Shallow Packings and their Applications

Kunal Dutta

INRIA, DATASHAPE group
(Based on joint works with Arijit Ghosh, Bruno Jartoux and Nabil Mustafa)

GUDHI Workshop, Porquerolles
October, 2016

Flow

Geometric Set Systems

$(X, \mathcal{R})$: X - a universe of “elements”, e.g. points in $\mathbb{R}^d$ or regions of $\mathbb{R}^d$.

$\mathcal{R}$: a collection of subsets of the universe, having some geometric property, e.g. induced by containment or intersections with a given family of geometric objects.

Geometric Set Systems

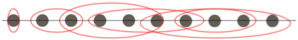

Figure: Points and intervals in $\mathbb{R}$

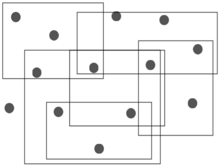

Figure: Points and axis-parallel rectangles in $\mathbb{R}^2$

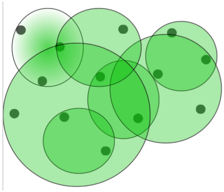

Figure: Points and disks in $\mathbb{R}^2$

Shallow Packings

Given a set system $(X, \mathcal{R})$, a δ-packing $\mathcal{P} \subset \mathcal{R}$ is a subcollection of sets such that

$$\{\forall S_1, S_2 \in \mathcal{P} : |S_1 \Delta S_2| > \delta\}.$$

A k-shallow δ-packing is a δ-packing, all of whose sets contain at most k points.

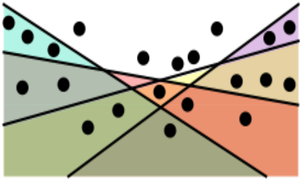

Figure: 5-packing with points and half-spaces (lines) in $\mathbb{R}^2$

Projections

Given a set system $(X, \mathcal{R})$, and a set $Y \subset X$, the *projection* of $\mathcal{R}$ on Y is

$$\mathcal{R}|_Y := \{S \cap Y | S \in \mathcal{R}\}.$$

Shallow Cell Complexity

A set system $(X, \mathcal{R})$ has *shallow cell complexity* $\varphi(.,.)$, if for any $Y \subseteq X$, the number of subsets in $\mathcal{R}|_Y$ of size at most k, is $O(|Y| \cdot \varphi(|Y|, k))$.

If the dependence of φ on k is polynomial, we'll drop the second parameter and call it $\varphi(.)$.

Primal shatter dimension d: If $\varphi(m, k) = O(m^{d-1})$, for any $1 \leq k \leq m$.

Clarkson-Shor property (d, d_1): If $\varphi(m, k) = O(m^{d_1 - 1} k^{d - d_1})$.

Shallow Cell Complexity

A set system $(X, \mathcal{R})$ has *shallow cell complexity* $\varphi(.,.)$, if for any $Y \subseteq X$, the number of subsets in $\mathcal{R}|_Y$ of size at most k, is $O(|Y| \cdot \varphi(|Y|, k))$.

If the dependence of φ on k is polynomial, we'll drop the second parameter and call it $\varphi(.)$.

Primal shatter dimension d: If $\varphi(m, k) = O(m^{d-1})$, for any $1 \leq k \leq m$.

Clarkson-Shor property (d, d_1): If $\varphi(m, k) = O(m^{d_1 - 1} k^{d - d_1})$.

Shallow Cell Complexity

A set system $(X, \mathcal{R})$ has *shallow cell complexity* $\varphi(.,.)$, if for any $Y \subseteq X$, the number of subsets in $\mathcal{R}|_Y$ of size at most k, is $O(|Y| \cdot \varphi(|Y|, k))$.

If the dependence of φ on k is polynomial, we'll drop the second parameter and call it $\varphi(.)$.

Primal shatter dimension d: If $\varphi(m, k) = O(m^{d-1})$, for any $1 \leq k \leq m$.

Clarkson-Shor property (d, d_1): If $\varphi(m, k) = O(m^{d_1 - 1} k^{d - d_1})$.

