Metric Dichotomies (lecture notes)

Manor Mendel

Edinburgh, April 2007

1 Problem Statement, Motivation

For a mapping between metric spaces $f : X \to H$ we define $\|f\|_{\text{Lip}} = \max_{x \neq y} d_H(f(x), f(y))/d_X(x, y)$, and $\text{dist}(f) = \|f\|_{\text{Lip}} \cdot \|f^{-1}\|_{\text{Lip}}$ the distortion of f. We next define $c_H(X) = \inf\{\text{dist}(f) | \ f : X \to H\}$. For a class of metric spaces $\mathcal{H}$, we define $c_{\mathcal{H}}(X) = \inf_{H \in \mathcal{H}} c_H(X)$, the distortion of the "best embedding" into $\mathcal{H}$.

In this talk we will be interested in theorems of the form:

Template Theorem($\mathcal{X}$). *For any host space $\mathcal{H}$ either*

- $\forall X \in \mathcal{X}, \ c_{\mathcal{H}}(X) = 1$*; or*
- $\exists_\infty X \in \mathcal{X}$ *for which $c_{\mathcal{H}}(X)$ is "large".*

The second option is made precise by either requiring $\sup_{X \in \mathcal{X}} c_{\mathcal{H}}(X) = \infty$, or more quantitatively: The sequence $D_N(\mathcal{H}, \mathcal{X}) = \sup\{c_{\mathcal{H}}(X) : \ X \in \mathcal{X}, \ |X| \leq N\}$ increases (relatively) rapidly with N.

In this talk $\mathcal{X}$ is one of the following class of finite metric spaces:

Metric class	Regular subclass
Finite subsets of $\mathbb{R}$	Path metrics
Finite subsets of L_1	Hamming cubes
Finite metric spaces	L_∞ grids
Finite tree metrics	Binary trees

1.1 Motivation

1. In CS embedding is very useful. One approach is

   > Given a "hard" algorithmic problem on a metric $X \in \mathcal{X}$, first embed X into a simpler metric space $H \in \mathcal{H}$, $e : X \to H \in \mathcal{H}$, and then solve the algorithmic problem on $e(X)$.

   For this approach to work we need:

   (a) H should be simple enough so the algorithmic problem becomes tractable.
   (b) $e(X)$ should be close to X.

   Metric dichotomies results draw limits for this approach in terms of the possible "closeness". This motivation has been pointed out by Arora, Lovász, Newman, Rabani, Rabinovich, and Vempala [ALNRRV].

2. A dichotomy theorem for $\mathcal{X}$ can be interpreted as a form of rigidity of $\mathcal{X}$: Small deformations of all the spaces in $\mathcal{X}$ is impossible.

3. A closely related notion is that of *Bounded-Distortion Ramsey problem*, studied by Matoušek. As observed in [ALNRRV], BD-Ramsey implies rigidity (and metric dichotomy). All the metric dichotomies I'm going to mention in the talk actually give BD-Ramsey results.

2 Dichotomy of the Line metrics

For line metrics we have

Theorem 1. *For every class of metric spaces $\mathcal{H}$, one of the following happens*
- $c_{\mathcal{H}}(L) = 1$, *$\forall$ finite line metric L.*
- $\exists \beta > 0$, *such that $c_{\mathcal{H}}(P_n) \geq \Omega(n^\beta)$, where P_n is n-point path metric.*

This theorem is essentially due to Matoušek. I'll show a different proof which conveys the approach to prove more complicated dichotomies. In a sense, this is a caricature of the proofs of dichotomies for finite subsets of L_1 and L_∞. It will serves us here to illustrate the outline of the proofs for L_1 finite metrics, and for all finite metrics, since those proofs are too lengthy.

The general approach in those proofs is to define an "isomorphic" inequality, and prove sub-multiplicativity. This approach is used in Banach space theory quite often, and is probably originated from the work of Pisier and Maurey-Pisier.

