VARIATIONAL INEQUALITIES FOR BILINEAR AVERAGING OPERATORS OVER CONVEX BODIES

YONG DING, GUIXIANG HONG, XINFENG WU

ABSTRACT. We study $q$-variation inequality for bilinear averaging operators over convex bodies $(G_t)_{t>0}$ defined by

$$A_t^G(f_1, f_2)(x) = \frac{1}{|G_t|} \int_{G_t} f_1(x + y_1)f_2(x + y_2) dy_1 dy_2, \quad x \in \mathbb{R}^d,$$

where $G_t$ are the dilates of a convex body $G$ in $\mathbb{R}^{2d}$. We prove that

$$\|V_q(A_t^G(f_1, f_2) : t > 0)\|_{L^p} \lesssim \|f_1\|_{L^{p_1}} \|f_2\|_{L^{p_2}},$$

for $2 < q < \infty$, $1 < p_1, p_2 \leq \infty$, $1/2 < p < \infty$ with $1/p = 1/p_1 + 1/p_2$. The target space $L^p$ should be replaced by $L^{p, \infty}$ for $p_1 = 1$ and/or $p_2 = 1$, and by dyadic BMO when $p_1 = p_2 = \infty$. As applications, we obtain variational inequalities for bilinear discrete averaging operators, bilinear averaging operators of Demeter-Tao-Thiele, and ergodic bilinear averaging operators. As a byproduct, we also obtain the same mapping properties for a new class of bilinear square functions involving conditional expectation, which are of independent interest.

1. Introduction

For $1 \leq q < \infty$, the $q$-variation seminorm (or $V_q$ seminorm) of a family $(a_t(x))_{t>0}$ of complex-valued functions is defined as

$$V_q(a_t(x) : t > 0) = \sup \left\{ \left( \sum_{k \geq 0} |a_{t_{k+1}}(x) - a_{t_k}(x)|^q \right)^{1/q} \right\},$$

where the supremum runs over all increasing sequences $(t_k)_{k \geq 0}$ of positive numbers. The $q$-variation inequality is a crucial tool in studying the pointwise convergence for a family of operators, due to the fact that it immediately implies the pointwise convergence of the underlying family of operators without using the Banach principle via the corresponding maximal inequality. Moreover, the variational inequality is stronger than the maximal inequality, in the following sense:

$$\sup_{t > 0} |a_t(x)| \leq |a_{t_0}(x)| + 2V_q(a_t(x) : t > 0)$$

for any $t_0 > 0$, and hence is of more interest.

The first $q$-variation inequality was proved by Lépingle [37] for martingales. Bourgain [3] proved variational inequalities for the ergodic averages, and directly deduced pointwise

2010 Mathematics Subject Classification. Primary 42B25, 47B38, 47A35, 47D07; Secondary 46E40, 46B20.

Key words and phrases. bilinear averaging operator, bilinear ergodic averaging operator, convex body, $q$-variation.

*Corresponding author.
convergence results. Bourgain’s work [3] has inaugurated a new research direction in harmonic analysis and ergodic theory. In [6, 7, 30, 31, 32], Jones and his collaborators systematically studied variational inequalities for ergodic averages and truncated singular integrals of homogeneous type. Since then many mathematicians established variational inequalities for various operators in ergodic theory and harmonic analysis (see [33, 34, 14, 49, 43, 44, 42, 36, 29, 40, 35, 39, 26, 9, 47, 45, 46, 48, 5, 4, 52], among others).

The purpose of this article is to establish variational inequalities for a class of multinear averaging operators over convex bodies. This is the first $q$-variation inequality in the multi-linear Calderón-Zygmund theory. In the linear case, similar results on averaging operators were studied first by Bourgain [3], and subsequently by Jones et al [30, 32]. The weighted case was investigated by Ma-Torrea-Xu [40, 41]. The dimension-free variational inequalities for averaging operators over convex bodies were recently established by Bourgain-Mirek-Stein-Wróbel in [5]. For relevant maximal inequalities, we refer the reader to the works of Bourgain [1, 2] and Carbery [8].

The present article is concerned with variational inequalities in the multilinear setting. For $t > 0$, let $\tilde{Q}_t = Q_t \times Q_t$ be the cube in $\mathbb{R}^{2d}$ centered at origin of side length $t$. The bilinear averaging operators $A_t^Q$ over cubes $\tilde{Q}$ are just the tensor product of two linear averaging operators $A_t^Q$:

$$A_t^Q(f_1, f_2)(x) = \left( \frac{1}{|Q_t|} \int_{Q_t} f_1(x + y_1) dy_1 \right) \left( \frac{1}{|Q_t|} \int_{Q_t} f_2(x + y_2) dy_2 \right) =: A_t^Q(f_1)(x) \cdot A_t^Q(f_2)(x), \quad x \in \mathbb{R}^d.$$

Applying the trivial estimate

$$V_q(a_t \cdot b_t : t > 0) \leq \left( \sup_{t > 0} |a_t| \right) \cdot V_q(b_t : t > 0) + \left( \sup_{t > 0} |b_t| \right) \cdot V_q(a_t : t > 0),$$

the desired variational inequalities for $A_t^\tilde{Q}$ follow easily from known variational inequalities for $A_t^Q$ (see [31, 26]) and the Hardy-Littlewood maximal inequality. However, for multilinear averaging operators over balls (and more generally, over convex bodies), the variational inequalities are not so simple, and cannot be deduced from results in the linear case. Moreover, in contrast to the case of maximal operators, the $q$-variation seminorms are not monotone, i.e., $a_t \leq b_t$ for all $t > 0$ does not imply $V_q(a_t : t > 0) \leq V_q(b_t : t > 0)$. Hence there is no a priori comparison between $q$-variations of averaging operators over cubes and the ones over balls (or convex bodies), and estimates for one average operator do not imply the same estimates for the other. Therefore, it would be interesting and non-trivial to establish variational inequalities for multilinear averaging operators over convex bodies, which is the main goal of the current paper.

To state our main result, we need to recall some definitions. For notational simplicity, in this article, we shall concentrate on the bilinear case. Let $G$ be a non-empty convex body in $\mathbb{R}^{2d}$, which means that $G$ is a bounded convex open subset of $\mathbb{R}^{2d}$. For $t > 0$ set

$$G_t = \\{(x, y) \in \mathbb{R}^{2d} : t^{-1} \cdot (x, y) \in G\}.$$

For $f, g \in \mathcal{L}^{1}_{loc}(\mathbb{R}^d)$, the bilinear averaging operator associated to $G_t$ is defined by

$$A_t^G(f_1, f_2)(x) = \frac{1}{|G_t|} \int_{G_t} f_1(x + y_1)f_2(x + y_2) dy_1 dy_2, \quad x \in \mathbb{R}^d.$$
The dyadic BMO space, $BMO_d$, is defined via the seminorm:

$$
\|f\|_{BMO_d} \simeq \sup_{Q \text{dyadic}} \inf_{a_Q} \frac{1}{|Q|} \int_Q |f(x) - a_Q| \, dx.
$$

The main result of this article is the following:

**Theorem 1.1.** Let $2 < q < \infty$. Then the following statements are true:

(i) For $1/2 < p < \infty$, $1 < p_1, p_2 \leq \infty$ satisfying $1/p = 1/p_1 + 1/p_2$ there exists a constant $C = C(p_1, p_2, q, d)$ such that for any $f_1 \in L^{p_1}(\mathbb{R}^d)$, $f_2 \in L^{p_2}(\mathbb{R}^d)$,

$$
\|V_q(A^G_t(f_1, f_2) : t > 0)\|_p \leq C\|f_1\|_{p_1}\|f_2\|_{p_2};
$$

For either $p_1 = \infty$ or $p_2 = \infty$, $L^\infty$ should be replaced by $L^\infty_c$.

(ii) For any $\alpha > 0$, $1 \leq p_2 \leq \infty$, $1/2 \leq p \leq 1$ satisfying $1 + 1/p_2 = 1/p$, there exists a constant $C = C(p_2, q, d)$ such that for any $f_1 \in L^1(\mathbb{R}^d)$, $f_2 \in L^{p_2}(\mathbb{R}^d)$,

$$
|\{V_q(A^G_t(f_1, f_2) : t > 0) > \alpha]\} \leq \frac{C}{\alpha^p}\|f_1\|_1^p\|f_2\|_{p_2}^p;
$$

When $p_2 = \infty$, $L^\infty$ should be replaced by $L^\infty_c$.

(iii) If $p_1 = p_2 = \infty$, then there exists a constant $C = C(d)$ such that for any $f_1, f_2 \in L^\infty_c(\mathbb{R}^d)$ (the space of all compactly supported $L^\infty$ functions),

$$
\|V_q(A^G_t(f_1, f_2) : t > 0)\|_{BMO_d} \leq C\|f_1\|_\infty\|f_2\|_\infty.
$$

**Remark 1.1.** (i) In contrast with the fact that linear variational inequalities are Banach valued, Theorem 1.1 contains both Banach and quasi-Banach valued inequalities (that is, the target space exponent $p$ could be smaller than 1). The constraint $q > 2$ coincides with the one in the linear case.

(ii) The operator $V_q(A^G_t)$ is not of restricted weak type $(\infty, \infty, \infty)$, that is, the $L^\infty \times L^\infty \rightarrow L^\infty$ variational inequality fails, even for characteristic functions of measurable sets of finite measures, as will be shown via a counterexample in Section 6 below. In this sense, the $L^\infty \times L^\infty \rightarrow BMO_d$ estimate in part (iii) cannot be strengthened.

Our strategy of the proof can be described as follows. We first establish the $L^\infty \times L^2 \rightarrow L^2$ variation bounds. We consider the long and short variation operators separately. For the long variation, matters are reduced to showing the $L^\infty \times L^2 \rightarrow L^2$ bound for a new bilinear square function:

$$
\mathcal{L}(f_1, f_2)(x) = \left( \sum_{k \in \mathbb{Z}} |A_{2k}(f_1, f_2)(x) - \mathcal{E}_k f_1(x) \cdot \mathcal{E}_k f_2(x)|^2 \right)^{1/2},
$$

where $\mathcal{E}_k f$ is the conditional expectation with respect to the $\sigma$-algebra generated by dyadic cubes of side length $2^k$ in $\mathbb{R}^d$. Such a square function falls outside the multilinear vector-valued Calderón-Zygmund theory studied in literature. To achieve our goal, we shall develop several new tools including bilinear almost orthogonality principles, a paramultiplicativity decomposition involving conditional expectation, uniform Carleson measure estimates, and new bilinear maximal function. These, together with an adaptation of the geometric arguments in [31] to the multilinear setting, allow us to conclude the $L^\infty \times L^2 \rightarrow L^2$ variational estimate.
Next, based on the $L^\infty \times L^2 \to L^2$ variation estimate, we further prove the $L^\infty \times L^\infty \to BMO_d$ estimate, which naturally extends the $L^\infty \to BMO_d$ estimate for linear averages in [26]. We also show, by constructing a counterexample, that the stronger $L^\infty \times L^\infty \to L^\infty$ variation estimates fails even for characteristic functions of measurable subsets of finite measures. Regarding the $L^1$ endpoints, we prove the $L^1 \times L^{p_2} \to L^{p, \infty}$ variation estimates for all $1 \leq p_2 \leq \infty$ by using a variant of multi-linear Calderón-Zygmund decomposition in [19].

Combining the above mentioned estimates, we obtain that the $q$-variation operator is of restricted weak types $(1, \infty, 1)$, $(\infty, 1, 1)$, $(1, 1, 1/2)$, $(\infty, s, s)$ and $(s, \infty, s)$, for any $2 < s < \infty$. Hence the proof of Theorem 1.1 are concluded by applying the multi-linear real interpolation of Grafakos-Kalton in [21] (see also [20]). As a byproduct of our arguments, we obtain the same estimates for the bilinear square operator $\mathcal{L}$, which are of independent interest (see Section 7 below).

Finally, we point out that the techniques used here can be extended to the multi-linear context as well as to the weighted setting, and are expected to be useful for other operators such as the multilinear truncated singular integrals studied by Grafakos-Torres [23]. We also mention that in [15] Do, Oberlin, and Palsson considered a different class of bilinear averaging operators, which are closely related to the bilinear Hilbert transform. Similar bilinear variational inequalities were proved for $2/3 < p < \infty$ by using techniques of time-frequency analysis, which are quite different from ours.

This article is organized as follows. In Section 2, we prove some lemmas which will be used in the proof of Theorem 1.1. Section 3 is devoted to establishing the $L^\infty \times L^2 \to L^2$ bound. We further prove the $L^\infty \times L^\infty \to BMO_d$ and $L^1 \times L^{p_2} \to L^{p, \infty}$ variational estimates in Sections 4 and 5 respectively. In Section 6, we first discuss the restricted weak type estimates and then conclude the proof of Theorem 1.1. The same bounds for the bilinear square function are established in Section 7. Finally, applications to various bilinear averaging operators are presented in Section 8.

2. Preliminaries and some lemmas

In this section, we give some lemmas that will be used in the proof of Theorem 1.1.

Lemma 2.1 (Bilinear almost orthogonality principle). Let $(S_n)_{n \in \mathbb{Z}}$ be a sequence of operators, and let $(\sigma(j))_{j \in \mathbb{Z}}$ be a sequence of positive numbers with $w := \sum_j \sigma(j) < \infty$.

(i) If $\|S_k(u_n, v_n)\|_2 \lesssim \sigma(n - k)a(u_n, v_n)$ for any $n, k \in \mathbb{Z}$, then

$$\sum_k \left\| \sum_n S_k(u_n, v_n) \right\|_2^2 \leq w \cdot \sum_n a(u_n, v_n)^2.$$  

(ii) If $\|S_k(u_n, v_n)\|_2 \leq \sigma(n - k)a_k(u_n, v_n)$ for any $n, k \in \mathbb{Z}$, then

$$\sum_k \left\| \sum_n S_k(u_n, v_n) \right\|_2^2 \leq w \cdot \sup_n \sum_k a_k(u_{k-n}, v_{k-n})^2.$$
Proof. We prove (i) first. By the triangle inequality in $L^2$, 
\[
\left\| \sum_n S_k(u_n, v_n) \right\|_2 \leq \sum_n \left\| S_k(u_n, v_n) \right\|_2 \leq \sum_n \sigma(k - n)a(u_n, v_n).
\]

The desired estimate then follows by taking the $\ell^2$ norm, and applying Young’s inequality for (discrete) convolution of sequences
\[
\|a *_d b\|_{\ell^2} \leq \|a\|_{\ell^2}\|b\|_{\ell^1},
\]
with $a = (a(u_n, v_n))_{n \in \mathbb{Z}}$ and $b = (\sigma(n))_{n \in \mathbb{Z}}$, where $*_d$ means convolution with respect to the counting measure.

The proof of (ii) is similar. By our hypothesis and triangle inequality,
\[
\left\| \sum_n S_k(u_n, v_n) \right\|_2 \leq \sum_n \left\| S_k(u_n, v_n) \right\|_2 \leq \sum_n \sigma(k - n)\|a_k(u_n, v_n)\|_2.
\]

Using Minkowski’s inequality, we have
\[
\sum_k \left\| \sum_n S_k(u_n, v_n) \right\|_2^2 \leq \sum_k \left( \sum_n \sigma(k - n)\|a_k(u_n, v_n)\|_2 \right)^2
\]
\[
= \sum_k \left( \sum_{n'} \sigma(n')\|a_k(u_{k-n'}, v_{k-n'})\|_2 \right)^2
\]
\[
\leq \left( \sum_{n'} \sigma(n') \cdot \left( \sum_k \|a_k(u_{k-n'}, v_{k-n'})\|_2^2 \right)^{1/2} \right)^2
\]
\[
\leq w^2 \cdot \sup_n \sum_k \|a_k(u_{k-n}, v_{k-n})\|_2^2,
\]
concluding the proof. \qed

For any cube in $\mathbb{R}^d$, denote by $\ell(Q)$ the side-length of $Q$. For $j \in \mathbb{Z}$, denote by $\mathcal{D}_j$ (resp. $\mathcal{D}_j$) the set of all dyadic cubes of side-length $2^j$ in $\mathbb{R}^d$ (resp. $\mathbb{R}^{2d}$). The conditional expectation of a locally integrable function $f$ with respect to the increasing family of $\sigma$-algebras generated by $\mathcal{D}_j$ is defined by
\[
\mathcal{E}_j f(x) = \sum_{Q \in \mathcal{D}_j} \left( \frac{1}{|Q|} \int_Q f(y)dy \right) \cdot 1_Q(x), \quad \forall j \in \mathbb{Z}.
\]

The dyadic martingale difference operator $d_j$ is defined by $d_j(f) = \mathcal{E}_{j-1} f - \mathcal{E}_j f$.

The following result is a paraproduct type decomposition involving conditional expectation.

