LIPSCHITZ SUBTYPE

R.M. CAUSEY

Abstract. We give necessary and sufficient conditions for a Lipschitz map, or more generally a uniformly Lipschitz family of maps, to factor the Hamming cubes. This is an extension to Lipschitz maps of a particular spatial result of Bourgain, Milman, and Wolfson [3].

1. Introduction

In 1969, M. Ribe [15] proved that two Banach spaces which are uniformly homeomorphic must be crudely finitely representable in each other. Since then, the Ribe program has attracted significant attention (see [13] for a survey on the Ribe program), with the goal of providing purely metric characterizations of local properties of Banach spaces. An important result in this area is that of Bourgain, Milman, and Wolfson, who defined one notion of metric type $p$ and proved that any family of metric spaces with no non-trivial type must contain almost isometric copies of the Hamming cubes. Another goal within the Ribe program is to find, for a given class of important linear operators between Banach spaces, natural metric analogues within the class of Lipschitz maps between metric spaces (see [5], [8], [11]). One such class is the class of super-Rosenthal operators, for which Beauzamy gave a linear characterization in terms of a sequence of subtype constants (see Section 2). The goal of this work is to undertake the process of proving the Lipschitz analogue of Beauzamy’s linear result for the super-Rosenthal operators. We define different notions of subtype constants for a Lipschitz map (or more generally, a uniformly Lipschitz collection of maps) between Banach spaces, which are the analogues of linear subtype constants appearing in the literature in the aforementioned work of Beauzamy and the work of Hinrichs [10]. We prove the non-linear analogues of the results found in the work of Beauzamy and Hinrichs, in that if the subtype constants of a uniformly Lipschitz family of maps exhibit the asymptotically worst possible behavior, then the Lipschitz maps preserve copies of the Hamming cubes. Our subtype constants are based on the Bourgain, Milman, Wolfson notion of metric type.

We next make these descriptions precise, and then state the main result.

We agree to the convention that $0^n = 0$. Given a map $g: (U, d_U) \to (V, d_V)$ between metric spaces, we let

$$\text{Lip}(g) = \sup_{x, y \in U} \frac{d_V(g(x), g(y))}{d_U(x, y)}.$$  

For a map $g: (U, d_U) \to (V, d_V)$, we let $\text{dist}(g) = \infty$ if $g$ is not injective, and otherwise we let $\text{dist}(g) = \text{Lip}(g)\text{Lip}(g^{-1})$, where $g^{-1}$ is understood to be defined on $g(U)$.

We let $2^n = \{\pm 1\}^n$ be the (vertex set of the) Hamming cube. Given $\varepsilon \in 2^n$, we denote the coordinates of $\varepsilon$ by $\varepsilon(1), \varepsilon(2), \ldots$. We endow $2^n$ with the normalized graph metric

$$\partial_n(\varepsilon, \delta) = \frac{1}{n} |\{i : \varepsilon(i) \neq \delta(i)\}|.$$  

When no confusion can arise, we will suppress the subscript $n$ and just write $\partial$. We also endow $2^n$ with the uniform probability measure $\mathbb{P}_n$, also suppressing the subscript when no confusion can arise. Given $1 \leq i \leq n$, we let $d_i$ denote the function on $2^n$ which changes the $i$th coordinate and leaves the other coordinates unchanged.

To avoid cumbersome notation, if $(X, d_X)$, $(Y, d_Y)$ are metric spaces and $f: 2^n \to X$, $F: X \to Y$ are functions, we let $\partial_X^f$, $\partial_Y^f$, respectively, denote the pseudometrics on $2^n$ given by

$$\partial_X^f(\varepsilon, \delta) = d_X(f(\varepsilon), f(\delta))$$  

and

$$\partial_Y^f(\varepsilon, \delta) = d_Y(F \circ f(\varepsilon), F \circ f(\delta)).$$
Now suppose we have $\lambda > 0$ fixed and a collection $\mathcal{F}$ of $\lambda$-Lipschitz functions between (possibly different) metric spaces. For $n \in \mathbb{N}$, we let $a_n(\mathcal{F})$ denote the infimum of those $a > 0$ such that for each $F : X \to Y \in \mathcal{F}$ and $f : 2^n \to X$,

$$E\rho^f_{\mathcal{F}}(\varepsilon, -\varepsilon) \leq a \operatorname{Lip}(f : 2^n \to X).$$

Note that $a \leq \lambda$. These are the Lipschitz analogues of the linear quantities appearing in [2]. We note that there appears a factor of $n^{-1}$ on the expectation. The reason is because this constant $n^{-1}$ has been subsumed by $\operatorname{Lip}(g)$ and our convention of using the normalized graph metric on $2^n$.

For $1 < p < \infty$, let $b_{p,n}(\mathcal{F})$ denote the infimum of those $b > 0$ such that for each $F : X \to Y \in \mathcal{F}$ and $f : 2^n \to X$,

$$E\rho^f_{\mathcal{F}}(\varepsilon, -\varepsilon)^p \leq b^p n^{p-1} \sum_{i=1}^n E\rho^f_X(\varepsilon, d_i \varepsilon)^p.$$

Let us note that by combining the triangle and Hölder inequalities, $b_{p,n}(\mathcal{F}) \leq \lambda$. In the case $p = 2$, these are the Lipschitz analogues of the linear quantities appearing in [10], as well as the generalization to maps of the metric type $1$ constant as defined by Bourgain, Milman, and Wolfson.

Let us say that $\mathcal{F}$ \textit{crudely factors the Hamming cubes} provided that there exist constants $c, D > 0$ such that for each $n \in \mathbb{N}$, there exist $F : X \to Y \in \mathcal{F}$ and $f : 2^n \to X$ and constants $a, b > 0$ such that for each $\varepsilon, \delta \in 2^n$,

$$\frac{a}{D}\partial(\varepsilon, \delta) \leq d_X(f(\varepsilon), f(\delta)) \leq aD\partial(\varepsilon, \delta),$$

$$\frac{b}{D}\partial(\varepsilon, \delta) \leq d_Y(F \circ f(\varepsilon), F \circ f(\delta)) \leq bD\partial(\varepsilon, \delta),$$

and $b \geq ac$. An important feature of this definition is that the scaling factors $a, b$ be uniformly equivalent (that is, $ac \leq b \leq a\lambda D^2$). Let us say that $\mathcal{F}$ \textit{factors the Hamming cubes} provided that there exists a constant $c > 0$ such that for each $D > 1$ and each $n \in \mathbb{N}$, there exist $F : X \to Y \in \mathcal{F}$ and $f : 2^n \to X$ and constants $a, b$ such that for each $\varepsilon, \delta \in 2^n$,

$$\frac{a}{D}\partial(\varepsilon, \delta) \leq d_X(f(\varepsilon), f(\delta)) \leq aD\partial(\varepsilon, \delta),$$

$$\frac{b}{D}\partial(\varepsilon, \delta) \leq d_Y(F \circ f(\varepsilon), F \circ f(\delta)) \leq bD\partial(\varepsilon, \delta),$$

and $b \geq ac$.

We now present the main theorem.

**Theorem.** The following are equivalent:

(i) $\mathcal{F}$ factors the Hamming cubes.

(ii) $\mathcal{F}$ crudely factors the Hamming cubes.

(iii) $\limsup_n a_n(\mathcal{F}) > 0$.

(iv) For each $1 < p < \infty$, $\limsup_n b_{p,n}(\mathcal{F}) > 0$.

(v) For some $1 < p < \infty$, $\limsup_n b_{p,n}(\mathcal{F}) > 0$.

It is obvious that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v). To see why (iii) $\Rightarrow$ (iv), note that for $F : X \to Y \in \mathcal{F}$ and $f : 2^n \to X$, since $g^f_X(\varepsilon, d_i \varepsilon) \leq \operatorname{Lip}(f)/n$ for each $\varepsilon \in 2^n$ and $1 \leq i \leq n$,

$$Eg^f_{\mathcal{F}}(\varepsilon, -\varepsilon) \leq \left[ E\rho^f_{\mathcal{F}}(\varepsilon, -\varepsilon)^p \right]^{1/p} \leq b_{p,n}(\mathcal{F}) \left[ n^{p-1} \sum_{i=1}^n g^f_X(\varepsilon, d_i \varepsilon)^p \right]^{1/p} \leq b_{p,n}(\mathcal{F}) \operatorname{Lip}(f),$$

so $a_n(\mathcal{F}) \leq b_{p,n}(\mathcal{F})$. Thus the main part of this work is concerned with proving the implication (v) $\Rightarrow$ (i).

We note that if $l/2 \leq k \leq l$, then $b_{p,k}(\mathcal{F}) \leq 2^{l-1/p} b_{p,l}(\mathcal{F})$. Indeed, for $F : X \to Y \in \mathcal{F}$ and $f : 2^k \to X$, we can extend $f : 2^k \to X$ to a function $g : 2^l \to X$ by $g(\varepsilon) = f(\varepsilon(1), \ldots, \varepsilon(k))$. Then

$$E2^k g^f_{\mathcal{F}}(\varepsilon, -\varepsilon)^p = E2^l g^f_{\mathcal{F}}(\varepsilon, -\varepsilon)^p \leq b_{p,l}(\mathcal{F})^p l^{p-1} \sum_{i=1}^l E2^l g^f_X(\varepsilon, d_i \varepsilon)^p \leq 2^{p-1} b_{p,l}(\mathcal{F})^p k^{p-1} \sum_{i=1}^k E2^k g^f_X(\varepsilon, d_i \varepsilon)^p.$$

Applying this to $k \in \mathbb{N}$ and $l = 2^{\left\lfloor \log_2(k) \right\rfloor}$, we deduce that $\limsup_n b_{p,n}(\mathcal{F}) > 0$ if and only if $\limsup_n b_{p,2^n}(\mathcal{F}) > 0$. Thus our goal, completed in the fourth section of this work, will be to show that if for some $1 < p < \infty$,
lim supₙ bₚ₂ⁿ(F) > 0, then F factors the Hamming cubes. In the fifth section, we use concentration of measure to provide a quantitatively sharp proof that (iii) ⇒ (i).

We note that the definition of our quantities bₚₙ(F) are reminiscent of metric type as defined by Bourgain, Milman, and Wolfson in [3]. One may also ask about Enflo’s [6] definition of non-linear type. In the subtype regime, however, the two notions coincide. We give the details of this in the next section.

The author wishes to thank B. Randrianantoanina for making him aware of the coarse differentiation method of Eskin, Fisher, and Whyte.

2. Spatial versus operator results; Subtype

We first recall a result implicitly shown in [3] in the particular case p = 2. The general case 1 < p < ∞ follows by substituting their Fact 2.5 with our Lemma [3.1]

Theorem 2.1. [3 Theorem 2.6] For 1 < p < ∞, l ∈ N, and D > 1, there exists a constant 0 < a < 1 such that if (Z, dₙ) is a metric space and h : 2ⁿ → Z is a function such that

\[ \mathbb{E} d_{Z}(h(\varepsilon), h(-\varepsilon))^{p} > (1 - a)^{p-1} \sum_{i=1}^{l} \mathbb{E} d_{Z}(h(\varepsilon), h(d_{i}\varepsilon))^{p} \]

and if

\[ t^{p} = \frac{1}{l} \sum_{i=1}^{l} \mathbb{E} d_{Z}(h(\varepsilon), h(d_{i}\varepsilon))^{p}, \]

then for any ε₁, ε₂ ∈ 2ⁿ,

\[ \frac{1}{D} \partial(\varepsilon_{1}, \varepsilon_{2}) \leq d_{Z}(h(\varepsilon_{1}), h(\varepsilon_{2}))/t \leq D \partial(\varepsilon_{1}, \varepsilon_{2}). \]

With the preceding remarkable result, in the case that F is a collection of identity operators (that is, in the spatial case), it is easy to complete the main theorem. This is because the function n ↦ bₚₙ(F) is submultiplicative in the spatial case. From this it follows that either bₚₙ(F) = 1 (worst possible value) for all n ∈ N, in which case we immediately finish by Theorem [2.1] or bₚₙ(F) → 0. But this method does not apply to the map case because of the lack of submultiplicativity of n ↦ bₚₙ(F) in the non-spatial case.

