Ergod. Th. & Dynam. Sys. (2018), **38**, 244–256 © *Cambridge University Press, 2016*
doi:10.1017/etds.2016.27

Weighted and vector-valued variational estimates for ergodic averages

BEN KRAUSE† and PAVEL ZORIN-KRANICH‡

† *UCLA Math Sciences Building, Los Angeles, CA 90095-1555, USA*
(e-mail: benkrause23@math.ucla.edu)
‡ *Universität Bonn, Mathematisches Institut, Endenicher Allee 60, 53115 Bonn, Germany*
(e-mail: pzorin@uni-bonn.de)

(*Received* 3 *March* 2015 *and accepted in revised form* 24 *February* 2016)

Abstract. We prove weighted and vector-valued variational estimates for ergodic averages on $\mathbb{R}^d$. The weighted square function estimate relating ergodic averages to the dyadic martingale is obtained using an ℓ^r version of a reverse Hölder inequality for variation seminorms.

1. *Introduction*

We denote ergodic averages on $\mathbb{R}^d$, $d \geq 1$, by

$$A_t f(x) = \fint_{B(0,t)} f(x+y)\, dy,$$

where $B(0, t)$ is the ball of radius t centered at zero and $\fint_A = |A|^{-1} \int_A$. The homogeneous r-variation norm is denoted by

$$\|a_t\|_{\tilde{V}^r_{t\in I}} := \sup_{t_0<t_1<\cdots<t_J\in I} \left(\sum_{j=1}^{J} |a_{t_j} - a_{t_{j-1}}|^r\right)^{1/r}$$

and the inhomogeneous r-variation norm by

$$\|a_t\|_{V^r_{t\in I}} := \sup_{t\in I} |a_t| + \|a_t\|_{\tilde{V}^r_{t\in I}}.$$

Variational estimates for ergodic averages of the form

$$\| \|A_t f(x)\|_{V^r_{t>0}} \|_{L^p_x} \lesssim \|f\|_{L^p}$$

have been introduced by Bourgain [**Bou89**] and a complete theory has been developed by a number of authors [**JRW03**, **JKRW98**, **JSW08**] covering the full possible range of exponents p, r including the end points. Our purpose is to obtain weighted and vector-valued versions of these results.

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

A *weight* is a positive function on $\mathbb{R}^d$. The A_1 constant of a weight is given by

$$[w]_{A_1} = \sup_x Mw(x)/w(x); \tag{1.1}$$

the A_p constant, $1 < p < \infty$, is given by

$$[w]_{A_p} = \sup_Q \left(|Q|^{-1} \int_Q w\right)\left(|Q|^{-1} \int_Q w^{-1/(p-1)}\right)^{p-1}, \tag{1.2}$$

and a weight w is A_∞ if

$$w(E)/w(Q) \le C_w(|E|/|Q|)^\delta \tag{1.3}$$

for some $\delta \in (0, 1]$ and all $E \subset Q$. Here and later, the letter Q denotes a cube in $\mathbb{R}^d$, not necessarily dyadic unless explicitly stated so. We refer to [**GCF85**, Ch. IV] for the basic properties of A_p weights. Note that all constants in the results from [**GCF85**, Ch. IV] that we use can be taken to depend only on the A_p constants of the weights involved in them and not otherwise on the weight. This kind of dependence is required for the Rubio de Francia extrapolation theorem to apply (see e.g. [**Duo11**, Theorem 3.1]).

The set of dyadic cubes in $\mathbb{R}^d$ with side length 2^k is denoted by $\mathscr{Q}_k$. For a point x, we write $\mathscr{Q}_k(x)$ for the unique (away from a set of measure zero) dyadic cube with side length 2^k containing x. The conditional expectation onto the σ-algebra generated by $\mathscr{Q}_k$ is denoted by E_k.

Ergodic averages will be compared with a dyadic martingale using the short variations

$$S_k f := \|A_t f - E_k f\|_{V^2_{2^k \le t \le 2^{k+1}}}.$$

Following [**Kra13**], we will in fact use the (larger) smoothed version of the short variations given by

$$\tilde{S}_k f(x) := \sup_{y \in 3\mathscr{Q}_k(x)} S_k f(y).$$

Here $3\mathscr{Q}_k(x)$ denotes the concentric cube to $\mathscr{Q}_k(x)$ with three times the side length. The smoothed square function is defined by

$$\tilde{S} f := \left(\sum_k (\tilde{S}_k f)^2\right)^{1/2}.$$

Weighted bounds for a discrete version of $\tilde{S}$ have been proved in [**Kra13**] using the Auscher–Martell extrapolation theorem with limited range of exponents (see [**Duo11**, Theorem 7.1]). We present a direct proof relying on an ℓ^r version of a reverse Hölder inequality from [**JRW03**] and another alternative proof using a good λ inequality.

THEOREM 1.4. *The operator $\tilde{S}$ is bounded:*
(1) *on $L^p(w)$ for every $w \in A_p$, $1 < p < \infty$; and*
(2) *as an operator $L^1(w) \to L^{1,\infty}(w)$ for every $w \in A_1$.*
In each case the operator norm is bounded in terms of the A_p constant of the weight.

For an $\ell^\infty \to$ BMO (the space of functions with bounded mean oscillation) estimate in the discrete setting, see also [**Kra13**, §3]. The above estimate for the square function together with weighted bounds for the dyadic martingale square function and Haar multipliers will be used to show weighted and vector-valued bounds for the r-variation of ergodic averages.

