COMBINATORICS OF PLANAR MEASURES AND
BI-PARAMETER CARLESON EMBEDDING

NICOLA ARCOZZI, PAVEL MOZOLYAKO, AND ALEXANDER VOLBERG

ABSTRACT. The main result below is Theorem 1.3 that shows an unexpected property of any positive planar measure. This property goes, on the first glance, against a famous Carleson’s counterexample, and against the obvious geometric property of huge overlap among dyadic rectangles. Lennart Carleson showed in 1974 that the natural generalization of Carleson measure from one parameter case (disc) to bi-parameter case (bi-disc) does not work. Sun-Yang A. Chang in 1979 found the necessary and sufficient condition for the Carleson embedding of bi-harmonic extension into bi-disc. In both works the underlying measure was Lebesgue measure on bi-torus, and “embedding” measure was a priori arbitrary. Sun-Yang A. Chang found that to have embedding it is necessary and sufficient for measure to satisfy Chang–Carleson condition. And as was mentioned above the example of Carleson showed that it is not enough to have a simple “box” condition on embedding measure. Here we switched the places of two measures involved: the embedding measure is now the simplest measure, and the initial measure on bi-torus is an arbitrary measure.

Embeddings of Dirichlet space on bi-disc lead to that kind of questions. Several equivalent necessary and sufficient conditions for such an embedding were found in [AMPS18], [AHMV18a]. These papers have shown that the analog of Chang–Carleson condition is again necessary and sufficient for embedding. This is also probably true on n-torus. However, a new and unexpected effect is exposed below: for n = 2 a simple “box” condition turns out to be equivalent to Chang–Carleson type of condition. This seems to be a new combinatorial fact about all positive measures on the plane. Also this is the note about what goes wrong with bi-parameter two weight Carleson embedding theorem: we construct a family of counterexamples to various statements that hold true in one parameter case.

1. Paraproducts, bi-parameter embedding by Sun-Yang A. Chang, Carleson’s counterexample

Let $Q_0$ be $[0,1]^2$, and let $m$ denote planar Lebesgue measure on $Q_0$. By $\mathcal{D}$ we denote all dyadic sub-rectangles of $Q_0$, and a common name for them will be $R, R_0, R_1, \ldots$.

2010 Mathematics Subject Classification. 42B20, 42B35, 47A30.
Key words and phrases. Carleson embedding on dyadic tree, bi-parameter Carleson embedding, Bellman function, capacity on dyadic tree and bi-tree.

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

AV is partially supported by the NSF grant DMS-160065 and DMS 1900268 and by Alexander von Humboldt foundation.
By \( h_R \) we denote Haar function on \( R \). If \( R = I \times J \), then \( h_R = h_I(x)h_J(y) \), where \( h_K \) denotes the Haar function of dyadic interval \( K \), and we always assume that \( h_K \) has norm 1 in \( L^2 \) with respect to Lebesgue measure on the line.

Consider the following operator named dyadic (bi-parameter) paraproduct. This operator and its numerous variants play one of the major parts in singular integral theory. We consider here the simplest dyadic (but bi-parameter) example of paraproduct:

\[
\Pi_b \varphi := \sum_{R \in \mathcal{D}} \langle \varphi \rangle_R (b, h_R) h_R.
\]

The question of boundedness of the paraproduct \( L^2(m) \to L^2(m) \) is equivalent to the following estimate:

\[
(1) \quad \sum_{R \in \mathcal{D}} |\langle \varphi \rangle_R|^2 (b, h_R)^2 \leq C \int \varphi^2 dm.
\]

This is the condition on function \( b \), or, which is the same, on its Haar coefficients

\[
\beta_R^2 := (b, h_R)^2.
\]

So (1) becomes the question of characterizing coefficients \( \{ \beta_R \}_{R \in \mathcal{D}} \) such that

\[
(2) \quad \sum_{R \in \mathcal{D}} |\langle \varphi \rangle_R|^2 \beta_R^2 \leq C \int_{Q_0} \varphi^2 dm.
\]

The reader can recognize here the dyadic version of Chang–Fefferman “product BMO”: \( b \in BMO^d_{CF} \). Here index \( d \) stays for “dyadic”.

The necessary and sufficient condition on \( \{ \beta_R \} \)’s was found by S.-Y. A. Chang in [Ch] (she considered bi-harmonic bi-disc embedding, not the dyadic one). To formulate Chang’s embedding theorem on bi-disc, let us recall that for an interval \( I \) (arc) on the circle, for any \( z = re^{i\theta_0} \), \( I_z \) denote the arc \( \{e^{i\theta} : |\theta - \theta_0| < 1 - r\} \) and for each interval \( I \) (arc) on the circle, \( |I| < 1 \), we can find \( z \) in the disc such that \( I = I_z \). For \( I \) we also introduce tent \( T(I) = \{ z : I_z \subset I \} \). Theorem below concerns bi-harmonic extension \( u \) of \( f \in L^2(T^2, dm) \) \((dm \text{ is Lebesgue measure on torus } T^2)\). Now if \( R = I \times J \) is a “rectangle” on \( T^2 \), then tent \( T(R) := T(I) \times T(J) \subset D^2 \). For any open subset \( O \) of \( T^2 \),

\[
T(O) := \bigcup_{R \in O} T(R) = \{ (z_1, z_2) \in D^2 : I_{z_1} \times I_{z_2} \subset O \},
\]

here \( R \)'s are not just dyadic but all rectangles.

**Theorem 1.1** (S.-Y. A. Chang, 1979) Let \( u \) be the bi-harmonic extension of \( f \) in \( L^2(T^2, dm) \). Let \( \mu \) be measure on \( D^2 \). Then embedding

\[
\int_{D^2} |u(z_1, z_2)|^2 d\mu(z_1, z_2) \leq C \int_{T^2} |u|^2 dm
\]

holds if and only if

\[
\forall O \quad \mu(T(O)) \leq C' m(O).
\]

**Remark 1.2.** This resembles (2) of course, and the fact that in (2) we have square \( Q_0 = [0, 1]^2 \) and in theorem one has \( T^2 \) has no significance.
The necessary and sufficient condition for (2) (in dyadic form) is below. This necessary and sufficient condition was later used by Chang–Fefferman to characterize the dual space to Hardy space on bi-disc $H^1(D_2^2)$ (which turned out to be $b \in BMO_{CF}$), the reader should notice that this space is not dyadic anymore, see [ChF].

For dyadic inequality (2) the necessary and sufficient condition turns out to be (3)
\[
\forall \Omega \in \mathcal{C}, \quad \sum_{R \in \Omega} \beta^2_R \leq C' m(\Omega),
\]
where $\mathcal{C}$ is the class of subsets of $Q_0$ that are the finite unions of dyadic rectangles: $
\Omega = \bigcup_{T \in R} R$, where $R$ runs over all finite subsets of $D$.

This condition is often called (bi-parameter) Carleson condition.

It is obvious that (2) implies (3) as we can choose $\phi = \chi_\Omega$ in (2) and so get (3).

1.1. The proof that (3) implies (2). We first adopt the proof from [Ch], and then we give a slightly different proof that requires a small geometric argument.

Notice that we can assign to any opens set $O \subset Q_0$ a tent $T(O)$ exactly as before
\[
T(O) = \bigcup_{R \subset O} T(R),
\]
where $R = I \times J$ are all (not just dyadic) rectangles inside $O$, and
\[
T(R) = \{(z, w) : z = x + is, w = u + it, (x - is, x + is) \subset I, (u - it, u + it) \subset J\}.
\]

Consider measure on $\mathbb{C}^2 \times \mathbb{C}^2$ given by the following rule. Given finite collection $\mathcal{R}$ of dyadic rectangles in $Q_0$ and numbers $\beta^2_R$, we assign to each $R = I \times J$ a point $P_R$ in $[0,1]^2 \times [0,1]^2 \subset \mathbb{C}^2 \times \mathbb{C}^2$, call it $(z_1, z_2)$ such that
\[
z_1 = c(I) + \frac{|I|}{2}, \quad z_2 = c(J) + i \frac{|J|}{2},
\]
where $c(\cdot)$ means the center of the interval
\[
\mu_R := \sum_{R \in \mathcal{R}} \beta^2_R \delta_{P_R}.
\]

Let us check that it satisfies
\[
\forall O \quad \mu(T(O)) \leq C' m(O).
\]

Fix an open set $O \subset Q_0$. Consider those $P_R$ that lie on $T(O)$. If a point $(z, w) \in T(O)$ then
\[
z = x + is, w = u + it, (x - is, x + is) \times (u - it, u + it) \subset O
\]
by definition of $T(O)$. In particular all dyadic rectangles $R \in \mathcal{R}$ such that $P_R \in T(O)$ will lie in $O$. Hence,
\[
\mu_{\mathcal{R}}(T(O)) = \sum_{R \in \mathcal{R}, P_R \in T(O)} \beta^2_R = \sum_{R \in \mathcal{R}, R \subset O} \beta^2_R \leq C' \mu(O).
\]

Thus we can use Theorem 1.1 and guarantee embedding (2) because average $|\langle \phi \rangle_R|$ is trivially bounded by the bi-harmonic extension $u_{|\phi|}(P_R)$.
Now we come to the second proof. The first remark is that dyadic rectangular maximal function
\[ M^d_r \varphi := \sup_{x \in R, R \in D} \langle |\varphi| \rangle_R \]
is majorized by the superposition of one dimensional dyadic maximal functions:
\[ M^d_x \varphi(x_0, y_0) \leq M^d_x M^d_y \varphi(x_0, y_0). \]
This is obvious by definition. Just by separation of variable the latter superposition is bounded in \( L^2(Q_0, m) \), so \( M^d_r : L^2(Q_0, m) \to L^2(Q_0, m) \) is bounded.

