Error Compensation in Leaf Root Problems

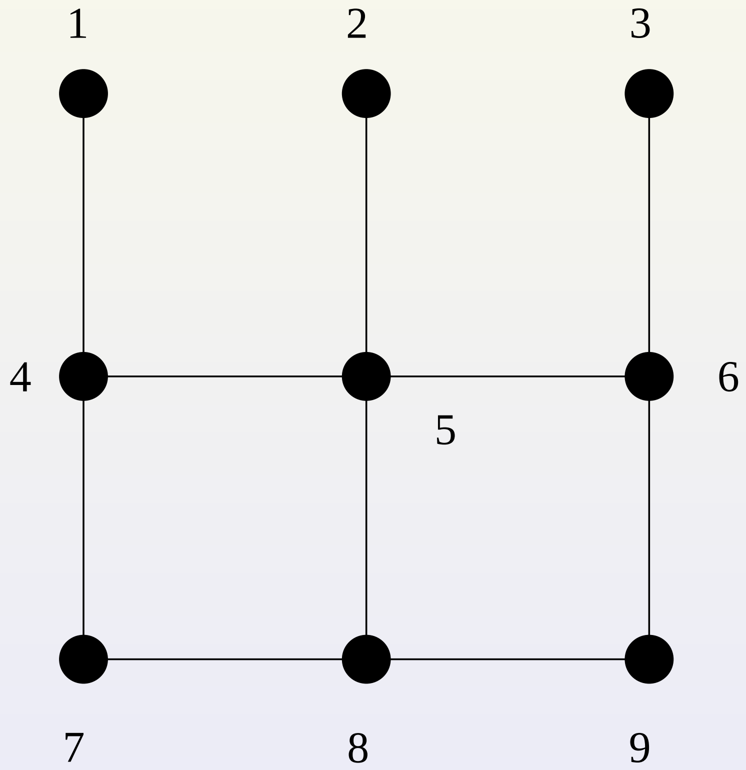
Michael Dom, Jiong Guo, Falk Hüffner, and Rolf Niedermeier

Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13,
D-72076 Tübingen, Fed. Rep. of Germany

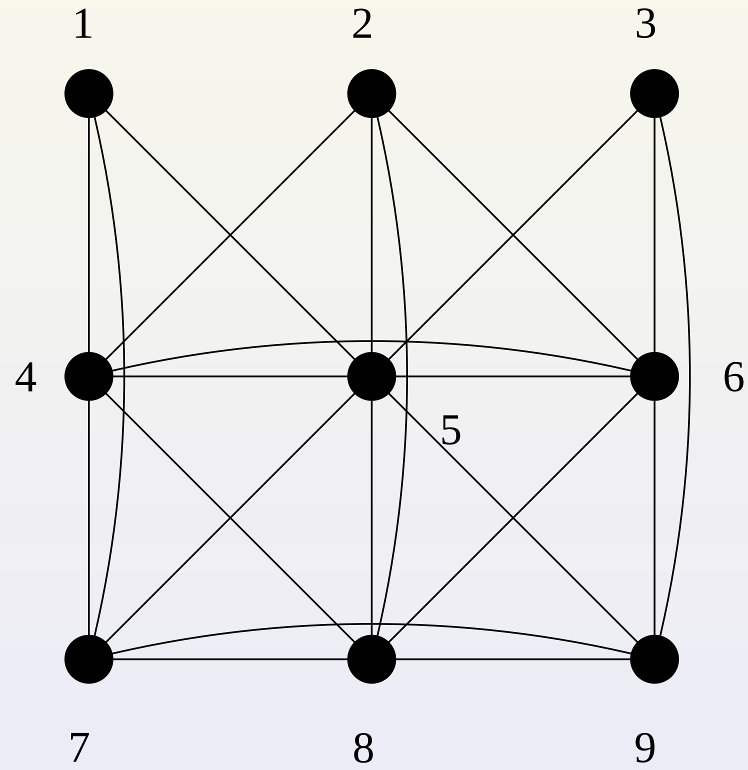
Structure of the Talk

- **Introduction: Leaf Roots and Leaf Root Problems**
- Forbidden subgraph characterization for 3-LEAF ROOT
- Fixed-parameter tractability of CLOSEST 3-LEAF ROOT