The Proof

We choose an appropriate inequality that captures the distortion of embedding P_n in $\mathcal{H}$. In this case, let $L_n(\mathcal{H})$ be the smallest L such that

$$\forall f : P_n \to H \in \mathcal{H}, \quad d_H(f_0, f_{n-1}) \leq L(n-1) \max_{i=0,\dots n-2} d_H(f_i, f_{i+1}). \tag{1}$$

Lemma 2. $L_m(\mathcal{H}) \leq 1$.

Proof. Triangle inequality. □

Lemma 3. $c_{\mathcal{H}}(P_n) \geq 1/L_n(\mathcal{H})$.

Proof. Fix $f : P_n \to H$. Plugging into (1)

$$\frac{n}{\|f^{-1}\|_{\text{Lip}}} \leq d_H(f_0, f_{n-1}) \leq L_n(\mathcal{H}) n \max_{i=0,\dots n-2} d_H(f_i, f_{i+1}) \leq L_n(\mathcal{H}) n \|f\|_{\text{Lip}},$$

So $\text{dist}(f) = \|f\|_{\text{Lip}} \cdot \|f^{-1}\|_{\text{Lip}} \geq 1/L_n(\mathcal{H})$. □

Lemma 4. *If $L_n(\mathcal{H}) = 1$, then $c_{\mathcal{H}}(P_n) = 1$.*

Proof. If $L_n(\mathcal{H}) = 1$, then $\exists f : P_n \to H \in \mathcal{H}$ for which

$$d_H(f_0, f_{n-1}) = n \max_{i=0,\dots n-2} d_H(f_i, f_{i+1}),$$

and this easily implies that for all $i \in \{0, \dots, n-2\}$, $d_H(f_0, f_{n-1}) = n d_H(f_i, f_{i+1})$.

*** Strictly speaking this is inaccurate, since we are only guaranteed that for every $\varepsilon > 0$, $d_H(f_0, f_{n-1}) \geq (1-\varepsilon) n \max_{i=0,\dots n-2} d_H(f_i, f_{i+1})$. □

Now come the simple (yet powerful) insight: sub-multiplicativity.

Lemma 5. *$L_{mn}(\mathcal{H}) \leq L_m(\mathcal{H}) \cdot L_n(\mathcal{H})$.*

Proof. Fix $f : P_{mn} \to H$. Define $g : P_n \to H$, by $g(i) = f(im)$. Applying (1) to g, we obtain

$$d_H(f_0, f_{mn-1}) \leq L_n(\mathcal{H}) n \max_{i=0\dots n-1} d_H(f_{im}, f_{(i+1)m}).$$

Next, define $h_i : P_m \to H$, $h_i(j) = f(im + j)$, and apply (1) for each h_i, and so

$$d_H(f_{im}, f_{(i+1)m}) \leq L_m(\mathcal{H}) \max_{j=0\dots m-1} d_H(f_{im+j}, f_{(i+1)m+j+1}).$$ □

We can now prove the dichotomy:

Proof of Dichotomy for line metrics.

- If $\forall n \in \mathbb{N}$, $L_n(\mathcal{H}) = 1$, then $c_{\mathcal{H}}(P_n) = 1$, Note that for any finite line metric L, $c_{(P_n)_n}(L) = 1$, and so $c_{\mathcal{H}}(L) = 1$.
- If $\exists n_0$ for which $L_{n_0}(\mathcal{H}) = \eta < 1$, then let $\beta > 0$ be such that $n_0^{-\beta} = \eta$, and from the submultiplicativity, $L_{n_0^k}(\mathcal{H}) \leq \eta^k = (n_0^k)^{-\beta}$, and so $c_{\mathcal{H}}(P_{n_0^k}) \geq (n_0^k)^{\beta}$. □

BTW This dichotomy is tight in the sense that for any $\beta \in (0, 1]$, there exists H_β such that any n-point line metric L embeds with distortion $O(n^\beta)$ and $c_{H_\beta}(P_n) \geq n^\beta$.