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

THEOREM 1.5. *Let $r > 2$ and $1 < q < \infty$. The operator*

$$f_i(x) \mapsto (A_t f_i)(x)$$

acting on sequence-valued functions on $\mathbb{R}^d$ is bounded as an operator:

$$L^1_x(\ell^q_i, w) \to L^{1,\infty}_x(\ell^q_i(V^r_t), w) \quad \text{for any } w \in A_1(\mathbb{R}^d), \tag{1.6}$$

$$L^\infty_x(\ell^q_i) \to \mathrm{BMO}_x(\ell^q_i(V^r_t)), \quad \text{and} \tag{1.7}$$

$$L^p_x(\ell^q_i, w) \to L^p_x(\ell^q_i(V^r_t), w) \quad \text{for any } 1 < p < \infty \text{ and } w \in A_p(\mathbb{R}^d). \tag{1.8}$$

In each case the operator norm is bounded in terms of the A_p constant of the weight and grows as $r/(r-2)$ for $r \to 2$.

Theorem 1.5 is a joint generalization of the Fefferman–Stein vector-valued maximal inequality [**FS71**] and the Bourgain–Lépingle inequality [**Bou89**, Lemma 3.11] (the latter is a real-line version of the Lépingle inequality for martingales [**Lép76**]; see also [**PX88**]). Our proof also works for dilates of neighborhoods of the origin other than balls, provided that their boundary is sufficiently smooth. We do not state the precise smoothness requirements, but note that this applies e.g. to cubes.

As we have been made aware of after the completion of this work, the one-dimensional case of Theorem 1.5 has been previously established by Ma *et al* [**MTX15a**]. On the other hand, the additional flexibility of Lemma 2.1 later found an application in the same authors' work on variation norm estimates for rough truncations of singular integrals in several dimensions [**MTX15b**].

Since bounds for a fixed t in Theorem 1.5 are immediate, it suffices to treat the homogeneous variation. Throughout the proof we assume that $r < \infty$, which is justified by the fact that the conclusion of the theorem becomes stronger as r decreases. We also assume that the coordinate functions f_i are smooth on $\mathbb{R}^d$; the general case follows by standard truncation arguments.

The p-maximal function and the sharp p-maximal function are denoted by

$$M_p f(x) = \sup_{Q \ni x} \left(|Q|^{-1} \int_Q |f|^p \right)^{1/p}, \quad M^\sharp_p f(x) = \sup_{Q \ni x} \inf_c \left(|Q|^{-1} \int_Q |f - c|^p \right)^{1/p}. \tag{1.9}$$

The subscript p is omitted if $p = 1$.

2. *A reverse Hölder inequality*

For $1 < r < \infty$, let

$$R_k b(x) := \|A_t b\|_{V^r_{t \in [2^k, 2^{k+1}]}}, \quad \tilde{R}_k b(x) = \sup_{y \in 3Q_k(x)} R_k b(y), \quad \tilde{R}(b) := \left(\sum_k \tilde{R}_k(b)^r \right)^{1/r}.$$

Denote the side length of a cube Q by $\ell(Q)$.

LEMMA 2.1. (Cf. [**JRW03**, Lemma 4.2]) *Let $1 < r < \infty$. Let $\mathcal{Q}$ be a collection of disjoint cubes Q of size $\lesssim 2^k$. For each Q, let b^Q be a scalar-valued function supported on Q with $\int b^Q = 0$. Then, for every $\alpha > (d-1)/r'$, we have*

$$R_k \left(\sum_Q b^Q(x) \right)^r \lesssim_{\alpha,r,d} (2^k)^{\alpha r} \sum_Q \ell(Q)^{-\alpha r} R_k(b^Q(x))^r. \tag{2.2}$$

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

Proof. We consider only the homogeneous variation; in order to get the inhomogeneous variation, it suffices to additionally consider an arbitrary (but fixed) t, which is similar but easier. Fix an arbitrary sequence $2^k \le t_1 < \cdots < t_J \le 2^{k+1}$ and split the sum over cubes in

$$\sum_j \Big| \sum_Q (A_{t_{j+1}} - A_{t_j}) b^Q(x) \Big|^r.$$

According to the size of the cubes,

$$\sum_j \Big| \sum_{i \le 0} \sum_{Q:\ell(Q)\approx 2^{k+i}} (A_{t_{j+1}} - A_{t_j}) b^Q(x) \Big|^r.$$

Let α be chosen later and apply Hölder's inequality in the sum over i to estimate this by

$$\sum_j \Big(\sum_{i \le 0} \sum_{Q:\ell(Q)\approx 2^{k+i}} |2^{-\alpha i} (A_{t_{j+1}} - A_{t_j}) b^Q(x)|^r \Big) \times \Big(\sum_{i \le 0} \sum_{Q:\ell(Q)\approx 2^{k+i}} |2^{\alpha i} 1_{Q: Q \cap (\partial B(x,t)) \neq \emptyset, t = t_j \text{ or } t_{j+1}}|^{r'} \Big)^{r/r'}. \tag{2.3}$$

In order to estimate the second bracket, note that $Q \cap (\partial B(x,t)) \neq \emptyset$ implies that Q is contained in a spherical shell of radius $\approx 2^k$ and thickness $\approx 2^{k+i}$. For a fixed i, there can be at most $O(2^{-i(d-1)})$ cubes of this kind. Hence, the second bracket in (2.3) is bounded by

$$\Big(\sum_{i \le 0} 2^{\alpha i r'} 2^{-i(d-1)} \Big)^{r/r'},$$

which is finite by the assumption on α. This gives for (2.3) the bound

$$\sum_j \Big(\sum_{i \le 0} \sum_{Q:\ell(Q)\approx 2^{k+i}} |2^{-\alpha i} (A_{t_{j+1}} - A_{t_j}) b^Q(x)|^r \Big) \lesssim \sum_j \sum_Q \ell(Q)^{-\alpha r} (2^k)^{\alpha r} |(A_{t_{j+1}} - A_{t_j}) b^Q(x)|^r.$$

Taking the supremum over sequences $(t_j)_j$, we obtain (2.2). □

Estimating the r-variation norm by the 1-variation norm and noting that only cubes with $\operatorname{dist}(x, Q) \lesssim 2^k$ contribute to the sum, we obtain the following result.