More generally, one is often interested in a sequence of composition submultiplicative seminorms \( (T_{n})_{n=1}^{\infty} \) defined on the class of bounded, linear operators between Banach spaces (such as Rademacher or gaussian type \( p \) [10], Haar or margina type \( p \) [11], or asymptotic notions of Rademacher or basic type \( p \) [3]). By “composition submultiplicative,” we mean that for any pair of operators A, B such that the composition AB is defined, \( T_{mn}(AB) \leq T_{m}(A)T_{n}(B) \) for any natural numbers m, n. In the case that A = Iₓ, we can apply this fact with B = A = Iₓ to deduce that \( T_{mn}(I_{x}) = T_{mn}(I_{x}^{2}) \leq T_{m}(I_{x})T_{n}(I_{x}) \). The standard procedure in this case is to use these inequalities to prove that either \( (T_{n}(I_{x}))_{n=1}^{\infty} \) exhibits the quantitatively worst possible behavior for each n and use this to prove the presence of certain structures (such as \( \ell_{n}^{q} \) subspaces in the Rademacher case), or to prove that \( (T_{n}(I_{x}))_{n=1}^{\infty} \) is growing/shrinking rapidly enough to ensure some non-trivial power type behavior. This “automatic power type” phenomenon fails for all examples in the non-spatial case. One example, which is relevant to the subject of this work, is the diagonal operator \( F : \ell_{1} \rightarrow \ell_{1} \) given by \( F \sum_{n=1}^{\infty} a_{n}e_{n} = \sum_{n=1}^{\infty} \frac{a_{n}}{\log(n+1)} e_{n} \). This is compact, and cannot factor the Hamming cubes. But one can check that \( F \) has no non-trivial Rademacher type, and therefore no non-trivial non-linear type in the sense of Bourgain, Milman, and Wolfson. This is because for \( 1 < p < \infty \),

\[ \left( \mathbb{E} \| F \sum_{i=1}^{n} \varepsilon(i)e_{i} \|^{p} \right)^{1/p} \geq \frac{n}{\log(n+1)} \]

and

\[ \left( \sum_{i=1}^{n} \| e_{i} \|^{p} \right)^{1/p} = n^{1/p} = o(n/\log(n+1)). \]

More generally, we can choose any \( 1 < p < \infty \) and a sequence \( (w_{n})_{n=1}^{\infty} \) of positive numbers vanishing as slowly as we like and define the diagonal operator \( F : \ell_{1} \rightarrow \ell_{1} \) by \( F \sum_{n=1}^{\infty} a_{n}e_{n} = \sum_{n=1}^{\infty} a_{n}w_{n}e_{n} \). Then bₚₙ(F) is necessarily
vanishing, but as slowly as we like. Examples such as this motivate the search for a characterization of when the worst possible behavior does not hold (in our case, worst possible behavior means factoring the Hamming cube, while in other cases it is crude finite representability/asymptotic crude finite representability of the identity operator of $\ell_1^n$, non-super weak compactness, or non-asymptotic uniform smoothability). A technique in this case is to define a sequence of subtype constants of the map (or family of maps). One instance of this approach is due to Beauzamy [2], who gave a characterization of when the identity on $\ell_1$ is crudely finitely representable in a linear operator using a sequence of constants which are the linear analogues of our $a_n(F)$. Hinrichs proved a similar result using constants which are the linear analogues of our $b_{2,n}(F)$. In [4], asymptotic analogues of the results of Beauzamy and Hinrichs were proven for both the asymptotic linear analogues of the $a_n(F)$ and $b_{p,n}(F)$ constants.

The general approach to subtype problems is as follows: Suppose we have a sequence $(T_{n})_{n=1}^{\infty}$ as in the previous paragraph and positive numbers $(c_{n})_{n=1}^{\infty}$ such that for each $\lambda > 0$, $\lambda c_{n}$ is the supremum of $T_{n}(A)$ as $A$ ranges over all bounded, linear operators with $\|A\| \leq \lambda$. Then one may ask if, for a given class $A$ of operators with norms not more than $\lambda$, does $(\sup_{A \in A} T_{n}(A))_{n=1}^{\infty}$ exhibit the essentially worst possible behavior with respect to the sequence $(c_{n})_{n=1}^{\infty}$ (that is, $\limsup_{n} \sup_{A \in A} T_{n}(A)/c_{n} > 0$)? We then say $A$ has subtype if it does not exhibit the worst possible behavior (that is, $\lim_{n} \sup_{A \in A} T_{n}(A)/c_{n} = 0$). This has been applied when $T_{n}$ is the Rademacher/gaussian/Haar/martingale type $p$ norms. More generally, we may isolate non-linear subtype properties by replacing continuous, linear operators with Lipschitz functions, replacing the $(T_{n})_{n=1}^{\infty}$ sequence with a sequence $(\tau_{n})_{n=1}^{\infty}$ defined on the class of Lipschitz maps between metric spaces, and by replacing operator norm with Lipschitz constant. One then says that a class $F$ has subtype if $\lim_{n} \sup_{F \in F} \tau_{n}(F)/c_{n} = 0$. This is the approach we take.

Now for a family $F$ of $\lambda$-Lipschitz maps, let us define $e_{p,n}(F)$ to be the smallest constant $t > 0$ such that for any $F : X \to Y \in F$ and $f : 2^{n} \to X$,

$$E_{\Phi}^{F}(\varepsilon, -\varepsilon)^{p} \leq t^{p} \sum_{i=1}^{n} E_{\Phi}^{F}(\varepsilon, d_{i}\varepsilon)^{p}.$$  

Note that $e_{p,n}(F) \leq \lambda n^{1-1/p}$. Therefore with $c_{n} = n^{1-1/p}$, we can say $F$ has Enflo subtype if $\lim_{n} e_{p,n}(F)/c_{n} = 0$. But $e_{p,n}(F)/n^{1-1/p} = b_{p,n}(F)$. Therefore the subtype approach applied to Enflo type recovers the same condition as the Bourgain, Milman, Wolfson approach.

3. Rigidity Results

**Lemma 3.1.** For $1 < p < \infty$, $n \in \mathbb{N}$, and $\Phi > 1$, there exists $\phi = \phi(p, n, \Phi) \in (0, 1)$ such that if $a = (a_{i})_{i=1}^{n} \in \ell_{p}^{n}$ satisfies $\|a\|_{\ell_{p}^{n}}^{p} > \phi^{p-1}|\|a\|_{\ell_{p}^{n}}^{p}$, then $\max_{i} |a_{i}| \leq \Phi \min_{i} |a_{i}|$.

**Proof.** By the uniform convexity of $\ell_{p}^{n}$, there exists $0 < \delta < 1$ such that if $x, y \in B_{\ell_{p}^{n}}$ are such that $\|x + y\|_{\ell_{p}^{n}}^{p} > 2(1-\delta)$, then $\|x - y\|_{\ell_{p}^{n}}^{p} < \frac{2}{2^{n} \Phi^{1-\delta}}$. Let $\phi = (1-\delta)^{1/p}$. Suppose $a = (a_{i})_{i=1}^{n} \in \ell_{p}^{n}$ satisfies $\|a\|_{\ell_{p}^{n}}^{p} > \phi^{p-1}|\|a\|_{\ell_{p}^{n}}^{p}$. Without loss of generality, let us assume that $a_{i} = |a_{i}|$ for all $1 \leq i \leq n$. Let $x_{i} = a_{i}/\|a\|_{\ell_{p}^{n}}$ and $x = (x_{i})_{i=1}^{n} \in S_{\ell_{p}^{n}}$. Let $y_{i} = n^{-1/p}$ for $1 \leq i \leq n$ and $y = (y_{i})_{i=1}^{n} \in S_{\ell_{p}^{n}}$. Note that $\|x\|_{\ell_{p}^{n}}^{p} > \phi^{1/p} n^{1-1/p} = (1-\delta)n^{1-1/p}$ and $\|y\|_{\ell_{p}^{n}}^{p} = n^{1-1/p}$. Then $\|x + y\|_{\ell_{p}^{n}}^{p} \geq \|x + y\|_{\ell_{p}^{n}}^{1-1/p} = \|x\|_{\ell_{p}^{n}}^{p} + \|y\|_{\ell_{p}^{n}}^{p} \geq (1-\delta) + 1 > 2(1-\delta)$. Therefore $\|x - y\|_{\ell_{p}^{n}}^{p} \leq \frac{1}{2n^{1/p}} \cdot \frac{\Phi - 1}{\Phi}$. Since $\max_{i} x_{i} \geq 1/n^{1/p}$, we deduce that

$$\max_{i} x_{i} - \min_{i} x_{i} \leq |n^{-1/p} - \max_{i} x_{i}| + |n^{-1/p} - \min_{i} x_{i}| \leq 2\|x - y\|_{\ell_{p}^{n}}^{p} \leq \frac{1}{n^{1/p}} \cdot \frac{\Phi - 1}{\Phi} \leq \left(\frac{\Phi - 1}{\Phi}\right) \max_{i} x_{i}.$$  

Rearranging yields that $\max_{i} x_{i} \leq \Phi \min_{i} x_{i}$, and we deduce the result by homogeneity.

**\square**

**Lemma 3.2.** Fix $1 < p < \infty$. Let $\Omega$ be a probability space and let $D_{V}, D_{X}, E_{V}, E_{X} : \Omega \to \mathbb{R}$, $\lambda, \Theta > 0$, $a, b, v, \mu \in (0, 1)$ be such that
(i) \( D_Y, D_X, E_Y, E_X \) are non-negative, measurable functions on \( \Omega \) such that \( D_Y \leq E_Y, D_X \leq E_X, D_Y \leq \lambda^p D_X, E_Y \leq \lambda^p E_X, D_Y \leq (1 + \nu) \Theta^p E_X, \)

(ii) \( \mathbb{E} D_Y > (1 - \nu) \Theta^p \mathbb{E} E_X, \mathbb{E} D_Y > (1 - \mu) \mathbb{E} E_Y, \mathbb{E} D_X > (1 - \mu) \mathbb{E} E_X, \)

(iii) \( \lambda^p \left( \frac{2 \nu}{a} + \frac{2 \nu}{b} \right) < \Theta^p (1 - \nu). \)

Then there exists \( \varepsilon \in \Omega \) such that \( D_Y(\varepsilon) > (1 - \alpha) E_Y(\varepsilon), D_X(\varepsilon) > (1 - \alpha) E_X(\varepsilon), \text{ and } D_Y(\varepsilon) > (1 - b) \Theta^p E_X(\varepsilon). \)

Proof. Let \( A_Y = (D_Y \leq (1 - a) E_Y), A_X = (D_X \leq (1 - a) E_X), \text{ and } B = (D_Y \leq (1 - b) \Theta^p E_X). \) Then the conclusion of the lemma is equivalent to \( A_Y \cap A_X \cap B^c \neq \emptyset. \) We work by contradiction. Assume \( A_Y \cap A_X \cap B^c = \emptyset, \) so \( \Omega = A_Y \cup A_X \cup B. \) Let us first note that

\[
(1 - \mu) \mathbb{E} E_Y < \mathbb{E} D_Y = \mathbb{E} 1_{A_Y} D_Y + \mathbb{E} 1_{A_X} D_Y \leq (1 - a) \mathbb{E} 1_{A_Y} E_Y + \mathbb{E} 1_{A_X} E_Y = \mathbb{E} E_Y - a \mathbb{E} 1_{A_Y} E_Y.
\]

From this it follows that

\[
\mathbb{E} 1_{A_Y} E_Y \leq \frac{\mu}{a} \mathbb{E} E_Y.
\]

By replacing each \( Y \) with \( X, \) we deduce that

\[
\mathbb{E} 1_{A_X} E_X \leq \frac{\mu}{a} \mathbb{E} E_X.
\]

Also,

\[
\Theta^p (1 - \nu) \mathbb{E} E_X < \mathbb{E} D_Y = \mathbb{E} 1_B D_Y + \mathbb{E} 1_{B^c} D_Y \leq \Theta^p (1 - b) \mathbb{E} 1_B E_X + \Theta^p (1 + \nu) \mathbb{E} 1_{B^c} E_X.
\]

Dividing by \( \Theta^p \) and rearranging yields that

\[
\mathbb{E} 1_B E_X \leq \frac{2 \nu}{b} \mathbb{E} E_X.
\]

Recalling that \( A_Y \cup A_X \cup B = \Omega, \) we deduce that

\[
\Theta^p (1 - \nu) \mathbb{E} E_X < \mathbb{E} D_Y \leq \mathbb{E} 1_{A_Y} D_Y + \mathbb{E} 1_{A_X} D_Y + \mathbb{E} 1_B D_Y
\]

\[
\leq \mathbb{E} 1_{A_Y} E_Y + \lambda^p \mathbb{E} 1_{A_X} E_X + \lambda^p \mathbb{E} 1_B E_X
\]

\[
\leq \frac{\mu}{a} \mathbb{E} E_Y + \frac{\mu \lambda^p}{a} \mathbb{E} E_X + \frac{2 \nu \lambda^p}{b} \mathbb{E} E_X
\]

\[
\leq \lambda^p \left( \frac{2 \mu}{a} + \frac{2 \nu}{b} \right) \mathbb{E} E_X.
\]

