)$ given by the formula $H(K) = \frac{1}{\langle h^{-1}\rangle^{1/2}_K}$ for $K \in \mathscr{D}(J)$. Then $H \in L^2(\mathscr{D}(J), \nu, S^\infty)$ with norm bounded by $2\|h\|_\nu$.*

Proof. Again we fix $\lambda > 0$ and denote by $\mathcal{H}_\lambda$ the collection of maximal dyadic intervals $K \subset J$ such that $H(K) > \lambda$. For such K we define

$$E_K := \{x \in K : h^{-1}(x) \leq 2\langle h^{-1}\rangle_K\}\ .$$

Then $\nu(E_K) \geq \frac{1}{2}\nu(K)$. On the other hand, on our maximal K, we have

$$x \in E_K \Rightarrow \lambda^2 < \frac{1}{\langle h^{-1}\rangle_K} \leq 2h(x)\ .$$

This inequality implies

$$\mu_\nu(S^\infty(H) > \lambda^2) \leq \nu(\cup_{K\in\mathcal{H}_\lambda} K) \leq 2\sum_{K\in\mathcal{H}_\lambda} \nu(E_K) = 2\nu(\cup_{K\in\mathcal{H}_\lambda} E_K) \leq 2\nu\{x : h(x) \geq \lambda^2/2\}\ .$$

Taking the Choquet integral on both sides proves the lemma. □

Applying the lemma, we have reduced (2.3) to showing

$$\left\| \rho_K^w \alpha_K^{w^{-1}} \langle w\rangle_K^{1/2} \langle w^{-1}\rangle_K^{1/2} \right\|_{L^\infty(\mathscr{D}(J),\nu,S^2_\nu)} \le AQ\ .$$

This means that we need to show for any subinterval K of J

$$\sum_{L\in\mathscr{D}(K)} (\rho_L^w \alpha_L^{w^{-1}})^2 \langle w\rangle_L \langle w^{-1}\rangle_L \nu(L) \le AQ\nu(K)\ . \tag{2.4}$$

We calculate the α-s and ρ-s. For any $I\in\mathscr{D}$ we denote by I_+ the left child and by I_- the right child. We observe that with the quantities

$$h_I^\nu(x) = \begin{cases} \sqrt{\frac{\nu(I_-)}{\nu(I_+)\nu(I)}} \text{ if } x\in I_+ \\ -\sqrt{\frac{\nu(I_+)}{\nu(I_-)\nu(I)}}, \text{ if } x\in I_- \end{cases} \tag{2.5}$$

the collection h_I^ν is indeed an orthonormal set in $L^2(d\nu)$. Note further that $\chi_I^\nu = 1_I \nu(I)^{-1/2}$. By taking inner products we calculate

$$\alpha_I^w = (h_I^\nu, h_I^w)_w = \sqrt{\frac{\nu(I_-)w(I_-)}{\nu(I_+)\nu(I)w(I_+)w(I)}}\,w(I_+) + \sqrt{\frac{\nu(I_+)w(I_+)}{\nu(I_-)\nu(I)w(I_-)w(I)}}\,w(I_-)$$

$$= \frac{\nu(I_-)+\nu(I_+)}{\nu(I)} \sqrt{\frac{w(I_+)w(I_-)\nu(I)}{\nu(I_+)\nu(I_-)w(I)}} = \langle w\rangle_{I_+}^{1/2} \langle w\rangle_{I_-}^{1/2} \langle w\rangle_I^{-1/2}\ ,$$

$$\rho_I^w = (h_I^\nu, \chi_I^w)_w = \sqrt{\frac{\nu(I_-)}{\nu(I_+)\nu(I)w(I)}}\,w(I_+) - \sqrt{\frac{\nu(I_+)}{\nu(I_-)\nu(I)w(I)}}\,w(I_-)$$

$$= \left(\frac{\nu(I_+)\nu(I_-)}{\nu(I)\nu(I)}\right)^{1/2} \left(\langle w\rangle_{I_+} - \langle w\rangle_{I_-}\right) \langle w\rangle_I^{-1/2}\ .$$

Analogous identities hold for w^{-1}. With these calculations we pause to complete the proof of case I by the observation that $\langle w\rangle_I \langle w^{-1}\rangle_I \ge 1$ and thus

$$\alpha_I^w \alpha_I^{w^{-1}} \le \left(\frac{\langle w\rangle_{I_+}\langle w^{-1}\rangle_{I_+}\langle w\rangle_{I_-}\langle w^{-1}\rangle_{I_-}}{\langle w\rangle_I \langle w^{-1}\rangle_I}\right)^{1/2} \le Q\ . \tag{2.6}$$

Continuing with term III, we calculate

$$(\rho_L^w \alpha_L^{w^{-1}})^2 \langle w\rangle_L \langle w^{-1}\rangle_L = \left(\frac{\nu(I_+)\nu(I_-)}{\nu(I)\nu(I)}\right) \langle w^{-1}\rangle_{I_+} \langle w^{-1}\rangle_{I_-} \left(\langle w\rangle_{I_+} - \langle w\rangle_{I_-}\right)^2\ .$$