k-Roots and k-Powers



2-root

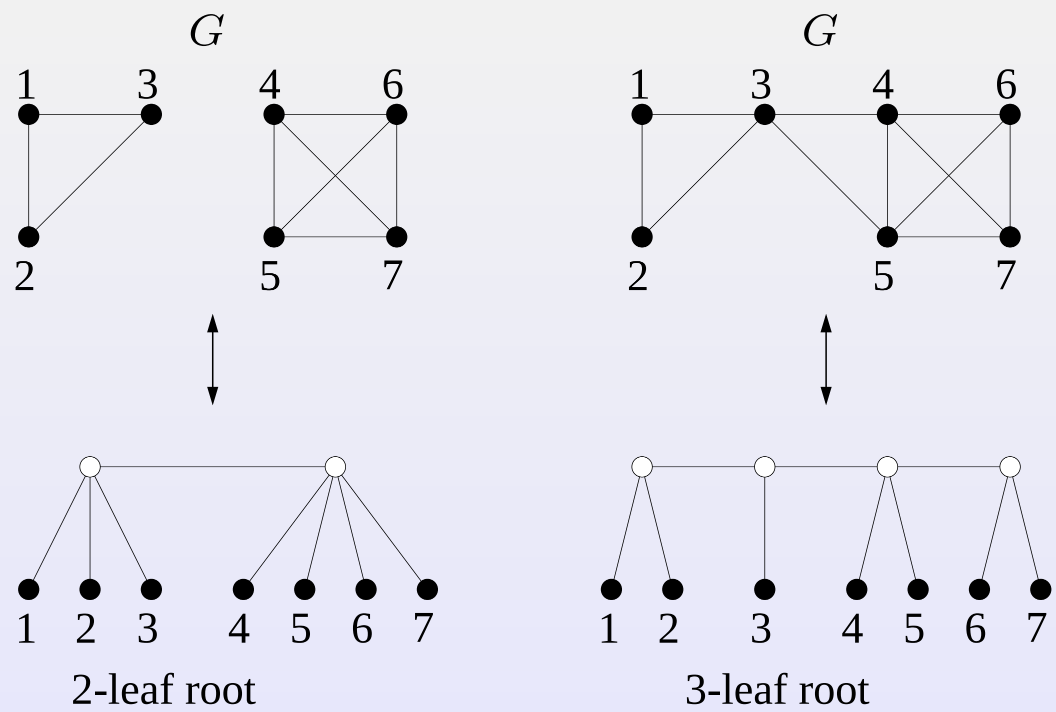


2-power

k-Leaf Roots

Given a graph $G = (V, E)$. A k-*leaf root* of G is a tree T with the following properties:

1. The leaves of T are the elements of V
2. $d_T(u, v) \leq k \Leftrightarrow (u, v) \in E$, where d_T denotes the distance between u and v in T



Problem: k-LEAF ROOT

k-LEAF ROOT (LRk)
Instance: A graph $G = (V, E)$.
Question: Is there a k-leaf root of G?

Complexity of k-LEAF ROOT:

- $O(|V| + |E|)$ for $k = 3$
- $O(|V|^3)$ for $k = 4$[a]
- unknown for $k \geq 5$

---

[a] N. Nishimura, P. Ragde, D. M. Thilikos, *J. Algorithms*, 2002

Problem: CLOSEST k-LEAF ROOT

CLOSEST k-LEAF ROOT (CLRk)
Instance: A graph $G = (V, E)$ and a nonnegative integer l.
Question: Is there a graph G' such that G' has a k-leaf root and that G' and G differ by at most l edges:
$|(E(G') \setminus E(G)) \cup (E(G) \setminus E(G'))| \leq l$

Variations:

- CLRk EDGE DELETION
- CLRk EDGE INSERTION
- CLRk VERTEX DELETION

Complexity

Complexity of closest k-leaf root problems.

	$k = 2$	$k \geq 3$
“Edge editing”	NP-complete[a]	NP-complete[b]
Edge deletion	NP-complete[c]	NP-complete[b]
Edge insertion	P	NP-complete[b]
Vertex deletion	NP-complete[d]	NP-complete[d]

