The Parameterized Landscape of Labeled Contractions

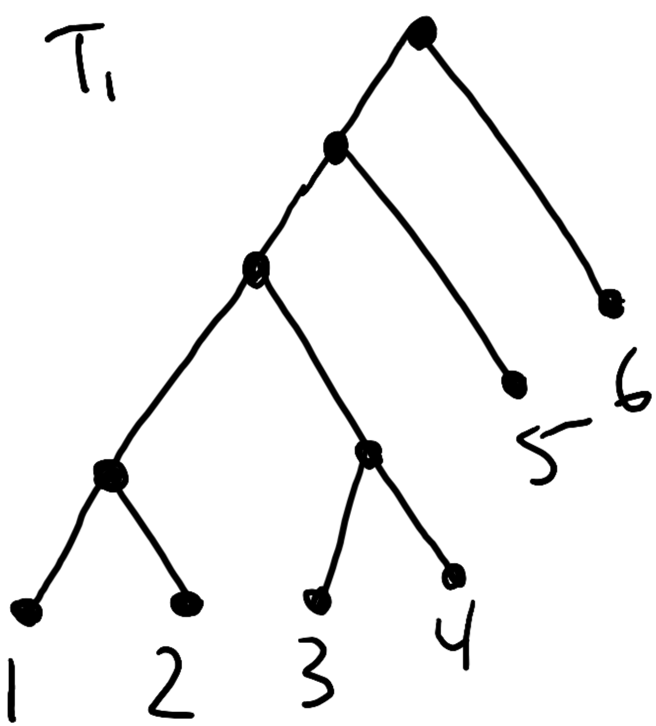
Manuel Lafond **Université de Sherbrooke**

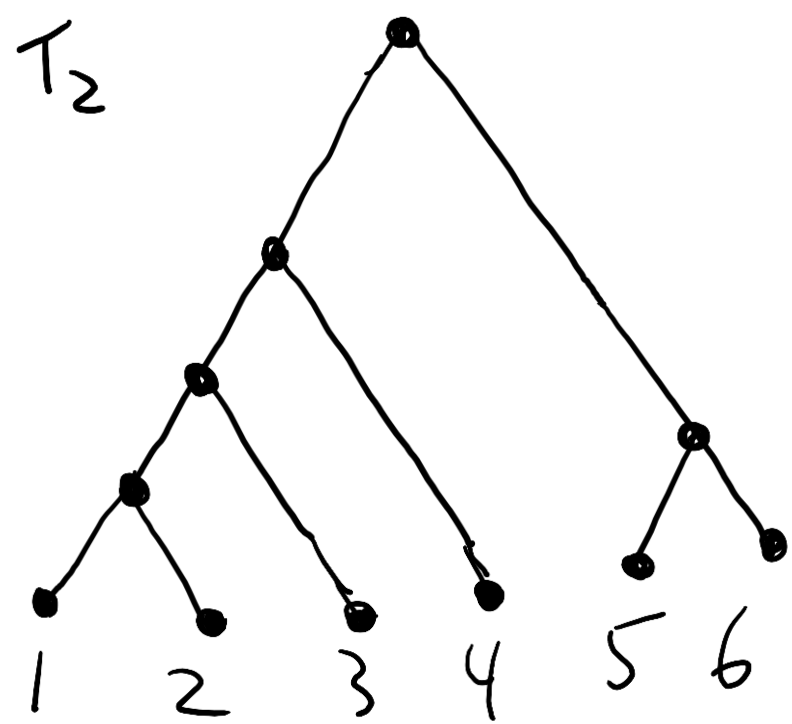
Bertrand Marchand Université de Sherbrooke

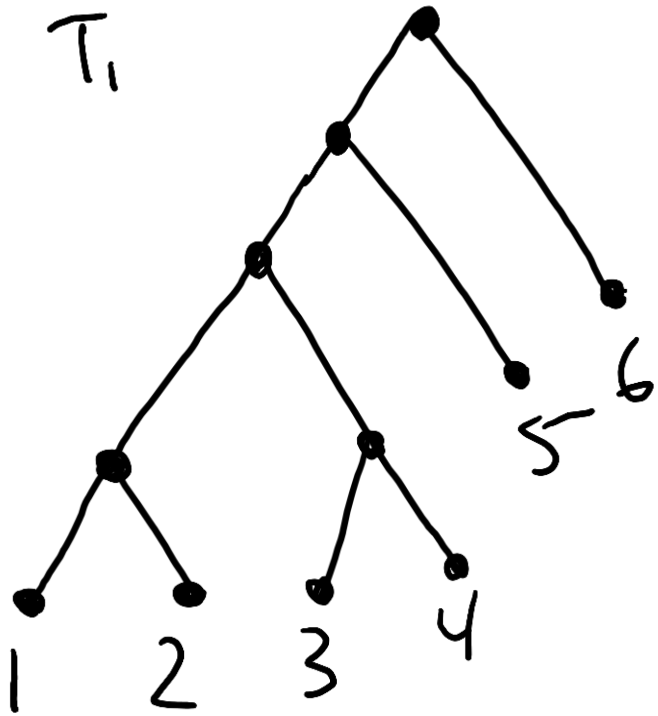
In this talk

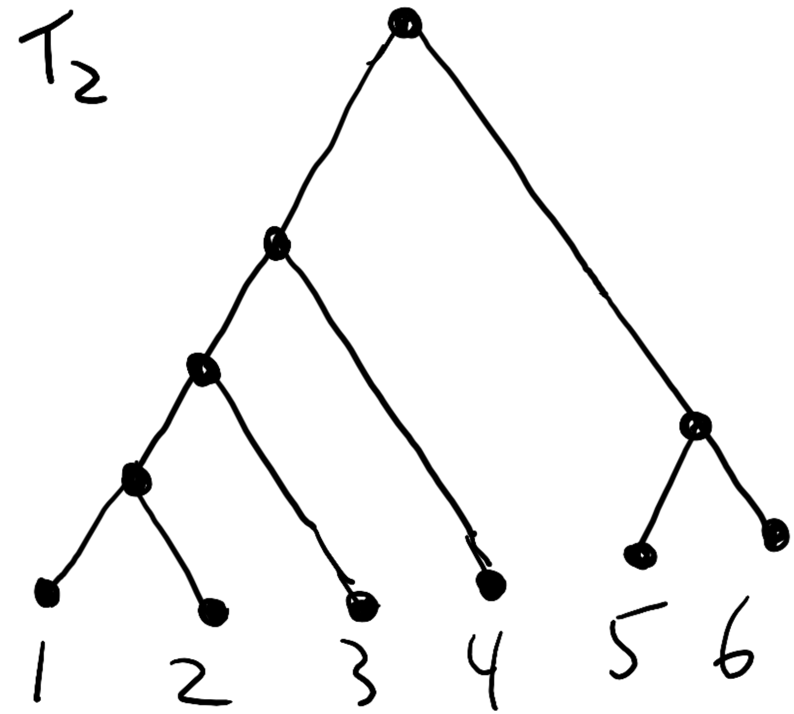
1) Contractions and motivations in Computational Biology
2) Problems: Max Common Labeled Contractions and Labeled Contractibility
3) FPT, W[1]-hardness, and para-NP-hardness

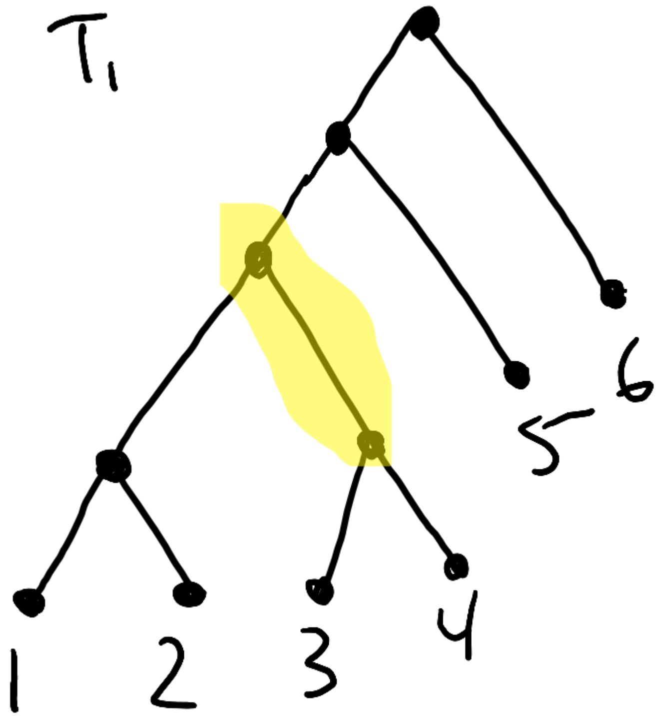
- In computational biology, we like to compare (rooted) evolutionary trees.
- Want: distance between trees + common structure.
- Leaves are uniquely labeled (not internal nodes).

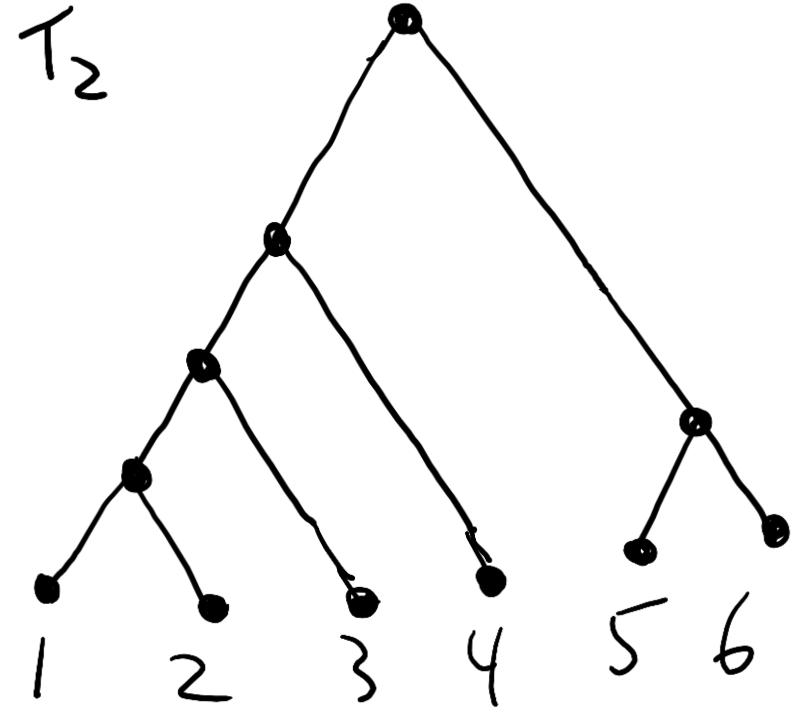


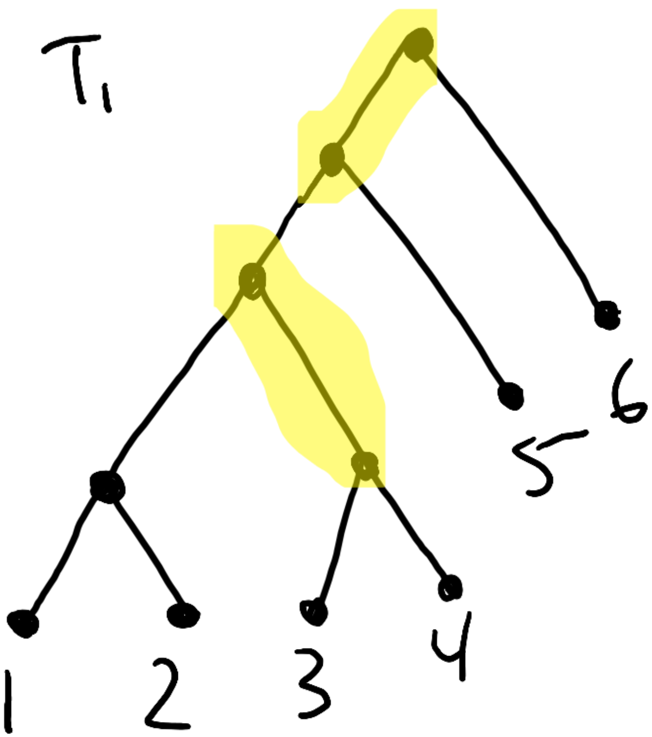


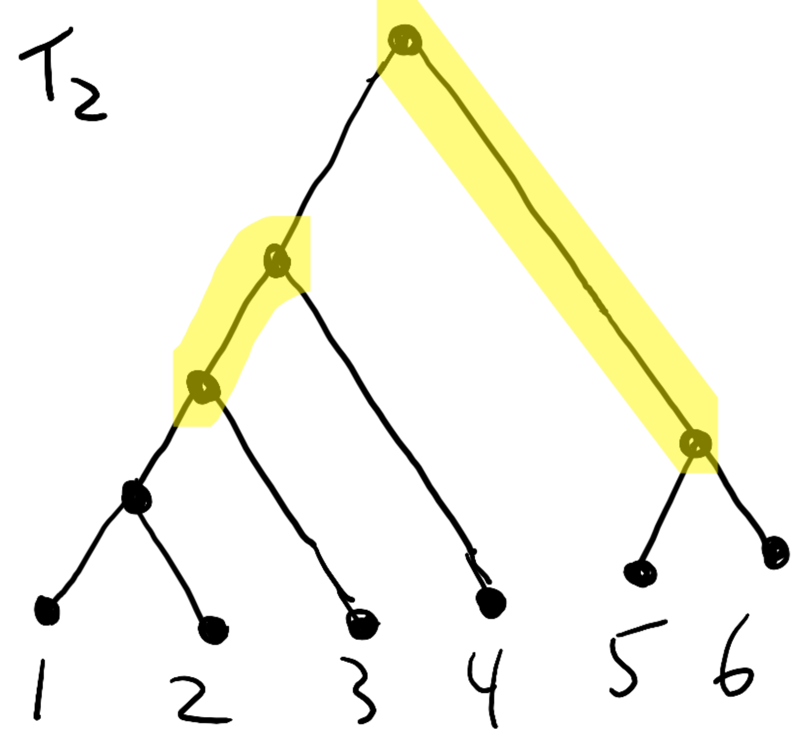

T1
1
2
3
4
5
6

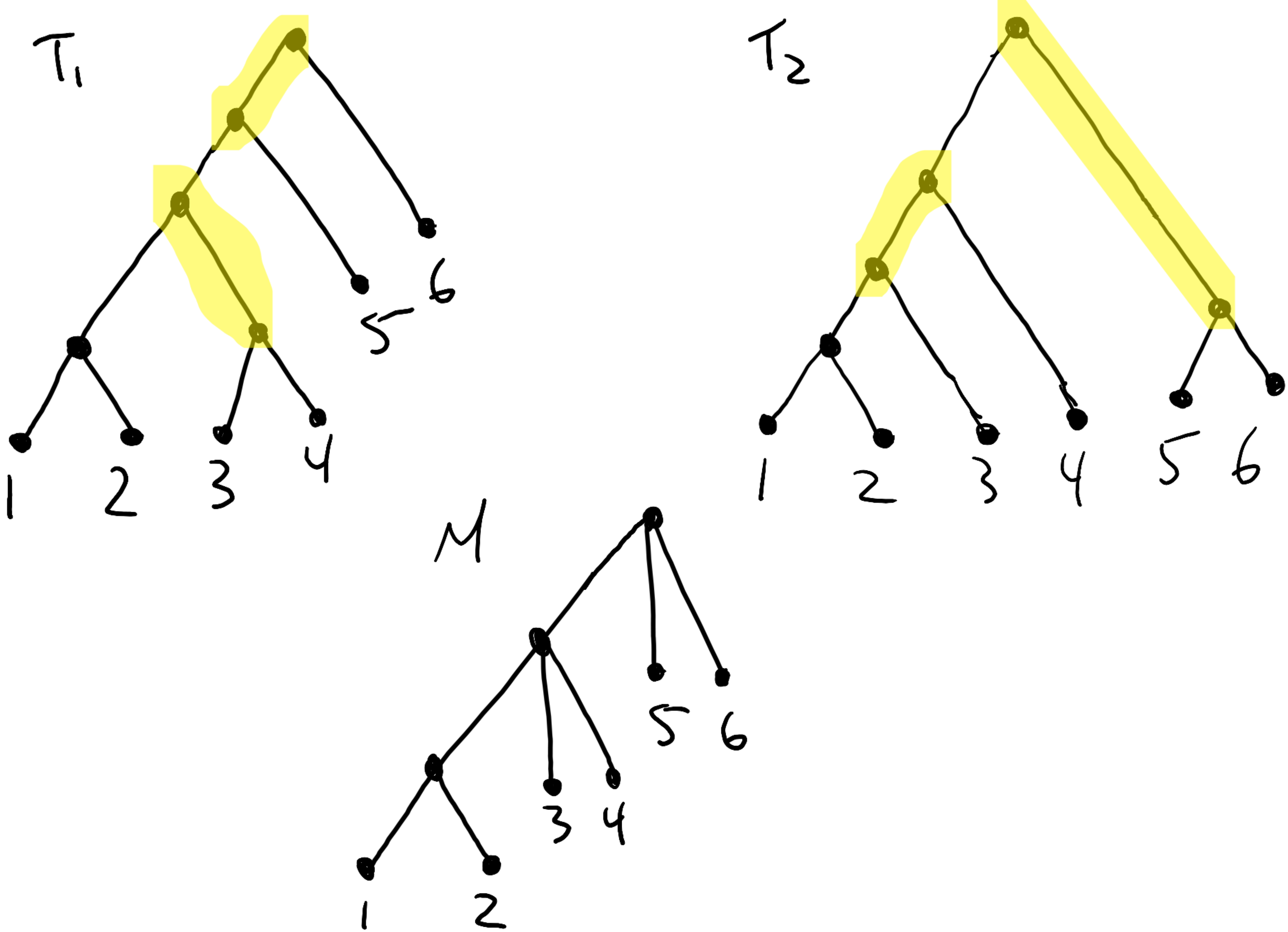

T2
1
2
3
4
5
6


T_1
1
2
3
4
5
6


T_2
1
2
3
4
5
6


T1
1
2
3
4
5
6


T2
1
2
3
4
5
6



T_1
1 2 3 4 5 6
T_2
1 2 3 4 5 6
M
1 2 3 4 5 6

Robinson-Foulds metric

Min # of edge contractions to apply to T_1 and T_2 to obtain the same tree.

