Hitting sets and VC-dimension

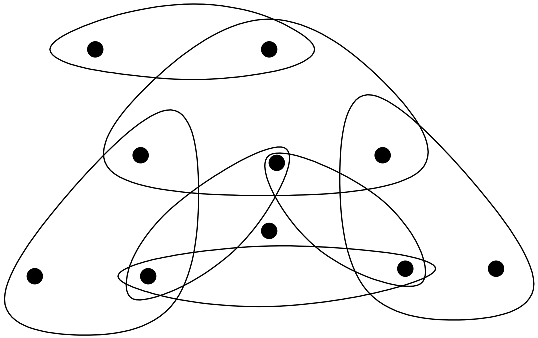
Nicolas Bousquet[1]

Joint work with Stéphan Thomassé[2]

[1]Université Montpellier II, LIRMM
[2]LIP, ENS Lyon

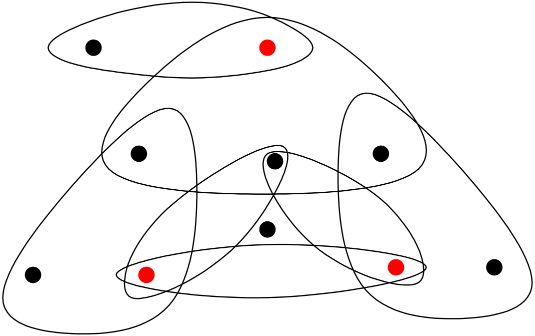
Definitions

A **hypergraph** is a pair (V, E) where V is a set of vertices and E is a set of hyperedges (subsets of vertices).



Definitions

A **hypergraph** is a pair (V, E) where V is a set of vertices and E is a set of hyperedges (subsets of vertices).
A **hitting set** is a subset of vertices intersecting all the hyperedges.
The **transversality** τ of a hypergraph is the minimum size of a hitting set.



Linear programming

Linear Programing

Variables: for each $v_i \in V$, associate x_i a non negative integer.
Constraints: for each $e \in E$,

$$\sum_{v_i \in e} x_i \geq 1$$

Objective function:

$$\tau = \min(\sum_{i=1}^{n} x_i)$$

Linear programming

Fractional Relaxation

Variables: for each $v_i \in V$, associate x_i a non negative real.

Constraints: for each $e \in E$,

$$\sum_{v_i \in e} x_i \geq 1$$

Objective function:

$$\tau^* = \min(\sum_{i=1}^{n} x_i)$$

Integrality gap between τ and τ^*

Inequality

$$\tau \geq \tau^*$$

Integrality gap between τ and τ^*

Inequality

$$\tau \geq \tau^*$$

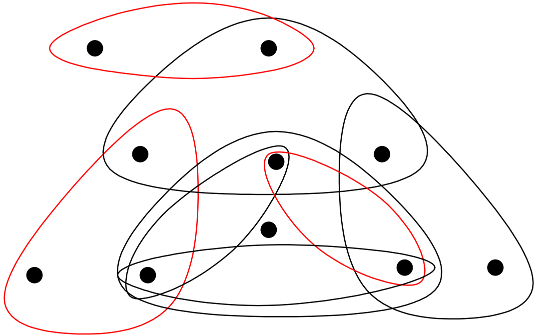
Integrality gap

$V = \{1, ..., 2n\}$
$e \in E$ iff $|e| = n$.

- $\tau^* = 2$: give the uniform weight $1/n$ to each vertex.
- $\tau = n + 1$, otherwise it remains one hyperedge in the complement of the hitting set.

Definition

The *packing number* ν *of a hypergraph* is the maximum number of disjoint hyperedges.



Definition

The *packing number* ν *of a hypergraph* is the maximum number of disjoint hyperedges.

Linear Programing

Variables: for each $e_i \in E$, associate x_i a non negative integer.
Constraints: for each $v \in V$,

$$\sum_{e_i / v \in e} x_i \leq 1$$

Objective function:

$$\nu = \max(\sum_{i=1}^{|E|} x_i)$$

Definition

The *packing number ν of a hypergraph* is the maximum number of disjoint hyperedges.

