Identifying codes and VC-dimension

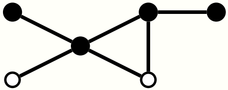
Aurélie Lagoutte
LIP, ENS Lyon

Joint work with:
N. Bousquet, Z. Li, A. Parreau and S. Tomassé

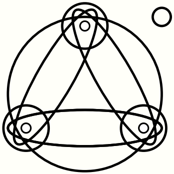
BGW - November 19, 2014

Contents

Identifying codes



VC-dimension



?

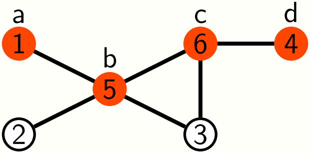
Part I

Identifying codes

Modelization with a graph

Identifying code C = subset of vertices which is

- dominating : $\forall u \in V, N[u] \cap C \neq \emptyset$,
- separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.



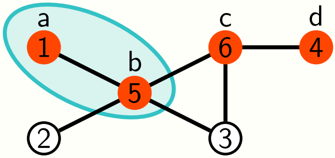
V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the size $\gamma^{ID}(G)$ of minimum identifying code ?

Modelization with a graph

Identifying code C = subset of vertices which is

- dominating : $\forall u \in V, N[u] \cap C \neq \emptyset$,
- separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.



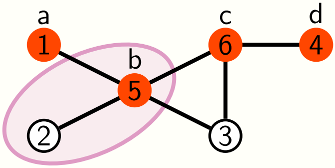
V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the size $\gamma^{ID}(G)$ of minimum identifying code ?

Modelization with a graph

Identifying code C = subset of vertices which is

- dominating : $\forall u \in V, N[u] \cap C \neq \emptyset$,
- separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.



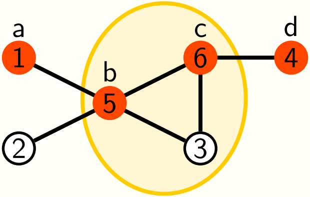
V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the size $\gamma^{ID}(G)$ of minimum identifying code ?

Modelization with a graph

Identifying code C = subset of vertices which is

- dominating : $\forall u \in V, N[u] \cap C \neq \emptyset$,
- separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.

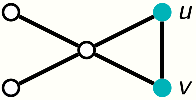


V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the size $\gamma^{ID}(G)$ of minimum identifying code ?

Some facts about identifying codes

- Introduced in 1998 by Karpvosky, Chakrabarty and Levitin
- Exists iff there is no twins



Twins: $N[u] = N[v]$

Some facts about identifying codes

- Introduced in 1998 by Karpvosky, Chakrabarty and Levitin
- Exists iff there is no twins
- NP-complete (Charon, Hudry, Lobstein, 2001)

Some facts about identifying codes

- Introduced in 1998 by Karpvosky, Chakrabarty and Levitin
- Exists iff there is no twins
- NP-complete (Charon, Hudry, Lobstein, 2001)
- Hard to approximate: best approximation factor $\log(|V|)$

Some facts about identifying codes

- Introduced in 1998 by Karpvosky, Chakrabarty and Levitin
- Exists iff there is no twins
- NP-complete (Charon, Hudry, Lobstein, 2001)
- Hard to approximate: best approximation factor $\log(|V|)$
- Lower bound:

  $\rightarrow$ A vertex is identified by a nonempty subset of $C \Rightarrow |V| \leq 2^{\gamma^{ID}(G)} - 1$

$$\gamma^{ID}(G) \geq \log(|V| + 1)$$

Tight example:

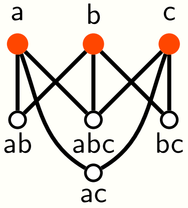
Some facts about identifying codes

- Introduced in 1998 by Karpvosky, Chakrabarty and Levitin
- Exists iff there is no twins
- NP-complete (Charon, Hudry, Lobstein, 2001)
- Hard to approximate: best approximation factor $\log(|V|)$
- Lower bound:

  $\rightarrow$ A vertex is identified by a nonempty subset of $C \Rightarrow |V| \leq 2^{\gamma^{ID}(G)} - 1$

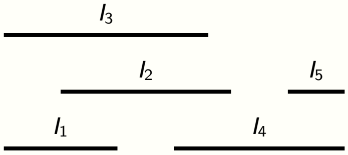
$$\gamma^{ID}(G) \geq \log(|V| + 1)$$

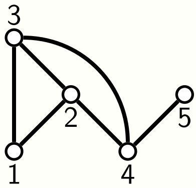
Tight example:



In restricted classes of graphs?

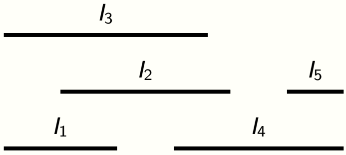
Example 1 : Class of interval graphs

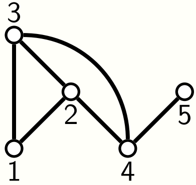




In restricted classes of graphs?

Example 1 : Class of interval graphs



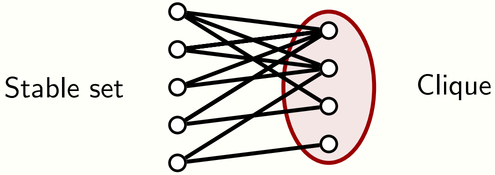


Proposition Foucaud, Naserasr, Parreau, Valicov, 2012+

If G is an interval graph, $\gamma^{ID}(G) \geq \sqrt{2|V|}$.

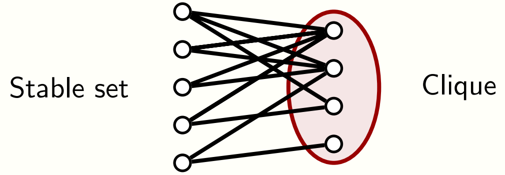
In restricted classes of graphs?

Example 2: Class of split graphs



In restricted classes of graphs?

Example 2: Class of split graphs



Proposition

For infinitely many split graphs G, $\gamma^{ID}(G) = \log(|V| + 1)$.

Proposition Foucaud, 2013

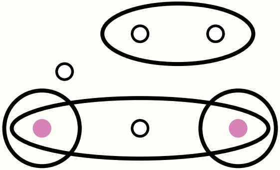
MIN-ID-CODE is log-APX-hard for split graphs.

Part II

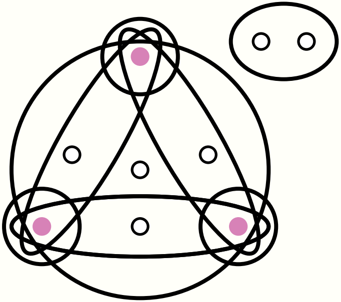
VC-dimension

Shattered set

- $\mathcal{H} = (V, \mathcal{E})$ an hypergraph
- A set $X \subseteq V$ is shattered if for all $Y \subseteq X$, there exists $e \in \mathcal{E}$, s.t $e \cap X = Y$.



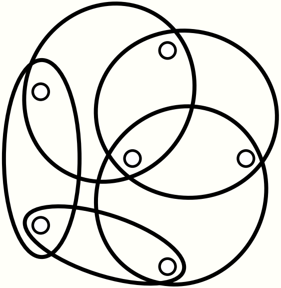
A 2-shattered set



A 3-shattered set

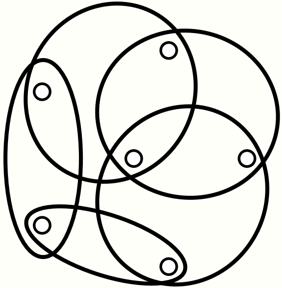
Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.