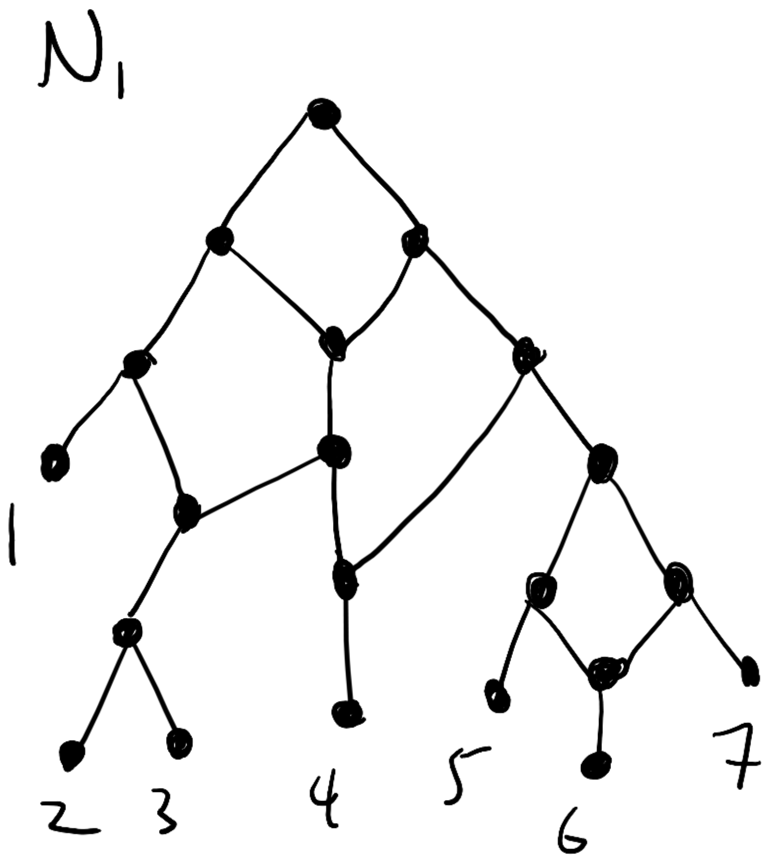
- All leaves must remain leaves.
- That "same tree" is a **common contraction** of max size.
- Can be found in **linear time**.

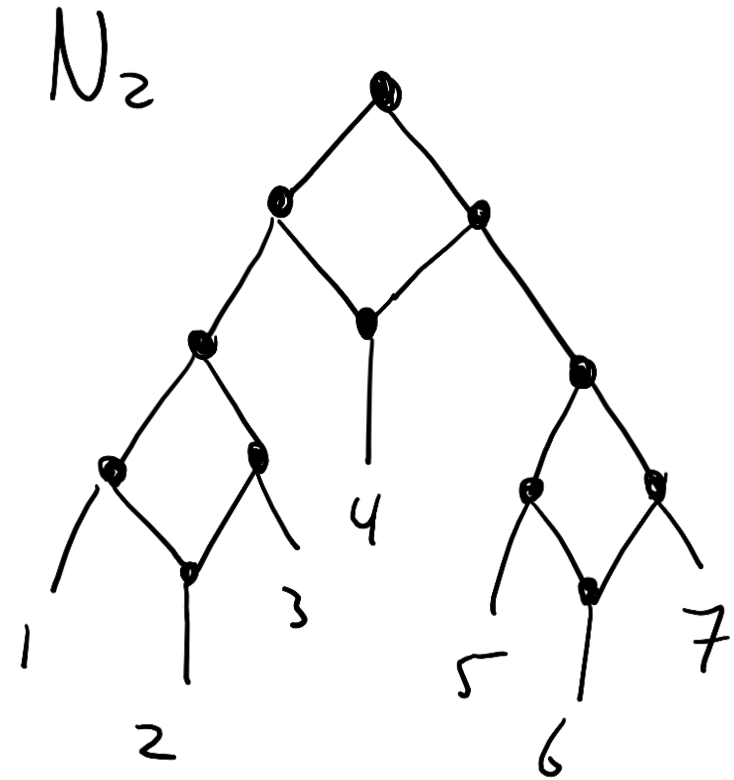
Robinson-Foulds metric

Min # of edge contractions to apply to T_1 and T_2 to obtain the same tree.

- All leaves must remain leaves.
- That "same tree" is a **common contraction** of max size.
- Can be found in **linear time**.

- Trees are nice and easy (when leaves are labeled).
- Attempts at extending to phylogenetic networks
  - Graphs with uniquely labeled leaves.


N_1
1
2
3
4
5
6
7


N_2
1
2
3
4
5
6
7

- Phylogenetic network isomorphism with leaf-label preservation is GI-hard [Cardona et al, 2009].
- Finding max common contraction cannot be easier.
- Leaf labels do not help much...

- Phylogenetic network isomorphism with leaf-label preservation is GI-hard [Cardona et al, 2009].
- Finding max common contraction cannot be easier.
- Leaf labels do not help much...

- Too hard...
- Let's just label **every vertex**!

- We assume that every vertex has a distinguishing identifier.
- We compare sets of vertices and edge directly – no isomorphism.
- Graphs G and H are *equal* if $V(G) = V(H)$ and $E(G) = E(H)$.
  - Deciding equality is trivial.
  - What about finding common contractions?

Labeled contraction

Given an (undirected) edge uv, the contraction (u, v) adds all neighbors of v as neighbors of u, then deletes v. (collapse parallel edges)

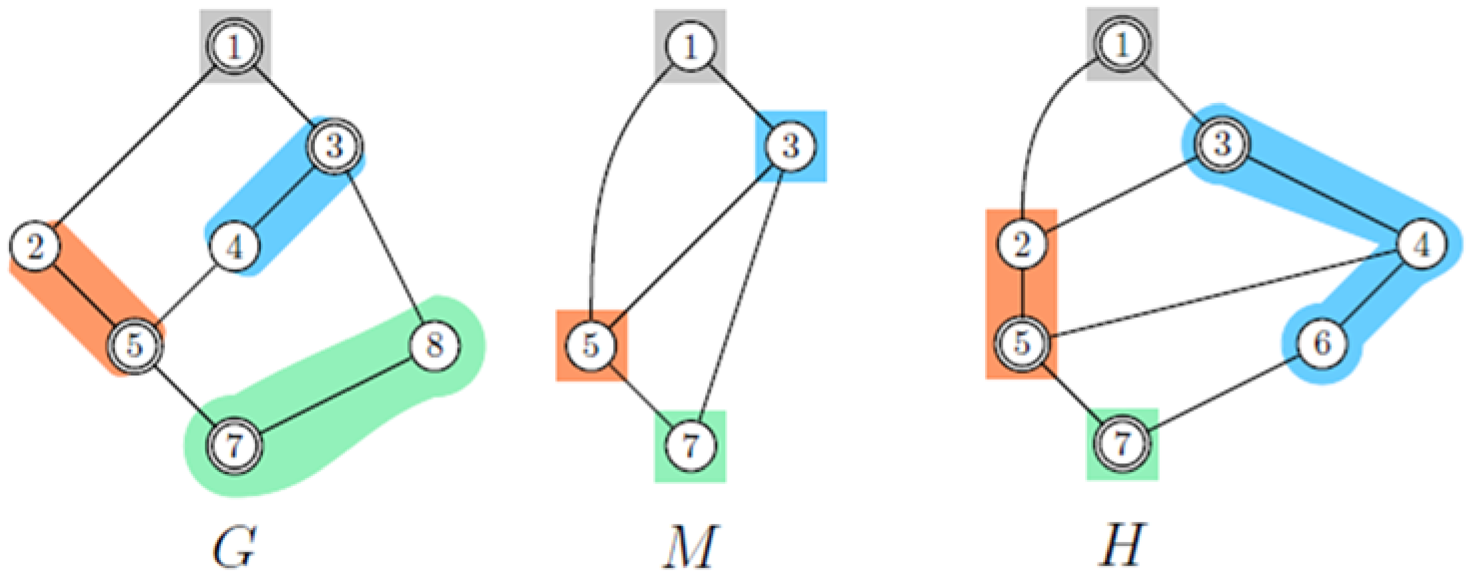
- u is kept => **not the same** result as contracting (v, u)

Labeled contraction of G

Graph obtainable from G by a sequence of labeled contractions.

Common Labeled Contraction (of G and H)

Labeled contraction of both graphs.


1
2
3
4
5
7
8
G
1
3
5
7
M
1
2
3
4
5
6
7
H

Maximum Common Labeled Contraction (MCLC)

Input: graphs G and H, integer s

Question: is there a common labeled contraction of G and H with at least s vertices?

We usually denote k = number of contractions to apply to G and H to achieve a common contraction.

Labeled Contractibility (LC)

Input: graphs G and H

Question: is H a labeled contraction of G?

We usually denote $k = |V(G) \setminus V(H)|$.

Hardness in LC => Hardness in MCLC.

Algorithm for MCLC => Algorithm for LC.

Past algorithmic results:

- On labeled graphs, not much... (?)
- Without labels:
  - Deciding if a **tree** can be contracted into a given **tree** = NP-hard [Matoušek & Thomas, 1992]
  - Contracting a graph $\boldsymbol{G}$ **into a** $\boldsymbol{P_4}$ = NP-hard [Brouwer et al., 1987].
  - Contractibility when $\boldsymbol{G}$ **is chordal**: W[1]-hard in $|V(H)|$ but XP in $|V(H)|$ [Belmonte et al., 2014]
  - Contractibility when $\boldsymbol{G}$ **is planar**: XP in $|V(H)|$ [Kamiński et al., 2010]
  - ...
- Max Common Contractions, not much...

Past algorithmic results:

- Contracting G into a **target graph class**
  - Contracting G into a **path or tree**: FPT, poly kernel for paths but not trees [Heggernes et al., 2011]
  - Contracting G into a graph of **max-degree $\leq d$**, FPT in $k + d$ but W[2]-hard in k, paraNP-hard in d [Belmonte et al., 2014]
    - also contracting into a 2-degenerate graph is NP-hard + W[1]-hard in k
  - Similar results for contracting G into a graph of **min-degree $\geq d$** [Golovach et al., 2013]
  - also contracting into a cactus, a grid graph, a bipartite graph, ...
  - ...

Maximum Common Labeled Contraction (MCLC)

Input: graphs G and H, integer s

Question: is there a common labeled contraction of G and H with at least s vertices?

We usually denote k = number of contractions to apply to G and H to achieve a common contraction.

Labeled Contractibility (LC)

Input: graphs G and H

Question: is H a labeled contraction of G?

We usually denote $k = |V(G) \setminus V(H)|$.

Hardness in LC => Hardness in MCLC.

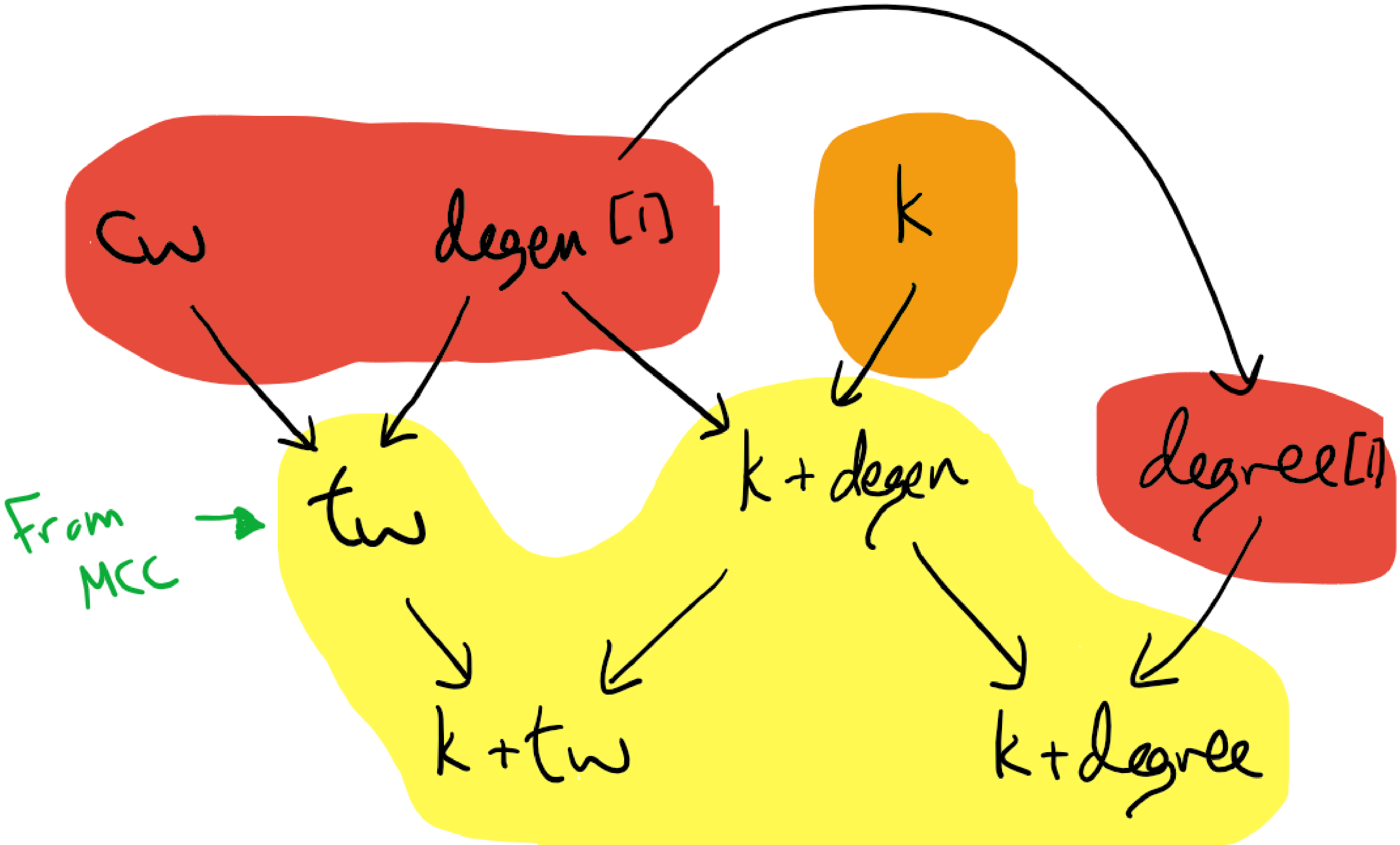
Algorithm for MCLC => Algorithm for LC.

