A LOCAL INEQUALITY FOR HANKEL OPERATORS ON THE SPHERE AND ITS APPLICATION

Quanlei Fang and Jingbo Xia

Abstract. Let $H^2(S)$ be the Hardy space on the unit sphere S in $\mathbf{C}^n$. We establish a local inequality for Hankel operators $H_f = (1-P)M_f|H^2(S)$. As an application of this local inequality, we characterize the membership of H_f in the Lorentz-like ideal $\mathcal{C}_p^+$, $2n < p < \infty$.

1. Introduction

Throughout the paper, let S denote the unit sphere $\{z \in \mathbf{C}^n : |z| = 1\}$ in $\mathbf{C}^n$. We assume that the complex dimension n is greater than or equal to 2. Let σ be the positive, regular Borel measure on S that is invariant under the orthogonal group $O(2n)$, i.e., the group of isometries on $\mathbf{C}^n \cong \mathbf{R}^{2n}$ which fix 0. As usual, the measure σ is normalized in such a way that $\sigma(S) = 1$.

Recall that the Hardy space $H^2(S)$ is the closure of $\mathbf{C}[z_1, \dots, z_n]$ in $L^2(S, d\sigma)$. Let P be the orthogonal projection from $L^2(S, d\sigma)$ onto $H^2(S)$. Then the Hankel operator $H_f : H^2(S) \to L^2(S, d\sigma) \ominus H^2(S)$ is defined by the formula

$$H_f = (1-P)M_f|H^2(S).$$

There is a rich literature on various kinds of Hankel operators. See, e.g., [1,2,4-6,9-13,19]. This paper falls within the so-called "one-sided" theory of Hankel operators. We remind the reader that the term "one-sided" theory refers to the study of the Hankel operator H_f alone, whereas the simultaneous study of the pair H_f and $H_{\bar f}$ is called "two-sided" theory. By virtue of the identity

$$[M_f, P] = H_f - H_{\bar f}^*,$$

"two-sided" theory is equivalent to the study of the commutator $[M_f, P]$. By contrast, "one-sided" theory must deal with Hankel operators H_f that cannot be expressed in the form of $[M_g, P]$, $g \in L^2(S, d\sigma)$, which poses a much greater challenge.

Let us write $\mathbf{B}$ for the open unit ball $\{z \in \mathbf{C}^n : |z| < 1\}$ in $\mathbf{C}^n$. Furthermore, write

$$k_z(w) = \frac{(1-|z|^2)^{n/2}}{(1-\langle w, z\rangle)^n}, \qquad |z| < 1, \quad |w| \le 1.$$

Then k_z is the normalized reproducing kernel for the Hardy space $H^2(S)$. In the study of Hankel operators, an extremely important role is played by the scalar quantity

$$\mathrm{Var}(f; z) = \|(f - \langle f k_z, k_z\rangle)k_z\|^2,$$

Keywords. Hankel operator, local inequality, norm ideal.

$f \in L^2(S, d\sigma)$, $z \in \mathbf{B}$. One can think of $\mathrm{Var}(f; z)$ as the "variance" of f with respect to the probability measure $|k_z|^2 d\sigma$ on S, hence the notation.

It was shown in [16] that for $f \in L^2(S, d\sigma)$, the Hankel operator H_f is bounded if and only if $f - Pf \in \mathrm{BMO}$, which is equivalent to the boundedness of the function $z \mapsto \mathrm{Var}(f - Pf; z)$ on $\mathbf{B}$. Also, H_f is compact if and only if $f - Pf \in \mathrm{VMO}$ [16], which is equivalent to

$$\lim_{|z|\uparrow 1} \mathrm{Var}(f - Pf; z) = 0.$$

In [5] we proved that H_f belongs to the Schatten class $\mathcal{C}_p$, $2n < p < \infty$, if and only if

$$\int \mathrm{Var}^{p/2}(f - Pf; z) d\lambda(z) < \infty,$$

where $d\lambda$ is the standard Möbius-invariant measure on $\mathbf{B}$.

What sets the "two-sided" theory of Hankel operators apart from the "one-sided" theory is just one thing: If one has both Hankel operators H_f and $H_{\bar{f}}$ available, then one has the local inequality

$$\mathrm{Var}(f; z) \le \|H_f k_z\|^2 + \|H_{\bar{f}} k_z\|^2 \tag{1.1}$$

for every $z \in \mathbf{B}$ (see [15,(6.4)]). Most of the difficulties that are particular to the "one-sided" theory of Hankel operators can be traced to the single fact that there is nothing comparable to (1.1) in the "one-sided" theory. Nothing, that is, up to this point.

One of the motivations for this paper is to find a local inequality analogous to (1.1) in the context of the "one-sided" theory of Hankel operators. This we manage to do. As it turns out, our "one-sided" analogue of (1.1) enables us to characterize the membership of the Hankel operator H_f in the Lorentz-like ideal $\mathcal{C}_p^+$, $2n < p < \infty$.

To state our result, we first need to introduce a sequence of contractions in the radial direction of the ball. For each pair of $j \in \mathbf{N}$ and $z \in \mathbf{B}$, define

$$\rho_j(z) = \begin{cases} (1 - 4^j(1 - |z|^2))^{1/2}(z/|z|) & \text{if} \quad 4^j(1 - |z|^2) < 1; \\ 0 & \text{if} \quad 4^j(1 - |z|^2) \ge 1. \end{cases} \tag{1.2}$$

To better understand these contractions, notice the following relations: we have

$$\begin{cases} \rho_j(z)/|\rho_j(z)| = z/|z| \qquad \text{and} \\ 1 - |\rho_j(z)|^2 = 4^j(1 - |z|^2) \end{cases} \tag{1.3}$$

if $4^j(1 - |z|^2) < 1$.

Recall form [5] that for each $z \in \mathbf{B}$, we define the Schur multiplier

$$m_z(w) = \frac{1 - |z|}{1 - \langle w, z \rangle}, \quad |w| \le 1. \tag{1.4}$$

Obviously, the corresponding multiplication operator M_{m_z} is a contraction on $L^2(S,d\sigma)$. Note in particular that m_0 is just the constant function 1. With all the ingredients explained, we can now present our "one-sided" analogue of (1.1).

Theorem 1.1. *Given any $0 < \delta \leq 1/2$, there exists a constant $0 < C(\delta) < \infty$ which depends only on δ and the complex dimension n such that the inequality*

$$\mathrm{Var}^{1/2}(f - Pf; z) \leq C(\delta) \sum_{j=1}^{\infty} \frac{1}{2^{(1-\delta)j}} \|M_{m_{\rho_j(z)}} H_f k_{\rho_j(z)}\| \tag{1.5}$$

holds for all $f \in L^2(S,d\sigma)$ and $z \in \mathbf{B}$.

Since $\|M_{m_z}\| = \|m_z\|_\infty = 1$, a slightly weaker, but perhaps aesthetically more pleasing version of (1.5) is

$$\mathrm{Var}^{1/2}(f - Pf; z) \leq C(\delta) \sum_{j=1}^{\infty} \frac{1}{2^{(1-\delta)j}} \|H_f k_{\rho_j(z)}\|, \tag{1.6}$$

$f \in L^2(S,d\sigma)$ and $z \in \mathbf{B}$. From (1.6) we see immediately that if $\sup_{|z|<1} \|H_f k_z\| < \infty$, then $f - Pf \in \mathrm{BMO}$. Similarly, it is a trivial exercise to deduce from (1.6) that the condition

$$\lim_{|z|\uparrow 1} \|H_f k_z\| = 0$$

implies $f - Pf \in \mathrm{VMO}$. In other words, local inequality (1.6) recaptures the main results in [16], and indeed explains why these results hold true. But for the application that we will present in this paper, we need (1.5), the stronger version of the local inequality.

Recall that, for each $1 \leq p < \infty$, the formula

$$\|A\|_p^+ = \sup_{k\geq 1} \frac{s_1(A) + s_2(A) + \cdots + s_k(A)}{1^{-1/p} + 2^{-1/p} + \cdots + k^{-1/p}}$$

defines a symmetric norm for operators [7,Section III.14], where $s_1(A), \ldots, s_k(A), \ldots$ are the s-numbers of A. On any separable Hilbert space $\mathcal{H}$, the set

$$\mathcal{C}_p^+ = \{A \in \mathcal{B}(\mathcal{H}) : \|A\|_p^+ < \infty\}$$

is a norm ideal [7,Section III.2]. It is well known that $\mathcal{C}_p^+$ contains the Schatten class $\mathcal{C}_p$ and that $\mathcal{C}_p^+ \neq \mathcal{C}_p$. An interesting property of $\mathcal{C}_p^+$ is that it is not separable with respect to the norm $\|.\|_p^+$.

Let us also recall the notion of *symmetric gauge functions*. Let $\hat{c}$ be the linear space of sequences $\{a_j\}_{j\in\mathbf{N}}$, where $a_j \in \mathbf{R}$ and for each sequence $a_j \neq 0$ only for a finite number of j's. A symmetric gauge function (also called *symmetric norming function*) is a map

$$\Phi : \hat{c} \to [0, \infty)$$

that has the following properties:

(a) Φ is a norm on $\hat{c}$.

(b) $\Phi(\{1, 0, \ldots, 0, \ldots\}) = 1$.

(c) $\Phi(\{a_j\}_{j\in\mathbf{N}}) = \Phi(\{|a_{\pi(j)}|\}_{j\in\mathbf{N}})$ for every bijection $\pi : \mathbf{N} \to \mathbf{N}$.

See [7,page 71]. Each symmetric gauge function Φ gives rise to the *symmetric norm*

$$\|A\|_\Phi = \sup_{k\geq 1} \Phi(\{s_1(A), \ldots, s_k(A), 0, \ldots, 0, \ldots\})$$

for operators. On any separable Hilbert space $\mathcal{H}$, the set of operators

$$\mathcal{C}_\Phi = \{A \in \mathcal{B}(\mathcal{H}) : \|A\|_\Phi < \infty\}$$

is a norm ideal [7,page 68]. This term refers to the following properties of $\mathcal{C}_\Phi$:

- For any B, $C \in \mathcal{B}(\mathcal{H})$ and $A \in \mathcal{C}_\Phi$, $BAC \in \mathcal{C}_\Phi$ and $\|BAC\|_\Phi \leq \|B\|\|A\|_\Phi\|C\|$.
- If $A \in \mathcal{C}_\Phi$, then $A^* \in \mathcal{C}_\Phi$ and $\|A^*\|_\Phi = \|A\|_\Phi$.
- For any $A \in \mathcal{C}_\Phi$, $\|A\| \leq \|A\|_\Phi$, and the equality holds when $\mathrm{rank}(A) = 1$.
- $\mathcal{C}_\Phi$ is complete with respect to $\|.\|_\Phi$.

For an unbounded operator, s-numbers are, of course, not defined. But it will be convenient to adopt the convention that $\|X\|_\Phi = \infty$ for any unbounded operator X.

Given a symmetric gauge Φ, it is a common practice to extend its domain of definition beyond the space $\hat{c}$. Suppose that $\{b_j\}_{j\in\mathbf{N}}$ is an arbitrary sequence of real numbers, i.e., the set $\{j \in \mathbf{N} : b_j \neq 0\}$ is not required to be finite. Then we define

$$\Phi(\{b_j\}_{j\in\mathbf{N}}) = \sup_{k\geq 1} \Phi(\{b_1, \ldots, b_k, 0, \ldots, 0, \ldots\}). \tag{1.7}$$

For our purpose we also need to deal with sequences indexed by sets other than $\mathbf{N}$. If W is a countable, infinite set, then we define

$$\Phi(\{b_\alpha\}_{\alpha\in W}) = \Phi(\{b_{\pi(j)}\}_{j\in\mathbf{N}}),$$

where $\pi : \mathbf{N} \to W$ is any bijection. The definition of symmetric gauge functions guarantees that the value of $\Phi(\{b_\alpha\}_{\alpha\in W})$ is independent of the choice of the bijection π. To be thorough, let us also mention the case of finite sequences. For a finite index set $F = \{x_1, \ldots, x_\ell\}$, we define

$$\Phi(\{b_x\}_{x\in F}) = \Phi(\{b_{x_1}, \ldots, b_{x_\ell}, 0, \ldots, 0, \ldots\}).$$

In particular, associated with the ideal $\mathcal{C}_p^+$ is the symmetric gauge function Φ_p^+, which is defined as follows. Let $1 \leq p < \infty$. For each $\{a_j\}_{j\in\mathbf{N}} \in \hat{c}$, define

$$\Phi_p^+(\{a_j\}_{j\in\mathbf{N}}) = \sup_{k\geq 1} \frac{|a_{\pi(1)}| + \cdots + |a_{\pi(k)}|}{1^{-1/p} + \cdots + k^{-1/p}}, \tag{1.8}$$

where $\pi : \mathbf{N} \to \mathbf{N}$ is any bijection such that $|a_{\pi(j)}| \geq |a_{\pi(j+1)}|$ for every $j \in \mathbf{N}$, which exists because $a_j = 0$ for all but a finite number of j's. Then $\mathcal{C}_p^+ = \mathcal{C}_{\Phi_p^+}$. More precisely, the relation between the norm $\|\cdot\|_p^+$ and symmetric gauge function Φ_p^+ is that

$$\|A\|_p^+ = \Phi_p^+(\{s_1(A), \dots, s_k(A), \dots\}).$$

Theorem 1.6 in [5] implies that if Φ is a symmetric gauge function and if $0 < \|H_f\|_\Phi < \infty$ for some $f \in L^2(S, d\sigma)$, then $\mathcal{C}_\Phi \supset \mathcal{C}_{2n}^+$.

Theorem 1.1 enables us to characterize the membership $H_f \in \mathcal{C}_p^+$, $2n < p < \infty$. Our characterization result involves the notion of lattice in $\mathbf{B}$, which is defined in terms of the Bergman metric, as follows. For each $z \in \mathbf{B}\backslash\{0\}$, we have the Möbius transform

$$\varphi_z(w) = \frac{1}{1 - \langle w, z\rangle}\left\{z - \frac{\langle w, z\rangle}{|z|^2}z - (1 - |z|^2)^{1/2}\left(w - \frac{\langle w, z\rangle}{|z|^2}z\right)\right\} \tag{1.9}$$

of the unit ball $\mathbf{B}$. Recall that each φ_z is an involution, i.e., $\varphi_z \circ \varphi_z = \mathrm{id}$ [14,Theorem 2.2.2]. Also, we define $\varphi_0(w) = -w$. It is well known that the formula

$$\beta(z, w) = \frac{1}{2}\log\frac{1 + |\varphi_z(w)|}{1 - |\varphi_z(w)|}$$

defines a metric on $\mathbf{B}$. For each $z \in \mathbf{B}$ and each $a > 0$, we define the corresponding β-ball

$$D(z, a) = \{w \in \mathbf{B} : \beta(z, w) < a\}.$$

Definition 1.2. [18,Definition 1.1] (i) Let a be a positive number. A subset Γ of $\mathbf{B}$ is said to be a-separated if $D(z, a) \cap D(w, a) = \emptyset$ for all distinct elements z, w in Γ.
(ii) Let $0 < a < b < \infty$. A subset Γ of $\mathbf{B}$ is said to be an a, b-lattice if it is a-separated *and* has the property $\cup_{z\in\Gamma} D(z, b) = \mathbf{B}$.

As it was mentioned in [18], the simplest example of such a lattice is the following. Take any positive number $0 < a < \infty$, and then take any subset M of $\mathbf{B}$ that is *maximal* with respect to the property of being a-separated. Then M is an $a, 2a$-lattice in $\mathbf{B}$.

We can now characterize the membership $H_f \in \mathcal{C}_p^+$, $2n < p < \infty$.

Theorem 1.3. *Let $2n < p < \infty$ be given. Let $0 < a < b < \infty$ be positive numbers such that $b \geq 2a$. Then there exist constants $0 < c \leq C < \infty$ which depend only on the given p, a, b and the complex dimension n such that the inequality*

$$c\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma}) \leq \|H_f\|_p^+ \leq C\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma})$$

holds for every $f \in L^2(S, d\sigma)$ and every a, b-lattice Γ in $\mathbf{B}$.

The rest of the paper is organized as follows. The proof of Theorem 1.1, which is based on a collection of estimates of mean oscillation, both old and new, will be presented in Section 2. Sections 3-5 are taken up by the proof of the lower bound in Theorem 1.3.

Specifically, we introduce a standard decomposition of the ball $\mathbf{B}$ and modified kernel function $\psi_{z,t}$. With this decomposition and $\psi_{z,t}$, we define *partial sampling* operators $R_F^{(t)}$. In Section 3 we prove that for each given $t > 0$, there is a bound $C_{3.3}(t)$ for the norms of the operators $R_F^{(t)}$. This is the key step in the proof of the lower bound in Theorem 1.3. Then in Section 4 we use these $R_F^{(t)}$ to "sample" operators of the form $(H_f^* H_f)^{p/2}$. The result of this "sampling" together with the local inequality provided by Theorem 1.1, plus some counting arguments that are standard in connection with the decomposition of $\mathbf{B}$, give us a lower bound for $\|(H_f^* H_f)^{p/2}\|_\Phi$ in Proposition 5.5. In fact, Proposition 5.5 is a rather general result by itself. To complete the proof of the lower bound in Theorem 1.3, we need a special property (Lemma 5.7) of the family of symmetric gauge functions Φ_p^+, $1 < p < \infty$. Because of this special property, the general lower bound given in Proposition 5.5 implies the lower bound in Theorem 1.3.

The proof of the upper bound in Theorem 1.3 takes up Sections 6-8. For the upper bound, it suffices to work with commutators $[P, M_g]$ rather than H_f. In other words, the upper bound in Theorem 1.3 can be treated as a "two-sided" problem, which is how one usually deals with upper bounds for Hankel operators. The two main steps in the proof of the upper bound in Theorem 1.3 are a "reverse Hölder's inequality" involving Φ_p^+ and an interpolation. These two steps are accomplished in Sections 6 and 7 respectively. The proof of the upper bound is then completed in Section 8.

2. Various mean oscillations

As one might expect, the proof of Theorem 1.1 is a collection of estimates of mean oscillations. Fortunately, some of these estimates have been established previously [5,16]. But new estimates will also be needed. We begin with a review of what has already been established, and then progress to new material.

First of all, we will follow the notation in [5,14,16]. It is well known that the formula

$$d(\zeta, \xi) = |1 - \langle \zeta, \xi \rangle|^{1/2}, \qquad \zeta, \xi \in S, \tag{2.1}$$

defines a metric on S [14,page 66]. Throughout the paper, we denote

$$B(\zeta, r) = \{x \in S : |1 - \langle x, \zeta \rangle|^{1/2} < r\}$$

for $\zeta \in S$ and $r > 0$. There is a constant $2^{-n} < A_0 < \infty$ such that

$$2^{-n} r^{2n} \le \sigma(B(\zeta, r)) \le A_0 r^{2n} \tag{2.2}$$

for all $\zeta \in S$ and $0 < r \le \sqrt{2}$ [14,Proposition 5.1.4]. Note that the upper bound actually holds when $r > \sqrt{2}$. For $f \in L^2(S, d\sigma)$, $\zeta \in S$ and $r > 0$, we define

$$\mathrm{SD}(f; \zeta, r) = \left(\frac{1}{\sigma(B(\zeta, r))} \int_{B(\zeta,r)} |f - f_{B(\zeta,r)}|^2 d\sigma \right)^{1/2}, \tag{2.3}$$

where

$$f_E = \frac{1}{\sigma(E)} \int_E f d\sigma$$

if E is any Borel set with $\sigma(E) > 0$.

Proposition 2.1. [16,Proposition 2.2] *There exists a constant* $0 < C_{2.1} < \infty$ *such that the inequality*

$$\mathrm{SD}(Pf; \zeta, a) \le C_{2.1} \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{SD}(f; \zeta, 2^k a)$$

holds for all $f \in L^2(S, d\sigma)$, $\zeta \in S$ *and* $a > 0$.

Lemma 2.2. [5,Lemma 2.2] *There exists a constant* $0 < C_{2.2} < \infty$ *such that for all* $f \in L^2(S, d\sigma)$ *and* $z \in \mathbf{B}\backslash\{0\}$, *we have*

$$\|(f - \langle fk_z, k_z\rangle)k_z\| \le C_{2.2} \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{SD}(f; \zeta, 2^k a),$$

where $a = (1 - |z|^2)^{1/2}$ *and* $\zeta = z/|z|$.

Recall that for $z \in \mathbf{B}\backslash\{0\}$, the Möbius transform φ_z is defined by (1.9). Moreover,

$$\langle g \circ \varphi_z, 1\rangle = \langle gk_z, k_z\rangle$$

for every $g \in L^2(S, d\sigma)$ [14,page 44].

Lemma 2.3. [5,Lemma 2.3] *There is a constant* $C_{2.3}$ *such that the following estimate holds: Let* $0 < a < 1$ *and* $\zeta \in S$, *and set* $z = (1 - a^2)^{1/2}\zeta$. *Let* $f \in L^2(S, d\sigma)$. *Then for each* $a \le b \le 4$,

$$\mathrm{SD}(f \circ \varphi_z; \zeta, b) \le C_{2.3} \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{SD}(f; \zeta, 2^{k+2}(a/b)).$$

Lemma 2.4. *There is a constant* $C_{2.4}$ *such that the following estimate holds: Let* $0 < a < 1$ *and* $\zeta \in S$. *Set* $z = (1 - a^2)^{1/2}\zeta$. *Then*

$$\sum_{k=N}^{\infty} \frac{1}{2^k} \mathrm{SD}((P(f \circ \varphi_z)) \circ \varphi_z; \zeta, 2^k a) \le C_{2.4} \frac{1}{2^{\epsilon N}} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{SD}(f; \zeta, 2^j a)$$

for all $f \in L^2(S, d\sigma)$, $N \in \mathbf{N}$ *and* $0 < \epsilon \le 1/2$.

Proof. The more substantive half of the lemma, namely the case where the natural number $N \in \mathbf{N}$ satisfies the condition $2^N a \le 4$, was proved as Lemma 2.4 in [5]. Therefore we only need to consider the other half of the lemma, i.e., the case where $2^N a > 4$.