Vapnik Chervonenkis (VC) dimension of an hypergraph

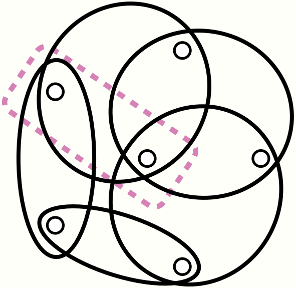
- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.



No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.

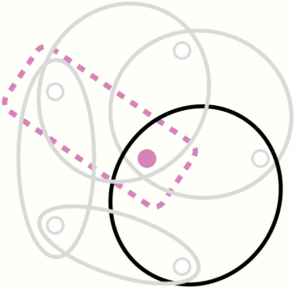


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

A 2-shattered set $\Rightarrow$ VC-dim $= 2$

Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.

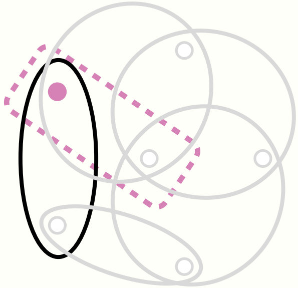


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

A 2-shattered set $\Rightarrow$ VC-dim $= 2$

Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.

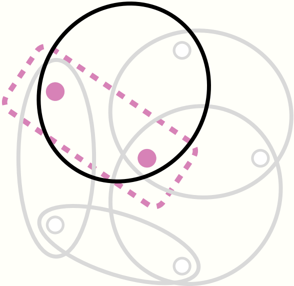


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

A 2-shattered set $\Rightarrow$ VC-dim $= 2$

Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.

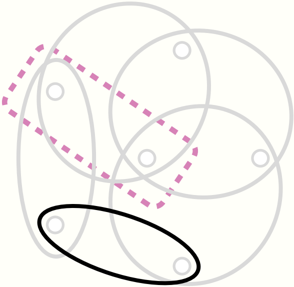


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

A 2-shattered set $\Rightarrow$ VC-dim $= 2$

Vapnik Chervonenkis (VC) dimension of an hypergraph

- A set X is shattered if $\forall Y \subseteq X$, $\exists e \in \mathcal{E}$, s.t $e \cap X = Y$.
- VC-dimension of $\mathcal{H}$: largest size of a shattered set.

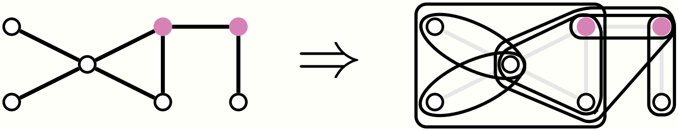


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

A 2-shattered set $\Rightarrow$ VC-dim $= 2$

VC dimension of a graph / of a class of graph

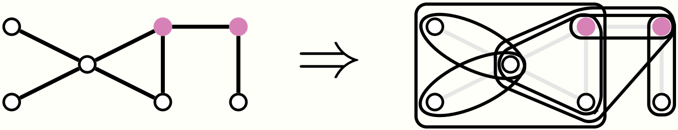
- VC-dimension of G: VC-dim of the hypergraph of closed neighborhoods



$$\text{VC-dim}(G) = 2$$

VC dimension of a graph / of a class of graph

- VC-dimension of G: VC-dim of the hypergraph of closed neighborhoods

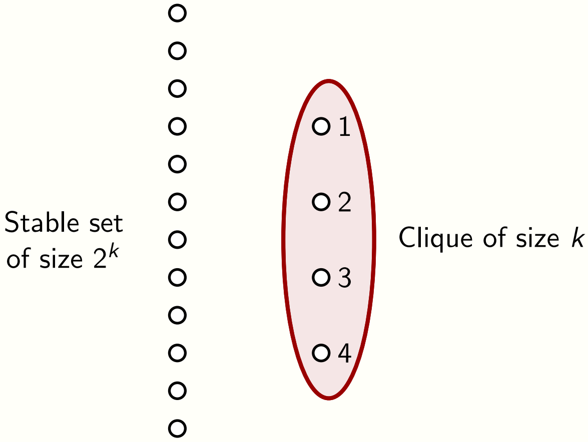


$$\text{VC-dim}(G) = 2$$

- VC-dimension of a class $\mathcal{C}$: maximal VC-dimension over $\mathcal{C}$
    - Class of interval graphs has VC-dimension 2.
    - Class of split graphs has infinite VC-dimension.

