Shallow Packing Lemma and its Applications in Combinatorial Geometry

Arijit Ghosh^1

¹Indian Statistical Institute Kolkata, India

- Kunal Dutta (INRIA, DataShape)
- Esther Ezra (Georgia Tech, Math Dep.)
- Bruno Jartoux (Ben-Gurion Univ., CS Dep.)
- Nabil H. Mustafa (Université Paris-Est, LIGM)

Talk will be based on the following papers

- Shallow packings, semialgebraic set systems, Macbeath regions and polynomial partitioning, with Bruno Jartoux, Kunal Dutta and Nabil Hassan Mustafa. *Discrete & Computational Geometry*, to appear.
- A Simple Proof of Optimal Epsilon Nets, with Kunal Dutta and Nabil Hassan Mustafa. *Combinatorica*, 38(5): 1269 – 1277, 2018.
- Two proofs for Shallow Packings, with Kunal Dutta and Esther Ezra. Discrete & Computational Geometry, 56(4): 910-939, 2016.

Situation map: three combinatorial structures



◆□▶ ◆□▶ ◆三▶ ◆三▶ ◆□ ◆ ◆ ◆

Situation map: three combinatorial structures



Geometric set systems

Point-disk incidences: an example of geometric set system



Geometric set systems

Point-disk incidences: an example of geometric set system



Typical applications: range searching, point set queries.









For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \ge \varepsilon$ includes one of them.

Mnets – for halfplanes

For a set *K* of *n* points and $\varepsilon > 0$, an **Mnet** is a collection of subsets of $\Theta(\varepsilon n)$ points such that any halfplane *h* with $|h \cap K| \ge \varepsilon n$ includes one of them.

For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \ge \varepsilon$ includes one of them.

Mnets – for disks

For a set *K* of *n* points and $\varepsilon > 0$, an **Mnet** is a collection of subsets of $\Theta(\varepsilon n)$ points such that any disk *h* with $|h \cap K| \ge \varepsilon n$ includes one of them.

For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \ge \varepsilon$ includes one of them.

Mnets – for [shapes]

For a set K of n points and $\varepsilon > 0$, an **Mnet** is a collection of subsets of $\Theta(\varepsilon n)$ points such that any [shape] h with $|h \cap K| \ge \varepsilon n$ includes one of them.

For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \ge \varepsilon$ includes one of them.

Mnets – for [shapes]

For a set K of n points and $\varepsilon > 0$, an **Mnet** is a collection of subsets of $\Theta(\varepsilon n)$ points such that any [shape] h with $|h \cap K| \ge \varepsilon n$ includes one of them.

Goal: discrete analogue of Macbeath's tool.

Question

What is the minimum size of an Mnet?

э

Question

What is the minimum size of an Mnet?

Theorem (D.–G.–J.–M. '17)

Semialgebraic set systems with VC-dim. $d < \infty$ and shallow cell complexity φ have an ε -Mnet of size

$$O\left(\frac{d}{\varepsilon}\cdot\varphi\left(\frac{d}{\varepsilon},d\right)\right).$$

Question

What is the minimum size of an Mnet?

Theorem (D.–G.–J.–M. '17)

Semialgebraic set systems with VC-dim. $d < \infty$ and shallow cell complexity φ have an ε -Mnet of size

$$O\left(\frac{d}{\varepsilon}\cdot\varphi\left(\frac{d}{\varepsilon},d\right)\right).$$

- 🗸 Disks
- \checkmark Rectangles
- Lines
- ✓ 'Fat' objects
- × General convex sets

Question

What is the minimum size of an Mnet?

Theorem (D.-G.-J.-M. '17)

Semialgebraic set systems with VC-dim. $d < \infty$ and shallow cell complexity φ have an ε -Mnet of size

$$O\left(\frac{d}{\varepsilon}\cdot\varphi\left(\frac{d}{\varepsilon},d\right)\right).$$

- 🗸 Disks
- Rectangles
- Lines
- ✓ 'Fat' objects
- \times General convex sets

Theorem (D.–G.–J.–M. '17)

This is tight for hyperplanes.