Shallow Cell Complexity

A set system $(X, \mathcal{R})$ has *shallow cell complexity* $\varphi(., .)$, if for any $Y \subseteq X$, the number of subsets in $\mathcal{R}|_Y$ of size at most k, is $O(|Y| \cdot \varphi(|Y|, k))$.

If the dependence of φ on k is polynomial, we'll drop the second parameter and call it $\varphi(.)$.

Primal shatter dimension d: If $\varphi(m, k) = O(m^{d-1})$, for any $1 \leq k \leq m$.

Clarkson-Shor property (d, d_1): If $\varphi(m, k) = O(m^{d_1 - 1} k^{d - d_1})$.

Packing Theorems

Theorem[Haussler, 1995] Given a set system having primal shatter dimension at most d (i.e. shallow cell complexity m^{d-1}), any δ-packing has size at most $O\left(\left(\frac{n}{\delta}\right)^d\right)$.

Theorem Ezra '14, D.-Ezra-Ghosh '15, Mustafa '15: Any set system having primal shatter dimension at most d and shallow cell complexity $\varphi(.,.)$, can have (k,δ) packings of size at most $\frac{24dn}{\delta}\varphi\left(\frac{4dn}{\delta},\frac{12dk}{\delta}\right)$.

Packing Theorems

Theorem[Haussler, 1995] Given a set system having primal shatter dimension at most d (i.e. shallow cell complexity m^{d-1}), any δ-packing has size at most $O\left(\left(\frac{n}{\delta}\right)^d\right)$.

Theorem Ezra '14, D.-Ezra-Ghosh '15, Mustafa '15: Any set system having primal shatter dimension at most d and shallow cell complexity $\varphi(.,.)$, can have (k,δ) packings of size at most $\frac{24dn}{\delta}\varphi\left(\frac{4dn}{\delta}, \frac{12dk}{\delta}\right)$.

Our results: Lower bounds, applications etc.

Lower bounds: Tight upto constants (in d).

Bounds on (all known classes of) ε-nets as functions of shallow cell complexity.
(Summarises 30 years of progress!)

Bounds on M-nets (combinatorial analogs of Macbeath regions) as functions of shallow cell complexity.

Our results: Lower bounds, applications etc.

Lower bounds: Tight upto constants (in d).

Bounds on (all known classes of) ε-nets as functions of shallow cell complexity.
(Summarises 30 years of progress!)

Bounds on M-nets (combinatorial analogs of Macbeath regions) as functions of shallow cell complexity.

Our results: Lower bounds, applications etc.

Lower bounds: Tight upto constants (in d).

Bounds on (all known classes of) ε-nets as functions of shallow cell complexity.
(Summarises 30 years of progress!)

Bounds on M-nets (combinatorial analogs of Macbeath regions) as functions of shallow cell complexity.

Flow

Lower Bound Result

(D.-Ezra-Ghosh, SoCG'15): The upper bound on shallow packings is tight upto constants, when $\varphi(m, k) = O(mk^2)$.

Theorem(D.-Ghosh-Jartoux-Mustafa, '16+) For any positive integers d, d_1, n, $d \geq d_1$, there exists a set system $(X, \mathcal{R})$ on n elements such that

(i) $(X, \mathcal{R})$ has shallow cell complexity $\psi(m, r) = m^{d_1 - 1} r^{d - d_1}$, and,

(ii) for any k and δ such that $\delta \leq \frac{k}{4d}$, $(X, \mathcal{R})$ has a k-shallow δ-packing of size $\Omega\left(\frac{n^{d_1 - 1} k^{d - d_1}}{\delta^d}\right)$.

Construction: $(1,1)$-Clarkson-Shor system

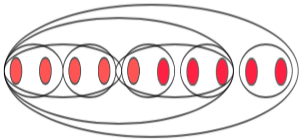

Figure: $(1,1)$-Clarkson-Shor system

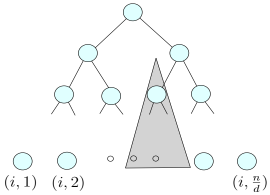

Figure: $(1,1)$-CS system: Alternate view

For each node v of the tree T, there is an edge in $\mathcal{P}$ consisting of the leaves of the sub-tree rooted at v.

