

Structural Parameterizations for COLORFUL COMPONENTS



IWOCA 2018

Neeldhara Misra
IIT GANDHINAGAR

Structural Parameterizations for COLORFUL COMPONENTS



IWOCA 2018

Neelghara Misra
IIT GANDHINAGAR

OUTLINE

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

COLOR CODING FOR MEC

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

COLOR CODING FOR MEC

LINEAR KERNEL FOR MEC ON TREES

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

COLOR CODING FOR MEC

LINEAR KERNEL FOR MEC ON TREES

CC BY VERTEX COVER IS FPT

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

LINEAR KERNEL FOR MEC ON TREES

CC BY VERTEX COVER IS FPT

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel [1]

LINEAR KERNEL FOR MEC ON TREES

CC BY VERTEX COVER IS FPT

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel [1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel [1]

CC BY VERTEX COVER IS FPT

OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel [1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel [1]

CC BY VERTEX COVER IS FPT

In contrast: para-NP-hard already for deletion sets to stars^[2] and paths^[1].

REFERENCES

[1] R. Dondi and F. Sikora.

Parameterized complexity and approximation issues for the colorful components problems.

Proceedings of the 12th Conference on Computability in Europe, (CiE), 2016

[2] S. Bruckner, F. Hüffner, C. Komusiewicz, R. Niedermeier, S. Thiel, and J. Uhlmann.

Partitioning into colorful components by minimum edge deletions.

Proceedings of the 23rd Annual Symposium of Combinatorial Pattern Matching (CPM), 2012

[3] A. Adamaszek, G. Blin, and A. Popa.

Approximation and hardness results for the maximum edges in transitive closure problem.

Proceedings of the 25th International Workshop on Combinatorial Algorithms (IWOCA), 2014.

[4] S. Bruckner, F. Hüffner, C. Komusiewicz, and R. Niedermeier.

Evaluation of ILP- based approaches for partitioning into colorful components.

Proceedings of the 12th International Symposium of Experimental Algorithms (SEA), 2013.

TALK OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel^[1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel^[1]

CC BY VERTEX COVER IS FPT

In contrast: para-NP-hard already for deletion sets to stars^[2] and paths^[1].

PROBLEM DEFINITIONS & MOTIVATION

The Setting

The Setting

G — A vertex-colored graph.

The Setting

G — A vertex-colored graph.

F — A subset of edges.

The Setting

G — A vertex-colored graph.

F — A subset of edges.

$G \setminus F$ — A graph with colorful components.

The Setting

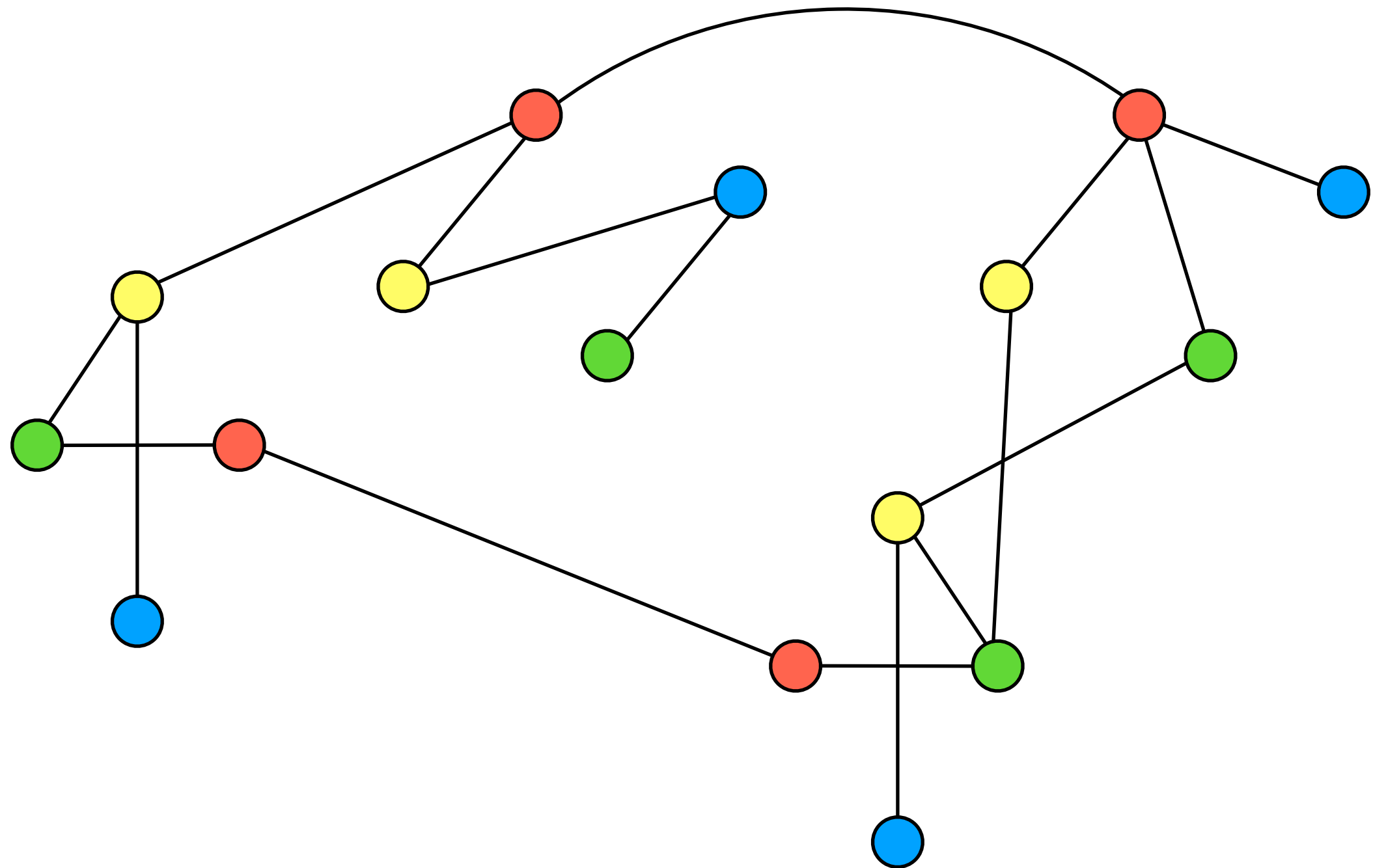
G — A vertex-colored graph.

F — A subset of edges.

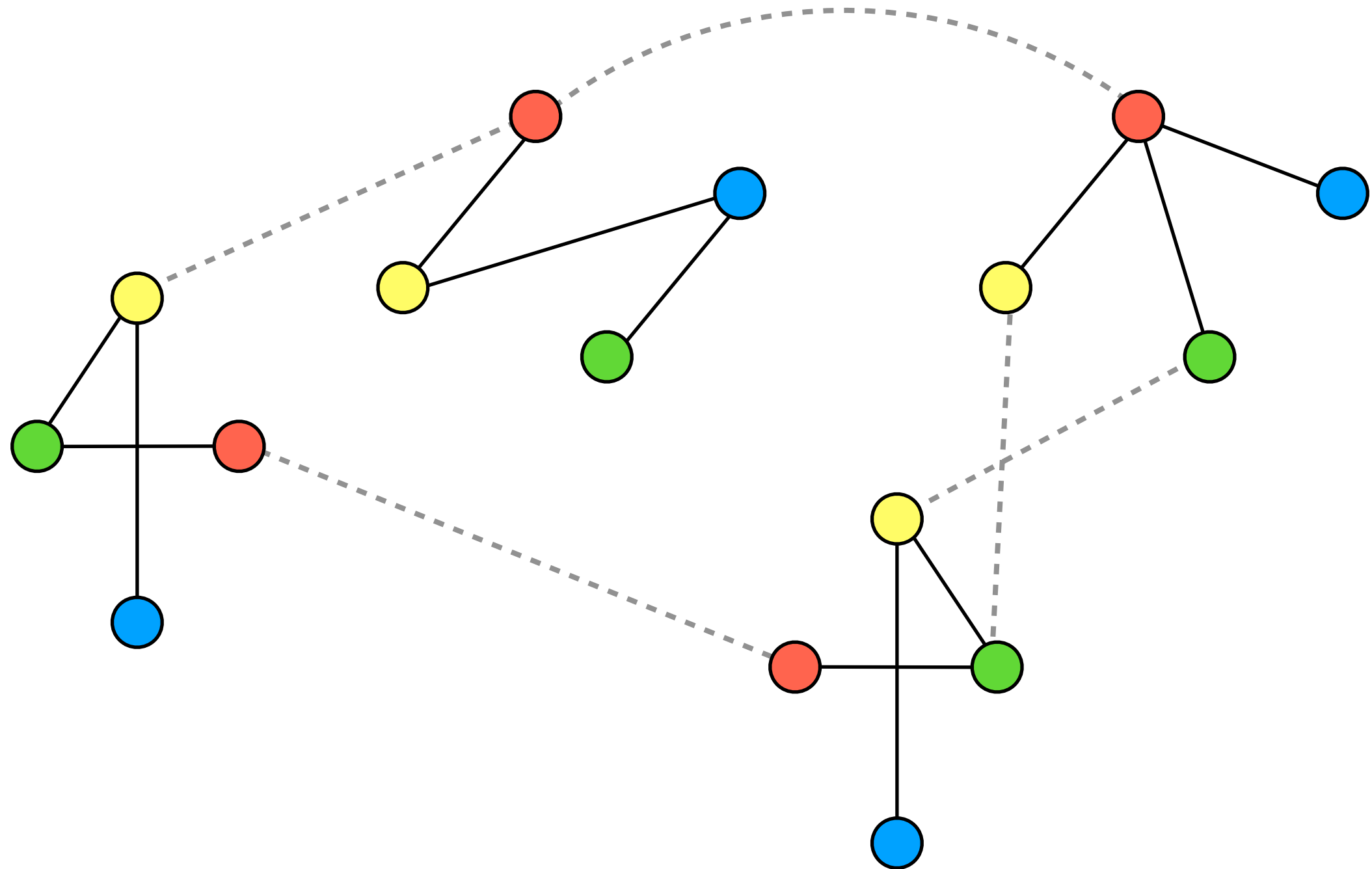
$G \setminus F$ — A graph with colorful components.

Typically: **maximize $G \setminus F$** or **minimize F** .

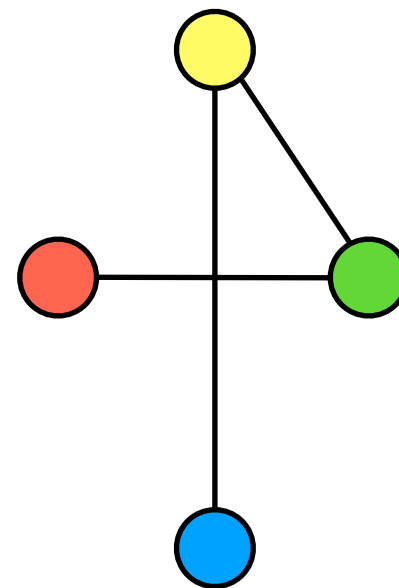
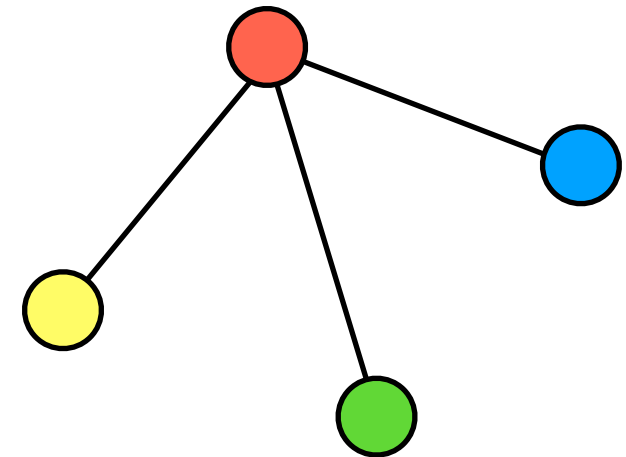
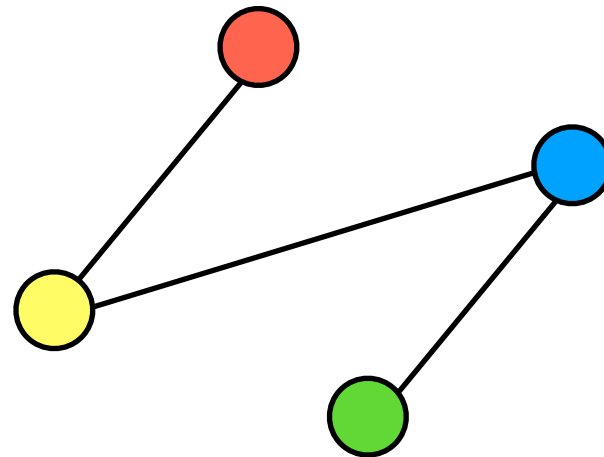
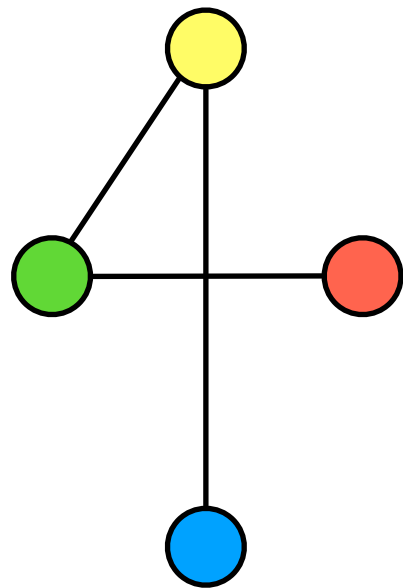
PROBLEM DEFINITIONS & MOTIVATION



PROBLEM DEFINITIONS & MOTIVATION

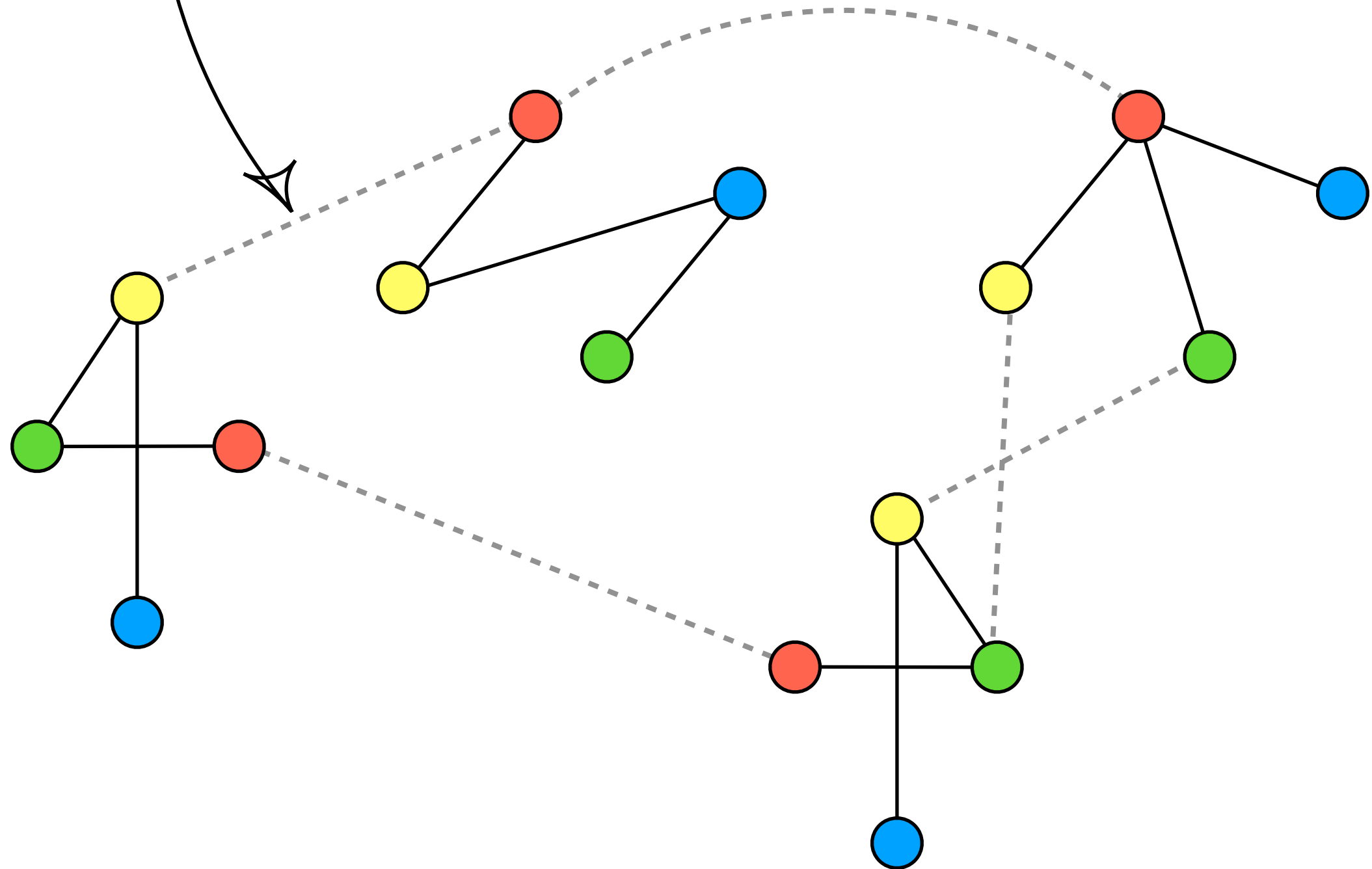


PROBLEM DEFINITIONS & MOTIVATION

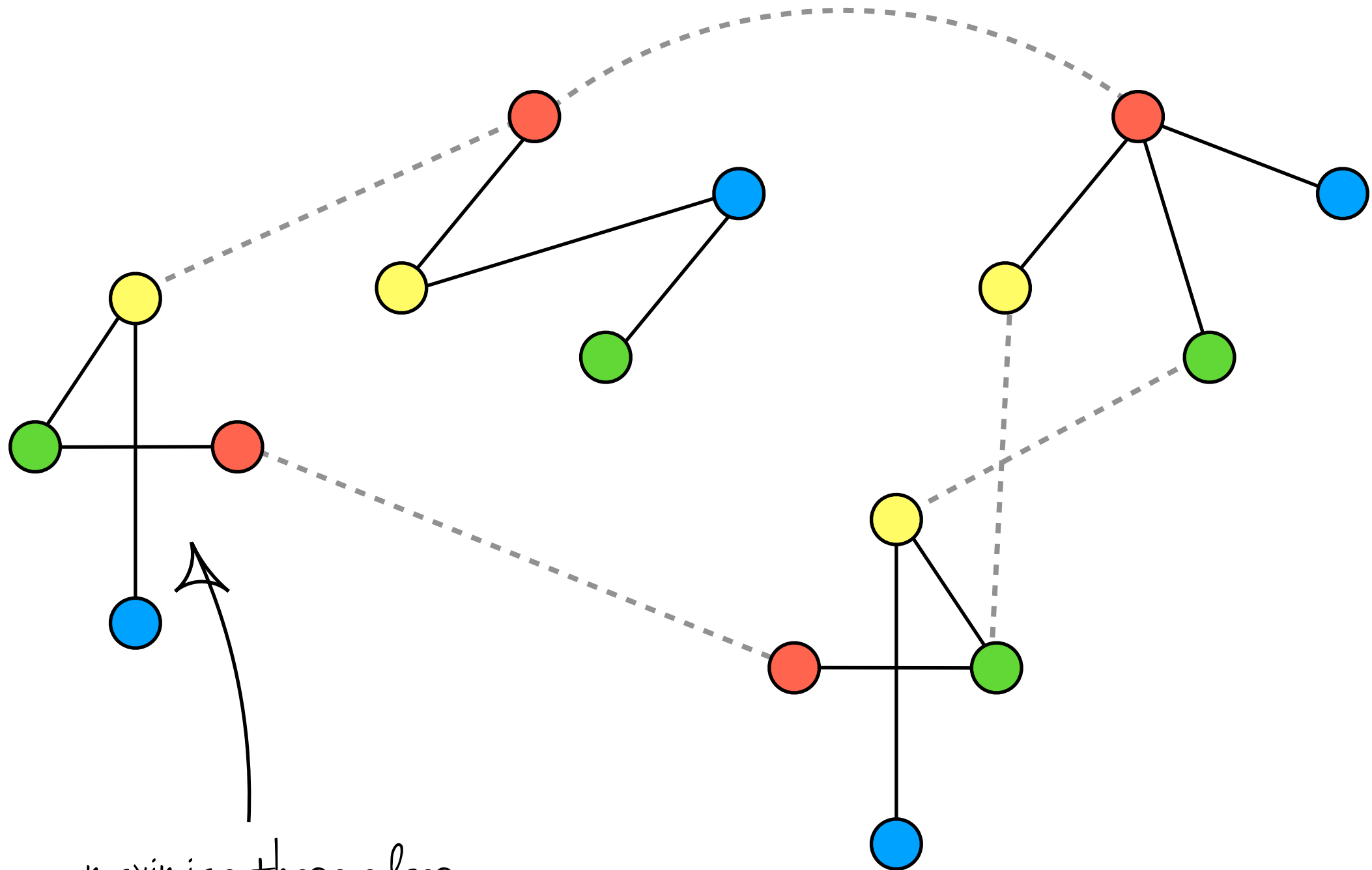


PROBLEM DEFINITIONS & MOTIVATION

minimize these edges,
AKA the cut



PROBLEM DEFINITIONS & MOTIVATION



maximize these edges,
traditionally: the transitive closure

PROBLEM DEFINITIONS & MOTIVATION

The Setting

The Setting

Parameterized Complexity

The Setting

Parameterized Complexity

Evaluate running times of algorithms as a function of the size of input **and** additional structural parameters.

The Setting

Parameterized Complexity

Evaluate running times of algorithms as a function of the size of input **and** additional structural parameters.

$$f(k)p(n)$$

The Setting

Parameterized Complexity

Evaluate running times of algorithms as a function of the size of input **and** additional structural parameters.

a structural parameter



$$f(k)p(n)$$

The Setting

Parameterized Complexity

Evaluate running times of algorithms as a function of the size of input **and** additional structural parameters.

a structural parameter

input size

$$f(k)p(n)$$

PROBLEM DEFINITIONS & MOTIVATION

The Motivation

The Motivation

*Comparative genomics.**

The Motivation

*Comparative genomics.**

[*] Zheng, Swenson, Lyons, and Sankoff, WABI 2011:
OMG! orthologs in multiple genomes - competing graph-theoretical formulations

The Motivation

*Comparative genomics.**

“...A fundamental task is to infer the relations between genes in different genomes and, more precisely, to infer which genes are orthologous.

[*] Zheng, Swenson, Lyons, and Sankoff, WABI 2011:
OMG! orthologs in multiple genomes - competing graph-theoretical formulations

The Motivation

*Comparative genomics.**

“...A fundamental task is to infer the relations between genes in different genomes and, more precisely, to infer which genes are orthologous.

Genes are orthologous when they originate via a speciation event from a gene of an ancestral genome.

[*] Zheng, Swenson, Lyons, and Sankoff, WABI 2011:
OMG! orthologs in multiple genomes - competing graph-theoretical formulations

The Motivation

*Comparative genomics.**

“...A fundamental task is to infer the relations between genes in different genomes and, more precisely, to infer which genes are orthologous.

Genes are orthologous when they originate via a speciation event from a gene of an ancestral genome.

A graph approach aiming to identify disjoint orthology sets, where each of such sets corresponds to a colorful component in the given graph and the colorful components associated with the orthology sets are disjoint.”

[*] Zheng, Swenson, Lyons, and Sankoff, WABI 2011:
OMG! orthologs in multiple genomes – competing graph-theoretical formulations

PROBLEM DEFINITIONS & MOTIVATION

Definitions

Definitions

The Official Problem Statements

Definitions

The Official Problem Statements

MAXIMIZING EDGES IN THE TRANSITIVE CLOSURE (MEC)

Input: A vertex-colored graph $(G = (V, E); c : V \rightarrow [\ell])$ and $k \in \mathbb{N}$.

Question: Does there exist a colorful subgraph H of G that has at least k edges in the transitive closure of its connected components?

Definitions

The Official Problem Statements

MAXIMIZING EDGES IN THE TRANSITIVE CLOSURE (MEC)

Input: A vertex-colored graph $(G = (V, E); c : V \rightarrow [\ell])$ and $k \in \mathbb{N}$.

Question: Does there exist a colorful subgraph H of G that has at least k edges in the transitive closure of its connected components?

COLORFUL COMPONENTS (CC)

Input: A vertex-colored graph $(G = (V, E); c : V \rightarrow [\ell])$ and $k \in \mathbb{N}$.

Question: Does there exist a subset $F \subseteq E$ of at most k edges such that every connected component in $G \setminus F$ is colorful?

Definitions

The Official Problem Statements

MAXIMIZING EDGES IN THE TRANSITIVE CLOSURE (MEC)

Input: A vertex-colored graph $(G = (V, E); c : V \rightarrow [\ell])$ and $k \in \mathbb{N}$.

Question: Does there exist a colorful subgraph H of G that has at least k edges in the transitive closure of its connected components?

COLORFUL COMPONENTS (CC)

Input: A vertex-colored graph $(G = (V, E); c : V \rightarrow [\ell])$ and $k \in \mathbb{N}$.

Question: Does there exist a subset $F \subseteq E$ of at most k edges such that every connected component in $G \setminus F$ is colorful?

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in the TC

Maximizing Edges in the TC

Literature Overview - [1] and [3]

Maximizing Edges in the TC

Literature Overview - [1] and [3]

NP-hard to approximate within $|V|^{1/3-\epsilon}$

Maximizing Edges in the TC

Literature Overview - [1] and [3]

NP-hard to approximate within $|V|^{1/3-\epsilon}$

APX-hard even for 3 colors

Maximizing Edges in the TC

Literature Overview - [1] and [3]

NP-hard to approximate within $|V|^{1/3-\epsilon}$

APX-hard even for 3 colors

$\sqrt{2} \cdot OPT$ -approximation algorithm

Maximizing Edges in the TC

Literature Overview - [1] and [3]

NP-hard to approximate within $|V|^{1/3-\epsilon}$

APX-hard even for 3 colors

$\sqrt{2} \cdot OPT$ -approximation algorithm

Exponential kernel in general,
quadratic kernel on trees

PROBLEM DEFINITIONS & MOTIVATION

Colorful Components

Colorful Components

Literature Overview - [2], [3], and others

Colorful Components

Literature Overview - [2], [3], and others

A special case of multi-multiway cut

Colorful Components

Literature Overview - [2], [3], and others

A special case of multi-multiway cut

NP-hard: 3 or more colors, on trees of diameter four, on graphs which are constant-away from a collection of paths

Colorful Components

Literature Overview - [2], [3], and others

A special case of multi-multiway cut

NP-hard: 3 or more colors, on trees of diameter four, on graphs which are constant-away from a collection of paths

Hard to approximate within $|V|^{1/2-\epsilon}$

Colorful Components

Literature Overview - [2], [3], and others

A special case of multi-multiway cut

NP-hard: 3 or more colors, on trees of diameter four, on graphs which are constant-away from a collection of paths

Hard to approximate within $|V|^{1/2-\epsilon}$

FPT algorithms, heuristics,
and approximation algorithms for special cases

TALK OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel^[1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel^[1]

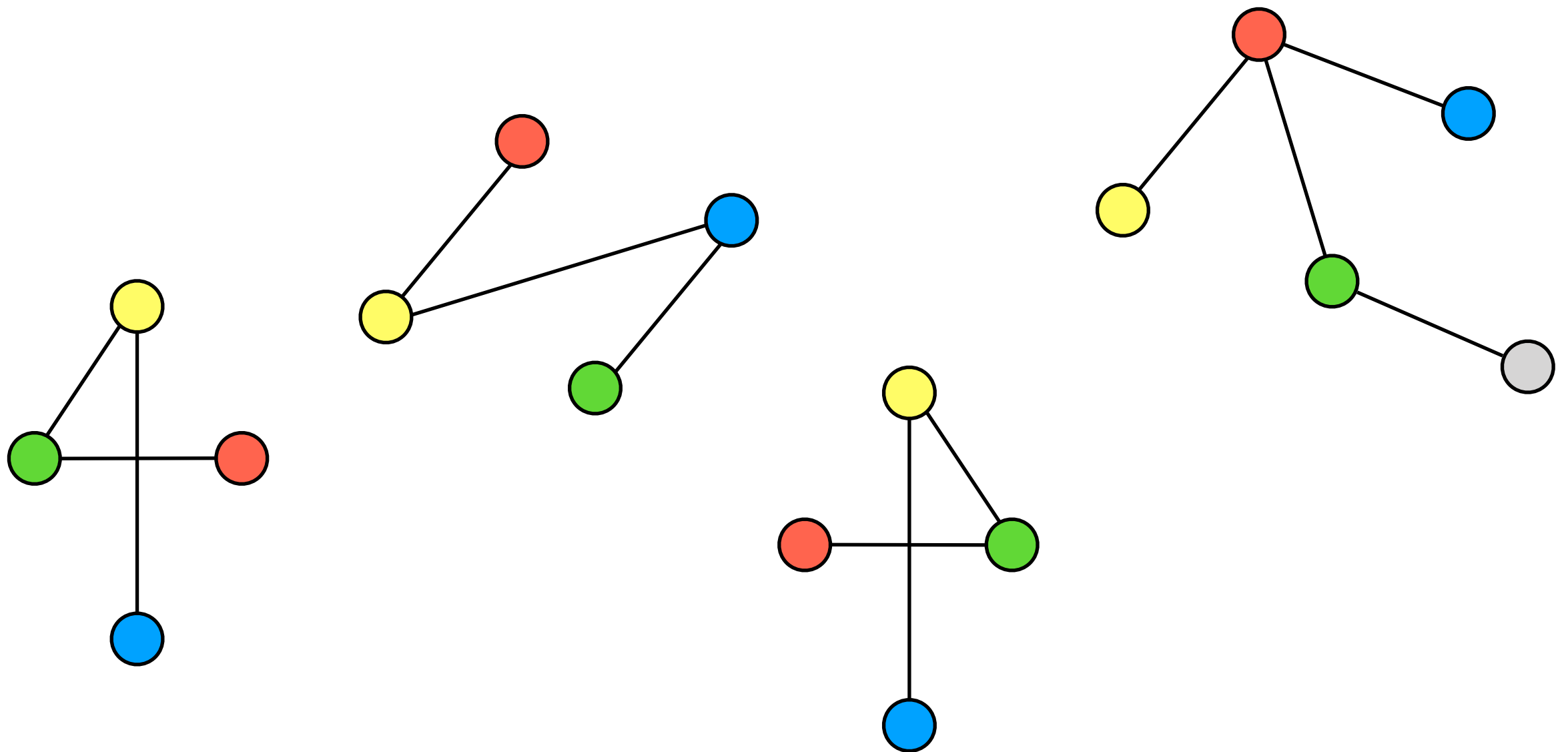
CC BY VERTEX COVER IS FPT

In contrast: para-NP-hard already for deletion sets to stars^[2] and paths^[1].

