ON RAMSEY TECHNIQUES IN QUANTITATIVE METRIC GEOMETRY: *THE MINIMUM DISTORTION NEEDED TO EMBED A BINARY TREE INTO ℓ_p*

DANIEL GALICER

ABSTRACT. It is commonly said that a result is typical of the Ramsey theory, if in any finite coloring of some mathematical object one can extract a sub-object (usually having some kind of desired structure), which is monochromatic. In this essay we discuss in detail a clever Ramsey-type argument due to Jiří Matoušek utilized in the context of embedding theory. Namely, to study the smallest constant $C = C(n)$ for which a complete binary tree of height n can be C-embedded into a given uniformly convex Banach space. As a consequence, the quantitative lower bound of $const \cdot (\log n)^{\min(1/2,1/p)}$ in the distortion needed to embed this space into ℓ_p (for $1 < p < \infty$) is explained.

1. A GLIMPSE TO RAMSEY-TYPE RESULTS AND BOURGAIN'S WORK ON BINARY TREES

Let us begin with a seemingly banal but enlightening question: How many people should be on a party to ensure that three of them are either mutual acquaintances (each one knows the other two) or mutual strangers (each one does not know either of the other two)? This query is usually known as the problem of friends and strangers. For our purposes, it is convenient to phrase this question in a graph-theoretic language. Denote by n the number of people at the party and suppose that each person is represented by a vertex of a complete graph (a simple undirected graph in which every pair of distinct vertices is connected by a unique edge) K_n of order n. Given two partygoers (or vertices), we paint in red the edge that links them if they know each other and in blue otherwise. Therefore, our problem translates into the following: How big n must be to assert the existence of a complete subgraph of order 3 in K_n painted entirely in red or blue?

Ramsey's classical theorem [Ram30] points in the same direction as this question. Colloquially speaking, it states that in any coloring of the edges (using a palette with a finite number of colors) of a sufficiently large complete graph, one will find monochromatic (i.e., of the same color) complete subgraphs. This foundational tool in combinatorics initiated a new perspective that is now framed as part of the Ramsey theory. But what exactly do people mean

Key words and phrases. Graph trees, bi-Lipschitz structure, Ramsey techniques, Uniformly convex Banach spaces.

when they refer to a statement as of Ramsey-type? Perhaps the most popular result of this type (although quite naive) is the well-known pigeonhole principle: if $f : \{1,\dots n\} \to \{1,\dots m\}$ and $n > m$ then f can not be injective (if you have fewer pigeon holes than pigeons and you put every pigeon in a pigeon hole, then there must result at least one pigeon hole with more than one pigeon). The typical scenario of the Ramsey theory starts with some mathematical object which is divided into several pieces. The question that arises in this context is how big should be the original object in order to affirm that at least one of the pieces has a given interesting property. Going back to our friends and strangers' example, we wanted to know how big had to be our study set (n = the number of partygoers) to ensure the existence of a certain structure (three "friends" or three "complete strangers"). By the way... the answer is $n \geq 6$ and it is an interesting challenge to prove this, but this is another matter.

Summarizing, a statement has essentially a Ramsey-type flavor if it ensures the existence of some kind of rigid substructure in a given set having enough members. Being a bit extreme, Ramsey-type results give certain regularity amid disorder. These techniques have proven to be extremely useful in various contexts alien to it (allowing to solve, for example, long-standing problems in analysis; see [AT06] for a proper treatment on several important applications). Of course, Ramsey theory may be labeled undoubtedly as a part of combinatorics or discrete mathematics, and in general these branches seem to be quite distant, at least at first glance, from embedding theory or metric geometry. The aim of this note is to show how to apply this kind of discrete techniques to study the smallest distortion needed in a particular embedding problem. Before going into details, let us start with a couple of definitions in order to clarify all the notions we deal with.

