Irina Holmes – Research Statement

October 2017

My research interests can be generally described to live in the realms of analysis and probability. My doctoral research was in probability, with a strong analysis flavor. Specifically, my main focus was developing a Radon transform for infinite-dimensional Banach spaces. This was as much a probability problem as it was a functional analysis problem. The merger between the two in infinite dimensions often occurs due to the absence of Lebesgue measure in this setting. Probability offers several options in this case, but the most useful one seems to be abstract Wiener spaces, first developed by Leonard Gross. The main idea is that one works with a very special Gaussian measure μ on an infinite-dimensional Banach space B; what makes this measure so special is that it comes from a Hilbert space H embedded in the larger Banach space – so an abstract Wiener space is really a triple (H, B, μ). Some of the geometrical properties of the underlying inner product norm transfer to the Gaussian measure in some sense, resulting in a beautiful interplay between probabilistic approaches and functional analysis approaches.

As a postdoc at Georgia Tech, I joined the harmonic analysis group and set off on an extensive project involving weighted inequalities for commutators with Calderón-Zygmund operators. These operators are of fundamental importance in harmonic analysis; the prototypical such operator, the Hilbert transform on $\mathbb{R}$, enjoys many connections with complex analysis which unfortunately do not transfer to higher-dimensional versions such as the Riesz transforms. The introduction of Calderón-Zygmund operators brought real analysis techniques into the arsenal of tackling such operators, thus opening them up to many other areas of mathematics, such as PDE theory and operator theory. My work in this area concerns commutators $[b, T]f := b \cdot Tf - T(b \cdot f)$, where b is a function on $\mathbb{R}^n$ and T is a Calderón-Zygmund operator on $\mathbb{R}^n$. Such commutators have many applications throughout various areas of analysis, such as complex analysis (where they capture the Hankel operators), PDE theory (the Div-Curl Lemma), or factorizations of spaces of functions.

Currently as a postdoc at Michigan State University, I am working on sharp inequalities for several types of operators using the Bellman function technique. Originating from stochastic optimal control, this technique is now widely used in harmonic analysis and is a stunning example of probabilistic ideas being applied in analysis.

One of my big goals for the future is to bring together the two directions I have so far studied independently, namely harmonic analysis and infinite-dimensional analysis via abstract Wiener spaces. From the very beginning, abstract Wiener spaces were developed for the purpose of carrying out harmonic analysis in infinite dimensions. Leonard Gross came up with abstract Wiener spaces while trying to make sense of the $\mathbb{R}^n$ Poisson equation $\frac{1}{2}\Delta u = -g$ in infinite dimensions. At this point I am armed with a solid background in both, and would like to pursue a research program that studies harmonic analysis ideas like BMO and Hardy spaces, maximal and square functions, Calderón-Zygmund operators, commutators, and weighted inequalities, in infinite-dimensional Banach spaces.

Contents

1 Weighted Inequalities in Harmonic Analysis

1.1 Two-Weight Inequalities for Commutators

If $\mathbf{X}$ and $\mathbf{P}$ are position and momentum operators, norm estimates for the commutator $[\mathbf{X}, \mathbf{P}] := \mathbf{XP} - \mathbf{PX}$ address the degree to which these commute, and yield the Heisenberg Uncertainty Principle. The Hilbert transform

$$Hf(x) := \text{p.v. } \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(y)}{x-y}\, dy,$$

the prototypical singular integral operator, translates this to the language of Fourier analysis: H is a multiplier operator with symbol $-i\text{sgn}(\xi)$, where ξ is momentum. If x denotes position, the commutator of interest becomes $[M_x, H]$, where M_x is the operator of multiplication by x. More generally, a Calderón-Zygmund operator (CZO) is an integral operator T,

$$Tf(x) := \int_{\mathbb{R}^n} K(x,y) f(y)\, dy, \quad x \notin \text{supp} f,$$

associated to a kernel $K(x, y)$ satisfying some standard size and smoothness estimates. The prototypes for this important class of operators are the Hilbert transform, in the one-dimensional case, and the Riesz transforms:

$$R_j(f)(x) := \text{p.v.} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \int_{\mathbb{R}^n} f(y) \frac{x_j - y_j}{|x-y|^{n+1}}\, dy, \quad j = 1, \ldots, n,$$

in the multidimensional case.

One is then concerned with commutators $[b, T] := [M_b, T]$, where b is a 'symbol' function. These commutators have deep connections to operator theory, complex function theory, and PDEs. A large part of my recent work focused on a two-weight extension of the foundational paper [12], where Coifman, Rochberg and Weiss characterize the boundedness of $[b, T]$ on $L^p(\mathbb{R}^n)$ in terms of the $BMO(\mathbb{R}^n)$ norm of b. Recall that, in harmonic analysis, the real Hardy space $\mathcal{H}^1(\mathbb{R}^n)$ and the space $BMO(\mathbb{R}^n)$ of functions with bounded mean oscillation are the appropriate replacements of

$L^1(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$, respectively. The duality $\mathcal{H}^1(\mathbb{R}^n)^* = BMO(\mathbb{R}^n)$ [19] is a cornerstone of the field. The $BMO(\mathbb{R}^n)$ norm is defined by

$$\|b\|_{BMO(\mathbb{R}^n)} := \sup_Q \frac{1}{|Q|} \int_Q |b(x) - \langle b\rangle_Q|\, dx,$$

where the supremum is over all cubes Q in $\mathbb{R}^n$, $|Q|$ denotes the Lebesgue measure of Q, and $\langle b\rangle_Q := \frac{1}{|Q|}\int_Q b(x)\,dx$ denotes the average of b over Q.

The result of [12] can be stated in two parts, as is typical in these situations: an "upper bound" result valid for all CZOs:

$$\|[b,T] : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)\| \lesssim \|b\|_{BMO(\mathbb{R}^n)},$$

and a "lower bound" result specialized to the Riesz transforms:

$$\|b\|_{BMO(\mathbb{R}^n)} \lesssim \sum_{j=1}^n \|[b, R_j] : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)\|,$$

where $1 < p < \infty$ and "$A \lesssim B$" denotes $A \leq CB$ for some constant C which, in this case, depends on the dimension. Later on, in weighted inequalities, this constant may also depend on the weights.

1.1.1 Weights and Bloom's Inequality.

The project I worked on in the last three years started with the goal of obtaining a two-weight extension of this result, using dyadic methods. A *weight* is a locally integrable, positive function w on $\mathbb{R}^n$. In terms of weighted inequalities, one is interested in operators on the spaces $L^p(w)$, the Lebesgue spaces corresponding to the measures $\int f\,dw := \int f(x)w(x)\,dx$. Of particular interest are the A_p weights – those weights for which the Hardy-Littlewood maximal function, or the Hilbert transform, is bounded on $L^p(w)$ [38,51]. A weight w on $\mathbb{R}^n$ is said to be a Muckenhoupt A_p weight, $1 < p < \infty$, provided that

$$[w]_{A_p} := \sup_Q \langle w\rangle_Q \langle w^{1-p'}\rangle^{p-1} < \infty,$$

where p' denotes the Hölder conjugate of p, and the supremum is over all cubes in $\mathbb{R}^n$. Note that A_2 weights are particularly nice to work with: $[w]_{A_2} := \sup_Q \langle w\rangle_Q \langle w^{-1}\rangle_Q$.

The one-weight theory, studying operators $L^p(w) \to L^p(w)$, is very well-developed at this point. However, the two-weight theory, studying operators $L^p(\mu) \to L^p(\lambda)$, is much more difficult, since most techniques in the one-weight theory do not work in this setting. In terms of commutators, the original characterization of Coifman, Rochberg and Weiss still stands in the one-weight setting – i.e. for any A_p weight w on $\mathbb{R}^n$, there holds

$$\|[b,T] : L^p(w) \to L^p(w)\| \lesssim \|b\|_{BMO(\mathbb{R}^n)}, \quad \text{and} \quad \|b\|_{BMO(\mathbb{R}^n)} \lesssim \sum_{j=1}^n \|[b, R_j] : L^p(w) \to L^p(w)\|.$$

In the two-weight setting however, the usual BMO space is no longer the correct space. Instead, as showed by Bloom [3] in 1985, one should work with a weighted BMO space.

For a weight w on $\mathbb{R}^n$, define the *weighted BMO* space by:

$$\|b\|_{BMO(w)} := \sup_Q \frac{1}{w(Q)} \int_Q |b(x) - \langle b\rangle_Q|\, dx.$$

Remark here that this definition is a little surprising, since the weight only appears in front of the integral. There are several other definitions of weighted BMO in the literature, involving the weight inside the integral in some form, and these spaces are not equivalent. For reasons which will become apparent in what follows, the definition above should be the "correct" definition of weighted BMO.

Bloom's result [3] characterizes the norm of the commutator with the Hilbert transform, acting between two weighted spaces $L^p(\mu)$ and $L^p(\lambda)$, where $1 < p < \infty$ and μ, λ are A_p weights, as follows:

$$\|[b, H] : L^p(\mu) \to L^p(\lambda)\| \simeq \|b\|_{BMO(\nu)}.$$

Here ν is a third weight, defined in terms of μ and λ as follows:

$$\nu := \mu^{1/p} \lambda^{-1/p}.$$

This was a beautiful result, a bit ahead of its time, having remained largely unremarked for almost 30 years. The project I joined, initiated by Michael Lacey and Brett Wick, had in mind to ultimately extend this result to multiparameter harmonic analysis – a notoriously difficult field in which dyadic methods have recently been shown to produce results. Bloom's original proof is not suited to this goal, as it makes use of several objects which are not available in the multiparameter setting (such as the sharp function). So, the first task was to reprove Bloom's result using modern dyadic methods (this was accomplished in [30]), followed by a generalization of this one-dimensional result to all CZOs on $\mathbb{R}^n$ (accomplished in [31]), and finally on to multiparameter analysis (this was accomplished for biparameter Journé operators – biparameter generalizations of CZOs – in [32]).

1.1.2 Bloom's Inequality via Dyadic Methods

There will be three main dyadic components for us in what follows: dyadic grids, Haar functions, and paraproducts. The main idea, started by Petermichl in [59], is to represent difficult singular integral operators as probabilistic averages of simpler, dyadic operators.

