comput. complex. **15** (2006), 20 – 39
1016-3328/06/010020–20
DOI 10.1007/s00037-006-0205-6

computational complexity

ON THE COMPLEXITY OF APPROXIMATING k-SET PACKING

Elad Hazan, Shmuel Safra, and Oded Schwartz

Abstract. Given a k-uniform hypergraph, the Maximum k-Set Packing problem is to find the maximum disjoint set of edges. We prove that this problem cannot be efficiently approximated to within a factor of $\Omega(k/\ln k)$ unless P = NP. This improves the previous hardness of approximation factor of $k/2^{O(\sqrt{\ln k})}$ by Trevisan. This result extends to the problem of k-Dimensional-Matching.

Keywords. Computational complexity, hardness of approximation, set packing.

Subject classification. 68Q17.

1. Introduction

This paper studies the following basic optimization problem: given a family of sets over a certain domain, find the maximum number of disjoint sets. We consider the case where all sets in the given family are of the same size k.

For the case where $k = 2$, we can view the sets as edges in a graph whose vertices are the domain, and hence the problem is exactly the famous maximal matching problem which is solvable in polynomial time (Papadimitriou 1994). For $k \geq 3$, again viewing the sets as hyper-edges in a hyper-graph, the problem of finding the maximum matching in k-uniform hyper-graphs is NP-hard. Hence, unless P = NP, the best hope is to obtain a polynomial time approximation algorithm with provably good approximation guaranty.

The simple greedy algorithm is the following: iteratively pick an arbitrary set and add it to the collection of sets maintained thus far, while removing all sets intersecting it. Continue as long as there remain edges in the graph. Obviously this algorithm returns a family of pairwise disjoint sets. It is easy to prove that this algorithm provides a k-approximation to the optimal solution. A constant improvement in the approximation ratio, to $k/2$ (Hurkens & Schrijver 1989), can be obtained by a simple local search heuristic, and is the best known approximation to date.

In this work we prove that the latter approximation guaranty is almost tight, proving the following:

THEOREM 1.1. *It is NP-hard to approximate k-SP to within $\Omega(k/\ln k)$.*

1.1. Previous results. The general MAXIMUM SET PACKING problem is as follows: given a family $\mathcal{F} = \{S_1, \ldots, S_m\}$ of sets over a certain domain $\mathcal{D} = \{x_1, \ldots, x_n\}$, the objective is to find a maximum packing, i.e. a maximum number of pairwise disjoint sets from the given family. This problem is often phrased in graph theory terminology, as a set system is in fact a hyper-graph where the vertices are the items in the domain and the edges are the given sets. In graph theory jargon, a disjoint set of edges is called a matching, hence the objective is to find a maximum matching.

Packing problems are among the fundamental combinatorial optimization problems. Variants of MAXIMUM SET PACKING, including the MAXIMUM INDEPENDENT SET and MAXIMUM CLIQUE problems, have been widely studied (Arora *et al.* 1998; Arora & Safra 1998; Bar-Yehuda & Moran 1984; Boppana & Halldórsson 1992; Feige *et al.* 1996; Håstad 1999; Wigderson 1983). These general formulations of packing problems are notoriously hard even to approximate: Håstad (1999) showed that MAXIMUM CLIQUE (and therefore MAXIMUM INDEPENDENT SET and MAXIMUM SET PACKING as well) cannot be approximated to within $O(N^{1-\varepsilon})$ unless NP $\subseteq$ ZPP (for every $\varepsilon > 0$). The best approximation algorithm for MAXIMUM INDEPENDENT SET achieves an approximation ratio of $O(N/\log^2 N)$ (Boppana & Halldórsson 1992).

In this paper we consider several natural variants of packing problems. The first, and perhaps most natural, is when the size of the hyper-edges is bounded by k. This problem is called MAXIMUM k-SET PACKING (for short k-SP). If in addition we bound the degree of the vertices by two, this becomes the problem of maximum independent set in graphs of degree at most k.

Another (stronger) natural restriction studied here is when we impose a bound on the colorability of the input graph. This is the problem of MAXIMUM k-DIMENSIONAL MATCHING (for short k-DM). It is a variant of MAXIMUM k-SET PACKING where the vertices of the input hyper-graph are a union of k disjoint sets, $V = V_1 \cup \cdots \cup V_k$, and each hyper-edge contains exactly one vertex from each set, i.e., $E \subseteq V_1 \times \cdots \times V_k$. In other words, the vertices of the hyper-graph can be colored using k colors, so that no hyper-edge contains the same color twice. A graph having this property is called *k-strongly-colorable*. Thus the color-bounded version of MAXIMUM k-SET PACKING is, given a k-uniform k-strongly-colorable hyper-graph, find a matching of maximum size.

These bounded variants of MAXIMUM SET PACKING are known to admit approximation algorithms better than their general versions, the quality of the approximation being a function of the bounds. As mentioned previously, the greedy algorithm guarantees a k-approximation for Maximum k-Set Packing. A simple local-search heuristic achieves an approximation ratio of $k/2$ (Hurkens & Schrijver 1989). This is, to date, the best approximation algorithm for Maximum k-Dimensional Matching as well.

For the special case where $k = 2$, both problems are solvable in polynomial time. Maximum 2-Dimensional Matching is just the problem of finding a maximum matching in a bipartite graph, and can be solved in polynomial time, say by a reduction to network flow problems (Papadimitriou 1994). Maximum 2-Set Packing is the problem of finding a maximum matching in a general graph, and for this problem polynomial time algorithms are also known (Edmonds 1965) (for recent efficient algorithms see Mucha & Sankowski 2004).

However, for all $k \geq 3$, Maximum k-Dimensional Matching is NP-hard (Karp 1972; Papadimitriou 1994). Furthermore, for $k = 3$, the problem is known to be APX-hard (Kann 1991). Alon *et al.* (1995) proved that for a suitably large k, Maximum k-Independent Set (finding an independent set of maximum size in k-regular graphs, for short k-IS) is NP-hard to approximate to within k^c for some $c > 0$. This was later improved to the currently best asymptotical inapproximability factor of $k/2^{O(\sqrt{\ln k})}$ (Trevisan 2001). All hardness factors for Maximum k-Independent Set hold in fact for Maximum $k+1$-Dimensional Matching as well (by a simple reduction).

