Two Weight Problems and Bellman Functions on filtered probability spaces

by

Jingguo Lai

B. S., Fudan University; Shanghai, China, 2008

M. S., Michigan State University; East Lansing, MI, 2010

A Dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Mathematics at Brown University

Providence, Rhode Island

May 2015

This dissertation by Jingguo Lai is accepted in its present form
by the Department of Mathematics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date ______________ ____

Sergei Treil, Advisor

Recommended to the Graduate Council

Date ______________ ____

Jill Pipher, Reader

Date ______________ ____

Brian Cole, Reader

Approved by the Graduate Council

Date ______________ ____

Peter M. Weber, Dean of the Graduate School

Curriculum Vita

Jingguo Lai was born in Shenyang, Liaoning, P.R. China on Feburary 11, 1985 to Changbin Lai and Lijie Xue. He completed the B.S. at Fudan University on July 2008 and the M.S. at Michigan State University on July 2010. After graduating, he continued the study of mathematics at Brown University. He married Maggie Fong in the summer of 2013. Jingguo completed this thesis under the supervision of Sergei Treil.

Dedicated to my beloved parents:
Changbin Lai and Lijie Xue

Acknowledgements

To my advisor, Prof. Sergei Treil for his invaluable mentoring over the past five years. He suggested these problems, taught me the right way of thinking process, and provided guidance and ideas for all the tough steps.

To my readers, Prof. Jill Pipher and Prof. Brian Cole for their careful reading, gentle criticism, and insightful edits.

To Prof. Justin Holmer for his help and support along the way.

To the wonderful mathematics department staff, particularly Audrey Aguiar, Larry Larivee, and Doreen Pappas.

To my parents Changbin Lai and Lijie Xue, and my wife Maggie Fong for their love and support.

To my cat Gigi for bringing me so much fun.

Contents

Abstract of "Two Weight Problems and Bellman Funcfions on filtered probability spaces"

by Jingguo Lai, Ph.D., Brown University, May 2015

Chapter 1 provides the necessary background and states the main results of the thesis. Two seperate topics are studied. The first topic is on two weight problems. The second topic is on Bellman functions on filtered probability spaces.

Chapter 2 proves a two weight estimation for a vector-valued positive operator. We consider two different cases of this theorem. The easier case $1 < p \leq q$ requires only one testing condition. However, we construct a counterexample showing this testing condtion alone is not sufficient for the case $q < p < \infty$. We apply the Rubio de Francia Algorithm to reduce our problem to the well-known two weight estimates for positive dyadic operators.

Chapter 3 proves a two weight estimation for paraproducts. We again consider two different cases seperately. The first few steps of the proof proceeds exactly the same as in Chapter 2. However, for the harder case $2 < p < \infty$, we need to characterize a two weight inequality for shifted bilinear forms, which takes up the majority of this chapter.

Chapter 4 considers the celebrated Dyadic Carleson Embedding Theorem. We streamline a way of finding a super-solution of the Bellman function via the Burkhölder's hull. We give an explicit formula of the Burkhölder's hull and hence a super-solution in this chapter.

Chapter 5 generalizes the Dyadic Carleson Embedding Theorem to the filtered probability spaces and proves the coincidence of the Bellman functions on an infinite refining filtration. The proof requires a remodeling of the Dyadic Carleson Embedding Theorem. Finally, we also consider the Bellman function of the Doob's Martingale Inequality.

CHAPTER 1

Introduction

In this chapter we provide some useful background on the topics of this thesis. First we establish the general setup of two weight problems and raise the questions of interest. Then we introduce a well-known application of the Bellman function techniques and pose the questions we want to solve. Further we provide a brief outline of this thesis.

1. Two Weight Problems

The original question about two weight estimates is to find a necessary and sufficient condition on the weights (non-negative locally integrable functions) w and v such that an operator $T : L^p(w) \to L^p(v)$ is bounded for all $1 < p < \infty$, i.e. the inequality

$$\int |Tf|^p v dx \leq C^p \cdot \int |f|^p w dx, \text{ for } f \in L^p(w). \tag{1.1}$$

Let $u = w^{-p/p}$. A symmetric formulation of (1.1), well-known from 80s, is

$$\int |T(uf)|^p v dx \leq C^p \cdot \int |f|^p u dx, \text{ for } f \in L^p(u). \tag{1.2}$$

(1.2) looks more natural than (1.1) in the two weight setting: in particular, if T is an integral operator, then the integration in the operator is performed with respect to the same measure udx as in the domain.

Denote $\mu = udx$ and $\nu = vdx$. Let $T^\mu(f) := T(\mu f)$. We can rewrite (1.2) into

$$\int |T^\mu(f)|^p d\nu \leq C^p \cdot \int |f|^p d\mu, \text{ for } f \in L^p(\mu). \tag{1.3}$$

Two weight problems are notoriously hard. The first few results are for T being

- Hardy operator by Muckenhoupt [1].
- Maximal operators by Sawyer [2], where a testing condition is introduced.

- Fractional integrals by Sawyer and Wheeden [3] and [4].

For all these special operators, a characterization for all $1 < p < \infty$ is given. In particlar, Fractional integrals are examples of *positive operators* which are relatively easier and more or less solved. To highlight some results of discrete positive operators, we have results for T being

- Positive dyadic operators and $p = 2$ in [5].
- Positive dyadic operators and $1 < p < \infty$ in [6].
- Vector-valued positive dyadic operators and $1 < p < \infty$ in [9].

Several simplied proofs for the results listed above are also found. For example,

- [7] and [8] simplifies the proof given in [6].
- [10] simplifies the proof given in [9].

In rescent years, there is a breakthrough on this problem for *singular operators*. Initiated by Nazarov, Treil and Volberg, and followed by Lacey, Sawyer, Uriarte-Tuero et al., we have the following results for T being

- Haar multipliers in [5], which is the first case of discrete singular operators.
- Well localized operators including Haar shifts in [11].
- Sufficient conditions for Calderón-Zygmund singular integral operators, and necessary and sufficient conditions for Calderón-Zygmund singular integral operators together with two maximal operators in [12].
- Hilbert Transform in [13] and [14].
- Cauchy Transform in [15].
- Riesz Transform in [16].

Two weight problems for general Calderón-Zygmund singular integral operators remain unsolved. Note the original Two Weight Problems (1.1), (1.2), (1.3) make sence for all $1 < p < \infty$. However, rescent results in [5], [11]-[16] only consider for $p = 2$.

The two weight problems we are interested in are both discrete ones. The first is a new two weight estimates for *Vector-valued positive operators* when $1 < p < \infty$.

The second is a two weight estimates for *Paraproducts* when $1 < p < \infty$. Our setup follows from the one in [17], which is more general than the dyadic case.

DEFINITION 1.1. For a measurable space $(\mathcal{X}, \mathcal{T})$, a *lattice* $\mathcal{L} \subseteq \mathcal{T}$ is a collection of measurable subsets of $\mathcal{X}$ with the following properties

(i) $\mathcal{L}$ is a union of *generations* $\mathcal{L}_n, n \in \mathbb{Z}$, where each generation is a collection of disjoint measurable sets (call them *intervals*), covering $\mathcal{X}$.

(ii) For each $n \in \mathbb{Z}$, the covering $\mathcal{L}_{n+1}$ is a countable refinement of the covering $\mathcal{L}_n$, i.e. each interval $I \in \mathcal{L}_n$ is a countable union of disjoint intervals $J \in \mathcal{L}_{n+1}$. We allow the situation where there is only one such interval J, i.e. $J = I$; this means that $I \in \mathcal{L}_n$ also belongs to the generation $\mathcal{L}_{n+1}$.

DEFINITION 1.2. For an interval $I \in \mathcal{L}$, let $\mathrm{rk}(I)$ be the *rank* of the interval I, i.e. the largest number n such that $I \in \mathcal{L}_n$. For an interval $I \in \mathcal{L}$, $\mathrm{rk}(I) = n$, a *child* of I is an interval $J \in \mathcal{L}_{n+1}$ such that $J \subseteq I$ (actually, $J \subsetneqq I$). The colletion of all children of I is denoted by $\mathrm{child}(I)$. Correspondingly, I is called the *parent* of J.

DEFINITION 1.3. For a positive measure μ on $(\mathcal{X}, \mathcal{T})$, define the *averaging operator* as

$$\mathbb{E}^{\mu}_{I} f = \langle f \rangle_{I,\mu} \mathbf{1}_{I} = \left(\mu(I)^{-1} \int_I f d\mu \right) \mathbf{1}_{I}. \tag{1.4}$$

where $\mathbf{1}_{I}$ is the indicator function of the interval I. The *martingale difference operator* is then defined to be

$$\Delta^{\mu}_{I} f = -\mathbb{E}^{\mu}_{I} f + \sum_{J \in \mathrm{child}(I)} \mathbb{E}^{\mu}_{J} f. \tag{1.5}$$

From now on, we assume $(\mathcal{X}, \mathcal{T})$ is a measurable space, $\mathcal{L} \subseteq \mathcal{T}$ is a lattice on $\mathcal{X}$, and μ, ν are two positive measures.

We denote the conjugate Hölder exponent of p by p', where $1/p + 1/p' = 1$. Here, and throughout the thesis, we use the notation $A \lesssim B$ meaning that there exists an absolute constant C, such that $A \leq CB$, and we write $A \approx B$ if $A \lesssim B \lesssim A$.

Definition 1.4. Let $\boldsymbol{\alpha} = \{\alpha_{_I} : I \in \mathcal{L}\}$ be non-negative constants associated to a lattice $\mathcal{L}$ on $(\mathcal{X}, \mathcal{T})$. Define a vector-valued operator

$$\boldsymbol{T}^{\mu}_{\boldsymbol{\alpha}} f = \{\alpha_{_I} \cdot \mathbb{E}^{\mu}_{_I} f\}_{I \in \mathcal{L}}. \tag{1.6}$$

Theorem 1.5 (Two weight estimates for a vector-valued positive operator [18]). *Let* $1 < p < \infty$ *and* $1 \leq q < \infty$.

$$\int_{\mathcal{X}} \left[\sum_{I \in \mathcal{L}} \left| \alpha_{_I} \cdot \mathbb{E}^{\mu}_{_I} f \right|^q \right]^{\frac{p}{q}} d\nu \leq C^p \cdot \int_{\mathcal{X}} |f|^p d\mu, \tag{1.7}$$

holds if and only if

(i) *for the case* $1 < p \leq q$, *we have*

$$\int_J \left| \sum_{I \in \mathcal{L} : I \subseteq J} \alpha^q_{_I} \cdot \mathbf{1}_{_I} \right|^{\frac{p}{q}} d\nu \leq C^p_1 \cdot \mu(J), \ J \in \mathcal{L} \tag{1.8}$$

(ii) *for the case* $q < p < \infty$, *we have both* (1.8) *and*

$$\int_J \left| \sum_{I \in \mathcal{L} : I \subseteq J} \alpha^q_{_I} \cdot \frac{\nu(I)}{\mu(I)} \cdot \mathbf{1}_{_I} \right|^{\left(\frac{p}{q}\right)'} d\mu \leq C_2^{q\left(\frac{p}{q}\right)'} \cdot \nu(J), \ J \in \mathcal{L}. \tag{1.9}$$

In particular, $C \approx C_1 + C_2$.

Remark 1.6. Vector-valued positive operators in the thesis are viewed as a simple model of the paraproducts defined below.

Definition 1.7. For a measurable function b, the *paraproduct operator with symbol* b is

$$\pi^{\mu}_b f = \sum_{I \in \mathcal{L}} \left(\mathbb{E}^{\mu}_{_I} f \right) \left(\Delta^{\nu}_{_I} b \right). \tag{1.10}$$

Paraproducts play an important role in the investigation of the weighted inequalities for the singular integral operators. The L^2-boundedness of paraproducts is easy, a necessary and sufficient condition (1.12) follows immediately from the Carleson Embedding Theorem.

This necessary and sufficient condition (1.12) can be stated as a *testing condition*, i.e. a paraproduct is bounded in L^2 if and only if there is a uniform estimate on all intervals. In the classical non-weighted situation, the L^2 boundedness is equivalent to the boundedness of the paraproduct in *all* L^p, $1 < p < \infty$.

The weighted situation is much more interesting. It was shown in [17] that in the one weight situation, the testing condition (1.12) is still necessary and sufficient, but it now depends on p: the boundedness in L^{p_0} implies the boundedness in L^p with $1 < p \leq p_0$, but not in L^p with $p_0 < p < \infty$.

Two weight case becomes even more interesting: while it is not hard to show that for $p \leq 2$ the testing condition is still sufficient for the boundedness, we will present a counterexample showing that for $p > 2$ the testing condition alone does not work.

THEOREM 1.8 (Two weight estimates for paraproducts [19]).

$$||\pi_b^\mu f||^p_{L^p(\nu)} \leq C^p \cdot ||f||^p_{L^p(\mu)}, \tag{1.11}$$

holds if and only if

(i) *for the case* $1 < p \leq 2$*, we have*

$$\int_J \left[\sum_{I \in \mathcal{L}, I \subseteq J} \left| \Delta^\nu_I b \right|^2 \right]^{\frac{p}{2}} d\nu \leq C_1^p \cdot \mu(J), \;\; J \in \mathcal{L}, \tag{1.12}$$

(ii) *for the case* $2 < p < \infty$*, we have both (1.12) and*

(1.13)

$$\int_{J'} \left[\sum_{I \in \mathcal{L}} \sum_{\substack{I' \in \text{child}(I) \\ I' \subsetneq J'}} \frac{\nu(I')}{\mu(I)} \mathbb{E}^\nu_I \left(\left| \Delta^\nu_I b \right|^2 \right) \right]^{(\frac{p}{2})'} d\mu \leq C_2^{2(\frac{p}{2})'} \cdot \nu(J'), \;\; J \in \mathcal{L}, J' \in \text{child}(J).$$

In particular, $C \approx C_1 + C_2$*.*

2. Bellman Functions on filtered probability spaces

Denote the Lebesgue measure of a set E by $|E|$, the average value of f on an interval I by $\langle f \rangle_I$. The celebrated dyadic Carleson Embedding Theorem states

Theorem 1.9 (Dyadic Carleson Embedding Theorem). *Let $\mathcal{D} = \{([0,1)+j)\cdot 2^k : j,k \in \mathbb{Z}\}$ be the standard dyadic lattice on $\mathbb{R}$, and let $\{\alpha_{_I}\}_{I\in\mathcal{D}}$ be a sequence of non-negative numbers satisfying the Carleson condition that: $\sum_{J\in\mathcal{D}, J\subseteq I} \alpha_{_I} \le C|I|$ holds for all dyadic intervals $I \in \mathcal{D}$. Then the embedding*

$$\sum_{I\in\mathcal{D}} \alpha_{_I} |\langle f\rangle_{_I}|^p \le C_p \cdot C\|f\|^p_{L^p} \text{ holds for all } f\in L^p, \text{ where } p>1. \tag{1.14}$$

Moreover, the constant $C_p = (p')^p$ is sharp (cannot be replaced by a smaller one).

An approach of proving Theorem 1.9 is the introduction of the Bellman function. Without loss of generality, we can assume $f \ge 0$. Following [20] and [21], we define the Bellman function in three variables $(F, \mathbf{f}, M)$ as

(1.15)

$$\mathcal{B}(F,\mathbf{f},M;C) = \sup\left\{|J|^{-1}\sum_{I\in\mathcal{D}, I\subseteq J} \alpha_{_I}\langle f\rangle^p_{_I} : f, \{\alpha_{_I}\}_{I\in\mathcal{D}} \text{ satisfy (i), (ii), (iii), and (iv)}\right\},$$

(i) $\langle f^p\rangle_{_J} = F$; (ii) $\langle f\rangle_{_J} = \mathbf{f}$; (iii) $|J|^{-1}\sum_{I\subseteq J}\alpha_{_I} = M$; (iv) $\sum_{I\subseteq J}\alpha_{_I} \le C|J|$ for all $J\in\mathcal{D}$.

Note that the Bellman function $\mathcal{B}(F,\mathbf{f},M;C)$ defined above dose not depend on the choice of the interval J.

In [20], (1.14) was first proved using the Bellman function method for the case $p = 2$, and in [21], the sharpness for the case $p = 2$ was also claimed. Later, A. Melas found in [22] the exact Bellman function for all $p > 1$ in a tree-like setting using combinatorial and covering reasoning. In [23], an alternative way of finding the exact Bellman function based on Monge-Ampère equation was also established.

The Bellman functions have deep connetions to the Stochastic Optimal Control theory [21]. Finding the exact Bellman functions is a difficult task. Both the combinatorial methods in [22] and the methods of solving the Bellman PDE in [23] are quite complicated. Luckily, the proof of Theorem 1.9 only needs a super-solution instead of the exact Bellman function, see [20], [21]. In this thesis, we will present a way of calculating a super-solution via the Burkhölder's hull.

On the other hand, computation of the exact Bellman functions usually reflects deeper structure of the corresponding harmonic analysis problem. It is interesting to note that the exact Bellman function of Theorem 1.9 is not restricted to the standard dyadic lattice. In [22], it also works for the tree-like structure. Let us consider a more general situation here.

Let $(\mathcal{X}, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mu)$ be a discrete-time filtered probability space. By a discrete-time filtration, we mean a sequence of non-decreasing σ-fields

$$\{\emptyset, \mathcal{X}\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_n \subseteq \ldots \subseteq \mathcal{F}.$$

We introduce notations $f_n = \mathbb{E}^\mu[f|\mathcal{F}_n]$ and $\langle f\rangle_{E,\mu} = \mu(E)^{-1}\int_E f\,d\mu$.

Definition 1.10. A sequence of non-negative random variables $\{\alpha_n\}_{n\geq 0}$ is called a *Carleson sequence*, if each α_n is $\mathcal{F}_n$-measurable and

$$\mathbb{E}^\mu\left[\sum_{k\geq n}\alpha_k|\mathcal{F}_n\right] \leq C \text{ for every } n \geq 0. \tag{1.16}$$

Definition 1.11. $\{\mathcal{F}_n\}_{n\geq 0}$ is called an *infinitely refining filtration*, if for every $\varepsilon > 0$, every $n \geq 0$ and every set $E \in \mathcal{F}_n$, there exists a real-valued $\mathcal{F}_k$-measurable $(k > n)$ random variable h, such that: (i) $|h\mathbf{1}_E| = \mathbf{1}_E$ and (ii) $\int_E |h_n| d\mu \leq \varepsilon$.

Theorem 1.12 (Martingale Carleson Embedding Theorem). *If $f \in L^p(\mathcal{X}, \mathcal{F}, \mu)$ and $\{\alpha_n\}_{n\geq 0}$ is a Carleson sequence, then*

$$\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n|f_n|^p\right] \leq C_p \cdot C \cdot \mathbb{E}^\mu\left[|f|^p\right]. \tag{1.17}$$

Moreover, if $\{\mathcal{F}_n\}_{n\geq 0}$ is an infinitely refining filtration, then the constant $C_p = (p')^p$ is sharp.

Here, again without loss of generality, we can assume $f \geq 0$. We define the Bellman function $\mathcal{B}^{\mathcal{F}}_\mu(F, \mathbf{f}, M; C)$ in the martingale setting by

(1.18)

$$\mathcal{B}^{\mathcal{F}}_\mu(F, \mathbf{f}, M; C) = \sup\left\{\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n f_n^p\right] : f, \{\alpha_n\}_{n\geq 0} \text{ satisfy (i), (ii), (iii) and (iv)}\right\},$$

(i) $\mathbb{E}^\mu[f^p] = F$; (ii) $\mathbb{E}^\mu[f] = \mathbf{f}$; (iii) $\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right] = M$; (iv) $\{\alpha_n\}_{n\geq 0}$ satisfies (1.16).

Now, we are ready to state the first main theorem.