To proceed further, we need to find a suitable concave function. Define

$$B_1(u,v) := (uv)^{1/2}$$

$$B_2(u,v) := -(uv)^2$$

and

$$B(u,v) := CQ^{3/2} B_1(u,v) + B_2(u,v)$$

for sufficiently large absolute constant C to be determined later. The crucial point of this function B is the following uniform concavity property on the domain

$$\Omega_Q = \{(u,v) : 0 < u, 0 < v, uv \le Q\}\ .$$

Lemma 2.3. *Let α_+, α_- be positive with $\alpha_+ + \alpha_- = 1$. For any triple*

$$x = (u, v), x^+ = (u^+, v^+), x^- = (u^-, v^-)$$

of points in Ω_Q with

$$x = \alpha_+ x^+ + \alpha^- x^-$$

we have

$$\alpha_- \alpha_+ v^+ v^- (u^+ - u^-)^2 \leq C(B(u, v) - \alpha_+ B(u^+, v^+) - \alpha_- B(u^-, v^-)) \ .$$

Before we prove this lemma, we will indicate how it finishes the estimate of term III. We apply the lemma for each $L \subset K$ with $\alpha_+ = \nu(L_+)/\nu(L)$ and $\alpha_- = \nu^-(L)/\nu(L)$ and

$$(u, v) = (\langle w \rangle_L, \langle w^{-1} \rangle_L) \ ,$$
$$(u^+, v^+) = (\langle w \rangle_{L_+}, \langle w^{-1} \rangle_{L_+}) \ ,$$
$$(u^-, v^-) = (\langle w \rangle_{L_-}, \langle w^{-1} \rangle_{L_-}) \ .$$

Notice that indeed

$$\langle w \rangle_I = \alpha_+(I) \langle w \rangle_{I_+} + \alpha_+(I) \langle w \rangle_{I_-}$$

and likewise for w^{-1}. Multiplying the inequality obtained by the lemma by $\nu(L)$ yields

$$\langle w^{-1} \rangle_{L_+} \langle w^{-1} \rangle_{L_-} (\langle w \rangle_{L_+}) - \langle w \rangle_{L_-})^2 \frac{\nu(L_+)\nu(L_-)}{\nu(L)\nu(L)} \nu(L)$$
$$\leq C\big(B(u, v)\nu(L) - B(u^+, v^+)\nu(L_+) - B(u^-, v^-)\nu(L_-)\big) \ .$$

Adding over all the inequalities obtained, noting cancellation of most terms on the right hand side since all intervals other than K appear once as parent and once as child yields formally

$$\sum_{L \in \mathscr{D}(K)} \rho_L^w \alpha_L^{w^{-1}} \langle w \rangle_L^{1/2} \langle w^{-1} \rangle_L^{1/2} \nu(L) \leq B(\langle w \rangle_K \langle w^{-1} \rangle_K) \nu(K) \leq AQ^2 \nu(K) \ .$$

This formal sum is infinite, however rigorous justification is easily done by truncating the sum at some small scale $|L|$ and using positivitiy of the function B to obtain the inequality for the truncated sum. Since the sum has only positive summands, the limit inequality is readily seen. This shows (2.3) and thus completes the proof of case III, save for the proof of the above lemma which we now present.

Proof. We note that $B_1(u, v)$ is concave in the quadrant $u > 0, v > 0$ in the sense that for any triple as in the lemma

$$B_1(u, v) - \alpha_- B_1(u^-, v^-) - \alpha_+ B_1(u^+, v^+) \geq 0 \ .$$

To see this, it suffices to observe that the Hessian of B_1 is negative semi–definite in this quadrant, namely

$$d^2(\sqrt{uv}) = \begin{pmatrix} -\frac{1}{4}\sqrt{\frac{v}{u^3}} & \frac{1}{4\sqrt{uv}} \\ \frac{1}{4\sqrt{uv}} & -\frac{1}{4}\sqrt{\frac{u}{v^3}} \end{pmatrix} ,$$

and this matrix has determinant zero and is therefore is negative semi–definite since the entries on the diagonal are negative in the given quadrant. We also note that $B^*(u, v) = \frac{1}{2} C Q^{3/2} B_1(u, v) + B_2(u, v)$ is concave in Ω_Q. To see this, let $\Delta_u := u^+ - u^-, \Delta_v = v^+ - v^-$ so that

$$u^+ = u + \alpha_- \Delta_u, \ u^- = u - \alpha_+ \Delta_u; \ v^+ = v + \alpha_- \Delta_v, \ v^- = v - \alpha_+ \Delta_v \ .$$

We calculate, using convexity of $x \to x^2$

$$\begin{aligned}
B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+)& \\
\geq -(uv)^2 + (\alpha_- u^- v^- + \alpha_+ u^+ v^+)^2& \\
= -(uv)^2 + (uv + \alpha_-\alpha_+\Delta_u\Delta_v)^2& \\
= 2\alpha_-\alpha_+ uv\Delta_u\Delta_v + \alpha_-\alpha_+(\Delta_u\Delta_v)^2& \\
\geq 2\alpha_-\alpha_+ uv\Delta_u\Delta_v\ .&
\end{aligned} \tag{2.7}$$