Local Rigidity. Here is the local rigidity corollary that we will need

Proposition 6. *For any $A > 1$, $\delta \in (0, 0.5)$, and for any $t \in \mathbb{N}$, there exists $n = n(t, A, \delta) < \infty$ such that for every $\mathcal{H}$, if $c_{\mathcal{H}}(P_n) \leq A$, then $c_{\mathcal{H}}(P_t) \leq 1 + \delta$. We can take $n \approx t^{\frac{2t \log A}{\delta}}$.*

Outline of the proof. Assuming that $c_{\mathcal{H}}(P_t) > 1 + \delta$, by applying an approximate version of Lemma 4, we deduce that $L_t(\mathcal{H}) \leq 1 - \delta/2t$. Then, by the submutliplicativity, taking $n = t^k$, we have

$$c_{\mathcal{H}}(P_n) \geq L_{t^k}(\mathcal{H})^{-1} \geq (1 - \tfrac{\delta}{2t})^{-k}.$$

So choosing $k = \log A / \log(1 + \frac{\delta}{2t}) \approx \frac{2t \log A}{\delta}$, guarantee that $c_{\mathcal{H}}(P_n) > A$. □

3 Dichotomy for finite L_1 metrics

3.1 L_1 metrics (Hamming cubes)

Matoušek has shown non quantitative dichotomy for subsets of L_1. Here we give a quantitative version. It is essentially restatement of a theorem from the paper of Bourgain, Milman, & Wolfson on non-linear type.

Theorem 7. *For every class of metric spaces $\mathcal{H}$, one of the following happens*

- $c_{\mathcal{H}}(X) = 1$, $\forall$ *finite* $X \subset L_1$.
- $\exists \beta > 0$, *such that* $c_{\mathcal{H}}(\{0,1\}^n) \geq \Omega(n^\beta)$, *where* $\{0,1\}^n$ *is n-dim' Hamming cube.*

The proof outline follows the the proof of the dichotomy for line metrics, but more complicated. Here the property/inequality we use is a variant of the metric-type inequality. Let $(e_i)_{i=1}^n$ the standard basis of $\{0,1\}^n$, and $\mathbf{1} = \sum e_i$. Let $T_n(\mathcal{H})$ be the smallest T such that $\forall f : \{0,1\}^n \to H \in \mathcal{H}$,

$$\mathop{\mathbb{E}}_{x\in\{0,1\}^n} d_H(f(x), f(x+\mathbf{1}))^2 \leq T^2 n \sum_{i=1,\dots n} \mathop{\mathbb{E}}_{x\in\{0,1\}^n} d_H(f(x), f(x+e_i))^2. \tag{2}$$

As for the finite line metric, we prove

Lemma 8.

1. $T_m(\mathcal{H}) \leq 1$.
2. $c_{\mathcal{H}}(\{0,1\}^n) \geq 1/T_n(\mathcal{H})$.
3. *If* $T_n(\mathcal{H}) = 1$, *then* $c_{\mathcal{H}}(\{0,1\}^n) = 1$.
4. $T_{mn}(\mathcal{H}) \leq T_m(\mathcal{H}) \cdot T_n(\mathcal{H})$.

Now the proof of the dichotomy is the same as in the case of line, noting the any finite L_1 metric is almost isometric to subset of $\{0,1\}^n$ for some n.

We don't know whether this Dichotomy is tight: For which $\beta \in (0,1]$ there exists H_β such that $\forall X \subset L_1$, $|X| = N$, $c_{H_\beta} \leq O^*((\log N)^\beta)$, and $\exists$ L_1 metrics X_N, $|X_N| = N$, for which $c_{H_\beta}(X_N) \geq \Omega(\log^\beta N)$.

We do know that $c_{L_2}(\{0,1\}^n) = \sqrt{n}$, and due to the work of Arora, Lee, Naor; Chwalla, Gupta, Räcke; Arora, Rao, Vazirani That any n-point L_1 metric is $O(\sqrt{\log N}\log\log N)$ embeddable in L_2.