THEOREM 1.13 (Coincidence of the Bellman functions [24]).

$$\mathcal{B}_\mu^{\mathcal{F}}(F,\mathbf{f},M;C) \leq \mathcal{B}(F,\mathbf{f},M;C). \tag{1.19}$$

Moreover, if $\{\mathcal{F}_n\}_{n\geq 0}$ is an infinitely refining filtration, then

$$\mathcal{B}_\mu^{\mathcal{F}}(F,\mathbf{f},M;C) = \mathcal{B}(F,\mathbf{f},M;C). \tag{1.20}$$

For the Doob's martingale inequality, recall the definition of the maximal function associated to a discrete-time filtration $\{\mathcal{F}_n\}_{n=0}^\infty$

$$f^*(x) = \sup_{n\geq 0}|f_n(x)|. \tag{1.21}$$

THEOREM 1.14 (Doob's Martingale Inequality). *For every $p>1$ and every $f\in L^p(\mathcal{X},\mathcal{F},\mu)$, we have*

$$||f^*||^p_{L^p(\mathcal{X},\mathcal{F},\mu)} \leq (p')^p\cdot ||f||^p_{L^p(\mathcal{X},\mathcal{F},\mu)}. \tag{1.22}$$

Moreover, if $\{\mathcal{F}_n\}_{n\geq 0}$ is an infinitely refining filtration, then the constant $(p')^p$ is sharp.

The study of the L^p-norm of the maximal function was initiated from the celebrated Doob's martingale inequality, e.g. in [30]. The sharpness of this inequality was shown in [26] and [27] if one looks at all martingales. For particular martingales including the dyadic case, see [22] and [28]. Theorem 1.14 covers all these results.

Assuming $f\geq 0$, we define the Bellman function $\widetilde{\mathcal{B}}_\mu^{\mathcal{F}}(F,\mathbf{f})$ associated to the Doob's martingale inequality by

$$\widetilde{\mathcal{B}}_\mu^{\mathcal{F}}(F,\mathbf{f}) = \sup\left\{\mathbb{E}^\mu\left[|f^*|^p\right] : \mathbb{E}^\mu[f^p] = F,\ \mathbb{E}^\mu[f] = \mathbf{f}\right\}. \tag{1.23}$$

The connection between the Carleson Embedding Theorem and the maximal theory has been known and exploited a lot, e.g. in [20] and [22]. Using this connection, we give a proof of the second main theorem.

THEOREM 1.15 (The Bellman function of the maximal operators [24]).

$$\widetilde{\mathcal{B}}_{\mu}^{\mathcal{F}}(F, \mathbf{f}) \leq \mathcal{B}_{\mu}^{\mathcal{F}}(F, \mathbf{f}, 1; C = 1). \tag{1.24}$$

Moreover, if $\{\mathcal{F}_n\}_{n\geq 0}$ *is an infinitely refining filtration, then*

$$\widetilde{\mathcal{B}}_{\mu}^{\mathcal{F}}(F, \mathbf{f}) = \mathcal{B}_{\mu}^{\mathcal{F}}(F, \mathbf{f}, 1; C = 1). \tag{1.25}$$

3. Outline of the thesis

In chapter 2, we prove Theorem 1.5. We discuss two cases $1 < p \leq q$ and $q < p < \infty$ seperately. The theorem is actually equivalent to two weight estimates for positive dyadic operators.

In chapter 3, we prove Theorem 1.8. Again, we discuss two cases $1 < p \leq q$ and $q < p < \infty$ seperately. The sufficiency part of the case $q < p < \infty$ needs to be done in greater detail.

In chapter 4, we find a super-solution of Theorem 1.9 via the Burkhölder's hull, which proves the existence of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$.

In chapter 5, we present a remodeling of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$ and use this to prove the two main results Theorem 1.13 and Theorem 1.15.

CHAPTER 2

Two weight estimates for a vector-valued positive operators

In this chapter, we prove Theorem 1.5. We first discuss the easier case $1 < p \leq q$. Then we present a counterexample that (1.8) itself dose not imply (1.7). Enventually, we reduce Theorem 1.5 to the well-known two weight estimates for positive dyadic operators and complete our proof. A comparison of our theorem and the main results in [9] and [10] is also given.

1. The case $1 < p \leq q$

We will see in this section that when $1 < p \leq q$, (1.8) is equivalent to (1.7). On one hand, (1.8) can be deduced from (1.7) by setting $f = \mathbf{1}_J$. On the other hand, consider the maximal function

$$M_\mu f(x) := \sup_{x \in I, I \in \mathcal{L}} \left| \mathbb{E}_I^\mu f(x) \right|. \tag{2.1}$$

The celebrated *Doob's martingale inequality* asserts

$$\left|\left| M_\mu f \right|\right|_{L^p(\mu)} \leq p' \cdot \left|\left| f \right|\right|_{L^p(\mu)}. \tag{2.2}$$

Let $E_k := \{x \in \mathcal{X} : M_\mu f(x) > 2^k\}$ and let $\mathcal{E}_k := \{I \in \mathcal{L} : I \in E_k\}$. Note that E_k is a disjoint union of maximal intervals in $\mathcal{E}_k$, maximal in the sense of inclusion. Denote these disjoint maximal intervals by $\mathcal{E}_k^*$. Hence, $E_k = \cup_{J \in \mathcal{E}_k^*} J$.

$$\int_{\mathcal{X}}\left[\sum_{I\in\mathcal{L}}\left|\alpha_{_I}\cdot\mathbb{E}^{\mu}_{I}f\right|^{q}\right]^{\frac{p}{q}}d\nu\le\sum_{k}\int_{E_k}\left[\sum_{I\in\mathcal{E}_k\setminus\mathcal{E}_{k+1}}\left|\alpha_{_I}\cdot\mathbb{E}^{\mu}_{I}f\right|^{q}\right]^{\frac{p}{q}}d\nu,\ 1<p\le q$$

$$\le\sum_{k}2^{(k+1)p}\int_{E_k}\left[\sum_{I\in\mathcal{E}_k\setminus\mathcal{E}_{k+1}}\alpha^{q}_{_I}\cdot\mathbf{1}_{_I}\right]^{\frac{p}{q}}d\nu$$

$$\le\sum_{k}2^{(k+1)p}\sum_{J\in\mathcal{E}^{*}_{k}}\int_{J}\left[\sum_{I\in\mathcal{L}:I\subseteq J}\alpha^{q}_{_I}\cdot\mathbf{1}_{_I}\right]^{\frac{p}{q}}d\nu$$

$$\le C_1^p\cdot\sum_{k}2^{(k+1)p}\cdot\mu(E_k),\ (1.8)$$

$$\lesssim C_1^p\cdot||M_\mu f||^p_{L^p(\mu)}$$

$$\le C_1^p\cdot(p')^p\cdot||f||^p_{L^p(\mu)},\ (2.2).$$

2. The case $q<p<\infty$: a counterexample

In this section, we see that (1.8) itself is not sufficient for (1.7) for the case $q<p<\infty$.

Consider the real line $\mathbb{R}$ with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. Let the lattice be all the tri-adic intervals. We specify the positive measures μ,ν, the non-negative constants $\boldsymbol{\alpha}=\{\alpha_{_I}:I\in\mathcal{L}\}$, and the functions f in the following way.

Let $C=\cap_{n\ge0}C_n$ be the 1/3-Cantor set, where $C_0=[0,1),C_1=[0,1/3)\cup[2/3,1)$ and, in general, $C_n=\cup\left\{[x,x+3^{-n}):x=\sum_{j=1}^{n}\varepsilon_j3^{-j},\varepsilon_j\in\{0,2\}\right\}$.

- (i) The measure μ is the Lebsgue measure restricted on $[0,1)$ and the measure ν is the Cantor measure, i.e. $\nu(I)=2^{-n}$ for each I belongs to a connect component of C_n.
- (ii) Define $\alpha_{_I}=(2/3)^{n/p}$ for each I belongs to a connect component of C_n.
- (iii) For the function f, consider the gap of C, i.e. $[0,1)\setminus C$. This is a disjoint union of tri-adic intervals. Let $f=(3/2)^{n/p}\cdot n^{-r}$ for each $I\in[0,1)\setminus C$ with length of I equals 3^{-n}, where r is to be chosen later.

Claim 2.1. The construction gives a counterexample with properly chosen r.

Proof. We begin with checking (1.8). It suffices to check for every J belongs to a connected component of C_n, and thus $\mu(J) = 3^{-n}$. Note that

$$\left|\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I^q \cdot \mathbf{1}_I\right|^{\frac{p}{q}} \leq \left|\sum_{k\geq n}\left(\frac{2}{3}\right)^{\frac{qk}{p}}\right|^{\frac{p}{q}} \asymp \left(\frac{2}{3}\right)^n.$$

Hence,

$$\int_J \left|\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I^q \cdot \mathbf{1}_I\right|^{\frac{p}{q}} d\nu \lesssim \left(\frac{2}{3}\right)^n \cdot \nu(J) = \mu(J).$$

Next, we show that (1.7) fails. This requires a careful choice of r in the definition of f. Picking $r > \frac{1}{p}$, we have

$$||f||^p_{L^p(\mu)} = \int_0^1 |f|^p dx = \sum_{n\geq 1}\left(\frac{3}{2}\right)^n n^{-pr} \cdot \frac{1}{3^n} \cdot 2^n = \sum_{n\geq 1} n^{-pr} < \infty.$$

Since $q < p < \infty$, we can pick r such that $\frac{1}{p} < r < \frac{1}{q}$. Note that for every I belongs to a connected component of C_n, we have

$$\mathbb{E}_I^\mu f \geq \frac{1}{3}\left(\frac{3}{2}\right)^{\frac{n+1}{p}} (n+1)^{-r}.$$

Hence, consider $\mathcal{I}_n = \{I : I \text{ is tri-adic with length less than or equal to } 3^{-n}\}$,

$$\sum_{I\in\mathcal{I}_n} \left|\alpha_I \cdot \mathbb{E}_I^\mu f\right|^q \cdot \mathbf{1}_{C_n} \geq \sum_{k\leq n}\left|\frac{1}{3}\left(\frac{3}{2}\right)^{\frac{1}{p}}(k+1)^{-r}\right|^q \gtrsim \sum_{k\leq n}(k+1)^{-qr}.$$

And so,

$$\int\left[\sum_{I\in\mathcal{I}_n}\left|\alpha_I \cdot \mathbb{E}_I^\mu f\right|^q\right]^{\frac{p}{q}} d\nu \gtrsim \left[\sum_{k\leq n}(k+1)^{-qr}\right]^{\frac{p}{q}} \cdot \nu(C_n) \to \infty \text{ as } n \to \infty.$$

We can see that the condition $q < p < \infty$ is crutial in our construction. □

3. The case $q < p < \infty$: necessity and sufficiency

We discuss the case $q < p < \infty$ of Theorem 1.5 in this section. In particular, we see that both (1.8) and (1.9) are testing conditions on some families of special functions.

To start, since

$$||\boldsymbol{T}^{\mu}_{\boldsymbol{\alpha}} f||^q_{L^p(l^q,\nu)} = \sup_{||g||_{L^{(p/q)'}(\nu)}=1} \int_{\mathcal{X}} \left[\sum_{I \in \mathcal{L}} \left| \alpha_{_I} \cdot \mathbb{E}^{\mu}_{_I} f \right|^q \right] g d\nu, \tag{2.3}$$

we can write

$$||\boldsymbol{T}^{\mu}_{\boldsymbol{\alpha}}||^q_{L^p(\mu)\to L^p(l^q,\nu)} = \sup_{||f||_{L^p(\mu)}=1} \sup_{||g||_{L^{(p/q)'}(\nu)}=1} \int_{\mathcal{X}} \left[\sum_{I \in \mathcal{L}} \left| \alpha_{_I} \cdot \mathbb{E}^{\mu}_{_I} f \right|^q \right] g d\nu. \tag{2.4}$$

Without loss of generality, we assume that both f and g are non-negative. The following lemma reduces us to the scalar-valued case.

Lemma 2.2.

$$||\boldsymbol{T}^{\mu}_{\boldsymbol{\alpha}}||^q_{L^p(\mu)\to L^p(l^q,\nu)} \approx \sup_{||f||_{L^p(\mu)}=1} \sup_{||g||_{L^{(p/q)'}(\nu)}=1} \int_{\mathcal{X}} \left[\sum_{I \in \mathcal{L}} \alpha^q_{_I} \cdot \mathbb{E}^{\mu}_{_I}(f^q) \right] g d\nu. \tag{2.5}$$

An easy application of Hölder's inequality shows that the LHS of (2.5) is no more than its RHS. The other half of this lemma depends on the following famous *Rubio de Francia Algorithm.*

Lemma 2.3 (Rubio de Francia Algorithm). *For every $q < p < \infty$ and $f \in L^p(\mu)$, there exists a function $F \in L^p(\mu)$, such that $f \le F$, $||F||_{L^p(\mu)} \approx |f||_{L^p(\mu)}$ and*

$$\mu(I)^{-1} \int_I F^q d\mu \lesssim \inf_{x\in I} F^q(x), \ \ I \in \mathcal{L}.$$

Proof. Consider the maximal operator M_μ defined in (2.1). Doob's martingale inequality (2.2) implies

$$||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \le \left(\frac{p}{q}\right)'. \tag{2.6}$$

Denote $M^{(0)}_\mu = Id$, $M^{(1)}_\mu = M_\mu$ and $M^{(k)}_\mu = M_\mu \circ M^{(k-1)}_\mu$. Define the function F by

$$F = \left[\sum_{k\ge 0} \left(2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{-k} M^{(k)}_\mu (f^q) \right]^{\frac{1}{q}}. \tag{2.7}$$

First we check the validity of the definition for F. Note that

$$
\begin{aligned}
||F||_{L^p(\mu)}^q &= \left\{ \int_{\mathcal{X}} \left[\sum_{k\geq 0} \left(2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{-k} M_\mu^{(k)}(f^q) \right]^{\frac{p}{q}} d\mu \right\}^{\frac{q}{p}} \\
&\leq \sum_{k\geq 0} \left(2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{-k} \left(\int_{\mathcal{X}} \left| M_\mu^{(k)}(f^q) \right|^{\frac{p}{q}} d\mu \right)^{\frac{q}{p}}, \text{ Minkowski inequality} \\
&\leq \sum_{k\geq 0} \left(2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{-k} \left(||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{k} ||f||_{L^p(\mu)}^q = 2||f||_{L^p(\mu)}^q.
\end{aligned}
$$

Hence, F is the $L^{p/q}(\mu)$-limit of the partial sums and thus well-defined. Moreover, we have also proved that $||F||_{L^p(\mu)} \lesssim ||f||_{L^p(\mu)}$.

Considering only $k = 0$ in the definition for F, we have $F \geq f$. And so $||F||_{L^p(\mu)} \approx ||f||_{L^p(\mu)}$. Finally, note that

$$
\mu(I)^{-1} \int_I F^q d\mu \leq \inf_{x\in I} \ M_\mu(F^q)(x) \tag{2.8}
$$

and

$$
\begin{aligned}
M_\mu(F^q) &= \sum_{k\geq 0} \left(2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} \right)^{-k} M_\mu^{(k+1)}(f^q) \\
&= 2||M_\mu||_{L^{p/q}(\mu)\to L^{p/q}(\mu)} (F^q - f^q) \lesssim F^q.
\end{aligned}
$$

Therefore, we deduce

$$
\mu(I)^{-1} \int_I F^q d\mu \lesssim \inf_{x\in I} \ F^q(x), \ I \in \mathcal{L}.
$$

□

Applying Rubio de Francia Algorithm, we obtain

$$
\begin{aligned}
\int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \alpha_I^q \cdot \mathbb{E}_I^\mu(f^q) \right] g d\nu &\leq \int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \alpha_I^q \cdot \mathbb{E}_I^\mu(F^q) \right] g d\nu \\
&\lesssim \int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \alpha_I^q \cdot \left(\mathbb{E}_I^\mu(F) \right)^q \right] g d\nu \\
&\leq ||\boldsymbol{T}_{\boldsymbol{\alpha}}^\mu||_{L^p(\mu)\to L^p(l^q,\nu)}^q \cdot ||F||_{L^p(\mu)} \cdot ||g||_{L^{(p/q)'}(\nu)}, \ (2.4) \\
&\lesssim ||\boldsymbol{T}_{\boldsymbol{\alpha}}^\mu||_{L^p(\mu)\to L^p(l^q,\nu)}^q, \ \left(||f||_{L^p(\mu)} = ||g||_{L^{(p/q)'}(\nu)} = 1 \right).
\end{aligned}
$$

Now that our problem is reduced to determine a necessary and sufficient condition of

$$\int_{\mathcal{X}} \left|\sum_{I\in\mathcal{L}} \alpha_I^q \cdot \mathbb{E}_I^\mu(f)\right|^{\frac{p}{q}} d\nu \lesssim C^p \int_{\mathcal{X}} |f|^{\frac{p}{q}} d\mu, \tag{2.9}$$

we may consult to the scalar-valued Theorem 2.4 below. Therefore, Theorem 1.8 follows from Theorem 2.4 for free, and both (1.8) and (1.9) are testing conditions with respect to this derived scalar-valued problem.

Consider the linear operator defined by

$$T_{\boldsymbol{\alpha}}^\mu f := \sum_{I\in\mathcal{L}} \alpha_I \cdot \mathbb{E}_I^\mu f. \tag{2.10}$$

THEOREM 2.4. *Let $1 < p < \infty$ and let $1/p + 1/p' = 1$. $T_{\boldsymbol{\alpha}}^\mu : L^p(\mu) \to L^p(\nu)$ if and only if*

$$\int_J \left|\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I \cdot \mathbf{1}_I\right|^p d\nu \le C_1^p \cdot \mu(J),\ \ J \in \mathcal{L} \tag{2.11}$$

$$\int_J \left|\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I \cdot \frac{\nu(I)}{\mu(I)} \cdot \mathbf{1}_I\right|^{p'} d\mu \le C_2^{p'} \cdot \nu(J),\ \ J \in \mathcal{L}. \tag{2.12}$$

In particular, $||T_{\boldsymbol{\alpha}}^\mu||_{L^p(\mu)\to L^p(\nu)} \approx C_1 + C_2$.

REMARK 2.5. Theorem 2.4 is originally proved for the dyadic case. This general version is explained in [7].

REMARK 2.6. In [9] and [10], to obtain the two testing conditions, they first rewrite (1.7) into

$$\sum_{I\in\mathcal{L}} \alpha_I \cdot \mathbb{E}_I^\mu f \cdot \mathbb{E}_I^\nu g_I \cdot \nu(I) \le C||f||_{L^p(\mu)} \cdot ||\{g_I\}_{I\in\mathcal{L}}||_{L^{p'}(l^q,\nu)}. \tag{2.13}$$

Setting $f = \mathbf{1}_J$, one deduces (1.8). For the second testing condition, one turns to consider the family of functions $\{g_I\}_{I\in\mathcal{L}}$ supported on $J \in \mathcal{L}$ with $L^\infty(l^q,\nu)$-norm equal to 1. This gives

$$\int_J \left|\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I \cdot \frac{\nu(I)}{\mu(I)} \cdot \mathbb{E}_I^\nu g_I\right|^{p'} d\mu \le C^{p'} \cdot \nu(J),\ \ J \in \mathcal{L}. \tag{2.14}$$

Compare Theorem 1.5 with the main results in [**9**] and [**10**]. We have a very different condition (1.9) than (2.14) with seemingly *'wrong'* exponents. However, both (1.8) and (1.9) are testing conditions on some families of special functions as we have shown in this chapter.

CHAPTER 3

Two weight estimates for paraproducts

In this chapter, we prove Theorem 1.8. We first follow the line of chapter 2. But for the sufficient part of the case $2 < p < \infty$, we need to be more careful.