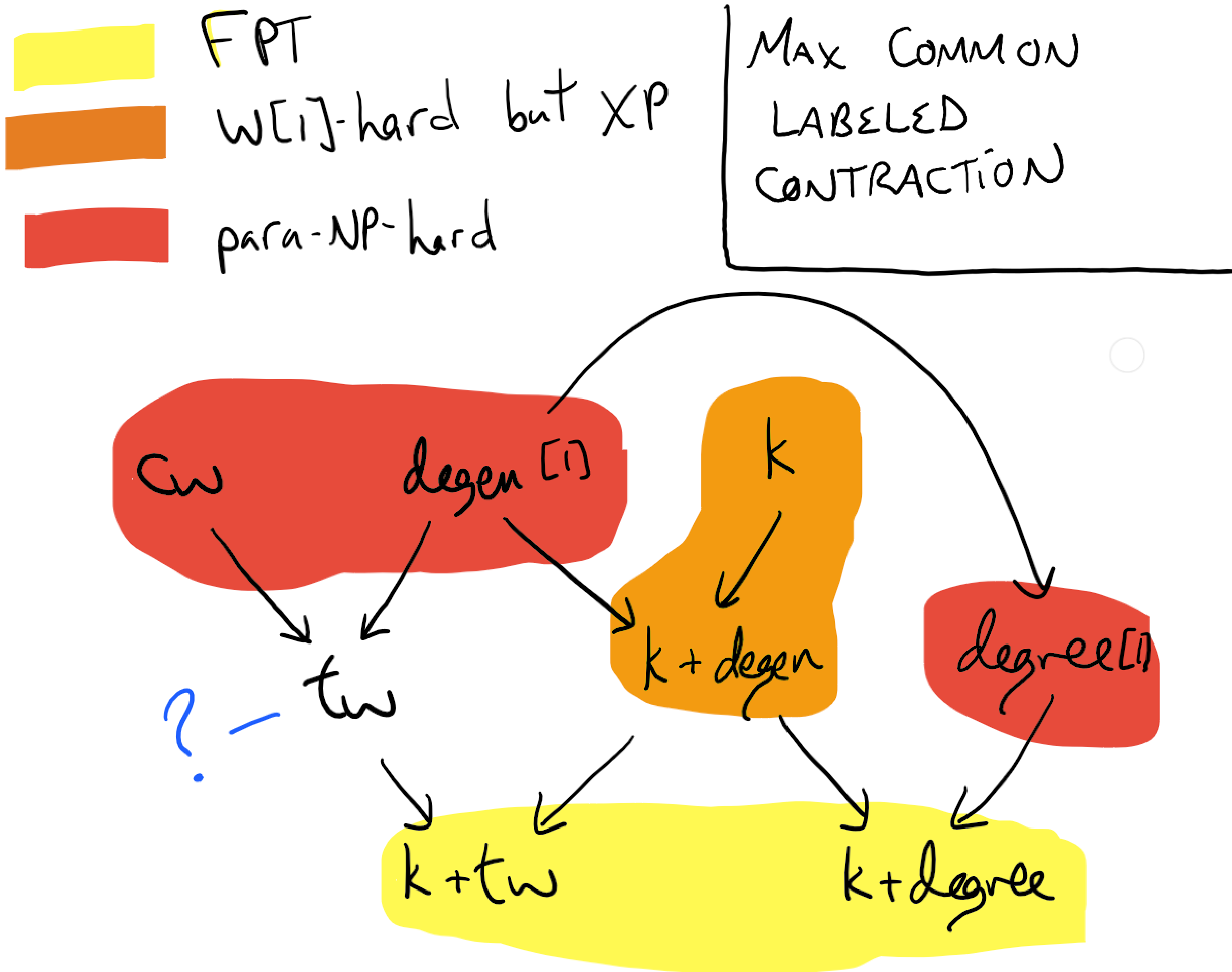
LABELED CONTRACTIBILITY

FPT

W[1]-hard but XP

para-NP-hard




FPT
W[1]-hard but XP
para-NP-hard
MAX COMMON LABELED CONTRACTION
cw
degen [1]
k
k + degen
degree[1]
tw
?
k + tw
k + degree

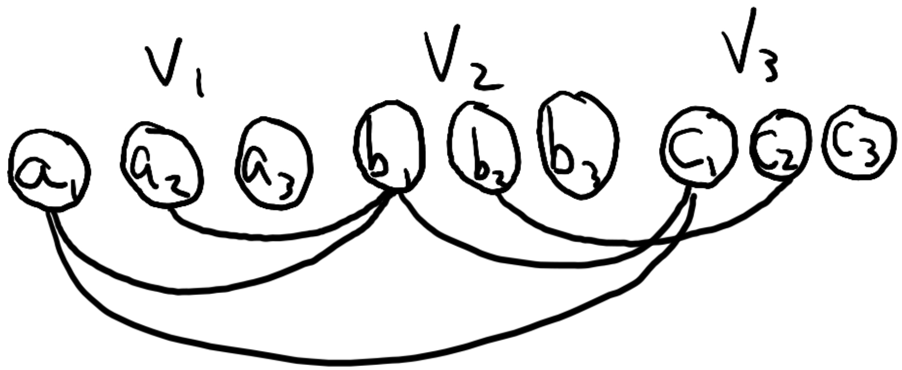
Theorem: Labeled Contractibility is W[1]-hard in parameter $k = |V(G) \setminus V(H)|$.

Relatively standard reduction from Multicolored Clique.

Multicolored Clique

Input: graph $G_C = (V_1 \cup \cdots \cup V_k, E_C)$

Question: is there a clique with one vertex per V_i?

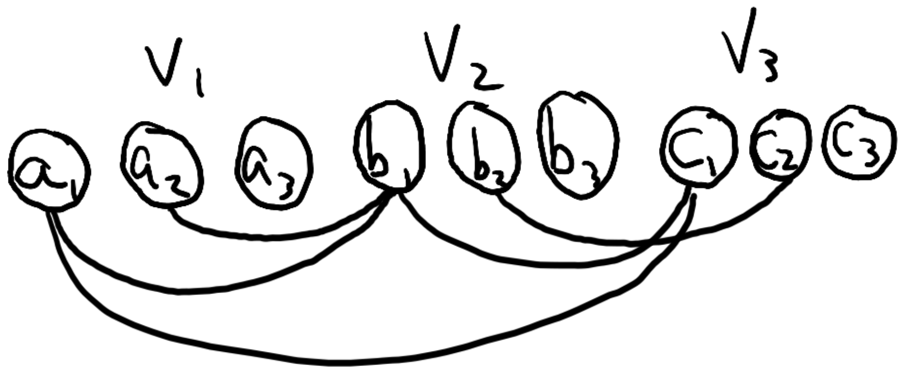


Multicolored Clique

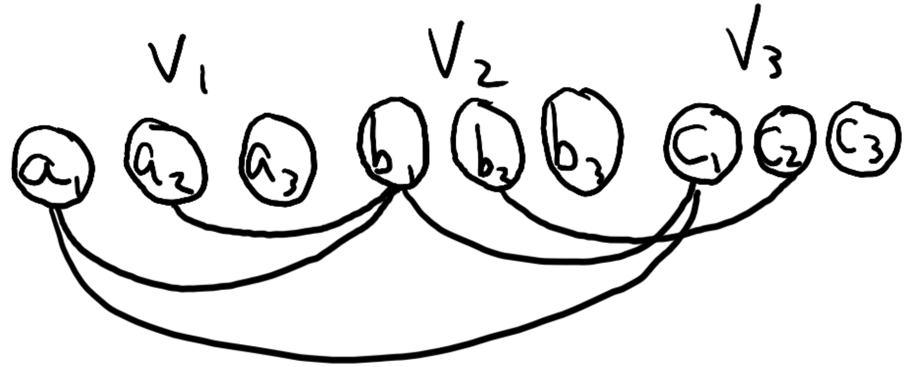
Input: graph $G_C = (V_1 \cup \cdots \cup V_k, E_C)$

Question: is there a clique with one vertex per V_i?

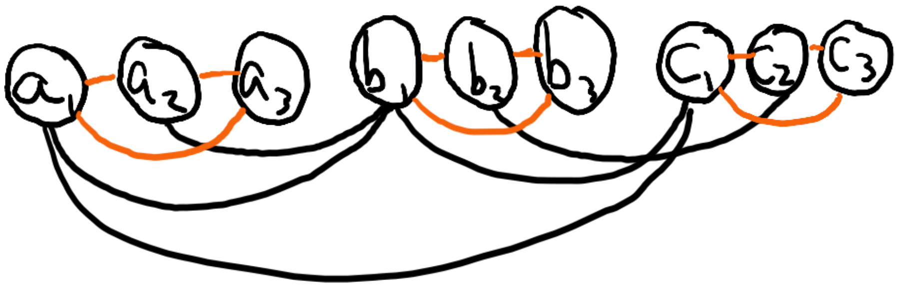
We assume wlog that each V_i has an isolated vertex.



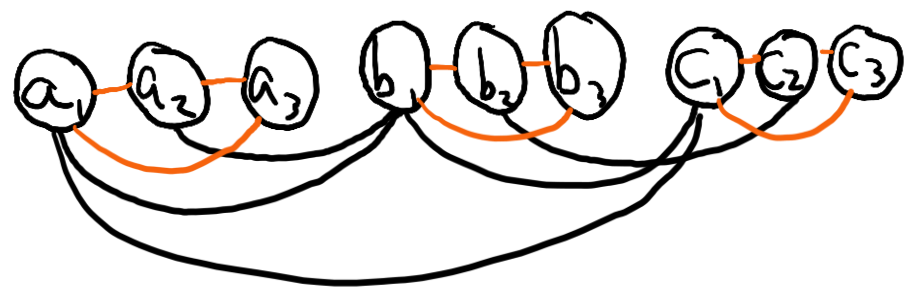
Gc



V1
V2
V3
a1
a2
a3
b1
b2
b3
c1
c2
c3



a1
a2
a3
b1
b2
b3
c1
c2
c3

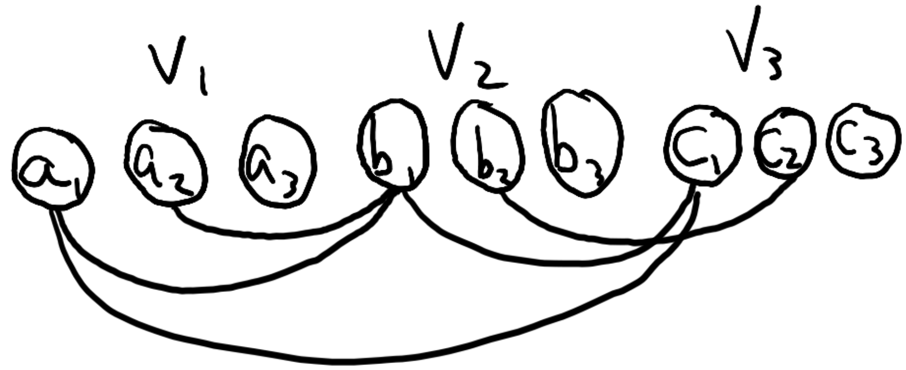


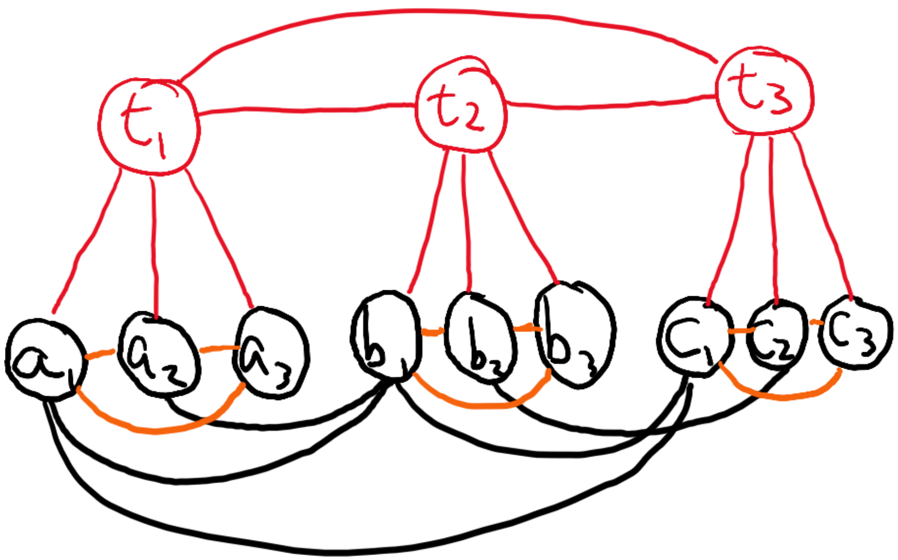
a1
a2
a3
b1
b2
b3
c1
c2
c3

G

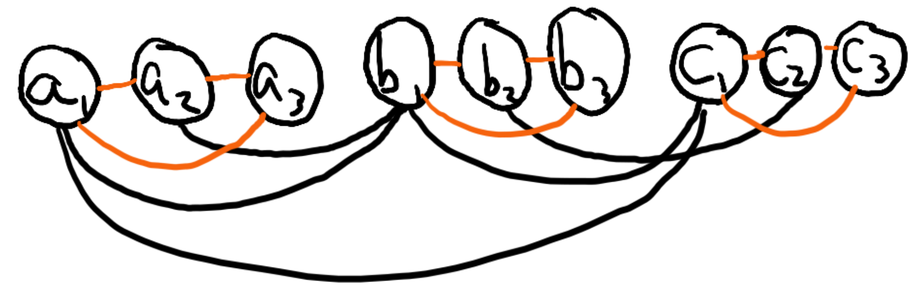
H

G_c


V_1
V_2
V_3
a_1 a_2 a_3 b_1 b_2 b_3 c_1 c_2 c_3

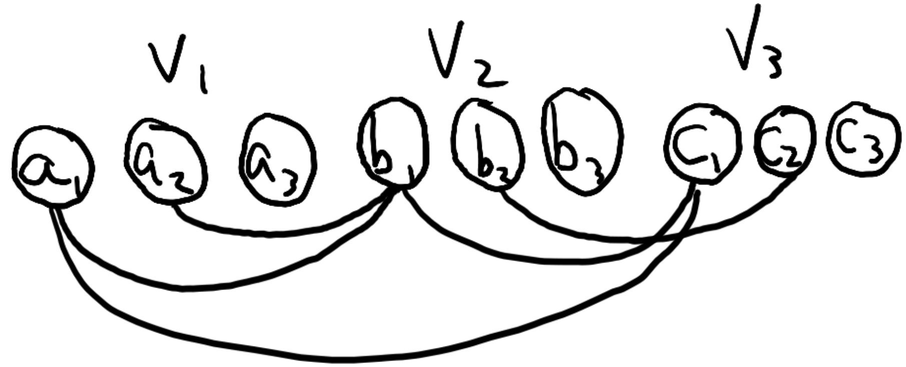

t_1 t_2 t_3
a_1 a_2 a_3 b_1 b_2 b_3 c_1 c_2 c_3

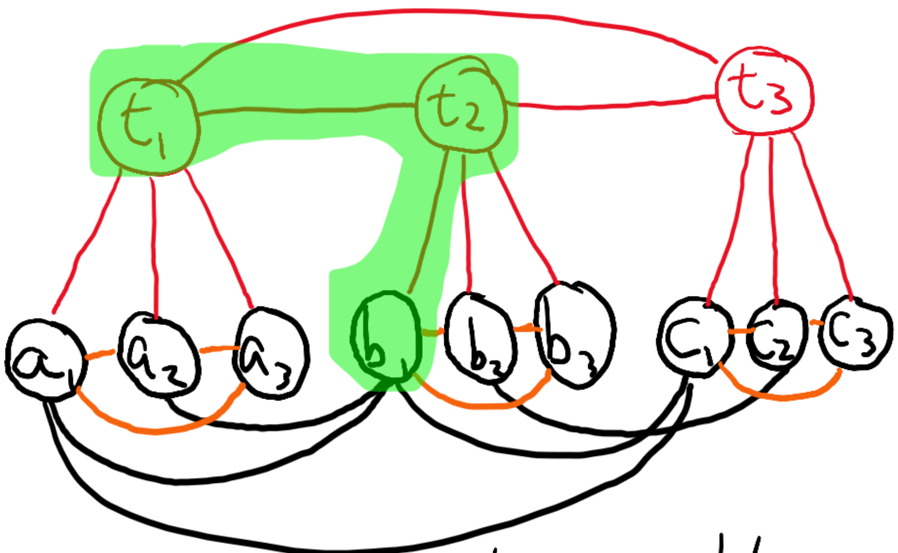
G

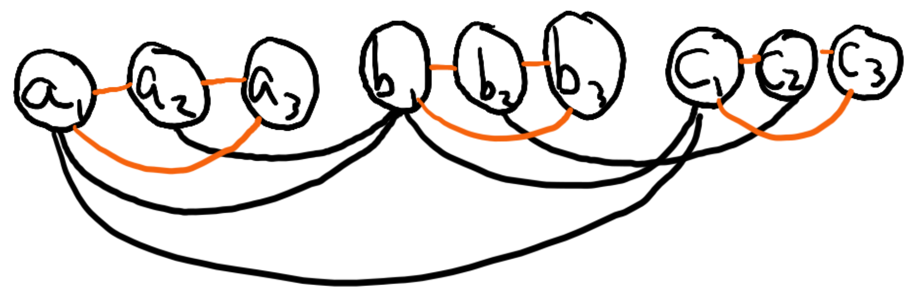

a_1 a_2 a_3 b_1 b_2 b_3 c_1 c_2 c_3

H

G_c





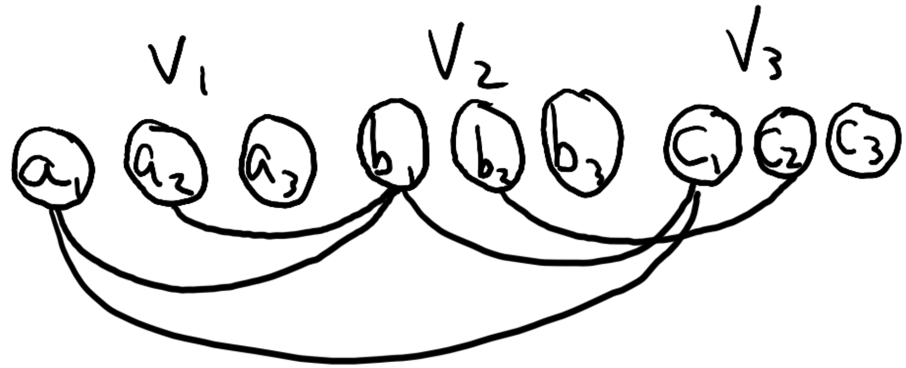


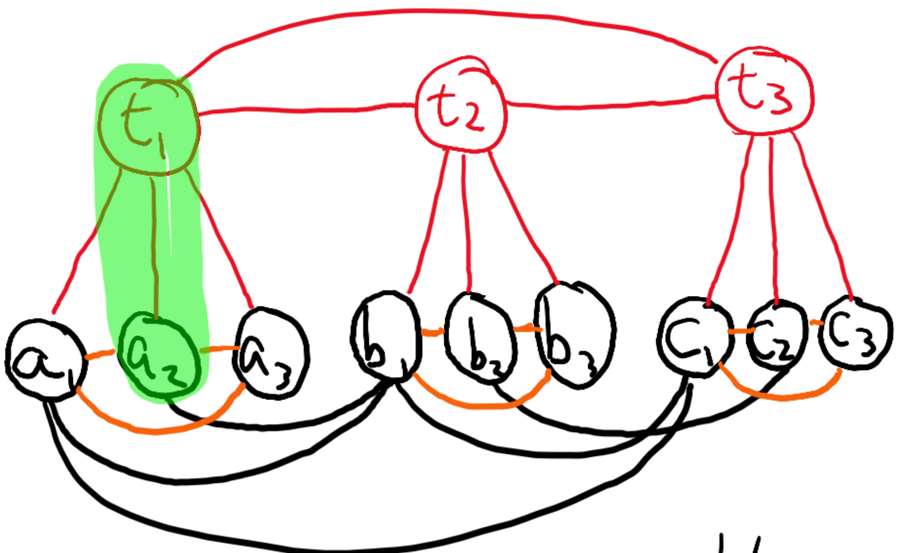
not possible:
makes b_1 a
neighbor of V_1