COROLLARY 2.4. *In the situation of Lemma 2.1, we have*

$$\tilde{R}_k \Big(\sum_Q b^Q(x) \Big)^r \lesssim_{\alpha,r,d} (2^k)^{\alpha r - dr} \sum_{Q:\operatorname{dist}(x,Q) \lesssim 2^k} \ell(Q)^{-\alpha r} \|b^Q\|_1^r.$$

3. *Strong-type bounds for the square function*

Recall that a Haar function on a dyadic cube is a function that is constant on each dyadic subcube and has integral zero. A Haar function is L^∞ normalized if it is bounded by 1.

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

LEMMA 3.1. *Let $k \in \mathbb{Z}$ and $j < 0$. Let h_Q be L^∞ normalized Haar functions supported on the cubes $Q \in \mathcal{Q}_{k+j+1}$. Let $1 < p < \infty$ and $w \in A_p$. Then*

$$\Big\| \tilde{S}_k \Big(\sum_{Q \in \mathcal{Q}_{k+j+1}} h_Q E_{k+j+1} f \Big) \Big\|_{L^p(w)} \lesssim 2^{-\epsilon|j|} \|f\|_{L^p(w)}, \tag{3.2}$$

where $\epsilon > 0$ and the implied constant depend only on p and $[w]_{A_p}$, but not on k, j, or the Haar functions.

Proof. By homogeneity, we may assume that $k = 0$. Let $w \in A_p$; then there exists $r \in (1, \min(2, p))$ such that $w \in A_{p/r}$. Write f_Q for the value of $E_{j+1} f$ on Q and observe that $E_0 \sum_{Q \in \mathcal{Q}_{j+1}} h_Q f_Q = 0$. Hence, we have

$$\Big\| \tilde{S}_0 \Big(\sum_{Q \in \mathcal{Q}_{j+1}} h_Q E_{j+1} f \Big) \Big\|_{L^p(w)} \lesssim \Big\| \tilde{R}_0 \Big(\sum_{Q \in \mathcal{Q}_{j+1}} h_Q f_Q \Big) \Big\|_{L^p(w)},$$

where $\tilde{R}_0$ is defined using the above value of r. Applying Corollary 2.4, we obtain

$$\tilde{R}_0 \Big(\sum_{Q \in \mathcal{Q}_{j+1}} h_Q f_Q \Big)^r (x) \lesssim 2^{-r\alpha j} \sum_{Q \in \mathcal{Q}_{j+1}, \operatorname{dist}(x,Q) \lesssim 1} \|h_Q f_Q\|_1^r.$$

By Hölder's inequality, this is bounded by

$$\begin{aligned} 2^{-r\alpha j} \sum_{Q \in \mathcal{Q}_{j+1}, \operatorname{dist}(x,Q) \lesssim 1} \|1_Q f_Q\|_r^r \|h_Q\|_{r'}^r &\le 2^{-r\alpha j} 2^{jdr/r'} \int_{B(x,C)} |E_{j+1} f|^r \\ &\lesssim 2^{r(d/r' - \alpha) j} M_r (E_{j+1} f)^r (x). \end{aligned}$$

Thus, we have established that

$$\tilde{R}_0 \Big(\sum_{Q \in \mathcal{Q}_{j+1}} h_Q f_Q \Big)(x) \lesssim 2^{\epsilon j} M_r (E_{j+1} f)(x)$$

for any $\epsilon < 1/r'$. We are done, since the maximal function M_r is bounded on $L^p(w)$. □

Proof of Theorem 1.4, part (1). By Rubio de Francia's extrapolation theorem, it suffices to consider $p = 2$. Write $f = \sum_j d_j$, $d_j = E_j f - E_{j+1} f$. The result will follow if we can show that

$$\|\tilde{S}_k(d_{k+j})\|_{L^2(w)} \lesssim 2^{-\epsilon|j|} \|d_{k+j}\|_{L^2(w)}.$$

Indeed,

$$\tilde{S} f = \Big(\sum_k \Big(\tilde{S}_k \Big(\sum_j d_{k+j} \Big) \Big)^2 \Big)^{1/2} \le \sum_j \Big(\sum_k (\tilde{S}_k(d_{k+j}))^2 \Big)^{1/2},$$

and taking $L^2(w)$ norms on both sides we obtain

$$\|\tilde{S} f\|_{L^2(w)} \le \sum_j \Big(\sum_k \|\tilde{S}_k(d_{k+j})\|_{L^2(w)}^2 \Big)^{1/2} \lesssim \sum_j 2^{-\epsilon|j|} \Big(\sum_k \|d_{k+j}\|_{L^2(w)}^2 \Big)^{1/2}.$$

The sum over k on the right-hand side is the dyadic martingale square function which is bounded on $L^2(w)$; see [**Buc93**, Theorem 3.6] or [**LPR10**].