[a] M. Křivánek and J. Morávek, *Acta Informatica*, 1986

[b] M. Dom, J. Guo, F. Hüffner, R. Niedermeier, *15th ISAAC*, 2004

[c] A. Natanzon, *Master Thesis*, 1999

[d] J. M. Lewis and M. Yannakakis, *JCSS*, 1980

Structure of the Talk

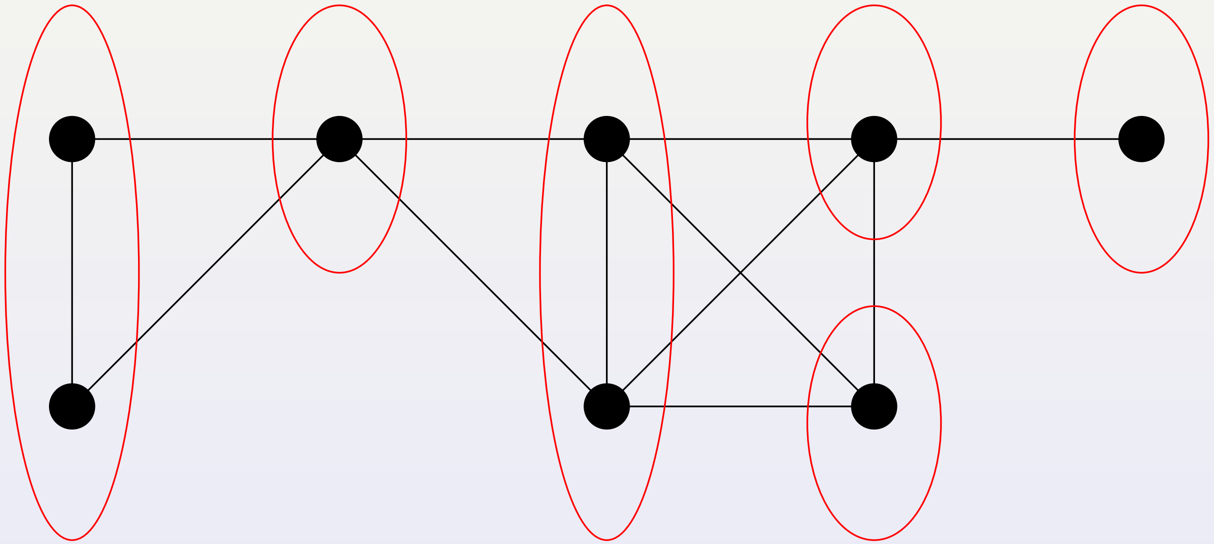
- Introduction: Leaf Roots and Leaf Root Problems
- **Forbidden subgraph characterization for 3-LEAF ROOT**
- Fixed-parameter tractability of CLOSEST 3-LEAF ROOT

Forbidden Subgraph Charact.

- Graph property Π
- Set F of forbidden subgraphs
- $G \in \Pi$
  $\Leftrightarrow$
  G does not contain any of the forbidden subgraphs as induced subgraph

Critical Cliques

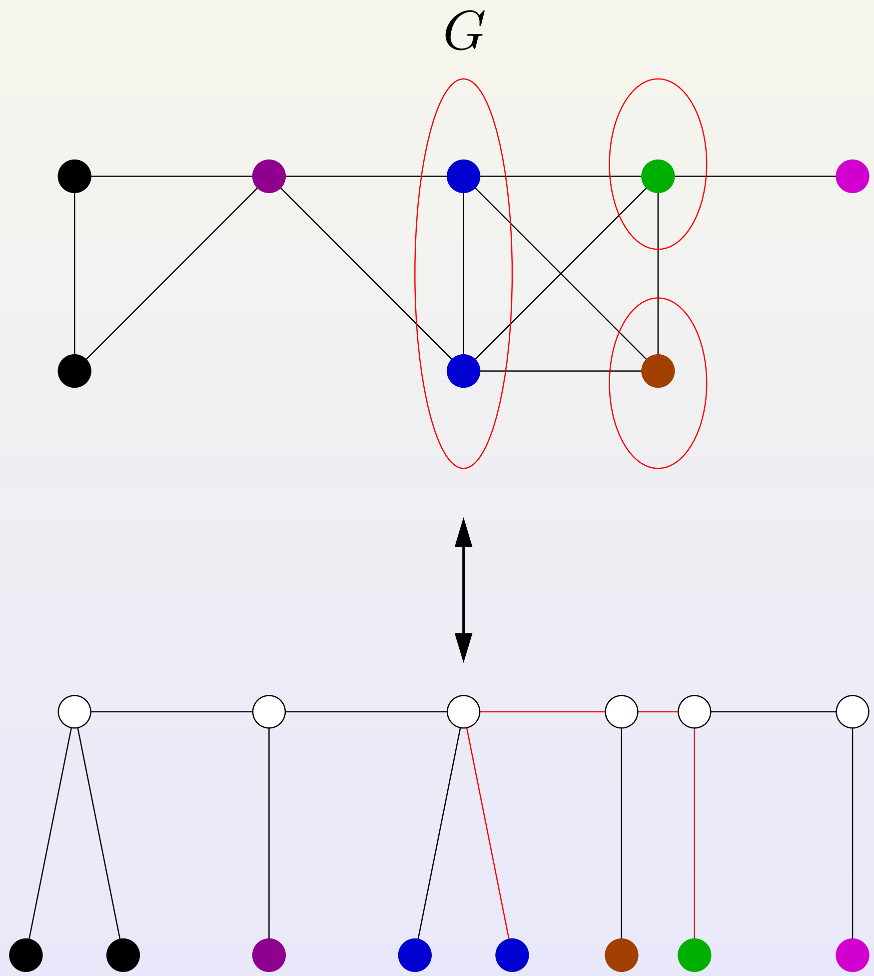
A *critical clique*[a] of a graph G is a clique K where the vertices of K all have the same set of neighbors in $G \setminus K$, and K is maximal under this property.