Let $f \in L^2(S, d\sigma)$. Since we now assume $2^N a > 4$, if $k \geq N$, then $B(\zeta, 2^k a) = S$. Consequently, with $c_z = \langle (P(f \circ \varphi_z)) \circ \varphi_z, 1 \rangle = \langle k_z P(f \circ \varphi_z), k_z \rangle$, we have

$$\sum_{k=N}^{\infty} \frac{1}{2^k} \mathrm{SD}((P(f \circ \varphi_z)) \circ \varphi_z; \zeta, 2^k a) = \frac{2}{2^N} \left\{ \int |(P(f \circ \varphi_z)) \circ \varphi_z - c_z|^2 d\sigma \right\}^{1/2}$$

$$= \frac{2}{2^N} \left\{ \int |P(f \circ \varphi_z) - c_z|^2 |k_z|^2 d\sigma \right\}^{1/2} = \frac{2}{2^N} \mathrm{Var}^{1/2}(P(f \circ \varphi_z); z). \tag{2.4}$$

Now we apply Lemma 2.2 and Proposition 2.1. This gives us

$$\frac{2}{2^N} \mathrm{Var}^{1/2}(P(f \circ \varphi_z); z) \leq \frac{2}{2^N} C_{2.2} \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{SD}(P(f \circ \varphi_z); \zeta, 2^k a)$$

$$\leq \frac{2C_{2.2}C_{2.1}}{2^N} \sum_{k=1}^{\infty} \frac{1}{2^k} \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^{k+\ell} a) \leq \frac{C_1}{2^N} \sum_{j=1}^{\infty} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a)$$

$$\leq \frac{C_2}{2^N} \sum_{j=1}^{J} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a), \tag{2.5}$$

where J is the smallest natural number satisfying the condition $2^J a \geq 4$ and $C_2 = C_1(1 + \sum_{j=1}^{\infty} (j+1)/2^j)$. Using Lemma 2.2 again, we have

$$\mathrm{SD}(f \circ \varphi_z; \zeta, 2^J a) = \mathrm{Var}^{1/2}(f \circ \varphi_z; 0) = \mathrm{Var}^{1/2}(f; z) \leq C_{2.2} \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{SD}(f; \zeta, 2^k a). \tag{2.6}$$

On the other hand, for each $1 \leq j \leq J-1$, the definition of J ensures $2^j a < 4$, consequently Lemma 2.3 can be applied with $b = 2^j a$. Therefore, by Lemma 2.3,

$$\sum_{j=1}^{J-1} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a) \leq C_{2.3} \sum_{j=1}^{J-1} \frac{j}{2^j} \sum_{i=1}^{\infty} \frac{1}{2^i} \mathrm{SD}\left(f; \zeta, 2^{i+2} \frac{a}{2^j a}\right)$$

$$= C_{2.3} \sum_{j=1}^{J-1} \frac{j}{2^j} \sum_{i=1}^{\infty} \frac{1}{2^i} \mathrm{SD}(f; \zeta, 2^{i-J+2+(J-j)}) \leq C_3 \sum_{j=1}^{J-1} \frac{j}{2^j} \sum_{i=1}^{\infty} \frac{1}{2^i} \mathrm{SD}(f; \zeta, 2^{i+(J-j)} a),$$

where the last $\leq$ follows from the inequality $2^{-J+2} \leq a < 2^{-J+3}$ and (2.2). Hence

$$\sum_{j=1}^{J-1} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a) \leq C_3 \sum_{m=1}^{\infty} \mathrm{SD}(f; \zeta, 2^m a) \sum_{C(i,j;m)} \frac{j}{2^{i+j}},$$

where $C(i, j; m)$ represents the following three constraints: $i + J - j = m$, $1 \leq j \leq J-1$, and $i \geq 1$. Note that $N \geq J$ by the definition of J. Thus for every $0 < \epsilon \leq 1/2$ we have

$$\frac{1}{2^N} \sum_{j=1}^{J-1} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a) \leq \frac{C_3}{2^{\epsilon N}} \sum_{m=1}^{\infty} \mathrm{SD}(f; \zeta, 2^m a) \sum_{C(i,j;m)} \frac{j}{2^{i+j+(1-\epsilon)J}}$$

$$\leq \frac{C_3}{2^{\epsilon N}} \sum_{m=1}^{\infty} \frac{1}{2^{(1-\epsilon)m}} \mathrm{SD}(f; \zeta, 2^m a) \sum_{C(i,j;m)} \frac{j}{2^{2(1-\epsilon)j}}. \tag{2.7}$$

Obviously, for each $m \geq 1$, if $i, i' \in \mathbf{N}$ and $j, j' \in \{1, \ldots, J-1\}$ satisfy the equations $i + J - j = m$ and $i' + J - j' = m$, and if $i \neq i'$, then $j \neq j'$. Since we assume $0 < \epsilon \leq 1/2$, it now follows that

$$\sum_{C(i,j;m)} \frac{j}{2^{2(1-\epsilon)j}} \leq \sum_{C(i,j;m)} \frac{j}{2^j} \leq \sum_{j=1}^{\infty} \frac{j}{2^j} = C_4.$$

Substituting this in (2.7), we find that

$$\frac{1}{2^N} \sum_{j=1}^{J-1} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a) \leq \frac{C_3 C_4}{2^{\epsilon N}} \sum_{m=1}^{\infty} \frac{1}{2^{(1-\epsilon)m}} \mathrm{SD}(f; \zeta, 2^m a).$$

Now if we let $C_5 = C_3 C_4 + C_{2.2}$ and if we take (2.6) into account, we see that

$$\frac{1}{2^N} \sum_{j=1}^{J} \frac{j}{2^j} \mathrm{SD}(f \circ \varphi_z; \zeta, 2^j a) \leq \frac{C_5}{2^{\epsilon N}} \sum_{m=1}^{\infty} \frac{1}{2^{(1-\epsilon)m}} \mathrm{SD}(f; \zeta, 2^m a).$$

Recalling (2.4) and (2.5), this completes the proof of the lemma in the case $2^N a > 4$. $\square$

Lemma 2.5. [5,Lemma 3.4] *Let $f \in L^2(S, d\sigma)$ and write $g = f - Pf$. Then for every $z \in \mathbf{B}\backslash\{0\}$ we have $H_f k_z = v_z k_z$, where $v_z = g - (P(g \circ \varphi_z)) \circ \varphi_z$.*

Lemma 2.6. [5,Lemma 5.1] *For each $k \geq 0$, there is a $C_{2.6}(k)$ which depends only on k and n such that*

$$\mathrm{SD}(v_z; z/|z|, 2^k(1 - |z|^2)^{1/2}) \leq C_{2.6}(k) \|M_{m_z} H_f k_z\|$$

for all $f \in L^2(S, d\sigma)$ and $z \in \mathbf{B}\backslash\{0\}$, where the relation between f and v_z is the same as in Lemma 2.5.

Lemma 2.7. *For each $N \in \mathbf{N}$, there exists a constant $C(N)$ which depends only on N and the complex dimension n such that if $f \in L^2(S, d\sigma)$ and $g = f - Pf$, then*

$$\mathrm{SD}(g; z/|z|, (1-|z|^2)^{1/2}) \leq C(N)\|M_{m_z} H_f k_z\| + \frac{C_{2.7}}{2^{\epsilon N}} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{SD}(g; z/|z|, 2^j(1-|z|^2)^{1/2})$$

for all $z \in \mathbf{B}\backslash\{0\}$ and $0 < \epsilon \leq 1/2$, where $C_{2.7} = (1 + C_{2.1})(1 + C_{2.4})$ and $C_{2.1}$, $C_{2.4}$ are the constants provided by Proposition 2.1 *and Lemma* 2.4 *respectively.*

Proof. Let $f \in L^2(S, d\sigma)$ and write $g = f - Pf$. Given a $z \in \mathbf{B}\backslash\{0\}$, we set $h_z = (P(g \circ \varphi_z)) \circ \varphi_z$ and $v_z = g - (P(g \circ \varphi_z)) \circ \varphi_z$. Since $g = v_z + h_z$ and $Pg = 0$, we have $h_z = -Pv_z$. Therefore

$$\begin{aligned}\mathrm{SD}(g; z/|z|, (1-|z|^2)^{1/2}) &\leq \mathrm{SD}(v_z; z/|z|, (1-|z|^2)^{1/2}) + \mathrm{SD}(Pv_z; z/|z|, (1-|z|^2)^{1/2}) \\ &\leq (1 + C_{2.1}) \sum_{k=0}^{\infty} \frac{1}{2^k} \mathrm{SD}(v_z; z/|z|, 2^k(1-|z|^2)^{1/2}),\end{aligned}$$

where the second $\leq$ follows from Proposition 2.1. Thus for each $N \in \mathbf{N}$ we have

$$\mathrm{SD}(g; z/|z|, (1-|z|^2)^{1/2}) \leq (1 + C_{2.1})(R_N + S_N + T_N),$$

where

$$R_N = \sum_{k=0}^{N-1} \frac{1}{2^k} \mathrm{SD}(v_z; z/|z|, 2^k(1-|z|^2)^{1/2}),$$
$$S_N = \sum_{k=N}^{\infty} \frac{1}{2^k} \mathrm{SD}(g; z/|z|, 2^k(1-|z|^2)^{1/2}),$$
$$T_N = \sum_{k=N}^{\infty} \frac{1}{2^k} \mathrm{SD}((P(g \circ \varphi_z)) \circ \varphi_z; z/|z|, 2^k(1-|z|^2)^{1/2}).$$

Applying Lemma 2.4 to T_N, we see that

$$S_N + T_N \leq \frac{1 + C_{2.4}}{2^{\epsilon N}} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{SD}(g; z/|z|, 2^j(1-|z|^2)^{1/2})$$

for every $0 < \epsilon \leq 1/2$. By Lemma 2.6, $R_N \leq M(N)\|M_{m_z} H_f k_z\|$, where $M(N)$ depends only on N and n. This completes the proof. $\square$

Lemma 2.8. *There exists a constant $C_{2.8}$ which depends only on n such that*

$$\mathrm{SD}(f; z/|z|, 2^j(1-|z|^2)^{1/2}) \leq C_{2.8} \mathrm{Var}^{1/2}(f; \rho_j(z))$$

for all $f \in L^2(S, d\sigma)$, $z \in \mathbf{B}\backslash\{0\}$ and $j \in \mathbf{N}$, where ρ_j is defined by (1.2).

Proof. It is easy to see that there is a C_1 which depends only on n such that

$$\mathrm{SD}(f; z/|z|, (1-|z|^2)^{1/2}) \leq C_1 \mathrm{Var}^{1/2}(f; z)$$

for all $f \in L^2(S, d\sigma)$ and $z \in \mathbf{B}\backslash\{0\}$. Thus in the case where $\rho_j(z) \neq 0$, the lemma follows from the above inequality and (1.3). On the other hand, if $\rho_j(z) = 0$, then by (1.2) we have $2^j(1-|z|^2)^{1/2} \geq 1$. Hence, recalling (2.2), in the case $\rho_j(z) = 0$ we have

$$\mathrm{SD}(f; z/|z|, 2^j(1-|z|^2)^{1/2}) \leq C_2 \mathrm{SD}(f; z/|z|, 2) = C_2 \mathrm{Var}^{1/2}(f; 0) = C_2 \mathrm{Var}^{1/2}(f; \rho_j(z)).$$

This completes the proof. $\square$

Corollary 2.9. *For each $N \in \mathbf{N}$, there exists a constant $C(N)$ which depends only on N and n such that if $f \in L^2(S, d\sigma)$ and $g = f - Pf$, then*

$$\mathrm{SD}(g; z/|z|, (1-|z|^2)^{1/2}) \leq C(N)\|M_{m_z} H_f k_z\| + \frac{C_{2.9}}{2^{\epsilon N}} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{Var}^{1/2}(g; \rho_j(z))$$

for all $z \in \mathbf{B}\backslash\{0\}$ and $0 < \epsilon \le 1/2$, where $C_{2.9} = C_{2.7}C_{2.8}$ and $C_{2.7}$, $C_{2.8}$ are the constants provided by Lemmas 2.7 and 2.8 respectively.

Proof. This follows immediately from Lemmas 2.7 and 2.8. □

Lemma 2.10. *There is a constant $C_{2.10}$ which depends only on n such that if $r \ge 1/2$, then the inequality*

$$\mathrm{SD}(f;\zeta,r) \le C_{2.10}\mathrm{Var}^{1/2}(f;0)$$

holds for all $f \in L^2(S,d\sigma)$ and $\zeta \in S$.

The proof of Lemma 2.10 is trivial and will be omitted.

Lemma 2.11. *Let $\gamma > 0$ and $0 < \epsilon \le 1/4$. There exists a constant $C_{2.11}(\gamma,\epsilon)$ which depends only on γ, ϵ and n such that if $f \in L^2(S,d\sigma)$ and $g = f - Pf$, then*

$$\mathrm{Var}^{1/2}(g;z) \le C_{2.11}(\gamma,\epsilon)\sum_{k=1}^{\infty}\frac{1}{2^k}\|M_{m_{\rho_k(z)}}H_f k_{\rho_k(z)}\| + \gamma\sum_{\ell=1}^{\infty}\frac{1}{2^{(1-2\epsilon)\ell}}\mathrm{Var}^{1/2}(g;\rho_\ell(z)), \tag{2.8}$$

for every $z \in \mathbf{B}$, where ρ_ℓ is the radial contraction defined by (1.2).

Proof. Let $f \in L^2(S,d\sigma)$ be given and write $g = f - Pf$. Let $z \in \mathbf{B}$ also be given.

(1) First we consider the case where $4(1-|z|^2) < 1$. For such a $z \in \mathbf{B}$, let $K(z)$ be the smallest natural number such that $4^{K(z)+1}(1-|z|^2) \ge 1$. Then, of course,

$$4^{K(z)}(1-|z|^2) < 1. \tag{2.9}$$

Applying first Lemma 2.2 and then Lemma 2.10, we have

$$\begin{aligned}
\mathrm{Var}^{1/2}(g;z) &\le C_{2.2}\sum_{k=1}^{\infty}\frac{1}{2^k}\mathrm{SD}(g;z/|z|,2^k(1-|z|^2)^{1/2}) \\
&= C_{2.2}\sum_{k=1}^{K(z)}\frac{1}{2^k}\mathrm{SD}(g;z/|z|,2^k(1-|z|^2)^{1/2}) \\
&\quad + C_{2.2}\sum_{k=K(z)+1}^{\infty}\frac{1}{2^k}\mathrm{SD}(g;z/|z|,2^k(1-|z|^2)^{1/2}) \\
&\le C_{2.2}\sum_{k=1}^{K(z)}\frac{1}{2^k}\mathrm{SD}(g;z/|z|,2^k(1-|z|^2)^{1/2}) + C_{2.2}\sum_{k=K(z)+1}^{\infty}\frac{C_{2.10}}{2^k}\mathrm{Var}^{1/2}(g;0) \\
&= C_{2.2}A + C_{2.2}C_{2.10}B, \qquad (2.10)
\end{aligned}$$

where

$$A = \sum_{k=1}^{K(z)}\frac{1}{2^k}\mathrm{SD}(g;z/|z|,2^k(1-|z|^2)^{1/2}) \quad \text{and} \quad B = \frac{1}{2^{K(z)}}\mathrm{Var}^{1/2}(g;0).$$

We estimate A and B separately.

For A, note that (2.9) ensures that $4^k(1-|z|^2) < 1$ for every $1 \le k \le K(z)$. Thus, by (1.2), we have

$$A = \sum_{k=1}^{K(z)} \frac{1}{2^k} \mathrm{SD}(g; \rho_k(z)/|\rho_k(z)|, (1-|\rho_k(z)|^2)^{1/2}).$$

Applying Corollary 2.9 to each term on the right-hand side, for $N \in \mathbf{N}$ and $0 < \epsilon \le 1/4$ we have

$$A \le C(N) \sum_{k=1}^{\infty} \frac{1}{2^k} \|M_{m_{\rho_k(z)}} H_f k_{\rho_k(z)}\| + \frac{C_{2.9}}{2^{\epsilon N}} \sum_{k=1}^{\infty} \frac{1}{2^k} \sum_{j=1}^{\infty} \frac{j}{2^{(1-\epsilon)j}} \mathrm{Var}^{1/2}(g; \rho_j(\rho_k(z))).$$

From (1.3) we see that $\rho_j(\rho_k(z)) = \rho_{j+k}(z)$. Therefore

$$\begin{aligned} A &\le C(N) \sum_{k=1}^{\infty} \frac{1}{2^k} \|M_{m_{\rho_k(z)}} H_f k_{\rho_k(z)}\| + \frac{C_{2.9}}{2^{\epsilon N}} \sum_{\ell=2}^{\infty} \mathrm{Var}^{1/2}(g; \rho_\ell(z)) \sum_{j+k=\ell} \frac{1}{2^k} \cdot \frac{j}{2^{(1-\epsilon)j}} \\ &\le C(N) \sum_{k=1}^{\infty} \frac{1}{2^k} \|M_{m_{\rho_k(z)}} H_f k_{\rho_k(z)}\| + \frac{C_2(\epsilon)}{2^{\epsilon N}} \sum_{\ell=1}^{\infty} \frac{1}{2^{(1-2\epsilon)\ell}} \mathrm{Var}^{1/2}(g; \rho_\ell(z)), \end{aligned}$$

where $C_2(\epsilon) = C_{2.9} \sup_{\ell \ge 1} 2^{-\epsilon \ell} \ell^2$. Let $\gamma > 0$ also be given. Then we set N to be the smallest natural number such that $C_2(\epsilon)/2^{\epsilon N} \le \gamma/C_{2.2}$. This determines N in terms of γ, ϵ and allows us to write $C(N) = C_1(\gamma, \epsilon)$. Hence the above becomes

$$A \le C_1(\gamma, \epsilon) \sum_{k=1}^{\infty} \frac{1}{2^k} \|M_{m_{\rho_k(z)}} H_f k_{\rho_k(z)}\| + \frac{\gamma}{C_{2.2}} \sum_{\ell=1}^{\infty} \frac{1}{2^{(1-2\epsilon)\ell}} \mathrm{Var}^{1/2}(g; \rho_\ell(z)). \tag{2.11}$$

For B, notice that since $g = f - Pf$, we have

$$\mathrm{Var}^{1/2}(g; 0) = \|g\| = \|H_f 1\| = \|H_f k_0\| = \|M_{m_{\rho_{K(z)+1}(z)}} H_f k_{\rho_{K(z)+1}(z)}\|,$$

where the last $=$ is due to the fact $\rho_{K(z)+1}(z) = 0$, which follows from (1.2) and the inequality $4^{K(z)+1}(1-|z|^2) \ge 1$. Thus

$$B = \frac{1}{2^{K(z)}} \mathrm{Var}^{1/2}(g; 0) = \frac{2}{2^{K(z)+1}} \|M_{m_{\rho_{K(z)+1}(z)}} H_f k_{\rho_{K(z)+1}(z)}\|. \tag{2.12}$$

Now if we combine (2.10) with (2.11) and (2.12), we see that (2.8) holds for the constant $C_{2.11}(\gamma, \epsilon) = C_{2.2} C_1(\gamma, \epsilon) + 2C_{2.2} C_{2.10}$ in the case $4(1-|z|^2) < 1$.

(2) Suppose that $4(1-|z|^2) \ge 1$. Then $|z|^2 \le 3/4$, and consequently $\|k_z g\| \le C_3 \|g\|$. Also, the condition $4(1-|z|^2) \ge 1$ implies $\rho_1(z) = 0$. Therefore in this case we have

$$\mathrm{Var}^{1/2}(g; z) \le \|g k_z\| \le C_3 \|g\| = C_3 \|H_f 1\| = C_3 \|H_f k_0\| = C_3 \|M_{m_{\rho_1(z)}} H_f k_{\rho_1(z)}\|.$$

This proves (2.8) in the case $4(1-|z|^2) \geq 1$. $\square$

Proof of Theorem 1.1. Let $0 < \delta \leq 1/2$ be given and set $\epsilon = \delta/3$. Write $A(\epsilon) = \sum_{m=1}^{\infty} 2^{-\epsilon m}$. Let $\gamma > 0$ be such that $\gamma A(\epsilon) \leq 1$. Let $f \in L^2(S, d\sigma)$ and denote

$$g = f - Pf.$$

In addition to the ρ_j defined by (1.2) in the case $j \in \mathbf{N}$, we also define $\rho_0(z) = z$ for every $z \in \mathbf{B}$. Obviously, we still have $\rho_k \circ \rho_j = \rho_{k+j}$ for all $k, j \geq 0$. Let $z \in \mathbf{B}$ be given. Then we can apply Lemma 2.11 to every $\mathrm{Var}^{1/2}(g; \rho_j(z))$, $j \geq 0$. This gives us

$$\begin{aligned}
\sum_{j=0}^{\infty} \frac{1}{2^{(1-3\epsilon)j}} \mathrm{Var}^{1/2}(g; \rho_j(z)) &\leq C_{2.11}(\gamma,\epsilon) \sum_{j=0}^{\infty} \frac{1}{2^{(1-3\epsilon)j}} \sum_{k=1}^{\infty} \frac{1}{2^k} \|M_{m_{\rho_k(\rho_j(z))}} H_f k_{\rho_k(\rho_j(z))}\| \\
&\quad + \gamma \sum_{j=0}^{\infty} \frac{1}{2^{(1-3\epsilon)j}} \sum_{\ell=1}^{\infty} \frac{1}{2^{(1-2\epsilon)\ell}} \mathrm{Var}^{1/2}(g; \rho_\ell(\rho_j(z))) \\
&= C_{2.11}(\gamma,\epsilon) \sum_{\nu=1}^{\infty} \|M_{m_{\rho_\nu(z)}} H_f k_{\rho_\nu(z)}\| \sum_{j+k=\nu} \frac{1}{2^{(1-3\epsilon)j} 2^k} \\
&\quad + \gamma \sum_{\nu=1}^{\infty} \mathrm{Var}^{1/2}(g; \rho_\nu(z)) \sum_{j+\ell=\nu} \frac{1}{2^{(1-3\epsilon)j} 2^{(1-2\epsilon)\ell}} \\
&\leq C_{2.11}(\gamma,\epsilon) A(\epsilon) \sum_{\nu=1}^{\infty} \frac{1}{2^{(1-3\epsilon)\nu}} \|M_{m_{\rho_\nu(z)}} H_f k_{\rho_\nu(z)}\| \\
&\quad + \gamma A(\epsilon) \sum_{\nu=1}^{\infty} \frac{1}{2^{(1-3\epsilon)\nu}} \mathrm{Var}^{1/2}(g; \rho_\nu(z)). \qquad (2.13)
\end{aligned}$$

Since $\mathrm{Var}^{1/2}(g; \rho_\nu(z)) \leq \|g k_{\rho_\nu(z)}\| \leq \|k_{\rho_\nu(z)}\|_\infty \|g\|$ and $|\rho_\nu(z)| \leq |z|$, it follows that

$$\sum_{\nu=1}^{\infty} \frac{1}{2^{(1-3\epsilon)\nu}} \mathrm{Var}^{1/2}(g; \rho_\nu(z)) < \infty.$$

Since $\gamma A(\epsilon) \leq 1$, we can cancel out $\sum_{\nu=1}^{\infty} 2^{-(1-3\epsilon)\nu} \mathrm{Var}^{1/2}(g; \rho_\nu(z))$ from both sides of (2.13) to obtain

$$\mathrm{Var}^{1/2}(g; z) = \mathrm{Var}^{1/2}(g; \rho_0(z)) \leq C_{2.11}(\gamma,\epsilon) A(\epsilon) \sum_{\nu=1}^{\infty} \frac{1}{2^{(1-3\epsilon)\nu}} \|M_{m_{\rho_\nu(z)}} H_f k_{\rho_\nu(z)}\|.$$

Since $3\epsilon = \delta$, this completes the proof of Theorem 1.1. $\square$

3. Partial sampling

Our next goal is to prove Theorem 1.3. For this we need the modified kernel function $\psi_{z,t}$ introduced in [5], whose definition we now recall. For each pair of $0 < t < \infty$ and $z \in \mathbf{B}$, we define

$$\psi_{z,t}(\zeta) = \frac{(1-|z|^2)^{(n/2)+t}}{(1-\langle \zeta, z\rangle)^{n+t}}, \tag{3.1}$$

$|\zeta| \le 1$. In terms of the Schur multiplier m_z defined by (1.4) and the normalized reproducing kernel k_z, we have the relation

$$\psi_{z,t} = (1+|z|)^t m_z^t k_z.$$

We think of $\psi_{z,t}$ as a modified version of k_z. This modification improves the "decaying rate" of the kernel, as is shown in the next proposition.