COLOR CODING FOR MEC

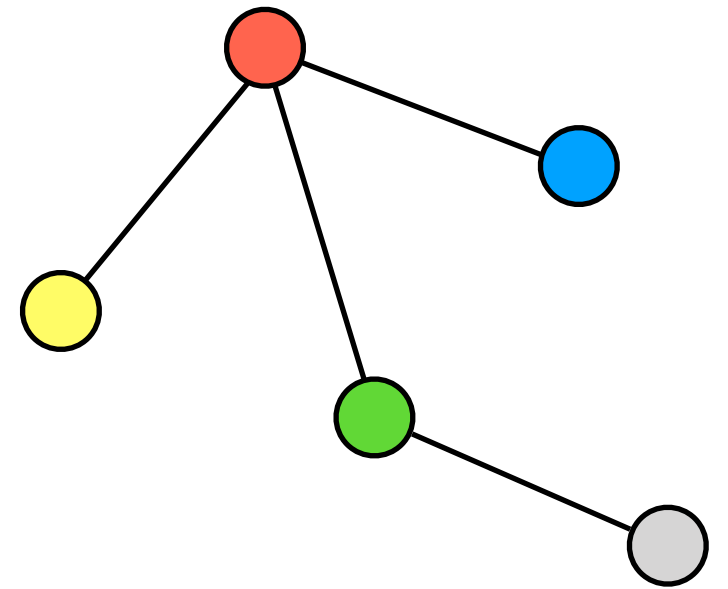
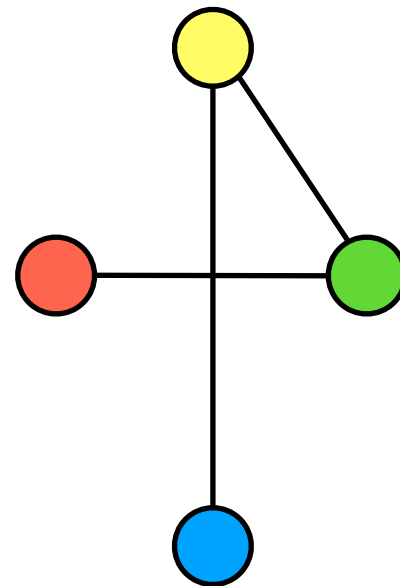
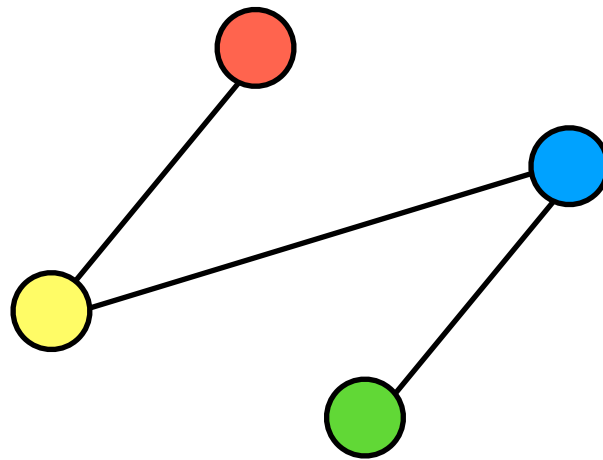
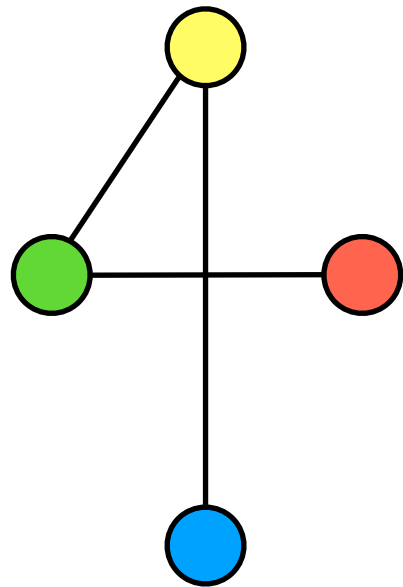
The signature of a solution with r components is a tuple of length r .

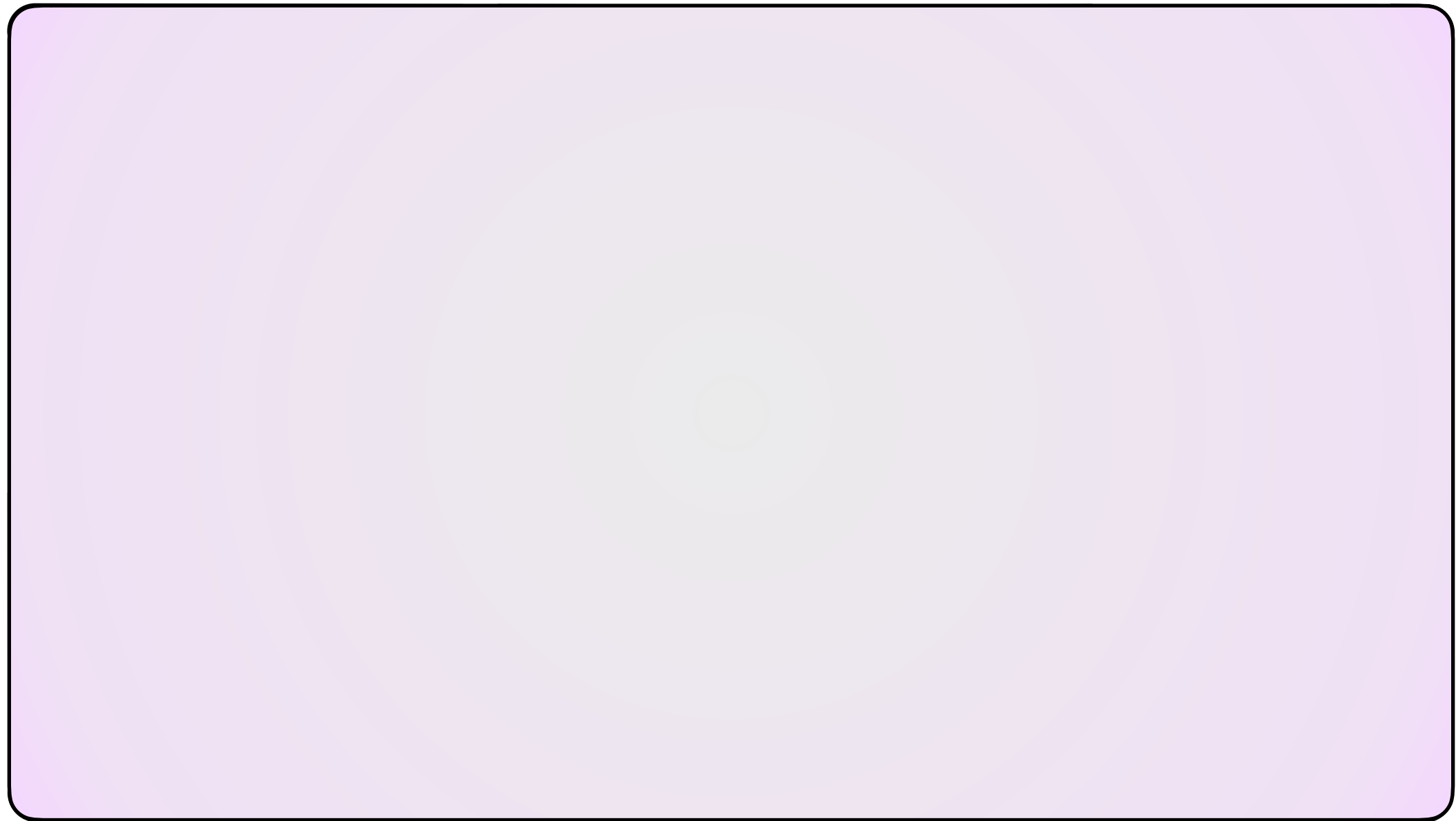
The signature of a solution with r components is a tuple of length r .



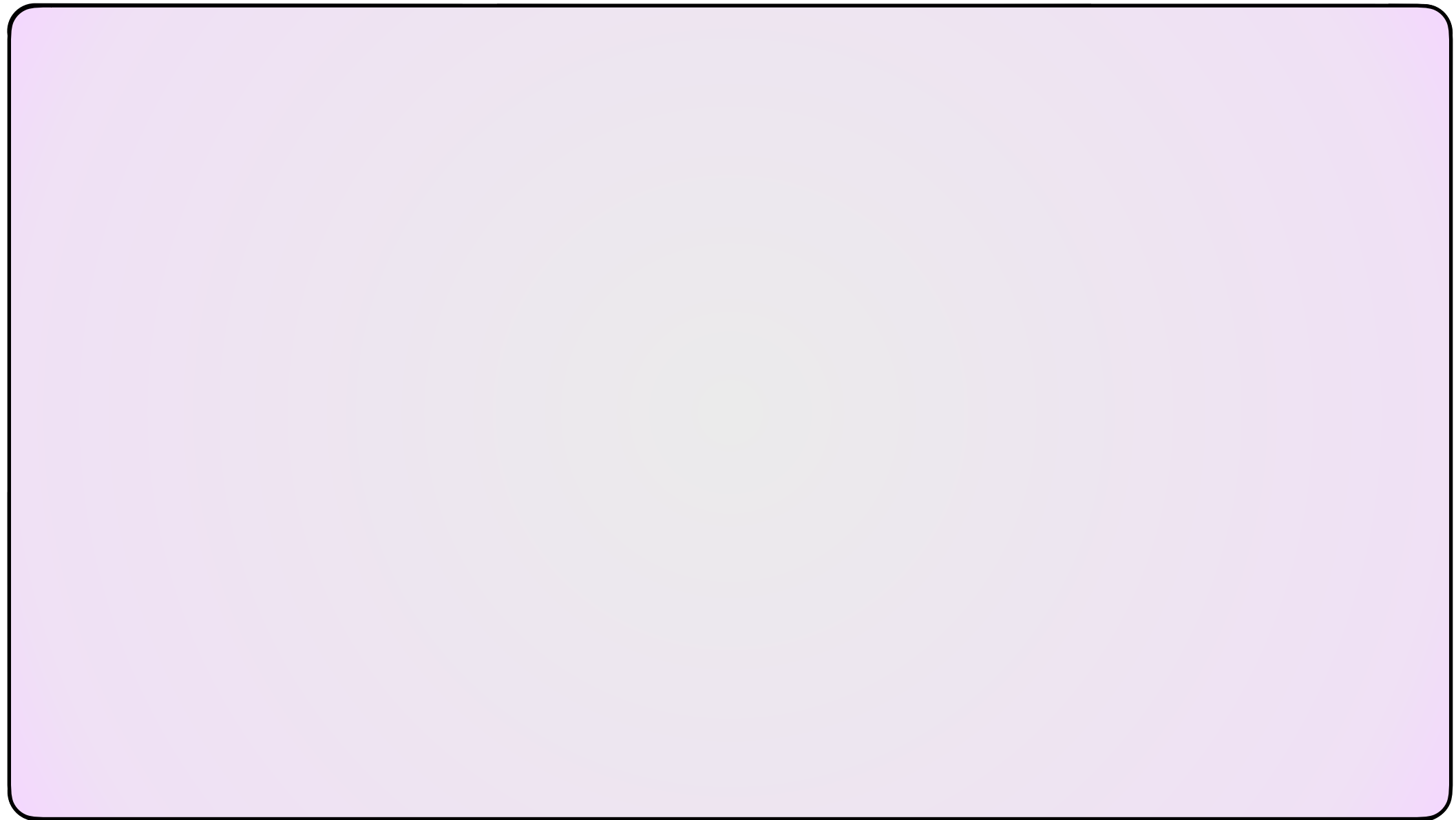
The signature of a solution with r components is a tuple of length r .

(6,6,6,10)

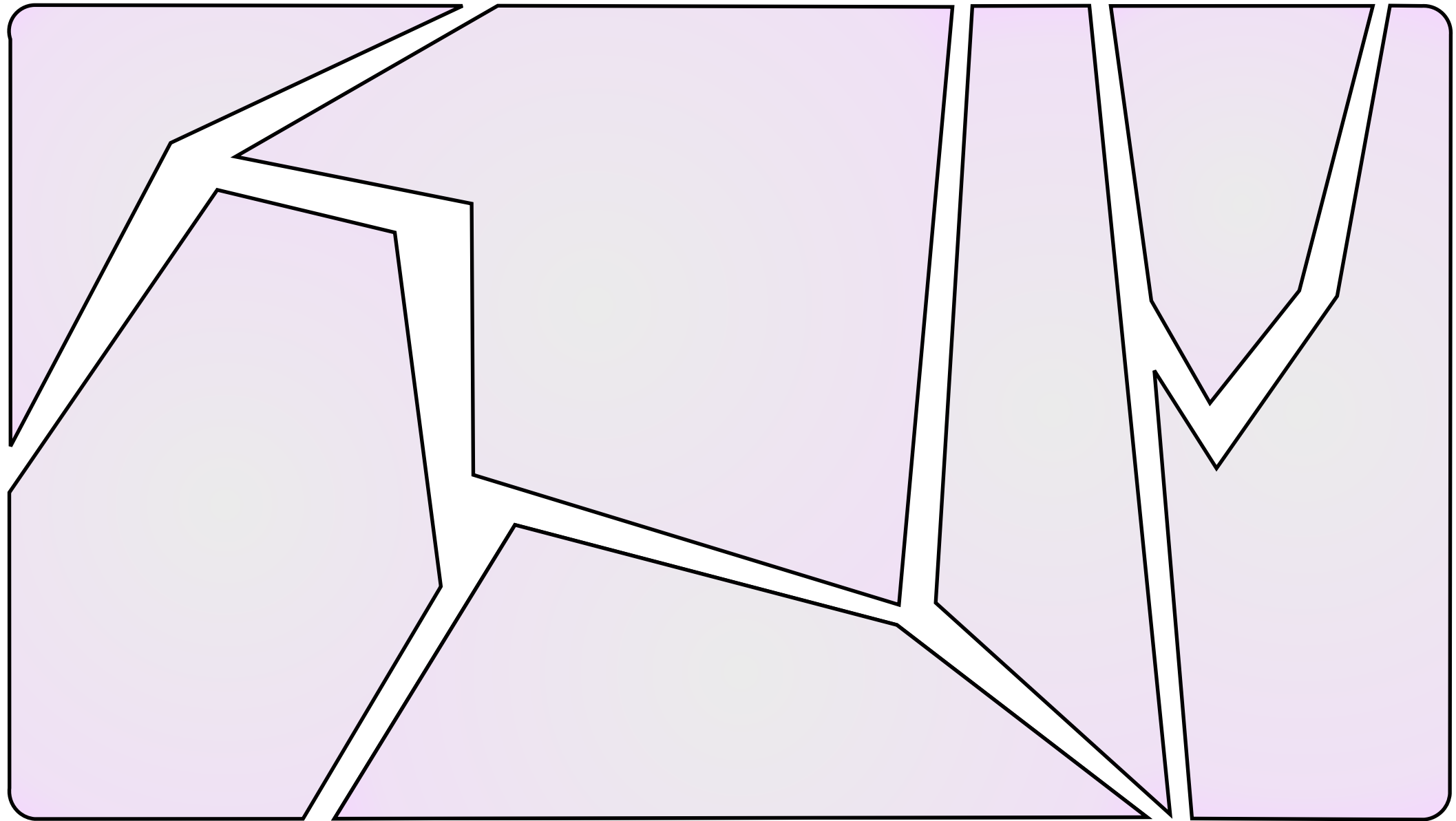




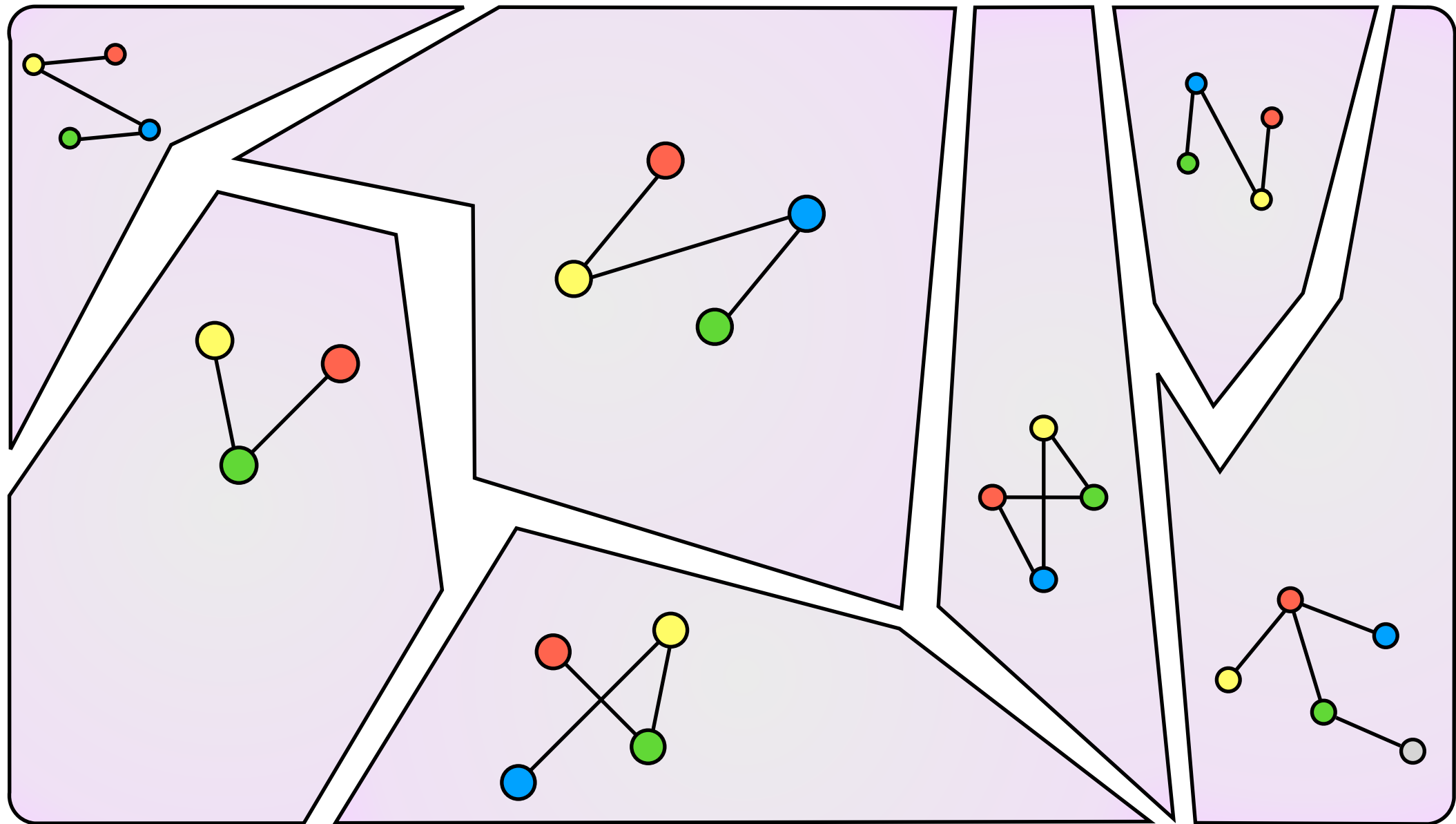
A random partition is good if all components of the solution go into different parts.



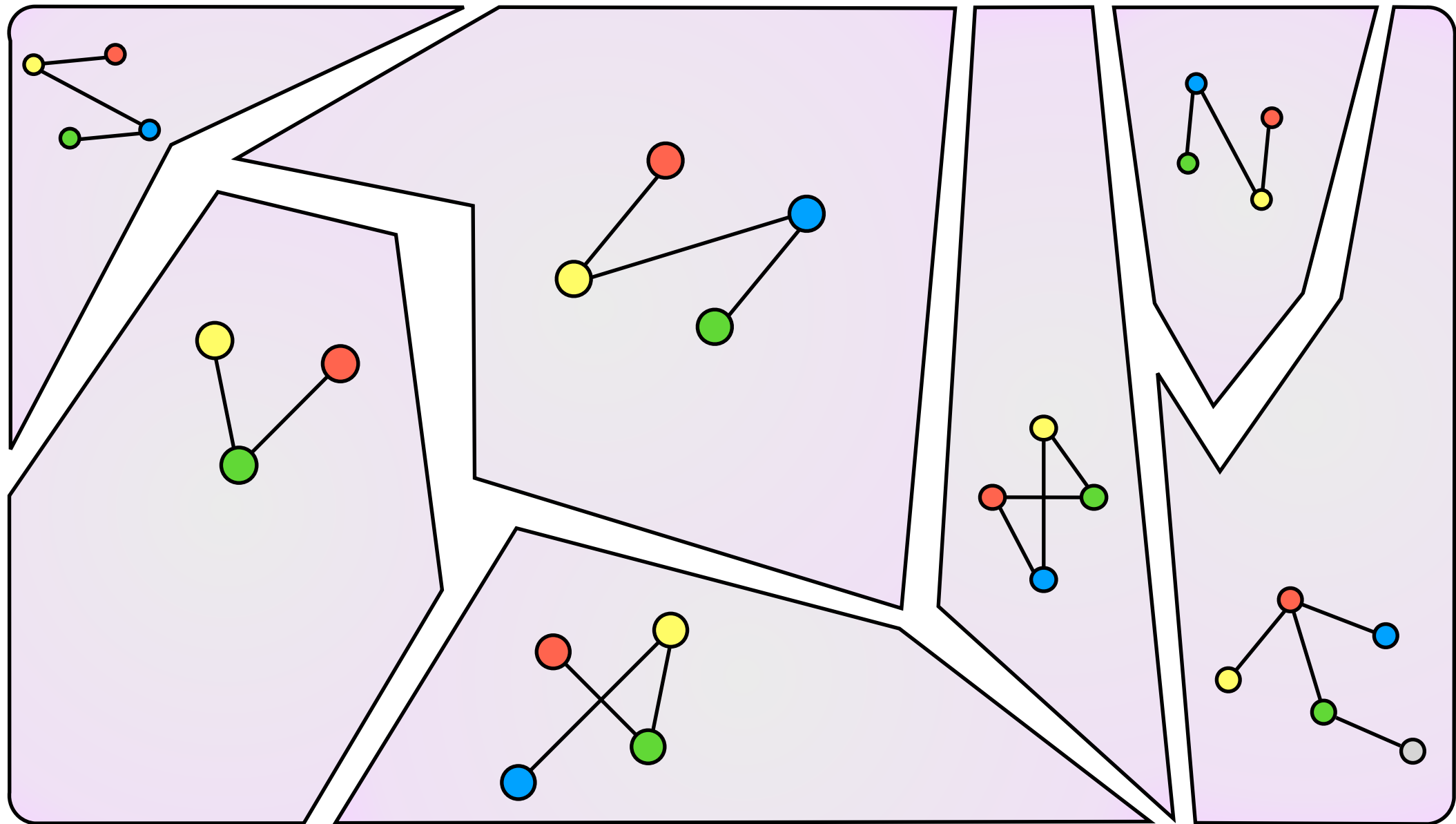
A random partition is good if all components of the solution go into different parts.



A random partition is good if all components of the solution go into different parts.



A random partition is good if all components of the solution go into different parts.



Solve each subpart by invoking **COLORFUL GRAPH MOTIF** as a subroutine.

TALK OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel^[1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel^[1]

CC BY VERTEX COVER IS FPT

In contrast: para-NP-hard already for deletion sets to stars^[2] and paths^[1].

Kernelization

Kernelization



Kernelization

 (x, k)  (x', k')

FAST

SMALL

SANE

Kernelization

(x, k) $\longrightarrow$ (x', k')

$|x|^{O(1)}$ time

SMALL

SANE

Kernelization

(x, k) $\longrightarrow$ (x', k')

$|x|^{O(1)}$ time

$|x'| = f(k)$ and $k \leq k'$

SANE

Kernelization

(x, k) $\longrightarrow$ (x', k')

$|x|^{O(1)}$ **time**

$|x'| = f(k)$ **and** $k \leq k'$

$(x, k) \equiv (x', k')$

Kernelization

(x, k) $\longrightarrow$ (x', k')