4 Dichotomy for all finite metric spaces

The BD-Ramsey theorem of Matoušek implies a dichotomy for the class of all finite metrics: Let $\mathcal{M}$ be the class of all finite metric spaces, then for every $\mathcal{H}$, either $\sup_{X\in\mathcal{M}} c_{\mathcal{H}}(X) = 1$ or $\sup_{X\in\mathcal{M}} c_{\mathcal{H}}(X) = \infty$.

We now show a stronger Dichotomy, due to A. Naor & myself.

Theorem 9. *For every class of metric spaces $\mathcal{H}$, one of the following happens*

- $\sup_{X\in\mathcal{M}} c_{\mathcal{H}}(X) = 1$.

- *$\exists\beta > 0$, such that $c_{\mathcal{H}}([n]_\infty^n) \ge \Omega(n^\beta)$, where $[m]_\infty^n$ is the $[m]^n$ grid with the L_∞ distance.*

Again we use appropriate inequality, this time metric cotype. The inequality we actually use is more complicated, as follows.

Denote by $\Gamma_n(\mathcal{H})$ the smallest Γ such that for every $m \in \mathbb{N}$, and every $f : \mathbb{Z}_m^n \to H \in \mathcal{H}$

$$\sum_{i=1}^{n} \mathop{\mathbb{E}}_{x\in\mathbb{Z}_m^n} d_H(f(x), f(x+ne_j))^2 \le \Gamma^2 \cdot n^2 \cdot n \mathop{\mathbb{E}}_{\varepsilon\in\{\pm1\}^n} \mathop{\mathbb{E}}_{x\in\mathbb{Z}_m^n} d_H(f(x), f(x+\varepsilon))^2. \qquad (3)$$

Note: This is roughly the reverse inequality for type, but on finer grid (or torus) with the "horizontal" distances scaled down by a factor of n.

Again we prove

Lemma 10.

1. *$\Gamma_n(\mathcal{H}) \le 1$, for all even n.*
2. *$c_{\mathcal{H}}([n/4]_\infty^n) \ge 1/\Gamma_n(\mathcal{H})$.*
3. *If $\Gamma_n(\mathcal{H}) = 1$, then $c_{\mathcal{H}}([n/4]_\infty^n) = 1$.*
4. *$\Gamma_{n_1 n_2}(\mathcal{H}) \le \Gamma_{n_1}(\mathcal{H}) \cdot \Gamma_{n_2}(\mathcal{H})$.*

Here the proofs are considerably more complicated (especially item 3., and 4.).

Again we conclude the dichotomy, noting that any finite metric space can be embedded in $[n/4]_\infty^n$ almost isometrically, for some large n.

5 Tree metrics

Lets summarize the approach we applied in the previous cases:

We want to prove a dichotomy result for a class of finite metric spaces $\mathcal{X}$.

1. Find a subclass $\mathcal{Y} \subset \mathcal{X}$ of highly "regular" metrics, that nonetheless satisfies $\sup_{X\in\mathcal{X}} c_{\mathcal{Y}}(X) = 1$.
2. Define a property (inequality) of class $\mathcal{H}$ of metric spaces which is "based" on the objects in $\mathcal{Y}$, and which bounds from below the distortion of $Y \in \mathcal{Y}$ when embedded in $\mathcal{H}$.
3. Prove the 4 lemmas above.

Following a question of C. Fefferman, A. Naor & myself tried to prove a dichotomy for tree metrics, according to the plan described above:

1. The natural candidate for regular subclass of metrics are the binary complete & balanced binary trees $(B_n)_{n\in\mathbb{N}}$.

2. Which inequality to use? A natural inequality that is based on binary trees is the Markov convexity inequality of Lee, Naor, & Peres, based on a proof of Bourgain. the inequality should be roughly the Markov p-convexity inequality; Let $K_{n,h}(\mathcal{H})$ be the smallest K, such that for every $f : B_n \to H \in \mathcal{H}$

   $$\sum_{k=1}^{h} \mathop{\mathbb{E}}_{t=1,\dots,n} \mathbb{E} \left(\frac{d_H(f(X_t), f(\widetilde{X}_t(t-2^k)))}{2^k} \right)^2 \leq K^2 \cdot h \cdot \mathop{\mathbb{E}}_{t=1,\dots,n} \mathbb{E}\, d_H(f(X_t), f(X_{t-1}))^2,$$

   Where $(X_t)_t = 1, \dots n$ is a directed random on B_n starting from the root downward, $\widetilde{X}_t(a)$ is a also random walk on B_n from the root downward and such that: (i) $X_t = X_t(a)$ for $t < a$, (ii) $X_a(a) \neq x_a$, and (iii) X_t, and $X_t(a)$ evolve according to independent random variable when $t > a$.

