Error Correction for of Tiling Self-Assemblies

John Reif
(from papers by Winfree, Goel, Chen, Reif and their coauthors)

- # Algorithmic Self-assembly (without errors):

- crystal growth is programmed by designing a set of tiles with binding interactions enforce specific local assembly rules.

- Growth begins from a nucleating structure and consists of a series of attachments of single tiles.

- Under slightly supersaturated conditions, the attachment of a tile to a growing crystal is energetically favorable only if it attaches to a growing crystal by at least two binding sites.

- The tiles are designed so that during correct assembly, at every step a tile in the pattern attaches by a particular set of two or more binding sites.

- These binding sites are the tile's *inputs:*
the identities of the binding sites together determine which tile can attach at a given site.

Theoretical and Algorithmic Issues for Tiling Self-Assemblies

- **Efficiently assembling basic shapes with precisely controlled size and pattern**
  - **Constructing N X N squares with Ω(log n/log log n) tiles**
    [Adleman, Cheng, Goel, Huang, 2001]
  - **Perform universal computation by simulating BCA**
    [Winfree’ 99]
- **Library of primitives to use in designing nano-scale structures** [Adleman, Cheng, Goel, Huang, 2001]
- **Automate the design process**
  [Adleman, Cheng, Goel, Huang, Kempe, Moisset de espanes, Rothemund 2001]
- **Robustness**

(Cheng, Goel, Cheng)

Robustness of Tiling Self-Assemblies

- In practice, self-assembly is a thermodynamic process. When T=2, tiles with 0 or 1 matches also attach; tiles held by total strength 2 also fall off at a small rate.
- Currently, there are 1-10% errors observed in experimental self-assembly.
  [Winfree, Bekbolatov, ’03]
- Possible schemes for error correction
  - Biochemistry tricks
  - Coding theory and error correction

[Cheng, Goel, Cheng]

Types of Errors of Algorithmic self-assembly:

- self-assembly is stochastic:

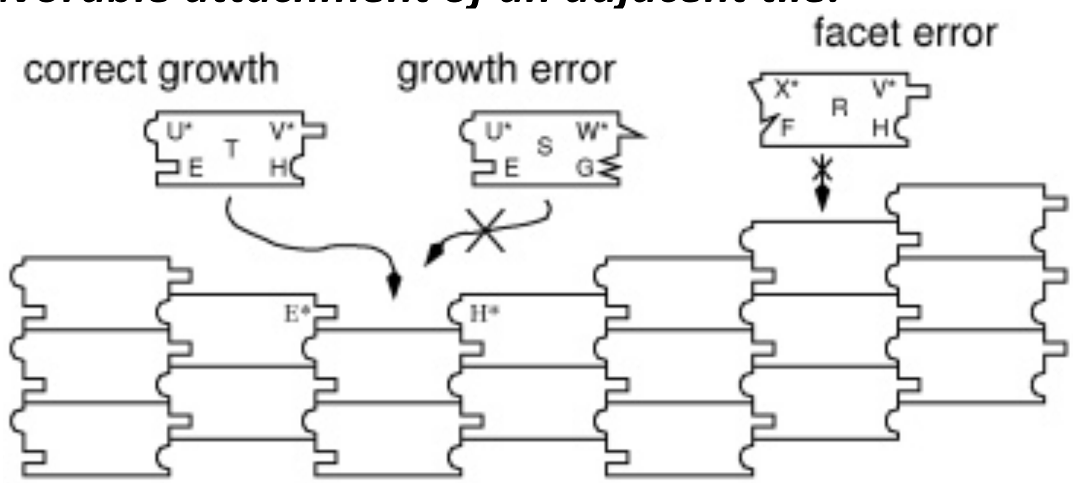
- unfavorable attachments of tiles with one or more incorrect or absent inputs also occur.

Tiling Error: ***when a tile that attaches unfavorably does not match some of its input binding sites, so may not be the correct tile in the desired pattern***

- Subsequent algorithmic pattern formation can be severely disrupted, resulting in a grossly malformed product.

Errors of insufficient attachment:

- While tiles that attach unfavorably usually fall off quickly, occasionally such a tile is locked in by the subsequent favorable attachment of an adjacent tile.



[Winfree 2007]

Illustration of Correct and Erroneous assembly steps.

Growth errors:

Insufficient attachments at sites where a correct tile could have attached:
involve both a correctly matching binding site and a mismatch.
Insufficient attachments on facets involve no mismatches; nonetheless, incorrect.

Modelling Self-Assembly Errors

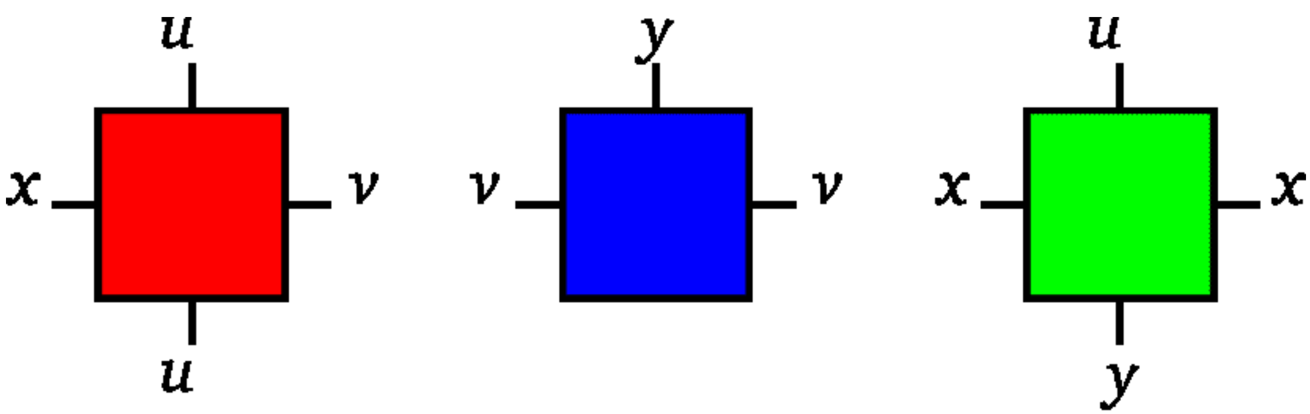
- **Temperature**: A positive integer giving number of attachements needed for assembly of a tile. (Usually 1 or 2)
- **A set of tile types**: Each tile is an oriented rectangle with glues on its corners. Each glue has a non-negative strength (0, 1 or 2).
- **An initial assembly (seed)**.

- **Rules**: A tile can attach to an assembly iff the combined strength of the "matched glues" is greater or equal than the temperature.
- Tiles with combined strength equal to temperature can fall off.
- **Errors**: Once a while, there will be two tiles attach at the same time and both are held by strength at least two after the attachment.
- We call this an "insufficient attachment".

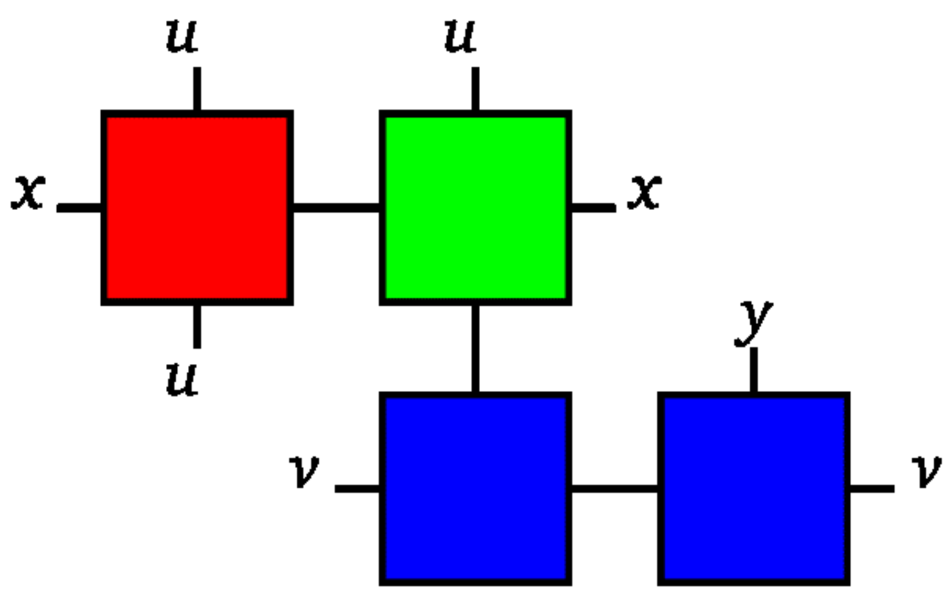
- **Our goal**: minimize the impact of insufficient attachments

[Chen]

Example of Error-Free Self-Assembly

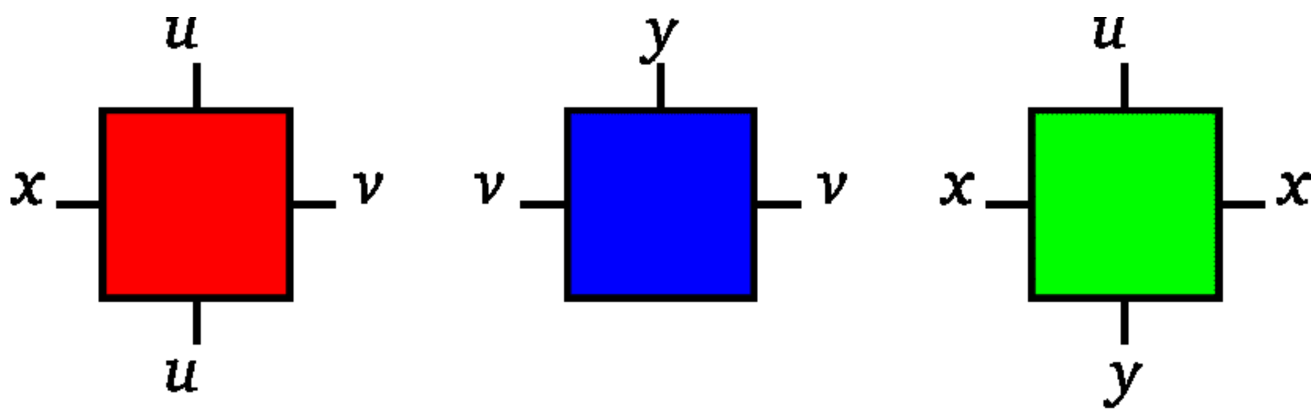


T=2

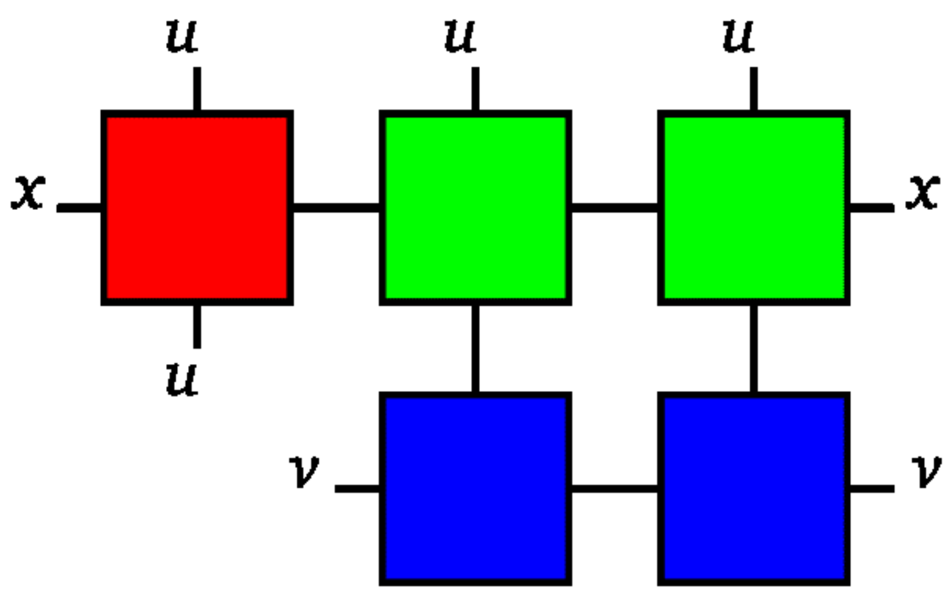


[Cheng, Goel, Cheng]

Example of Error-Free Self-Assembly

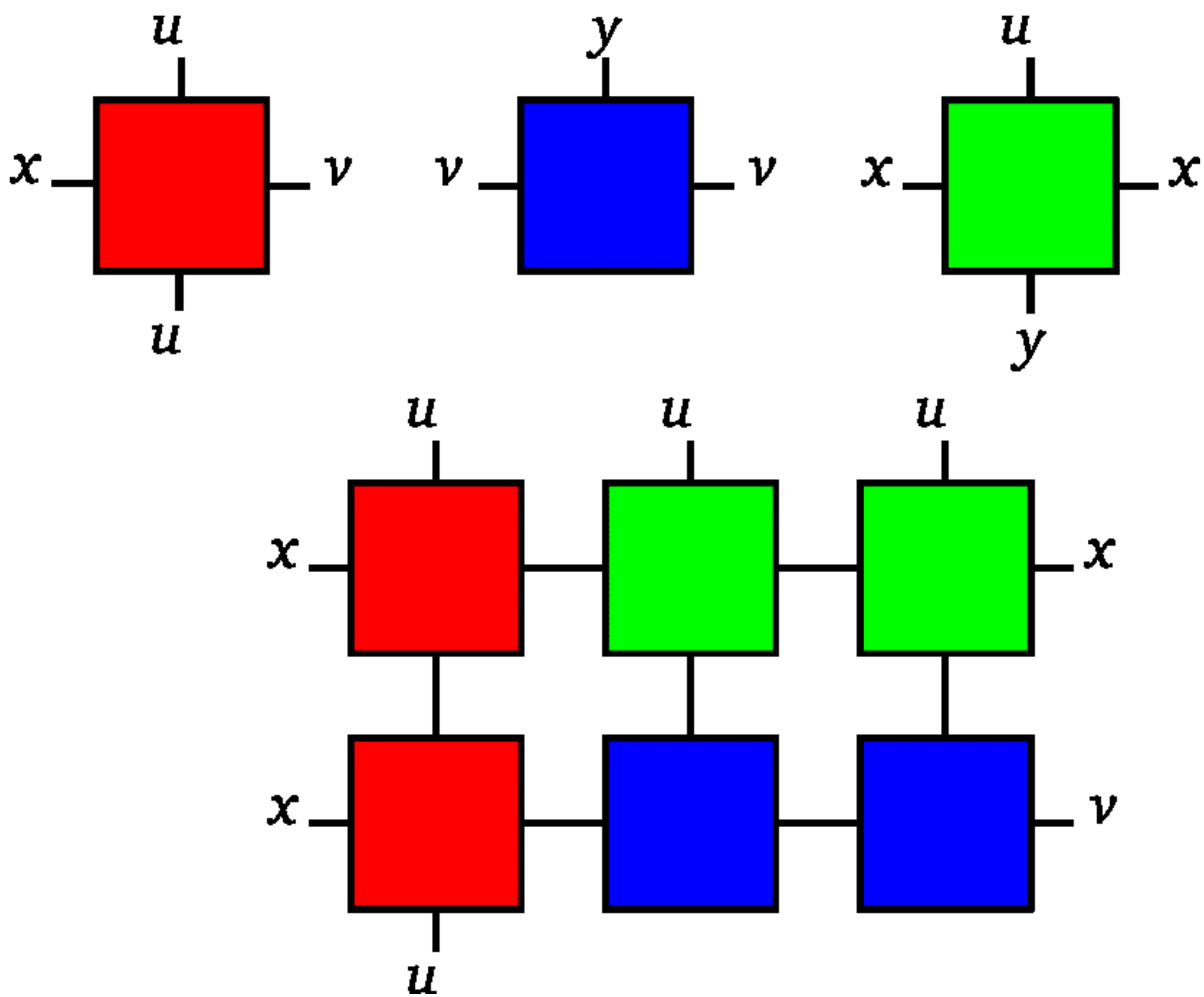


T=2



[Cheng, Goel, Cheng]

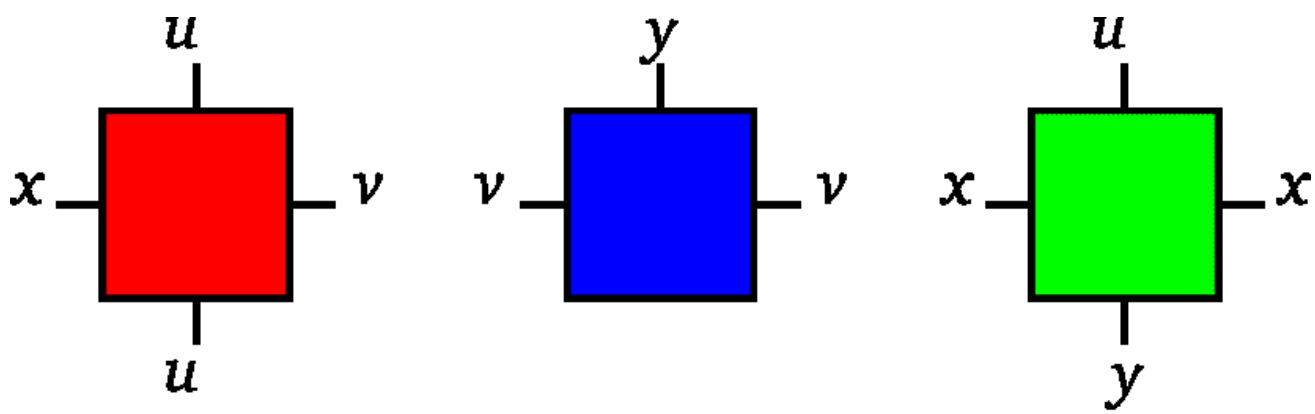
Example of Error-Free Self-Assembly



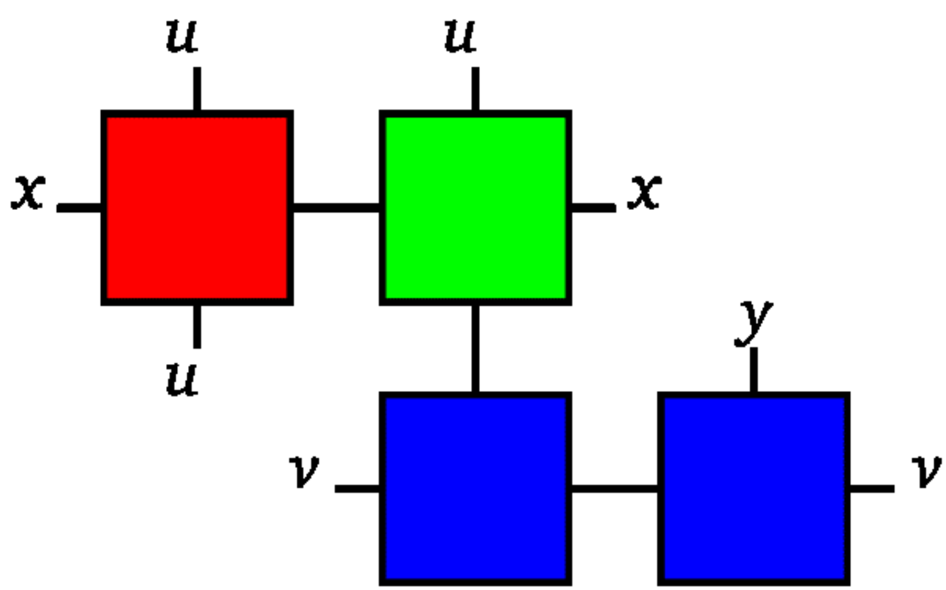
T=2

[Cheng, Goel, Cheng]

What can go wrong in a Self-Assembly?

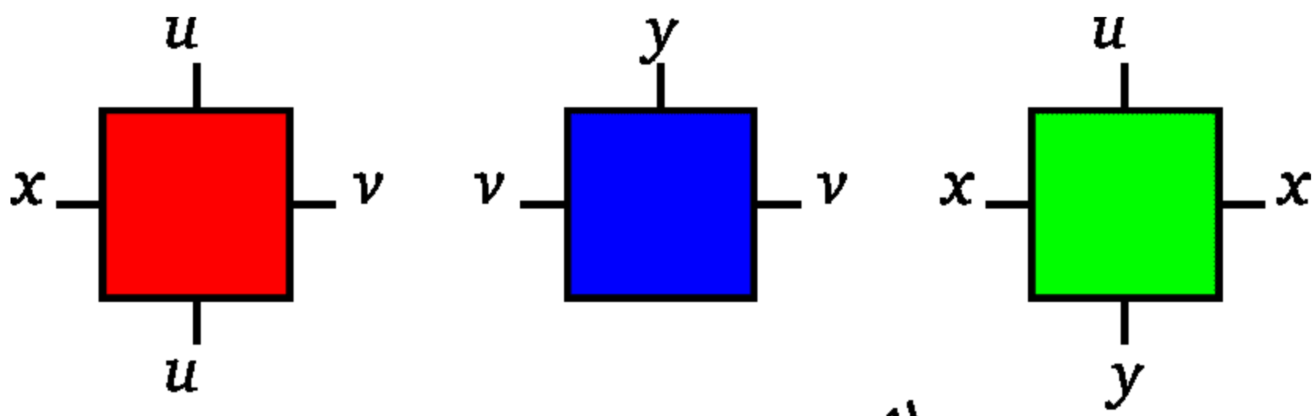


T=2

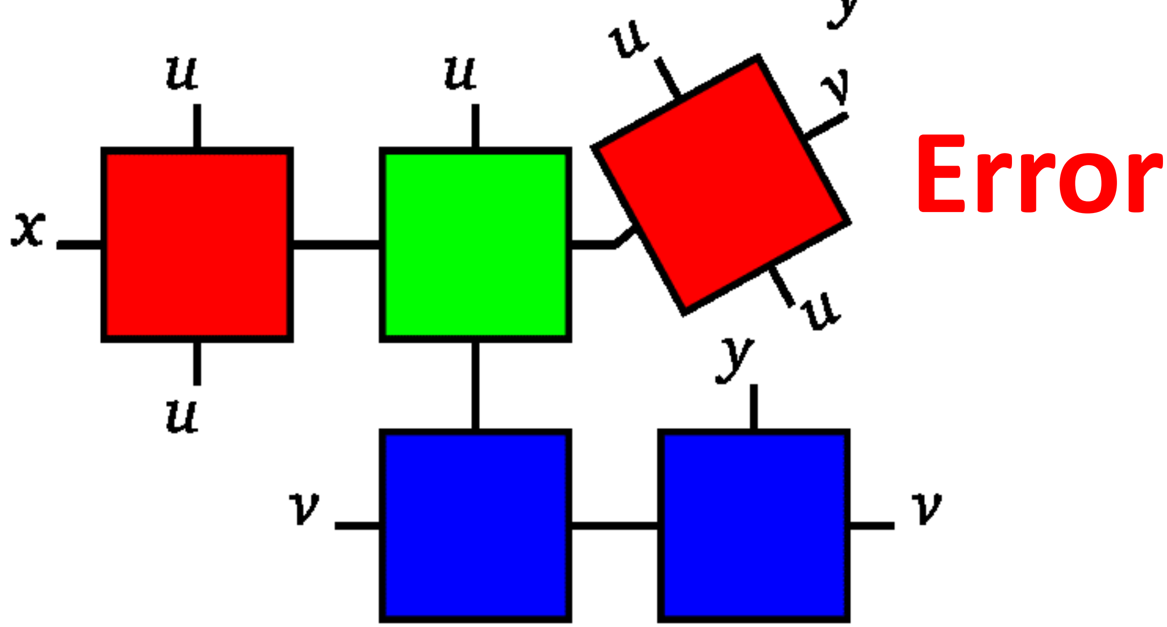


[Cheng, Goel, Cheng]

What can go wrong in a Self-Assembly?

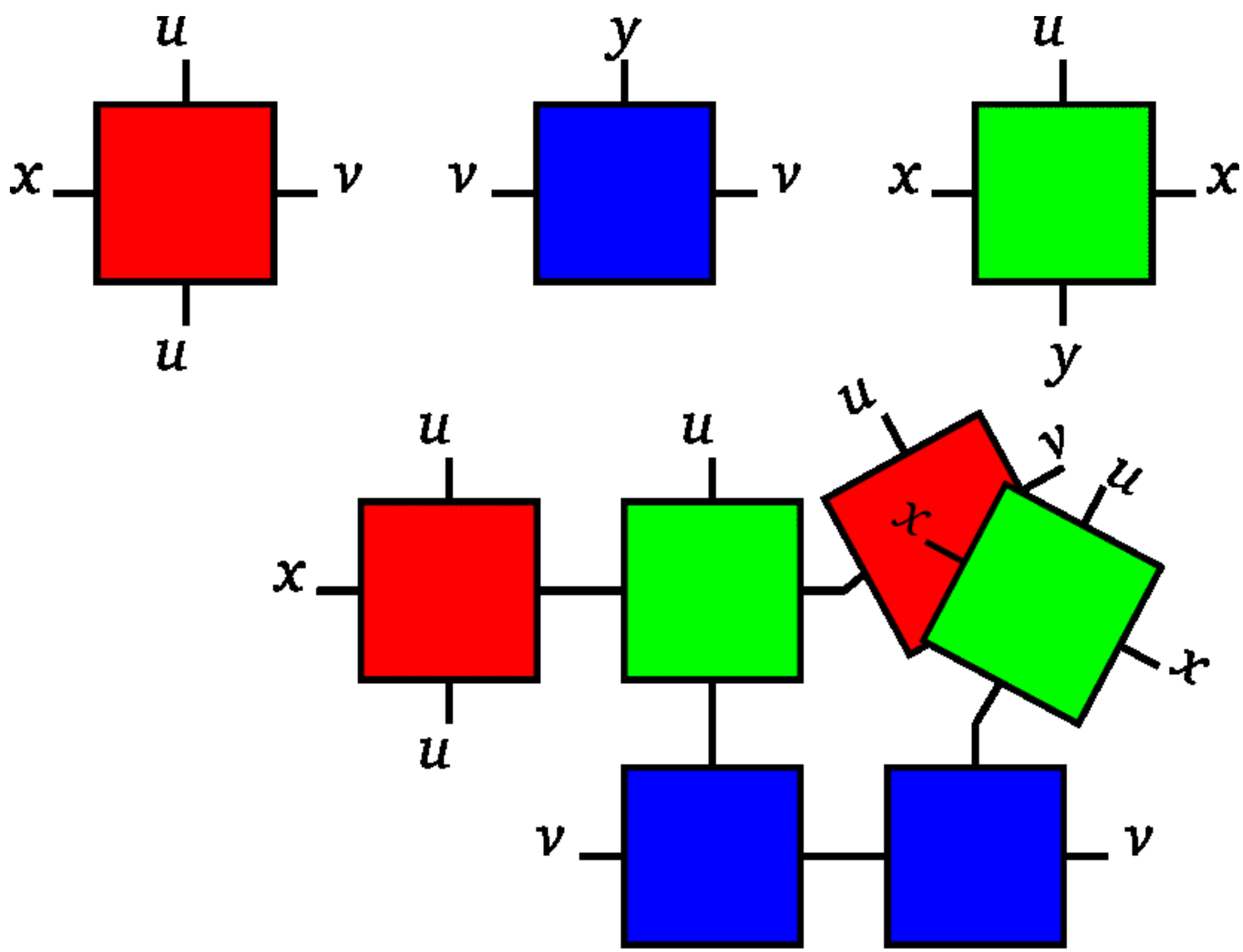


T=2



[Cheng, Goel, Cheng]

Why it may not matter:

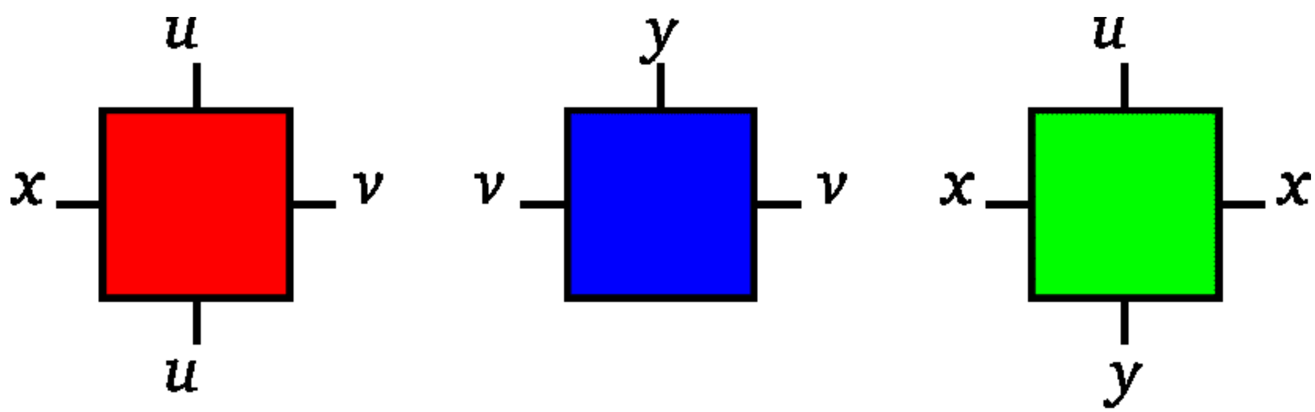


T=2

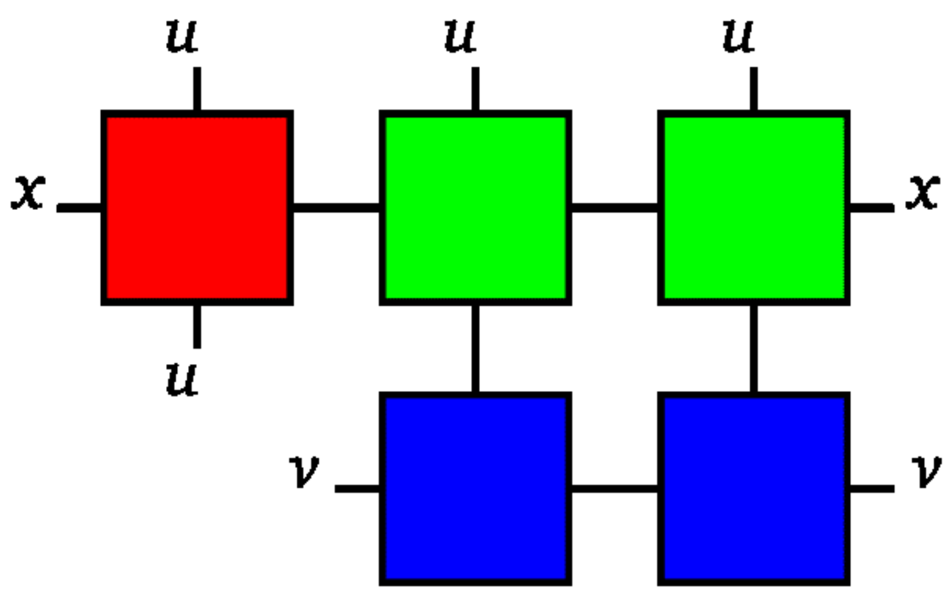
More Errors

[Cheng, Goel, Cheng]

Why errors of Self-assembly may not matter:

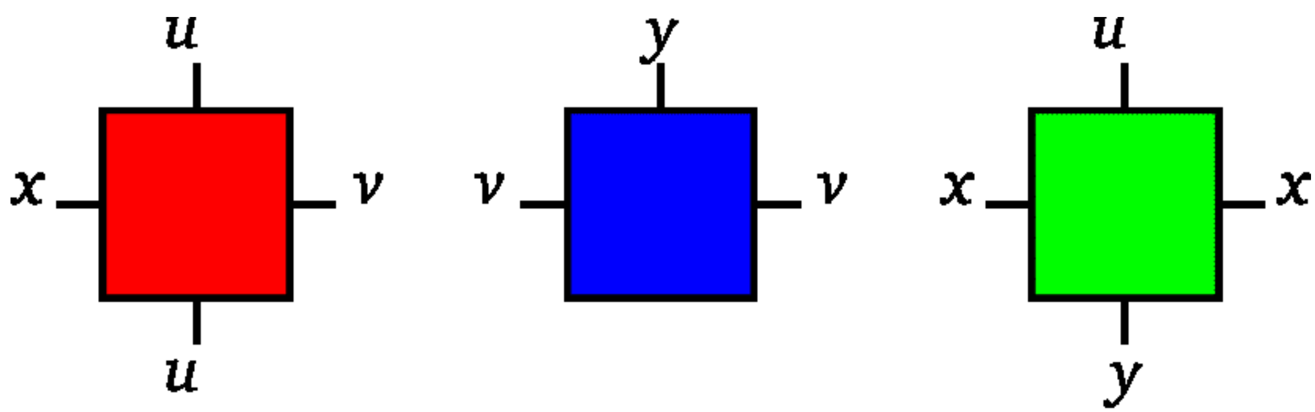


T=2

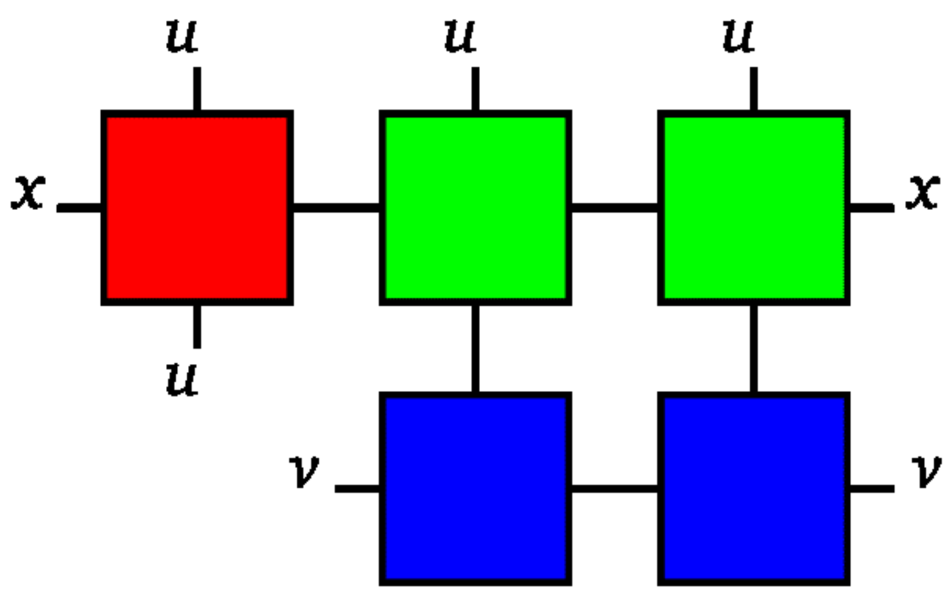


[Cheng, Goel, Cheng]

What can go really wrong in a Self-Assembly?



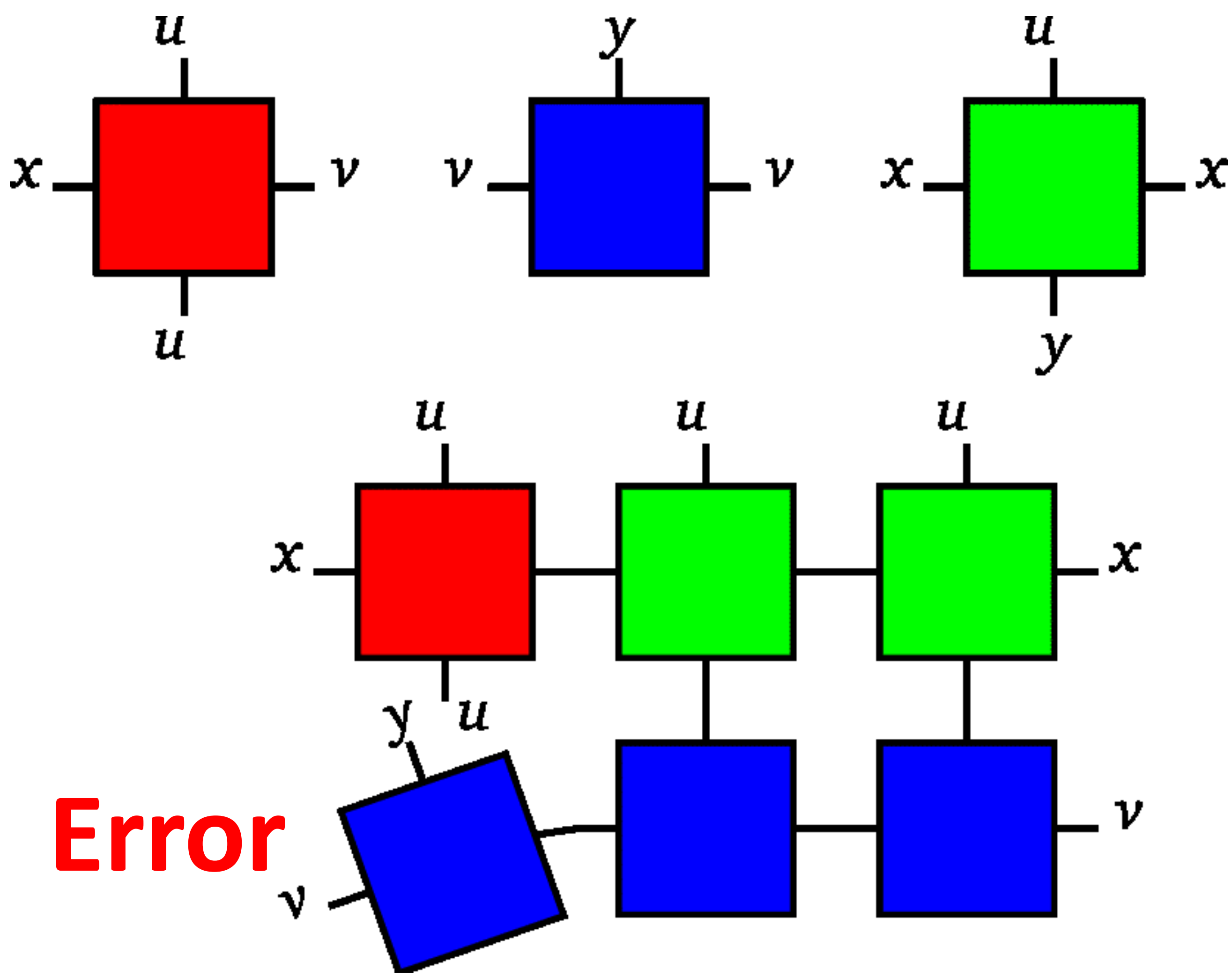
T=2



[Cheng, Goel, Cheng]

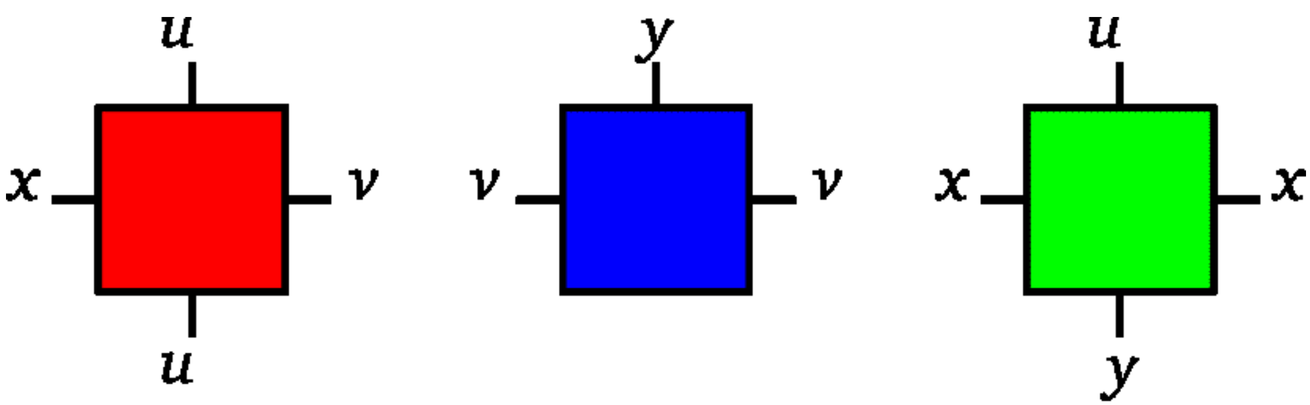
What can go really wrong in a Self-Assembly?

T=2

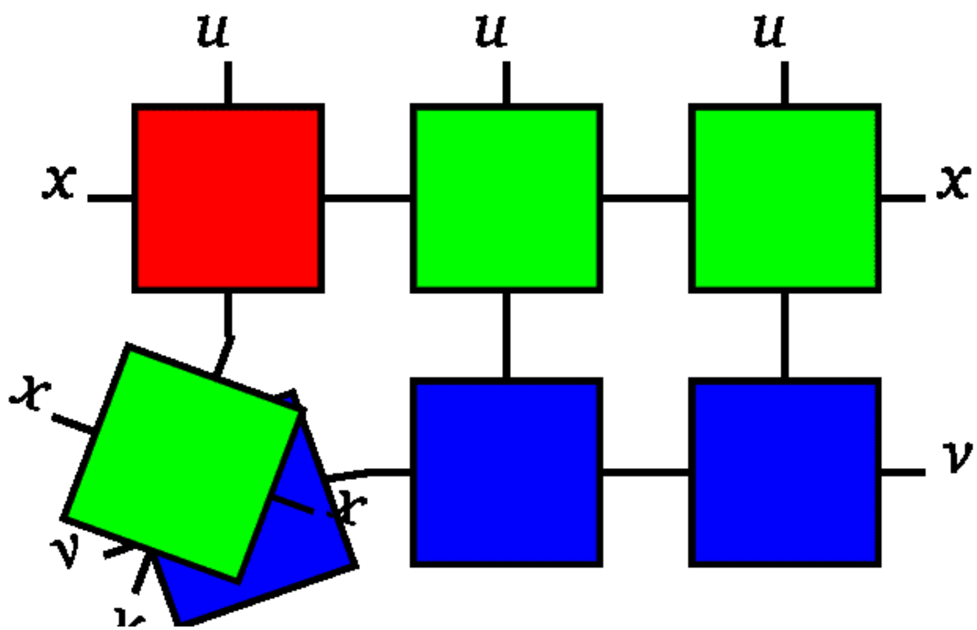


[Cheng, Goel, Cheng]

What can go *really* wrong?



T=2



More Errors

[Cheng, Goel, Cheng]

Error-Reducing Designs

- Biochemistry tricks
  - Strand Invasion mechanisms

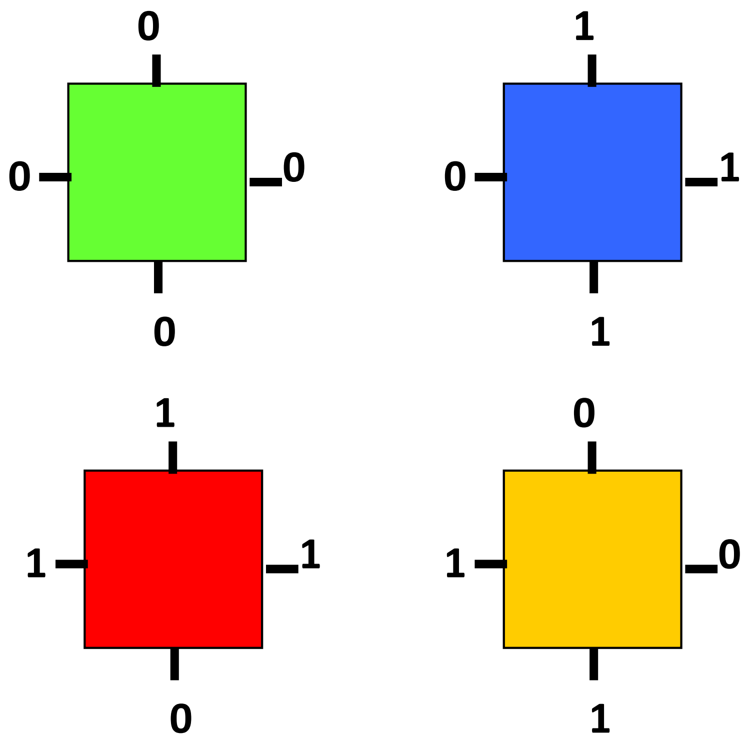
    [Chen, Cheng, Goel, Huang, Moisset de espanes, 2004]
- Coding theory and error correction
  - Proofreading tiles

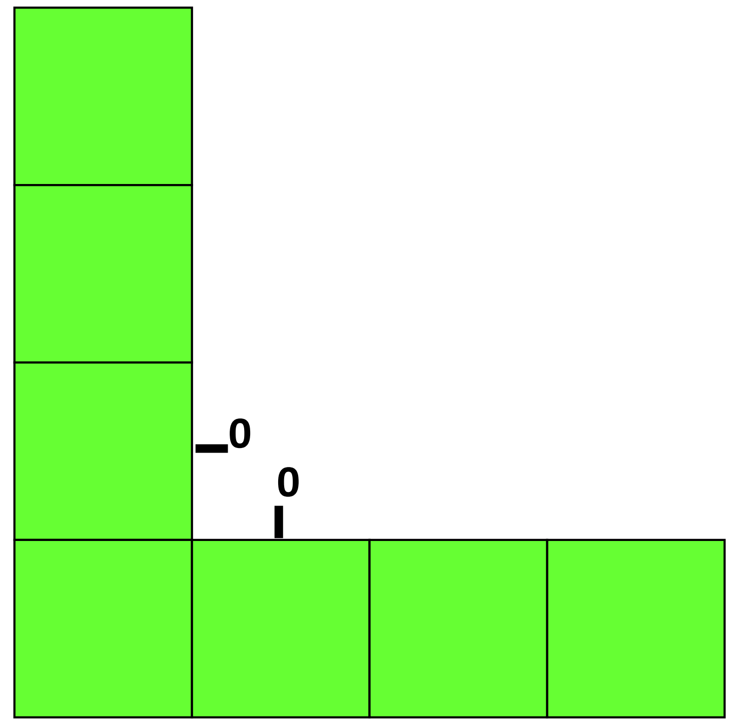
    [Winfree, Bekbolatov,2003]
  - Snake tiles

    [Chen, Goel 2004]
  - Compact Redundant Tiles

    [Reif,Sahu,Yin 2004]

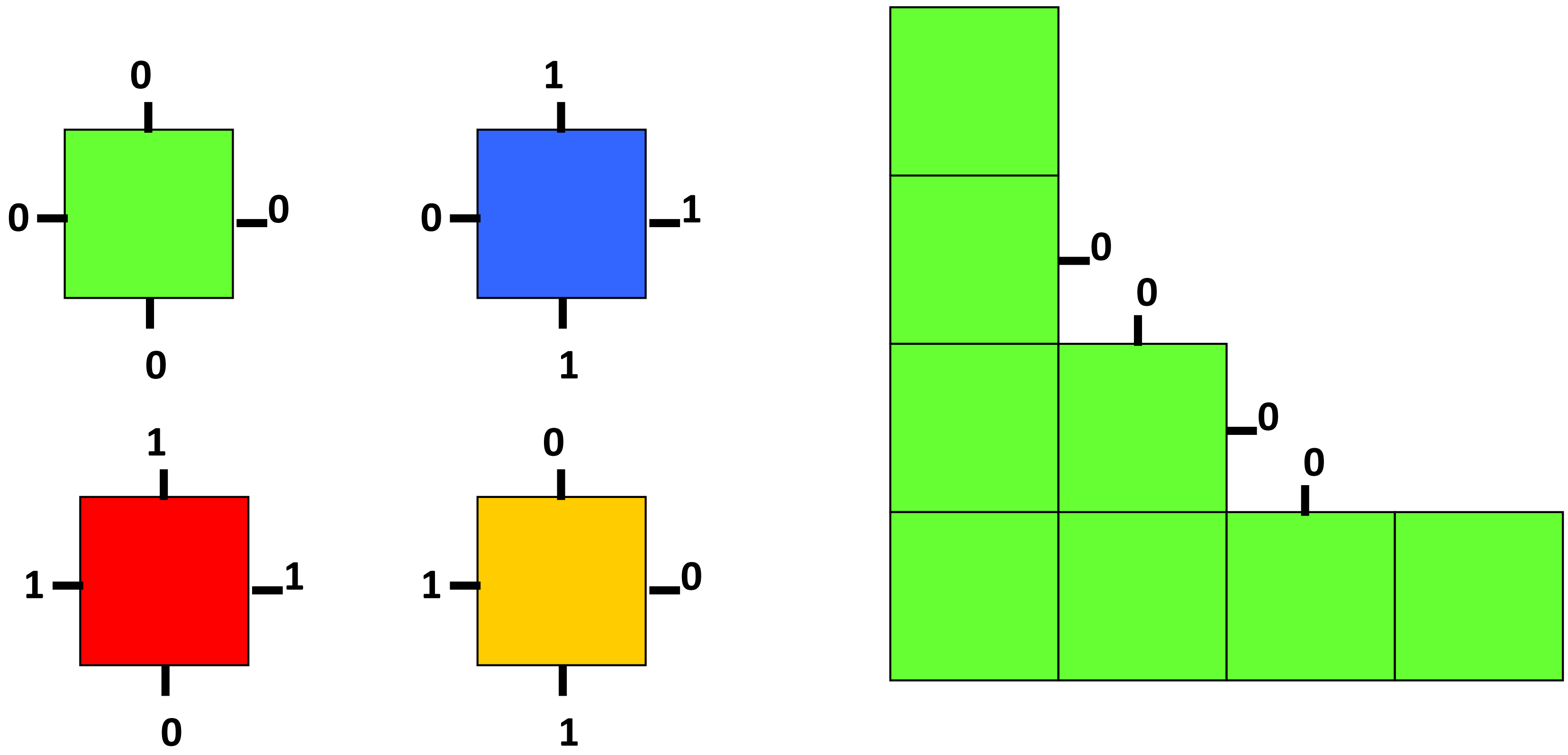
Example: Sierpinski Tile System





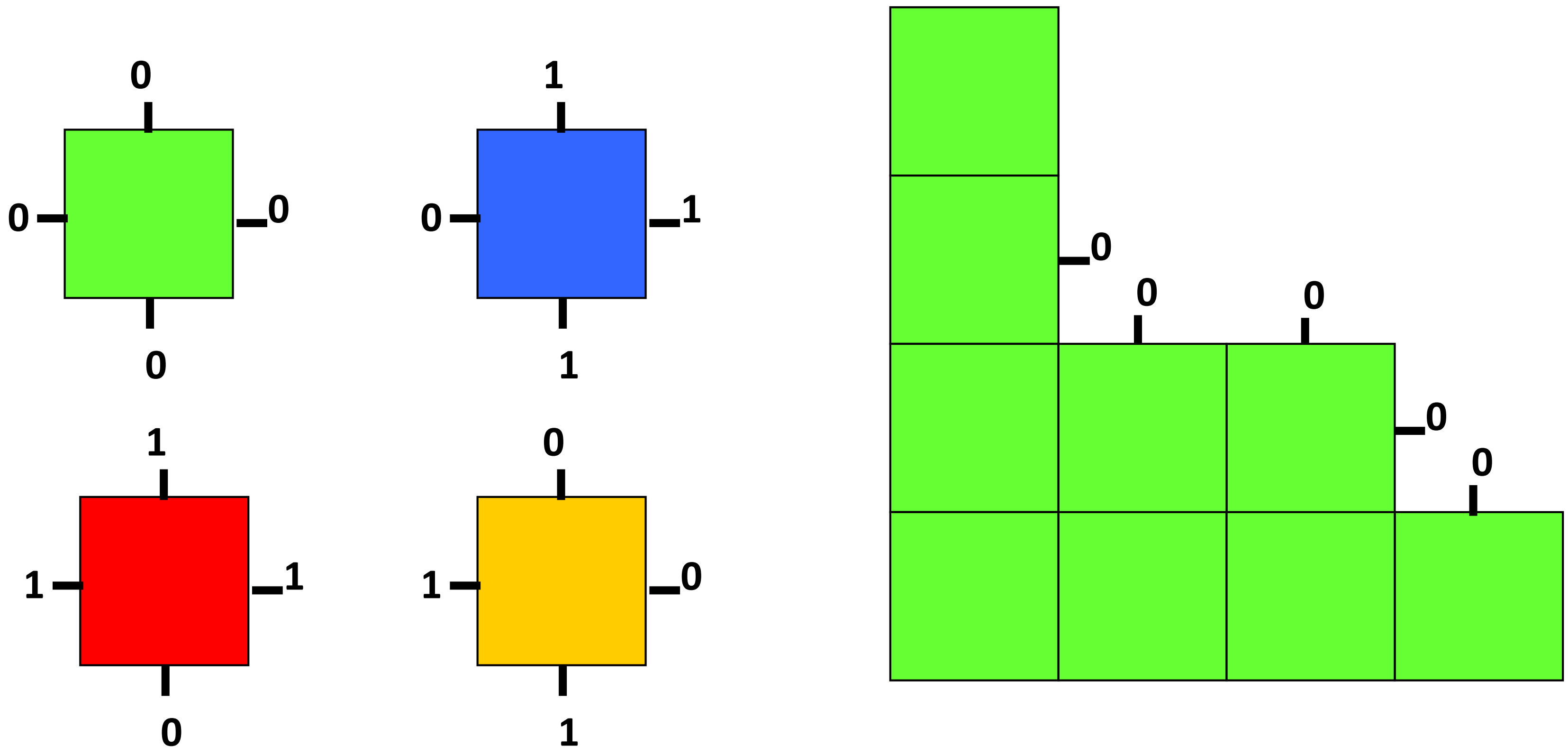
[Cheng, Goel, Cheng]

Example: Sierpinski Tile System



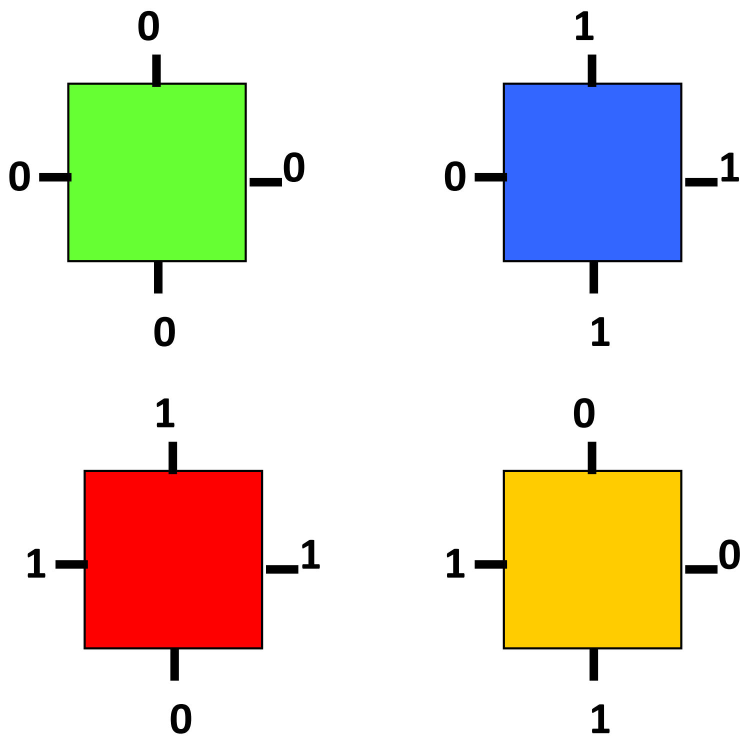
[Cheng, Goel, Cheng]

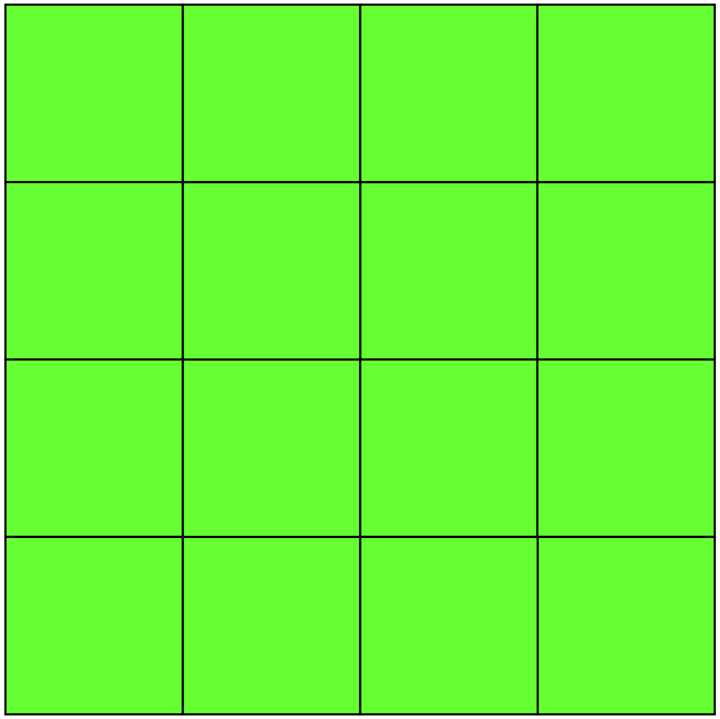
Example: Sierpinski Tile System



[Cheng, Goel, Cheng]