Since \( \mathbb{E} E_X > 0, \) this contradicts (iii) and finishes the proof.

\[\square\]

For a natural number \( n, \) we let \([n] = \{1, \ldots, n\}\) denote the integer interval. Fix natural numbers \( l_1, \ldots, l_d \) and let \( L = \prod_{j=1}^d l_j. \) We define \( T = \bigcup_{i=0}^d \Lambda_i \) as follows. We let \( \Lambda_0 = \{[L]\} \) consist of a single integer interval. Now suppose that for \( i < d, \) \( \Lambda_i \) has been defined and consists of pairwise disjoint subintervals of \([L]\) each of which has cardinality \( \prod_{j=i+1}^d l_j. \) For each \( I \in \Lambda_i, \) let \( I_I = \{J_1, \ldots, J_{l_{i+1}}\} \) be a partition of \( I \) into subintervals of equal cardinality (and therefore of cardinality \( \prod_{j=i+1}^d l_j. \)) Now let \( \Lambda_{i+1} = \bigcup_{I \in \Lambda_i} I_I. \) This completes the recursive definition of \( \Lambda_0, \ldots, \Lambda_d. \) Now let \( T = \bigcup_{i=0}^d \Lambda_i. \) We refer to \( T \) as the \((l_1, \ldots, l_d)\) interval tree. For \( 0 < j \leq d \) and \( J \in \Lambda_j, \) let \( J^- \) be the member \( I \in \Lambda_{j-1} \) such that \( J \subset I. \) That is, \( J^- \) is the interval \( I \in \Lambda_j \) such that \( J \in I_I. \)

Remark 3.3. Suppose \( l_1, \ldots, l_{d+1} \) are natural numbers and \( T \) is the \((l_1, \ldots, l_{d+1})\) interval tree. Suppose that \((t_I)_{I \in T}\) is a collection of non-negative numbers such that for each \( 0 \leq j \leq d \) and \( I \in \Lambda_j, \) \( t_I \leq \sum_{J^-=-I} t_J. \) Then using this fact repeatedly yields that for any \( 0 \leq i < j \leq d + 1 \) and \( I \in \Lambda_i, \)

\[
t_I \leq \sum_{J \supseteq I} t_J.
\]

Also, by Hölder’s inequality, it follows that for any such \( j \) and \( I, \)

\[
t_I^p \leq \left( \sum_{J \supseteq I} t_J^p \right)^{\frac{p}{p-1}} \sum_{J \supseteq I} t_J^p,
\]
and more generally,
\[
t_I^p \leq \left( \prod_{m=i+1}^{d} l_m^{p-1} \right) \sum_{I \supset J \in \Lambda_j} t_J^p
\]
for any \(0 \leq i < j \leq d+1\) and \(I \in \Lambda_i\). We will use this fact frequently in this section.

**Lemma 3.4.** Fix \(1 < p < \infty\). Fix natural numbers \(l_1, \ldots, l_d, 0 < \mu < 1, \lambda, \Theta > 0\), and \(M > \lambda/\Theta\). Then for any \(0 < \eta_1 < 1\), there exists \(0 < \eta < \eta_1\) with the following property: Suppose \(l_{d+1}\) is a natural number, \(T\) is the \((l_1, \ldots, l_{d+1})\) interval tree, and \((r_I)_{I \in T}, (s_I)_{I \in T}\) are non-negative numbers such that

(i) for each \(I \in T\), \(r_I \leq \lambda s_I\),

(ii) for each \(I \in T \setminus \Lambda_{d+1}\), \(r_I \leq \prod_{J \supset I} r_J\) and \(s_I \leq \prod_{J \supset I} s_J\),

(iii) for each \(I \in \Lambda_d\), \(r_I^p \leq (1 + \eta)\Theta^{p(I_i^p-1)} \prod_{J \supset I} s_J^p\),

(iv) \(r_{[I]} > (1 - \eta)\Theta^{\prod_{i=1}^{d+1} l_i^{p-1}} \sum_{J \supset I} s_J^p\).

Then for any \(0 \leq j \leq d\), \(0 \leq i < d\), and \(I_1 \in \Lambda_j\), \(\max_{I \in \Lambda_j} s_I \leq \min_{I \in \Lambda_j} r_I\) and \(r_I^p > (1 - \mu)\Theta^{\prod_{i=1}^{d+1} l_i^{p-1}} \sum_{J \supset I} s_J^p\).

**Proof.** First fix \(\Phi > 1\) such that \(M > \Phi^3 \lambda/\Theta\). Now let \(0 < \phi < 1\) be such that for any \(1 \leq n \leq \prod_{i=1}^{d} l_i\) and any \(v = (v_i)_{i=1}^{n} \in \ell_p^n\) with \(\max_i |v_i| > \Phi \min_i |v_i|\), \(|v|^p_{\ell_p^n} < \phi n^{p-1}|v|^p_{\ell_p^n}\). Such a \(\phi\) exists by Lemma 3.1. Now fix \(0 < \eta < \eta_1\) so small that

(a) \(\phi(1 + \mu) < 1 - \eta\),

(b) \(\frac{\phi^p}{1 + \eta} \leq \left( \frac{1}{1 - \eta} - \frac{1}{1 + \eta} \right) \left( \prod_{i=1}^{d} l_i \right) \Phi^p\),

(c) \(M > \Phi^{3 \lambda/\Theta}\),

(d) \((1 - \frac{\phi^p}{1 + \eta})^{1 - \eta} (1 + \eta) < (1 - \eta)\).

Now suppose that \(l_{d+1} = (r_I)_{I \in T}, (s_I)_{I \in T}\) are as in the lemma.

Step 1: For any \(0 \leq j \leq d\), \(\max_I r_I \leq \Phi \min_I r_I\).

If it were not so, then by the choice of \(\phi\) applied to the vector \((r_I)_{I \in \Lambda_j} \in \ell_p(\Lambda_j)\),

\[
\left( \sum_{I \in \Lambda_j} r_I \right)^p \leq \phi \left( \prod_{i=1}^{j} l_i^{p-1} \right) \sum_{I \in \Lambda_j} r_I^p.
\]

Then

\[
r_I^p \leq \left( \sum_{I \in \Lambda_j} r_I \right)^p \leq \phi \left( \prod_{i=1}^{j} l_i^{p-1} \right) \sum_{I \in \Lambda_j} r_I^p \leq \phi \left( \prod_{i=1}^{d} l_i^{p-1} \right) \sum_{I \in \Lambda_d} r_I^p.
\]

But since \(r_I^p > (1 - \eta)\Theta^{\prod_{i=1}^{d+1} l_i^{p-1}} \sum_{J \supset I} s_J^p\), we contradict item (a) of our choice of \(\eta\). This completes Step 1.

Step 2: For any \(0 \leq j \leq d\),

\[
\max_{I \in \Lambda_j} \sum_{J \supset I} s_J^p \leq \Phi^{2p} \min_{I \in \Lambda_j} \sum_{J \supset I} s_J^p.
\]

If it were not so, we could find \(0 \leq j \leq d\) and \(I_1, I_2 \in \Lambda_j\) such that

\[
\frac{1}{\Phi^{2p}} \sum_{I \supset J \in \Lambda_{d+1}} s_J^p > \sum_{I_1 \supset J \in \Lambda_{d+1}} s_J^p.
\]
Then by Step 1,
\[
r_i^p \leq \Phi^p r_{i+1}^p \leq \Phi^p \left( \prod_{i=j+1}^{d} t_i^{p-1} \right) \sum_{I \supset J \in \Lambda_d} r_j^p \leq \Phi^p (1 + \eta) \Theta^p \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \supset J \in \Lambda_{d+1}} s_j^p.
\]

Note that for any \( I \in \Lambda_j \),
\[
r_j^p \leq \left( \prod_{i=j+1}^{d} t_i^{p-1} \right) \sum_{I \supset J \in \Lambda_d} r_j^p \leq (1 + \eta) \Theta^p \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_{d+1}} s_j^p.
\]

Since
\[
(1 - \eta) \Theta^p \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_{d+1}} s_j^p \leq r_j^p \left( \prod_{i=1}^{j} t_i^{p-1} \right) \sum_{I \in \Lambda_j} r_j^p,
\]
we see that
\[
\frac{1}{(1 + \eta) \Theta^p} \left[ \Phi^p r_i^p + \sum_{I \in \Lambda_j \setminus \Lambda_j} r_j^p \right] \leq \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_{d+1}} s_j^p \leq \frac{1}{(1 - \eta) \Theta^p} \sum_{I \in \Lambda_j} r_j^p.
\]

Manipulating the first and last terms of this inequality, we deduce that
\[
\left( \frac{\Phi^p}{1 + \eta} - 1 \right) r_i^p \leq \left( \frac{1 - \eta}{1 - \eta} - 1 \right) \sum_{I \in \Lambda_{d+1}} r_i^p \leq \left( \frac{1 - \eta}{1 - \eta} - 1 \right) \Phi^p|\Lambda_j| r_i^p,
\]
\[
\leq \left( \frac{1 - \eta}{1 - \eta} - 1 \right) \Phi^p \left( \prod_{i=1}^{d} t_i \right) r_i^p.
\]

Since \( r_i^p > 0 \), we reach a contradiction of (b) of our choice of \( \eta \).

Step 3: For any \( 0 \leq j \leq d \), \( \max_{I \in \Lambda_j} s_I \leq M \min_{I \in \Lambda_j} s_I \). Fix such a \( j \) and let
\[
R = \max_{I \in \Lambda_j} r_I \quad S = \max_{I \in \Lambda_j} s_I \quad S_1 = \max_{I \in \Lambda_j} \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \supset J \in \Lambda_{d+1}} s_j^p
\]
and
\[
r = \min_{I \in \Lambda_j} r_I \quad s = \min_{I \in \Lambda_j} s_I \quad s_1 = \min_{I \in \Lambda_j} \left( \prod_{i=j+1}^{d+1} t_i^{p-1} \right) \sum_{I \supset J \in \Lambda_{d+1}} s_j^p.
\]

We know from Step 1 that \( R \leq \Phi r \). We know from Step 2 that \( S_1 \leq \Phi^2 s_1 \). We know from hypothesis that \( R \leq \lambda S \), and we know from Remark 3.3 that \( S^p \leq S_1 \). Moreover,
\[
\left( \prod_{i=1}^{j} t_i^{p-1} \right) s_1 \lambda_j \leq \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_j} \sum_{I \supset J \in \Lambda_{d+1}} s_j^p = \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_{d+1}} s_j^p < \left[ (1 - \eta) \Theta^p \right]^{-1} r_j^p \left( \prod_{i=1}^{j} t_i^{p-1} \right)
\]
\[
\leq \left[ (1 - \eta) \Theta^p \right]^{-1} \Phi^p \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) \sum_{I \in \Lambda_j} r_j^p \leq \left[ (1 - \eta) \Theta^p \right]^{-1} \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) R \lambda_j
\]
\[
\leq \left[ (1 - \eta) \Theta^p \right]^{-1} \Phi^p \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) r_j^p \lambda_j \leq \left[ (1 - \eta) \Theta^p \right]^{-1} \Phi^p \lambda^p \left( \prod_{i=1}^{d+1} t_i^{p-1} \right) s_j \lambda_j.
\]

Therefore \( s_1 \leq \left[ (1 - \eta) \Theta^p \right]^{-1} \Phi^p \lambda^p s_1 \). Also, \( S^p \leq S_1 \leq \Phi^2 s_1 \), so
\[
S^p \leq \Phi^2 s_1 \leq \left[ (1 - \eta) \Theta^p \right]^{-1} \Phi^p \lambda^p s_1.
\]
Taking \( p^{th} \) roots and appealing to (c) finishes Step 3.
Step 4: For any $0 \leq j < d$ and $I \in \Lambda_j$, $r^p_I > (1 - \mu)p^{-1}_{j+1} \sum_{j-I} r^p_J$. If it were not so, then for some $I_1 \in \Lambda_j$, $r^p_{I_1} \leq (1 - \mu)p^{-1}_{j+1} \sum_{j-I_1} r^p_J$. Let us note that

$$\sum_{j-I_1} r^p_j \geq \frac{1}{\Phi p_{|\Lambda_j+1|}} \sum_{j \in \Lambda_{j+1}} r^p_j \geq \frac{1}{\Phi p \prod_{i=1}^d l_i} \sum_{j \in \Lambda_{j+1}} r^p_j,$$

so

$$\sum_{I \in \Lambda_j} r^p_I \leq (1 - \mu)p^{-1}_{j+1} \sum_{j-I_1} r^p_j + \frac{P}{1 - \mu} \sum_{I \in \Lambda_j \setminus \Lambda_{j+1}} r^p_I \leq (1 - \mu)p^{-1}_{j+1} \sum_{I \in \Lambda_{j+1}} r^p_I \leq (1 - \mu)p^{-1}_{j+1} \mu \sum_{I \in \Lambda_{j+1}} r^p_I \leq \left(1 - \frac{\mu}{\Phi p \prod_{i=1}^d l_i} \right) \sum_{I \in \Lambda_{j+1}} r^p_I.$$

Then

$$r^p_{I_1} \leq \left(\prod_{i=1}^d \frac{1}{l_i} \right) \sum_{I \in \Lambda_j} r^p_I \leq \left(1 - \frac{\mu}{\Phi p \prod_{i=1}^d l_i} \right) \left(\prod_{i=1}^d \frac{1}{l_i} \right) \sum_{I \in \Lambda_{j+1}} r^p_I \leq \left(1 - \frac{\mu}{\Phi p \prod_{i=1}^d l_i} \right) \left(1 + \eta \right) \Theta \left(\prod_{i=1}^{d+1} \frac{1}{l_i} \right) \sum_{I \in \Lambda_{d+1}} s^p_I,$$

But since $r^p_{I_1} > (1 - \eta)\Theta \left(\prod_{i=1}^{d+1} \frac{1}{l_i} \right) \sum_{I \in \Lambda_{d+1}} s^p_I$, this contradicts (d) and finishes the proof.