We can normalize Carleson condition and think that
\[ \forall \Omega \in \mathcal{C}, \quad \sum_{R \in \Omega} \beta^2_R < m(\Omega). \]
Now we need a nice geometric lemma, see [Verb], [TH], [AB] and also Section 6 below.

**Lemma 1** Let the finite collection of numbers \( \beta_R, R \in \mathcal{D} \), satisfy (4). Then we can assign a subset \( E_R \) to every dyadic rectangle \( R \) in such a way that 1) sets \( E_R \) are pairwise disjoint, 2) \( m(E_R) = \beta^2_R \).

Given this lemma we can write (for any finite sequence of \( \beta_R \)'s):
\[
\sum_{R \in \mathcal{D}} |\langle \varphi \rangle_R|^2 \beta^2_R \leq \sum_{R \in \mathcal{D}} \int_{E_R} (M^d_r \varphi(x, y))^2 \, dm \leq \int_{Q_0} (M^d_r \varphi(x, y))^2 \, dm \leq C \int \varphi^2 \, dm.
\]
And we are done.

1.2. A very natural question and Carleson counterexample. A very natural question arises: consider one box condition.

\[ \forall R_0 \in \mathcal{D} \quad \sum_{R \subset R_0} \beta^2_R \leq C' m(R_0), \]
it is like (3), but complicated sets (so-called “dyadic open sets”) \( \Omega \) are replaced by just all dyadic rectangles \( R_0 \).

May be (5) is equivalent to (3), and, thus, is a simple necessary and sufficient condition for paraproduct inequality
\[
\sum_{R \in \mathcal{D}} |\langle \varphi \rangle_R|^2 \beta^2_R \leq C \int \varphi^2 \, dm?
\]
The answer is “no”, Carleson constructed counterexample, see [Car], [Tao]. He constructed the family \( \mathcal{R} \) of dyadic rectangles and numbers \( \beta_R, R \in \mathcal{R} \), such that (5) is valid with constant 1, but for \( \Omega = \bigcup_{R \in \mathcal{R}} R \) inequality
\[ \sum_{R \subset \Omega} \beta^2_R \geq C' m(\Omega) \]
with \( C' \) as large as one wishes.

**Remark 1.3.** In Carleson’s construction \( \beta^2_R = m(R) \).
1.3. From Lebesgue measure on the square to Lebesgue measure on bi-tree. Let us reformulate slightly the paraproduct estimate. Instead of asking when
\[ \sum_{R \in D} |\langle \varphi \rangle_R|^2 \beta^2_R \leq C \int \varphi^2 \, dm \]
holds true, let us ask when
\[ (6) \quad \sum_{R \in D} |\langle \varphi \rangle_R|^2 m(R)^2 \alpha_R \leq C \int \varphi^2 \, dm \]
holds true? We just did nothing, we changed the notations:
\[ \beta^2_R =: m(R)^2 \alpha_R . \]

But now it is natural to ask: why only Lebesgue measure \( m \)? Let us replace it by arbitrary finite positive measure \( \mu \) on \( Q_0 \). Such problem appears immediately from certain question on spaces of holomorphic functions in bi-disc, see [AMPST18].

So here is a generalization. When
\[ (7) \quad \sum_{R \in D} |\langle \varphi \rangle_{R,\mu}|^2 \mu(R)^2 \alpha_R \leq C \int \varphi^2 \, d\mu ? \]
We need to characterize pairs \( (\mu, \alpha := \{ \alpha_R \}_{R \in D}) \) such that this embedding of \( L^2(Q_0, \mu) \) to \( l^2(T^2, \alpha) \) holds. Here \( T^2 \) is the graph of all dyadic rectangles in \( D \), and \( \alpha \) is the measure on this graph (=bi-tree).

Now measure \( \mu \) is arbitrary. Before \( \mu \) was Lebesgue measure on \( Q_0 \). Before \( \alpha \) as measure on bi-tree was arbitrary. Let us now make it “Lebesgue measure” on bi-tree, meaning to choose the simplest possible choice:
\[ \alpha_R \equiv 1 . \]

So before we had pair of measures on \( Q_0, T^2 \) correspondingly (Lebesgue \( m \), arbitrary \( \alpha \)). Now let us consider another “extremal case”: (arbitrary \( \mu \), \( \alpha \equiv 1 \)). The choice of \( \alpha \) is dictated by the problem of embedding of Dirichlet space of analytic functions on bi-disc.

One has all reasons to expect that the same Carleson type counterexample holds true: there should be positive measure on \( Q_0 \) that satisfies box condition (8) below but does not satisfy Carleson condition (9) below.

\[ (8) \quad \forall R_0 \in D \quad \sum_{R \subset R_0} \mu(R)^2 \leq C' \mu(R_0) . \]

\[ (9) \quad \forall \Omega \in \mathcal{C} \quad \sum_{R \subset \Omega} \mu(R)^2 \leq C' \mu(\Omega) . \]

There is no counterexample as we show in this paper. Our main theorem is

**Theorem 1.4** For any positive finite measure \( \mu \) on the square (8) implies (9).
1.4. Main definitions. In addition to proving Theorem 1.4, we prove in this note that there is no $T1$ theorem for bi-parameter dyadic Carleson embedding that starts from space $L^2(\mu)$ and embeds it into space $\ell^2(T^2, \alpha)$ even when $\alpha$ is a rather nice 0,1 sequence enumerated by dyadic rectangles. Notice that if the space where we embed has a special structure, then $T1$ theorem holds, see [?], [AHMV18b] for the case $\alpha \equiv 1$, and [AMVZ19] for tensor $\alpha$’s.

Let $T^2$ be a finite (but very deep) bi-tree. Bi-tree is the directed graph of all dyadic rectangles in the square $Q_0 = [0,1]^2$. We assume that it terminates at small squares of size $2^{-N} \times 2^{-N}$, we call them generically by symbol $\omega$. The boundary $\partial(T^2)$ is this collection of $\omega$’s. Often we identify $T^2$ with dyadic rectangles, whose family is called $D$. When we write $E \subset (\partial T)^2$ we mean any subset of $\omega$’s. It is convenient to think of $E$ as the union of “$N$-coarse” dyadic rectangles.

Box condition
\begin{equation}
\sum_{Q \in T^2, Q \subset R} \mu^2(Q) \alpha_Q \leq C \mu(E), \quad \text{for any } R \in T^2.
\end{equation}

Carleson condition
\begin{equation}
\sum_{Q \in T^2, Q \subset E} \mu^2(Q) \alpha_Q \leq C \mu(E), \quad \text{for any } E \subset (\partial T)^2.
\end{equation}

Restricted Energy Condition
\begin{equation}
\sum_{Q \in D} \mu^2(Q \cap E) \alpha_Q \leq C \mu(E), \quad \text{for any } E \subset (\partial T)^2
\end{equation}

Embedding
\begin{equation}
\sum_{Q \in D} \left( \int_Q \varphi \, d\mu \right)^2 \alpha_Q \leq C \int_{Q_0} \varphi^2 \, d\mu \quad \text{for any } \varphi \in L^2(Q_0, d\mu).
\end{equation}

1.5. The plan of paper. In the next section Theorem 1.4 is proved. Then we change $\alpha$. Instead of $\alpha \equiv 1$ we consider more general weights. We show then all kind of counterexamples—even when $\alpha$ takes only values 1,0 and even if its support is a really nice subgraph of $T^2$. The natural question arises, what other weights $\alpha$, apart from $\alpha \equiv 1$, would have the analog of Theorem 1.4 and, in general, would emulate the result of S.-Y. A. Chang? The partial answer to this question can be found in [AMVZ19].
2. FROM BOX CONDITION TO CARLESON CONDITION: COMBINATORICS OF ARBITRARY POSITIVE MEASURE

The result of this section seem to be $2D$ results that might not be extendable to more dimensions. This effect is not unlike the one described—in a different situation—in [RF1], [RF2], [JJJ].

We always assume that the rectangles are dyadic and $(N+1)$-coarse.

**Theorem 2.1** Assume $\alpha \equiv 1$. Then the bi-parameter box condition implies Carleson condition.

**Scheme of the proof.**

- We try to replicate the argument from [AHMV18a] (Section 3). There it was shown that Carleson condition is equivalent to the so-called Restricted Energy Condition (REC). We therefore argue by contradiction, namely we show that if REC does not hold for a measure $\mu$, then the box condition also fails.

