

IMPROVED LOWER BOUNDS FOR EMBEDDINGS INTO L_1 *

ROBERT KRAUTHGAMER[†] AND YUVAL RABANI[‡]

Abstract. We improve upon recent lower bounds on the minimum distortion of embedding certain finite metric spaces into L_1 . In particular, we show that for every $n \geq 1$, there is an n -point metric space of negative type that requires a distortion of $\Omega(\log \log n)$ for such an embedding, implying the same lower bound on the integrality gap of a well-known semidefinite programming relaxation for *sparsest cut*. This result builds upon and improves the recent lower bound of $(\log \log n)^{1/6-o(1)}$ due to Khot and Vishnoi [*The unique games conjecture, integrality gap for cut problems and the embeddability of negative type metrics into l_1* , in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2005, pp. 53–62]. We also show that embedding the edit distance metric on $\{0, 1\}^n$ into L_1 requires a distortion of $\Omega(\log n)$. This result improves a very recent $(\log n)^{1/2-o(1)}$ lower bound by Khot and Naor [*Nonembeddability theorems via Fourier analysis*, in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2005, pp. 101–112].

Key words. metric embeddings, semidefinite programming relaxation, integrality gap, approximation algorithms, negative type metrics, graph partitioning, edit distance

AMS subject classifications. 68W25, 90C22, 05C85

DOI. 10.1137/060660126

1. Introduction. In recent years, low distortion embeddings of finite metric spaces into L_1 have become a powerful tool in an algorithm designer’s arsenal. Such embeddings are extremely useful in two very different contexts, which we discuss below.

In combinatorial optimization, cuts in an n -vertex graph correspond to n -point cut semimetrics (see, e.g., [5, 16, 28]).¹ These semimetrics span the cone of n -point semimetrics that are subsets of L_1 . Polynomial-time computable relaxations of NP-hard cut problems are often expressed as optimization over larger sets of semimetrics. Thus, using (“rounding”) a relaxed solution to approximate the optimal solution to the original problem often boils down to embedding the relaxed solution into L_1 .

In data analysis, proximity and classification problems are often easier to perform when data sets are subsets of L_1 . In fact, for many such problems L_1 behaves as well as Euclidean space (see, for example, [19, 24]). Therefore, a common approach to solving such problems in other input families is first to embed the distance function into a well-behaved normed space such as L_1 , and then to use known solutions for the chosen target space.

We consider in this paper the embedding into L_1 of two types of metrics that have attracted much attention recently. First, we consider negative type metrics.

*Received by the editors May 17, 2006; accepted for publication (in revised form) January 5, 2009; published electronically April 10, 2009. A preliminary version of this paper appeared in *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM, New York, SIAM, Philadelphia, 2006, pp. 1010–1017.

<http://www.siam.org/journals/sicomp/38-6/66012.html>

[†]Weizmann Institute of Science, Rehovot 76100, Israel and IBM Almaden Research Center, San Jose, CA 95120 (robert.krauthgamer@weizmann.ac.il). Part of this author’s work was done while visiting the Technion.

[‡]Computer Science Department, Technion—Israel Institute of Technology, Haifa 32000, Israel (rabani@cs.technion.ac.il). Part of this author’s work was done while visiting IBM Almaden Research Center. This author’s research was supported in part by ISF grant 52/03 and BSF grant 02-00282.

¹A semimetric satisfies all the properties of a metric, except that points may be collocated.

A metric is said to be of *negative type* if the metric derived by taking the square roots of the original distances is Euclidean (i.e., embeds isometrically into L_2). We show that n -point metrics of negative type may require $\Omega(\log \log n)$ distortion to embed into L_1 . Second, we consider the edit distance on $\{0, 1\}^n$, which is the minimum number of character insert/delete/substitute operations required to transform one string into the other. We show that this metric requires $\Omega(\log n)$ distortion to embed into L_1 . Both results are proved using simple tools from Fourier analysis of boolean functions. The lower bound for negative type metrics improves upon the $(\log \log n)^{1/6-o(1)}$ bound by Khot and Vishnoi [23]. (They also indicate that the bound improves to $(\log \log n)^{1/4-o(1)}$ using [29].) The lower bound for edit distance improves upon the $(\log n)^{1/2-o(1)}$ lower bound by Khot and Naor [22]. Both previous bounds were proved very recently using Bourgain's deep result [9] on noise insensitive boolean functions. In contrast with the use of ε -noise in [23, 22], we use hypercube edges in both of our lower bound constructions. As it turns out, this leads to stronger bounds.

Note added in proof. Several generalizations and/or further improvements were published after a preliminary version of our paper appeared in SODA 2006. Our bound for negative type metrics was improved (qualitatively, e.g., to average-distortion) in [15]. A very different construction of negative type metrics (based on the Heisenberg group), given in [25, 13], is likely to give a lower bound of $(\log n)^{\Omega(1)}$. For edit distance, a different lower bound based on a communication complexity approach was given in [2] and answers some of the questions posed in section 3.1.

Sparsest cut and negative type metrics. The most striking connection between L_1 embeddings and combinatorial approximation algorithms is exhibited by the sparsest cut problem. In this problem the input is an undirected graph $G = (V, E)$ with positive edge capacities $c : E \rightarrow \mathbb{N}$, a set of pairs of vertices $D = \{s_i, t_i\}_{i=1}^k$ (called terminals or demand pairs), and a positive demand function $h : \{1, 2, \dots, k\} \rightarrow \mathbb{N}$. The goal is to find a cut (S, \bar{S}) in G (where $\emptyset \neq S \subsetneq V$, $S \cap \bar{S} = \emptyset$, and $S \cup \bar{S} = V$) that minimizes the ratio

$$\frac{\sum_{e \in E: |S \cap e|=1} c(e)}{\sum_{i: |S \cap \{s_i, t_i\}|=1} h(i)}.$$

In the uniform-demand case, D consists of all pairs of vertices and $h(i) = 1$ for all i . Our lower bound requires a nonuniform choice of demands.