   A Markov p-convexity has been shown to hold in uniform p-convex Banach spaces. To add credibility to this approach (and of independent interest), we proved the reverse direction and those established

   Theorem 11. *A Banach space is Markov p-convex iff it is isomorphic to uniform p-convex space.*

3. Proving analogue for the 4 lemmas above:

   (a) $K_{n,h}(\mathcal{H}) \leq 1$ is easy to prove.

   (b) $c_{\mathcal{H}}(B_h) \geq a/K_{4h,h}(\mathcal{H})$, for some universal constant $a > 0$ is also easy.

   (c) If $\sup_n K_{h,n}(\mathcal{H}) = 1$ then $c_{\mathcal{H}}(B_h) = 1$: Probably correct.

   (d) What about an analog to sub-multiplicativity? We worked hard, but couldn't prove it.

As it turned out, tree metrics do not have the relevant rigidity

Theorem 12. *For any $B \geq 1$, there exists a metric space H such that $\sup_T c_H(T) = B$, where T ranges over the finite tree metrics.*

Furthermore,

Theorem 13. *For any $\delta \in (0, 0.001)$, and for any sequence (s_n) satisfying (i) $s(n)$ is non decreasing; (ii) $s(n)/n$ is non-increasing (iii) $s(n) \leq O(\delta \log n / \log \log n)$, there exist a metric space $H$and n_0 such that for every $n \geq n_0$, $(1-\delta)s(n) \leq c_H(B_n) \leq s(n)$.*

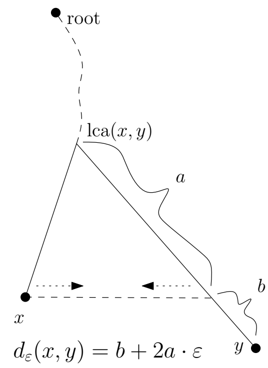
Theorem 12 is a corollary of Theorem 13, when substituting $s(n) = B$,
From the perspective of Banach space theory this result is somewhat surprising, because when the host space H is a Banach space then there is a dichotomy:

- Either H is isomorphic to a uniformly convex space (superreflexiv), and then by Pisier this space has modulus of unform convexity of the form ε^p, which implies by a theorem of Bourgain/Matoušekthat $c_H(B_n) \geq (\log n)^{1/p}$,
- Or H is not superreflexive. In this case Bourgain has shown that H contains biLipschitz all finite trees, but his proof actually give distortion $1 + \varepsilon$, for every $\varepsilon > 0$.

5.1 Proof of Theorem 13

We take the infinite Binary tree B_∞, with the tree metric on that tree, and contract the "horizontal" distances. Let $h(x)$ be the depth of $x \in B_\infty$, i.e. the distance from the root of B_∞. Pick $\varepsilon \in (0,1]$, and the distance $d_\varepsilon(x,y)$, for $x, y \in B_\infty$ is defined as

$$d_\varepsilon(x,y) = h(u) - h(x) + 2(h(x) - h(\text{lca}(x,y)) \cdot \varepsilon. \tag{4}$$



It's not hard to check that

1. d_ε is a metric.

2. $c_{d_\varepsilon}(B_\infty) \leq 1/\varepsilon$. Indeed the identity mapping does not expand distance, and contracts them by at most $1/\varepsilon$.

Thus d_ε is our candidate host space H for proving Theorem 12.

To prove Theorem 13, the contraction should get larger as we embed larger and larger binary trees.