Proposition 3.1. *Given any positive number $0 < t < \infty$, there is a constant $C_{3.1}(t)$ that depends only on t and n such that the inequality*

$$|\langle \psi_{z,t}, \psi_{w,t}\rangle| \le C_{3.1}(t)\left(\frac{(1-|z|^2)^{1/2}(1-|w|^2)^{1/2}}{|1-\langle w, z\rangle|}\right)^{n+t}$$

holds for all $z, w \in \mathbf{B}$.

Proof. Let $0 < t < \infty$. By [14,Proposition 1.4.10], there is a constant $C(t)$ such that

$$\int \frac{d\sigma(\zeta)}{|1-\langle \zeta, \gamma\rangle|^{n+t}} \le \frac{C(t)}{(1-|\gamma|^2)^t} \tag{3.2}$$

for every $\gamma \in \mathbf{B}$. Let $z, w \in \mathbf{B}$ be given. The key idea is to express $\langle w, z\rangle$ in the form

$$\langle w, z\rangle = v^2 \tag{3.3}$$

for some $v \in \mathbf{C}$ with $|v| < 1$. On the open unit disc $\{u \in \mathbf{C} : |u| < 1\}$, we have the power series expansion

$$\frac{1}{(1-u)^{n+t}} = \sum_{j=0}^{\infty} b_j u^j.$$

If $j \ne k$, then $\langle \zeta, z\rangle^j$ and $\langle \zeta, w\rangle^k$ are orthogonal to each other in $L^2(S, d\sigma)$. Therefore

$$\int \frac{d\sigma(\zeta)}{(1-\langle \zeta, z\rangle)^{n+t}(1-\langle w, \zeta\rangle)^{n+t}} = \sum_{j=0}^{\infty} b_j^2 \int \langle \zeta, z\rangle^j \langle w, \zeta\rangle^j d\sigma(\zeta).$$

By an obvious change of variables or by a direct calculation using [14,Proposition 1.4.9],

$$\int \langle \zeta, z\rangle^j \langle w, \zeta\rangle^j d\sigma(\zeta) = \langle w, z\rangle^j \int |\zeta_1|^{2j} d\sigma(\zeta),$$

where ζ_1 denotes the first component of ζ. Using (3.3), we now have

$$\int \frac{d\sigma(\zeta)}{(1-\langle\zeta,z\rangle)^{n+t}(1-\langle w,\zeta\rangle)^{n+t}} = \sum_{j=0}^{\infty} b_j^2 \int (v\zeta_1)^j (v\bar{\zeta}_1)^j d\sigma(\zeta)$$
$$= \int \frac{d\sigma(\zeta)}{(1-v\zeta_1)^{n+t}(1-v\bar{\zeta}_1)^{n+t}},$$

consequently

$$\left|\int \frac{d\sigma(\zeta)}{(1-\langle\zeta,z\rangle)^{n+t}(1-\langle w,\zeta\rangle)^{n+t}}\right| \le \int \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}|1-v\bar{\zeta}_1|^{n+t}}. \tag{3.4}$$

It is elementary that if c is a complex number with $|c| \le 1$ and if $0 \le r \le 1$, then

$$2|1-rc| \ge |1-c|.$$

By (3.3) and this inequality, we have

$$|1-\langle w,z\rangle| = |1-v^2| \le 2|1-v\zeta_1 \cdot v\bar{\zeta}_1| \le 2|1-v\zeta_1| + 2|1-v\bar{\zeta}_1|.$$

Thus if we set

$$A = \{\zeta \in S : |1-v\zeta_1| \ge (1/4)|1-\langle w,z\rangle|\} \quad \text{and}$$
$$B = \{\zeta \in S : |1-v\bar{\zeta}_1| \ge (1/4)|1-\langle w,z\rangle|\},$$

then $A \cup B = S$. Hence it follows from (3.4) that

$$\left|\int \frac{d\sigma(\zeta)}{(1-\langle\zeta,z\rangle)^{n+t}(1-\langle w,\zeta\rangle)^{n+t}}\right|$$
$$\le \frac{4^{n+t}}{|1-\langle w,z\rangle|^{n+t}} \left(\int_A \frac{d\sigma(\zeta)}{|1-v\bar{\zeta}_1|^{n+t}} + \int_B \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}}\right).$$

Applying (3.2) to the vectors $(v,0,\dots,0)$ and $(\bar{v},0,\dots,0)$ in $\mathbf{B}$, we have

$$\int \frac{d\sigma(\zeta)}{|1-v\bar{\zeta}_1|^{n+t}} \le \frac{C(t)}{(1-|v|^2)^t} \quad \text{and} \quad \int \frac{d\sigma(\zeta)}{|1-v\zeta_1|^{n+t}} \le \frac{C(t)}{(1-|v|^2)^t}.$$

But $|v|^2 = |\langle w,z\rangle| \le |w||z|$. Therefore $1-|v|^2 \ge (1/2)(1-|z|^2)$ and $1-|v|^2 \ge (1/2)(1-|w|^2)$. Combining the above, we find that

$$\left|\int \frac{d\sigma(\zeta)}{(1-\langle\zeta,z\rangle)^{n+t}(1-\langle w,\zeta\rangle)^{n+t}}\right| \le \frac{4^{n+t}2^{t+1}C(t)}{|1-\langle w,z\rangle|^{n+t}(1-|z|^2)^{t/2}(1-|w|^2)^{t/2}}.$$

Recalling the definition (3.1) of $\psi_{z,t}$ and $\psi_{w,t}$, the proof is complete. □

In [5], the modified kernel functions $\psi_{z,t}$ were used to construct a certain "quasi-resolution" of the identity operator. In this paper, $\psi_{z,t}$ will be used differently; we will use these functions to construct certain "partial sampling" operators. We will see that the estimate provided by Proposition 3.1 leads to a uniform bound for these partial sampling operators, which is a crucial step in the proof of the lower bound in Theorem 1.3. But first we need to define these operators, which involves a decomposition of the unit ball.

For each integer $k \geq 0$, let $\{u_{k,1}, \ldots, u_{k,m(k)}\}$ be a subset of S which is *maximal* with respect to the property

$$B(u_{k,j}, 2^{-k-1}) \cap B(u_{k,j'}, 2^{-k-1}) = \emptyset \quad \text{for all} \;\; 1 \leq j < j' \leq m(k). \tag{3.5}$$

The maximality of $\{u_{k,1}, \ldots, u_{k,m(k)}\}$ implies that

$$\cup_{j=1}^{m(k)} B(u_{k,j}, 2^{-k}) = S. \tag{3.6}$$

For each pair of $k \geq 0$ and $1 \leq j \leq m(k)$, define

$$T_{k,j} = \{ru : 1 - 2^{-2k} \leq r^2 < 1 - 2^{-2(k+1)}, u \in B(u_{k,j}, 2^{-k})\}. \tag{3.7}$$

We also define the index set

$$I = \{(k,j) : k \geq 0, 1 \leq j \leq m(k)\}.$$

Definition 3.2. (a) A *partial sampling* set is a finite subset F of the open unit ball $\mathbf{B}$ with the property that $\mathrm{card}(F \cap T_{k,j}) \leq 1$ for every $(k,j) \in I$.
(b) For any partial sampling set F and any $t > 0$, denote

$$R_F^{(t)} = \sum_{z \in F} \psi_{z,t} \otimes \psi_{z,t}.$$

Proposition 3.3. *For each $t > 0$, there is a constant $C_{3.3}(t)$ such that $\|R_F^{(t)}\| \leq C_{3.3}(t)$ for every partial sampling set F.*

To prove this proposition, we need to recall the following counting lemma:

Lemma 3.4. [17,Lemma 4.1] *Let X be a set and let E be a subset of $X \times X$. Suppose that m is a natural number such that*

$$\mathrm{card}\{y \in X : (x,y) \in E\} \leq m \quad \text{and} \quad \mathrm{card}\{y \in X : (y,x) \in E\} \leq m$$

for every $x \in X$. Then there exist pairwise disjoint subsets E_1, E_2, ..., E_{2m} of E such that

$$E = E_1 \cup E_2 \cup \cdots \cup E_{2m}$$

and such that for each $1 \leq j \leq 2m$, the conditions $(x,y), (x',y') \in E_j$ and $(x,y) \neq (x',y')$ imply both $x \neq x'$ and $y \neq y'$.

Proof of Proposition 3.3. Let $t > 0$ and partial sampling set F be given. Then pick an arbitrary orthonormal set $\{e_z : z \in F\}$ indexed by F. We factor $R_F^{(t)}$ in the form

$$R_F^{(t)} = B^*B, \quad \text{where} \ \ B = \sum_{z\in F} e_z \otimes \psi_{z,t}.$$

Since $\|B^*B\| = \|BB^*\|$, it suffices to estimate the latter. Obviously,

$$BB^* = \sum_{z,w\in F} \langle \psi_{w,t}, \psi_{z,t}\rangle e_z \otimes e_w.$$

For each $k \geq 0$, define

$$F_k = F \cap (\cup_{j=1}^{m(k)} T_{k,j}) = \{z \in F : 1 - 2^{-2k} \leq |z|^2 < 1 - 2^{-2(k+1)}\}. \tag{3.8}$$

Then it is easy to see that

$$BB^* = A_0 + \sum_{\ell=1}^{\infty}(A_\ell + A_\ell^*), \tag{3.9}$$

where

$$A_\ell = \sum_{k=0}^{\infty} \sum_{w\in F_k} \sum_{z\in F_{k+\ell}} \langle \psi_{w,t}, \psi_{z,t}\rangle e_z \otimes e_w$$

for every $\ell \geq 0$. From (3.8) we see that for each $\ell \geq 0$, if $k \neq k'$, then we have both $\{e_w : w \in F_k\} \cap \{e_w : w \in F_{k'}\} = \emptyset$ and $\{e_z : z \in F_{k+\ell}\} \cap \{e_z : z \in F_{k'+\ell}\} = \emptyset$. Since $\{e_z : z \in F\}$ is an orthonormal set, it follows that

$$\|A_\ell\| = \sup_{k\geq 0} \|A_{\ell,k}\| \tag{3.10}$$

for every $\ell \geq 0$, where

$$A_{\ell,k} = \sum_{w\in F_k} \sum_{z\in F_{k+\ell}} \langle \psi_{w,t}, \psi_{z,t}\rangle e_z \otimes e_w,$$

$k \geq 0$. Obviously, our task is to estimate $\|A_{\ell,k}\|$. To do this, it is crucial that we group and re-enumerate the terms in the above sum properly.

Fix a pair of $k \geq 0$ and $\ell \geq 0$ for the moment. First of all, (3.6) tells us that for each $\gamma \in \{1, \ldots, m(k + \ell)\}$, we have $u_{k+\ell,\gamma} \in B(u_{k,i(\gamma)}, 2^{-k})$ for some $i(\gamma) \in \{1, \ldots, m(k)\}$. Hence there is a partition

$$\{1, \ldots, m(k + \ell)\} = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_{m(k)}$$

such that $\mathcal{I}_i \cap \mathcal{I}_{i'} = \emptyset$ if $i \neq i'$ and such that for every $i \in \{1, \ldots, m(k)\}$, we have

$$u_{k+\ell,\gamma} \in B(u_{k,i}, 2^{-k}) \quad \text{whenever} \ \ \gamma \in \mathcal{I}_i. \tag{3.11}$$

By (3.5) and (2.2), this implies

$$\operatorname{card}(\mathcal{I}_i) \le C_1 2^{2n\ell} \tag{3.12}$$

for every i, where the constant C_1 depends only on the complex dimension n. By (3.8), corresponding to the above is a partition

$$F_{k+\ell} = G_1 \cup \cdots \cup G_{m(k)}$$

such that $G_i \cap G_{i'} = \emptyset$ if $i \neq i'$ and such that for every $i \in \{1, \ldots, m(k)\}$, we have

$$G_i \subset \cup_{\gamma \in \mathcal{I}_i} T_{k+\ell,\gamma}. \tag{3.13}$$

From Definition 3.2(a) and (3.8), we see that there is a subset J of $\{1, \ldots, m(k)\}$ such that

$$F_k = \{w_j : j \in J\},$$

where $w_j \neq w_{j'}$ if $j \neq j'$, and such that $w_j \in T_{k,j}$ for every $j \in J$. Now, for each pair of $j \in J$ and $i \in \{1, \ldots, m(k)\}$, we define

$$\varphi_{j,i} = \sum_{z \in G_i} \langle \psi_{w_j,t}, \psi_{z,t} \rangle e_z, \tag{3.14}$$

keeping in mind the usual convention that summing over the empty set results in 0. Since $G_i \cap G_{i'} = \emptyset$ whenever $i \neq i'$ and since $\{e_z : z \in F\}$ is an orthonormal set, we have

$$\langle \varphi_{j,i}, \varphi_{j',i'} \rangle = 0 \quad \text{whenever} \quad i \neq i', \tag{3.15}$$

regardless of what j and j' may be.

By the preceding paragraph, we can rewrite $A_{\ell,k}$ as

$$A_{\ell,k} = \sum_{j \in J} \sum_{i=1}^{m(k)} \varphi_{j,i} \otimes e_{w_j}.$$

We need to further decompose $A_{\ell,k}$ according to the relation between j and i. Define

$$E_1 = \{(j,i) : j \in J, 1 \le i \le m(k), d(u_{k,j}, u_{k,i}) \le 2^{-k+2}\} \quad \text{and}$$
$$E_q = \{(j,i) : j \in J, 1 \le i \le m(k), 2^{-k+q} < d(u_{k,j}, u_{k,i}) \le 2^{-k+q+1}\} \quad \text{for} \quad q \ge 2.$$

Then

$$A_{\ell,k} = \sum_{q=1}^{\infty} Y_q, \quad \text{where} \quad Y_q = \sum_{(j,i) \in E_q} \varphi_{j,i} \otimes e_{w_j}. \tag{3.16}$$

By (3.5) and (2.2), there is a C_2 that depends only on the complex dimension n such that

$$\mathrm{card}\{i : 1 \le i \le m(k), d(u_{k,j}, u_{k,i}) \le 2^{-k+q+1}\} \le C_2 2^{2nq}$$

for every $q \in \mathbf{N}$ and every $j \in \{1, \dots, m(k)\}$. Now we apply Lemma 3.4 to each E_q. By that lemma, for each $q \in \mathbf{N}$ we have a partition

$$E_q = E_{q,1} \cup \cdots \cup E_{q,2\alpha(q)}, \tag{3.17}$$

where $\alpha(q) \le C_2 2^{2nq}$, such that for every $1 \le \nu \le 2\alpha(q)$, the conditions $(j,i), (j',i') \in E_{q,\nu}$ and $(j,i) \ne (j',i')$ imply both $j \ne j'$ and $i \ne i'$.

According to (3.17), we can further decompose each Y_q in the form

$$Y_q = \sum_{\nu=1}^{2\alpha(q)} Y_{q,\nu}, \quad \text{where} \quad Y_{q,\nu} = \sum_{(j,i) \in E_{q,\nu}} \varphi_{j,i} \otimes e_{w_j}. \tag{3.18}$$

The property of $E_{q,\nu}$ ensures that both projections $(j,i) \mapsto j$ and $(j,i) \mapsto i$ are injective on $E_{q,\nu}$. Thus by (3.15) and the orthogonality $\langle e_{w_j}, e_{w_{j'}} \rangle = 0$ for $j \ne j'$, we have

$$\|Y_{q,\nu}\| = \sup_{(j,i) \in E_{q,\nu}} \|\varphi_{j,i}\| \tag{3.19}$$

for every pair of $q \in \mathbf{N}$ and $1 \le \nu \le 2\alpha(q)$.

Obviously, our next task is to estimate $\|\varphi_{j,i}\|$. Since $\{e_z : z \in F\}$ is an orthonormal set, it follows from (3.14) that

$$\|\varphi_{j,i}\| \le \{\mathrm{card}(G_i)\}^{1/2} \sup_{z \in G_i} |\langle \psi_{w_j,t}, \psi_{z,t} \rangle|.$$

By Definition 3.2(a) and (3.13), we have $\mathrm{card}(G_i) \le \mathrm{card}(\mathcal{I}_i)$. Recalling (3.12), this means

$$\|\varphi_{j,i}\| \le C_1^{1/2} 2^{n\ell} \sup_{z \in G_i} |\langle \psi_{w_j,t}, \psi_{z,t} \rangle|.$$

Applying Proposition 3.1, we further obtain

$$\|\varphi_{j,i}\| \le C_1^{1/2} C_{3.1}(t) 2^{n\ell} \sup_{z \in G_i} \left(\frac{(1-|z|^2)^{1/2}(1-|w_j|^2)^{1/2}}{|1-\langle z, w_j \rangle|} \right)^{n+t}. \tag{3.20}$$

The key observation at this point is to factor the fraction inside the $(\cdots)^{n+t}$ in the form

$$\frac{(1-|z|^2)^{1/2}(1-|w_j|^2)^{1/2}}{|1-\langle z, w_j \rangle|} = \left\{ \frac{1-|z|^2}{1-|w_j|^2} \right\}^{1/2} \cdot \frac{1-|w_j|^2}{|1-\langle z, w_j \rangle|}.$$

As we will see momentarily, the two factors on the right-hand side provide decay in the radial direction and the spherical direction respectively. Obviously, the decaying rates in these two directions are different. Much of the complicated decomposition above is dictated by this disparity between the two directions.

Since $G_i \subset F_{k+\ell}$, it follows from (3.7) that $1-|z|^2 \le 2^{-2(k+\ell)}$ for every $z \in G_i$. On the other hand, (3.7) tells us that $1-|w_j|^2 > 2^{-2k-2}$ for every $j \in J$. Thus $\{(1-|z|^2)/(1-|w_j|^2)\}^{1/2} \le 2^{-\ell+1}$ if $z \in G_i$, $i \in \{1,\dots,m(k)\}$, and $j \in J$. Bringing this into (3.20), after the obvious simplification we find that

$$\|\varphi_{j,i}\| \le 2^{n+t}C_1^{1/2}C_{3.1}(t)2^{-t\ell}\sup_{z\in G_i}\left(\frac{1-|w_j|^2}{|1-\langle z,w_j\rangle|}\right)^{n+t}. \tag{3.21}$$

Since $|\langle z,w_j\rangle| \le |w_j|$, this immediately gives us the simple estimate

$$\|\varphi_{j,i}\| \le C_3(t)2^{-t\ell} \quad \text{in the case } \ (j,i) \in E_{1,\nu},$$

where $C_3(t) = 4^{n+t}C_1^{1/2}C_{3.1}(t)$. By (3.19), we have $\|Y_{1,\nu}\| \le C_3(t)2^{-t\ell}$ for every $1 \le \nu \le 2\alpha(1)$. Since $\alpha(1) \le 2^{2n}C_2$, by (3.18) we have

$$\|Y_1\| \le 2^{2n+1}C_2C_3(t)2^{-t\ell}. \tag{3.22}$$

But to estimate $\|Y_{q,\nu}\|$ in the case $q \ge 2$, we must analyze $(1-|w_j|^2)/|1-\langle z,w_j\rangle|$.

Consider a (j,i) in some $E_{q,\nu}$ with $q \ge 2$, and suppose that $z \in G_i$. Then by (3.13) there is a $\gamma \in \mathcal{I}_i$ such that $z \in T_{k+\ell,\gamma}$. By (3.7), this means $z = |z|\zeta$ for some $\zeta \in B(u_{k+\ell,\gamma}, 2^{-(k+\ell)})$. On the other hand, we have $w_j \in T_{k,j}$ by the definition of w_j. Therefore, by (3.7), there is a $\xi \in B(u_{k,j}, 2^{-k})$ such that $w_j = |w_j|\xi$. Thus

$$2|1-\langle z,w_j\rangle| \ge |1-\langle \zeta,\xi\rangle| = d^2(\zeta,\xi).$$

Since $\gamma \in \mathcal{I}_i$, we have $u_{k+\ell,\gamma} \in B(u_{k,i}, 2^{-k})$ by (3.11). Since $(j,i) \in E_{q,\nu}$, $q \ge 2$, we have

$$\begin{aligned} 2^{-k+q} < d(u_{k,j},u_{k,i}) &\le d(u_{k,j},\xi) + d(\xi,\zeta) + d(\zeta,u_{k+\ell,\gamma}) + d(u_{k+\ell,\gamma},u_{k,i}) \\ &\le 2^{-k} + d(\xi,\zeta) + 2^{-(k+\ell)} + 2^{-k} \\ &\le d(\xi,\zeta) + (3/4)2^{-k+q}. \end{aligned}$$

Cancelling out $(3/4)2^{-k+q}$ from both sides, we find that $d(\zeta,\xi) \ge (1/4)2^{-k+q}$. Hence

$$|1-\langle z,w_j\rangle| \ge (1/32)2^{-2k+2q} = 2^{-2k+2q-5}.$$

By (3.7), the membership $w_j \in T_{k,j}$ also means $1-|w_j|^2 \le 2^{-2k}$. Thus we conclude that

$$\frac{1-|w_j|^2}{|1-\langle z,w_j\rangle|} \le 2^{-2q+5} \tag{3.23}$$

if $z \in G_i$, $(j, i) \in E_{q,\nu}$, and $q \geq 2$.