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

For $j \geq -10^d$ (say), we note that S_k is invariant under constant addition to use the estimate

$$S_k f(x) \lesssim \inf_c 2^{-kd} \int_{B_{C2^k}(x)} |f - c|.$$

It follows that an estimate of the same kind holds for $\tilde{S}_k$. Writing $d_{k+j} = \sum_{Q \in \mathcal{Q}_{k+j}} a_Q 1_Q$, we obtain

$$\tilde{S}_k d_{k+j} \lesssim \sum_{Q \in \mathcal{Q}_{k+j}} |a_Q| 1_{\partial Q + B_{C2^k}}.$$

The overlap of the characteristic functions on the right-hand side is bounded by an absolute constant, so

$$(\tilde{S}_k d_{k+j})^2 \lesssim \sum_{Q \in \mathcal{Q}_{k+j}} |a_Q|^2 1_{\partial Q + B_{C2^k}}.$$

Integrating this, we obtain

$$\|\tilde{S}_k d_{k+j}\|_{L^2(w)}^2 \lesssim \sum_{Q \in \mathcal{Q}_{k+j}} |a_Q|^2 w(\partial Q + B_{C2^k}).$$

Since the A_2 weight w also satisfies the A_∞ condition, this is bounded by

$$\sum_{Q \in \mathcal{Q}_{k+j}} |a_Q|^2 w(Q + Q_{C2^k})(|\partial Q + Q_{C2^k}|/|Q + Q_{C2^k}|)^\delta$$

for some $\delta > 0$ and, since w is doubling, this is bounded by

$$\sum_{Q \in \mathcal{Q}_{k+j}} |a_Q|^2 w(Q)(2^{-j})^\delta = 2^{-j\delta} \|d_{k+j}\|_{L^2(w)}^2.$$

For $j < -10^d$, write

$$d_{k+j} = \sum_{Q \in \mathcal{Q}_{k+j+1}} h_Q E_{k+j+1} f,$$

where $h_Q = d_{k+j} 1_Q / \|d_{k+j} 1_Q\|_\infty$ are normalized Haar functions (with the convention that $h_Q = 0$ if $d_{k+j} 1_Q = 0$) and $f = \sum_{Q \in \mathcal{Q}_{k+j+1}} 1_Q \|d_{k+j} 1_Q\|_\infty$. Note that $\|f\|_{L^2(w)} \lesssim \|d_{k+j}\|_{L^2(w)}$, since w is doubling and d_{k+j} is constant at scale 2^{k+j}. The claim now follows from Lemma 3.1. □

4. *The jump inequality*

Let $(a_t)_t$ be an arbitrary function. The *jump counting function* $N_\lambda(a_t)$ is the supremum over all J such that there exist $t_0 < t_1 < \cdots < t_J$ with $|a_{t_j} - a_{t_{j-1}}| > \lambda$ for all $j = 1, \ldots, J$. See [**JSW08**] for the basic properties of the jump counting function and its relation to the variational norms.

PROPOSITION 4.1. *Let $w \in A_p(\mathbb{R}^d)$, $1 < p < \infty$. Then*

$$\sup_{\lambda > 0} \lambda^{-1} \left\| \sqrt{N_\lambda(A_t f)} \right\|_{L^p(w)} \lesssim \|f\|_{L^p(w)}.$$

Proposition 4.1 implies the scalar-valued case of (1.8) in Theorem 1.5 by the interpolation argument in [**JSW08**, §2] (note that an A_p weight, $1 < p < \infty$, is also in A_q for all q in a neighborhood of p). Extrapolation then yields the vector-valued case of (1.8).

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

Proof. In view of the weighted bound for the square function (part (1) of Theorem 1.4), it suffices to show that

$$\sup_{\lambda>0} \lambda^{-1} \left\| \sqrt{N_\lambda(E_t f)} \right\|_{L^p(w)} \lesssim \|f\|_{L^p(w)},$$

where t takes dyadic values (here we use the convention that E_k is the conditional expectation at scale 2^{-k}); see e.g. [**JSW08**, Lemma 1.3]. By the greedy selection argument, see e.g. [**JKRW98**, Lemma 6.7], this follows from

$$\left\| \left(\sum_j |E_{t_{j+1}(x)} f(x) - E_{t_j(x)} f(x)|^2 \right)^{1/2} \right\|_{L^p_x(w)} \lesssim \|f\|_{L^p(w)},$$

where $t_1 \le t_2 \le \cdots$ are stopping times and the bound does not depend on the stopping times. By truncation, we may assume that $t_0 = -\infty$ and $t_J = +\infty$ for some J. Writing r_j for the Rademacher functions, the left-hand side can be estimated by

$$\begin{aligned} &\left\| \left\| \sum_j r_j(s) (E_{t_{j+1}(x)} f(x) - E_{t_j(x)} f(x)) \right\|_{L^p_s} \right\|_{L^p_x(w)} \\ &= \left\| \left\| \sum_j r_j(s) \sum_{k=t_j(x)}^{t_{j+1}(x)-1} d_k(x) \right\|_{L^p_x(w)} \right\|_{L^p_s}, \end{aligned}$$

where $d_k = E_{k+1} f - E_k f$. Now the $L^p_x(w)$ norm on the right-hand side can be written as

$$\left\| \sum_k r_{\max\{j : t_j(x) \le k\}}(s) d_k(x) \right\|_{L^p_x(w)}.$$

The function $x \mapsto r_{\max\{j:t_j(x)\le k\}}(s)$ is constant at scale 2^{-k} since t_j are stopping times, so this is bounded by $\|f\|_{L^p(w)}$ uniformly in s in view of the weighted bound for the Haar multipliers (which extends to all $1 < p < \infty$ by extrapolation); see [**Wit00**, Theorem 3.1] for the case $d = 1$ and [**LPR10**, Theorem 1.6] for the general case. Integrating over s, we obtain the claim. □

5. *Weak-type* (1, 1) *bounds*

PROPOSITION 5.1. *Let $1 < q < \infty$, let w be an A^1 weight and let T be a sublinear operator that is bounded on $L^q(w)$ and pointwise bounded by a finite linear combination of operators of the form $\tilde{R}$ for some $1 < r \le q$. Then $T : L^1(\ell^q, w) \to L^{1,\infty}(\ell^q, w)$.*

This allows us to deduce part (2) of Theorem 1.4 from part (1) of that theorem and part (1.6) of Theorem 1.5 from part (1.8) of that theorem.

