Near-Optimal Generalisations of a Theorem of Macbeath

Nabil Mustafa, Saurabh Ray

Near-Optimal Generalisations of a Theorem of Macbeath

Nabil H. Mustafa[1] **and Saurabh Ray**[2]

1 **Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, Equipe A3SI, ESIEE Paris, France.**
`mustafan@esiee.fr`

2 **Indian Institute of Science Bangalore, C.V.Raman Road, Mathikere, Bangalore 560012, India.**
`saurabh@csa.iisc.ernet.in`

Abstract

The existence of Macbeath regions is a classical theorem in convex geometry ("A Theorem on non-homogeneous lattices", Annals of Math, 1952). We refer the reader to the survey of I. Barany for several applications [3]. Recently there have been some striking applications of Macbeath regions in discrete and computational geometry.

In this paper, we study Macbeath's problem in a more general setting, and not only for the Lebesgue measure as is the case in the classical theorem. We prove near-optimal generalizations for several basic geometric set systems. The problems and techniques used are closely linked to the study of ϵ-nets for geometric set systems.

1 Introduction

The goal of this paper is to study small, uniform-sized decompositions of geometric range spaces which approximate the range space. This can be seen as a discrete analogue and extension of the classical result of Macbeath [12] in convex geometry, as well as having several basic connections to well-studied problems in discrete geometry.

Classical Macbeath Regions. Given a convex body K in $\mathbb{R}^d$ of unit volume, and a parameter $\epsilon > 0$, the theorem of Macbeath states the existence of disjoint convex bodies of K, each of volume $\Theta(\epsilon)$, called *Macbeath regions*, such that any halfspace containing at least ϵ volume of K completely contains one of these convex objects. Formally, the following theorem follows from their work:

▶ **Theorem 1** (Macbeath Regions). *Given a convex body $K \subset \mathbb{R}^d$ of unit volume, and a parameter $0 < \epsilon < 1/(2d)^{2d}$, there exists a set of convex objects $\mathcal{M}$, $|\mathcal{M}| = O((1/\epsilon)^{1-2/(d+1)})$, such that for any halfspace h with $vol(h \cap K) \geq \epsilon$, there exists $K_i \in \mathcal{M}$ such that $K_i \subset h \cap K$ and*

$$vol(K_i) \geq \frac{1}{(6d)^{3d}} vol(h \cap K)$$

The existence of Macbeath regions was first proven in the paper of Macbeath [12], with several later applications to geometric problems. Edwald, Larmen and Rogers [8] used it for cap coverings, which was later further extended by Barany and Larman [4] (also see Barany [3] for a survey of this and several other results). It was used for lower-bounds on range-searching by Bronniman, Chazelle and Pach [9]. And very recently, Macbeath regions were used in an elegant way by Arya, Fonseca and Mount [2] for computing near-optimal Hausdorff approximations to polytopes.

Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

A fundamental and powerful result in computational geometry is the existence of small-sized ϵ-nets: given a set system $(X, \mathcal{R})$, and a parameter ϵ, an ϵ-net is a subset $X' \subset X$ such that $X' \cap R \neq \emptyset$ for all $R \in \mathcal{R}$ where $|R| \geq \epsilon|X|$. The famous theorem of Haussler-Welzl [10] shows that ϵ-nets of size $O(d/\epsilon \log d/\epsilon)$ exist for set systems $(X, \mathcal{R})$, where d is the VC-dimension of the set system $(X, \mathcal{R})$. This bound was later improved in [11] to an optimal bound of $(1 + o(1))(\frac{d}{\epsilon} \log(1/\epsilon))$. By now ϵ-nets are an indispensable tool in combinatorics and algorithms [17, 5, 7, 1, 15, 16, 2, 11, 13, 6].

Note that Macbeath's original theorem immediately implies an ϵ-net kind of a result: for any convex body C in $\mathbb{R}^d$ of volume V, it is possible to pick $O(\frac{1}{\epsilon})$ points in C which stab all halfspaces containing an ϵ-th fraction of the volume of C. However, the statement itself is much stronger than that: instead of just points, it gives us $O(\frac{1}{\epsilon})$ regions of volume $\Theta(\epsilon V)$ so that each halfspace containing an ϵ fraction of the volume of C contains one of the regions. The same kind of result is not true in general in a discrete setting (with counting measure instead of Lebesgue measure) for halfspaces in $\mathbb{R}^d$. However, it is true for halfspaces in $\mathbb{R}^3$. Given n points in $\mathbb{R}^3$, one can find $O(\frac{1}{\epsilon})$ groups containing $\Theta(\epsilon n)$ points each so that any halfspace containing ϵn points contains one of the groups. This is much stronger than just the existence of ϵ-nets of size $O(\frac{1}{\epsilon})$.

This raises the intriguing question: of the large number of results known for ϵ-nets, which can be optimally strengthened like above?

Combinatorial Macbeath Regions. Given the existence of decomposition of a convex set K into roughly equal-volume subsets with respect to halfspaces, the natural question is to prove the existence of a small-sized set of Macbeath regions for the counting measure (instead of the Lebesgue measure).

So the problem is: given a set P of n points in $\mathbb{R}^d$ and a parameter $\epsilon > 0$, one would like to construct sets $\mathcal{P} = \{P_1, \ldots, P_m\}$, $P_i \subset P$, such that each set P_i has size $\Omega(\epsilon n)$, and any halfspace containing at least ϵn points contains a set in $\mathcal{P}$.

It turns out that this is implied by a classical result in discrete geometry, called *shallow cuttings*, which states the following [13, 6]. Given a set of n regions $\mathcal{S}$ in $\mathbb{R}^d$ and two parameters r, l, a $(1/r, l)$-shallow cutting w.r.t. $\mathcal{S}$ is a partition of $\mathbb{R}^d$ into cells (of constant descriptive complexity) such that i) each cell is intersected by the boundary of at most n/r regions of $\mathcal{S}$, and ii) the number of cells containing points of depth smaller than l is at most $O((rl/n + 1)^d \cdot n/l \cdot \phi(n/l))$. A set of regions is said to have union complexity $\phi(\cdot)$ if the combinatorial complexity of the union of any r of the regions is at most $r\phi(r)$.

