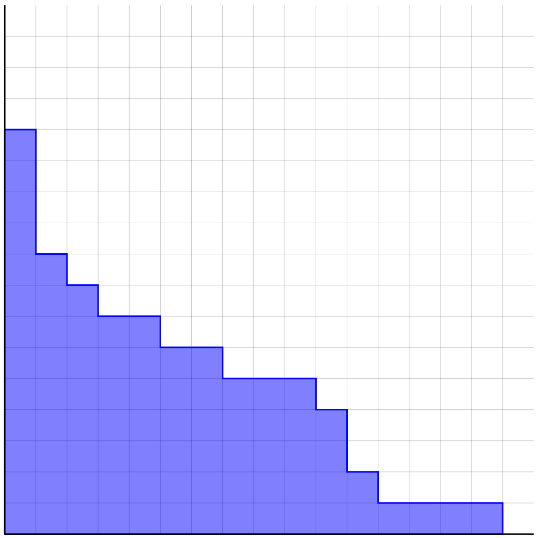
Sampling Algorithms and Phase Transitions

Dana Randall

Georgia Institute of Technology

Monotone surfaces

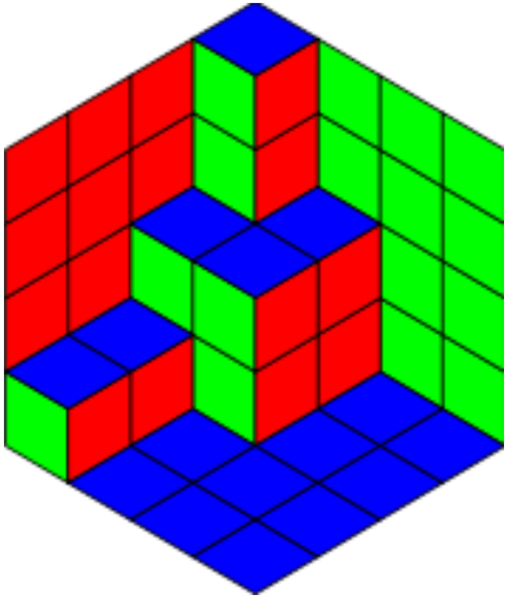
1)



2-d

- Exclusion processes
- Permutations
- Dyck Paths (> diagonal)
- Integer partitions

2)

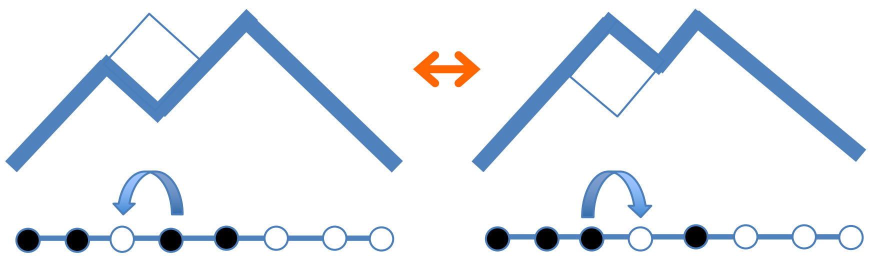


3-d

- Plane partitions
- Lozenge tilings
- Dimer model on hexagonal lattice

One path

1) Simple Exclusion Process:



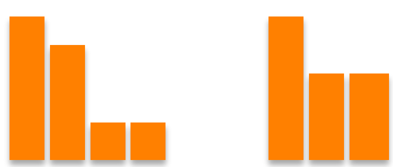
("Mountain-valley chain")

2) Permutations with Nearest Neighbor Transpositions:

0	1	0	0	0	0	0
0	1	0	0	1	0	0
0	1	0	1	1	0	0
0	1	0	1	1	1	0
0	1	0	1	1	1	1
1	1	0	1	1	1	1

3) Integer Partitions: unordered sums of positive integers

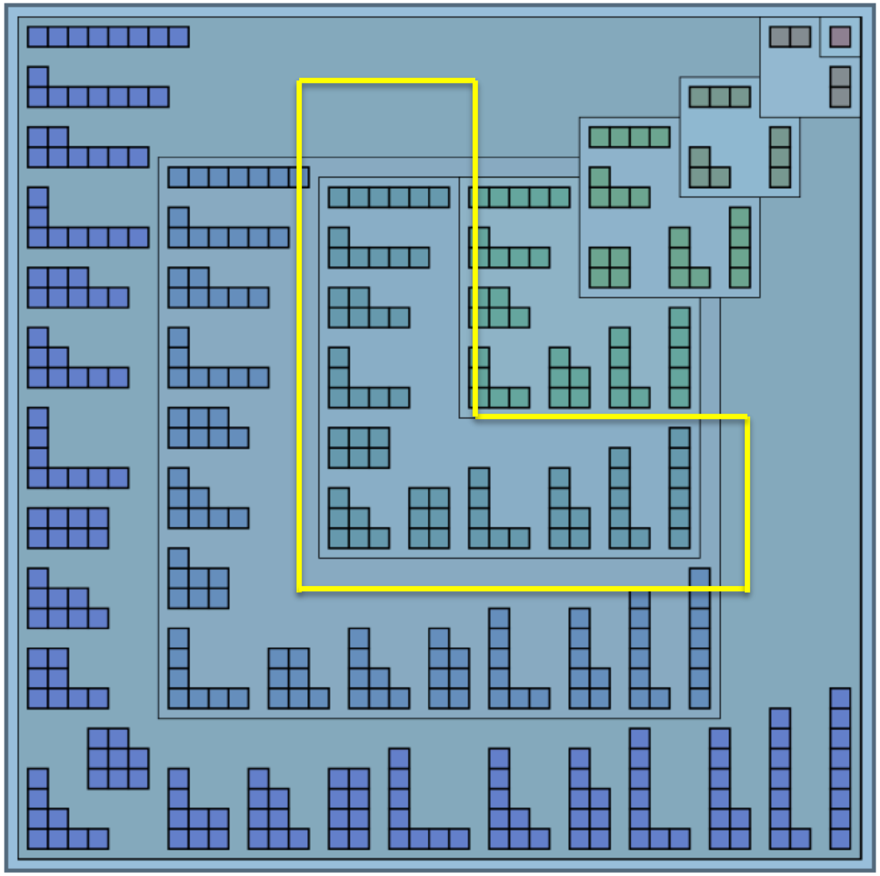
Ex: 5+4+1+1, 5 + 3 + 3, & 11 are all partitions of 11.



Integer partitions as paths

Integer Partitions

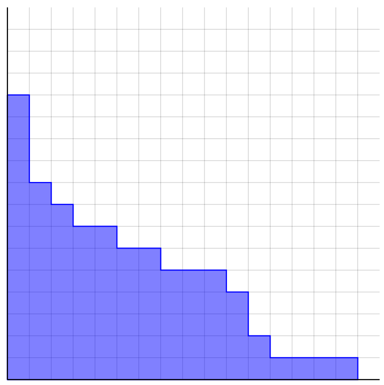
Ferrers Diagrams:



* Sampling integer partitions of n is the same as sampling
lattice paths bounding regions of area n.

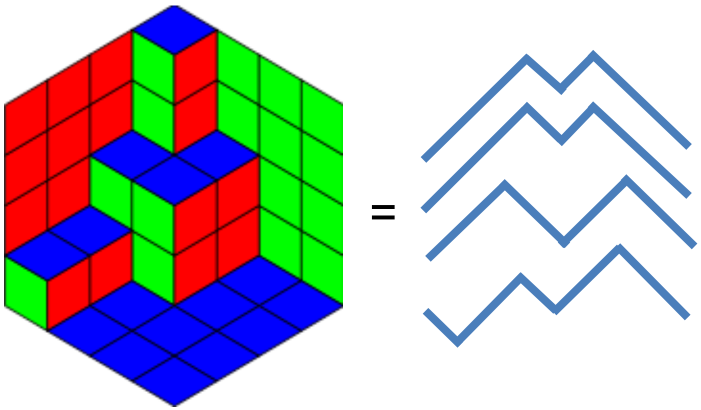
Monotone surfaces and height functions

1)



2-d

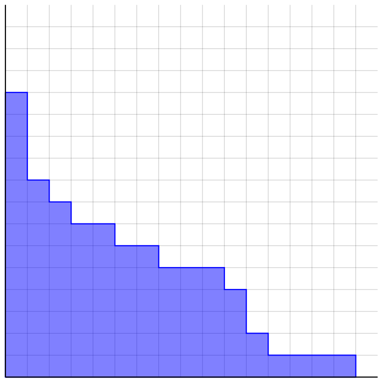
2)



3-d: Disjoint paths

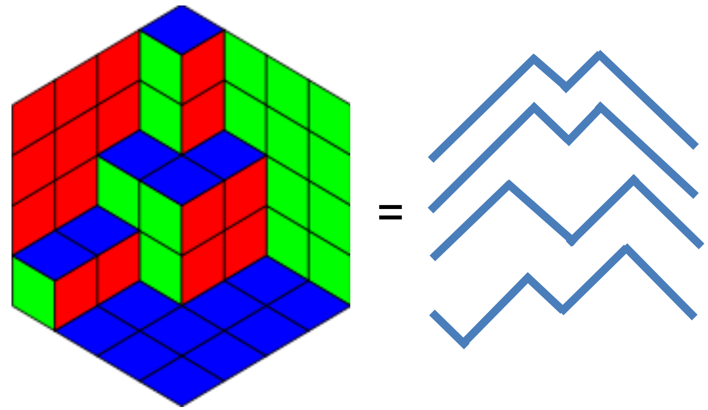
Monotone surfaces and height functions

1)



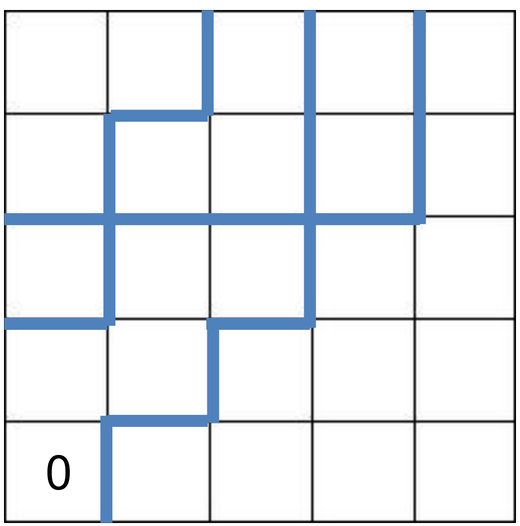
2-d

2)



3-d: Vertex disjoint paths

3)



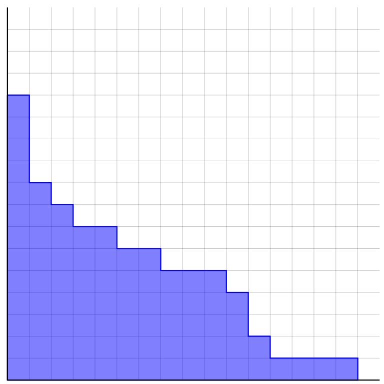
Edge-disjoint paths

Assign colors using the rules:

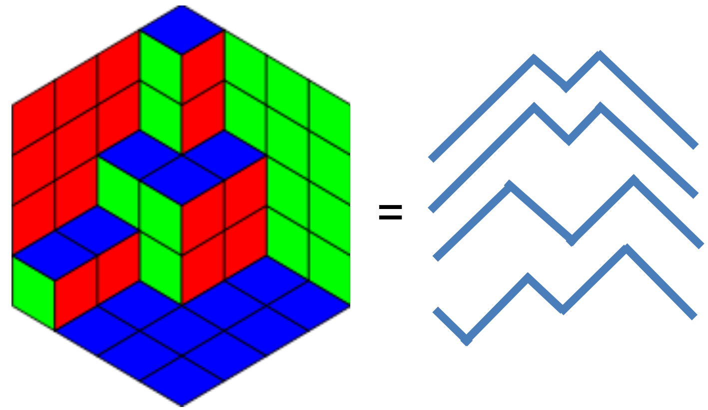
i | i+1 (mod 3) or i over i+1 (mod 3)

i | i-1 (mod 3) or i over i-1 (mod 3)