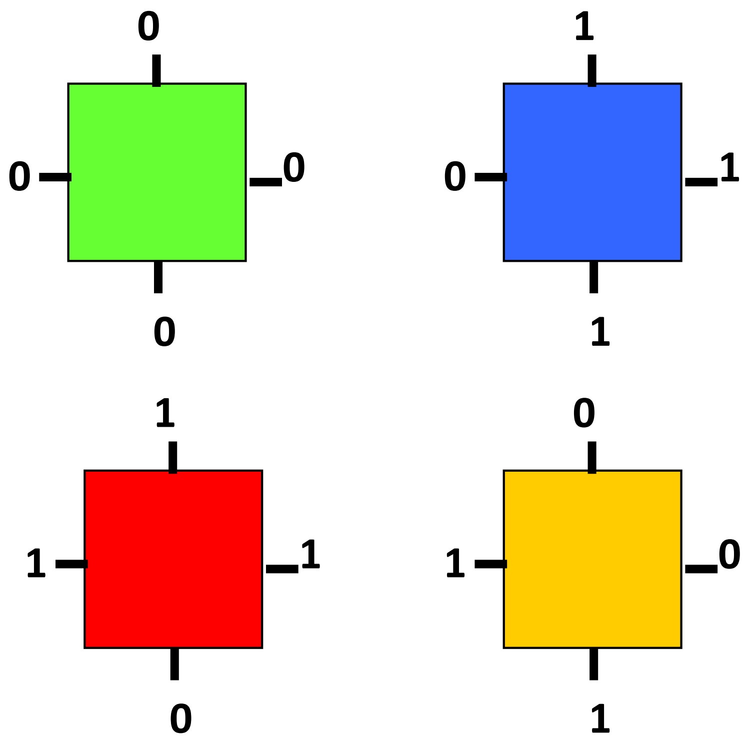
Example: Sierpinski Tile System correct self-assembly

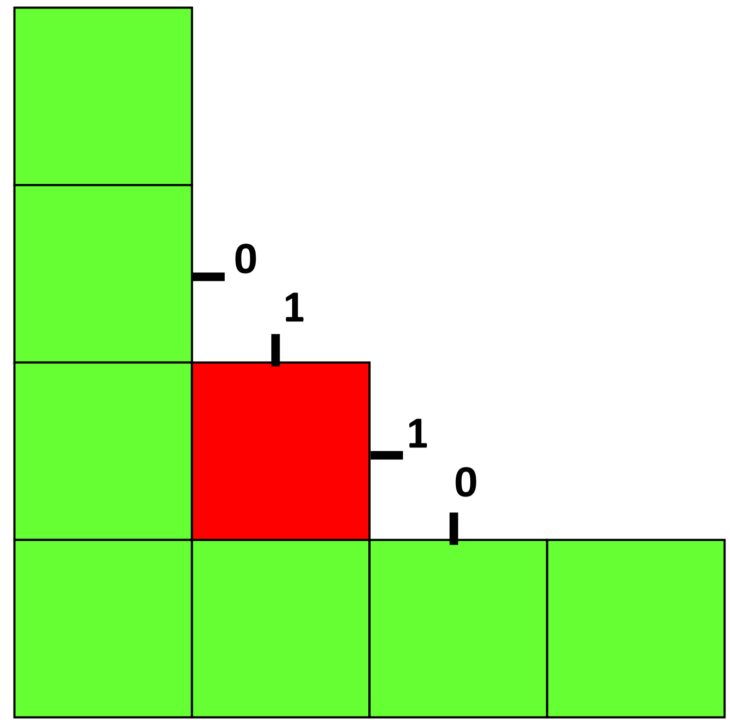




[Cheng, Goel, Cheng]

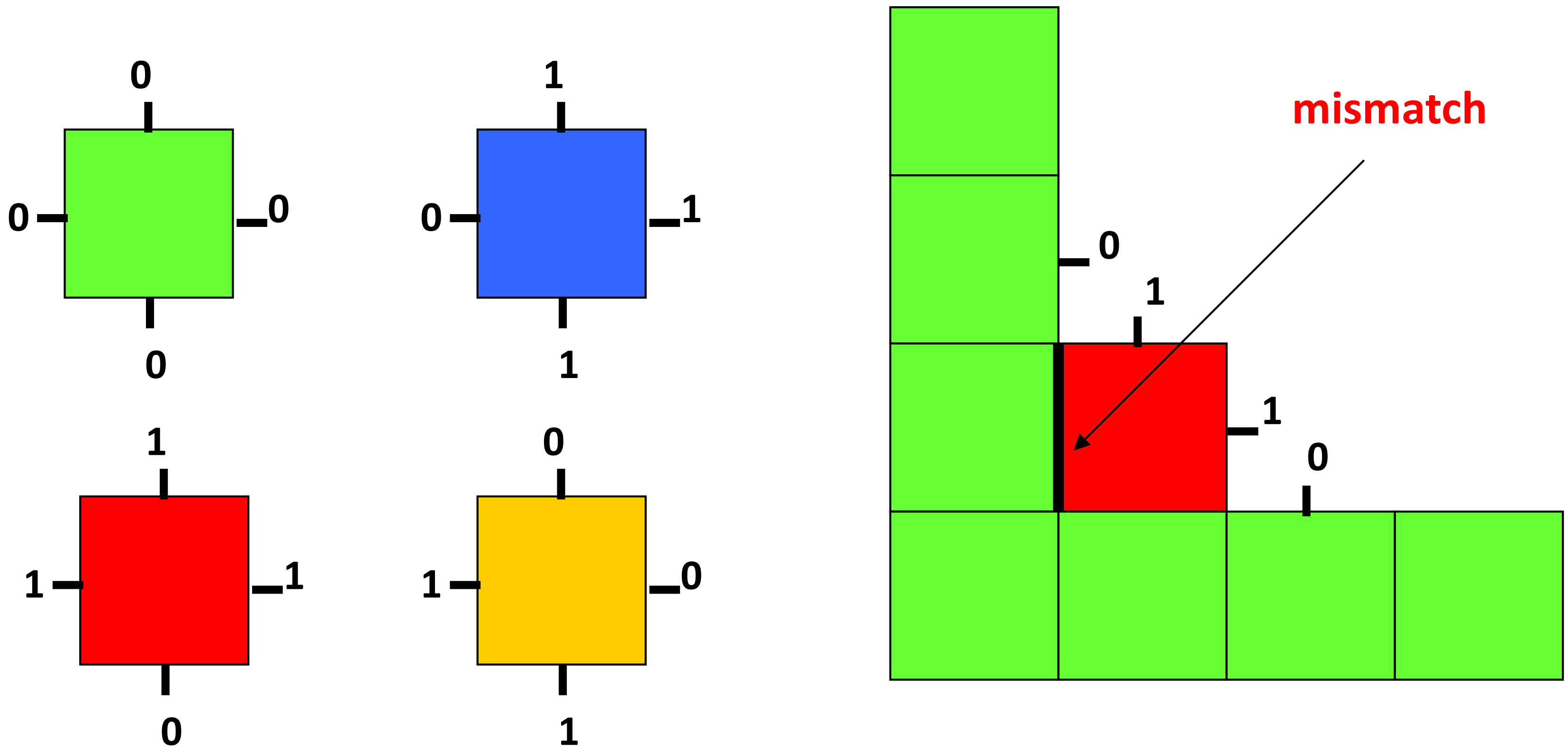
Crystallization Errors





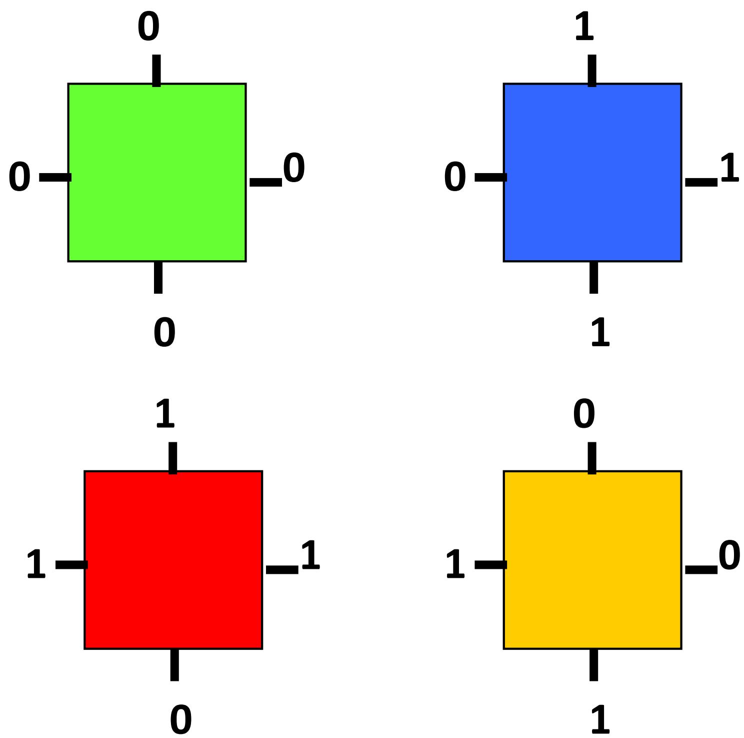
[Cheng, Goel, Cheng]

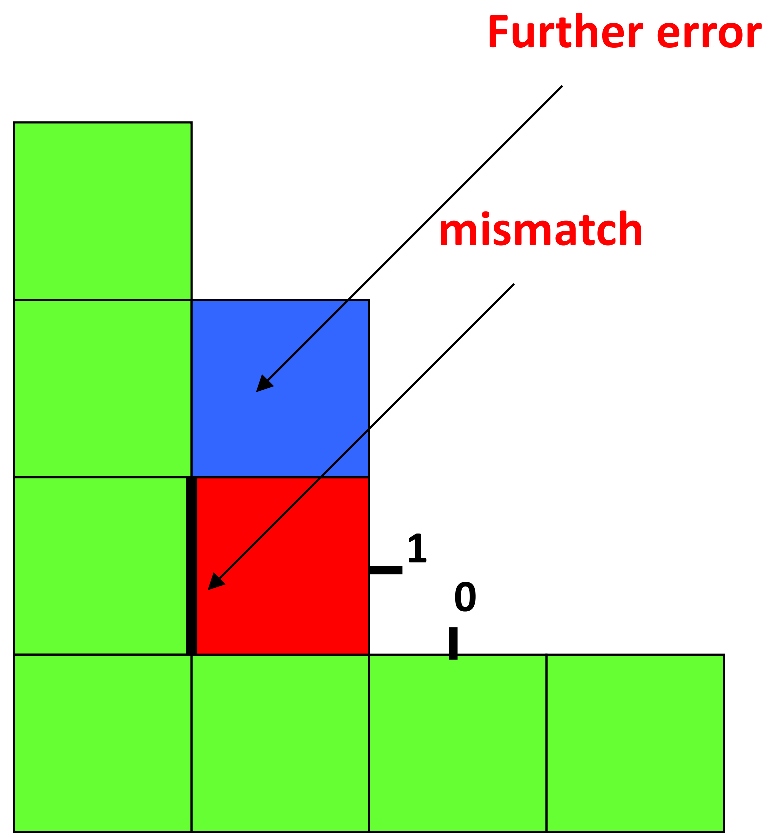
Crystallization Errors



[Cheng, Goel, Cheng]

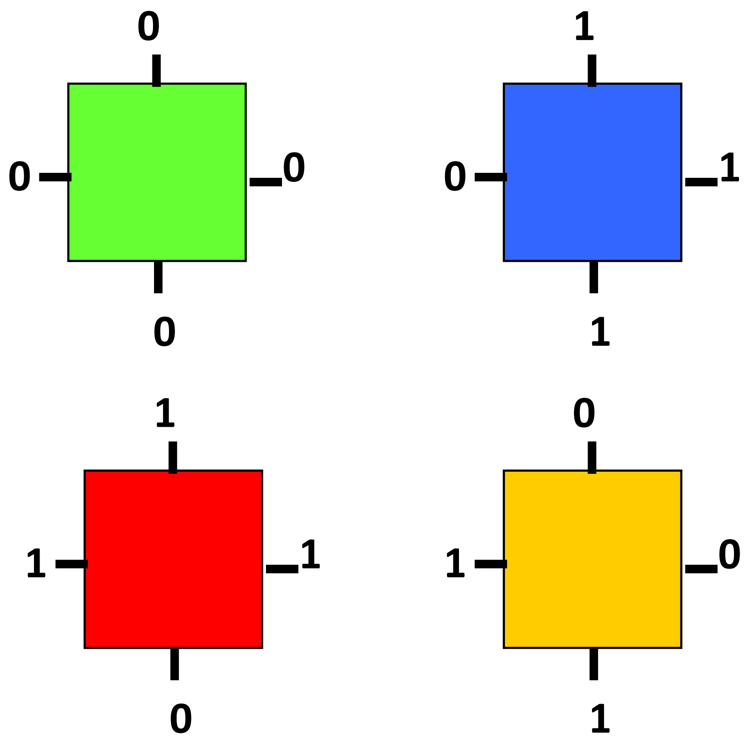
Crystallization Errors

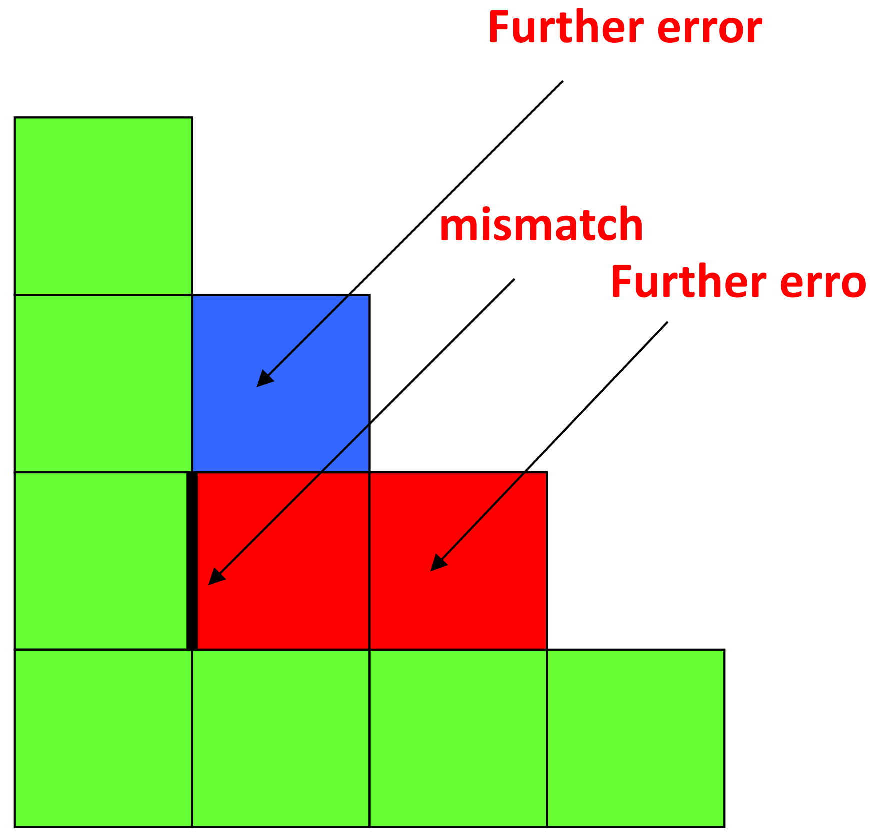




[Cheng, Goel, Cheng]

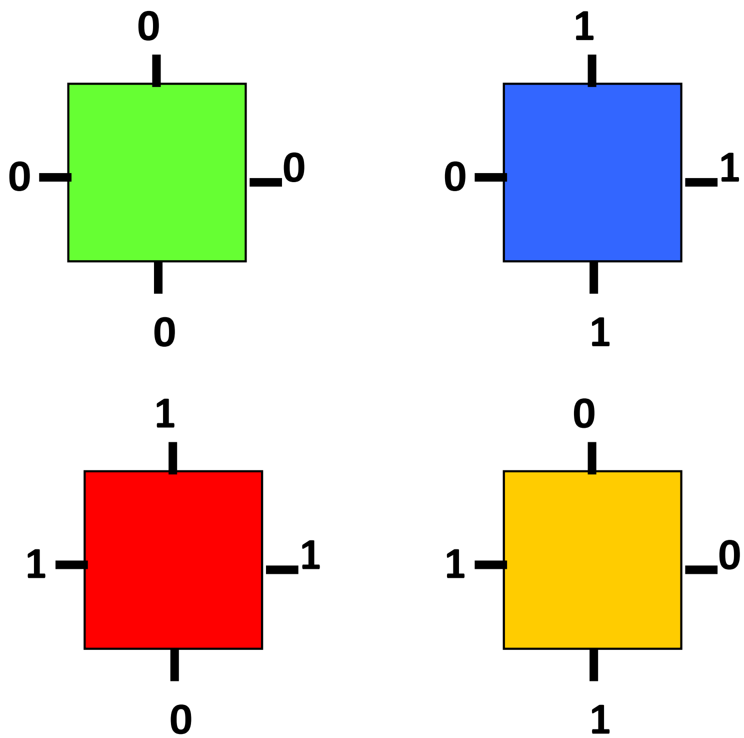
Crystallization Errors

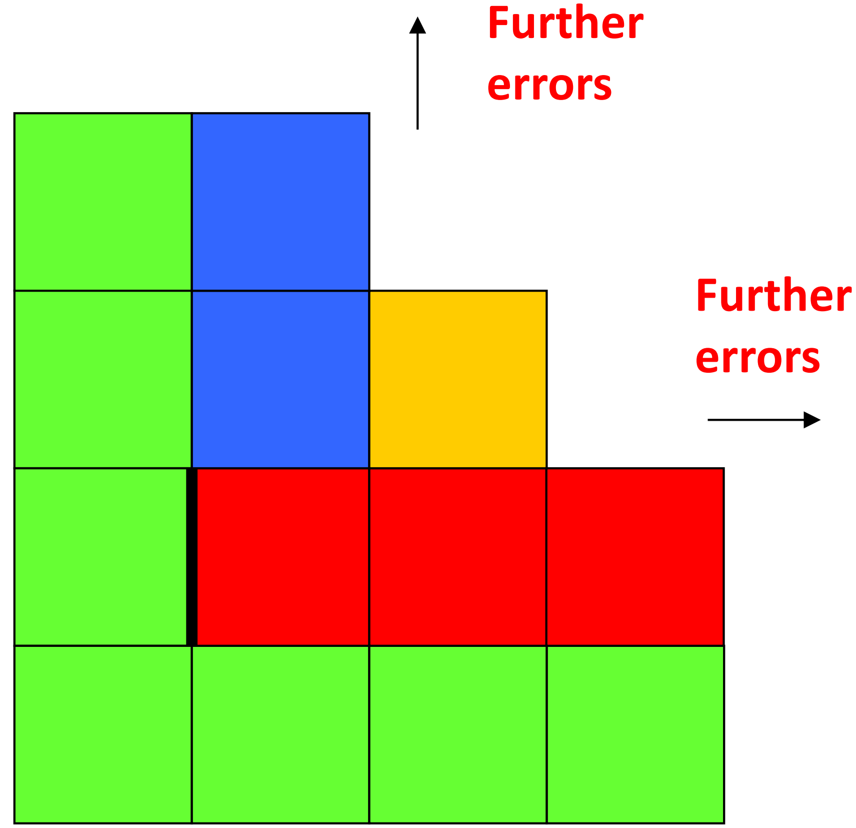




[Cheng, Goel, Cheng]

Crystallization Errors





[Cheng, Goel, Cheng]

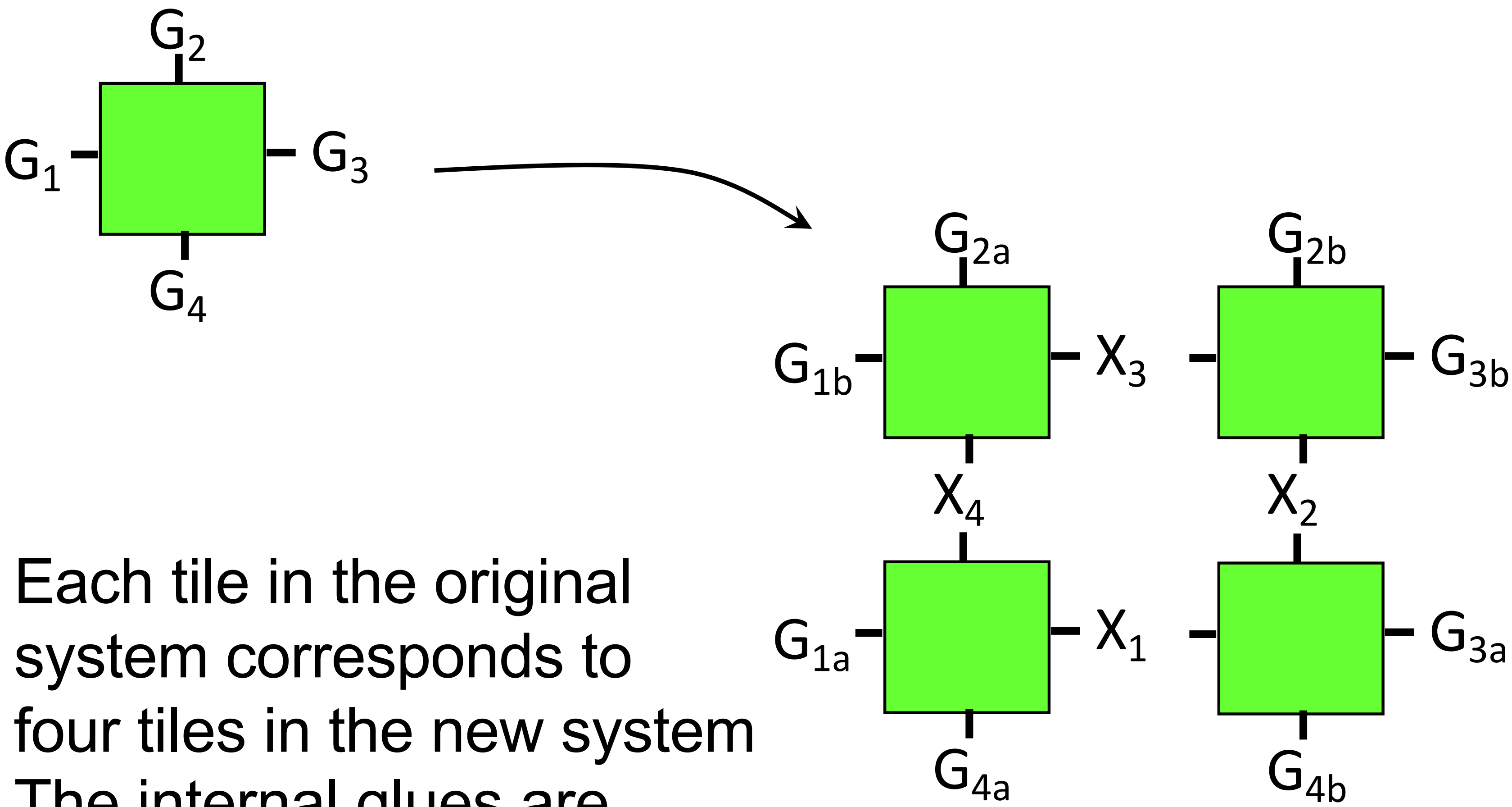
Proofreading Tile Sets: Error Correction for Algorithmic Self-Assembly

Erik Winfree and Renat Bekbolatov

2003

Proofreading Tiles

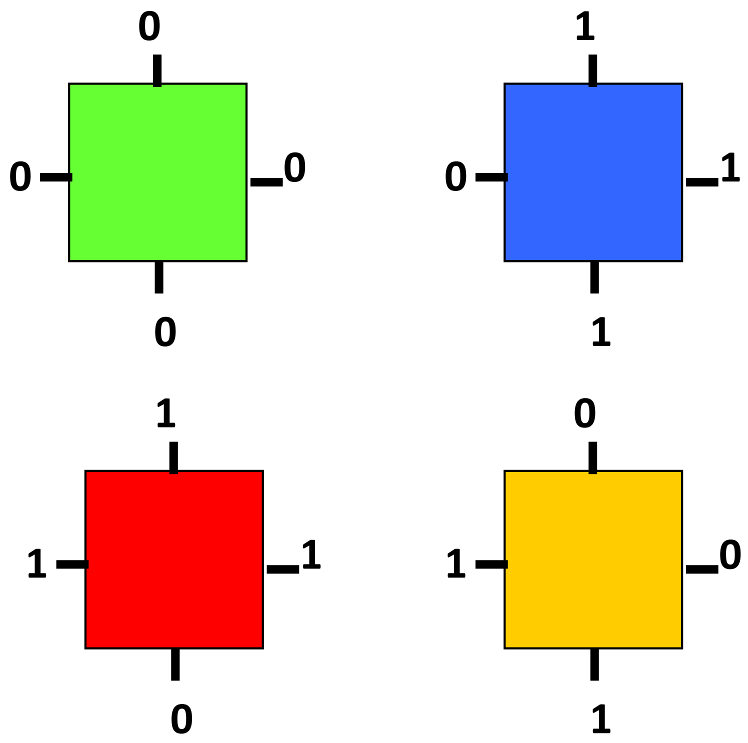
[Winfree, Bekbolatov, 2003]

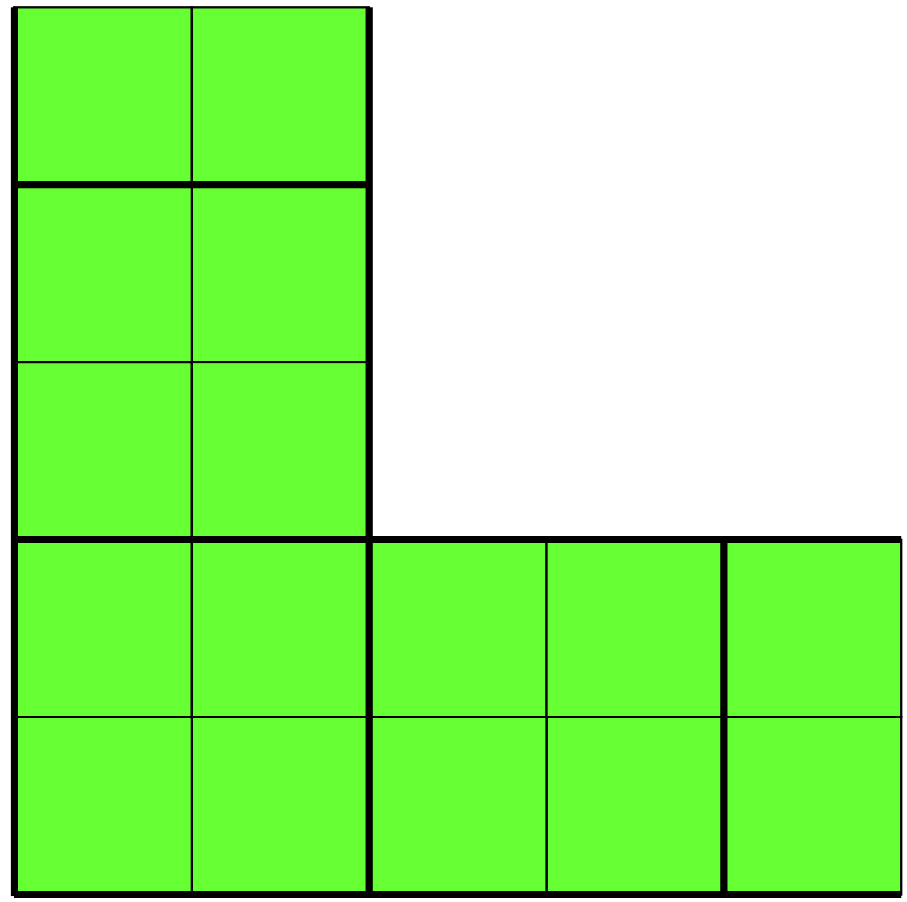


- Each tile in the original system corresponds to four tiles in the new system
- The internal glues are unique to this block

[Cheng, Goel, Cheng]

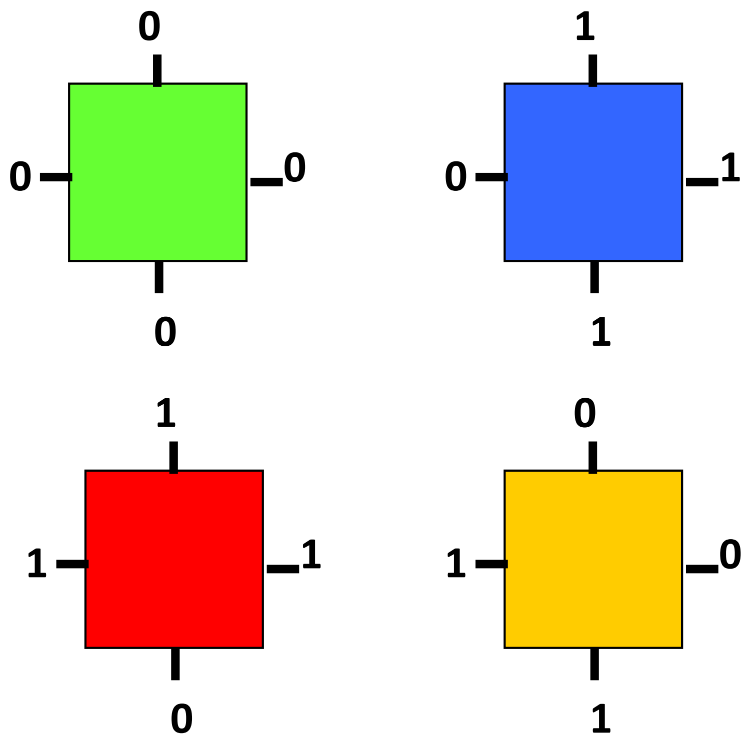
How Proofreading Tiles Reduce Errors of self-assembly

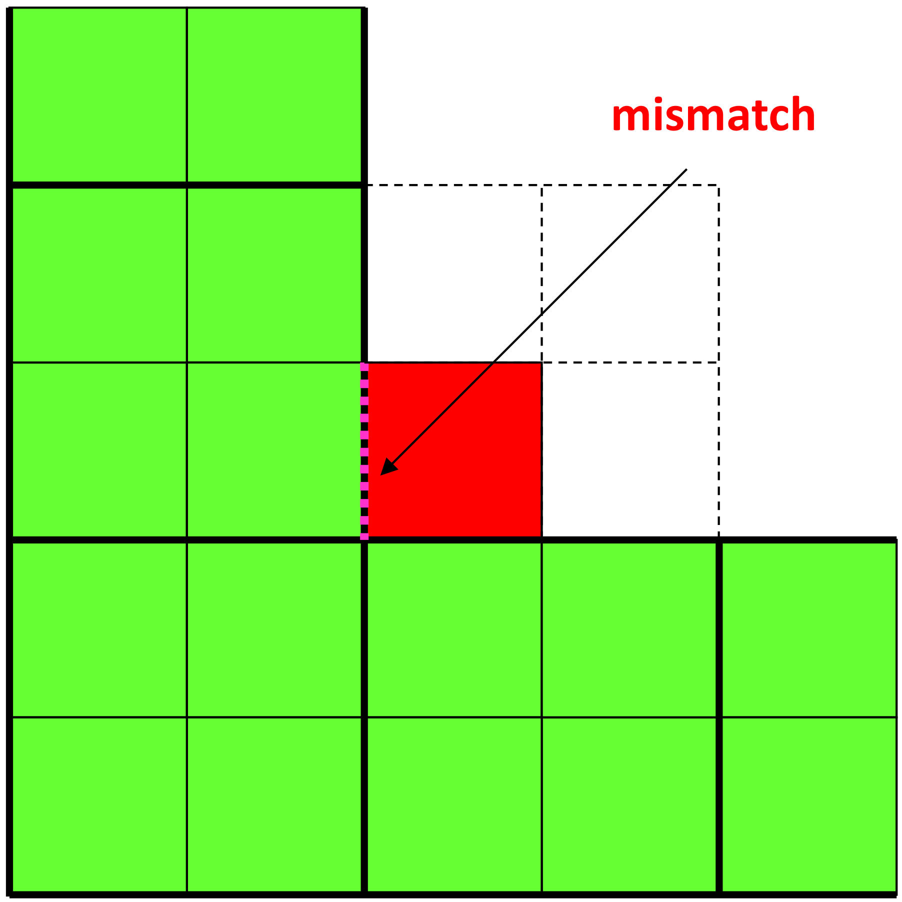




[Cheng, Goel, Cheng]

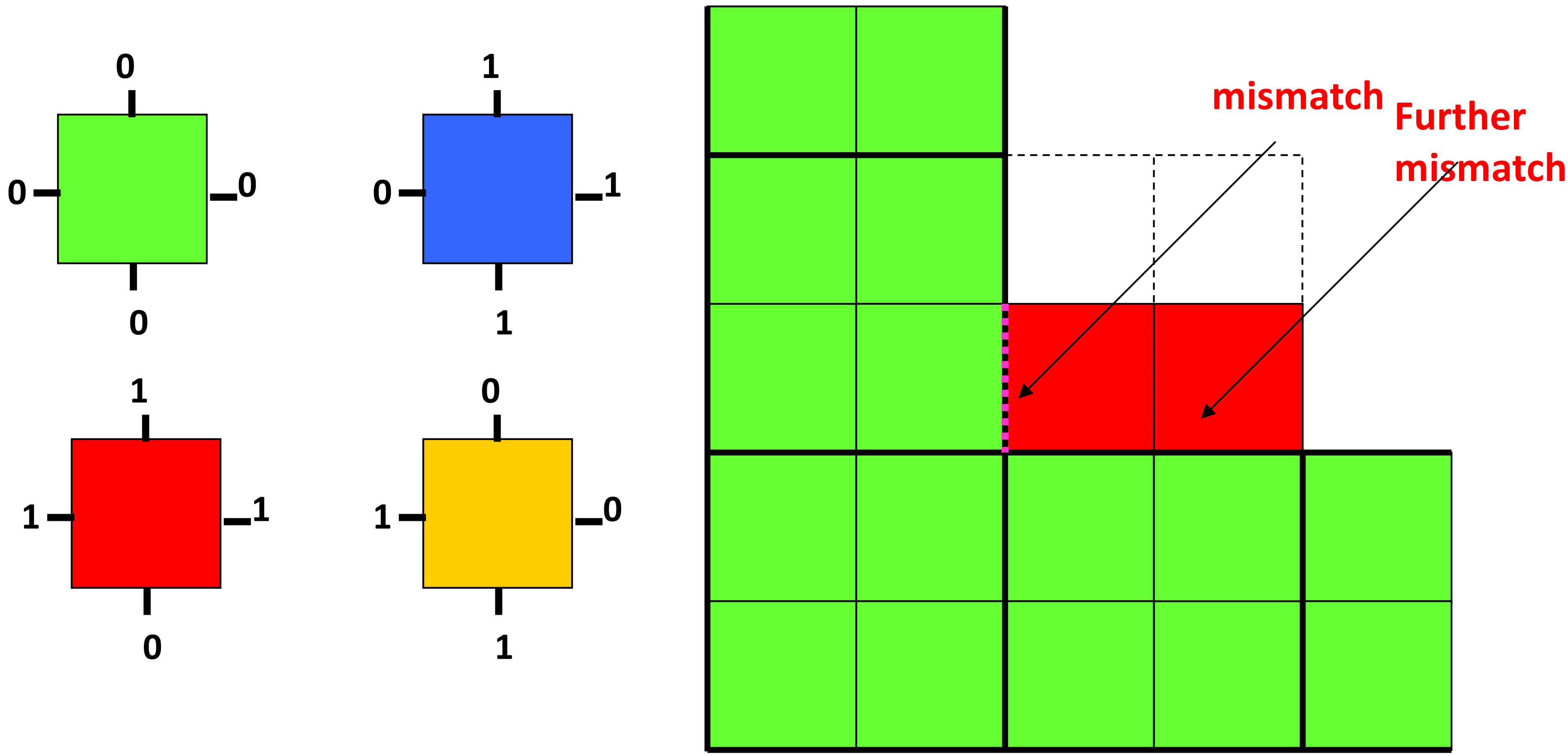
How Proofreading Tiles Reduce Errors of self-assembly





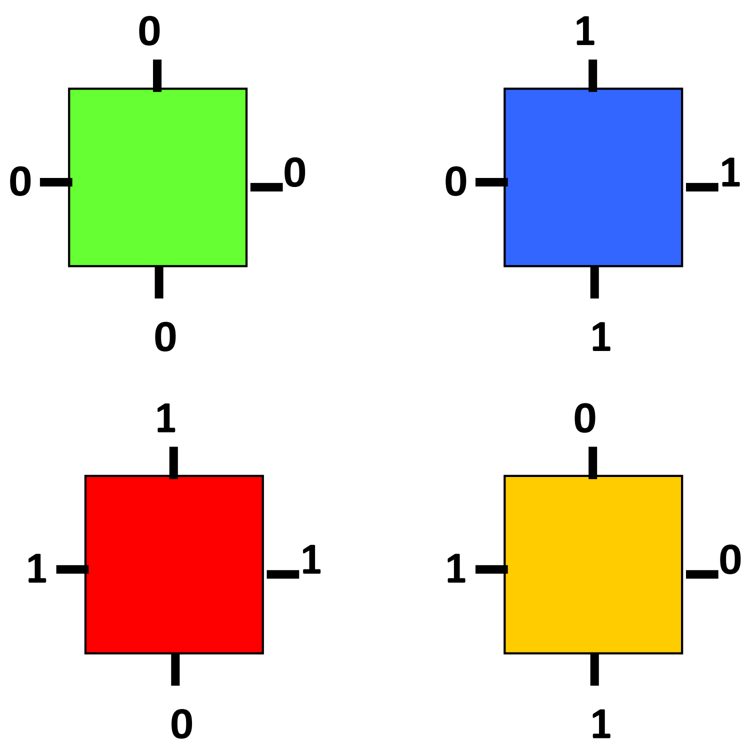
[Cheng, Goel, Cheng]

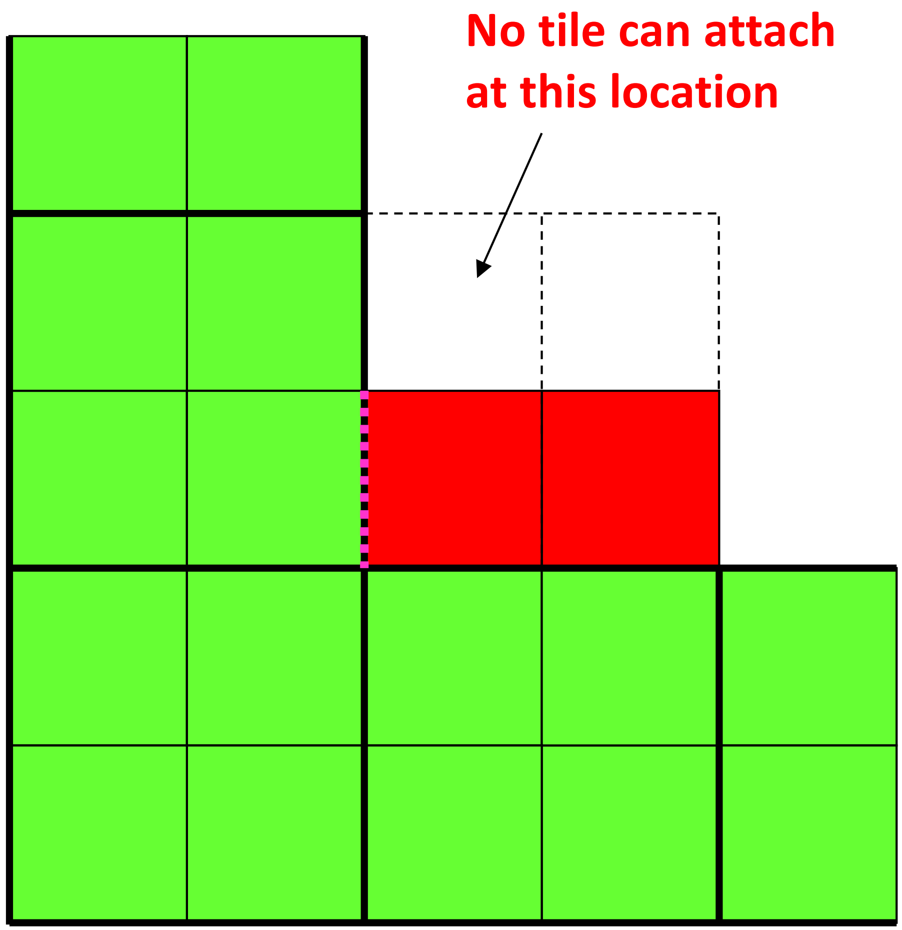
How Proofreading Tiles Reduce Errors of self-assembly



[Cheng, Goel, Cheng]

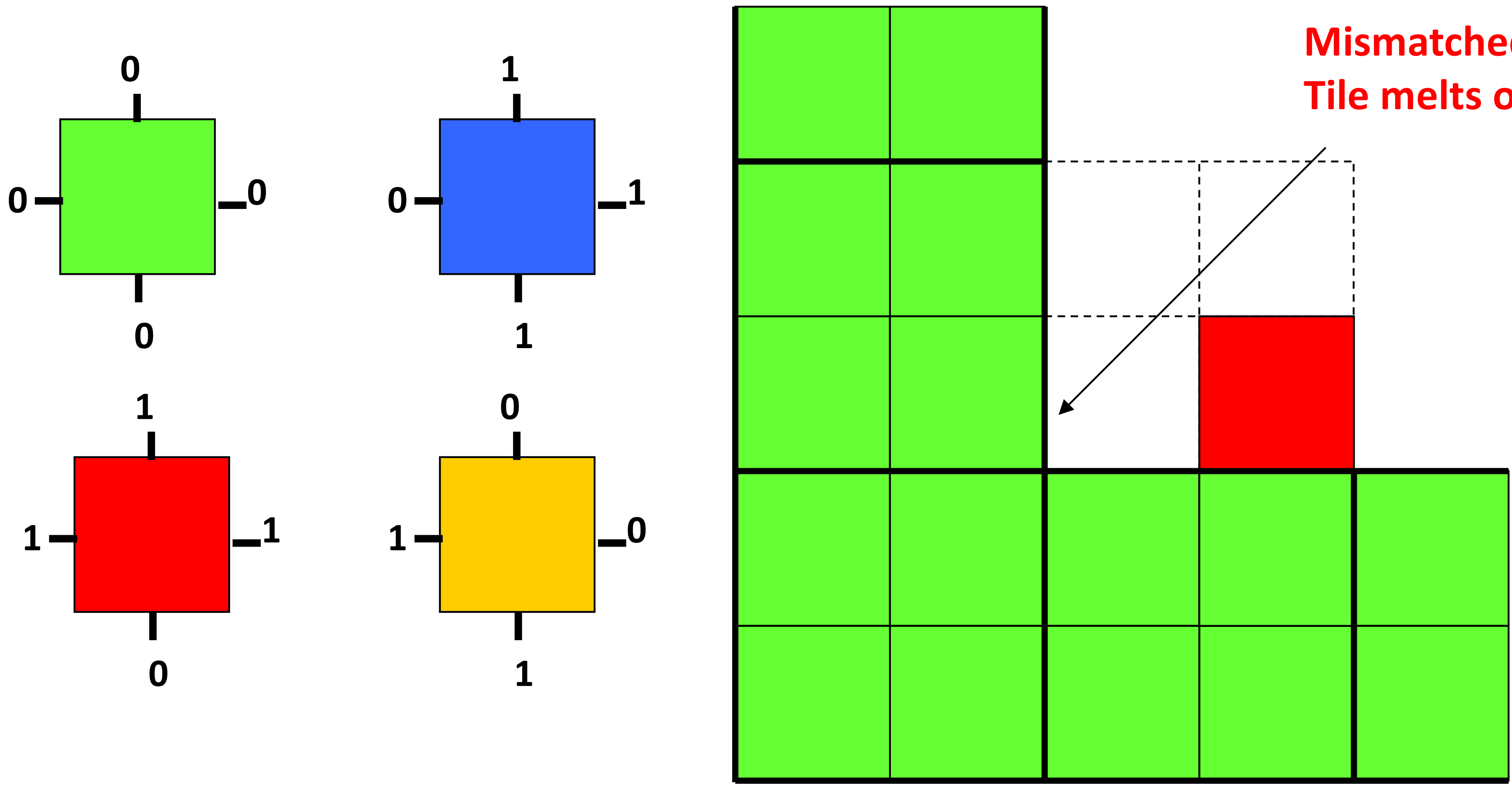
How Proofreading Tiles Reduce Errors of self-assembly





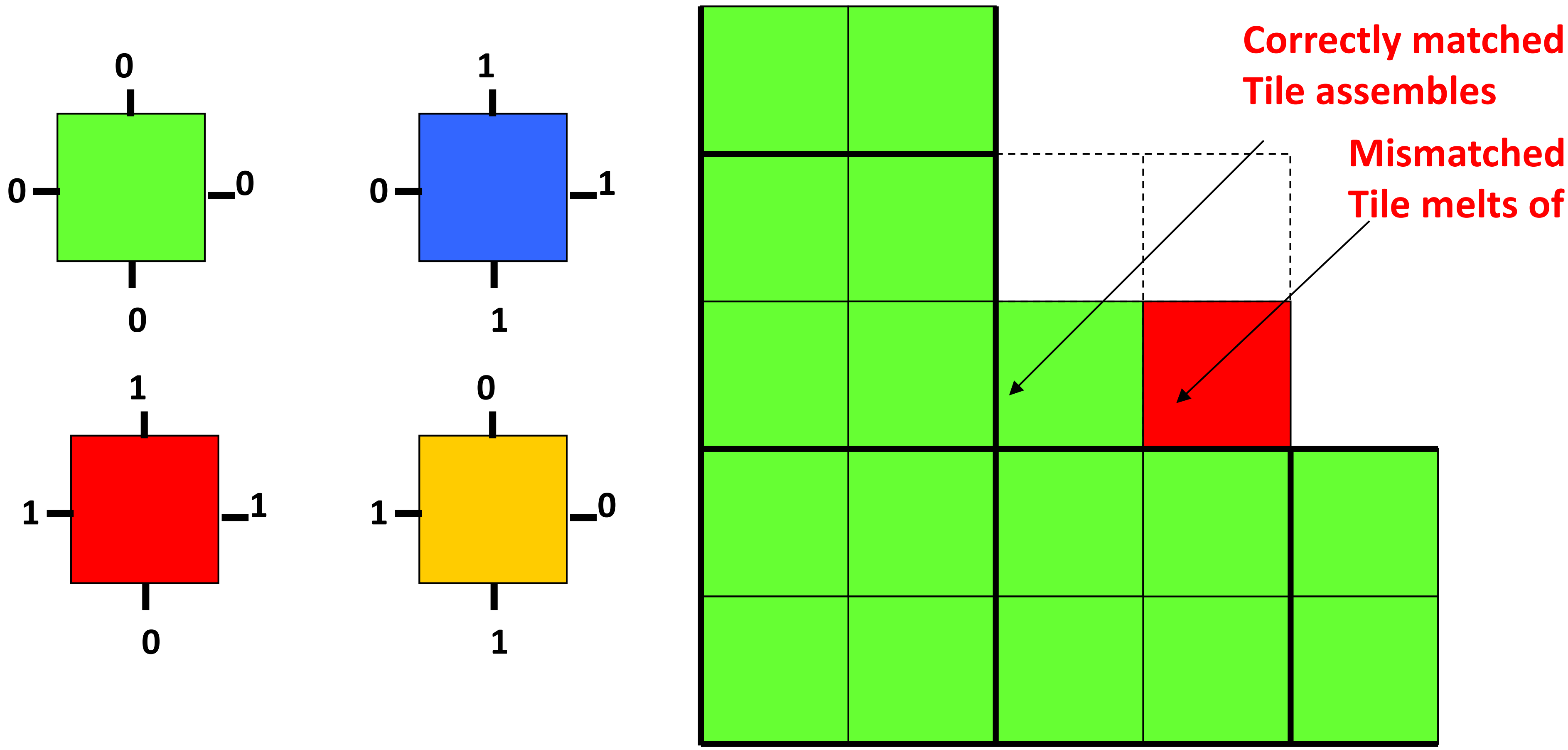
[Cheng, Goel, Cheng]

How Proofreading Tiles Reduce Errors of self-assembly



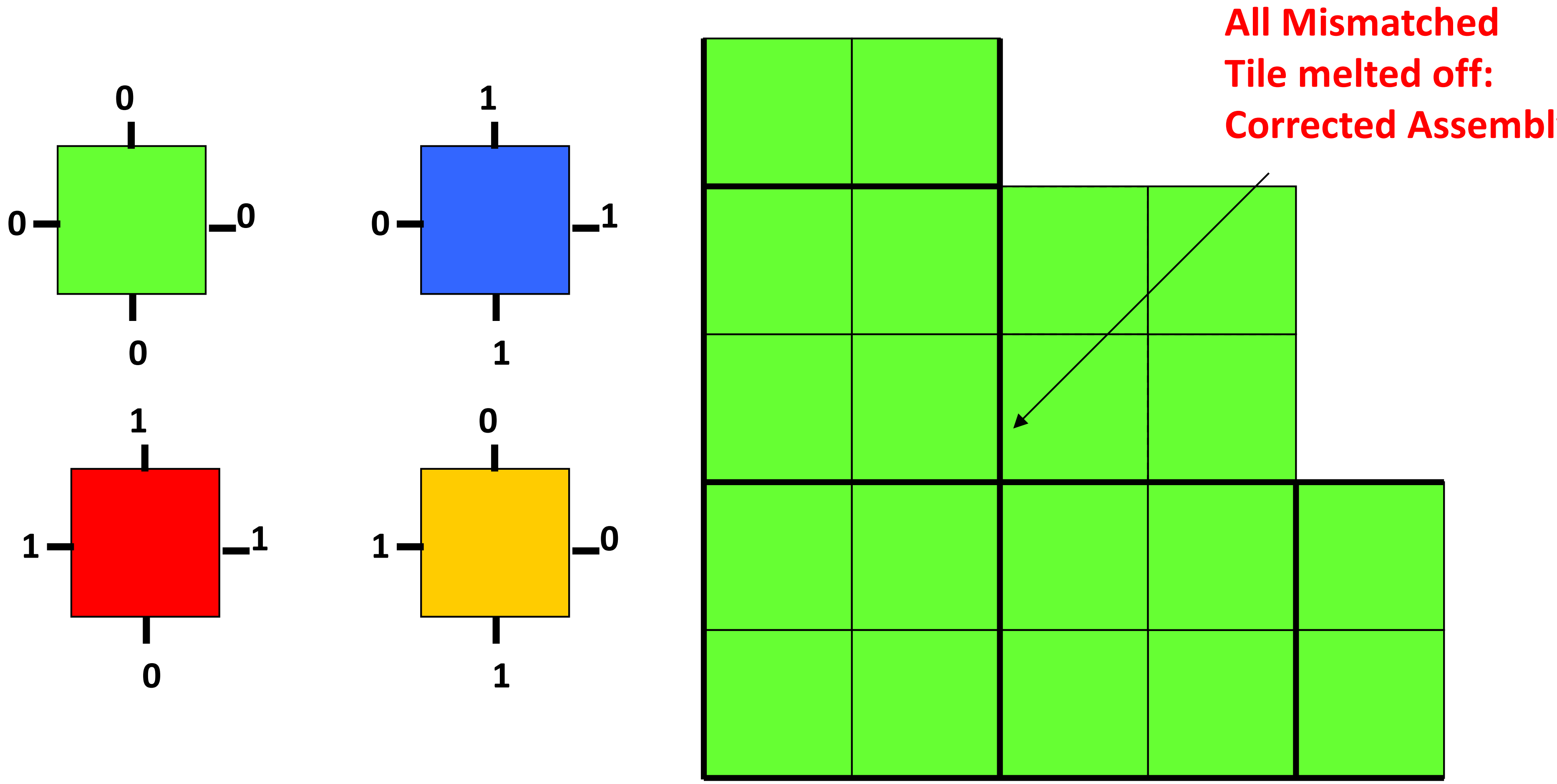
[Cheng, Goel, Cheng]

How Proofreading Tiles Reduce Errors of self-assembly



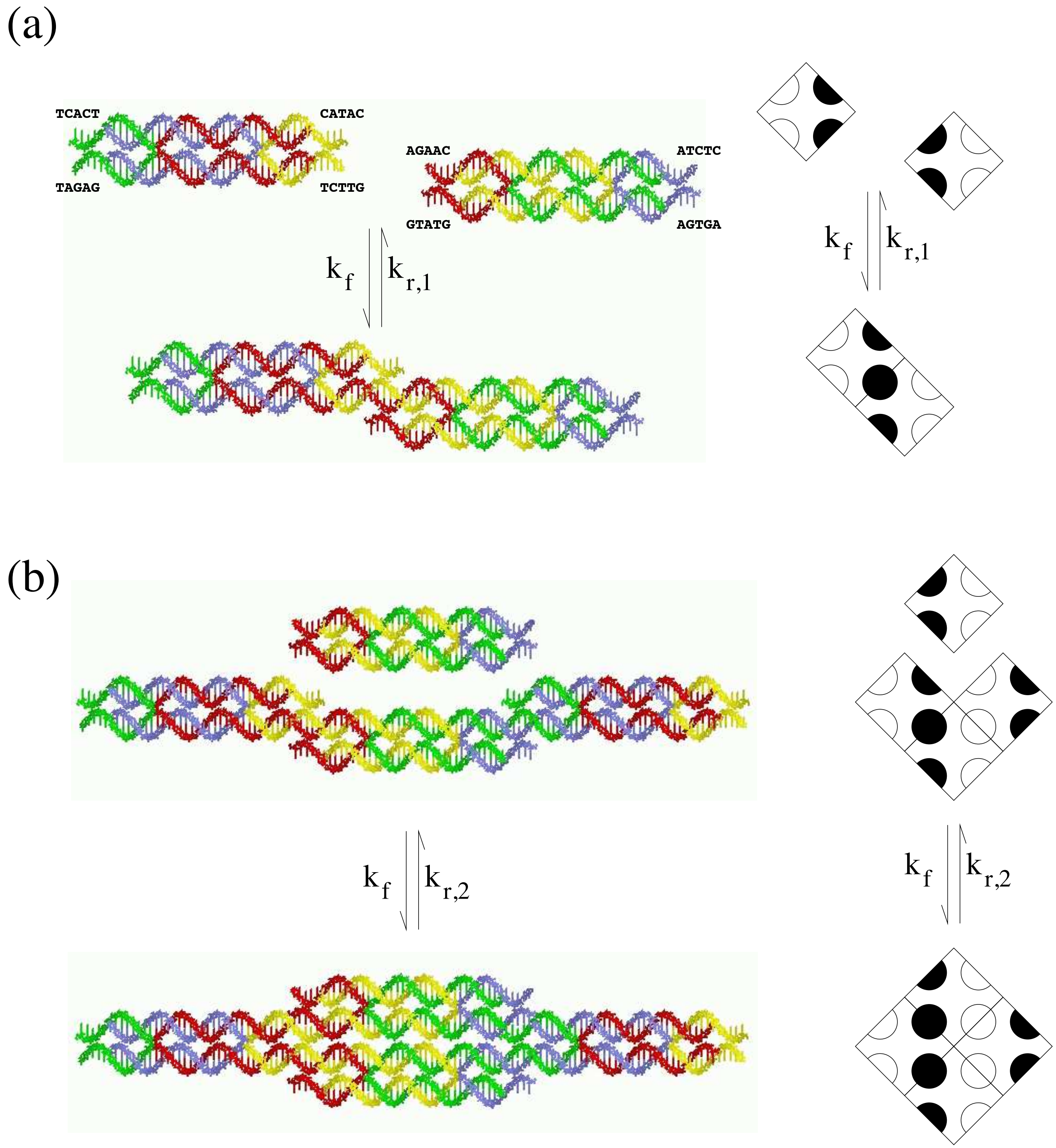
[Cheng, Goel, Cheng]

How Proofreading Tiles Reduce Errors of self-assembly



[Cheng, Goel, Cheng]

Self-assembly of Double Crossover Tiles used to form a Sierpinski Triangle



[Winfree,Bekbolatov2003]

Kinetic Analysis of Original Assembly without Proofreading Tiles:

- ***Assume a continuous-time Markov process (satisfying detailed balance) for modeling the 3 D growth of a single crystal in a solution of free monomer tiles.***
- ***Assume a monomer tile whose interactions (in the strength units of the aTAM) with the crystal sum to***
  b= number of unit-strength sticky ends binding the tile to the crystal.