The groundbreaking work of [26] gave an $O(\log n)$ -approximation for uniform demand sparsest cut. The algorithm is based on a linear programming relaxation that produces a solution in the form of a (general) metric on the vertex set V , and this metric is used to generate a cut by a region growing argument. Follow-up work applied this method to a similar relaxation for the nonuniform case. Eventually, tight $O(\log k)$ bounds (asymptotically matching the integrality gap) were derived in [5, 28] by relating the sparsest cut problem to embedding into L_1 and applying a theorem of Bourgain [8] that shows that every n -point metric space embeds into L_1 with distortion $O(\log n)$.

An obvious direction for improving the approximation guarantee is to use a more restricted set of metrics in the relaxation. A natural candidate for this is the set of n -point negative type metrics, which includes all n -point subsets of L_1 and can be optimized over in polynomial time using semidefinite programming (within arbitrarily small error). Formally, associate to every vertex $x \in V$ a vector v_x , and then the relaxation can be stated as given in Figure 1.

It has been conjectured that this relaxation can be used to get a constant factor

<p>Minimize $\sum_{e=\{x,y\}\in E} c(e)\ v_x - v_y\ _2^2$</p> <p>s.t. $\ v_x - v_z\ _2^2 \leq \ v_x - v_y\ _2^2 + \ v_y - v_z\ _2^2 \quad \forall x, y, z \in V$</p> <p>$\sum_{\{s_i, t_i\} \in D} h(i)\ v_{s_i} - v_{t_i}\ _2^2 = 1$</p>
--

FIG. 1. A semidefinite programming relaxation for sparsest cut.

approximation (see [4]). A recent breakthrough result [4] obtained an $O(\sqrt{\log n})$ -approximation for the uniform demand case. Follow-up work [11, 3] led to a bound of $O(\sqrt{\log k} \log \log k)$ for the nonuniform case by showing an embedding of negative type metrics into L_1 . (In fact, the embedding is into L_2 , which embeds isometrically into L_1 .)

On the other hand, two recent results [12, 23] independently proved that under Khot’s unique games conjecture [21] the sparsest cut problem is NP-hard to approximate within any constant factor. Quantitatively stronger, yet plausible, versions of the conjecture imply inapproximability factors of up to $\Omega(\sqrt{\log \log n})$. The second paper [23] further showed that the integrality ratio of the above semidefinite relaxation can be as large as $(\log \log N)^{1/6-o(1)}$, exhibiting in particular an n -point negative type metric that requires a distortion of $(\log \log N)^{1/6-o(1)}$ to embed into L_1 . We prove the following theorem.

THEOREM 1.1. *The semidefinite relaxation in Figure 1 has integrality ratio $\Omega(\log \log n)$.*

Using standard arguments (see [23]), this theorem implies the following corollary.

COROLLARY 1.2. *For every integer $n \geq 1$ there is an n -point negative type metric space that requires a distortion of $\Omega(\log \log n)$ to embed into L_1 .*

Our proof of Theorem 1.1 is based on the argument in [23]. Our input instance differs in the choice of edges—we use the regular hypercube edges instead of the ε -noise model used in [23]. As a result, we are able to use Friedgut’s theorem for approximating low average sensitivity boolean functions by juntas [17].

Edit distance. The edit distance (a.k.a. Levenshtein distance [27]) between two strings is the minimum number of character insertions, deletions, and substitutions needed to transform one string into the other. Edit distance is a fundamental measure of similarity between strings and appropriately weighted variants of it play a central role in several domains of data analysis. In particular, efficient algorithms for dealing with variations of the edit distance (often referred to as sequence alignment) are among the most investigated computational problems in molecular biology, dating back to 1970 [30].

We restrict our discussion to strings in $\{0, 1\}^n$, focusing on the asymptotic behavior as n goes to infinity. For $x, y \in \{0, 1\}^n$, let $\text{ed}(x, y)$ denote their edit distance. It is a simple observation that $(\{0, 1\}^n, \text{ed})$ forms a metric space. Our concern is with low distortion embeddings of this metric space into L_1 , motivated, as explained above, by applications in data analysis such as near-neighbor searching (see [31]).

Until recently, no nontrivial bounds were known for embedding edit distance into L_1 . In fact, isometric embedding was only recently ruled out by the $3/2$ lower bound of [1]. On the other hand, an L_1 embedding of edit distance with distortion bounded

by $2^{O(\sqrt{\log n \log \log n})}$ was recently achieved in [31], and this indeed yields the best approximate nearest neighbor scheme currently known for edit distance. (An upper bound of $O(n^{2/3})$ can be derived from the results in [6].) Lastly, the lower bound was significantly strengthened to $(\log n)^{1/2-o(1)}$ in [22], the paper that motivated our work on edit distance. We show the following theorem.

THEOREM 1.3. *The distortion of embedding the metric space $(\{0, 1\}^n, \text{ed})$ into L_1 is $\Omega(\log n)$.*

Although inspired by the lower bound of [22], our proof of the theorem requires more basic machinery, namely, a result due to Kahn, Kalai, and Linial [20]. Technically, the result is obtained by analyzing *separately* the different types of edit operations, namely, a random bit flip and a cyclic shift. For edit distance, this analysis leads to a stronger distortion lower bound because it can be applied also to a random bit flip (i.e., the hypercube edges) compared with the ε -noise used in [22].

2. Integrality gap for the sparsest cut relaxation. Our proof actually analyzes the (nonuniform demand) balanced cut problem instead of analyzing the sparsest cut problem. In the B -balanced cut problem, the input is similar to that of the sparsest cut problem with an additional parameter $0 < B < 1$. The goal is to find a minimum-capacity subset of the edges $F \subseteq E$ whose removal disconnects at least a B -fraction of the demand. The above-mentioned semidefinite relaxation for sparsest cut can be adapted to B -balanced cut as given in Figure 2. (As we do not need the full generality of the problem, we restrict our attention to the case of unit edge capacities and unit, though nonuniform, demands.)

$\text{Minimize } \frac{1}{4} \sum_{(i,j) \in E} \ v_i - v_j\ _2^2$
$\text{s.t. } \begin{array}{ll} \ v_i\ _2^2 = 1 & \forall i \in V \\ \ v_i - v_k\ _2^2 \leq \ v_i - v_j\ _2^2 + \ v_j - v_k\ _2^2 & \forall i, j, k \in V \\ \frac{1}{4} \sum_{\{i,j\} \in D} \ v_i - v_j\ _2^2 \geq B \cdot D \end{array}$

FIG. 2. A semidefinite programming relaxation for B -balanced cut.