Fractional Relaxation

Variables: for each $e_i \in E$, associate x_i a non negative real.
Constraints: for each $v \in V$,

$$\sum_{e_i / v \in e} x_i \leq 1$$

Objective function:

$$\nu^* = \max(\sum_{i=1}^{|E|} x_i)$$

Integrality gap between ν and ν^*

Inequality

$$\nu \leq \nu^*$$

Integrality gap between ν and ν^*

Inequality

$$\nu \leq \nu^*$$

Integrality gap

The vertices of H are the edges of a clique on n vertices.
The hyperedges are the maximum stars of the clique.

- $\nu = 1$
- $\nu^* = n/2$

Erdős-Pósa property

Duality Theorem of Linear Programing

$$\tau^* = \nu^*$$

Erdős-Pósa property

Duality Theorem of Linear Programing

$$\tau^* = \nu^*$$

Inequalities

$$\nu \leq \nu^* = \tau^* \leq \tau$$

Erdős-Pósa property

Duality Theorem of Linear Programing

$$\tau^* = \nu^*$$

Inequalities

$$\nu \leq \nu^* = \tau^* \leq \tau$$

Erdős-Pósa property

A class $\mathcal{H}$ of hypergraphs has the Erdős-Pósa property iff there exists a function f such that for all $H \in \mathcal{H}$, $\tau \leq f(\nu)$.

Theorem (Erdős-Pósa)

The cycle hypergraph of a graph has the Erdős-Pósa property.

Our goal

$$\nu \leq \nu^* = \tau^* \leq \tau$$

- Under which conditions can we bound τ by a function of τ^*?

Our goal

$$\nu \leq \nu^* = \tau^* \leq \tau$$

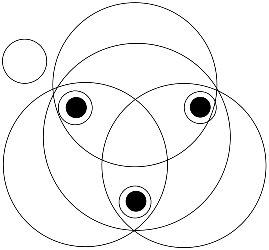
- Under which conditions can we bound τ by a function of τ^*?
- And τ by a function of ν?

VC-dimension

Definition

A set $X \subseteq V$ is **shattered** iff for all $Y \subseteq X$, there exist $e \in E$ such that $e \cap X = Y$.

The **VC-dimension** of a hypergraph is the maximum size of a shattered set.



Theorem

Theorem (Haussler, Welzl '73)

Every hypergraph H of VC-dimension d satisfies

$$\tau \leq 2d\tau^* log(11\tau^*).$$

Theorem

Theorem (Haussler, Welzl '73)

Every hypergraph H of VC-dimension d satisfies

$$\tau \leq 2d\tau^* log(11\tau^*).$$

- Randomized proof but some proofs can be derandomized.
- Constructive proof: provides an approximation algorithm.
- Based on the fact that a hypergraph of VC-dimension d has at most n^{d+1} hyperedges.

Applications

k-majority tournament

$V = \{1, ..., n\}$. Let $\prec_1, \ldots, \prec_{2k-1}$ be total orders on V.
The tournament realized by $\prec_1, \ldots, \prec_{2k-1}$ has an arc from i to j iff $i \succ j$ in at least k orders.
A tournament is a k-**majority tournament** if it can be realized by $2k - 1$ total orders.

Applications

k-majority tournament

$V = \{1, ..., n\}$. Let $\prec_1, \ldots, \prec_{2k-1}$ be total orders on V.
The tournament realized by $\prec_1, \ldots, \prec_{2k-1}$ has an arc from i to j iff $i \succ j$ in at least k orders.
A tournament is a k-**majority tournament** if it can be realized by $2k-1$ total orders.

Theorem (Alon, Brightwell, Kierstead, Kotochka, Winkler '04)

Each k-majority tournament has a dominating set of size $\mathcal{O}(k \cdot log(k))$.