We generalize the construction a bit. Now ε is a sequence $\varepsilon = (\varepsilon_n)_{n=0}^{\infty}$ of numbers $\varepsilon_n \in (0,1]$ satisfying (i) ε_n non increasing, and (ii) $n\varepsilon_n$ non-decreasing. the distance $d_\varepsilon(x,y)$, for $x, y \in B_\infty$ is defined as

$$d_\varepsilon(x,y) = h(u) - h(x) + 2(h(x) - h(\text{lca}(x,y)) \cdot \varepsilon_{h(x)}.$$

Again from the definition it's clear that the identity mapping gives $c_{d_\varepsilon}(B_n) = 1/\varepsilon_n$. It can also be proved that d_ε is a metric, although this requires more work. The rest of talk will discuss how to prove that $c_{d_\varepsilon}(B_n) \geq \Omega(1/\varepsilon_n)$.

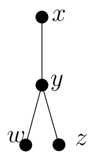
Our approach follows Matoušek's proof of

Theorem 14. $C_{L_2}(B_n) \geq \Omega(\sqrt{\log n})$.

As we have seen before this theorem was first proved by Bourgain, and his proof motivated the definition of Markov p-convexity.

For the current purposes Matoušek's approach seems more suitable. His proof's outline is as follows:

Define a δ-fork to be a quadruple (x, y, z, w) such that both (x, y, z), and (x, y, w) are $1 + \delta$ equivalent to $(0, 1, 2)$. It's not hard to see that in Hilbert space (and more generally, uniform 2-convex spaces), if (x, y, z, w) is a δ-fork then $\|z - w\| \leq O(\sqrt{\delta})\|x - y\|$. Matoušek's approach is to assume toward a contradiction that $\exists f : B_n \to L_2$ such that $\mathrm{dist}(f) \leq c\sqrt{\log n}$, and use this assumption to find a 3-leaves star (x, y, z, w) in B_n whose center is y such that $(f(x), f(y), f(z), f(w))$ is a δ fork for $\delta \approx 1/\log n$, and this will imply a large contraction of the distance between z and w.



Finding a δ-fork in $f(B_n)$: Apply Ramsey-type theorems as follows

1. The first step is to find in $C \subset B_n$ which is a $(1 + \delta)$ metric approximation of $B_{n'}$, where n' is not much smaller than n, and such that in C for every to vertical paths p,q from the root to a leaf, $f(p)$, and $f(q)$ have the same metric up to multiplicative factor of $1 + \delta/10$.

2. Now we have that every "vertical" path p in $B_{n'}$ looks the same as $f(p)$. We also knows that it is A approximation of the path metrics, so applying the "local rigidity" of the path we can find in $f(p)$ a 3 point path metric (x, y, z), such that $(f(x), f(y), f(z))$ is $1 + \delta/10$ approximation of $(0, 1, 2)$ (provided that n' is large enough).

   Now, let w be another descendant of y of the same height as z such that $y = \mathrm{lca}(z, w)$. The metric (x, y, w) is also $1 + \delta/10$ close to $(0, 1, 2)$, so (x, y, z, w) is δ fork.

Note that the part of finding a δ-fork is independent of the range of the embedding. We use the geometric properties of the range (uniform convexity) only to argue about large contractions in δ-forks. Hence it make sense to try this approach on the $(B_\infty, d_\varepsilon)$.

To this end, we need to show that every δ fork in $(B_\infty, d_\varepsilon)$ has large contraction.

1. Indeed, δ forks that looks like case (p$\|$p) in Fig. 1 has by the design of d_ε, a contraction of $1/(O(\delta) + \varepsilon_h)$.
2. However, this is not the only way a δ-fork can be embedded in d_ε. Case (t$\|$t) in Fig. 1 is another type of embedding. Fortunately here the contraction is even larger: $1/O(\delta)$.
3. However this is not the end of story! there other types of δ forks embedded the d_ε. For example type II in Fig. 2, can be made even 0-fork, but with very small contraction of the tips! This means that the approach that attempts to show large contract of δ forks is simply not true in d_ε..
4. This is not yet the end of the bad news, there are more bad forks, such as type I, III, and IV in FIg. 2.