It can be observed that this statement implies a Macbeath-type statement for halfspaces, and more generally, for the following problem for regions of small union complexity: given a set of regions $\mathcal{S}$ of union complexity $\phi(\cdot)$, the objective is to compute a family $\mathcal{U}$ of subsets of $\mathcal{S}$, each of size $\Omega(\epsilon n)$, such that any point contained in at least ϵn objects of $\mathcal{S}$ hits all elements of some set in $\mathcal{U}$.

To construct the Macbeath sets $\mathcal{U}$ for regions in $\mathcal{S}$, fix $l = 2\epsilon n$, $r = 2/\epsilon$, and construct a $(1/r, l)$-shallow cutting for $\mathcal{S}$. For a cell C in the shallow cutting, let $r(C)$ be the set of regions in $\mathcal{S}$ that completely contain C, i.e., $s_i \in r(C)$ iff $C \subset s_i$. Now, for all cells C that contain a point of depth at most $2\epsilon n$ (called *shallow cells*), add $r(C)$ to $\mathcal{U}$. By the shallow-cutting theorem, the number of cells containing a point of depth ϵn is

$$O((rl/n + 1)^d \cdot n/l \cdot \phi(n/l)) = O(4^d \cdot 1/\epsilon \cdot \phi(2/\epsilon))$$

and so $|\mathcal{U}| = O(1/\epsilon \cdot \phi(2/\epsilon))$. To show that sets in $\mathcal{U}$ form the required Macbeath regions, recall that the cutting partitions $\mathbb{R}^d$ into a set of cells such that each cell intersects the boundary of at most $n/r = \epsilon n/2$ objects in $\mathcal{S}$. For a point p hitting ϵn regions, let C be the shallow cell containing p. By the property of shallow-cuttings, of the ϵn regions containing p, at most $\epsilon n/2$ regions can intersect C.

The remaining at least $\epsilon n/2$ regions must then completely contain C, and so for the $r(C)$ added to $\mathcal{U}$ that contains p, we have $|r(C)| \geq \epsilon n/2$.

The above shows the existence of $O(1/\epsilon\phi(1/\epsilon))$ sets such that any point contained in $\Theta(\epsilon n)$ sets of $\mathcal{S}$ must hit one of the constructed sets. To make it work for all sets of size at least ϵn, we can iteratively construct sets for increasing values of ϵ, i.e., $\epsilon, 2\epsilon, \ldots, 2^i\epsilon$, and take the union, still obtaining $O(1/\epsilon\phi(1/\epsilon))$ sets.

For our problem of halfspaces, simply dualize each point in P to a halfspace, and apply the above construction. For halfspaces, $r\phi(r) = O(r^{\lfloor d/2 \rfloor})$, and so we get the following combinatorial version of Macbeath: given a set P of n points in $\mathbb{R}^d$, there exists a set $\mathcal{P} = \{P_1, \ldots, P_m\}, P_i \subset P$, with $m = \frac{1}{\epsilon^{\lfloor d/2 \rfloor}}$, such that $i)$ all sets in $\mathcal{P}$ have size $\Omega(\epsilon n)$, and $ii)$ any halfspace containing ϵn points of P contains at least one set in $\mathcal{P}$.

The existence of a number of other 'Macbeath-type' statements for several other range spaces is also implied by the above proof. In particular, for regions with linear union complexity, i.e., $\phi(r) = O(1)$, there exist linear-sized Macbeath regions. This points to the possibility of the existence of such structural partitioning properties for a wide range of sets derived from geometric objects. In this paper we initiate a systematic study of the analogues of Macbeath regions for other commonly studied geometric set-systems.

Our results.

Given a set system $(X, \mathcal{R})$ and $\epsilon > 0$, we say that a set system $\mathcal{U}$ over X is an ϵ-Macbeath net (or ϵ-Mnet for short) of $(X, \mathcal{R})$ if $i)$ each set in $\mathcal{U}$ has size $\Omega(\epsilon|X|)$, and $ii)$ for every set $R \in \mathcal{R}$ of size at least $\epsilon|X|$, there exists a set $U \in \mathcal{U}$ such that $U \subseteq R$. The *size* of an ϵ-Mnet $\mathcal{U}$ is $|\mathcal{U}|$. Parameterizing the problem a little further, if each set in $\mathcal{U}$ has size at least $\epsilon n/k$, we call it a $\frac{1}{k}$-heavy ϵ-Mnet .

In the study of ϵ-nets for geometric set systems, there are two types of set-systems that are frequently studied: each of these are defined by a set of points P, and a set of regions $\mathcal{S}$. In the so-called 'primal' set-systems, P is taken as the ground set, and the subsets are induced by the regions in $\mathcal{S}$, where a region R induces the subset $R \cap P$. In the so-called 'dual' set-systems, $\mathcal{S}$ is taken as the ground set, and the subsets are induced by points in P, where a point p induces the subset consisting of the regions in $\mathcal{S}$ containing p. Often in the dual setting P is not mentioned, and is assumed to be the entire Euclidean space.

Earlier we pointed out the existence of ϵ-Mnets for halfspaces of size $O(1/\epsilon^{\lfloor d/2 \rfloor})$. Unfortunately this bound cannot be improved substantially: in Section 3, we show that it is not very far from optimal, that is for any d, there exist a set of n points in $\mathbb{R}^d$ where any $\mathcal{P}$ satisfying conditions $i)$ and $ii)$ has size at least $\Omega(\frac{1}{\epsilon^{\lceil (d-1)/3 \rceil}})$.

The earlier statement in fact proved the existence of Macbeath sets for the dual problem for general regions in terms of their union complexity. Namely, it showed:

► Theorem (ϵ-Mnets for dual set-systems). Let $\mathcal{S}$ be a set of n regions in $\mathbb{R}^d$ with union complexity $\phi(r)$[1]. Then there exists an ϵ-Mnet for the dual set-system defined by $\mathcal{S}$ and $\mathbb{R}^d$ (i.e., subsets of $\mathcal{S}$ hit by a point in the plane for the set system), of size $O(\frac{1}{\epsilon}\phi(\frac{1}{\epsilon}))$.