The best known approximation algorithm for k-IS achieves an approximation ratio of $O(k \log\log k / \log k)$ (Vishwanathan 1996). For k-IS of low k values, the best approximation algorithm achieves an approximation ratio of $(k+3)/5$ for $k \geq 3$ (Berman & Fujito 1995; Berman & Furer 1994). Berman & Karpinski (2003) showed an inapproximability factor of 98/97 for Maximum 3-Dimensional Matching. For more on low-degree inapproximability results see Hazan (2002).

1.2. Our contribution. We improve the inapproximability factor for Maximum k-Set Packing, and prove that it is NP-hard to approximate k-SP to within $\Omega(k/\ln k)$. We extend this result to Maximum k-Dimensional Matching. These results also imply the same bound for $(k+1)$-claw-free graphs (see Halldórsson 1998 for definition of this problem and relation to k-SP). They do not hold, however, for k-IS.

The proof of these lower bounds introduces a combinatorial object called hyper-edge-disperser, and we present a randomized construction of such an object. This object may be of independent interest.

1.3. Outline. Some preliminaries are given in Section 2. Section 2.2 presents the notion of hyper-edge-dispersers. Section 3 contains the proof of the asymptotic hardness of approximation for k-SP. Section 4 extends the proof to hold for k-DM. The existence of a good hyper-disperser is proved in Section 5. The optimality of its parameters is shown in Section 5.1. Section 6 contains a discussion on the implications of our results, the techniques used and some open problems.

2. Preliminaries

In order to prove inapproximability of a maximization problem, one usually defines a corresponding gap problem.

Definition 2.1. Let A be a maximization problem. gap-A-$[a, b]$ is the following decision problem: Given an input instance, decide whether

- there exists a solution of fractional size at least b, or
- every solution of the given instance is of fractional size smaller than a.

If the size of the solution is located between these values, then the output is unconstrained.

Clearly, for any maximization problem, if gap-A-$[a, b]$ is NP-hard, than it is NP-hard to approximate A to within any factor smaller than b/a.

Our main result in this paper is derived by a reduction from the following problem.

Definition 2.2. MAX-3-LIN-q is the following optimization problem:
Input: A set Φ of linear equations modulo an integer q, each depending on three variables.
Problem: Find an assignment that satisfies the maximum number of equations.

The following central theorem stems from a long line of research, using the PCP theorem (Arora *et al.* 1998; Arora & Safra 1998) and the parallel repetition theorem (Raz 1998) as a starting point:

Theorem 2.3 (Håstad 2001). *For every* $q \in \mathbb{N}$ *and* $\varepsilon > 0$, gap-MAX-3-LIN-q-$[1/q + \varepsilon, 1 - \varepsilon]$ *is NP-hard. Furthermore, the result holds for instances of* MAX-3-LIN-q *in which the number of occurrences of each variable is a constant (depending on* ε *and on* q*).*

We denote an instance of MAX-3-LIN-q by $\Phi = \{\varphi_1, \ldots, \varphi_m\}$. Φ is over the set of variables $X = \{x_1, \ldots, x_n\}$. Let $\Phi(x)$ be the (multi) set of all equations in Φ depending on $x \in X$ (i.e. it can be seen as all the occurrences of x). Denote by $\mathrm{Sat}(\Phi, A)$ the set of all equations in Φ satisfied by the assignment A. For an assignment A, we denote by $A_{|x}$ the value $a \in [q]$ that A assigns to x.

2.1. Hyper-graphs. A hyper-graph $H = (V, E)$ consists of a set of vertices V and a collection E of subsets of V called hyper-edges (for short, edges).

As usual, the degree of a vertex is the number of edges it appears in. A hyper-graph H is called *d-regular* if the degree of each of its vertices is exactly d, and *k-uniform* if the size of each of its edges is exactly k.

A *matching* is a subset M of E such that all edges of M are pairwise disjoint.

We use the following non-standard definition of an independent set in hyper-graphs:

DEFINITION 2.4. Let $H = (V, E)$ be a hyper-graph. A subset of vertices $I \subseteq V$ is called an *independent set* if any edge $e \in E$ contains at most one vertex from I.

From it we derive the corresponding (but non-standard) definition of colorability:

DEFINITION 2.5. The hyper-graph $H = (V, E)$ is said to be *k-strongly-colorable* if there is a partition of V into k sets such that each part is an independent set.

Hence, a k-uniform k-strongly-colorable hyper-graph H may be denoted by $H = (V^1, \ldots, V^k, E)$, where $E \subseteq V^1 \times \cdots \times V^k$. An analogous notion to strong colorability applies to the edges of a hyper-graph:

DEFINITION 2.6. A hyper-graph $H = (V, E)$ is said to be *d-strongly-edge-colorable* if there exists a coloring of the edges $f : E \to [d]$ so that every vertex participates in at most one edge of each color.

Using these definitions we can formally define the related packing problem studied here:

DEFINITION 2.7. MAXIMUM k-SET PACKING is the following optimization problem:
Input: A k-uniform hyper-graph $H = (V^1, \ldots, V^k, E)$.
Problem: Find a matching of maximum size in H.

Maximum k-Dimensional Matching is the same problem, where the input graph is k-strongly colorable.

2.2. Hyper-dispersers. The following definition is a generalization of disperser graphs. For definitions and results regarding dispersers see Radhakrishnan & Ta-Shma (2000).

Definition 2.8. A hyper-graph $H = (V, E)$ is a (q, δ)-*hyper-edge-disperser* if there exists a partition of its edges: $E = E_1 \cup \cdots \cup E_q$ with $|E_1| = \cdots = |E_q|$, such that every large matching M of H is (almost) concentrated in one part of the edges. Formally, for every M there exists i so that

$$|M \setminus E_i| \leq \delta |E|.$$

Lemma 2.9. *For every $q > 1$ and $t > c(q)$ (where $c(q)$ is a constant depending only on q) there exists a hyper-graph $H = (V, E)$ such that*

- $V = [t] \times [d]$, *where* $d = \Theta(q \ln q)$.
- H *is a* $(q, 1/q^2)$*-hyper-edge-disperser.*
- H *is* d*-uniform,* d*-strongly-colorable.*
- H *is* q*-regular,* q*-strongly-edge-colorable.*

Henceforth we denote such a graph by $\mathcal{D}[t, q]$. Because of the regularity, uniformity and colorability conditions on this hyper-graph, the number of edges is exactly qt, and they can be partitioned into q disjoint color sets. We therefore name its edges $e[i, j]$ where $j \in [q]$ is the color of the edge by an arbitrary strong edge coloring (a coloring where no two edges of the same color share a vertex) and $i \in [t]$ is an arbitrary indexing of the t edges of each color. Note that the t edges of any single color cover all the vertices of $\mathcal{D}[t, q]$.

