

Testing Embeddability between Metric Spaces

Ching-Lueh Chang¹

Yuh-Dauh Lyuu²

Yen-Wu Ti³

¹ Department of Computer Science and Information Engineering
National Taiwan University,
Taipei, Taiwan,
Email: d95007@csie.ntu.edu.tw

² Department of Computer Science and Information Engineering
National Taiwan University,
Taipei, Taiwan,
Email: lyuu@csie.ntu.edu.tw

³ Department of Computer Science and Information Engineering
National Taiwan University,
Taipei, Taiwan,
Email: d91010@csie.ntu.edu.tw

Abstract

Let $L \geq 1, \epsilon > 0$ be real numbers, (M, d) be a finite metric space and (N, ρ) be a metric space (Rudin 1976). The metric space (M, d) is said to be L -bilipschitz embeddable into (N, ρ) if there is an injective function $f : M \rightarrow N$ with

$$1/L \cdot d(x, y) \leq \rho(f(x), f(y)) \leq L \cdot d(x, y)$$

for all $x, y \in M$ (Farb & Moshier 1999, David & Semmes 2000, Croom 2002). In this paper, we also say that (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) if the above inequality fails on at least an ϵ fraction of pairs $(x, y) \in M \times M$ for every injective function $f : M \rightarrow N$.

Below, a query to a metric space consists of asking for the distance between a pair of points chosen for that query. We study the number of queries to metric spaces (M, d) and (N, ρ) needed to answer whether (M, d) is L -bilipschitz embeddable into (N, ρ) or ϵ -far from being L -bilipschitz embeddable into (N, ρ) . When (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) , we allow an $o(1)$ probability of error (i.e., returning the wrong answer “ L -bilipschitz embeddable”). However, we allow no error when (M, d) is L -bilipschitz embeddable into (N, ρ) . That is, algorithms with only one-sided errors are considered in this paper. When $|M| \leq |N|$ are finite,

we give an upper bound of $O(\sqrt{\frac{\ln |N|}{\epsilon |M|}} (|M|^2 + |N|^2))$ on the number of queries for determining with one-sided error whether (M, d) is L -bilipschitz embeddable into (N, ρ) or ϵ -far from being L -bilipschitz embeddable into (N, ρ) . For the special case of finite $|M| = |N|$, the above upper bound evaluates to $O(|N|^{3/2} \sqrt{\frac{\ln |N|}{\epsilon}})$. We also prove a lower bound of $\Omega(|N|^{3/2})$ even for the special case when $|M| = |N|$

are finite and $L = 1$, which coincides with testing isometry between finite metric spaces (Croom 2002). For finite $|M| = |N|$, the upper and lower bounds thus

match up to a multiplicative factor of at most $\sqrt{\frac{\ln |N|}{\epsilon}}$, which depends only sublogarithmically in $|N|$. We also investigate the case when (N, ρ) is not necessarily finite. Our results are based on techniques developed in an earlier work on testing graph isomorphism (Fischer & Matsliah 2006).

1 Introduction

The ability to analyze metric spaces is of growing importance across diverse disciplines as huge bodies of data await analysis. In bioinformatics, for example, enormous amounts of data such as DNA sequences and protein sequences are constantly being produced. Efficient algorithms and standard computer programs have been developed over the years for calculating the distances between DNA or protein sequences, and this turns the collection of all known DNA and protein sequences into a huge metric space. As pointed out by Linial (Linial 2002), proper analysis of this space is of great significance to the biological sciences.

The analysis of metric spaces often requires various notions of similarity and embeddability between metric spaces. The philosophy is that, when a metric space is embedded into another metric space such that the original space is similar to the embedded one, understandings of the original space may be achieved through analysis of the embedded one. For example, when a metric space is embedded into the Euclidean plane while roughly preserving the distances between pairs of points, many efficient geometric algorithms that are not available for general metric spaces become applicable and of great help (Goodman & O’Rourke 2004). Another advantage is that embeddings into the Euclidean plane make possible more succinct representations of the original space (Goodman & O’Rourke 2004). In respect of these benefits of metric embedding that preserves similarity between the original and the embedded spaces, it comes without surprise that such embeddings have found tremendous applications in graph theory, combinatorial optimization, learning theory and computational geometry (Deza & Laurent 1997, Gupta 2000, Linial 2002, Matoušek 2002, Johnson & Lindenstrauss 2003, Indyk 2001, Kenyon et al. 2004). Besides, studies on the structures of metric spaces

Research supported in part by NSC grant 95-2213-E-002-044.

Copyright ©2008, Australian Computer Society, Inc. This paper appeared at the Fourteenth Computing: The Australasian Theory Symposium (CATS2008), Wollongong, NSW, Australia. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 77, James Harland and Prabhu Manyem, Ed. Reproduction for academic, not-for profit purposes permitted provided this text is included.

also have applications to isometry groups and mathematical biology (Ganyushkin et al. 1994, Ganyushkin & Tsvirkunov 1994), quadratic forms and phylogenetic analysis (Dress et al. 2001), theory of quaternions (Weston 2001), scalable biological databases (Miranker 2003), sequences homology (Mao et al. 2005) and approximate string matching (Chávez & Navarro 2006), to name a few.

Unfortunately, as we will see in Section 3, it is computationally hard to determine whether one metric spaces is similar to another one under various notions of metric similarity. First, deciding whether two input metric spaces are isometric (Croom 2002) is as hard as the graph isomorphism problem (Papadimitriou 1994), for which no polynomial-time algorithms are known despite extensive research. Second, consider the problem of deciding, on input $L \geq 1$ and finite metric spaces (M, d) and (M, ρ) , whether or not these spaces are L -bilipschitz equivalent (Farb & Mosher 1999, David & Semmes 2000). That is, we want to decide whether the metric spaces (M, d) and (M, ρ) exhibit a bijective map between them that preserves distances up to multiplicative factors ranging from $1/L$ to L . We observe that the results of Kenyon, Rabani and Sinclair (Kenyon et al. 2004) imply that it is hard even to approximate the least value of L such that (M, d) and (M, ρ) are L -bilipschitz equivalent. This may be interpreted as saying that it is hard to approximately compute the level of bilipschitz similarity even between finite metric spaces with the same ground set. Given the above hardness results, a randomized approximation algorithm with a reasonable complexity can be an attractive alternative to attack the problem of determining metric similarity or even metric embeddability.

