

# WEIGHTED $L^2$ INEQUALITIES FOR SQUARE FUNCTIONS

RODRIGO BAÑUELOS AND ADAM OSEKOWSKI

ABSTRACT. Using Bellman function approach, we present new proofs of weighted  $L^2$  inequalities for square functions, with the optimal dependence on the  $A_2$  characteristics of the weight and further explicit constants. We study the estimates both in the analytic and probabilistic context, and, as application, obtain related estimates for the classical Lusin and Littlewood-Paley square functions.

## 1. INTRODUCTION

Square function inequalities play an important role in both classical and noncommutative probability theory, harmonic analysis, potential theory and many other areas of mathematics. The purpose of this paper is to establish sharp bounds in the dyadic case, which are closely related to the works of Bollobás [10], Davis [15], John and Nirenberg [23], Littlewood [27], Marcinkiewicz [29], Paley [36], Slavín and Vasyunin [42], Wang [48] and many others.

Let us start the paper with introducing some background and notation. In what follows, the interval  $[0, 1]$  will be denoted by  $\mathcal{J}$ . Let  $(h_n)_{n \geq 0}$  be the Haar system on  $\mathcal{J}$ , that is, the family of functions given by

$$\begin{aligned} h_0 &= \chi_{[0,1]}, & h_1 &= \chi_{[0,1/2)} - \chi_{[1/2,1)}, \\ h_2 &= \chi_{[0,1/4)} - \chi_{[1/4,1/2)}, & h_3 &= \chi_{[1/2,3/4)} - \chi_{[3/4,1)}, \\ h_4 &= \chi_{[0,1/8)} - \chi_{[1/8,1/4)}, & h_5 &= \chi_{[1/4,3/8)} - \chi_{[3/8,1/2)}, \end{aligned}$$

and so on. For any dyadic subinterval  $I$  of  $\mathcal{J}$  and any integrable function  $\varphi : \mathcal{J} \rightarrow \mathbb{R}$ , we will write  $\langle \varphi \rangle_I$  for the average of  $\varphi$  over  $I$ : that is,  $\langle \varphi \rangle_I = \frac{1}{|I|} \int_I \varphi$  (unless stated otherwise, the integration is with respect to Lebesgue measure). Furthermore, for any such  $\varphi$  and any nonnegative integer  $n$ , we will write

$$\varphi_n = \sum_{k=0}^{2^n-1} \frac{1}{|I_k|} \int_{I_k} \varphi(s) h_k(s) ds h_k$$

for the projection of  $\varphi$  on the subspace generated by the first  $2^n$  Haar functions ( $I_k$  is the support of  $h_k$ ). We define the dyadic square function of  $\varphi$  by the formula

$$S(\varphi)(x) = \left( \sum \left| \frac{1}{|I_n|} \int_{I_n} \varphi(s) h_n(s) ds \right|^2 \right)^{1/2},$$

where the summation runs over all nonnegative integers  $n$  such that  $x \in I_n$ .

---

2010 *Mathematics Subject Classification*. Primary: 42B20. Secondary: 46E30.

*Key words and phrases*. Square function, dyadic, Bellman function, best constants.

R. Bañuelos is supported in part by NSF Grant # 0603701-DMS.

A. Osękowski is supported in part by the NCN grant DEC-2014/14/E/ST1/00532.

The inequalities comparing the sizes of  $\varphi$  and its square function  $S(\varphi)$  are of importance in analysis and probability, and have been studied intensively in the literature. A classical result of Paley [36] and Marcinkiewicz [33] states that there are finite absolute constants  $c_p$  ( $0 < p < \infty$ ) and  $C_p$  ( $1 < p < \infty$ ), such that for any  $\varphi : \mathcal{J} \rightarrow \mathbb{R}$ ,

$$(1.1) \quad \|\varphi\|_{L^p(\mathcal{J})} \leq c_p \|S(\varphi)\|_{L^p(\mathcal{J})}$$

and

$$(1.2) \quad \|S(\varphi)\|_{L^p(\mathcal{J})} \leq C_p \|\varphi\|_{L^p(\mathcal{J})}.$$

The question about the optimal values of  $c_p$  and  $C_p$  was studied by Davis [15]. For  $0 < p < \infty$ , let  $\nu_p$  denote the smallest positive zero of a confluent hypergeometric function  $M_p$  (see Abramovitz and Stegun [1] for details). Using a related estimate for continuous-time martingales and Skorokhod embedding theorems, Davis [15] showed that if  $0 < p \leq 2$ , then the best choice for  $c_p$  is  $\nu_p$ , while for  $p \geq 2$ , the optimal value of  $C_p$  is  $\nu_p^{-1}$ . See also Wang [48] for the vector-valued analogues of these results.

In recent years, a question about the weighted version of (1.1) and (1.2) gathered a lot of interest. In what follows, the word ‘‘weight’’ refers to an integrable, positive function on  $\mathcal{J}$ , which will usually be denoted by  $w$ . Given  $p \in (1, \infty)$ , we say that  $w$  belongs to the dyadic Muckenhoupt  $A_p$  class (or, in short, that  $w$  is an  $A_p$  weight), if the  $A_p$  characteristics  $[w]_{A_p}$ , given by

$$[w]_{A_p} := \sup_I \left( \frac{1}{|I|} \int_I w \right) \left( \frac{1}{|I|} \int_I w^{-1/(p-1)} \right)^{p-1},$$

is finite. Here the supremum runs over all dyadic subintervals of  $\mathcal{J}$ . One can also define the appropriate versions of this condition for  $p = 1$  and  $p = \infty$ , by passing above with  $p$  to the appropriate limit (see e.g. [18], [19]). However, we omit the details, as in this paper we will be mainly concerned with the case  $1 < p < \infty$ . The condition  $A_p$  arises naturally in the study of weighted estimates for the Hardy-Littlewood maximal operator, as Muckenhoupt showed in [33].

Coming back to square function estimates, the first weighted bound in this setting is due to Buckley [11], who showed the  $L^2$  inequality

$$(1.3) \quad \|S(\varphi)\|_{L_w^2(\mathcal{J})} \leq C [w]_{A_2}^{3/2} \|\varphi\|_{L_w^2(\mathcal{J})},$$

with  $C$  being a universal constant. Here, of course, the weighted  $L^2$  norm is given by

$$\|\varphi\|_{L_w^2(\mathcal{J})} = \left( \int_{\mathcal{J}} \varphi^2 w \right)^{1/2}.$$

Can the exponent  $3/2$  in (1.3) be decreased? This question was studied by Hukovic [20] and Hukovic, Treil and Volberg [21]. It turns out that the sharp dependence is linear, i.e., the best exponent is 1. This result was later reproved by Wittwer [49] and Petermichl and Pott in [37] using a different approach. Actually, the latter paper contains also the proof of the reverse inequality

$$\|\varphi\|_{L_w^2(\mathcal{J})} \leq C [w]_{A_2}^{1/2} \|S(\varphi)\|_{L_w^2(\mathcal{J})},$$

in which the exponent  $1/2$  is also optimal. A considerable extension of these results was obtained recently by Cruz-Uribe, Martell and Pérez in [13], who showed the weighted  $L^p$  bound

$$\|S(\varphi)\|_{L_w^p(\mathcal{J})} \leq C [w]_{A_p}^{\max\{(p-1)^{-1}, 1/2\}} \|\varphi\|_{L_w^p(\mathcal{J})}$$

and proved that the exponent  $\max\{(p-1)^{-1}, 1/2\}$  is the best.

One of our objectives is to give a yet another proof of the weighted  $L^2$  estimate for square functions. Our reasoning will rest on the construction of certain special functions which enjoy appropriate majorization and concavity properties. This type of approach, called the Bellman function technique, originates from the theory of optimal stochastic control and has turned out to be very efficient in various problems in analysis and probability.

One of our main results is the following.

**Theorem 1.1.** *Suppose that  $w$  is an  $A_2$  weight and  $\varphi$  is a function belonging to  $L^2_w(\mathcal{I})$ . Then we have the estimates*

$$(1.4) \quad \|\varphi\|_{L^2_w(\mathcal{I})} \leq (160[w]_{A_2})^{1/2} \|S(\varphi)\|_{L^2_w(\mathcal{I})}$$

and

$$(1.5) \quad \|S(\varphi)\|_{L^2_w(\mathcal{I})} \leq 8\sqrt{2}[w]_{A_2} \|\varphi\|_{L^2_w(\mathcal{I})}.$$

Furthermore,

$$(1.6) \quad \|S(\varphi)\|_{L^2_w(\mathcal{I})} \leq \inf_{1 < r < 2} \left( \frac{2r}{2-r} [w]_{A_r} \right)^{1/2} \|\varphi\|_{L^2_w(\mathcal{I})}.$$

The reason why we have included (1.6) in the above statement is that this estimate implies the weighed  $L^2$  bound with the linear dependence on  $[w]_{A_2}$ , and hence can be regarded as an improvement of (1.5). To see the implication, recall the following classical fact, due to Coifman and Fefferman [12].

**Lemma 1.2.** *There is a constant  $\kappa_p$  depending only on  $p$  such that the following holds. If  $w$  is an  $A_p$  weight ( $1 < p < \infty$ ) on an interval, then  $w$  is an  $A_{p-\varepsilon}$  weight, where  $\varepsilon = \kappa_p^{-1} [w]_{A_p}^{-1/(p-1)}$ . Moreover, we have  $[w]_{A_{p-\varepsilon}} \leq \kappa_p [w]_{A_p}$ .*

We apply this lemma with  $p = 2$ . Taking  $r = 2 - \varepsilon = 2 - \kappa_2^{-1} [w]_{A_2}^{-1}$ , the estimate (1.6) yields

$$\|S(\varphi)\|_{L^2_w(\mathcal{I})} \leq \left( \frac{2r}{2-r} [w]_{A_r} \right)^{1/2} \|\varphi\|_{L^2_w(\mathcal{I})} \leq 2\kappa_2 [w]_{A_2} \|\varphi\|_{L^2_w(\mathcal{I})},$$

as desired.

We have organized the paper as follows. In the next section we introduce the Bellman functions corresponding to the estimates (1.4), (1.5) and (1.6), and study their properties. §3 is devoted to the proof of Theorem 1.1. In §4, we establish an appropriate probabilistic analogue of Theorem 1.1 for continuous time martingales and then, in §5 show how this yields similar results for the classical Lusin and Littlewood-Paley square functions. In section §6 we further elaborate on extensions to more general Markovian semigroups.

## 2. SPECIAL FUNCTIONS

Throughout this section,  $c > 1$  is a fixed parameter. For any  $1 < r < \infty$ , the symbol  $\Omega_c^r$  will denote the hyperbolic domain given by

$$\Omega_c^r = \{(w, v) \in \mathbb{R}_+ \times \mathbb{R}_+ : 1 \leq wv^{r-1} \leq c\}.$$

**2.1. Bellman function corresponding to (1.4).** The key role in the proof of the estimate (1.4) is played by the function  $B_c : \mathbb{R} \times [0, \infty) \times \Omega_c^2 \rightarrow \mathbb{R}$ , given by

$$B_c(x, y, w, v) = x^2 w \varphi(wv) - 40cyw,$$

where

$$\varphi(t) = 2 - \frac{1}{t} - \frac{\ln t}{2c}, \quad t \in [1, c].$$

In what follows, usually we will skip the lower index and write  $B$  instead of  $B_c$ , but keep in mind that the function does depend on the parameter  $c$ . Let us study some simple majorization properties of  $B$ . We start with the trivial observation that  $\varphi(t) \leq 2$  for all  $t$ , which implies the bound

$$(2.1) \quad B(x, x^2, w, v) \leq 2x^2 w - 40cx^2 w \leq 0 \quad \text{for all } x \in \mathbb{R}, (w, v) \in \Omega_c^2.$$

Next, note that for any  $t \in [1, c]$  we have  $\varphi(t) \geq 2 - 1 - \frac{1}{2} = \frac{1}{2}$ , and hence

$$(2.2) \quad B(x, y, w, v) \geq \frac{1}{2}w(x^2 - 80cy) \quad \text{for all } (x, y, w, v) \in \mathbb{R} \times [0, \infty) \times \Omega_c^2.$$

We turn our attention to the crucial property of  $B$ . It can be regarded as a concavity-type condition.

**Lemma 2.1.** *Suppose that  $(x, y, w, v) \in \mathbb{R} \times [0, \infty) \times \Omega_c^2$  is a given point and assume further that  $e, f$  are real numbers such that the line segment with endpoints  $(w \pm e, v \pm f)$  is entirely contained in  $\Omega_c^2$ . Then for any  $d \in \mathbb{R}$  we have*

$$(2.3) \quad \begin{aligned} & 2B(x, y, z, w) \\ & \geq B(x - d, y + d^2, w - e, v - f) + B(x + d, y + d^2, w + e, v + f). \end{aligned}$$

*Proof.* Introduce the function  $b : \mathbb{R} \times \Omega_c^2 \rightarrow [0, \infty)$ , given by  $b(x, w, v) = x^2 w \varphi(wv)$ . Of course, we have the identity

$$B(x, y, w, v) = b(x, w, v) - 40cyw.$$

Since  $-c(y + d^2)(w - e) - c(y + d^2)(w + e) = -2cyw - 2cd^2w$ , we see that the claim is equivalent to

$$(2.4) \quad 2b(x, w, v) \geq b(x - d, w - e, v - f) + b(x + d, w + e, v + f) - 80cd^2w.$$

To show this, we will prove that the matrix

$$(2.5) \quad \mathbb{A}(x, w, v) = D^2 b(x, w, v) - \begin{bmatrix} 80cw & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is nonpositive-definite. To see how this yields (2.4), consider the function

$$F(t) = b(x + td, w + te, v + tf) - 40ct^2 d^2 w, \quad t \in [-1, 1].$$

Note that for each such  $t$ , the point  $(x + td, w + te, v + tf)$  lies in the domain of  $b$ : see the assumption in the statement of the lemma above. Denoting the vector  $(d, e, f)$  by  $\Delta$ , we

compute that

$$\begin{aligned}
F''(t) + F''(-t) &= (D^2b(x + td, w + te, v + tf)\Delta, \Delta) - 80c wd^2 \\
&\quad + (D^2b(x - td, w - te, v - tf)\Delta, \Delta) - 80c wd^2 \\
&= (D^2b(x + td, w + te, v + tf)\Delta, \Delta) - 80c(w + te)d^2 \\
&\quad + (D^2b(x - td, w - te, v - tf)\Delta, \Delta) - 80c(w - te)d^2 \\
&= (\mathbb{A}(x + td, w + te, v + tf)\Delta, \Delta) \\
&\quad + (\mathbb{A}(x - td, w - te, v - tf)\Delta, \Delta) \\
&\leq 0.
\end{aligned}$$