G

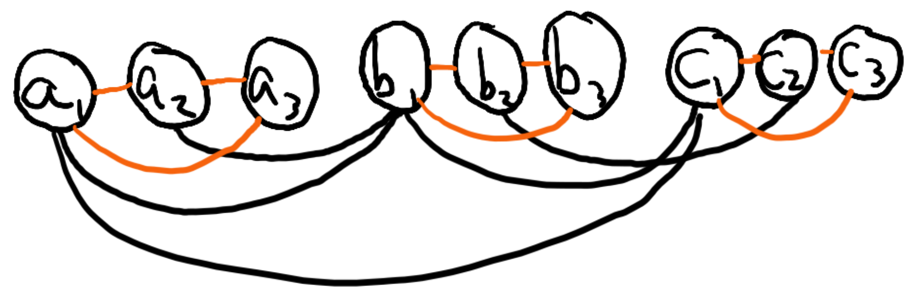
H

G_c





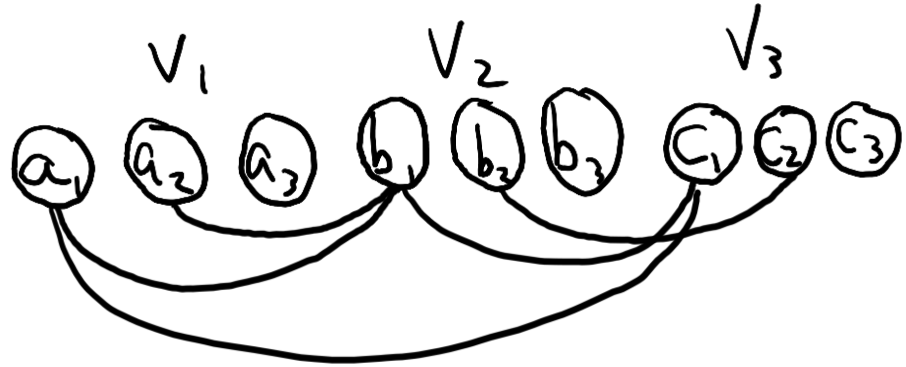
possible

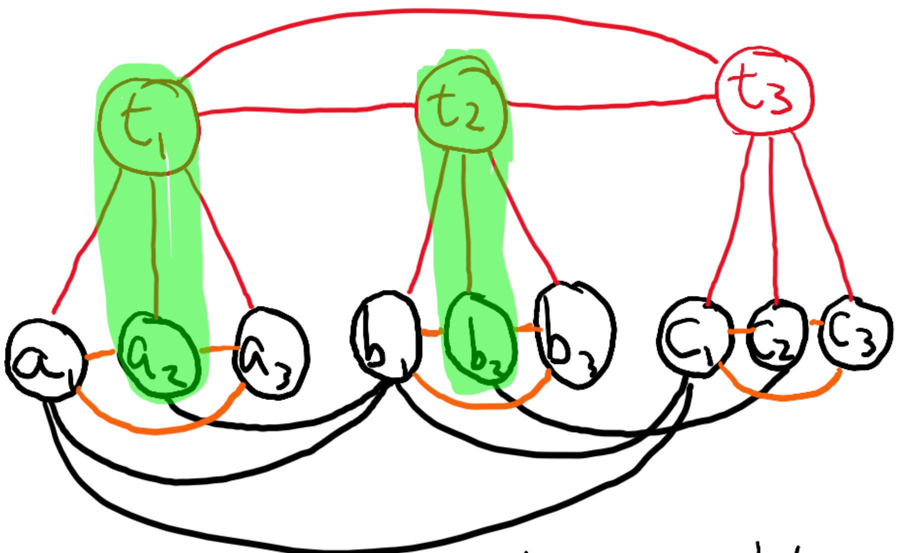


G

H

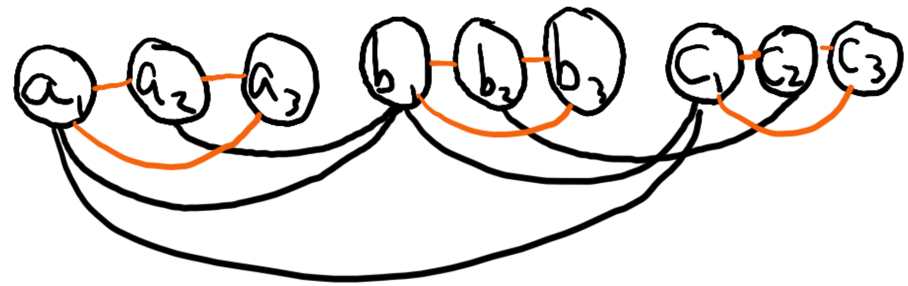
G_c





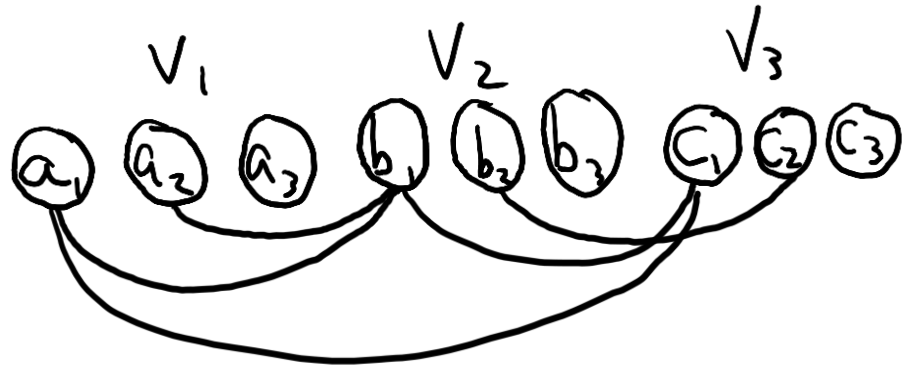
G

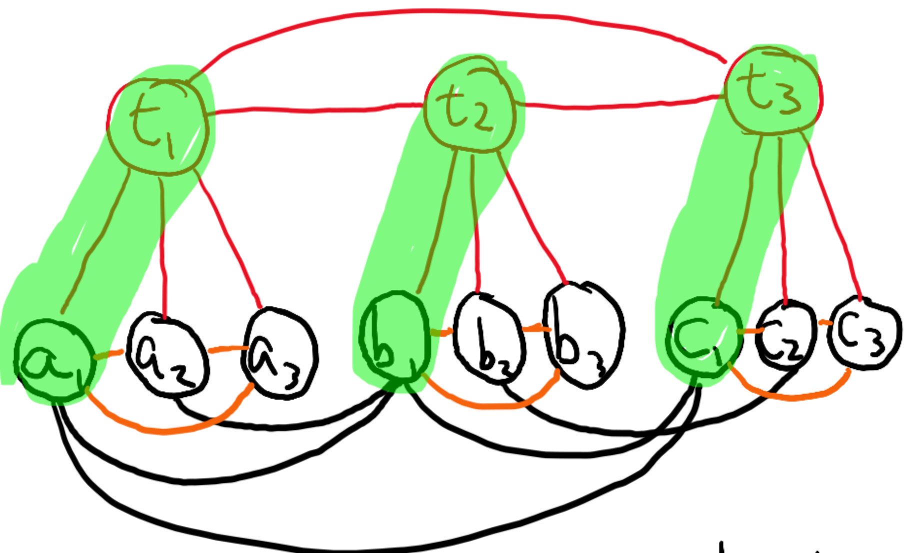
not possible:
creates accidental
edge $a_2 b_2$



H

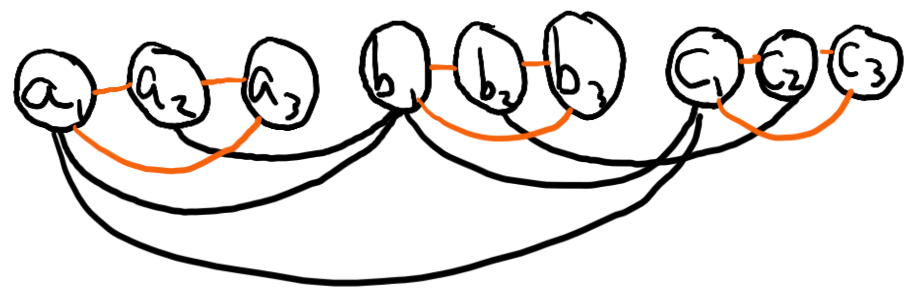
G_c





G

t_i's must choose v_i's of a multicolored clique



H

Theorem: Labeled Contractibility is NP-hard on graphs of clique-width 4.

Slightly less standard reduction from Unary Perfect Bin Packing.

Unary Perfect Bin Packing

Input: multiset of unary-encoded integers $A = \{a_1, \dots, a_n\}$, bin capacity C, bin count k.

Question: can we assign items of A to k bins so that they each sum to exactly C.

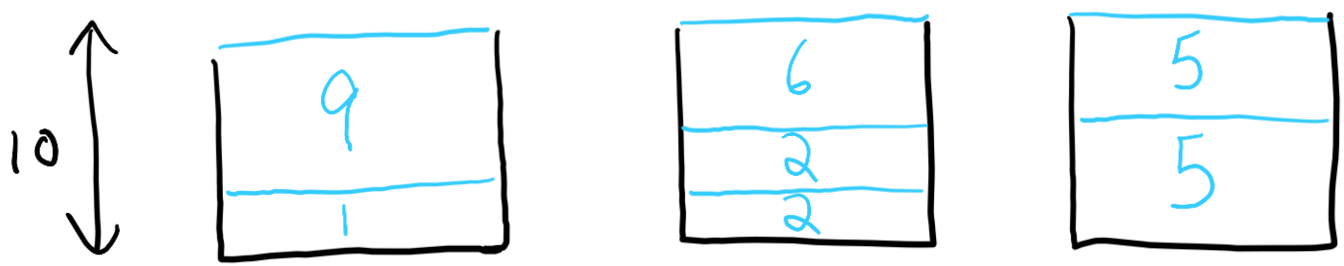
$A = \{1, 2, 2, 5, 5, 6, 9\}$ $C = 10$ $k = 3$

Unary Perfect Bin Packing

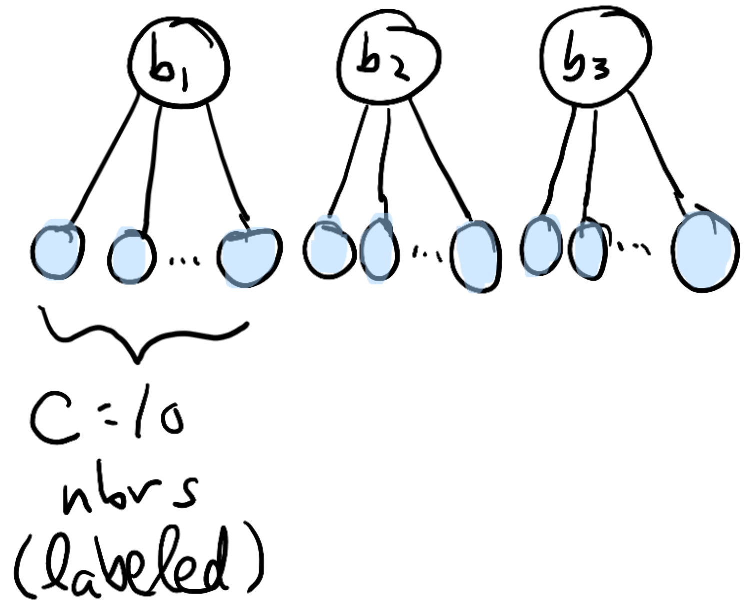
Input: multiset of unary-encoded integers $A = \{a_1, \ldots, a_n\}$, bin capacity C, bin count k.

Question: can we assign items of A to k bins so that they each sum to exactly C.