A proof of the above lemma appears in Section 5. Note that $\mathcal{D}[t, 2]$ is the dual graph of a standard disperser.

3. Proof of the asymptotic inapproximability factor for k-SP

This section provides a polynomial time reduction from MAX-3-LIN-q to k-SP. Given an instance Φ of MAX-3-LIN-q, that is, a set of equations modulo integer q, we construct an instance of k-SP, namely a k-uniform hyper-graph.

For the hyper-graph we construct, we add hyper-edges corresponding to the equations of Φ and satisfying assignments to them. The main idea of the reduction is to construct the hyper-graph in such a way that a large matching corresponds to a consistent satisfying assignment to Φ. For this purpose, the hyper-graph has common vertices for edges that correspond to assignments that are inconsistent.

In general, the sparsity and uniformity of the constructed graph are strongly related to the quality of the hardness result. In order to obtain a sparse graph with small edge size, while retaining edge-intersection properties, we utilize the hyper-edge-disperser graphs defined in the previous section.

3.1. The construction. Let $\Phi = \{\varphi_1, \dots, \varphi_n\}$ be an instance of MAX-3-LIN-q over the set of variables X, where each variable $x \in X$ occurs a constant number of times $c(x) = O(1)$ (as in Theorem 2.3). Let us now describe how to construct, in polynomial time, an instance of k-SP, the hyper-graph $H_\Phi = (V, E)$.

Let us fix, for every variable $x \in X$, a one-to-one mapping between all indices $i_x \in [c(x)]$ and all occurrences of x in Φ.

Recall Lemma 2.9 that asserts the hyper-edge-disperser $\mathcal{D}[t, q]$ (Definition 2.8) exists. For every variable $x \in X$ consider the graph $\mathcal{D}[c(x), q]$. Each vertex in $V(\mathcal{D}[c(x), q])$ corresponds to an occurrence of x in Φ, and a number in $[d]$, where $d = \Theta(q \log q)$. Each of the edges in $E(\mathcal{D}[c(x), q])$ is in a one-to-one correspondence with an occurrence $i_x \in [c(x)]$ and a value $a \in [q]$ according to the strong edge coloring of $\mathcal{D}[c(x), q]$, so let us denote these edges by $e\langle x, i_x, a\rangle$.

The set of vertices V of H_Φ consists of one copy of the vertices of the disperser graph $\mathcal{D}[c(x), q]$ for every variable $x \in X$. Namely,

$$V \triangleq \{v\langle x, i, j\rangle \mid x \in X,\, i \in [c(x)],\, j \in [d]\}.$$

Henceforth, for any variable $x \in X$, the copy of $\mathcal{D}[c(x), q]$ over the vertices $V_x \triangleq \{v\langle x, i, j\rangle \mid i \in [c(x)],\, j \in [d]\}$ will be denoted $\mathcal{D}_x$.

Let us now define the set of edges E of H_Φ. The edges of H_Φ will be composed of several edges from the hyper-graphs $\mathcal{D}_x$. The set E consists of one edge for every equation $\varphi \in \Phi$ over variables x, y, z and assignment A to x, y, z that satisfies φ. Denote by $A|_x, A|_y, A|_z \in [q]$ the values that A assigns to these variables (notice there are q^2 such satisfying assignments). Denote by i_x, i_y, i_z the indices of the occurrences of x, y, z respectively in φ. The edge corresponding to φ and A is a union of three edges from the copies of hyper-graphs $\mathcal{D}_x, \mathcal{D}_y, \mathcal{D}_z$ of the variables in the equation φ:

$$e\langle \varphi, A\rangle = e\langle x, i_x, A|_x\rangle \cup e\langle y, i_y, A|_y\rangle \cup e\langle z, i_z, A|_z\rangle.$$

Clearly, the cardinality of each edge $e\langle \varphi, A\rangle$ is $3d$, as it is the disjoint union of three edges of cardinality d. Note that each of the three edges

$e\langle x, i_x, A|_x\rangle, e\langle y, i_y, A|_y\rangle, e\langle z, i_z, A|_z\rangle$, which compose $e\langle\varphi, A\rangle$, participates in q edges of the hyper-graph H_Φ.

Altogether, the edges of H_Φ are

$$E = \{e\langle\varphi, A\rangle \mid \varphi \in \Phi,\ A \text{ is a satisfying assignment to } \varphi\}.$$

This concludes the construction of the k-SP instance H_Φ.

Notice that for every constant q, the construction can be carried out in deterministic polynomial time. To this end, each disperser $\mathcal{D}_x$ should be constructed in deterministic polynomial time. As each variable x occurs $c(x) = O(1)$ times (by Theorem 2.3), the size of $\mathcal{D}_x$ is constant as well. According to Lemma 2.9, we know that $\mathcal{D}_x$ exists. Therefore, we can enumerate all possible hyper-graphs of the required size and verify whether they are indeed hyper-edge-dispersers with the required parameters.

3.2. Proof of correctness. We next show that the size of a maximum matching in H_Φ is proportional to the maximum number of equations of Φ that can be simultaneously satisfied. That is, if there exists an assignment that satisfies almost all equations of Φ then there exists a matching that covers almost all vertices of H_Φ. On the other hand, if every assignment satisfies at most a small fraction of the equations of Φ, then every matching of H_Φ is small.