If $\Delta_u\Delta_v \geq 0$ then the concavity property of B^* follows from this calculation and concavity of B_1. Consider $\Delta_u\Delta_v < 0$, then we need to more carefully invoke the term B_1. Using concavity of $x \to \sqrt{x}$ we calculate

$$\begin{aligned}
B_1(u,v) - \alpha_- B_1(u^-,v^-) - \alpha_+ B_1(u^+,v^+)& \\
\geq \sqrt{uv} - \sqrt{\alpha_- u^- v^- + \alpha_+ u^+ v^+}& \\
= \sqrt{uv} - \sqrt{uv + \alpha_-\alpha_+\Delta_u\Delta_v}& \\
\geq -\frac{1}{2\sqrt{uv}}\alpha_-\alpha_+\Delta_u\Delta_v\ .&
\end{aligned}$$

This together with $uv \leq Q$ shows that for sufficiently large C we have concavity of B^* in case $\Delta_u\Delta_v < 0$. Collecting cases, we have even more shown that for sufficiently large C and arbitrary sign of $\Delta_u\Delta_v$:

$$B(u,v) - \alpha_- B(u^-,v^-) - \alpha_+ B(u^+,v^+) \geq cuv\alpha_-\alpha_+|\Delta_u\Delta_v|\ . \tag{2.8}$$

We pause our main line of reasoning to observe for later purpose that the same arguments work for a modified function $\tilde{B}$ with linear rather than quadratic term in uv. Namely, set $\tilde{B}_2(u,v) = -uv$ and

$$\tilde{B}(u,v) = CQ^{1/2}B_1(u,v) + \tilde{B}_2(u,v)\ .$$

Then we have

$$\tilde{B}(u,v) - \alpha_-\tilde{B}(u^-,v^-) - \alpha_+\tilde{B}(u^+,v^+) \geq cuv\alpha_-\alpha_+|\Delta_u\Delta_v|\ , \tag{2.9}$$

which follows the same way as above by expanding in lieu of (2.7)

$$\tilde{B}_2(u,v) - \alpha_-\tilde{B}_2(u^-,v^-) - \alpha_+\tilde{B}_2(u^+,v^+) = uv\alpha_-\alpha_+\Delta_u\Delta_v\ .$$

We continue the main line of reasoning for the proof of the lemma and claim :

Lemma 2.4. *For triples in the domain Ω_Q we have that*

$$B(u,v) - \alpha_- B(u^-,v^-) - \alpha_+ B(u^+,v^+) \geq c\alpha_-\alpha_+ v^+ v^- \Delta_u^2$$

$$B(u,v) - \alpha_- B(u^-,v^-) - \alpha_+ B(u^+,v^+) \geq c\alpha_-\alpha_+ u^+ u^- \Delta_v^2 \tag{2.10}$$

for sufficiently small absolute c.

We prove Lemma (2.4). By symmetry, it suffices to prove (2.10). By invariance of the inequality under the scaling $u \to \lambda u$ and $v \to \lambda^{-1}v$ we may assume throughout that $u = v$. We split into several cases.

Case 1: We consider $|\Delta_u| > Nu$ for some large constant N.

Case 1a: We consider in addition $|\Delta_v| > Nv$. Then expanding B_2 we obtain

$$\begin{aligned}
B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+)& \\
= -(uv)^2 + \alpha_-(u^-v^-)^2 + \alpha_+(u^+v^+)^2 =&
\end{aligned}$$

$$-(uv)^2 + \alpha_-(uv - \alpha_+ u\Delta_v - \alpha_+\Delta_u v + \alpha_+^2\Delta_u\Delta_v)^2 + \alpha_+(uv + \alpha_- u\Delta_v + \alpha_-\Delta_u v + \alpha_-^2\Delta_u\Delta_v)^2$$

$$= \alpha_+\alpha_-(\alpha_+^3 + \alpha_-^3)(\Delta_u\Delta_v)^2(1 + O(N^{-1}))$$
$$\geq \frac{1}{8}\alpha_+\alpha_-(\Delta_u\Delta_v)^2$$

for sufficiently large N. But we have under the given case assumptions $u^+ \leq 2|\Delta_u|$ and $u^- \leq 2|\Delta_u|$ and hence

$$\frac{1}{8}\alpha_+\alpha_-(\Delta_u\Delta_v)^2 \geq c\alpha_-\alpha_+u^+u^-\Delta_v^2 \ .$$

Together with concavity of B_1 this completes the proof in Case 1a.

Case 1b: We now consider $|\Delta v| \leq Nv$, still in addition to $|\Delta u| > Nu$. Then we simply have

$$u^-u^+ \leq \min(u^-, u^+)(u + |\Delta_u|)| \leq 2u|\Delta u|$$

and hence

$$c\alpha_-\alpha_+u^-u^+\Delta_v^2 \leq 2cN\alpha_-\alpha_+uv|\Delta_u\Delta_v| \ .$$

The desired inequality (2.10) now follows from (2.8) for small c depending on N. This completes the Case 1b.