Lemma 2.2. For $k \in \mathbb{Z}$ and $f_1, f_2 \in L^2$, define bilinear operators $\mathcal{L}_k$ by
\[
\mathcal{L}_k(f_1, f_2)(x) = A_{2^k}(f_1, f_2)(x) - \mathcal{E}_{k} f_1(x) \cdot \mathcal{E}_{k} f_2(x).
\]

Then, for almost every $x \in \mathbb{R}^d$,
\[
\mathcal{L}_k(f_1, f_2)(x) = \sum_{n \in \mathbb{Z}} \mathcal{L}_k(d_{1,n}, \mathcal{E}_{n-1} f_2)(x) + \sum_{n \in \mathbb{Z}} \mathcal{L}_k(\mathcal{E}_n f_1, d_{2,n})(x).
\]
Proof. By the Cauchy-Schwarz inequality, for any given \( x \in \mathbb{R}^d \),
\[
|\mathcal{E}_j f_i(x)| \leq |\mathcal{E}_j |f_i|^2(x)|^{\frac{1}{2}} \leq 2^{-jd\bar{d}} \|f_i\|_2 \to 0, \quad i = 1, 2,
\]
as \( j \to +\infty \). Moreover, for any \((y_1, y_2) \in \mathbb{R}^{2d}\) and any \( x \in \mathbb{R}^d \),
\[
\sup_j |\mathcal{E}_j f_i(y_1)\mathcal{E}_j f_2(y_2)| \leq \mathcal{M}(f_1)(y_1)\mathcal{M}(f_2)(y_2) \in L^1((x, x) + G_{2^k}).
\]
It follows by the Lebesgue dominated convergence theorem that
\[
\lim_{j \to +\infty} \mathcal{L}_k(\mathcal{E}_j f_1, \mathcal{E}_j f_2)(x) = 0, \quad \text{a.e. } x \in \mathbb{R}^d.
\]
On the other hand, Lebesgue differentiation theorem yields \( \mathcal{E}_j f_i(x) \to f(x) \) a.e. as \( j \to -\infty \), which, by the Lebesgue dominated convergence theorem, implies
\[
\mathcal{L}_k(f_1, f_2)(x) = \lim_{j \to -\infty} \mathcal{L}_k(\mathcal{E}_j f_1, \mathcal{E}_j f_2)(x), \quad \text{a.e. } x \in \mathbb{R}^d.
\]
Hence, by the bi-linearity of \( \mathcal{L}_k \), we can express \( \mathcal{L}_k(f_1, f_2) \) as
\[
\mathcal{L}_k(f_1, f_2)(x) = \lim_{l \to -\infty} \mathcal{L}_k(\mathcal{E}_{l-1} f_1, \mathcal{E}_{l-1} f_2)(x) - \lim_{j \to +\infty} \mathcal{L}_k(\mathcal{E}_j f_1, \mathcal{E}_j f_2)(x)
\]
\[
= \lim_{j \to +\infty} \sum_{n=l}^j \left[ \mathcal{L}_k(\mathcal{E}_{n-1} f_1, \mathcal{E}_{n-1} f_2)(x) - \mathcal{L}_k(\mathcal{E}_n f_1, \mathcal{E}_n f_2)(x) \right]
\]
\[
= \lim_{j \to +\infty} \sum_{n=l}^j \left[ \mathcal{L}_k(\mathcal{E}_{n-1} f_1, \mathcal{E}_{n-1} f_2)(x) - \mathcal{L}_k(\mathcal{E}_n f_1, \mathcal{E}_n f_2)(x) \right]
\]
\[
+ \mathcal{L}_k(\mathcal{E}_n f_1, \mathcal{E}_n f_2)(x) - \mathcal{L}_k(\mathcal{E}_n f_1, \mathcal{E}_n f_2)(x) \right]
\]
\[
= \sum_{n \in \mathbb{Z}} \mathcal{L}_k(d_{1,n}, \mathcal{E}_{n-1} f_2)(x) + \sum_{n \in \mathbb{Z}} \mathcal{L}_k(\mathcal{E}_n f_1, d_{2,n})(x),
\]
where all series converge for a.e. \( x \in \mathbb{R}^d \).

Lemma 2.3. Let \( b \in BMO(\mathbb{R}^d) \), \( \delta_{2^k}(t) \) be the Dirac mass at the point \( t = 2^k \), and let \( \zeta(x) = (1 + |x|)^{-d-\epsilon} \) for some \( \epsilon > 0 \). Then
(i) \( (d\mu_n(x, t))_{n \in \mathbb{N}} = (\sum_{k \in \mathbb{Z}} |\mathcal{E}_{k+1-n} b(x) - \mathcal{E}_{k-n} b(x)|^2 \, dx \delta_{2^k}(t))_{n \in \mathbb{N}} \) is a sequence of Carleson measures on \( \mathbb{R}^{d+1}_+ \) with norm bounded by \( C\|b\|_{BMO}^2 \) uniformly in \( n \).
(ii) \( (d\nu_n(x, t))_{n \in \mathbb{N}} = \left( \sum_{k \in \mathbb{Z}} (\zeta_k * |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2)(x) \, dx \delta_{2^k}(t) \right)_{n \in \mathbb{N}} \) is a sequence of Carleson measures on \( \mathbb{R}^{d+1}_+ \) with norm bounded by \( C\|b\|_{BMO}^2 \) uniformly in \( n \). Here \( \zeta_k(x) = 2^{kd} \zeta(2^{-k}x) \).

Proof. For a cube \( Q \) in \( \mathbb{R}^d \) we let \( Q^* \) be the cube with the same center and orientation whose side-length is \( 100\sqrt{d} \ell(Q) \), where \( \ell(Q) \) denotes the side-length of \( Q \). Fix a cube \( Q \) in \( \mathbb{R}^d \), split \( b \) as
\[
b = (b - b_{2Q}) \mathbf{1}_{Q^*} + (b - b_{2Q}) \mathbf{1}_{(Q^*)^c} + b_{2Q},
\]
where \( b_{2Q} = |2Q|^{-1} \int_{2Q} b(y) \, dy \). Let \( T(Q) = Q \times (0, \ell(Q)) \). Since \( \mathcal{E}_{k+1-n} b_{2Q} = \mathcal{E}_{k-n} b_{2Q} = b_{2Q} \),
\[
\mathcal{E}_{k+1-n} b_{2Q} - \mathcal{E}_{k-n} b_{2Q} = 0.
\]
Thus,
\[
\mu(T(Q)) := \sum_{2^k \leq \ell(Q)} \int_Q |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2 \, dx \leq 2I_1 + 2I_2,
\]
where
\[
I_1 = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^d} |\mathcal{E}_{k+1}[(b-b_{2Q})1_Q^c](x) - \mathcal{E}_k[(b-b_{2Q})1_Q^c](x)|^2 \, dx
\]
and
\[
I_2 = \sum_{2^k \leq \ell(Q)} \int_Q |\mathcal{E}_{k+1-n}[(b-b_{2Q})1_Q^c](x) - \mathcal{E}_{k-n}[(b-b_{2Q})1_Q^c](x)|^2 \, dx.
\]

By the $L^2$ boundedness of martingale square function (see e.g. [33]), we have
\[
I_1 \lesssim \int_{Q^c} |b(x) - b_{2Q}|^2 \, dx \lesssim |Q| \|b\|_{BMO}^2,
\]
where the implicit constant is independent of $n$.

Next, let us show that $I_2 = 0$. Indeed, for any $x \in Q$ and for any $k$ with $2^k \leq \ell(Q)$,
\[
\mathcal{E}_{k+1-n}[(b-b_{2Q})1_Q^c](x) = \frac{1}{|Q(k-n,x)|} \int_{Q(k-n,x) \cap (Q^c)^c} (b(y) - b_{2Q}) \, dy,
\]
where $Q(k-n,x)$ is the unique dyadic cube in $\mathbb{R}^d$ of side-length $2^{k+1-n}$ that contains $x$. For all $n \in \mathbb{N}$, $\ell(Q(k-n,x)) \leq 2\ell(Q)$, which implies $Q(k-n,x) \cap (Q^c)^c = \emptyset$. Hence $I_2 = 0$ for all $n \in \mathbb{N}$.

Altogether, we obtain
\[
\mu(T(Q)) \leq C|Q| \|b\|_{BMO}^2,
\]
where $C$ is independent of $n$. This means $(\mu_n)_{n \in \mathbb{N}}$ is a collection of Carleson measures with norms at most $C\|b\|_{BMO}^2$, uniformly in $n$.

Let us now prove part (ii). We need to show that, for any cube $Q = Q(x_0, \ell(Q))$ and any $n \in \mathbb{N}$,
\[
\int_Q \sum_{2^k \leq \ell(Q)} (\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2)(x) \, dx \lesssim \|b\|_{BMO}^2,
\]
uniformly in $Q$ and $n \in \mathbb{N}$. Define
\[
F(y, 2^k) = \int_{\mathbb{R}^d} \zeta_k(y-z)1_Q(z) \, dz, \quad \forall y \in \mathbb{R}^d, \, k \in \mathbb{Z};
\]
\[
F^*(x) = \sup_{(y,k): |x-y| \leq 2^k} |F(y, 2^k)|, \quad \forall y \in \mathbb{R}^d.
\]
Clearly $F(y, 2^k) \leq \|\zeta\|_{L^1}$. Moreover, for $|y-x_0| \geq 2\sqrt{d} \ell(Q)$ and $z \in Q$, the triangle inequality implies $|y-z| \geq |y-x_0| - |z-x_0| \geq |y-x_0|/2$; therefore
\[
F(y, 2^k) \leq C \frac{\ell(Q)^d}{2^{kd}(2^{k})^{d+\epsilon}} \leq C \left( \frac{|y-x_0|}{\ell(Q)} \right)^{-d-\epsilon} \leq C \left( \frac{|x-x_0|}{\ell(Q)} \right)^{-d-\epsilon}
\]
for all \( x \) satisfying \(|x - y| \leq 2^k \), where the constants \( C \) is independent of \( k \). Hence \( F^*(x) \leq C(1 + \frac{|x - x_0|}{(Q)} )^{-d-\varepsilon} \), and therefore \( F^* \in L^1(\mathbb{R}^d) \). By part (i) and the Carleson inequality (see [16] or [51]), we have
\[
\int_Q \sum_{2^k \leq |Q|} (\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2)(x) \, dx \lesssim \|b\|^2_{BMO} \|F^*\|_{L^1} \lesssim \|b\|^2_{BMO},
\]
uniformly in \( Q \) and \( n \).

For a locally integrable function \( f \) on \( \mathbb{R}^d \), the uncentered Hardy-Littlewood maximal function \( \mathcal{M}(f) \) is defined as
\[
\mathcal{M}(f)(x) = \sup_B \frac{1}{|B|} \int_B f(y) \, dy,
\]
where the supremum is taken over all balls \( B \) in \( \mathbb{R}^d \) which contain \( x \).

The following lemma establishes a uniform Carleson measure estimate, which can be viewed as a martingale version of Grafakos-Miyachi-Tomita’s result in [22] and will be a crucial tool for our proof of the theorem.

**Lemma 2.4.** Let \( 1 < l < 2 \) and \( \zeta \) be defined as above. Then there exists a constant \( C \) depending on \( d, \zeta \) and \( l \), but not on \( n \), such that
\[
\sum_k \int_{\mathbb{R}^d} (\zeta_k \ast |f|^l)^{2/l}(x) (\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^l)(x) \, dx \leq C\|f\|^2 \|\mathcal{M}(f)\|^2_{BMO}, \quad \forall n \in \mathbb{N}.
\]

**Proof.** First, the Hölder inequality gives
\[
(\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^l)^{2/l} \leq \left( \|\zeta\|^1_{L^1} [\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2(x)]^{1/2} \right)^{2/l} \lesssim (\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2)(x),
\]
which implies
\[
\sum_k \int_{\mathbb{R}^d} (\zeta_k \ast |f|^l)^{2/l}(\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^l)^{2/l} \, dx \lesssim \sum_k \int_{\mathbb{R}^d} (\zeta_k \ast |f|^l)^{2/l}(\zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2)(x) \, dx.
\]
By Lemma 2.3 (ii), \( \sum_{k \in \mathbb{Z}} \zeta_k \ast |\mathcal{E}_{k+1-n} b - \mathcal{E}_{k-n} b|^2(x) \, dx \delta_t(t) \) is a sequence of Carleson measures on \( \mathbb{R}^{d+1}_+ \) with norm bounded by \( C\|b\|^2_{BMO} \), uniformly in \( n \). By Carleson’s inequality and the \( L^2 \) boundedness of \( \mathcal{M} \), the last term above is dominated by a constant multiple of
\[
\|b\|^2_{BMO} \int_{\mathbb{R}^d} \sup_{|z-x| \leq 2^k} [(\zeta_k \ast |f|^l)(x)]^{2/l} \, dx \lesssim \|b\|^2_{BMO} \|\mathcal{M}(f^l)\|_{L^{2/l}}^{2/l} \lesssim \|b\|^2_{BMO} \|f\|_{L^2}^2,
\]
where the implicit constant does not depend on \( n \). This completes the proof. \( \square \)

The following easy variational inequality was stated in [10, pp. 13-14].

**Lemma 2.5.** [10]. For \( q > 2 \), we have
\[
\|V_q^j(\mathcal{E}_j f \ast \mathcal{E}_j f) : j \in \mathbb{Z}\|_{L^2(\mathbb{R}^d)} \lesssim \min(\|f_1\|_{L^2(\mathbb{R}^d)} \|f_2\|_{L^\infty(\mathbb{R}^d)}, \|f_1\|_{L^\infty(\mathbb{R}^d)} \|f_2\|_{L^2(\mathbb{R}^d)}).
\]
We now introduce a new maximal function for $k$-measurable functions, which will be used to pointwise dominate the bilinear averages. Assume that $h_1$ and $h_2$ are $(n - 1)$-measurable functions on $\mathbb{R}^d$, which means that they are constant on each atom $Q \in \mathcal{D}_{n-1}$. For every $Q \in \mathcal{D}_{n-1}$, denote by $M_i^Q$, $i = 1, 2$, the maximum of $|h_i|$ on $Q$ and the cubes neighbouring $Q$ (i.e., contained in $3Q$). Define the maximal functions $h_i^*$ on $\mathbb{R}^d$ by setting $h_i^*(x) = M_i^Q$, where $Q$ is the unique atom in $\mathcal{D}_{n-1}$ containing $x$. It was observed in [32] that the maximal function $h_i^*$ dominates the linear ball averages $A_{2k}(h_i)$ if $k \leq n - 1$, and is bounded on $L^2$: 

$$\int_{\mathbb{R}^d} [h_i^*]^2 \lesssim \int_{\mathbb{R}^d} h_i^2.$$

In the bilinear contexts, it is easy to see that $|A_{2k}^G(h_1, h_2)|$ is smaller than $h_1^*(x) \cdot h_2^*(x)$ when $k < n$ (see (3.1) below). However, the desired $L^2$ bound

$$\int_{\mathbb{R}^d} (h_1^*)^2 \cdot (h_2^*)^2 \lesssim \int_{\mathbb{R}^d} |h_1|^2 \cdot |h_2|^2$$

fails. Indeed, in 1-dimensional case, let $I_1, I_2$ be two adjacent dyadic intervals, suppose $h_1 = 1$ in $I_2$, and $h_1 = 0$ elsewhere, and let $h_2 = 1$ on $I_1$ and $h_2 = 0$ elsewhere. Then it is easy to see that the above inequality fails.

To fix this issue, we observe that the following inequalities hold

$$\int_{\mathbb{R}^d} [\cdot [h_1 \cdot |h_2|]^2 \lesssim \int_{\mathbb{R}^d} [h_1 \cdot |h_2|]^2 \lesssim \int_{\mathbb{R}^d} [h_1 \cdot |h_2|]^2,$$

$$\int_{\mathbb{R}^d} [\cdot |h_2|^2 \lesssim \int_{\mathbb{R}^d} [h_1 \cdot |h_2|^2 \lesssim \int_{\mathbb{R}^d} [h_1 \cdot |h_2|^2].$$

This observation leads to the following definition of bi-sublinear maximal function:

$$[h_1, h_2]^+(x) = \max\{ (h_1^* \cdot |h_2|)^*, (h_1 \cdot h_2^*)^* \}.$$ 

Then

$$(2.2) \int_{\mathbb{R}^d} \{ [h_1, h_2]^+ \}^2 \lesssim \int_{\mathbb{R}^d} |h_1|^2 \cdot |h_2|^2.$$

Moreover, this maximal function dominates the bilinear averages, which is given by

**Lemma 2.6.** Let $k < n$. Assume that $h_1$ and $h_2$ are $(n - 1)$-measurable functions on $\mathbb{R}^d$, which means that they are constant-valued on each atom $Q \in \mathcal{D}_{n-1}$. Then

$$|A_{2k}^G(h_1, h_2)(x)| \leq [h_1, h_2]^+(x), \quad \forall x \in \mathbb{R}^d.$$

**Proof.** Let $k \leq n - 1$. For any $x \in \mathbb{R}^d$, by (3.1), $(x, x) + G_{2k} \subset (x, x) + \tilde{B}_{2k}$. Therefore, $(x, x) + G_{2k}$ can intersect with at most $2^{2d}$ atoms in $\mathcal{D}_{n-1}$. Recall that $\mathcal{D}_j$ denotes the set of all dyadic cubes in $\mathbb{R}^{2d}$ of side-length $2^j$. We only deal with the case where $(x, x) + G_{2k}$ intersects with exactly $2^{2d}$ atoms in $\mathcal{D}_{n-1}$, as the other cases are easier and can be treated in the same way. We denote these $2^{2d}$ atoms by $\tilde{Q}_{(I, J)}$, where $(I, J) = (i_1, \ldots, i_d, j_1, \ldots, j_d) \in \mathbb{Z}^d \times \mathbb{Z}^d$ denotes the coordinates of their lower-left corner. Without loss of generality, we may assume that $n = 1$ and the lower-left corner of the union of these atoms is the origin so that $i_k, j_\ell \in \{0, 1\}$ for all $k, \ell = 1, \ldots, d$. In other words, $(I, J) \in \{0, 1\}^d \times \{0, 1\}^d$. 
Writing

\[ |A^G_{\eta}(h_1, h_2)(x)| \leq \frac{1}{|G_{\eta}|} \sum_{I,J \in \{0,1\}^d} \int_{[(x,x)+G_{\eta}] \cap \tilde{Q}_{I,J}} |h_1(y_1)h_2(y_2)| dy_1 dy_2, \]

we see that it suffices to show the pointwise bound: For every \((I, J) \in \{0,1\}^d\),

\[ |h_1(y_1)h_2(y_2)| \leq [h_1, h_2]^+(x), \quad \forall (y_1, y_2) \in \tilde{Q}_{(I,J)} = Q_I \times Q_J. \]

A key observation is that, for any \(I, J \in \{0,1\}^d\), either \(Q_I = Q_J\) or \(Q_I\) neighbours \(Q_J\); so, in either case, we have \(|h(z)| \leq h^*(w)\) for all \((n-1)\)-measurable functions \(h\) and all \(z \in Q_I\) and \(w \in Q_J\). Hence, for any \((I, J) \in \{0,1\}^d\) and \((y_1, y_2) \in \tilde{Q}_{(I,J)} = Q_I \times Q_J,

\[ |h_1(y_1)h_2(y_2)| \leq (|h_1| \cdot h_2^*)(y_1) \leq (|h_1| \cdot h_2^*)(x) \leq [h_1, h_2]^+(x), \]

concluding the proof. \(\Box\)

Let \(S(\mathbb{R}^d)\) denote the space of functions of the form \(\sum_{i=1}^N c_i \mathbb{1}_{E_i}\), where each measurable subset \(E_i\) of \(\mathbb{R}^d\) has finite measure. The following lemma is a multilinear version of Marcinkiewicz interpolation theorem with initial restricted weak type conditions (cf. [21, Theorem 4.6]), which will be used to prove Theorem 1.1.

