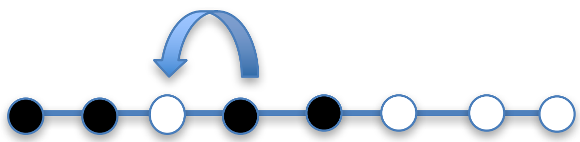
Sampling Paths, Permutations and Lattice Structures

Dana Randall

Georgia Institute of Technology

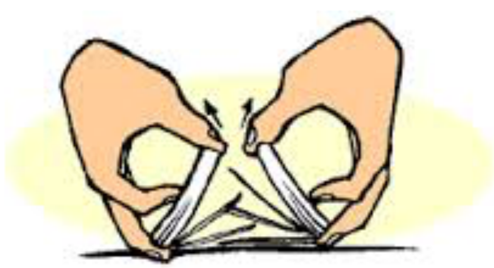
Sampling Problems



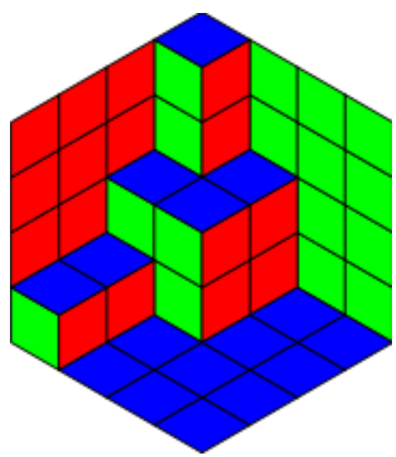
Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

Integer Partitions



Card Shuffling

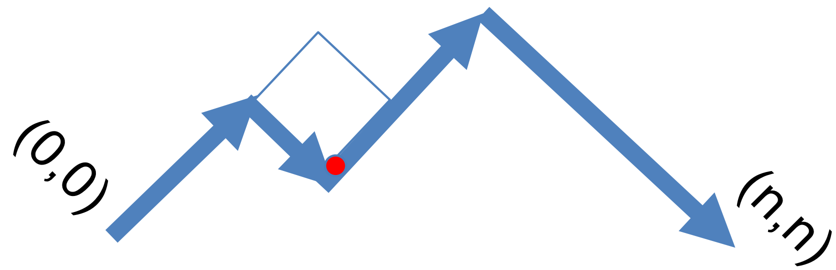


Lozenge Tilings

0	2	0	1	2
2	0	2	0	1
0	1	0	1	0
1	0	1	0	2
0	2	0	2	1

3-colorings

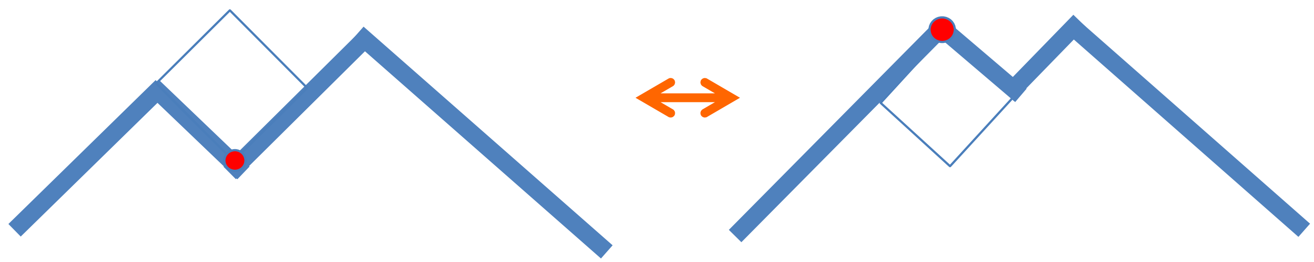
Lattice Paths



State space: All monotonic lattice paths from (0,0) to (n,n).

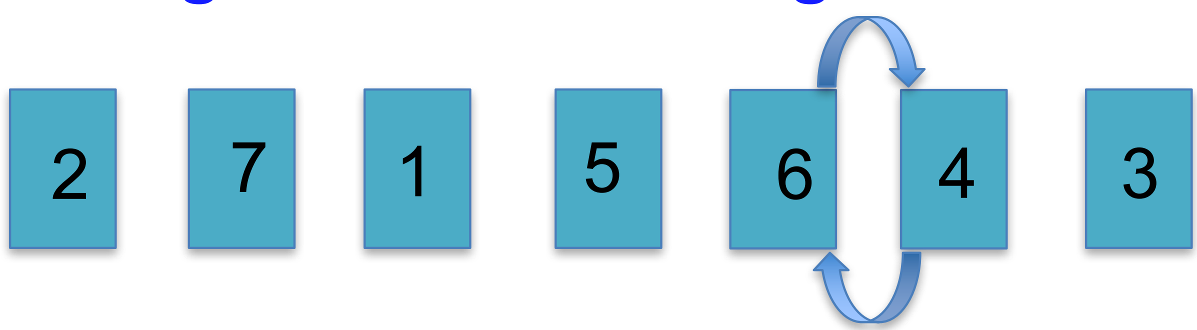
Local dynamics: Switch between "mountains" and "valleys".

Lattice Paths



Simple Exclusion Processes

Card Shuffling with Nearest Neighbor Transpositions

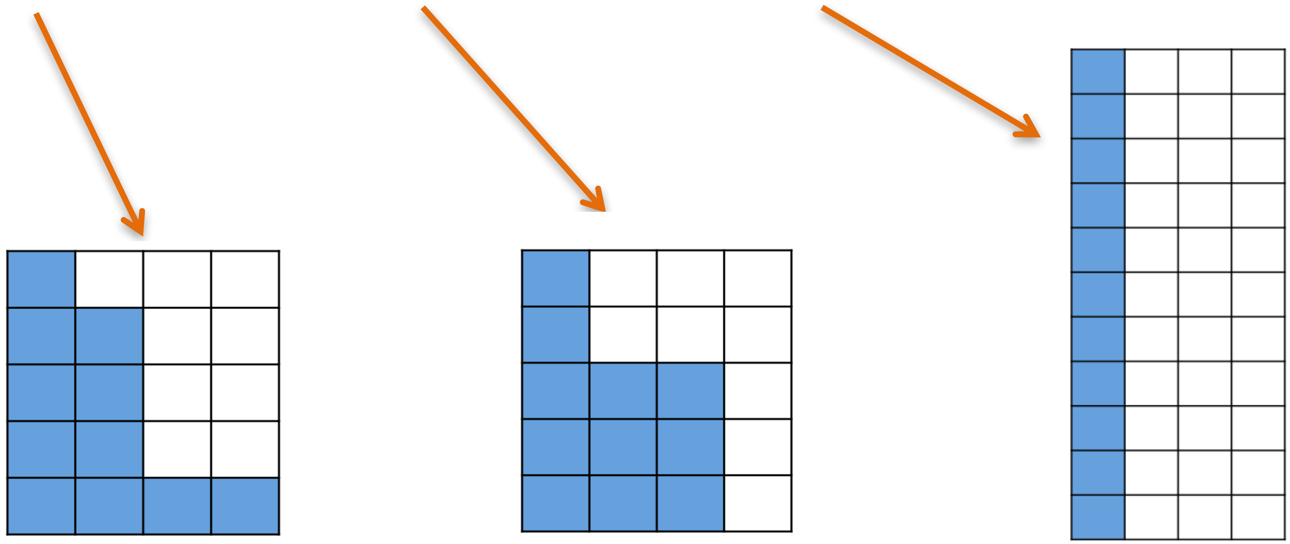


0	1	0	0	0	0	0
0	1	0	0	1	0	0
0	1	0	1	1	0	0
0	1	0	1	1	1	0
0	1	0	1	1	1	1
1	1	0	1	1	1	1

Integer Partitions

An Integer Partition of n is a sum of positive integers where order doesn't matter   (4 + 6 is the same as 6 + 4).

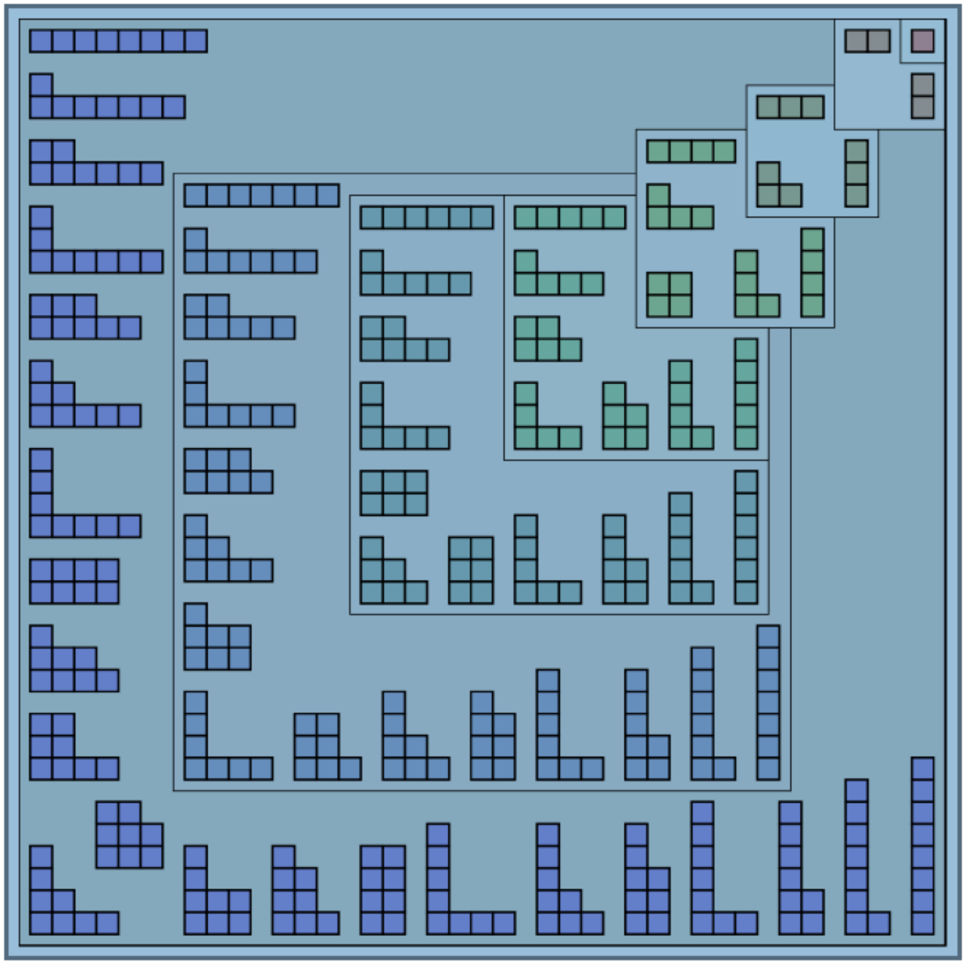
Ex: 1 + 1 + 4 + 5,   3 + 3 + 5, and 11 are partitions of 11.



Integer Partitions

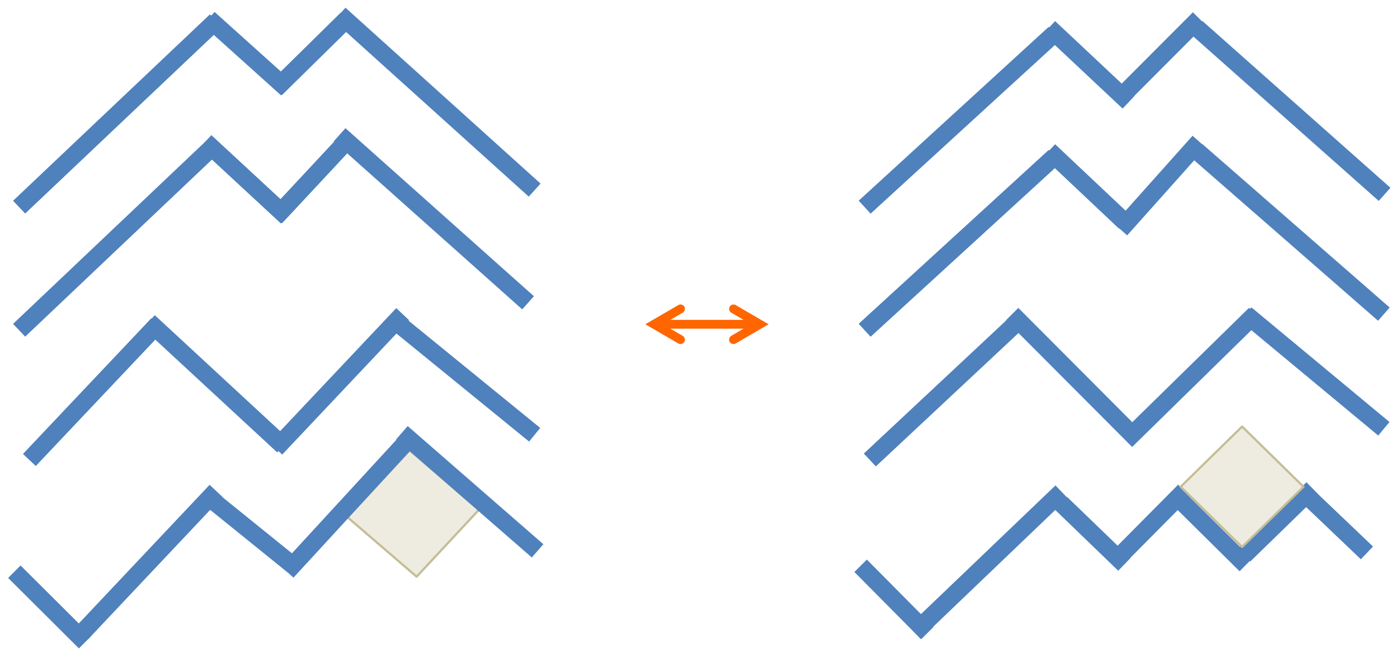
Ferrers (Young) Diagrams:

Each piece of the partition is represented as an ordered "stack" of squares.