Substituting (3.23) in (3.21), we find that

$$\|\varphi_{j,i}\| \leq C_4(t) 2^{-t\ell} 2^{-2(n+t)q} \quad \text{if} \quad (j,i) \in E_{q,\nu}, \quad q \geq 2,$$

where $C_4(t) = 2^{6(n+t)} C_1^{1/2} C_{3.1}(t)$. Recalling (3.19), we obtain

$$\|Y_{q,\nu}\| \leq C_4(t) 2^{-t\ell} 2^{-2(n+t)q} \quad \text{for all} \quad q \geq 2 \quad \text{and} \quad 1 \leq \nu \leq 2\alpha(q).$$

Since $\alpha(q) \leq C_2 2^{2nq}$, this and (3.18) together give us

$$\|Y_q\| \leq C_4(t) 2^{-t\ell} 2^{-2(n+t)q} \cdot 2C_2 2^{2nq} = C_5(t) 2^{-t\ell} 2^{-2tq}$$

for $q \geq 2$, where $C_5(t) = 2C_2 C_4(t)$. Now set $C_6(t) = \max\{C_5(t), 2^{2n+1+2t} C_2 C_3(t)\}$. Combining the above inequality with (3.22) and (3.16), we conclude that

$$\|A_{\ell,k}\| \leq \sum_{q=1}^{\infty} \|Y_q\| \leq C_6(t) \sum_{q=1}^{\infty} 2^{-t\ell} 2^{-2tq} = C_7(t) 2^{-t\ell},$$

where $C_7(t) = C_6(t) \sum_{q=1}^{\infty} 2^{-2tq}$. By (3.10), this means

$$\|A_\ell\| \leq C_7(t) 2^{-t\ell}$$

for every $\ell \geq 0$. Finally, returning to (3.9), we reach the conclusion that

$$\|BB^*\| \leq \|A_0\| + 2 \sum_{\ell=1}^{\infty} \|A_\ell\| \leq 2C_7(t) \sum_{\ell=0}^{\infty} 2^{-t\ell} = C_{3.3}(t).$$

Since $\|BB^*\| = \|B^*B\|$ and $B^*B = R_F^{(t)}$, this completes the proof of the proposition. $\square$

4. Symmetric norm

The significance of Proposition 3.3 is that it brings the symmetric norm $\|(H_f^* H_f)^{p/2}\|_\Phi$ into our estimates:

Lemma 4.1. *Let $0 < t < \infty$ and $2 \leq p < \infty$. If $f \in L^2(S, d\sigma)$ and if H_f is bounded, then the inequality*

$$\Phi(\{\|H_f \psi_{z,t}\|^p\}_{z \in F}) \leq 2^{t(p-2)} C_{3.3}(t) \|(H_f^* H_f)^{p/2}\|_\Phi$$

holds for every symmetric gauge function Φ and every partial sampling set F, where $C_{3.3}(t)$ is the constant provided by Proposition 3.3.

Proof. Let Φ be a symmetric gauge function. Then it has the following property: For non-negative numbers $a_1 \geq \cdots \geq a_\nu \geq 0$ and $b_1 \geq \cdots \geq b_\nu \geq 0$ in descending order, if

$$a_1 + \cdots + a_j \leq b_1 + \cdots + b_j \quad \text{for every} \quad 1 \leq j \leq \nu,$$

then

$$\Phi(\{a_1,\dots,a_\nu,0,\dots,0,\dots\}) \le \Phi(\{b_1,\dots,b_\nu,0,\dots,0,\dots\}).$$

See Lemma III.3.1 in [7]. We will use this property to prove the lemma.

Let t, p, f and F be given as in the statement of the lemma. Suppose that $\mathrm{card}(F) = m$. Then we can enumerate the elements in F as $z_1,\dots,z_m$ in such a way that

$$\|H_f\psi_{z_1,t}\| \ge \|H_f\psi_{z_2,t}\| \ge \cdots \ge \|H_f\psi_{z_m,t}\|.$$

For each $1 \le k \le m$, let $F_k = \{z_1,\dots,z_k\}$. Being a subset of F, each F_k is, of course, a partial sampling set. Hence it follows from Proposition 3.3 that $\|R^{(t)}_{F_k}\| \le C_{3.3}(t)$. In terms of s-numbers, this implies that

$$s_j((H_f^*H_f)^{p/2}R^{(t)}_{F_k}) \le C_{3.3}(t)s_j((H_f^*H_f)^{p/2})$$

for every $j \ge 1$ (see page 61 in [7]). By Definition 3.2(b), $\mathrm{rank}(R^{(t)}_{F_k}) \le k$. Writing $\|\cdot\|_1$ for the norm of the trace class, we have

$$\begin{aligned}\mathrm{tr}((H_f^*H_f)^{p/2}R^{(t)}_{F_k}) &\le \|(H_f^*H_f)^{p/2}R^{(t)}_{F_k}\|_1\\ &= s_1((H_f^*H_f)^{p/2}R^{(t)}_{F_k}) + \cdots + s_k((H_f^*H_f)^{p/2}R^{(t)}_{F_k})\\ &\le C_{3.3}(t)\{s_1((H_f^*H_f)^{p/2}) + \cdots + s_k((H_f^*H_f)^{p/2})\}.\end{aligned} \tag{4.1}$$

On the other hand,

$$\mathrm{tr}((H_f^*H_f)^{p/2}R^{(t)}_{F_k}) = \langle (H_f^*H_f)^{p/2}\psi_{z_1,t},\psi_{z_1,t}\rangle + \cdots + \langle (H_f^*H_f)^{p/2}\psi_{z_k,t},\psi_{z_k,t}\rangle. \tag{4.2}$$

Since $p/2 \ge 1$, it follows from the spectral decomposition of $(H_f^*H_f)^{p/2}$ and Hölder's inequality that

$$\begin{aligned}\|H_f\psi_{z_j,t}\|^p = \langle H_f^*H_f\psi_{z_j,t},\psi_{z_j,t}\rangle^{p/2} &\le \langle (H_f^*H_f)^{p/2}\psi_{z_j,t},\psi_{z_j,t}\rangle\|\psi_{z_j,t}\|^{p-2}\\ &\le 2^{t(p-2)}\langle (H_f^*H_f)^{p/2}\psi_{z_j,t},\psi_{z_j,t}\rangle.\end{aligned}$$

Combining this with (4.2) and (4.1), we obtain the inequality

$$\|H_f\psi_{z_1,t}\|^p + \cdots + \|H_f\psi_{z_k,t}\|^p \le 2^{t(p-2)}C_{3.3}(t)\{s_1((H_f^*H_f)^{p/2}) + \cdots + s_k((H_f^*H_f)^{p/2})\}$$

for every $1 \le k \le m$. By the first paragraph of the proof, this implies

$$\begin{aligned}\Phi(\{\|H_f\psi_{z,t}\|^p\}_{z\in F}) &\le 2^{t(p-2)}C_{3.3}(t)\Phi(\{s_j((H_f^*H_f)^{p/2})\}_{j\in\mathbf{N}})\\ &= 2^{t(p-2)}C_{3.3}(t)\|(H_f^*H_f)^{p/2}\|_\Phi.\end{aligned}$$

This completes the proof of the lemma. □

Having applied Proposition 3.3, our next step is to apply the local inequality (Theorem 1.1) in the proof of the lower bound in Theorem 1.3. But this requires some preparation because of the involvement of ρ_ℓ and the possible overlap between $T_{k,j}$ and $T_{k,j'}$.

Definition 4.2. Let $N \in \mathbf{N}$. An *N-partial sampling map* is a map φ from a finite set X into $\mathbf{B}$ that has the property $\mathrm{card}\{x \in X : \varphi(x) \in T_{k,j}\} \le N$ for every $(k,j) \in I$.

Recall that for each $\ell \in \mathbf{N}$, the radial contraction ρ_ℓ is defined by (1.2). In addition, as we did in the proof of Theorem 1.1, we define $\rho_0(z) = z$, $z \in \mathbf{B}$.

Lemma 4.3. *There exists a natural number $M_{4.3}$ determined by the complex dimension n such that the following holds true: Let I' be a finite subset of I and let $z : I' \to \mathbf{B}$ be a map such that $z(k,j) \in T_{k,j}$ for every $(k,j) \in I'$. Then for each $\ell \in \mathbf{Z}_+$ the map $\rho_\ell \circ z : I' \to \mathbf{B}$ is $M_{4.3}2^{2n\ell}$-partial sampling.*

Proof. By (3.5) and (2.2), there is a natural number M_1 such that

$$\mathrm{card}\{i : B(u_{k+\ell,i}, 2^{-k-\ell}) \cap B(u_{k,j}, 2^{-k}) \ne \emptyset, 1 \le i \le m(k+\ell)\} \le M_1 2^{2n\ell} \tag{4.3}$$

for all $\ell \in \mathbf{Z}_+$ and $(k,j) \in I$. Also by (3.5) and (2.2), there is a natural number M_2 such that $m(k) \le M_2 2^{2nk}$ for every $k \ge 0$. Consequently

$$m(0) + m(1) + \cdots + m(\ell) \le 2M_2 2^{2n\ell} \tag{4.4}$$

for every $\ell \ge 0$. Suppose that $z : I' \to \mathbf{B}$ is a map such that $z(k,j) \in T_{k,j}$ for every $(k,j) \in I'$ and suppose that $\ell \in \mathbf{Z}_+$. Let us estimate $\mathrm{card}\{(k',i) \in I' : \rho_\ell(z(k',i)) \in T_{k,j}\}$ for each $(k,j) \in I$.

(1) First we consider the case $k \ge 1$. Then $0 \notin T_{k,j}$. By (1.3) and (3.7), if $\rho_\ell(z(k',i)) \in T_{k,j}$, then

$$1 - |z(k',i)|^2 = 2^{-2\ell}(1 - |\rho_\ell(z(k',i))|^2) \in (2^{-2(k+\ell+1)}, 2^{-2(k+\ell)}].$$

Since it is assumed that $z(\kappa,\nu) \in T_{\kappa,\nu}$ for every $(\kappa,\nu) \in I'$, the above implies $k' = k + \ell$. Therefore

$$\mathrm{card}\{(k',i) \in I' : \rho_\ell(z(k',i)) \in T_{k,j}\} = \mathrm{card}\{i : (k+\ell,i) \in I', \rho_\ell(z(k+\ell,i)) \in T_{k,j}\}.$$

Since the membership $\rho_\ell(z(k+\ell,i)) \in T_{k,j}$ implies $\rho_\ell(z(k+\ell,i))/|\rho_\ell(z(k+\ell,i))| \in B(u_{k,j}, 2^{-k})$ and since $z(k+\ell,i)/|z(k+\ell,i)| = \rho_\ell(z(k+\ell,i))/|\rho_\ell(z(k+\ell,i))|$, we have

$$\mathrm{card}\{i : (k+\ell,i) \in I', \rho_\ell(z(k+\ell,i)) \in T_{k,j}\} \le \\ \mathrm{card}\{i : B(u_{k+\ell,i}, 2^{-k-\ell}) \cap B(u_{k,j}, 2^{-k}) \ne \emptyset, 1 \le i \le m(k+\ell)\}.$$

Recalling (4.3), we now have

$$\mathrm{card}\{(k',i) \in I' : \rho_\ell(z(k',i)) \in T_{k,j}\} \le M_1 2^{2n\ell} \tag{4.5}$$

in the case $k \geq 1$.

(2) Consider the case $k = 0$. That is, we need to estimate $\mathrm{card}\{(k', i) \in I' : \rho_\ell(z(k', i)) \in T_{0,j}\}$, $j \in \{1, \ldots, m(0)\}$. First of all, by (1.2) and by the argument in (1), we have

$$\mathrm{card}\{(k', i) \in I' : \rho_\ell(z(k', i)) \in T_{0,j}, \rho_\ell(z(k', i)) \neq 0\} \leq M_1 2^{2n\ell}$$

for every $j \in \{1, \ldots, m(0)\}$. Thus it remains to estimate $\mathrm{card}\{(k', i) \in I' : \rho_\ell(z(k', i)) = 0\}$. By (1.2), if $|w|^2 > 1 - 2^{-2\ell}$, then $\rho_\ell(w) \neq 0$. Hence if $(k', i) \in I'$ is such that $\rho_\ell(z(k', i)) = 0$, then $|z(k', i)|^2 \leq 1 - 2^{-2\ell} < 1 - 2^{-2(\ell+1)}$. Combining this with the assumption $z(\kappa, \nu) \in T_{\kappa,\nu}$ for every $(\kappa, \nu) \in I'$ and with (3.7), we conclude that the condition $\rho_\ell(z(k', i)) = 0$ implies $k' \leq \ell$. Thus by (4.4) we have

$$\mathrm{card}\{(k', i) \in I' : \rho_\ell(z(k', i)) = 0\} \leq m(0) + m(1) + \cdots + m(\ell) \leq 2M_2 2^{2n\ell}.$$

Therefore

$$\mathrm{card}\{(k', i) \in I' : \rho_\ell(z(k', i)) \in T_{0,j}\} \leq (M_1 + 2M_2)2^{2n\ell} \tag{4.6}$$

for every $j \in \{1, \ldots, m(0)\}$.

Finally, from (4.5) and (4.6) we see that if we set $M_{4.3} = M_1 + 2M_2$, then the map $\rho_\ell \circ z : I' \to \mathbf{B}$ is $M_{4.3} 2^{2n\ell}$-partial sampling. $\square$

Lemma 4.4. *There exists a natural number $M_{4.4}$ determined by the complex dimension n such that the following holds true: Suppose that $N \in \mathbf{N}$ and that $\varphi : X \to \mathbf{B}$ an N-partial sampling map as defined in Definition 4.2. Then there is a partition*

$$X = X_1 \cup \cdots \cup X_{M_{4.4}N}$$

such that for every $i \in \{1, \ldots, M_{4.4}N\}$, the map $\varphi : X_i \to \mathbf{B}$ is 1-partial sampling.

Proof. We use a standard maximality argument, as follows. By (3.5), (3.7) and (2.2), there is a natural number $M_{4.4}$ such that the inequality

$$\mathrm{card}\{\nu : T_{k,\nu} \cap T_{k,j} \neq \emptyset, 1 \leq \nu \leq m(k)\} \leq M_{4.4}$$

holds for every $(k, j) \in I$. Moreover, if $k \neq k'$, then $T_{k,j} \cap T_{k',j'} = \emptyset$ for all possible j and j'. Suppose that $\varphi : X \to \mathbf{B}$ is an N-partial sampling map. Then we define X_1 to be a subset of X that is maximal with respect to the property that the restricted map $\varphi : X_1 \to \mathbf{B}$ is 1-partial sampling. Suppose that $m \geq 1$ and that we have defined $X_1, \ldots, X_m$. Then we define X_{m+1} to be a subset of $X \backslash \{\cup_{i=1}^m X_i\}$ that is maximal with respect to the property that the restricted map $\varphi : X_{m+1} \to \mathbf{B}$ is 1-partial sampling. Inductively, this defines all X_i, $i \geq 1$. We need to show that

$$X_1 \cup \cdots \cup X_{M_{4.4}N} = X.$$

If not, then there would be some $x^* \in X \backslash \{\cup_{i=1}^{M_{4.4}N} X_i\}$. By the maximality of each X_i, the restricted map $\varphi : X_i \cup \{x^*\} \to \mathbf{B}$ must fail to be 1-partial sampling, $1 \le i \le M_{4.4}N$. That is, for each $i \in \{1, \dots, M_{4.4}N\}$, there would be an $x_i \in X_i$ and a $(k_i, j_i) \in I$ such that

$$T_{k_i,j_i} \supset \{\varphi(x_i), \varphi(x^*)\}.$$

Let $(k^*, j^*) \in I$ be such that $\varphi(x^*) \in T_{k^*,j^*}$. Then we have $k_i = k^*$ and $T_{k^*,j_i} \supset \{\varphi(x_i), \varphi(x^*)\}$ for every $i \in \{1, \dots, M_{4.4}N\}$. This in particular implies that $T_{k^*,j^*} \cap T_{k^*,j_i} \neq \emptyset$. Therefore

$$\operatorname{card}(\{j^*\} \cup \{j_i : 1 \le i \le M_{4.4}N\}) \le M_{4.4}.$$

Since φ is an N-partial sampling map, it follows from this inequality that the inverse image of the set $T_{k^*,j^*} \cup \{\cup_{i=1}^{M_{4.4}N} T_{k^*,j_i}\}$ under φ contains at most $M_{4.4}N$ elements. But the above also tells us that this inverse image contains $x_1, \dots, x_{M_{4.4}N}$ and x^*, $M_{4.4}N + 1$ distinct elements. This is a contradiction. □

For each $m \in \mathbf{N}$, define $I_m = \{(k, j) \in I : k \le m\}$.

Corollary 4.5. *Suppose that $z_{k,j} \in T_{k,j}$ for every $(k, j) \in I$. Let $0 < t < \infty$ and $2 \le p < \infty$. If $f \in L^2(S, d\sigma)$ and if H_f is bounded, then the inequality*

$$\Phi(\{\|H_f \psi_{z_{k,j},t}\|^p\}_{(k,j)\in I_m}) \le 2^{t(p-2)} C_{3.3}(t) M_{4.3} M_{4.4} \|(H_f^* H_f)^{p/2}\|_\Phi$$

holds for every symmetric gauge function Φ and every $m \ge 1$, where $C_{3.3}(t)$, $M_{4.3}$ and $M_{4.4}$ are the constants provided by Proposition 3.3 *and Lemmas* 4.3, 4.4 *respectively.*

Proof. First of all, a symmetric gauge function Φ has the following obvious property: If A is any countable set and if $A = A_1 \cup \dots \cup A_N$, then for every map $\varphi : A \to [0, \infty)$ we have

$$\Phi(\{\varphi(\alpha)\}_{\alpha\in A}) \le \Phi(\{\varphi(\alpha)\}_{\alpha\in A_1}) + \dots + \Phi(\{\varphi(\alpha)\}_{\alpha\in A_N}). \tag{4.7}$$

Let $m \ge 1$ be given and consider the map $(k, j) \mapsto z_{k,j}$ from I_m into $\mathbf{B}$. Since $z_{k,j} \in T_{k,j}$, if we apply Lemma 4.3 to the case $\ell = 0$, we see that this map is $M_{4.3}$-partial sampling. Therefore, by Lemma 4.4, there is a partition

$$I_m = E_1 \cup \dots \cup E_{M_{4.4}M_{4.3}}$$

such that for every $1 \le i \le M_{4.4}M_{4.3}$, the map $(k, j) \mapsto z_{k,j}$ is 1-partial sampling on E_i. That is, the map $(k, j) \mapsto z_{k,j}$ is injective on E_i and $\{z_{k,j} : (k, j) \in E_i\}$ is a partial sampling set as defined in Definition 3.2. Hence Lemma 4.1 gives us

$$\Phi(\{\|H_f \psi_{z_{k,j},t}\|^p\}_{(k,j)\in E_i}) \le 2^{t(p-2)} C_{3.3}(t) \|(H_f^* H_f)^{p/2}\|_\Phi$$

for every $1 \le i \le M_{4.4}M_{4.3}$. By (4.7), we also have

$$\Phi(\{\|H_f \psi_{z_{k,j},t}\|^p\}_{(k,j)\in I_m}) \le \sum_{i=1}^{M_{4.4}M_{4.3}} \Phi(\{\|H_f \psi_{z_{k,j},t}\|^p\}_{(k,j)\in E_i}).$$

Obviously, the corollary follows from the above two inequalities. □

Lemma 4.6. *Let $h : \mathbf{B} \to [0, \infty)$ be a map such that $\sup_{w \in T_{k,j}} h(w) < \infty$ for every $(k, j) \in I$. For each $(k, j) \in I$, let $w_{k,j} \in T_{k,j}$ be such that*

$$h(w_{k,j}) \geq \frac{1}{2} \sup_{w \in T_{k,j}} h(w). \tag{4.8}$$

Suppose that $z_{k,j} \in T_{k,j}$ for every $(k, j) \in I$. Then the inequality

$$\Phi(\{h(\rho_\ell(z_{k,j}))\}_{(k,j)\in I_m}) \leq 2M_{4.3}M_{4.4}2^{2n\ell}\Phi(\{h(w_{k,j})\}_{(k,j)\in I_m}) \tag{4.9}$$

holds for every pair of $m, \ell \in \mathbf{N}$ and every symmetric gauge function Φ, where $M_{4.3}$ and $M_{4.4}$ are the natural numbers provided by Lemmas 4.3 *and* 4.4 *respectively.*

Proof. Let $m, \ell \in \mathbf{N}$ be given. Define the map $\varphi : I_m \to \mathbf{B}$ by the formula $\varphi(k, j) = \rho_\ell(z_{k,j})$, $(k, j) \in I_m$. Then Lemma 4.3 tells us that φ is $M_{4.3}2^{2n\ell}$-partial sampling. By Lemma 4.4, there is a partition

$$I_m = I_{m,1} \cup \cdots \cup I_{m,M_{4.4}M_{4.3}2^{2n\ell}}$$

such that for every $1 \leq i \leq M_{4.4}M_{4.3}2^{2n\ell}$, the restricted map $\varphi : I_{m,i} \to \mathbf{B}$ is 1-partial sampling. By (1.2), we have $\varphi(I_{m,i}) \subset \cup_{(k,j)\in I_m} T_{k,j}$. Thus for every $(k, j) \in I$,

$$\begin{aligned}
&\operatorname{card}\{(\kappa, \nu) \in I_{m,i} : \varphi(\kappa, \nu) \in T_{k,j}\} \leq 1 \quad \text{if} \quad 0 \leq k \leq m, \\
&\operatorname{card}\{(\kappa, \nu) \in I_{m,i} : \varphi(\kappa, \nu) \in T_{k,j}\} = 0 \quad \text{if} \quad k > m.
\end{aligned}$$

This implies that for each $1 \leq i \leq M_{4.4}M_{4.3}2^{2n\ell}$, there exist a subset J_i of I_m and a bijection $\pi_i : J_i \to I_{m,i}$ such that $\varphi(\pi_i(k, j)) \in T_{k,j}$ for every $(k, j) \in J_i$. By (4.8), this implies $h(\varphi(\pi_i(k, j))) \leq 2h(w_{k,j})$ for every $(k, j) \in J_i$. Since π_i is a bijection, we have

$$\Phi(\{h(\varphi(\kappa, \nu))\}_{(\kappa,\nu)\in I_{m,i}}) = \Phi(\{h(\varphi(\pi_i(k, j)))\}_{(k,j)\in J_i}) \leq 2\Phi(\{h(w_{k,j})\}_{(k,j)\in I_m})$$

for every $1 \leq i \leq M_{4.4}M_{4.3}2^{2n\ell}$. By (4.7), we have

$$\begin{aligned}
\Phi(\{h(\varphi(\kappa, \nu))\}_{(\kappa,\nu)\in I_m}) &\leq \sum_{i=1}^{M_{4.4}M_{4.3}2^{2n\ell}} \Phi(\{h(\varphi(\kappa, \nu))\}_{(\kappa,\nu)\in I_{m,i}}) \\
&\leq 2M_{4.4}M_{4.3}2^{2n\ell}\Phi(\{h(w_{k,j})\}_{(k,j)\in I_m}).
\end{aligned}$$

Since $\varphi(\kappa, \nu) = \rho_\ell(z_{\kappa,\nu})$, this proves the lemma. □

With the above preparation, the local inequality in Theorem 1.1 can now be applied:

Proposition 4.7. *Given any $2n < p < \infty$, there is a constant $C_{4.7}(p)$ such that the following holds true: Let $f \in L^2(S, d\sigma)$. For each (k, j), let $w_{k,j} \in T_{k,j}$ be such that*

$$\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\| \geq \frac{1}{2} \sup_{w \in T_{k,j}} \|M_{m_w} H_f k_w\|. \tag{4.10}$$

Let $z_{k,j} \in T_{k,j}$, $(k,j) \in I$. Then we have

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j)\in I_m}) \le C_{4.7}(p)\Phi(\{\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\|^p\}_{(k,j)\in I_m})$$

for every symmetric gauge function Φ and every $m \in \mathbf{N}$.

