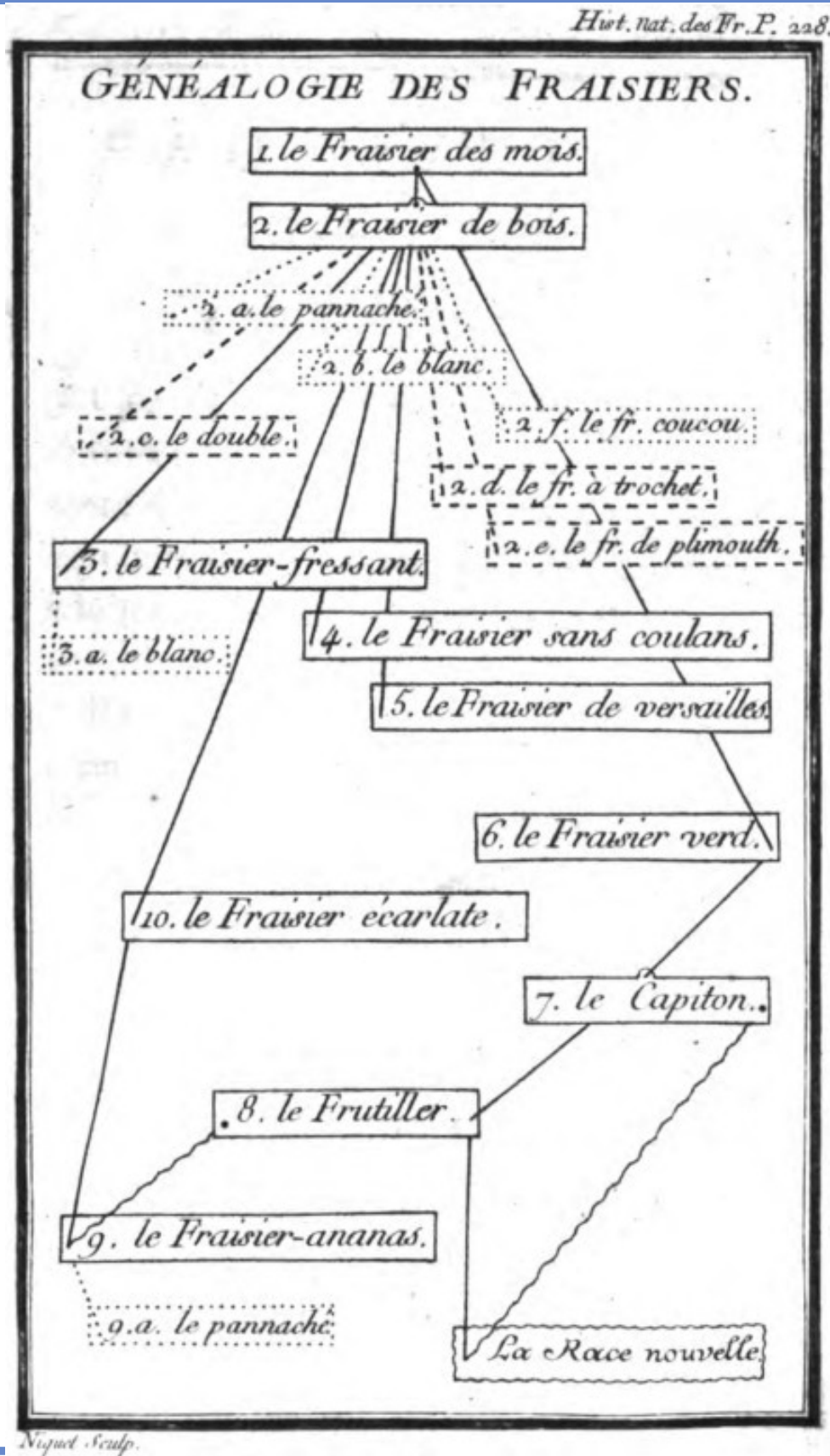

Rearrangement moves on rooted phylogenetic networks

Leo van Iersel

Workshop on Algebraic and Combinatorial Phylogenetics
Barcelona, June 30, 2017

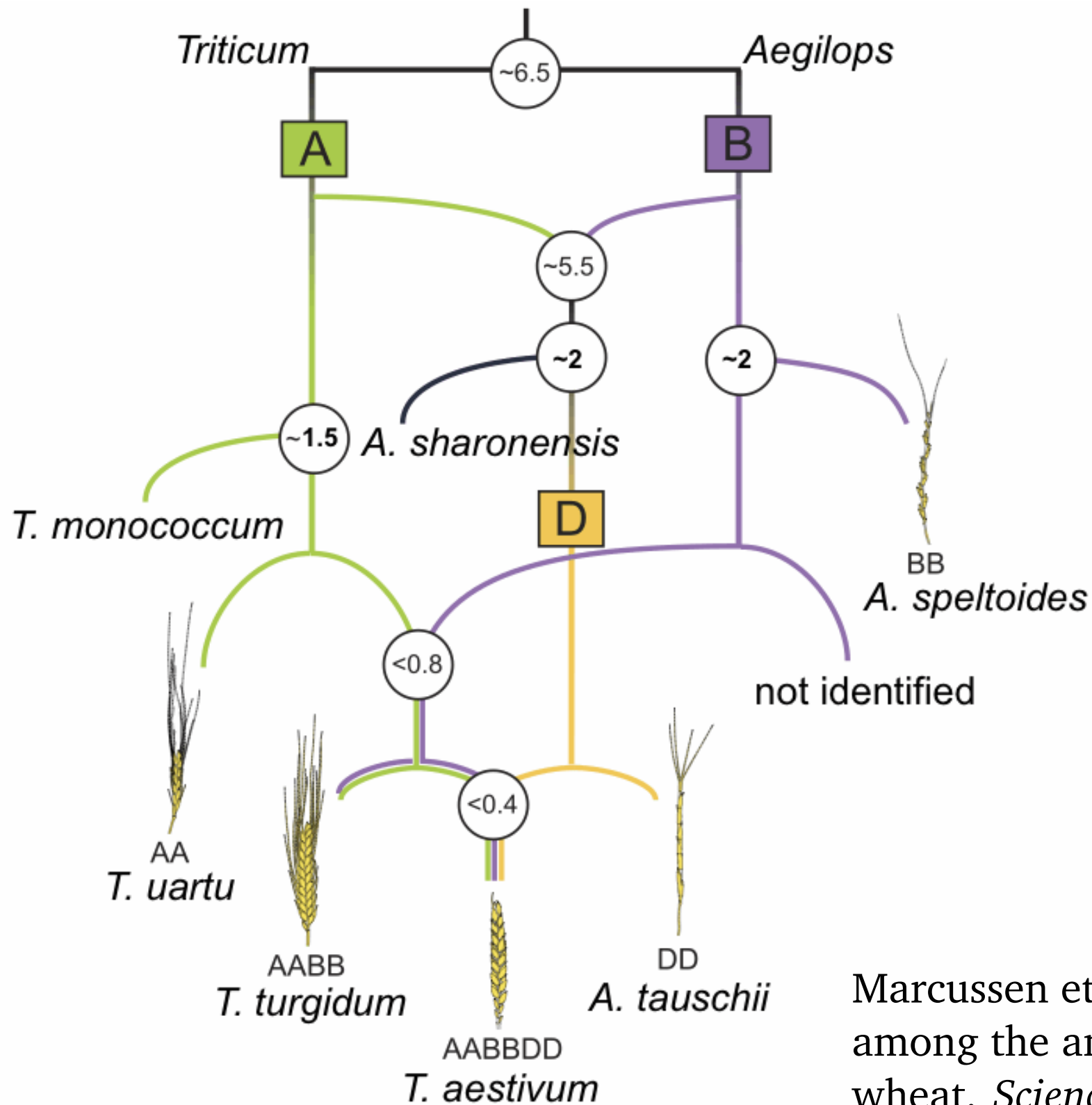
Joint work with Péter L. Erdős, Philippe Gambette, Remie Janssen,
Mark Jones, Manuel Lafond, Fabio Pardi and Celine Scornavacca

Rooted phylogenetic networks



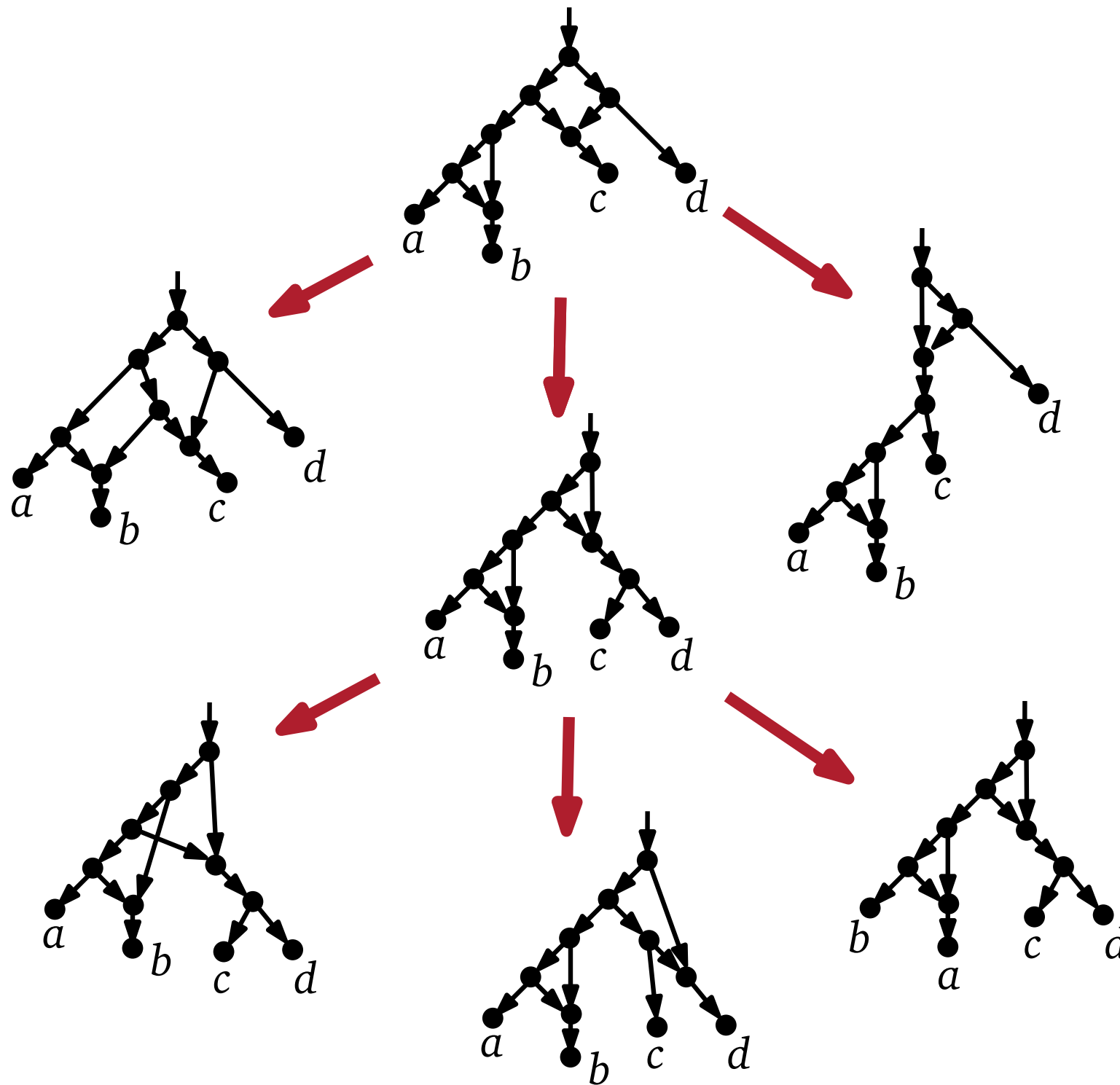
Strawberries
Duchesne, 1766

Rooted phylogenetic networks



Marcussen et al., Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, 2014.

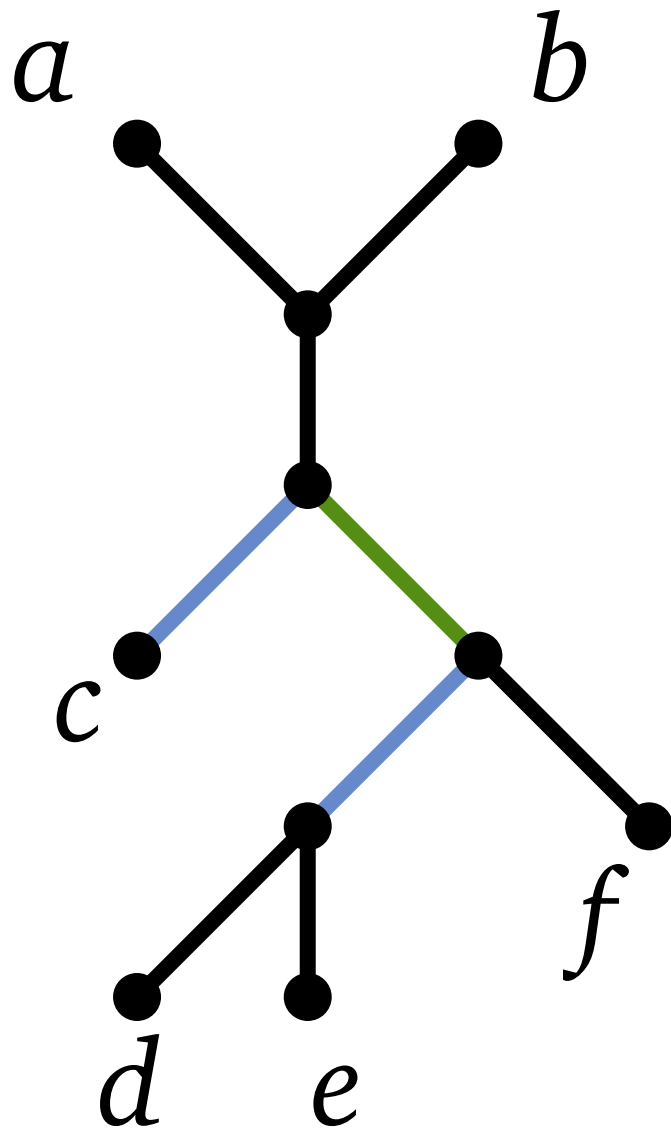
Searching through network space



- Relevant for:
- Maximum Parsimony
 - Maximum Likelihood
 - Bayesian

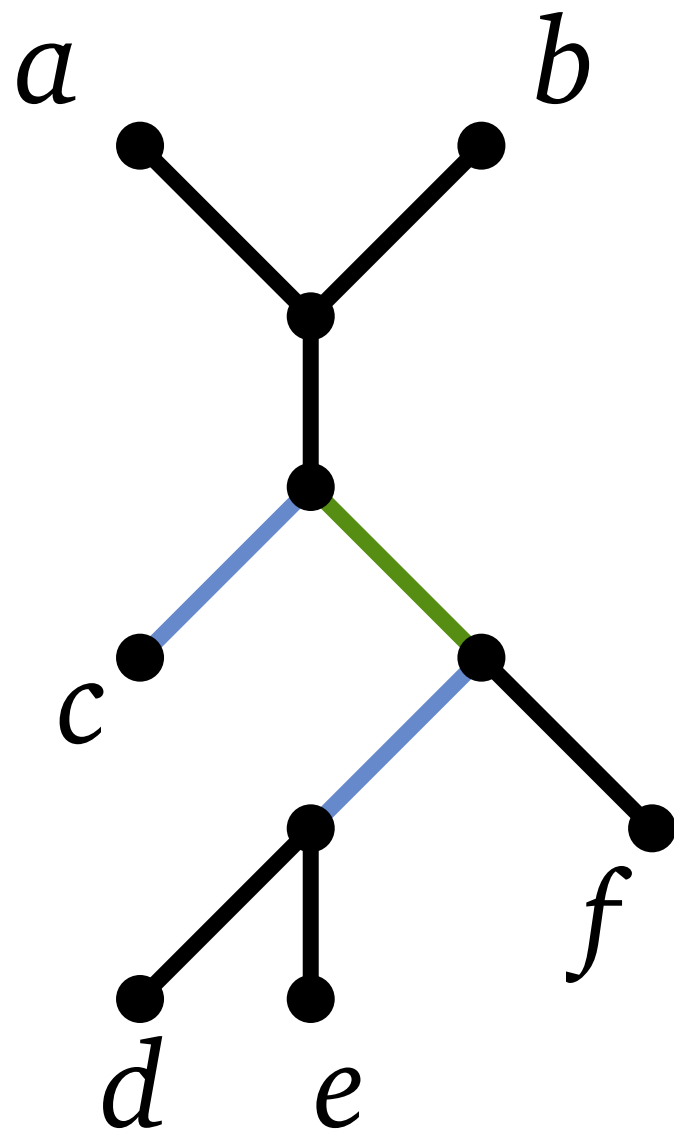
NNI moves on trees

Example. Interchange the subtrees on c and d, e .

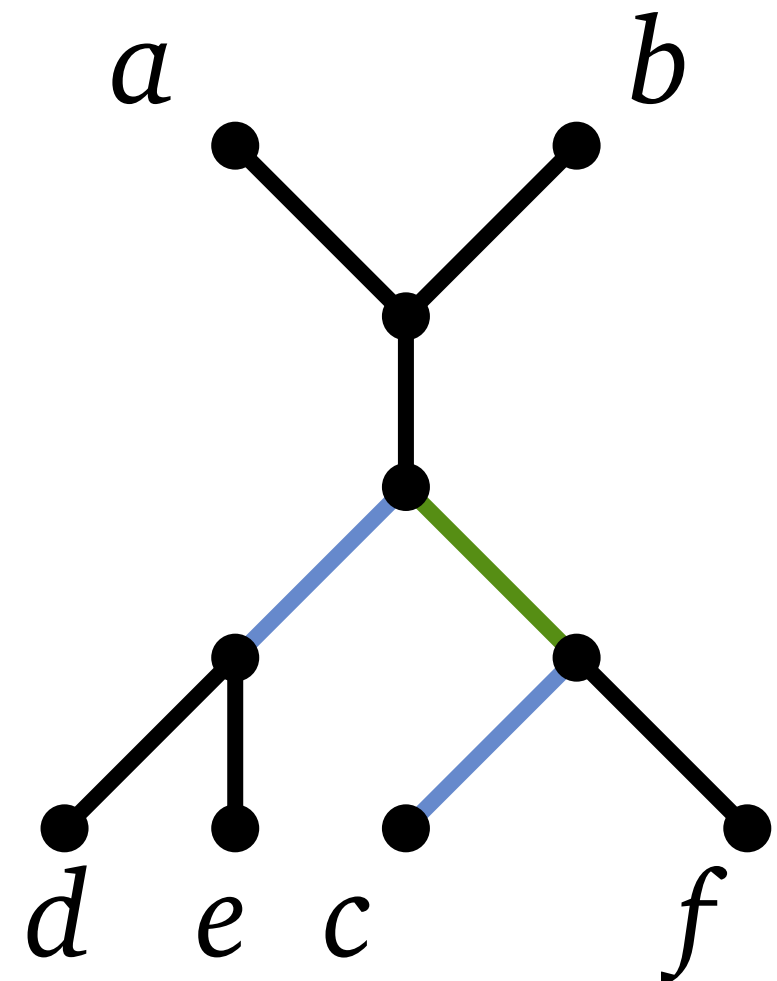


NNI moves on trees

Example. Interchange the subtrees on c and d, e .



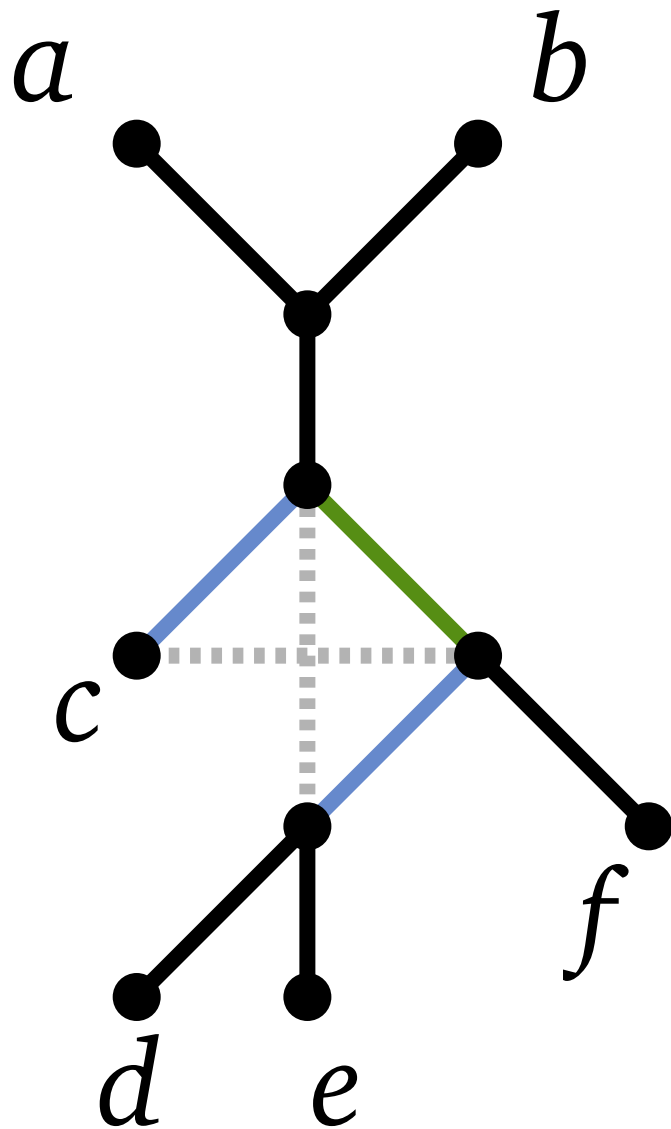
NNI
→



Note: All trees/networks in this talk will be *binary*.