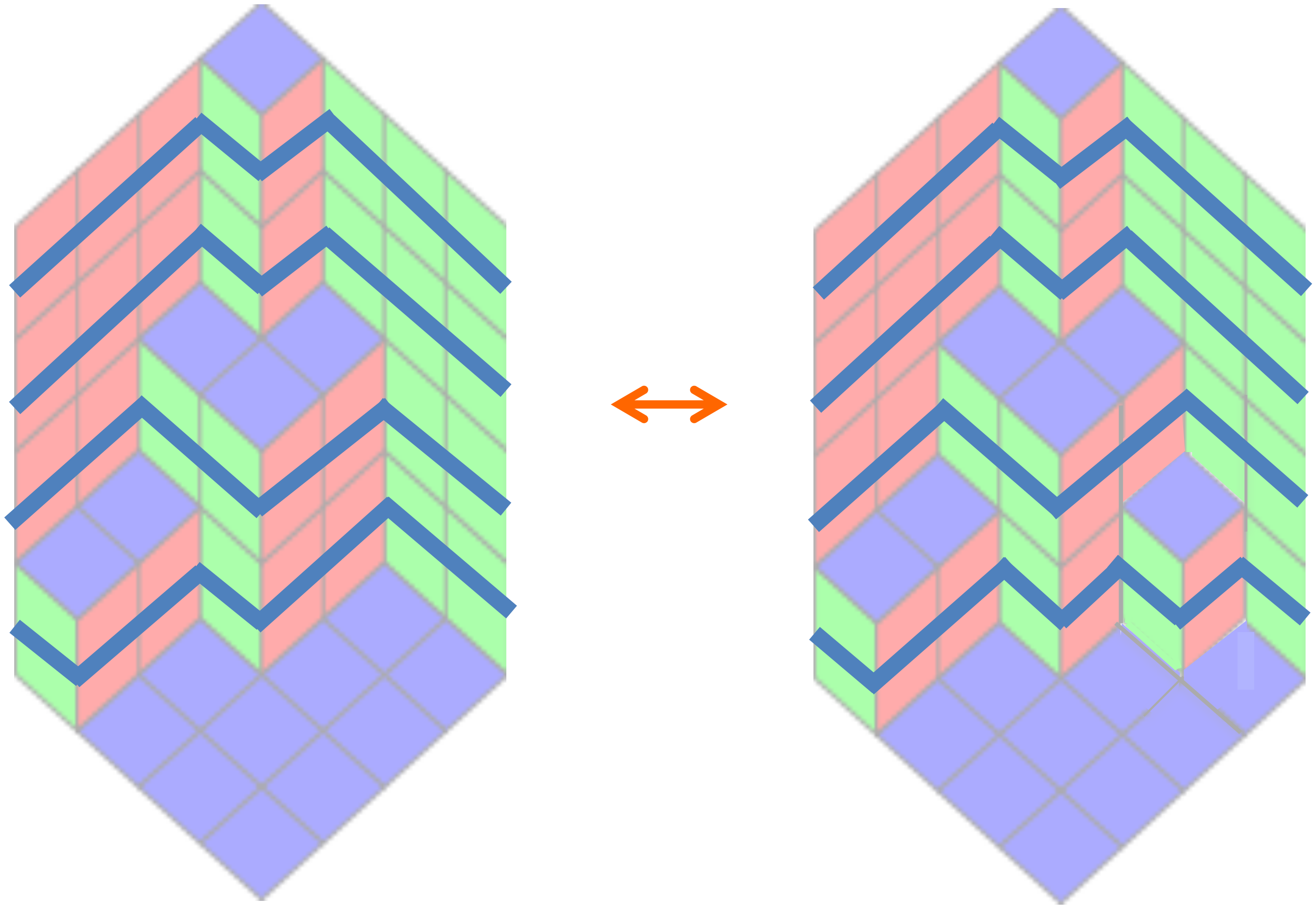


Sampling integer partitions of n is the same as sampling lattice paths bounding regions of area n.

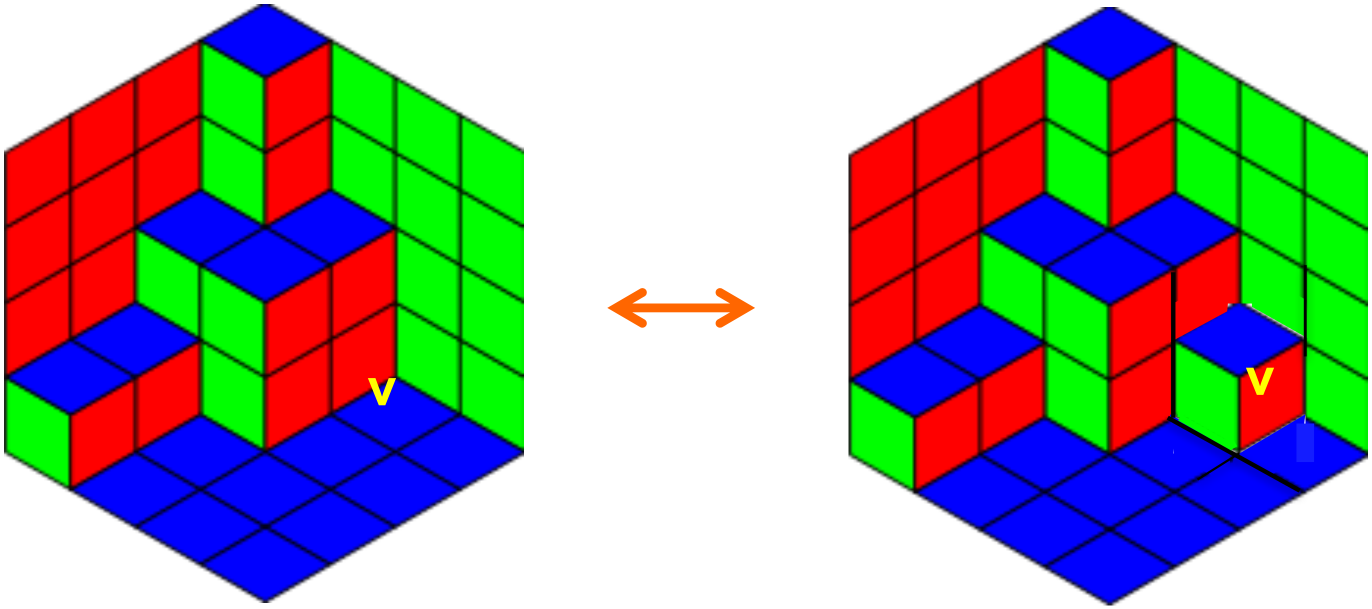
Multiple Nonintersecting Paths



Multiple Nonintersecting Paths



Vertex Disjoint Paths = Lozenge Tilings

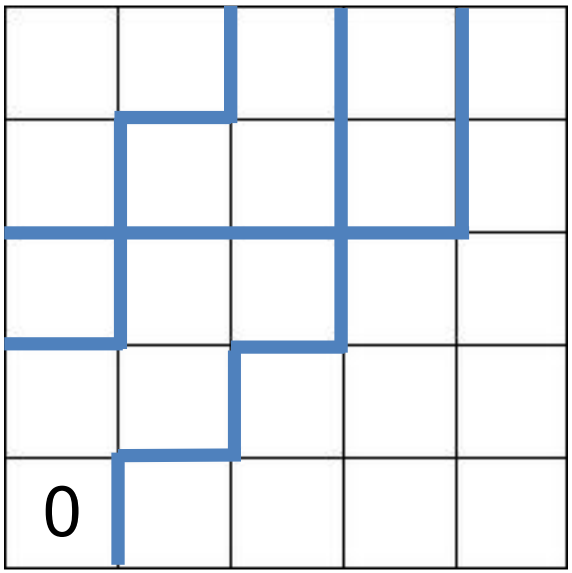


Repeat:

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

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

Or, If ***Edge*** Disjoint…



<u>Crossing path:</u>

D, R: +1 (mod 3)

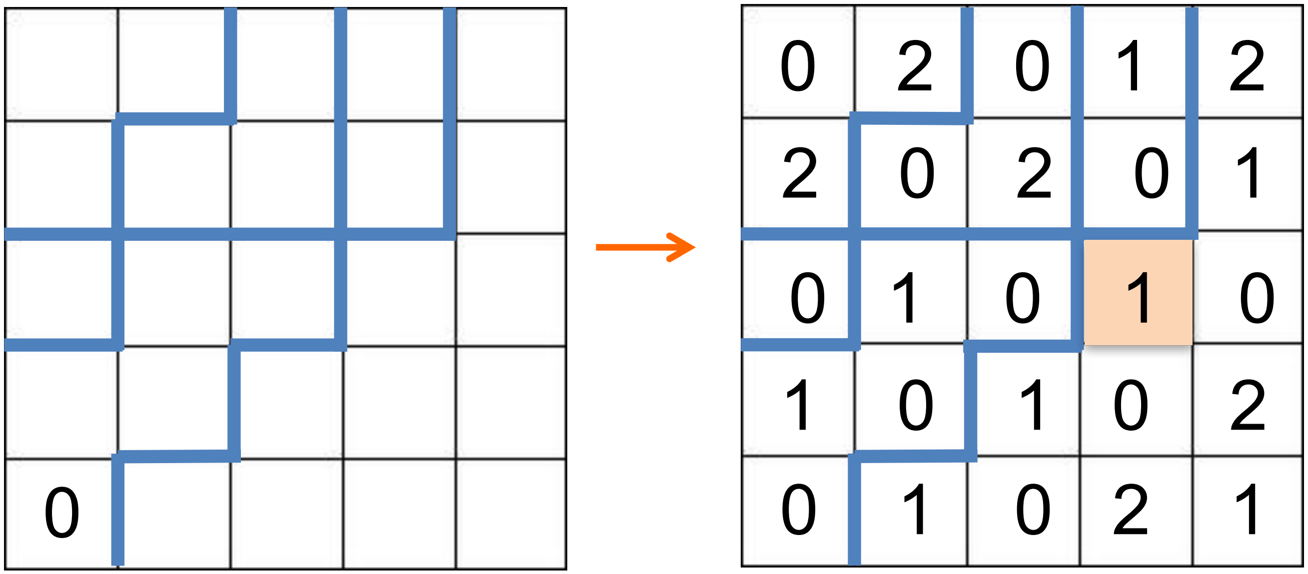
U, L: -1

<u>No path:</u>

D, R: -1

U, L: +1

3-Colorings (or Eulerian Orientations)



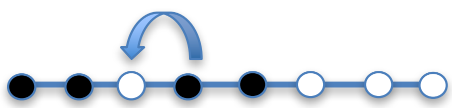
Crossing path:
D, R: +1 (mod 3)
U, L: -1

No path:
D, R: -1
U, L: +1

There is a bijection between edge disjoint lattice paths and proper **3-colorings of Z^2** (and the "**6-vertex model**").

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

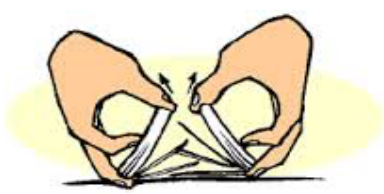
Q: How do we sample lattice paths?



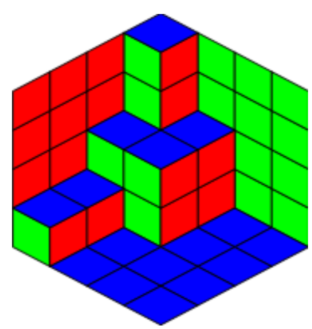
Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

Integer Partitions



Card Shuffling



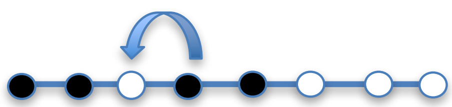
Lozenge Tilings

0	2	0	1	2
2	0	2	0	1
0	1	0	1	0
1	0	1	0	2
0	2	0	2	1

3-colorings

Outline

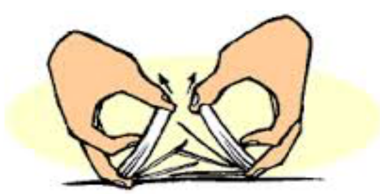
- Sampling paths uniformly
- Paths with uniform bias
- Paths with non-uniform bias



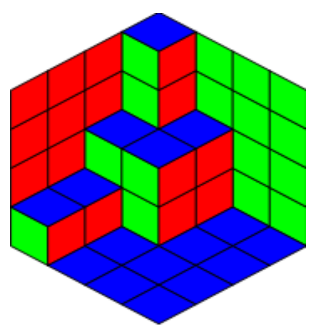
Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

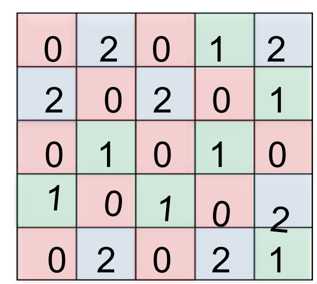
Integer Partitions



Card Shuffling



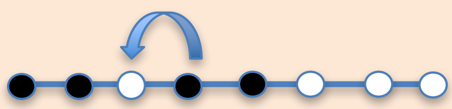
Lozenge Tilings



3-colorings

Outline

- Sampling paths uniformly: **One path**
- Paths with uniform bias
- Paths with non-uniform bias