Lemma 3.1 (Completeness). *If there is an assignment to Φ which satisfies $1-\varepsilon$ of its equations, then there is a matching in H_Φ of size $((1-\varepsilon)/q^2)|E|$.*

Proof. Let $A\colon X \to [q]$ be an assignment that satisfies $1-\varepsilon$ of the equations. Consider the matching $M \subseteq E$ consisting of all edges corresponding to A, i.e.

$$M = \{e\langle\varphi, A\rangle \mid \varphi \in \mathrm{Sat}(\Phi, A)\}.$$

As M contains one edge corresponding to each satisfied equation, and for each equation there are q^2 satisfying assignments, we have $|M| = ((1-\varepsilon)/q^2)|E|$. To see that these edges are indeed a matching, consider any two edges of M. If they do not relate to the same variables then they do not contain vertices from a joint hyper-edge-disperser. On the other hand, if they do relate to a joint variable $x \in X$, then they relate to different occurrences $i_{x,1}, i_{x,2} \in [c(x)]$, but the same assignment $a \in [q]$ to it. Hence they contain vertices of the same hyper-edge-disperser $\mathcal{D}_x$, but from two distinct edges of the same color, therefore they do not share a vertex. □

Lemma 3.2 (Soundness). *If every assignment to Φ satisfies at most a $1/q + \varepsilon$ fraction of its equations, then every matching in H_Φ is of size $O(q^{-3}|E|)$.*

The proof idea is as follows: given a matching M, each edge in it corresponds to an assignment to three variables. Given a matching M, we use it to define a global assignment A_{maj} to the variables of Φ: every variable is assigned the value which agrees with the maximal number of hyper-edges of M. We then partition the edges of M into two sets: those that agree with the global assignment (named M_{maj}) and the complement set (named M_{min}). The size of M_{maj} is bounded, as it corresponds to the set equations satisfied by A_{maj} (which is small). We then proceed to bound the size of M_{min} using the expansion property of the hyper-edge-dispersers.

Proof. Denote by E_x the edges of H_Φ corresponding to equations that contain the variable x,

$$E_x = \{e\langle\varphi, A\rangle \mid \varphi \in \Phi(x),\, e\langle\varphi, A\rangle \in E\}.$$

Denote by $E_{x=a}$ the subset of E_x corresponding to an assignment of a to x, that is,

$$E_{x=a} = \{e\langle\varphi, A\rangle \mid e\langle\varphi, A\rangle \in E_x,\, A_{|x} = a\}$$

Let M be a matching of maximum size in H_Φ. According to the matching M we define the majority assignment A_{maj} as follows: for every $x \in X$, the assignment $A_{\mathrm{maj}}(x)$ is the value $a \in [q]$ such that $|E_{x=a} \cap M|$ is maximized. Let M_{maj} be the set of edges in M that agree with A_{maj}, and M_{min} be all the other edges in M:

$$M_{\mathrm{maj}} = M \cap \{e\langle\varphi, A_{\mathrm{maj}}\rangle \mid \varphi \in \Phi\}, \qquad M_{\mathrm{min}} = M \setminus M_{\mathrm{maj}}.$$

As the number of equations satisfied by A_{maj} satisfies $|\mathrm{Sat}(\Phi, A_{\mathrm{maj}})| \le 1/q + \varepsilon$, and for each equation there are q^2 edges corresponding to all satisfying assignments for this equation, we have

$$|M_{\mathrm{maj}}| < \left(\frac{1}{q} + \varepsilon\right)\frac{|E|}{q^2}. \tag{3.3}$$

We next bound the size of M_{min}. The idea is as follows: we decompose each edge in M_{min} into the three constructing edges. At least one of those three edges corresponds to an assignment other than the majority assignment. Hence it suffices to bound the number of "constructing edges" that correspond to the minority assignments. This is accomplished using the disperser property.

Consider a certain variable $x \in X$. Then $\mathcal{D}_x$ is a $(q, 1/q^2)$-hyper-edge-disperser (recall Definition 2.8). That is, in any subset of edges of $\mathcal{D}_x$ which is a matching, all but at most a $1/q^2$ fraction of the edges are of one color (which corresponds to a single assignment to the variable x). Clearly, if two edges of $\mathcal{D}_x$ intersect, then so do any pair of edges of H_Φ containing these two edges.

Therefore,

$$\sum_{a \neq A_{\mathrm{maj}}(x)} |M_{\min} \cap E_{x=a}| \leq \frac{1}{q^2} |E(\mathcal{D}_x)| \tag{3.4}$$

where $|E(\mathcal{D}_x)|$ is the number of edges of $\mathcal{D}_x$.

Consider an edge $e\langle x, i_x, a\rangle$ of $\mathcal{D}_x$, where i_x is the index of x when it appears in the equation $\varphi \in \Phi$. This edge was used in the construction of q edges of H_Φ, namely those that correspond to the satisfying assignments of φ that assign a to x.

Hence, every edge of $\mathcal{D}_x$ is a subset of q hyper-edges in E_x. However, no more than one of these q edges may participate in M (as M is a matching). Plugging this observation into (3.4) we obtain

$$\sum_{a \neq A_{\mathrm{maj}}(x)} |M_{\min} \cap E_{x=a}| \leq \frac{1}{q^3} |E_x|. \tag{3.5}$$

Summing up over the variables of Φ yields

$$|M_{\min}| \leq \sum_{x \in X,\, a \neq A_{\mathrm{maj}}(x)} |M_{\min} \cap E_{x=a}| \leq \frac{1}{q^3} \sum_{x \in X} |E_x| = \frac{3}{q^3} |E| \tag{3.6}$$

where the last equality follows from the fact that each equation contains three variables. Thus, from (3.3) and (3.6),

$$|M| = |M_{\min}| + |M_{\mathrm{maj}}| \leq \left(\frac{4}{q^3} + \varepsilon\right) |E|. \qquad \square$$

By Lemmas 3.1 and 3.2 we showed that Gap-k-SP-$[4/q^3+\varepsilon, 1/q^2-\varepsilon]$ is NP-hard. Since each edge is of size $k = 3d = \Theta(q \log q)$ it is NP-hard to approximate k-SP to within $\Omega(k/\ln k)$.

4. Extending the proof for k-DM

The proof for k-DM is similar to the k-SP proof, yet we take additional care to ensure that the graph H_Φ we construct has the required structure (namely, that H_Φ is not only k-uniform, but also k-strongly-colorable).

The construction for k-DM takes into account the *location* of the variables in the equations they appear in. As there are three variables per equation, there are three possible locations. We use the following notation: $\Phi(x, l)$ is the subset of $\Phi(x)$ where x is the lth variable in the equation ($l \in [3]$). We may assume that every variable appears the same number of times in every location

$(\Phi(x,1) = \Phi(x,2) = \Phi(x,3))$, as we can take three copies of each equation, and shift the location of the variables.

Similar to the k-SP construction, we associate a vertex with each appearance of a variable. For every variable $x \in X$, we now have three copies of a hyper-edge-disperser (instead of just one we had for k-SP): a different disperser for each location in the equations. For every location $l \in [3]$, we have a hyper-disperser $\mathcal{D}[c(x)/3, q]$ which is denoted by $\mathcal{D}_{x,l}$. The vertices of H_Φ are the union of the vertices of all these hyper-dispersers corresponding to all variables in the equation set and all locations.

