

ON STRONG WEIGHTED AND RESTRICTED WEAK WEIGHTED ESTIMATES OF THE SQUARE FUNCTION

PAATA IVANISVILI, PAVEL MOZOLYAKO, AND ALEXANDER VOLBERG

ABSTRACT. In this note we give a sharp weighted estimate for square function from $L^2(w)$ to $L^2(w)$, $w \in A_2$. This has been known. We also give sharp weighted estimate *on test functions* for square function from $L^2(w)$ to $L^{2,\infty}(w)$, $w \in A_2$. This has not been known. We also reduce the full estimate for square function from $L^2(w)$ to $L^{2,\infty}(w)$, $w \in A_2$ to a the estimate of the solution of a certain partial differential inequality.

1. INTRODUCTION

In this note we give a sharp weighted estimate for square function from $L^2(w)$ to $L^2(w)$, $w \in A_2$. This has been known. We also give sharp weighted estimate *on test functions* for square function from $L^2(w)$ to $L^{2,\infty}(w)$, $w \in A_2$. This has not been known.

We also reduce the full estimate for square function from $L^2(w)$ to $L^{2,\infty}(w)$, $w \in A_2$ to a the estimate of the solution of a certain partial differential inequality.

2. SHARP WEIGHTED ESTIMATE $S : L^2(w) \rightarrow L^2(w)$. THE LEGO CONSTRUCTION.

We use the approach quite different from the approach in [9], where such sharp estimate was first obtained. Notice that more general sharp weighted estimates for square function in $L^p(w)$ were obtained since then by Lerner [12].

$$(2.1) \quad \frac{1}{|J|} \sum_{I \in D(J)} |\Delta_I w^{-1}|^2 \langle w \rangle_I |I| \leq C_{s,T}([w]_{A_2}) \langle w^{-1} \rangle_J.$$

$$(2.2) \quad \frac{1}{|J|} \sum_{I \in D(J)} |\Delta_I(\varphi w^{-1})|^2 \langle w \rangle_I |I| \leq C_s([w]_{A_2}) \langle \varphi^2 w^{-1} \rangle_J.$$

Remark. We are working with operator $S : L^2(w) \rightarrow L^{2,\infty}(w)$ or $S : L^2(w) \rightarrow L^2(w)$. However, it is more convenient to work with isomorphic

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

PM is supported by the Russian Science Foundation grant 17-11-01064.

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

objects: $S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^{2,\infty}(w)$ or $S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^2(w)$, here $S_{w^{-1}}$ denotes the product $SM_{w^{-1}}$, where $M_{w^{-1}}$ is the operator of multiplication.

2.1. Sharp estimate for $C_{s,T}([w]_{A_2})$. Definition. For a smooth function B of d real variables (x_1, \dots, x_d) we denote by $d^2B(x)$ the second differential form of B , namely,

$$d^2B(x) = (H_B(x)dx, dx)_{\mathbb{R}^d},$$

where vector $dx = (dx_1, \dots, dx_d)$ is an arbitrary vector in \mathbb{R}^d , and $H_B(x)$ is the $d \times d$ matrix of the second derivatives of B (Hessian matrix) at point $x \in \mathbb{R}^d$.

For brevity we write A_2 in this Section, but we mean the dyadic class A_2 .

Theorem 2.1. $C_{s,T} \leq A[w]_{A_2}^2$ and this estimate is sharp.

We introduce the following function of 2 real variables

$$(2.3) \quad \mathbf{B}(u, v) := \mathbf{B}_Q(u, v) := \sup \frac{1}{|J|} \sum_{I \in D(J)} |\Delta_I w^{-1}|^2 \langle w \rangle_I |I|,$$

where supremum is taken over all $w \in A_2$, $[w]_{A_2} \leq Q$, such that

$$\langle w \rangle_J = u, \langle w^{-1} \rangle_J = v.$$

Notice that by scaling argument our function does not depend on J but depends on $Q = [w]_{A_2}$. Notice also that function \mathbf{B}_Q is defined in the domain O_Q , where

$$O_Q := \{(u, v) \in \mathbb{R}^2 : 1 < uv \leq Q\}.$$

Function \mathbf{B}_Q is the Bellman function of our problem. In particular, it is very easy to observe that to prove the estimate in Theorem 2.1 is equivalent to proving $\mathbf{B}_Q(u, v) \leq A Q^2$, and the sharpness in Theorem 2.1 is just the claim that $\sup_{(u,v) \in O_Q} \frac{1}{v} \mathbf{B}_Q(u, v) \geq c > 0$.

Remark. Therefore, Theorem 2.1 can be proved by finding the explicit formula for \mathbf{B}_Q . To do that we obviously need to solve an infinite dimensional optimization problem of finding the (almost) best possible $w \in A_2$ (the reader is recalled that it is a dyadic class) such that $\langle w \rangle_J = u, \langle w^{-1} \rangle_J = v$. This can be done, and we give this formula in the Addendum. But here we adapt a slightly different approach to proving Theorem 2.1.

Proof. Below, A, a are positive absolute constants. Instead of finding precisely \mathbf{B}_Q , we will find another function B_Q such that the following properties are satisfied:

- B_Q is defined in O_{4Q} ,
- $0 \leq B_Q(u, v) \leq A Q^2 v$,
- $B_Q(u, v) - \frac{B_Q(u_+, v_+) + B_Q(u_-, v_-)}{2} \geq a(v_+ - v_-)^2 u$, if $x = (u, v) \in O_Q$, $x_{\pm} = (u_{\pm}, v_{\pm}) \in O_Q$, and $x = \frac{x_+ + x_-}{2}$

For example, it is not difficult to check that \mathbf{B}_{4Q} satisfies all the first and the third properties. We leave it as an exercise to the reader. However, as we said, the second property is not easy, it can be observed when the complicated formula for \mathbf{B} is written down (see Addendum). Here we will write down an explicit (and rather easy) form of *some* B_Q that satisfies all three properties.

Firs let us observe that if the existence of such a B_Q is proved, the inequality in Theorem 2.1 gets proved. In fact, fix $I \in D(J)$, and introduce $x_I = (\langle w \rangle_I, \langle w^{-1} \rangle_I)$, $x_{I_{\pm}} = (\langle w \rangle_{I_{\pm}}, \langle w^{-1} \rangle_{I_{\pm}})$. Of course $\{x_I\}_{I \in D(J)}$ is a martingale. Now we compose this martingale with B_Q (notice that $x_I \in O_Q$, so $B_Q(x_I)$ is well-defined). The resulting object is not a martingale anymore, but it is a super-martingale, moreover, by the third property

$$\langle w \rangle_I |\Delta_I w^{-1}|^2 |I| \leq |I| B_Q(x_I) - |I_+| B_Q(x_{I_+}) - |I_-| B_Q(x_{I_-}).$$

Now the reader knows what happens next: we use the telescopic nature of the sum in the right hand side to observe that the summation in all $I \in D(J)$ cancels all the terms except $|I| B_Q(x_J)$, which, by the second property is at most $A Q^2 \langle w^{-1} \rangle_J |J|$. Hence, we obtained

$$\sum_{I \in D(J)} |\Delta_I w^{-1}|^2 \langle w \rangle_I |I| \leq A Q^2 \langle w^{-1} \rangle_J |J|,$$

We leave as an exercise for the reader to explain, where we used the positivity of B in this reasoning.

The last estimate is precisely inequality (2.1) and Theorem 2.1 gets proved (apart from the sharpness) as soon as any function B_Q as above is proved to exist.

The construction of a certain B_Q with above mentioned three properties is split to two steps.

Step 1: the reduction to non-linear ODE.

First of all we wish to find a smooth $B(u, v)$ in the domain

$$O := O_Q := \{(u, v) > 0 : 1 \leq uv \leq Q\},$$

such that the following quadratic forms inequality holds in O_Q :

$$(2.4) \quad -\frac{1}{2} d^2 B \geq u (dv)^2.$$

We will be searching for homogeneous B : $B(u/t, tv) = tB(u, v)$. Hence

$$B(u, v) = \frac{1}{u} \phi(uv)$$

Then (2.4) becomes

$$\begin{bmatrix} x^2 \phi''(x) - 2x \phi'(x) + 2\phi, & x \phi''(x) \\ x \phi''(x), & \phi''(x) + 2 \end{bmatrix} \leq 0.$$

To have this it is enough to satisfy for all $x \in [1, Q]$

$$(2.5) \quad \begin{aligned} \phi''(x) + 2 &\leq 0, \quad -x\phi'(x) + \phi(x) \leq 0, \\ \phi'' \cdot (-x\phi' + \phi) + x^2\phi'' - 2x\phi' + 2\phi &= 0. \end{aligned}$$

The last equation is just making the determinant of our matrix to vanish. Let us start with this equation and put $g = \phi(x)/x$. Then we know that $-x^2g' = \phi - x\phi' \leq 0$, so g is increasing. Also $xg'' + 2g' = \phi'' \leq -2$, hence $g'' \leq 0$ as g was noticed to be increasing.

In terms of g we have equation

$$x(-g'g'' + g'') - 2(g')^2 = 0.$$

This is a first order non-linear ODE on $h := g'$ of which we know that $h \geq 0, h' \leq 0$:

$$x(-hh' + h') - 2h^2 = 0.$$

Variables separate and we get

$$(2.6) \quad \frac{1-h}{h^2}h' = \frac{2}{x}.$$

The condition we saws that $h' = g''$ is negative and x here is positive, so $h \geq 1$, and the condition $\phi - x\phi' \leq 0$ is the same as $h \geq 0$. Thus any solution $h \geq 1$ of (2.6) gives the desired result.

We want to solve this for $x \in [1, Q]$:

$$-\log h - \frac{1}{h} = 2 \log x + c \quad \Leftrightarrow \quad -\frac{1}{h}e^{-\frac{1}{h}} = -x^2C, \quad C > 0.$$

Notice that Lambert W function (which is multivalued) solves the equation $z = W(z)e^{W(z)}$. Thus we must have $W(-x^2C) = -\frac{1}{h(x)}$. The condition $-1 \leq -\frac{1}{h(x)} \leq 0$ requires that $-1 \leq W \leq 0$, and this gives single-valued function $W_0(y)$ defined on the interval $[-1/e, 0]$ such that $W_0(-1/e) = -1$, $W_0(0) = 0$ and $W_0(y)$ is increasing. So $h(x) = -\frac{1}{W_0(-x^2C)}$. The condition $-1/e \leq -x^2C \leq 0$ for $x \in [1, Q]$ gives the range for constant C i.e., $0 < C \leq \frac{1}{Q^2e}$. Going back to the functions ϕ and B we obtain:

$$\begin{aligned} \varphi(x) &= -x \int_1^x \frac{dt}{W_0(-t^2C)} + x\varphi(1) \\ \text{and thus } B(u, v) &= -v \int_1^{uv} \frac{dt}{W_0(-t^2C)} + v\varphi(1). \end{aligned}$$

Let us also see how bounded is B . Choosing $C = \frac{1}{Q^2e}$ gives minimal B . $B(u, v)$ can be assumed to be 0 if $uv = 1$. Hence $\phi(1) = 0$. Then

$$(2.7) \quad \begin{aligned} B^Q(u, v) &:= B(u, v) = -v \int_1^{uv} \frac{dt}{W_0(-\frac{t^2}{Q^2e})} \leq \\ v \int_1^{uv} \frac{Q^2e}{t^2} &= eQ^2v \left(1 - \frac{1}{uv}\right), \quad 1 \leq uv \leq Q. \end{aligned}$$

Here we used the fact that Lambert function $W_0(x) \leq x$ for $x \in [-1/e, 0]$. Actually one can get better estimates by using the series expansion for W_0 i.e.,

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n, \quad |x| < \frac{1}{e}.$$

Step 2: from infinitesimal inequality on d^2B^Q to global concavity property of B^Q .

