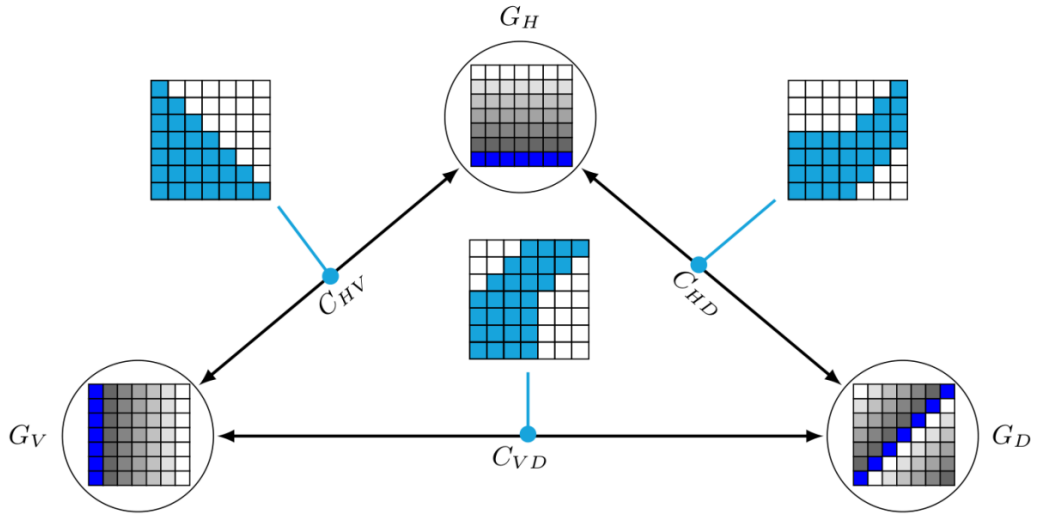


Digging (energy) wells for unknown substances:

Programming substrate-independent energy barriers in catalytic chemical reactions

David Doty

Maynooth University CS Colloquium, May 8th, 2019

Slides credit: David Haley

Joint work with: Keenan Breik, Cameron Chalk, David Haley, David Soloveichik

Acknowledgements

David Haley

Keenan Breik

Cameron Chalk

David Soloveichik

Goal

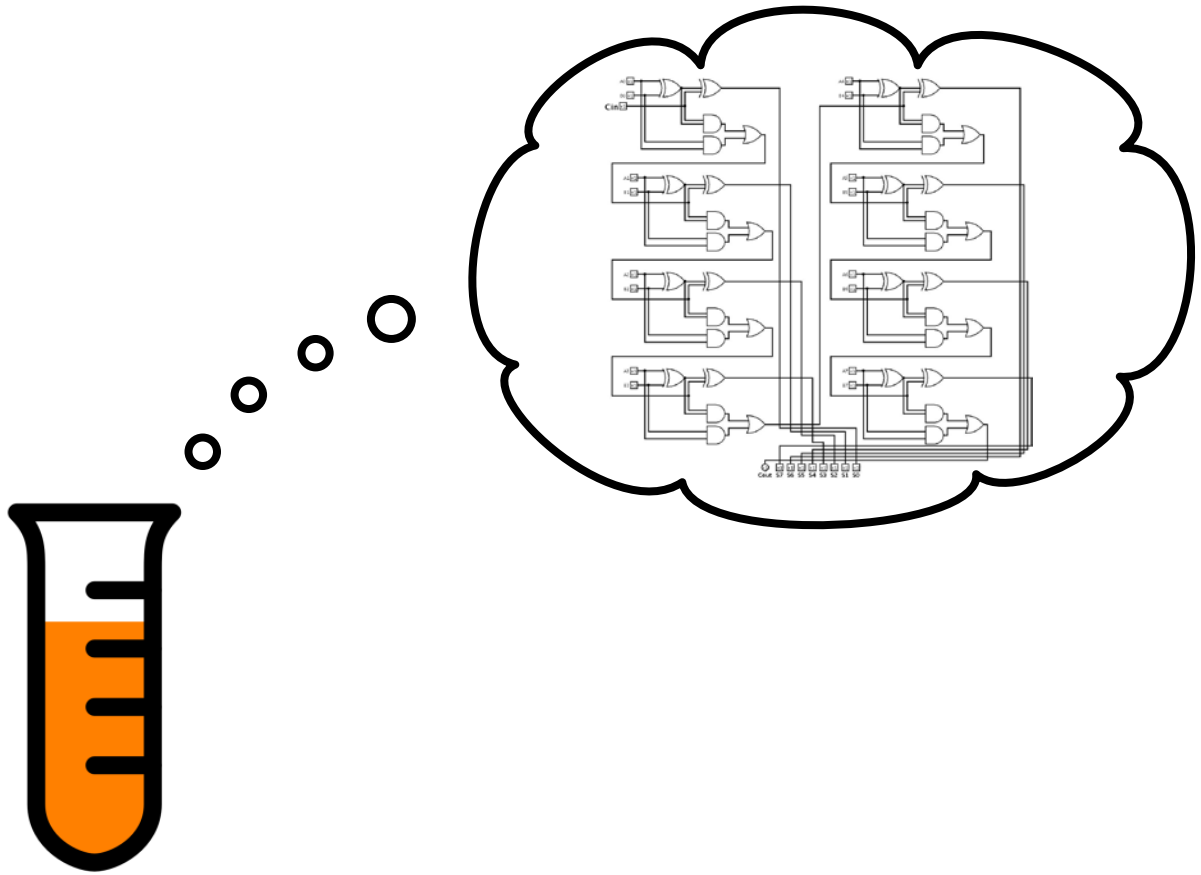
Build molecular computing systems that are reliable

Goal

Build molecular computing systems that are reliable

Goal

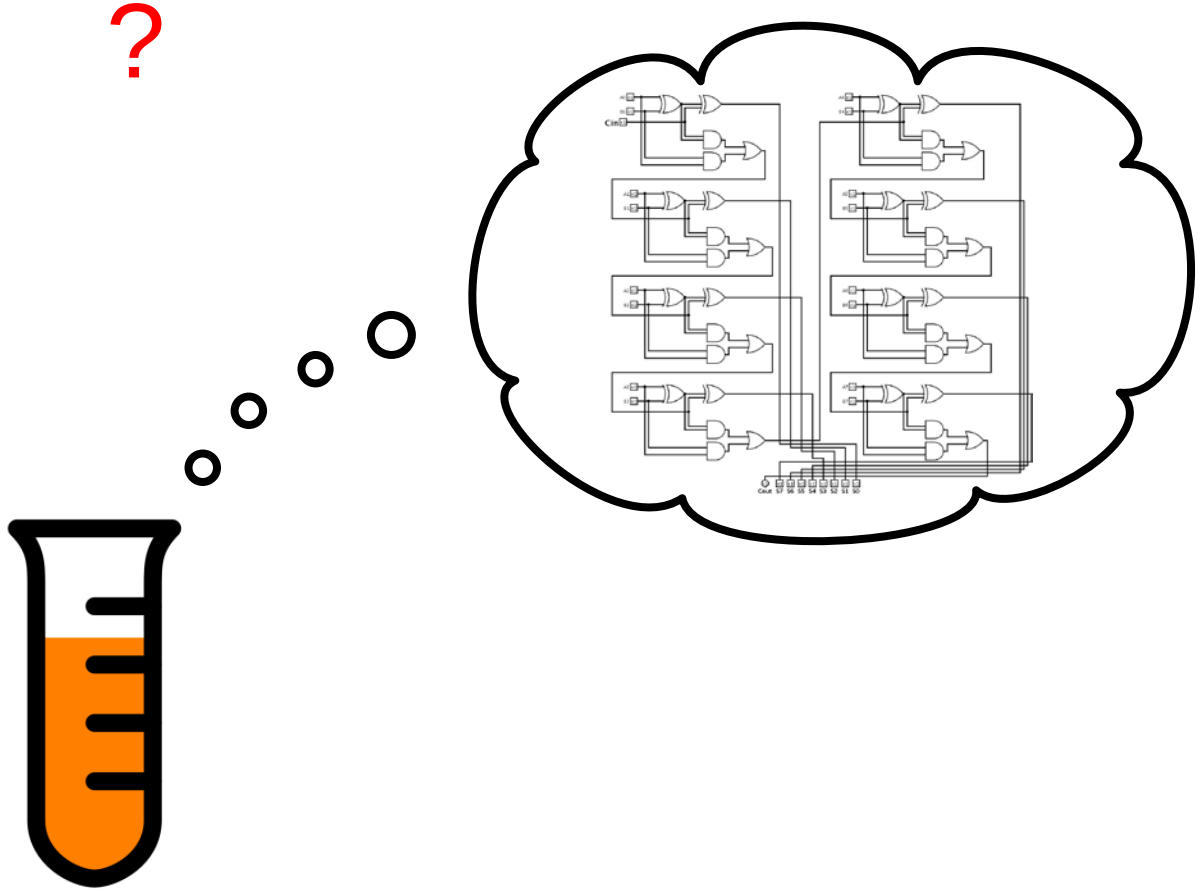
Build molecular computing systems that are reliable



Goal

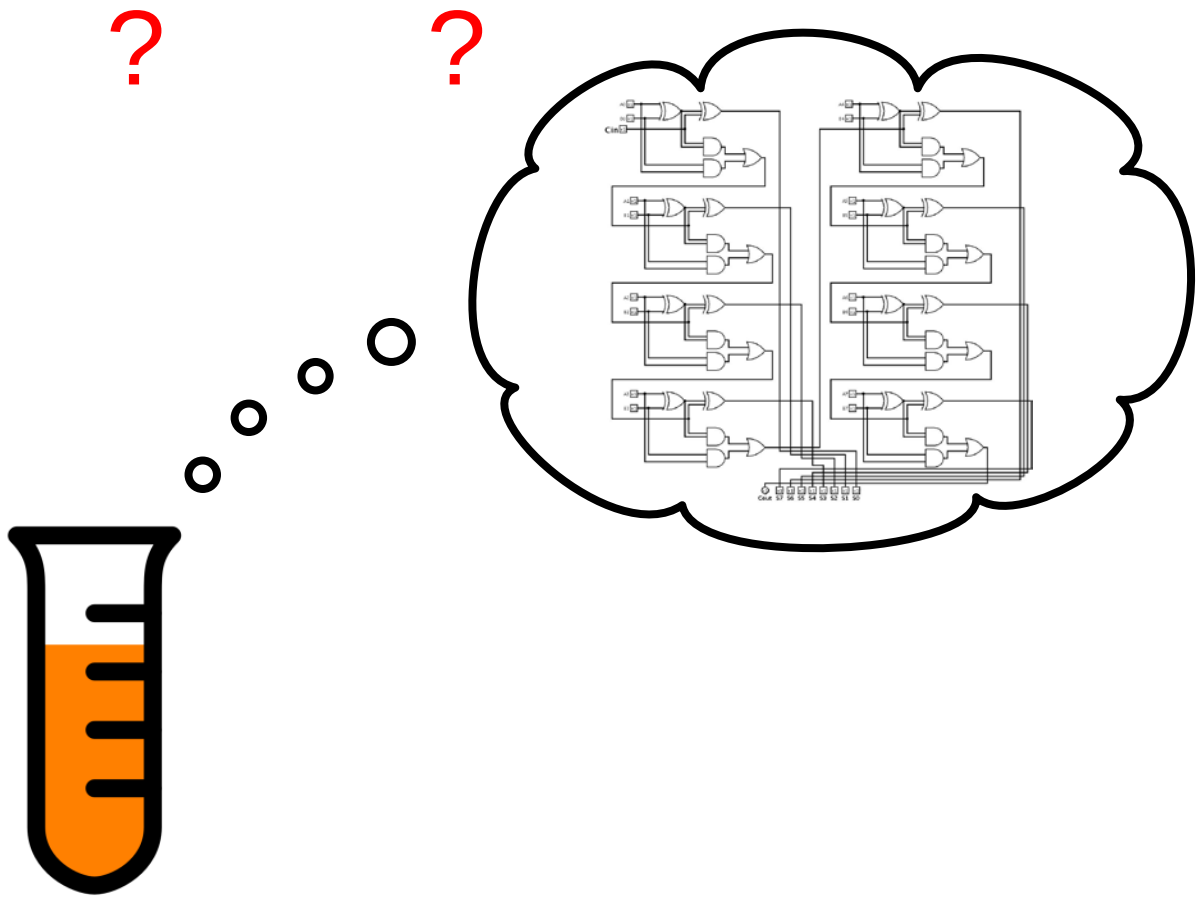
Build molecular computing systems that are reliable

?



Goal

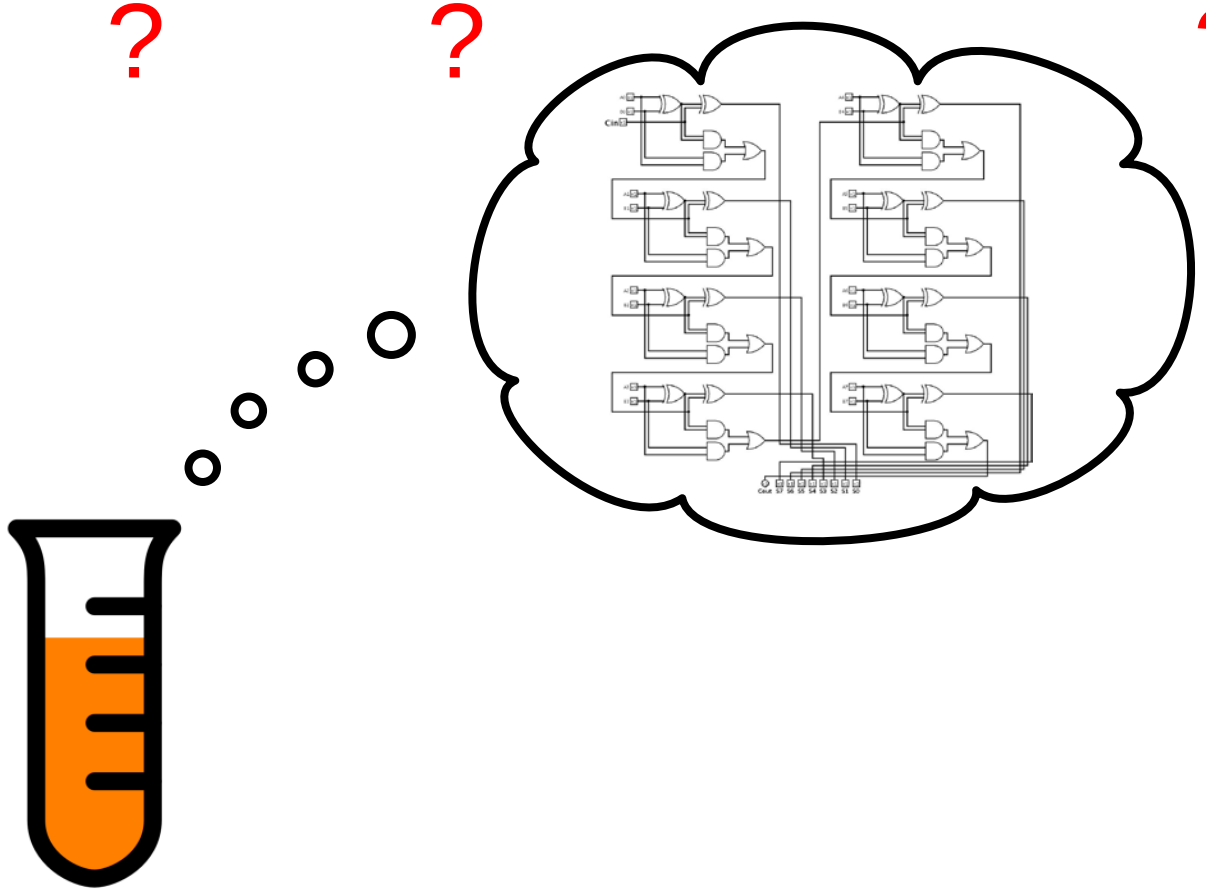
Build molecular computing systems that are reliable



Goal

Build molecular computing systems that are reliable

? ? ?!?

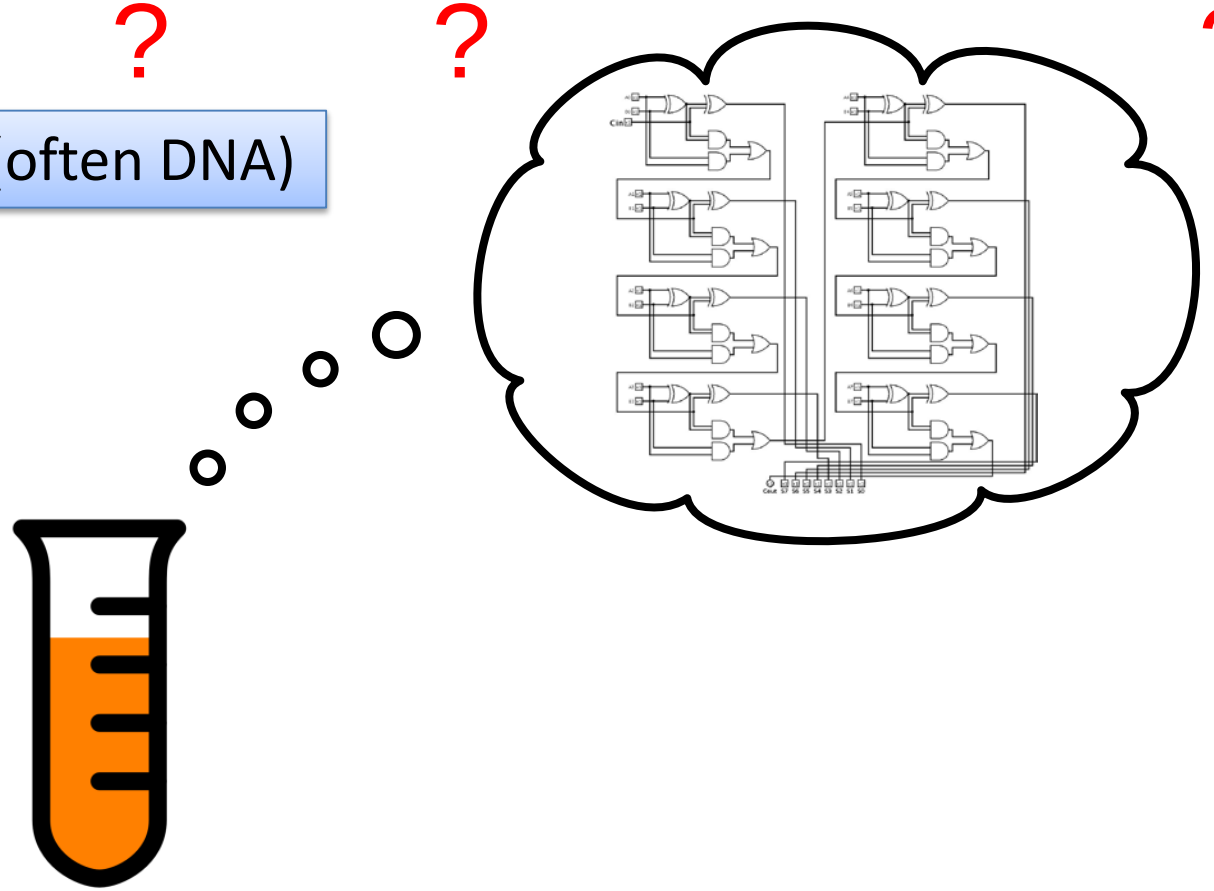


Goal

Build molecular computing systems that are reliable

? ? ?!?

(often DNA)



DNA nanotechnology applications

nonbiological:

- nanoscale resolution surface placement
- X-ray crystallization scaffolding
- molecular motors
- super-resolution imaging
- molecular circuits

biological:

- smart drugs
- mRNA detection
- cell surface marker detection
- genetically encoded structures

DNA nanotechnology applications

nonbiological:

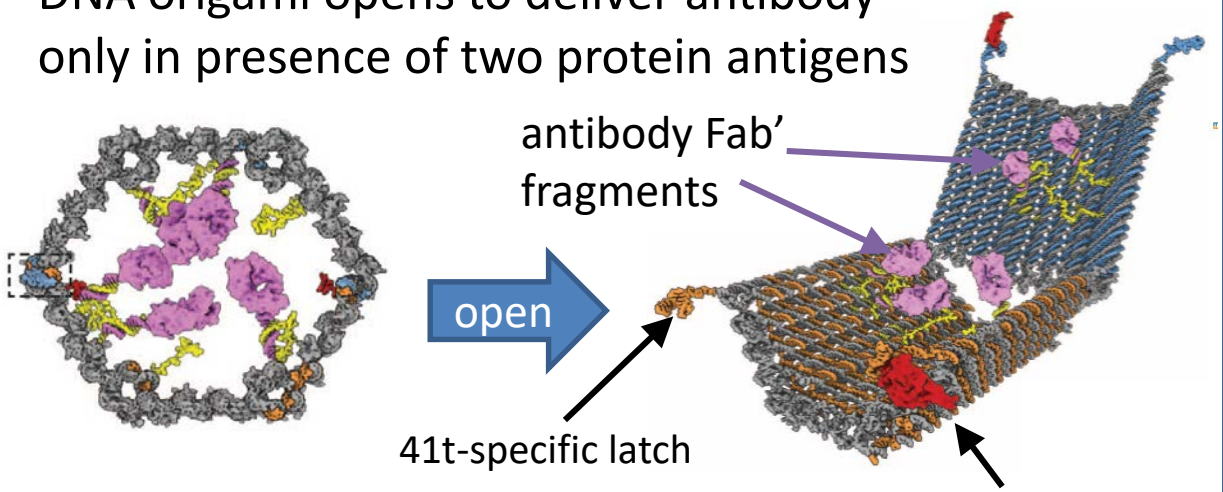
- nanoscale resolution surface placement
- X-ray crystallization scaffolding
- molecular motors
- super-resolution imaging
- molecular circuits

biological:

- smart drugs
- mRNA detection
- cell surface marker detection
- genetically encoded structures

example

DNA origami opens to deliver antibody only in presence of two protein antigens

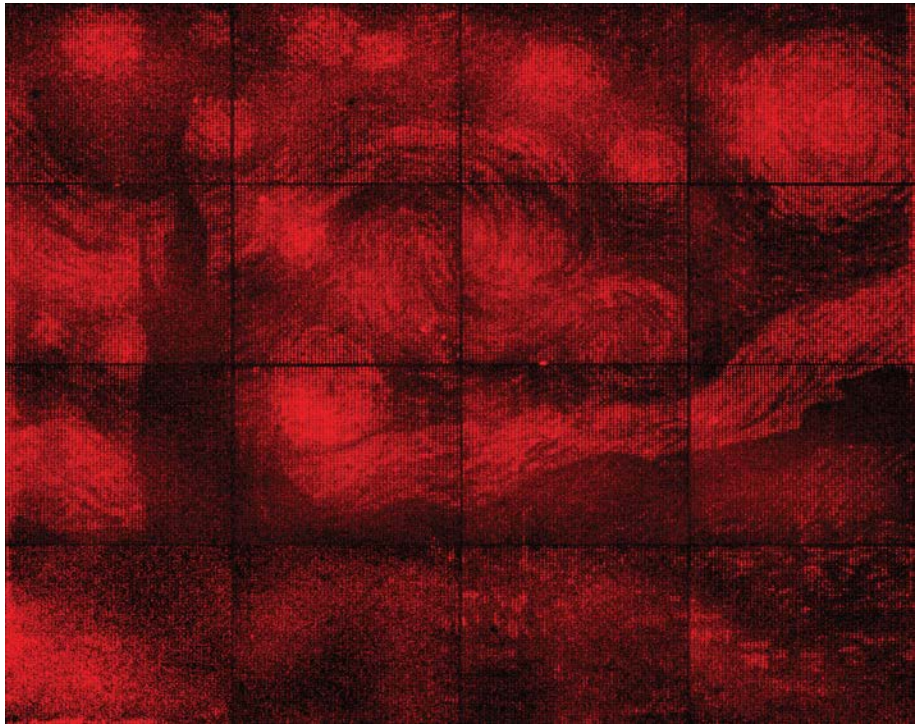


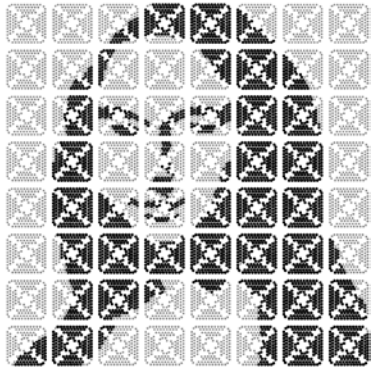
Shawn Douglas, Ido Bachelet, George Church, *A logic-gated nanorobot for targeted transport of molecular payloads*, Science 2012

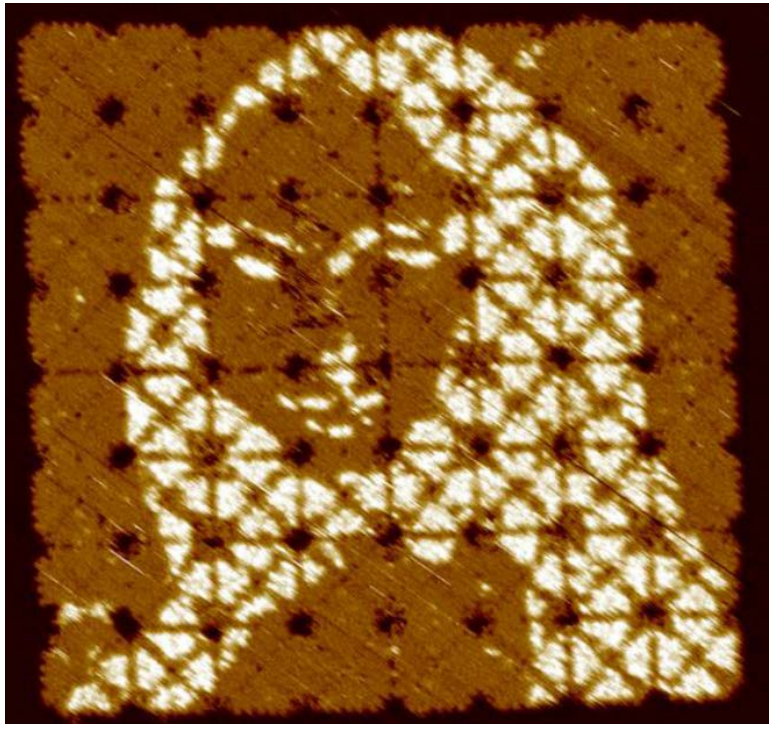
DNA nanotechnology applications

nonbiological:

- **art**





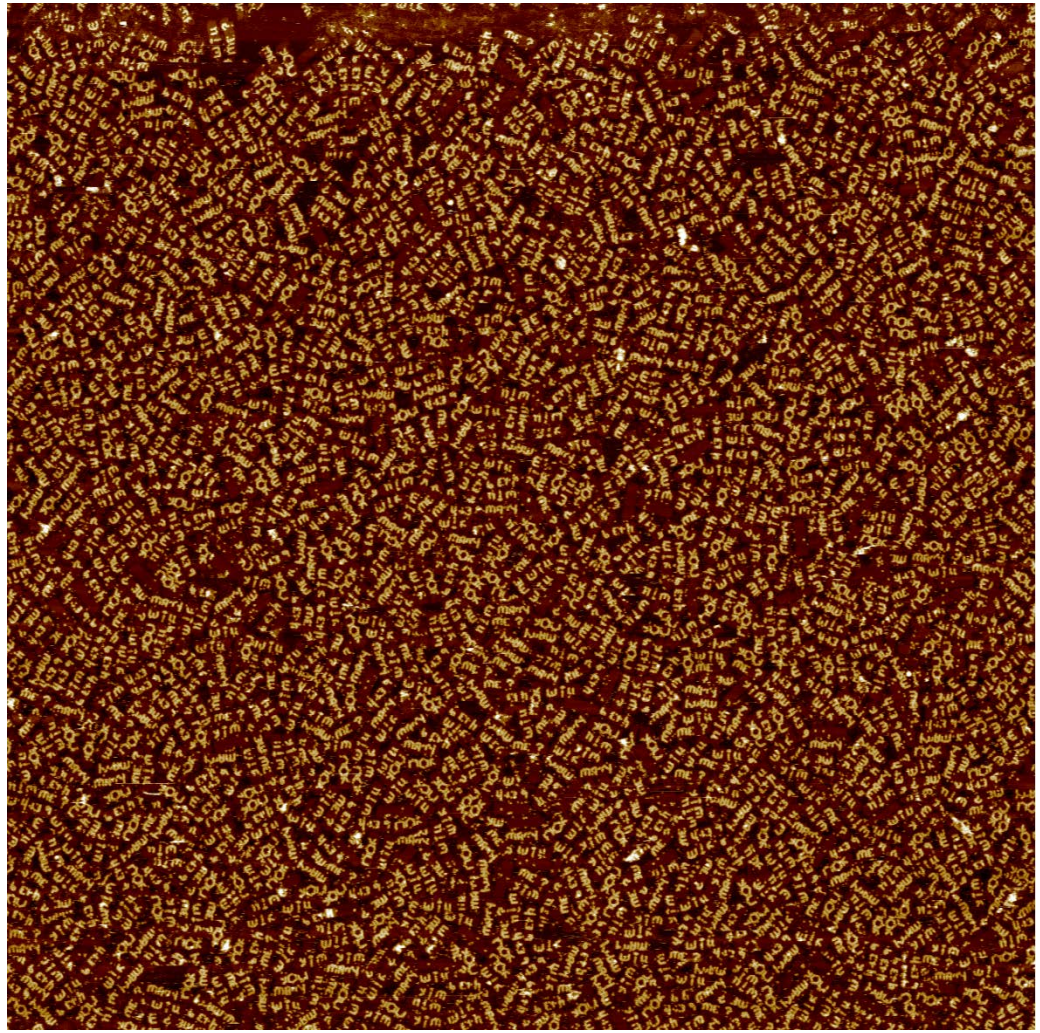


Ashwin Gopinath, Evan Miyazono, Andrei Faraon, Paul Rothemund, *Engineering and mapping nanocavity emission via precision placement of DNA origami*, Nature 2016

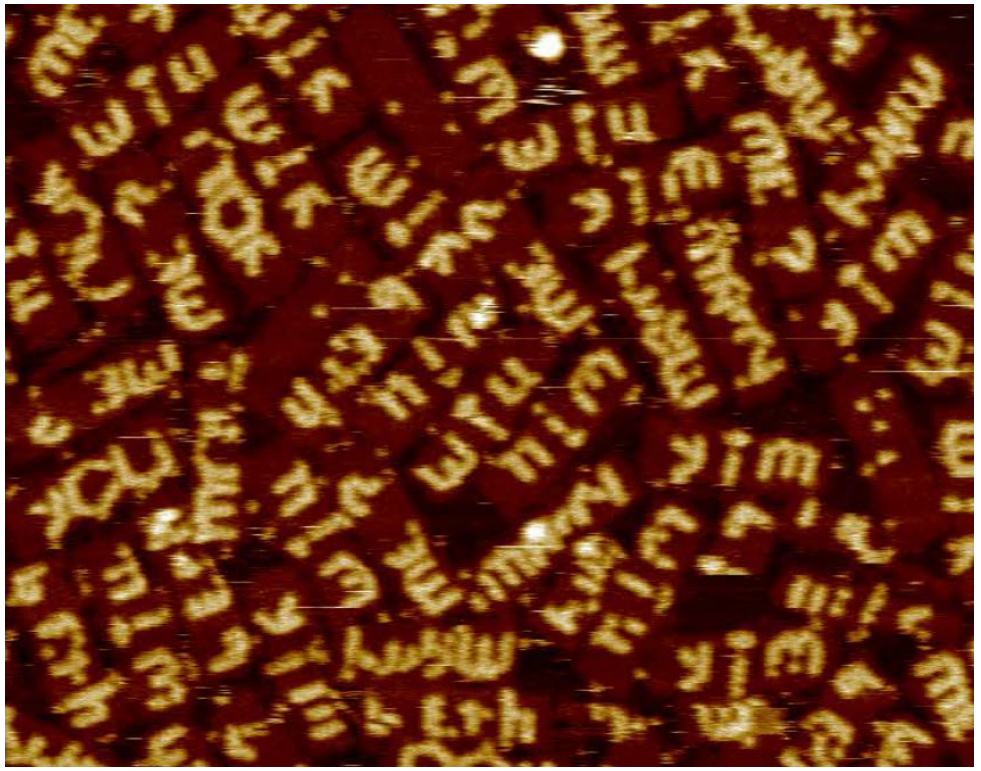
Grigory Tikhomirov, Philip Petersen, and Lulu Qian. *Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns*. Nature 2017.