$A = \{1, 2, 2, 5, 5, 6, 9\}$ $\quad C = 10$ $\quad k = 3$

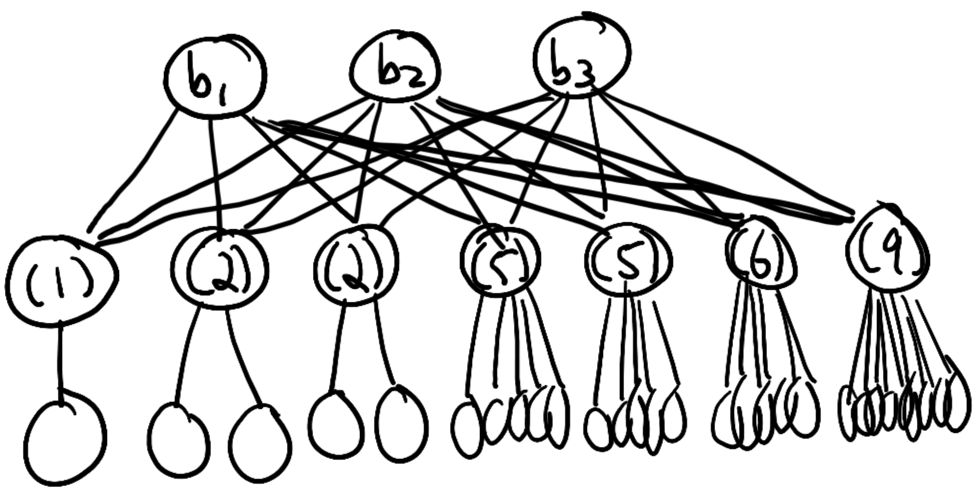


$A = \{1, 2, 2, 5, 5, 6, 9\}$ $\quad C = 10 \quad k = 3$

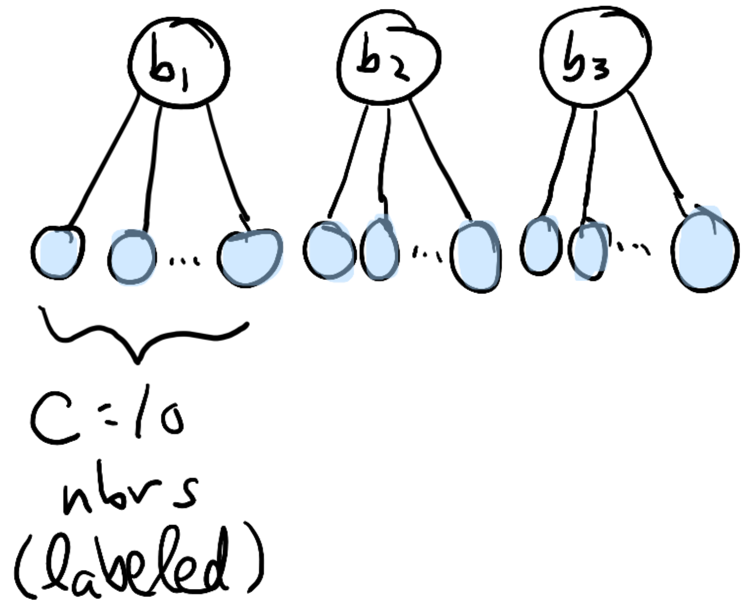


H

$A = \{1, 2, 2, 5, 5, 6, 9\}$ $\quad C = 10 \quad k = 3$

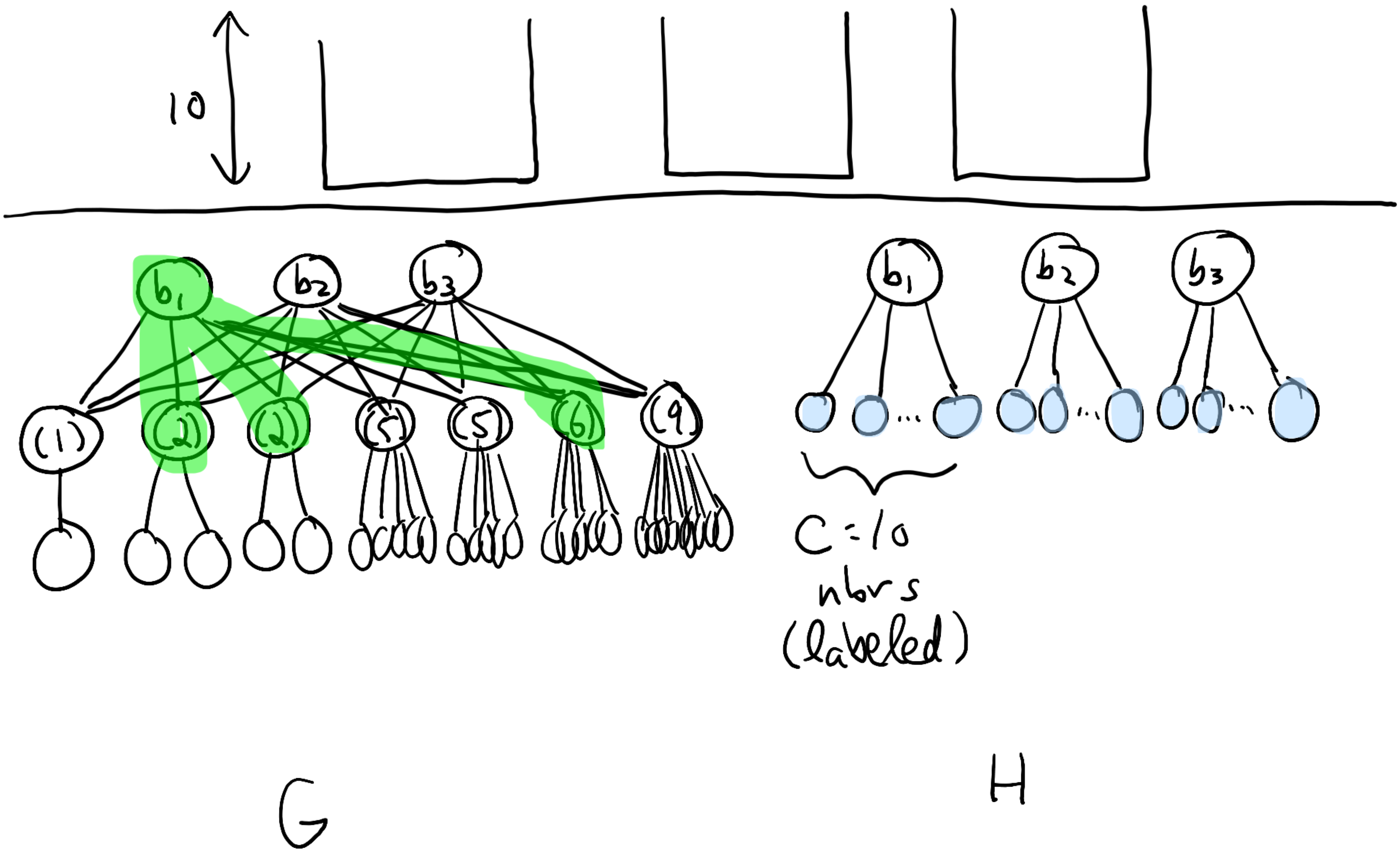


G

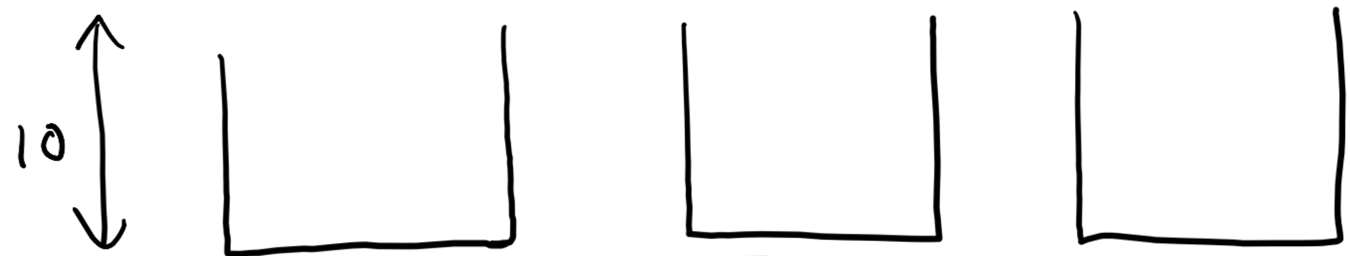


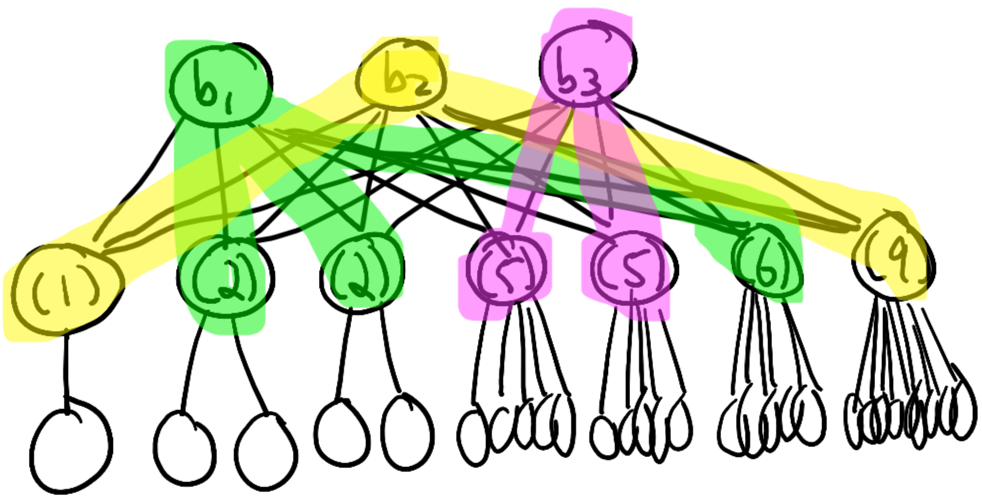
H

$A = \{1, 2, 2, 5, 5, 6, 9\} \quad C = 10 \quad k = 3$

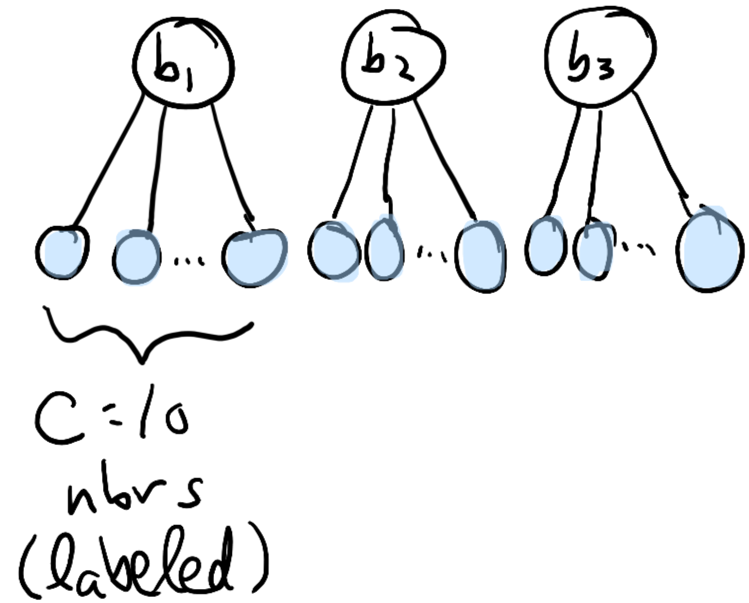


$A = \{1, 2, 2, 5, 5, 6, 9\}$ $\quad C = 10 \quad k = 3$



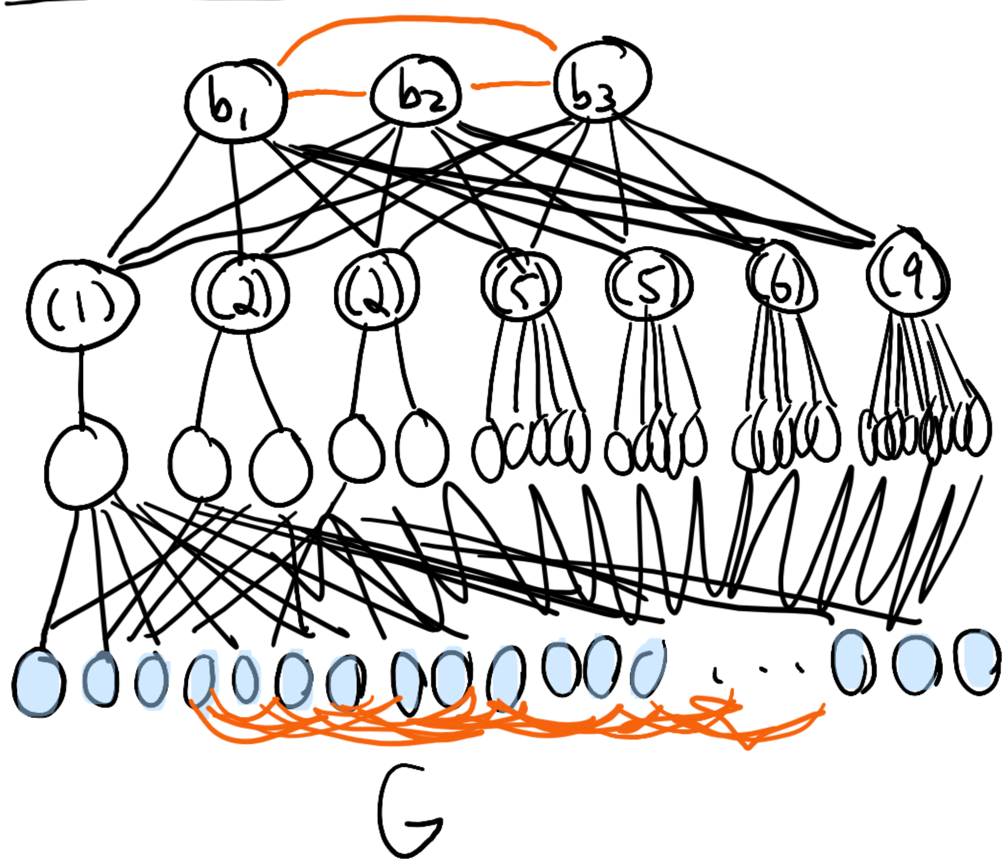


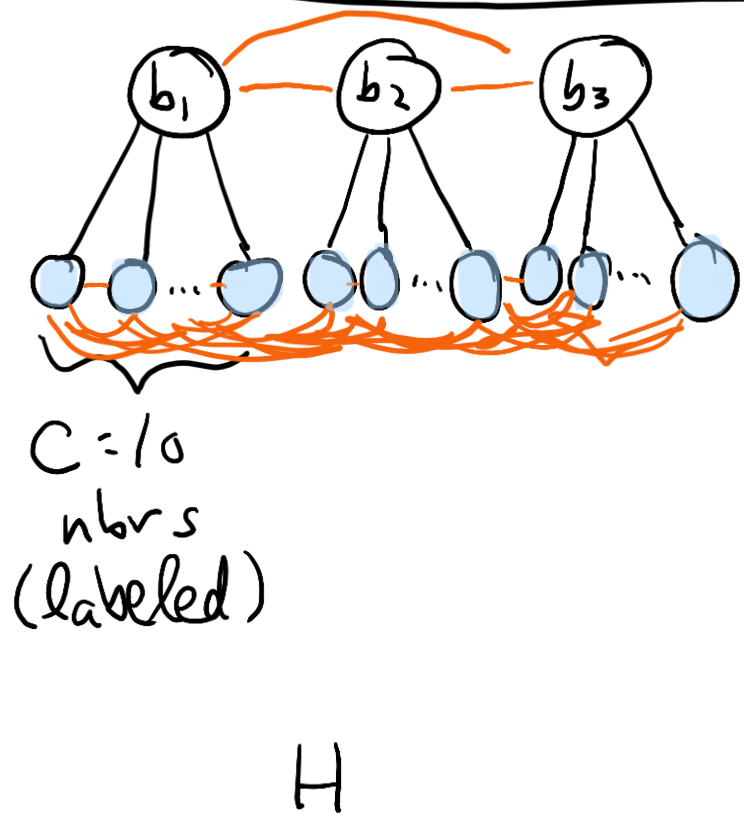
G



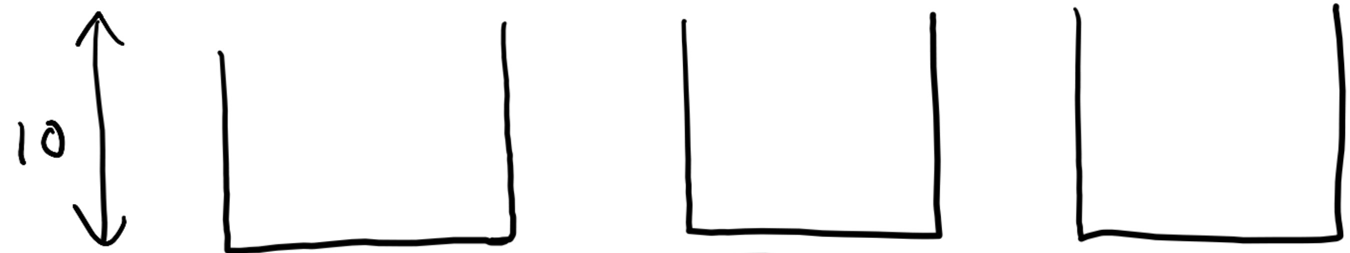
H

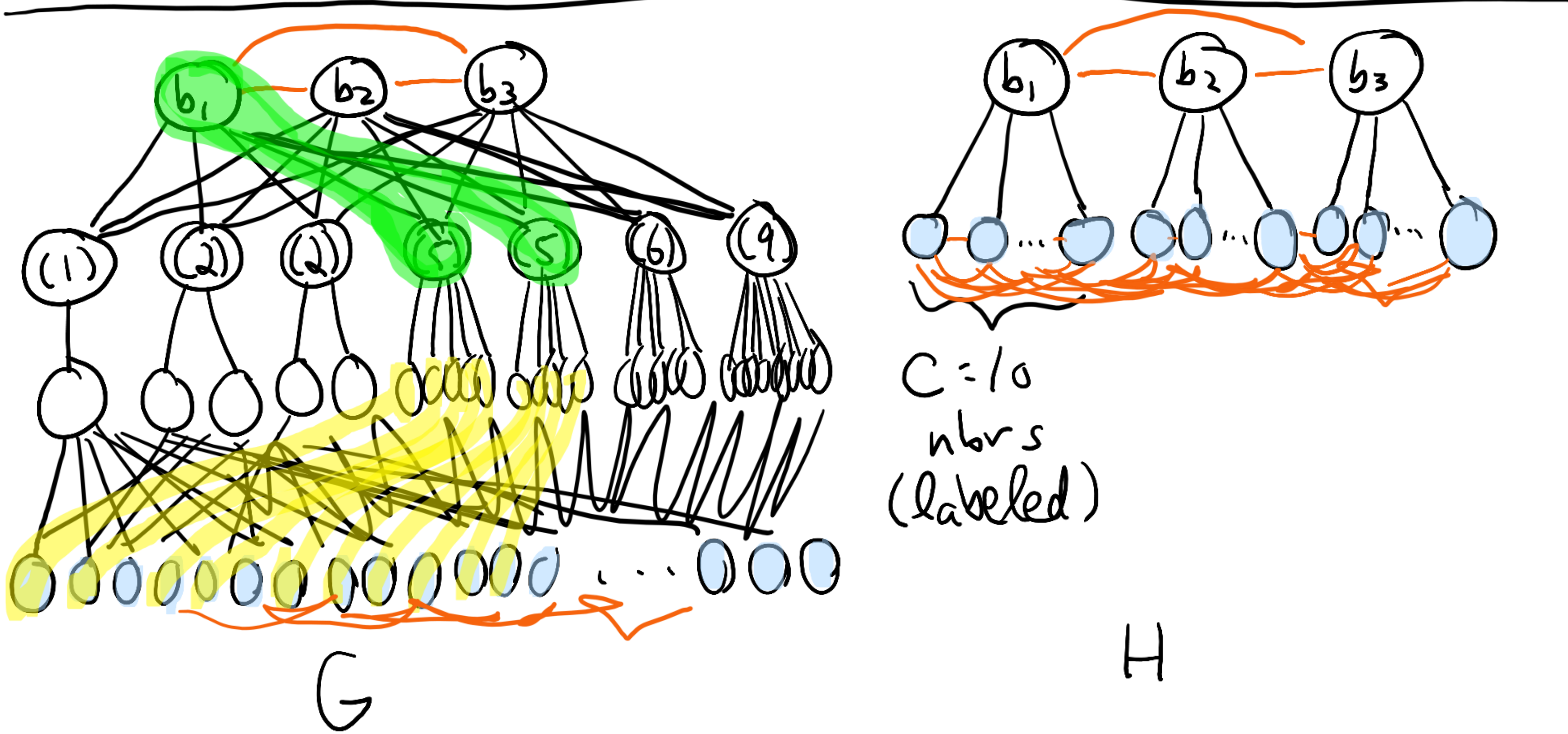
$A = \{1, 2, 2, 5, 5, 6, 9\} \qquad C = 10 \qquad k = 3$