Recall the standard dyadic grid on $\mathbb{R}^n$:

$$\mathcal{D}^0 := \left\{ 2^{-k} \left([0,1)^n + m \right) : k \in \mathbb{Z}; m \in \mathbb{Z}^n \right\}.$$

For every $\omega = (\omega_j)_{j \in \mathbb{Z}} \in (\{0,1\}^n)^{\mathbb{Z}}$ we may translate $\mathcal{D}^0$ by letting $\mathcal{D}_\omega := \{Q \dot{+} \omega : Q \in \mathcal{D}^0\}$, where $Q \dot{+} \omega := Q + \sum_{j : 2^{-j} < l(Q)} 2^{-j} \omega_j$. Here $l(Q)$ will denote the side length of a cube Q in $\mathbb{R}^n$. We will only need to pay attention to ω when dealing with $\mathbb{E}_\omega$, which denotes expectation with respect to the standard probability measure on the set of parameters ω. We denote a generic dyadic grid $\mathcal{D}_\omega$ on $\mathbb{R}^n$ by $\mathcal{D}$. Any such $\mathcal{D}$ has the standard nestedness properties:

- For every $P, Q \in \mathcal{D}$, $P \cap Q$ is one of P, Q, and $\emptyset$.
- All $Q \in \mathcal{D}$ with $l(Q) = 2^{-k}$ for some fixed $k \in \mathbb{Z}$ partition $\mathbb{R}^n$.

Recall that every dyadic interval $I \subset \mathbb{R}$ is associated with two Haar functions:

$$h_I := h_I^0 := \frac{1}{\sqrt{|I|}} \left(\mathbb{1}_{I_-} - \mathbb{1}_{I_+} \right), \text{ and } h_I^1 := \frac{1}{\sqrt{|I|}} \mathbb{1}_I.$$

Note that h_I^0 is cancellative, while h_I^1 is non-cancellative. The cancellative Haar functions associated to a dyadic system on $\mathbb{R}$ form an orthonormal basis for $L^2(\mathbb{R})$, so we write $f = \sum_{I \in \mathcal{D}} \widehat{f}(I) h_I$, where

$\widehat{f}(I) := \langle f, h_I \rangle_{L^2(dx)}$ denotes the Haar coefficient of f on I. For any dyadic lattice $\mathcal{D}_\omega$ on $\mathbb{R}$, and every $I \in \mathcal{D}$, Petermichl's dyadic shift is defined by

$$Ш_\omega f := \frac{1}{\sqrt{2}} \widehat{f}(I)(h_{I_-} - h_{I_+}).$$

Looking at the action of this shift on a Haar function, we may think of it as turning a sine wave into a cosine wave. Petermichl's representation theorem [59] states that

$$Hf \simeq \mathbb{E}_\omega(Ш_\omega f).$$

From here, it is straightforward to see that one also has a representation for commutators $[b, H]f \simeq \mathbb{E}_\omega([b, Ш_\omega]f)$. This was the strategy employed in [30] to reprove Bloom's inequality in dimension 1: using the representation theorem, it suffices to prove the bound for $[b, Ш_\omega]$ instead of $[b, H]$. If the bound holds uniformly in ω, i.e. the choice of dyadic grid does not matter, then we are taking expectation of a constant in

$$\langle [b, T]f, g \rangle \simeq \mathbb{E}_\omega \langle [b, Ш_\omega]f, g \rangle,$$

and the result is proved.

The way we deal with $[b, Ш]$, where from now on we suppress the subscript ω, is paraproducts. In dimension 1, the paraproducts Π_b and Π_b^* with symbol function b are defined as

$$\Pi_b f := \sum_{I \in \mathcal{D}} \widehat{b}(I) \langle f \rangle_I h_I \quad \text{and} \quad \Pi_b^* f = \sum_{I \in \mathcal{D}} \widehat{b}(I) \widehat{f}(I) |I|^{-1} \mathbb{1}_I.$$

The usefulness of these operators comes from the fact that one can split the product of two functions as

$$bf = \Pi_b f + \Pi_b^* f + \Pi_f b.$$

In terms of commutators, this yields

$$[b, Ш]f = \Pi_b \, Ш f + \Pi_b^* \, Ш f - Ш \, \Pi_b f - Ш \, \Pi_b^* f + (\Pi_{Ш f} b - Ш \, \Pi_f b). \tag{1}$$

Since one-weight bounds for the dyadic shift are well-known [40], the first four terms will be bounded once we prove two-weight bounds for paraproducts. The "remainder" term in parentheses needs to be handled separately. In [30] and [31] we prove two-weight bounds for the paraproducts, of the form

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| \lesssim \|b\|_{BMO(\nu)}.$$

I will only highlight here the following interesting feature of this proof: in the end, it boils down to applying two separate one-weight inequalities for well-known operators like the maximal function and the dyadic square function. This is a little unusual in two-weight theory, and also it highlights the beauty of the weight ν as originally defined by Bloom. First, I made the rather simple observation that, if μ and λ are A_p weights, then $\nu := \mu^{1/p} \lambda^{-1/p}$ is an A_2 weight. This, in turn, opened the door to a very nice expression of weighted $H^1 - BMO$ duality valid for A_2 weights [46, 66]:

$$|\langle b, \Phi \rangle| \lesssim [\nu]_{A_2} \|b\|_{BMO(\nu)} \|S_{\mathcal{D}} \Phi\|_{L^1(\nu)}, \tag{2}$$

where $S_{\mathcal{D}}$ denotes the dyadic square function adapted to the dyadic grid $\mathcal{D}$:

$$S_{\mathcal{D}}^2 f := \sum_{I \in \mathcal{D}} |\widehat{f}(I)|^2 \frac{\mathbb{1}_I}{|I|}. \tag{3}$$

Then, for instance in the case of Π_b, one really needs to show that

$$|\langle \Pi_b f, g\rangle| \lesssim \|b\|_{BMO(\nu)} \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')},$$

where $\lambda' := \lambda^{1-p'}$ is the $A_{p'}$ weight "dual" to $\lambda \in A_p$. Isolate the b out as follows:

$$\langle \Pi_b f, g\rangle = \sum_{I\in\mathcal{D}} \widehat{b}(I)\langle f\rangle_I \widehat{g}(I) = \langle b, \Phi\rangle,$$

where $\Phi = \sum_{I\in\mathcal{D}} \langle f\rangle_I \widehat{g}(I) h_I$. From (2), we now have $|\langle \Pi_b f, g\rangle| \lesssim \|b\|_{BMO(\nu)} \|S_{\mathcal{D}}\Phi\|_{L^1(\nu)}$. Finally, it is easy to see that

$$S^2_{\mathcal{D}}\Phi = \sum_{I\in\mathcal{D}} \langle f\rangle_I^2 |\widehat{g}(I)|^2 \frac{\mathbb{1}_I}{|I|} \leq (Mf)^2 \cdot S^2_{\mathcal{D}} g,$$

where $Mf := \sup_I \langle |f|\rangle_I \mathbb{1}_I$ is the Hardy-Littlewood maximal function. This is where a simple application of Hölder's inequality finishes the proof by separating out μ and λ from ν:

$$\|S_{\mathcal{D}}\Phi\|_{L^1(\nu)} \leq \int Mf \cdot S_{\mathcal{D}} g \, d\mu^{1/p}\lambda^{-1/p} \leq \|Mf\|_{L^p(\mu)} \|S_{\mathcal{D}}\Phi\|_{L^{p'}(\lambda')},$$

and we may now apply classical one-weight bounds for M and $S_{\mathcal{D}}$. While this strategy becomes very technical when moving to multidimensional operators in [31], and even more complicated when moving to the biparameter case [32], at its core it remains the same, and runs through many of the results we have obtained.

We state next the more general result obtained later in [31], valid not only for the Hilbert transform, but for all CZOs in $\mathbb{R}^n$.

Theorem 1 (H., Lacey, Wick [31]). *Let T be a Calderón-Zygmund operator on $\mathbb{R}^n$ and $\mu, \lambda \in A_p$ with $1 < p < \infty$. Suppose $b \in BMO(\nu)$, where $\nu = \mu^{\frac{1}{p}}\lambda^{-\frac{1}{p}}$. Then*

$$\|[b, T] : L^p(\mu) \to L^p(\lambda)\| \leq c\|b\|_{BMO(\nu)},$$

where c is a constant depending on the dimension n, the operator T, and μ, λ, and p.

The underlying dyadic strategy is the same in the multidimensional case, using Hytönen's representation theorem [39] instead – this is a celebrated generalization of Petermichl's theorem to all CZOs on $\mathbb{R}^n$. The dyadic shifts become much more complicated, of course, since we are now dealing with n-dimensional cubes instead of intervals.

1.1.3 The Lower Bound Result and a Two-Weight John-Nirenberg Inequality

Also in [31] we obtain the lower bound result in the two-weight setting.

Theorem 2 (H., Lacey, Wick [31]). *For $1 < p < \infty$, and $\mu, \lambda \in A_p$, set $\nu = \mu^{\frac{1}{p}}\lambda^{-\frac{1}{p}}$. Then there are constants $0 < c < C < \infty$, depending only on n, p, μ and λ, for which*

$$c\|b\|_{BMO(\nu)} \leq \sum_{i=1}^{n} \|[b, R_i] : L^p(\mu) \to L^p(\lambda)\| \leq C\|b\|_{BMO(\nu)}. \tag{4}$$

Similar to [12], the equivalence above yields a weak-factorization result for weighted Hardy spaces.

Corollary 3 (H., Lacey, Wick [31]). *Under the hypotheses and notation of Theorems 1 and 2, let $\lambda' := \lambda^{1-p'}$ and T be a Calderón-Zygmund operator on $\mathbb{R}^n$. We have the inequality*

$$\|g_1(Tg_2) - (T^* g_1)g_2\|_{H^1(\nu)} \leq c\|g_1\|_{L^q(\lambda')}\|g_2\|_{L^p(\mu)},$$

where c is a constant depending on the dimension, the operator T, and on μ, λ, and p. Conversely, there exists a constant c so that every $f \in H^1(\nu)$ can be written as

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{\infty} g_j^i(x) R_i h_j^i(x) + h_j^i(x) R_i g_j^i(x), \tag{5}$$

where R_i is the Riesz transform in the ith variable, and $g_j^i \in L^q(\lambda')$, $h_j^i \in L^p(\mu)$ with

$$\sum_{i=1}^{n}\sum_{j=1}^{\infty} \|g_j^i\|_{L^q(\lambda')}\|h_j^i\|_{L^p(\mu)} \leq c\|f\|_{H^1(\nu)}.$$

The proof of Theorem 2 follows closely the original unweighted proof of Coifman, Rochberg and Weiss, using spherical harmonics. This proof was easy to adapt, once we obtained the following characterization of Bloom's BMO:

Theorem 4 (H., Lacey, Wick [31]). *Let $\mathcal{D}$ be a fixed dyadic grid on $\mathbb{R}^n$. The following are equivalent:*

(1) $b \in BMO^2_{\mathcal{D}}(\nu)$.