Other applications of DNA nanotechnology

4 μm wide scan



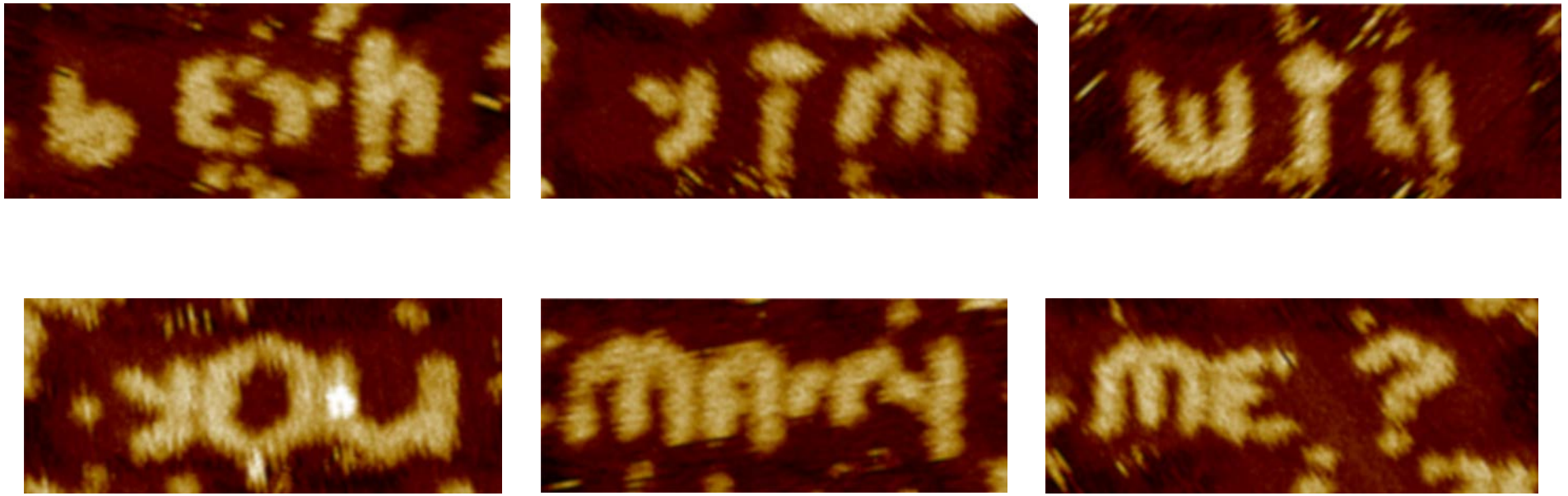
zoom in



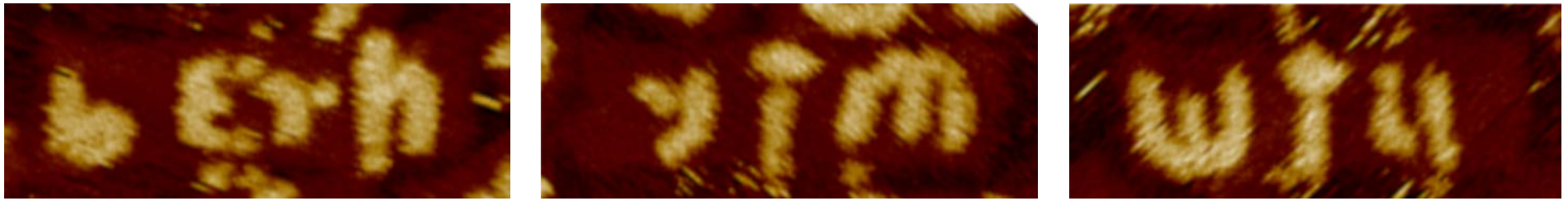
Cherry-picked samples

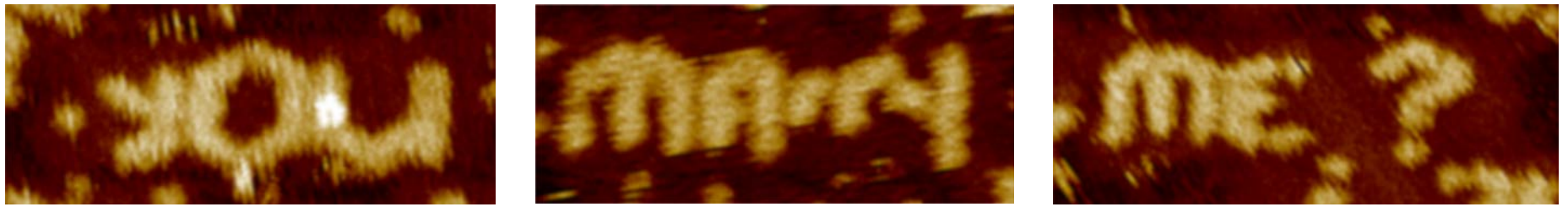
Beth Yim

Cherry-picked samples



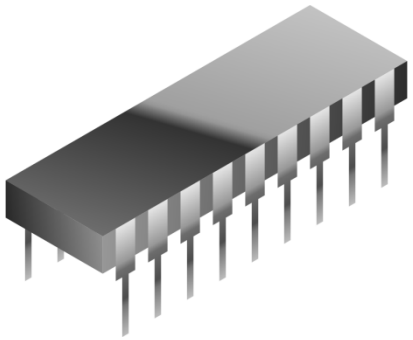
Cherry-picked samples

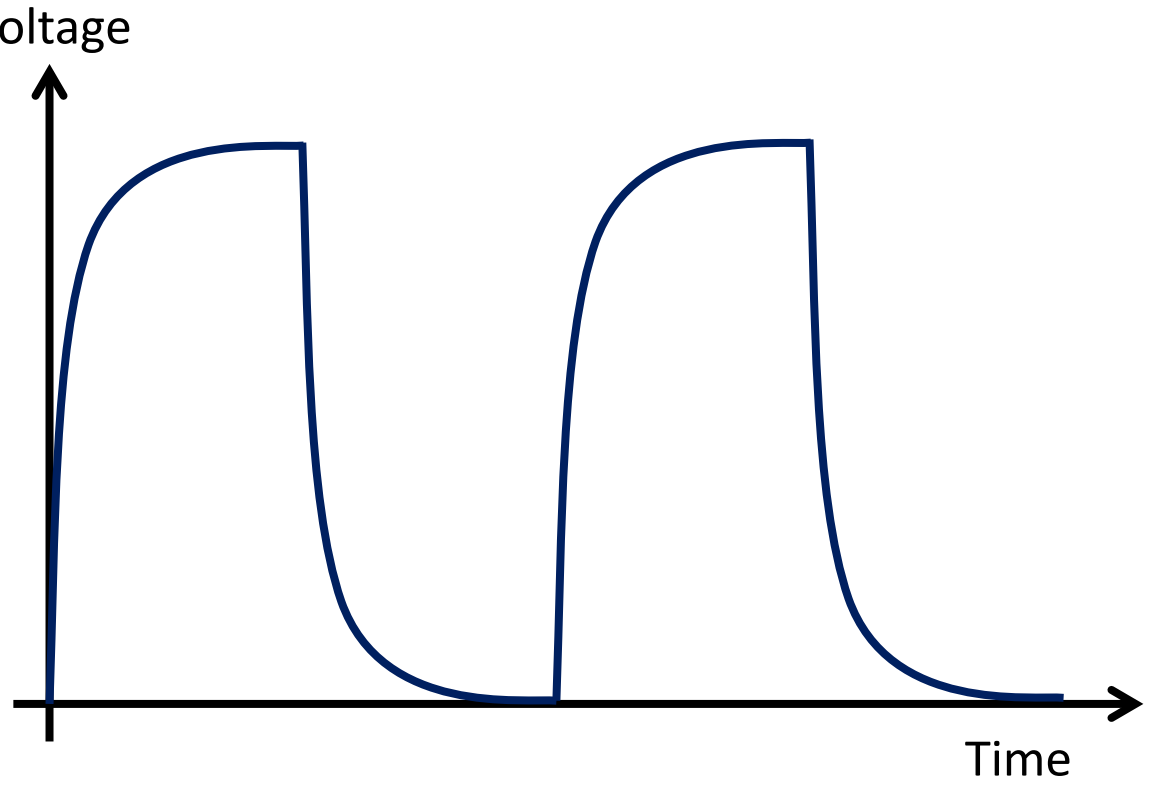




Beth Yim

Representing Information with Molecules

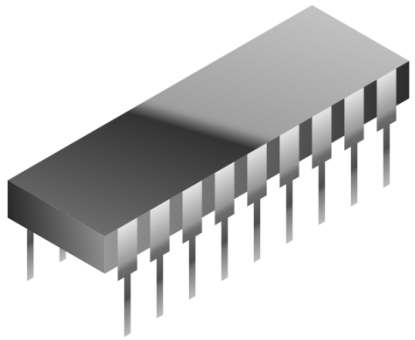


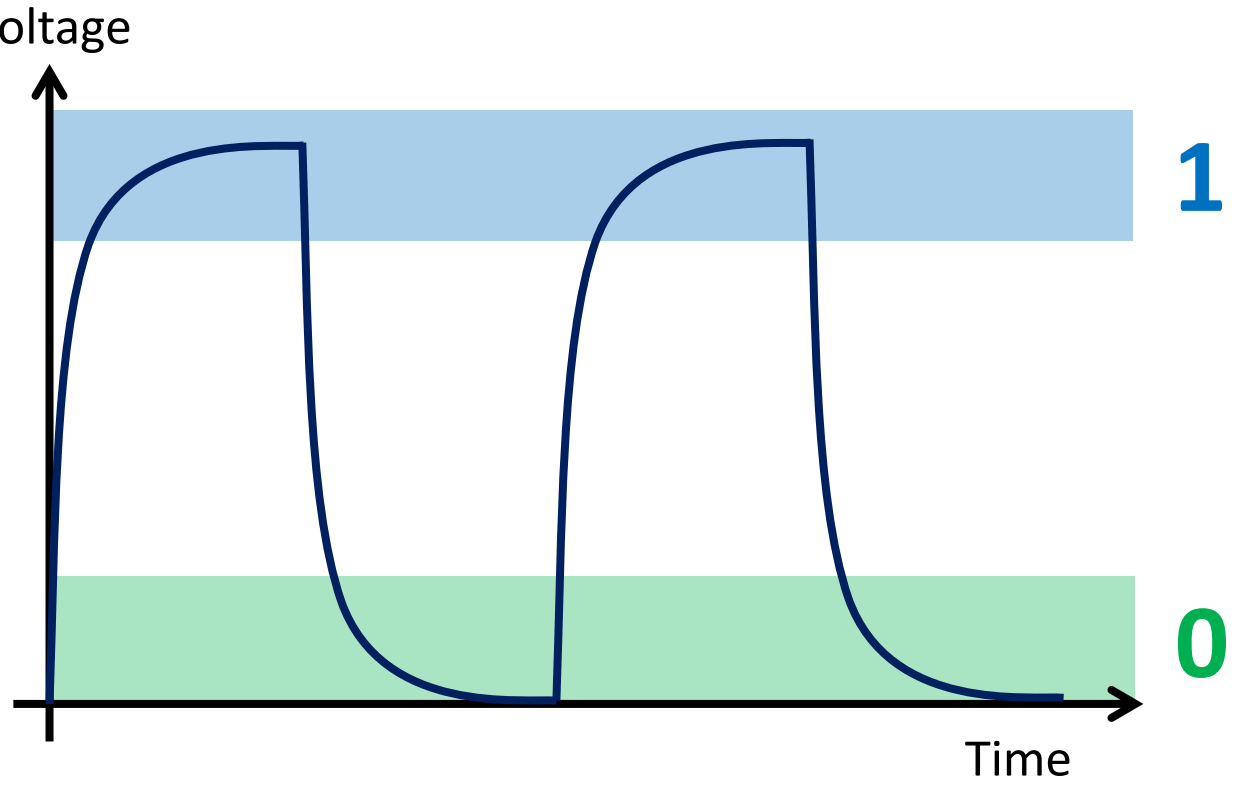


In an electronic circuit,

voltage can represent Boolean input

Representing Information with Molecules

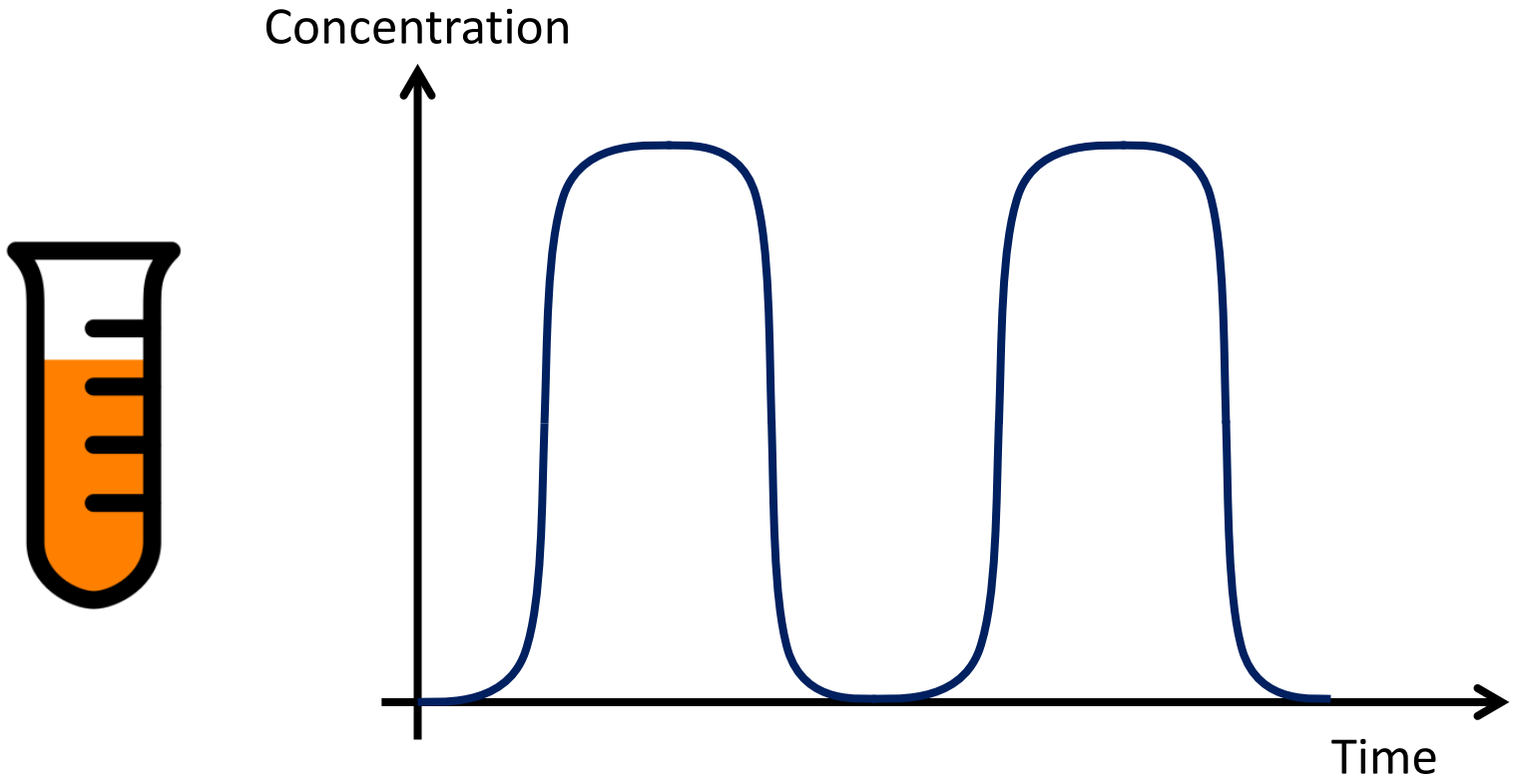




In an electronic circuit,

voltage can represent Boolean input

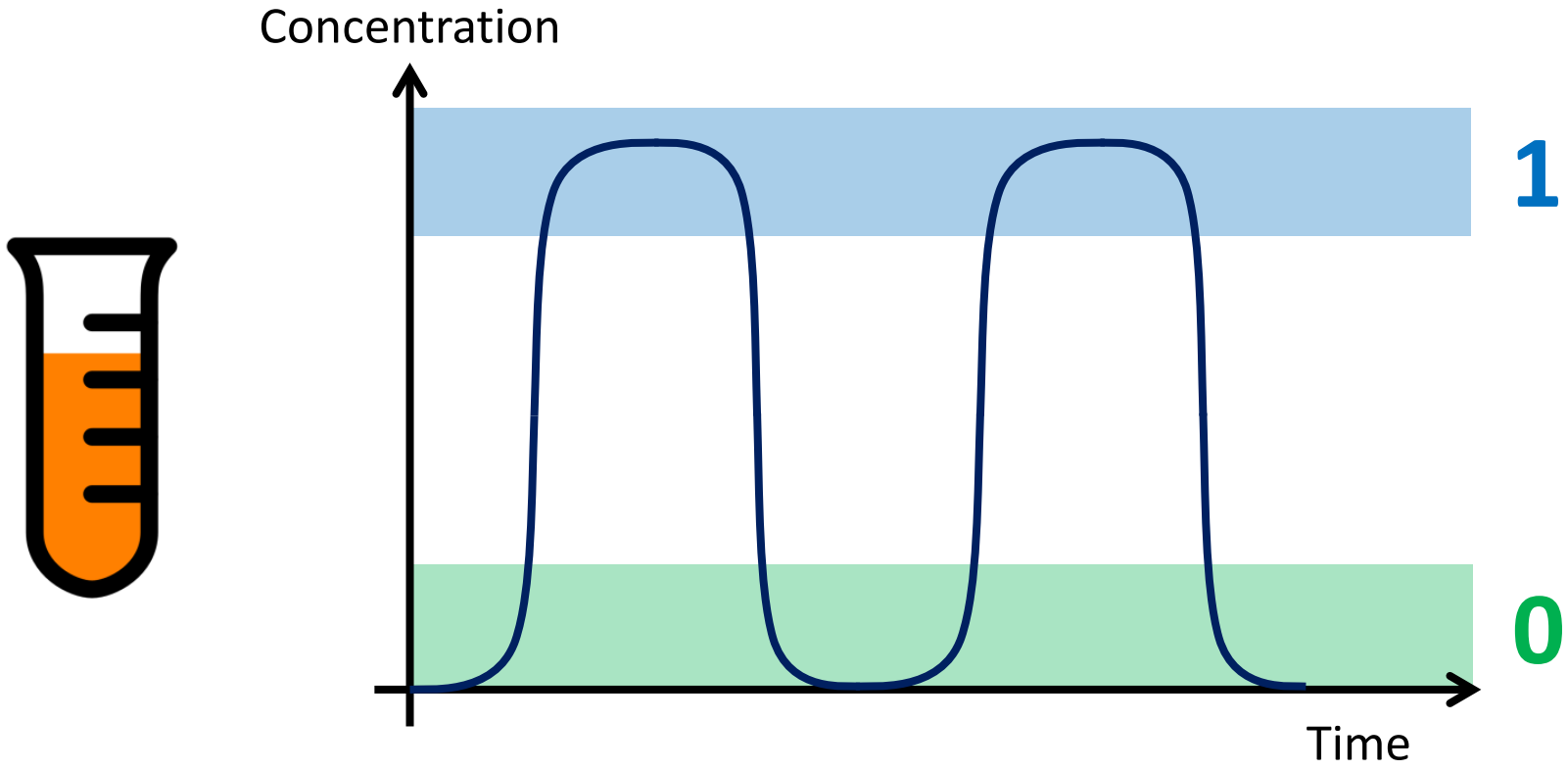
Representing Information with Molecules



In a well-mixed solution,

concentration can represent Boolean input

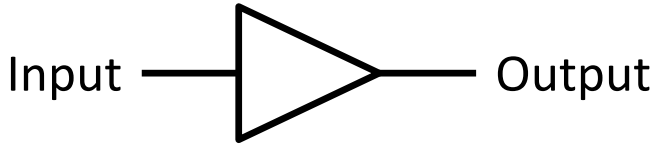
Representing Information with Molecules



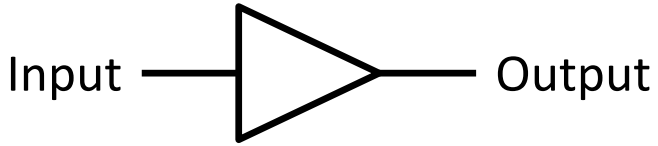
In a well-mixed solution,

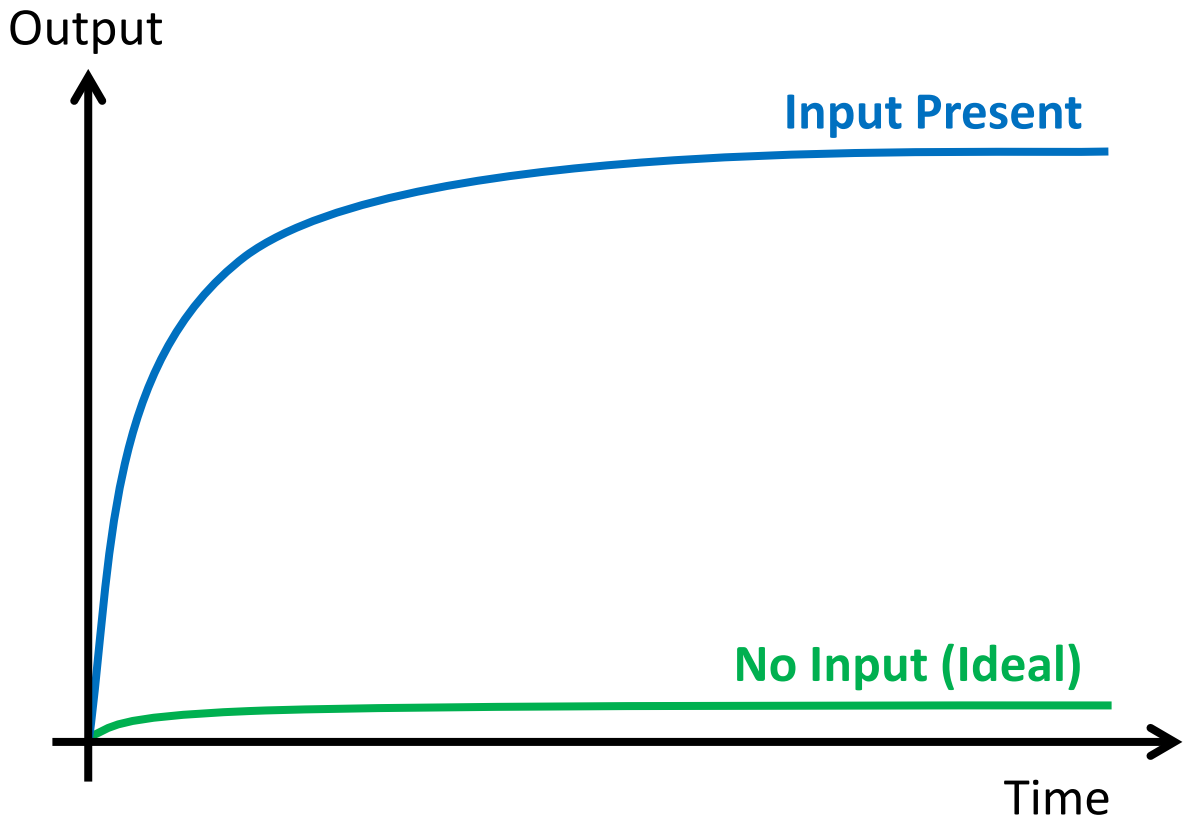
concentration can represent Boolean input

Chemical Identity Gate: Idealized vs. Actual Behavior



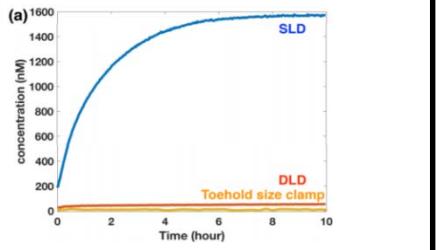
Chemical Identity Gate: Idealized vs. Actual Behavior



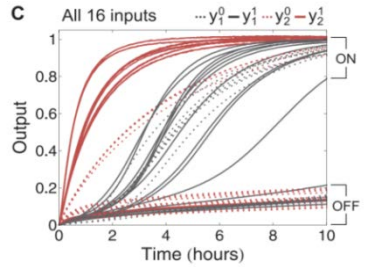


Chemical Identity Gate: Idealized vs. Actual Behavior

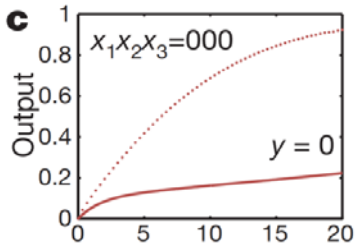
Experimental Implementation of Chemical Logic



Wang et al. ***DNA*** 23, 2017

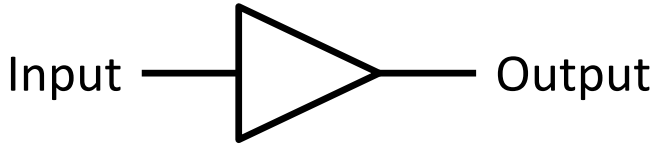


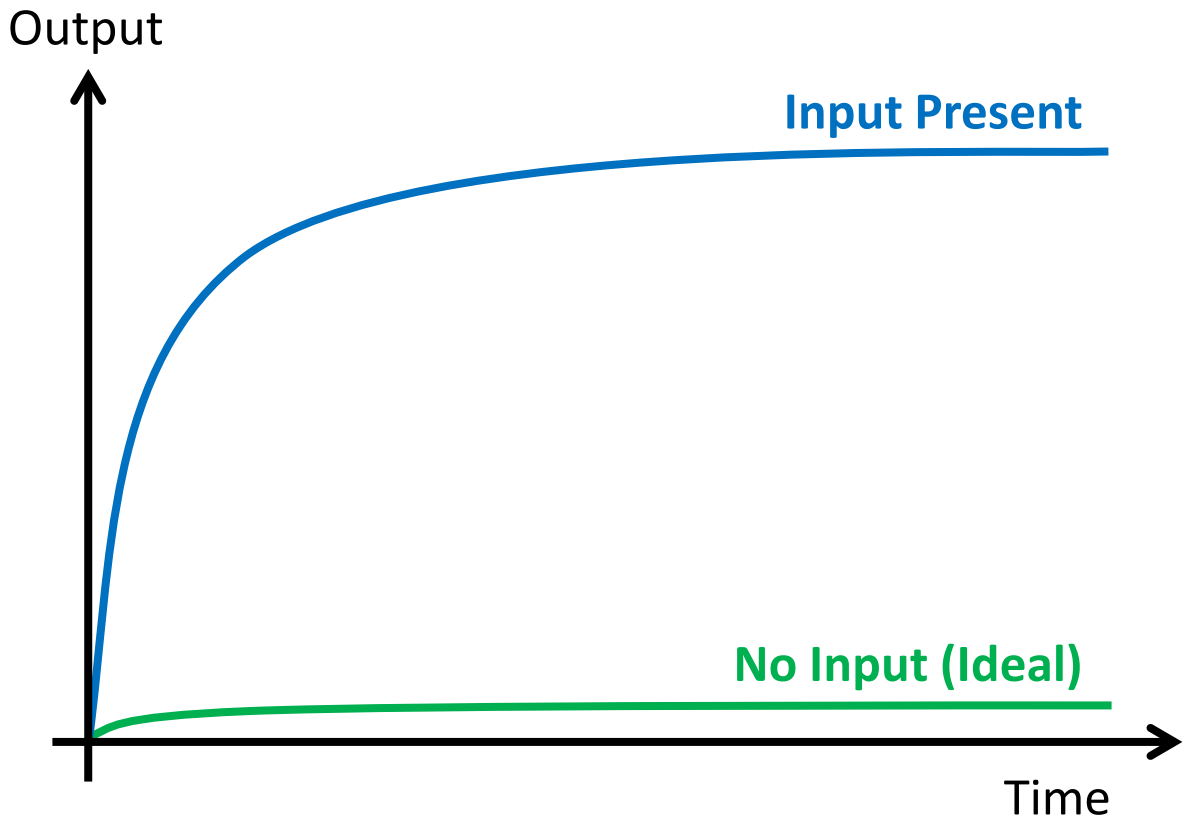
Qian et al. ***Science*** 332, 2011



Qian et al. ***Nature*** 475, 2011

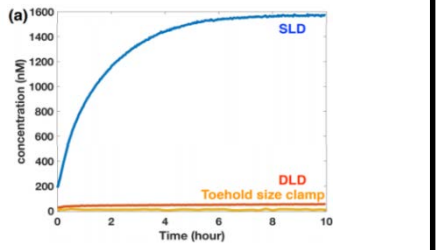
(Don't worry about the details in the pictures above)



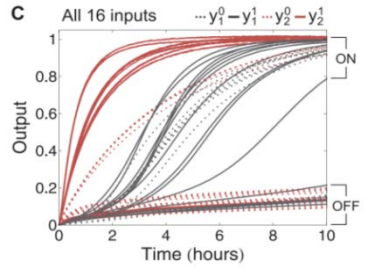


Chemical Identity Gate: Idealized vs. Actual Behavior

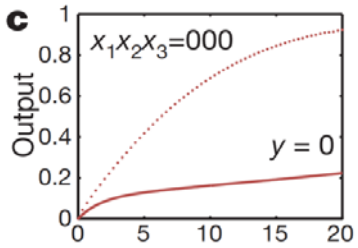
Experimental Implementation of Chemical Logic



Wang et al. ***DNA*** 23, 2017

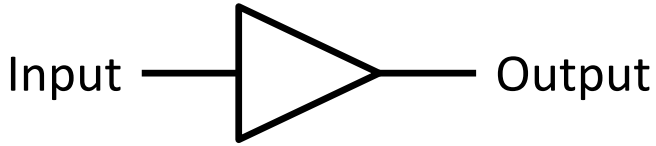


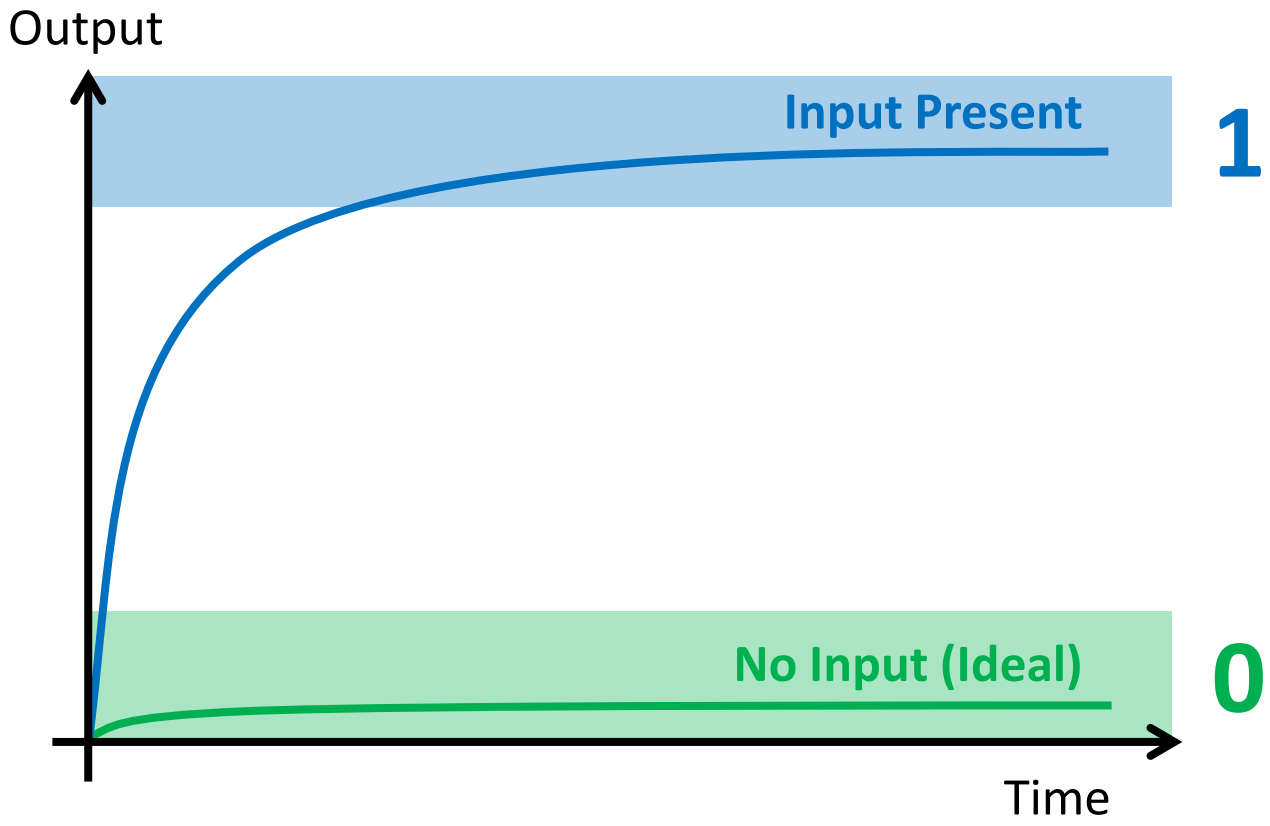
Qian et al. ***Science*** 332, 2011



Qian et al. ***Nature*** 475, 2011

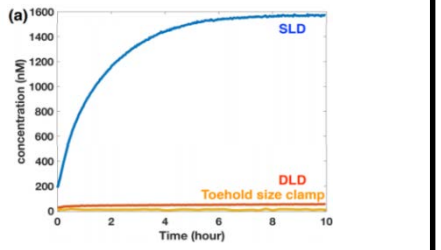
(Don't worry about the details in the pictures above)



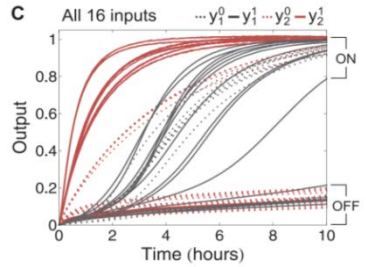


Chemical Identity Gate: Idealized vs. Actual Behavior

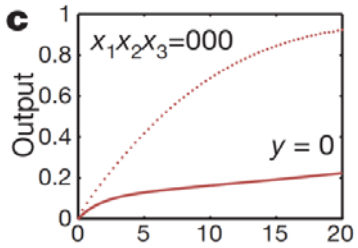
Experimental Implementation of Chemical Logic



Wang et al. ***DNA*** 23, 2017

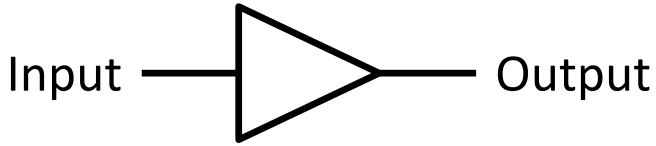


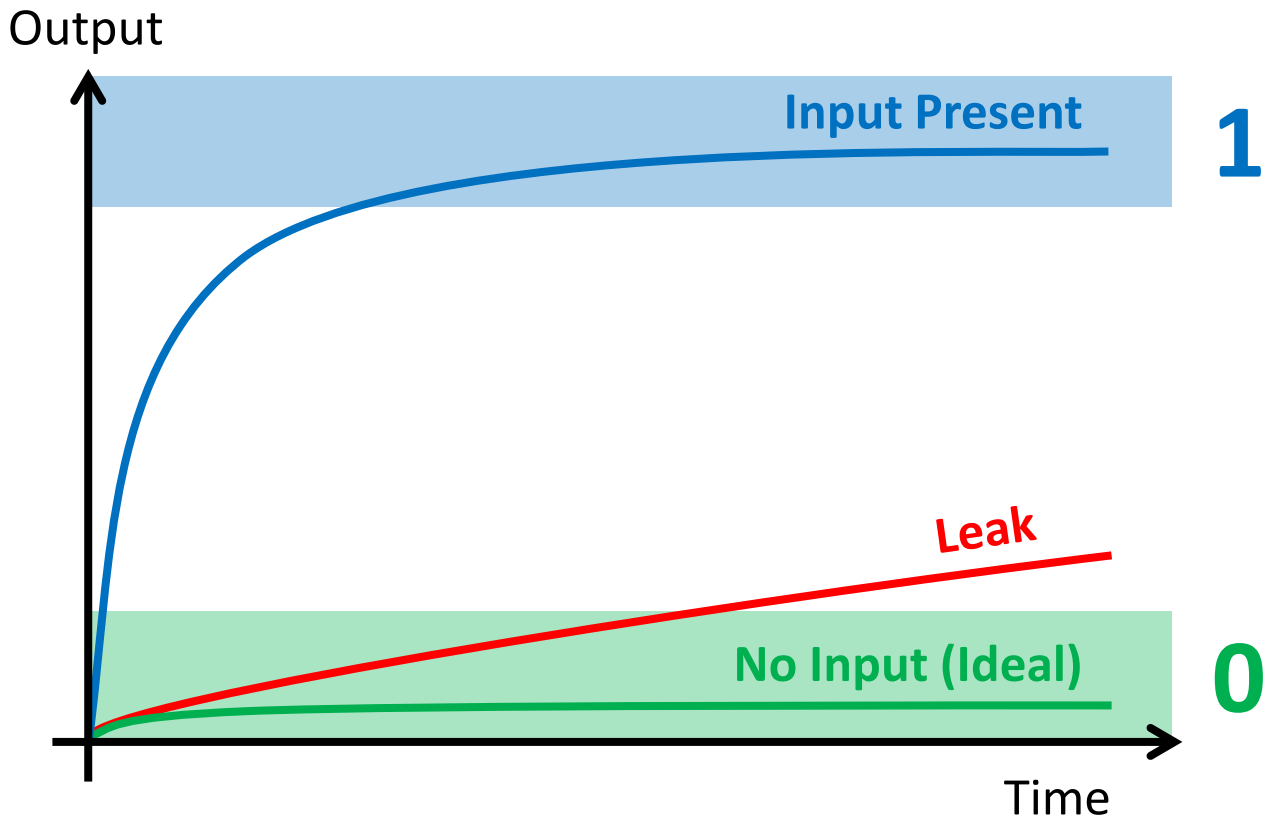
Qian et al. ***Science*** 332, 2011



Qian et al. ***Nature*** 475, 2011

(Don't worry about the details in the pictures above)





Levels of Abstraction

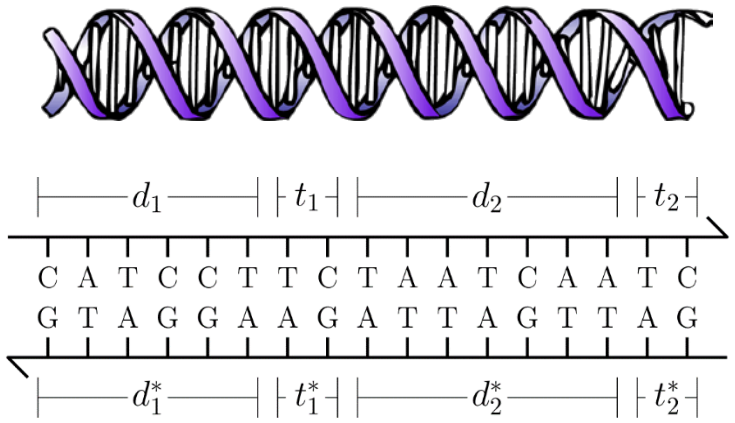
Levels of Abstraction

DNA

Levels of Abstraction

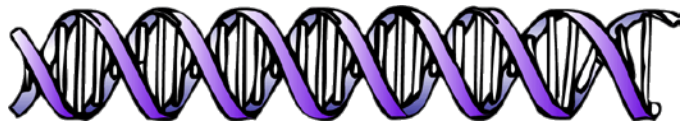
DNA

Base Pairs

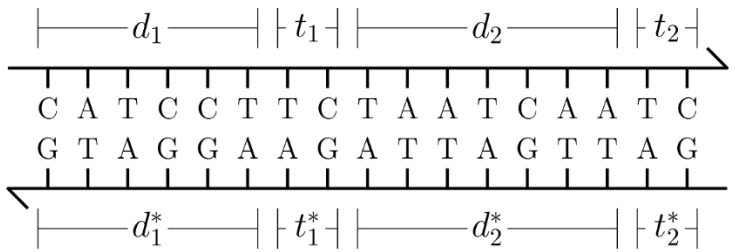


Levels of Abstraction

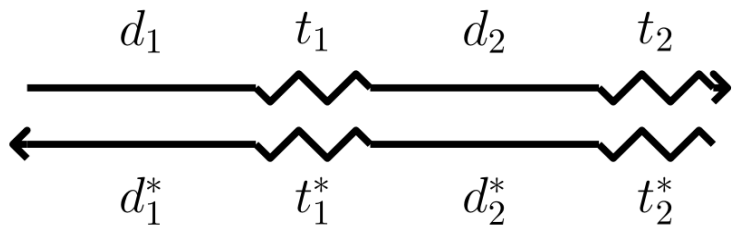
DNA



Base Pairs



Strands



Levels of Abstraction

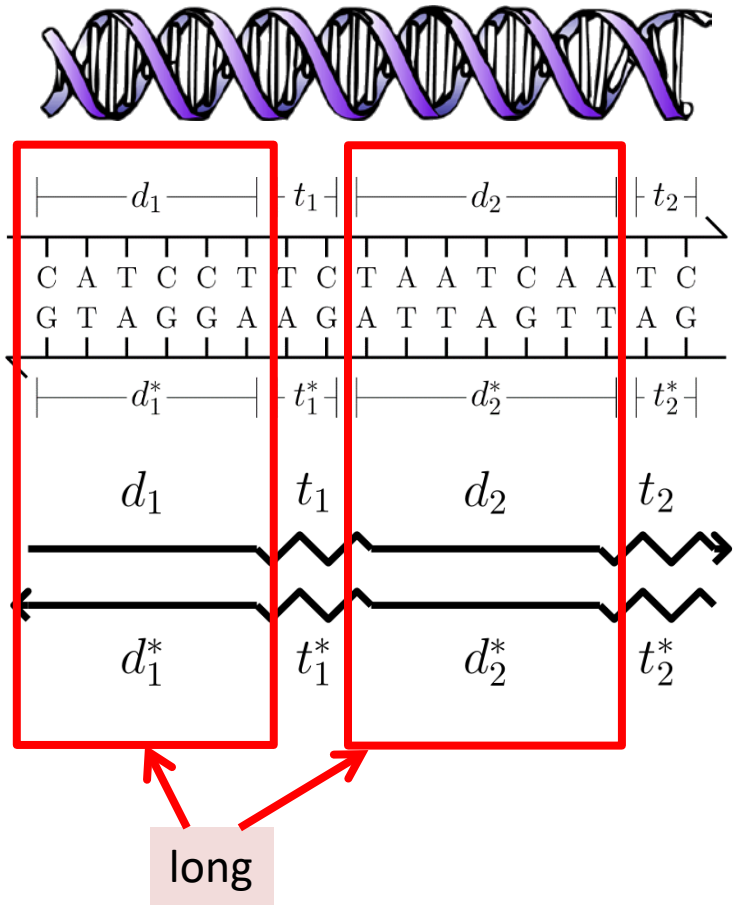
DNA

⬇

Base Pairs

⬇

Strands



Levels of Abstraction

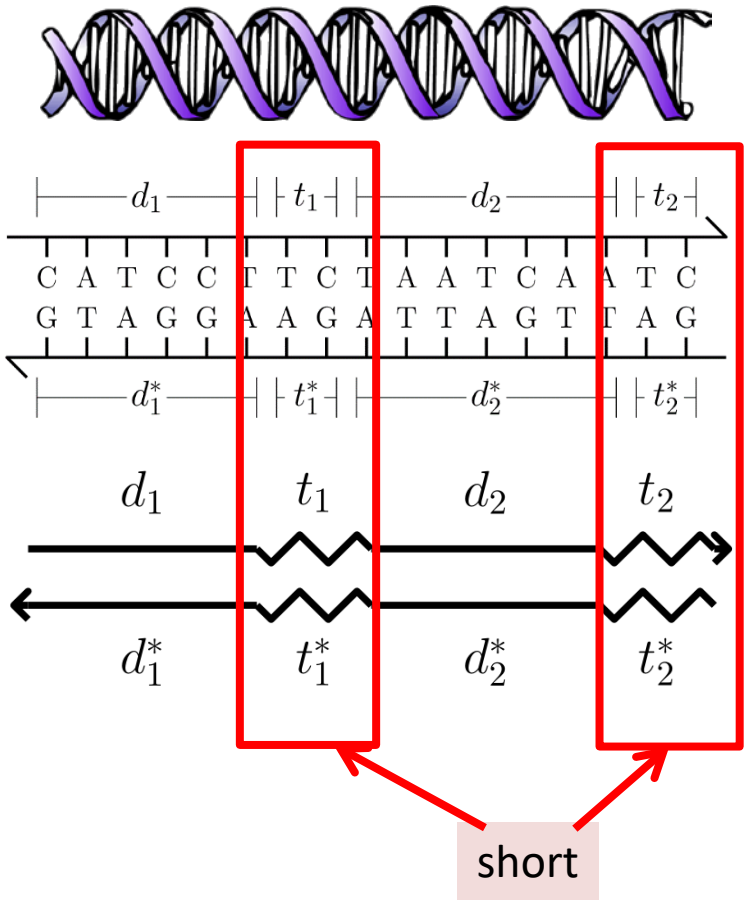
DNA

⬇

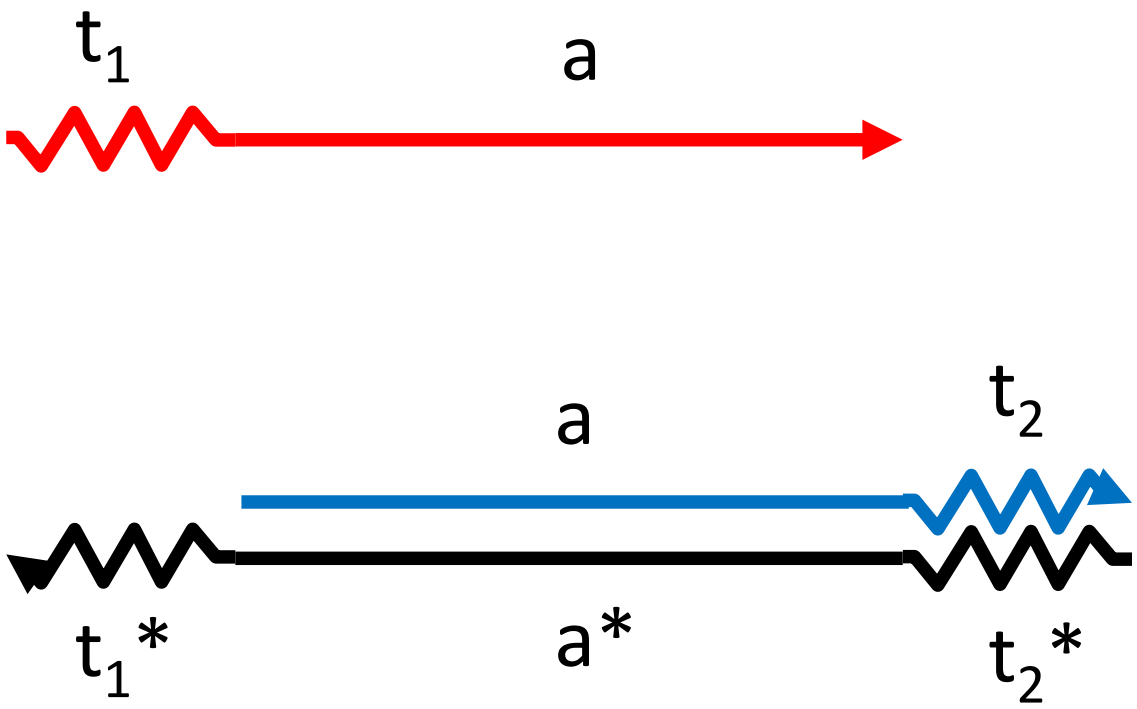
Base Pairs

⬇

Strands

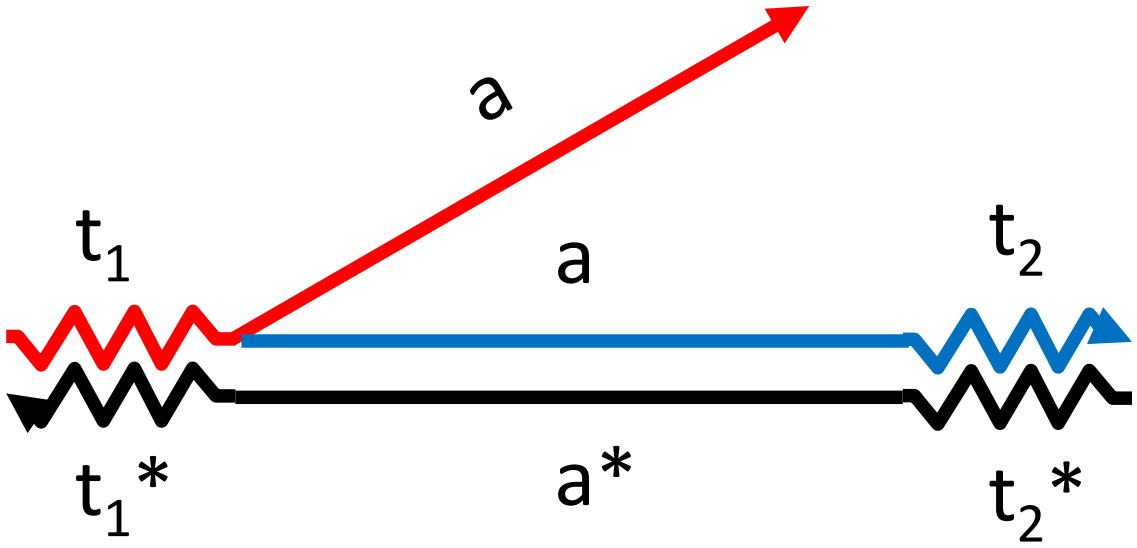


DNA strand displacement



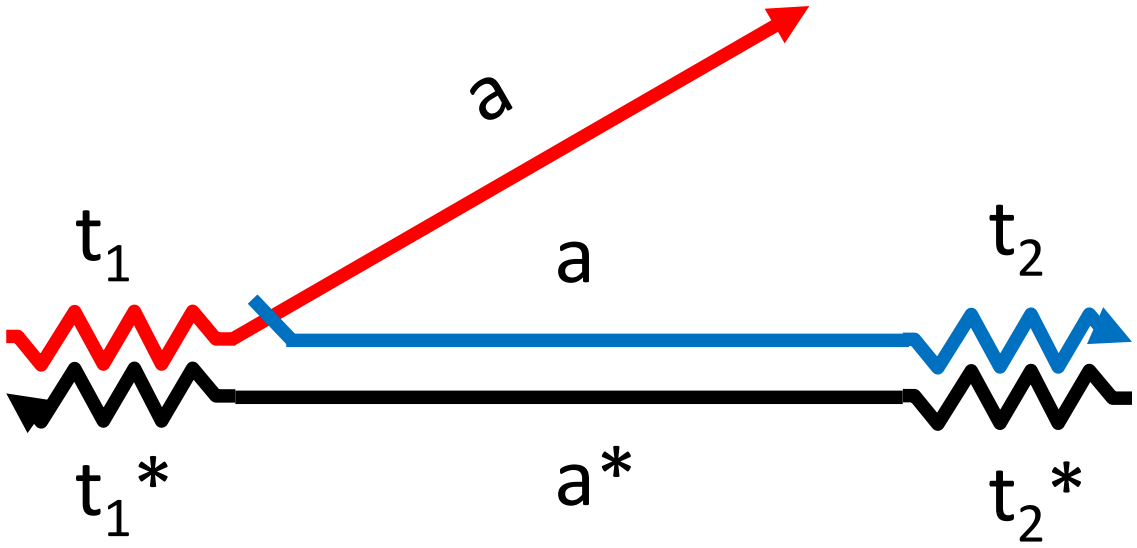
DNA strand displacement

Bind



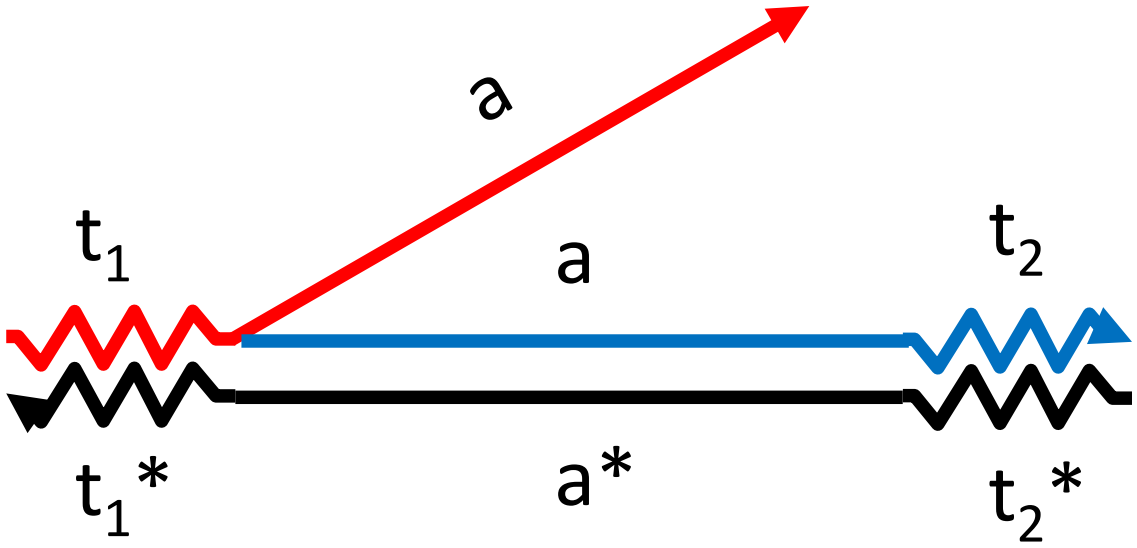
DNA strand displacement

Bind

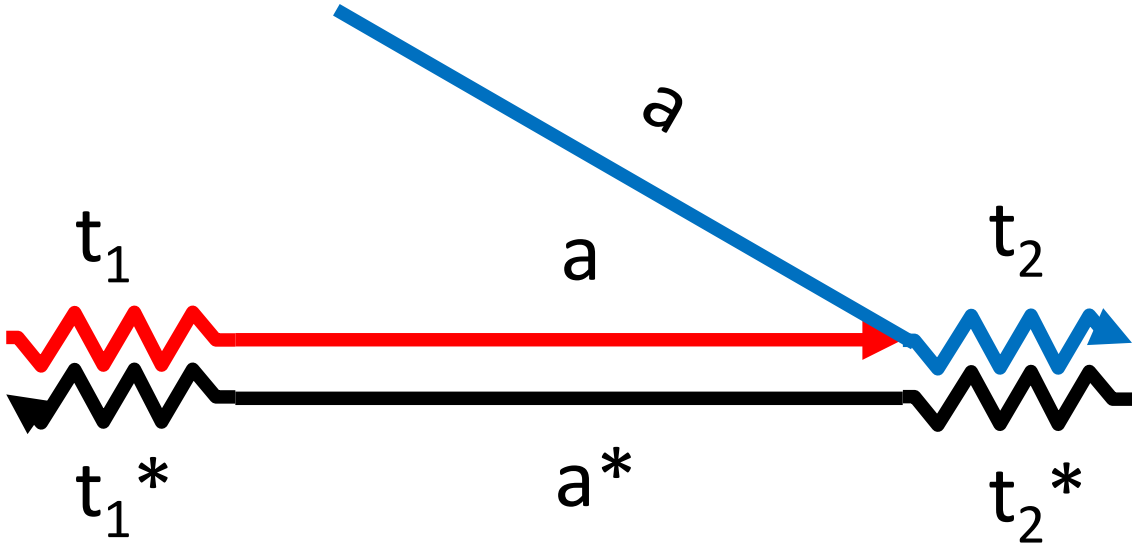


DNA strand displacement

Bind



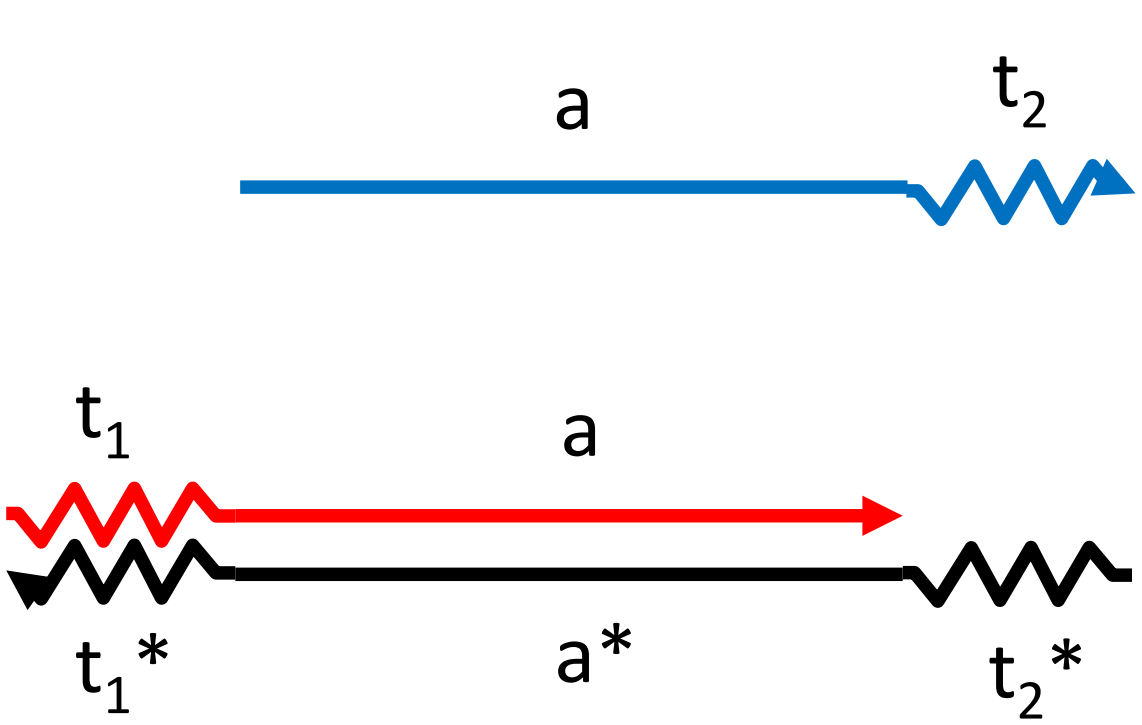
DNA strand displacement



Bind

Displace

DNA strand displacement



Bind

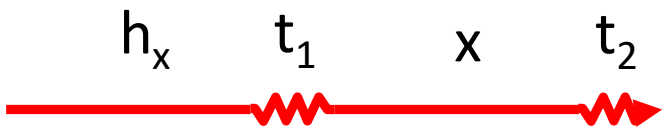
Displace

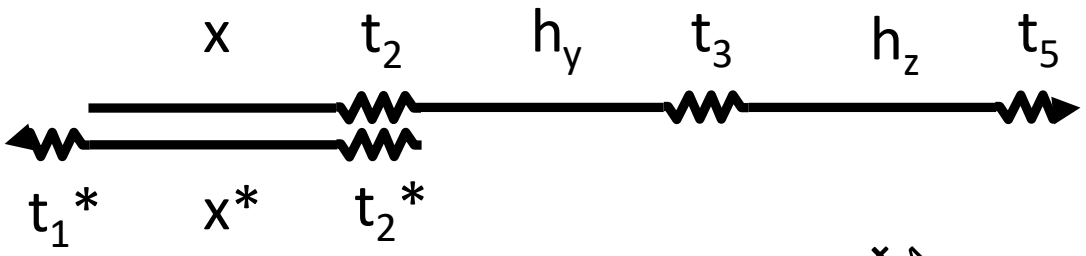
Release

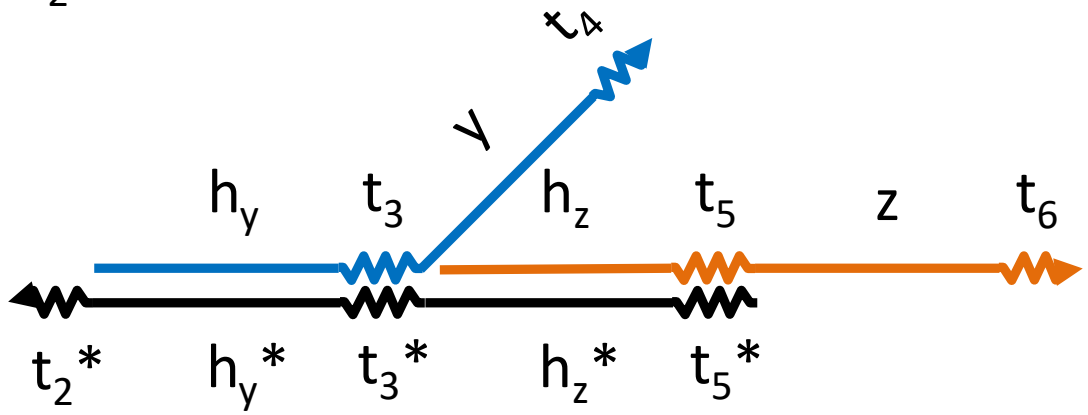
DNA strand displacement implementing A+B → C