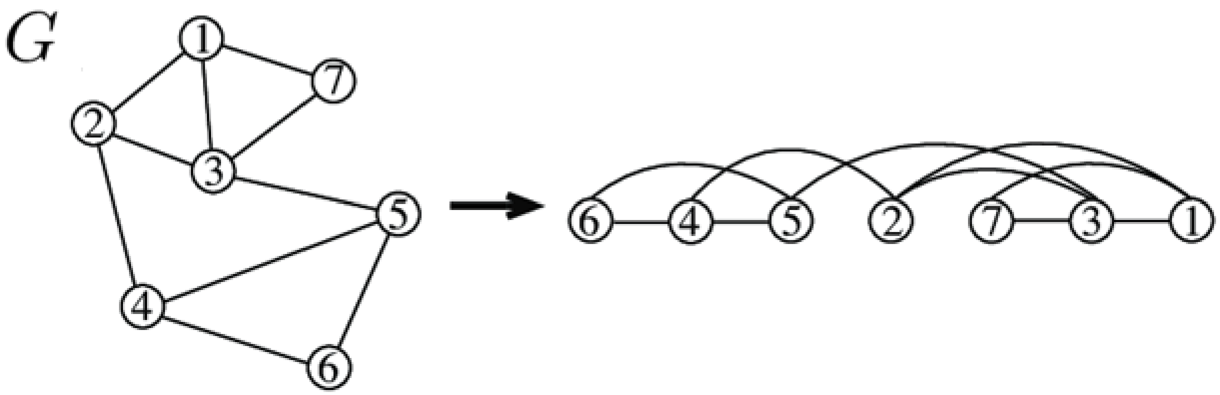
$A = \{1, 2, 2, 5, 5, 6, 9\} \quad C = 10 \quad k = 3$





Theorem: Labeled Contractibility is FPT in parameter $k + \delta(G)$.

$\delta(G)$ = degeneracy of G: each subgraph of G has a vertex of degree at most $\delta(G)$. Can order vertices so that each vertex has $\leq \delta(G)$ neighbors ahead.



Theorem: Labeled Contractibility is FPT in parameter $k + \delta(G)$.

$\delta(G)$ = degeneracy of G: each subgraph of G has a vertex of degree at most $\delta(G)$. Can order vertices so that each vertex has $\leq \delta(G)$ neighbors ahead.

Lemma: if G can be contracted into H using k contractions, then $\delta(H) \leq \delta(G) + k$.

Idea: start with ordering. After a contraction (u, v), put u at the end. Each other vertex has at most 1 more neighbor ahead.

Lemma: if G can be contracted into H using k contractions, then $\delta(H) \leq \delta(G) + k$.

Branching algorithm uses the fact that any contraction (u, v) must have $v \in V(G) \setminus V(H)$, i.e., must make an extra label disappear.

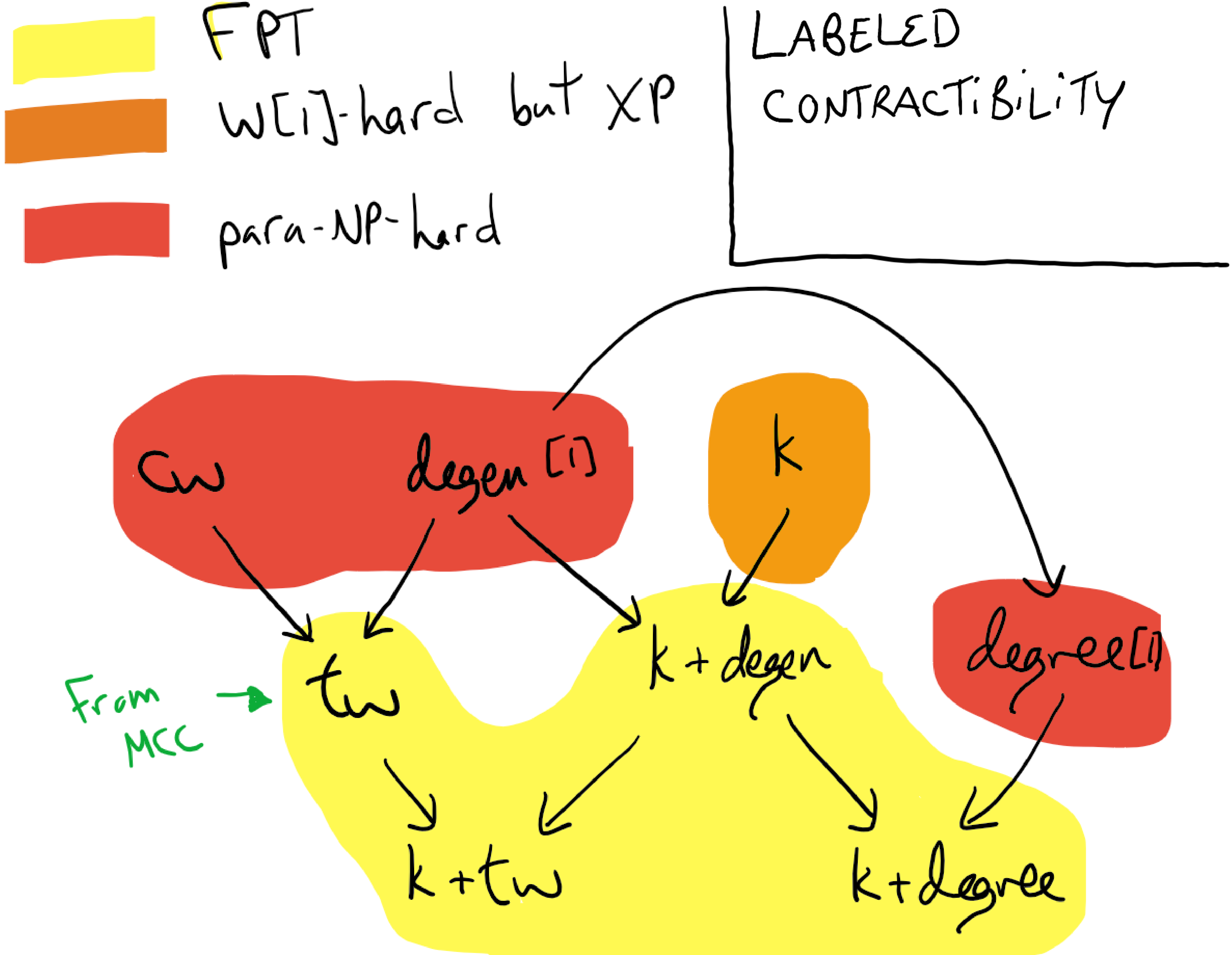
There are at most k vertices in $V(G) \setminus V(H)$ and they are known. Take a u of min-degree and branch into possible contractions.

Lemma: if G can be contracted into H using k contractions, then $\delta(H) \leq \delta(G) + k$.

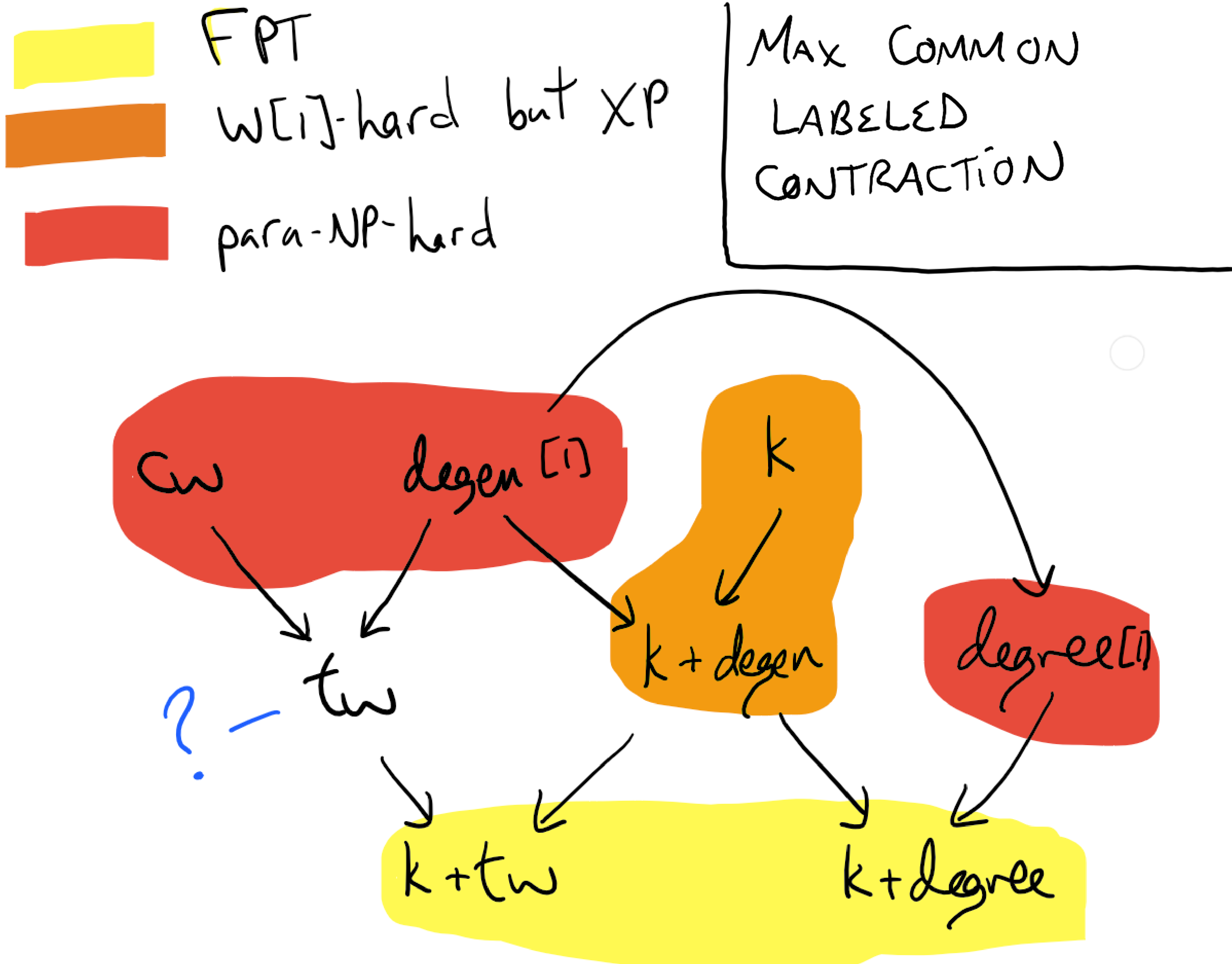
Algorithm (simplified)

Find u of minimum degree, at most $\delta(G) + k$.

1. If $u \notin V(H)$, u must disappear. Branch into all ways of contracting u with a neighbor.
2. If $N_G(u) = N_H(u)$, delete u.
3. Otherwise, $u \in V(H)$ but $N_G(u)$ is incorrect.
   - If u needs a neighbor $v \in V(H)$, branch into all contractions involving u or v (with a vertex in $V(G) \setminus V(H)$).
   - If u has a extra neighbor w, then w cannot be in $V(H)$. Branch into ways of contracting u or w with a vertex in $N_G(u)$.


FPT
W[1]-hard but XP
para-NP-hard
LABELED CONTRACTIBILITY
cw
degen [1]
k
degree[1]
From MCC
tw
k + degen
k+tw
k+degree

MAX COMMON LABELED CONTRACTION



Theorem: Max Common Labeled Contraction is W[1]-hard in parameter k, even on graphs of degeneracy 4.

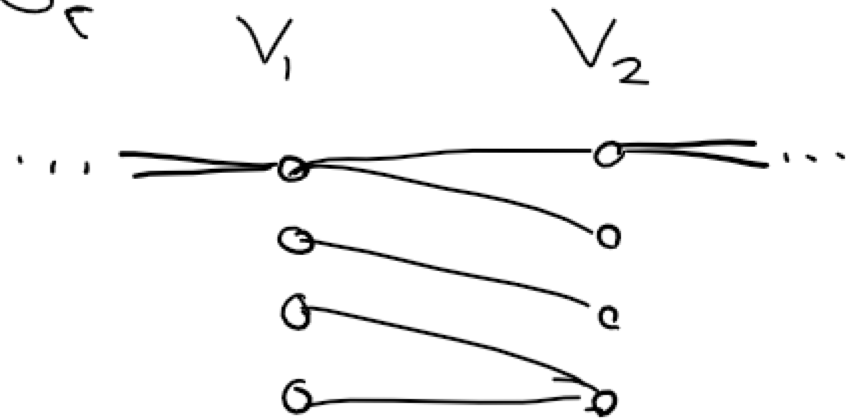
So little hope for an FPT algorithm in $k + \delta(G)$, or even $f(k)n^{g(\delta(G))}$.

Theorem: Max Common Labeled Contraction is W[1]-hard in parameter k, even on graphs of degeneracy 4.

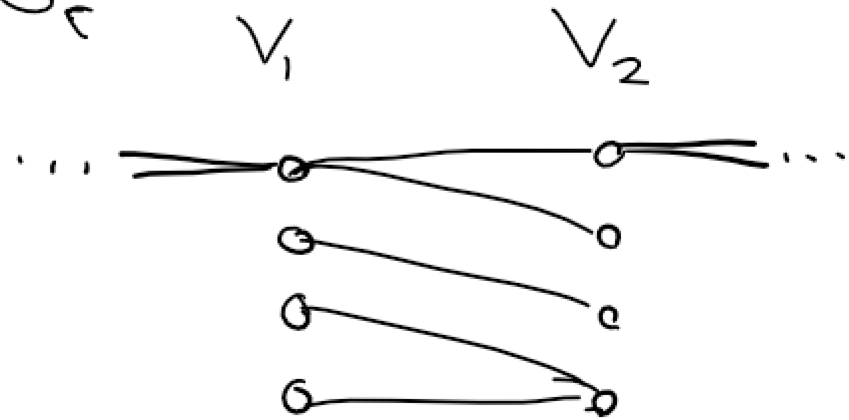
So little hope for an FPT algorithm in $k + \delta(G)$, or even $f(k)n^{g(\delta(G))}$.

More involved reduction from Multicolored Clique.