We start with some useful reductions. Obviously, it sufficies to constider only for non-negative functions $f \geq 0$ in Theorem 1.8. Moreover, recall the following version of Littlewood-Paley theorem

THEOREM 3.1 (Littlewood-Paley). *If a function f has a Littlewood-Paley decomposition $f = \sum_{I\in\mathcal{L}} \Delta^{\nu}_{I} f$, and define its square function to be $S^{\nu} f = \left[\sum_{I\in\mathcal{L}} \left|\Delta^{\nu}_{I} f\right|^2\right]^{\frac{1}{2}}$, then $||f||_{L^p(\nu)} \asymp ||S^{\nu} f||_{L^p(\nu)}$ for all $1 < p < \infty$.*

PROOF. See [17]. □

Applying this theorem, (1.11) is equivalent to

$$||S^{\nu}(\pi^{\mu}_{b} f)||^{p}_{L^p(\nu)} = \int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \left|\mathbb{E}^{\mu}_{I} f\right|^2 \left|\Delta^{\nu}_{I} b\right|^2\right]^{\frac{p}{2}} d\nu \lesssim C^p \int_{\mathcal{X}} |f|^p d\mu. \tag{3.1}$$

We will consider (3.1) instead of (1.11) in the following.

1. Construction of the stopping intervals

Let us construct a collection $\mathcal{G} \subseteq \mathcal{F} \subseteq \mathcal{L}$ of stopping intervals as follows. Given a non-negative function $f \geq 0$. For $J \in \mathcal{F}$, let $\mathcal{G}^*(J)$ be the collection of maximal intervals $I \subseteq \mathcal{F}$, $I \in J$ such that

$$\langle f \rangle_{I,\mu} > 2\langle f \rangle_{J,\mu}.$$

In case there are more than one such intervals I, we choose the one from the smallest generation. Note that intervals from $\mathcal{G}^*(J)$ are pairwisely disjoint. Let $\mathcal{F}(J) = \{I \in$

$\mathcal{F} : I \subseteq J\}$ and let $G(J) = \cup_{I \in \mathcal{G}^*(J)} I$. Define also $\mathcal{E}(J) = \mathcal{F}(J) \setminus \cup_{I \in \mathcal{G}^*(J)} \mathcal{F}(I)$. Then we have the following properties

(i) For any $I \in \mathcal{E}(J)$, $\langle f \rangle_{I,\mu} \leq 2 \langle f \rangle_{J,\mu}$,

(ii) $\mu(G(J)) < \frac{1}{2}\mu(J)$.

To construct a collection $\mathcal{G}$, fix some large integer $N \in \mathbb{Z}$ and consider all maximal intervals J from $\{\mathcal{L}_k\}_{k \geq -N}$ and $J \in \mathcal{F}$. These intervals form the first generation $\mathcal{G}_1^*$ of stopping intervals. Inductively define the $(n+1)$-th generation of stopping intervals by $\mathcal{G}_{n+1}^* = \cup_{I \in \mathcal{G}_n^*} \mathcal{G}^*(I)$ and we define the collection of stopping intervals by $\mathcal{G} = \cup_{n \geq 1} \mathcal{G}_n^*$.

Property (ii) implies that the collection $\mathcal{G}$ of stopping intervals satisfies the famous *Carleson measure condition*

$$\sum_{I \in \mathcal{G}, I \subseteq J} \mu(I) < 2\mu(J), \quad J \in \mathcal{L}. \tag{3.2}$$

A special form of the Martingale Carleson Embedding Theorem 1.12 says

THEOREM 3.2. *Let μ be a measure on $(\mathcal{X}, \mathcal{T})$ and let $\alpha_I \geq 0$, $I \in \mathcal{L}$ satisfy the Carleson measure condition*

$$\sum_{I \in \mathcal{G}, I \subseteq J} \alpha_I \leq C \cdot \mu(J). \tag{3.3}$$

Then for any measurable function f and any $1 < p < \infty$

$$\sum_{I \in \mathcal{L}} \alpha_I \left| \langle f \rangle_{I,\mu} \right|^p \leq C \cdot (p')^p \cdot ||f||^p_{L^p(\mathcal{X}, \mathcal{T}, \mu)}. \tag{3.4}$$

2. The case $1 < p \leq 2$

We will see in this section that when $1 < p \leq 2$, (1.12) is equivalent to (3.1). On one hand, (1.12) can be deduced from (3.1) by setting $f = \mathbf{1}_J$. On the other hand, we can apply the stopping intervals with $\mathcal{F} = \mathcal{L}$ constructed in the previous section to obtain

$$\int_{\mathcal{X}}\left[\sum_{I\in\mathcal{L},I\subseteq\mathcal{L}_{-N}}\left|\mathbb{E}^{\mu}_{I}f\right|^2\left|\Delta^{\nu}_{I}b\right|^2\right]^{\frac{p}{2}}d\nu=\int_{\mathcal{X}}\left[\sum_{J\in\mathcal{G}}\sum_{I\subseteq\mathcal{E}(J)}\left|\mathbb{E}^{\mu}_{I}f\right|^2\left|\Delta^{\nu}_{I}b\right|^2\right]^{\frac{p}{2}}d\nu,$$

$$\le\int_{\mathcal{X}}\left[\sum_{J\in\mathcal{G}}4\langle f\rangle^2_{J,\mu}\sum_{I\subseteq\mathcal{E}(J)}\left|\Delta^{\nu}_{I}b\right|^2\right]^{\frac{p}{2}}d\nu,\ 1<p\le 2$$

$$\le 2^p\sum_{J\in\mathcal{G}}\langle f\rangle^p_{J,\mu}\int_J\left[\sum_{I\subseteq\mathcal{E}(J)}\left|\Delta^{\nu}_{I}b\right|^2\right]^{\frac{p}{2}}d\nu,\ (1.12)$$

$$\le 2^p\sum_{J\in\mathcal{G}}\langle f\rangle^p_{J,\mu}\cdot C_1^p\cdot\mu(J),\ (3.4)$$

$$\le C_1^p\cdot 2^{p+1}\cdot(p')^p\cdot||f||^p_{L^p(\mathcal{X},\mathcal{T},\mu)}.$$

Letting $N\to\infty$, we prove exactly (3.1). We can see $1<p\le 2$ plays an important role in this argument. There is no analogue for the case $2<p<\infty$.

3. The case $2<p<\infty$: a counterexample

In this section, we see that (1.12) itself is not sufficient for (3.1) for the case $2<p<\infty$.

Consider the real line $\mathbb{R}$ with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. Let the lattice be all the tri-adic intervals. We specify the admissible measures μ,ν and the functions b,f in the following way.

Let $C=\cap_{n\ge 0}C_n$ be the 1/3-Cantor set, where $C_0=[0,1), C_1=[0,1/3)\cup[2/3,1)$ and, in general, $C_n=\cup\left\{[x,x+3^{-n}):x=\sum_{j=1}^n\varepsilon_j 3^{-j},\varepsilon_j\in\{0,2\}\right\}$.

(i) The measure μ is the Lebsgue measure restricted on $[0,1)$ and the measure ν is the Cantor measure, i.e. $\nu(I)=2^{-n}$ for each I belongs to a connect component of C_n.

(ii) For the function b, we specify its martingale differences $\Delta^{\nu}_{I}b$. Let $|\Delta^{\nu}_{I}b|=(2/3)^{n/p}$ for each I belongs to a connect component of C_n such that $\int_I(\Delta^{\nu}_{I}b)d\nu=0$.

(iii) For the function f, consider the gap of C, i.e. $[0,1)\setminus C$. This is a disjoint union of tri-adic intervals. Let $f = (3/2)^{n/p}\cdot n^{-r}$ for each $I \in [0,1)\setminus C$ with length of I equals 3^{-n}, where r is to be chosen later in the proof.

Claim 3.3. The above construction gives a counterexample.

Proof. We begin with checking (1.12). It suffices to check for every J belongs to a connected component of C_n, and thus $\mu(J) = 3^{-n}$. Note that

$$\left[\sum_{I\subseteq J}\left|\Delta_I^\nu b\right|^2\right]^{\frac{p}{2}} \le \left[\sum_{k\ge n}\left(\frac{2}{3}\right)^{\frac{2k}{p}}\right]^{\frac{p}{2}} \lesssim \left(\frac{2}{3}\right)^n.$$

Hence,

$$\int_J\left[\sum_{I\subseteq J}\left|\Delta_I^\nu b\right|^2\right]^{\frac{p}{2}} d\nu \lesssim \left(\frac{2}{3}\right)^n \nu(J) = \left(\frac{2}{3}\right)^n\cdot\left(\frac{1}{2}\right)^n = \mu(J).$$

Next, we show that (3.1) fails. This requires a careful choice of r in the definition of f. Picking $r > \frac{1}{p}$, we have

$$||f||^p_{L^p(\mu)} = \int_0^1 |f|^p dx = \sum_{n\ge 1}\left(\frac{3}{2}\right)^n n^{-pr}\cdot\frac{1}{3^n}\cdot 2^n = \sum_{n\ge1} n^{-pr} < \infty.$$

Since $2 < p < \infty$, we can pick r such that $\frac{1}{p} < r < \frac{1}{2}$. Note that for every I belongs to a connected component of C_n, we have

$$\mathbb{E}_I^\mu f \ge \frac{1}{3}\left(\frac{3}{2}\right)^{\frac{n+1}{p}}(n+1)^{-r}.$$

Hence, consider $\mathcal{I}_n = \{I : I \text{ is tri-adic with length less than or equal to } 3^{-n}\}$,

$$\sum_{I\in\mathcal{I}_n}\left|\mathbb{E}_I^\mu f\right|^2\left|\Delta_I^\nu b\right|^2\cdot\mathbf{1}_{C_n} \ge \sum_{k\le n}\left|\frac{1}{3}\left(\frac{3}{2}\right)^{\frac{1}{p}}(k+1)^{-r}\right|^2 \gtrsim \sum_{k\le n}(k+1)^{-2r}.$$

And so,

$$\int\left[\sum_{I\in\mathcal{I}_n}\left|\mathbb{E}_I^\mu f\right|^2\left|\Delta_I^\nu b\right|^2\right]^{\frac{p}{2}} d\nu \gtrsim \left[\sum_{k\le n}(k+1)^{-2r}\right]^{\frac{p}{2}}\cdot\nu(C_n)\to\infty \text{ as } n\to\infty.$$

□

4. The case $2 < p < \infty$: trilinear forms and necessity

We discuss the necessity of Theorem 1.8 in this section. In particular, we see that both (1.12) and (1.13) are testing conditions on some families of special functions. To make our explanations more clear and also for later purpose, we generalize Theorem 1.8 to a trilinear form.

(1.12) is a simple testing condtion on functions of the form $f = \mathbf{1}_{J}$, but (1.13) is not that clear. To deduce (1.13) from (3.1), we first note that $\Delta^{\nu}_{I} b$ is constant on each $I' \in \text{child}(I)$. Let $\beta_{II'} = \left(\mu(I)^{-1}\left|\Delta^{\nu}_{I} b\right| \cdot \mathbf{1}_{I'}\right)^2$ for each $I' \in \text{child}(I)$. (3.1) becomes

$$\int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \sum_{I'\in\text{child}(I)} \beta_{II'} \left(\int_I f d\mu\right)^2 \mathbf{1}_{I'}\right]^{\frac{p}{2}} d\nu \lesssim C^p \int_{\mathcal{X}} |f|^p d\mu. \tag{3.5}$$

Consider the following generalization of Theorem 1.8 to a trilinear form.

Theorem 3.4 (Two weight estimates for a trilinear form). *For every sequence of non-negative constants* $\left\{\beta_{II'}\right\}_{I\in\mathcal{L}, I'\in\text{child}(I)}$*, define the trilinear operator*

$$\Pi(f,g,h) = \sum_{I\in\mathcal{L}} \sum_{I'\in\text{child}(I)} \beta_{II'} \left(\int_I f d\mu\right)\left(\int_I g d\mu\right)\left(\int_{I'} h d\nu\right). \tag{3.6}$$

$$\Pi(f,g,h) \le C||f||_{L^p(\mu)}||g||_{L^p(\mu)}||h||_{L^{(\frac{p}{2})'}(\nu)} \tag{3.7}$$

holds if and only if

(i)

$$\int_J \left[\sum_{I\in\mathcal{L}, I\subseteq J} \sum_{I'\in\text{child}(I)} \beta_{II'} \cdot \mu(I)^2 \cdot \mathbf{1}_{I'}\right]^{\frac{p}{2}} d\nu \le C_1^{\frac{p}{2}} \cdot \mu(J), \;\; J \in \mathcal{L}, \tag{3.8}$$

(ii)

(3.9)

$$\int_{J'} \left[\sum_{I\in\mathcal{L}} \sum_{I'\in\text{child}(I), I'\subsetneq J'} \beta_{II'} \cdot \mu(I)\cdot\nu(I') \cdot \mathbf{1}_{I}\right]^{(\frac{p}{2})'} d\mu \le C_2^{(\frac{p}{2})'} \cdot \nu(J'), \;\; J \in \mathcal{L}, J' \in \text{child}(J).$$

In particular, $C \approx C_1 + C_2$.

Remark 3.5. Note that Theorem 3.4 is written in duality form. In Theorem 3.4, if we choose $g = f$ and $\beta_{II'} = \left(\mu(I)^{-1}\left|\Delta_I^\nu b\right| \cdot \mathbf{1}_{I'}\right)^2$, and take care of the powers of the constants, then we recover Theorem 1.8.

All amounts to deduce (3.9) from (3.7). The argument depends on again the Rubio de Francia Algorithm Lemma 2.3. Let $f = g = F$ in (3.7), we obtain

$$\Pi(F, F, h) = \sum_{I\in\mathcal{L}} \sum_{I'\in \mathrm{child}(I)} \beta_{II'} \left(\int_I F d\mu\right)^2 \left(\int_{I'} h d\nu\right)$$
$$= \int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \sum_{I'\in \mathrm{child}(I)} \beta_{II'} \left(\int_I F d\mu\right)^2 \cdot \mathbf{1}_{I'}\right] h d\nu \le C||F||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

By Lemma 2.3 with $q = 2$, we have $||F||_{L^p(\mu)} \approx ||f||_{L^p(\mu)}$ and

$$\left(\int_I F d\mu\right)^2 \ge \mu(I)^2 \cdot \inf_{x\in I} F^2(x) \gtrsim \mu(I) \cdot \int_I F^2 d\mu \ge \mu(I) \cdot \int_I f^2 d\mu,$$

thus we deduce

$$\int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \sum_{I'\in \mathrm{child}(I)} \beta_{II'} \cdot \mu(I) \left(\int_I f^2 d\mu\right) \cdot \mathbf{1}_{I'}\right] h d\nu \lesssim C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}, \tag{3.10}$$

which implies

$$\int_{\mathcal{X}} \left[\sum_{I\in\mathcal{L}} \sum_{I'\in \mathrm{child}(I), I'\subsetneq J'} \beta_{II'} \cdot \mu(I) \left(\int_{I'} h d\nu\right) \cdot \mathbf{1}_{I}\right] f^2 d\nu \lesssim C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

Testing on $h = \mathbf{1}_{J'}$, we get exactly (3.9).

5. The case $2 < p < \infty$: from trilinear forms to shifted bilinear forms

In this section, we give an equivalent statement of Theorem 3.4 in terms of a shifted positive operator. Based on this, we will prove the sufficiency in the next section.

We start to understand Theorem 3.4 by two claims.

Claim 3.6.

$$\Pi(f, g, h) \le C||f||_{L^p(\mu)} ||g||_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}$$

is equivalent to

$$\Pi(f,f,h) \le C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

PROOF. Only need to see the later implies the former. Since by definition (3.6), we have

$$2\Pi(f,g,h) \le \Pi(f,f,h) + \Pi(g,g,h) \le C\left(||f||^2_{L^p(\mu)} + ||g||^2_{L^p(\mu)}\right)||h||_{L^{(\frac{p}{2})'}(\nu)}$$

Hence, by homogeniety, for every $t > 0$,

$$\Pi(f,g,h) = \Pi(tf, \frac{1}{t}g, h) \le C\left(t^2||f||^2_{L^p(\mu)} + \frac{1}{t^2}||g||^2_{L^p(\mu)}\right)||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

Taking $t^2 = ||g||_{L^p(\mu)}/||f||_{L^p(\mu)}$, we conclude that

$$\Pi(f,g,h) \le C||f||_{L^p(\mu)}||g||_{L^p(\mu)}||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

□

CLAIM 3.7. $\Pi(f,f,h) \le C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}$ holds, if and only if

(3.11)

$$\sum_{I\in\mathcal{L}} \sum_{I'\in \text{child}(I)} \beta_{II'} \cdot \mu(I) \left(\int_I f^2 d\mu\right)\left(\int_{I'} h d\nu\right) \le C||f^2||_{L^{(\frac{p}{2})}(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)} \text{ holds.}$$

PROOF. In the last section, we have deduced that

$$\Pi(f,f,h) \le C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}$$

implies (3.10), which is equivalent to (3.11). On the other hand, since $\mu(I)\cdot\int_I f^2 d\mu \ge \left(\int_I f d\mu\right)^2$, we know (3.11) implies

$$\Pi(f,f,h) \le C||f||^2_{L^p(\mu)} ||h||_{L^{(\frac{p}{2})'}(\nu)}.$$

□

Because of the two claims, if we supress notation $\alpha_{II'} = \beta_{II'} \cdot \mu(I)$ in (3.11) and instead of assuming $2 < p < \infty$ and considering $p/2$, we still let $1 < p < \infty$ and consider p. Theorem 3.4 can be restated into the following.

THEOREM 3.8 (Two weight estimates for shifted positive operator). *For every sequence of non-negative constants* $\left\{\alpha_{II'}\right\}_{I\in\mathcal{L},I'\in\text{child}(I)}$*, define the shifted positive operator*

$$T_\alpha(f,g)=\sum_{I\in\mathcal{L}}\sum_{I'\in\text{child}(I)}\alpha_{II'}\left(\int_I fd\mu\right)\left(\int_{I'}gd\nu\right). \tag{3.12}$$

$$T_\alpha(f,g)\le C||f||_{L^p(\mu)}||g||_{L^{p'}(\nu)} \tag{3.13}$$

holds if and only if

(i)

$$\int_J\left[\sum_{I\in\mathcal{L},I\subseteq J}\sum_{I'\in\text{child}(I)}\alpha_{II'}\cdot\mu(I)\cdot\mathbf{1}_{I'}\right]^p d\nu\le C_1^p\cdot\mu(J),\ \ J\in\mathcal{L}, \tag{3.14}$$

(ii)

$$\int_{J'}\left[\sum_{I\in\mathcal{L}}\sum_{I'\in\text{child}(I),I'\subsetneq J'}\alpha_{II'}\cdot\nu(I')\cdot\mathbf{1}_I\right]^{p'} d\mu\le C_2^{p'}\cdot\nu(J'),\ \ J\in\mathcal{L},J'\in\text{child}(J). \tag{3.15}$$

In particular, $C\approx C_1+C_2$.

6. The case $2<p<\infty$: sufficiency

This section is dedicated to prove Theorem 3.8 and hence the sufficiency of Theorem 1.8 for the case $2<p<\infty$. The idea of the proof is from [7] with some new twists. It suffices to consider only for $f\ge 0$ and $g\ge 0$.

We split the estimate into two parts according to the following splitting condition: $\mathcal{L}=\mathcal{A}\cup\mathcal{B}$, where

$$\mathcal{A}=\left\{I\in\mathcal{L}:\langle f\rangle_{I,\mu}^p\cdot\mu(I)\ge\langle g\rangle_{I,\nu}^{p'}\cdot\nu(I)\right\}\text{ and }\mathcal{B}=\mathcal{L}\setminus\mathcal{A}. \tag{3.16}$$

Standard approximation reasoning allows us to assume that only finitely many terms $\alpha_{II'}$ are non-zero, so all the sums are finite.