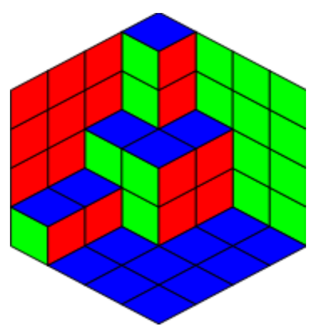


Simple Exclusion Processes

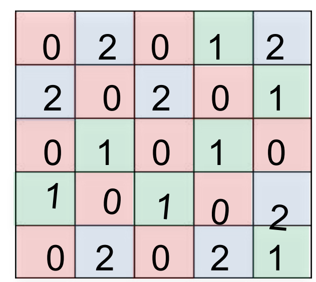
1 + 1 + 4 + 5 = 11

Integer Partitions

Card Shuffling

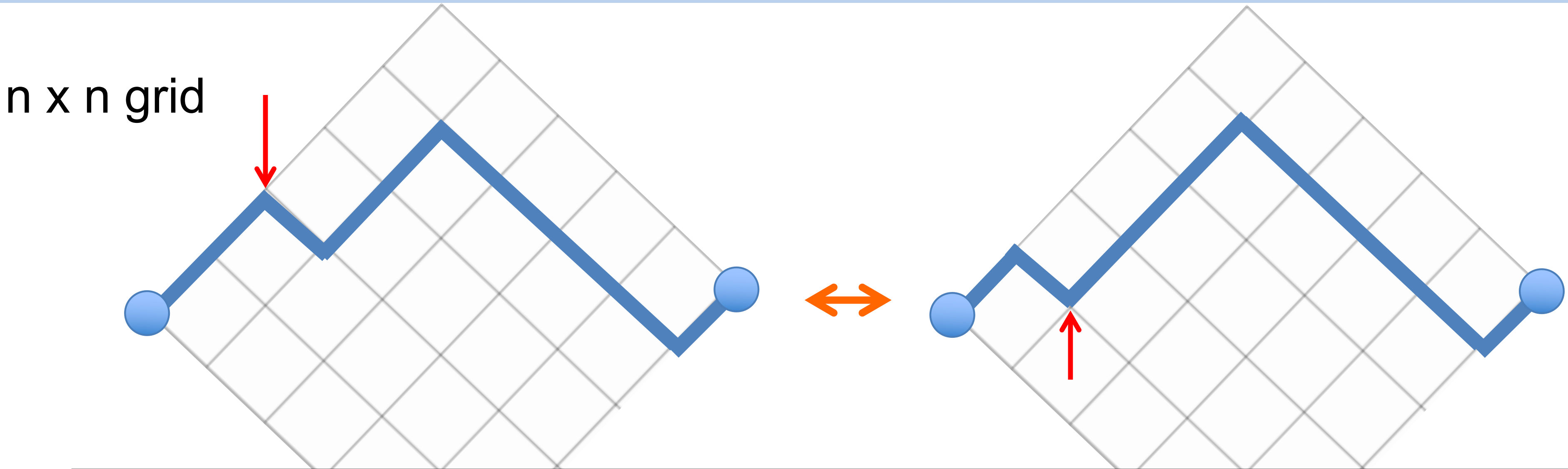


Lozenge Tilings



3-colorings

Lattice Paths in $\mathbf{Z}^2$



To sample, repeat:

- Pick **v** on the path;
- If **v** is a mountain/valley, invert w.p. ½ (if possible).

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]

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).

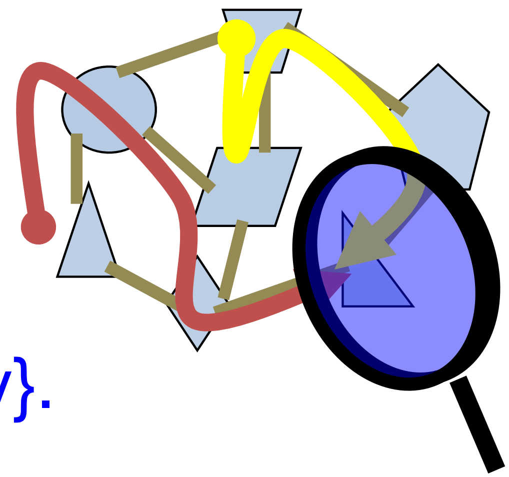
Coupling

Definition: A coupling is a MC on $\Omega \times \Omega$:

1) Each process $\{X_t\}$, $\{Y_t\}$ is a faithful copy of the original MC;
2) If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$.

The coupling time T is:

$$T = \max_{x,y} (E [T^{x,y}]),$$



where $T^{x,y} = \min \{t: X_t=Y_t \mid X_0=x, Y_0=y\}$.

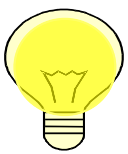
Thm: $\tau(\varepsilon) \leq T \, e \ln \varepsilon^{-1}$. [Aldous]

Path Coupling

Coupling: Show for all $x,y \in \Omega$, $E[\Delta(\mathrm{dist}(x,y))] \leq 0$.

Path coupling: Show for all u,v s.t. $\mathrm{dist}(u,v)=1$, that

$E[\Delta(\mathrm{dist}(u,v))] \leq 0$. [Bubley, Dyer, Greenhill]



Consider a shortest path:

$$x = z_0, \quad z_1, \quad z_2, \quad . \; . \; . \; , \quad z_r = y,$$

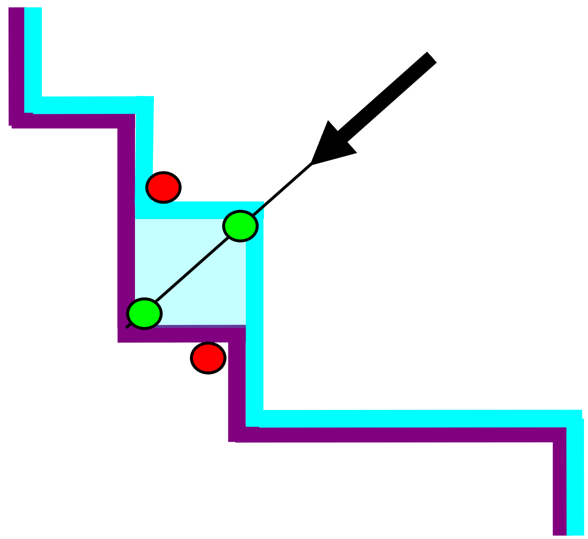
$\mathrm{dist}(z_i, z_{i+1}) = 1$,

$\mathrm{dist}(x,y) = r$.

$$\Rightarrow \quad E[\Delta(\mathrm{dist}(x,y))] \leq \sum_i E[\Delta(\mathrm{dist}(z_i, z_{i+1}))]$$

$$\leq 0.$$

Coupling the Unbiased Chain



Coupling:

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

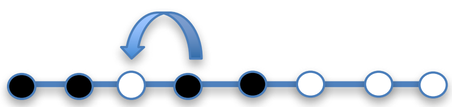
The **distance** Ψ_t at time t is the unsigned area between the two configurations.

- $E[\Delta(\Psi_t)] = p\,(-\#G + \#B) \leq 0$
- Var > 0 if $\Psi_t > 0$;
- $0 \leq \Psi_t \leq n^2$;
- $\Psi_t = 0$ implies $\Psi_{t+1} = 0$.

Then the paths couple quickly, so the MC is rapidly mixing.

Outline

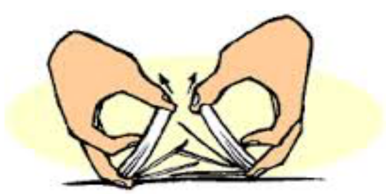
- Sampling paths uniformly: **Multiple paths**
- Paths with uniform bias
- Paths with non-uniform bias



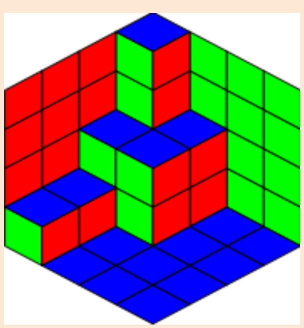
Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

Integer Partitions



Card Shuffling

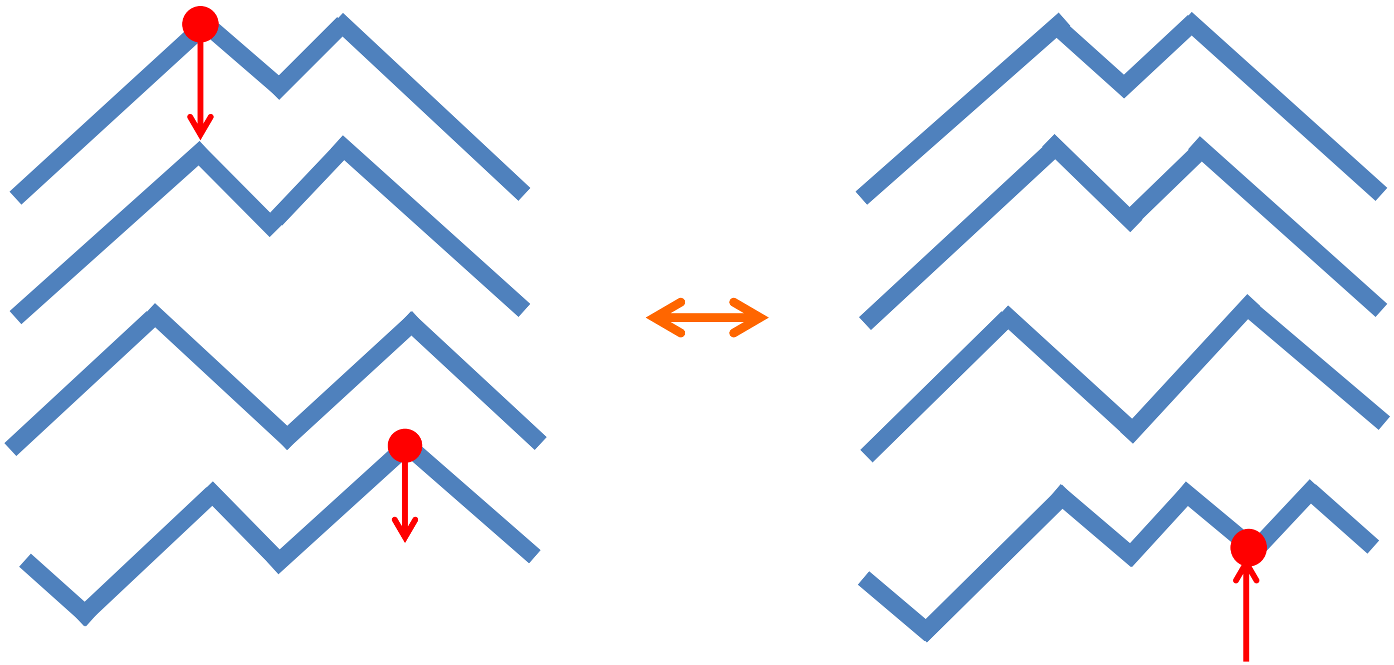


Lozenge Tilings

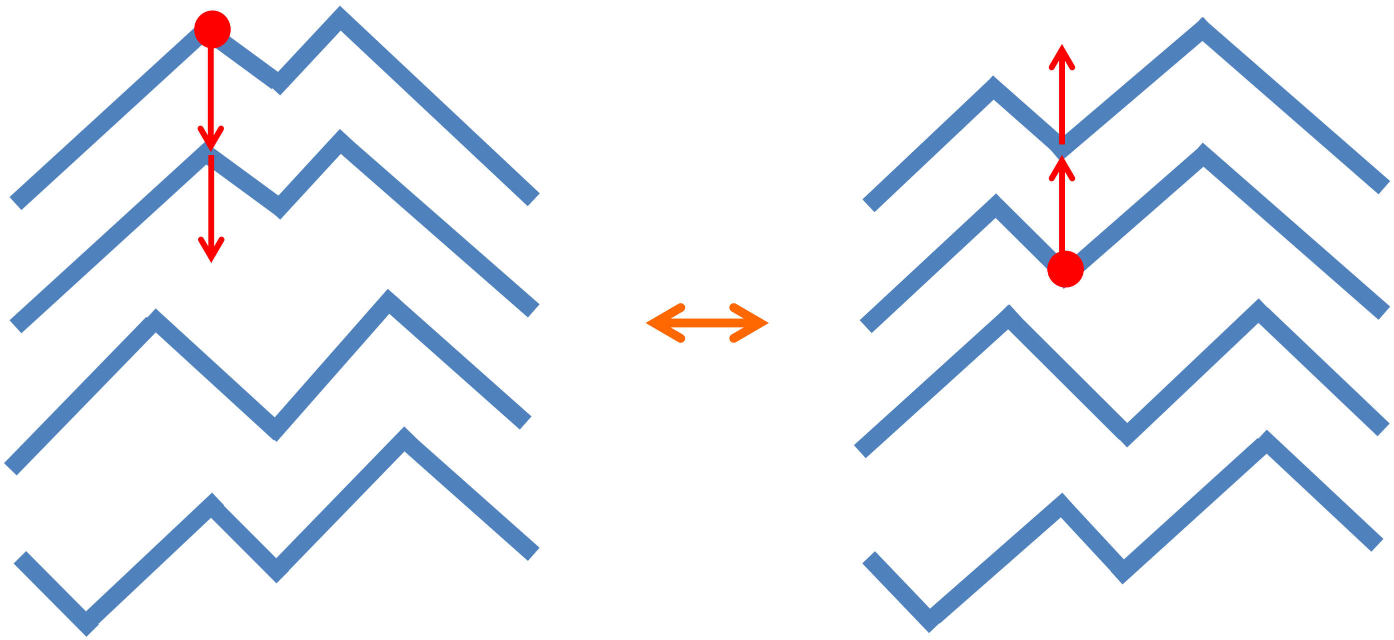
0	2	0	1	2
2	0	2	0	1
0	1	0	1	0
1	0	1	0	2
0	2	0	2	1

3-colorings

Markov chain for Lozenge Tilings



“Tower chain” for Lozenge Tilings



The “tower chain” [Luby, R., Sinclair]

Also couples (and mixes) quickly for lozenge tilings (and similarly for 3-colorings).

Higher Dimensions?

When does the MC converge in poly time on $\mathbf{Z}^d$?

Lattice paths, lozenge tilings and "space partitions" in $\mathbf{Z}^d$:

d=2: Yes (simple coupling)

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

d≥4: ???

3-colorings of $\mathbf{Z}^d$:

d=2: Yes (simple coupling)

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

d=4: ???

Higher Dimensions?

When does the MC converge in poly time on $\mathbf{Z}^d$?

Lattice paths, lozenge tilings and "space partitions" in $\mathbf{Z}^d$:

d=2: Yes (simple coupling)

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

d≥4: ???