At this point the whole business seems hopeless.

However it turns out that the situation is not that bad. The four types of "bad forks" are the only ones that exist!

Lemma 15. *Every δ fork in $(B_\infty, d_\varepsilon)$ is close to one of the 6 types of forks in Figures 1 and 2, up to distortion of* $1 + O(\delta)$.

To show that we first analyze approximate $\{0, 1, 2\}$ metric. Checking it carefully we find that there are only two different possible approximate $(0, 1, 2)$ configurations, as

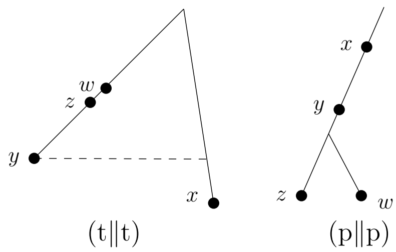


Figure 1: Forks with tips contracted

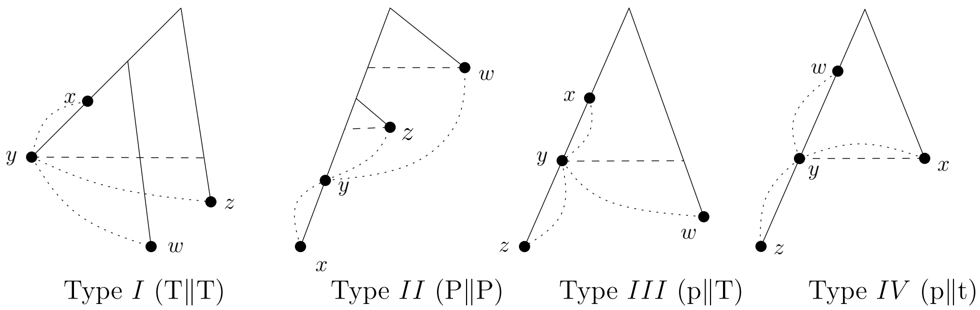


Figure 2: Forks in which the tips do not contract

depicted in Fig. 3. These two configuration are actually four if we also count symmetric configurations, and therefore there are $\binom{5}{2} = 10$ plausible ways to connect two $\{0, 1, 2\}$ metric into forks. Four of them turn out to be impossible, and we are left with the 6 depicted in Figures 1 and 2.

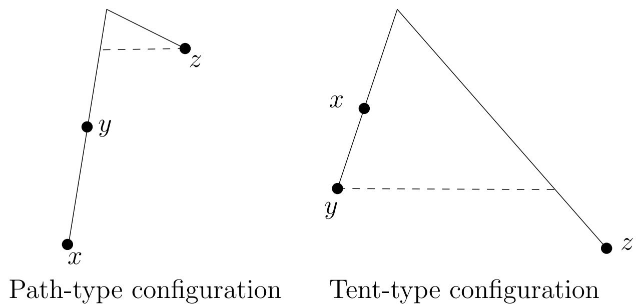


Figure 3: Forks in which the tips do not contract

Having this small set of problematic forks, we now want to prove a generalization of δ fork embedding.

Definition 1. *An embedding of a rooted tree will be called* $1 + \delta$ *vertically faithful if all vertical distances in the tree are preserved up to distortion* $1 + \delta$*.*

Matoušek's argument above showed that any mild embedding of B_n for large enough n gives $1+\delta$ vertically faithful embedding of B_2, But we can use the local rigidity argument to conclude (for a bit larger n) that it is also implies vertically faithful embedding of B_5. We now show that vertically faithful embedding of B_4 in d_ε necessarily has a large contraction.

Lemma 16. *Any* $1+\delta$ *vertically faithful embedding of* B_4 *in* $(B_\infty \setminus B_h, d_\varepsilon)$ *must be distorted by at least* $\min\{O(\delta^{-1}), \varepsilon_h^{-1}\}$.

