Combinatorial Macbeath Regions for Semi-Algebraic Set Systems

Arijit Ghosh[1]

[1]Indian Statistical Institute
Kolkata, India

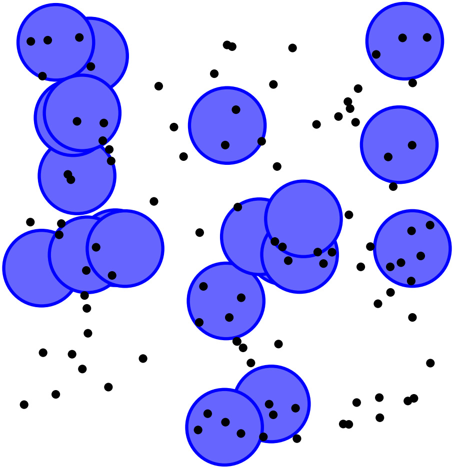
College de France 2017

Coauthors

- Kunal Dutta (DataShape, INRIA)
- Bruno Jartoux (Université Paris-Est, LIGM)
- Nabil H. Mustafa (Université Paris-Est, LIGM)

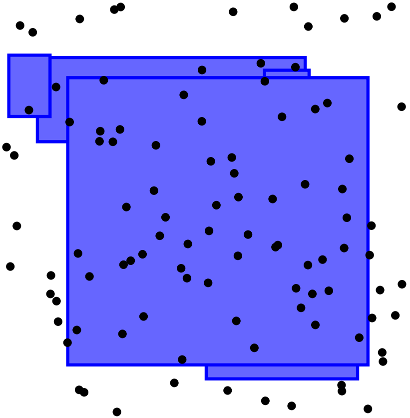
Geometric set systems

Point-disk incidences: an example of geometric *set system*.



Geometric set systems

Typical applications: range searching, point set queries.



Map

Macbeath regions

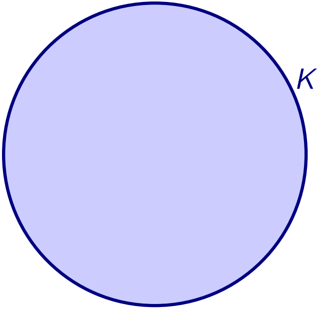
Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ contains one of them.

Macbeath regions

Macbeath decomposition (Macbeath 1952)

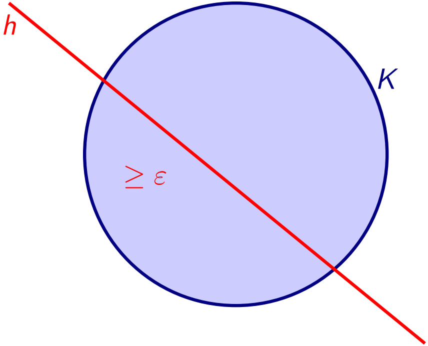
For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \geq \varepsilon$ contains one of them.



Macbeath regions

Macbeath decomposition (Macbeath 1952)

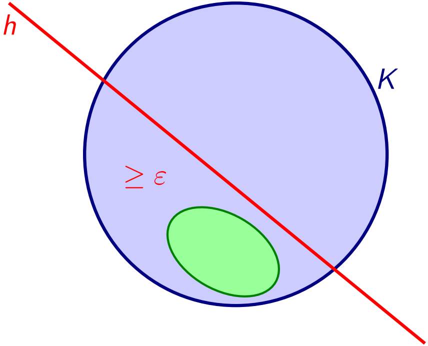
For any convex body K with unit volume and $\varepsilon > 0$, there is a *small* collection of convex subsets of K with volume $\Theta(\varepsilon)$ such that any halfplane h with $vol(h \cap K) \geq \varepsilon$ contains one of them.



Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ contains one of them.



Mnets, or combinatorial Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ includes one of them.

Mnets, or combinatorial Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ 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 $|h \cap K| \geq \varepsilon n$ includes one of them.

Mnets, or combinatorial Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ 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 $|h \cap K| \geq \varepsilon n$ includes one of them.