Interestingly, the dependence of $\phi(\cdot)$ in general cannot be reduced to, for example, $\log \phi(\cdot)$, as is the bound for ϵ-nets. A set-system which provides a counter-example follows from our first main result, which considers ϵ-Mnets for the commonly-studied range-space induced by axis-parallel rectangles

[1] And satisfying certain technical conditions of bounded algebraic complexity. See [7] for a broader discussion on this.

in the plane. For these, we optimally tighten the result produced by shallow-cuttings by improving the upper-bound, and then providing a matching lower-bound. We prove the following in Section 2:

▶ **Theorem 1** (ϵ-Mnet for rectangles in $\mathbb{R}^2$)**.** *Let R be a set of n rectangles, P a set of n points in $\mathbb{R}^2$, $\epsilon > 0$ a parameter, and $k \geq 2$ an integer:*

1. *There exists a $\frac{1}{2^k}$-heavy ϵ-Mnet for the dual set-system defined by R and $\mathbb{R}^2$, of size $O(4^k/\epsilon^{1+1/k})$. Furthermore, this cannot be significantly improved: there exists a set $\mathcal{R}$ of n axis-parallel rectangles such that any $\frac{1}{k}$-heavy ϵ-Mnet for the dual set-system defined by $\mathcal{R}$ and $\mathbb{R}^2$ has size $\Omega((1/\epsilon)^{1+1/(k-1)})$.*
2. *There exists an ϵ-Mnet for the primal set-system defined by P and R, of size $O((1/\epsilon)\log 1/\epsilon)$. Furthermore, this cannot be significantly improved: there exists a set P of n points in $\mathbb{R}^2$ such that any $\frac{1}{k}$-heavy ϵ-$Mnet$ for the primal set-system defined by P and R has size $\Omega(\frac{1}{\epsilon}\log_k \frac{1}{\epsilon})$.*

Our second main result is to consider the primal case, i.e., where the input is a set of points P, and the ranges are defined by geometric objects such as circles, k-sided polygons, and in general, objects of some fixed descriptive complexity. We prove the following in Section 3:

▶ **Theorem 2.** *Let P be a set of n points in $\mathbb{R}^2$. Then one can construct ϵ-Mnets of size[2]:*

- *$O(1/\epsilon)$ for sets induced by disks in the plane,*
- *$O(1/\epsilon)$ for sets induced by rectangles all intersecting a fixed line l,*
- *$\tilde{O}(1/\epsilon^2)$ for sets induced by lines, cones, strips in the plane,*
- *$\tilde{O}(1/\epsilon^3)$ for sets induced by triangles, and in general k-sided polygons in the plane (the constant in the asymptotic notation depends on k).*

We further show in Section 3 that near-linear bounds (like those achieved for halfspaces in 2 and 3 dimensions, or for the dual set-systems of linear union-complexity) are not possible for even simple primal set-systems: there exist a set P of n points in the plane such that any ϵ-Mnet for lines must have size $\Omega(1/\epsilon^2)$. This implies that for strips or cones in the plane, the same bound holds, ruling out near-linear bounds for even the simplest type of geometric objects.

We conclude our study by observing that the above series of results, while their proofs use different techniques, indicate an intriguing relation between the size of ϵ-nets and the size of ϵ-Mnets . In all cases, they obey the following: if for a range-space (dual, or primal), the ϵ-nets have size $O(1/\epsilon f(1/\epsilon))$, then the size of ϵ-Mnets for the same range-space is $O(1/\epsilon c^{f(1/\epsilon)})$, where c is constant. So for all spaces known to have linear-sized ϵ-nets (which is optimal), our proofs prove the existence of linear-sized ϵ-Mnets (which is optimal). For the primal set-systems of axis-parallel rectangles in the plane, the ϵ-nets have size $O(1/\epsilon \log\log 1/\epsilon)$ (shown to be optimal) [1, 15], and our result show ϵ-Mnets of size $O(1/\epsilon \log 1/\epsilon)$ (which we show to be optimal). And for the remaining ranges which have ϵ-nets of size $O(1/\epsilon \log 1/\epsilon)$, we show the existence of ϵ-Mnets of size $O(1/\epsilon^c)$. It would be interesting to see if there is any connection with the (still) open problem of finding the right bound on the size of ϵ-nets for lines in the plane.

2 Proof of Theorem 1

In this section we prove Theorem 1, which completely resolves the case for rectangles. We start by giving the lower-bounds for the primal and the dual problem, and then give the matching upper-bounds

[2] $\tilde{O}(\cdot)$ ignores polylogarithmic factors.

for both.

2.1 Lower Bounds

The following Lemma gives the key insight to studying ϵ-Mnet for rectangles.

▶ **Lemma 3.** *For any integers $r, d \geq 1$, consider the grid $G = [r]^d$ in $\mathbb{R}^d$. Then the set system on G induced by incidences with axis-parallel lines can be realized by point-rectangle incidences in $\mathbb{R}^2$.*

Proof. Let $r \geq 1$ be any integer and let $[r]$ represent the set $\{0, \cdots, r-1\}$. Let $G = [r]^d$ which can be thought of as a finite d-dimensional grid of side length r. For some fixed integers $a_1, \cdots, a_{i-1}, a_{i+1}, \cdots a_d \in [r]$, consider the set of points $S_i(a_1, \cdots, a_{i-1}, a_{i+1}, \cdots a_d) = \{(a_1, \cdots, a_{i-1}, x, a_{i+1}, \cdots a_d) : x \in [r]\}$. We call such a set a *line* in direction i. There are dr^{d-1} lines : r^{d-1} in each of the d directions.

We will show that there exists a mapping $\pi : G \mapsto \mathbb{R}^2$ s.t. for each line l (in any direction i), the smallest (inclusion minimal) axis parallel rectangle containing the image $\pi(l)$ of the points in l does not contain the image of any other points of G. Here is the mapping π that we will use: $\pi((a_1, \cdots, a_d)) = \sum_j a_j \vec{v}_j$, where $\vec{v}_j = (r^j, r^{d+1-j})$. For any point $z \in G$, we will interpret $p = \pi(z)$ both as a vector and as a point, as suitable. When treating it as a vector we will denote it as $\vec{p}$.

For any point $p = (a_1, \cdots, a_d) \in G$, let $\vec{V}_{<i}(p)$ denote the vector $\sum_{j<i} a_j \vec{v_j}$ and $\vec{V}_{>i}(p)$ denote the vector $\sum_{j>i} a_j \vec{v_j}$. Thus we can write $\pi(p)$ as $\vec{V}_{<i}(p) + a_i \vec{v}_i + \vec{V}_{>i}(p)$.