Given two metric spaces (M, d_M), (N, d_N), and a mapping $f : M \to N$, we denote the Lipschitz constant of f by $\|f\|_{\mathrm{Lip}} := \sup\{\frac{d_N(f(x), f(y))}{d_M(x,y)} : x \neq y\}$. If f is injective then the (bi-Lipschitz) distortion of f is defined as $\mathsf{dist}(f) = \|f\|_{\mathrm{Lip}} \cdot \|f^{-1}\|_{\mathrm{Lip}}$. Informally, the distortion is a measure of the amount by which a function warps distances. Note that a function with distortion 1 does not necessarily preserve mutual distances but it may re-scale them in the same ratio. We write $M \overset{C}{\hookrightarrow} N$ if there exist an embedding $f : M \to N$ with $\mathsf{dist}(f) \leq C$ (such an embedding is called a C-embedding or a C-isomorphism). The smallest distortion with which M embeds into N is denoted $c_N(M)$, namely,

$$c_N(M) = \inf\{C : \ M \overset{C}{\hookrightarrow} N\}.$$

We say that $f : M \hookrightarrow N$ is non-contracting if $d_M(x, y) \leq d_N(f(x), f(y))$ for every $x, y \in M$ (i.e., $\|f^{-1}\|_{\mathrm{Lip}} \leq 1$). In this working we focus on the case where the target space N is a Banach space $(X, \|\cdot\|)$, so we can compute $c_X(M)$ (by re-scaling if necessary) as $\inf\{\|f\|_{\text{Lip}} : f : M \hookrightarrow X \text{ non-contracting}\}$. If $X = \ell_p$ for some $p \geq 1$ we use the shorter notation $c_p(M) = c_{\ell_p}(M)$. The parameter $c_2(M)$ is usually known as the Euclidean distortion of X.

Lipschitz (or uniform and coarse) embeddings of metric spaces into Banach spaces with "good geometrical properties" have found many significant applications, specially in computer science and topology. The advantages of low distortion embeddings are based on the fact that for spaces with "good properties" one can apply several geometric tools which are generally not available for typical metric spaces. The most significant accomplishments throughout these lines were obtained in the design of algorithms (the information obtained from concrete geometric representations of finite spaces is used to obtain efficient approximation algorithms and data structures). In this context, the spaces with "good geometrical features" are mostly separable Hilbert spaces (or certain classical Banach spaces such as L_p spaces).

The bi-Lipschitz structure of arbitrary trees and its applications to different context have been studied extensively during the last years. We refer to [Dre84, Mat90, Bar98, JLPS02, LS03, Dra03, FRT03, BS05, NPS$^+$06] and the references therein for a detailed treatment. Recall that a (graph-theoretical) tree is an undirected graph $T = (V, E)$ in which any two vertices are connected by exactly one path. In other words, any connected graph without simple cycles is a tree. The present essay is devoted to the study of the Euclidean (and L_p) distortion of complete binary trees.

Just to be in tune, we denote by B_n the complete rooted binary tree of height (or depth) n. This is a graph defined as follows: B_0 is a single vertex (the root), and B_{n+1} arises by taking one vertex (the root) and connecting it to the roots of two disjoint copies of B_n. We also consider k-ary trees of height h (each non-leaf vertex has k successors), which we denote by $T_{k,h}$. These spaces are metric space endowed with the path-metric: the distance between two vertices is the number of edges in the path connecting them (i.e., we consider the graph-theoretic distance on the vertex set, with edges of unit length).

A famous result in embedding theory due to Bourgain [Bou86] states the following.

Theorem 1.1. *Let $1 < p < \infty$, for any embedding $f : B_n \hookrightarrow \ell_p$ we have $\mathsf{dist}(f) \geq c \log(n)^{\min(1/2,1/p)}$, where c is a constant depending only on p.*

In other words, he showed that $c_p(B_n) = \Omega_p\left(\log(n)^{\min(1/2,1/p)}\right)$. Among Bourgain's contributions we find a noteworthy characterization (in terms of their metric structure) of a linear property of Banach spaces. Namely, he showed that a Banach space X is superreflexive (see definition below) if and only if $\lim_{m\to\infty} c_X(B_n)=\infty$. He also established the following interesting dichotomy: For a Banach space X either $c_X(B_n)=1$ for all n, or there exists $\alpha>0$ such that $c_X(B_n)=\Omega\left((\log n)^{\alpha}\right)$. Bourgain used this result to solve a question posed by Gromov, showing that the hyperbolic plane does not admit a bi-Lipschitz Euclidean embedding. The arguments involved in his work are based on the use of some technical probabilistic tools (diadic Walsh-Paley martingales). We highlight that Bourgain derived Theorem 1.1 as a particular case of a much more general result involving structural properties of Banach spaces. We recall some classic definitions from Banach space theory in order to state all this.