(2) *The operator $\Pi_b : L^p(\mu) \to L^p(\lambda)$ is bounded.*

(3) *The operator $\Pi_b^* : L^p(\mu) \to L^p(\lambda)$ is bounded.*

(4) *The operators Π_b and Π_b^* are bounded $L^2(\nu) \to L^2(\nu^{-1})$.*

(5) $\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) := \sup_{Q\in\mathcal{D}} \left(\frac{1}{\mu(Q)} \int_Q |b - \langle b\rangle_Q|^p \, d\lambda\right)^{\frac{1}{p}} < \infty$.

(6) $\mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda') := \sup_{Q\in\mathcal{D}} \left(\frac{1}{\lambda'(Q)} \int_Q |b - \langle b\rangle_Q|^q \, d\mu'\right)^{\frac{1}{q}} < \infty$.

(7) $b \in BMO_{\mathcal{D}}(\nu)$.

Above, $BMO_{\mathcal{D}}(\nu)$ denotes the dyadic weighted BMO space (where the supremum is only over all cubes $Q \in \mathcal{D}$), and

$$\|b\|_{BMO^2_{\mathcal{D}}(\nu)} := \left(\sup_{Q\in\mathcal{D}} \frac{1}{\nu(Q)} \int_Q |b - \langle b\rangle_Q|^2 \, d\nu^{-1}\right)^{1/2}.$$

Recall that the main application of the John-Nirenberg inequality (in the unweighted setting) is that the $BMO(\mathbb{R}^n)$ norm can be expressed not only in terms of an L^1 norm, but as any L^p norm:

$$\|b\|_{BMO(\mathbb{R}^n)} := \sup_Q \frac{1}{|Q|} \int_Q |b - \langle b\rangle_Q| \, dx \simeq \sup_Q \left(\frac{1}{|Q|} \int_Q |b - \langle b\rangle_Q|^p \, dx\right)^{1/p},$$

Therefore, the equivalences in Theorem 4 between (1) and (5), (6), (7), are two-weight versions of the John-Nirenberg theorem for weighted BMO.

1.1.4 Sparse Operators

The proof of Theorem 4 relied on, and was inspired by, a result of Muckenhoupt and Wheeden [53] which we state below – also a two-weight John-Nirenberg type theorem, involving an A_p weight w and its dual weight w'. In their original work, Muckehoupt and Wheeden did not track the A_p characteristics, and we used it also without tracking constants in [31]. Later, I gave a different proof of this result, using sparse operators and yielding sharp powers of the A_p characteristics. This was my main contribution in the paper [18]:

Theorem 5 (Muckenhoupt and Wheeden [53]; Duong, H., Li, Wick, Yang [18])**.** *Suppose* $1 < p < \infty$ *and* $w \in A^p(\mathbb{R}^n)$*. Let* $b \in \mathrm{BMO}_w(\mathbb{R}^n)$*. Then for any* $1 \leq r \leq p'$*, we have*

$$\|b\|_{\mathrm{BMO}_w(\mathbb{R}^n)} \approx \|b\|_{\mathrm{BMO}_{w,r}(\mathbb{R}^n)} := \left(\sup_Q \frac{1}{w(Q)} \int_Q \left| b(x) - \langle b \rangle_Q \right|^r w^{1-r}(x) dx \right)^{\frac{1}{r}}.$$

In particular, we have

$$\|b\|_{\mathrm{BMO}_w(\mathbb{R}^n)} \leq \|b\|_{\mathrm{BMO}_{w,r}(\mathbb{R}^n)} \leq C_{n,p,r} [w]_{A^p}^{\max\{1, \frac{1}{p-1}\}} \|b\|_{\mathrm{BMO}_w(\mathbb{R}^n)}.$$

Given $0 < \eta < 1$, a collection $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes is said to be η-sparse provided that for every $Q \in \mathcal{S}$, there is a measurable subset $E_Q \subset Q$ such that $|E_Q| \geq \eta |Q|$ and the sets $\{E_Q\}_{Q \in \mathcal{S}}$ are pairwise disjoint. Given $\Lambda > 1$, a family $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes is said to be Λ-Carleson provided that

$$\sum_{P \in \mathcal{S}, P \subset Q} |P| \leq \Lambda |Q|,$$

for all $Q \in \mathcal{S}$.

A remarkable property is that $\mathcal{S}$ is η-sparse if and only if it is $1/\eta$-Carleson [45, 47]. Given a sparse collection $\mathcal{S} \subset \mathcal{D}$, a sparse operator is one of the form:

$$\mathcal{A}_{\mathcal{S}} f(x) := \sum_{Q \in \mathcal{S}} \langle f \rangle_Q \mathbb{1}_Q(x).$$

Sparse operators are known to satisfy linear A_2 inequalities [45, 47]. The use of sparse operators is a very active research area at the moment, since this technique is incredibly powerful to prove sharp weighted inequalities. One usually wants to do a "domination by sparse," that is the operator T to be bounded is dominated pointwise by a sparse operator. More recently, it has been observed [15] that it is easier to dominate the bilinear forms $\langle Tf, g \rangle$, using the Calderón-Zygmund decompositions of f and g. This usually all happens locally inside some fixed cube Q_0, proving the desired estimate for compactly supported functions.

This was the strategy employed in my new proof of Theorem 5, with the following interesting twist: the bilinear form in this case, $\langle \mathbb{1}_{Q_0}(b - \langle b \rangle_{Q_0}), f \rangle$, involves a *weighted* BMO function b. This creates a problem for sums of the form

$$\sum_{Q \subset Q_0} |\widehat{b}(Q)| |\widehat{g}(Q)| \leq \big(\sum_{Q \subset Q_0} |\widehat{b}(Q)|^2 \big)^{1/2} \big(\sum_{Q \subset Q_0} |\widehat{g}(Q)|^2 \big)^{1/2},$$

since there is no nice way to estimate the first sum in terms of the weighted BMO norm of b (one could employ the square function, but that will bring in additional powers of $[w]_{A_p}$, and we are looking for sharp powers).

As it turns out, a modified CZ-decomposition for b lets us ultimately work with a function a in *unweighted* BMO, where one has the nice expression

$$\|a\|_{BMO(\mathbb{R}^n)} \simeq \sup_{Q_0} \Big(\frac{1}{|Q_0|} \sum_{Q \subset Q_0} |\widehat{a}(Q)|^2 \Big)^{1/2}.$$

To obtain this decomposition, one starts with the usual CZ decomposition *for the weight* w,

$$\mathcal{E} := \{\text{Maximal subcubes } R \subset Q_0 \text{ such that } \langle w \rangle_R > \alpha \langle w \rangle_{Q_0}\},$$

then, instead of defining the usual "good" and "bad" functions for w, we let

$$a := \mathbb{1}_{Q_0} b - \sum_{R \in \mathcal{E}} (b - \langle b \rangle_R) \mathbb{1}_R.$$

Then $\|a\|_{BMO(\mathbb{R}^n)} \leq 2\alpha \langle w \rangle_{Q_0} \|b\|_{BMO(w)}$, and $\widehat{b}(Q) = \widehat{a}(Q)$ for all $Q \subset Q_0$ with $Q \not\subset \cup_{R \in \mathcal{E}} R$. So, this is a sort of CZ-decomposition for weighted BMO functions, where the "good" function is in unweighted BMO.

1.1.5 Commutators with Fractional Integral Operators

Recall the classical fractional integral operator, or Riesz potential, on $\mathbb{R}^n$: let $0 < \alpha < n$ be fixed and, for a Schwartz function f define the fractional integral operator (or Riesz potential) I_α by

$$I_\alpha f(x) := \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}} \, dy.$$

These operators have been studied since 1949, when they were introduced by Marcel Riesz, and have since found many applications in analysis – such as Sobolev embedding theorems and PDEs.

To contrast these with CZOs, note for example that fractional integral operators are positive, which in many cases makes them easier to work with. On the other hand, the fractional integral operators do not commute with dilations and therefore can never boundedly map $L^p(dx)$ to itself. Additionally, the kernel of the fractional integral operator does not satisfy the standard estimates required for CZO kernels. Therefore, the theory of fractional integral operators is not just a subset of the theory of Calderón–Zygmund operators, and results which are known for CZOs also need to be proved for the fractional integral operators.

Commutators with Riesz potentials were first studied in [8]. The Bloom two-weight theory was extended to commutators with fractional integral operators in [33], joint work with graduate students Robert Rahm and Scott Spencer. In this case one works with the $A_{p,q}$ classes of weights, a generalization of A_p classes for the fractional integral setting introduced in [52]: we say that a weight w belongs to the $A_{p,q}$ class provided that

$$[w]_{A_{p,q}} := \sup_Q \langle w^q \rangle_Q \langle w^{-p'} \rangle_Q^{q/p'} < \infty.$$

We take a similar dyadic approach in this paper, where the role of the dyadic shifts will be played by the dyadic version of the fractional integral operator I_α, given by: $I_\alpha^{\mathcal{D}} f := \sum_{Q \in \mathcal{D}} |Q|^{\alpha/n} \langle f \rangle_Q \mathbb{1}_Q$. The main result in this paper is

Theorem 6 (H., Rham, Spencer [33]). *Suppose that* $\alpha/n + 1/q = 1/p$ *and* $\mu, \lambda \in A_{p,q}$. *Let* $\nu := \mu\lambda^{-1}$. *Then:*

$$\|[b, I_\alpha] : L^p(\mu^p) \to L^q(\lambda^q)\| \simeq \|b\|_{BMO(\nu)}.$$

1.1.6 Generalizations to Iterated Commutators in the One-Parameter Setting

The next natural extension was to consider higher iterates of the commutator $[b, T]$, which are also treated some in [12]. For some Calderón–Zygmund operator T, let $C_b^1(T) := [b, T]$, and

$$C_b^k(T) := [b, C_b^{k-1}(T)], \text{ for all integers } k \geq 1.$$

At this point, a few natural questions arise: (1) What is the norm of the kth iterate as a function of the norm of $b \in BMO(\mathbb{R}^n)$? (2) What happens if we attempt to compute the norm of this operator when it acts on $L^p(w)$ for a weight $w \in A_p$? (3) Is there an extension of Theorem 1 for the iterates? In the paper [10] Chung, Pereyra, and Perez provide answers to questions (1) and (2) and show that:

$$\left\| C_b^k(T) : L^2(w) \to L^2(w) \right\| \leq c \|b\|^k_{BMO(\mathbb{R}^n)} [w]_{A_2}^{k+1},$$

where c is a constant depending on n, k and T. However, a two-weight extension lies outside the scope of the results in [10]. Additionally, in [10, pg. 1166] they ask if it is possible to provide a proof of the norm of the iterates of commutators with Calderón–Zygmund operators via the methods of dyadic analysis.