Proof. The argument is adapted from [**JKRW98**, **JRW03**]. By homogeneity, it suffices to prove that

$$w\left\{ x : \left(\sum_i |Tf_i|^q \right)^{1/q} > 1 \right\} \lesssim \left\| \left(\sum_i |f_i|^q \right)^{1/q} \right\|_{L^1(w)}.$$

We use the Calderón–Zygmund decomposition as in [**FS71**]. Let $F = (\sum_i |f_i|^q)^{1/q}$; then there exist disjoint cubes (denoted by Q) such that $\|F\|_{L^\infty(\mathbb{R} \setminus \bigcup Q)} \le 1$, $\sum_Q |Q| \lesssim \|F\|_1$,

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

and $\|F\|_{L^1(Q)} \lesssim |Q|$. Let

$$g_j(x) = \begin{cases} |Q|^{-1} \displaystyle\int_Q f_j, & x \in Q, \\ f_j(x), & x \notin \bigcup Q \end{cases}$$

and

$$b_j = \sum_Q b_j^Q, \quad b_j^Q(x) = \begin{cases} f_j(x) - |Q|^{-1} \displaystyle\int_Q f_j, & x \in Q, \\ 0, & x \notin \bigcup Q. \end{cases}$$

Let $G = (\sum_j |g_j|^q)^{1/q}$; then $G(x) = F(x)$ for $x \notin \bigcup Q$ and

$$G(x) = \Big(\sum_j \Big||Q|^{-1} \int_Q f_j\Big|^q\Big)^{1/q} \le |Q|^{-1} \int_Q \Big(\sum_j |f_j|^q\Big)^{1/q} = |Q|^{-1}\|F\|_{L^1(Q)} \lesssim 1$$

for $x \in Q$. Hence, $\|G\|_{L^1(w)} \le \|F\|_{L^1(w)}$ and $\|G\|_\infty \lesssim 1$, so we get the required weak-type bound for $(g_j)_j$ from the strong-type (q, q) bound.

Note that

$$w(5Q) \le \int_Q |F| w(5Q)/|Q| \lesssim \int_Q |F| Mw \lesssim \int_Q |F| w,$$

so, with $\tilde{E} := \bigcup_Q 5Q$, it suffices to show that

$$w\Big\{x \notin \tilde{E} : \Big(\sum_i |Tb_i|^q\Big)^{1/q} > 1\Big\} \lesssim \sum_Q w(Q). \tag{5.2}$$

By the hypothesis, it suffices to show that

$$\int_{\mathbb{R}^d \setminus \tilde{E}} \Big(\sum_i \Big(\sum_k (\tilde{R}_k b_i)^r\Big)^{q/r}\Big)^{r/q} w \lesssim \sum_Q w(Q). \tag{5.3}$$

Let $\alpha > (d-1)/r'$ be chosen later. For $x \notin E$, only cubes with $\ell(Q) \lesssim 2^k$ contribute to $\tilde{R}_k$. Thus, by Corollary 2.4, we have

$$(\tilde{R}_k b_i)^r(x) \lesssim 2^{k\alpha r - kdr} \sum_{Q : \operatorname{dist}(x,Q) \lesssim 2^k} \ell(Q)^{-\alpha r} \|b_i^Q\|_1^r.$$

Hence, the left-hand side of (5.3) can be estimated by

$$\int_{\mathbb{R}^d \setminus \tilde{E}} \Big(\sum_i \Big(\sum_k 2^{k(\alpha-d)r} \sum_Q 1_{\operatorname{dist}(\cdot,Q) \lesssim 2^k}(x) \ell(Q)^{-\alpha r} \|b_i^Q\|_1^r\Big)^{q/r}\Big)^{r/q} w(x)\, dx. \tag{5.4}$$

By Minkowski's inequality, this is bounded by

$$\sum_Q \ell(Q)^{-\alpha r} \Big(\sum_i \|b_i^Q\|_1^q\Big)^{r/q} \sum_k 2^{k(\alpha-d)r} \int_{\mathbb{R}^d \setminus \tilde{E}} 1_{\operatorname{dist}(\cdot,Q) \lesssim 2^k}(x) w(x)\, dx.$$

Note that

$$\sum_i \|b_i^Q\|_1^q \le \Big(\int_Q \Big(\sum_i |b_Q^i|^q\Big)^{1/q}\Big)^q \lesssim \|F\|_{L^1(Q)}^q \lesssim |Q|^q \tag{5.5}$$

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

by Minkowski's inequality. Hence, we obtain the bound

$$\sum_Q \ell(Q)^{-\alpha r}|Q|^r \sum_{k:2^k\gtrsim\ell(Q)} 2^{k(\alpha-d)r}\int_{\mathbb{R}^d} 1_{\operatorname{dist}(\cdot,Q)\lesssim 2^k}(x)w(x)\,dx.$$