The following result is similar in spirit to the coarse differentiation result of Eskin, Fisher, and Whyte [7].

**Lemma 3.5.** Fix $1 < p < \infty$. Fix natural numbers $(l_1, \ldots, l_{d+1})$ and let $T$ be the $(l_1, \ldots, l_{d+1})$ interval tree. Suppose $0 < \Delta, \mu < 1$, $M > 1$, $m \in \mathbb{N}$, $\lambda > 0$, and $(s_I)_{I \in T} \subset (0, \infty)$ are such that

(i) for each $I \in T \setminus \Lambda_{d+1}$, $s_I \leq \sum_{j-I} s_J$,

(ii) for all $0 \leq j < d$, $\max_{I \in \Lambda_j} s_I \leq M \min_{I \in \Lambda_j} s_I$,

(iii) $s^p_{|I|} > (1 - \nu/2) \Theta (\prod_{i=1}^{d+1} \frac{1}{l_i} \prod_{j-I} s^p_J \prod_{j-I} s^p_J)$,

(iv) $(1 - \mu \Delta/M^p)^m < (1 - \nu/2)^{\Theta}$.

For each $0 \leq j < d$, let $I_j = \{I \in \Lambda_j : s^p_I \leq (1 - \mu)p^{-1}_{j+1} \sum_{j-I} s^p_J\}$ and let $B = \{j < d : |I_j| \geq \Delta/|\Lambda_j|\}$. Then $|B| \leq m$.

**Proof.** First suppose $j \in B$. Let $s = \min_{I \in \Lambda_{j+1}} s_I$ and $S = \max_{I \in \Lambda_{j+1}} s_I \leq s M$. Let $A = \{j \in \Lambda_{j+1} : J^- \in I_j\}$ and note that $|A/|\Lambda_{j+1}| = |I_j/|\Lambda_j| \geq \Delta$. Then

$$\sum_{I \in \Lambda_j} s^p_I = \sum_{I \in I_j} s^p_I + \sum_{I \in \Lambda_j \setminus I_j} s^p_I \leq (1 - \mu)p^{-1}_{j+1} \sum_{I \in A} s^p_J + \sum_{I \in \Lambda_{j+1}} s^p_J \leq (1 - \mu)p^{-1}_{j+1} \sum_{I \in \Lambda_{j+1}} s^p_J - \mu p^{-1}_{j+1} s^p_J |A| \leq (1 - \mu)p^{-1}_{j+1} \sum_{I \in \Lambda_{j+1}} s^p_I - \mu p^{-1}_{j+1} \frac{M^p s^p_J}{M^p} |A| \leq \mu p^{-1}_{j+1} \Delta |\Lambda_{j+1}| \leq \mu p^{-1}_{j+1} \sum_{I \in \Lambda_{j+1}} s^p_I \leq (1 - \mu \Delta/M^p)p^{-1}_{j+1} \sum_{I \in \Lambda_{j+1}} s^p_I.$$


Now suppose that \(|B| > m\) and fix 0 ≤ \(j_0 < \ldots < j_m\), \(j_i \in B\) for each 0 ≤ \(i \leq m\). Then
\[
\sum_{I \in \Lambda_{j_0}} s^p_I \leq \left(1 - \frac{\mu \Delta}{M^p}\right) \sum_{I \in \Lambda_{j_0+1}} s^p_I \\
\leq \left(1 - \frac{\mu \Delta}{M^p}\right) \left(\prod_{i=1}^{j_0} \frac{l_i^{p-1}}{l_{i-1}^{p-1}}\right) \sum_{I \in \Lambda_{j_1}} s^p_I \\
\leq \left(1 - \frac{\mu \Delta}{M^p}\right) \left(\prod_{i=1}^{j_1} \frac{l_i^{p-1}}{l_{i-1}^{p-1}}\right) \sum_{I \in \Lambda_{j_2}} s^p_I \\
\leq \ldots \leq \left(1 - \frac{\mu \Delta}{M^p}\right)^m \left(\prod_{i=1}^{j_m+1} \frac{l_i^{p-1}}{l_{i-1}^{p-1}}\right) \sum_{I \in \Lambda_{j_m+1}} s^p_I \\
\leq \left(1 - \frac{\mu \Delta}{M^p}\right)^m \left(\prod_{i=1}^{d+1} \frac{l_i^{p-1}}{l_{i-1}^{p-1}}\right) \sum_{I \in \Lambda_{d+1}} s^p_I.
\]
This contradicts \((iii)\) and \((iv)\) and finishes the proof.

**Lemma 3.6.** Fix 1 < \(p < \infty\). Fix natural numbers \((l_1, \ldots, l_{d+1})\) and let \(T\) be the \((l_1, \ldots, l_{d+1})\) interval tree. Suppose \(0 < \Delta, \nu < 1, \lambda > 0, M > 1\), and \((r_I)_{I \in T}, (s_I)_{I \in T} \subset [0, \infty)\) are such that

(i) for each \(I \in T\), \(r_I \leq \lambda s_I\),
(ii) for each \(I \in T \setminus \Lambda_{d+1}\), \(r_I \leq \sum_{J \subseteq I} r_J, s_I \leq \sum_{J \subseteq I} s_J\),
(iii) for all 0 ≤ \(j \leq d\), \(\max_{I \in \Lambda_j} s_I \leq M \min_{I \in \Lambda_j} s_I\),
(iv) \(\Delta M^p \leq \nu \Theta^p / 2 \lambda^p\),
(v) \(r^p_{[L]} > (1 - \nu / 2) \Theta^p \left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \sum_{I \in \Lambda_{d+1}} s^p_I\).

Then for any 0 ≤ \(j \leq d\), \(|\{I \in \Lambda_j : r^p_I \leq (1 - \nu) \Theta^p \sum_{J \subseteq I} s^p_J\}| < (1 - \Delta)|\Lambda_j|\).

**Proof.** Suppose not. Fix 0 ≤ \(j < d\) such that \(B = \{I \in \Lambda_j : r^p_I \leq (1 - \nu) \Theta^p \sum_{J \subseteq I} s^p_J\}\) has cardinality at least \((1 - \Delta)|\Lambda_j|\). Let \(s = \min_{I \in \Lambda_{j+1}} s_I\) and \(S = \max_{I \in \Lambda_{j+1}} s_I \leq sM\). Let \(A = \{J \in \Lambda_{j+1} : J \subseteq B\}\) and note that \(|A|/|\Lambda_{j+1}| = |B|/|\Lambda_j| \geq 1 - \Delta\). Now
\[
\sum_{J \in \Lambda_{j+1} \setminus A} s^p_J \leq |\Lambda_{j+1} \setminus A| S^p \leq \Delta |\Lambda_{j+1}| S^p \leq \Delta M^p \sum_{I \in \Lambda_{j+1}} s^p_I \leq \frac{\nu \Theta^p}{2 \lambda^p} \sum_{I \in \Lambda_{j+1}} s^p_I.
\]
Then
\[
\sum_{I \in \Lambda_j} r^p_I \leq \left(\sum_{i=1}^{j+1} l_i^{p-1}\right) \sum_{I \in \Lambda_j} \left[\sum_{I \subseteq B} r^p_I + \sum_{I \in \Lambda_j \setminus B} r^p_I\right]
\leq \left(\sum_{i=1}^{j+1} l_i^{p-1}\right) \left[(1 - \nu) \Theta^p \sum_{I \subseteq B} s^p_I + \lambda^p \sum_{I \in \Lambda_j \setminus B} s^p_I\right]
\leq \left(\sum_{i=1}^{j+1} l_i^{p-1}\right) \left[(1 - \nu) \Theta^p \sum_{I \subseteq A} s^p_I + \lambda^p \sum_{I \in \Lambda_{j+1} \setminus A} s^p_I\right]
\leq (1 - \nu / 2) \Theta^p \left(\sum_{i=1}^{j+1} l_i^{p-1}\right) \sum_{I \in \Lambda_{j+1}} s^p_I \\
\leq (1 - \nu / 2) \Theta^p \left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \sum_{I \in \Lambda_{d+1}} s^p_I.
\]
This contradiction finishes the proof.
4. Proof of $(v) \Rightarrow (i)$

Fix $1 < p < \infty$. Let $\lambda = \sup_{F \subseteq X} \text{Lip}(F) \in (0, \infty)$ and let $\Theta = \lim \sup_n b_{p,2^n}(F) \in (0, \lambda]$. Fix $0 < \vartheta < \Theta$.

Fix $0 < b < 1$ such that $(1 - b)^{1/p} \Theta > \vartheta$. Next fix $0 < \nu < b$ such that $\nu/b < (1 - \nu)\Theta^p/4\lambda^p$. Fix $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $b_{p,2^n}(F) \leq (1 + \nu)^{1/p} \Theta$. Fix $n \geq n_0$ and $D > 1$, and let $l = 2^n$. By Theorem 2.1 there exists $0 < a < 1$ such that if $(Z,d_Z)$ is any metric space and $h : 2^l \to Z$ satisfies

$$E d_Z(h(\varepsilon), h(-\varepsilon))^p > (1 - a)^{p-1} \sum_{i=1}^l d_Z(h(\varepsilon), h(d_i \varepsilon))^p$$

and if

$$\nu^p = \frac{1}{l} \sum_{i=1}^l d_Z(h(\varepsilon), h(d_i \varepsilon))^p,$$

then for any $\varepsilon_1, \varepsilon_2 \in 2^l$,

$$\frac{1}{D} \partial(\varepsilon_1, \varepsilon_2) \leq \frac{d_Z(h(\varepsilon_1), h(\varepsilon_2))}{l} \leq D \partial(\varepsilon_1, \varepsilon_2).$$

Now fix $0 < \mu < a$ such that $\mu/a < (1 - \nu)\Theta^p/4\lambda^p$. Fix $M > \lambda/\Theta$. Fix $0 < \Delta < 1$ such that $\Delta M^p < \nu \Theta^p/2\lambda^p$. Fix $m \in \mathbb{N}$ such that

$$\left(1 - \frac{\Delta \mu}{M^p}\right)^m < (1 - \nu/2) \Theta^p.$$

Fix $d > m+1$ and let $l_1 = \ldots = l_d = l$. Fix $0 < \eta < \nu$ according to Lemma 3.4 with all of these choices of parameters. Fix $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, $b_{p,2^n}(F) \leq (1 + \eta)^{1/p} \Theta$. Fix $n_2 > n_1 + nd$ such that $b_{p,2^{n_2}}(F) > (1 - \eta/2)^{1/p} \Theta$. Let $l_{d+1} = 2^{n_2-nd} > 2^{n_1}$ and let $L = \prod_{i=1}^{d+1} l_i = 2^{n_2}$. Let $T$ be the $(l_1, \ldots, l_{d+1})$ interval tree. Fix $F : X \to Y \in F$ and $f : 2^L \to X$ such that

$$E \Theta^f_\varepsilon(\varepsilon, -\varepsilon)^p > (1 - \eta/2) \Theta^p \left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \sum_{i=1}^{L} E \Theta^X_{d_i \varepsilon}(\varepsilon, d_i \varepsilon)^p.$$ 

For an interval $I \in T$ and $\varepsilon \in 2^L$, let $I \varepsilon \in 2^L$ be the member of $2^L$ given by

$$I \varepsilon(i) = \begin{cases} \varepsilon(i) & : i \in [L] \setminus I \\ -\varepsilon(i) & : i \in I. \end{cases}$$

For each $I \in T$, let

$$r_I = \left[E \Theta^f_\varepsilon(\varepsilon, I \varepsilon)^p\right]^{1/p}$$

and

$$s_I = \left[E \Theta^X_{d \varepsilon}(\varepsilon, d \varepsilon)^p\right]^{1/p}.$$

Claim 1. (i) $r_{[L]}^p > (1 - \eta/2) \left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \Theta^p \sum_{I \in \Lambda_{d+1}} s_I^p$.