Consequently,  $F(1) + F(-1) \leq 2F(0)$ , which is precisely (2.4). To show that  $\mathbb{A}(x, w, v)$  is nonpositive-definite, we compute directly that the matrix is equal to

$$\begin{bmatrix} 2w\varphi(t) - 80cw & 2x\varphi + 2xt\varphi'(t) & 2xw^2\varphi'(t) \\ 2x\varphi(t) + 2xt\varphi'(t) & 2x^2v\varphi'(t) + x^2tv\varphi''(t) & 2x^2w\varphi'(t) + x^2tw\varphi''(t) \\ 2xw^2\varphi'(t) & 2x^2w\varphi'(t) + x^2tw\varphi''(t) & x^2w^3\varphi''(t) \end{bmatrix},$$

where  $t = wv$ . By well-known facts from linear algebra, it is enough to show that

$$(2.6) \quad x^2w^3\varphi''(t) \leq 0,$$

$$(2.7) \quad \det \begin{bmatrix} 2x^2v\varphi'(t) + x^2tv\varphi''(t) & 2x^2w\varphi'(t) + x^2tw\varphi''(t) \\ 2x^2w\varphi'(t) + x^2tw\varphi''(t) & x^2w^3\varphi''(t) \end{bmatrix} \geq 0$$

and

$$(2.8) \quad \det \mathbb{A}(x, w, v) \leq 0.$$

To establish (2.6), observe that  $t = wv \in [1, c]$ , by the definition of  $\Omega_c^2$ , and therefore

$$x^2w^3\varphi''(t) = -\frac{x^2w^3}{2ct^3}(4c - t) \leq 0.$$

The inequality (2.7) is equivalent to  $\varphi'(t)(2\varphi'(t) + t\varphi''(t)) \geq 0$ , and follows from the estimates

$$\varphi'(t) = \frac{1}{2ct^2}(2c - t) \geq 0, \quad 2\varphi'(t) + t\varphi''(t) = -\frac{1}{2ct} \leq 0.$$

Finally, we turn our attention to (2.8). Let us simplify the matrix  $\mathbb{A}$ , by carrying out some elementary operations. Dividing the second row and the second column by  $x$ , and then the third row and column by  $xw$ , we see that the determinant of  $\mathbb{A}$  has the same sign as

$$\det \begin{bmatrix} 2w\varphi(t) - 80cw & 2\varphi + 2t\varphi'(t) & 2w\varphi'(t) \\ 2\varphi(t) + 2t\varphi'(t) & 2v\varphi'(t) + tv\varphi''(t) & 2\varphi'(t) + t\varphi''(t) \\ 2w\varphi'(t) & 2\varphi'(t) + t\varphi''(t) & w\varphi''(t) \end{bmatrix}.$$

Next, multiply the third row by  $v$  and subtract it from the second row; then multiply the second row by  $w$  and subtract it from the first row. As the result, we obtain that the sign of the determinant of  $\mathbb{A}$  is the same as that of

$$\begin{aligned}
&\det \begin{bmatrix} -80cw & 2\varphi + 2t\varphi'(t) & 0 \\ 2\varphi(t) & 0 & 2\varphi'(t) \\ 2w\varphi'(t) & 2\varphi'(t) + t\varphi''(t) & w\varphi''(t) \end{bmatrix} \\
&= 4w \left[ (2(\varphi'(t))^2 - \varphi(t)\varphi''(t))(\varphi(t) + t\varphi'(t)) + 40c\varphi'(t)(2\varphi'(t) + t\varphi''(t)) \right].
\end{aligned}$$

However, we compute that

$$\begin{aligned}\varphi(t) + t\varphi'(t) &= 2 - \frac{\ln t}{2c} - \frac{1}{2c} \leq 2, \\ 2(\varphi'(t))^2 &= \varphi'(t) \cdot \frac{2c-t}{ct^2} \leq \frac{2\varphi'(t)}{t}\end{aligned}$$

and, since  $\varphi(t) \leq 2$ ,

$$-\frac{\varphi(t)\varphi''(t)}{\varphi'(t)} \leq \frac{2\left(\frac{2}{t^3} - \frac{1}{2ct^2}\right)}{\frac{1}{t^2} - \frac{1}{2ct}} \leq \frac{8}{t}.$$

Consequently,

$$\left(2(\varphi'(t))^2 - \varphi(t)\varphi''(t)\right)(\varphi(t) + t\varphi'(t)) \leq \frac{20\varphi'(t)}{t}$$

and since

$$40c\varphi'(t)(2\varphi'(t) + t\varphi''(t)) = -\frac{20\varphi'(t)}{t},$$

the inequality (2.8) is satisfied. This completes the proof.  $\square$

**2.2. Bellman function corresponding to (1.6).** The Bellman function associated with the  $A_r$ -estimate is slightly simpler. Let  $r$  be an arbitrary number belonging to  $(1, 2)$  and define  $B = B_{c,r} : \mathbb{R} \times [0, \infty) \times \Omega_c^r \rightarrow \mathbb{R}$  by

$$B(x, y, w, v) = yw - \frac{rc}{2-r} \frac{x^2}{v^{r-1}}.$$

As previously, we will first establish the appropriate majorizations for  $B$ . By the definition of  $\Omega_c^r$ , we have  $cv^{1-r} \geq w$  and hence

$$(2.9) \quad B(x, x^2, w, v) \leq x^2w \left(1 - \frac{r}{2-r}\right) \leq 0 \quad \text{for all } x \in \mathbb{R}, (w, v) \in \Omega_c^r,$$

where in the last bound we used the estimate  $r > 1$ . Furthermore, the inequality  $v^{1-r} \leq w$  implies

$$(2.10) \quad B(x, y, w, v) \geq yw - \frac{rc}{2-r} x^2w \quad \text{for all } (x, y, w, v) \in \mathbb{R} \times [0, \infty) \times \Omega_c^r.$$

We turn to the analogue of Lemma 2.1.

**Lemma 2.2.** *Suppose that  $(x, y, w, v) \in \mathbb{R} \times [0, \infty) \times \Omega_c^r$  is a given point and assume further that  $e, f$  are real numbers such that the line segment with endpoints  $(w \pm e, v \pm f)$  is entirely contained in  $\Omega_c^r$ . Then for any  $d \in \mathbb{R}$  we have*

$$(2.11) \quad \begin{aligned} &2B(x, y, z, w) \\ &\geq B(x-d, y+d^2, w-e, v-f) + B(x+d, y+d^2, w+e, v+f). \end{aligned}$$

*Proof.* Repeating the reasoning from the proof of Lemma 2.1, we see that it is enough to show that the matrix

$$\mathbb{A}(x, w, v) = D^2b(x, w, v) + \begin{bmatrix} 2w & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is nonpositive-definite. Here  $b$  is a function given on  $\mathbb{R} \times \Omega_c^r$  by the formula

$$b(x, w, v) = -\frac{rc}{2-r} \frac{x^2}{v^{r-1}}.$$

We compute directly that

$$\mathbb{A}(x, w, v) = \begin{bmatrix} 2w - \frac{2rc}{(2-r)v^{r-1}} & 0 & \frac{2r(r-1)cx}{(2-r)v^r} \\ 0 & 0 & 0 \\ \frac{2r(r-1)cx}{(2-r)v^r} & 0 & -\frac{r^2(r-1)c}{2-r} \frac{x^2}{v^{r+1}} \end{bmatrix}.$$

We see that the entry in the lower-right corner is nonpositive and the determinant is equal to 0. Thus it is enough to show that

$$\det \begin{bmatrix} 2w - \frac{2rc}{(2-r)v^{r-1}} & \frac{2r(r-1)cx}{(2-r)v^{r-2}} \\ \frac{2r(r-1)cx}{(2-r)v^{r-2}} & -\frac{r^2(r-1)c}{2-r} \frac{x^2}{v^{r+1}} \end{bmatrix} \geq 0.$$

But  $w \leq cv^{1-r}$ , so the above determinant is not smaller than

$$\begin{aligned} \det & \begin{bmatrix} \frac{2c}{v^{r-1}} - \frac{2rc}{(2-r)v^{r-1}} & \frac{2r(r-1)cx}{(2-r)v^{r-2}} \\ \frac{2r(r-1)cx}{(2-r)v^{r-2}} & -\frac{r^2(r-1)c}{2-r} \frac{x^2}{v^{r+1}} \end{bmatrix} \\ &= \frac{2c^2x^2}{v^{2r}} \left[ -\left(1 - \frac{r}{2-r}\right) \frac{r^2(r-1)}{2-r} - \frac{2r^2(r-1)^2}{(2-r)^2} \right] = 0. \end{aligned}$$

This completes the proof.  $\square$

**2.3. The special function corresponding to (1.5).** Finally, we turn our attention to the second weighted estimate of Theorem 1.1. The Bellman function is slightly more complicated than that studied in the preceding section, but it has the advantage that it produces a „self-contained” proof of (1.5) (i.e., it does not refer to the self-improving properties of  $A_2$  weights). Define  $B = B_c : \mathbb{R} \times [0, \infty) \times \Omega_c^2 \rightarrow \mathbb{R}$  given by the formula

$$B(x, y, w, v) = yw - \frac{16c^2x^2w}{(wv - 1/2)^\alpha},$$

where  $\alpha = 1 - (4c)^{-1}$ . Let us now establish the appropriate majorizations for this object. First, note that

$$B(x, x^2, w, v) \leq x^2w \left[ 1 - \frac{16c^2}{(c - 1/2)^\alpha} \right] \leq x^2w(1 - 16c) \leq 0.$$

Furthermore, for any  $(x, y, w, v)$  from the domain of  $B$ , we clearly have

$$B(x, y, w, v) \geq yw - 2^\alpha \cdot 16c^2x^2w \geq yw - 32c^2x^2w.$$

Finally, we will show that  $B$  enjoys the property described in Lemma 2.1. We have

$$B(x, y, w, v) = 16c^2 \left[ \frac{yw}{16c^2} + b(x, w, v) \right],$$

where  $b(x, w, v) = -x^2w(wv - 1/2)^{-\alpha}$ . Arguing as in the preceding subsections, we see that it is enough to prove that the matrix

$$\mathbb{A}(x, w, v) = D^2b(x, w, v) + \begin{bmatrix} 2w/(16c^2) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is nonpositive-definite. Substituting  $t = wv$ , we compute that

$$\mathbb{A}(x, w, v) = \begin{bmatrix} \frac{w}{8c^2} - \frac{2w}{(t-1/2)^\alpha} & -\frac{2x((1-\alpha)t-1/2)}{(t-1/2)^{\alpha+1}} & \frac{2\alpha xw^2}{(t-1/2)^{\alpha+1}} \\ -\frac{2x((1-\alpha)t-1/2)}{(t-1/2)^{\alpha+1}} & \frac{\alpha x^2v((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & \frac{\alpha x^2w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} \\ \frac{2\alpha xw^2}{(t-1/2)^{\alpha+1}} & \frac{\alpha x^2w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & -\frac{\alpha(\alpha+1)x^2w^3}{(t-1/2)^{\alpha+2}} \end{bmatrix}.$$

As previously, we apply Sylvester's criterion. Obviously, we have

$$-\frac{\alpha(\alpha+1)x^2w^3}{(t-1/2)^{\alpha+2}} \leq 0.$$

Furthermore,

$$\det \begin{bmatrix} \frac{\alpha x^2 v((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & \frac{\alpha x^2 w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} \\ \frac{\alpha x^2 w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & -\frac{\alpha(\alpha+1)x^2w^3}{(t-1/2)^{\alpha+2}} \end{bmatrix} = \frac{2\alpha^2 x^4 w^2 (1 - (1-\alpha)t)(t-1/2)}{(t-1/2)^{2\alpha+4}} \geq 0,$$

since  $1 - (1-\alpha)t = 1 - t/(4c) > 0$ . Thus, to show that  $\mathbb{A}(x, w, v)$  is nonpositive-definite, it suffices to show that its determinant is nonpositive. To do this, let us conduct some operations on the rows of this matrix. First, multiply the third row by  $v/w$  and add it to the second row. Then

$$\det \mathbb{A}(x, w, v) = \det \begin{bmatrix} \frac{w}{8c^2} - \frac{2w}{(t-1/2)^\alpha} & -\frac{2x((1-\alpha)t-1/2)}{(t-1/2)^{\alpha+1}} & \frac{2\alpha x w^2}{(t-1/2)^{\alpha+1}} \\ -\frac{2x}{(t-1/2)^\alpha} & 0 & \frac{2\alpha x^2 w}{(t-1/2)^{\alpha+1}} \\ \frac{2\alpha x w^2}{(t-1/2)^{\alpha+1}} & \frac{\alpha x^2 w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & -\frac{\alpha(\alpha+1)x^2 w^3}{(t-1/2)^{\alpha+2}} \end{bmatrix}.$$

Next, multiply the second row by  $\alpha w^2/(t-1/2)$  and add it to the third row; furthermore, multiply the second row by  $w/x$  and subtract it from the first row. As the result, we see that

$$\det \mathbb{A}(x, w, v) = \det \begin{bmatrix} \frac{w}{8c^2} & -\frac{2x((1-\alpha)t-1/2)}{(t-1/2)^\alpha} & 0 \\ -\frac{2x}{(t-1/2)^\alpha} & 0 & \frac{2\alpha x^2 w}{(t-1/2)^{\alpha+1}} \\ 0 & \frac{\alpha x^2 w((1-\alpha)t-1)}{(t-1/2)^{\alpha+2}} & \frac{\alpha(\alpha-1)x^2 w^3}{(t-1/2)^{\alpha+2}} \end{bmatrix},$$

so the sign of  $\det \mathbb{A}(x, w, v)$  is the same as the sign of

$$\begin{aligned} \det \begin{bmatrix} \frac{w}{16c^2} & -((1-\alpha)t-1/2) & 0 \\ -1 & 0 & 1 \\ 0 & \alpha((1-\alpha)t-1) & \frac{(\alpha-1)w}{(t-1/2)^\alpha} \end{bmatrix} \\ = \frac{((1-\alpha)t-1/2)(1-\alpha)w}{(t-1/2)^\alpha} + \frac{\alpha(1-(1-\alpha)t)w}{16c^2}. \end{aligned}$$

But the above expression is nonpositive; this can be equivalently rewritten in the form

$$\alpha(1-(1-\alpha)t)(t-1/2)^\alpha < (1-\alpha)(1/2-(1-\alpha)t) \cdot 16c^2$$

and follows from the observations that  $\alpha < 1$ ,  $0 < 1-(1-\alpha)t < 1$ ,  $(t-1/2)^\alpha < c^\alpha \leq c$  and  $(1-\alpha)(1/2-(1-\alpha)t) \cdot 16c^2 = 4c(1/2-t/(4c)) \geq 4c(1/2-1/4) = c$ .

### 3. PROOF OF THEOREM 1.1

We start with the following geometric fact.