For an interval $I \in \mathcal{L}$, let $\widehat{I}$ denote its parents. Using the splitting condition (3.16), we can write $T_\alpha(f,g) = T_1 + T_2$, where

$$T_1 = \sum_{I \in \mathcal{A}} \alpha_{\widehat{I}I} \left(\int_{\widehat{I}} f d\mu \right) \left(\int_I g d\nu \right), \tag{3.17}$$

$$T_2 = \sum_{I \in \mathcal{B}} \alpha_{\widehat{I}I} \left(\int_{\widehat{I}} f d\mu \right) \left(\int_I g d\nu \right). \tag{3.18}$$

6.1. A modified stopping interval construction. To estimate T_1 we need to modify a bit the construction of stopping intervals from Section 1. The main feature of the construction is that the stopping intervals well be the intervals $I \in \mathcal{A}$, but the stopping criterion will be checked on their parents $\widehat{I}$.

We start with some interval J (not necessarily in $\mathcal{A}$). For the interval J we define the primary *preliminary stopping intervals* to be the maximal by inclusion intervals $\widehat{I} \subseteq J$, $I \in \mathcal{A}$, such that

$$\langle f \rangle_{\widehat{I},\mu} > 2\langle f \rangle_{J,\mu}. \tag{3.19}$$

Note that different $I \in \mathcal{A}$ can give the same $\widehat{I}$, but this $\widehat{I}$ is counted only once.

It is obvious that these primary preliminary stopping intervals are disjoint and their total μ-measure is at most $\mu(J)/2$.

For each such preliminary stopping interval pick all its children L that belong to $\mathcal{A}$ (there is at least one such L), and declare these children to be the stopping intervals.

For the children $K \notin \mathcal{A}$ we continue the process: we will find the maximal by inclusion intervals $\widehat{I} \subseteq K$, $I \in \mathcal{A}$ satisfying (3.19), and declare these $\widehat{I}$ to be the secondary preliminary stopping intervals (note that in the stopping criterion (3.19) we still compare with the average over the original interval J).

For these secondary preliminary stopping intervals we add their children $L \in \mathcal{A}$ to the stopping intervals, and for the children $K \notin \mathcal{A}$ we continue the precess (again, still comparing the averages with the average over the original interval J).

We assumed that the collection $\mathcal{A}$ is finite, so at some point the process will stop (no $I \in \mathcal{A}$, $\widehat{I} \subseteq K$). We end up with the disjoint collection $\mathcal{G}^*(J)$ of stopping intervals.

Since all the stopping intervals are inside the primary *preliminary stopping intervals*, we can conclude that

$$\sum_{I\in\mathcal{G}^*(J)}\mu(I)<\frac{1}{2}\mu(J). \tag{3.20}$$

Let $G(J):=\bigcup_{I\in\mathcal{G}^*(J)}I$. Define $\mathcal{E}(J)=\mathcal{A}(J)\setminus\bigcup_{I\in\mathcal{G}^*(J)}\mathcal{A}(I)$, where $\mathcal{A}(J)=\{I\in\mathcal{A}:I\subseteq J\}$. It easily follows from the construction that for any $I\in\mathcal{E}(J)$

$$\langle f\rangle_{I,\mu}\le 2\langle f\rangle_{J,\mu}. \tag{3.21}$$

To construct a collection $\mathcal{G}$, we start with $\mathcal{G}_0$ of disjoint intervals covering the set $\bigcup_{I\in\mathcal{A}}\widehat{I}$. For each $J\in\mathcal{G}_0$ we run the stopping intervals construction to get the collection $\mathcal{G}^*(J)$. The union $\bigcup_{J\in\mathcal{G}_0}\mathcal{G}^*(J)$ give us the first generation of stopping intervals $\mathcal{G}_1^*$. Define inductively $\mathcal{G}_{k+1}^*=\bigcup_{J\in\mathcal{G}_k^*}\mathcal{G}^*(J)$ and put $\mathcal{G}=\bigcup_{k\ge1}\mathcal{G}_k^*$.

Note that the condition (3.20) implies that the collection $\mathcal{G}$ satisfies the following *Carleson measure condition*

$$\sum_{I\in\mathcal{G},I\subseteq J}\mu(I)<2\mu(J),\ \ J\in\mathcal{L}. \tag{3.22}$$

we also can replace $\mathcal{G}$ by $\mathcal{G}\cup\mathcal{G}_0$ here, and still have the same estimate.

6.2. Estimation of T_1. We start with the estimation of T_1. Using the modified stopping intervals constructed in the previous subsection and remember that $J\in\mathcal{G}_0$ is chosen such that $J\notin\mathcal{A}$, we obtain

$$\begin{aligned}\sum_{I\in\mathcal{A}}\alpha_{\widehat{I}I}\left(\int_{\widehat{I}}fd\mu\right)\left(\int_I gd\nu\right)&=\sum_{J\in\mathcal{G}\cup\mathcal{G}_0}\sum_{I\in\mathcal{E}(J)}\alpha_{\widehat{I}I}\left(\int_{\widehat{I}}fd\mu\right)\left(\int_I gd\nu\right)\\&=\text{Ⓐ}+\text{Ⓑ}\end{aligned}$$

$$\text{Ⓐ}=\sum_{J\in\mathcal{G}\cup\mathcal{G}_0}\sum_{I\in\mathcal{E}(J),I\neq J}\alpha_{\widehat{I}I}\left(\int_{\widehat{I}}fd\mu\right)\left(\int_I gd\nu\right)$$

$$\text{Ⓑ}=\sum_{I\in\mathcal{G}}\alpha_{\widehat{I}I}\left(\int_{\widehat{I}}fd\mu\right)\left(\int_I gd\nu\right)$$

For piece Ⓐ, by (3.21), we have

$$\begin{aligned}
Ⓐ &\leq \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \left(\int_I g d\nu\right) \\
&= \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \int_J \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I\right] g d\nu \\
&= ① + ②
\end{aligned}$$

$$① = \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \int_{J\setminus G(J)} \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I\right] g d\nu,$$

$$② = \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \int_{J\cap G(J)} \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I\right] g d\nu.$$

To estimate ①, since the sets $J\setminus G(J)$ are pairwisely disjoint,

$$\begin{aligned}
① &\leq \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \left[\int_J \left|\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I\right|^p d\nu\right]^{\frac{1}{p}} \left[\int_{J\setminus G(J)} |g|^{p'} d\nu\right]^{\frac{1}{p'}}, \text{ (3.14)} \\
&\leq \sum_{J\in\mathcal{G}\cup\mathcal{G}_0} 2\langle f\rangle_{J,\mu} \cdot C_1 \cdot \mu(J)^{\frac{1}{p}} \cdot \left[\int_{J\setminus G(J)} |g|^{p'} d\nu\right]^{\frac{1}{p'}}, \text{ Hölder's inequality} \\
&\leq C_1 \cdot 2 \left[\sum_{J\in\mathcal{G}\cup\mathcal{G}_0} \langle f\rangle_{J,\mu}^p \cdot \mu(J)\right]^{\frac{1}{p}} \left[\sum_{J\in\mathcal{G}\cup\mathcal{G}_0} \int_{J\setminus G(J)} |g|^{p'} d\nu\right]^{\frac{1}{p'}}, \text{(3.4) and disjointness} \\
&\leq C_1 \cdot 2^{1+\frac{1}{p}} \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.
\end{aligned}$$

To estimate ②, since the sets from $\mathcal{G}^*(J)$ are pairwisely disjoint and $G(J) = \cup_{K\in\mathcal{G}^*(J)} K$,

$$② = \sum_{J\in\mathcal{G}} 2\langle f\rangle_{J,\mu} \sum_{K\in\mathcal{G}^*(J)} \int_K \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I\right] g d\nu.$$

Note that the integrand is constant on every $K \in \mathcal{G}^*(J)$. Hence, we obtain

$$
\begin{aligned}
② &= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} 2\langle f \rangle_{J,\mu} \int_J \left[\sum_{I \in \mathcal{E}(J), I \neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I \right] \left[\sum_{K \in \mathcal{G}^*(J)} \langle g \rangle_{K,\nu} \cdot \mathbf{1}_K \right] d\nu \\
&\leq \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} 2\langle f \rangle_{J,\mu} \left[\int_J \left| \sum_{I \in \mathcal{E}(J), I \neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I \right|^p d\nu \right]^{\frac{1}{p}} \left[\int_J \left| \sum_{K \in \mathcal{G}^*(J)} \langle g \rangle_{K,\nu} \cdot \mathbf{1}_K \right|^{p'} d\nu \right]^{\frac{1}{p'}} \\
&= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} 2\langle f \rangle_{J,\mu} \left[\int_J \left| \sum_{I \in \mathcal{E}(J), I \neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I \right|^p d\nu \right]^{\frac{1}{p}} \left[\sum_{K \in \mathcal{G}^*(J)} \langle g \rangle_{K,\nu}^{p'} \cdot \nu(K) \right]^{\frac{1}{p'}}.
\end{aligned}
$$

Using the splitting condition (3.16) for the definition $\mathcal{A}$, we can estimate

$$
\begin{aligned}
② &\leq \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} 2\langle f \rangle_{J,\mu} \left[\int_J \left| \sum_{I \in \mathcal{E}(J), I \neq J} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \mathbf{1}_I \right|^p d\nu \right]^{\frac{1}{p}} \left[\sum_{K \in \mathcal{G}^*(J)} \langle f \rangle_{K,\mu}^{p} \cdot \mu(K) \right]^{\frac{1}{p'}} \\
&\leq \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} 2\langle f \rangle_{J,\mu} \cdot C_1 \cdot \mu(J)^{\frac{1}{p}} \cdot \left[\sum_{K \in \mathcal{G}^*(J)} \langle f \rangle_{K,\mu}^{p} \cdot \mu(K) \right]^{\frac{1}{p'}} \\
&\leq C_1 \cdot 2 \left[\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J,\mu}^p \cdot \mu(J) \right]^{\frac{1}{p}} \left[\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \sum_{K \in \mathcal{G}^*(J)} \langle f \rangle_{K,\mu}^{p} \cdot \mu(K) \right]^{\frac{1}{p'}}, \ (3.4) \\
&\leq C_1 \cdot 4(p')^p \cdot ||f||^p_{L^p(\mu)}.
\end{aligned}
$$

Combine the estimation of ① and ②. We conclude

$$
Ⓐ \leq C_1 \cdot 2^{1+\frac{1}{p}} \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_1 \cdot 4(p')^p \cdot ||f||^p_{L^p(\mu)}.
$$

For piece Ⓑ, note that

$$\text{Ⓑ} = \sum_{I\in\mathcal{G}} \alpha_{\widehat{I}I} \cdot \mu(\widehat{I}) \cdot \nu(I) \cdot \langle f\rangle_{\widehat{I},\mu} \cdot \langle g\rangle_{I,\nu}$$

$$\leq \left[\sum_{I\in\mathcal{G}} \alpha^p_{\widehat{I}I} \cdot \mu(\widehat{I})^p \cdot \nu(I) \cdot \langle f\rangle^p_{\widehat{I},\mu}\right]^{\frac{1}{p}} \left[\sum_{I\in\mathcal{G}} \langle g\rangle^{p'}_{I,\nu} \cdot \nu(I)\right]^{\frac{1}{p'}}, \ (3.16)$$

$$\leq \left[\sum_{I\in\mathcal{G}} \alpha^p_{\widehat{I}I} \cdot \mu(\widehat{I})^p \cdot \nu(I) \cdot \langle f\rangle^p_{\widehat{I},\mu}\right]^{\frac{1}{p}} \left[\sum_{I\in\mathcal{G}} \langle f\rangle^{p}_{I,\nu} \cdot \mu(I)\right]^{\frac{1}{p'}}, \ (3.4)$$

$$\leq 2^{\frac{1}{p'}} \cdot p \cdot ||g||_{L^{p'}(\nu)} \cdot \left[\sum_{I\in\mathcal{G}} \alpha^p_{\widehat{I}I} \cdot \mu(\widehat{I})^p \cdot \nu(I) \cdot \langle f\rangle^p_{\widehat{I},\mu}\right]^{\frac{1}{p}}.$$

To finish, we need the following lemma.

Lemma 3.9. *The sequence* $\{\alpha_I\}_{I\in\mathcal{L}}$,

$$\alpha_I = \sum_{I'\in\text{child}(I)} \alpha^p_{II'} \cdot \mu(I)^p \cdot \nu(I')$$

satisfies the Carleson measure condition.

Proof. For $J \in \mathcal{L}$, we have

$$\sum_{I\in\mathcal{L}:I\subseteq J} \alpha_I = \left|\left|\left[\sum_{I\subseteq J}\sum_{I'\in\mathcal{G},I'\in\text{child}(I)} \alpha^p_{II'} \cdot \mu(I)^p \cdot \mathbf{1}_{I'}\right]^{\frac{1}{p}}\right|\right|^p_{L^p(\nu)}, \ ||\cdot||_{l^p} \leq ||\cdot||_{l^1}$$

$$\leq \left|\left|\sum_{I\subseteq J}\sum_{I'\in\mathcal{G},I'\in\text{child}(I)} \alpha_{II'} \cdot \mu(I) \cdot \mathbf{1}_{I'}\right|\right|^p_{L^p(\nu)}$$

$$= \int_J \left|\sum_{I\subseteq J}\sum_{I'\in\mathcal{G},I'\in\text{child}(I)} \alpha_{II'} \cdot \mu(I) \cdot \mathbf{1}_{I'}\right|^p d\nu, \ (3.14)$$

$$\leq C_1^p \cdot \mu(J).$$

□

Hence, we can estimate

$$\text{Ⓑ} \leq C_1 \cdot 2^{\frac{1}{p'}} \cdot p \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)},$$

which, together with the estimation of Ⓐ, imply that

$$T_1 \leq C_1 \cdot 2^{1+\frac{1}{p}} \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_1 \cdot 4(p')^p \cdot ||f||^p_{L^p(\mu)}$$
$$+ C_1 \cdot 2^{\frac{1}{p'}} \cdot p \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.$$

6.3. Estimation of T_2. Now we take care of the estimation of T_2. The estimation proceeds similar as in subsection 6.2. Using the stopping intervals constructed in section 1 with $\mathcal{F} = \mathcal{B}$, we obtain

$$\sum_{I\in\mathcal{B}} \alpha_{\widehat{I}I} \left(\int_{\widehat{I}} f d\mu\right)\left(\int_I g d\nu\right) = \sum_{J\in\mathcal{G}} \sum_{I\in\mathcal{E}(J)} \alpha_{\widehat{I}I} \left(\int_{\widehat{I}} f d\mu\right)\left(\int_I g d\nu\right)$$
$$= Ⓐ + Ⓑ$$

$$Ⓐ = \sum_{J\in\mathcal{G}} \sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \left(\int_{\widehat{I}} f d\mu\right)\left(\int_I g d\nu\right),$$

$$Ⓑ = \sum_{J\in\mathcal{G}} \alpha_{\widehat{J}J} \left(\int_{\widehat{J}} f d\mu\right)\left(\int_J g d\nu\right).$$

Note here $I \neq J$, we can write

$$Ⓐ \leq \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot \int_J \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I) \cdot \mathbf{1}_{\widehat{I}}\right] f d\mu = ① + ②$$

$$① = \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot \int_{J\setminus G(J)} \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I) \cdot \mathbf{1}_{\widehat{I}}\right] f d\mu,$$

$$② = \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot \int_{J\cap G(J)} \left[\sum_{I\in\mathcal{E}(J), I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I) \cdot \mathbf{1}_{\widehat{I}}\right] f d\mu.$$

To estimate ①, again since the sets $J \setminus G(J)$ are pairwisely disjoint,

$$①\leq \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \left[\int_J \left|\sum_{I\in\mathcal{E}(J),I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I)\cdot \mathbf{1}_{\widehat{I}}\right|^{p'} d\nu\right]^{\frac{1}{p'}} \left[\int_{J\setminus G(J)} |f|^p d\mu\right]^{\frac{1}{p}}, \ (3.15)$$

$$\leq C_2 \cdot 2 \left[\sum_{J\in\mathcal{G}} \langle g\rangle^{p'}_{J,\nu} \cdot \nu(J)\right]^{\frac{1}{p'}} \left[\sum_{J\in\mathcal{G}} \int_{J\setminus G(J)} |f|^p d\mu\right]^{\frac{1}{p}}, \ (3.4) \text{ and disjointness}$$

$$\leq C_2 \cdot 2^{1+\frac{1}{p'}} \cdot p \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.$$

To estimate ②, note again that the sets from $\mathcal{G}^*(J)$ are pairwisely disjoint and $G(J) = \cup_{K\in\mathcal{G}^*(J)} K$. Hence,

$$②= \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \sum_{K\in\mathcal{G}^*(J)} \int_K \left[\sum_{I\in\mathcal{E}(J),I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I) \cdot \mathbf{1}_{\widehat{I}}\right] f d\mu.$$

Since the integrand is constant on K, we have

$$②= \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot \int_J \left[\sum_{I\in\mathcal{E}(J),I\neq J} \alpha_{\widehat{I}I} \cdot \nu(I) \cdot \mathbf{1}_{\widehat{I}}\right] \left[\sum_{K\in\mathcal{G}^*(J)} \langle f\rangle_{K,\mu} \cdot \mathbf{1}_K\right] d\mu$$

$$\leq \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot C_2 \cdot \nu(J)^{\frac{1}{p'}} \cdot \left[\int_J \left|\sum_{K\in\mathcal{G}^*(J)} \langle f\rangle_{K,\mu} \cdot \mathbf{1}_K\right|^p d\mu\right]^{\frac{1}{p}}, \ \text{disjointness}$$

$$= \sum_{J\in\mathcal{G}} 2\langle g\rangle_{J,\nu} \cdot C_2 \cdot \nu(J)^{\frac{1}{p'}} \cdot \left[\sum_{K\in\mathcal{G}^*(J)} \langle f\rangle^p_{K,\mu} \cdot \nu(K)\right]^{\frac{1}{p}}, \ \text{splitting condition } (3.16)$$

$$\leq C_2 \cdot 2 \left[\sum_{J\in\mathcal{G}} \langle g\rangle^{p'}_{J,\nu} \cdot \nu(J)\right]^{\frac{1}{p'}} \left[\sum_{J\in\mathcal{G}} \sum_{K\in\mathcal{G}^*(J)} \langle g\rangle^{p'}_{K,\nu} \cdot \nu(K)\right]^{\frac{1}{p}}, \ (3.4)$$

$$\leq C_2 \cdot 4(p)^{p'} \cdot ||g||^{p'}_{L^{p'}(\nu)}.$$

Combine the estimation of ① and ②, we have

$$Ⓐ \leq C_2 \cdot 2^{1+\frac{1}{p'}} \cdot p \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_2 \cdot 4(p)^{p'} \cdot ||g||^{p'}_{L^{p'}(\nu)}.$$

Finally, to estimate Ⓑ, note again

$$
\begin{aligned}
Ⓑ &= \sum_{J\in\mathcal{G}} \alpha_{\widehat{J}J} \cdot \mu(\widehat{J}) \cdot \nu(J) \cdot \langle f \rangle_{\widehat{J},\mu} \cdot \langle g \rangle_{J,\nu} \\
&\leq \left[\sum_{J\in\mathcal{G}} \alpha^p_{\widehat{J}J} \cdot \mu(\widehat{J})^p \cdot \nu(J) \cdot \langle f \rangle^p_{\widehat{J},\mu}\right]^{\frac{1}{p}} \left[\sum_{J\in\mathcal{G}} \langle g \rangle^{p'}_{J,\nu} \cdot \nu(J)\right]^{\frac{1}{p'}}, \ (3.4) \\
&\leq 2^{\frac{1}{p'}} \cdot p \cdot ||g||_{L^{p'}(\nu)} \cdot \left[\sum_{J\in\mathcal{G}} \alpha^p_{\widehat{J}J} \cdot \mu(\widehat{J})^p \cdot \nu(J) \cdot \langle f \rangle^p_{\widehat{J},\mu}\right]^{\frac{1}{p}}, \ \text{Lemma } 3.9 \\
&\leq C_1 \cdot 2^{\frac{1}{p'}} \cdot p \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.
\end{aligned}
$$

Hence, we deduce that

$$
\begin{aligned}
T_2 \leq\ & C_2 \cdot 2^{1+\frac{1}{p'}} \cdot p \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_2 \cdot 4(p)^{p'} \cdot ||g||^{p'}_{L^{p'}(\nu)} \\
& + C_1 \cdot 2^{\frac{1}{p'}} \cdot p \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.
\end{aligned}
$$

Eventually, we conclude that

$$
\begin{aligned}
T_\alpha(f,g) &= T_1 + T_2 \\
&\leq C_1 \cdot 2^{1+\frac{1}{p}} \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_1 \cdot 4(p')^p \cdot ||f||^p_{L^p(\mu)} \\
&+ C_2 \cdot 2^{1+\frac{1}{p'}} \cdot p \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + C_2 \cdot 4(p)^{p'} \cdot ||g||^{p'}_{L^{p'}(\nu)} \\
&+ C_1 \cdot 2^{1+\frac{1}{p'}} \cdot p \cdot p' \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.
\end{aligned}
$$

By homogeniety, for every $t > 0$,

$$
T_\alpha(f,g) = T_\alpha(tf, \frac{1}{t}g) \lesssim (C_1 + C_2) \cdot \left(t^p ||f||^p_{L^p(\mu)} + ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)} + \frac{1}{t^{p'}} ||g||^{p'}_{L^{p'}(\nu)} \right).
$$

Taking $t = ||g||^{\frac{1}{p}}_{L^{p'}(\nu)} / ||f||^{\frac{1}{p'}}_{L^p(\mu)}$, we prove exactly

$$
T_\alpha(f,g) \lesssim (C_1 + C_2) \cdot ||f||_{L^p(\mu)} ||g||_{L^{p'}(\nu)}.
$$

CHAPTER 4

Bellman functions on filtered probability spaces I: Burkhölder's hull and Super-solutions

In this chapter, we find explicitly a super-solution of the dyadic Carleson Embedding Theorem 1.9 via the Burkhölder's hull. We start with some properties of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$, in particular, we prove the main inequality. Then, we define and discuss the super-solutions in detail. In the last section, we introduce the Burkhölder's hull and solve for a super-solution of Theorem 1.9 via the Burkhölder's hull. This chapter proves the existence of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$.