[a] Introduced by G.-H. Lin, P. E. Kearney, T. Jiang, *11th ISAAC*, 2000

Forbidden Subgraphs (1)

Lemma. *If a graph G has a 3-leaf root, then every clique in G consists of at most two critical cliques.*



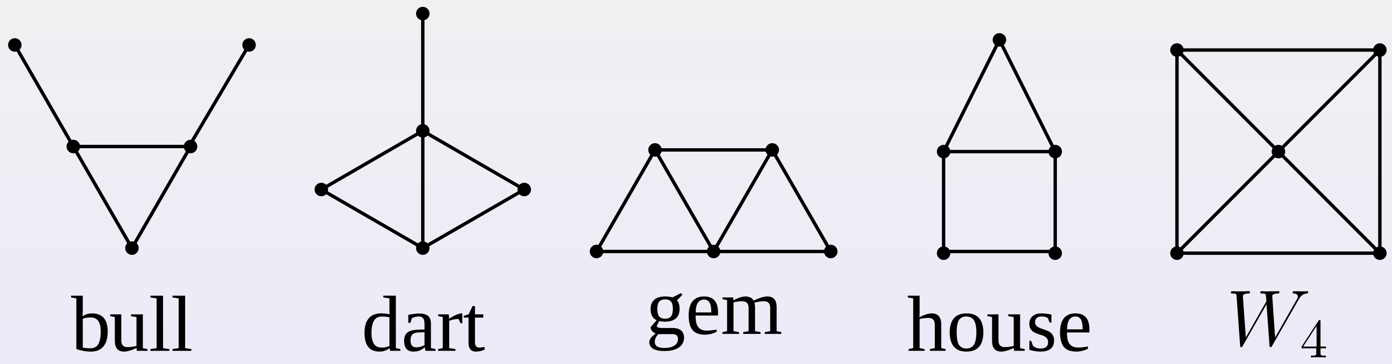
There is no 3-leaf root!

Forbidden Subgraphs (2)

Lemma. *For a graph G, the following statements are equivalent:*

(1) There is a clique K in G that consists of at least three critical cliques.

(2) G contains a bull, dart, gem, house or W_4 as induced subgraph.



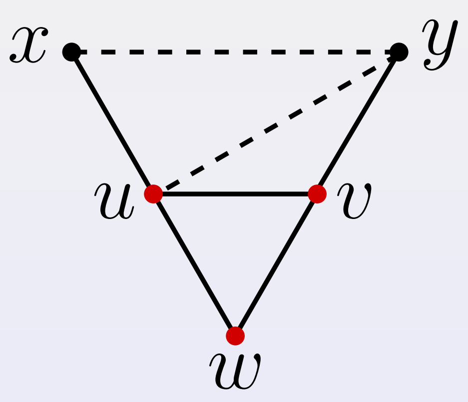
Forbidden Subgraphs (3)

Proof.
($1 \Rightarrow 2$): Let u, v and w be vertices of the same clique that belong to three different critical cliques.

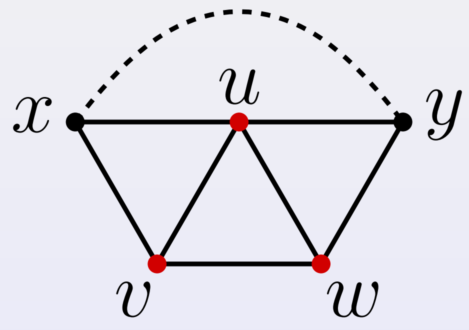
(a) There is a vertex (x) in G which is connected to exactly one of the three vertices u, v or w, or

(b) there is no such vertex in G

(a)



(b)



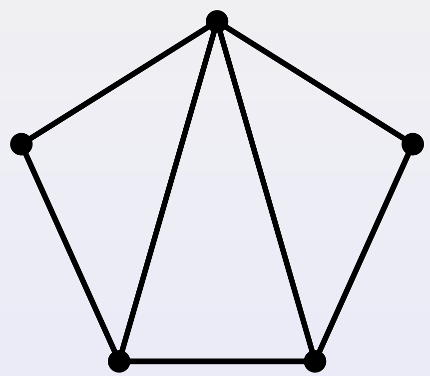
($2 \Rightarrow 1$): Easy to see.

Forbidden Subgraphs (4)

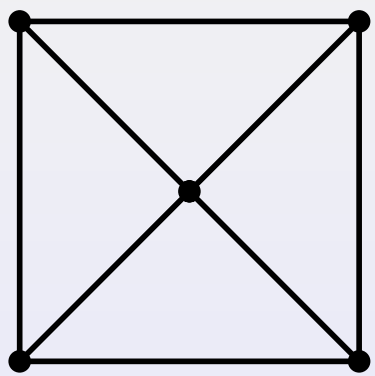
The following lemma is well-known and easy to see:

Lemma. *If a graph G has a k-leaf root for any k, then G is chordal.*

(A graph G is chordal, iff it contains no induced chordless cycle.)



chordal



not chordal

Forbidden Subgraphs (4)

The following lemma is well-known and easy to see:

Lemma. *If a graph G has a k-leaf root for any k, then G is chordal.*

Summary of the lemmas:

- 3-leaf root $\Rightarrow \leq 2$ crit. cliques per max. clique
- 3-leaf root $\Rightarrow$ chordal
- chordal and ≤ 2 crit. cliques per max. clique $\Rightarrow$ no bull, dart or gem

Forbidden Subgraphs (5)

Therefore, one direction of this theorem is clear:

Theorem. *For a graph G, the following statements are equivalent:*

(1) G has a 3-leaf root.

