On the Approximability of Partial VC Dimension

Cristina Bazgan[1] Florent Foucaud[2] Florian Sikora[1]

[1]LAMSADE, Université Paris Dauphine, CNRS – France
[2]LIMOS, Université Blaise Pascal, Clermont-Ferrand – France

COCOA 2016

On the Approximability of Partial VC Dimension

Cristina Bazgan[1] Florent Foucaud[2] Florian Sikora[1]

[1]LAMSADE, Université Paris Dauphine, CNRS – France
[2]LIMOS, Université Blaise Pascal, Clermont-Ferrand – France

COCOA 2016

On the Approximability of Partial VC Dimension

Cristina Bazgan[1] Florent Foucaud[2] Florian Sikora[1]

[1]LAMSADE, Université Paris Dauphine, CNRS – France
[2]LIMOS, Université Blaise Pascal, Clermont-Ferrand – France

COCOA 2016

Outline

Introduction

Distinguishing Transversal

VC-dimension

Studied Problem: Partial VC-dimension

Positive results

Negative results

Outline

Fire detection in a museum?

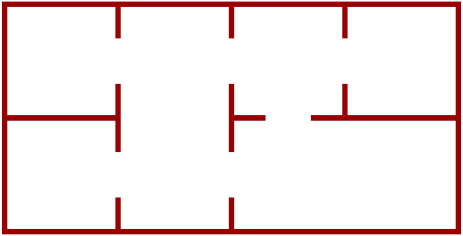
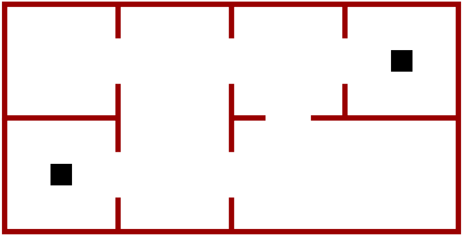


Figure and example: A. Parreau

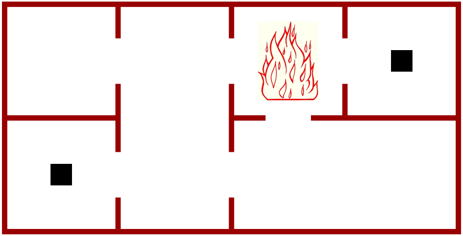
Fire detection in a museum?



- Detector can detect fire in their room or in their neighborhood.

Figure and example: A. Parreau

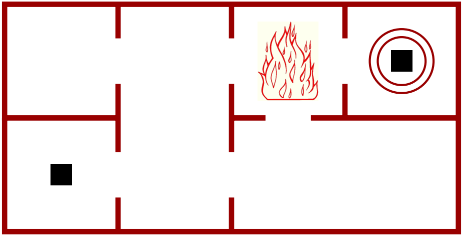
Fire detection in a museum?



- Detector can detect fire in their room or in their neighborhood.

Figure and example: A. Parreau

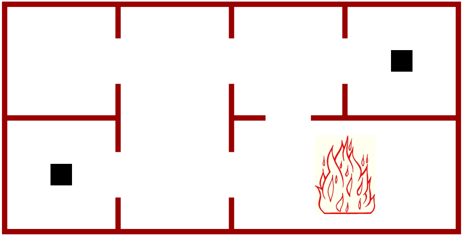
Fire detection in a museum?



- Detector can detect fire in their room or in their neighborhood.

Figure and example: A. Parreau

Fire detection in a museum?



- Detector can detect fire in their room or in their neighborhood.

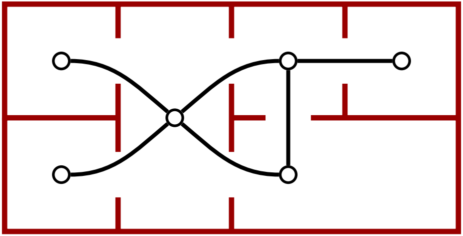
Figure and example: A. Parreau

Fire detection in a museum?

- Detector can detect fire in their room or in their neighborhood.
- Each room must contain a detector or have a detector in a neighboring room.