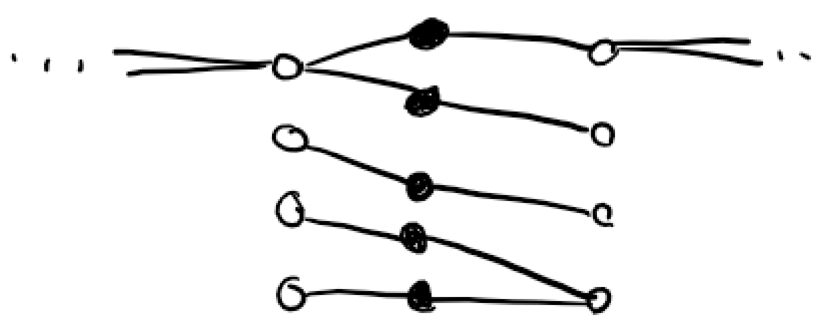
G_r



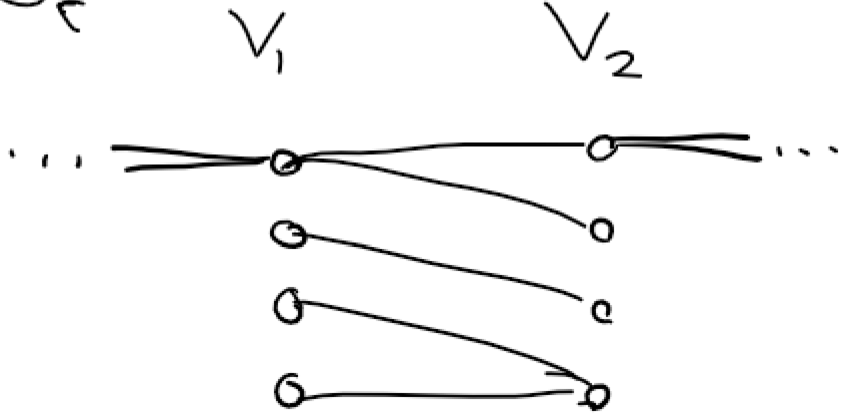
G_r



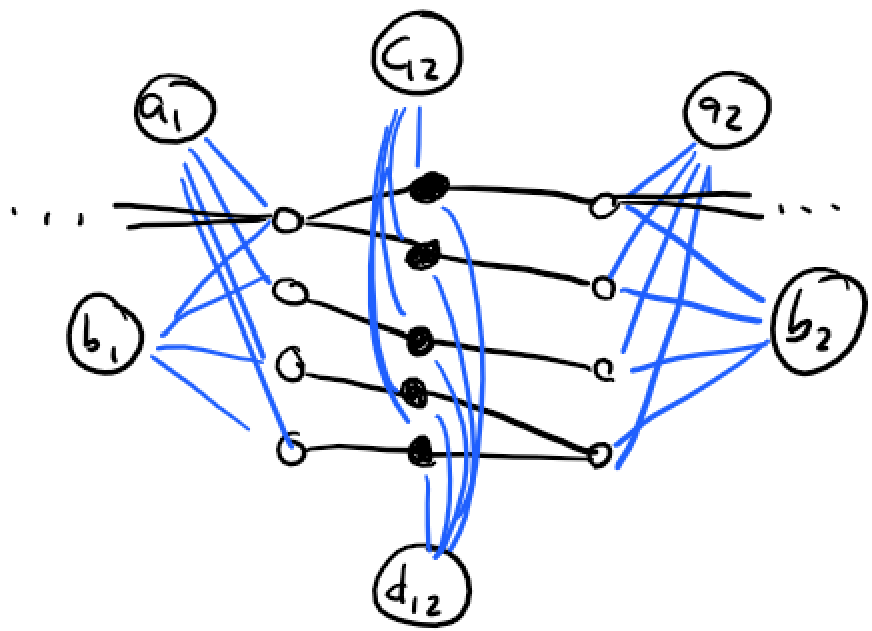
G

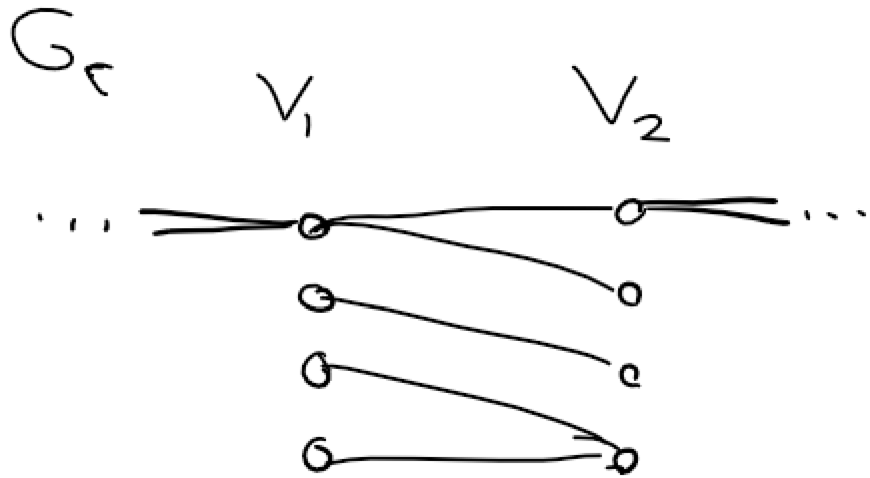


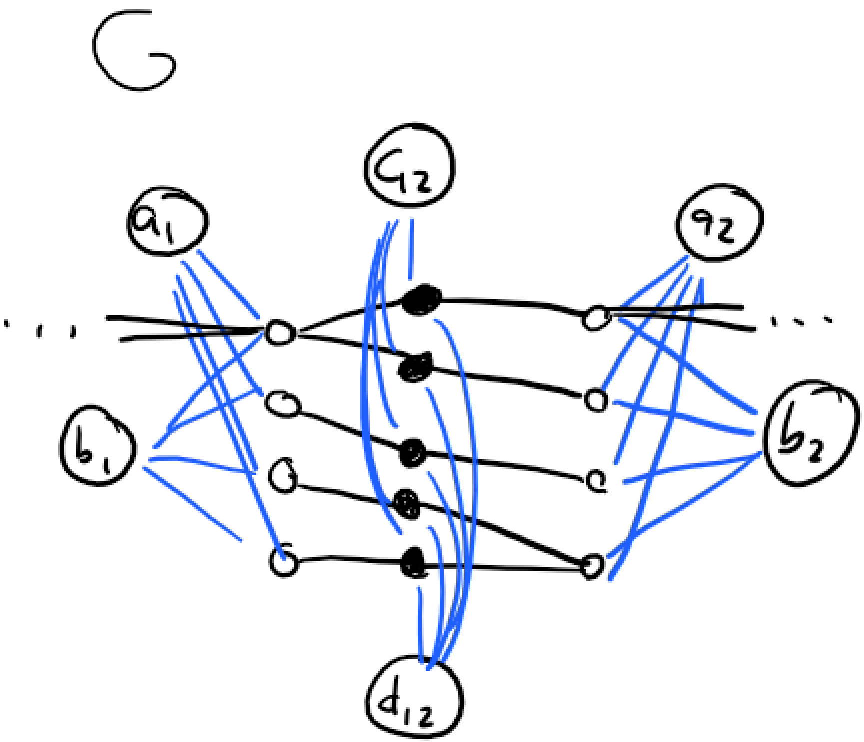
G_r

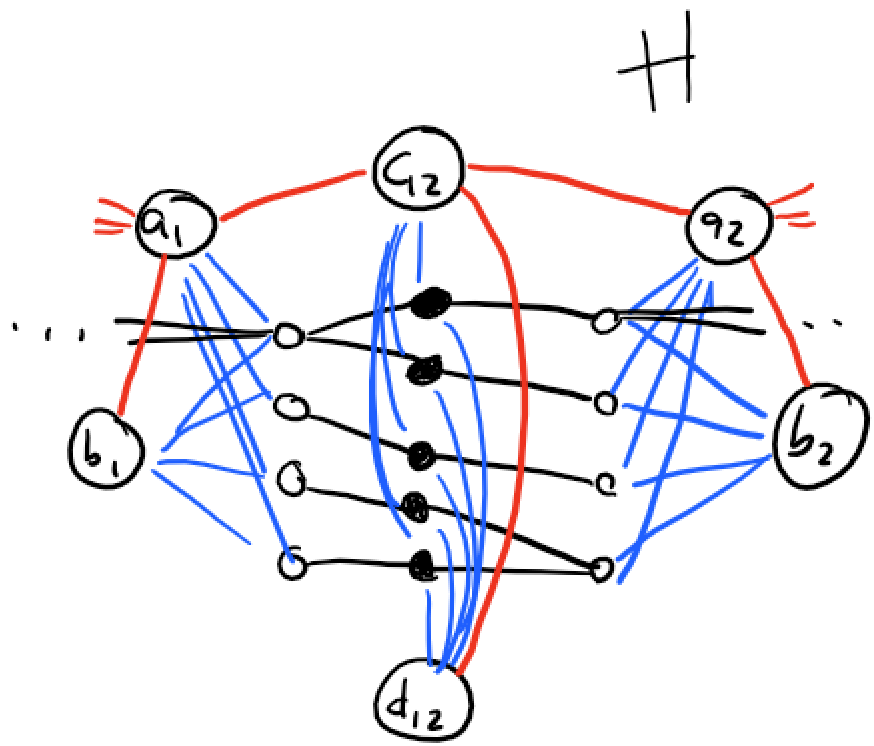


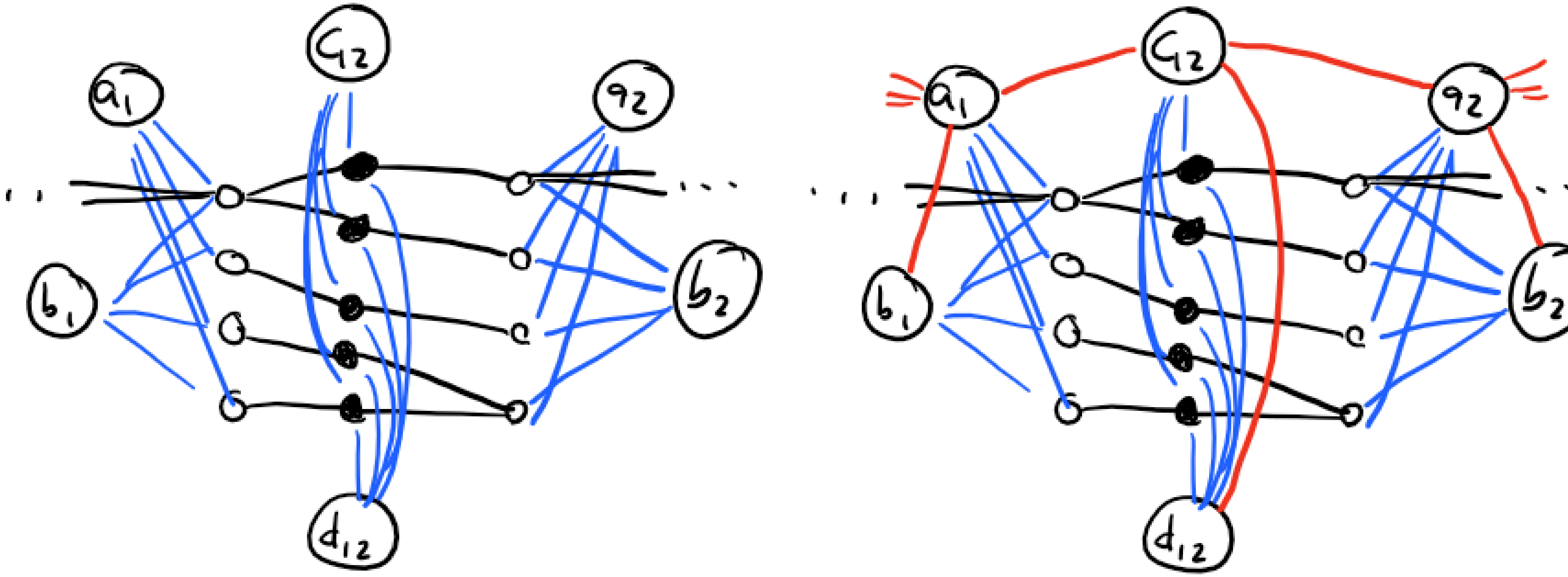
G



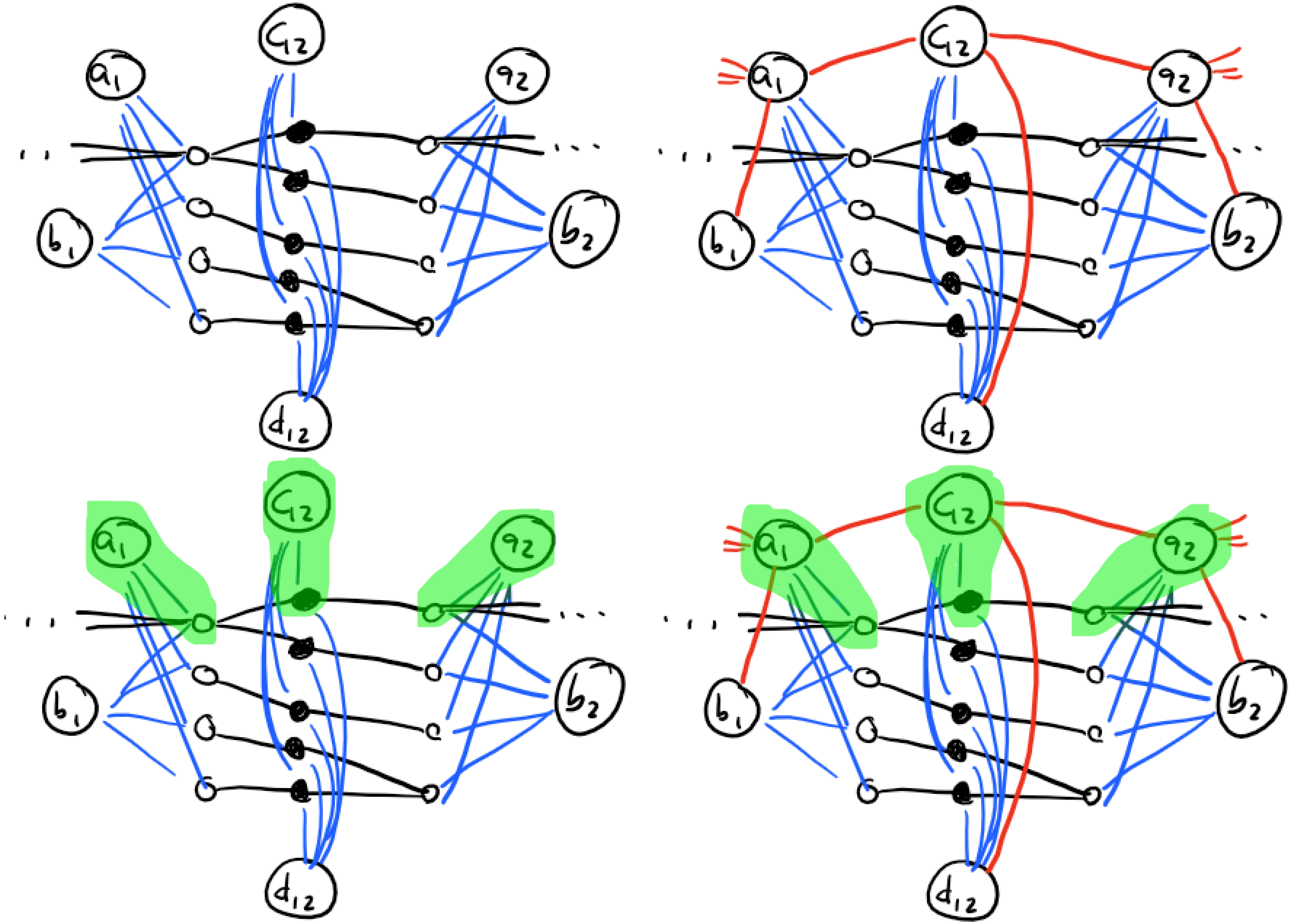

G_r
V_1
V_2


G
a_1
b_1
c_{12}
d_{12}
a_2
b_2


H
a_1
b_1
c_{12}
d_{12}
a_2
b_2



G_C has a multicolored clique
iff
G, H have a common contraction achievable in at most $2(k + \binom{k}{2})$ total contractions in G and H.



a_1
c_{12}
a_2
b_1
b_2
d_{12}

Theorem: Max Common Labeled Contraction is FPT in $k + \max(\Delta(G), \Delta(H))$, where Δ is the max degree.

Theorem: Max Common Labeled Contraction is FPT in $k + \max(\Delta(G), \Delta(H))$, where Δ is the max degree.

Exercise: show that max degrees remain bounded by $O(k\Delta)$ after k contractions, use that to design a branching algorithm.

Treewidth parameterization

Open problem: is MCLC FPT in $\max(tw(G), tw(H))$?

(difficulty: DP over two tree decompositions)

Treewidth parameterization

Open problem: is MCLC FPT in $\max(tw(G), tw(H))$?

Theorem: Max Common Labeled Contraction is FPT in $tw(G \cup H)$.

$G \cup H$: take union of vertex sets and edge sets

Treewidth parameterization

Open problem: is MCLC FPT in $\max(tw(G), tw(H))$?

Theorem: Max Common Labeled Contraction is FPT in $tw(G \cup H)$.

$G \cup H$: take union of vertex sets and edge sets

Idea 1: use dynamic programming over a tree decomposition. Probably gives $O^*(2^{tw^2})$, but **horrible recurrences** and not worth it.

Treewidth parameterization

Open problem: is MCLC FPT in $\max(tw(G), tw(H))$?

Theorem: Max Common Labeled Contraction is FPT in $tw(G \cup H)$.

$G \cup H$: take union of vertex sets and edge sets

Idea 1: use dynamic programming over a tree decomposition. Probably gives $O^*(2^{tw^2})$, but **horrible recurrences** and not worth it.

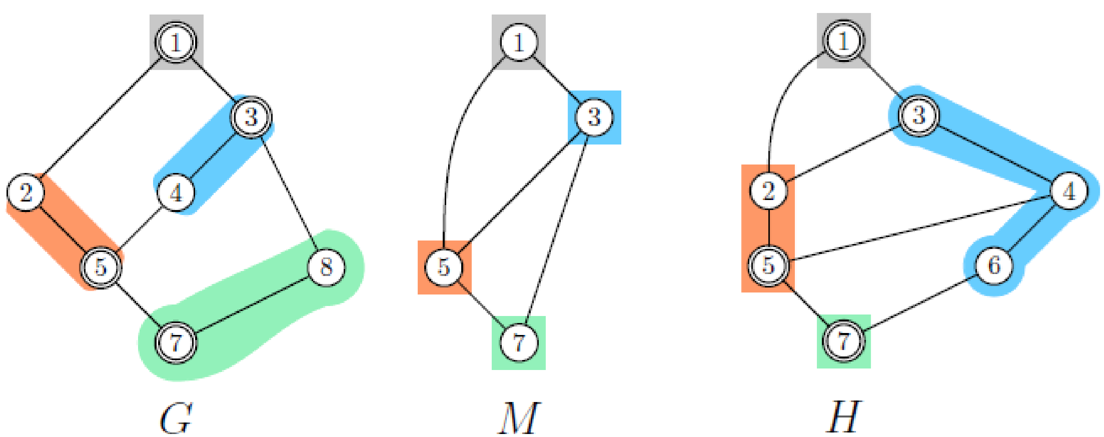
Idea 2: use Courcelle.