The function B^Q defined in (2.7) is not function B_Q with three properties formulated at the beginning of the proof of this theorem. However, let us prove that $B_Q := B^{4Q}$ has all these three properties. The first property is just by definition, and the second property is because we just proved in (2.7) that

$$B_Q(u, v) = B^{4Q}(u, v) \leq 4eQ^2v.$$

To prove the third property let us fix $x = (u, v) \in O_Q$, $x_{\pm} = (u_{\pm}, v_{\pm}) \in O_Q$ such that $x = \frac{x_+ + x_-}{2}$. Introduce two functions defined on $[-1, 1]$:

$$U(t) = \frac{1+t}{2}u_+ + \frac{1-t}{2}u_-, \quad V(t) = \frac{1+t}{2}v_+ + \frac{1-t}{2}v_-.$$

Compose the vector function $X(t) = (U(t), V(t))$ and function $B_Q = B^{4Q}$, namely, put

$$b(t) := B_Q(X(t)).$$

Then $X(\pm 1) = x_{\pm}$, $X(0) = x$. It is important to notice that b is well-defined because

$$\forall t \in [-1, 1] \quad X(t) \in O_{4Q}.$$

The latter is an elementary geometric observation saying that if three points X_{\pm}, X belong to O_Q , and $X = \frac{X_+ + X_-}{2}$, then the whole segment with end points X_{\pm} lie in O_{4Q} . (But not, in general, on $O_{Q'}$ with $Q' < 4Q$.) Now we differentiate twice function b . The chain rule gives us immediately

$$b''(t) = (H_{B_Q}(X(t))(x_+ - x_-), (x_+ - x_-))_{\mathbb{R}^2},$$

where H_{B_Q} denotes as always the Hessian matrix of function B_Q . Therefore, the use of (2.4) gives us

$$-b''(t) \geq 2U(t)(v_+ - v_-)^2.$$

On the other hand,

$$B_Q(x) - \frac{B_Q(x_+) + B_Q(x_-)}{2} = b(0) - \frac{b(1) + b(-1)}{2} = \frac{1}{2} \int_{-1}^1 -b''(t)(1 - |t|) dt$$

Notice that the integrand is always nonnegative by the previous display formula. By the same formula, the integrand is at least $u = U(0)$ for $t \in$

$[0, 1/2]$ because on this interval $U(t) \geq \frac{1}{2}U(0)$ by the obvious geometric reason. Combing that we obtain

$$B_Q(x) - \frac{B_Q(x_+) + B_Q(x_-)}{2} \geq \frac{3}{8}u(v_+ - v_-)^2.$$

We established all three properties for B_Q , and we have already shown that this is enough to prove the inequality in Theorem 2.1. The sharpness is not difficult to see for a weight with one singular point, see [9] for example. \square

2.2. Proving the instance of $T1$ theorem using its Bellman function.

Theorem 2.2. $C_s \leq AQ^2$ and this estimate is sharp.

Let us deduce this result from Theorem 2.1. This the occasion of the so-called weighted $T1$ theorem.

We use the notation h_I for a standard Haar function supported on dyadic interval I , it is given by

$$h_I = \begin{cases} \frac{1}{\sqrt{I}}, & \text{on } I+ \\ -\frac{1}{\sqrt{I}}, & \text{on } I- . \end{cases}$$

It is an orthonormal basis in unweighted L^2 . Now consider the same type of Haar basis but in weighted $L^2(w^{-1})$: functions $h_I^{w^{-1}}$ are orthogonal to constants in $L^2(w^{-1})$, normalized in $L^2(w^{-1})$, assume constant value on each child of I , and are supported on I . For dyadic intervals on the line we get

We need a couple of lemmas.

Lemma 2.3. *The following holds $h_I = \alpha_I^{w^{-1}} h_I^{w^{-1}} + \beta_I^{w^{-1}} \frac{\mathbf{1}_I}{\sqrt{I}}$ with*

$$(2.8) \quad \alpha_I^{w^{-1}} = \frac{\langle w^{-1} \rangle_{I+}^{1/2} \langle w^{-1} \rangle_{I-}^{1/2}}{\langle w^{-1} \rangle_I^{1/2}}, \quad \beta_I^{w^{-1}} = \frac{\langle w^{-1} \rangle_{I+} - \langle w^{-1} \rangle_{I-}}{\langle w^{-1} \rangle_I}$$

Proof. This is a direct calculation, we need to define two constants $\alpha_I^{w^{-1}}, \beta_I^{w^{-1}}$ and we have two conditions: $\|h_I^{w^{-1}}\|_{L^2(w^{-1})} = 1$ and $(h_I^{w^{-1}}, \mathbf{1})_{L^2(w^{-1})} = 0$. \square

The second lemma is just the instance of the chain rule.

Lemma 2.4. *Let $\Phi(x'), B(x'')$ be smooth functions of $x' = (x_1, \dots, x_n, x_0)$, $x'' = (x_{n+1}, \dots, x_m)$. Then we compute the second differential form of the composition function*

$$\mathcal{B}(x_1, \dots, x_n, x_{n+1}, \dots, x_m) = \Phi(x_1, \dots, x_n, B(x_{n+1}, \dots, x_m))$$

by the following formula:

$$d^2 \mathcal{B} = d^2 \Phi + \frac{\partial \Phi}{\partial x_0} d^2 B.$$

Remark. We understand the left hand side as $(H_{\mathcal{B}}(x)dx, dx)_{\mathbb{R}^m}$, where

$$x := (x_1, \dots, x_n, x_{n+1}, \dots, x_m), \quad dx = (dx_1, \dots, dx_n, dx_{n+1}, \dots, dx_m).$$

We understand $d^2\Phi$ in the right hand side as $(H_{\Phi}(x_1, \dots, x_n, B(x''))dy, dy)_{\mathbb{R}^{n+1}}$, where

$$dy = (dx_1, \dots, dx_n, dB), \quad dB = (\nabla B(x''), dx'')_{\mathbb{R}^{m-n}}.$$

Now we are ready to prove Theorem 2.2.

Proof. The sum we want to estimate in (2.2)

$$\frac{1}{|J|} \sum_{I \in D(J)} (\langle \varphi w^{-1} \rangle_{I_+} - \langle \varphi w^{-1} \rangle_{I_-})^2 \langle w \rangle_I |I|$$

is of course

$$\Sigma = \frac{2}{|J|} \sum_{I \in D(J)} (\varphi w^{-1}, h_I)^2 \langle w \rangle_I.$$

We can plug the decomposition of Lemma 2.3 and take into account that for dyadic lattice obviously $\alpha_I^{w^{-1}} \leq 2\langle w^{-1} \rangle_I^{1/2}$. Then we obtain

$$\begin{aligned} \Sigma &\leq \frac{8}{|J|} \sum_{I \in D(J)} (\varphi w^{-1}, h_I^{w^{-1}})^2 \langle w^{-1} \rangle_I \langle w \rangle_I + \\ &\frac{2}{|J|} \sum_{I \in D(J)} \langle \varphi w^{-1} \rangle_I^2 (\beta_I^{w^{-1}})^2 \langle w \rangle_I |I| =: \Sigma_1 + \Sigma_2. \end{aligned}$$

The system $\{h_I^{w^{-1}}\}_{I \in D(J)}$ is orthonormal in $L^2(w^{-1})$, and $\langle w^{-1} \rangle_I \langle w \rangle_I \leq Q$. Hence, immediately we have

$$(2.9) \quad \Sigma_1 \leq 8Q \|f\|_{L^2(w^{-1})}^2.$$

We are left to estimate Σ_2 . To do that let us rewrite Σ_2 :

$$\Sigma_2 = \frac{2}{|J|} \sum_{I \in D(J)} \left(\frac{\langle \varphi w^{-1} \rangle_I}{\langle w^{-1} \rangle_I} \right)^2 \gamma_I |I| = \frac{2}{|J|} \sum_{I \in D(J)} (\langle \varphi w^{-1} \rangle_{I, w^{-1}})^2 \gamma_I |I|,$$

where $\langle \cdot \rangle_{I, w^{-1}}$ means the average with respect to measure $\mu := w^{-1}(x)dx$ and

$$\gamma_I := (\langle w^{-1} \rangle_{I_+} - \langle w^{-1} \rangle_{I_-})^2 \langle w \rangle_I.$$

We are going to prove now that with some absolute constant A

$$(2.10) \quad \frac{1}{|J|} \sum_{I \in D(J)} (\langle \varphi w^{-1} \rangle_{I, w^{-1}})^2 \gamma_I |I| \leq A Q^2 \langle \varphi^2 w^{-1} \rangle_J.$$

This of course finishes the proof of Theorem 2.2.

To prove 2.10 we will construct a special function of 4 real variables $\mathcal{B}(X) = \mathcal{B}_Q(X)$, $X = (F, f, u, v)$, that possesses the following properties

- 1) \mathcal{B}_Q is defined in a non-convex domain O_{4Q} , where $O_Q := \{(F, f, u, v) \in \mathbb{R}_+^4 : f^2 < Fv, 1 < uv < Q\}$,
- 2) $0 \leq \mathcal{B}_Q \leq F$,

- 3) $\mathcal{B}_Q(X) - \frac{\mathcal{B}_Q(X_+) + \mathcal{B}_Q(X_-)}{2} \geq aQ^{-2} \frac{f^2}{v^2} u(v_+ - v_-)^2$, where $X = (F, f, u, v)$, $X_{\pm} = (F_{\pm}, f_{\pm}, u_{\pm}, v_{\pm})$ belong to O_Q , and $X = \frac{X_+ + X_-}{2}$.

As soon as such a function constructed (2.10) and Theorem 2.2 follow immediately. In fact we repeat our telescopic consideration. We set the vector martingale $X_I := (F_I, f_I, u_I, v_I)$, where

$$F_I = \langle \varphi w^{-1} \rangle_I, f_I = \langle \varphi w^{-1} \rangle_I, u_I = \langle w \rangle_I, v_I = \langle w^{-1} \rangle_I.$$

It is obvious that vector martingale $\{X_I\}_{I \in D(J)}$ is always inside O_Q , and so the superposition of this martingale and \mathcal{B}_Q is well defined: $\mathcal{B}_Q(X_I)$. Then, the property 3) claims that $\{\mathcal{B}_Q(X_I)\}_{I \in D(J)}$ is a super-martingale, and moreover

$$|I| \frac{f_I^2}{v_I^2} u_I (v_{I_+} - v_{I_-})^2 \leq AQ^2 (|I| \mathcal{B}(X_I) - |I_+| \mathcal{B}(X_{I_+}) - |I_-| \mathcal{B}(X_{I_-}))$$

We use the telescopic nature of the term in the right hand side, and summing these terms for all $I \in D(J)$, we then notice that all of them will cancel each other, except $AQ^2 |J| \mathcal{B}(X_J)$, which is bounded by $AQ^2 |J| F_J = AQ^2 |J| \langle \varphi^2 w^{-1} \rangle_J$. We proved (2.10) provided that the existence of \mathcal{B}_Q is validated.