$|x|^{O(1)}$ time

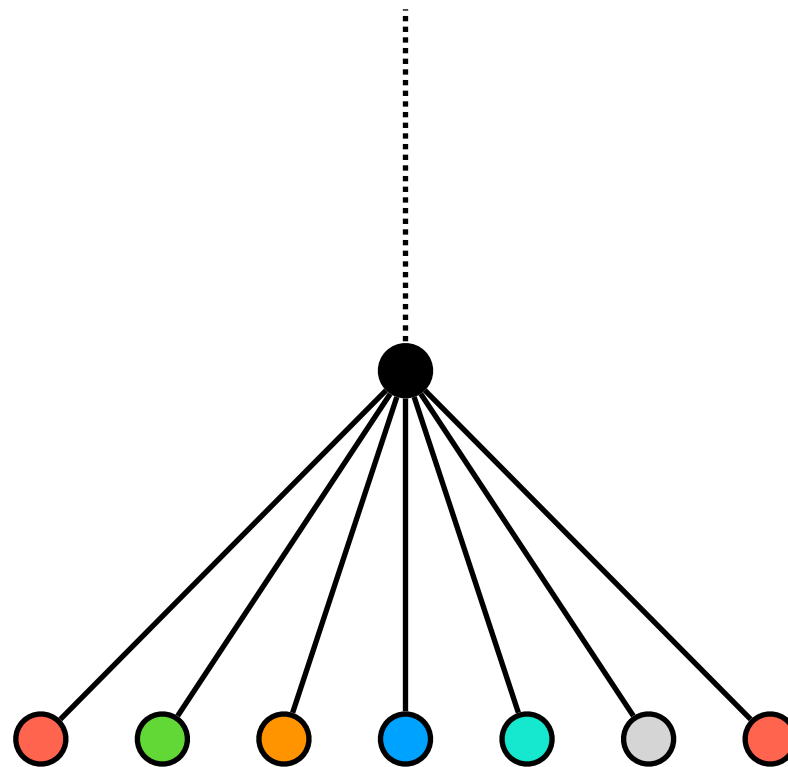
$|x'| = f(k)$ and $k \leq k'$

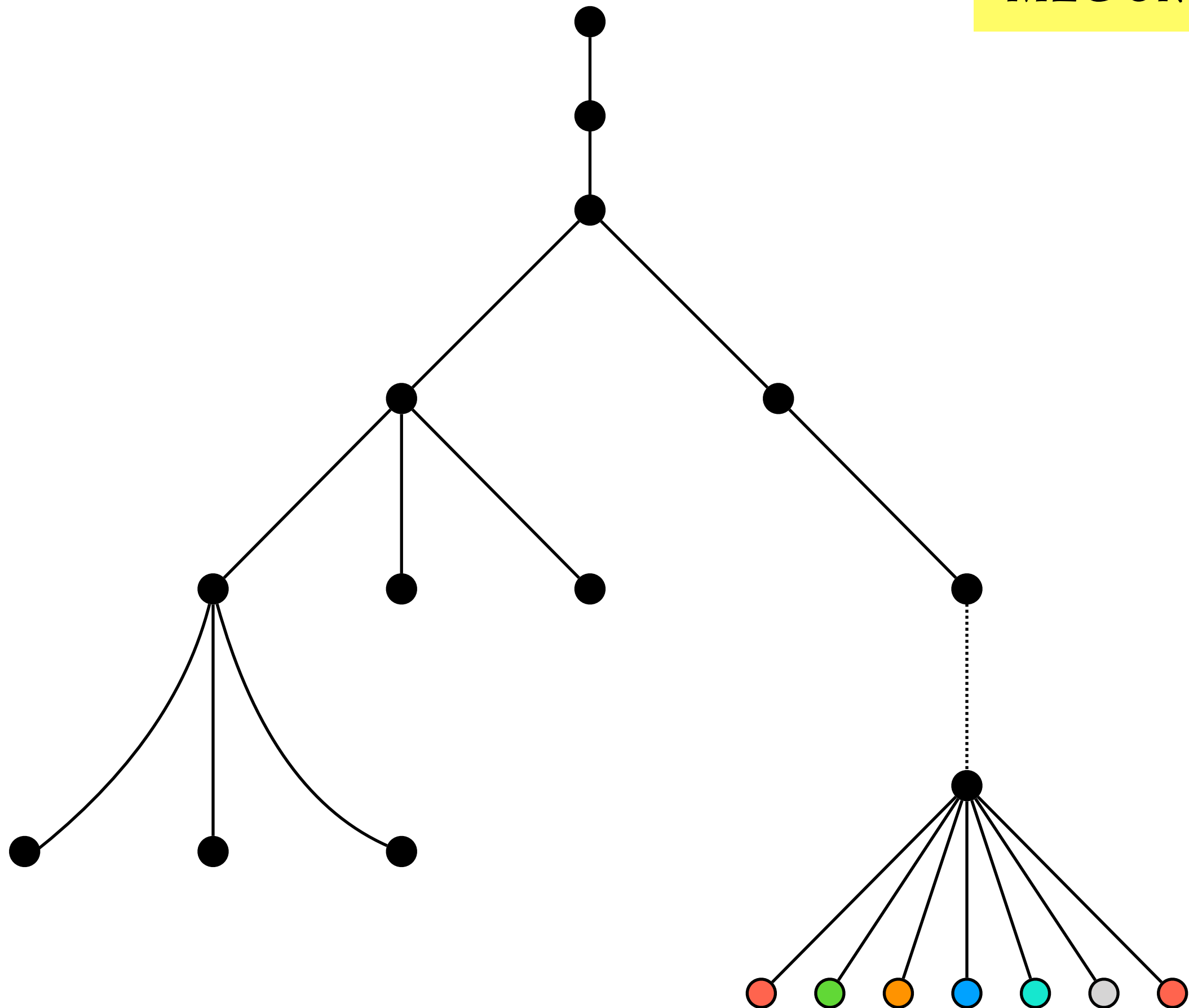
$(x, k) \equiv (x', k')$

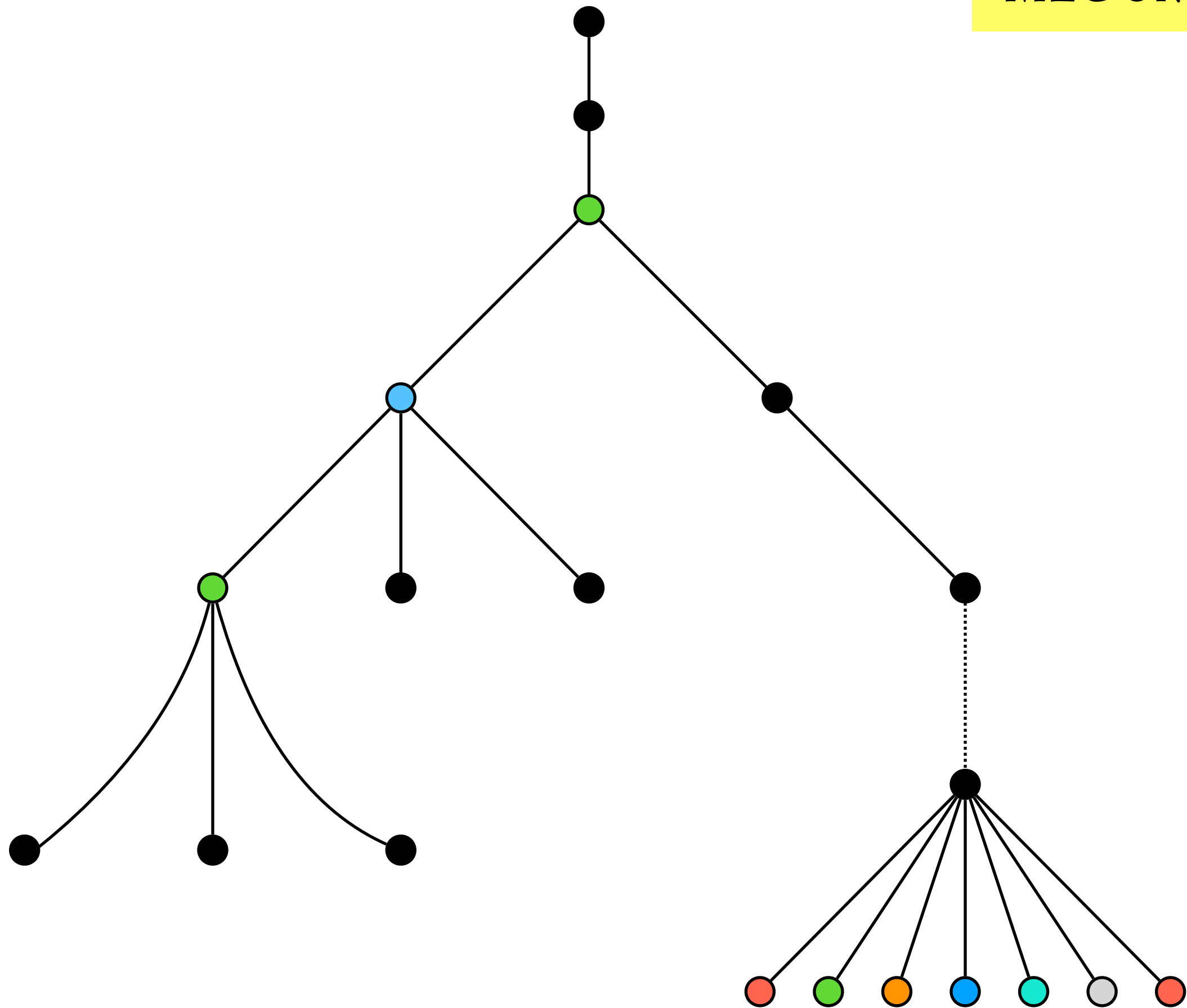
$f(k)$ -kernel

Start with the leaves

Start with the leaves

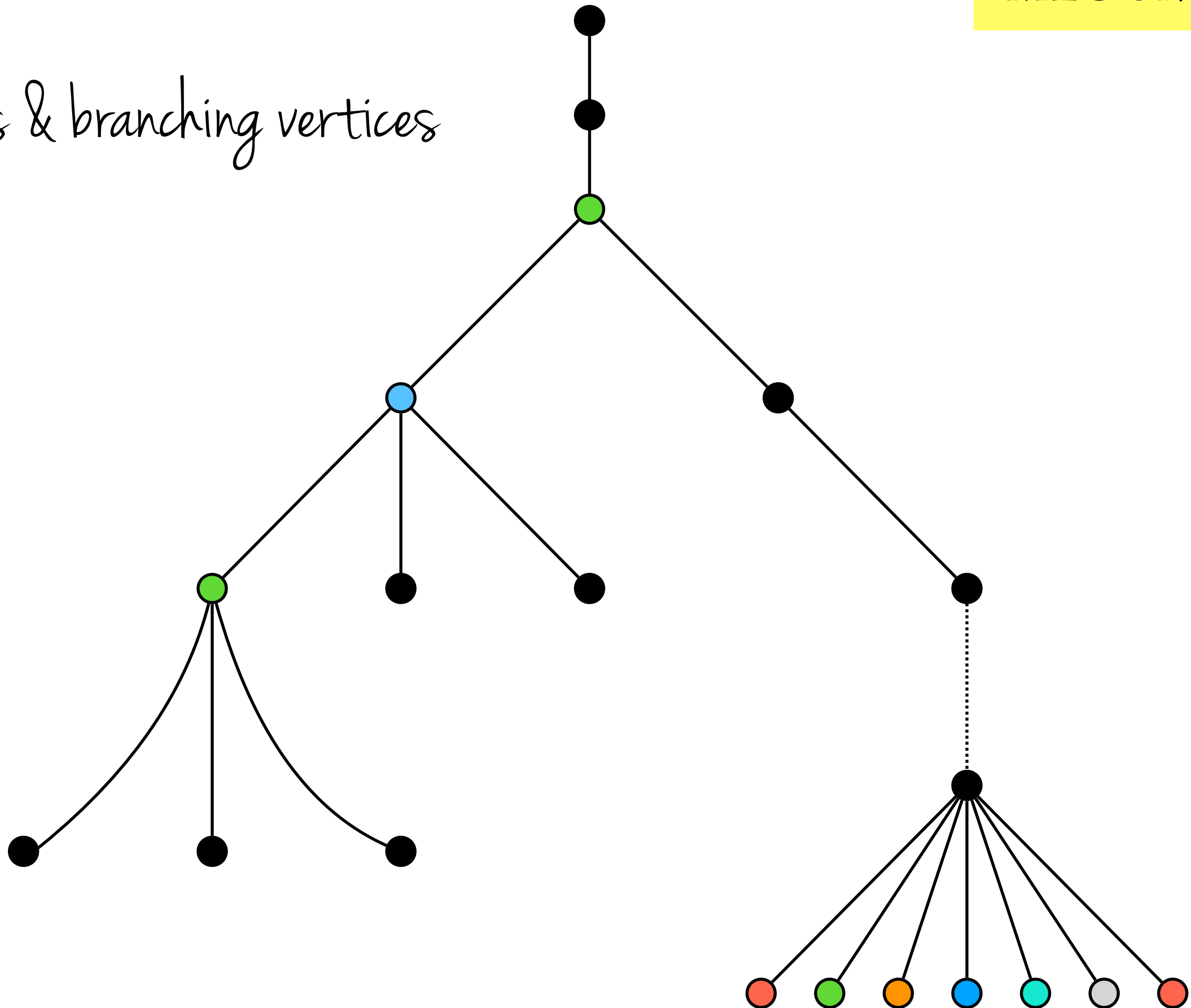




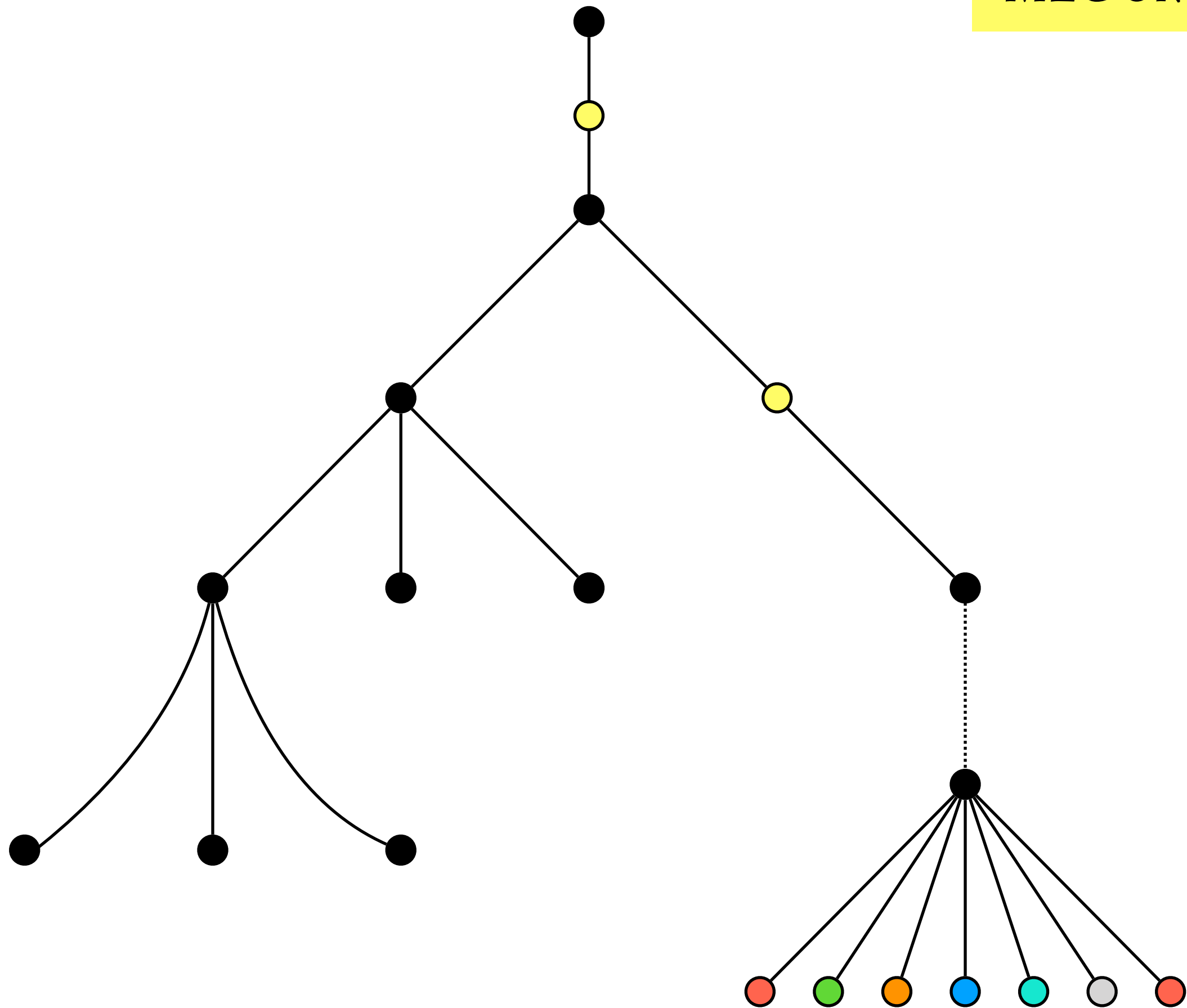


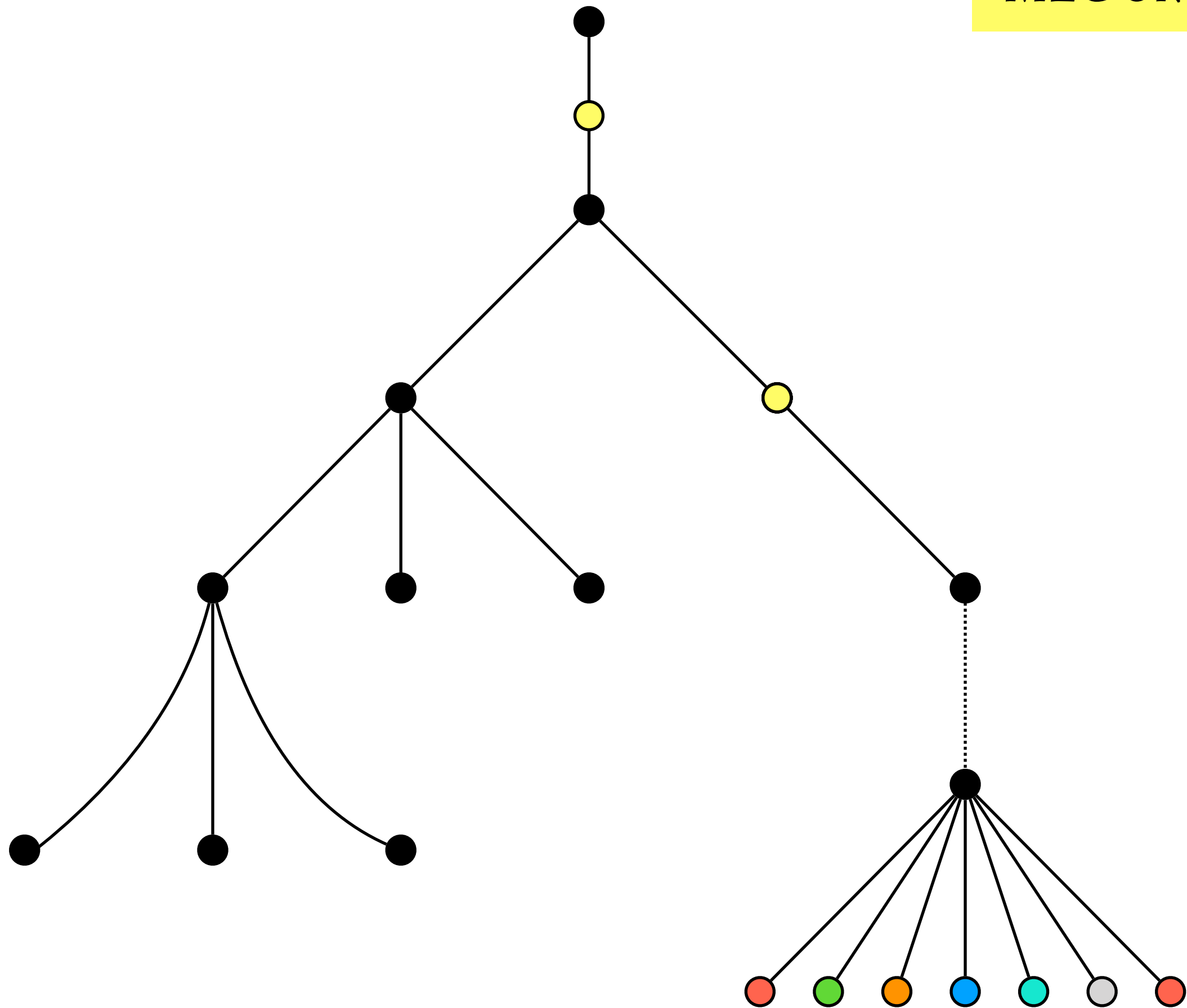
MEC ON TREES

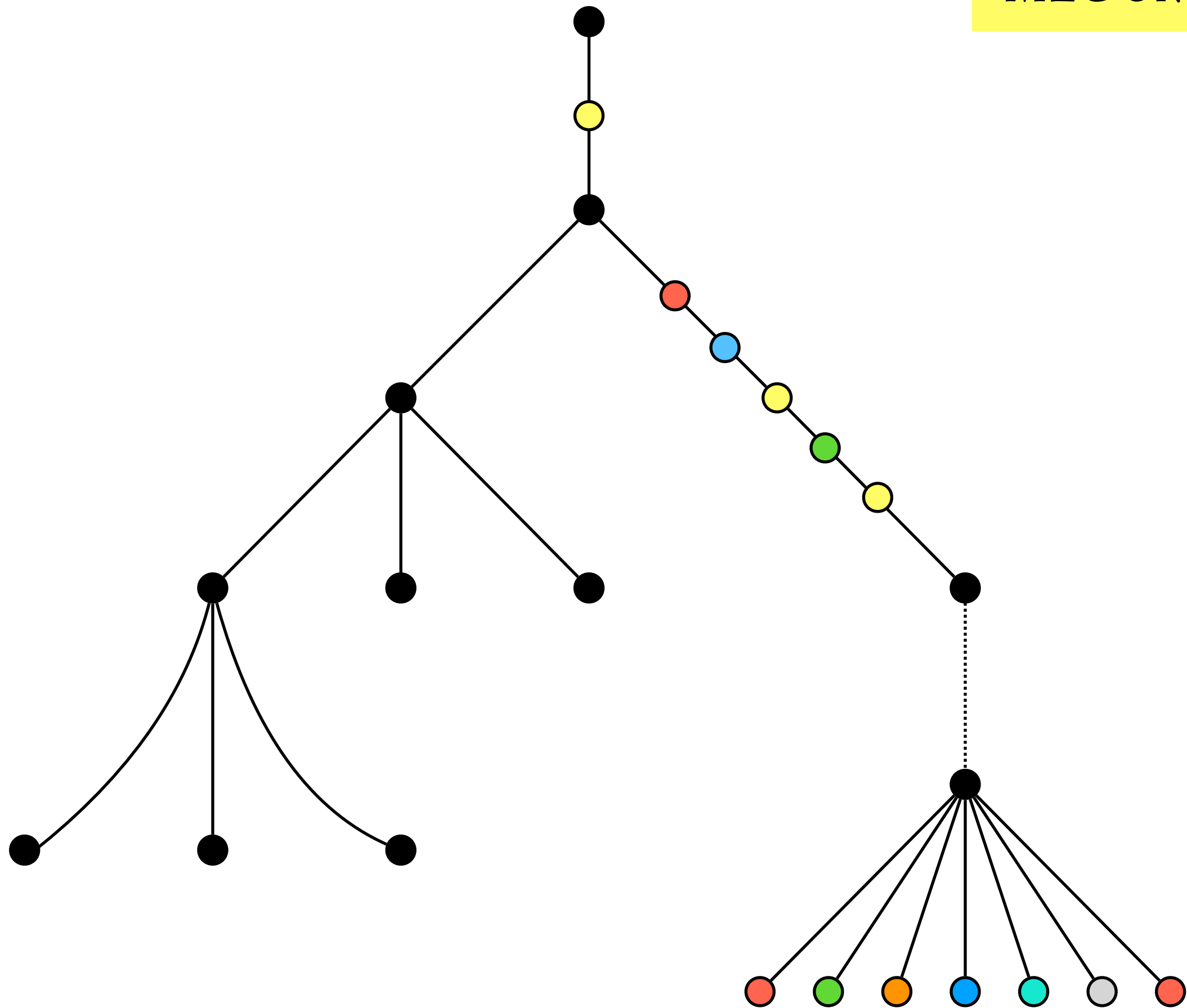
leaves & branching vertices



MEC ON TREES

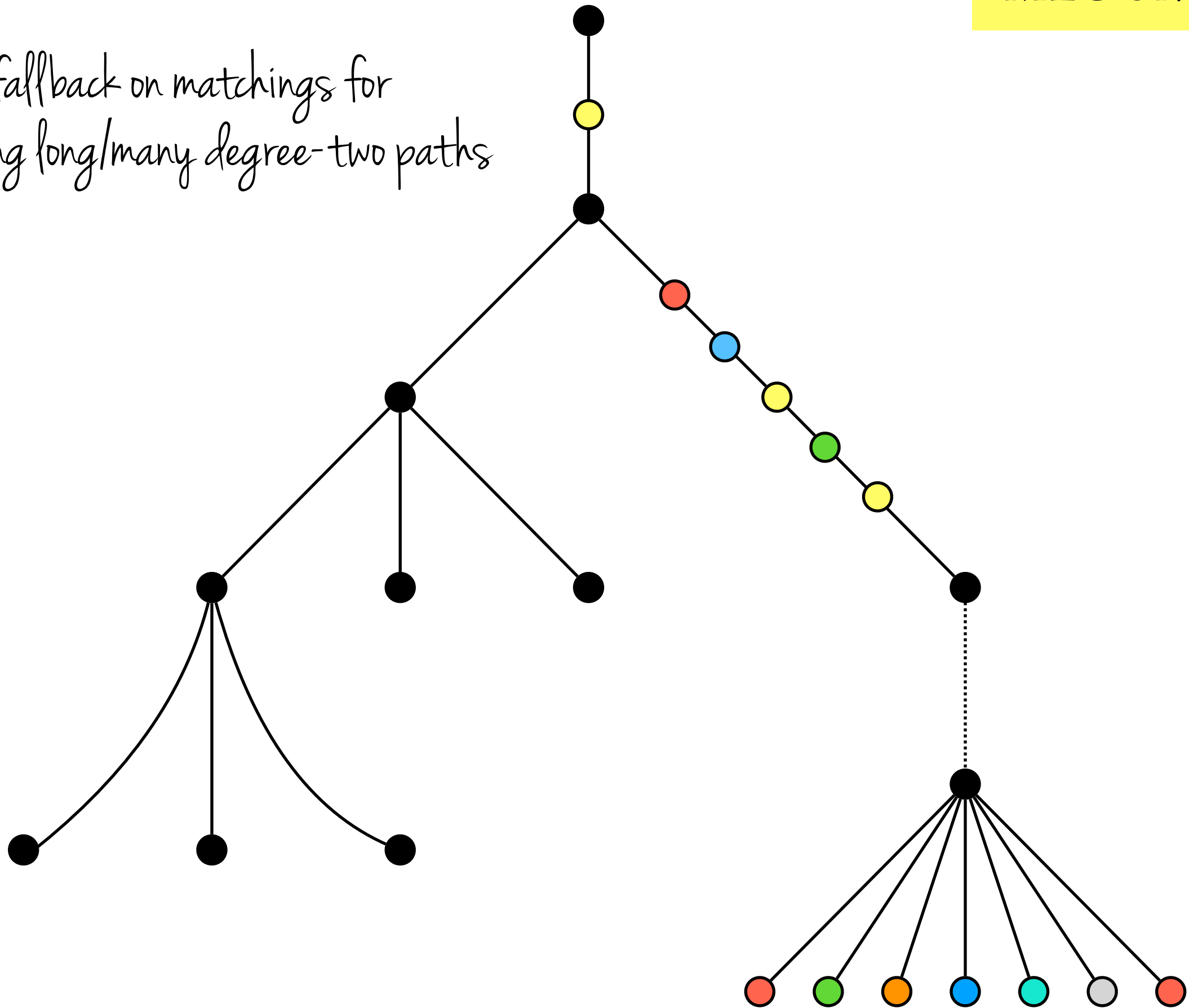






MEC ON TREES

fallback on matchings for
handling long/many degree-two paths



TALK OUTLINE

PROBLEM DEFINITIONS & MOTIVATION

Maximizing Edges in Transitive Closure, and Colorful Components

COLOR CODING FOR MEC

Previously: brute-force on an exponential kernel^[1]

LINEAR KERNEL FOR MEC ON TREES

Previously: a quadratic kernel^[1]

CC BY VERTEX COVER IS FPT

In contrast: para-NP-hard already for deletion sets to stars^[2] and paths^[1].

COLORFUL COMPONENTS | VERTEX COVER

The Setting

The Setting

G — A vertex-colored graph.

The Setting

G — A vertex-colored graph.

T — A vertex cover of size t .

The Setting

G — A vertex-colored graph.

T — A vertex cover of size t .

F — A subset of at most k edges.

The Setting

G — A vertex-colored graph.

T — A vertex cover of size t .

F — A subset of at most k edges.

$G \setminus F$ — A graph with colorful components.

What we have

G — A vertex-colored graph.

T — A vertex cover of size t .

F — A subset of at most k edges.

$G \setminus F$ — A graph with colorful components.

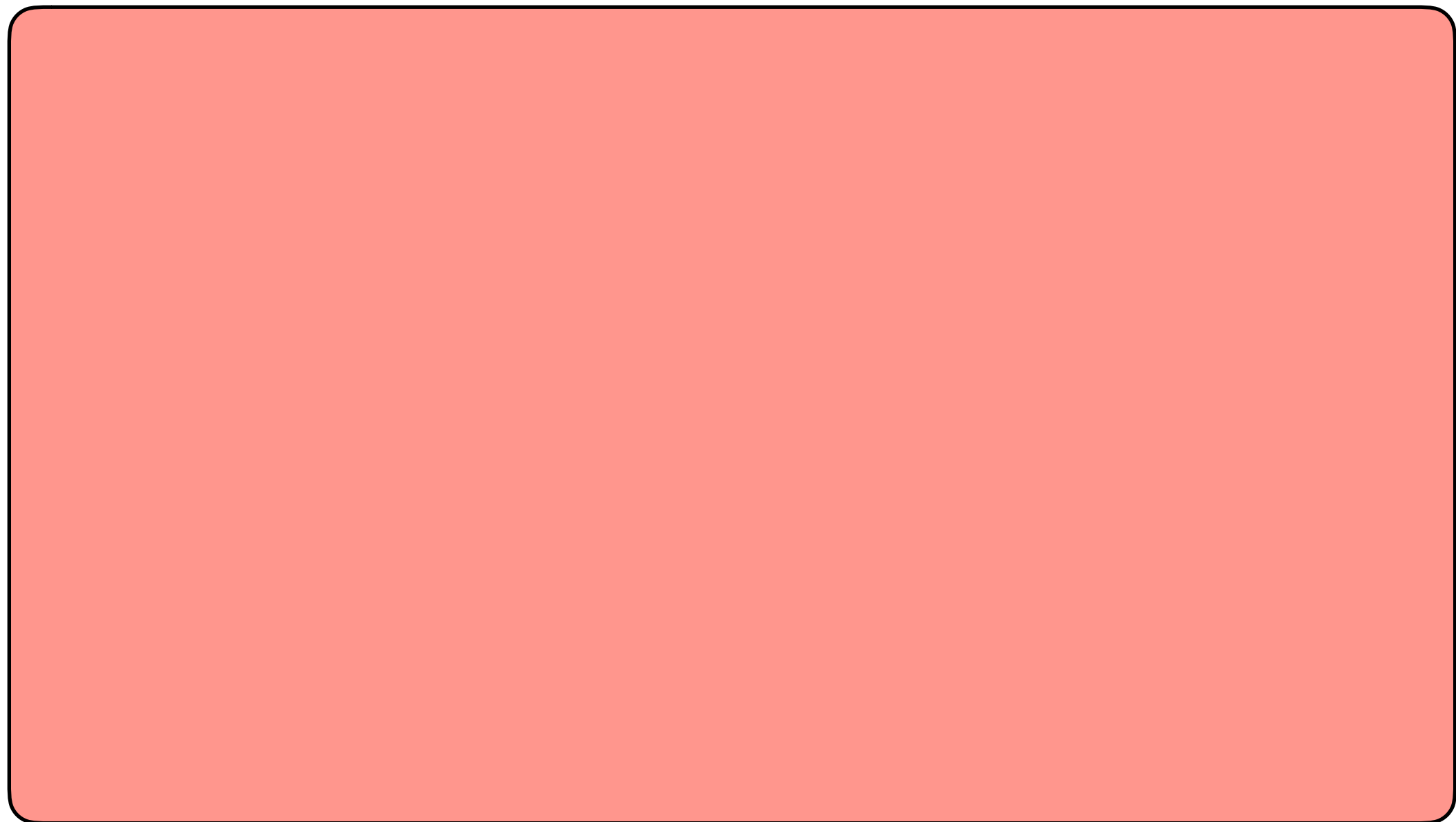
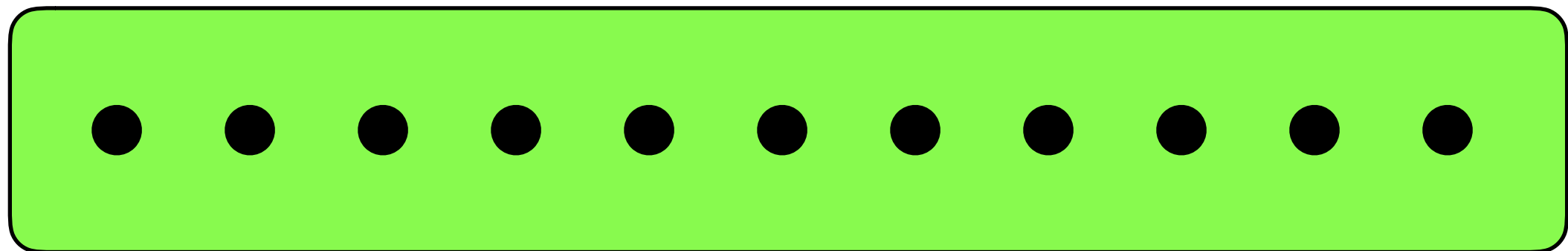
What we want

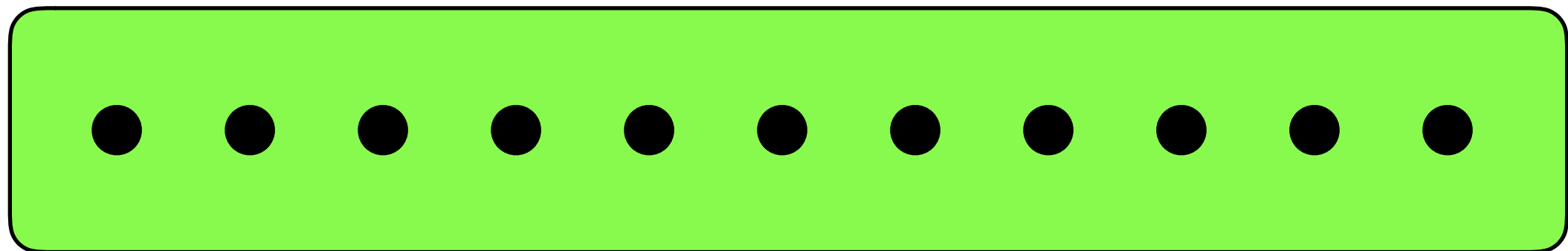
G — A vertex-colored graph.

T — A vertex cover of size t .

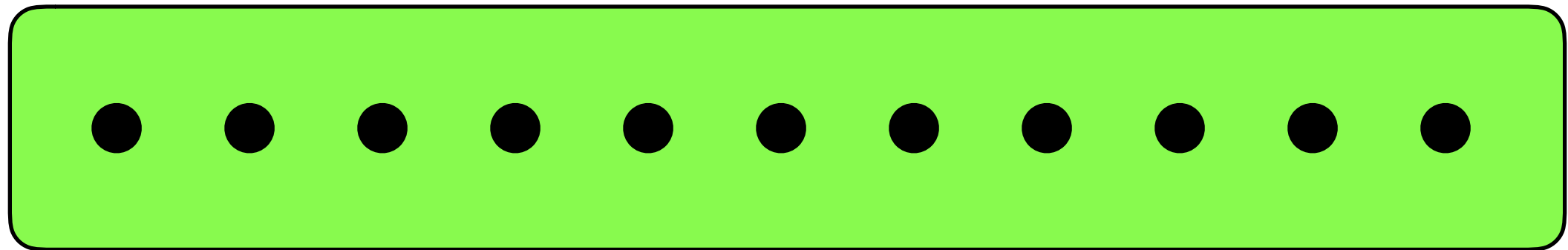
F — A subset of at most k edges.

$G \setminus F$ — A graph with colorful components.



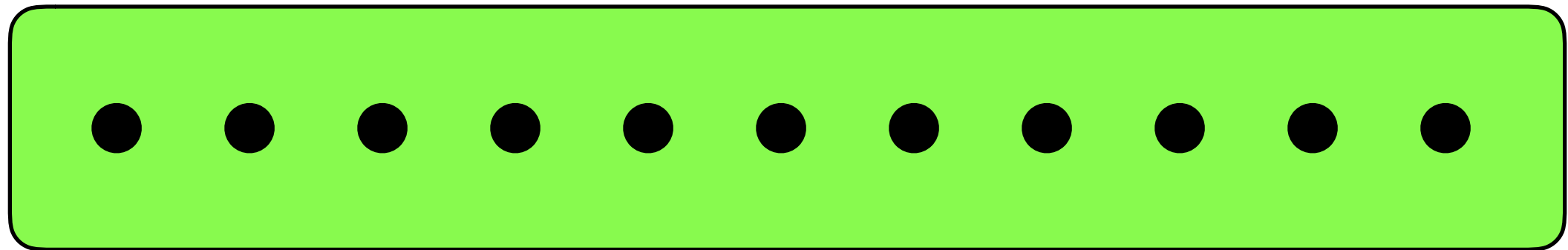


The first claim.



The first claim.