Entropic Cost of Tile Assembly*:*

Gmc **= the entropic cost of putting a tile at a binding site**
(depends on the *monomer* tile concentration)

Free Energy Cost of Tile Disassembly*:*

Gse = **free energy cost of breaking a single strength-1 bond**

[Winfree,Bekbolatov2003]

Kinetic Analysis of Original Assembly ***without Proofreading Tiles:***

Entropic Cost of Tile Assembly*:*

G_{mc} = the entropic cost of putting a tile at a binding site (depends on the *monomer* tile concentration)

Free Energy Cost of Tile Disassembly*:*

G_{se} = free energy cost of breaking a single strength-1 bond

Absolute *Rates:*

Rate Association of a new monomer tile at any given site:

$$r_f = k_f\,[\text{monomer tile}] = k_f\, e^{-G_{mc}}$$

[Winfree,Bekbolatov2003]

Rate Disassociation:

$$r_{r,b} = k_{r,b} = k_f\, e^{-bG_{se}} = k_f\, e^{-2G_{se}} \quad \textit{for case } b = 2$$

b= number of unit-strength sticky ends binding the tile to the crystal.

Kinetic Analysis of Modified Assembly ***without Proofreading Tiles:***

Gmc **= the entropic cost of putting a tile at a binding site**
(depends on the *monomer* tile concentration)
Gse = **free energy cost of breaking a single strength-1 bond**

Optimal Growth Rates are near Melting Temp of crystals:
When $Gmc \approx b\ Gse \approx 2\ Gse$ *when* $b=2$

ΔG = difference in *free energy between an assembly with mismatched tile and assembly with a correct tile.*

Assume thermodynamic limit: ***Error Rate*** $\varepsilon \approx e^{-\Delta G/RT} \approx e^{-\Delta Gse}$

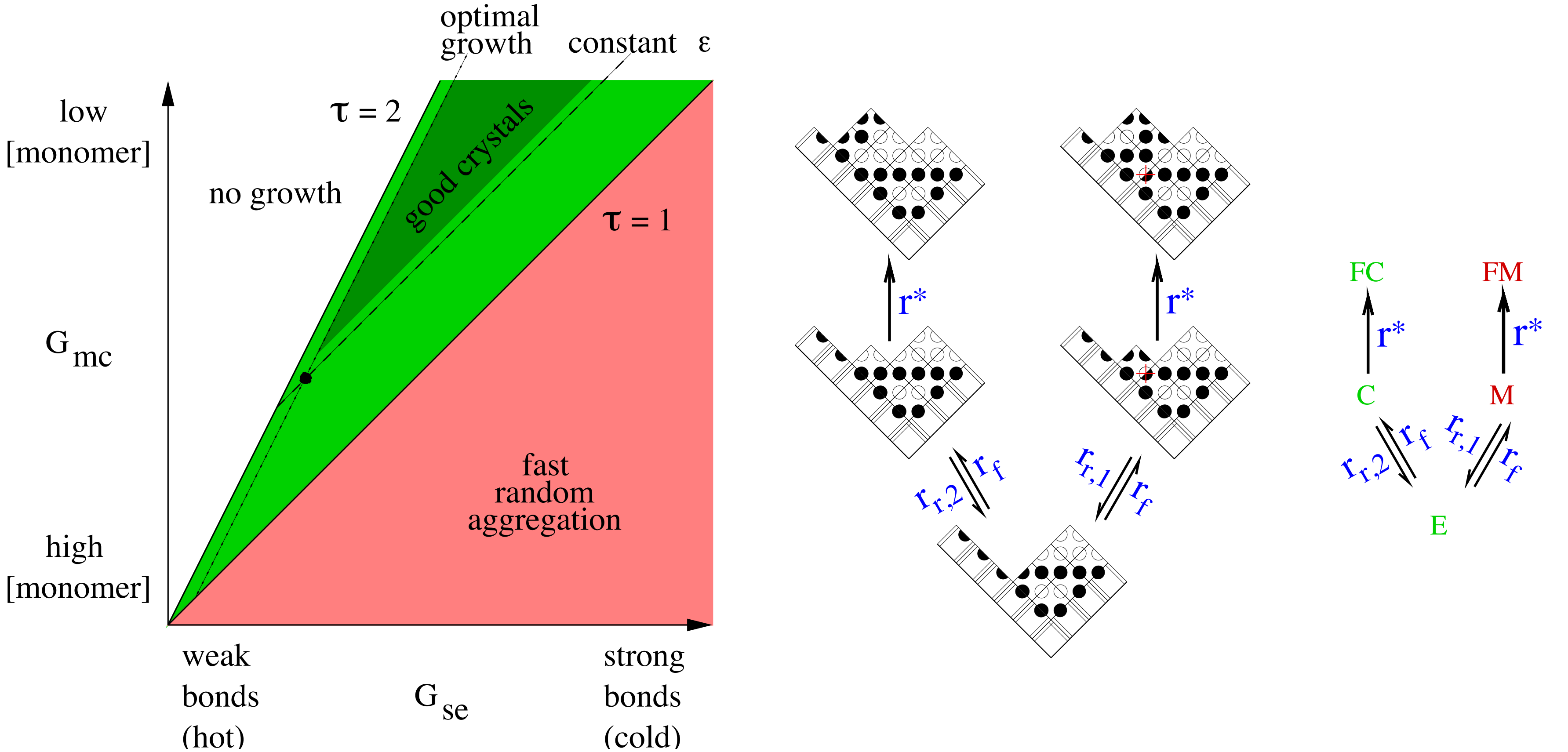
Growth Rate of Assembly:

$r \approx$ ***monomer tile concentration***

$= \beta\ [monomer\ tile]\ = \beta\ e^{-Gmc} \approx \beta\ e^{-2Gse} = \beta\varepsilon^2$

[Winfree,Bekbolatov2003]

Kinetic Analysis of Original Assembly without Proofreading Tiles



(a) (b)

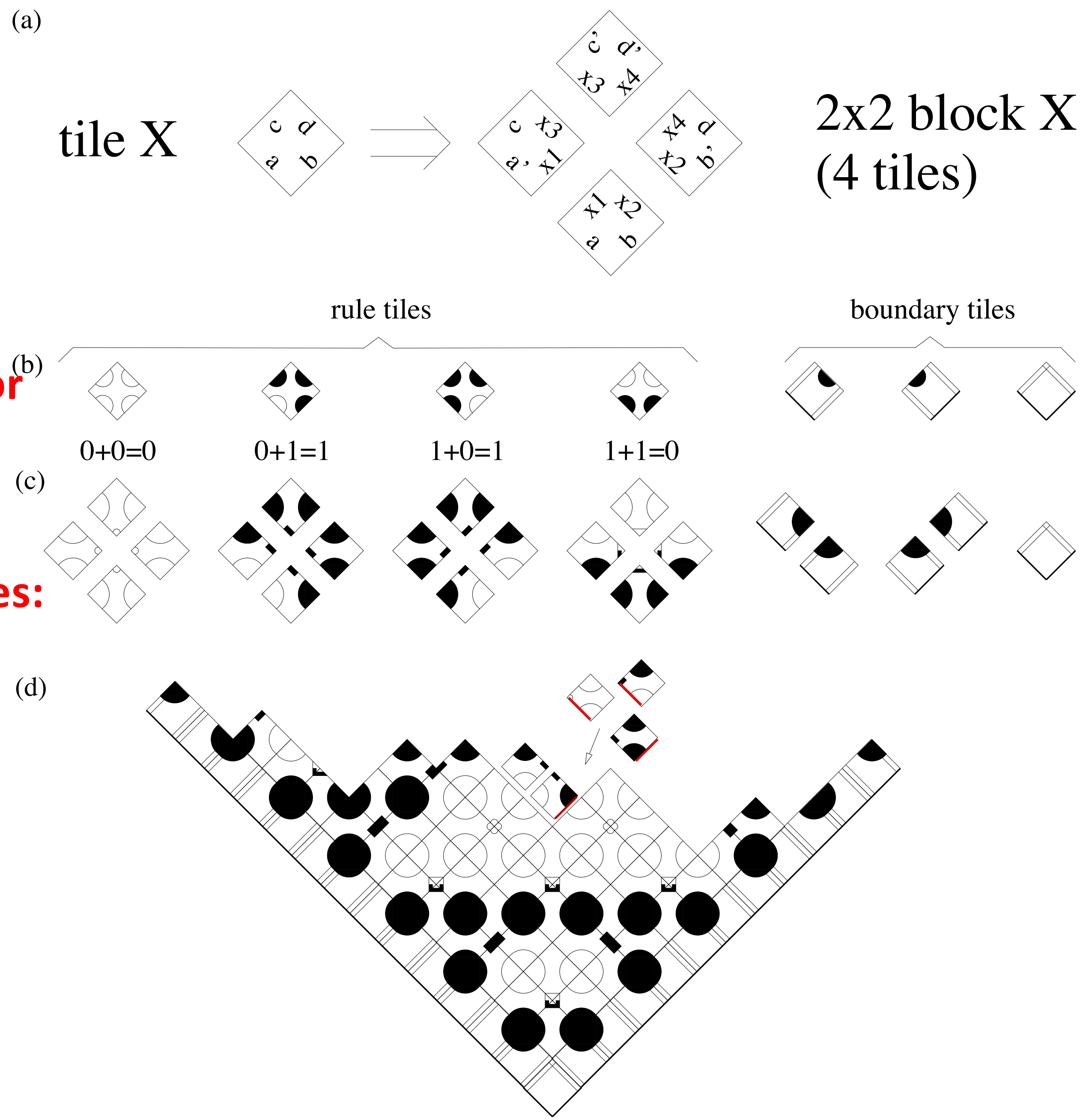
(a) Phase diagram [28] for crystal growth of tiles implementing a BCA, under the kTAM. “Good crystals” (growth rate comparable to $k_f[DX]$ and error rate smaller than ε) are obtained for large G_{se} and G_{mc}, below the $\tau = 2$ boundary marking the melting transition where $G_{mc} = 2G_{se}$. **(b)** Model for kinetic trapping. The growth site may (E) be empty; (C) contain a correct tile; (M) contain a mismatched tile; (FC) be “frozen” wit rrect tile in place; or (FM) be “frozen” with the mismatched tile. r^* represents the rate at which tiles on the growth front are covered. The error rate is taken to be the probability that, starting in E, the system reaches FM.

[Winfree,Bekbolatov2003]

Design of original tiles for Sierpinski Triangle:

Redesign of original tiles to 2 x 2 Proofreading Tiles:

Using 2 x 2 Proofreading Tiles to assemble a Sierpinski Triangle With less errors

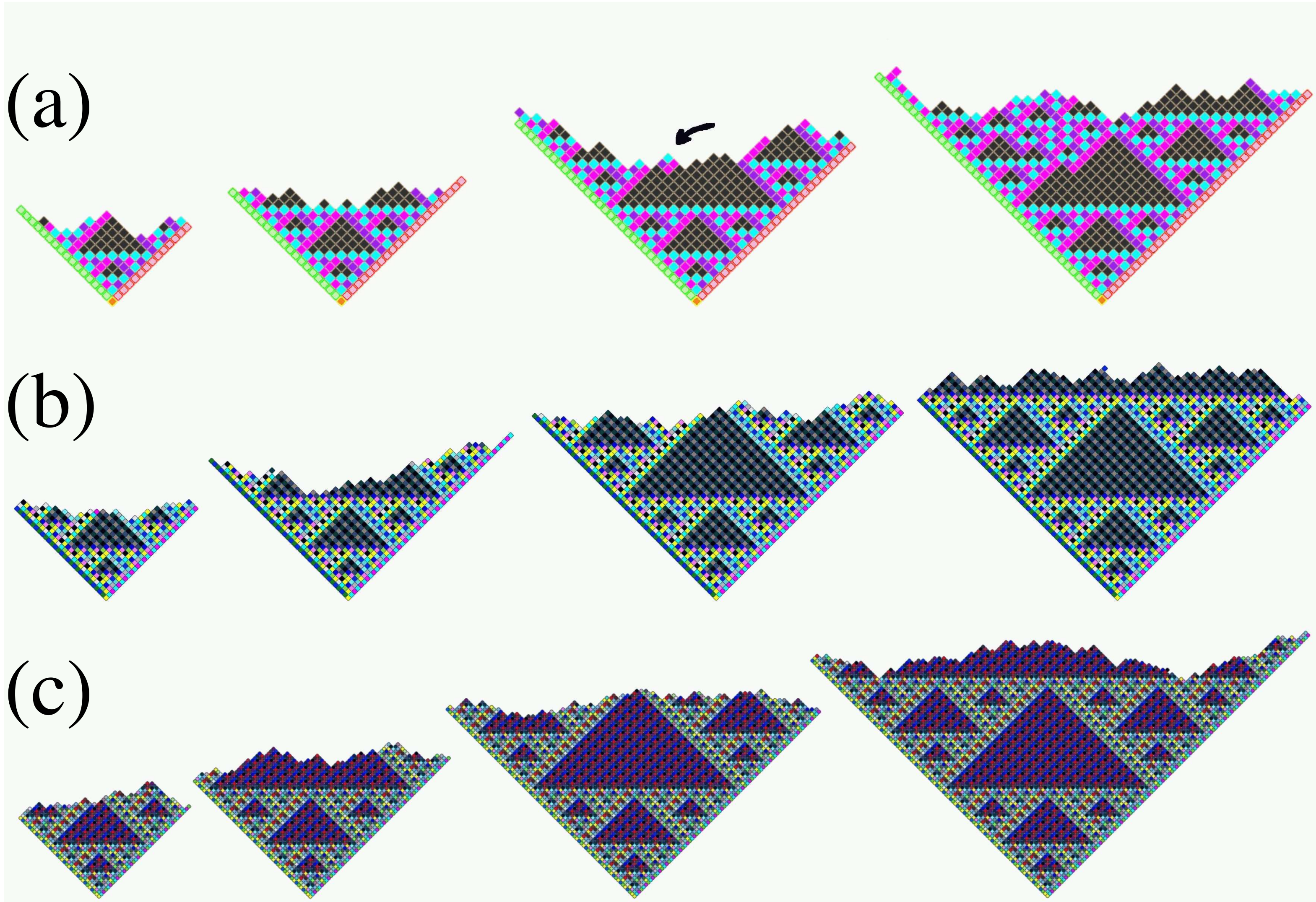


(a) The general 2×2 proofreading construction for rule tiles. **(b)** The original Sierpinski tiles. **(c)** The 2×2 proofreading Sierpinski tiles. **(d)** Growth of the proofreading Sierpinski tiles. Small tiles illustrate that when a mismatched tile is incorporated, further growth on one side must involve a second mismatch. *[Winfree,Bekbolatov2003]*

Using Original Tiles to assemble a Sierpinski Triangle with many errors

Using 2 x 2 Proofreading Tiles to assemble a Sierpinski Triangle with less errors

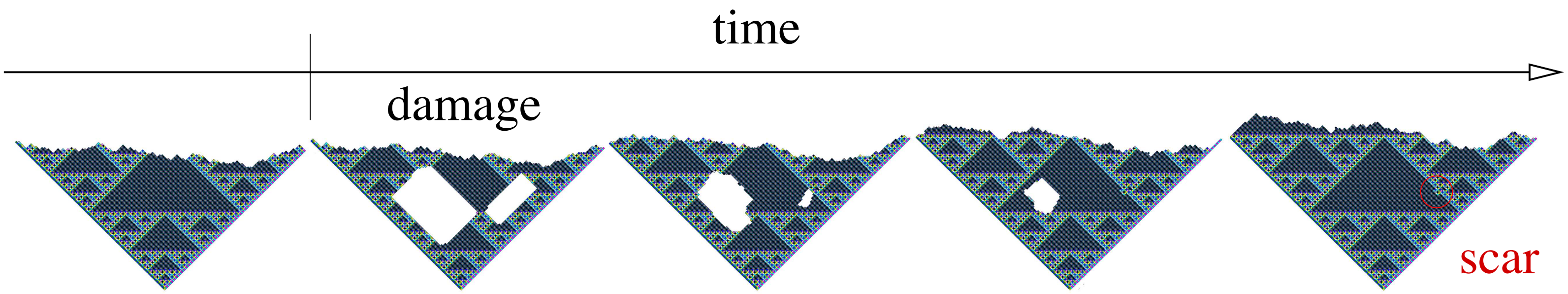
Using 3 x 3 roofreading Tiles to ssemble a Sierpinski Triangle vith much less errors



(a) Growth of the original (1×1) Sierpinski tile set at $G_{mc} = 13.9$ and $G_{se} = 7.0$, to a size of ~ 32 layers in ~ 530 simulated seconds. Two errors can be seen; the first occurs in the third frame and is indicated by an arrow. Subsequent error-free growth correctly propagates the erroneous information. **(b)** Growth of the 2×2 proofreading tiles at $G_{mc} = 12.9$ and $G_{se} = 6.5$, to a size of ~ 64 layers in ~ 460 simulated seconds. **(c)** Growth of the 3×3 proofreading tiles at $G_{mc} = 11.9$ and $G_{se} = 6.0$, to a size of ~ 96 layers in ~ 310 simulated seconds.

[Winfree,Bekbolatov2003]

Using Proofreading Tiles to heal punctures in Sierpinski Triangle



Proofreading tile sets are often able to heal a puncture in the crystal. Sometimes, as in this case, some of the tiles that fill in the puncture do not perfectly match their neighbors – a form of "scar tissue."

[Winfree,Bekbolatov2003]

Kinetic Analysis of k x kProof-Reading Assembly:

Again assume a continuous-time Markov process:
(satisfying detailed balance) for modeling the 3D growth of a single crystal in a solution of free monomer tiles.

In Proof-Reading Assembly each monomer tile is replaced with a K x K subassembly.

Again Optimal Growth Rates are near Melting Temp of crystals, where $Gmc \approx 2\ Gse$

Assume thermodynamic limit: error rate for an entire block: now determined by K mismatched tiles

Error Rate $\varepsilon \approx e^{-\Delta G/RT} \approx e^{-\Delta Gse} \approx e^{-KGse}$ so $e^{-Gse} \approx \varepsilon^{1/k}$

Growth Rate of Proof-Reading Assembly:

$$r \approx \text{monomer tile concentration}$$
$$= \beta\ [\text{monomer tile}]$$
$$= \beta\ e^{-Gmc} \approx \beta\ e^{-2Gse} = \beta\varepsilon^{2/K}$$

[Winfree,Bekbolatov2003]

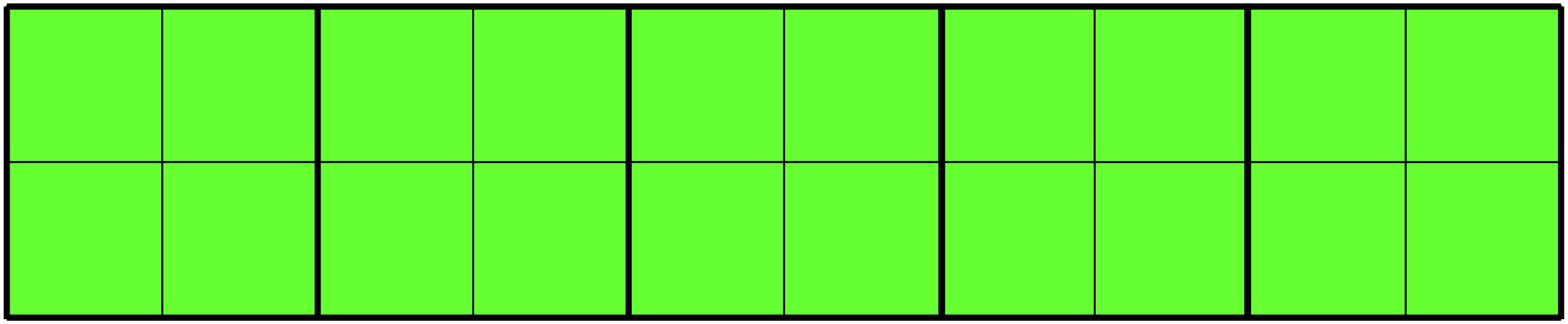
Error Free Self-assembly Using Snaked Proof-Reading Tiles

Ho-Lin Chen & Ashish Goel

2005

(Chen,Goel)

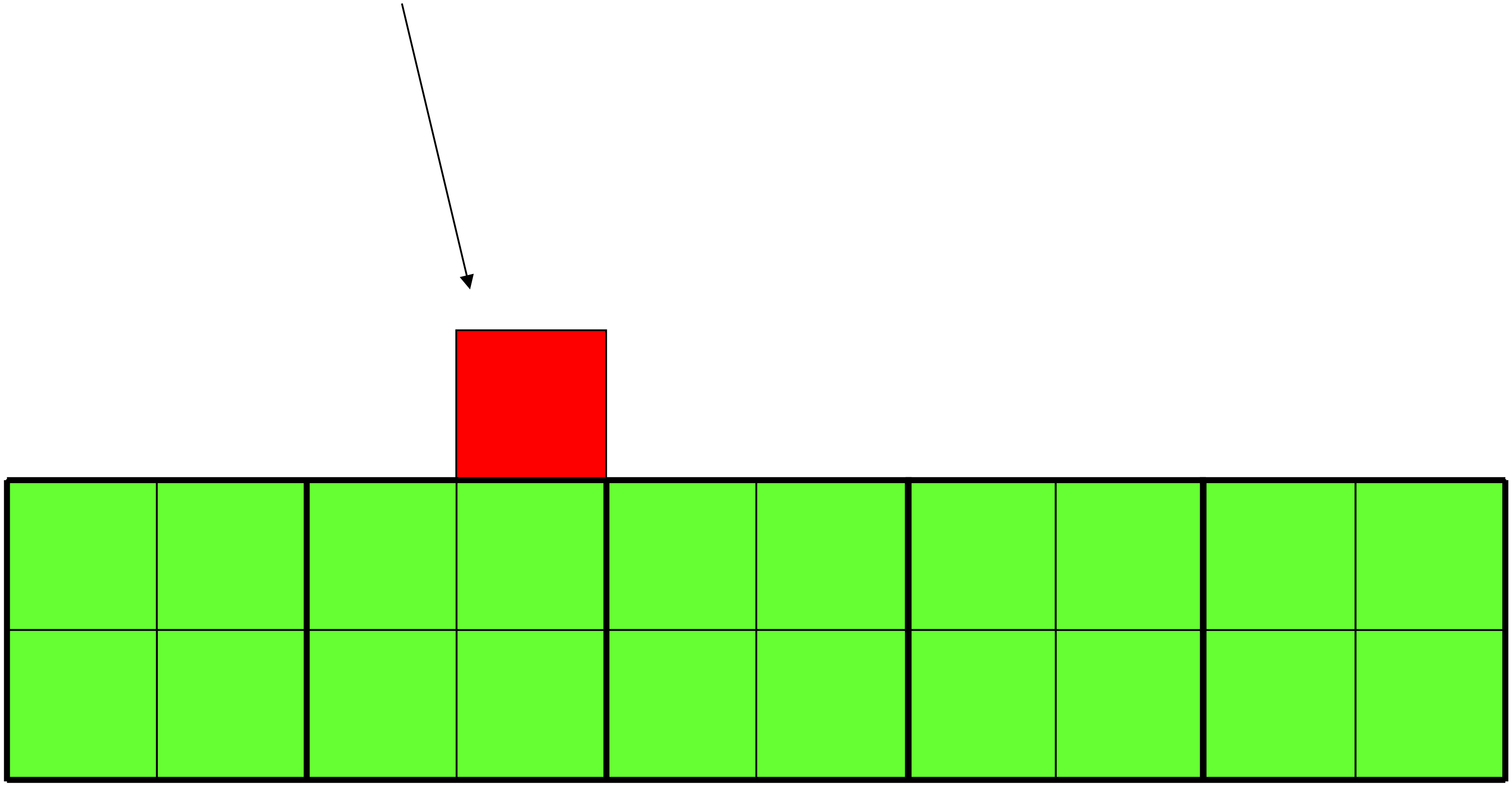
Initial Error-Free Assembly



[Cheng, Goel, Cheng]

Example of Nucleation Errors

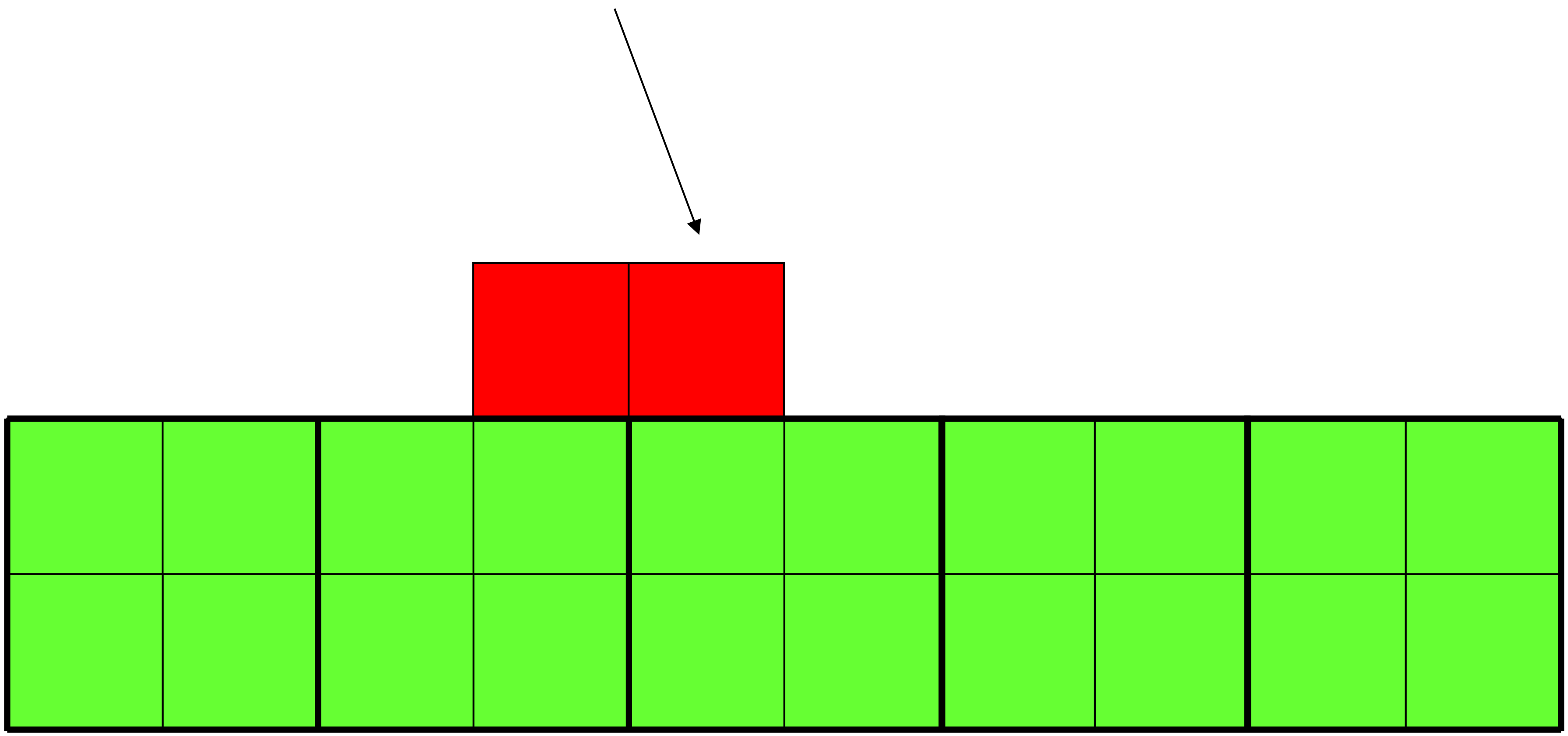
- First tile attaches with a weak binding strength



[Cheng, Goel, Cheng]

Example of Nucleation Errors

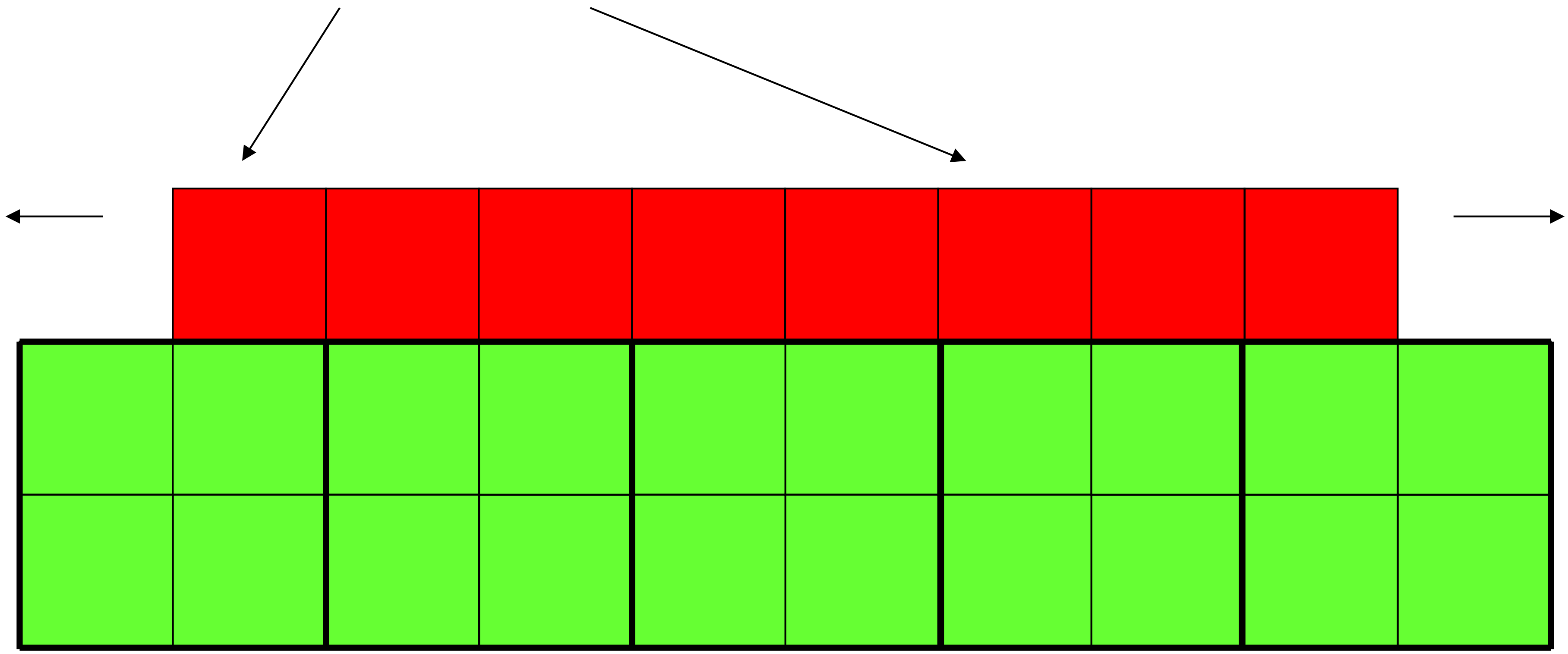
- First tile attaches with a weak binding strength
- Second tile attaches and secures the first tile



[Cheng, Goel, Cheng]

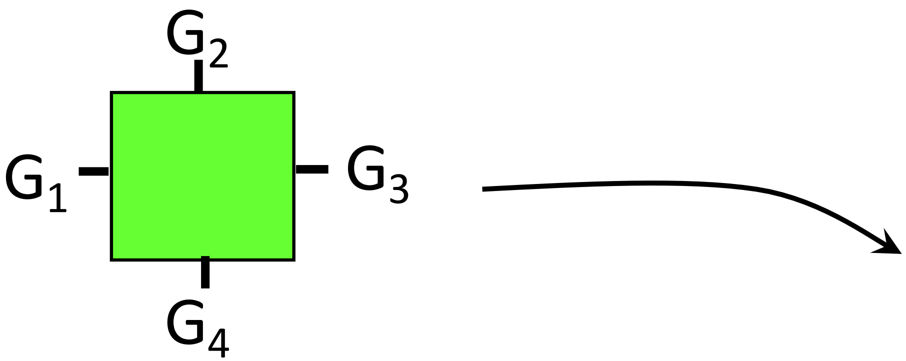
Example of Nucleation Errors

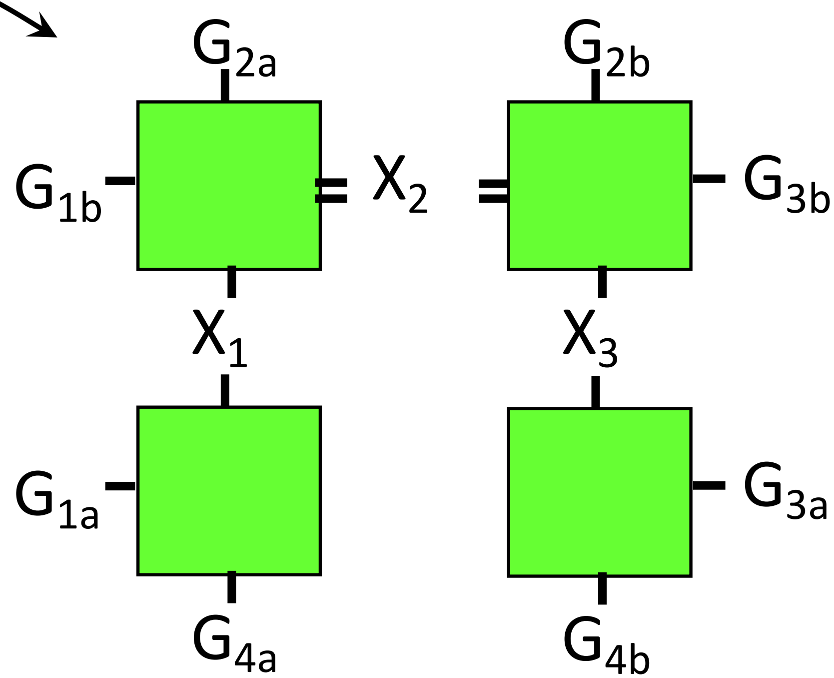
- First tile attaches with a weak binding strength
- Second tile attaches and secures the first tile
- Other tiles can attach and forms a layer of (possibly incorrect) tiles.



[Cheng, Goel, Cheng]

Snake Tiles



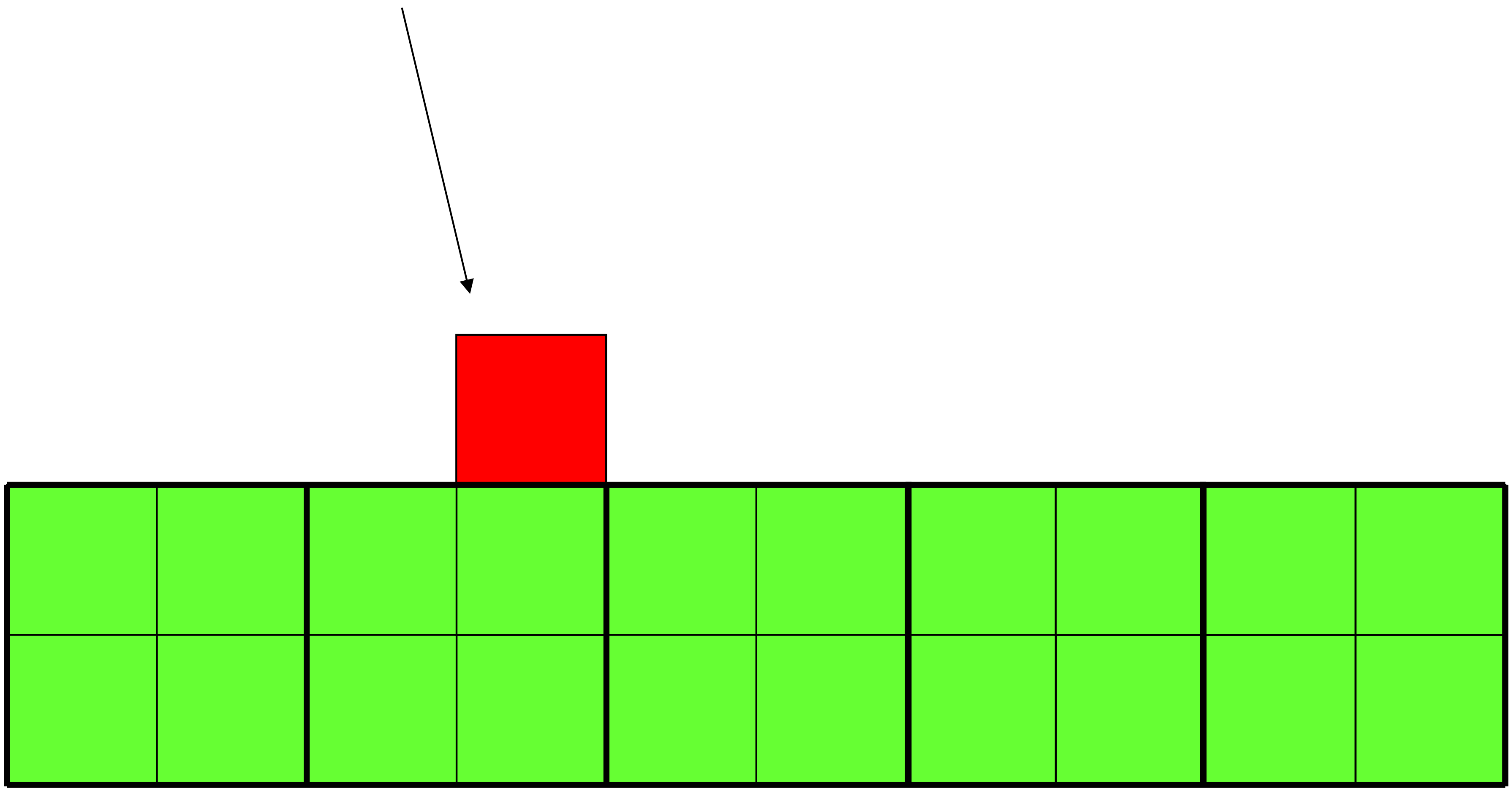


- Each tile in the original system corresponds to four tiles in the new system
- The internal glues are unique to this block

[Cheng, Goel, Cheng]

How Snake Tiles reduce Nucleation Errors

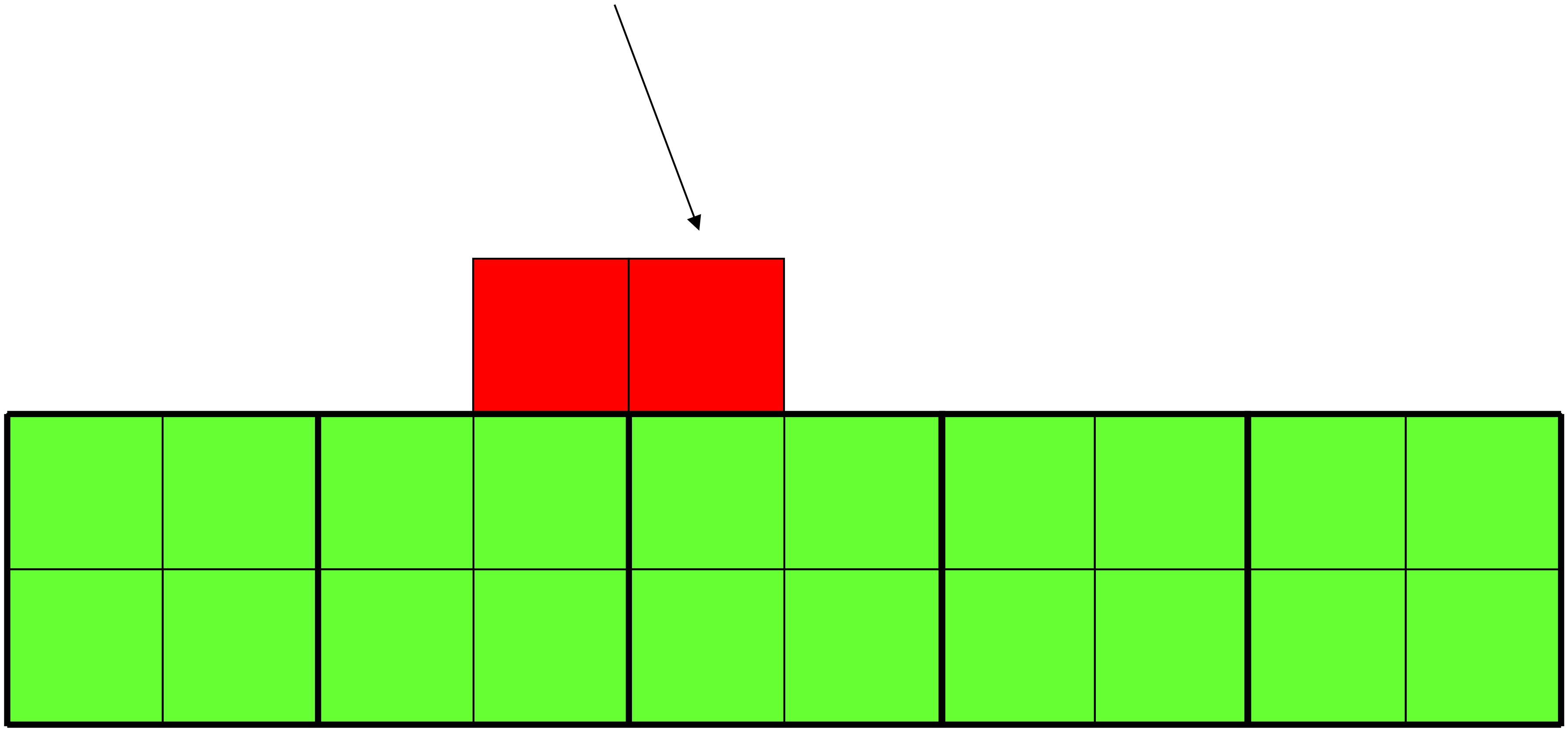
- First tile attaches with a weak binding strength



[Cheng, Goel, Cheng]

How Snake Tiles reduce Nucleation Errors

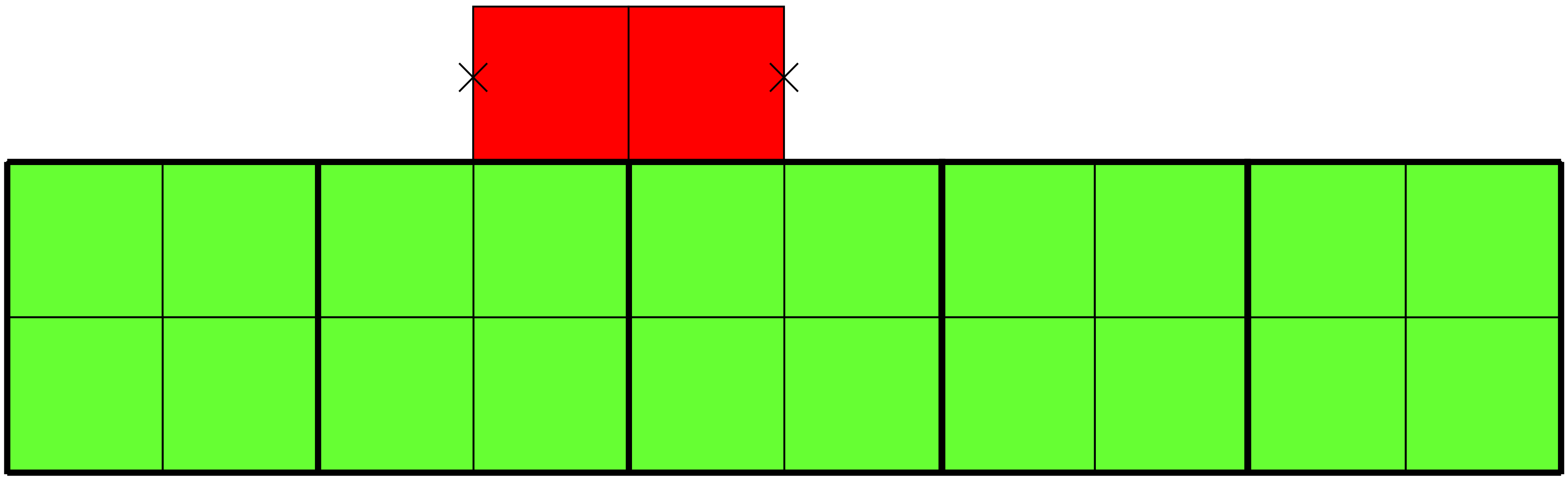
- First tile attaches with a weak binding strength
- Second tile attaches and secures the first tile



[Cheng, Goel, Cheng]

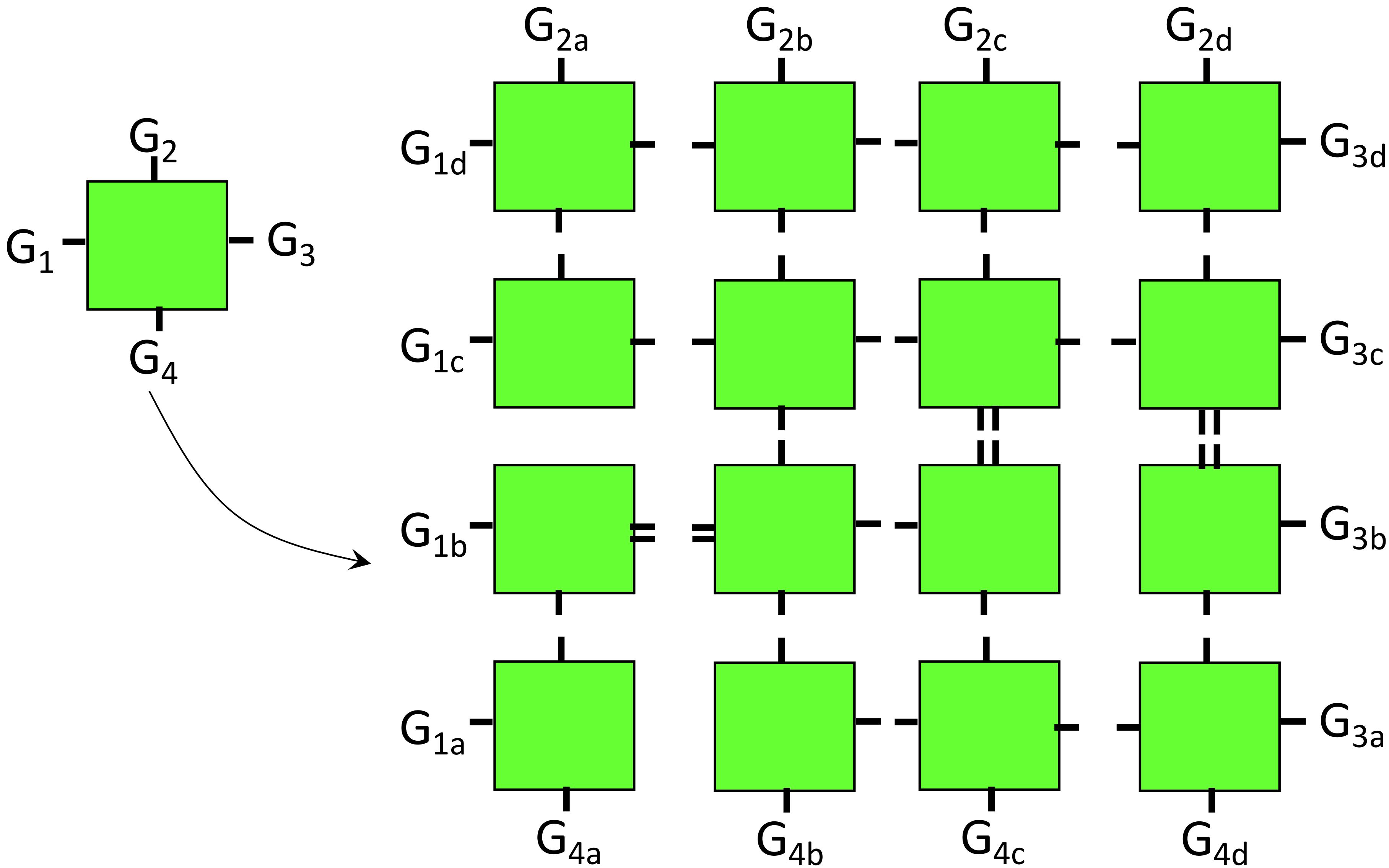
How Snake Tiles reduce Nucleation Errors

- First tile attaches with a weak binding strength
- Second tile attaches and secures the first tile
- No Other tiles can attach without another nucleation error



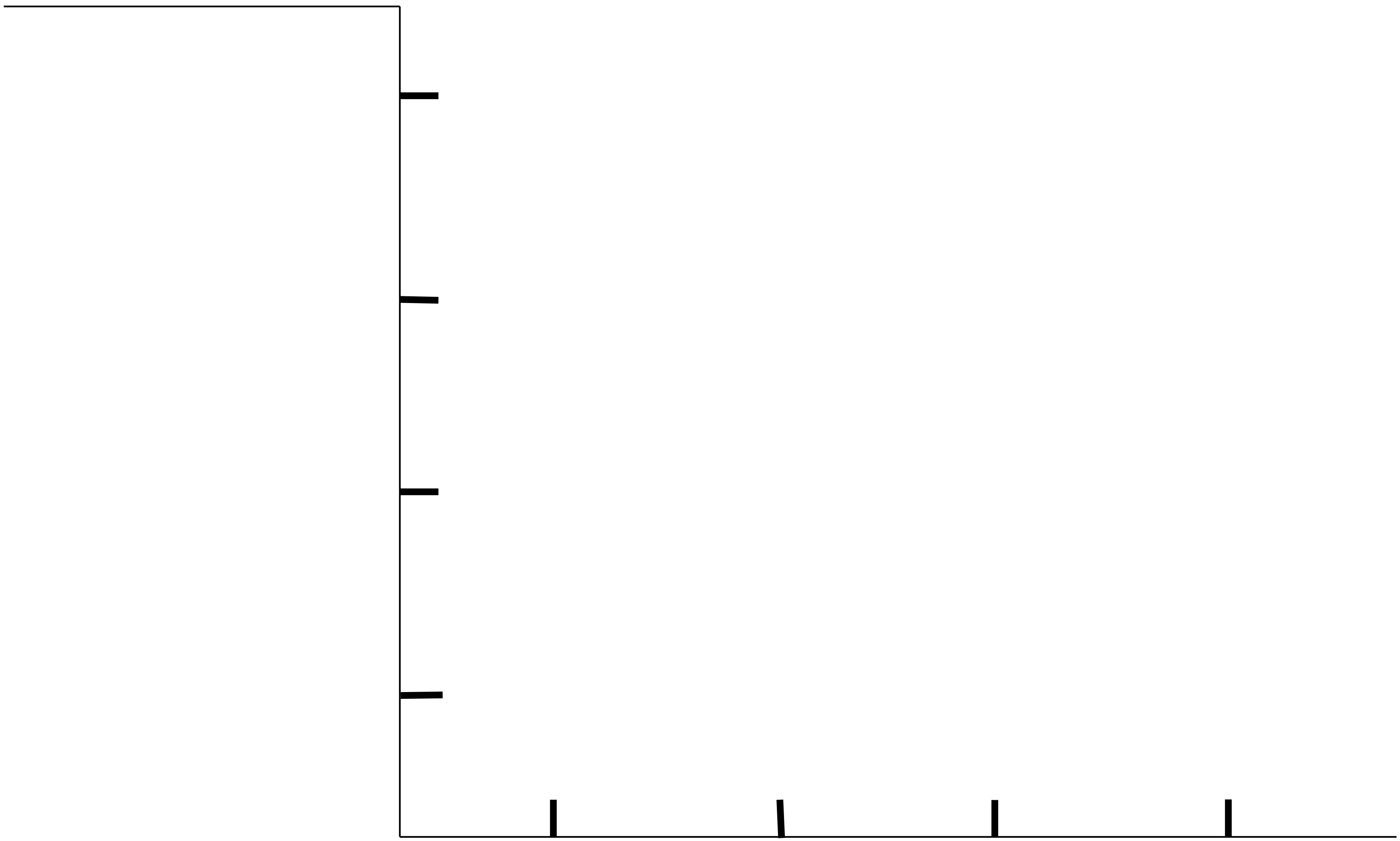
[Cheng, Goel, Cheng]

Four by Four Snake Tiles



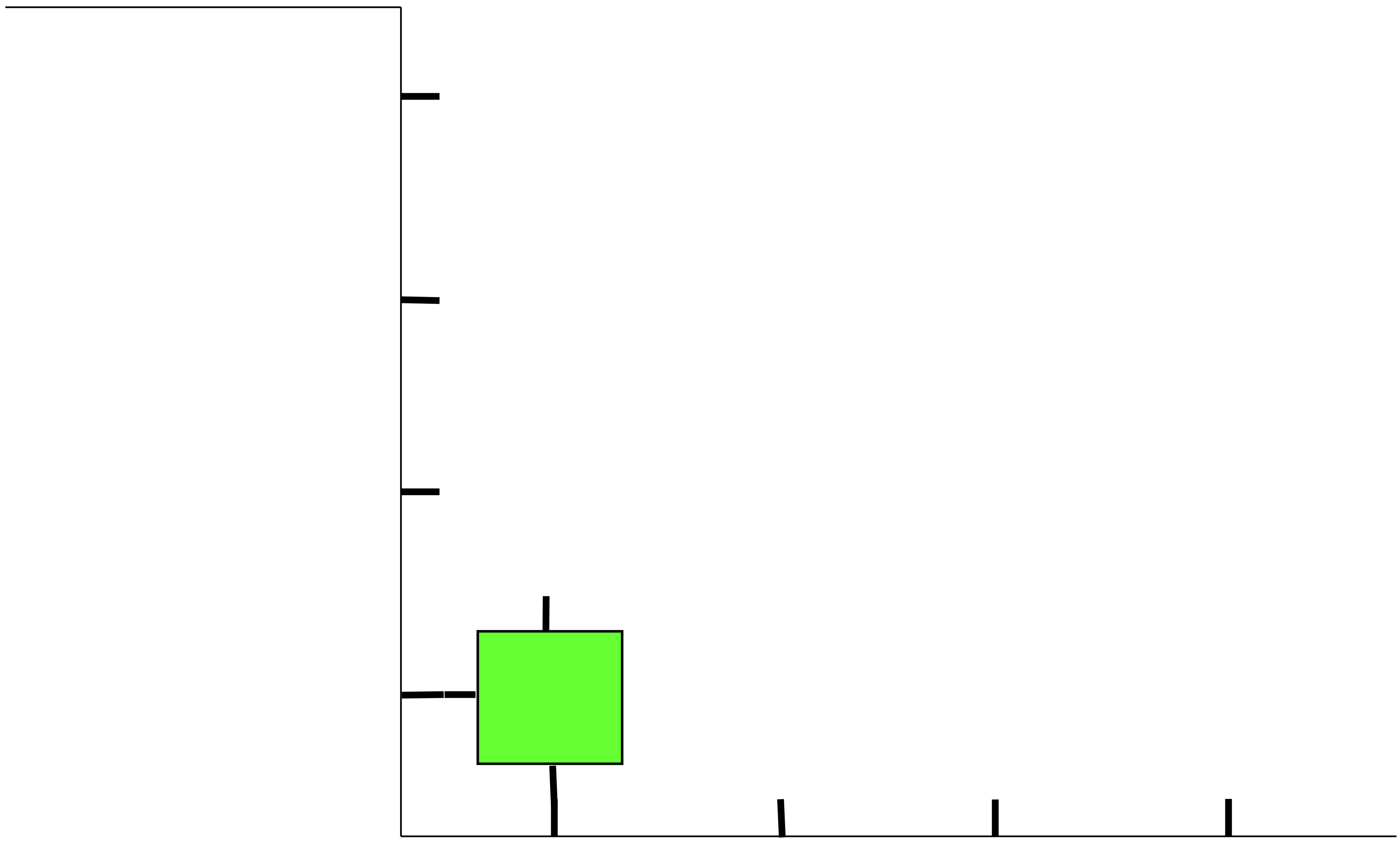
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