**Lemma 3.1.** *Assume that  $c > 1$  and  $r \in (1, 2]$ . Suppose that points  $P$ ,  $Q$  and  $R = (P+Q)/2$  lie in  $\Omega_c^r$ . Then the whole line segment  $PQ$  is contained within  $\Omega_{2c}^r$ .*

*Proof.* Using a simple geometrical argument, it is enough to consider the case when the points  $P$  and  $R$  lie on the curve  $wv^{r-1} = c$  (the upper boundary of  $\Omega_c^r$ ) and  $Q$  lies on the curve  $wv^{r-1} = 1$  (the lower boundary of  $\Omega_c^r$ ). Then the line segment  $RQ$  is contained within  $\Omega_c^r$ , and hence also within  $\Omega_{2c}^r$ , so it is enough to ensure that the segment  $PR$  is



contained in  $\Omega_{2c}^r$ . Let  $P = (P_x, P_y)$ ,  $Q = (Q_x, Q_y)$  and  $R = (R_x, R_y)$ . We consider two cases. If  $P_x < R_x$ , then

$$P_y = 2R_y - Q_y < 2R_y,$$

so the segment  $PR$  is contained in the quadrant  $\{(x, y) : x \leq R_x, y \leq 2R_y\}$ . Consequently,  $PR$  lies below the hyperbola  $xy^{r-1} = 2^{r-1}c$  passing through  $(R_x, 2R_y)$ , and hence also below the hyperbola  $xy^{r-1} = 2c$ . This proves the assertion in the case  $P_x < R_x$ . In the case  $P_x \geq R_x$  the reasoning is similar. Indeed, we check easily that the line segment  $PR$  lies below the hyperbola  $xy^{r-1} = 2c$  passing through  $(2R_x, R_y)$ .  $\square$

We are ready to establish the inequalities of Theorem 1.1.

*Proof of (1.4).* Let us start with introducing some auxiliary objects and notation. Let  $(\mathcal{I}^n)_{n \geq 0}$  denote the dyadic filtration of  $\mathfrak{J}$ : given a nonnegative integer  $n$ ,  $\mathcal{I}^n$  denotes the  $\sigma$ -algebra generated by all dyadic intervals contained within  $\mathfrak{J}$ , which are of measure  $2^{-n}$ . Let  $w$  be an  $A_2$  weight with  $c = [w]_{A_2}$  and let  $\varphi$  be a function belonging to  $L_w^2(\mathfrak{J})$ . For any  $x \in \mathfrak{J}$  and any nonnegative integer  $n$ , define

$$(3.1) \quad \varphi_n(x) = \langle \varphi \rangle_I, \quad \mathfrak{w}_n(x) = \langle w \rangle_I \quad \text{and} \quad \mathfrak{v}_n(x) = \langle w^{-1} \rangle_I,$$

where  $I = I(x)$  is the atom of  $\mathcal{I}^n$  which contains  $x$  (such an atom is unique for almost all  $x \in \mathfrak{J}$ , so the above equalities give functions which are well-defined on the subset of  $\mathfrak{J}$  of full measure). Note that this definition of  $\varphi_n$  is consistent with that given in the introductory section. Furthermore, define the truncated square function of  $\varphi$  by

$$S_n(\varphi) = S(\varphi_n), \quad n = 0, 1, 2, \dots$$

Now, let  $B = B_{2c}$  be the function introduced in §2.1 (note that we take the index  $2c$ ). A crucial fact which exhibits the interplay between  $B$  and the sequences  $\varphi_n$ ,  $S_n(\varphi)$ ,  $\mathfrak{w}_n$  and  $\mathfrak{v}_n$ , is that

$$(3.2) \quad \int_{\mathfrak{J}} B(\varphi_n(s), S_n^2(\varphi)(s), \mathfrak{w}_n(s), \mathfrak{v}_n(s)) ds \\ \geq \int_{\mathfrak{J}} B(\varphi_{n+1}(s), S_{n+1}^2(\varphi)(s), \mathfrak{w}_{n+1}(s), \mathfrak{v}_{n+1}(s)) ds$$

for  $n \geq 0$ . To prove this estimate, fix an  $n$  and pick an atom  $I$  of  $\mathcal{I}^n$ . We will prove a slightly stronger statement than (3.2), in which  $\mathfrak{J}$  is replaced by  $I$ , i.e.,

$$(3.3) \quad \int_I B(\varphi_n(s), S_n^2(\varphi)(s), \mathfrak{w}_n(s), \mathfrak{v}_n(s)) ds \\ \geq \int_I B(\varphi_{n+1}(s), S_{n+1}^2(\varphi)(s), \mathfrak{w}_{n+1}(s), \mathfrak{v}_{n+1}(s)) ds.$$

Denote the left and right half of  $I$  by  $I^-$  and  $I^+$ , respectively. Then the functions  $\varphi_n$ ,  $S_n^2(\varphi)$ ,  $\mathfrak{w}_n$  and  $\mathfrak{v}_n$  are constant on  $I$ ; let us denote the corresponding values by  $x$ ,  $y$ ,  $w$  and  $v$ . Similarly,  $\varphi_{n+1}$ ,  $S_{n+1}^2(\varphi)$ ,  $\mathfrak{w}_{n+1}$  and  $\mathfrak{v}_{n+1}$  are constant on each of  $I^-$ ,  $I^+$ : denote the appropriate values by  $x_{\pm}$ ,  $y_{\pm}$ ,  $w_{\pm}$  and  $v_{\pm}$ . From the very definition of the sequences  $(\varphi_n)_{n \geq 0}$ ,  $(\mathfrak{w}_n)_{n \geq 0}$  and  $(\mathfrak{v}_n)_{n \geq 0}$ , we infer that

$$x = (x_- + x_+)/2, \quad w = (w_- + w_+)/2 \quad \text{and} \quad v = (v_- + v_+)/2,$$

so there are  $d, e, f \in \mathbb{R}$  such that

$$x_{\pm} = x \pm d, \quad w_{\pm} = w \pm e \quad \text{and} \quad v_{\pm} = v \pm f.$$

In addition, by the very definition of  $(S_n(\varphi))_{n \geq 0}$ , we see that  $y_- = y_+ = y + d^2$ . If we plug all these facts into (3.3) and divide both sides by  $|I|/2$ , we get an estimate which

is equivalent to (2.3). Thus, (3.3) will be established if we show that the assumption of Lemma 2.1 is satisfied. However, the points

$$\begin{aligned}(w_-, v_-) &= \left( \frac{1}{|I^-|} \int_{I^-} w, \frac{1}{|I^-|} \int_{I^-} w^{-1} \right), \\ (w_+, v_+) &= \left( \frac{1}{|I^+|} \int_{I^+} w, \frac{1}{|I^+|} \int_{I^+} w^{-1} \right), \\ (w, v) &= \frac{(w_-, v_-) + (w_+, v_+)}{2} = \left( \frac{1}{|I|} \int_I w, \frac{1}{|I|} \int_I w^{-1} \right)\end{aligned}$$

belong to  $\Omega_c^2$ , by the very definition of  $A_2$  weights. Consequently, by Lemma 3.1, the line segment with endpoints  $(w_-, v_-)$ ,  $(w_+, v_+)$  is entirely contained in  $\Omega_{2c}^2$ , which is precisely the requirement of Lemma 2.1 (recall that the special function  $B$  we use corresponds to the parameter  $2c$ ). This yields (3.3), and summing over all atoms  $I$  of  $\mathcal{I}^n$ , we obtain (3.2). The remainder of the proof is straightforward. By induction, (3.2) gives

$$(3.4) \quad \int_{\mathcal{J}} B(\varphi_n(s), S_n^2(\varphi)(s), \mathfrak{w}_n(s), \mathfrak{v}_n(s)) ds \leq \int_{\mathcal{J}} B(\varphi_0(s), S_0^2(\varphi)(s), \mathfrak{w}_0(s), \mathfrak{v}_0(s)) ds.$$

However, we have

$$\varphi_0(s) = \frac{1}{|\mathcal{J}|} \int_{\mathcal{J}} \varphi, \quad \text{and} \quad S_0(\varphi) = \left| \frac{1}{|\mathcal{J}|} \int_{\mathcal{J}} \varphi \right|,$$

so (2.1) implies that the right-hand side of (3.4) is nonpositive. To deal with the left-hand side, we exploit (2.2). As the result, we obtain the estimate

$$\int_{\mathcal{J}} (\varphi_n(s))^2 \mathfrak{w}_n(s) ds \leq 160c \int_{\mathcal{J}} S_n^2(\varphi)(s) \mathfrak{w}_n(s) ds,$$

which, by the very definition of  $\mathfrak{w}_n$ , implies

$$\int_{\mathcal{J}} (\varphi_n(s))^2 w(s) ds \leq 160c \int_{\mathcal{J}} S_n^2(\varphi)(s) w(s) ds.$$

However, if we let  $n$  go to infinity, then  $S_n(\varphi) \uparrow S(\varphi)$  and  $\varphi_n \rightarrow \varphi$  almost surely, by Lebesgue's differentiation theorem. Consequently, Fatou's lemma and Lebesgue's monotone convergence theorem imply

$$\|\varphi\|_{L_w^2(\mathcal{J})} \leq (160c)^{1/2} \|S(\varphi)\|_{L_w^2(\mathcal{J})},$$

which is the desired claim.  $\square$

*Proof of (1.5).* The arguments go along the same lines as above. We omit the straightforward repetitions.  $\square$

*Proof of (1.6).* Here the reasoning is essentially the same as above, but we have decided to include some details due to the appearance of the  $A_r$  weights. Suppose that  $w$  is an  $A_2$  weight. Then, by Lemma 1.2,  $w$  is an  $A_r$  weight for some  $r < 2$ . Let  $c = [w]_{A_r}$  and assume that  $B = B_{2c,r}$  is the function introduced in §2.2 (again, note that we use the doubled index  $2c$ ). Let  $\varphi$  be a function belonging to  $L_w^2(\mathcal{J})$  and define  $(\varphi_n)_{n \geq 0}$ ,  $(S_n(\varphi))_{n \geq 0}$ ,  $(\mathfrak{w}_n)_{n \geq 0}$  as previously. The corresponding sequence  $(\mathfrak{v}_n)_{n \geq 0}$  is slightly different, as it captures the fact that  $w$  is an  $A_r$  weight: let

$$\mathfrak{v}_n(x) = \langle w^{-1/(r-1)} \rangle_{I(x)}$$

(recall that  $I(x)$  is the element of  $\mathcal{I}^n$  which contains the point  $x$ ). By Muckenhoupt's condition  $A_r$ , we see that the sequence  $((w_n, v_n))_{n \geq 0}$  is  $\Omega_c^r$ -valued. Therefore, repeating the arguments from the preceding proof, we get that

$$\int_{\mathcal{J}} B(\varphi_n(s), S_n^2(\varphi)(s), w_n(s), v_n(s)) ds \leq \int_{\mathcal{J}} B(\varphi_0(s), S_0^2(\varphi)(s), w_0(s), v_0(s)) ds$$

for any  $n \geq 0$ . Consequently, by (2.9) and (2.10), we get

$$\int_{\mathcal{J}} S_n^2(\varphi) w_n(s) ds \leq \frac{2rc}{2-r} \int_{\mathcal{J}} (\varphi_n(s))^2 w_n(s) ds$$

and hence

$$\int_{\mathcal{J}} S_n^2(\varphi) w(s) ds \leq \frac{2rc}{2-r} \int_{\mathcal{J}} (\varphi_n(s))^2 w(s) ds.$$

If we let  $n \rightarrow \infty$ , the left-hand side converges to  $\|S(\varphi)\|_{L_w^2(\mathcal{J})}^2$ , by Lebesgue's monotone convergence theorem. To deal with the right-hand side, recall that  $\varphi$  belongs to  $L_w^2(\mathcal{J})$  and hence, by Muckenhoupt's inequality, so does the dyadic maximal function  $\mathcal{M}_d \varphi$ . Therefore, by Lebesgue's differentiation and dominated convergence theorems, we see that

$$\int_{\mathcal{J}} (\varphi_n(s))^2 w(s) ds \xrightarrow{n \rightarrow \infty} \|\varphi\|_{L_w^2(\mathcal{J})}^2.$$

This gives the claim.  $\square$

#### 4. INEQUALITIES FOR CONTINUOUS TIME MARTINGALES

All the results studied above have their counterparts in the martingale theory, as we will prove now. Suppose that  $(\Omega, \mathcal{F}, \mathbb{P})$  is a complete probability space, equipped with a filtration  $(\mathcal{F}_t)_{t \geq 0}$ , i.e., a nondecreasing sequence of sub- $\sigma$ -algebras of  $\mathcal{F}$ . Assume further that  $\mathcal{F}_0$  contains all the events of probability 0. Let  $X = (X_t)_{t \geq 0}$  be an adapted, uniformly integrable continuous-path martingale and let  $\langle X \rangle = (\langle X_t \rangle)_{t \geq 0}$  stand for its quadratic covariance process (square bracket). See e.g. Dellacherie and Meyer [16] for the detailed exposition and properties of this object. Then  $X$  and  $\langle X \rangle^{1/2}$  are the probabilistic versions of the function  $\varphi$  and its square function  $S(\varphi)$ . To introduce the appropriate analogue of  $A_p$  weights, assume that  $Y$  is a nonnegative, uniformly integrable martingale with continuous trajectories, satisfying  $Y_0 = \mathbb{E}[Y_\infty] = 1$ . (Note that this normalization is not an essential assumption as multiplying by a constant does not affect the  $A_p$  condition.) Following Izumisawa and Kazamaki [22], we say that  $Y$  satisfies Muckenhoupt's condition  $A_p(\text{mart})$  (where  $1 < p < \infty$  is a fixed parameter), if

$$(4.1) \quad \|Y\|_{A_p(\text{mart})} := \sup_{t > 0} \left\| Y_t \left( \mathbb{E} \left[ \left( \frac{1}{Y_\infty} \right)^{1/(p-1)} \middle| \mathcal{F}_t \right] \right)^{p-1} \right\|_\infty < \infty,$$

where  $Y_t = \mathbb{E}[Y_\infty | \mathcal{F}_t]$ . Any process  $Y$  as above gives rise to the probability measure  $\mathbb{Q}$  defined by the equation  $d\mathbb{Q} = Y_\infty d\mathbb{P}$ , and thus it can be regarded as a weight. Put  $Z_t = \mathbb{E}\{Y_\infty^{-1/(p-1)} | \mathcal{F}_t\}$  for  $t \geq 0$ . If  $\|Y\|_{A_p} \leq c$ , then we have  $Y_t Z_t^{p-1} \leq c$  for all  $t$ , that is, the process  $(Y, Z)$  takes values in the set  $\Omega_c^p$ . This process is precisely the martingale analogue of the sequence  $((w_n, v_n))_{n \geq 0}$  studied in Section 3 above.