Now we will write the explicit formula for \mathcal{B}_Q . Exactly as in Theorem 2.1 we first construct, by an explicit formula, an auxiliary function \mathcal{B}^Q . Here it is

$$(2.11) \quad \mathcal{B}^Q(F, f, u, v) := F - \frac{f^2}{v + aQ^{-2} \mathcal{B}^Q(u, v)},$$

where B^Q was defined in (2.7). It is clear that it satisfies property 2). It “almost” satisfies property 1), but it is defined only in O_Q , not in a larger domain O_{4Q} . As to the property 3) it does satisfy its infinitesimal version: at point $X = (F, f, u, v) \in O_Q$

$$(2.12) \quad -d^2 \mathcal{B}^Q \geq aQ^{-2} \frac{f^2}{v} u (dv)^2.$$

Let us prove this. This follows from Lemma 2.4. In fact, consider $\Phi(x_1, x_2, x_3, x_0) := x_1 - \frac{x_2^2}{x_3 + x_0}$. By a direct simple calculation one can see that it is concave in \mathbb{R}_+^4 , so $d^2 \Phi \leq 0$. Now we see that

$$\mathcal{B}^Q(F, f, u, v) = \Phi(F, f, v, aQ^{-2} \mathcal{B}^Q(u, v)),$$

and by Lemma 2.4

$$d^2 \mathcal{B}^Q = d^2 \Phi + aQ^{-2} \frac{\partial \Phi}{\partial x_0} \cdot (d^2 \mathcal{B}^Q) \leq \frac{a}{Q^2} \frac{f^2}{(v + aQ^{-2} \mathcal{B}^Q(u, v))^2} (d^2 \mathcal{B}^Q).$$

But in Theorem 2.1 we proved that $B^Q \leq AQ^2 v$, hence, choosing small absolute constant a we guarantee that $v + aQ^{-2} \mathcal{B}^Q(u, v) \leq 2v$. We also

proved in Theorem 2.1 that $-d^2B^Q \geq u(dv)^2$. Combining these facts with the last display inequality we obtain

$$-d^2\mathcal{B}^Q \geq \frac{a}{4Q^2} \frac{f^2}{v^2} u(dv)^2,$$

which is precisely (2.12).

We need function with property 3) and defined in the domain O_{4Q} . So let us put $\mathcal{B}_Q := \mathcal{B}^{4Q}$. Exactly as before as in Step 2 of Theorem 2.1 we can prove now that not only infinitesimal special concavity holds in the form $-d^2\mathcal{B}_Q \geq \frac{a}{64Q^2} \frac{f^2}{v^2} u(dv)^2$, but also we have with some small positive a_0

$$(2.13) \quad \mathcal{B}_Q(X) - \frac{\mathcal{B}_Q(X_+) + \mathcal{B}_Q(X_-)}{2} \geq a_0Q^{-2} \frac{f^2}{v^2} u(v_+ - v_-)^2,$$

for all triple of points $(X, X_+, X_-) \in O_Q^3$ such that $X = (F, f, u, v)$, $X_\pm = (F_\pm, f_\pm, u_\pm, v_\pm)$, and $X = \frac{X_+ + X_-}{2}$. \square

Remark. The constant in the right hand side of the main inequality (2.10) is AQ^2 . It is important to notice that the constant in the right hand side of (2.10) depends only on the constant in Theorem 2.1.

In fact, these two constants are equal up to an absolute constant. So if, for example, the constant in Theorem 2.1 is denoted by K , then the constant in (2.10) becomes AK , and the constant in Theorem 2.2 would also become AK (with another absolute constant A). This principle of passing from uniform testing conditions (Theorem 2.1) to full boundedness of the operator is quite general, and, as we said already, is often called *T1* theorem.

Here this principle was proved by showing that *the formula* that proves Theorem 2.2 can be obtained by totally formal manipulations from *the formula* that prove testing condition (Theorem 2.1). See the next remark.

Remark. The reader should pay attention to a following curious formula.

$$(2.14) \quad \mathcal{B}_Q(F, f, u, v) = F - \frac{f^2}{v + aQ^{-2}B^Q(u, v)},$$

This is the function \mathcal{B}_Q that proves Theorem 2.2, and B^Q is the function that proved Theorem 2.1.

So we see another instance of transference by use of Bellman function. By formula (2.14) we transfer the claim of Theorem 2.1 to Theorem 2.2. A Bellman function of Theorem 2.1 was used as “a lego piece” to construct a Bellman function for Theorem 2.2. In this instance this “lego construction” proved for us the *T1* theorem for the weighted square function operator.

3. A SMALL SHARPENING OF THE *T1* THEOREM FOR DYADIC SQUARE FUNCTION

Recall that

$$\gamma_I = (\langle w^{-1} \rangle_{I_+} - (\langle w^{-1} \rangle_{I_-})^2 \langle w \rangle_I).$$

A natural question arises: Let $[w]_{A_2} = Q \gg 1$ and at the same time

$$(3.1) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \gamma_I |I| \leq q^2 \langle w^{-1} \rangle_J, \quad \forall J \in \mathcal{D},$$

where $q \ll Q$, does this mean that the norm $\|S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^2(w)\|$ is bounded by Cq ? That would be quite expected because this statement reminds the statement of the so-called $T1$ theorems. In fact, (3.1) is precisely the testing condition and can be rewritten as

$$(3.2) \quad \|S_{w^{-1}} \chi_J\|_w^2 \leq q^2 \|\chi_J\|_{w^{-1}}^2, \quad \forall J \in \mathcal{D}.$$

We are quite sure that this is wrong. If so the next natural question is the following: if, however, $q \ll Q$, is it true that one can give a better estimate than the one we proved in the previous section:

$$\|S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^2(w)\| \lesssim Q?$$

The answer to this question is positive. We have the following result.

Theorem 3.1. *Let $w \in A_2$ (dyadic as always), let $[w]_{A_2} = Q$, and let (3.1) holds with $q \ll Q$, then we have the following improved estimate on $\|S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^2(w)\|$:*

$$\|S_{w^{-1}} : L^2(w^{-1}) \rightarrow L^2(w)\| \leq C(Q^{1/2} + q).$$

Proof. Let us consider the following aggregate, which is quite akin to the function from (2.11) (constant a is again a small positive constant):

$$(3.3) \quad \mathcal{B}(F, f, A, v) := F - \frac{f^2}{v + aq^{-2}A},$$

Given $I \in \mathcal{D}$, we put

$$A_I := \frac{1}{|I|} \sum_{\ell \in \mathcal{D}(I)} \gamma_\ell |\ell|.$$

As before we set $Y_I := (F_I, f_I, A_I, v_I)$, where is defined above, and

$$F_I = \langle \varphi w^{-1} \rangle_I, \quad f_I = \langle \varphi w^{-1} \rangle_I, \quad v_I = \langle w^{-1} \rangle_I.$$

We also denote $y_I := (F_I, f_I, \frac{1}{2}(A_{I_+} + A_{I_-}), v_I)$. We can estimate

$$\begin{aligned} & |I| \mathcal{B}(Y_I) - |I_+| \mathcal{B}(Y_{I_+}) - |I_-| \mathcal{B}(Y_{I_-}) \\ &= |I| \left((\mathcal{B}(Y_I) - \mathcal{B}(y_I)) + (\mathcal{B}(y_I) - \frac{1}{2} \mathcal{B}(Y_{I_+}) - \frac{1}{2} \mathcal{B}(Y_{I_-})) \right) \\ &\geq \frac{|I|}{4} \frac{f_I^2}{v_I^2} (A_I - \frac{1}{2}(A_{I_+} + A_{I_-})) \frac{a}{q^2} = \frac{a}{4q^2} \frac{f_I^2}{v_I^2} \gamma_I |I| \\ &= \frac{a}{4q^2} \frac{\langle \varphi w^{-1} \rangle_I^2}{\langle w^{-1} \rangle_I^2} (\langle w^{-1} \rangle_{I_+} - \langle w^{-1} \rangle_{I_-})^2 \langle w \rangle_I |I| \end{aligned}$$

Summing up over all $I \in \mathcal{D}(J)$ and using the telescopic nature of terms, we get

$$(3.4) \quad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} (\langle \varphi w^{-1} \rangle_{I, w^{-1}})^2 \gamma_I |I| \leq Aq^2 \langle \varphi^2 w^{-1} \rangle_J,$$

which is like (2.10), but with q replacing Q . In the inequality above we used the concavity of function $\mathcal{B}(F, f, A, v) = F - \frac{f^2}{v+aq^{-2}A}$, which, in particular, gives us that

$$\mathcal{B}(y_I) - \frac{1}{2}\mathcal{B}(Y_{I_+}) - \frac{1}{2}\mathcal{B}(Y_{I_-}) \geq 0,$$

and we use the estimate from below for $\frac{\partial}{\partial A}\mathcal{B}(F, f, A, v)$:

$$\min_{\frac{1}{2}(A_{I_+} + A_{I_-}) \leq A \leq A_I} \frac{\partial}{\partial A}\mathcal{B}(F, f, A, v) \geq \frac{f^2}{(v + aq^{-2}A_I)^2} \geq \frac{f^2}{4v^2}.$$

The last inequality is clear if we use (3.1).

Now combining estimate (2.9) of the previous section and (3.4) we get the claim of Theorem 3.1. \square

4. WEAK WEIGHTED ESTIMATES FOR THE SQUARE FUNCTION

Our measure space will be (X, \mathfrak{A}, dx) , where σ -algebra \mathfrak{A} is generated by a standard dyadic filtration $\mathcal{D} = \cup_k \mathcal{D}_k$ on \mathbb{R} . We considered the weak weighted estimate for the martingale transform. In [20] the end-point exponent was $p = 1$, and critical weights belong to A_1 class.

Now we are going to consider weighted weak estimates for the dyadic square function. The end-point exponent is now $p = 2$, and critical weights belong to A_2 class. See the estimates for subcritical and supercritical exponents in [11] and [6].

The weak estimate of the square function in $L^2(wdx)$, $w \in A_1$ is well known, see, e. g. [30]. The fact that the end-point exponent is now at $p = 2$ is explainable by the bi-linear nature of square function transform.

Recall that the symbol $\text{ch}(J)$ denotes the dyadic children of J . Recall that the martingale difference operator Δ_J was defined as follows:

$$\Delta_J f := \sum_{I \in \text{ch}(J)} \mathbf{1}_I (\langle f \rangle_I - \langle f \rangle_J).$$

For our case of dyadic lattice on the line we have that $|\Delta_J f|$ is constant on J , and

$$\Delta_J f = \frac{1}{2} [(\langle f \rangle_{J_+} - \langle f \rangle_{J_-}) \mathbf{1}_{J_+} + (\langle f \rangle_{J_-} - \langle f \rangle_{J_+}) \mathbf{1}_{J_-}].$$

The square function operator is

$$(Sf)^2(x) = \sum_{J \in \mathcal{D}} |\Delta_J f|^2 \mathbf{1}_J(x).$$

In this Section we work only with the dyadic A_2 classes of weights, but we skip the word dyadic, because we consider here only dyadic operators. We consider a positive function $w(x)$, and as before we call it A_2 weight if

$$(4.1) \quad Q := [w]_{A_2} := \sup_{J \in \mathcal{D}} \langle w \rangle_J \langle w^{-1} \rangle_J < \infty.$$

We are going to consider restricted weak estimate, when the operator is applied to $f = \mathbf{1}_E$, but only for set E being itself a dyadic interval.

$$(4.2) \quad \frac{1}{|I|} w \left\{ x \in I : \sum_{J \in \mathcal{D}(I)} |\Delta_J w^{-1}|^2 \mathbf{1}_J(x) > \lambda \right\} \leq C_T([w]_{A_2}) \frac{\langle w^{-1} \rangle_I}{\lambda}.$$

We give some information on full weak estimate for the square function operator:

$$(4.3) \quad \frac{1}{|I|} w \left\{ x \in I : \sum_{J \in \mathcal{D}(I)} |\Delta_J(\varphi w^{-1})|^2 \mathbf{1}_J(x) > \lambda \right\} \leq C([w]_{A_2}) \frac{\langle \varphi^2 w^{-1} \rangle_I}{\lambda}.$$

Here φ runs over all functions such that $\text{supp } \varphi \subset I$ and $\varphi \in L^2(I, w dx)$, $w \in A_2$.