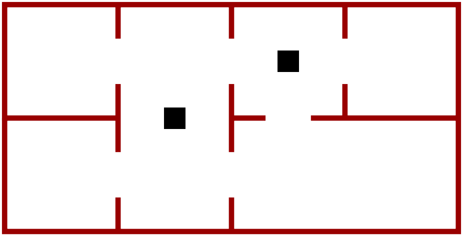
Figure and example: A. Parreau

Modelization with a graph

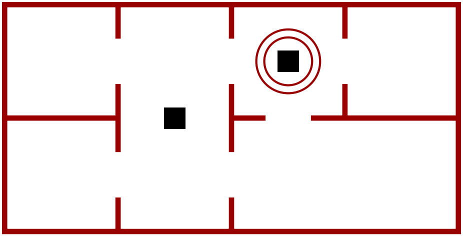


- Vertices V: rooms
- Edges E: between two neighboring rooms

Back to the museum

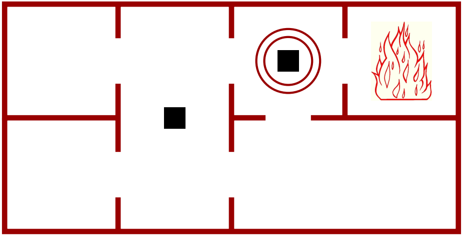


Back to the museum



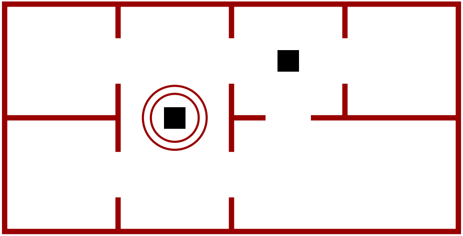
Where is the fire ?

Back to the museum



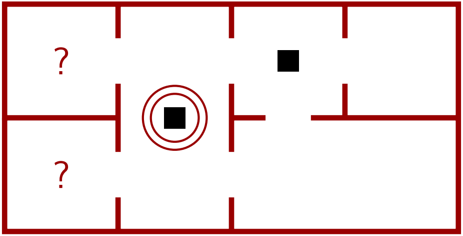
Where is the fire ?

Back to the museum



Where is the fire ?

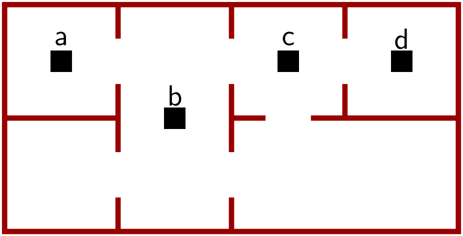
Back to the museum



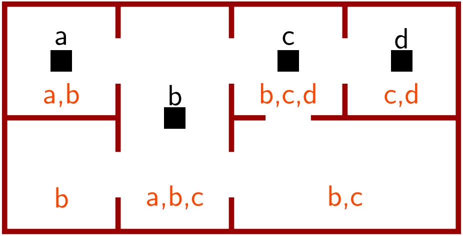
Where is the fire ?

To locate the fire, we need more detectors.

Locate the fire



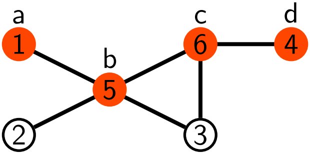
Locate the fire



In each room, the set of detectors in the neighborhood is unique.

Modelization with a graph

Distinguishing Transversal C = subset of vertices which is separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.

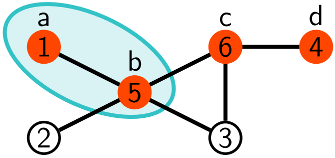


V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the minimum size of C?

Modelization with a graph

Distinguishing Transversal C = subset of vertices which is separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.

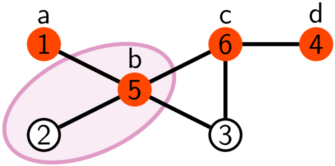


V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the minimum size of C?