An algorithm in the flavor of property testing (Fischer 2001) is one such alternative. It determines whether a problem instance has a certain property or is ϵ -far (under a certain distance measure) from having such a property, while allowing a small probability of error. In this paper, we seek an algorithm T that, when given as input $\epsilon > 0, L \geq 1$ and given oracle access to finite metric spaces (M, d) and (N, ρ) with $|M| \leq |N|$, has the following two properties. First, T accepts if

$$1/L \cdot d(x, y) \leq \rho(f(x), f(y)) \leq L \cdot d(x, y)$$

holds for some injection $f : M \rightarrow N$ and all $(x, y) \in M \times M$, that is, T accepts if (M, d) is L -bilipschitz embeddable into (N, ρ) (Farb & Mosher 1999, David & Semmes 2000, Croom 2002). Second, T rejects with high probability if the above inequality fails on at least an ϵ fraction of pairs $(x, y) \in M \times M$ for every injection $f : M \rightarrow N$. Such an algorithm T is called a one-sided tester for bilipschitz embeddability in this paper. Its query complexity is measured in terms of the number of times that it queries the metric spaces, where each query asks for the distance between a pair of points chosen for that query.

We give a one-sided tester for bilipschitz embeddability with query complexity at most $O(\sqrt{\frac{\ln |N|}{\epsilon |M|}} (|M|^2 + |N|^2))$. We also show an $\Omega(|N|^{3/2})$ lower bound on the query complexity of any one-sided tester for bilipschitz embeddability even for the special case of finite $|M| = |N|$ and $L = 1$. If (N, ρ) is known in advance, queries need only go to (M, d) and the query complexity is shown to be $O(\frac{|M| \ln |N|}{\epsilon})$. Our results utilize techniques developed by Fischer and Matsliah (Fischer & Matsliah 2006) in an earlier work on testing graph isomorphism.

We also give an extension to the case where the metric space (N, ρ) is known in advance but is not necessarily finite. When (N, ρ) is a totally bounded

(Croom 2002) metric space known a priori, we devise an algorithm that in a technical sense tests whether a finite metric space (M, d) is (κ, C) quasi-isometrically embeddable (Ghys & de la Harpe 1991, Farb 1997, Farb & Mosher 1999, 2000) into (N, ρ) , for input parameters $\kappa \geq 1$ and $C \geq 0$. The exact statement of this result is given in Section 6.

This paper is organized as follows. Section 2 gives the definitions. Section 3 gives the hardness results, which motivate switching to a property-testing flavored model. Sections 4-5 present upper bound and lower bounds on the query complexity of one-sided testers for bilipschitz embeddability. Section 6 extends the results to testing embeddability of a finite metric space into a totally bounded metric space. Section 7 discusses definitional issues and concludes the paper.

2 Definitions

Let S be an arbitrary set and t be a positive integer. We write S^t for the t -dimensional Cartesian product of S , and any pair $(x, y) \in S \times S$ is understood as an ordered pair unless otherwise specified. A function $d_S : S \times S \rightarrow \mathbb{R}$ is a metric on S if for all $x, y, z \in S$, we have $d_S(x, y) \geq 0$, $d_S(x, y) = 0$ if and only if $x = y$, $d_S(x, y) = d_S(y, x)$ and $d_S(x, y) \leq d_S(x, z) + d_S(z, y)$. A metric space is a set (called its ground set) endowed with a metric on it (Rudin 1976).

Let $L \geq 1$, (M, d) be a finite metric space and (N, ρ) be a metric space. We say that (M, d) is L -bilipschitz embeddable into (N, ρ) if there is an injective function $f : M \rightarrow N$ satisfying

$$1/L \cdot d(x, y) \leq \rho(f(x), f(y)) \leq L \cdot d(x, y) \quad (1)$$

for all $(x, y) \in M \times M$ (Apostol 1974, Farb & Mosher 1999, David & Semmes 2000). Clearly, Eq. (1) could also be written equivalently as

$$1/L \cdot \rho(f(x), f(y)) \leq d(x, y) \leq L \cdot \rho(f(x), f(y)).$$

In this paper, we also say that (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) if, for every injection $f : M \rightarrow N$, there are at least $\epsilon |M|^2$ pairs $(x, y) \in M \times M$ violating Eq. (1). Similarly, for $\kappa \geq 1$ and $C \geq 0$, we say that (M, d) is (κ, C) quasi-isometrically embeddable into (N, ρ) if there is a function $f : M \rightarrow N$ satisfying

$$1/\kappa \cdot d(x, y) - C \leq \rho(f(x), f(y)) \leq \kappa \cdot d(x, y) + C \quad (2)$$

for all $(x, y) \in M \times M$ (Ghys & de la Harpe 1991, Farb 1997, Farb & Mosher 1999, 2000). If for every function $f : M \rightarrow N$, Eq. (2) fails on at least $\epsilon |M|^2$ pairs $(x, y) \in M \times M$, then (M, d) is said to be ϵ -far from being (κ, C) quasi-isometrically embeddable into (N, ρ) .

For finite $|M| = |N|$, we say that (M, d) is L -bilipschitz equivalent to (N, ρ) if (M, d) is L -bilipschitz embeddable into (N, ρ) (Farb & Mosher 1999, David & Semmes 2000). Since for finite $|M| = |N|$, every injection from M to N is a bijection, it is easy to see that (M, d) is L -bilipschitz equivalent to (N, ρ) if and only if Eq. (1) holds for some bijection $f : M \rightarrow N$ and all $(x, y) \in M \times M$. Clearly, L -bilipschitz equivalence is a reflexive and symmetric relation between metric spaces. For finite $|M| = |N|$, the minimum value of $L \geq 1$ for which (M, d) and (N, ρ) are L -bilipschitz equivalent can be thought of as a measure on the similarity between (M, d) and (N, ρ) . The smaller this value, the more similar the metric spaces are. In the extreme case, (M, d) and (N, ρ) are 1-bilipschitz equivalent if and only if

they are isometric, that is, there exists a distance-preserving bijective map (called an isometry) between them (Croom 2002). For $\epsilon > 0$ and finite $|M| = |N|$, we say that (M, d) and (N, ρ) are ϵ -far from being L -bilipschitz equivalent if (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) . This is the same as saying that Eq. (1) fails on at least an ϵ fraction of pairs $(x, y) \in M \times M$ for every bijection f . If (M, d) and (N, ρ) are ϵ -far from being 1-bilipschitz equivalent, they are said to be ϵ -far from being isometric.