Consider the line $l = S_i(a_1, \cdots, a_{i-1}, a_{i+1}, \cdots a_d)$. The minimal rectangle R containing $\pi(l)$ is defined by the two opposite corners $\pi(u)$ and $\pi(v)$, where $u = (a_1, \cdots, a_{i-1}, 0, a_{i+1}, \cdots a_d)$ and $v = (a_1, \cdots, a_{i-1}, r-1, a_{i+1}, \cdots a_d)$ are the extreme points in l. The width of R is $(r-1)r^i$ and its height is $(r-1)r^{d+1-i}$.

Consider any point $z = (b_1, \cdots, b_d) \in G \backslash l$. Let $p = \pi(z)$ and let q be the point $\sum_{j \neq i} a_j \vec{v}_j + b_i \vec{v}_i \in l$. Now, $\vec{p} - \vec{q} = (\vec{V}_{<i}(p) - \vec{V}_{<i}(q)) + (\vec{V}_{>i}(p) - \vec{V}_{>i}(q))$. Since $\vec{p} \neq \vec{q}$, one of the summands must be non-zero. Without loss of generality assume that the latter summand is non-zero. The other case is symmetric.

Since the vector $\vec{V}_{>i}(p) - \vec{V}_{>i}(q)$ is an integral combinations of the vectors v_j , $j > i$, its x-coordinate has magnitude at least r^{i+1}. On the other hand the x-coordinate of $(\vec{V}_{<i}(p) - \vec{V}_{<i}(q))$ has magnitude at most $\sum_{1 \leq j < i} (r-1)r^j = r^i - r^{i-1}$. Therefore, the horizontal distance between p and q is at least $r^{i+1} - (r^i - r^{i-1})$ which is greater than the width of R. Hence, $p \notin R$. When $(\vec{V}_{<i}(p) - \vec{V}_{<i}(q)) \neq 0$, a similar argument holds for the y-coordinates of p and q showing that their vertical distance is larger than the height of R. ◀

Proof of Theorem 1 part 1) lower-bound. We now show that for any integer constant $d \geq 2$, there exists a set $\mathcal{R}$ of n axis-parallel rectangles such that any $\frac{1}{k}$-heavy ϵ-Mnet for $\mathcal{R}$ w.r.t. points has size $\Omega((1/\epsilon)^{1+1/(k-1)})$.

Proof. Now apply Lemma 3 with $d = k$ and $r = \epsilon^{-\frac{1}{d-1}}$. Let G be the grid $[r]^d$ as before. We set $P = \{\pi(p) : p \in G\}$ and we take $\mathcal{R}$ to be the set of rectangles with $\epsilon n/d$ copies of each of the set $\mathcal{R}'$ of dr^{d-1} rectangles corresponding to the dr^{d-1} lines in G. Note that $|\mathcal{R}| = \epsilon n/d \cdot dr^{d-1} = n$. Since each of the points in G is contained in d lines (one in each direction), the points in P are contained in d rectangles of $\mathcal{R}'$ and consequently ϵn rectangles of $\mathcal{R}$. Since there is at most one line through two points in G there is at most one rectangle in $\mathcal{R}'$, and hence at most $\epsilon n/d$ rectangles of $\mathcal{R}$ that contain any pair of points $p, q \in P$. Since for any $\frac{1}{k}$-heavy ϵ-Mnet $\mathcal{U}$, each $U \in \mathcal{U}$ has size more than $\epsilon n/k$, it must be that no set in $\mathcal{U}$ can be contained in two sets $\mathcal{R}(p)$ and $\mathcal{R}(q)$ induced by two distinct points p and q in P. Therefore $|\mathcal{U}| \geq |P| = r^d = \epsilon^{-\frac{k}{k-1}}$. ◀

Proof of Theorem 1 part 2) lower-bound. We now show that for any integer constant $k \geq 2$, there exists a set P of n points in $\mathbb{R}^2$ such that any $\frac{1}{k}$-heavy ϵ-$Mnet$ of P, w.r.t. axis-parallel rectangles, has size $\Omega(\frac{1}{\epsilon}\log_k \frac{1}{\epsilon})$.

Proof. Apply Lemma 3 with $r = k$, and d such that $r^{d-1} = \frac{1}{\epsilon}$. Let $\mathcal{R}$ be the set of $dr^{d-1} = 1/\epsilon \log_k 1/\epsilon$ rectangles corresponding to the lines of G, and let P be the set of points with $\epsilon n/r$ copies of each $\pi(p), \forall p \in G$. Each of the rectangles in $\mathcal{R}$ contains $r\epsilon n/r = \epsilon n$ points of P. Any two rectangles of $\mathcal{R}$ share at most $\epsilon n/r = \epsilon n/k$ points of P. Thus no two rectangles in $\mathcal{R}$ may share the same set $U \in \mathcal{U}$ of a $\frac{1}{k}$-heavy ϵ-$Mnet$ $\mathcal{U}$. Since each of them must contain some $U \in \mathcal{U}$, we have $|\mathcal{U}| \geq |\mathcal{R}|$ and the result follows. ◀

2.2 Upper Bounds

We now give constructions which match the preceding lower-bounds to complete the proof of Theorem 1 part 1. Our argument is based essentially on the technique of boot-strapping; at the cost of worse constant factors, we give a simple exposition below.

Construct a hierarchical subdivision of the rectangles in R by vertical lines, with an integer $k = 1/\epsilon^{1/d}$, as follows. Let $n_i = n/k^i$, and $\epsilon_i = \epsilon(k/2)^i$. At the 0-th level ($i = 0$), let $l_1^0, \ldots, l_k^0$ by a set of k vertical lines such that the number of rectangles of R lying between two consecutive lines ('a slab') is at most n/k. Let R_j^0 be the set of rectangles lying entire in the j-th slab. For each line l_j^0, construct a $\epsilon_i/4$-Mnet for all of the (at most) n rectangles of R intersecting it. Furthermore, recursively construct a ϵ_{i+1}-Mnet for the rectangles in R_j^0 for each j. The recursive construction continues for d steps, where at the i-level, there are k^i total subproblems, each subproblem has at most $n_i = n/k^i$ rectangles, and with $\epsilon_i = \epsilon(k/2)^i$. Finally we use a direct $O(1/\epsilon_d^2)$-sized construction for the ϵ_d-Mnet of the final k^d subproblems at level $i = d$: construct $10/\epsilon_d$ vertical and horizontal lines so that each vertical and horizontal slab contains at most $\epsilon_d n/10$ rectangles, and for each grid cell c, add to $\mathcal{U}$ any $\epsilon_d n_d/2$ rectangles containing c (if possible). Now notice that any point in $\epsilon_d n_d$ rectangles must have at most $\epsilon_d n_d/5$ rectangles intersecting the cell boundary in which it lies, and so at least $\epsilon_d n_d/2$ of the remaining ones would form a set in $\mathcal{U}$. The next two claims show that all these Mnet together form a ϵ-Mnet $\mathcal{U}$ for R of the required size, and we're done.