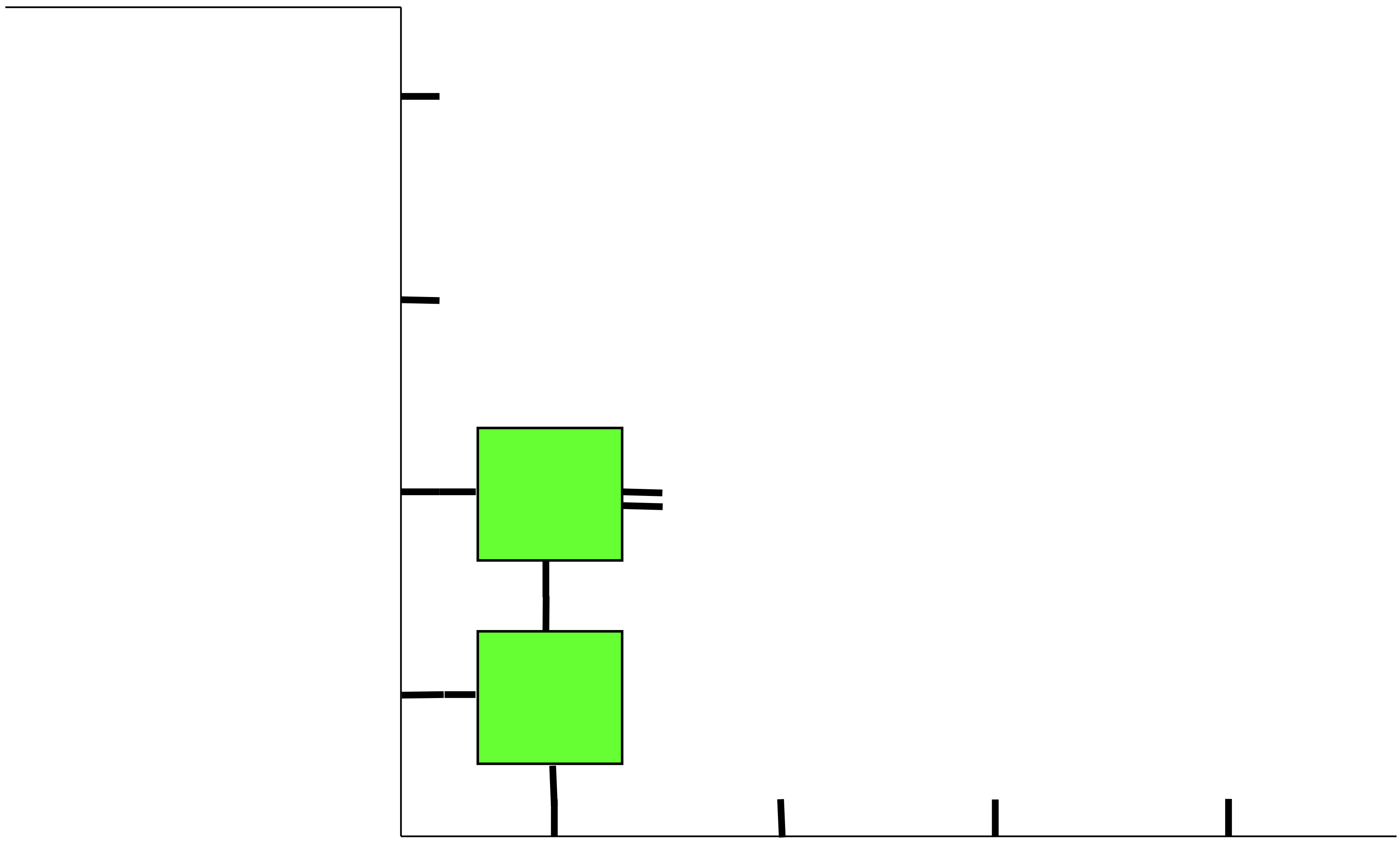
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



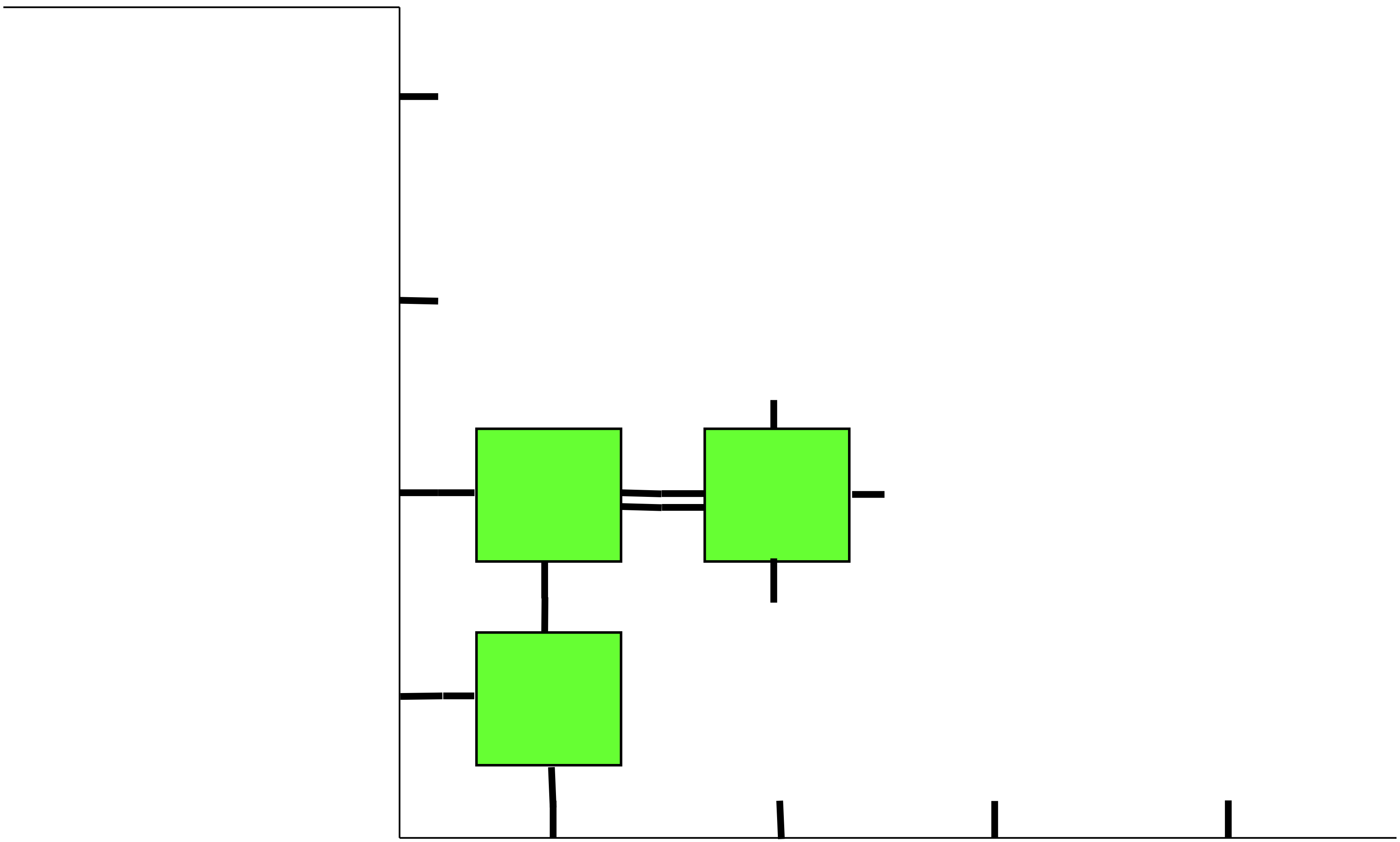
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



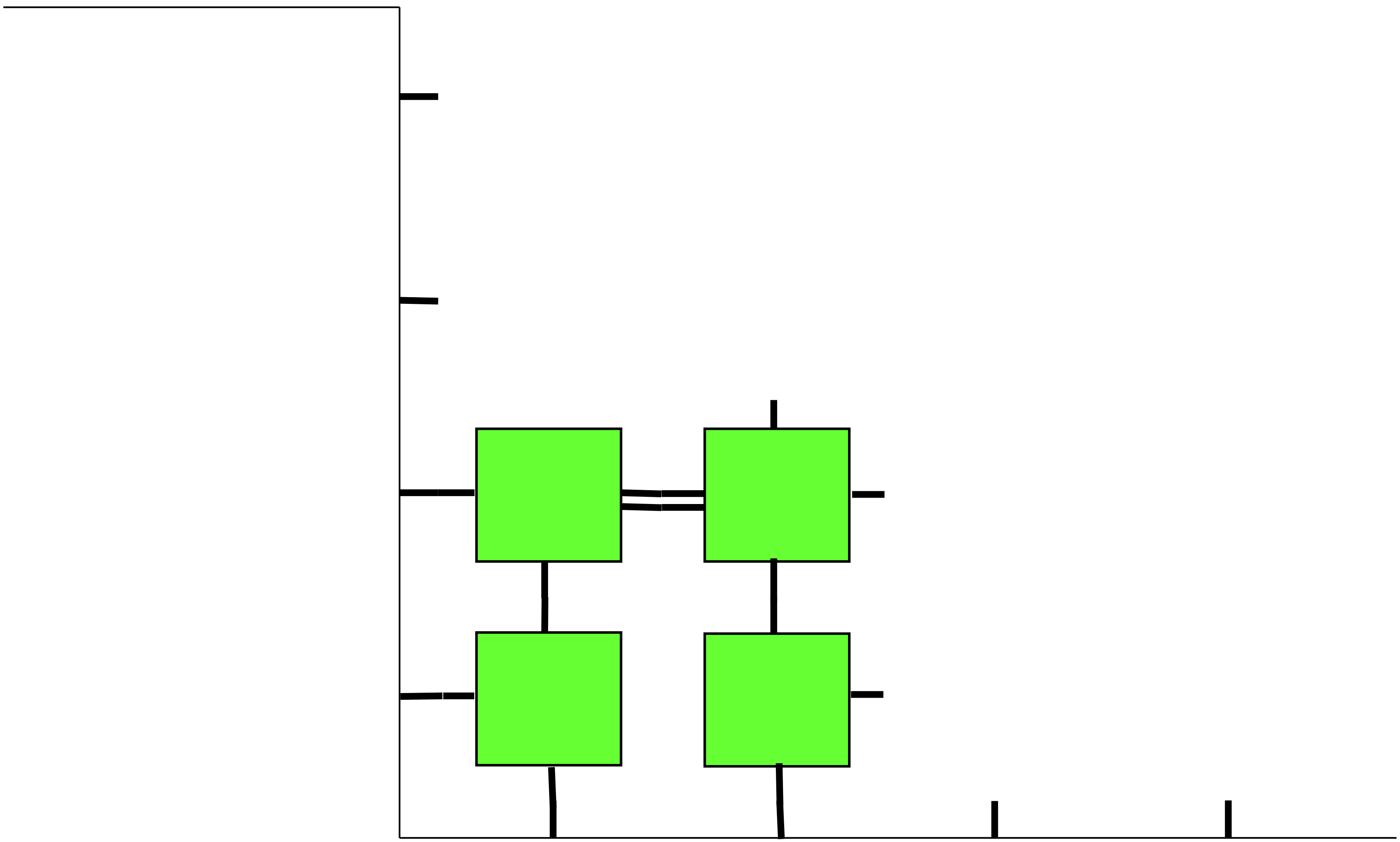
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



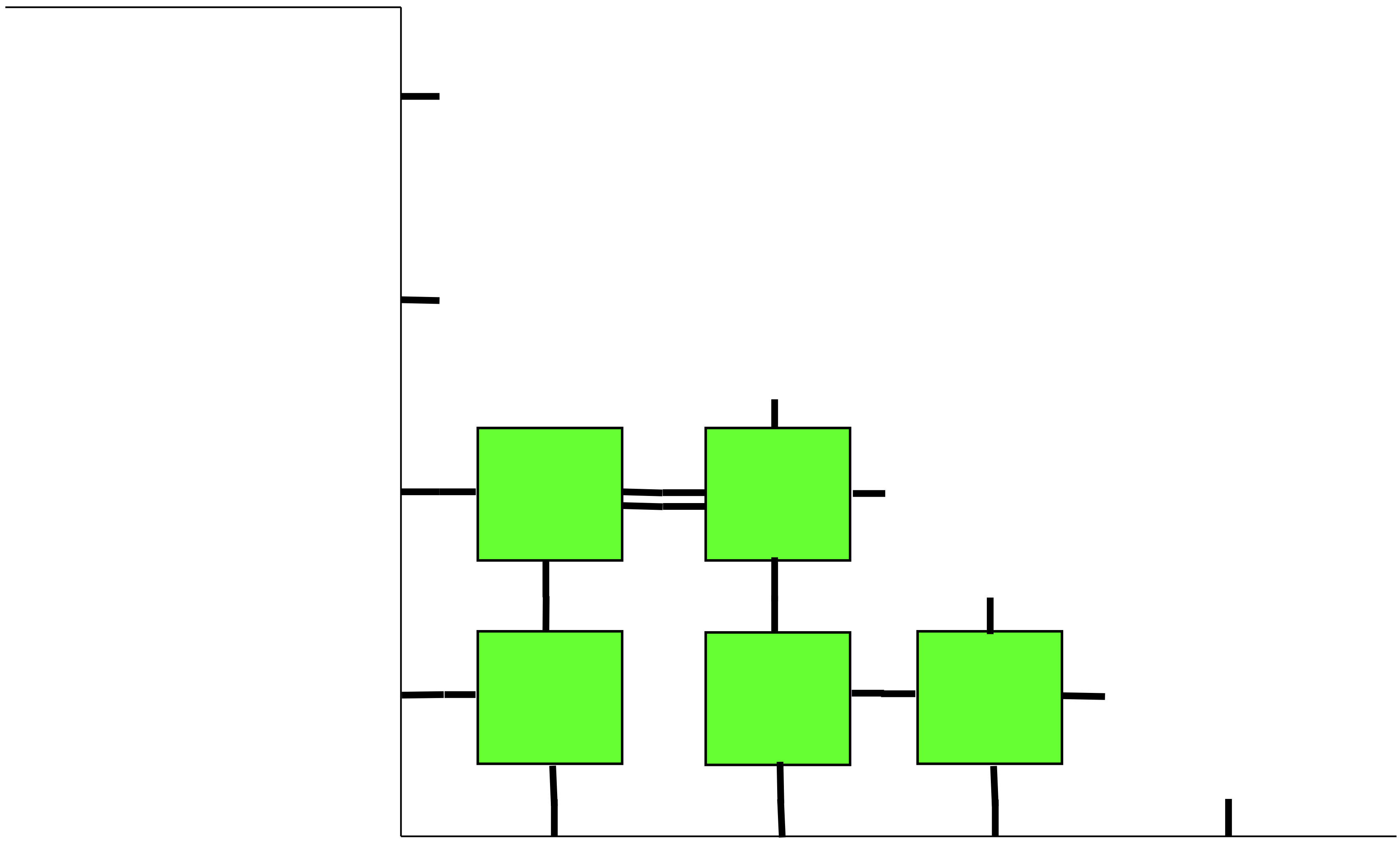
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



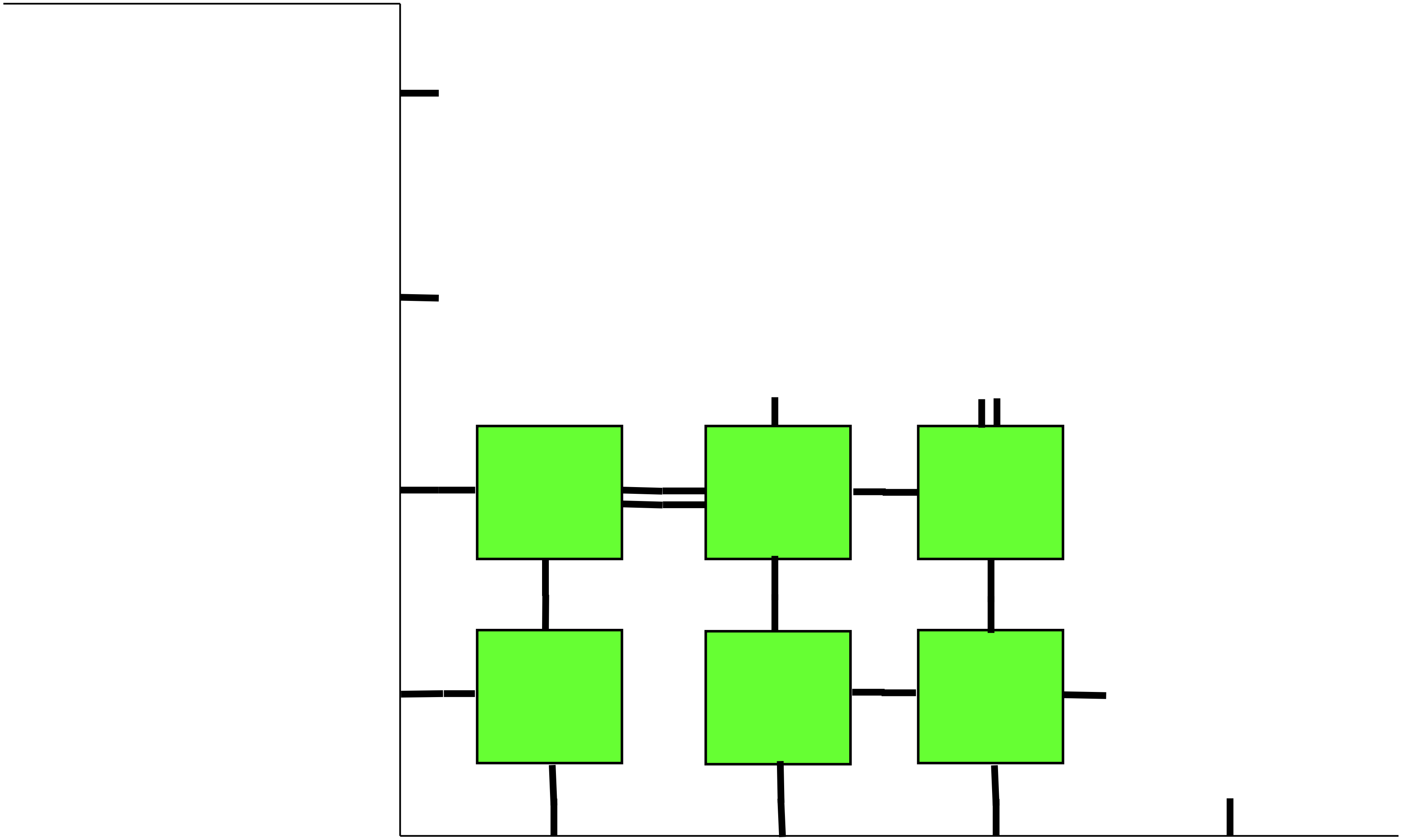
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



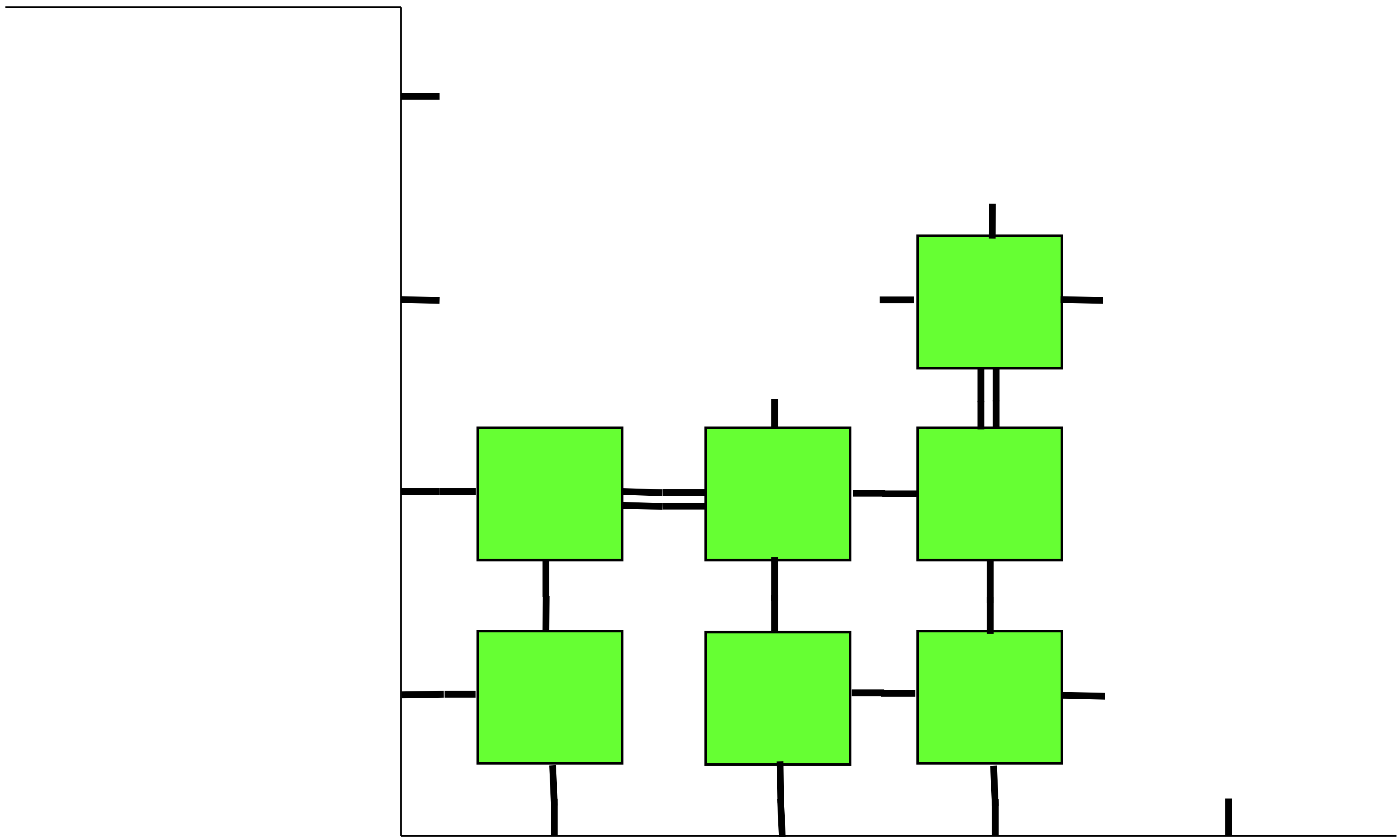
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



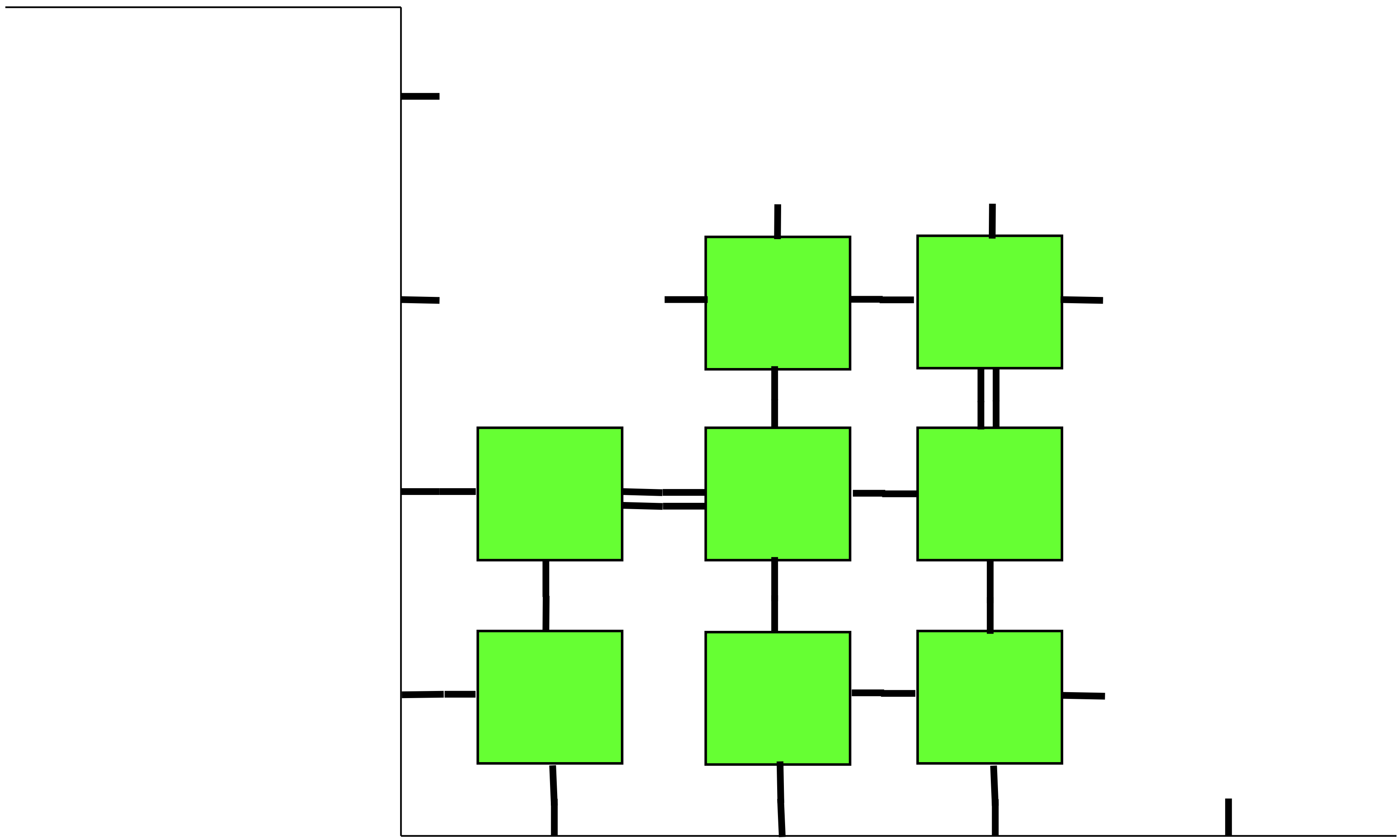
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



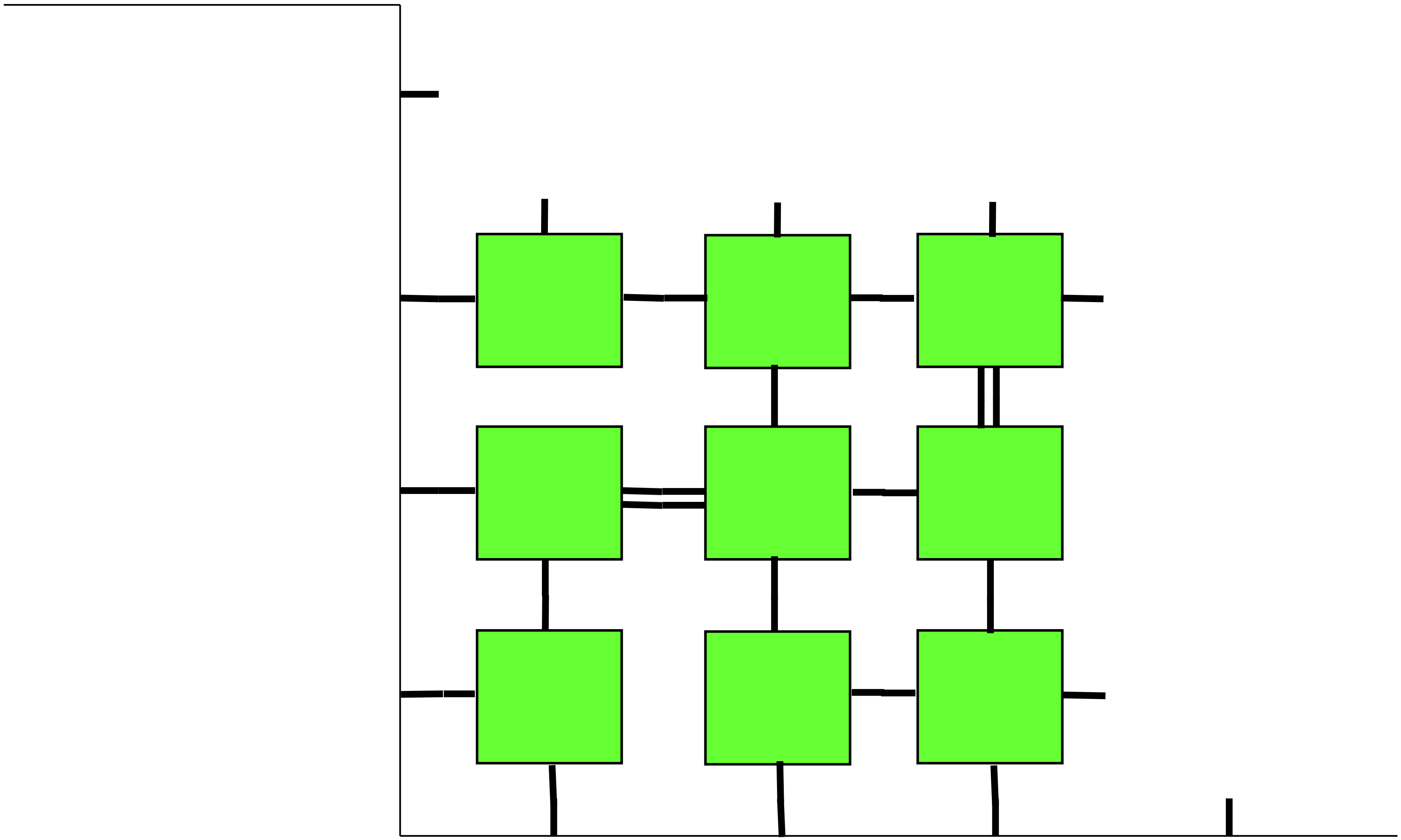
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



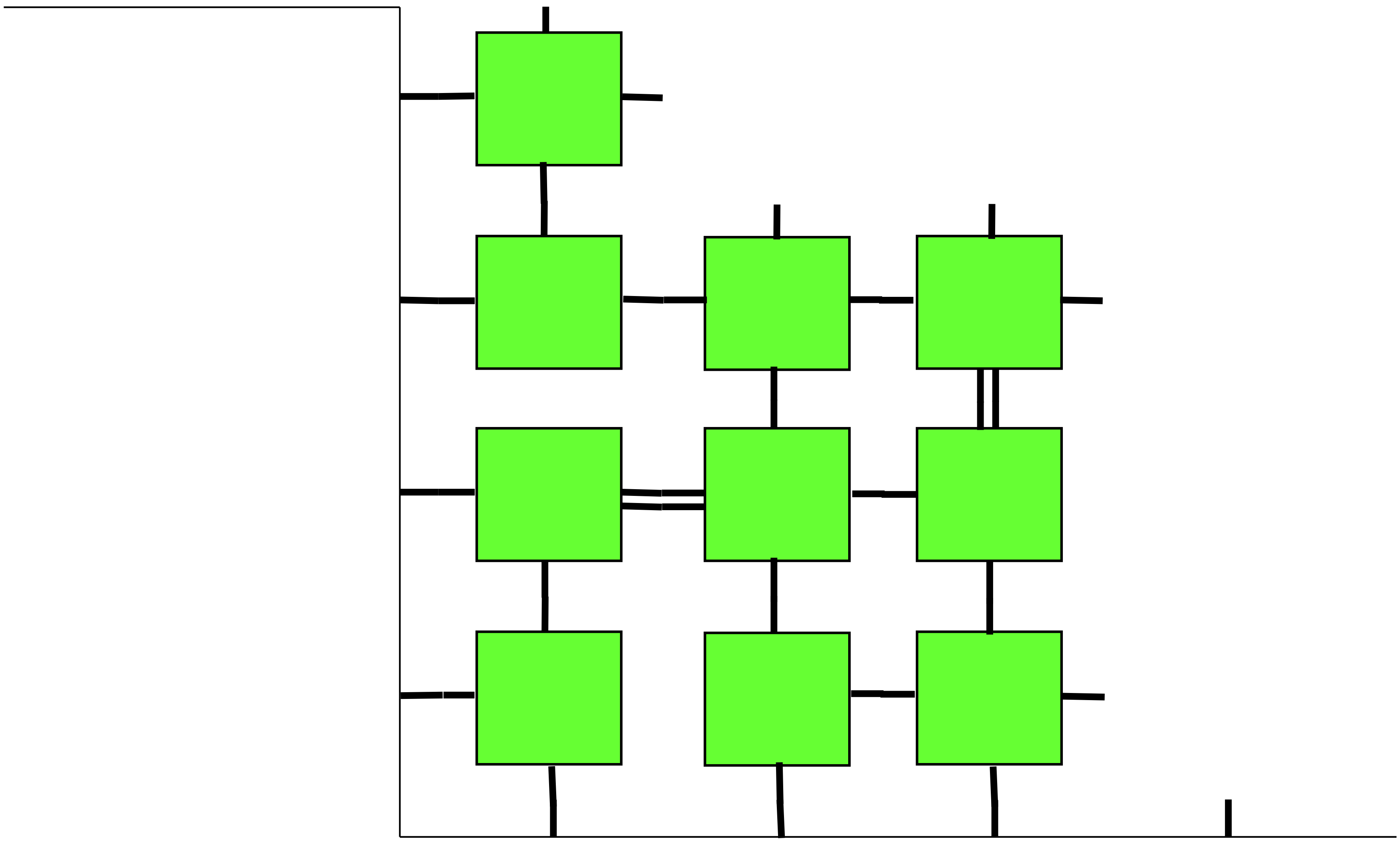
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



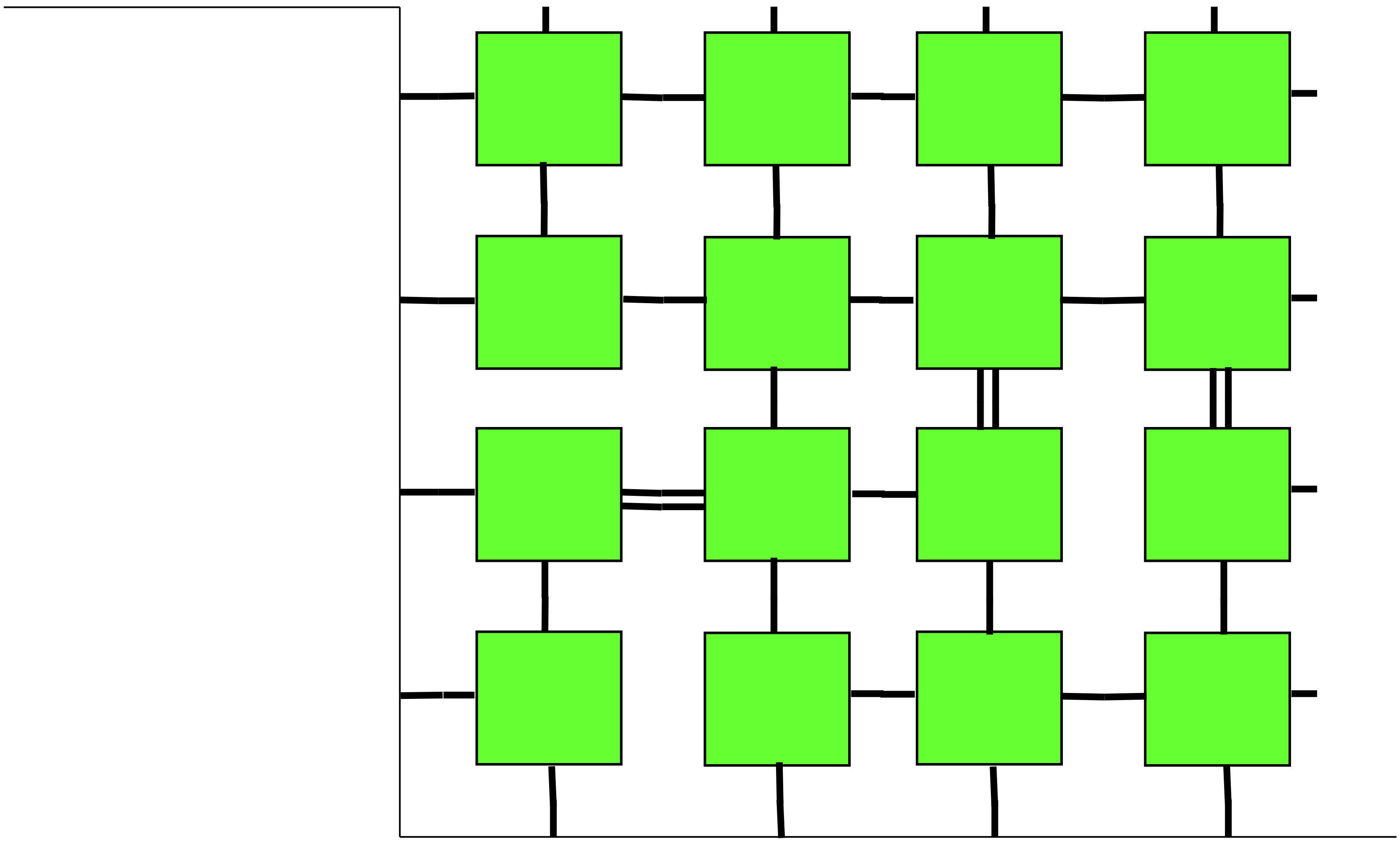
[Cheng, Goel, Cheng]

Example: Four by Four Snake Tiles



[Cheng, Goel, Cheng]

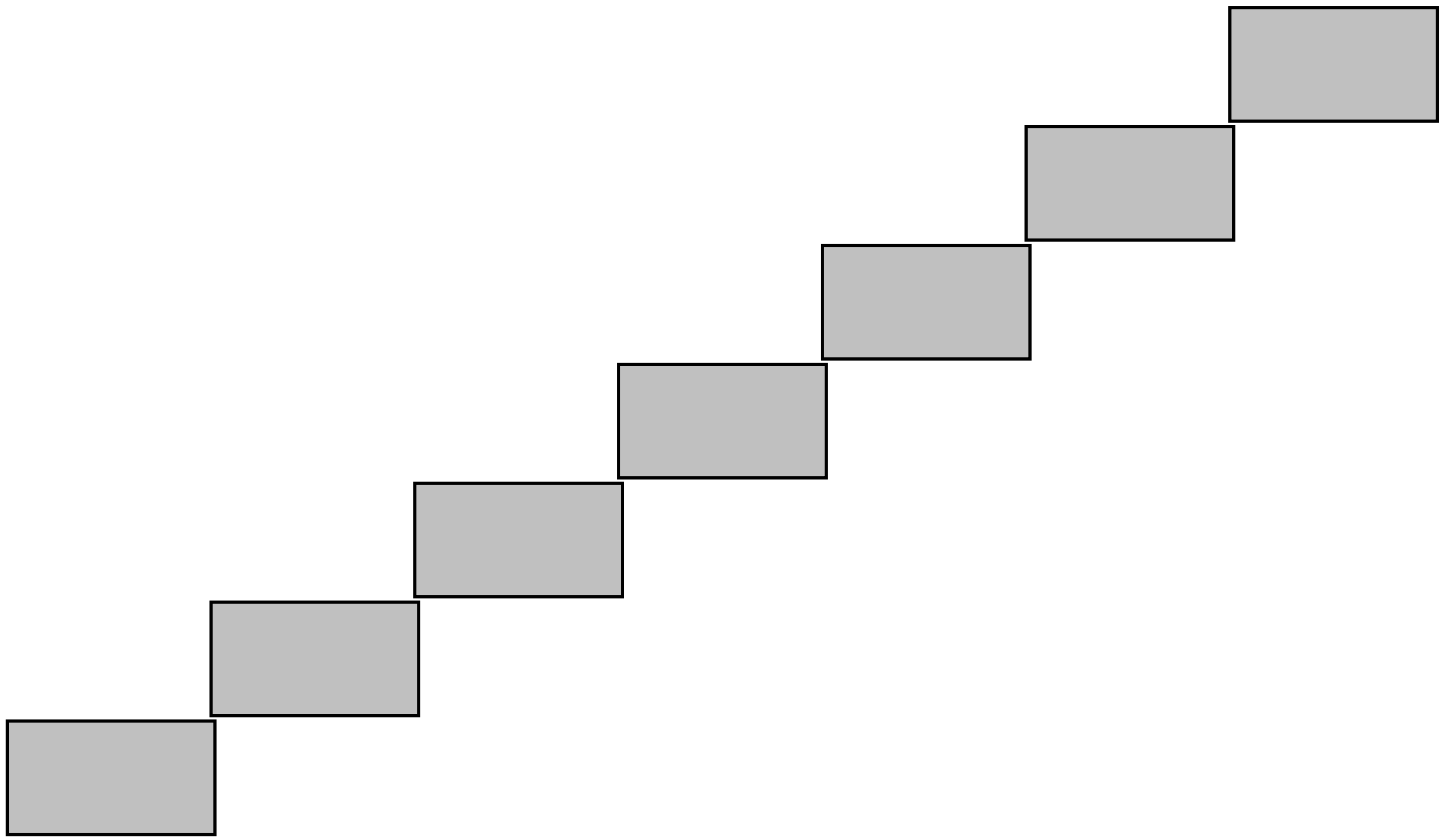
Example: Four by Four Snake Tiles



[Cheng, Goel, Cheng]

Example: Nucleation Errors (T=2) for Diagonal Tile Assemblies

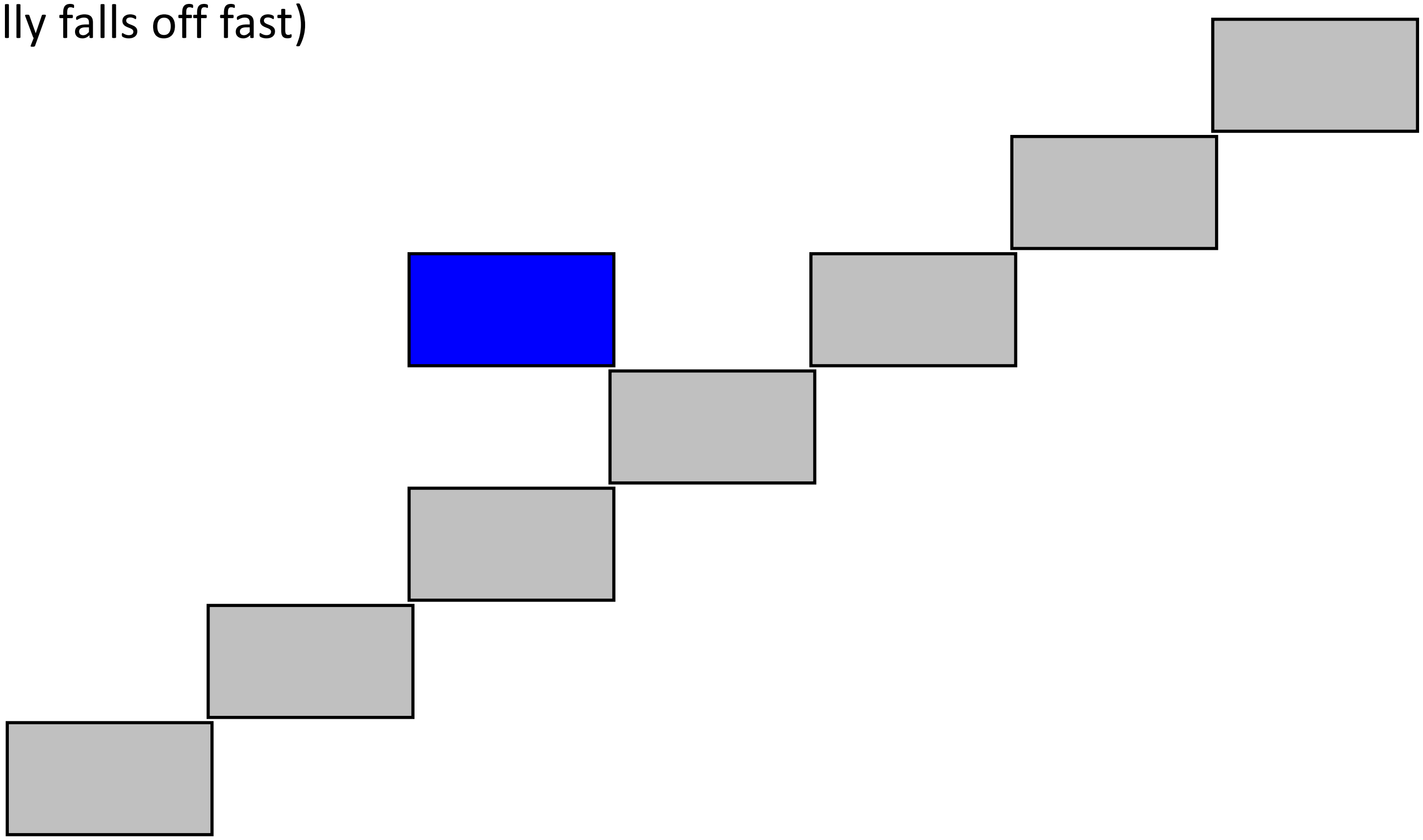
Starting from an initial assembly



[Cheng&Goel,2004]

Example: Nucleation Errors (T=2) for Diagonal Tile Assemblies

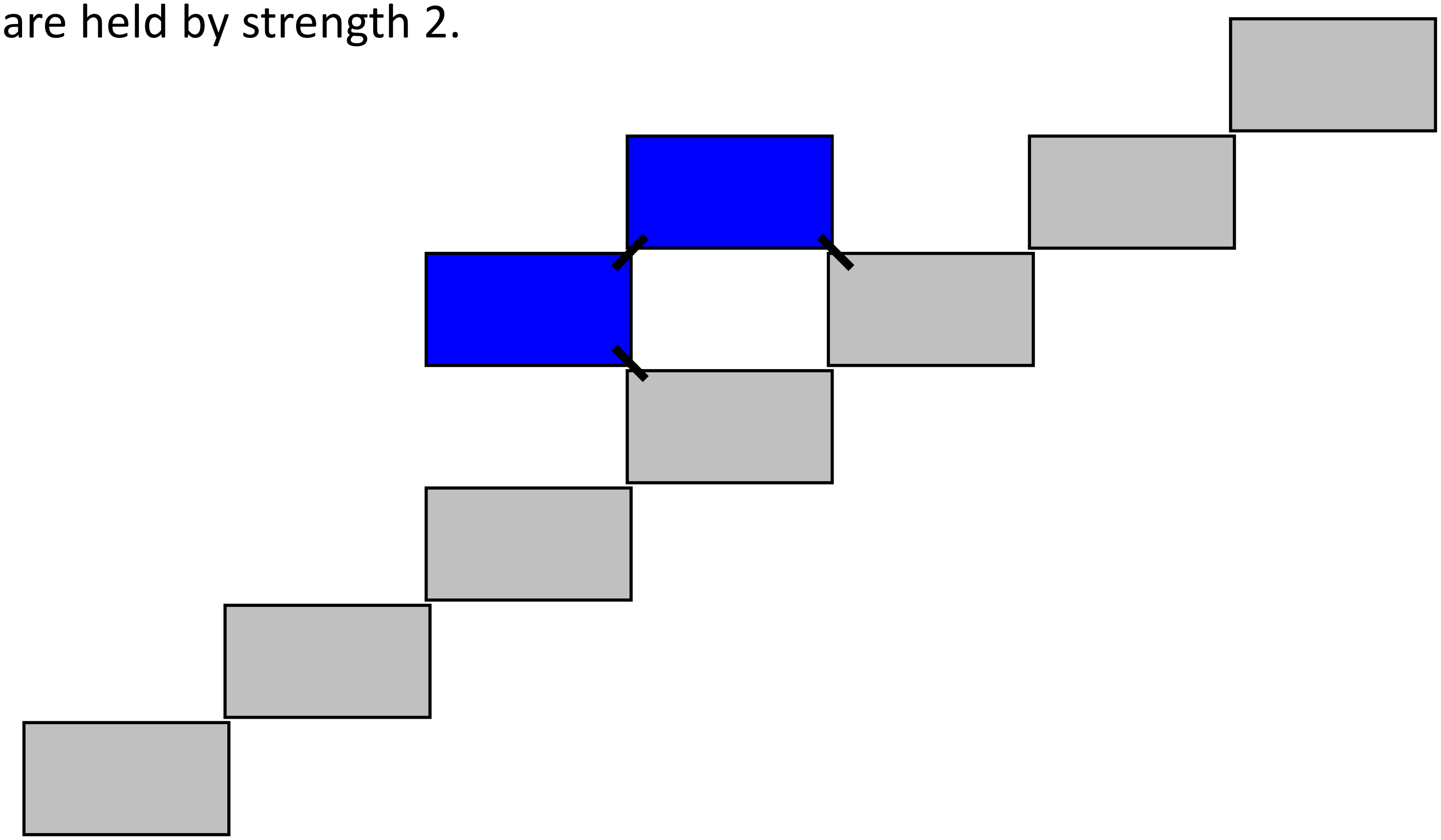
The first tile attaches with strength 1.
(usually falls off fast)



[Cheng&Goel,2004]

Example: Nucleation Errors (T=2) for Diagonal Tile Assemblies

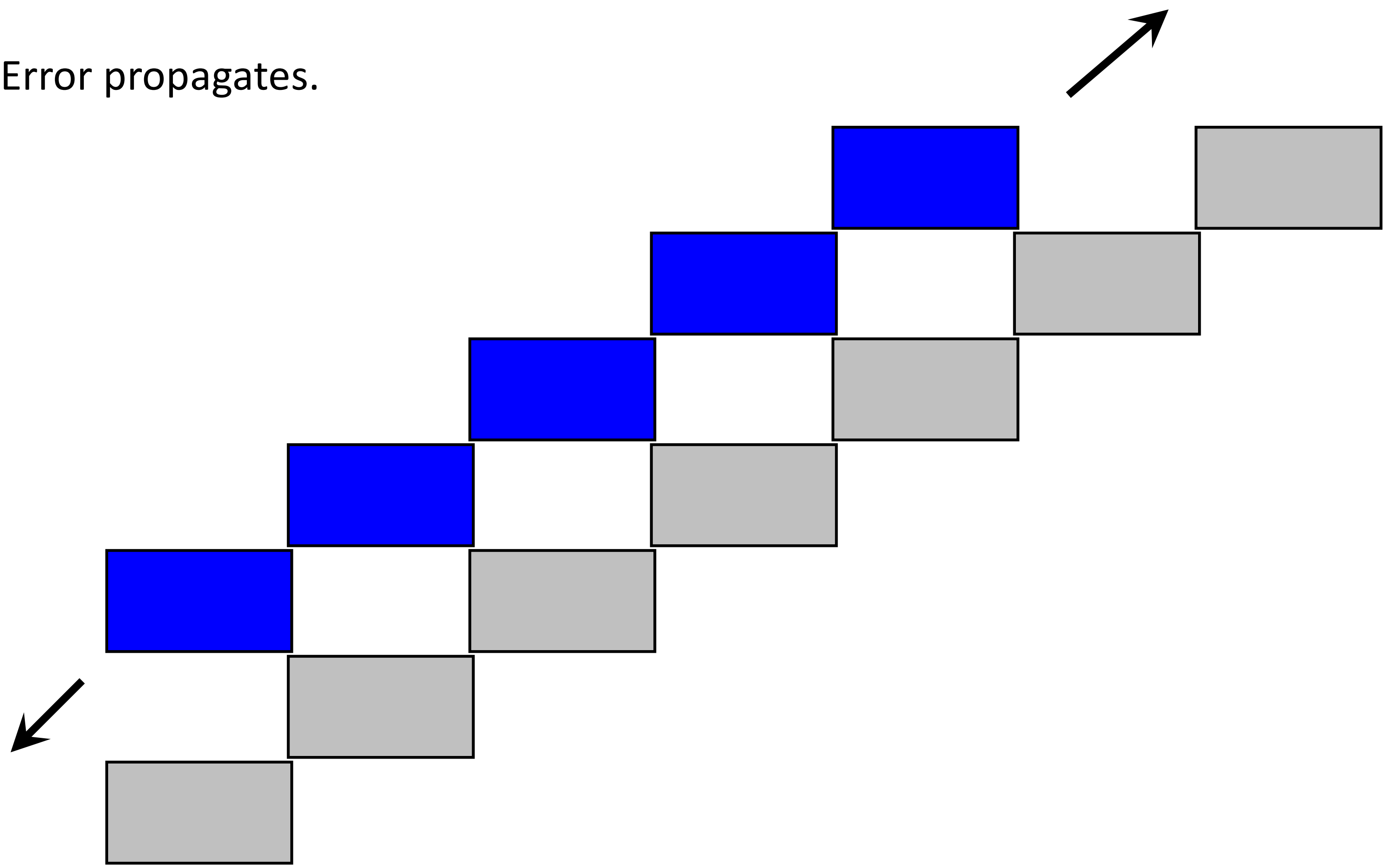
The second tile attaches and now both tiles are held by strength 2.



[Cheng&Goel,2004]

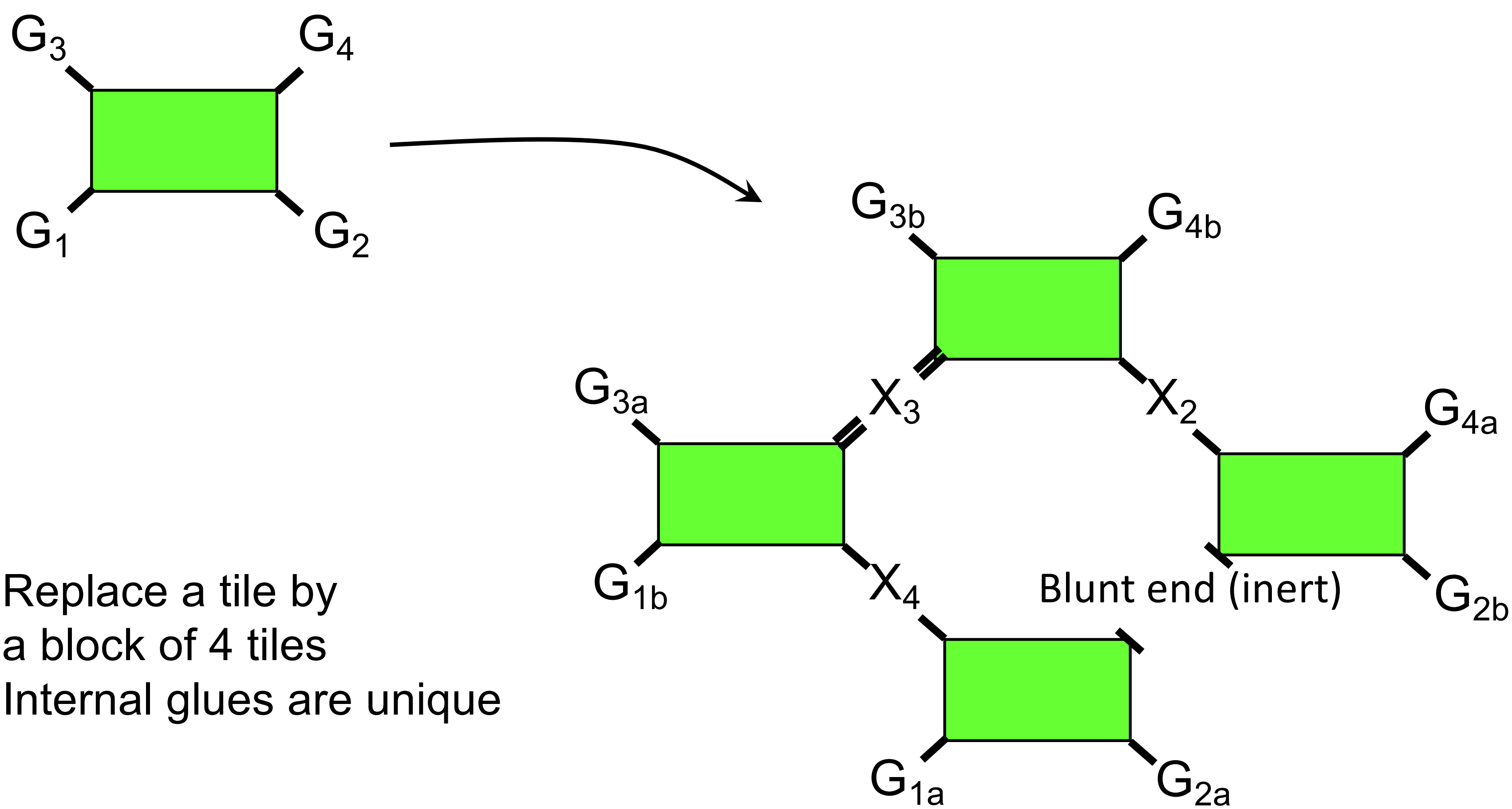
Example: Nucleation Errors (T=2) for Diagonal Tile Assemblies

Error propagates.