Modelization with a graph

Distinguishing Transversal C = subset of vertices which is separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.

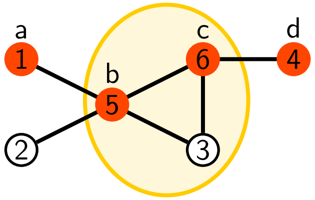


V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the minimum size of C?

Modelization with a graph

Distinguishing Transversal C = subset of vertices which is separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.

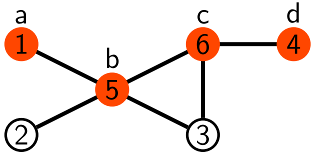


V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

Given a graph G, what is the minimum size of C?

Modelization with a graph

Distinguishing Transversal C = subset of vertices which is separating : $\forall u, v \in V, N[u] \cap C \neq N[v] \cap C$.



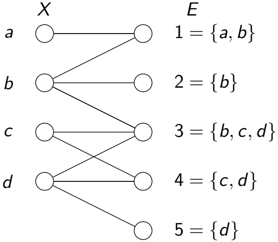
V \ C	a	b	c	d
1	•	•	-	-
2	-	•	-	-
3	-	•	•	-
4	-	-	•	•
5	•	•	•	-
6	-	•	•	•

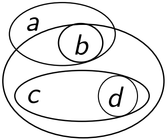
Given a graph G, what is the minimum size of C?

- Each neighborhood equivalence class is of size 1.

Distinguishing Transversal (a.k.a. Test Cover)

- Generalization on **hypergraphs** $H = (X, E)$:
  - find k vertices $C \subseteq X$ inducing $|E|$ classes $(e \cap C \neq f \cap C, \forall e, f \in E)$.

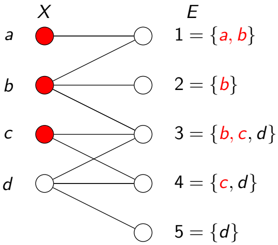


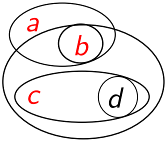


- Distinguish **all pairs of hyperedges** with minimum vertices.
- Each neighborhood equivalence class must have size 1.

Distinguishing Transversal (a.k.a. Test Cover)

- Generalization on **hypergraphs** $H = (X, E)$:
  - find k vertices $C \subseteq X$ inducing $|E|$ classes ($e \cap C \neq f \cap C, \forall e, f \in E$).



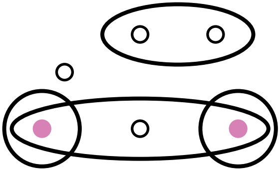


- Distinguish **all pairs of hyperedges** with minimum vertices.
- Each neighborhood equivalence class must have size 1.

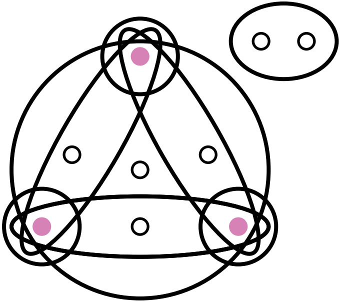
Outline

Shattered set

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



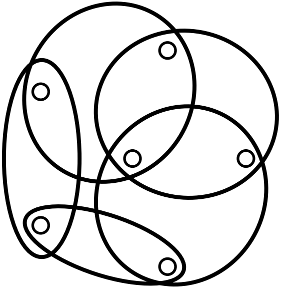
A 2-shattered set



A 3-shattered set

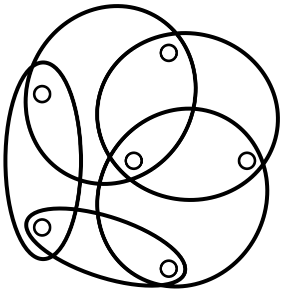
Vapnik Chervonenkis (VC) dimension of a hypergraph

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



Vapnik Chervonenkis (VC) dimension of a hypergraph

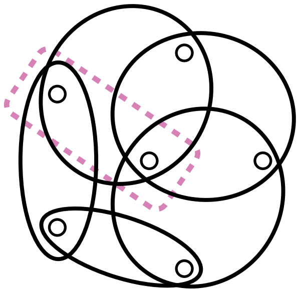
- A set C is shattered if $\forall Y \subseteq C$, $\exists e \in E$, s.t $e \cap C = Y$.
- VC-dimension of H: largest size of a shattered set.