▶ Claim 1. Each set in $\mathcal{U}$ has size $\Theta(\epsilon n/2^d)$. The size of $\mathcal{U}$ is $O(4^d/\epsilon^{1+1/d})$.

Proof. At the i-level there are k^i subproblems, each of size at most $n_i = n/k^i$ with $\epsilon_i = \epsilon(k/2)^i$. For each such subproblem, we partition its n_i rectangles by k lines, and construct a $\epsilon_i/4$-Mnet for the rectangles intersecting of these k lines. Note that the set of rectangles intersecting any line, and clipped to one side of the line have linear union complexity and by our Theorem on the dual set-systems, there exists a $\epsilon_i/4$-Mnet of size $O(1/\epsilon_i)$. Hence the total size over all internal subproblems is:

$$\sum_{i=0}^{d} k^i \cdot k \cdot O(\frac{1}{\epsilon_i}) = \sum_{i=0}^{d} k^{i+1} \cdot O(\frac{2^i}{\epsilon k^i}) = \sum_{i=0}^{d} O(\frac{2^i}{\epsilon^{1+1/d}}) = O(\frac{2^d}{\epsilon^{1+1/d}})$$

After d steps, we have k^d subproblems, each with at most n/k^d rectangles, and $\epsilon_d = \epsilon(k/2)^d$. Now just use a direct construction which constructs an ϵ-Mnet of size $O(1/\epsilon^2)$, to get the total size of Mnet at the last step to be $O(\frac{k^d}{\epsilon_d^2}) = O(\frac{4^d}{\epsilon^2 k^d}) = O(4^d/\epsilon)$.

At any level i, we construct a ϵ_i-Mnet on a set of at most n/k^i rectangles. So each set in the Mnet has size $\epsilon_i \cdot n/k^i = O(\epsilon n/2^i)$. ◀

▶ Claim 2. Each point in at least ϵn rectangles of R contains a set of $\mathcal{U}$.

Proof. Take a point q lying in at least ϵn rectangles of R. At the 0-th level, say q lies in the vertical slab defined by lines l_j^0 and l_{j+1}^0. If q hits at least $\epsilon n/4$ rectangles intersected by either l_j^0 or l_{j+1}^0, say l_j^0, then it hits at least $\epsilon n/4$ rectangles out of a total of at most n rectangles intersected by l_j^0. So the $(\epsilon_i/4 = \epsilon/4)$-Mnet for l_j^0 will have a set contained by q. Otherwise q hits at least $\epsilon_0 n_0/2 = \epsilon n/2 = \epsilon(k/2)(n/k) = \epsilon_1 n_1$ rectangles of the set R_j^0 of size $n_1 = n_0/k$, and we proceed to this subproblem.

In general, at the i-level, each subproblem has $n_i = n/k^i$ rectangles, with $\epsilon_i = \epsilon(k/2)^i$. Then either q hits at least $\epsilon_i \cdot n_i/4$ rectangles intersecting one of the lines, and so will contain a set from the $\epsilon_i/4$-Mnet constructed for each of the k vertical lines. Or q contains at least $\epsilon_i n_i/2$ rectangles out of a total of n_i/k rectangles lying in one of the slabs defined by the k vertical lines. But as

$$\epsilon_i n_i/2 = \epsilon/2 \cdot (k/2)^i \cdot n/k^i = \epsilon(k/2)^{i+1} n/k^{i+1} = \epsilon_{i+1} n_{i+1}$$

q will be covered inductively by the ϵ_{i+1}-Mnet constructed for the $n_{i+1} = n/k^{i+1}$ rectangles in one of the resulting subproblems at level $i+1$. ◀

Finally we present the tight upper-bound for the primal case of axis-parallel rectangles in Theorem 1 part 2.

Assume $P = \{p_1, \ldots, p_n\}$ are sorted by their x-coordinates. Given P, construct the balanced binary subdivision of P with vertical lines: divide P by a vertical line into two equal-sized subsets P_0^0, P_1^0, and then recursively divide each of these sets for $\log 1/\epsilon$ levels. Let P_j^i be the j-th resulting subset of P at level i, i.e., $P_j^i = \{p_{jn/2^i}, \ldots, p_{(j+1)n/2^i-1}\}$.

For each set P_j^i, and for each of its two bounding lines l_0 and l_1 in the binary subdivision above, construct a $2^{i-1}\epsilon$-Mnet for the following primal set-system: the base set is P_j^i, and given the line $l \in \{l_0, l_1\}$, the sets are induced by rectangles intersecting the line l. Note that all points of P_j^i lie on the same side of l. Let $\mathcal{U}$ be the union of all these Mnets . By Theorem 2, a ϵ-Mnet for such a set-system has size $O(1/\epsilon)$.

We now prove that $\mathcal{U}$ is an ϵ-Mnet of P, w.r.t. axis-parallel rectangles, of size $O(1/\epsilon \log 1/\epsilon)$.

▶ Claim 3. Each set in $\mathcal{U}$ has size $\Theta(\epsilon n)$, and size of $\mathcal{U}$ is $O(1/\epsilon \log 1/\epsilon)$.