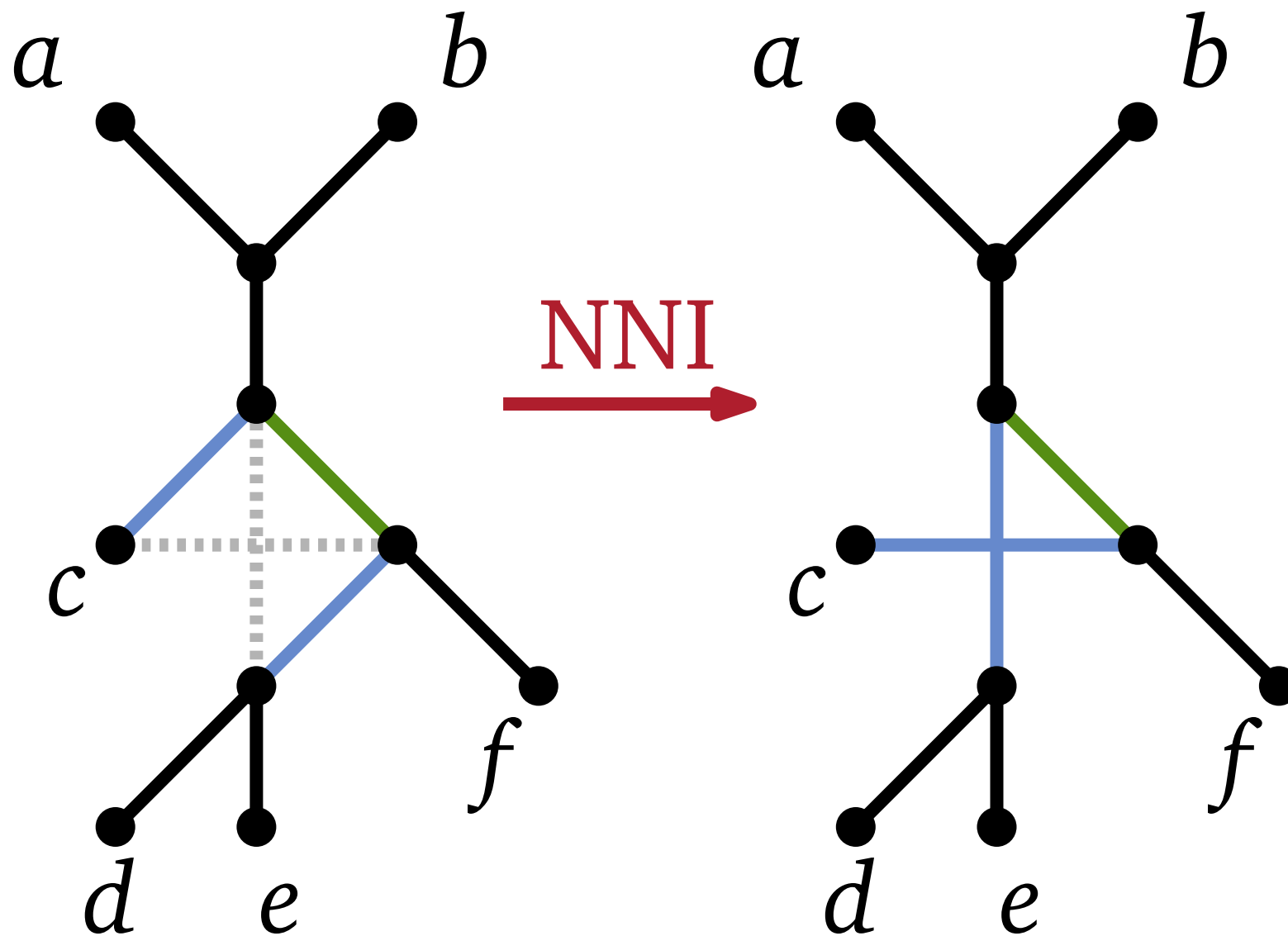
NNI moves on trees *by swapping edges*

Delete the blue edges and replace them by the dashed edges



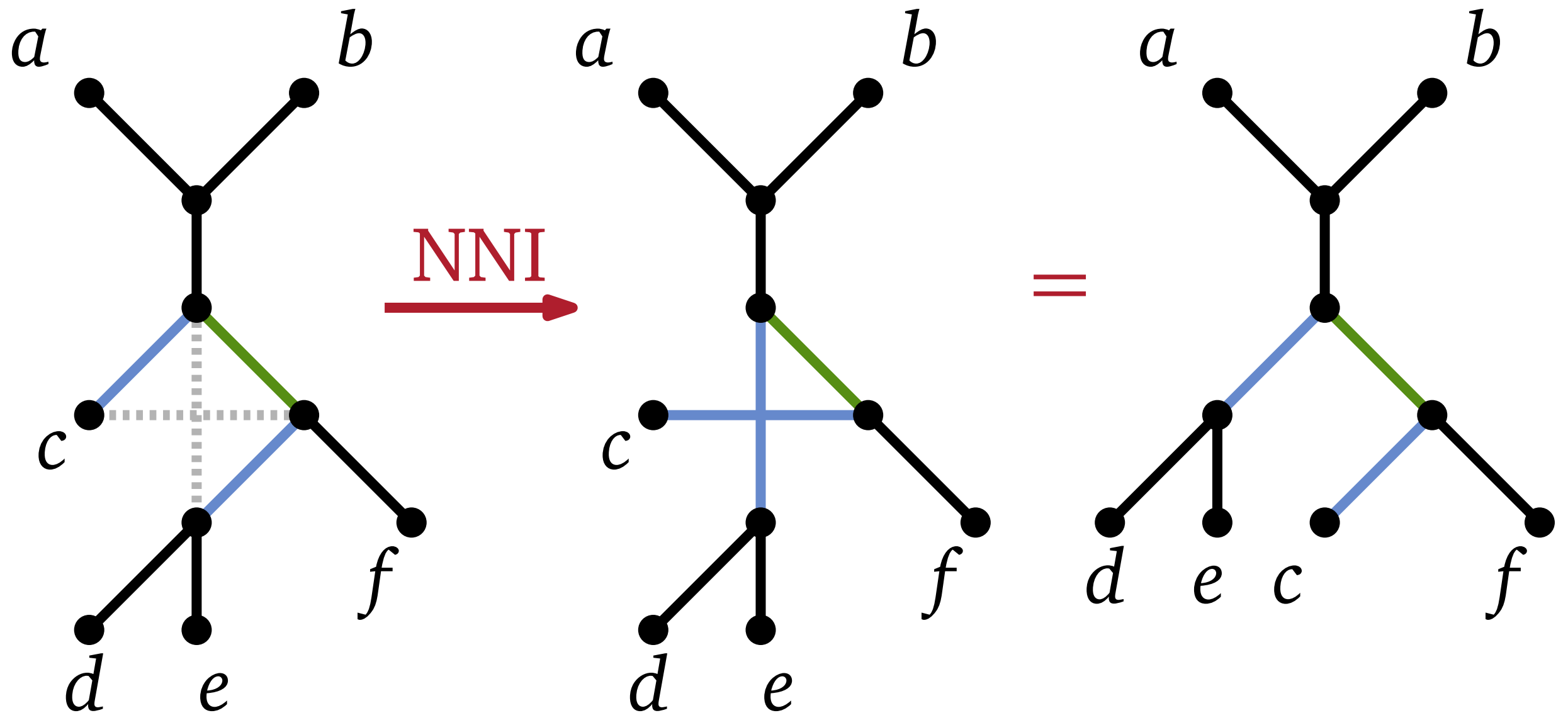
NNI moves on trees *by swapping edges*

Delete the blue edges and replace them by the dashed edges



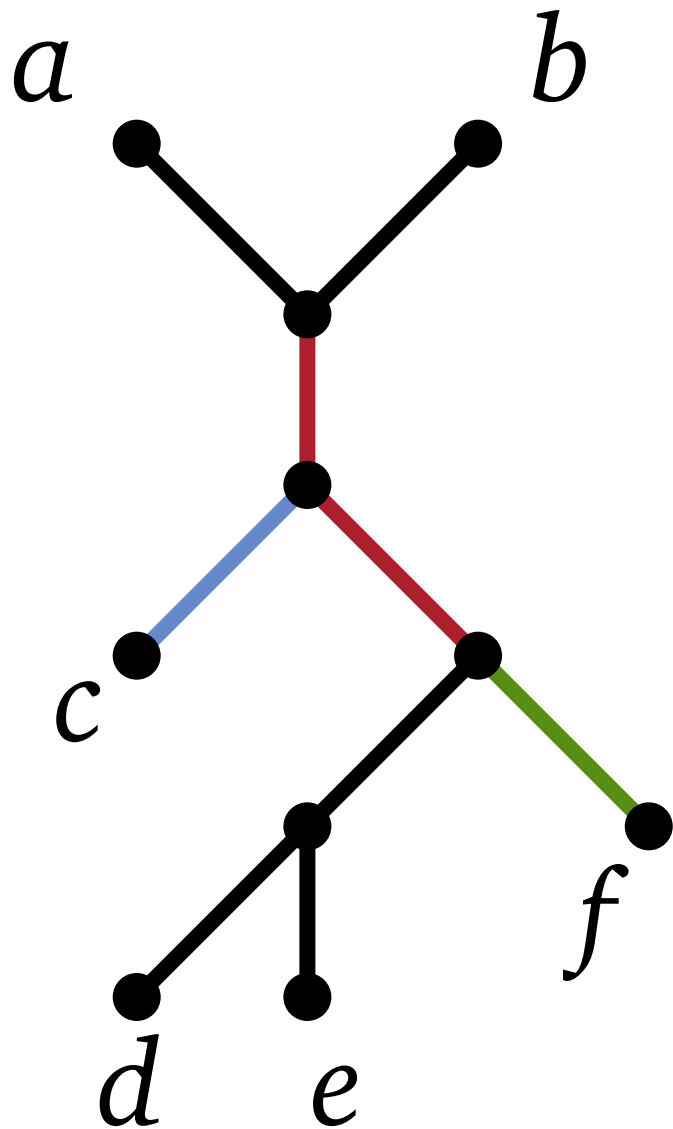
NNI moves on trees *by swapping edges*

Delete the blue edges and replace them by the dashed edges



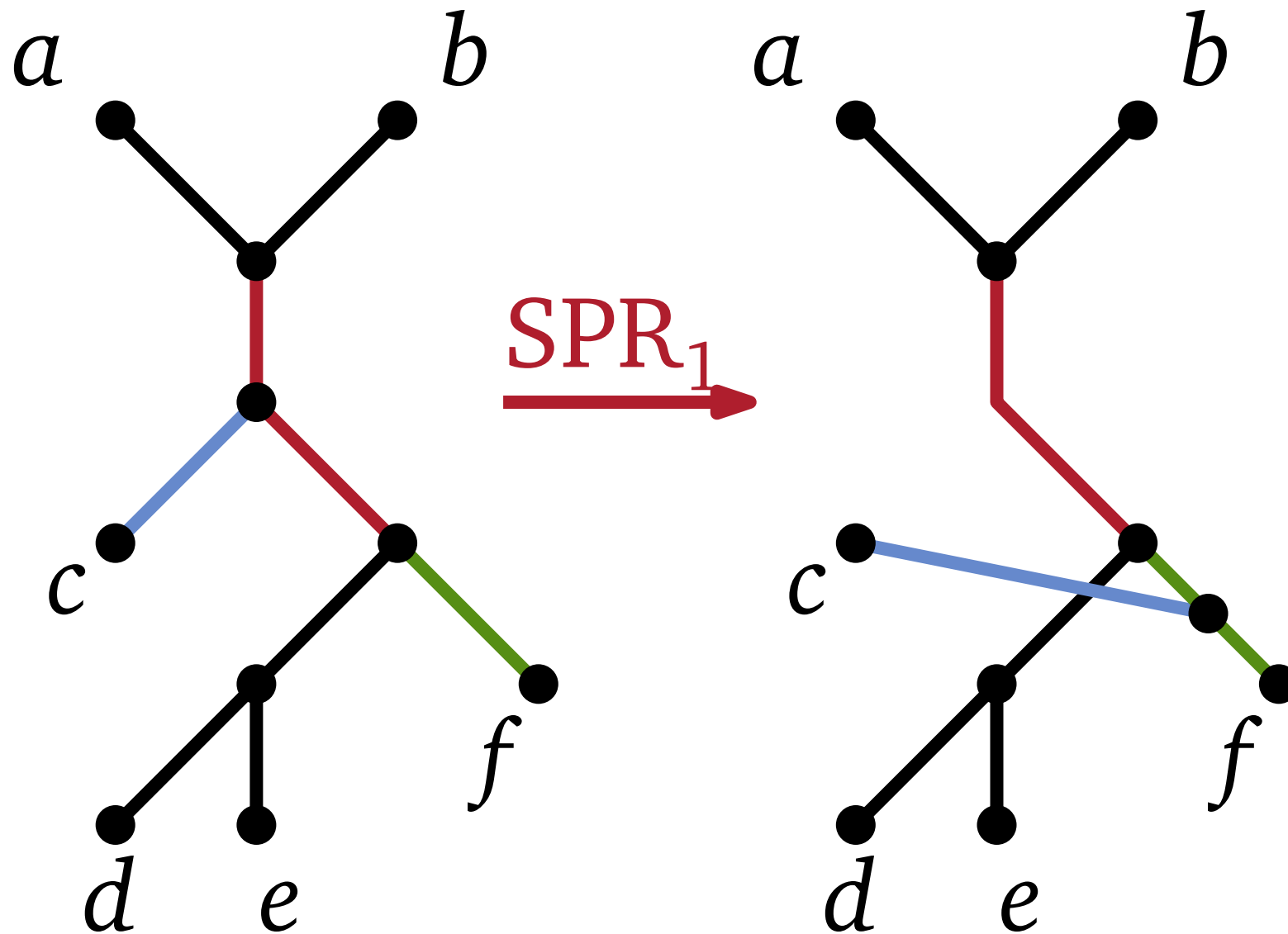
NNI moves on trees *as local SPR moves*

Prune the subtree on c off the green edges and regraft it onto the red edge



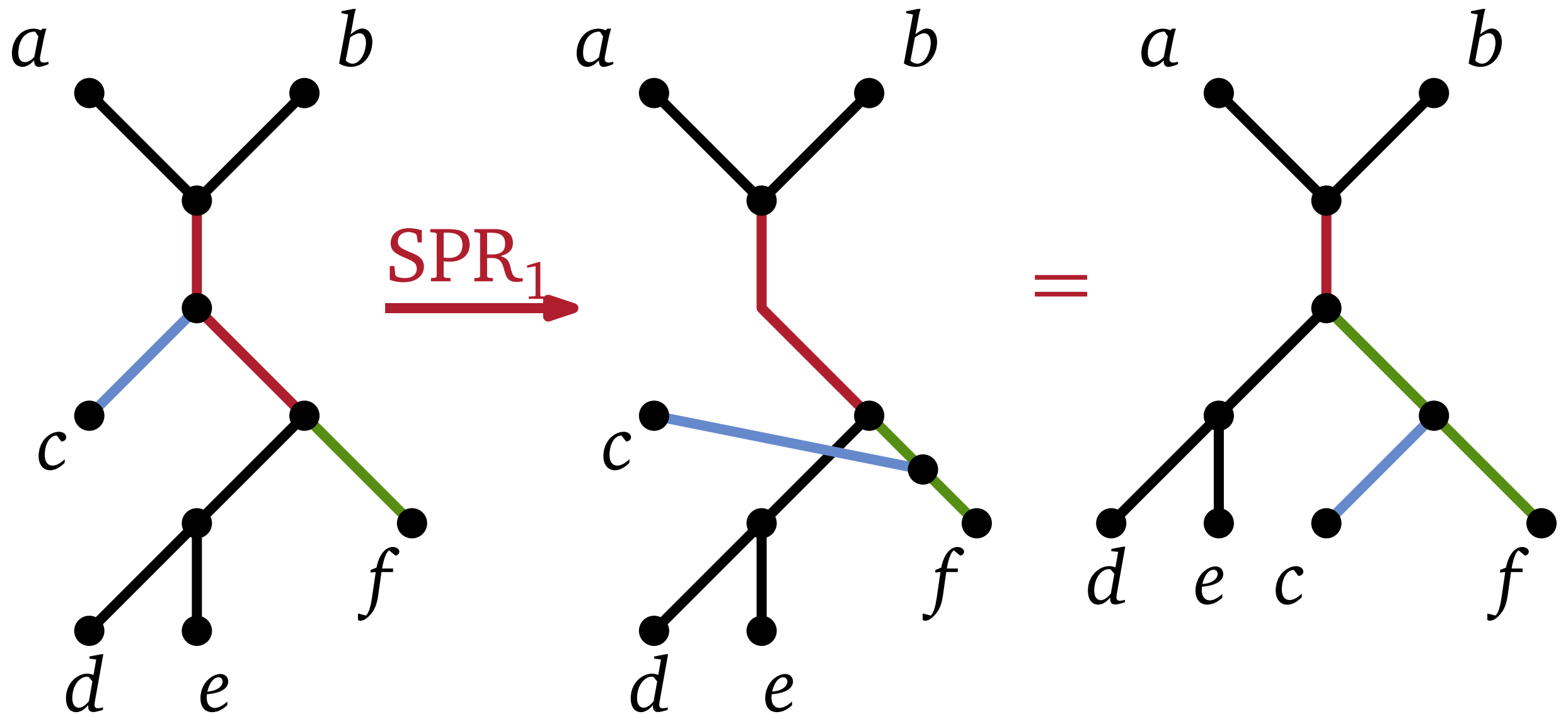
NNI moves on trees *as local SPR moves*

Prune the subtree on c off the green edges and regraft it onto the red edge



NNI moves on trees *as local SPR moves*

Prune the subtree on c off the green edges and regraft it onto the red edge



The move is an SPR_1 if the red edge is incident to one of the green edges

NNI moves on unrooted networks *by swapping edges*

Definition. If an **unrooted** network has edges su, uv, vt but neither ut nor sv , then an **NNI** move consists of replacing the edges su, vt by edges ut, sv .

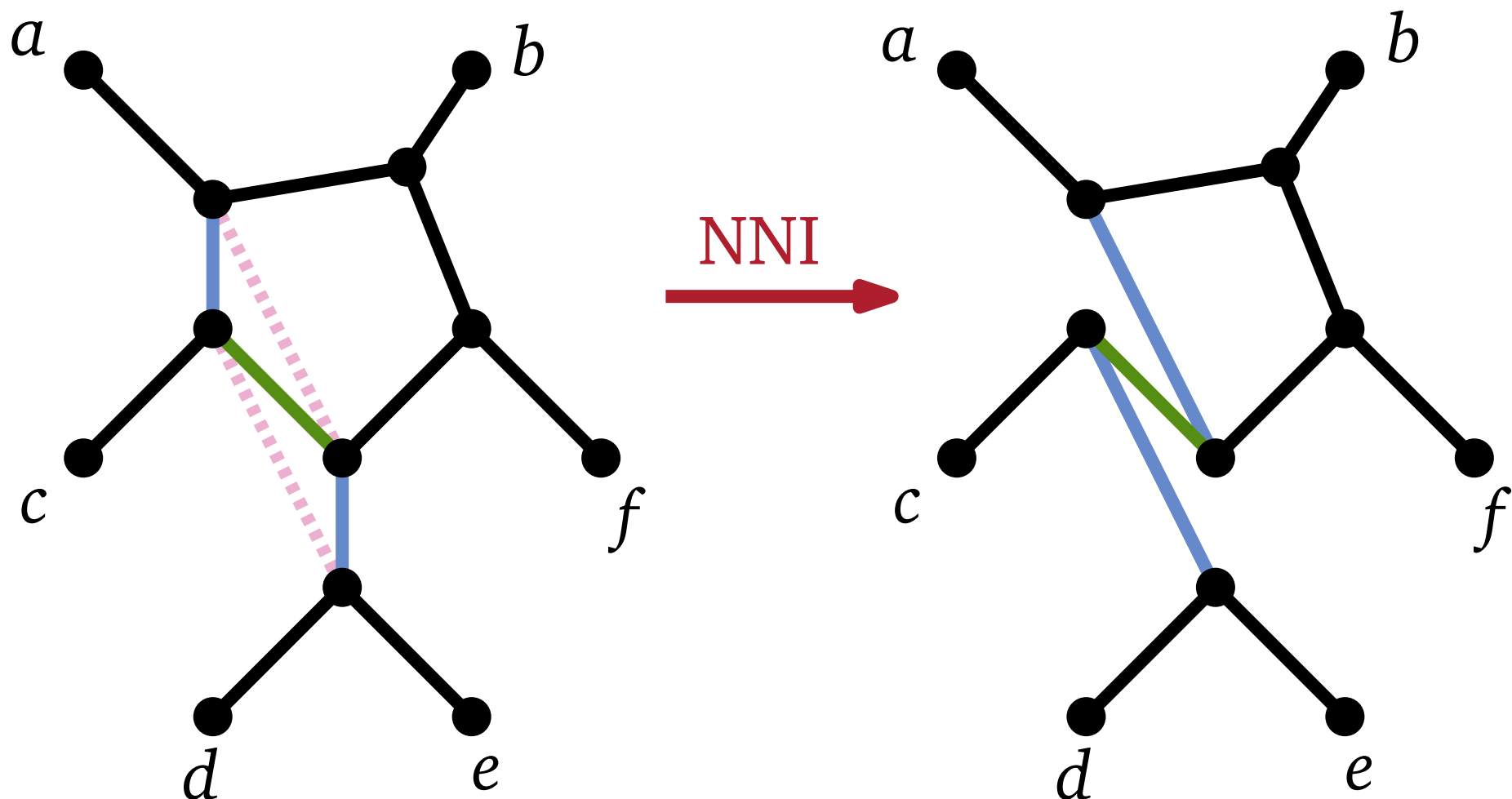


NNI moves on unrooted networks *by swapping edges*

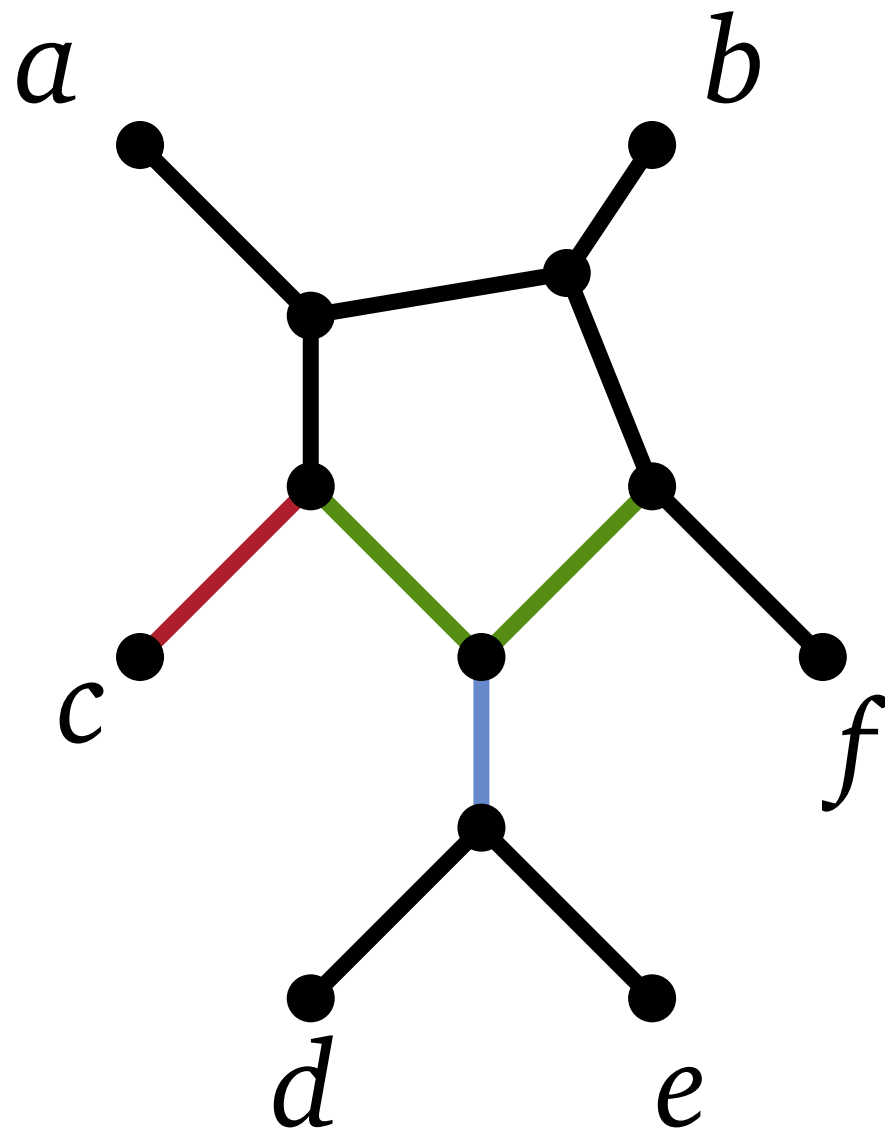
Definition. If an **unrooted** network has edges su, uv, vt but neither ut nor sv , then an **NNI** move consists of replacing the edges su, vt by edges ut, sv .



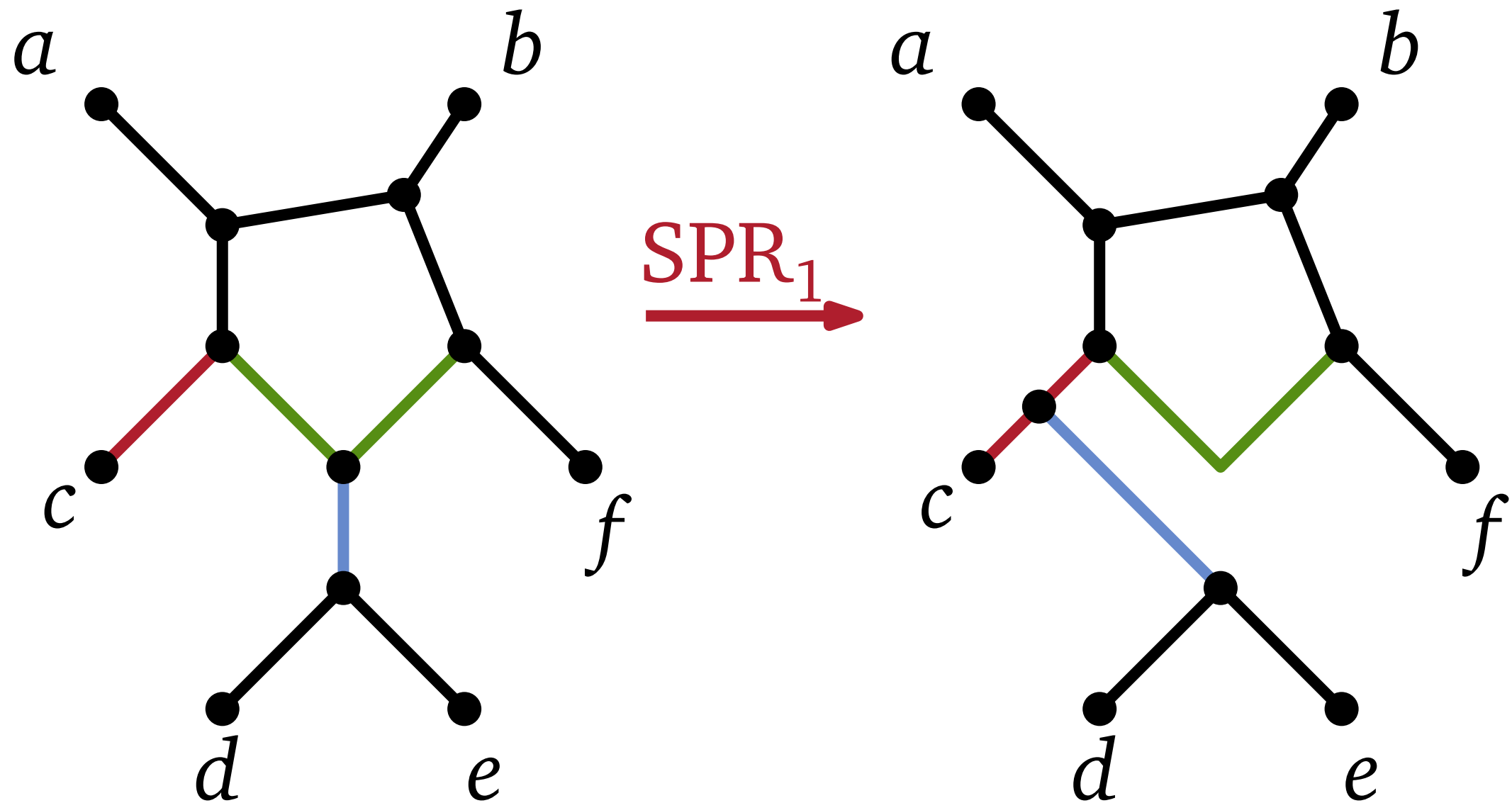
Example.