When a metric space is given as an oracle, it means that we can query the oracle for the distance between any pair of points. Given as input $L \geq 1, \epsilon > 0$, positive integers $m \leq n$ and given oracle access to finite metric spaces $(M, d), (N, \rho)$ with $|M| = m$ and $|N| = n$, we are interested in the number of queries (to (M, d) and (N, ρ)) needed to determine whether (M, d) is L -bilipschitz embeddable into (N, ρ) or ϵ -far from being L -bilipschitz embeddable into (N, ρ) . In particular, we seek an algorithm T that accepts when (M, d) is L -bilipschitz embeddable into (N, ρ) , and rejects with high probability when (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) . Such an algorithm T is said to be a one-sided tester for bilipschitz embeddability in this paper. Similarly, an algorithm is a one-sided tester for bilipschitz equivalence (respectively, isometry) if, when we restrict to finite $|M| = |N|$, it accepts when (M, d) and (N, ρ) are L -bilipschitz equivalent (respectively, isometric) and rejects with high probability when (M, d) and (N, ρ) are ϵ -far from being L -bilipschitz equivalent (respectively, isometric). Finally, a one-sided tester for quasi-isometric embeddability is given as input $\kappa \geq 1, C \geq 0$, positive integers $m \leq n$ and given oracle access to metric spaces (M, d) and (N, ρ) with $|M| = m$ and $|N| = n$. It is required to accept if (M, d) is (κ, C) quasi-isometrically embeddable into (N, ρ) and reject with high probability if (M, d) is ϵ -far from being (κ, C) quasi-isometrically embeddable into (N, ρ) .

For $\epsilon > 0$, positive integers $m \leq n$ and a one-sided tester T for bilipschitz embeddability, the query complexity of T with respect to ϵ, m and n is its worst-case number of queries when it is given ϵ, m, n , any $L \geq 1$ and oracle access to any metric spaces (M, d) and (N, ρ) with $|M| = m$ and $|N| = n$. Here the worst case is taken over all $L \geq 1$ and all metric spaces (M, d) and (N, ρ) (of sizes m and n) given as oracles. The query complexity (with respect to ϵ, m and n) of a one-sided tester for isometry is defined similarly except that L is fixed to 1 and m is fixed to equal n .

Let $G_1 = (V, E_1)$ and $G = (V, E_2)$ be undirected simple graphs (West 2001). An isomorphism between G_1 and G_2 is a bijection $\pi : V \rightarrow V$ such that for all $x, y \in V$, we have $(x, y) \in E_1$ if and only if $(\pi(x), \pi(y)) \in E_2$ (Papadimitriou 1994). The graph isomorphism problem asks whether two undirected simple graphs exhibit an isomorphism between them (Papadimitriou 1994). For $\epsilon > 0$, we say that G_1 and G_2 are ϵ -far from being isomorphic if for every bijection $\pi : V \rightarrow V$, there are at least $\epsilon \binom{|V|}{2}$ unordered pairs $(x, y) \in V \times V$ such that $(x, y) \in E_1$ but $(\pi(x), \pi(y)) \notin E_2$, or $(x, y) \in E_2$ but $(\pi(x), \pi(y)) \notin E_1$ (Fischer & Matsliah 2006).

When an algorithm is given oracle access to an undirected simple graph $G = (V, E)$, it means that the algorithm may query the oracle on any $(x, y) \in V \times V$ and be informed of whether $(x, y) \in E$. A one-sided tester for graph isomorphism receives as input $\epsilon > 0$, a positive integer n and is given oracle access to two undirected simple graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ with $|V| = n$. It must accept if G_1 is isomorphic to G_2 and reject with high probability if G_1 is ϵ -far from being isomorphic to G_2 .

3 Hardness

In this section, we show that the problem of deciding whether two input metric spaces are isometric is polynomial-time reducible to and from the graph isomorphism problem, for which no polynomial-time algorithm has been known despite extensive research. Furthermore, we show that it is hard even to approximate the least $L \geq 1$ for which two input finite metric spaces are L -bilipschitz equivalent. In contrast to these hardness results, we will show in the Section 4 that there is an efficient one-sided tester for bilipschitz embeddability.

We state the following theorem. For its proof please refer to Appendix I.

Theorem 1. *The problem of testing whether two input metric spaces with the same finite ground set are isometric is polynomial-time reducible to and from the graph isomorphism problem.*

The problem of approximating the least $L \geq 1$ for which two input finite metric spaces with the same ground set are L -bilipschitz equivalent is even harder, provided that $\mathcal{NP} \neq \mathcal{P}$. This is stated in the following theorem, which is implicit in the work of Kenyon, Rabani and Sinclair (Kenyon et al. 2004) (see Proposition 2.2 and Proposition 2.4 in their paper).

Theorem 2. ((Kenyon et al. 2004)) *If there is an algorithm that, on input any two finite metric spaces (M, d) and (M, ρ) , outputs a number $L^* \geq 1$ such that (M, d) is L^* -bilipschitz equivalent to (M, ρ) and*

$$L^* < \sqrt{\frac{4}{3}} \cdot \min\{L \geq 1 \mid (M, d) \text{ is } L\text{-bilipschitz equivalent to } (M, \rho)\},$$

then $\mathcal{NP} = \mathcal{P}$.

That is, it is hard to approximate to within a multiplicative $\sqrt{4/3}$ the minimum value of $L \geq 1$ for which two input finite metric spaces are L -bilipschitz equivalent.

4 An upper bound on the query complexity

In this section, we give a one-sided tester for bilipschitz embeddability. Clearly, this also gives one-sided testers for bilipschitz equivalence and isometry. For convenience, we make the following definition.

Definition 1. *Let $L \geq 1, (M, d)$ be a finite metric space, (N, ρ) be a metric space and $f : M \rightarrow N$ be a function. A quadruple $(x, y, u, v) \in M^2 \times N^2$ refutes f for the L -bilipschitz embeddability of (M, d) into (N, ρ) if $u = f(x), v = f(y)$ but*

$$1/L \cdot d(x, y) \leq \rho(f(x), f(y)) \leq L \cdot d(x, y)$$

fails to hold. A set $S \subseteq M^2 \times N^2$ of quadruples refutes f for the L -bilipschitz embeddability of (M, d) into (N, ρ) if at least one element of S does. When $L, (M, d)$ and (N, ρ) are clear from the context, we may simply say that a quadruple $(x, y, u, v) \in M^2 \times N^2$ or a set of quadruples refutes f without explicitly referring to $L, (M, d)$ and (N, ρ) .

The following lemma states that the algorithm TEST-BILIP in Figure 1 is a one-sided tester for bilipschitz embeddability.