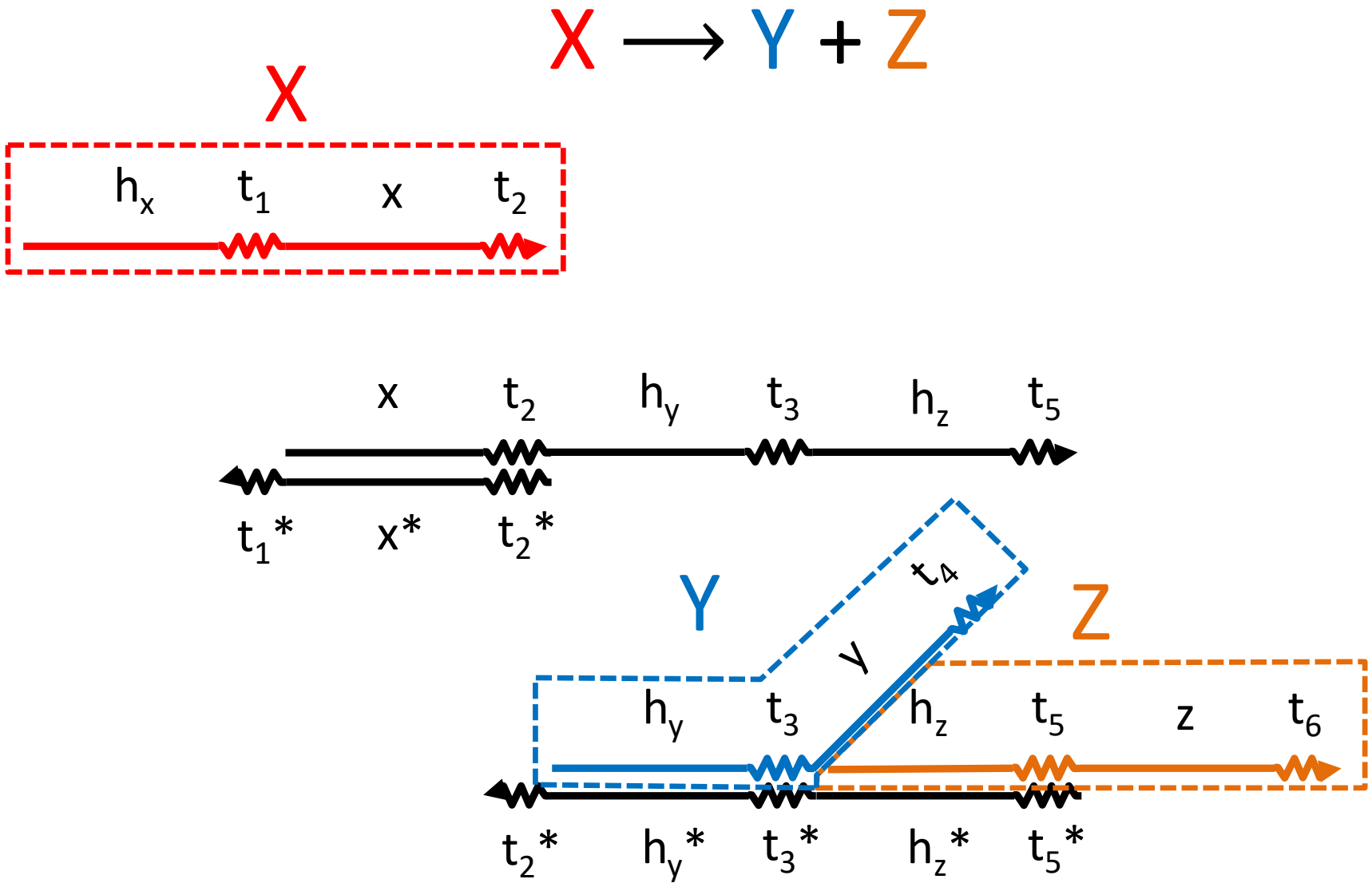
video: Microsoft Research Cambridge

$$X \longrightarrow Y + Z$$

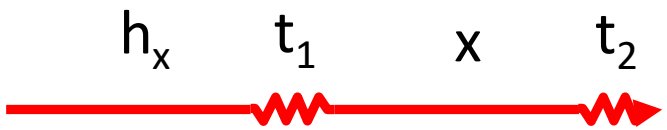


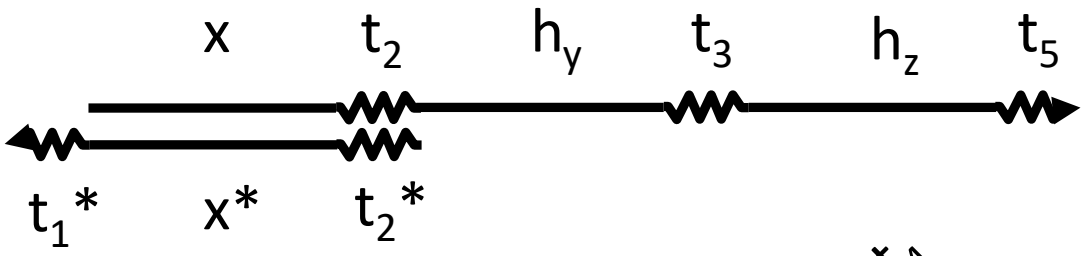


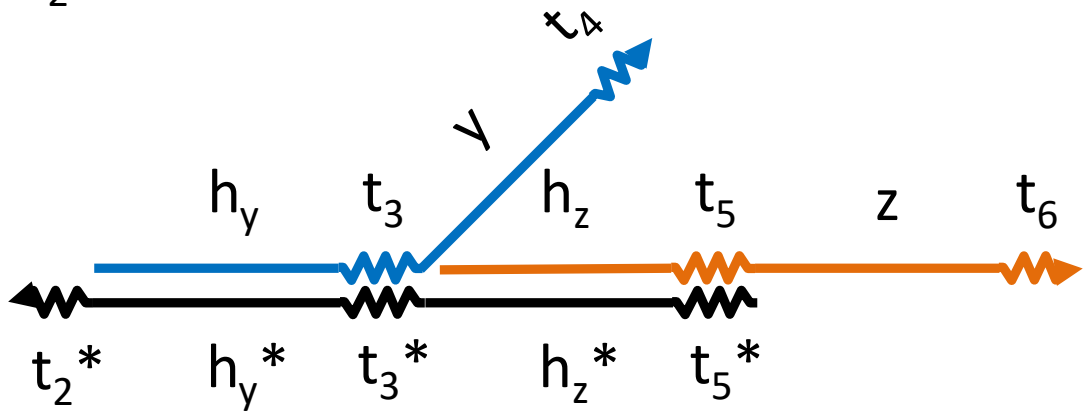



X ⟶ Y + Z
X
h_x
t_1
x
t_2
x
t_2
h_y
t_3
h_z
t_5
t_1*
x*
t_2*
Y
t_4
Z
y
h_y
t_3
h_z
t_5
z
t_6
t_2*
h_y*
t_3*
h_z*
t_5*

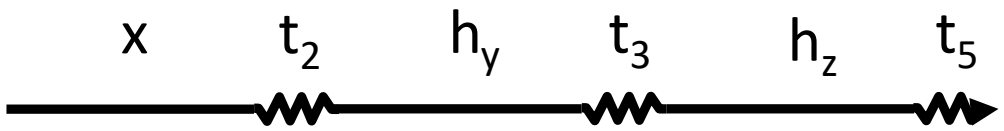
$$X \longrightarrow Y + Z$$

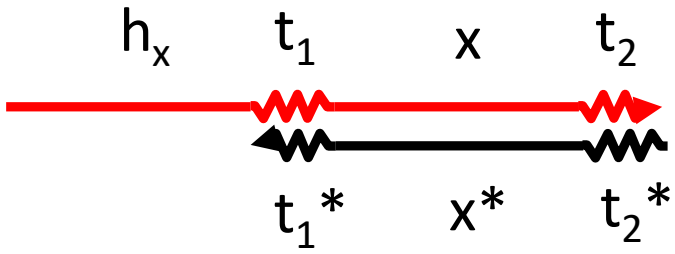


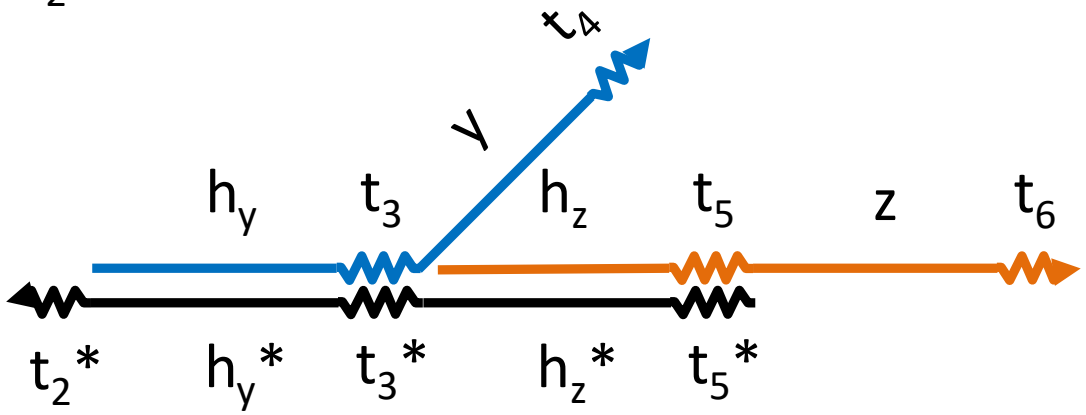


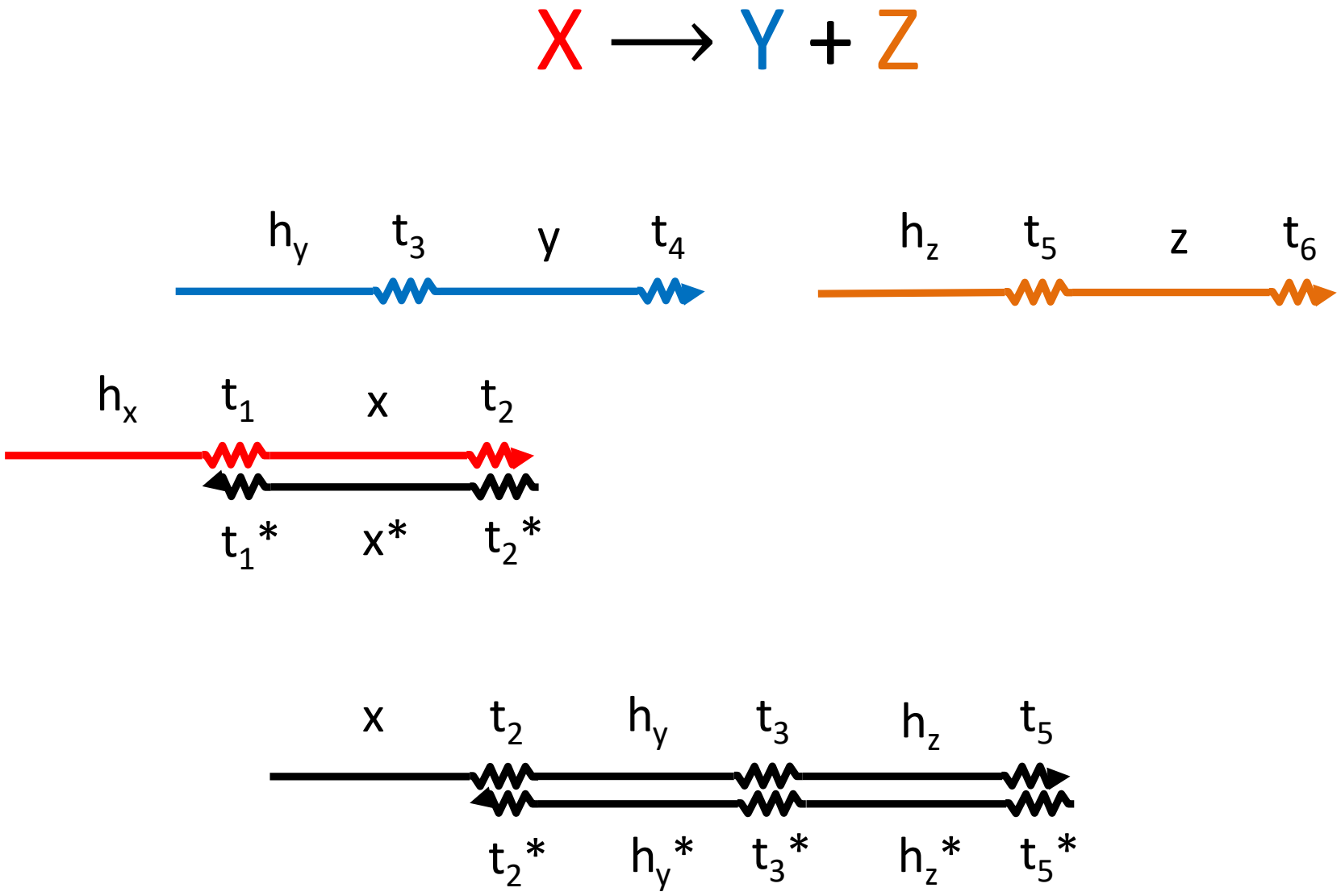


$$X \longrightarrow Y + Z$$

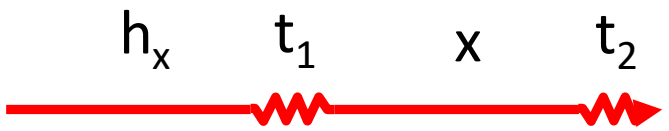


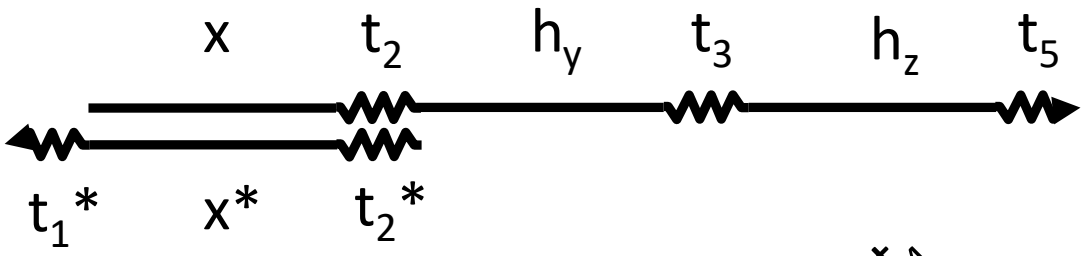


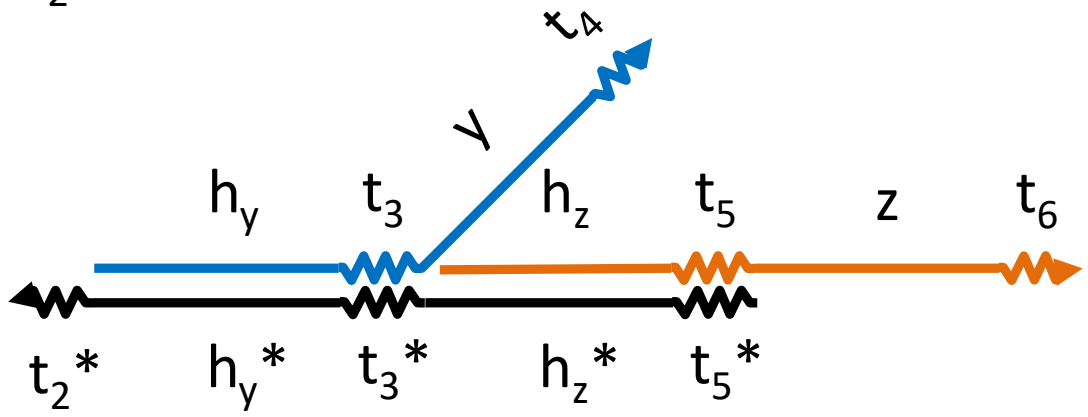



X ⟶ Y + Z
h_y t_3 y t_4
h_z t_5 z t_6
h_x t_1 x t_2
t_1* x* t_2*
x t_2 h_y t_3 h_z t_5
t_2* h_y* t_3* h_z* t_5*

$$X \longrightarrow Y + Z$$

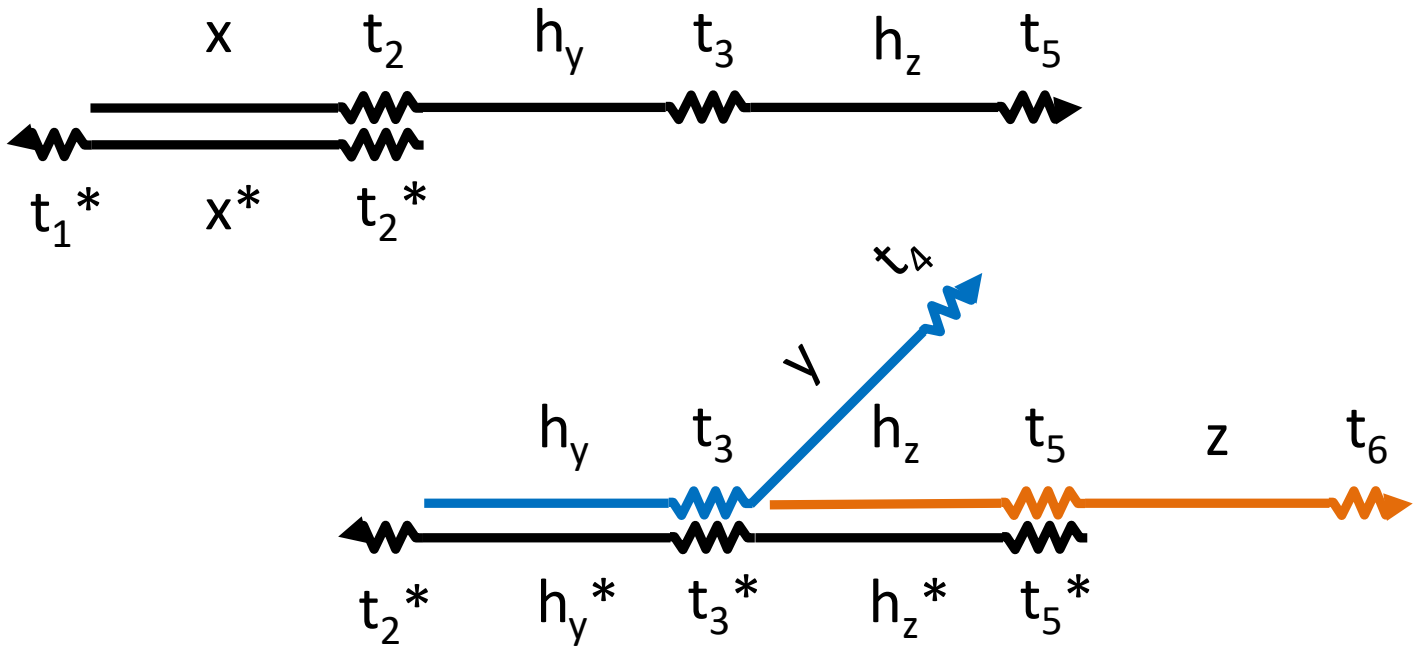






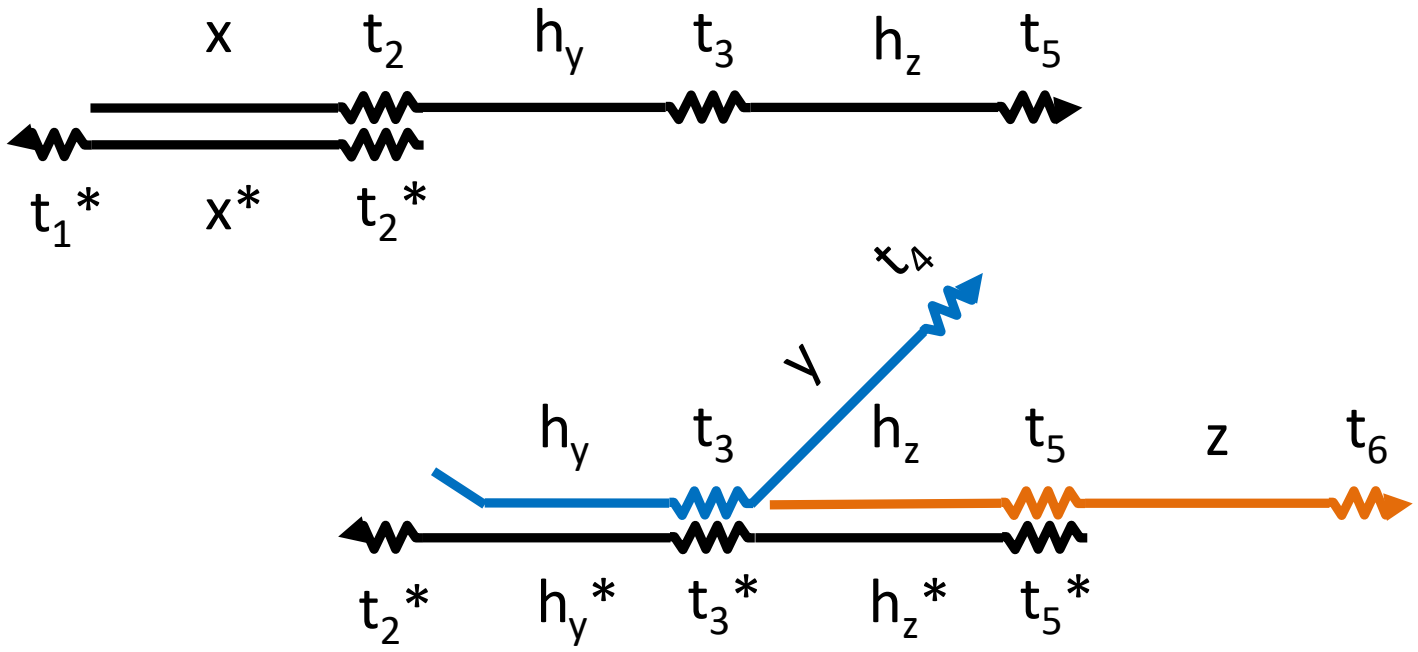
Leak: A Source of Error

$$\emptyset \longrightarrow Y + Z$$



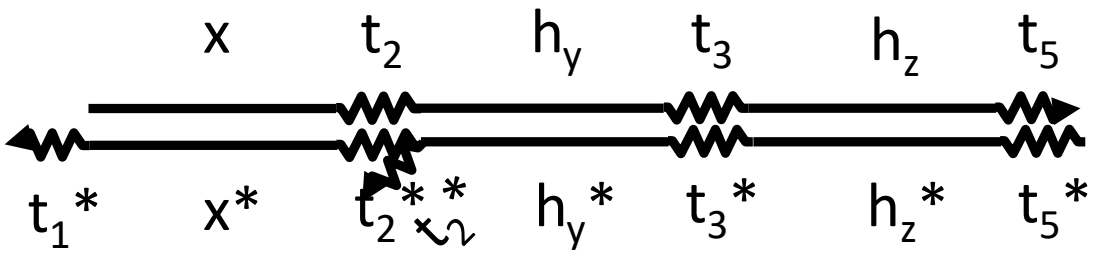
Leak: A Source of Error

$$\emptyset \longrightarrow Y + Z$$



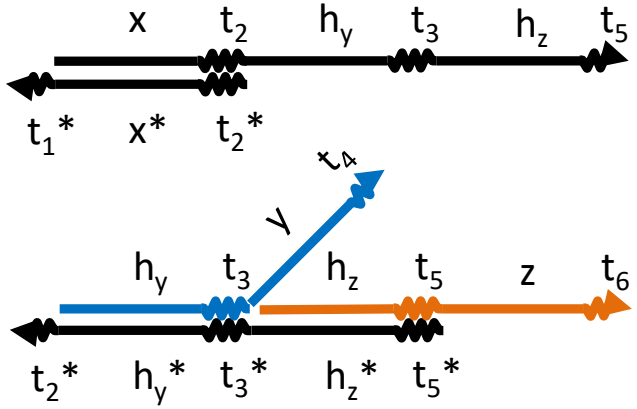
Leak: A Source of Error

$$\emptyset \longrightarrow Y + Z$$

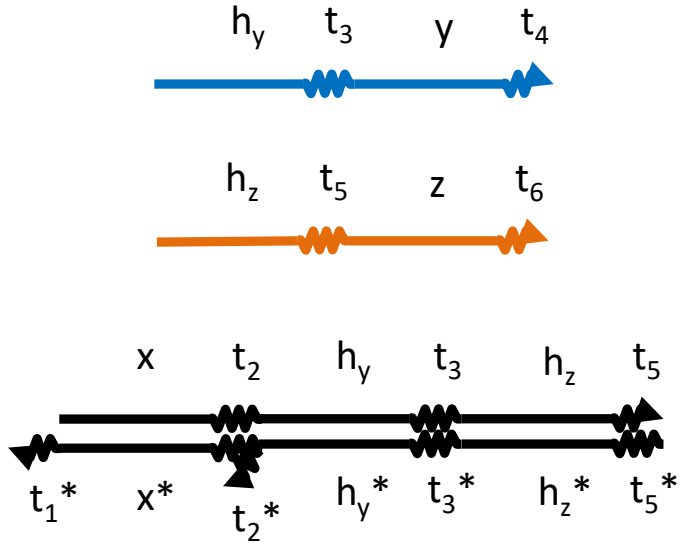


A Driving Force for Leak

Before:

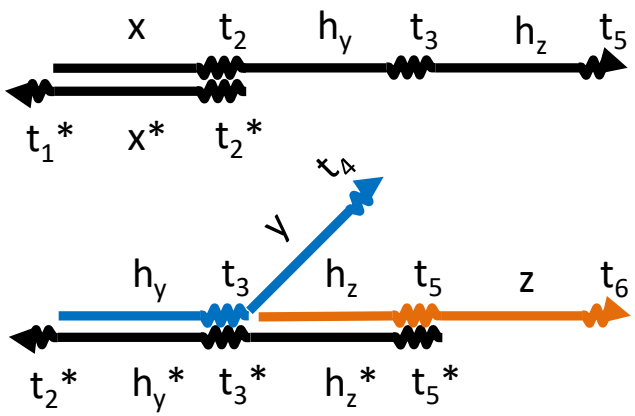


After:



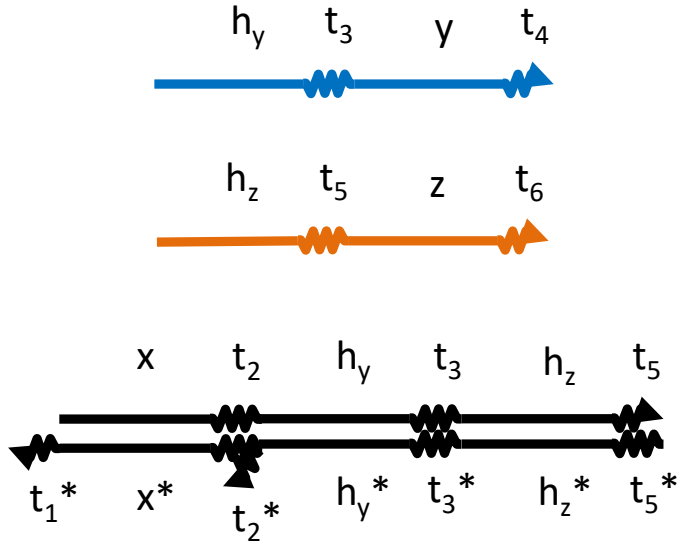
A Driving Force for Leak

Before:



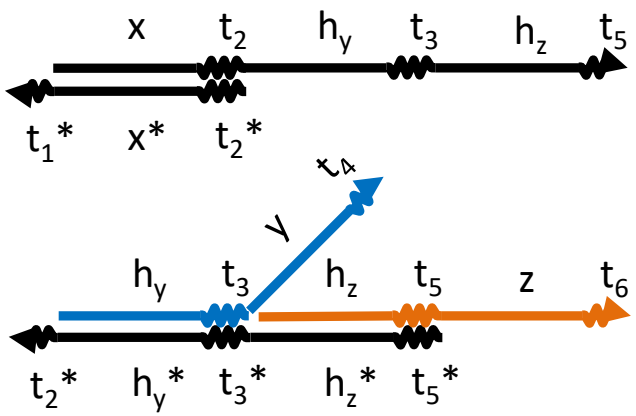
slow

After:



A Driving Force for Leak

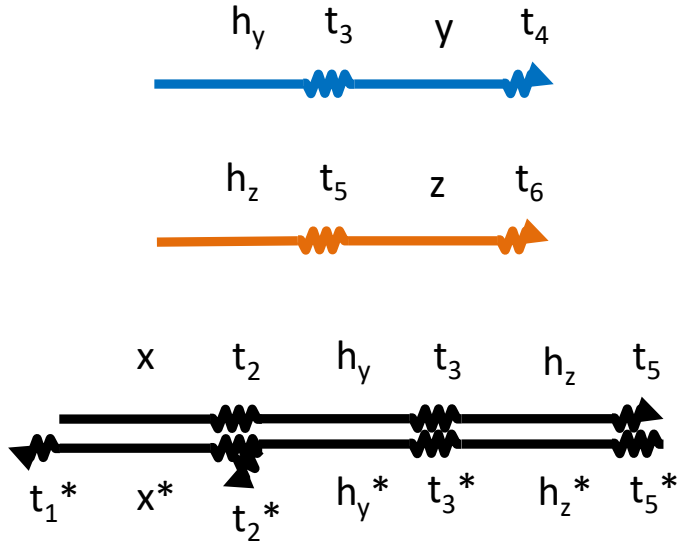
Before:



slow

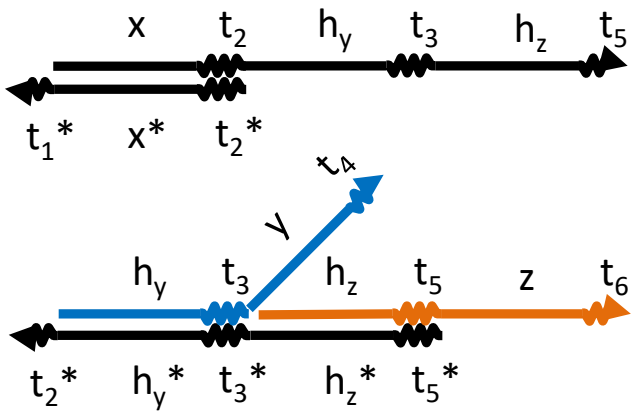
very slow

After:



A Driving Force for Leak

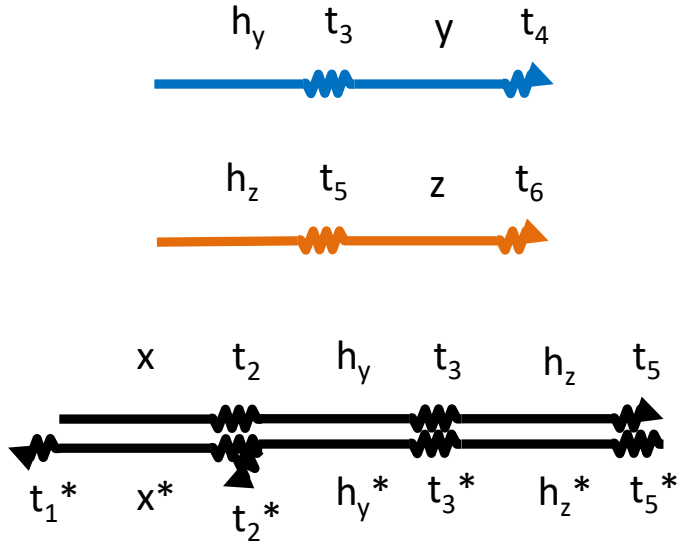
Before:



slow

very slow

After:

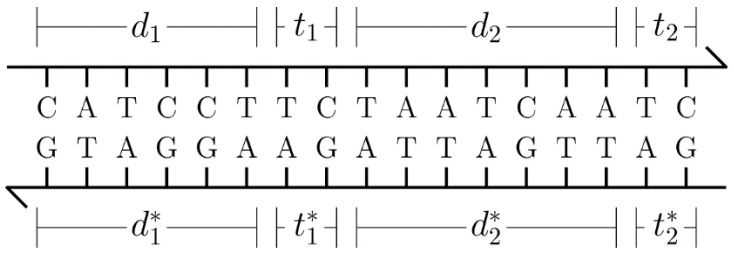


less favorable

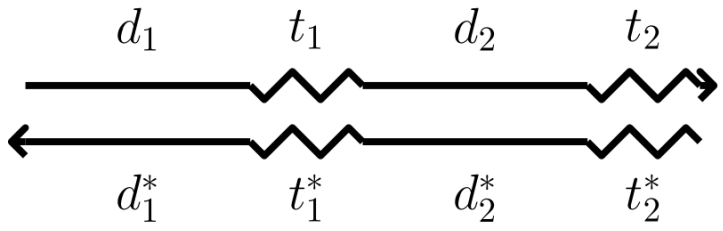
more favorable

Levels of Abstraction

DNA



Base Pairs

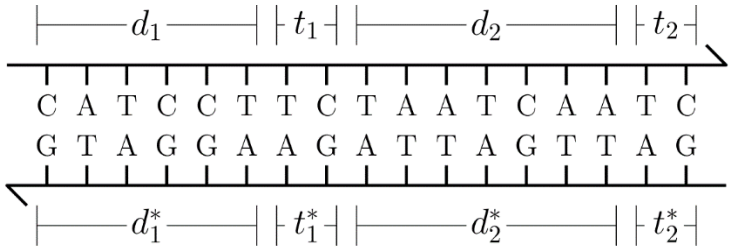


Strands

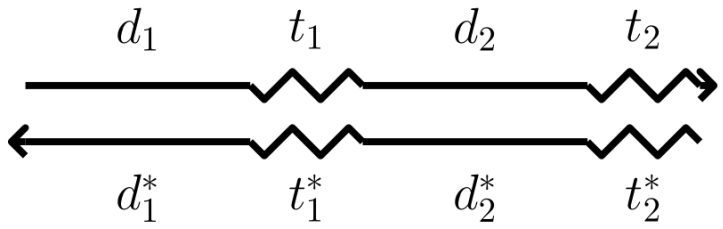
Levels of Abstraction

DNA

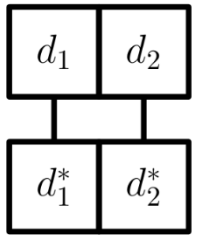
Base Pairs



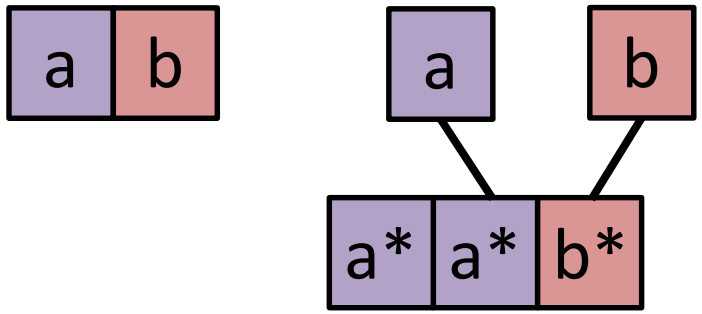
Strands



Thermodynamic Binding Network



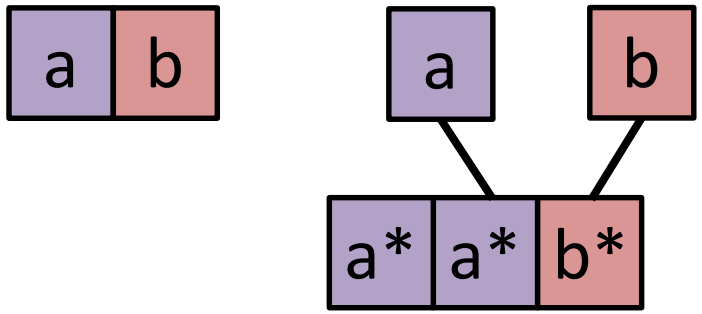
Thermodynamic Binding Networks



Monomer = collection of domains

Configuration = how monomers are bound

Thermodynamic Binding Networks



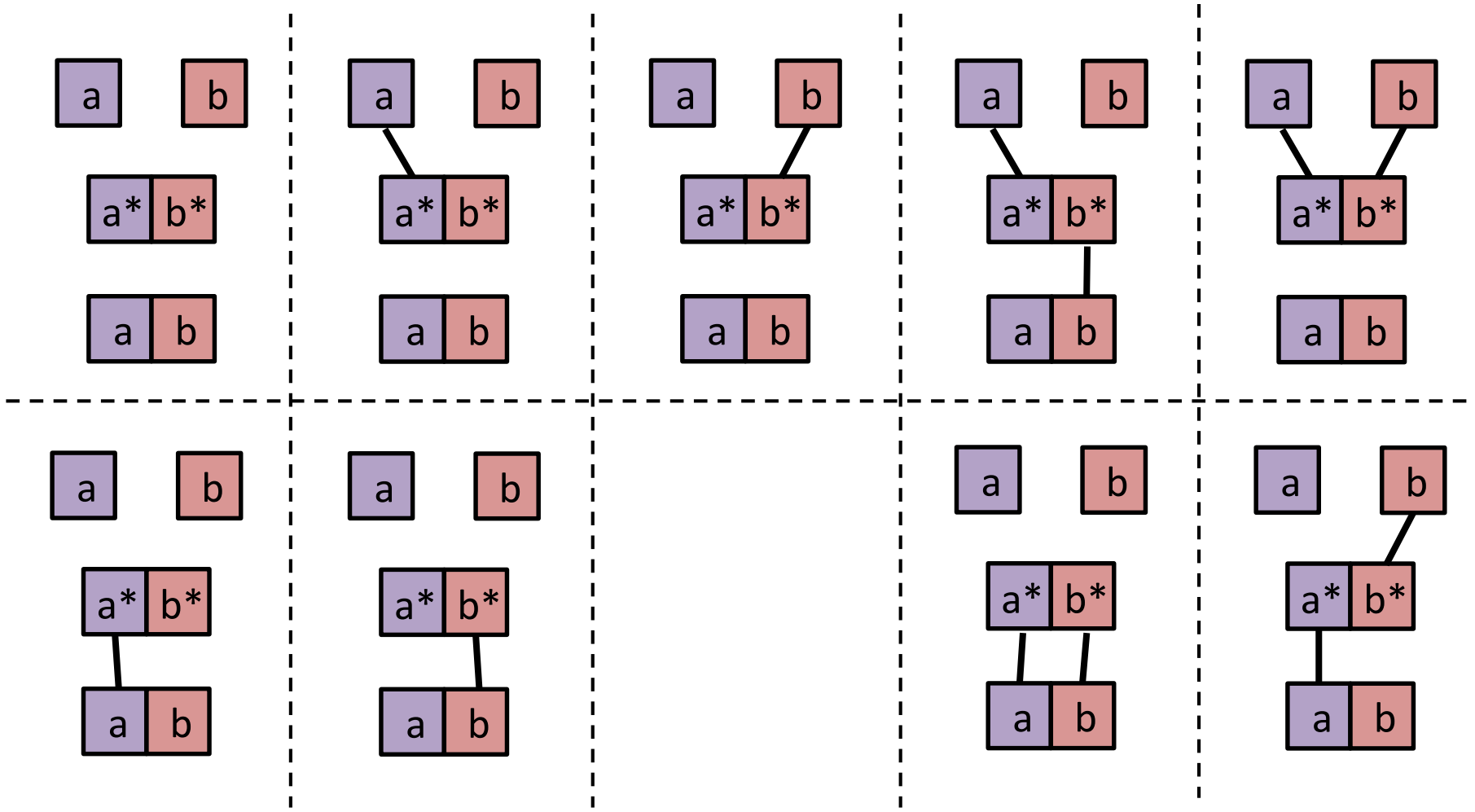
Geometry-Free Model:

The domains within a monomer are unordered

Monomer = collection of domains

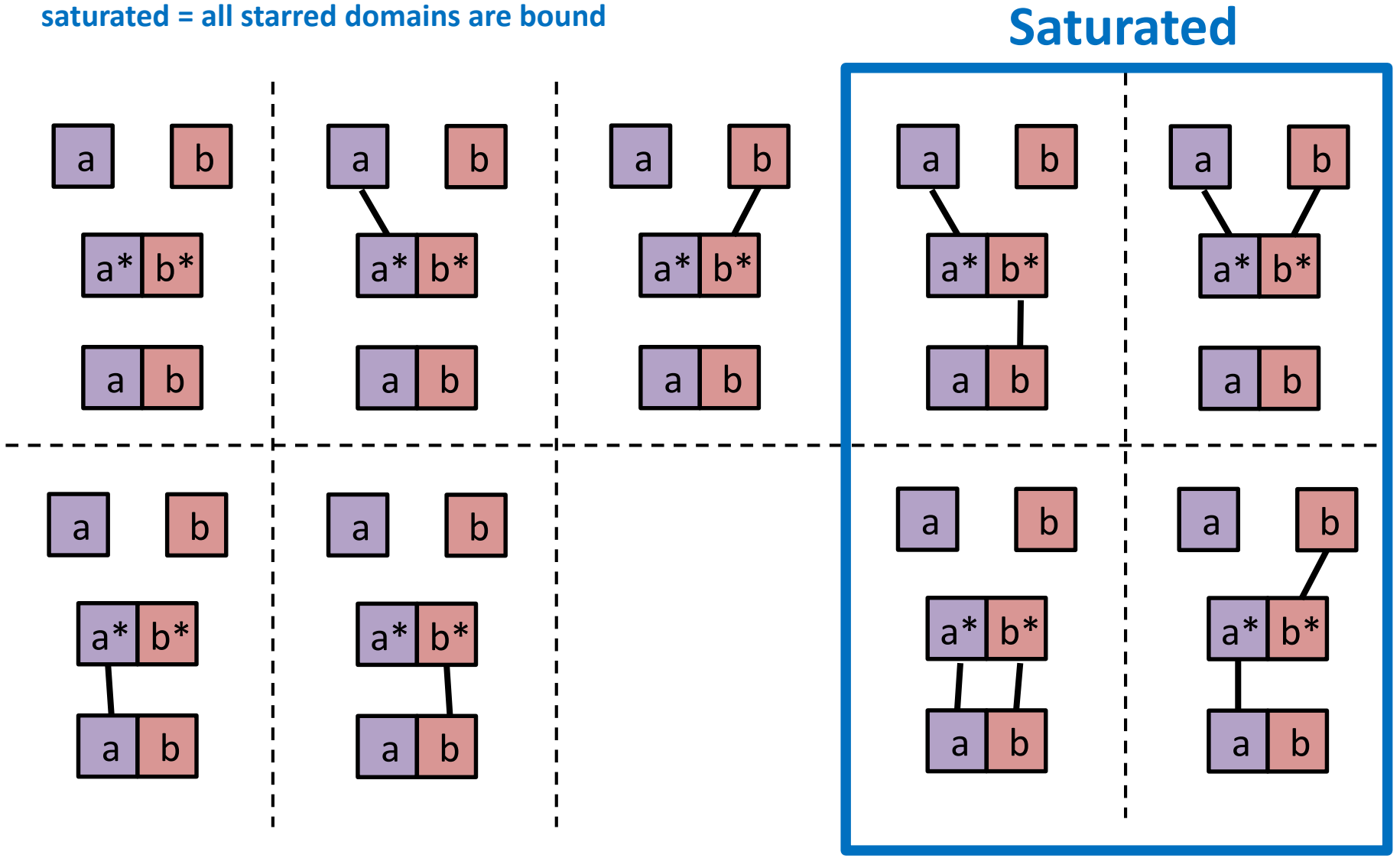
Configuration = how monomers are bound

Thermodynamic Binding Networks



Thermodynamic Binding Networks

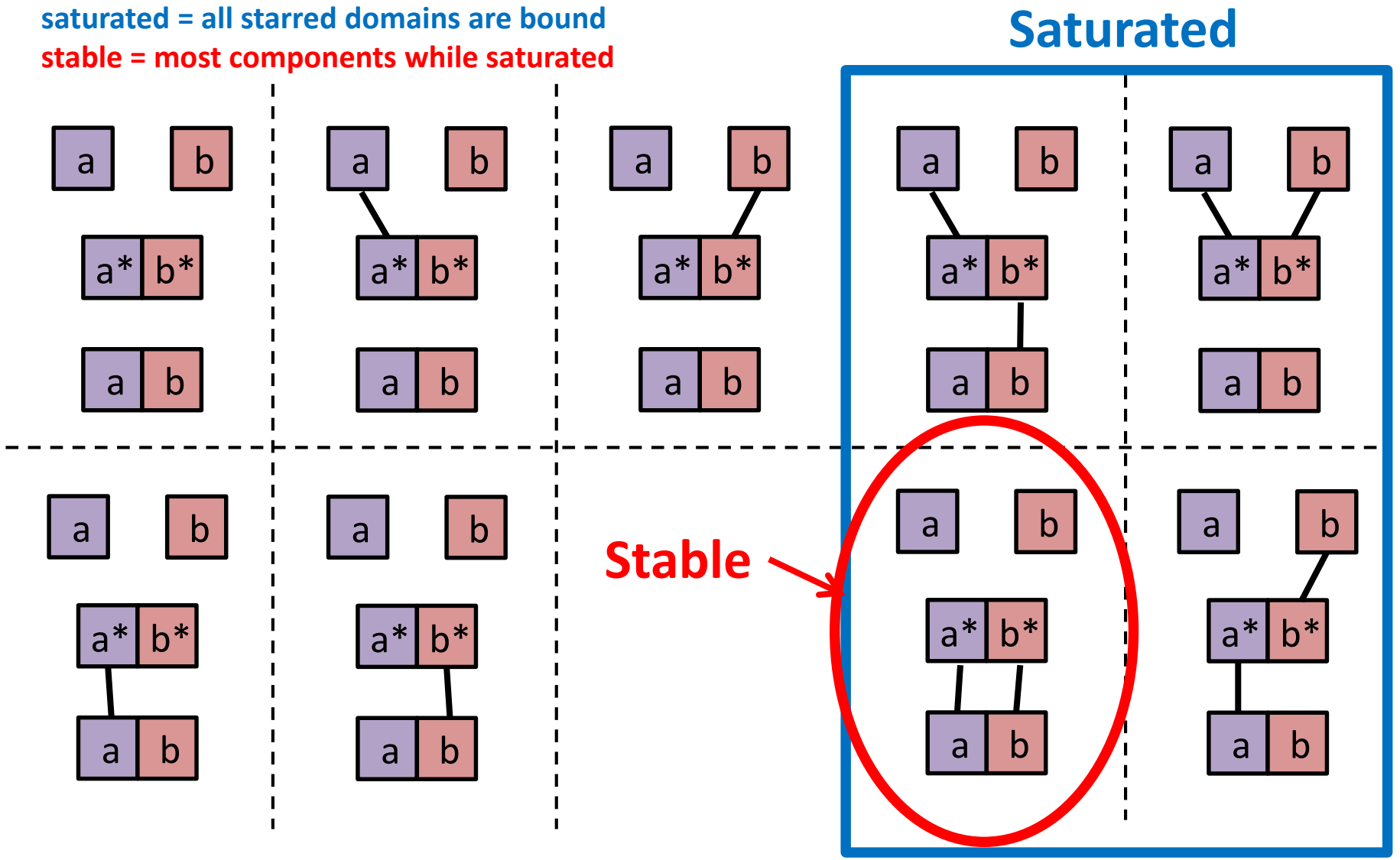
saturated = all starred domains are bound



Thermodynamic Binding Networks

saturated = all starred domains are bound

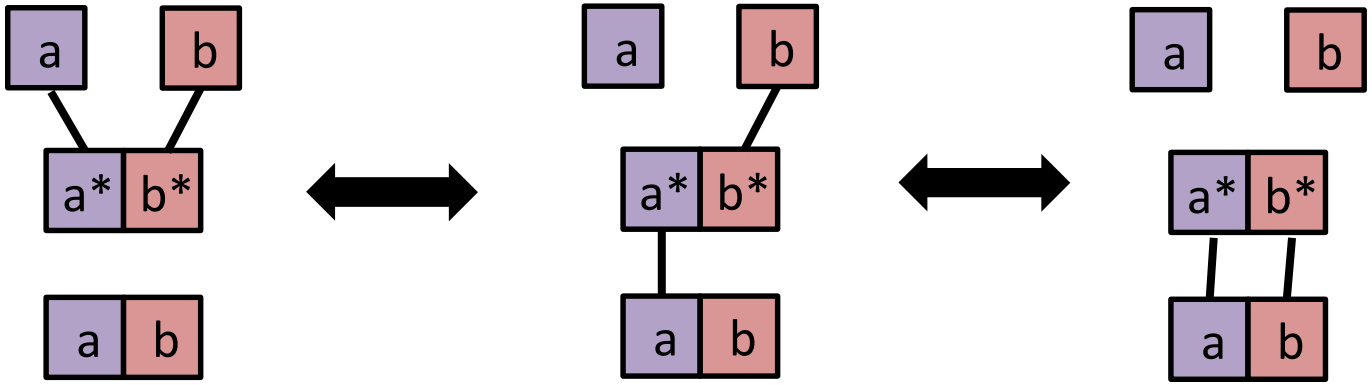
stable = most components while saturated



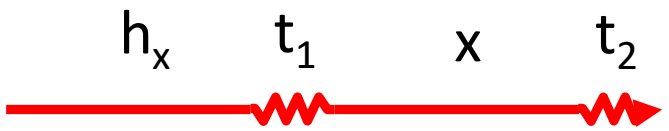
Pathways

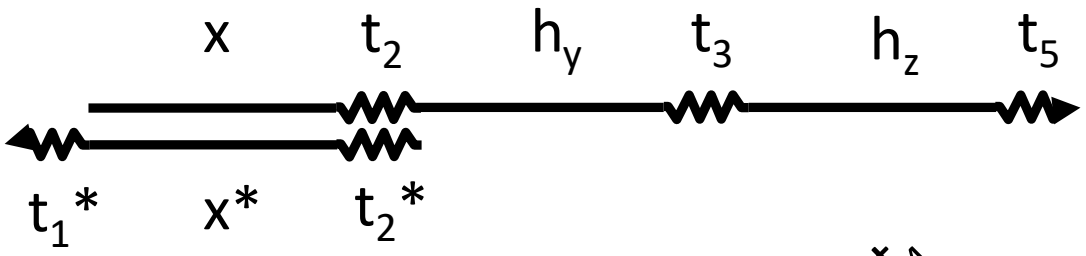
Thermodynamics: Which configurations are energetically favorable

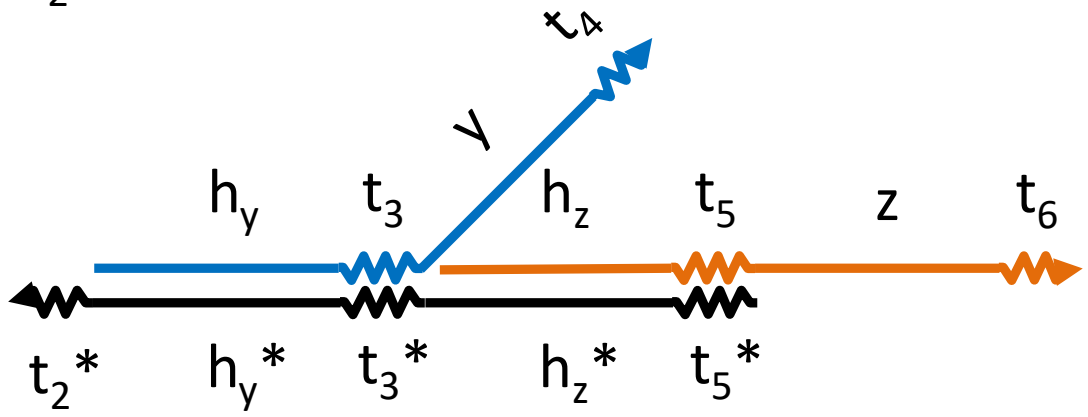
Kinetics: How a system moves between configurations over time



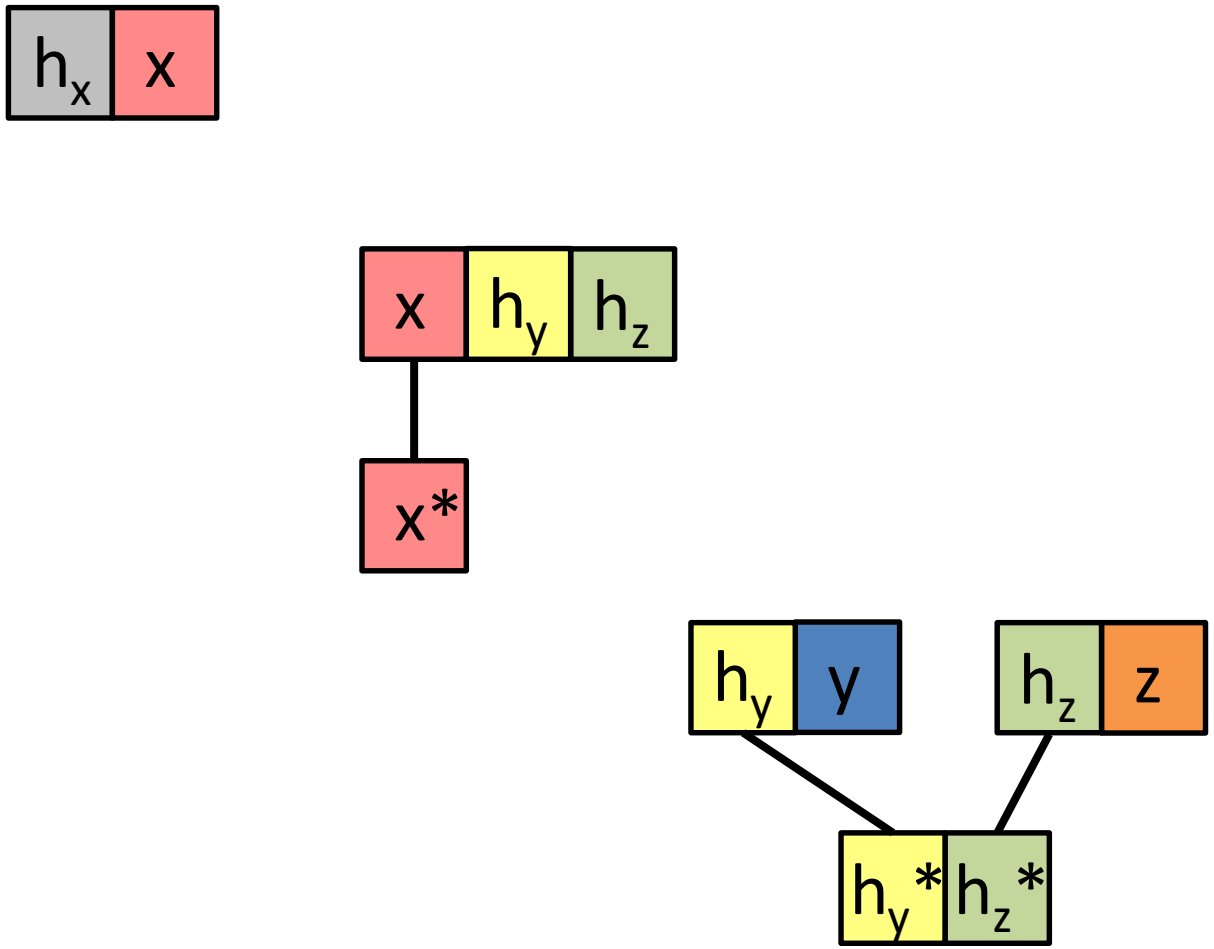
$$X \longrightarrow Y + Z$$

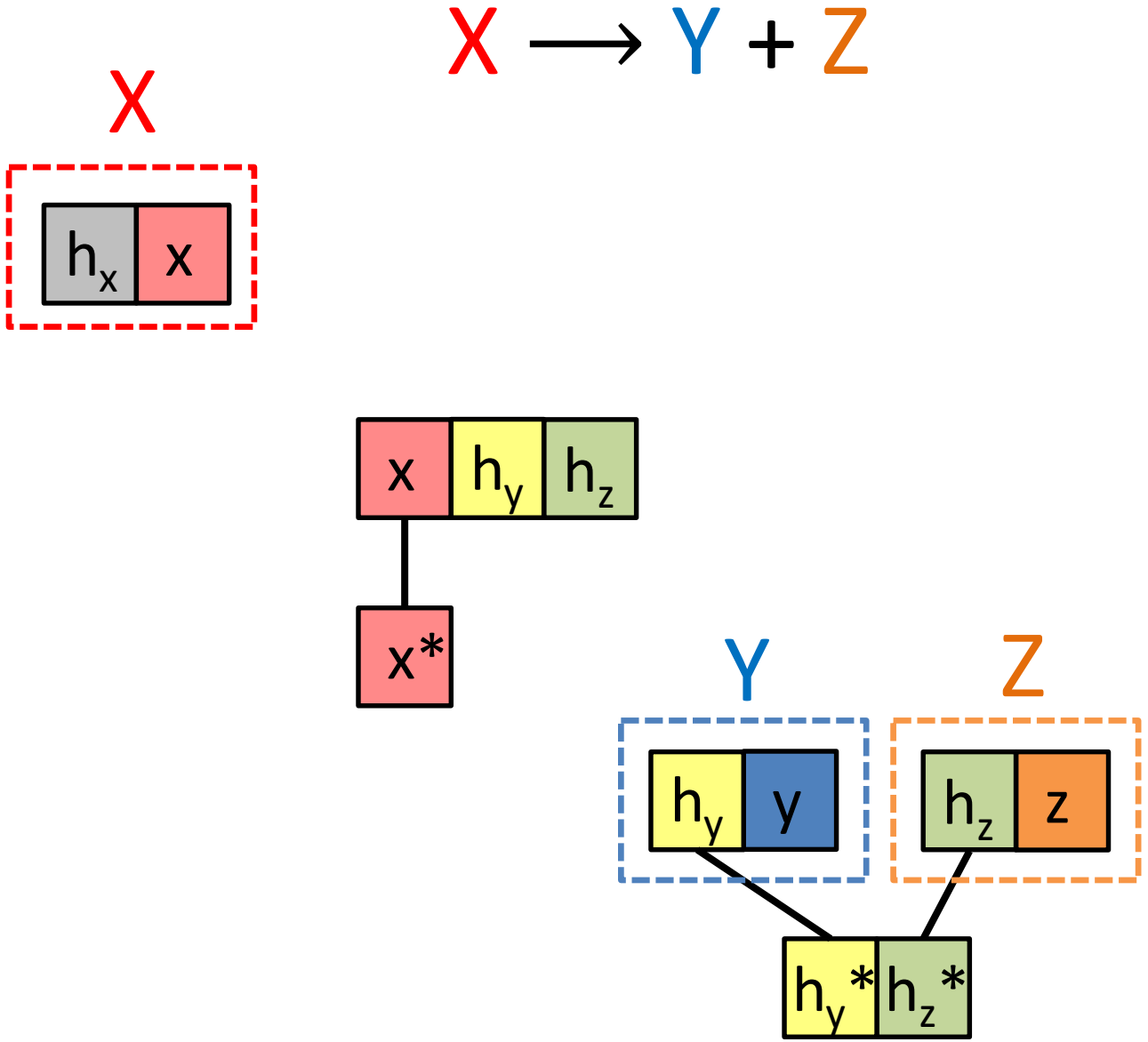




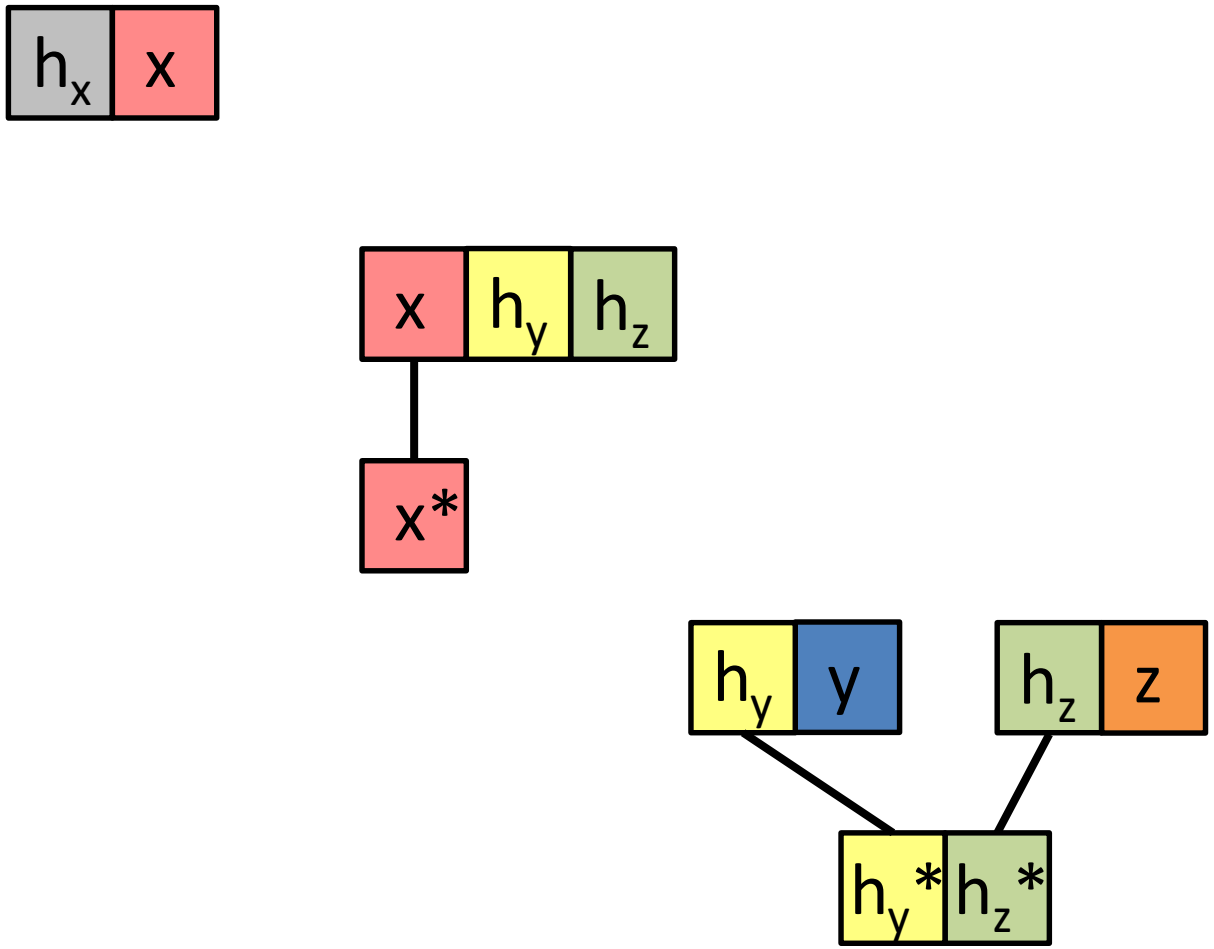


$$X \longrightarrow Y + Z$$

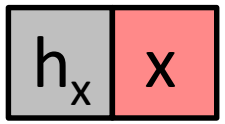


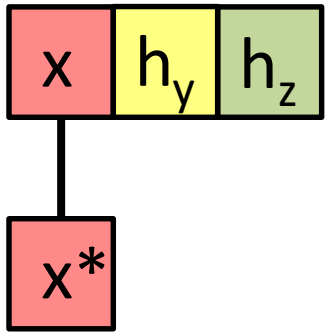

X ⟶ Y + Z
X
h_x
x
x
h_y
h_z
x*
Y
Z
h_y
y
h_z
z
h_y*
h_z*

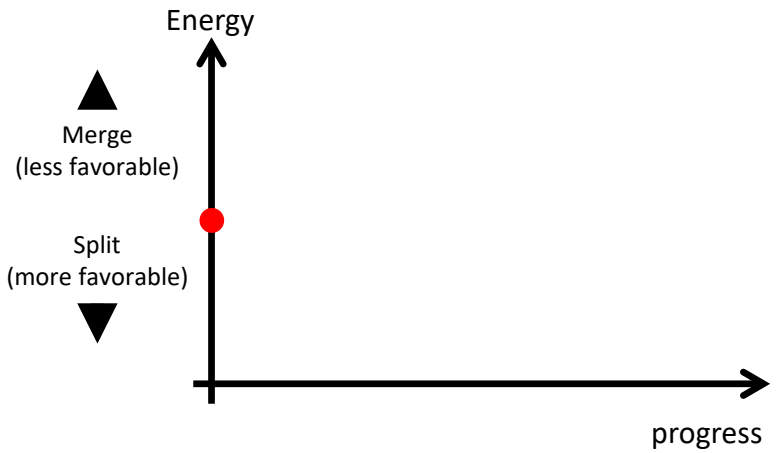
$$X \longrightarrow Y + Z$$

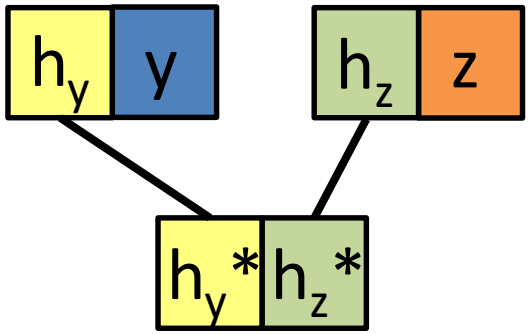


$$X \longrightarrow Y + Z$$

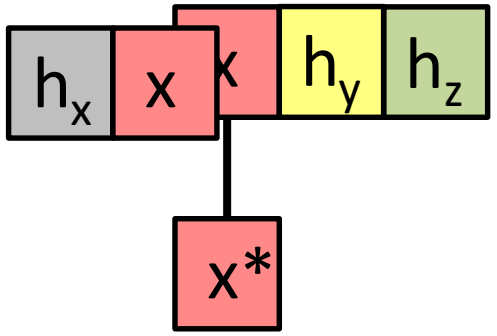


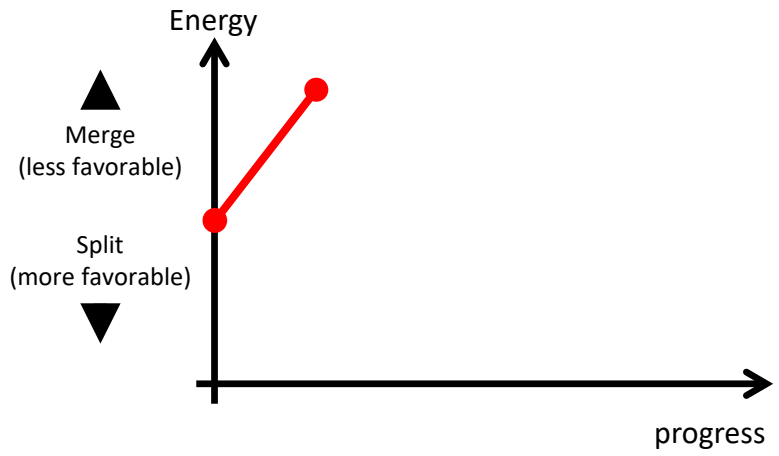


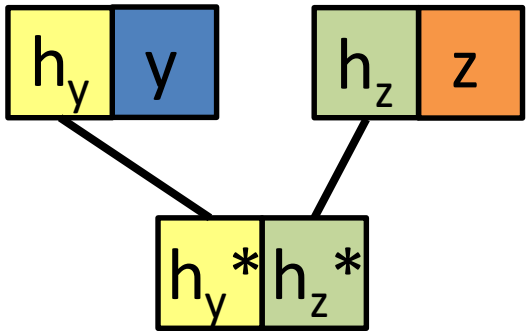




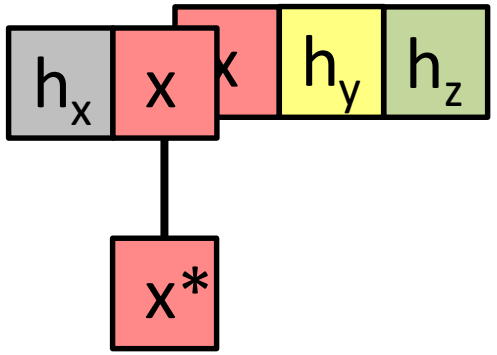
$$X \longrightarrow Y + Z$$

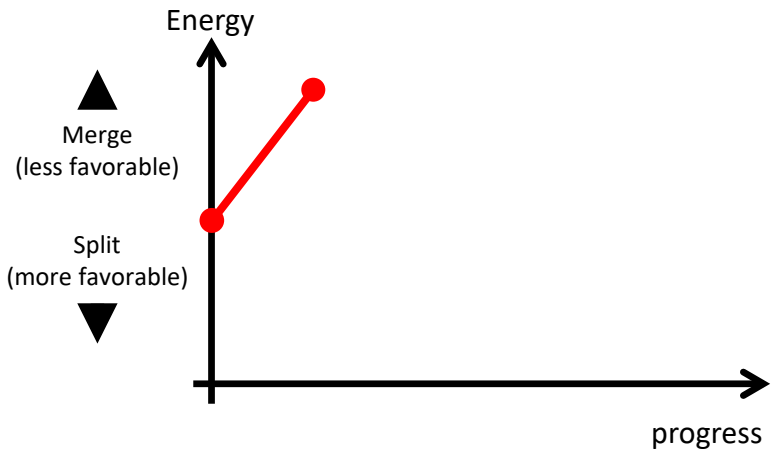


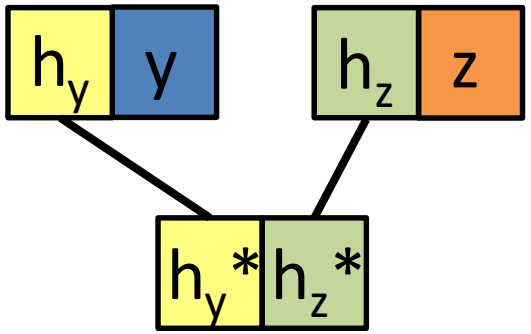




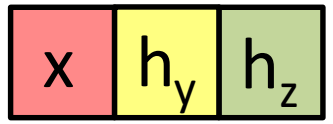
$$X \longrightarrow Y + Z$$

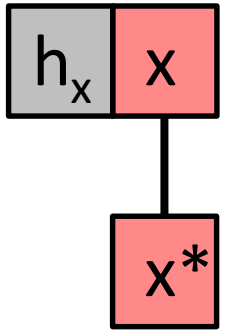


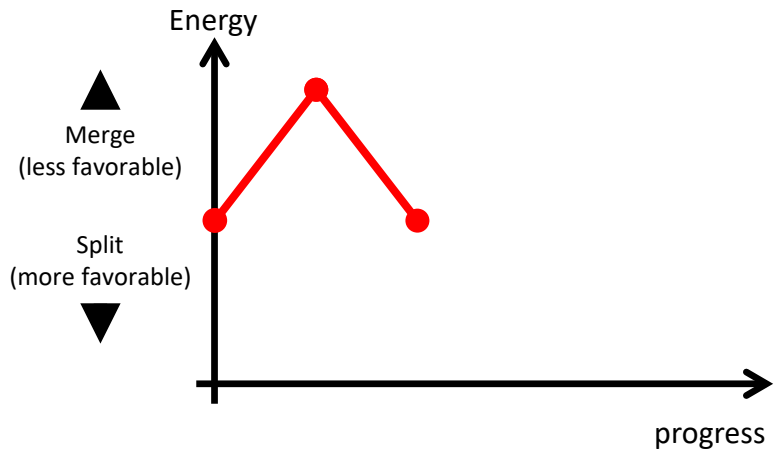


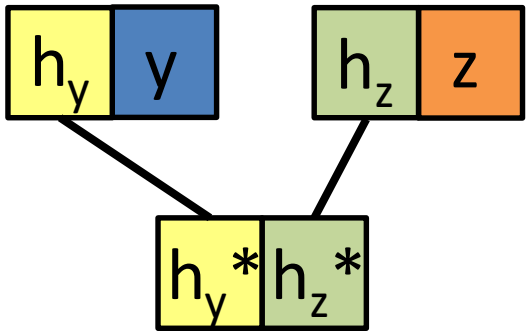


$$X \longrightarrow Y + Z$$

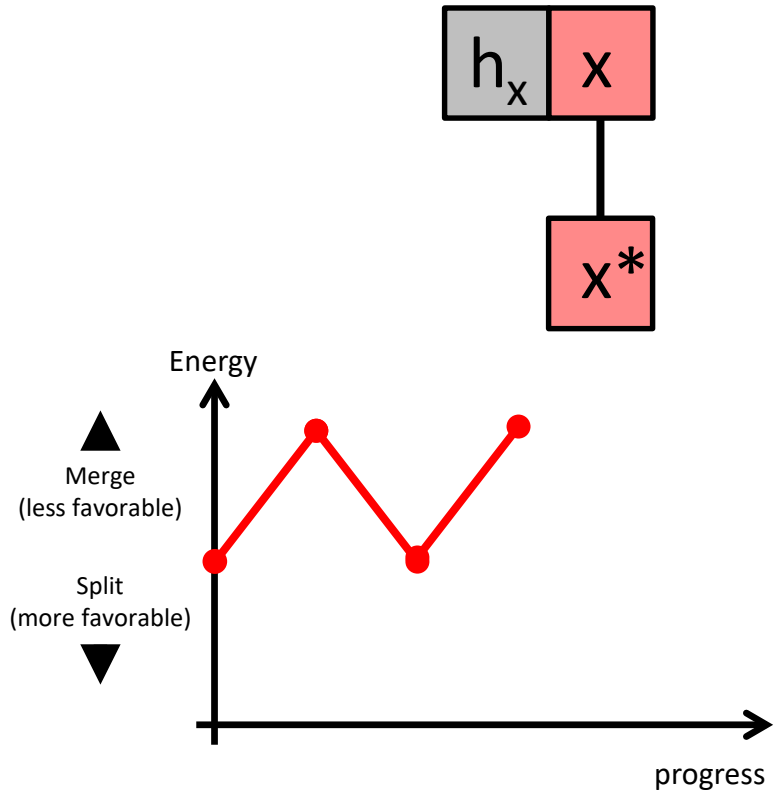


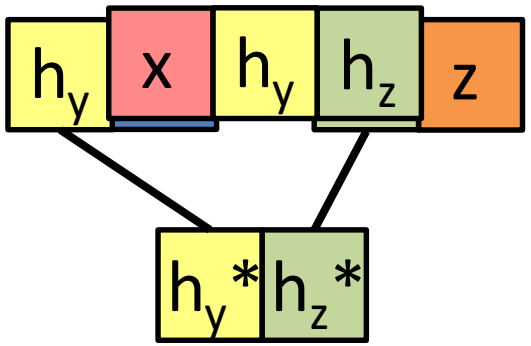




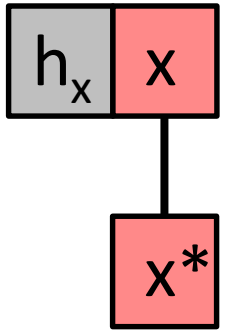


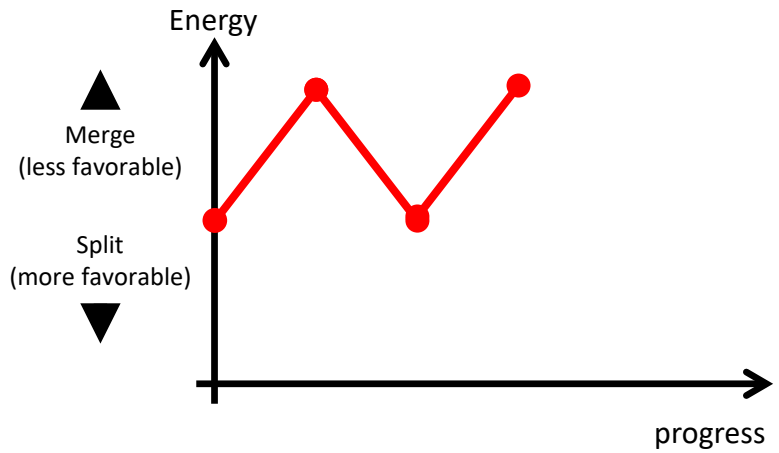
$$X \longrightarrow Y + Z$$

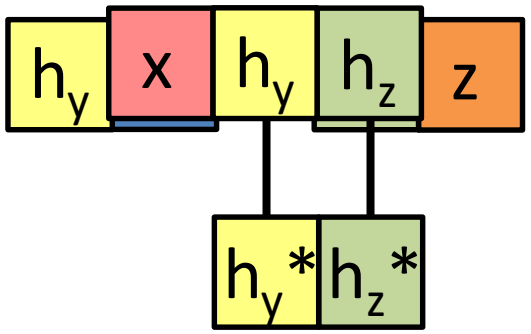




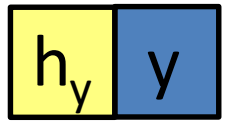
$$X \longrightarrow Y + Z$$

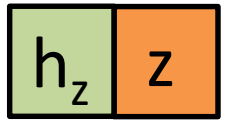


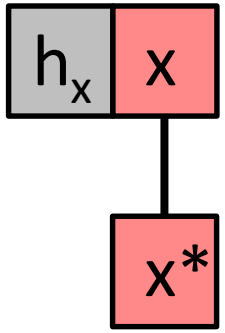


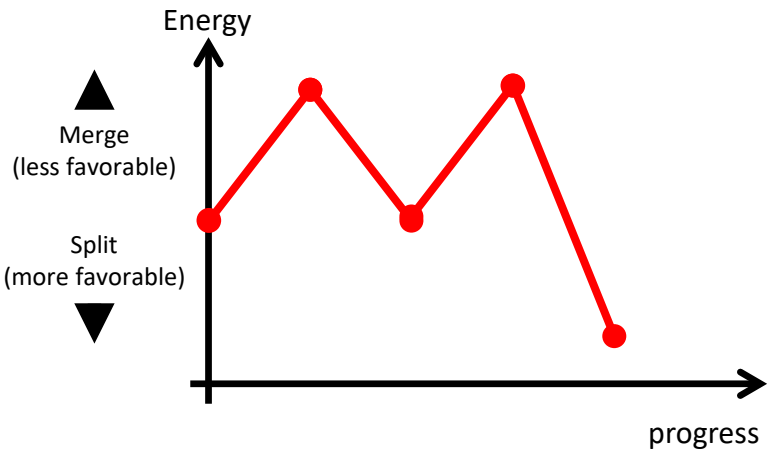


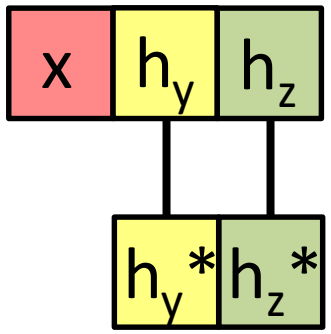
$$X \longrightarrow Y + Z$$





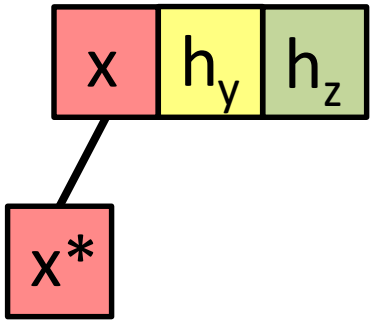


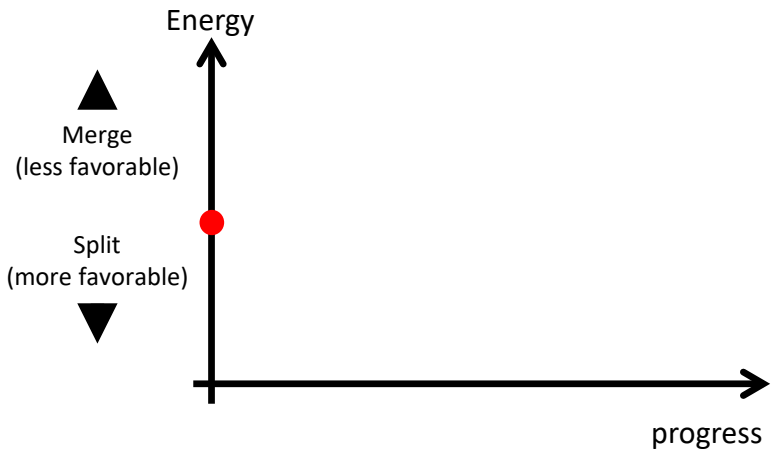


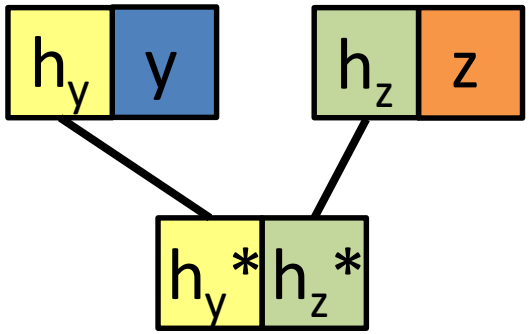


Leak: A Source of Error

$$\emptyset \longrightarrow Y + Z$$

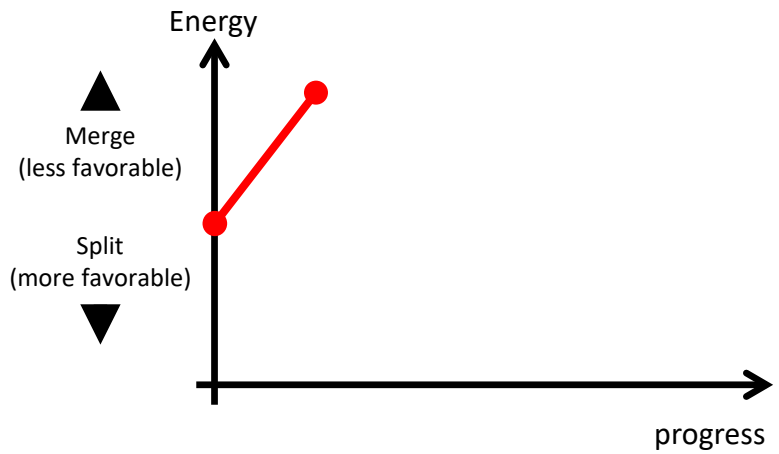


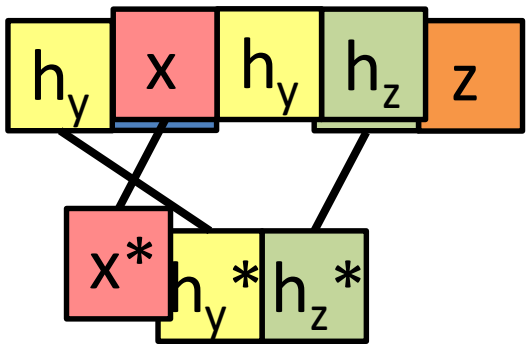




Leak: A Source of Error

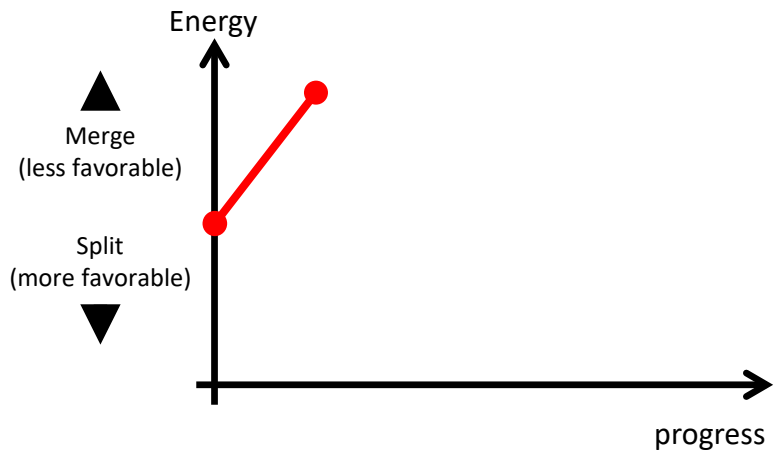
$$\emptyset \longrightarrow Y + Z$$

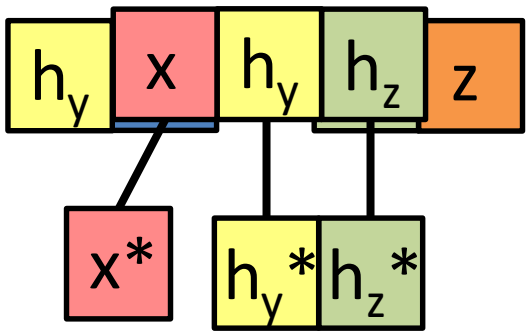




Leak: A Source of Error

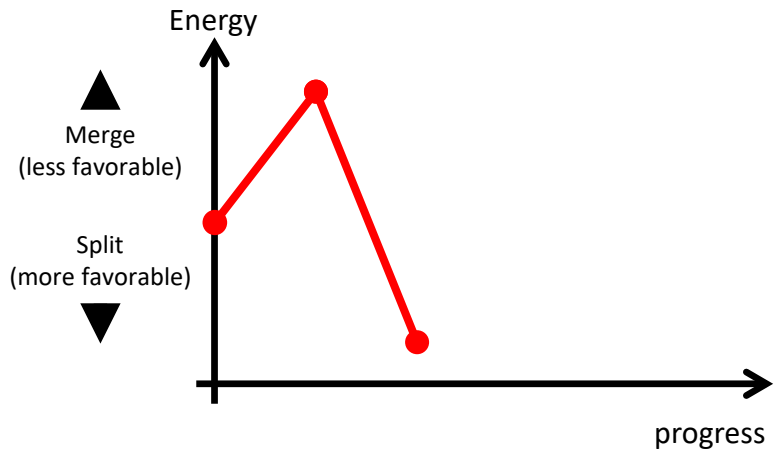
$$\emptyset \longrightarrow Y + Z$$

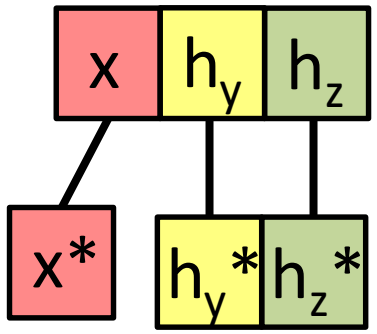




Leak: A Source of Error

$$\emptyset \longrightarrow Y + Z$$





Kinetic Binding Networks

Kinetic Binding Networks

- Favorability is a combination of bond count and component count

Kinetic Binding Networks

- Favorability is a combination of bond count and component count

<u>Weighted average:</u>

Energy := $-w_H(\#\text{ bonds}) - (\#\text{ components})$

Kinetic Binding Networks

- Favorability is a combination of bond count and component count

Weighted average:

Energy := $-w_H(\#\text{ bonds}) - (\#\text{ components})$

- Define pathways to consist of merges and splits

Kinetic Binding Networks

- Favorability is a combination of bond count and component count

Weighted average:

Energy := $-w_H(\#\text{ bonds}) - (\#\text{ components})$

- Define pathways to consist of merges and splits
- But for $w_H \geq 2$, only saturated pathways need be considered[2]

[2]Keenan Breik, Cameron Chalk, David Doty, David Haley, David Soloveichik. *Programming Substrate-Independent Kinetic Barriers with Thermodynamic Binding Networks*. CMSB 2018

Kinetic Binding Networks

- Favorability is a combination of bond count and component count

Weighted average:

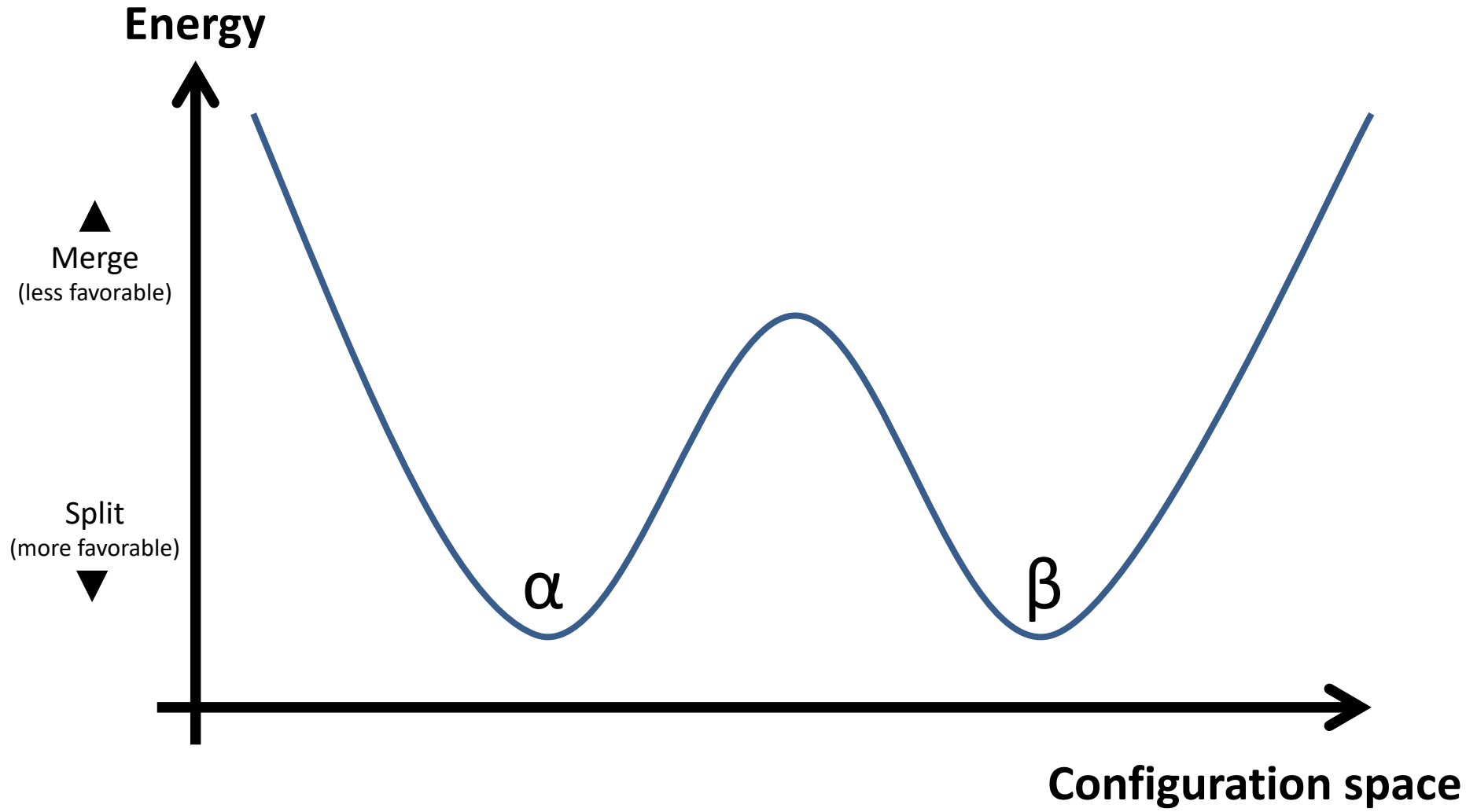
$$\text{Energy} := -w_H(\#\text{ bonds}) - (\#\text{ components})$$

- Define pathways to consist of merges and splits
- But for $w_H \geq 2$, only saturated pathways need be considered[2]

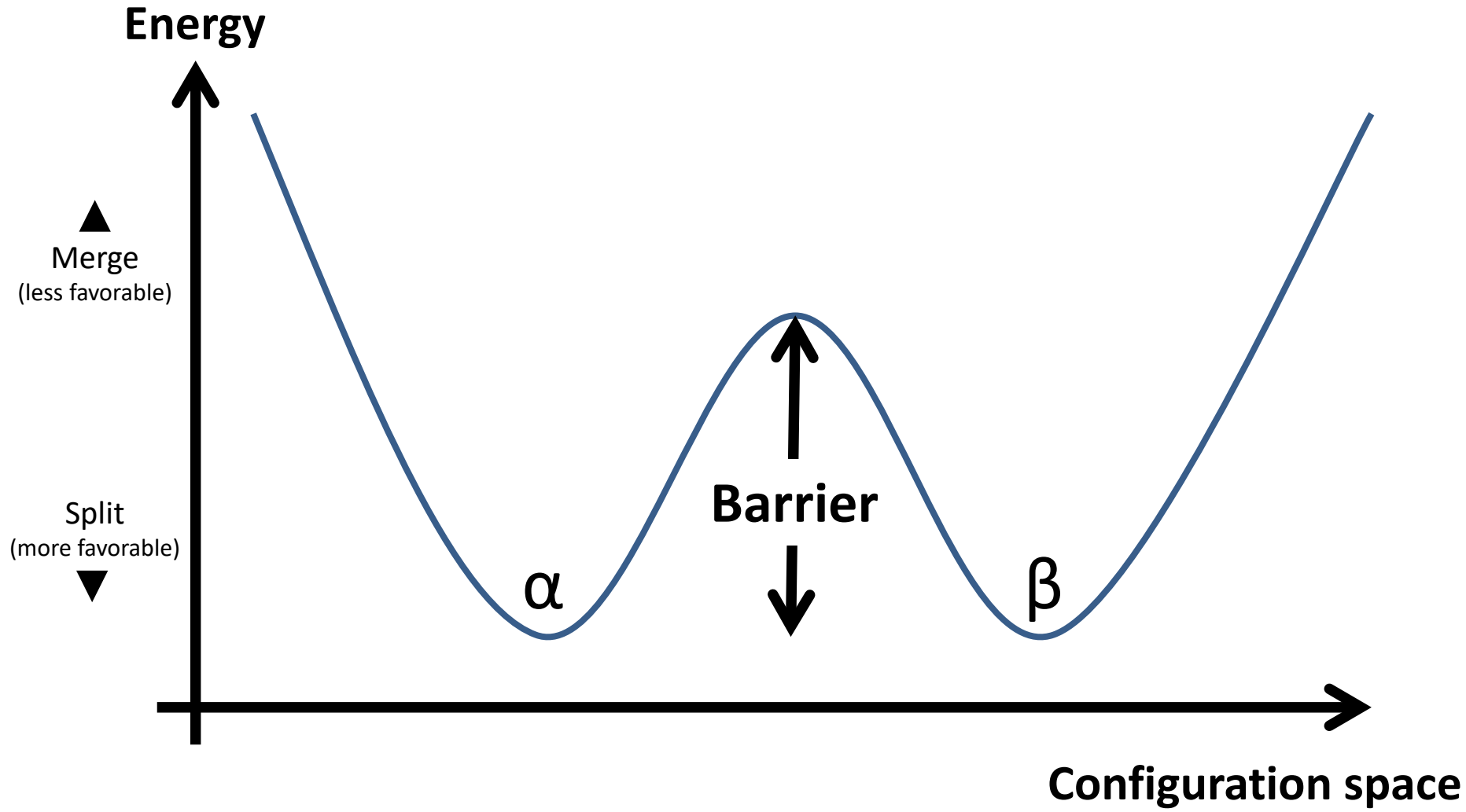
Since all saturated configurations have an equal number of bonds, we can focus solely on the number of components

[2]Keenan Breik, Cameron Chalk, David Doty, David Haley, David Soloveichik. *Programming Substrate-Independent Kinetic Barriers with Thermodynamic Binding Networks*. CMSB 2018

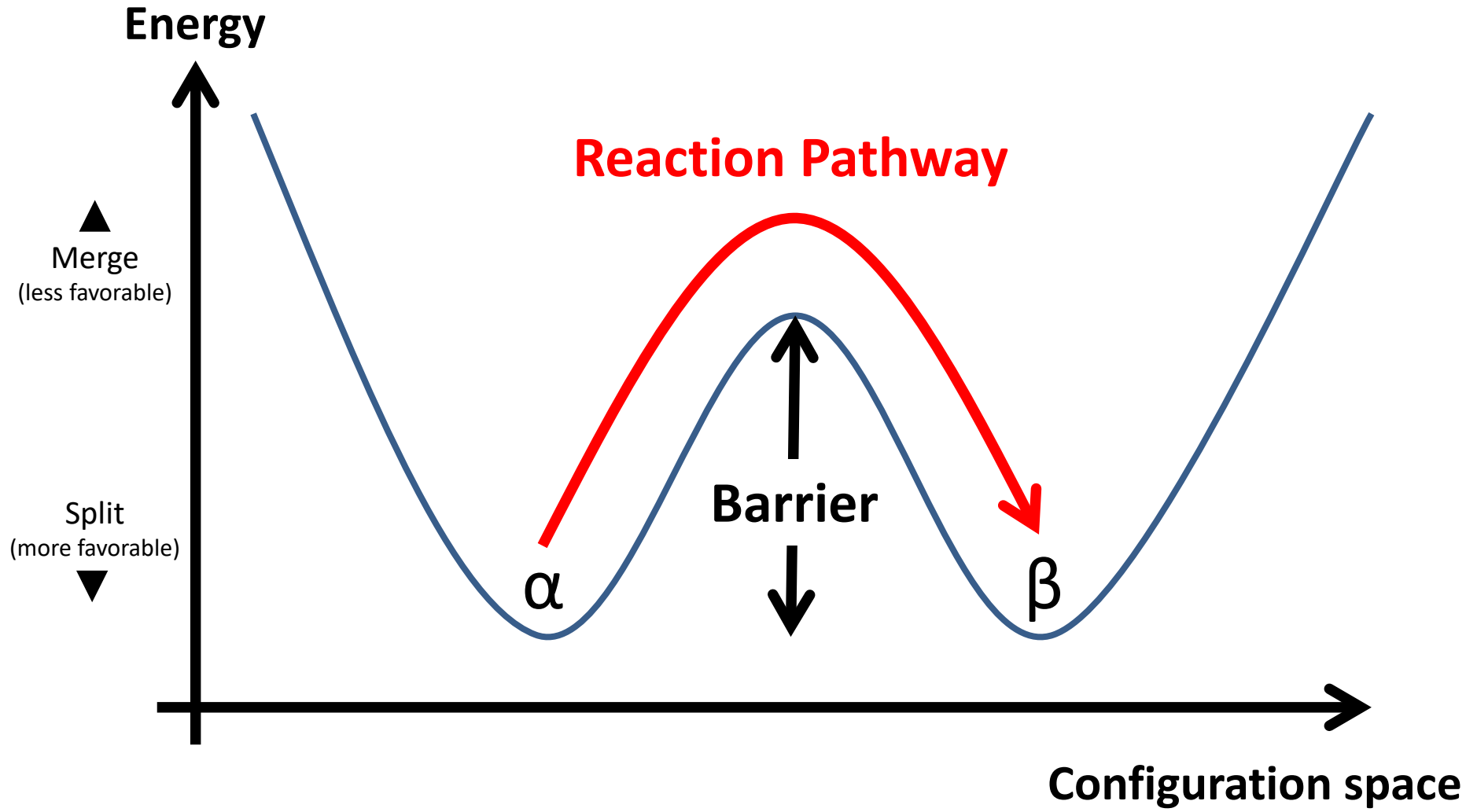
Large Energy Barriers



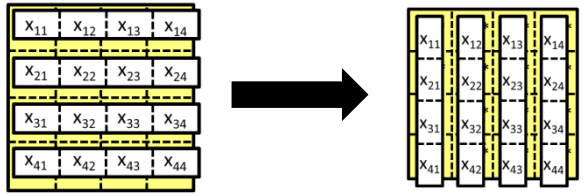
Large Energy Barriers

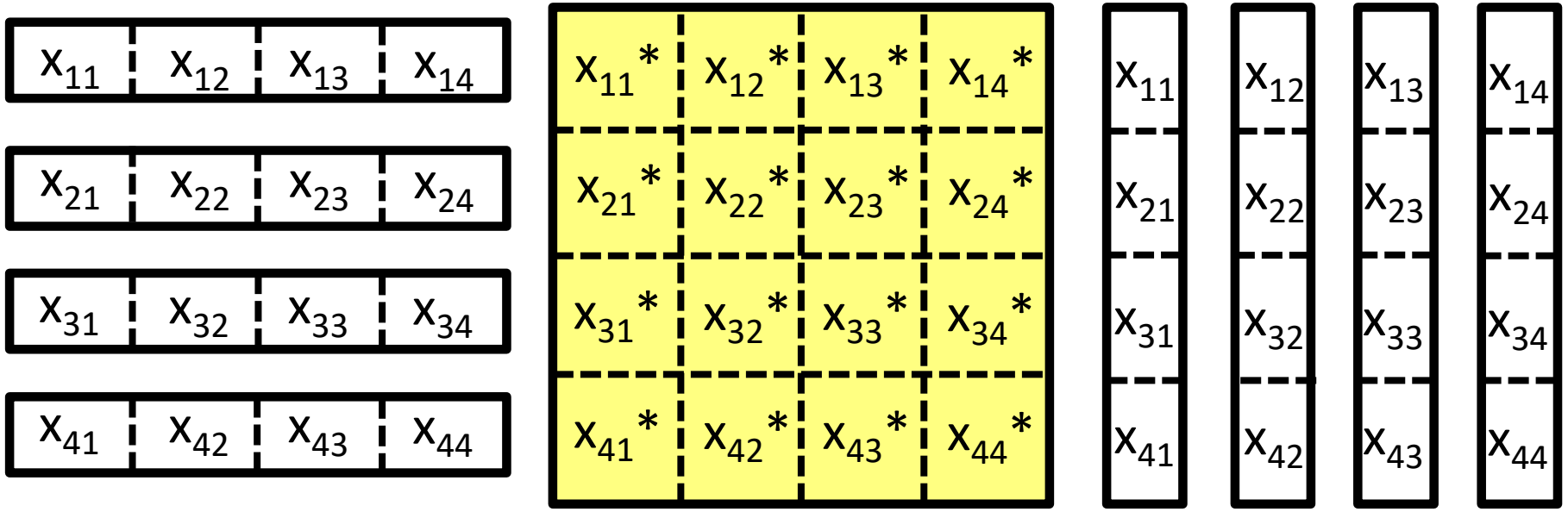


Large Energy Barriers

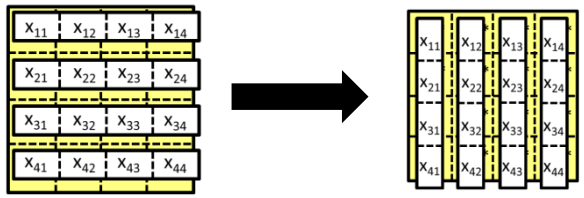


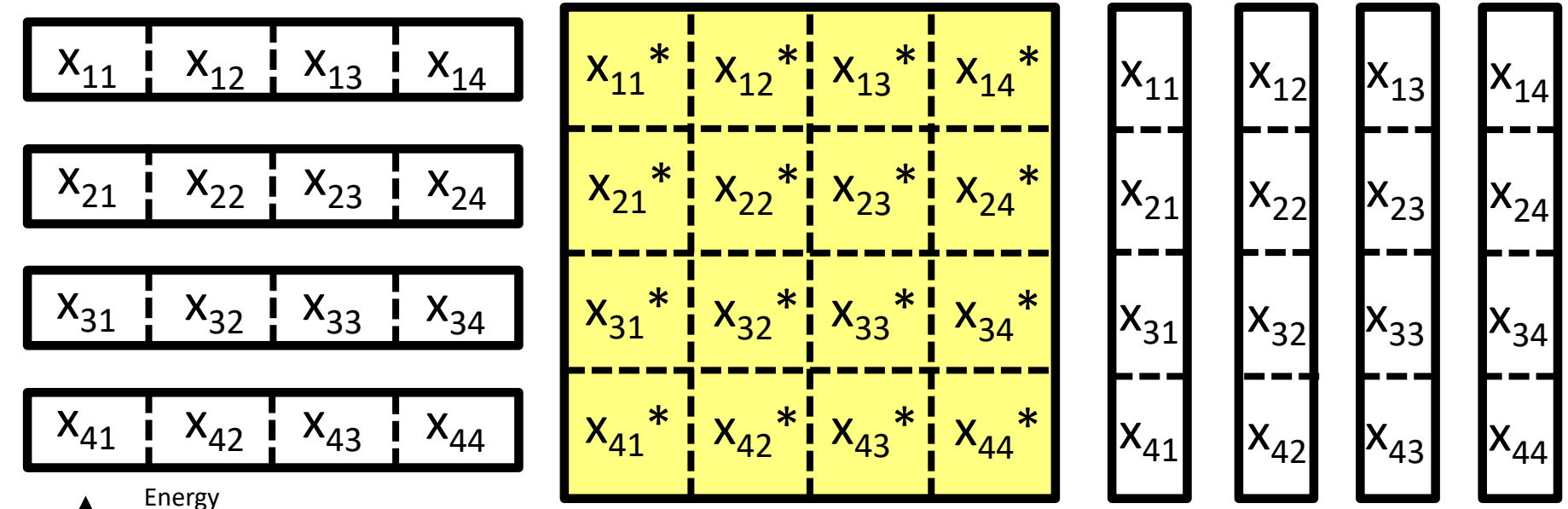
A Network with a Programmable Energy Barrier

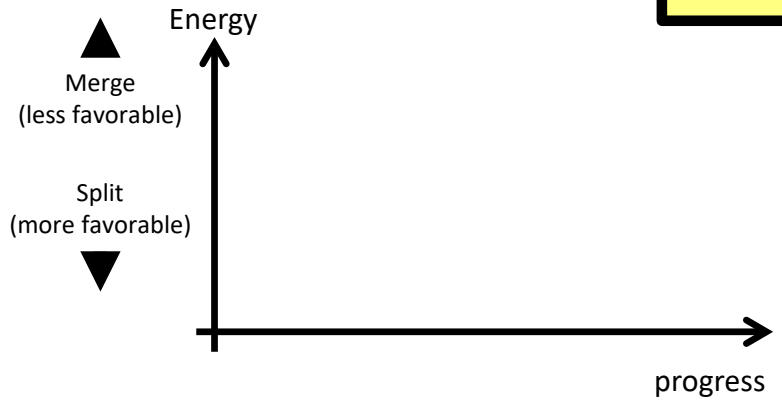




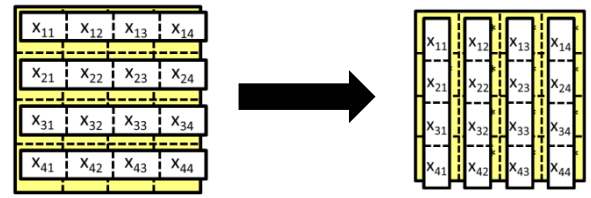
A Network with a Programmable Energy Barrier

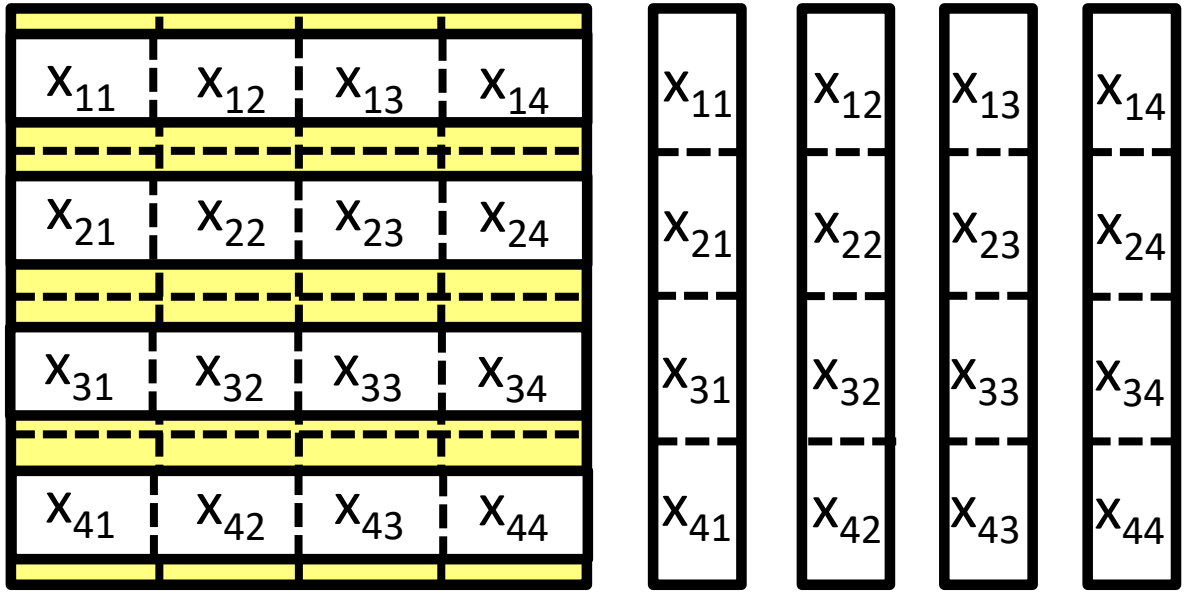


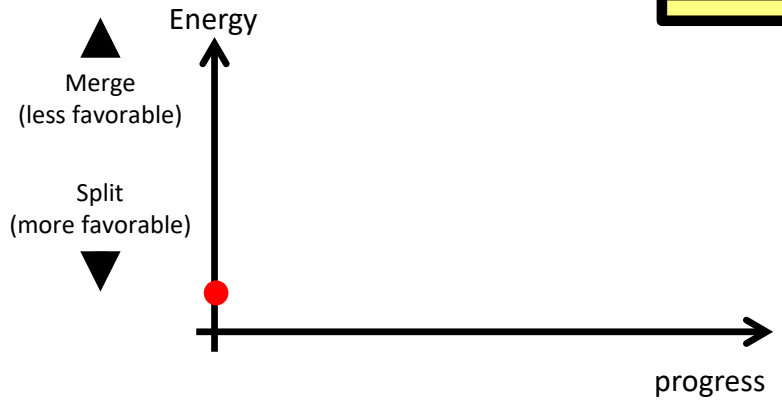




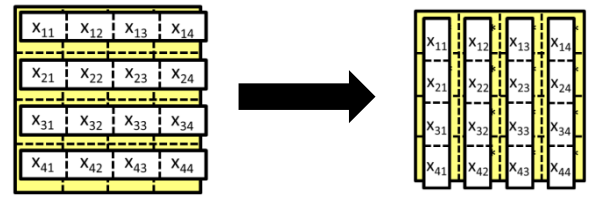
A Network with a Programmable Energy Barrier

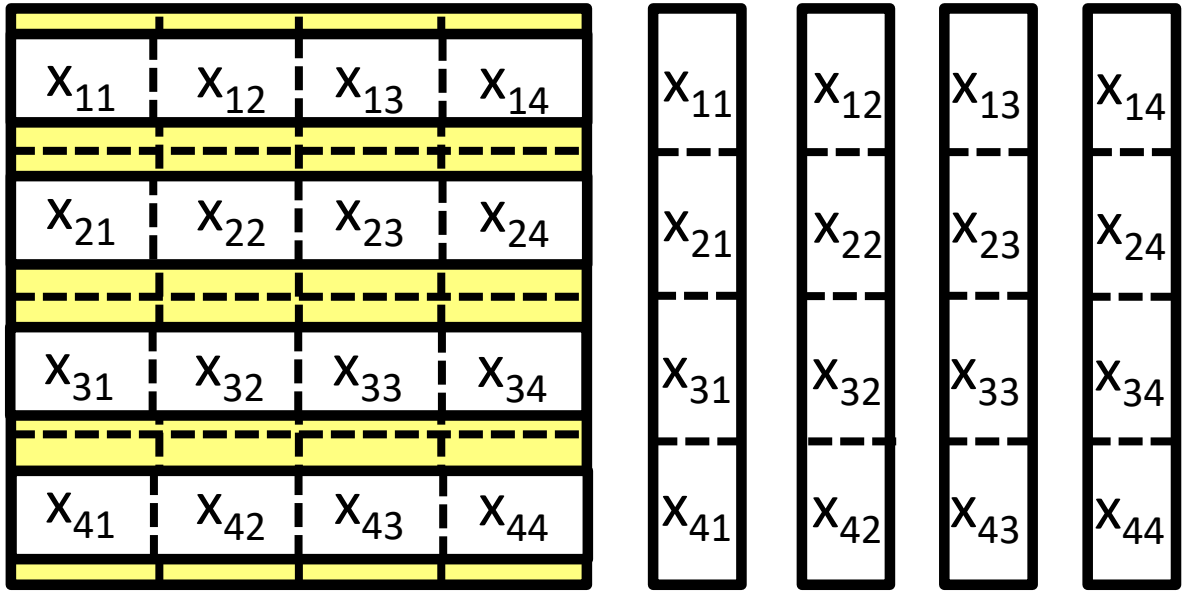


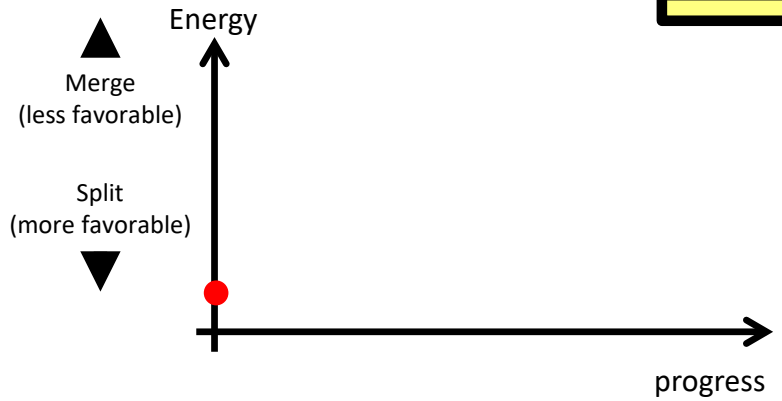




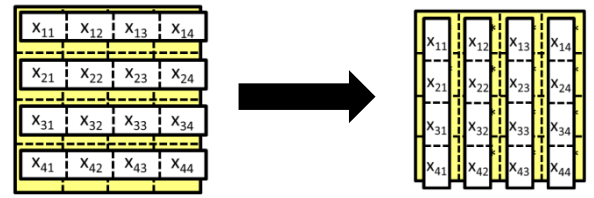
A Network with a Programmable Energy Barrier

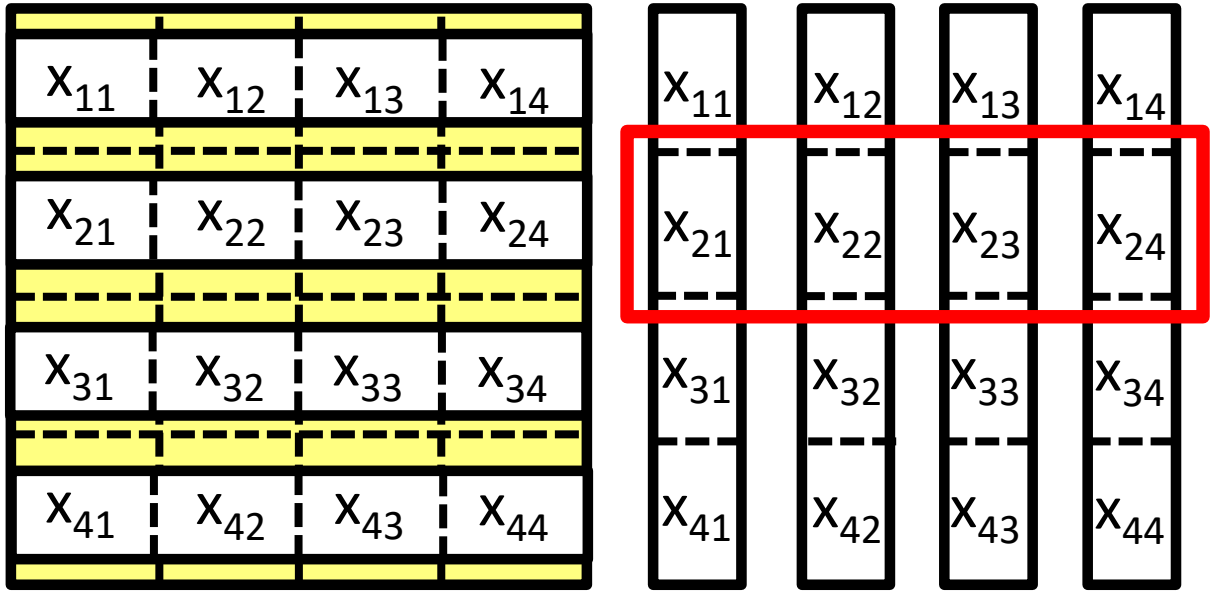


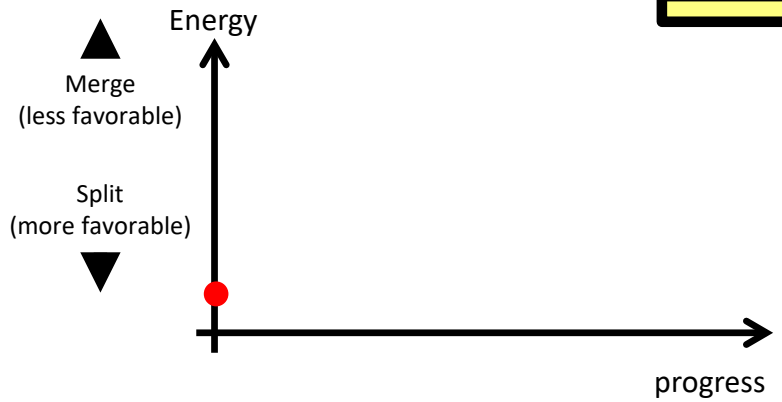




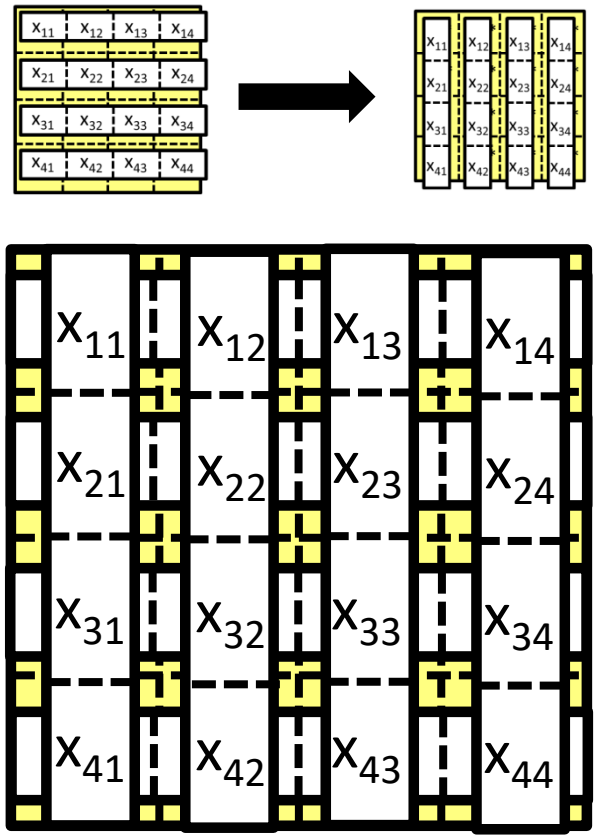
A Network with a Programmable Energy Barrier

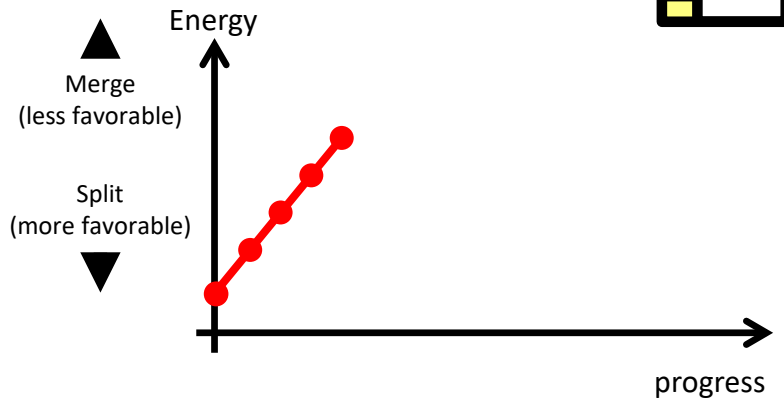




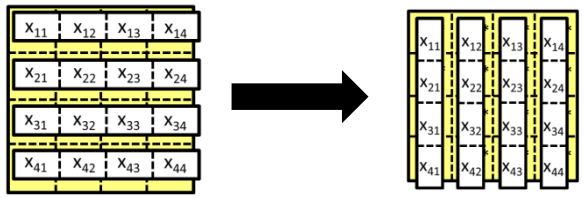


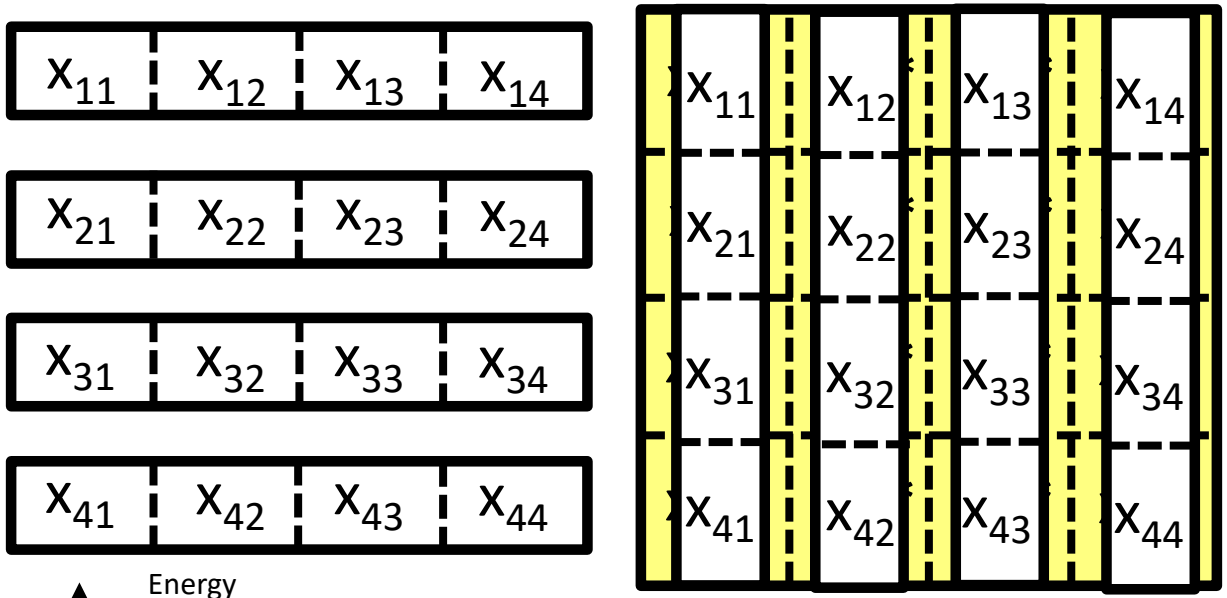
A Network with a Programmable Energy Barrier



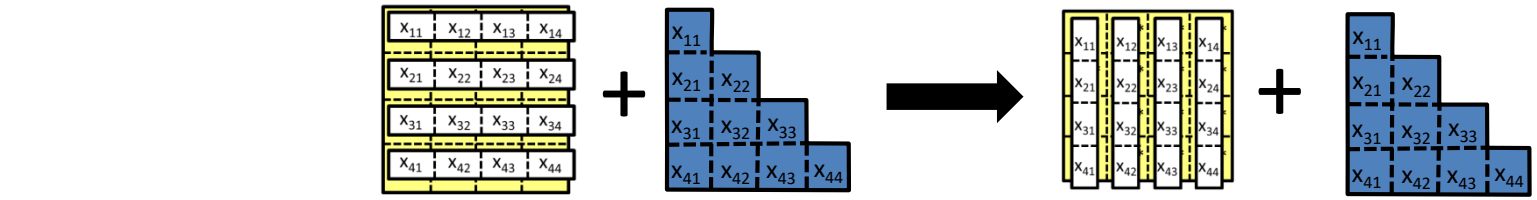


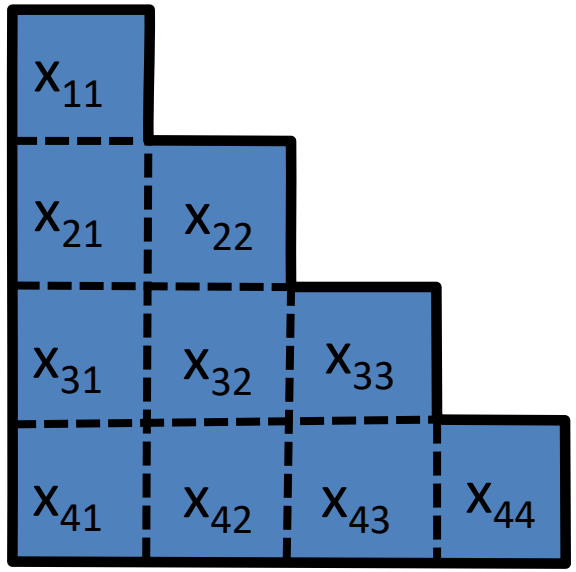
A Network with a Programmable Energy Barrier