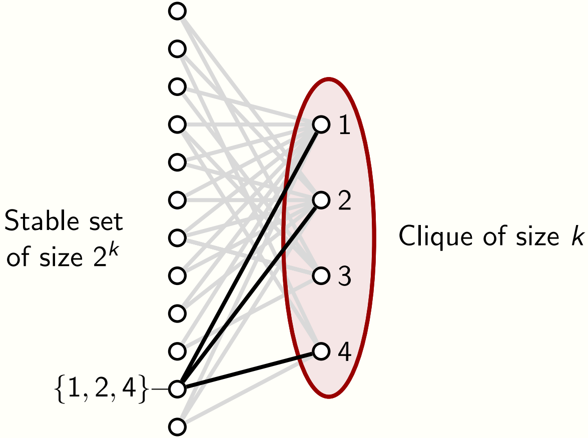
Split graphs have infinite VC-dimension

For any k, there is a split graph with VC-dimension k.



Split graphs have infinite VC-dimension

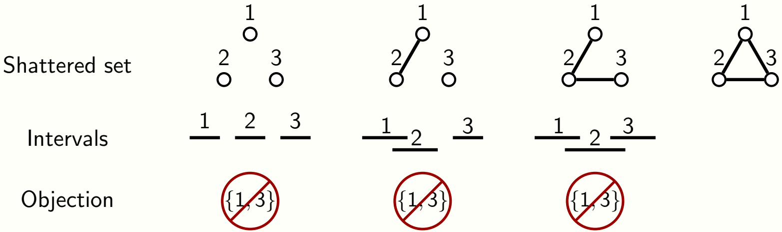
For any k, there is a split graph with VC-dimension k.



Intervals have finite VC-dimension

There is no interval graph with VC-dimension 3.

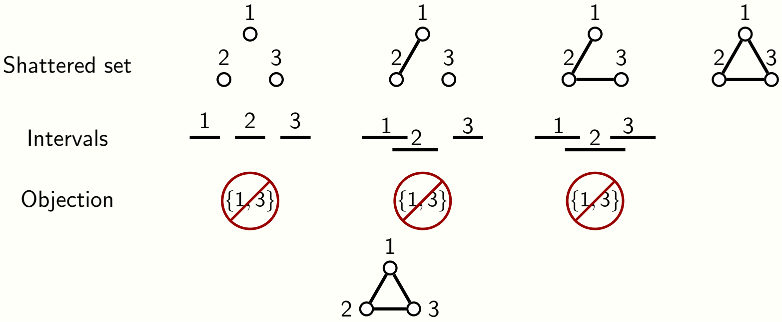
Assume there is a shattered set $\{1, 2, 3\}$.



Intervals have finite VC-dimension

There is no interval graph with VC-dimension 3.

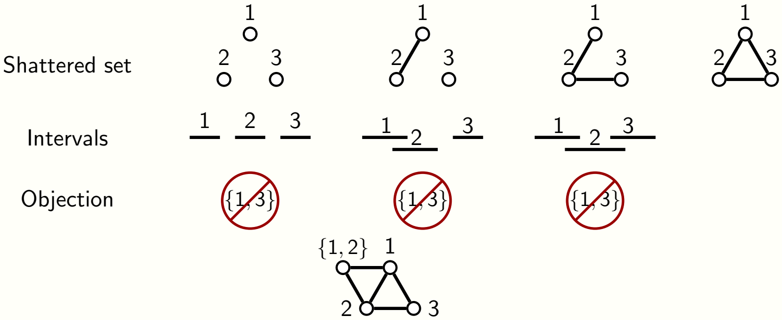
Assume there is a shattered set $\{1, 2, 3\}$.