The main goal of the paper [37], joint with B. Wick, is to extend Theorem 1 to the case of iterates, addressing question (3), and in the process show how to answer the question raised in [10]. The main result of the paper is:

Theorem 7 (H., Wick [37])**.** *Let T be a Calderón-Zygmund operator on $\mathbb{R}^n$ and $\mu, \lambda \in A_p$ with $1 < p < \infty$. Suppose $b \in BMO(\mathbb{R}^n) \cap BMO(\nu)$, where $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$. Then for all integers $k \geq 1$:*

$$\left\| C_b^k(T) : L^p(\mu) \to L^p(\lambda) \right\| \leq c \, \|b\|^{k-1}_{BMO} \|b\|_{BMO(\nu)},$$

where c is a constant depending on n, k, T, μ, λ, and p.

In particular, if $\mu = \lambda = w \in A_2$:

$$\left\| C_b^k(T) : L^2(w) \to L^2(w) \right\| \leq c \|b\|^k_{BMO} [w]_{A_2}^{k+1},$$

where c is a constant depending on n, k and T.

1.1.7 Bloom's Inequality for Biparameter Journé Operators

Perhaps the crowning achievement of this extensive project is the recent paper [32], which develops the Bloom theory for biparameter Calderón-Zygmund operators, known as Journé operators, and characterizes their norms in terms of a weighted version of the little bmo space of Cotlar and Sadosky [14]. This completes part of the initial goal to extend this theory to the setting of multi-parameter analysis. The main results are:

Theorem 8 (H., Petermichl, Wick [32])**.** *Let T be a biparameter Journé operator on $\mathbb{R}^{\vec{n}} = \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$. Let μ and λ be $A_p(\mathbb{R}^{\vec{n}})$ weights, $1 < p < \infty$, and define $\nu := \mu^{1/p} \lambda^{-1/p}$. Then*

$$\|[b, T] : L^p(\mu) \to L^p(\lambda)\| \lesssim \|b\|_{bmo(\nu)},$$

where $\|b\|_{bmo(\nu)}$ denotes the norm of b in the weighted little bmo(ν) space on $\mathbb{R}^{\vec{n}}$.

At its core, the strategy is the same as in [31], and may be roughly stated as:

1. Use a representation theorem to reduce the problem from bounding the norm of $[b, T]$ to bounding the norm of $[b, \text{Dyadic Shift}]$.

2. Prove the two-weight bound for $[b, \text{Dyadic Shift}]$ by decomposing into paraproducts.

However, the biparameter case presents some significant new obstacles. In [31], T was a Calderón-Zygmund operator on $\mathbb{R}^n$, and the representation theorem was that of Hytönen [39]. In [32], T is a biparameter Journé operator on $\mathbb{R}^{\vec{n}} = \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$, and we use Martikainen's representation theorem [48] to reduce the problem to commutators $[b, \mathbb{S}_{\boldsymbol{\mathcal{D}}}]$, where $\mathbb{S}_{\boldsymbol{\mathcal{D}}}$ is now a *biparameter* dyadic shift. These can be cancellative, i.e. all Haar functions have mean zero, or non-cancellative.

The main difficulty arises from the structure of the biparameter dyadic shifts. At first glance, the cancellative shifts are "almost" compositions of two one-parameter shifts $\mathbb{S}_{\mathcal{D}_1}$ and $\mathbb{S}_{\mathcal{D}_2}$ applied in each variable – if this were so, many of the results would follow trivially by iteration of the one-parameter results. Unfortunately, there is no reason for this to happen, so many of the inequalities needed for biparameter shifts must be proved from scratch.

Even more difficult is the case of non-cancellative shifts. These are really paraproducts, and there are three possible types that arise from the representation theorem:

1. Full standard paraproducts;

2. Full mixed paraproducts;

3. Partial paraproducts.

These methods were considered previously in [55] and [56] for the unweighted, $p = 2$ case. In [55] it was shown that

$$\|[b, T] : L^2(\mathbb{R}^{\vec{n}}) \to L^2(\mathbb{R}^{\vec{n}})\| \lesssim \|b\|_{bmo(\mathbb{R}^{\vec{n}})}, \tag{6}$$

where T is a *paraproduct-free* Journé operator. This restriction essentially means that all the dyadic shifts in the representation of T are *cancellative*, so the case of non-cancellative shifts remained open. This gap was partially filled in [56], which treats the case of non-cancellative shifts of standard paraproduct type. So the case of general Journé operators, which includes non-cancellative shifts of mixed and partial type in the representation, remained open even in the unweighted, $p = 2$ case. These types of paraproducts are notoriously difficult – see also [49] for a wonderful discussion of this issue. We fill this gap in [32], where we prove two-weight bounds of the type

$$\|[b, \mathbb{S}_{\boldsymbol{\mathcal{D}}}] : L^p(\mu) \to L^p(\lambda)\| \lesssim \|b\|_{bmo(\nu)},$$

where $\mathbb{S}_{\boldsymbol{\mathcal{D}}}$ is a non-cancellative shift. The same is proved for cancellative shifts.

At the backbone of all these proofs will be the biparameter paraproducts and a variety of biparameter square functions. For instance, in the case of the cancellative shifts, one can decompose the commutator as

$$[b, \mathbb{S}_{\boldsymbol{\mathcal{D}}}^{\vec{i},\vec{j}}]f = \sum [\mathsf{P_b}, \mathbb{S}_{\boldsymbol{\mathcal{D}}}^{\vec{i},\vec{j}}]f + \sum [\mathsf{p_b}, \mathbb{S}_{\boldsymbol{\mathcal{D}}}^{\vec{i},\vec{j}}]f + \mathcal{R}_{\vec{i},\vec{j}}f.$$

Here $\mathsf{P_b}$ runs through nine paraproducts associated with *product BMO*, and $\mathsf{p_b}$ runs through six paraproducts associated with *little bmo*, so we are dealing with *fifteen* paraproducts in total in the biparameter case. Some of these are straightforward generalizations of the one-parameter paraproducts, while some are more complicated "mixed" paraproducts. Two-weight bounds are proved for all these paraproducts, building on two essential blocks: the biparameter square functions developed in Section 3 of [32], and the weighted $H^1 - BMO$ duality in the product setting. In fact, Section 4 of [32] is a self-contained presentation of large parts of the weighted biparameter BMO theory.

We point out that in our quest to prove Theorem 8, we also obtain a much simplified proof of the following one-weight result for Journé operators, originally due to R. Fefferman:

Theorem 9 (Weighted Inequality for Journé Operators)**.** *Let T be a biparameter Journé operator on $\mathbb{R}^{\vec{n}} = \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$. Then T is bounded $L^p(w) \to L^p(w)$ for all $w \in A_p(\mathbb{R}^{\vec{n}})$, $1 < p < \infty$.*

A version of Theorem 9 first appeared in R. Fefferman and E. M. Stein [20], with restrictive assumptions on the kernel. Subsequently the kernel assumptions were weakened significantly by R. Fefferman in [21], at the cost of assuming the weight belongs to the more restrictive class $A_{p/2}$. This was due to the use of his sharp function $T^{\#}f = M_S(f^2)^{1/2}$, where M_S is strong maximal function. Finally, R. Fefferman improved his own result in [22], where he showed that the A_p class sufficed and obtained the full statement of Theorem 9. This was achieved by an involved bootstrapping argument based on his previous result [21]. Our proof in [32] of this result is significantly simpler, as it is only based on one-weight bounds for the biparameter dyadic shifts.

Finally, we also have the lower bound:

Theorem 10 (H., Petermichl, Wick [32])**.** *Let μ, λ be $A_p(\mathbb{R}^n \times \mathbb{R}^n)$ weights, and set $\nu = \mu^{1/p}\lambda^{-1/p}$. Then*

$$\|b\|_{bmo(\nu)} \lesssim \sup_{1 \leqslant k,l \leqslant n} \|[b, R_k^1 R_l^2]\|_{L^p(\mu) \to L^p(\lambda).}, \tag{7}$$

where R_k^1 and R_l^2 are the Riesz transforms acting in the first and second variable, respectively.

1.2 Future Directions: Commutators

1.2.1 Two-Weight Inequalities in Multi-Parameter Analysis.

The main difficulty in this theory, which is concerned with product spaces of the form $\mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$, is the definition of Chang-Fefferman product BMO [6,7]. One would like to maintain the classical duality and define $BMO(\mathbb{R} \otimes \mathbb{R}) := \mathcal{H}^1(\mathbb{R} \otimes \mathbb{R})^*$. A famous example of Carleson [5] shows that one cannot then simply mimic the one-parameter BMO condition by replacing intervals I with rectangles $I \times J$. Instead, one defines $\|b\|_{BMO(\mathbb{R}\otimes\mathbb{R})}$ in terms of *open* sets with *finite measure.*

For iterated multi-parameter commutators, one is interested in inequalities of the form:

$$c\|b\|_{BMO} \leq \|[\ldots[[b, T_1], T_2]\ldots, T_t]\|_{L^p \to L^p} \leq C\|b\|_{BMO},$$

where each T_i is a Calderón-Zygmund operator on $\mathbb{R}^{d_i}$. Such (unweighted) inequalities have been extensively studied in the literature [16, 17, 23, 24, 42–44].

Problem 11 (Upper bound)**.** *Let $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$ and for every $i = 1, \ldots, t$, let T_i be a Calderón-Zygmund operator on $\mathbb{R}^{d_i}$. Let $\mu, \lambda \in A_p(\mathbb{R}^{\vec{d}})$ and set $\nu := \mu^{1/p}\lambda^{-1/p}$. Then:*

$$\|[\ldots[[b, T_1], T_2]\ldots, T_t]\|_{L^p(\mu) \to L^p(\lambda)} \leq c\|b\|_{BMO(\nu)}, \tag{8}$$

where $BMO(\nu)$ is the weighted Chang-Fefferman product BMO. and c is a constant depending on $\vec{d}$, p, the operators T_i, and the weights μ, λ.