Proof. Given a $2n < p < \infty$, we pick a $0 < \delta < 1/2$ such that $(1-2\delta)p > 2n$. This allows us to write $(1-2\delta)p = 2n + \epsilon$ with some $\epsilon > 0$. With this δ, Theorem 1.1 gives us

$$\mathrm{Var}^{1/2}(f - Pf; z_{k,j}) \le C(\delta)\sum_{\ell=1}^{\infty}\frac{1}{2^{(1-\delta)\ell}}\|M_{m_{\rho_\ell(z_{k,j})}} H_f k_{\rho_\ell(z_{k,j})}\|$$

for every $(k,j) \in I$. Now factor the $2^{-(1-\delta)\ell}$ above in the form $2^{-(1-\delta)\ell} = 2^{-(1-2\delta)\ell}\cdot 2^{-\delta\ell}$, apply Hölder's inequality to the sum $\sum_{\ell=1}^{\infty}$ with conjugate exponents p and $p/(p-1)$, and then raise both sides to the power p. The result of this is

$$\mathrm{Var}^{p/2}(f - Pf; z_{k,j}) \le C_1\sum_{\ell=1}^{\infty}\frac{1}{2^{(2n+\epsilon)\ell}}\|M_{m_{\rho_\ell(z_{k,j})}} H_f k_{\rho_\ell(z_{k,j})}\|^p,$$

where $C_1 = C^p(\delta)(2^{\delta p/(p-1)} - 1)^{-(p-1)}$. Since any symmetric gauge function Φ is a norm on $\hat{c}$, the above inequality implies

(4.11)

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j)\in I_m}) \le C_1\sum_{\ell=1}^{\infty}\frac{1}{2^{(2n+\epsilon)\ell}}\Phi(\{\|M_{m_{\rho_\ell(z_{k,j})}} H_f k_{\rho_\ell(z_{k,j})}\|^p\}_{(k,j)\in I_m}),$$

$m \ge 1$. Now consider the map $h : \mathbf{B} \to [0,\infty)$ defined by the formula

$$h(z) = \|M_{m_z} H_f k_z\|^p,$$

$z \in \mathbf{B}$. Then (4.10) means that h and $w_{k,j}$ satisfy condition (4.8). Now apply Lemma 4.6. For this particular h, (4.9) translates to

$$\Phi(\{\|M_{m_{\rho_\ell(z_{k,j})}} H_f k_{\rho_\ell(z_{k,j})}\|^p\}_{(k,j)\in I_m}) \le 2M_{4.3}M_{4.4}2^{2n\ell}\Phi(\{\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\|^p\}_{(k,j)\in I_m}),$$

$\ell \ge 1$. Substituting this inequality in (4.11), we see that the proposition holds for the constant $C_{4.7}(p) = 2M_{4.3}M_{4.4}C_1\sum_{\ell=1}^{\infty}2^{-\epsilon\ell}$. □

5. Lower bound

We need to bridge the gap between $\|M_{m_z} H_f k_z\|$ and $\|H_f \psi_{z,t}\|$. For this, we recall

Lemma 5.1. [5,Lemma 3.3] *There exists a constant $C_{5.1}$ which depends only on the complex dimension n such that the inequality $\|[P, M_{m_z^t}]\| \le C_{5.1}t$ holds for all $z \in \mathbf{B}$ and $t > 0$.*

Lemma 5.2. *For all* $f \in L^2(S, d\sigma)$, $z \in \mathbf{B}$ *and* $0 < t \le 1$, *we have*

$$\|M_{m_z} H_f k_z\| \le \|H_f \psi_{z,t}\| + C_{5.1} t \mathrm{Var}^{1/2}(f - Pf; z),$$

where $C_{5.1}$ *is the constant mentioned in Lemma* 5.1.

Proof. On page 3100 in [5], we proved that

$$\|M_{m_z} H_f k_z\| \le \|H_f \psi_{z,t}\| + C_{5.1} t \|H_f k_z\|.$$

But $\|H_f k_z\| \le \mathrm{Var}^{1/2}(f - Pf; z)$. To see this, set $\varphi = f - Pf - \langle (f - Pf)k_z, k_z \rangle$. Then $\|H_f k_z\| = \|H_\varphi k_z\| \le \|\varphi k_z\| = \mathrm{Var}^{1/2}(f - Pf; z)$. This proves the lemma. □

Proposition 5.3. *Given any* $2n < p < \infty$, *there is a constant* $C_{5.3}(p)$ *such that the following holds true: Let* $f \in L^2(S, d\sigma)$ *and suppose that the Hankel operator* H_f *is bounded. For each* (k, j), *let* $z_{k,j} \in T_{k,j}$ *satisfy the condition*

$$\mathrm{Var}(f - Pf; z_{k,j}) \ge \frac{1}{2} \sup_{z \in T_{k,j}} \mathrm{Var}(f - Pf; z). \tag{5.1}$$

Then the inequality

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j) \in I}) \le C_{5.3}(p) \|(H_f^* H_f)^{p/2}\|_\Phi$$

holds for every symmetric gauge function Φ.

Proof. For the given $f \in L^2(S, d\sigma)$, pick $w_{k,j} \in T_{k,j}$ such that

$$\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\| \ge \frac{1}{2} \sup_{w \in T_{k,j}} \|M_{m_w} H_f k_w\|,$$

$(k, j) \in I$. Then by Proposition 4.7 we have

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j) \in I_m}) \le C_{4.7}(p) \Phi(\{\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\|^p\}_{(k,j) \in I_m})$$

for every symmetric gauge function Φ and every $m \in \mathbf{N}$. Applying Lemma 5.2 to each $\|M_{m_{w_{k,j}}} H_f k_{w_{k,j}}\|$, for $0 < t \le 1$ we have

$$\begin{aligned}\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j) \in I_m}) &\le 2^{p-1} C_{4.7}(p) \Phi(\{\|H_f \psi_{w_{k,j},t}\|^p\}_{(k,j) \in I_m}) \\ &\quad + 2^{p-1} C_{4.7}(p) (C_{5.1} t)^p \Phi(\{\mathrm{Var}^{p/2}(f - Pf; w_{k,j})\}_{(k,j) \in I_m}).\end{aligned}$$

Since $w_{k,j} \in T_{k,j}$, it follows from (5.1) that $\mathrm{Var}(f - Pf; w_{k,j}) \le 2\mathrm{Var}(f - Pf; z_{k,j})$. Hence

$$\begin{aligned}\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j) \in I_m}) &\le 2^{p-1} C_{4.7}(p) \Phi(\{\|H_f \psi_{w_{k,j},t}\|^p\}_{(k,j) \in I_m}) \\ &\quad + 2^{p-1} 2^{p/2} C_{4.7}(p) (C_{5.1} t)^p \Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j) \in I_m}).\end{aligned}$$

Now, for the given $2n < p < \infty$, we pick $0 < t \le 1$ such that $2^{p-1}2^{p/2}C_{4.7}(p)(C_{5.1}t)^p \le 1/2$. This fixes the value of t in terms of p, and from the above inequality we obtain

$$\begin{aligned}\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I_m}) \le 2^{p-1}C_{4.7}(p)\Phi(\{\|H_f\psi_{w_{k,j},t}\|^p\}_{(k,j)\in I_m}) \\ + (1/2)\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I_m}).\end{aligned}$$

Since I_m is a finite set, the quantity $\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I_m})$ is finite. Therefore after the obvious cancellation the above inequality becomes

$$\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I_m}) \le 2^pC_{4.7}(p)\Phi(\{\|H_f\psi_{w_{k,j},t}\|^p\}_{(k,j)\in I_m}).$$

Finally, an application of Corollary 4.5 to the right-hand side gives us

$$\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I_m}) \le 2^pC_{4.7}(p)2^{t(p-2)}C_{3.3}(t)M_{4.3}M_{4.4}\|(H_f^*H_f)^{p/2}\|_\Phi.$$

Since this holds for every $m \in \mathbf{N}$, by (1.7) we have

$$\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z_{k,j})\}_{(k,j)\in I}) \le 2^pC_{4.7}(p)2^{t(p-2)}C_{3.3}(t)M_{4.3}M_{4.4}\|(H_f^*H_f)^{p/2}\|_\Phi.$$

This completes the proof of the proposition. □

We need to bring the Bergman metric into the proof of the lower bound.

Lemma 5.4. [18,Lemma 2.4] *Given any $0 < a < \infty$, there exists a natural number K which depends only on a and the complex dimension n such that the following holds true: Suppose that Γ is an a-separated subset of $\mathbf{B}$. Then there exist pairwise disjoint subsets $\Gamma_1, \dots, \Gamma_K$ of Γ such that $\cup_{\mu=1}^K \Gamma_\mu = \Gamma$ and such that* $\mathrm{card}(\Gamma_\mu \cap T_{k,j}) \le 1$ *for all $\mu \in \{1, \dots, K\}$ and $(k,j) \in I$.*

We would like to alert the reader to the fact that although Lemma 5.4 looks identical to Lemma 2.4 in [18], these two lemmas actually differ slightly. This is due to the fact that the set $T_{k,j}$ is defined slightly differently in [18]. Compare [18,(2.5)] with (3.7) in this paper. But this slight difference in definition does not change the fact that there is a $0 < C < \infty$ such that the β-diameter of each $T_{k,j}$ does not exceed $2C$. Therefore the proof of [18,Lemma 2.4] works verbatim here for Lemma 5.4.

Proposition 5.5. *Given $a > 0$, let K be the natural number provided by Lemma 5.4. Let $2n < p < \infty$. Let $f \in L^2(S, d\sigma)$ and suppose that the Hankel operator H_f is bounded. Then the inequality*

$$\Phi(\{\mathrm{Var}^{p/2}(f-Pf;z)\}_{z\in\Gamma}) \le 2^{p/2}KC_{5.3}(p)\|(H_f^*H_f)^{p/2}\|_\Phi$$

holds for every symmetric gauge function Φ and every a-separated set Γ in $\mathbf{B}$, where $C_{5.3}(p)$ is the constant provided by Proposition 5.3.

Proof. For each a-separated set Γ in $\mathbf{B}$, Lemma 5.4 provides the partition

$$\Gamma = \Gamma_1 \cup \dots \cup \Gamma_K \tag{5.2}$$

such that $\mathrm{card}(\Gamma_i \cap T_{k,j}) \le 1$ for all $i \in \{1, \dots, K\}$ and $(k,j) \in I$. This means that for each $i \in \{1, \dots, K\}$, there is a subset $I^{(i)}$ of I such that

$$\Gamma_i = \{z_{k,j}^{(i)} : (k,j) \in I^{(i)}\}$$

and such that $z_{k,j}^{(i)} \in T_{k,j}$ for every $(k,j) \in I^{(i)}$.

Let $f \in L^2(S, d\sigma)$ and suppose that H_f is bounded. Then for each $(k,j) \in I$ pick $z_{k,j} \in T_{k,j}$ such that (5.1) holds. Since $z_{k,j}^{(i)} \in T_{k,j}$, (5.1) implies

$$\mathrm{Var}^{p/2}(f - Pf; z_{k,j}^{(i)}) \le 2^{p/2}\mathrm{Var}^{p/2}(f - Pf; z_{k,j}), \quad (k,j) \in I^{(i)}. \tag{5.3}$$

Proposition 5.3 tells us that

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j})\}_{(k,j)\in I}) \le C_{5.3}(p)\|(H_f^* H_f)^{p/2}\|_\Phi.$$

Combining this with (5.3), we find that

$$\begin{aligned}\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z)\}_{z\in\Gamma_i}) &= \Phi(\{\mathrm{Var}^{p/2}(f - Pf; z_{k,j}^{(i)})\}_{(k,j)\in I^{(i)}}) \\ &\le 2^{p/2}C_{5.3}(p)\|(H_f^* H_f)^{p/2}\|_\Phi.\end{aligned}$$

By (5.2) and (4.7) we have

$$\Phi(\{\mathrm{Var}^{p/2}(f - Pf; z)\}_{z\in\Gamma}) \le \sum_{i=1}^{K} \Phi(\{\mathrm{Var}^{p/2}(f - Pf; z)\}_{z\in\Gamma_i}).$$

Obviously, the proposition follows from the above two inequalities. □

One should view Proposition 5.5 as a rather general result, a result that may have implications beyond this paper. But our immediate goal is to deduce the lower bound in Theorem 1.3 from Proposition 5.5. As it turns out, this involves an interesting special property of the family of symmetric gauge functions Φ_p^+, $1 < p < \infty$.

Lemma 5.6. *Suppose that $1 < p < \infty$. Let $\alpha = \{\alpha_1, \dots, \alpha_k, \dots\}$ be a sequence such that $\alpha_k \ge 0$ for every $k \ge 1$ and*

$$\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_k \ge \cdots . \tag{5.4}$$

Define

$$F_p(\alpha) = \sup_{k\ge1} k^{1/p}\alpha_k.$$

Then

$$\frac{p-1}{p}F_p(\alpha) \le \Phi_p^+(\alpha) \le F_p(\alpha).$$

Proof. By (1.8), the upper bound $\Phi_p^+(\alpha) \le F_p(\alpha)$ is obvious. To prove the lower bound, note that for every $k \in \mathbf{N}$,

$$\sum_{j=1}^{k} \frac{1}{j^{1/p}} \le 1 + \int_1^k \frac{1}{x^{1/p}} dx \le \frac{p}{p-1} k^{1-(1/p)}.$$

Therefore by (1.8) and (5.4), we have

$$\Phi_p^+(\alpha) \ge \frac{\alpha_1 + \cdots + \alpha_k}{1^{-1/p} + \cdots + k^{-1/p}} \ge \frac{k\alpha_k}{\{p/(p-1)\}k^{1-(1/p)}} = \frac{p-1}{p} k^{1/p}\alpha_k.$$

Since this holds for every $k \in \mathbf{N}$, we have $\Phi_p^+(\alpha) \ge \{(p-1)/p\}F_p(\alpha)$ as promised. □

Lemma 5.7. *Let $1 < r < \infty$, $1 < \rho < \infty$ and $p = \rho r$. Then for every sequence $\alpha = \{\alpha_1, \dots, \alpha_k, \dots\}$ of non-negative numbers we have*

$$\frac{\rho-1}{\rho}\left(\Phi_p^+(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r \le \Phi_\rho^+(\{\alpha_k^r\}_{k\in\mathbf{N}}) \le \left(\frac{p}{p-1}\Phi_p^+(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r.$$

Proof. If the sequence α fails to converge to 0, then both $\Phi_\rho^+(\{\alpha_k^r\}_{k\in\mathbf{N}})$ and $\Phi_p^+(\{\alpha_k\}_{k\in\mathbf{N}})$ are infinity, and the desired inequality holds trivially. Therefore we may assume $\alpha_k \to 0$ as $k \to \infty$. Since $\alpha_k \to 0$ as $k \to \infty$, discarding zeros and re-enumerating the other terms if necessary, it suffices to consider the case where

$$\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_k \ge \cdots.$$

Applying Lemma 5.6 and the relation $\rho r = p$, we have

$$\begin{aligned}\Phi_\rho^+(\{\alpha_k^r\}_{k\in\mathbf{N}}) &\le F_\rho(\{\alpha_k^r\}_{k\in\mathbf{N}}) = \sup_{k\ge1} k^{1/\rho}\alpha_k^r = \left(\sup_{k\ge1} k^{1/p}\alpha_k\right)^r \\ &= \left(F_p(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r \le \left(\frac{p}{p-1}\Phi_p^+(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r,\end{aligned}$$

proving the upper bound. To obtain the lower bound, we also apply Lemma 5.6 and the relation $\rho r = p$. We have

$$\begin{aligned}\Phi_\rho^+(\{\alpha_k^r\}_{k\in\mathbf{N}}) &\ge \frac{\rho-1}{\rho}F_\rho(\{\alpha_k^r\}_{k\in\mathbf{N}}) = \frac{\rho-1}{\rho}\sup_{k\ge1} k^{1/\rho}\alpha_k^r = \frac{\rho-1}{\rho}\left(\sup_{k\ge1} k^{1/p}\alpha_k\right)^r \\ &= \frac{\rho-1}{\rho}\left(F_p(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r \ge \frac{\rho-1}{\rho}\left(\Phi_p^+(\{\alpha_k\}_{k\in\mathbf{N}})\right)^r.\end{aligned}$$

This completes the proof. □

Proof of the lower bound in Theorem 1.3. Let $2n < p < \infty$ and $a > 0$ be given. We need to find a $c > 0$ that depends only on p, a and n such that the inequality

$$\|H_f\|_p^+ \ge c\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z\in\Gamma}) \tag{5.5}$$

holds for every $f \in L^2(S, d\sigma)$ and every a-separated set Γ in $\mathbf{B}$.

Since $p > 2n$, we can pick an r such that $2n < r < p$. Write $\rho = p/r$. Then

$$1 < \rho < \infty \quad \text{and} \quad \rho r = p.$$

Given an $f \in L^2(S, d\sigma)$, to prove (5.5), we may assume that $\|H_f\|_p^+ < \infty$, for otherwise there is nothing to prove. Note that the s-numbers of $(H_f^* H_f)^{r/2}$ are $(s_1(H_f))^r, \dots, (s_k(H_f))^r, \dots$. By Lemma 5.7 and the relation $\rho r = p$, we have

$$\begin{aligned} \|(H_f^* H_f)^{r/2}\|_\rho^+ &= \Phi_\rho^+(\{(s_k(H_f))^r\}_{k \in \mathbf{N}}) \\ &\le \left(\frac{p}{p-1} \Phi_p^+(\{s_k(H_f)\}_{k \in \mathbf{N}}) \right)^r = \left(\frac{p}{p-1} \|H_f\|_p^+ \right)^r. \end{aligned} \tag{5.6}$$

Now we apply Proposition 5.5 to r and the symmetric gauge function Φ_ρ^+. For any a-separated set Γ, we have

$$\|(H_f^* H_f)^{r/2}\|_\rho^+ = \|(H_f^* H_f)^{r/2}\|_{\Phi_\rho^+} \ge (2^{r/2} K C_{5.3}(r))^{-1} \Phi_\rho^+(\{\mathrm{Var}^{r/2}(f - Pf; z)\}_{z \in \Gamma}), \tag{5.7}$$

where K is determined by a. Applying Lemma 5.7 and the relation $\rho r = p$ again,

$$\Phi_\rho^+(\{\mathrm{Var}^{r/2}(f - Pf; z)\}_{z \in \Gamma}) \ge \frac{\rho - 1}{\rho} \left(\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z \in \Gamma}) \right)^r.$$

Combining this with (5.6) and (5.7), we obtain

$$\left(\frac{p}{p-1} \|H_f\|_p^+ \right)^r \ge \frac{\rho - 1}{2^{r/2} \rho K C_{5.3}(r)} \left(\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf; z)\}_{z \in \Gamma}) \right)^r.$$

Obviously, this implies (5.5), completing the proof of the lower bound in Theorem 1.3. □

6. A reverse Hölder's inequality

Having proved the lower bound in Theorem 1.3, we will now turn our attention to the upper bound. In the study of Hankel operators, upper bounds are inherently "two-sided" problems, as it suffices to consider commutators of the form $[P, M_g]$. In our estimate of $\|[P, M_g]\|_p^+$, the duality between symmetric gauge functions will play a crucial role. Let us first recall this duality.

Given a symmetric gauge function Φ, the formula

$$\Phi^*(\{b_j\}_{j \in \mathbf{N}}) = \sup \left\{ \left| \sum_{j=1}^{\infty} a_j b_j \right| : \{a_j\}_{j \in \mathbf{N}} \in \hat{c}, \Phi(\{a_j\}_{j \in \mathbf{N}}) \le 1 \right\}, \quad \{b_j\}_{j \in \mathbf{N}} \in \hat{c},$$

defines the symmetric gauge function that is dual to Φ [7,page 125]. For any $A \in \mathcal{C}_\Phi$ and $B \in \mathcal{C}_{\Phi^*}$, we have

$$|\mathrm{tr}(AB)| \le \|A\|_\Phi \|B\|_{\Phi^*}. \tag{6.1}$$

This follows from inequality (7.9) on page 63 of [7]. Moreover, we have the relation $\Phi^{**} = \Phi$ [7,page 125]. This relation implies that

$$\Phi(\{a_j\}_{j\in\mathbf{N}}) = \sup\left\{\left|\sum_{j=1}^{\infty} a_j b_j\right| : \{b_j\}_{j\in\mathbf{N}} \in \hat{c}, \Phi^*(\{b_j\}_{j\in\mathbf{N}}) \le 1\right\} \tag{6.2}$$

for each $\{a_j\}_{j\in\mathbf{N}} \in \hat{c}$.

For each $1 < p < \infty$, define

$$\Phi_p^-(\{a_j\}_{j\in\mathbf{N}}) = \sum_{j=1}^{\infty} \frac{|a_{\pi(j)}|}{j^{(p-1)/p}}, \quad \{a_j\}_{j\in\mathbf{N}} \in \hat{c},$$

where $\pi : \mathbf{N} \to \mathbf{N}$ is any bijection such that $|a_{\pi(1)}| \ge |a_{\pi(2)}| \ge \cdots \ge |a_{\pi(j)}| \ge \cdots$, which exists because each $\{a_j\}_{j\in\mathbf{N}} \in \hat{c}$ only has a finite number of nonzero terms. Then Φ_p^- is a symmetric gauge function. Indeed it is well known that the pair of symmetric gauge functions $\Phi_{p/(p-1)}^+$ and Φ_p^- are dual to each other [7,pages 148-149]. We need the following special property of Φ_p^-.