NNI moves on unrooted networks *as local SPR moves*

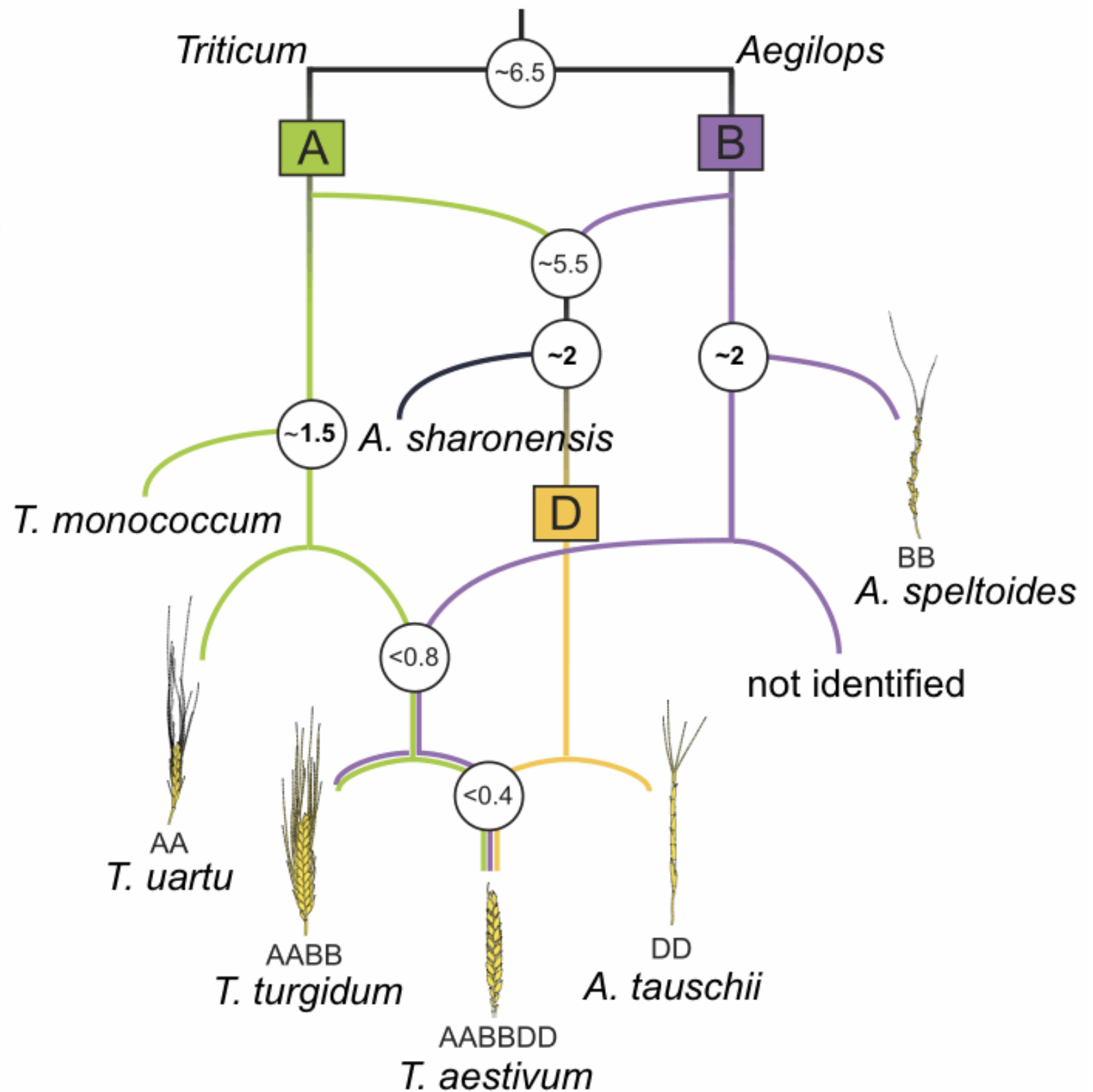
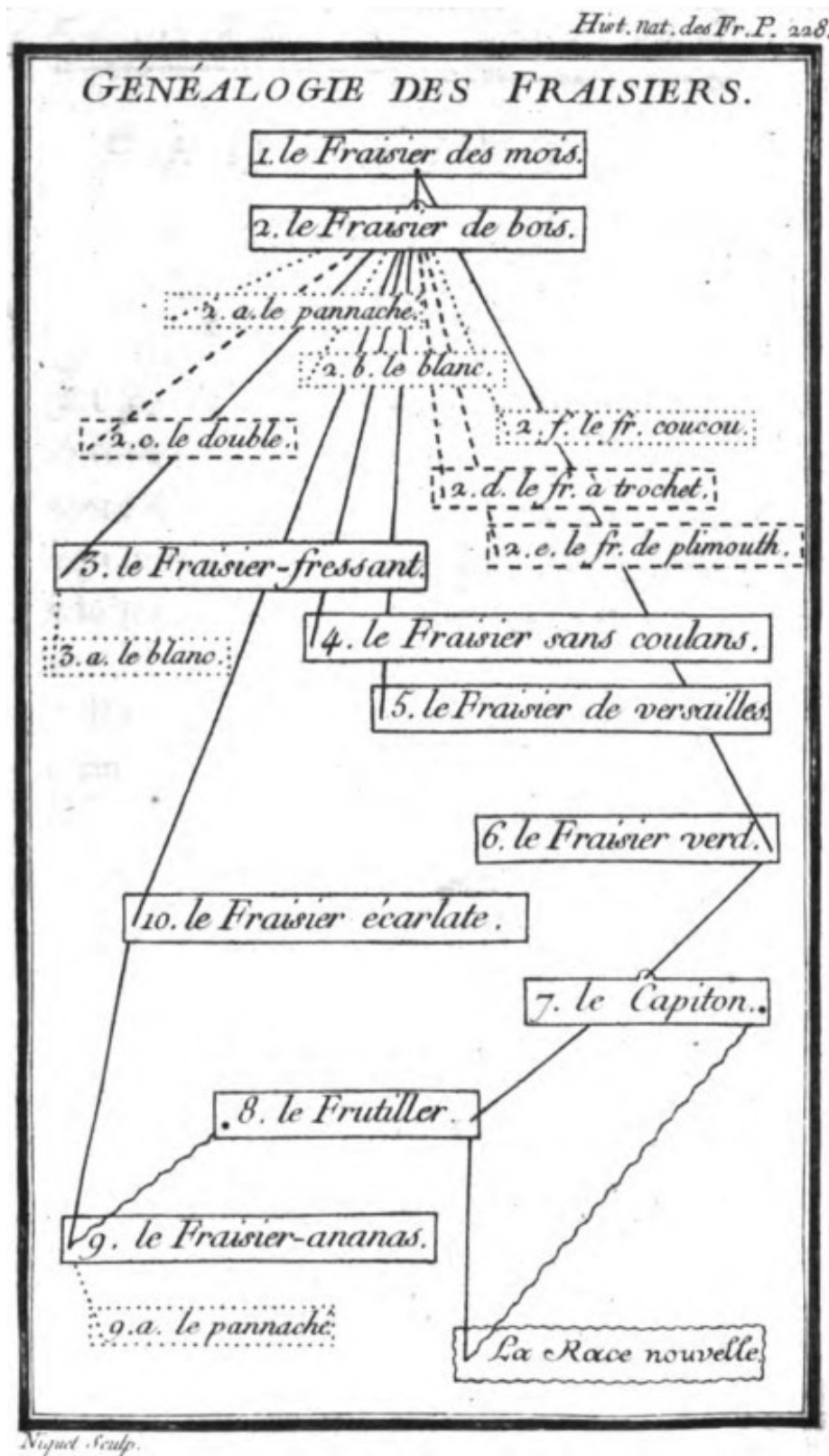


NNI moves on unrooted networks *as local SPR moves*



Only allowed when the green edges do not form a triangle with some third edge.

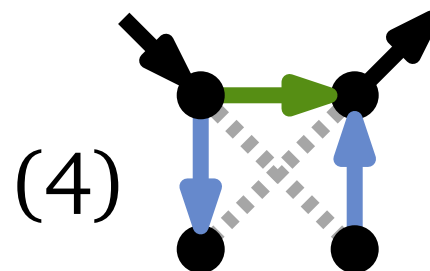
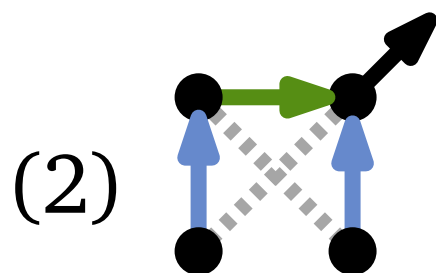
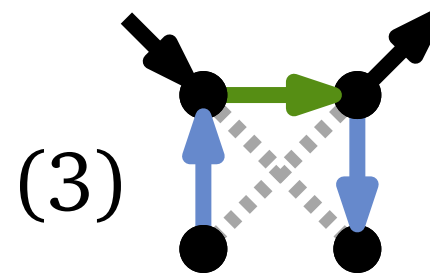
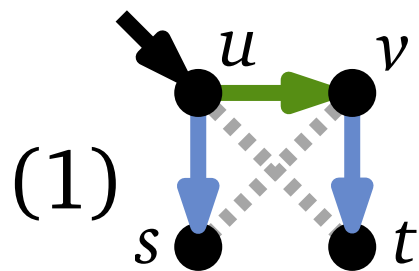
Rooted networks



rNNI moves on *rooted* networks

Definition. If a **rooted** network has arcs on su, uv and vt (in either direction) but not on ut and sv , then an *rNNI* move consists of replacing the arcs on su and vt by arcs on ut and sv such that:

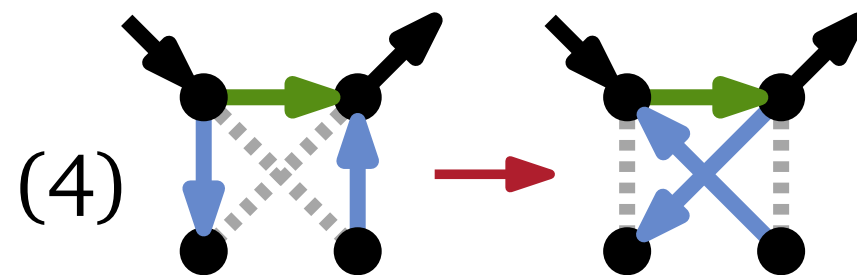
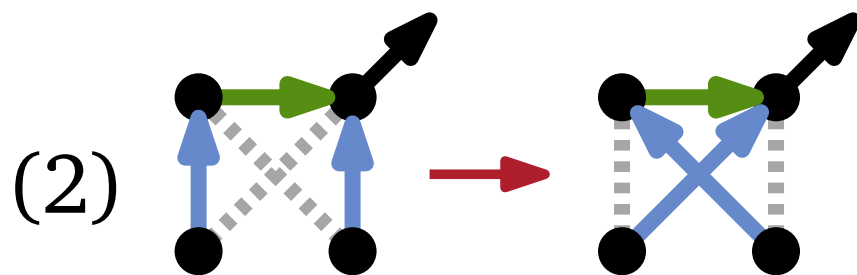
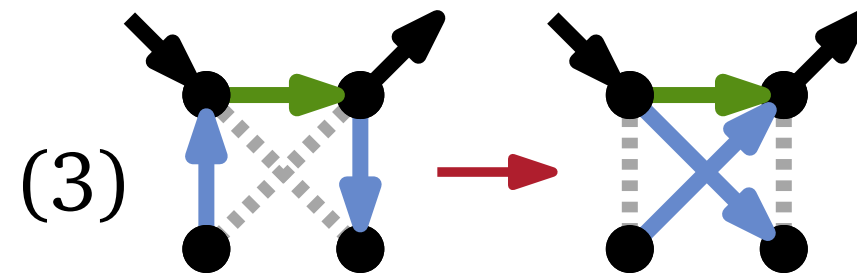
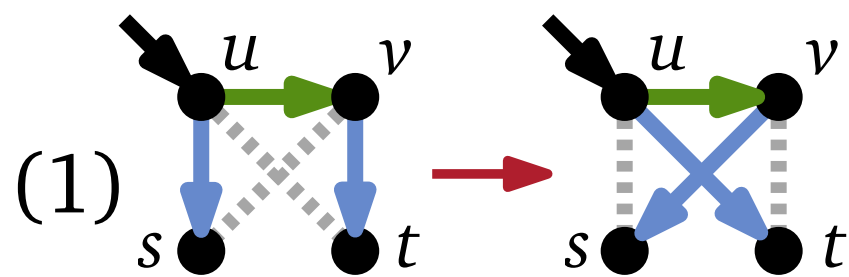
1. the in- and outdegrees of s and t are not affected by the move;
2. the in- and outdegrees of u and v remain at most 2 and
3. the obtained network is *acyclic*.



rNNI moves on *rooted* networks

Definition. If a **rooted** network has arcs on su, uv and vt (in either direction) but not on ut and sv , then an *rNNI* move consists of replacing the arcs on su and vt by arcs on ut and sv such that:

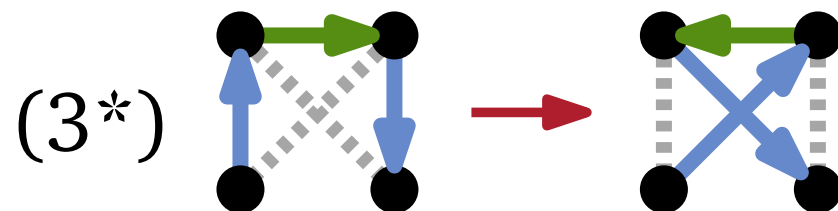
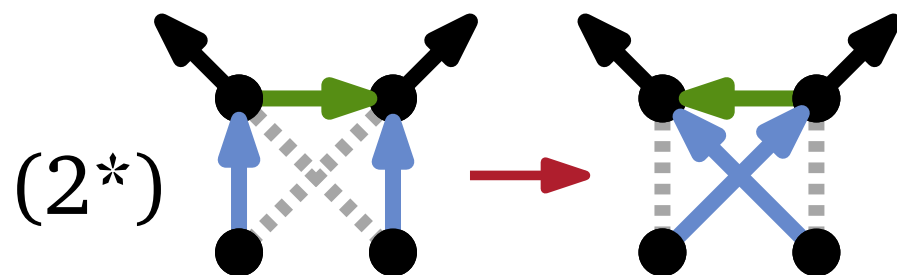
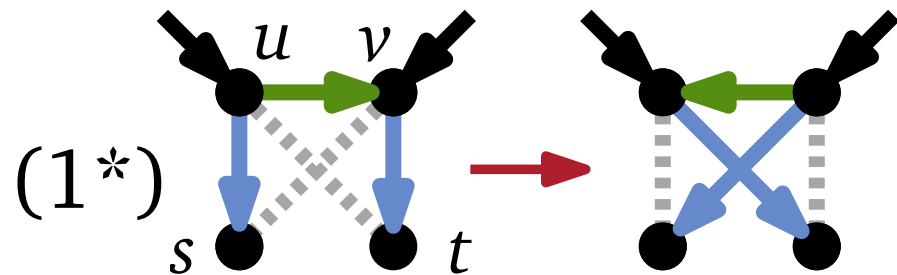
1. the in- and outdegrees of s and t are not affected by the move;
2. the in- and outdegrees of u and v remain at most 2 and
3. the obtained network is *acyclic*.



These moves are *only valid* if the resulting network is acyclic.

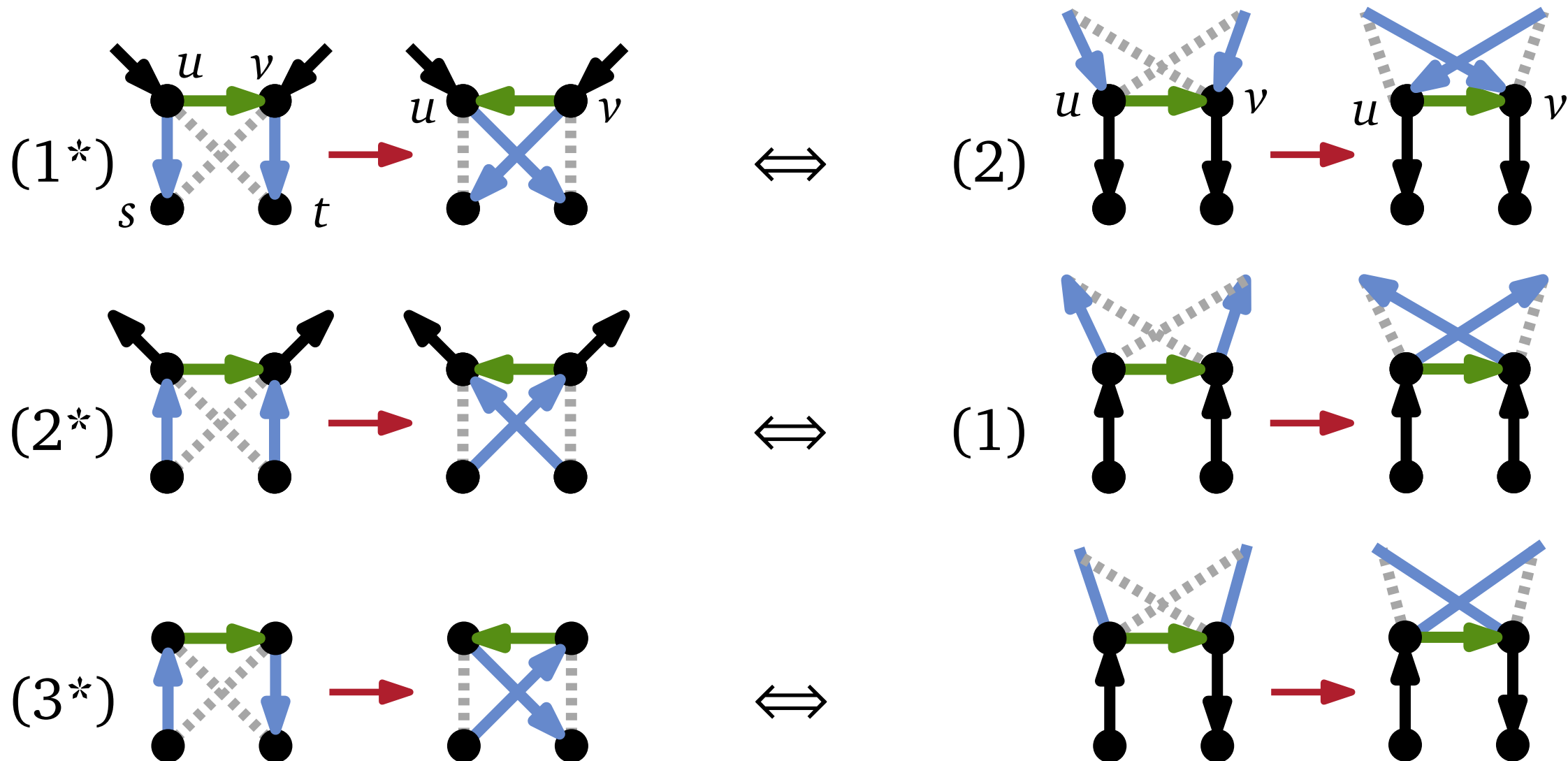
Additional possibilities

It will be useful to also allow reversing the arc on uv . Such moves are equivalent to one of the moves (1), (2), (3) or (4):



Additional possibilities

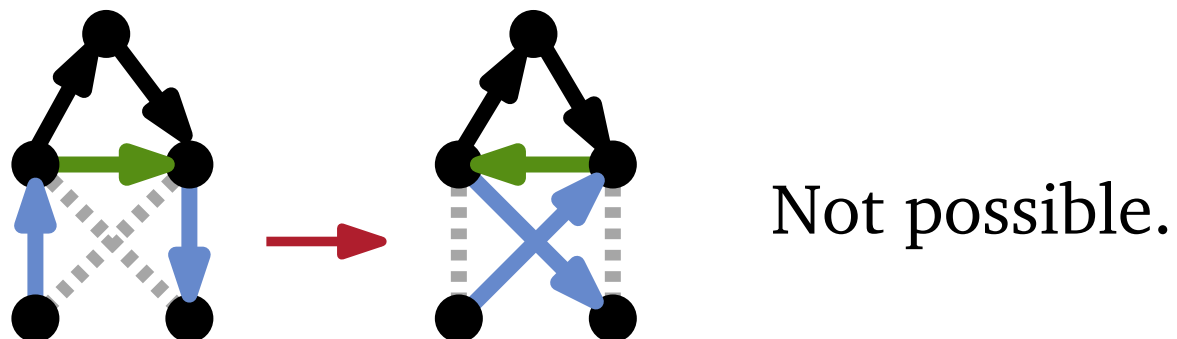
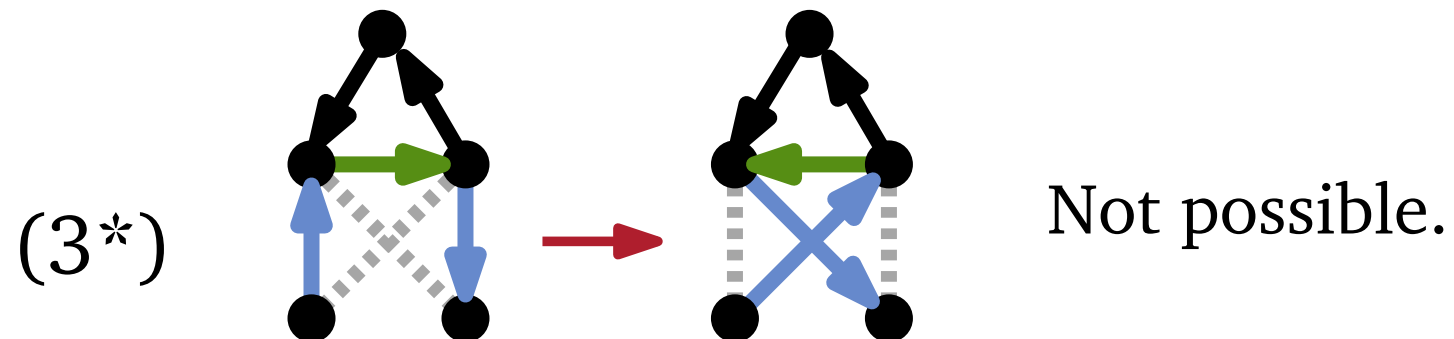
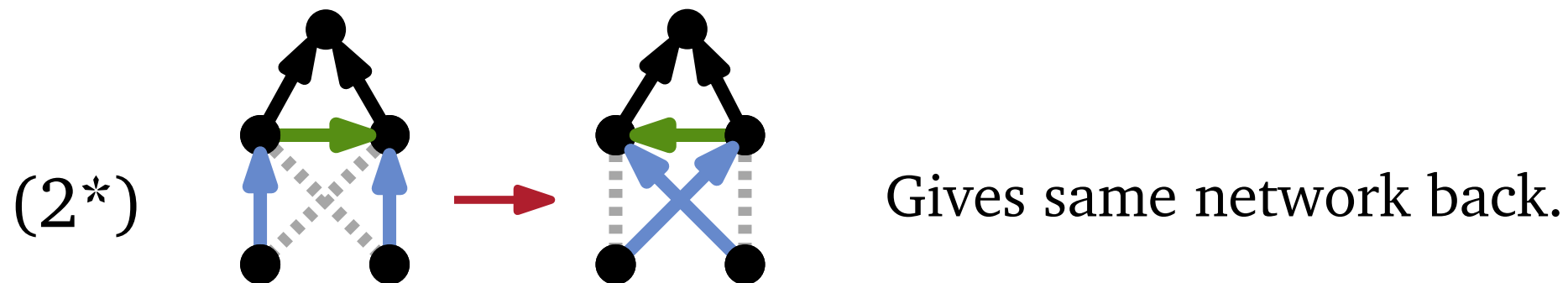
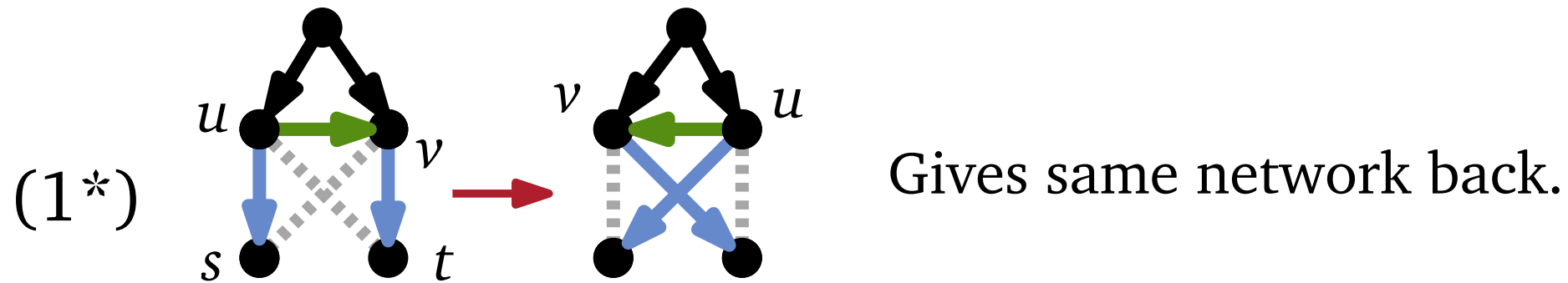
It will be useful to also allow reversing the arc on uv . Such moves are equivalent to one of the moves (1), (2), (3) or (4):



- In case (4) it is not possible to reverse arc uv and keep a binary network.
- If u, v have a common neighbour then these moves are not possible or give the same network back (see next slide).