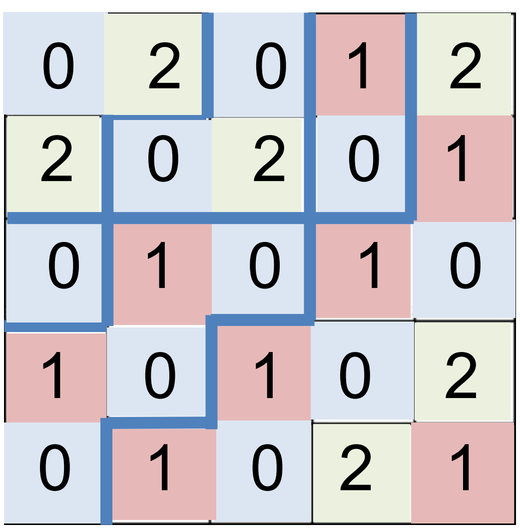
Monotone surfaces and height functions



2-d



3-d: Vertex disjoint paths



Edge-disjoint paths = 3-colorings

Assign colors using the rules:

i | i+1 (mod 3) or $\frac{i}{i+1 \pmod 3}$

i | i-1 (mod 3) or $\frac{i}{i-1 \pmod 3}$

Talk Outline

1. Sampling *unweighted* monotone surfaces and colorings

2. Sampling *biased* monotonic surfaces
   - Integer partitions of size n
   - Biased lozenge tilings

3. Sampling *weighted* 3-colorings: the “six-vertex model”
   - The antiferroelectric phase
   - The ferroelectric phase

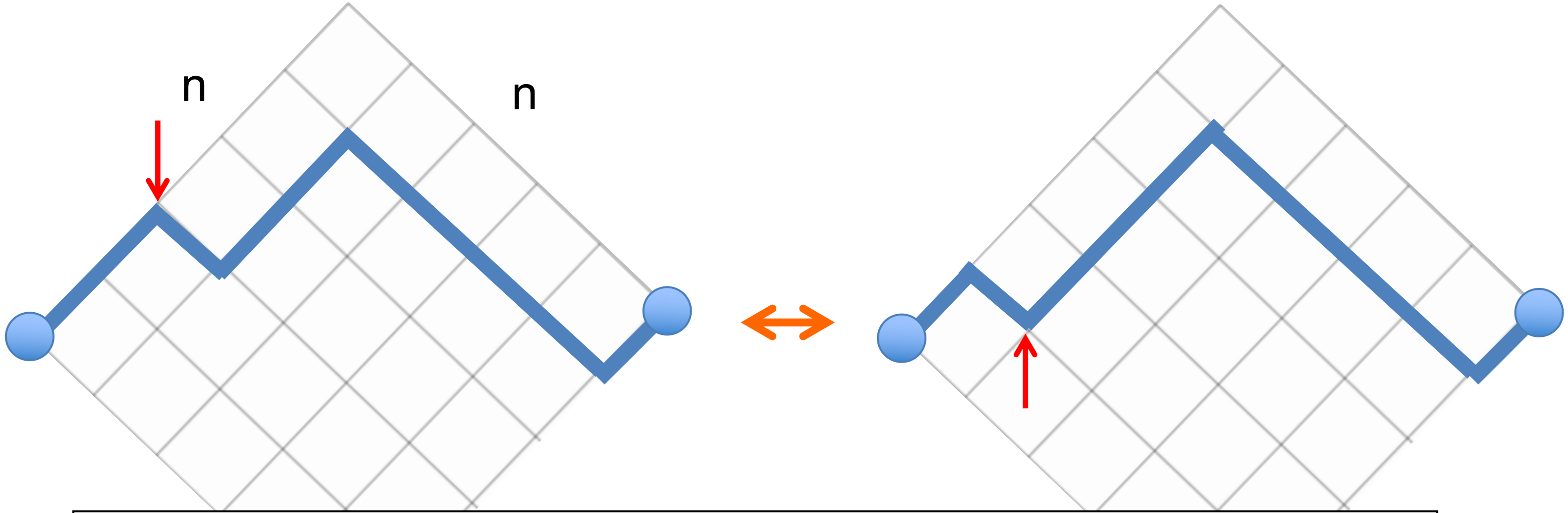
Talk Outline

1. Sampling *unweighted* monotone surfaces and colorings

2. Sampling *biased* monotonic surfaces
   - Integer partitions of size n
   - Biased lozenge tilings

3. Sampling *weighted* 3-colorings: the “six-vertex model”
   - The antiferroelectric phase
   - The ferroelectric phase

The Mountain-Valley Markov Chain



To sample, repeat:

- Pick **v** on the path;
- If **v** marks a mountain/valley, invert w.p. 1/2.

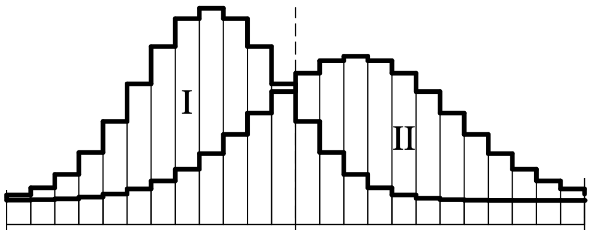
This Markov chain is reversible and ergodic, so it converges to the uniform distribution over lattice paths.

How long?

The mixing time

Def: The total variation distance is

$$||P^t,\pi|| = \max_{x\in\Omega} \frac{1}{2} \sum_{y\in\Omega} |P^t(x,y) - \pi(x)|.$$



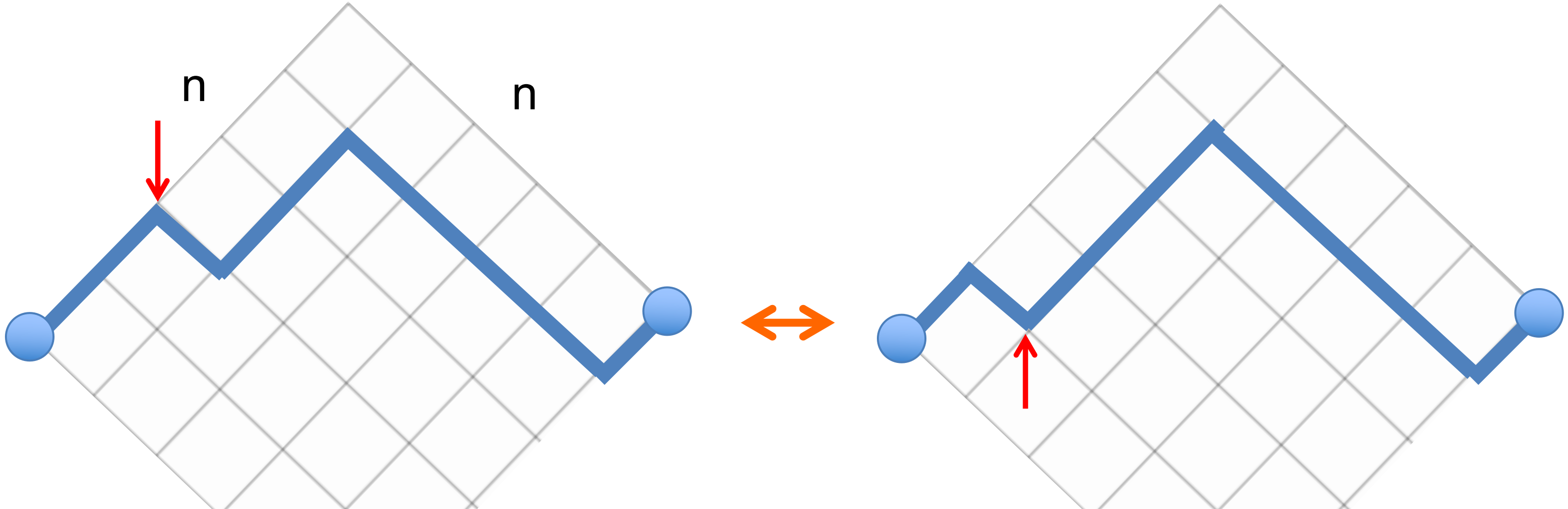
Def: Given ε, the mixing time is

$$\tau(\varepsilon) = \min\{t: ||P^{t'},\pi|| < \varepsilon,\ \forall\, t' \geq t\}.$$

A Markov chain is rapidly mixing if $\tau(\varepsilon)$ is poly(n, log(ε^{-1})).
(or polynomially mixing)

A Markov chain is slowly mixing if $\tau(\varepsilon)$ is at least exp(n).

The Mountain Valley Markov Chain



To sample, repeat:

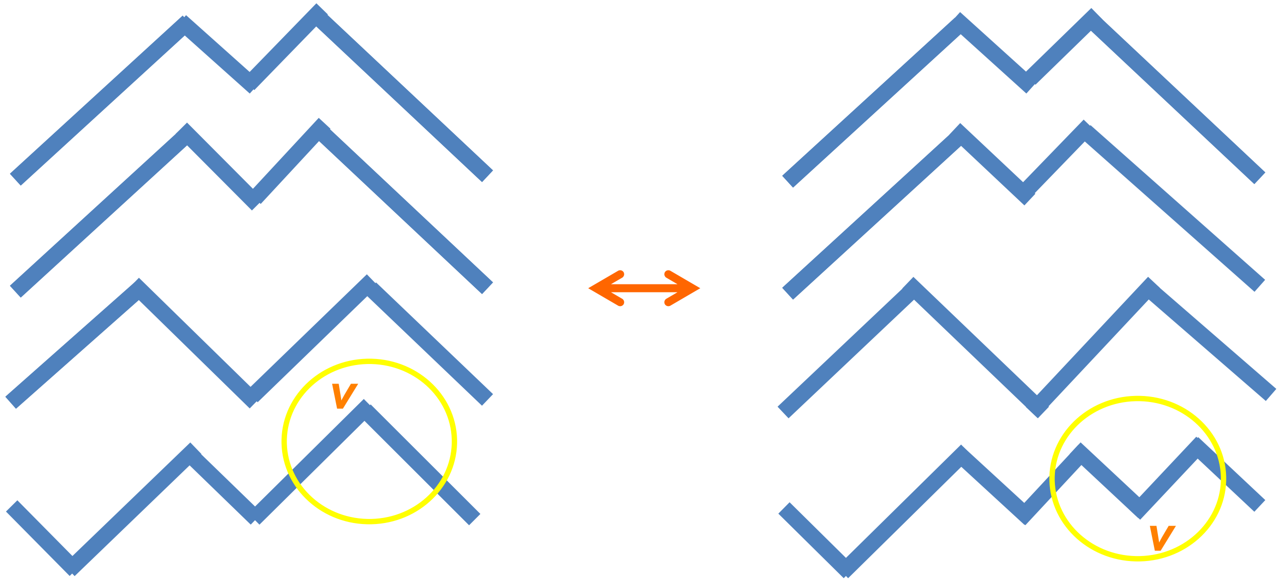
- Pick **v** on the path;
- If **v** marks a mountain/valley, invert w.p. 1/2.

This Markov chain is reversible and ergodic, so it converges to the uniform distribution over lattice paths.

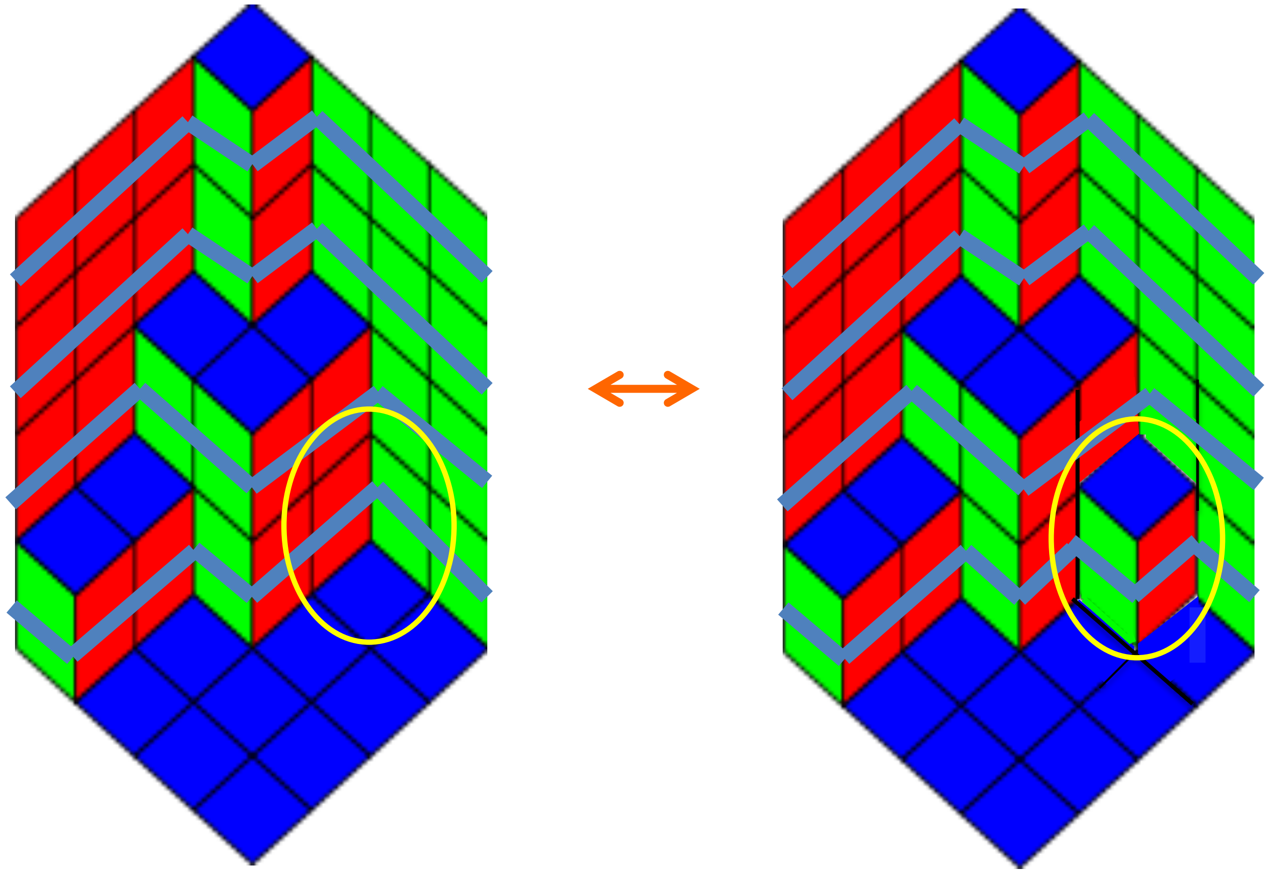
How long? Answer: $\Theta(n^3 \log n)$ [Wilson]

Multiple paths

The mountain/valley chain on multiple (vertex disjoint) paths.

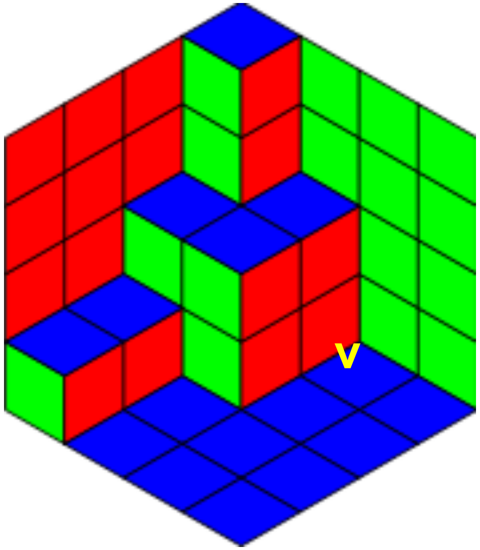


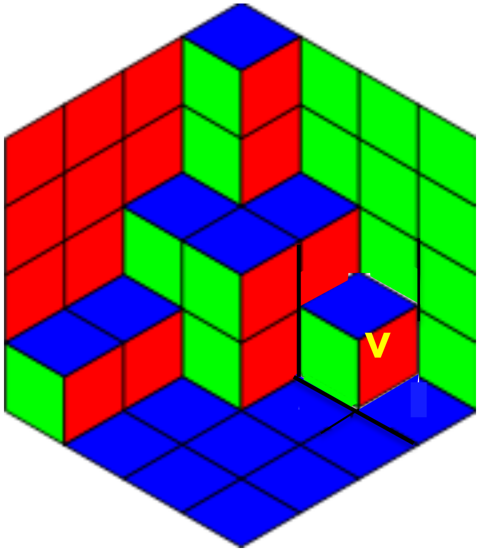
Multiple paths



There is a bijection between nonintersecting lattice paths and lozenge tilings (or dimer coverings).