Proof. P_j^i has $n/2^i$ points, and so a $(2^{i-1}\epsilon)$-Mnet of P_j^i has sets of size $O(2^{i-1}\epsilon \cdot n/2^i) = O(\epsilon n)$. Each such $2^{i-1}\epsilon$-Mnet has size $O(1/2^i\epsilon)$, there are 2^i sets P_j^i at level i, and a total of $\log 1/\epsilon$ levels. Hence the size of $\mathcal{U}$ is $O(1/2^i\epsilon \cdot 2^i \cdot \log 1/\epsilon) = O(1/\epsilon \log 1/\epsilon)$. ◀

▶ Claim 4. Each axis-parallel rectangle containing ϵn points of P contains a set of $\mathcal{U}$.

Proof. Let R be an axis-parallel rectangle containing ϵn points of P. Let i be the smallest index such that R intersects exactly one vertical line separarting two sets P_j^i and P_{j+1}^i at level i. Say R intersects the line l separating P_j^i and P_{j+1}^i. Then R must contain at least $\epsilon n/2$ points from either

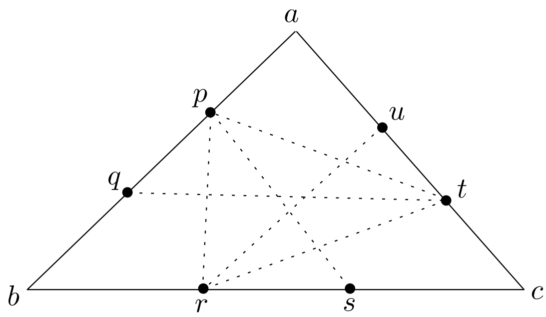

Figure 1 The union of triangles aqt,bsp, cur and prt covers the triangle abc.

P^i_j or P^i_{j+1}, say P^i_j. Let R' be the part of R on the side of l towards P^i_j. All such R' form a set of psuedo-disks, and so R' must contain at least one set of the $2^{i-1}\epsilon$-Mnet for P^i_j, as

$$|R \cap P^i_j| = |R' \cap P^i_j| \geq \epsilon n/2 = 2^{i-1}\epsilon \cdot n/2^i = 2^{i-1}\epsilon \cdot |P^i_j|$$

◀

3 Proof of Theorem 2

In this section we give the proof of Theorem 2. Given a set P of n points, first we give the proof for the most difficult case, that of the primal set-system induced by triangles, and in general, k-sided polygons in the plane. At the end we sketch out the modifications required for the rest of the cases of strips, cones and disks.

So we are given a set P of n points, and its subsets induced by the family of all k-sided polygons. The objective, as before, is to compute a small-sized ϵ-Mnet . We will assume P to be in general position.

Since a k-sided polygon can be triangulated with k triangles, any k-sided polygon containing ϵn points of P also contains a triangle containing $\epsilon n/k$ points. Hence an ϵ/k-Mnet with respect to triangles is an ϵ-Mnet with respect to k-sided polygons. We can therefore restrict ourselves to triangles.

Consider any triangle T in the plane that contains ϵn points of P. By moving the sides of the triangle we can ensure that each side of T contains at least two points of P and this can be done in such a way that no point outside T enters the interior of P. Some points in the interior of T may have moved to its boundary and some point outside T may also have moved to the boundary. Since at most 6 points may be on the boundary of T, due to P being in general position, the interior of T still contains at least $\epsilon n/2$ points assuming $\epsilon n \geq 12$. For $\epsilon n < 12$, the set P itself is an ϵ-Mnet . Thus we can further restrict ourselves to the interior of triangles each of whose sides contain at least two points. Figure 1 shows a triangle with each side containing two points of P. The points q and r could be identical, they could both be equal at the corner b of the triangle. Similarly s and t could be at c and u and p could be at a. Observe that the triangles aqt, bsp, cur and prt cover the triangle T and therefore one of them must contain at least $\epsilon n/4$ points of P. Each of these triangles are of the following type: at least two of the corners are in P and all sides contain at least two points of P. We call such triangles *anchored* triangles. Thus we can again restrict ourselves to the problem of anchored triangles containing ϵn points.

Let $\mathcal{O}$ be the set of all anchored triangles. Let $\mathcal{O}' = \{o_1, \ldots, o_t\}$ be a maximal set of t triangles from

$\mathcal{O}$ such that $o_i \cap P = \epsilon n$ and $|o_i \cap o_j \cap P| \leq \epsilon n/2$.

▶ Claim 5. $|\mathcal{O}'| \leq 2 \cdot f(\frac{c}{\epsilon} \cdot \log 1/\epsilon, 2c \log 1/\epsilon)$, where $f(n, l)$ is the maximum number of $\leq l$-sized subsets induced by objects in $\mathcal{O}$ given any set of n points, and c is some fixed constant.

Proof. The proof is via the probabilistic method. Pick each point of P independently at random with probability $p = c/(2\epsilon n) \cdot \log 1/\epsilon$ to get a random sample S.

▶ Fact 1. With probability at least $1/2$, the sets $o_i \cap S$, $i = 1 \ldots t$, are distinct and $|S| \leq c/\epsilon \cdot \log 1/\epsilon$.

Proof. Consider the range space $(P, \mathcal{R}')$, where $\mathcal{R}' = \{(o_i \setminus o_j) \cap P \mid \forall 1 \leq i < j \leq t\}$. First note that from the definition of $\mathcal{O}'$, we get that each set in $\mathcal{R}'$ has size at least $\epsilon n - \epsilon n/2 = \Theta(\epsilon n)$. Second, we use the fact that ranges induced by polygons with k sides have VC-dimension at most $2k+1$ [14]; it is easy to see that $\mathcal{R}'$ is a subset of the ranges induced by polygons (or union of polygons) with at most 9 sides (overall), and so the VC-dimension of $\mathcal{R}'$ is at most 19. Then by the Haussler-Welzl theorem [10], for $c > 19 \cdot 4$, with probability at least $3/4$, S is an ϵ-net for $(P, \mathcal{R}')$. Now observe that if $o_i \cap S = o_j \cap S$, then the set $(o_i \setminus o_j) \cap S$ is empty, a contradiction to the fact that S is an ϵ-net for $\mathcal{R}'$.

Finally, from standard concentration estimates from Chernoff bounds, it follows that $|S| \geq c/\epsilon \cdot \log 1/\epsilon$ with probability at most $1/4$. ◀

For each $o_i \in \mathcal{O}'$, let X_i be the random variable which is 1 if $|o_i \cap S| \geq 2c \log 1/\epsilon$, and 0 otherwise. Then

▶ Fact 2. With probability greater than $1/2$, $\sum X_i \leq t/2$.