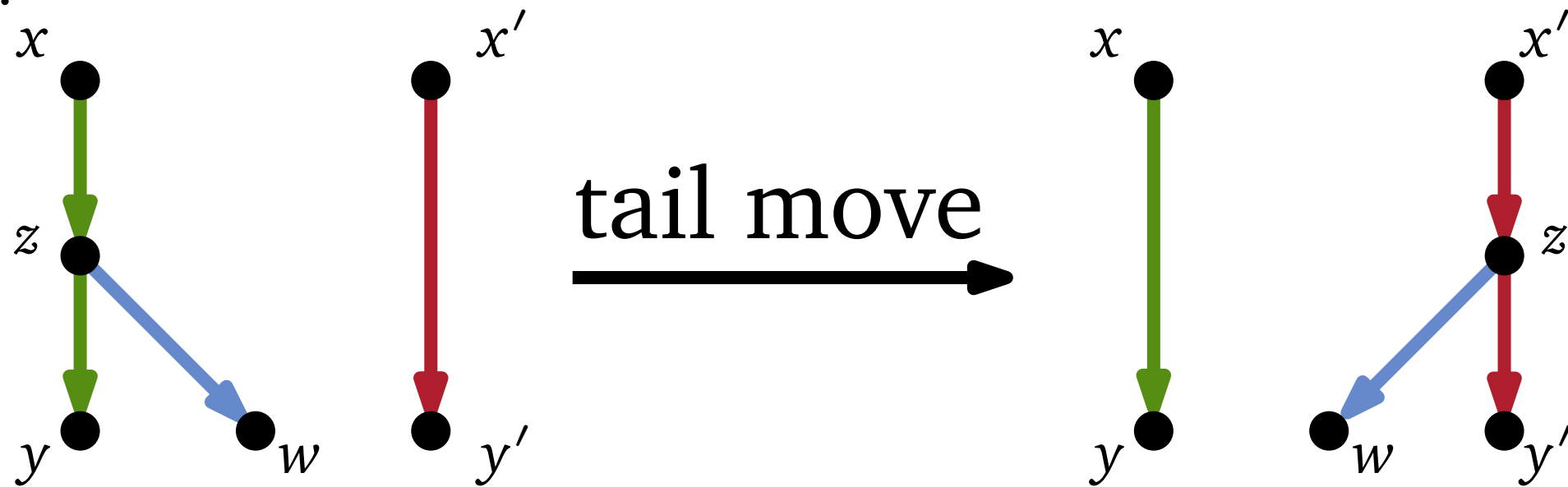
Additional possibilities

The case that u and v have a common neighbour:



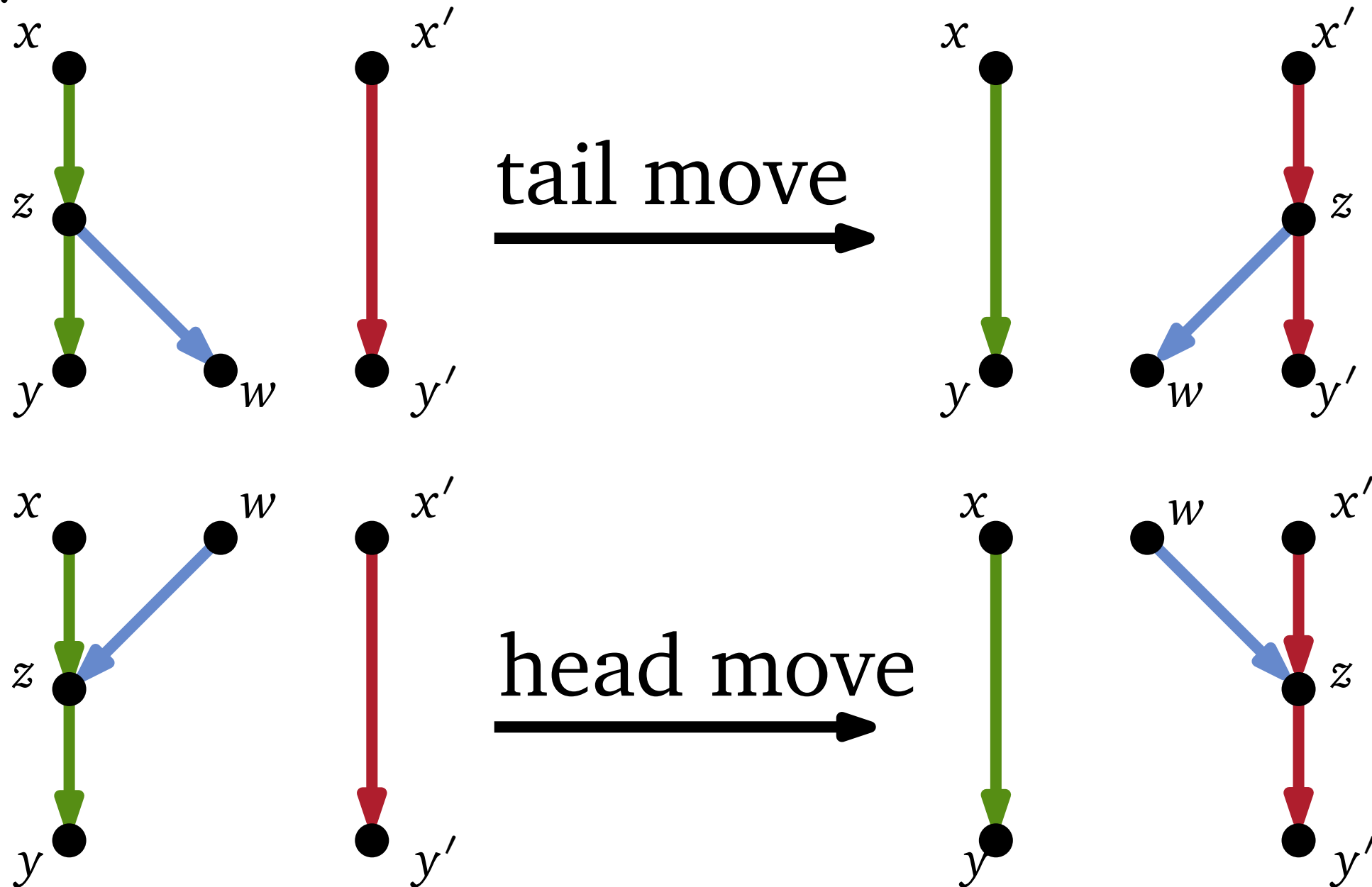
rSPR moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



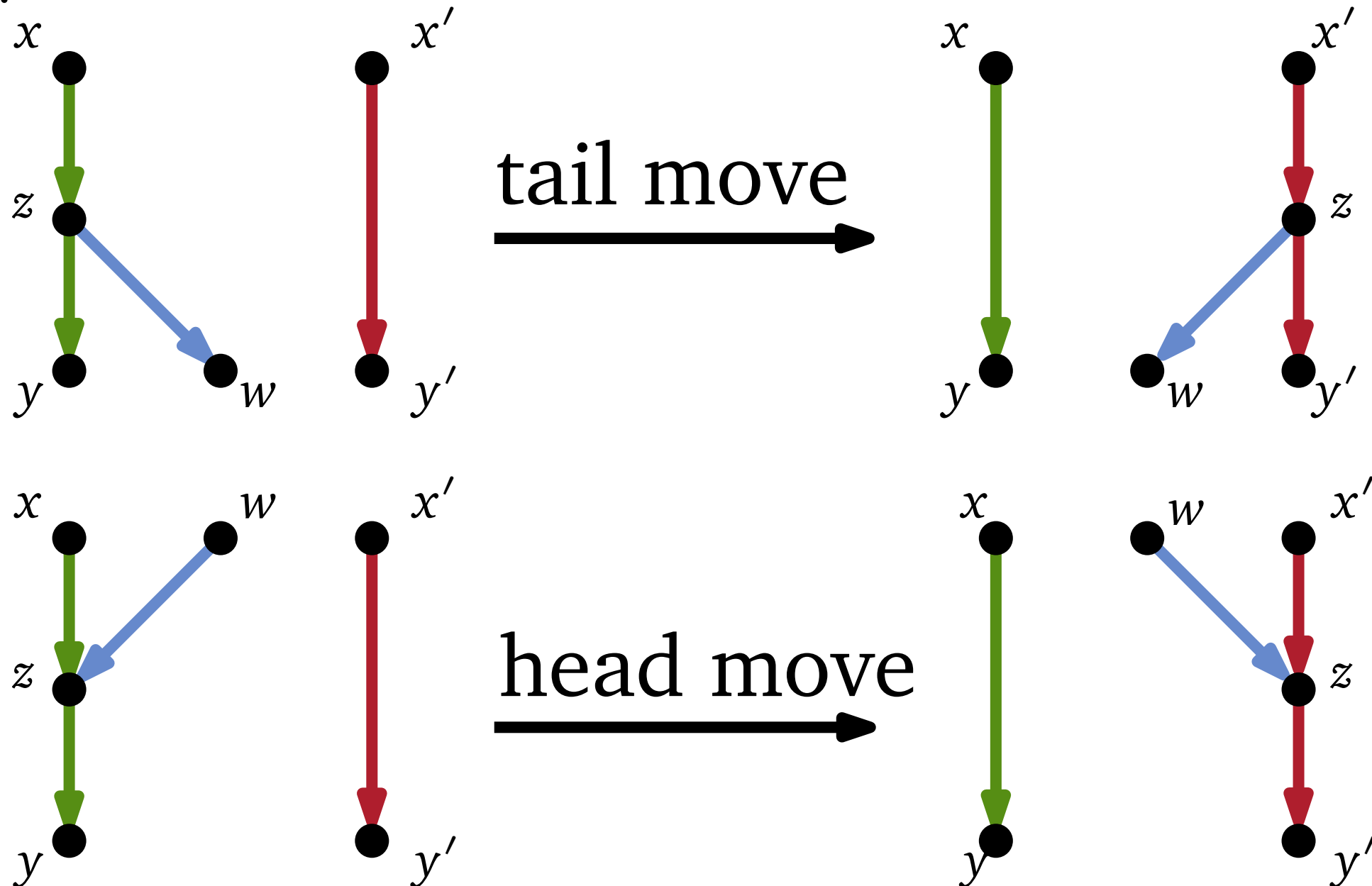
rSPR moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



rSPR moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



The green arcs zx, zy are the *donor arcs* and the red arc $x'y'$ is the *recipient arc*.

rNNI moves as local rSPR moves

An $rSPR_1$ move (or *distance-1 rSPR move*) is an rSPR move where the recipient arc is incident to one of the donor arcs.

rNNI moves as local rSPR moves

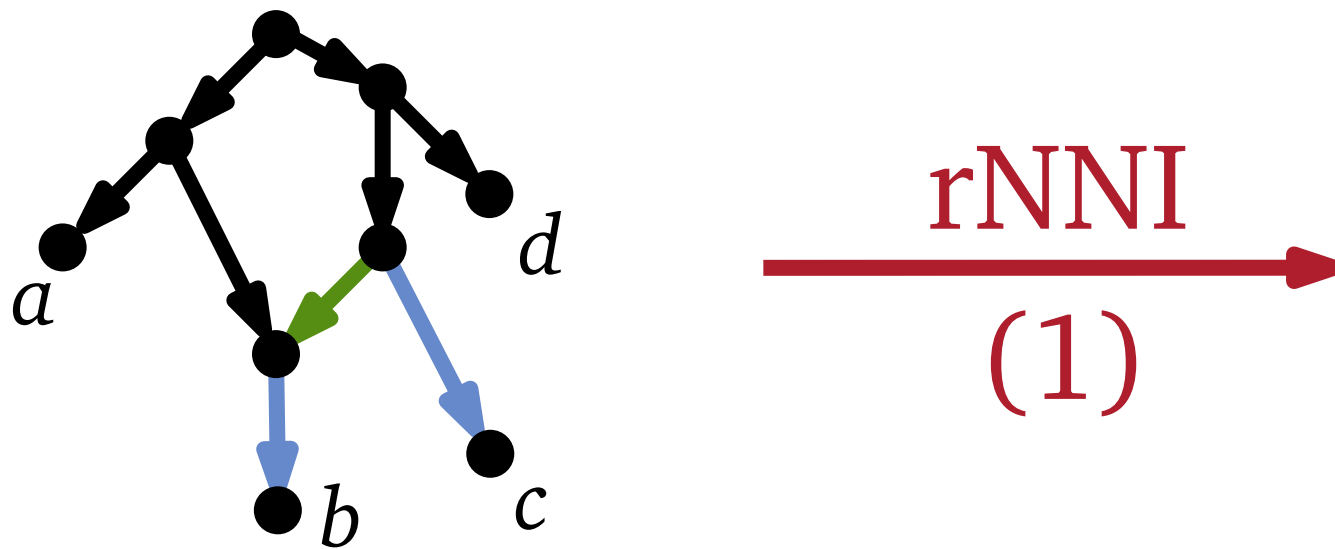
An ***rSPR₁*** move (or *distance-1 rSPR move*) is an rSPR move where the recipient arc is incident to one of the donor arcs.

Theorem (Gambette et al., 2017). N can be turned into N' by one rNNI move if and only if N can be turned into N' by one rSPR₁ move.

rNNI moves as local rSPR moves

An *rSPR₁* move (or *distance-1 rSPR move*) is an rSPR move where the recipient arc is incident to one of the donor arcs.

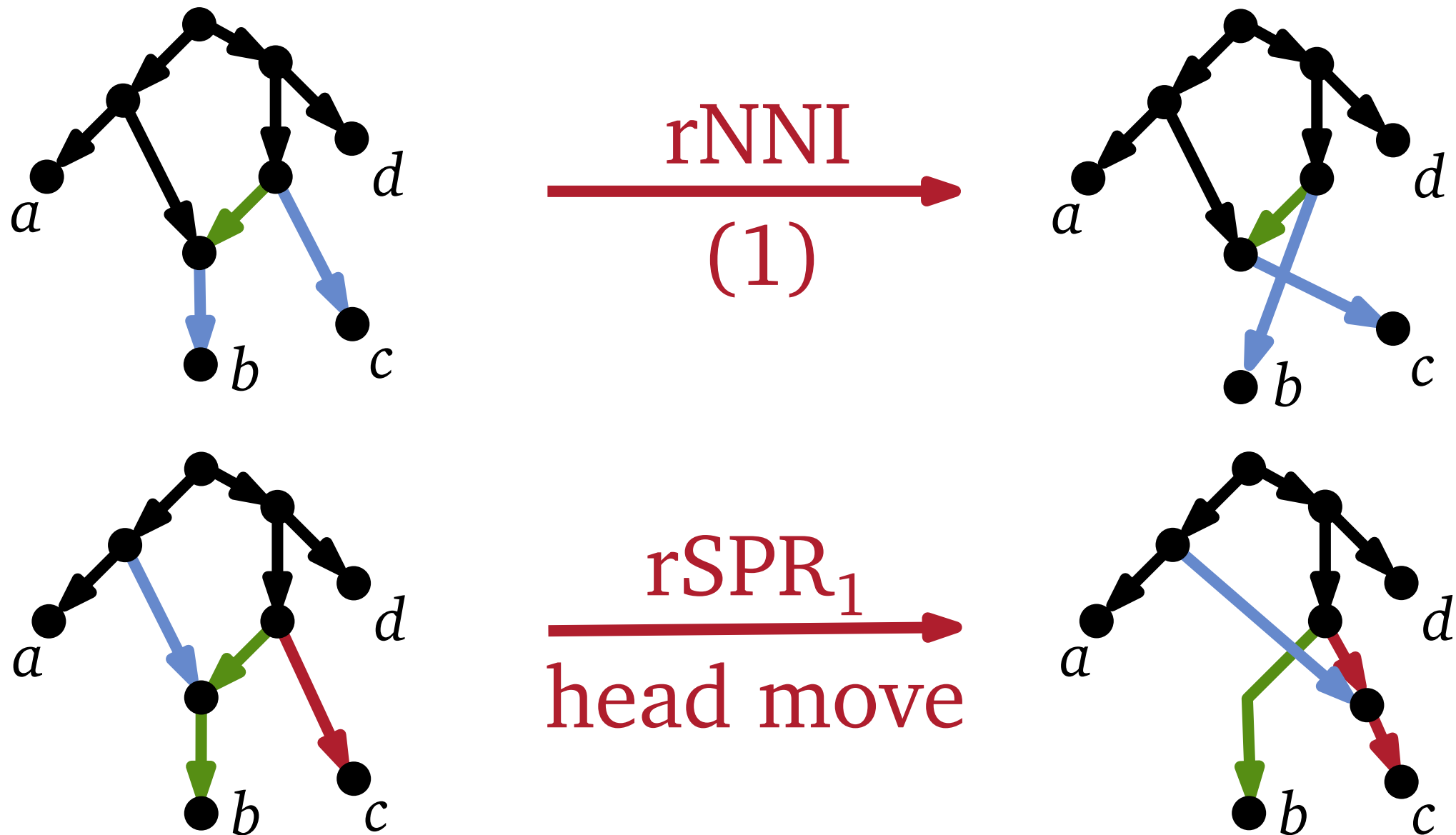
Theorem (Gambette et al., 2017). N can be turned into N' by one rNNI move if and only if N can be turned into N' by one rSPR₁ move.



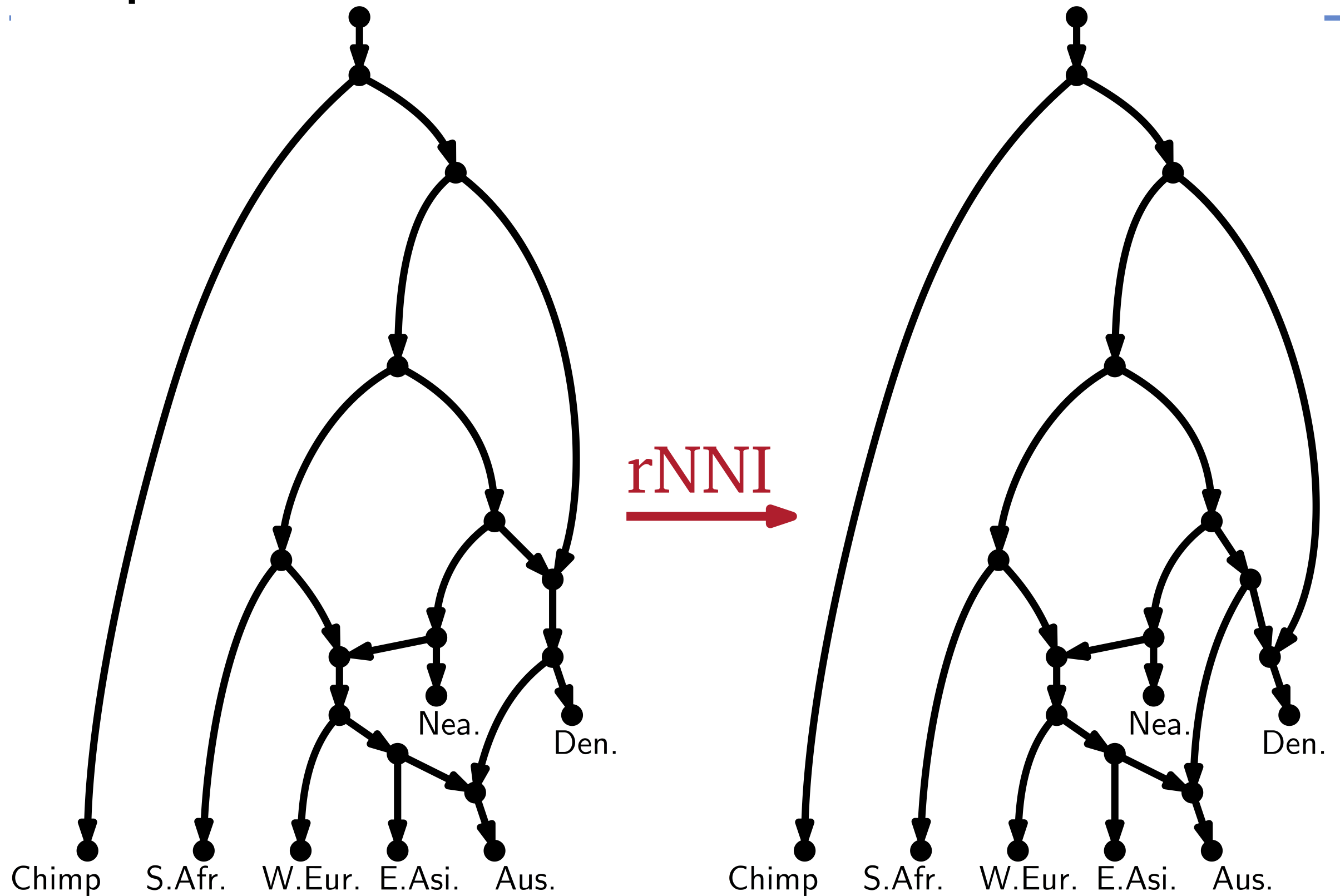
rNNI moves as local rSPR moves

An *rSPR*₁ move (or *distance-1 rSPR move*) is an rSPR move where the recipient arc is incident to one of the donor arcs.

Theorem (Gambette et al., 2017). N can be turned into N' by one rNNI move if and only if N can be turned into N' by one rSPR₁ move.



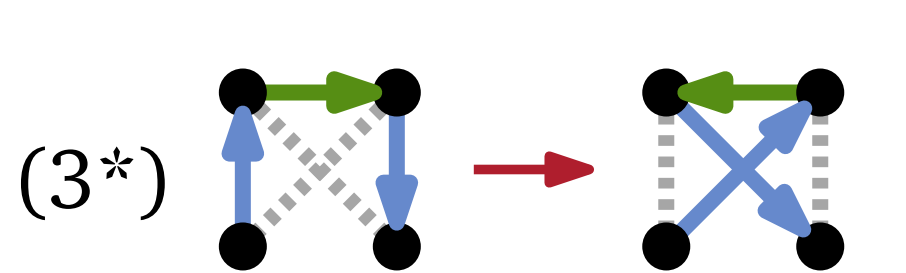
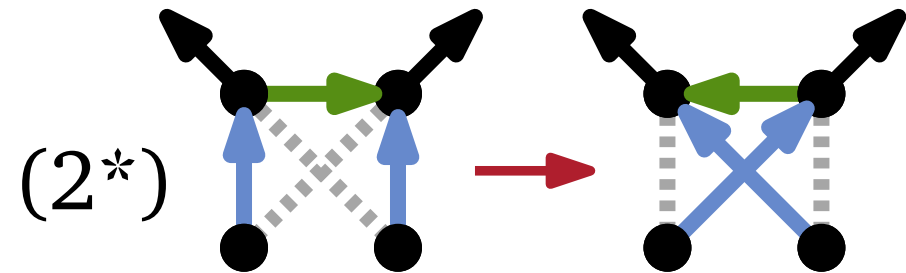
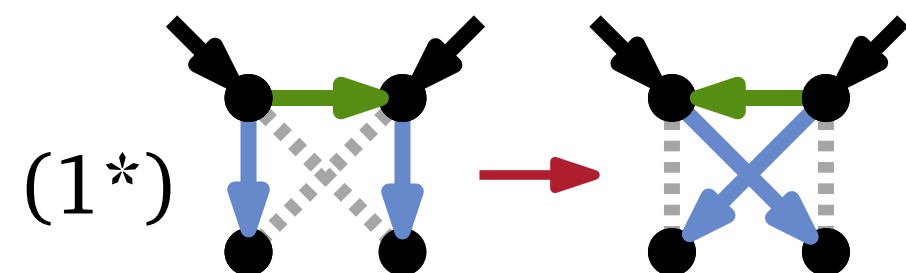
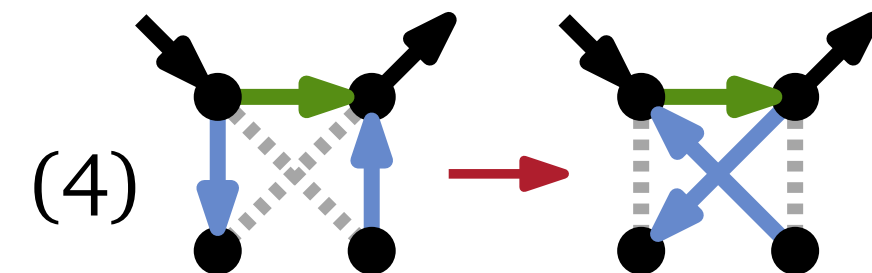
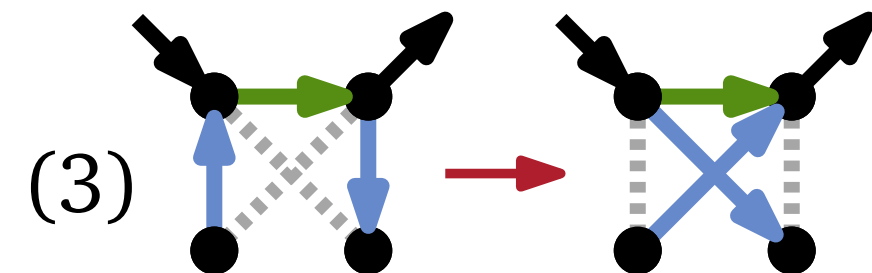
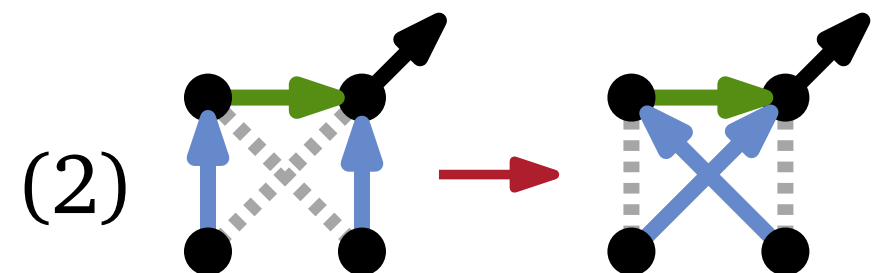
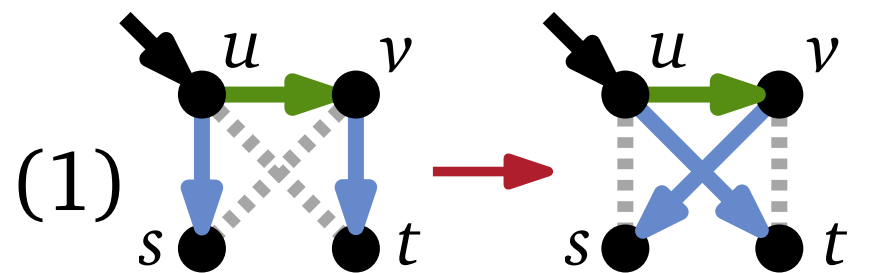
Example



Mallick, Li, Lipson, Mathieson, Gymrek, Racimo, et al. *Nature* (2016).