Proof

- Consider the hypergraph H with hyperedges the in-neighborhoods of the vertices of T: a transversal of H is a dominating set of T.

Proof

- Consider the hypergraph H with hyperedges the in-neighborhoods of the vertices of T: a transversal of H is a dominating set of T.
- τ^* is bounded (by 2).

Proof

- Consider the hypergraph H with hyperedges the in-neighborhoods of the vertices of T: a transversal of H is a dominating set of T.
- τ^* is bounded (by 2).
- The VC-dimension is bounded (by $\mathcal{O}(k \cdot log(k))$).

Proof

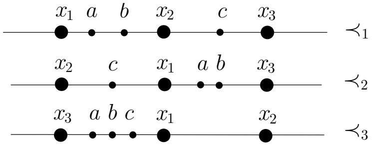
- Consider the hypergraph H with hyperedges the in-neighborhoods of the vertices of T: a transversal of H is a dominating set of T.
- τ^* is bounded (by 2).
- The VC-dimension is bounded (by $\mathcal{O}(k \cdot log(k))$).

Theorem (Haussler, Welzl '72)

For every hypergraph H of VC-dimension d:

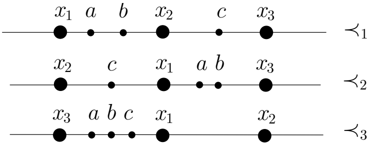
$$\tau \leq 2d\tau^* log(11\tau^*).$$

VC-dimension of the in-neighborhood hypergraph



$X = \{x_1, x_2, x_3\}$. Two vertices a, b are **non equivalent** if there are an order i and an element x_k of X such that $a \prec_i x_k \prec_i b$.

VC-dimension of the in-neighborhood hypergraph

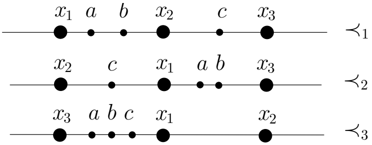


$X = \{x_1, x_2, x_3\}$. Two vertices a, b are **non equivalent** if there are an order i and an element x_k of X such that $a \prec_i x_k \prec_i b$.

Observation

Two equivalent vertices have the same neighborhood in X.

VC-dimension of the in-neighborhood hypergraph



$X = \{x_1, x_2, x_3\}$. Two vertices a, b are **non equivalent** if there are an order i and an element x_k of X such that $a \prec_i x_k \prec_i b$.

Observation

Two equivalent vertices have the same neighborhood in X.

- At most $(|X| + 1)^{2k-1}$ non equivalent vertices, so at most $(|X| + 1)^{2k-1}$ neighborhoods in X.
- And X is shattered if there are $2^{|X|}$ neighborhoods in X.

Conjecture

Definition

A set of k disjoint partial orders $\prec_1, ..., \prec_k$ *cover* a directed graph D if

$x_i x_j$ is an arc iff there is an order ℓ such that $x_i \prec_\ell x_j$.

Conjecture

Definition

A set of k disjoint partial orders $\prec_1, \ldots, \prec_k$ *cover* a directed graph D if

$x_i x_j$ is an arc iff there is an order ℓ such that $x_i \prec_\ell x_j$.

Conjecture

Every tournament which can be covered by at most k disjoint partial orders has a dominating set of size at most $f(k)$.

Conjecture

Definition

A set of k disjoint partial orders $\prec_1, ..., \prec_k$ *cover* a directed graph D if
$x_i x_j$ is an arc iff there is an order ℓ such that $x_i \prec_\ell x_j$.

Conjecture

Every tournament which can be covered by at most k disjoint partial orders has a dominating set of size at most $f(k)$.

- Related with a conjecture of Erdős Sands Sauer Woodraw.
- The same method does not immediately holds.