Proof. For a fixed i, by linearity of expectation:

$$E[|o_i \cap S|] = c/2 \cdot \log 1/\epsilon$$

By Markov's inequality applied to each X_i,

$$Pr[X_i = 1] = Pr[|o_i \cap S| \geq 2c \cdot \log 1/\epsilon] = Pr[|o_i \cap S| \geq 4 \cdot E[|o_i \cap S|]] \leq 1/4$$

Hence the expected value of $Y = \sum X_i$ is:

$$E[\sum X_i] = \sum E[X_i] = \sum Pr[X_i = 1] \leq t/4$$

By Markov's inequality applied to Y, we get that

$$Pr[\sum X_i \geq t/2] \leq E[\sum X_i]/(t/2) \leq 1/2$$

So with probability greater than $1/2$, at least half the sets of $\mathcal{O}'$ contain at most $2c \log 1/\epsilon$ points of S. ◀

Therefore, putting together Fact 1 and Fact 2, there exists a subset S of size $(c/\epsilon) \log 1/\epsilon$ such that $o_i \cap S$ are distinct for all objects in $\mathcal{O}'$, and for at least $|\mathcal{O}'|/2$ of the objects in $\mathcal{O}'$, we have $|o_i \cap S| \leq 2c \log 1/\epsilon$.

Let $f(n, l)$ be the number of distinct subsets of size at most l that can be achieved by intersection with objects in $\mathcal{O}$. These are called $\leq l$-sets (the most extensively studied case is for halfspaces in $\mathbb{R}^d$). So in our case above, each $o_i \cap S$ formed by these $|\mathcal{O}'|/2$ objects is a $\leq l$-set of S, where $l = 2c \log 1/\epsilon$. By the bound on number of $\leq l$-sets for k-sided polygons, we get

$$|\mathcal{O}'|/2 = f(|S|, l) = f((c/\epsilon) \log 1/\epsilon, 2c \log 1/\epsilon)$$

This gives the required bound on $|\mathcal{O}'|$. ◀

Take this set $\mathcal{O}'$ of maximal objects, each containing ϵn points of P, and every pair of objects in $\mathcal{O}'$ intersecting in less than $\epsilon n/2$ points. For each object $o_i \in \mathcal{O}$, do the following: apply the simplicial partition theorem to $o_i \cap P$ with the parameter t, which is a large enough constant, to get a partition of $o_i \cap P$ into t sets of size $\Theta(|o_i \cap P|/t)$. Add each of these t sets to the ϵ-Mnet $\mathcal{U}$ for P.

▶ Claim 6. $\mathcal{U}$ is an ϵ-Mnet for the primal set-system defined by P and $\mathcal{O}$, of size $O(|\mathcal{O}'|)$.

Proof. First note that each set added to $\mathcal{U}$ had size $\Theta(|o_i \cap P|/t) = \Theta(\epsilon n)$. And the number of such sets is $O(|\mathcal{O}'| \cdot t) = O(|\mathcal{O}'|)$. It remains to show that any object containing ϵn points of P contain one set of $\mathcal{U}$.

Take any object o containing ϵn points of P. By the maximality of $\mathcal{O}'$, there exists $o_i \in \mathcal{O}'$ such that $|o \cap o_i| \geq \epsilon n/2$. Furthermore, of all the sets in the simplicial partition of o_i, each edge of ∂o can intersect only $O(\sqrt{t})$ sets; so in total the boundary of o can intersect at most $O(3\sqrt{t})$ sets. Each of these sets has $O(|o_i \cap P|/t)$ points. So these sets can contribute at most $O(3\sqrt{t} \cdot |o_i \cap P|/t)$ points of o_i to the object o. Setting t to be a large-enough constant (say, $t = 37$), this is less than $\epsilon n/2$. Therefore o must contain a point in o_i which lies in a partition for o_i not intersecting ∂o, i.e., the partition lies completely inside o. ◀

▶ Claim 7. $f(n, l) \leq ln^3$.

Proof. The proof is folklore, and follows by standard application of the Clarkson-Shor method [14]. For completeness we sketch it here. An anchored triangle abc can be of two types - either all corners are in P or exactly two corners, say a and b, are in P and there is a point $p \in P$ on ac and another point $q \in P$ on bc. The number of anchored triangles of the first type is clearly at most $\binom{n}{3}$. Thus we only need to bound the number of anchored triangles of the second type with at most l points in the interior. We first consider the case when $l = 0$, i.e., anchored triangles of the second type with no point of P in the interior. For such triangles, observe that fixing the points a, b and p determines q. If there were two points q and q' then it can be easily shown that one of anchored triangles T_1 determined by a, b, p and q and T_2 determined by a, b, p and q' is non-empty - either T_1 contains q' or T_2 contains q. Thus the number of such triangles is at most $\binom{n}{3}$.

Let N denote the number of anchored triangles of the second type with at most l points in the interior. Let Q be a subset of P obtained by picking each point of P independently with probability $p = 1/l$. The expected number of empty anchored triangles of the second type determined by Q is at most the expected number of triples in Q which is $p^3\binom{n}{3}$ since every triple in P appears as a triple in Q with probability p^3. At the same time, each of the N anchored triangles with at most l points in the interior becomes an empty anchored triangles in Q with probability $p^4(1-p)^l$. Thus the expected number of empty anchored triangles in Q is at least $Np^4(1-p)^l$. Thus $Np^4(1-p)^l \leq p^3\binom{n}{3}$. Since $p = 1/l$, it follows that $N = O(ln^3)$. ◀

Triangles and k-sided polygons. Finally, the proof for the size of ϵ-Mnet for triangles and k-sided polygons follows from Claims 5, 6 and 7. We now sketch the proof of the other cases along the above lines.

Lines, strips, cones. For sets induced by lines, strips, cones in the plane, one can follow the above proof. The function $f(n, l)$ correspondingly denotes the number of subsets of size l induced by the objects of the appropriate type (lines, strips, cones). For lines, $f(n, l) = O(n^2)$, for strips it is $O(n^2 l)$ and for cones it is $O(n^2 l^2)$. The proof then follows from the above claims.