(ii) For any $I \in T$, $r_I \leq \lambda s_I$.

(iii) For any $0 \leq j \leq d$ and $I \in \Lambda_j$, $r_I \leq \sum_{J=-1}^{J=1} r_J$ and $s_I \leq \sum_{J=-1}^{J=1} s_J$.

(iv) For all $I \in T$, $r_I^{p_2} \leq (1 + \nu) \Theta^{p_2} \sum_{J=-1}^{J=1} s_J^p$.

(v) For any $I \in \Lambda_d$, $r_I^{p_2} \leq (1 + \eta) l_{d+1}^{p_2-1} \Theta^p \sum_{J=-1}^{J=1} s_J^p$.

Proof. (i) This follows from our choice of $F$, $f$, and the fact that

$$r_{[L]}^p = E \Theta^f_\varepsilon(\varepsilon, -\varepsilon)^p$$

and

$$\left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \sum_{J \in \Lambda_{d+1}} s_J^p = \left(\prod_{i=1}^{d+1} l_i^{p-1}\right) \sum_{i=1}^{L} E \Theta^X_{d \varepsilon}(\varepsilon, d \varepsilon)^p.$$

(ii) This follows from the fact that $\text{Lip}(F) \leq \lambda$. 

(iii) Fix $0 \leq j \leq d$ and $I \in \Lambda_j$. Enumerate $\{J \in T : J^+ = I\}$ as $(I_l)_{l=1}^{l_{j+1}}$. For $i = 0, \ldots, l_{j+1}$, define $J_l : 2^L \to 2^L$ by letting $J_0$ be the identity and $J_i = I_i J_{i-1}$. For any $\varepsilon \in 2^L$, $J_0 \varepsilon = \varepsilon$ and $J_{l_{j+1}} \varepsilon = I \varepsilon$. Then for any $\varepsilon \in 2^L$,

$$g_Y^l(\varepsilon, I \varepsilon) \leq \sum_{i=1}^{l_{j+1}} g_Y^l(J_{i-1} \varepsilon, J_i \varepsilon).$$

Now the triangle inequality on $L_p(2^L)$ yields that

$$r_I \leq \sum_{i=1}^{l_{j+1}} \left[ \mathbb{E} g_Y^l(J_{i-1} \varepsilon, J_i \varepsilon)^p \right]^{1/p}.$$

But $g_Y^l(J_{i-1} \varepsilon, J_i \varepsilon)$ and $g_Y^l(\varepsilon, I \varepsilon)$ have the same distribution, so

$$r_I \leq \sum_{i=1}^{l_{j+1}} \left[ \mathbb{E} g_Y^l(J_{i-1} \varepsilon, J_i \varepsilon)^p \right]^{1/p} = \sum_{J=-I}^{l_{j+1}} r_J.$$

Replacing each $Y$ with $X$ yields that $s_I \leq \sum_{J=-I} s_J$.

(iv) and (v) Let $I$ and $(I_l)_{l=1}^{l_{j+1}}$ be as in the proof of (iii). Define $g : 2^L \times 2^{l_{j+1}} \to 2^L$ by letting

$$g(\varepsilon, \delta)(i) = \begin{cases} \varepsilon(i) & : i \in [L] \setminus I \\ \delta(k) \varepsilon(i) & : i \in I_k. \end{cases}$$

For fixed $\varepsilon \in 2^L$, let $f_\varepsilon : 2^{l_{j+1}} \to X$ be given by $f_\varepsilon(\delta) = f(g(\varepsilon, \delta))$. Note that $g(\varepsilon, -\delta) = I g(\varepsilon, \delta)$ and $g(\varepsilon, d_i \delta) = I_i g(\varepsilon, \delta)$ for all $\varepsilon \in 2^L$, $\delta \in 2^{l_{j+1}}$, and $1 \leq i \leq l_{j+1}$. From this it follows that $f_\varepsilon(-\delta) = f(I g(\varepsilon, \delta))$ and for $1 \leq i \leq l_{j+1}$, $f_\varepsilon(d_i \varepsilon) = f(I_i g(\varepsilon, \delta))$. Then

$$r_I^p = \mathbb{E} g_Y^l(\varepsilon, I \varepsilon)^p = \mathbb{E}_\varepsilon \mathbb{E}_\delta g_Y^l(g(\varepsilon, \delta), I g(\varepsilon, \delta))^p$$

$$= \mathbb{E}_\varepsilon \mathbb{E}_\delta g_Y^l(\varepsilon, \delta)^p \leq \mathbb{E}_\varepsilon b_{p,l_{j+1}}(F)^p l_{j+1}^{-1} \sum_{i=1}^{l_{j+1}} \mathbb{E}_\delta g_X^l(\delta, d_i \delta)^p$$

$$= \mathbb{E}_\varepsilon b_{p,l_{j+1}}(F)^p l_{j+1}^{-1} \sum_{i=1}^{l_{j+1}} \mathbb{E}_\varepsilon \mathbb{E}_\delta g_X^l(g(\varepsilon, \delta), I_i g(\varepsilon, \delta))^p$$

$$= b_{p,l_{j+1}}(F)^p l_{j+1}^{-1} \sum_{i=1}^{l_{j+1}} \mathbb{E}_\varepsilon \mathbb{E}_\delta g_X^l(\varepsilon, \delta)^p = b_{p,l_{j+1}}(F)^p l_{j+1}^{-1} \sum_{J=-I} s_J^p.$$
By Lemma 3.5, there are at most $m$ values of $j < d$ such that $|A_j| \geq \Delta |A_j|$. Here we note that
\[
 s^p_{[L]} \geq \frac{r^p_{[L]}}{\lambda^p} > (1 - \eta/2) \Theta^p \left( \prod_{i=1}^{d+1} p_i^{p-1} \right) \sum_{l \in \Lambda_{d+1}} s^p_l.
\]
Since $m + 1 < d$, there exists at least one value $j_0 < d$ such that $|A_{j_0}| < \Delta |A_{j_0}|$. By Lemma 3.6 for this $j_0$, $|B_{j_0}| < (1 - \Delta) |A_{j_0}|$. Thus
\[
 |A_{j_0}| + |B_{j_0}| < |A_{j_0}|,
\]
whence there exists $I \in \Lambda_{j_0} \setminus (A_{j_0} \cup B_{j_0})$. Since $I \notin \Lambda_{j_0} \setminus A_{j_0}$, (ii) is satisfied for this $I$. Since $I \in \Lambda_{j_0} \setminus B_{j_0}$ and since $p_i^{p-1} \sum_{j=I}^{p_i^p} r^p_j \geq r^p_I$, (iii) is satisfied for this $I$. Since $j_0 < d$ and $I \in \Lambda_{j_0}$, (i) is satisfied for this $I$ by Lemma 3.4.

For the remainder of the proof, $I$ is the fixed interval from Claim 2. Enumerate $\{J \in T : J^- = I\}$ as $(I_i)_{i=1}^l$ and define $g : 2^L \times 2^l \rightarrow 2^L$ by letting
\[
g(\varepsilon, \delta)(i) = \begin{cases} \varepsilon(i) & : i \in [L] \setminus I \\ \delta(j) \varepsilon(i) & : i \in I_j. \end{cases}
\]
We also define the functions $J_0, \ldots, J_l : 2^L \rightarrow 2^L$ by letting $J_0$ be the identity function and $J_i = I_i, I_{i-1}$. Note that $\varepsilon = J_0 \varepsilon$ and $J_{l} \varepsilon = I_{l} \varepsilon$.

Now define $D_Y, E_Y, D_X, E_X : 2^L \rightarrow [0, \infty)$ by
\[
 D_Y(\varepsilon) = \mathbb{E}_\delta g^Y_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p,
 E_Y(\varepsilon) = \Theta^p \sum_{i=1}^l \mathbb{E}_\delta g^Y_\varepsilon(J_{i-1}g(\varepsilon, \delta), J_i g(\varepsilon, \delta))^p,
 D_X(\varepsilon) = \mathbb{E}_\delta g^X_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p,
 E_X(\varepsilon) = \Theta^p \sum_{i=1}^l \mathbb{E}_\delta g^X_\varepsilon(J_{i-1}g(\varepsilon, \delta), J_i g(\varepsilon, \delta))^p.
\]

**Claim 3.** The functions $D_Y, E_Y, D_X, E_X$ satisfy the hypotheses of Lemma 3.8.

**Proof.** $D_Y \leq \lambda^p D_X$ and $E_Y \leq \lambda^p E_X$ follow from the fact that $\text{Lip}(F) \leq \lambda$. For a fixed $\delta \in 2^l$ and $\varepsilon \in 2^L$,
\[
g^X_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p \leq \Theta^p \sum_{i=1}^l g^X_\varepsilon(J_{i-1}g(\varepsilon, \delta), J_i g(\varepsilon, \delta))^p
\]
follows from the triangle and Hölder inequalities. Taking expectations over $\delta$ with $\varepsilon$ held fixed yields that $D_Y \leq E_Y$. Replacing $Y$ with $X$ yields that $D_X \leq E_X$. Now fix $\varepsilon \in 2^L$ and define $f_\varepsilon(\delta) = g(\varepsilon, \delta)$. Define $d_{<i}, d_{\leq i} : 2^l \rightarrow 2^l$ by letting $d_{<i} \delta$ be the member of $2^l$ by replacing $\delta(k)$ with $-\delta(k)$ for each $k < i$ and leaving the other coordinates of $\delta$ unchanged. Let $d_{\leq i} = d_d d_{<i}$. Note that $g^Y_X(\delta, d_{<i} \delta)$ and $g^X_X(d_{<i} \delta, d_{\leq i} \delta)$ have the same distribution as functions of $\delta \in 2^l$. Note also that $g(\varepsilon, d_{<i} \delta) = J_{i-1} g(\varepsilon, \delta)$ and $g(\varepsilon, d_{\leq i} \delta) = J_i g(\varepsilon, \delta)$. Then
\[
 D_Y(\varepsilon) = \mathbb{E}_\delta g^Y_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p = \mathbb{E}_\delta g^X_\varepsilon(\delta, -\delta)^p
\]
\[
 \leq b_{p, l}(\mathcal{F})^{p-1} \sum_{i=1}^l \mathbb{E}_\delta g^X_\varepsilon(\delta, d_{<i} \delta)^p = b_{p, l}(\mathcal{F})^{p-1} \sum_{i=1}^l \mathbb{E}_\delta g^X_\varepsilon(d_{<i} \delta, d_{\leq i} \delta)^p
\]
\[
 \leq (1 + \nu) \Theta^p \sum_{i=1}^l \mathbb{E}_\delta g^X_\varepsilon(J_{i-1} g(\varepsilon, \delta), J_i g(\varepsilon, \delta))^p = (1 + \nu) \Theta^p E_X(\varepsilon).
\]
This yields that $D_Y, E_Y, D_X, E_X$ satisfy hypothesis (i) of Lemma 3.2.

For a fixed $\delta \in 2^l$ and $1 \leq i \leq l$, $g^Y_X(\varepsilon, I_\varepsilon)$ and $g^X_X(g(\varepsilon, \delta), Ig(\varepsilon, \delta))$ have the same distribution as functions of $\varepsilon \in 2^L$, as do $g^Y_X(\varepsilon, I_\varepsilon)$ and $g^Y_X(J_{i-1} g(\varepsilon, \delta), J_i g(\varepsilon, \delta))$. The analogous statements hold with each $Y$ replaced by $X$.

By exchanging order of integration of $\varepsilon$ and $\delta$, we see that
\[
 \mathbb{E}_\varepsilon D_Y(\varepsilon) = \mathbb{E}_\varepsilon \mathbb{E}_\delta g^Y_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p = \mathbb{E}_\delta \mathbb{E}_\varepsilon g^Y_\varepsilon(g(\varepsilon, \delta), Ig(\varepsilon, \delta))^p = \mathbb{E}_\varepsilon g^Y_\varepsilon(\varepsilon, I_\varepsilon)^p = r^p_1.
\]
We similarly deduce that $E Y (\varepsilon) = l p^{-1} \sum_{j=1}^{p} E Y (\varepsilon) = s_p$, and $E X (\varepsilon) = l p^{-1} \sum_{j=1}^{p} s_p$. Thus hypothesis (ii) of Lemma 3.2 is satisfied because $I$ satisfies the conclusions of Claim 2.