Hitting sets and packings

Integrality gap

VC-dimension

k-majority tournaments

Erdős-Pósa property

2VC-dimension and dual VC-dimension

Graph coloring and Scott's conjecture

Domination at distance ℓ

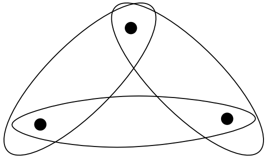
(p, q)-property

Conclusion

2VC-dimension

Definition

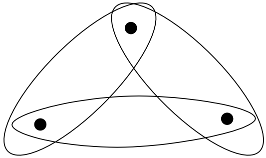
A set $X \subseteq V$ is 2-*shattered* iff for all $Y \subseteq X$ with $|Y| = 2$, there exist $e \in E$ such that $e \cap X = Y$.
The 2*VC-dimension* of a hypergraph is the maximum size of a 2-shattered set.



2VC-dimension

Definition

A set $X \subseteq V$ is 2-*shattered* iff for all $Y \subseteq X$ with $|Y| = 2$, there exist $e \in E$ such that $e \cap X = Y$.
The 2*VC-dimension* of a hypergraph is the maximum size of a 2-shattered set.

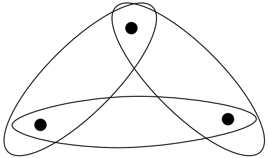


- VC $\leq$ 2VC.

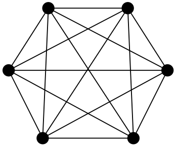
2VC-dimension

Definition

A set $X \subseteq V$ is 2-*shattered* iff for all $Y \subseteq X$ with $|Y| = 2$, there exist $e \in E$ such that $e \cap X = Y$.
The 2*VC-dimension* of a hypergraph is the maximum size of a 2-shattered set.



- VC $\leq$ 2VC.
- The reverse inequality does not holds.



Dual hypergraph

Bipartite incidence graph

A hypergraph $H = (V, E)$ can be seen as a *bipartite incidence graph* G with vertex set (V, E) where (v, e) in an edge iff $v \in e$.

Dual hypergraph

Bipartite incidence graph

A hypergraph $H = (V, E)$ can be seen as a *bipartite incidence graph* G with vertex set (V, E) where (v, e) in an edge iff $v \in e$.

Dual hypergraph

The pair (V, E) is oriented: the hypergraph associated to the pair (E, V) is the **dual hypergraph** denoted by H^d.

Dual hypergraph

Bipartite incidence graph

A hypergraph $H = (V, E)$ can be seen as a *bipartite incidence graph* G with vertex set (V, E) where (v, e) in an edge iff $v \in e$.

Dual hypergraph

The pair (V, E) is oriented: the hypergraph associated to the pair (E, V) is the **dual hypergraph** denoted by H^d.

Definition

The **dual VC-dimension** is the VC-dimension of the dual hypergraph.

Duality gap

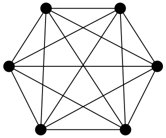
VC-dimension and dual VC-dimension

- VC-dimension and dual VC-dimension are linked by an exponential function.

Duality gap

VC-dimension and dual VC-dimension

- VC-dimension and dual VC-dimension are linked by an exponential function.
- An arbitrarily large gap is possible between the 2VC-dimension and the dual 2-VC-dimension.



Theorem

Theorem (Ding, Seymour, Winkler '91)

Let H be a hypergraph of dual 2VC-dimension d then:

$$\tau \leq 11d^2(\nu + d + 3) \cdot \binom{d+\nu}{d}.$$

Theorem

Theorem (Ding, Seymour, Winkler '91)

Let H be a hypergraph of dual 2VC-dimension d then:

$$\tau \leq 11d^2(\nu + d + 3) \cdot \binom{d + \nu}{d}.$$

Hint of the proof:

- τ bounded by a function of τ^*: the VC-dimension.
- ν^* bounded by a function of ν: a quite magical argument using duality of hypergraphs.

Scott's conjecture

Conjecture (Scott '97)

Let H be a fixed graph. Every triangle free graph with no induced copy of H has a bounded chromatic number.

- Several cases are already known. (paths, trees, stars, graphs on at most 4 vertices...).