(2) G is chordal and contains no bull, dart, or gem as induced subgraph.

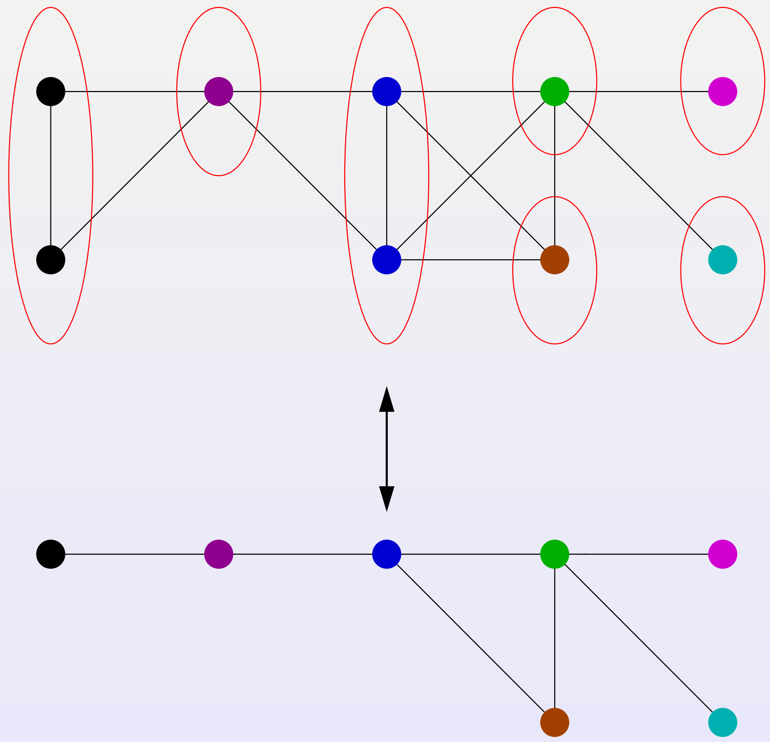
We still have to show: $(2) \Rightarrow (1)$.

We do this constructively by showing how to construct the 3-leaf root of a chordal, bull-, dart-, and gem-free graph.

The Critical Clique Graph

Given a graph G. The critical clique graph $\mathrm{CC}(G)$ has the critical cliques of G as nodes, and two nodes are connected iff the corresponding critical cliques form a larger clique in G.