Since $c(x)$ is exactly the number of appearances of the variable x in the equation set Φ, we can enumerate the vertices of H_Φ according to the variable $x \in X$ and equation $\varphi \in \Phi$ they correspond to (and these two parameters determine the location of the variable in the equation as well):

$$V = \{v\langle x, \varphi, j\rangle \mid x \in X,\ \varphi \in \Phi(x),\ j \in [d]\}.$$

The construction of the edges of H_Φ is almost identical to that of the k-SP instance, the difference being the distinction between the three dispersers for each variable. Notice there is a bijection between an occurrence of a variable in a certain equation and the corresponding vertex in one of the three hypergraphs corresponding to this variable. Therefore, there is no ambiguity in the edge construction process, which is otherwise identical to the one for k-SP.

The notation we use for the edges is identical to the k-SP construction as well: the edges correspond to the satisfying assignments to the equations, and are composed of three disperser edges each, $e\langle \varphi, A\rangle = e\langle x, i_x, A|_x\rangle \cup e\langle y, i_y, A|_y\rangle \cup e\langle z, i_z, A|_z\rangle$ (where these three edges are taken from $\mathcal{D}_{x,1}$, $\mathcal{D}_{y,2}$ and $\mathcal{D}_{z,3}$ respectively). The set of all edges is denoted

$$E = \{e\langle \varphi, A\rangle \mid \varphi \in \Phi,\ A \text{ is a satisfying assignment to } \varphi\}.$$

This concludes the construction for k-DM. We first show that the graph constructed is indeed a k-DM instance:

Proposition 4.1. *H_Φ is $3d$-strongly-colorable.*

Proof. We show how to partition V into $3d$ independent sets of equal size. Let the sets be $P_{l,i}$ for $i \in [d]$ and $l \in [3]$, where

$$P_{l,i} = \{v\langle x, \varphi, i\rangle \mid x \in X,\ \varphi \in \Phi(x, l)\}.$$

$P_{l,i}$ is clearly a partition of the vertices, as each vertex belongs to exactly one part.

We now explain why each part is an independent set. Let $P_{l,i}$ be an arbitrary part, and let $e\langle\varphi, A\rangle \in E$ be an arbitrary edge, where the equation φ depends on the variables x, y, z. By construction, this edge is a disjoint union of three hyper-disperser edges corresponding to the three variables x, y, z,

$$e\langle\varphi, A\rangle = e\langle x, i_x, A|_x\rangle \cup e\langle y, i_y, A|_y\rangle \cup e\langle z, i_z, A|_z\rangle.$$

$P_{l,i} \cap e\langle\varphi, A\rangle$ may contain vertices corresponding only to one of the variables x, y, z, since it contains variables corresponding to a single location (first, second or third). Let that variable be, say, x. Since the hyper-graph $D_{x,1}$ is d-uniform and d-strongly-colorable, the edge $e\langle x, i_x, A_{|x}\rangle$ (and hence $e\langle\varphi, A\rangle$) contains exactly one vertex from each of the d parts. Therefore, the set $P_{l,i} \cap e\langle\varphi, A\rangle$ contains exactly one vertex. Since $|P_{l,i} \cap e\langle\varphi, A\rangle| = 1$ for every edge and every set $P_{l,i}$, the graph H_Φ is $3d$-strongly-colorable. □

We proceed to prove the gap in maximum matching size between the cases in which the equation set Φ is almost satisfiable and very unsatisfiable. For the case in which Φ is almost satisfiable, the completeness lemma (Lemma 3.1) is valid for the current construction as well. We prove the appropriate soundness lemma, which is very similar to Lemma 3.2.

LEMMA 4.2 (Soundness). *If every assignment to Φ satisfies at most a $1/q + \varepsilon$ fraction of its equations, then every matching in H_Φ is of size $O(q^{-3}|E|)$.*

PROOF. Denote by $E_{x,l}$ the edges of H_Φ corresponding to equations φ containing the variable x in location l,

$$E_{x,l} = \{e\langle\varphi, A\rangle \mid \varphi \in \Phi(x, l),\ A \in [q^2]\}.$$

Denote by $E_{x=a,l}$ the subset of $E_{x,l}$ corresponding to an assignment of $a \in [q]$ to x, that is,

$$E_{x=a,l} = \{e\langle\varphi, A\rangle \mid \varphi \in \Phi(x, l),\ A_{|x} = a\}.$$

Let M be a matching of maximum size in H_Φ. According to the matching M we define the majority assignment A_{maj}, taking into account the locations of the variables, as follows: for every $x \in X$, let $\hat{l}(x)$ be the location for which $|E_{x,\hat{l}(x)} \cap M|$ is maximized. The assignment $A_{\text{maj}}(x)$ is the value $a \in [q]$ such that $|E_{x=a,\hat{l}(x)} \cap M|$ is maximized.

As before, let M_{maj} be the set of edges in M that agree with A_{maj}, and M_{min} be all the other edges in M:

$$M_{\text{maj}} = M \cap \{e\langle\varphi, A_{\text{maj}}\rangle | \varphi \in \Phi\}, \qquad M_{\text{min}} = M \setminus M_{\text{maj}}.$$

For the exact same reasons as in the previous soundness proof (Lemma 3.2), we have the estimates analogous to (3.3) and (3.5):

$$|M_{\mathrm{maj}}| < \left(\frac{1}{q} + \varepsilon\right)\frac{|E|}{q^2}, \tag{4.3}$$

$$\forall x \in X \quad \sum_{a \neq A_{\mathrm{maj}}(x)} |M_{\mathrm{min}} \cap E_{x=a,\hat{l}(x)}| \leq \frac{1}{q^3}|E_{x,\hat{l}(x)}|. \tag{4.4}$$