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

Vapnik Chervonenkis (VC) dimension of a hypergraph

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

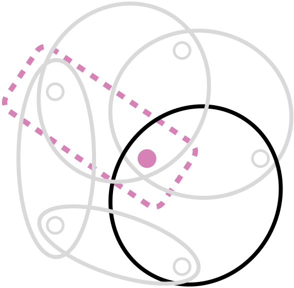


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

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

Vapnik Chervonenkis (VC) dimension of a hypergraph

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

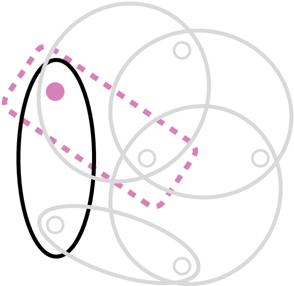


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

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

Vapnik Chervonenkis (VC) dimension of a hypergraph

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

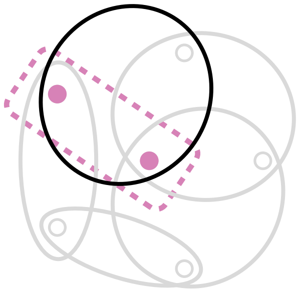


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

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

Vapnik Chervonenkis (VC) dimension of a hypergraph

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

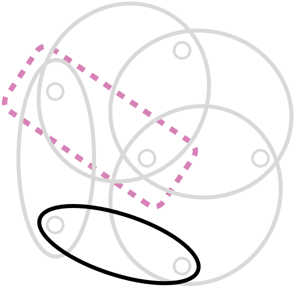


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

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

Vapnik Chervonenkis (VC) dimension of a hypergraph

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

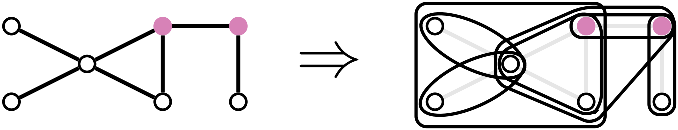


No 3-shattered set $\Rightarrow$ VC-dim ≤ 2

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

VC dimension of a graph

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



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

Outline

Cardinality Constraint Problems

- Find a solution of **cardinality** k (given in the input) s.t. an objective is **maximized** (or minimized).
- Examples:
  - MAX VERTEX COVER: Find k vertices s.t. the number of covered edges is maximum.
    - Classical VERTEX COVER is FPT.
    - Decision version of MAX VERTEX COVER is W[1]-hard.
  - MAX DOMINATING SET.
  - Same problems with minimization.
  - ...

Cardinality Constraint Versions

- PARTIAL VC DIMENSION (decision).
  - Generalize VC DIMENSION and DISTINGUISHING TRANSVERSAL.
- Given a Hypergraph $H = (X, E)$ and integers k and ℓ, find a **set $C \subseteq X$ of size k** inducing **at least ℓ equivalence classes**.

Cardinality Constraint Versions

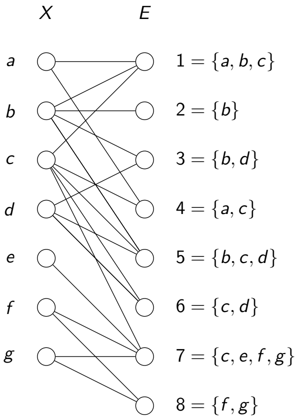
- PARTIAL VC DIMENSION (decision).
  - Generalize VC DIMENSION and DISTINGUISHING TRANSVERSAL.
- Given a Hypergraph $H = (X, E)$ and integers k and ℓ, find a **set $C \subseteq X$ of size k** inducing **at least ℓ equivalence classes**.