G



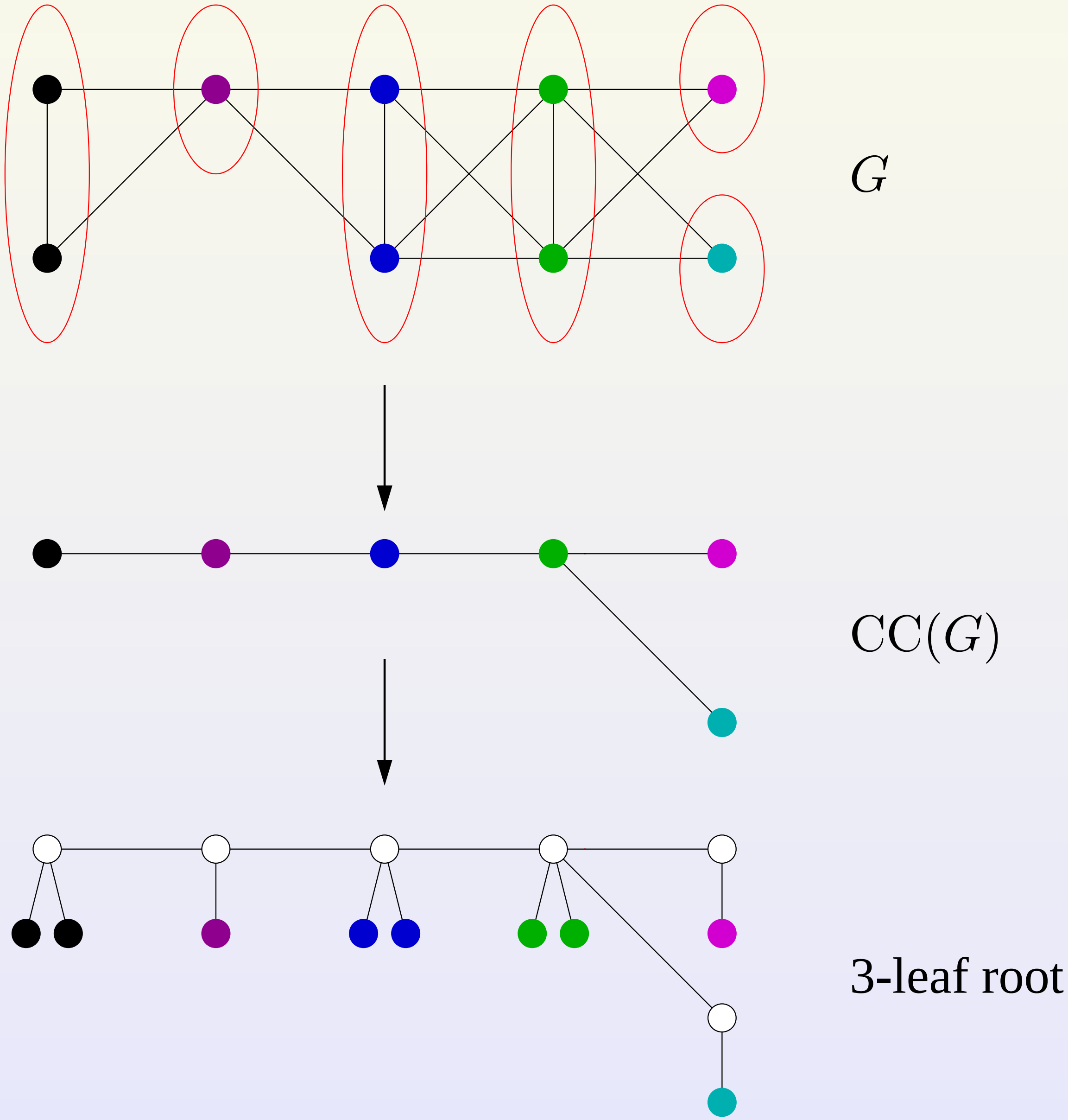
$\mathrm{CC}(G)$

How to construct a 3-leaf root

We need some more lemmas:

- **Lemma.** *A graph G is chordal iff $\mathrm{CC}(G)$ is chordal.*
- **Lemma.** *Every clique of a graph G consists of at most two critical cliques iff $\mathrm{CC}(G)$ contains no cliques of size > 2 (i.e. no triangles).*
- **Corollary.** *If a graph G is chordal, bull-, dart-, and gem-free, then $\mathrm{CC}(G)$ is a forest.*

How to construct a 3-leaf root



Structure of the Talk

- Introduction: Leaf Roots and Leaf Root Problems
- Forbidden subgraph characterization for 3-LEAF ROOT
- **Fixed-parameter tractability of CLOSEST 3-LEAF ROOT**

Fixed-Parameter Tractability

Definition of Fixed-Parameter Tractability:

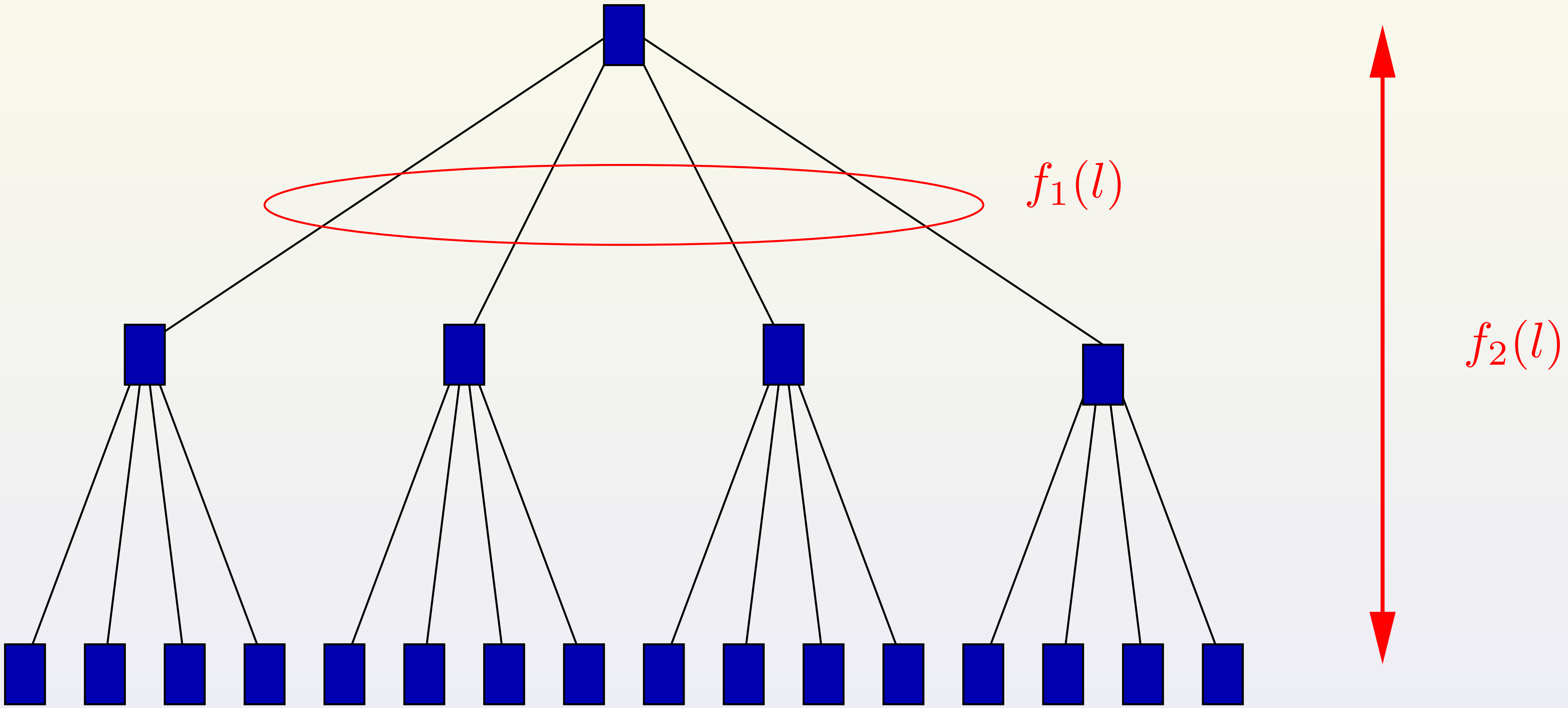
- Problem instance (x, l) with $|x| = n$
- Runtime $f(l) \cdot n^{O(1)}$ (instead of $f(n)$)