Scott's conjecture

Conjecture (Scott '97)

Let H be a fixed graph. Every triangle free graph with no induced copy of H has a bounded chromatic number.

- Several cases are already known. (paths, trees, stars, graphs on at most 4 vertices...).
- Contains a conjecture of Erdős... which was disproved last year by Pawlik et al.

Scott's conjecture

Conjecture (Scott '97)

Let H be a fixed graph. Every triangle free graph with no induced copy of H has a bounded chromatic number.

- Several cases are already known. (paths, trees, stars, graphs on at most 4 vertices...).
- Contains a conjecture of Erdős... which was disproved last year by Pawlik et al.

A **maximal triangle-free graph** is a graph such that the addition of any edge creates a triangle.

Theorem (B., Thomassé '12)

Every maximal triangle free graph with no induced subdivision of H has a bounded chromatic number.

Neighborhood hypergraphs

Neighborhood hypergraph of $G = (V, E)$

Vertex set: V.
Hyperedge set: $N(v)$ for every $v \in V$.

Neighborhood hypergraphs

Neighborhood hypergraph of $G = (V, E)$

Vertex set: V.
Hyperedge set: $N(v)$ for every $v \in V$.

Observation

The dual of a neighborhood hypergraph H is the hypergraph H:
$x \in B(y, l)$ iff $y \in B(x, l)$.

Sketch of the proof

Theorem (B., Thomassé '12)

Every maximal triangle free graph with no induced subdivision of H has a bounded chromatic number.

- The neighborhood hypergraph satisfies $\nu = 1$: otherwise one an edge can be added without creating any triangle.

Sketch of the proof

Theorem (B., Thomassé '12)

Every maximal triangle free graph with no induced subdivision of H has a bounded chromatic number.

- The neighborhood hypergraph satisfies $\nu = 1$: otherwise one an edge can be added without creating any triangle.
- If $\tau \leq k$, then $\chi(G) \leq 2k$: every neighborhood is a stable set.

Sketch of the proof

Theorem (B., Thomassé '12)

Every maximal triangle free graph with no induced subdivision of H has a bounded chromatic number.

- The neighborhood hypergraph satisfies $\nu = 1$: otherwise one an edge can be added without creating any triangle.
- If $\tau \leq k$, then $\chi(G) \leq 2k$: every neighborhood is a stable set.

Theorem (Ding, Seymour, Winkler '91)

Every hypergraph of dual 2VC-dimension d satisfies:

$$\tau \leq 11d^2(\nu + d + 3) \cdot \binom{d+\nu}{d}.$$

Covering of a planar graph

Theorem (Chepoi, Estellon, Vaxès '07)

There exists a constant m such that, for every planar graph of diameter 2ℓ, there are m balls of radius ℓ which cover the graph.

Covering of a planar graph

Theorem (Chepoi, Estellon, Vaxès '07)

There exists a constant *m* such that, for every planar graph of diameter 2ℓ, there are *m* balls of radius ℓ which cover the graph.

This result can be generalized as follows:

- ν_ℓ: number of disjoint balls of radius ℓ.
- τ_ℓ: size of a dominating set at distance ℓ.

Covering of a planar graph

Theorem (Chepoi, Estellon, Vaxès '07)

There exists a constant m such that, for every planar graph of diameter 2ℓ, there are m balls of radius ℓ which cover the graph.

This result can be generalized as follows:

- ν_ℓ: number of disjoint balls of radius ℓ.
- τ_ℓ: size of a dominating set at distance ℓ.

Theorem

There exists a polynomial function f such that, for every planar graph and every

$$\tau_\ell \leq f(\nu_\ell).$$

Iterated neighborhood hypergraph

Definition

We denote by $B(x, \ell)$ the set of the vertices at distance at most l from x in the graph.
The **iterated neighborhoods hypergraphs** of G is the hypergraph on V with hyperedges $B(x, \ell)$ for all x and ℓ.