[Cheng&Goel,2004]

Snaked Tile System for Diagonal Tile Assemblies

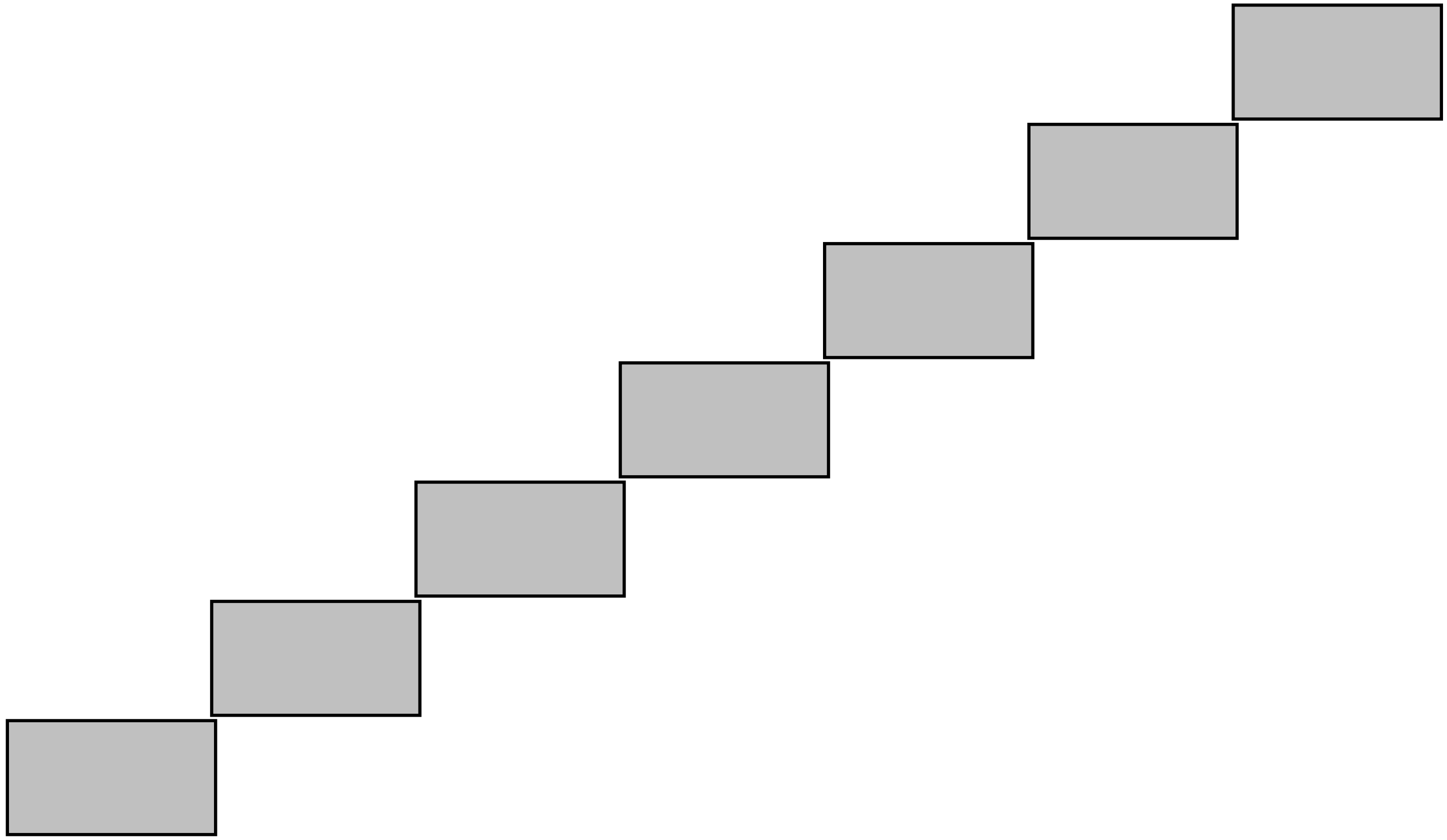


- Replace a tile by a block of 4 tiles
- Internal glues are unique

[Cheng&Goel,2004]

Example: Nucleation Errors (T=2) reduced for Diagonal Tile Assemblies using Snaked Tiles

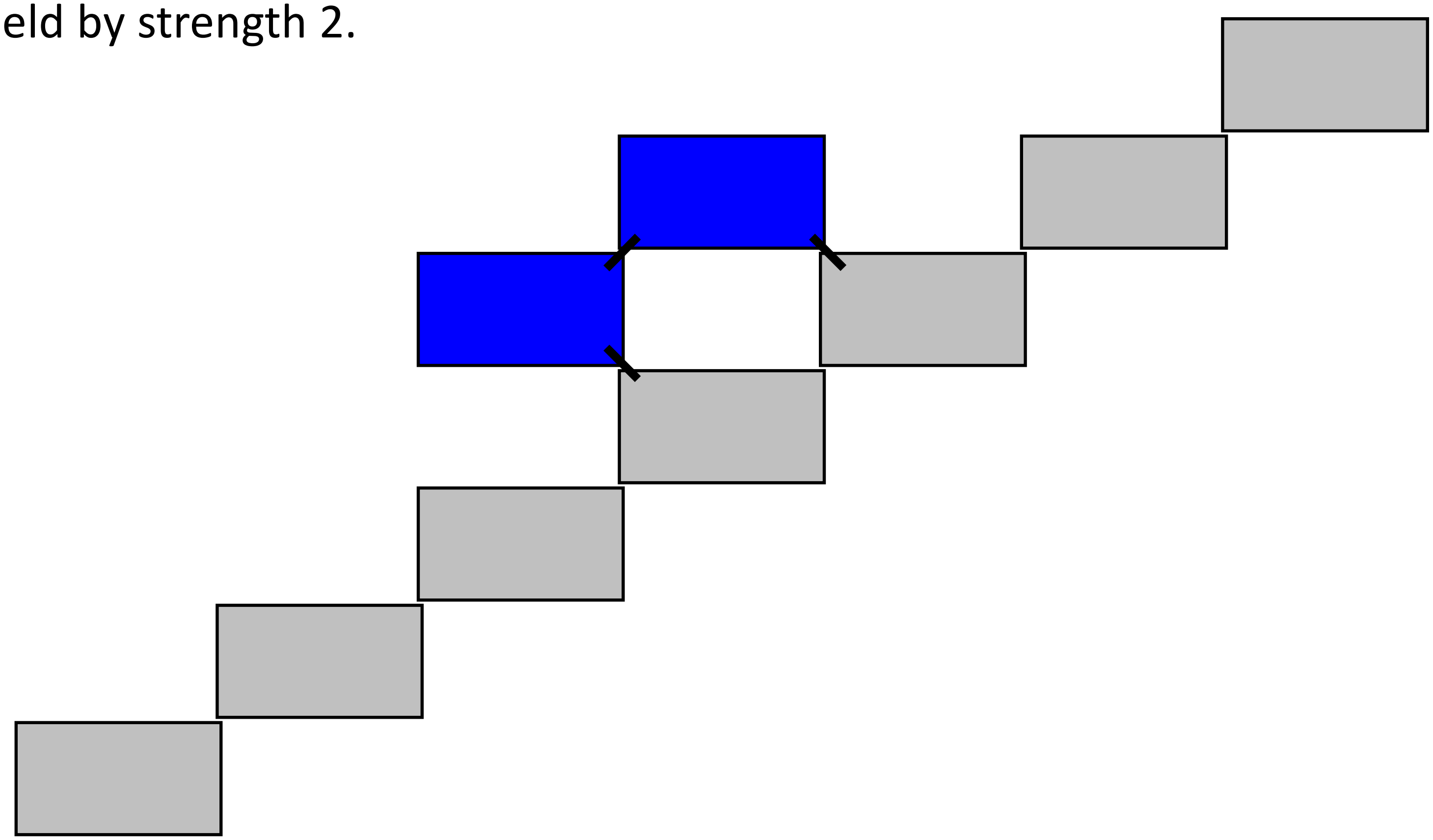
Starting from an initial assembly



[Cheng&Goel,2004]

Example: Nucleation Errors (T=2) reduced for Diagonal Tile Assemblies using Snaked Tiles

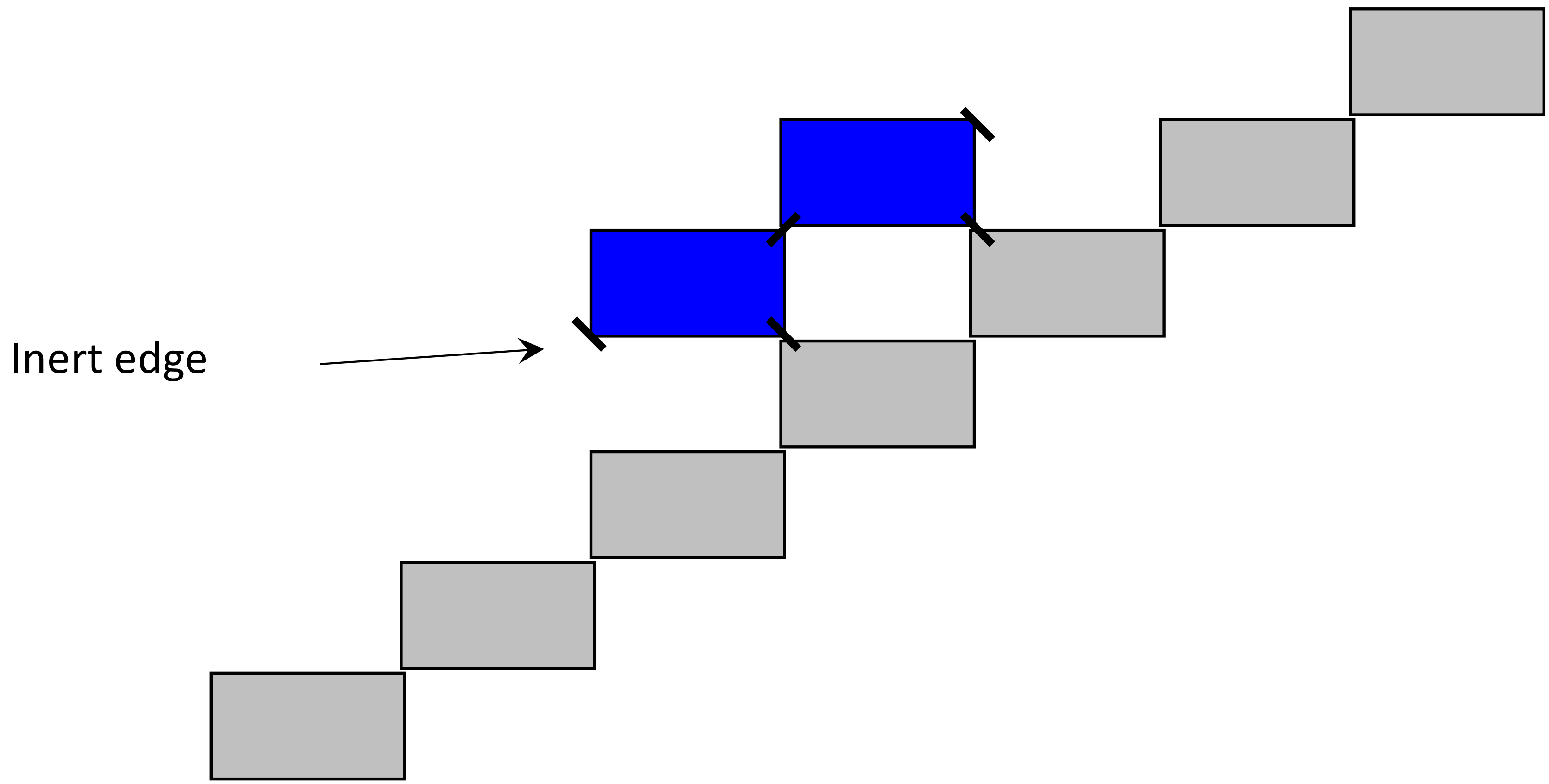
Two tiles attach and both tiles are held by strength 2.



[Cheng&Goel,2004]

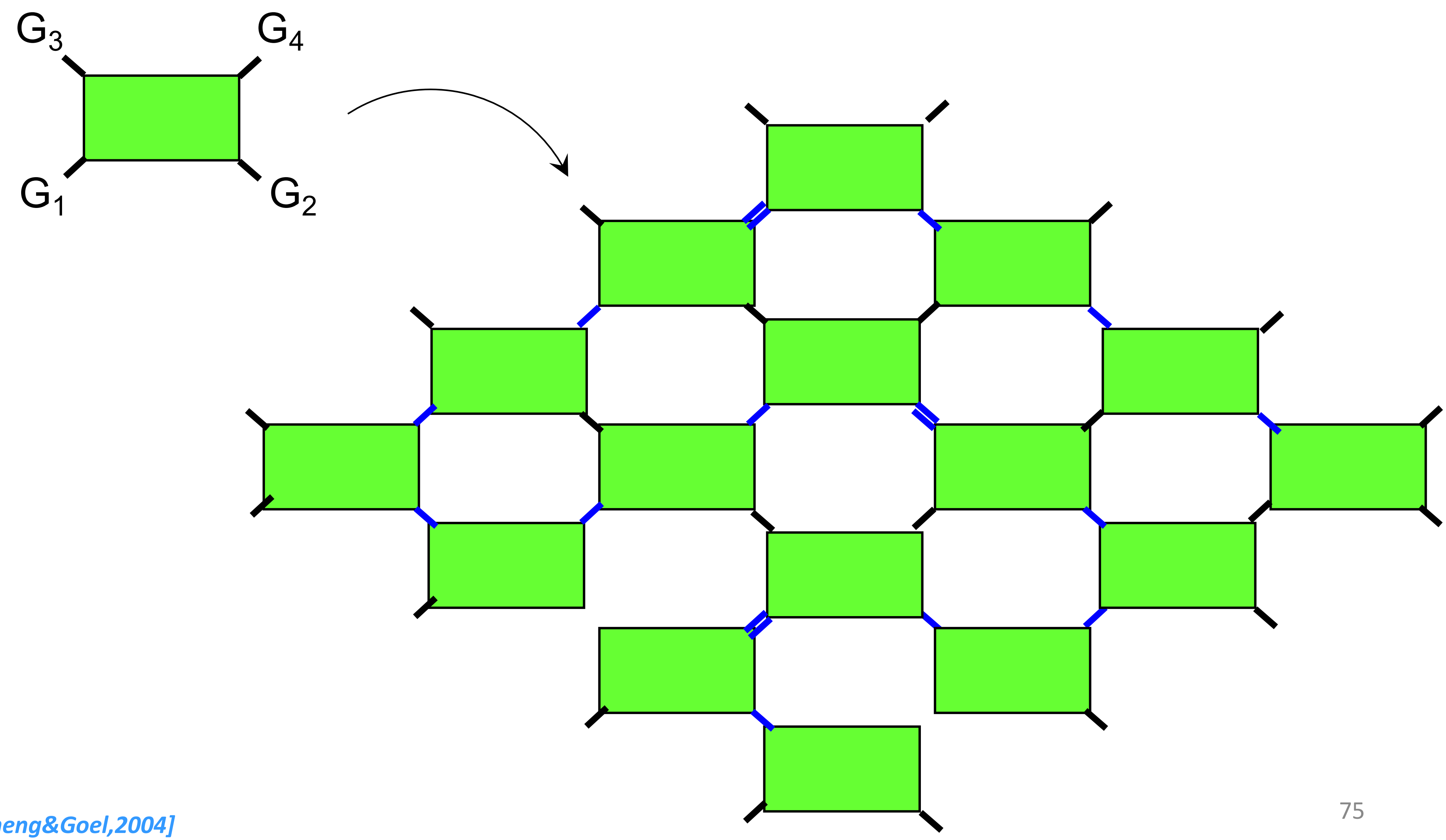
Example: Nucleation Errors (T=2) reduced for Diagonal Tile Assemblies using Snaked Tiles

No other tiles can attach.



[Cheng&Goel,2004]

Generalization of Snaked Tiles to Diagonal Tile Assemblies



[Cheng&Goel,2004]

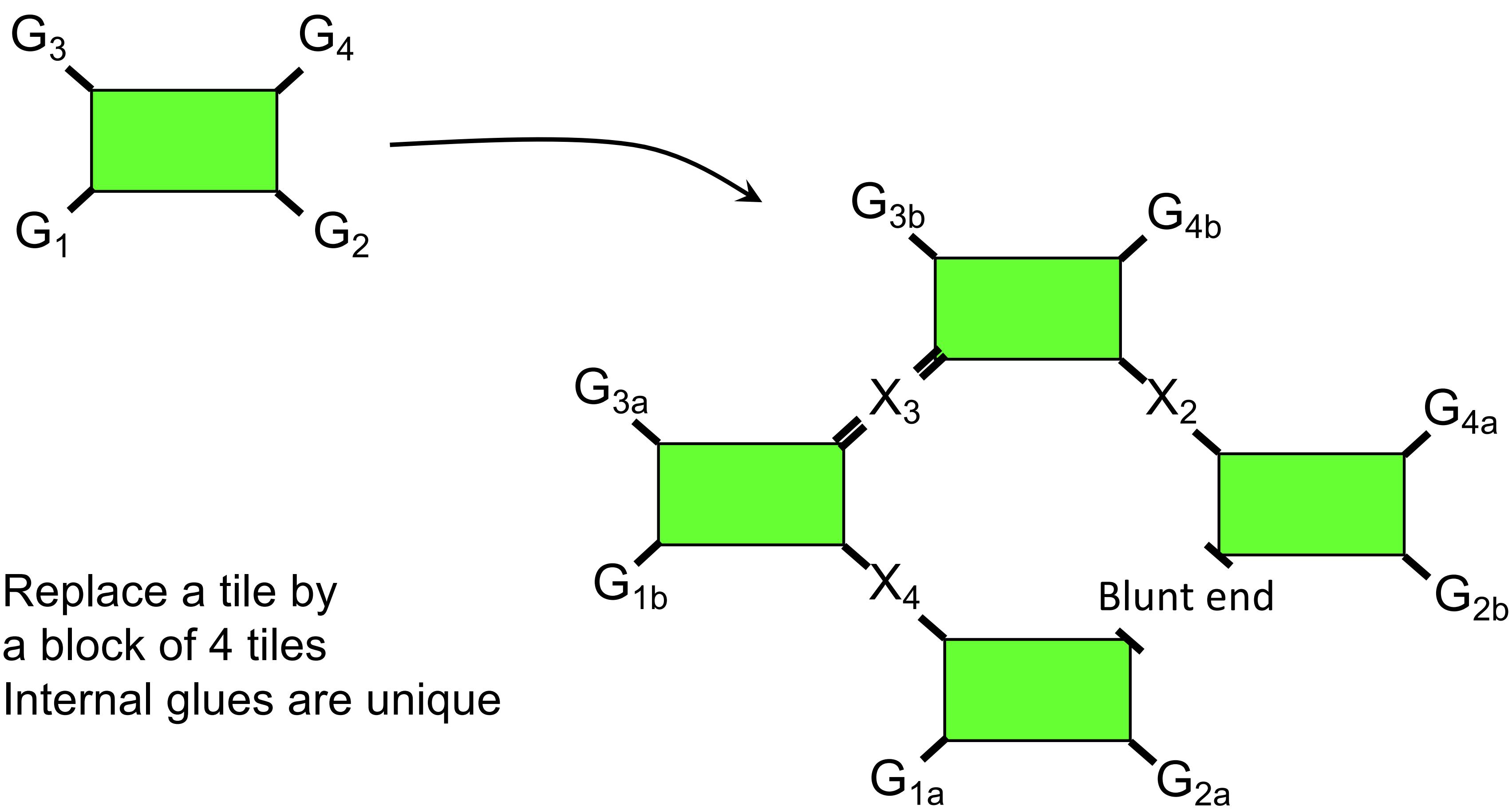
Experimental Verification of Proof-Reading & Snaked Tiles for Reducing Nucleation Errors

Ashish Goel
Rebecca Schulman
Erik Winfree

[Cheng,Schulman,Winfree]

Snaked Tile System for Diagonal Tile Systems

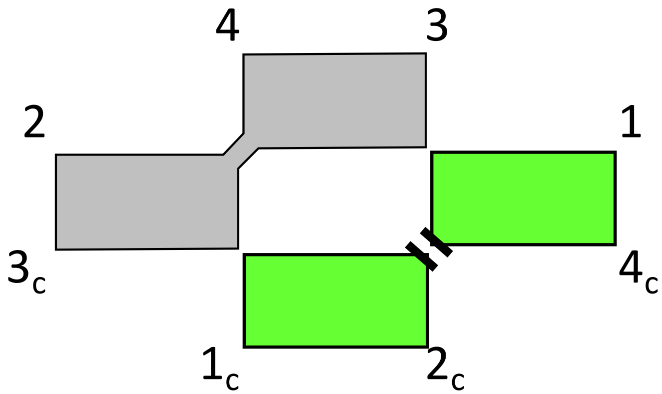
[Chen, Goel, 2004]

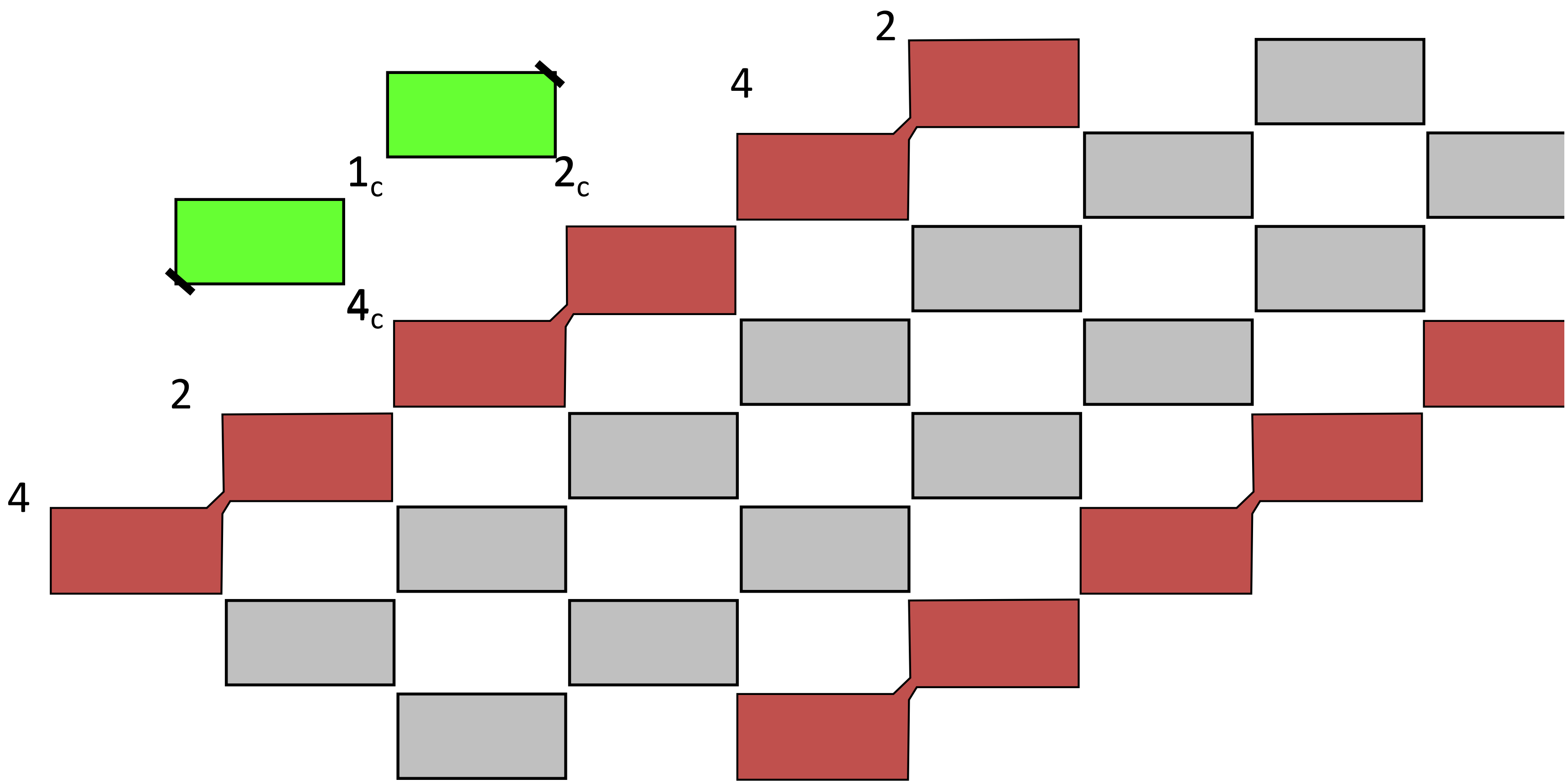


- Replace a tile by a block of 4 tiles
- Internal glues are unique

[Cheng,Schulman,Winfree]

Snaked Tile System for Diagonal Tile Systems

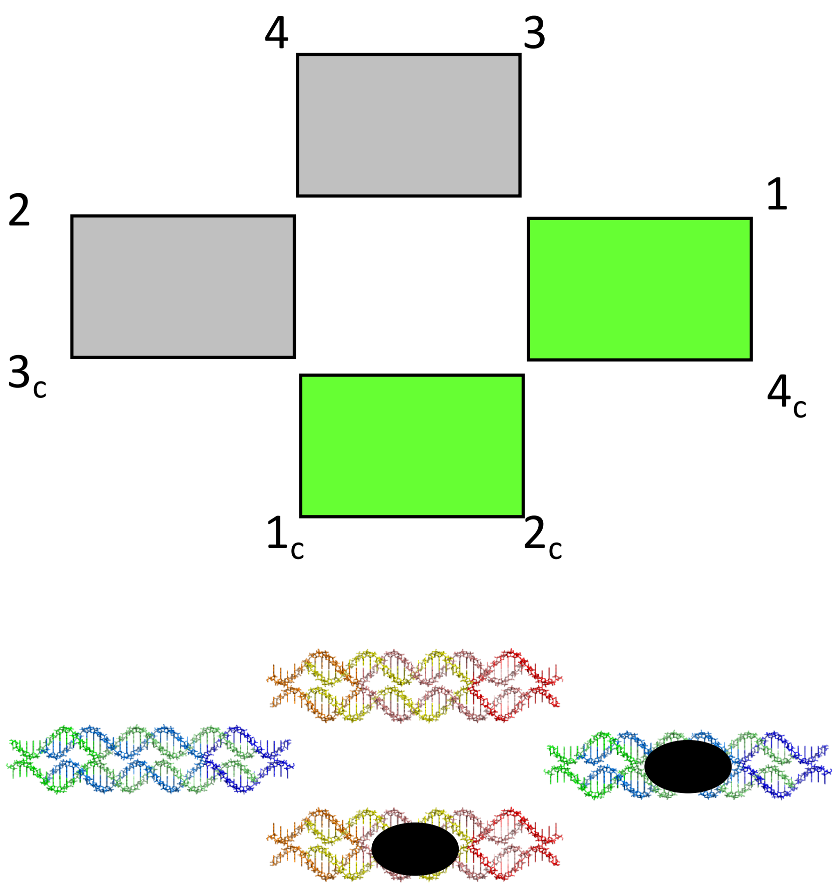




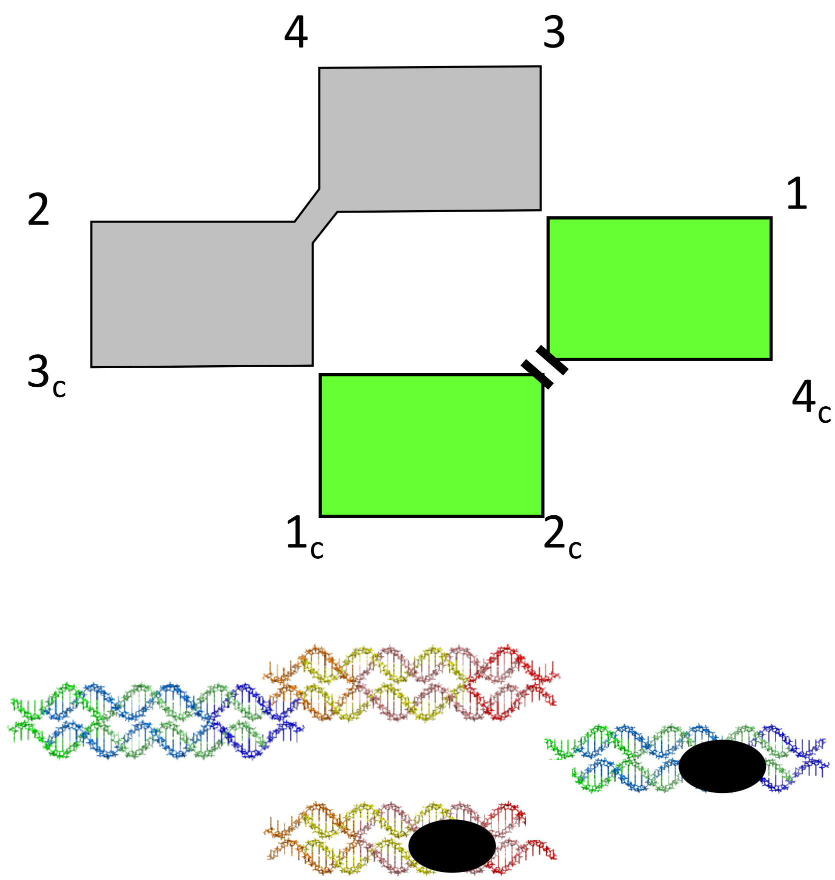
[Cheng,Schulman,Winfree]

Tile sets used in experiments

Proofreading block



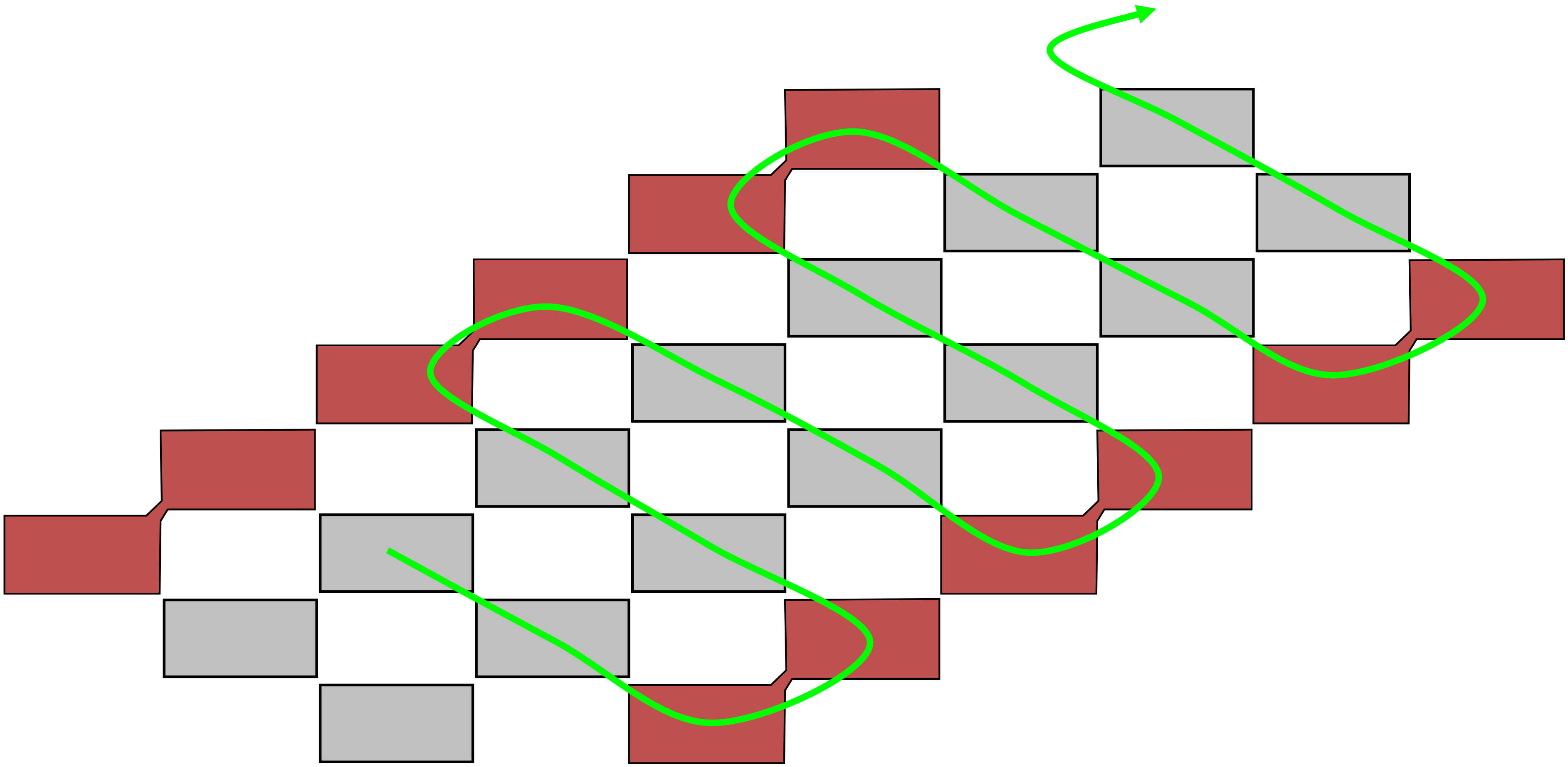
Snaked block



[Cheng,Schulman,Winfree]

Width-4 Zig-Zag Ribbon using Snaked Tiles

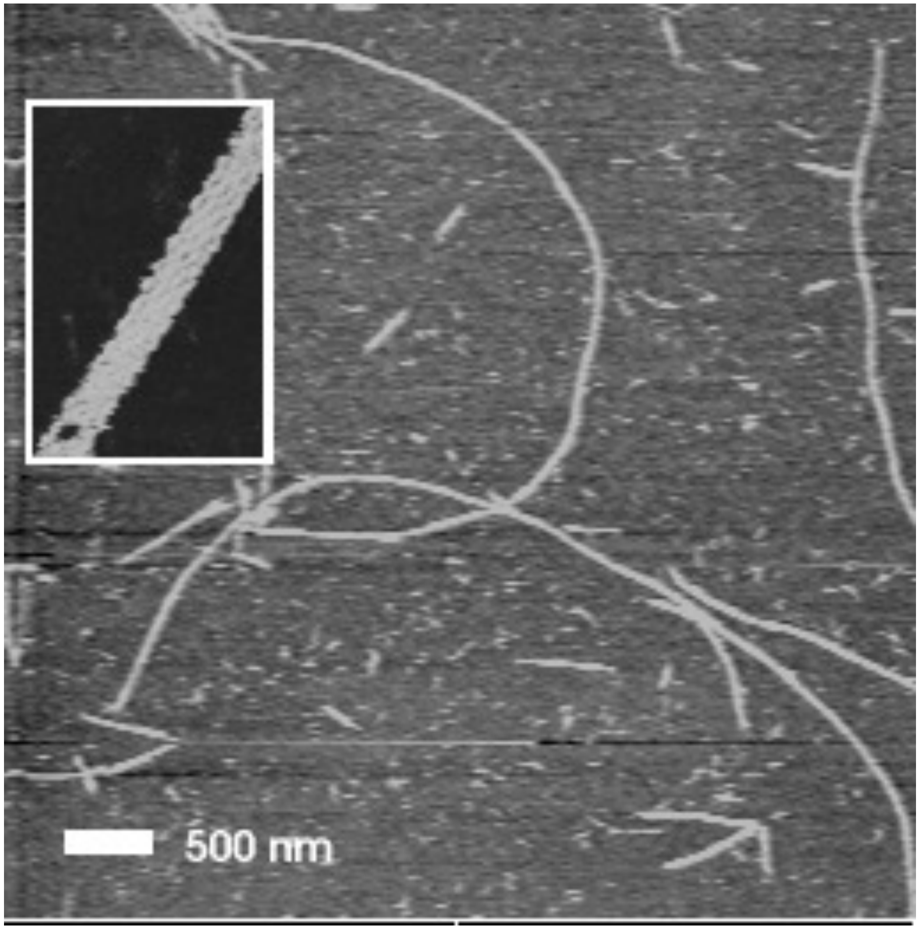
[Schulman, Winfree, DNA 10, 2004]



6 tile types

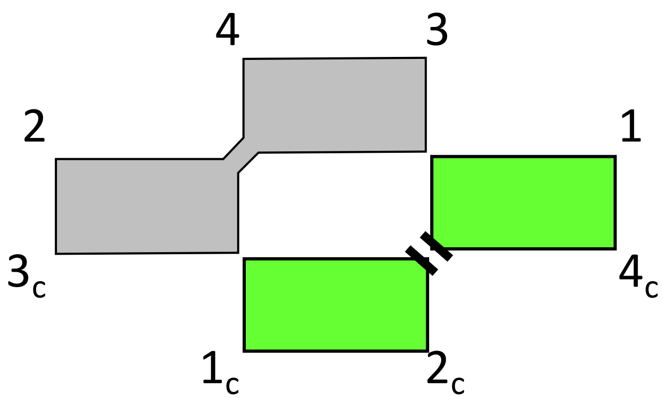
[Cheng,Schulman,Winfree]

AFM of Zig-Zag Ribbons

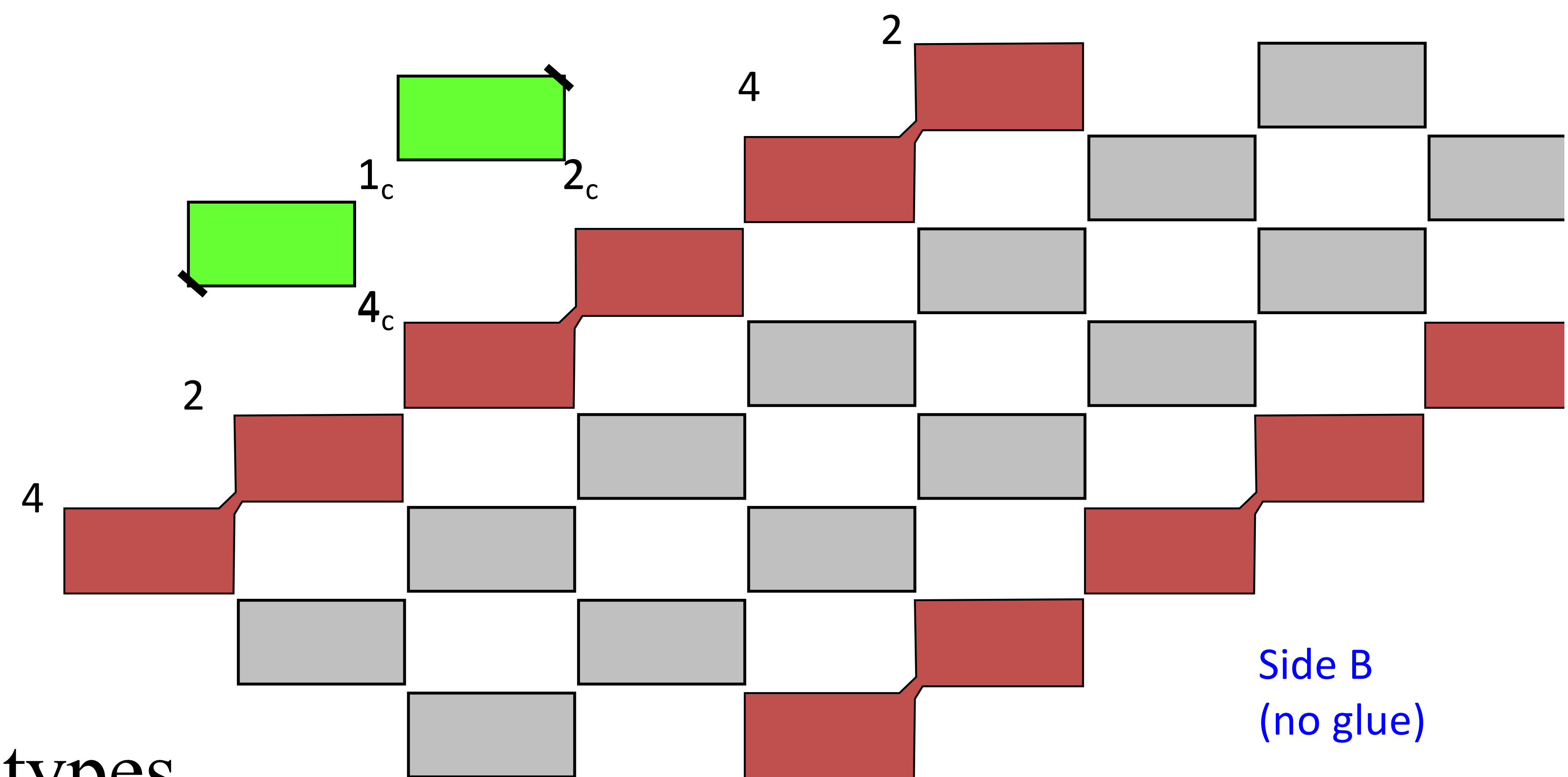


[Cheng,Schulman,Winfree]

ZZ + Snake Tiles



Side A



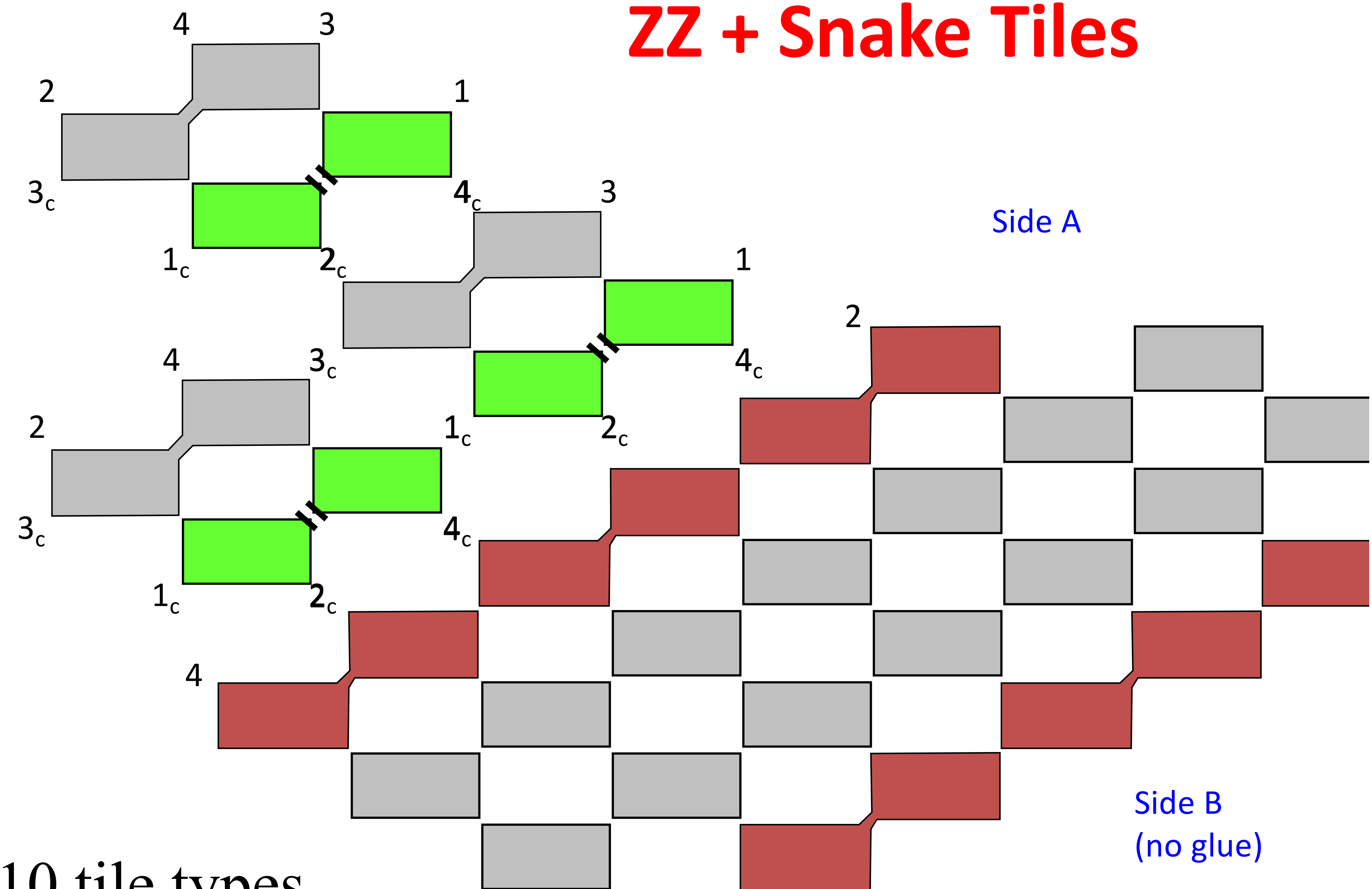
Side B
(no glue)

10 tile types

[Cheng,Schulman,Winfree]

Slow nucleation and growth!

ZZ + Snake Tiles

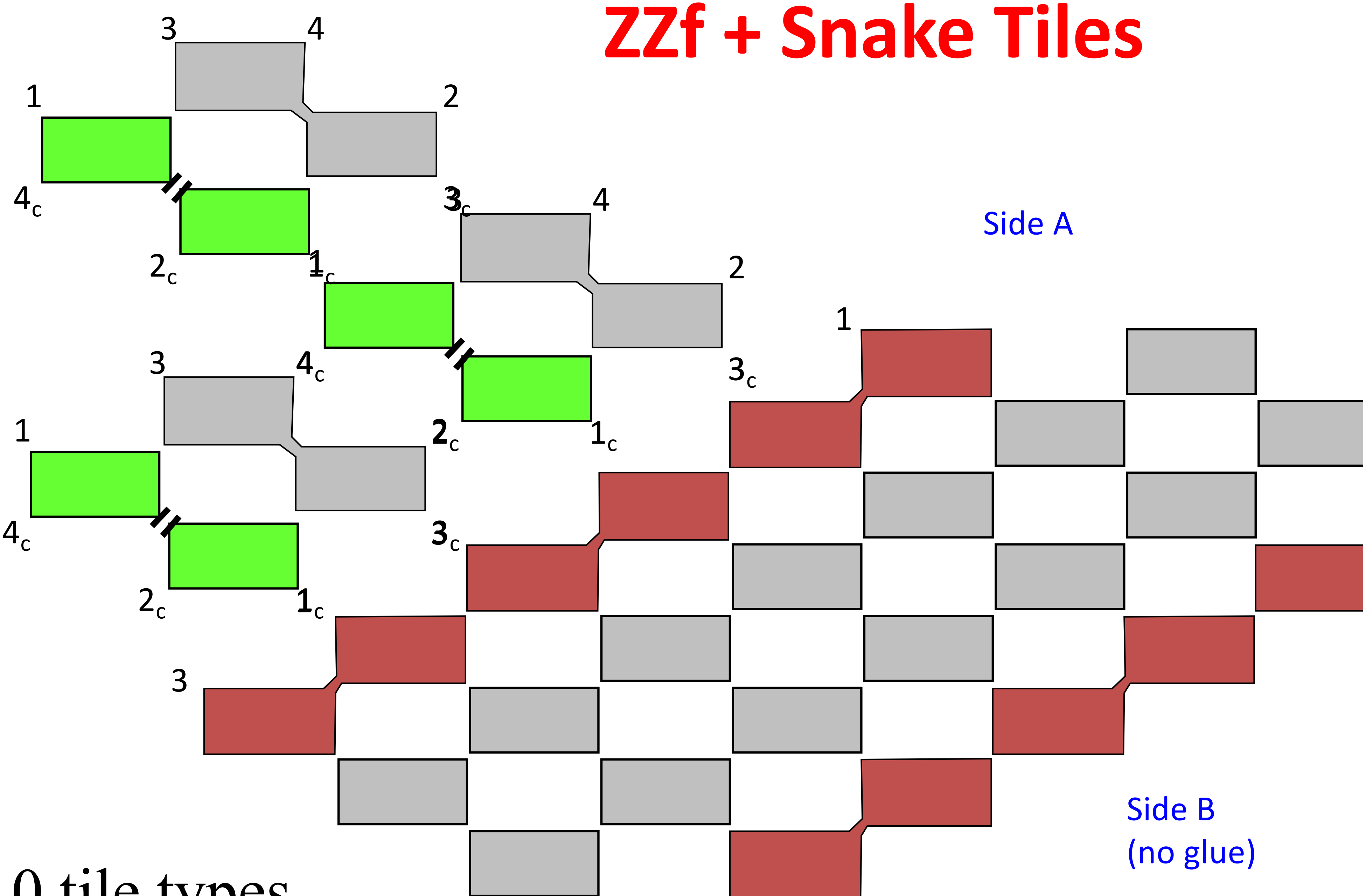


10 tile types

[Cheng,Schulman,Winfree]

Fast nucleation and growth!

ZZf + Snake Tiles



10 tile types

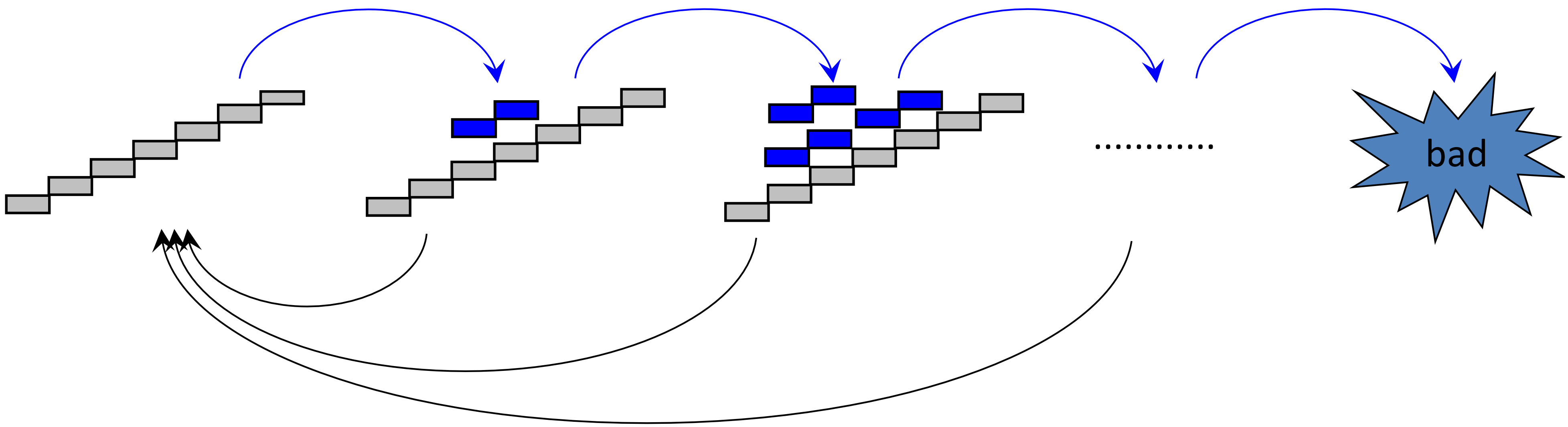
[Cheng,Schulman,Winfree]

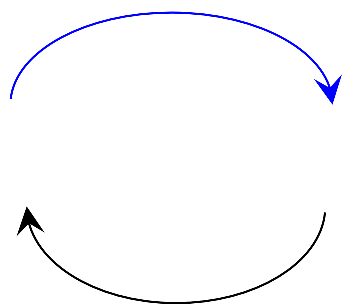
Theoretical Analysis of Snaked Tiles

- The snake tile design can be extended to 2k x 2k blocks.
- Prevents tile propagation even after k-1 insufficient attachments happen.

[Cheng,Schulman,Winfree]

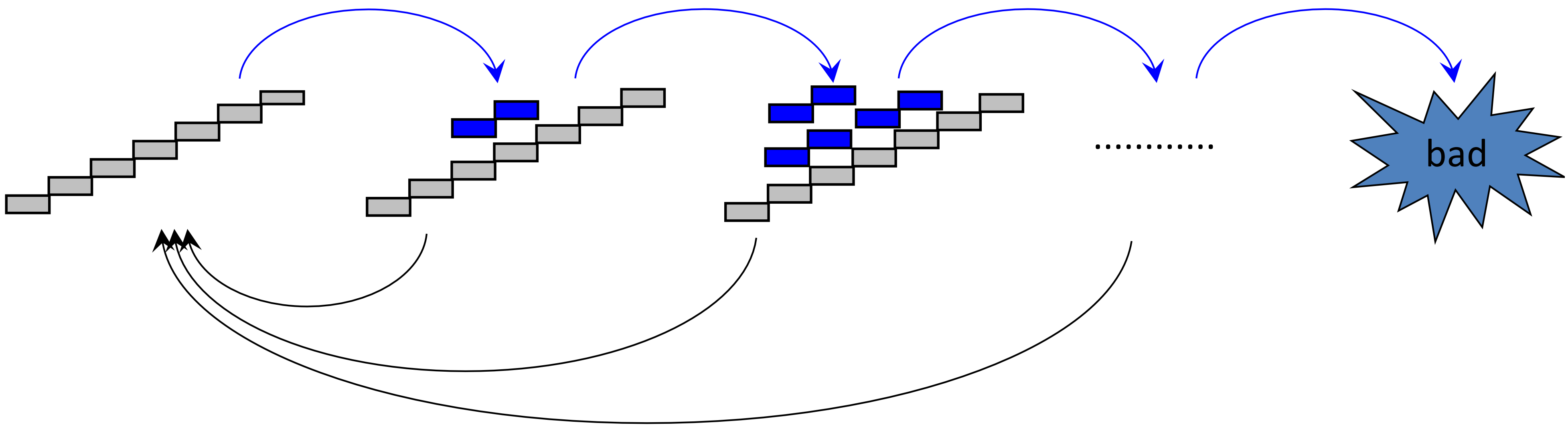
How do Snaked Tiles work?

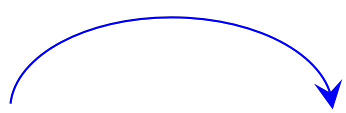




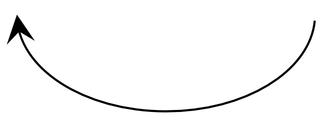
[Cheng,Schulman,Winfree]

How do Snaked Tiles work?





: happens with rate $O(e^{-G}) * r_f$



: erroneous tiles falling off

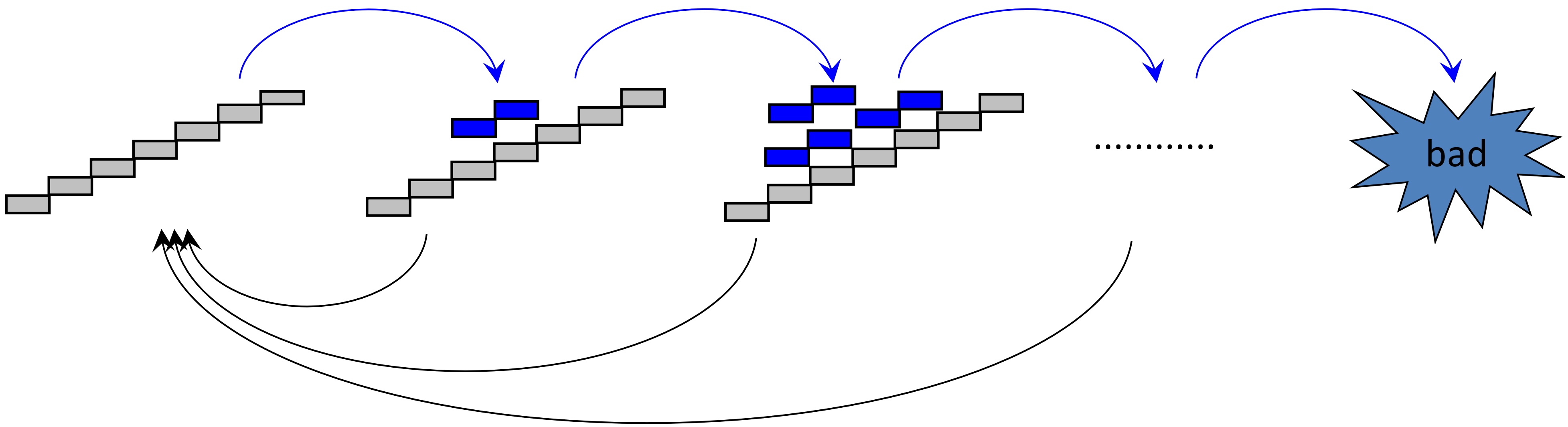
[Cheng,Schulman,Winfree]

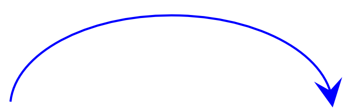
Theoretical Analysis of Snaked Tiles

- The snake tile design can be extended to 2k x 2k blocks.
- Prevents tile propagation even after k-1 insufficient attachments happen.
- When < k insufficient attachments happened locally, all the erroneous tiles are expected to fall off in time poly(k).