3-colorings of $\mathbf{Z}^d$:

d=2: Yes (simple coupling)

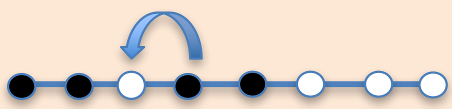
d=3: Yes [LRS], [Martin, Goldberg, Patterson], [R., Tetali]

d=4: ???

d>>1: No! [Galvin, Kahn, R., Sorkin], [Peled]

Outline

- Sampling paths uniformly
- Paths with uniform bias
- Paths with non-uniform bias

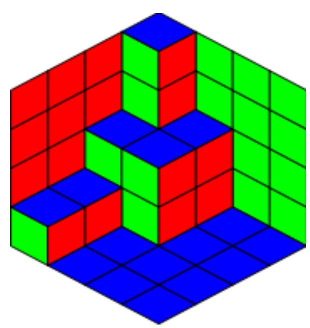


Simple Exclusion Processes

$1 + 1 + 4 + 5 = 11$

Integer Partitions

Card Shuffling



Lozenge Tilings

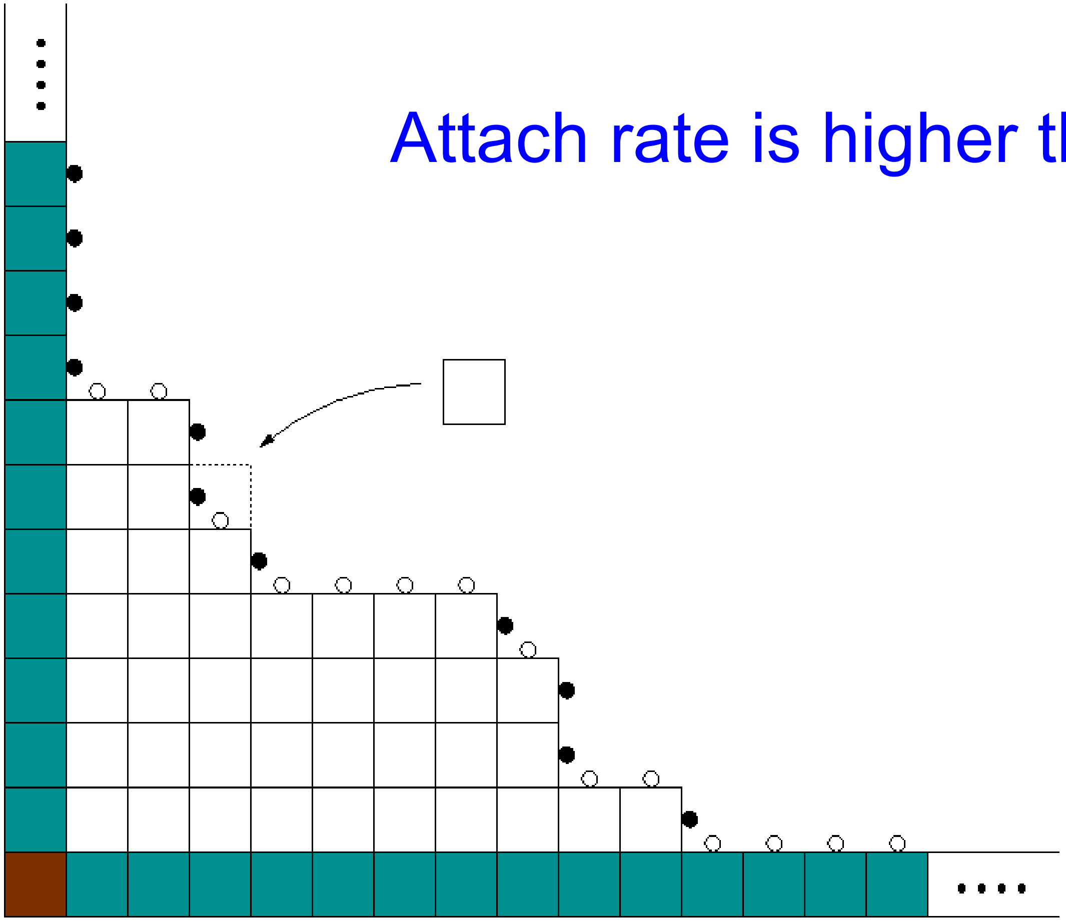
0	2	0	1	2
2	0	2	0	1
0	1	0	1	0
1	0	1	0	2
0	2	0	2	1

3-colorings

Lattice Paths with **Uniform Bias**

<u>Tile-based self-assembly</u> (a growth model):

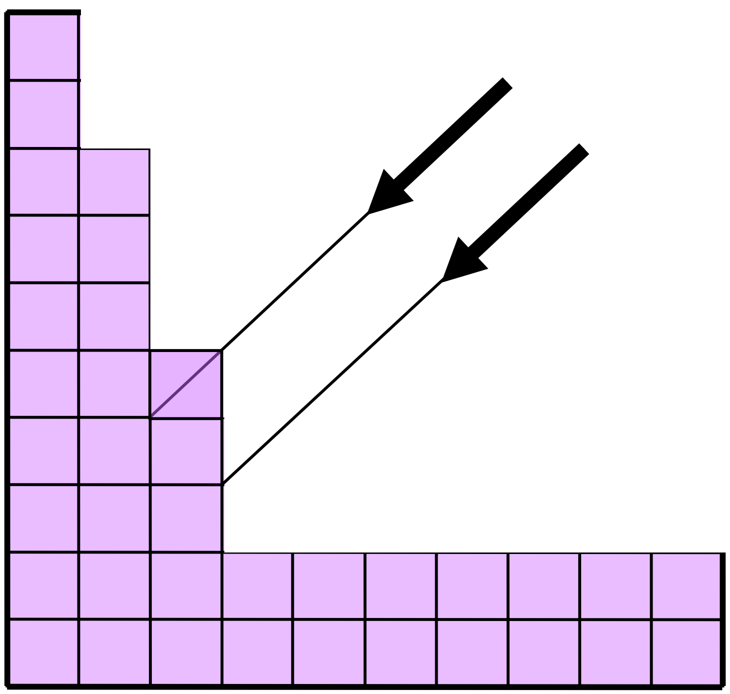
A tile can attach if 2 neighbors are present and detach if 2 neighbors are missing.



Attach rate is higher than the detach rate.

Generating Biased Surfaces

Given $\lambda > 1$:



Repeat:

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

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

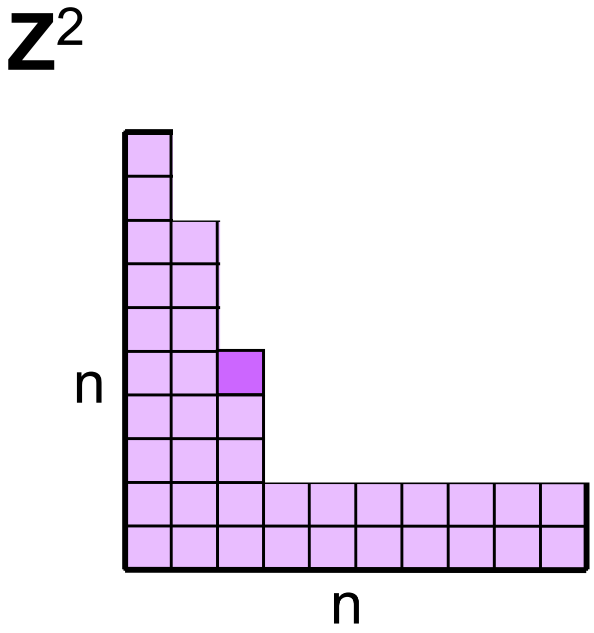
If a square can be removed at v,
and d=-, remove it w.p. λ^{-1};

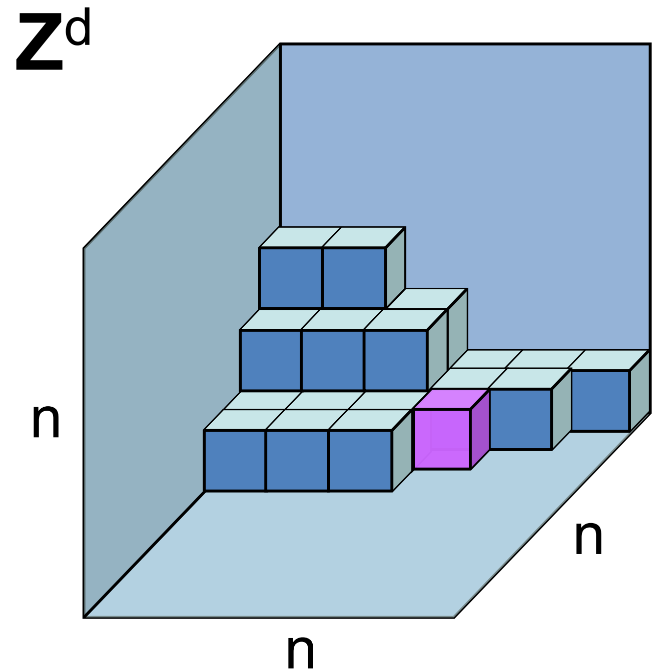
Otherwise do nothing.

Converges to the distribution:

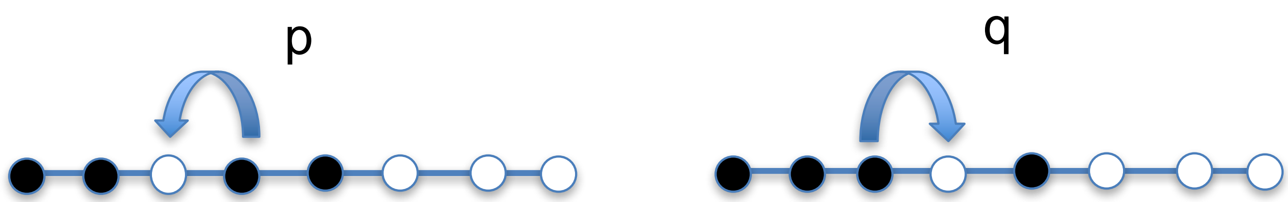
$$\pi(S) = \lambda^{area(S)} / Z.$$

Generating Biased Surfaces





ASEPs: Asymmetric Simple Exclusion Process:



How fast?

Biased Surfaces in Z^d

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

[Benjamini, Berger, Hoffman, Mossel]

$d = 2$; $\lambda > 1$ const, $O(n^2)$ mixing time (optimal).

Biased Surfaces in Z^d

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

[Benjamini, Berger, Hoffman, Mossel]

$d = 2$; $\lambda > 1$ const, $O(n^2)$ mixing time (optimal).

[Majumder, Sahu, Reif]

$d = 2,3$; $\lambda = \Theta(n)$, poly time.

Biased Surfaces in Z^d

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

[Benjamini, Berger, Hoffman, Mossel]

$d = 2$; $\lambda > 1$ const, $O(n^2)$ mixing time (optimal).

[Majumder, Sahu, Reif]

$d = 2,3$; $\lambda = \Theta(n)$, poly time.

[Greenberg, Pascoe, R.]

$d = 2$, $\lambda > 1$ const
$d \geq 3$, $\lambda > d^2$
} $O(n^d)$ mixing time.

Biased Surfaces in Z^d

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

[Benjamini, Berger, Hoffman, Mossel]

$d = 2$; $\lambda > 1$ const, $O(n^2)$ mixing time (optimal).

[Majumder, Sahu, Reif]

$d = 2,3$; $\lambda = \Theta(n)$, poly time.

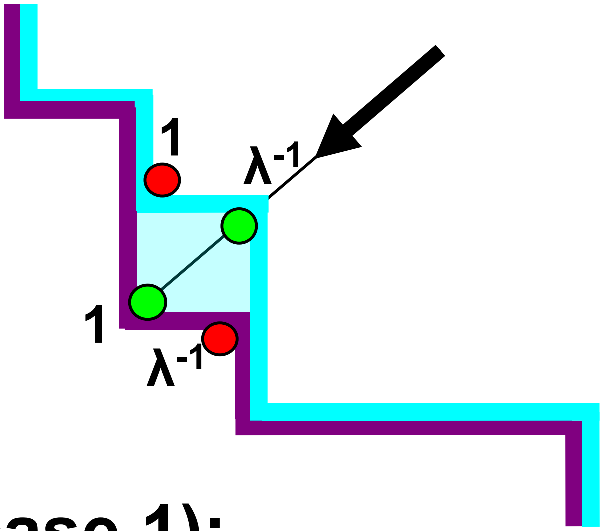
[Greenberg, Pascoe, R.]

$d = 2$, $\lambda > 1$ const
$d \geq 3$, $\lambda > d^2$
} $O(n^d)$ mixing time.

[Caputo, Martinelli, Toninelli]

$d = 3$, $\lambda > 1$ $O(n^3)$ mixing time.