We show fixed-parameter tractability with respect to the number of editing operations l for all CLRk variations:

- CLR3 EDGE DELETION,
- CLR3 EDGE INSERTION and
- CLR3 ("Edge Editing").

(For CLR3 VERTEX DELETION see M. Dom, J. Guo, F. Hüffner, R. Niedermeier, *Proc. 15th ISAAC*, 2004.)

Search Trees



Fixed-Parameter Algorithms (1)

From the previous section, we know:
If a graph G has a 3-leaf root, then $\mathrm{CC}(G)$ is a forest.

To solve CLR3, modify the given graph so that its critical clique graph becomes a forest!

Fixed-Parameter Algorithms (2)

Basic scheme for our FPT-algorithms:

(1) Edit G to get rid of the forbidden subgraphs bull, dart, gem, house, and W_4.
Runtime[a]: $O(c^l \cdot |V|^d)$
(c, d: constants, l: number of allowed edge modifications.)
After step (1), $\mathrm{CC}(G)$ contains no triangles.

(2) Edit G to make it chordal.
After step (2), $\mathrm{CC}(G)$ is a forest.

---

[a] L. Cai, *Information Processing Letters*, 1996

CLR3 Edge Deletion

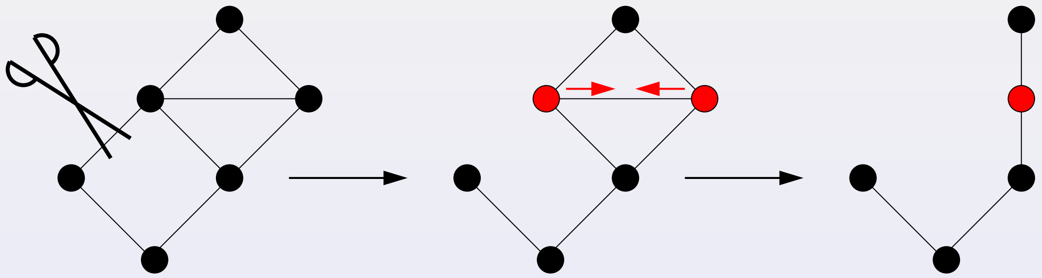
Difficulty: How can we make G chordal?

- Operate on $\mathrm{CC}(G)$ instead of G:
  Lemma. *There is always an optimal solution that does not delete any edges within a critical clique and that deletes either all or no edges between two critical cliques.*

CLR3 Edge Deletion

Difficulty: How can we make G chordal?

- Operate on $\mathrm{CC}(G)$ instead of G:
  Lemma. *There is always an optimal solution that does not delete any edges within a critical clique and that deletes either all or no edges between two critical cliques.*



“Collapsing” triangles.

CLR3 Edge Deletion

Difficulty: How can we make G chordal?

- Operate on $\mathrm{CC}(G)$ instead of G:
  Lemma. *There is always an optimal solution that does not delete any edges within a critical clique and that deletes either all or no edges between two critical cliques.*
- After step (1) $\mathrm{CC}(G)$ contains no triangles.

$\Rightarrow$ The remaining cycles in $\mathrm{CC}(G)$ cannot "collapse".

$\Rightarrow$ At least one edge of every cycle in $\mathrm{CC}(G)$ has to be deleted.

$\Rightarrow$ Step (2) is finding a maximum spanning tree on the critical clique graph.

CLR3 Edge Insertion

Main idea for step (2): Every cycle of length ≥ 4 in $\mathrm{CC}(G)$ has to be triangulated.