We wish to understand the sharp order of magnitude of the weak estimates constants $C_T([w]_{A_2})$, $C([w]_{A_2})$ from (4.2), (4.2) correspondingly ($[w]_{A_2}$ is supposed to be large). We wish to compare them to the similar estimates of strong type: (2.1), (2.2) from Section 2.

Constant $C([w]_{A_2})$ gives the dependence on $[w]_{A_2}$ of the norm of the sub linear operator $S\varphi := (\sum_{J \in \mathcal{D}} |\Delta_J \varphi|^2 \mathbf{1}_J(x))^{1/2}$ from $L^2(w)$ to $L^{2,\infty}(w)$. Constant $C_T([w]_{A_2})$ gives the dependence on $[w]_{A_2}$ of the same weak norm but measured only on test functions φ that are just characteristic functions of dyadic intervals: $\varphi = \mathbf{1}_I, I \in \mathcal{D}$.

Let M_a be an operator of multiplication by function a . Notice a convenient change of variable $\varphi \rightarrow \varphi W^{-1}$, which reduces the problem of estimating the square function operator $S : L^2(w) \rightarrow L^{2,\infty}(w)$ to the estimate of $S_{w^{-1}} := SM_{w^{-1}}$ from $L^2(w^{-1})$ to $L^{2,\infty}(w)$.

Of course, there are several obvious estimates: 1) weak constants are smaller than strong constants:

$$C_T([w]_{A_2}) \leq C_{s,T}([w]_{A_2}), \quad C([w]_{A_2}) \leq C_s([w]_{A_2})$$

just by Chebyshev inequality; also 2) test function estimates are trivially at least as good as the full estimates:

$$C_T([w]_{A_2}) \leq C([w]_{A_2}), \quad C_{s,T}([w]_{A_2}) \leq C_s([w]_{A_2}).$$

It is quite well known (although far from being trivial) that there is at least one converse inequality:

$$(4.4) \quad C_s([w]_{A_2}) \leq A(d)C_{s,T}([w]_{A_2}).$$

This is an instance of the celebrated $T1$ theorem, on this occasion applied to square function transform in the weighted situation. We have proved this statement in Theorem 2.2 above, see Remark 2.2.

Remark. Weak type estimates with sharp dependence on $[w]_{A_2}$ are rather difficult, in part because their sharp constants are much smaller than the sharp constants in the corresponding strong type estimates.

Remark. This is in a big contrast with the usual singular integrals T of Calderón–Zygmund type. We know that for $w \in A_2$ the strong norm $\|T : L^2(w) \rightarrow L^2(w)\|$ is exactly equivalent to $\|T : L^2(w) \rightarrow L^{2,\infty}(w)\| + \|T^* : L^2(w^{-1}) \rightarrow L^{2,\infty}(w^{-1})\|$, see [21], [10] for example.

In view of Remark 4 it was important to review the strong sharp weighted estimate $S : L^2(w) \rightarrow L^2(w)$ in Section 2.

Remark. The reader can see that (4.2) estimate is a particular case of (4.3) estimate for a special choice of test function $f = \mathbf{1}_I$. The same remark holds for (2.1) and (2.2). We already noted above that test function estimate (2.1) (plus its symmetric counterpart for w exchanged by w^{-1}) imply the strong estimate (2.2). This, as we already mentioned, is the essence of weighted $T1$ theorem (testing condition theorem in the terminology of E. Sawyer).

Remark. Unfortunately there is no $T1$ principle in general for weak type estimates. See the papers [4], [1], [5].

4.1. Sharp constant in weak testing estimate. An “isoperimetric” problem.

Theorem 4.1. $C_{w,T} \leq AQ$ and this estimate is sharp.

We will prove the estimate, the sharpness is well known just for one-point-singularity weights.

As always in this Section A_2 means dyadic A_2 . We introduce the following function of 3 real variables

$$(4.5) \quad B_Q(u, v, \lambda) := \sup \frac{1}{|J|} w \left\{ x \in J : \sum_{I \in D(J)} |\Delta_I w^{-1}|^2 \chi_I(x) > \lambda \right\},$$

where the supremum is taken over all $w \in A_2, [w]_{A_2} \leq Q$, such that

$$\langle w \rangle_J = u, \langle w^{-1} \rangle_J = v.$$

Notice that by scaling argument our function does not depend on J but depends on $Q = [w]_{A_2}$. For brevity we can skip Q : $B := B_Q$.

Remark. Ideally we want to find the formula for this function. Notice that this is similar to solving a problem of “isoperimetric” type, where the solution of certain non-linear PDE is a common tool, see e. g. [2].

4.1.1. *Properties of B and the main inequality.* Notice several properties of B :

- B is defined in $\Omega := \{(u, v, \lambda) : 1 \leq uv \leq Q, u > 0, v > 0, 0 \leq \lambda < \infty\}$.
- If $P = (u, v, \lambda), P_+ = (u_+, v_+, \lambda_+), P_- = (u_-, v_-, \lambda_-)$ belong to Ω , and $u = \frac{1}{2}(u_+ + u_-), v = \frac{1}{2}(v_+ + v_-), \lambda = \min(\lambda_+, \lambda_-)$, then **the main inequality** holds with constant $c = 1$:

$$B(u, v, \lambda + c(v_+ - v_-)^2) - \frac{B(P_+) + B(P_-)}{2} \geq 0.$$

- B is decreasing in λ .
- Homogeneity: $B(ut, v/t, \lambda/t^2) = tB(u, v, \lambda), t > 0$.
- Obstacle condition: for all points (u, v, λ) such that $10 \leq uv \leq Q, \lambda \geq 0$, if $\lambda \leq \delta v^2$ for a positive absolute constant δ , one has $B(u, v, \lambda) = u$.
- The boundary condition $B(u, v, \lambda) = 0$ if $uv = 1$.

All these properties are very simple consequences of the definition of B . However, let us explain a bit the second and the fifth bullet. The second bullet is the consequence of the scale invariance of B . We consider data P_+ and find weight w_+ that almost supremizes $B(P_+)$. By definition of B , we have it on J . But by scale invariance we can think that w_+ lives on J_+ . Then we consider data P_- and find weight w_- that almost supremizes $B(P_-)$. Again we are supposed to have it on J . But by scale invariance we can think that w_- lives on J_- . The next step is to consider the contatination of w_+ and w_- :

$$w_c := \begin{cases} w_+, & \text{on } J_+ \\ w_-, & \text{on } J_- . \end{cases}$$

Clearly this new weight is a competitor for giving the supremum for date P on J . But it is only a competitor, the real supremum in (4.5) is bigger. This implies the second bullet above (*the main inequality*).

Now let us explain the fifth bullet above, we call it *the obstacle condition*. Let us consider a special weight w_s in J : it is one constant on J_- and just another constant on J_+ . Moreover, we wish to have $\langle w_s^{-1} \rangle_{J_+} = 4 \langle w_s^{-1} \rangle_{J_-}$. Notice that then $b \langle w_s^{-1} \rangle_J \leq |\Delta_J w_s^{-1}|$ with some positive absolute constant b .

Now it is obvious that if $\lambda \leq \delta^2 \langle w_s^{-1} \rangle_J^2$ then $\{x \in J : S_{w_s^{-1}}^2(\mathbf{1}_J) \geq \lambda\} = J$ and so $\frac{1}{|J|} w_s \{x \in J : S_{w_s^{-1}}^2(\mathbf{1}_J) \geq \lambda\} = \langle w_s \rangle_J$. Notice now that w_s is just one admissible weight, and that we have to take supremum over all such admissible weights.

We get the fifth bullet above (=the obstacle condition): $B(u, v, \lambda) = u$ for those points (u, v, λ) in the domain of definition of B , where the corresponding w_s with $\langle w_s \rangle = u, \langle w_s^{-1} \rangle_J = v$ exists. It is obvious that for all sufficiently large Q and for any pair (u, v) such that $10 \leq uv \leq Q$ one can construct a just “two-valued” w_s as above with $[w_s]_{A_2} \leq Q$ (we recall that we deal only with dyadic A_2 weights).

Notice that the main inequality above transforms into a **partial differential inequality** if considered infinitesimally (and if we tacitly assume that B is smooth):

$$(4.6) \quad -\frac{1}{2}d_{u,v}^2 B + c \frac{\partial B}{\partial \lambda} (dv)^2 \geq 0.$$

We get it with $c = 1$ for the function B defined above (if B happens to be smooth).

We are not going to find B defined in (4.5), but instead we will construct smooth \mathcal{B} that satisfies all the properties above (and of course (4.6)) except for the boundary condition (the last bullet above). It will satisfy even a slightly stronger properties, for example, the obstacle condition (the fifth bullet) will be satisfied with 1 instead of 10:

$$(4.7) \quad \begin{aligned} &\forall (u, v, \lambda) \text{ such that } 1 \leq uv \leq Q, \lambda \geq 0, \\ &\text{if } \lambda \leq \delta v^2 \text{ for some } \delta > 0, \text{ then } B(u, v, \lambda) = u. \end{aligned}$$

Here a will be some positive absolute constant (it will not depend on Q).

Using our usual telescopic sums consideration it will be very easy to prove the following

Theorem 4.2. *Suppose we have a smooth function \mathcal{B} satisfying all the conditions above except the boundary condition, but satisfying the obstacle condition in the form (4.7). We also allow c to be a small positive constant (say, $c = \frac{1}{8}$). And suppose it also satisfies*

$$(4.8) \quad \mathcal{B}(u, v, \lambda) \leq A Q \frac{v}{\lambda}.$$

Then the constant $C_{w,T}$ in (4.2) is at most AQ .

Proof. It is a stopping time reasoning. It is enough to think that w is constant on some very small dyadic intervals and to prove the estimate on $C_{w,T}$ uniformly. Then we start with any such $w, [w]_{A_2} \leq Q$, and we use the main inequality with $u = \langle w \rangle_J, v = \langle w^{-1} \rangle_J, u_{\pm} = \langle w \rangle_{J_{\pm}}, v_{\pm} = \langle w^{-1} \rangle_{J_{\pm}},$

$$\lambda - c(\langle w^{-1} \rangle_{J_+} - \langle w^{-1} \rangle_{J_-})^2 =: \lambda_{J_{\pm}}$$

to obtain

$$(4.9) \quad \begin{aligned} &|J_+| \mathcal{B}(\langle w \rangle_{J_+}, \langle w^{-1} \rangle_{J_+}, \lambda_{J_+}) + |J_-| \mathcal{B}(\langle w \rangle_{J_-}, \langle w^{-1} \rangle_{J_-}, \lambda_{J_-}) \leq \\ &\mathcal{B}(\langle w \rangle_J, \langle w^{-1} \rangle_J, \lambda) |J| \leq \frac{A Q \langle w^{-1} \rangle_J}{\lambda} |J| \end{aligned}$$

We continue to use the main inequality (because J_{\pm} are not at all different from J) and finally after large but finite number of steps, on certain collection \mathcal{I} of small intervals $I = J_{\pm\pm\dots\pm}$ we come to the situation that

$$(4.10) \quad \lambda_{J_{\pm\pm\dots\pm}} < c(\langle w^{-1} \rangle_{J_{\pm\pm\dots+}} - \langle w^{-1} \rangle_{J_{\pm\pm\dots-}})^2.$$

Collection \mathcal{I} may be empty of course, but we know that $I \in \mathcal{I}$ if on I the following holds for $x \in I$:

$$c \sum_{L \in D(J), I \subset L} |\Delta_I w^{-1}|^2 \mathbf{1}_L(x) > \lambda.$$

Let us combine (4.10) with an obvious inequality

$$c(\langle w^{-1} \rangle_{J_{\pm\pm\dots+}} - \langle w^{-1} \rangle_{J_{\pm\pm\dots-}})^2 \leq \delta \langle w^{-1} \rangle_{J_{\pm\pm\dots}}^2.$$

At this moment we use the property 5 of B called obstacle condition. On intervals $I \in \mathcal{I}$ the obstacle condition will provide us with $\mathcal{B}(\langle w \rangle_I, \langle w^{-1} \rangle_I, \lambda_I) = \langle w \rangle_I$. So on a certain large finite step N we get from the iteration of (4.9) N times the following estimate

$$\sum_{I \in D_N(J): I \in \mathcal{I}} |I| \langle w \rangle_I \leq \frac{A Q \langle w^{-1} \rangle_J}{\lambda} |J|.$$

Therefore, we proved

$$\frac{1}{|J|} w \{x \in J : \sum_{I \in D(J)} |\Delta_I w^{-1}|^2 \chi_I(x) > \lambda\} \leq A Q \frac{\langle w^{-1} \rangle_J}{\lambda},$$

which is (4.2). □

4.1.2. *Formula for the function \mathcal{B} . Monge–Ampère equation with a drift.* Here is the formula for \mathcal{B} that satisfies all the properties in 4.1.1 (except for the last one, the boundary condition):

$$(4.11) \quad \mathcal{B}(u, v, \lambda) = \frac{1}{\sqrt{\lambda}} \Theta(u\sqrt{\lambda}, \frac{v}{\sqrt{\lambda}}),$$

where $\Theta(\gamma, \tau) := \min \left(\gamma, Q e^{-\tau^2/2} \int_0^\tau e^{s^2/2} ds \right).$

Notice that the fact that \mathcal{B} has the form $\mathcal{B}(u, v, \lambda) = \frac{1}{\sqrt{\lambda}} \Theta(u\sqrt{\lambda}, \frac{v}{\sqrt{\lambda}})$ is trivial, this follows from property 4 called homogeneity.