Throughout this section, we will use N (rather than n) to denote the number of vertices in a graph. A standard argument, which can be found in [23], shows that for all $0 < B' < B \leq 1$, if there is an N -vertex instance in which the size of every B' -balanced cut exceeds by at least a ρ factor the value of the B -balanced cut relaxation in Figure 2, then the integrality ratio of the sparsest cut relaxation in Figure 1 on N -vertex instances is at least $\Omega(\rho/(B - B'))$. Therefore, our goal in the rest of this section is to show instances satisfying the above premise with $B = 1/2$, $B' = 1/4$, and a gap of $\rho = \Omega(\log \log N)$.

We will first define a graph G^* that is based on a folding of the $2n$ -dimensional hypercube, and define for it a set of demand pairs D to make it an instance of $\frac{1}{2}$ -balanced cut. Next, we will demonstrate a solution of the semidefinite relaxation in Figure 2 for this instance. The solution vectors will be identical to those of [23], so the solution's feasibility follows from the proof in that paper, and we will only need to evaluate the value of this solution. (Notice that the set of feasible solutions is

independent of the structure of the input graph, which is why we can use the same solution as in [23], even though we use a different graph on the same set of vertices.) Finally, we will lower bound the cost of a 1/4-balanced cut, which will conclude the proof of Theorem 1.1.

Note. We employ the notation used in an early version of [23]. It is substantially different from the newer version of [23].

2.1. The instance. Fix an integer $k > 0$ and set $n = 2^k$. Let \mathcal{F} be the family of all functions $f : \{-1, +1\}^k \rightarrow \{-1, +1\}$; hence $|\mathcal{F}| = 2^n$. For $S \subseteq [k]$, define

$$\chi_S : \{-1, +1\}^k \rightarrow \{-1, +1\}, \quad \chi_S(z) = \prod_{i \in S} z_i.$$

The Walsh functions χ_S are characters of the group \mathbb{F}_2^k . They form an orthonormal basis for the space of functions mapping $\{-1, +1\}^k$ to the reals.

For $g, h \in \mathcal{F}$, let gh denote the function that is the pointwise multiplication of g and h . For $z \in \{-1, +1\}^k$ and $S \subseteq [k]$, let $z \circ S$ denote the vector obtained from z by flipping the bits in S . For $g \in \mathcal{F}$ and $S \subseteq [k]$, define

$$g \circ S : \{-1, +1\}^k \rightarrow \{-1, +1\}, \quad (g \circ S)(z) = g(z \circ S).$$

Notice that for every $S \subseteq [k]$, both $f \mapsto f\chi_S$ and $g \mapsto g \circ S$ are one-to-one functions.

For $f, g, f', g' \in \mathcal{F}$, write $(f, g) \equiv (f', g')$ if there exists $S \subseteq [k]$ such that $f' = f\chi_S$ and $g' = g \circ S$. It is easy to see that \equiv is an equivalence relation on $\mathcal{F} \times \mathcal{F}$, and the size of every equivalence class is exactly $2^k = n$. Let $[(f, g)]$ denote the equivalence class of (f, g) .

Let $d_{\mathcal{H}}(f, f')$ denote the Hamming distance between $f, f' \in \mathcal{F}$, i.e., the number of $z \in \{-1, +1\}^k$ such that $f(z) \neq f'(z)$. Notice that for all $f \in \mathcal{F}$ and $\emptyset \neq S \subseteq [k]$, we have $d_{\mathcal{H}}(f, f\chi_S) = n/2$.

We are now ready to define the graph $G^* = (V^*, E^*)$. (See the remark below for a more informal description.) The vertex set is

$$V(G^*) = \{[(f, g)] : f, g \in \mathcal{F}\},$$

i.e., the collection of equivalence classes in $\mathcal{F} \times \mathcal{F}$; hence $N = |V(G^*)| = 2^{2n}/n$. The edge set is

$$E(G^*) = \left\{ (u, u') : \exists (f, g) \in u, \exists (f', g') \in u', \text{ such that } d_{\mathcal{H}}(f, f') + d_{\mathcal{H}}(g, g') = 1 \right\},$$

i.e., two vertices $u, u' \in V(G^*)$ are connected by an edge if we can write them as $u = [(f, g)]$ and $u' = [(f', g')]$ such that either (i) $f = f'$ and the functions g, g' differ in exactly one coordinate, or (ii) $g = g'$ and the functions f, f' differ in exactly one coordinate. It is easy to verify that for all $f, f' \in \mathcal{F}$ and $S \subseteq [k]$, we have $d_{\mathcal{H}}(f, f') = d_{\mathcal{H}}(f\chi_S, f'\chi_S)$ and $d_{\mathcal{H}}(f, f') = d_{\mathcal{H}}(f \circ S, f' \circ S)$. Thus, all vertices in G^* have degree $2n$.

Similarly to [23], we define the demands to be uniform among the vertices corresponding to each single $f \in \mathcal{F}$, i.e., $D = \{(u, u') : \exists f, g, g' \in \mathcal{F}, \text{ such that } u = [(f, g)], u' = [(f, g')]\}$. We note that we could have also defined them similarly to [12] to be $D = \{(u, u') : \exists f, g \in \mathcal{F}, \text{ such that } u = [(f, g)], u' = [(f, -g)]\}$.

Remark. The graph G^* is the $2n$ -dimensional hypercube with the standard edges (i.e., connecting vertices at Hamming distance 1), “folded” by merging together every

vertex (f, g) with $(f\chi_S, g \circ S)$ for all $S \subseteq [k]$. The demands are then uniform inside the n -dimensional cubes corresponding to every single $f \in \mathcal{F}$. It is instructive to think of $f \in \mathcal{F}$ and $g \in \mathcal{F}$ as vectors in $\{-1, +1\}^n$. The map $f \mapsto f\chi_S$ then corresponds to flipping $n/2$ bits in the vector f (as determined by S), and there are n different flip patterns as generated by the Hadamard code. The map $g \mapsto g \circ S$ corresponds to permuting the n coordinates in the vector g (as determined by S), and the n possible permutations form a 1-transitive group, meaning that for every $i, j \in [n]$ there is $S \subseteq [k]$ such that $g_j \equiv (g \circ S)_i$.