Mnets, or combinatorial Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ 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 $|h \cap K| \geq \varepsilon n$ includes one of them.

Mnets, or combinatorial Macbeath regions

Macbeath decomposition (Macbeath 1952)

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 $vol(h \cap K) \geq \varepsilon$ 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 $|h \cap K| \geq \varepsilon n$ includes one of them.

Goal: discrete analogue of Macbeath's tool.

Bounds on Mnets

Question

What is the minimum size of an Mnet?

Bounds on Mnets

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*

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

Bounds on Mnets

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
- × General convex sets

Bounds on Mnets

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
- × General convex sets

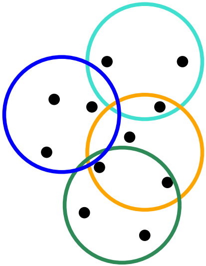
Theorem (D.–G.–J.–M. '17)

This is tight for hyperplanes.

Abstract set systems

X := arbitrary n-point set
Σ := 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*



Abstract set systems

Projection:

For $Y \subseteq X$,

$$\Sigma_Y := \{S \cap Y : \; S \in \Sigma\}$$

and

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

Primal shatter dimension and shallow cell complexity

Primal Shatter function Given (X, Σ), primal shatter function is defined as

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

Primal shatter dimension and shallow cell complexity

Primal Shatter function Given (X, Σ), primal shatter function is defined as

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

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

Primal shatter dimension and shallow cell complexity

Primal Shatter function Given (X, Σ), primal shatter function is defined as

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

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

Primal Shatter dimension: A set system has primal shatter dimension d if for all $m \leq n$, $\pi_\Sigma(m) \leq O(m^d)$.

Primal shatter dimension and shallow cell complexity

Primal Shatter function Given (X, Σ), primal shatter function is defined as

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

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

Primal Shatter dimension: A set system has primal shatter dimension d if for all $m \leq n$, $\pi_\Sigma(m) \leq O(m^d)$.

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

$$\left| \Sigma_Y^k \right| \leq |Y| \times \varphi(|Y|, k).$$

Shallow cell complexity of some geometric set systems

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 $\mathbb{R}^d$ — $O(|Y|^{\lfloor (d+1)/2 \rfloor - 1} k^{\lceil (d+1)/2 \rceil})$
3. $(d-1)$-variate polynomial function of constant degree and points in $\mathbb{R}^d$ — $|Y|^{d-2+\varepsilon} k^{1-\varepsilon}$

δ-packing number

δ-packing number

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.

Shallow packing result

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 $\delta \geq 1$ and $k \leq n$ be two integer parameters such that:

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

Then

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

Shallow packing result

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 $\delta \geq 1$ and $k \leq n$ be two integer parameters such that:

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

Then

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

We can show that the above bound is tight.

Polynomial partitioning lemma (incorrect version)

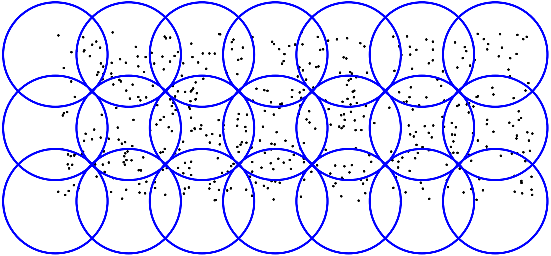
Theorem (Matoušek-Patáková 15)

Given a set of n-points $P \subset \mathbb{R}^m$. Then there exists a polynomial $f(X_1, \ldots, X_m)$ of degree at most $r^{1/m}$ such that

