A Polynomial Kernel for Paw-Free Editing

William Lochet, Univeristy of Bergen.

Joint work with E. Eiben, and S. Saurabh

The result

Theorem
The paw-free modification problem has a kernel on:

- $O(k^3)$ *vertices for deletion/addition.*
- $O(k^6)$ *vertices for* ***edition****.*

We use the following structural result:

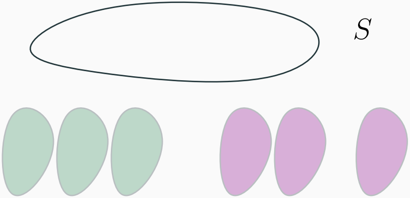
Proposition
If G *is paw-free, then the connected components of* G *are either:*

- *Triangle-free*
- *Complete-multipartite*

General approach

Let $\mathcal{H}$ be a **maximal** packing of edge-disjoint paws, either:

- $|\mathcal{H}| \geq k + 1$ and the instance is a NO-instance
- There is a set $S \subseteq V(G)$ of size at most $4k$ s.t $G - S$ is paw-free



Goal is to bound:

- triangle-free components
- complete-multipartite components

Complete multipartite components (CMC)

Twins

Reduction rule 1 (RR1)
If X is an independent set of $2k+5$ vertices with the same neighborhood, remove one vertex $x \in X$.

Twins

Reduction rule 1 (RR1)
If X is an independent set of $2k + 5$ vertices with the same neighborhood, remove one vertex $x \in X$.

Let G' denote the reduced instance.

- If (G, k) has a solution, then so does (G', k) because G' is a **subgraph** of G

Claim
Suppose A is a set of less than k pairs of vertices, such that $G \Delta A$ has a paw, then $G' \Delta A$ has one.

- If x not in the paw $\rightarrow$ easy.
- Only $2k$ vertices of X can be adjacent to A.
- In $G' \Delta A$ there at least 4 vertices of X have the same neighborhood as x in $G \Delta A \rightarrow$ can replace x in the paw.

Number of parts

A very similar arguments shows safeness of the following rule.

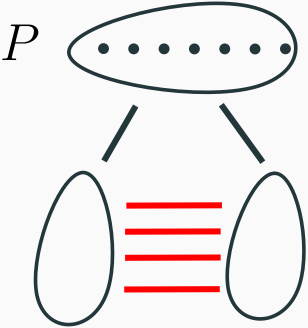
Reduction rule 2 (RR2)
If there is a complete multipartite subgraph C of G with $2k + 5$ parts having the same neighborhood **outside** of C, then remove one of these parts.

- RR1 is easy to apply.
- RR2 not so obvious $\rightarrow$ specific situations.

Size of the parts

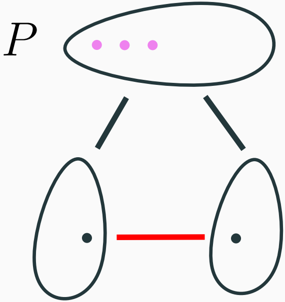
Suppose RR 1 cannot be applied anymore.

Reduction rule 3 (RR3)
If C is a CMC of $G - S$ and P a part of size at least $4k + 5$, removed all the edges between the other parts of C, and **decrease** k accordingly.



Size of the parts

Suppose there is solution A which **does not use** one of these edges.



- There are $2k + 5$ vertices of P not adjacent to A.
- These vertices belong to a CCM of $G \Delta A$.
- They are **twins** in G and RR1 could have been applied.

Total size

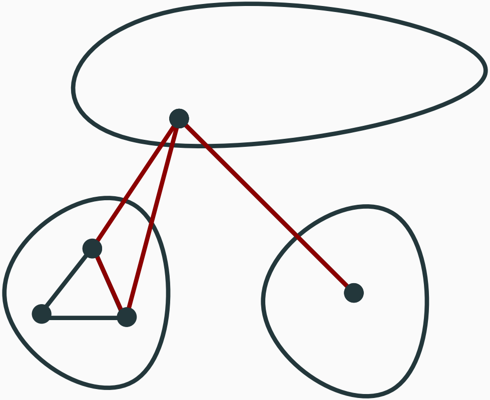
Suppose RR1 and RR3 cannot be applied anymore.

Lemma
If C is a CMC of $G - S$ and $|C| > (4k+5)^2$, then either RR2 can be applied ***in polynomial time****, or (G, k) is a NO instance.*

- RR3 cannot be applied $\Rightarrow$ no part is bigger than $4k+5$
- If there is more than $4k+5$ parts, $2k+5$ won't be touched by a solution $\Rightarrow$ we can apply RR2.

Number of CMC

Lemma
For any $s \in (S \cup S')$, s is adjacent to at most one CMC of $G - S$



Summary

Lemma
If RR1-2-3 cannot be applied, then either (G, k) is a NO-instance or the set of vertices in CMCs of $G - S$ has size $O(k^3)$.