X := arbitrary *n*-point set $\Sigma :=$ collection of subsets of X, i.e., $\Sigma \subseteq 2^X$ The pair (X, Σ) is called a *set system* Set systems (X, Σ) are also referred to as *hypergraphs*, *range spaces*





and

$$\Sigma_Y^k := \{S \cap Y : S \in \Sigma \text{ and } |S \cap Y| \le k\}$$

VC dimension and shallow cell complexity

<u>Primal Shatter function</u> Given (X, Σ) , primal shatter function is defined as

$$\pi_{\Sigma}(m) := \max_{Y \subseteq X, |Y|=m} |\Sigma_Y|$$

VC dimension and shallow cell complexity

<u>Primal Shatter function</u> Given (X, Σ) , primal shatter function is defined as

$$\pi_{\Sigma}(m) := \max_{Y \subseteq X, |Y|=m} |\Sigma_Y|$$

VC dimension: $d_0 := \max \{m \mid \pi_{\Sigma}(m) = 2^m\}$

<u>Primal Shatter function</u> Given (X, Σ) , primal shatter function is defined as

$$\pi_{\Sigma}(m) := \max_{Y \subseteq X, \, |Y|=m} |\Sigma_Y|$$

VC dimension: $d_0 := \max \{m | \pi_{\Sigma}(m) = 2^m\}$ Sauer-Shelah Lemma: VC dim. d_0 implies for all $m \le n$, $\pi_{\Sigma}(m) \le O(m^{d_0})$. <u>Primal Shatter function</u> Given (X, Σ) , primal shatter function is defined as

$$\pi_{\Sigma}(m) := \max_{Y \subseteq X, \, |Y|=m} |\Sigma_Y|$$

VC dimension: $d_0 := \max \{m \mid \pi_{\Sigma}(m) = 2^m\}$ Sauer-Shelah Lemma: VC dim. d_0 implies for all $m \le n$, $\pi_{\Sigma}(m) \le O(m^{d_0})$.

Shallow cell complexity $\varphi(\cdot, \cdot)$ If $\forall Y \subseteq X$,

$$\left|\Sigma_{Y}^{k}\right| \leq |Y| \times \varphi(|Y|, k).$$

1. Points and half-spaces or orthants in \mathbb{R}^d

$$O(|Y|^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil})$$

2. Points and balls in \mathbb{R}^d

 $O(|Y|^{\lfloor (d+1)/2 \rfloor - 1} k^{\lceil (d+1)/2 \rceil})$

 $|Y|^{d-2+\varepsilon}k^{1-\varepsilon}$



Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, \ Y \cap S \neq \emptyset$

Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, \ Y \cap S \neq \emptyset$$



Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, Y \cap S \neq \emptyset$$



Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, \ Y \cap S \neq \emptyset$$



Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, \ Y \cap S \neq \emptyset$$

Theorem (Haussler-Welzl'87)

For a set system with VC-dimen d there exists an ε -net of size $O\left(\frac{d}{\varepsilon}\log\frac{d}{\varepsilon}\right)$

Epsilon-nets: For a set system (X, Σ) , $Y \subseteq X$ is an ε -net if

$$\forall S \in \Sigma \text{ with } |S| \geq \varepsilon n, \ Y \cap S \neq \emptyset$$

Theorem (Haussler-Welzl'87)

For a set system with VC-dimen d there exists an ε -net of size $O\left(\frac{d}{\varepsilon}\log\frac{d}{\varepsilon}\right)$

 $O\left(\frac{1}{\varepsilon}\right)$ -size ε -nets are known for special set systems Half-spaces in \mathbb{R}^2 and \mathbb{R}^3 , pseudo-disks, homothetic copies of convex objects, α -fat wedges etc ... (Matousek-Seidel-Welzl, Buzaglo-Pinchasi-Rote, Pyra-Ray, ...)

Theorem (Varadarajan'10, Aronov et al.'10, Chan et al.'12)

Let (X, Σ) be a set system with constant VC-dimen and shallow cell complexity $\varphi(\cdot)$. Then there exists an ε -net of (X, Σ) of size

$$O\left(\frac{1}{\varepsilon}\log\varphi\left(\frac{1}{\varepsilon}\right)\right).$$

Theorem (Varadarajan'10, Aronov et al.'10, Chan et al.'12)

Let (X, Σ) be a set system with constant VC-dimen and shallow cell complexity $\varphi(\cdot)$. Then there exists an ε -net of (X, Σ) of size

$$O\left(\frac{1}{\varepsilon}\log\varphi\left(\frac{1}{\varepsilon}\right)\right).$$

Remark: This result gives optimal size nets for all known geometric set systems.

Theorem (Varadarajan'10, Aronov et al.'10, Chan et al.'12)

Let (X, Σ) be a set system with constant VC-dimen and shallow cell complexity $\varphi(\cdot)$. Then there exists an ε -net of (X, Σ) of size

$$O\left(\frac{1}{\varepsilon}\log\varphi\left(\frac{1}{\varepsilon}\right)
ight).$$

Remark: This result gives optimal size nets for all known geometric set systems. For example the above result implies $\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}$ size nets for points and rectangles in plane.