Case 2: $|\Delta_u| \leq Nu$. We consider first

Case 2a: $|\Delta_v| > N^2v$. Then we have with $u^+v^+ \leq Q$ and $u^-v^- \leq Q$ and $\max(u^-, u^+) \leq 2Nu$

$$u^-u^+\max(v^+, v^-) \leq \max(u^-, u^+)Q \leq 2NuQ \ ,$$

and hence with $|\Delta_v| \leq \max(v^+, v^-)$

$$c\alpha_-\alpha_+u^-u^+\Delta_v^2 \leq 2cNQ\alpha_-\alpha_+u|\Delta_v| \ .$$

By concavity of B^* it suffices to show for some appropriate c

$$cN\alpha_-\alpha_+u|\Delta_v| \leq Q^{1/2}(B_1(u, v) - \alpha_-B_1(u^-, v^-) - \alpha_+B_1(u^+, v^+)) \ .$$

Using $Q > uv = u^2$ it suffices to show

$$cN\alpha_-\alpha_+|\Delta_v| \leq B_1(u, v) - \alpha_-B_1(u^-, v^-) - \alpha_+B_1(u^+, v^+) \ .$$

Assume without loss of generality $v^- < v < v^+$. So $\Delta_v > 0$. Then we have

$$B_1(u, v) - \alpha_-B_1(u^-, v^-) - \alpha_+B_1(u^+, v^+)$$
$$= (B_1(u, v) - B_1(u^-, v^-)) + \alpha_+(B_1(u^-, v^-) - B_1(u^+, v^+)) \ .$$

We will establish that the first term dominates the second. To estimate the second term from above we note

$$\alpha_+(|B_1(u^-, v^-)| + |B_1(u^+, v^+)|) \leq 2\alpha_+\sqrt{(u + |\Delta_u|)(v + |\Delta_v|)}$$
$$\leq 2\alpha_+\sqrt{(v + Nv)2|\Delta_v|} \leq 2\alpha_+\sqrt{4(|\Delta_v|/N)|\Delta_v|}$$
$$\leq 4\alpha_+N^{-1/2}\Delta_v \ .$$

Now we estimate the first term. By the mean value theorem we have for m a certain value between uv and u^-v^-.

$$B_1(u, v) - B_1(u^-, v^-) = \sqrt{uv} - \sqrt{u^-v^-}$$
$$= \frac{uv - u^-v^-}{2\sqrt{m}} = \frac{u(v - v^-) - (u^- - u)v^-}{2\sqrt{m}} \ .$$

The first summand in the numerator dominates the second. Namely, since the vector $x_+ - x_-$ has the same direction and $x - x_-$, we have $|u - u^-| < N^{-1}|v - v^-|$ and

$$|u - u^-|v^- < |u - u^-|v = |u - u^-|u \leq N^{-1}u|v - v^-| \ .$$

Hence

$$|B_1(u,v) - B_1(u^-,v^-)| \geq$$
$$\frac{u|v-v^-|}{4\sqrt{m}} = \frac{u\alpha_+\Delta_v}{4\sqrt{m}}$$

Since the direction $x - x_-$ is facing nearly north, we have $u^-v^- < uv$ and hence $m < uv$ and hence with $u = v$

$$|B_1(u,v) - B_1(u^-,v^-)| \geq (1/4)\alpha_+\Delta_v \ .$$

This is the desired lower bound on the first term. We may therefore estimate for large N

$$B_1(u,v) - \alpha_- B_1(u^-,v^-) - \alpha_+ B_1(u^+,v^+) \geq c\alpha_+\Delta_v \ ,$$

where we have used that the left-hand-side is positive. This finishes Case 2a with appropriate choice of c depending on N.

Case 2b: $|\Delta_v| \leq N^2 v$, still in addition to $|\Delta_u| \leq Nu$. We make a further case distinction into four subcases i, ii , iii, iv, depending on the angle of the vector (Δ_u, Δ_v), which might be either almost vertical or almost horizontal or in Northwest/Southeast or Northeast/Southwest direction. Assume first

Case 2bi: $\Delta_u\Delta_v > 0$ and the vector Δ_u, Δ_v has angle at least N^{-5} from each the horizontal and vertical direction. We shall prove

$$B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+) \geq c\alpha_-\alpha_+ u^+u^-\Delta_v^2 \ ,$$

which will clearly finish this case. Note that the second derivative of B_2 at (x,y) in the first quadrant in direction of a unit vector (a,b) with $ab > 0$ is given by

$$2a^2y^2 + 8abxy + 2b^2x^2 \geq 2a^2y^2 + 2b^2x^2 \ . \tag{2.11}$$

Note that

$$B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+)$$

is dominated below by a constant times the integral over the line segment from midpoint of (u^-,v^-) and (u,v) to midpoint of (u,v) and (u^+,v^+) of the second partial of B in direction of the line segment. Namely:

$$\begin{aligned}
&B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+) \\
&= \alpha_-[B_2(u,v) - B_2(u^-,v^-)] - \alpha_+[B_2(u^+,v^+) - B_2(u,v)] \\
&= \alpha_- \int_{(u^-,v^-)}^{(u,v)} \partial_{(\alpha,\beta)} B_2 ds - \alpha_+ \int_{(u,v)}^{(u^+,v^+)} \partial_{(\alpha,\beta)} B_2 ds \\
&= \alpha_-\alpha_+\sqrt{\Delta_u^2 + \Delta_v^2} \int_0^1 \partial_{(\alpha,\beta)} B_2(u^- + t(u-u^-), v^- + t(v-v^-))dt \\
&\quad -\alpha_+\alpha_-\sqrt{\Delta_u^2 + \Delta_v^2} \int_0^1 \partial_{(\alpha,\beta)} B_2(u + t(u^+-u), v + t(v^+-v))dt \\
&= \alpha_+\alpha_-\sqrt{\Delta_u^2 + \Delta_v^2} \int_0^1 \int_{(u^-+t(u-u^-),v^-+t(v-v^-))}^{(u+t(u^+-u),v+t(v^+-v))} \partial^2_{(\alpha,\beta)} B_2 dsdt
\end{aligned}$$

Now assume the second partial is positive in the domain of question and the partial derivative is bounded from below by m on the line segment from $(u^- + (u-u^-)/2, v^- + (u-u^-)/2)$ to $(u + (u^+-u)/2, v + t(v^+-v)/2)$. Then the double integral can be estimated below using Fubini by

$$\geq c\alpha_+\alpha_-\sqrt{\Delta_u^2 + \Delta_v^2}\left[\int_{(u^-+(u-u^-)/2,v^-+(v-v^-)/2)}^{(u,v)} \int_0^{1/2} mdtds\right.$$

$$+\int_{(u,v)}^{(u+(u^+-u)/2,v+t(v^+-v)/2)} \int_{1/2}^{1} m dt ds\Big]$$
$$\geq cm\alpha_+\alpha_-(\Delta_u^2+\Delta_v^2)\ .$$

On this line segment, the second derivative (2.11) is bounded below by cuv. Hence

$$B_2(u,v)-\alpha_-B_2(u^-,v^-)-\alpha_+B_2(u^+,v^+)\geq c\alpha_+\alpha_-uv(\Delta_u^2+\Delta_v^2)\ .$$

This proves the desired inequality in this case 2bi.

Now consider case $2bii$ which is the direction of $\Delta_u\Delta_v<0$ and again the angle between the vector (Δ_u,Δ_v) to x-axis is at least N^{-5}. We proceed similarly as before, but now with the function B_1. Note that the second derivative of B_1 in direction $(a,b), a>0, b<0$, is

$$c(\frac{v^{1/4}}{u^{3/4}}a-\frac{u^{1/4}}{v^{3/4}}b)^2$$

and in particular is positive. Moreover, since the mixed term in the expansion of the square has favourable sign, the last display is bounded from below as

$$c(\frac{v^{1/4}}{u^{3/4}}a-\frac{u^{1/4}}{v^{3/4}}b)^2\geq c(\frac{v^{1/2}}{u^{3/2}}a^2+\frac{u^{1/2}}{v^{3/2}}b^2)\ .$$

Using the arithmetic/geometric mean inequality we bound the latter from below by

$$c\frac{1}{\sqrt{uv}}|ab|\geq\frac{c}{Q^{3/2}}uv|ab|\geq\frac{c'}{Q^{3/2}}uv(a^2+b^2)$$

on the line segment in question, where in the last inequality we have used that the entries of (a,b) are bounded below depending on N and that the coordinates of points of our line segment are comparable with u and v correspondingly (with constants depending on N only). This completes the proof in this case in analogy to the previous case.

Now we consider case $2biii$ where the vector (Δ_u,Δ_v) has small angle (less than N^{-5}) to the y-axis, that is $\Delta_u\leq N^{-5}\Delta_v$. We shall prove in this case

$$B_2(u,v)-\alpha_-B_2(u^-,v^-)-\alpha_+B_2(u^+,v^+)\geq c\alpha_+\alpha_-u^-u^+\Delta_v^2\ ,$$

which suffices by concavity of B_1. We compare with a vertical line segment and thus write

$$B_2(u,v)-\alpha_-B_2(u^-,v^-)-\alpha_+B_2(u^+,v^+)$$
$$=B_2(u,v)-\alpha_-B_2(u,v^-)-\alpha_+B_2(u,v^+)$$
$$+\alpha_-(B_2(u,v^-)-B_2(u^-,v^-))+\alpha_+(B_2(u,v^+)-B_2(u^+,v^+))$$
$$=\alpha_-\alpha_+u^2\Delta_v^2+\alpha_-(B_2(u,v^-)-B_2(u^-,v^-))+\alpha_+(B_2(u,v^+)-B_2(u^+,v^+))\ ,$$

where we have simply expanded B_2 to evaluate the terms on the vertical segment. The remaining terms we consider as error terms and observe

$$\alpha_-(B_2(u,v^-)-B_2(u^-,v^-))+\alpha_+(B_2(u,v^+)-B_2(u^+,v^+))$$
$$=\alpha_-(u-u^-)(u+u^-)(v^-)^2+\alpha_+(u-u^+)(u+u^+)(v^+)^2$$
$$=\alpha_-\alpha_+\Delta_u[(2u-\alpha_+\Delta_u)(v-\alpha_+\Delta_v)^2-(2u+\alpha_-\Delta_u)(v+\alpha_-\Delta_v)^2]\ .$$