Glauber dynamics for lozenges

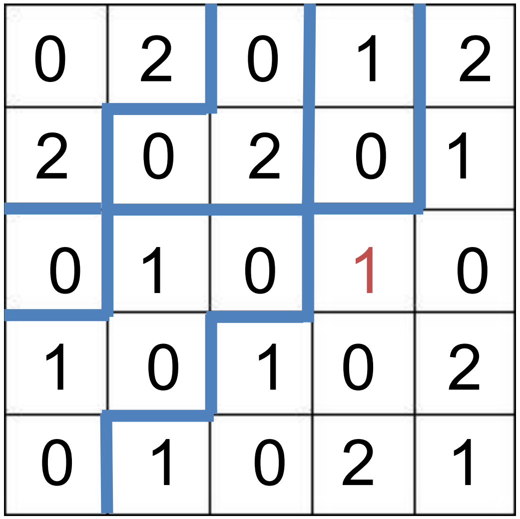


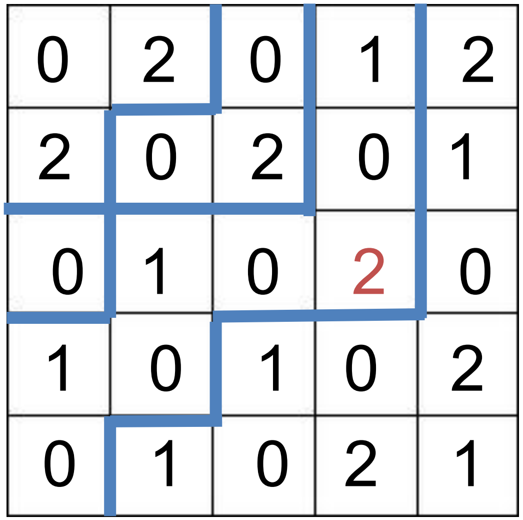


Repeat:

- Pick **v** in the lattice region;
- Add / remove the "cube" at **v** w.p. ½, if possible.

Glauber dynamics for 3-colorings





Repeat:

- Pick a cell uniformly;
- Recolor the cell w.p. ½, if possible.

This is also a “mountain-valley” move!

Sampling Monotonic Surfaces in Z^d

Do the Mountain-Valley Chains converge in poly time?

Lattice paths, lozenge tilings and plane/space partitions:

d=2: Yes (simple coupling)

d=3: Yes [Luby, R., Sinclair], [Wilson]

d≥4: ???

3-colorings of Z^d:

d=2: Yes (simple coupling)

d=3: Yes [Luby, R., Sinclair], [Martin, Goldberg, Patterson], [R., Tetali]

d=4: ???

Sampling Monotonic Surfaces in Z^d

Do the Mountain-Valley Chains converge in poly time?

Lattice paths, lozenge tilings and plane/space partitions:

d=2: Yes (simple coupling)

d=3: Yes [Luby, R., Sinclair], [Wilson]

d≥4. ???

3-colorings of Z^d:

d=2: Yes (simple coupling)

d=3: Yes [Luby, R., Sinclair], [Martin, Goldberg, Patterson], [R., Tetali]

d=4: ???

d >> 4: No! [Galvin, Kahn, R., Sorkin], [Peled]

Talk Outline

1. Sampling *unweighted* monotone surfaces and colorings

2. Sampling *biased* monotonic surfaces

    - Integer partitions of size n
    - Biased lozenge tilings

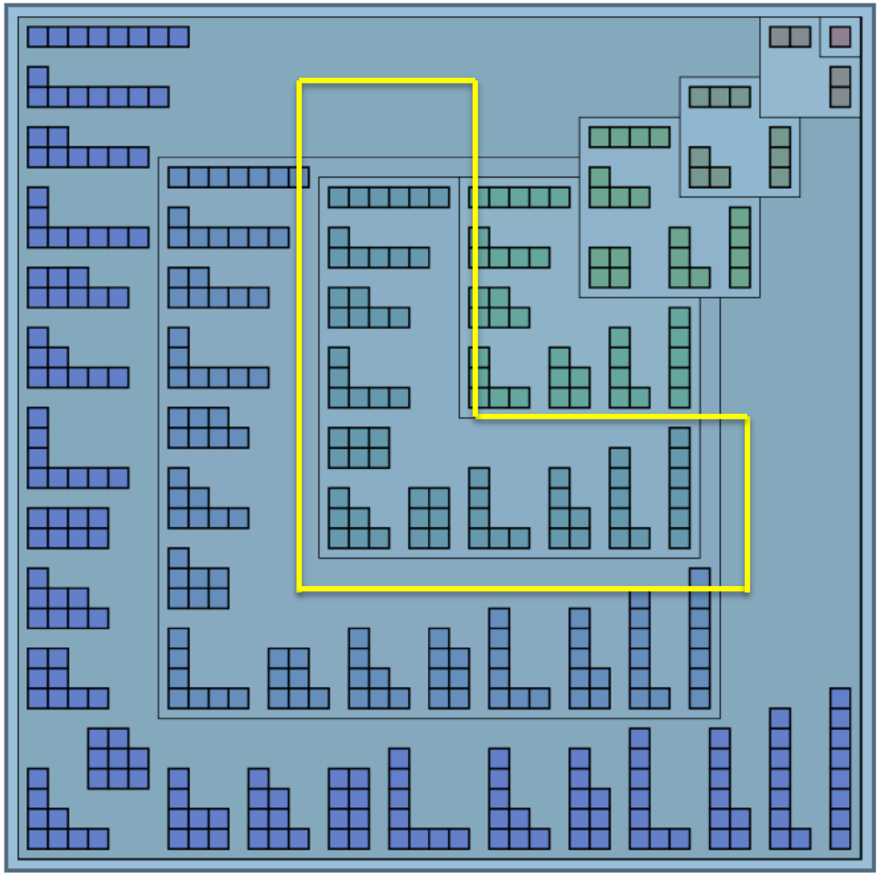
3. Sampling *weighted* 3-colorings: the “six-vertex model”

    - The antiferroelectric phase
    - The ferroelectric phase

Integer partitions as paths

Integer Partitions

Ferrers Diagrams:

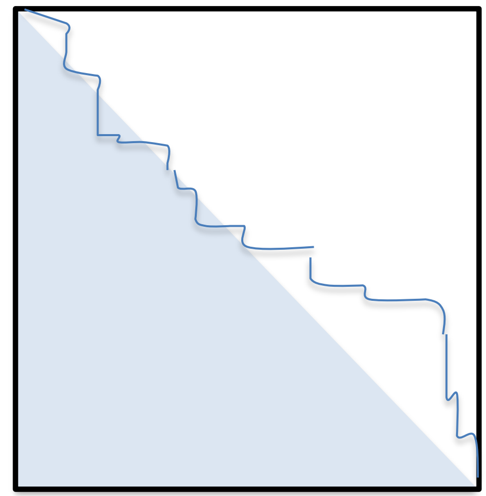


* Sampling integer partitions of n is the same as sampling lattice paths bounding regions of area n.

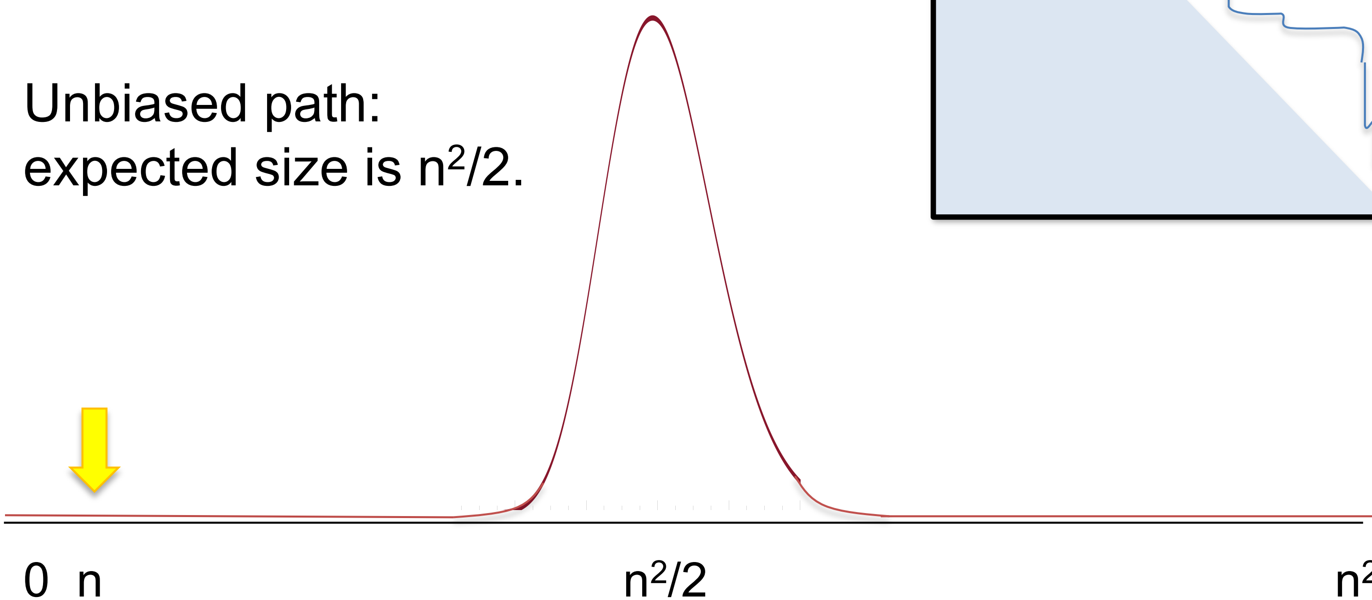
Sampling Ferrers Diagrams

We want to sample region of area n in an n x n lattice region.

Sample from *all* paths in the n x n grid and output if its area is n.



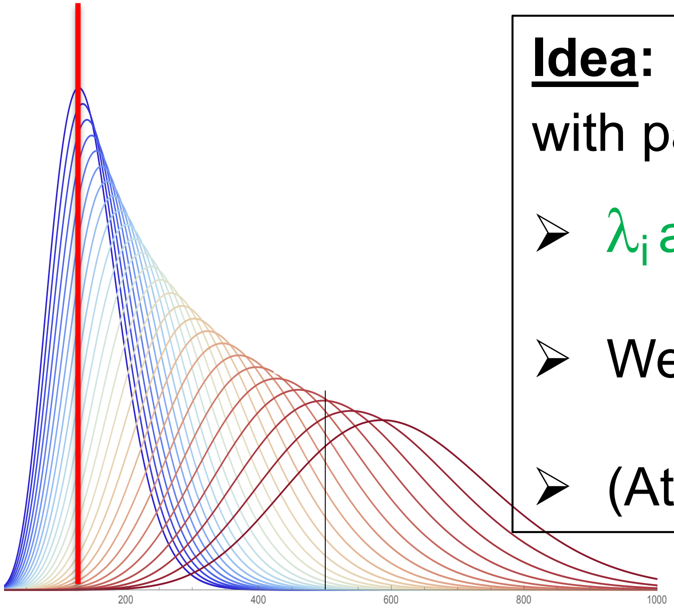
Unbiased path:
expected size is $n^2/2$.



Boltzmann Sampling

Solution: Use *biased* walks: [Cousins, Bhakta, Fahrbach, R. '17]

Given λ, the Gibbs (or Boltzmann) distribution gives
a path with area k weight λ^k.



Idea: Make a polynomial number of distributions with parameters $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$ so that:

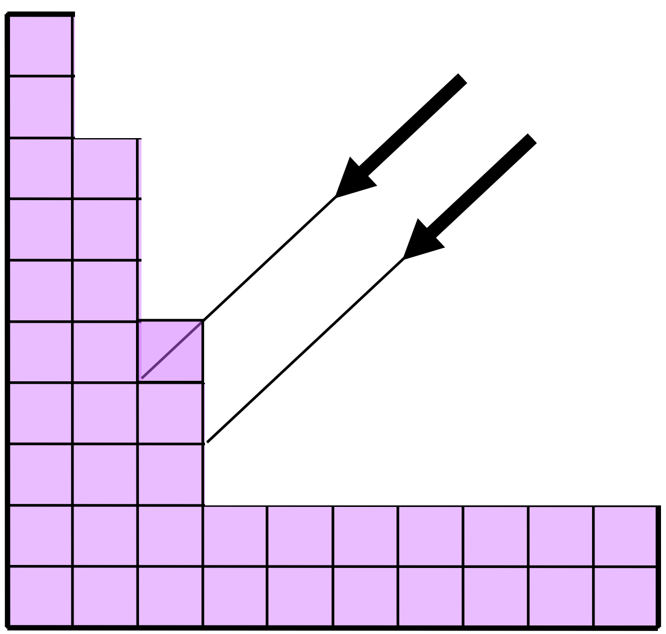
- λ_i and λ_{i+1} are close (their distributions overlap a lot)
- We can sample efficiently from each distribution
- (At least) one has nontrivial mass at n.

Then sample from each until there is an output a region with area n.

(Each of these area n regions are equally likely for each distribution!)

Sampling Biased Surfaces

Given $0 < \lambda < 1$:



Repeat:

Choose (v,d) in S x {+,-}.