Hypothesis (iii) of Lemma 3.2 is satisfied by our choices of $\mu, a, \nu$, and $b$.

Now by Lemma 3.2 there exists $\varepsilon_0 \in 2^l$ such that

$$D Y (\varepsilon_0) > (1-a) E Y (\varepsilon_0),$$

$$D X (\varepsilon_0) > (1-a) E X (\varepsilon_0),$$

and

$$E Y (\varepsilon_0) \geq D Y (\varepsilon_0) > (1-b) \Theta_p E X (\varepsilon_0).$$

Now define $h : 2^l \rightarrow X$ by letting $h (\delta) = f (g (\varepsilon_0, \delta))$. Let

$$r^p = \frac{1}{l} \sum_{i=1}^{l} E \delta^Y (\delta, d_i \delta)^p = E Y (\varepsilon_0)/l^p$$

and

$$s^p = \frac{1}{l} \sum_{i=1}^{l} E \delta^X (\delta, d_i \delta)^p = E X (\varepsilon_0)/l^p \leq (1-b) \Theta_p r^p.$$ 

Then since

$$E \delta^Y (\delta, -\delta)^p = D Y (\varepsilon_0) > (1-a) E Y (\varepsilon_0) = (1-a) l p^{-1} \sum_{i=1}^{l} E \delta^Y (\delta, d_i \delta)^p$$

and

$$E \delta^X (\delta, -\delta)^p = D X (\varepsilon_0) > (1-a) E X (\varepsilon_0) = (1-a) l p^{-1} \sum_{i=1}^{l} E \delta^X (\delta, d_i \delta)^p,$$

it follows from our choice of $a$ and Theorem 2.1 that for any $\delta_1, \delta_2 \in 2^l$,

$$\frac{1}{D} \partial (\delta_1, \delta_2) \leq \frac{d Y (h (\delta_1), h (\delta_2))}{r} \leq D \partial (\delta_1, \delta_2)$$

and

$$\frac{1}{D} \partial (\delta_1, \delta_2) \leq \frac{d X (F \circ h (\delta_1), F \circ h (\delta_2))}{s} \leq D \partial (\delta_1, \delta_2).$$

Since $r \geq \partial s$, this finishes the proof.

Remark 4.1. We observe the following quantitative consequence of the previous proof and our remark from the introduction. If we define $c (F)$ to be the supremum of those $c > 0$ such that for each $D > 1$ and $n \in \mathbb{N}$, there exist $F : X \rightarrow Y \in F, f : 2^n \rightarrow X$, and $a, b > 0$ such that $b \geq ac$ and for each $\varepsilon_1, \varepsilon_2 \in 2^n$,

$$\frac{a}{D} \partial (\varepsilon_1, \varepsilon_2) \leq d X (f (\varepsilon_1), f (\varepsilon_2)) \leq a D \partial (\varepsilon_1, \varepsilon_2)$$

and

$$\frac{b}{D} \partial (\varepsilon_1, \varepsilon_2) \leq d Y (F \circ f (\varepsilon_1), F \circ f (\varepsilon_2)) \leq b D \partial (\varepsilon_1, \varepsilon_2),$$

then

$$\limsup_{n} b_{p,n} (F) \geq c (F) \geq \limsup_{n} b_{p,2^n} (F) \geq 2^{1/p-1} \limsup_{n} b_{p,n} (F).$$
5. THE QUANTITIES \( a_n(F) \)

The goal of this section is to prove the implication \((iii) \Rightarrow (i)\) with the additional quantitative information: If \((i)\) is satisfied and \(c(F)\) is as defined in Remark 1, then \(c(F) = \limsup_n a_n(F) = \lim_n a_n(F)\). It is obvious that \(c(F) \leq \limsup_n a_n(F)\), so we establish the following criterion for obtaining the reverse inequality. The basis of this criterion is to use standard self-improvement arguments for embeddings into \(X\) without worsening the scaling factors between the embeddings of the cube into \(X\) via some \(f\) and the corresponding embedding of the cube into \(Y\) via \(F \circ f\).

**Lemma 5.1.** Suppose \(\lambda = \sup_{F \in \mathcal{F}} \text{Lip}(F) \in (0, \infty)\). If \(\Theta > 0\) is such that for any \(0 < \vartheta < \Theta, D > 1,\) and \(l \in \mathbb{N},\) there exist \(F : X \to Y \in \mathcal{F}\) and \(h : 2^l \to X\) such that \(\text{dist}(F \circ h) \leq D\) and \(\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \geq \vartheta d^l_X(\varepsilon_1, \varepsilon_2)\) for all \(\varepsilon_1, \varepsilon_2 \in 2^l\), then \(c(F) \geq \Theta\).

**Proof.** First fix \(0 < \vartheta < \Theta\). For each \(l \in \mathbb{N},\) let \(\xi_n = \xi_n(F)\) be the supremum of those constants \(\xi > 0\) such that for all \(D > 1,\) there exist \(F : X \to Y \in \mathcal{F}\) and \(f : 2^n \to X\) such that \(\text{dist}(F \circ f) \leq D\), for each \(\xi_1, \xi_2 \in 2^n,\)

\[
\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \geq \vartheta d^l_X(\varepsilon_1, \varepsilon_2)\text{, and}
\]

\[
\mathbb{E}\vartheta^l_X(\varepsilon, -\varepsilon)^2 > n\xi^2 \sum_{i=1}^{n} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2.
\]

Let us observe the following facts:

(i) For all \(n \in \mathbb{N},\) \(\vartheta/\lambda \leq \xi_n \leq 1\).

(ii) For all \(m, n \in \mathbb{N},\) \(\xi_{mn} \leq \xi_m \xi_n\).

For the first fact, \(\xi_n \leq 1\) follows as usual from the triangle and Hölder inequalities. By hypothesis, for each \(D_1, D_2 > 1,\) there exist \(F : X \to Y \in \mathcal{F}\) and \(f : 2^n \to X\) such that \(\text{dist}(F \circ f) \leq D = \min\{D_1, D_2\}\) and \(\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \leq \vartheta d^l_X(\varepsilon_1, \varepsilon_2)\) for all \(\varepsilon_1, \varepsilon_2 \in 2^n\). Then

\[
\lambda^2 \mathbb{E}\vartheta^l_X(\varepsilon, -\varepsilon)^2 \geq \mathbb{E}\vartheta^l_Y(\varepsilon_1, -\varepsilon_1)^2 \geq D^{-4} \sum_{i=1}^{n} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2 \geq \vartheta^2 D^{-4} \sum_{i=1}^{n} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2.
\]

Now applying this as \(D_1 \downarrow 1\) with \(D_2 > 1\) held fixed, we deduce that \(\xi_n \geq \vartheta/\lambda D_2^2\). Unfixing \(D_2 > 1\) yields that \(\xi_n \geq \vartheta/\lambda\).

For the second item, suppose \(\xi_{mn} > \xi_m \xi_n\) for some \(m, n \in \mathbb{N}\). Fix \(\alpha > \xi_m\) and \(\beta > \xi_n\) such that \(\alpha \beta < \xi_{mn}\). By definition of \(\xi_m\), there exists \(D_1 > 1\) such that for all \(F : X \to Y \in \mathcal{F}\) and \(f : 2^m \to X\) with \(\text{dist}(F \circ f) \leq D_1\) and \(\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \leq \vartheta d^l_X(\varepsilon_1, \varepsilon_2)\) for all \(\varepsilon_1, \varepsilon_2 \in 2^m\),

\[
\mathbb{E}\vartheta^l_X(\varepsilon, -\varepsilon)^2 \leq m\alpha^2 \sum_{i=1}^{m} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2.
\]

Similarly, there exists \(D_2 > 1\) such that for any \(F : X \to Y \in \mathcal{F}\) and \(f : 2^n \to X\) with \(\text{dist}(F \circ f) \leq D_2\) and \(\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \leq \vartheta d^l_Y(\varepsilon_1, \varepsilon_2)\) for all \(\varepsilon_1, \varepsilon_2 \in 2^n\),

\[
\mathbb{E}\vartheta^l_X(\varepsilon, -\varepsilon)^2 \leq n\beta^2 \sum_{i=1}^{n} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2.
\]

Let \(D = \min\{D_1, D_2\} > 1\) and fix \(\xi, F, f\) such that \(\alpha \beta < \xi < \xi_{mn}\) and \(F : X \to Y \in \mathcal{F}, f : 2^{mn} \to X\) satisfy \(\text{dist}(F \circ f) \leq D\), \(\vartheta d^l_X(\varepsilon_1, \varepsilon_2) \leq \vartheta d^l_Y(\varepsilon_1, \varepsilon_2)\) for all \(\varepsilon_1, \varepsilon_2 \in 2^{mn}\), and

\[
\mathbb{E}\vartheta^l_X(\varepsilon, -\varepsilon)^2 \geq mn\xi^2 \sum_{i=1}^{mn} \mathbb{E}\vartheta^l_X(\varepsilon, d_i\varepsilon)^2.
\]

Now as usual, let \(I_1, \ldots, I_m\) be a partition of \([mn]\) into intervals of cardinality \(n\) and define

\[
I_j = \begin{cases} 
\varepsilon(i) & : i \in [mn] \setminus I_j \\
-\varepsilon(i) & : i \in I_j
\end{cases}
\]

Define \(g : 2^{mn} \times 2^m \to 2^{mn}\) be defined by \(g(\varepsilon, \delta)(i) = \delta(j) \varepsilon(i)\), where \(i \in I_j\). We identify \(\varepsilon\) with \((\varepsilon_i)_{i=1}^{m}\), where \((\varepsilon((i-1)n+1), \ldots, \varepsilon(in)) = \varepsilon_i \in 2^m\). For \(\varepsilon \in 2^{mn}\), we let \(\varepsilon_{-i} \in 2^{(m-1)n}\) be defined by \(\varepsilon_{-i} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_m)\). Note
that for a fixed \( \varepsilon \) and a fixed \( \varepsilon' \in 2^{mn} \), if \( h \) is the map from \( 2^n \) to \( X \) given by \( \varepsilon \mapsto f((\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon, \varepsilon_{i+1}, \ldots, \varepsilon_m)) \) or if \( h \) is the map from \( 2^{mn} \) to \( X \) given by \( \delta \mapsto f(g(\varepsilon', \delta)) \), then dist\((F \circ h) \leq \text{dist}(F \circ f) \leq D \) and \( \partial g_X^f(\varepsilon_1, \varepsilon_2) \leq g_Y^f(\varepsilon, \varepsilon') \) for all \( \varepsilon_1, \varepsilon_2 \in 2^n \) (resp. \( \varepsilon_1, \varepsilon_2 \in 2^{mn} \)). Then

\[
\begin{align*}
\min_{\varepsilon} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbb{E}_{\varepsilon, \delta} \partial^2 g_X^f(\varepsilon, \delta) &< \mathbb{E}_{\varepsilon} \partial^2 g_X^f(\varepsilon, -\varepsilon)^2 = \mathbb{E}_{\varepsilon} \mathbb{E}_{\delta} \partial^2 g_X^f(\varepsilon, \delta, g(\varepsilon, -\delta))^2 \\
&\leq \max_{\varepsilon} \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \mathbb{E}_{\varepsilon} \partial^2 g_X^f(\varepsilon, \delta, g(\varepsilon, \delta))^2 = \max_{\varepsilon} \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \mathbb{E}_{\varepsilon} \mathbb{E}_{\delta} \partial^2 g_X^f(\varepsilon, \delta, I_i g(\varepsilon, \delta))^2 \\
&= \max_{\varepsilon} \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \partial^2 g_X^f(\varepsilon, \delta, I_i g(\varepsilon, \delta))^2 = \max_{\varepsilon} \sum_{i=1}^{m} \mathbb{E}_{\varepsilon} \partial^2 g_X^f(\varepsilon, I_i g(\varepsilon, \delta))^2 \\
&\leq \min_{\varepsilon} \beta^2 \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \partial^2 g_X^f(\varepsilon, \delta, I_i g(\varepsilon, \delta))^2 = \min_{\varepsilon} \beta^2 \sum_{i=1}^{m} \mathbb{E}_{\varepsilon} \partial^2 g_X^f(\varepsilon, I_i g(\varepsilon, \delta))^2.
\end{align*}
\]

This contradiction yields \((ii)\).