Lemma 3. *On input $L \geq 1, \epsilon > 0$, positive integers $m \leq n$ and given oracle access to finite metric spaces (M, d) and (N, ρ) with $|M| = m$ and $|N| = n$,*

```

1: if  $\frac{\ln n}{\epsilon m} \geq 1/4$  then
2:   Query  $(M, d)$  and  $(N, \rho)$  for the distances between all pairs of points;
3:   if  $(M, d)$  is  $L$ -bilipschitz embeddable into  $(N, \rho)$  then
4:     Accept;
5:   else
6:     Reject;
7:   end if
8: else if  $2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} \geq 1$  then
9:    $p_M \leftarrow 1$ ;
10:   $p_N \leftarrow 4 \cdot \frac{\ln n}{\epsilon m}$ ;
11: else
12:   $p_M \leftarrow 2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}}$ ;
13:   $p_N \leftarrow 2 \cdot \frac{m}{n} \sqrt{\frac{\ln n}{\epsilon m}}$ ;
14: end if
15: Construct  $Q_M \subseteq M \times M$  by choosing each pair in  $M \times M$  into  $Q_M$  independently with probability  $p_M$ ;
16: Construct  $Q_N \subseteq N \times N$  by choosing each pair in  $N \times N$  into  $Q_N$  independently with probability  $p_N$ , using random coin tosses independent from those used to construct  $Q_M$ ;
17: if  $|Q_M| > 1000 p_M m^2$  or  $|Q_N| > 1000 p_N n^2$  then
18:   Accept without making any queries;
19: else
20:   Query every element of  $Q_M$  to  $(M, d)$ ;
21:   Query every element of  $Q_N$  to  $(N, \rho)$ ;
22:   if all injections from  $M$  to  $N$  are refuted by  $Q_M \times Q_N \subseteq M^2 \times N^2$  then
23:     Reject;
24:   else
25:     Accept;
26:   end if
27: end if

```

Figure 1: Algorithm TEST-BILIP. The inputs are $L \geq 1, \epsilon > 0$ and positive integers $m \leq n$. The metric spaces (M, d) and (N, ρ) are given as oracles and satisfy $|M| = m$ and $|N| = n$.

TEST-BILIP *accepts* if (M, d) is L -bilipschitz embeddable into (N, ρ) , and *rejects* with high probability if (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) .

Proof. If $\frac{\ln n}{\epsilon m} \geq 1/4$, then TEST-BILIP does exhaustive queries and accepts exactly when (M, d) is L -bilipschitz embeddable into (N, ρ) . Hence, we may assume that

$$\frac{\ln n}{\epsilon m} < 1/4 \quad (3)$$

in the following.

It is clear that TEST-BILIP accepts whenever (M, d) is L -bilipschitz embeddable into (N, ρ) .

Now assume that (M, d) is ϵ -far from being L -bilipschitz embeddable into (N, ρ) and let $f : M \rightarrow N$ be an arbitrary injection. Denote by S_f the set of all pairs $(x, y) \in M \times M$ for which the inequality

$$1/L \cdot d(x, y) \leq \rho(f(x), f(y)) \leq L \cdot d(x, y)$$

fails to hold. By assumption we have $|S_f| \geq \epsilon |M|^2 = \epsilon m^2$. For any $(x, y) \in S_f$, the probability,

taken over the random coin tosses of TEST-BILIP, that both $(x, y) \in Q_M$ and $(f(x), f(y)) \in Q_N$ is $p_M p_N$ (although there are two possible assignments to p_M and p_N by TEST-BILIP). Now write $S_f = \{(x_1, y_1), \dots, (x_t, y_t)\}$. Since f is injective, the pairs $(f(x_1), f(y_1)), \dots, (f(x_t), f(y_t))$ are different. Hence, the $2t$ events

$$\begin{aligned} (x_1, y_1) &\in Q_M \\ &\vdots \\ (x_t, y_t) &\in Q_M \\ (f(x_1), f(y_1)) &\in Q_N \\ &\vdots \\ (f(x_t), f(y_t)) &\in Q_N \end{aligned}$$

are independent. From this it is not hard to see that with probability

$$\prod_{(x,y) \in S_f} (1 - p_M p_N) \leq (1 - p_M p_N)^{\epsilon m^2},$$

none of $(x, y) \in S_f$ satisfies both $(x, y) \in Q_M$ and $(f(x), f(y)) \in Q_N$. Since $Q_M \times Q_N$ refutes f when there is a pair $(x, y) \in S_f$ satisfying both $(x, y) \in Q_M$ and $(f(x), f(y)) \in Q_N$, the probability taken over the random coin tosses of TEST-BILIP that $Q_M \times Q_N$ refutes f is at least

$$1 - (1 - p_M p_N)^{\epsilon m^2}.$$

By the union bound and the fact that there are n^m functions from M to N , with probability at least

$$1 - n^m (1 - p_M p_N)^{\epsilon m^2} \quad (4)$$

over the random coin tosses of TEST-BILIP, every injection from M to N is refuted by $Q_M \times Q_N$.

Now there are two cases to consider. The first is when $2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} \geq 1$. In this case, TEST-BILIP sets $p_M = 1$ and $p_N = 4 \cdot \frac{\ln n}{\epsilon m}$ where $p_N < 1$ is guaranteed by Eq. (3). The second case is when $2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} < 1$. In this case, TEST-BILIP sets $p_M = 2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} < 1$ and $p_N = 2 \cdot \frac{m}{n} \sqrt{\frac{\ln n}{\epsilon m}}$ where $p_N < 1$ is guaranteed by the facts that $p_N = (\frac{m}{n})^2 p_M$ and $m \leq n$. In both cases, we have $p_M p_N = 4 \cdot \frac{\ln n}{\epsilon m}$, resulting in Eq. (4) to be

$$\begin{aligned} &1 - n^m (1 - p_M p_N)^{\epsilon m^2} \\ &= 1 - \exp(m \ln n) \cdot (1 - p_M p_N)^{\frac{1}{p_M p_N} \cdot p_M p_N \epsilon m^2} \\ &\geq 1 - \exp(m \ln n) \exp(-p_M p_N \epsilon m^2) \\ &= 1 - \exp(-3m \ln n). \end{aligned}$$

By the Chernoff bound (Chernoff 1952) and the fact that $m \leq n$, it can be verified that in both the aforementioned cases of setting p_M and p_N , the event $|Q_M| > 1000 p_M m^2$ happens with probability $\exp(-\Omega(n\sqrt{m}))$ over the random coin tosses of TEST-BILIP (in fact, for the case of $p_M = 1$, the probability that $|Q_M| > 1000 p_M m^2$ is zero). Similarly, the event $|Q_N| > 1000 p_N n^2$ happens with probability $\exp(-\Omega(n))$. Finally, if $|Q_M| \leq 1000 p_M m^2, |Q_N| \leq 1000 p_N n^2$ and every injection from M to N is refuted by $Q_M \times Q_N$, then TEST-BILIP clearly rejects. The union bound therefore

shows that TEST-BILIP rejects with probability at least

$$1 - \exp(-3m \ln n) - \exp(-\Omega(n\sqrt{m})) - \exp(-\Omega(n)),$$

which is close to 1 for sufficiently large $n \in \mathbb{N}$. \square

We now turn to analyze the query complexity of TEST-BILIP.