Proceeding along the lines of Lemma 3.2, we obtain

$$\begin{aligned}
|M| &\leq \sum_{x,l} |M \cap E_{x,l}| \\
&\leq \sum_{x,l} |M_{\mathrm{maj}} \cap E_{x,l}| + \sum_{x,l,a\neq A_{\mathrm{maj}}(x)} |M_{\mathrm{min}} \cap E_{x=a,l}| \\
&\leq 3\sum_{x} |M_{\mathrm{maj}} \cap E_{x,\hat{l}(x)}| + 3\sum_{x,a\neq A_{\mathrm{maj}}(x)} |M_{\mathrm{min}} \cap E_{x=a,\hat{l}(x)}| \\
&\qquad\qquad \text{(since } |E_{x,\hat{l}(x)} \cap M| \text{ is maximized by } \hat{l}(x)\text{)} \\
&\leq 3|M_{\mathrm{maj}}| + 3\sum_{x,a\neq A_{\mathrm{maj}}(x)} |M_{\mathrm{min}} \cap E_{x=a,\hat{l}(x)}| \\
&< 3\left(\frac{1}{q} + \varepsilon\right)\frac{|E|}{q^2} + \frac{3}{q^3}\sum_{x} |E_{x,\hat{l}(x)}| \quad \text{(by (4.3) and (4.4))} \\
&\leq \left(\frac{6}{q^3} + 3\varepsilon\right)|E|. \qquad \square
\end{aligned}$$

Finally, by the completeness lemma from the previous section and Lemma 4.2 we conclude that Gap-k-DM-$[6/q^3+3\varepsilon, 1/q^2-\varepsilon]$ is NP-hard, thus it is NP-hard to approximate k-DM to within $\Omega(k/\ln k)$.

5. Hyper-dispersers

In this section, we prove Lemma 2.9. As stated before, hyper-dispersers are generalizations of disperser graphs. In Section 5.1, we prove that the parameters given below are the best (up to a constant) for a hyper-disperser one can hope to achieve.

Lemma 2.9. *For every $q > 1$ and $t > c(q)$ (where $c(q)$ is a constant depending only on q) there exists a hyper-graph $H = (V, E)$ such that*

- ◦ $V = [t] \times [d]$, *where* $d = \Theta(q \ln q)$.

- *H is a $(q, 1/q^2)$-hyper-edge-disperser.*
- *H is d-uniform, d-strongly-colorable.*
- *H is q-regular, q-strongly-edge-colorable.*

We denote this graph by $\mathcal{D}[t, q]$.

Proof. We prove that the probability that a randomly generated graph is not a $\mathcal{D}[t, q]$ graph is strictly smaller than 1, which yields the existence of such graphs. Let

$$V = [t] \times [d]$$

and define $V_i = [t] \times \{i\}$. We next randomly construct the edges of the hypergraph, so that it is d-uniform and q-regular. Let S_t be the set of all permutations over t elements. For every $(i_1, i_2) \in [q] \times [d]$ choose a permutation from S_t uniformly at random:

$$\Pi_{i_1,i_2} \in_R S_t.$$

Define

$$e[i, j] = \{(\Pi_{j,1}(i), 1), (\Pi_{j,2}(i), 2), \ldots, (\Pi_{j,d}(i), d)\} \tag{5.1}$$

and let

$$E = \{e[i, j] \mid (i, j) \in [t] \times [q]\},$$

so $|E| = tq$. Define a partition of the edges as follows: $E_i = \{e[j, i] \mid j \in [t]\}$. Thus $|E_1| = \cdots = |E_q| = t$ and each set E_j of edges covers every vertex exactly once. Therefore, H is q-strongly-edge-colorable. On the other hand, every edge contains exactly one vertex from each set of vertices V_i. Thus H is d-strongly-colorable.

We next show that, with high probability, H has the disperser property, namely, every matching M of H is concentrated on a single part of the edges, except for maybe $(1/q^2)|E| = t/q$ edges of M. Denote by P the probability that H does *not* have the disperser property.

Definition 5.2. Define

$$\mathcal{M}_k = \left\{ M \subseteq E \;\middle|\; |M \cap E_k| \leq \frac{t}{q},\ |M \setminus E_k| = \frac{t}{q},\ \forall i,\ |M \cap E_k| \geq |M \cap E_i| \right\}.$$

Proposition 5.3. *If H is not a $(q, 1/q^2)$-hyper-edge-disperser, then there exists a $k \in [q]$ and a set $M \in \mathcal{M}_k$ that is a matching.*

Proof. Suppose that H is not a $(q, 1/q^2)$-hyper-edge-disperser. Then there exists a matching $M' \subseteq E$ that is not concentrated on one color of edges: $\forall i$, $|M' \setminus E_i| > (1/q^2)|E| = t/q$. Let $k \in [q]$ be such that $|M' \cap E_k|$ is maximal. As any subset of a matching is a matching, we can remove edges from $M' \setminus E_k$ until we are left with exactly t/q edges. Likewise, we can remove edges of $M' \cap E_k$ until this set contains at most t/q edges. Note that the property $\forall i$, $|M' \cap E_k| \geq |M' \cap E_i|$ cannot be violated by the deletion of those edges. Thus the new set obtained is a matching in $\mathcal{M}_k$. □

Having the above proposition, we proceed with the proof considering only sets in $\mathcal{M}_1$. Denote by $\Pr[M]$ the probability (over the random choice of H) that M is a matching. By union bound, symmetry with respect to k, and the above proposition,

$$\begin{aligned} P &\leq \Pr_H[\exists k,\, M \in \mathcal{M}_k,\, M \text{ is a matching}] \\ &\leq q \sum_{M \in \mathcal{M}_1} \Pr[M] \leq q|\mathcal{M}_1| \Pr[\hat{M}] \end{aligned} \tag{5.4}$$

where $\hat{M} \in \mathcal{M}_1$ is the set which maximizes $\Pr[M]$. The size of $\mathcal{M}_1$ is bounded from above by the number of possibilities to choose at most t/q edges from E_1 and another t/q edges from the rest of the edge color sets. Therefore, using the known inequality $\binom{n}{k} \leq (en/k)^k$ and assuming $t \gg q \gg 2$ we obtain

$$|\mathcal{M}_1| \leq \binom{(q-1)t}{t/q}\binom{t+1}{t/q} \leq (eq^2)^{t/q}(eq)^{2t/q} \leq (eq)^{4t/q}. \tag{5.5}$$

We next bound $\Pr[\hat{M}]$. Denote by $\hat{M}_i$ the event that $\hat{M}$ restricted to the vertices of V_i is a matching (that is, the edges of $\hat{M}$ do not share a vertex in V_i). According to the independent choice of permutations in the construction of H (recall (5.1)), the events $\hat{M}_i$ are independent and identically distributed. Hence,

$$\Pr[\hat{M}] = \prod_{i=1}^{d} \Pr[\hat{M}_i] \tag{5.6}$$

and we proceed to bound $\Pr[\hat{M}_1]$. Henceforth we shall only consider vertices of V_1.