Lemma 6.1. *Let $1 < p < \infty$. Let X, Y be countable sets and let $N \in \mathbf{N}$. Suppose that $T : X \to Y$ is a map that is at most N-to-1. That is,* $\operatorname{card}\{x \in X : T(x) = y\} \le N$ *for every $y \in Y$. Then for every set of real numbers $\{a_y\}_{y\in Y}$ we have*

$$\Phi_p^-(\{a_{T(x)}\}_{x\in X}) \le \max\{p, 2\} N^{1/p} \Phi_p^-(\{a_y\}_{y\in Y}).$$

Proof. By (1.7), it suffices to consider the case $\{a_i\}_{i\in\mathbf{N}} \in \hat{c}$ where the terms are non-negative and in the descending order: $a_1 \ge a_2 \ge \cdots \ge a_i \ge \cdots$. Let $N \in \mathbf{N}$ and let $T : X \to \mathbf{N}$ be a map that is at most N-to-1. Now we define $\{b_j\}_{j\in\mathbf{N}}$ by the rule that

$$b_j = a_i \quad \text{if} \quad (i-1)N + 1 \le j \le iN, \quad i \in \mathbf{N}.$$

That is, one can regard $\{b_j\}_{j\in\mathbf{N}}$ as the "direct sum" of N copies of $\{a_i\}_{i\in\mathbf{N}}$. Since T is at most N-to-1, there is a subset E of $\mathbf{N}$ such that $\Phi_p^-(\{a_{T(x)}\}_{x\in X}) = \Phi_p^-(\{b_j\}_{j\in E})$. Therefore

$$\Phi_p^-(\{a_{T(x)}\}_{x\in X}) \le \Phi_p^-(\{b_j\}_{j\in\mathbf{N}}) = \sum_{j=1}^{\infty} \frac{b_j}{j^{p/(p-1)}} = \sum_{i=1}^{\infty} a_i \sum_{j=1}^{N} \frac{1}{((i-1)N + j)^{(p-1)/p}}. \tag{6.3}$$

For each $i \ge 2$, we have

$$\sum_{j=1}^{N} \frac{1}{((i-1)N + j)^{(p-1)/p}} \le \frac{N}{((i-1)N)^{(p-1)/p}} = \frac{\{i/(i-1)\}^{(p-1)/p} N^{1/p}}{i^{(p-1)/p}} \le \frac{2N^{1/p}}{i^{(p-1)/p}}.$$

For the term $i = 1$, we have $\sum_{j=1}^{N} j^{-(p-1)/p} \le pN^{1/p} = pN^{1/p}/1^{(p-1)/p}$. Substituting these in (6.3), the lemma follows. □

Recall from Section 3 that for each $k \ge 0$, $\{u_{k,1}, \dots, u_{k,m(k)}\}$ is a subset of S that is maximal with respect to property (3.5). For each $(k, j) \in I$, we now define the sets

$$B_{k,j} = B(u_{k,j}, 2^{-k+2}), \quad C_{k,j} = B(u_{k,j}, 2^{-k+3}) \quad \text{and} \quad D_{k,j} = B_{k,j} \times B_{k,j}.$$

Previously we introduced the finite subsets $I_m = \{(k, j) \in I : k \le m\}$, $m \in \mathbf{N}$, of I. But now we need subsets of I of a different kind. From now on, for each $\nu \in \mathbf{Z}_+$ we write

$$I^{(\nu)} = \{(k, j) \in I : k \ge \nu\}.$$

Lemma 6.2. *There is a constant $C_{6.2}$ that depends only on the complex dimension n such that the following estimate holds: Suppose that $\{b_{k,j}\}_{(k,j)\in I}$ is a set of non-negative numbers and let $\nu \ge 0$. Define*

$$b^{(\nu)}_{k+\nu,i} = \max\{b_{k,j} : B_{k,j} \cap B_{k+\nu,i} \ne \emptyset\}$$

for every $(k + \nu, i) \in I^{(\nu)}$. Then for every $2 \le p < \infty$ we have

$$\Phi_p^-(\{b^{(\nu)}_{k+\nu,i}\}_{(k+\nu,i)\in I^{(\nu)}}) \le C_{6.2}p2^{2n\nu/p}\Phi_p^-(\{b_{k,j}\}_{(k,j)\in I}).$$

Proof. By (2.2) and (3.5), there is a natural number $C_{6.2}$ such that for every triple of $k \ge 0$, $1 \le j \le m(k)$ and $\nu \ge 0$, we have

$$\operatorname{card}\{i : B_{k,j} \cap B_{k+\nu,i} \ne \emptyset, 1 \le i \le m(k + \nu)\} \le C_{6.2}2^{2n\nu}.$$

Let $\{b_{k,j}\}_{(k,j)\in I}$ and $\nu \ge 0$ be given. We define a map $T : I^{(\nu)} \to I$ in the following way. For each $(k + \nu, i) \in I^{(\nu)}$, there is an $h(k + \nu, i) \in \{1, \dots, m(k)\}$ which has the properties that $B_{k,h(k+\nu,i)} \cap B_{k+\nu,i} \ne \emptyset$ and that $b_{k,h(k+\nu,i)} = \max\{b_{k,j} : B_{k,j} \cap B_{k+\nu,i} \ne \emptyset\}$. Set $T(k + \nu, i) = (k, h(k + \nu, i))$. The second property ensures that $\Phi_p^-(\{b^{(\nu)}_{k+\nu,i}\}_{(k+\nu,i)\in I^{(\nu)}}) = \Phi_p^-(\{b_{T(k+\nu,i)}\}_{(k+\nu,i)\in I^{(\nu)}})$, while the first property guarantees that T is at most $C_{6.2}2^{2n\nu}$-to-1. Applying Lemma 6.1, we have

$$\Phi_p^-(\{b^{(\nu)}_{k+\nu,i}\}_{(k+\nu,i)\in I^{(\nu)}}) = \Phi_p^-(\{b_{T(k+\nu,i)}\}_{(k+\nu,i)\in I^{(\nu)}}) \le p(C_{6.2}2^{2n\nu})^{1/p}\Phi_p^-(\{b_{k,j}\}_{(k,j)\in I}).$$

Since $C_{6.2} \ge 1$ and $1/p < 1$, we have $C_{6.2}^{1/p} \le C_{6.2}$. Hence the lemma holds. □

In addition to the duality of between symmetric gauge functions, the estimate of $\|[P, M_g]\|_p^+$ also involves more mean oscillations in connection with our particular spherical decomposition. For $g \in L^2(S, d\sigma)$ and $(k, j) \in I$ we define

$$J(g; k, j) = \frac{1}{\sigma(C_{k,j})}\int_{C_{k,j}} |g - g_{C_{k,j}}|d\sigma \quad \text{and}$$

$$V_t(g; k, j) = \left\{2^{4nk}\iint_{D_{k,j}} |g(x) - g(y)|^t d\sigma(x)d\sigma(y)\right\}^{1/t}, \quad 1 < t < \infty.$$

These mean oscillations play a significant role in the proof of the upper bound.

Lemma 6.3. *Suppose that $1 < r < \infty$, $1 < t < \infty$, $1 < \rho < 2$ and that these numbers satisfy the condition $1 > r/t > (\rho - 1)/\rho$. Then there is a constant $C_{6.3} = C_{6.3}(r, t, \rho, n)$ such that*

$$\Phi_\rho^+(\{V_t^r(g; k, j)\}_{(k,j)\in I}) \le C_{6.3}\Phi_\rho^+(\{J^r(g; k, j)\}_{(k,j)\in I})$$

for every $g \in L^2(S, d\sigma)$.

Proof. Let $g \in L^2(S, d\sigma)$, and consider a large $m \in \mathbf{N}$. By (6.2) and by the duality between Φ_ρ^+ and $\Phi_{\rho/(\rho-1)}^-$, there exists a set of non-negative numbers $\{c_{k,j} : (k, j) \in I_m\}$ such that

$$\Phi_{\rho/(\rho-1)}^-(\{c_{k,j}\}_{(k,j)\in I_m}) \le 1 \quad \text{and}$$
$$\sum_{(k,j)\in I_m} c_{k,j}V_t^r(g; k, j) \ge \frac{1}{2}\Phi_\rho^+(\{V_t^r(g; k, j)\}_{(k,j)\in I_m}).$$

For convenience we define $c_{k,j} = 0$ when $k > m$. Then

$$\Phi_{\rho/(\rho-1)}^-(\{c_{k,j}\}_{(k,j)\in I}) = \Phi_{\rho/(\rho-1)}^-(\{c_{k,j}\}_{(k,j)\in I_m}) \le 1.$$

Let us analyze the sum

$$\mathcal{S} = \sum_{(k,j)\in I_m} c_{k,j}V_t^r(g; k, j).$$

For this we introduce the functions

$$g_k(\zeta) = \frac{1}{\sigma(B(\zeta, 2^{-k-2}))}\int_{B(\zeta,2^{-k-2})} g d\sigma, \quad \zeta \in S,$$

$k \ge 0$. Since $|g(\zeta) - g(\xi)| \le |g(\zeta) - g_k(\zeta)| + |g_k(\zeta) - g_k(\xi)| + |g_k(\xi) - g(\xi)|$, we have

$$\begin{aligned}\mathcal{S} \le C_1 &\sum_{(k,j)\in I_m} c_{k,j}\left\{2^{4nk}\iint_{D_{k,j}} |g(\zeta) - g_k(\zeta)|^t d\sigma(\zeta)d\sigma(\xi)\right\}^{r/t} \\ + C_1 &\sum_{(k,j)\in I_m} c_{k,j}\left\{2^{4nk}\iint_{D_{k,j}} |g_k(\zeta) - g_k(\xi)|^t d\sigma(\zeta)d\sigma(\xi)\right\}^{r/t} \\ + C_1 &\sum_{(k,j)\in I_m} c_{k,j}\left\{2^{4nk}\iint_{D_{k,j}} |g_k(\xi) - g(\xi)|^t d\sigma(\zeta)d\sigma(\xi)\right\}^{r/t},\end{aligned} \tag{6.4}$$

where $C_1 = 3^{r-1}$. By (2.2),

$$\begin{aligned}\iint_{D_{k,j}} |g(\zeta) - g_k(\zeta)|^t d\sigma(\zeta)d\sigma(\xi) &= \sigma(B_{k,j})\int_{B_{k,j}} |g - g_k|^t d\sigma \\ &\le A_0 2^{2n(-k+2)}\int_{B_{k,j}} |g - g_k|^t d\sigma.\end{aligned}$$

Substituting this in (6.4), we find that

$$\mathcal{S} \leq C_1\{2(A_0 2^{4n})^{r/t} S_1 + S_2\}, \tag{6.5}$$

where

$$S_1 = \sum_{(k,j)\in I_m} c_{k,j} \left\{ 2^{2nk} \int_{B_{k,j}} |g - g_k|^t d\sigma \right\}^{r/t} \qquad \text{and}$$

$$S_2 = \sum_{(k,j)\in I_m} c_{k,j} \left\{ 2^{4nk} \iint_{D_{k,j}} |g_k(\zeta) - g_k(\xi)|^t d\sigma(\zeta) d\sigma(\xi) \right\}^{r/t}.$$

We analyze S_1 and S_2 separately.

For S_1, note that for each pair of integers $k \geq 0$ and $s \geq 0$,

$$2^{2nk} \int_{B_{k,j}} |g_{k+s} - g_{k+s+1}|^t d\sigma \leq \frac{C_2}{\sigma(B_{k,j})} \int_{B_{k,j}} (|g_{k+s} - g_{C_{k,j}}|^t + |g_{C_{k,j}} - g_{k+s+1}|^t) d\sigma. \tag{6.6}$$

If $\zeta \in B_{k,j} = B(u_{k,j}, 2^{-k+2})$, then $B(\zeta, 2^{-k-s-2}) \subset B(u_{k,j}, 2^{-k+3}) = C_{k,j}$ and

$$\begin{aligned} |g_{k+s}(\zeta) - g_{C_{k,j}}| &\leq \frac{1}{\sigma(B(\zeta, 2^{-k-s-2}))} \int_{B(\zeta,2^{-k-s-2})} |g - g_{C_{k,j}}| d\sigma \\ &\leq \frac{C_3 2^{2ns}}{\sigma(C_{k,j})} \int_{C_{k,j}} |g - g_{C_{k,j}}| d\sigma = C_3 2^{2ns} J(g;k,j). \end{aligned} \tag{6.7}$$

A similar inequality holds for $|g_{C_{k,j}} - g_{k+s+1}(\zeta)|$, $\zeta \in B_{k,j}$. Substituting these estimates in (6.6), we obtain

$$2^{2nk} \int_{B_{k,j}} |g_{k+s} - g_{k+s+1}|^t d\sigma \leq C_4 2^{2nst} J^t(g;k,j), \tag{6.8}$$

$k \geq 0$, $s \geq 0$. Obviously, $|g_{B_{k,j}} - g_{C_{k,j}}| \leq C_5 J(g;k,j)$. Combining this with (6.7), we have

$$2^{2nk} \int_{B_{k,j}} |g_{B_{k,j}} - g_k|^t d\sigma \leq C_6 J^t(g;k,j).$$

Hence

$$\begin{aligned} 2^{2nk} \int_{B_{k,j}} |g - g_k|^t d\sigma &\leq 2^{t-1} 2^{2nk} \int_{B_{k,j}} |g - g_{B_{k,j}}|^t d\sigma + 2^{t-1} 2^{2nk} \int_{B_{k,j}} |g_{B_{k,j}} - g_k|^t d\sigma \\ &\leq C_7 V_t^t(g;k,j) + C_8 J^t(g;k,j). \end{aligned} \tag{6.9}$$

Suppose that $\nu \in \mathbf{N}$. If $j, j' \in \{1, \ldots, m(k)\}$ and $i \in \{1, \ldots, m(k+\nu)\}$ are such that both intersections $B_{k,j} \cap B_{k+\nu,i}$ and $B_{k,j'} \cap B_{k+\nu,i}$ are non-empty, then $d(u_{k,j}, u_{k,j'}) \leq 3 \cdot 2^{-k+2}$. Thus by (3.5) and (2.2), there is an $M \in \mathbf{N}$ such that the inequality

$$\mathrm{card}\{j : B_{k,j} \cap B_{k+\nu,i} \neq \emptyset, 1 \leq j \leq m(k)\} \leq M$$

holds for all $k \geq 0$, $\nu \in \mathbf{N}$ and $i \in \{1, \ldots, m(k+\nu)\}$. Now, for each triple of $k \geq 0$, $\nu \in \mathbf{N}$ and $i \in \{1, \ldots, m(k+\nu)\}$, define

$$\alpha_{k+\nu,i}^{(\nu)} = M \max\{c_{k,j} : B_{k,j} \cap B_{k+\nu,i} \neq \emptyset\}. \tag{6.10}$$

Then for each pair of $k \geq 0$ and $\nu \geq 1$, we have

$$\begin{aligned}
&\sum_{j=1}^{m(k)} c_{k,j} \left\{ 2^{2nk} \int_{B_{k,j}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} \\
&\quad \leq 2^{-2n\nu r/t} \sum_{j=1}^{m(k)} c_{k,j} \sum_{B_{k+\nu,i} \cap B_{k,j} \neq \emptyset} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} \\
&\quad = 2^{-2n\nu r/t} \sum_{i=1}^{m(k+\nu)} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} \sum_{B_{k,j} \cap B_{k+\nu,i} \neq \emptyset} c_{k,j} \\
&\quad \leq 2^{-2n\nu r/t} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t}, \qquad (6.11)
\end{aligned}$$

where the assumption $r/t < 1$ is used to justify the first $\leq$. For each $\nu \in \mathbf{N}$, we have

$$S_1 \leq 2^{r-1}(T_1(\nu) + T_2(\nu)), \tag{6.12}$$

where

$$T_1(\nu) = \sum_{(k,j) \in I_m} c_{k,j} \left\{ 2^{2nk} \int_{B_{k,j}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} \quad \text{and}$$

$$T_2(\nu) = \sum_{(k,j) \in I_m} c_{k,j} \left\{ 2^{2nk} \int_{B_{k,j}} |g_{k+\nu} - g_k|^t d\sigma \right\}^{r/t}.$$

Denote

$$T_3(\nu) = \sum_{k=m-\nu+1}^{m} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t}.$$

By (6.11) and (6.9), we have

$$
\begin{aligned}
T_1(\nu) &\le 2^{-2n\nu r/t} \sum_{k=0}^{m} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} \\
&= 2^{-2n\nu r/t} \sum_{k=0}^{m-\nu} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} \left\{ 2^{2n(k+\nu)} \int_{B_{k+\nu,i}} |g - g_{k+\nu}|^t d\sigma \right\}^{r/t} + 2^{-2n\nu r/t} T_3(\nu) \\
&\le C_9 2^{-2n\nu r/t} \sum_{k=0}^{m-\nu} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} V_t^r(g; k+\nu, i) \\
&\qquad + C_{10} 2^{-2n\nu r/t} \sum_{k=0}^{m-\nu} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} J^r(g; k+\nu, i) + 2^{-2n\nu r/t} T_3(\nu).
\end{aligned}
$$

Since $\{(k+\nu, i) \in I : 0 \le k \le m - \nu\}$ is a subset of both I_m and $I^{(\nu)}$, by (6.2) we have

$$
\sum_{k=0}^{m-\nu} \sum_{i=1}^{m(k+\nu)} \alpha_{k+\nu,i}^{(\nu)} V_t^r(g; k+\nu, i) \le \Phi_{\rho/(\rho-1)}^{-}(\{\alpha_{k+\nu,i}^{(\nu)}\}_{(k+\nu,i)\in I^{(\nu)}}) \Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}).
$$

Therefore

$$
\begin{aligned}
T_1(\nu) &\le C_9 2^{-2n\nu r/t} \Phi_{\rho/(\rho-1)}^{-}(\{\alpha_{k+\nu,i}^{(\nu)}\}_{(k+\nu,i)\in I^{(\nu)}}) \Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) \\
&\quad + C_{10} 2^{-2n\nu r/t} \Phi_{\rho/(\rho-1)}^{-}(\{\alpha_{k+\nu,i}^{(\nu)}\}_{(k+\nu,i)\in I^{(\nu)}}) \Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) + 2^{-2n\nu r/t} T_3(\nu).
\end{aligned}
$$

By (6.10) and Lemma 6.2, we have

$$
\Phi_{\rho/(\rho-1)}^{-}(\{\alpha_{k+\nu,i}^{(\nu)}\}_{(k+\nu,i)\in I^{(\nu)}}) \le M C_{6.2}\{\rho/(\rho-1)\} 2^{2n\nu(\rho-1)/\rho} \Phi_{\rho/(\rho-1)}^{-}(\{c_{k,j}\}_{(k,j)\in I}).
$$

Write $\delta = (r/t) - \{(\rho-1)/\rho\}$. Recall that we have $\Phi_{\rho/(\rho-1)}^{-}(\{c_{k,j}\}_{(k,j)\in I}) \le 1$. Thus

$$
\begin{aligned}
T_1(\nu) &\le C_9' 2^{-2n\nu\delta} \Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) \\
&\quad + C_{10}' 2^{-2n\nu\delta} \Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) + T_3(\nu),
\end{aligned}
$$

where $C_9' = C_9 M C_{6.2}\{\rho/(\rho-1)\}$ and $C_{10}' = C_{10} M C_{6.2}\{\rho/(\rho-1)\}$. Substituting this in (6.12) and then recalling (6.5), with $C_{11} = 2^r C_1 (A_0 2^{4n})^{r/t}$, we have

$$
\begin{aligned}
\frac{1}{2}\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) \le \mathcal{S} &\le C_{11} C_9' 2^{-2n\nu\delta} \Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) \\
&\quad + C_{11} C_{10}' 2^{-2n\nu\delta} \Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) \\
&\quad + C_{11}(T_2(\nu) + T_3(\nu)) + C_1 S_2.
\end{aligned}
$$

Now comes the most crucial step in the proof. Since $\delta = (r/t) - \{(\rho-1)/\rho\}$ is greater than 0 by assumption, we can fix a $\nu \in \mathbf{N}$ such that $C_{11} C_9' 2^{-2n\nu\delta} \le 1/4$. Since C_{11} and

C_9' depend only on n, r, t and ρ, this fixes the value of ν in terms of these numbers. With $C_{11}C_9'2^{-2n\nu\delta} \le 1/4$, an obvious cancellation in the above leads to

$$\begin{aligned}\frac{1}{4}\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) &\le C_{12}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) \\ &\quad + C_{11}(T_2(\nu)+T_3(\nu)) + C_1S_2,\end{aligned} \tag{6.13}$$

where $C_{12} = C_{11}C_{10}'$. Next we estimate $T_2(\nu)$, $T_3(\nu)$ and S_2.