**Lemma 2.7.** [21, 20] Let \(m\) be a positive integer, and let \(T\) be a multi-sublinear operator defined on \(S(\mathbb{R}^d) \times \cdots \times S(\mathbb{R}^d)\) taking values in the set of measurable functions on \(\mathbb{R}^d\). For \(1 \leq k \leq m+1\) and \(1 \leq j \leq m\) we are given \(p_{k,j}\), with \(0 < p_{k,j} \leq \infty\) and \(0 < q_k \leq \infty\). Suppose that the open convex hull of the points

\[ \bar{P}_k = \left( \frac{1}{p_{k,1}}, \ldots, \frac{1}{p_{k,m}} \right) \]

is an open set in \(\mathbb{R}^d\), and \(T\) is of restricted weak types \((p_{k,1}, \ldots, p_{k,m}, q_k)\) for all \(1 \leq k \leq m+1\), that is,

\[ \|T(\mathbb{1}_{E_1}, \ldots, \mathbb{1}_{E_m})\|_{L^{q_k,\infty}} \leq \theta_k \prod_{j=1}^m |E_j|^{\frac{1}{p_{k,j}}} \]

for all \(1 \leq k \leq m+1\) and for all subsets \(E_j\) of \(\mathbb{R}^d\) with \(|E_j| < \infty\). Let

\[ \bar{P} = \left( \frac{1}{p_1}, \ldots, \frac{1}{p_m} \right) = \sum_{k=1}^{m+1} \eta_k \bar{P}_k \]

for some \(\eta_k \in (0,1)\) such that \(\sum_{k=1}^{m+1} \eta_k = 1\), and define \(1/q = \sum_{k=1}^{m+1} \eta_k/q_k\). If \(\gamma_j \neq 0\) for all \(j = 1, \ldots, m\) and \(1/q \leq 1/p_1 + \cdots + 1/p_m\), then \(T\) is bounded from \(L^{p_1} \times \cdots \times L^{p_m}\) to \(L^q\), and moreover,

\[ \|T(f_1, \ldots, f_m)\|_{L^q} \lesssim \left( \prod_{k=1}^{m+1} \theta_k^{\eta_k} \right) \left( \prod_{j=1}^m \|f_j\|_{L^{p_j}} \right) \]

for all \(f_j \in L^{p_j}, j = 1, \ldots, m\).
3. $L^\infty \times L^2 \to L^2$ Estimate

Since $V_q(A^G_t(f_1, f_2) : t > 0) = V_q(A^{G_{t_0}}_t(f_1, f_2) : t > 0)$ for any $t_0 > 0$, we may assume throughout that $G$ is normalized such that
\[ (3.1) \quad \tilde{B}_\tau \subset G \subset \tilde{B}_1 \quad \text{for some } 0 < \tau < 1. \]

We shall show that the $L^\infty \times L^2 \to L^2$ estimate holds when $f_1 \in L^\infty_c$ and $f_2 \in L^2$. To proceed, we first divide the $q$-variation into long and short variations. Let $(t_i)_{i \geq 0}$ be an increasing sequence. For each interval $I_i = (t_i, t_{i+1}]$, we consider two cases.

- Case 1: $I_i$ does not contain any power of 2;
- Case 2: $I_i$ contains powers of 2. In this case, we decompose $I_i$ further as disjoint union of intervals like $(t_i, 2^{m_i}] \cup (2^{m_i}, 2^{n_i}] \cup (2^{n_i}, t_{i+1}]$ such that $(t_i, 2^{m_i}]$ or $(2^{n_i}, t_{i+1}]$ do not contain any more powers of 2.

Let $\mathcal{S}$ be the set of all “short intervals” consisting of all intervals that is contained in $(2^k, 2^{k+1}]$ for some $k \in \mathbb{Z}$, that is the intervals in Case 1 and the ones of the form $(t_i, 2^{m_i}]$ or $(2^{n_i}, t_{i+1}]$ in Case 2. Let $\mathcal{L}$ consist of all disjoint intervals of the form $(2^{m_j}, 2^{n_j}]$ in Case 2.

Clearly, $\mathcal{S} \cup \mathcal{L}$ is a disjoint family of intervals. We use the following pointwise inequality:
\[
\left( \sum_{i \geq 0} |(A^{G}_{t_{i+1}} - A^{G}_{t_{i}})(f_1, f_2)|^q \right)^{\frac{1}{q}} 
\leq \sup_{(t_i)_i} \left( \sum_{i \in \mathcal{S}} |(A^{G}_{t_{i+1}} - A^{G}_{t_{i}})(f_1, f_2)|^q \right)^{\frac{1}{q}} + \sup_{(t_i)_i} \left( \sum_{i \in \mathcal{L}} |(A^{G}_{t_{i+1}} - A^{G}_{t_{i}})(f_1, f_2)|^q \right)^{\frac{1}{q}} 
\leq V_q(A^{G}_{2^k}(f_1, f_2) : k \in \mathbb{Z}) + \left( \sum_{k \in \mathbb{Z}} V_q(A^{G}_{t_{i}}(f_1, f_2) : t \in (2^k, 2^{k+1}]) \right)^{\frac{1}{q}} 
=: \mathcal{L}V_q(A^{G}_{t_{i}}(f_1, f_2)) + \mathcal{S}V_q(A^{G}_{t_{i}}(f_1, f_2)).
\]

We call $\mathcal{L}V_q(A^{G}_{t_{i}})$ and $\mathcal{S}V_q(A^{G}_{t_{i}})$ long and short variation operators respectively.

The bounds of $\mathcal{L}V_q(A^{G}_{t_{i}})$ and $\mathcal{S}V_q(A^{G}_{t_{i}})$ will be proved in the following two subsections.

3.1. Long variation estimates. To prove the long variation estimates, we shall compare the bilinear averaging operators with conditional expectation, which reduces matters to bounding a certain bilinear square function. More specifically, define a bilinear operator $L_k$ by
\[ L_k(f_1, f_2)(x) = A^{G}_{2^k}(f_1, f_2)(x) - \mathcal{E}_k f_1(x) \cdot \mathcal{E}_k f_2(x), \]
and the square operator $\mathcal{L}$ by
\[ \mathcal{L}(f_1, f_2)(x) = \left( \sum_{k \in \mathbb{Z}} |L_k(f_1, f_2)(x)|^2 \right)^{\frac{1}{2}}. \]

Then the following estimate holds pointwise
\[ \mathcal{L}V_q(A^{G}_{t_{i}})(f_1, f_2) \lesssim \mathcal{L}(f_1, f_2) + \sup_{(n_i)_i} \left( \sum_{i} |(\mathcal{E}_{n_{i+1}} - \mathcal{E}_{n_{i}})(f_1, f_2)|^q \right)^{\frac{1}{q}}. \]
The second term is just the $q$-variation for martingales, for which the desired bound follows from (1.1) and the known $q$-variation inequalities for martingales (cf. [50]). It suffices to establish
\[ \| L(f_1, f_2) \|_2 \lesssim \| f_1 \|_{L^\infty} \| f_2 \|_2. \]

Since $f_1, f_2 \in L^2(\mathbb{R}^d)$, we may write $f_1 = \sum_{n \in \mathbb{Z}} d_{1,n}$ and $f_2 = \sum_{m \in \mathbb{Z}} d_{2,m}$, where $d_{1,n} = d_n(f_1)$ and $d_{2,n} = d_n(f_2)$ are martingale differences and both series converge almost everywhere and in the topology of $L^2$. Moreover,
\[ \sum_n \| d_{1,n} \|_2^2 = \| f_1 \|_2^2, \quad \sum_n \| d_{2,n} \|_2^2 = \| f_2 \|_2^2. \]

Using Lemma 2.2, it thus suffices to show
\[ \sum_{k \in \mathbb{Z}} \| \sum_{n+k} L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2^2 \lesssim \| f_1 \|_{L^\infty}^2 \| f_2 \|_2^2; \]
\[ \sum_{k \in \mathbb{Z}} \| \sum_{n-k} L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2^2 \lesssim \| f_1 \|_{L^\infty}^2 \| f_2 \|_2^2; \]
\[ \sum_{k \in \mathbb{Z}} \| \sum_{n>k} L_k(\mathcal{E}_n f_1, d_{2,n}) \|_2^2 \lesssim \| f_1 \|_{L^\infty}^2 \| f_2 \|_2^2; \]
\[ \sum_{k \in \mathbb{Z}} \| \sum_{n<k} L_k(\mathcal{E}_n f_1, d_{2,n}) \|_2^2 \lesssim \| f_1 \|_{L^\infty}^2 \| f_2 \|_2^2. \]

We first prove (3.2) and (3.4). We assume $n > k$, then $\mathcal{E}_k d_{1,n} = d_{1,n}$. Write
\[ \| L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2^2 = \sum_{Q \in \mathcal{G}_{n-1}} \int_Q | L_k(d_{1,n}, \mathcal{E}_{n-1} f_2)(x) |^2 \, dx. \]
Since $d_{1,n}$ and $\mathcal{E}_{n-1} f_2$ are both constants on the atom $Q \in \mathcal{G}_{n-1}$, we have
\[ L_k(d_{1,n}, \mathcal{E}_{n-1} f_2)(x) = A_2^G(d_{1,n}, \mathcal{E}_{n-1} f_2)(x) - d_{1,n}(x) \cdot \mathcal{E}_{n-1} f_2(x) = 0 \]
if $(x, x) + G_{2k} \subset Q \times Q$. Thus, for $x \in Q$, $L_k(d_{1,n}, \mathcal{E}_{n-1} f_2)(x)$ is nonzero only if $(x, x) + G_{2k}$ intersects with the complement of $Q \times Q$. For any subset $E$ of $\mathbb{R}^{2d}$ containing the origin, denote
\[ \mathcal{H}(E, Q) = \{ x \in Q : [(x, x) + E] \cap (Q \times Q)^c \neq \emptyset \}. \]
By (3.1), $\mathcal{H}(G_{2k}, Q) \subset \mathcal{H}(\tilde{B}_{2k}, Q)$, and
\[ | \mathcal{H}(G_{2k}, Q) | \leq | \mathcal{H}(\tilde{B}_{2k}, Q) | \lesssim 2^{(d-1)n} \cdot 2^k. \]
From this, the Cauchy-Schwarz inequality and Lemma 2.6, it follows that
\[ \int_Q | L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) |^2 \leq \int_{\mathcal{H}(G_{2k}, Q)} | L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) |^2 \]
\[ \lesssim 2^{(d-1)n} \cdot 2^k \cdot \{ [d_{1,n}, \mathcal{E}_{n-1} f_2]^+ \}^2 = 2^{-|k-n|} \int_Q \{ [d_{1,n}, \mathcal{E}_{n-1} f_2]^+ \}^2. \]
Summing this over $Q \in \mathcal{G}_{n-1}$ and using (2.2), we get
\[ \| L_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2 \lesssim 2^{-\frac{|k-n|}{2}} \| [d_{1,n}, \mathcal{E}_{n-1} f_2]^+ \|_2 \lesssim 2^{-\frac{|k-n|}{2}} \| d_{1,n} \cdot \mathcal{E}_{n-1} f_2 \|_2. \]
Applying Lemma 2.1, (i) with $S_k = L_k$, $u_n = d_{1,n}$, $v_n = E_{n-1} f_2$, $a(d_{1,n}, E_{n-1} f_2) = \|d_{1,n} \cdot E_{n-1} f_2\|_2$ and $\sigma(n) = -|n|/2$, Lemma 2.3, (i), and the Carleson inequality, we obtain
\[
\sum_k \left( \sum_{n > k} \|L_k(d_{1,n}, E_{n-1} f_2)\|_2 \right)^2 \lesssim \left( \sum_n 2^{-|n|/2} \right)^2 \cdot \left( \sum_n \int_{\mathbb{R}^d} |d_{1,n}|^2 \cdot |E_{n-1} f_2|^2 \right)
\lesssim \|f_1\|_{BMO}^2 \|M(f_2)\|_2^2 \lesssim \|f_1\|_{\infty}^2 \|f_2\|_2^2.
\]
This concludes the proof of (3.2).

To show (3.4), by the same arguments as above, we get
\[
\sum_k \left( \sum_{n > k} \|L_k(E_{n} f_1, d_{2,n})\|_2 \right)^2 \lesssim \sum_n \int_{\mathbb{R}^d} |E_{n} f_1|^2 \cdot |d_{2,n}|^2.
\]
Now, using the simple estimate $\sup_n \|E_{n} f_1\|_\infty \leq \|f_1\|_\infty$, the last term above is majorized by $\|f_1\|_\infty^2 \sum_n \|d_{2,n}\|_2^2 = \|f_1\|_{\infty}^2 \|f_2\|_2^2$, which gives (3.4).

We first prove (3.3) and (3.5). Assume $n \leq k$. Since $E_k d_{1,n} = 0$ in this case, (3.3) and (3.5) will follow from the following pointwise estimates respectively: For $1 < l < 2$,
\[
|A_{k}^G(d_{1,n}, E_{n-1} f_2)|^2 \lesssim 2^{-|k-n|(2-2/l)} A_{k}^G(|d_{1,n}|^l, |E_{n-1} f_2|^l)^{2/l},
\]
and
\[
|A_{k}^G(E_{n} f_1, d_{2,n})|^2 \lesssim 2^{-|k-n|(2-2/l)} A_{k}^G(|E_{n} f_1|^l, |d_{2,n}|^l)^{2/l}.
\]
Indeed, assume that (3.6) and (3.7) hold for the moment, let us prove (3.3) and (3.5). Denote by $Q_{2k}$ the cube centered at origin and having side-length $2^k$ in $\mathbb{R}^d$ and let $\tilde{I}_{Q_{2k}} = |Q_{2k}|^{-1} I_{Q_{2k}}$. Recalling that $G_{2k} \subset \tilde{B}_{2k}$, we have
\[
A_{2k}^G(|g_1|, |g_2|)(x) \lesssim \left( \frac{1}{|Q_{2k}|} \int_{Q_{2k}} |g_1(x - y_1)| dy_1 \right) \left( \frac{1}{|Q_{2k}|} \int_{Q_{2k}} |g_2(x - y_2)| dy_2 \right)
= \left( \tilde{I}_{Q_{2k}} \ast |g_1|(x) \right) \cdot \left( \tilde{I}_{Q_{2k}} \ast |g_2|(x) \right) \lesssim (\zeta_k \ast |g_1|(x)) \cdot (\zeta_k \ast |g_2|(x)),
\]
where $\zeta(x) = (1 + |x|)^{-d-\epsilon}$ for some $\epsilon > 0$. Applying (3.6) and the above estimate with $g_1 = |d_{1,n}|^l$ and $g_2 = |E_{n-1} f_2|^l$, we get
\[
\|A_{2k}^G(d_{1,n}, E_{n-1} f_2)\|_2 \lesssim 2^{-(1-1/l)|k-n|} \|A_{2k}^G(|d_{1,n}|^l, |E_{n-1} f_2|^l)^{1/l}\|_2
\lesssim 2^{-(1-1/l)|k-n|} \|\zeta_k \ast |d_{1,n}|^l\|^{1/l} \cdot \|\zeta_k \ast |E_{n-1} f_2|^l\|^{1/l},
\]
Applying Lemma 2.1, (ii) with $a_k(u_n, v_n) = \|\zeta_k \ast |u_n|^l\|^{1/l} \cdot \|\zeta_k \ast |v_n|^l\|^{1/l}_2$, $u_n = d_{1,n}$ and $v_n = E_{n-1} f_2$, and Lemma 2.4 with $b = f_1$ and $f = M f_2$, we deduce that
\[
\sum_k \|\sum_{n \leq k} A_{2k}^G(d_{1,n}, E_{n-1} f_2)\|_2^2
\lesssim \sup_{n' \in \mathbb{N}} \sum_k \int_{\mathbb{R}^d} (\zeta_k \ast |d_{1,k-n'}|^{l})^{2/l}(x) (\zeta_k \ast |E_{k-n'} f_2|^l)^{2/l}(x) \, dx
\lesssim \sup_{n' \in \mathbb{N}} \sum_k \int_{\mathbb{R}^d} (\zeta_k \ast |d_{1,k-n'}|^{l})^{2/l}(x) (\zeta_k \ast |M f_2|^l)^{2/l}(x) \, dx
\lesssim \|f_1\|_{BMO}^2 \|M f_2\|_2^2 \lesssim \|f_1\|_{\infty}^2 \|f_2\|_2^2,
\]
which gives (3.3).
For (3.5), integrating both sides of (3.7) and using the same arguments as above, we have

\[
\sum_k \left\| \sum_{n \leq k} A^{G}_{2^k} (E_n f_1, d_{2^k}, n) \right\|_2^2 
\leq \sup_{n' \in \mathbb{N}} \sum_k \int_{\mathbb{R}^d} (\zeta_k * |d_{2^k, n'-1}|)^{2/l}(x)(\zeta_k * |E_{k-n'} f_1|^l)^{2/l}(x) \, dx 
\leq \sup_{n' \in \mathbb{N}} \sum_k \int_{\mathbb{R}^d} M(|d_{2^k, n'-1}|^{2/l}(x) \cdot M(M f_1^{l})^{2/l}(x) \, dx 
\leq \|M(M f_1^{l})^{2/l}\|_\infty \sup_{n' \in \mathbb{N}} \sum_k \|M(|d_{2^k, n'-1}|^{l})\|_{L^2/l}^2 
\leq \|f_1\|_\infty^2 \cdot \left(\sup_{n' \in \mathbb{N}} \sum_k \|d_{2^k, n'}\|_2^2\right) \lesssim \|f_1\|_\infty^2 \|f_2\|_2^2,
\]

which gives (3.5).