Coupling the *Biased* Chain



Coupling:

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

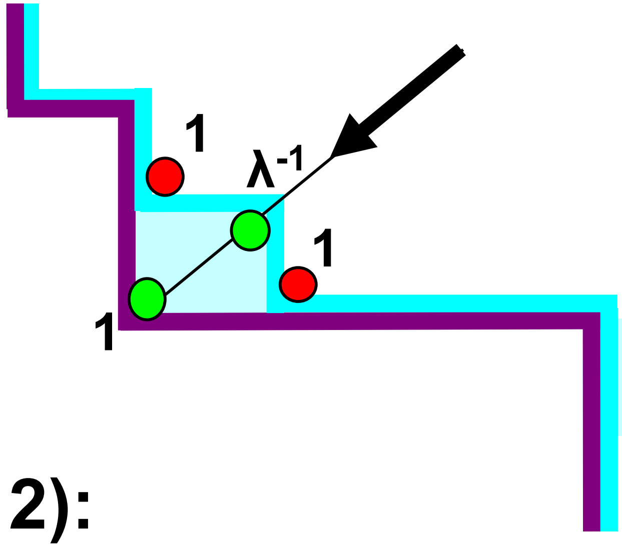
(case 1):

- $E[\Delta(\Psi_t)] = p\,(-\,wt(G) + wt(B)\,)$

$= p\,(-1 - \lambda^{-1} + 1 + \lambda^{-1})$

≤ 0

Coupling the *Biased* Chain



Coupling:

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

(case 2):

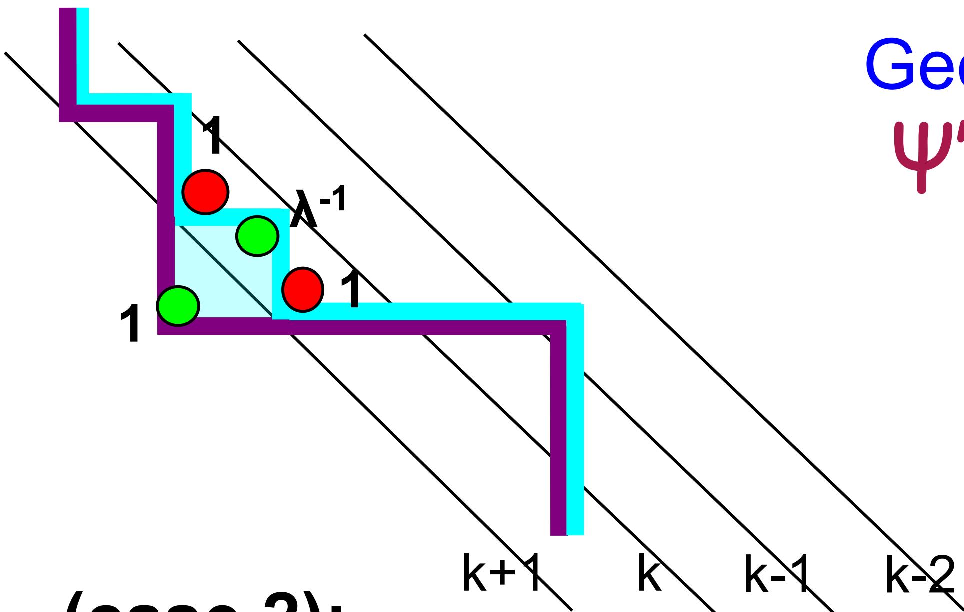
- $E[\Delta(\Psi_H)] = p\,(-\,wt(G) + wt(B))$

$= p\,(-1 - \lambda^{-1} + 1 + 1)$

> 0

Introduce a different metric.

Introduce a New Metric



Geometric distance function:

$$\Psi'(\sigma,\tau) = \sum_{x \text{ in } \tau\oplus\sigma} (\sqrt{\lambda})^{\mathbf{diag(x)}}$$

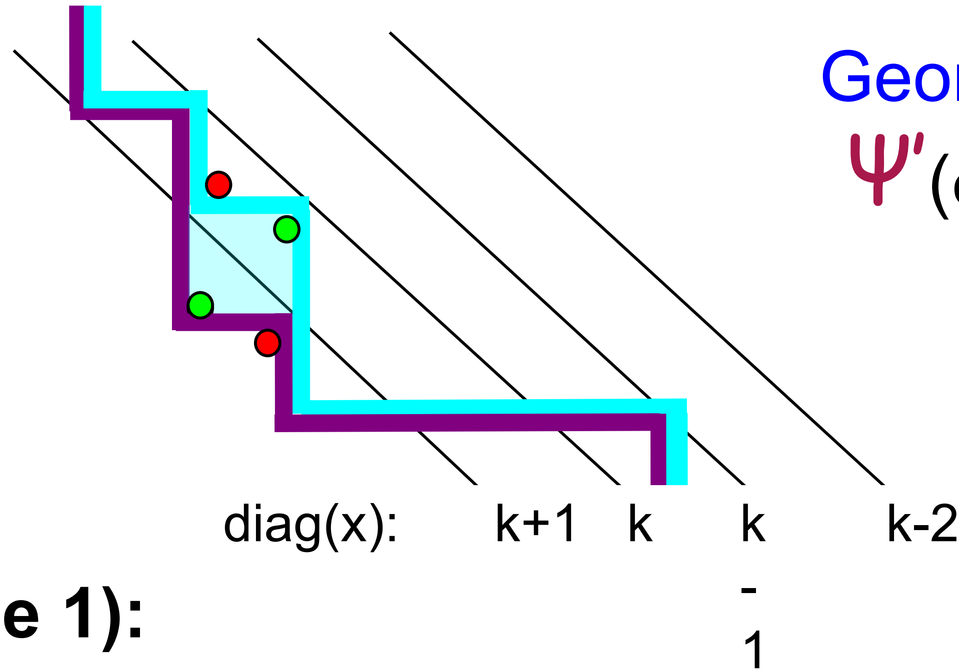
diag(x):

(case 2):

- $E[\Delta(\Psi)] = p\,(-\,wt(G) + wt(B))$

$$= p\,\lambda^{(k+1)/2}\,(-1 - \lambda^{-1} + \lambda^{-1/2} + \lambda^{-1/2}) \;<\; 0$$

Introduce a New Metric



Geometric distance function:

$$\Psi'(\sigma,\tau) = \sum_{x \text{ in } \tau\oplus\sigma} (\sqrt{\lambda})^{\mathbf{diag(x)}}$$

(case 1):

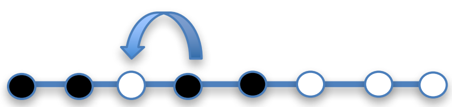
- $E[\Delta(\Psi')] = p\,(-\,wt(G) + wt(B)\,)$

$$= p\,\lambda^{(\kappa+1)/2}\left(-1 - \lambda^{-1} + \lambda^{-1/2} + (\lambda^{-1})\,\lambda^{1/2}\right) \;\mathbf{< 0}$$

The distance Ψ'_t is always nonincreasing (in expectation), and by path coupling the chain is rapidly mixing.

Outline

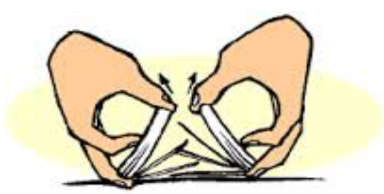
- Sampling paths uniformly
- Paths with uniform bias
- Paths with non-uniform bias



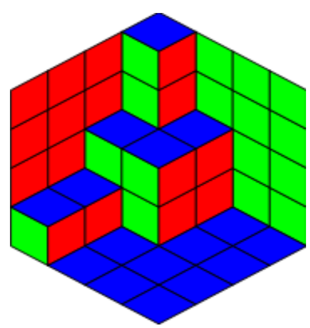
Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

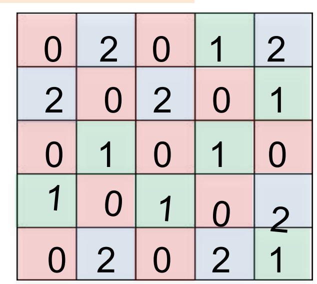
Integer Partitions



Card Shuffling



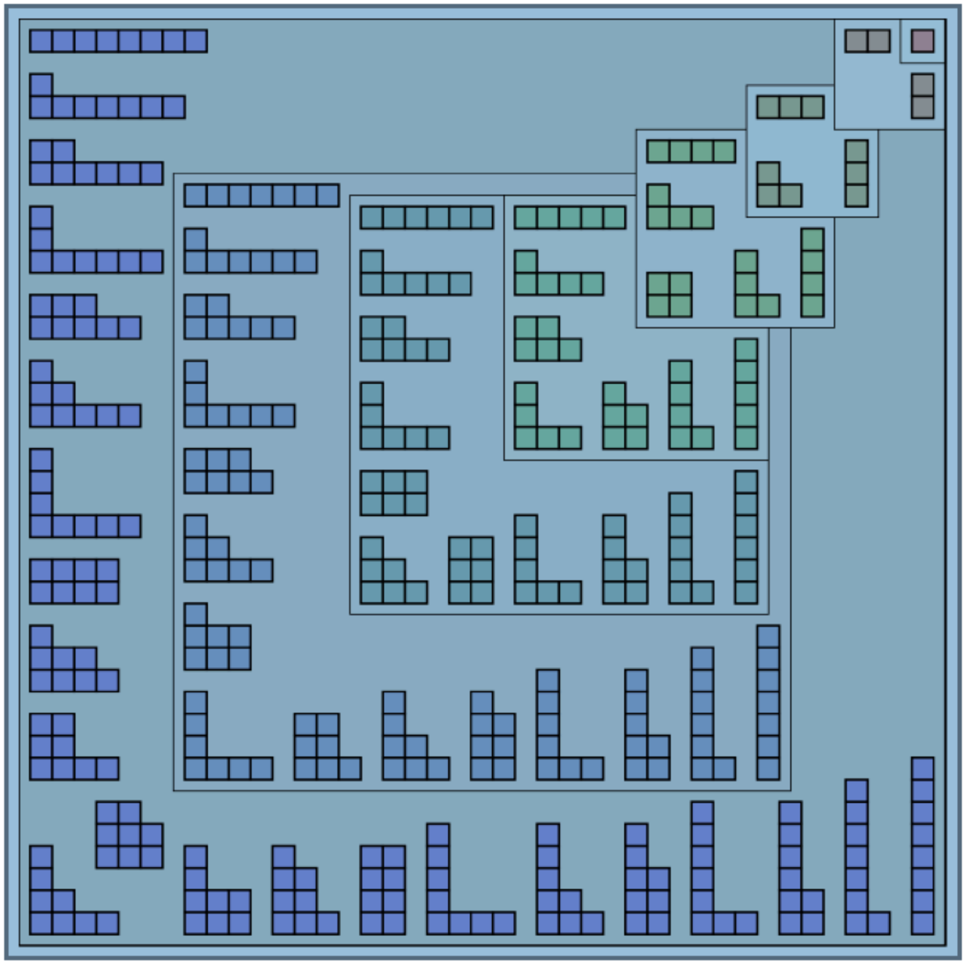
Lozenge Tilings



3-colorings

Integer Partitions

Ferrers Diagrams:



Let the partition number p(n) be the number of partitions of size n:

p(i): 1, 2, 3, 5, 7, 11, 15, 20,…

Asymptotic Estimate: $p(n) \sim \frac{1}{4\sqrt{3n}} e^{\pi\sqrt{2n/3}}$

[Hardy, Ramanujan 1918]

Sampling Integer Partitions

Dynamic Programming:

Restricted Partition Number p(n,k):

number of ways to partition n into at most k pieces.

Simple recurrence relation: p(n,k) = p(n-k, k) + p(n, k-1).

Thus we can exactly sample partitions of n using dynamic programming and self-reducibility.

However, *space* requirements are very large:

partition numbers grow as $\approx e^{O(\sqrt{n})}$

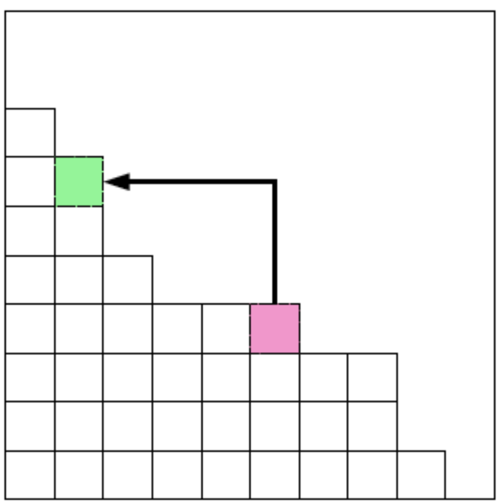
Almost a petabyte for $n \approx 1000000$.

Markov Chains?

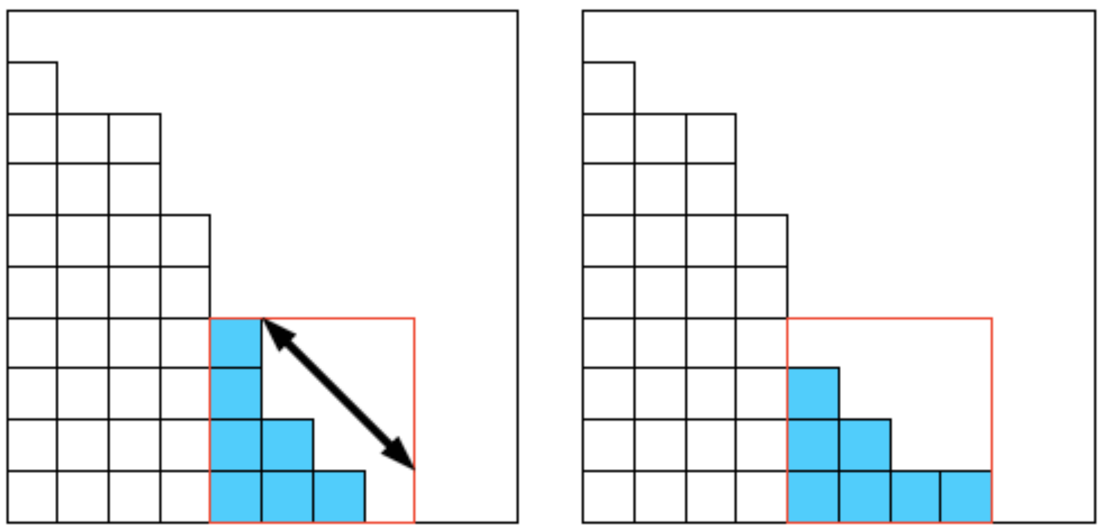
Many chains with simple rules converge to the uniform dist'n, but the mixing time remains open for all of them.

Ex:

Chain 1: move a square



Chain 2: pick a sub-square and flip

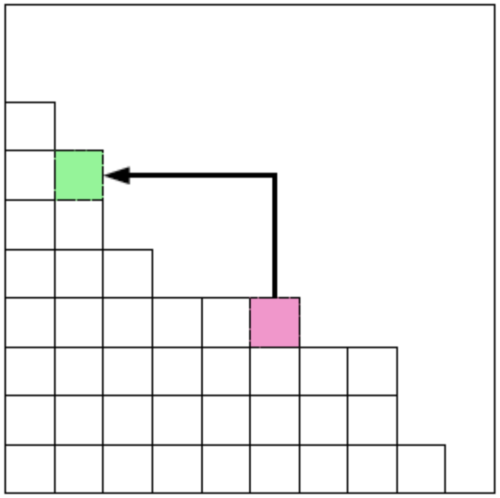


[Aldous 1999], [Berestycki, Pitman 2007]

And many others….