Iterated neighborhood hypergraph

Definition

We denote by $B(x, \ell)$ the set of the vertices at distance at most l from x in the graph.
The **iterated neighborhoods hypergraphs** of G is the hypergraph on V with hyperedges $B(x, \ell)$ for all x and ℓ.

Definition

The VC-dimension of a graph is equal to the VC-dimension of the hypergraph of iterated neighborhoods.

Iterated neighborhood hypergraph

Definition

We denote by $B(x, \ell)$ the set of the vertices at distance at most l from x in the graph.
The **iterated neighborhoods hypergraphs** of G is the hypergraph on V with hyperedges $B(x, \ell)$ for all x and ℓ.

Definition

The VC-dimension of a graph is equal to the maximum, over all induced subgraph, of the VC-dimension of the iterated neighborhoods hypergraph.

Iterated neighborhood hypergraph

Definition

We denote by $B(x, \ell)$ the set of the vertices at distance at most l from x in the graph.
The **iterated neighborhoods hypergraphs** of G is the hypergraph on V with hyperedges $B(x, \ell)$ for all x and ℓ.

Definition

The VC-dimension of a graph is equal to the maximum, over all induced subgraph, of the VC-dimension of the iterated neighborhoods hypergraph.

Theorem (Ding, Seymour, Winkler '91)

For every ℓ and every graph of 2VC-dimension d, we have:

$$\tau\ell \leq 11d^2(\nu_\ell + d + 3) \cdot \binom{d + \nu_\ell}{d}.$$

VC-dimensionof planar graphs

Theorem

Planar graphs have 2VC-dimension at most 4.

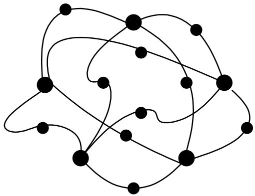
VC-dimensionof planar graphs

Theorem

Planar graphs have 2VC-dimension at most 4.

Proof:

- Assume that the dual VC-dimension is at least 5.



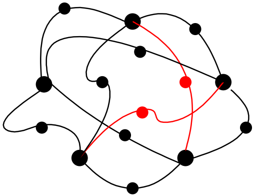
VC-dimensionof planar graphs

Theorem

Planar graphs have 2VC-dimension at most 4.

Proof:

- Assume that the dual VC-dimension is at least 5.
- Two paths must intersect (otherwise the planar graph must contain a K_5-minor).



A proof of the result of Chepoi et al.

Theorem (Chepoi, Estellon, Vaxès)

Every planar graph of diameter 2ℓ admits a dominating set at distance ℓ of size at most 880000.

A proof of the result of Chepoi et al.

Theorem (Chepoi, Estellon, Vaxès)

Every planar graph of diameter 2ℓ admits a dominating set at distance ℓ of size at most 880000.

- The dual 2VC-dimension is bounded by 4.
- The packing number equals to 1.

A proof of the result of Chepoi et al.

Theorem (Chepoi, Estellon, Vaxès)

Every planar graph of diameter 2ℓ admits a dominating set at distance ℓ of size at most 880000.

- The dual 2VC-dimension is bounded by 4.
- The packing number equals to 1.

Ding, Seymour, Winkler

Let H be a hypergraph of dual 2VC-dimension d then:

$$\tau \leq 11d^2(\nu + d + 3) \cdot \binom{d + \nu}{d}$$

A proof of the result of Chepoi et al.

Theorem (Chepoi, Estellon, Vaxès)

Every planar graph of diameter 2ℓ admits a dominating set at distance ℓ of size at most 880000.

- The dual 2VC-dimension is bounded by 4.
- The packing number equals to 1.

Ding, Seymour, Winkler

Let H be a hypergraph of dual 2VC-dimension d then:

$$\tau \leq 11d^2(\nu + d + 3) \cdot \binom{d+\nu}{d}$$

Known lower bound: ≈ 10

Conjecture

Every planar graph satisfies $\tau \leq \mathcal{O}(\nu^5)$.