- Assume that there exists a measure $\nu$ on $(\partial T)^2$ that satisfies the box condition with some very small constant $\delta > 0$, but the constant in the corresponding restricted energy condition is much larger, i.e. there exists a set $\tilde{E}$ such that

$$\sum_{Q \in \mathcal{D}} \nu^2(Q \cap \tilde{E}) \geq 10\nu(\tilde{E}).$$

- We proceed by truncating $\nu$ on $\tilde{E}$ (see [AHMV18a] (Lemma 3.1)) to obtain a set $E \subset \tilde{E}$ such that $\mu := \nu|_E$ is almost equilibrium

$$\forall \mu \geq 1, \quad \text{on supp } \mu.$$

- The resulting measure $\mu$ is fairly “unbalanced”, in the sense that the box condition for this measure fails for many rectangles $Q \subset Q_0$, therefore the rest of $\nu$ should ’rebalance’ $\mu$ on those rectangles, so that $\nu$ would satisfy the box condition with desired constant.

- We next show that such rectangles are quite numerous, moreover some of them generate the set $F$ such that $\nu$ is even more unbalanced on $F$,

$$\sum_{Q \in \mathcal{D}} \nu^2(Q \cap F) \geq 100\nu(F).$$

- We repeat the procedure until we hit the whole bi-tree, or we run out of energy.

**Lemma 2** Assume $\nu$ is a measure on $(\partial T)^2$ that satisfies

$$\sum_{R \subset Q} \nu^2(R) \leq \delta \nu(Q), \quad \text{for any } Q \subset Q_0, \quad \text{(14)}$$

for some number $0 < \delta \leq 1$, but also there exists a set $E \subset (\partial T)^2$ such that

$$\forall \mu(\omega) \geq 1, \quad \omega \in E, \quad \text{(15)}$$
where $\mu = \nu|_{E}$. Then there exists a number $\delta_0 > 0$ such that, if $\delta \leq \delta_0$, then one can find a set $F \subset (\partial T)^2$ satisfying

$$E[\sigma] \geq M_\delta|\sigma|,$$

where $\sigma = (\nu - \mu)|_{F}$ and $M_\delta \geq 100$.

3. Proof of Lemma 2

3.1. Part 1. Let $0 < \theta, \varepsilon \leq \frac{1}{100}$, and $0 < \delta \leq \delta_0 \leq \frac{\theta^{10} \varepsilon^{10}}{100}$. Define a “layer”

$$R_\varepsilon := \{Q : \varepsilon \leq \mathcal{V}^\mu(Q) \leq \varepsilon^{10}\}.$$

Consider a “$(\theta, \varepsilon)$-unbalanced” part of $R_\varepsilon$

$$R_{\theta, \varepsilon} := \{Q \in R_\varepsilon : E_Q[\mu] := \sum_{R \subset Q} \mu^2(R) \geq \varepsilon^\theta \frac{\mu(Q)}{20} \}.$$

Let

$$F := \bigcup_{Q \in R_{\theta, \varepsilon}} Q,$$

and let $\sigma := (\nu - \mu)|_{F}$. We do not know a priori if the set $F$ or the collection $R_{\theta, \varepsilon}$ are non-empty, however, if $Q \in R_{\theta, \varepsilon}$, then, by (14) one has

$$\varepsilon^\theta \frac{\mu(Q)}{20} \leq E_Q[\mu] \leq E_Q[\mu + \sigma] \leq \delta(\mu + \sigma)(Q).$$

We will choose later

$$\theta = \varepsilon, \delta = \varepsilon^{30}.$$

This, combined with (17) it asserts that

$$\mu(Q) \leq 30\varepsilon^{28}\sigma(Q), \quad \forall Q \in R_{\theta, \varepsilon}.$$

Since $\text{supp} \sigma \subset F$, we have

$$\forall^\mu(\omega) \geq \varepsilon^{10}, \quad \omega \in \text{supp} \sigma.$$

We aim to show that if $\varepsilon \leq \varepsilon_0$ is small enough and (18) holds, then (16) holds.

3.2. Part 2. Our main goal here is to prove that $\sigma$ has a lot of mass (compared to $\mu$). We know that $\sigma$ is relatively large on the rectangles from $R_{\theta, \varepsilon}$, hence we intent to see how numerous such rectangles are. The measuring stick here is the $\mu$-energy.

Lemma 3 ($R_{\theta, \varepsilon}$-lemma) There exist absolute constants $c, \varepsilon_0 > 0$ such that for $0 < \theta, \varepsilon \leq \varepsilon_0$ one has

$$E_{R_{\theta, \varepsilon}}[\mu] := \sum_{Q \in R_{\theta, \varepsilon}} \mu^2(Q) \geq c\varepsilon E[\mu].$$

Assume for a moment that this lemma holds. Then, as a corollary, we have the following statement.
Lemma 4 (Rebalancing lemma) One has

\[
\int \nabla \sigma \, d\mu \leq C_{\theta, \varepsilon, \delta} \int \nabla \sigma \, d\sigma = C_{\theta, \varepsilon, \delta} E[\sigma],
\]

where

\[
C_{\theta, \varepsilon, \delta} = \frac{C\delta}{\theta \varepsilon^2}
\]

for some absolute constant \( C \).

The inequality (16) follows almost immediately. Indeed, by (20) we have

\[
\varepsilon^{10} |\sigma| \leq \int \nabla \sigma \, d\mu \leq C_{\theta, \varepsilon, \delta} \varepsilon \|\mu\| = C_{\theta, \varepsilon, \delta} E[\sigma],
\]

and we have (16) with \( M_{\delta} = M(\delta, \theta, \varepsilon) = \frac{c_{\theta, \varepsilon}}{c_{\theta, \varepsilon}} \). Taking \( \delta \) to be as in (18) and \( \varepsilon \) to be small enough, given a dyadic we make \( \delta \geq 100 \).

3.3. Proof of Lemma 4 modulo \( \mathcal{R}_{\theta, \varepsilon} \)-lemma. Since

\[
\int \nabla \sigma \, d\mu = \sum_{Q} \mu(Q) \sigma(Q) \leq \left( \sum_{Q} \mu(Q) \right) \left( \sum_{Q} \sigma(Q) \right) = E[\sigma] E[\sigma],
\]

the estimate (22) holds, if we show that

\[
\mathcal{E}[\mu] \leq C_{\theta, \varepsilon, \delta}^2 E[\sigma],
\]

with \( C_{\theta, \varepsilon, \delta} \) satisfying (23). From (19) and \( \mathcal{R}_{\theta, \varepsilon} \)-lemma we deduce that

\[
\mathcal{E}[\mu] \leq \frac{1}{c\varepsilon} \mathcal{E}_{\mathcal{R}_{\theta, \varepsilon}}[\mu] = \frac{1}{c\varepsilon} \sum_{Q \in \mathcal{R}_{\theta, \varepsilon}} \mu^2(Q) \leq \frac{1}{c\varepsilon} \left( \frac{40\delta}{\theta \varepsilon} \right)^2 \sum_{Q \in \mathcal{R}_{\theta, \varepsilon}} \sigma^2(Q) \leq \frac{40^2}{c\theta^2 \varepsilon^3} \delta^2,
\]

where

\[
C_{\theta, \varepsilon, \delta} = \frac{40^2}{c\theta^2 \varepsilon^3} \delta^2,
\]

and Lemma 4 follows.

3.4. Proof of \( \mathcal{R}_{\theta, \varepsilon} \)-lemma. To prove that

\[
\mathcal{E}[\mu] \leq 2\varepsilon^{-1} \mathcal{E}_{\mathcal{R}_{\theta, \varepsilon}}[\mu]
\]

it is enough to consider \( \hat{\mathcal{R}}_{\theta, \varepsilon} := \{ Q : \nabla \mu(Q) \leq \varepsilon, \mathcal{E}_Q[\mu] \geq \frac{\varepsilon^2}{c\theta^2 \varepsilon^3} \mu(Q) \} \), and to prove

\[
\mathcal{E}[\mu] \leq \varepsilon^{-1} \mathcal{E}_{\hat{\mathcal{R}}_{\theta, \varepsilon}}[\mu].
\]

In fact, we know that (see by Lemma 6.9 of [AHMV18b])

\[
\mathcal{E}_{\hat{\mathcal{R}}_{\theta, \varepsilon}}[\mu] - \mathcal{E}_{\mathcal{R}_{\theta, \varepsilon}}[\mu] \leq \varepsilon^{10} \mathcal{E}[\mu] \leq \varepsilon^5 \mathcal{E}[\mu],
\]

so (26) clearly implies (25). The difference of the definition of \( \hat{\mathcal{R}}_{\theta, \varepsilon} \) with the definition of \( \mathcal{R}_{\theta, \varepsilon} \) is that the restriction \( \varepsilon^{10} \leq \nabla \mu(Q) \) is not there anymore.
Let us introduce the relationship between small (boundary of the bi-tree) squares $\omega$ and rectangles $Q$. We call $(\omega, Q) \in P_\varepsilon$ if

$$\sum_{R \omega \subset R \subset Q} \mu(R) \leq \varepsilon.$$  

We also subdivide the family of $Q$, such that $\forall^\mu(Q) \geq \varepsilon$ to bad and good ones. Such rectangle $Q$ is bad, $Q \in \text{Bad}_\varepsilon$, if

$$\mathcal{E}_Q[\mu] < \frac{\varepsilon^2}{20} \mu(Q).$$

The rest is collected to family $\text{Good}_\varepsilon$.