The modulus of (uniform) convexity $\delta_X(\varepsilon)$ of a Banach space X endowed with norm $\|\cdot\|$ is defined as

$$\delta_X(\varepsilon):=\inf\left\{1-\left\|\frac{x+y}{2}\right\| \,\middle|\, \|x\|=\|y\|=1 \text{ and } \|x-y\|\geqslant\varepsilon\right\},$$

for $\varepsilon\in(0,2]$. The space X is said to be uniformly convex of type $q\geqslant 2$ if $\delta_X(\varepsilon)\geqslant c\varepsilon^q$ for some $c>0$. Put simply, the modulus of convexity measures how deep inside (in the unit ball of X) must lie the midpoint of a line segment with extremes in the sphere of X in terms of the length of the segment. Intuitively, if a space has a "big" modulus of convexity then the center of a line segment included in the unit ball must lie very deep inside the ball (i.e., has small norm) unless the segment is short. If the function $\delta_X(\cdot)$ is never zero, we say that X is uniformly convex (or uniformly rotund). Spaces with this property are common examples of reflexive Banach spaces (this is a consequence of the classical Milman-Pettis theorem [Mil38, Pet39]). Since the converse does not hold, this justifies the name given to those spaces that are isomorphic to uniformly convex spaces; that is, superreflexive Banach space.

It is well-known that, for $1<p<\infty$, the ℓ_p space (or any L_p-space) is uniformly convex. The asymptotic behavior of its moduli (as computed by Hanner [Han56]) is given by

$$\delta_p(\varepsilon)=\begin{cases}\frac{(p-1)\varepsilon^2}{8}+o(\varepsilon^2) & \text{if } 1<p\leq 2,\\ \frac{\varepsilon^p}{p2^p}+o(\varepsilon^p) & \text{if } 2\leq p<\infty.\end{cases} \tag{1}$$

In particular, $\delta_p(\varepsilon)\geq c\varepsilon^{\max(2,p)}$ where $c=c(p)$.

Now that we have the definition of uniform convexity in mind, we are able to state Bourgain's embedding theorem on binary trees.

Theorem 1.2. *Let X be a uniformly convex Banach space whose modulus of uniform convexity satisfies $\delta_X(\varepsilon)\geqslant c\varepsilon^q$ for some $q\geq 2$ and $c>0$ (i.e., X uniformly convex of type $q\geqslant 2$). Then* *the minimum distortion necessary for embedding B_n into X is at least c_1 $(\log n)^{1/q}$ for some $c_1 = c_1(c,q) > 0$.*

Observe that Theorem 1.1 becomes a direct consequence of this theorem since, by Equation (1), any L^p-space ($1 < p < \infty$) is uniformly convex of type $q = \max(2,p)$. It should be noted that Bourgain's bound in Theorem 1.1 is optimal, as proven by Bourgain himself in his seminal work for the Euclidean case ($p = 2$) and by Matoušek [Mat99] for every $1 < p < \infty$. Thus, $c_p(B_n) = \Theta\left(\log(n)^{\min(1/2,1/p)}\right)$.

Several proofs of Theorem 1.2 have been published over the years (e.g., [Bou86, Mat99, LS03, LNP09, MN13, Klo14]). This note aims to present an elementary proof (due to Matoušek in [Mat99]), where a shrewd use of a Ramsey-type result is displayed.

2. Matoušek's proof or Theorem 1.2

Matoušek's argument has a geometric ingredient and a combinatorial one. The former is the simplest and relates uniform convexity to embeddings of some special trees. Consider the four-vertices tree with one root v_0 which has one son v_1 and two grandchildren v_2, v_2'. We denote by S this tree (with edges of unit length). We say that a subset $F = \{x_0, x_1, x_2, x_2'\}$ of a metric space (M, d_M) is an δ-*fork* if there exist a function $f : S \to F$ mapping v_i to x_i (for $i = 0,1,2$) and v_2' to x_2', such that the restricted functions $f|_{\{v_0,v_1,v_2\}} : \{v_0, v_1, v_2\} \to \{x_0, x_1, x_2\}$ and $f|_{\{v_0,v_1,v_2'\}} : \{v_0, v_1, v_2'\} \to \{x_0, x_1, x_2'\}$ are $(1+\delta)$-isomorphisms. Qualitatively, for small δ, the mutual distances between elements of the sets $\{x_0, x_1, x_2\}$ and $\{x_0, x_1, x_2'\}$ are similar to those of $\{0,1,2\} \subset \mathbb{R}$. It should be noted that in this definition, no information about the distance between the vertices v_2 and v_2' is inherited by F. We call the points x_2 and x_2' the *tips* of F. The name "fork" (which is obviously given by mnemonic purposes) comes by understanding how this object should look like in the Euclidean space $\mathbb{R}^3$ for a small δ. The following lemma states that if a fork F in a uniformly convex Banach space has a rigid structure (i.e. δ is small) then its tips are very close to each other.