Lemma 4. *On input $L \geq 1, \epsilon > 0$, positive integers $m \leq n$ and given oracle access to finite metric spaces (M, d) and (N, ρ) with $|M| = m$ and $|N| = n$, the query complexity (with respect to ϵ, m and n) of TEST-BILIP is $O(\sqrt{\frac{\ln n}{\epsilon m}}(m^2 + n^2))$.*

Proof. If $\frac{\ln n}{\epsilon m} \geq 1/4$, then TEST-BILIP does exhaustive queries. The query complexity is $m^2 + n^2 = O(\sqrt{\frac{\ln n}{\epsilon m}}(m^2 + n^2))$. Hence, we may assume that

$$\frac{\ln n}{\epsilon m} < 1/4 \quad (5)$$

in the following.

The query complexity of TEST-BILIP is at most $1000 p_M m^2 + 1000 p_N n^2 = O(p_M m^2 + p_N n^2)$. Again, there are two cases to consider. The first is when

$$2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} \geq 1. \quad (6)$$

In this case, TEST-BILIP sets $p_M = 1$ and $p_N = 4 \cdot \frac{\ln n}{\epsilon m}$. The second is when Eq. (6) does not hold. In this case, TEST-BILIP sets $p_M = 2 \cdot \frac{n}{m} \sqrt{\frac{\ln n}{\epsilon m}} < 1$ and $p_N = 2 \cdot \frac{m}{n} \sqrt{\frac{\ln n}{\epsilon m}}$.

In the first case,

$$\begin{aligned} & p_M m^2 + p_N n^2 \\ &= m^2 + \frac{\ln n}{\epsilon m} \cdot 4n^2 \\ &\stackrel{\text{Eq. (6)}}{\leq} \frac{\ln n}{\epsilon m} \cdot 4n^2 + \frac{\ln n}{\epsilon m} \cdot 4n^2 \\ &\stackrel{\text{Eq. (5)}}{\leq} \sqrt{\frac{\ln n}{\epsilon m}} (4n^2 + 4n^2) \\ &= O\left(\sqrt{\frac{\ln n}{\epsilon m}}(m^2 + n^2)\right). \end{aligned}$$

In the second case,

$$p_M m^2 + p_N n^2 = O(mn \sqrt{\frac{\ln n}{\epsilon m}}) = O\left(\sqrt{\frac{\ln n}{\epsilon m}}(m^2 + n^2)\right). \quad \square$$

Combining Lemmas 3–4, we finally arrive at the main result for this section.

Theorem 5. *TEST-BILIP is a one-sided tester for bilipschitz embeddability with query complexity $O(\sqrt{\frac{\ln n}{\epsilon m}}(m^2 + n^2))$ with respect to any $\epsilon > 0$ and any positive integers $m \leq n$.*

When the space (N, ρ) is not too large, or more specifically when $n = \exp(o(\epsilon m))$, Theorem 5 implies that TEST-BILIP has a query complexity of $O(m^2 + n^2)$ with respect to ϵ, m and n . That is, most

distances between pairs need not be queried for one-sided testing of bilipschitz embeddability, provided that the host space is not excessively large.

When (N, ρ) is known in advance, a one-sided tester for bilipschitz embeddability needs only query the other space (M, d) . Equivalently, we could consider one-sided testers for bilipschitz embeddability that may still make queries to both metric spaces, while counting only its query complexity concerning (M, d) . That is, queries to (N, ρ) are regarded as dummy queries. This gives the following easy extension of Theorem 5, whose sketch of proof is given in Appendix II. But this time we use quasi-isometric embeddability for illustration and to be used later in Section 6.

Theorem 6. *There is a one-sided tester for quasi-isometric embeddability which, on input $\kappa \geq 1, C \geq 0, \epsilon > 0$, positive integers $m \leq n$ and given oracle access to metric spaces (M, d) and (N, ρ) with $|M| = m$ and $|N| = n$, makes $O(\frac{m \ln n}{\epsilon})$ queries to (M, d) .*

5 A lower bound on the query complexity

In this section we show a lower bound on the query complexity of any one-sided tester for isometry. This will imply the same lower bound for any one-sided tester of bilipschitz equivalence. For this purpose, we relate the testing of isometry to testing graph isomorphism. The following theorem is due to Fischer and Matsliah (Fischer & Matsliah 2006).

Theorem 7. ((Fischer & Matsliah 2006)) *Let $\epsilon \in (0, \frac{1}{100})$ and n be a positive integer. For every one-sided tester T for graph isomorphism, there are undirected simple n -vertex graphs G_1 and G_2 such that given $\epsilon \in (0, \frac{1}{100})$, n and oracle access to G_1 and G_2 , T makes at least $\frac{n^{3/2}}{200}$ queries.*

Using Theorem 7, it is not hard to give the following $\frac{n^{3/2}}{200}$ lower bound on the query complexity of any one-sided tester for isometry.

Theorem 8. *Let $\epsilon \in (0, \frac{1}{200})$ and n be a positive integer. The query complexity of any one-sided tester for isometry is at least $\frac{n^{3/2}}{200}$ with respect to ϵ and n .*

The interested reader is referred to Appendix III for the proof of Theorem 8

6 Embeddability into possibly infinite spaces

So far we have been dealing with the embeddability of a finite metric space into another finite one. In this section, we are interested in testing the embeddability of a finite metric space (M, d) into a totally bounded (Croom 2002) metric space (N, ρ) that is known in advance. Examples of totally bounded metric spaces include all compact metric spaces (Croom 2002), which in turn include all closed and bounded sets in the Euclidean space by the Heine-Borel theorem (Rudin 1976).

Definition 2. ((Croom 2002)) *Let (X, d) be any metric space. For $\delta > 0$, a δ -net A_δ of (X, d) is a finite subset of X such that for every point $x \in X$, there is an $y \in A_\delta$ with $d(x, y) < \delta$. If (X, d) has a δ -net for every $\delta > 0$, then (X, d) is totally bounded.*

We are now ready to state our main theorem for this section.

Theorem 9. *Let (N, ρ) be a totally bounded metric space. Assume there is an algorithm that outputs a δ -net A_δ of (N, ρ) on input any $\delta > 0$. Then there is an algorithm T that, on input $\kappa \geq 1, 0 \leq C' < C, \epsilon > 0$, a positive integer m and given oracle access to a metric space (M, d) with $|M| = m$, satisfies the following conditions.*

1. *If (M, d) is (κ, C') quasi-isometrically embeddable into (N, ρ) , then T accepts.*
2. *If (M, d) is ϵ -far from being (κ, C) quasi-isometrically embeddable into (N, ρ) , then T rejects with high probability.*
3. *T makes $O(\frac{m \ln |A_{(C-C')/2}|}{\epsilon})$ queries to (M, d) .*

Proof. Denote by QUASI-ISO the algorithm implied in Theorem 6. The algorithm T first selects a $(C - C')/2$ -net $A_{(C-C')/2}$. Then T runs QUASI-ISO on input $\kappa, C, \epsilon, m, |A_{(C-C')/2}|$ and supplies QUASI-ISO with oracle access to (M, d) and $(A_{(C-C')/2}, \rho)$. Clearly, T could satisfy each query of QUASI-ISO by turning the same query to the corresponding metric space. Finally, T accepts if and only if QUASI-ISO accepts. The intuition is that T uses QUASI-ISO to test (M, d) for (κ, C) quasi-isometric embeddability into $(A_{(C-C')/2}, \rho)$.