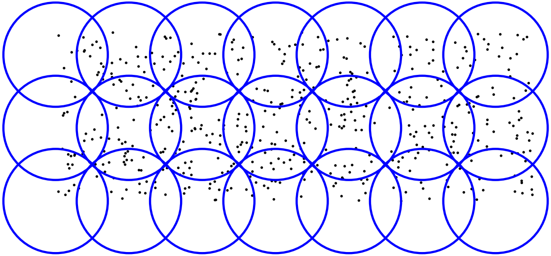
1. $\mathbb{R}^m \setminus Z(f)$ *has at most r maximally connected components, i.e,* $\mathbb{R}^m \setminus Z(f) = \omega_1 \sqcup \cdots \sqcup \omega_t$ *where* ω_i *are maximally connected components and* $t \leq r$.
2. $|\omega_i \cap P| \leq \frac{n}{r}$ *and* $|Z(f) \cap P| = 0$
3. *Any semialgebraic set* $\mathcal{O}$ *crosses at most* $r^{1-\frac{1}{m}}$ *connected components of* $\mathbb{R}^m \setminus Z(f)$.

(Def. of "Crossing") We say a set A crosses a set B if $A \cap B \notin \{\emptyset, B\}$.

Proof of Mnets bound

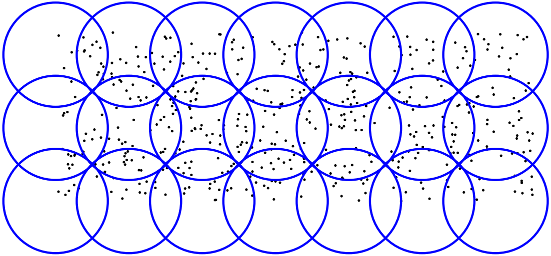


Proof of Mnets bound



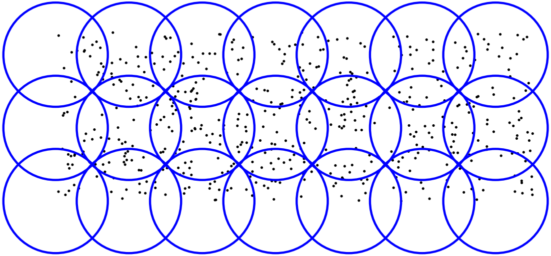
- Assume all the large sets are of size exactly εn

Proof of Mnets bound



- Assume all the large sets are of size exactly εn
- Build a maximal packing with $k = \varepsilon n$ and $\delta = \varepsilon n/2$. Number of sets in the packing is $\leq \frac{d}{\varepsilon}\varphi\left(\frac{d}{\varepsilon}, d\right)$ (via Shallow packing lemma)

Proof of Mnets bound



- Assume all the large sets are of size exactly εn
- Build a maximal packing with $k = \varepsilon n$ and $\delta = \varepsilon n/2$. Number of sets in the packing is $\leq \frac{d}{\varepsilon}\varphi\left(\frac{d}{\varepsilon}, d\right)$ (via Shallow packing lemma)
- Any set of size εn in the set system is either in the packing or has a large intersection with a set in the packing (size of intersection $\geq \varepsilon n/2$)

Proof: How to get a candidate Mnet set?

- Partition each set in the maximal packing with a r-partitioning polynomial (satisfying those magical properties), where r is a large constant to be fixed later.

Proof: How to get a candidate Mnet set?

- Partition each set in the maximal packing with a r-partitioning polynomial (satisfying those magical properties), where r is a large constant to be fixed later.
- Include in your Mnets set all the subsets $\omega_i \cap P$ where $|\omega_i \cap P| \geq \frac{\varepsilon n}{r^2}$.

Proof: How to get a candidate Mnet set?

- Partition each set in the maximal packing with a r-partitioning polynomial (satisfying those magical properties), where r is a large constant to be fixed later.
- Include in your Mnets set all the subsets $\omega_i \cap P$ where $|\omega_i \cap P| \geq \frac{\varepsilon n}{r^2}$.
- Do this for all the sets in the maximal packing and you will get an Mnet satisfying the bound in the theorem.

Proof: How to get a candidate Mnet set?

- Partition each set in the maximal packing with a r-partitioning polynomial (satisfying those magical properties), where r is a large constant to be fixed later.
- Include in your Mnets set all the subsets $\omega_i \cap P$ where $|\omega_i \cap P| \geq \frac{\varepsilon n}{r^2}$.
- Do this for all the sets in the maximal packing and you will get an Mnet satisfying the bound in the theorem.