#(non-singleton components in $G \setminus F$) is at most t



The size of the
Vertex Cover

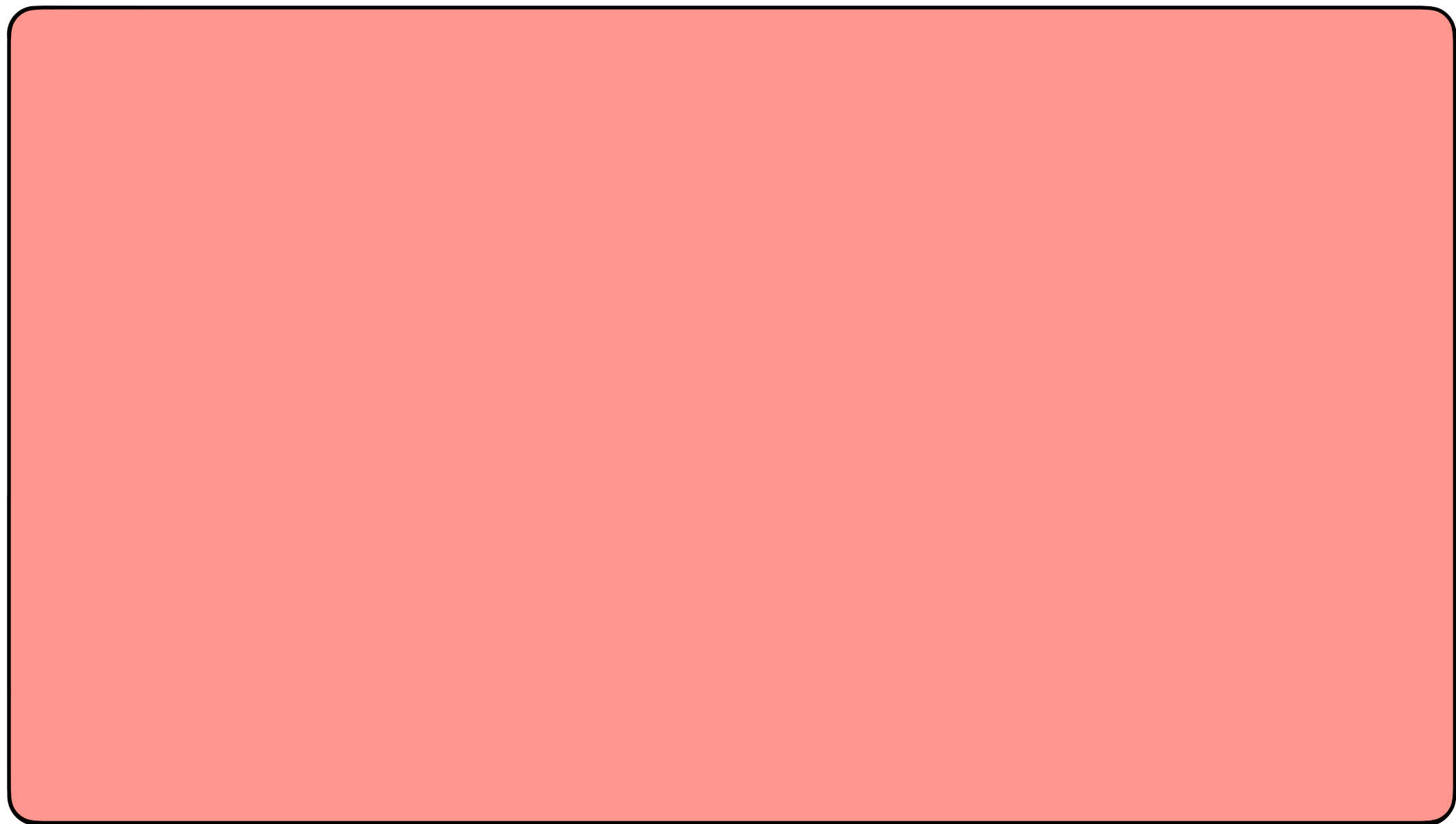
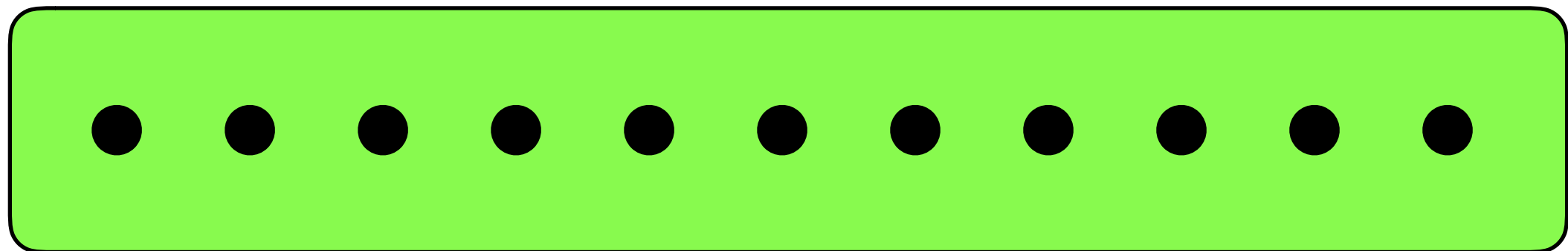


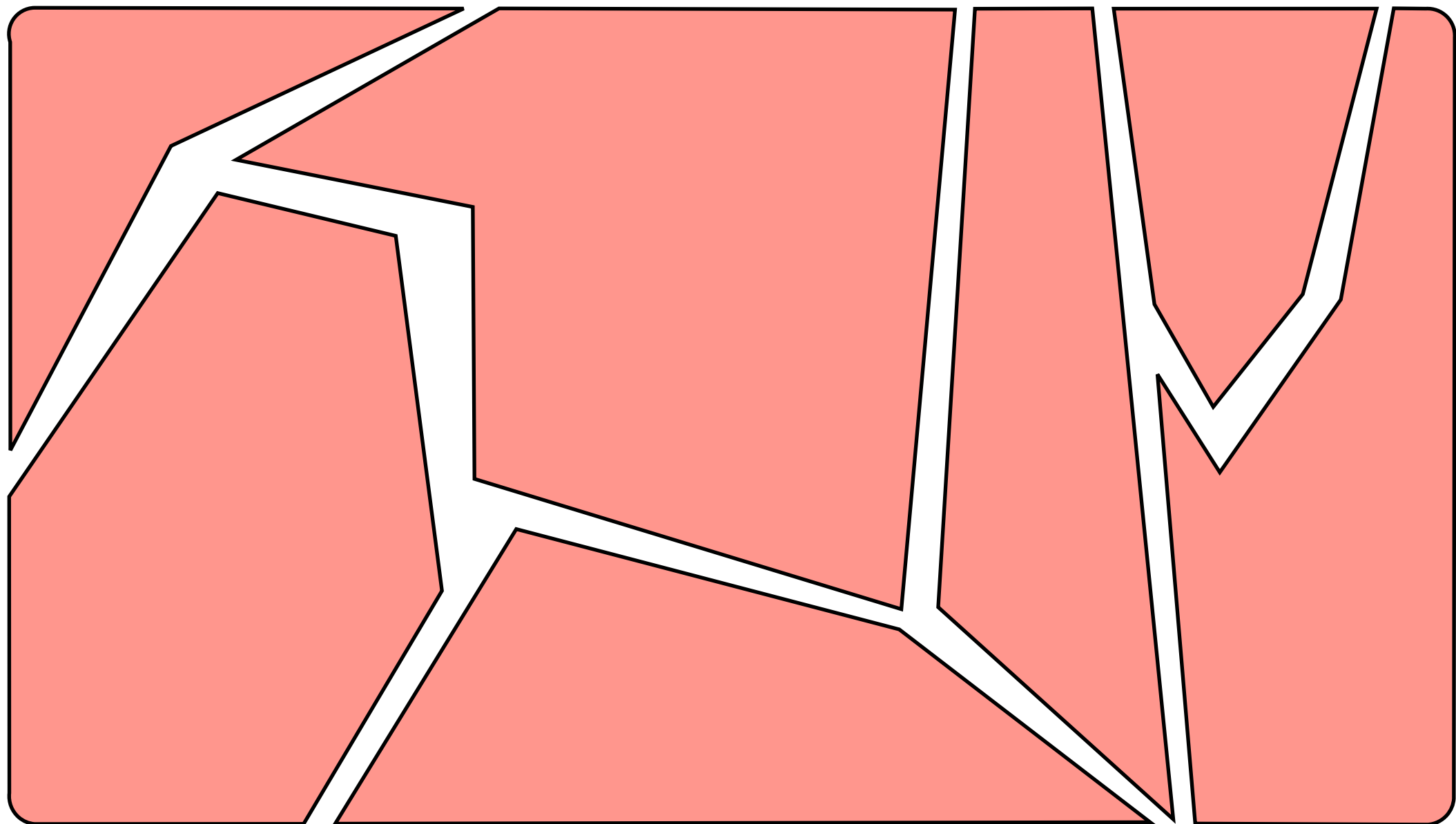
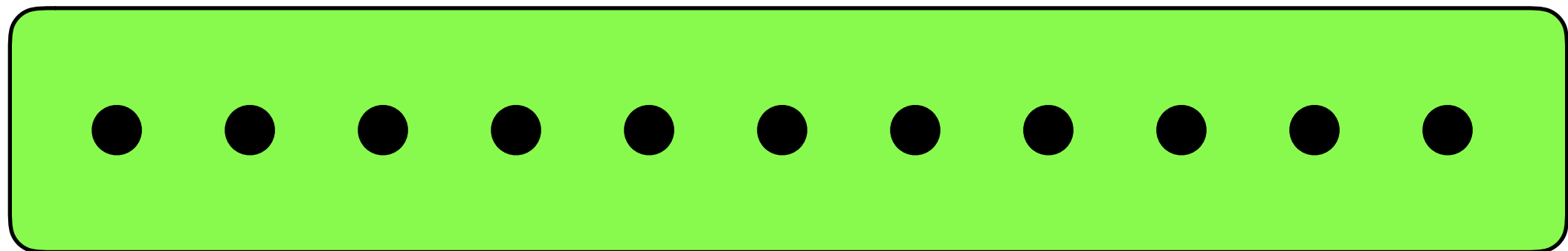
The first claim.

#(non-singleton components in $G \setminus F$) is at most t

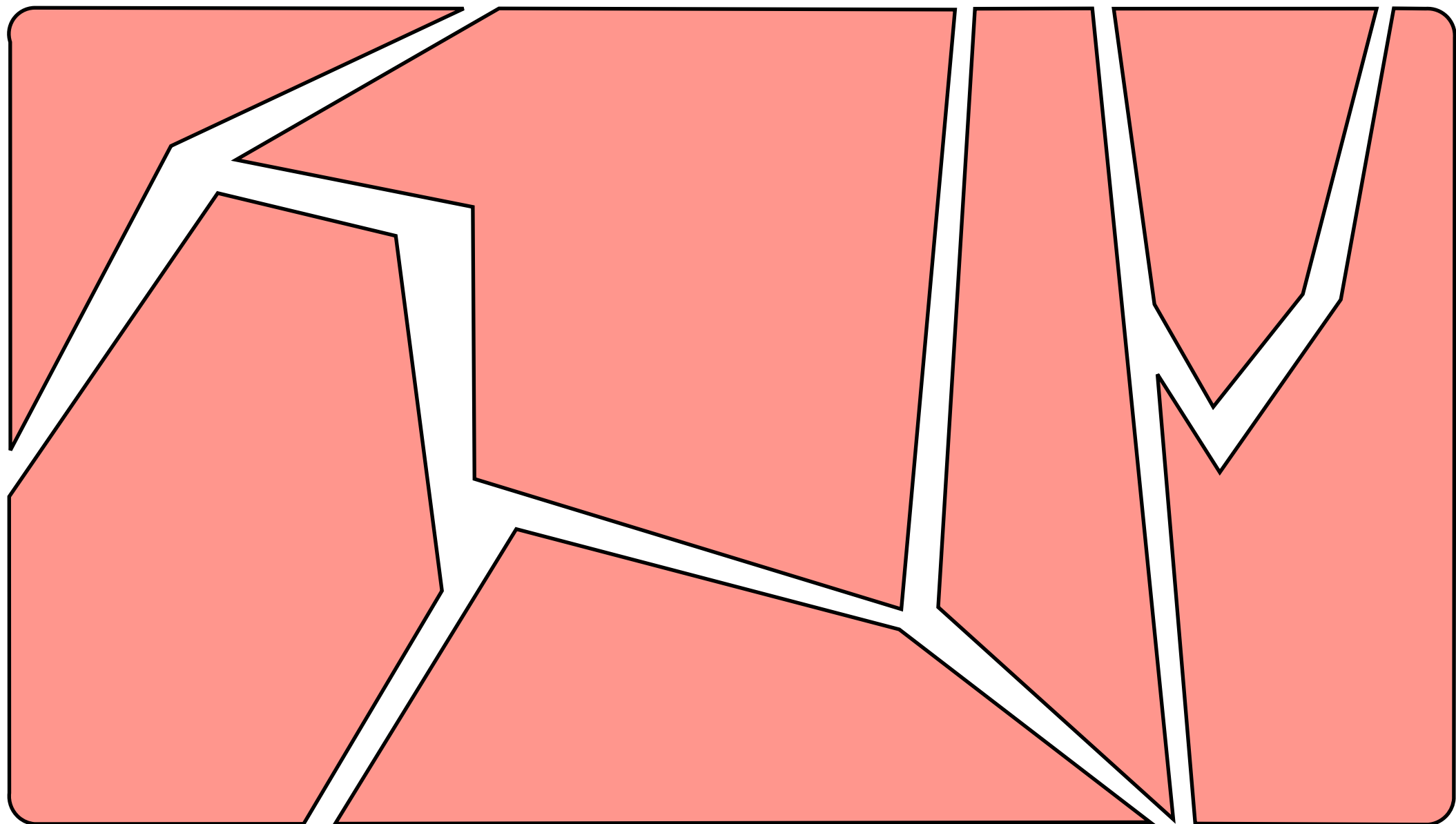
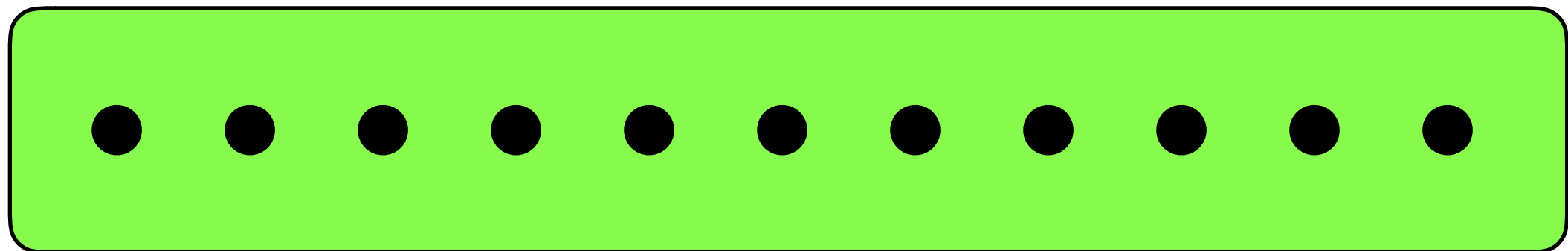
A solution

A hand-drawn arrow pointing upwards from the text "A solution" to the variable t in the claim.

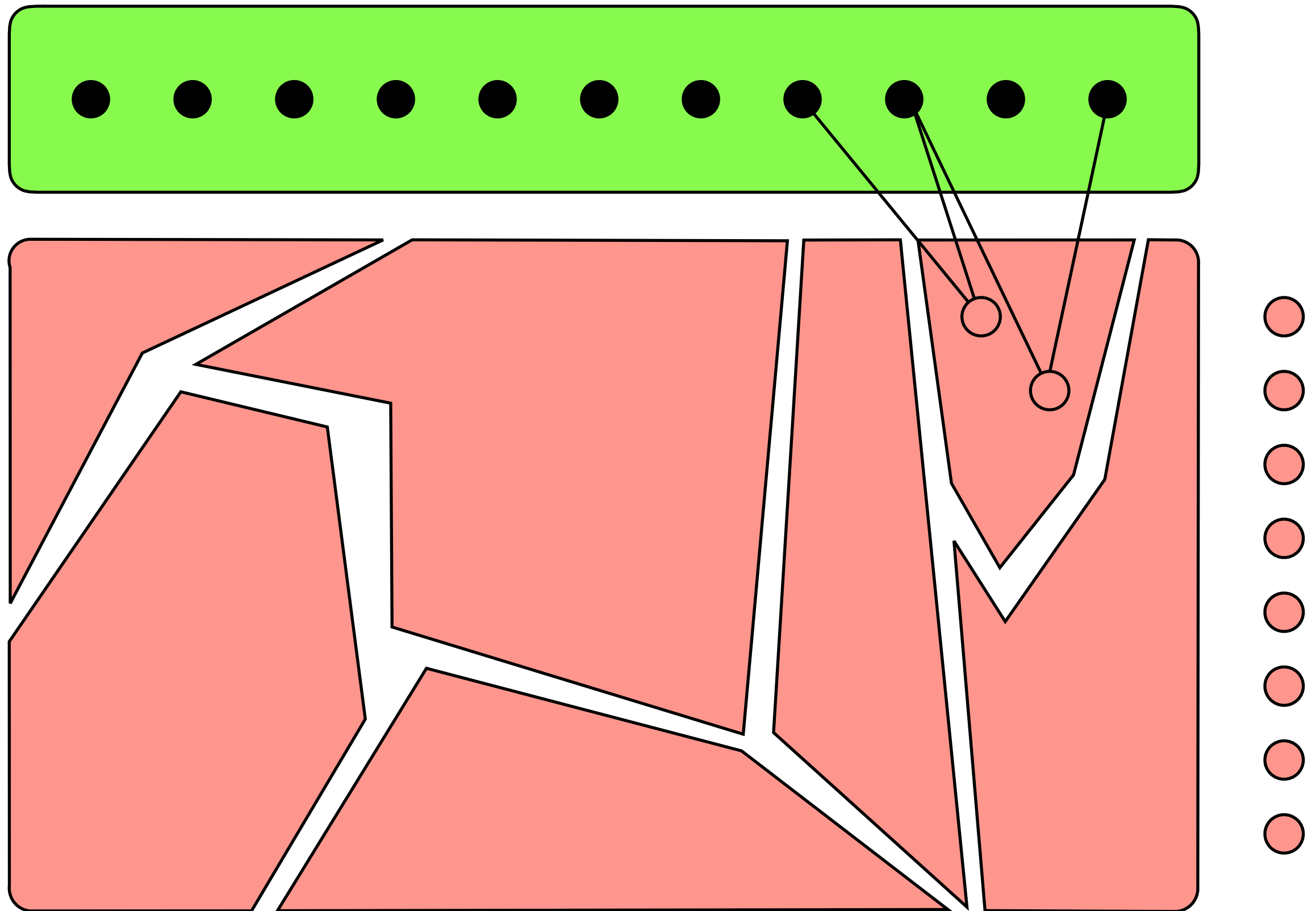




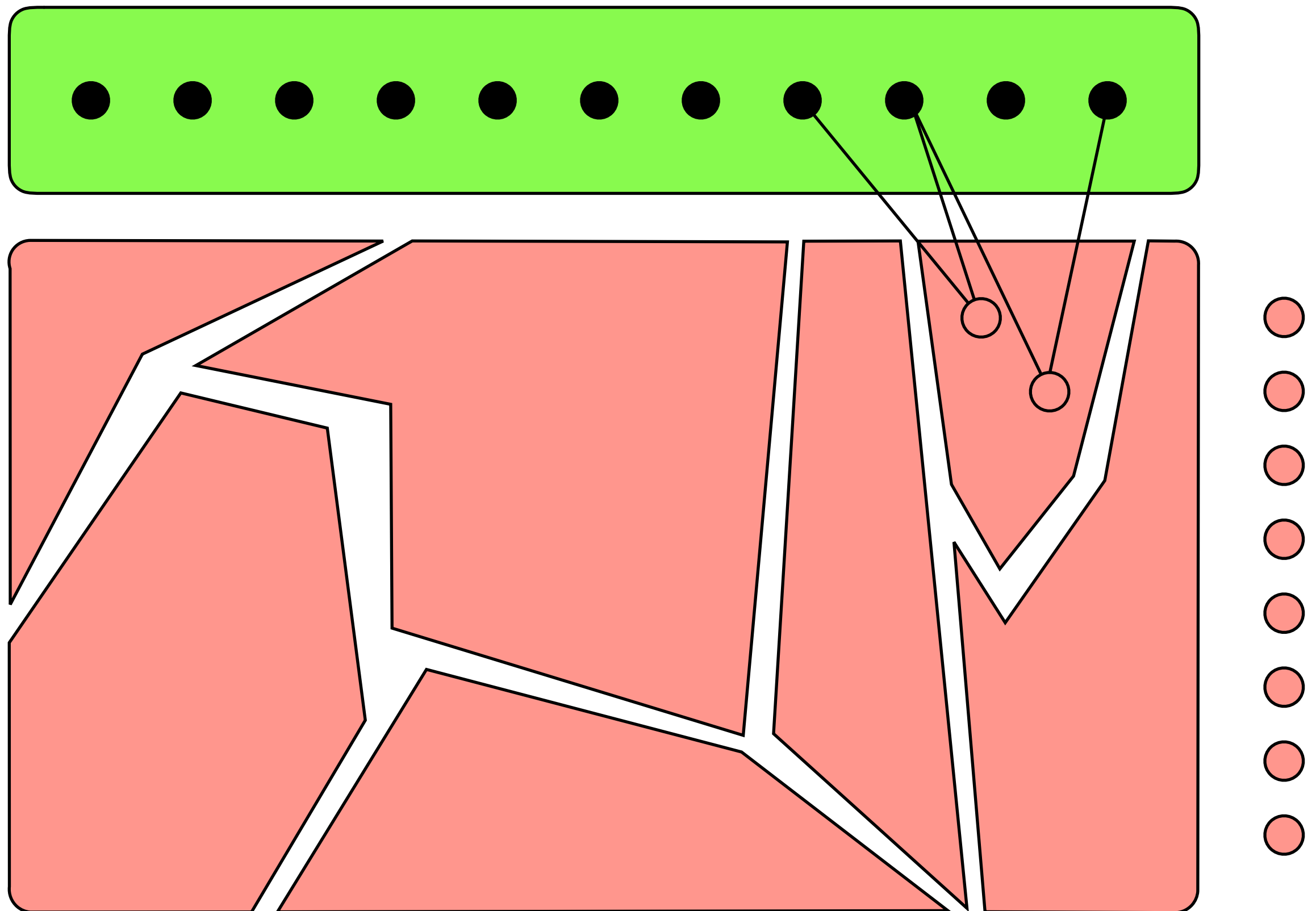
COLORFUL COMPONENTS | VERTEX COVER



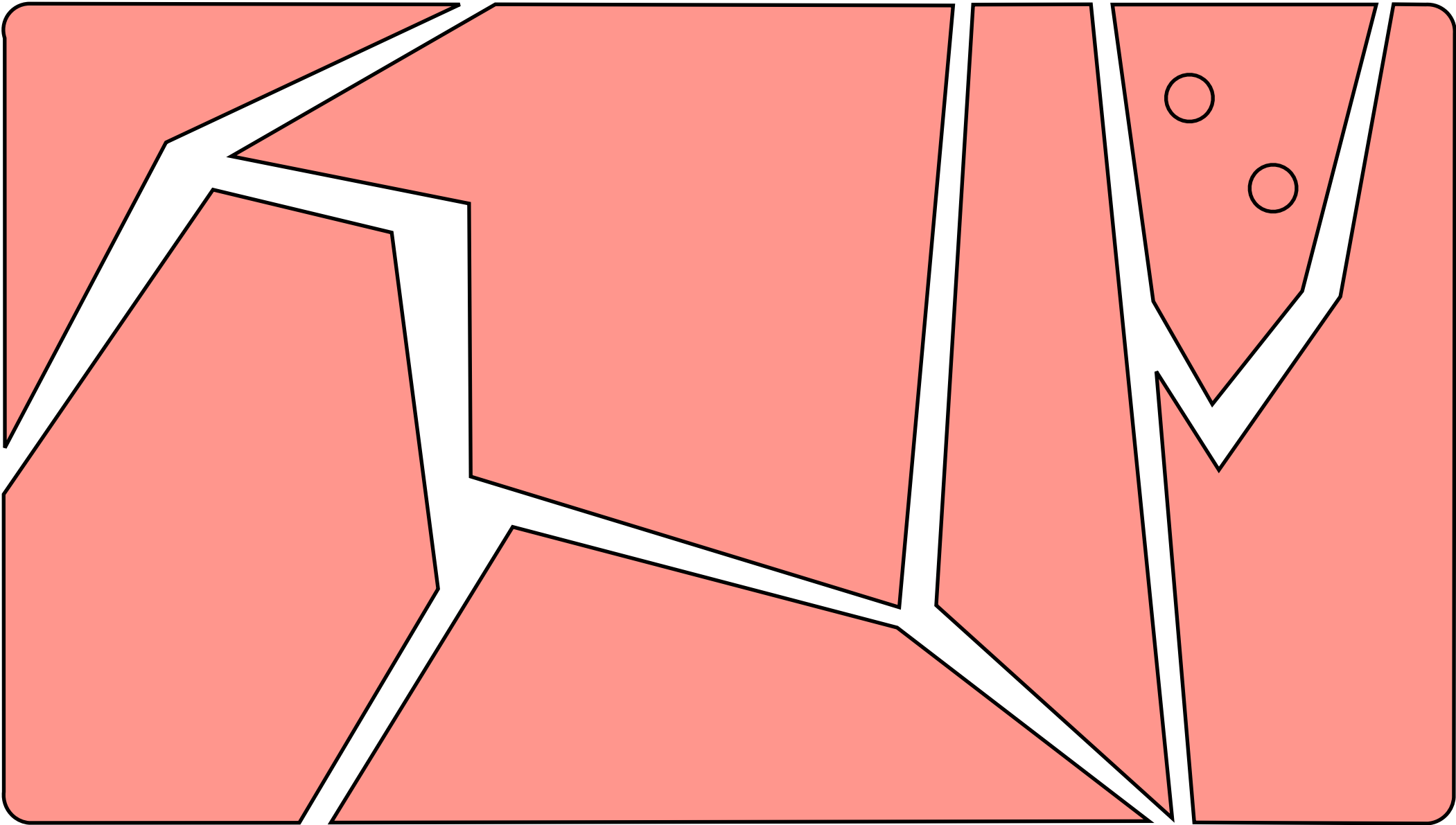
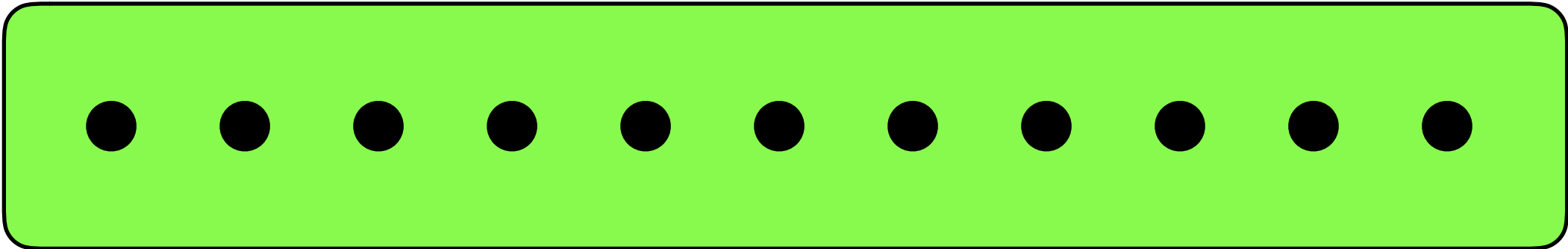
COLORFUL COMPONENTS | VERTEX COVER



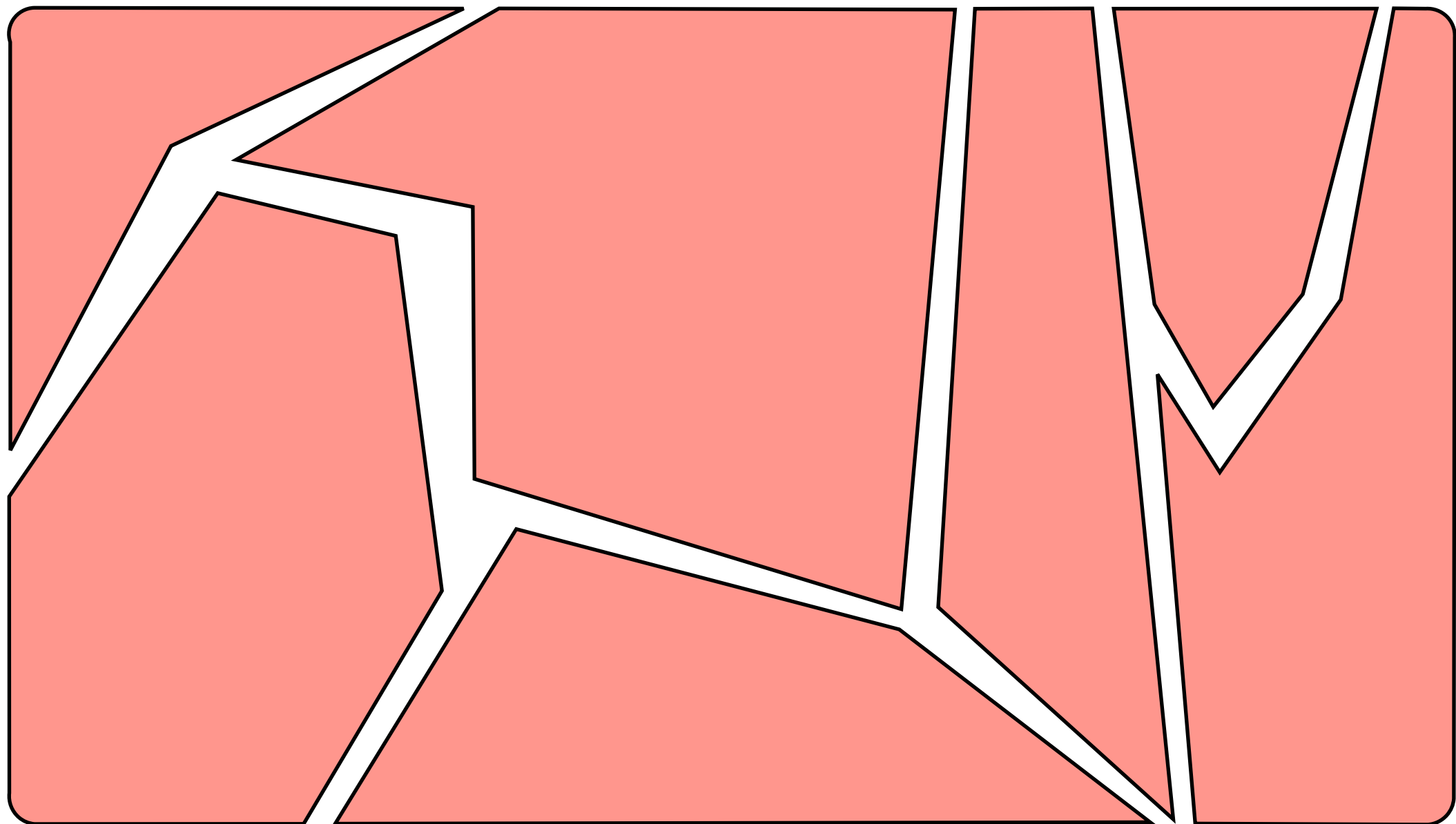
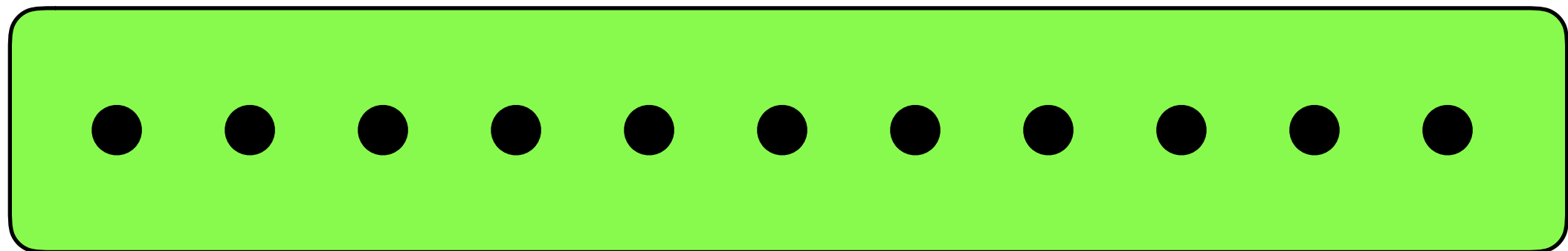
COLORFUL COMPONENTS | VERTEX COVER