2.2. The solution to the semidefinite relaxation. For a function $f \in \mathcal{F}$, define a unit-length vector $\psi(f) \in \mathbb{R}^n$ by setting $(\psi(f))_z = \frac{f(z)}{\sqrt{n}}$ for all $z \in \{-1, +1\}^k$ (this is the truth table of f , scaled to be a unit vector). Let $\mathbf{1}$ denote a function (in \mathcal{F}) that is constant 1. For a vector u and an integer $r \geq 0$, let $u^{\otimes r}$ denote the tensor product of u with itself r times. We associate every $f, g \in \mathcal{F}$ with the following vector in $\mathbb{R}^{n^{st}}$:

$$\varphi(f, g) = \left(\frac{1}{\sqrt{n}} \sum_{S \subseteq [k]} g(\mathbf{1} \circ S) \psi(f\chi_S)^{\otimes s} \right)^{\otimes t},$$

where $s = 8, t = 2^{240} + 1$. (Notice that $\mathbf{1} \circ S$ is the characteristic vector of the set S .) It can be easily verified that for all $f, g \in \mathcal{F}$ and $T \subseteq [k]$, we have $\varphi(f, g) = \varphi(f\chi_T, g \circ T)$, because $(g \circ T)(\mathbf{1} \circ S) = g((\mathbf{1} \circ S) \circ T) = g(\mathbf{1} \circ (S \Delta T))$ and $(f\chi_T)\chi_S = f\chi_{S \Delta T}$. Therefore, φ can be viewed as a function from $V(G^*)$ to $\mathbb{R}^{n^{st}}$, and we can take $(\varphi(u))_{u \in V(G^*)}$ as a solution to the semidefinite relaxation in Figure 2. Since these vectors are exactly the vectors used as the semidefinite solution in [23], and also the demands are the same as in [23], it follows from that paper that all the relaxation’s constraints are satisfied. However, our choice of edges is different and thus the value of this semidefinite solution is different.

LEMMA 2.1. *Let $f, g, f', g' \in \mathcal{F}$ and suppose that $d_{\mathcal{H}}(f, f') + d_{\mathcal{H}}(g, g') = 1$. Then*

$$\frac{1}{n} \sum_{S, S' \subseteq [k]} g(\mathbf{1} \circ S) g'(\mathbf{1} \circ S') (\psi(f\chi_S) \cdot \psi(f'\chi_{S'}))^s \geq 1 - O(s/n).$$

Proof. It is easily seen that $\{\psi(f\chi_S)\}_{S \subseteq [k]}$ is an orthonormal basis of \mathbb{R}^n . For every $h \in \mathcal{F}$, we also have $\psi(h) \in \mathbb{R}^n$, and thus $\sum_{S \subseteq [k]} (\psi(f\chi_S) \cdot \psi(h))^2 = \|\psi(h)\|^2 = 1$.

Suppose first that $g = g'$ and $d_{\mathcal{H}}(f, f') = 1$, and fix $S \subseteq [k]$. We then have $d_{\mathcal{H}}(f\chi_S, f'\chi_S) = 1$, and thus $(\psi(f\chi_S) \cdot \psi(f'\chi_S)) = 1 - 2/n$. It follows that

$$\begin{aligned} \sum_{S' \subseteq [k], S' \neq S} (\psi(f\chi_S) \cdot \psi(f'\chi_{S'}))^s &\leq (1 - (1 - 2/n)^2)^{s/2} \\ &\leq (4/n + 4/n^2)^{s/2}. \end{aligned}$$

Thus, the left-hand side (LHS) in the statement of the lemma is lower bounded by $(1 - 2/n)^{s/2} - (4/n + 4/n^2)^{s/2} \geq 1 - O(s/n)$.

Suppose next that $f = f'$ and $d_{\mathcal{H}}(g, g') = 1$. It follows that for all $S', S \subseteq [k]$ with $S' \neq S$ we have $(\psi(f\chi_S) \cdot \psi(f'\chi_{S'})) = 0$. Thus, the LHS in the statement of the lemma is given by $\frac{1}{n} \sum_{S \subseteq [k]} g(\mathbf{1} \circ S) g'(\mathbf{1} \circ S) = 1 - 2/n$. \square

LEMMA 2.2. *The solution $(\varphi(u))_{u \in V(G^*)}$ for the semidefinite relaxation in Figure 2 has value $O(st/n) \cdot |E(G^*)|$.*

Proof. Let $(u, u') \in E(G^*)$. Let $f, g, f', g' \in \mathcal{F}$ be such that $u = [(f, g)]$, $u' = [(f', g')]$, and $d_{\mathcal{H}}(f, f') + d_{\mathcal{H}}(g, g') = 1$. By Lemma 2.1, $\|\varphi(u) - \varphi(u')\|^2 = 2 - 2(\varphi(u) \cdot \varphi(u')) \leq 2 - 2(1 - O(s/n))^{\frac{1}{2}} \leq O(st/n)$, and we conclude that $\sum_{(u, u') \in E(G^*)} \|\varphi(u) - \varphi(u')\|^2 \leq O(st/n) \cdot |E(G^*)|$. \square

2.3. Balanced cuts. We use the following theorem of Friedgut. The *total influence* (a.k.a. *average sensitivity*) of a boolean function $b : \{0, 1\}^n \rightarrow \{0, 1\}$ is $\sum_{i=1}^n \Pr_{x \in \{0, 1\}^n} [f(x) \neq f(x^i)]$, where $x^i \in \{0, 1\}^n$ is derived from x by flipping the i th bit.

THEOREM 2.3 (Friedgut [17]). *Let $b : \{0, 1\}^n \rightarrow \{0, 1\}$ be a boolean function with average sensitivity k . Let $\epsilon > 0$. Then, there exists a boolean function $b' : \{0, 1\}^n \rightarrow \{0, 1\}$ depending only on $2^{O(k/\epsilon)}$ variables, such that b' differs from b on at most $\epsilon 2^n$ inputs.*

The following lemma shows that every $1/4$ -balanced cut in G^* cuts $\Omega(\frac{\log n}{n})|E(G^*)|$ edges.