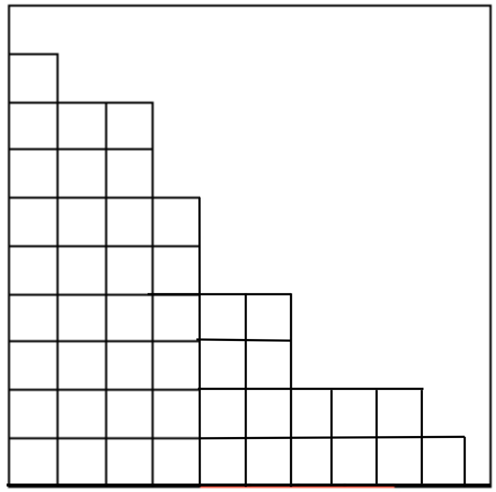
Approaches

Try 1:



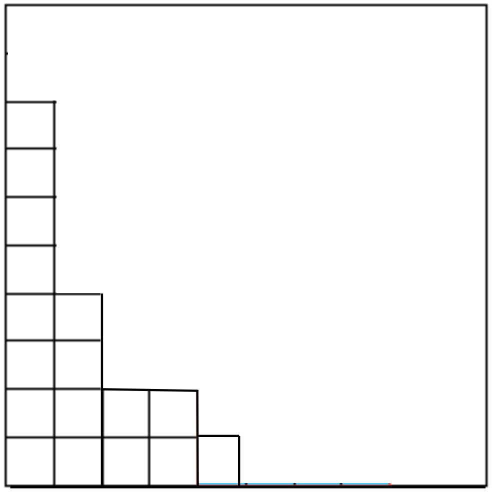
Ω_n

Try 2:



$\Omega = \bigcup_{i=1}^{n} \Omega_i$

Try 3:



$\Omega = \bigcup_{i=1}^{n} \Omega_i, \quad \lambda < 1$

Need: 1. The chain is rapidly mixing.

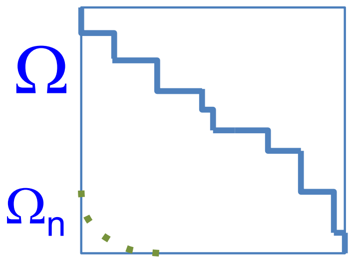
2. Rejection sampling is efficient.

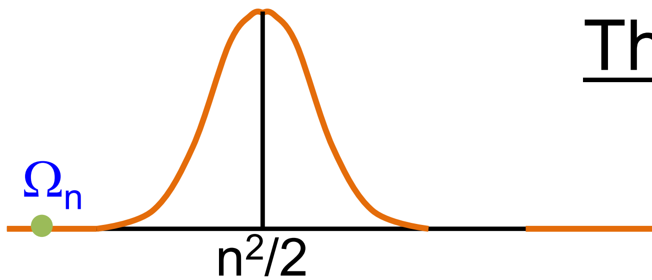
Boltzmann Sampling

[Bhakta, Cousins, Fahrbach, R.]

Let Ω be the set of all lattice paths in an n x n region;

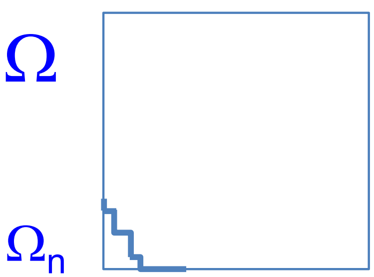
and Ω_n the set of lattice paths with area n.

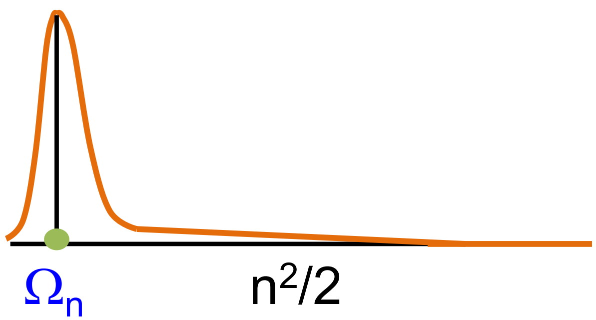




<u>Thm:</u> $p_i = |\Omega_i|$ is logconcave ($n > 25$). [DeSalvo, Pak]

- Generate samples σ of Ω with prob. proportional to $\lambda^{area(\sigma)}$.





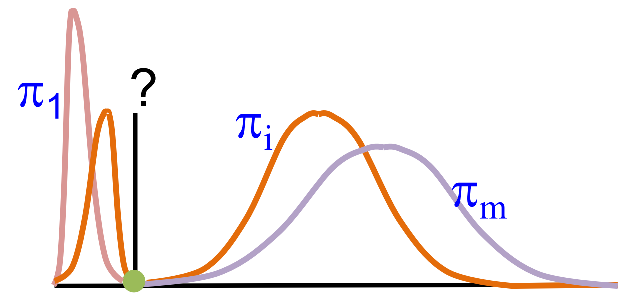
So $q_i = p_i \lambda^i$ is also logconcave (and therefore unimodal).

- Setting $\lambda = p_n / p_{n+1}$ gives $q_n = q_{n+1}$.
- So n and $n+1$ must be the modes of the dist'n.

Boltzmann Sampling

[Bhakta, Cousins, Fahrbach, R.]

What about partition classes where we do not know if the sequence is logconcave?
(e.g., partitions with at most k pieces,…)



Need: 1. The chain is rapidly mixing.

2. Rejection sampling is efficient.

<u>Thm:</u> If the Markov chain is rapidly mixing for all λ,

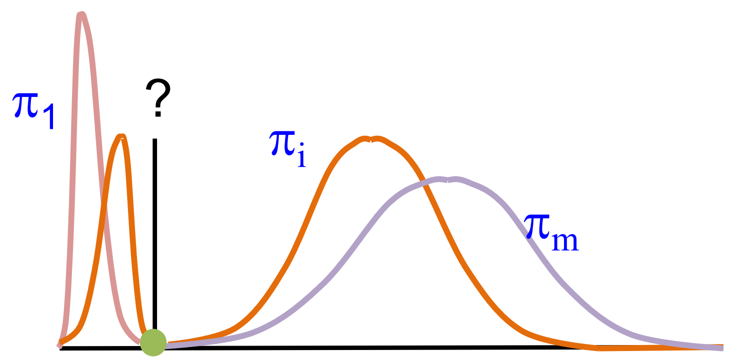
then rejection sampling is also efficient for some λ !

Boltzmann Sampling

[Bhakta, Cousins, Fahrbach, R.]

Define $\lambda_1, \ldots, \lambda_m$ and let π_i is the distribution with bias λ_i s.t.:

- $|| \pi_i , \pi_{i+1} ||$ small, for all i;
- π_1 is concentrated on configurations of size < n;
- π_m is concentrated on configurations of size > n;
- MC is rapidly mixing, for all λ_i.

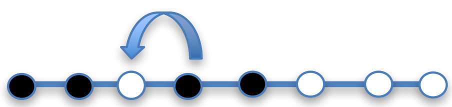


Then there exists a λ_i s.t.

Pr(π_i outputs a sample of size n) > poly(n).

Outline

- Sampling paths uniformly
- Paths with uniform bias
- Paths with non-uniform bias

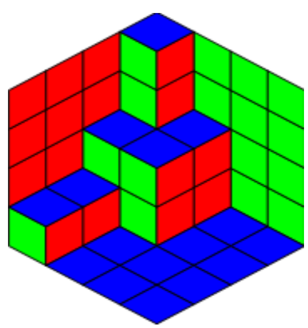


Simple Exclusion Processes

1 + 1 + 4 + 5 = 11

Integer Partitions

Card Shuffling



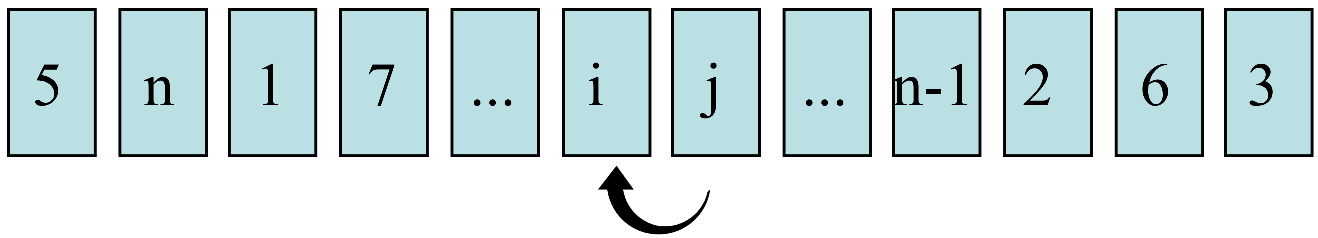
Lozenge Tilings

0	2	0	1	2
2	0	2	0	1
0	1	0	1	0
1	0	1	0	2
0	1	0	2	1

3-colorings

Biased Card Shuffling

- Pick a pair of adjacent cards uniformly at random
- Put j ahead of i with probability $p_{j,i} = 1 - p_{i,j}$



Converges to: $\pi(\sigma) = \prod_{i<j:\ \sigma(i)>\sigma(j)} \frac{p_{ij}}{p_{ji}} / Z$

This is related to "Move-Ahead-One" for self-organizing lists.

[Fill]

Biased Permutations

Question: If the $\{p_{ij}\}$ are positively biased ($p_{ij} \geq 1/2 \quad \forall\, i < j$), is M always rapidly mixing?

Recall, with constant bias:

If $p_{ij} = p \quad \forall\, i < j$, $p>1/2$, then M mixes in $\theta(n^2)$ time. [BBHM]

Linear extensions of a partial order: If $p_{i,j} = 1/2$ or $1 \quad \forall\, i < j$, then M mixes in $O(n^3 \log n)$ time. [Bubley and Dyer]

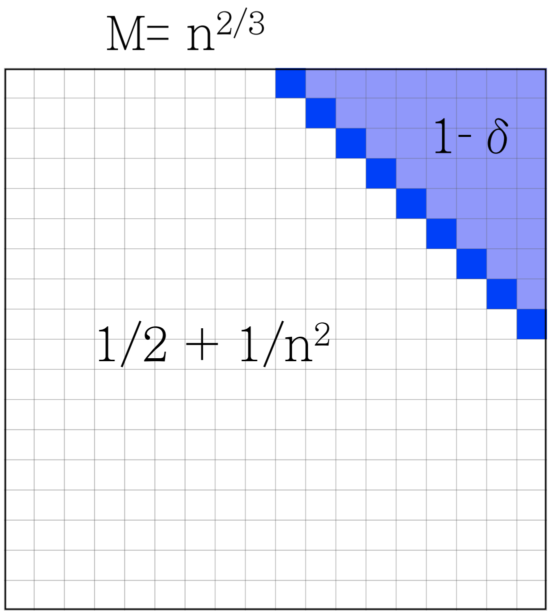
Fast for two classes: "Choose your weapon" and "league hierarchies" (if *weakly regular*). [Bhakta, Miracle, R., Streib]

Biased Permutations

Question: If the $\{p_{ij}\}$ are positively biased ($p_{ij} \geq 1/2 \quad \forall\ i < j$), is M always rapidly mixing?

No !!!

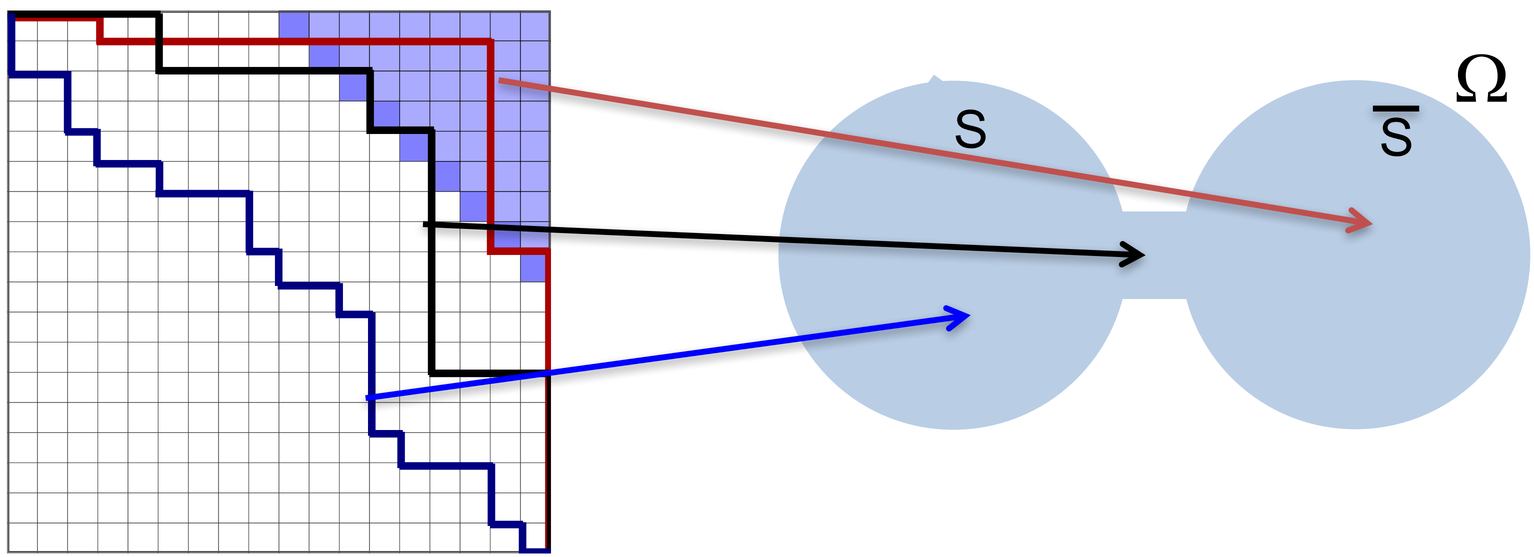
[BMRS]



Biased Permutations

Question: If the $\{p_{ij}\}$ are positively biased ($p_{ij} \geq 1/2 \quad \forall\, i < j$), is M always rapidly mixing?

No !!!



The state space has a "bad cut" so M requires exponential time.