Parameter: Let $\delta > 0$ be a integer parameter

 δ -separated: A set system (X, Σ) is δ -separated if for all S_1 , S_2 in Σ , if the size of the symmetric difference (Hamming distance) $S_1 \Delta S_2$ is greater than δ , i.e. $|S_1 \Delta S_2| > \delta$.

 δ -packing number: The cardinality of the largest δ -separated subcollection of Σ is called the δ -packing number of Σ .

This is analogous to packing maximum number of Euclidean balls of radius $\delta/2$ in a box with edge length *n*.



This is analogous to packing maximum number of Euclidean balls of radius $\delta/2$ in a box with edge length *n*.



In \mathbb{R}^d , this packing number is $O\left(\left(\frac{n}{\delta}\right)^d\right)$

This is analogous to packing maximum number of Euclidean balls of radius $\delta/2$ in a box with edge length *n*.



In \mathbb{R}^d , this packing number is $O\left(\left(\frac{n}{\delta}\right)^d\right)$

We have the same bound for the case of set systems with VC dimension d. (due to Haussler, Chazelle and Wernisch)

Theorem (Dutta-Ezra-G.'15 and Mustafa'16)

Let (X, Σ) be a set system with VC-dim d and shallow cell complexity $\varphi(\cdot)$ on a n-point set X. Let $\delta \ge 1$ and $k \le n$ be two integer parameters such that:

- 1. $\forall S \in \sum, |S| \leq k$, and
- 2. \sum is δ -packed.

Then

$$|\Sigma| \leq \frac{dn}{\delta} \varphi\left(\frac{dn}{\delta}, \frac{dk}{\delta}\right)$$

We can show that the above bound is tight.

Theorem (Dutta-G.-Mustafa'17)

Let (X, Σ) be a set system with VC-dim d and shallow cell complexity $\varphi(\cdot)$ on a n-point set X. Then there exists an ε -net of size

$$rac{1}{arepsilon}\logarphi\left(rac{d}{arepsilon},d
ight)+rac{d}{arepsilon}$$

Theorem (Dutta-G.-Mustafa'17)

Let (X, Σ) be a set system with VC-dim d and shallow cell complexity $\varphi(\cdot)$ on a n-point set X. Then there exists an ε -net of size

$$rac{1}{arepsilon}\logarphi\left(rac{d}{arepsilon},d
ight)+rac{d}{arepsilon}$$

Remark: The proof just uses the Shallow Packing Lemma and the Alteration Technique from *The Probabilistic Method*.

Semialgebraic set systems with VC-dim. $d < \infty$ and shallow cell complexity φ have an ε -Mnet of size

$$O\left(\frac{d}{\varepsilon}\cdot\varphi\left(\frac{d}{\varepsilon},d\right)\right)$$

Semialgebraic set systems with VC-dim. $d < \infty$ and shallow cell complexity φ have an ε -Mnet of size

$$O\left(\frac{d}{\varepsilon}\cdot\varphi\left(\frac{d}{\varepsilon},d\right)\right)$$

Remark: The proof just uses Guth-Katz's Polynomial Partitioning Theorem together with Shallow Packing Lemma.

$$egin{pmatrix} arepsilon - Mnet \ of \ size \ M \ with \ sets \ of \ size \ \geq au arepsilon n \end{pmatrix} \implies arepsilon - net \ of \ size \ rac{\log(arepsilon M)/ au + 1}{arepsilon} \end{cases}$$

This gives ε -nets of size $\frac{d}{\varepsilon} \log \varphi\left(\frac{d}{\varepsilon}, d\right)$ for semialgebraic set systems.

▲ □ ▶ ▲ □ ▶ ▲

$$egin{pmatrix} arepsilon - {\cal M}$$
 net of size $M \ {
m with}$ sets of size $\geq au arepsilon n \end{pmatrix} \implies arepsilon$ -net of size $rac{\log(arepsilon {\cal M})/ au+1}{arepsilon}$

This gives ε -nets of size $\frac{d}{\varepsilon} \log \varphi\left(\frac{d}{\varepsilon}, d\right)$ for semialgebraic set systems.

• Yields best known bounds on $\varepsilon\text{-nets}$ for geometric set systems with bounded VC-dim.

$$egin{pmatrix} arepsilon - {\cal M}$$
 net of size $M \ {
m with}$ sets of size $\geq au arepsilon n \end{pmatrix} \implies arepsilon$ -net of size $rac{\log(arepsilon {\cal M})/ au+1}{arepsilon}$

This gives ε -nets of size $\frac{d}{\varepsilon} \log \varphi\left(\frac{d}{\varepsilon}, d\right)$ for semialgebraic set systems.

• Yields best known bounds on ε -nets for geometric set systems with bounded VC-dim.