Rectangles intersecting a common line l. As each rectangle contains ϵn points and intersects l, for each rectangle R, take the portion of the rectangle on the side of l that contains at least $\epsilon n/2$ points. We can construct $\epsilon n/2$-Mnets for the two sides of l separately.

So consider the rectangles anchored on l lying on the same side containing $\epsilon n/2$ points of P. Call this set $\mathcal{O}$. As before, let $\mathcal{O}'$ be the maximal subset of $\mathcal{O}$ such that $i)$ every pair of rectangles in $\mathcal{O}'$ share at most $\epsilon n/10$ points, and $ii)$ each rectangle contains $\epsilon n/2$ points. These form *pseudo-disks* (i.e., no two rectangles pierce each other) and by the result of [16], $|\mathcal{O}'| = O(1/\epsilon)$. Now Claim 6 implies that one can construct $\epsilon/2$-Mnet of size $O(1/\epsilon)$.

Disks. By standard Veronese map, points P and disks D can be lifted to halfspaces H in $\mathbb{R}^3$ such that each point is lifted to a point in $\mathbb{R}^3$ and each disk is lifted to a halfspace in $\mathbb{R}^3$ in such a way that their incidences are preserved. Now the required bound follows from the result for halfspaces in $\mathbb{R}^3$.

Lower-bounds.

▶ **Theorem 4.** *For every $\epsilon > 0$ and k an integer, there exists a set P of n points in the plane, and a set $\mathcal{D}$ of $\Omega(\frac{1}{\epsilon^{d+1}})$ curves of degree at most d, such that $i)$ each curve contains ϵn points of P and, $ii)$ no two curves share more than $\epsilon n/k$ points of P.*

Proof. For the lower-bound on the size of $\frac{1}{k}$ heavy ϵ-$Mnet$, consider the grid $G = [dk] \times [\frac{1}{\epsilon}]$ in the plane for some $d \geq 1$, where $[r]$ denotes the set $\{0, \cdots, \lceil r \rceil - 1\}$. Now, consider univariate functions of x of the form $y = \sum_{i=0}^{d} a_i x^i$ where each a_i is an integer in $[\frac{1}{\epsilon(d+1)(dk)^i}]$. Clearly there are at least $\Omega(\prod_{i=0}^{d} \frac{1}{\epsilon(d+1)(dk)^i}) = \Omega(\frac{1}{\epsilon^{d+1}})$ of these polynomials. Since for each value of $x \in [dk]$, the value of y is in $[\frac{1}{\epsilon}]$, each of these curves contain dk points of G. Also, since these are curves of degree at most d, no two intersect in more than d points. Let P be the set of n points containing $\epsilon n/dk$ copies of each of the points in G. We thus get a set of $\Omega(\frac{1}{\epsilon^{d+1}})$ curves of degree at most d, each of which contain ϵn points of P and no two of which share more than $\epsilon n/k$ points of P. ◀

▶ **Corollary 5.** *This gives a lower bound of $\Omega(\frac{1}{\epsilon^{d+1}})$ for $\frac{1}{k}$-heavy ϵ-$Mnets$ for range spaces induced by curves of degree at most d in the plane.*

Note that this immediately implies that for sets induced by lines in the plane, ϵ-Mnets must have size $\Omega(1/\epsilon^2)$. Which in turn is a special case for strips and cones in the plane.

▶ **Corollary 6.** *Any ϵ-Mnet for sets induced by lines, strips and cones in the plane must have size $\Omega(1/\epsilon^2)$.*

Finally, using standard linearization [14] (with Veronese maps), it is possible to lift a set of polynomial curves of degree d and a set of points to R^{3d+2} so that each point in the plane is lifted to a point in R^{3d+2} and each curve is lifted to a halfspace (i.e., the curve $y = f(x)$ becomes $(y - f(x))^2 \leq 0$, and each monomial of this expansion can be treated as a different coordinate for the linearization). Thus we have the following:

▶ **Corollary 7.** *Any ϵ-Mnet for sets induced by halfspaces in $\mathbb{R}^d$ must have size $\Omega(\frac{1}{\epsilon^{\lceil (d-1)/3 \rceil}})$.*

References

1 Boris Aronov, Esther Ezra, and Micha Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.

2 Sunil Arya, Guilherme Dias da Fonseca, and David M. Mount. Optimal area-sensitive bounds for polytope approximation. In *Symposium on Computational Geometry*, pages 363–372, 2012.

3 I. Barany. Random polytopes, convex bodies, and approximation. In W. Weil, editor, *Stochastic Geometry*, pages 77–118. Springer, 2007.

4 I. Barany and D. G. Larman. Convex bodies, economic cap coverings, random polytopes. *Mathematika*, 35:274—-291, 1988.

5 Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *SODA*, 2012.

6 Chandra Chekuri, Kenneth L. Clarkson, and Sariel Har-Peled. On the set multi-cover problem in geometric settings. In *Symposium on Computational Geometry*, pages 341–350, 2009.

7 Kenneth L. Clarkson and Kasturi R. Varadarajan. Improved approximation algorithms for geometric set cover. In *Symposium on Computational Geometry*, pages 135–141, 2005.

8 G. Ewald, D. G. Larman, and C. A. Rogers. The directions of the line segments and of the r-dimensional balls on the boundary of a convex body in euclidean space. *Mathematika*, 17:1—-20, 1970.

9 B. Chazelle H. Bronnimann and J. Pach. How hard is halfspace range searching. *Discrete Comput. Geom.*, 10:143—155, 1993.

10 D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

11 J. Komos, J. Pach, and G. Woeginger. Almost tight bounds for epsilon nets. *Discrete & Computational Geometry*, pages 163–173, 1992.

12 A. M. Macbeath. A theorem on non-homogeneous lattices. *Annals of Math*, 56:269—-293, 1952.

13 Jiri Matousek. Reporting points in halfspaces. *Computational Geometry*, 2(3):169 – 186, 1992.

14 Jirí Matousek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.

15 János Pach and Gábor Tardos. Tight lower bounds for the size of epsilon-nets. In *Symposium on Computational Geometry*, pages 458–463, 2011.

16 Evangelia Pyrga and Saurabh Ray. New existence proofs epsilon-nets. In *Symposium on Computational Geometry*, pages 199–207, 2008.

17 Kasturi Varadarajan. Weighted geometric set cover via quasi-uniform sampling. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 641–648, 2010.