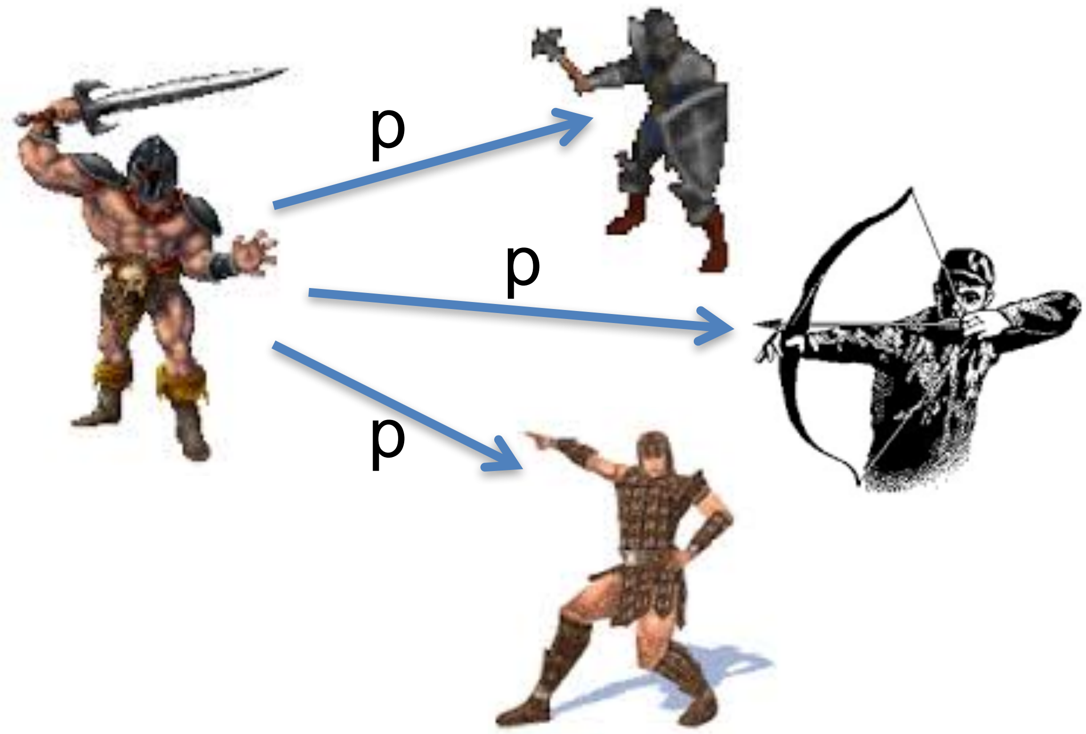
However, most cases do seem fast….

“Choose your weapon”

[BMRS]

Given $r_{1, \ldots ,} r_{n-1,}$ $r_i \geq 1/2$.

<u>Thm 1</u>: Let $p_{ij} = r_i$ $\forall\ i < j$. Then M_{NN} is rapidly mixing.

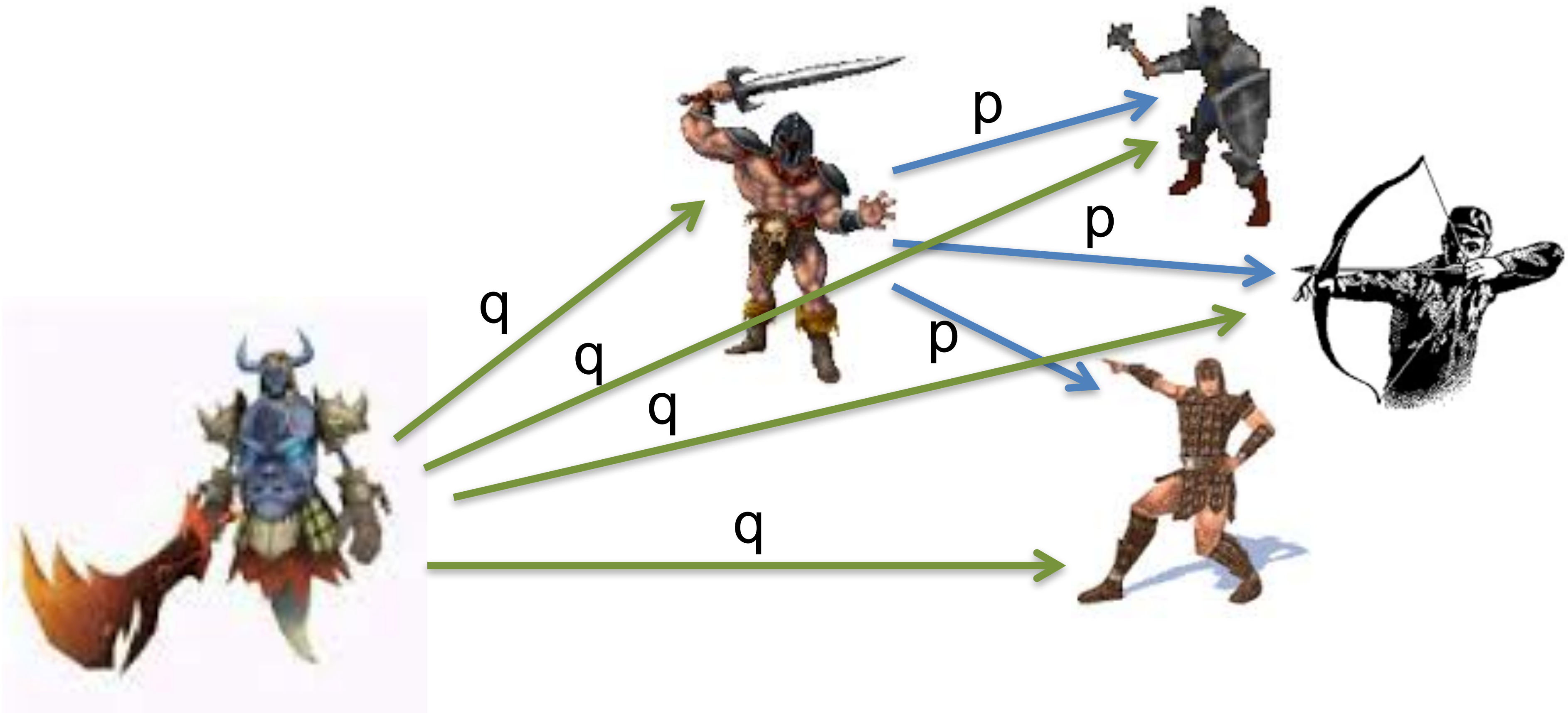


“Choose your weapon”

[BMRS]

Given $r_{1, \ldots ,} r_{n-1,}$ $r_i \geq 1/2$.

<u>Thm 1</u>: Let $p_{ij} = r_i$ $\forall$ $i < j$. Then M_{NN} is rapidly mixing.

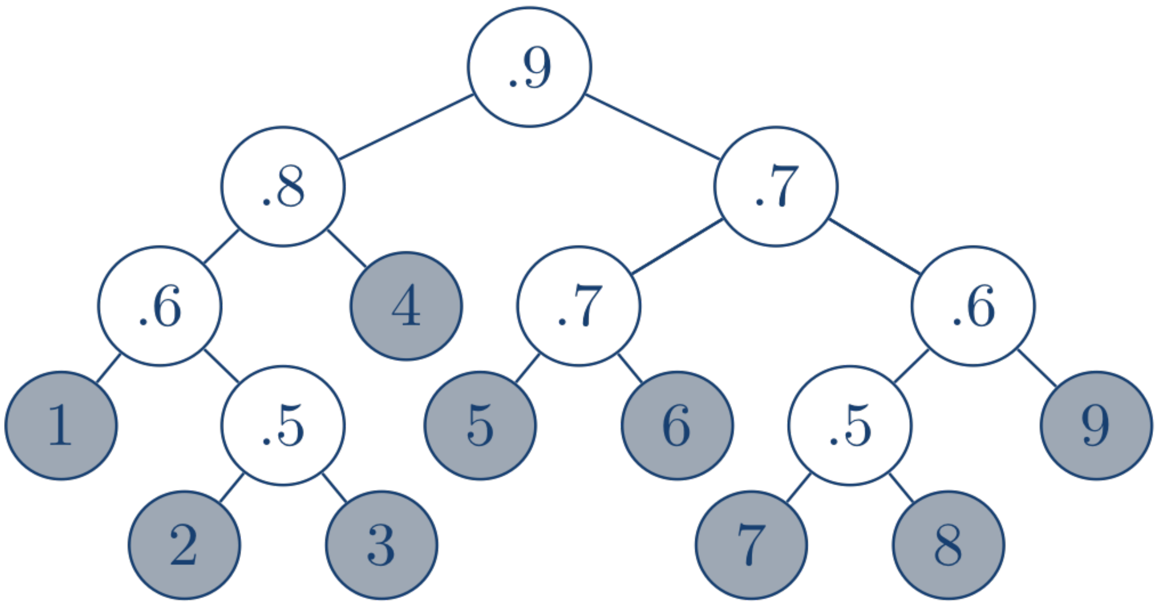


"League Hierarchies"

Let T be a binary tree with leaves labeled $\{1,\ldots,n\}$.

Given $q_v \geq 1/2$ for each *internal* vertex v.

<u>Thm 2:</u> Let $p_{i,j} = q_{i \wedge j}$ for all $i < j$. Then M_{NN} is rapidly mixing.

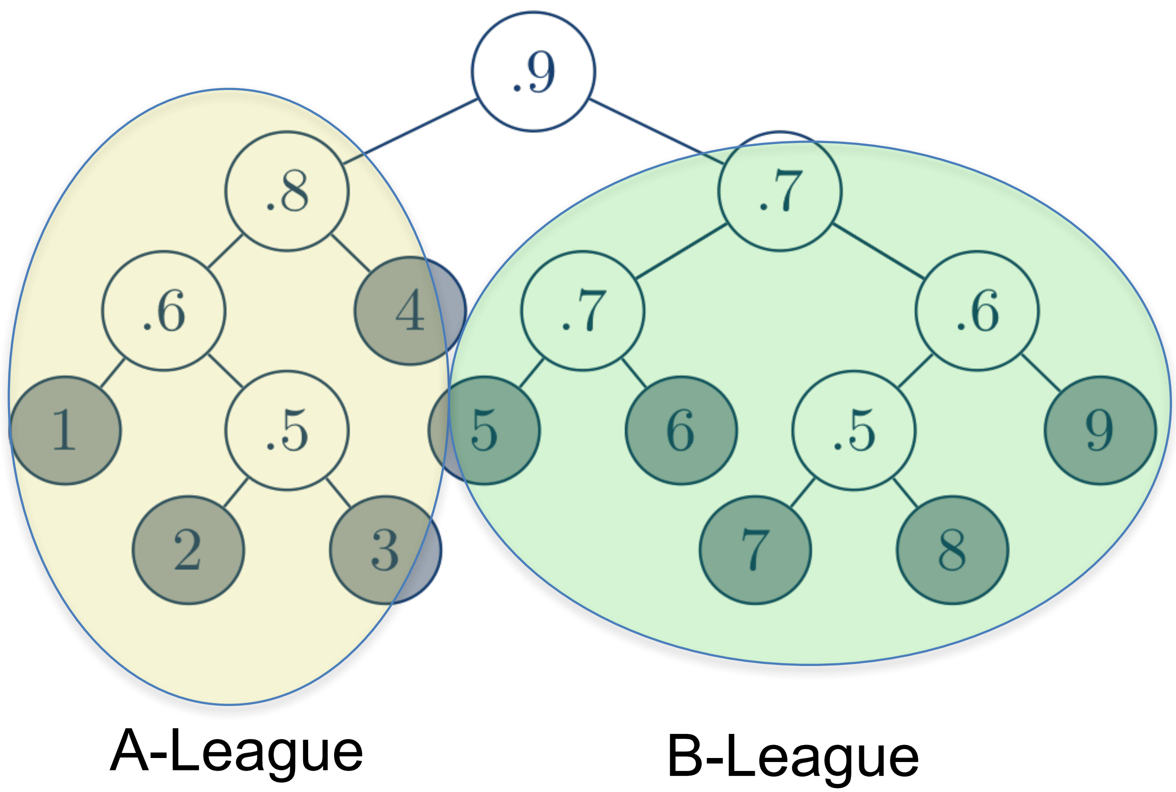


League Hierarchies

Let T be a binary tree with leaves labeled {1,…,n}.

Given $q_v \geq 1/2$ for each *internal* vertex v.

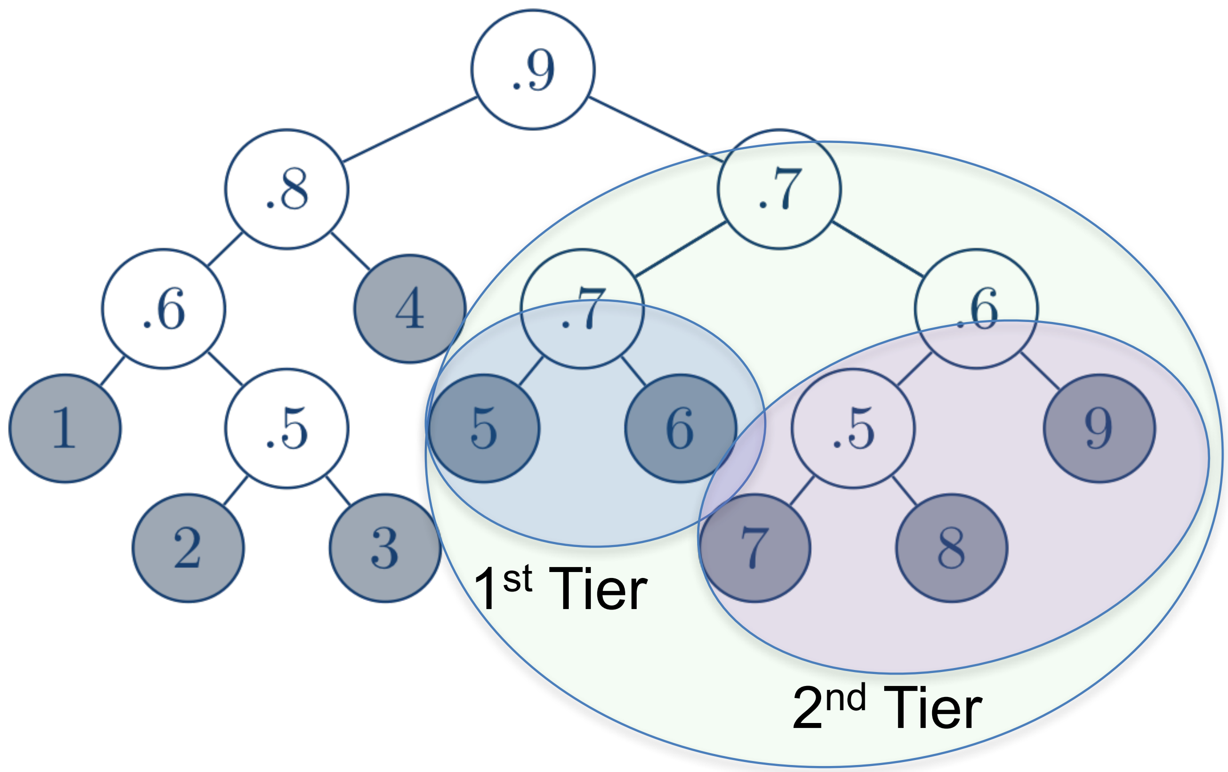
Thm 2: Let $p_{i,j} = q_{i \wedge j}$ for all $i < j$. Then M_{NN} is rapidly mixing.