If $Q \in \text{Bad}_\varepsilon$ the measure of all $\omega$ in it such that $(\omega, Q) \in P_\varepsilon$ is at least $(1 - \frac{\varepsilon}{20}) \mu(Q)$. We call this measure $m(Q)$. In fact, this is just Tchebyshov inequality as

$$\int_Q V^\mu_{\varepsilon} d\mu = \mathcal{E}_Q[\mu] \leq \frac{\varepsilon^2}{20} \mu(Q).$$

Notice that for the collection $\tilde{R}_{\varepsilon, \theta}$ from (26) we can write

$$\tilde{R}_{\varepsilon, \theta} = \text{Good}_\varepsilon.$$

Now to prove (26) let us start the estimate from below

$$\mathcal{E}_{\tilde{R}_{\varepsilon, \theta}}[\mu] = \mathcal{E}_{\text{Good}_\varepsilon}[\mu] = \sum_{Q \in \text{Good}_\varepsilon} \mu(Q)^2$$

$$= \sum_{Q : \forall^\mu(Q) \geq \varepsilon} \mu(Q)^2 - \sum_{Q \in \text{Bad}_\varepsilon} \mu(Q)^2$$

$$\geq \int V^\mu_{\varepsilon} d\mu - (1 - \frac{\varepsilon}{20}) \sum_{Q \in \text{Bad}_\varepsilon} \mu(Q) m(Q)$$

If we denote

$$V^\mu_{\text{bad}, \varepsilon}(\omega) := \sum_{Q : \omega \subset Q, (\omega, Q) \in P_\varepsilon} \mu(Q); \quad V^\mu_{\text{good}, \varepsilon}(\omega) := V^\mu_{\varepsilon}(\omega) - V^\mu_{\text{bad}, \varepsilon}(\omega),$$

then Fubini’s theorem shows that

$$\sum_{Q \in \text{Bad}_\varepsilon} \mu(Q) m(Q) = \int V^\mu_{\text{bad}, \varepsilon} d\mu.$$

Hence, we can continue

$$\mathcal{E}_{\tilde{R}_{\varepsilon, \theta}}[\mu] \geq \int V^\mu_{\varepsilon} d\mu - (1 + \frac{\varepsilon}{30}) \int V^\mu_{\text{bad}, \varepsilon} d\mu = \int V^\mu_{\text{good}, \varepsilon} d\mu - \frac{\varepsilon}{30} \int V^\mu_{\text{bad}, \varepsilon} d\mu.$$

Now it is enough to prove the following Lemma.

**Lemma 5** Either $V^\mu_{\text{good}, \varepsilon}(\omega) \geq \frac{\varepsilon}{60}$, or $V^\mu_{\varepsilon}(\omega) \geq \frac{1}{5} V^\mu(\omega)$.

Assume lemma is proved. As the second alternative can happen only on collection of $\omega$’s of total measure $\leq C \varepsilon^{1/2}$ by Lemma 6.9 of [AHMV18b], we have $V^\mu_{\text{good}, \varepsilon}(\omega) \geq \frac{\varepsilon}{60} V^\mu(\omega) \geq \frac{1}{5} V^\mu(\omega)$ at least $1 - C \varepsilon^{1/2}$ portion of measure $\mu$. Then we continue.

$$\mathcal{E}_{\tilde{R}_{\varepsilon, \theta}}[\mu] \geq (1 - C \varepsilon^{1/2}) \frac{\varepsilon}{5} \|\mu\| - \frac{\varepsilon}{30} \int V^\mu_{\varepsilon} d\mu \geq \frac{\varepsilon}{7} \int V^\mu_{\varepsilon} d\mu,$$
since we assumed $\|\mu\| = \int \forall \nu \, d\mu$.

Before proving the lemma we need some notations since the proof employs a certain representation of the ancestors of a given point $\omega \in (\partial T)^2$. Assume we have the depth $(N+1)$ of the bi-tree fixed (we remind that the actual value of $N$ bears no importance, other than that it should be very large). Let

$$q = q_N := [0, 1]^2 \cap \left(\frac{1}{N} \mathbb{Z}\right)^2,$$

in other words, $q = \left\{ (\frac{j}{N}, \frac{k}{N}) \mid 0 \leq j, k \leq N \right\}$. Now fix any point $\omega \in (\partial T)^2$. It has exactly $(N+1)^2$ ancestors, in particular to every $Q \ni \omega$ one can associate a point $\left(\frac{j_0}{N}, \frac{k_0}{N}\right) \in q$ in such a unique way that $|Q_x| = 2^{j_0-N}$ and $|Q_y| = 2^{k_0-N}$, where $Q_x$ and $Q_y$ are the sides of the rectangle $Q$. Vice-versa, every point $(x, y) \in q$ has a uniquely defined associated rectangle $Q_{(x, y)} \ni \omega$. It also means that to every $(x, y) \in q$ one can attach a mass $\Phi(x, y) := \mu(Q)$ with $Q = Q_{(x, y)}$.

Now given $(x_0, y_0) = \left(\frac{j_0}{N}, \frac{k_0}{N}\right) \in q$ one clearly has that

$$\forall^\mu(Q_{(x_0, y_0)}) = \sum_{R \ni Q_{(x_0, y_0)}} \mu(R) = \sum_{j=j_0}^N \sum_{k=k_0}^N \Phi \left(\frac{j}{N}, \frac{k}{N}\right);$$

$$\forall^\mu_{Q_{(x_0, y_0)}}(\omega) = \sum_{\omega \in R \ni Q_{(x_0, y_0)}} \mu(R) = \sum_{j=0}^{j_0} \sum_{k=0}^{k_0} \Phi \left(\frac{j}{N}, \frac{k}{N}\right).$$

We say that $(x_0, y_0) = \left(\frac{j_0}{N}, \frac{k_0}{N}\right) \preceq (x_1, y_1) = \left(\frac{j_1}{N}, \frac{k_1}{N}\right)$, if $j_0 \leq j_1$ and $k_0 \leq k_1$, this corresponds to $Q_{(x_0, y_0)} \subset Q_{(x_1, y_1)}$.

**Proof.** Fix $\omega$. Let us consider all points of $q$ (they represent rectangles $R$) such that $\omega \subset R$, and $\forall^\mu(R) \leq \varepsilon$. On the boundary (let us call it red line) we have rectangles such that

$$\varepsilon/2 \leq \forall^\mu(R) \leq \varepsilon.$$

All points right and top of the red line will be called $RT_{red}$; all the points on the left and bottom are called $LB_{red}$.

Now let us consider all points of $q$ (they represent rectangles $Q$) such that $\omega \subset Q$ and such that $(\omega, Q) \in \mathcal{P}_\varepsilon$. The boundary is called green line and we have $RT_{green}$ and $LB_{green}$ collections of points.

One more notation. For any point $P \in q$ (it corresponds to a dyadic rectangle) we consider collection $R_P$ of points in $q$ having the form of rectangle with sides parallel to the axis, such that $P$ is left bottom vertex. Also $R_P'$ is the collection of points in $q$ having the form of rectangle with sides parallel to the axis, such that $P$ is the right top point.

Suppose $RT_{green}$ is not inside $RT_{red}$. Then there is a point $P$ on the red line such that $R_P \subset RT_{red}$ and $R_P'$ contains no points corresponding to $Q$ which is in relationship $\mathcal{P}_\varepsilon$ with $\omega$. Then

$$\forall^\mu_{\text{good, } \varepsilon}(\omega) = \sum_{Q \text{ corresponding to } q \in RT_{green}} \mu(Q) \geq \frac{\varepsilon}{2},$$

...
and we are done.

So we can assume that \( RT_{\text{green}} \subset RT_{\text{red}} \), that is, the green line is always upper and to the right of the red line.

Suppose \( \mathcal{V}_{\mu}^{\mu}(\omega) \) is much smaller than \( \varepsilon \), say \( \leq \varepsilon/10 \). Then the following is true:

\[
(28) \quad \forall P \in \text{red line}, \quad \sum_{Q \text{ corresponding to } R_P \cap LB_{\text{green}}} \mu(Q) \geq \frac{\varepsilon}{2} - \frac{\varepsilon}{10} = \frac{2}{5} \varepsilon.
\]

Now for any point \( P \in q \) let us denote \( F_P := \{ q : q \in R_P \cap LB_{\text{green}} \} \). We can start with the first northeast point \( P_1 \) on red line and build \( F_{P_1} \).

On the boundary of \( F_{P_1} \) there is the rightest vertical segment (it is green and may consist on one point). Its top point is called \( B_1 \).

Follow this vertical green segment vertically. Follow it even when it stops to be green, but up to the moment of intersection with red line. This is \( P_2 \). Construct \( F_{P_2} \). By construction \( F_{P_1} \cap F_{P_2} = \emptyset \).

Repeat the procedure: on \( F_{P_2} \) there is the rightest vertical segment (it is green and may consist on one point). Its top point is called \( B_2 \). Follow this vertical green segment vertically. Follow it even when it stops to be green, but up to the moment of intersection with red line. This is \( P_3 \). Construct \( F_{P_3} \).