1. Properties of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$

PROPOSITION 4.1 (Properties of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$).

(i) Domain: $\mathbf{f}^p \le F$ and $0 \le M \le C$.

(ii) Range: $0 \le \mathcal{B}(F, \mathbf{f}, M; C) \le C_p \cdot C \cdot F$.

(iii) The main inequality: For all triples $(F, \mathbf{f}, M)$ and $(F_\pm, \mathbf{f}_\pm, M_\pm)$ belong to the domain $\mathbf{f}^p \le F$, $0 \le M \le C$, with $F = \frac{1}{2}(F_+ + F_-)$, $\mathbf{f} = \frac{1}{2}(\mathbf{f}_+ + \mathbf{f}_-)$ and $M = \Delta M + \frac{1}{2}(M_+ + M_-)$, where $0 \le \Delta M \le M$, we have

$$\mathcal{B}(F, \mathbf{f}, M; C) \ge \frac{1}{2}\{\mathcal{B}(F_+, \mathbf{f}_+, M_+; C) + \mathcal{B}(F_-, \mathbf{f}_-, M_-; C)\} + \Delta M \cdot \mathbf{f}^p. \tag{4.1}$$

PROOF. (i) follows from the Hölder's inequality and that $\{\alpha_{_I}\}_{I\in\mathcal{D}}$ is a Carleson sequence. (ii) holds if we assume Theorem 1.9 is true. We explain (iii) in more detail.

Split the sum in the definition (1.15) of $\mathcal{B}(F, \mathbf{f}, M)$ into three pieces

$$|I|^{-1}\sum_{J\subseteq I}\alpha_{_J}\langle f\rangle_{_J}^p = \frac{1}{2}|I_+|^{-1}\sum_{J\subseteq I_+}\alpha_{_J}\langle f\rangle_{_J}^p + \frac{1}{2}|I_-|^{-1}\sum_{J\subseteq I_-}\alpha_{_J}\langle f\rangle_{_J}^p + |I|^{-1}\alpha_{_I}\langle f\rangle_{_I}^p,$$

where $I_\pm$ means the right and left halves of I, respectively.

Now, we choose $f^{\pm}$ on the interval $I_{\pm}$ that almost give the supremum in the definition (1.15) of $\mathcal{B}(F_{\pm}, \mathbf{f}_{\pm}, M_{\pm})$, i.e. for small $\varepsilon > 0$,

$$|I_{\pm}|^{-1} \sum_{J \subseteq I\pm} \alpha_{_J} \langle f^{\pm} \rangle_{_J}^p \geq \mathcal{B}(F_{\pm}, \mathbf{f}_{\pm}, M_{\pm}; C) - \frac{\varepsilon}{2},$$

and note that $|I|^{-1} \alpha_{_I} \langle f \rangle_{_I}^p = \Delta M \cdot \mathbf{f}$, we conclude

$$|I|^{-1} \sum_{J \subseteq I} \alpha_{_J} \langle f \rangle_{_J}^p \geq \frac{1}{2} \{ \mathcal{B}(F_+, \mathbf{f}_+, M_+; C) + \mathcal{B}(F_-, \mathbf{f}_-, M_-; C) \} - \varepsilon + \Delta M \cdot \mathbf{f}^p,$$

which yields exactly (4.1). □

REMARK 4.2. From (1.15), we know that the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$ exists and $0 \leq \mathcal{B}(F, \mathbf{f}, M; C) \leq C_p \cdot C \cdot F$ if and only if Theorem 1.9 is true. The sharpness is explained as

$$\sup_{\mathbf{f}^p \leq F,\ 0 \leq M \leq C} \frac{\mathcal{B}(F, \mathbf{f}, M; C)}{C \cdot F} = (p')^p. \tag{4.2}$$

2. Properties of the Super-solutions

2.1. The super-solutions and the dyadic Caleson Embedding Theorem.

DEFINITION 4.3. A function satisfies Propositon 4.1 is called a *super-solution*. We denote a super-solution by $\mathbb{B}(F, \mathbf{f}, M; C)$.

We have seen that the dyadic Carleson Embedding Theorem 1.9 gives rise to a super-solution $\mathbb{B}(F, \mathbf{f}, M; C)$. On the other hand, to prove (1.14) and actually Theorem 1.9, it suffices to find any super-solution.

Indeed, pick $f \geq 0$ and $\{\alpha_{_I}\}_{I \in \mathcal{D}}$ satisfying the Carleson condition. For every dyadic interval $I \in \mathcal{D}$, let $F_{_I}, \mathbf{f}_{_I}, M_{_I}$ be the corresponding averages

$$F_{_I} = \langle f^p \rangle_{_I}, \mathbf{f}_{_I} = \langle f \rangle_{_I}, M_{_I} = |I|^{-1} \sum_{J \subseteq I} \alpha_{_J}.$$

Note that $F_{_I} = \frac{1}{2}(F_{_{I_+}} + F_{_{I_-}})$, $\mathbf{f}_{_I} = \frac{1}{2}(\mathbf{f}_{_{I_+}} + \mathbf{f}_{_{I_-}})$ and $M_{_I} = \Delta M_{_I} + \frac{1}{2}(M_{_{I_+}} + M_{_{I_-}})$, where $0 \leq \Delta M_{_I} = |I|^{-1} \alpha_{_I} \leq M_{_I}$. For the interval I, the main inequality (4.1) implies

$$\alpha_{_I} \langle f \rangle_{_I}^p \leq |I| \mathbb{B}(F_{_I}, \mathbf{f}_{_I}, M_{_I}; C) - |I_+| \mathbb{B}(F_{_{I_+}}, \mathbf{f}_{_{I_+}}, M_{_{I_+}}; C) - |I_-| \mathbb{B}(F_{_{I_-}}, \mathbf{f}_{_{I_-}}, M_{_{I_-}}; C).$$

Going n levels down, we get the inequality

$$\sum_{J\subseteq I,|J|>2^{-n}|I|}\alpha_{_J}\langle f\rangle_{_J}^p \leq |I|\mathbb{B}(F_{_I},\mathbf{f}_{_I},M_{_I};C)-\sum_{J\subseteq I,|J|=2^{-n}|I|}|J|\mathbb{B}(F_{_J},\mathbf{f}_{_J},M_{_J};C)$$
$$\leq |I|\mathbb{B}(F_{_I},\mathbf{f}_{_I},M_{_I};C)\leq C_p\cdot C\cdot|I|F_{_I}=C_p\cdot C\cdot\int_I f^p.$$

Applying the above estimate for the intervals $[-2^n,0)$ and $[0,2^n)$ and taking the limit as $n\to\infty$, we prove exactly (1.14).

Remark 4.4. To prove Theorem 1.9, all amounts to finding a super-solution $\mathbb{B}(F,\mathbf{f},M;C)$. We will see in section 3 that the least possible constant for which $\mathbb{B}(F,\mathbf{f},M;C)$ exists is $C_p=(p')^p$.

2.2. Further properties of $\mathcal{B}(F,\mathbf{f},M;C)$. We start with the following celebrated theorem in convex analysis. We will give a proof for the sake of completeness, for more details, see [**29**].

Theorem 4.5. *Let $f:\Omega\to\mathbb{R}$ be a locally bounded function defined on some convex domain $\Omega\in\mathbb{R}^n$ and f satisfies the midpoint concavity: $f(\frac{x+y}{2})\geq\frac{f(x)+f(y)}{2}$ for all $x,y\in\Omega$. Then f is concave and locally Lipschitz.*

Proof. For concavity: If f is not concave, then there exist two points $a,b\in\Omega$, as well as the line segment connecting them $[a,b]=\{\lambda a+(1-\lambda)b:0\leq\lambda\leq1\}\subseteq\Omega$, such that the function $\varphi(\lambda)=f(\lambda a+(1-\lambda)b)-\lambda f(a)-(1-\lambda)f(b)$ verifies

$$-\infty<C=\inf\{\varphi(\lambda):0\leq\lambda\leq1\}<0.$$

Note that we have used Ω being convex and f being locally bounded here. Furthermore, $\varphi(0)=\varphi(1)=0$ and a direct computation shows that φ is also midpoint concave. Take $0<\delta<-\frac{C}{2}$ and let $0\leq\lambda_0\leq1$, such that $\varphi(\lambda_0)\leq C+\delta$, without loss of generality, further assuming $0<\lambda_0<\frac{1}{2}$, hence we have $\varphi(0)=0$ and $\varphi(2\lambda_0)\geq C$, however

$$\varphi(\lambda_0)\leq C+\delta<\frac{C}{2}=\frac{\varphi(0)+\varphi(2\lambda_0)}{2},\ \text{ a contradiction!}$$

For locally Lipschitz continuity: Given $a \in \Omega$, we can find a ball $B(a, 2r) \subseteq \Omega$ on which f is bounded by a constant M. For $x \neq y$ in $B(a, 2r)$, put $z = y + (\frac{r}{\alpha})(y - x)$, where $\alpha = ||y - x||$. Clearly, $z \in B(a, 2r)$. Moreover, since $y = \frac{r}{r+\alpha}x + \frac{\alpha}{r+\alpha}z$, from the concavity of f we infer that $f(y) \geq \frac{r}{r+\alpha}f(x) + \frac{\alpha}{r+\alpha}f(z)$. So $|f(y) - f(x)| \leq \frac{\alpha}{r+\alpha}|f(z) - f(x)| \leq \frac{||y-x||}{r} \cdot 2M$. □

In the case of our main inequality (4.1), first put $F = \frac{1}{2}(F_+ + F_-)$, $\mathbf{f} = \frac{1}{2}(\mathbf{f}_+ + \mathbf{f}_-)$ and $M = \frac{1}{2}(M_+ + M_-)$ (i.e. $\Delta M = 0$) and assume all triples $(F, \mathbf{f}, M)$, $(F_\pm, \mathbf{f}_\pm, M_\pm)$ are in the convex domain: $\mathbf{f}^p \leq F, 0 \leq M \leq C$, then we obtain the midpoint concavity of $\mathcal{B}(F, \mathbf{f}, M; C)$. Apply Theorem 4.5 to the function $\mathcal{B}$, so $\mathcal{B}$ is itself concave and locally Lipschitz. In particular, $\mathcal{B}$ is a continuous function.

Now let $0 \leq \lambda \leq 1$ and $F = \lambda F_+(1-\lambda)F_-$, $\mathbf{f} = \lambda \mathbf{f}_+(1-\lambda)\mathbf{f}_-$, $M = \Delta M + \lambda M_+ + (1-\lambda)M_-$. The main inequality (4.1) and concavity of $\mathcal{B}$ imply that

$$\begin{aligned}\Delta M \cdot \mathbf{f}^p &\leq \mathcal{B}(F, \mathbf{f}, M; C) - \mathcal{B}(F, \mathbf{f}, M - \Delta M; C) \\ &\leq \mathcal{B}(F, \mathbf{f}, M; C) - \{\lambda \mathcal{B}(F_+, \mathbf{f}_+, M_+; C) + (1-\lambda)\mathcal{B}(F_-, \mathbf{f}_-, M_-; C)\}.\end{aligned}$$

Hence, the Bellman function $\mathcal{B}(F, \mathbf{f}, M)$ is continuous and satisfies

$$\mathcal{B}(F, \mathbf{f}, M; C) \geq \lambda \mathcal{B}(F_+, \mathbf{f}_+, M_+; C) + (1-\lambda)\mathcal{B}(F_-, \mathbf{f}_-, M_-; C) + \Delta M \cdot \mathbf{f}^p. \tag{4.3}$$

2.3. Regularization of the super-solutions. As we have seen, the Bellman function $\mathcal{B}$ is concave and locally Lipschitz, and thus continuous, but hardly any better than that. Fortunately, we know that the proof of Theorem 1.9 boils down to finding just a super-solution $\mathbb{B}$. We recall the trick of regularization of the super-solutions from [25].

Given a super-solution $\mathbb{B}(F, \mathbf{f}, M; C)$ satisfying Proposition 4.1. Let $\phi_\varepsilon, \psi_\varepsilon : (0, \infty) \to [0, \infty)$ be any two nonnegative C^∞ functions, such that $\operatorname{supp}(\phi_\varepsilon) \subseteq [1, (1+\varepsilon)^p]$, $\operatorname{supp}(\psi_\varepsilon) \subseteq [1+\varepsilon, 1+2\varepsilon]$ and $\int_0^\infty \phi_\varepsilon(t)\frac{dt}{t} = \int_0^\infty \psi_\varepsilon(t)\frac{dt}{t} = 1$. Define

$$\begin{aligned}\mathbb{B}_\varepsilon(F, \mathbf{f}, M; C) &= \iiint_{(0,\infty)^3} \mathbb{B}\left(\frac{F}{u}, \frac{\mathbf{f}}{v}, \frac{M}{w}; C\right) \phi_\varepsilon(u)\psi_\varepsilon(v)\phi_\varepsilon(w)\frac{du dv dw}{uvw} \\ &= \iiint_{(0,\infty)^3} \mathbb{B}(u, v, w; C)\phi_\varepsilon\left(\frac{F}{u}\right)\psi_\varepsilon\left(\frac{\mathbf{f}}{v}\right)\phi_\varepsilon\left(\frac{M}{w}\right)\frac{du dv dw}{uvw}\end{aligned}$$

Note that the second representation shows $\mathbb{B}_\varepsilon \in C^\infty$. Since $\mathbb{B}$ is continuous, the family of smooth functions $\{\mathbb{B}_\varepsilon : \varepsilon > 0\}$ converges to $\mathbb{B}$ pointwisely as $\varepsilon \to 0$.

To check Proposition 4.1 for $\mathbb{B}_\varepsilon$. Note that the supports of ϕ_ε and ψ_ε guarantee that $\mathbb{B}_\varepsilon$ is well-defined in the region $\{\mathbf{f}^p \le F,\ 0 \le M \le C\}$ and an easy calculation shows that $0 \le \mathbb{B}_\varepsilon \le C_p \cdot C \cdot F$. For the main inequality, the first representation and (4.1) imply that

$$\begin{aligned}\mathbb{B}_\varepsilon(F,\mathbf{f},M;C) &- \frac{1}{2}\left\{\mathbb{B}_\varepsilon(F_+,\mathbf{f}_+,M_+;C)+\mathbb{B}_\varepsilon(F_-,\mathbf{f}_-,M_-;C)\right\}\\ &\ge \Delta M\cdot \mathbf{f}^p \int_1^{(1+\varepsilon)^p}\int_{1+\varepsilon}^{1+2\varepsilon}\int_1^{(1+\varepsilon)^p}\frac{1}{v^p w}\phi_\varepsilon(u)\psi_\varepsilon(v)\phi_\varepsilon(w)\frac{dudvdw}{uvw}\\ &\ge \frac{1}{(1+2\varepsilon)^p(1+\varepsilon)^p}\Delta M\cdot \mathbf{f}^p \to \Delta M \cdot \mathbf{f}^p \text{ as } \varepsilon\to 0.\end{aligned}$$

Hence, the proof of (1.14) given in subsection 2.1 works for the smooth function $\mathbb{B}_\varepsilon(F,\mathbf{f},M;C)$ as well. In what follows, it suffices to consider only for smooth supersolutions $\mathbb{B}(F,\mathbf{f},M;C)$.

2.4. The main inequality in its infinitesimal version. For a smooth supersolution $\mathbb{B}(F,\mathbf{f},M;C)$, being concave means the second differential $d^2\mathbb{B}\le 0$. By the main inequality (4.1), we have: $\mathbb{B}(F,\mathbf{f},M)-\mathbb{B}(F,\mathbf{f},M-\Delta M)\ge \Delta M\cdot\mathbf{f}^p$, and thus $\frac{\partial \mathbb{B}}{\partial M}\ge \mathbf{f}^p$.

Therefore, the main inequality (4.1) implies the following two infinitesimal ones

$$d^2\mathbb{B}(F,\mathbf{f},M;C)\le 0 \quad\text{and}\quad \frac{\partial\mathbb{B}}{\partial M}(F,\mathbf{f},M;C)\ge \mathbf{f}^p. \tag{4.4}$$

Actually, (4.4) is equivalent to the main inequality (4.1). Since by (4.4), we can deduce

$$\begin{aligned}\Delta M\cdot\mathbf{f}^p &\le \mathbb{B}(F,\mathbf{f},M;C)-\mathbb{B}(F,\mathbf{f},M-\Delta M;C)\\ &\le \mathbb{B}(F,\mathbf{f},M;C)-\frac{1}{2}\left\{\mathbb{B}(F_+,\mathbf{f}_+,M_+;C)+\mathbb{B}(F_-,\mathbf{f}_-,M_-;C)\right\}.\end{aligned}$$

3. Finding a super-solution via the Burkhölder's hull

3.1. Burkhölder's hull and some reductions. Assume $\mathbb{B}(F, \mathbf{f}, M; C)$ is a smooth super-solution. In this section, we present an explicit function $\mathbb{B}(F, \mathbf{f}, M; C)$ with the help of the Burkhölder's hull.

DEFINITION 4.6. The *Burkhölder's hull* of $\mathbb{B}(F, \mathbf{f}, M; C)$ is defined by

$$u(\mathbf{f}, M; C) = \sup_F \{\mathbb{B}(F, \mathbf{f}, M; C) - C_p \cdot C \cdot F\}, \ \mathbf{f} \geq 0, \ 0 \leq M \leq C. \tag{4.5}$$

REMARK 4.7. This trick of eliminating one variable is due to D. Burkhölder [27].