Why is this a valid Mnet?

Proof: Why the construction is a valid Mnet?

- Let $S \in \Sigma$ with $|S| = \varepsilon n$.

Proof: Why the construction is a valid Mnet?

- Let $S \in \Sigma$ with $|S| = \varepsilon n$.
- Since the set system is a semialgebraic set system, there exists a semialg. object $\mathcal{O}$ such that $\mathcal{O} \cap P = S$.

Proof: Why the construction is a valid Mnet?

- Let $S \in \Sigma$ with $|S| = \varepsilon n$.
- Since the set system is a semialgebraic set system, there exists a semialg. object $\mathcal{O}$ such that $\mathcal{O} \cap P = S$.
- Since we have a maxi. packing, $\exists S_i$ from the packing such that $|S \cap S_i| \geq \varepsilon n/2$.

Proof: Why the construction is a valid Mnet?

- Let $S \in \Sigma$ with $|S| = \varepsilon n$.
- Since the set system is a semialgebraic set system, there exists a semialg. object $\mathcal{O}$ such that $\mathcal{O} \cap P = S$.
- Since we have a maxi. packing, $\exists S_i$ from the packing such that $|S \cap S_i| \geq \varepsilon n/2$.
- Maximum contribution to $|S \cap S_i|$ from connected regions crossed by $\mathcal{O}$ and the small set is at most

$$\frac{\varepsilon n}{r} \times r^{1-\frac{1}{m}} + \frac{\varepsilon n}{r^2} \times r = \varepsilon n \left(\frac{1}{r^{1/m}} + \frac{1}{r} \right) \ll \frac{\varepsilon n}{2}$$

From Mnets to ε-nets

Theorem (D.–G.–J.–M. '17)

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

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

From Mnets to ε-nets

Theorem (D.–G.–J.–M. '17)

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

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.

From Mnets to ε-nets

Theorem (D.–G.–J.–M. '17)

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

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.

Table: Upper bounds on Mnets and ε-nets

	Mnet	ε-net
Disks	ε^{-1}	ε^{-1}
Rectangles	$\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$	$\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}$
Halfspaces ($\mathbf{R}^d$)	$O\left(\varepsilon^{-\lfloor d/2 \rfloor}\right)$	$\frac{d}{\varepsilon} \log \frac{1}{\varepsilon}$

Probabilistic proof

Theorem (D.–G.–J.–M. '17)

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

Proof.

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

□

Probabilistic proof

Theorem (D.–G.–J.–M. '17)

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

Proof.

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.

$\square$

Probabilistic proof

Theorem (D.–G.–J.–M. '17)

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

Proof.

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.
3. $\forall m \in \mathcal{M}, \quad \Pr[S \cap m = \emptyset] = (1-p)^{|m|} \leq e^{-p|m|} \leq \frac{1}{\varepsilon M}$

$\square$

Probabilistic proof

Theorem (D.–G.–J.–M. '17)

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

Proof.

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.
3. $\forall m \in \mathcal{M}, \quad \Pr[S \cap m = \emptyset] = (1-p)^{|m|} \leq e^{-p|m|} \leq \frac{1}{\varepsilon M}$
4. In expectation, $|S| + |m \in \mathcal{M} : S \cap m = \emptyset| \leq np + \frac{1}{\varepsilon}$.

$\square$

Probabilistic proof

Theorem (D.–G.–J.–M. '17)

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

Proof.

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.
3. $\forall m \in \mathcal{M}, \quad \Pr[S \cap m = \emptyset] = (1-p)^{|m|} \leq e^{-p|m|} \leq \frac{1}{\varepsilon M}$
4. In expectation, $|S| + |m \in \mathcal{M} : S \cap m = \emptyset| \leq np + \frac{1}{\varepsilon}$.
5. so there is an ε-net of size $\leq np + \frac{1}{\varepsilon}$ (why?).

$\square$

Conclusion

- 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.