Lemma 2.1. *(Fork Lemma) Let X be a uniformly convex Banach space whose modulus of uniform convexity satisfies $\delta_X(\varepsilon) \geqslant c\varepsilon^q$ for some $q \geq 2$ and $c > 0$, and let $F = \{x_0, x_1, x_2, x_2'\} \subset X$ be an δ-fork. Then $\|x_2 - x_2'\| = O(\delta^{1/q})\|x_0 - x_1\|$.*

Proof. (of Lemma 2.1) By translating and re-scaling if necessary, we may presume that $x_0 = 0$ and $\|x_1\| = 1$.

Set

$$z := x_1 + \frac{x_2 - x}{\|x_2 - x_1\|}.$$

Obvious computations show $\|z - x_1\| = 1$, $\|x_2 - x_1\| \le 1 + 2\delta$, $\|z - x_2\| \le 2\delta$ and $\|z\| \ge 2 - 2\delta$. Put $u = z - 2x_1$. Observe that $x = x_1$ and $y := x_1 + u$ lie on the unit sphere of X, and for the midpoint

$$\frac{x+y}{2} = x + \frac{u}{2} = \frac{z}{2}$$

we have

$$\left\| \frac{x+y}{2} \right\| \ge 1 - \delta.$$

Using the uniform convexity condition we obtain $\|y - x\| = \|u\| = O(\delta^{1/q})$, thus

$$\|x_2 - 2x_1\| \le \|x_2 - z\| + \|\underbrace{z - 2x_1}_{u}\| \le 2\delta + O(\delta^{1/q}) = O(\delta^{1/q}),$$

for δ small. Note that the constant of proportionality in the last $O(\cdot)$ notation depends only on c and q. Analogously (by symmetry), we also get $\|x_2' - 2x_1\| = O(\delta^{1/q})$, hence $\|x_2 - x_2'\| = O(\delta^{1/q})$, concluding the proof. $\square$

For our purposes it will be useful to compute the smallest distortion needed to embed complete k-ary trees into uniformly convex spaces instead of dealing with complete binary trees. Any complete k-ary tree can be 2-embedded into a complete binary tree of height large enough. This is stated in the following lemma. Recall that the level of a vertex of $T_{k,h}$ is just its distance from the root.

Lemma 2.2. *Let $T_{k,h}$ be a complete k-ary tree of height h. Then $T_{k,h}$ can be 2-embedded into the complete binary tree B_n for height $n = 2h\lceil \log_2 k \rceil$.*

Proof. Note that it is sufficient to demonstrate this for powers of 2 (i.e., $k = 2^s$). We obviously map the root of $T_{2^s,h}$ into the root of B_{2hs}. We now follow an inductive procedure. If a vertex v of $T_{2^s,h}$ has already being mapped to a vertex u at some level l of B_{2hs}, we map the 2^s successors of v to 2^s vertices above u at level $l + 2s$ whose mutual distances are all between $2s$ and $4s$. Indeed, without loss of generality we can assume that $l = 0$, now note that B_{2s} is constructed by gluing to each leaf of B_s another disjoint copy of B_s. For each of these copies we select a leaf and map the successors of v to them. $\square$

Given a rooted tree T, we denote by $SP(T)$ the set all pairs of vertices $\{x, y\}$ of T such that x lies in between the way from y to the root. The following Ramsey-type result, whose proof is simple and short, can be regarded as the cornerstone towards the proof of Theorem 1.2.

Lemma 2.3. *Suppose that each of the pairs of the set $SP(T_{k,h})$ is painted with a color from a palette of r colors. If $k \geq r^{(h+1)^2}$, then there exist a subtree $T' \subset T_{k,h}$ which is a copy of the complete binary tree B_h, such that the color of any pair $\{x, y\} \in SP(T')$ depends exclusively on the levels of x and y.*

Proof. We start proving the following simple claim: Suppose that all the leaves of $T_{k,h}$ (i.e., vertices at level h) are colored by r' colors and $k > r'$ then there exist a copy of B_h in $T_{k,h}$ such that all its leaves have the same color. The case $h = 0$ is trivial. For $h \geq 1$, consider all the k subtrees isomorphic to $T_{k,h-1}$ connected to the root of $T_{k,h}$. By inductive hypothesis we can pick a copy of B_{h-1} with monochromatic leaves. Since $k > r'$ by the pigeonhole principle, two of this copies have the same color of leaves. If we connect this copies to the root we get the copy of B_h with the desired property.