Continue in this fashion. There will be \( F_{P_i}, i = 1, \ldots, n \). From (28) we know that

\[
(29) \quad \forall i \quad \frac{2}{5} \varepsilon \leq \sum_{Q \text{ corresponding to } q \in F_{P_i}} \mu(Q) \leq \varepsilon.
\]

Now notice that \( \bigcup_i R_{B_i} \) and \( \bigcup_i F_{P_i} \) and \( RT_{\text{green}} \) contain all \( q \)'s.

\[
\frac{\varepsilon}{10} + n \varepsilon + n \cdot \varepsilon \geq \mathcal{V}_{\mu}^{\mu}(\omega).
\]

The term \( n \cdot \varepsilon \) represents the contribution of \( R_{B_i}, i = 1, \ldots, n \), in fact, points \( B_i \) are green and so \( \sum_{Q \text{ corresponding to } q \in R_{B_i}} \mu(Q) \leq \varepsilon \). The term \( n \varepsilon \) stands here because of the right inequality of (29). The term \( \frac{\varepsilon}{10} \) correspond to summing \( \mu(Q) \) over \( Q \) corresponding to \( q \in RT_{\text{green}} \), by our assumption on \( \mathcal{V}_{\mu}^{\mu}(\omega) \) it is smaller than \( \varepsilon/10 \).

Now let us use the left inequality (29). Then we have

\[
\mathcal{V}_{\mu}^{\mu}(\omega) \geq \sum_i \sum_{Q \text{ corresponding to } q \in F_{P_i}} \mu(Q) \geq \frac{2}{5} \varepsilon \cdot n.
\]

Comparing the last two display formulas, we get

\[
\mathcal{V}_{\mu}^{\mu}(\omega) \geq c \mathcal{V}_{\mu}^{\mu}(\omega).
\]

Lemma is proved.

\( \square \)
4. Examples of having Carleson condition but not restricted energy condition REC.

Our aim here is to show that if we do not restrict ourselves to the constant weights as in [AMPS18], [AHMV18b], [AMVZ19], then the Carleson condition (11) is no longer sufficient for the embedding (13). In fact even the Restricted Energy Condition (12) is not necessarily implied by (11). Namely we prove the following statement.

**Proposition 4.1** For any $\delta > 0$ there exists number $N$, a weight $\alpha : T^2_N \mapsto \mathbb{R}_+$ and a measure $\mu$ on $(\partial T)^2$ such that $\mu$ satisfies Carleson condition (11) with the constant $C_\mu = \delta$, $\sum_{Q \subset E} \mu^2(Q)\alpha_Q \leq \delta \mu(E)$, for any $E \subset (\partial T)^2$.

but one can also find a set $F$ such that

$$\sum_{Q \in \mathcal{D}} \mu^2(Q \cap F)\alpha_Q \geq \mu(F),$$

hence the constant in (12) is at least 1.

We intend to give two examples of this kind. Both of them rely on the fact that one can basically consider this problem on a cut bi-tree by letting $\alpha$ to be either 1 or 0. This approach clearly does not work on the tree (see [AHMV18b] Theorem 1.1), but the bi-tree has richer geometric structure.

While globally (i.e. for $\alpha \equiv 1$) it looks at least somewhat similar to the tree (this similarity is implicit in the proof of [AHMV18b] Theorem 1.5), one can remove some vertices (which is what essentially happens when we put $\alpha_Q := 0$) in such a way that the remaining part looks nothing like the full bi-tree (or a tree for that matter). In particular this allows us to create a significant difference between the amount of “available” rectangles that lie inside $E$ or just intersect $E$ for a certain choice of the weight $\alpha$ and the set $E$ (this corresponds to the difference between Carleson and REC conditions).

The first example is quite simple and is inspired by the counterexample for $L^2$-boundedness of the bi-parameter maximal function. The weight $\alpha$ in this case cuts most of the bi-tree, and the resulting set differs greatly from the original graph. The second example is somewhat more involved, on the other hand the weight there leaves a much bigger portion of the bi-tree, actually it has a certain monotonicity property: $\alpha_R \geq \alpha_Q$ for $R \supset Q$. The structure of the “available” set is more rich in this case (it looks more like $\mathbb{Z}^2$ in a sense), nevertheless there is not enough rectangles to have the Carleson-REC equivalence.

4.1. A simple of having Carleson condition but not restricted energy condition. Let $N \in \mathbb{N}$ be some large number (to be specified later), and let $T^2 = T^2_N$ be a bi-tree of depth $N$. We use the dyadic rectangle representation of $T^2$.

Let $\omega$ be our $[0, 2^{-N}]^2$ left lower corner. Given $R = [a, b] \times [c, d] \in \mathcal{D}$ let $R^{++} := [\frac{a+b}{2}, b] \times [\frac{c+d}{2}, d]$ be the upper right quadrant of $R$. Consider $Q_1 =$
\([0, 1] \times [0, 2^{-N+1}]\) and its \(Q_1^{++}\), and \(Q_2 = [0, 2^{-1}] \times [0, 2^{-N+2}]\) and its \(Q_2^{++}\), \(Q_3 = [0, 2^{-2}] \times [0, 2^{-N+3}]\) and its \(Q_3^{++}\), et cetera. . . . In total, \(N\) of them.

Put measure \(\mu\) to have mass 
\[
\tau_0 := 1/\sqrt{N}
\]
on \(\omega\), and uniformly distribute mass \(\tau_i\) on \(Q_i^{++}\), \(Q_3 = [0, 2^{-2}] \times [0, 2^{-N+3}]\) and its \(Q_3^{++}\), et cetera. . . . In total, \(N\) of them.

Now \(\alpha_R\) is always zero except when \(R = \omega, Q_1, Q_2, \ldots\) For those \(\alpha = 1\), so we have \(N + 1\) alphas equal to 1.

Now choose set \(E = \omega\). When we calculate \(E[\mu|\Omega]\) we sum up
\[
\tau_0^2 + \sum_{i=1}^{N} \mu(\omega \cap Q_i)^2 = (N + 1)\tau_0^2 = \frac{N + 1}{N} \geq \sqrt{N} \frac{1}{\sqrt{N}} = \sqrt{N} \mu(E).
\]

So REC condition has a big constant.

Since
\[
\mathcal{E}_\Omega[\mu] = \sum_{R \subset \Omega, \alpha_R \neq 0} \mu(R)^2,
\]
then, denoting \(Q_0 := \omega\), we have
\[
\mathcal{E}_\Omega[\mu] = \sum_{Q_i \subset \Omega} \mu(Q_i)^2 =: \sum_{j \in J(\Omega)} \mu(Q_j)^2.
\]

Let \(\tau_0 = 1/\sqrt{N} \leq \frac{1}{4} = \tau_1 = \tau_2 = \ldots = \tau_N\). Then
\[
\mathcal{E}_\Omega[\mu] = \sum_{j \in J(\Omega)} (\tau_0 + \tau_j)^2 \leq 4 \sum_{j \in J(\Omega)} \tau_j^2.
\]

And this is \(\leq \sum_{j \in J(\Omega)} \tau_j \leq \mu(\Omega).\) So Carleson condition holds with constant 1.

4.2. The lack of maximal principle matters. All measures and dyadic rectangles below will be \(N\)-coarse.

In this section we build another example when Carleson condition holds, but restricted energy condition fails. But the example is more complicated (and more deep) than the previous one. In it the weight \(\alpha\) again has values either 1 or 0, but the support \(S\) of \(\alpha\) is an up-set, that is, it contains every ancestor of every rectangle in \(S\).

The example is based on the fact that potentials on bi-tree may not satisfy maximal principle. So we start with constructing \(N\)-coarse \(\mu\) such that given a small \(\delta > 0\)
\[
\mathcal{V}_\mu \lesssim \delta \quad \text{on supp } \mu,
\]
but with an absolute strictly positive \(c\)
\[
\max \mathcal{V}_\mu \geq \mathcal{V}_\mu(\omega_0) \geq c \log N,
\]
where \(\omega_0 := [0, 2^{-N}] \times [0, 2^{-N}]\).

We define a collection of rectangles
\[
Q_j := [0, 2^{-2^j}] \times [0, 2^{-2^{-j}N}], \quad j = 1 \ldots M \approx \log N,
\]
and we let
\[ Q_j^{++} := [2^{-2j-1}, 2^{-2j}] \times [2^{-2j}N^{-1}, 2^{-2j}N] \]
\[ Q_j^- := [0, 2^{-2j-1}] \times [0, 2^{-2j}N], \quad j = 1 \ldots M \]
(35)
\[ Q_j^+ := Q_j \setminus Q_j^- \]
\[ Q_j^- := Q_j \setminus Q_j^+ \]
to be their upper right quadrants, lower halves, top halves, right halves, and lower quadrant respectively. Now we put
\[ R := \{ R : Q_j \subset R \text{ for some } j = 1 \ldots M \} \]
\[ \alpha_Q := \chi_R(Q) \]
(36)
\[ \mu(\omega) := \frac{\delta}{N} \sum_{j=1}^{M} \frac{1}{|Q_j^{++}|} \chi_{Q_j^{++}}(\omega), \]
\[ P_j = (2^{-2j}, 2^{-2j}N) \]

here \(|Q|\) denotes the total amount of points \(\omega \in (\partial T)^2 \cap Q\), i.e. the amount of the smallest possible rectangles (of the size \(2^{-2N}\)) in \(Q\).