Notice also that function Θ is given in the domain enclosed by two hyperbolas

$$H := \{(\gamma, \tau) > 0 : 1 \leq \gamma\tau \leq Q\}.$$

All properties listed at the beginning of Section 4.1.1 (except for the sixth bullet, which is boundary condition, but we do not use it anywhere) follow by direct computation. In the next section we explain how to get this formula.

4.1.3. *Explanation of how to find such a function Θ .* The main inequality (with $c = \frac{1}{8}$) in terms of Θ becomes a “drift concavity condition”:

$$(4.12) \quad \frac{1}{\sqrt{1 + \frac{(\Delta\tau)^2}{8}}} \Theta \left(\sqrt{1 + \frac{(\Delta\tau)^2}{8}} \frac{\gamma_- + \gamma_+}{2}, \frac{1}{\sqrt{1 + \frac{(\Delta\tau)^2}{8}}} \frac{\tau_- + \tau_+}{2} \right) \geq \frac{\Theta(\gamma_-, \tau_-) + \Theta(\gamma_+, \tau_+)}{2},$$

where $(\gamma_-, \tau_-), (\gamma_+, \tau_+) \in H$, $0 < \tau_- < \tau_+$, $\Delta\tau := \tau_+ - \tau_-$.

Assuming that Θ is smooth (we will find a smooth function), the infinitesimal version appears, it is a sort of Monge–Ampère relationship with a drift. Namely, the following matrix relationship must hold

$$(4.13) \quad \begin{bmatrix} \Theta_{\gamma\gamma} & \Theta_{\gamma\tau} \\ \Theta_{\gamma\tau} & \Theta_{\tau\tau} + \Theta + \tau\Theta_\tau - \gamma\Theta_\gamma \end{bmatrix} \leq 0.$$

A direct calculation shows that this property is equivalent to the following one. On any curve $\gamma = \phi(\tau)$ lying in the domain H

$$(4.14) \quad \begin{aligned} &\text{and such that } \phi'' + \tau\phi' + \phi = 0 \\ &\text{we have } (\Theta(\phi(\tau), \tau))'' + \tau(\Theta(\phi(\tau), \tau))' + \Theta(\phi(\tau), \tau) \leq 0. \end{aligned}$$

This hints at a possibility to have a change of variables $(\gamma, \tau) \rightarrow (\Gamma, T)$ such that condition (4.13) transforms to a simple concavity. To some extent this is what happens. Namely, notice the following simple

Lemma 4.3. *Consider the following change of variable: $T = \int_0^\tau e^{s^2/2} ds$. Then $\phi''(\tau) + \tau\phi'(\tau) + \phi(\tau) \leq 0$ if and only if $(e^{\tau^2/2}\phi(\tau))_{TT} \leq 0$ and $\phi''(\tau) + \tau\phi'(\tau) + \phi(\tau) = 0$ if and only if $(e^{\tau^2/2}\phi(\tau))_{TT} = 0$.*

Proof. The proof is a direct differentiation. \square

This Lemma hints that the right change of variable should look like

$$(4.15) \quad \begin{cases} \Gamma := \gamma e^{\tau^2/2}, \\ T = \int_0^\tau e^{s^2/2} ds, (\gamma, \tau) \in \mathbb{R}_+^1 \times \mathbb{R}_+^1 \end{cases}$$

then in the new coordinates the family of curves $\gamma = \phi(\tau)$ such that $\phi'' + \tau\phi' + \phi = 0$ becomes a family of all straight lines $\Gamma = CT + D$. (Notice that both families depend on two arbitrary constants.)

Denote

$$O := \{(\Gamma, T) : (\gamma, \tau) \in G\}.$$

The condition $(\Theta(\phi(\tau), \tau))'' + \tau(\Theta(\phi(\tau), \tau))' + \Theta(\phi(\tau), \tau) \leq 0$ on any of these curves becomes

$$(4.16) \quad \left(e^{\tau^2/2} \Theta(\phi(\tau), \tau) \right)_{TT} \leq 0, \quad \Gamma = CT + D, \quad (\gamma, \tau) \in G.$$

which is the concavity of $e^{\tau^2/2}\Theta(\gamma, \tau)$ in a new coordinate T along the line $\Gamma = CT + D$. Let us rewrite two functions in the new coordinates:

$$\Phi(\Gamma, T) := \Theta(\gamma, \tau), \quad U(T) := e^{\tau^2/2}.$$

Then (4.16) transforms into

$$(4.17) \quad \forall C, D \in \mathbb{R}, (U(T)\Phi(CT + D, T))_{TT} \leq 0, (\Gamma, T) \in O.$$

This is just a concavity of $U(T)\Phi(\Gamma, T)$ on O of course. Notice that neither H nor O are convex, so we should understand (4.17) as a local concavity in O : just the negativity of its second differential form

$$d_{\Gamma, T}^2(U(T)\Phi(\Gamma, T)) \leq 0, (\Gamma, T) \in O.$$

So we reduce the question to finding a concave function in new coordinates. Now we choose a simplest possible concave function:

$$U(T)\Phi(\Gamma, T) := \min(\Gamma, KT),$$

where the constant $K = K(Q)$ will be chosen momentarily.

If we write down now $\Theta(\gamma, \tau) = \Phi(\Gamma, T)$ in the old coordinates, we get exactly function Θ from (4.11) (we need to define constant K yet), namely,

$$(4.18) \quad \Theta(\gamma, \tau) := \min\left(\gamma, K e^{-\tau^2/2} \int_0^\tau e^{s^2/2} ds\right).$$

Recall that now we can consider

$$(4.19) \quad B(u, v, \lambda) = \frac{1}{\sqrt{\lambda}} \Theta(u\sqrt{\lambda}, \frac{v}{\sqrt{\lambda}})$$

and we are going to apply Theorem 4.2 to it. But we need to choose K to satisfy all the conditions (except the last one) at the beginning of Section 4.1.1.

First of all it is now very easy to understand why the form of the domain $H = \{1 \leq \gamma\tau \leq Q\}$ plays the role. In fact, by choosing

$$K = AQ$$

with some absolute constant A , we guarantee that in this domain our function Θ satisfies the obstacle condition

$$(4.20) \quad \Theta(\gamma, \tau) = \gamma \text{ as soon as } \tau \geq a_0 > 0,$$

where a_0 is an absolute positive constant. In fact, for all sufficiently small τ , $e^{-\tau^2/2} \int_0^\tau e^{s^2/2} ds \asymp \tau$, and therefore, for all sufficiently small τ (smaller than a certain absolute constant)

$$\Theta(\gamma, \tau) := \min(\gamma, K\tau).$$

The fifth condition at the beginning of Section 4.1.1 (the obstacle condition) requires then that $\min(\gamma, K\tau) = \gamma$ if $\tau \geq a_0 > 0$. But on the upper hyperbola then $\gamma = Q/a_0$ for $\tau = a_0$. We see that the smallest possible K we can choose to satisfy the obstacle condition is $K \asymp Q$.

Secondly, function Θ satisfies the infinitesimal condition (4.13) by construction. But we need to check that the main inequality (4.12) is satisfied as well.

This can be done by the following lemma.

Lemma 4.4. *Inequality (4.12) for function Θ built above holds if and only if the following inequality is satisfied for $\phi(\tau) := e^{-\tau^2/2} \int_0^\tau e^{s^2/2} ds$:*

$$(4.21) \quad \frac{1}{\sqrt{1 + \frac{(\Delta\tau)^2}{100}}} \phi\left(\frac{\tau_1 + \tau_2}{2\sqrt{1 + \frac{(\Delta\tau)^2}{100}}}\right) \geq \frac{\phi(\tau_1) + \phi(\tau_2)}{2}, \quad \forall 0 < \tau_1 \leq \tau_2 \leq \tau_0,$$

with some absolute positive small constant τ_0 .

Proof. Lemma is easy, because we can immediately see that the main inequality (4.12) commutes with the operation of minimum. \square

We are left to prove (4.21). We leave this as exercise to the reader. It is a direct if rather involved calculation.

4.2. A Partial Differential Inequality for the sharp weak estimate of weighted square function. In this Section we are going to reduce the harmonic analysis problem of sharp estimate of weak norm of weighted square function to a certain Partial Differential Inequality (PDI) with an obstacle condition in a rather simple subdomain of \mathbb{R}^3 .

We introduce the following function of 5 real variables (compare with (4.5)):

$$(4.22) \quad \mathbf{B}_Q(F, f, u, v, \lambda) := \sup \frac{1}{|J|} w \left\{ x \in J : \sum_{I \in D(J)} |\Delta_I(\phi w^{-1})|^2 \chi_I(x) > \lambda \right\},$$

where supremum is taken over all $w \in A_2$, $[w]_{A_2} \leq Q$, such that

$$\langle w \rangle_J = u, \langle w^{-1} \rangle_J = v,$$

and over all functions ϕ such that $\langle \phi w^{-1} \rangle_J = f$, $\langle \phi^2 w^{-1} \rangle_J = F$.

Remark. Our usual disclaimer: the reader should not be confused by notations, but should be warned: of course F, f, u, v stand for numbers, not functions.

Notice that by scaling argument function \mathbf{B}_Q does not depend on J but depends on $Q = [w]_{A_2}$. Notice also that numbers F, f and u, v are not independent, in particular,

$$(4.23) \quad f^2 \leq F \cdot v, \quad 1 \leq u \cdot v \leq Q, \quad \lambda \geq 0.$$

This describes the domain $\Omega = \Omega_Q$ of definition of function B . Let

$$(4.24) \quad \sup_{\Omega_Q} \frac{\lambda B(F, f, u, v, \lambda)}{F} =: R.$$

By simple one-point-singularity weight we know that for an absolute positive constant κ the following holds

$$R \geq \kappa Q.$$

The question is of course whether one of the following is true

$$(4.25) \quad \begin{cases} \exists A < \infty \ R \leq AQ \\ \limsup_{Q \rightarrow \infty} R/Q = \infty \\ \exists \varepsilon > 0, a_\varepsilon > 0 : R \geq a_\varepsilon Q (\log Q)^\varepsilon ? \end{cases}$$

From [11], [6] we know (recall that \hat{Q} is not the weak norm but is the square of the weak norm $\|S : L^2(w) \rightarrow L^{2,\infty}(w)\|$):

$$(4.26) \quad R \leq AQ \log Q.$$

The goal of this section is to reduce the estimate of R to a solution of a PDI in a certain simple domain in \mathbb{R}^3 .