$(1,1)$-Clarkson-Shor system (contd.)

Observation: For any set $S \subseteq X$, and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{S,\leq r}| < 2|S|$.

Proof: easy if the leaves in $\mathcal{P}|_S$ form a connected component.

Otherwise: induct on the number of connected components.

Crucial: independent of r.

$(1,1)$-Clarkson-Shor system (contd.)

Observation: For any set $S \subseteq X$, and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{S,\leq r}| < 2|S|$.

Proof: easy if the leaves in $\mathcal{P}|_S$ form a connected component.

Otherwise: induct on the number of connected components.

Crucial: independent of r.

$(1, 1)$-Clarkson-Shor system (contd.)

Observation: For any set $S \subseteq X$, and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{S,\leq r}| < 2|S|$.

Proof: easy if the leaves in $\mathcal{P}|_S$ form a connected component.

Otherwise: induct on the number of connected components.

Crucial: independent of r.

$(1,1)$-Clarkson-Shor system (contd.)

Observation: For any set $S \subseteq X$, and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{S,\leq r}| < 2|S|$.

Proof: easy if the leaves in $\mathcal{P}|_S$ form a connected component.

Otherwise: induct on the number of connected components.

Crucial: independent of r.

$(1, 0)$-Clarkson-Shor system

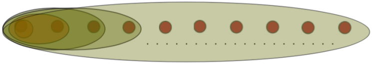

Figure: $(1, 0)$-Clarkson-Shor system

$$\mathcal{Q} = \{\{1\}, \{1, 2\}, \{1, 2, 3\}, \ldots, \{1, 2, \ldots, n/d\}\}.$$

Combining $(1,1)$- and $(1,0)$- CS systems

Take d_1 copies of $(X,\mathcal{P})_{i=1}^{d_1}$, and $d-d_1$ copies of $(Y,\mathcal{Q})_{i=d_1+1}^{d}$:

$Z = \prod_{i=1}^{d_1} X_i \times \prod_{d_1+1}^{d} Y_i$.

$\mathcal{R} = \{r_1 \cup r_2 \cup \ldots r_d | (r_1,\ldots,r_d) \in \mathcal{P}_1 \times \ldots \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \ldots \mathcal{Q}_d\}$.

For any $S \subseteq Z$,

$$\begin{aligned} \mathcal{R}|_{S,\le l}| &\le \left(\prod_{i=1}^{d_1} \mathcal{P}_i|_{X_i,\le l}\right) \times \left(\prod_{d_1+1}^{d} \mathcal{Q}_i|_{Y_i,\le l}\right) \\ &\le (2m)^{d_1} \times (l)^{d-d_1} \end{aligned}$$

Combining $(1,1)$- and $(1,0)$- CS systems

Take d_1 copies of $(X, \mathcal{P})_{i=1}^{d_1}$, and $d - d_1$ copies of $(Y, \mathcal{Q})_{i=d_1+1}^{d}$:

$Z = \prod_{i=1}^{d_1} X_i \times \prod_{d_1+1}^{d} Y_i$.

$\mathcal{R} = \{r_1 \cup r_2 \cup \dots r_d | (r_1, \dots, r_d) \in \mathcal{P}_1 \times \dots \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \dots \mathcal{Q}_d\}$.

For any $S \subseteq Z$,

$$\begin{aligned} \mathcal{R}|_{S,\le l}| &\le \left(\prod_{i=1}^{d_1} \mathcal{P}_i|_{X_i,\le l}\right) \times \left(\prod_{d_1+1}^{d} \mathcal{Q}_i|_{Y_i,\le l}\right) \\ &\le (2m)^{d_1} \times (l)^{d-d_1} \end{aligned}$$

Combining $(1,1)$- and $(1,0)$- CS systems

Take d_1 copies of $(X,\mathcal{P})_{i=1}^{d_1}$, and $d-d_1$ copies of $(Y,\mathcal{Q})_{i=d_1+1}^{d}$:

$Z = \prod_{i=1}^{d_1} X_i \times \prod_{d_1+1}^{d} Y_i$.