If a square can be added at v, and d = +, add it w.p. λ;

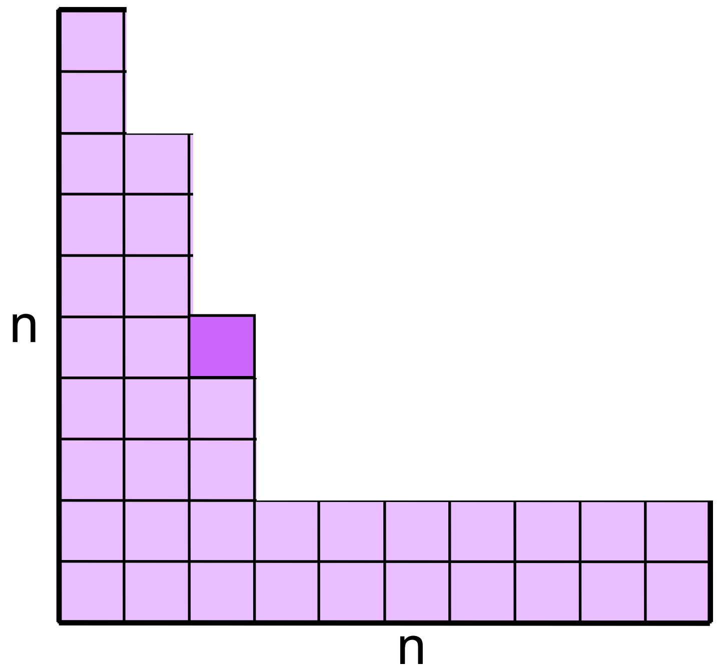
If a square can be removed at v, and d = -, remove it w.p. 1;

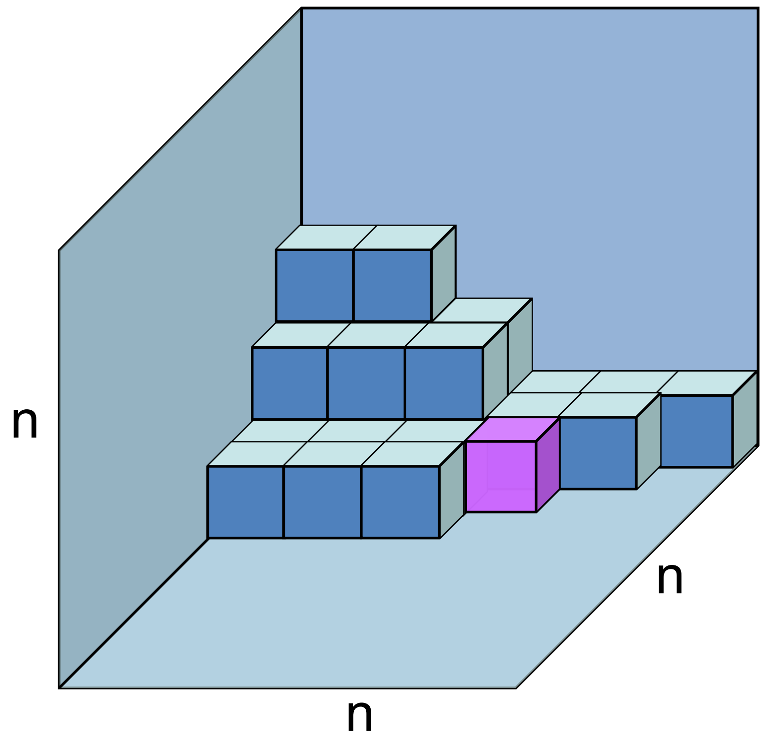
Otherwise do nothing.

Converges to the distribution:

$$\pi(S) = \lambda^{\mathrm{area}(S)} / Z.$$

Generating Biased Surfaces





How fast?

Biased Surfaces in Z^d

Q: How long does the biased MC take to converge?

[Benjamini, Berger, Hoffman, Mossel '05].
$d = 2$; $\lambda > 1$ const, $O(n^2)$ mixing time (optimal).

[Greenberg, Pascoe, R. '09]
$$\left.\begin{array}{ll} d = 2, & \lambda > 1 \text{ const} \\ d \geq 3, & \lambda > d^2 \end{array}\right\} O(n^d) \text{ mixing time.}$$

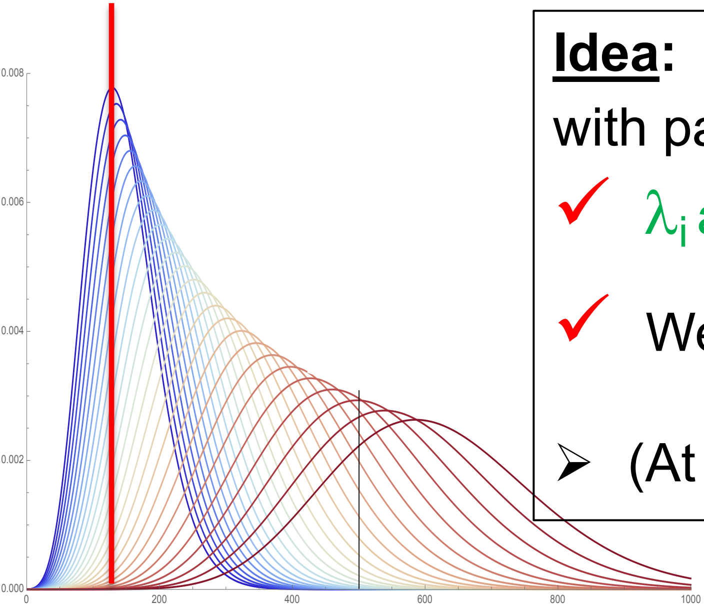
[Levin, Peres '16]
$d = 2$; $\lambda > 1$, $O(n^2)$ mixing time. $\Rightarrow$ The weighted chains for Boltzmann sampling are fast for all λ.

[Caputo, Martinelli, Toninelli '11]
$d = 3$; $\lambda > 1$, poly mixing time.

Boltzmann Sampling

Solution: Use *biased* walks:

Given λ, the Gibbs (or Boltzmann) distribution gives
a path with area k weight λ^k.



Idea: Make a polynomial number of distributions with parameters $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$ (m poly) so that:

- ✓ λ_i and λ_{i+1} are close (their distributions overlap a lot)
- ✓ We can sample efficiently from each distribution

➢ (At least) one has > 1/poly mass at n.

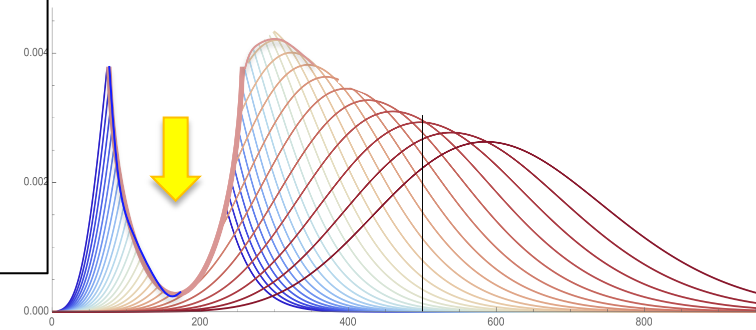
Then sample from each until there is an output a region with area n.

(Each of these area n regions are equally likely for each distribution!)

Boltzmann Sampling

We have a polynomial number of distributions with parameters $\lambda_0, \lambda_1, \dots, \lambda_{m-1}$.

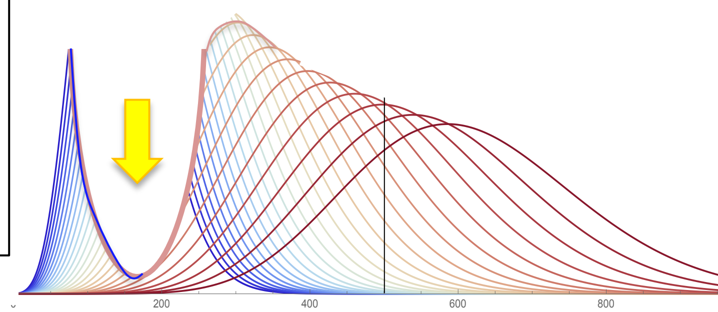
- Does (at least) one have > 1/poly mass at n?



Boltzmann Sampling

We have a polynomial number of distributions with parameters $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$.

- Does (at least) one have > 1/poly mass at n?



Yes!

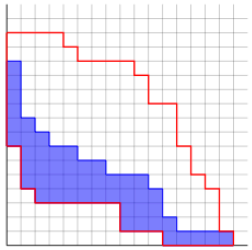
Method 1: **Thm**: $\{p(k)\}_{n=26}^{\infty}$ is logconcave. [Desalvo, Pak '14].

Thus $\{p(k)\,\lambda^k\}_{n=26}^{\infty}$ is also logconcave (and hence unimodal) for all λ.

(# partitions of size k)

Setting $\lambda = p(n)/p(n+1)$ skews the distribution so that the mode is at n.

But won't easily generalize:

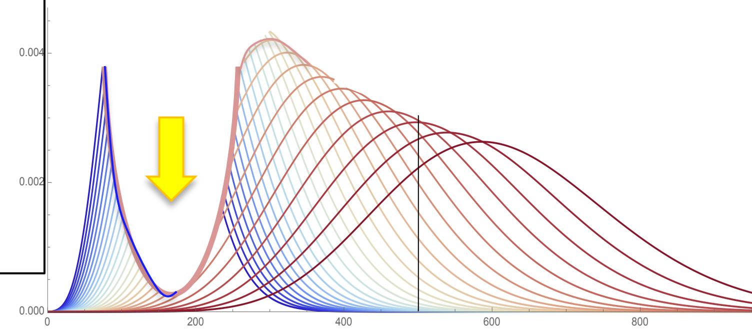


Skew partitions

Boltzmann Sampling

We have a polynomial number of distributions with parameters $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$.

➢ Does (at least) one have > 1/poly mass at n?



Yes!

Method 2 We have:

- $\| \Pi_i , \Pi_{i+1} \| > 1/2$ for all i
- Π_0 is concentrated below n.
- Π_{m-1} is concentrated above n.

So there exists j s.t. Pr [x < n] > 1/3 and Pr [x > n] > 1/3.

But the Mountain-Valley chain for bias λ_j is always rapidly mixing!

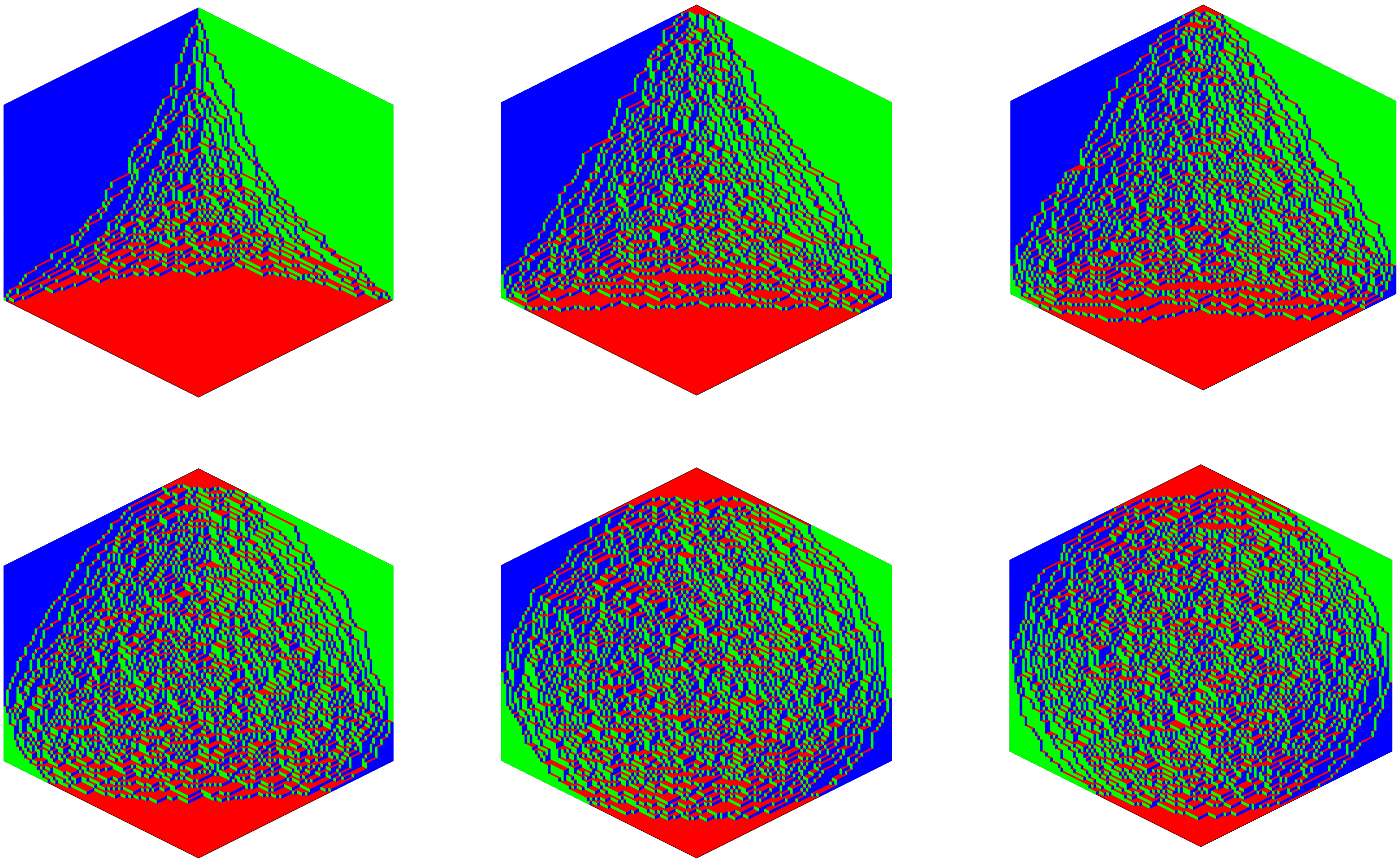
Therefore Pr [x=n] > 1/poly on that distribution. [BCFR '17]