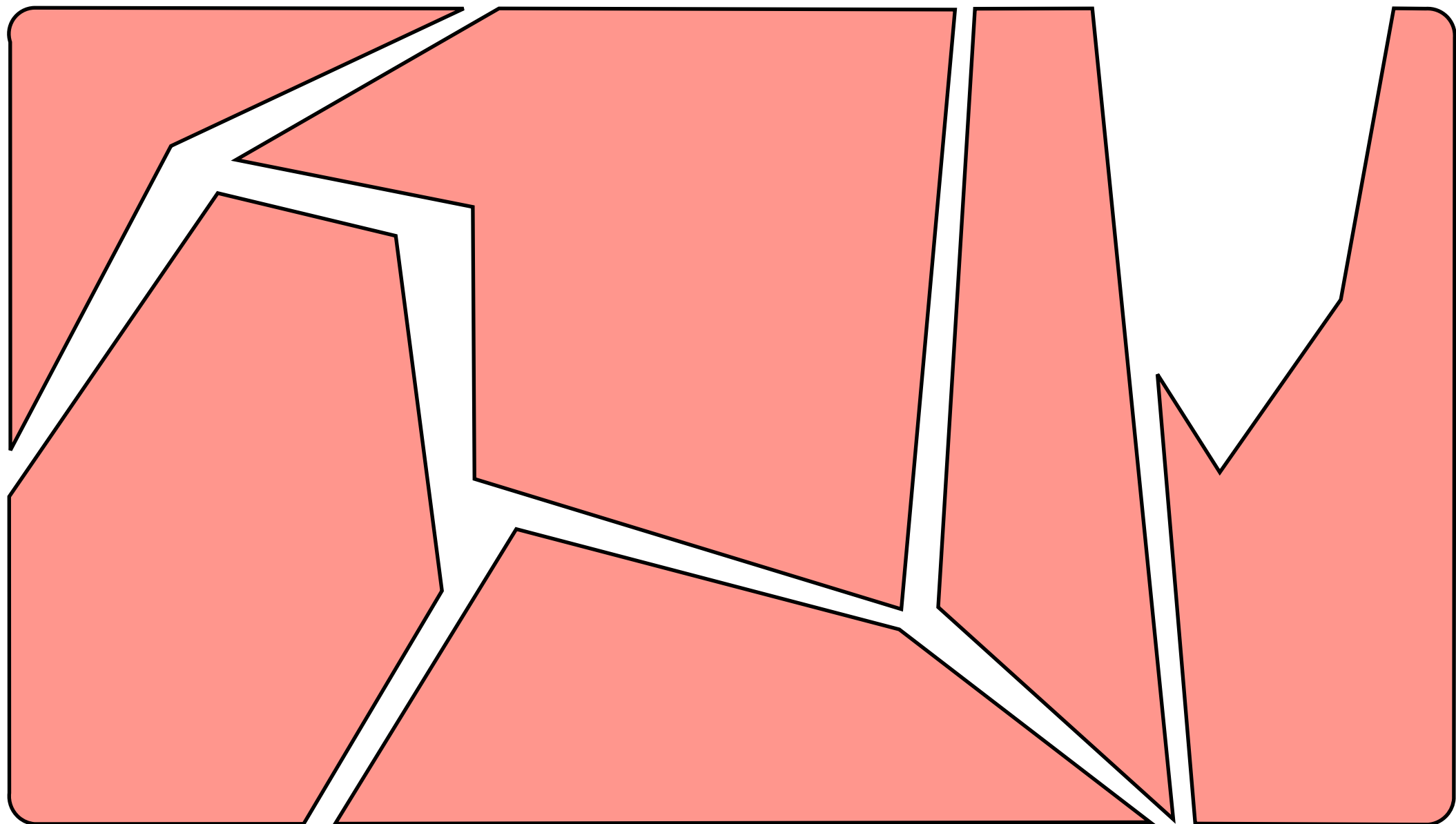
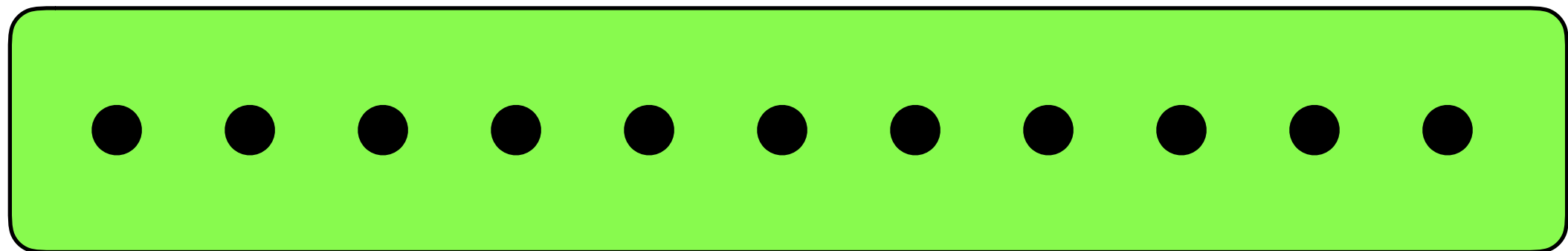
COLORFUL COMPONENTS | VERTEX COVER

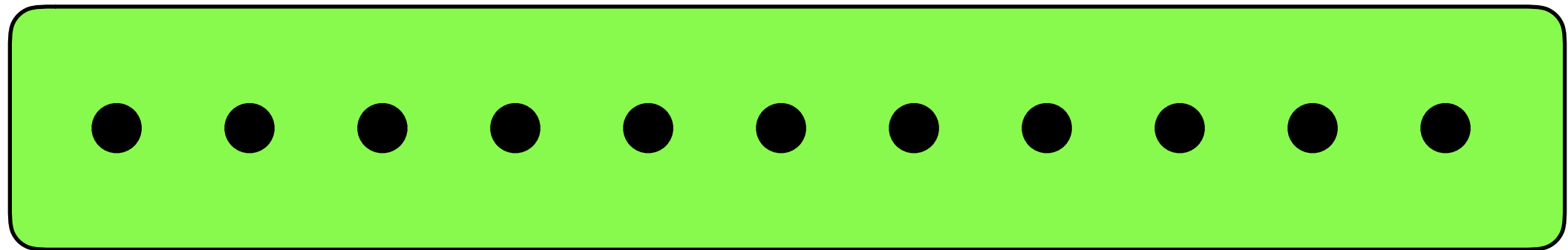


COLORFUL COMPONENTS | VERTEX COVER



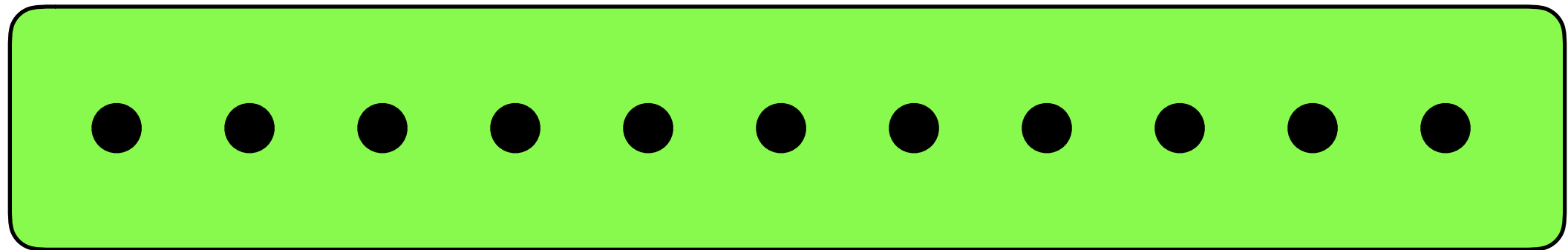
COLORFUL COMPONENTS | VERTEX COVER





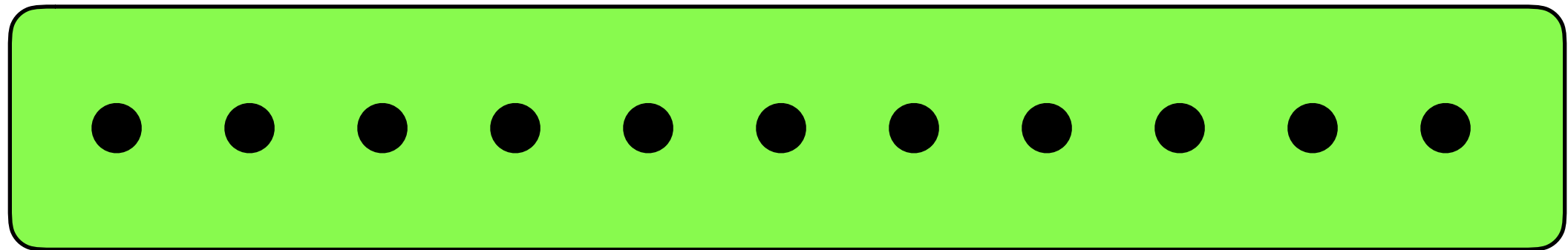
Useful to remember.

#(non-singleton components in $G \setminus F$) is at most t



We can afford to...

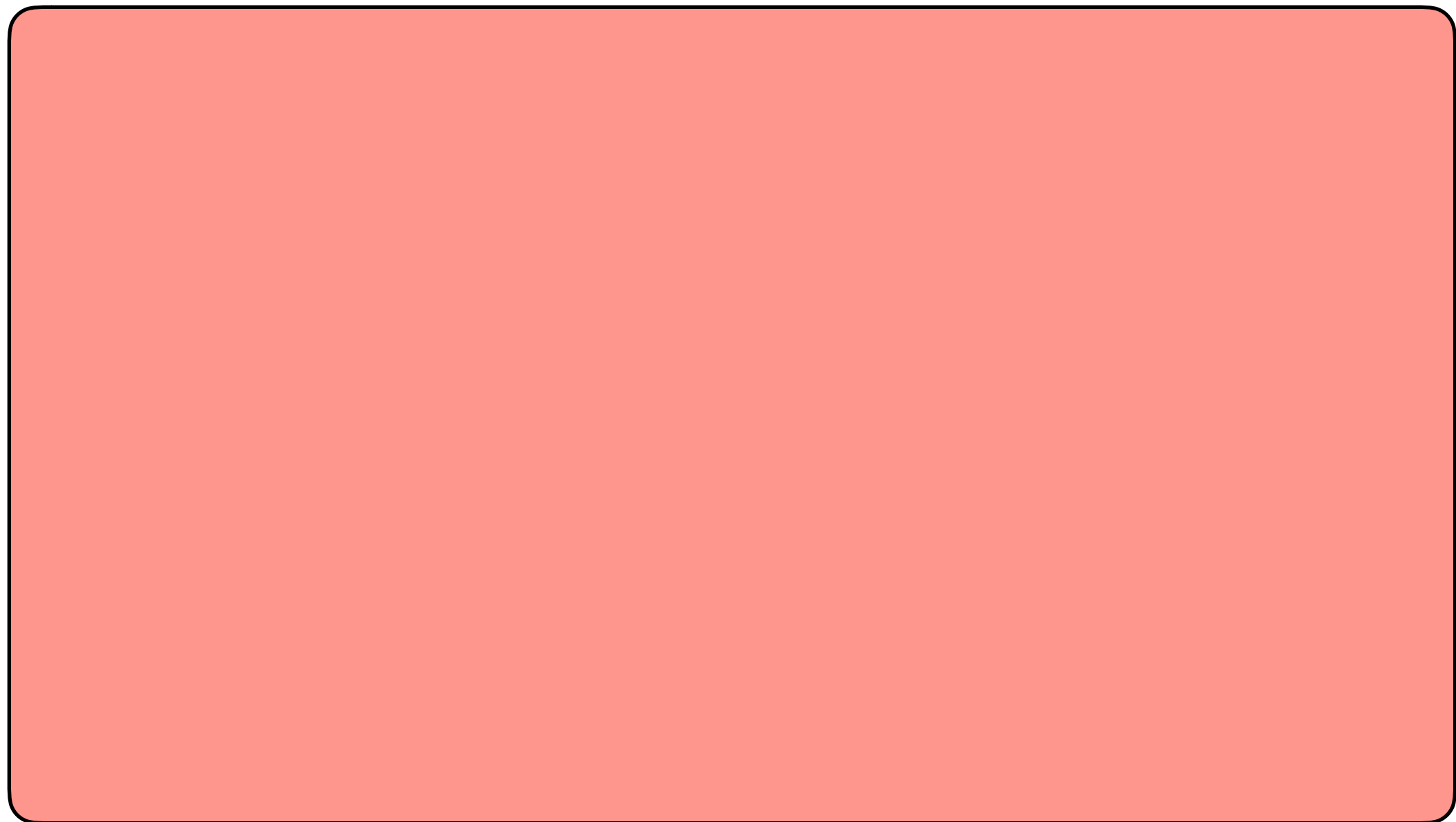
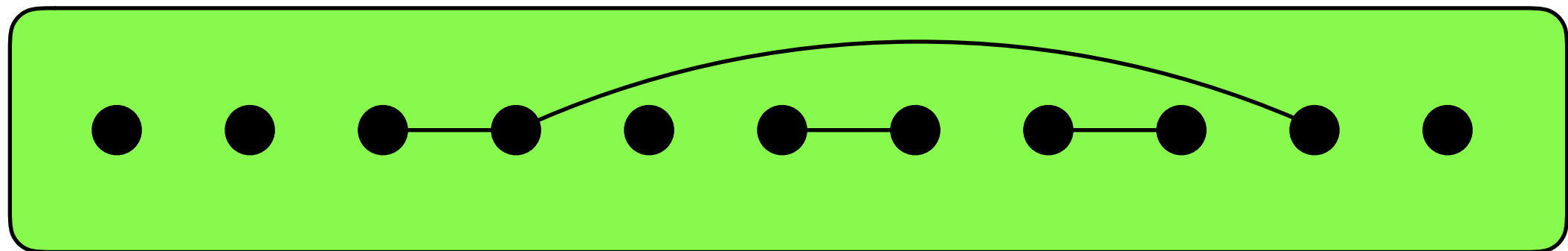
Guess the exact number of components, say h .

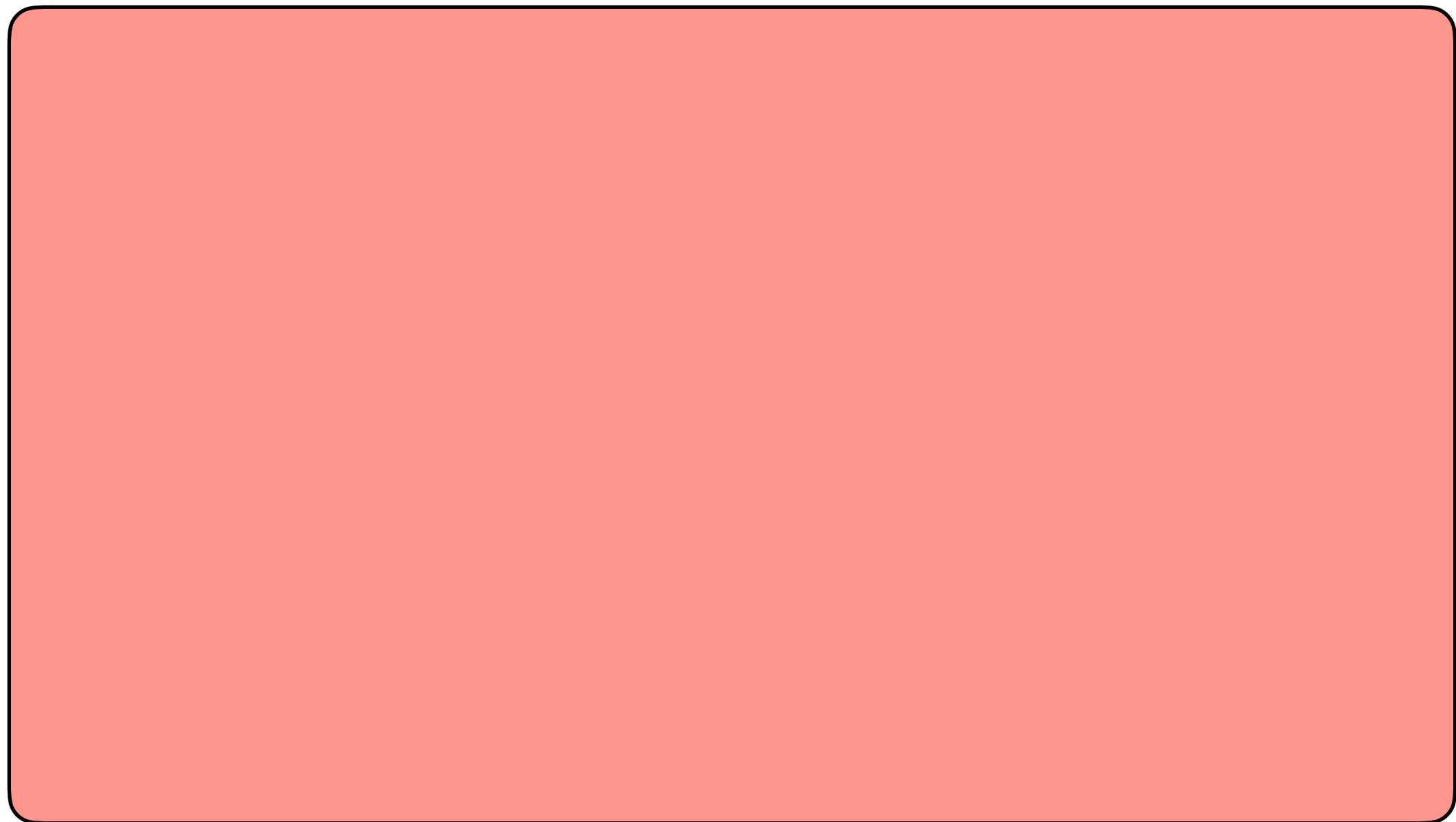
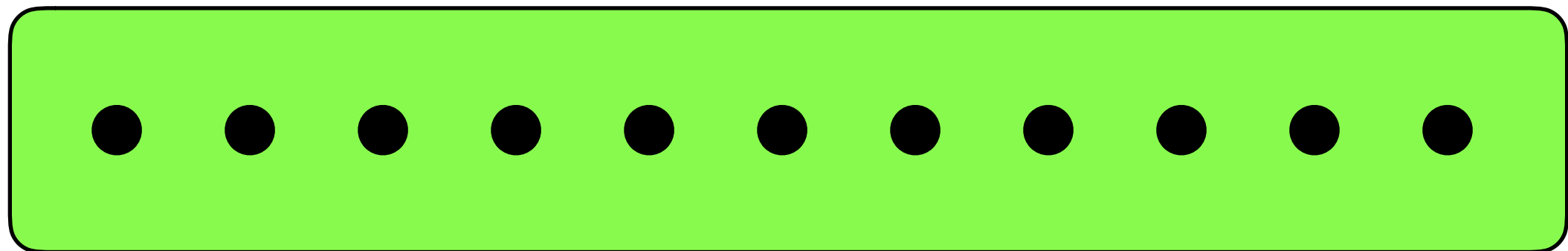


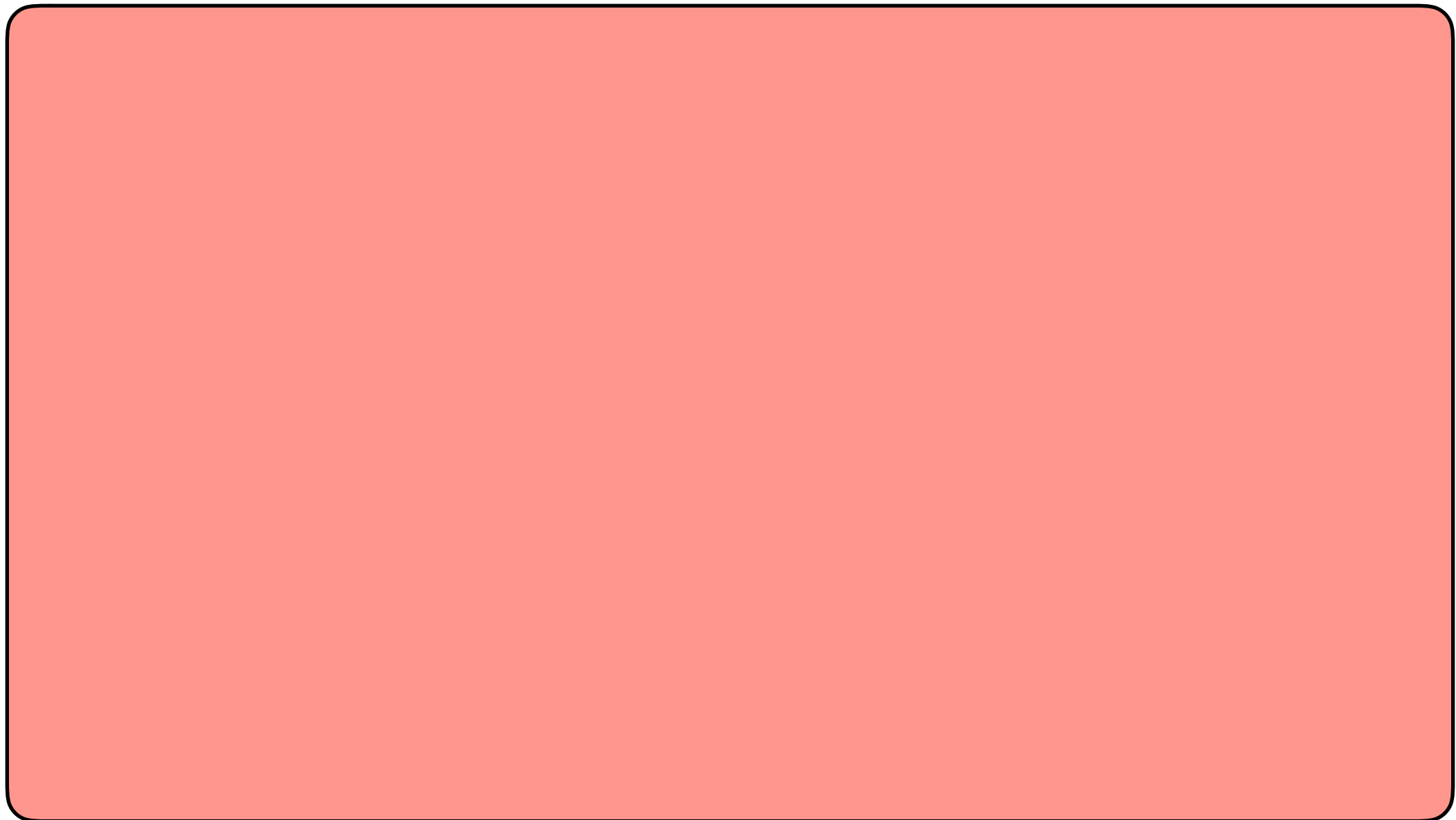
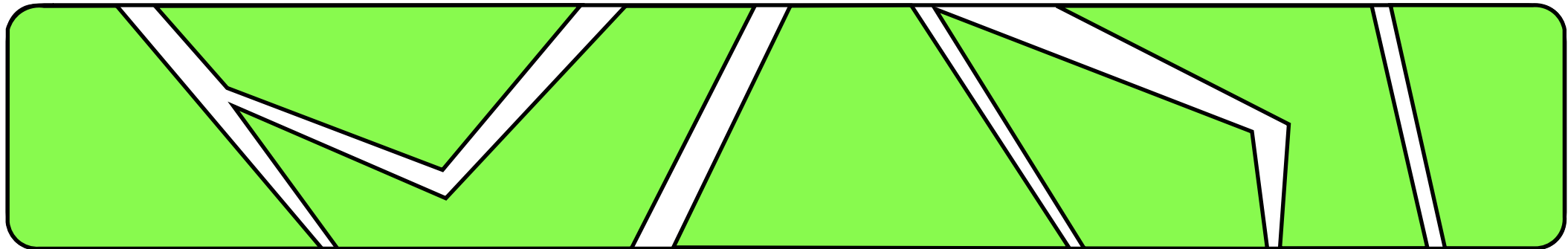
The first step (A)

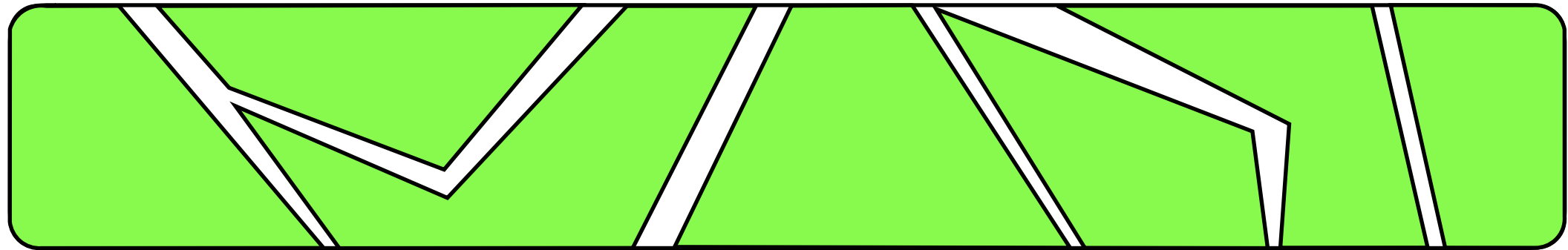
Guess the intersection of F and the edges in T .

COLORFUL COMPONENTS | VERTEX COVER



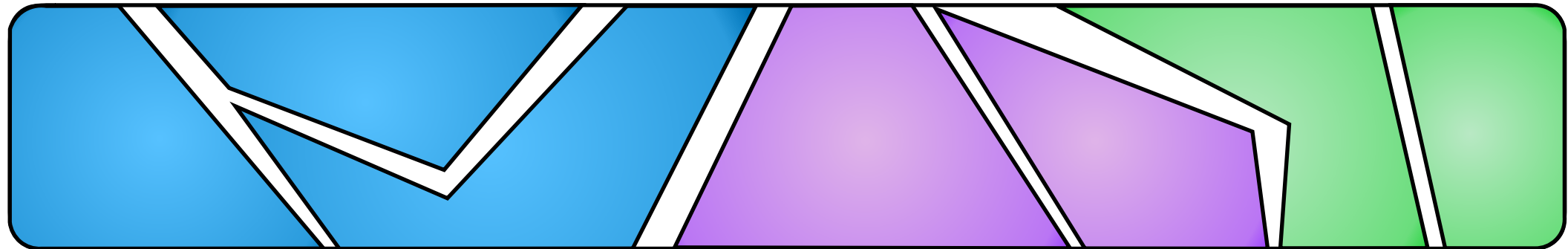






The first step (B)

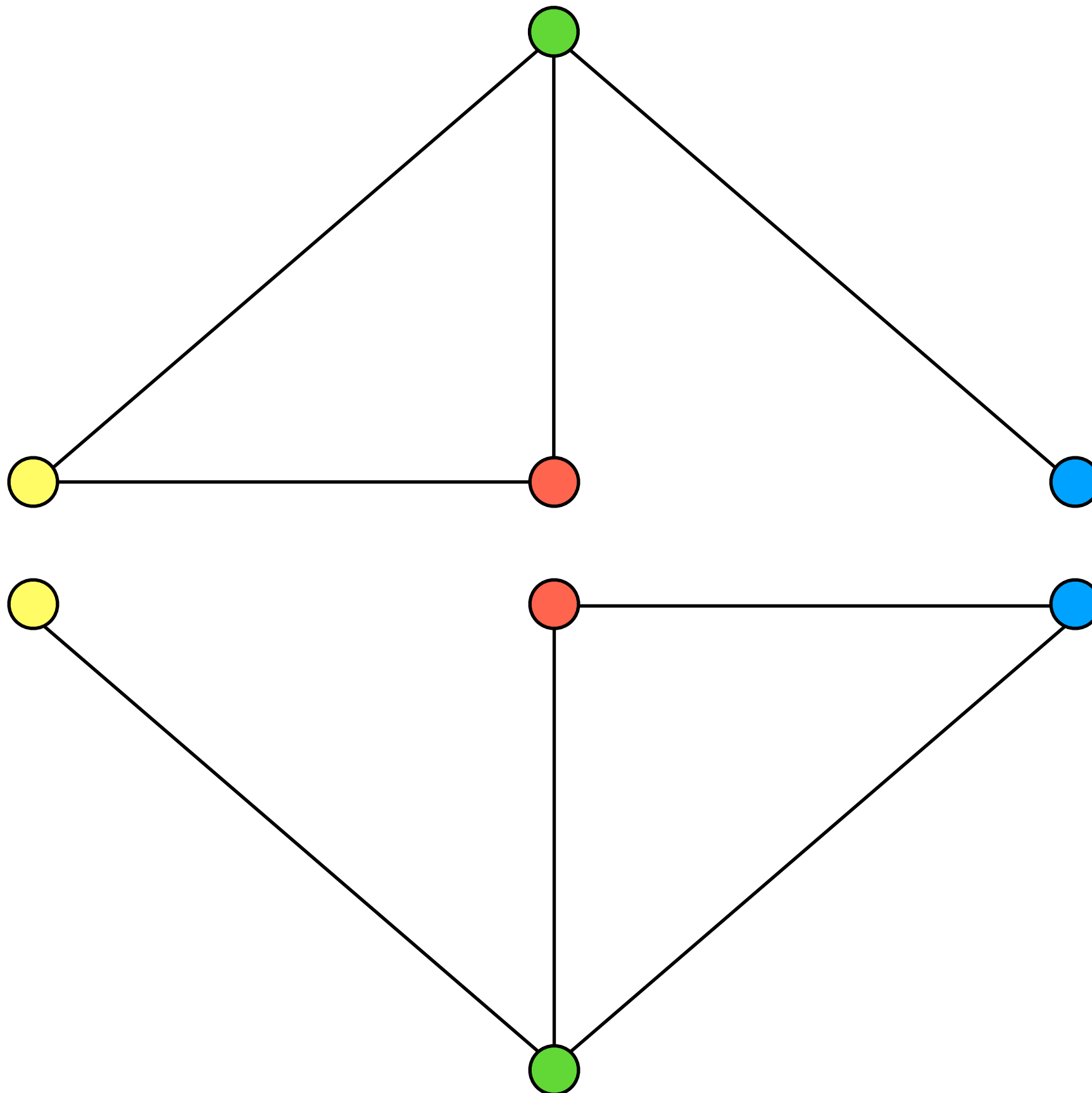
Guess the “component structure” of $G \setminus F$ inside T .