Now we prove item 1. The premise of item 1 translates to the existence of a function $f : M \rightarrow N$ such that

$$1/\kappa \cdot d(x, y) - C' \leq \rho(f(x), f(y)) \leq \kappa \cdot d(x, y) + C' \quad (7)$$

holds for all $(x, y) \in M \times M$. Below we define a function $g : M \rightarrow A_{(C-C')/2}$. For each $x \in M$, let $g(x)$ be the point in $A_{(C-C')/2}$ that is closest to $f(x)$, breaking ties arbitrarily. Clearly, we have $\rho(g(x), f(x)) < (C - C')/2$ for each $x \in M$. Therefore,

$$\begin{aligned} & \rho(f(x), f(y)) \\ & \leq \rho(f(x), g(x)) + \rho(g(x), g(y)) + \rho(g(y), f(y)) \\ & < \rho(g(x), g(y)) + C - C' \end{aligned}$$

for all $x, y \in M$, and in fact $|\rho(f(x), f(y)) - \rho(g(x), g(y))| < C - C'$ for all $x, y \in M$ by a similar argument. This and Eq. (7) give

$$1/\kappa \cdot d(x, y) - C \leq \rho(g(x), g(y)) \leq \kappa \cdot d(x, y) + C$$

for all $x, y \in M$. Therefore, (M, d) is (κ, C) quasi-isometrically embeddable into $(A_{(C-C')/2}, \rho)$ and thus T accepts.

Item 2 is easily justified because its premise trivially implies that (M, d) is ϵ -far from being (κ, C) quasi-isometrically embeddable into $(A_{(C-C')/2}, \rho)$, which results in rejection of T with high probability.

Item 3 is established by directly invoking Theorem 6 and calculating the query complexity. \square

We briefly justify the applicability of Theorem 9. It is meant to deal with the case where (M, d) is to be embedded into an already-known (N, ρ) . In this case, queries to (N, ρ) can be answered without actually making a query. Since (N, ρ) is known beforehand and since we usually want to embed metric spaces into a host metric space with a simple structure, it is not strange to assume that we can find δ -nets for (N, ρ) . For example, if (N, ρ) is a closed ball of radius $R > 0$ in the 3-dimensional Euclidean space, then it is easy to find a δ -net of cardinality $O(R^3/\delta^3)$ for (N, ρ) .

7 Concluding remarks

We have defined bilipschitz embeddability and ϵ -farness from bilipschitz embeddability using injective functions. Such a definition is justifiable for the following reasons. First, Eq. (1) could be satisfied for all $(x, y) \in M \times M$ only if $f : M \rightarrow N$ is injective. Second and more importantly, one usually defines embeddings between metric spaces using injections, and in fact in many (if not most) areas of mathematics, embeddings are defined using injections (see, e.g., (*Embedding n.d.*, *Croom 2002*, *Goodman & O'Rourke 2004*, *Kenyon et al. 2004*)). In contrast, quasi-isometric embeddability is defined via functions that are not necessarily injective (*Ghys & de la Harpe 1991*, *Farb 1997*, *Farb & Mosher 1999*, *2000*), as we did in Section 2. We could also define the notions of quasi-isometric embeddability and ϵ -farness from quasi-isometric embeddability using injections by modifying the corresponding definitions in Section 2 to concern only with injections $f : M \rightarrow N$. That is, we could define (M, d) to be (κ, C) quasi-isometrically embeddable into (N, ρ) under injections if Eq. (2) holds for some injection $f : M \rightarrow N$ and all $(x, y) \in M \times M$. We could also say that (M, d) is ϵ -far from being (κ, C) quasi-isometrically embeddable into (N, ρ) under injections if Eq. (2) fails on at least an ϵ fraction of pairs $(x, y) \in M \times M$ for every injection $f : M \rightarrow N$. Theorems 5–6 and 9 can be easily adapted to give the corresponding tests for quasi-isometric embeddability under injections. The proofs are mostly the same except for a few trivial modifications to Definition 1 and algorithm TEST-BILIP. The query complexities remain the same. A minor point is that we have treated pairs selected from a metric space as ordered ones. They could also be treated as unordered since the distance function of any metric space is symmetric. Again, this does not change our results.

Our definition of ϵ -farness from L -bilipschitz embeddability is directly concerned with the least possible (over all injections $f : M \rightarrow N$) fraction of pairs $(x, y) \in M \times M$ violating Eq. (1), which is naturally interpreted as the quality of the best possible embedding $f : M \rightarrow N$. This seems as intuitively appealing feature of our definition. However, other definitions of ϵ -farness from L -bilipschitz embeddability may also be worth studying. For example, we may adopt one of the following definitions for (M, d) to be ϵ -far from being L -bilipschitz embeddable into (N, ρ) .

1. At least an ϵ fraction of (ordered or unordered) pairs $(x, y) \in M \times M$ need to have their d -distance changed to obtain a metric space that is L -bilipschitz embeddable into (N, ρ) .
2. Among all (ordered or unordered) pairs in $(M \times M) \cup (N \times N)$, at least an ϵ fraction of them need to have their d -distance or ρ -distance changed so that the modified metric space (M, d) is L -bilipschitz embeddable into the modified metric space (N, ρ) .
3. For a reasonable set of edit operations on metric spaces, the least number of edit operations to turn (M, d) into a metric space that is L -bilipschitz embeddable into (N, ρ) is at least $\epsilon|M|^2$ (or $\epsilon|M|$, depending on whichever is more relevant).
4. For a reasonable set of edit operations on metric spaces, the least number of edit operations on (M, d) and (N, ρ) to turn (M, d) into being L -bilipschitz embeddable into (N, ρ) is at least $\epsilon(|M|^2 + |N|^2)$ (or $\epsilon(|M| + |N|)$, depending on whichever is more relevant).

Although in these definitions, fairness from L -bilipschitz embeddability may no longer correspond to the quality of the best possible embedding, tests for L -bilipschitz embeddability under these definitions may still be worth studying and may provide new insights.

Appendix I: Proof of Theorem 1

Proof of Theorem 1. We first show the easy reduction from the graph isomorphism problem to the problem of testing isometry between finite metric spaces. Given two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$, the reduction outputs two metric spaces (V, d) and (V, ρ) described below. For distinct $x, y \in V$, $d(x, y) = 2$ if $(x, y) \in E_1$ and $d(x, y) = 3$ otherwise. Also, set $d(x, x) = 0$ for each $x \in V$. The metric ρ is defined similarly with E_2 in place of E_1 . It is not hard to verify that (V, d) and (V, ρ) are metric spaces and they are isometric if and only if G_1 is isomorphic to G_2 .