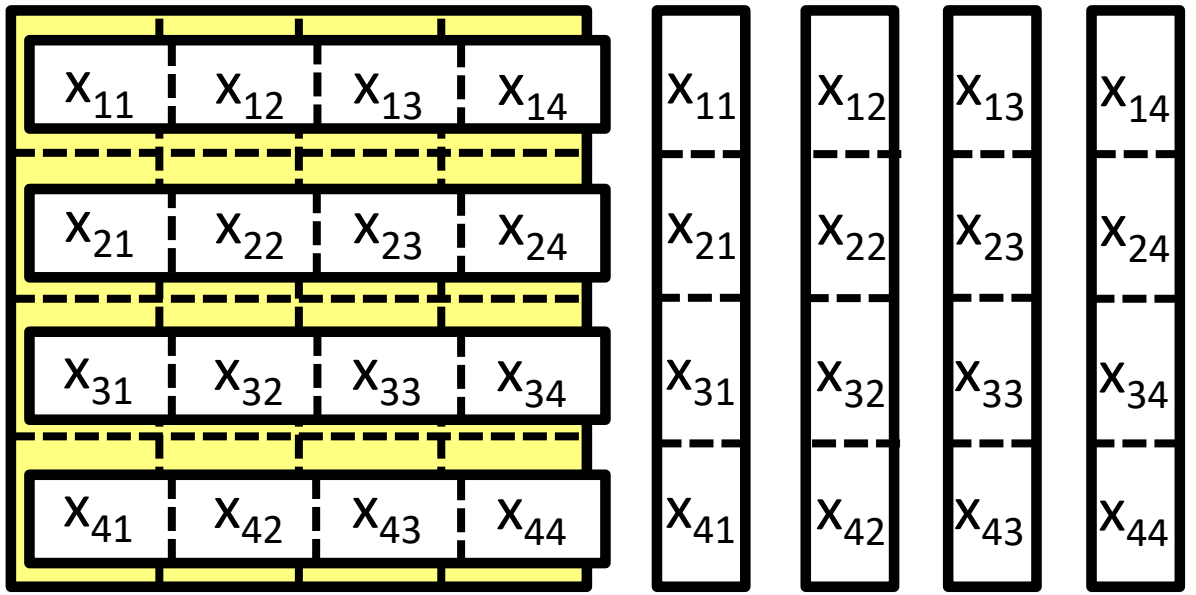


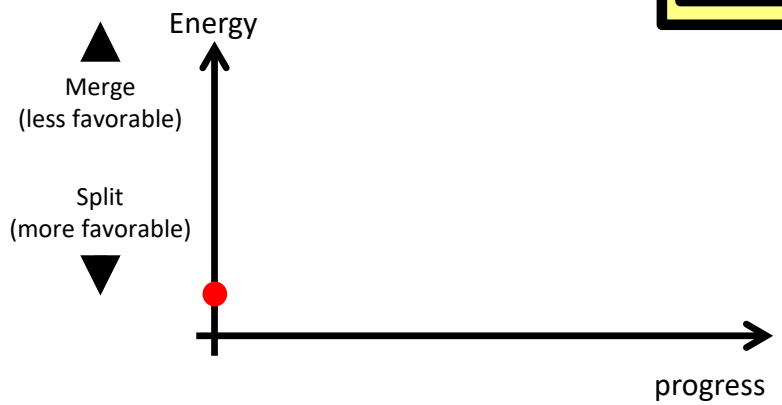


Catalysis

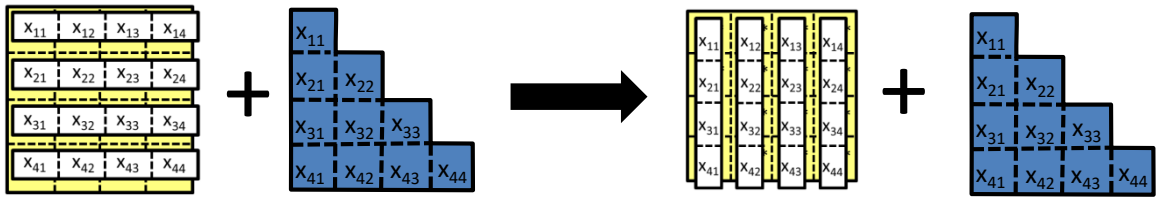

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

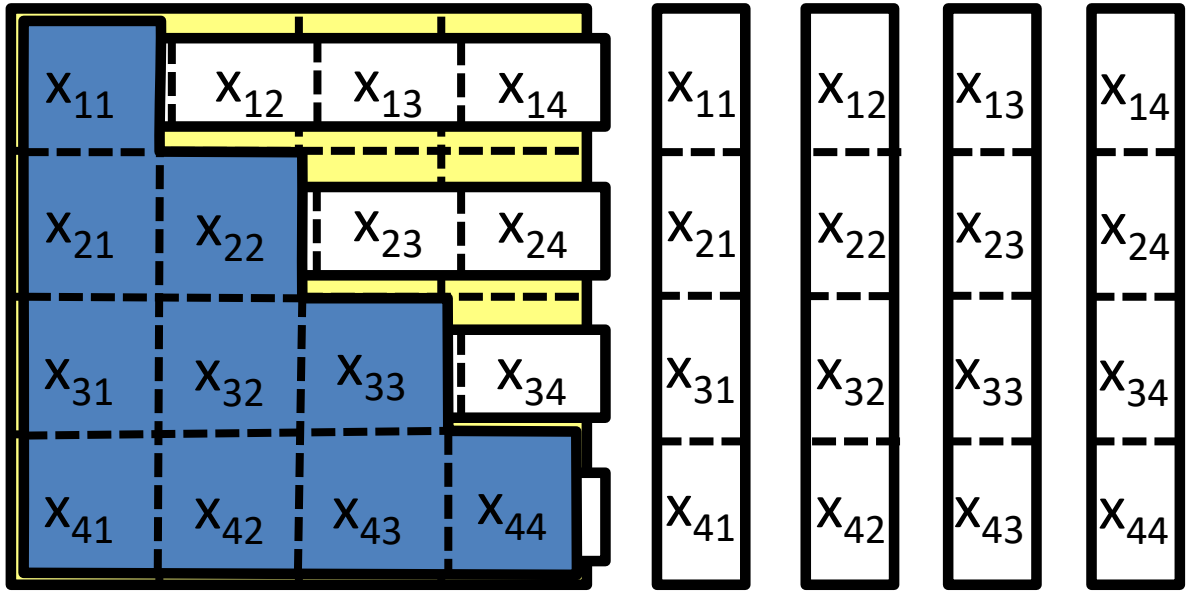

x11
x21 x22
x31 x32 x33
x41 x42 x43 x44

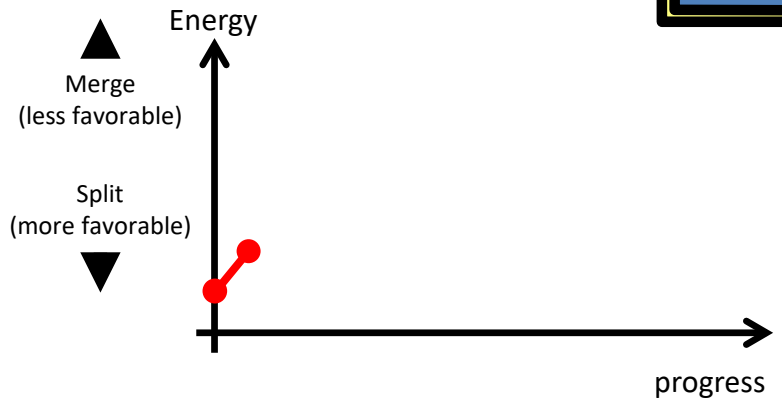

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x21 x31 x41
x12 x22 x32 x42
x13 x23 x33 x43
x14 x24 x34 x44


Energy
Merge
(less favorable)
Split
(more favorable)
progress

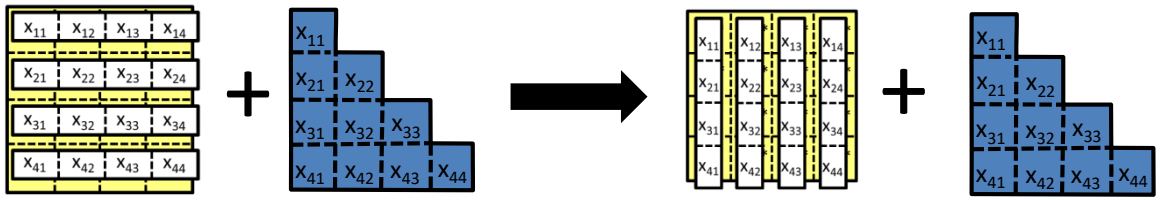
Catalysis

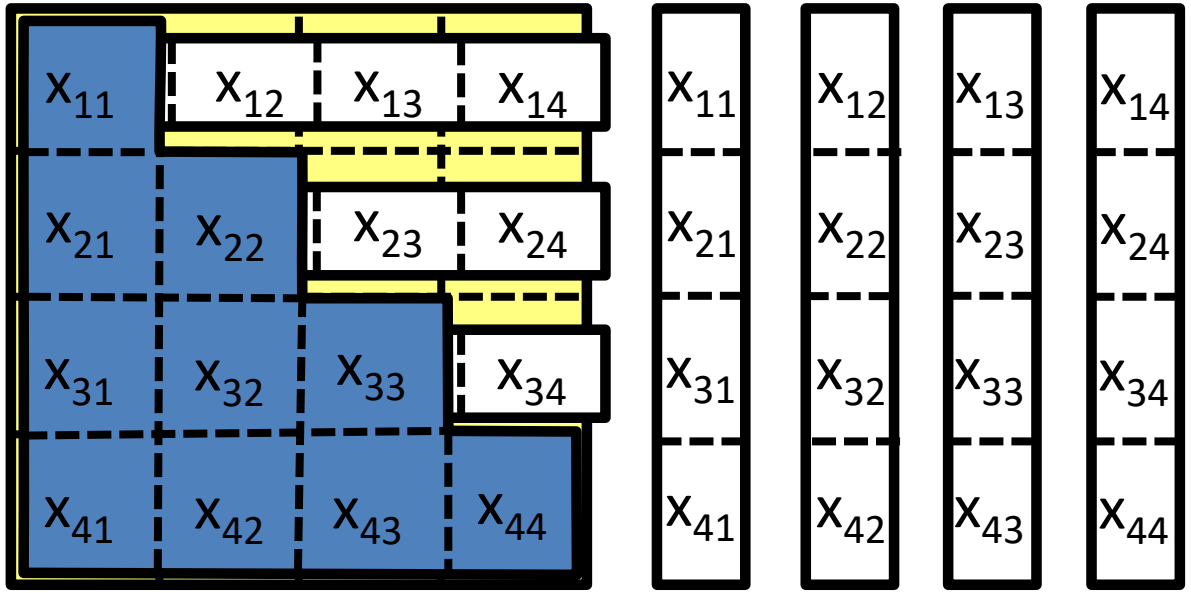

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

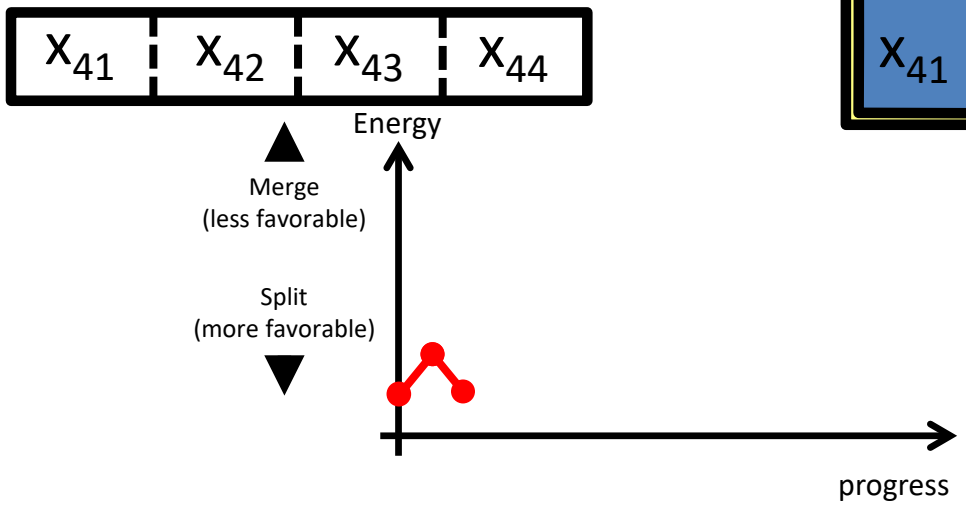

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x21 x31 x41
x12 x22 x32 x42
x13 x23 x33 x43
x14 x24 x34 x44


Energy
Merge
(less favorable)
Split
(more favorable)
progress

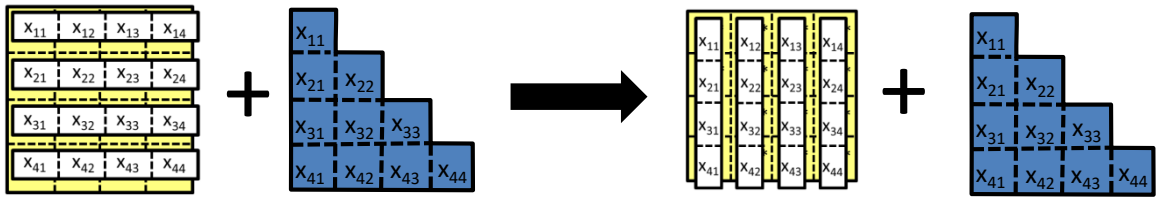
Catalysis

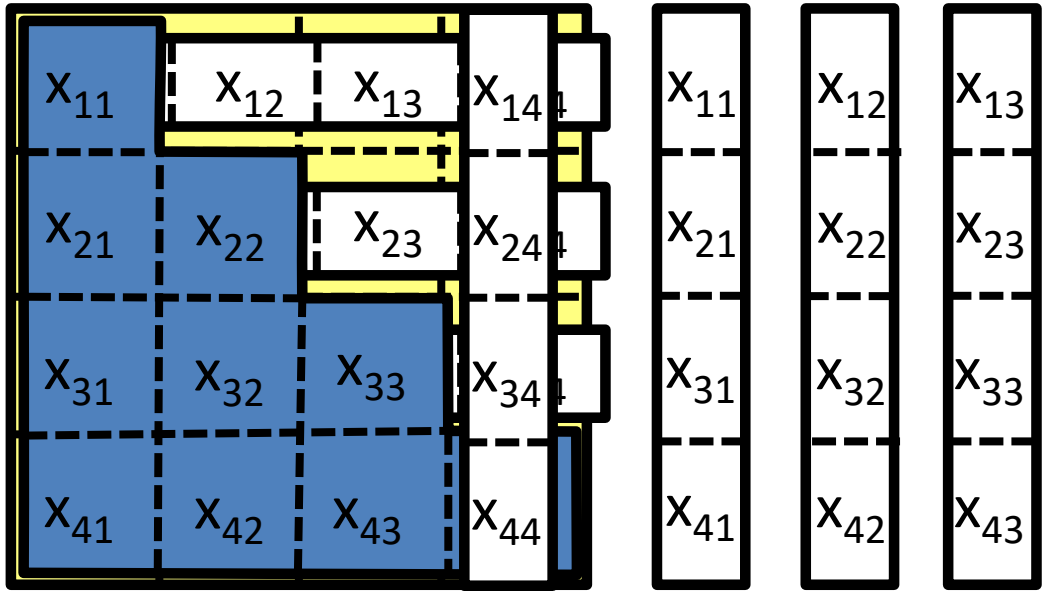

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

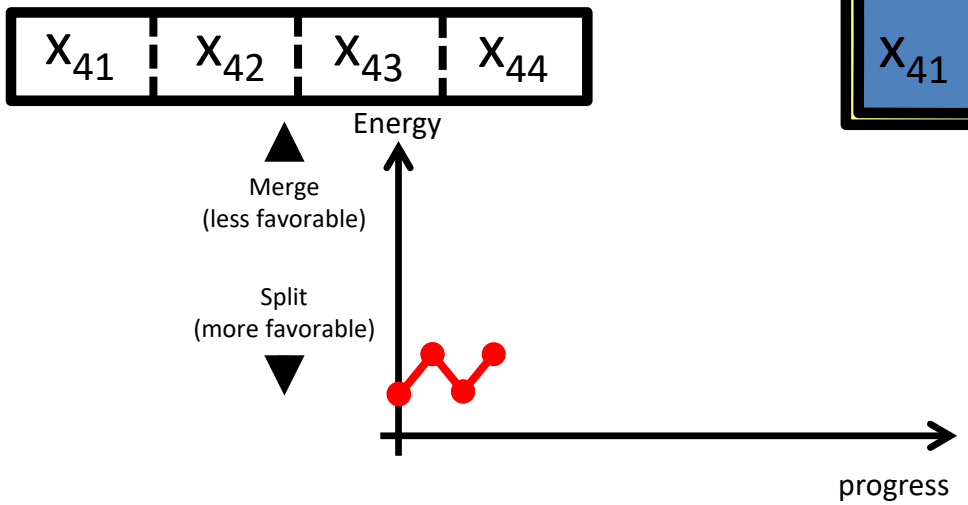

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x21 x31 x41
x12 x22 x32 x42
x13 x23 x33 x43
x14 x24 x34 x44


x41 x42 x43 x44
Energy
Merge
(less favorable)
Split
(more favorable)
progress

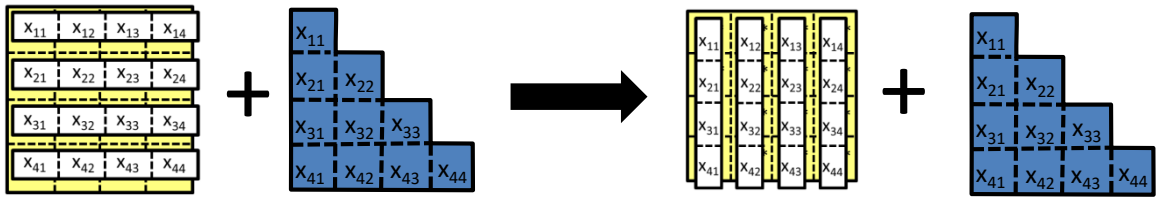
Catalysis

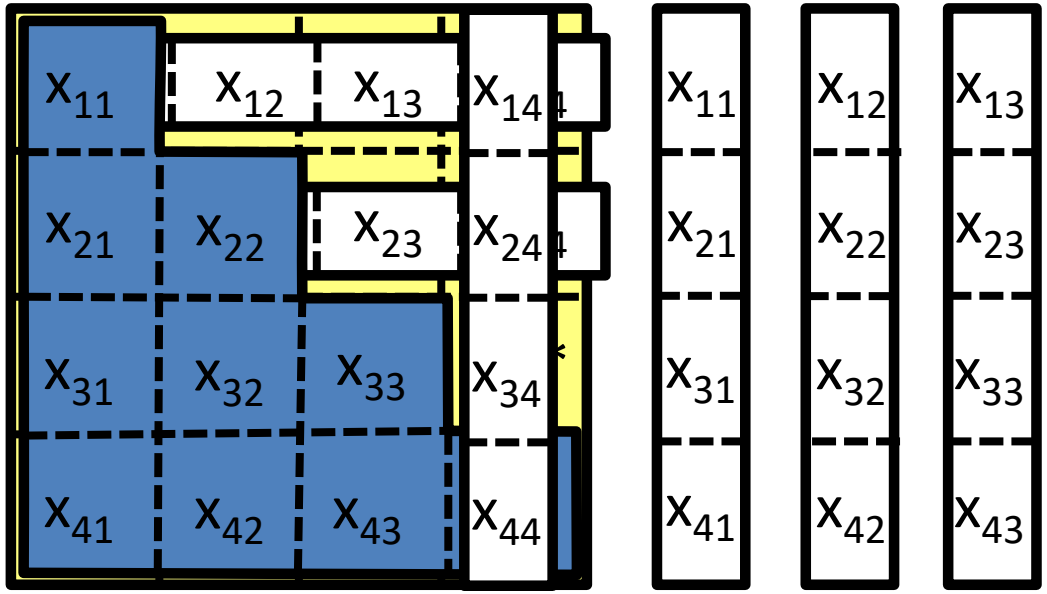

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

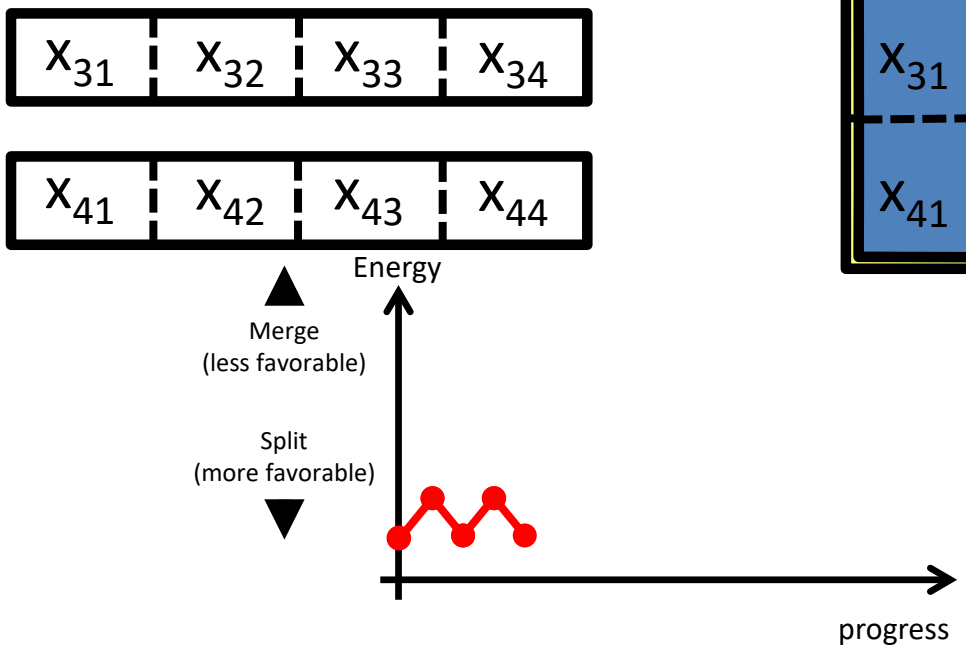

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x12 x13
x21 x22 x23
x31 x32 x33
x41 x42 x43
x44
x11 x21 x31 x41
x12 x22 x32 x42
x13 x23 x33 x43


x41 x42 x43 x44
Energy
Merge
(less favorable)
Split
(more favorable)
progress

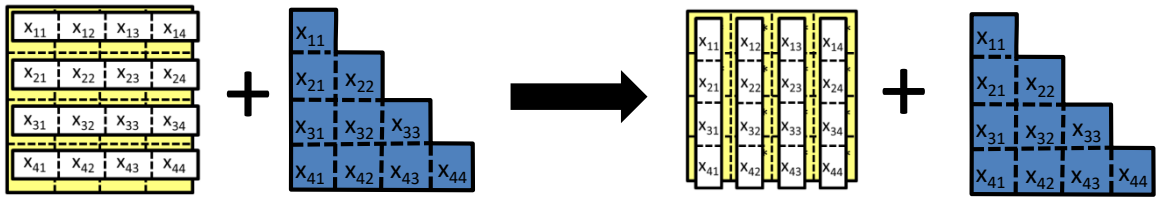
Catalysis

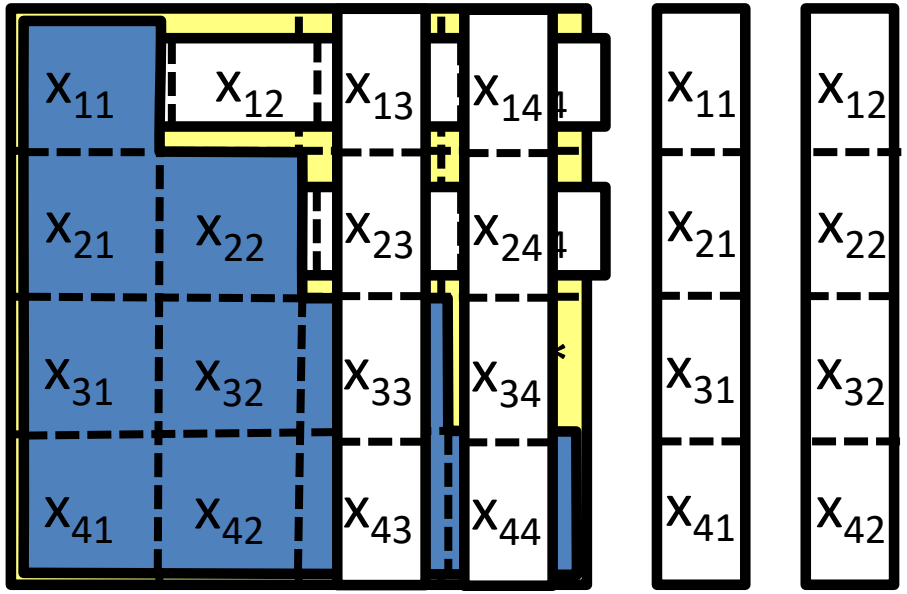

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

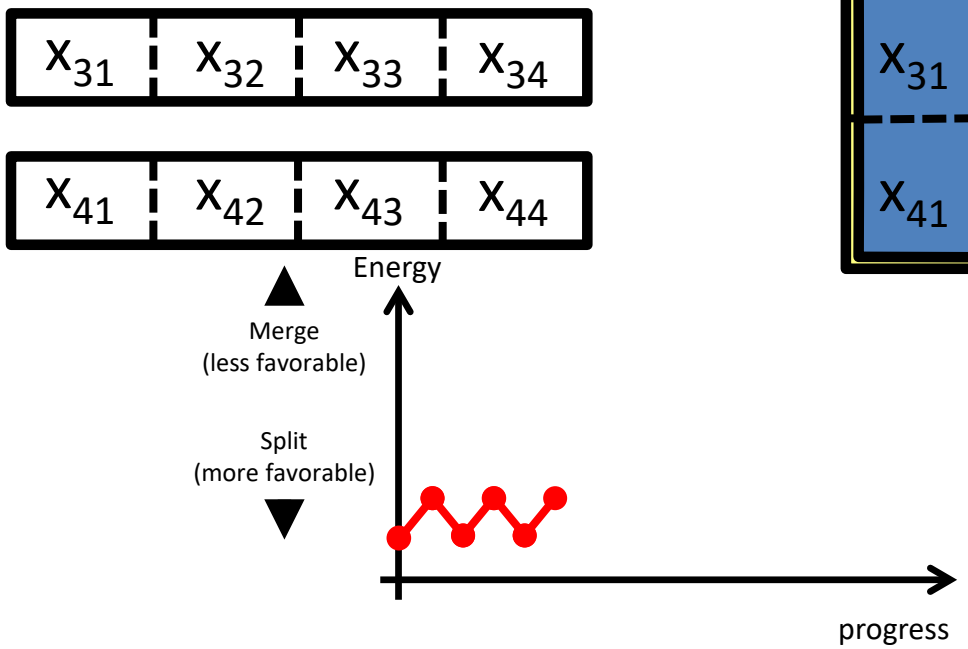

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x12 x13
x21 x22 x23
x31 x32 x33
x41 x42 x43


x31 x32 x33 x34
x41 x42 x43 x44
Energy
Merge
(less favorable)
Split
(more favorable)
progress

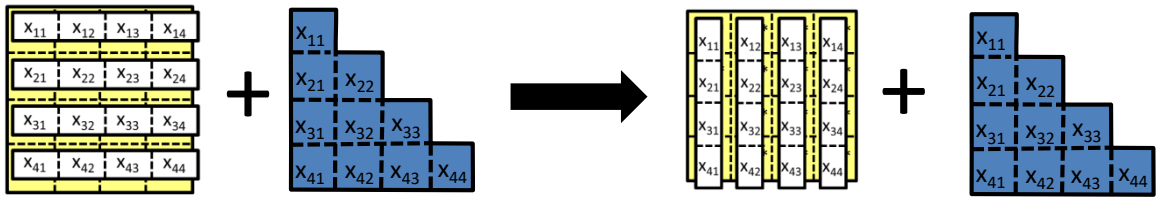
Catalysis

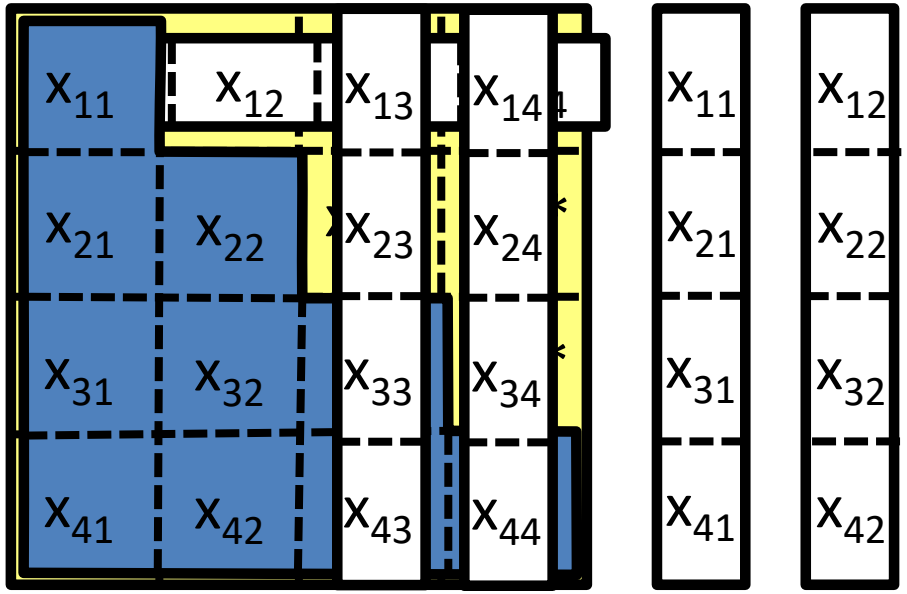

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

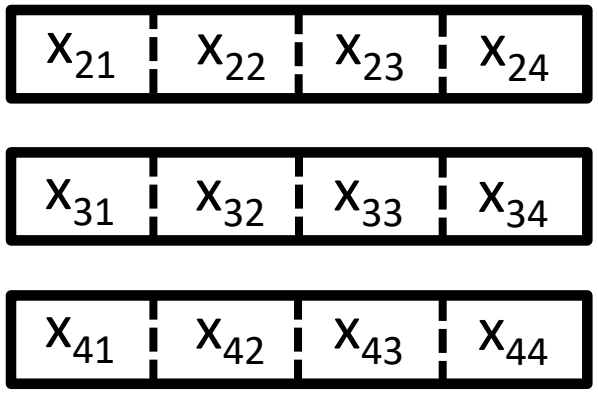

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x12
x21 x22
x31 x32
x41 x42

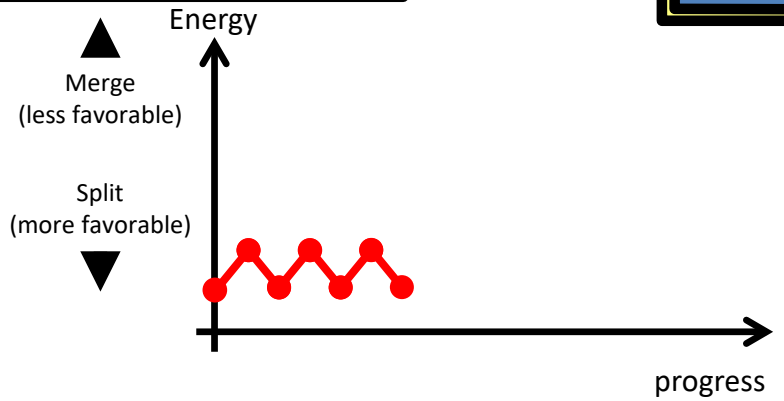

x31 x32 x33 x34
x41 x42 x43 x44
Energy
Merge
(less favorable)
Split
(more favorable)
progress

Catalysis

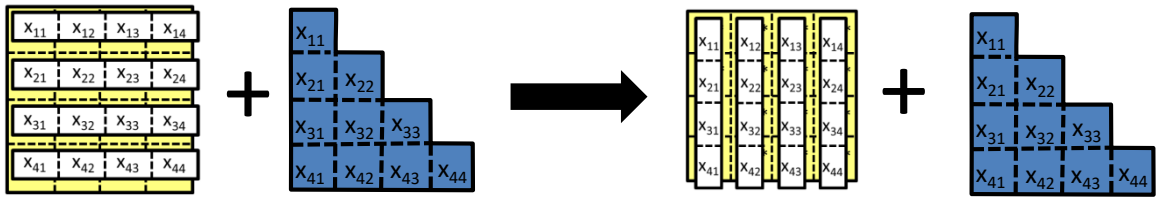

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

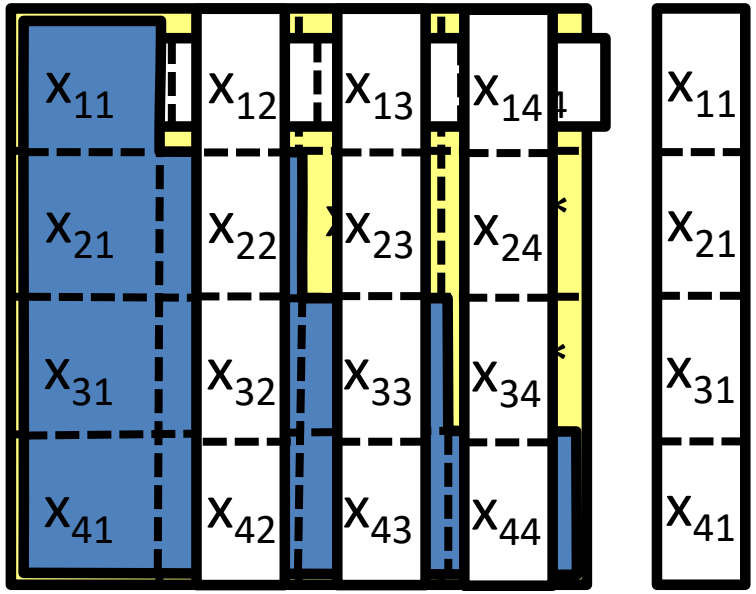

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11 x12
x21 x22
x31 x32
x41 x42

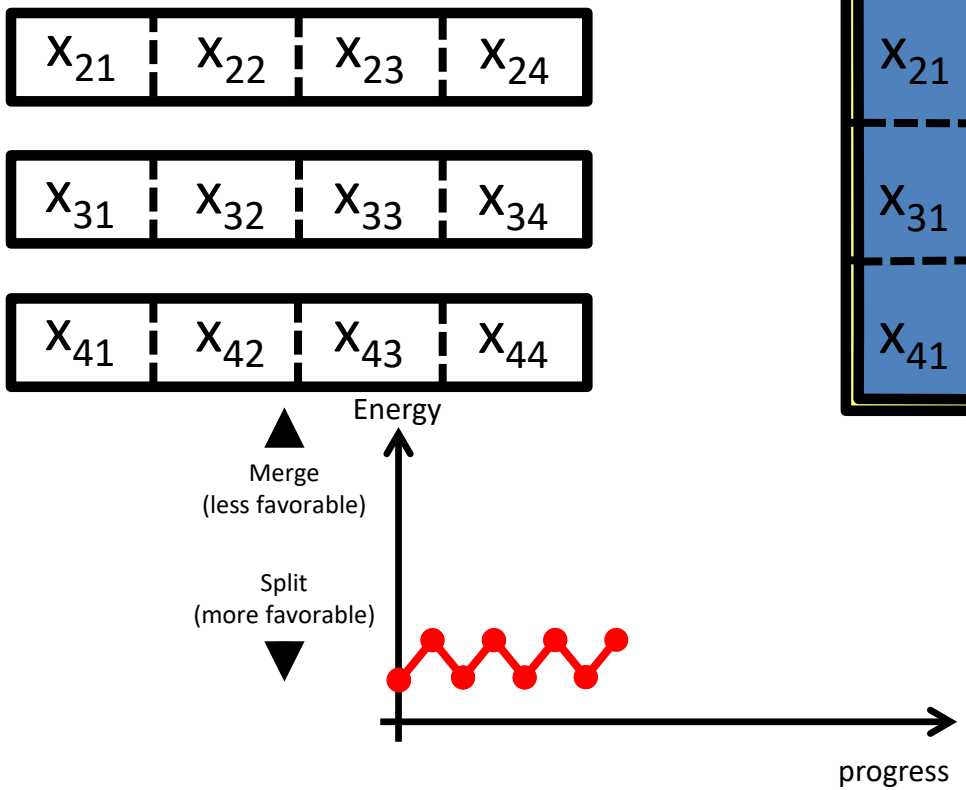

x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44


Energy
Merge
(less favorable)
Split
(more favorable)
progress

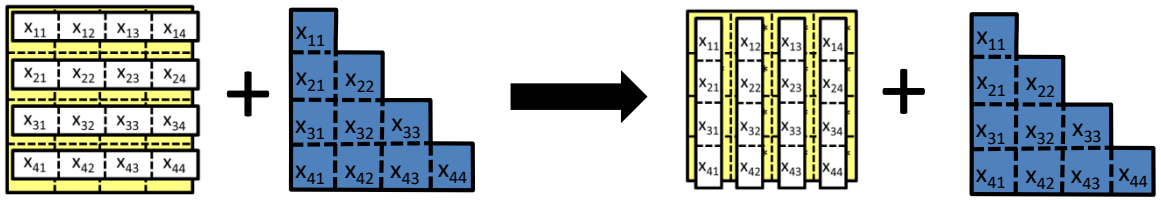
Catalysis

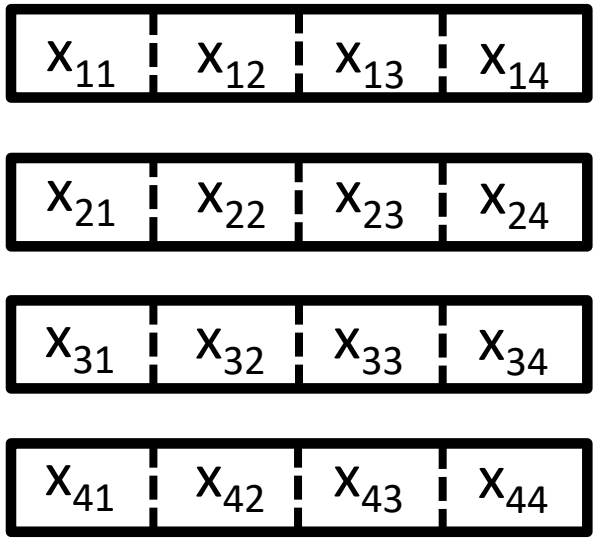

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

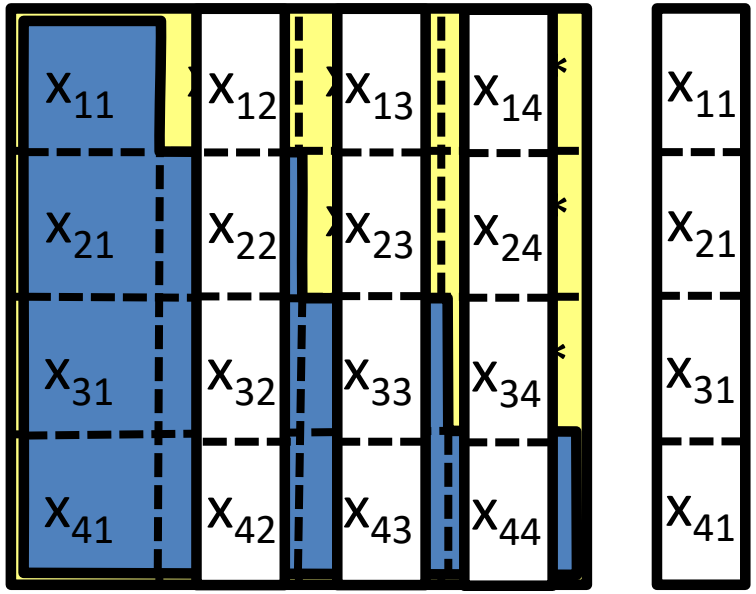

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11
x21
x31
x41

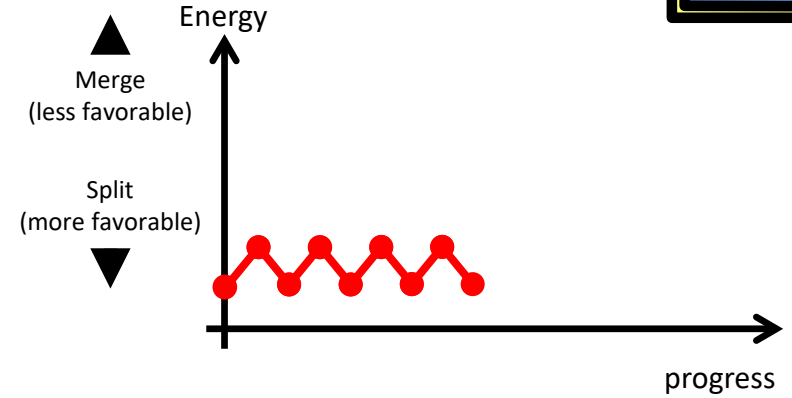

x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
Energy
Merge
(less favorable)
Split
(more favorable)
progress