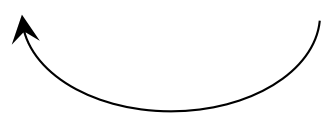
[Cheng,Schulman,Winfree]

Theoretical Analysis of Snaked Tiles





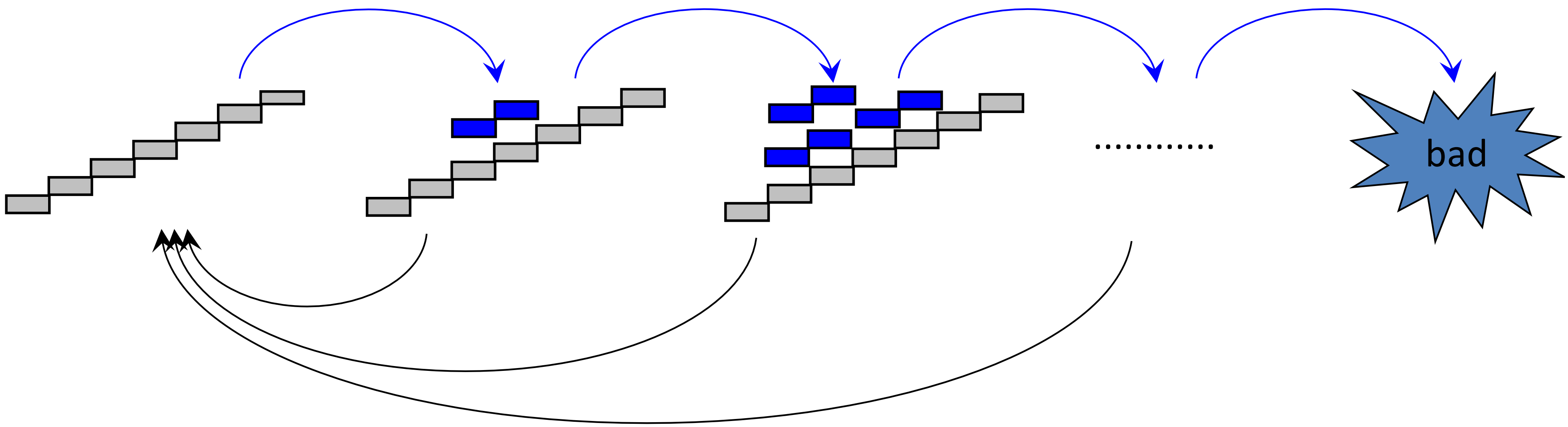
: happens with rate $O(e^{-G}) * r_f$



: happens with rate 1/poly(k)

[Cheng,Schulman,Winfree]

Theoretical Analysis of Snaked Tiles



: happens with rate $O(e^{-kG})$

if backward rate >> forward rate

[Cheng,Schulman,Winfree]

Theoretical Analysis of Snaked Tiles

If we want to assemble a structure with size N, we can use Snaked Tile System with block size k=O(log N).

The assembly process is expected to finish within time Õ(N) and be error free with high probability.

[Cheng,Schulman,Winfree]

Example Parity Tile system:

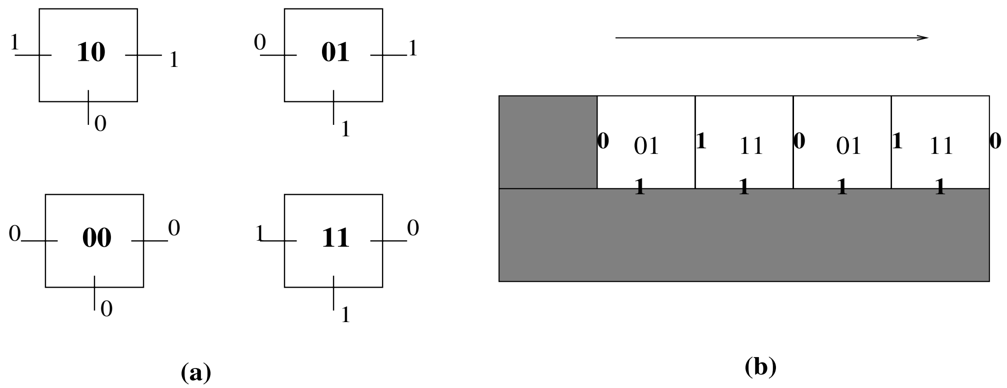


Fig. 1. (a) The parity tile system. (b) Illustrating the action of the parity tile system on the "input" string 1111. The arrow at the top represents the order in which tiles must attach in the absence of errors

Example Parity Tile system with Proof-reading Tiles and Snaked Tiles:

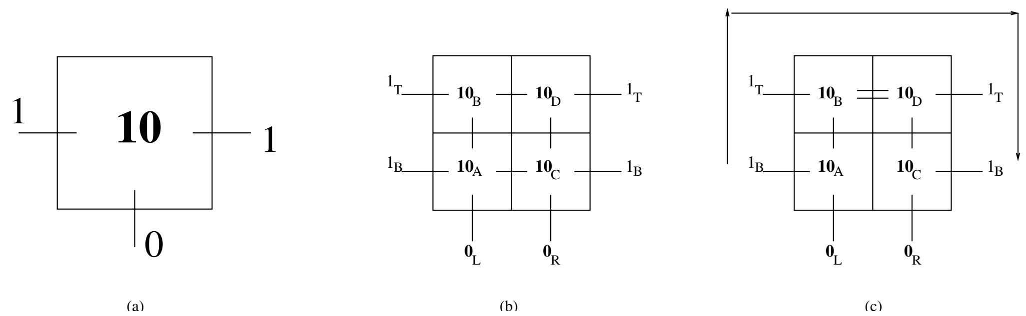


Fig. 2. (a) The original 10 tile. (b) The four proof-reading tiles for the 10 tile, using the construction of Winfree and Bekbolatov [11]. (c) The snaked proof-reading tiles for the parity tile system. The internal glues are all unique to the 2×2 block corresponding to the 10 tile. Notice that there is no glue on the right side of 10_A or the left side of 10_C and that the glue between the top two tiles is of strength 2. This means that the assembly process doubles or "snakes" back onto itself, as demonstrated by the arrow

[Chen&Goel2005]

Theoretical Analysis of Snake Tiling (can skip)

- The snake tile design can be extended to 2k x 2k blocks.
  - Prevents tile propagation even after k+1 nucleation error happen.
- With 2k x 2k snake tile system, we can assemble an N by N square of blocks with time $\tilde{O}(N^{1+4/k+1})$ and (with high probability) remain stable for time $\tilde{O}(N^{1+4/k+1})$.
  - Assuming tiles held by strength 3 does not fall off
  - The error probability at each location changes from p to p^k

[Chen&Goel2005]

Analysis of Errors of Insufficient Attachment

Lemma 1. *The rate at which an insufficient attachment happens at any location in a growing assembly is* $\frac{f^2}{r}e^{-G_{se}} = O(e^{-3G_{se}})$.

Proof. The rate of an insufficient attachment can be modeled as the Markov Chain shown in figure 3. For a nucleation error to happen, first a single tile must attach (at rate f). The fall-off rate of the first tile is $re^{G_{se}}$ and the rate at which a second tile can come and attach to the first tile is f. After the second tile attaches, an insufficient attachment has happened. So the overall rate of an insufficient attachment is $f * \frac{f}{f+re^{G_{se}}} \approx \frac{f^2}{r}e^{-G_{se}}$

[Chen&Goel2005]

Nucleation Errors

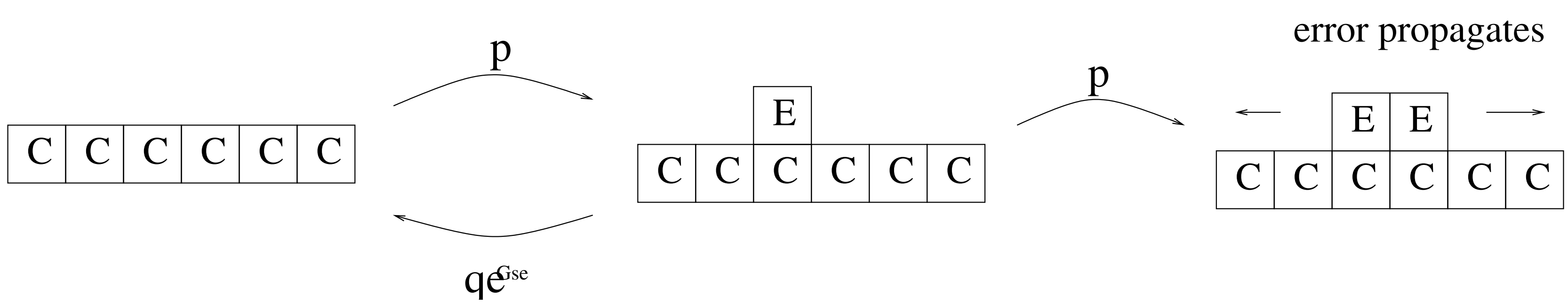


Fig. 3. The C tiles represent the existing assembly, and the E tiles are new erroneous tiles

[Chen&Goel2005]

Analysis of Nucleation Errors

Lemma 2. *The rate at which a nucleation error takes place in our snaked proof-reading system is* $O(e^{-4G_{se}})$.

Proof. In the snaked system, two insufficient attachments need to happen next to each other for a nucleation error to occur. According to lemma 1, the first insufficient attachment happens at rate $O(e^{-3G_{se}})$. After the first insufficient attachment, the error will eventually be corrected unless another insufficient attachment happens next to the first. The second insufficient attachment happens at rate $O(e^{-3G_{se}})$; but the earlier insufficient attachment gets "corrected" at rate $O(e^{-2G_{se}})$ (remember that $a \approx 1$ and hence a tile attached with strength 2 falls off at roughly the growth rate). Hence, the probability of another insufficient attachment taking place before the previous insufficient attachment gets reversed is $O(e^{-G_{se}})$, bringing the nucleation error rate down to $O(e^{-4G_{se}})$.

[Chen&Goel2005]

Snaked Assembly Growth

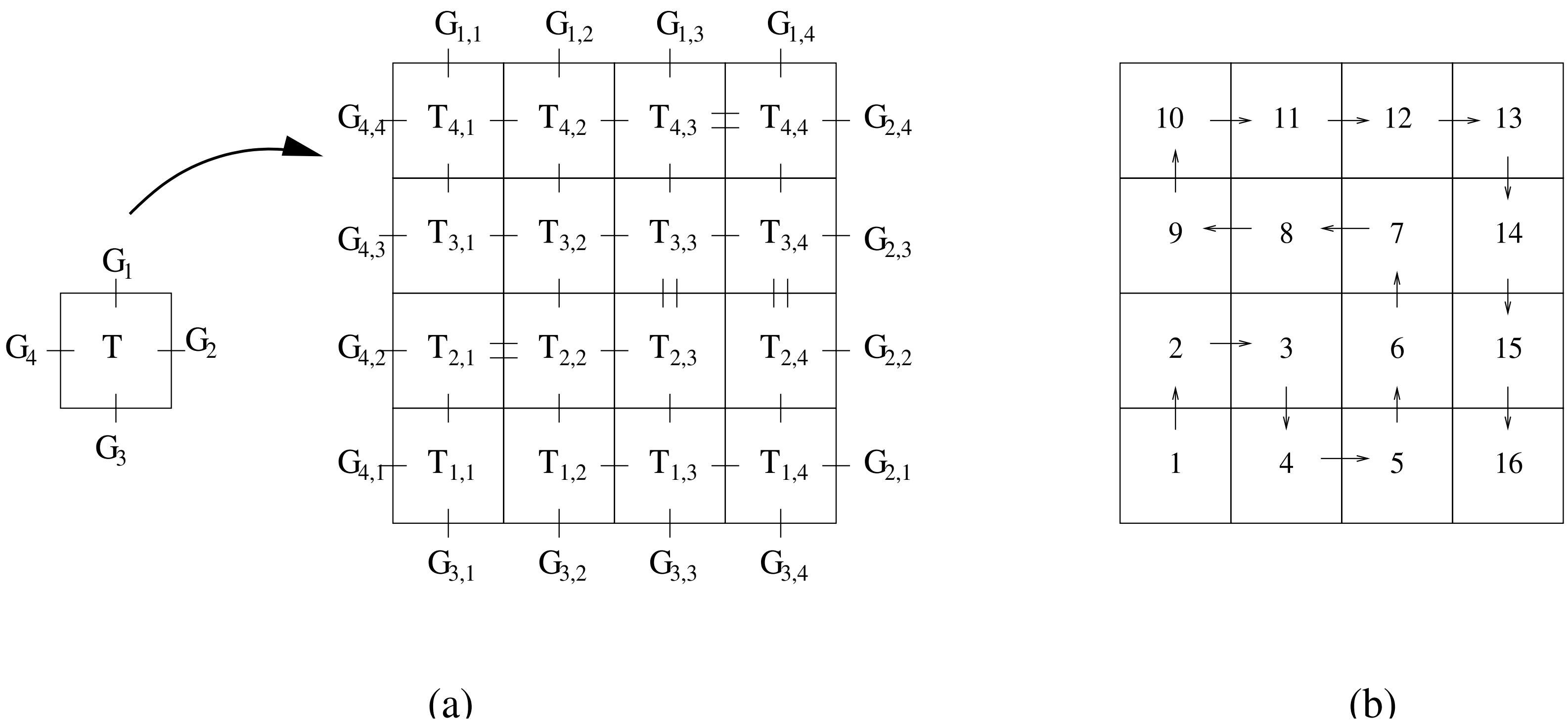


Fig. 4. (a) The structure of 4x4 block. (b) The order of the growth

[Chen&Goel2005]

Analysis of Snaked Assembly Growth

Theorem 2. *With a $2k \times 2k$ snaked tile system, $k = O(\log n)$, assuming that we can set $e^{G_{se}}$ to be $O(k^6)$, an $n \times n$ square of blocks can be assembled in time $\widetilde{O}(n)$ and with high probability, no block errors happen for $\widetilde{\Omega}(n)$ time after that.*

[Chen&Goel2005]

Analysis of Snaked Assembly Growth

Lemma 3. *Consider any connected structure caused by m insufficient attachments ($1 \le m \le k$). Then the width of the structure can be at most $2m$, and the height of the structure can be at most $2k$ (i.e., this connected structure can only span two blocks). This structure will fall off in expected time $O(\frac{k^5}{r})$ unless there's a block error somewhere in the assembly or an insufficient attachment happens within the (at most two) blocks spanned by the structure.*

PROOF OUTLINE: The proof of this lemma involves a lot of technical details. Due to space constraints, we only present a sketch in this version. In the structure of $2k \times 2k$ snaked tiles, all the glues between the $(2i)$-th row and $(2i+1)$-th row have strength 1 for all i. So, to increase the width from $2i$ to $2i+1$, we must have at least one insufficient attachment. So, with m insufficient attachments, the width of the structure can be at most $2m$. Using similar arguments, the height of the structure can be at most $2k$. Also, the attached tiles can be partitioned into $O(k)$ parts. Each of these parts can be viewed as a $2 \times O(k)$ rectangle with every internal glue having strength 1. The process of tiles attaching to and detaching from each rectangle can be modeled using two orthogonal random walks and hence, each rectangle will fall off in expected time $O(\frac{k^4}{r})$. The different rectangles can fall off sequentially, and after one rectangle falls off completely, none of its tiles will attach again unless an insufficient attachment happens. Thus, the structure will fall off in expected time $O(\frac{k^5}{r})$ unless there's a block error (anywhere in the assembly) or an insufficient attachment happens (within the two blocks) before the structure has a chance to fall off. □

[Chen&Goel2005]

Analysis of Snaked Assembly Growth

Theorem 3. *Assume that we use a $2k \times 2k$ snaked tile system and $G_{mc} = 2G_{se}$. Then for any ϵ, there exists a constant c such that, with probability $1-\epsilon$, no k-bottleneck will happen at a specific location within time $c\frac{1}{f}e^{G_{se}}(\frac{e^{-G_{se}}+1/k^6}{e^{-G_{se}}})^{k-1}$.*

Proof. By definition, k insufficient attachments are required before a k-bottleneck happens. After i insufficient attachments take place, one of the following is going to happen:

- One more insufficient attachment. Consider any structure X caused by i insufficient attachments. By Lemma 3, the size of X cannot exceed two blocks, hence the number of insufficient attachment locations that can cause this structure to grow larger is at most $4k$. So, the rate of the $(i+1)$-th insufficient attachment happening is at most $4kfe^{-G_{se}}$.
- All the attached tiles fall off. By Lemma 3, the expected time for all the attached tiles to fall off is $O(\frac{k^5}{r})$

So, after i insufficient attachments happen, the probability of the $(i+1)$-th insufficient attachment happening before all tiles fall off is $O(\frac{kfe^{-G_{se}}}{kfe^{-G_{se}}+r/k^5}) = O(\frac{e^{-G_{se}}}{e^{-G_{se}}+1/k^6})$. So, after the first insufficient attachment takes place, the probability of a k-bottleneck happening before all the attached tiles fall off is less than $O((\frac{e^{-G_{se}}}{e^{-G_{se}}+1/k^6})^{k-1})$. As shown in Lemma 1, the expected time for the first insufficient attachment is $O(\frac{1}{f}e^{G_{se}})$. So, the expected time for a k-bottleneck to happen at a certain location is at most $O(\frac{1}{f}e^{G_{se}}(\frac{e^{-G_{se}}+1/k^6}{e^{-G_{se}}})^{k-1})$. Hence, for any small ϵ, we can find a constant c such that, with probability $1-\epsilon$, no k-bottleneck will happen at a specific location within time $c\frac{1}{f}e^{G_{se}}(\frac{e^{-G_{se}}+1/k^6}{e^{-G_{se}}})^{k-1}$.

[Chen&Goel2005]

Analysis of Snaked Assembly Growth

Theorem 4. *If we assume there are no k-bottlenecks, and the rate of insufficient attachments is at most $O(\frac{f}{k^6})$, then an $n \times n$ square of $2k \times 2k$ snaked tile blocks can be assembled in expected time $O(\frac{k^5 n}{f})$.*

Proof. With the snaked tile system, after all the tiles in a block attach, all the tiles are held by strength at least 3 and will never fall off. Using the running time analysis technique of Adleman *et al.* [2], the system finishes in expected time $O(n \times T_B)$, where n is the size of the terminal shape and T_B is the expected time for a block to assemble. Without presence of k-bottlenecks, when we want to assemble a block, the erroneous tiles that currently occupy that block are formed by at most $k-1$ insufficient attachments. By Lemma 3, without any further insufficient attachments happening, the erroneous tiles will fall off in time $O(\frac{k^5}{f})$ and the correct block can attach within time $O(\frac{k^4}{f})$. By assumption, the rate of insufficient attachment happening is at most $O(\frac{f}{k^6})$, and there are at most $O(k)$ locations for insufficient attachments to happen and affect this process. So, there's a constant probability that no insufficient attachments will happen during the whole process and thus the time required to assemble a block, T_B, is at most $O(\frac{k^5}{f})$.

[Chen&Goel2005]

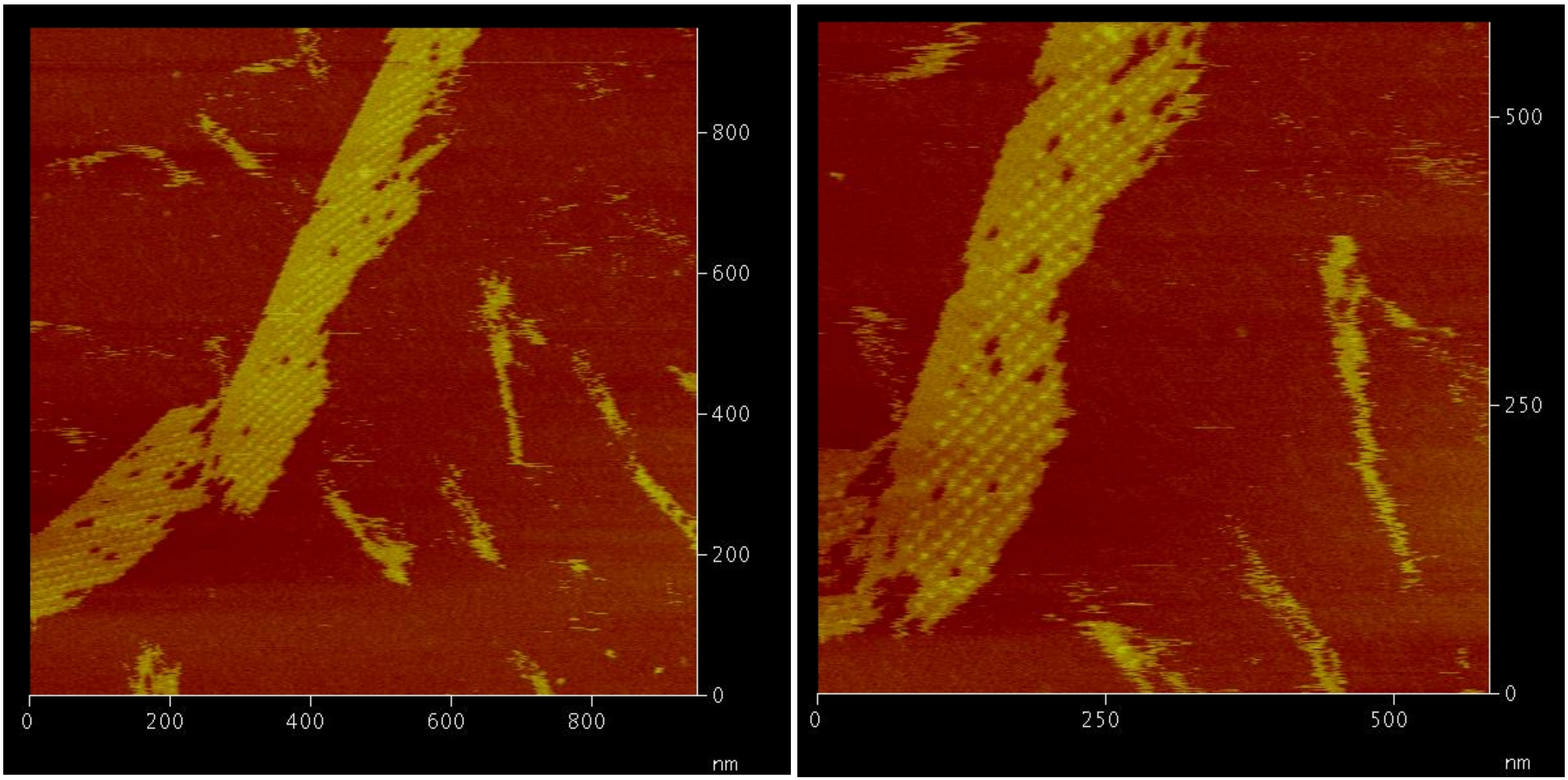
Zig-Zag ribbons tested in experiments

name	description
Normal Zig-Zag (ZZ)	Side A: Glues 2, 4 slow Side B: blunt
Flipped (ZZf)	Side A: Glues 1, 3 fast Side B: blunt
Double_sided (ZZ_DS)	Side A: Glues 2, 4 slow Side B: Glues 1, 3 fast
Flipped + double_sided (ZZ_DSf)	Side A: Glues 1, 3 fast Side B: Glues 2, 4 slow

snake tiles

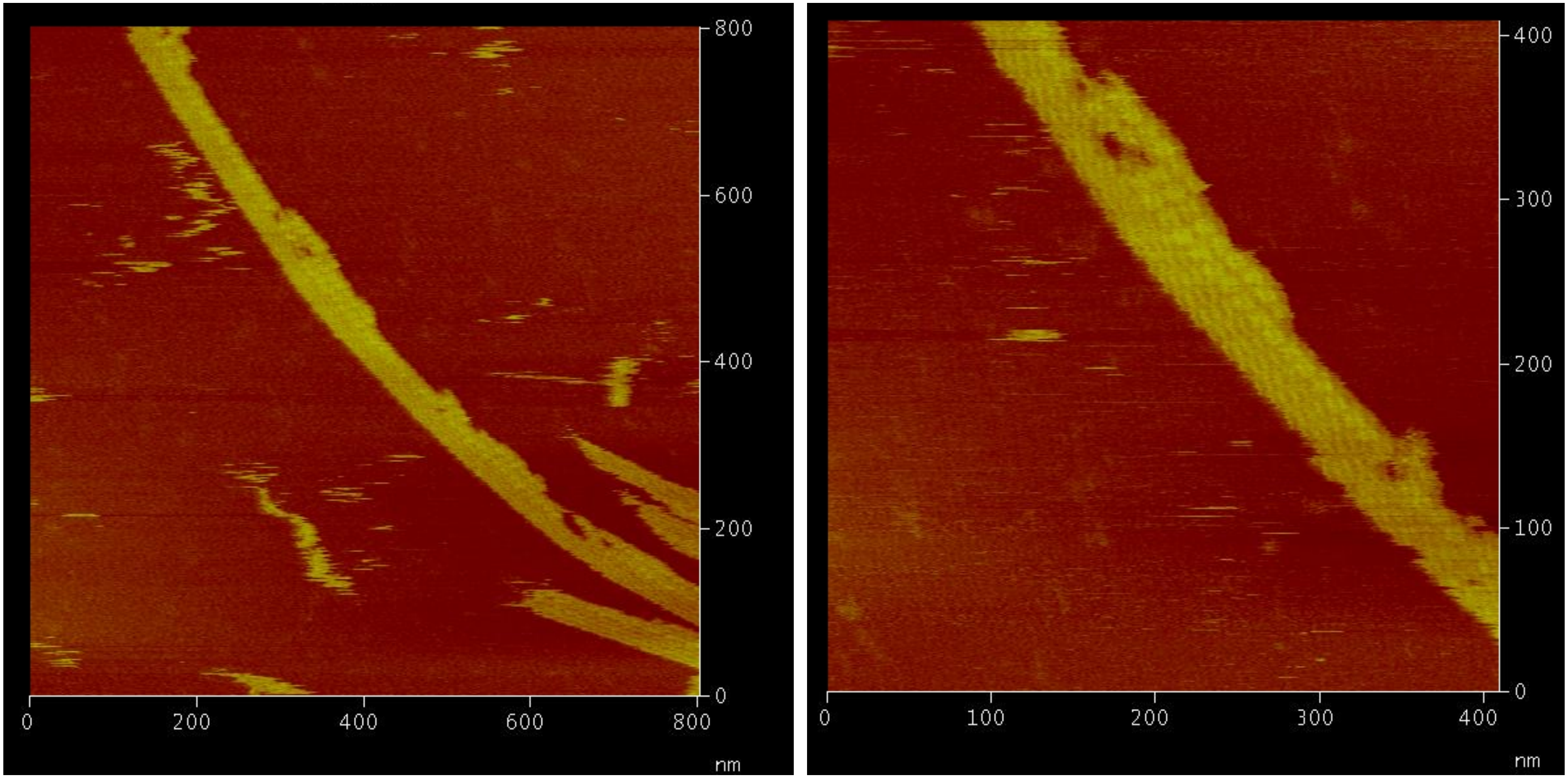
[Cheng,Schulman,Winfree]

AFM Imaging of ZZf + 100 nM Snaked block



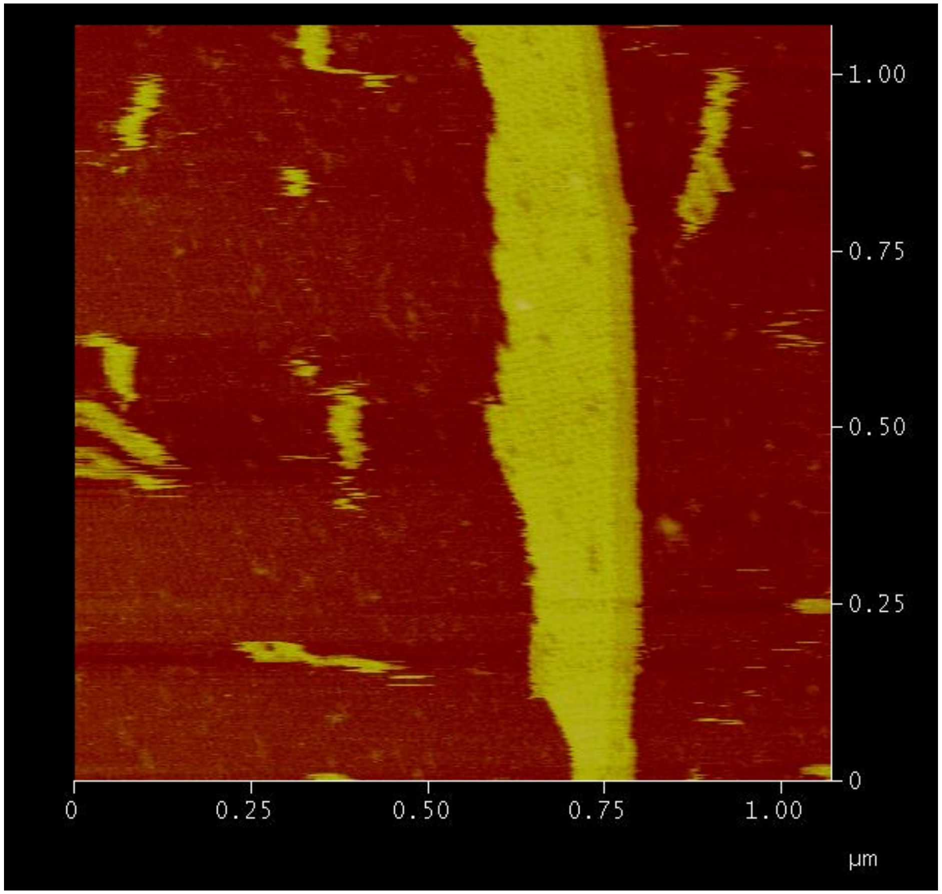
[Cheng,Schulman,Winfree]

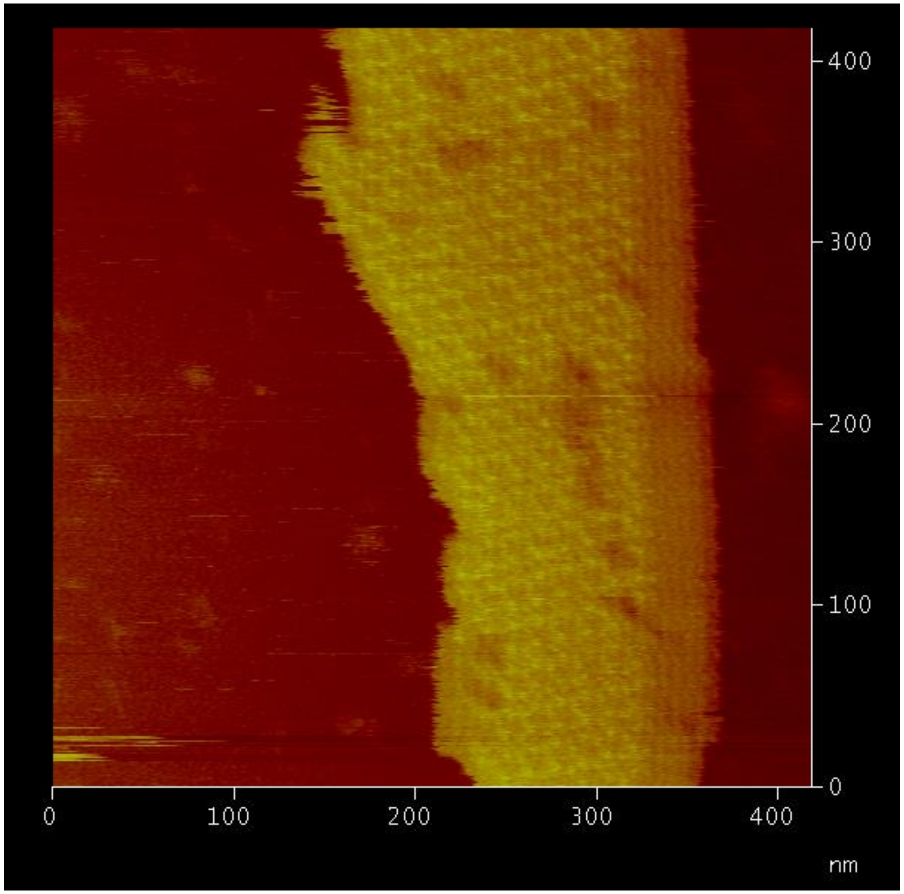
AFM Imaging of Zig-Zag + 100 nM Snaked block



[Cheng,Schulman,Winfree]

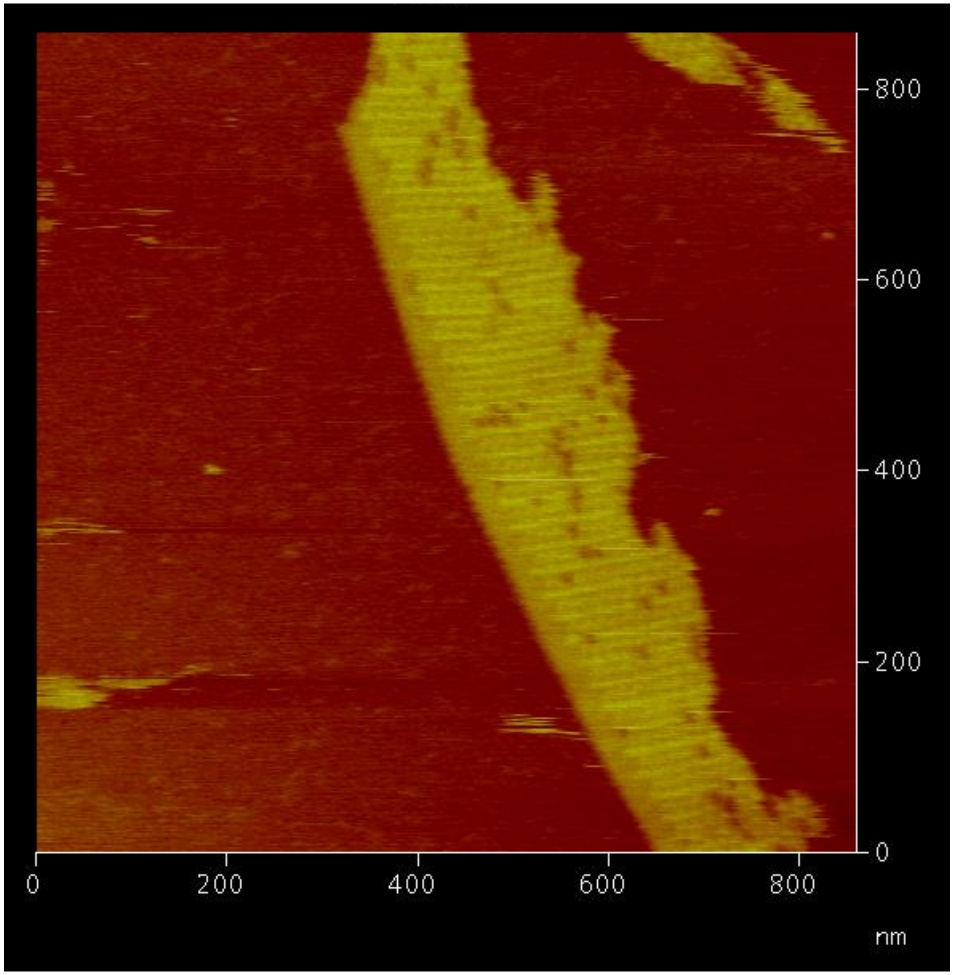
AFM Imaging of Zig-Zag + 100 nM Proofreading

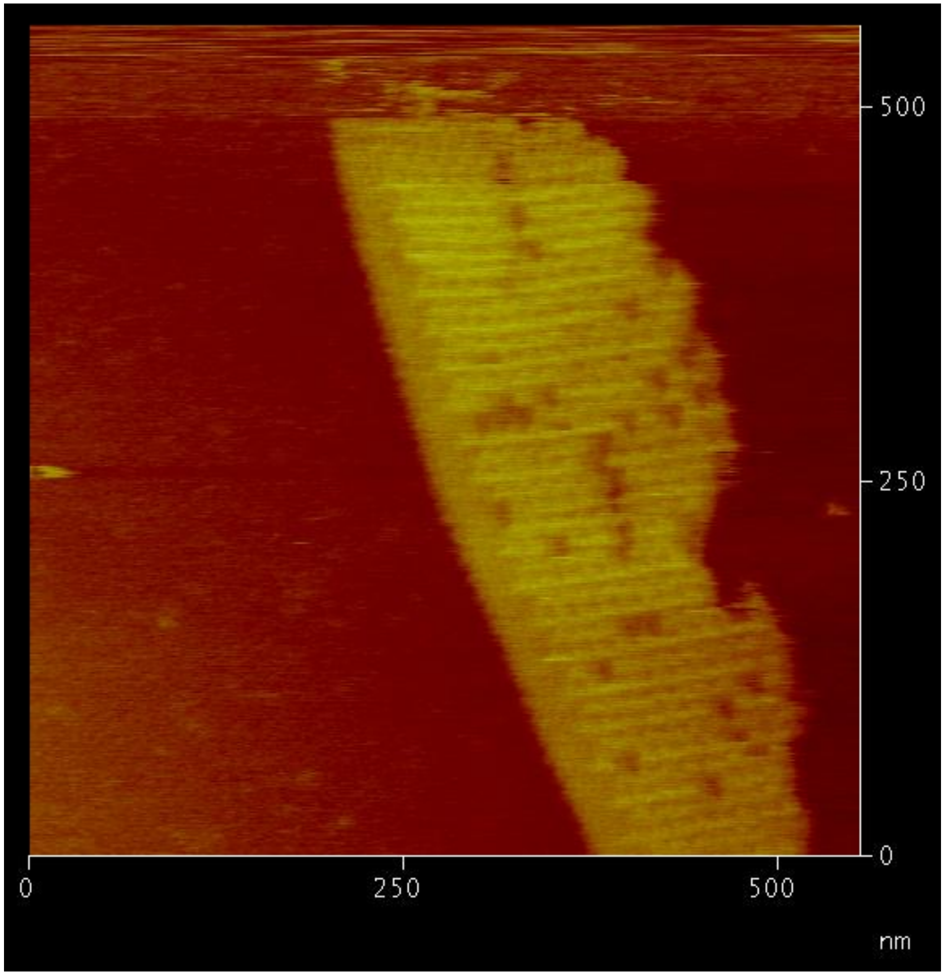




[Cheng,Schulman,Winfree]

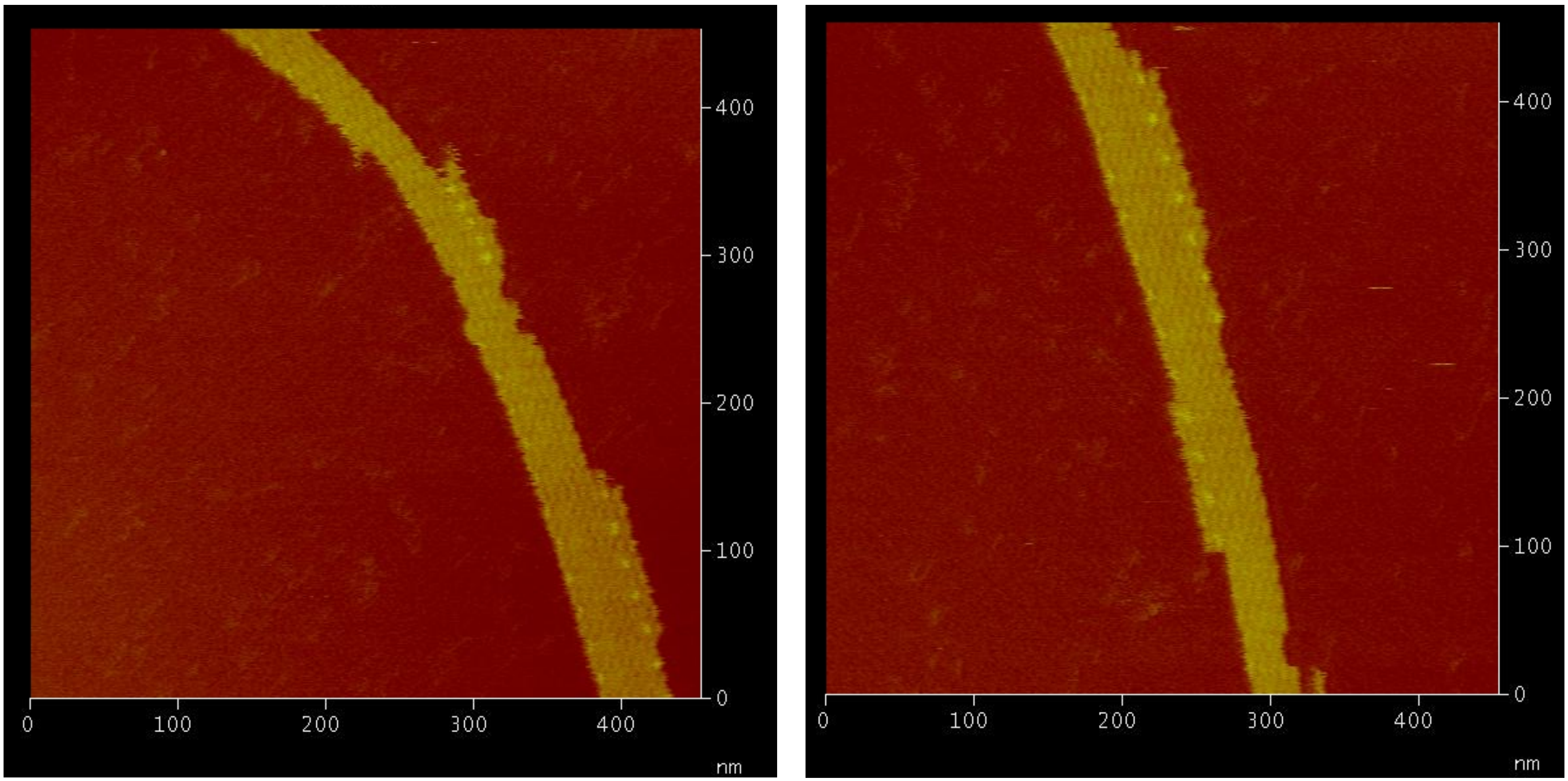
AFM Imaging of
ZZf + 100 nM Proofreading





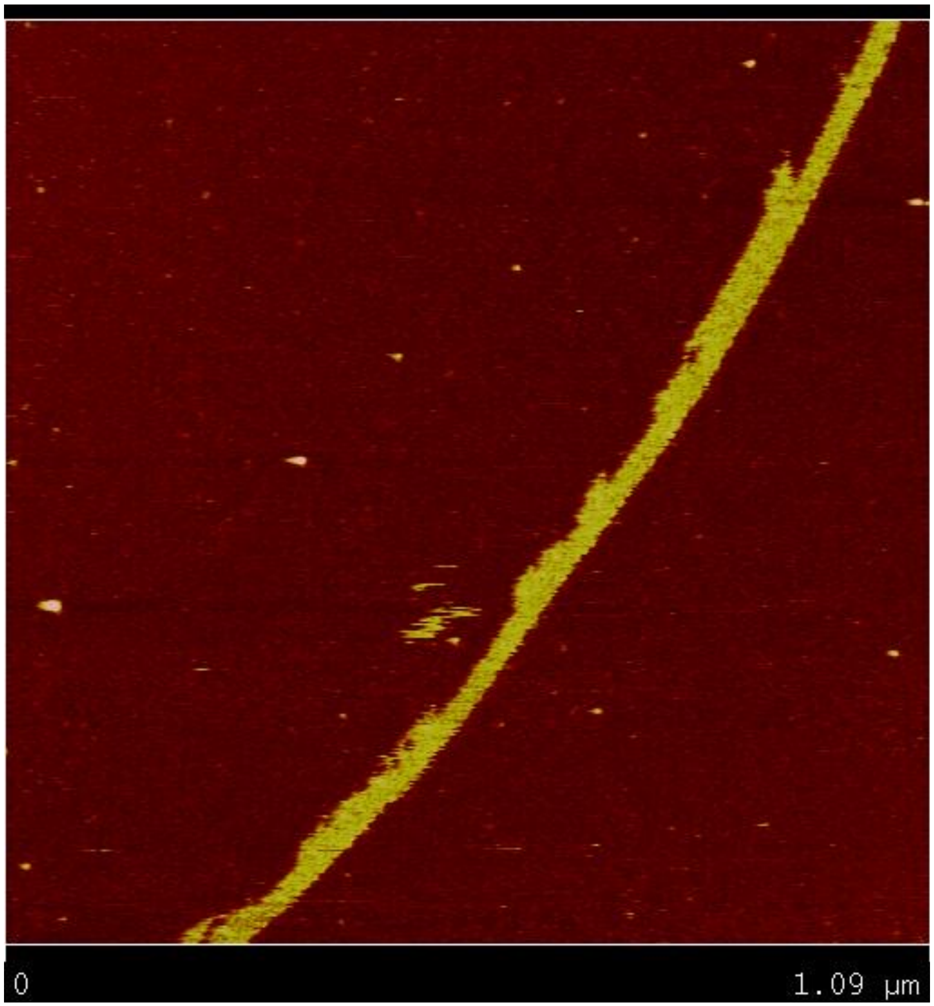
[Cheng,Schulman,Winfree]

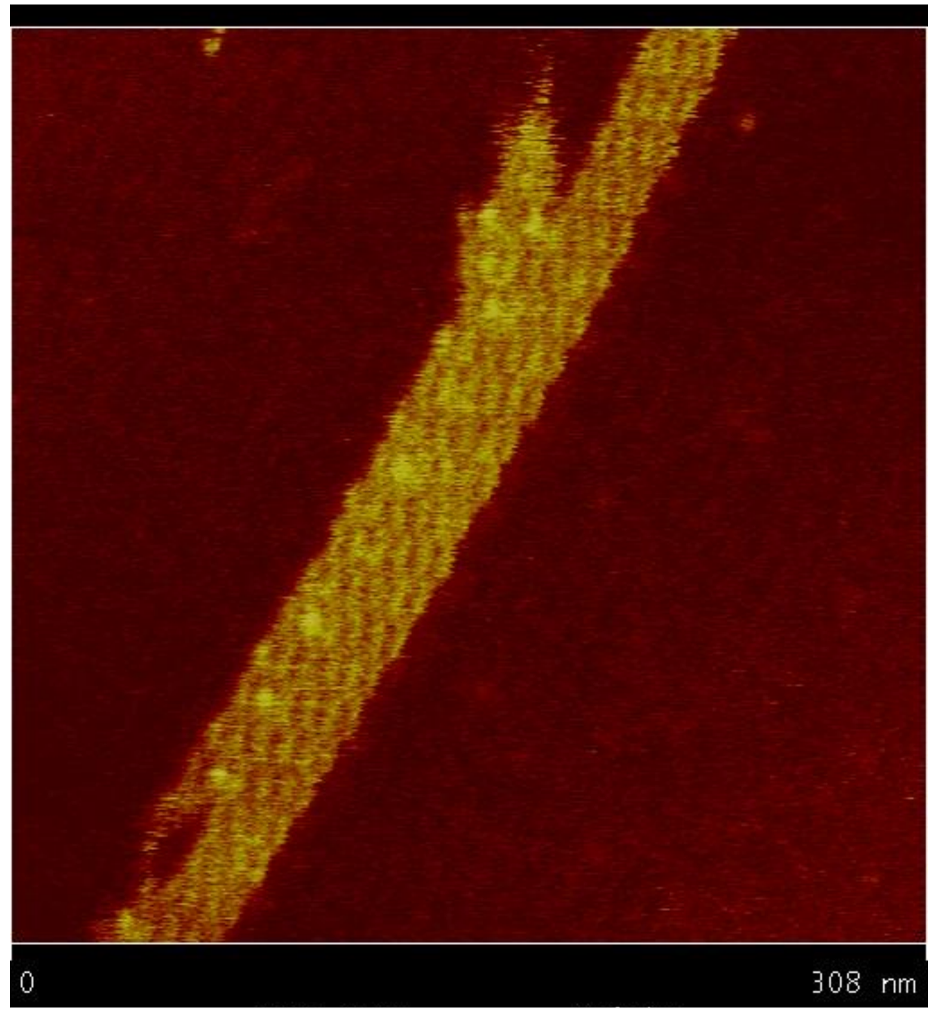
AFM Imaging of
ZZ_DS + 10 nM Proofreading



[Cheng,Schulman,Winfree]

AFM Imaging of ZZ_DS + 10 nM Snaked block





[Cheng,Schulman,Winfree]

Statistics comparing Snaked Blocks with Proofreading Blocks

Ratio of chunks on each side	Zig-Zag Side A: glues 2, 4 Side B: glues 1, 3	Zig-Zag Side A: glues 1, 3 Side B: glues 2, 4
Snaked block	4.7	4.2
Proofreading block	1.1	1.5

[Cheng,Schulman,Winfree]

Statistics comparing Snaked Blocks with Proofreading Blocks

Ratio of tiles on each side	Zig-Zag Side A: glues 2, 4 Side B: glues 1, 3	Zig-Zag Side A: glues 1, 3 Side B: glues 2, 4
Snaked block	4.3	3.9
Proofreading block	1.0	1.2

[Cheng,Schulman,Winfree]

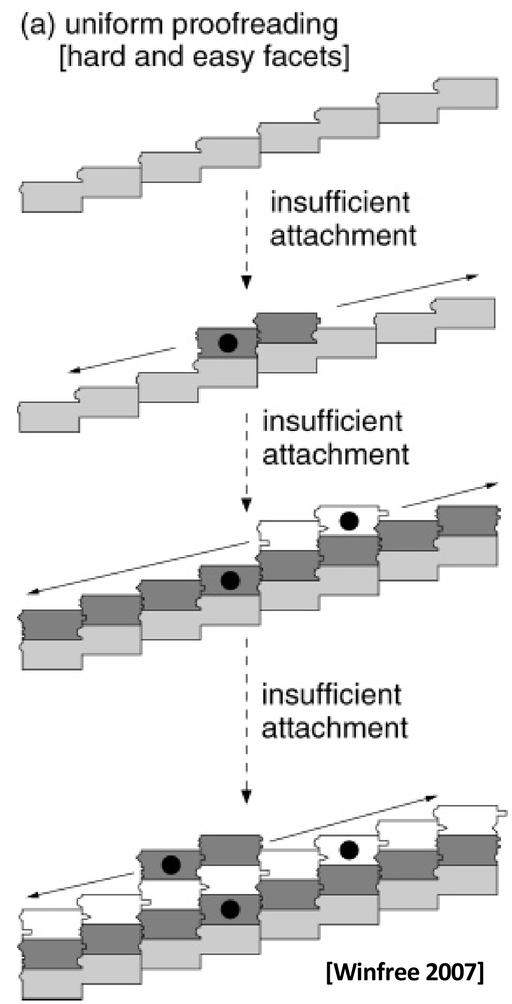
Reducing Facet Nucleation during Algorithmic Self-Assembly

Ho-Lin Chen,Rebecca Schulman,Ashish Goel,and Erik Winfree

NANOLETTERS 2007

[Chen,Schulman,Goel,Winfree2007]

Errors of Algorithmic self-assembly:
- Insufficient attachments
- Facet Nucleation Errors



-Insufficient attachments on assembly facets that involve no mismatches

- The added tiles may be incorrect for the pattern.

In standard crystal growth:

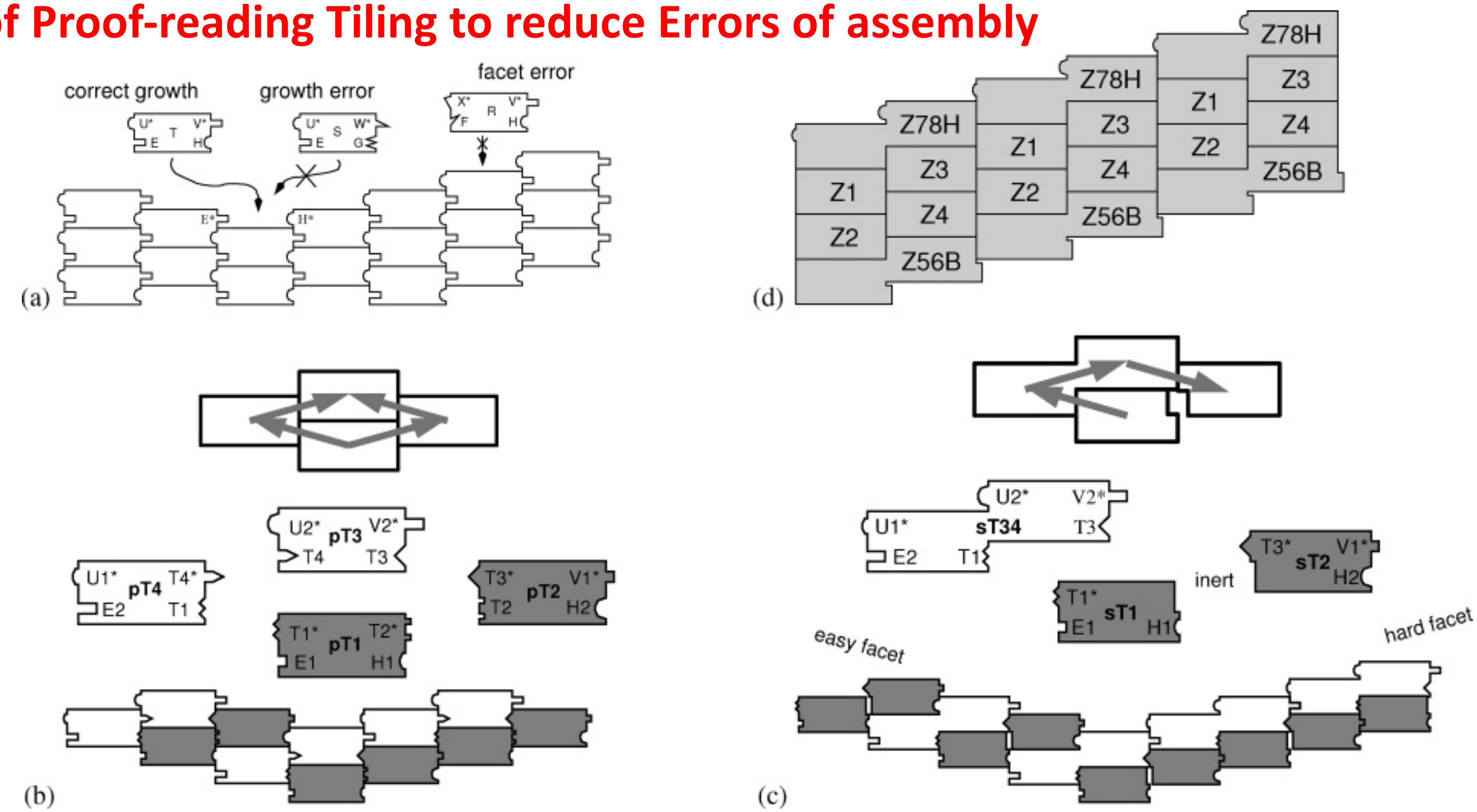
- nucleation on facets is part of the desired growth process,

In a proper algorithmic self-assembly:

- every tile attaches by two or more binding sites, so nucleation on facets may cause errors of assembly

[Chen,Schulman,Goel,Winfree2007]

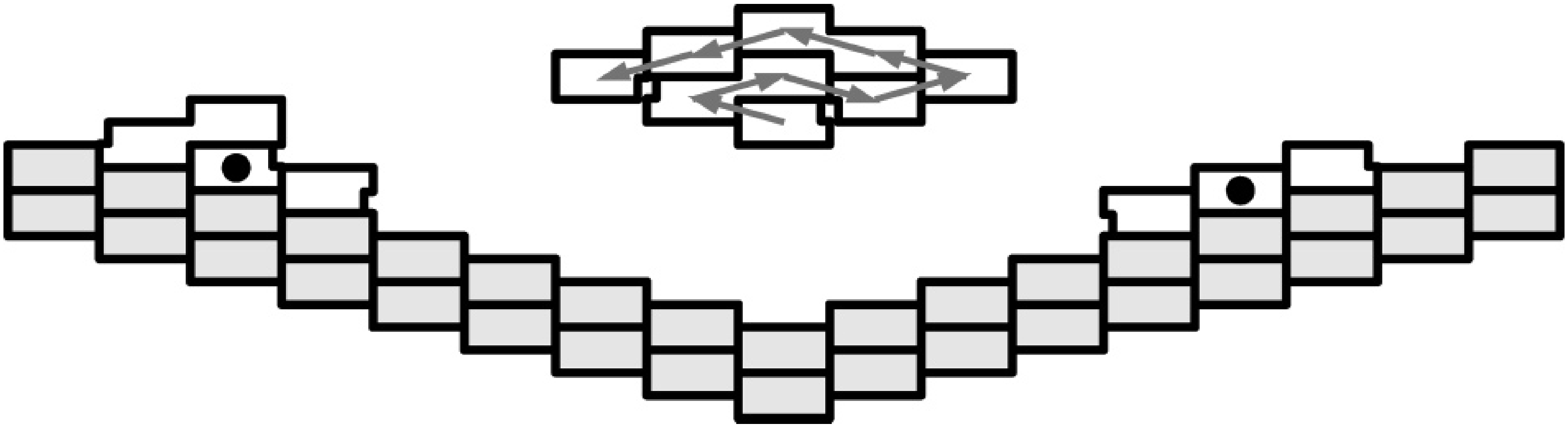
Use of Proof-reading Tiling to reduce Errors of assembly



Algorithmic self-assembly and proofreading blocks. (a) During algorithmic self-assembly, a tile attaches to a growing crystal by binding domains on its edges. Here, the four labels on a tile's corners indicate specific binding domains; asterisk indicates complementary domains (X binds to X*). The attachment of a tile where both its input (bottom) edges match the available edges on the crystal is preferred over the attachment of a tile where a single (or no) match occurs. Growth errors occur when a tile attaches by one matching bond and one nonmatching bond. Facet errors occur when a tile attaches by only one matching bond. In both cases, for an error to occur, the incorrect tile must be "locked in" by a second tile before it detaches. (b) The logical structure of a 2 × 2 uniform proofreading block. Each tile in the original tile set is converted into four tiles that, as a logical block, redundantly encode the same input and output information on the perimeter of the block, while binding domains on the interior of the block encode the identity of the original tile. Correct assembly at a growth site proceeds one tile at a time, either in the order pT1−pT2−pT4−pT3 or in the order pT1−pT4−pT2−pT3. The unique labels inside a proofreading block reduce the rate of growth errors because for a block to be completed, one of the tiles that attaches on top of an incorrect tile must also be incorrect—it cannot match both the label inside the block and the label presented by the crystal. (c) The structure of a 2 × 2 snaked proofreading block. Binding labels on the perimeter of the block are the same as in a uniform proofreading block, but the interior has two modifications: there is an inert interaction between sT1 and sT2, and the other two tiles are fused to create the "double tile" sT34. This forces correct assembly to proceed in the order sT1−sT34−sT2. Snaked proofreading tile sets, like uniform proofreading tile sets, force a subsequent tile that attaches after an incorrect attachment to be incorrect also. (d) Zigzag ribbons. While only three repeat units are shown, ribbons can be arbitrarily long. The 6 tiles interact through 12 distinct pairwise binding domains, all shown as flat sides, as their logic is not essential here. In addition to the double tiles shown, we use variants of the double tiles that present binding domains to create a desired facet (e.g., Z78H presents H1* and H2* for the hard facet) or present inert "blunt ends" (e.g., Z56B) to which nothing may bind.