The proof idea is to list the constrains inherited from concatenating forks with an overlap of an edge, and find out that after 4 such concatenations, we must encounter a contracting fork. The analysis begins by analyzing concatenation of two $\{0,1,2\}$ metric (2-path) into $\{0,1,2,3\}$ metric (3-path) with an overlapping edge. Again there are 10 plausible way to do that (not counting symmetries), and only three can be realized geometrically. These are depicted in Fig. 4

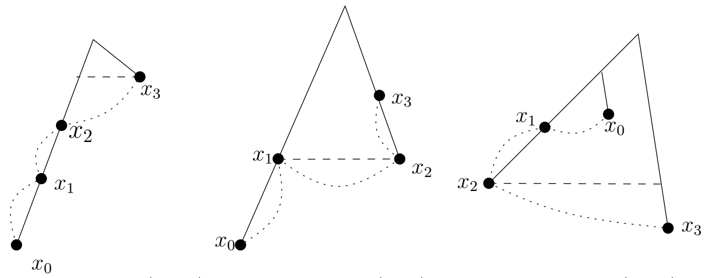


Figure 4: The 3 possible types of 3-paths.

With these classifications we now reach the conclusion by syntactic case analysis.

- Suppose first that a top fork is of type I or III. One of its paths is of tent-type. The only 3-path type that begins with tent-type 2-path is concatenated to reverse path-type 2-path. Hence, the from the tent-type leg, the only fork type that can be formed is built of two path-type, but this type of fork is contracting.
- Next suppose that a top fork is of type IV. Then one of its legs is reverse path-type configuration. There are two types of 3-paths that begins with tent-type 2-path: One end with reverse path-type, the other ends with tent-type. There are three forks that can be made of gluing together two out of reverse path-type and tent-type:
  - gluing two reverse path-types, gives contracting fork.
  - gluing two tent-types, gives fork of type I, which leads to contraction as above.
  - gluing tent-type with reverse path type, gives fork of type III, which leads to contraction as above.
- The last option, suppose the top fork is of type II (two path type glued together). There are two 3-path types that begins with path-type. One ends with path-type, the other with reverse tent-type. There are two forks types that can be formed from gluing together two out of path-type and reverse tent-type:
  - Two reverse tent-type gives a contracting fork.
  - two path-type gives type II. Lets (x,y,z,w) be the top fork of type II. Let (y,z,v_1,v_2), and (y,w,u_1,u_2) be the concatenated fork, also of type II. Since (y,z,v_1,v_2) is of type II, z is an ancestor of y, and since (y,w,u_1,u_2) is type II, w is an ancestor of y, and since $d(y,z) \approx d(y,w)$, we conclude that $d(z,w) \approx 0$, which means large contraction.

6 Some Open Questions

- Most interesting: Tightness of the metric dichotomy.
  - Currently we know: For every host space H, the distortion is either 1, or $\geq \log^{\beta} n$, for some $\beta = \beta(H) \in (0, 1]$.
  - However all H we know how to analyze and which the distortion is not 1, we have a lower bound of $\Omega(\log n)$.
  - What's the right answer?

  Both answers seem to require interesting machinery. Improving the dichotomy to $\Omega(\log n)$ may require universal badly embeddable metrics. Expanders are natural candidate here. Can their inemebddability be generalized? Would this give new (combinatorial) properties of expanders? On the other hand, a metric space such that any n-point metric embeds in it with distortion $O(\log^{0.9} n)$ (but not $1 + \varepsilon$) would be a new creature in the neighborhood.

- Beyond the first gap, it would be nice to map all accessible distortions of the class of all finite metrics. For example, there exist $0 < \varepsilon_1 < \varepsilon_2 < 1$ and host space H such that any n-point metric space embeds in H with distortion at most n^{e_2}, and there are n-point metrics that embed with distortion at least n^{ε_1}. (Think of H being constant dimensional Euclidean space). Is it true for any $0 < \varepsilon_1 < \varepsilon_2 \leq 1$? Is there a space such that any n-point metric space embeds with distortion $O(\log^2 n)$, but there exists metrics whose distortion is $\omega(\log n)$?

- All those questions can be asked for other class of metric spaces, of course.