Going back to our problem... Label each leaf $z \in T_{k,h}$ by a vector having the colors of the pairs $\{x, y\} \in SP(T_{k,h})$ lying on the path form z to the root (we write the coordinates of the vectors using a predetermined order common for all leaves). We want show the existence of subtree $T' \subset T_{k,h}$, which is a copy of B_h, such that the color of the pair $\{x, y\} \in SP(T')$ depends only on the levels of x and y. This can be rephrased into finding a copy of B_h in $T_{k,h}$ whose all leaves are labeled with the same vector. Note that each of this vector have $\binom{h+1}{2} < (h+1)^2$ coordinates; hence, the leaves of $T_{k,h}$ are colored with $r' < r^{(h+1)^2}$ possible colors. The result now follows from our preliminary claim. □

The following lemma states that if a copy of the metric space $P_h = \{0, 1, \ldots, h\} \subset \mathbb{R}$ is embedded with a constant-bounded distortion into a given metric space, and h is large enough, then we can find a 3-term arithmetic progression such that the restriction of our embedding to this set has distortion near 1.

Lemma 2.4. *(Path embedding Lemma) For any given constants $\alpha > 0$ and $\beta \in (0, 1)$ the exists a constant $C = C(\alpha, \beta)$ with the following property: for every non-contracting mapping f defined in the metric space $P_h = \{0, 1, \ldots, h\} \subset \mathbb{R}$ into some metric space (M, d_M) with $h \geq 2^{CK^\alpha}$, for $K = \|f\|_{\mathrm{Lip}}$, there exists an arithmetic progression $Z = \{x, x+a, x+2a\} \subset P_h$ such that the restriction of f onto Z is a $(1+\varepsilon)$-isomorphism with*

$$\varepsilon = \beta \left(\frac{d_M(f(x), f(x+a))}{a} \right)^{-\alpha}.$$

The proof of the lemma is a bit cumbersome. Maybe should put it aside on a first read.

Proof. (of Lemma 2.4) We define, for $a \in \{1,\dots,h\}$ the number

$$K(a) := \max\left\{\frac{d_M(f(x),f(y))}{|x-y|} : x,y \in P_h, |x-y| = a\right\}.$$

By the triangle inequality $K(a) \geq K(2a)$ for every a.

We also define an decreasing sequence of numbers $x_0 > x_1 > x_2 > \dots$ by setting $x_0 = K$ and $x_{j+1} = \frac{x_j}{\left(1+\frac{\beta}{4x_j^\alpha}\right)}$. We denote by t the first index with $x_t \leq 1$. It can be seen that $t = O(K^\alpha)$, and therefore we can assume that $2^t \leq h$ (by picking C large enough). Observe that in the sequence $K(2^0) \geq K(2^1) \geq K(2^2) \geq \dots \geq K(2^t)$, there must be two consecutive values, say $K(2^i)$ and $K(2^{i+1})$, belonging to the same interval $[x_{j+1}, x_j)$. Thus,

$$1 \leq \frac{K(2^i)}{K(2^{i+1})} \leq 1+\eta,$$

where $\eta = \frac{\beta}{4K(2^i)^\alpha}$. We consider the number $a := 2^i$ and we fix the points $x, x+2a \in P_h$ such that $K(2a) = K(2^i)$ is attained. This means that, $d_M(f(x), f(x+2a)) = 2aK(2a)$. We therefore have

$$d_M(f(x), f(x+a)) \leq aK(a) \leq a(1+\eta)K(2a),$$

and also

$$d_M(f(x+a), f(x+2a)) \leq a(1+\eta)K(2a).$$

In addition, we have

$$\begin{aligned} d_M(f(x), f(x+a)) &\geq d_M(f(x), f(x+2a)) - d_M(f(x+a), f(x+2a)) \\ &\geq 2aK(2a) - a(1+\eta)K(2a) \\ &= a(1-\eta)K(2a). \end{aligned}$$