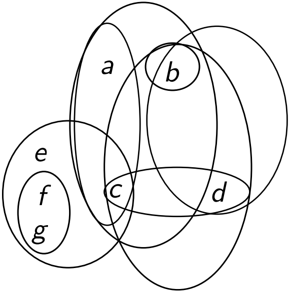
- If $\ell = 2^k$: VC DIMENSION.
- If $\ell = |E|$: DISTINGUISHING TRANSVERSAL.

Cardinality Constraint Versions

- PARTIAL VC DIMENSION (decision).
  - Generalize VC DIMENSION and DISTINGUISHING TRANSVERSAL.
- Given a Hypergraph $H = (X, E)$ and integers k and ℓ, find a **set $C \subseteq X$ of size k** inducing **at least ℓ equivalence classes**.
- If $\ell = 2^k$: VC DIMENSION.
- If $\ell = |E|$: DISTINGUISHING TRANSVERSAL.
- MAX PARTIAL VC DIMENSION: maximum number of equivalence classes with k vertices.

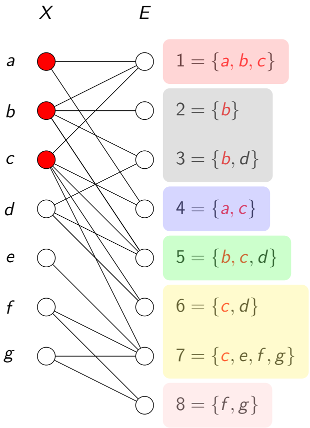
Partial VC-Dimension

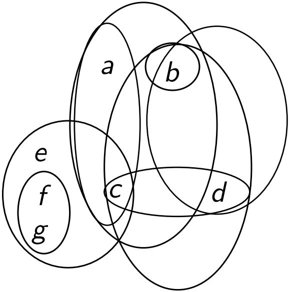




- $\ell \leqslant \min\{|E|, 2^k\}$
- $k = 3$, $\ell = 6$: YES.

Partial VC-Dimension





- $\ell \leqslant \min\{|E|, 2^k\}$
- $k = 3$, $\ell = 6$: YES.

Known Results for Partial VC Dimension

- Generalization of DISTINGUISHING TRANSVERSAL:
  - **NP-hard** on many restricted classes (hypergraphs where each vertex belongs to at most 2 hyperedges,...).
- Generalization of VC DIMENSION:
  - **W[1]-complete** w.r.t. k.

Known Results for Partial VC Dimension

- Generalization of DISTINGUISHING TRANSVERSAL:
  - **NP-hard** on many restricted classes (hypergraphs where each vertex belongs to at most 2 hyperedges,...).
- Generalization of VC DIMENSION:
  - **W[1]-complete** w.r.t. k.
- Approximation of MAX PARTIAL VC DIMENSION open.

Outline

Greedy Approximation

- On a twin-free hypergraph $H = (X, E)$.
  - 2 twins are always in the same neighborhood equivalence class.
- Find k **vertices** inducing $k + 1$ **equivalence classes**.

Greedy Approximation

- On a twin-free hypergraph $H = (X, E)$.
  - 2 twins are always in the same neighborhood equivalence class.
- Find k **vertices** inducing $k+1$ **equivalence classes**.
- Add k vertices iteratively, choosing the one maximizing the number of classes.
  - At least one class is added at each step, or we have $|E|$ classes
    - Less than $|E|$ classes: there is a class with at least 2 edges e_1, e_2.
    - There is a vertex $x \in e_1, x \notin e_2$ (twin-free).
    - Adding x increase by one.

Greedy Approximation

- On a twin-free hypergraph $H = (X, E)$.
  - 2 twins are always in the same neighborhood equivalence class.
- Find k **vertices** inducing $k + 1$ **equivalence classes**.

- Add k vertices iteratively, choosing the one maximizing the number of classes.
  - At least one class is added at each step, or we have $|E|$ classes
    - Less than $|E|$ classes: there is a class with at least 2 edges e_1, e_2.
    - There is a vertex $x \in e_1, x \notin e_2$ (twin-free).
    - Adding x increase by one.