Intervals have finite VC-dimension

There is no interval graph with VC-dimension 3.

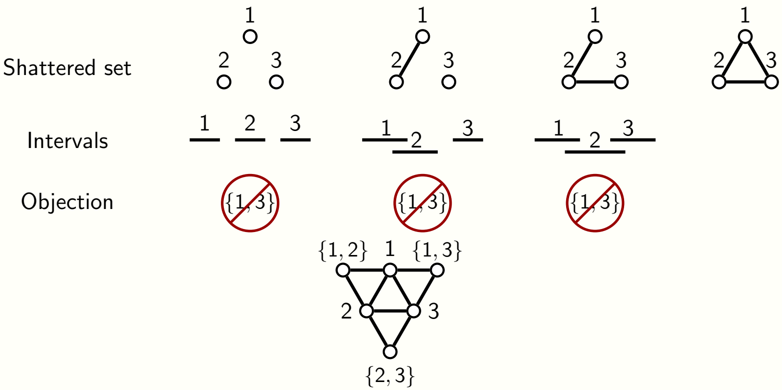
Assume there is a shattered set $\{1, 2, 3\}$.



Intervals have finite VC-dimension

There is no interval graph with VC-dimension 3.

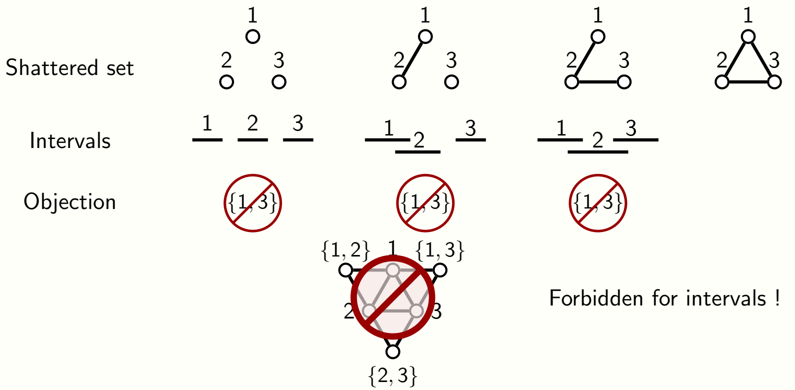
Assume there is a shattered set $\{1, 2, 3\}$.



Intervals have finite VC-dimension

There is no interval graph with VC-dimension 3.

Assume there is a shattered set $\{1,2,3\}$.



Interval graphs have VC-dimension at most 2.

Part III

Identifying codes and VC-dimension

Back to identifying codes

Graph class	Lower bound (order)	Approx
All	$\log n$	log APX-h
Split	$\log n$	log APX-h
Interval	$n^{1/2}$	open
Unit Interval	n	2
Bipartite	$\log n$	log APX-h
Line graphs	$n^{1/2}$	4
Chordal	$\log n$	log APX-h
Planar	n	7
Cograph	n	1

Back to identifying codes

Graph class	Lower bound (order)	Approx	VC dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Bipartite	$\log n$	log APX-h	∞
Line graphs	$n^{1/2}$	4	4
Chordal	$\log n$	log APX-h	∞
Planar	n	7	4
Cograph	n	1	2

Back to identifying codes

Graph class	Lower bound (order)	Approx	VC dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Bipartite	$\log n$	log APX-h	∞
Line graphs	$n^{1/2}$	4	4
Chordal	$\log n$	log APX-h	∞
Planar	n	7	4
Cograph	n	1	2

Back to identifying codes

Graph class	Lower bound (order)	Approx	VC dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2

A dichotomy result

Theorem

VC-dim of $\mathcal{C}$?