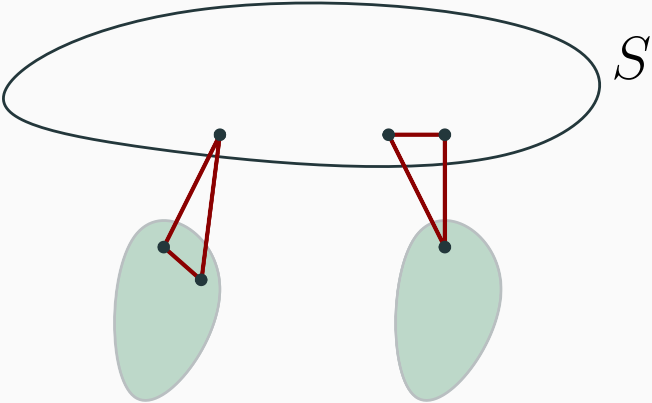
- We know there is at most $|S|$ of components
- Each has size at most $O(k^2)$

Triangle Free Components

Removing triangles

Claim

There exists a set S' of size $O(k^2)$ such that no vertex of TFC of $G - (S' \cup S)$ belongs to a triangle.



First observation

Lemma
If $x \in G$ has $6k+10$ neighbors in TFC of $G-S$ and A is a solution, then x is in a TFC of $G\Delta A$.

Suppose x belongs to some CMC C in $G\Delta A$.

- $4k+10$ of the neighbors won't be adjacent to the solution
- They belong to at most **two** parts of C
- One part is bigger than $2k+5$ and we can apply RR1

Kernel for deletion

Theorem
Paw-free deletion admits a $O(k^3)$ kernel.

- At most $O(k^3)$ vertices belongs to CMC of $G' - S$.
- For every $s \in (S' \cup S)$, **mark** $6k + 10$ vertices in TFC.
- Remove all the unmarked vertices.

The graph G' induced by all the marked vertices is the kernel.

Proof

Claim
For any set A of less than k pairs, $G\Delta A$ is paw-free $\iff$ $G'\Delta A$ is paw-free

$\Rightarrow$ G' is a **subgraph** of G

$\Leftarrow$ Suppose $G'\Delta A$ is paw free, but $G\Delta A$ has a paw $x_1x_2x_3 - x_4$

- The triangle $x_1x_2x_3$ is a triangle of G and thus G'.
- Thus x_4 is an **unmarked vertex** of some TFC.
- x_3 is a vertex of $(S \cup S')$ with $6k + 10$ adjacent vertices.
- In $G'\Delta A$, x_3 has to be in a TFC, a contradiction.

Edge-editing

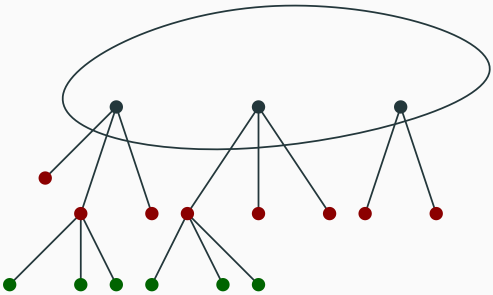
- Marking $6k + 10$ vertices $\rightarrow$ know which vertices of $(S \cup S')$ must belong to TFC in the solution.

Edge-editing

- Marking $6k + 10$ vertices $\rightarrow$ know which vertices of $(S \cup S')$ must belong to TFC in the solution.
- In the **modification problem**, vertices in TFC of $G - (S \cup S')$ can end up in CMC of $G \Delta A$.

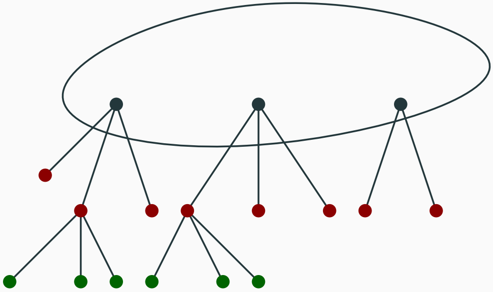
Edge-editing

- Marking $6k + 10$ vertices $\rightarrow$ know which vertices of $(S \cup S')$ must belong to TFC in the solution.
- In the **modification problem**, vertices in TFC of $G - (S \cup S')$ can end up in CMC of $G \Delta A$.