[Chen,Schulman,Goel,Winfree2007]

Use of Snakes-tiling Blocks to reduce Errors of facet growth



3 × 3 snaked blocks reduce facet growth on both facet types. On either facet, an isolated insufficient attachment (initiated by the tiles marked with dots) can grow by favorable attachment to a maximal size of three tiles, at which point it is no longer possible to attach a tile by two binding sites. However, at a proper growth site, the series of exclusively favorable assembly steps following the snaked path shown can complete the block quickly.

[Chen,Schulman,Goel,Winfree2007]

Zigzag Ribbons of Proof-Reading Tiles have to deal with various Facet Types:

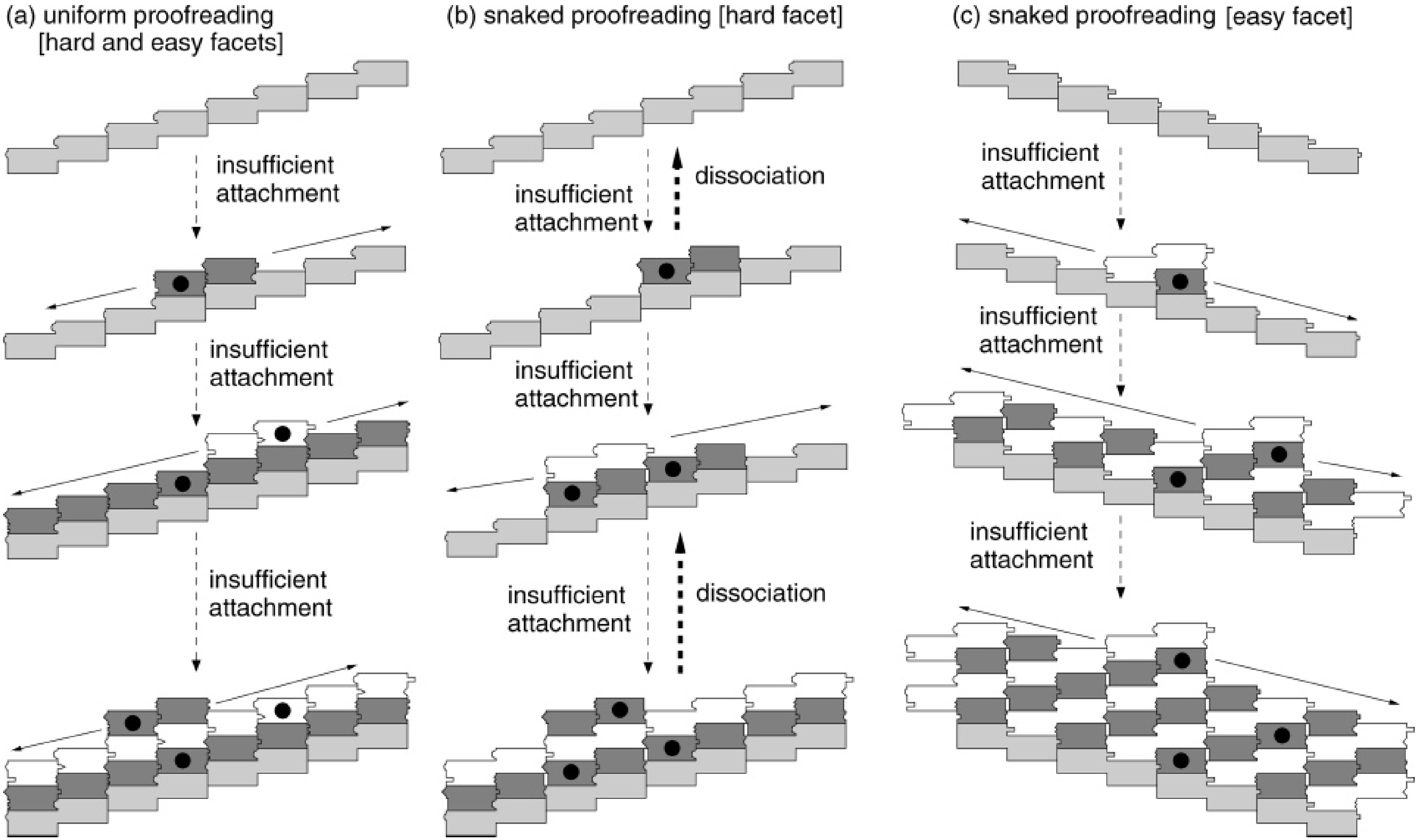
Easy: where snaked proofreading tiles can nucleate growth with just one insufficient attachment

Hard: where two adjacent insufficient attachments are required for snaked proofreading tiles to nucleate facet growth

Blunt: which contain no binding sites and therefore do not allow growth.

[Chen,Schulman,Goel,Winfree2007]

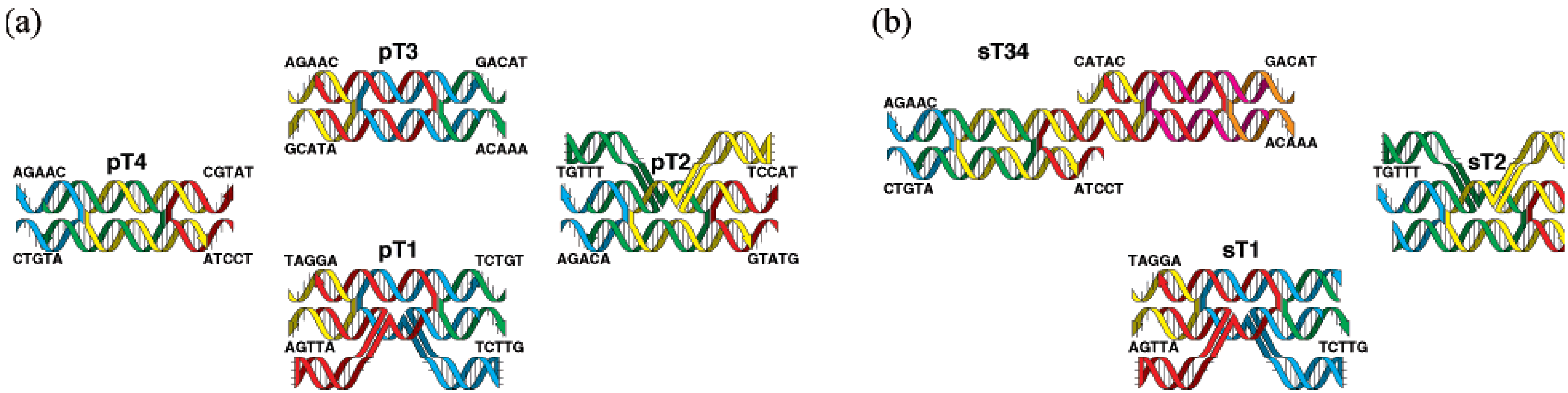
Tiling Errors: Facet Nucleation and growth



Facet nucleation and growth. (a) Facet nucleation of uniform proofreading blocks. Following a single insufficient attachment (tiles with dots indicate the unfavorable attachments that were locked in), subsequent growth by favorable attachments can grow an entire layer of tiles. Subsequent rows are each nucleated by a single insufficient attachment event. (b) Facet nucleation for snaked proofreading blocks along the hard facet. Here, a single insufficient attachment results in a pair of tiles on the facet, but further favorable growth steps are impossible because of the inert bonds interior to the snaked blocks. Two adjacent insufficient attachments are necessary to nucleate two layers of facet growth. Each additional two layers of growth requires another two adjacent insufficient attachments. (c) Facet nucleation for snaked proofreading blocks along the easy facet. Here, an insufficient attachment consists of a single tile and an adjacent double tile. Thus, two layers of snaked proofreading tiles can be nucleated by just one insufficient attachment.

[Chen,Schulman,Goel,Winfree2007]

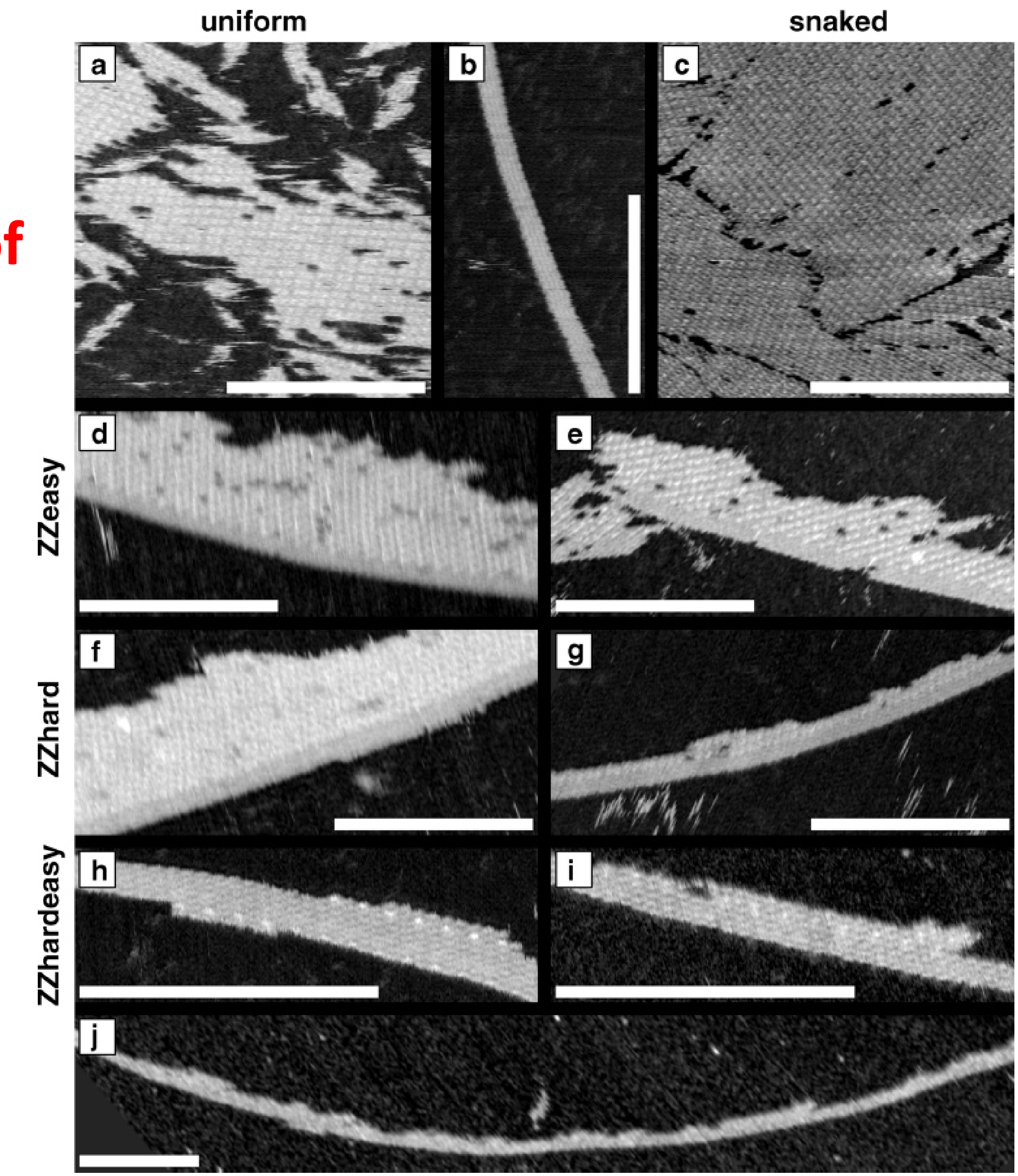
DNA implementation of Snaked Proofreading Blocks



DNA implementation of uniform and snaked proofreading blocks. (a) DNA tiles for a uniform proofreading block. Each tile shown is a double-crossover molecule known as the DAO-E molecule.[25] A DAO-E molecule is composed of four strands of DNA. While the "core" of the molecule is double stranded, there are four five-nucleotide single-stranded regions (sticky ends) on each molecule that can bind to complementary sticky ends on other tiles. Two hairpins are present on each of the shaded tiles in Figure 1b to provide AFM contrast. (b) DNA tiles for a 2 × 2 snaked proofreading block, which in the rest of the paper will be referred to simply as a "snaked proofreading block". To make an inert bond between sT1 and sT2, the sticky ends of the tiles are double stranded and truncated by two and three bases, respectively. The double tile sT34 is implemented by a larger molecule which has the structure of two DAO-E tiles fused together. Hairpins are used on tiles shaded in Figure 1c.

[Chen,Schulman,Goel,Winfree2007]

AFM Images of DNA implementation of Snaked Proofreading Blocks



AFM images. Missing tiles are due to damage during AFM scanning. Scale bars are 300 nm. (a) Uniform proofreading lattices (100 nM). (b) Zigzag ribbons, ZZhardeasy (10 nM). (c) Snaked proofreading lattices (100 nM). (d) ZZeasy (10 nM) with uniform proofreading (50 nM). (e) ZZeasy (10 nM) with snaked proofreading (50 nM). (f) ZZhard (10 nM) with uniform proofreading (50 nM). (g) ZZhard (10 nM) with snaked proofreading (50 nM). (h) ZZhardeasy (10 nM) with uniform proofreading (10 nM). (i) ZZhardeasy (10 nM) with snaked proofreading (10 nM). (j) A long ZZhardeasy ribbon with snaked proofreading.

[Chen,Schulman,Goel,Winfree2007]

Self-healing Tile Systems

[Winfree,2005]

- Goal: When a big portion of the lattice is removed, it should be able to grow back correctly.
- Method: For each tile in the original system, we create a unique block in the new system.
- Idea: Use the block to prevent tile from growing backwards.

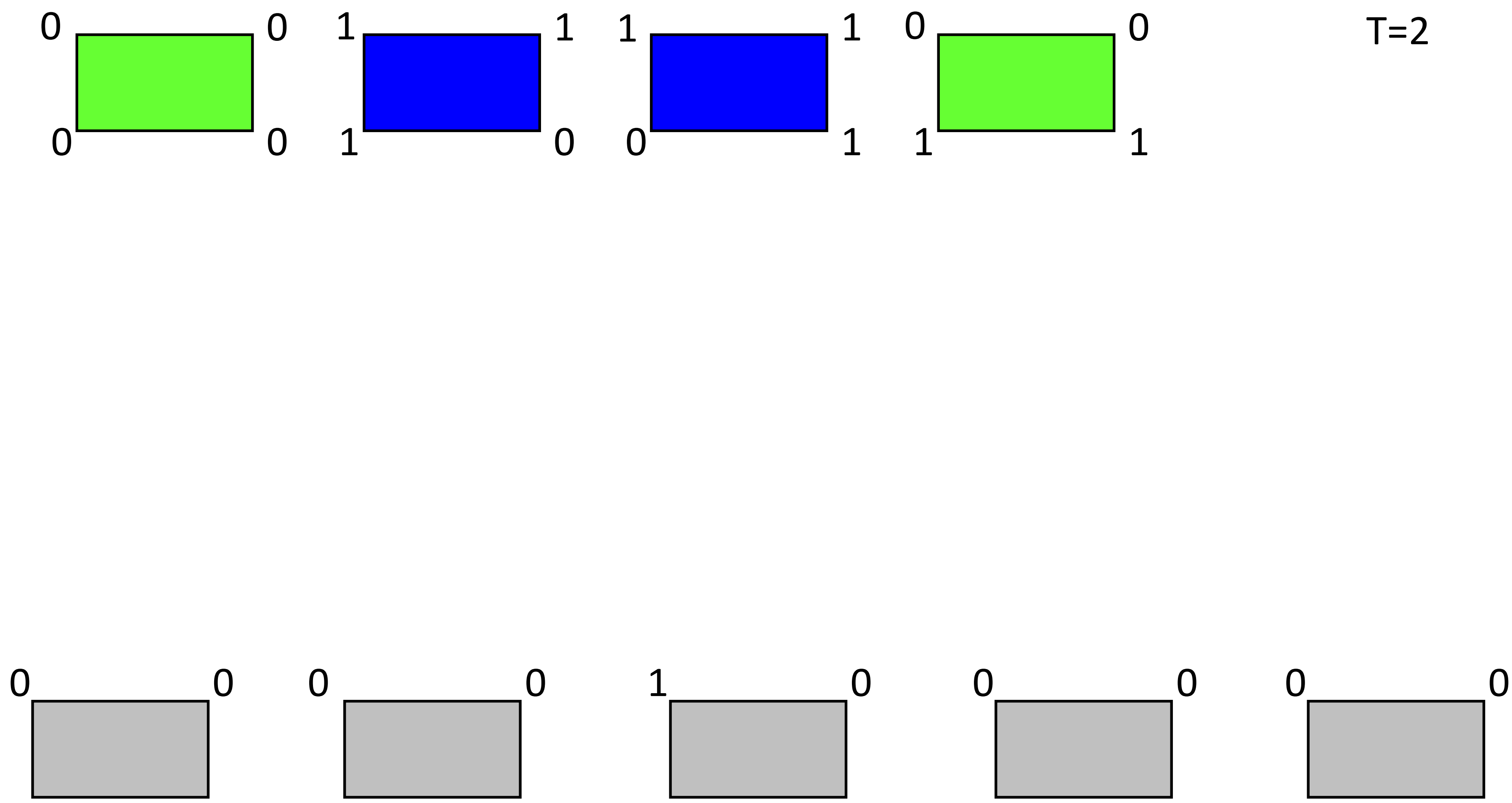
[Winfree,2005]

Assumptions

- **Use abstract tile assembly model (TAM).**
- **Requires a fix set of incoming and outgoing edges for each tile in the original system.**

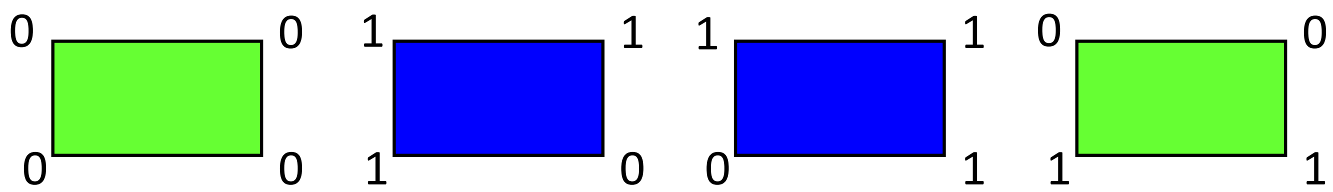
[Winfree,2005]

Example: Sierpinski System

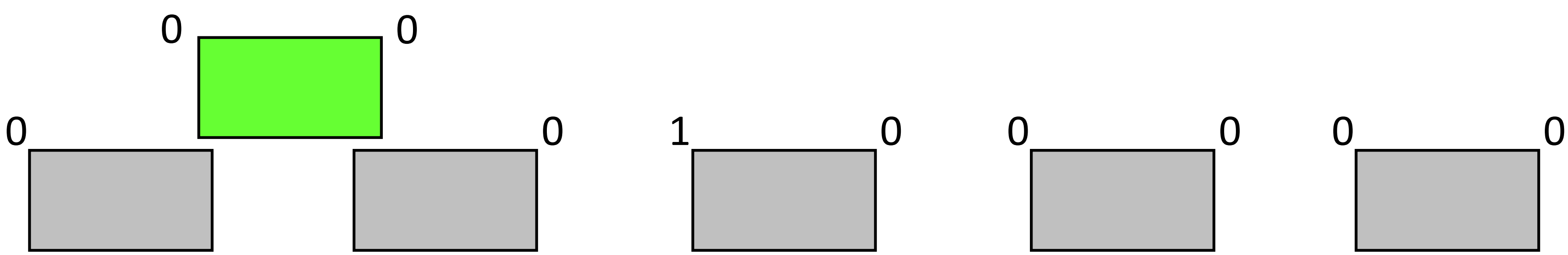


[Winfree,2005]

Example: Sierpinski System

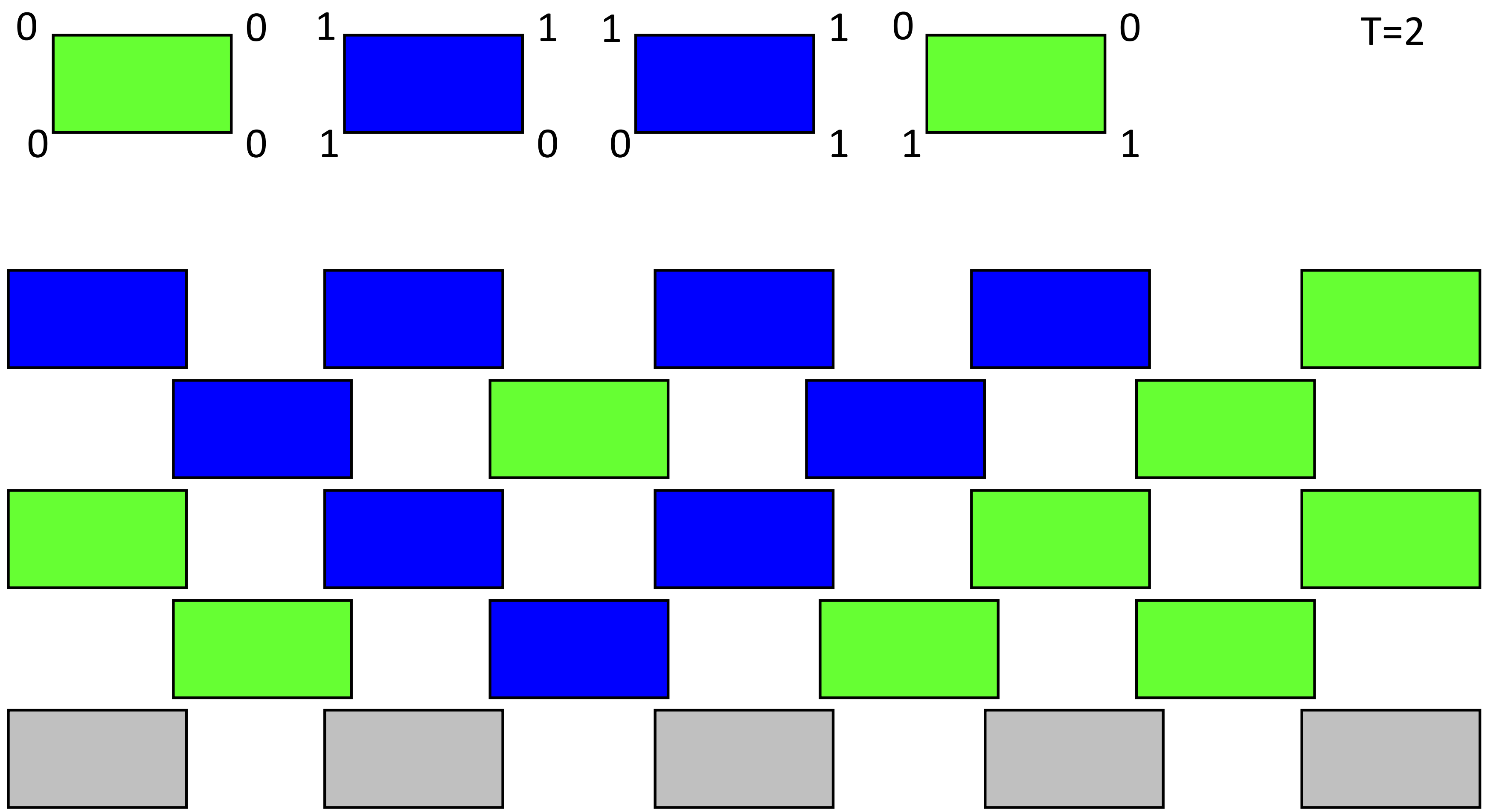


T=2



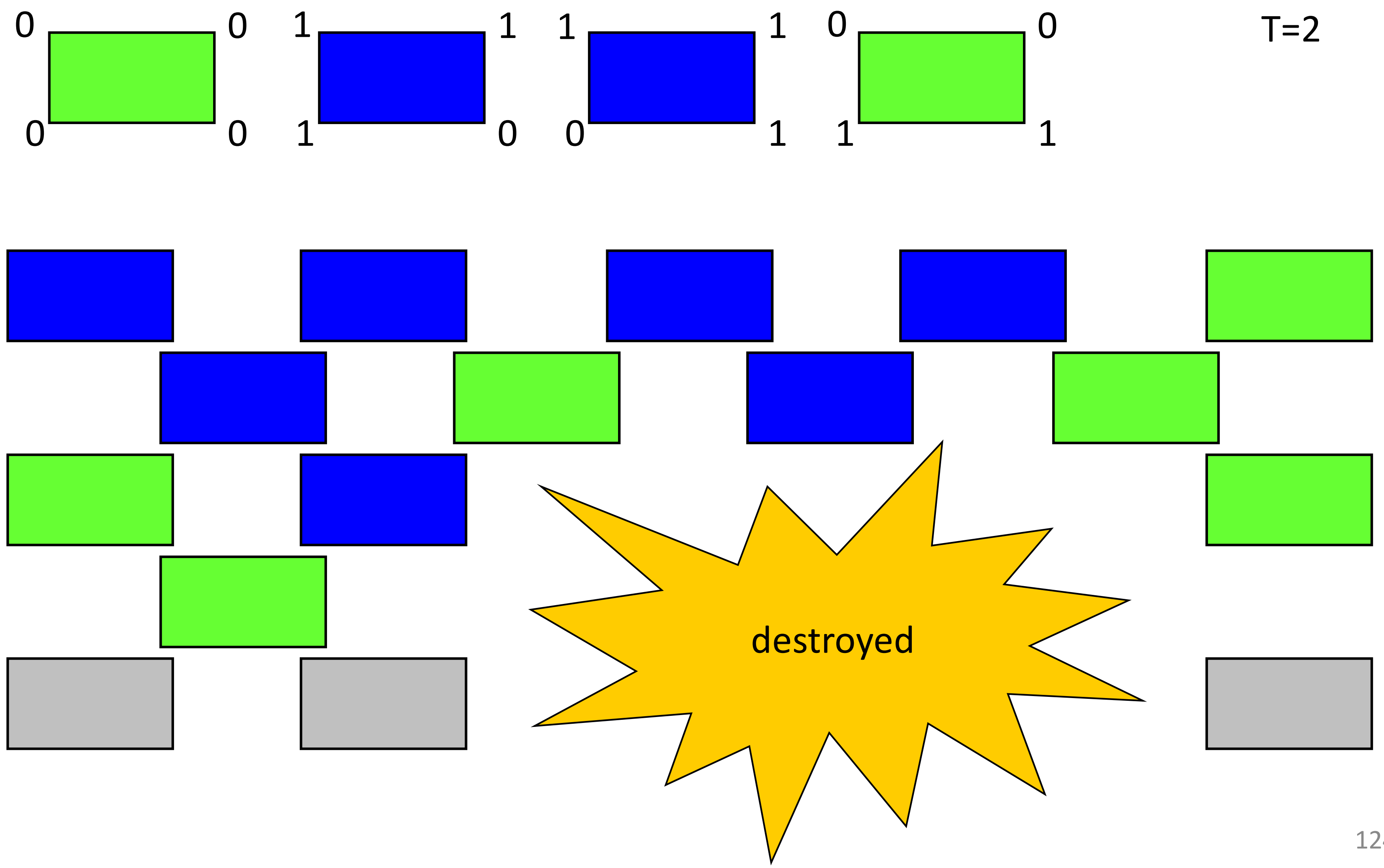
[Winfree,2005]

Example: Sierpinski System



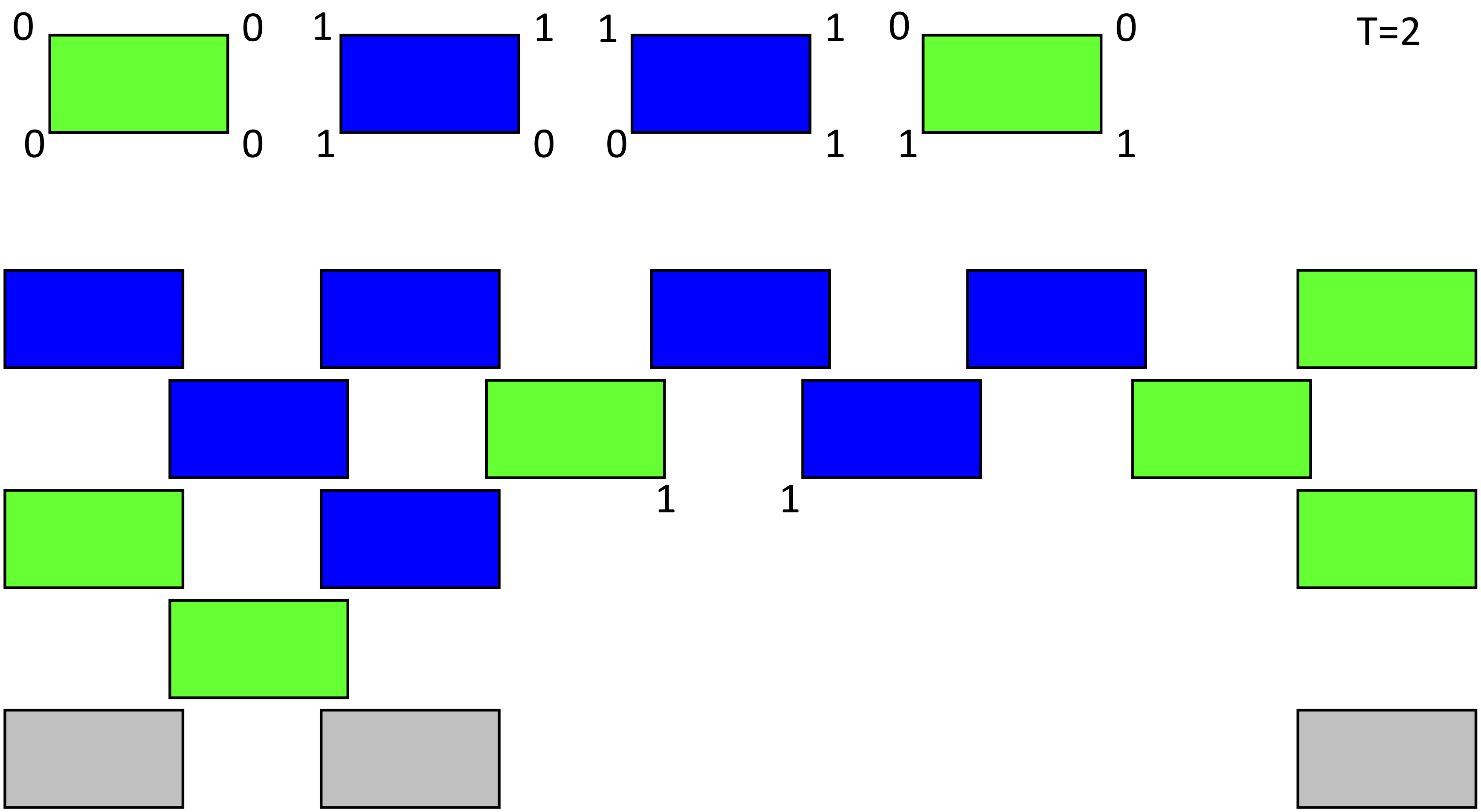
[Winfree,2005]

Example: Sierpinski System



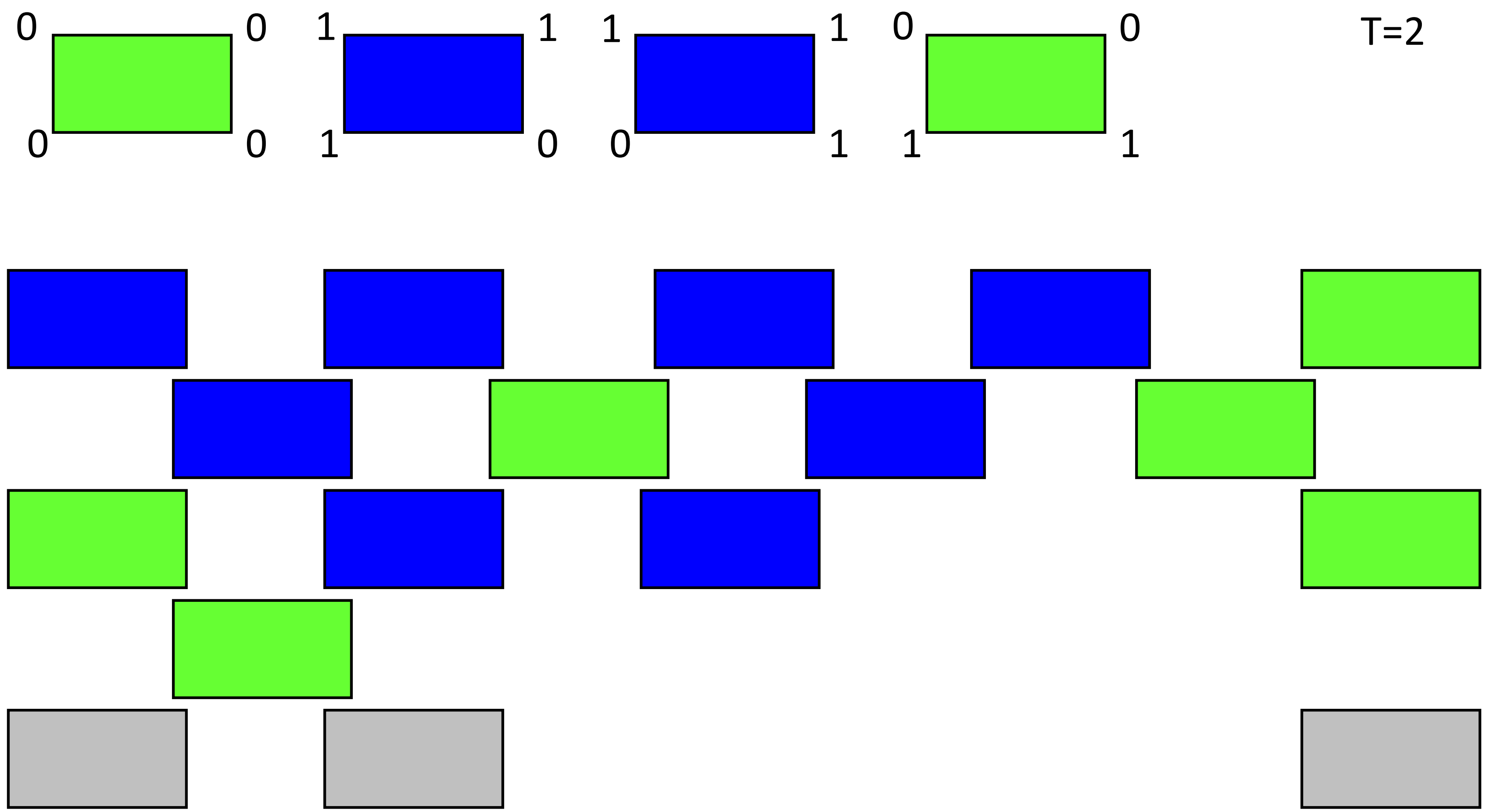
(Chen)

Example: Sierpinski System



(Chen)

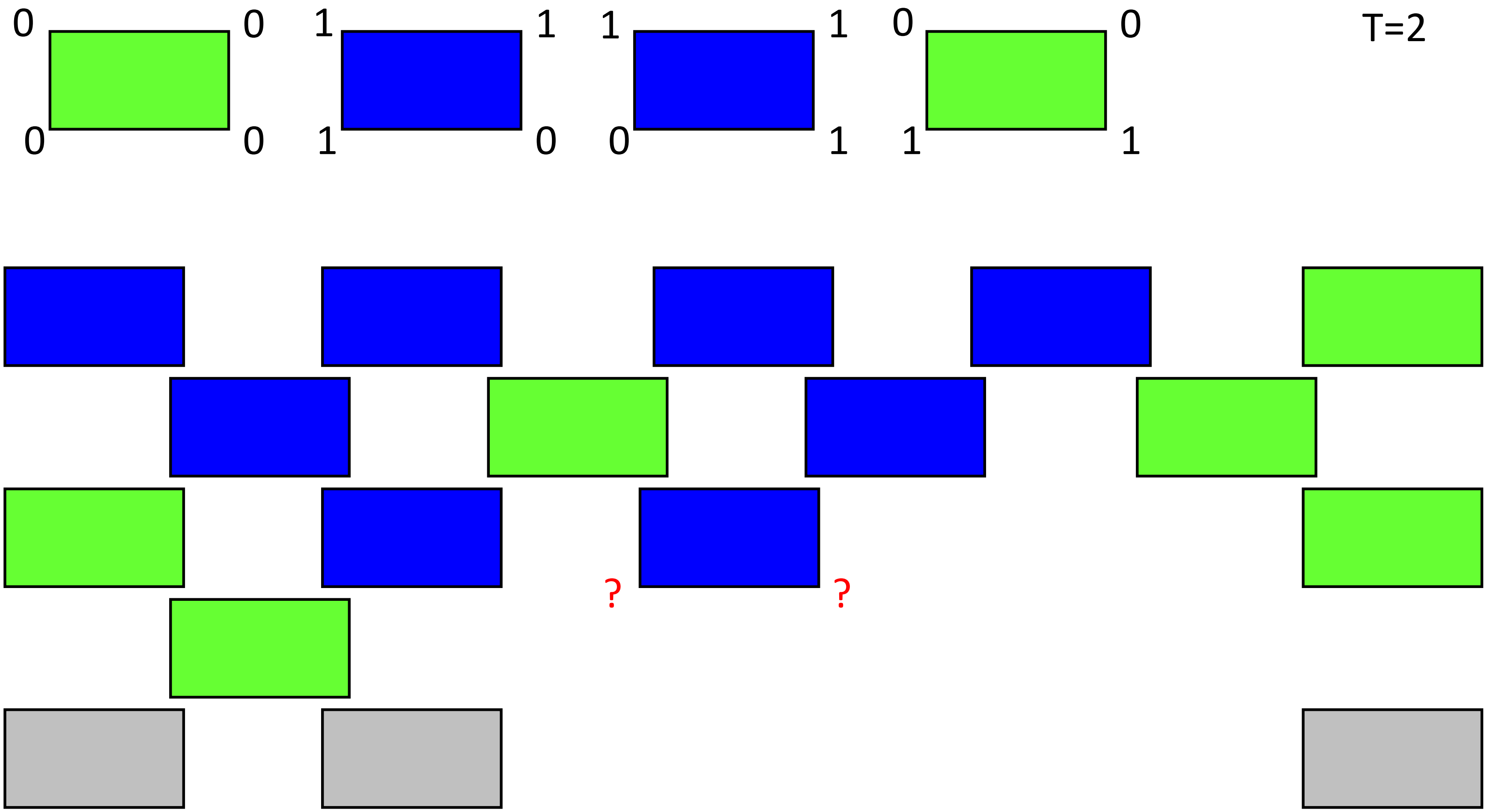
Example: Sierpinski System



[Winfree,2005]

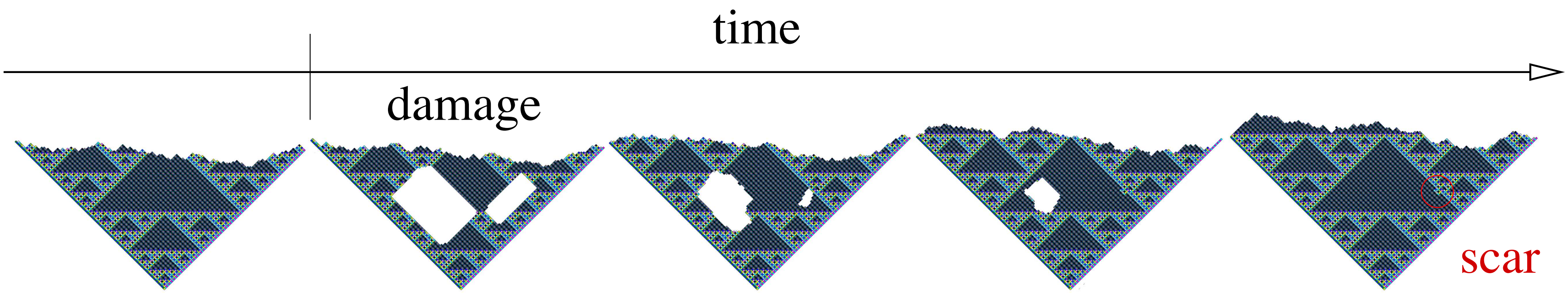
Example: Sierpinski System

Note: the inputs of Sierpinski tiles are not reversible.



[Winfree,2005]

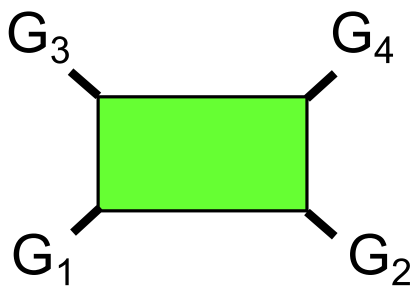
Example: Using Proofreading Tiles sets in Sierpinski System to Heal



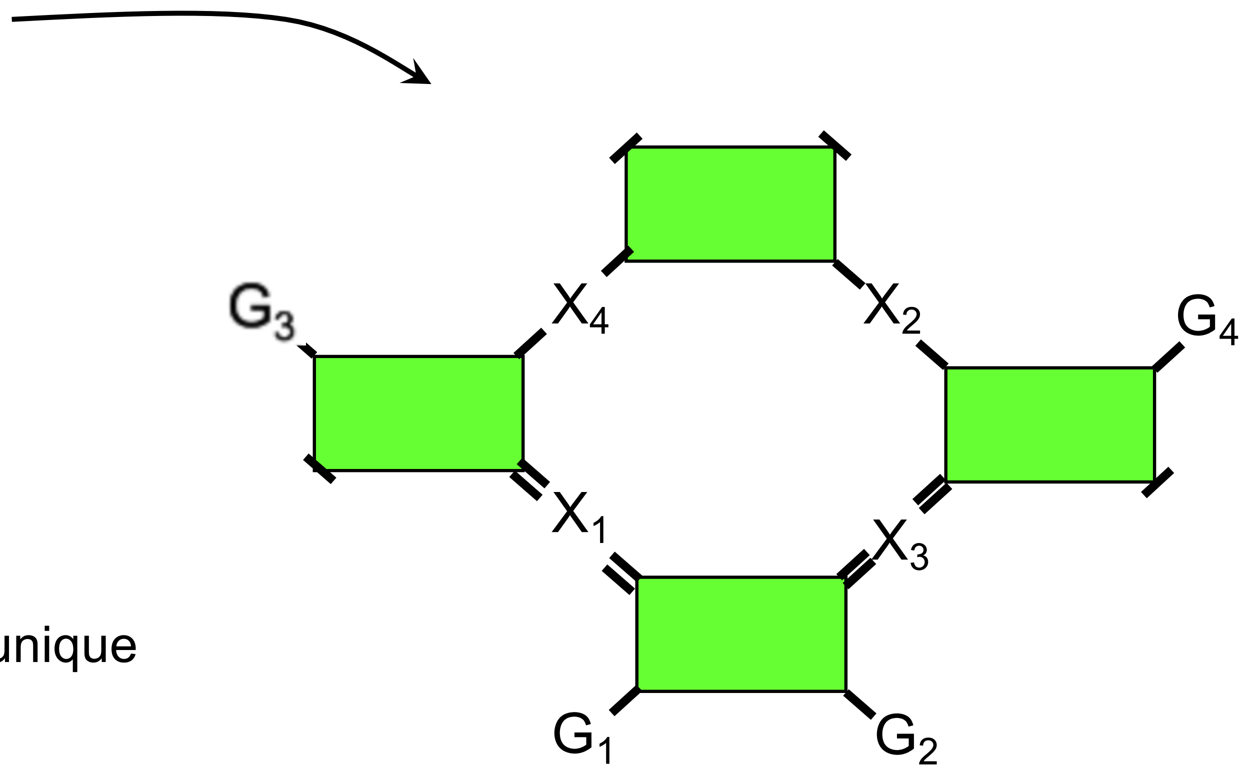
. Proofreading tile sets are often able to heal a puncture in the crystal. Sometimes, as in this case, some of the tiles that fill in the puncture do not perfectly match their neighbors – a form of “scar tissue.”

[Winfree & Bekbolatov2003]

Using Blocks of Tiles to Promote Healing (T=2)



Original Tile



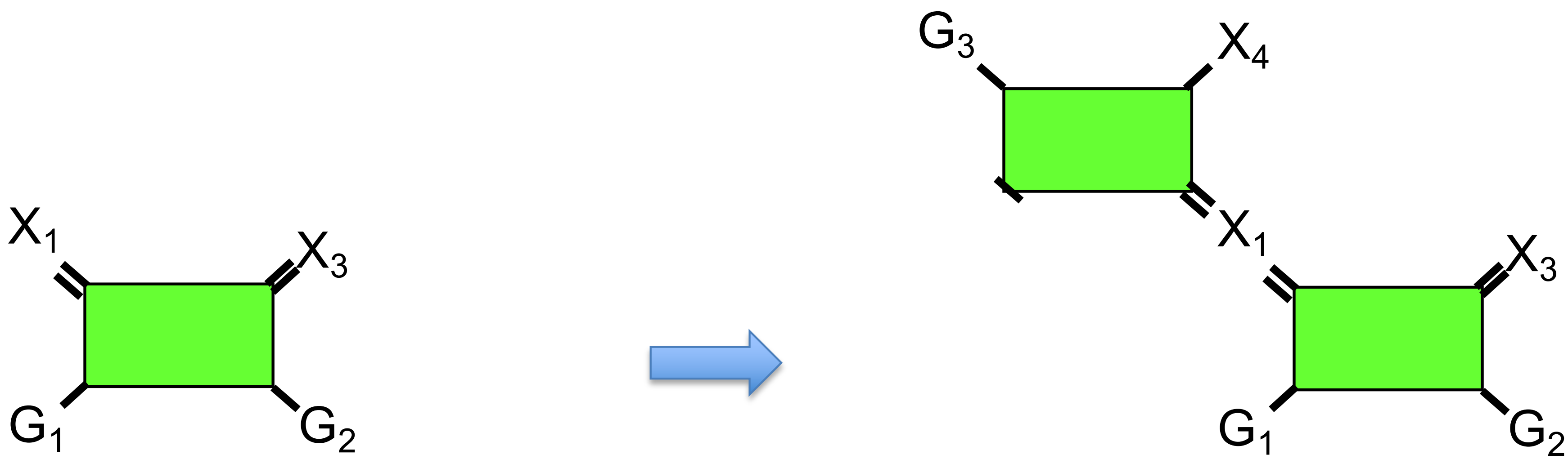
- Replace a tile by a block of 4 tiles
- Internal glues are unique

Block of Tiles

(Chen)

Using Blocks of Tiles to Promote Healing (T=2)

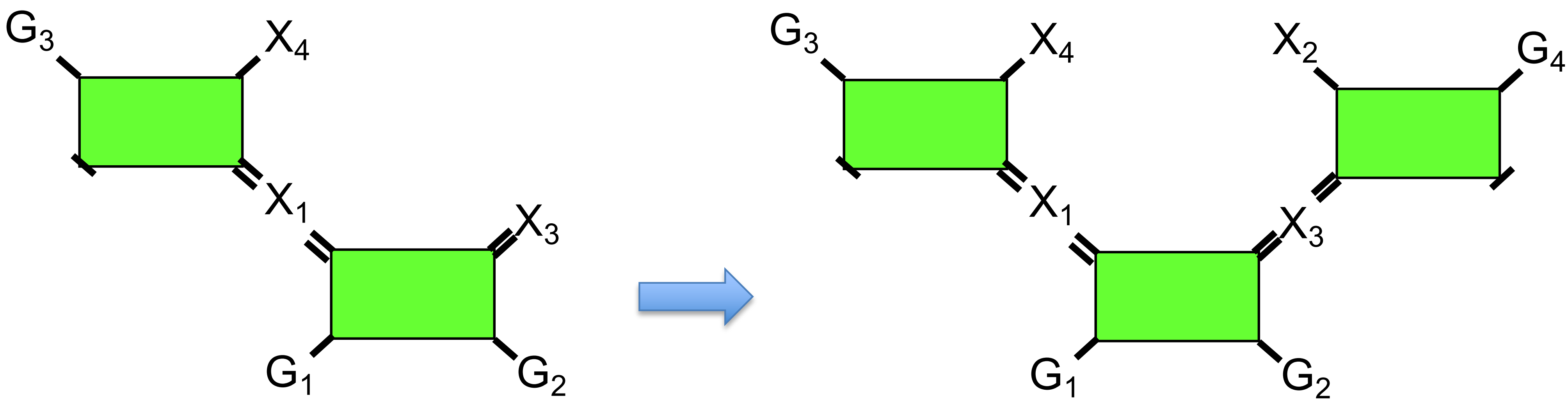
- The system is safe even when several tiles can form a bigger block before attaching to the assembly.



(Chen)

Using Blocks of Tiles to Promote Healing (T=2)

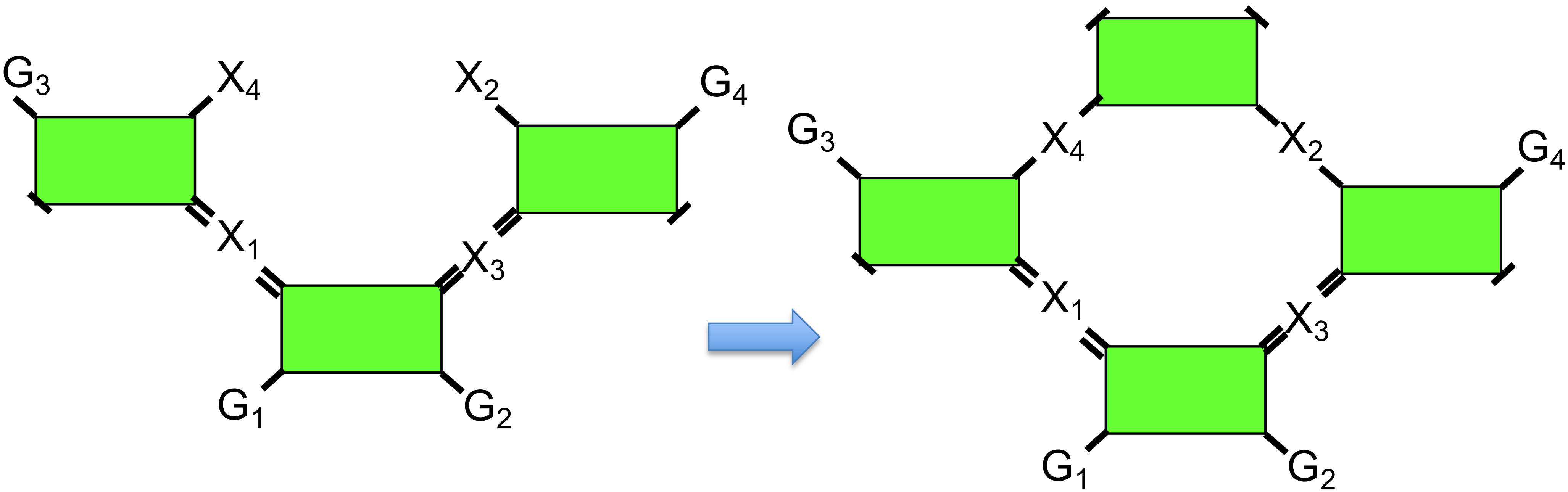
- The system is safe even when several tiles can form a bigger block before attaching to the assembly.



(Chen)

Using Blocks of Tiles to Promote Healing (T=2)

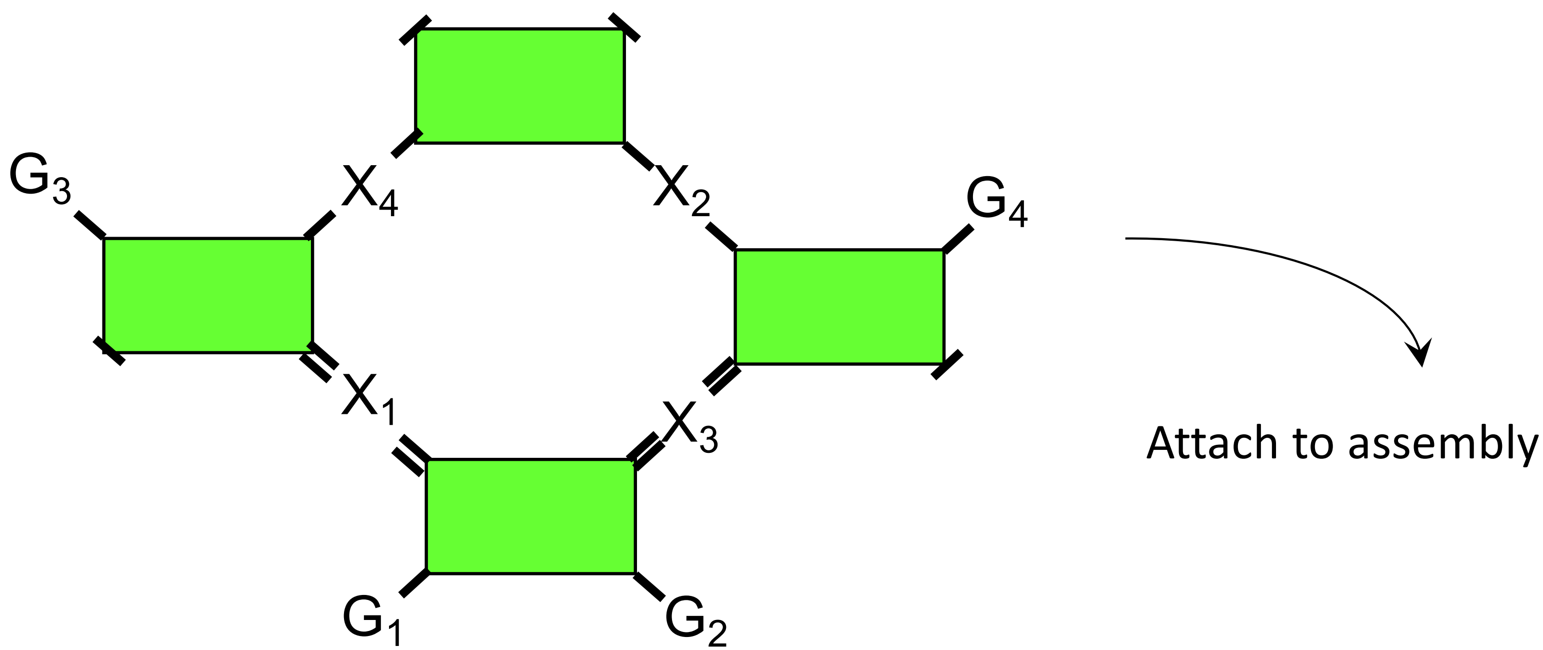
- The system is safe even when several tiles can form a bigger block before attaching to the assembly.



(Chen)

Using Blocks of Tiles to Promote Healing (T=2)

- The system is safe even when several tiles can form a bigger block before attaching to the assembly.