Lemma 4.5. *Function \mathbf{B}_Q introduced above is 1) even in variable f , 2) decreases in variable λ , 3) it is not greater than u .*

Proof. All claims of this lemma follow immediately from the definition of \mathbf{B}_Q . \square

4.3. The main inequality and the Partial Differential version of it.

Exactly like in Section 4.1.1 we can write the following *main inequality*. If $P = (F, f, u, v, \lambda)$, $P_+ = (F_+, f_+, u_+, v_+, \lambda_+)$, $P_- = (F_-, f_-, u_-, v_-, \lambda_-)$ belong to Ω_Q , and $F = \frac{1}{2}(F_+ + F_-)$, $f = \frac{1}{2}(f_+ + f_-)$, $u = \frac{1}{2}(u_+ + u_-)$, $v = \frac{1}{2}(v_+ + v_-)$, $\lambda = \min(\lambda_+, \lambda_-)$, then *the main inequality* holds with constant $\kappa = 1$:

$$(4.27) \quad \mathbf{B}(F, f, u, v, \lambda + \kappa(f_+ - f_-)^2) - \frac{\mathbf{B}(P_+) + \mathbf{B}(P_-)}{2} \geq 0.$$

In particular, function \mathbf{B} is convex in (F, f, u, v) variables and monotone decreasing (and left continuous) in λ . In infinitesimal form this inequality becomes

$$(4.28) \quad -\frac{1}{2}d_{F,f,u,v}^2 \mathbf{B} + \kappa \frac{\partial \mathbf{B}}{\partial \lambda} \cdot (df)^2 \geq 0.$$

By Aleksandrov's theorem [8] we have this inequality pointwise a. e. in Ω , if by $\frac{\partial \mathbf{B}}{\partial \lambda}$ we denote upper derivative in λ .

4.4. Reduction of the number of variables. Properties of $\mathbf{B} = \mathbf{B}_Q$ and $\Psi = \Psi_Q$. By the construction of $\mathbf{B} = \mathbf{B}_Q$ we can easily see two homogeneity properties:

$$\begin{aligned}\mathbf{B}\left(\frac{F}{t}, \frac{f}{t}, tu, \frac{v}{t}, \frac{\lambda}{t^2}\right) &= t\mathbf{B}(F, f, u, v, \lambda), \\ \mathbf{B}\left(\frac{F}{s^2}, \frac{f}{s}, u, v, \frac{\lambda}{s^2}\right) &= \mathbf{B}(F, f, u, v, \lambda).\end{aligned}$$

Denote $\Phi := \Phi_Q(\cdot, \cdot, \cdot, \cdot) := \mathbf{B}(\cdot, \cdot, \cdot, \cdot, 1)$, then from these relationships we obtain

$$\begin{aligned}\Phi\left(\frac{F}{\sqrt{\lambda}}, \frac{f}{\sqrt{\lambda}}, \sqrt{\lambda}u, \frac{v}{\sqrt{\lambda}}\right) &= \sqrt{\lambda}\mathbf{B}(F, f, u, v, \lambda), \\ \Phi\left(\frac{F}{\lambda}, \frac{f}{\sqrt{\lambda}}, u, v\right) &= \mathbf{B}(F, f, u, v, \lambda).\end{aligned}$$

Hence, with obvious notations for A, c , $\Phi(\sqrt{\lambda}A, c, \sqrt{\lambda}u, \frac{v}{\sqrt{\lambda}}) = \sqrt{\lambda}\Phi(A, c, u, v)$. Denote by $\Psi := \Psi_Q(\cdot, \cdot, *) = \Phi(\cdot, *, \cdot, 1)$. Then denoting by

$$(4.29) \quad a = \sqrt{\lambda}A = \frac{Fv}{\lambda}, \quad b = uv, \quad c = \frac{f}{\sqrt{\lambda}}$$

we can write

$$(4.30) \quad \Psi(a, b, c) = v\Phi\left(\frac{a}{v}, c, u, v\right) = v\mathbf{B}(F, f, u, v, \lambda).$$

We can rewrite all information on \mathbf{B} in terms of Ψ . For example, the domain $\Omega_Q = \{(F, f, u, v, \lambda) : f^2 \leq Fv, 1 < uv < Q, \lambda > 0\}$ of definition of \mathbf{B} is mapped to the domain of definition of Ψ :

$$G := G_Q := \{(a, b, c) : c^2 \leq a, 1 < b < Q\}.$$

(4.24) becomes

$$(4.31) \quad \sup_{(a,b,c) \in G_Q} \frac{\Psi(a, b, c)}{a} = R.$$

By the formula

$$(4.32) \quad \Psi\left(\frac{vF}{\lambda}, uv, \frac{f}{\sqrt{\lambda}}\right) = v\mathbf{B}(F, f, u, v, \lambda)$$

we see that Ψ is concave in all its three variables (just vary (F, f, u) to see this).

The main property 4.28 becomes the following a. e. pointwise in G property (it is a direct calculation with the change of variables (4.29)):

$$(4.33) \quad \begin{bmatrix} \Psi_{aa}, & \Psi_{ab} & \Psi_{ac} & 0 \\ \Psi_{ab} & \Psi_{bb} & \Psi_{bc} & 0 \\ \Psi_{ac} & \Psi_{bc} & \Psi_{cc} + \kappa(c\Psi_c + 2a\Psi_a) & -\Psi_c \\ 0 & 0 & -\Psi_c & 2\Psi - 2a\Psi_a - 2b\Psi_b \end{bmatrix} \leq 0$$

in $G = G_Q$ and with $\kappa = 1$.

We also have obstacle condition exactly as in Section 4.1.1. By the same example of the same weight w_s as in Section 4.1.1 we can see that

$$\mathbf{B}(t^2v, tv, u, v, \lambda) = u, \text{ as soon as } \lambda \leq \delta v^2$$

with a certain small absolute positive δ and any (positive or negative) t as soon as $uv \geq 10$. In terms of Ψ this reads as follows:

$$\Psi(c^2, b, c) = b, \text{ as soon as } \delta^{-1} < c, b \geq 10.$$

We have, thus, an obstacle condition on the intersection of the parabolic boundary

$$(4.34) \quad \partial_{\text{parab}}G = \{(a, b, c) : a = c^2, 1 \leq b \leq Q\}.$$

with $b \geq 10$. But function Ψ is concave in its three variables, therefore

$$(4.35) \quad \Psi(a, b, c) = b, \text{ as soon as } \delta^{-1} < c, b \geq 10.$$

We used here the fact that by (4.30)

$$\Psi(a, b, c) \leq b.$$

Function of (F, f, u, v, λ) in the left hand side of (4.32) is absolutely continuous in each of its variables (we reasoned that Ψ is concave), so then is the function in the right hand side of (4.32). Let us differentiate this formula in λ :

$$-\frac{\partial \Psi}{\partial a} \frac{vF}{\lambda^2} - \frac{1}{2} \frac{\partial \Psi}{\partial c} \frac{f}{\lambda^{3/2}} \leq 0,$$

because \mathbf{B} decreases in λ . We obtained one more property of Ψ :

$$(4.36) \quad 2a\Psi_a + c\Psi_c \geq 0.$$

From (4.33) one see that

$$(4.37) \quad a\Psi_a + b\Psi_b \geq \Psi.$$

This implies that function $\phi(t) := \Psi(ta, tb, c)$ defined on $t \in [\max(\frac{c^2}{a}, \frac{1}{b}), \frac{Q}{b}]$ is such that $\frac{\phi(t)}{t}$ is increasing. But ϕ is convex. So

$$(4.38) \quad Rat - \Psi(ta, tb, c) = Rat - \phi(t) \leq Ra\left(\frac{c^2}{a} + \frac{1}{b}\right) = R\left(c^2 + \frac{a}{b}\right).$$

4.5. The boundary conditions of \mathbf{B}_Q and Ψ_Q . Let us consider $\mathbf{B}_Q(F, f, u, v, \lambda)$ on $F = v, f = v$. This corresponds to the testing function φ to be just $\mathbf{1}_J$. We obtain exactly the function B_Q from Section 4.1.1:

$$B_Q(u, v, \lambda) = \mathbf{B}_Q(v, v, u, v, \lambda).$$

By homogeneity we can write it as

$$B_Q(u, v, \lambda) =: \frac{1}{\sqrt{\lambda}} \Theta_Q(u\sqrt{\lambda}, \frac{v}{\sqrt{\lambda}}),$$

where this equation serves as the definition of Θ_Q . By (4.32) we have a chain of equalities:

$$\frac{1}{\sqrt{\lambda}}\Theta_Q(u\sqrt{\lambda}, \frac{v}{\sqrt{\lambda}}) = B_Q(v, v, u, v, \lambda) = \frac{1}{v}\Psi_Q(\frac{v^2}{\lambda}, uv, \frac{v}{\sqrt{\lambda}}).$$

Thus, on the boundary G_{parab} from (4.34) we obtain

$$(4.39) \quad \Psi_Q(c^2, b, c) = c\Theta_Q(\frac{b}{c}, c).$$

Recall that we found the formula for certain function Θ in (4.11). And we actually proved that

$$(4.40) \quad \Theta_Q \leq A\Theta$$

in the domain $H = \{(x, y) \in \mathbb{R}_+^2 : 1 < xy < Q\}$. But by formula (4.11) $\Theta(x, y) \leq Qy$ at least for $y \leq Q^{-1/2}$.

So for small $c, c \leq Q^{-1/2}$ we have

$$\Psi_Q(c^2, b, c) \leq AQc^2.$$

And this is true for all $b \in (1, Q)$.

But there is another obstacle condition that shows that

$$\Psi_Q(c^2, b, 0) \geq A'Qc^2.$$

4.6. Full weak weighted estimate of square function. First we provide information on full weak estimate for square function, namely on the asymptotics of constant $C([w]_{A_2})$ from (4.3).

The following result proved by C. Domingo-Salazar, M. Lacey, G. Rey [6].

Theorem 4.6. *Let $w \in A_2$, then the norm of the square function operator $S: L^2(w) \rightarrow L^{2,\infty}(w)$ is bounded by $C[w]_{A_2}^{1/2} \log^{1/2}(1 + [w]_{A_\infty})$, where C is an absolute constant.*

The claim of the theorem is equivalent to the following inequality that should be proved for an arbitrary function $\varphi \in L^2((0, 1), w)$:

$$(4.41) \quad w\{x \in (0, 1): S\varphi(x) > 1\} \leq C[w]_{A_2} \log(1 + [w]_{A_\infty}) \|\varphi\|_w^2.$$

Let us make several remarks.

Remark. In the previous section we studied the case of special functions φ , namely, we proved that if $\varphi = w^{-1}\mathbf{1}_I, I \in \mathcal{D}$, then inequality (4.41) can be strengthened. By this we mean that for such special test functions the estimate above has the right hand side $C[w]_{A_2} \|\varphi\|_w^2$. There is no logarithmic blow-up.

Remark. Exponent $p = 2$ is critical for Theorem 4.6. By this we mean that one can quite easily deduce from this theorem the result for $p > 2$: if $w \in A_p$, then the norm of the square function operator $S: L^p(w) \rightarrow L^{p,\infty}(w)$ is bounded by $C[w]_{A_p}^{1/2} \log^{1/2}(1 + [w]_{A_\infty})$, where C is an absolute constant. For that reduction to the case $p = 2$ the reader can look at [6]. It is important

to note that for $1 \leq p < 2$, and $w \in A_p$, [6] proves the estimate $C[w]_{A_p}^{1/2}$ for the weak norm of the square function operator.