League Hierarchies

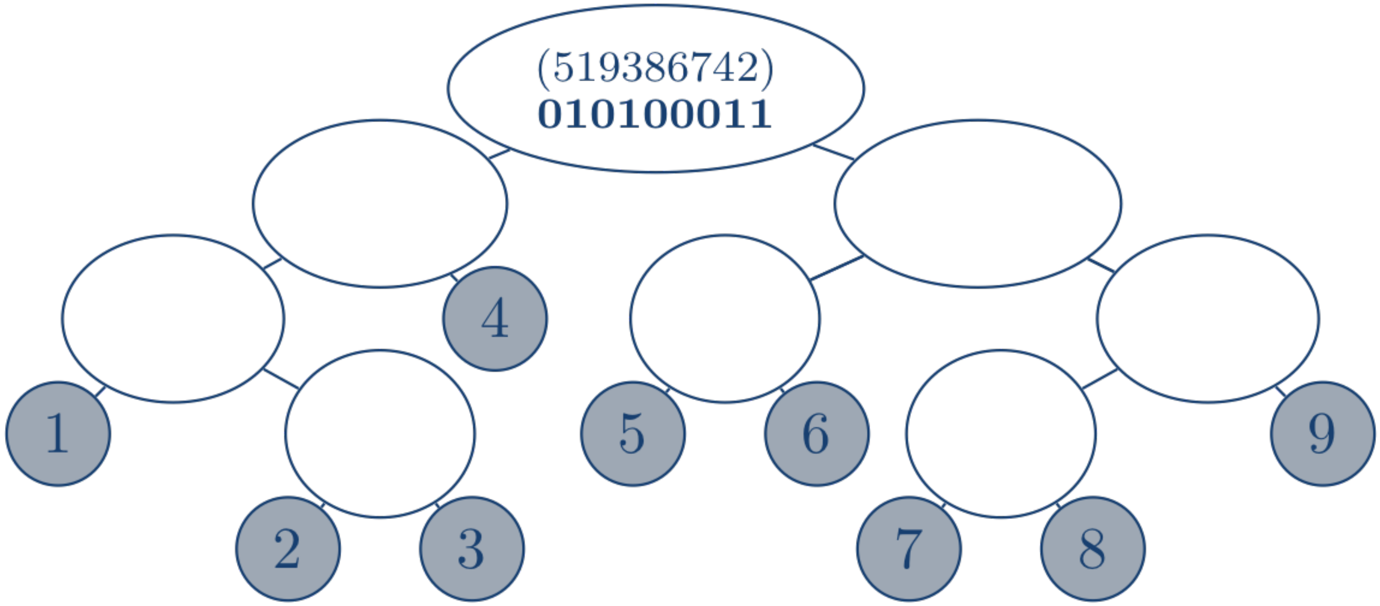
Let T be a binary tree with leaves labeled $\{1,\ldots,n\}$.
Let $q_v \geq ½$ be assigned to each *internal* vertex v.

Thm 2: Suppose $p_{ij} = q_{i\wedge j}$ for all $i < j$ for some labeled binary tree.
Then M_{NN} is rapidly mixing.



Thm 2: Proof sketch

Thm 2: Let T be a binary tree with leaves labeled $\{1,\ldots,n\}$.
Let $q_v \geq ½$ be assigned to each *internal* vertex v.
Let $p_{i,j} = q_{i \wedge j}$ for all $i < j$. Then M_{NN} is rapidly mixing.

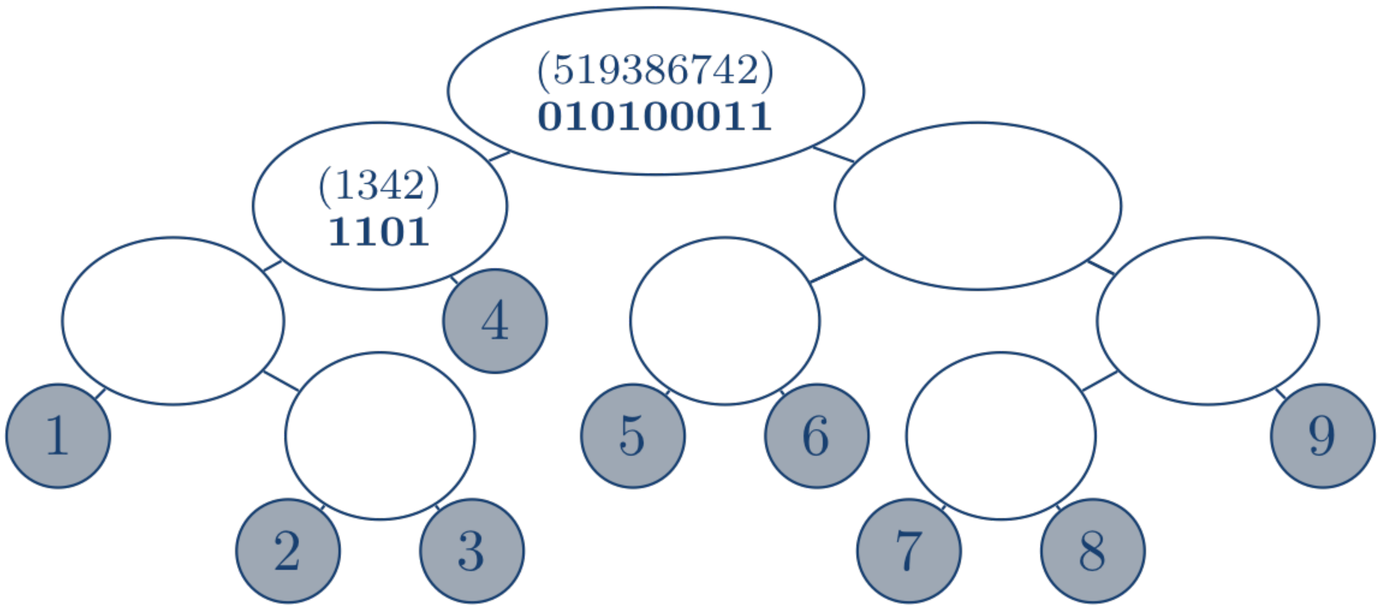


Thm 2: Proof sketch

Thm 2: Let T be a binary tree with leaves labeled $\{1,\ldots,n\}$.

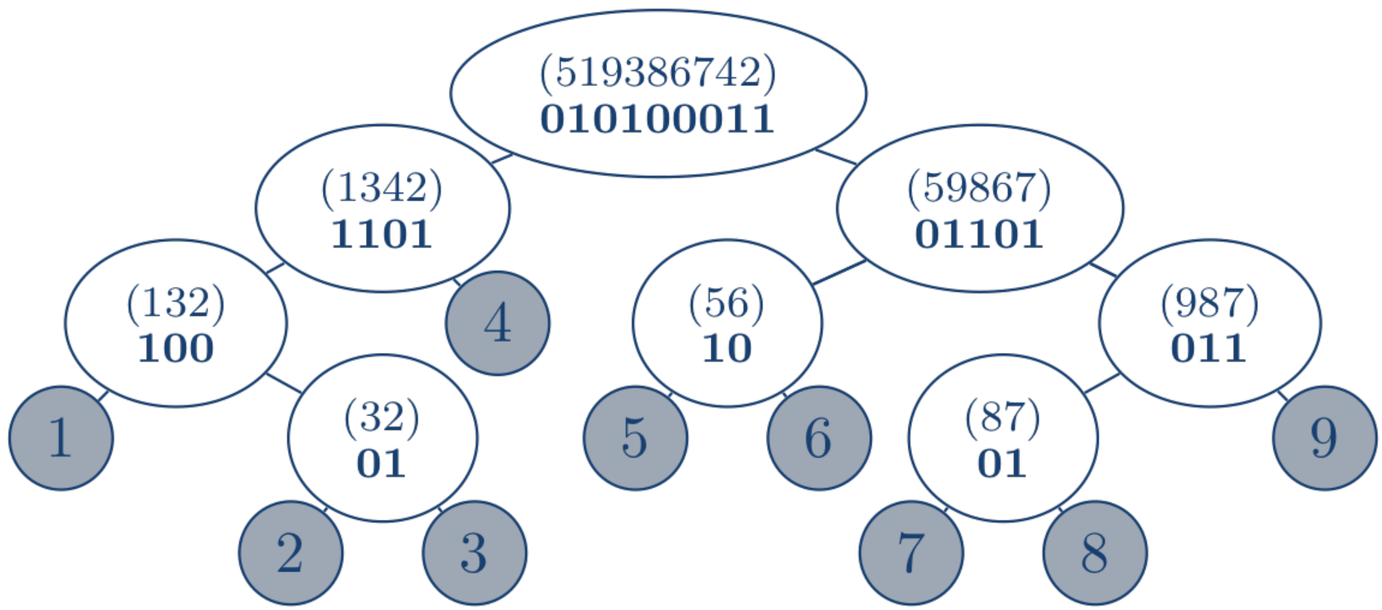
Let $q_v \geq ½$ be assigned to each *internal* vertex v.

Let $p_{i,j} = q_{i \wedge j}$ for all $i < j$. Then M_{NN} is rapidly mixing.



Thm 2: Proof sketch

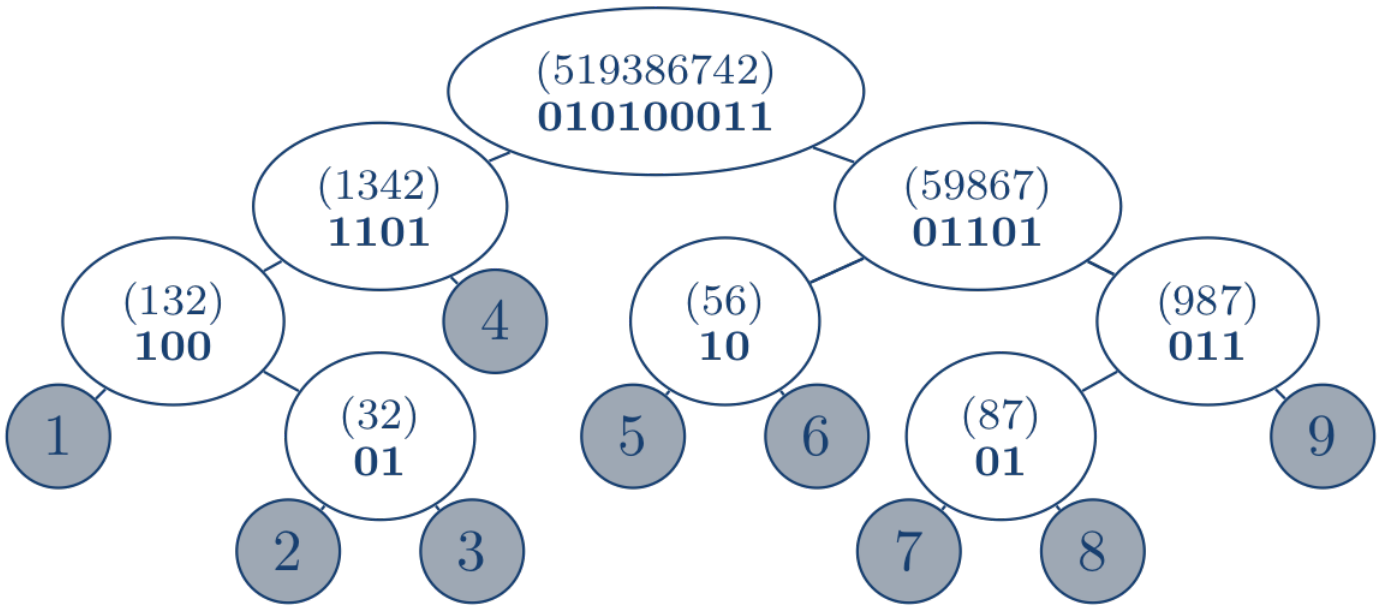
Thm 2: Let T be a binary tree with leaves labeled {1,…,n}.
Let $q_v \geq ½$ be assigned to each *internal* vertex v.
Let $p_{i,j} = q_{i∧j}$ for all $i < j$. Then M_{NN} is rapidly mixing.



Thm 2: Proof sketch

Markov chain M′ allows a transposition if it corresponds to an ASEP move on one of the internal vertices.

Each ASEP is rapidly mixing $\Rightarrow$ M′ is rapidly mixing.



M_{NN} is also rapidly mixing if {p} is *weakly regular*.

i.e., for all i, $p_{i,j} < p_{i,j+1}$ if $j > i$. (by comparison)

Open Problems

1. **Fill's conjecture**: is M_{NN} always rapidly mixing when $\{p_{ij}\}$ are positively biased and regular?

   (i.e., $p_{ij} > ½$ and p_{ij} is monotonic in i and j)

1′. What about the **special case**:

   Given $a_1,\ldots,a_n$ "strengths", with $a_i > 0$, let $p_{ij} = a_i / (a_i + a_j)$.

2. When does bias speed up or slow down a chain?

Thank you!