$\mathcal{R} = \{r_1 \cup r_2 \cup \ldots r_d | (r_1,\ldots,r_d) \in \mathcal{P}_1 \times \ldots \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \ldots \mathcal{Q}_d\}$.

For any $S \subseteq Z$,

$$\begin{aligned}\mathcal{R}|_{S,\le l}| &\le \left(\prod_{i=1}^{d_1} \mathcal{P}_i|_{X_i,\le l}\right) \times \left(\prod_{d_1+1}^{d} \mathcal{Q}_i|_{Y_i,\le l}\right) \\ &\le (2m)^{d_1} \times (l)^{d-d_1}\end{aligned}$$

Combining $(1,1)$- and $(1,0)$- CS systems

Take d_1 copies of $(X,\mathcal{P})_{i=1}^{d_1}$, and $d-d_1$ copies of $(Y,\mathcal{Q})_{i=d_1+1}^{d}$:

$Z = \prod_{i=1}^{d_1} X_i \times \prod_{d_1+1}^{d} Y_i$.

$\mathcal{R} = \{r_1 \cup r_2 \cup \ldots r_d | (r_1, \ldots, r_d) \in \mathcal{P}_1 \times \ldots \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \ldots \mathcal{Q}_d\}$.

For any $S \subseteq Z$,

$$\begin{aligned} \mathcal{R}|_{S,\leq l}| &\leq \left(\prod_{i=1}^{d_1} \mathcal{P}_i|_{X_i,\leq l}\right) \times \left(\prod_{d_1+1}^{d} \mathcal{Q}_i|_{Y_i,\leq l}\right) \\ &\leq (2m)^{d_1} \times (l)^{d-d_1} \end{aligned}$$

Combining $(1,1)$- and $(1,0)$- CS systems (contd.):

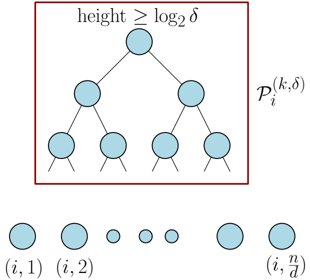

Figure: Height: $\geq \log_2 \delta$

$\mathcal{P}_i^{k,\delta} = \{\{i\} \times \{2^\alpha\beta + 1, \ldots, 2^\alpha(\beta+1)\} | \log_2 \delta \leq \alpha \leq \log_2(k/d), 0 \leq \beta \leq 2^{-\alpha}(n/d)\}$.

$\mathcal{Q}_i^{k,\delta} = \{\{i\} \times \{1, 2, \ldots, \gamma\delta\} | 1 \leq \gamma \leq k/(d\delta)\}$.

Combining $(1,1)$- and $(1,0)$- CS systems (contd.):

$$
\begin{aligned}
|\mathcal{R}| &\geq (2\delta)^{-d_1}\left(\frac{n}{d}\right)^{d_1}\left(\frac{k}{d\delta}\right)^{d-d_1} \\
&\geq 2^{-d_1} d^{-d}\left(\frac{n^{d_1}k^{d-d_1}}{\delta^{d}}\right)
\end{aligned}
$$

Flow

What are ε-nets?

ε-nets: Subset of elements hitting all sets having at least εn elements.

Haussler-Welzl ('87): Let $(X, \mathcal{R})$ be a set system with $\varphi(m) = O(m^{d-1})$ for some constant d,
let $\varepsilon > 0$ be a given parameter.
Let N be a random sample constructed by sampling each point of X randomly and independently with probability

$$\frac{4}{\varepsilon|X|}\log\frac{2}{\gamma} + \frac{8d}{\varepsilon|X|}\log\frac{8d}{\varepsilon}.$$

Then N is an ε-net with probability at least $1 - \gamma$.