Infinite

$\Downarrow$

There are infinitely many G,

$$\gamma^{ID}(G) \approx \log |V|$$

Finite d

$\Downarrow$

For all G,

$$\gamma^{ID}(G) \geq (|V| - 1)^{1/d}$$

Proof - Case with finite VC dimension

Proposition

If $\mathcal{C}$ has finite VC-dimension d, $\forall G \in \mathcal{C}$, $\gamma^{ID}(G) \geq (|V| - 1)^{1/d}$.

Proof - Case with finite VC dimension

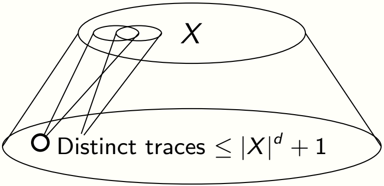
Proposition

If $\mathcal{C}$ has finite VC-dimension d, $\forall G \in \mathcal{C}$, $\gamma^{ID}(G) \geq (|V|-1)^{1/d}$.

Proof: direct consequence of:

Sauer's Lemma

Let X be a subset of vertices of graph G of VC-dimension d. The number of distinct traces on X is at most $\sum_{i=1}^{d} \binom{|X|}{i} \leq |X|^d + 1$.



Proof - Case with finite VC dimension

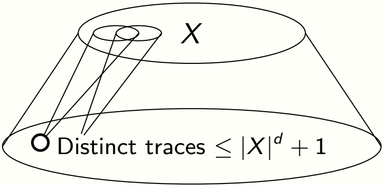
Proposition

If $\mathcal{C}$ has finite VC-dimension d, $\forall G \in \mathcal{C}$, $\gamma^{ID}(G) \geq (|V|-1)^{1/d}$.

Proof: direct consequence of:

Sauer's Lemma

Let X be a subset of vertices of graph G of VC-dimension d. The number of distinct traces on X is at most $\sum_{i=1}^{d} \binom{|X|}{i} \leq |X|^d + 1$.



Identifying code $\quad \gamma^{ID}(G)$

All vertices $|V| \quad \leq \gamma^{ID}(G)^d + 1$

Back to the table

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

- Lower bound not optimal (ex: Line graphs)

Back to the table

Graph class	Lower bound ✓	Approx ?	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

- Lower bound not optimal (ex: Line graphs)
- What about approximation ?

Inapproximability in infinite VC dimension

Theorem

If $\mathcal{C}$ has ∞ VC-dimension, $\textsc{Min-Id-Code}$ is log-APX-hard on $\mathcal{C}$.

Inapproximability in infinite VC dimension

Theorem

If $\mathcal{C}$ has ∞ VC-dimension, $\textsc{Min-Id-Code}$ is log-APX-hard on $\mathcal{C}$.

Consequence of:

Proposition

If $\mathcal{C}$ has infinite VC-dimension, $\mathcal{C}$ contains:

- all bipartite graphs, or
- all split graphs, or
- all cobipartite graphs.

and

Theorem Foucaud, 2013

$\textsc{Min-Id-Code}$ is log-APX-hard on bipartite, split and cobipartite graphs.

In the finite case ?

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

Is there a constant approximation in finite VC-dimension?

A class of finite VC-dimension with no good approximation

Theorem

MIN-ID-CODE cannot be approximated within a $o(\log |V|)$ factor in polynomial time for the class of bipartite C_4-free graphs.

- Class of VC-dimension 2
- Reduction from SET COVERING WITH INTERSECTION 1.

What about open approximation ?

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

What about open approximation ?

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	open	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

What about open approximation ?

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	6	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

What about open approximation ?

Graph class	Lower bound	Approx	VC-dim
All	$\log n$	log APX-h	∞
Split	$\log n$	log APX-h	∞
Bipartite	$\log n$	log APX-h	∞
Chordal	$\log n$	log APX-h	∞
Interval	$n^{1/2}$	6	2
Unit Interval	n	2	2
Line graphs	$n^{1/2}$	4	4
Planar	n	7	4
Cograph	n	1	2
Permutation	$n^{1/3}$	open	3
Unit disk graphs	$n^{1/3}$	open	3

Thank you for your attention !