A proof here has several stages, each of which will be new in this level of generality. First, reduce the problem to commutators with dyadic shifts by appealing to the Hytönen representation theorem. As seen in [16, 42, 43], this idea can be carried out in the multi-parameter case for the type of commutator considered in Problem 11. The biggest challenge I see next is obtaining an inequality as in (2), for product BMO. The real difficulty here, due to the product BMO definition, will be proving a multi-parameter version of Theorem 5 [18, 53].

Problem 12 (Lower bound)**.** *Let $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$. For a vector $\vec{j} = (j_1, \ldots, j_t)$ with $1 \leq j_i \leq d_i$ for $i = 1, \ldots, t$, let R_{i,j_i} denote the j_i^{th} Riesz transform acting on $\mathbb{R}^{d_i}$. Let $\mu, \lambda \in A_p(\mathbb{R}^{\vec{d}})$. Then*

$$\|b\|_{BMO(\nu)} \leq c \sup_{\vec{j}} \|[\ldots[[b, R_{1,j_1}], R_{2,j_2}] \ldots, R_{t,j_t}]\|_{L^p(\mu) \to L^p(\lambda)},$$

where $\nu := \mu^{1/p}\lambda^{-1/p}$ and c is a constant depending on $\vec{d}$, p, and the weights μ and λ.

The one-parameter proof of this result in [31] relies on a computation in [12] in combination with Theorem 4. Unfortunately, the Coifman-Rochberg-Weiss approach is known to fail in the multi-parameter case. Instead, the strategies in [23, 44] seem to be the right direction to follow. It is unclear at the moment if a multi-parameter version of Theorem 4 will be needed. Regardless, proving such a result may be valuable and interesting in itself, as it characterizes a one-weight BMO norm in terms of two-weight operator norms of paraproducts.

A resolution to these problems would be valuable to the field because it allows one to determine whether or not a function belongs to the weighted Chang-Fefferman BMO space by testing the norm of an operator, instead of using the difficult product BMO definition.

1.2.2 Operator Theoretic Properties of Commutators.

Having discussed the boundedness of the operator $[b, T] : L^p(\mu) \to L^p(\lambda)$, a natural next question is: when does $[b, T]$ belong to certain important subclasses of bounded operators, such as the compact operators? A classic result in this direction is due to Uchiyama [65], who showed that, for T in a certain class of Calderón-Zygmund operators, $[b, T] : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is compact if and only if b is in the space $VMO(\mathbb{R}^n)$ of functions with *vanishing mean oscillation* – defined as the closure in $BMO(\mathbb{R}^n)$ of $\mathcal{C}_c^\infty(\mathbb{R}^n)$ [13, 63].

In the weighted setting, compactness of commutators was studied in [11], where it is shown that if $b \in VMO(\mathbb{R}^n)$, then $[b, T] : L^p(w) \to L^p(w)$ is compact, where $w \in A_p$ and T is any Calderón-Zygmund operator. As indicated by Uchiyama, the opposite implication, a sort of 'lower bound' result, likely requires specialization to some subclass of 'nice' Calderón-Zygmund operators, such as the Riesz transforms. Moreover, a two-weight characterization of compactness likely requires a *weighted* space of functions with vanishing mean oscillation, which has not been well studied.

Problem 13. *Let μ, λ be A_p weights on $\mathbb{R}^n$ and $\nu = \mu^{1/p}\lambda^{-1/p}$. Define $VMO(\nu)$ to be the closure of $\mathcal{C}_c^\infty$ in $BMO(\nu)$. Show that if $b \in VMO(\nu)$, then the commutator $[b, T] : L^p(\mu) \to L^p(\lambda)$ is compact, where T is a Calderón-Zygmund operator. Conversely, show that if the commutators $[b, R_j] : L^p(\mu) \to L^p(\lambda)$ with the Riesz transforms R_j are compact for $j = 1, \ldots, n$, then $b \in VMO(\nu)$. Obtain as a corollary for $\mu = \lambda$ the converse to the result in* [11].

An obvious initial strategy is to follow Uchiyama's proof, which rests on a supporting lemma that provides three mean oscillation decay properties equivalent to $VMO(\mathbb{R}^n)$.

Restricting one's attention to L^2-spaces, some questions can be asked about Schatten class operators, i.e. those with l^p-summable singular values. The space usually paired with S_p, the $L^2(\mathbb{R}^n)$ Schatten class operators, is the Besov space $B_p(\mathbb{R}^n)$, a subspace of $BMO(\mathbb{R}^n)$. Commutators and Schatten classes have been studied in the literature mostly in the context of Hankel operators [57, 58, 61, 64]. These are important objects in complex analysis which have deep connections to the commutator $[b, H]$ with the Hilbert transform, and their properties show that $[b, H] \in S_p(\mathbb{R})$ if and only if $b \in B_p(\mathbb{R})$. For a Calderón-Zygmund operator T on $\mathbb{R}^n$, a result of Janson and Wolff [41] presents a very interesting split phenomenon: If $p > n$, then $[b, T] \in S_p(\mathbb{R}^n)$ if and only if $b \in B_p(\mathbb{R}^n)$, but if $p \leq n$, then $[b, T] \in S_p(\mathbb{R}^n)$ if and only if b is constant, so $[b, T]$ vanishes.

Interestingly, this argument does not work for dimension $n = 1$ and the Hilbert transform. One can then ask under what conditions $[b, T]$, acting between some weighted L^2 spaces, is a Schatten class operator, for a Calderón-Zygmund operator T?

Problem 14. *Is there a unifying proof of the behavior of $[b, T]$ as a Schatten class operator on $L^2(\mathbb{R}^n)$? A proof using real analysis techniques which yields the result in all dimensions?*

The split phenomenon in Janson-Wolff [41] should also occur somehow in the weighted case. A treatment of Schatten class commutators with general Calderón-Zygmund operators in the weighted setting appears to be missing from the literature, and would be a valuable addition:

Problem 15. *If $[b, T]$ is Schatten class on $L^2(w)$, must b belong to the Besov space $B_p(\mathbb{R}^n)$? In the two-weight case: if $[b, T] : L^2(\mu) \to L^2(\lambda)$ is Schatten class, must b belong to some weighted Besov space, perhaps in a sense analogous to Bloom? What is the correct definition of weighted Besov space in this context? Does the converse hold?*

A good place to start both Problems 14 and are papers like [9, 60, 62], which treat this problem for paraproducts and other dyadic operators, such as the martingale transform.

1.3 Sharp Inequalities via Bellman Functions

As a postdoc at Michigan State, I am working on several projects involving sharp inequalities for operators in harmonic analysis using the Bellman function technique. An interesting work in progress involves the dyadic square function. It was conjectured by Bollobas [4] that the best constant C in the weak-type inequality for the dyadic square function S:

$$|\{x \in [0, 1] : Sf(x) \geq \lambda\}| \leq \frac{C}{\lambda}\|f\|_{L^1},$$

is given by $C = e^{-1/2} + \int_0^1 e^{-t^2/2}\, dt$. A proof of this sharpness result was later given in [54]. A recent project [29], joint with A. Volberg and P. Ivanisvili, gives an interesting new proof of this sharpness result. The method makes use of what would be the standard Bellman function for this problem:

$$\mathbb{M}(f, F, \lambda) := \sup \frac{1}{|J|}|\{x \in J : S\varphi(x) \geq \lambda\}|,$$

where supremum is over all functions φ supported in a dyadic interval J with $\langle\varphi\rangle_J = f$ and $\langle|\varphi|\rangle_J = F$. The twist is that we define a second Bellman function $\mathbb{L}$ (inspired by Bollobas' work [4]),

$$\mathbb{L}(f, p, \lambda) := \inf\langle|\varphi|\rangle_J,$$

where the infimum is over all functions φ, supported in $J \in \mathcal{D}$, such that

$$\langle\varphi\rangle_J = f \quad \text{and} \quad \frac{1}{|J|}|\{x \in J : S_J^2\varphi(x) \geq \lambda\}| = p.$$

Being defined as an infimum, this function will have most of the mirrored properties of $\mathbb{M}$ – replace concavity with convexity for example. Also mirroring $\mathbb{M}$, we show that $\mathbb{L}$ is the so-called "greatest subsolution" for its main inequality (instead of the usual "least subsolution").

In particular, we discover very interesting relationships between $\mathbb{M}$ and $\mathbb{L}$ – for instance, $\mathbb{L}(f, p, \lambda)$ is the smallest value of F for which $\mathbb{M}(f, F, \lambda) = p$:

$$\mathbb{L}(f, p, \lambda) = \inf\{F \geq |f| : \mathbb{M}(f, F, \lambda) = p\},$$

and $\mathbb{M}(f, F, \lambda)$ is the largest value of p such that $\mathbb{L}(f, p, \lambda) = F$:

$$\mathbb{M}(f, F, \lambda) = \sup\{p \in [0, 1] : \mathbb{L}(f, p, \lambda) = F\}.$$

Then also from the above, we have

$$\mathbb{M}\big(f, \mathbb{L}(f, p, \lambda), \lambda\big) = p \quad \text{and} \quad \mathbb{L}\big(f, \mathbb{M}(f, F, \lambda), \lambda\big) = F,$$

which are later used in the new proof of the sharpness of C.

As we briefly outline next, $\mathbb{M}$ and $\mathbb{L}$ yield each other's "sharp" obstacle conditions, and this intertwining between the two Bellman functions is also used to reprove the sharpness of the constant C. This approach seems to be new, and very intriguing to pursue in many other Bellman type problems.

Using the standard methods, we obtain so-called "obstacle conditions" for $\mathbb{M}$ and $\mathbb{L}$, namely

$$\mathbb{M}(f, F, \lambda) = 1,\ \forall F \geq \sqrt{\lambda} \ \ \text{and} \ \ \mathbb{L}(f, p, \lambda) = |f|,\ \forall |f| \geq \sqrt{\lambda}.$$

While these suffice, as expected, to prove the least supersolution and greatest subsolution results, there is no reason to believe they are optimal. That is, $\mathbb{M}$ could very well be equal to 1 for some points where $F < \sqrt{\lambda}$, for instance. As it turns out, we may obtain the optimal obstacle condition for $\mathbb{M}$ from information about $\mathbb{L}$, and vice versa.