Observe that on \(Q_j\) the measure is basically a uniform distribution of the mass \(\frac{\delta}{N}\) over the upper right quarter \(Q_j^{++}\) of the rectangle \(Q_j\) (and these quadrants are disjoint).

To prove (32) we fix \(\omega \in Q_j^{++}\) and split \(\nabla^\mu(\omega) = \nabla^\mu_{Q_j^{++}}(\omega) + \mu(Q_j^+) + \mu(Q_j^-) + \nabla^\mu(Q_j^{++})\), where the first term sums up \(\mu(Q)\) for \(Q\) between \(\omega\) and \(Q_j^{++}\). This term obviously satisfies \(\nabla^\mu_{Q_j^{++}}(\omega) \lesssim \frac{\delta}{N}\). Trivially \(\mu(Q_j^+) + \mu(Q_j^-) \leq \frac{2M}{N}\). The non-trivial part is the estimate
(37)
\[ \nabla^\mu(Q_j^{++}) \lesssim \delta. \]

To prove (37), consider the sub-interval of interval \([1, n]\) of integers. We assume that \(j \in [m, m+k]\). We call by \(C_j^{[m,m+k]}\) the family of dyadic rectangles containing \(Q_j^{++}\) along with all \(Q_i^{++}, \quad i \in [m, m+k]\) (and none of the others). Notice that \(C_j^{[m,m+k]}\) are not disjoint families, but this will be no problem for us as we wish to estimate \(\nabla^\mu(Q_j^{++})\) from above.

Notice that, for example, \(C_j^{[m,m+1]}\) are exactly the dyadic rectangles containing point \(P_j\). It is easy to calculate that the number of such rectangles is
\[ (2^j + 1) \cdot (2^{-J}N + 1) \lesssim N. \]

Analogously, dyadic rectangles in family \(C_j^{[m,m+k]}\) have to contain points \(P_m, P_{m+k}\). Therefore, each of such rectangles contains point \((2^{-2m}, 2^{-2m-k}N).\) The number of such rectangles is obviously at most \(\lesssim 2^{-k}N\). The number of classes \(C_j^{[m,m+k]}\) is at most \(k+1\).
Therefore, $\mathcal{V}^\mu((Q_j^{++})$ involves at most $(k + 1)2^{-k}N$ times the measure in the amount $k \cdot \frac{\delta}{N}$. Hence

$$\mathcal{V}^\mu((Q_j^{++}) \leq \sum_{k=1}^n k(k + 1)2^{-k}N \cdot \frac{\delta}{N},$$

and (37) is proved. Inequality (32) is also proved.

We already denoted

$$\omega_0 := [0, 2^{-N}] \times [0, 2^{-N}],$$

calculate now $\mathcal{V}^\mu(\omega_0)$. In fact, we will estimate it from below. The fact that $C_{j}^{[m,m+k]}$ are not disjoint may represent the problem now because we wish estimate $\mathcal{V}^\mu(\omega_0)$ from below.

To be more careful for every $j$ we denote now by $c_j$ the family of dyadic rectangles containing the point $P_j$ but not containing any other point $P_i, i \neq j$. Rectangles in $c_j$ contain $Q_j^{++}$ but do not contain any of $Q_i^{++}, i \neq j$. There are $(2^{j-1} - 2^j - 1) \cdot (2^j - N - 2^{-j}N - 1), j = 2, \ldots, M - 2$. This is at least $\frac{1}{8}N$.

But now families $c_j$ are disjoint, and rectangles of class $c_j$ contribute at least $\frac{1}{8}N \cdot \frac{\delta}{N}$ into the sum that defines $\mathcal{V}^\mu(\omega_0)$. We have $m_4$ such classes $c_j$, as $j = 2, \ldots, M - 2$. Hence,

$$\mathcal{V}^\mu(\omega_0) \geq \frac{1}{8}N \cdot \frac{\delta}{N} \cdot (M - 4) \geq \frac{1}{9}\delta M.$$

Choose $\delta$ to be a small absolute number $\delta_0$. Then we will have (see 38)

$$\mathcal{V}^\mu \leq 1, \text{ on supp } \mu.$$

But (38) proves also (33) as $M \asymp \log N$.

**Remark 4.1.** Notice that in this example $\mathcal{V}^\mu \leq 1$ on supp $\mu$, and

$$\text{cap}\{\omega : \mathcal{V}^\mu \geq \lambda\} \leq c e^{-2\lambda}.$$

Here capacity is the bi-tree capacity defined e. g. in [AMPS18]. So there is no maximal principle for the bi-tree potential, but the set, where the maximal principle breaks down, has small capacity.

Now we construct the example of $\nu$ and $\alpha$ with $\alpha = 1$ on an up-set (and zero otherwise), and such that Carleson condition is satisfied but REC (restricted energy condition) is not satisfied. We use the same measure $\mu$ we have just constructed, and we put

$$\nu := \mu + \nu|\omega_0,$$

where $\nu|\omega_0$ is the uniformly distributed over $\omega_0$ measure of total mass $\frac{1}{MN}$. Weight $\alpha$ is chosen as in (30).

**Warning.** The meaning of $\mathcal{V}$ changes from now on. Before $\mathcal{V} = \Pi^* (\cdot)$. Everywhere below,

$$\mathcal{V}^\nu := I[\alpha \Pi^* (\nu)].$$
Let us first check that REC constant is bad. We choose $F = \cup_j Q_j^-$. Then $\nu_F := \nu|_{F} = \nu|_{\omega_0}$. On the hand, and this is the main feature,

$$\omega_0 \text{ lies in } M \text{ rectangles } Q_j. \tag{40}$$

Hence,

$$\text{(41)} \quad \exists \geq cNM \text{ dyadic rectangles } R \text{ such that } \alpha_R = 1 \text{ and } \omega_0 \subset R. \tag{41}$$

In fact, consider dyadic rectangles in $\cup c_j$, where families $c_j$ were built above. For each $R \in \cup c_j$ we have $\alpha_R = 1$, see (36). And there are at least $\frac{1}{8}NM$ of them. We conclude

$$\forall \rho^\nu(\omega_0) \geq \frac{1}{8}MN \cdot \nu(\omega_0). \tag{42}$$

Therefore,

$$\int \rho^\nu|_F \ d\nu|_F \geq \nu(\omega_0)^2 \cdot \frac{1}{8}NM = \frac{1}{8} \frac{1}{MN} \geq c_0 \nu(\omega_0).$$

This means that constant of REC is at least absolute constant $c_0$. Let us show that the Carleson constant is $\lesssim c \cdot \delta$. But $c_0$ has nothing to do with $\delta$ that can be chosen as small as we wish.

**Remark 4.2.** We do not need the following claim now, we will need it only later, but notice that in a fashion completely similar to the one that just proved (42), one can also prove

$$\forall \mu(\omega_0) \geq \frac{1}{8}MN \cdot \frac{\delta}{N} \geq c \delta M. \tag{43}$$

Moreover, we already proved it in (33). This holds because $\omega_0$ is contained in exactly $M$ rectangles $Q_j^{++}$.

**Definition 4.3.** Dyadic rectangles whose left lower corner is $(0,0)$ will be called hooked rectangles.

To check the Carleson condition with small constant we fix any finite family $A$ of dyadic rectangles, and let

$$A = \cup_{R \in A} R. \tag{44}$$

We are interested in subfamily $A'$ of $R$ such that $\alpha_R = 1$. Other elements of $A$ do not give any contribution to $E_A[\nu]$ as $\mu(Q)^2 \alpha_Q = \mu(Q)^2 \cdot 0 = 0$ for any $Q \subset Q', Q' \in A \setminus A'$, as the support of $\alpha$ is an up-set.

All rectangles from $A'$ are hooked rectangles. As we noticed, we can think that $A' = A$. In other words, without the loss of generality, we can think that $A$ consists only of hooked rectangles. Any hooked rectangle generates a closed interval $J$ in the segment $[1, n]$ of integers: interval $J$ consists of $j$, $1 \leq j \leq n$, such that point $P_j$ lies in this hooked rectangle. This is the same as to say that $Q_j$, $j \in J$, is a subset of this hooked rectangle.

So family $A$ generates the family of closed intervals in the segment $[1, n]$ of integers. Let us call $J_A$ this family of intervals in the segment $[1, n]$ of integers. Intervals of family $J_A$ can be not disjoint. But we can do the following, if intervals
intersect, or even if these closed intervals are adjacent, we unite them to a new interval. The new system (of disjoint and not even adjacent) closed intervals corresponds to another initial system \(\tilde{\mathcal{A}}\), and we can think that \(\tilde{\mathcal{A}}\) consists of hooked rectangles. We call a system of hooked rectangles a clean system if it gives rise to not adjacent disjoint family of closed intervals inside the set \([1, n]\) of integers. The relationship between rectangles in \(\mathcal{A}\) and \(\tilde{\mathcal{A}}\) is the following: each rectangle of \(\tilde{\mathcal{A}}\) is a common ancestor of a group of rectangles in \(\mathcal{A}\).