Catalysis

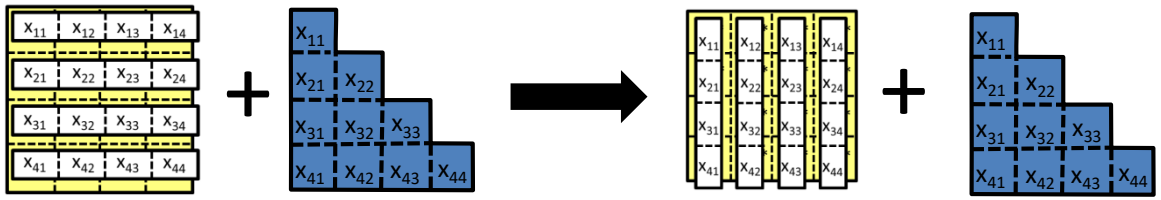

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

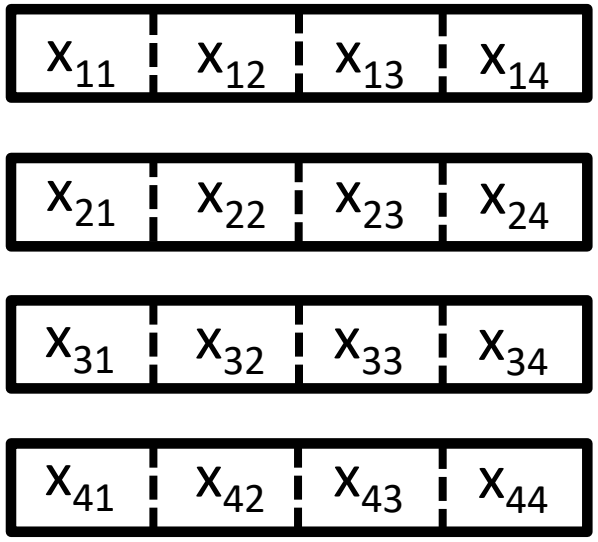

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

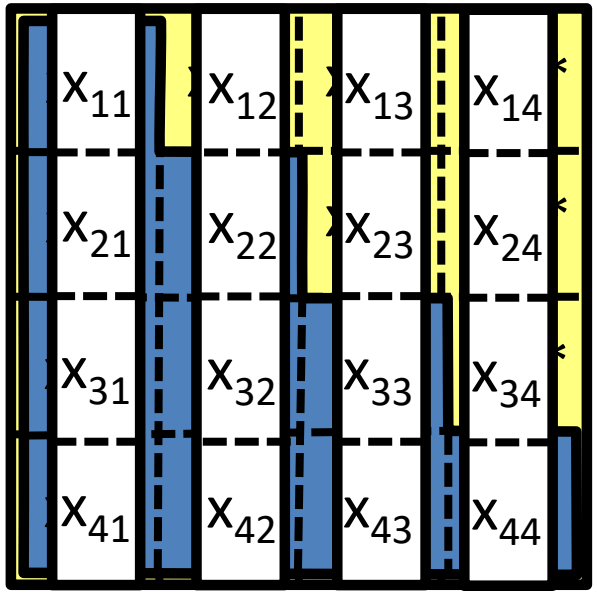

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44
x11
x21
x31
x41

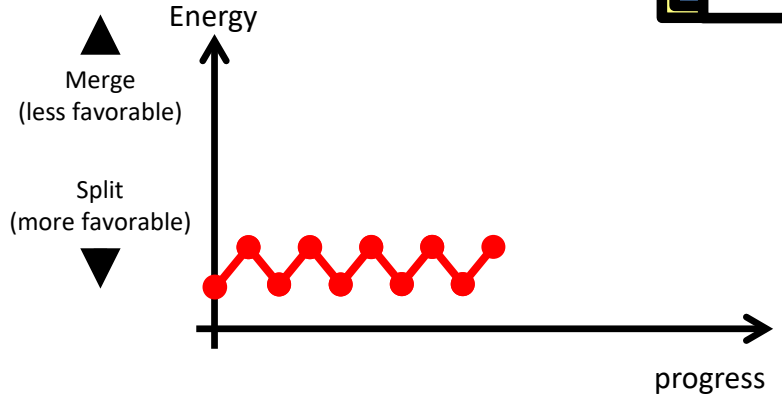

Energy
Merge
(less favorable)
Split
(more favorable)
progress

Catalysis

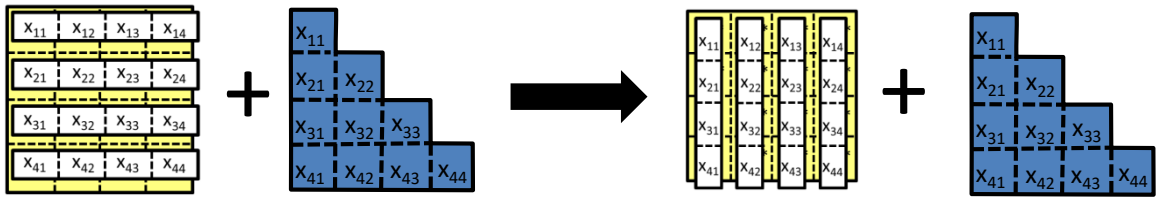

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

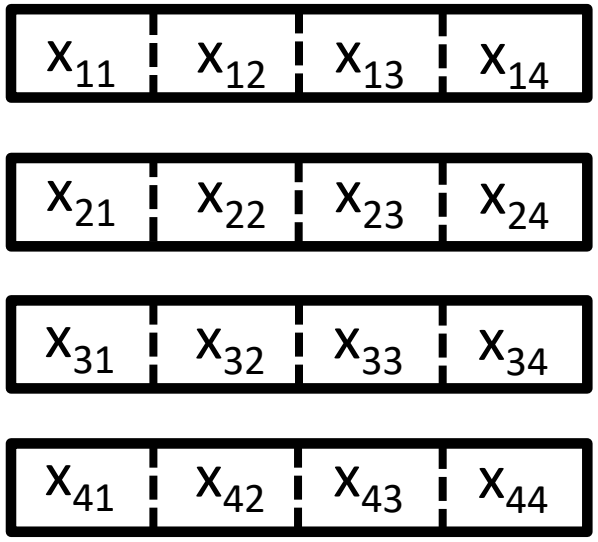

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

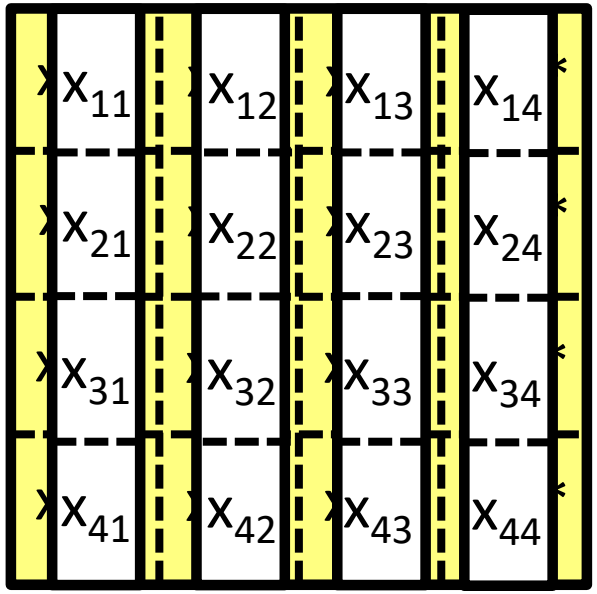

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

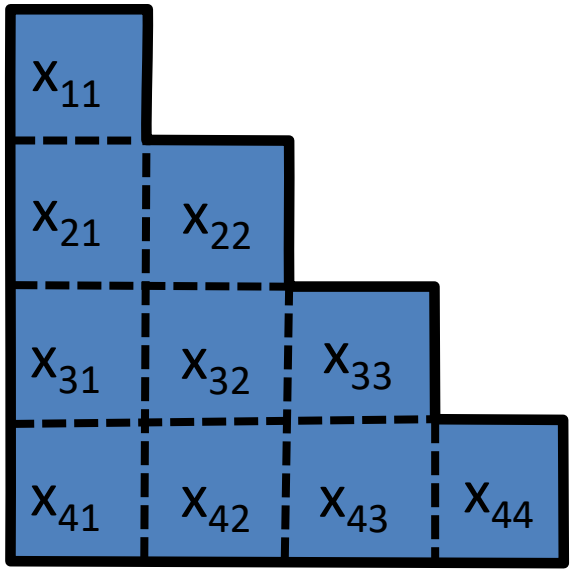

Energy
Merge
(less favorable)
Split
(more favorable)
progress

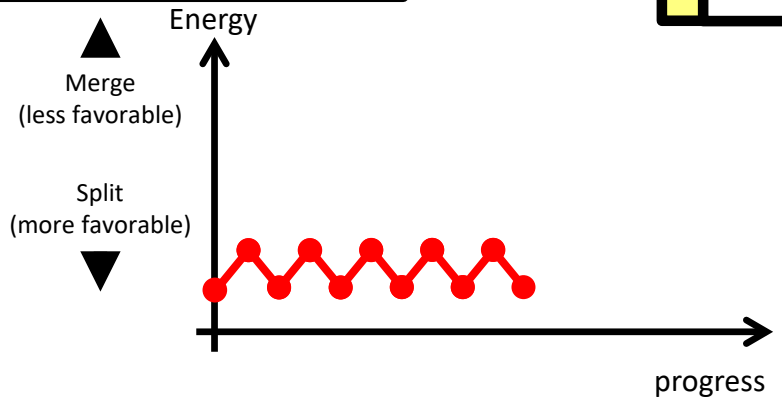
Catalysis


x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

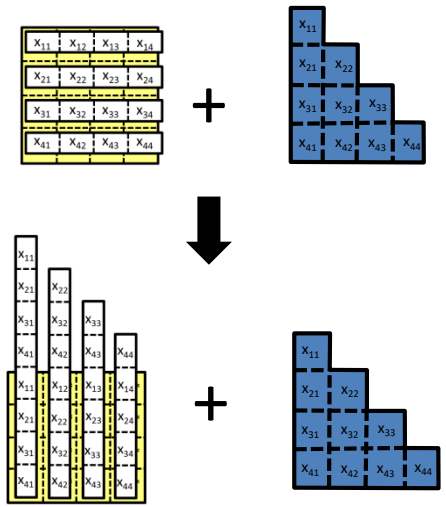

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

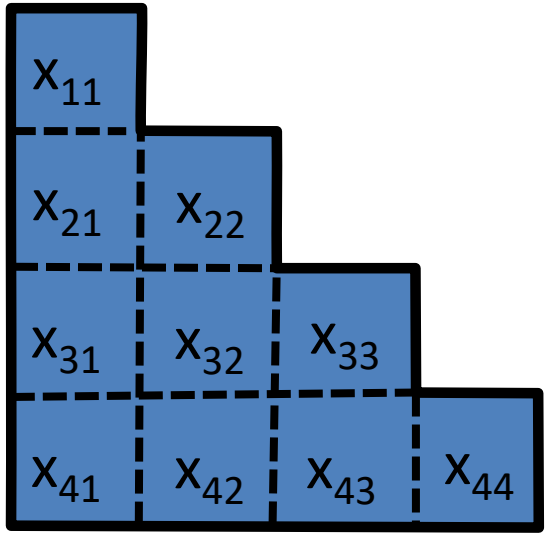

x11 x12 x13 x14
x21 x22 x23 x24
x31 x32 x33 x34
x41 x42 x43 x44

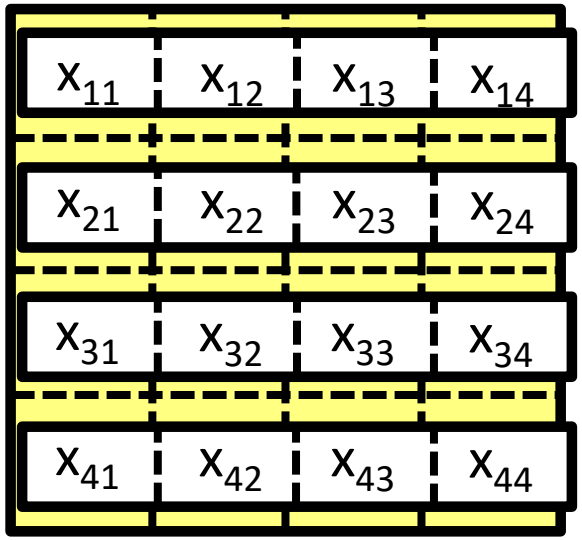

x11
x21 x22
x31 x32 x33
x41 x42 x43 x44

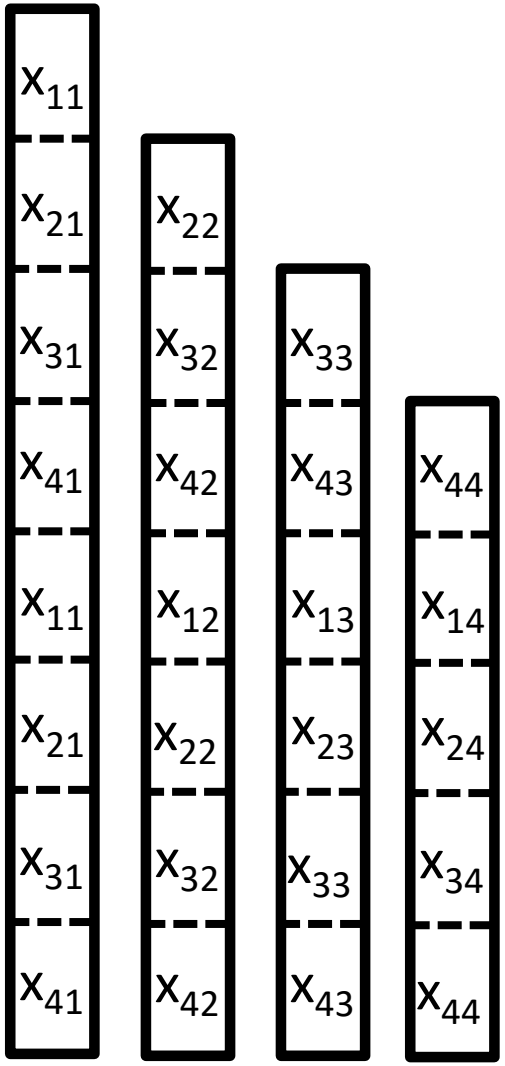

Energy
Merge
(less favorable)
Split
(more favorable)
progress

Autocatalysis

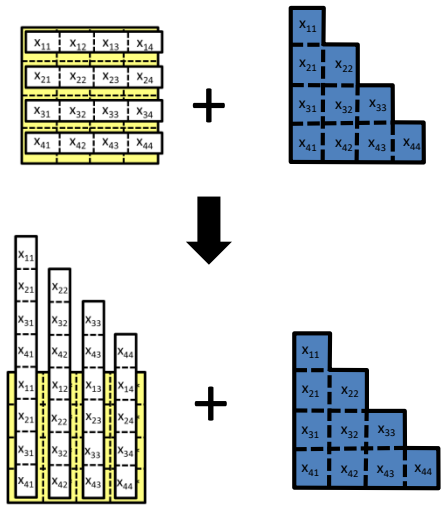


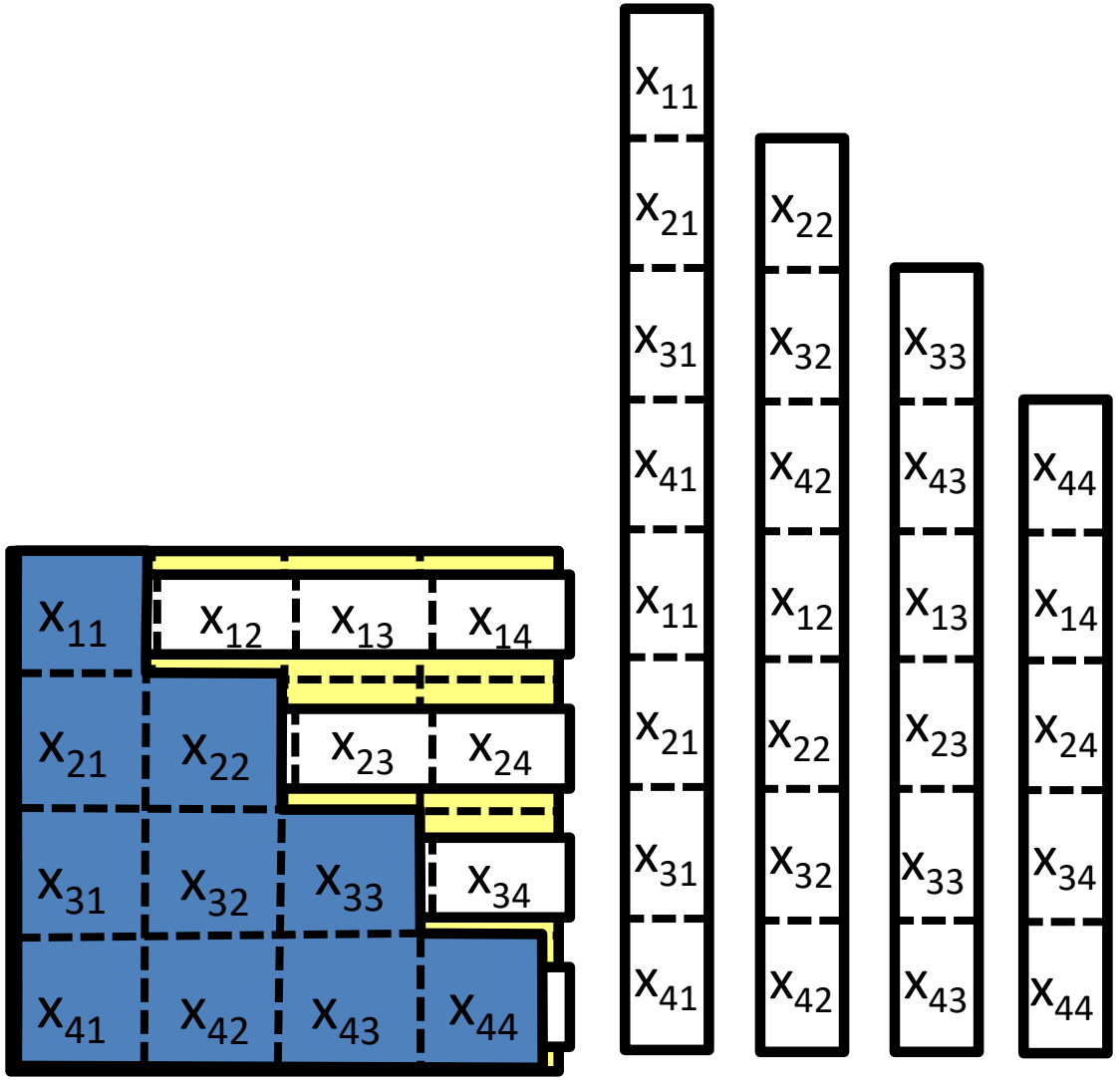




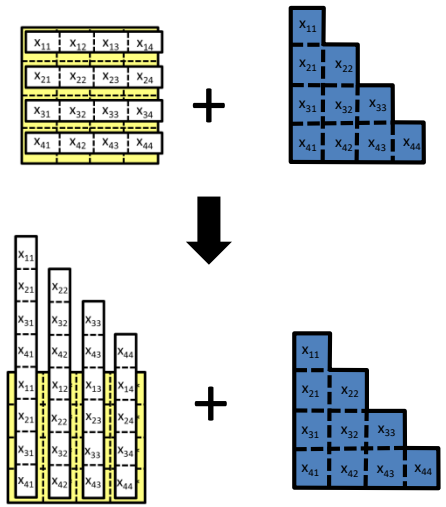


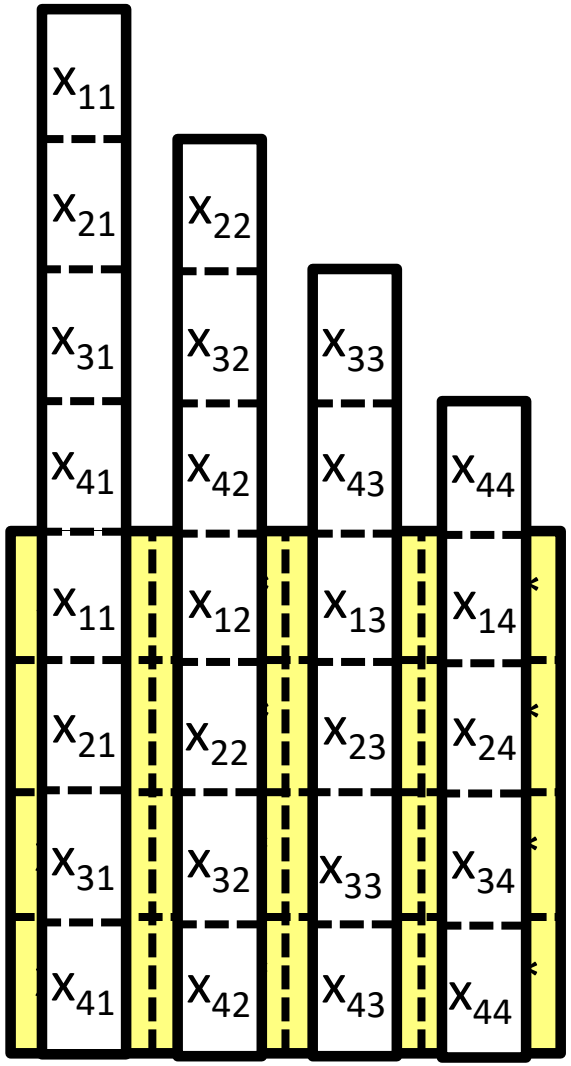
Autocatalysis

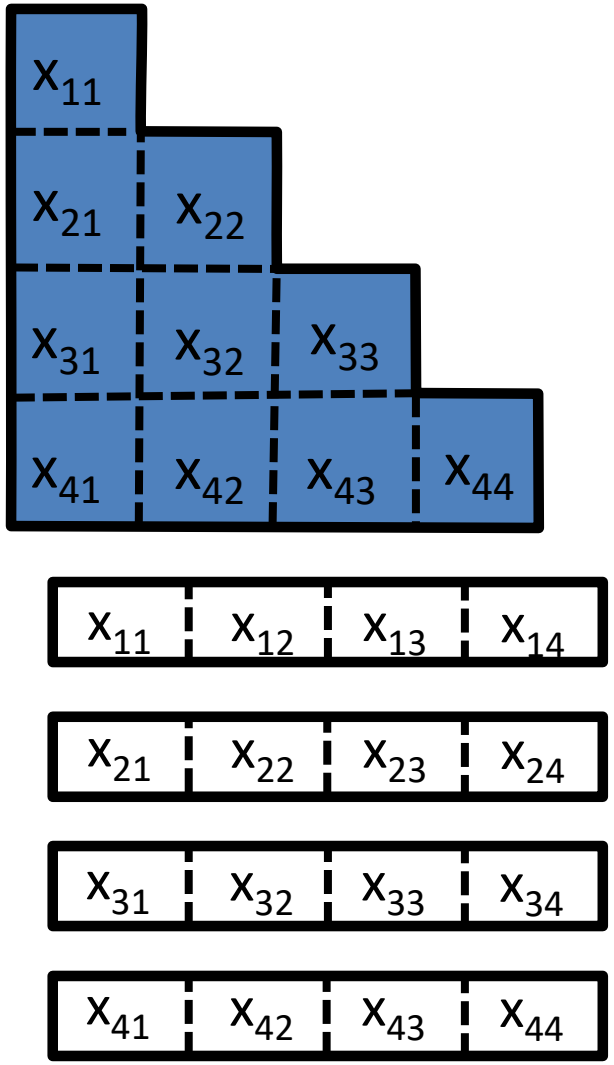




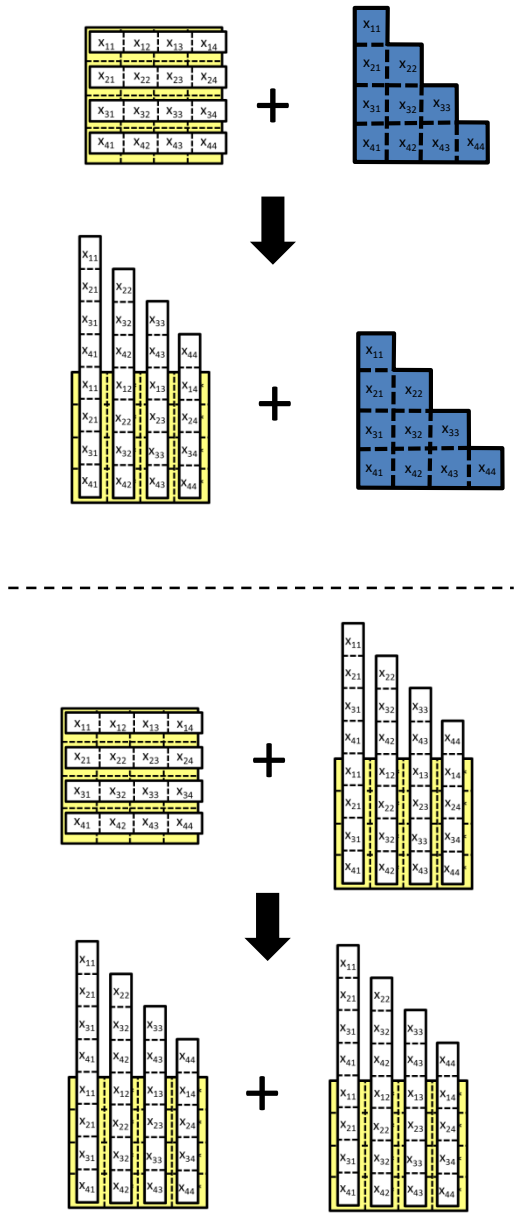
Autocatalysis

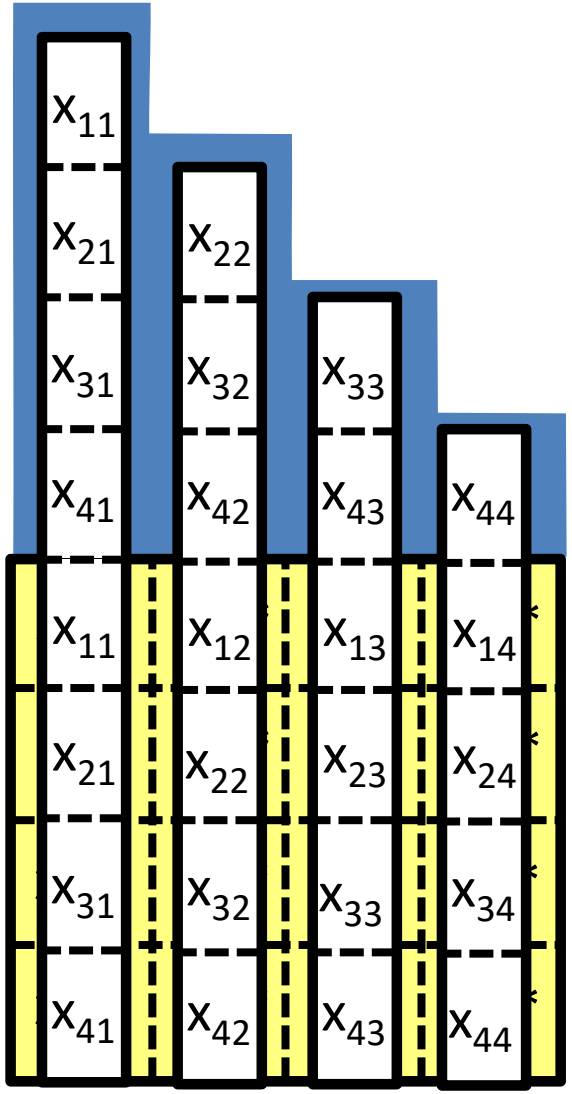


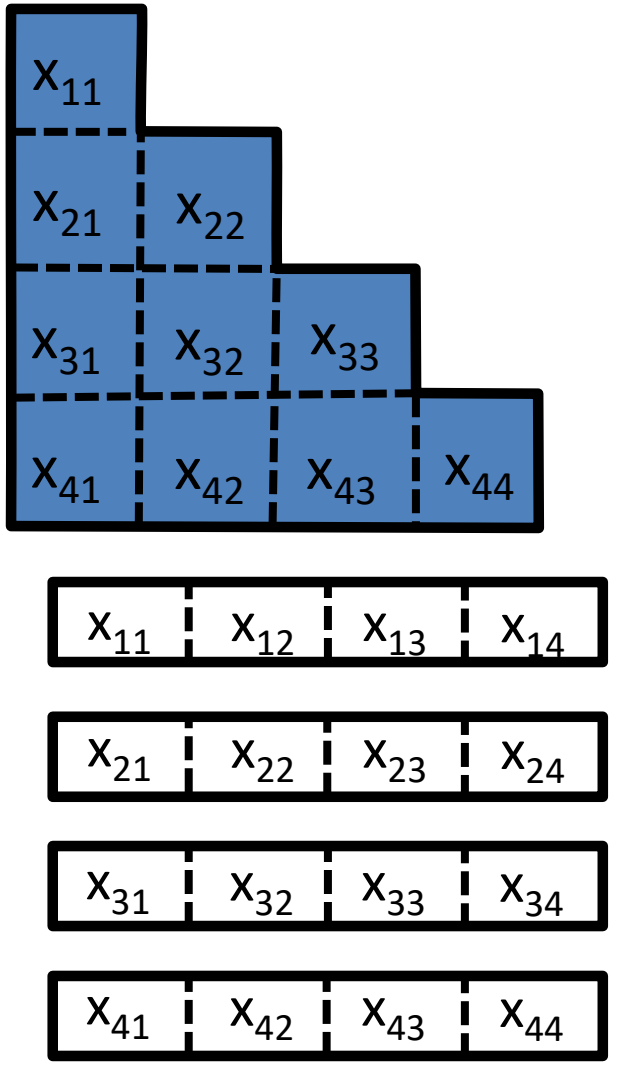




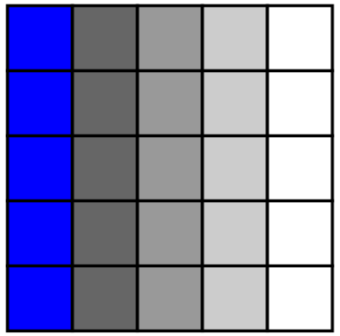
Autocatalysis

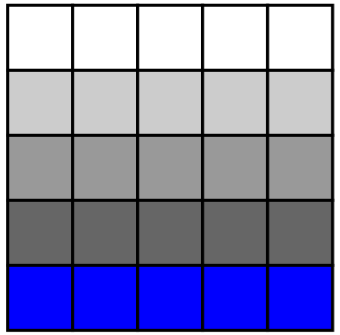




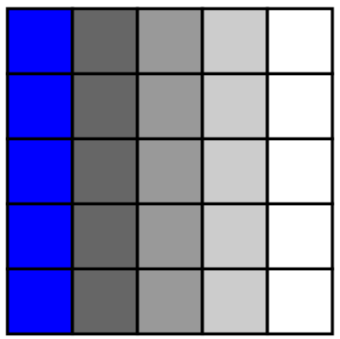


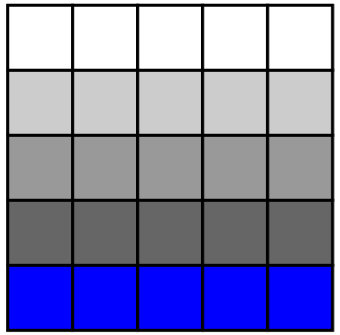
Multiple Stable Configurations

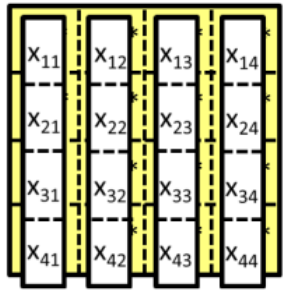


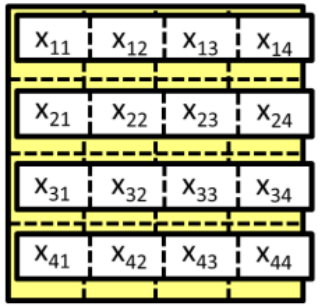


Multiple Stable Configurations

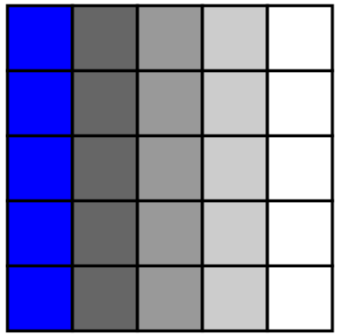


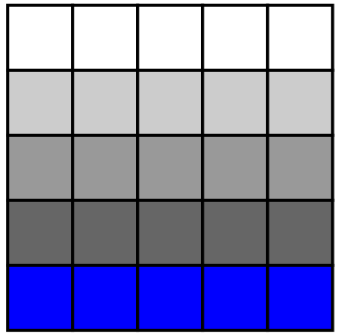




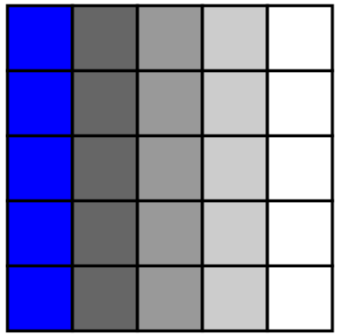


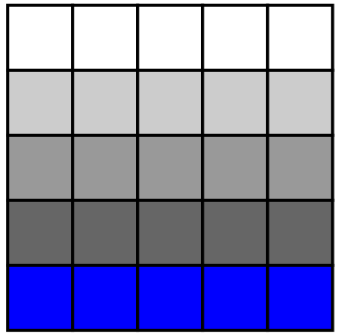
Multiple Stable Configurations

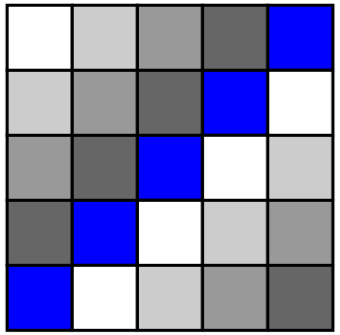




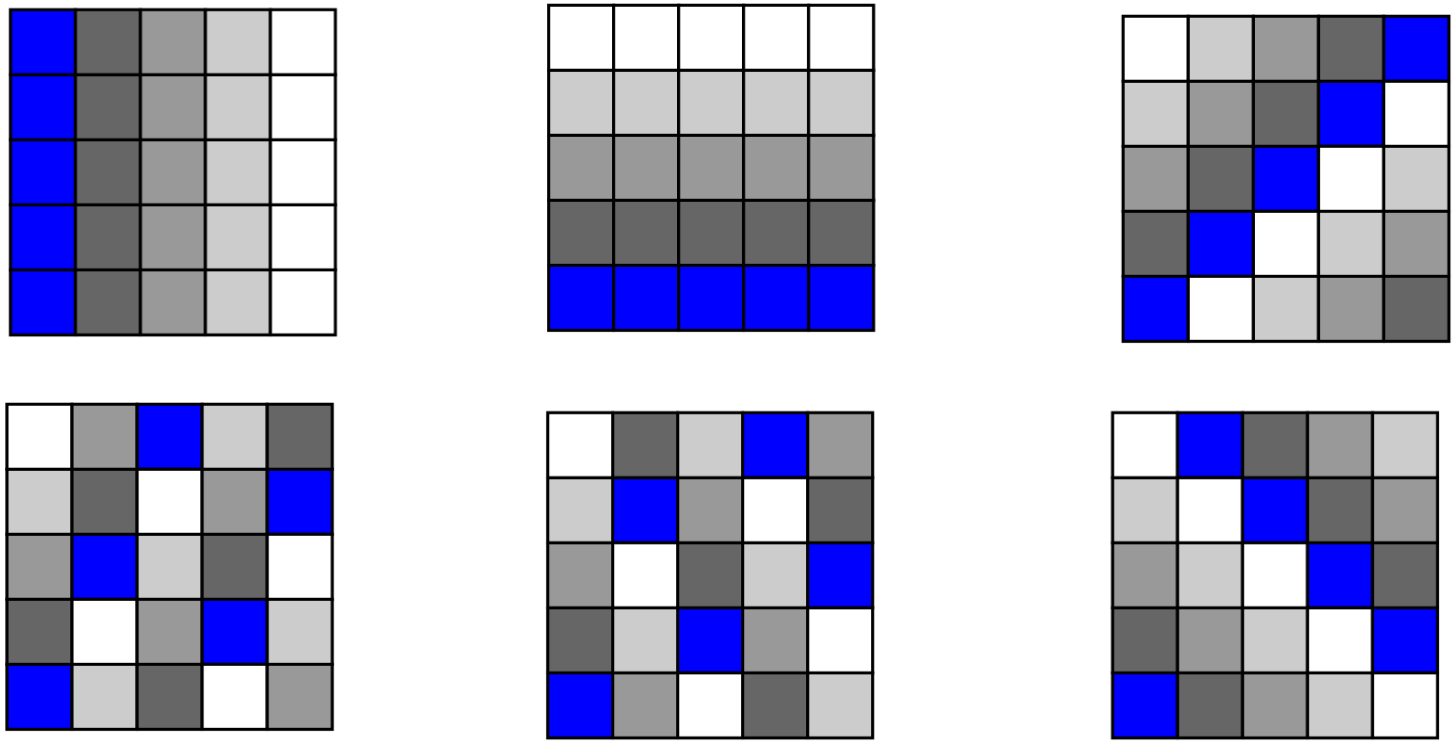
Multiple Stable Configurations





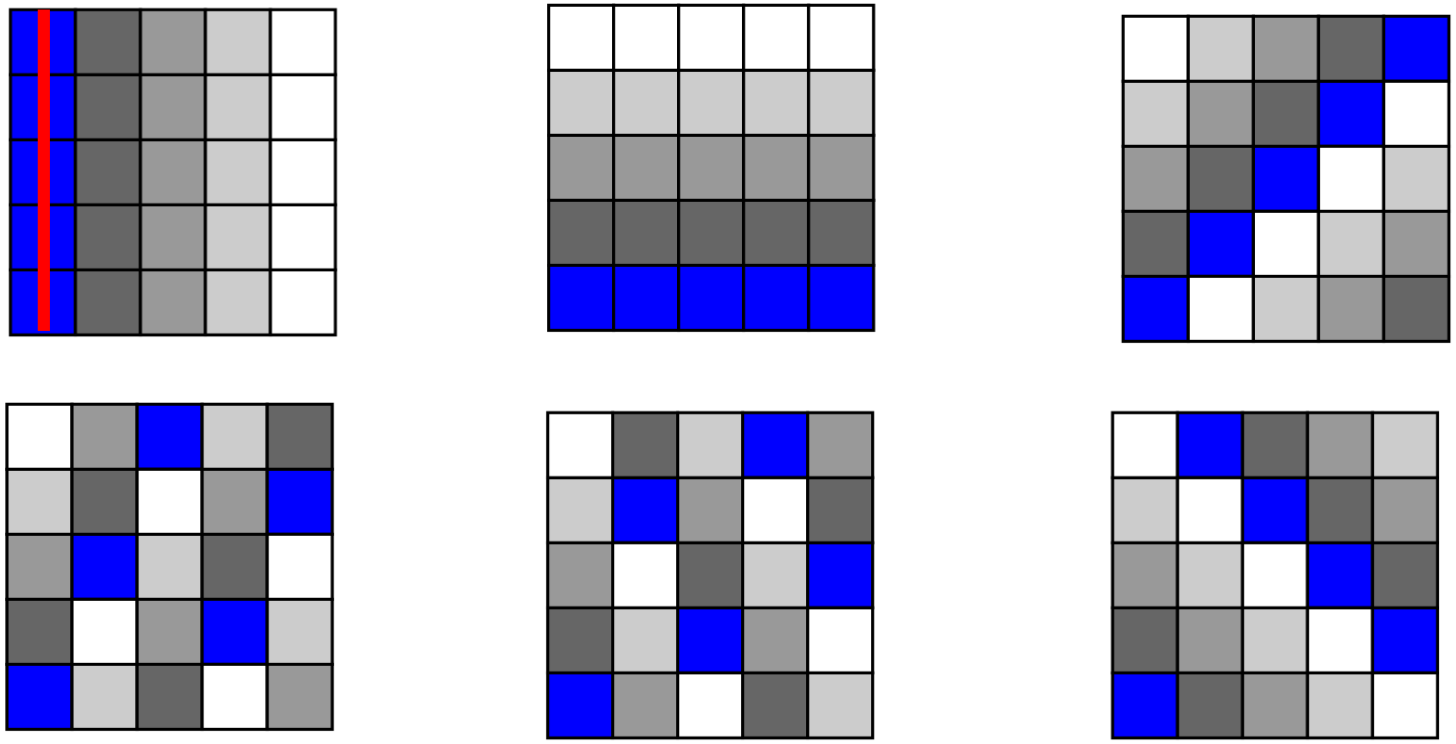


Multiple Stable Configurations



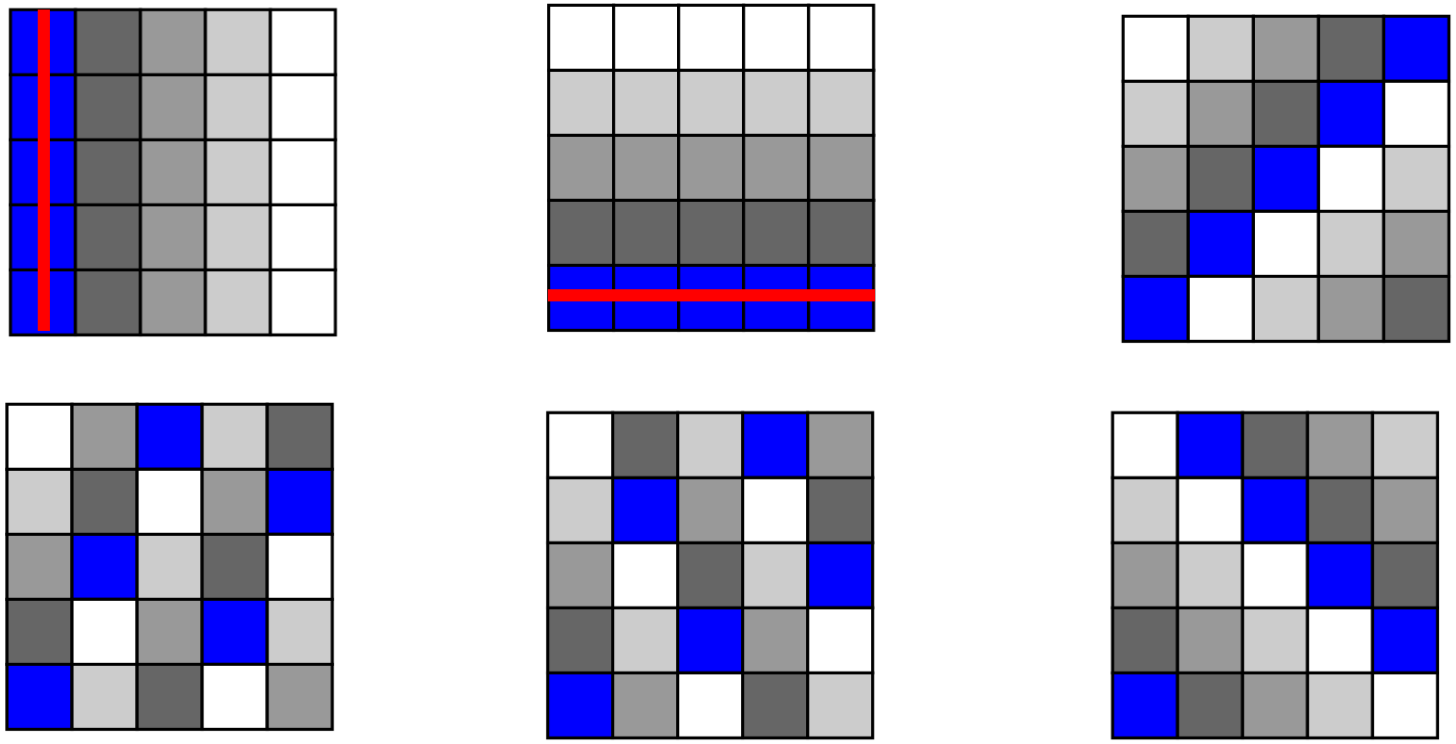
For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



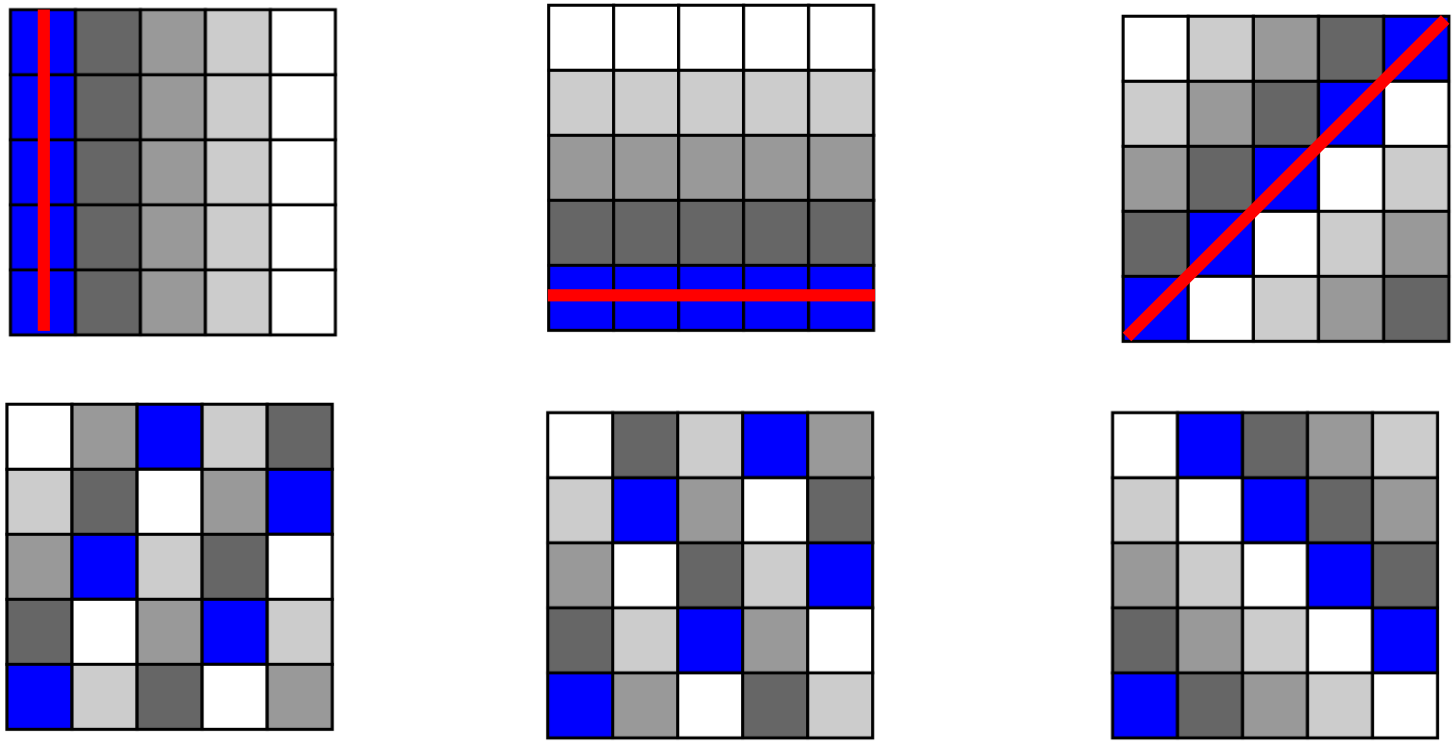
For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



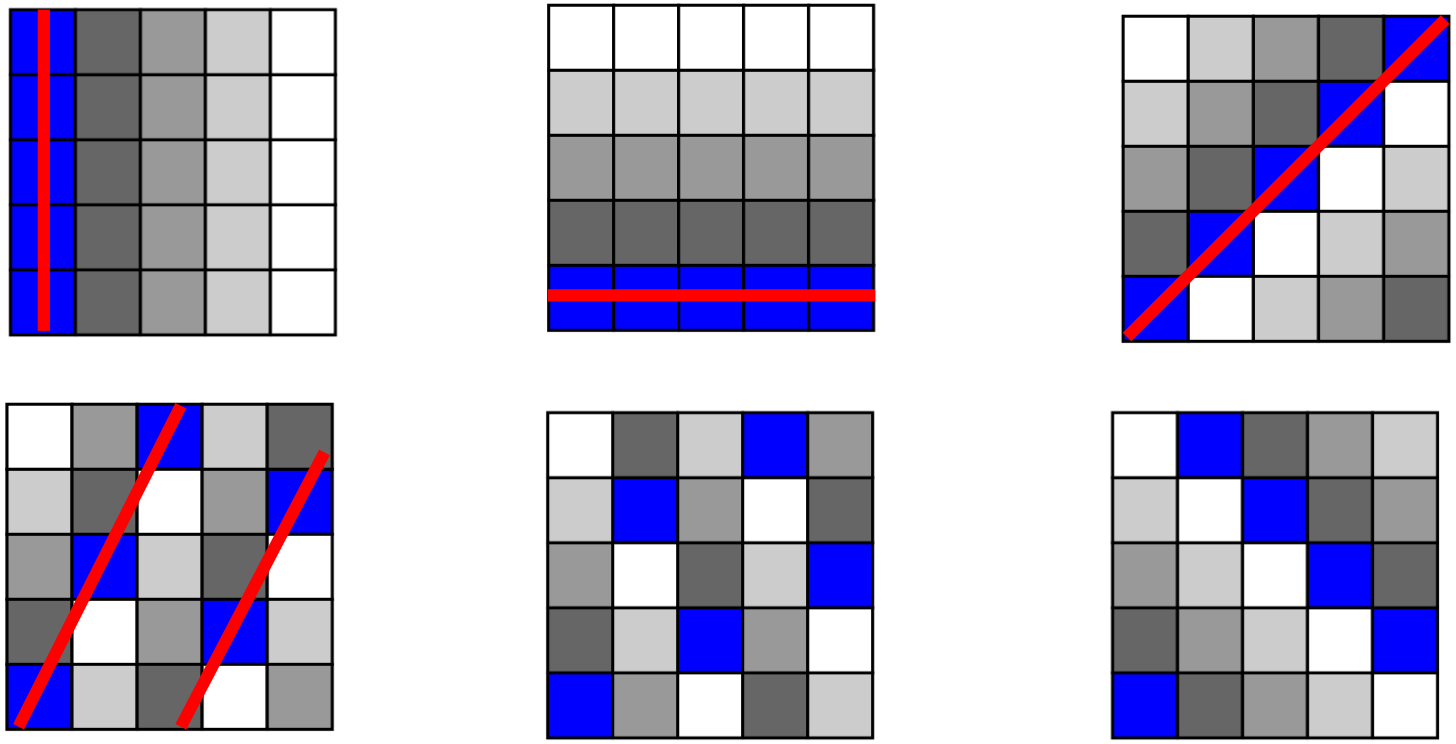
For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