5. RESTRICTED WEAK WEIGHTED ESTIMATE OF THE SQUARE FUNCTION

5.0.1. *Sparse square function operators.* **Definition.** A family \mathcal{S} of intervals of \mathcal{D} is called ε -sparse if the following condition is satisfied

$$(5.1) \quad \sum_{\substack{I \in \mathcal{S}, \\ I \subsetneq J}} |I| \leq \varepsilon |J|, \quad \forall J \in \mathcal{S}.$$

Definition. Sparse square function operator is defined for each sparse family \mathcal{S} as follows

$$S^{\text{sp}} \varphi \stackrel{\text{def}}{=} S_{\mathcal{S}}^{\text{sp}} \varphi \stackrel{\text{def}}{=} \left(\sum_{I \in \mathcal{S}} \langle \varphi \rangle_I^2 \mathbf{1}_I \right)^{1/2}.$$

Theorem 5.1. *For any $\varepsilon > 0$ and any $\varphi \in L^1$ there exists a constant $C = C(\varepsilon)$ independent of φ and a sparse family \mathcal{S} (depending on ε and on φ) such that pointwisely almost everywhere*

$$S\varphi \leq CS^{\text{sp}}\varphi.$$

Proof. It is well known (see e.g. [30]) that the square function operator is weakly bounded in unweighted L^1 . Let us call A the norm of the operator S from L^1 to $L^{1,\infty}$. Fix ε and let $C = 100A/\varepsilon$. We start with interval $I_0 = (0, 1)$, put $\mathcal{S}_0 \stackrel{\text{def}}{=} \{I_0\}$, and define the first generation of stopping intervals \mathcal{S}_1 as follows: $Q \in \mathcal{S}_1$ if it is the maximal interval in I_0 such that

$$S_Q^{I_0} \varphi \stackrel{\text{def}}{=} \left(\sum_{\substack{I \in \mathcal{D}, \\ Q \subset I \subset I_0}} (\Delta_I \varphi)^2 \right)^{1/2} > C \langle |\varphi| \rangle_{I_0};$$

The second generations of stopping intervals \mathcal{S}_2 will be nested inside the first generation \mathcal{S}_1 . For every $I \in \mathcal{S}_1$ we define its subintervals from \mathcal{D} by the same rule as before, but with I playing the rôle of I_0 . Namely, we define the first generation of stopping intervals \mathcal{S}_2 inside $I \in \mathcal{S}_1$ as follows: $Q \in \mathcal{S}_2$ if it is the maximal interval in I such that

$$S_Q^I \varphi \stackrel{\text{def}}{=} \left(\sum_{\substack{J \in \mathcal{D}, \\ Q \subset J \subset I}} (\Delta_J \varphi)^2 \right)^{1/2} > C \langle |\varphi| \rangle_I;$$

We continue the construction of generations of intervals $\mathcal{S}_3, \mathcal{S}_4, \dots$ recursively, and we put $\mathcal{S} \stackrel{\text{def}}{=} \cup_{k=0}^{\infty} \mathcal{S}_k$.

Notice that by the fact that operator S and dyadic maximal operator M are weakly bounded in unweighted L^1 and from our choice of constant C at the beginning of the proof, we get that $\sum_{Q \in \mathcal{S}_1} |Q| \leq \frac{\varepsilon}{50} |I_0|$, and similarly,

$$\sum_{\substack{Q \in \mathcal{S}_{k+1}, \\ Q \subset I}} |Q| \leq \frac{\varepsilon}{50} |I|, \quad \forall I \in \mathcal{S}_k.$$

Obviously, and with a good margin, we obtained that \mathcal{S} is ε -sparse.

Now to see the pointwise estimate of the theorem, let us notice that given $x \in I_0$, which is not an end-point of any dyadic interval, we will be able to find the tower of intervals $\cdots \subsetneq I_k \subsetneq I_1 \subsetneq I_0$ such that x is contained in all of them and such that $I_k \in \mathcal{S}_k$. This tower may degenerate to just one interval I_0 , or it can be an infinite tower. But the set of points for which the tower is infinite has Lebesgue measure zero. This is clear from the fact that \mathcal{S} is sparse. In any case,

$$\begin{aligned} S^2\varphi(x) &= \sum_{\substack{I \in \mathcal{D}, \\ I_1 \subsetneq I \subset I_0}} (\Delta_I \varphi)^2 + \sum_{\substack{I \in \mathcal{D}, \\ I_2 \subsetneq I \subset I_1}} (\Delta_I \varphi)^2 \\ &+ \sum_{\substack{I \in \mathcal{D}, \\ I_3 \subsetneq I \subset I_2}} (\Delta_I \varphi)^2 + \dots \end{aligned}$$

But then, using our stopping criterion we see that the last expression is bounded by

$$C^2(\langle |\varphi| \rangle_{I_0}^2 + \langle |\varphi| \rangle_{I_1}^2 + \langle |\varphi| \rangle_{I_2}^2 + \dots),$$

which proves the theorem. □

Remark. Now we will see what can be changed in the reasoning of [6] in order to obtain the estimate better than (??) for a special choice of $\varphi = w^{-1}$. In fact, the Bellman function technique of the previous section gave us a better estimate in this particular case $\varphi = w^{-1}$:

$$(5.2) \quad w\{x \in I_0 : S^{\text{sp}}w^{-1} > 3\} \leq A[w]_{A_2} \int_{I_0} w^{-1} dx.$$

Theorem 5.2. *Let $w \in A_2$ and \mathcal{S} be a collection of sparse dyadic intervals in $\mathcal{D}(I_0)$. Let S^{sp} be the sparse square function operator built on this collection. Then the restricted weak type of the operator $S^{\text{sp}}_{w^{-1}}$ from $L^2(w^{-1})$ to $L^{2,\infty}(w)$ is bounded by $A[w]_{A_2}^{1/2}$, where A is an absolute constant.*

We need the following well-known result:

Lemma 5.3. *Let M be the dyadic maximal operator, and w be in dyadic A_2 . Then the norm of M from $L^2(w)$ to $L^{2,\infty}(w)$ is bounded by $[w]_{A_2}^{1/2}$.*

Proof. Given a test function $\varphi \geq 0$, let $\{I\}$ be the maximal dyadic intervals for which $M\varphi > 1$. Then

$$\begin{aligned} \sum_I w(I) &\leq \int \sum \frac{w(I)}{|I|} \varphi \mathbf{1}_I dx = \int \sum_I \frac{w(I)}{|I|} \varphi \mathbf{1}_I w^{-1/2} w^{1/2} dx \\ &\leq \left(\sum_I \left(\frac{w(I)}{|I|} \right)^2 w^{-1}(I) \right)^{1/2} \|f\|_w \leq [w]_{A_2}^{1/2} \left(\sum_I w(I) \right)^{1/2} \|\varphi\|_w \end{aligned}$$

Hence,

$$(w\{x: M\varphi > 1\})^{1/2} = \left(\sum_I w(I)\right)^{1/2} \leq [w]_{A_2}^{1/2} \|\varphi\|_w,$$

which is precisely what the lemma claims. \square

Remark. One can skip the word “dyadic” everywhere in the statement of this lemma. Also one can generalize the statement to \mathbb{R}^n . Then lemma remains true, only the estimate becomes by $C_n[w]_{A_2}^{1/2}$. See [30].

Let us assume that $0 \leq \varphi \leq w^{-1}\mathbf{1}_{I_0}$.

As in [6] let \mathcal{S} be an ε -sparse system of dyadic sub-intervals of I_0 , and S^{sp} be a corresponding sparse square function. We split $\mathcal{S} = \mathcal{S}^0 \cup \bigcup_{m=0}^{\infty} \mathcal{S}_m$, where \mathcal{S}^0 consists of intervals of \mathcal{S} such that $\langle \varphi \rangle_I > 1$, and \mathcal{S}_{m+1} consists of intervals of \mathcal{S} such that

$$(5.3) \quad 2^{-m-1} < \langle |\varphi| \rangle_I \leq 2^{-m}, \quad m = 0, 1, \dots$$

We denote

$$S_m^{\text{sp}}\varphi \stackrel{\text{def}}{=} \left(\sum_{I \in \mathcal{S}_m} \langle \varphi \rangle_I^2 \mathbf{1}_I \right)^{1/2}.$$

We are going to estimate these measures

$$\begin{aligned} W_0 &\stackrel{\text{def}}{=} w\{x \in I_0: (S_0^{\text{sp}}\varphi)^2 > 1\}, \\ W &\stackrel{\text{def}}{=} w\{x \in I_0: \sum_{m=0}^{\infty} (S_m^{\text{sp}}\varphi)^2 > 4\}. \end{aligned}$$

The estimate of W_0 is easy. The sum $(S_0^{\text{sp}}\varphi)^2$ is supported on intervals where the dyadic maximal function of φ is bigger than 1. The set, where the dyadic maximal function is bigger than 1 has w -measure bounded by $[w]_{A_2} \|\varphi\|_w^2$ by Lemma 5.3.

Now we work with W , and we start exactly as in the estimate of W above. Namely, the support of $S_m^{\text{sp}}\varphi$ is in $\cup Q_i^{(m)}$, where $Q_i^{(m)} \in \mathcal{S}_m$ that are the maximal dyadic intervals Q with the property

$$(5.4) \quad \sum_{I \in \mathcal{S}_m, Q \subset I \subset I_0} \langle |\varphi| \rangle_I^2 \mathbf{1}_I \geq 2^{-m}$$

By maximality and by (5.3)

$$(5.5) \quad \sum_{I \in \mathcal{S}_m, Q_i^{(m)} \subset I \subset I_0} \langle |\varphi| \rangle_I^2 \mathbf{1}_I \leq 2^{-m} + 2^{-2m}.$$

To estimate W we use a union bound.

$$\begin{aligned}
 & w\{x \in I_0 : \sum_{m=0}^{\infty} (S_m^{\text{SP}}\varphi)^2 > 4\} \\
 & \leq w\{x \in I_0 : \sum_{m=0}^{\infty} (S_m^{\text{SP}}\varphi)^2 > \sum_{m=0}^{\infty} 2^{-m/2}\} \\
 & \leq \sum_{m=0}^{\infty} w\{x \in \cup_i Q_i^{(m)} : (S_m^{\text{SP}}\varphi)^2 > 2^{-m/2}\} = \\
 & \leq \sum_{m=0}^{\infty} \sum_i w\{x \in Q_i^{(m)} : (S_m^{\text{SP}}\varphi)^2 > 2^{-m/2}\}.
 \end{aligned}$$

Now let us fix $Q_i^{(m)}$. Then

$$\sum_{I \in \mathcal{S}_m, Q_i^{(m)} \subset I \subset I_0} \langle |\varphi| \rangle_I^2 \mathbf{1}_I \geq 2^{-m}$$

On this interval $\langle |\varphi| \rangle_I^2 \leq 2^{2m}$ by the definition of \mathcal{S}_m . Hence,

$$\sum_{I \in \mathcal{S}_m, Q_i^{(m)} \subset I \subset I_0} \mathbf{1}_I \geq 2^m.$$

Therefore, every $Q_i^{(m)}$ lies at least 2^m generations down in \mathcal{S} . By sparse choice

$$(5.6) \quad \left| \cup_i Q_i^{(m)} \right| \leq e^{-a2^m}.$$

Consider maximal dyadic intervals $\ell \in \mathcal{S}_m$, $\ell \subset Q_i^{(m)}$, on which $(S_m^{\text{SP}}\varphi)^2 > 2^{-m/2}$. Call them $\mathcal{L}(Q_i^{(m)})$. Let us estimate $\sum_{\ell \in \mathcal{L}(Q_i^{(m)})} |\ell|$.