The analogue of Theorem 1.1 is the following.

**Theorem 4.1.** *Suppose that  $Y$  is an  $A_2$  weight and  $X$  is a martingale bounded in  $L^2(\mathbb{Q})$ . Then we have the estimates*

$$(4.2) \quad \|X_\infty\|_{L^2(\mathbb{Q})} \leq (80\|Y\|_{A_2(\text{mart})})^{1/2} \|\langle X \rangle_\infty^{1/2}\|_{L^2(\mathbb{Q})}$$

and

$$(4.3) \quad \|\langle X \rangle_\infty^{1/2}\|_{L^2(\mathbb{Q})} \leq 4\sqrt{2}\|Y\|_{A_2(\text{mart})}\|X_\infty\|_{L^2(\mathbb{Q})}.$$

Furthermore,

$$(4.4) \quad \|\langle X \rangle_\infty^{1/2}\|_{L^2(\mathbb{Q})} \leq \inf_{1 < r < 2} \left( \frac{r}{2-r} \|Y\|_{A_r(\text{mart})} \right)^{1/2} \|X_\infty\|_{L^2(\mathbb{Q})}.$$

Note that the constants are slightly better: this is due to the fact that we will not require Lemma 3.1, or any probabilistic counterpart of that statement (see the beginning of the proof below).

*Proof.* We will focus on (4.2), the reasoning leading to (4.4) is essentially the same. Let  $c = \|Y\|_{A_2(\text{mart})}$  and let  $B = B_c$  be the function of Subsection §2.1. (Note that in contrast with the analytic setting, here we use the function  $B_c$ , not  $B_{2c}$ . This will give the aforementioned improvement of the constants). The function  $B$  is of class  $C^\infty$ ; actually, it can be extended to a  $C^\infty$  function on a certain open set containing  $\mathbb{R} \times [0, \infty) \times \Omega_c^2$ . Furthermore, by the probabilistic  $A_2$  condition, we see that the process  $S = (X, \langle X \rangle, Y, Z)$  takes values in the domain of  $B$ . Thus, an application of Itô's formula gives that for each  $t \geq 0$ ,

$$(4.5) \quad B(S_t) = B(S_0) + I_1 + I_2 + I_3/2,$$

where

$$\begin{aligned} I_1 &= \int_{0+}^t B_x(S_s) dX_s + \int_{0+}^t B_w(S_s) dY_s + \int_{0+}^t B_v(S_s) dZ_s, \\ I_2 &= \int_{0+}^t B_y(S_s) d\langle X \rangle_s, \\ I_3 &= \int_{0+}^t D_{x,w,v}^2 B(S_s) \cdot d\langle (X, Y, Z), (X, Y, Z) \rangle_s. \end{aligned}$$

Here

$$D_{x,w,v}^2 B(x, y, w, v) = \begin{bmatrix} B_{xx} & B_{xw} & B_{xv} \\ B_{xw} & B_{ww} & B_{wv} \\ B_{xv} & B_{wv} & B_{vv} \end{bmatrix}$$

and the integral  $I_3$  is actually a sum of the integrals

$$\int_{0+}^t B_{xx}(S_s) d\langle X \rangle_s, \quad 2 \int_{0+}^t B_{xw}(S_s) d\langle X, Y \rangle_s, \quad 2 \int_{0+}^t B_{xv}(S_s) d\langle X, Z \rangle_s, \dots$$

and so on. By the properties of stochastic integrals, we see that the term  $I_1$  has expectation 0. Furthermore, the sum  $I_2 + I_3$  is nonpositive: this follows directly from some standard approximation and the fact that the matrix  $\mathbb{A}$ , introduced in (2.5), is nonpositive-definite. Consequently, integrating both sides of (4.5) gives  $\mathbb{E}B(S_t) \leq \mathbb{E}B(S_0)$ . However,  $\langle X \rangle_0 = |X_0|^2$ , so (2.1) implies  $B(S_0) \leq 0$ ; furthermore, by (2.2), we have

$$B(S_t) \geq \frac{Y_t}{2}(X_t^2 - 80c\langle X \rangle_t).$$

Combining these facts, we get the estimate

$$\mathbb{E}X_t^2 Y_t \leq 80c\mathbb{E}\langle X \rangle_t Y_t,$$

which, by the martingale property of  $Y$ , implies

$$\mathbb{E}X_t^2 Y_\infty \leq 80c \mathbb{E}\langle X \rangle_t Y_\infty \leq 80c \mathbb{E}\langle X \rangle_\infty Y_\infty.$$

This clearly implies (4.2), in view of Fatou's lemma. As we have mentioned above, the proofs of (4.3) and (4.4) are similar, so we leave them to the interested reader.  $\square$

The estimate (4.4) leads to the following improvement of (4.3).

**Corollary 4.1.** *Suppose that  $Y$  is an  $A_2$  weight and  $X$  is a martingale bounded in  $L^2(\mathbb{Q})$ . Then we have the estimate*

$$(4.6) \quad \|\langle X \rangle^{1/2}\|_{L^2(\mathbb{Q})} \leq 2^{7/4} \|Y\|_{A_2(\text{mart})} \|X_\infty\|_{L^2(\mathbb{Q})}.$$

*Proof.* We will need an appropriate probabilistic version of Coifman-Fefferman  $A_{p-\varepsilon}$ -result (see Lemma 1.2 above). As shown by Uchiyama [45], we have the identity

$$(4.7) \quad \|Y\|_{A_r(\text{mart})} = \sup \left[ \lambda^r \mathbb{Q}(\xi^* > \lambda) / \|\xi_\infty\|_{L^r(\mathbb{Q})}^r \right],$$

where the supremum is taken over all  $\lambda > 0$  and all adapted cadlag martingales  $\xi = (\xi_t)_{t \geq 0}$ . On the other hand, it follows from the results of Osękowski [35] that if  $Y$  is an  $A_2$  weight satisfying  $\|Y\|_{A_2} = c$ , then the expression on the right of (4.7) does not exceed

$$\left[ \left( 1 - \frac{\sqrt{1-c^{-1}}}{r-1} \right)^{r-1} \left( 1 + \sqrt{1-c^{-1}} \right) \right]^{-1/r},$$

provided  $1 + \sqrt{1-c^{-1}} < r \leq 2$ . As one easily verifies, the latter double bound is satisfied by  $r = 2 - (2c)^{-1}$ , and then (4.7) implies

$$\|Y\|_{A_r} \leq \left[ \left( 1 - \frac{\sqrt{1-c^{-1}}}{1-(2c)^{-1}} \right)^{1-(2c)^{-1}} \left( 1 + \sqrt{1-c^{-1}} \right) \right]^{-1/(2-(2c)^{-1})}.$$

However,  $1 + \sqrt{1-c^{-1}} \geq 1$  and

$$1 - \frac{\sqrt{1-c^{-1}}}{1-(2c)^{-1}} = \frac{(4c^2)^{-1}}{(1-(2c)^{-1})(1-(2c)^{-1} + \sqrt{1-c^{-1}})} \geq \frac{1}{8c^2}.$$

Consequently, we obtain the upper bound

$$\|Y\|_{A_r} \leq (8c^2)^{(1-(2c)^{-1})/(2-(2c)^{-1})} \leq (8c^2)^{1/2} = 2\sqrt{2}c.$$

Furthermore, we have

$$\frac{r}{2-r} = \frac{2-(2c)^{-1}}{(2c)^{-1}} \leq 4c.$$

Plugging the above two estimates into (4.4) gives the claim.  $\square$

## 5. LITTLEWOOD-PALEY SQUARE FUNCTIONS AND THE LUSIN AREA INTEGRAL

Our goal in this section is to prove versions of Theorems 1.1 and 4.1 for the Littlewood-Paley square functions  $g_*$  and  $G_*$  for harmonic and parabolic functions, respectively. These operators arise as conditional expectations of square functions of martingales obtained by composing harmonic functions in the upper half-space with Brownian motion and from martingales obtained by composing solutions of the heat equation with space-time Brownian motion. These constructions are quite general and apply in the wide setting of general symmetric Markovian semigroups and their Poisson semigroups obtained by Bochner's 1/2-subordination.

Before proceeding further, we mention here that there is a vast literature on weighted  $L^p$  inequalities for the classical Littlewood-Paley square functions and their many variants. These include estimates with the sharp dependence on the characteristics  $[w]_{A_p}$  of the weight. For some of this literature we refer the reader to A. Lerner [24, 25, 26] and the many references given in those papers. Our inequalities below provide information not only on the  $[w]_{A_2}$  dependence but also give  $L^2$  bounds with very explicit constants. This raises questions of obtaining sharp bounds on weighted norm inequalities for classical operators not only with respect to  $[w]_{A_p}$  but also with respect to  $p$  when the weights are the probabilistic Poisson or heat (and even more general symmetric Markovian semigroup) as defined below. Of particular interest would be the case of the Hilbert transform, first and second order Riesz transforms, and the Beurling-Ahlfors operator. A first step in these problems would be to obtain such sharp weighted norm inequalities for martingale transforms of stochastic integrals with the probabilistic weights as defined by (4.1). Unfortunately, as of now we have not been able to make progress on these problems.

**5.1. Littlewood-Paley and Lusin square functions for harmonic functions.** In this section we will derive versions of Theorems 1.1 and 4.1 for the Littlewood-Paley  $g_*$  function for harmonic and parabolic functions. This function dominates, pointwise, the classical Littlewood-Paley  $g$  function and the Lusin area integral. We will use the standard construction of  $g_*$  as the conditional expectation of the martingale square function. For this construction we refer the reader to [5] which is a special case of the more general Poisson semigroup construction presented in Varopoulos [46]. (See also Meyer [30, 31, 32].)

For any  $f \in L^p(\mathbb{R}^n)$ ,  $1 < p < \infty$ , we will denote by  $K_y(f)(x)$  its harmonic extension to the upper half-space  $\mathbb{R}_+^{n+1} = \{(x, y) : x \in \mathbb{R}^n, y > 0\}$  obtained by convolving  $f$  with the Poisson kernel

$$(5.1) \quad k_y(x) = \frac{c_n y}{(y^2 + |x|^2)^{(n+1)/2}}, \quad c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}},$$

where  $c_n$  is chosen so that  $k_y(x)$  has integral 1 for all  $y > 0$ . The cone in  $\mathbb{R}_+^{n+1}$  with vertex at  $x$  and aperture  $\alpha > 0$  is defined by

$$\Gamma_\alpha(x) = \{(z, y) : |z - x| < \alpha y, z \in \mathbb{R}^n, y > 0\}.$$

The Littlewood-Paley functions  $g(f)$ ,  $g_*(f)$ , and Lusin area integral  $A_\alpha(f)$  are defined, respectively, by

$$(5.2) \quad g(f)(x) = \left( \int_0^\infty y |\nabla K_y(f)(x)|^2 dy \right)^{1/2},$$

$$(5.3) \quad g_*(f)(x) = \left( \int_0^\infty \int_{\mathbb{R}^d} y k_y(x-z) |\nabla K_y(f)(z)|^2 dz dy \right)^{1/2}$$

and

$$(5.4) \quad A_\alpha(f)(x) = \left( \int_{\Gamma_\alpha(x)} y^{1-n} |\nabla K_y(f)(x)|^2 dz dy \right)^{1/2},$$

where for any  $u$  in the upper half-space,

$$\nabla u = \left( \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial u}{\partial y} \right)$$

is the full gradient.

As in the case of the dyadic square function, the inequalities (1.1) and (1.2) hold for all  $f \in L^p(\mathbb{R}^n)$ , for  $1 < p < \infty$ , for both  $g$  and  $A_\alpha$ . For  $g_*$  we have  $\|g_*(f)\|_p \leq c_p \|f\|_p$  for  $2 \leq p < \infty$  and it is well-known that the inequality fails for  $1 < p < 2$ ; see [8] for an explicit example. We refer the reader to Stein [44] for these classical results and where it is also shown that there are constants  $c_{\alpha,n}, C_{\alpha,n}$ , depending only on  $\alpha$  and  $n$ , such that

$$(5.5) \quad g(f)(x) \leq c_{\alpha,n} A_\alpha(f)(x) \leq C_{\alpha,n} g_*(f)(x), \text{ for all } x \in \mathbb{R}^n.$$

In fact, the second inequality is trivial since

$$y^{-n} \leq \frac{(\alpha^2 + 1)^{\frac{n+1}{2}}}{c_n} k_y(z - x), \quad (z, y) \in \Gamma_\alpha(x),$$

where  $c_n$  is the constant in (5.1). This gives

$$(5.6) \quad A_\alpha(f)(x) \leq \frac{(\alpha^2 + 1)^{\frac{n+1}{4}}}{\sqrt{c_n}} g_*(f)(x).$$

Also, the semigroup property of  $K_y$  gives directly (see also the proof of (5.16) below) that

$$(5.7) \quad g(f)(x) \leq 2g_*(f)(x).$$

An important property for our purpose in this paper is the fact that  $g_*(f)$  can be expressed as the conditional expectation of the quadratic variation (square function) of the martingale obtained by composing the harmonic function  $K_y f(x)$  with the Brownian motion in  $\mathbb{R}_+^{n+1}$ . Let us explain this further. Let  $B_t = (X_t, Y_t)$ ,  $t \geq 0$ , be Brownian motion in  $\mathbb{R}_+^{n+1}$  starting at the point  $(z, a)$  so that  $(X_t)_{t \geq 0}$  is an  $n$ -dimensional Brownian motion and  $(Y_t)_{t \geq 0}$  is a one-dimensional Brownian motion. We denote the corresponding starting probability measure and expectation by  $\mathbb{P}_{(z,a)}$  and  $\mathbb{E}_{(z,a)}$ , respectively. Let  $\tau$  be its exit time from  $\mathbb{R}_+^{n+1}$  so that  $\tau = \inf\{s > 0 : Y_s = 0\}$ . Since the density of the distribution of the random variable  $X_\tau$  under the probability measure  $\mathbb{P}_{(z,a)}$  is given by the Poisson kernel  $k_a(z - x)$ , Fubini's theorem gives that for nonnegative (or integrable) functions  $F$  on  $\mathbb{R}^n$  we have

$$(5.8) \quad \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)} F(X_\tau) dz = \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} k_a(z - x) F(x) dx \right) dz = \int_{\mathbb{R}^n} F(x) dx,$$

for all  $a > 0$ . This simple formula is used below multiple times to convert integrals over  $\mathbb{R}^n$  with respect to the Lebesgue measure to expectations which then permits the application of martingale inequalities.