Sampling Integer Partitions

Algorithm	Runtime	Space	Skew Partitions
Dynamic Programming [Nijinhuis-Wilf '78]	$O(n^{2.5})$	$O(n^{2.5})$	poly
Generating Function Boltzmann Sampler [Arratia-DeSalvo '15]	$\widetilde{O}(n^{0.5})$	$\widetilde{O}(n^{0.5})$	N/A
Markov Chain Monte Carlo [BCFR '17]	$O(n^{2.25})$	$\widetilde{O}(n^{0.5})$	$O(n^{16})$

Sampling lozenge tilings of fixed height

Sampling biased lozenge tilings is also fast. [Caputo, Martinelli, Toninelli ’11]



Talk Outline

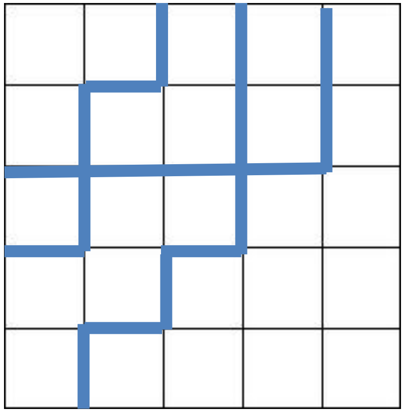
1. Sampling *unweighted* monotone surfaces and colorings

2. Sampling *biased* monotonic surfaces
   - Integer partitions of size n
   - Biased lozenge tilings

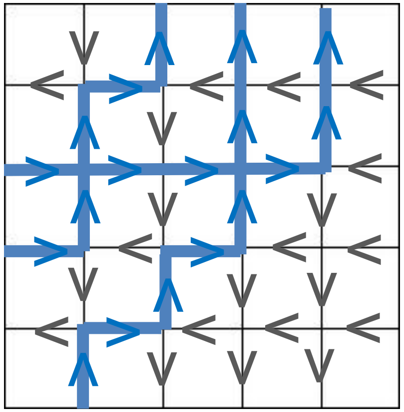
3. Sampling *weighted* 3-colorings: the “six-vertex model”
   - The antiferroelectric phase
   - The ferroelectric phase

The 6-Vertex Model

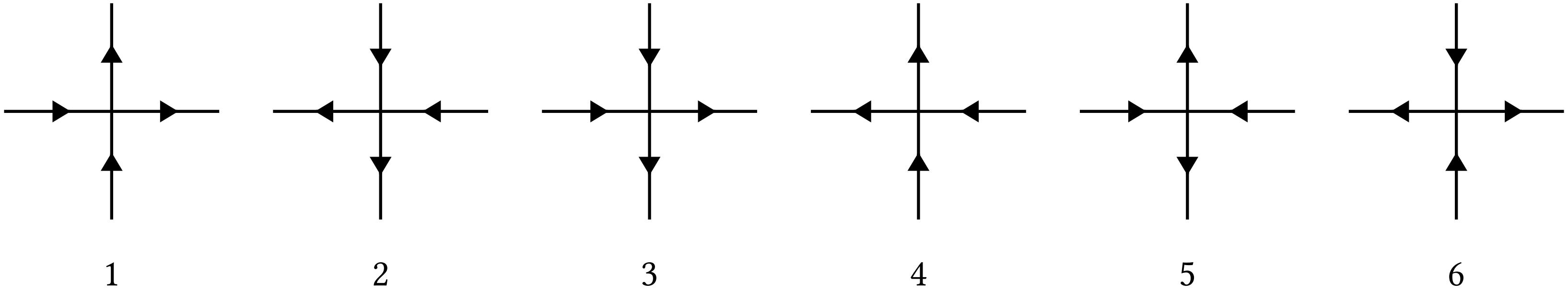
View 3-colorings of the grid (edge-disjoint paths) as an Eulerian orientation:



$\Rightarrow$

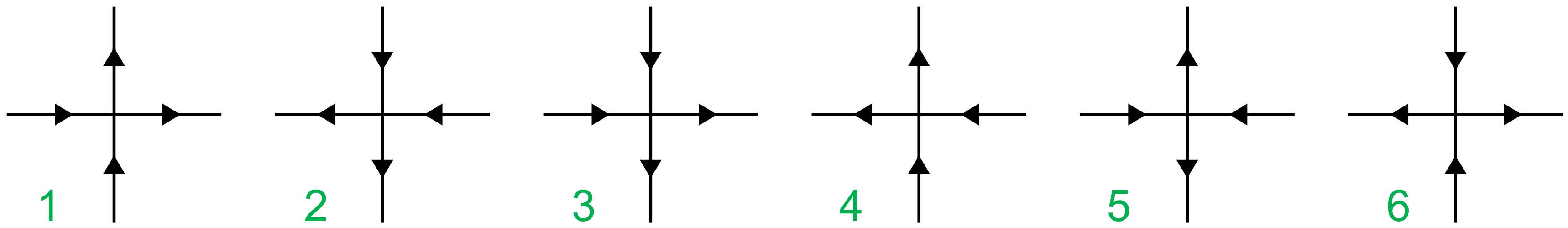


There are 6 possibilities for each internal vertex:



The 6-Vertex Model

Assign the Boltzmann weights $w_1, w_2, \ldots, w_6 > 0$ to the 6 types:



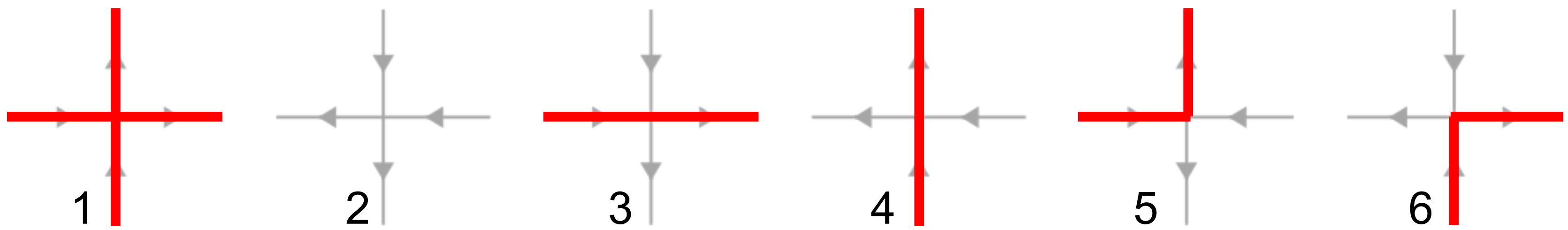
The weight of a configuration $x \in \Omega$ is $\pi(x) = \prod_{i=1}^{6} w_i^{n_i(x)} / Z$,

where $n_i(x)$ is the number of vertices in x of type w_i

and Z is the partition function.

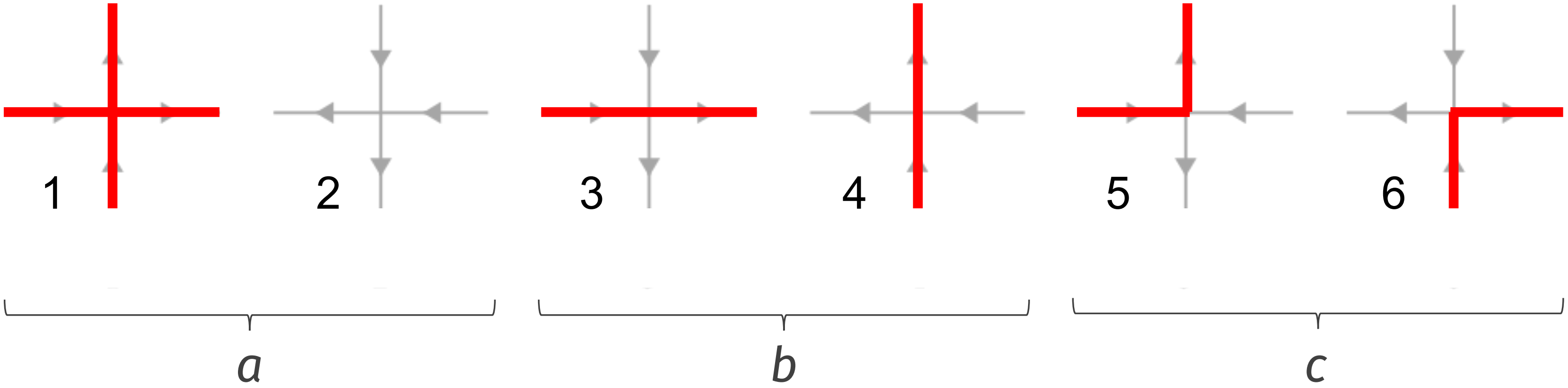
The 6-Vertex Model

In the paths representation, keeping only up and right arrows, we have:



The 6-Vertex Model

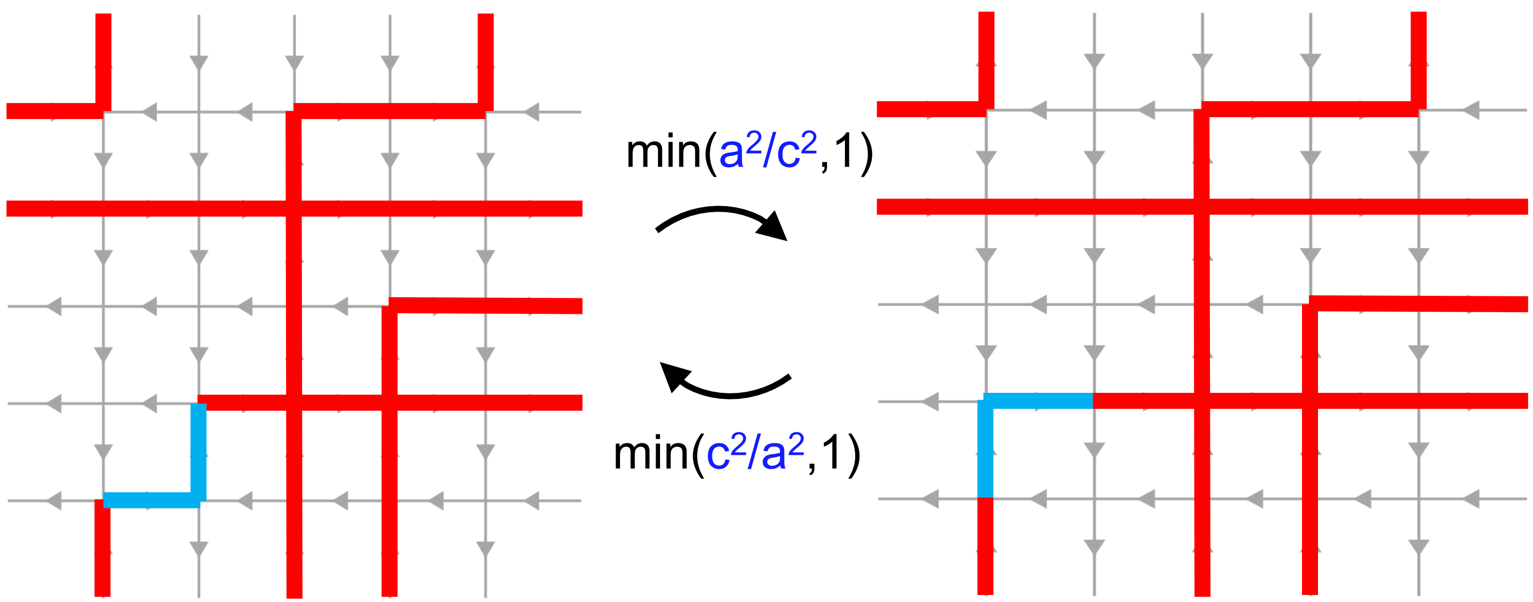
In the paths representation, keeping only up and right arrows, we have:



Due to invariants, there are really only 2 parameters: a/c and b/c.