Expanding we notice that the terms independent of the vector (Δ_u,Δ_v) inside the bracket cancel. Thus we obtain , using $|\Delta_u|\leq|\Delta_v|$ and $|\Delta_v|\leq N^2u$ that the terms in the bracket are bounded by $100N^4|\Delta_v|u^2$. Hence we can estimate

$$|\alpha_-(B(u,v^-)-B(u^-,v^-))+\alpha_+(B(u,v^+)-B(u^+,v_)))|$$
$$\leq 100\alpha_-\alpha_+N^4u^2|\Delta_u\Delta_v|$$

$$\leq 100\alpha_-\alpha_+N^{-1}u^2\Delta_v^2 \ .$$

For large enough N this is dominated by the above main term and thus

$$B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+) \geq c\alpha_-\alpha_+u^2\Delta_v^2 \ .$$

This completes the proof in Case $2biii$.

Now we consider case $2biv$ where the vector (Δ_u, Δ_v) has small angle (less than N^{-5}) to x-axis. We proceed in symmetric fashion (B_2 is symmetric under the exchange of u and v) to the previous case to obtain

$$B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+) \geq c\alpha_-\alpha_+v^2\Delta_u^2 \ .$$

Using $u = v$ and $\Delta_v \leq \Delta_u$ gives

$$B_2(u,v) - \alpha_- B_2(u^-,v^-) - \alpha_+ B_2(u^+,v^+) \geq c\alpha_-\alpha_+u^2\Delta_v^2 \ .$$

This completes the proof in this case. □

3. Boundedness of term IV

This section is devoted to proving the desired bound for term IV, that is

$$\sum_{J\in\mathscr{D}(I)} |\rho_J^w \rho_J^{w^{-1}}(f,\chi_J^w)_\nu (g,\chi_J^{w^{-1}})_\nu| \leq AQ\|f1_I\|_{w^{-1}}\|g1_I\|_w \ ,$$

where Q is as in (2.1). We use outer measure spaces as in the previous section. Applying the Radon-Nikodym property estimates the left-hand-side by

$$\|\rho_J^w \rho_J^{w^{-1}}(f,\chi_J^w)_\nu (g,\chi_J^{w^{-1}})_\nu \nu(J)^{-1}\|_{L^1(\mathscr{D}(I),\nu,S^1_\nu)} \ .$$

Applying the outer Hölder inequality estimates we get the bound

$$\|\rho_J^w \rho_J^{w^{-1}}\langle w\rangle_J^{1/2}\langle w^{-1}\rangle_J^{1/2}\|_{L^\infty(\mathscr{D}(I),\nu,S^1_\nu)}$$

$$\times\|(f,\chi_J^w)_\nu (g,\chi_I^{w^{-1}})_\nu \langle w\rangle_J^{-1/2}\langle w^{-1}\rangle_J^{-1/2}\nu(J)^{-1}\|_{L^1(\mathscr{D}(I),\nu,S^\infty)} \ .$$

To estimate the first factor by Q we identify the required estimate as

$$\sum_{K\subset J} \rho_K^w \rho_K^{w^{-1}}\langle w\rangle_K^{1/2}\langle w^{-1}\rangle_K^{1/2}\nu(K) \leq CQ\nu(J)$$

for every $J \subset I$. Applying the formulae for ρ-s we obtain

$$\rho_L^w \rho_L^{w^{-1}}\langle w\rangle_L^{1/2}\langle w^{-1}\rangle_L^{1/2} = \left(\frac{\nu(I_+)\nu(I_-)}{\nu(I)\nu(I)}\right)^{1/2} \big(\langle w\rangle_{I_+} - \langle w\rangle_{I_-}\big)\big(\langle w^{-1}\rangle_{I_+} - \langle w^{-1}\rangle_{I_-}\big) \ .$$

The desired estimate is then proved analogously to the estimate (2.3) using the function $\tilde{B}$ and the remarks near inequality (2.9). It remains to prove

$$\|(f,\chi_J^w)_\nu (g,\chi_J^{w^{-1}})_\nu \langle w\rangle_J^{-1/2}\langle w^{-1}\rangle_J^{-1/2}\nu(J)^{-1}\|_{L^1(\mathscr{D}(I),\nu,S^\infty)} \leq \|f1_I\|_{w^{-1}d\nu}\|g1_I\|_{wd\nu} \ .$$

With the usual identification $(f,\chi_J^w)_\nu = (fw^{-1},\chi_J^w)_w$ and similarly for g, this will follow from the following bilinear embedding theorem.

Theorem 3.1. *Denote*

$$F_w(J) = w(J)^{-1}\int_J fwd\nu,\ G_{w^{-1}}(J) = w^{-1}(J)^{-1}\int_J gw^{-1}d\nu \ .$$

The operator $f\times g \to F_wG_{w^{-1}}$ is bounded from $L^2(w)\times L^2(w^{-1})$ into $L^1(\mathscr{D}(I),\nu,S^\infty)$.