Now since \((\xi_n)_{n=1}^\infty\) is submultiplicative and lies in \([0, \lambda, 1]\), it must be that \( \xi_n = 1 \) for all \( n \in \mathbb{N} \). Indeed, if \( \xi_n < 1 \), then for large enough \( t \in \mathbb{N} \), \( \xi_n < \theta / \lambda \). Now fix \( n \in \mathbb{N} \) and \( D > 1 \). By Theorem 2.1 there exists \( 0 < \mu < 1 \) such that if \( f : 2^n \to X \) satisfies

\[
\mathbb{E}_{g_X^f}(\varepsilon, -\varepsilon)^2 > (1 - \mu) n \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \partial^2 g_X^f(\varepsilon, \delta, I_i g(\varepsilon, \delta))^2,
\]

then dist\((f) \leq D \). By the definition of \( \xi_n \) and since \( \xi_n = 1 > 1 - \mu \), there exist \( F : X \to Y \in \mathcal{F} \) and \( f : 2^n \to X \) such that dist\((F \circ f) \leq D \), \( \partial g_X^f(\varepsilon_1, \varepsilon_2) \leq g_Y^f(\varepsilon_1, \varepsilon_2) \) for all \( \varepsilon_1, \varepsilon_2 \in 2^n \), and

\[
\mathbb{E}_{g_X^f}(\varepsilon, -\varepsilon)^2 > (1 - \mu) n \sum_{i=1}^{m} \mathbb{E}_{\varepsilon, \delta} \partial^2 g_X^f(\varepsilon, \delta, I_i g(\varepsilon, \delta))^2.
\]

From this it follows that dist\((f) \leq D \). Moreover, if \( a, b > 0 \) are such that

\[
\frac{a}{D} \partial(\varepsilon_1, \varepsilon_2) \leq g_X^f(\varepsilon_1, \varepsilon_2) \leq aD \partial(\varepsilon_1, \varepsilon_2)
\]

and

\[
\frac{b}{D} \partial(\varepsilon_1, \varepsilon_2) \leq g_X^f(\varepsilon_1, \varepsilon_2) \leq bD \partial(\varepsilon_1, \varepsilon_2)
\]

for all \( \varepsilon_1, \varepsilon_2 \), then \( a\partial / D \leq bD \). Since \( D > 1 \), \( n \in \mathbb{N} \) are arbitrary, \( c(\mathcal{F}) \geq \theta \). Now we unfix \( 0 < \theta < \Theta \) and deduce that \( c(\mathcal{F}) \geq \Theta \).

Proposition 5.2. If \( \mathcal{F} \) is a uniformly Lipschitz collection of maps, then \( \lim_n a_n(\mathcal{F}) = \inf_n a_n(\mathcal{F}) \).

Proof. Let \( \lambda = \sup_{F \in \mathcal{F}} \text{Lip}(F) \in (0, \infty) \). Fix \( k, l \in \mathbb{N} \). Let \( I_1, \ldots, I_l \) be a partition of \([kl]\) into subintervals of cardinality \( k \). Let \( I_j : 2^{kl} \to 2^{kl} \) be such that

\[
I_j \varepsilon(i) = \begin{cases} 
\varepsilon(i) & : i \in [kl] \setminus I_j \\
-\varepsilon(i) & : i \in I_j.
\end{cases}
\]

Define \( g : 2^{2kl} \times 2^l \to 2^{kl} \) by \( g(\varepsilon, \delta)(i) = \delta(j) \varepsilon(i) \), where \( j \) is such that \( i \in I_j \).

Now fix any \( F : X \to Y \in \mathcal{F} \) and \( f : 2^{kl} \to X \). Note that for fixed \( \varepsilon \in 2^{kl} \), the function \( \delta \mapsto g(\varepsilon, \delta) \) is distance preserving. Therefore for each fixed \( \varepsilon \in 2^l \), the function \( f_\varepsilon : 2^l \to X \) given by \( f_\varepsilon(\delta) = f(g(\varepsilon, \delta)) \) satisfies \( \text{Lip}(f_\varepsilon) \leq \text{Lip}(f) \). Then

\[
\begin{align*}
\mathbb{E}_{\varepsilon} \partial^2 g_Y^f(\varepsilon, -\varepsilon, -\varepsilon) &\leq \mathbb{E}_{\varepsilon} \mathbb{E}_{\delta} \partial^2 g_Y^f(\varepsilon, \delta, g(\varepsilon, -\varepsilon, -\varepsilon)) = \mathbb{E}_{\varepsilon} \mathbb{E}_{\delta} \partial^2 g_Y^f(\delta, -\delta) \\
&\leq a_1(\mathcal{F}) \mathbb{E}_{\varepsilon} \text{Lip}(f_\varepsilon) \leq a_1(\mathcal{F}) \text{Lip}(f).
\end{align*}
\]

From this it follows that \( a_{kl}(\mathcal{F}) \leq a(\mathcal{F}) \).

Now fix \( l \in \mathbb{N} \). For \( m > l \), write \( m = k_m l + r_m \) where \( k_m \in \mathbb{N} \) and \( 0 \leq r_m < k \). For the moment, we suppress the subscript \( m \) and simply write \( k_m = k \) and \( r_m = r \). Now fix \( F : X \to Y \in \mathcal{F} \) and \( f : 2^{mn} \to X \). Let
\[ \delta = (\delta(1), \ldots, \delta(r)) \in 2^r \] be arbitrary and define \( g : 2^l \to 2^m \) by \( g(\varepsilon) = (\varepsilon(1), \ldots, \varepsilon(lk), \delta(1), \ldots, \delta(r)) \). Define \( h : 2^m \to 2^l \) by \( h(\varepsilon) = (\varepsilon(1), \ldots, \varepsilon(lk)) \). Note that for each \( \varepsilon_1, \varepsilon_2 \in 2^l \),

\[
\frac{1}{lk} \partial(\varepsilon_1, \varepsilon_2) = \frac{1}{m} \partial(g(\varepsilon_1), g(\varepsilon_2)).
\]

Therefore the map \( G : 2^l \to X \) given by \( G(\varepsilon) = f(g(\varepsilon)) \) has \( \text{Lip}(G) \leq \frac{m}{lk} \partial(\varepsilon_1, \varepsilon_2) \). Define \( H : 2^m \to X \) by \( H(\varepsilon) = G(h(\varepsilon)) \). Let us also note that \( \mathbb{E} g_Y^f(\varepsilon, -\varepsilon) = \mathbb{E} g_Y^f(\varepsilon, -\varepsilon) \). For any \( \varepsilon \in 2^m \), since \( \varepsilon \) and \( g(h(\varepsilon)) \) differ in at most \( r \) coordinates,

\[
g_Y^f(\varepsilon, g(h(\varepsilon))) \leq \lambda \text{Lip}(f) r/m.
\]

Then formally writing

\[
1 = \frac{\partial(\varepsilon_1, \varepsilon_2)}{m} \partial\bigl(g(\varepsilon_1), g(\varepsilon_2)\bigr) = \frac{\partial(\varepsilon_1, \varepsilon_2)}{2^r} \partial\bigl(g(\varepsilon_1), g(\varepsilon_2)\bigr).
\]

Remark 5.3. If \( L, l \in \mathbb{N} \) are two natural numbers, \( g : 2^L \times 2^l \to 2^L \) is a function such that for each \( \delta \in 2^l \), \( \varepsilon \mapsto g(\varepsilon, \delta) \) is a bijection, and \( \Omega \subset 2^L \) is such that \( \mathbb{P}(\Omega) < 1/2^l \), then there exists \( \varepsilon \in 2^L \) such that \( \{g(\varepsilon, \delta) : \delta \in 2^l \} \subset \Omega^c \). Indeed, for each \( \varepsilon_0 \in 2^L \) and \( \delta_0 \in 2^l \), define \( \Omega_{\varepsilon_0} = \{\delta \in 2^l : g(\varepsilon_0, \delta) \in \Omega\} \), \( \Omega_{\delta_0} = \{\varepsilon \in 2^L : g(\varepsilon, \delta_0) \in \Omega\} \), and \( \Omega_1 = \{(\varepsilon, \delta) \in 2^L \times 2^l : g(\varepsilon, \delta) \in \Omega\} \). Then if for each \( \varepsilon \), there exists \( \delta \in 2^l \) such that \( g(\varepsilon, \delta) \in \Omega^c \),

\[
\mathbb{P}(\Omega) = \frac{1}{2^l} \sum_{\delta \in 2^l} \mathbb{P}(\Omega) = \frac{1}{2^l} \sum_{\delta \in 2^l} \mathbb{P}(\Omega_1) = \frac{1}{2^l} \sum_{\varepsilon \in 2^L} \mathbb{P}(\Omega_{\varepsilon_0}) \geq \frac{1}{2^l} \sum_{\varepsilon \in 2^L} 1/2^l = 1/2^l.
\]

Let us also recall the following simple consequence of the reverse triangle inequality, which we use as a substitute for Theorem 2.1 in this section.

**Proposition 5.4.** For each \( l \in \mathbb{N} \) and \( D > 1 \), there exists \( 0 < a < 1 \) such that if \( (Z, d_Z) \) is any metric space and \( h : 2^l \to Z \) is a map such that \( (1 - a) \text{Lip}(h) < \min_{\delta \in 2^l} d_Z(h(\delta), h(-\delta)) \), then \( h \) is an embedding with distortion not more than \( D \).

**Proof.** Fix \( 0 < a < 1/l \) so small that \( 1 - al > 1/D \). Fix \( \delta \neq \delta_1 \in 2^l \) and let \( m = l \partial(\delta, \delta_1) \). Then

\[
d_Z(h(\delta), h(\delta_1)) \geq d_Z(h(\delta), h(-\delta)) - d_Z(h(-\delta), h(\delta_1)) \geq (1 - a) \text{Lip}(h) - \frac{l - m}{l} \text{Lip}(h)
\]

\[
= \left( \frac{m}{l} - a \right) \text{Lip}(h) \geq \frac{m}{l} (1 - al) \text{Lip}(h) > \frac{\text{Lip}(h)}{D} \partial(\delta, \delta_1).
\]

From this it follows that \( \text{Lip}(h^{-1}) \leq D/\text{Lip}(h) \), so \( \text{Lip}(h) \text{Lip}(h^{-1}) \leq D \).

We next recall the concentration of measure for the Hamming cube.

**Lemma 5.5.** There exist constants \( \alpha, \beta > 0 \) such that for any \( n \in \mathbb{N} \) and \( \lambda_1 > 0 \), if \( \Phi : 2^n \to \mathbb{R} \) is \( \lambda_1 \)-Lipschitz and if \( \phi \) is a median of \( \Phi \), then for any \( t > 0 \),

\[
\mathbb{P}\left( |\Phi - \phi| > t\lambda_1 \right) \leq \alpha \exp(-\beta tn).
\]
Proof of (iii) ⇒ (i). Let \( \lambda = \sup_{F \in \mathcal{F}} \text{Lip}(F) \in (0, \infty) \) and \( \Theta = \lim_{\alpha \to 0} a_{\alpha}(\mathcal{F}) \in [0, \lambda] \). Fix \( 0 < \vartheta < \Theta \). Fix \( l \in \mathbb{N} \) and \( D > 1 \) such that \( \vartheta < \Theta/D \). Let \( 0 < a < 1 \) be chosen according to Proposition [5.4]. Now fix \( 0 < \mu < 1/2 \) such that \( \frac{1+\mu}{1-a} > 1-a \) and \( \frac{(1-2\mu)}{D} \Theta > \vartheta \). Fix \( 0 < \eta < \mu \) such that

\[
\frac{1-\mu}{l} + (1+\eta)\left(\frac{l-1}{l}\right) < 1-\eta.
\]

Fix \( t > 0 \) such that \( t < \mu \Theta/l \) and

\[
(l+1)t < \mu(1-2\mu) \Theta.
\]

Note that the second inequality implies that for any \( M \geq (1-2\mu) \Theta \),

\[
(l+1)t + M \leq (1+\mu)M.
\]

Fix \( k \in \mathbb{N} \) so large that for all \( m \geq k \), \( a_m(\mathcal{F}) \in ((1-\eta)\Theta, (1+\eta)\Theta) \), \( \lambda \exp(-\beta tm/8\lambda) < t/4\lambda \), and \( (l+1)\alpha \exp(-\beta tm/8\lambda) < 1/2^l \). Fix \( F : X \to Y \in \mathcal{F} \) and \( f : 2^{lk} \to X \) such that

\[
\mathcal{E} \varphi_f^I(\varepsilon, -\varepsilon) > (1-\eta)\text{Lip}(f).
\]

Let \( T \) be the \((l,k)\) interval tree and for each \( I \in T \), define \( \Phi_I : 2^{lk} \to \mathbb{R} \) by \( \Phi_I(\varepsilon) = \varphi_f^I(\varepsilon, \varepsilon) \), where \( \varepsilon \) is obtained by changing the signs of the coordinates of \( \varepsilon \) which lie in \( I \) and leaving the other coordinates unchanged. Let \( \phi_I \) be a median of \( \Phi_I \) and let \( r_I = \mathbb{E} \Phi_I \). For the remainder of the proof, fix a partition of \([lk]\) into intervals \( I_1, \ldots, I_l \), where \( I_j = \{(j-1)k + 1, \ldots, jk\} \).