Now we turn to the other direction of the reduction. Given two finite metric spaces (M, d) and (M, ρ) , the reduction computes the sets (not multisets) $\{d(x, y) \mid x, y \in M, x \neq y\}$ and $\{\rho(x, y) \mid x, y \in M, x \neq y\}$. Let $\alpha_1 < \dots < \alpha_t$ be an enumeration of the first set in strictly increasing order and $\beta_1 < \dots < \beta_{t'}$ be that of the second. Assume that $t = t'$ and $\alpha_i = \beta_i$ for $1 \leq i \leq t$, for otherwise the reduction just outputs any two non-isomorphic graphs.

The reduction outputs two undirected simple graphs G_1 and G_2 defined below. It begins with G_1 having vertex set M and the empty edge set, and proceeds by adding to G_1 new vertices and new edges. For each pair of distinct $x, y \in M$, denote by $i(x, y, d)$ the unique value of $i \in \{1, \dots, t\}$ satisfying $d(x, y) = \alpha_i$. The reduction adds $3i(x, y, d)$ new vertices $v_{x,y,1}, \dots, v_{x,y,3i(x,y,d)}$ and also adds new edges

$$(x, v_{x,y,1}), \dots, (x, v_{x,y,3i(x,y,d)})$$

and

$$(v_{x,y,1}, y), \dots, (v_{x,y,3i(x,y,d)}, y)$$

to G_1 . After adding new vertices and edges as above for each pair of distinct $x, y \in M$, the graph G_1 is finally formed. The graph G_2 is formed similarly with ρ in place of d .

Clearly, if (M, d) is isometric to (M, ρ) , then G_1 and G_2 are isomorphic.

Now assume that G_1 is isomorphic to G_2 . We are to show that (M, d) is isometric to (M, ρ) . The set of vertices of G_1 is $M \cup S_1$ where

$$S_1 = \{v_{x,y,j} \mid x, y \in M, x \neq y, 1 \leq j \leq 3i(x, y, d)\}$$

is the set of newly added vertices to G_1 . Similarly, the set of vertices of G_2 is denoted $M \cup S_2$ where S_2 is the set of newly added vertices to G_2 . We may assume without loss of generality that $|M| \geq 2$. From the way we add edges to G_1 (respectively, G_2), it is not hard to see that every vertex in S_1 (respectively, S_2) has degree exactly two in G_1 (respectively, G_2), and every vertex in M has degree at least 3 in G_1 (respectively, G_2). An isomorphism f from G_1 to G_2 must therefore map M one-to-one and onto to M , and S_1 one-to-one and onto to S_2 . Now fix distinct $x, y \in M$ arbitrarily. We are to show that $d(x, y) = \rho(f(x), f(y))$, which implies that f itself (when restricted on M) is an isometry from (M, d) to (M, ρ) . That f is an isomorphism implies

$$\begin{aligned} & |\{v \mid (x, v), (v, y) \text{ are edges of } G_1 \text{ and } v \text{ has} \\ & \quad \text{degree exactly 2 in } G_1\}| \\ = & |\{u \mid (f(x), u), (u, f(y)) \text{ are edges of } G_2 \text{ and} \\ & \quad u \text{ has degree exactly 2 in } G_2\}|. \end{aligned}$$

The fact that S_1 (respectively, S_2) consists of exactly those vertices in G_1 (respectively, G_2) with degree two then implies

$$\begin{aligned} & |\{v \in S_1 \mid (x, v), (v, y) \text{ are edges of } G_1\}| \\ = & |\{u \in S_2 \mid (f(x), u), (u, f(y)) \text{ are edges of } G_2\}|, \end{aligned}$$

which in turn implies that $d(x, y) = \rho(f(x), f(y))$. \square

Appendix II: Proof of Theorem 6

Sketch of proof of Theorem 6. We modify TEST-BILIP slightly to prove the theorem. If $\frac{\ln n}{\epsilon m} \leq 1/4$, the modified TEST-BILIP still does exhaustive queries. Otherwise, TEST-BILIP sets $p_M = 4 \cdot \frac{\ln n}{\epsilon m}$ and $p_N = 1$ (we let TEST-BILIP do exhaustive queries to (N, ρ)). These are different from the original assignments of TEST-BILIP to p_M and p_N . Also modify TEST-BILIP so that after querying Q_M and Q_N to (M, d) and (N, ρ) , it rejects if all functions (not necessarily injective) from M to N are refuted by $Q_M \times Q_N$.

Clearly, when $\frac{\ln n}{\epsilon m} \geq 1/4$, the modified TEST-BILIP does exhaustive queries and the query complexity also follows. It is also clear that the modified TEST-BILIP accepts if (M, d) is (κ, C) quasi-isometrically embeddable into (N, ρ) .

Now assume that (M, d) is ϵ -far from being (κ, C) quasi-isometrically embeddable into (N, ρ) and $\frac{\ln n}{\epsilon m} < 1/4$. It is clear that the modified assignment of $p_M = 4 \cdot \frac{\ln n}{\epsilon m}$ does not exceed 1. Now fix an arbitrary function $f : M \rightarrow N$. Similar to in Lemma 3, we define S_f to be the set of pairs $(x, y) \in M \times M$ violating

$$1/\kappa \cdot d(x, y) - C \leq \rho(f(x), f(y)) \leq \kappa \cdot d(x, y) + C.$$

We have $|S_f| \geq \epsilon m^2$. Since we do exhaustive queries to (N, ρ) , this time f can be refuted by $Q_M \times Q_N$ if some pair in S_f is put into Q_M . The probability that $Q_M \times Q_N$ does not refute f is therefore at most

$$(1 - p_M)^{|S_f|} \leq (1 - 4 \cdot \frac{\ln n}{\epsilon m})^{\epsilon m^2}.$$

By the union bound, the probability that every function from M to N is refuted by $Q_M \times Q_N$ is at least $1 - n^m (1 - 4 \cdot \frac{\ln n}{\epsilon m})^{\epsilon m^2} = 1 - o(1)$. The probability that $Q_M > 1000 p_M m^2$ is small, and $Q_N > 1000 p_N n^2$ happens with probability zero. Therefore, with high probability $Q_M \times Q_N$ refutes every function from M to N , and the whole Q_M and Q_N are queried to (M, d) and (N, ρ) , respectively, resulting in rejection of the modified TEST-BILIP.

The number of queries to (M, d) is at most $1000 p_M m^2$, which is easily verified to obey the desired bound. \square

Appendix III: Proof of Theorem 8

Proof of Theorem 8. Let T be a one-sided tester for isometry with query complexity $q(\epsilon, n)$ with respect to ϵ and n . Using T , we develop a one-sided tester T' for graph isomorphism with query complexity at most $q(\epsilon/2, n)$ with respect to ϵ and n . The theorem is then immediate from Theorem 7.