The value of $\mathbb{M}$ along the boundary $F = |f|$:

$$\mathbb{M}_b(f, \lambda) := \mathbb{M}(f, |f|, \lambda) = \sup\{p \in [0, 1] :\ \mathbb{L}(f, p, \lambda) = |f|\},$$

yields the optimal obstacle condition for $\mathbb{L}$, and the value of $\mathbb{L}$ along the boundary $p = 1$:

$$\mathbb{L}_b(f, \lambda) := \mathbb{L}(f, 1, \lambda) = \inf\{F \geq |f| :\ \mathbb{M}(f, F, \lambda) = 1\},$$

yields the optimal obstacle condition for $\mathbb{M}$. We find $\mathbb{M}_b$ and $\mathbb{L}_b$ explicitly. We show that the function $\mathbb{M}_b$ is given by

$$\mathbb{M}_b(|f|, \lambda) = \mathbb{M}(f, |f|, \lambda) = \begin{cases} \frac{\Phi\left(\frac{|f|}{\sqrt{\lambda}}\right)}{\Phi(1)}, & |f| < \sqrt{\lambda} \\ 1, & |f| \geq \sqrt{\lambda}. \end{cases} = \min\left(\frac{\Phi(|f|/\sqrt{\lambda})}{\Phi(1)}, 1\right),$$

where

$$\Phi(\tau) := \int_0^\tau e^{-x^2/2}\, dx,$$

for all $\tau \geq 0$. The function $\mathbb{L}_b$ is given by

$$\mathbb{L}_b(f, \lambda) = \mathbb{L}(f, 1, \lambda) = \begin{cases} \frac{\sqrt{\lambda}\Psi\left(\frac{|f|}{\sqrt{\lambda}}\right)}{\Psi(1)}, & 0 \leq |f| < \sqrt{\lambda} \\ |f|, & |f| \geq \sqrt{\lambda}. \end{cases} = \sqrt{\lambda} \max\left(\frac{\Psi(|f|/\sqrt{\lambda})}{\Psi(1)}, \frac{|f|}{\sqrt{\lambda}}\right),$$

where

$$\Psi(\tau) = \tau\Phi(\tau) + e^{-\tau^2/2},$$

for all $\tau \geq 0$.

To visualize the optimal obstacle conditions induced by $\mathbb{M}_b$ and $\mathbb{L}_b$ for $\mathbb{L}$ and $\mathbb{M}$, respectively, we restrict our attention to $f \geq 0$ and use homogeneity to express $\mathbb{M}$ and $\mathbb{L}$ as functions of two variables. Specifically, we write

$$\mathbb{M}(f, F, \lambda) = \mathbb{M}(f/\sqrt{\lambda}, F/\sqrt{\lambda}, 1) =: \theta(\tau, \gamma) \ \ \text{and} \ \ \mathbb{L}(f, p, \lambda) =: \sqrt{\lambda}\eta(\tau, p),$$

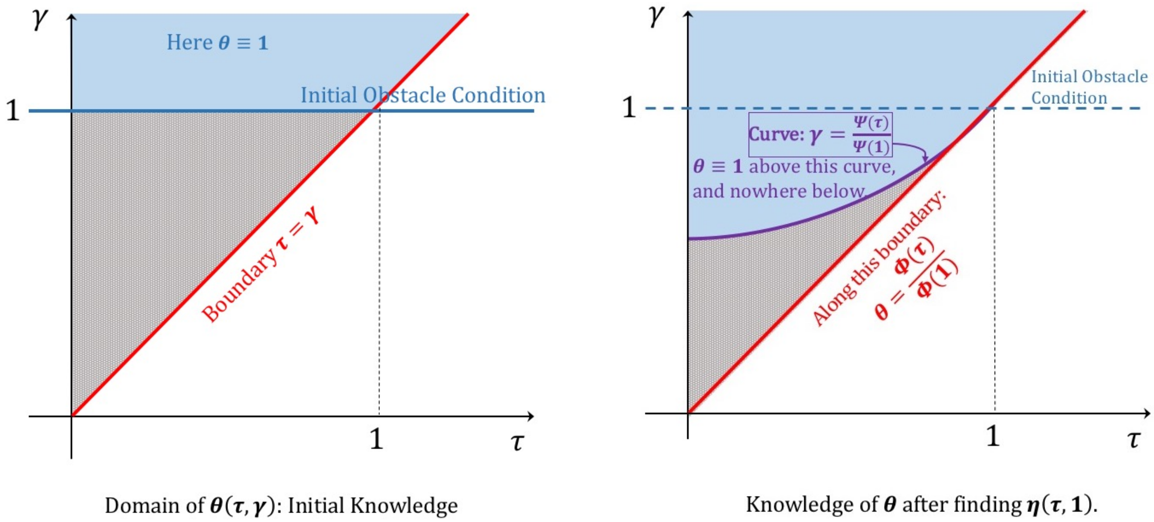

Figure 1: Initial and optimal Obstacle Conditions for θ.

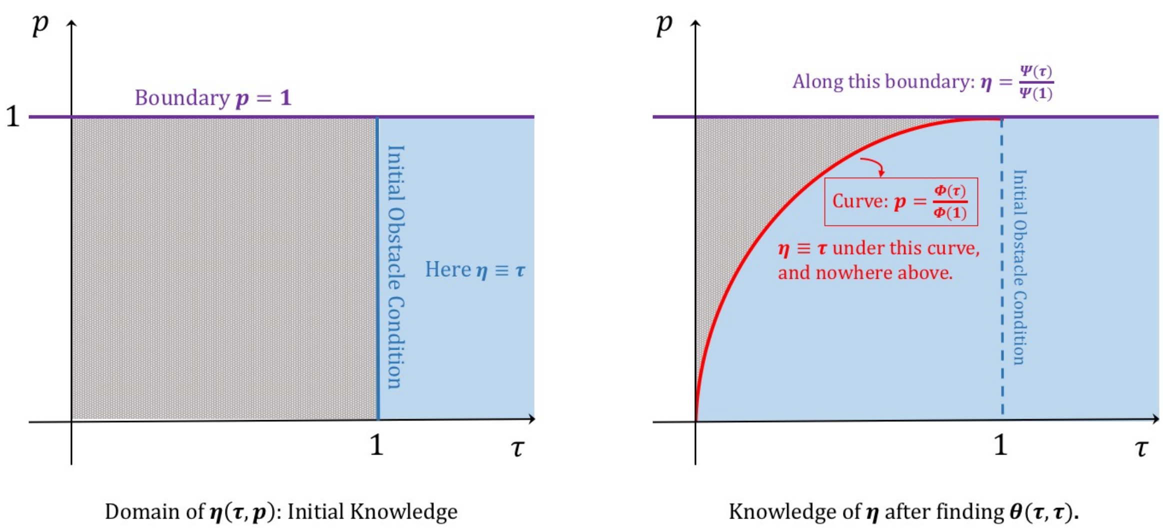

Figure 2: Initial and optimal Obstacle Conditions for η.

where $\tau = f/\sqrt{\lambda}$ and $\gamma = F/\sqrt{\lambda}$ Thus θ is defined on $\{0 \le \tau \le \gamma\}$, and η is defined on $\{0 \le p \le 1;\ \tau \ge 0\}$. The original obstacle conditions for $\mathbb{M}$ and $\mathbb{L}$ translate to

$$\theta(\tau, \gamma) = 1,\ \forall \gamma \ge 1 \quad \text{and} \quad \eta(\tau, p) = \tau,\ \forall \tau \ge 1.$$

Moreover,

$$\eta(\tau, 1) = \inf\{\gamma \ge \tau : \theta(\tau, \gamma) = 1\} \quad \text{and} \quad \theta(\tau, \tau) = \sup\{p : \eta(\tau, p) = \tau\}.$$

The expression for $\mathbb{L}_b$ gives that

$$\eta(\tau,1)=\begin{cases}\frac{\Psi(\tau)}{\Psi(1)}, & 0\leq\tau<1\\ |\tau|, & |\tau|\geq 1.\end{cases},$$

which yields the optimal obstacle condition for θ (see Figure 1). Similarly, $\mathbb{M}_b$ gives that

$$\theta(\tau,\tau)=\begin{cases}\frac{\Phi(\tau)}{\Phi(1)}, & 0\leq\tau\leq 1\\ 1, & \tau\geq 1.\end{cases},$$

which yields the optimal obstacle condition for η (see Figure 2).

2 Future Goals: A Confluence Between the Two Research Directions: Harmonic Analysis and Infinite-Dimensional Analysis

My graduate research was focused on developing the Gaussian Radon transform, an infinite dimensional version of the Radon transform within the framework of abstract Wiener spaces - first introduced by Leonard Gross in [25]. Ambar Sengupta, my doctoral adviser, was a student of Gross, and equipped me with a uniquely advantageous insight into this theory. In collaboration with my adviser, I established an infinite dimensional version of the Helgason support theorem, an inversion theorem for the Gaussian Radon transform, and applications of this transform to machine learning.

The main force driving my desired future efforts is that Gross's inspiration for developing the theory of abstract Wiener spaces was exactly his quest in studying potential theory in infinite dimensions. In order to wield this elegantly powerful theory in its entirety, I can bring my abstract Wiener space expertise full-circle, having acquired the scope of harmonic analysis.

2.1 A Gaussian Radon Transform in Infinite Dimensions

The Radon transform Rf of a function $f:\mathbb{R}^n\to\mathbb{R}$ assigns each hyperplane P in $\mathbb{R}^n$ the value $Rf(P)=\int_P f(x)\,dx$, where integration is with respect to Lebesgue measure on P. The ability of Rf to reconstruct the density of an n-dimensional object from its $(n-1)$-dimensional cross-sections in different directions makes the Radon transform the mathematical backbone of CT scans, tomography, and other image reconstruction applications. In our setting, the analogous goal will be the ability to extract information about a function defined on an infinite dimensional space from its conditional expectations. Generalizations of the Radon transform to infinite dimensional spaces had been studied before in [1,2,50], but not yet in the context of abstract Wiener spaces - currently the most useful setting in infinite dimensional analysis and probability. This was accomplished in my thesis [27] and related publications [28,34–36].

2.1.1 Abstract Wiener Spaces.

An *abstract Wiener space* is a triple (H,B,μ), where $(H,\|\cdot\|)$ is a real separable Hilbert space, $(B,|\cdot|)$ is the Banach space obtained by completing H with respect to a *measurable norm* $|\cdot|$, and μ is *Wiener measure*. A norm $|\cdot|$ on H is said to be a *measurable norm* provided that for every $\epsilon>0$, there is a finite-dimensional subspace F_ϵ of H such that $\gamma_F\{x\in F:|x|>\epsilon\}<\epsilon$, for all finite-dimensional subspaces F of H that are orthogonal to F_ϵ, where γ_F denotes standard Gaussian measure on F.