LEMMA 2.4. *Let $A : V(G^*) \rightarrow \{0, 1\}$ be a function such that $\Pr_{(u, u') \in D} [A(u) \neq A(u')] \geq 1/4$. Then $\Pr_{(u, u') \in E(G^*)} [A(u) \neq A(u')] \geq \Omega(\frac{\log n}{n})$.*

Proof. We first define an “unfolded” version of G^* . Formally, let G be the $2n$ -dimensional hypercube whose vertex set is $V(G) = \mathcal{F} \times \mathcal{F}$ and edge set is $E(G) = \{((f, g), (f', g')) : d_{\mathcal{H}}(f, f') + d_{\mathcal{H}}(g, g') = 1\}$. To see that this is indeed a hypercube, write every function $f \in \mathcal{F}$ as a vector in $\{-1, +1\}^n$. We will need some specialized terminology regarding the $2n$ variables (dimensions) of this hypercube. In particular, we associate each of the last n variables (i.e., variables in the g -part when writing a vertex as $(f, g) \in V(G)$) with a distinct set $S \subseteq [k]$, as follows: Fix a function $\sigma_0 : \{-1, +1\}^k \rightarrow \{-1, +1\}$ that takes the value -1 exactly once, and for every $S \subseteq [k]$, let $\sigma_S : \{-1, +1\}^k \rightarrow \{-1, +1\}$ be the function $\sigma_S = \sigma_0 \circ S$. Notice that $\{\sigma_S\}_{S \subseteq [k]}$ is the family of functions in \mathcal{F} that take the value -1 exactly once, and thus $d_{\mathcal{H}}((f, g), (f, g\sigma_S)) = 1$.

We extend the function $A : V(G^*) \rightarrow \{0, 1\}$ in a straightforward way to a function $B : V(G) \rightarrow \{0, 1\}$ by defining $B(u) = A([u])$. By definition,

$$(1) \quad B(f, g) = B(f\chi_S, g \circ S) \quad \forall f \in \mathcal{F}, g \in \mathcal{F}, S \subseteq [k].$$

Assume towards contradiction that $\Pr_{(u, u') \in E(G^*)} [A(u) \neq A(u')] \leq \frac{c \log n}{n}$, where $c > 0$ is a constant that will be determined shortly. It follows immediately that the total influence of B is

$$2n \cdot \Pr_{(u, u') \in E(G)} [B(u) \neq B(u')] = 2n \cdot \Pr_{(u, u') \in E(G^*)} [A(u) \neq A(u')] \leq 2c \log n.$$

Using Theorem 2.3 we get that B can be ϵ -approximated by a junta function \hat{B} that depends only on variables in \hat{J} for $|\hat{J}| \leq 2^{O(c \log n)/\epsilon}$. Fixing $\epsilon = \frac{1}{20}$, we see that if $c > 0$ is a sufficiently small constant, then $|\hat{J}| \leq n^{1/3}$. We further restrict our attention to J , which contains the variables in \hat{J} that are among the last n variables (i.e., are in the g -part in G). For $T \subseteq [k]$, define $J \circ T$ to be the set of variables in the g -part that contains the variable indexed by S if and only if J contains the variable indexed by $S \Delta T$; formally, $J \circ T = \{S \subseteq [k] : S \Delta T \in J\}$.

We first claim that there exists $S^* \subseteq [k]$ such that $J \cap (J \circ S^*) = \emptyset$. To prove the claim, observe that for every S with $J \cap (J \circ S) \neq \emptyset$ there is in J a distinct pair of variables, namely variables indexed by T_1, T_2 such that $S = T_1 \Delta T_2$. Hence, at most $n^{2/3}$ choices of $S \subseteq [k]$ may yield a nonempty intersection between J and $J \circ S$. The claim follows since the number of choices for S is $2^k = n > n^{2/3}$.

We now wish to upper bound $\Pr_{(u,u') \in D}[A(u) \neq A(u')] = \Pr_{f,g,h \in \mathcal{F}}[B(f,g) \neq B(f,gh)]$. To this end, notice that every $h \in \mathcal{F}$ can be written as $h = h_J h_{\bar{J}}$, where $h_J \in \mathcal{F}$ is negative only on coordinates corresponding to variables in J , and $h_{\bar{J}} \in \mathcal{F}$ is negative only on coordinates corresponding to variables (in the g -part) not in J . Thus,

$$\begin{aligned} \Pr_{f,g,h \in \mathcal{F}}[B(f,g) \neq B(f,gh)] &\leq \Pr_{f,g,h_{\bar{J}}} [B(f,g) \neq B(f,gh_{\bar{J}})] \\ &\quad + \Pr_{f,g,h_{\bar{J}},h_J} [B(f,gh_{\bar{J}}) \neq B(f,gh)] \\ &\leq 2\varepsilon + \Pr_{f,g,h_J} [B(f,g) \neq B(f,gh_J)] \\ &= 2\varepsilon + \Pr_{f,g,h_J} [B(f\chi_{S^*}, g \circ S^*) \neq B(f\chi_{S^*}, (gh_J) \circ S^*)] \\ &\leq 4\varepsilon; \end{aligned}$$

the first inequality follows by a union bound; the second one holds because $h_{\bar{J}}$ is negative only in variables that \hat{B} does not depend on; the subsequent equality holds because of (1); the last inequality follows by observing that $(gh_J) \circ S^* = (g \circ S^*)h_{J \circ S^*}$ differs from $g \circ S^*$ only on coordinates where $h_{J \circ S^*}$ is negative, and that \hat{B} does not depend on variables in $J \circ S^*$.

We arrive at the conclusion that the fraction of demands separated by this cut is $\Pr_{(u,u') \in D}[A(u) \neq A(u')] \leq 4\varepsilon = 1/5$, contradicting our assumption that the cut is $1/4$ -balanced. \square

Combining Lemmas 2.2 and 2.4 we see that every $1/4$ -balanced cut exceeds by at least a $\Omega(\log n) = \Omega(\log \log N)$ factor the value of the $1/2$ -balanced cut semidefinite relaxation in Figure 2. We conclude that the integrality ratio of the semidefinite relaxation in Figure 1 is $\Omega(\log \log N)$, completing the proof of Theorem 1.1.