Let us now show (3.6) and (3.7) under the assumption \( n \leq k \). We divide \( \mathbb{R}^{2d} \) into all atoms \( \tilde{Q} \) in \( \mathcal{D}_n \), and write

\[
\int_{(x,x)+G_{2^k}} (d_{1,n} \otimes E_{n-1} f_2) = \sum_{\tilde{Q} \in \mathcal{D}_n} \int_{[(x,x)+G_{2^k}] \cap \tilde{Q}} (d_{1,n} \otimes E_{n-1} f_2).
\]

Writing \( \tilde{Q} = Q_1 \times Q_2 \in \mathcal{D}_n \times \mathcal{D}_n \), then \( \int_{Q_1} d_{1,n} = 0 \) if \( \tilde{Q} \subset (x,x) + G_{2^k} \), which implies

\[
\int_{\tilde{Q}} (d_{1,n} \otimes E_{n-1} f_2) = \left(\int_{Q_1} d_{1,n}\right) \left(\int_{Q_2} E_{n-1} f_2\right) = 0.
\]

Hence \( \int_{[(x,x)+G_{2^k}] \cap \tilde{Q}} (d_{1,n} \otimes E_{n-1} f_2) \) is non-zero only if \( \tilde{Q} \) intersects with the boundary of \( (x, x) + G_{2^k} \). Recall that

(3.8) \( I((x, x) + G_{2^k}, n) = \bigcup \{(x, x) + G_{2^k} \cap \tilde{Q} : \tilde{Q} \in \mathcal{D}_n, [(x, x) + \partial G_{2^k}] \cap \tilde{Q} \neq \emptyset\} \).

Then

\[
\int_{(x,x)+G_{2^k}} (d_{1,n} \otimes E_{n-1} f_2) = \int_{I((x, x) + G_{2^k}, n)} (d_{1,n} \otimes E_{n-1} f_2)
\]

and

(3.9) \( |I((x, x) + G_{2^k}, n)| \lesssim 2^{n-k} 2^{kd} \).

(See [31] for more discussions on (3.9)).

Using Hölder’s inequality, (3.9) and (3.1), we get that for \( 1 < l < 2 \),

\[
|A^{G}_{2^k} (d_{1,n}, E_{n-1} f_2)|^l = \frac{1}{|G_{2^k}|^l} \left| \int_{I((x, x) + G_{2^k}, n)} (d_{1,n} \otimes E_{n-1} f_2) \right|^l 
\lesssim \frac{1}{(2^{kd})^l} (2^{n+2d-l-1})^\frac{l-1}{l} \int_{(x,x)+G_{2^k}} (|d_{1,n}|^l \otimes |E_{n-1} f_2|^l) 
\lesssim 2^{(n-k)(l-1)} \frac{1}{|G_{2^k}|} \int_{(x,x)+G_{2^k}} (|d_{1,n}|^l \otimes |E_{n-1} f_2|^l),
\]

which gives (3.8).
which implies (3.6). Inequality (3.7) can be proved analogously, the details being omitted. This concludes the proof of the \( L^\infty \times L^2 \rightarrow L^2 \) long variation estimates.

3.2. Short variation estimates. By the embedding \( \ell^2 \hookrightarrow \ell^q \), it suffices to bound \( \mathcal{S}\mathcal{V}_2(A_t^G) \) instead of \( \mathcal{S}\mathcal{V}_q(A_t^G) \). Define a bi-subadditive operator \( G_k \) by

\[
G_k(f_1, f_2)(x) = \left( \sum_{i \in \mathcal{I}_k} |(A_{t,i+1}^G - A_{t_i}^G)(f_1, f_2)(x)|^2 \right)^{1/2},
\]

where we recall that \( I_i \in \mathcal{I}_k \) means that \( I_i = (t_i, t_{i+1}] \subset (2^k, 2^{k+1}] \). We need to show

\[
\left\| \left( \sum_{k \in \mathbb{Z}} [G_k(f_1, f_2)]^2 \right)^{1/2} \right\|_2 \lesssim \|f_1\|_\infty \|f_2\|_2.
\]

Similar to (2.1), we can express \( (A_{t,i+1}^G - A_{t_i}^G)(f_1, f_2) \) as

\[
(A_{t,i+1}^G - A_{t_i}^G)(f_1, f_2) = \sum_{n \in \mathbb{Z}} (A_{t,i+1}^G - A_{t_i}^G)(d_{1,n}, \mathcal{E}_{n-1} f_2) + \sum_{n \in \mathbb{Z}} (A_{t,i+1}^G - A_{t_i}^G)(\mathcal{E}_n f_1, d_{2,n}),
\]

where the series converges pointwise. Taking the \( \ell^2 \) norm over \( I_i \in \mathcal{I}_k \) and using triangle inequality, we get

\[
G_k(f_1, f_2) \lesssim \sum_{n \in \mathbb{Z}} G_k(d_{1,n}, \mathcal{E}_{n-1} f_2) + \sum_{n \in \mathbb{Z}} G_k(\mathcal{E}_n f_1, d_{2,n}).
\]

To show (3.10), we consider two cases \( n > k \) and \( n \leq k \) separately, and it then suffices to show

\[
\sum_{k \in \mathbb{Z}} \| \sum_{n > k+1} G_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2 \lesssim \|f_1\|_\infty \|f_2\|_2^2,
\]

\[
\sum_{k \in \mathbb{Z}} \| \sum_{n \leq k+1} G_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2 \lesssim \|f_1\|_\infty \|f_2\|_2^2,
\]

\[
\sum_{k \in \mathbb{Z}} \| \sum_{n > k+1} G_k(\mathcal{E}_n f_1, d_{2,n}) \|_2 \lesssim \|f_1\|_\infty \|f_2\|_2^2,
\]

\[
\sum_{k \in \mathbb{Z}} \| \sum_{n \leq k+1} G_k(\mathcal{E}_n f_1, d_{2,n}) \|_2 \lesssim \|f_1\|_\infty \|f_2\|_2^2.
\]

We only show the first two inequalities as the others can be handled similarly.

Let us show the first inequality of (3.12). We assume \( n > k + 1 \). By the almost orthogonality principle, matters are reduced to showing

\[
\|G_k(d_{1,n}, \mathcal{E}_{n-1} f_2)\|_2^2 \lesssim 2^{-(n-k)} \|d_{1,n} \cdot \mathcal{E}_{n-1} f_2\|_2^2.
\]

In fact, once (3.13) is established, it then follows from Lemma 2.1, (i) with \( S_k = G_k \), \( u_n = d_{1,n} \), \( v_n = \mathcal{E}_{n-1} f_2 \) and \( a(u_n, v_n) = \|d_{1,n} \cdot \mathcal{E}_{n-1} f_2\|_2 \) that

\[
\sum_{k \in \mathbb{Z}} \| \sum_{n > k+1} G_k(d_{1,n}, \mathcal{E}_{n-1} f_2) \|_2^2 \lesssim \sum_{n} \|d_{1,n} \cdot \mathcal{E}_{n-1} f_2\|_2^2,
\]

which is bounded by \( C\|f_1\|_\infty^2 \|f_2\|_2^2 \), as shown in the proof of (3.2) and (3.4).
Let us show (3.13). We write
\[ \|G_k(d_{1,n}, \mathcal{E}_{n}f_2)\|_2^2 = \sum_{Q \in \mathcal{D}_{n-1}} \int_Q \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)(x)|^2 \, dx. \]

We have
\[ \left( \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)|^2 \right)^\frac{1}{2}(x) \]
\[ \leq \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)|(x) \]
\[ \leq \sum_{I_i \in \mathcal{I}_k} \frac{1}{|G_{i+1}|} \int_{(x,x)+G_{i+1}\setminus G_i} |d_{1,n}| \otimes |\mathcal{E}_{n-1}f_2| \]
\[ + \sum_{I_i \in \mathcal{I}_k} \left( \frac{1}{|G_i|} - \frac{1}{|G_{i+1}|} \right) \int_{(x,x)+G_{i+1}} |d_{1,n}| \otimes |\mathcal{E}_{n-1}f_2| \]
\[ \leq \frac{1}{|G_{2k}|} \sum_{I_i \in \mathcal{I}_k} \int_{(x,x)+G_{2k+1}\setminus G_i} |d_{1,n}| \otimes |\mathcal{E}_{n-1}f_2| \]
\[ + \left( \sum_{I_i \in \mathcal{I}_k} \frac{1}{|G_i|} - \frac{1}{|G_{i+1}|} \right) \int_{(x,x)+G_{2k+1}} |d_{1,n}| \otimes |\mathcal{E}_{n-1}f_2| \]
\[ \lesssim \frac{1}{|G_{2k}|} \int_{(x,x)+G_{2k+1}} |d_{1,n}| \otimes |\mathcal{E}_{n-1}f_2| \]
\[ \lesssim A_{2k+1}^G([d_{1,n}|, \mathcal{E}_{n-1}f_2])(x) \leq [d_{1,n}, \mathcal{E}_{n-1}f_2]^+(x), \]
where we used Lemma 2.6 in the last inequality. For each \( Q \in \mathcal{D}_{n-1} \) and \( x \in Q \), since \( d_{1,n} \otimes \mathcal{E}_{n-1}f_2 \) is constant on \( Q \times Q \), we see that
\[ \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)(x)|^2 \]
is nonzero only if, for some \( I_i \in \mathcal{I}_k \), the convex body \((x, x) + G_{i+1}\) (hence \((x, x) + G_{2k+1}\)) intersects with the complement of \( Q \times Q \). Hence
\[ \int_Q \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)|^2 \lesssim \int_{\mathcal{H}(G_{2k+1}, Q)} \{[d_{1,n}, \mathcal{E}_{n-1}f_2]^+\}^2 \]
\[ \lesssim 2^{(d-1)n}2^k \{[d_{1,n}, \mathcal{E}_{n-1}f_2]^+\}^2 \]
\[ = 2^{k-n} \int_Q \{[d_{1,n}, \mathcal{E}_{n-1}f_2]^+\}^2, \]
where
\[ \mathcal{H}(G_{2k+1}, Q) = \{ x \in Q : [(x, x) + G_{2k+1}] \cap (Q \times Q)^c \neq \emptyset \}, \]
and its measure is no more than \( 2^{(d-1)n}2^k \). Summing the above over \( Q \in \mathcal{D}_{n-1} \), and using (2.2), we obtain
\[ \sum_{Q \in \mathcal{D}_{n-1}} \int_Q \sum_{I_i \in \mathcal{I}_k} |(A_{i+1}^G - A_i^G)(d_{1,n}, \mathcal{E}_{n-1}f_2)|^2 \]
\[ \lesssim 2^{k-n} \int_{\mathbb{R}^d} \left\{ [d_{1,n}, \mathcal{E}_{-1} f_2]^+ \right\}^2 \lesssim 2^{k-n} \int_{\mathbb{R}^d} |d_{1,n}|^2 \cdot |\mathcal{E}_{-1} f_2|^2, \]

which verifies (3.13), and hence the first inequality in (3.12).

Finally, let us prove the second inequality in (3.12). We assume \( k \geq n-1 \). By arguments similar to that in the long variation case, (3.12) will follow from the pointwise estimate:

For \( 1 \leq l \leq 2 \),

\[ \sum_{I_i \in \mathcal{I}_k} |(A_{n,k+1}^G - A_{l,i}^G)(d_{1,n}, \mathcal{E}_{-1} f_2)|^2 \leq 2^{(n-k)(2-l/|E|)} A_{2k+1}^G (|d_{1,n}|^l, |\mathcal{E}_{-1} f_2|^l)^{2/l}. \]

To show (3.14), we write

\[ \sum_{I_i \in \mathcal{I}_k} |(A_{n,k+1}^G - A_{l,i}^G)(d_{1,n}, \mathcal{E}_{-1} f_2)|^2 \]

\[ \lesssim \sum_{I_i \in \mathcal{I}_k} \left| \frac{1}{|G_{t_i}|} \int_{(x,x)+G_{t_{i+1}} \setminus G_{t_i}} (d_{1,n} \otimes \mathcal{E}_{-1} f_2) \right|^2 \]

\[ + \sum_{I_i \in \mathcal{I}_k} \left( \frac{1}{|G_{t_i}|} - \frac{1}{|G_{t_{i+1}}|} \right)^2 \left| \int_{(x,x)+G_{t_{i+1}}} d_{1,n} \otimes \mathcal{E}_{-1} f_2 \right|^2 \]

\[ \lesssim \frac{1}{|G_{2k+1}|^2} \left( \sum_{I_i \in \mathcal{I}_k} \left| \int_{(x,x)+G_{t_{i+1}} \setminus G_{t_i}} (d_{1,n} \otimes \mathcal{E}_{-1} f_2) \right|^l \right)^{2/l} \]

\[ + \left( \sum_{I_i \in \mathcal{I}_k} \left( \frac{1}{|G_{t_i}|} - \frac{1}{|G_{t_{i+1}}|} \right)^l \sup_{I_i \in \mathcal{I}_k} \left| \int_{(x,x)+G_{t_{i+1}}} d_{1,n} \otimes \mathcal{E}_{-1} f_2 \right|^l \right)^{2/l} \]

\[ =: I + II. \]

For the first term \( I \), by the fact that

\[ \int_{\bar{Q}} (d_{1,n} \otimes \mathcal{E}_{-1} f_2) = \left( \int_{Q_1} d_{1,n} \right) \cdot \left( \int_{Q_2} \mathcal{E}_{-1} f_2 \right) = 0 \]

for any \( \bar{Q} = Q_1 \times Q_2 \in \mathcal{D}_n \), we have

\[ \int_{(x,x)+G_{t_{i+1}} \setminus G_{t_i}} (d_{1,n} \otimes \mathcal{E}_{-1} f_2) = \int_{I((x,x)+G_{t_{i+1}} \setminus G_{t_i}, n)} (d_{1,n} \otimes \mathcal{E}_{-1} f_2), \]

where \( I((x,x) + G_{t_{i+1}} \setminus G_{t_i}, n) \) was defined as in (3.8). From Hölder’s inequality, the estimate \( |I((x,x) + G_{t_{i+1}} \setminus G_{t_i}, n)| \lesssim 2^{n} 2^{(2d-1)k} \) and (3.1), it follows that

\[ I^{l/2} \lesssim \frac{1}{(2k^d)^l} (2^{n} 2^{(2d-1)k})^{l-1} \sum_{I_i \in \mathcal{I}_k} \int_{(x,x)+G_{t_{i+1}} \setminus G_{t_i}} (|d_{1,n}|^l \otimes |\mathcal{E}_{-1} f_2|^l) \]

\[ \lesssim 2^{(n-k)(l-1)} \frac{1}{|G_{2k+1}|} \int_{(x,x)+G_{2k+1}} (|d_{1,n}|^l \otimes |\mathcal{E}_{-1} f_2|^l), \]

which gives the required bound for \( I \).
For the second term $I I$, we have
\[
\sum_{l \in \mathcal{Y}_k} \left( \frac{1}{|G_l|} - \frac{1}{|G_{l+1}|} \right)^{l} \leq \left( \frac{1}{|G_2|} - \frac{1}{|G_{2^k+1}|} \right)^{l} \lesssim 2^{-2dkl}.
\]
Using Hölder’s inequality, the estimate $|I((x, x) + G_l, n)| \lesssim 2^n 2^{(2d-1)k}$, and (3.1), we deduce that
\[
II^{1/2} \lesssim \frac{1}{(2^{dk})^t} (2^n 2^{(2d-1)k})^{-1} \sup_{l \in \mathcal{Y}_k} \int_{(x, x) + G_{l+1}} (|d_{1, n}| \otimes |\mathcal{E}_{n-1} f_2|)
\lesssim 2^{(n-d)(t-1)} \frac{1}{|G_{2^k+1}|} \int_{(x, x) + G_{2^k+1}} (|d_{1, n}| \otimes |\mathcal{E}_{n-1} f_2|),
\]
which gives the required bound for $II$. This verifies (3.14) and hence completes the proof of the $L^\infty \times L^2 \to L^2$ short variation estimate.

**Remark 3.1.** The hypothesis $f_1 \in L^\infty_c$ can be replaced by $f_1 \in L^\infty \cap L^2$, and the $L^\infty \times L^2 \to L^2$ bound still holds true. Interchanging the roles of $f_1$ and $f_2$, we can conclude that the $L^2 \times L^\infty \to L^2$ variational inequality holds for $f_1 \in L^2$ and $f_2 \in L^\infty \cap L^2$.

## 4. $L^\infty \times L^\infty \to BMO_d$ estimate

Let us prove the long variation inequality first, namely, $L V_q(A^G_t)$ is bounded from $L^\infty \times L^\infty$ to $BMO_d$ for $f_1, f_2 \in L^\infty_c$.

For $f_i \in L^\infty_c$ and a dyadic cube $Q$ in $\mathbb{R}^d$, we divide $f_i$ into the local and global parts:
\[
f_i = f_i^{1 Q^*} + f_i^{0 \mathbb{R}^d \setminus Q^*} = f_i^{0} + f_i^{\infty},
\]
where $Q^* = 3\sqrt{d}Q$. By the triangle inequality,
\[
\frac{1}{|Q|} \int_Q |L V_q(A^G_t)(f_1, f_2) - a_Q|
\leq \frac{1}{|Q|} \int_Q |L V_q(A^G_t)(f_1^{0}, f_2^{0})| + \frac{1}{|Q|} \int_Q |L V_q(A^G_t)(f_1^{0}, f_2^{\infty})|
\quad + \frac{1}{|Q|} \int_Q |L V_q(A^G_t)(f_1^{\infty}, f_2^{0})| + \frac{1}{|Q|} \int_Q |L V_q(A^G_t)(f_1^{\infty}, f_2^{\infty}) - a_Q|
\leq I_1 + I_2 + I_3 + I_4,
\]
where $a_Q$ is a constant depending on $Q$ which will be determined later.

Since the first three terms $I_1, I_2, I_3$ can be estimated in the same way, we just estimate $I_2$. Since $f_1^{0} \in L^2$ and $f_2^{\infty} \in L^\infty \cap L^2$, by the $L^\infty \times L^2 \to L^2$ boundedness of $L V_q(A^G_t)$ (see Remark 3.1),
\[
I_2 \leq \left( \frac{1}{|Q|} \int_Q |V_q(A^G_{2^k})(f_1^{0}, f_2^{\infty})|^2 \right)^{1/2}
\lesssim \frac{1}{\sqrt{|Q|}} \|f_1^{0}\|_{L^2(Q^*)} \|f_2^{\infty}\|_{L^\infty(Q^*)} \leq \|f_1\|_{\infty} \|f_2\|_{\infty}.
\]
It remains to estimate $I_4$. Choosing $a_Q = \mathcal{L}V_q(A_i^G(f_1^2, f_2^2))(c_Q)$, by triangle inequality in $\ell^q$ and the embedding $\ell^q \hookrightarrow \ell^2$, it suffices to show that for any $x \in Q$,

$$
\left( \sum_{k \in \mathbb{Z}} |A_{2^k}^G(f_1^\infty, f_2^\infty)(x) - A_{2^k}^G(f_1^\infty, f_2^\infty)(c_Q)|^2 \right)^{1/2} \lesssim \|f_1\|_\infty \|f_2\|_\infty,
$$

where the implicit constant is independent of $x$ and $Q$.