The inspiration behind the definition of measurable norms is the following. Standard Gaussian measure γ on a finite-dimensional Hilbert space H can be defined as the distribution measure of a random variable $Z = Z_1e_1 + \cdots + Z_ne_n : \Omega \to H$, where $\{e_1, \ldots, e_n\}$ is an orthonormal basis for H and $Z_1, \ldots, Z_n$ are independent standard Gaussian random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Specifically, $\gamma(E) = \mathbb{P}[Z \in E] = (2\pi)^{-\frac{n}{2}} \int_E e^{-\frac{\|x\|^2}{2}}\, dx$, an expression which clearly makes little sense if $n \to \infty$. If we try the same procedure and take an orthonormal basis $\{e_k\}_{k\geq 1}$ of a real separable infinite-dimensional Hilbert space H and an independent sequence $\{Z_k\}_{k\geq 1}$ of standard Gaussian random variables, the random variable $Z = \sum_{k=1}^{\infty} Z_ke_k$ on Ω, which we would like to be H-valued, does *not* converge almost everywhere.

In an effort to "force" this random variable to converge, Gross thought perhaps Z would converge with respect to a weaker norm than the original Hilbert norm. This turned out to be true, under certain conditions on this new norm: If we complete H with respect to a measurable norm $|\cdot|$ and obtain the Banach space B, properties of $|\cdot|$ reveal that the random variable $Z = \sum_{k=1}^{\infty} Z_ke_k$, thought of as B-valued instead of H-valued, converges almost everywhere for some orthonormal basis. The resulting distribution measure μ is a centered Gaussian measure on B.

2.1.2 The Gaussian Radon Transform.

As a first step in defining a Radon transform on (H, B, μ), we set out to define appropriate measures on hyperplanes of B. More generally, we constructed in [34] Gaussian measures concentrated on closed affine subspaces of B.

Theorem 16 (Construction of Gaussian Measures; H., Sengupta [34])**.** *For every closed subspace M_0 of H and $p \in M_0^{\perp}$, there is a unique Gaussian measure μ_{M_p} on B with mean $\langle p, h_{x^*}\rangle$ and variance $\|P_{M_0}h_{x^*}\|^2$, where P_{M_0} denotes orthogonal projection onto M_0, and for every $x^* \in B^*$, we denote by h_{x^*} the unique element in H determined by $x^*(h) = \langle h, h_{x^*}\rangle_H$ for all $h \in H$. Specifically:*

$$\int_B e^{ix^*}\, d\mu = e^{i\langle p, h_{x^*}\rangle - \frac{1}{2}\|P_{M_0}h_{x^*}\|^2},$$

for all $x^* \in B^*$. Moreover, the measure μ_{M_p} is concentrated on the closure of $M_p = p + M_0$ in B.

We then define the *Gaussian Radon transform* Gf of a Borel function f on B by:

$$Gf(p + M_0) = \int_B f\, d\mu_{M_p}.$$

Restricting the definition above to subspaces of codimension 1 gives a Radon transform on B.

Exploring the hyperplanes of B in more depth, I discovered that every hyperplane in B is, in fact, the closure of a hyperplane in H. However, a hyperplane in H determined by an element $u \in H^*$ that is not continuous with respect to $|\cdot|$ is dense in B, so, in a sense, there are "more" hyperplanes in H than in B. My thesis [27] proved a similar result for subspaces of finite codimension. The following result was also proved in [34]:

Theorem 17 (Infinite Dimensional Helgason Theorem; H., Sengupta [34])**.** *If f is a bounded continuous function on H and Gf is 0 on all hyperplanes in H that do not intersect a closed, bounded, convex set $K \subset H$, then f vanishes off of K.*

Theorem 18 (Disintegration of Wiener Measure; H. [28])**.** *Let (H, B, μ) be an abstract Wiener space, M_0 be a closed subspace of H and $\overline{M_0}$ denote the closure of M_0 in B. Then:*

$$\int_B f\, d\mu = \int_{\overline{M_0^{\perp}}} \left(\int_B f\, d\mu_{M_p} \right) d\mu_{M_0^{\perp}}(p)$$

for all Borel functions $f : B \to \mathbb{C}$ for which the left side exists.

This result, proved in [28], gives the connection of Gf to conditional expectation. Loosely speaking, the expectation of a Borel function f on B, conditioned by $\langle \cdot, h_j\rangle = y_j$, where $h_j \in H$, $1 \leq j \leq n$ are linearly independent and $y_j \in \mathbb{R}$, is the Gaussian Radon transform of f on the closed affine subspace determined by $\bigcap_{j=1}^{n} [\langle \cdot, h_j\rangle = y_j]$. These properties were used later in [28] to find an inversion procedure for Gf, utilizing the machinery of the Segal-Bargmann transform:

Theorem 19 (The Inversion Formula; H. [28])**.** *Let (H, B, μ) be an abstract Wiener space, $f \in L^2(B, \mu)$ and Q_0 be a closed subspace of H of finite codimension. Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis for $Q_0^{\perp}$. Then for all $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$:*

$$S_{Q_0^{\perp}}\left(Gf(p + Q_0)\right)(\alpha) = \left(S_B f\right)(\alpha \cdot u)$$

where $\alpha \cdot u = \alpha_1 u_1 + \ldots + \alpha_n u_n \in H_{\mathbb{C}}$ and $S_{Q_0^{\perp}}$ and S_B are the Segal-Bargmann transforms on $L^2(Q_0^{\perp}, \mu_{Q_0^{\perp}})$ and $L^2(B, \mu)$, respectively.

As a consequence, if we know $Gf(P)$ for all hyperplanes P in H, then we know $S_B f$ and can obtain f from the inverse Segal-Bargmann transform.

Finally, [36] provides computations of Gaussian Radon transforms on the *classical Wiener space*. Specifically, we derive formulas for transforms of Brownian functionals specified by stochastic integrals, and establish a Fock space decomposition for the L^2-spaces of the measures μ_{M_p}.

2.2 Future Directions: Machine Learning Applications

In [35] we explored connections between the Gaussian Radon transform and support vector machines. These methods involve projecting data into a reproducing kernel Hilbert space (RKHS), and their probabilistic interpretations have become increasingly interesting to the machine learning community. Our focus was the *ridge regression* method, which has a perfectly valid stochastic interpretation when the RKHS is finite dimensional, but the absence of Lebesgue measure breaks this interpretation down when the RKHS is infinite dimensional - a case which occurs quite frequently in practice. We showed that, through the lens of the Gaussian Radon transform, ridge regression has a probabilistic interpretation in infinite dimensions.

Specifically, the classical ridge regression solution $\hat{f}_\lambda(p)$, the prediction of the output at a future input p, can be expressed as $G\tilde{K}_p(L)$, the Gaussian Radon transform of a sort of "point evaluation" functional $\tilde{K}_p$ on the abstract Wiener space, evaluated on a closed subspace of finite codimension L determined by the training data. This result raised an interesting question, which I am currently exploring: using this stochastic approach, could we perhaps predict more than just the value at a future input p? For instance, could we predict the maximal value over a set S of future inputs by $GF(L)$, where $F = \sup_{p \in S} \tilde{K}_p$? Note that this will usually be different from predicting all the outputs over S first, and then taking the supremum.

2.3 Future Directions: Harmonic Analysis in Abstract Wiener Spaces

There is a rich theory of harmonic functions on abstract Wiener spaces. This goes back to the following fundamental idea of Gross, who set out to make sense of the $\mathbb{R}^n$ Poisson equation $\frac{1}{2}\Delta u = -g$ in infinite dimensions. In $\mathbb{R}^n$, an important solution is given by the potential of g, which is the convolution of g with Green measure F on $\mathbb{R}^n$:

$$F(E) = \frac{1}{(n-2)w_n} \int_E \frac{1}{|x|^{n-2}}\, dx, \tag{9}$$

for all Borel subsets E, where w_n is the surface area of the unit sphere. Clearly the formula above makes no sense, as is, if $n \to \infty$. However, (9) can also be expressed as:

$$F(E) = \int_0^\infty p_t(E)\,dt, \tag{10}$$

where for every $t > 0$, p_t is the probability measure on $\mathbb{R}^n$ given by:

$$p_t(A) = (2\pi t)^{-\frac{n}{2}} \int_A e^{-\frac{|x|^2}{2t}}\,dt,$$

for all Borel subsets A. Using (10) instead, Gross [26] was able to use the abstract Wiener space construction to define a meaningful analogue of the Laplacian in infinite dimensional spaces. Specifically, if one starts the process of constructing Wiener measure μ outlined in Section 2.1.1 by choosing the random variables Z_k to have variance $t > 0$ instead of variance 1, one easily obtains another centered Gaussian measure μ_t on B, with variance t. These measures form a strongly continuous contraction semigroup, and:

$$F(E) = \int_0^\infty \mu_t(E)\,dt$$

is the analogue of Green's measure in infinite dimensions. From here, with the aid of Frechet derivatives, all the regular notions of Laplacian, Hilbert, and Riesz transforms have valid interpretations for abstract Wiener spaces, thus paving the way for a plethora of other notions still to be explored, such as BMO and H^1 spaces, weights, and commutators.

References

[1] Jeremy J. Becnel, *The support theorem for the Gauss-Radon transform*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **15** (2012), no. 2, 1250006, 21, DOI 10.1142/S0219025712500063. MR2957133

[2] Jeremy J. Becnel and Ambar N. Sengupta, *A support theorem for a Gaussian Radon transform in infinite dimensions*, Trans. Amer. Math. Soc. **364** (2012), no. 3, 1281–1291, DOI 10.1090/S0002-9947-2011-05365-1, available at `https://www.math.lsu.edu/~sengupta/papers/RGBecSenDec09.pdf`. MR2869177

[3] Steven Bloom, *A commutator theorem and weighted BMO*, Trans. Amer. Math. Soc. **292** (1985), no. 1, 103–122.

[4] B. Bollobás, *Martingale inequalities*, Math. Proc. Camb. Phil. Soc. **87** (1980).

[5] Lennart Carleson, *A counterexample for measures bounded on H^p spaces for the bidisk*, Mittag-Leffler Rep. **7** (1974).

[6] Sun-Yung A. Chang and Robert Fefferman, *A continuous version of duality of H^1 with BMO on the bidisc*, Ann. of Math. (2) **112** (1980), no. 1, 179–201.