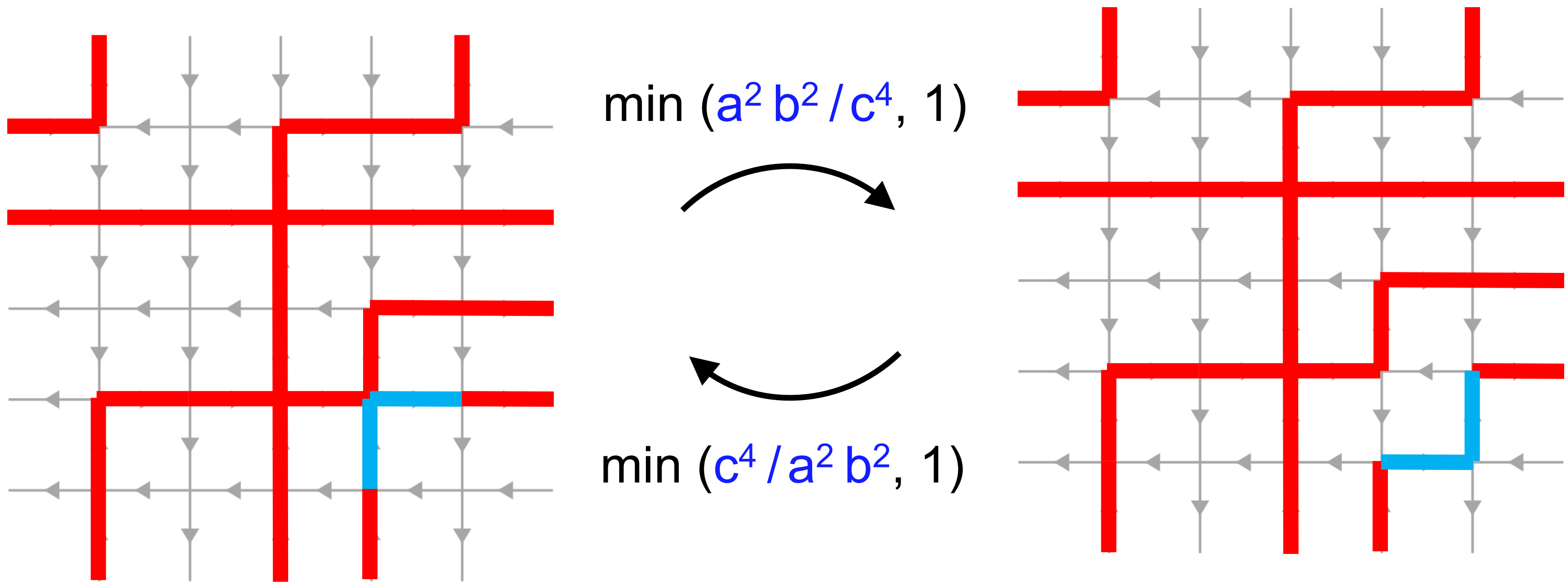
Glauber Dynamics

Glauber dynamics picks a cell and reverses all 4 edges if they form an oriented cycle (*with Metropolis transition probs).*



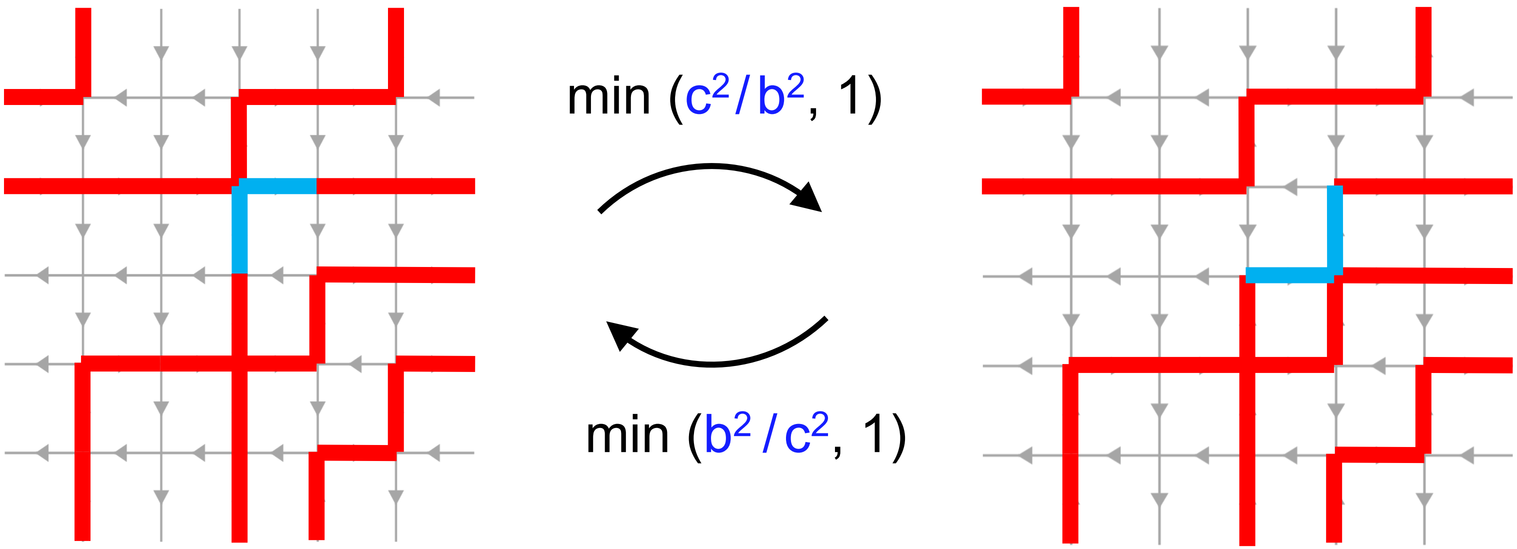
Glauber Dynamics

Glauber dynamics picks a cell and reverses all 4 edges if they form an oriented cycle. (with Metropolis transition probs).



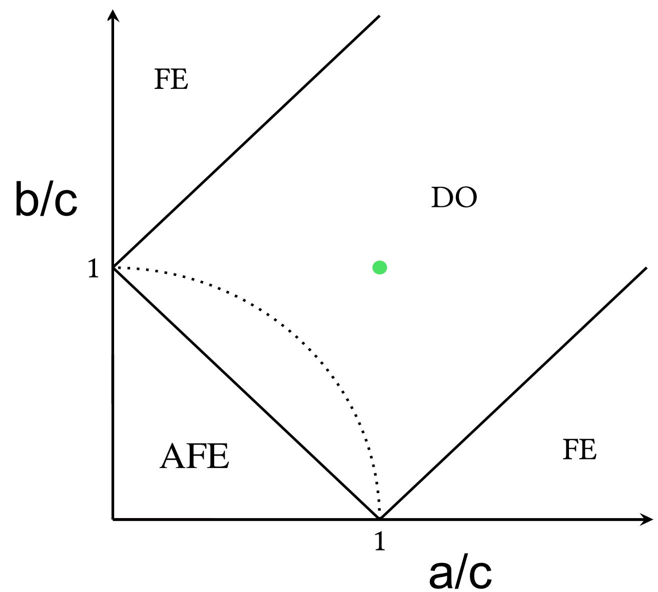
Glauber Dynamics

Glauber dynamics picks a cell and reverses all 4 edges if they form an oriented cycle.   (with Metropolis transition probs).



The Phase Diagram

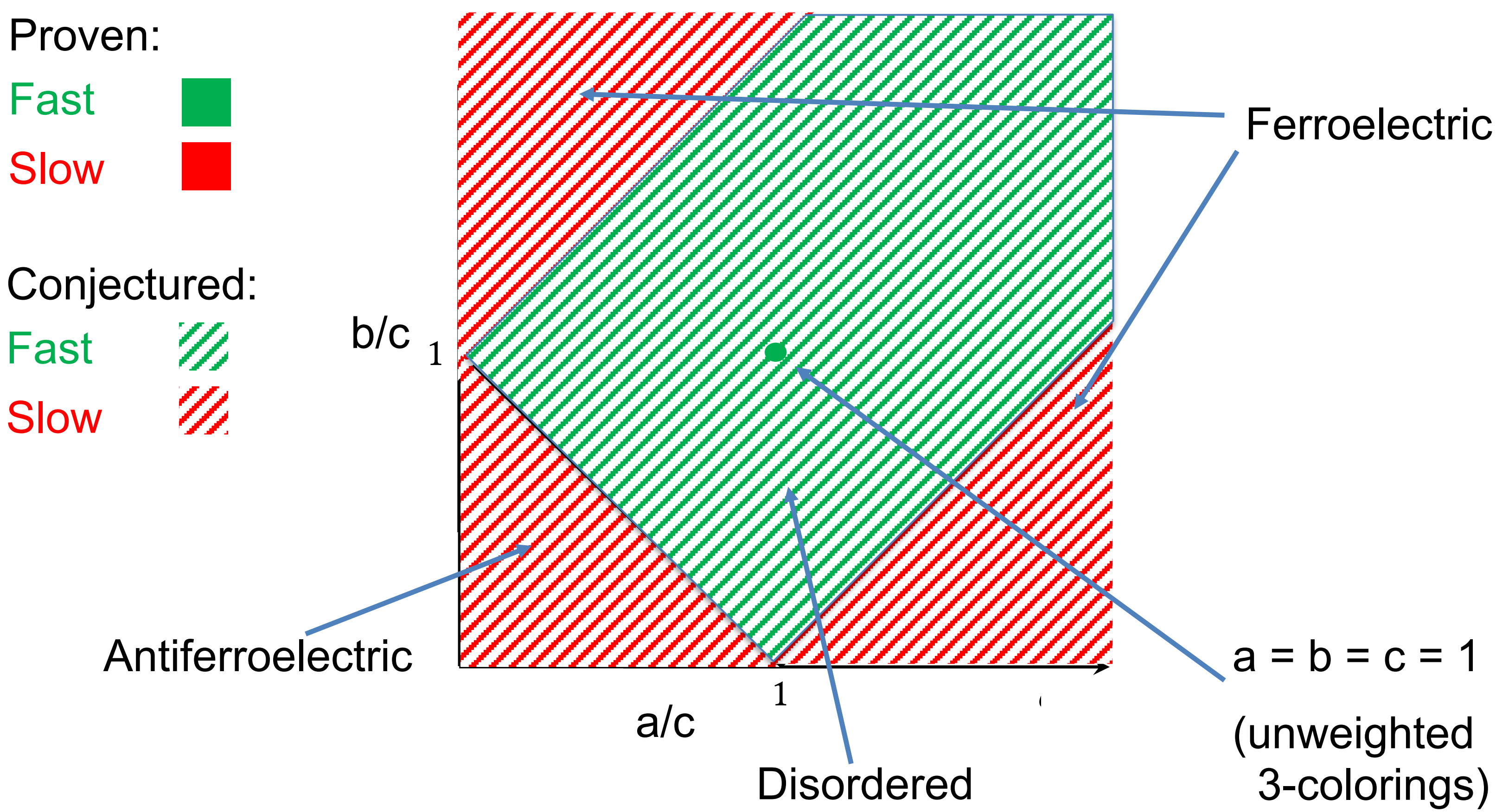
The (conjectured) phase diagram from physics:



- The configuration is ferroelectric if $a > b+c$ or $b > a + c$.
- The configuration is antiferroelectric if $c > a + b$.
- Otherwise the configuration is disordered.

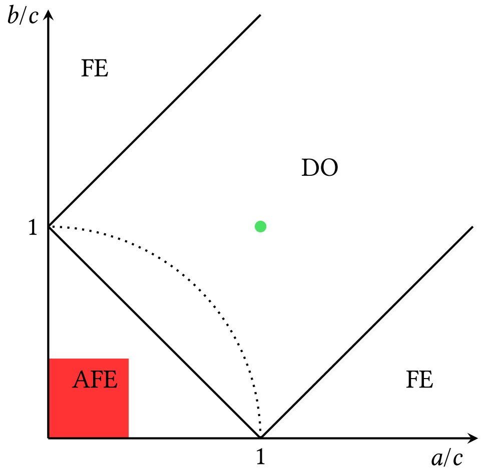
The Phase Diagram

The (conjectured) phase diagram from physics:

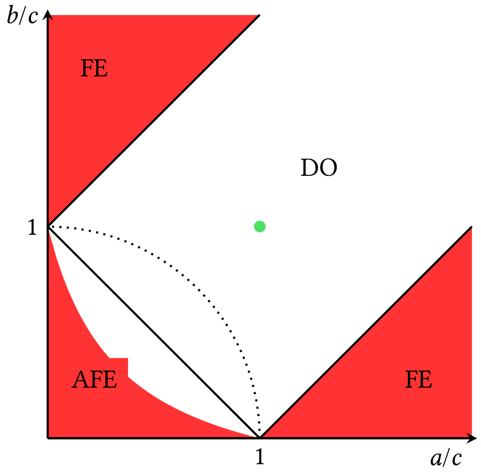


The Phase Diagram: Results

Proven results:



[Liu ‘18]



[Fahrbach, R. ‘19]

Results

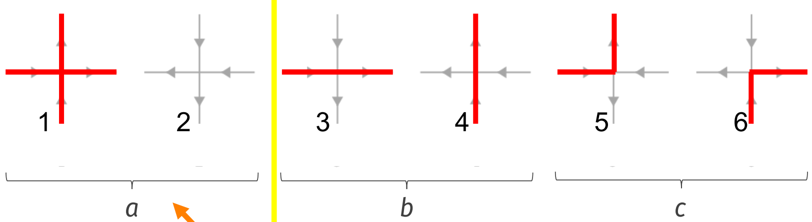
[Fahrbach, R. 19]

<u>Thm</u> (Ferroelectric): If $a > b + c$ or $b > a + c$ then there exist boundary conditions for which Glauber dynamics mixes exponentially slowly.

<u>Thm</u> (Antiferroelectric): If $ac + bc + 3ab < c^2$ then Glauber dynamics with free boundary conditions mixes exponentially slowly.

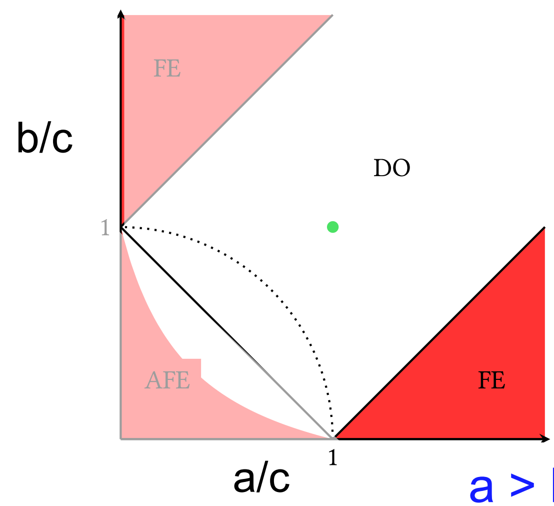
The Phase Diagram: Results

Proven results:



Favoring intersections

over straights or turns



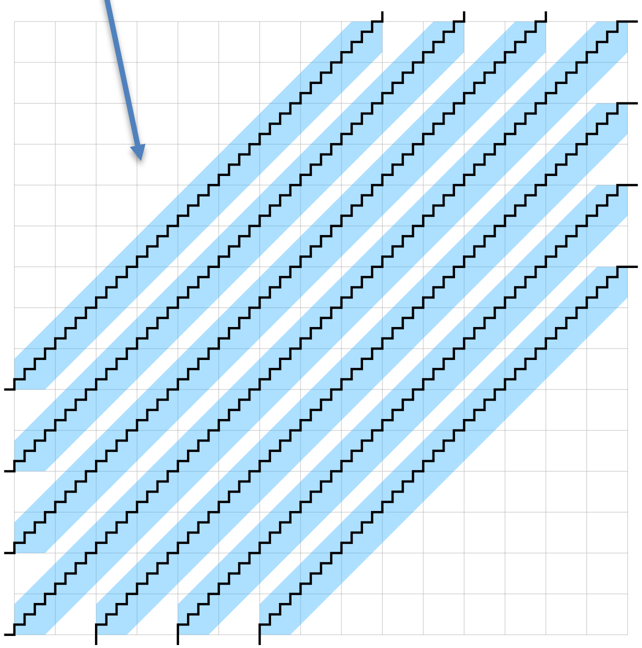
$a > b + c$

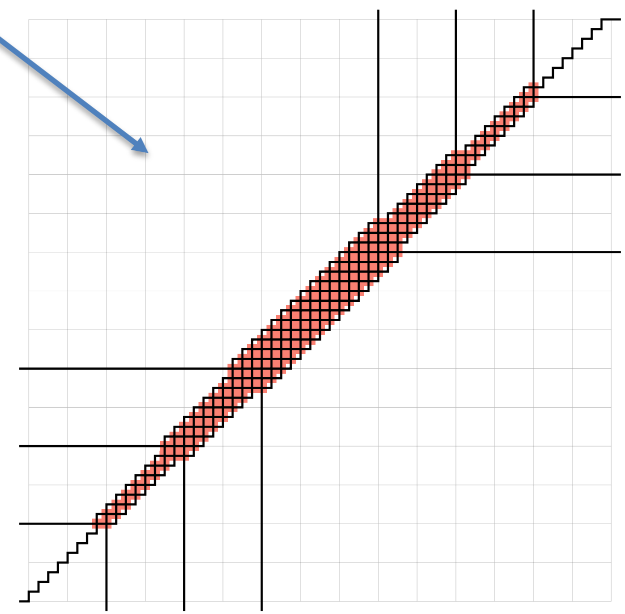
[Fahrbach, R. '19]

The Ferroelectric Region

<u>Thm</u> (Ferroelectric): If $a > b + c$ or $b > a + c$ then there exist boundary conditions for which Glauber dynamics mixes exponentially slowly.