We now consider the martingale  $M(f)_t = K_{Y_{\tau \wedge t}}(f)(X_{\tau \wedge t})$ ,  $t \geq 0$ . By the Itô formula,

$$M(f)_t = K_a f(z) + \int_0^{\tau \wedge t} \nabla K_{Y_s}(f)(X_s) \cdot dB_s$$

and the quadratic variation of the martingale is given by

$$\langle M(f) \rangle_t = |K_a f(z)|^2 + \int_0^{\tau \wedge t} |\nabla K_{Y_s}(f)(X_s)|^2 ds.$$

Setting

$$\mathbb{E}_{(z,a)}^x \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds \right) = \mathbb{E}_{(z,a)} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds \mid B_\tau = (x, 0) \right),$$

it is proved in Bañuelos [5, p. 663]) that

$$(5.9) \quad \begin{aligned} g_{*,a}^2(f)(x) &= \int_0^a \int_{\mathbb{R}^n} y k_y(x-z) |\nabla K_y(f)(z)|^2 dz dy \\ &= \frac{1}{2} \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)}^x \left( \int_0^{\tau} |\nabla K_{Y_s}(f)(X_s)|^2 ds \right) k_a(x-z) dz. \end{aligned}$$

**5.2. Littlewood-Paley and Lusin square functions for parabolic functions.** As it is well known, the Poisson kernel in the above construction can be replaced by more general volume preserving dilations of approximations to the identity and these give rise to more general Littlewood-Paley and Lusin square functions that have similar  $L^p$  boundedness properties. See for example, [7]. A probabilistic way (also well known by now) to generalize the Poisson kernel construction above is to replace it with a more general Poisson semigroup obtained by the Bochner 1/2-subordination of a symmetric Markovian semigroup. This is what is done in Varopoulos [46] and Meyer [30]. A variation of this construction applies to space-time martingales arising from the Markovian semigroup itself and not just by its 1/2-subordination. This construction was used, in for example, [6], for applications to the Beurling-Ahlfors operator and second order Riesz transforms. For our purpose here, and to connect to the classical  $A_p$ -weights and the classical parabolic square functions in analysis, we present the construction for the heat (Gaussian kernel) semigroup on  $\mathbb{R}^n$ . For  $t > 0$ ,  $x \in \mathbb{R}^n$ , denote the heat (Gaussian) kernel by

$$(5.10) \quad p_t(x) = \frac{1}{(2\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{2t}}$$

and this time consider the heat extension  $P_t f(x) = (p_t * f)(x)$  of  $f$ . The parabolic ‘‘cone’’ in  $\mathbb{R}_+^{n+1}$  with vertex at  $x$  and aperture  $\alpha$  is defined by

$$\mathcal{P}_\alpha(x) = \{(z, t) \in \mathbb{R}_+^{n+1} : |z - x| < \alpha\sqrt{t}\}.$$

The corresponding parabolic Littlewood-Paley functions  $G(f)$ ,  $G_*(f)$  and parabolic Lusin area function  $\mathcal{P}A_\alpha f(x)$  are given, receptively, by

$$(5.11) \quad G(f)(x) = \left( \int_0^\infty |\nabla_x P_t f(x)|^2 dt \right)^{1/2},$$

$$(5.12) \quad G_*(f)(x) = \left( \int_0^\infty \int_{\mathbb{R}^d} p_t(x-z) |\nabla_x P_t f(z)|^2 dz dt \right)^{1/2},$$

and

$$(5.13) \quad \mathcal{P}A_\alpha(f)(x) = \left( \int_{\mathcal{P}_\alpha(x)} t^{-\frac{n}{2}} |\nabla_x P_t f(z)|^2 dz dt \right)^{1/2},$$

where for any function  $U(x, t)$  in the upper half-space

$$\nabla_x U = \left( \frac{\partial U}{\partial x_1}, \frac{\partial U}{\partial x_2}, \dots, \frac{\partial U}{\partial x_n} \right)$$

denotes its ‘‘horizontal’’ gradient.

These square functions have also been widely studied in the literature. We refer the reader to [7] (and references given there) for some of their basic properties. As in the case of harmonic functions we have  $\|G(f)(x)\|_p \approx \|\mathcal{P}A_\alpha f(x)\|_p \approx \|f\|_p$ , for  $1 < p < \infty$ ,



and  $\|G_*(f)\|_p \leq c_p \|f\|_p$ ,  $2 \leq p < \infty$ . Similarly, the the following pointwise inequality holds:

$$(5.14) \quad G(f)(x) \leq c_{\alpha,n} \mathcal{P}A_\alpha(f)(x) \leq C_{\alpha,n} G_*(f)(x), \text{ for all } x \in \mathbb{R}^n,$$

for some constants  $c_{\alpha,n}$  and  $C_{\alpha,n}$  depending only on  $\alpha$  and  $n$ . In fact, since

$$t^{-n/2} \leq (2\pi)^{n/2} e^{\alpha^2/2} p_t(x-z), \text{ for } z \in \mathcal{P}_\alpha(x),$$

we have

$$(5.15) \quad \mathcal{P}A_\alpha(f)(x) \leq (2\pi)^{n/4} e^{\alpha^2/4} G_*(f)(x)$$

Similarly, by the semigroup property,

$$P_t f(x) = P_{t/2}(P_{t/2}f)(x) = \int_{\mathbb{R}^n} p_{t/2}(x-z) P_{t/2}f(z) dz$$

and by Jensen's inequality,

$$|\nabla_x P_t f(x)|^2 \leq \int_{\mathbb{R}^n} p_{t/2}(x-z) |\nabla_x P_{t/2}f(z)|^2 dz.$$

This gives

$$(5.16) \quad \begin{aligned} G^2(f)(x) &= \int_0^\infty |\nabla_x P_t f(z)|^2 dt \\ &\leq \int_0^\infty \int_{\mathbb{R}^n} p_{t/2}(z-x) |\nabla_x P_{t/2}f(z)|^2 dz dt \\ &= 2G_*^2(f)(x). \end{aligned}$$

As in the case of the Poisson kernel, the Littlewood-Paley  $G_*$  function is the conditional expectation of the corresponding martingale square function. This time, however, the martingales arise from space-time Brownian motion. This fact is proved exactly as in [46] or [5, p. 663], once the space-time martingale is identified. For the sake of completeness, we briefly explain this. For the space-time martingale details as used here, see for example [6].

Let  $(B_t)_{t \geq 0}$  be the Brownian motion in  $\mathbb{R}^n$  starting at  $z$  and let  $\mathbb{P}_z$  and  $\mathbb{E}_z$  be the probability and expectation for  $B$ . Fix  $0 < T < \infty$ . Then

$$M(f)_t = P_{T-t}f(B_t), \quad 0 \leq t \leq T,$$

is a martingale and by the Itô's formula,

$$M(f)_t = P_T f(z) + \int_0^t \nabla_x P_{T-s}f(B_s) \cdot dB_s.$$

The quadratic variation (square function) of this martingale is given by

$$\langle M(f) \rangle_t = |P_T f(z)|^2 + \int_0^t |\nabla_x P_{T-s}f(B_s)|^2 ds.$$

Setting

$$(5.17) \quad G_{*,T}(f)(x) = \left( \int_0^T \int_{\mathbb{R}^d} p_t(x-z) |\nabla_x P_t f(z)|^2 dz dt \right)^{1/2},$$

we claim that

$$(5.18) \quad G_{*,T}^2 f(x) = \int_{\mathbb{R}^n} \mathbb{E}_z^x \left( \int_0^T |\nabla_x P_{T-s}f(B_s)|^2 ds \right) p_T(x-z) dz,$$

where

$$\mathbb{E}_z^x \left( \int_0^T |\nabla_x P_{T-s} f(B_s)|^2 ds \right) = \mathbb{E}_z \left( \int_0^T |\nabla_x P_{T-s} f(B_s)|^2 ds \mid B_T = x \right).$$

To prove (5.18) recall that the conditional distribution of  $B_s$  given  $B_T = x$  (the Brownian bridge with  $B_0 = z, B_T = x$ ) is

$$\frac{p_s(z-w)p_{T-s}(w-x)}{p_T(z-x)} dw.$$

Thus,

$$\begin{aligned} & \mathbb{E}_z^x \left( \int_0^T |\nabla_z P_{T-s} f(B_s)|^2 ds \right) \\ &= \int_0^T \int_{\mathbb{R}^n} \frac{p_s(z-w)p_{T-s}(w-x)}{p_T(z-x)} |\nabla_x P_{T-s} f(w)|^2 dw ds. \end{aligned}$$

Multiplying both sides of this equality by  $p_T(z-x)$  and integrating on  $z$  we obtain

$$\begin{aligned} & \int_{\mathbb{R}^d} \mathbb{E}_z^x \left( \int_0^T |\nabla_x P_{T-s} f(B_s)|^2 ds \right) p_T(z-x) dz \\ &= \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^n} p_s(z-w)p_{T-s}(w-x) |\nabla_x U_f(w, T-s)|^2 dw dz ds \\ &= \int_0^T \int_{\mathbb{R}^d} p_{T-s}(y-z) |\nabla P_{T-s} f(y)|^2 dy, ds \\ &= \int_0^T \int_{\mathbb{R}^d} p_s(w-x) |\nabla_x P_{T-s} f(w)|^2 dw ds = G_{*,T}^2 f(x), \end{aligned}$$

which verifies (5.18).

**5.3. Poisson  $A_p$  weights in the Disc.** We now explore the connections between the martingale  $A_p$  weights studied in §4 and various classes previously studied in analysis. This connections are more transparent for the unit disk in the plane (or unit ball in  $\mathbb{R}^n$ ) where the Brownian motion has a natural place to start, namely the origin. For further clarity and to connect to the classical Littlewood-Paley square functions, we first treat this case. Let  $\mathbb{D} = \{x \in \mathbb{C} : |z| < 1\}$  be the unit disc in the complex plain with the circle  $\mathbb{T} = \partial D$  as its boundary and Poisson kernel given by

$$P_z(e^{i\theta}) = \frac{1-|z|^2}{|z-e^{i\theta}|^2}, \quad z \in \mathbb{D}.$$

For the rest of this section we assume that  $w$  is a positive and integrable function on the unit circle  $\mathbb{T}$ . Let

$$u_w(z) = \frac{1}{2\pi} \int_{\mathbb{T}} P_z(e^{i\theta}) w(e^{i\theta}) d\theta$$

be the Poisson integral of  $w$ .

**Definition 5.1.** We say that  $w \in A_p(\text{Poisson}, \mathbb{T})$  if

$$(5.19) \quad \|w\|_{A_p, \mathbb{T}} = \|u_w(z) (u_{w^{-1/p-1}}(z))^{p-1}\|_{L^\infty(\mathbb{D})} < \infty.$$

Now, let  $(B_t)_{t \geq 0}$  be Brownian motion in  $\mathbb{D}$  starting at the origin and let  $\tau$  be its first exit time from  $\mathbb{D}$ . Since  $u_w(z)$  is harmonic, the process  $Y_t = u_w(B_{\tau \wedge t})$ ,  $t \geq 0$ , is a martingale with  $Y_\infty = w(B_\tau)$ . By the strong Markov property,

$$\mathbb{E}_0(Y_\infty | \mathcal{F}_{\tau \wedge t}) = \mathbb{E}_{B_{\tau \wedge t}}(w(B_\tau)) = Y_t$$

and similarly for  $Y_\infty^{-1/(p-1)}$ :

$$\mathbb{E}_0\left(Y_\infty^{-1/(p-1)} | \mathcal{F}_{\tau \wedge t}\right) = \mathbb{E}_0\left(u_{w^{-1/(p-1)}}(B_\tau) | \mathcal{F}_{\tau \wedge t}\right) = \mathbb{E}_{B_{\tau \wedge t}}\left(u_{w^{-1/(p-1)}}(B_\tau)\right).$$

Recalling the martingale  $A_p$  weights defined in (4.1), we see that  $Y \in A_p(\text{mart})$  if and only if  $w \in A_p(\text{Poisson}, \mathbb{T})$  and in fact, we have

$$(5.20) \quad \|Y\|_{A_p(\text{mart})} = \|w\|_{A_{p,\mathbb{T}}}.$$

**Remark 5.1.** The  $A_p(\text{Poisson}, \mathbb{T})$  weights have been studied in recent years in connection with the  $L_w^2(\mathbb{T})$  boundedness of the conjugate function (Hilbert transform) with the correct dependence on the constant  $\|w\|_{A_{p,\mathbb{T}}}$ . For this, we refer the reader to S. Petermichl and J. Wittwer [39]. The fact that these weights are probabilistic weights for the corresponding martingales has been known for many years. The first author learned this from R. Durrett in the early 1980's.

The Littlewood-Paley  $g_*$  function on the circle  $\mathbb{T}$  is defined by

$$(5.21) \quad g_*(f)(e^{i\theta}) = \left( \frac{1}{\pi} \int_{\mathbb{D}} \frac{1 - |z|^2}{|z - e^{i\theta}|^2} \log \frac{1}{|z|} |\nabla u_f(z)|^2 dz \right)^{1/2},$$

where  $dz$  denotes the area measure in the plane. This version, which is pointwise comparable to the classical Zygmund [50]  $g_*$  function, was introduced in [5]. As in the case of  $\mathbb{R}^n$ , this square function is the conditional expectation of the square function of the martingale  $u_f(B_{\tau \wedge t})$ . That is, by the Itô formula, we have

$$u_f(B_{\tau \wedge t}) = u_f(0) + \int_0^{\tau \wedge t} \nabla u_f(B_s) \cdot dB_s$$

for all  $t$ , and hence this martingale has the square function given by

$$\langle u_f(B) \rangle_{\tau \wedge t} = |u_f(0)|^2 + \int_0^{\tau \wedge t} |\nabla u_f(B_s)|^2 ds.$$

Now, we have

$$g_*^2(f)(e^{i\theta}) = \mathbb{E}_0^\theta \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \right) = \mathbb{E}_0 \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \mid B_\tau = e^{i\theta} \right).$$

We refer the reader to [5, p. 650] for the details on this formula which is proved using the transition probabilities for the Doob  $h$ -process for Brownian motion starting at 0, conditioned to exit  $\mathbb{D}$  at  $e^{i\theta}$ . Since  $B_\tau$  is uniformly distributed on  $\mathbb{T}$  under  $\mathbb{P}_0$ , we have

$$(5.22) \quad \begin{aligned} \frac{1}{2\pi} \int_{\mathbb{T}} g_*^2(f)(e^{i\theta}) w(e^{i\theta}) d\theta &= \mathbb{E}_0 \left( \mathbb{E}_0 \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \mid B_\tau \right) w(B_\tau) \right) \\ &= \mathbb{E}_0 \left( \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \right) w(B_\tau) \right). \end{aligned}$$

**Theorem 5.1.** *Suppose  $w \in A_2(\text{Poisson}, \mathbb{T})$  and  $f \in C(\mathbb{T})$ , the space of continuous functions in  $\mathbb{T}$ . Then,*

$$(5.23) \quad \|f - u_f(0)\|_{L_w^2(\mathbb{T})} \leq (80 \|w\|_{A_{2,\mathbb{T}}})^{1/2} \|g_*(f)(e^{i\theta})\|_{L_w^2(\mathbb{T})},$$

$$(5.24) \quad \|g_*(f)(e^{i\theta})\|_{L_w^2(\mathbb{T})} \leq \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_{r,\mathbb{T}}} \right)^{1/2} \|f\|_{L_w^2(\mathbb{T})}$$

and

$$(5.25) \quad \|g_*(f)(e^{i\theta})\|_{L_w^2(\mathbb{T})} \leq 2^{7/4} \|w\|_{A_{2,\mathbb{T}}} \|f\|_{L_w^2(\mathbb{T})},$$

where  $u_f(0) = \frac{1}{2\pi} \int_{\mathbb{T}} f(e^{i\theta}) d\theta$ .