Edge-editing

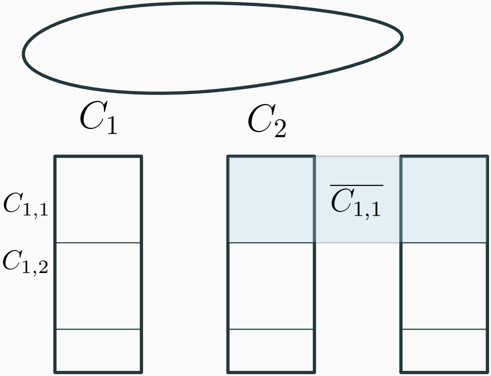
- Marking $6k + 10$ vertices $\rightarrow$ know which vertices of $(S \cup S')$ must belong to TFC in the solution.
- In the **modification problem**, vertices in TFC of $G - (S \cup S')$ can end up in CMC of $G \Delta A$.



Lemma
Let (G, k) be a YES instance, there exists a solution A s.t no CMC of $G \Delta A$ contains a vertex at distance 5 from S in G.

Distance to S

- Let A be a solution **minimizing** the CMC of $D\Delta A$
- $C_1, \ldots, C_l$ the parts of a CMC C of $G\Delta A$
- $C_{i,j}$ the set of vertices of C_i at distance j from S
- $\overline{C_{i,j}} = \bigcup_{t \neq i} C_{t,j}$



Proof

Lemma
For any $j > 3$, $i \in [l]$, *if* $C_{i,0} \cup C_{i,1}$ *is non empty, then* $C_{i,j}$ *is.*

Suppose it is not:

- A must contain all the edges in $C_{i,j} \times (\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}})$
- Removing $C_{i,j}$ from C costs $|E(C_{i,j}, (\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}}))|$

Proof

Lemma
For any $j > 3$, $i \in [l]$, *if* $C_{i,0} \cup C_{i,1}$ *is non empty, then* $C_{i,j}$ *is.*

Suppose it is not:

- A must contain all the edges in $C_{i,j} \times (\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}})$
- Removing $C_{i,j}$ from C costs $|E(C_{i,j}, (\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}}))|$
- This means $|\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}}| \geq |\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}}|$

Proof

Lemma
For any $j > 3$, $i \in [l]$, if $C_{i,0} \cup C_{i,1}$ is non empty, then $C_{i,j}$ is.

Suppose it is not:

- A must contain all the edges in $C_{i,j} \times (\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}})$
- Removing $C_{i,j}$ from C costs $|E(C_{i,j}, (\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}}))|$
- This means $|\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}}| \geq |\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}}|$
- A must contain $(C_{i,0} \cup C_{i,1}) \times (\overline{C_{i,j-1}} \cup \overline{C_{i,j}} \cup \overline{C_{i,j+1}})$
- Removing $(C_{i,0} \cup C_{i,1})$ from C costs $(C_{i,0} \cup C_{i,1}) \times (\overline{C_{i,0}} \cup \overline{C_{i,1}} \cup \overline{C_{i,2}})$.

Contradiction to the minimality of $|C|$!

Proof

For any j, let $S_j = \bigcup_{i\in[r]} C_{i,j}$. Previous result implies that:

Lemma
If S_i non-empty for $i > 3$, then $(S_1 \cup S_0) \times S_i \subset A$.

Indeed S_j and $(S_0 \cup S_1)$ belong to different parts of C.

Lemma
If S_5 is non empty,$|S_4| \geq |S_1 \cup S_0|$

- S_4 is not empty
- A contains $S_5 \times (S_0 \cup S_1)$
- Disconnecting S_5 from C costs $|E_G(S_4, S_5)|$
- This means that $|S_4||S_5| \geq |E_G(S_4, S_5)| \geq |S_5||S_1 \cup S_0|$

Proof

Lemma
S_j is empty for $j \geq 5$

- S_4 is not empty
- A contains $S_4 \times (S_0 \cup S_1)$
- Disconnecting S_1 from S_0 costs less than $|S_1 \cup S_0|^2$
- Previous lemma $\Rightarrow |S_4||S_1 \cup S_0| \geq |S_1 \cup S_0|^2$

Therefore the solution A' obtained from A by disconnecting S_1 from S_0 and removing all the pairs of A of the form (x, y) with $x \in S_0$, $y \in S_j$ for $i \geq 0$ and $j \geq 1$ is a better solution.

Proof

Lemma
S_j is empty for $j \geq 5$

- S_4 is not empty
- A contains $S_4 \times (S_0 \cup S_1)$
- Disconnecting S_1 from S_0 costs less than $|S_1 \cup S_0|^2$
- Previous lemma $\Rightarrow |S_4||S_1 \cup S_0| \geq |S_1 \cup S_0|^2$

Therefore the solution A' obtained from A by disconnecting S_1 from S_0 and removing all the pairs of A of the form (x, y) with $x \in S_0$, $y \in S_j$ for $i \geq 0$ and $j \geq 1$ is a better solution.

Lemma
Let (G, k) be a YES instance, there exists a solution A s.t no CMC of $G \Delta A$ contains a vertex at distance 5 from S in G.

Thank you!