Types of rNNI moves

All possible moves:



We may restrict (equivalently) to either:

- moves (1), (2), (3) and (4); or
- moves (1), (2) and (3*)

Reticulate complexity

Proposition. For binary rooted networks N_1, N_2 on X , the following are equivalent:

1. N_1 and N_2 have the same number of reticulations (indegree-2 vertices);
 2. N_1 and N_2 have the same number of vertices;
 3. N_1 and N_2 require the same number of arc removals to be turned into a rooted phylogenetic tree.
-

Reticulate complexity

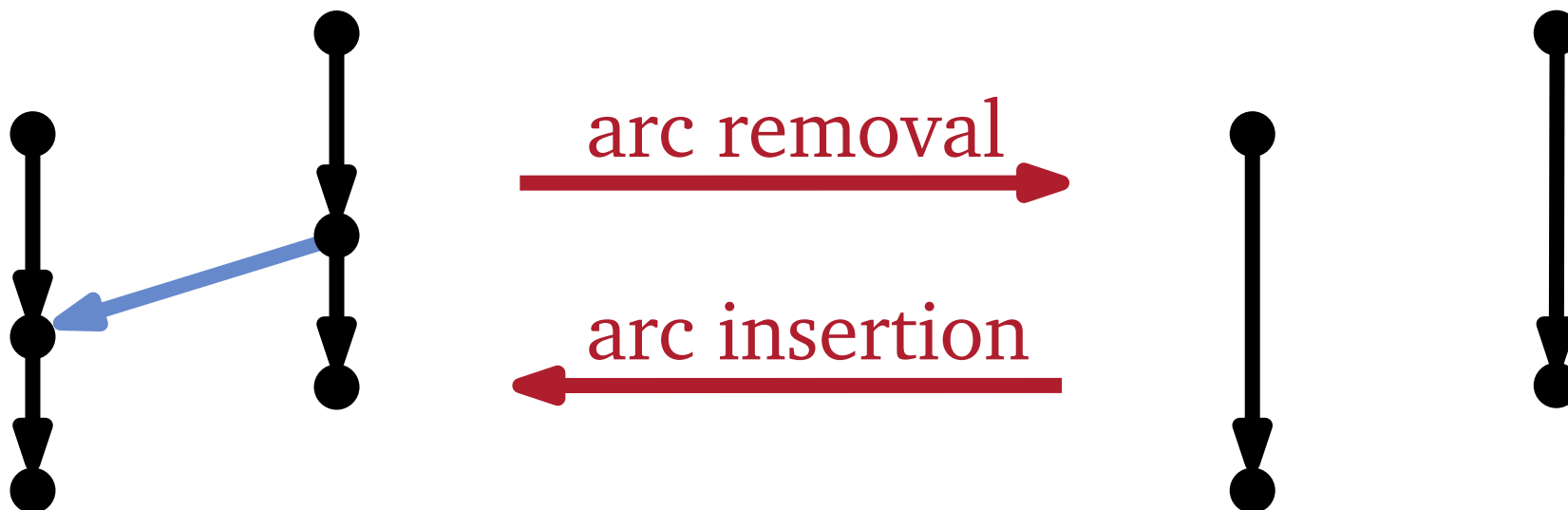
Proposition. For binary rooted networks N_1, N_2 on X , the following are equivalent:

1. N_1 and N_2 have the same number of reticulations (indegree-2 vertices);
 2. N_1 and N_2 have the same number of vertices;
 3. N_1 and N_2 require the same number of arc removals to be turned into a rooted phylogenetic tree.
- The **reticulate complexity** of a binary network can be defined as the number of reticulations.
 - rNNI and rSPR are **horizontal moves**, i.e., they do not change the reticulate complexity.
 - **Vertical moves** do change the reticulate complexity, for example:

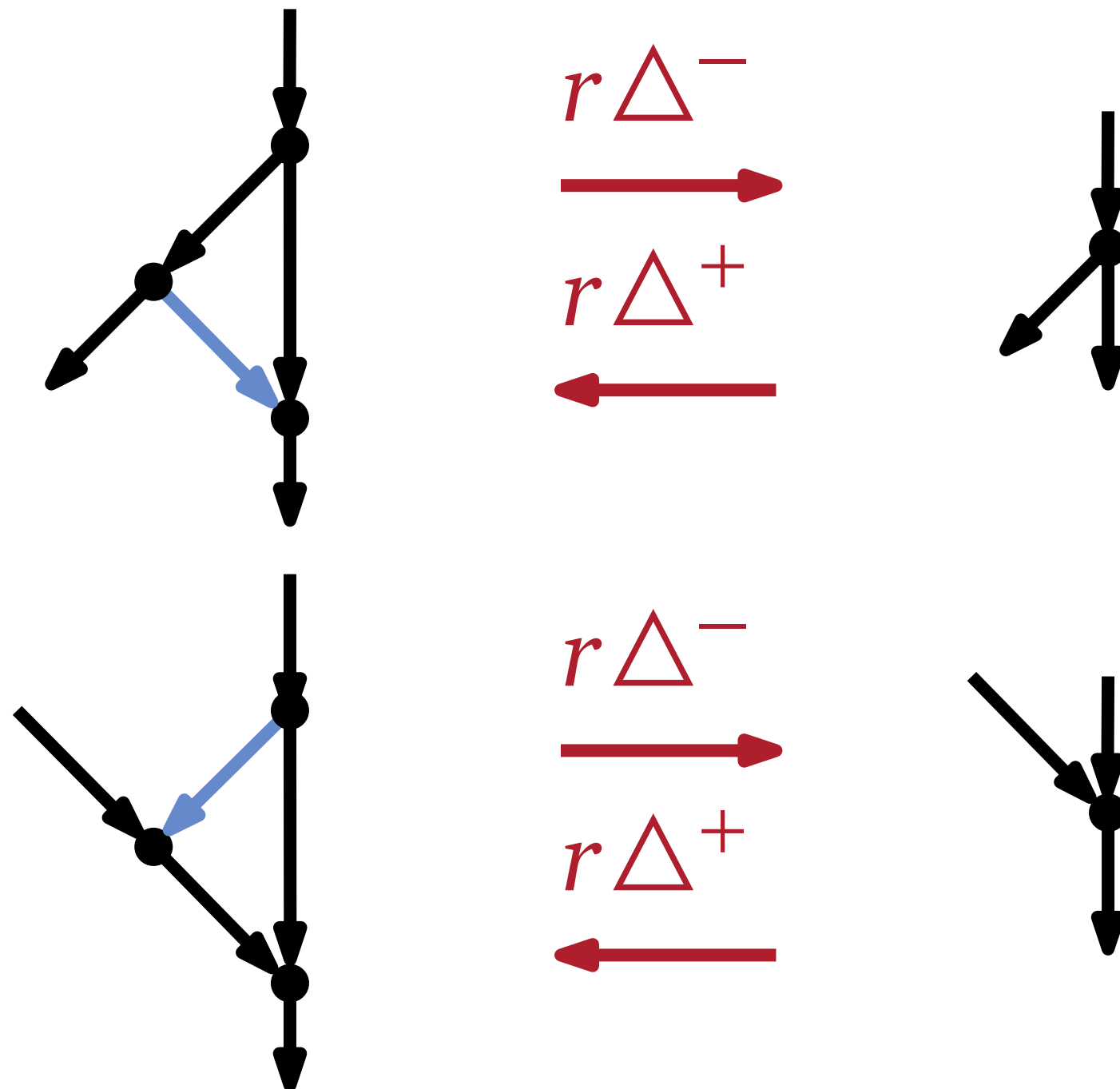
Reticulate complexity

Proposition. For binary rooted networks N_1, N_2 on X , the following are equivalent:

1. N_1 and N_2 have the same number of reticulations (indegree-2 vertices);
 2. N_1 and N_2 have the same number of vertices;
 3. N_1 and N_2 require the same number of arc removals to be turned into a rooted phylogenetic tree.
- The **reticulate complexity** of a binary network can be defined as the number of reticulations.
 - rNNI and rSPR are **horizontal moves**, i.e., they do not change the reticulate complexity.
 - **Vertical moves** do change the reticulate complexity, for example:



Other example of *vertical moves*

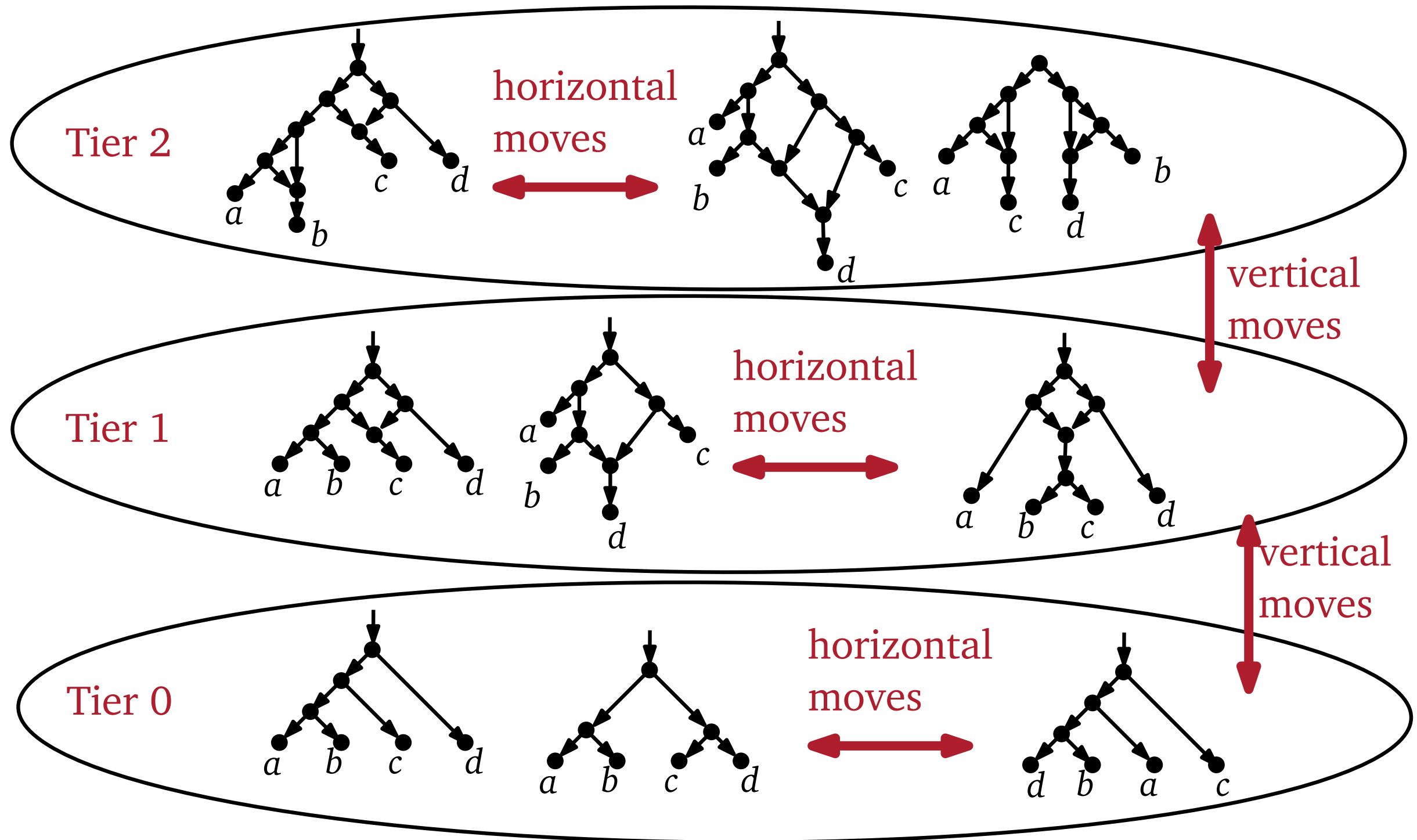


$r\Delta^+$ moves are a *local* version of arc insertions.

$r\Delta^-$ moves are a *local* version of arc removals.

Phylogenetic network space

The *tiers* of phylogenetic network space consist of networks of the same reticulate complexity.



Results on *rNNI* moves

Theorem (Gambette et al. 2017). For any two rooted networks N, N' on X with the same reticulate complexity, there exists a sequence of rNNI moves turning N into N' .

Results on *rNNI* moves

Theorem (Gambette et al. 2017). For any two rooted networks N, N' on X with the same reticulate complexity, there exists a sequence of rNNI moves turning N into N' .

Corollary (Gambette et al. 2017). For any two rooted networks N, N' and for any vertical move C^+ that can turn any network into a network with one higher reticulate complexity and has a reverse C^- , there exists either

- a sequence of rNNI and C^+ moves turning N into N' ; or
 - a sequence of rNNI and C^- moves turning N into N' .
-

Results on *rNNI* moves

Theorem (Gambette et al. 2017). For any two rooted networks N, N' on X with the same reticulate complexity, there exists a sequence of rNNI moves turning N into N' .

Corollary (Gambette et al. 2017). For any two rooted networks N, N' and for any vertical move C^+ that can turn any network into a network with one higher reticulate complexity and has a reverse C^- , there exists either

- a sequence of rNNI and C^+ moves turning N into N' ; or
- a sequence of rNNI and C^- moves turning N into N' .

Note: The moves C^+, C^- can for example be arc insertions/removals or $r\Delta^+, r\Delta^-$ moves.

Both results automatically also hold for $rSPR_1$, and hence for rSPR.

Results on *rNNI* moves

Theorem (Gambette et al. 2017). For any two rooted networks N, N' on X with the same reticulate complexity, there exists a sequence of rNNI moves turning N into N' .

Corollary (Gambette et al. 2017). For any two rooted networks N, N' and for any vertical move C^+ that can turn any network into a network with one higher reticulate complexity and has a reverse C^- , there exists either

- a sequence of rNNI and C^+ moves turning N into N' ; or
- a sequence of rNNI and C^- moves turning N into N' .

Note: The moves C^+, C^- can for example be arc insertions/removals or $r\Delta^+, r\Delta^-$ moves.

Both results automatically also hold for $rSPR_1$, and hence for rSPR.

Observation. rNNI moves are *reversible*, i.e. if N' can be obtained from N by one rNNI move, then it is also possible to obtain N from N' by a single rNNI move (same for rSPR).

A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

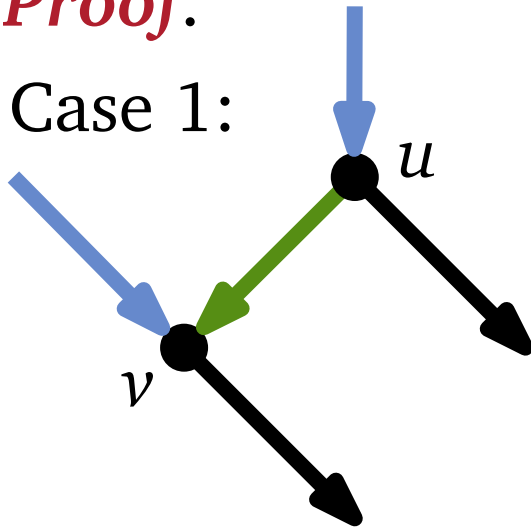
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

Proof.

Case 1:



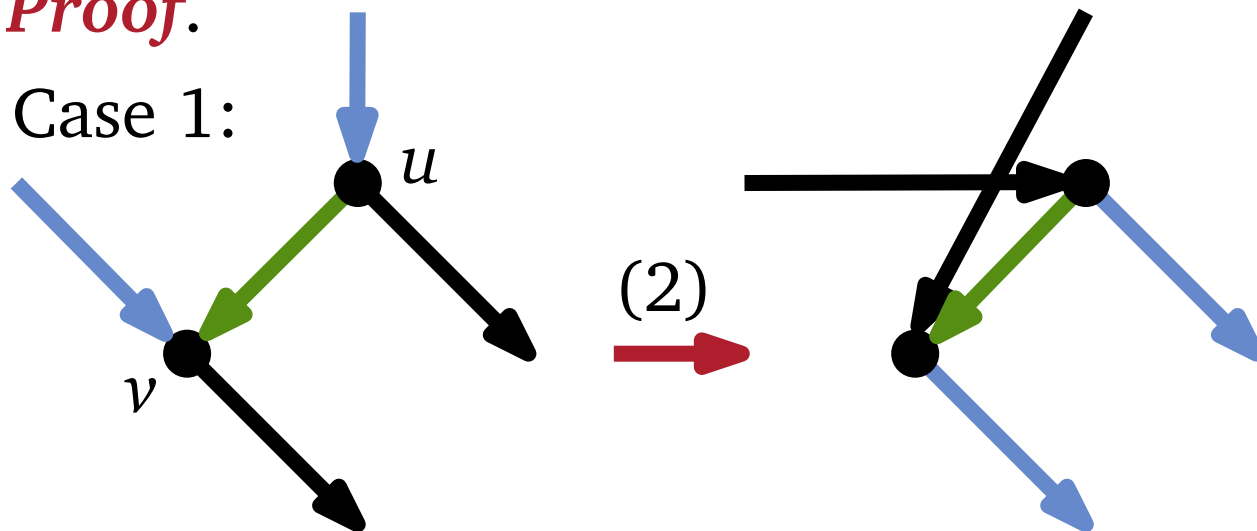
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

Proof.

Case 1:



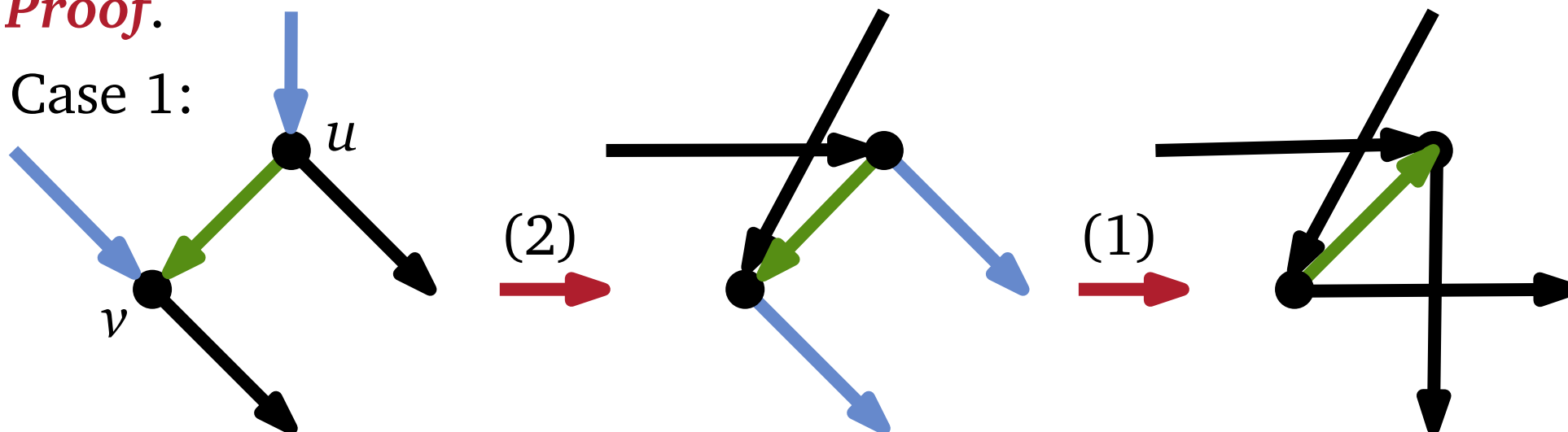
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

Proof.

Case 1:



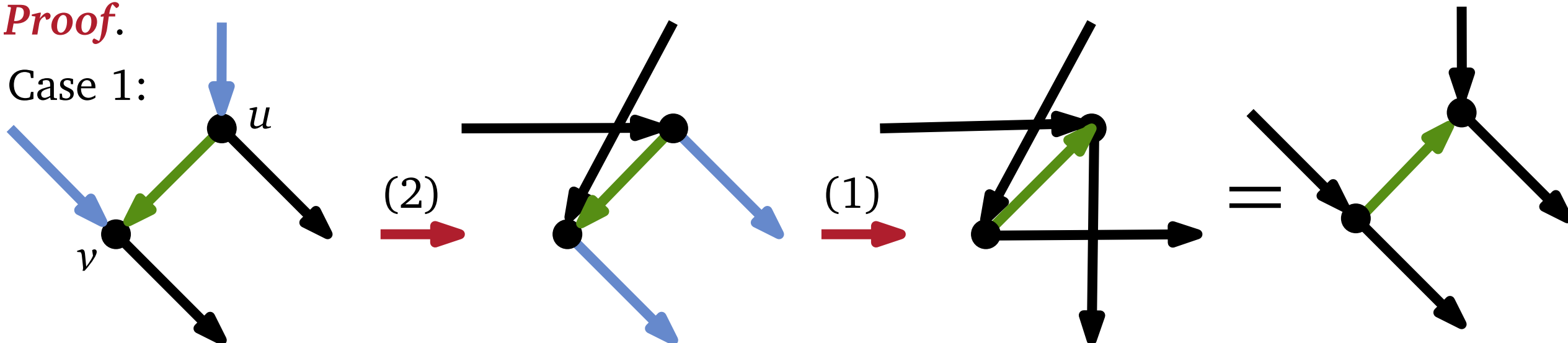
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

Proof.

Case 1:



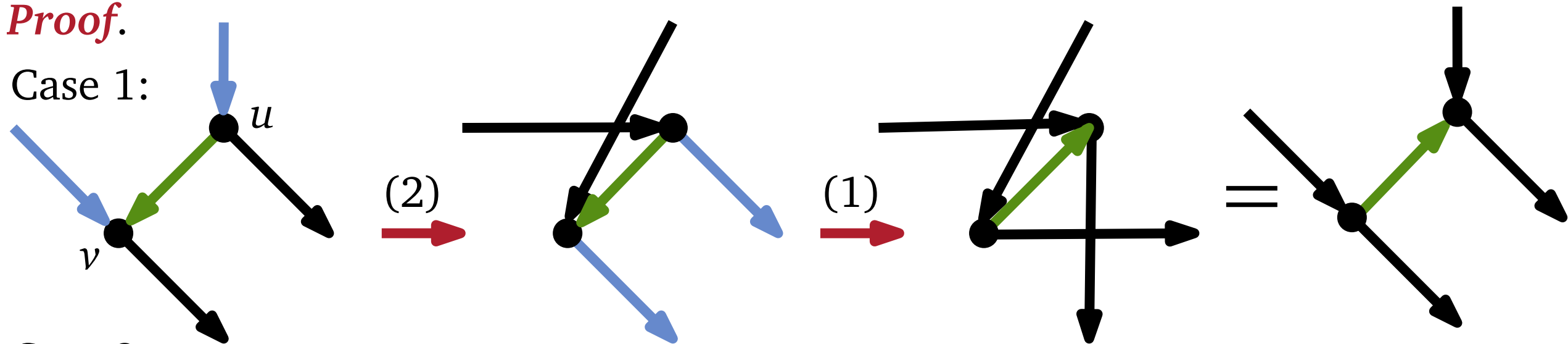
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

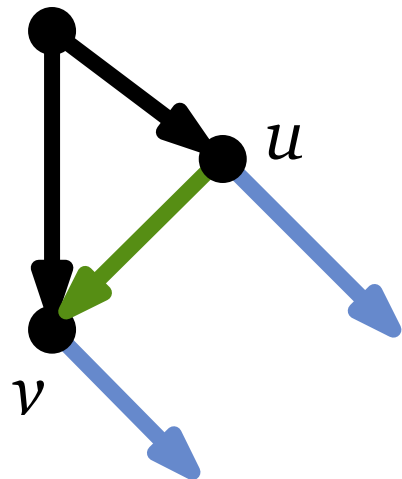
Lemma. If N, N' are not isomorphic and N' can be obtained from N by an *arc-flip*, then N' can be obtained from N by at most *two rNNI* moves.

Proof.

Case 1:



Case 2:



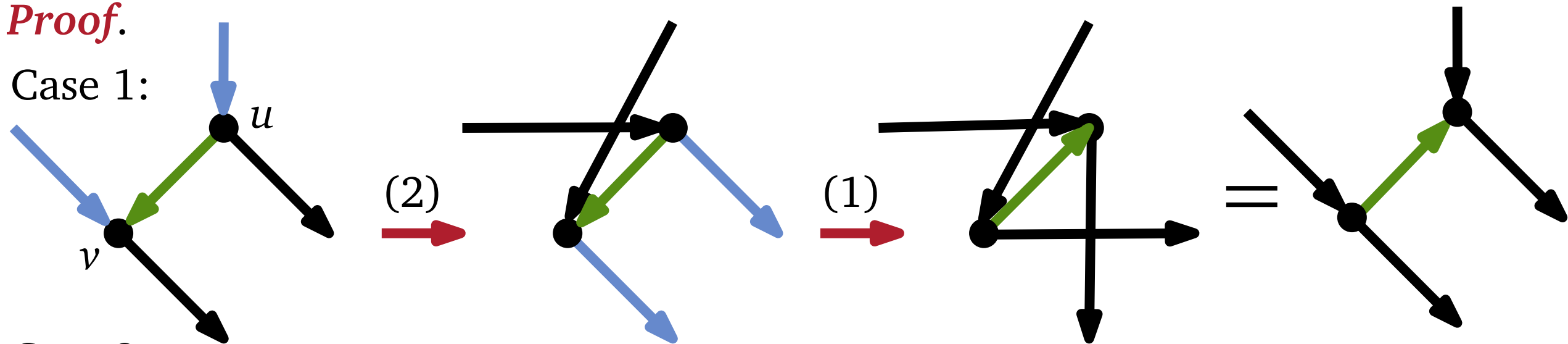
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two **unrooted** networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

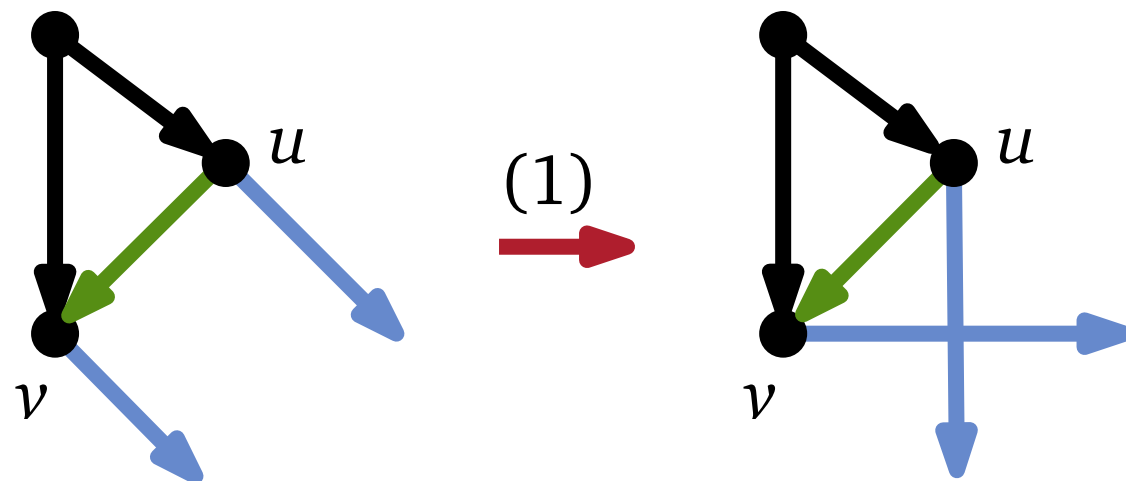
Lemma. If N, N' are not isomorphic and N' can be obtained from N by an **arc-flip**, then N' can be obtained from N by at most **two rNNI** moves.

Proof.

Case 1:



Case 2:



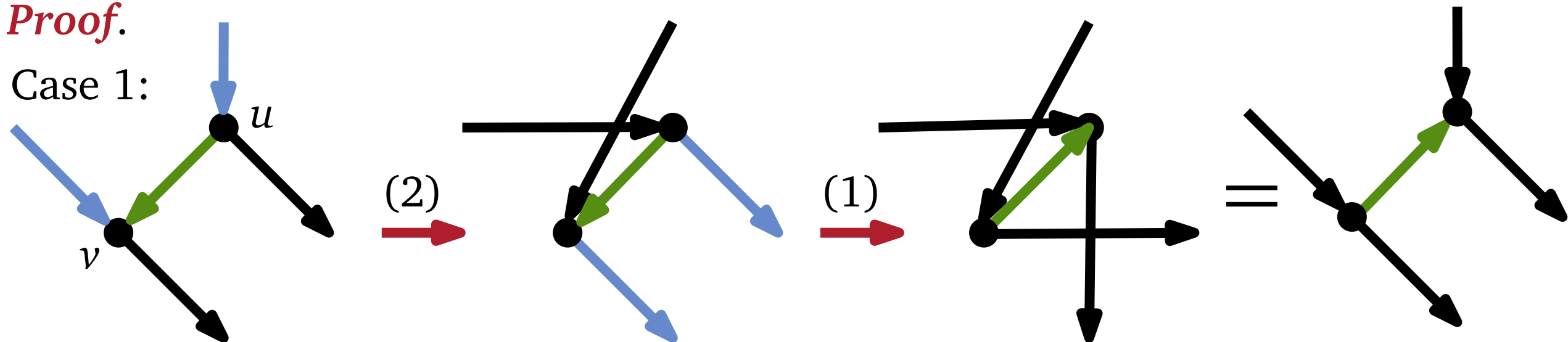
A recent result and a useful lemma

Theorem (Huber et al. 2016). For any two **unrooted** networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

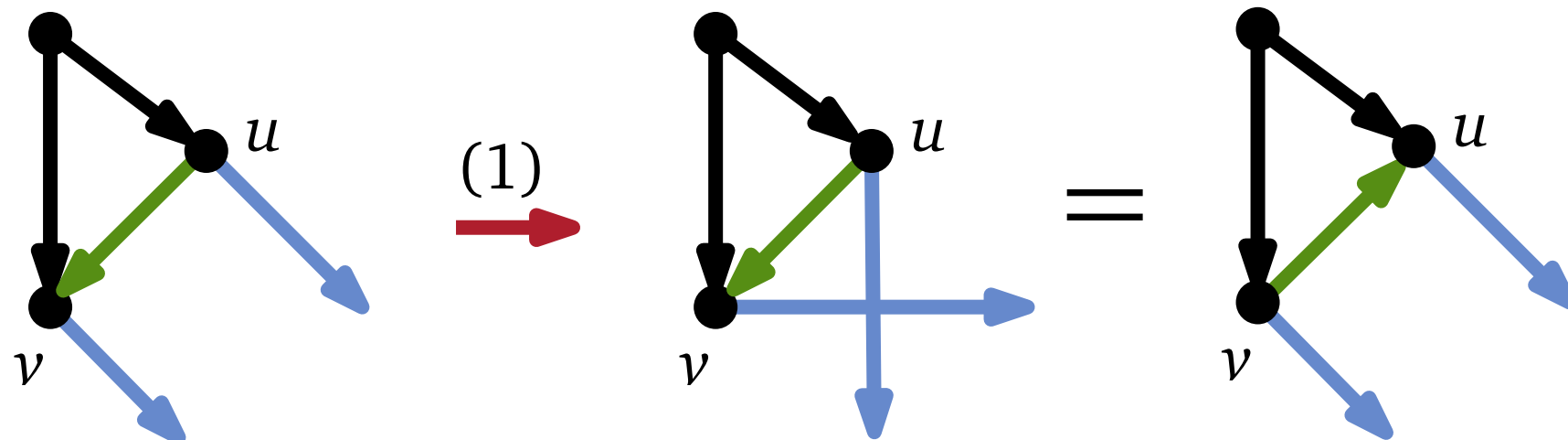
Lemma. If N, N' are not isomorphic and N' can be obtained from N by an **arc-flip**, then N' can be obtained from N by at most **two rNNI** moves.