More on ε-nets: Uses

Approximation algorithms,

Discrete and computational geometry,

Meshing theory,

Discrepancy theory,

Learning theory, etc.

Results on ε-nets

Papers published roughly each year since 1987, showing existence of "small" ε-nets for specific geometric set systems.

Matoušek: Multiple results in '90s, including half-planes, bounded primal shatter dimension, etc.

Clarkson-Varadarajan (2007)): Size of ε-nets for dual systems in terms of union complexity.

Ray-Pyrga (2008): Pseudo-disks have $O(1/\varepsilon)$-sized ε-nets.

Aronov-Ezra-Sharir (2010): Dual set systems actually have $\varepsilon^{-1} \log \varphi(\varepsilon^{-1})$-sized nets.

Aronov-Ezra-Sharir (2010): Axis-parallel rectangles have $\varepsilon^{-1} \log \log(\varepsilon^{-1})$-sized nets.

Result

Theorem[Chan et al '10, Varadarajan '07] (Mustafa-D.-Ghosh '16+): Shallow cell complexity $\varphi(.)$ (where $\varphi(.) = O(n^d)$) implies there exists an ε-net of size $O\left(\frac{1}{\varepsilon}\log\varphi\left(\frac{1}{\varepsilon}\right)\right)$.

Proof

Want to hit all sets of size $\geq \varepsilon n$, (assume sets have exactly εn elements).

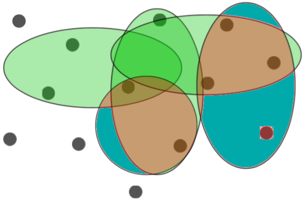

Figure: Nearest neighbours in $(3, 4)$-packing

Nearest neighbour property:
Each set S in the system has a neighbour in any maximal δ-packing $\mathcal{D}$, whose distance from S is at most δ.

Proof

So if $\delta = \varepsilon n/2$, then each set S of size εn has a neighbour in $\mathcal{D}$, which shares $\varepsilon n/2$ of its elements with S.

Suffices to take $1/2$-nets over all sets in a maximal $\varepsilon n/2$-packing!

Choose $k = \varepsilon n$ to get appropriate number of sets.

Proof

So if $\delta = \varepsilon n/2$, then each set S of size εn has a neighbour in $\mathcal{D}$, which shares $\varepsilon n/2$ of its elements with S.

Suffices to take $1/2$-nets over all sets in a maximal $\varepsilon n/2$-packing!

Choose $k = \varepsilon n$ to get appropriate number of sets.

Proof

So if $\delta = \varepsilon n/2$, then each set S of size εn has a neighbour in $\mathcal{D}$, which shares $\varepsilon n/2$ of its elements with S.

Suffices to take $1/2$-nets over all sets in a maximal $\varepsilon n/2$-packing!

Choose $k = \varepsilon n$ to get appropriate number of sets.

Proof

So if $\delta = \varepsilon n/2$, then each set S of size εn has a neighbour in $\mathcal{D}$, which shares $\varepsilon n/2$ of its elements with S.

Suffices to take $1/2$-nets over all sets in a maximal $\varepsilon n/2$-packing!

Choose $k = \varepsilon n$ to get appropriate number of sets.

Proof

Note: by Shallow Packing lemma, $(\varepsilon n, \varepsilon n/2)$-packing has size $O\left(\frac{n\varphi(n/(\varepsilon n/2))((\varepsilon n/2)^c)}{(\varepsilon n/2)^{c+1}}\right) = O\left(\frac{\varphi(1/\varepsilon)}{\varepsilon}\right)$.

Doesn't work: Too many sets in packing! :(

Idea: Do random sampling,

then use a *maximal* εn-shallow, $\varepsilon n/2$-packing $\mathcal{P}$ for correction sets (only if required!).

Proof

Note: by Shallow Packing lemma, $(\varepsilon n, \varepsilon n/2)$-packing has size $O\left(\frac{n\varphi(n/(\varepsilon n/2))((\varepsilon n/2)^c)}{(\varepsilon n/2)^{c+1}}\right) = O\left(\frac{\varphi(1/\varepsilon)}{\varepsilon}\right)$.