It follows from the definition (1.15) of $\mathcal{B}(F, \mathbf{f}, M; C)$ that

$$\mathcal{B}(F, \mathbf{f}, M; C) = C \cdot \mathcal{B}(F, \mathbf{f}, M/C; 1). \qquad \text{Scaling Property} \tag{4.6}$$

Thus, it suffices to consider only for $C = 1$. We adopt the notations $\mathcal{B}(F, \mathbf{f}, M) = \mathcal{B}(F, \mathbf{f}, M; C = 1)$, $\mathbb{B}(F, \mathbf{f}, M) = \mathbb{B}(F, \mathbf{f}, M; C = 1)$ and $u(\mathbf{f}, M) = u(\mathbf{f}, M; C = 1)$.

PROPOSITION 4.8. *The Bukhölder's hull $u(\mathbf{f}, M)$ satisfies the following properties*

$$\text{(i) } \frac{\partial u}{\partial M}(\mathbf{f}, M) \geq \mathbf{f}^p \quad \text{and} \quad \text{(ii) } u(\mathbf{f}, M) \text{ is concave.}$$

PROOF. The proof follows from the definition (4.5).

(i) From $\frac{\partial \mathbb{B}}{\partial M}(F, \mathbf{f}, M) \geq \mathbf{f}^p$, we conclude that $\mathbb{B}(F, \mathbf{f}, M + \Delta M) - \mathbb{B}(F, \mathbf{f}, M) \geq \Delta M \cdot \mathbf{f}^p$. Choose F_0 that almost gives the supremum in the definition of $u(\mathbf{f}, M)$, i.e. for small $\varepsilon > 0$, $\mathbb{B}(F_0, \mathbf{f}, M) - C_p \cdot F_0 > u(\mathbf{f}, M) - \varepsilon$, then

$$\begin{aligned}
& u(\mathbf{f}, M + \Delta M) - u(\mathbf{f}, M) \\
\geq \ & [\mathbb{B}(F_0, \mathbf{f}, M + \Delta M) - C_p \cdot F_0] - [\mathbb{B}(F_0, \mathbf{f}, M) - C_p \cdot F_0 + \varepsilon] \\
= & [\mathbb{B}(F_0, \mathbf{f}, M + \Delta M) - \mathbb{B}(F_0, \mathbf{f}, M)] - \varepsilon \\
\geq & \Delta M \cdot \mathbf{f}^p - \varepsilon.
\end{aligned}$$

Letting $\varepsilon \to 0$, so $\frac{\partial u}{\partial M}(\mathbf{f}, M) \geq \mathbf{f}^p$.

(ii) We need the following simple lemma.

LEMMA 4.9. *Let $\varphi(x, y)$ be a convex function and let $\Phi(x) = \sup_y \varphi(x, y)$, then $\Phi(x)$ is also a concave function.*

PROOF. We need to see $\Phi(\lambda x_1 + (1-\lambda)x_2) \geq \lambda\Phi(x_1) + (1-\lambda)\Phi(x_2)$ for all x_1, x_2 and $0 \leq \lambda \leq 1$. Again choose y_1 and y_2 in the definition of $\Phi(x)$, such that for small $\varepsilon > 0$, $\varphi(x_1, y_1) > \Phi(x_1) - \varepsilon$ and $\varphi(x_2, y_2) > \Phi(x_2) - \varepsilon$. Then

$$\begin{aligned}\lambda\Phi(x_1) + (1-\lambda)\Phi(x_2) &< \lambda\varphi(x_1, y_1) + (1-\lambda)\varphi(x_2, y_2) + \varepsilon \\ &\leq \varphi(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) + \varepsilon \\ &\leq \Phi((\lambda x_1 + (1-\lambda)x_2) + \varepsilon,\end{aligned}$$

which proves the lemma. □

A direct application of this lemma gives (ii).

□

REMARK 4.10. From Proposition 4.8 and (4.5), if the dyadic Carleson Embedding Theorem 1.9 holds with constant C_p, then there exists a concave function $u(\mathbf{f}, M)$ satisfying $\frac{\partial u}{\partial M}(\mathbf{f}, M) \geq \mathbf{f}^p$ and $-C_p \cdot \mathbf{f}^p \leq u(\mathbf{f}, M) \leq 0$. On the other hand, if such a $u(\mathbf{f}, M)$ exists, then we can define $\mathbb{B}(F, \mathbf{f}, M) = u(\mathbf{f}, M) + C_p \cdot F$ for $F \geq \mathbf{f}^p, 0 \leq M \leq 1$, and so $\mathbb{B}$ is a super-solution that proves the dyadic Carleson Embedding Theorem with the same constant C_p. Hence, the best constant in the dyadic Carleson Embedding Theorem is exactly the best constant for which the fuction $u(\mathbf{f}, M)$ exists.

Now, note the definition (1.15) of $\mathcal{B}(F, \mathbf{f}, M; C)$ implies that

(4.7) $$\mathcal{B}(t^p F, t\mathbf{f}, M; C) = t^p \cdot \mathcal{B}(F, \mathbf{f}, M; C) \text{ for all } t \geq 0. \qquad \text{Homogeniety}$$

Hence, $u(t\mathbf{f}, M) = t^p \cdot u(\mathbf{f}, M)$, which means $u(\mathbf{f}, M)$ can be represented as $u(\mathbf{f}, M) = \mathbf{f}^p \cdot \varphi(M)$. For such a function $u(\mathbf{f}, M)$, the Hessian equals

$$\begin{pmatrix} p(p-1)\mathbf{f}^{p-2}\varphi(M) & p\mathbf{f}^{p-1}\varphi'(M) \\ p\mathbf{f}^{p-1}\varphi'(M) & \mathbf{f}^p\varphi''(M) \end{pmatrix},$$

so the concavity of $u(\mathbf{f}, M)$ is equivalent to the following two inequalities

$$\varphi(M) \leq 0 \text{ and } \varphi\varphi'' - (p')(\varphi')^2 \geq 0 \text{ for } 0 \leq M \leq 1.$$

The inequality $\frac{\partial u}{\partial M}(\mathbf{f}, M) \geq \mathbf{f}^p$ means $\varphi'(M) \geq 1$ and $\varphi(M)$ also satisfies $-C_p \leq \varphi(M) \leq 0$.

Hence, our task is to find $\varphi(M)$, such that

(i) $0 \leq M \leq 1$

(ii) $-C_p \leq \varphi(M) \leq 0$

(iii) $\varphi'(M) \geq 1$

(iv) $\varphi\varphi'' - (p')(\varphi')^2 \geq 0$,

and the least possible constant is $C_p = \inf_\varphi \sup_{0 \leq M \leq 1} \{-\varphi(M)\}$.

3.2. The formula of the Burkhölder's hull and an explicit super-solution. We first introduce $\phi(M) = -\varphi(M) \geq 0$, then $\phi(M)$ satisfies

$$\text{(i) } 0 \leq M \leq 1; \text{ (ii) } 0 \leq \phi(M) \leq C_p; \text{ (iii) } \phi'(M) \leq -1; \text{ (iv) } \phi\phi'' - (p')(\phi)^2 \geq 0,$$

and we need to consider $C_p = \inf_\phi \sup_{0 \leq M \leq 1} \{\phi(M)\}$.

Rewrite $\phi\phi'' - (p')(\phi)^2 \geq 0$ as $\phi^{p'+1} \cdot \left(\phi'/\phi^{p'}\right)' \geq 0$ or equivalently $\left(\phi'/\phi^{p'}\right)' \geq 0$. Let $\left(\phi'/\phi^{p'}\right)' = g(M) \geq 0$ and denote $G(M) = \int_0^M g$, we can solve

$$\phi(M) = \left[\frac{p-1}{C_2 M + C_1 - \int_0^M G}\right]^{p-1},$$

where C_1 and C_2 are some constants, such that $C_2 M + C_1 - \int_0^M G \geq 0$ for $0 \leq M \leq 1$.

Note that $\phi'(M) \leq -1$, so $\sup_{0 \leq M \leq 1} \phi(M) = \phi(0) = \left[\frac{p-1}{C_1}\right]^{p-1}$. All we need to do now is to minimize $\left[\frac{p-1}{C_1}\right]^{p-1}$ among all possible $\phi(M)$.

To this end, we compute

$$\phi'(M) = -\left[\frac{p-1}{C_2 M + C_1 - \int_0^M G}\right]^p \cdot \left[C_2 - G(M)\right],$$

and use again $\phi'(M) \leq -1$ with $M = 1$, which yields

$$C_1 \leq -C_2 + \int_0^1 G + (p-1) \cdot \left[C_2 - G(1)\right]^{\frac{1}{p}}.$$

Remember that $G'(M) = g(M) \geq 0$, thus $G(M)$ is increasing, in particular, $\int_0^1 G \leq G(1)$, so $C_1 \leq -[C_2 - G(1)] + (p-1) \cdot [C_2 - G(1)]^{\frac{1}{p}}$. An easy calculation gives the maximum of the right hand side equals $(p-1) \cdot (p')^{-p'}$ when $C_2 = G(1) + (p')^{-p'}$, therefore, C_1 is at most $(p-1) \cdot (p')^{-p'}$ and thus $\left[\frac{p-1}{C_1}\right]^{p-1} \geq (p')^p$.

To write down an explicit super-solution, simply take $G(M) = 0$, $C_2 = (p')^{-p'}$ and $C_1 = (p-1) \cdot (p')^{-p'}$, then

$$\phi(M) = \left[\frac{p-1}{C_2 M + C_1 - \int_0^M G}\right]^{p-1} = \frac{p^p}{(p-1) \cdot [M + (p-1)]^{p-1}},$$

and recall the relation $\mathbb{B}(F, \mathbf{f}, M) = u(\mathbf{f}, M) + C_p F = (p')^p F - \mathbf{f}^p \cdot \phi(M)$, we obtain

$$u(\mathbf{f}, M) = -\frac{(p\mathbf{f})^p}{(p-1) \cdot [M + (p-1)]^{p-1}}, \tag{4.8}$$

$$\mathbb{B}(F, \mathbf{f}, M) = (p')^p F - \frac{(p\mathbf{f})^p}{(p-1) \cdot [M + (p-1)]^{p-1}}. \tag{4.9}$$

In the general case, we have $u(\mathbf{f}, M; C) = C \cdot u(\mathbf{f}, M/C)$ and $\mathbb{B}(F, \mathbf{f}, M; C) = C \cdot \mathbb{B}(F, \mathbf{f}, M/C)$. Therefore, we have proved the following theorem.

Theorem 4.11. *The Burkhölder's hull of the dyadic Carleson Embedding Theorem 1.9 is given by*

$$u(\mathbf{f}, M; C) = -\frac{C \cdot (p\mathbf{f})^p}{(p-1) \cdot [\frac{M}{C} + (p-1)]^{p-1}}. \tag{4.10}$$

A super-solution that gives the sharpness $C_p = (p')^p$ is

$$\mathbb{B}(F, \mathbf{f}, M; C) = (p')^p F - \frac{C \cdot (p\mathbf{f})^p}{(p-1) \cdot [\frac{M}{C} + (p-1)]^{p-1}}. \tag{4.11}$$

Remark 4.12. Now that the dyadic Carleson Embedding Theorem 1.9 is proved, the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$ exists with $C_p = (p')^p$. However, the super-solution $\mathbb{B}(F, \mathbf{f}, M; C)$ obtained above is not the real Bellman function, since on the boundary $F = \mathrm{f}^p$ the real Bellman function must satisfy the boundary condition $\mathcal{B}(F, \mathbf{f}, M; C) = M\mathbf{f}^p = MF$, but the function we constructed does not satisfy this condition. So, this super-solution only touches the real one along some set. For the exact Bellman function $\mathcal{B}(F, \mathbf{f}, M; C)$, see [22] and [23].

CHAPTER 5

The Bellman functions on filtered probability spaces II: Remodeling and proof of the main theorems

In this chapter, we prove the two main results: Theorem 1.13 and Theorem 1.15. The proof depends on a remodeling of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C = 1)$ for an infinitely refining filtration.

1. Properties of the Bellman function $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M, C)$

The Bellman function $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C)$ associated to the martingale Carleson Embedding Theorem 1.13 does not formally have the main inequality. But it still satisfies the following properties.

Proposition 5.1 (Properties of the Bellman function $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C)$).

(i) Domain: $\mathbf{f}^p \leq F$ and $0 \leq M \leq C$.
(ii) Range: $0 \leq \mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C) \leq C_p \cdot C \cdot F$.
(iii) Homogeniety: $\mathcal{B}_\mu^\mathcal{F}(t^p F, t\mathbf{f}, M; C) = t^p \cdot \mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C)$ for all $t \geq 0$.
(iv) Scaling Property: $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C) = C \cdot \mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M/C; 1)$.
(v) $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C) \geq \mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M - \Delta M; C) + \Delta M \cdot \mathbf{f}^p$ for $0 \leq \Delta M \leq M$.
In particular, $\mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M; C)$ is increasing in M.

Proof. (i) follows from the Hölder's inequality and that $\{\alpha_n\}_{n\geq 0}$ is a Carleson sequence. (ii) holds if we assume Theorem 1.12 is true. (iii) and (iv) are obtained directly from definition (1.18). We explain (v) in more detail. Choose $f \geq 0$ and $\{\alpha_n\}_{n\geq 0}$ that almost give the supremum in the definition (1.18), i.e. for small $\varepsilon > 0$,

$$\mathbb{E}^\mu \left[\sum_{n\geq 0} \alpha_n f_n^p \right] \geq \mathcal{B}_\mu^\mathcal{F}(F, \mathbf{f}, M - \Delta M; C) - \varepsilon,$$

where $\mathbb{E}^\mu[f^p] = F$, $\mathbb{E}^\mu[f] = \mathbf{f}$, $\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right] = M - \Delta M$ and $\mathbb{E}^\mu\left[\sum_{k\geq n}\alpha_k|\mathcal{F}_n\right] \leq C$ for every $n \geq 0$. Since $0 \leq M \leq C$, if we increase α_0 to $\alpha_0 + \Delta M$ then everything is retained except we have now $\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right] = M$ and

$$\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n f_n^p\right] \geq \mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M - \Delta M; C) - \varepsilon + \Delta M \cdot \mathbf{f}^p.$$

Letting $\varepsilon \to 0$, we obtain $\mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M; C) \geq \mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M - \Delta M; C) + \Delta M \cdot \mathbf{f}^p$. □

2. Remodeling of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C = 1)$ for an infinitely refining filtration

In this section, we present a remodeling of the Bellman function $\mathcal{B}(F, \mathbf{f}, M; C = 1)$ for an infinitely refining filtration, which is central to the proof of Theorem 1.13 and Theorem 1.15. We use the notation $\mathcal{B}(F, \mathbf{f}, M) = \mathcal{B}(F, \mathbf{f}, M; C = 1)$ in this and later sections.

Consider the unit interval $I = [0, 1] \in \mathcal{D}$, let $\{I_j^k : 1 \leq j \leq 2^k\}$ be its k-th generation desendant by subdividing I into 2^k congruent dyadic intervals and denote $I_1^0 = I$.

Starting from the definition (1.15) of the Bellman function $\mathcal{B}(F, \mathbf{f}, M)$, we can find a function $f \geq 0$ with $\langle f^p\rangle_I = F$, $\langle f\rangle_I = \mathbf{f}$ and a sequence $\{\alpha_J\}_{J\subseteq I}$, $\sum_{J\subseteq I}\alpha_J = M$ satisfying the Carleson condition with constant $C = 1$, such that the sum $\sum_{J\subseteq I}\alpha_J\langle f\rangle_J^p$ (almost) attains $\mathcal{B}(F, \mathbf{f}, M)$.

To proceed, we further assume that the sequence $\{\alpha_J\}_{J\subseteq I}$ has only *finitely many* non-zero terms. Hence, the indices of $\{\alpha_J\}_{J\subseteq I}$ belong to the collection $\{I_j^k : 1 \leq k \leq N, 1 \leq j \leq 2^k\}$ for some fixed integer N, i.e. for all $J \notin \{I_j^k : 1 \leq k \leq N, 1 \leq j \leq 2^k\}$, we have $\alpha_J = 0$. As a consequence, we can think the function f being piecewise constant on all $\{I_j^N : 1 \leq j \leq 2^N\}$.

Now, let us do the remodeling. Fix a small ε, $0 < \varepsilon < 1$. Consider a discrete-time filtered probability space $(\mathcal{X}, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mu)$. The initial construction is $\mathcal{X}_1^0 = \mathcal{X}$, and this is $\mathcal{F}_{n_0}$-measurable, where $n_0 = 0$. Assume that the $\mathcal{F}_{n_k}$-measurable sets $\mathcal{X}_j^k, 1 \leq j \leq 2^k$ are constructed. We want to inductively construct $\mathcal{F}_{n_{k+1}}$-measurable sets $\mathcal{X}_j^{k+1}, 1 \leq j \leq 2^{k+1}$. Take a $\mathcal{F}_{n_k}$-measurable set $\mathcal{X}_j^k$. Our construction consists of two steps.

The first step is a modification of the set $\mathcal{X}_j^k$. For the given $\varepsilon > 0$ and $\mathcal{X}_j^k \in \mathcal{F}_{n_k}$, Definition 1.11 guarantees the existence of a real-valued $\mathcal{F}_{n_j^k}$-measurable random variable h ($n_j^k > n_k$), such that: (i) $|h\mathbf{1}_E| = \mathbf{1}_E$ and (ii) $\int_{\mathcal{X}_j^k} |h_{n_k}| d\mu \leq \frac{\varepsilon^2}{4}\mu(X_j^k)$. The condition (ii) is chosen in such a way that

$$\mu\left(\left\{x \in \mathcal{X}_j^k : |h_{n_k}| > \frac{\varepsilon}{2}\right\}\right) \leq \frac{\varepsilon}{2}\mu(\mathcal{X}_j^k). \tag{5.1}$$

Let $\widetilde{\mathcal{X}_j^k} = \mathcal{X}_j^k \setminus \{x \in \mathcal{X}_j^k : |h_{n_k}| > \varepsilon/2\}$. So we can conclude $|h_{n_k}| \leq \varepsilon/2$ on $\widetilde{\mathcal{X}_j^k}$, and moreover, $(1-\varepsilon/2)\,\mu(\mathcal{X}_j^k) \leq \mu(\widetilde{\mathcal{X}_j^k}) \leq \mu(\mathcal{X}_j^k)$.

In the second step, we set $\mathcal{X}_{2j-1}^{k+1} = \widetilde{X_j^k} \cap \{h = 1\}$ and $\mathcal{X}_{2j}^{k+1} = \widetilde{X_j^k} \cap \{h = -1\}$. Since $\left|\int_{\widetilde{\mathcal{X}_j^k}} h d\mu\right| \leq \int_{\widetilde{\mathcal{X}_j^k}} |h_{n_k}| d\mu \leq \frac{\varepsilon}{2}\mu(\widetilde{\mathcal{X}_j^k})$, which gives $\left|\mu(\mathcal{X}_{2j-1}^{k+1}) - \mu(\mathcal{X}_{2j}^{k+1})\right| \leq \frac{\varepsilon}{2}\mu(\widetilde{\mathcal{X}_j^k}) \leq \frac{\varepsilon}{2}\mu(\mathcal{X}_j^k)$, we have

$$\frac{1}{2}(1-\varepsilon) \leq \max\left\{\frac{\mu(\mathcal{X}_{2j-1}^{k+1})}{\mu(\mathcal{X}_j^k)}, \frac{\mu(\mathcal{X}_{2j}^{k+1})}{\mu(\mathcal{X}_j^k)}\right\} \leq \frac{1}{2}(1+\varepsilon). \tag{5.2}$$

Do this for all $\mathcal{X}_j^k$, $1 \leq j \leq 2^k$ and let $n_{k+1} = \max\{n_j^k : 1 \leq j \leq 2^k\}$. Hence, we construct $\mathcal{F}_{n_{k+1}}$-measurable sets $\mathcal{X}_j^{k+1}, 1 \leq j \leq 2^{k+1}$. Our construction stops when $k = N$.