A very important geometric property of \(\tilde{\mathcal{A}}\) is the following. Let \(Q \in \tilde{\mathcal{A}}\) and let \(R_1, \ldots, R_s\) be all rectangles from \(\mathcal{A}\) such that \(R_i \subset Q\), \(i = 1, \ldots, s\). Then
\[
\nu(Q \cup \bigcup_{i=1}^{m} R_i) = 0.
\]

In particular, (44) implies
\[
\nu(A) = \nu(\tilde{\mathcal{A}}).
\]

When checking the Carleson condition
\[
E_A[\nu] \lesssim \delta \nu(A),
\]
we can always think about \(A\) being replaced by \(\tilde{\mathcal{A}}\) and \(A\) being replaced by \(\tilde{\mathcal{A}}\) because in (46) the RHS stays the same, but the LHS can jump up with passage from \(A\) to \(\tilde{\mathcal{A}}\). Therefore, checking (46) for clean systems of rectangles is the same as to check it for all systems of rectangles. From now on \(A\) is supposed to be clean.

To prove (46) is the same as to prove (since \(\nu_A(Q) = (\mu_A(Q) + \nu_{\omega_0}(Q))\))
\[
\sum_{Q \subset A} \nu_A(Q)^2 \alpha_Q + \sum_{Q \subset A} \nu_{\omega_0}(Q)^2 \alpha_Q \lesssim \delta \nu(A),
\]

The first sum is bounded by \(\int \nabla^{\mu_A} d\mu|A\). But by (52) (which follows from (57)) we have
\[
\sum_{Q \subset A} \nu_A(Q)^2 \alpha_Q = \int \nabla^{\mu_A} d\mu|A \lesssim \delta \|\mu_A\| = \delta \mu(A) \leq \delta \nu(A),
\]
and this means that the part of (47) is proved.

To estimate \(\sum_{Q \subset A} \nu_{\omega_0}(Q)^2 \alpha_Q = \frac{1}{\log N} \sum_{\alpha_R = 1, R \subset A} \{R\} \subset \mathcal{J}_k\) we take one interval \(\mathcal{J}_k\) from the family generated by the clean \(\mathcal{A}\) in \([1, n]\), we denote
\[
m_k := \# \mathcal{J}_k,
\]
and we estimate how many dyadic rectangles \(R\) contain one of \(Q_j, j \in \mathcal{J}_k\). We even do not care now whether \(R\) is a subset of \(A\) or not. The number of such rectangles in at most \(m_k \cdot N\). On the other hand,
\[
\nu(A) \geq \sum_k m_k \cdot \frac{\delta}{N} + \frac{1}{MN} \sum_k m_k \geq \sum_k m_k \cdot \frac{\delta}{N}.
\]
So to prove the estimate for the second sum in (47), we need to see that
\[
\frac{1}{(MN)^2} \sum_k m_k \cdot N \lesssim \frac{1}{M} \sum_k m_k \cdot \frac{\delta}{N},
\]
which is obviously true if we choose \(\delta \geq \frac{1}{M} = \frac{1}{\log N}\). So (47) is proved.
Remark 4.4. The system \((\nu, \alpha)\) constructed above gives the example of measure and \(\alpha = 1, 0\) with support on an up-set (so on a nice subgraph of a bi-tree) such that Carleson condition holds, but REC condition does not hold. Notice that the support of \(\alpha\) consists of \(m \approx \log N\) sets \(S_1, \ldots, S_M\) in bi-tree, where each set \(S_i\) can be given by the equation
\[
S_i = \{Q = I \times J : \tau_i(I) \cdot \eta_i(J) = 1\},
\]
where \(\tau, \eta\) are characteristic functions of certain up-sets in a simple tree \(T\). So
\[
\alpha = \sum_{i=1}^{M} \tau_i \otimes \eta_i.
\]
The discrepancy between Carleson constant and REC constant is at least of the order \(1/M\). But its at least, nobody said that it is equivalent to \(1/M\), where \(M\) is the least number in display formula above. To understand better the relationship between the least number in display formula above and the REC-Carleson discrepancy one can ask what happens if \(M\) is small, like \(M = 1\) (this means that the weight \(\alpha\) has a product structure), or \(M \simeq 1\)? It turns out that in this case REC and Carleson conditions are equivalent, see [AMVZ19].

5. Restricted energy condition holds but no embedding

In this section we emulate the previous construction, we start with \(\{Q_j\}\) and measure \(\mu\) but instead of adding \(\omega_0\) we will add a more sophisticated piece of measure.

Let us start with recalling the system \(\{Q_j\}, j = 1, \ldots, M\) and measure \(\mu\) from the previous section. We continue with denoting
\[
Q_{0,j} := Q_j, \quad \mu_0 := \mu \text{ from the previous section}.
\]
Next we continue with defining a a sequence of collections \(Q_k, k = 0 \ldots K\), of dyadic rectangles as follows
\[
Q_k := \left\{ Q_{k,j} = \bigcap_{i=j}^{j+2^k} Q_{0,i}, \quad j = 1 \ldots M - 2^k \right\}, \quad k = 1 \ldots K.
\]
In other words, \(Q_0\) is the basic collection of rectangles, and \(Q_k\) consists of the intersections of \(2^k\) consecutive elements of \(Q_0\).
The total amount of rectangles in \(Q_k\) is denoted by \(M_k = M - 2^k + 1\).

We also denoted by \(R\) the collection of rectangles lying above \(Q_0\)
\[
R := \{ R : Q_{0,j} \subset R \text{ for some } 1 \leq j \leq M\},
\]
and we let
\[
S_k := \bigcup_{Q \in Q_k} Q.
\]
The weight \(\alpha\) was defined as follows:
\[
\alpha_Q := 1, \quad \text{if } Q \in R
\]
\[
\alpha_Q := 0, \quad \text{otherwise}.
\]
Now we construct the measure $\mu$, whose main part will be already constructed $\mu_0$. Let

$$Q_{k,j}^{++} := \left[2^{-2^j+2^k-1}, 2^{-2^j+2^k}\right] \times \left[2^{-2^jN-1}, 2^{-2^jN}\right]$$

be the upper right quadrant of $Q_{k,j}$. For every $k = 0 \ldots K$ we distribute the mass $2^{-2^k} \delta \frac{M}{N}$ over the rectangles $Q_{k,j}^{++}$. Namely, for every $j = 1 \ldots M_k$ we attach a mass $2^{-2^k} \delta \frac{M}{N}$ to the rectangle $Q_{k,j}^{++}$ that is uniformly distributed over the quadrant $Q_{k,j}^{++}$.

We note that all these quadrants $Q_{k,j}^{++}$ are disjoint.

Measure $\mu_0$ is the “main” part of $\mu$, in the sense that $\mu_0$ is generated by the masses on $Q_0$,

$$\mu_0(\omega) := \frac{\delta}{N} \sum_{j=1}^{M} \frac{1}{|Q_{0,j}^{++}|} \chi_{Q_{0,j}^{++}}(\omega), \quad \omega \in (\partial T)^2,$$

and let $\mu_k$ be the corresponding mass on $Q_k$

$$\mu_k(\omega) := \frac{2^{-2^k} \delta}{N} \frac{M_k}{N} \sum_{j=1}^{M_k} \frac{1}{|Q_{k,j}^{++}|} \chi_{Q_{k,j}^{++}}(\omega), \quad \omega \in (\partial T)^2,$$

so that

$$\mu = \mu_0 + \sum_{k=1}^{K} \mu_k.$$

Finally we define the function $f$, and we do it in such a way that it is ‘congruent’ with the distribution of $\mu_0$ over $\Omega$, namely we let

$$f(R) := \mu_0(R) \cdot \alpha_R.$$

5.1. **Main idea.** Notice that

$$\nabla \mu_0 = \mathbb{I} f = \mathbb{I} [\alpha^* \mu_0], \quad \int \nabla \mu_0 \, d\mu_0 = \sum \mathbb{I}^* \mu_0 \cdot \mathbb{I}^* \mu_0 \cdot \alpha = \sum f^2 \cdot \alpha.$$

To prove that embedding has a bad constant, it is sufficient to show that the dual inequality has a bad constant:

$$\int (\mathbb{I} f)^2 \, d\mu >> \sum f^2 \cdot \alpha,$$

which becomes

$$\int (\nabla \mu_0)^2 \, d\mu >> \int \nabla \mu_0 \, d\mu_0.$$

Let us look at Remark 4.2 at (40), (41), (43) and compare (40) with (51). The conclusion is: since every $Q_{k,j}$ lies in $2^k$ of $Q_{0,j}$ (number $2^k$ replaces $M$ in (40), (41), (43)), then

$$\nabla \mu_0 \geq c 2^k \frac{\delta}{N} = c\delta 2^k$$

don each $Q_{k,j}$.

We already saw that $\nabla \mu_0 \leq \delta$ on $\mu_0$, so

$$\int \nabla \mu_0 \, d\mu_0 \lesssim \delta \frac{M}{N}.$$
Now, using (55) we get

\[
\int (V_{\mu_0})^2 \, d\mu = \sum_{k=1}^{K} \int (V_{\mu_0})^2 \, d\mu_k \geq c^2 \delta^2 \sum_{k=1}^{K} 2^{2k} \|\mu_k\| = c^2 \delta^3 \frac{M \log M}{N}
\]

For example, let

\[
\delta = \frac{1}{\log M}.
\]

Then the constant of embedding is \( \approx 1 \).