If $2^k < \ell(Q)$, then for any $x \in Q$ and any $(y, y') \in G_{2^k}$,

$$
|x + y - (c_Q)| \leq |x - (c_Q)| + |y| < \sqrt{d} \ell(Q),
$$

which implies that $(x, x) + G_{2^k}$ is contained in $Q^* \times Q^*$. Hence, by the support of $f_1^\infty$,

$$
A_{2^k}^G(f_1^\infty, f_2^\infty)(x) = A_{2^k}^G(f_1^\infty, f_2^\infty)(c_Q) = 0
$$

for all $x \in Q$.

If $2^k \geq \ell(Q)$, then

$$
\left| A_{2^k}^G(f_1^\infty, f_2^\infty)(x) - A_{2^k}^G(f_1^\infty, f_2^\infty)(c_Q) \right| = \frac{1}{|G_{2^k}|} \left| \int_{(x,x)+G_{2^k}} f_1^\infty \otimes f_2^\infty \right| 
$$

$$
= \frac{1}{2^{kd}} \left| \int_{\mathbb{R}^{2d}} (f_1^\infty \otimes f_2^\infty)(1_{([c_Q,c_Q]+G_{2^k})\setminus ([x,x]+G_{2^k})}) - 1_{([x,x]+G_{2^k})\setminus ([c_Q,c_Q]+G_{2^k})}) \right| 
$$

$$
\leq \frac{1}{2^{kd}} \|f_1\|_\infty \|f_2\|_\infty \|([c_Q,c_Q] + G_{2^k}) \Delta ([x,x] + G_{2^k})| 
$$

$$
\lesssim \ell(Q)^{2k(2d-1)} \|f_1\|_\infty \|f_2\|_\infty,
$$

where $E \Delta F$ denotes the symmetric difference of two sets $E$ and $F$. In the last inequality we used the observation that, by (3.1), the measure of projection of the symmetric difference $\{|c_Q,c_Q] + G_{2^k} \Delta ([x,x] + G_{2^k})\}$ to a hyperplane perpendicular to the vector $(x, x) - (c_Q, c_Q)$ is not more than the “volume” of $(2d-1)$-dimensional ball $B_{2^{d-1}}(0, 2^k)$, and therefore

$$
\|([c_Q,c_Q] + G_{2^k}) \Delta ([x,x] + G_{2^k})| \lesssim \ell(Q) \cdot 2^{k(2d-1)}.
$$

Summing over $\{k : 2^k \geq \ell(Q)\}$, we obtain

$$
\left( \sum_{k : 2^k \geq \ell(Q)} |A_{2^k}^G(f_1^\infty, f_2^\infty)(x) - (A_{2^k}^G(f_1^\infty, f_2^\infty))(c_Q)|^2 \right)^{1/2} \lesssim \sum_{k : 2^k \geq \ell(Q)} \ell(Q)^{2k(2d-1)} \|f_1\|_\infty \|f_2\|_\infty \lesssim \|f_1\|_\infty \|f_2\|_\infty,
$$

which concludes the proof of the long variation inequality.

Let us now treat the short variation operator. By arguments similar to that given in the long variation case, it suffices to show that, for any $x \in Q$ and any increasing sequence $(t_i)_i$,

$$
\sum_{2^k \geq \ell(Q)} \sum_{t_i \in \mathcal{F}_k} |(A_{t_{i+1}}^G - A_{t_i}^G)(f_1^\infty, f_2^\infty)(x) - (A_{t_{i+1}}^G - A_{t_i}^G)(f_1^\infty, f_2^\infty)(c_Q)|^2 \lesssim \|f_1\|_\infty^2 \|f_2\|_\infty^2,
$$

where $\mathcal{F}_k$ is the set of all increasing sequences $(t_i)_i$.
where the implicit constant is independent of $Q$ and the sequence $(t_i)_i$. By comparing $t_{i+1} - t_i$ with $\ell(Q)$, we divide the summation in $I_i \in \mathcal{S}_k$ into two pieces according to whether $t_{i+1} - t_i$ is bigger or smaller than $\ell(Q)$.

We consider the case $t_{i+1} - t_i > \ell(Q)$ first. Since the family $(G_i)_i$ of convex bodies are nested, we have that for any $z \in Q$,

$$|A^G_{t_{i+1}}(f_1^\infty, f_2^\infty)(z) - A^G_{t_i}(f_1^\infty, f_2^\infty)(z)| = \left| \left( \frac{1}{|G_{t_{i+1}}|} - \frac{1}{|G_{t_i}|} \right) \int_{(z,z) + G_{t_i}} f_1^\infty \otimes f_2^\infty + \frac{1}{|G_{t_{i+1}}|} \int_{(z,z) + [G_{t_{i+1}} \setminus G_{t_i}]} f_1^\infty \otimes f_2^\infty \right|$$

$$\leq \left( \frac{1}{|G_{t_i}|} - \frac{1}{|G_{t_{i+1}}|} \right) |G_{t_i}| \|f_1\|_\infty \|f_2\|_\infty + \frac{1}{|G_{t_{i+1}}|} (|G_{t_{i+1}}| - |G_{t_i}|) \|f_1\|_\infty \|f_2\|_\infty$$

$$= 2(|G_{t_{i+1}}| - |G_{t_i}|) \frac{1}{|G_{t_{i+1}}|} \|f_1\|_\infty \|f_2\|_\infty \int_{|G_{t_i}|}^{2|G_{t_{i+1}}|} \frac{1}{u} \, du.$$

Applying the Cauchy-Schwarz inequality and using $t_{i+1} - t_i < \ell(Q)$ and $t_i \sim t_{i+1}$, we can continue the above estimates by

$$\lesssim \|f_1\|_\infty \|f_2\|_\infty \ell(Q) \frac{1}{|G_{t_i}|} \left( \int_{|G_{t_i}|}^{2|G_{t_{i+1}}|} \frac{1}{u^2} \, du \right)^{1/2}$$

Hence

$$\left( \sum_{2^k \geq \ell(Q)} \sum_{t_i \in \mathcal{S}_k} |A^G_{t_{i+1}}(f_1^\infty, f_2^\infty)(x) - A^G_{t_i}(f_1^\infty, f_2^\infty)(x)|^2 \right)^{1/2} \lesssim \|f_1\|_\infty \|f_2\|_\infty \ell(Q) \frac{1}{|G_{t_i}|} \left( \sum_{2^k \geq \ell(Q)} 2^{k(2d-1)} \int_{|G_{t_i}|}^{2|G_{t_{i+1}}|} \frac{1}{u^2} \, du \right)^{1/2}$$

$$\leq \|f_1\|_\infty \|f_2\|_\infty \ell(Q) \frac{1}{|G_{t_i}|} \left( \sum_{2^k \geq \ell(Q)} 2^{k(2d-1)} \int_{(2^k)^{2d}} \frac{1}{u^2} \, du \right)^{1/2}$$

$$\leq \|f_1\|_\infty \|f_2\|_\infty \ell(Q) \frac{1}{|G_{t_i}|} \left( \sum_{2^k \geq \ell(Q)} \tau^{-2d} 2^{-k} \right)^{1/2}$$

$$\lesssim \|f_1\|_\infty \|f_2\|_\infty \ell(Q),$$

where $\tau$ is the constant in (3.1).
It remains to treat the case $t_{i+1} - t_i \geq \ell(Q)$. By an argument similar to (4.2), we can get
\[
|A_{t_i}^G(f_1^\infty, f_2^\infty)(x) - A_{t_i}^G(f_1^\infty, f_2^\infty)(c_Q)| \lesssim |Q|^2 \|f_1\|_\infty \|f_2\|_\infty \frac{1}{|G_{t_i}|}.
\]
Since the number of intervals $I_i \in \mathcal{J}$ satisfying $t_{i+1} - t_i \geq \ell(Q)$ is not more than $C 2^k/\ell(Q)$, we conclude that
\[
II_2 = \left( \sum_{2^k \geq \ell(Q)} \sum_{I_i \in \mathcal{J}_k} |A_{t_{i+1}}^G(f_1^\infty, f_2^\infty)(x) - A_{t_{i+1}}^G(f_1^\infty, f_2^\infty)(c_Q)|^2 \right)^{1/2}
\lesssim |Q|^2 \|f_1\|_\infty \|f_2\|_\infty \left( \sum_{2^k \geq \ell(Q)} \sum_{I_i \in \mathcal{J}_k} \frac{1}{|G_{t_i}|^2} \right)^{1/2}
\lesssim |Q|^2 \|f_1\|_\infty \|f_2\|_\infty \left( \sum_{2^k \geq \ell(Q)} \frac{2^k}{\ell(Q)^2 24kd} \right)^{1/2} \lesssim \|f_1\|_\infty \|f_2\|_\infty.
\]
This completes the proof of the short variational inequality.

5. $L^1 \times L^{p_2} \to L^{p,\infty}$ estimates, $1 + 1/p_2 = 1/p$

This section is devoted to the proof of $L^1 \times L^{p_2} \to L^{p,\infty}$ estimates for $1 \leq p_2 \leq \infty, 1/2 \leq p \leq 1$ with $1 + 1/p_2 = 1/p$. We distinguish two cases $1 \leq p_2 < \infty$ and $p_2 = \infty$, as the Calderón-Zygmund decomposition does not hold for functions in $L^\infty$. In the following two subsections, we treat the two cases $1 \leq p_2 < \infty$ and $p_2 = \infty$ respectively.

5.1. $L^1 \times L^{p_2} \to L^{p,\infty}$ estimates with $1 \leq p_2 < \infty$. Fix an $\alpha$ and assume $f_1 \in L^1$ and $f_2 \in L^{p_2}$, both with compact supports. Without loss of generality, we may assume $\|f_1\|_1 = \|f_2\|_{p_2} = 1$. Set $p_1 = 1$. For $i = 1, 2$, we perform a variant of Calderón-Zygmund decomposition given in [19] for $L^p$ functions $f_i$ at height $\alpha^{p/p_i}$ to obtain “good” and “bad” functions $g_i$ and $b_i$, and families of cubes $\{Q_{i,j}\}$ with disjoint interiors such that
(i) $f_i = g_i + b_i$,
(ii) $b_i = \sum_j b_{i,j}$,
(iii) $\text{supp } (b_{i,j}) \subset Q_{i,j}$,
(iv) $\int b_{i,j}(x) \, dx = 0$,
(v) $\|b_{i,j}\|_{p_i} \leq 2^{d+p_i} \alpha^p |Q_{i,j}|$,
(vi) $\bigcup_j |Q_{i,j}| \leq C \alpha^{-p} \|f_i\|_{p_i}$,
(vii) $\|b_{i,j}\|_{p_i} \leq 2^{d+p_i} \|f_i\|_{p_i}$,
(viii) $\|g_{i,j}\|_{p_i} \leq \|f_i\|_{p_i} \|g_i\|_\infty \leq 2^{d/p_i} \alpha^{p_i}$.

Indeed, to show the above decomposition, one can imitate the same idea of classical Calderón-Zygmund decomposition, but select a cube $Q$ when $(\frac{1}{|Q|} \int_Q |f_i|^{p_i})^{1/p_i} > \alpha^{p/p_i}$; the functions $g_i$ and $b_i$ are defined as in the classical case. Now, let us begin to treat the long and short variation operators separately.
5.1.1. Long variation estimates. By the Calderón-Zygmund decomposition mentioned above, we have

\[ |\{ \mathcal{L}V_q(A^G_I)(f_1, f_2)(x) > \alpha \}| \]
\[ \leq |\{ \mathcal{L}V_q(A^G_I)(g_1, g_2)(x) > \alpha/4 \}| + |\{ \mathcal{L}V_q(A^G_I)(g_1, b_2)(x) > \alpha/4 \}| \]
\[ + |\{ \mathcal{L}V_q(A^G_I)(b_1, g_2)(x) > \alpha/4 \}| + |\{ \mathcal{L}V_q(A^G_I)(b_1, b_2)(x) > \alpha/4 \}|, \]

where \( \mathcal{L}V_q(A^G_I) \) represents the long variation operator defined in Section 3. Next we estimate each of these terms separately.

**Estimate for the first term.** We note that \( g_2 \in L^\infty_c \) and \( \| g_1 \|^2_2 \leq \| g_1 \|_\infty \| g_1 \|_1 \lesssim \alpha^p \) by property (viii) of the Calderón-Zygmund decomposition. Now the \( L^2 \times L^\infty \rightarrow L^2 \) boundedness of \( \mathcal{L}V_q(A^G_I) \) gives

\[ |\{ \mathcal{L}V_q(A^G_I)(g_1, g_2)(x) > \alpha/4 \}| \lesssim \alpha^{-2} \| \mathcal{L}V_q(A^G_I)(g_1, g_2) \|_2^2 \]
\[ \lesssim \alpha^{-2} \| g_1 \|_2^2 \| g_2 \|_\infty^2 \lesssim \alpha^{-2} \cdot \alpha^p \cdot (\alpha^{p/2})^2 = \alpha^{-p}. \]

**Estimate for the second term.** By the trivial pointwise estimate

\[ \mathcal{L}V_q(A^G_I)(g_1, b_2)(x) \leq 2 \left( \sum_{k \in \mathbb{Z}} |A^G_{2k}(g_1, b_2)(x)|^2 \right)^{1/2}, \]

it suffices to show

\[ |\{ x \in \mathbb{R}^d : (\sum_{k} |A^G_{2k}(g_1, b_2)(x)|^2)^{1/2} > \alpha/8 \}| \lesssim \alpha^{-p}. \]

We use \( Q_{i,j} \) (or \( Q_{i,j'} \)) to denote the cubes stemming from the Calderón-Zygmund decomposition, and \( Q^*_I \) to denote the cubes in \( \mathcal{D}_n \). Denote by \( Q^* \) a certain dimensional dilate of a cube \( Q \) in \( \mathbb{R}^d \), and let

\[ \Omega = \bigcup_{j,j'} (Q_{1,j} \cup Q_{2,j'}), \quad \Omega^* = \bigcup_{j,j'} (Q^*_{1,j} \cup Q^*_{2,j'}). \]

By property (vi) of the Calderón-Zygmund decomposition,

\[ |\Omega^*| \leq \sum_{j,j'} (|Q^*_1| + |Q^*_2|) \lesssim \sum_{j,j'} (|Q_{1,j}| + |Q_{2,j'}|) \lesssim \alpha^{-p}. \]

It thus suffices to show

\[ |\{ x \in \mathbb{R}^d \setminus \Omega^* : (\sum_{k} |A^G_{2k}(g_1, b_2)(x)|^2)^{1/2} > \alpha/8 \}| \lesssim \alpha^{-p}. \]

By Chebyshev’s inequality, we have

\[ |\{ x \in \mathbb{R}^d \setminus \Omega^* : (\sum_{k} |A^G_{2k}(g_1, b_2)(x)|^2)^{1/2} > \alpha/8 \}| \lesssim \alpha^{-2} \int_{\mathbb{R}^d \setminus \Omega^*} \sum_{k} |A^G_{2k}(g_1, b_2)(x)|^2 \, dx. \]

For each \( n \in \mathbb{Z} \) and \( i = 1, 2 \), let \( Q^i_n \) be the collection of cubes \( Q_{i,j} \) of side length \( 2^n \). Define

\[ h_{i,n} = \sum_{j : Q_{i,j} \in \mathcal{Q}_n} b_{i,j}, \quad i = 1, 2. \]
Then
\[ b_i = \sum_{n \in \mathbb{Z}} \sum_{j: Q_{i,j} \in \mathcal{Q}_n} b_{i,j} = \sum_{n \in \mathbb{Z}} h_{i,n}. \]

Notice that \( A_{G_{2k}}^G(g_1, h_{2,n})(x) = 0 \) if \( n \geq k \) since \((x, x) + G_{2k} \) is disjoint from the strip \( \mathbb{R}^d \times Q_{2,j} \) for any \( Q_{2,j} \) contained in the support of \( h_{2,n} \). It thus suffices to show
\[
(5.3) \quad \int_{\mathbb{R}^d \setminus \Gamma} \sum_{k} |A_{G_{2k}}^G(g_1, \sum_{n < k} h_{2,n})(x)|^2 \, dx \lesssim \alpha^{2-p}.
\]

Let \( 1_{i,n} = \sum_{Q_{i,j} \in \mathcal{Q}_n} 1_{Q_{i,j}}, i = 1, 2. \) Then (5.3) will follow if we show that
\[
(5.4) \quad |A_{G_{2k}}^G(g_1, h_{2,n})(x)|^2 \lesssim 2^{n-k} \cdot \alpha^{2-p} A_{G_{2k}}^G(|g_1|, |1_{2,n}|)(x), \quad \forall x \in \mathbb{R}^d, n < k,
\]
where \( c = \sqrt{d+1} \). Indeed, integrating both sides of (5.4), applying Minkowski’s inequality and properties (vi) and (viii) of Calderón-Zygmund decomposition, we get
\[
\|A_{G_{2k}}^G(g_1, h_{2,n})\|^2 \lesssim 2^{n-k} \cdot \alpha^{2-p} \|A_{G_{2k}}^G(|g_1|, |1_{2,n}|)\|_1
\]
\[
(5.5) \quad \leq 2^{n-k} \cdot \alpha^{2-p} \frac{1}{G_{2k}} \int_{G_{2k}} \|g_1(\cdot + y_1)1_{2,n}(\cdot + y_2)\|_1 \, dy_1 dy_2
\]
\[
\leq 2^{n-k} \cdot \alpha^{2-p} \|g_1\|_{\infty} \|1_{2,n}\|_1.
\]

Applying Lemma 2.1, (ii) with \( S_k = A_{G_{2k}}^G, \sigma(j) = 2^{-\frac{|j|}{2}}, u_n = g_1, v_n = h_{2,n} \) and \( a(u_n, v_n) = (\alpha^{2-p}\|u_n\|_{\infty}\|v_n\|_1)^{1/2} \), the fact that the cubes in \( \mathcal{Q}_2 \) are pairwise disjoint, and property (vi) of the Calderón-Zygmund decomposition, we conclude that
\[
\text{LHS of (5.3)} \leq \alpha^{2-p} \|g_1\|_{\infty} \sum_{n \in \mathbb{Z}} \|1_{2,n}\|_1 \lesssim \alpha^2 \sum_{n \in \mathbb{Z}} \sum_{j: Q_{2,j} \in \mathcal{Q}_n} |Q_{2,j}| \lesssim \alpha^{2-p}.
\]