*Proof.* Applying (5.20) and the inequality (4.2) of Theorem 4.1, we have

$$\begin{aligned} \frac{1}{2\pi} \int_{\mathbb{T}} |f(e^{i\theta}) - u_f(0)|^2 w(e^{i\theta}) d\theta &= \mathbb{E}_0 (|f(B_\tau) - u_f(0)|^2 w(B_\tau)) \\ &= E_0 \left( \left| \int_0^\tau \nabla u_f(B_s) dB_s \right|^2 w(B_\tau) \right) \\ &\leq 80 \|w\|_{A_{2,\mathbb{T}}} \mathbb{E}_0 \left( \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \right) w(B_\tau) \right) \\ &= 80 \|w\|_{A_{2,\mathbb{T}}} \left( \frac{1}{2\pi} \int_{\mathbb{T}} g_*^2(f)(e^{i\theta}) w(e^{i\theta}) d\theta \right), \end{aligned}$$

where the last equality follows from (5.22). This proves (5.23). To establish (5.24) we apply (4.4) to obtain

$$\begin{aligned} \frac{1}{2\pi} \int_{\mathbb{T}} g_*^2(f)(e^{i\theta}) w(e^{i\theta}) d\theta &= \mathbb{E}_0 \left( \left( \int_0^\tau |\nabla u_f(B_s)|^2 ds \right) w(B_\tau) \right) \\ &\leq \mathbb{E}_0 \left( \left( |u_f(0)|^2 + \int_0^\tau |\nabla u_f(B_s)|^2 ds \right) w(B_\tau) \right) \\ &\leq \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_{r,\mathbb{T}}} \right) \mathbb{E}_0 (|f(B_\tau)|^2 w(B_\tau)) \\ &= \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_{r,\mathbb{T}}} \right) \left( \frac{1}{2\pi} \int_{\mathbb{T}} |f(e^{i\theta})|^2 w(e^{i\theta}) d\theta \right). \end{aligned}$$

Finally, (5.25) is proved the same way applying Corollary 4.1.  $\square$

For  $0 < \alpha < 1$  the Stoltz domain, denoted here by  $\Gamma_\alpha(\theta)$ , is the interior of the smallest convex set containing the disc  $\{z \in \mathbb{C} : |z| < \alpha\}$  and the point  $e^{i\theta}$ . The Lusin area function (area integral) of  $f$  is

$$A_\alpha(f)(e^{i\theta}) = \left( \int_{\Gamma_\alpha(\theta)} |\nabla u_f(z)|^2 dz \right)^{1/2}.$$

Similarly, the Littlewood-Paley function  $g$  is defined by

$$g(f)(e^{i\theta}) = \left( \int_0^1 (1-r) |\nabla u_f(re^{i\theta})|^2 dr \right)^{1/2}.$$

As before, it is easy to show that there are universal constant  $C_\alpha$  and  $C$  such that the pointwise inequalities  $A_\alpha(f)(e^{i\theta}) \leq C_\alpha g_*(f)(e^{i\theta})$  and  $g(f)(e^{i\theta}) \leq C g_*(f)(e^{i\theta})$  hold. This gives the following Corollary.

**Corollary 5.1.** *Suppose  $w \in A_2(\text{Poisson}, \mathbb{T})$  and  $f \in C(\mathbb{T})$ . Then*

$$(5.26) \quad \|A_\alpha(f)(e^{i\theta})\|_{L^2(w)} \leq 2^{7/4} C_\alpha \|w\|_{A_{2,\mathbb{T}}} \|f\|_{L^2(w)}$$

and

$$(5.27) \quad \|g(f)(e^{i\theta})\|_{L^2(w)} \leq 2^{7/4}C\|w\|_{A_{2,T}}\|f\|_{L^2(w)}.$$

**5.4. Poisson and heat  $A_p$  weights on  $\mathbb{R}^n$ .** In this section we carry out the computations in  $\mathbb{R}^n$  done above for the disc. We follow the notation of §5.1.

**Definition 5.2.** Let  $w$  be a positive locally integrable function defined on  $\mathbb{R}^n$  and  $1 < p < \infty$ .

(i) We will say  $w \in A_p(\text{Poisson}, \mathbb{R}^n)$  if

$$(5.28) \quad \|w\|_{A_p(\text{Poisson}, \mathbb{R}^n)} := \|K_y w(x) \left( K_y(w^{-1/(p-1)})(x) \right)^{p-1}\|_{L^\infty(\mathbb{R}_+^{n+1})} < \infty.$$

(ii) We will say  $w \in A_p(\text{heat}, \mathbb{R}^n)$  if

$$(5.29) \quad \|w\|_{A_p(\text{heat}, \mathbb{R}^n)} := \|P_t w(x) \left( P_t(w^{-1/(p-1)})(x) \right)^{p-1}\|_{L^\infty(\mathbb{R}_+^{n+1})} < \infty.$$

We remark here that these  $A_p$  weights can be defined for any Markovian semigroup and not just for the Poisson or heat semigroup in  $\mathbb{R}^n$ . Such  $A_p$  weights are nothing more than the martingale  $A_p(\text{mart})$  weights arising from the stochastic process associated with the semigroup. Before we explain this more precisely, we recall that both classes of  $A_p$  weights defined above have been studied before in connection to weight problems and applications. Indeed, it was proved by S. Petermichl and A. Volberg in [38] that there are constants  $a$  and  $b$ , depending only on the dimension  $n$ , such that

$$(5.30) \quad aA_p(\text{heat}, \mathbb{R}^n) \leq \|w\|_{A_p} \leq bA_p(\text{heat}, \mathbb{R}^n),$$

where  $\|w\|_{A_p}$  is as in the original definition of Muckenhoupt. That is,  $w \in A_p$  if

$$\|w\|_{A_p} = \sup_Q \left( \frac{1}{|Q|} \int_Q w \right) \left( \frac{1}{|Q|} \int_Q w^{-1/(p-1)} \right)^{p-1} < \infty,$$

where the sup is taken over all cubes  $Q \subset \mathbb{R}^n$ . As for  $A_p(\text{Poisson}, \mathbb{R}^n)$ , it is known that when  $n = 1$ ,  $A_2(\text{Poisson}) = A_2$  and that in fact there are universal constants  $a$  and  $b$  such that

$$(5.31) \quad a\|w\|_{A_2} \leq \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \leq b\|w\|_{A_2}^2.$$

On the other hand, for  $n > 1$  there are weights  $w$  for which we have  $\|w\|_{A_2} < \infty$ , but  $\|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} = \infty$ . Thus, for  $n > 1$ ,  $A_2 \neq A_2(\text{Poisson}, \mathbb{R}^n)$ . For these results, as well as the boundedness of the classical Riesz transforms on  $L_w^2(\mathbb{R}^n)$ ,  $w \in A_2(\text{Poisson})$  with constants independent of the dimension  $n$ , we refer the reader to Huković [21], Petermichl [39] and K. Domelevo, Petermichl and Wittwer [40].

**Remark 5.2.** We remark here that while  $A_2 \neq A_2(\text{Poisson}, \mathbb{R}^n)$ , it is easy to see that  $A_2(\text{Poisson}, \mathbb{R}^n) \subset A_2$  for all  $n \geq 1$ . Indeed, given a cube  $Q$  centered at  $z$  and length  $l_Q$ , we pick  $a \approx l_Q$  to obtain that  $\frac{1}{|Q|}\chi_Q \leq C_n k_a(z-x)$  for all  $x \in Q$  for some universal constant  $C_n$  depending only on  $n$ . This immediately shows that  $A_2(\text{Poisson}, \mathbb{R}^n) \subset A_2$  for all  $1 < p < \infty$  and  $n \geq 1$ . The same argument (picking this time  $t^2 \approx l_Q$ ) shows that  $\frac{1}{|Q|}\chi_Q \leq C'_n p_t(z-x)$  for all  $x \in Q$ . This gives that  $A_p(\text{heat}, \mathbb{R}^n) \subset A_p$  for all  $1 < p < \infty$  and  $n \geq 1$ .

Our aim now is to prove versions of Theorem 4.1 and Corollary 4.1 for the Littlewood-Paley functions  $g_*$  and  $G_*$  with respect to weights in  $A_2(\text{Poisson}, \mathbb{R}^n)$  and  $A_2(\text{heat}, \mathbb{R}^n)$ .

**Lemma 5.1.** *Suppose  $w \in A_p(\text{Poisson}, \mathbb{R}^n)$ . Fix  $(z, a) \in \mathbb{R}_+^{n+1}$  and let  $B_t = (X_t, Y_t)$ ,  $t \geq 0$ , be Brownian motion  $\mathbb{R}_+^{n+1}$  starting at  $(z, a)$  and denote by  $\tau$  its exit time from  $\mathbb{R}_+^{n+1}$ . Let  $\tilde{Y}_t = K_{Y_{\tau \wedge t}} w(X_{\tau \wedge t})$ ,  $t \geq 0$ , be the martingale under the measure  $\mathbb{P}_{(z, a)}$ . Then  $\tilde{Y}_\infty = w(X_\tau) \in A_p(\text{mart})$  and  $\|\tilde{Y}_\infty\|_{A_p(\text{mart})} \leq \|w\|_{A_p(\text{Poisson}, \mathbb{R}^n)}$*

*Proof.* With  $\tilde{Y}_\infty = w(X_\tau)$  and  $K_a w(z) = \mathbb{E}_{(z, a)} w(X_\tau)$ , the Strong Markov property gives

$$\begin{aligned} \mathbb{E}_{(z, a)} \left( \left( \frac{1}{\tilde{Y}_\infty} \right)^{1/(p-1)} \middle| \mathcal{F}_{\tau \wedge t} \right) &= \mathbb{E}_{(z, a)} \left( \left( \frac{1}{w(X_\tau)} \right)^{1/(p-1)} \middle| \mathcal{F}_{\tau \wedge t} \right) \\ &= \mathbb{E}_{B_{\tau \wedge t}} \left( \left( \frac{1}{w(X_\tau)} \right)^{1/(p-1)} \right) \\ &= K_{Y_{\tau \wedge t}} (w^{-1/(p-1)})(X_{\tau \wedge t}). \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{Y}_t \left( \mathbb{E}_{(z, a)} \left( \left( \frac{1}{\tilde{Y}_\infty} \right)^{1/(p-1)} \middle| \mathcal{F}_{\tau \wedge t} \right) \right)^{p-1} \\ = K_{Y_{\tau \wedge t}} w(X_{\tau \wedge t}) \left( K_{Y_{\tau \wedge t}} (w^{-1/(p-1)})(X_{\tau \wedge t}) \right)^{p-1}. \end{aligned}$$

It follows from this that for all  $(z, a) \in \mathbb{R}_+^{n+1}$ ,

$$\begin{aligned} \|\tilde{Y}_\infty\|_{A_p(\text{mart})} &= \sup_{t \geq 0} \left\| \tilde{Y}_t \left( \mathbb{E}_{(z, a)} \left( \left( \frac{1}{\tilde{Y}_\infty} \right)^{1/(p-1)} \middle| \mathcal{F}_{\tau \wedge t} \right) \right)^{p-1} \right\|_\infty \\ &= \sup_{t > 0} \|K_{Y_{\tau \wedge t}} w(X_{\tau \wedge t}) \left( K_{Y_{\tau \wedge t}} (w^{-1/(p-1)})(X_{\tau \wedge t}) \right)^{p-1}\|_\infty \\ &\leq \|w\|_{A_p(\text{Poisson}, \mathbb{R}^n)}. \end{aligned}$$

This completes the proof.  $\square$

**Remark 5.3.** It is important to note here, for our applications below, that the above inequality  $\|\tilde{Y}_\infty\|_{A_p(\text{mart})} \leq \|w\|_{A_p(\text{Poisson}, \mathbb{R}^n)}$  holds for all starting points  $(z, a)$ . While not needed for the purpose of this paper, we note that here we actually have equality. That is,  $\|\tilde{Y}_\infty\|_{A_p(\text{mart})} = \|w\|_{A_p(\text{Poisson}, \mathbb{R}^n)}$ . This follows from the fact that if  $F(x, y)$  is a continuous bounded function in the upper half-space then

$$\sup_t \|F(X_{\tau \wedge t}, Y_{\tau \wedge t})\|_{L^\infty(\mathbb{P}_{(z, a)})} = \|F(x, y)\|_{L^\infty(\mathbb{R}_+^{n+1})},$$

since given any ball  $B$  in the upper-half space there will be a time  $t > 0$  such that  $\mathbb{P}_{(z, a)}\{(X_{\tau \wedge t}, Y_{\tau \wedge t}) \in B\} > 0$ . Indeed, this quantity is given by the integral of the Dirichlet heat kernel in the upper half-space (which is just the product of the heat kernel in  $\mathbb{R}^n$  and heat kernel for the half line) over the ball  $B$ .