It suffices to prove this for positive functions f and g. Let us consider the collection $\mathcal{J}_\lambda$ of all maximal dyadic intervals J such that

$$F_w(J)G_{w^{-1}}(J) > \lambda \ .$$

This implies that any such J and any $x, y \in J$:

$$(M^d_w f)(x)(M^d_{w^{-1}} g)(y) \geq F_w(J)G_{w^{-1}}(J) > \lambda \ .$$

Hence, if $J \in \mathcal{J}_\lambda$

$$\inf_J \left(M^d_w f \cdot M^d_{w^{-1}} g\right) \geq \inf_J M^d_w f \cdot \inf_J M^d_{w^{-1}} g > \lambda \ .$$

Denote $\Phi := M^d_w f \cdot M^d_{w^{-1}} g$, then we have seen

$$\mu_\nu(S^\infty(F_w G_{w^{-1}}) > \lambda) \leq \sum_{J \in \mathcal{J}_\lambda} \nu(J) = \nu(\cup_{J \in \mathcal{J}_\lambda} J) \leq \nu\{x \in \mathbb{R} : \Phi(x) > \lambda\} \ .$$

Applying the Choquet integral on both sides, we are reduced to estimating the classical $L^1(\mathbb{R}, d\nu)$ norm of Φ. But this follows from the general Hardy–Littlewood maximal theorem with respect to the weights w and w^{-1}:

$$\int \Phi d\nu = \int (M^d_w f \cdot M^d_{w^{-1}} g) d\nu = \int (M^d_w f) w^{1/2} (M^d_{w^{-1}} g) w^{-1/2} d\nu$$

$$\leq \left(\int (M^d_w f)^2 w d\nu\right)^{1/2} \left(\int (M^d_{w^{-1}} g)^2 w^{-1} d\nu\right)^{1/2} \leq A\|f\|_w \|g\|_{w^{-1}}.$$

This completes the proof of Theorem 3.1.

4. Appendix on outer measures

We recall two inequalities for outer L^p spaces from [1]. We first prove an exemplary form of an inequality which is referred to several times as Radon–Nikodym property in this paper,

$$\left(\sum_{J \in \mathscr{D}(I)} |F(J)|^2 \nu(J)\right)^{1/2} \leq C\|F\|_{L^2(\mathscr{D}(I), \nu, S^2_\nu)} \ . \tag{4.1}$$

It is used as well with an exponent 1 in place of the exponent 2 everywhere, and it is used with weights w or w^{-1} in place of ν. All these generalizations are straightforward. The name Radon–Nikodym stems from [1], where this inequality in case of exponents 1 is based on the interpretation that a standard measure entering on the left-hand-side has bounded Radon–Nikodym derivative relative to an outer measure on the right-hand-side. Fix $2^k = \lambda$ and consider a set $\mathcal{K}$ with the property that for every $J \in \mathscr{D}(I)$ we have

$$S^2_\nu(F 1_{\mathscr{D}(I) \setminus \mathcal{K}}, \mathscr{D}(J)) \leq \lambda$$

and we have

$$\mu_\nu(\mathcal{K}) \leq 2\mu_\nu(S^2_\nu(F) > \lambda) \ .$$

That is, $\mathcal{K}$ is almost an extremizer for the infimum in the definition of the super level measure. Now pick a collection $\mathcal{J}$ of dyadic intervals such that

$$\mathcal{K} \subset \bigcup_{J \in \mathcal{J}} \mathscr{D}(J)$$

and

$$\sum_{J \in \mathcal{J}} \nu(J) \leq 2\mu_\nu(\mathcal{K}) \ .$$

That is, $\mathcal{J}$ is almost an extremizer of the infimum in the definition of the outer measure. Now write $\mathcal{K}_k := \mathcal{K}$ and $\mathcal{J}_k := \mathcal{J}$ and let k vary. Then we have

$$\sum_{H\in\mathscr{D}(I)} |F(H)|^2\nu(H) \le \sum_k \left(\sum_{H\in\mathcal{K}_k\setminus\bigcup_{l>k}\mathcal{K}_l} |F(H)|^2\nu(H)\right)$$

$$\le \sum_k \left(\sum_{J\in\mathcal{J}_k}\left(\sum_{H\in\mathscr{D}(J)\setminus\bigcup_{l>k}\mathcal{K}_l} |F(H)|^2\nu(H)\right)\right)$$

$$\le \sum_k \left(\sum_{J\in\mathcal{J}_k} S^2_\nu(F1_{\mathscr{D}(I)\setminus\bigcup_{l>k}\mathcal{K}_l}, \mathscr{D}(J))^2\nu(J)\right)$$

$$\le \sum_k C2^{2k}\left(\sum_{J\in\mathcal{J}_k}\nu(J)\right) \le \sum_k C2^{2k}\mu_\nu(S^2_\nu(F) > 2^k) \le C\|F\|^2_{L^2(\mathscr{D}(I),\nu,S^2_\nu)} \ .$$