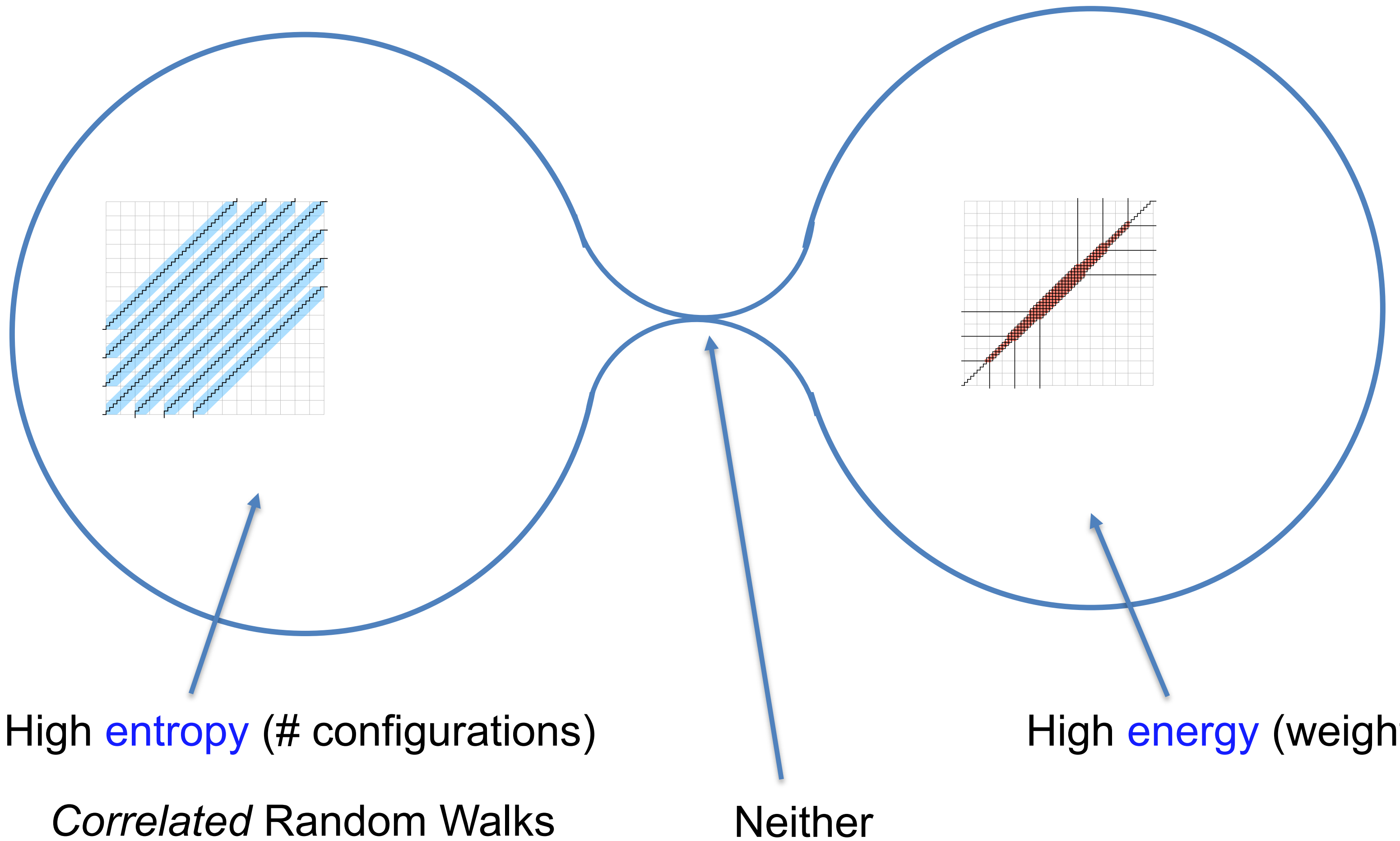
- Induce well-separated paths from boundary conditions.
- Adjust parameters so that the ground state has large weight.





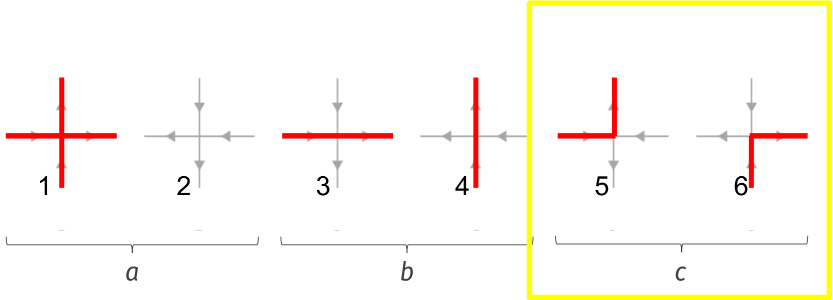
A Bad Cut in Ω

Thm (Ferroelectric): If a> b+c or b > a+c then there exist boundary conditions for which Glauber dynamics mixes exponentially slowly.



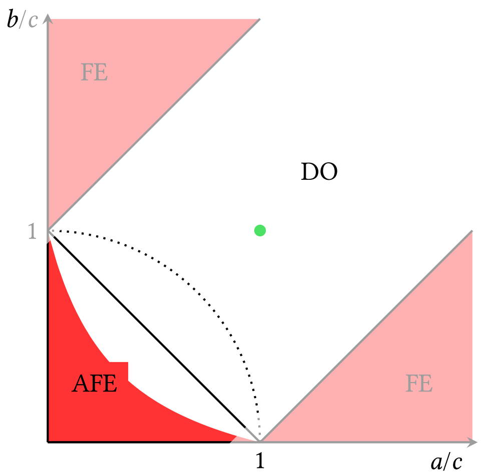
The Antiferroelectric Region

Proven results:



Favoring turns edges

over intersections or straights

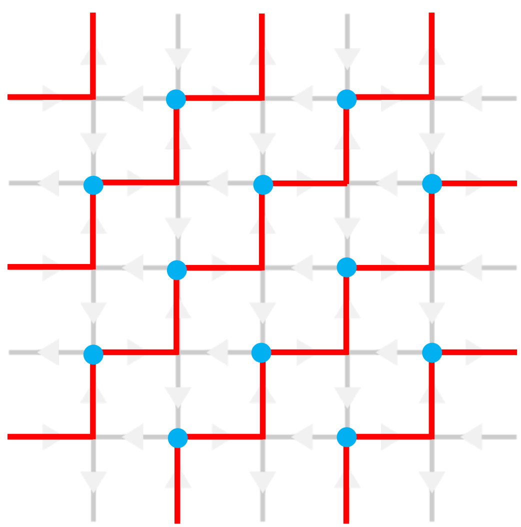


[Fahrbach, R. '19]

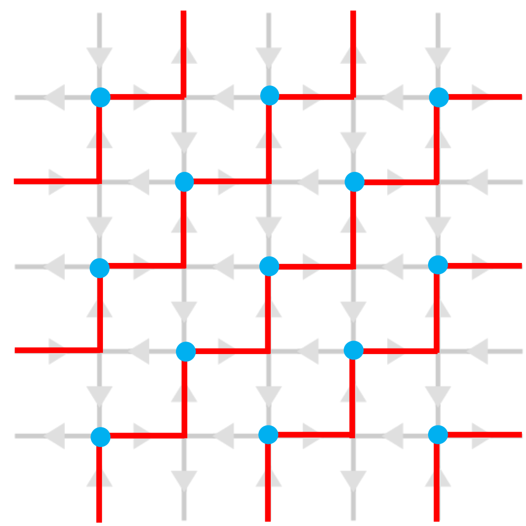
The Antiferroelectric Region

Thm (Antiferroelectric): If $ac + bc + 3ab < c^2$ then Glauber dynamics with free boundary conditions mixes exponentially slowly.

Main idea: for large values of c, the ground state behaves like the Independent set model at low temperature.



Odd vertices

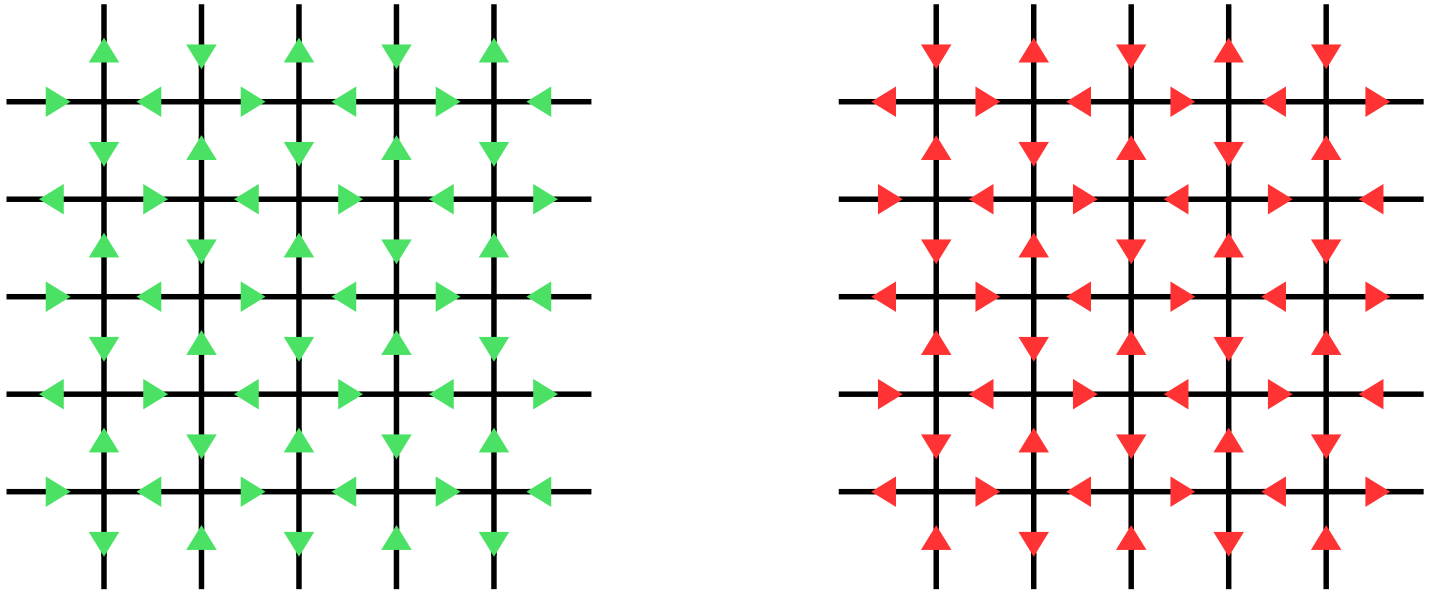


Even vertices

The Antiferroelectric Region

Thm (Antiferroelectric): If $ac + bc + 3ab < c^2$ then Glauber dynamics with free boundary conditions mixes exponentially slowly.