From the equations above, the result easily follows. □

We are now able to display Matoušek's proof of Theorem 1.2. First we sketch the main steps of his argument. We pick k,h adequate natural numbers such that $T_{k,h} \overset{2}{\hookrightarrow} B_n$ (according to Lemma 2.2) and consider a non-contracting mapping $f : T_{k,h} \to X$ such that $\|f\|_{\text{Lip}}$ is is smaller than our expected bound (i.e., $\|f\|_{\text{Lip}} = c_1(\log n)^{1/q}$, for c_1 small enough), we will get to an absurdity. Using cunningly Lemma 2.3 we are able to find a complete binary tree inside $T_{k,h}$ for which f embeds "identically" every path between a root to a leaf (this is the key point and is based heavily on mixing combinatorics with distortion). This fact, together with Lemma 2.4, allow us to find a 0-fork in $T_{k,h}$ (for some $a \in \mathbb{N}$) mapped by f to an δ-fork in X for δ small. But, according to Lemma 2.1, this can not happen: the tips of the 0-fork in $T_{k,h}$ are far apart.

A rigorous and detailed proof is the following.

Proof. (of Theorem 1.2) First we declare the parameters involved in the proof, their values will be fixed later. Let $\beta > 0$ be small enough (depending on q and c), suppose n is large and let k, h be natural numbers (depending on n). Fix $f : T_{k,h} \to X$ a non-contracting mapping with $\|f\|_{\mathrm{Lip}} = K = c_1(\log n)^{1/q}$ for c_1 small, we will get a contradiction.

Let $r = \left\lceil \frac{2K^{q+1}}{\beta} \right\rceil$, and suppose $k \geq r^{(h+1)^2}$. We label the pairs in $SP(T_{k,h})$ according to the distortion of their distance by f; that is to say, each pair $\{x, y\} \in SP(T_{k,h})$ is colored with the number

$$\left\lfloor \frac{K^q}{\beta} \frac{\|f(x) - f(y)\|}{d_{T_{k,h}}(x, y)} \right\rfloor \in \{0, 1, \ldots, r-1\},$$

where $d_{T_{k,h}}$ stands for the path-metric in $T_{k,h}$. By our Ramsey-type result, Lemma 2.3, we can find a subtree T', which is a copy of B_h, inside $T_{k,h}$ such that the color of each pair $\{x, y\} \in SP(T')$ depends exclusively on the levels of x and y. This is the core of Matoušek's argument: we manage to find a binary tree on which the mutual distances induced by f only depend on the position of the vertices.

Fix P a path from a root to a leaf in T' (note that this path, is isometric to $P_h = \{0, \ldots, h\} \subset \mathbb{R}$). If h is big enough, say $h = 2^{CK^q}$ where $C = C(q, \beta)$ is as in Lemma 2.4, we can pick three vertices y_0, y_1, y_2 of P whose levels form an arithmetic progression with common difference a (i.e., this vertices are at levels $l, l+a, l+2a$, respectively), such that the restriction of f to this triple becomes a $(1+\delta)$ isomorphism for

$$\delta = \beta \left(\frac{\|f(y_0) - f(y_1)\|}{a} \right)^{-q}.$$

Let y_2' be a vertex at T' at the same level as y_2 (i.e, $l+2a$) and at distant $2a$ from y_2 (note that this also implies that y_2' is at distant a and $2a$ from y_1 and y_0, respectively). By the level dependence of the colors we have that the pairs $\{y_i, y_2\}$ and $\{y_i, y_2'\}$ in $SP(T')$ are equally labeled $(i = 0, 1)$. Precisely, for $i = 0, 1$ we have

$$\left\lfloor \frac{K^q}{\beta} \frac{\|f(y_i) - f(y_2)\|}{d_{T_{k,h}}(y_i, y_2)} \right\rfloor = \left\lfloor \frac{K^q}{\beta} \frac{\|f(y_i) - f(y_2')\|}{d_{T_{k,h}}(y_i, y_2')} \right\rfloor. \tag{2}$$

This implies that the restriction of f to the triple $\{y_0, y_1, y_2'\}$ is a $(1+2\delta)$-isomorphism (a priori we can not ensure to be a $(1+\delta)$ isomorphism since the equality in Equation (2) is given only for the integer parts). Therefore, the set $\{f(y_0), f(y_1), f(y_2), f(y_2')\}$ is a 3δ-fork in X. By Lemma 2.1, we obtain

$$2a \leq \|f(y_2) - f(y_2')\| = O(\delta^{1/q} a) = O(\beta^{1/q}) \|f(y_0) - f(y_1)\| = O(\beta^{1/q} a).$$

Recall that the constant of proportionality in the last $O(\cdot)$ notation depends only on c and q (and not on β). Thus, by choosing β small enough we have a contradiction.