To finish the estimate for the second term, it remains to show (5.4). Assume \( n < k \). It follows from property (v) of Calderón-Zygmund decomposition and Hölder’s inequality that
\[
(5.6) \quad \|b_{2,j}\|_{L^1(Q_{2,j})} \leq \|b_{2,j}\|_{L^{p_2}} |Q_{2,j}|^{1-\frac{1}{p_2}} \lesssim \alpha^{\frac{p_2}{p_2-1}} |Q_{2,j}|.
\]

Write the support of \( g_1 \odot h_{2,n} \) as
\[
\mathbb{R}^d \times \left( \bigcup_{Q_{2,j} \in \mathcal{Q}_2} Q_{2,j} \right) = \bigcup_{Q_{1,t} \times Q_{2,j} \in \mathcal{S}_n \times \mathcal{Q}_2} Q_{1,t} \times Q_{2,j} =: \bigcup_{\tilde{Q}_{t,j} \in \mathcal{Q}_n \times \mathcal{Q}_2} \tilde{Q}_{t,j},
\]
where \( \tilde{Q}_{t,j} = Q_{1,t} \times Q_{2,j} \), which leads to the decomposition
\[
\int_{(x,x) + G_{2k}} g_1(1_{2,n}(\cdot)) \, dy_1 dy_2
\]
\[
= \sum_{\tilde{Q}_{t,j} \in \mathcal{Q}_n \times \mathcal{Q}_2} \int_{[(x,x) + G_{2k}] \cap \tilde{Q}_{t,j}} g_1(1_{2,n}(\cdot)) \, dy_1 dy_2.
\]

Since each \( b_{2,j} \) has mean zero, the integral on the right hand side is nonzero only if \( \tilde{Q}_{t,j} \) intersects with the boundary of the ball \((x, x) + G_{2k}\). Denote by \( \mathcal{J}(G_{2k}) \) the collection of such cubes \( \tilde{Q}_{t,j} \), i.e.,
\[
\mathcal{J}(\partial G_{2k}) = \{\tilde{Q}_{t,j} \in \mathcal{Q}_n \times \mathcal{Q}_2 : [(x,x) + \partial G_{2k}] \cap \tilde{Q}_{t,j} \neq \emptyset\}. \]
For $n < k$, the union of the cubes $\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})$ is contained in $(x, x) + G_{2^k}$ for $c = \sqrt{d} + 1$. Hence

$$\left| \int_{(x, x) + G_{2^k}} g_1(y_1) h_{2, n}(y_2) \, dy_1 dy_2 \right|$$

$$\leq \sum_{\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})} \left( \int_{Q'_{l,j}} |g_1(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |b_{2,j}(y_2)| \, dy_2 \right)$$

$$(5.7)$$

$$\lesssim \sum_{\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})} \left( \int_{Q'_{l,j}} |g_1(y_1)| \, dy_1 \right) \cdot (\alpha \frac{\|b\|}{|Q_{2,j}|})$$

$$= \alpha \frac{\|b\|}{|Q_{2,j}|} \sum_{\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})} \int_{\tilde{Q}_{l,j}} |g_1(y_1)| \mathbb{1}_{Q_{2,j}}(y_2) \, dy_1 dy_2$$

$$\leq \alpha \frac{\|b\|}{|Q_{2,j}|} \int_{(x, x) + G_{2^k}} |g_1(y_1)| \mathbb{1}_{2, n}(y_2) \, dy_1 dy_2,$$

where in the second inequality we used (5.6). Therefore,

$$|A_{2^k}^G(g_1, h_{2, n})(x)| \lesssim \frac{\|b\|}{|G_{2^k}|} A_{2^k}^G(|g_1|, \mathbb{1}_{2, n})(x).$$

(5.8)

On the other hand, using $\|g_1\|_\infty \lesssim \alpha^p$ we have

$$|A_{2^k}^G(g_1, h_{2, n})(x)|$$

$$\leq \frac{1}{|G_{2^k}|} \sum_{\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})} \left( \int_{Q'_{l,j}} |g_1(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |h_{2, n}(y_2)| \, dy_2 \right)$$

$$\lesssim \alpha \cdot \frac{1}{|G_{2^k}|} \sum_{\tilde{Q}_{l,j} \in \mathcal{S}(\partial G_{2^k})} \tilde{Q}_{l,j} \lesssim \alpha \cdot \frac{2^{n/2}(2d-1)k}{2^{2kd}} = 2^{n-k} \alpha.$$

Combining (5.8) with (5.9), we obtain (5.4). This completes all estimates for the second term.

**Estimate for the third term.** The estimate of the third term is similar to that of the second one, the details being omitted.

**Estimate for the last term.** We write

$$b_1 = \sum_{m \in \mathbb{Z}} \sum_{Q_{1,i} \subseteq Q_{2,i}} b_{1,i} =: \sum_{m \in \mathbb{Z}} h_{1,m},$$

and similarly for $b_2$. Define

$$\tilde{h}_{1,n} = \sum_{m \leq n} h_{1,m}, \quad \tilde{\mathbb{1}}_{i,n} = \sum_{m \leq n} \mathbb{1}_{i,m}, \quad i = 1, 2.$$

As above, $A_{2^k}^G(h_{1,m}, h_{2,n})(x) = 0$ if $k \leq m$ or $k \leq n$ since $(x, x) + G_{2^k}$ is disjoint from any of $Q_{1,j} \times Q_{2,j'}$ in the support of $h_{1,m} \otimes h_{2,n}$. These discussions together with Chebyshev’s inequality yields

$$|\{x \in \mathbb{R}^d \setminus \Omega^* : V_q(A_{2^k}^G(b_1, b_2)(x) : k \in \mathbb{Z}) > \alpha/8\}|$$
\begin{align*}
\lesssim & \alpha^2 \int_{\mathbb{R}^d \setminus \Omega^*} \sum_k |A_{2^k}^G(\sum_{m<k} h_{1,m}, \sum_{n<k} h_{2,n})(x)|^2 \, dx \\
\lesssim & \alpha^2 \int_{\mathbb{R}^d \setminus \Omega^*} \sum_k |\sum_{n<k} A_{2^k}^G(\sum_{m<n} h_{1,m}, h_{2,n})(x)|^2 \, dx \\
& + \alpha^2 \int_{\mathbb{R}^d \setminus \Omega^*} \sum_k |\sum_{n<k} A_{2^k}^G(h_{1,m}, \sum_{m<n} h_{2,n})(x)|^2 \, dx \\
= & \alpha^2 \int_{\mathbb{R}^d \setminus \Omega^*} \sum_k |\sum_{m<k} A_{2^k}^G(\tilde{h}_{1,n}, h_{2,n})(x)|^2 \, dx \\
& + \alpha^2 \int_{\mathbb{R}^d \setminus \Omega^*} \sum_k |\sum_{m<k} A_{2^k}^G(h_{1,m}, \tilde{h}_{2,m-1})(x)|^2 \, dx.
\end{align*}

Since these two terms can be handled in the same way, we only treat the first. We need to show

\begin{equation}
\sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^d \setminus \Omega^*} |\sum_{n<k} A_{2^k}^G(\tilde{h}_{1,n}, h_{2,n})(x)|^2 \, dx \lesssim \alpha^{-p}.
\end{equation}

We claim that there exists a constant $c > 1$ such that

\begin{equation}
|A_{2^k}^G(\tilde{h}_{1,n}, h_{2,n})(x)|^2 \lesssim 2^{n-k} \cdot \alpha^2 A_{2^k}^G(\tilde{1}_{1,n}, \tilde{1}_{2,n})(x).
\end{equation}

Assume the claim for the moment. Let us prove (5.10). Integrating both sides and using the trivial estimate $\tilde{1}_{1,n} \leq 1$ yield

\begin{align*}
\|A_{2^k}^G(\tilde{h}_{1,n}, h_{2,n})\|^2 & \lesssim 2^{n-k} \cdot \alpha^2 \|A_{2^k}^G(|\tilde{d}_{1,n}|, |d_{2,n}|)\|_1 \\
& \leq 2^{n-k} \cdot \alpha^2 \frac{1}{|G_{2^k}|} \int_{G_{2^k} \setminus Q_{2,j}} |Q_{2,j}| \, dy_1 dy_2 \\
& = 2^{n-k} \cdot \alpha^2 \sum_{Q_{2,j} \in \mathbb{Q}_n^2} |Q_{2,j}|.
\end{align*}

Then (5.10) follows by utilizing Lemma 2.1, (i) with $S_k = A_{2^k}^G$, $\sigma(j) = 2^{-|j|}$ and $a(u_n, v_n) = \alpha(\sum_{Q_{2,j} \in \mathbb{Q}_n^2} |Q_{2,j}|)^{1/2}$.

Finally, let us prove (5.11). Noting that $n < k$ and using property (vii) of Calderón-Zygmund decomposition $\|b_{1,j}\|_{L^1(Q_{1,j})} \lesssim \alpha^p |Q_{1,j}|$ and $\|b_{2,j}\|_{L^1(Q_{2,j})} \lesssim \alpha^p |Q_{2,j}|$, we have

\begin{align*}
& \left| \int_{(x,x)+G_{2^k}} \tilde{h}_{1,n}(y_1)h_{2,n}(y_2) \, dy_1 dy_2 \right| \\
\leq & \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(\partial G_{2^k})} \int_{\tilde{Q}_{1,j}} |\tilde{h}_{1,n}(y_1)b_{2,j}(y_2)| \, dy_1 dy_2 \\
\leq & \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(\partial G_{2^k})} \left( \int_{Q_{1,j}^l} |\tilde{h}_{1,n}(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |b_{2,j}(y_2)| \, dy_2 \right)
\end{align*}
by Minkowski’s inequality for series, it suffices to show
\[ G \]
Recalling that the bi-subadditive operator
\[ L \]
The first term can be treated using the
\[ 5.1.2. \text{Short variation estimates.} \]
Combining the above two estimates yields (5.13), and hence concluding the proof of
\[ L^1 \times L^{p^2} \rightarrow L^{p, \infty} \] long variation estimate.

5.1.2. Short variation estimates. By the triangle inequality,
\[
\left\{ SV_q(A_i^G)(f_1, f_2) > \alpha \right\}
\leq \left\{ SV_q(A_i^G)(g_1, g_2) > \alpha/4 \right\} + \left\{ SV_q(A_i^G)(g_1, b_2) > \alpha/4 \right\}
\]
\[
+ \left\{ SV_q(A_i^G)(b_1, g_2) > \alpha/4 \right\} + \left\{ SV_q(A_i^G)(b_1, b_2) > \alpha/4 \right\}.
\]
The first term can be treated using the \( L^2 \times L^\infty \rightarrow L^2 \) estimate as in the long variation case.

For the second term, by an argument similar to that used in the proof of long variation estimates for the second term, we only need to show
\[
\sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^d \setminus \Omega^*} \left| \sum_{n \leq k} [A^G_{i+1}(g_1, h_{2,n}) - A^G_{i}(g_1, h_{2,n})] \right|^2 \lesssim \alpha^{2-p}.
\]
Recalling that the bi-subadditive operator \( G_k \) was defined by
\[
G_k(f_1, f_2)(x) = \left( \sum_{I_i \in \mathcal{A}_k} |A^G_{i+1}(f_1, f_2)(x) - A^G_{i}(f_1, f_2)(x)|^2 \right)^{1/2},
\]
by Minkowski’s inequality for series, it suffices to show
\[
\sum_{k \in \mathbb{Z}} \left| \sum_{n \leq k} G_k(g_1, h_{2,n}) \right|^2_{L^2(\mathbb{R}^d \setminus \Omega^*)} \lesssim \alpha^{2-p}.
\]
Applying Lemma 2.1, (i) and using the simple estimate
\[ \|A_{c_{2k+1}}^G(|g_1|, \mathbb{1}_{2,n})\|_1 \lesssim \|g_1\|_\infty \|\mathbb{1}_{2,n}\|_1, \]
matters are reduced to showing the pointwise estimate
\[ G_k(g_1, h_{2,n})(x)^2 \lesssim \alpha^{2-p} 2^{n-k} A_{c_{2k+1}}^G(|g_1|, \mathbb{1}_{2,n})(x), \quad \forall \ x \in \mathbb{R}^d. \]
Since the left hand side is majorized by a constant multiple of \( J_1 + J_2 \), where
\[ J_1 = \sum_{I_i \in \mathcal{I}_k} \frac{1}{|G_{t_i+1}|} \int_{(x,x)+(G_{t_i+1}\setminus G_{t_i})} (g_1 \otimes h_{2,n})^2, \]
and
\[ J_2 = \sum_{I_i \in \mathcal{I}_k} \left( \frac{1}{|G_{t_i}|} - \frac{1}{|G_{t_i+1}|} \right)^2 \int_{(x,x)+G_{t_i}} (g_1 \otimes h_{2,n})^2, \]
it suffices to bound \( J_1 \) and \( J_2 \) by the right hand side of (5.15).
We treat first \( J_2 \) which is easier. Denote
\[ \mathcal{I}(\partial G_{t_i}) = \{ \tilde{Q}_{i,j} = Q'_{1,l} \times Q'_{2,j} \in \mathcal{D}_n \times Q^2_n : (x, x) + \partial G_{t_i} \cap \tilde{Q}_{i,j} \neq \emptyset \}. \]
By similar arguments to that in (5.7), we have
\[
\sqrt{J_2} \lesssim \left( \sum_{I_i \in \mathcal{I}_k} \left( \frac{1}{|G_{t_i}|} - \frac{1}{|G_{t_i+1}|} \right) \right) \sup_{I_i \in \mathcal{I}_k} \left| \int_{(x,x)+G_{t_i}} (g_1 \otimes h_{2,n}) \right| \lesssim \frac{1}{2^{kd}} \sup_{I_i \in \mathcal{I}_k} \sum_{\tilde{Q}_{i,j} \in \mathcal{I}(\partial G_{t_i})} \left( \int_{Q'_{1,l}} |g_1(y_1)| \ dy_1 \right) \cdot \left( \int_{Q_{2,j}} |b_{2,j}(y_2)| \ dy_2 \right) \lesssim \frac{1}{2^{kd}} \sum_{\tilde{Q}_{i,j} \in \mathcal{I}(\partial G_{t_i})} \left( \alpha p \right) |Q'_{1,l}| \cdot (\alpha^{\frac{p}{2}} |Q_{2,j}|) \lesssim \alpha^{\frac{p}{2}} \cdot A_{c_{2k+1}}^G(|g_1|, \mathbb{1}_{2,n})(x),
\]
and
\[
\sqrt{J_2} \lesssim \frac{1}{2^{kd}} \sup_{I_i \in \mathcal{I}_k} \sum_{\tilde{Q}_{i,j} \in \mathcal{I}(\partial G_{t_i})} \left( \int_{Q'_{1,l}} |g_1(y_1)| \ dy_1 \right) \cdot \left( \int_{Q_{2,j}} |b_{2,j}(y_2)| \ dy_2 \right) \lesssim \frac{1}{2^{kd}} \sum_{\tilde{Q}_{i,j} \in \mathcal{I}(\partial G_{t_i})} (\alpha^p |Q'_{1,l}|) \cdot (\alpha^{\frac{p}{2}} |Q_{2,j}|) \lesssim \alpha \cdot \frac{2^{k(2d-1)} \cdot 2^n}{2^{kd}} = \alpha 2^{n-k}.
\]
Combining the above two estimates, we obtain (5.15).
Combining these two inequalities, we obtain the desired bound for $J_5.2$. estimate of the second term. The other terms can be similarly treated, the details being omitted.

For $J_1$, we have

$$
sup_{I_i \in \mathcal{Y}_k} \left| \frac{1}{|G_{t_i+1}|} \int_{(x,x)+(G_{t_i+1}\setminus G_{t_i})} g_1(y_1)h_{2,n}(y_2) \, dy_1 \, dy_2 \right|$$

$$\leq \frac{1}{2^{2kd}} \sup_{I_i \in \mathcal{Y}_k} \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(\partial(G_{t_i+1}\setminus G_{t_i}))} \left( \int_{Q_{1,j}} |g_1(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |h_{2,n}(y_2)| \, dy_2 \right)$$

$$\lesssim \frac{1}{2^{2kd}} \sum_{I_i \in \mathcal{Y}_k} \int_{(x,x)+(G_{t_i+1}\setminus G_{t_i})} |g_1(y_1)h_{2,n}(y_2)| \, dy_1 \, dy_2$$

$$\leq \frac{1}{2^{2kd}} \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(G_{2k+1})} \left( \int_{\tilde{Q}_{1,j}} |g_1(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |h_{2,n}(y_2)| \, dy_2 \right)$$

$$\lesssim \alpha^{2p_2} \frac{1}{2^{2kd}} \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(G_{2k+1})} \left( \int_{\tilde{Q}_{1,j}} |g_1(y_1)| \, dy_1 \right) \cdot \left( \int_{Q_{2,j}} |h_{2,n}(y_2)| \, dy_2 \right)$$

$$= \alpha^{2p_2} \frac{1}{2^{2kd}} \sum_{\tilde{Q}_{1,j} \in \mathcal{F}(G_{2k+1})} \int_{\tilde{Q}_{1,j}} |g_1(y_1)1_{2,n}(y_2)| \, dy_1 \, dy_2$$

$$\leq \alpha^{2p_2} \frac{1}{2^{2kd}} \int_{G_{2k+1}} |g_1(y_1)1_{2,n}(y_2)| \, dy_1 \, dy_2$$

$$\approx \alpha^{2p_2} A^G_{c2k+1}(|g_1|, 1_{2,n})(x).$$

Combining these two inequalities, we obtain the desired bound for $J_1$. This completes the estimate of the second term. The other terms can be similarly treated, the details being omitted.