[7] Sun-Yung A. Chang and Robert Fefferman, *Some recent developments in Fourier analysis and H^p-theory on product domains*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), no. 1, 1–43.

[8] S. Chanillo, *A note on commutators*, Indiana Univ. Math. J. **31** (1982), no. 1, 7–16.

[9] J.-A. Chao and Lizhong Peng, *Schatten classes and commutators on simple martingales*, Colloq. Math. **71** (1996), no. 1, 7–21.

[10] Daewon Chung, M. Cristina Pereyra, and Carlos Perez, *Sharp Bounds for General Commutators On Weighted Lebesgue Spaces*, Transactions of the American Mathematical Society **364** (2012), no. 3, 1163–1177.

[11] Albert Clop and Victor Cruz, *Weighted estimates for Beltrami equations*, Ann. Acad. Sci. Fenn. Math. **38** (2013), no. 1, 91–113.

[12] R. R. Coifman, R. Rochberg, and Guido Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635.

[13] Ronald R. Coifman and Guido Weiss, *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 4, 569–645.

[14] M. Cotlar and C. Sadosky, *Two distinguished subspaces of product BMO and Nehari-AAK theory for Hankel operators on the torus*, Integr. Equ. Oper. Theory **26** (1996), no. 3.

[15] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *Domination of multilinear singular integrals by positive sparse forms* (2016), available at `https://arxiv.org/abs/1603.05317`.

[16] Laurent Dalenc and Yumeng Ou, *Upper Bound for Multi-parameter Iterated Commutators* (2014), 1–25 pp., available at `http://arxiv.org/abs/1401.5994`.

[17] Laurent Dalenc and Stefanie Petermichl, *A lower bound criterion for iterated commutators*, J. Funct. Anal. **266** (2014), no. 8, 5300–5320.

[18] Xuan Thinh Duong, Irina Holmes, Ji Li, Brett D. Wick, and Dongyong Yang, *Two weight Commutators in the Dirichlet and Neumann Laplacian settings* (2017), available at `https://arxiv.org/abs/1705.06858`.

[19] Charles Fefferman, *Characterizations of bounded mean oscillation*, Bull. Amer. Math. Soc. **77** (1971), 587–588.

[20] R. Fefferman and E. M. Stein, *Singular integrals on product spaces*, Adv. Math. **45** (1982), no. 2.

[21] R. Fefferman, *Harmonic Analysis on Product Spaces*, Ann. of Math. **126** (1987), no. 1.

[22] R. Fefferman, *A^p weights and singular integrals*, Amer. J. Math. **110** (1988), no. 5.

[23] Sarah H. Ferguson and Michael T. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143–160.

[24] Sarah H. Ferguson and Cora Sadosky, *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures*, J. Anal. Math. **81** (2000), 239–267.

[25] Leonard Gross, *Abstract Wiener spaces*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42. MR0212152 (35 #3027)

[26] Leonard Gross, *Abstract Wiener measure and infinite dimensional potential theory*, Lectures in Modern Analysis and Applications, II, Lecture Notes in Mathematics, Vol. 140. Springer, Berlin, 1970, pp. 84–116.

[27] Irina Holmes, *The Gaussian Radon Transform for Banach Spaces*, Ph.D. Thesis, Louisiana State University, 2014, `etd.lsu.edu/docs/available/etd-06272014-122519/unrestricted/Holmesdiss.pdf`.

[28] Irina Holmes, *An Inversion Formula for the Gaussian Radon Transform for Banach Spaces*, Infinite Dimensional Analysis, Quantum Probability and Related Topics **16** (2013), no. 4, available at `http://arxiv.org/abs/1308.1392`.

[29] Irina Holmes, Paata Ivanisvili, and Alexander Volberg, *The Sharp Constant in the Weak (1,1) Inequality for the Square Function: A New Proof*, available at `https://arxiv.org/abs/1710.01346`.

[30] Irina Holmes, Michael T. Lacey, and Brett D. Wick, *Bloom's Inequality: Commutators in a Two-Weight Setting*, Archiv der Mathematik **106** (2016), no. 1, available at `http://arxiv.org/abs/1505.07947`.

[31] Irina Holmes, Michael T. Lacey, and Brett D. Wick, *Commutators in the Two-Weight Setting*, Mathematische Annalen **367** (2017), no. 1–2, available at `http://arxiv.org/abs/1506.05747`.

[32] Irina Holmes, Stefanie Petermichl, and Brett D. Wick, *Weighted little bmo and two-weight inequalities for Journé commutators* (2017), available at `https://arxiv.org/abs/1701.06526`.

[33] Irina Holmes, Robert Rahm, and Scott Spencer, *Two-Weight Inequalities for Commutators with Fractional Integral Operators*, Studia Math **233** (2016), no. 3, available at `https://arxiv.org/abs/1510.05331`.

[34] Irina Holmes and Ambar Sengupta, *A Gaussian Radon transform for Banach spaces*, Journal of Functional Analysis **263** (2012), no. 11, available at `http://arxiv.org/abs/1208.5743`.

[35] Irina Holmes and Ambar Sengupta, *The Gaussian Radon Transform and Machine Learning*, Infinite Dimensional Analysis, Quantum Probability and Related Topics **18** (2015), no. 3, available at `http://arxiv.org/abs/1310.4794`.

[36] Irina Holmes and A. Sengupta, *The Gaussian Radon Transform in Classical Wiener Space*, Communications on Stochastic Analysis **8** (2014), no. 2, available at `https://www.math.lsu.edu/cosa/8-2-07[444].pdf`.

[37] Irina Holmes and Brett D. Wick, *Two Weight Inequalities for Iterated Commutators with Calderón-Zygmund Operators*, Journal of Operator Theory (to appear) (2017).

[38] Richard Hunt, Benjamin Muckenhoupt, and Richard Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. **176** (1973).

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

[40] T. P. Hytönen, M. T. Lacey, H. Martikainen, T. Orponen, M Reguera, E. T. Sawyer, and I. Uriarte-Tuero, *Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on A_p weighted spaces*, J. Anal. Math. **118** (2012), no. 1, 177–220.

[41] Svante Janson and Thomas H. Wolff, *Schatten classes and commutators of singular integral operators*, Ark. Mat. **20** (1982), no. 2, 301–310.

[42] Michael T. Lacey, Stefanie Petermichl, Jill C. Pipher, and Brett D. Wick, *multi-parameter Riesz commutators*, Amer. J. Math. **131** (2009), no. 3, 731–769.

[43] Michael T. Lacey, Stefanie Petermichl, Jill C. Pipher, and Brett D. Wick, *Iterated Riesz commutators: a simple proof of boundedness*, Harmonic analysis and partial differential equations, Contemp. Math., vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 171–178.

[44] Michael T. Lacey and Erin Terwilleger, *Hankel operators in several complex variables and product BMO*, Houston J. Math. **35** (2009), no. 1, 159–183.

[45] M. T. Lacey, *An elementary proof of the A^2 Bound*, Israel Journal of Mathematics **217** (2017), no. 15, 181–195.

[46] Ming-Yi Lee, Chin-Cheng Lin, and Ying-Chieh Lin, *A wavelet characterization for the dual of weighted Hardy spaces*, Proc. Amer. Math. Soc. **137** (2009), no. 12, 4219–4225.

[47] A. K. Lerner and F. Nazarov, *Intuitive dyadic calculus: the basics*, arXiv:1508.05639v1, 1–53.

[48] H. Martikainen, *Representation of bi-parameter singular integrals by dyadic operators*, Adv. Math. **229** (2012), no. 3.

[49] H. Martikainen and T. Orponen, *Some obstacles in characterising the boundedness of bi-parameter singular integrals*, Math. Z. **282** (2016).

[50] Vochita Mihai and Ambar N. Sengupta, *The Radon-Gauss transform*, Soochow J. Math. **33** (2007), no. 3, 415–433.

[51] Benjamin Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226.

[52] B. Muckenhoupt and R. L. Wheeden, *Weighted norm inequalities for fractional integrals*, Trans. Amer. Math. Soc. **192** (1974), 261–237.

[53] B. Muckenhoupt and R. L. Wheeden, *Weighted bounded mean oscillation and the Hilbert transform*, Studia Math. **54** (1975/76), no. 3, 221–237.

[54] A. Osekowski, *On the best constant in the weak type inequality for the square function of a conditionally symmetric martingale*, Statist. Probab. Lett. **79** (2009), 1536–1538.

[55] Y. Ou, S. Petermichl, and E. Strouse, *Higher order Journé commutators and characterizations of multi-parameter BMO*, Adv. Math. **291** (2016), 24–58.

[56] Y. Ou and S. Petermichl, *Little bmo and Journé commutators*, available at `http://yumengou.mit.edu/sites/default/files/documents/LittleBMOJourneYOSP.pdf`.

[57] V. V. Peller, *Smooth Hankel operators and their applications (ideals $\mathfrak{S}_p$, Besov classes, random processes)*, Dokl. Akad. Nauk SSSR **252** (1980), no. 1, 43–48 (Russian).

[58] V. V. Peller, *Description of Hankel operators of the class $\mathfrak{S}_p$ for $p > 0$, investigation of the rate of rational approximation and other applications*, Mat. Sb. (N.S.) **122(164)** (1983), no. 4, 481–510 (Russian).

[59] Stefanie Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris Sér. I Math. **330** (2000), no. 6, 455–460.

[60] Sandra Pott and Martin P. Smith, *Paraproducts and Hankel operators of Schatten class via p-John-Nirenberg theorem*, J. Funct. Anal. **217** (2004), no. 1, 38–78.

[61] Richard Rochberg and Stephen Semmes, *A decomposition theorem for BMO and applications*, J. Funct. Anal. **67** (1986), no. 2, 228–263.

[62] Richard Rochberg and Stephen Semmes, *Nearly weakly orthonormal sequences, singular value estimates, and Calderon-Zygmund operators*, J. Funct. Anal. **86** (1989), no. 2, 237–306.

[63] Donald Sarason, *Functions of vanishing mean oscillation*, Trans. Amer. Math. Soc. **207** (1975), 391–405.

[64] Stephen Semmes, *Trace ideal criteria for Hankel operators, and applications to Besov spaces*, Integral Equations Operator Theory **7** (1984), no. 2, 241–281.

[65] Akihito Uchiyama, *On the compactness of operators of Hankel type*, Tôhoku Math. J. (2) **30** (1978), no. 1, 163–171.

[66] S. Wu, *A wavelet characterization for weighted Hardy Spaces*, Rev. Mat. Iberoamericana **8** (1992), no. 3, 329–349.