- There are at most $\min\{2^k, |E|\}$ possible classes.
- $\frac{\min\{2^k,|E|\}}{k+1}$ approximation.

Approximation - Corollaries

- If $d(H) \leqslant \Delta$: at most $(k(\Delta+1)+2)/2$ classes [LOTS OF PEOPLE]
- Ratio is then: $\frac{k(\Delta+1)+2}{2(k+1)} \leq (\Delta+1)/2$.

Approximation - Corollaries

- If $d(H) \leqslant \Delta$: at most $(k(\Delta+1)+2)/2$ classes [Lots of people]
- Ratio is then: $\frac{k(\Delta+1)+2}{2(k+1)} \leq (\Delta+1)/2$.

- If $\mathrm{VC}(H) \leqslant d$: at most k^d+1 classes. [Sauer-Shelah lemma]
- Ratio is then: $\frac{k^d+1}{k+1} \leqslant k^{d-1}$.
- Can be extended to $|E|^{d-1/d}$.

Approximation - Corollaries

- If $d(H) \leqslant \Delta$: at most $(k(\Delta+1)+2)/2$ classes [Lots of people]
- Ratio is then: $\frac{k(\Delta+1)+2}{2(k+1)} \leq (\Delta+1)/2$.

- If $\mathrm{VC}(H) \leqslant d$: at most k^d+1 classes. [Sauer-Shelah lemma]
- Ratio is then: $\frac{k^d+1}{k+1} \leqslant k^{d-1}$.
- Can be extended to $|E|^{d-1/d}$.
  - Hypergraphs with no 4-cycles in its bipartite incidence graph: $\mathrm{VC} \leqslant 3 \rightarrow |E|^{2/3}$-approx.
  - Hypergraphs with maximum edge size d: $\mathrm{VC} \leqslant d$.
  - Neighborhood hypergraphs of interval graphs: VC $\leqslant 2$: $E^{1/2}$-approx.
  - Many other (but not bipartite or split).

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.
- Two cases:

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.
- Two cases:
  1. $k < \log_2 \log_2(n)$:
     - 2^k-approximation $\rightarrow 2^{\log_2 \log_2 n}$: $\log_2 n$**-approximation (in polynomial time)**.

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.
- Two cases:
  1. $k < \log_2 \log_2(n)$:
     - 2^k-approximation $\rightarrow 2^{\log_2 \log_2 n}$: $\log_2 n$**-approximation (in polynomial time)**.
  2. $k > \log_2 \log_2(n) \Rightarrow n < 2^{2^k}$.
     - Brute force $\binom{n}{k} = \binom{2^{2^k}}{k}$: **Exact algorithm in fpt-time**.

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.
- Two cases:
  1. $k < \log_2 \log_2(n)$:
     - 2^k-approximation $\rightarrow 2^{\log_2 \log_2 n}$: $\log_2 n$-**approximation (in polynomial time)**.
  2. $k > \log_2 \log_2(n) \Rightarrow n < 2^{2^k}$.
     - Brute force $\binom{n}{k} = \binom{2^{2^k}}{k}$: **Exact algorithm in fpt-time**.
- All together: **approximation algorithm in fpt-time** ($\log_2(n)$-approximation in time $O^*(2^{k2^k})$).

fpt-time approximation within arbitrarily small ratios

- Known: 2^k-approximation in polynomial-time.
- Goal: $\log_2(n)$-approximable in fpt-time.
- Two cases:
  1. $k < \log_2 \log_2(n)$:
     - 2^k-approximation $\rightarrow 2^{\log_2 \log_2 n}$: $\log_2 n$-**approximation (in polynomial time)**.
  2. $k > \log_2 \log_2(n) \Rightarrow n < 2^{2^k}$.
     - Brute force $\binom{n}{k} = \binom{2^{2^k}}{k}$: **Exact algorithm in fpt-time**.
- All together: **approximation algorithm in fpt-time** ($\log_2(n)$-approximation in time $O^*(2^{k2^k})$).
- Generalization: replace $\log_2(n)$ by any strictly increasing function of n.
  - A **worse running time** implies a **better ratio**.