Now that we have constructed $\{\mathcal{X}_j^k : 0 \leq k \leq N, 1 \leq j \leq 2^k\}$. We can define a new sequence $\{\alpha_n\}_{n\geq 0}$ on the space $(\mathcal{X}, \mathcal{F}, \mu)$ as

$$\alpha_{n_k} = \begin{cases} \mu(\mathcal{X}_j^k)^{-1}\alpha_{I_j^k}, \text{ if } x \in \mathcal{X}_j^k \\ 0, \text{ if } x \in \mathcal{X} \setminus \bigcup_{j=1}^{2^k} \mathcal{X}_j^k \end{cases}$$

and $\alpha_n = 0$ for all n's different from $n_k, 1 \leq k \leq N$.

Finally, set the new function $\widetilde{f}$ as $\widetilde{f}\mathbf{1}_{\mathcal{X}_j^N} = f\mathbf{1}_{I_j^N}, 1 \leq j \leq 2^N$, and set $\widetilde{f} = 0$ on $\mathcal{X} \setminus \bigcup_{j=1}^{2^N} \mathcal{X}_j^N$. Note that the function $\widetilde{f}$ is also piecewise constant on all $\{\mathcal{X}_j^k : 0 \leq k \leq N, 1 \leq j \leq 2^k\}$.

REMARK 5.2. This construction guarantees that $\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right] = \sum_{J\subseteq I}\alpha_J = M$ and $\mathbb{E}^\mu\left[\widetilde{f}\right] = \langle f\rangle_I = \mathbf{f}$. Later in subsection 2.2 and subsection 3.2, we use a slightly modified version of this construction.

We will frequently consult to the following proposition.

PROPOSITION 5.3. (i) $\frac{1}{2}(1-\varepsilon) \leq \max\left\{\frac{\mu(\mathcal{X}_{2j-1}^{k+1})}{\mu(\mathcal{X}_j^k)}, \frac{\mu(\mathcal{X}_{2j}^{k+1})}{\mu(\mathcal{X}_j^k)}\right\} \leq \frac{1}{2}(1+\varepsilon)$.

(ii) *For every subset $E \in \mathcal{F}_{n_k}$ and $\mu(E\cap\mathcal{X}_j^k) > 0$, we have*

$$\max\left\{\frac{\mu(E\cap\mathcal{X}_{2j-1}^{k+1})}{\mu(E\cap\mathcal{X}_j^k)}, \frac{\mu(E\cap\mathcal{X}_{2j}^{k+1})}{\mu(E\cap\mathcal{X}_j^k)}\right\} \leq \frac{1}{2}(1+\varepsilon). \tag{5.3}$$

Combined with (i), *we have*

$$\max\left\{\frac{\mu(E\cap\mathcal{X}_{2j-1}^{k+1})}{\mu(\mathcal{X}_{2j-1}^{k+1})}, \frac{\mu(E\cap\mathcal{X}_{2j}^{k+1})}{\mu(\mathcal{X}_{2j}^{k+1})}\right\} \leq \frac{1+\varepsilon}{1-\varepsilon}\cdot\frac{\mu(E\cap\mathcal{X}_j^k)}{\mu(\mathcal{X}_j^k)}. \tag{5.4}$$

(iii) $(1-\varepsilon)^k \leq \frac{\mu(\mathcal{X}_j^k)}{|I_j^k|} \leq (1+\varepsilon)^k$ *for all* $0\leq k\leq N, 1\leq j\leq 2^k$.

(iv) $(1-\varepsilon)^k\langle f\rangle_{I_j^{N-k}} \leq \langle\widetilde{f}\rangle_{X_j^{N-k},\mu} \leq (1+\varepsilon)^k\langle f\rangle_{I_j^{N-k}}$ *for all* $0\leq k\leq N, 1\leq j\leq 2^k$.

PROOF. (i) This is (5.2) from our construction.

(ii) This is an important extension of (i). But we only have the upper bound estimation in this general case. Recall that our construction gives $|h_{n_k}| \leq \varepsilon/2$ on $\widetilde{\mathcal{X}_j^k}$, so $\left|\int_{E\cap\widetilde{\mathcal{X}_j^k}} h d\mu\right| \leq \int_{E\cap\widetilde{\mathcal{X}_j^k}} |h_{n_k}| d\mu \leq \frac{\varepsilon}{2}\mu(E\cap\widetilde{\mathcal{X}_j^k})$, which is $\left|\mu(E\cap\mathcal{X}_{2j-1}^{k+1}) - \mu(E\cap\mathcal{X}_{2j}^{k+1})\right| \leq \frac{\varepsilon}{2}\mu(E\cap\widetilde{\mathcal{X}_j^k}) \leq \frac{\varepsilon}{2}\mu(E\cap\mathcal{X}_j^k)$. So we obtain (5.3). (5.4) follows from (5.3) and (i).

(iii) We prove this by induction. For $k=0$, we have $\mu(\mathcal{X}_1^0) = |I_1^0| = 1$. Assuming (iii) holds for k, by (i) we can estimate for $\mathcal{X}_{2j}^{k+1}$ (same for $\mathcal{X}_{2j-1}^{k+1}$) that

$$(1-\varepsilon)^{k+1} \leq \frac{\mu(\mathcal{X}_{2j}^{k+1})}{|I_{2j}^{k+1}|} = 2\cdot\frac{\mu(\mathcal{X}_{2j}^{k+1})}{\mu(\mathcal{X}_j^k)}\cdot\frac{\mu(\mathcal{X}_j^k)}{|I_j^k|} \leq (1+\varepsilon)^{k+1}.$$

(iv) Again by induction, for $k=0$, since $\widetilde{f}\mathbf{1}_{\mathcal{X}_j^N} = f\mathbf{1}_{I_j^N}, 1\leq j\leq 2^N$ and $\widetilde{f}=0$ on $\mathcal{X}\setminus\bigcup_{j=1}^{2^N}\mathcal{X}_j^N$, we have $\langle\widetilde{f}\rangle_{X_j^N,\mu} = \langle f\rangle_{I_j^N}$. Assuming (iv) holds for k, by

(i) we have

$$\begin{aligned}(1-\varepsilon)^{k+1}\langle f\rangle_{I_j^{N-(k+1)}} &\leq \langle \widetilde{f}\rangle_{X_j^{N-(k+1)},\mu} \\ &= \frac{\mu(\mathcal{X}_{2j-1}^{N-k})}{\mu(\mathcal{X}_j^{N-(k+1)})}\langle \widetilde{f}\rangle_{\mathcal{X}_{2j-1}^{N-k},\mu} + \frac{\mu(\mathcal{X}_{2j}^{N-k})}{\mu(\mathcal{X}_j^{N-(k+1)})}\langle \widetilde{f}\rangle_{\mathcal{X}_{2j}^{N-k},\mu} \\ &\leq (1+\varepsilon)^{k+1}\langle f\rangle_{I_j^{N-(k+1)}}.\end{aligned}$$

□

3. The Bellman function $\mathcal{B}_\mu^{\mathcal{F}}(F,\mathbf{f},M;C)$ of Theorem 1.13

3.1. $\mathcal{B}_\mu^{\mathcal{F}}(F,\mathbf{f},M) \leq \mathcal{B}(F,\mathbf{f},M)$. We show (1.19) for the case $C=1$ and the general case follows from the scaling property. Take the Bellman function $\mathcal{B}(F,\mathbf{f},M)$ of the dyadic Carleson Embedding Theorem. Consider an arbitrary function $f\geq 0$ and an arbitrary Carleson sequence $\{\alpha_n\}_{n\geq 0}$ with $C=1$. Set for every $n\geq 0$,

$$X^n = (F^n,\mathbf{f}^n,M^n) = \left(\mathbb{E}^\mu\left[f^p|\mathcal{F}_n\right],\mathbb{E}^\mu\left[f|\mathcal{F}_n\right],\mathbb{E}^\mu\left[\sum_{k\geq n}\alpha_k|\mathcal{F}_n\right]\right).$$

Fix the initial step

$$X^0 = \left(\mathbb{E}^\mu[f^p],\mathbb{E}^\mu[f],\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right]\right) = (F,\mathbf{f},M).$$

By (1.16), we have $0\leq M^n\leq 1$. Also, $\mathbf{f}^n = f_n$ and when $n\geq 1$, $F^n,\mathbf{f}^n$ and M^n are random variables.

Lemma 5.4. *For every $n\geq 0$, we have*

$$\mathbb{E}^\mu\left[\mathcal{B}(X^n)\right] - \mathbb{E}^\mu\left[\mathcal{B}(X^{n+1})\right] \geq \mathbb{E}^\mu\left[\alpha_n f_n^p\right],$$

where $\mathcal{B}(X^n) = \mathcal{B}(F^n,\mathbf{f}^n,M^n)$.

Proof. Recall that the Bellman function $\mathcal{B}(F,\mathbf{f},M)$ satisfies (4.3). Note also we have

$$X^n = \mathbb{E}^\mu\left[X^{n+1}|\mathcal{F}_n\right] + (0,0,\alpha_n).$$

By (4.3) and the Jensen's inequality, we deduce

$$\mathcal{B}(X^n) \geq \mathcal{B}\left(\mathbb{E}^\mu\left[X^{n+1}|\mathcal{F}_n\right]\right) + \alpha_n f_n^p \geq \mathbb{E}^\mu\left[\mathcal{B}(X^{n+1})|\mathcal{F}_n\right] + \alpha_n f_n^p.$$

Taking expectation, we prove exactly

$$\mathbb{E}^\mu\left[\mathcal{B}(X^n)\right] - \mathbb{E}^\mu\left[\mathcal{B}(X^{n+1})\right] \geq \mathbb{E}^\mu\left[\alpha_n f_n^p\right].$$

□

Summing up, we get the inequality

$$\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n f_n^p\right] \leq \sum_{n\geq 0}\left(\mathbb{E}^\mu[\mathcal{B}(X^n)] - \mathbb{E}^\mu[\mathcal{B}(X^{n+1})]\right) \leq \mathcal{B}(X^0).$$

Hence, we conclude that $\mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M) \leq \mathcal{B}(F, \mathbf{f}, M)$.

3.2. $\mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M) = \mathcal{B}(F, \mathbf{f}, M)$ **for an infinitely refining filtration.** To show (1.20), again we consider $C = 1$. Note first that on the boundary $\mathbf{f}^p = F$, we have $\mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, M) = \mathcal{B}(F, \mathbf{f}, M) = MF$. For the case $\mathbf{f}^p < F$, we need to apply the remodeling from section 1.

For technical issues, we slightly modify our remodeling here. First, by the continuity of $\mathcal{B}$, there exists $\delta_1 > 0$, such that $\mathbf{f}^p < F - \delta_1$ and $\mathcal{B}(F - \delta_1, \mathbf{f}, M)$ is close to $\mathcal{B}(F, \mathbf{f}, M)$. Next, by the definition of $\mathcal{B}$, we can find a non-negative function f on the unit interval $I = [0, 1]$ with $\langle f^p\rangle_I = F - \delta_1$, $\langle f\rangle_I = \mathbf{f}$ and a sequence $\{\alpha_J\}_{J\subseteq I}$, $\sum_{J\subseteq I}\alpha_J = M$ satisfying the Carleson condition with constant $C = 1$, such that the sum $\sum_{J\subseteq I}\alpha_J\langle f\rangle_J^p$ (almost) equals $\mathcal{B}(F, \mathbf{f}, M)$. Moreover, by again the continuity, we can choose a *finite* subset of $\{\alpha_J\}_{J\subseteq I}$ such that $\sum_{J\subseteq I}\alpha_J = M - \delta_2$ for some $\delta_2 > 0$ and $\sum_{J\subseteq I}\alpha_J\langle f\rangle_J^p$ still (almost) equals $\mathcal{B}(F, \mathbf{f}, M)$. For simplicity, we assume exactly

$$\sum_{J\subseteq I}\alpha_J\langle f\rangle_J^p = \mathcal{B}(F, \mathbf{f}, M). \tag{5.5}$$

Let the indices of $\{\alpha_J\}_{J\subseteq I}$ belong to the collection $\{I_j^k : 1 \leq k \leq N, 1 \leq j \leq 2^k\}$ for some fixed integer N. Choose $\varepsilon > 0$, such that $F - \delta_1 \leq F/(1+\varepsilon)^N$. We do the remodeling with this $\varepsilon > 0$ to construct $\{\mathcal{X}_j^k : 0 \leq k \leq N, 1 \leq j \leq 2^k\}$, $\{\alpha_n\}_{n\geq 0}$ and $\widetilde{f}$ on the space $(\mathcal{X}, \mathcal{F}, \mu)$. To proceed, we observe that

Lemma 5.5.

$$\mathbb{E}^\mu\left[\widetilde{f^p}\right] \le (1+\varepsilon)^N \langle f^p\rangle_I. \tag{5.6}$$

Proof. By (iii) of Proposition 5.3,

$$\mathbb{E}^\mu\left[\widetilde{f^p}\right] = \sum_{j=1}^{2^N}\langle\widetilde{f^p}\rangle_{\mathcal{X}_j^N,\mu}\mu(\mathcal{X}_j^N) \le \sum_{j=1}^{2^N}\langle f^p\rangle_{I_j^N}\cdot(1+\varepsilon)^N|I_j^N| = (1+\varepsilon)^N\langle f^p\rangle_I.$$

□

So (5.6) and $\langle f^p\rangle_I = F - \delta_1 \le F/(1+\varepsilon)^N$ imply that $\mathbb{E}^\mu\left[\widetilde{f^p}\right] \le F$. Also recall from the remodeling, we know $\mathbb{E}^\mu\left[\widetilde{f}\right] = \langle f\rangle_I = \mathbf{f}$. Let us further modify the function $\widetilde{f}$ in the following way. Note that we are working on an infinitely refining filtration (see definition 1.11). There exists a simple function g behaving like a Haar function, such that g is supported on $\mathcal{X}_1^N$, $\langle g\rangle_{\mathcal{X}_1^N,\mu} = 0$ and $0 < \mathbb{E}^\mu\left[|g|^p\right] < \infty$. Consider the continuous function

$$a(t) = \mathbb{E}^\mu\left[\left|\widetilde{f}+tg\right|^p\right].$$

Thus, $a(0) \le F$ and $\lim_{t\to\infty} a(t) = \infty$. Hence, we can find $t_0 \ge 0$, such that $\mathbb{E}^\mu\left[\left|\widetilde{f}+t_0g\right|^p\right] = F$. Update $\widetilde{f}$ to $\widetilde{f}+t_0g$. We have then $\mathbb{E}^\mu\left[\left|\widetilde{f}\right|^p\right] = F$ and $\mathbb{E}^\mu\left[\widetilde{f}\right] = \mathbf{f}$. Note here the updated function $\widetilde{f}$ might be negative, however, all the relevant average values we will use are still non-negative.

Now, let us discuss the properties of the Carleson sequence $\{\alpha_n\}_{n\ge 0}$. Directly from the remodeling, we know $\mathbb{E}^\mu\left[\sum_{n\ge 0}\alpha_n\right] = \sum_{J\subseteq I}\alpha_J = M - \delta_2$. Moreover, we can prove

Lemma 5.6. *The non-negative sequence $\{\alpha_n\}_{n\ge 0}$ satisfies each α_n is $\mathcal{F}_n$-measurable and*

$$\mathbb{E}^\mu\left[\sum_{k\ge n}\alpha_k|\mathcal{F}_n\right] \le \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}} \text{ for every } n \ge 0. \tag{5.7}$$

Proof. From the construction, it is clear that each α_n is non-negative and $\mathcal{F}_n$-measurable. So we need to show for every $\mathcal{F}_n$-measurable set E, we have

$$\mathbb{E}^\mu\left[\sum_{k\ge n}\alpha_k\mathbf{1}_E\right] \le \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}}\cdot\mu(E).$$

Denote by $k_0 = \min\{k : n_k \geq n\}$. Since $\mathbb{E}^\mu\left[\sum_{k\geq n} \alpha_k \mathbf{1}_E\right] = \mathbb{E}^\mu\left[\mathbb{E}^\mu\left[\sum_{k\geq k_0} \alpha_{n_k} | \mathcal{F}_{n_{k_0}}\right] \mathbf{1}_E\right]$, it suffices to show

$$\mathbb{E}^\mu\left[\sum_{k\geq k_0} \alpha_{n_k} | \mathcal{F}_{n_{k_0}}\right] \leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}},$$

or equivalently, for every $\mathcal{F}_{n_{k_0}}$-measurable set E, we have

$$\mathbb{E}^\mu\left[\sum_{k\geq k_0} \alpha_{n_k} \mathbf{1}_E\right] \leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}} \cdot \mu(E).$$

Now the explicit computation shows

$$\mathbb{E}^\mu\left[\sum_{k\geq k_0} \alpha_{n_k} \mathbf{1}_E\right] = \sum_{k\geq k_0} \sum_{j=1}^{2^k} \alpha_{I_j^k} \frac{\mu(E\cap \mathcal{X}_j^k)}{\mu(\mathcal{X}_j^k)}.$$

An iteration of (5.4) gives

$$\frac{\mu(E\cap \mathcal{X}_j^k)}{\mu(\mathcal{X}_j^k)} \leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^N} \cdot \frac{\mu(E\cap \mathcal{X}_l^{k_0})}{\mu(\mathcal{X}_l^{k_0})}, \text{ whenever } \mathcal{X}_j^k \subseteq \mathcal{X}_l^{k_0}.$$

So we can estimate

$$\begin{aligned}
\mathbb{E}^\mu\left[\sum_{k\geq k_0} \alpha_{n_k} \mathbf{1}_E\right] &\leq \sum_{l=1}^{2^{k_0}} \frac{(1+\varepsilon)^N}{(1-\varepsilon)^N} \cdot \frac{\mu(E\cap \mathcal{X}_l^{k_0})}{\mu(\mathcal{X}_l^{k_0})} \sum_{k,j:\mathcal{X}_j^k\subseteq \mathcal{X}_l^{k_0}} \alpha_{I_j^k}, \ \{\alpha_I\} \text{ Carleson sequence} \\
&\leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^N} \sum_{l=1}^{2^{k_0}} \frac{\mu(E\cap \mathcal{X}_l^{k_0})}{\mu(\mathcal{X}_l^{k_0})} \cdot \left|I_l^{k_0}\right|, \text{ Proposition 5.3 (iii)} \\
&\leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}} \sum_{l=1}^{2^{k_0}} \mu(E\cap \mathcal{X}_l^{k_0}) \leq \frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}} \cdot \mu(E).
\end{aligned}$$

□

To finish, we need one final lemma.