**Theorem 5.2.** *Suppose  $w \in A_2(\text{Poisson}, \mathbb{R}^n)$  and  $f \in C_0(\mathbb{R}^n)$ , the space of continuous functions of compact support. Then,*

$$(5.32) \quad \|f\|_{L_w^2(\mathbb{R}^n)} \leq (320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)})^{1/2} \|g_*(f)\|_{L_w^2(\mathbb{R}^n)},$$

$$(5.33) \quad \|g_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq \frac{1}{\sqrt{2}} \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_r(\text{Poisson}, \mathbb{R}^n)} \right)^{1/2} \|f\|_{L_w^2(\mathbb{R}^n)}$$

and

$$(5.34) \quad \|g_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq 2^{5/4} \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}.$$

*Proof.* Let  $R$  be large enough so that the support of  $f$  is contained in the ball  $B = B(0, R)$ . By (5.8) we have

$$(5.35) \quad \begin{aligned} \int_{\mathbb{R}^n} |f(x)|^2 w(x) dx &= \int_{\mathbb{R}^n} \mathbf{1}_B(x) |f(x)|^2 w(x) dx \\ &= \int_{\mathbb{R}^n} E_{(z,a)} (\mathbf{1}_B(X_\tau) |f(X_\tau)|^2 w(X_\tau)) dz \\ &\leq 2 \int_{\mathbb{R}^n} E_{(z,a)} (\mathbf{1}_B(X_\tau) |f(X_\tau) - K_a f(z)|^2 w(X_\tau)) dz \\ &\quad + 2 \int_{\mathbb{R}^n} |K_a f(z)|^2 \mathbb{E}_{(z,a)} \mathbf{1}_B(X_\tau) w(X_\tau) dz. \end{aligned}$$

We now estimate the first term under the integral on the right hand side of the above inequality. By Theorem 4.1 and Lemma 5.1 we have

$$\begin{aligned} &2 \mathbb{E}_{(z,a)} (|f(X_\tau) - K_a f(z)|^2 w(X_\tau)) \\ &\leq 160 \|w\|_{A_2(\text{Poisson})} \mathbb{E}_{(z,a)} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds w(X_\tau) \right) \\ &= 160 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \mathbb{E}_{(z,a)} \left( \mathbb{E}_{(z,a)} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds w(X_\tau) \middle| X_\tau \right) \right) \\ &= 160 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \mathbb{E}_{(z,a)} \left( \mathbb{E}_{(z,a)} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds \right) \middle| X_\tau \right) w(X_\tau) \\ &= 160 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \mathbb{E}_{(z,a)} \left( \mathbb{E}_{(z,a)}^{X_\tau} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds \right) w(X_\tau) \right) \\ &= 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \mathbb{E}_{(z,a)} [g_{*,a}^2(f)(X_\tau) w(X_\tau)]. \end{aligned}$$

Integrating both sides of this inequality in  $z$  gives

$$\begin{aligned} &2 \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)} (|f(X_\tau) - K_a f(z)|^2 w(X_\tau)) dz \\ &\leq 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)} [g_{*,a}^2(f)(X_\tau) w(X_\tau)] dz \\ &= 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \int_{\mathbb{R}^n} g_{*,a}^2(f)(x) w(x) dx \\ &\leq 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \int_{\mathbb{R}^n} g_*^2(f)(x) w(x) dx. \end{aligned}$$

Combining this with (5.35) we obtain

$$(5.36) \quad \begin{aligned} \int_{\mathbb{R}^n} |f(x)|^2 w(x) dx &\leq 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \int_{\mathbb{R}^n} g_*^2(f)(x) w(x) dx \\ &\quad + 2 \int_{\mathbb{R}^n} |K_a f(z)|^2 \mathbb{E}_{(z,a)} (\mathbf{1}_B(X_\tau) w(X_\tau)) dz. \end{aligned}$$

Since  $f \in C_0(\mathbb{R}^n)$ , we have

$$|K_a f(z)| = \left| \int_{\mathbb{R}^n} k_a(x-z) f(x) dx \right| \leq \frac{c_n}{a^n} \int_{\mathbb{R}^n} |f(x)| dx.$$

Thus,

$$\begin{aligned} & 2 \int_{\mathbb{R}^n} |K_a f(z)|^2 \mathbb{E}_{(z,a)}(\mathbf{1}_B(X_\tau) w(X_\tau)) dz \\ & \leq \frac{2c_n}{a^n} \left( \int_{\mathbb{R}^n} |f(x)| dx \right) \left( \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)}(\mathbf{1}_B(X_\tau) w(X_\tau)) dz \right) \\ & = \frac{2c_n}{a^n} \left( \int_{\mathbb{R}^n} |f(x)| dx \right) \left( \int_B w(x) dx \right). \end{aligned}$$

Combining this with (5.36) and letting  $a \rightarrow \infty$  gives

$$\int_{\mathbb{R}^n} |f(x)|^2 w(x) dx \leq 320 \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \int_{\mathbb{R}^n} g_*^2(f)(x) w(x) dx,$$

which is the announced inequality.

Similarly, inequality (4.4) in Theorem 4.1 gives

$$\begin{aligned} & \int_{\mathbb{R}^n} g_{*,a}^2(f)(x) w(x) dx \\ & = \frac{1}{2} \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)} \left( \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds w(X_\tau) \right) dz \\ & \leq \frac{1}{2} \int_{\mathbb{R}^n} E_{(z,a)} \left( \left( |K_a f(z)|^2 + \int_0^\tau |\nabla K_{Y_s}(f)(X_s)|^2 ds \right) w(X_\tau) \right) dz \\ & \leq \frac{1}{2} \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_r(\text{Poisson}, \mathbb{R}^n)} \right) \int_{\mathbb{R}^n} \mathbb{E}_{(z,a)} |f(X_\tau)|^2 w(X_\tau) dz \\ & = \frac{1}{2} \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_r(\text{Poisson}, \mathbb{R}^n)} \right) \int_{\mathbb{R}^n} |f(x)|^2 w(x) dx. \end{aligned}$$

Combining the above arguments with Corollary 4.1 we obtain (5.34) and this completes the proof of the theorem.  $\square$

From the pointwise inequalities (5.6) and (5.7), combined with (5.34), we obtain

**Corollary 5.2.** *Suppose  $w \in A_2(\text{Poisson}, \mathbb{R}^n)$  and  $f \in C_0(\mathbb{R}^n)$ . Then*

$$(5.37) \quad \|A_\alpha(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq \frac{(\alpha^2 + 1)^{\frac{n+1}{4}}}{\sqrt{c_n}} 2^{5/4} \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}$$

and

$$(5.38) \quad \|g(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq 2^{9/4} \|w\|_{A_2(\text{Poisson}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}.$$

Our results for  $A_p(\text{heat}, \mathbb{R}^n)$  weights parallel those for  $A_p(\text{Poisson}, \mathbb{R}^n)$ . We start with the corresponding lemma which shows the identification of these weights with the martingale weights arising from the semigroup.

**Lemma 5.2.** *Suppose  $w \in A_p(\text{heat}, \mathbb{R}^n)$ . Fix  $0 < T < \infty$  and  $z \in \mathbb{R}^n$ . Let  $(B_t)_{t \geq 0}$  be Brownian motion in  $\mathbb{R}^n$  starting at  $z$ . Consider the martingale  $Y_t = P_{T-t} w(B_t)$ ,  $0 \leq t \leq T$ , under the measure  $\mathbb{P}_z$ . Then  $Y_T = w(B_T) \in A_p(\text{mart})$  and  $\|Y_T\|_{A_p(\text{mart})} \leq \|w\|_{A_p(\text{heat}, \mathbb{R}^n)}$ .*



*Proof.* Recall that  $P_t w(z) = \mathbb{E}_z w(B_t)$ . As before, we apply the Strong Markov property to obtain that for any  $0 < t < T$ ,

$$\begin{aligned} \mathbb{E}_z \left( \left( \frac{1}{Y_T} \right)^{1/(p-1)} \middle| \mathcal{F}_t \right) &= \mathbb{E}_z \left( \left( \frac{1}{w(B_T)} \right)^{1/(p-1)} \middle| \mathcal{F}_t \right) \\ &= \mathbb{E}_{B_t} \left( \left( \frac{1}{w(X_{T-t})} \right)^{1/(p-1)} \right) \\ &= P_{T-t}(w^{-1/(p-1)})(B_t). \end{aligned}$$

Thus,

$$Y_t \left( \mathbb{E}_z \left( \left( \frac{1}{Y_T} \right)^{1/(p-1)} \middle| \mathcal{F}_t \right) \right)^{p-1} = P_{T-t} w(B_t) \left( P_{T-t}(w^{-1/(p-1)})(B_t) \right)^{p-1}$$

and

$$\|Y_T\|_{A_p(\text{mart})} = \sup_{0 < t < T} \|Y_t \left( \mathbb{E}_z \left( \left( \frac{1}{Y_T} \right)^{1/(p-1)} \middle| \mathcal{F}_t \right) \right)^{p-1}\|_{\infty} \leq \|w\|_{A_p(\text{heat}, \mathbb{R}^n)},$$

as claimed.  $\square$

As before, a remark similar to Remark 5.3 applies. With this lemma established, we can repeat the above argument for the space-time martingales and obtain similar results for  $\|w\|_{A_p(\text{heat}, \mathbb{R}^n)}$  weights.

**Theorem 5.3.** *Suppose  $w \in A_2(\text{heat}, \mathbb{R}^n)$  and  $f \in C_0(\mathbb{R}^n)$ . Then*

$$(5.39) \quad \|f\|_{L_w^2(\mathbb{R}^n)} \leq (160 \|w\|_{A_2(\text{heat}, \mathbb{R}^n)})^{1/2} \|G_*(f)\|_{L_w^2(\mathbb{R}^n)},$$

$$(5.40) \quad \|G_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq \inf_{1 < r < 2} \left( \frac{r}{2-r} \|w\|_{A_r(\text{heat}, \mathbb{R}^n)} \right)^{1/2} \|f\|_{L_w^2(\mathbb{R}^n)}$$

and

$$(5.41) \quad \|G_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq 2^{7/4} \|w\|_{A_2(\text{heat}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}.$$

We remark that the reason the constants here are slightly different than those for the Poisson case is that the representation for  $G_{*,T}$  in terms of the conditional expectation of the corresponding martingale square function given in (5.18) does not have the  $\frac{1}{2}$  factor as in (5.9).

From the inequalities  $\mathcal{P}A_\alpha(f)(x) \leq (2\pi)^{n/4} e^{\alpha^2/4} G_*(f)(x)$  and  $G(f)(x) \leq \sqrt{2} G_*(f)(x)$  already proved in (5.15) and (5.16), we obtain

**Corollary 5.3.** *Suppose  $w \in A_2(\text{heat}, \mathbb{R}^n)$  and  $f \in C_0(\mathbb{R}^n)$ . Then*

$$(5.42) \quad \|\mathcal{P}A_\alpha(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq (2\pi)^{n/4} e^{\alpha^2/4} 2^{7/4} \|w\|_{A_2(\text{heat}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}$$

and

$$(5.43) \quad \|G(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq 2^{9/4} \|w\|_{A_2(\text{heat}, \mathbb{R}^n)} \|f\|_{L_w^2(\mathbb{R}^n)}.$$

Theorem 5.3 and Corollary 5.3, combined with the Petermichl–Volberg inequality (5.30) proving the equivalence of the classical Muckenhoupt  $A_p$  and  $A_p(\text{heat})$ , give

**Corollary 5.4.** *Suppose  $w \in A_2$  is in the classical Muckenhoupt class and  $f \in C_0(\mathbb{R}^n)$ . Then*

$$(5.44) \quad \|f\|_{L_w^2(\mathbb{R}^n)} \leq \left( \frac{160\|w\|_{A_2}}{a} \right)^{1/2} \|G_*(f)\|_{L_w^2(\mathbb{R}^n)},$$

$$(5.45) \quad \|G_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq \inf_{1 < r < 2} \left( \frac{r}{2-r} \frac{\|w\|_{A_r}}{a} \right)^{1/2} \|f\|_{L_w^2(\mathbb{R}^n)},$$

$$(5.46) \quad \|G_*(f)\|_{L_w^2(\mathbb{R}^n)} \leq \frac{2^{7/4}\|w\|_{A_2}}{a} \|f\|_{L_w^2(\mathbb{R}^n)},$$

$$(5.47) \quad \|\mathcal{P}A_\alpha(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq \frac{(2\pi)^{n/4} e^{\alpha^2/4} 2^{7/4} \|w\|_{A_2}}{a} \|f\|_{L_w^2(\mathbb{R}^n)}$$

and

$$(5.48) \quad \|G(f)(x)\|_{L_w^2(\mathbb{R}^n)} \leq \frac{2^{9/4}\|w\|_{A_2}}{a} \|f\|_{L_w^2(\mathbb{R}^n)},$$

where  $a$  is the constant in (5.30).

## 6. SYMMETRIC MARKOVIAN SEMIGROUPS

In [46], Varopoulos defines the  $g_*$  function in the general setting of Poisson semigroups. However, due to the lack of gradient in this general setting, he only considers the time derivative of the semigroup in the definition of his square functions, for both his  $g$  and  $g_*$ . This construction can be applied to obtain versions of the above inequalities for semigroups which yield martingales with continuous paths. In this section we aim to define a Littlewood-Paley function  $G_*$  for general Markovian and the corresponding  $A_p$  weights. Since our martingale results require continuous trajectories, our  $A_2$  inequality will only be stated for Riemannian manifolds of non-negative Ricci curvature, using Meyer's "carre du champ."

Let  $(M, \mathcal{M}, \mu)$  be a metric measure space. That is, a measure space (equipped with a countably generated  $\sigma$ -algebra)  $\mathcal{M}$  which is also a metric space with metric  $\rho$ . The measure  $\mu$  is assumed to be  $\sigma$ -finite. Let  $(P_t, t \geq 0)$  be a family of Markovian linear operators which acts as a  $C_0$ -contraction semigroup on  $L^p(M)$  for all  $1 \leq p \leq \infty$ . We further assume that  $P_t$  is self-adjoint on  $L^2(M)$  for all  $t \geq 0$  and that it is given by an integral kernel

$$P_t f(x) = \int_M p_t(x, y) f(y) d\mu(y)$$

which is symmetric. That is,  $p_t(x, y) = p_t(y, x)$  and

$$\int_M p_t(x, y) d\mu(y) = 1.$$

It follows from [14] that  $T_t = e^{-t\mathcal{L}}$  where  $\mathcal{L}$  is a positive self-adjoint operator on  $L^2(M)$ . If we denote by  $D(\mathcal{L}) \subset L^2(M)$  the domain of  $\mathcal{L}$  the for  $f, h \in D(\mathcal{L})$ , define the operator "carré du champ" is defined by

$$(6.1) \quad \Gamma(f, h) = \mathcal{L}(fh) - f\mathcal{L}h - h\mathcal{L}f.$$

By the definition of  $\Gamma$ , we have

$$\begin{aligned}
\int_0^\infty \int_M \Gamma(P_t f, P_t f)(x) d\mu(x) dt &= -2 \int_0^\infty \int_M P_t f(x) \mathcal{L} P_t f(x) d\mu(x) dt \\
&= -2 \int_0^\infty \int_M P_t f(x) \frac{d}{dt} P_t f(x) d\mu(x) dt \\
&= - \int_0^\infty \int_M \frac{d}{dt} (P_t f(x))^2 d\mu(x) dt \\
&= \int_M |f(x)|^2 d\mu(x).
\end{aligned}$$

Defining, respectively, the Littlwood-Paley  $G$  and  $G_*$  by

$$(6.2) \quad G(f)(x) = \left( \int_0^\infty \Gamma(P_t f, P_t f)(x) dt \right)^{1/2}$$

and

$$(6.3) \quad G_*(f)(x) = \left( \int_0^\infty \int_M \Gamma(P_t f, P_t f)(x) p_t(x, z) d\mu(z) dt \right)^{1/2},$$

we see that

$$\|G_* f\|_2 = \|G(f)\|_2 = \|f\|_2.$$

We now denote by  $(X_t)_{t \geq 0}$  the Markov process associated with this semigroup so that  $P_t f(x) = \mathbb{E}_x[f(X_t)]$ , and consider the martingale

$$(6.4) \quad M_t(f) = P_{T-t} f(X_t), \quad 0 \leq t \leq T.$$

Under quite general conditions on the Markovian semigroup (as those imposed on [46]), the process  $(X_t)_{t \geq 0}$  has càdlàg paths,  $G_*$  is the conditional expectation of the square function for this martingale (see Bakry Emery [4, p. 181] or Revuz and Yor, [43, p. 326]) and it follows from the Burkholder-Gundy inequalities that  $\|G_* f(x)\|_p \leq C_p \|f\|_p$ ,  $2 \leq p < \infty$ , where  $C_p$  depends only on  $p$ .