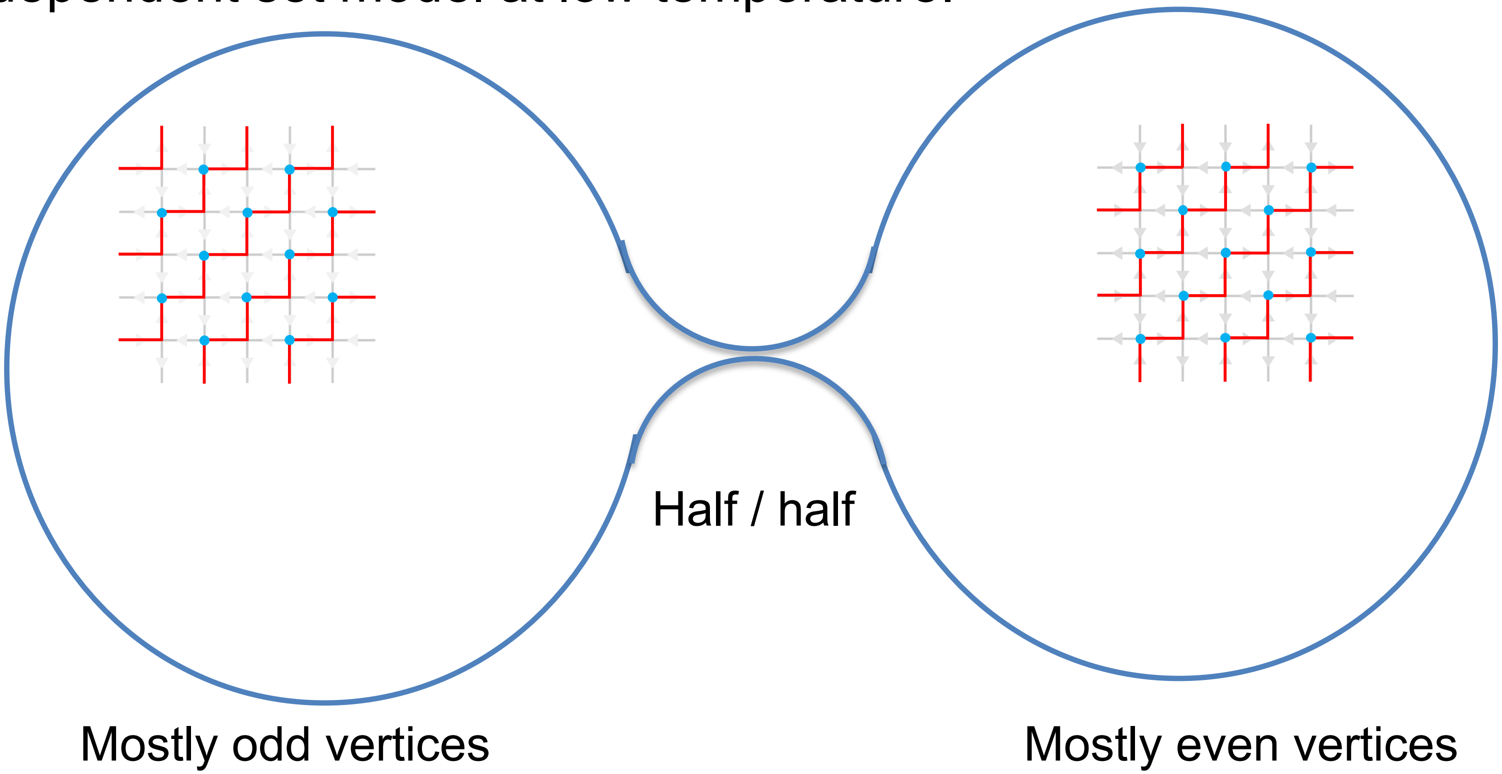
Main idea: for large values of c, the ground state behaves like the Independent set model at low temperature.



The Antiferroelectric Region

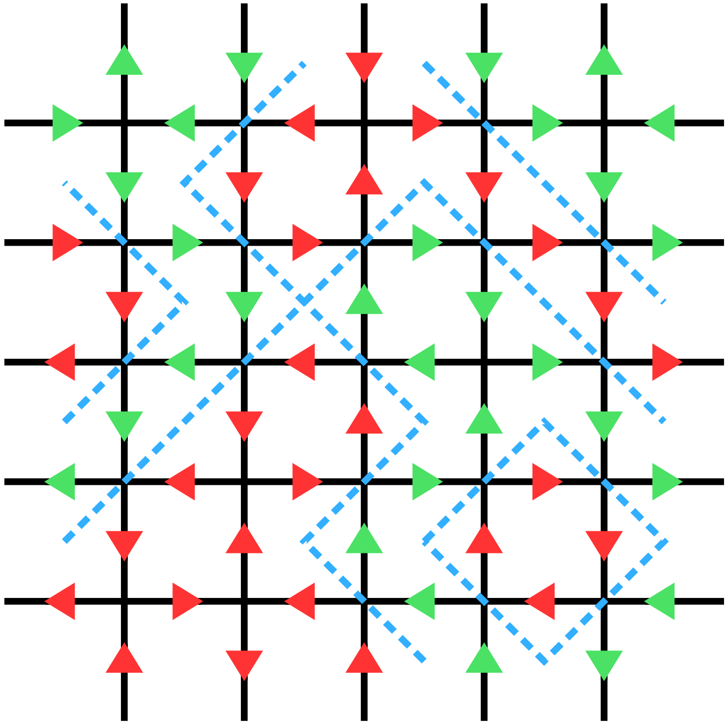
Thm (Antiferroelectric): If $ac + bc + 3ab < c^2$ then Glauber dynamics with free boundary conditions mixes exponentially slowly.

Main idea: for large values of c, the ground state behaves like the Independent set model at low temperature.



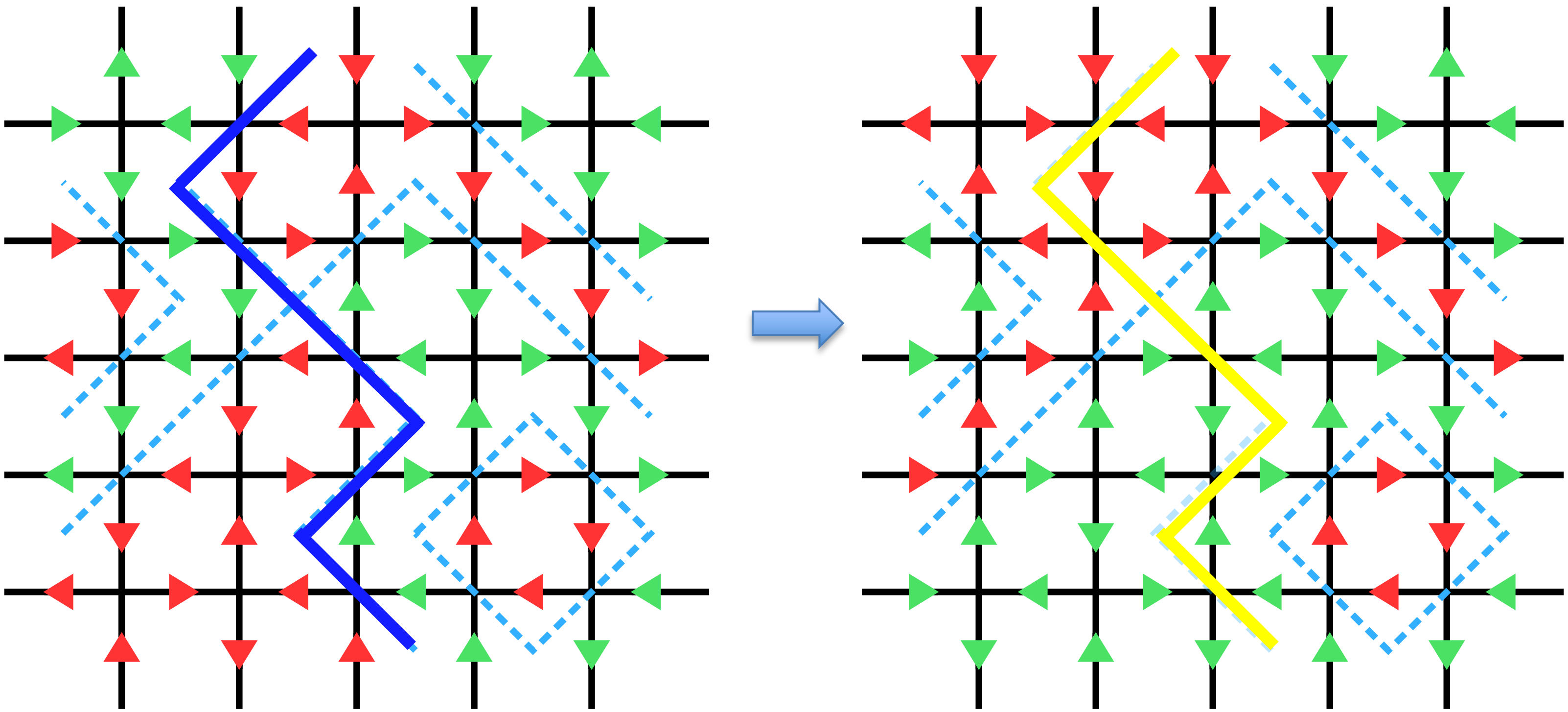
The Antiferroelectric Region

Use a Peierls Argument: alter “cut” configurations to increase weight exponentially s.t. the information needed to undo the map is a smaller exponential.



Peierls Argument

Use a Peierls Argument: alter "cut" configurations to increase weight exponentially s.t. the information needed to undo the map is a smaller exponential.

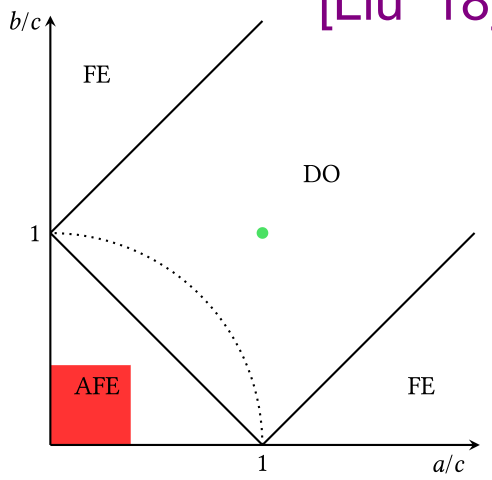


For each point on the "fault line", we gain weight c/a or c/b.

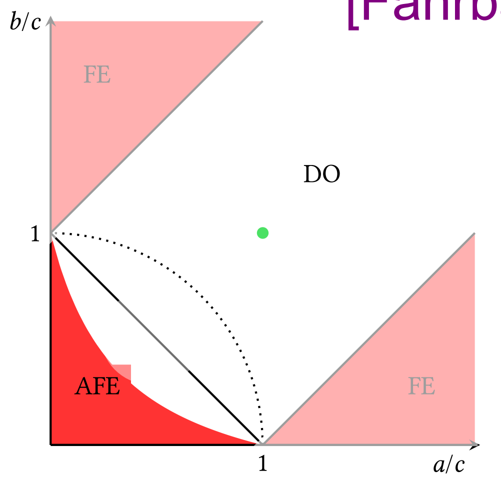
The Phase Diagram: Results

Proven results:

[Liu '18]



[Fahrbach, R. '19]



- c/a, b/c > μ^n

  (SAW connective const.)

- Count *non-backtracking walks*
- Use a generating function to allow a or b to be big if the other is small

Recap and Open Questions

1. Sampling *unweighted* monotone surfaces and colorings

   Higher dimensions?

2. Sampling *biased* monotonic surfaces

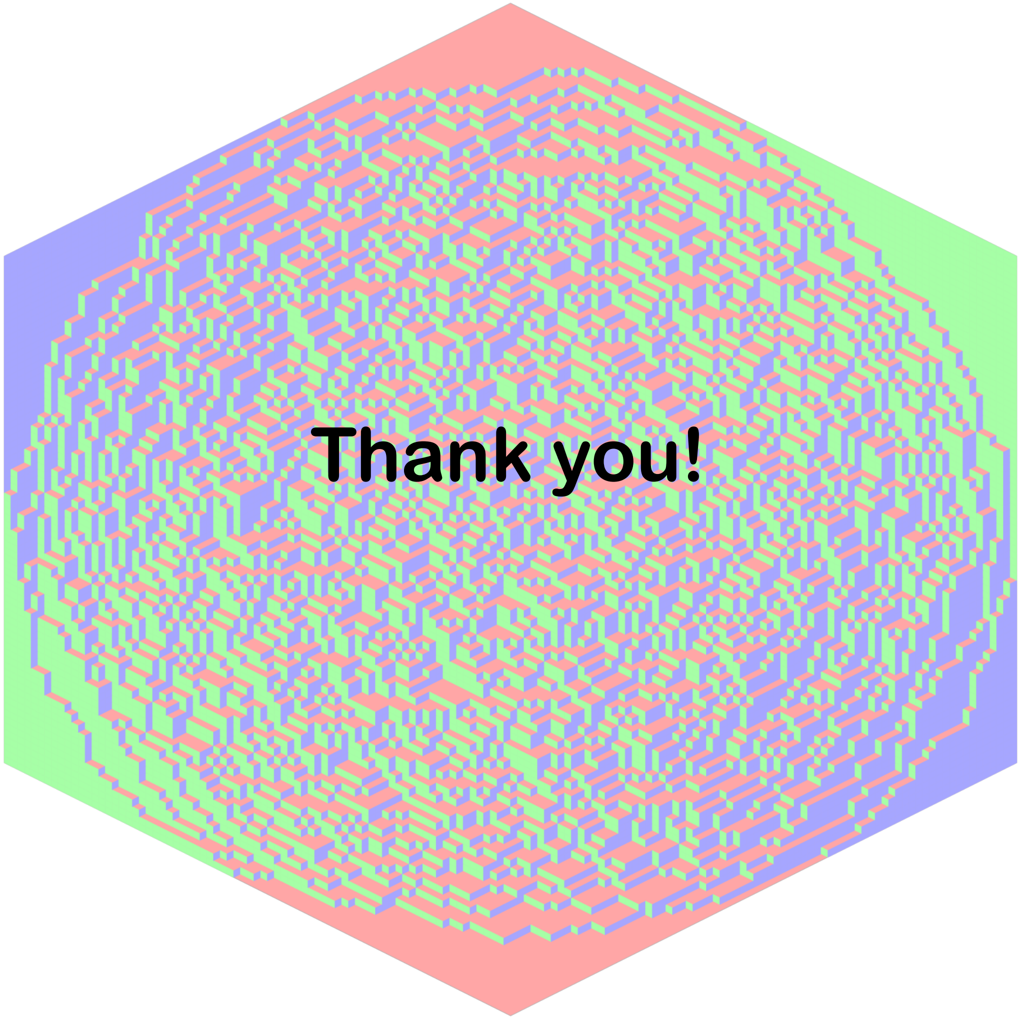
   - Integer partitions of size n
   - Biased lozenge tilings

   Biased 3-colorings in high dimensions?

3. Sampling *weighted* 3-colorings: the “six-vertex model”

   - The antiferroelectric phase
   - The ferroelectric phase

   The disordered phase?


Thank you!