Doesn't work: Too many sets in packing! :(

Idea: Do random sampling,

then use a *maximal* εn-shallow, $\varepsilon n/2$-packing $\mathcal{P}$ for correction sets (only if required!).

Proof

Note: by Shallow Packing lemma, $(\varepsilon n, \varepsilon n/2)$-packing has size $O\left(\frac{n\varphi(n/(\varepsilon n/2))((\varepsilon n/2)^c)}{(\varepsilon n/2)^{c+1}}\right) = O\left(\frac{\varphi(1/\varepsilon)}{\varepsilon}\right)$.

Doesn't work: Too many sets in packing! :(

Idea: Do random sampling,

then use a *maximal* εn-shallow, $\varepsilon n/2$-packing $\mathcal{P}$ for correction sets (only if required!).

Proof

Note: by Shallow Packing lemma, $(\varepsilon n, \varepsilon n/2)$-packing has size $O\left(\frac{n\varphi(n/(\varepsilon n/2))((\varepsilon n/2)^c)}{(\varepsilon n/2)^{c+1}}\right) = O\left(\frac{\varphi(1/\varepsilon)}{\varepsilon}\right)$.

Doesn't work: Too many sets in packing! :(

Idea: Do random sampling,

then use a *maximal* εn-shallow, $\varepsilon n/2$-packing $\mathcal{P}$ for correction sets (only if required!).

Proof: Random sampling step

Randomly sample each element in X with probability:
$\frac{4}{(1/2)\varepsilon n} \log 2\varphi\left(\frac{1}{\varepsilon}\right) + \frac{8d}{(1/2)\varepsilon n} \log \frac{8d}{(1/2)}$.

By Haussler-Welzl theorem, get a $(1/2)$-net for each $S \in \mathcal{P}$, with probability at least $1 - \frac{1}{\varphi(1/\varepsilon)}$.

Proof: Random sampling step

Randomly sample each element in X with probability:
$\frac{4}{(1/2)\varepsilon n}\log 2\varphi\left(\frac{1}{\varepsilon}\right)+\frac{8d}{(1/2)\varepsilon n}\log\frac{8d}{(1/2)}$.

By Haussler-Welzl theorem, get a $(1/2)$-net for each $S\in\mathcal{P}$, with probability at least $1-\frac{1}{\varphi(1/\varepsilon)}$.

Proof: Correction step

Compute the expected size of the union of the correction sets:

$$\begin{aligned}\mathbb{E}[|M|] &= \sum_{S \in \mathcal{P}} Pr[N \text{ is not a } (1/2)\text{-net for } S] \cdot (\text{size of } (1/2)\text{-net for } S) \\ &\leq \sum_{S \in \mathcal{P}} \frac{1}{\varphi(1/\varepsilon)} \cdot O(1) = O\left(\frac{1}{\varepsilon}\right).\end{aligned}$$

Therefore $\mathbb{E}[H] = \mathbb{E}[N] + \mathbb{E}[M] = O\left(\frac{\log(\varphi(1/\varepsilon))}{\varepsilon} + \frac{1}{\varepsilon}\right)$.

Consequences: Some well-known results

Set System	Primal/Dual	Size of ϵ-nets	
Halfspaces in $\mathbb{R}^d$, $d = 2, 3$	P	$O(\frac{1}{\varepsilon})$	Pach-Woegir
Pseudodisks in $\mathbb{R}^2$	P	$O(\frac{1}{\varepsilon})$	
Union complexity $\varphi(.)$	D	$O(\frac{1}{\varepsilon}\log\varphi(\frac{1}{\varepsilon}))$	Clark
Axis parallel rectangles	P	$O(\frac{1}{\varepsilon}\log\log(\frac{1}{\varepsilon}))$	

Flow

Remarks

Also prove existence of Mnets in terms of shallow cell complexity.

Easier proof of Haussler's theorem?

Other applications?

Thanks!