For the remaining of this paper we will make the further assumption that our Markovian semigroup corresponds to Brownian motion on a complete Riemannian manifold of non-negative Ricci curvature and therefore the process has continuous paths. To be precise, we let  $M$  be a complete Riemannian manifold of dimension  $n$  with non-negative Ricci curvature. Let  $\Delta$  be the Laplace-Beltrami operator and  $\mu$  be the Riemannian volume measure. Then the heat equation  $\frac{\partial u}{\partial t} = \Delta u(t)$  has a fundamental solution  $p \in C^\infty((0, \infty) \times M \times M)$  which we call the *heat kernel* and this gives the kernel generating our semigroup  $(P_t)_{t \geq 0}$  above. The following heat kernel bounds of Li and Yau [28] are important for many applications. For all  $t > 0, x, z \in M$ :

$$(6.5) \quad \frac{C_1}{V(x, \sqrt{t})} \exp\left(-\frac{\rho(x, y)^2}{c_1 t}\right) \leq p_t(x, y) \leq \frac{C_2}{V(x, \sqrt{t})} \exp\left(-\frac{\rho(x, y)^2}{c_2 t}\right),$$

where  $\rho$  is the Riemannian metric and for  $r > 0, V(x, r) = \mu(B(x, r))$  is the volume of the ball  $B(x, r)$  of radius  $r$  centered at  $x$ . It is also well-known (cf. [9]) that for all  $x \in M$ ,

$$(6.6) \quad V(x, r) \leq v(n)r^n,$$

where  $v(n)$  is the volume of the unit ball in  $\mathbb{R}^d$ .

With the Laplacian as the generator, the carré du champ has the familiar form

$$\Gamma(P_t f, P_t f)(x) = |\nabla P_t f(x)|^2$$

and the square function of the martingale  $M_f(f)$  is given by

$$\langle M(f) \rangle_t = |P_T f(x)|^2 + \int_0^t |\nabla_x P_{T-s} f(X_s)|^2 ds, \quad t \geq 0.$$

With this, the exact same argument as in  $\mathbb{R}^n$  gives that

$$\begin{aligned} G_{*,T}(f)(x) &= \int_0^T \int_{\mathbb{R}^d} |\nabla_x P_t f(z)|^2 p_t(x, z) d\mu(z) dt \\ (6.7) \quad &= \int_{\mathbb{R}^n} \mathbb{E}_z^x \left( \int_0^T |\nabla_x P_{T-s} f(X_s)|^2 ds \right) p_T(x, z) d\mu(z). \end{aligned}$$

As before we have the pointwise inequality

$$(6.8) \quad G(f)(x) \leq \sqrt{2} G_*(f)(x).$$

To prove this we recall that under the assumption of non-negative Ricci curvature, the ‘‘Bakry  $\Gamma_2 \geq 0$ ’’ holds. That is, we have the inequality  $\Gamma(P_t f, P_t f) \leq P_t \Gamma(f, f)$  (see [3] for details). From this and the semigroup property, we obtain

$$\begin{aligned} \Gamma(P_t f, P_t f)(x) &= \Gamma(P_{t/2} T_{t/2} f, P_{t/2} P_{t/2} f)(x) \\ &\leq P_{t/2} \Gamma(T_{t/2} f, P_{t/2} f)(x) \\ &= \int_M \Gamma(T_{t/2} f, P_{t/2} f)(y) p_{t/2}(x, y) d\mu(y). \end{aligned}$$

Integrating both sides of this inequality in  $t$  gives (6.8).

Next, we introduce the parabolic cone using the metric on the manifold by

$$\mathcal{P}_\alpha(x) = \{(z, t) \in \mathbb{R}_+^{n+1} : d(x, z) < \alpha \sqrt{t}\}$$

and define the Lusina area integral by

$$(6.9) \quad \mathcal{P}A_\alpha(f)(x) = \left( \int_{\mathcal{P}_\alpha(x)} t^{-n/2} |\nabla P_t f(z)|^2 d\mu(z) dt \right)^{1/2}.$$

By (6.5) and (6.6) we have

$$(6.10) \quad \mathcal{P}A_\alpha(f)(x) \leq \sqrt{\frac{v(n)e^{\frac{\alpha^2}{2c_1}}}{C_1}} G_*(f)(x).$$

Given a positive and  $\mu$ -locally integrable function  $w$  on  $M$ , we will write  $w \in A_p(\text{heat}, M)$  if

$$(6.11) \quad \|w\|_{A_p(\text{heat}, M)} := \|P_t w(x) \left( P_t(w^{-1/(p-1)})(x) \right)^{p-1}\|_{L^\infty(M \times (0, \infty))} < \infty.$$

The same argument as that in Lemma 5.2 shows that

$$\|Y_T\|_{A_p(\text{mart})} \leq \|w\|_{A_p(\text{heat}, M)},$$

where  $Y$  stands for the martingale  $Y_t = P_{T-t} w(X_t)$ ,  $0 \leq t \leq T$ .

Similarly, we say that  $w \in A_p(M)$  (the classical Muckenhoupt  $A_p$ -class) if

$$\|w\|_{A_p(M)} = \sup_B \left( \frac{1}{\mu(B)} \int_B w(z) d\mu(z) \right) \left( \frac{1}{\mu(B)} \int_B w(z)^{-1/(p-1)} d\mu(z) \right)^{p-1} < \infty,$$

where the supremum is taking over balls. Because of the bound on the heat kernel  $p_t(x, z)$  given in (6.5), the observations of Remark 5.2 show that

$$A_p(M) \leq aA_p(\text{heat}, M),$$

for some constant  $a$  depending on  $c_1, C_1$ .

With the above definitions in place, we now state the following version of Theorem 5.3, whose proof is exactly the same as the proof of that theorem.

**Theorem 6.1.** *Let  $M$  be a complete Riemannian manifold of non-negative Ricci curvature. Assume further that (\*)  $\sup_{x \in M} p_t(x, x) = c_t \rightarrow 0$ , as  $t \rightarrow \infty$ , holds true. Suppose  $w \in A_2(\text{heat}, M)$  and  $f \in C_0(M)$ . Then inequalities (5.39), (5.40), (5.41), (5.42) and (5.43) hold for the functions  $G(f)$ ,  $G_*(f)$  and  $\mathcal{P}A_\alpha(f)$  as defined in (6.2), (6.3) and (6.9).*

**Remark 6.1.** If in addition we assume that  $V(x, r) \geq c_n r^n$ , then (\*) is automatically satisfied. For various known conditions that guarantee this lower bound volume growth, see [47, p. 255]) and [2].

#### ACKNOWLEDGMENT

The research was initiated during the fall semester of 2013 when the second-named author visited Purdue University.

#### REFERENCES

- [1] M. Abramowitz and I. A. Stegun, editors, Handbook of Mathematical Functions with formulas, graphs and mathematical tables, Reprint of the 1972 edition, Dover Publications, Inc., New York, 1992.
- [2] P. Auscher, T. Coulhon, X.T. Duong and S. Hofmann, *Riesz transforms on manifolds and heat kernel regularity*, Ann. Scient. Éc. Norm. Sup. **37**, 911-957 (2004)
- [3] D. Bakry, *Functional inequalities for Markov semigroups*, In Probability measures on groups: Recent directions and trends, 91-147. Proceedings of the CIMPA-TIFR School (2002), Tata Institute of Foundational Research. Narosha Publishing House, New Delhi (2006).
- [4] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités (Strasbourg), **19** (1985), 177-206.
- [5] R. Bañuelos, *Brownian motion and area functions*, Indiana Univ. Math. J. **35** (1986), 643-668.
- [6] R. Bañuelos and P. Méndez-Hernández, *Space-time Brownian motion and the Beurling-Ahlfors transform*, Indiana University Math J. **52** (2003), 981-990.
- [7] R. Bañuelos and C. Moore, *Probabilistic behavior of harmonic functions*, Birkhäuser, 1999.
- [8] A. Bennett, *Probabilistic square functions and a priori estimates*, Trans. Amer. Math. Soc. **291** (1985), 159-166.
- [9] R. L. Bishop and R. J. Crittenden, *Geometry on Manifolds*, Academic Press, New York (1964).
- [10] B. Bollobás, *Martingale inequalities*, Math. Proc. Cambridge Phil. Soc. **87** (1980), 377-382.
- [11] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253-272.
- [12] R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. **51** (1974), 241-250.
- [13] D. Cruz-Uribe, J. M. Martell and C. Pérez, *Sharp weighted estimates for classical operators*, Adv. in Math. **229** (2012), 408-441.
- [14] E. B. Davies, *Heat Kernels and Spectral Theory*, Cambridge University Press, Cambridge, 1989.
- [15] B. Davis, *On the  $L^p$  norms of stochastic integrals and other martingales*, Duke Math. J. **43** (1976), 697-704.
- [16] C. Dellacherie and P.-A. Meyer, *Probabilities and potential B: Theory of martingales*, North Holland, Amsterdam, 1982.
- [17] J. L. Doob, *Stochastic processes*, John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953.
- [18] L. Grafakos, *Classical Fourier analysis*, Second Edition, Springer, New York, 2008.
- [19] S. V. Hrusčev, *A description of weights satisfying the  $A_\infty$  condition of Muckenhoupt*, Proc. Amer. Math. Soc. **90** (1984), 253-257.
- [20] S. Hukovic, *Singular integral operators in weighted spaces and Bellman functions*, Doctoral Thesis.

- [21] S. Hukovic, S. Treil and A. Volberg, *The Bellman functions and sharp weighted inequalities for square functions*, Complex analysis, operators, and related topics, 97–113, Oper. Theory Adv. Appl., **113**, Birkhäuser, Basel, 2000.
- [22] M. Izumisawa and N. Kazamaki, *Weighted norm inequalities for martingales*, Tôhoku Math. Journ. **29** (1977), 115–124.
- [23] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure and Appl. Math. **14** (1961), 415–426.
- [24] A. K. Lerner, *On sharp aperture-weighted estimates for square functions*, J. Fourier Anal. Appl. **20** (2014), 784–800.
- [25] A. K. Lerner, *Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals*, Adv. Math. **226** (2011), 3912–3926.
- [26] A. K. Lerner, *On some weighted norm inequalities for Littlewood-Paley operators*, Illinois J. Math. **52** (2008), 653–666.
- [27] J. E. Littlewood, *On bounded bilinear forms in an infinite number of variables*, Quart. J. Math. Oxford **1** (1930), 164–174.
- [28] P. Li, S-T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** 153-201 (1986).
- [29] J. Marcinkiewicz, *Quelques théorèmes sur les séries orthogonales*, Ann. Soc. Polon. Math., **16** (1937), 84–96.
- [30] P. A. Meyer, *Retour sur la théorie de Littlewood-Paley*, Séminaire de probabilités, XV, Lect. Notes in Math., **850**, Springer, Berlin (1979/20), 151-166.
- [31] P. A. Meyer, *Démonstration probabiliste de certaines inégalités de Littlewood-Paley*, in “Sém. Probab. X” Lecture Notes in Mathematics **511**, Springer-Verlag, Berlin, New York.
- [32] P. A. Meyer, *Transformations de Riesz pour les lois gaussiennes*, Séminaire de Probab., XVIII, Lect. Notes in Math., **1059**, Springer, Berlin (1984), 179–193.
- [33] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226.
- [34] F. L. Nazarov and S. R. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis*, St. Petersburg Math. J. **8** (1997), 721–824.
- [35] A. Osękowski, *Weighted maximal inequalities for martingales*, Tohoku Math. J. **65** (2013), 75–91.
- [36] R. E. A. C. Paley, *A remarkable series of orthogonal functions I*, Proc. London Math. Soc., **34** (1932), 241–264.
- [37] S. Petermichl and S. Pott, *An estimate for weighted Hilbert transform via square functions*, Trans. Amer. Math. Soc. **354** (2002), 1699–1703.
- [38] S. Petermichl and A. Volberg, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. **112** (2002), 281–305.
- [39] S. Petermichl and J. Wittwer, *A Sharp Estimate for the Weighted Hilbert Transform via Bellman Functions*, Michigan Math. **50** (2002), 71–87.
- [40] K. Domelevo, S. Petermichl, and J. Wittwer, *A linear dimensionless bound for the weighted Riesz vector*, <http://arxiv.org/pdf/1501.07880.pdf>
- [41] G. Pisier and Q. Xu, *Non-commutative martingale inequalities*, Commun. Math. Phys. **189** (1997), 667–698.
- [42] L. Slavín and V. Vasyunin *Sharp results in the integral-form John-Nirenberg inequality*, Trans. Amer. Math. Soc. **363**(2011), 4135–4169.
- [43] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, **293**, 1980.
- [44] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, New Jersey (1970).
- [45] A. Uchiyama, *Weight functions on probability spaces*, Tôhoku Math. Journ. **30** (1978), 463–470.
- [46] N. Th. Varopoulos, *Aspects of probabilistic Littlewood-Paley theory*, J. Funct. Anal. **38**, 25–60.(1980).
- [47] N. Th. Varopoulos, *Hardy-Littlewood theory for semigroups*, J. Funct. Anal. **63**, 240-60 (1985).
- [48] G. Wang, *Sharp inequalities for the conditional square function of a martingale*, Ann. Probab. **19** (1991), 1679–1688.
- [49] J. Wittwer, *A sharp bound for the martingale transform*, Math. Res. Lett. **7** (2000), 1–12.
- [50] A. Zygmund, *Trigonometrical Series*. Annals of Mathematics Studies, No. 63. Cambridge University Press, Cambridge, 1959.

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907, USA  
*E-mail address:* banuelos@math.purdue.edu

DEPARTMENT OF MATHEMATICS, INFORMATICS AND MECHANICS, UNIVERSITY OF WARSAW, BANACHA 2, 02-097 WARSAW, POLAND  
*E-mail address:* ados@mimuw.edu.pl