On input ϵ, n and given oracle access to two undirected simple graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ with $|V| = n$, the algorithm T' simulates T on input $n, \epsilon/2$ and provides T with oracle access to two metric spaces (V, d) and (V, ρ) described below. The

metric space (V, d) is defined by $d(x, x) = 0$ for $x \in V$, $d(x, y) = 2$ for $(x, y) \in E_1$ and $d(x, y) = 3$ for distinct $x, y \in V$ with $(x, y) \notin E_1$. The metric space (V, ρ) is defined similarly except that E_1 is replaced by E_2 . Whenever T makes a query $(x, y) \in V \times V$ to the metric space (V, d) (respectively, (V, ρ)), T' asks G_1 (respectively, G_2) whether $(x, y) \in E_1$ (respectively, $(x, y) \in E_2$) and then computes $d(x, y)$ (respectively, $\rho(x, y)$) to satisfy the query of T . The query complexity of T' is clearly at most $q(\epsilon/2, n)$. Finally, T' accepts (respectively, rejects) if and only if T accepts (respectively, rejects).

It is clear that if G_1 and G_2 are isomorphic, then (V, d) and (V, ρ) are isometric. Hence T and thus T' accepts.

Now assume that G_1 and G_2 are ϵ -far from being isomorphic and let $\pi : V \rightarrow V$ be any bijection. There are at least $\epsilon \binom{|V|}{2}$ unordered pairs $(x, y) \in V \times V$ such that either $(x, y) \in E_1$ and $(\pi(x), \pi(y)) \notin E_2$, or $(x, y) \in E_2$ and $(\pi(x), \pi(y)) \notin E_1$, and it is clear that any such pair satisfies $x \neq y$. This implies the existence of at least $2\epsilon \binom{|V|}{2}$ ordered pairs $(x, y) \in V \times V$ with $d(x, y) \neq \rho(\pi(x), \pi(y))$. Since the bijection π is arbitrary, (V, d) and (V, ρ) must be $\frac{2\epsilon \binom{|V|}{2}}{|V|^2} > \epsilon/2$ far from being isometric, resulting in the rejection of T and thus T' with high probability. \square

References

- Apostol, T. M. (1974), *Mathematical Analysis*, Addison Wesley.
- Chávez, E. & Navarro, G. (2006), ‘A metric index for approximate string matching’, *Theoretical Computer Science* **352**, 266–279.
- Chernoff, H. (1952), ‘A measure of the asymptotic efficiency of tests of a hypothesis based on the sum of observations’, *Annals of Mathematical Statistics* **23**, 493–507.
- Croom, F. H. (2002), *Principles of Topology*, 1st edn, Thomson Learning Asia.
- David, G. & Semmes, S. (2000), ‘Regular mappings between dimensions’, *Publicacions Matemàtiques* **44**, 369–417.
- Deza, M. & Laurent, M. (1997), *Geometry of Cuts and Metrics*, Vol. 15 of *Algorithms and Combinatorics*, Springer.
- Dress, A., Huber, K. T. & Moulton, V. (2001), Metric spaces in pure and applied mathematics, in ‘Quadratic Forms and Related Topics’, pp. 121–139.
- Embedding* (n.d.), Wikipedia: The Free Encyclopedia. <http://en.wikipedia.org/wiki/Embedding>.
- Farb, B. (1997), ‘The quasi-isometry classification of lattices in semisimple Lie groups’, *Mathematical Research Letters* **4**, 705–717.
- Farb, B. & Mosher, L. (1999), ‘Quasi-isometric rigidity for the solvable Baumslag-Solitar groups, II’, *Inventiones Mathematicae* **137**(3), 613–649.
- Farb, B. & Mosher, L. (2000), ‘On the asymptotic geometry of abelian-by-cyclic groups’, *Acta Mathematica* **184**(2), 145–202.
- Fischer, E. (2001), ‘The art of uninformed decisions: A primer to property testing’, *Bulletin of the European Association for Theoretical Computer Science* **75**, 97–126.
- Fischer, E. & Matsliah, A. (2006), Testing graph isomorphism, in ‘Proceedings of the 17th annual ACM-SIAM Symposium on Discrete Algorithms’, pp. 299–308.
- Ganyushkin, A. G., Sushchanskii, V. I. & Tsvirkunov, V. V. (1994), ‘Computations in isometry groups of finite metric spaces’, *Cybernetics and Systems Analysis* **30**(3), 331–347.
- Ganyushkin, A. G. & Tsvirkunov, V. V. (1994), ‘On classification of finite metric spaces’, *Mathematical Notes* **56**(4), 1023–1029.
- Ghys, E. & de la Harpe, P. (1991), *Infinite groups as geometric objects (after Gromov)*, Ergodic theory, symbolic dynamics and hyperbolic space, Oxford University Press.
- Goodman, J. E. & O’Rourke, J., eds (2004), *Handbook of discrete and computational geometry*, 2nd edn, CRC Press, Inc.
- Gupta, A. (2000), Embeddings of Finite Metrics, PhD thesis, University of California, Berkeley.
- Indyk, P. (2001), Algorithmic applications of low-distortion geometric embeddings, in ‘Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science’, pp. 10–33.
- Johnson, W. B. & Lindenstrauss, J., eds (2003), *Handbook of the Geometry of Banach Spaces*, North Holland.
- Kenyon, C., Rabani, Y. & Sinclair, A. (2004), Low distortion maps between point sets, in ‘Proceedings of the 36th annual ACM Symposium on Theory of Computing’, pp. 272–280.
- Linial, N. (2002), ‘Finite metric spaces — combinatorics, geometry and algorithms’, <http://www.cs.huji.ac.il/~nati/PAPERS/icm.ps.gz>.
- Mao, R., Xu, W., Singh, N. & Miranker, D. P. (2005), ‘An assessment of a metric space database index to support sequence homology’, *International Journal on Artificial Intelligence Tools* **14**(5), 867–885.
- Matoušek, J. (2002), *Lectures on Discrete Geometry*, Springer-Verlag New York, Inc.
- Miranker, D. P. (2003), ‘Metric-space indexes as a basis for scalable biological databases’, *OMICS: A Journal of Integrative Biology* **7**(1), 57–60.
- Papadimitriou, C. H. (1994), *Computational Complexity*, Addison Wesley.
- Rudin, W. (1976), *Principles of Mathematical Analysis*, 3rd edn, McGraw-Hill.
- West, D. B. (2001), *Introduction to Graph Theory*, 2nd edn, Prentice-Hall.
- Weston, J. D. (2001), ‘Vectors as quaternions: A corner of linear algebra’, *The Mathematical Gazette* **85**(502), 25–35.