For $T_2(\nu)$, note that

$$\begin{aligned}\left\{2^{2nk}\int_{B_{k,j}}|g_{k+\nu}-g_k|^t d\sigma\right\}^{r/t} &\le \nu^{r-1}\sum_{s=0}^{\nu-1}\left\{2^{2nk}\int_{B_{k,j}}|g_{k+s+1}-g_{k+s}|^t d\sigma\right\}^{r/t} \\ &\le \nu^r C_4^{r/t}2^{2n\nu r}J^r(g;k,j),\end{aligned}$$

where the second $\le$ follows from (6.8). Since the natural number ν is now fixed, we can write $C_{13} = \nu^r C_4^{r/t}2^{2n\nu r}$. Recalling the definition of $T_2(\nu)$, we have

$$\begin{aligned}T_2(\nu) \le C_{13}\sum_{(k,j)\in I_m} c_{k,j}J^r(g;k,j) &\le C_{13}\Phi^-_{\rho/(\rho-1)}(\{c_{k,j}\}_{(k,j)\in I_m})\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I_m}) \\ &\le C_{13}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}).\end{aligned} \tag{6.14}$$

The estimate of $T_3(\nu)$ is more interesting. Here, the main observation is that the total number of k's involved in $T_3(\nu)$ is only ν: $m-\nu+1 \le k \le m$. Thus

$$T_3(\nu) = Y_{m+1} + \cdots + Y_{m+\nu}, \tag{6.15}$$

where

$$Y_\ell = \sum_{i=1}^{m(\ell)} \alpha_{\ell,i}^{(\nu)}\left\{2^{2n\ell}\int_{B_{\ell,i}}|g-g_\ell|^t d\sigma\right\}^{r/t},$$

$m+1 \le \ell \le m+\nu$. To estimate Y_ℓ, define

$$Y_{\ell,L} = \sum_{i=1}^{m(\ell)} \alpha_{\ell,i}^{(\nu)}\left\{2^{2n\ell}\int_{B_{\ell,i}}|g_{\ell+L}-g_\ell|^t d\sigma\right\}^{r/t}$$

for $L \in \mathbf{N}$. Write $\epsilon = \delta/2$ (recall that $\delta = (r/t) - \{(\rho-1)/\rho\}$). By Hölder's inequality,

$$|g_{\ell+L}-g_\ell|^t \le \frac{1}{(1-2^{-\epsilon/(t-1)})^{t-1}}\sum_{q=0}^{L-1}2^{\epsilon q}|g_{\ell+q+1}-g_{\ell+q}|^t.$$

Since $r/t < 1$, this leads to

$$Y_{\ell,L} \le C_{14}\sum_{q=0}^{L-1}2^{\epsilon q}Z_{\ell,q}, \quad \text{where} \quad Z_{\ell,q} = \sum_{i=1}^{m(\ell)}\alpha_{\ell,i}^{(\nu)}\left\{2^{2n\ell}\int_{B_{\ell,i}}|g_{\ell+q+1}-g_{\ell+q}|^t d\sigma\right\}^{r/t}.$$

For each triple of $\ell \geq m+1$, $q \geq 0$ and $\gamma \in \{1, \dots, m(\ell+q)\}$, define

$$h_{\ell+q,\gamma} = M \max\{\alpha^{(\nu)}_{\ell,i} : B_{\ell,i} \cap B_{\ell+q,\gamma} \neq \emptyset\}. \tag{6.16}$$

Then

$$\begin{aligned}
Z_{\ell,q} &\leq \sum_{i=1}^{m(\ell)} \alpha^{(\nu)}_{\ell,i} \sum_{B_{\ell,i}\cap B_{\ell+q,\gamma}\neq\emptyset} \left\{ 2^{2n\ell} \int_{B_{\ell+q,\gamma}} |g_{\ell+q+1} - g_{\ell+q}|^t d\sigma \right\}^{r/t} \\
&\leq 2^{-2nqr/t} \sum_{\gamma=1}^{m(\ell+q)} h_{\ell+q,\gamma} \left\{ 2^{2n(\ell+q)} \int_{B_{\ell+q,\gamma}} |g_{\ell+q+1} - g_{\ell+q}|^t d\sigma \right\}^{r/t} \\
&\leq C_4^{r/t} 2^{-2nqr/t} \sum_{\gamma=1}^{m(\ell+q)} h_{\ell+q,\gamma} J^r(g; \ell+q, \gamma),
\end{aligned}$$

where for the last $\leq$ we use (6.8) in the case $s = 0$. By (6.16) and Lemma 6.2, we have

$$\begin{aligned}
\Phi^-_{\rho/(\rho-1)}(\{h_{\ell+q,\gamma}\}_{\gamma=1}^{m(\ell+q)}) &\leq MC_{6.2}\{\rho/(\rho-1)\} 2^{2nq(\rho-1)/\rho} \Phi^-_{\rho/(\rho-1)}(\{\alpha^{(\nu)}_{k+\nu,i}\}_{(k+\nu,i)\in I^{(\nu)}}) \\
&\leq \{MC_{6.2}\{\rho/(\rho-1)\}\}^2 2^{2n\nu(\rho-1)/\rho} \cdot 2^{2nq(\rho-1)/\rho} = C_{15} 2^{2nq(\rho-1)/\rho}.
\end{aligned}$$

By (6.2), we have

$$\begin{aligned}
Z_{\ell,q} &\leq C_4^{r/t} 2^{-2nqr/t} \Phi^-_{\rho/(\rho-1)}(\{h_{\ell+q,\gamma}\}_{\gamma=1}^{m(\ell+q)}) \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}) \\
&\leq C_{16} 2^{-2nq\delta} \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}).
\end{aligned}$$

Since $\delta = 2\epsilon$, we now have

$$Y_{\ell,L} \leq C_{14}C_{16} \sum_{q=0}^\infty 2^{\epsilon q} \cdot 2^{-2nq\delta} \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}) = C_{17} \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}).$$

Obviously, $g_{\ell+L}(\zeta) \to g(\zeta)$ as $L \to \infty$ for σ-a.e. $\zeta \in S$. Thus, by Fatou's lemma,

$$Y_\ell \leq \liminf_{L\to\infty} Y_{\ell,L} \leq C_{17} \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}).$$

Recalling (6.15), if we write $C_{18} = \nu C_{17}$, then

$$T_3(\nu) \leq C_{18} \Phi^+_\rho(\{J^r(g;k,j)\}_{(k,j)\in I}). \tag{6.17}$$

For S_2, note that

$$\begin{aligned}
2^{4nk} \iint_{D_{k,j}} |g_k(\zeta) - g_k(\xi)|^t d\sigma(\zeta) d\sigma(\xi) &\leq 2^t 2^{4nk} \sigma(B_{k,j}) \int_{B_{k,j}} |g_k - g_{C_{k,j}}|^t d\sigma \\
&\leq \frac{C_{19}}{\sigma(B_{k,j})} \int_{B_{k,j}} |g_k - g_{C_{k,j}}|^t d\sigma \leq C_{19} C_3^t J^t(g;k,j),
\end{aligned}$$

where the last $\leq$ follows from (6.7). Writing $C_{20} = (C_{19}C_3^t)^{r/t}$, we have

$$S_2 \leq C_{20} \sum_{(k,j)\in I_m} c_{k,j} J^r(g;k,j) \leq C_{20}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) \tag{6.18}$$

where the second $\leq$ follows from (6.2) and the fact that $\Phi^-_{\rho/(\rho-1)}(\{c_{k,j}\}_{(k,j)\in I_m}) \leq 1$.

Now if we set

$$C_{6.3} = 4\{C_{12} + C_{11}(C_{13} + C_{18}) + C_1C_{20}\},$$

then the combination of (6.13), (6.14), (6.17) and (6.18) gives us

$$\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I_m}) \leq C_{6.3}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}).$$

Since $C_{6.3}$ is independent of m, by (1.7) the above implies

$$\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I}) \leq C_{6.3}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}).$$

This completes the proof of the lemma. $\square$

We now have the following "reverse Hölder's inequality" involving Φ_p^+:

Proposition 6.4. *Suppose that $1 < p \leq t < \infty$. Then there is a constant $C_{6.4} = C_{6.4}(p,t,n)$ such that*

$$\Phi_p^+(\{V_t(g;k,j)\}_{(k,j)\in I}) \leq C_{6.4}\Phi_p^+(\{J(g;k,j)\}_{(k,j)\in I})$$

for every $g \in L^2(S, d\sigma)$.

Proof. Given $1 < p \leq t < \infty$, we pick an $r \in (\max\{1, p/2\}, p)$ such that $r/t > (p-r)/p$. With r so chosen, we set $\rho = p/r$. Then $1 < \rho < 2$, $\rho r = p$, and $(\rho-1)/\rho = (p-r)/p < r/t < 1$. Applying Lemma 6.3 to these r, t and ρ, we obtain

$$\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I}) \leq C_{6.3}\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}),$$

$g \in L^2(S, d\sigma)$. On the other hand, since $\rho r = p$, Lemma 5.7 tells us that

$$(\Phi_p^+(\{V_t(g;k,j)\}_{(k,j)\in I}))^r \leq \{\rho/(\rho-1)\}\Phi_\rho^+(\{V_t^r(g;k,j)\}_{(k,j)\in I})$$

and that

$$\Phi_\rho^+(\{J^r(g;k,j)\}_{(k,j)\in I}) \leq \{p/(p-1)\}^r(\Phi_p^+(\{J(g;k,j)\}_{(k,j)\in I}))^r.$$

Obviously, the proposition follows from the combination of these three inequalities. $\square$

7. Interpolation

Recall that for an operator A, we write $s_1(A)$, $s_2(A), \cdots$, $s_j(A), \cdots$ for its s-numbers. Furthermore, for each $t > 0$ we denote

$$N_A(t) = \text{card}\{j \in \mathbf{N} : s_j(A) > t\}$$

as in [5]. Also recall from [5,(7.1)] that we have the inequality

$$N_{A+B}(t) \le N_A(t/2) + N_B(t/2).$$

A common tool in the study of commutators is interpolation. See, e.g., [3,8]. But the estimate of $\|[P, M_g]\|_p^+$ involves an interpolation of a rather ad hoc kind. This particular interpolation requires quite a bit of work, including all the work in Section 6.

For this section we introduce the measure

$$d\mu(x, y) = \frac{d\sigma(x)d\sigma(y)}{|1 - \langle x, y\rangle|^{2n}}$$

on $S \times S$. For each $1 < p < \infty$, let $L^p_{\text{sym}}(S \times S, d\mu)$ be the collection of functions F on $S \times S$ which are L^p with respect to $d\mu$ and which satisfy the condition

$$|F(x, y)| = |F(y, x)|, \qquad (x, y) \in S \times S.$$

For each $F \in L^p_{\text{sym}}(S \times S, d\mu)$, define T_F to be the integral operator on $L^2(S, d\sigma)$ with the kernel function

$$K_F(x, y) = \frac{F(x, y)}{(1 - \langle x, y\rangle)^n}.$$

Our interpolation begins with the requisite weak-type inequality.

Proposition 7.1. *Given any $2 < p < \infty$, there is a constant $C_{7.1} = C_{7.1}(p, n)$ such that*

$$N_{T_F}(t) \le \frac{C_{7.1}}{t^p} \int |F|^p d\mu = \frac{C_{7.1}}{t^p} \iint \frac{|F(x, y)|^p}{|1 - \langle x, y\rangle|^{2n}} d\sigma(x)d\sigma(y)$$

for all $F \in L^p_{\text{sym}}(S \times S, d\mu)$ and $t > 0$.

Proof. Given $2 < p < \infty$, we denote $q = p/(p-1)$. Then $1 < q < 2$. Let $F \in L^p_{\text{sym}}(S\times S, d\mu)$ be given. For each $x \in S$, define $\eta(x)$ to be the smallest quantity (non-negative number or infinity) such that the inequality

$$\sigma(\{y \in S : |K_F(x, y)|^q > \lambda\}) \le \eta^q(x)/\lambda$$

holds for all $\lambda > 0$. The proof is divided into two steps. The first step is to show that

$$N_{T_F}(t) \le \left(\frac{4}{2 - q} + 2\right)\left(\frac{2}{q - 1}\right)^p \frac{1}{t^p} \int \eta^p d\sigma \tag{7.1}$$

for all $t > 0$. Fix a $t > 0$ for the moment. To prove (7.1), define the set

$$\mathcal{G} = \mathcal{G}_t = \{(x, y) \in S \times S : |K_F(x, y)| \le t^{-p+1} \max\{\eta^p(x), \eta^p(y)\}\}.$$

We have the factorization

$$K_F(x, y) = \theta(x, y)|K_F(x, y)|, \qquad \text{where} \ \ |\theta(x, y)| = 1,$$

on $S \times S$. Now define the functions

$$G(x, y) = \begin{cases} K_F(x, y) & \text{if} \quad (x, y) \in \mathcal{G} \\ t^{-p+1} \max\{\eta^p(x), \eta^p(y)\}\theta(x, y) & \text{if} \quad (x, y) \notin \mathcal{G} \end{cases} \qquad \text{and}$$

$$B(x, y) = \begin{cases} 0 & \text{if} \quad (x, y) \in \mathcal{G} \\ (|K_F(x, y)| - t^{-p+1} \max\{\eta^p(x), \eta^p(y)\})\theta(x, y) & \text{if} \quad (x, y) \notin \mathcal{G} \end{cases}.$$

Then

$$K_F = G + B \quad \text{and} \quad T_F = T_1 + T_2,$$

where T_1 and T_2 are the integral operators in $L^2(S, d\sigma)$ which have G and B as kernel functions respectively. If $x \in S$ is such that $0 < \eta(x) < \infty$, then

$$\begin{aligned} \int |B(x, y)|d\sigma(y) &\le \int_{|K_F(x,y)|-t^{-p+1}\eta^p(x)>0} (|K_F(x, y)| - t^{-p+1}\eta^p(x))d\sigma(y) \\ &= \int_0^\infty \sigma(\{y : |K_F(x, y)| - t^{-p+1}\eta^p(x) > s\})ds \\ &= \int_{\eta^p(x)/t^{p-1}}^\infty \sigma(\{y : |K_F(x, y)| > s\})ds \le \int_{\eta^p(x)/t^{p-1}}^\infty \frac{\eta^q(x)}{s^q} ds \\ &= \frac{\eta^q(x)}{q-1} \cdot \frac{1}{(\eta^p(x)/t^{p-1})^{q-1}} = \frac{t}{q-1}. \end{aligned}$$

A chase of the definitions shows that $\int |B(x, y)|d\sigma(y) = 0$ if either $\eta(x) = 0$ or $\eta(x) = \infty$. Therefore

$$\int |B(x, y)|d\sigma(y) \le \frac{t}{q-1} \tag{7.2}$$

for every $x \in S$. Since $|K_F(x, y)| = |K_F(y, x)|$ we also have

$$\int |B(x, y)|d\sigma(x) \le \frac{t}{q-1}. \tag{7.3}$$

It is well known that (7.2) and (7.3) together imply

$$\|T_2\| \le \frac{t}{q-1}. \tag{7.4}$$

Next we estimate $N_{T_1}(t)$. Writing $\|\cdot\|_2$ for the Hilbert-Schmidt norm, we have

$$\begin{aligned}
t^2 N_{T_1}(t) \le \|T_1\|_2^2 &= \iint |G|^2 d\sigma d\sigma = \iint_{\mathcal{G}} |G|^2 d\sigma d\sigma + \iint_{(S\times S)\setminus\mathcal{G}} |G|^2 d\sigma d\sigma \\
&\le \iint_{|K_F(x,y)|\le \eta^p(x)/t^{p-1}} |K_F(x,y)|^2 d\sigma(x) d\sigma(y) \\
&+ \iint_{|K_F(x,y)|\le \eta^p(y)/t^{p-1}} |K_F(x,y)|^2 d\sigma(x) d\sigma(y) \\
&+ \int \left(\frac{\eta^p(x)}{t^{p-1}}\right)^2 \sigma(\{y : |K_F(x,y)| > \eta^p(x)/t^{p-1}\}) d\sigma(x) \\
&+ \int \left(\frac{\eta^p(y)}{t^{p-1}}\right)^2 \sigma(\{x : |K_F(x,y)| > \eta^p(y)/t^{p-1}\}) d\sigma(y) \\
&\le 2\int \left(\int_0^{\eta^p(x)/t^{p-1}} 2s\sigma(\{y : |K_F(x,y)| > s\}) ds\right) d\sigma(x) \\
&+ 2\int \left(\frac{\eta^p(x)}{t^{p-1}}\right)^2 \sigma(\{y : |K_F(x,y)| > \eta^p(x)/t^{p-1}\}) d\sigma(x). \qquad (7.5)
\end{aligned}$$

Using the definition of $\eta(x)$ again, and using the fact that $2-q>0$, we have

$$\begin{aligned}
\int_0^{\eta^p(x)/t^{p-1}} 2s\sigma(\{y : |K_F(x,y)| > s\}) ds &\le \int_0^{\eta^p(x)/t^{p-1}} \frac{2s\eta^q(x)}{s^q} ds \\
&= \frac{2}{2-q}\eta^q(x)(\eta^p(x)/t^{p-1})^{2-q} = \frac{2}{2-q}\cdot\frac{\eta^p(x)}{t^{p-2}}. \qquad (7.6)
\end{aligned}$$

On the other hand,

$$\left(\frac{\eta^p(x)}{t^{p-1}}\right)^2 \sigma(\{y : |K_F(x,y)| > \eta^p(x)/t^{p-1}\}) \le \left(\frac{\eta^p(x)}{t^{p-1}}\right)^2 \frac{\eta^q(x)}{(\eta^p(x)/t^{p-1})^q} = \frac{\eta^p(x)}{t^{p-2}}. \qquad (7.7)$$

Substituting (7.6) and (7.7) back in (7.5), we have

$$t^2 N_{T_1}(t) \le \left(\frac{4}{2-q} + 2\right)\frac{1}{t^{p-2}}\int \eta^p d\sigma.$$

Equivalently,

$$N_{T_1}(t) \le \left(\frac{4}{2-q} + 2\right)\frac{1}{t^p}\int \eta^p d\sigma. \qquad (7.8)$$

Now the relation $T_F = T_1 + T_2$ leads to

$$N_{T_F}(2t/(q-1)) \le N_{T_1}(t/(q-1)) + N_{T_2}(t/(q-1)).$$

But (7.4) tells us that $N_{T_2}(t/(q-1)) = 0$. Combining this with the fact that $t/(q-1) > t$ and with (7.8), we now have

$$N_{T_F}(2t/(q-1)) \le N_{T_1}(t/(q-1)) \le N_{T_1}(t) \le \left(\frac{4}{2-q} + 2\right) \frac{1}{t^p} \int \eta^p d\sigma.$$

Thus (7.1) follows from this by an obvious substitution.

The second step in the proof of the proposition is to show that

$$\int \eta^p d\sigma \le 2^{p/q} A_0^{p-2} \iint \frac{|F(x,y)|^p}{|1-\langle x,y\rangle|^{2n}} d\sigma(x) d\sigma(y), \tag{7.9}$$

where A_0 is the constant in (2.2). To prove this, set $\alpha = p/(p-2)$. For a Borel function φ on S we define $\|\varphi\|_{\alpha,\infty}$ to be the smallest quantity (non-negative number or infinity) such that the inequality

$$\sigma(\{y : |\varphi(y)|^\alpha > \lambda\}) \le \|\varphi\|_{\alpha,\infty}^\alpha / \lambda$$

holds for all $\lambda > 0$. Fix an $x \in S$ for the moment and define the functions

$$g(y) = \frac{F(x,y)}{(1-\langle x,y\rangle)^{2n/p}} \qquad \text{and} \qquad h(y) = \frac{1}{(1-\langle x,y\rangle)^{n(1-(2/p))}}.$$

Note that if $\|g\|_p = 0$, then x has the property $\eta(x) = 0$. Since $F \in L^p_{\text{sym}}(S \times S, d\mu)$, to prove (7.9), it suffices to consider the case where $0 < \|g\|_p < \infty$. By (2.2), we have

$$\begin{aligned}\sigma(\{y : |h(y)|^\alpha > \lambda\}) &= \sigma(\{y : |1-\langle x,y\rangle| < \lambda^{-(1/n)}\}) \\ &= \sigma(B(x, \lambda^{-(1/2n)})) \le A_0 (\lambda^{-(1/2n)})^{2n} = A_0/\lambda.\end{aligned}$$

Thus

$$\|h\|_{\alpha,\infty} \le A_0^{1/\alpha} < \infty. \tag{7.10}$$

Now suppose that

$$\|g\|_p \|h\|_{\alpha,\infty} = a.$$

Then define

$$A = a^{1/(p-1)} \quad \text{and} \quad B = a^{(p-2)/(p-1)}.$$

We have

$$A^p = a^q = B^\alpha. \tag{7.11}$$

Since $AB = a$, there is a $c > 0$ such that if we set

$$G = g/c \quad \text{and} \quad H = ch,$$

then

$$\|G\|_p \|H\|_{\alpha,\infty} = a, \quad \|G\|_p = A, \quad \text{and} \quad \|H\|_{\alpha,\infty} = B. \tag{7.12}$$

For each $\lambda > 0$, consider the set

$$E_\lambda = \{y : |K_F(x,y)|^q > \lambda\} = \{y : |g(y)h(y)|^q > \lambda\} = \{y : |G(y)H(y)|^q > \lambda\}.$$

If $z \in E_\lambda \cap \{y : |H(y)|^\alpha \le \lambda\}$, then $\lambda < |G(z)H(z)|^q \le |G(z)|^q \lambda^{q/\alpha}$, i.e.,

$$|G(z)|^q > \lambda^{1-(q/\alpha)} = \lambda^{1/(p-1)}, \quad \text{which means} \quad |G(z)|^p > \lambda.$$

Hence

$$E_\lambda \subset \{y : |H(y)|^\alpha > \lambda\} \cup \{y : |G(y)|^p > \lambda\}$$

for every $\lambda > 0$. Recalling (7.11) and (7.12), we now have

$$\begin{aligned} \sigma(E_\lambda) &\le \sigma(\{y : |H(y)|^\alpha > \lambda\}) + \sigma(\{y : |G(y)|^p > \lambda\}) \\ &\le \lambda^{-1}(\|H\|_{\alpha,\infty}^\alpha + \|G\|_p^p) = 2\lambda^{-1}a^q = 2\lambda^{-1}(\|g\|_p\|h\|_{\alpha,\infty})^q. \end{aligned}$$

Recalling the definitions of $\eta(x)$ and g and recalling (7.10), this gives us

$$\eta(x) \le 2^{1/q}\|g\|_p\|h\|_{\alpha,\infty} \le 2^{1/q}A_0^{1/\alpha}\left(\int \frac{|F(x,y)|^p}{|1-\langle x,y\rangle|^{2n}}d\sigma(y)\right)^{1/p}.$$

This inequality holds for every x for which the right-hand side is finite. Since we assume $F \in L^p_{\text{sym}}(S \times S, d\mu)$, the above inequality holds for σ-a.e. $x \in S$. This clearly implies (7.9). Finally, the proposition follows from the combination of (7.1) and (7.9). $\square$

For $g \in L^2(S, d\sigma)$ and $(k,j) \in I$, in addition to the $J(g;k,j)$ introduced in Section 6, let us also define

$$J_2(g;k,j) = \left\{\frac{1}{\sigma(C_{k,j})}\int_{C_{k,j}} |g - g_{C_{k,j}}|^2 d\sigma\right\}^{1/2}.$$

By the Cauchy-Schwarz inequality, $J(g;k,j) \le J_2(g;k,j)$ for all $g \in L^2(S, d\sigma)$ and $(k,j) \in I$. Also, it is obvious that there is a constant C such that

$$2^{4nk}\iint_{D_{k,j}} |g(y) - g(x)|^2 d\sigma(x)d\sigma(y) \le CJ_2^2(g;k,j)$$

for all $g \in L^2(S, d\sigma)$ and $(k,j) \in I$. Now comes the main step in the proof of the upper bound in Theorem 1.3:

Proposition 7.2. *Let $2 < p < \infty$. Then there is a constant $C_{7.2} = C_{7.2}(p,n)$ such that*

$$\|[P, M_g]\|_p^+ \le C_{7.2}\Phi_p^+(\{J_2(g;k,j)\}_{(k,j)\in I}) \tag{7.13}$$

for every $g \in L^2(S, d\sigma)$.