5.2. $L^1 \times L^\infty \to L^{1,\infty}$ estimate. Since the proof is similar to that in the case $1 < p_2 < \infty$, we only sketch the proof of the long variation estimate. Assume that $f_1 \in L^1(\mathbb{R}^d)$ and $f_2 \in L^\infty_c(\mathbb{R}^d)$ with $\|f_1\|_1 = \|f_2\|_\infty = 1$. We need to show

$$|\{\mathcal{L}V_q(A^G_q)(f_1, f_2) > \alpha\}| \lesssim \alpha^{-1}.$$
We adopt the same notations as used in previous subsection. By the Calderón-Zygmund decomposition \( f_1 = g_1 + b_1 \), we have
\[
|\{ \mathcal{L}V_q(A^G_t)(f_1, f_2) > \alpha \}| \leq |\{ \mathcal{L}V_q(A^G_t)(g_1, f_2) > \alpha/2 \}| + |\{ \mathcal{L}V_q(A^G_t)(b_1, f_2) > \alpha/2 \}|
\]

By the \( L^2 \times L^\infty \rightarrow L^2 \) boundedness of \( \mathcal{L}V_q(A^G_t) \) and the estimate \( \|g_1\|_2 \leq \|g_1\|_\infty \|g_1\|_1 \lesssim \alpha \), we have
\[
|\{ \mathcal{L}V_q(A^G_t)(g_1, f_2) > \alpha/2 \}| \lesssim \alpha^{-2} \| \mathcal{L}V_q(A^G_t)(g_1, f_2) \|_2^2 \\
\lesssim \alpha^{-2} \|g_1\|_2^2 \|f_2\|_\infty^2 \lesssim \alpha^{-2} \cdot \alpha = \alpha^{-1}.
\]

For the second term, by an argument similar to that used in Section 5.1.1, we can reduce the matters to showing
\[
(5.16) \quad \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^d, \Omega^*} |A^G_{2k}(\sum_{n<k} h_{1,n}, f_2)(x)|^2 \, dx \lesssim \alpha.
\]

Assuming \( n < k \), we claim that
\[
(5.17) \quad |A^G_{2k}(h_{1,n}, f_2)(x)|^2 \lesssim 2^{n-k} \cdot \alpha^2 |A^G_{2k}(1_{1,n}, |f_2|)(x)|, \quad \forall \, x \in \mathbb{R}^d.
\]

Assume this claim holds for the moment, let us show (5.16). Integrating both sides and using the hypothesis that \( \|f_2\|_{\infty} = 1 \), we get
\[
\|A^G_{2k}(h_{1,n}, f_2)\|_2^2 \lesssim 2^{n-k} \cdot \alpha^2 \|A^G_{2k}(1_{1,n}, |f_2|)\|_1 \lesssim 2^{n-k} \cdot \alpha^2 \sum_{j:Q_{1,j} \in \mathcal{Q}_n^1} |Q_{1,j}|.
\]

Then (5.16) follows by applying Lemma 2.1 with \( S_k = A^G_{2k}, \, \sigma(j) = 2^{-\frac{j}{2}}, \) and \( a(h_{1,n}, f_2) = \alpha(\sum_{j:Q_{1,j} \in \mathcal{Q}_n^1} |Q_{1,j}|)^{1/2} \).

It suffices to establish claim (5.17). By arguments similar to those given in (5.7) and (5.9) (with the good function \( g_2 \) replaced by the function \( f_2 \) itself in the current case), one can derive that
\[
\frac{1}{|G_{2k}|} \left| \int_{(x,x)+G_{2k}} h_{1,n}(y_1) f_2(y_2) \, dy_1 dy_2 \right| \lesssim \frac{\alpha}{|G_{2k}|} \int_{(x,x)+G_{2k}} 1_{1,n}(y_1) |f_2(y_2)| \, dy_1 dy_2
\]
and
\[
\frac{1}{|G_{2k}|} \left| \int_{(x,x)+G_{2k}} h_{1,n}(y_1) f_2(y_2) \, dy_1 dy_2 \right| \lesssim 2^{n-k} \cdot \alpha.
\]

Combining these two estimates yields (5.17), and hence concludes the proof of long variation inequality.

The short variation estimate can be proved analogously, the details being omitted.

6. Restricted weak type estimates and concluding the proof of Theorem 1.1

In this section, we will conclude our proof of Theorem 1.1 by multilinear real interpolation of Grafakos-Kalton [21] based on the previous estimates. To begin with, let us recall some definitions.
Denote by $S(\mathbb{R}^d)$ the space of functions of the form $\sum_{i=1}^N c_i \mathbb{1}_{E_i}$, where each measurable subset $E_i$ of $X$ has finite measure. An operator $T$ defined on $S(\mathbb{R}^d) \times \cdots \times S(\mathbb{R}^d)$ and taking values in the set of complex-valued measurable functions on $\mathbb{R}^d$ is called multilinear if for all $1 \leq j \leq m$, all $f_j, g_j$ in $S(\mathbb{R}^d)$ and all $\lambda \in \mathbb{C}$ the statements
\[
|T(f_1, \ldots, \lambda f_j, \ldots, f_m)| = |\lambda| |T(f_1, \ldots, f_j, \ldots, f_m)|,
\]
\[
|T(\ldots, f_j + g_j, \ldots)| \leq |T(\ldots, f_j, \ldots)| + |T(\ldots, g_j, \ldots)|
\]
hold almost everywhere.

**Definition 6.1.** Let $0 < p_1, p_2, \cdots, p_m, p \leq \infty$. We say that an $m$-sublinear operator $T$ is of restricted weak type $(p_1, \ldots, p_m, p)$ if there is a constant $C = C(p_1, \ldots, p_m, p)$ such that for all measurable subsets $A_1, \ldots, A_m$ of finite measure we have
\[
\sup_{\lambda > 0} \lambda \{ x \in \mathbb{R}^d : |T(1_{A_1}, \ldots, 1_{A_m})| > \lambda \} \frac{1}{\lambda^p} \leq C |A_1|^{\frac{1}{p_1}} \cdots |A_m|^{\frac{1}{p_m}}
\]
when $p < \infty$ and
\[
\|T(1_{A_1}, \ldots, 1_{A_m})\|_\infty \leq C |A_1|^{\frac{1}{p_1}} \cdots |A_m|^{\frac{1}{p_m}}
\]
when $p = \infty$.

**Proof of Theorem 1.1.** In order to prove the main result by using multilinear real interpolation (cf. [21]), we need to show that $V_q(A_t^G : t > 0)$ is of restricted weak types $(1, \infty, 1), (\infty, 1, 1), (1, 1, 1/2), (\infty, s, s)$ and $(s, \infty, s)$, for any $2 < s < \infty$. The first three follow from the corresponding estimates obtained in previous sections by taking $f_1 = \mathbb{1}_A$ and $f_2 = \mathbb{1}_B$. It remains to show the restricted weak types $(\infty, s, s)$ and $(s, \infty, s)$ for any $2 < s < \infty$.

For any fixed $f_1 \in L^\infty \cap L^2$, define a sublinear operator $T_{f_1}$ by
\[
T_{f_1}(g)(x) = V_q(A_t^G(f_1, g) : t > 0)(x), \quad \forall g \in L^p.
\]
From the boundedness results of $V_q(A_t^G : t > 0)$ proved in previous sections, it follows that $T_{f_1}$ is bounded from $L^2$ to $L^2$, from $L^\infty$ to $BMO_d$, and from $L^1$ to $L^{1,\infty}$, and the operator norms are both bounded by a constant multiple of $\|f_1\|_\infty$. By classical Marcinkiewicz interpolation for sublinear operators, $T_{f_1}$ is bounded on $L^s$ for all $1 < s < \infty$, with operator norm at most $C_{p,d}\|f_1\|_\infty$. This implies that
\[
\|V_q(A_t^G(f_1, f_2) : t > 0)\|_s = \|T_{f_1}(f_2)\|_s \lesssim \|f_1\|_\infty \|f_2\|_s
\]
for any $1 < s < \infty$, any $f_1 \in L^\infty \cap L^2$, $f_2 \in L^s$. Applying this with $f_1 = \mathbb{1}_A$ and $f_2 = \mathbb{1}_B$, we see that $V_q(A_t^G : t > 0)$ is of restricted weak type $(\infty, s, s)$ for any $1 < s < \infty$. By symmetry, $V_q(A_t^G : t > 0)$ is of restricted weak type $(s, \infty, s)$ for any $1 < s < \infty$.

Finally, for $1 < p_1, p_2 < \infty$ and $1/p = 1/p_1 + 1/p_2$, $(p_1, p_2, p)$ lies in the convex hull of $(1, \infty, 1), (\infty, 1, 1), (1, 1, 1/2), (\infty, s, s)$ and $(s, \infty, s)$ for sufficiently large constant $s$. The required $L^{p_1} \times L^{p_2} \rightarrow L^p$ bound then follows by applying the multilinear real interpolation in Lemma 2.7. This concludes the proof of Theorem 1.1. \hfill \Box

We conclude this section by showing that $V_q(A_t^G(f_1, f_2))$ is, in general, not of the restricted weak type $(\infty, \infty, \infty)$. This means that the $L^\infty \times L^\infty \rightarrow BMO_d$ estimate proved in Section 4 cannot be strengthened to $L^\infty \times L^\infty \rightarrow L^\infty$ estimate even for $f_1, f_2$ being characteristic functions of measurable sets of finite measures. We do this in the case
Let $G = \tilde{B}(0, 1) := \tilde{B}$, the unit ball of $\mathbb{R}^{2d}$. To this end, we construct two sequences $(E_n)_n$ and $(F_n)_n$ of measurable sets of finite measures in $\mathbb{R}^d$ such that

$$\|V_q(A_t^\tilde{B} : t > 0)(1_{E_n}, 1_{F_n})\|_\infty \gtrsim n, \quad \forall n \in \mathbb{N}.$$ 

Take $\alpha > 1$ large enough so that $|[(B_1)^c \times \mathbb{R}^d] \cap \tilde{B}_\alpha| > \frac{4}{5} |\tilde{B}_\alpha|$. There exists $\epsilon_0 > 0$ such that $|[(B_1)^c \times \mathbb{R}^d] \cap [(x, x) + \tilde{B}_\alpha]| > \frac{3}{4} |\tilde{B}_\alpha|$ for all $x \in B_{\epsilon_0}$. For each $n \in \mathbb{N}$, define

$$E_n = \bigcup_{i=1}^{n} B_{\alpha^{2n+1}} \setminus B_{\alpha^{2n}} \quad \text{and} \quad F_n = B_{\alpha^{2n+2}}.$$ 

Then for $1 \leq i \leq 2n + 1$ and any $x \in B_{\epsilon_0}$,

$$A_{\alpha^i} \tilde{B}(1_{E_n}, 1_{F_n})(x) = \frac{|[(x, x) + \tilde{B}_\alpha] \cap [E_n \times F_n]|}{|B_\alpha|} \begin{cases} > \frac{3}{4}, & i \text{ odd}; \\ < \frac{1}{4}, & i \text{ even}. \end{cases}$$

Let $t_i = \alpha^i$ for $i \in \mathbb{N}$. Then for $0 \leq i \leq 2n$ and $x \in B_{\epsilon_0}$,

$$\left| A_{t_{i+1}}^\tilde{B}(1_{E_n}, 1_{F_n}) - A_{t_i}^\tilde{B}(1_{E_n}, 1_{F_n}) \right| > \frac{1}{2},$$

which implies

$$V_q(A_t^\tilde{B}(1_{E_n}, 1_{F_n}) : t > 0)(x) > 2^{-q+1} n, \quad \forall x \in B_{\epsilon_0}.$$ 

This shows that $V_q(A_t^\tilde{B} : t > 0)$ is not of restricted weak type $(\infty, \infty, \infty)$.

7. Bounds for new bilinear square functions involving conditional expectation

Recall that the square function $L$ is defined by

$$L(f_1, f_2)(x) = \left( \sum_{k \in \mathbb{Z}} |L_k(f_1, f_2)(x)|^2 \right)^{\frac{1}{2}},$$

where

$$L_k(f_1, f_2)(x) = A_{2k}^G(f_1, f_2)(x) - \mathcal{E}_k f_1(x) \cdot \mathcal{E}_k f_2(x).$$

As a byproduct of our arguments, we obtain the $L^{p_1} \times L^{p_2} \to L^p$ boundedness of the bilinear square function $L$, which is the content of the following

**Theorem 7.1.** The square function $L$ is bounded from $L^{p_1} \times L^{p_2}$ to $L^p$ for all $1 < p_1, p_2 < \infty$ and $1/p = 1/p_1 + 1/p_2$, bounded from $L_\infty^c \times L_\infty^c$ to $BMO_\alpha$, and bounded from $L^1 \times L^{p_2}$ to $L^{p, \infty}$ for all $1 \leq p_2 \leq \infty$ and $1/p = 1 + 1/p_2$.

**Proof.** By arguments similar to that given in the proof of Theorem 1.1, it suffices to show the $L^2 \times L^\infty \to L^2$, $L^\infty \times L^\infty \to BMO$, and $L^1 \times L^{p_2} \to L^{p, \infty}$ bounds of $L$. The $L^2 \times L^\infty \to L^2$ has been established in Section 3.
Let us now show the \( L^\infty \times L^\infty \to BMO \) bound. Following similar arguments as in Section 4, we have

\[
\frac{1}{|Q|} \int_Q |\mathcal{L}(f_1, f_2) - a_Q| \leq \frac{1}{|Q|} \int_Q |\mathcal{L}(f_1^0, f_2^0)| + \frac{1}{|Q|} \int_Q |\mathcal{L}(f_1^\infty, f_2^\infty)| \\
+ \frac{1}{|Q|} \int_Q |\mathcal{L}(f_1^\infty, f_2^0)| + \frac{1}{|Q|} \int_Q |\mathcal{L}(f_1^\infty, f_2^\infty) - \mathcal{L}(f_1^\infty, f_2^\infty)(c_Q)| \\
=: II_1 + II_2 + II_3 + II_4.
\]

The required bounds for \( II_1, II_2 \) and \( II_3 \) follows easily from the \( L^\infty \times L^2 \to L^2 \) boundedness of \( \mathcal{L} \) as before. The estimate for \( II_4 \) will follow if we show that, for any \( x \in Q \),

\[
\left( \sum_{k \in \mathbb{Z}} \left| \left[ A_{2^k}^{G}(f_1^\infty, f_2^\infty)(x) - \mathcal{E}_k f_1^\infty(x) \cdot \mathcal{E}_k f_2^\infty(x) \right] - \left[ A_{2^k}^{G}(f_1^\infty, f_2^\infty)(c_Q) - \mathcal{E}_k f_1^\infty(c_Q) \cdot \mathcal{E}_k f_2^\infty(c_Q) \right] \right|^2 \right)^{1/2} \lesssim \| f_1 \|_\infty \| f_2 \|_\infty,
\]

where the implicit constant is independent of \( x \) and \( Q \).

We consider two cases as before. If \( 2^k < \ell(Q) \), then for any \( x \in Q \), both \( \mathcal{E}_k f_1^\infty \) and \( \mathcal{E}_k f_2^\infty \) are supported in \( \mathbb{R}^d \setminus Q \), so

\[
\mathcal{E}_k f_1^\infty(x) \cdot \mathcal{E}_k f_2^\infty(x) = \mathcal{E}_k f_1^\infty(c_Q) \cdot \mathcal{E}_k f_2^\infty(c_Q) = 0.
\]

If \( 2^k \geq \ell(Q) \), then in this case, there exists an atom in \( \mathcal{D}_k \) that contains \( Q \), so \( \mathcal{E}_k f_1^\infty(x) \cdot \mathcal{E}_k f_2^\infty(x) = \mathcal{E}_k f_1^\infty(c_Q) \cdot \mathcal{E}_k f_2^\infty(c_Q) \).

In both cases, the conditional expectations on the left side of (7.1) are equal to zero, and hence (7.1) follows from (4.1). This completes the proof of the \( L^\infty \times L^\infty \to BMO \) bound of \( \mathcal{L} \).

Finally, let us show the \( L^1 \times L^{p^2} \to L^{p,\infty} \) bounds of \( \mathcal{L} \). We will see that again the conditional expectation will play no role and the estimates follows from corresponding estimates in Section 4. By the Calderón-Zygmund decomposition, we have

\[
|\{ \mathcal{L}(f_1, f_2)(x) > \alpha \}| \leq |\{ \mathcal{L}(g_1, g_2)(x) > \alpha/4 \}| + |\{ \mathcal{L}(g_1, b_2)(x) > \alpha/4 \}| + |\{ \mathcal{L}(b_1, g_2)(x) > \alpha/4 \}| + |\{ \mathcal{L}(b_1, b_2)(x) > \alpha/4 \}|,
\]

The first term can be treated as in Section 4 by using the \( L^2 \times L^\infty \to L^2 \) bound of \( \mathcal{L} \), together with the properties of good functions \( g_1 \) and \( g_2 \).

To estimate the second term, by (5.1), it suffices to prove

\[
|\{ x \in \mathbb{R}^d \setminus \Omega^* : \sum_{k} | A_{2^k}^{G}(g_1, b_2)(x) - \mathcal{E}_k g_1(x) \cdot \mathcal{E}_k b_2(x) |^2 \}^{1/2} > \alpha/8 \} | \lesssim \alpha^{-p}.
\]

Write

\[
b_2 = \sum_{n \in \mathbb{Z}} \sum_{j: Q_{2,j} \in \mathcal{Q}^2_n} b_{2,j} = \sum_{n \in \mathbb{Z}} h_{2,n}.
\]

We shall show that for \( x \notin \Omega^* \), \( \mathcal{E}_k b_2(x) = 0 \) for every \( k \). Indeed, if \( k \leq n \), then \( \mathcal{E}_k h_{2,n}(x) = 0 \) since the atom of \( \mathcal{D}_k \) containing \( x \) is disjoint from the support of \( h_{2,n} \); if \( n < k \), for each \( Q_{2,j} \in \mathcal{Q}^2_n \), there exists an atom of \( \mathcal{D}_k \) containing \( Q_{2,j} \). Then \( \mathcal{E}_k b_{2,j} = 0 \) due to the
cancellation condition of $b_{2,j}$, and hence $\mathcal{E}_kh_{2,n}(x) = 0$. Now (7.2) is equivalent to (5.2) established in Section 5. The other terms can be treated in the same way.

This concludes the proof of the $L^1 \times L^{p_2} \to L^{p,\infty}$ bounds of $\mathcal{L}$, and hence Theorem 7.1.