	Mnet	$\varepsilon extsf{-net}$			
Disks	ε^{-1}	ε^{-1}			
Rectangles	$\frac{1}{\epsilon} \log \frac{1}{\epsilon}$	$\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}$			
Halfspaces (\mathbf{R}^d)	$O\left(\varepsilon^{-\lfloor d/2 \rfloor} ight)$	$\frac{d}{\varepsilon}\log\frac{1}{\varepsilon}$		_	
G	CAALM Work	shop 2010	1121	-	Φ) Q

Table: Upper bounds on Mnets and ε -nets

Theorem (D.-G.-J.-M. '17)

$$\begin{pmatrix} \varepsilon \text{-Mnet of size } M \\ \text{with sets of size} \ge \tau \varepsilon n \end{pmatrix} \implies \varepsilon \text{-net of size } \frac{\log(\varepsilon M)/\tau + 1}{\varepsilon}$$

Proof.

• \mathcal{M} is such an Mnet. Let $p = \frac{1}{\tau \in n} \log(\varepsilon M)$.

▲ 同 ▶ → ● 三

Theorem (D.-G.-J.-M. '17)

$$\begin{pmatrix} \varepsilon \text{-Mnet of size } M \\ \text{with sets of size} \ge \tau \varepsilon n \end{pmatrix} \implies \varepsilon \text{-net of size } \frac{\log(\varepsilon M)/\tau + 1}{\varepsilon}$$

- \mathcal{M} is such an Mnet. Let $p = \frac{1}{\tau \varepsilon n} \log(\varepsilon M)$.
- **2** Pick every point into a sample S with probability p.

Theorem (D.-G.-J.-M. '17)

$$\begin{pmatrix} \varepsilon \text{-Mnet of size } M \\ \text{with sets of size} \ge \tau \varepsilon n \end{pmatrix} \implies \varepsilon \text{-net of size } \frac{\log(\varepsilon M)/\tau + 1}{\varepsilon}$$

- **1** \mathcal{M} is such an Mnet. Let $p = \frac{1}{\tau \varepsilon n} \log(\varepsilon M)$.
- 2 Pick every point into a sample S with probability p.

Theorem (D.-G.-J.-M. '17)

$$\begin{pmatrix} \varepsilon \text{-Mnet of size } M \\ \text{with sets of size} \ge \tau \varepsilon n \end{pmatrix} \implies \varepsilon \text{-net of size } \frac{\log(\varepsilon M)/\tau + 1}{\varepsilon}$$

- **1** \mathcal{M} is such an Mnet. Let $p = \frac{1}{\tau \varepsilon n} \log(\varepsilon M)$.
- 2 Pick every point into a sample S with probability p.
- In expectation, $|S| + |m \in \mathcal{M} : S \cap m = \emptyset| \le np + \frac{1}{\varepsilon}$.

Theorem (D.-G.-J.-M. '17)

$$\begin{pmatrix} \varepsilon \text{-Mnet of size } M \\ \text{with sets of size} \ge \tau \varepsilon n \end{pmatrix} \implies \varepsilon \text{-net of size } \frac{\log(\varepsilon M)/\tau + 1}{\varepsilon}$$

- **1** \mathcal{M} is such an Mnet. Let $p = \frac{1}{\tau \varepsilon n} \log(\varepsilon M)$.
- 2 Pick every point into a sample S with probability p.
- In expectation, $|S| + |m \in \mathcal{M} : S \cap m = \emptyset| \le np + \frac{1}{\varepsilon}$.
- so there is an ε -net of size $\leq np + \frac{1}{\varepsilon}$ (why?).

Theorem follows from the following result:

Theorem

Let (X, \mathcal{R}) be a δ -separated set system with VC dimension at most d. Then

 $|\mathcal{R}| \leq 2\mathbb{E}\left[|\mathcal{R}_{A'}|\right]$

where A' is an uniformaly random subset of X of size $\frac{4dn}{\delta} - 1$.

Theorem follows from the following result:

Theorem

Let (X, \mathcal{R}) be a δ -separated set system with VC dimension at most d. Then

 $|\mathcal{R}| \leq 2\mathbb{E}\left[|\mathcal{R}_{A'}|\right]$

where A' is an uniformaly random subset of X of size $\frac{4dn}{\delta} - 1$.

For the proof of the main result consider:

$$\mathcal{R}' := \left\{ \sigma \in \mathcal{R} : |\sigma \cap A'| > 3 \times \frac{4dk}{\delta} \right\}$$

$$\mathcal{R}'' := \mathcal{R} \setminus \mathcal{R}'$$

$$(1)$$

- Ideally we want a combinatorial proof of the Mnets bound for set systems.
- Improve the current lower bound.
- Find more applications/connections of Mnets in combinatorial geometry.

Thank you.

⊡ ► < ≣ ►