The latter integral can be estimated by $2^{kd}\min_{x\in Q} Mw(x)$, and the sum over k is finite provided $(\alpha-d)r+d<0$. In this case we obtain the bound

$$\begin{aligned}&\sum_Q \ell(Q)^{-\alpha r}|Q|^r\ell(Q)^{(\alpha-d)r+d}\min_{x\in Q} Mw(x)\\ &\lesssim \sum_Q |Q|^r\ell(Q)^{d(1-r)}\min_{x\in Q} w(x)\\ &\approx \sum_Q |Q|\min_{x\in Q} w(x)\\ &\le \sum_Q w(Q),\end{aligned}$$

as required. Note that the restrictions on α imposed here are equivalent to $(d-1)/r'<\alpha<d/r'$, so an appropriate α can be chosen whenever $r>1$. □

The above proof also yields an A_1 weighted weak-type $(1,1)$ estimate for the jump counting function.

6. *The BMO bound*

Finally we prove (1.7) in Theorem 1.5. Suppose that $\|(\sum_i |f_i|^q)^{1/q}\|_\infty\le 1$ and let Q be a cube. It suffices to show that

$$\fint_Q\Big(\sum_i \|A_t f_i(x)-c_{t,i,Q}\|_{V_t^r}^q\Big)^{1/q}dx\lesssim 1$$

for some choice of functions $c_{t,i,Q}$. We split $f=g+b$, where $g=f1_{3Q}$. For the local part g, we have $\|g\|_{L^q(\ell^q)}\lesssim|Q|^{1/q}$, and it follows from the scalar-valued case of the Bourgain–Lépingle inequality (1.8) that

$$\Big\|\Big(\sum_i \|A_t g_i(x)\|_{V_t^r}^q\Big)^{1/q}\Big\|_{L_x^q}\lesssim\|g\|_{L^q(\ell^q)}.$$

The required estimate for the local part then follows by an application of the Cauchy–Schwarz inequality in the x variable.

It remains to treat the global part b. We set $c_{t,i,Q}=\fint_Q A_t b_i(y)\,dy$. We have

$$\fint_Q\Big(\sum_i \|A_t b_i(x)-c_{t,i,Q}\|_{V_t^r}^q\Big)^{1/q}dx\le\fint_Q\fint_Q\Big(\sum_i \|A_t b_i(x)-A_t b_i(y)\|_{V_t^r}^q\Big)^{1/q}dx\,dy,$$

so it suffices to show that

$$\Big(\sum_i \|A_t b_i(x)-A_t b_i(y)\|_{V_t^r}^q\Big)^{1/q}\lesssim 1 \tag{6.1}$$

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

uniformly in $x, y \in Q$. Estimating the r-variational norm by the $\min(r, q)$-variational norm, we may assume that $r \le q$ for the remaining part of the proof. By Lemma B.1, we have

$$\|A_t b_i(x) - A_t b_i(y)\|_{V_t^r} \lesssim \|A_t b_i(x) - A_t b_i(y)\|_{L_t^r}^{1-1/r} \left\| \frac{d}{dt}(A_t b_i(x) - A_t b_i(y)) \right\|_{L_t^r}^{1/r}.$$

By Cauchy–Schwarz, the left-hand side of (6.1) can be estimated by

$$\left(\sum_i \|A_t b_i(x) - A_t b_i(y)\|_{L_t^r}^q \right)^{(1-1/r)/q} \left(\sum_i \left\| \frac{d}{dt}(A_t b_i(x) - A_t b_i(y)) \right\|_{L_t^r}^q \right)^{(1/r)/q}.$$

By Minkowski's integral inequality, this is bounded by

$$\left\| \left(\sum_i |A_t b_i(x) - A_t b_i(y)|^q \right)^{1/q} \right\|_{L_t^r}^{1-1/r} \left\| \left(\sum_i \left| \frac{d}{dt}(A_t b_i(x) - A_t b_i(y)) \right|^q \right)^{1/q} \right\|_{L_t^r}^{1/r}. \tag{6.2}$$

Note that $A_t b_i(x) = A_t b_i(y) = 0$ if $t < \ell(Q)$. For $t > \ell(Q)$, we write

$$A_t b_i(x) - A_t b_i(y) = \frac{C}{t^d} \int (1_{B(x,t)} - 1_{B(y,t)}) b_i$$

and use the fact that

$$|B(x, t) \Delta B(y, t)| \lesssim |x - y| t^{d-1} \lesssim \ell(Q) t^{d-1}. \tag{6.3}$$

This gives the estimate

$$\begin{aligned} &\left\| \left(\sum_i |A_t b_i(x) - A_t b_i(y)|^q \right)^{1/q} \right\|_{L^r_{t>\ell(Q)}} \\ &\lesssim \left\| t^{-d} \int |1_{B(x,t)} - 1_{B(y,t)}| \left(\sum_i |b_i|^q \right)^{1/q} \right\|_{L^r_{t>\ell(Q)}} \\ &\lesssim \ell(Q) \|t^{-1}\|_{L^r_{t>\ell(Q)}} \lesssim \ell(Q)^{1/r} \end{aligned}$$

for the first factor in (6.2). In the second factor, we split

$$\begin{aligned} \left| \frac{d}{dt}(A_t b_i(x) - A_t b_i(y)) \right| \lesssim\ & t^{-d-1} \left| \int_{B(x,t)\Delta B(y,t)} b_i \right| + t^{-d} \left| \frac{d}{dt} \int_{B(x,t)} b_i \right| \\ &+ t^{-d} \left| \frac{d}{dt} \int_{B(y,t)} b_i \right| \end{aligned}$$

and estimate the contributions of the first and the second summands separately (the third summand is entirely analogous to the second). The contribution of the first summand is estimated as above and gives $\|\cdots\|_{L^r} \lesssim \ell(Q)^{-1+1/r}$. In the second summand, note that

$$\frac{d}{dt} \int_{B(x,t)} b_i \approx \int_{\partial B(x,t)} b_i,$$

and analogously to the above case we again obtain $\|\cdots\|_{L^r} \lesssim \ell(Q)^{-1+1/r}$. Combining these estimates, we see that (6.2) is bounded by an absolute constant, and this concludes the proof.