8. Applications

8.1. Discrete bilinear averaging operators. Let $1 < p_1, p_2 < \infty$. Consider the discrete bilinear averaging operator defined by

$$A_t^{G,d}(f_1, f_2)(j) = \frac{1}{\sharp(G_t \cap \mathbb{Z}^{2d})} \sum_{(k,m) \in G_t \cap \mathbb{Z}^{2d}} f_1(j-k) f_2(j-m), \quad t > 0, \ j \in \mathbb{Z}^d$$

for $f_1 \in \ell^{p_1}(\mathbb{Z}^d)$ and $f_2 \in \ell^{p_2}(\mathbb{Z}^d)$, where $\sharp(S)$ is the cardinality of the set $S$.

By arguments similar to that used in the proof of Theorem 1.1, we can prove

Corollary 8.1. Let $2 < q < \infty$. Then for any $1 < p_1, p_2 < \infty$, $1/2 < p < \infty$ with $1/p = 1/p_1 + 1/p_2$, there exists a constant $C_{p,q}$ such that

$$\|V_q((A_t^{G,d})_{t>0})(f_1, f_2)\|_{\ell^q} \leq C_{p,q} \|f_1\|_{\ell^{p_1}} \|f_2\|_{\ell^{p_2}}.$$ 

Remark 8.1. This result provides a bilinear analogue of the result obtained by Jones-Rosenblatt-Wierdl [32].

8.2. Bilinear averaging operators of Demeter-Tao-Thiele. Demeter, Tao, and Thiele [12] studied a general class of multilinear averaging operators and proved maximal inequalities for them. We consider a special case where the averaging operator takes the form

$$M_{\Lambda,t}(f_1, f_2)(x) = \frac{1}{t^{2d}} \int_{\mathbb{R}^d} f_1(x + \lambda_1 u_1 + \lambda_2 u_2) f_2(x + \lambda_2 u_1 + \lambda_2 u_2) du_1 du_2.$$ 

Here $\Lambda = \{\lambda_{i,j}\}_{2 \times 2}$ is assumed to be nonsingular.

This operator can be expressed as an averaging operator over convex bodies. Indeed, let $\Gamma$ be the inverse of $\Lambda$, which is also nonsingular. Then the following set

$$G_\Gamma = \{(y_1, y_2) \in \mathbb{R}^{2d} : |\gamma_{1,1} y_1 + \gamma_{1,2} y_2| < 1, |\gamma_{2,1} y_1 + \gamma_{2,2} y_2| < 1\}$$

is a symmetric convex body in $\mathbb{R}^{2d}$, whose “volume” is given by the absolute value of its determinant $|\det(\Gamma)|$. Changing the variable we have

$$M_{\Lambda,t}(f_1, f_2)(x) = \frac{1}{|\det(\Gamma)| \cdot t^{2d}} \int_{G_\Gamma} f_1(x + y_1) f_2(x + y_2) dy_1 dy_2$$

$$= \frac{1}{|(G_\Gamma)_t|} \int_{(G_\Gamma)_t} f_1(x + y_1) f_2(x + y_2) dy_1 dy_2 = A_t^{G_\Gamma}(f_1, f_2)(x).$$

Then the following corollary follows immediately from Theorem 1.1.

Corollary 8.2. Let $2 < q < \infty$. If $\Lambda$ is nonsingular, then for any $1 < p_1, p_2 < \infty$, $1/2 < p < \infty$ with $1/p = 1/p_1 + 1/p_2$, there exists a constant $C_{p,q}$ such that

$$\|V_q(M_{\Lambda,t})_{t>0})(f_1, f_2)\|_{L^p(\mathbb{R}^d)} \leq C_{p,q} \|f_1\|_{L^{p_1}(\mathbb{R}^d)} \|f_2\|_{L^{p_2}(\mathbb{R}^d)}.$$
8.3. Ergodic bilinear averaging operators. Let \((\Omega, \mu)\) be a measurable space. For \(x \in \mathbb{R}^d\), let \(T^x : \Omega \to \Omega\) be a family of measure preserving transformations on \((\Omega, \mu)\) which satisfy the following semigroup properties: \(T^{x+y} = T^x \cdot T^y\), \(\forall x, y \in \mathbb{R}^d\).

For \(t > 0\), \(f_1, f_2 \in L^1_{\text{loc}}(\Omega)\), define a family of ergodic bilinear averaging operators by:

\[
\mathcal{A}_t(f_1, f_2)(\omega) = \frac{1}{|G_t|} \int_{G_t} f_1(T^x \omega) \cdot f_2(T^y \omega) \, dx \, dy, \quad \omega \in \Omega.
\]

**Corollary 8.3.** Let \(2 < q < \infty\). Then for any \(1 < p_1, p_2 < \infty\), \(1/2 < p < \infty\) with \(1/p = 1/p_1 + 1/p_2\), there exists a constant \(C_{p,q}\) such that

\[
\|V_q((\mathcal{A}_t)_{t>0})(f_1, f_2)\|_{L^p(\Omega)} \leq C_{p,q}\|f_1\|_{L^{p_1}(\Omega)}\|f_2\|_{L^{p_2}(\Omega)}.
\]

**Proof.** The proof uses the transference principle. For a family \((f_t)_{t>0}\) of functions, we use the notation \(\|f_t\|_{V_q(t \in (0,R))} := V_q((f_t) : t \in (0,R))\). For any fixed \(R > 0\), we have

\[
\left\| \left( \frac{1}{|G_t|} \int_{G_t} f_1(T^x \omega) f_2(T^y \omega) \, dx \, dy \right) \right\|_{V_q(t \in (0,R))}^p \leq \left\| \frac{1}{|G_t|} \int_{G_t} f_1(T^{xz} \omega) f_2(T^{yz} \omega) \, dx \, dy \right\|_{V_q(t \in (0,R))}^p \left( \int_{|x| \leq J} |\tilde{z}| \left( \int_{G_t} \left| f_1(T^{xz} \cdot) f_2(T^{yz} \cdot) \right| (\omega) \, dx \, dy \right) \right)^{\frac{p}{p}} \, dz.
\]

Notice that \(\|\tilde{z} \circ id\|_{L^p(\Omega, V_q(t \in (0,R)))} \to L^p(\Omega, V_q(t \in (0,R))) = 1\).

Therefore the last term in (8.1) is equal to

\[
\frac{1}{Jd} \int_{|x| \leq J} \left\| \frac{1}{|G_t|} \int_{G_t} \left| f_1(T^{xz} \cdot, f_2(T^{yz} \cdot) \right| (\omega) \, dx \, dy \right\|_{L^p(\Omega, V_q(t \in (0,R)))}^p \, dz
\]

\[
= \frac{1}{Jd} \int_{|x| \leq J} \left\| A_t[f_1(T^{xz} \cdot, f_2(T^{yz} \cdot)](\omega) \right\|_{L^p(\Omega, V_q(t \in (0,R)))}^p \, dz
\]

\[
= \frac{1}{Jd} \int_{|x| \leq J} \left\| A^G_t[F_1(\cdot, \omega), F_2(\cdot, \omega)](\omega) \right\|_{L^p(\Omega, V_q(t \in (0,R)))}^p \, d\omega,
\]

where \(F_1(z, \omega) = f_1(T^{-z} \omega)1_{|z| \leq J+R}\) and \(F_2(z, \omega) = f_2(T^{-z} \omega)1_{|z| \leq J+R}\).

Applying Theorem 1.1, the last term above is bounded by

\[
\frac{1}{Jd} \int_{\Omega} \left( \int_{\mathbb{R}^d} |F_1(z, \omega)|^{p_1} \, dz \, d\omega \right)^{\frac{p}{p_1}} \left( \int_{\mathbb{R}^d} |F_2(z, \omega)|^{p_2} \, dz \, d\omega \right)^{\frac{p}{p_2}} \leq \frac{1}{Jd} \left( \int_{|x| \leq J+R} \int_{\Omega} |f_1(T^{-z} \omega)|^{p_1} \, d\omega \, dz \right)^{\frac{p}{p_1}} \left( \int_{|x| \leq J+R} \int_{\Omega} |f_2(T^{-z} \omega)|^{p_2} \, d\omega \, dz \right)^{\frac{p}{p_2}}
\]
\[ \lesssim \frac{(J + R)^d}{J^d} \| f_1 \|_{L^p_2(\Omega)}^{p} \| f_2 \|_{L^p_2(\Omega)}^{p}. \]

Since the estimates holds for all \( J > 0 \), letting \( J \to +\infty \), we obtain
\[
\frac{1}{J^d} \int_{\Omega} \| F_1(\cdot, \omega) \|_{L^p_{p_1}(\mathbb{R}^d)}^{p} \| F_2(\cdot, \omega) \|_{L^p_{p_2}(\mathbb{R}^d)}^{p} d\omega \lesssim \| f_1 \|_{L^p_2(\Omega)}^{p} \| f_2 \|_{L^p_2(\Omega)}^{p}.
\]

Putting all together we have for any \( R > 0 \),
\[
\left\| \frac{1}{|G_t|} \int_{G_t} f_1(T^x \omega) f_2(T^y \omega) dx dy \right\|_{L^p_{p}(\Omega)}^{p} \lesssim \| f_1 \|_{L^p_2(\Omega)}^{p} \| f_2 \|_{L^p_2(\Omega)}^{p}.
\]

Letting \( R \to +\infty \) and applying monotone convergence theorem, Corollary 8.3 follows. \( \square \)

Acknowledgement. The first author is partially supported by the NSF of China(Nos. 11571160, 11871096), the second author partially supported by the NSF of China(Nos. 11601396, 11431011), and the third author partially supported by NSF of China(No. 11671397).

References

[1] J. Bourgain, On high dimensional maximal functions associated to convex bodies. Amer. J. Math., 108 (1986), 1467-1476.
[2] J. Bourgain, On \( L^p \) bounds for maximal functions associated to convex bodies in \( \mathbb{R}^n \). Israel J. Math., 54 (1986), 257-265.
[3] J. Bourgain, Pointwise ergodic theorems for arithmetic sets. Publ. Math. IHES 69 (1989), 5-41.
[4] J. Bourgain, M. Mirek, E.M. Stein and B. Wróbel, Dimension-free estimates for discrete Hardy-Littlewood averaging operators over the cubes in \( \mathbb{Z}^d \). Amer. J. Math. 141 (2019), 857-905.
[5] J. Bourgain, M. Mirek, E.M. Stein and B. Wróbel, On dimension-free variational inequalities for averaging operators in \( \mathbb{R}^d \). Geom. Funct. Anal. 28 (2018), 58-99.
[6] J. Campbell, R. Jones, K. Reinhold and M. Wierdl, Oscillation and variation for the Hilbert transform. Duke Math. J. 105 (2000), 59-83.
[7] J. Campbell, R. Jones, K. Reinhold and M. Wierdl, Oscillation and variation for singular integrals in higher dimensions. Trans. Amer. Math. Soc. 355 (2003), 2115-2137.
[8] A. Carbery. An almost-orthogonality principle with applications to maximal functions associated to convex bodies. Bull. Amer. Math. Soc., 14 (1986), 269-274.
[9] Y. Chen, Y. Ding, G. Hong and H. Liu, Weighted jump and variational inequalities for rough operators. J. Funct. Anal. 274 (2018), 2446-2475.
[10] Y. Chen, Y. Ding, G. Hong and J. Xiao, Some jump and variational inequalities for the Calderón commutators and related operators. Preprint available at arXiv:1709.03127v1.
[11] M. Cowling and M. Leinert, Pointwise convergence and semigroups acting on vector-valued functions. Bull. Aust. Math. Soc. 84 (2011), 44-48.
[12] C. Demeter, T. Tao, and C. Thiele, Maximal multilinear operators. Trans. Amer. Math. Soc. 360 (2008), 4989-5042.
[13] Y. Ding, G. Hong and H. Liu, Jump and variational inequalities for rough operators. J. Fourier Anal. Appl. 23 (2017), 679-711.
[14] Y. Do, C. Muscalu and C. Thiele, Variational estimates for paraproducts. Rev. Mat. Iberoam. 28 (2012), 857-878.
[15] Y. Do, R. Oberlin and E.A. Palsson, Variation-norm and fluctuation estimates for ergodic bilinear averages. Indiana Univ. Math. J. 66 (2017), 55-99.
[16] C. Fefferman and E.M. Stein, Some maximal inequalities. Amer. J. Math. 93 (1971), 107-115.
[17] G. Fendler, Dilations of one parameter semigroups of positive contractions on \( L^p \)-spaces. Canad. J. Math. 49 (1997), 736-748.
[18] J. García-Cuerva and J.L. Rubio de Francia, Weighted Norm Inequalities and Related Topics. North-Holland Publishing Co, Amsterdam (1985).
[19] T.A. Gillespie and J.L. Torrea, Dimension free estimates for the oscillation of Riesz transforms. Israel J. Math. 141 (2004), 125-144.
[20] L. Grafakos, Modern Fourier Analysis. Third edition. Graduate Texts in Mathematics, 249. Springer, New York, 2014.
[21] L. Grafakos and N. Kalton, Some remarks on multilinear maps and interpolation. Math. Ann. 319 (2001), 151-180.
[22] L. Grafakos, A. Miyachi, and N. Tomita, On multilinear Fourier multipliers of limited smoothness. Canad. J. Math. 65 (2013), no. 2, 299-330.
[23] L. Grafakos and R.H. Torres, Maximal operator and weighted norm inequalities for multilinear singular integrals. Indiana Univ. Math. J. 51 (2002), 1261-1276.
[24] L. Grafakos and R.H. Torres, Multilinear Calderón-Zygmund theory. Adv. Math. 165 (2002), 124-164.
[25] G. Hong, The behaviour of square functions from ergodic theory in $L^\infty$. Proc. Amer. Math. Soc. 143 (2015), 4797-4802.
[26] G. Hong and T. Ma, Vector valued $q$-variation for differential operators and semigroups I. Math. Z. 286 (2017), 89-120.
[27] G. Hong and T. Ma, Vector valued $q$-variation for ergodic averages and analytic semigroups. J. Math. Anal. Appl. 437 (2016), 1084-1100.
[28] K. Hughes, B. Krause and B. Trojan, The maximal function and square function control the variation: an elementary proof. Proc. Amer. Math. Soc. 144 (2016), 3583-3588.
[29] T.P. Hytönen, M.T. Lacey and I. Parissis, A variation norm Carleson theorem for vector-valued Walsh-Fourier series. Rev. Mat. Iberoam. 30 (2014), 979-1014.
[30] R.L. Jones, R. Kaufman, J.M. Rosenblatt and M. Wierdl, Oscillation in ergodic theory. Ergod. Theory Dyn. Syst. 18 (1998), 889-935.
[31] R. Jones, J. Rosenblatt and M. Wierdl, Oscillation inequalities for rectangles. Proc. Amer. Math. Soc 129 (2000), 1349-1358.
[32] R. Jones, J. Rosenblatt and M. Wierdl, Oscillation in ergodic theory: higher dimensional results. Israel J. Math. 135 (2003), 1-27.
[33] R. Jones, A. Seeger and J. Wright, Strong variational and jump inequalities in harmonic analysis. Trans. Amer. Math. Soc. 360 (2008), 6711-6742.
[34] R.L. Jones and G. Wang, Variation inequalities for the Fejér and Poisson kernels. Trans. Amer. Math. Soc. 356 (2004), 4493-4518.
[35] B. Krause and P. Zorin-Kranich, Weighted and vector-valued variational estimates for ergodic averages. arXiv:1409.7120.
[36] C. Le Merdy and Q. Xu, Strong $q$-variation inequalities for analytic semigroups. Ann. Inst. Fourier 62 (2012), 2069-2097.
[37] D. Lépingle, La variation d’ordre $p$ des semi-martingales. Z. Wahrscheinlichkeitstheorie Verwandte Geb. 36 (1976), 295-31.
[38] A. Lerner, S. Ombrosi, C. Pérez, R. Torres and R. Trujillo-González, New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory. Forum Math. 27 (2015), 77-97.
[39] 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.
[40] T. Ma, J.L. Torrea and Q. Xu, Weighted variation inequalities for differential operators and singular integrals in higher dimensions. Sci. China Math. 60 (2017), 1419-1442.
[41] A. Mas, Variation for singular integrals on Lipschitz graphs: $L^p$ and endpoint estimates. Trans. Amer. Math. Soc. 365 (2013), 5759-5781.
[42] A. Mas and X. Tolsa, Variation and oscillation for singular integrals with odd kernel on Lipschitz graphs. Proc. Lond. Math. Soc. 105 (2012), 49-86.
[43] A. Mas and X. Tolsa, Variation for the Riesz transform and uniform rectifiability. J. Eur. Math. Soc. 16 (2014), 2267-2321.
[45] M. Mirek, E.M. Stein and B. Trojan, $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: variational estimates. Invent. Math. 209 (2017), 665-748.

[46] M. Mirek, E. M. Stein, and B. Trojan, $\ell^p(\mathbb{Z}^d)$ -estimates for discrete operators of Radon types I: Maximal functions and vector-valued estimates. J. Funct. Anal. 277 (2019), 2471-2521.

[47] M. Mirek and B. Trojan, Discrete maximal functions in higher dimensions and applications to ergodic theory. Amer. J. Math. 138 (2016), no. 6, 1495-1532.

[48] M. Mirek, B. Trojan and P. Zorin-Kranich, Variational estimates for averages and truncated singular integrals along the prime numbers. Trans. Amer. Math. Soc. 369 (2017), 5403-5423.

[49] R. Oberlin, A. Seeger, T. Tao, C. Thiele and J. Wright, A variation norm Carleson theorem. J. Eur. Math. Soc. 14 (2012), 421-464.

[50] G. Pisier and Q. Xu, The strong p-variation of martingales and orthogonal series. Probab. Theory Relat. Fields 77 (1988), 497-514.

[51] E.M. Stein, Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals. Princeton University Press, Princeton, NJ, 1993.

[52] P. Zorin-Kranich, Variation estimates for averages along primes and polynomials. J. Funct. Anal. 268 (2015), 210-238.

School of Mathematical Sciences, Laboratory of Mathematics and Complex Systems (BNU), Ministry of Education, China, Beijing Normal University, Beijing 100875, China

E-mail address: dingy@bnu.edu.cn

School of Mathematics and Statistics, and Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan 430072, The People’s Republic of China

E-mail address: guixiang.hong@whu.edu.cn

Department of Mathematics, China University of Mining & Technology, Beijing 100083, The People’s Republic of China

E-mail address: wuxf@cumtb.edu.cn