Conjecture

Every planar graph satisfies $\tau \leq \mathcal{O}(\nu^5)$.

Conjecture (Chepoi, Estellon, Vaxès)

There exists a constant c such that the hypergraph of the balls of radius ℓ of a planar graph satisfies:

$$\tau_\ell \leq c \cdot \nu_\ell.$$

Conjecture

Every planar graph satisfies $\tau \leq \mathcal{O}(\nu^5)$.

Conjecture (Chepoi, Estellon, Vaxès)

There exists a constant c such that the hypergraph of the balls of radius ℓ of a planar graph satisfies:

$$\tau_\ell \leq c \cdot \nu_\ell.$$

Note that in both cases, the constant does not depend on ℓ.

Conjecture

Every planar graph satisfies $\tau \leq \mathcal{O}(\nu^5)$.

Conjecture (Chepoi, Estellon, Vaxès)

There exists a constant c such that the hypergraph of the balls of radius ℓ of a planar graph satisfies:

$$\tau_\ell \leq c \cdot \nu_\ell.$$

Note that in both cases, the constant does not depend on ℓ.

Theorem (Dvorak)

For every planar graph of diameter R and every integer ℓ, there exists a constant $\rho_{R,\ell}$ such that:

$$\tau_\ell \leq \rho_{R,\ell} \cdot \nu_\ell.$$

Graphs of bounded VC-dimension

Theorem (B.,Thomassé)

- Planar graphs have VC-dimension at most 4.
- K_n-minor free graphs have VC-dimension at most $n-1$.

Graphs of bounded VC-dimension

Theorem (B.,Thomassé)

- Planar graphs have VC-dimension at most 4.
- K_n-minor free graphs have VC-dimension at most $n - 1$.
- Graphs of rankwidth k have VC-dimension at most $3 \cdot 2^{2k+1} + 2$.

Graphs of bounded VC-dimension

Theorem (B., Thomassé)

- Planar graphs have VC-dimension at most 4.
- K_n-minor free graphs have VC-dimension at most $n-1$.
- Graphs of rankwidth k have VC-dimension at most $3 \cdot 2^{2k+1} + 2$.

Consequence

For all these classes, $\tau_\ell \leq f(\nu_\ell)$.

A weaker condition for the Edős-Pósa property

Definition

A hypergraph satisfies the (p, q)-property if for every set of p hyperedges, at least q of them have a non-empty intersection.

A weaker condition for the Edős-Pósa property

Definition

A hypergraph satisfies the (p, q)-property if for every set of p hyperedges, at least q of them have a non-empty intersection.

Theorem (Matousek)

Let H be a hypergraph of dual VC-dimension d. There exists a function f such that if H has the $(p, d+1)$ property then $\tau \leq f(p, d)$.

Domination at large distance

Theorem (B., Thomassé)

There exists a function f such that if G has VC-dimension d then, the hypergraph of the balls of radius l satisfies:

$$\tau \leq f(\nu, d).$$

Domination at large distance

Theorem (B., Thomassé)

There exists a function f such that if G has VC-dimension d then, the hypergraph of the balls of radius l satisfies:

$$\tau \leq f(\nu, d).$$

- The VC-dimension is bounded by definition.
- We have to verify that the $(p, d+1)$ property is verified for p which depends only of ν and d.

Conclusion

Open problems

- Triangle-free circle graphs have bounded VC-dimension.

Conclusion

Open problems

- Triangle-free circle graphs have bounded VC-dimension.
- A class of graphs of bounded VC-dimension is χ-bounded.

Conclusion

Open problems

- Triangle-free circle graphs have bounded VC-dimension.
- A class of graphs of bounded VC-dimension is χ-bounded.
- A class of graphs of bounded 2VC-dimension is χ-bounded.

Conclusion

Open problems

- Triangle-free circle graphs have bounded VC-dimension.
- A class of graphs of bounded VC-dimension is χ-bounded.
- A class of graphs of bounded 2VC-dimension is χ-bounded.

Merci