Proof. Given $2 < p < \infty$, we fix a $\lambda \in (p, \infty)$ for the proof. Let $g \in L^2(S, d\sigma)$ also be given. Then we only need to consider the case where $\Phi_p^+(\{J_2(g; k, j)\}_{(k,j)\in I}) < \infty$, for otherwise (7.13) holds for trivial reasons. Since $J(g; k, j) \le J_2(g; k, j)$ for every $(k, j) \in I$, it follows from Proposition 6.4 that

$$\Phi_p^+(\{V_\lambda(g; k, j)\}_{(k,j)\in I}) \le C_{6.4}\Phi_p^+(\{J_2(g; k, j)\}_{(k,j)\in I}). \tag{7.14}$$

Let us estimate $N_{[P,M_g]}(t)$, $t > 0$. This is where interpolation comes in.

The ideal is to decompose the commutator in the form $[P, M_g] = A_t + B_t$ and take advantage of the inequality

$$N_{[P,M_g]}(t) \le N_{A_t}(t/2) + N_{B_t}(t/2).$$

We will then estimate $N_{A_t}(t/2)$ by Proposition 7.1 and estimate $N_{B_t}(t/2)$ by using the Hilbert-Schmidt norm $\|B_t\|_2$. But before any estimates, we need to define A_t and B_t first, which has to be done carefully.

For convenience let us write

$$R = 2^{1/p}\frac{p}{p-1}C_{6.4}\Phi_p^+(\{J_2(g; k, j)\}_{(k,j)\in I}).$$

By (7.14) and Lemma 5.6, there is a bijection $\pi : \mathbf{N} \to I$ such that

$$V_\lambda(g; \pi(i)) \le R/i^{1/p} \quad \text{for every} \quad i \in \mathbf{N}. \tag{7.15}$$

Let $I(t) = \{\pi(i) : 1 \le i < (R/t)^p\}$. For each $k \ge 0$, let

$$E_k = \{(x, y) \in S \times S : 2^{-k} \le d(x, y) < 2^{-k+1}\}.$$

Then $E_k \subset \cup_{j=1}^{m(k)}(B_{k,j} \times B_{k,j}) = \cup_{j=1}^{m(k)} D_{k,j}$, $k \ge 0$. For each $k \ge 0$, we define

$$W_k(t) = \cup_{(k,j)\in I(t)} D_{k,j}.$$

Finally, we define

$$F(t) = \cup_{k=0}^\infty \{E_k \backslash W_k(t)\} \quad \text{and} \quad W(t) = \cup_{k=0}^\infty \{E_k \cap W_k(t)\}.$$

Then $F(t) \cap W(t) = \emptyset$ and $F(t) \cup W(t) = (S \times S)\backslash\{(x, x) : x \in S\}$. Now we let A_t and B_t be the integral operators on $L^2(S, d\sigma)$ with the kernel functions

$$\chi_{F(t)}(x, y)\frac{g(y) - g(x)}{(1 - \langle x, y\rangle)^n} \quad \text{and} \quad \chi_{W(t)}(x, y)\frac{g(y) - g(x)}{(1 - \langle x, y\rangle)^n}$$

respectively. We first estimate $N_{A_t}(t/2)$.

For each k, since $E_k \subset \cup_{j=1}^{m(k)} D_{k,j}$, we have $E_k \backslash W_k(t) \subset \cup_{(k,j)\in I\backslash I(t)} \{E_k \cap D_{k,j}\}$. Hence

$$\begin{aligned}\iint_{F(t)} \frac{|g(y)-g(x)|^\lambda}{|1-\langle x,y\rangle|^{2n}} d\sigma(x)d\sigma(y) &\leq \sum_{(k,j)\in I\backslash I(t)} \iint_{E_k\cap D_{k,j}} \frac{|g(y)-g(x)|^\lambda}{|1-\langle x,y\rangle|^{2n}} d\sigma(x)d\sigma(y) \\ &\leq \sum_{(k,j)\in I\backslash I(t)} 2^{4nk} \iint_{D_{k,j}} |g(y)-g(x)|^\lambda d\sigma(x)d\sigma(y) = \sum_{(k,j)\in I\backslash I(t)} \{V_\lambda(g;k,j)\}^\lambda \\ &= \sum_{i\geq (R/t)^p} \{V_\lambda(g;\pi(i))\}^\lambda \leq \sum_{i\geq (R/t)^p} (R/i^{1/p})^\lambda \qquad \text{(by (7.15))} \\ &\leq R^\lambda \cdot C_1(\max\{1,R/t\})^{p(1-(\lambda/p))} = C_1 R^\lambda (\max\{1,R/t\})^{p-\lambda}.\end{aligned}$$

Form the definition of $F(t)$ it is clear that $(x,y)\in F(t)$ if and only if $(y,x)\in F(t)$. Thus we can apply Proposition 7.1 to obtain

$$\begin{aligned}N_{A_t}(t/2) &\leq C_{7.1}(2/t)^\lambda \iint_{F(t)} \frac{|g(y)-g(x)|^\lambda}{|1-\langle x,y\rangle|^{2n}} d\sigma(x)d\sigma(y) \\ &\leq C_{7.1}(2/t)^\lambda \cdot C_1 R^\lambda (\max\{1,R/t\})^{p-\lambda} \leq 2^\lambda C_1 C_{7.1} R^p t^{-p},\end{aligned} \tag{7.16}$$

where the last $\leq$ uses the assumption $\lambda > p$.

To estimate $N_{B_t}(t/2)$, note that

$$\begin{aligned}\|B_t\|_2^2 &= \iint_{W(t)} \frac{|g(y)-g(x)|^2}{|1-\langle x,y\rangle|^{2n}} d\sigma(x)d\sigma(y) \leq \sum_{(k,j)\in I(t)} \iint_{E_k\cap D_{k,j}} \frac{|g(y)-g(x)|^2}{|1-\langle x,y\rangle|^{2n}} d\sigma(x)d\sigma(y) \\ &\leq \sum_{(k,j)\in I(t)} 2^{4nk} \iint_{D_{k,j}} |g(y)-g(x)|^2 d\sigma(x)d\sigma(y) \leq C \sum_{(k,j)\in I(t)} J_2^2(g;k,j).\end{aligned} \tag{7.17}$$

If we write

$$R' = 2^{1/p} \frac{p}{p-1} \Phi_p^+(\{J_2(g;k,j)\}_{(k,j)\in I}),$$

then by Lemma 5.6 there is a bijection $b : \mathbf{N} \to I$ such that

$$J_2(g;b(i)) \leq R'/i^{1/p} \quad \text{for every} \quad i\in\mathbf{N}.$$

Since $\operatorname{card}(I(t)) < (R/t)^p$, continuing with (7.17), we have

$$\begin{aligned}\|B_t\|_2^2 &\leq C \sum_{b(i)\in I(t)} J_2^2(g;b(i)) \leq C \sum_{b(i)\in I(t)} (R'/i^{1/p})^2 \leq C \sum_{1\leq i<(R/t)^p} (R'/i^{1/p})^2 \\ &\leq C_2(R')^2 \cdot (R/t)^{p(1-(2/p))} = C_2(R')^2 R^{p-2} t^{-p+2}.\end{aligned}$$

Therefore

$$N_{B_t}(t/2) \leq (2/t)^2 \|B_t\|_2^2 \leq 4C_2(R')^2 R^{p-2} t^{-p}.$$

Combining this with (7.16), we have

$$N_{[P,M_g]}(t) \le \{2^\lambda C_1 C_{7.1} R^p + 4C_2 (R')^2 R^{p-2}\} t^{-p} = C_3 \{\Phi_p^+ (\{J_2(g;k,j)\}_{(k,j)\in I})\}^p t^{-p}.$$

If $\nu \in \mathbf{N}$ and $t_\nu > 0$ are such that $N_{[P,M_g]}(t_\nu) < \nu$, then $s_\nu([P,M_g]) \le t_\nu$. Hence it follows from the above inequality that the s-numbers of $[P,M_g]$ satisfy the condition

$$s_\nu([P,M_g]) \le (2C_3)^{1/p} \Phi_p^+ (\{J_2(g;k,j)\}_{(k,j)\in I}) \nu^{-1/p}$$

for every $\nu \in \mathbf{N}$. Therefore

$$\|[P,M_g]\|_p^+ \le (2C_3)^{1/p} \Phi_p^+ (\{J_2(g;k,j)\}_{(k,j)\in I}).$$

This completes the proof of the proposition. □

8. Upper bound

Proposition 7.2 represents the essential part of the proof of the upper bound in Theorem 1.3. The remaining task in the proof of the upper bound is to bring Bergman lattice into the picture, which is a rather routine exercise. The reader will find that the material in this section closely resembles the second half of Section 2 in [18]. As such, an omission of the proofs in this section is well justified. But we decide to retain the proofs for those readers who are interested in details. Those who are not interested in details may wish to go directly to the proof of the upper bound at the end of the section.

By [14,Theorem 2.2.2], we have

$$1 - |\varphi_w(z)|^2 = \frac{(1-|w|^2)(1-|z|^2)}{|1-\langle z,w\rangle|^2}, \tag{8.1}$$

$z, w \in \mathbf{B}$. For each $(k,j) \in I$, we define

$$w(k,j) = (1-2^{-2k})^{1/2} u_{k,j},$$

which is an element in the set $T_{k,j}$ defined by (3.7).

Lemma 8.1. *Given any $0 < b < \infty$, there is a constant $C_{8.1}$ which depends only on b and n such that if $z \in \mathbf{B}$ and $(k,j) \in I$ satisfy the condition $w(k,j) \in D(z,b)$, then*

$$J_2(g;k,j) \le C_{8.1} \mathrm{Var}^{1/2}(g;z)$$

for every $g \in L^2(\mathbf{B}, dv)$.

Proof. A review of the definitions of J_2 and $\mathrm{Var}^{1/2}$ tells us that it suffices to show that there is a C such that the inequality

$$\frac{1}{\sigma(C_{k,j})} \chi_{C_{k,j}} \le C|k_z|^2 \tag{8.2}$$

holds on S whenever $(k,j) \in I$ and $z \in \mathbf{B}$ satisfy the condition $w(k,j) \in D(z,b)$. Since $\sigma(C_{k,j}) \geq c2^{-2nk}$, (8.2) will follow if we can show that there are $0 < c_1 < \infty$ and $0 < C_2 < \infty$ such that for $(k,j) \in I$ and $z \in \mathbf{B}$ satisfying the condition $w(k,j) \in D(z,b)$, we have

$$1 - |z|^2 \geq c_1 2^{-2k} \quad \text{and} \quad |1 - \langle \zeta, z\rangle| \leq C_2 2^{-2k} \quad \text{for each} \quad \zeta \in C_{k,j}. \tag{8.3}$$

To prove this, suppose that $D(z,b)$ contains some $w(k,j)$. Suppose that $1 - |z|^2 < \epsilon 2^{-2k}$ for some $\epsilon > 0$. Then by (8.1) we have

$$\begin{aligned} 1 - |\varphi_{w(k,j)}(z)| \leq 1 - |\varphi_{w(k,j)}(z)|^2 &\leq \frac{(1 - |w(k,j)|^2) \cdot \epsilon 2^{-2k}}{|1 - \langle z, w(k,j)\rangle|^2} \leq \frac{(1 - |w(k,j)|^2) \cdot \epsilon 2^{-2k}}{(1 - |w(k,j)|)^2} \\ &= \frac{(1 + |w(k,j)|)^2 \cdot \epsilon 2^{-2k}}{1 - |w(k,j)|^2} \leq 4\epsilon. \end{aligned}$$

Hence $b \geq \beta(w(k,j), z) \geq (1/2)\log\{(4\epsilon)^{-1}\}$. Solving this inequality, we find that $\epsilon \geq (1/4)e^{-2b}$. Therefore if we set $c_1 = (1/8)e^{-2b}$, then $1 - |z|^2 \geq c_1 2^{-2k}$.

To prove the other half of (8.3), we need an upper bound for $1 - |z|$. Note that $|1 - \langle z, w(k,j)\rangle| \geq 1 - |z|$. Using (8.1) again, we have

$$1 - |\varphi_{w(k,j)}(z)| \leq 1 - |\varphi_{w(k,j)}(z)|^2 \leq \frac{2(1 - |z|) \cdot (1 - |w(k,j)|^2)}{|1 - \langle z, w(k,j)\rangle|^2} \leq \frac{2 \cdot 2^{-2k}}{1 - |z|}.$$

Thus $b \geq (1/2)\log\{(1 - |z|)/(2 \cdot 2^{-2k})\}$, which implies $1 - |z| \leq 2e^{2b}2^{-2k} = C_3 2^{-2k}$.

Let us write $z = |z|\xi$, where $\xi \in S$. We need an upper bound for $d(u_{k,j}, \xi)$. Suppose that $|1 - \langle u_{k,j}, \xi\rangle| > A2^{-2k}$ for some $A > 0$. Then $2|1 - \langle w(k,j), z\rangle| \geq |1 - \langle u_{k,j}, \xi\rangle| \geq A2^{-2k}$. Another application of (8.1) now gives us

$$1 - |\varphi_{w(k,j)}(z)| \leq \frac{(1 - |w(k,j)|^2) \cdot 2(1 - |z|)}{|1 - \langle z, w(k,j)\rangle|^2} \leq \frac{2^{-2k} \cdot 2C_3 2^{-2k}}{((1/2)A2^{-2k})^2} = \frac{8C_3}{A^2}.$$

Hence $b \geq (1/2)\log\{A^2/(8C_3)\}$. That is, $A \leq 2\sqrt{2}C_3^{1/2}e^b$. Thus if we set $C_4 = 4C_3^{1/2}e^b$, then $|1 - \langle u_{k,j}, \xi\rangle| \leq C_4 2^{-2k}$. That is, $d(u_{k,j}, \xi) \leq C_4^{1/2}2^{-k}$.

Let $\zeta \in C_{k,j} = B(u_{k,j}, 2^{-k+3})$. Then $d(\zeta, \xi) \leq d(\zeta, u_{k,j}) + d(u_{k,j}, \xi) \leq (8 + C_4^{1/2})2^{-k}$. Thus if we set $C_5 = (8 + C_4^{1/2})^2$, then $|1 - \langle \zeta, \xi\rangle| \leq C_5 2^{-2k}$. Hence

$$|1 - \langle \zeta, z\rangle| \leq (1 - |z|) + |1 - \langle \zeta, \xi\rangle| \leq C_3 2^{-2k} + C_5 2^{-2k}.$$

This proves the second half of (8.3) and completes the proof of the lemma. □

Lemma 8.2. *Given any* $0 < b < \infty$, *there is a natural number* N *such that for every* $z \in \mathbf{B}$, *we have* $\operatorname{card}\{(k,j) \in I : w(k,j) \in D(z,b)\} \leq N$.

Proof. In the proof of Lemma 8.1 we showed that if $w(k,j) \in D(z,b)$, then $(c_1/2)2^{-2k} \le 1-|z| \le C_3 2^{-2k}$, where c_1 and C_3 depend only on b. In other words, there is an $m \in \mathbf{N}$ which depends only on b such that

$$2^{-2(k+m)} \le 1-|z| \le 2^{-2(k-m)}$$

if $w(k,j) \in D(z,b)$. If $w(k',j')$ also belongs to $D(z,b)$, then $2^{-2(k+m)} \le 1-|z| \le 2^{-2(k'-m)}$. Solving the inequality, we find that $k' \le k+2m$ if $w(k,j), w(k',j') \in D(z,a)$.

As in the previous proof, write $z = |z|\xi$, where $\xi \in S$. The previous proof tells us that $d(u_{k,j},\xi) \le C_4^{1/2}2^{-k}$ if $w(k,j) \in D(z,b)$. Hence if both $w(k,j)$ and $w(k,\nu)$ belong to $D(z,b)$, then $d(u_{k,j},u_{k,\nu}) \le 2C_4^{1/2}2^{-k}$. By (3.5) and (2.2), there is an N_1 which is determined by n and C_4 such that

$$\operatorname{card}\{j \in \{1,\dots,m(k)\} : w(k,j) \in D(z,b)\} \le N_1$$

for all $k \ge 0$ and $z \in \mathbf{B}$. Combining this with the conclusion of the preceding paragraph, we see that $\operatorname{card}\{(k,j) \in I : w(k,j) \in D(z,b)\} \le (2m+1)\cdot N_1$. □

Recall that Lemma 6.1 concerns the particular symmetric gauge function Φ_p^-, $1 < p < \infty$. In contrast, our next lemma provides a more crude estimate, but has the advantage that it holds for all symmetric gauge functions.

Lemma 8.3. [18,Lemma 2.2] *Suppose that X and Y are countable sets and that N is a natural number. Suppose that $T : X \to Y$ is a map that is at most N-to-1. That is, for every $y \in Y$, $\operatorname{card}\{x \in X : T(x) = y\} \le N$. Then for every set of real numbers $\{b_y\}_{y\in Y}$ and every symmetric gauge function Φ, we have*

$$\Phi(\{b_{T(x)}\}_{x\in X}) \le N\Phi(\{b_y\}_{y\in Y}).$$

Proposition 8.4. *Given any positive number $0 < b < \infty$, there is a constant $C_{8.4}$ which depends only on b and n such that if Γ is a countable subset of $\mathbf{B}$ with the property that $\cup_{z\in\Gamma}D(z,b) = \mathbf{B}$, then*

$$\Phi(\{J_2(g;k,j)\}_{(k,j)\in I}) \le C_{8.4}\Phi(\{\mathrm{Var}^{1/2}(g;z)\}_{z\in\Gamma})$$

for every $g \in L^2(S,d\sigma)$ and every symmetric gauge function Φ.

Proof. Given b, let N be the natural number provided by Lemma 8.2. Suppose that Γ has the property that $\cup_{z\in\Gamma}D(z,b) = \mathbf{B}$. Then for each $(k,j) \in I$, pick a $z(k,j) \in \Gamma$ such that $w(k,j) \in D(z(k,j),b)$. By Lemma 8.1, for each $g \in L^2(S,d\sigma)$ and each symmetric gauge function Φ we have

$$\Phi(\{J_2(g;k,j)\}_{(k,j)\in I}) \le C_{8.1}\Phi(\{\mathrm{Var}^{1/2}(g;z(k,j))\}_{(k,j)\in I}).$$

Lemma 8.2 tells us that the map $(k,j) \mapsto z(k,j)$ is at most N-to-1. Thus, by Lemma 8.3,

$$\Phi(\{\mathrm{Var}^{1/2}(g;z(k,j))\}_{(k,j)\in I}) \le N\Phi(\{\mathrm{Var}^{1/2}(g;z)\}_{z\in\Gamma}).$$

Hence the constant $C_{8.4} = NC_{8.1}$ suffices for the proposition. $\square$

Proof of the upper bound in Theorem 1.3. Given an $f \in L^2(S, d\sigma)$, write $g = f - Pf$. Then $H_f = H_g$. Let $2 < p < \infty$ and $b > 0$. Let Γ be a countable subset of $\mathbf{B}$ such that $\cup_{z\in\Gamma} D(z,b) = \mathbf{B}$. Applying Propositions 7.2 and 8.4, we have

$$\begin{aligned}\|H_f\|_p^+ = \|H_g\|_p^+ &\le \|[P, M_g]\|_p^+ \le C_{7.2}\Phi_p^+(\{J_2(g;k,j)\}_{(k,j)\in I}) \\ &\le C_{7.2}C_{8.4}\Phi_p^+(\{\mathrm{Var}^{1/2}(g;z)\}_{z\in\Gamma}) = C_{7.2}C_{8.4}\Phi_p^+(\{\mathrm{Var}^{1/2}(f - Pf;z)\}_{z\in\Gamma}).\end{aligned}$$

This completes the proof of Theorem 1.3. $\square$

References

1. J. Arazy, S. Fisher and J. Peetre, *Hankel operators on weighted Bergman spaces*, Amer. J. Math. **110** (1988), 989-1054.
2. D. Békollé, C. Berger, L. Coburn and K. Zhu, BMO *in the Bergman metric on bounded symmetric domains*, J. Funct. Anal. **93** (1990), 310-350.
3. J. Bergh and J. Löström, *Interpolation spaces. An introduction*, Springer-Verlag, Berlin, 1976.
4. R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. **103** (1976), 611-635.
5. Q. Fang and J. Xia, *Schatten class membership of Hankel operators on the unit sphere*, J. Funct. Anal. **257** (2009), 3082-3134.
6. M. Feldman and R. Rochberg, *Singular value estimates for commutators and Hankel operators on the unit ball and the Heisenberg group*, Analysis and partial differential equations, 121-159, Lecture Notes in Pure and Appl. Math., **122**, Dekker, New York, 1990.
7. I. Gohberg and M. Krein, *Introduction to the theorey of linear nonselfadjoint operators*, Amer. Math. Soc. Translations of Mathematical Monographs **18**, Providence, 1969.
8. S. Janson and T. Wolff, *Schatten classes and commutators of singular integral operators*, Ark. Mat. **20** (1982), 301-310.
9. H. Li, *Hankel operators on the Bergman spaces of strongly pseudoconvex domains*, Integral Equations Operator Theory **19** (1994), 458-476.
10. H. Li and D. Luecking, *Schatten class of Hankel operators and Toeplitz operators on the Bergman space of strongly pseudoconvex domains*, in "Contemporary Mathematics," Vol. 185 (Multivariable Operator Theory, R. Curto et al eds.), Amer. Math. Soc., Providence, RI, 1995.
11. D. Luecking, *Characterization of certain classes of Hankel operators on the Bergman spaces of the unit disc*, J. Funct. Anal. **110** (1992), 247-271.
12. V. Peller, *Hankel operators of class S_p and their applications* (*rational approximation, Gaussian processes, the problem of majorization of operators*), Math. USSR Sbornik **41** (1982), 443-479.
13. V. Peller, *Hankel operators and their applications*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2003.
14. W. Rudin, *Function theory in the unit ball of* $\mathbf{C}^n$, Springer-Verlag, New York, 1980.
15. J. Xia, *Bounded functions of vanishing mean oscillation on compact metric spaces*, J. Funct. Anal. **209** (2004), 444-467.

16. J. Xia, *Boundedness and compactness of Hankel operators on the sphere*, J. Funct. Anal. **255** (2008), 25-45.
17. J. Xia, *On certain quotient modules of the Bergman module*, Indiana Univ. Math. J. **57** (2008), 545-575.
18. J. Xia, *Bergman commutators and norm ideals*, J. Funct. Anal. **263** (2012), 988-1039.
19. K. Zhu, *Operator theory in function spaces.* Second edition. Mathematical Surveys and Monographs, **138**, American Mathematical Society, Providence, RI, 2007.

Department of Mathematics and Computer Science, Bronx Community College, CUNY, Bronx, NY 10453

E-mail: quanlei.fang@bcc.cuny.edu

Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260

E-mail: jxia@acsu.buffalo.edu