Let M_i be the set of edges in $\hat{M} \cap E_i$ restricted to the vertices of V_1. Let A_i be the event that the sets of edges $\{M_j \mid j \leq i\}$ are all disjoint. Then

$$\Pr[\hat{M}_1] = \Pr\left[\bigcap_{i=2}^{q} A_i\right] = \prod_{i=2}^{q} \Pr[A_i \mid A_{i-1}]. \tag{5.7}$$

The probability of the event $A_i \mid A_{i-1}$ is the probability of picking at random $|M_i|$ different vertices from a set of t vertices (the set V_1), and avoiding all vertices from $\bigcup_{l=1}^{i-1} M_l$. Naturally, this probability is smaller than the probability of picking $|M_i|$ vertices from a set of t vertices with repetition (one is allowed to choose the same vertex more than once). The assumption A_{i-1} implies that the sets M_l for all $l < i$ are disjoint, and hence $|\bigcup_{l=1}^{i-1} M_l| = \sum_{l=1}^{i-1} |M_l|$. Therefore,

$$\Pr[A_i \mid A_{i-1}] \le \left(1 - \frac{\sum_{l<i} |M_l|}{t}\right)^{|M_i|} \le e^{-t^{-1}|M_i| \sum_{l<i} |M_l|}$$

where for the last inequality we used $1 - x \le e^{-x}$. Thus by (5.6) and (5.7) we have

$$\Pr[\hat{M}] \le e^{-(d/t)\sum_{i=2}^{q}(|M_i|\sum_{j=1}^{i-1}|M_j|)} = e^{-(d/t)\sum_{i<j}|M_i||M_j|}. \tag{5.8}$$

We need an upper bound on the previous probability, and that is obtained when the term $\sum_{i<j} |M_i||M_j|$ is minimized. In our case, the constraint that $\hat{M} \in \mathcal{M}_1$ implies that $|M_1| \ge \max_{i=2}^{q} |M_i|$ and $\sum_{i=2}^{q} |M_i| = t/q$. Lemma 5.10 below shows that the minimum of this expression under these constraints is at least $t^2/4q^2$. Therefore, from 5.8, we obtain the following bound on the probability:

$$\Pr[\hat{M}] \le e^{-\frac{d}{t}\cdot\frac{t^2}{4q^2}} = e^{-\frac{dt}{4q^2}}. \tag{5.9}$$

Therefore by (5.4), (5.5), (5.9),

$$P \le q(eq)^{4t/q} e^{-dt/4q^2}$$

Any d which guarantees that $q(eq)^{4t/q} e^{-dt/4q^2} \ll 1$ suffices to conclude that $P < 1$, and therefore that there exists H with the disperser properties. Simple calculations show that we require $d > (4q^2 \ln q)/t + 12q(1 + \ln q)$. Since $t > q \ge 2$, any $d \ge 100q \ln q$ suffices. □

It remains to prove the following technical lemma:

Lemma 5.10. *Under the constraints*

$$\forall i \in [m],\ x_i \ge 0, \quad x_1 \ge \max_{i=2}^{q} x_i, \quad \sum_{i=2}^{q} x_i = T,$$

we have

$$\sum_{1 \le i < j \le q} x_i x_j \ge \frac{1}{4} T^2.$$

Proof. If $x_1 \geq T/2$, then we directly obtain

$$\sum_{1\leq i<j\leq q} x_i x_j \geq x_1 \sum_{i=2}^{q} x_i \geq \frac{T}{2}\cdot T \geq \frac{1}{4}T^2.$$

Otherwise, we know that

$$\sum_{1\leq i<j\leq q} x_i x_j \geq \sum_{2\leq i<j\leq q} x_i x_j = \frac{1}{2}\Big(\sum_{i=2}^{q} x_i\Big)^2 - \frac{1}{2}\sum_{i=2}^{q} x_i^2 \geq \frac{1}{2}T^2 - \sum_{i=2}^{q} x_i^2.$$

The function $\sum_{i=2}^{q} x_i^2$ is convex, and hence under the constraints $\sum_{i=2}^{q} x_i = T$ and $\max_{i=2}^{q} x_i \leq T/2$, it is maximized where $x_2 = x_3 = T/2$ and the rest of the variables are zero. We obtain $\sum_{i=2}^{q} x_i^2 \leq \frac{1}{4}T^2$, and finally

$$\sum_{1\leq i<j\leq q} x_i x_j \geq \frac{1}{2}T^2 - \sum_{i=2}^{q} x_i^2 \geq \frac{T^2}{4}. \qquad \square$$

5.1. Optimality of hyper-disperser construction. We now turn to see why the hyper-disperser from Lemma 2.9 has optimal parameters. We base our observation on the lemma below from Radhakrishnan & Ta-Shma (2000):

Definition 5.11. A bipartite graph $G = (V_1, V_2, E)$ is called a δ-*disperser* if for every $U_1 \subseteq V_1$ and $U_2 \subseteq V_2$ with $|U_1|, |U_2| \geq \delta|V_1| = \delta|V_2|$, the subset $U_1 \cup U_2$ is *not* an independent set.

Lemma 5.12 (Radhakrishnan & Ta-Shma 2000). *Every bipartite d-regular $1/k$-disperser must satisfy $d = \Omega(k \ln k)$.*

Using this lemma we prove:

Lemma 5.13. *Every d-uniform, d-strongly-colorable, q-regular, q-strongly-edge-colorable $(q, 1/q^2)$-hyper-edge-disperser must satisfy $d = \Omega(q \ln q)$.*

Proof. Consider a hyper-disperser $H = (V_H, E_1, \ldots, E_q)$ as in the statement. Let us construct a bipartite graph $G = (V_1, V_2, E_G)$ as follows. Let

$$V_1 \triangleq E_1, \qquad V_2 \triangleq E_2,$$
$$E_G = \{(e_i, e_j) \mid e_i \in E_1,\ e_j \in E_2,\ e_i \cap e_j \neq \emptyset\}.$$

The graph G is bipartite since the edge sets E_1 and E_2 of H are non-intersecting (as H is q-strongly-edge-colorable). To conclude that G is also d-regular observe the following: H is d-uniform, therefore, every edge $e \in E_1 \cup E_2$

contains d vertices. Moreover, H is q-regular, q-strongly-edge-colorable, thus every vertex of H is contained once in an edge of E_1 and once in an edge of E_2. Therefore, every edge of E_1 intersects one edge of E_2 for each of its d vertices (and vice versa). Thus, every vertex of G is of degree d.