(Chen)

Compact Error Resilient Computational DNA Tiling Assemblies

John Reif, Sudheer Sahu, Peng Yin

Department of Computer Science, Duke University

DNA 2004 Conference

[Reif,Sahu,Yin2004]

Self-Assembly of DNA Tiles

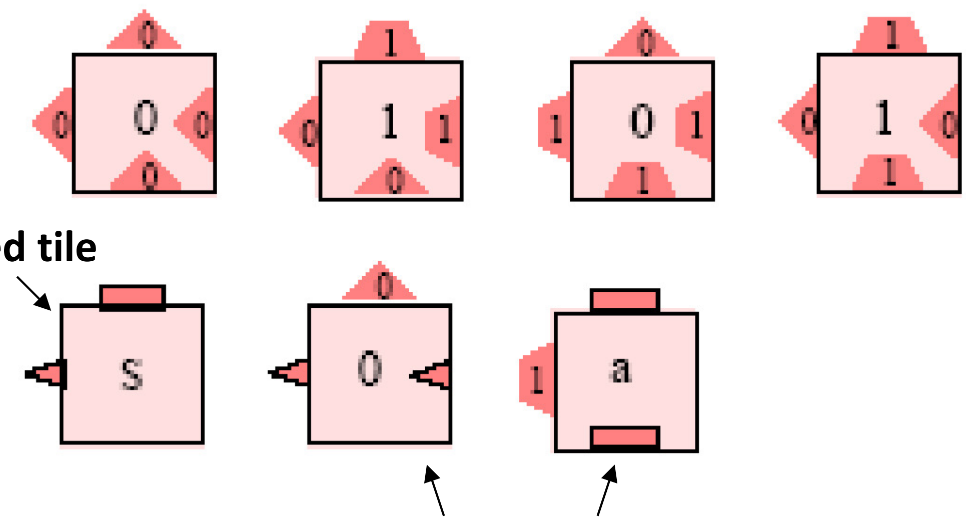
- **Perform universal computation.**
- **Manufacture patterned nanostructures from smaller unit nanostructures.**

[Reif,Sahu,Yin2004]

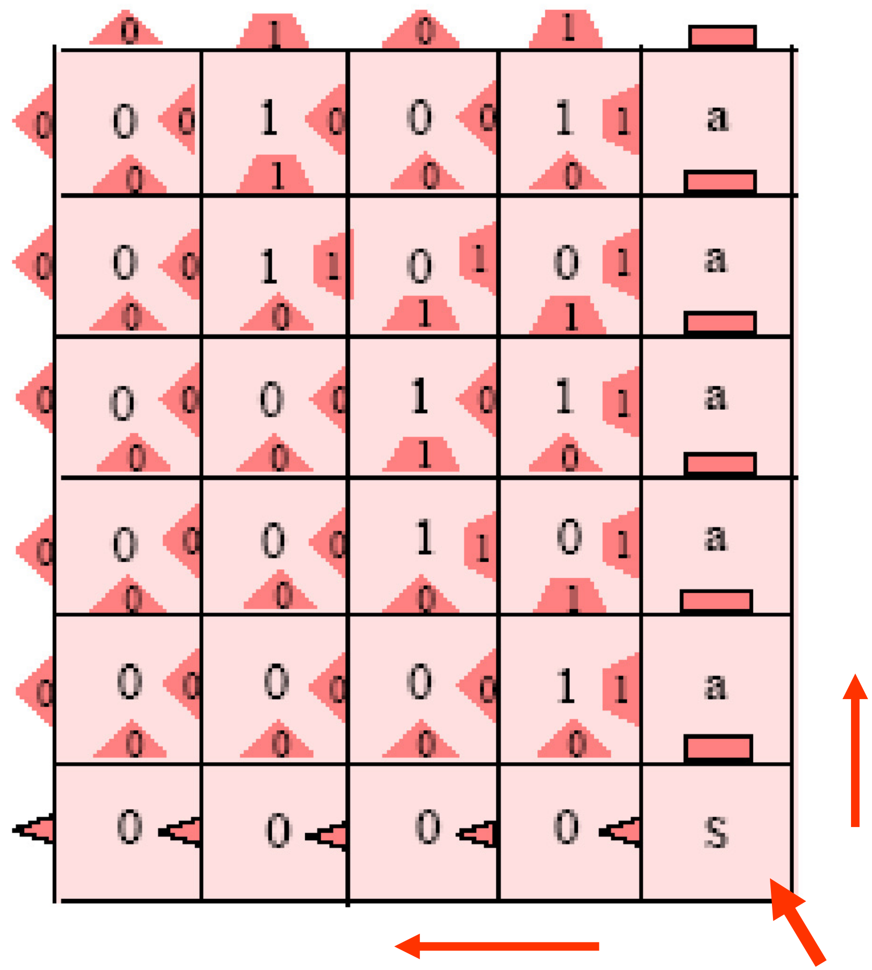
Assembly of Binary Counter (Winfree)

Computational tiles

Seed tile



Frame tiles



Binary counter

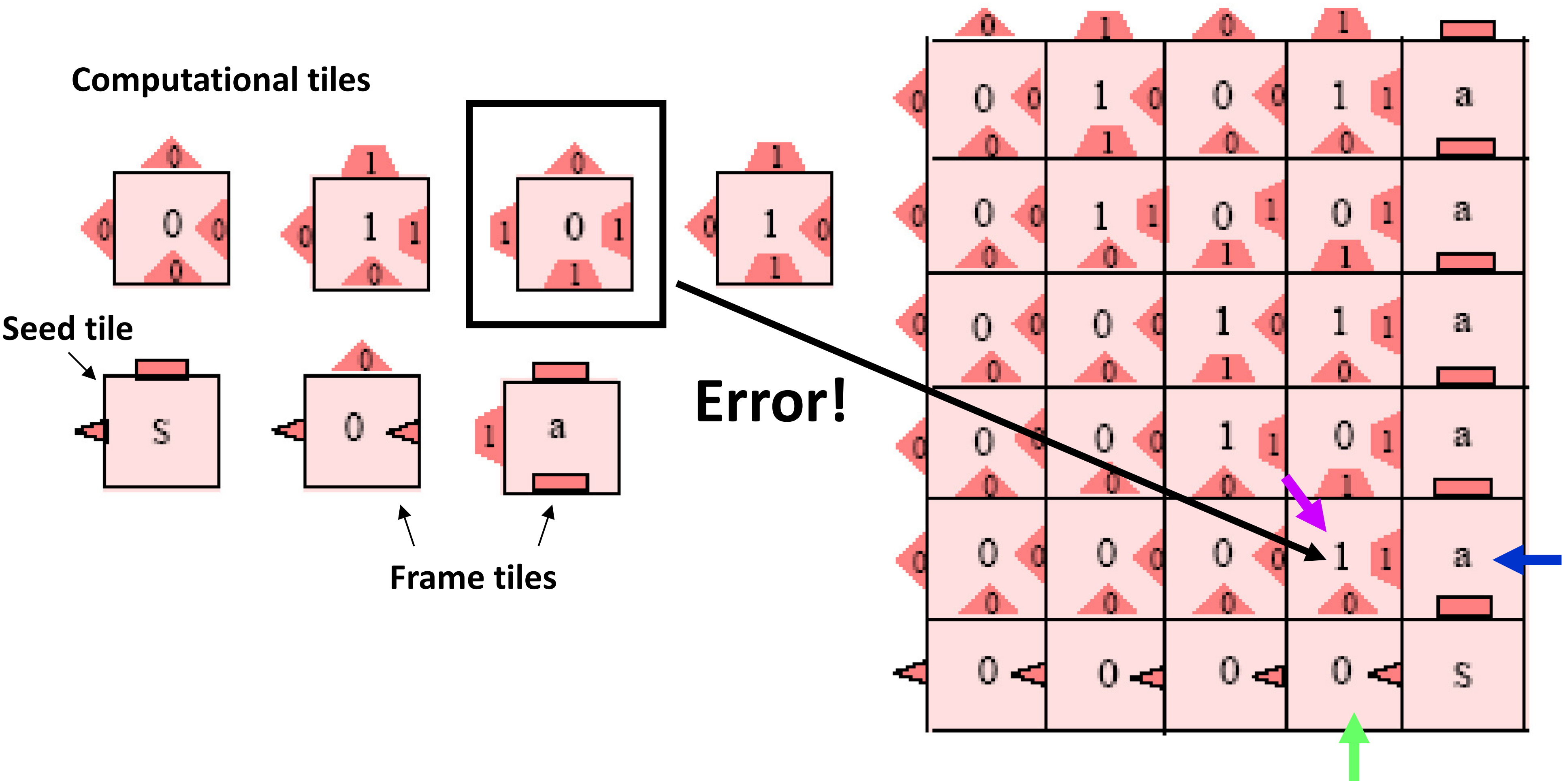
[Reif,Sahu,Yin2004]

Errors in Self-Assembly of DNA Tiles

- **Binding rules are not strict.**
- **A tile might get assembled to a binding site where it was not supposed to go.**

[Reif,Sahu,Yin2004]

Example of a Computational Error



[Reif,Sahu,Yin2004]

How to Decrease Errors?

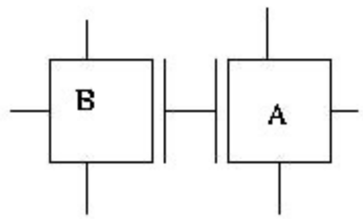
- Errors can be arbitrarily decreased by
  - Decreasing concentration of tiles.
  - Increasing binding strengths.
  - Drawback : Reduce speed.

- Another approach:
  - Change the logical design of the tiles.

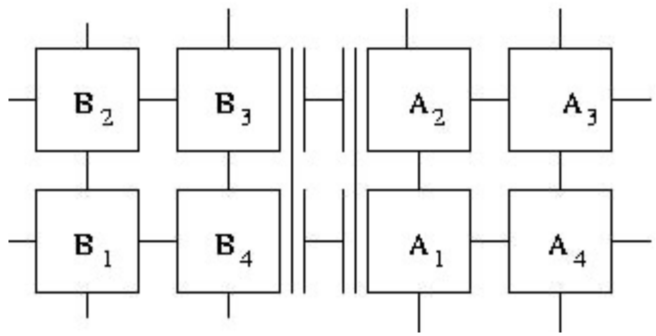
[Reif,Sahu,Yin2004]

Error Resilient Tilings by Winfree

Original tiles:



Error resilient tiles:



(Excerpted from Winfree 03)

Winfree's Construction:

Exchange each Tile with

2 x 2 array of tiles:

- Error rate reduced from $\epsilon \Rightarrow \epsilon^2$
- Assembly area increased by 4 times

Winfree's Generalized Construction:

Exchange each Tile with

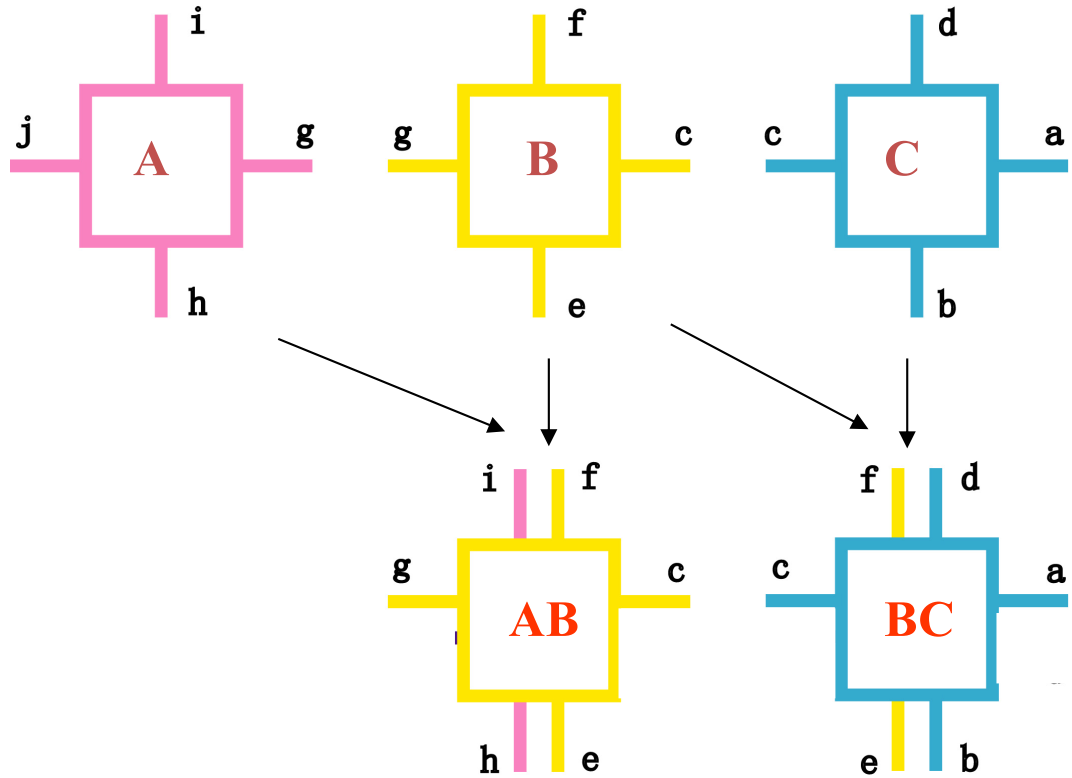
k x k array of tiles:

- Error rate reduced from $\epsilon \Rightarrow \epsilon^k$
- Assembly area increased by k^2

Error Resilient Tilings by [Reif,Sahu,&Yin 2004]

Original tiles:

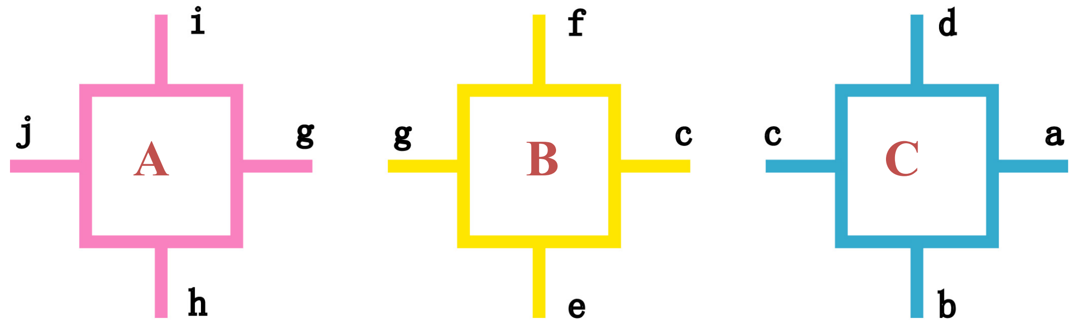
Error resilient tiles:



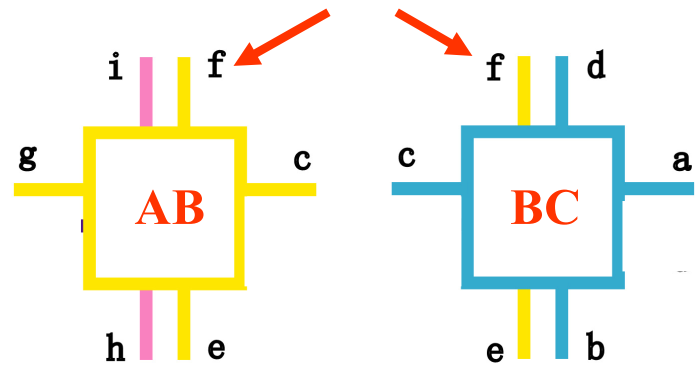
[Reif,Sahu,Yin2004]

Error Resilient Tilings by [Reif,Sahu,&Yin 2004]

Original tiles:



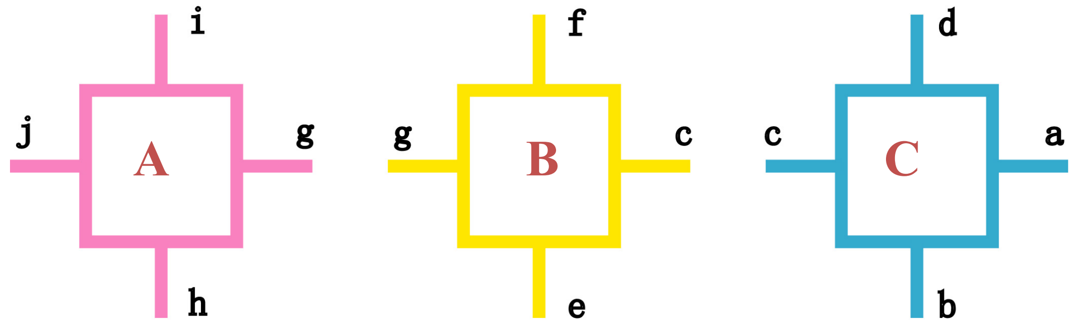
Error resilient tiles:



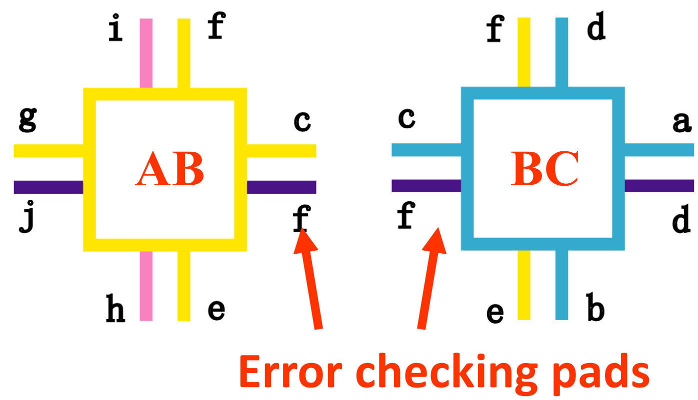
[Reif,Sahu,Yin2004]

Error Resilient Tilings by [Reif,Sahu,&Yin 2004]

Original tiles:

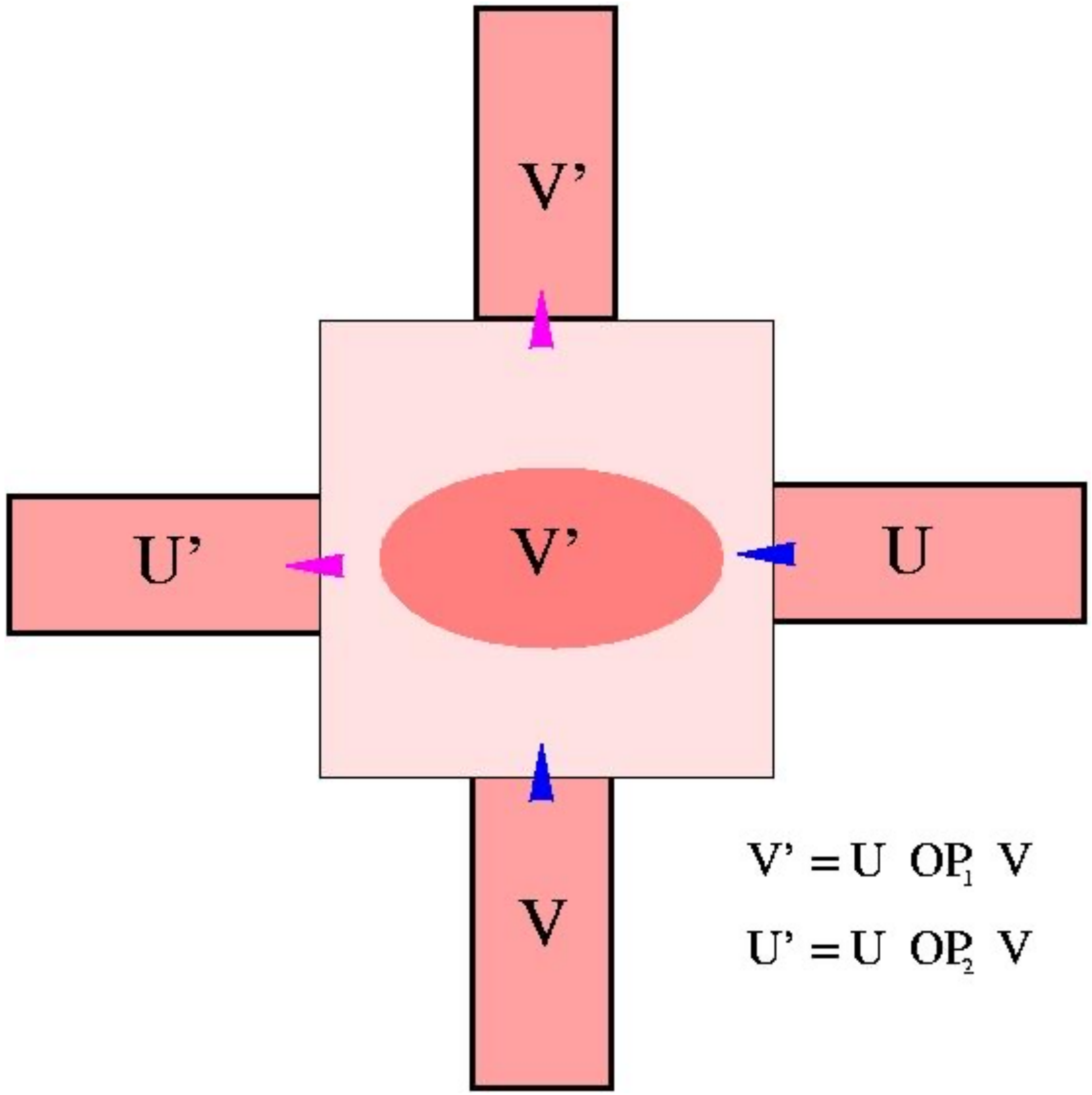


Error resilient tiles:



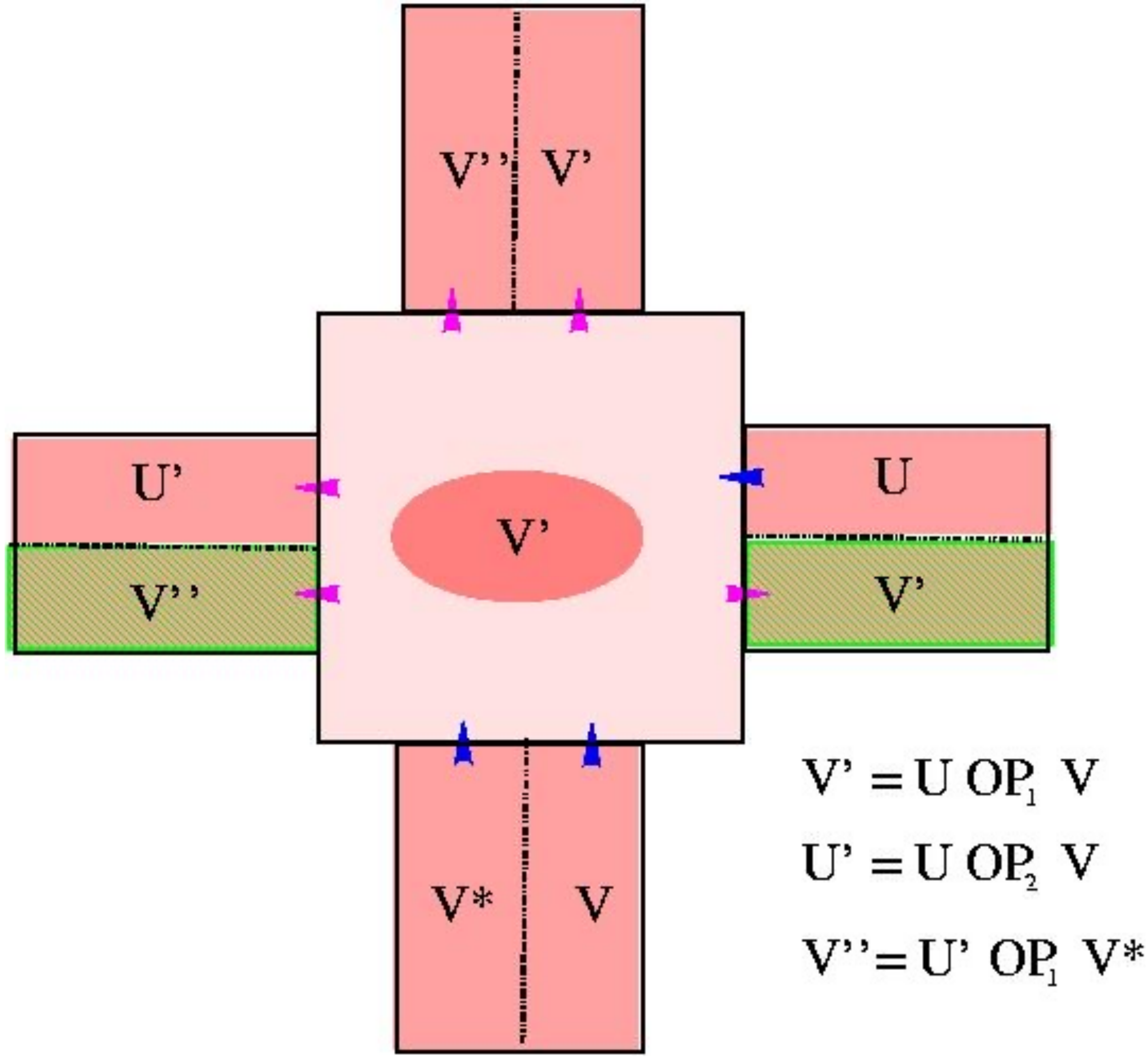
[Reif,Sahu,Yin2004]

A Computational Tile



[Reif,Sahu,Yin2004]

Compact Error Resilient Construction

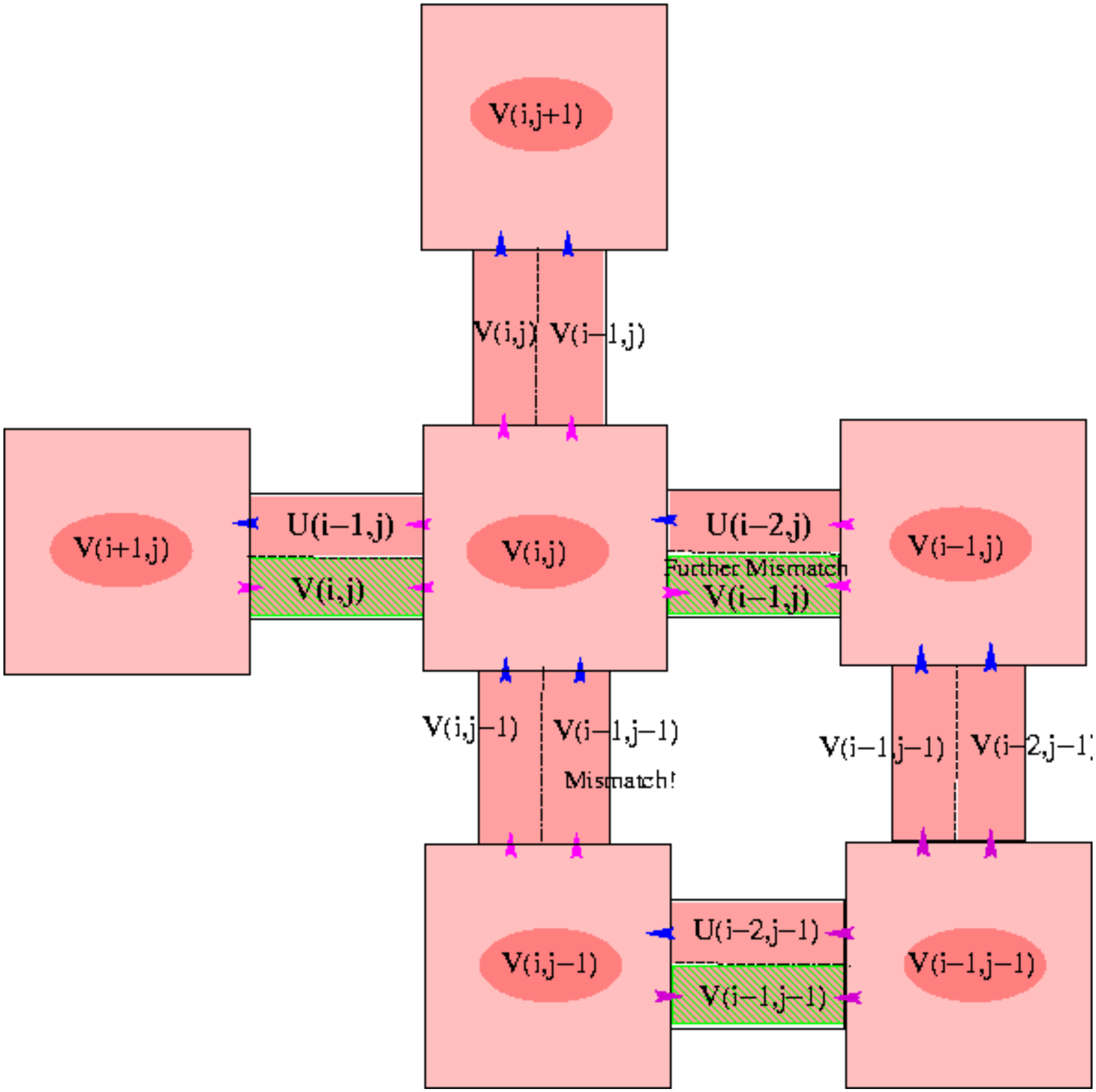


- Wholeness of pad: Single pad per side.

[Reif,Sahu,Yin2004]

One Mismatch causes more Mismatch

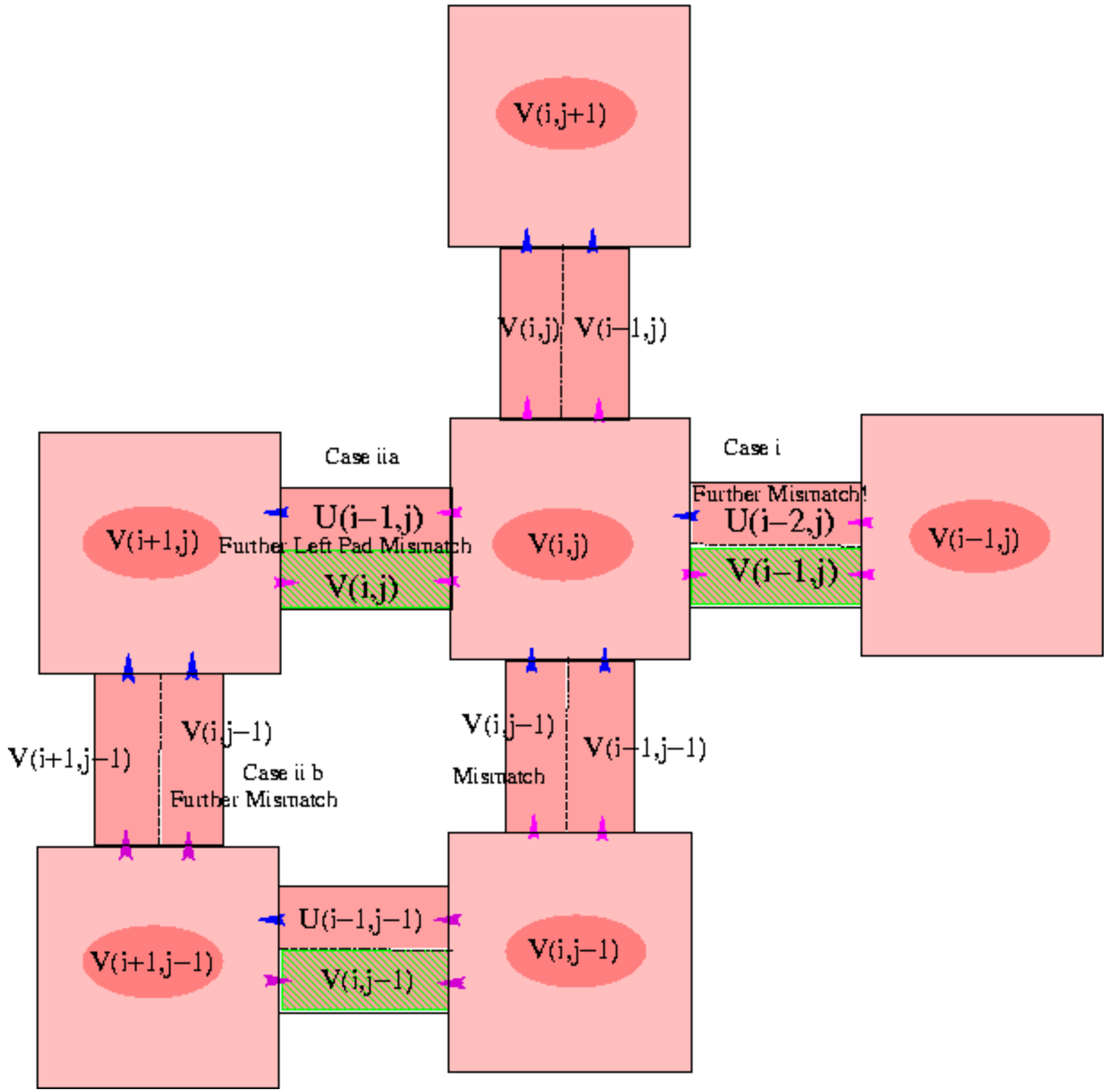
Case 1



[Reif,Sahu,Yin2004]

One Mismatch causes more Mismatch

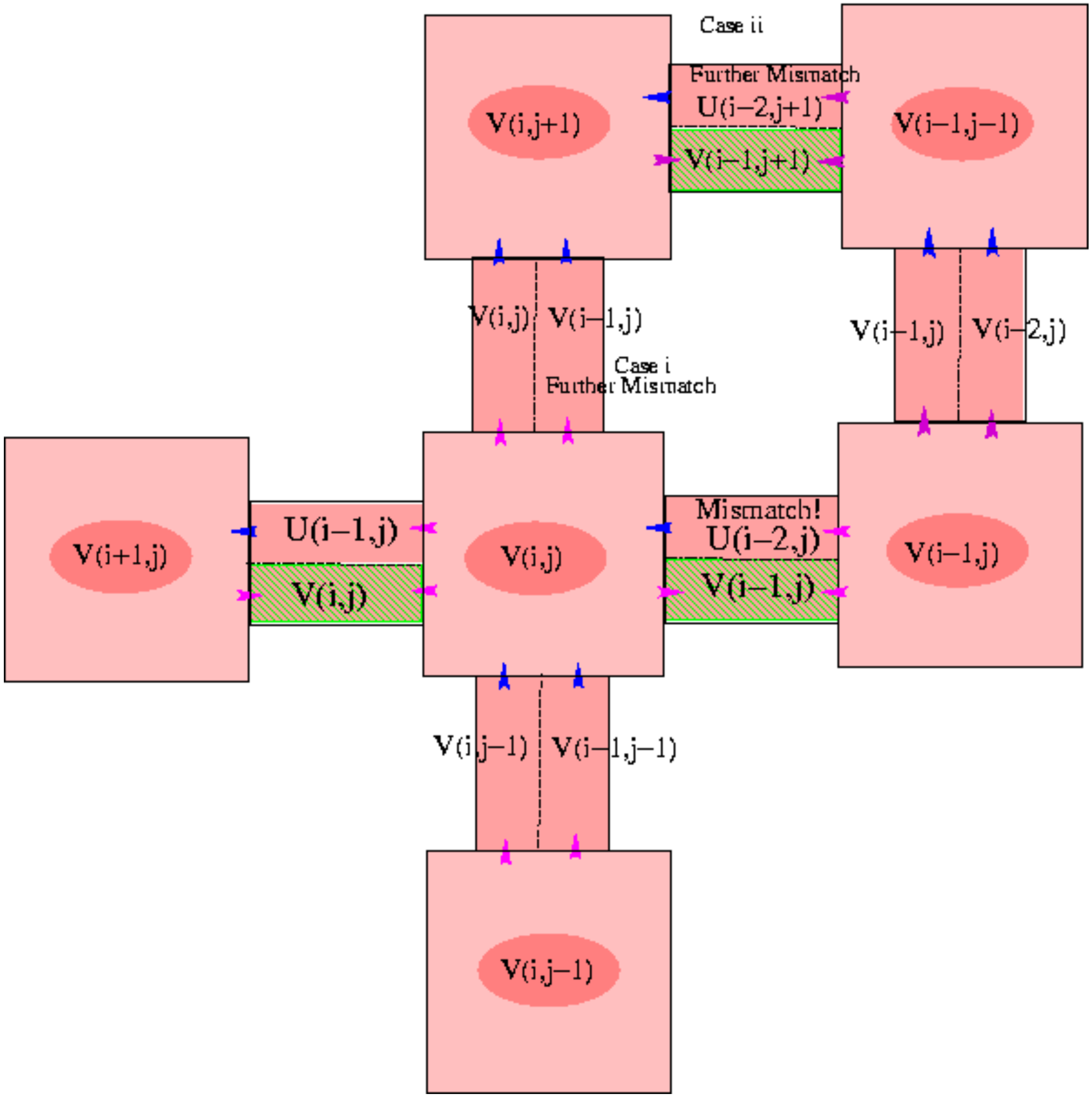
Case 2



[Reif,Sahu,Yin2004]

One Mismatch causes more Mismatch

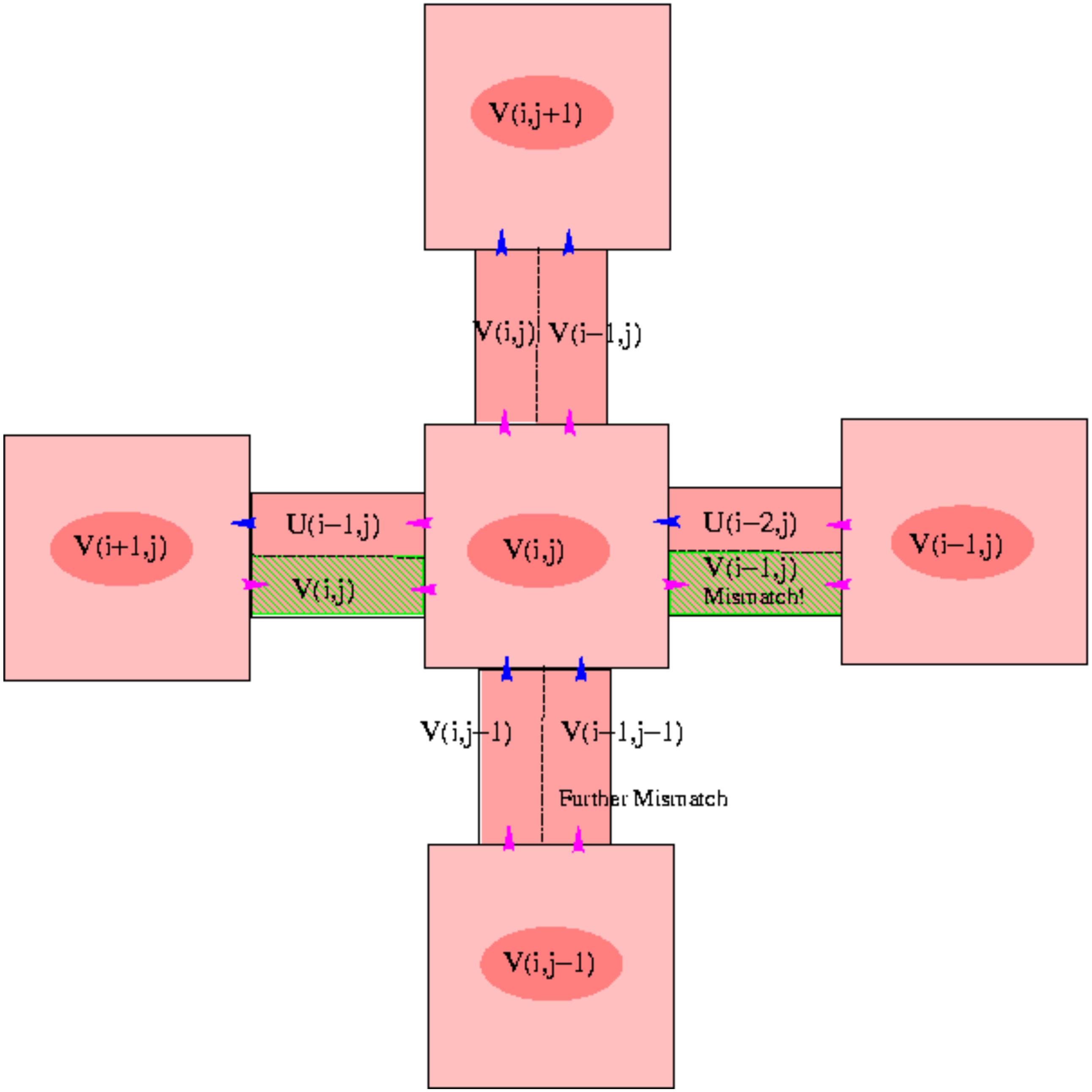
Case 3



(Reif,Sahu,Yin)

One Mismatch causes more Mismatch

Case 4



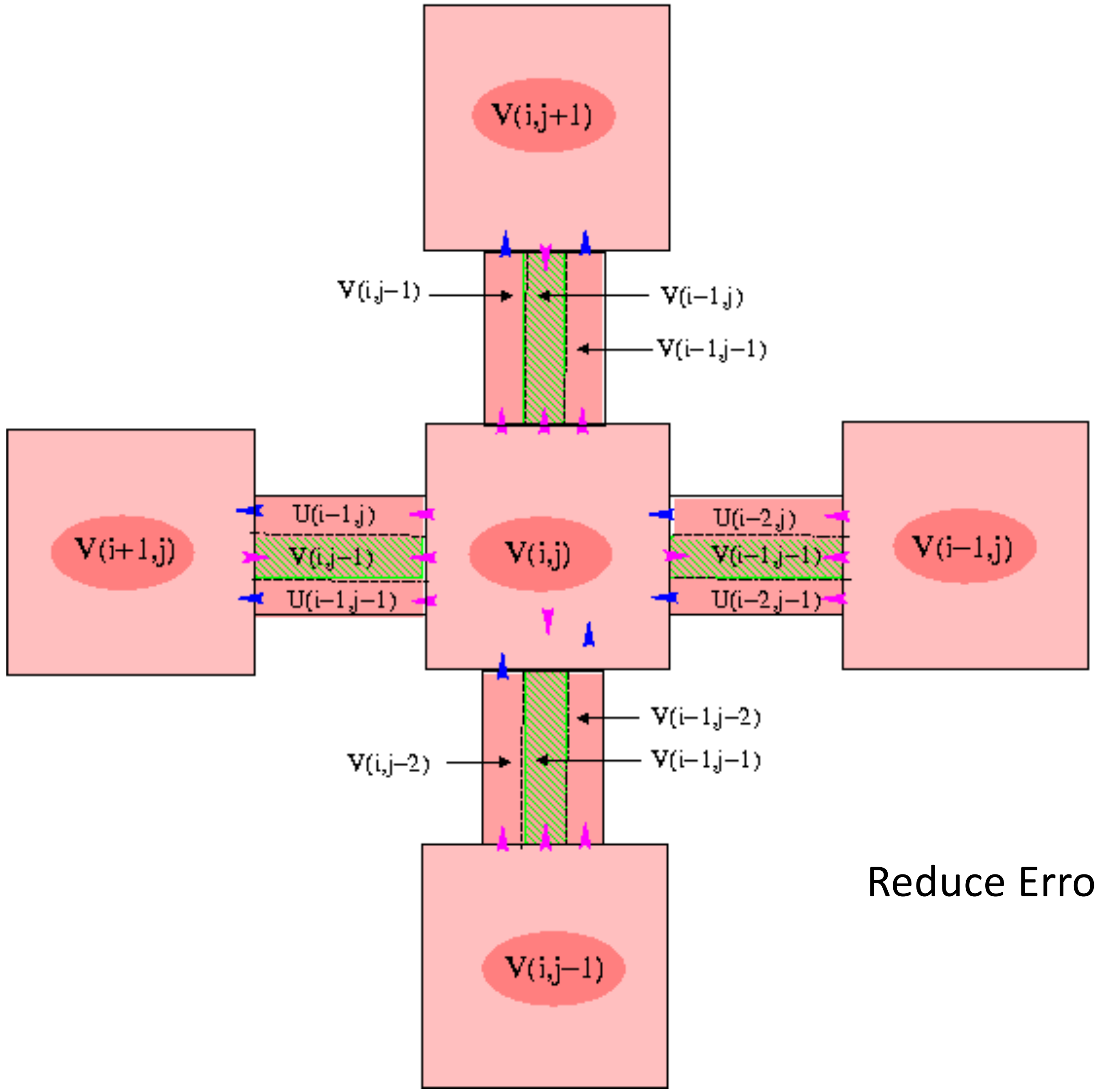
[Reif,Sahu,Yin2004]

Result of Compact Error Resilient Scheme

- We saw:
  - Two way overlay scheme.
  - One mismatch caused at least one more mismatch.
  - Error is reduced from ϵ to ϵ^2.
- Next we will see:
  - Three way overlay scheme.
  - One mismatch will cause at least two more mismatches.
  - Error can be reduced from ϵ to ϵ^3.

[Reif,Sahu,Yin2004]

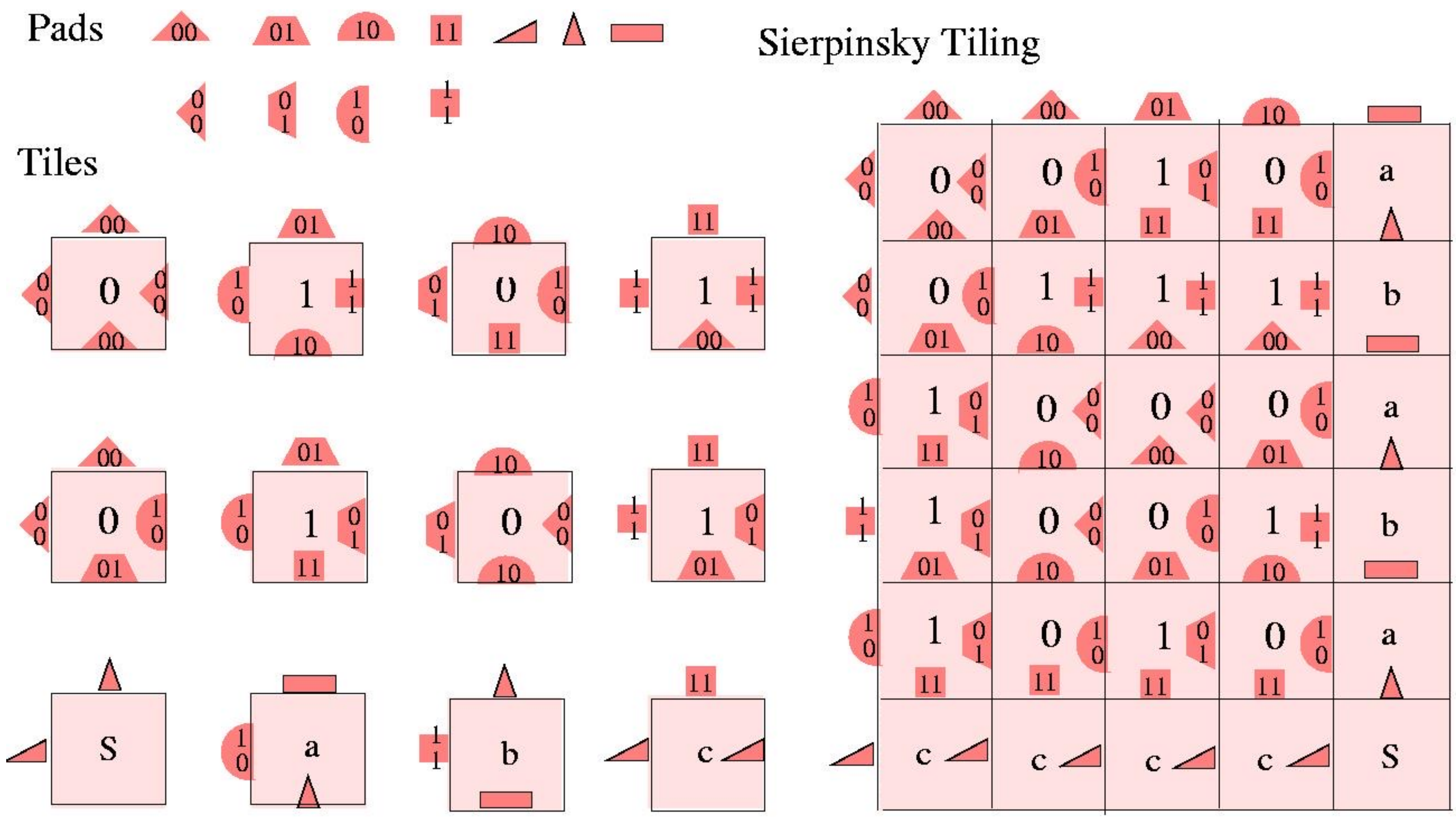
Compact Error Resilient Tiles (3-way overlay)



Reduce Error from $\in$ to $\in^3$

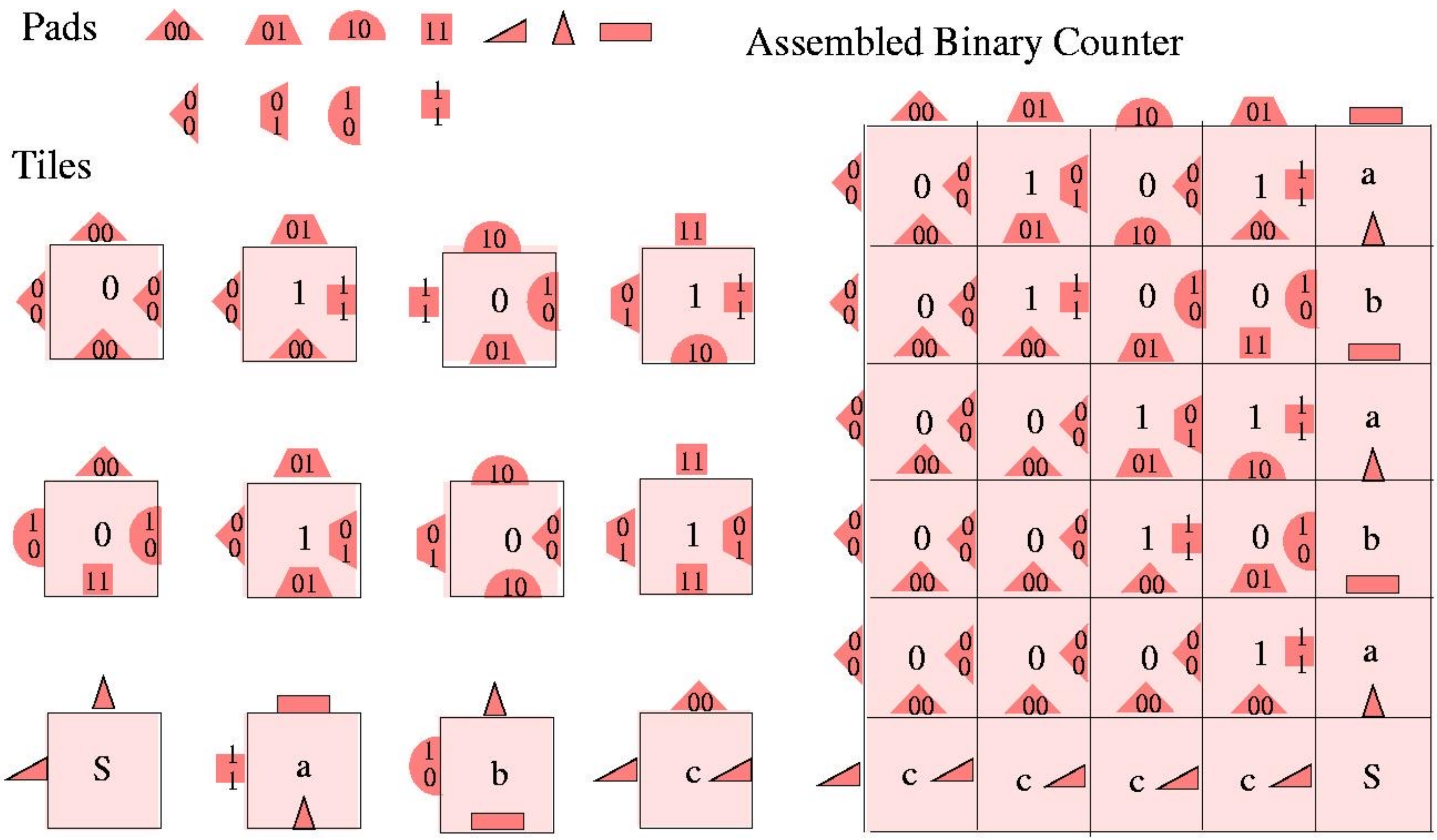
[Reif,Sahu,Yin2004]

Examples of Error Resilient Assembly



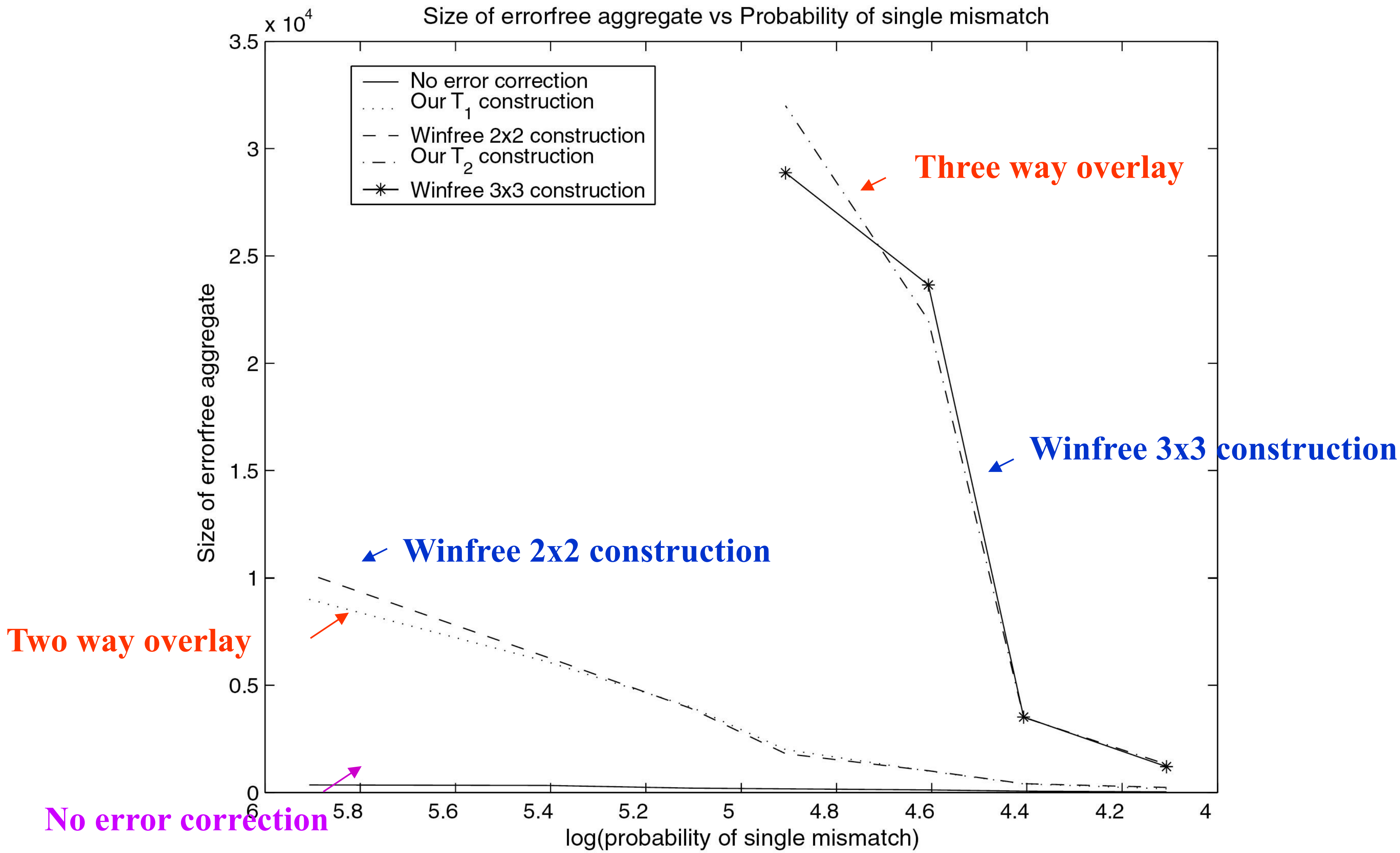
[Reif,Sahu,Yin2004]

Examples of Error Resilient Assembly



[Reif,Sahu,Yin2004]

Computer Simulation (Xgrow, Winfree)



[Reif,Sahu,Yin2004]

Conclusions

- **Assembly size not increased by this error-resilient tile design.**
- **Two way overlay: error rate ϵ (5%) => ϵ^2(0.25%).**
- **Three way overlay: error rate ϵ (5%) => ϵ^3 (0.0125%).**
- **Open question: Can we reduce error rate ϵ => ϵ^k ?**

[Reif,Sahu,Yin2004]