In the last line we have used a standard argument to compare the Choquet integral with a sum. This completes the proof of (4.1). We next prove the frequently used outer Hölder's inequality in the examplary form

$$\|FG\|_{L^2(\mathscr{D}(I),\nu,S^2_\nu)} \le C\|F\|_{L^2(\mathscr{D}(I),\nu,S^\infty_\nu)} \times \|G\|_{L^\infty(\mathscr{D}(I),\nu,S^2_\nu)} \ . \tag{4.2}$$

Again this inequality is also used with measures w and w^{-1} in place of ν, and with the exponent 2 replaced by 1, generalizations that are straightforward. Consider a set $\mathcal{L}_\lambda$ such that

$$\sup_{J\in\mathscr{D}(I)\setminus\mathcal{L}_\lambda} S^\infty_\nu(F,\mathscr{D}(J)) \le \lambda$$

and

$$\mu_\nu(\mathcal{L}_\lambda) \le (1+\epsilon)\mu_\nu(S^\infty_\nu(F) > \lambda) \ ,$$

which means that $\mathcal{L}_\lambda$ is an almost extremizer of the infimum in the definition of the super level measure. For $J \in \mathscr{D}(I) \setminus \mathcal{L}_\lambda$ we have from the definitions of the sizes

$$S^2_\nu(FG,\mathscr{D}(J)) \le S^\infty_\nu(F,\mathscr{D}(J)) \times S^2_\nu(G,\mathscr{D}(J) \ .$$

Note that this is a classical Hölder inequality. We obtain

$$S^2_\nu(FG,\mathscr{D}(J)) \le S^\infty_\nu(F,\mathscr{D}(J))\|G\|_{L^\infty(\mathscr{D}(I),\nu,S^2_\nu)}$$

by the definition of the outer L^∞ norm. Setting $\tilde\lambda = \lambda\|G\|_{L^\infty(\mathscr{D}(I),\nu,S^2_\nu)}$ we obtain

$$\mu(S^2_\nu(FG) > \tilde\lambda) \le (1+\epsilon)\mu_\nu(S^\infty_\nu(F) > \lambda) \ .$$

Since ϵ was arbitrary,

$$\int_0^\infty \tilde\lambda\mu(S^2_\nu(FG) > \tilde\lambda)d\tilde\lambda$$

$$\le C\int_0^\infty \tilde\lambda\mu(S^\infty_\nu(F) > \lambda)d\tilde\lambda$$

$$\le C\|G\|^2_{L^\infty(\mathscr{D}(I),\nu,S^2_\nu)}\int_0^\infty \lambda\mu(S^\infty_\nu(F) > \lambda)d\lambda \ .$$

This proves (4.2) by the definition of the outer L^2 norm.

References

[1] Y. Do, C. Thiele, L^p theory for outer measures and two themes of Lennart Carleson united, to appear in Bulletin AMS, arXiv:1309.0945.

[2] T. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators. Ann. of Math. (2) 175 (2012), no. 3, 1473–1506.

[3] T. Hytönen, C. Pérez, S. Treil, A. Volberg, Sharp weighted estimates for dyadic shifts and the A_2 conjecture. J. Reine Angew. Math. 687 (2014), 43–86.

[4] A. Lerner, A simple proof of A_2 conjecture, Intern. Math. Res. Notices **IMRN**, 2013, no. 14, 3159–3170.

[5] L. D. lópez-Sánchez, J.-M. Martell, J. Parcet, Dyadic harmonic analysis beyond doubling measures, arXiv:1211.6291, to appear in Advances in Math.

[6] F. Nazarov, S. Treil, A. Volberg, The Bellman functions and two-weight inequalities for Haar multipliers, J. Amer. Math. Soc. 12 (1999), no. 4, 909–928.

[7] F. Nazarov, A. Volberg, A simple sharp weighted estimate of the dyadic shifts on metric space with geometric doubling, arxiv:1104.4893, Int. Math. Res. Not. **IMRN** 2013, no. 16, 3771–3789.

[8] S. Petermichl, A. Volberg, Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. Duke Math. J. 112 (2002), no. 2, 281–305.

[9] S. Petermichl, The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. Amer. J. Math. 129 (2007), no. 5, 1355–1375.

[10] S. Treil, A. Volberg, Entropy conditions in two weight inequalities for singular integral operators, arxiv:1405.0385.

[11] V. Vasyunin, A. Volberg, *The Bellman functions for the simlpest two-weight inequality: an investigation of a particular case*, St. Petersburg Math. J. 18 (2007), no. 2, 201–222.

[12] J. Wittwer, A sharp estimate on the norm of martingale transform, Math. Res. Letters, v. 7 (2000), pp. 1–12.

Mathematisches Institut, Universität Bonn
Endenicher Allee 60, D - 53115 Bonn, Germany

and

Department of Mathematics, UCLA, Los Angeles, CA 90095-1555
E-mail address: thiele@math.uni-bonn.de (C. Thiele)

Department of Mathematics, Brown University, 151 Thayer St., Box 1917, Providence, RI 02912, USA
E-mail address: treil@math.brown.edu (S. Treil)

Department of Mathematics, Michigan Sate University, East Lansing, MI. 48823
E-mail address: volberg@math.msu.edu (A. Volberg)