Lemma 5.7.

$$\mathbb{E}^\mu\left[\sum_{n\geq 0} \alpha_n \left|\widetilde{f}_n\right|^p\right] \geq (1-\varepsilon)^{pN} \sum_{J\subseteq I} \alpha_J \langle f\rangle_J^p. \tag{5.8}$$

Proof.

$$\mathbb{E}^{\mu}\left[\sum_{n\geq 0}\alpha_n\left|\widetilde{f}_n\right|^p\right]=\mathbb{E}^{\mu}\left[\sum_{k\geq 0}\alpha_{n_k}\left|\widetilde{f}_{n_k}\right|^p\right]=\sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\left\langle\left|\widetilde{f}_{n_k}\right|^p\right\rangle_{\mathcal{X}_j^k,\mu}$$

$$\geq\sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\langle\widetilde{f}_{n_k}\rangle^p_{\mathcal{X}_j^k,\mu}=\sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\langle\widetilde{f}\rangle^p_{\mathcal{X}_j^k,\mu},\ \text{Proposition 5.3 (iv)}$$

$$\geq(1-\varepsilon)^{pN}\sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\langle f\rangle^p_{I_j^k}=(1-\varepsilon)^{pN}\sum_{J\subseteq I}\alpha_J\langle f\rangle^p_J.$$

□

Summarizing, we have constructed a function $\widetilde{f}$ and a Carleson sequence $\{\alpha_n\}_{n\geq 0}$ satisfying (5.7) with $\mathbb{E}^{\mu}\left[\left|\widetilde{f}\right|^p\right]=F$, $\mathbb{E}^{\mu}\left[\widetilde{f}\right]=\mathbf{f}$ and $\mathbb{E}^{\mu}\left[\sum_{n\geq 0}\alpha_n\right]=\sum_{J\subseteq I}\alpha_J=M-\delta_2$. By (5.5) and (5.8), we deduce

$$\mathcal{B}_{\mu}^{\mathcal{F}}\left(F,\mathbf{f},M-\delta_2;C=\frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}}\right)\geq(1-\varepsilon)^{pN}\sum_{J\subseteq I}\alpha_J\langle f\rangle^p_J=(1-\varepsilon)^{pN}\mathcal{B}\left(F,\mathbf{f},M\right).$$

And Proposition 5.1 (iv) and (v) imply that

$$\mathcal{B}_{\mu}^{\mathcal{F}}\left(F,\mathbf{f},M-\delta_2;C=\frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}}\right)=\frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}}\mathcal{B}_{\mu}^{\mathcal{F}}\left(F,\mathbf{f},\frac{(1-\varepsilon)^{2N}}{(1+\varepsilon)^N}(M-\delta_2)\right)$$
$$\leq\frac{(1+\varepsilon)^N}{(1-\varepsilon)^{2N}}\mathcal{B}_{\mu}^{\mathcal{F}}\left(F,\mathbf{f},M\right).$$

Letting $\varepsilon\to 0$, we prove exactly $\mathcal{B}_{\mu}^{\mathcal{F}}(F,\mathbf{f},M)\geq\mathcal{B}(F,\mathbf{f},M)$. The other inequality is proved in the subsection 3.1.

4. The Bellman function $\widetilde{\mathcal{B}}_{\mu}^{\mathcal{F}}(F,\mathbf{f})$ of the maximal operators

4.1. $\widetilde{\mathcal{B}}_{\mu}^{\mathcal{F}}(F,\mathbf{f})\leq\mathcal{B}_{\mu}^{\mathcal{F}}(F,\mathbf{f},1)$. Let us relate the maximal function (1.21) to the Bellman function $\mathcal{B}_{\mu}^{\mathcal{F}}(F,\mathbf{f},M)$. Define a sequence of sets

$$E_n=\{x\in\mathcal{X}:n\text{ is the smallest non-negative integer, such that }f^*(x)=|f_n(x)|\}.$$

Obviously, $\{E_n\}_{n\geq 0}$ forms a disjoint partition of $\mathcal{X}$. We can compute

$$\begin{aligned}||f^*||^p_{L^p(\mathcal{X},\mathcal{F},\mu)} &= \mathbb{E}^\mu\left[|f^*|^p\right] = \mathbb{E}^\mu\left[\sum_{n\geq 0}|f_n|^p \mathbf{1}_{E_n}\right] \\ &= \mathbb{E}^\mu\left[\sum_{n\geq 0}\mathbb{E}^\mu[\mathbf{1}_{E_n}|\mathcal{F}_n]\cdot|f_n|^p\right].\end{aligned}$$

Let $\alpha_n = \mathbb{E}^\mu[\mathbf{1}_{E_n}|\mathcal{F}_n]$, $n\geq 0$. The connection between the maximal function (1.21) and $\mathcal{B}^{\mathcal{F}}_\mu(F,\mathbf{f},M)$ relies on the following simple fact.

Lemma 5.8. *$\{\alpha_n\}_{n\geq 0}$ is a Carleson sequence with $C=1$. (see Definition 1.10).*

Proof. It is clear that each α_n is non-negative and $\mathcal{F}_n$-measurable. Moreover, for every set $E\in\mathcal{F}_n$, we have $\mathbb{E}^\mu\left[\sum_{k\geq n}\alpha_k\mathbf{1}_E\right] = \mathbb{E}^\mu\left[\sum_{k\geq n}\mathbf{1}_{E_k\cap E}\right]\leq \mu(E)$. So we prove the claim. □

To prove (1.24), fix $\mathbb{E}^\mu[f^p]=F$ and $\mathbb{E}^\mu[f]=\mathbf{f}$. Since $\{\alpha_n\}_{n\geq 0}$ is a Carleson sequence with $C=1$ and $\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\right]=1$, we conclude that $\widetilde{\mathcal{B}}^{\mathcal{F}}_\mu(F,\mathbf{f})\leq \mathcal{B}^{\mathcal{F}}_\mu(F,\mathbf{f},1)$.

4.2. $\widetilde{\mathcal{B}}^{\mathcal{F}}_\mu(F,\mathbf{f}) = \mathcal{B}^{\mathcal{F}}_\mu(F,\mathbf{f},1)$ for an infinitely refining filtration. Again, we appeal to the modified remodeling from subsection 2.2, but only with $M=1$. Note that we have

$$\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\left|\widetilde{f}_n\right|^p\right] = \mathbb{E}^\mu\left[\sum_{k\geq 0}\alpha_{n_k}\left|\widetilde{f}_{n_k}\right|^p\right] = \sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I^k_j}\left\langle\left|\widetilde{f}_{n_k}\right|^p\right\rangle_{\mathcal{X}^k_j,\mu}.$$

To proceed, we observe that

Lemma 5.9. *For every $0\leq k\leq N$ and $1\leq j\leq 2^k$, we have*

$$\left|\widetilde{f}_{n_{N-k}}\right|\mathbf{1}_{\mathcal{X}^{N-k}_j}\leq\frac{(1+\varepsilon)^k}{(1-\varepsilon)^k}\langle\widetilde{f}\rangle_{\mathcal{X}^{N-k}_j,\mu}. \tag{5.9}$$

Proof. First note that $\left|\widetilde{f}_{n_{N-k}}\right|\mathbf{1}_{\mathcal{X}^{N-k}_j} = \widetilde{f}_{n_{N-k}}\mathbf{1}_{\mathcal{X}^{N-k}_j}$ for every $0\leq k\leq N$ and $1\leq j\leq 2^k$. Induction on k, for $k=0$, the construction of $\widetilde{f}$ immediately gives $\widetilde{f}_{n_N}\mathbf{1}_{\mathcal{X}^N_j} = \langle\widetilde{f}\rangle_{\mathcal{X}^N_j,\mu}$, $1\leq j\leq 2^N$. Assuming (5.9) holds for k, then for every

$\mathcal{F}_{n_{N-(k+1)}}$-measurable set E, $E \subseteq \mathcal{X}_j^{N-(k+1)}$ and $\mu(E) > 0$, we can estimate

$$\int_E \widetilde{f}_{n_{N-(k+1)}} \mathbf{1}_{\mathcal{X}_j^{N-(k+1)}} d\mu = \int_{E\cap\mathcal{X}_j^{N-(k+1)}} \widetilde{f} d\mu = \int_{E\cap\mathcal{X}_{2j-1}^{N-k}} \widetilde{f}_{n_{N-k}} d\mu + \int_{E\cap\mathcal{X}_{2j}^{N-k}} \widetilde{f}_{n_{N-k}} d\mu$$
$$\leq \frac{(1+\varepsilon)^k}{(1-\varepsilon)^k}\left[\langle\widetilde{f}\rangle_{\mathcal{X}_{2j-1}^{N-k},\mu}\mu(E\cap\mathcal{X}_{2j-1}^{N-k}) + \langle\widetilde{f}\rangle_{\mathcal{X}_{2j}^{N-k},\mu}\mu(E\cap\mathcal{X}_{2j}^{N-k})\right].$$

And hence, we deduce

$$\mu(E)^{-1}\int_E \widetilde{f}_{n_{N-(k+1)}} \mathbf{1}_{\mathcal{X}_j^{N-(k+1)}} d\mu, \ (E \subseteq \mathcal{X}_j^{N-(k+1)})$$
$$\leq \frac{(1+\varepsilon)^k}{(1-\varepsilon)^k}\left[\langle\widetilde{f}\rangle_{\mathcal{X}_{2j-1}^{N-k},\mu}\frac{\mu(E\cap\mathcal{X}_{2j-1}^{N-k})}{\mu(E\cap\mathcal{X}_j^{N-k+1})} + \langle\widetilde{f}\rangle_{\mathcal{X}_{2j}^{N-k},\mu}\frac{\mu(E\cap\mathcal{X}_{2j}^{N-k})}{\mu(E\cap\mathcal{X}_j^{N-k+1})}\right], \ (5.3)$$
$$\leq \frac{1}{2}\frac{(1+\varepsilon)^{k+1}}{(1-\varepsilon)^k}\left[\langle\widetilde{f}\rangle_{\mathcal{X}_{2j-1}^{N-k},\mu} + \langle\widetilde{f}\rangle_{\mathcal{X}_{2j}^{N-k},\mu}\right], \ \text{Proposition 5.3 (i)}$$
$$\leq \frac{(1+\varepsilon)^{k+1}}{(1-\varepsilon)^{k+1}}\langle\widetilde{f}\rangle_{\mathcal{X}_j^{N-(k+1)},\mu}.$$

Since this is true for every $\mathcal{F}_{n_{N-(k+1)}}$-measurable set E, $E \subseteq \mathcal{X}_j^{N-(k+1)}$ and $\mu(E) > 0$, we prove (5.9) for $k+1$. □

Applying (5.9), we have

$$\mathbb{E}^\mu\left[\sum_{n\geq 0}\alpha_n\left|\widetilde{f}_n\right|^p\right] = \sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\left\langle\left|\widetilde{f}_{n_k}\right|^p\right\rangle_{\mathcal{X}_j^k,\mu} \leq \frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{pN}}\sum_{k\geq 0}\sum_{j=1}^{2^k}\alpha_{I_j^k}\langle\widetilde{f}\rangle^p_{\mathcal{X}_j^k,\mu}.$$

And note that Proposition 5.3 (iii) implies

$$\sum_{k,j:I_j^k\subseteq I_{j_0}^{k_0}}\alpha_{I_j^k} \leq \left|I_{j_0}^{k_0}\right| \leq \frac{1}{(1-\varepsilon)^N}\mu(\mathcal{X}_{j_0}^{k_0}) \text{ for every } 0\leq k_0\leq N, 1\leq j_0\leq 2^{k_0}.$$

Now, let us recall a useful lemma established in [22], formulated in our language,

Lemma 5.10. *Suppose* $\alpha_{I_j^k} \geq 0$*, where* $0\leq k\leq N, 1\leq j\leq 2^k$*, satisfies*

$$\sum_{k,j:I_j^k\subseteq I_{j_0}^{k_0}}\alpha_{I_j^k} \leq C\mu(\mathcal{X}_{j_0}^{k_0}) \tag{5.10}$$

for some constant $C > 0$*, then we can choose pairwise disjoint measurable* $\mathcal{A}_j^k \subseteq \mathcal{X}$ *such that* $\mathcal{A}_j^k \subseteq \mathcal{X}_j^k$ *and* $\alpha_{I_j^k} = C\mu(\mathcal{A}_j^k)$*.*

Proof. Without loss of generality, we can assume $C = 1$. We start at the level $k = N$. Since (5.10) with $C = 1$ implies $\alpha_{I_j^N} \le \mu(\mathcal{X}_j^N)$ for every $1 \le j \le 2^N$, we can choose $\mathcal{A}_j^N \subseteq \mathcal{X}_j^N$ such that $\alpha_{I_j^N} = \mu(\mathcal{A}_j^N)$. Assuming that we have chosen pairwise disjoint measurable $\mathcal{A}_j^k$ for all $k \ge k_0 + 1$ and $1 \le j \le 2^k$, note that (5.10) with $C = 1$ gives

$$\alpha_{I_{j_0}^{k_0}} + \sum_{k,j: I_j^k \subsetneq I_{j_0}^{k_0}} \alpha_{I_j^k} \le \mu(\mathcal{X}_{j_0}^{k_0}), \text{ so } \alpha_{I_{j_0}^{k_0}} \le \mu\left(\mathcal{X}_{j_0}^{k_0} \setminus \bigcup_{k,j: I_j^k \subsetneq I_{j_0}^{k_0}} \mathcal{A}_j^k\right),$$

and thus we can choose measurable set $\mathcal{A}_{j_0}^{k_0} \subseteq \mathcal{X}_{j_0}^{k_0} \setminus \bigcup_{k,j: I_j^k \subsetneq I_{j_0}^{k_0}} \mathcal{A}_j^k$, such that $\alpha_{I_{j_0}^{k_0}} = \mu(\mathcal{A}_{j_0}^{k_0})$. Continue this process for all the indices. This proves the lemma. □

By Lemma 5.10, we can estimate

$$\begin{aligned}
\mathbb{E}^\mu\left[\sum_{n\ge 0} \alpha_n \left|\widetilde{f}_n\right|^p\right] &\le \frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{pN}} \sum_{k\ge 0}\sum_{j=1}^{2^k} \alpha_{I_j^k} \langle \widetilde{f}\rangle^p_{\mathcal{X}_j^k,\mu} \\
&= \frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{pN}} \sum_{k\ge 0}\sum_{j=1}^{2^k} \frac{1}{(1-\varepsilon)^N} \mu(\mathcal{A}_j^k) \langle \widetilde{f}_{n_k}\rangle^p_{\mathcal{X}_j^k,\mu} \\
&\le \frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{(p+1)N}} \sum_{k\ge 0}\sum_{j=1}^{2^k} \mathbb{E}^\mu\left[\left|\widetilde{f}^*\right|^p \mathbf{1}_{\mathcal{A}_j^k}\right], \text{ disjointness} \\
&\le \frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{(p+1)N}} \mathbb{E}^\mu\left[\left|\widetilde{f}^*\right|^p\right].
\end{aligned}$$

Applying (5.5) and (5.8) with $M = 1$, together with Theorem 1.13, we have

$$\begin{aligned}
\frac{(1+\varepsilon)^{pN}}{(1-\varepsilon)^{(p+1)N}} \mathbb{E}^\mu\left[\left|\widetilde{f}^*\right|^p\right] &\ge \mathbb{E}^\mu\left[\sum_{n\ge 0} \alpha_n \left|\widetilde{f}_n\right|^p\right] \ge (1-\varepsilon)^{pN} \sum_{J\subseteq I} \alpha_J \langle f\rangle_J^p \\
&= (1-\varepsilon)^{pN} \mathcal{B}(F, \mathbf{f}, 1) = (1-\varepsilon)^{pN} \mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, 1).
\end{aligned}$$

Recall that $\mathbb{E}^\mu\left[\left|\widetilde{f}\right|^p\right] = F$ and $\mathbb{E}^\mu\left[\widetilde{f}\right] = \mathbf{f}$. Letting $\varepsilon \to 0$, we prove exactly $\widetilde{\mathcal{B}}_\mu^{\mathcal{F}}(F, \mathbf{f}) \ge \mathcal{B}_\mu^{\mathcal{F}}(F, \mathbf{f}, 1)$. The other inequality is proved in the subsection 4.1.

Bibliography

[1] B. Muckenhoupt, *Hardy's inequality with weights*, Studia Math. 44 (1972), 3138. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I.

[2] E. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. **75** (1982), no. 1, 1-11

[3] E. Sawyer, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), no. 2, 533-545.

[4] E. Sawyer and R. Wheeden, *Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces*, Amer. J. Math. **114** (1992), no. 4, 813-874

[5] F. Nazarov, S. Treil and A. Volberg, *The Bellman function and two weight inequalities for Haar multipliers*, J. Amer. Math. Soc. **12** (1999), no. 4, 909-928

[6] M. Lacey, E. Sawyer and I. Uriarte-Tuero, *Two weight inequalities for discrete positive operators*, ArXiv:0911.3437 [math.CA] (2009)

[7] S. Treil, *A remark on two weight estimates for positive dyadic operators*, ArXiv:1201.1145 [math.CA] (2012)

[8] Tuomas P. Hytönen, *The A_2 theorem: remarks and complements*, ArXiv:1212.3840 [math.CA] (2012)

[9] J. Scurry, *A characterization of two-weight inequalities for a vector-valued operator*, ArXiv:1007.3089 [math.CA] (2010)

[10] Timo S. Hänninen, *Another proof of Scurry's characterization of a two weight norm inequality for a sequence-valued positive dyadic operator*, ArXiv:1304.7759 [math.CA] (2013)

[11] F. Nazarov, S. Treil and A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583-597

[12] F. Nazarov, S. Treil and A. Volberg, *Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures*, ArXiv:1003.1596[math.CA] (2010).

[13] M. Lacey, E. Sawyer, C.Y. Shen and I. Uriarte-Tuero, *Two Weight Inequality for the Hilbert Trans-form: A Real Variable Characterization I*, Duke Math. J. **163** (2014), no. 15, 2795-2820

[14] M. Lacey, *Two Weight Inequality for the Hilbert Transform: A Real Variable Characterization II*, Duke Math. J. **163** (2014), no. 15, 2821-2840

[15] M. Lacey, E. Sawyer, C.Y. Shen, I. Uriarte-Tuero and B. Wick, *Two Weight Inequalities for the Cauchy Transform from $\mathbb{R}$ to $\mathbb{C}_+$*, ArXiv:1310.4820 [math.CV] (2013)

[16] M. Lacey and B. Wick, *Two Weight Inequalities for Riesz Transforms: Uniformly Full Dimension Weights*, ArXiv:1312.6163 [math.CA] (2013)

[17] S. Treil, *Commutators, Paraproducs and BMO in non-homogeneous martingale settings*, Rev. Mat. Iberoam. **29** (2013), no. 4, 1325-1372

[18] J. Lai, *A new two weight estimates for a vector-valued positive operator*, ArXiv:1503.06778v2 [math.CA] (2015)

[19] J. Lai and S. Treil, *Two weight estimates for paraproducts in non-homogeneous settings*, Preprint

[20] F. Nazarov, S. Treil, *The hunt for a Bellman function: applications to estimates of singular integral operators and to other classical problems in harmonic analysis*, Algebra i analiz, **8** (1996), no. 5, 32-162 (in Russian); English transl. in St. Petersburg Math. J. **8** (1997), no. 5, 721-824.

[21] F. Nazarov, S. Treil, A. Volberg, *Bellman function in stochastic control and harmonic analysis*, Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 393-423, Oper. Theory Adv. Appl., **129**, Birkhäuser, Basel, 2001.

[22] A. Melas, *The Bellman functions of dyadic-like maximal operators and related inequalities*, Adv. in Math., **192** (2005), no. 2, 310-340.

[23] V. Vasyunin, A. Volberg, *Monge-Ampère equation and Bellman optimization of Carleson Embedding Theorems*, Amer. Math. Soc. Transl. (2) **226** (2009), 195-238

[24] J. Lai, *The Bellman functions of the Carleson Embedding Theorem and the Doob's martingale inequality*, ArXiv:1411.5408v3 [math.CA] (2014)

[25] F. Nazarov, *A counterexample to Sarason's conjecture*, Preprint, Michigan State Univ., 1997, pp. 1-17.

[26] D. Burkhölder, *Martingales and Fourier analysis in Banach spaces*, C.I.M.E. Lectures Varenna Como, Italy, 1985, Lecture Notes Mathematics, **1206**(1986), 61-108.

[27] D. Burkhölder, *Explorations in martingale theory and its applications*, Ecole d'Eté de Probabilité de Saint-Flour XIX-1989, 1-66, Lecture Notes in Mathematics, **1464**(1991), Springer, Berlin.

[28] G. Wang, *Sharp maximal inequalities for conditionally symmetric martingales and Brownian motion*, Proc. Amer. Math. Soc. **112** (1991) 579-586.

[29] C. Niculescu, L. Persson, *Convex functions and their applications: A contemporary approach*, CMS Books in Mathematics, 1st edition, Springer, 2005.

[30] J. L. Doob, *Stochastic Processes*, John Wiley & Sons. ISBN 0-471-52369-0, 1953.