Proof.

Case 1:



Case 2:



First proof idea

Given two *rooted* networks N, N' on X with the same reticulate complexity, we could try to do the following:

1. turn N into a network M with the *same underlying unrooted network* as N' ;
2. turn M into N' by a sequence of *arc flips*;
3. replace each arc flip by at most two rNNI moves.

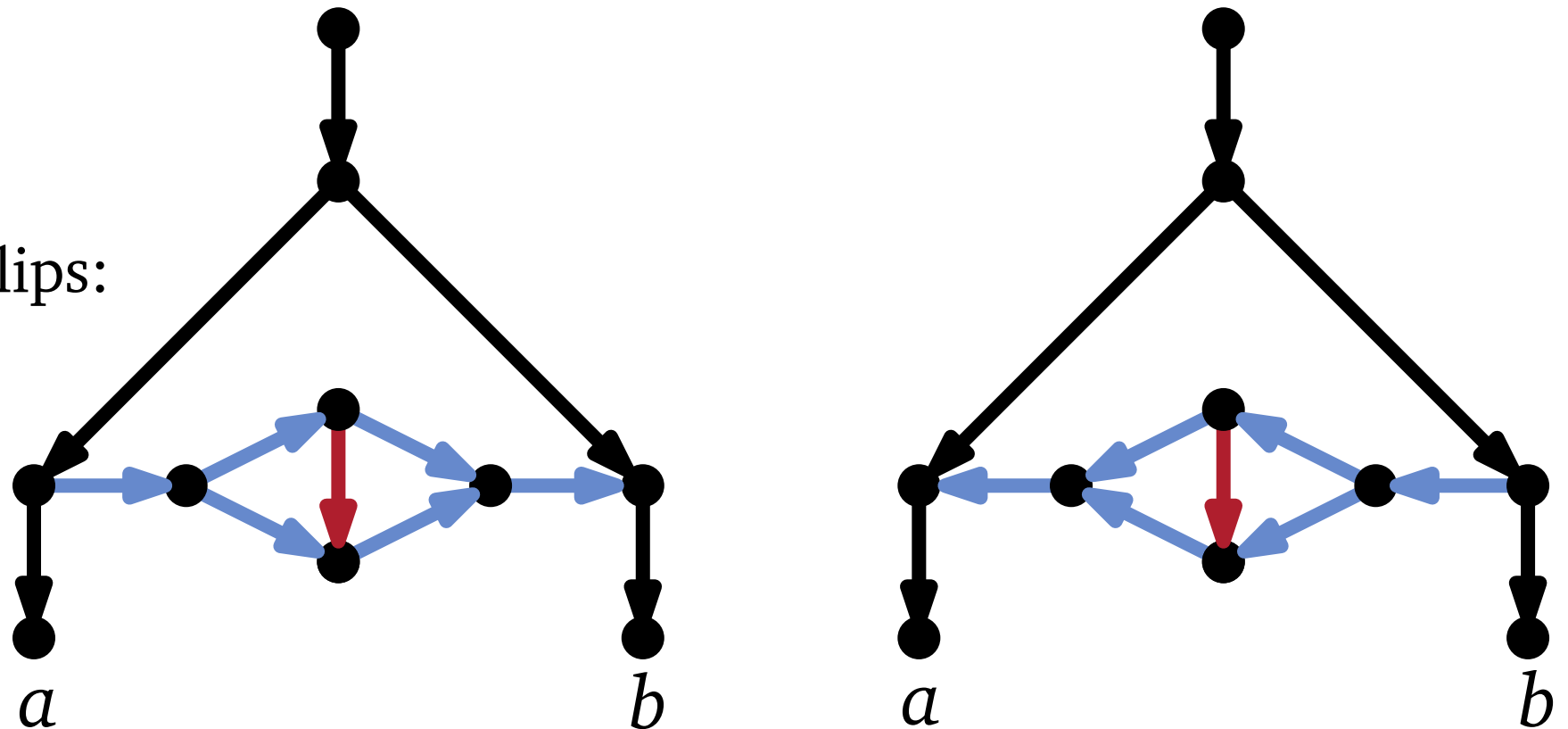
First proof idea

Given two *rooted* networks N, N' on X with the same reticulate complexity, we could try to do the following:

1. turn N into a network M with the *same underlying unrooted network* as N' ;
2. turn M into N' by a sequence of *arc flips*;
3. replace each arc flip by at most two rNNI moves.

Unfortunately, Step 2.
is not always possible.

These networks are
not connected by arc flips:



New idea

Definition. We call a rooted network *flip-friendly* if it can be turned into any other rooted network with the same underlying unrooted network by a sequence of arc-flips.

New idea

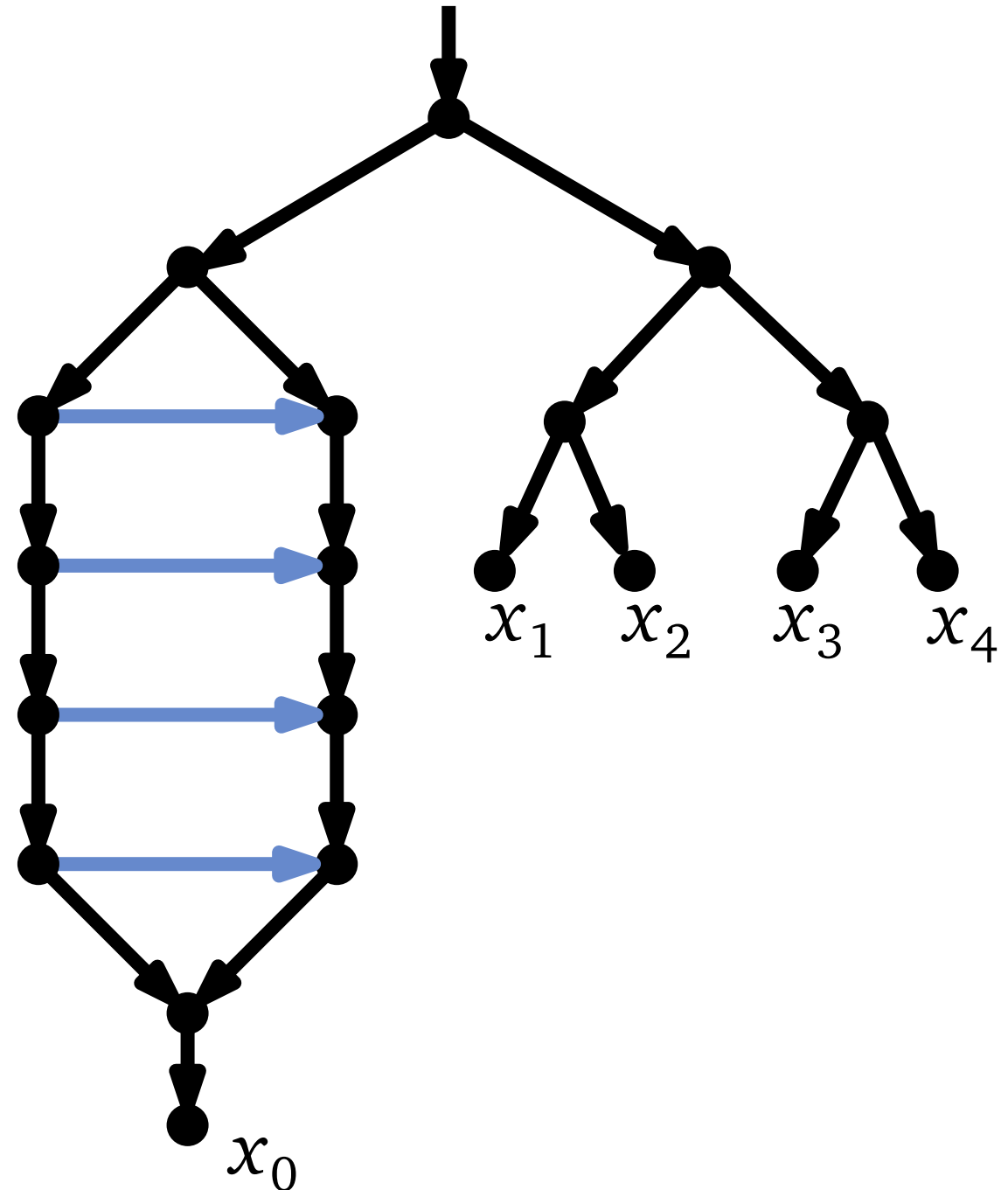
Definition. We call a rooted network *flip-friendly* if it can be turned into any other rooted network with the same underlying unrooted network by a sequence of arc-flips.

Lemma. For every X and r there exists a flip-friendly network on X with r reticulations.

New idea

Definition. We call a rooted network *flip-friendly* if it can be turned into any other rooted network with the same underlying unrooted network by a sequence of arc-flips.

Lemma. For every X and r there exists a flip-friendly network on X with r reticulations.



New proof idea

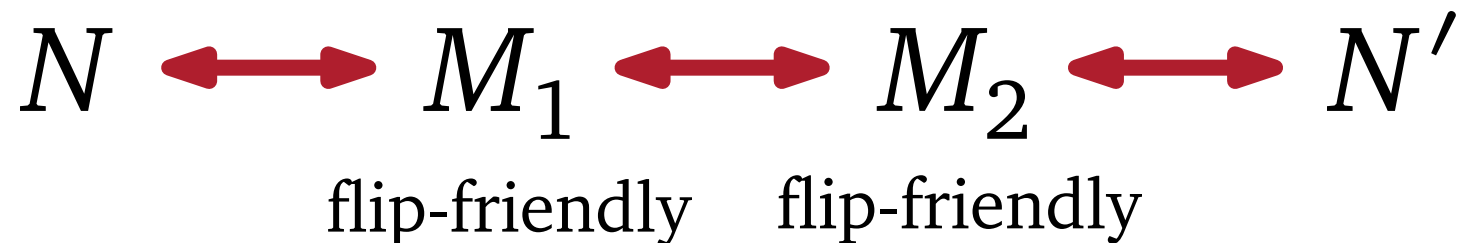
Given two rooted networks N, N' on X with r reticulations, we try to do the following:

1. let M be some *flip-friendly* network on X with r reticulations;
2. turn N into a network M_1 with the *same underlying unrooted network as M* using rNNI moves;
3. turn N' into a network M_2 with the same underlying unrooted network as M using rNNI moves;
4. turn M_1 into M_2 using rNNI moves.
5. turn M_2 into N' using rNNI moves (the reverses of the moves in 3.).

New proof idea

Given two rooted networks N, N' on X with r reticulations, we try to do the following:

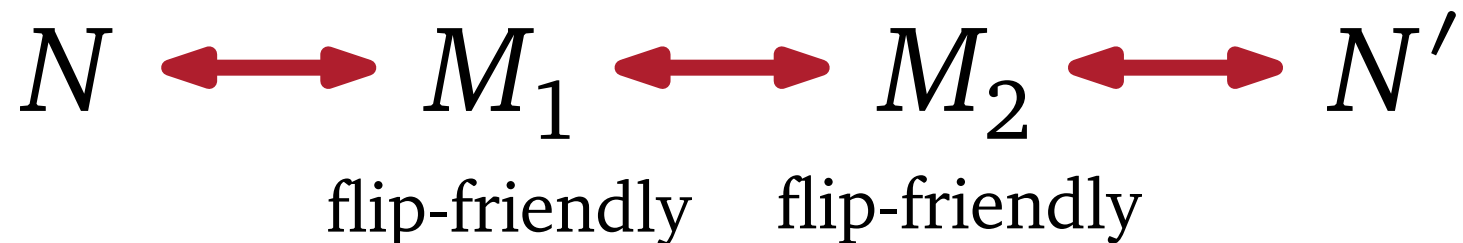
1. let M be some *flip-friendly* network on X with r reticulations;
2. turn N into a network M_1 with the *same underlying unrooted network as M* using rNNI moves;
3. turn N' into a network M_2 with the same underlying unrooted network as M using rNNI moves;
4. turn M_1 into M_2 using rNNI moves.
5. turn M_2 into N' using rNNI moves (the reverses of the moves in 3.).



New proof idea

Given two rooted networks N, N' on X with r reticulations, we try to do the following:

1. let M be some *flip-friendly* network on X with r reticulations;
2. turn N into a network M_1 with the *same underlying unrooted network as M* using rNNI moves;
3. turn N' into a network M_2 with the same underlying unrooted network as M using rNNI moves;
4. turn M_1 into M_2 using rNNI moves.
5. turn M_2 into N' using rNNI moves (the reverses of the moves in 3.).



To prove: we can turn N into a network with the same underlying unrooted network as M .

Changing the underlying unrooted network

Recall that we have the Theorem of Huber et al.:

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Changing the underlying unrooted network

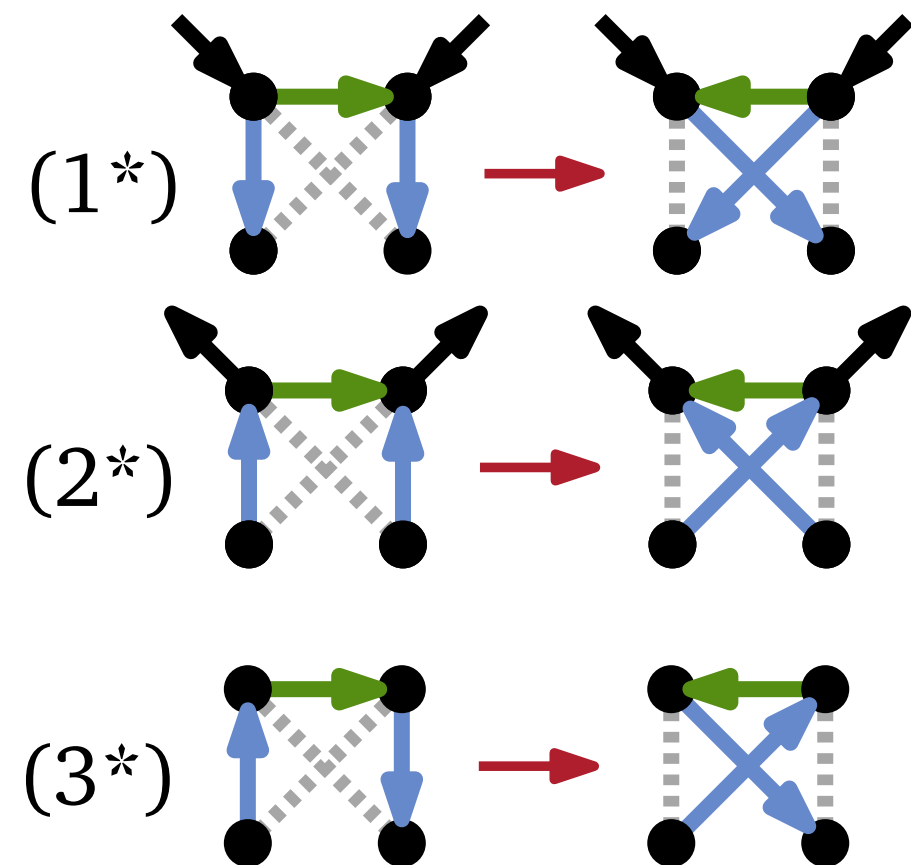
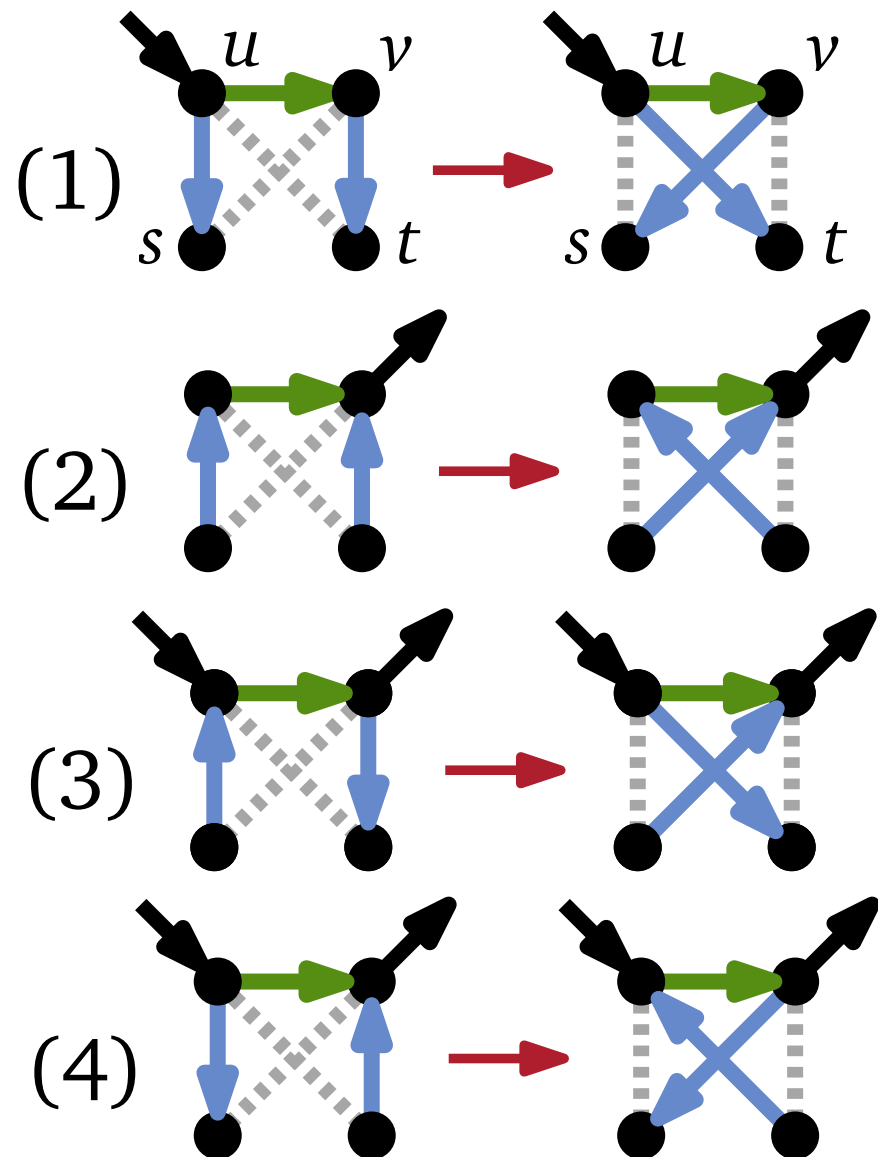
Recall that we have the Theorem of Huber et al.:

Theorem (Huber et al. 2016). For any two *unrooted* networks N, N' on X with the same number of vertices, there exists a sequence of NNI moves turning N into N' .

Hence, it remains to show that we can do an unrooted *NNI* move (on the underlying unrooted network) by doing a *sequence of rNNI* moves.

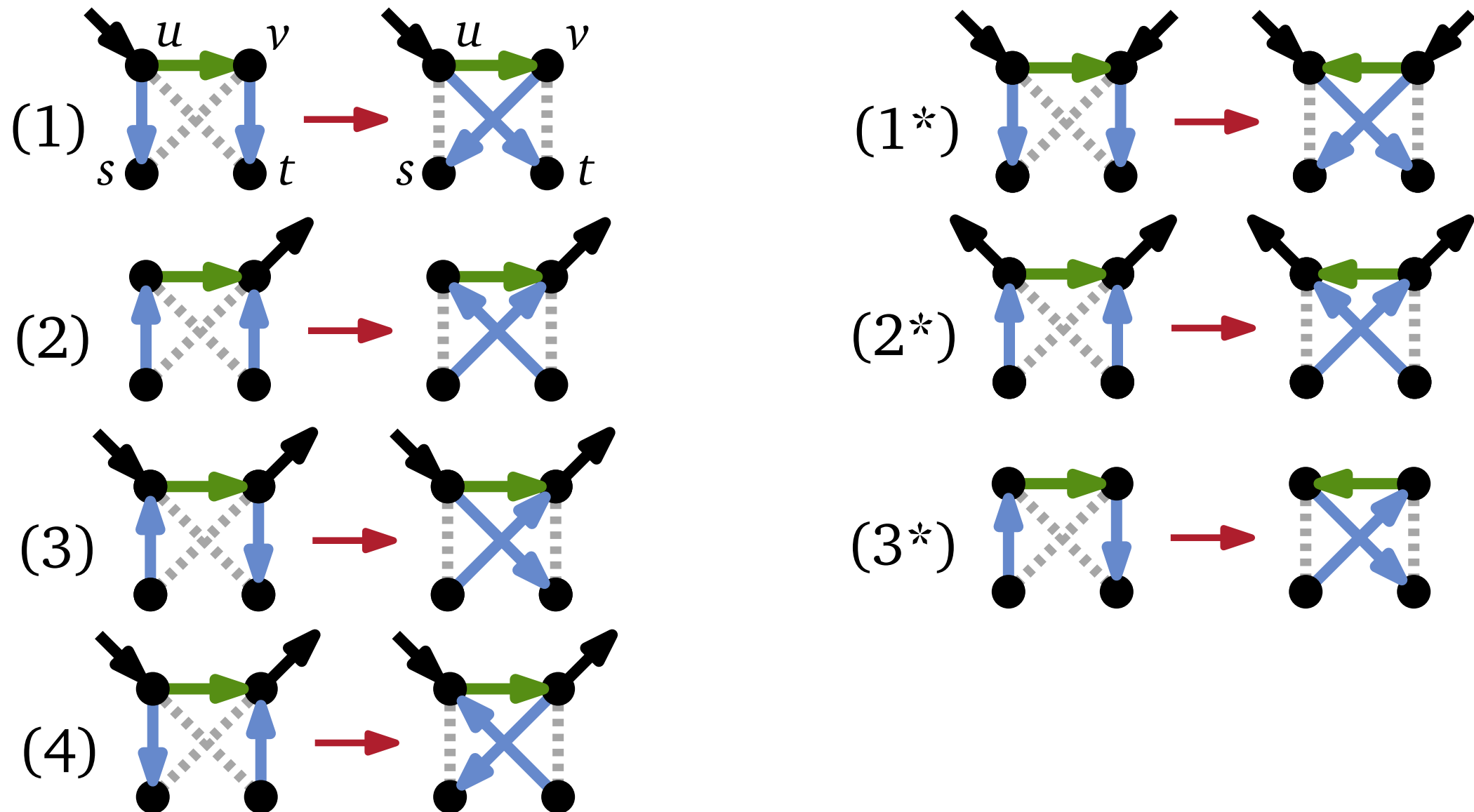
Mimicking NNI moves by rNNI moves

What is the problem? Each unrooted NNI move has a rooted counterpart:



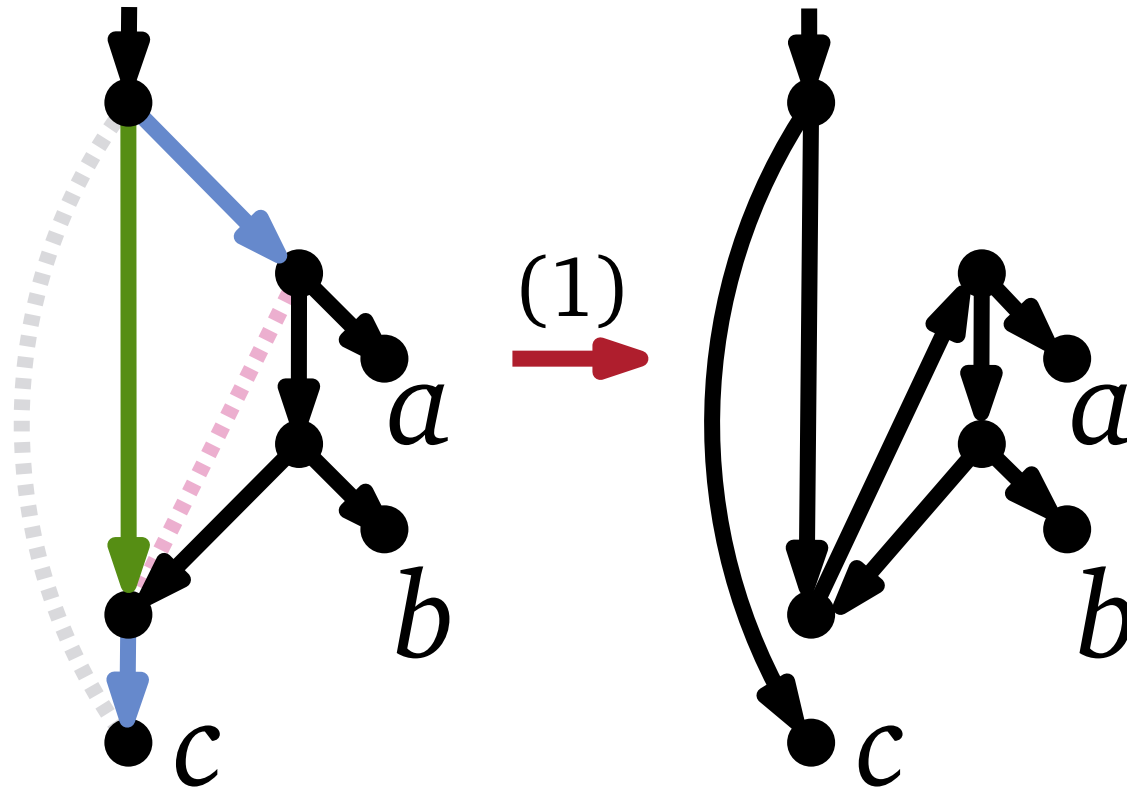
Mimicking NNI moves by rNNI moves

What is the problem? Each unrooted NNI move has a rooted counterpart:



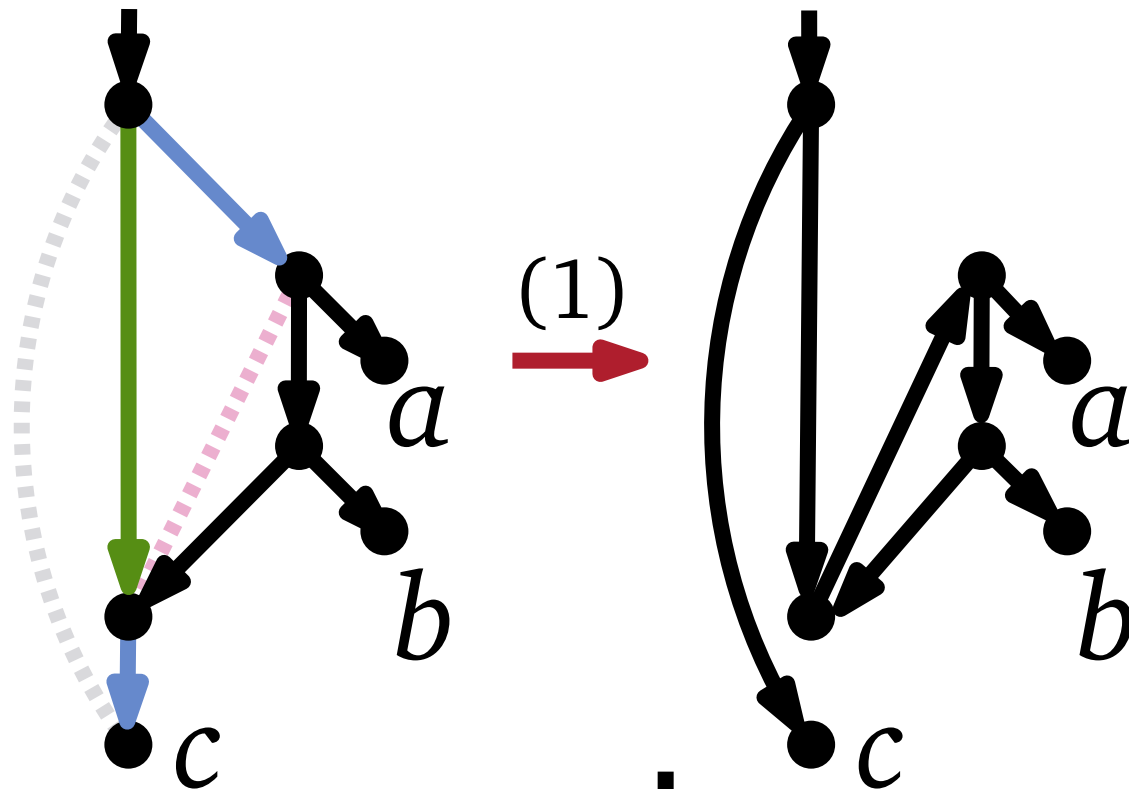
Problem: What if this move creates a *directed cycle*?