It can be easily seen using standard arguments (see [23]), that the semidefinite programming solution presented in section 2.2 forms an N -point negative type metric that requires distortion $\Omega(\log \log N)$ to embed into L_1 , which proves Corollary 1.2.

3. Lower bound for embedding edit distance into L_1 . To prove Theorem 1.3, fix an integer n and let $V = \{0, 1\}^n$. We will need a few weight functions over $V \times V$; for simplicity, we define them as probability distributions. First, let

$$E_H = \{(x, y) : x, y \in V, \|x - y\|_1 = 1\},$$

and let τ_H be a probability distribution over $V \times V$ that has a uniform support over E_H and assigns probability 0 for pairs not in E_H . (Notice that if $(x, y) \in E_H$, then $\text{ed}(x, y) = 1$.) Next, define $S : \{0, 1\}^n \rightarrow \{0, 1\}^n$ to be the cyclic left-shift operation, i.e., $S(x_1, \dots, x_n) = (x_2, \dots, x_n, x_1)$, let

$$E_S = \{(x, S(x)) : x \in V\},$$

and let τ_S be a probability distribution over $V \times V$ that has a uniform support over E_S and assigns probability 0 for pairs not in E_S . (Notice that if $(x, y) \in E_S$, then $\text{ed}(x, y) \leq 2$.) Let the distribution τ be the average of τ_H and τ_S , i.e., $\tau(x, y) = (\tau_H(x, y) + \tau_S(x, y))/2$. Let σ be the uniform distribution over V . For $A \subseteq V$, let $\bar{A} = V \setminus A$. For $A, B \subset V$, we denote by $\tau(A, B)$ the symmetric sum $\sum_{x \in A, y \in B} [\tau(x, y) + \tau(y, x)]$, and similarly define $\tau_H(A, B)$ and $\tau_S(A, B)$. It is well known that for two

strings $x, y \in V$ drawn independently at random, with probability $\Omega(1)$ they satisfy $\text{ed}(x, y) \geq \Omega(n)$. (Indeed, for every $x \in \{0, 1\}^n$, the number of strings $y \in \{0, 1\}^n$ within edit distance (say) $n/10$ from x is $(3n)^{n/10} = o(2^n)$; see [7, Lemma 8] and [22, Lemma 4.4].) We thus get

$$\frac{\mathbb{E}_{(x,y) \in \tau}[\text{ed}(x, y)]}{\mathbb{E}_{x \in \sigma, y \in \sigma}[\text{ed}(x, y)]} \leq \frac{2}{\Omega(n)} \leq O(1/n).$$

Using the cut cone representation of L_1 metrics (see, e.g., [5, 16, 28]), if (V, ed) embeds into L_1 with distortion $D \geq 1$, then there must exist $A \subseteq V$ such that

$$\frac{\tau(A, \bar{A})}{\sigma(A)\sigma(\bar{A})} \leq O(D/n).$$

The following key lemma would then complete the proof of Theorem 1.3.

LEMMA 3.1. *For every $A \subseteq V$,*

$$\tau(A, \bar{A}) \geq \Omega\left(\frac{\log n}{n}\right) \sigma(A)\sigma(\bar{A}).$$

Proof. Fix $A \subseteq V$, and assume without loss of generality that $|A| \leq |\bar{A}|$, i.e., $\sigma(A) \leq 1/2$. Define accordingly a boolean function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ by $f(x) = 1_{\{x \in A\}}$. Let I_j be the influence of the j th variable in f , i.e., $I_j = \Pr_{x \in V}[f(x) \neq f(x \oplus e_j)]$, where e_j is the j th unit vector and \oplus represents coordinatewise addition modulo 2. We shall soon require the following bound that is implicit in [20]; see also [18, Lemma 3.4] or [12, Lemma 2.3] for details.

LEMMA 3.2 (Kahn, Kalai, and Linial [20]). *Let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be a boolean function with balance $p = \Pr_{x \in \{0,1\}^n}[f(x) = 1] \leq 1/2$, and let $I_j = \Pr_{x \in \{0,1\}^n}[f(x) \neq f(x \oplus e_j)]$ be the influence of the j th variable. Then for all $\delta > 0$,*

$$\max_{j \in [n]} I_j \leq \delta \quad \Rightarrow \quad \sum_{j \in [n]} I_j \geq \Omega(p) \log(1/\delta).$$

Let $c > 0$ be a constant to be determined later, and assume towards contradiction that $\tau(A, \bar{A}) < \frac{c \log n}{n} \cdot \sigma(A)$. Observe that the total influence of f is

$$\sum_{j \in [n]} I_j = n \cdot \tau_H(A, \bar{A}) \leq 2n \cdot \tau(A, \bar{A}) < 2c \log n \cdot \sigma(A),$$

and its balance is $\sigma(A)$. We thus get from Lemma 3.2 that there exists $\ell \in [n]$ such that $I_\ell \geq 2/n^{1/8}$, if only $c > 0$ is chosen to be a sufficiently small (depending only on the hidden constant in Lemma 3.2) and assuming n is sufficiently large.

We now claim that for every $k \in \{1, \dots, n^{1/4}\}$, we have $I_{\ell+k} \geq 1/n^{1/8}$; here and in what follows we let I_t for $t > n$ be short hand for I_{t-n} . (The constant $1/4$ is somewhat arbitrary.) Indeed, by our assumption above,

$$\begin{aligned} \Pr_{x \in V}[f(x) \neq f(S(x))] &= \tau_S(A, \bar{A}) \\ &\leq 2\tau(A, \bar{A}) \\ &\leq \frac{2c \log n}{n} \cdot \sigma(A) \\ &\leq \frac{c \log n}{n}. \end{aligned}$$

Observe that if x is chosen uniformly at random from V , then $S(x)$ is also uniformly distributed over V . For every $k \in \{1, \dots, n^{1/4}\}$, it thus follows by a union bound that

$$\Pr_{x \in V}[f(x) \neq f(S^k(x))] \leq \sum_{i=0}^{k-1} \Pr_{x \in V}[f(S^i(x)) \neq f(S^{i+1}(x))] \leq \frac{ck \log n}{n}.$$

Observe that $S^k(x) \oplus e_\ell = S^k(x \oplus e_{\ell+k})$, and that if x is uniformly distributed over V , then so do $S^k(x)$ and $x \oplus e_{\ell+k}$. Using the last inequality above, we thus get

$$\begin{aligned} I_\ell &= \Pr_{x \in V}[f(S^k(x)) \neq f(S^k(x) \oplus e_\ell)] \\ &\leq \Pr_{x \in V}[f(S^k(x)) \neq f(x)] + \Pr_{x \in V}[f(x) \neq f(x \oplus e_{\ell+k})] \\ &\quad + \Pr_{x \in V}[f(x \oplus e_{\ell+k}) \neq f(S^k(x \oplus e_{\ell+k}))] \\ &\leq I_{\ell+k} + \frac{2ck \log n}{n}. \end{aligned}$$

The claim now follows from the bound $I_\ell \geq 2/n^{1/8}$ (assuming n is sufficiently large).