We first claim that

\[
(1-\eta)\text{Lip}(f) < r_{[lk]} < (1+\eta)\text{Lip}(f)
\]

and for each \( I \in \Lambda_1 \),

\[
(1-\mu)\text{Lip}(f) \leq l r_I \leq (1+\eta)\text{Lip}(f).
\]

The first pair of inequalities follows from the fact that \( r_{[lk]} = \mathcal{E} \varphi_f^I(\varepsilon, -\varepsilon) \leq a_{lk}(\mathcal{F}) \text{Lip}(f) < (1+\eta)\text{Lip}(f) \) and \( F, f \) were chosen such that \( (1-\eta)\text{Lip}(f) < \mathcal{E} \varphi_f^I(\varepsilon, -\varepsilon) \). Now fix \( 1 \leq j \leq l \) and define \( g : 2^{lk} \times 2^k \to 2^{lk} \) by letting

\[
g(\varepsilon, \delta)(i) = \begin{cases} 
\varepsilon(i) & : i \in [lk] \setminus I_j \\
\delta(m)\varepsilon(i) & : i = (j-1)k + m.
\end{cases}
\]

Note that for a fixed \( \varepsilon \in 2^{lk} \), the map \( f_\varepsilon : 2^l \to 2^{lk} \) given by \( f_\varepsilon(\delta) \mapsto g(\varepsilon, \delta) \) scales distances by a factor of \( 1/l \). From this it follows that for a fixed \( \varepsilon \), the map \( \delta \mapsto f(g(\varepsilon, \delta)) \) has Lipschitz constant not more than \( \text{Lip}(f)/l \). Therefore

\[
r_I = \mathcal{E} \varphi_f^I(\varepsilon, I_\varepsilon \varepsilon) = \mathcal{E} \mathcal{E} \delta f(g(\varepsilon, \delta), g(\varepsilon, -\delta)) \leq a_{lk}(\mathcal{F}) \mathcal{E} \text{Lip}(f) \leq (1+\eta)\text{Lip}(f)/l.
\]

From this we deduce that

\[
\max_{I \in \Lambda_1} lr_I \leq (1+\eta)\text{Lip}(f).
\]

To see that \( (1-\mu)\text{Lip}(f) \leq lr_I \) for all \( I \in \Lambda_1 \), suppose that there exists \( I_0 \in \Lambda_1 \) such that \( r_{I_0} < (1-\mu)\text{Lip}(f)/l \). Then

\[
(1-\eta)\text{Lip}(f) < r_{[lk]} \leq \sum_{I \in \Lambda_1} r_I < (1-\mu)\text{Lip}(f)/l + (l-1)(1+\eta)\text{Lip}(f)/l
\]

\[
= \frac{1-\mu}{l} + (1+\eta)\left(\frac{l-1}{l}\right)\text{Lip}(f) < (1-\eta)\text{Lip}(f).
\]

This is a contradiction and yields the remaining inequality. Here we are using the fact that \( r_{[lk]} \leq \sum_{I \in \Lambda_1} r_I \), which follows from the triangle inequality as in the proof from the previous section.

Let \( \Upsilon_I = (\Phi_I - \phi_I) > \text{Lip}(f)/4 \) and \( \Omega_I = (|r_I - \phi_I| > \text{Lip}(f)) \). We claim that \( \Phi_I \) is \( 2\lambda \text{Lip}(f) \)-Lipschitz taking values in \([0, \lambda \text{Lip}(f)]\), so \( \mathbb{P}(\Upsilon_I) \leq t/4\lambda \), \( |r_I - \phi_I| \leq t/2 \), and \( \mathbb{P}\left(\bigcup_{I \in T \setminus \Lambda_2} \Omega_I\right) < 1/2^l \). Since \( \text{diam}(2^{lk}) = 1 \) and \( \text{Lip}(F \circ f) \leq \lambda \text{Lip}(f) \), we deduce that \( \Phi_I \) takes values in \([0, \lambda \text{Lip}(f)]\). Next let us show that \( \Phi_I \) is \( 2\lambda \text{Lip}(f) \)-Lipschitz.

Fix \( \varepsilon_1, \varepsilon_2 \in 2^{lk} \) and note that

\[
g_f^I(\varepsilon_1, I \varepsilon_1) - g_f^I(\varepsilon_2, I \varepsilon_2) \leq g_f^I(\varepsilon_1, \varepsilon_2) + g_f^I(\varepsilon_2, I \varepsilon_2) + g_f^I(I \varepsilon_2, I \varepsilon_1) - g_f^I(\varepsilon_2, I \varepsilon_2)
\]

\[
= g_f^I(\varepsilon_1, \varepsilon_2) + g_f^I(I \varepsilon_1, I \varepsilon_2) = 2g_f^I(\varepsilon_1, \varepsilon_2) \leq 2\lambda \text{Lip}(f)\text{Lip}(f)\Theta(\varepsilon_1, \varepsilon_2).\]
By symmetry, we deduce that $\Phi_I$ is $2\lambda \text{Lip}(f)$-Lipschitz. From this it follows that

$$
P(\mathcal{T}_I) = P\left( |\phi_I - \phi_I'| \geq \frac{t}{8\lambda}(2\lambda \text{Lip}(f)) \right) \leq \alpha \exp(-\beta t / 8\lambda) < t/4\lambda. $$

Therefore

$$|\phi_I - r_I| \leq E_1(\mathcal{T}_I,|\phi_I - \Phi_I| + E_1(\mathcal{T}_I,|\phi_I - \Phi_I|) \leq \lambda \text{Lip}(f)P(\mathcal{T}_I) + t\text{Lip}(f)/4 \leq t\text{Lip}(f)/2.$$ 

Therefore $\Omega_I \subset \mathcal{T}_I$ and

$$P\left( \bigcup_{I \in T \setminus \Lambda_2} \Omega_I \right) \leq \sum_{I \in T \setminus \Lambda_2} P(\mathcal{T}_I) \leq (l+1)\alpha \exp(-\beta t / 8\lambda) < 1/2^l.$$ 

From this and Remark 5.3 we may define $g : 2^k \times 2^l \to 2^k$ by $g(\varepsilon, \delta) = \delta(j)\varepsilon(i)$ when $i \in I_j$ and choose $\varepsilon_0 \in 2^k$ such that $\{g(\varepsilon_0, \delta) : \delta \in 2^l\} \subset \bigcap_{I \in T \setminus \Lambda_2} \Omega_I.$

Now define $h : 2^l \to X$ by $h(\delta) = f(g(\varepsilon_0, \delta)).$ Note that

$$(1 - \mu)l \max_{i \in [l], \delta \in 2^l} \phi_X^\delta_i(\delta, d_0) \leq \frac{1 - \eta}{1 + \eta} l \max_{I \in \Lambda_2, \delta \in 2^l} \Phi_I(g(\varepsilon_0, \delta)) \leq lt\text{Lip}(f) + \frac{1 - \eta}{1 + \eta} lr_I$$

$$\leq lt\text{Lip}(f) + (1 - \eta)\Theta\text{Lip}(f) \leq lt\text{Lip}(f) + r_{[lk]}$$

$$\leq (l + 1)t\text{Lip}(f) + \min_{\delta \in 2^l} \Phi_{[lk]}(g(\varepsilon_0, \delta))$$

$$< (1 + \mu) \min_{\delta \in 2^l} \Phi_{[lk]}(g(\varepsilon_0, \delta)) = (1 + \mu) \min_{\delta \in 2^l} \delta_X^\delta_i(\delta, -\delta).$$

Here we are using the fact that since

$$M\text{Lip}(f) := \min_{\delta \in 2^l} \Phi_{[lk]}(g(\varepsilon_0, \delta)) \geq r_{[lk]} - t\text{Lip}(f) \geq ((1 - \mu)\Theta - t)\text{Lip}(f) \geq (1 - 2\mu)\Theta\text{Lip}(f),$$

it follows from our choice of $t$ that

$$(l + 1)t\text{Lip}(f) + M\text{Lip}(f) < (1 + \mu)M\text{Lip}(f).$$

Now our choice of $a$ and $\mu$ combined with Proposition 5.4 yield that $\text{dist}(F \circ h) \leq D.$

We next show that for any $\delta_1, \delta_2 \in 2^l,$ $\partial \phi_X^\delta_i(\delta_1, \delta_2) \leq \phi_X^\delta_i(\delta_1, \delta_2).$ First observe that the map $\delta \mapsto g(\varepsilon_0, \delta)$ is length preserving, so $\text{Lip}(h) \leq \text{Lip}(f).$ Therefore

$$\phi_X^\delta_i(\delta_1, \delta_2) \leq \text{Lip}(f)\partial(\delta_1, \delta_2).$$

Now let us observe that

$$\text{Lip}(F \circ h) = \max_{i \in [l], \delta \in 2^l} l\phi_X^\delta_i(\delta, d_0) \geq \min_{I \in \Lambda_2, \delta \in 2^l} \Phi_I(g(\varepsilon_0, \delta))$$

$$\geq \min lr_I - lt\text{Lip}(f) \geq (1 - \mu)\Theta\text{Lip}(f) - lt\text{Lip}(f) \geq (1 - 2\mu)\Theta\text{Lip}(f).$$

Since $h$ has distortion at most $D,$ for any $\delta_1, \delta_2 \in 2^l,$

$$\phi_X^\delta_i(\delta_1, \delta_2) \geq \frac{(1 - 2\mu)}{D} \cdot \Theta\text{Lip}(f)\partial(\delta_1, \delta_2) \geq \partial\text{Lip}(f)\partial(\delta_1, \delta_2) \geq \partial \phi_X^\delta_i(\delta_1, \delta_2).$$

An appeal to Lemma 5.1 finishes the proof.

\[\square\]

**References**

[1] D. Amir and V. D. Milman, *Unconditional and symmetric sets in n-dimensional normed spaces*, Israel J. Math. **37** (1980), 3-20.

[2] B. Beauzamy, *Opérateurs de type Rademacher entre espaces de Banach*, Sém Maurey-Schwartz, Exposés 6 et 7, École Polytechnique, Paris (1975/76).

[3] J. Bourgain, V. Milman, H. Wolfson, *On type of metric spaces*, Trans. Amer. Math. Soc., **294** (1986), no. 1, 295-317.
[4] R.M. Causey, S. Draga, T. Kochanek, *Operator ideals and three-space properties of asymptotic ideal seminorms*, submitted.

[5] J.A. Chávez-Domínguez, *Lipschitz (q,p)-mixing operators*, Proc. Amer. Math. Soc., **140** (2012), no. 9, 3101-3115.

[6] P. Enflo, *On infinite dimensional topological groups*, Séminaire sur la Géométrie des Espaces de Banach (1977–1978), Exp. No. 10–11, 11 pp., École Polytech., Palaiseau, 1978.

[7] A. Eskin, D. Fisher, K. Whyte, *Coarse differentiation of quasi-isometries I: Spaces not quasi-isometric to Cayley graphs*, Ann. of Math., **176** (2012), no. 1, 221-260.

[8] J. Farmer, W.B. Johnson, *Lipschitz p-summing operators*, Proc. Amer. Math. Soc., **137** (2009), no. 9, 2989-2995.

[9] L. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Combinatorial Theory **1** (1966), 385-393.

[10] A. Hinrichs, *Operators of Rademacher and Gaussian subcotype*, J. Lond. Math. Soc., **63** (2001), no. 2, 453-468.

[11] W.B. Johnson, G. Schechtman, *Diamond graphs and super-reflexivity*, J. Topol. Anal., **1** (2009), no. 2, 177-189.

[12] H. Knaust, Th. Schlumprecht, E. Odell, *On Asymptotic structure, the Szlenk index and UKK properties in Banach spaces*, Positivity **3** (1999), 173-199.

[13] A. Naor, *An introduction to the Ribe program*, Jpn. J. Math., **7**, (2012), no. 2, 167-233.

[14] G. Pisier, *Martingales with values in uniformly convex Banach spaces*, Israel J. Math., **20**(1975), no. 3, 326-350.

[15] M. Ribe, *On uniformly homeomorphic normed spaces*, Ark. Mat., **14** (1976), no. 2, 237-244.

[16] J. Wenzel, *Uniformly convex operators and martingale type*, Rev. Mat. Iberoam., **18** (2002), no. 1, 211-230.