In addition, G is a disperser: consider any two sets $S_1 \subseteq V_1$ and $S_2 \subseteq V_2$ of size $|S_1| = (1/q)|V_1|$ and $|S_2| = (1/q)|V_2|$. The corresponding sets (of edges) in H are of fractional size $1/q^2$ each, thus, as H is a $(q, 1/q^2)$-hyper-edge-disperser, they contain intersecting edges implying that $S_1 \cup S_2$ is not an independent set in G. Since this is true for any $S_1 \subseteq V_1$ and $S_2 \subseteq V_2$, G is a $1/q$-disperser.

As G is a bipartite d-regular $1/q$-disperser, $d = \Omega(q \ln q)$ by Lemma 5.12. □

6. Discussion

An interesting property of our construction is the *almost perfect completeness.* This property refers to the fact that the matching proved to exist in the completeness lemma 3.1 is almost perfect, that is, it covers $1 - \varepsilon$ of the vertices. Knowing the location of a gap is interesting in its own right and may prove useful (in particular if it is extreme on either the completeness or the soundness parameters, see for example Petrank 1994). In fact, applying our reduction on other PCP variants instead of Max-3-Lin-q (e.g. parallel repetition of 3-SAT) yields perfect completeness for k-SP and for k-DM (but with weaker hardness factors).

The ratio between the asymptotic inapproximability factor presented herein for k-SP and k-DM, and the tightest approximation algorithm known was reduced to $O(\ln k)$. The open question of where in the range, from $k/2$ to $O(k/\ln k)$ is the approximability threshold, is of independent interest, as its implications to the difference between k-DM and k-IS. The current asymptotic inapproximability factor of $\Omega(k/\ln k)$ for k-DM approaches the tightest approximation ratio known for k-IS, namely $O(k \log\log k/\log k)$ by Vishwanathan (1996). Thus, a small improvement in either the approximation ratio or the inapproximability factor will show these problems to be of inherently different complexity.

Acknowledgements

We would like to thank Adi Akavia and Dana Moshkovitz for their helpful comments.

References

N. Alon, U. Fiege, A. Wigderson & D. Zuckerman (1995). Derandomized graph products. *Comput. Complexity* **5**, 60–75.

S. Arora, C. Lund, R. Motwani, M. Sudan & M. Szegedy (1998). Proof verification and intractability of approximation problems. *J. ACM* **45**, 501–555.

S. Arora & S. Safra (1998). Probabilistic checking of proofs: a new characterization of NP. *J. ACM* **45**, 70–122.

R. Bar-Yehuda & S. Moran (1984). On approximation problems related to the independent set and vertex cover problems. *Discrete Appl. Math.* **9**, 1–10.

P. Berman & T. Fujito (1995). On approximation properties of the independent set problem for degree 3 graphs. In *Algorithms and Data Structures* (Kingston, ON), Lecture Notes in Comput. Sci. 955, Springer, 449–460.

P. Berman & M. Furer (1994). Approximating maximum independent set in bounded degree graphs. In *Proc. 5th Annual ACM-SIAM Symposium on Discrete Algorithms* (Arlington, VA), 365–371.

P. Berman & M. Karpinski (2003). Improved approximation lower bounds on small occurrence optimization. *Electronic Colloq. on Comput. Complexity (ECCC)* 10 (008).

R. Boppana & M. M. Halldórsson (1992). Approximating maximum independent sets by excluding subgraphs. *BIT* **32**, 180–196.

J. Edmonds (1965). Paths, trees and flowers. *Canad. J. Math.* **17**, 449–467.

U. Feige, S. Goldwasser, L. Lovász, S. Safra & M. Szegedy (1996). Interactive proofs and the hardness of approximating cliques. *J. ACM* **43**, 268–292.

M. M. Halldórsson (1998). Approximations of independent sets in graphs. In *Approximation Algorithms for Combinatorial Optimization* (Aalborg), Lecture Notes in Comput. Sci. 1444, Springer, 1–13.

J. Håstad (1999). Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Math.* **182**, 105–142.

J. Håstad (2001). Some optimal inapproximability results. *J. ACM* **48**, 798–859.

E. Hazan (2002). On the hardness of approximating k-dimensional matching. Master's thesis, Tel-Aviv Univ.

C. A. J. Hurkens & A. Schrijver (1989). On the size of systems of sets every t of which have an SDR, with an application to the worst-case ratio of heuristics for packing problems. *SIAM J. Discrete Math.* **2**, 68–72.

V. Kann (1991). Maximum bounded 3-dimensional matching is MAXSNP-complete. *Inform. Process. Lett.* **37**, 27–35.

R. M. Karp (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations* (Yorktown Heights, NY), Plenum, 85–103.

M. Mucha & P. Sankowski (2004). Maximum matchings via gaussian elimination. In *Proc. 45nd IEEE Symp. on Foundations of Computer Science*, 248–255.

C. Papadimitriou (1994). *Computational Complexity*. Addison, Wesley.

E. Petrank (1994). The hardness of approximation: gap location. *Comput. Complexity* **4**, 133–157.

J. Radhakrishnan & A. Ta-Shma (2000). Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM J. Discrete Math.* **13**, 2–24.

R. Raz (1998). A parallel repetition theorem. *SIAM J. Comput.* **27**, 763–803.

L. Trevisan (2001). Non-approximability results for optimization problems on bounded degree instances. In *Proc. 33rd ACM Symp. on Theory of Computing*, 453–461.

S. Vishwanathan (1996). Personal communication to M. Halldórsson cited in Halldórsson (1998).

A. Wigderson (1983). Improving the performance guarantee for approximate graph coloring. *J. ACM* **30**, 729–735.

Manuscript received 24 December 2003

Elad Hazan
Computer Science Department
Princeton University
Princeton, NJ 08540, U.S.A.
www.cs.princeton.edu/~ehazan

Shmuel Safra
School of Computer Science
Tel Aviv University
Tel Aviv 69978, Israel
safra@post.tau.ac.il

Oded Schwartz
School of Computer Science
Tel Aviv University
Tel Aviv 69978, Israel
odedsc@post.tau.ac.il