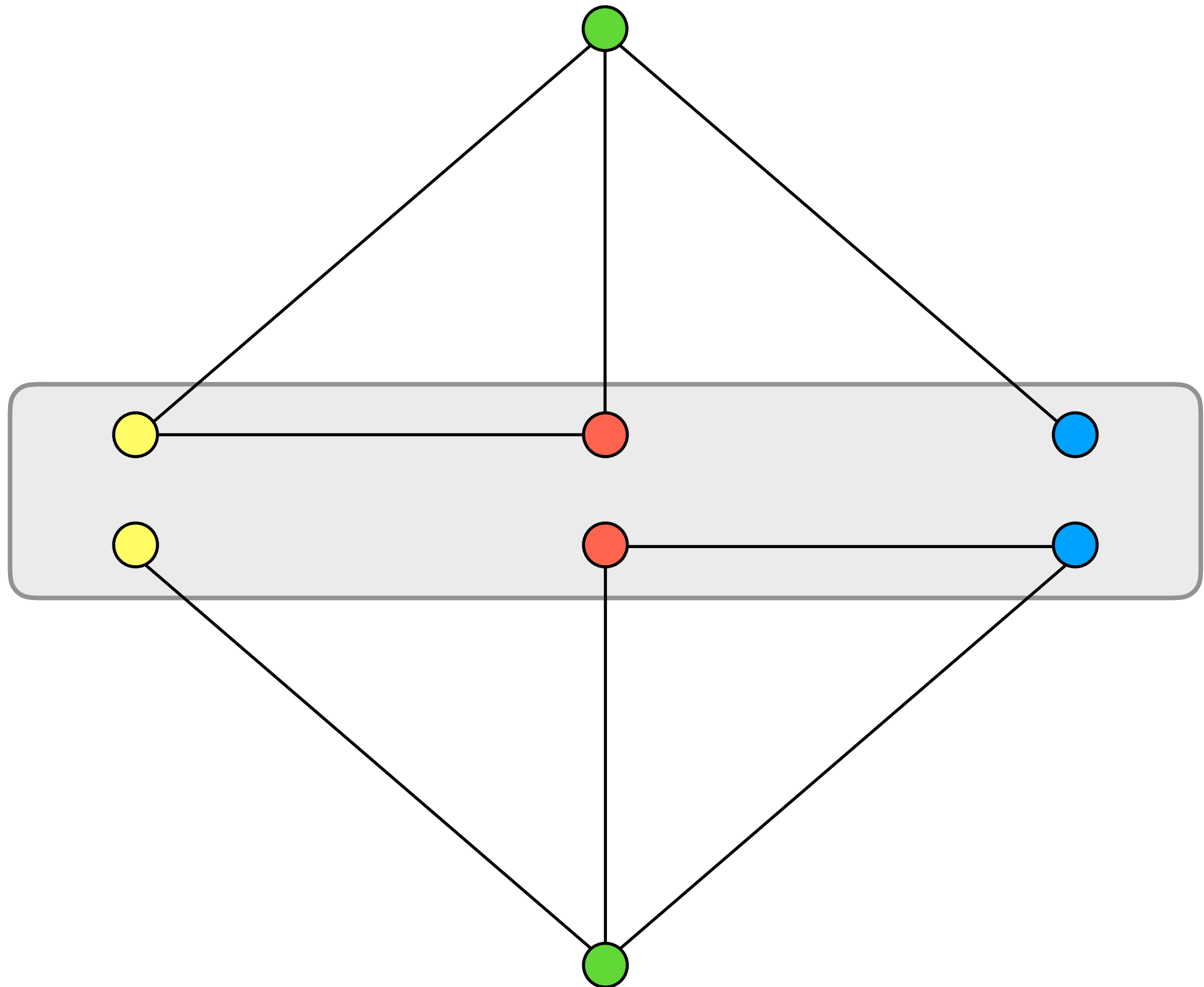
The first step (B)

Guess the “component structure” of $G \setminus F$ inside T .

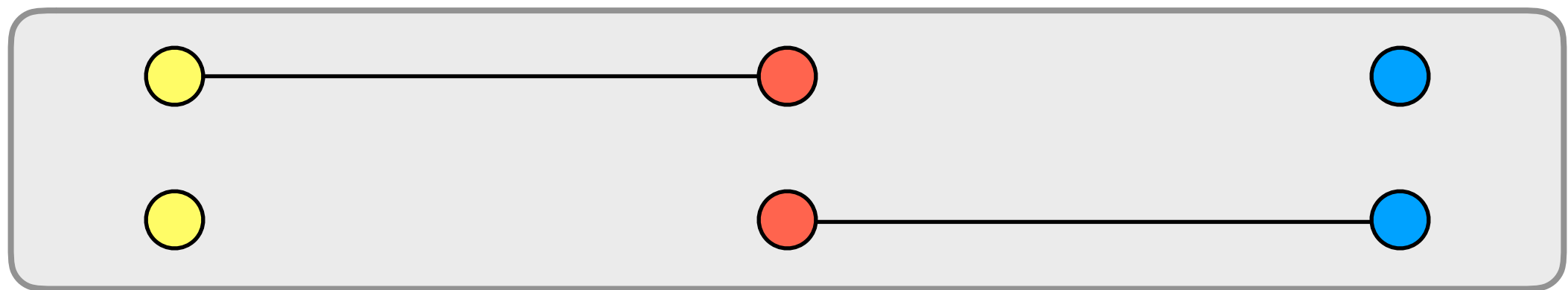
COLORFUL COMPONENTS | VERTEX COVER



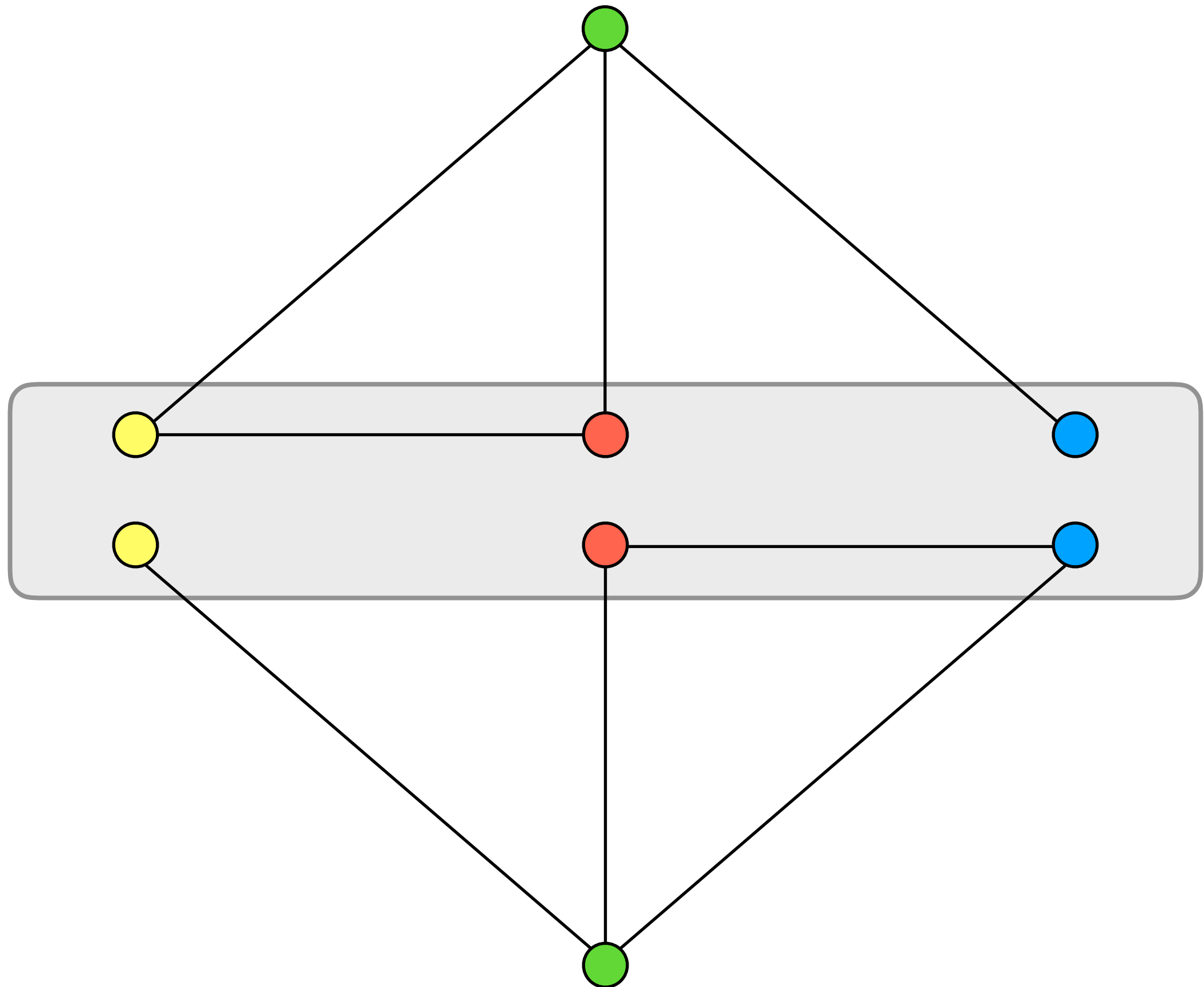
COLORFUL COMPONENTS | VERTEX COVER



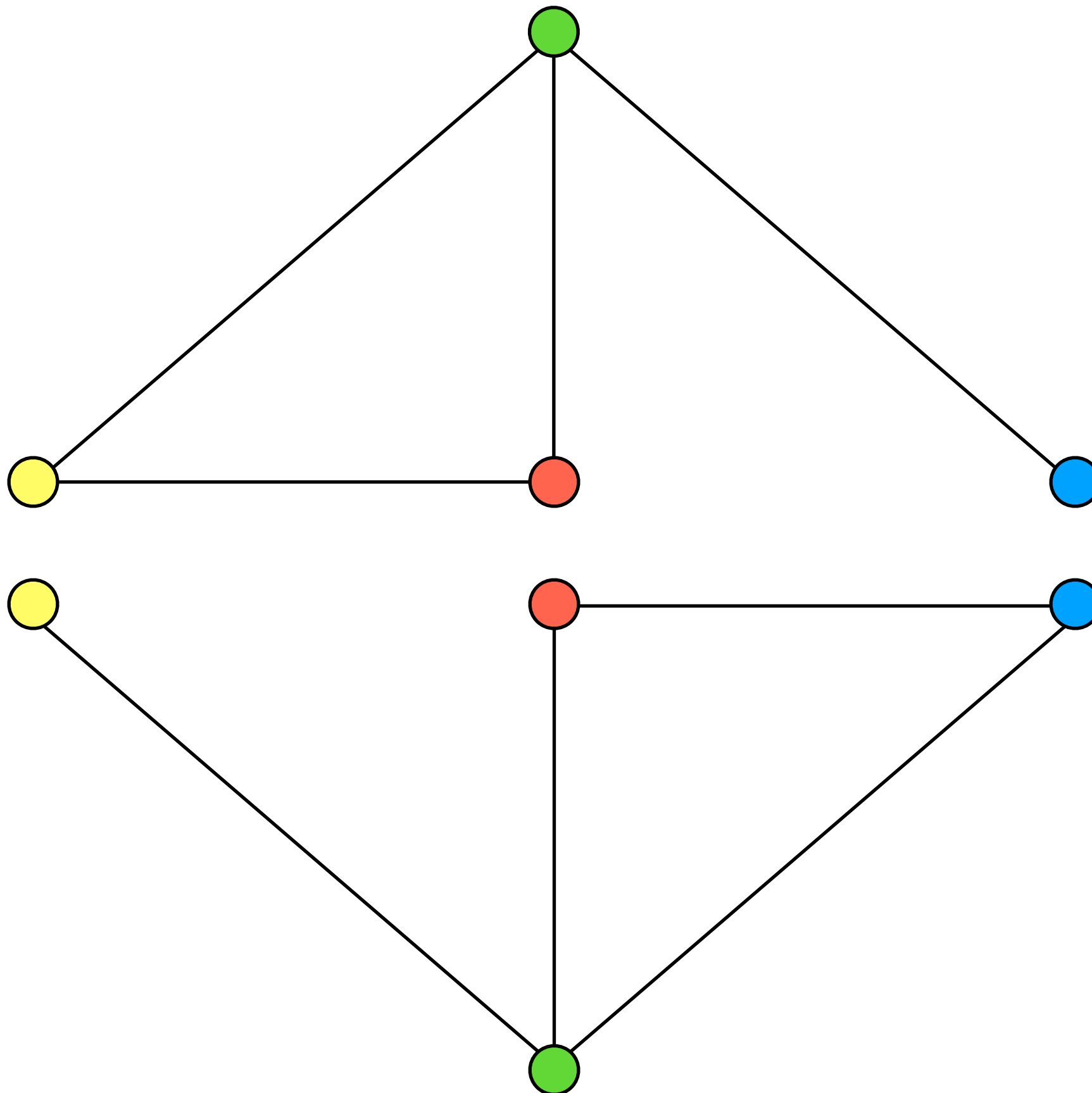
COLORFUL COMPONENTS | VERTEX COVER



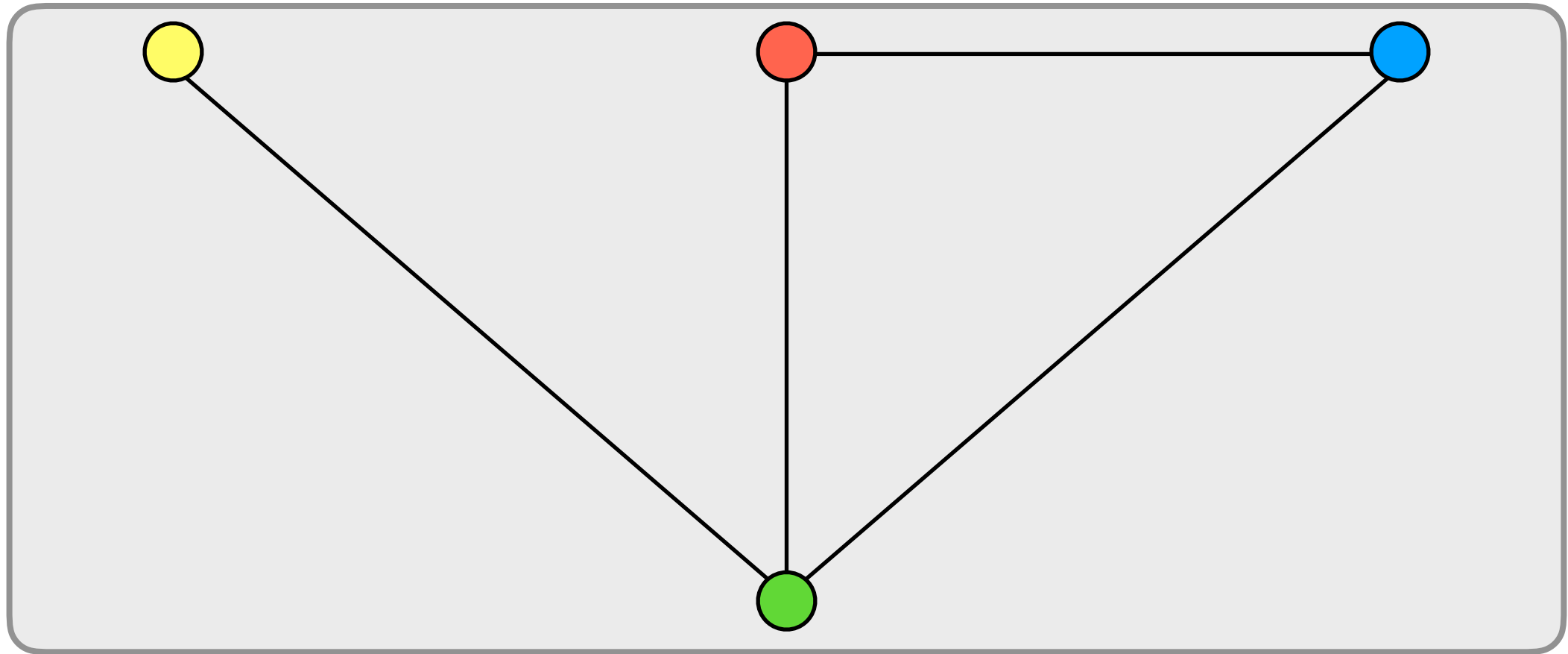
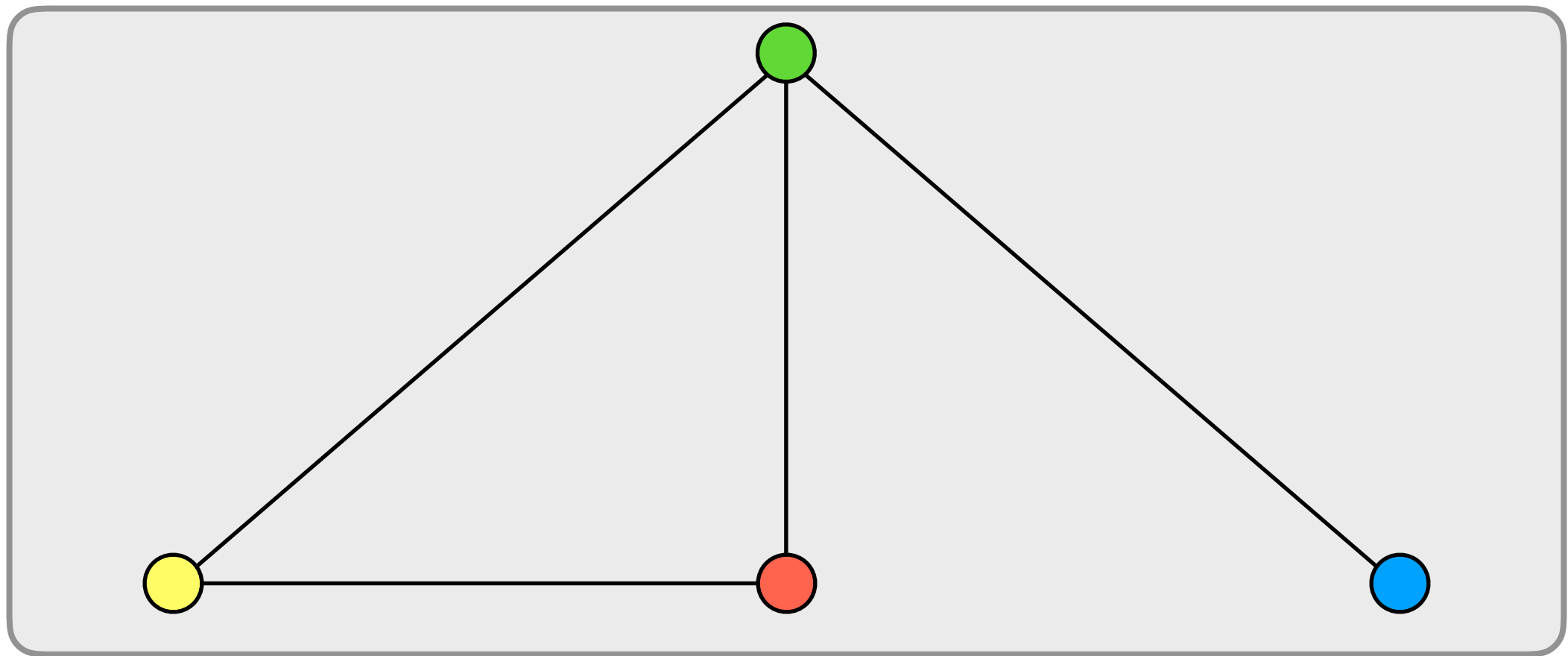
COLORFUL COMPONENTS | VERTEX COVER

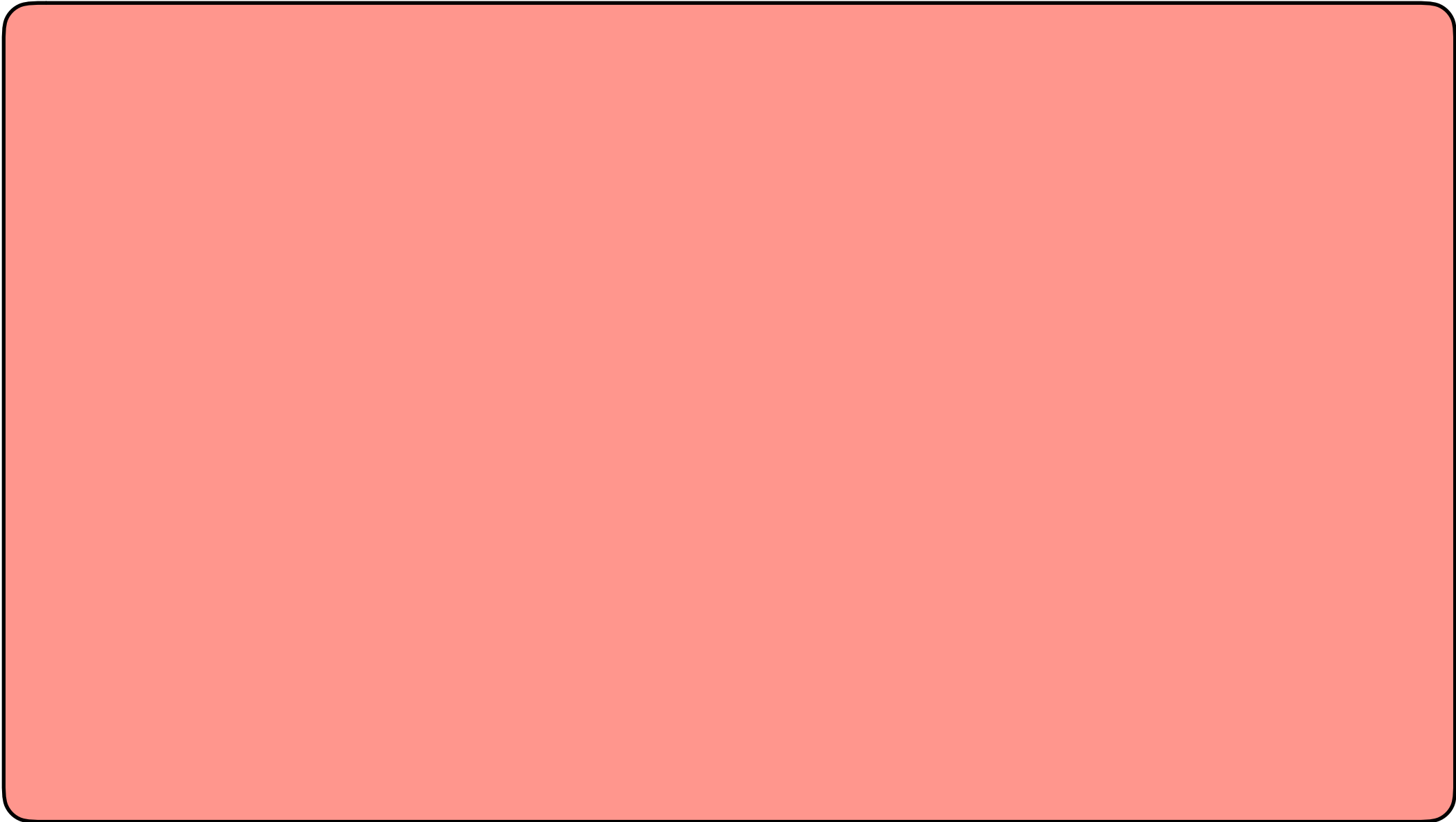
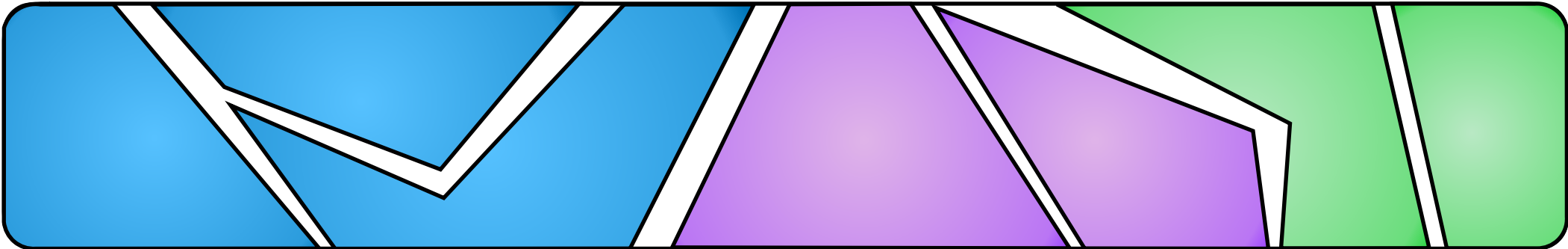


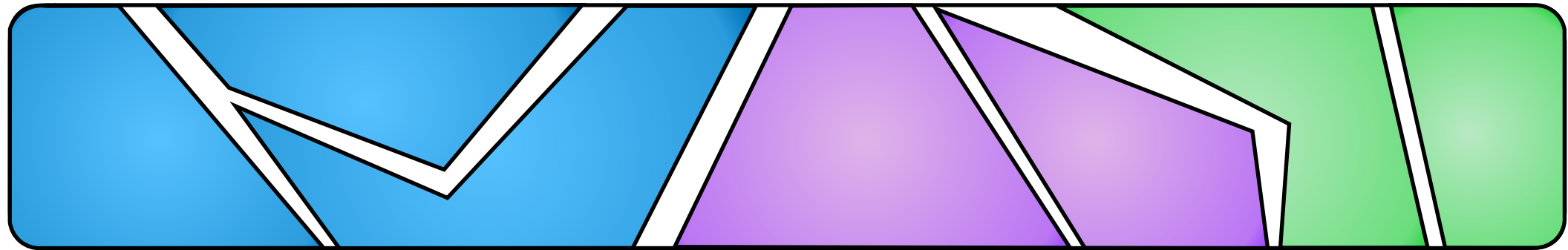
COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER

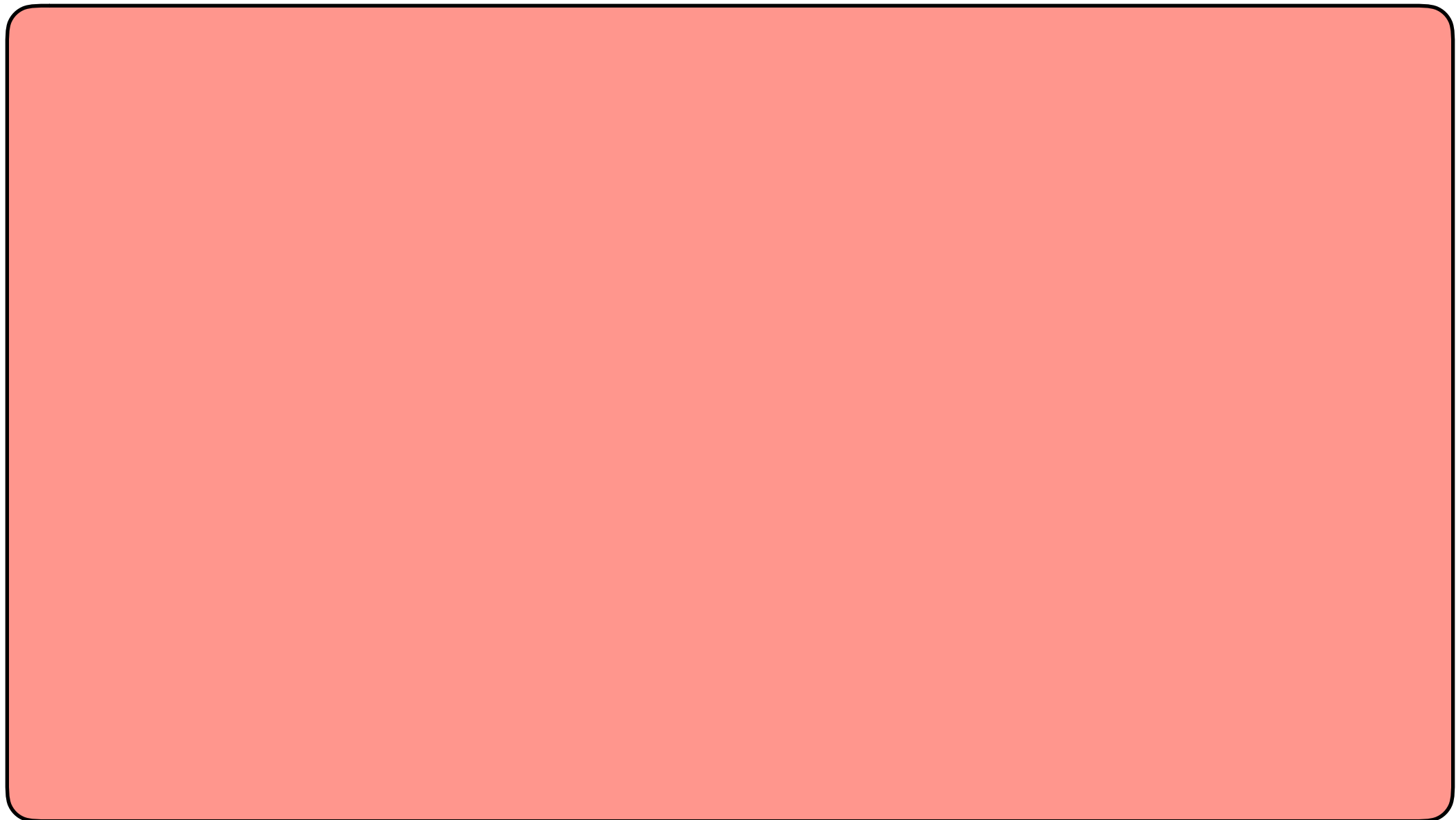
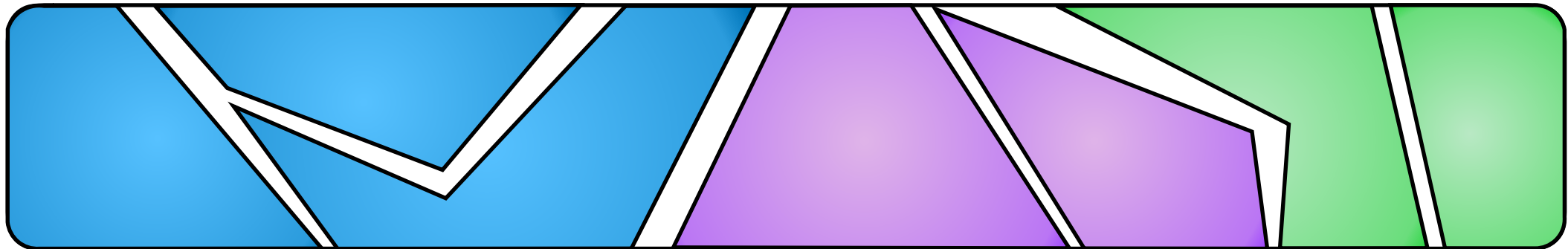