Notice that for each $\ell \in \mathcal{L}(Q_i^{(m)})$

$$\sum_{I \in \mathcal{S}_m, \ell \subset I \subset Q_i^{(m)}} \langle |\varphi| \rangle_I^2 \mathbf{1}_I \geq 2^{-m/2} - 2^{-m} - 2^{-2m} \geq \frac{1}{2} 2^{-m/2}.$$

This is by (5.5). Now we use again that $\langle |\varphi| \rangle_I^2 \leq 2^{2m}$ for all terms of this sum. Hence,

$$\sum_{I \in \mathcal{S}_m, \ell \subset I \subset Q_i^{(m)}} \mathbf{1}_I \geq 2^{-m/2} - 2^{-m} - 2^{-2m} \geq 2^{-m}.$$

Therefore, every ℓ from $\mathcal{L}(Q_i^{(m)})$ lies at least 2^m generations down from $Q_i^{(m)}$ in \mathcal{S}_m , and, thus, in \mathcal{S} . By sparse construction

$$(5.7) \quad \sum_{\ell \in \mathcal{L}(Q_i^{(m)})} |\ell| \leq e^{-a2^m} |Q_i^{(m)}|.$$

Recall that we assumed

$$\varphi \leq w^{-1},$$

Hence,

$$\langle w^{-1} \rangle_\ell^{-1} \leq \langle \varphi \rangle_\ell^{-1}.$$

By definition $\langle \varphi \rangle_\ell \geq 2^{-m-1}$, hence

$$\begin{aligned} \langle w \rangle_\ell &\leq [w]_{A_2} \langle w^{-1} \rangle_\ell^{-1} \leq [w]_{A_2} \langle \varphi \rangle_\ell^{-1} \\ &\leq 2^{m+1} [w]_{A_2}, \quad \forall \ell \in \mathcal{L}(Q_i^{(m)}). \end{aligned}$$

Therefore, by (5.7) we obtain

$$\begin{aligned} w\{x \in Q_i^{(m)} : (S_m^{\text{sp}} \varphi)^2 > 2^{-m/2}\} &= \sum_{\ell \in \mathcal{L}(Q_i^{(m)})} w(\ell) \\ &= \sum_{\ell \in \mathcal{L}(Q_i^{(m)})} \langle w \rangle_\ell |\ell| \leq 2^{m+1} [w]_{A_2} \sum_{\ell \in \mathcal{L}(Q_i^{(m)})} |\ell| = 2^{m+1} [w]_{A_2} e^{-a2^m} |Q_i^{(m)}|. \end{aligned}$$

We can conclude now that

$$\begin{aligned} w\{x \in Q_i^{(m)} : (S_m^{\text{sp}} \varphi)^2 > 2^{-m/2}\} &\leq 2^{m+1} [w]_{A_2} e^{-a2^m} |Q_i^{(m)}| \leq \\ &2^{2m+2} [w]_{A_2} e^{-a2^m} \langle \varphi \rangle_{Q_i^{(m)}} |Q_i^{(m)}| = \\ &2^{2m+2} [w]_{A_2} e^{-a2^m} \int_{Q_i^{(m)}} \varphi \, dx, \end{aligned}$$

where the last inequality follows from the definition of \mathcal{S}_m and the fact that $Q_i^{(m)} \in \mathcal{S}_m$.

Gathering things together, we obtain (5.2):

$$\begin{aligned} w\{x \in I_0 : \sum_{m=0}^{\infty} (S_m^{\text{sp}} \varphi)^2 > 4\} &\leq \\ &\sum_{m=0}^{\infty} \sum_i w\{x \in Q_i^{(m)} : (S_m^{\text{sp}} \varphi)^2 > 2^{-m/2}\} \\ &\leq [w]_{A_2} \sum_{m=0}^{\infty} \sum_i 2^{2m+2} e^{-a2^m} \int_{Q_i^{(m)}} \varphi \, dx \\ &\leq [w]_{A_2} \sum_{m=0}^{\infty} 2^{2m+2} e^{-a2^m} \int_{\cup Q_i^{(m)}} \varphi \, dx \\ &\leq [w]_{A_2} \int_{I_0} \varphi \, dx \sum_{m=0}^{\infty} 2^{2m+2} e^{-a2^m} \leq A [w]_{A_2} \int_{I_0} \varphi \, dx. \end{aligned}$$

Estimate

$$(5.8) \quad w\{x \in Q_i^{(m)} : \sum_{m=0}^{\infty} (S_m^{\text{sp}} \varphi)^2 > 4\} \leq A [w]_{A_2} \int_{I_0} \varphi \, dx$$

was just obtained by using one assumption: $\varphi \leq w^{-1}$. We want to exchange in the right hand side of (5.8) the integral $\int_{I_0} \varphi dx$ for the integral $\int_{I_0} \varphi^2 w dx$. This is trivial if one more property of φ holds, namely, if pointwisely

$$\varphi \leq C\varphi^2 w.$$

This is compatible with $\varphi \leq w^{-1}$ if for a. e. point $x \in I_0$ one of the following properties occurs: for every $x \in I_0$ either 1) $\frac{1}{C}w^{-1}(x) \leq \varphi(x) \leq w^{-1}(x)$, or 2) $\varphi(x) = 0$.

We conclude that (5.8) holds for every φ of the form $\varphi = w^{-1}\mathbf{1}_E$, where E is a measurable subset of I_0 .

In particular, Theorem 5.2 is proved.

Now we can use sparse domination Theorem 5.1. Then this theorem gives us the following one.

Theorem 5.4. *Let $w \in A_2$ and \mathcal{S} . Let S be the dyadic square function operator. Then the restricted weak type of the operator $S_{w^{-1}}$ from $L^2(w^{-1})$ to $L^{2,\infty}(w)$ is bounded by $A[w]_{A_2}^{1/2}$, where A is an absolute constant.*

REFERENCES

- [1] I. ASSANI, Z. BUCZOLICH, AND R.D. MAULDIN, *An L^1 counting problem in ergodic theory*, J. Anal. Math. 95 (2005), 221–241.
- [2] F. BARTHE, N. HUET, *On Gaussian Brunn–Minkowski inequalities*. Stud. Math. **191**, (2009), pp. 283–304 .
- [3] F. Barthe, *On a reverse form of the Brascamp–Lieb inequality*, Invent. Math., **134** (2), (1998) 235–361.
- [4] M. J. CARRO, *From restricted weak type to strong type estimates*, J. London Math. Soc. 70 (2004), no. 3, 750–762.
- [5] M. CARRO, L. GRAFAKOS, *WEIGHTED WEAK TYPE (1,1) ESTIMATES VIA RUBIO DE FRANCIA EXTRAPOLATION*, Journal of Functional Analysis, 269 (2015), no. 5, 1203–1233.
- [6] C. DOMINGO-SALAZAR, M. LACEY, G. REY, *Borderline Weak Type Estimates for Singular Integrals and Square Functions*, arXiv:1505.01804.
- [7] CRUZ-URIBE, DAVID V.; MARTELL, JOSÉ MARIA; PÉREZ, CARLOS, *Weights, extrapolation and the theory of Rubio de Francia*. Operator Theory: Advances and Applications, 215. Birkhäuser/Springer Basel AG, Basel, 2011. xiv+280 pp.
- [8] L. C. EVANS, R. F. GARIEPY, *Measure Theory and Fine Properties of Functions.*, 1992.
- [9] S. HUKOVIC, S. TREIL, 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.
- [10] T. HYTÖNEN, C. PÉREZ, S. TREIL, A. VOLBERG, *Sharp weighted estimates for dyadic shifts and the A_2 conjecture*, arXiv:1010.0755v2, to appear in J. für die reine und angew. Math.
- [11] M. LACEY, J. SCURRY, *Weighted weak type estimates for square function*, arXiv:1211.4219.
- [12] A. LERNER, *Sharp weighted norm inequalities for Littlewood–Paley operators and singular integrals*, Adv. Math., 226 (2011), 3912–3926.
- [13] A. LERNER, *On sharp aperture-weighted estimates for square functions*, J. Fourier Anal. Appl., 20 (2014), no. 4, 784–800.

- [14] A. LERNER, *Mixed $A_p - A_r$ inequalities for classical singular integrals and Littlewood–Paley operators*, J. Geom. Anal., 23 (2013), no. 3, 1343–1354.
- [15] A. LERNER, *A simple proof of the A_2 conjecture*, Int. Math. Res. Not., IMRN, 14 (2013), 3159–3170.
- [16] A. LERNER, *On an estimate of Calderón-Zygmund operators by dyadic positive operators*, J. Anal. Math., 121 (2013), no. 1, 141–161.
- [17] A. LERNER, S. OMBROSI, C. PÉREZ, *A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden*, Math. Res. Lett., 16 (2009) no. 1, 149–156.
- [18] A. LERNER, F. DI PLINIO, *On weighted norm inequalities for the Carleson and Walsh-Carleson operators*, J. London Math. Soc. 90 (2014), no. 3, 654–674.
- [19] N. Nadirashvili, V. Tkachev, S Vladut, *Nonlinear Elliptic Equations and Nonassociative Algebras*, AMS Math. Surveys and Monographs, v. 200, 2014.
- [20] F. NAZAROV, A. REZNIKOV, V. VASYUNIN, A. VOLBERG, *A Bellman function counterexample to the A_1 conjecture: the blow-up of weak norm estimates of weighted singular operators*, arXiv:1506.04710v1, pp. 1–23.
- [21] C. PÉREZ, S. TREIL, A. VOLBERG, *On A_2 conjecture and corona decomposition of weights*, arXiv:1006.2630.
- [22] M.C. REGUERA, *On Muckenhoupt-Wheeden conjecture*, Adv. Math. 227 (2011), no. 4, 1436–1450.
- [23] M.C. REGUERA, C. THIELE, *The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$* , arXiv:1011.1767, Math. Res. Lett. 19 (2012), no. 1, 1–7.
- [24] A. REZNIKOV, V. VASYUNIN, A. VOLBERG, *Extremizers and the Bellman function for the weak type martingale inequality*, arxiv: arXiv:1311.2133.
- [25] D. M. STOLYAROV, P. B. ZATITSKIY *Theory of locally concave functions and its applications to sharp estimates of integral functionals* preprint, 2014, <http://arxiv.org/abs/1412.5350>.
- [26] C. THIELE, S. TREIL, A. VOLBERG, *Weighted martingale multipliers in the non-homogeneous setting and outer measure spaces*, Adv. Math. 285 (2015), 1155–1188.
- [27] V. VASYUNIN, *The sharp constant in the John–Nirenberg inequality*, Preprint POMI no. 20, 2003.
- [28] V. VASYUNIN AND A. VOLBERG, *The Bellman functions for a certain two-weight inequality: A case study*, Algebra i analiz, 18 (2006), no. 2, 24–56 (in Russian); English transl. in: St. Petersburg Math. J., 18 (2007), 201–222.
- [29] V. VASYUNIN AND A. VOLBERG, *Monge–Ampère Equation and Bellman Optimization of Carleson Embedding Theorems*, Amer. Math. Soc. Transl. Ser. 2, Vol. 226 (2009), 195–238.
- [30] M. WILSON, *Weighted Littlewood–Paley Theory and Exponential-Square Integrability*, Lecture Notes in Math, Vol 1924, Springer-Verlag, 2007.

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI. 48823

Email address: ivanisvi@math.msu.edu (P. Ivanisvili)

UNIVERSITÀ DI BOLOGNA, DEPARTMENT OF MATHEMATICS, PIAZZA DI PORTA S. DONATO, 40126 BOLOGNA (BO)

Email address: pavel.mozolyako@unibo.it

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI. 48823

Email address: volberg@math.msu.edu (A. Volberg)