Finally, notice that by the above claim, the total influence is at least $\sum_{k=1}^{n^{1/4}} I_{\ell+k} \geq n^{1/8} \gg c \log n$, which contradicts our assumption above and completes the proof of Lemma 3.1. \square

3.1. Discussion. One obvious challenge left open from section 3 is to determine $c_1(\{0, 1\}^n, \text{ed})$, as there is still a large gap between the upper bound of [31] and the lower bound of Theorem 1.3. We note, however, that Lemma 3.1 is tight: There exists $A \subseteq V$ such that $1/4 \leq \sigma(A) \leq 3/4$ and $\tau(A, \bar{A}) \leq O(\frac{\log n}{n})$. Indeed, let $k = \Theta(\log n)$ and consider the collection of strings containing, possibly with a wrap-around, 0^k as a substring; namely,

$$A = \{x \in \{0, 1\}^n : \exists i \text{ such that } x_{i+1} = \dots = x_{i+k} = 0\},$$

where x_t for $t > n$ is short hand for x_{t-n} . Clearly, k can be chosen such that the cut (A, \bar{A}) is $1/4$ -balanced, i.e., $1/4 \leq \sigma(A) \leq 3/4$. For every $x \in A$ the probability that flipping a random bit in x yields a string not in A is at most k/n , hence $\tau_H(A, \bar{A}) = O(k/n)$, and similarly, for every $x \in A$ clearly $S(x) \in A$, and hence $\tau_S(A, \bar{A}) = 0$. It follows that $\tau(A, \bar{A}) \leq O(\frac{\log n}{n})$.

It is also interesting to note that there exists a subset V' containing a $1 - o(1)$ fraction of the strings in $\{0, 1\}^n$, such that $\Omega(\log n) \leq c_1(V', \text{ed}) \leq O(\log^2 n)$. (The conference version erroneously stated an $O(\log n)$ upper bound.) Specifically, set $t = 4 \log n$, and let V' contain every string $x \in \{0, 1\}^n$ that is t -nonrepetitive in the sense that the length t substrings of x are all distinct. It is easily seen that $\sigma(V') \geq 1 - 1/n^2$. The upper bound on $c_1(V', \text{ed})$ follows from [10, Theorem 3.1], where the t -nonrepetitiveness is used to show an embedding with distortion $O(t \log n) = O(\log^2 n)$. The lower bound on $c_1(V', \text{ed})$ is a consequence of section 3, as follows. Observe that for all $x \in V$, we have $\tau(\{x\}, V \setminus \{x\}) = 2^{-n+1} = 2\sigma(\{x\})$, and therefore $\tau(V', V \setminus V') \leq \sum_{x \in V \setminus V'} \tau(\{x\}, V \setminus \{x\}) = 2\sigma(V \setminus V') \leq 2/n^2$. Now suppose that (V', ed) embeds into L_1 with distortion $D \geq 1$. Then, similar to the argument in Theorem 1.3, there exists $A \subseteq V'$ such that

$$\frac{\tau(A, V' \setminus A)}{\sigma(A)\sigma(V' \setminus A)} \leq O(D/n).$$

By applying this argument repeatedly, it can be further guaranteed that $1/4 \leq \sigma(A) \leq 3/4$ (see, e.g., [22]). It follows that $\tau(A, \bar{A}) \leq \tau(A, V' \setminus A) + \tau(V', V \setminus V') \leq O(D/n) \cdot \sigma(A)\sigma(V \setminus A)$, and then Lemma 3.1 implies that $D \geq \Omega(\log n)$.

It was pointed out by A. Naor that Lemma 3.1 can be cast as the following Poincaré inequality: For every $f : \{0, 1\}^n \rightarrow L_1$,

$$\sum_{j=1}^n \mathbb{E}_{x \in \sigma} \|f(x) - f(x \oplus e_j)\|_1 + n \cdot \mathbb{E}_{x \in \sigma} \|f(x) - f(S(x))\|_1 \geq \Omega(\log n) \cdot \mathbb{E}_{x \in \sigma, y \in \sigma} \|f(x) - f(y)\|_1.$$

Furthermore, it can be generalized to some other operations on the coordinates (other than the cyclic shift).

Another intriguing question is: What is the least distortion required for embedding $(\{0, 1\}^n, \text{ed})$ into L_2 -squared (equivalently, embedding $(\{0, 1\}^n, \sqrt{\text{ed}})$ into L_2)? While a L_2 -squared embedding is technically weaker than an L_1 -embedding (i.e., it follows from but generally does not imply an L_1 -embedding), it is as useful for many applications like nearest neighbor search. We remark that an exact analogue to Lemma 3.1 (based on the same τ and σ) is not true, because by letting $h(x)$ be the Hamming weight of x we obtain $\mathbb{E}_{(x,y) \in \tau} |h(x) - h(y)|^2 \leq O(1/n) \mathbb{E}_{x \in \sigma, y \in \sigma} |h(x) - h(y)|^2$.

Finally, we note that our results do not give a lower bound for several related problems, such as embedding into L_1 of the Ulam metric (cf. [10]), or of the metric induced by edit distance with moves; cf. [14]. See also the note added in proof.

Acknowledgments. We thank James R. Lee, Elchanan Mossel, Assaf Naor, and the anonymous reviewers for useful comments. The first author also wishes to thank the participants of the Spring 2005 Berkeley theory reading group for stimulating discussions that partly led to this research.