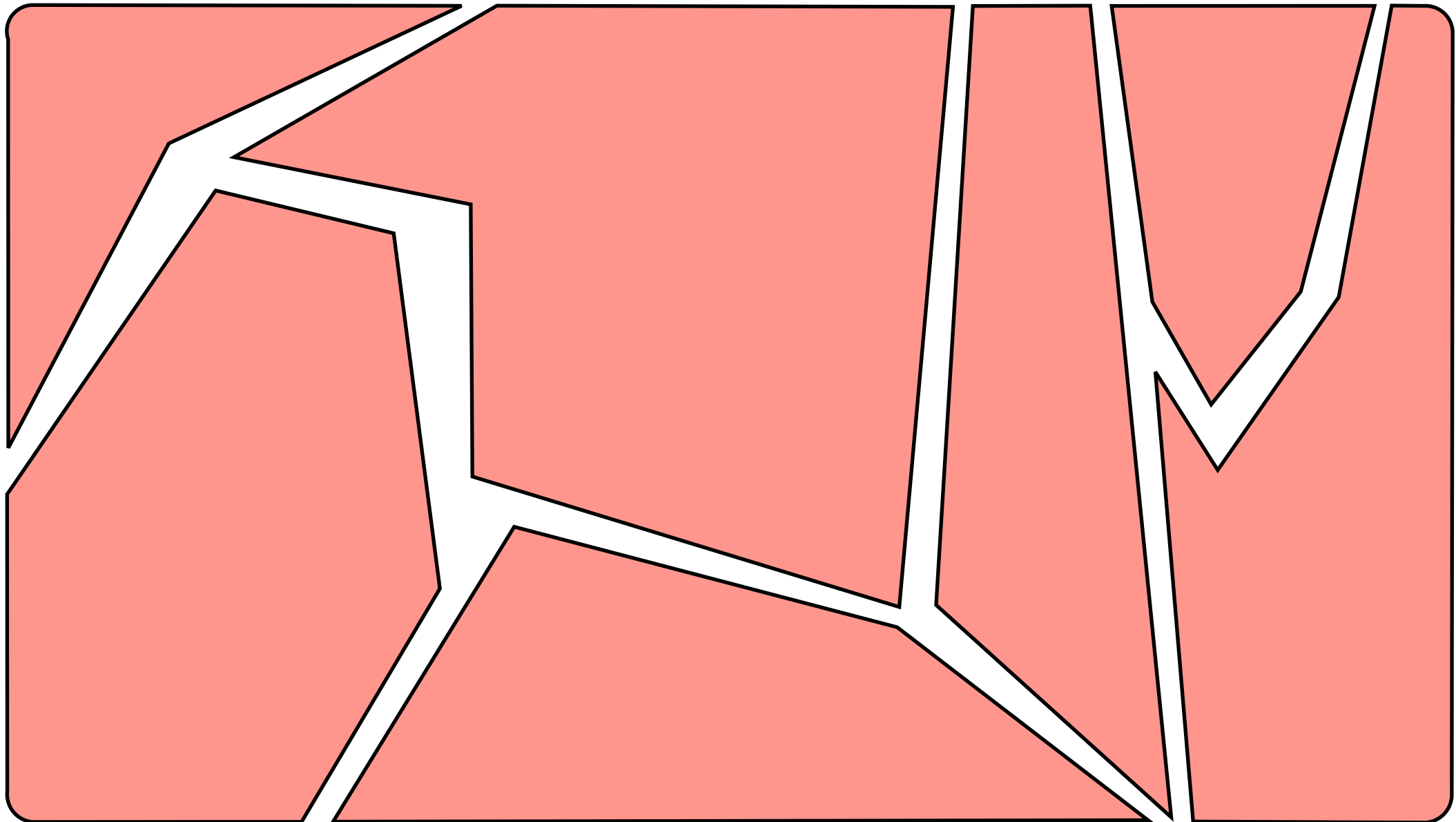
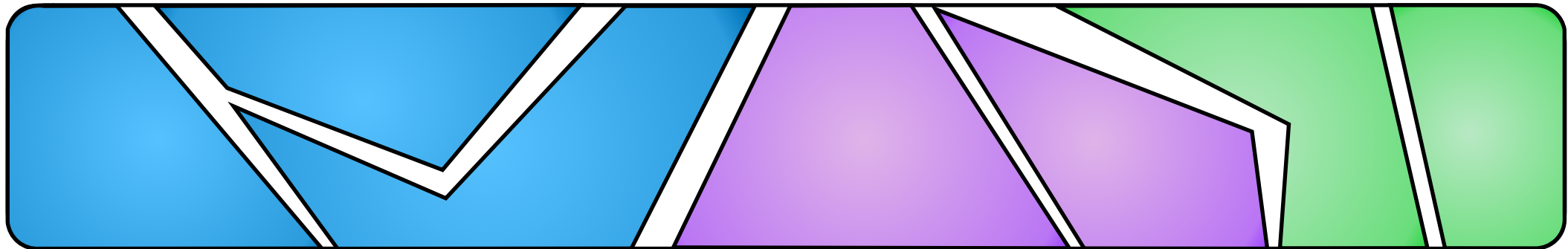


The second step (A)

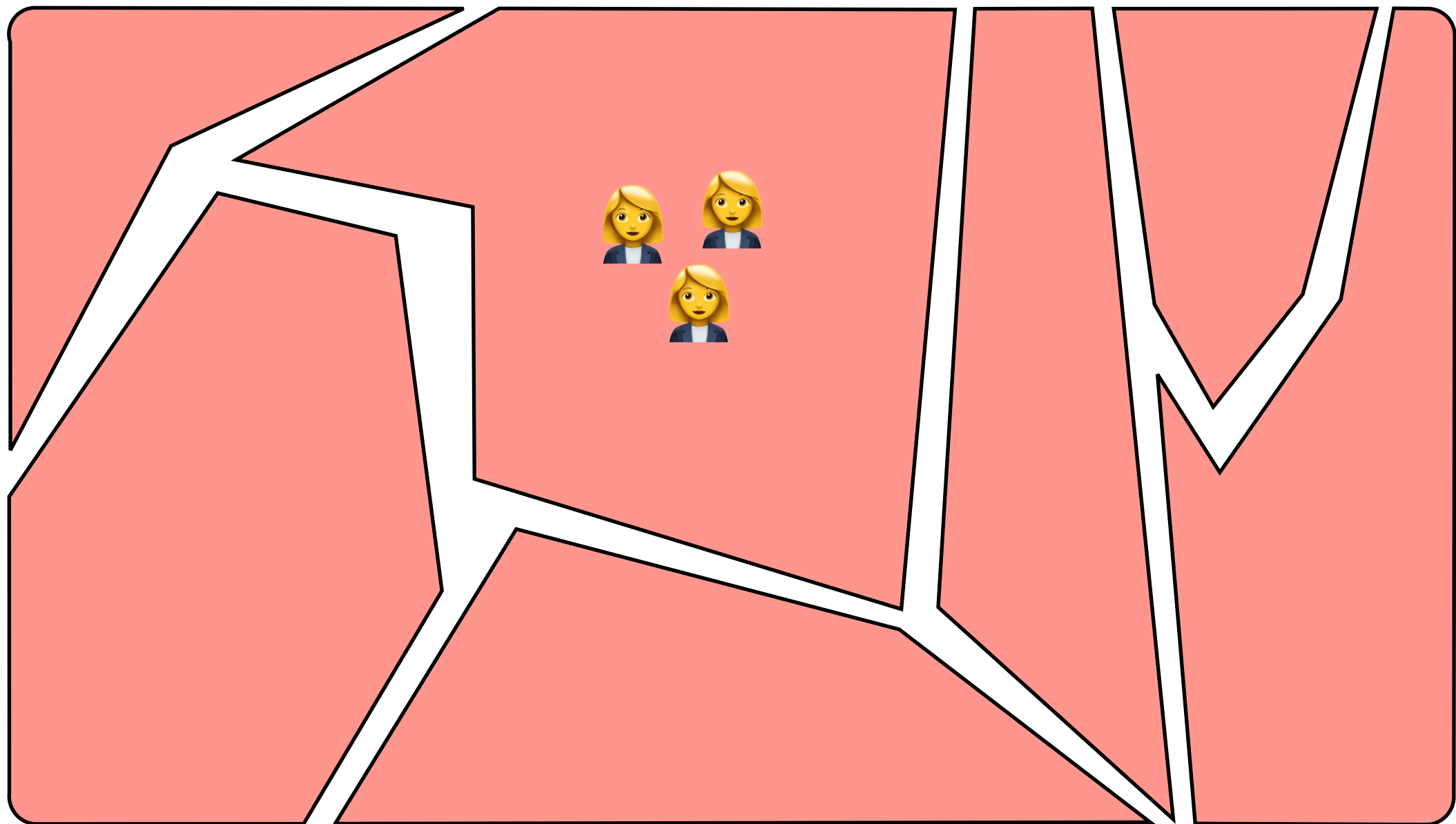
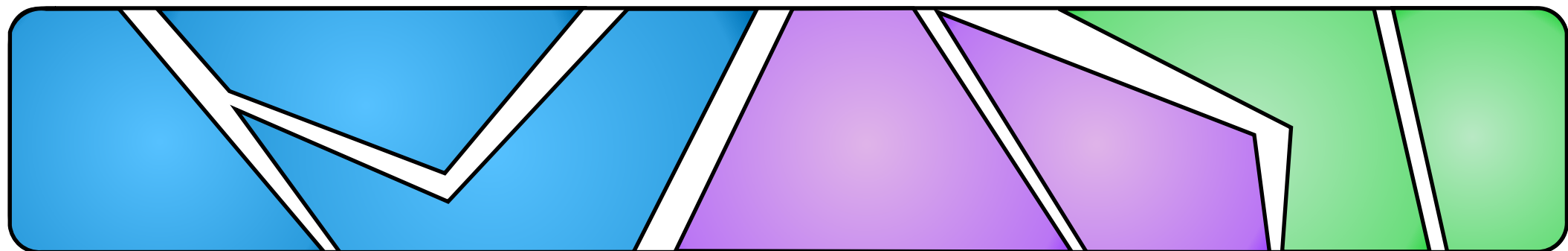
Partition $G \setminus T$ based on neighborhoods in T .

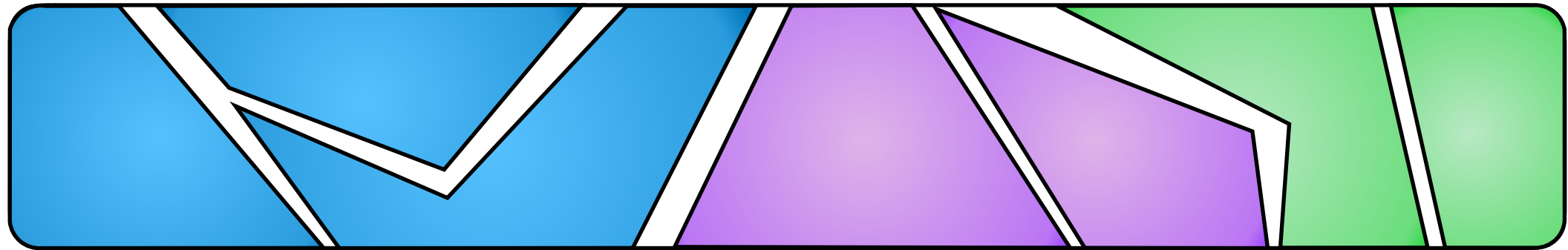


COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER

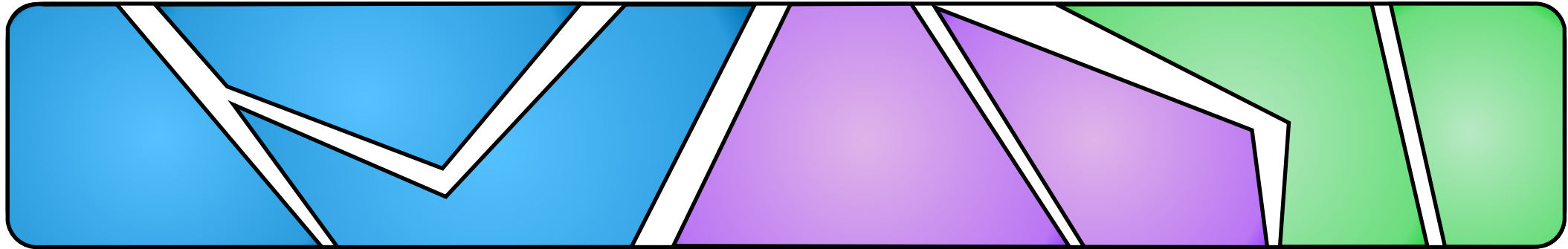




The second step (A)

Partition $G \setminus T$ based on neighborhoods in T .

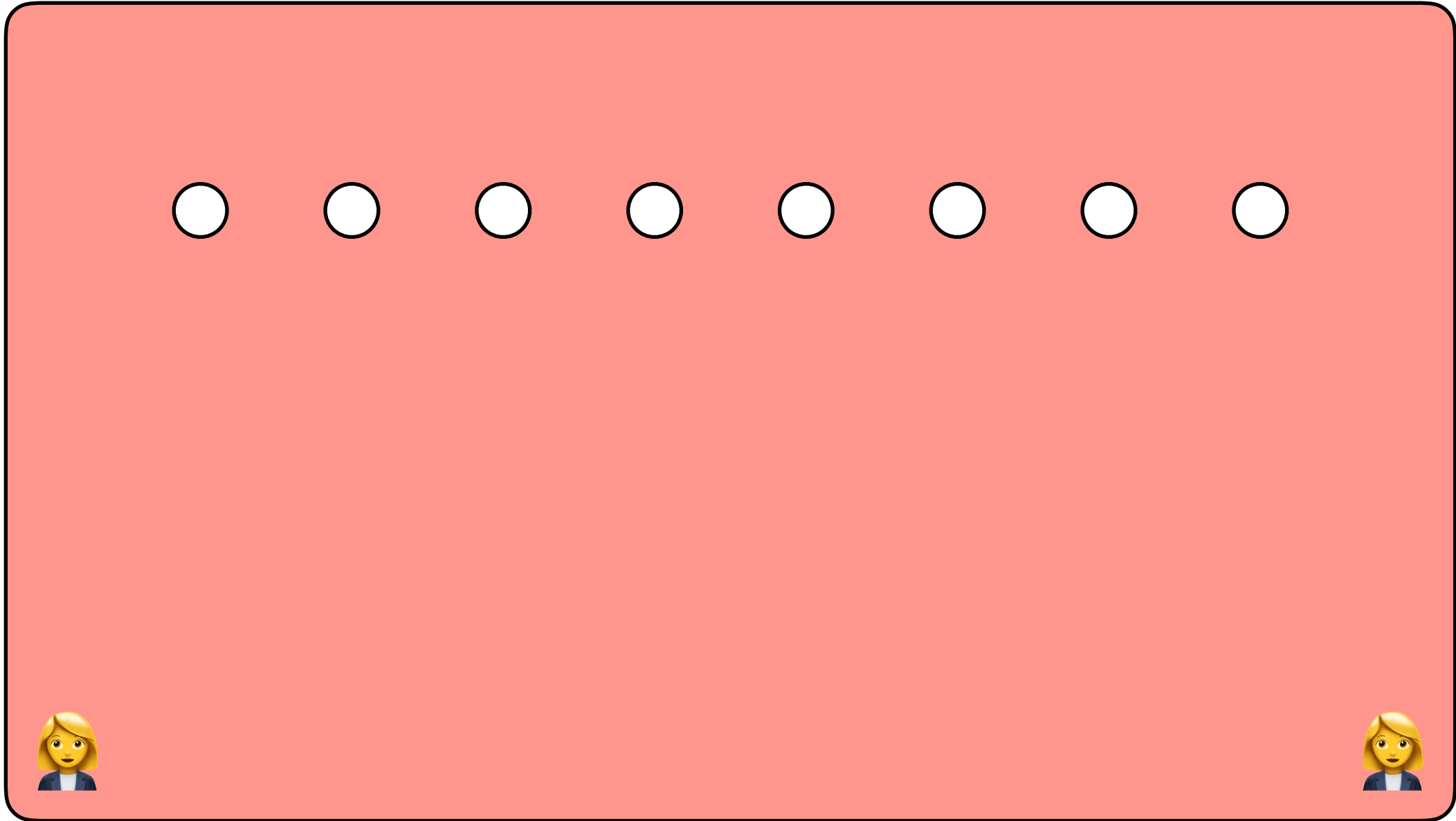
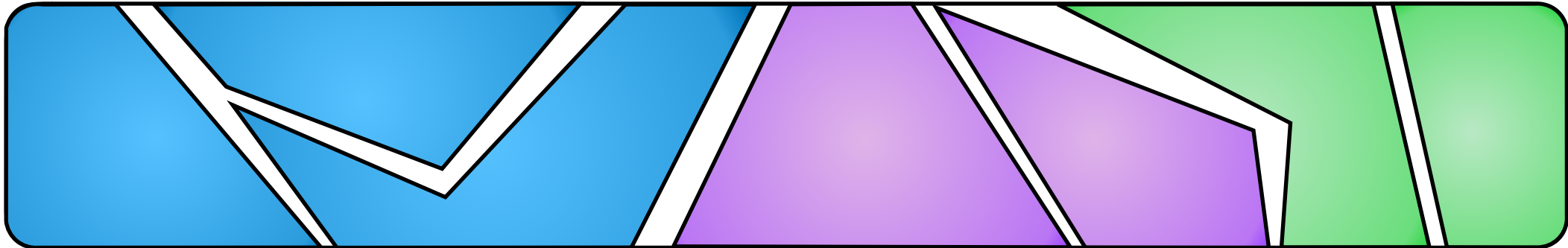
→ Note that there are a bounded number of twin classes.



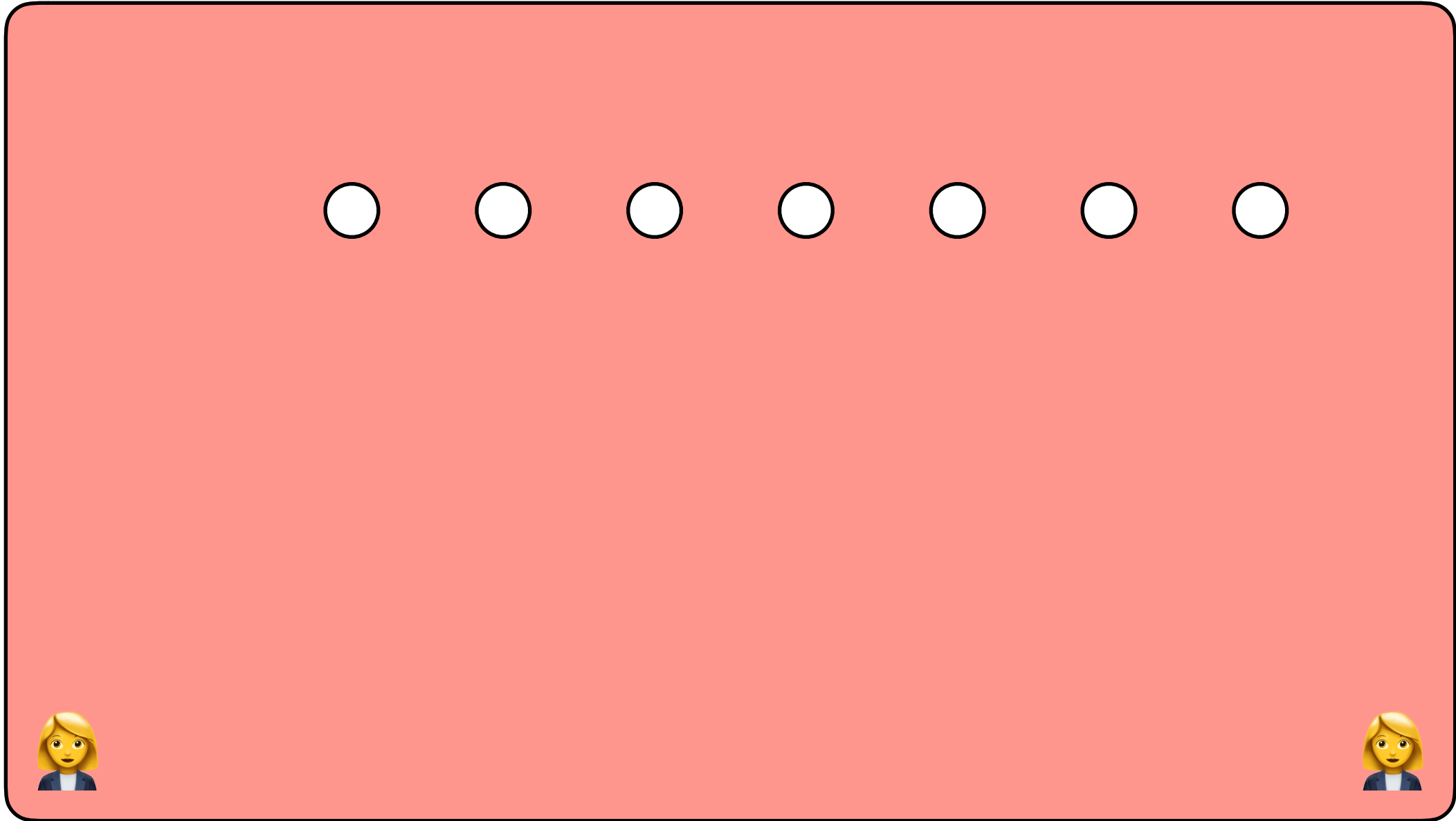
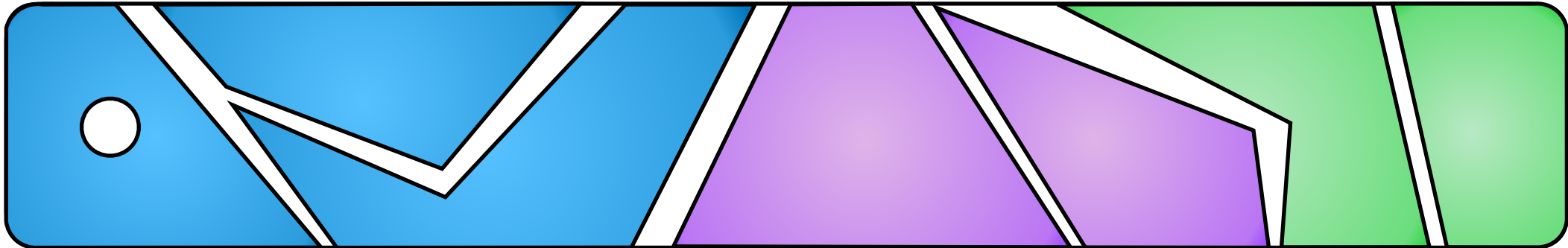
The second step (B)

For any color c , retain at most h twins of color c .

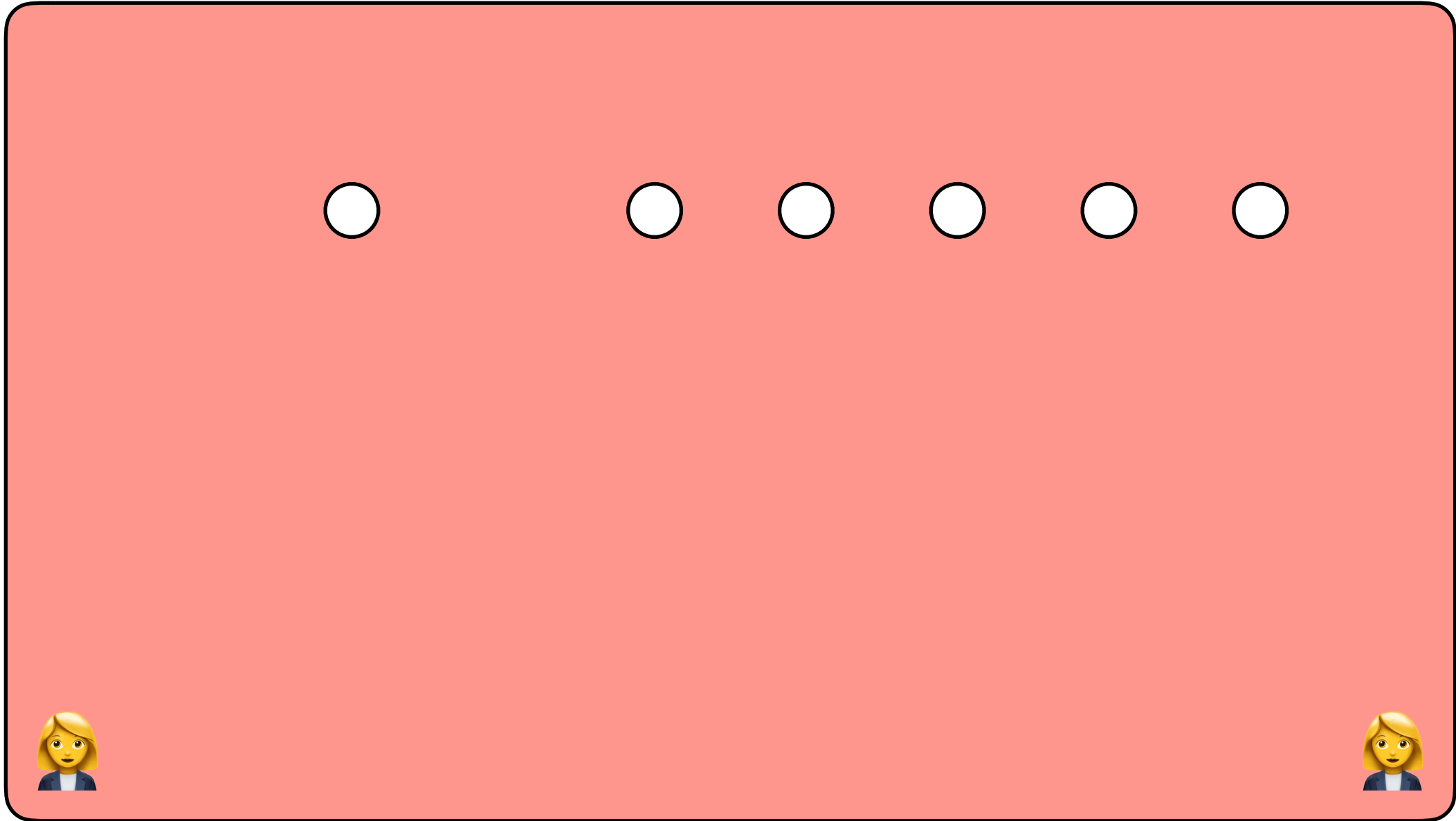
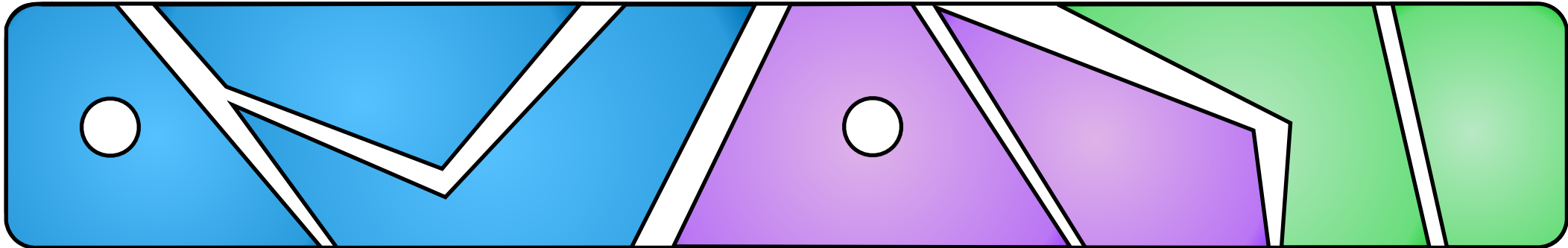
COLORFUL COMPONENTS | VERTEX COVER



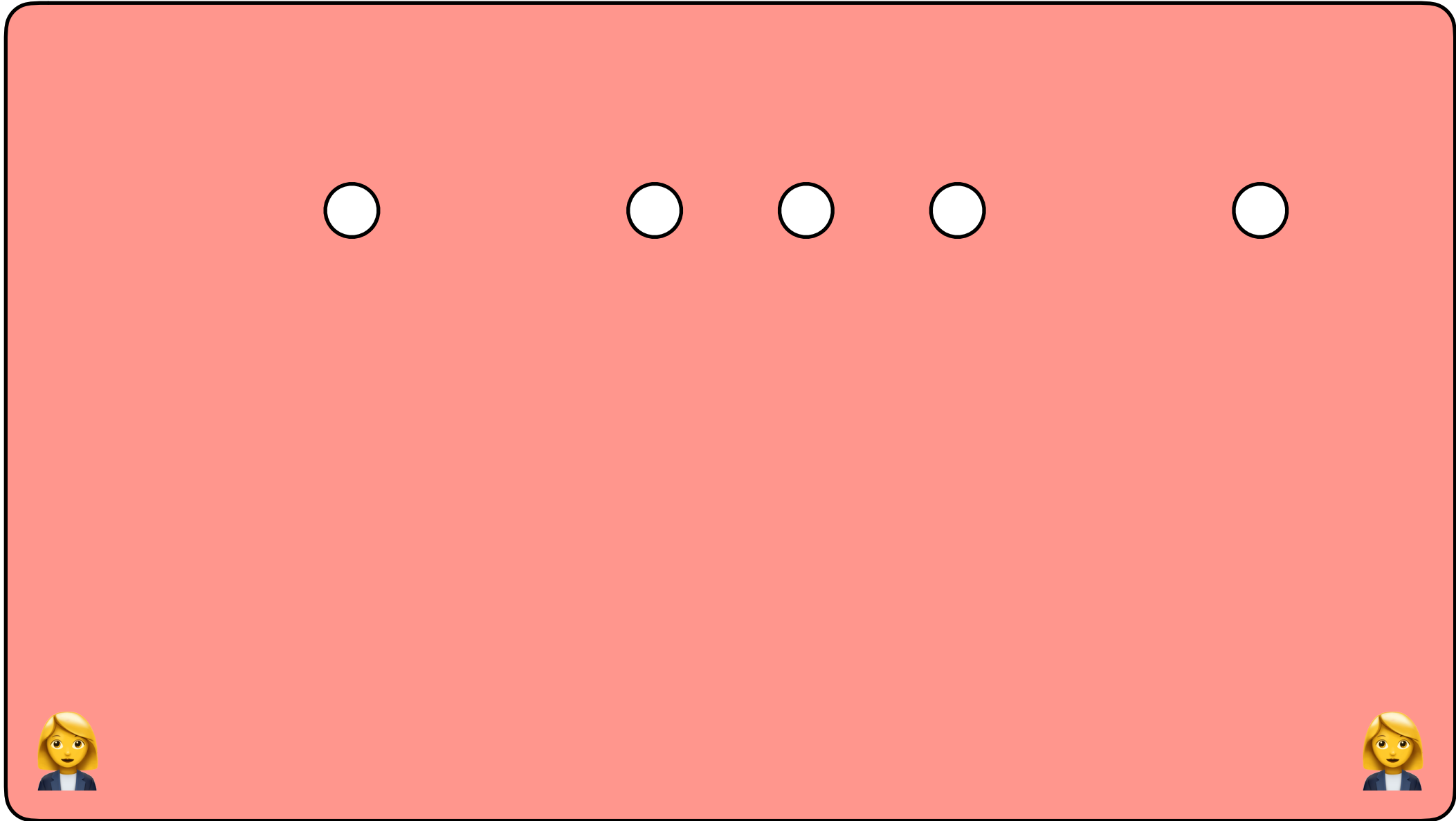
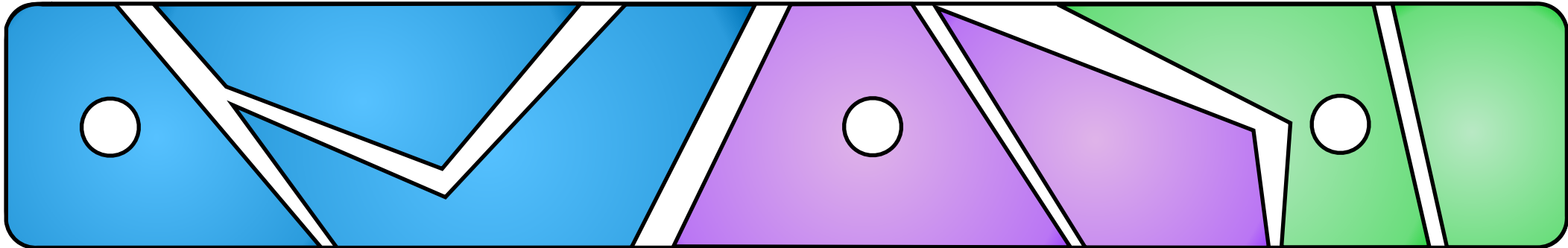
COLORFUL COMPONENTS | VERTEX COVER