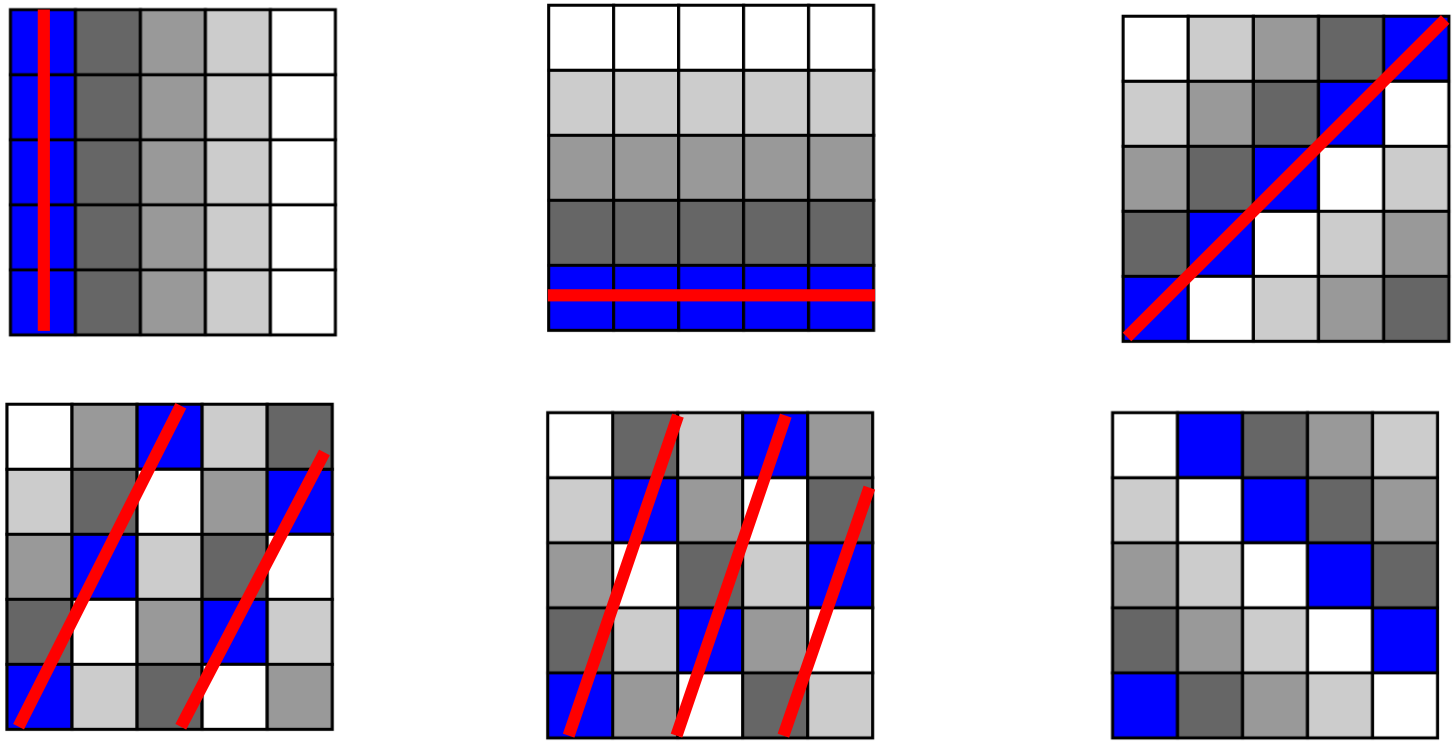
For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



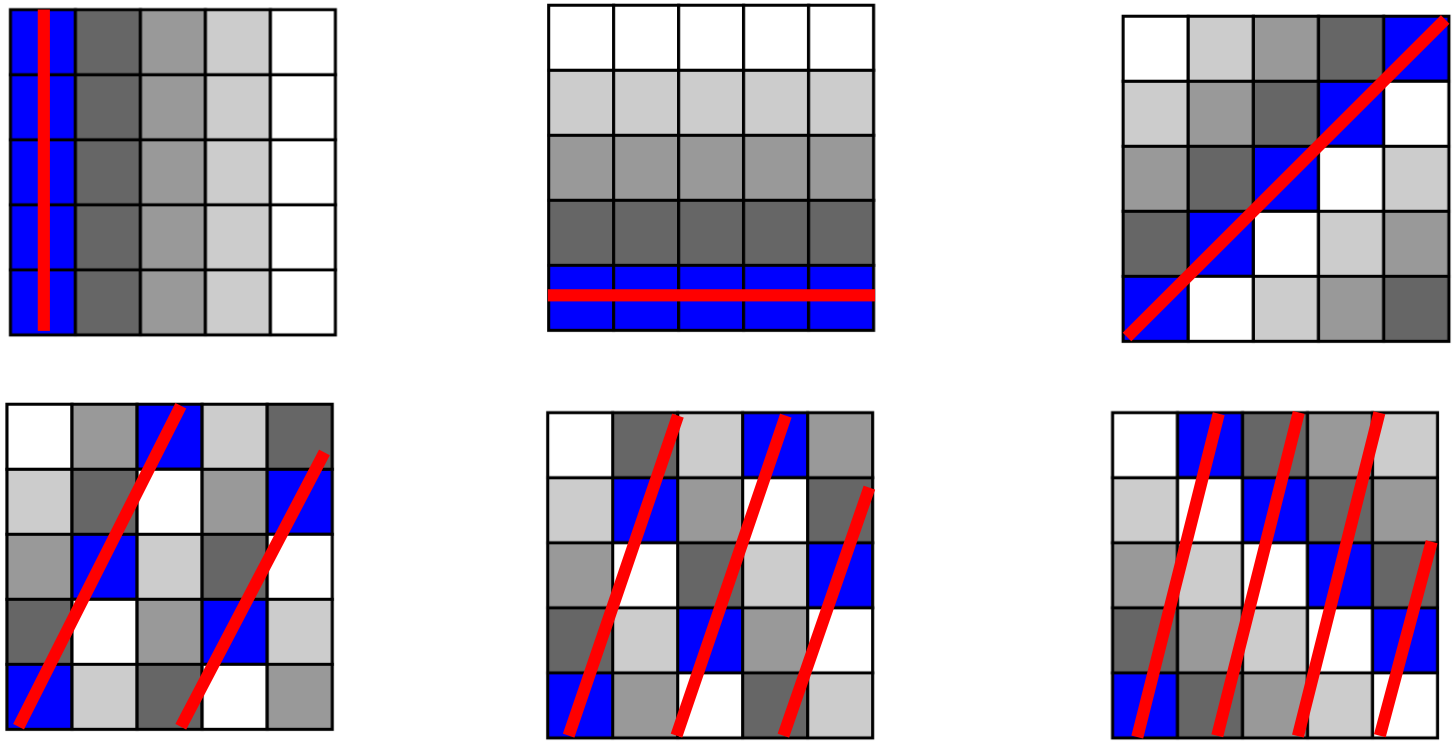
For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Multiple Stable Configurations



For a grid of size *n* x *n*, there can be at most *n*+1 different stable configurations with barrier *n* to pass between any of them

Social Golfer Problem

- Can 25 (n^2) golfers play in 5-somes (n-somes) for 6 (n+1) days, so that no two golfers play together more than once?

Social Golfer Problem

- Can 25 (n^2) golfers play in 5-somes (n-somes) for 6 (n+1) days, so that no two golfers play together more than once?
- First studied by Euler.

Social Golfer Problem

- Can 25 (n^2) golfers play in 5-somes (n-somes) for 6 (n+1) days, so that no two golfers play together more than once?
- First studied by Euler.
- True if n is a prime power (2,3,4,5,7,8,9,11,13,…)

Social Golfer Problem

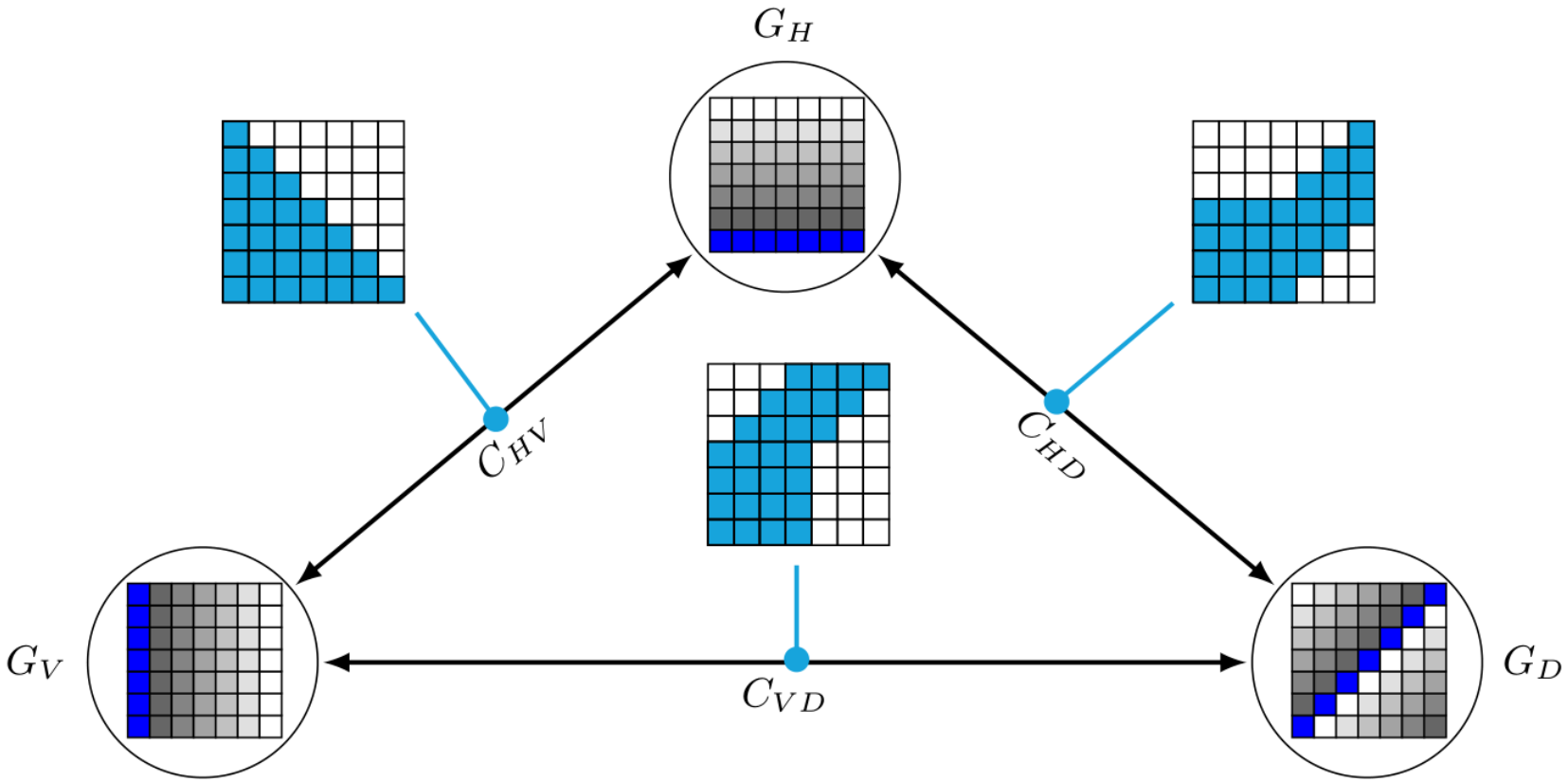
- Can 25 (n^2) golfers play in 5-somes (n-somes) for 6 (n+1) days, so that no two golfers play together more than once?
- First studied by Euler.
- True if n is a prime power (2,3,4,5,7,8,9,11,13,…)
- False for smallest prime power n=6: can only play for 3 days!

  [Gaston Tarry (1901). "Le Probléme des 36 Officiers". *Compte Rendu de l'Association Française pour l'Avancement des Sciences*. Secrétariat de l'Association. 2: 170–203.]

Social Golfer Problem

- Can 25 (n^2) golfers play in 5-somes (n-somes) for 6 (n+1) days, so that no two golfers play together more than once?
- First studied by Euler.
- True if n is a prime power (2,3,4,5,7,8,9,11,13,…)
- False for smallest prime power n=6: can only play for 3 days!
  [Gaston Tarry (1901). "Le Probléme des 36 Officiers". *Compte Rendu de l'Association Française pour l'Avancement des Sciences*. Secrétariat de l'Association. 2: 170–203.]
- Unknown for next prime power n=10:
  - trivial upper bound is 11 days
  - best known lower bound is 3

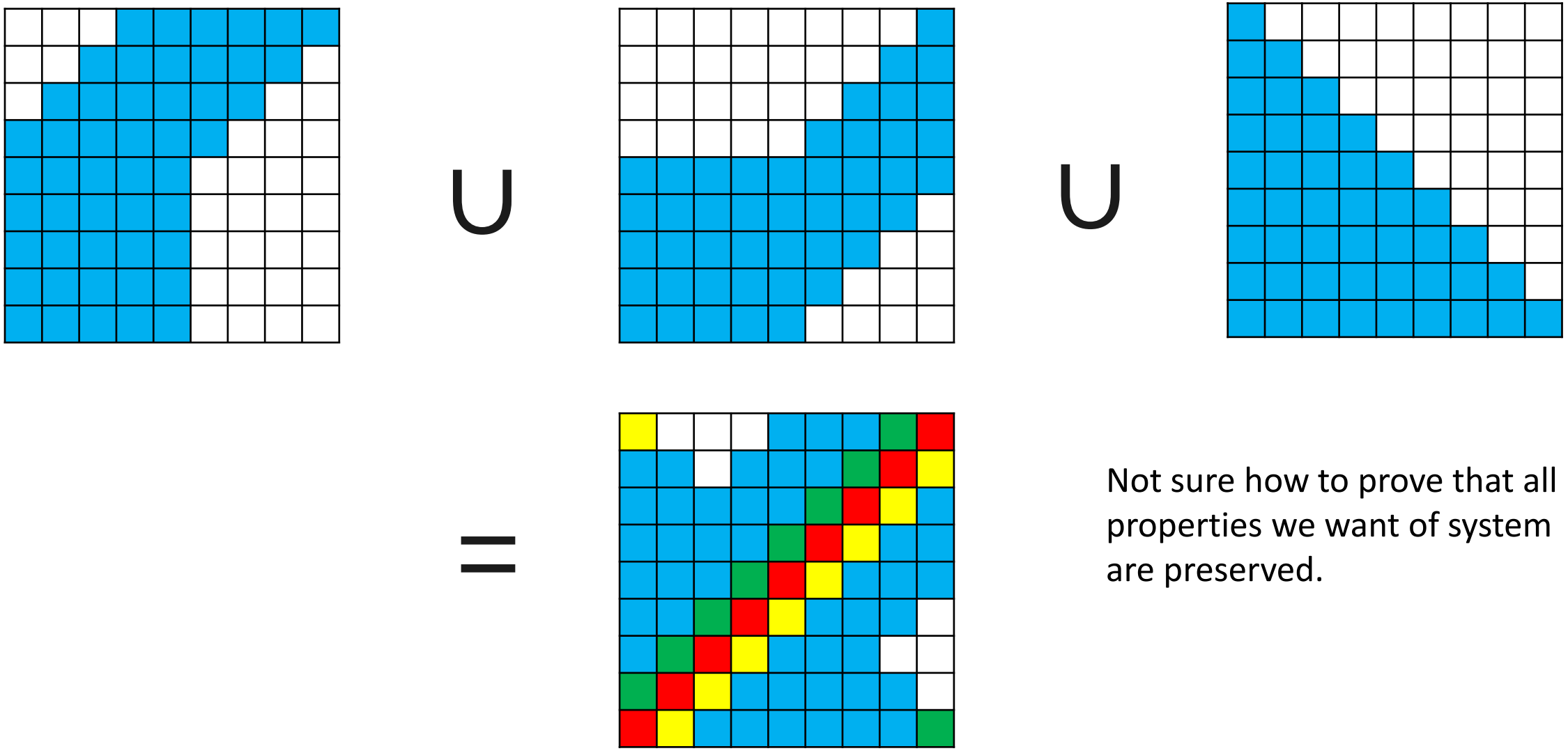
Directed Catalysis



Along a catalyzed pathway, the barrier is 1

Otherwise the barrier is $n/2$

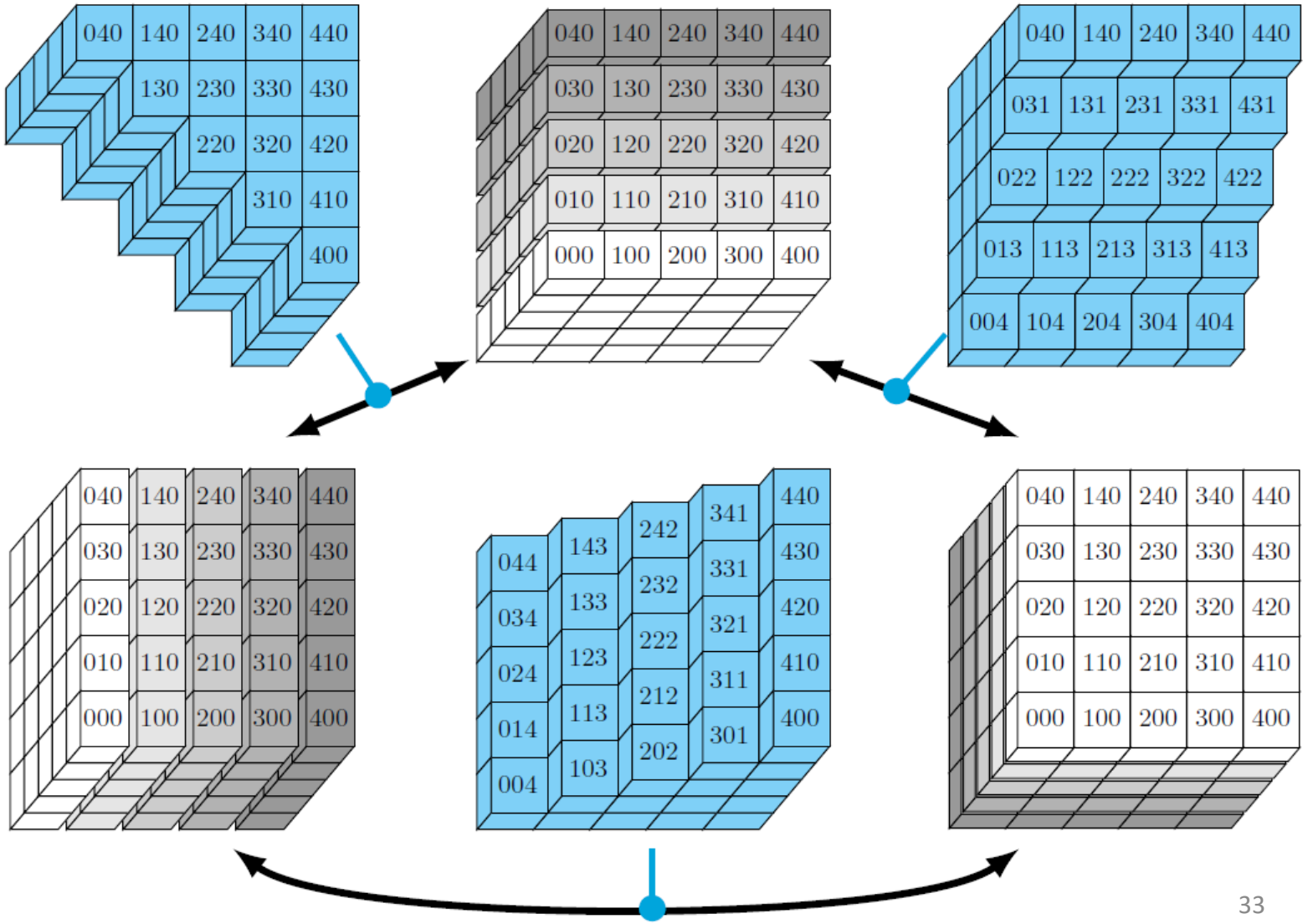
Allowing more than one catalyst at once



Not sure how to prove that all properties we want of system are preserved.

Allowing more than one catalyst at once

- To have *d* stable states, set of domains is discrete *d*-dimensional hypercube.
- Each monomer is (*d*-1)-dimensional hyperplane
- Maintains energy barrier *n*, not just Ω(*n*).



Future directions

Future directions

- Simultaneous presence of multiple directed catalysts (without exponential blowup)

Future directions

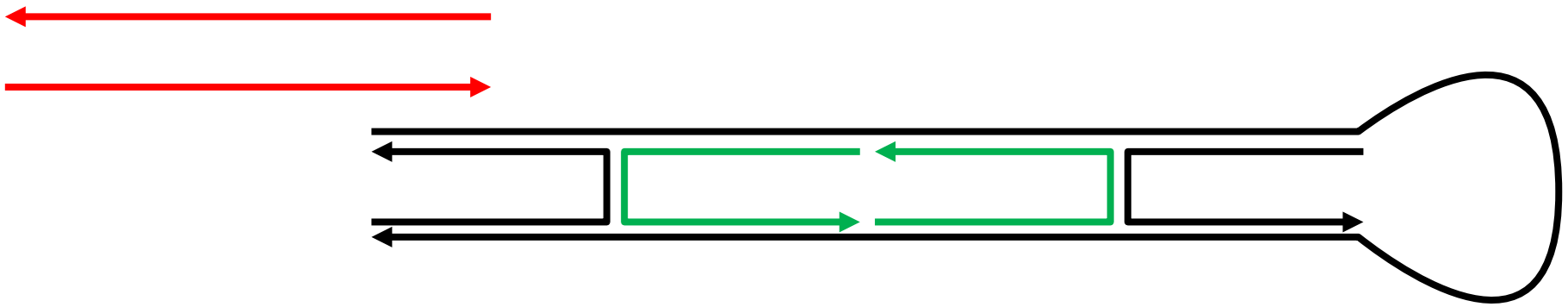
- Simultaneous presence of multiple directed catalysts (without exponential blowup)
- Minimum number of domain types to be robust to varying relative amounts of the monomers

Future directions

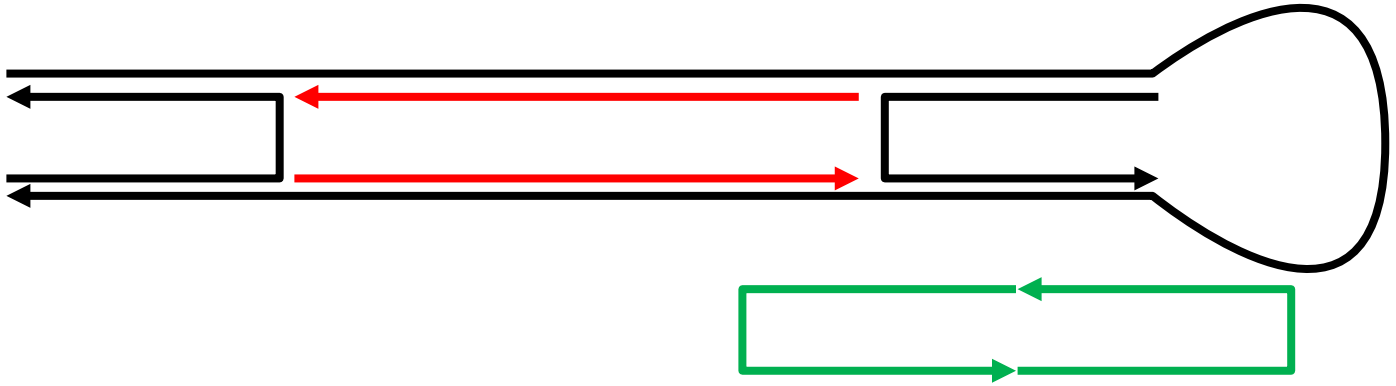
- Simultaneous presence of multiple directed catalysts (without exponential blowup)
- Minimum number of domain types to be robust to varying relative amounts of the monomers
- Constrained catalytic networks using overhangs

Thank you!

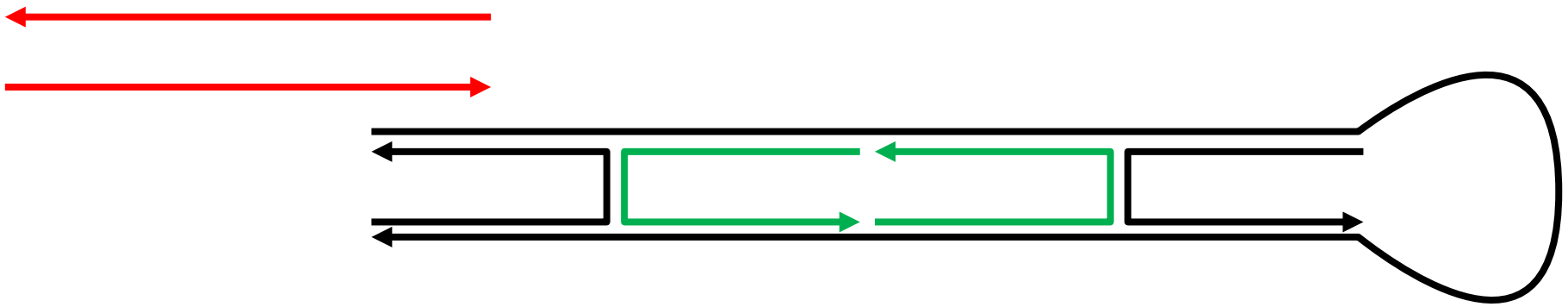
Feasible DNA implementation



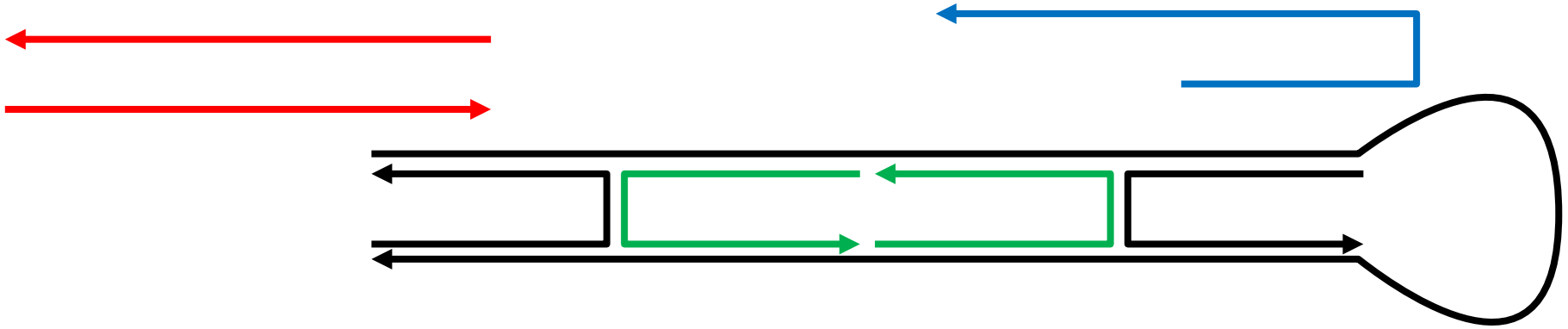
Feasible DNA implementation



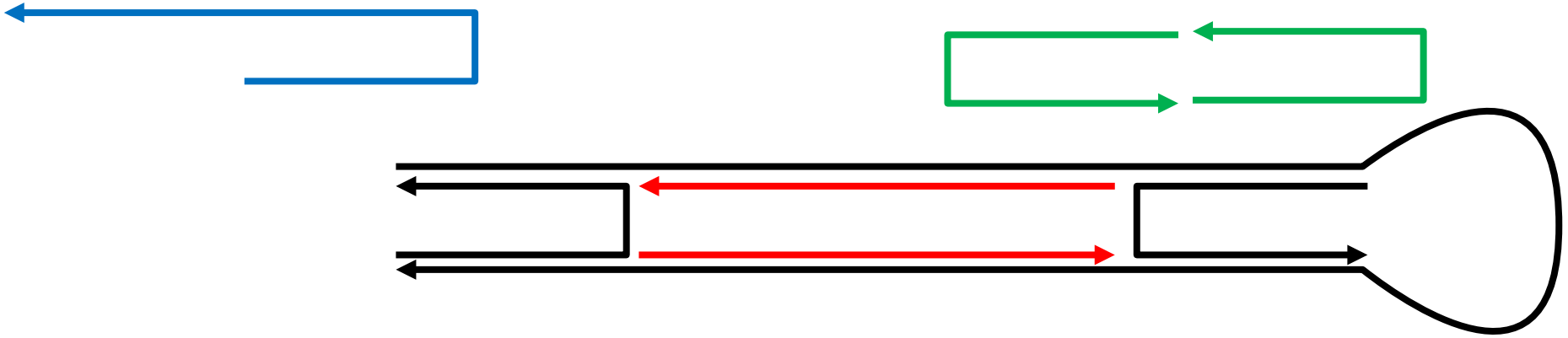
Feasible DNA implementation



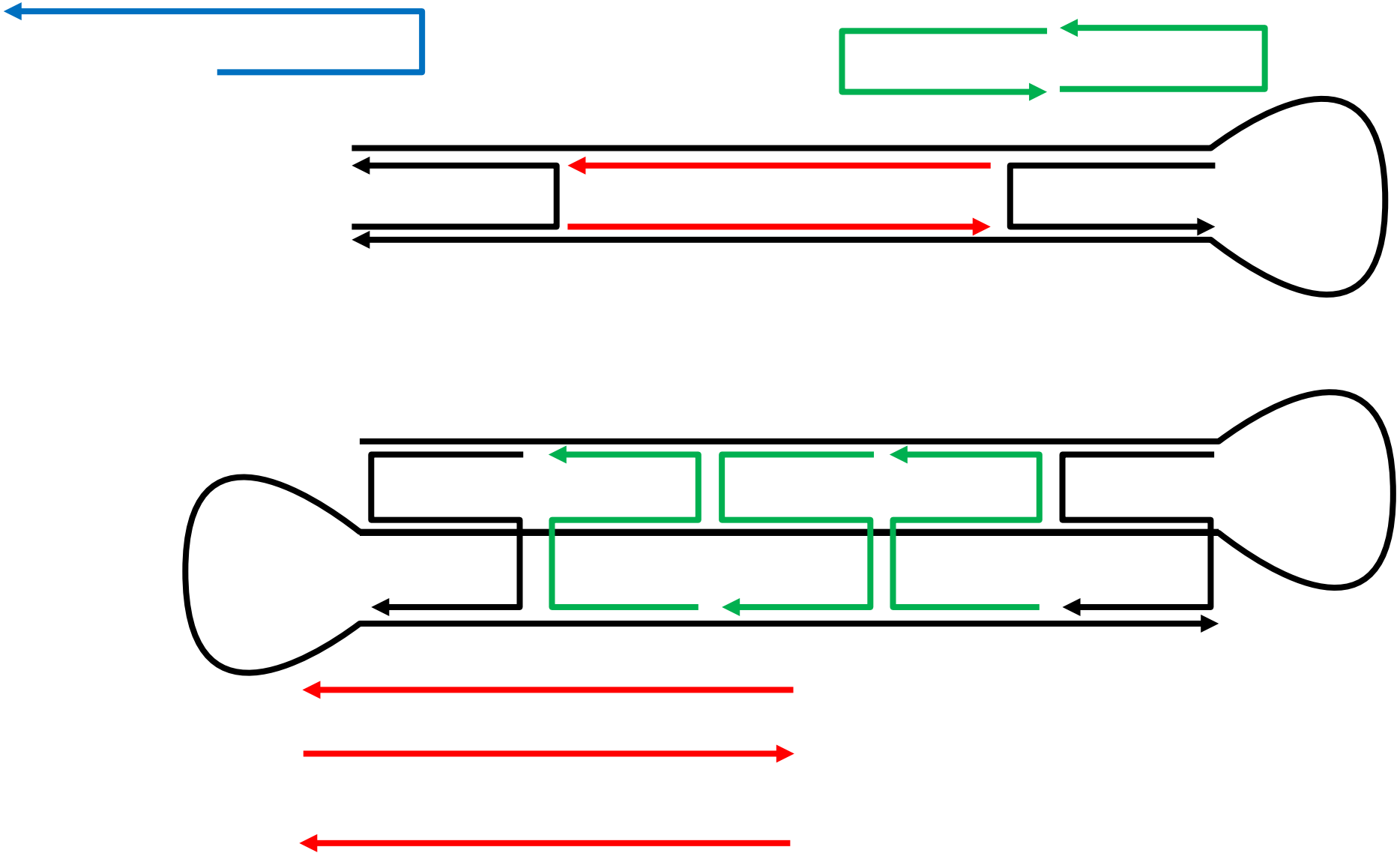
Feasible DNA implementation



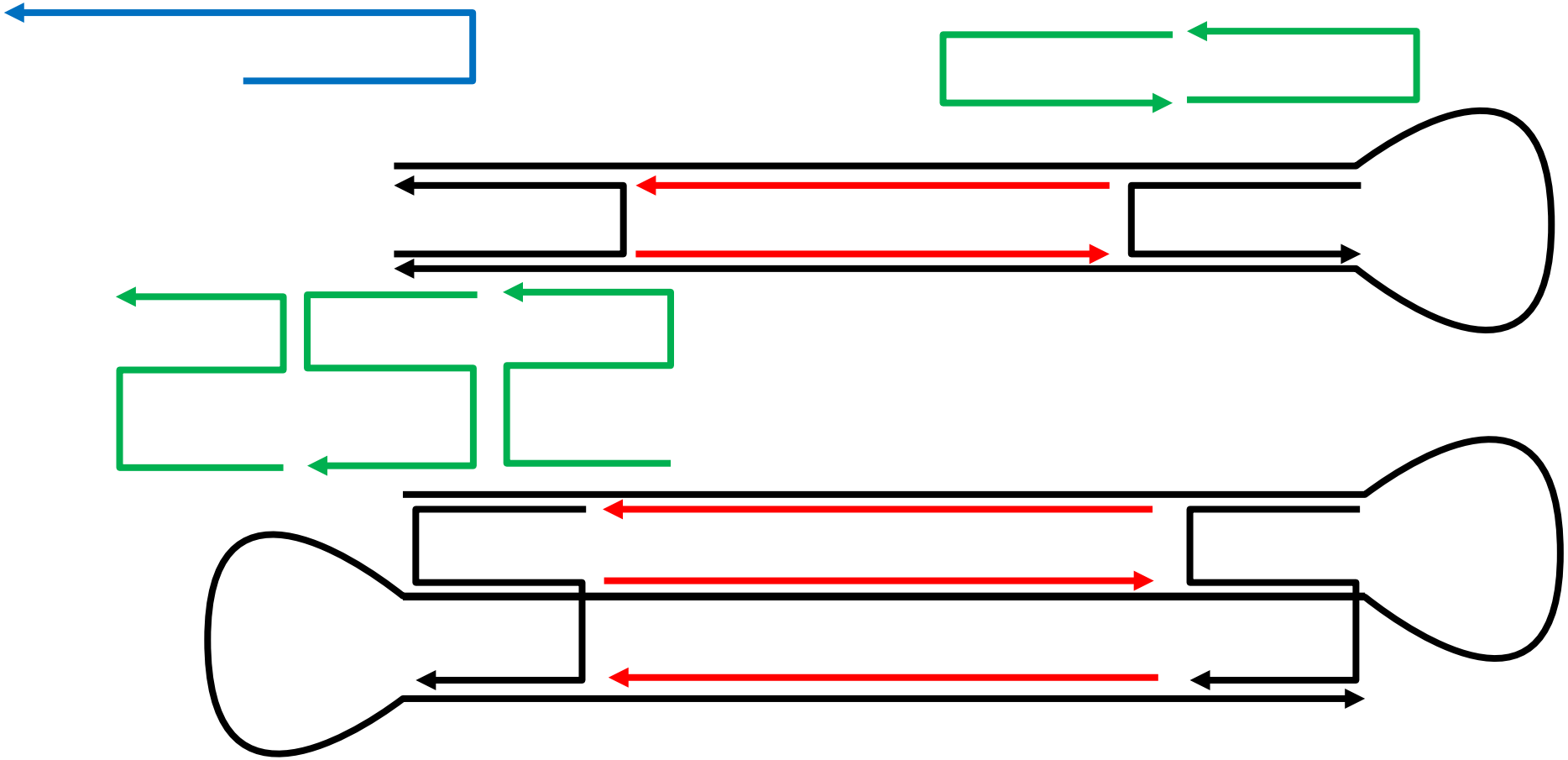
Feasible DNA implementation



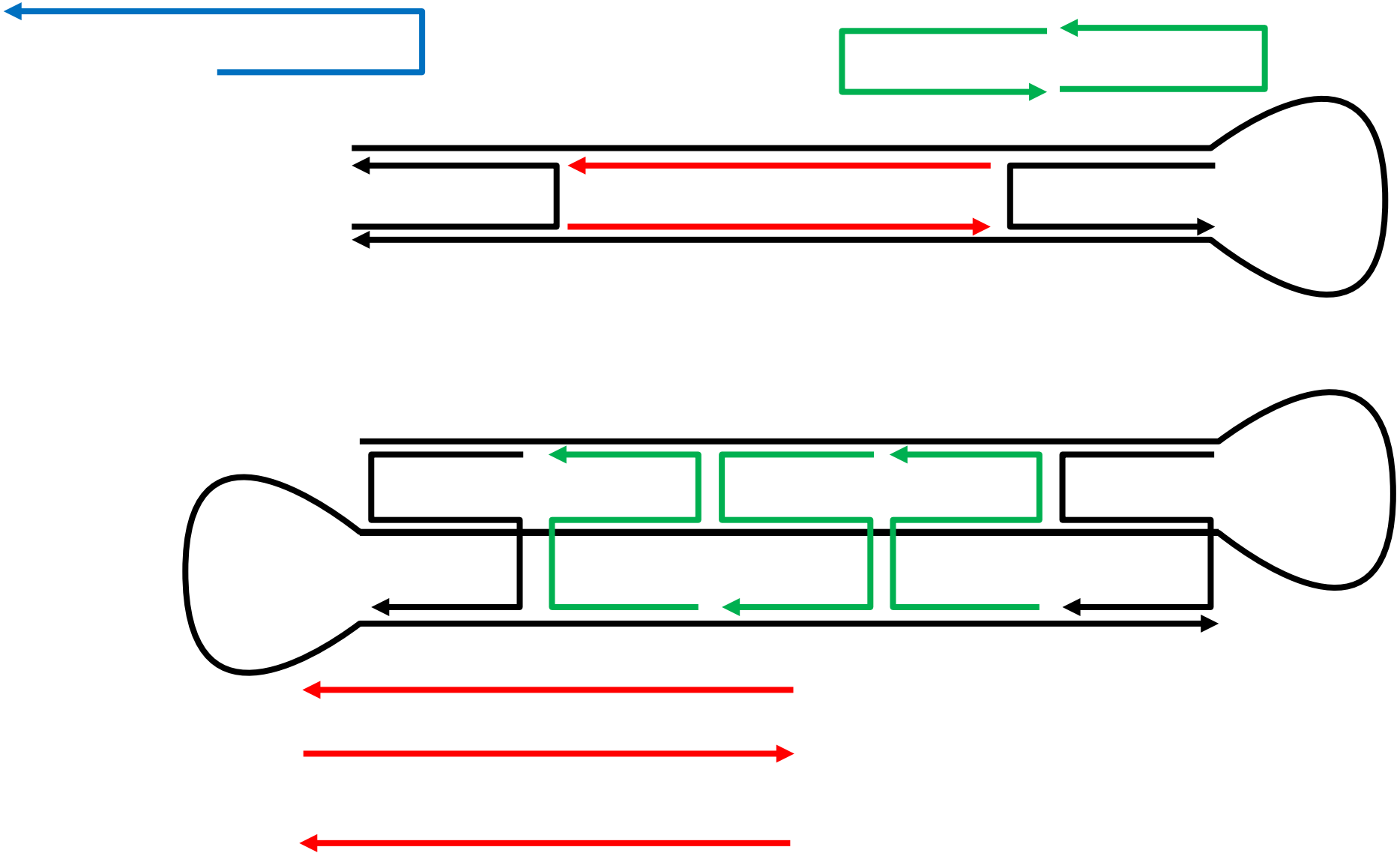
Feasible DNA implementation



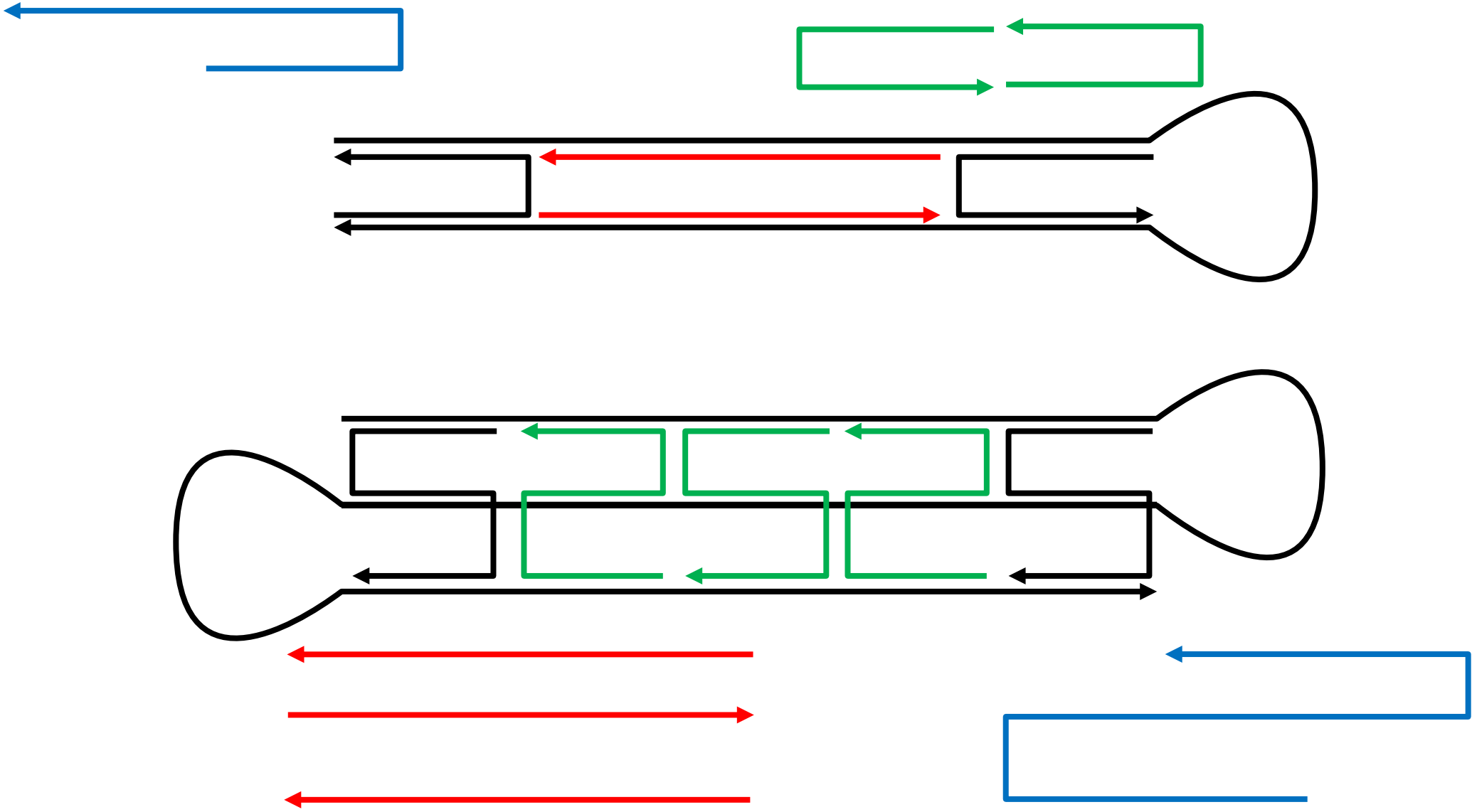
Feasible DNA implementation



Feasible DNA implementation



Feasible DNA implementation



Feasible DNA implementation