Idea 2: use Courcelle.

Minimization version that allows having a 'color' on vertices and edges (for a fixed # of colors). [Arnborg et al, 1991]

Given a tree decomposition of $G \cup H$ and knowledge of $G = (V_1, E_1), H = (V_2, E_2)$:

Use MSO2 to model:



contractions on G: $(5,2), (3,4), (7,8)$
contractions on H: $(5,2), (3,4), (3,6)$

$\mathcal{X} = \{\{1\}, \{2,5\}, \{3,4\}, \{6,7\}\}$

$\mathcal{Y} = \{\{1\}, \{2,5\}, \{3,4,6\}, \{7\}\}$

◯: representatives

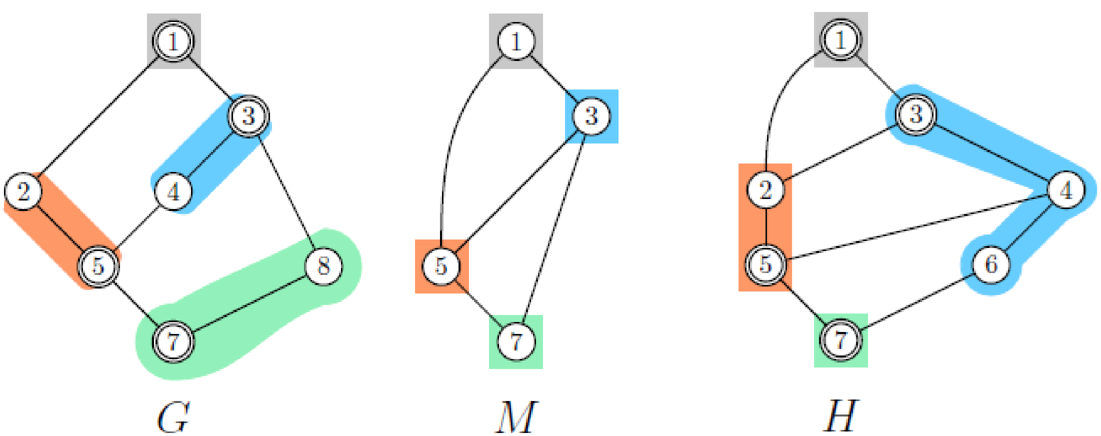
$G = (V_1, E_1), H = (V_2, E_2)$:

min |R| where

$\mathrm{R} \subseteq V_1 \cap V_2, \exists E_1' \subseteq E_1, E_2' \subseteq E_2$ such that

1) each connected component in $G[E_1']$ and $H[E_2']$ has a unique representative in R; (the label that remains)

2) if components C_1, C_2 of $G[E_1']$ have the same respective representative as components D_1, D_2 of $H[E_2']$, then C_1, C_2 share an edge in G iff D_1, D_2 share an edge in H.



contractions on G: $(5, 2), (3, 4), (7, 8)$

contractions on H: $(5, 2), (3, 4), (3, 6)$

$\mathcal{X} = \{\{1\}, \{2, 5\}, \{3, 4\}, \{6, 7\}\}$

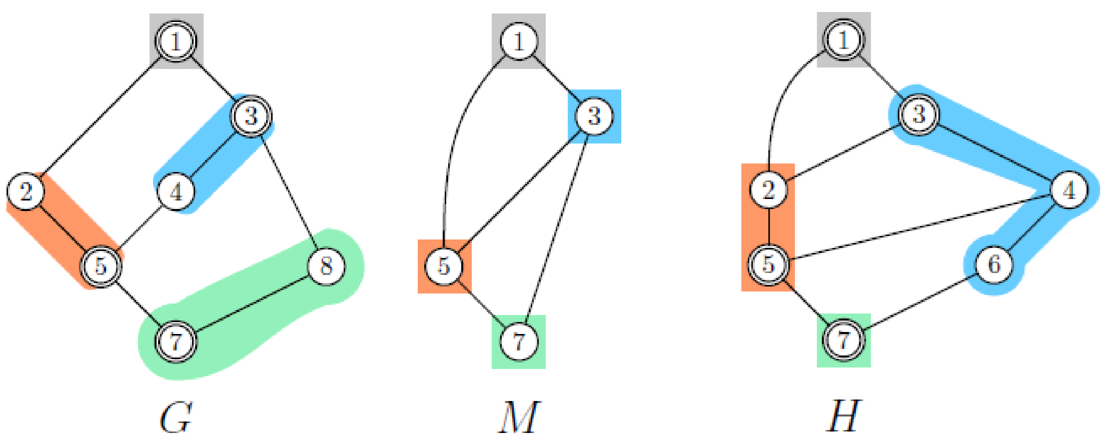
$\mathcal{Y} = \{\{1\}, \{2, 5\}, \{3, 4, 6\}, \{7\}\}$

◯: representatives

$$\exists S_1 \subseteq E_1, \exists S_2 \subseteq E_2, \exists R \subseteq V_1 \cap V_2 \text{ s.t}$$
$$\mathbf{contraction}_1(S_1, R) \text{ and } \mathbf{contraction}_2(S_2, R) \text{ and}$$
$$\forall r, r' \in R\ \mathbf{edge}_1(S_1, r, r') \text{ iff } \mathbf{edge}_2(S_2, r, r')$$

$$\mathbf{contraction}_i(S_i, R) = \forall r, r' \in R, r \neq r', \neg\mathbf{path}(r, r', S_i) \text{ and } \forall u \in V_i \setminus R\ \exists r \in R \text{ s.t.}$$
$$\mathbf{path}(u, r, S_i) \text{ and } \forall r' \neq r\ \neg\mathbf{path}(u, r', S_i)$$

$$\mathbf{edge}_i(S_i, r, r') = \exists x, y \in V_i \text{ s.t. } \mathbf{path}(x, r, S_i) \text{ and } \mathbf{path}(y, r, S_i) \text{ and } \{x, y\} \in E_i$$



contractions on G: $(5, 2), (3, 4), (7, 8)$
contractions on H: $(5, 2), (3, 4), (3, 6)$

$\mathcal{X} = \{\{1\}, \{2, 5\}, \{3, 4\}, \{6, 7\}\}$

$\mathcal{Y} = \{\{1\}, \{2, 5\}, \{3, 4, 6\}, \{7\}\}$

◎: representatives

Lemma: if H is a contraction of G, then $tw(G \cup H) \leq 2tw(G)$.

Lemma: if H is a contraction of G, then $tw(G \cup H) \leq 2tw(G)$.

Corollary: Labeled Contractibility is FPT in $\max(tw(G), tw(H))$.

Lemma: if H is a contraction of G, then
$tw(G \cup H) \leq 2tw(G)$.

Lemma: if G and H have a common contraction achievable in a total of k contractions, then
$tw(G \cup H) \leq \max\big(tw(G), tw(H)\big) + k$.

Corollary: Labeled Contractibility is FPT in $\max(tw(G), tw(H))$.

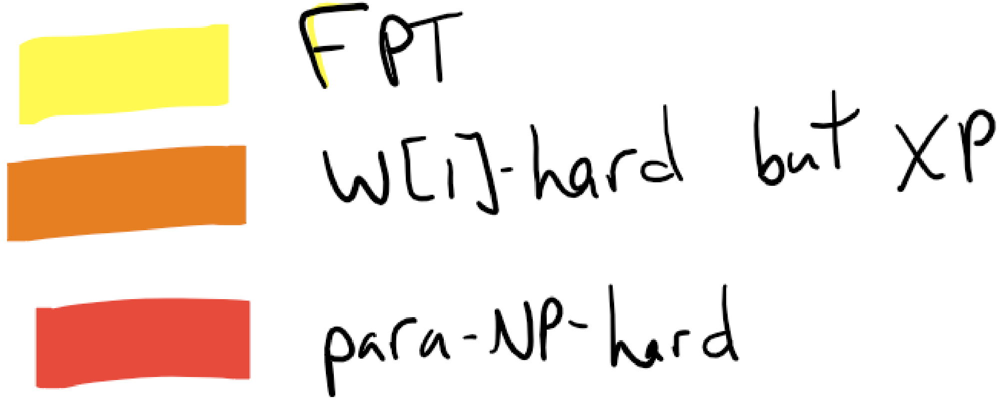
Lemma: if H is a contraction of G, then $tw(G \cup H) \leq 2tw(G)$.

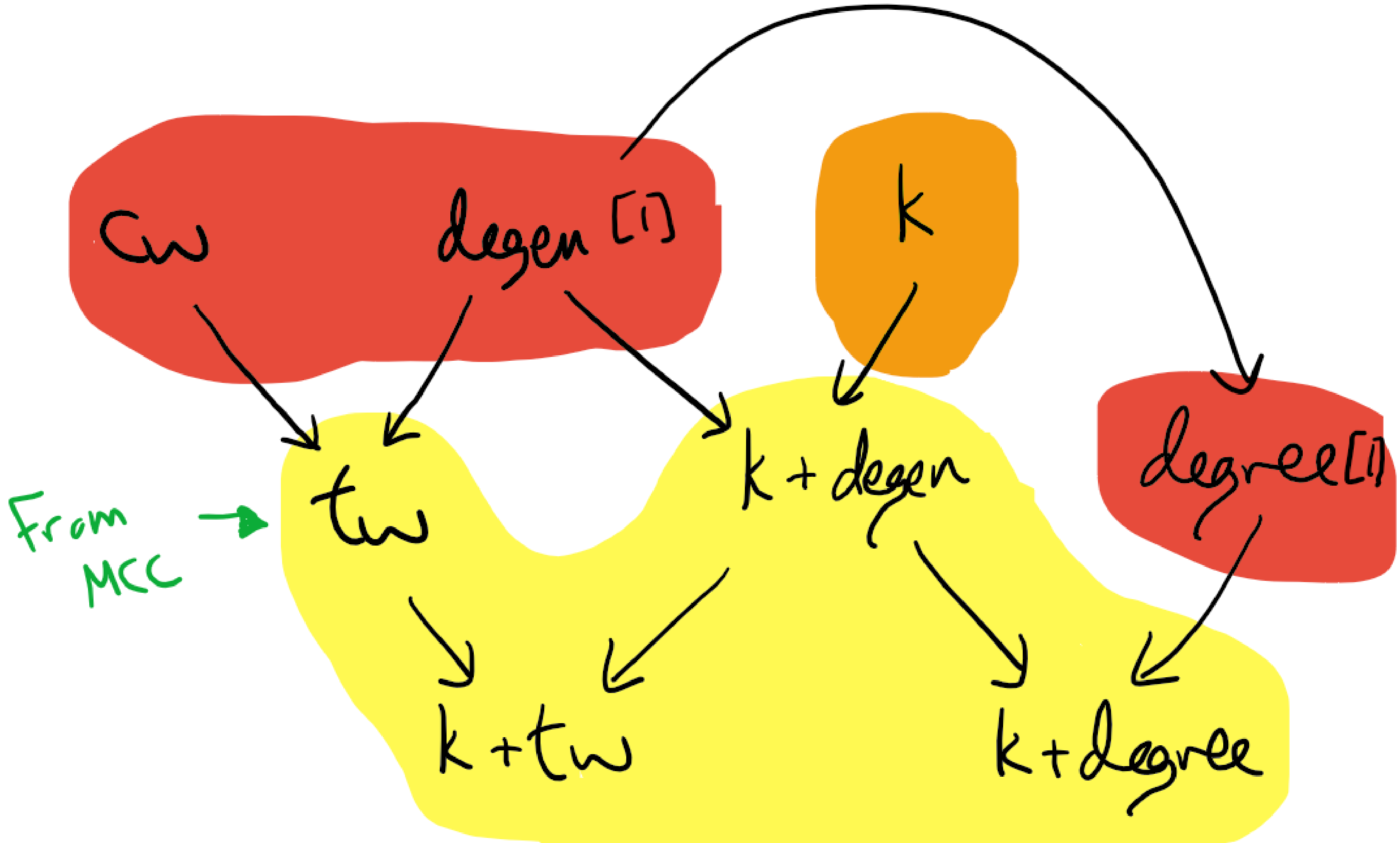
Lemma: if G and H have a common contraction achievable in a total of k contractions, then $tw(G \cup H) \leq \max\big(tw(G), tw(H)\big) + k$.

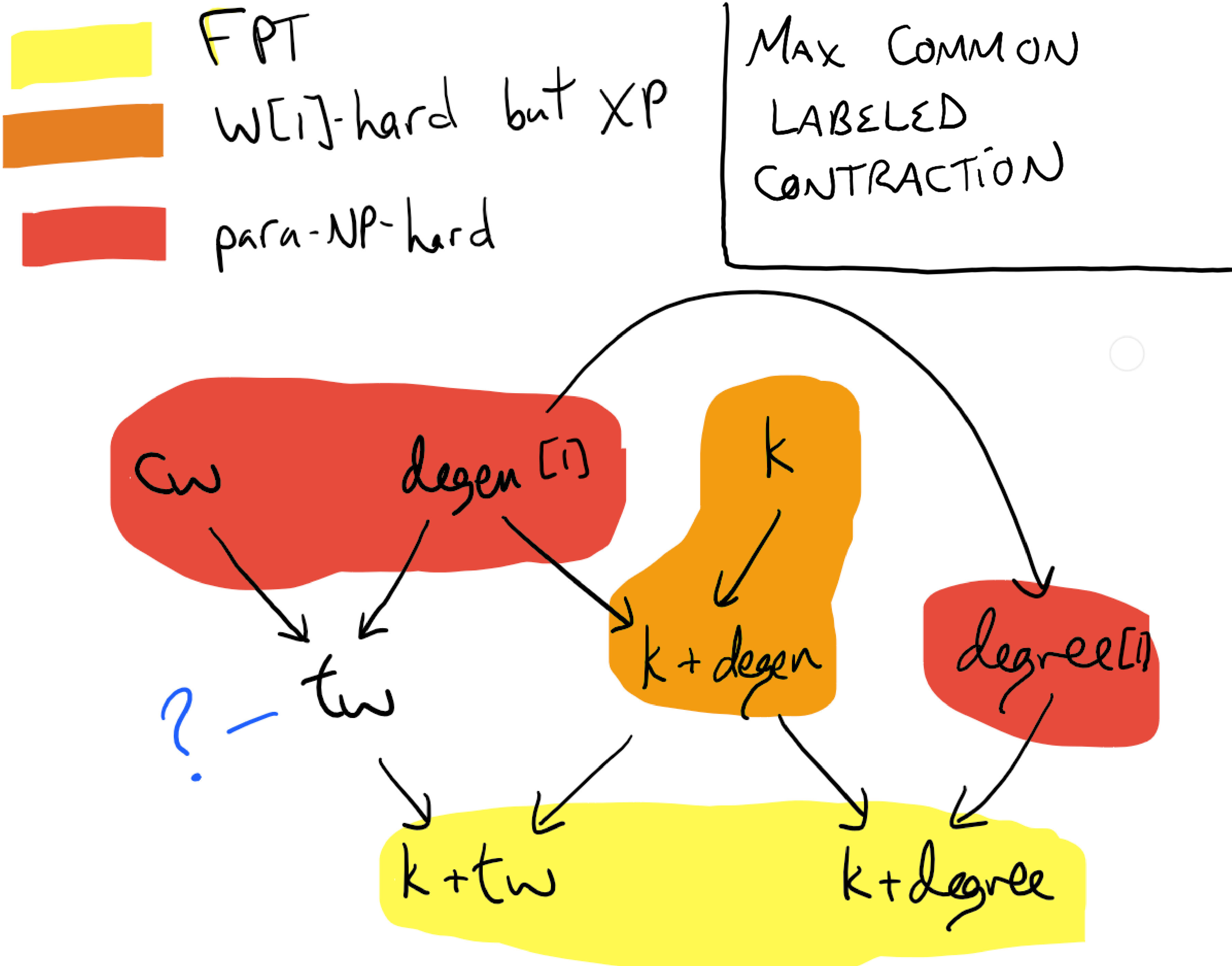
Corollary: Labeled Contractibility is FPT in $\max(tw(G), tw(H))$.

Maximum Common Labeled Contraction is FPT in $k + \max(tw(G), tw(H))$.

LABELED CONTRACTIBILITY







FPT
W[1]-hard but XP
para-NP-hard
MAX COMMON LABELED CONTRACTION
cw
degen [1]
k
k + degen
degree[1]
tw
?
k + tw
k + degree

Open problems

- Is LC NP-hard on cographs? (cw 2)
- Is MCLC NP-hard on 2-degenerate graphs?
- Is MCLC FPT in $\max(tw(G), tw(H))$?

- What about partially labeled graphs?
  - Need more research on leaf-labeled graphs
- Or graph in which labels have multiple but bounded occurrences?