Avoiding directed cycles: example

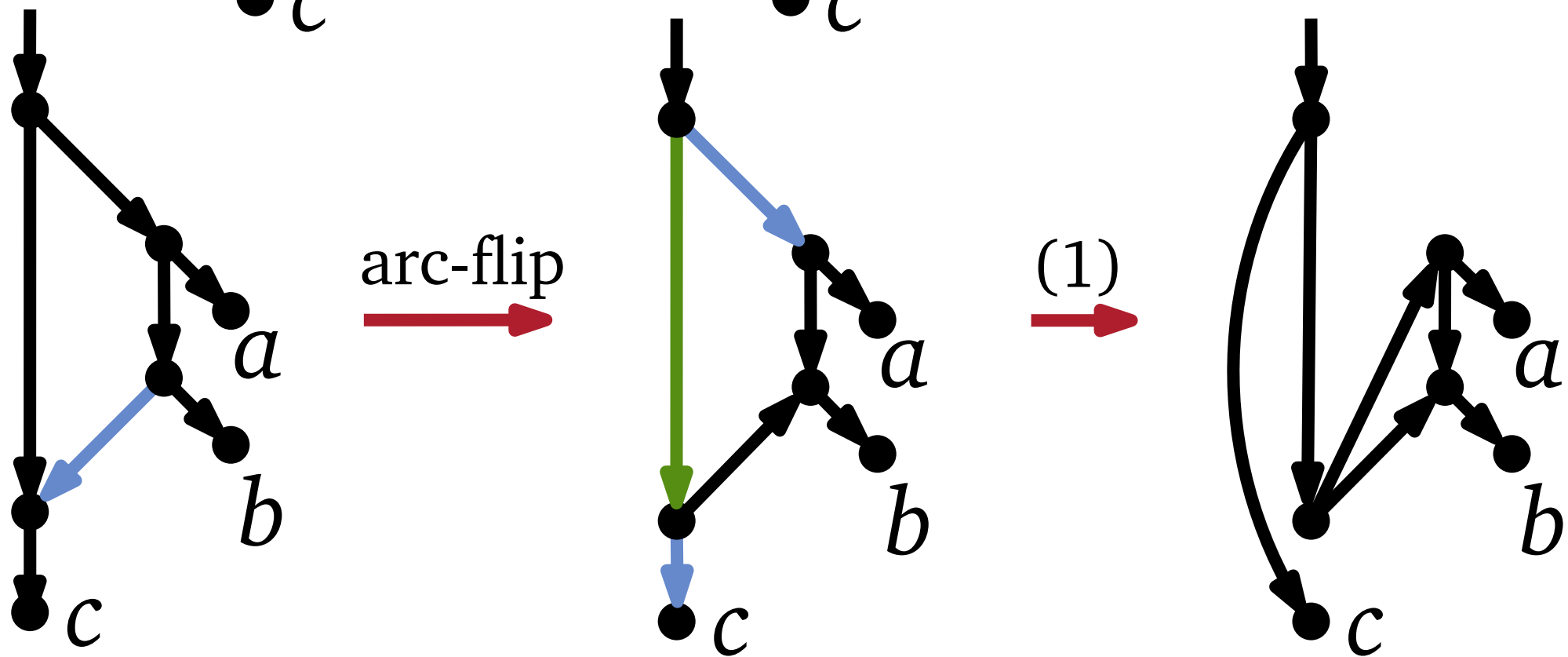


Avoiding directed cycles: example

Instead of

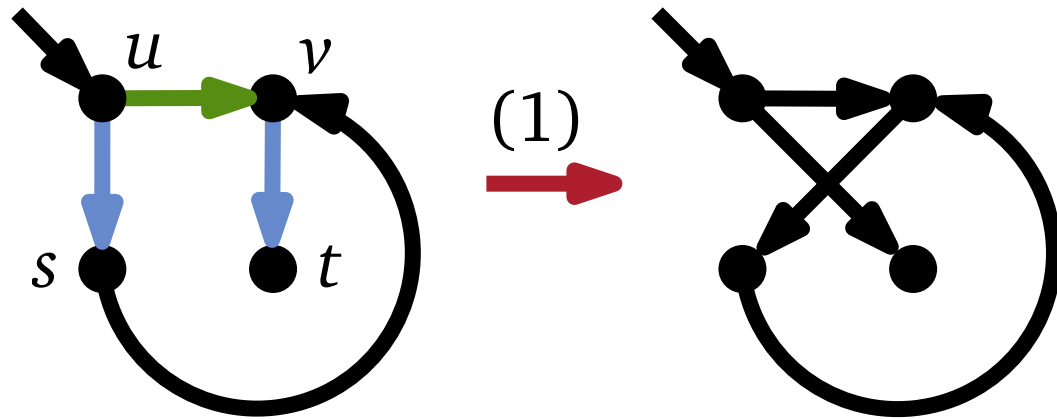


we can do



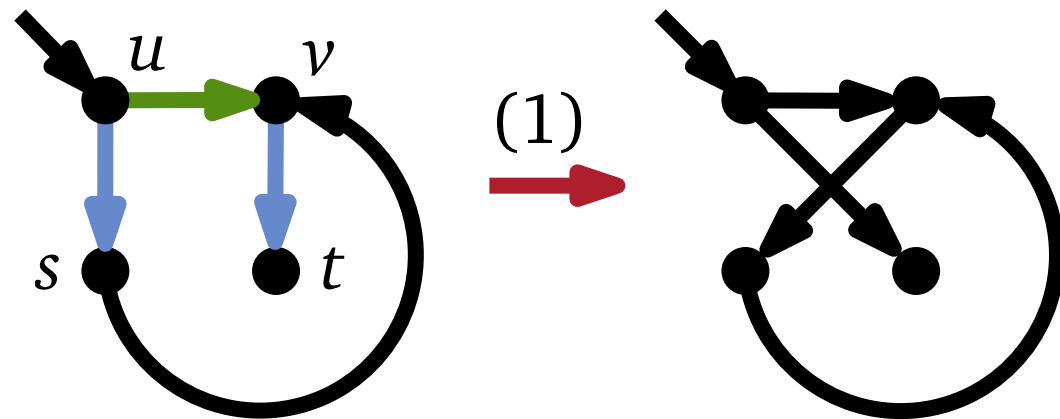
Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

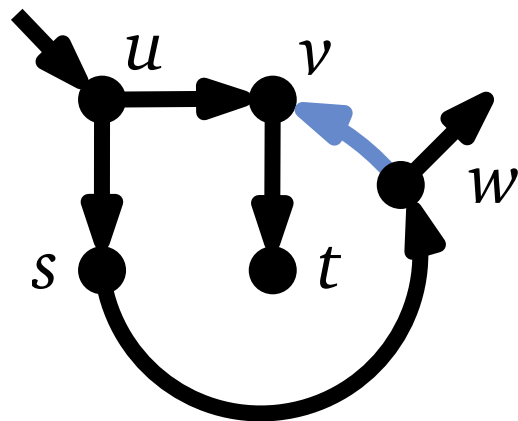


Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

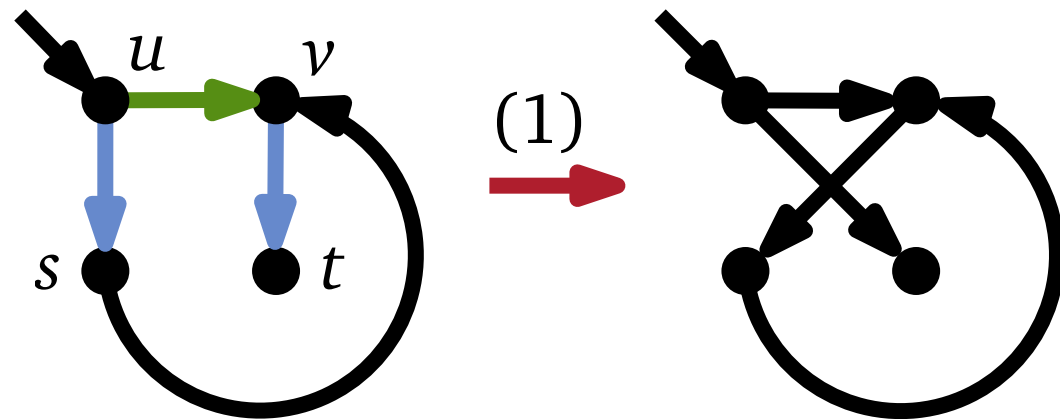


Instead of move (1), we do the following:

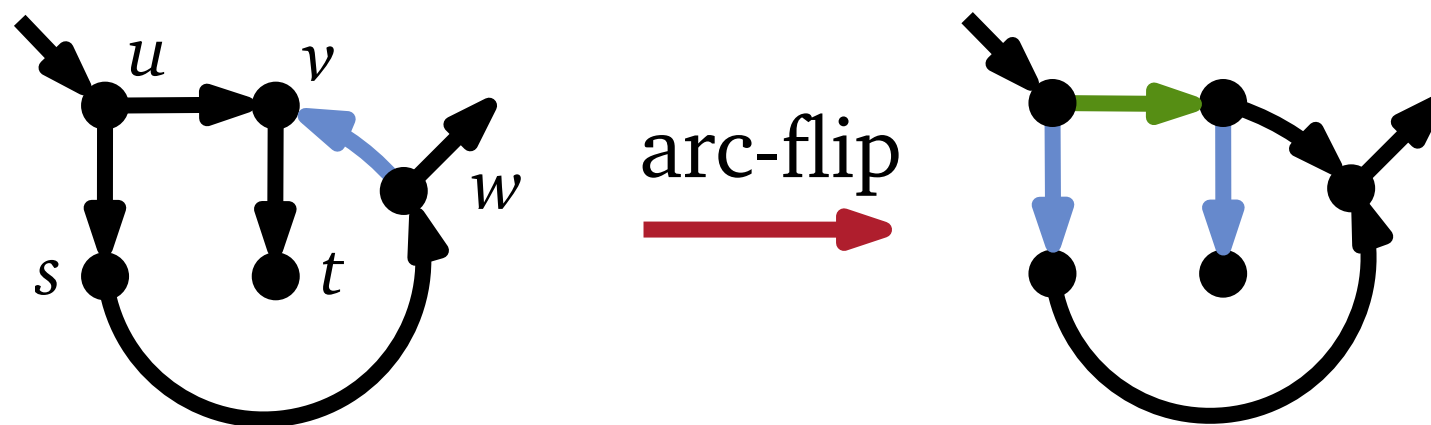


Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

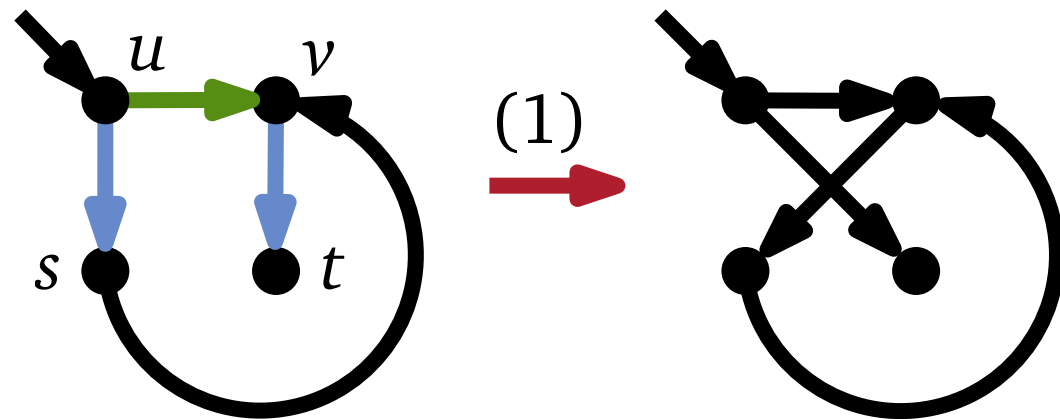


Instead of move (1), we do the following:

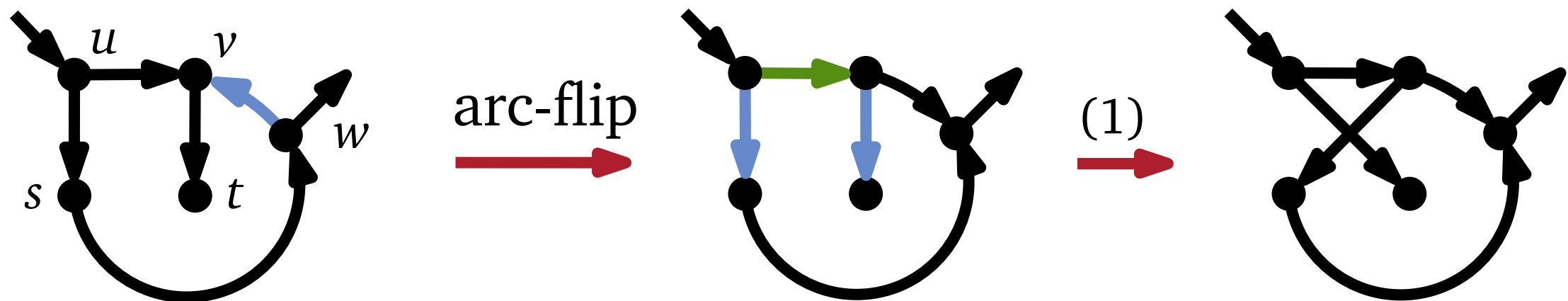


Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

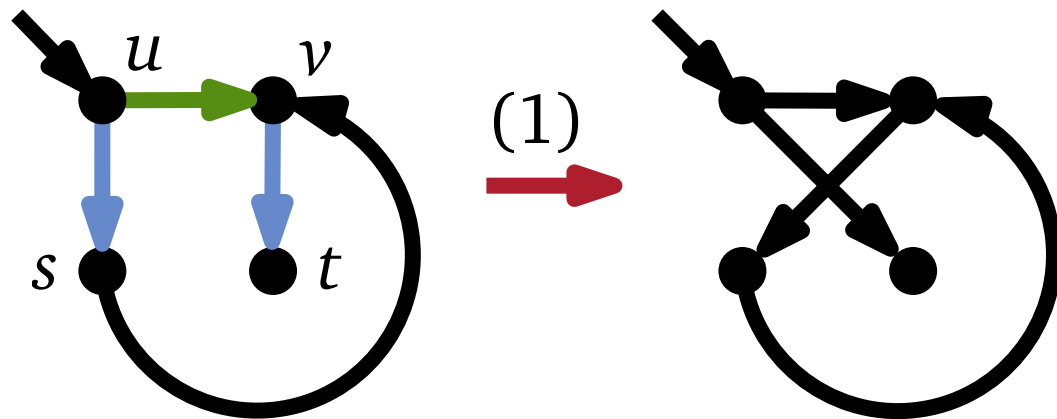


Instead of move (1), we do the following:

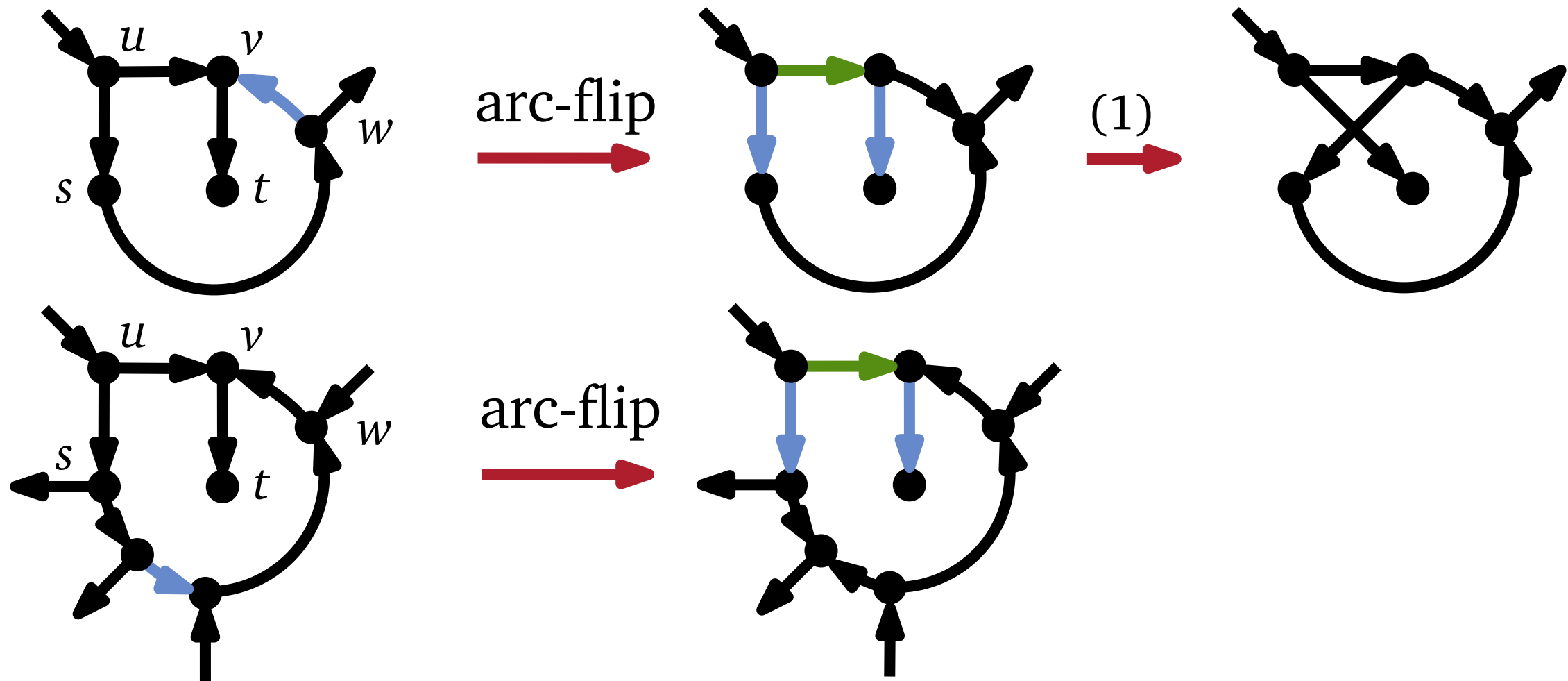


Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

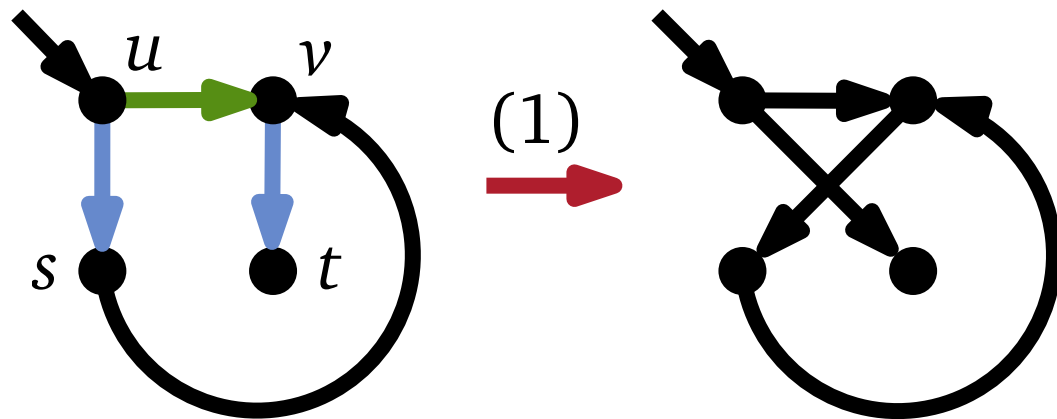


Instead of move (1), we do the following:

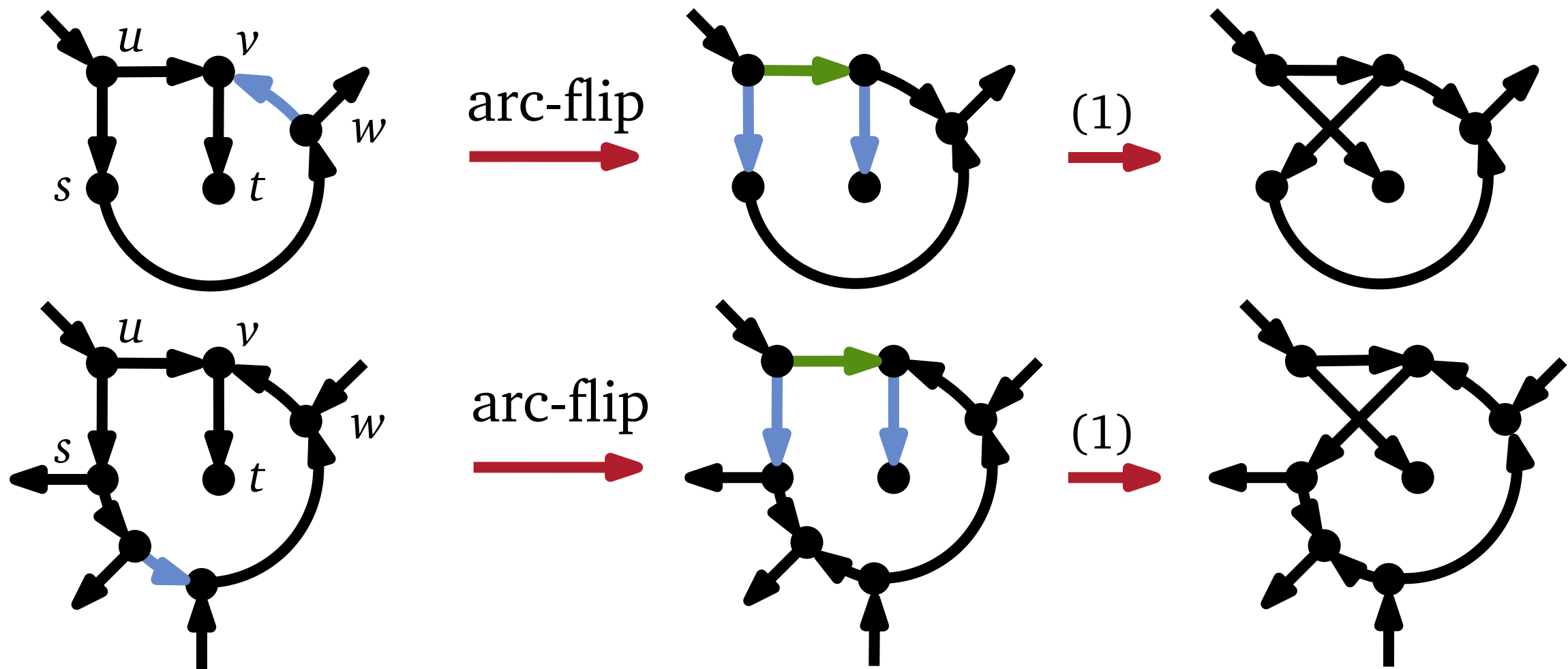


Avoiding directed cycles: more generally (but only case (1))

Case (1): A directed cycle can appear as follows:

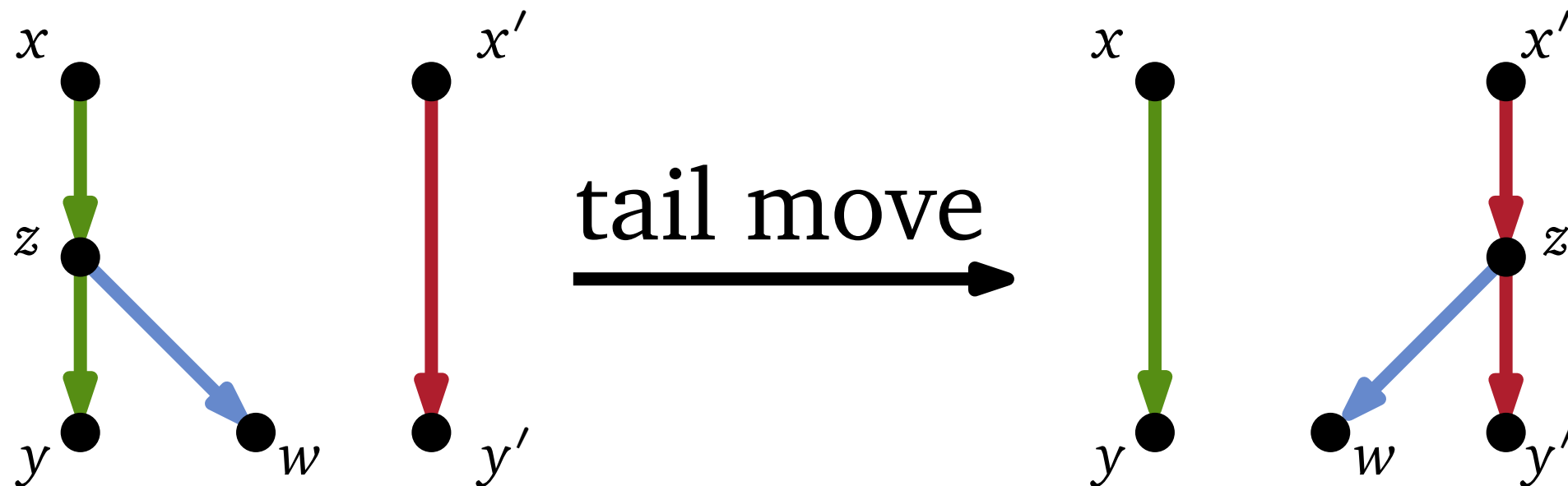


Instead of move (1), we do the following:



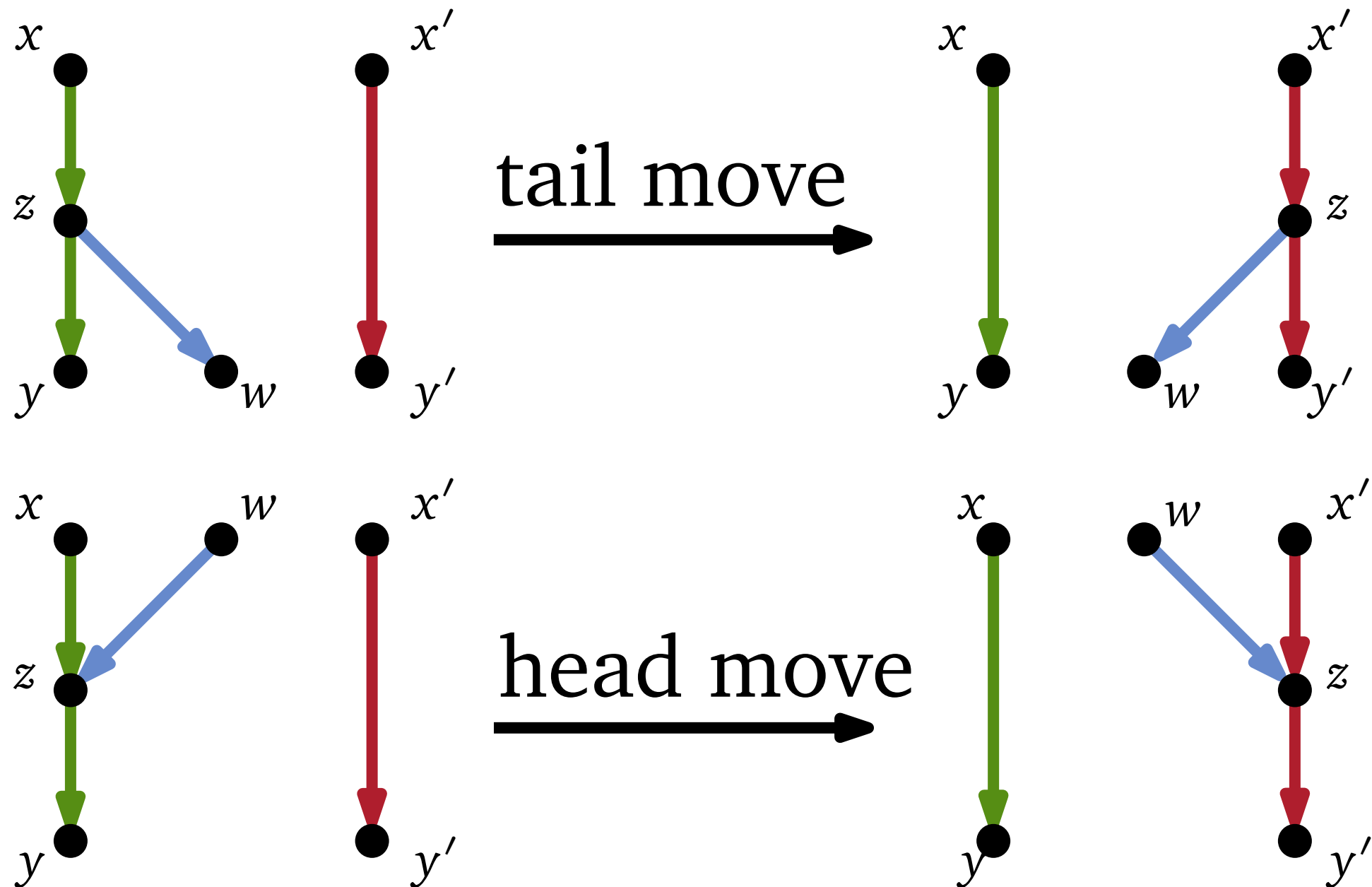
Head and tail moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



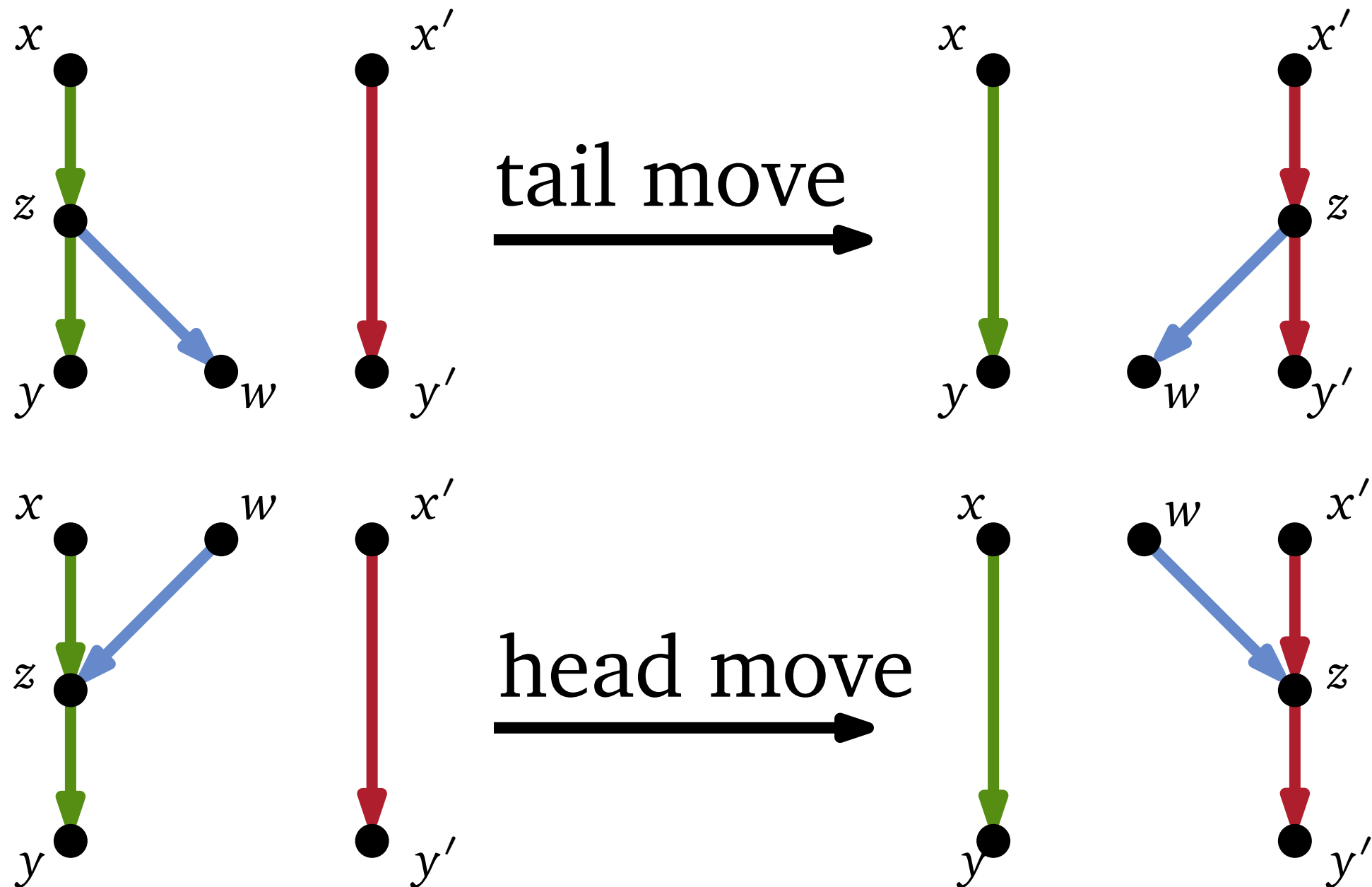
Head and tail moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



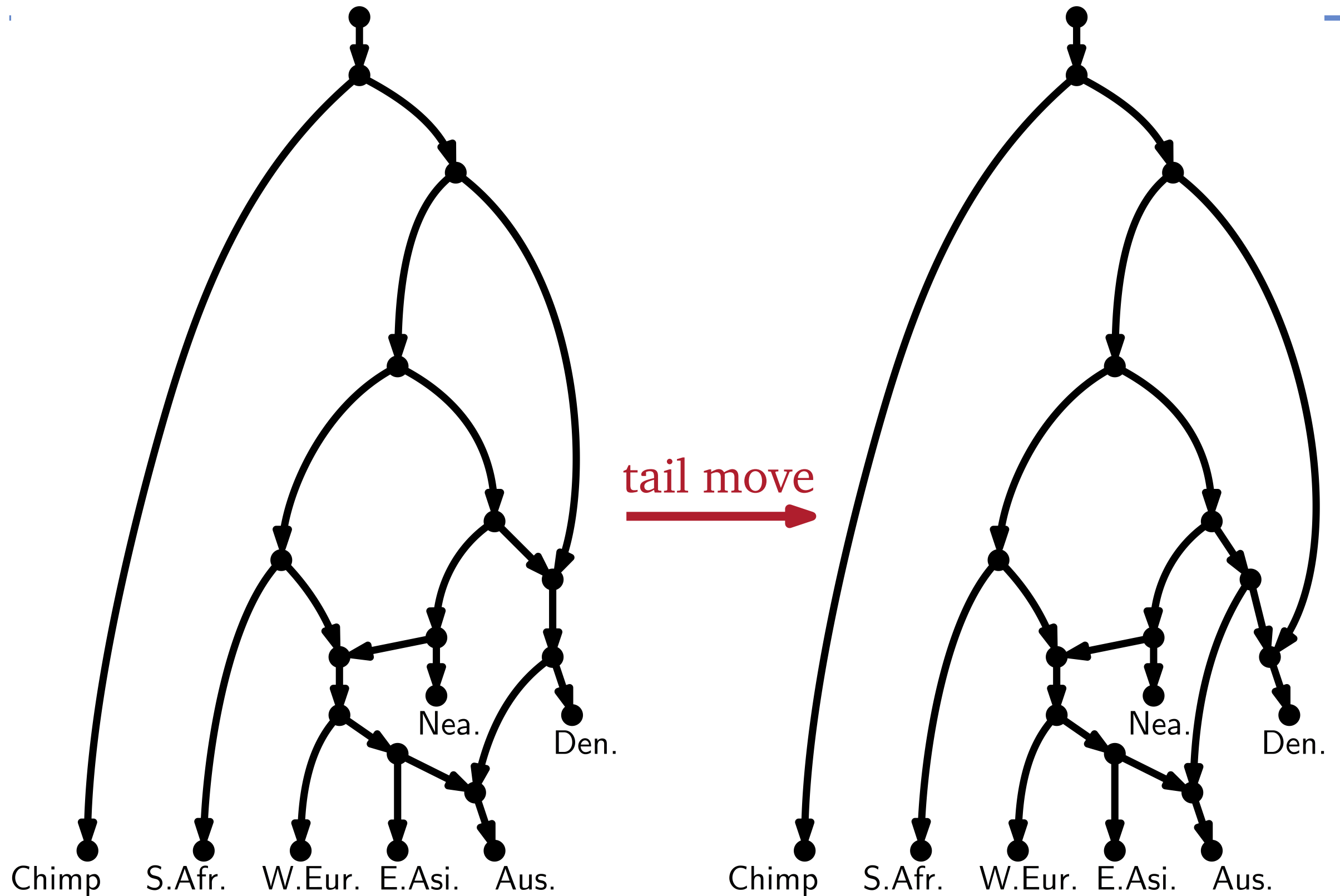
Head and tail moves

An *rSPR* move on a rooted network consists of relocating the head or tail of an arc.



Can we use only tail moves?

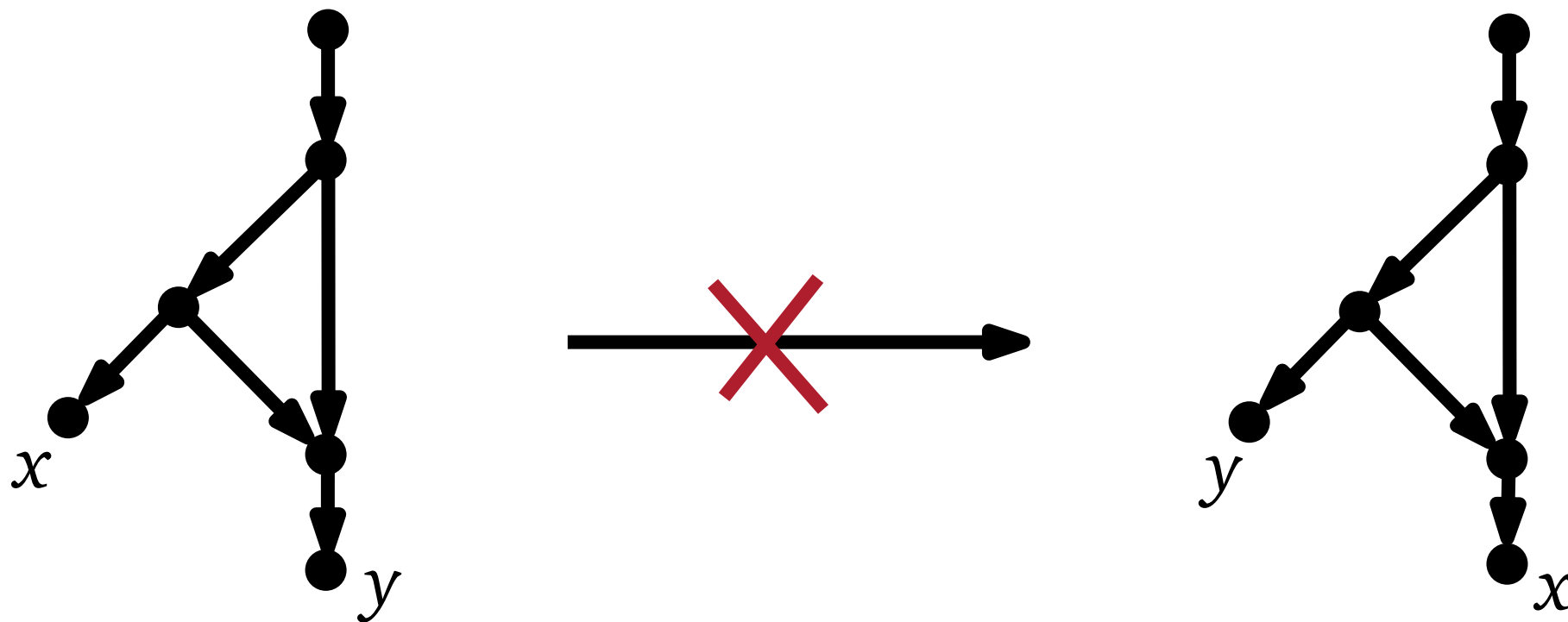
Tail moves



Mallick, Li, Lipson, Mathieson, Gymrek, Racimo, et al. *Nature* (2016).

Very recent results on *tail moves*

Theorem. For any two rooted networks N, N' on X with k reticulations, there exists a sequence of *tail moves* turning N into N' , *unless* $k = 1$ and $|X| = 2$.

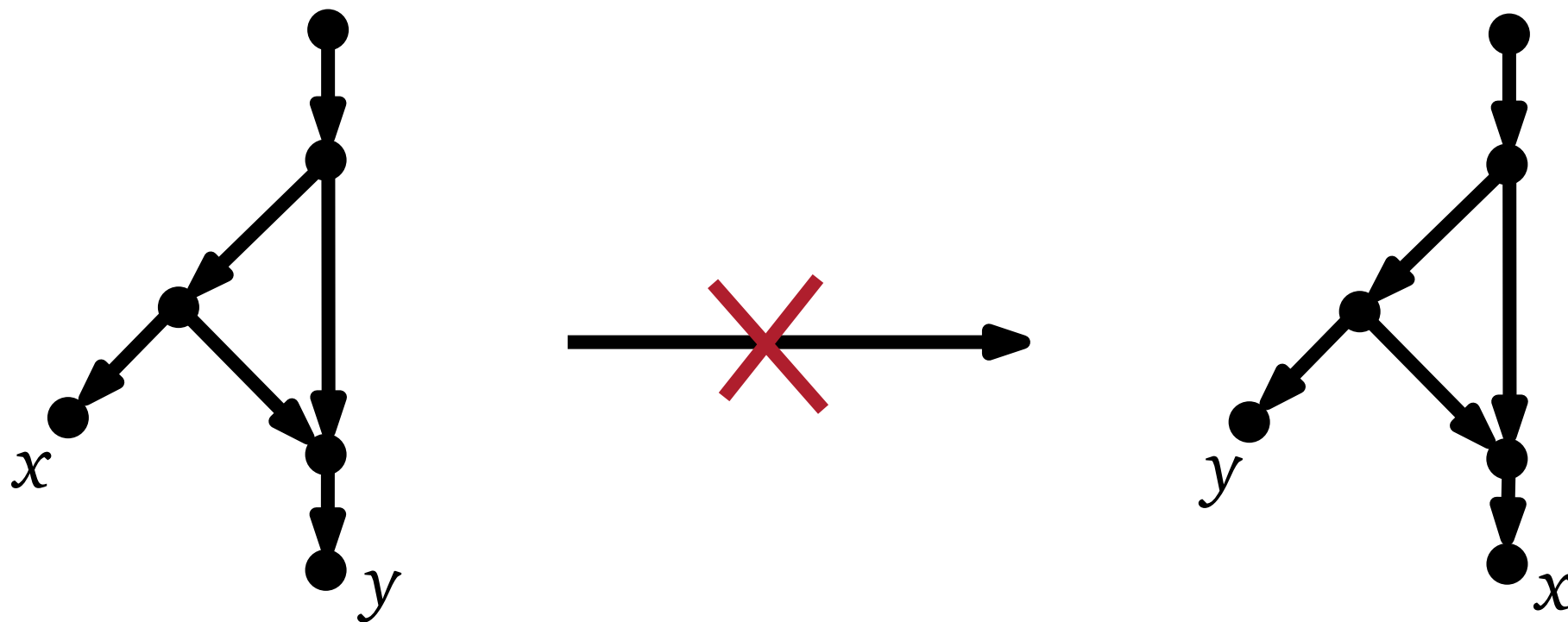


Remie Janssen, Mark Jones, Péter Erdős, Leo van Iersel and Celine Scornavacca.

Exploring the tiers of rooted phylogenetic network space using tail moves

Very recent results on *tail moves*

Theorem. For any two rooted networks N, N' on X with k reticulations, there exists a sequence of **at most** $4(|V| - |X|)$ *tail moves* turning N into N' , **unless** $k = 1$ and $|X| = 2$.



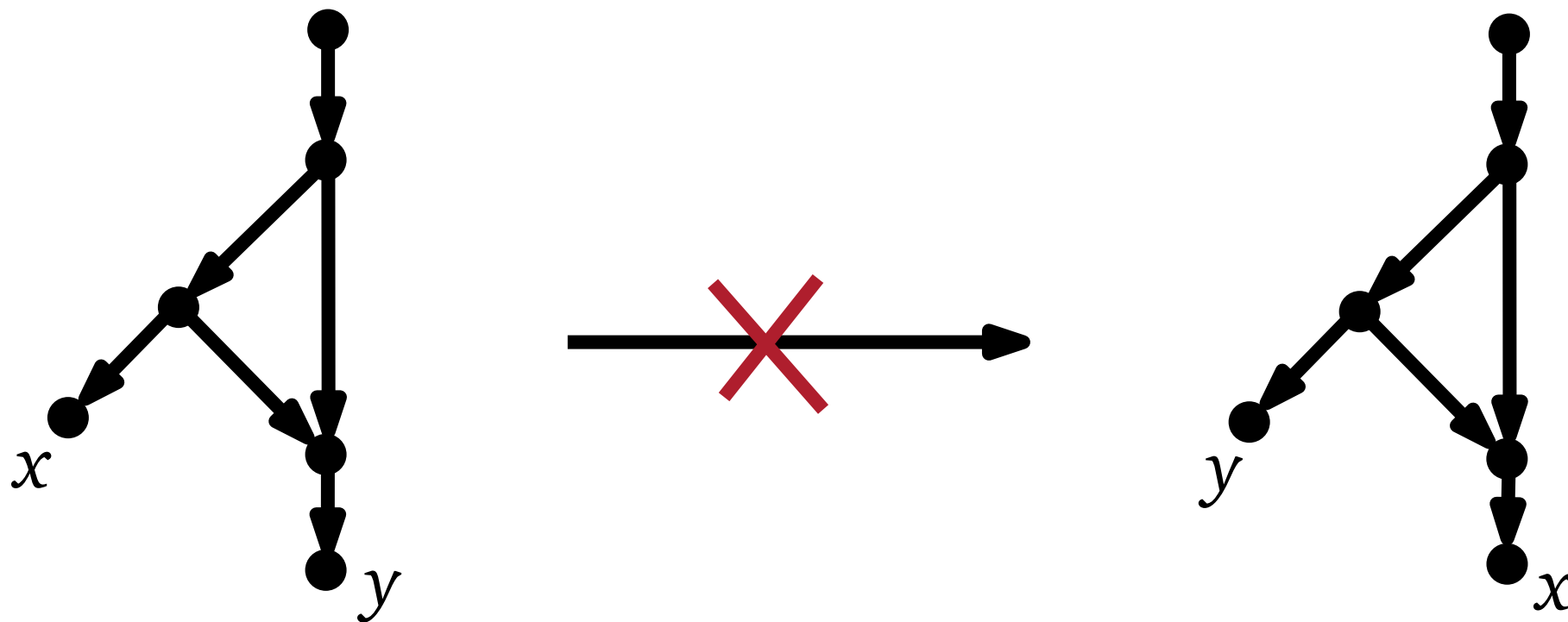
Remie Janssen, Mark Jones, Péter Erdős, Leo van Iersel and Celine Scornavacca.

Exploring the tiers of rooted phylogenetic network space using tail moves

Very recent results on *tail moves*

Theorem. For any two rooted networks N, N' on X with k reticulations, there exists a sequence of **at most $4(|V| - |X|)$ tail moves** turning N into N' , **unless** $k = 1$ and $|X| = 2$.

Theorem. For any two rooted networks N, N' on X with k reticulations, there exists a sequence of **at most $4(|V| - |X|)^2$ distance-1 tail moves** turning N into N' , unless $k = 1$ and $|X| = 2$.



Remie Janssen, Mark Jones, Péter Erdős, Leo van Iersel and Celine Scornavacca.

Exploring the tiers of rooted phylogenetic network space using tail moves

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
- *rSPR* moves are *head moves* and *tail moves*.

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
 - *rSPR* moves are *head moves* and *tail moves*.
 - *rNNI* moves are *local head moves* and *local tail moves*.
-

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
 - *rSPR* moves are *head moves* and *tail moves*.
 - *rNNI* moves are *local head moves* and *local tail moves*.
 - Each *tier* of phylogenetic network space is *connected* by *rNNI* moves (and hence by rSPR moves)
-

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
 - *rSPR* moves are *head moves* and *tail moves*.
 - *rNNI* moves are *local head moves* and *local tail moves*.
 - Each *tier* of phylogenetic network space is *connected* by *rNNI* moves (and hence by rSPR moves)
 - The *space* of phylogenetic networks is *connected by rNNI* (or rSPR) moves combined with reasonable *vertical moves*.
-

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
 - *rSPR* moves are *head moves* and *tail moves*.
 - *rNNI* moves are *local head moves* and *local tail moves*.
 - Each *tier* of phylogenetic network space is *connected* by *rNNI* moves (and hence by rSPR moves)
 - The *space* of phylogenetic networks is *connected by rNNI* (or rSPR) moves combined with reasonable *vertical moves*.
 - The *tail-move diameter* is at most $4(|V| - |X|)$.
-

Summary

- We studied rearrangement moves on *rooted phylogenetic networks*
 - *rSPR* moves are *head moves* and *tail moves*.
 - *rNNI* moves are *local head moves* and *local tail moves*.
 - Each *tier* of phylogenetic network space is *connected* by *rNNI* moves (and hence by *rSPR* moves)
 - The *space* of phylogenetic networks is *connected* by *rNNI* (or *rSPR*) moves combined with reasonable *vertical moves*.
 - The *tail-move diameter* is at most $4(|V| - |X|)$.
 - The *rNNI diameter* is at most $4(|V| - |X|)^2$.
-

Questions and references

Questions:

- Is the *rNNI diameter* $\theta(|V|^2)$ or $\theta(|V|\log(|V|))$?
- Can the bounds on the *rSPR* and *tail move* diameter be improved?

References:

- Philippe Gambette, Leo van Iersel, Mark Jones, Manuel Lafond, Fabio Pardi, Celine Scornavacca. *Rearrangement Moves on Rooted Phylogenetic Networks*. To appear in *PLOS Computational Biology*.
 - Remie Janssen, Mark Jones, Péter Erdős, Leo van Iersel and Celine Scornavacca. *Exploring the tiers of rooted phylogenetic network space using tail moves*, in preparation.
 - Katharina Huber, Vincent Moulton, Taoyang Wu. *Transforming phylogenetic networks: Moving beyond tree space*. *Journal of Theoretical Biology* (2016).
-