We have made several assumptions... It is time to see how to choose properly the parameters involved. We had $h = 2^{CK^q}$, hence if c_1 in the expression $K = c_1(\log n)^{1/q}$ is small enough, we can ensure that $h < n^{1/4}$. On the other hand, we had $k = r^{(h+1)^2}$, thus $\log_2 k = (h+1)^2 \log_2 r = O\left(\sqrt{n}(\log\log n)\right)$, therefore

$$h \log_2 k = O(n^{5/6}) < n, \tag{3}$$

for n large enough. Equation (3) and Lemma 2.2 ensures that the tree $T_{k,h}$ with which we have dealt can be embedded with distortion at most 2 into the complete binary B_n. This completes the proof. □

References

[AT06] S. A. Argyros and S. Todorcevic. *Ramsey Methods in Analysis*. Springer, 2006.

[Bar98] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 161–168. ACM, 1998.

[Bou86] J. Bourgain. The metrical interpretation of superreflexivity in banach spaces. *Israel Journal of Mathematics*, 56(2):222–230, 1986.

[BS05] Sergei Buyalo and Viktor Schroeder. Embedding of hyperbolic spaces in the product of trees. *Geometriae Dedicata*, 113(1):75–93, 2005.

[Dra03] A Dranishnikov. On hypersphericity of manifolds with finite asymptotic dimension. *Transactions of the American Mathematical Society*, 355(1):155–167, 2003.

[Dre84] A. W. Dress. Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces. *Advances in Mathematics*, 53(3):321–402, 1984.

[FRT03] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 448–455. ACM, 2003.

[Han56] O. Hanner. On the uniform convexity of l_p and l_p. *Arkiv för Matematik*, 3(3):239–244, 1956.

[JLPS02] William B Johnson, Joram Lindenstrauss, David Preiss, and Gideon Schechtman. Lipschitz quotients from metric trees and from banach spaces containing l< sub> 1</sub>. *Journal of Functional Analysis*, 194(2):332–346, 2002.

[Klo14] Benoît R Kloeckner. Yet another short proof of bourgainŠs distortion estimate for embedding of trees into uniformly convex banach spaces. *Israel Journal of Mathematics*, pages 1–4, 2014.

[LNP09] James R Lee, Assaf Naor, and Yuval Peres. Trees and markov convexity. *Geometric and Functional Analysis*, 18(5):1609–1659, 2009.

[LS03] Nathan Linial and Michael Saks. The euclidean distortion of complete binary trees. *Discrete and Computational Geometry*, 29(1):19–22, 2003.

[Mat90] Jiří Matoušek. Extension of lipschitz mappings on metric trees. *Commentationes Mathematicae Universitatis Carolinae*, 31(1):99–104, 1990.

[Mat99] J. Matoušek. On embedding trees into uniformly convex banach spaces. *Israel Journal of Mathematics*, 114(1):221–237, 1999.

[Mil38] D. Milman. On some criteria for the regularity of spaces of the type (b). *Doklady Akad. Nauk SSSR*, 20:243–246, 1938.

[MN13] Manor Mendel and Assaf Naor. Markov convexity and local rigidity of distorted metrics. *Journal of the European Mathematical Society*, 15(1):287–337, 2013.

[NPS+06] Assaf Naor, Yuval Peres, Oded Schramm, Scott Sheffield, et al. Markov chains in smooth banach spaces and gromov-hyperbolic metric spaces. *Duke Mathematical Journal*, 134(1):165–197, 2006.

[Pet39] B. J. Pettis. A proof that every uniformly convex space is reflexive. *Duke Mathematical Journal*, 5(2):249–253, 1939.

[Ram30] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 2(1):264–286, 1930.

DEPARTAMENTO DE MATEMÁTICA - PAB I, FACULTAD DE CS. EXACTAS Y NATURALES, UNIVERSIDAD DE BUENOS AIRES, (1428) BUENOS AIRES, ARGENTINA AND CONICET

E-mail address: dgalicer@dm.uba.ar