A minimal triangulation of a cycle with n edges consists of $n-3$ chords.[a]

$\Rightarrow$ If only l edge insertions are allowed, there cannot be a chordless cycle of length more than $l+3$.

$\Rightarrow$ $\frac{(l+3)\cdot(l+2)}{2}$-branching.

$\Rightarrow$ Runtime: $O(l^{2l} \cdot |V| \cdot |E|)$.

[a] H. Kaplan, R. Shamir, R. E. Tarjan, *SIAM J. Computing*, 1999

CLR3 (“Edge Editing”)

Cycles of length $> l + 3$ in the critical clique graph cannot be destroyed only with edge insertions.

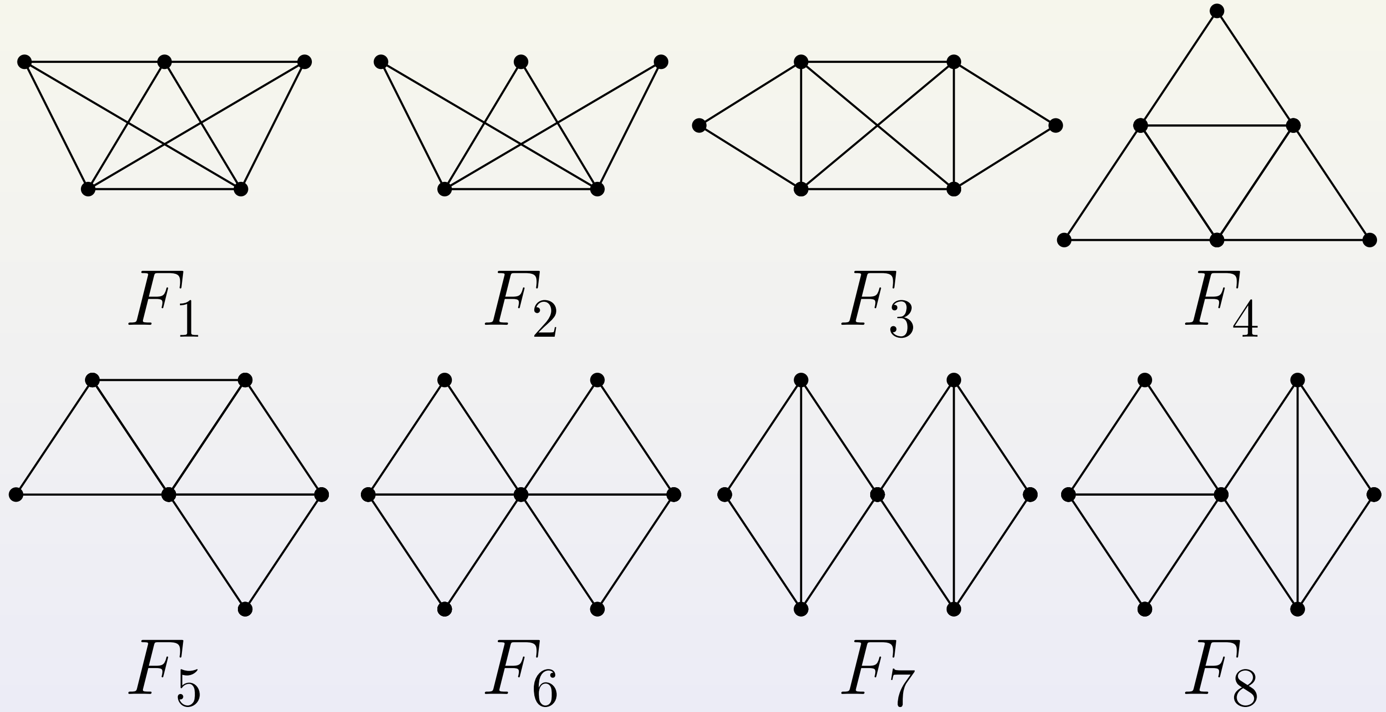
$\Rightarrow$ We first eliminate all cycles of length $\leq l + 3$, which leads to a $l + 3 + \frac{(l+3)\cdot(l+2)}{2}$-branching.

Thereafter, we eliminate the bigger cycles only with edge deletions (maximum spanning tree).

Runtime: $O(l^{2l} \cdot |V| \cdot |E|)$

Results for $k = 4$

- Forbidden subgraphs in the critical clique graph:



- FPT-algorithms for all variants of CLR4.

Open questions

- Generalization to CLOSEST k-LEAF ROOT for $k > 4$:
  - Can graphs that have a k-leaf root be recognized in polynomial time?
  - Is there a useful characterization by a small set of forbidden subgraphs?
- Extension to the closely related CLOSEST PHYLOGENETIC k-TH ROOT, where all inner nodes of the leaf root (then called “phylogenetic root”) must have degree ≥ 3?
- How small can the combinatorial explosion for CLR3, CLR4 and their variants in the parameter l (number of modifications) be made?