Scheme

- Using Baker: MAX PARTIAL VC DIMENSION admits an **EPTAS** on neighborhood hypergraph of **planar graphs**.

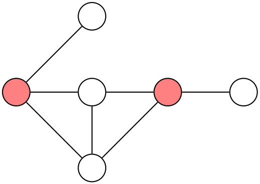
Outline

Bounded degree graphs

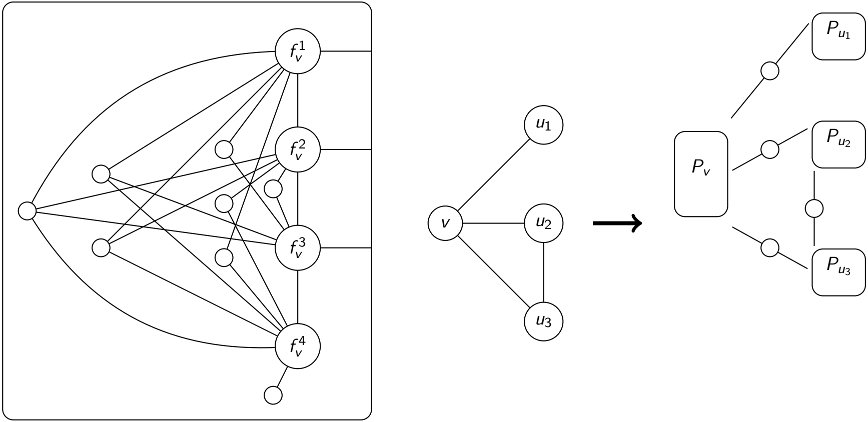
- MAX PARTIAL VC DIMENSION in APX for neighborhood hypergraphs of bounded degree graphs.
- MAX PARTIAL VC DIMENSION APX-hard for neighborhood hypergraphs of graphs with degree at most 7.

Bounded degree graphs

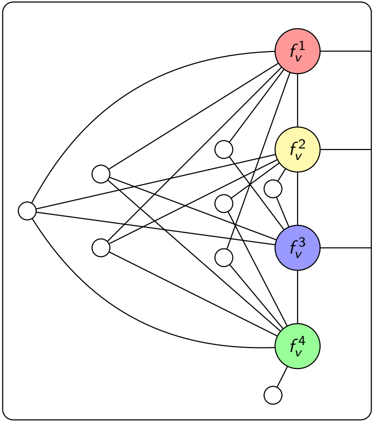
- Max Partial VC Dimension in APX for neighborhood hypergraphs of bounded degree graphs.
- Max Partial VC Dimension APX-hard for neighborhood hypergraphs of graphs with degree at most 7.
- L-reduction from Max Partial Vertex Cover in cubic graphs (APX-complete).
  - Select k vertices to cover the max number of edges.

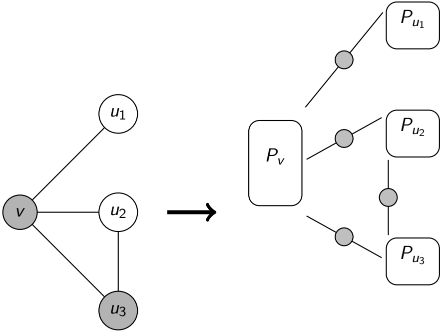


$k = 2$



f_v^1
f_v^2
f_v^3
f_v^4
v
u_1
u_2
u_3
P_v
P_{u_1}
P_{u_2}
P_{u_3}





- Add the 4 f for each selected vertex of VC.
  - $2^4 - 4$ classes in the vertex-gadget.
  - The edge vertex of covered edges is alone in its class.

Final words

- Generalization of VC-DIMENSION & DISTINGUISHING TRANSVERSAL.
- Constant-ratio approximation?

谢谢