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

Acknowledgement. The authors thank the Hausdorff Research Institute for Mathematics for hospitality during the Trimester Program 'Harmonic Analysis and Partial Differential Equations'.

A. *Appendix. A good λ inequality*

The good λ inequality below provides an alternative way to deduce part (1) of Theorem 1.4 from the special case $p = 2$, $w \equiv 1$.

PROPOSITION A.1. *Suppose that $\tilde{S} : L^p \to L^{p,\infty}$ (unweighted) for some $1 \le p < \infty$. Let w be a weight satisfying the A_∞ condition (1.3). Then, for each $\lambda > 0$ and $A > 1$, we have*

$$w\{\tilde{S}f > A\lambda,\, M_p^\sharp f \le \gamma\lambda\} \lesssim_p C_w \left(\frac{\gamma}{A-1}\right)^{p\delta} w\{\tilde{S}f > \lambda\}.$$

Proof. We decompose

$$\{\tilde{S}f > \lambda\} = \bigcup_{\mathscr{Q}} Q$$

into a disjoint union of maximal dyadic cubes Q. We have

$$\begin{aligned}\frac{w\{\tilde{S}f > A\lambda,\, M_p^\sharp f \le \gamma\lambda\}}{w\{\tilde{S}f > \lambda\}} &\le \sup_{Q\in\mathscr{Q}} \frac{w\{\tilde{S}f > A\lambda,\, M_p^\sharp f \le \gamma\lambda,\, Q\}}{w(Q)} \\ &\le C_w \sup_{Q\in\mathscr{Q}} \left(\frac{|Q \cap \{\tilde{S}f > A\lambda,\, M_p^\sharp f \le \gamma\lambda\}|}{|Q|}\right)^\delta\end{aligned}$$

by the A_∞ condition. Fix $Q \in \mathscr{Q}$; we need to find a bound for the ratio on the right-hand side. Now, if $M_p^\sharp f > \gamma\lambda$ on Q, then the ratio is zero, so we assume that there exists some $z = z_Q$ such that

$$M_p^\sharp f(z) \le \gamma\lambda.$$

Suppose that $Q \in \sigma_j$. Now, for $x \in Q$, we may express

$$(\tilde{S}f)^2(x) \lesssim \sum_{k\le j} \sum_{Q'\sim\mathscr{Q}_k(x)} S_k f(x_{Q'})^2 + \sum_{k>j} \sum_{Q'\sim\mathscr{Q}_k(x)} S_k f(x_{Q'})^2.$$

Here $Q' \sim \mathscr{Q}_k(x)$ means that the sum is taken over all neighbors Q' of $\mathscr{Q}_k(x)$, the unique dyadic cube with side length 2^k containing x, and $x_{Q'}$ is a point at which the supremum over Q is almost attained.

By maximality of Q, the second term is $\le \lambda^2$. On the other hand, the first term in the sum may be expressed as

$$\sum_{k\le j} \sum_{Q'\sim\mathscr{Q}_k(x)} S_k((f-c)1_{CQ})(x_{Q'})^2 \le (\tilde{S}((f-c)1_{CQ}))^2$$

for any c. Thus,

$$Q \cap \{\tilde{S}f > A\lambda,\, M_p^\sharp f \le \gamma\lambda\} \subset Q \cap \{\tilde{S}((f-c)\cdot 1_{CQ}) > (A-1)\lambda\}.$$

Using the weak-type (p, p) boundedness of $\tilde{S}$, we may estimate the measure of this final set by

$$\frac{1}{(A-1)^p\lambda^p} \|\tilde{S}\|_{L^p\to L^{p,\infty}}^p \int_{CQ} |f-c|^p.$$

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

Minimizing over c, we obtain the estimate

$$\frac{1}{(A-1)^p\lambda^p}|CQ|M_p^\sharp f(z)^p \lesssim \frac{\gamma^p}{(A-1)^p}|Q|. \qquad \square$$

B. *Appendix. A Sobolev embedding theorem for r-variation*

We recall a simple estimate for the r-variational norm of a function on $\mathbb{R}$ in terms of L^r norms. Note that the conclusion does not explicitly depend on T, so the estimate remains true for functions defined on infinite intervals.

LEMMA B.1. *Let X be a normed space, $1 \le r < \infty$, and $a : [0, T] \to X$. Then*

$$\|a\|_{\tilde V^r} \le 8\|a\|_{L^r(X)}^{1-1/r}\|a'\|_{L^r(X)}^{1/r}.$$

This is a direct consequence of interpolation between homogeneous Sobolev spaces [**BL76**, Theorem 6.3.1], embedding of Besov spaces into variation norm spaces [**BP74**], and the estimate for real interpolation norms by the J-functional [**BL76**, Theorem 3.2.2]. We include a short proof for completeness.