5.2. **REC condition holds with a small constant.** Let \( A \) be a collection of (hooked) rectangles, \( A = \cap_{R \in A} R \). Let \( \nu_k := \mu_k | A, k = 1, \ldots, K, \nu := \mu | A = \sum \nu_k \). We need to prove

\[
E_A[\nu] \lesssim \delta \|\nu\|.
\]

Let \( n > k \), we wish to estimate \( V_{\mu_n}(Q_{k,j}) \). This is a certain sum over a system \( S \) of rectangles of the form

\[
\sum_{R \in S} \mu_n(R),
\]

where dyadic rectangles are a) contain \( Q_{k,j} \), b) \( \alpha_R = 1 \). Notice that this system depends on \( Q_{k,j} \) but totally independent of \( n \). So if we manage to estimate \( \mu_n(R) \) via \( \mu_0(R) \), then we compare \( V_{\mu_n}(Q_{k,j}) \) to \( V_{\mu_0}(Q_{k,j}) \).

But let the number of \( Q_{s,j} \) in \( R \) be denoted by \( m^*_R \). Then it is very easy to see that

\[
m^*_R \leq m^0_R + 2^n + 1.
\]

Then

\[
2^{-2n} m^*_R \leq 2^{-n} (m^0_R + 1 + 2^{-n}) \leq 3 \cdot 2^{-n} m^0_R.
\]

Then

\[
V_{\mu_n}(Q_{k,j}) = \sum_{R \in S} \mu_n(R) \leq 3 \cdot 2^{-n} \sum_{R \in S} \mu_0(R) = 3 \cdot 2^{-n} V_{\mu_0}(Q_{k,j}) \lesssim \delta 2^{k-n}.
\]

Therefore,

\[
\sum_{n \geq k} V_{\mu_n} \, d\nu_k \lesssim \delta \sum_{n \geq k} 2^{k-n} \|\nu_k\| = 2\delta \|\nu_k\|.
\]

And so

\[
\sum_k \sum_{n \geq k} V_{\mu_n} \, d\nu_k \lesssim \delta \sum_k \|\nu_k\| = \delta \|\nu\|.
\]

Inequality (59) is proved.

**Remark 5.1.** We have

\[
\alpha = \sum_{i=1}^{M} \tau_i \otimes \eta_i.
\]

The discrepancy between REC constant and embedding constant is at least of the order \( \frac{1}{\log M} \).
6. Strong dyadic maximal function and counterexamples

**Definition.** Let $S$ be a family of dyadic sub-rectangles of $Q_0$ (may be $S = \mathcal{D}$, the family of all dyadic sub-rectangles). We call the sequence of $\{\beta_Q\}_{Q \in S}$ Carleson if

\[
\forall S' \subset S, \quad \sum_{Q \in S'} \beta_Q \mu(Q) \leq C \mu(\cup_{Q \in S'} Q).
\]

The best $C$ is called the Carleson norm of the sequence.

**Definition.** Abusing the language we say that the weight $\alpha := \{\alpha_Q\}_{Q \in S}$ satisfies Carleson condition if the sequence $\beta_Q := \alpha_Q \cdot \mu(Q)$ is a Carleson sequence:

\[
\forall S' \subset S, \quad \sum_{Q \in S'} \alpha_Q \mu(Q)^2 \leq C \mu(\cup_{Q \in S'} Q).
\]

We already know, see [AHMV18a] e.g., that bi-parameter Carleson embedding

\[
\sum_{R \in \mathcal{D}} \left( \int_R \psi \, d\mu \right)^2 \alpha_R = \sum_{R \in \mathcal{D}} \|I^* (\psi \mu)\|^2 \alpha_R \leq C' \int_{Q_0} \psi^2 \, d\mu
\]

is equivalent to the Carleson condition above if we have $\alpha_R = 1$, $\forall R \in \mathcal{D}$.

Understanding the general two-weight bi-parameter situation (that is $\alpha \neq 1$) seems to be super hard, as the examples above show. Notice that in two-weight one-parameter situation the answer is known, see, e.g. [NTV99]. And the answer is given in terms of Carleson condition. However, in bi-parameter situation this is far from being so as the following theorem shows. First we give

**Definition.** A finite positive measure $\mu$ on $Q_0$ is called a “bad” measure if there exists weight $\alpha := \{\alpha_Q\}_{Q \in \mathcal{D}}$ that satisfies the Carleson condition but such that the embedding

\[
\sum_{R \in \mathcal{D}} \left( \int_R \psi \, d\mu \right)^2 \alpha_R \leq C' \int_{Q_0} \psi^2 \, d\mu
\]

does not hold.

The strong maximal function with respect to $\mu$ is

\[\mathcal{M}_\mu \psi(x) = \sup_{R \in \mathcal{D}} \frac{1}{\mu(R)} \int_R |\psi| \, d\mu,\]

where $0/0 = 0$. The supremum is taken over all dyadic sub-rectangles of $Q_0$.

**Theorem 6.1** Let $\mu$ be atom free. Then the measure $\mu$ is bad if and only if $\mathcal{M}_\mu$ is not a bounded operator in $L^2(\mu)$.

This theorem is proved in [AMV19] with the proof that is based on the ideas from [TH], [Verb], [Dor]. It is also proved in a different fashion in [AMVZ19].

**References**

[AH96] D. R. Adams and L. I. Hedberg. Function spaces and potential theory. Vol. 314. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1996. pp. xii+366. mr: 1411441 (cit. on p. 7).

[AHMV18a] N. Arcozzi, I. Holmes, P. Mozolyako, and A. Volberg. Bellman function sitting on a tree. Preprint. Sept. 2018. arXiv: 1809.03397 [math.CA] (cit. on p. 1).
[AHMV18b] N. Arcozzi, I. Holmes, P. Mozolyako, and A. Volberg. Bi-parameter embedding and measures with restriction energy condition. Preprint. Nov. 2018. arXiv: 1811.00978 [math.CA] (cit. on p. 1).

[AMPS18] N. Arcozzi, P. Mozolyako, K.-M. Perfekt, and G. Sarfatti. Bi-parameter Potential theory and Carleson measures for the Dirichlet space on the bidisc. Preprint. Nov. 2018. arXiv: 1811.04990 [math.CV] (cit. on p. 1).

[AMV19] N. Arcozzi, P. Mozolyako, A. Volberg. “Counterexamples for bi-parameter Carleson embedding”, Preprint, 2019, pp. 1–15.

[AMVZ19] N. Arcozzi, P. Mozolyako, A. Volberg, P. Zorin-Kranich. Bi-parameter Carleson embeddings with product weights. Preprint, 2019, pp. 1–16.

[AB] A. Barron, Sparse bounds in harmonic analysis and semiperiodic estimates, Thesis, Brown University, 2019.

[Car] L. Carleson, A counterexample for measures bounded on $H^p$ for the bi-disc. Mittag-Leffler Report (1974), no. 7.

[Ch] Sun-Yung A. Chang Carleson measure on the bi-disc. Ann. of Math. (2) 109 (1979), no. 3, 613-620.

[ChF] Sun-Yung A. Chang, R. Fefferman, A continuous version of duality of $H^1$ with $BMO$ on the bidisc. Ann. of Math. (2) 112 (1980), no. 1, 179–201.

[Dor] L. Dor. On projections in $L^1$. Ann. of Math. (2), 102(3):463–474, 1975.

[NTV99] F. Nazarov, S. Treil, and A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. In: J. Amer. Math. Soc. 12.4 (1999), pp. 909928. mr: 1685781 (cit. on p. 1).

[RF] R. Fefferman, Strong differentiation with respect to measures , Amer. J. Math. 103(1981).

[RF1] R. Fefferman, Calderón-Zygmund theory for product domains: $H^p$ spaces. Proc. Nat. Acad. Sci. U.S.A. 83 (1986), no. 4, 840–843.

[RF2] R. Fefferman, Some recent developments in Fourier analysis and $H^p$ theory on product domains. II. Function spaces and applications (Lund, 1986), 4451, Lecture Notes in Math., 1302, Springer, Berlin, 1988.

[TH] T. Hanninen, Equivalence of sparse and Carleson coefficients for general sets, arXiv:1709.10457

[JLJ] J.-L. Journé, Two problems of Calderón-Zygmund theory on product-spaces. Ann. Inst. Fourier (Grenoble) 38 (1988), no. 1, 111132.

[Tao] T. Tao, Dyadic product $H^1$, $BMO$, and Carleson’s counterexample, preprint, pp. 1–12.

[Verb] Igor E. Verbitsky, embedding and multiplier theorems for discrete Littlewood-Paley spaces. Pacific J. Math., 176(2):529–556, 1996.

Universitá di Bologna, Department of Mathematics, Piazza di Porta S. Donato, 40126 Bologna (BO)
E-mail address: nicola.arcozzi@unibo.it

Universitá di Bologna, Department of Mathematics, Piazza di Porta S. Donato, 40126 Bologna (BO)
E-mail address: pavel.mozolyako@unibo.it

Department of Mathematics, Michigan State University, East Lansing, MI. 48823
E-mail address: volberg@math.msu.edu (A. Volberg)