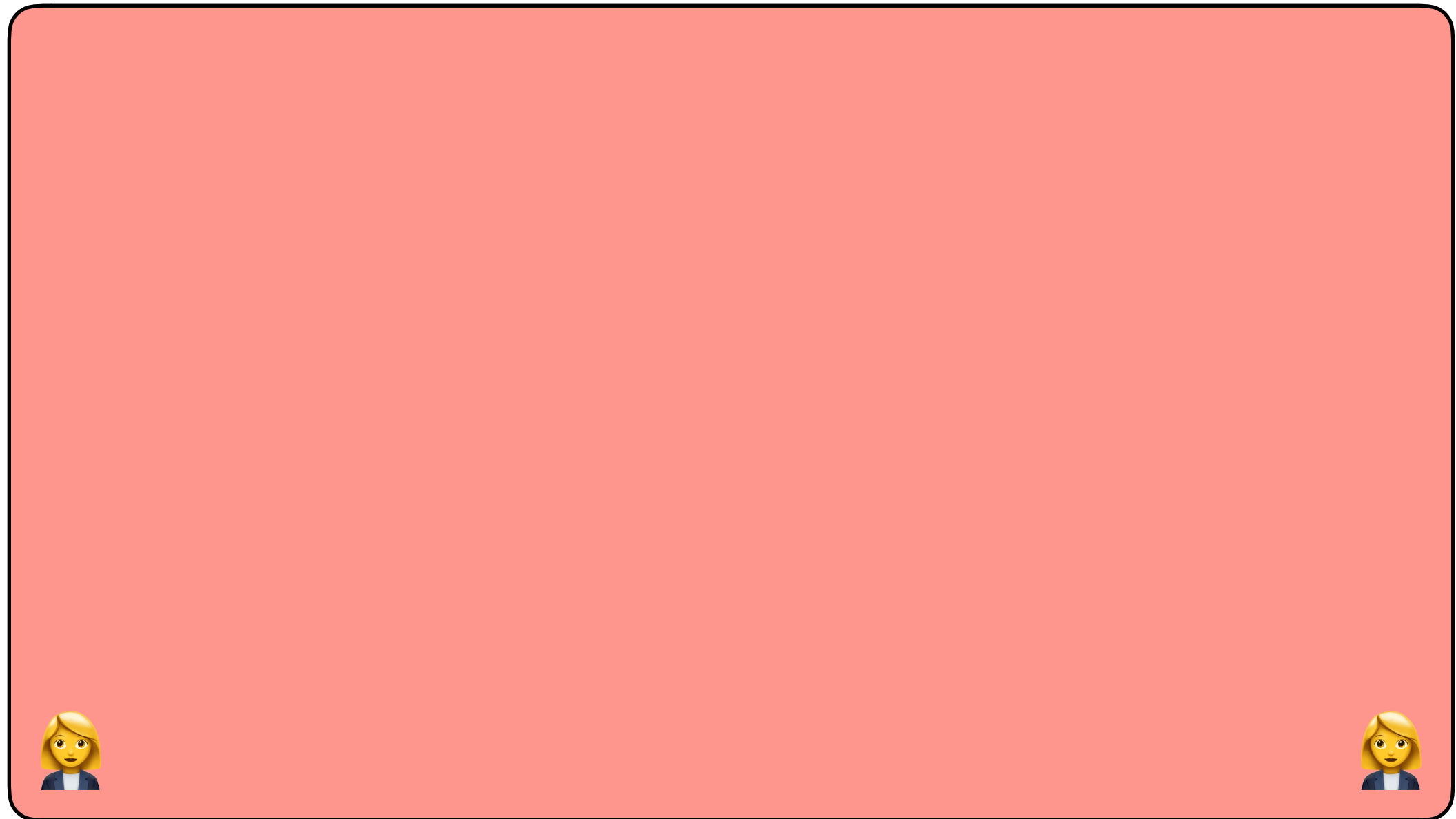
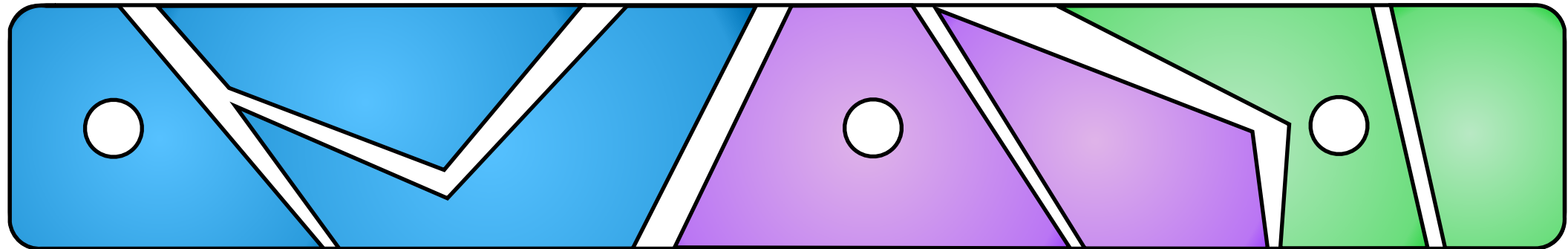
COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER



🎉 Already done 🎉

 Already done 

when the total number of colors is bounded.

 Already done 

when the total number of colors is bounded.

Otherwise...

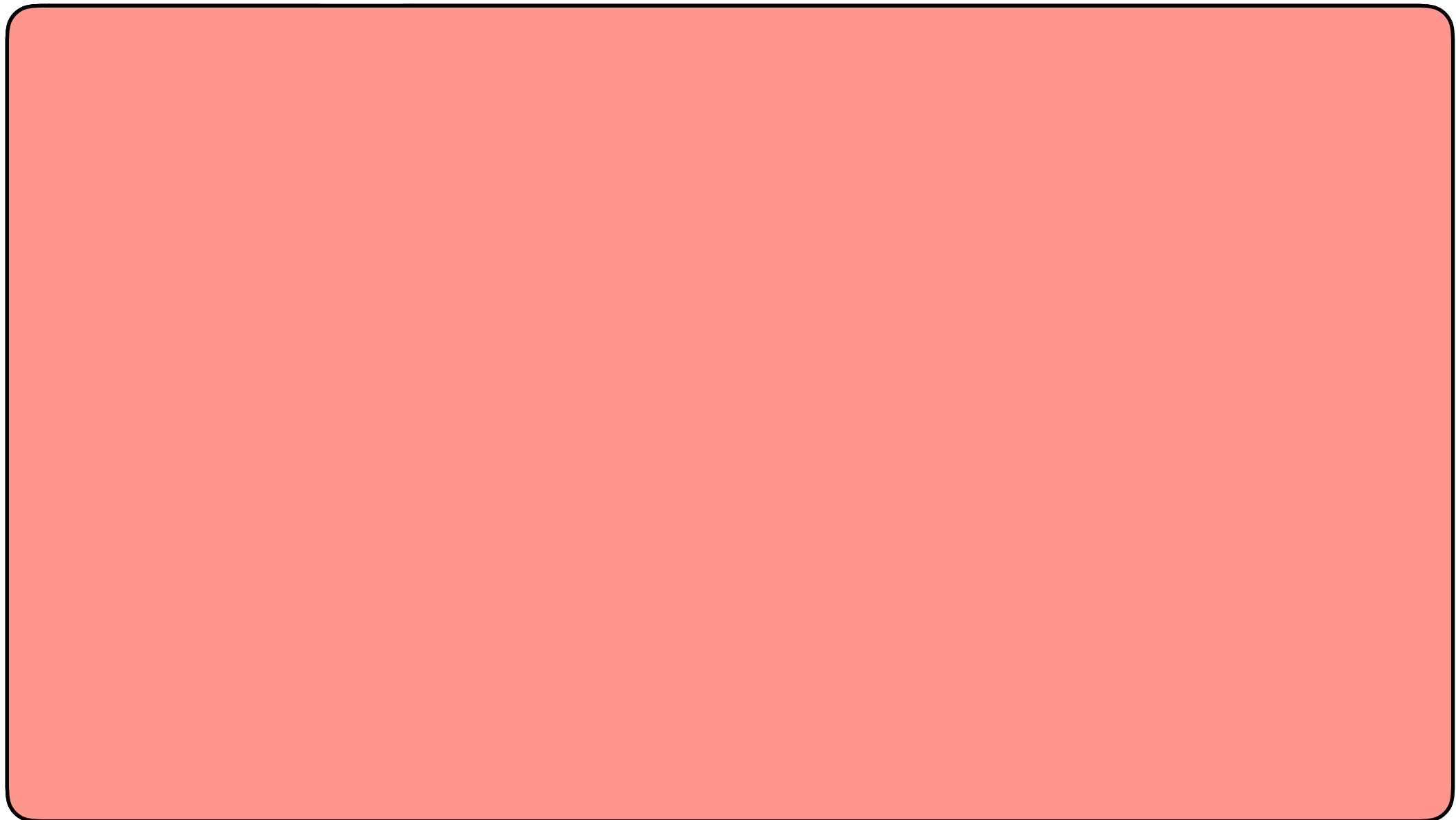
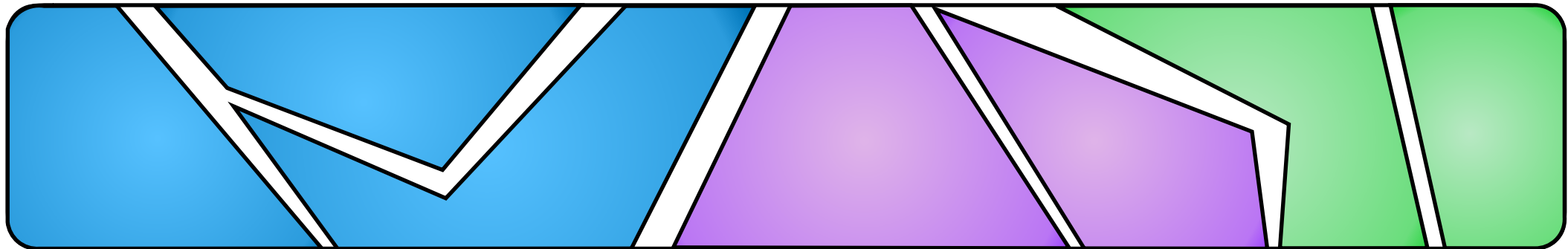
 Already done 

when the total number of colors is bounded.

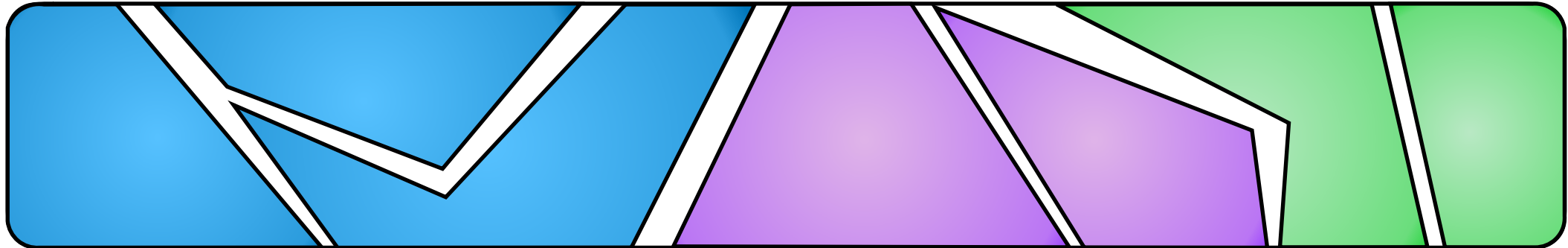
Otherwise...

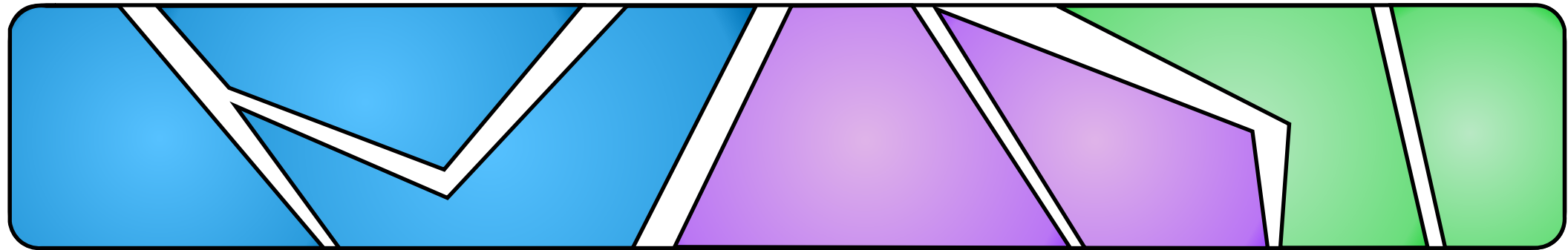


COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER

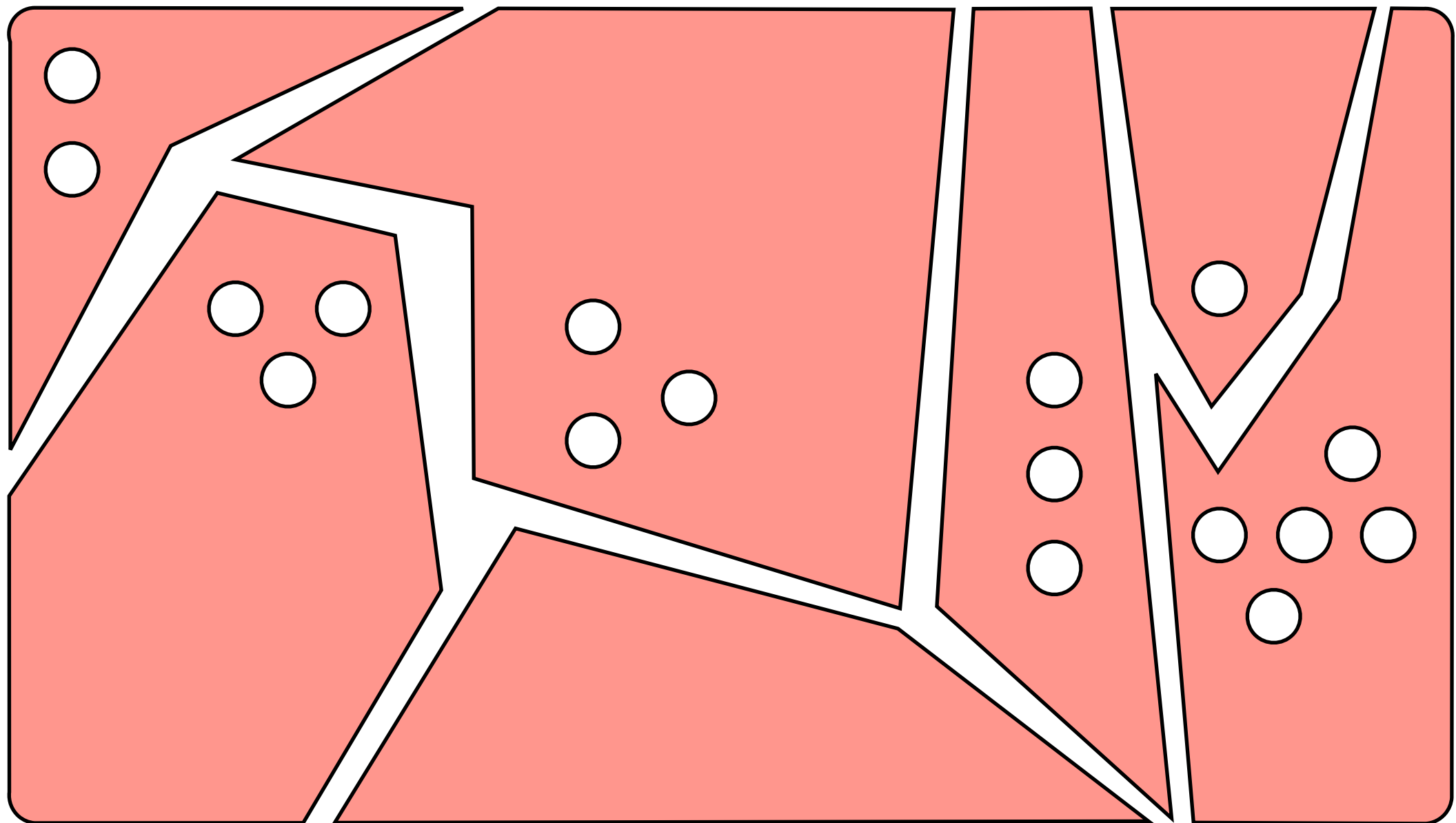
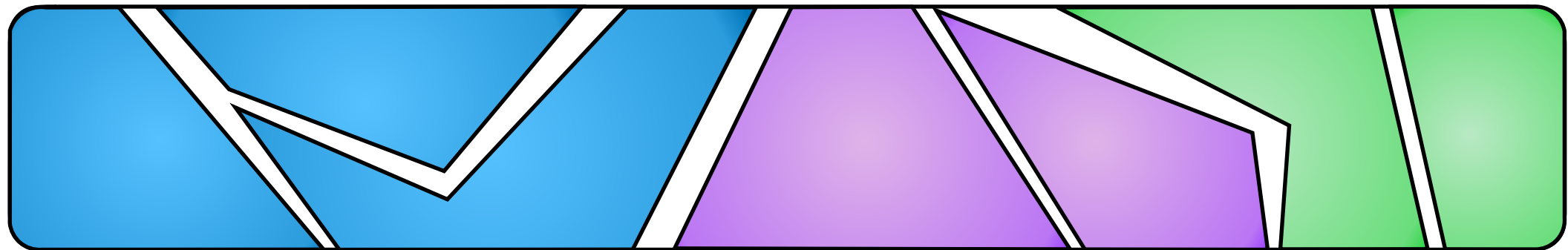




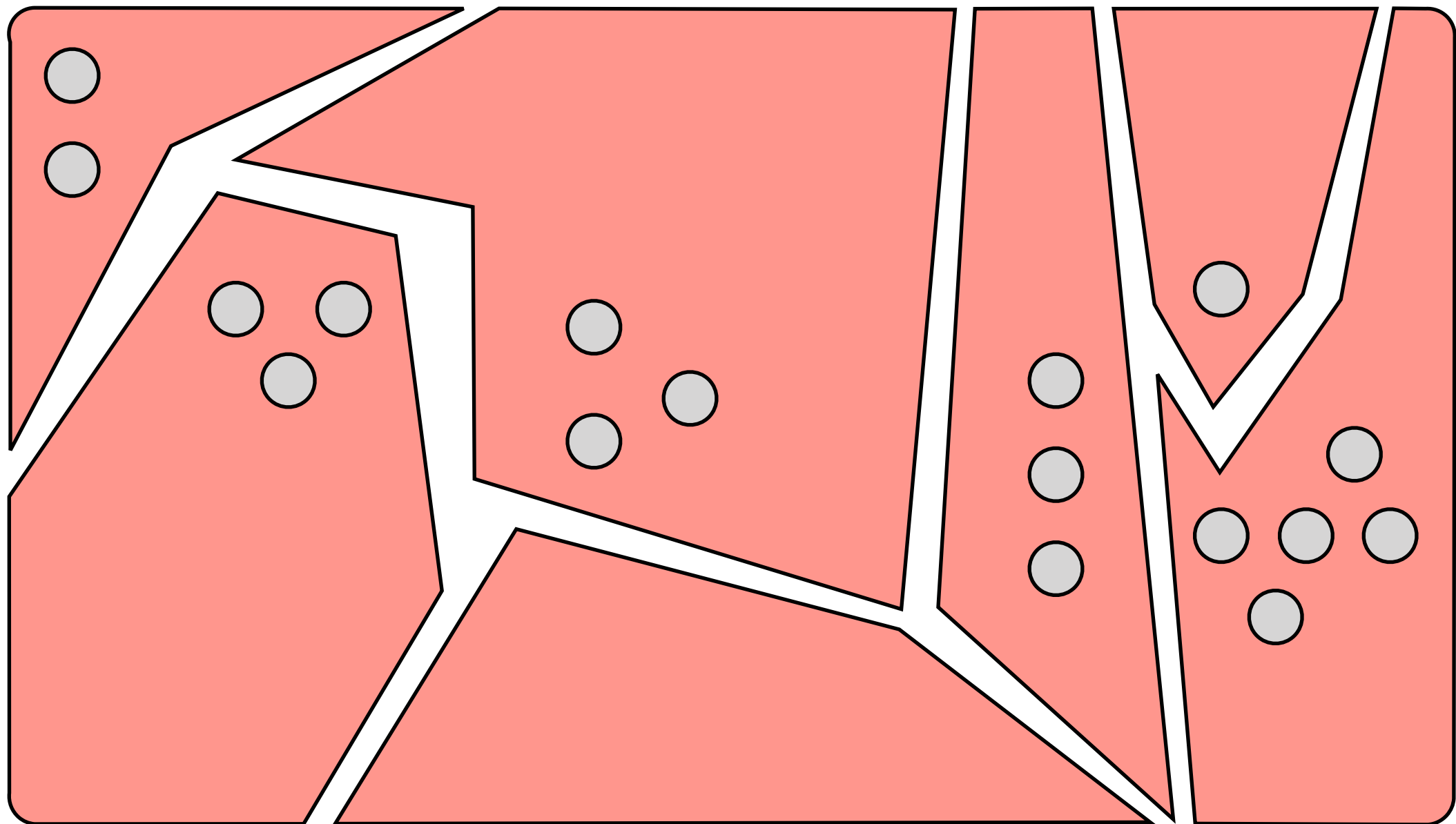
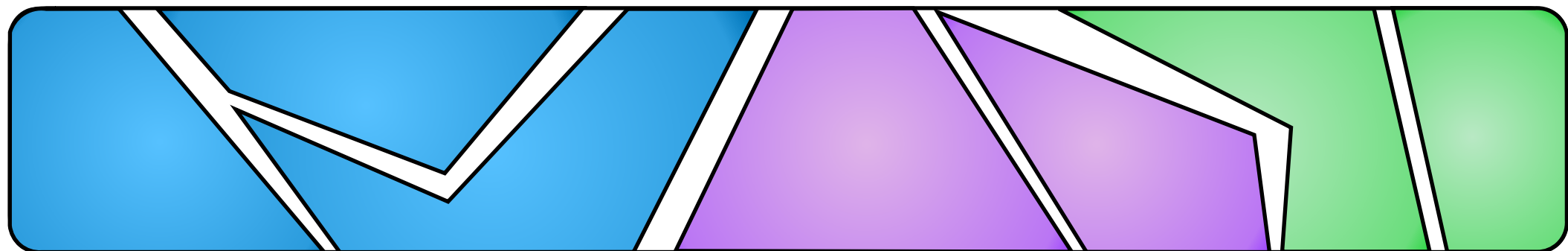
Type II Twins

We need another notion of “same”-ish vertices.

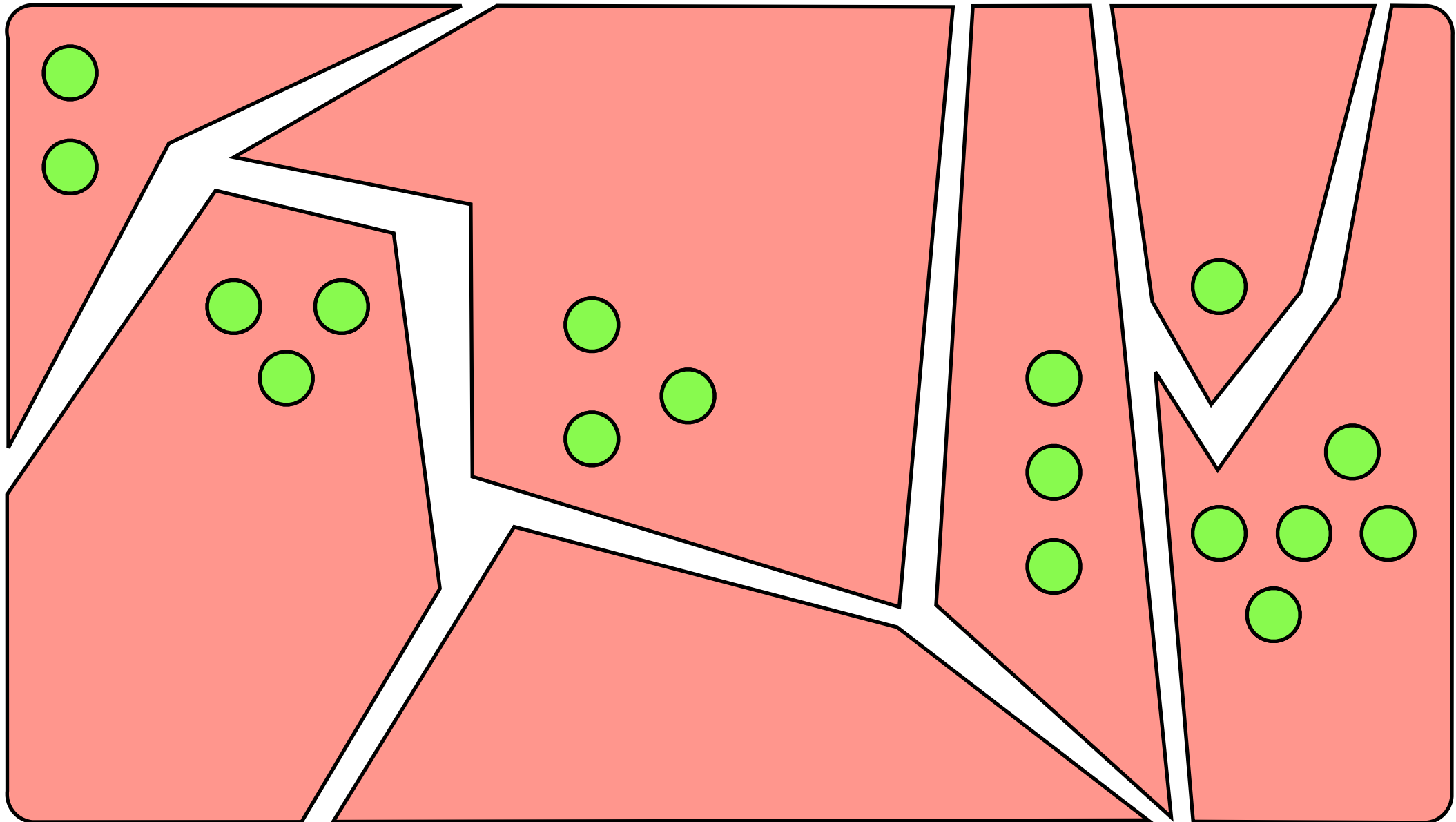
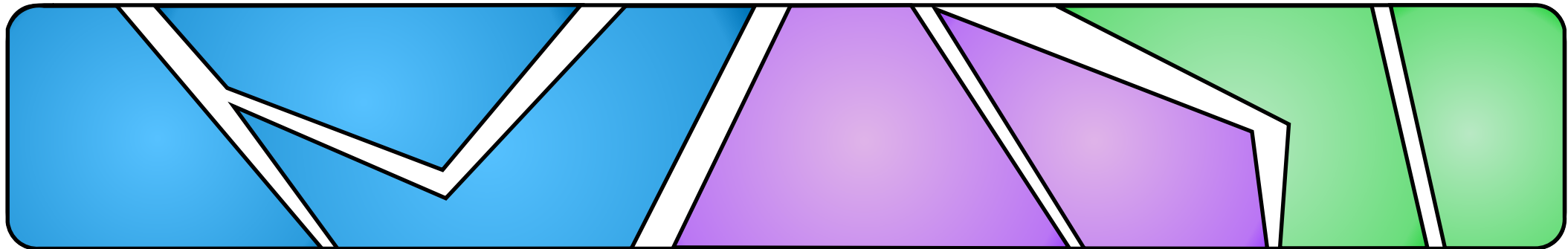
COLORFUL COMPONENTS | VERTEX COVER

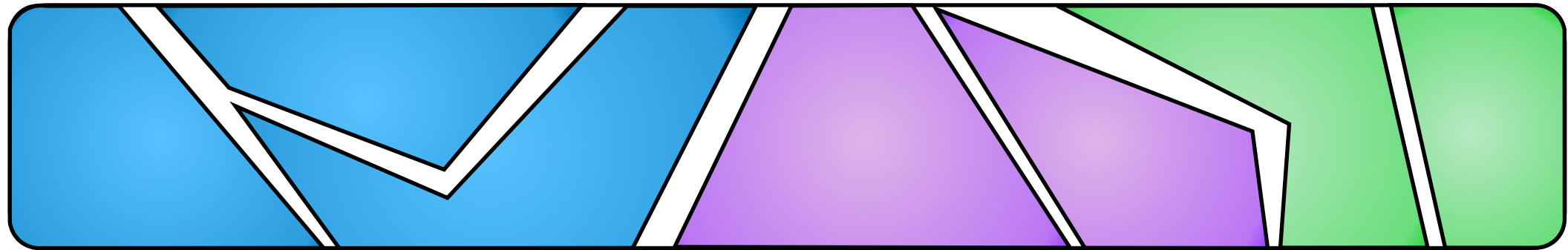


COLORFUL COMPONENTS | VERTEX COVER



COLORFUL COMPONENTS | VERTEX COVER





The third step



extra copies of vertices with the same signature.

what we want

G — A vertex-colored graph.

T — A vertex cover of size t .

F — A subset of at most k edges.

$G \setminus F$ — A graph with colorful components.

...and that gives us **what we wanted**

G — A vertex-colored graph.

T — A vertex cover of size t .

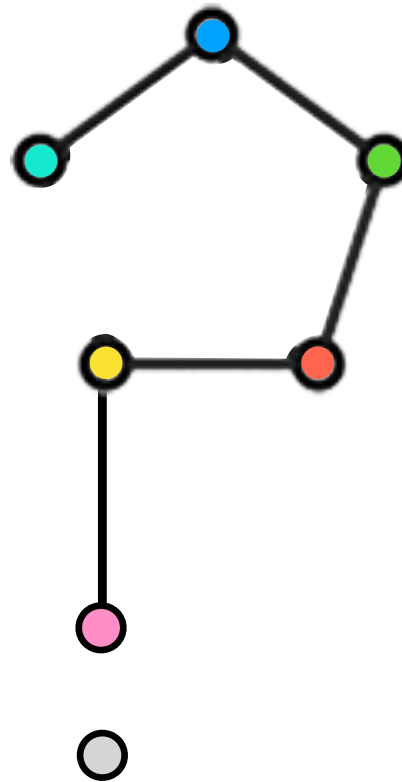
F — A subset of at most k edges.

$G \setminus F$ — A graph with colorful components.





Thank You!



Questions?