REFERENCES

- [1] A. ANDONI, M. DEZA, A. GUPTA, P. INDYK, AND S. RASKHODNIKOVA, *Lower bounds for embedding edit distance into normed spaces*, in Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, SIAM, Philadelphia, 2003, pp. 523–526.
- [2] A. ANDONI AND R. KRAUTHGAMER, *The computational hardness of estimating edit distance*, in Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2007, pp. 724–734.
- [3] S. ARORA, J. R. LEE, AND A. NAOR, *Euclidean distortion and the sparsest cut*, in Proceedings of the 37th Annual ACM Symposium on Theory of Computing, ACM, New York, 2005, pp. 553–562.
- [4] S. ARORA, S. RAO, AND U. VAZIRANI, *Expander flows, geometric embeddings, and graph partitionings*, in Proceedings of the 36th Annual ACM Symposium on Theory of Computing, ACM, New York, 2004, pp. 222–231.
- [5] Y. AUMANN AND Y. RABANI, *An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm*, SIAM J. Comput., 27 (1998), pp. 291–301.
- [6] Z. BAR-YOSSEF, T. S. JAYRAM, R. KRAUTHGAMER, AND R. KUMAR, *Approximating edit distance efficiently*, in Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2004, pp. 550–559.
- [7] T. BATU, F. ERGÜN, J. KILIAN, A. MAGEN, S. RASKHODNIKOVA, R. RUBINFELD, AND R. SAMI, *A sublinear algorithm for weakly approximating edit distance*, in Proceedings of the 35th Annual ACM Symposium on Theory of Computing, ACM, New York, 2003, pp. 316–324.
- [8] J. BOURGAIN, *On Lipschitz embedding of finite metric spaces in Hilbert space*, Israel J. Math., 52 (1985), pp. 46–52.
- [9] J. BOURGAIN, *On the distributions of the Fourier spectrum of Boolean functions*, Israel J. Math., 131 (2002), pp. 269–276.

- [10] M. CHARIKAR AND R. KRAUTHGAMER, *Embedding the Ulam metric into ℓ_1* , Theory of Computing, 2 (2006), pp. 207–224.
- [11] S. CHAWLA, A. GUPTA, AND H. RÄCKE, *Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut*, in Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, (Vancouver, BC), ACM, New York, SIAM, Philadelphia, 2005, pp. 102–111.
- [12] S. CHAWLA, R. KRAUTHGAMER, R. KUMAR, Y. RABANI, AND D. SIVAKUMAR, *On the hardness of approximating multicut and sparsest-cut*, Comput. Complexity, 15 (2006), pp. 94–114.
- [13] J. CHEEGER AND B. KLEINER, *Differentiating Maps into L^1 and the Geometry of BV Functions*, 2006. Available online at www.arxiv.org/abs/math/0611954.
- [14] G. CORMODE AND S. MUTHUKRISHNAN, *The string edit distance matching problem with moves*, in Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, SIAM, Philadelphia, 2002, pp. 667–676.
- [15] N. R. DEVANUR, S. A. KHOT, R. SAKET, AND N. K. VISHNOI, *Integrality gaps for sparsest cut and minimum linear arrangement problems*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing, ACM, New York, 2006, pp. 537–546.
- [16] M. M. DEZA AND M. LAURENT, *Geometry of Cuts and Metrics*, Springer-Verlag, Berlin, 1997.
- [17] E. FRIEDGUT, *Boolean functions with low average sensitivity depend on few coordinates*, Combinatorica, 18 (1998), pp. 27–35.
- [18] E. FRIEDGUT AND G. KALAI, *Every monotone graph property has a sharp threshold*, Proc. Amer. Math. Soc., 124 (1996), pp. 2993–3002.
- [19] P. INDYK AND R. MOTWANI, *Approximate nearest neighbors: Towards removing the curse of dimensionality*, in Proceedings of the 30th Annual ACM Symposium on Theory of Computing, ACM, New York, 1998, pp. 604–613.
- [20] J. KAHN, G. KALAI, AND N. LINIAL, *The influence of variables on boolean functions*, in Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 1988, pp. 68–80.
- [21] S. KHOT, *On the power of unique 2-prover 1-round games*, in Proceedings of the 34th Annual ACM Symposium on Theory of Computing, ACM, New York, 2002, pp. 767–775.
- [22] S. KHOT AND A. NAOR, *Nonembeddability theorems via Fourier analysis*, in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2005, pp. 101–112.
- [23] S. KHOT AND N. K. VISHNOI, *The unique games conjecture, integrality gap for cut problems and the embeddability of negative type metrics into ℓ_1* , in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2005, pp. 53–62.
- [24] E. KUSHILEVITZ, R. OSTROVSKY, AND Y. RABANI, *Efficient search for approximate nearest neighbor in high dimensional spaces*, SIAM J. Comput., 30 (2000), pp. 457–474.
- [25] J. R. LEE AND A. NAOR, *L_p metrics on the Heisenberg group and the Goemans-Linial conjecture*, in Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2006, pp. 99–108.
- [26] T. LEIGHTON AND S. RAO, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, J. ACM, 46 (1999), pp. 787–832.
- [27] V. I. LEVENSHTAIN, *Binary codes capable of correcting deletions, insertions, and reversals*, Soviet Physics Dokl., 10 (1965), pp. 707–710.
- [28] N. LINIAL, E. LONDON, AND Y. RABINOVICH, *The geometry of graphs and some of its algorithmic applications*, Combinatorica, 15 (1995), pp. 215–245.
- [29] E. MOSSEL, R. O'DONNELL, AND K. OLESZKIEWICZ, *Noise stability of functions with low influences: Invariance and optimality*, in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Piscataway, NJ, 2005, pp. 21–30.
- [30] S. B. NEEDLEMAN AND C. D. WUNSCH, *A general method applicable to the search for similarities in the amino acid sequence of two proteins*, J. Molecular Biology, 48 (1970), pp. 443–453.
- [31] R. OSTROVSKY AND Y. RABANI, *Low distortion embeddings for edit distance*, in Proceedings of the 37th Annual ACM Symposium on Theory of Computing, ACM, New York, 2005, pp. 218–226.