Proof. Let $L > 0$ be an integer chosen later and set $D := \{0, T/L, \dots, (L-1)(T/L)\}$. Let also $\delta \in [0, T/L]$ be chosen later. Splitting the variation into a short and a long part, we get

$$\begin{aligned}\|a\|_{\tilde V^r} &\le \|a(t)\|_{\tilde V^r_{t\in\delta+D}} + 2\Bigg(\sum_{j=-1}^{L}\|a(t)\|^r_{\tilde V^r_{t\in[0,T]\cap j(T/L)+\delta+[0,T/L]}}\Bigg)^{1/r}\\ &\le 2\|a(t)\|_{\ell^r_{t\in\delta+D}} + 2\Bigg(\sum_{j=-1}^{L}\|a(t)\|^r_{\tilde V^1_{t\in[0,T]\cap j(T/L)+\delta+[0,T/L]}}\Bigg)^{1/r}.\end{aligned}$$

We have $\|a\|_{L^r}^r = \int_{\delta=0}^{T/L}\|a(t)\|^r_{\ell^r_{t\in\delta+D}}$, so, for some δ, the first summand is bounded by $2(T/L)^{-1/r}\|a\|_{L^r}$. In the second summand, we have estimated the r-variation norm by the 1-variation norm, which can be estimated by the L^1 norm of the derivative. Hence, we obtain

$$\|a\|_{\tilde V^r} \le 2(T/L)^{-1/r}\|a\|_{L^r} + 2\Bigg(\sum_{j=-1}^{L}\|a'\|^r_{L^1_{t\in[0,T]\cap j(T/L)+\delta+[0,T/L]}}\Bigg)^{1/r}.$$

We estimate the L^1 norm on the right-hand side by the L^r norm using the Hölder inequality; this gives

$$\begin{aligned}\|a\|_{\tilde V^r} &\le 2(T/L)^{-1/r}\|a\|_{L^r} + 2(T/L)^{1-1/r}\Bigg(\sum_{j=-1}^{L}\|a'\|^r_{L^r_{t\in[0,T]\cap j(T/L)+\delta+[0,T/L]}}\Bigg)^{1/r}\\ &= 2(T/L)^{-1/r}\|a\|_{L^r} + 2(T/L)^{1-1/r}\|a'\|_{L^r}.\end{aligned}$$

The optimal choice $L = \lfloor T\|a'\|_{L^r}/\|a\|_{L^r}\rfloor$ gives the desired estimate unless $T\|a'\|_{L^r} < \|a\|_{L^r}$. However, in the latter case we have the easy bound

$$\|a\|_{\tilde V^r} \le \|a\|_{\tilde V^1} \le \|a'\|_{L^1} \le T^{1-1/r}\|a'\|_{L^r} \le \|a\|_{L^r}^{1-1/r}\|a'\|_{L^r}^{1/r}. \qquad \square$$

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press

References

[**BL76**] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction (Grundlehren der Mathematischen Wissenschaften, 223)*. Springer, Berlin, 1976.

[**Bou89**] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publ. Math. Inst. Hautes Études Sci.* **69** (1989), 5–45. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.

[**BP74**] J. Bergh and J. Peetre. On the spaces V_p $(0 < p \leq \infty)$. *Boll. Unione. Mat. Ital. (4)* **10** (1974), 632–648.

[**Buc93**] S. M. Buckley. Summation conditions on weights. *Michigan Math. J.* **40**(1) (1993), 153–170.

[**Duo11**] J. Duoandikoetxea. Extrapolation of weights revisited: new proofs and sharp bounds. *J. Funct. Anal.* **260**(6) (2011), 1886–1901.

[**FS71**] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.* **93** (1971), 107–115.

[**GCF85**] J. García-Cuerva and J. L. Rubio de Francia. *Weighted Norm Inequalities and Related Topics,* Vol. 116 *(North-Holland Mathematics Studies. Notas de Matemática [Mathematical Notes], 104)*. North-Holland, Amsterdam, 1985.

[**JKRW98**] R. L. Jones, R. Kaufman, J. M. Rosenblatt and M. Wierdl. Oscillation in ergodic theory. *Ergod. Th. & Dynam. Sys.* **18**(4) (1998), 889–935.

[**JRW03**] R. L. Jones, J. M. Rosenblatt and M. Wierdl. Oscillation in ergodic theory: higher dimensional results. *Israel J. Math.* **135** (2003), 1–27.

[**JSW08**] R. L. Jones, A. Seeger and J. Wright. Strong variational and jump inequalities in harmonic analysis. *Trans. Amer. Math. Soc.* **360**(12) (2008), 6711–6742.

[**Kra13**] B. Krause. On higher-dimensional oscillation in ergodic theory. *Preprint*, 2013. *New York J. Math.* to appear.

[**Lép76**] D. Lépingle. La variation d'ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **36**(4) (1976), 295–316.

[**LPR10**] M. T. Lacey, S. Petermichl and M. C. Reguera. Sharp A_2 inequality for Haar shift operators. *Math. Ann.* **348**(1) (2010), 127–141.

[**MTX15a**] T. Ma, J. L. Torrea and Q. Xu. Weighted variation inequalities for differential operators and singular integrals. *J. Funct. Anal.* **268**(2) (2015), 376–416.

[**MTX15b**] T. Ma, J. L. Torrea and Q. Xu. Weighted variation inequalities for differential operators and singular integrals in higher dimensions. *Preprint*, 2015.

[**PX88**] G. Pisier and Q. H. Xu. The strong p-variation of martingales and orthogonal series. *Probab. Theory Related Fields* **77**(4) (1988), 497–514.

[**Wit00**] J. Wittwer. A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.* **7**(1) (2000), 1–12.

https://doi.org/10.1017/etds.2016.27 Published online by Cambridge University Press