Edge Flip Chain for Unbiased Dyadic Tilings

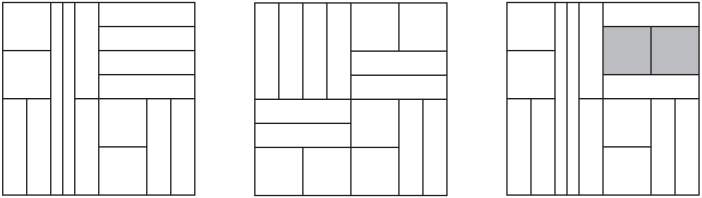
Sarah Cannon, *Georgia Tech*, ,
David A. Levin, *U. Oregon*

Alexandre Stauffer, *U. Bath*,

March 30, 2018

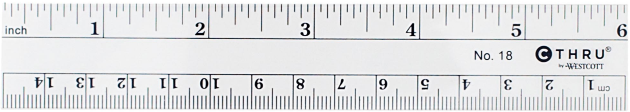
Dyadic Tilings

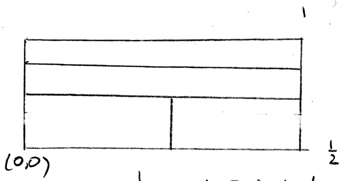
A **dyadic tiling** is a tiling on $[0,1]^2$ by $n = 2^k$ **dyadic rectangles**, rectangles which are cartesian products of intervals of the form $[a2^{-s}, (a+1)2^{-s}]$, each with area $\frac{1}{n}$.



(a) (b) (c)

Endpoints don't straddle big ticks:

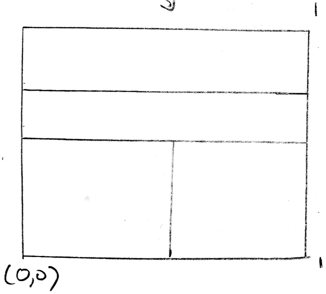




AREA $\frac{1}{8}$

TILING BY 2^{-3} AREA RECTANGLES

$$(x,y) \longmapsto (x,2y)$$



AREA $\frac{1}{4}$

tilings $= A_4$

Counting and sampling

- Self-similarity: By scaling by 2, a tiling of the right-half by dyadic tiles of area 2^{-k} corresponds bijectively with a tiling of the unit square by tiles of size 2^{-k+1}.
- Let Ω_k be the collection of dyadic tilings of size $n = 2^k$. Write $A_k = |\Omega_k|$.
- The first point means the number of tilings of the right-half by tilings of area 2^{-k} equals A_{k-1}.

Recursion

- Every tiling has either a vertical or a horizontal bisector, or both, whence
$$A_k = 2A_{k-1}^2 - A_{k-2}^4$$
- $A_k \sim \phi^{-1}\omega^{2^k}$, where $\phi = (1+\sqrt{5})/2$ and $\omega = 1.8445...$ [Lagarias, Spencer, Vinson provide asymptotics for this recursion for arbitrary initial conditions.]
- More on the combinatorics in Janson, Randall, Spencer (2002) and Angel, Holroyd, Kozma, Wästlund, Winkler (2014)

Other contexts

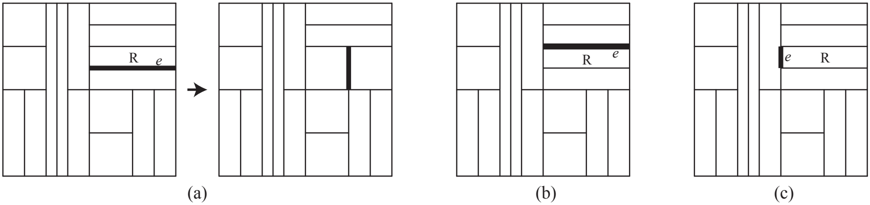
Tilings in which all rectangles are dyadic, but may have different areas, have been used as a basis for

- subdivision algorithms to solve problems such as approximating singular algebraic curves, and
- classifying data using decision trees.

In both of these examples, the unit square is repeatedly subdivided into smaller and smaller dyadic rectangles until the desired approximation or classification is achieved, with more subdivisions in the areas of the most interest (e.g., near the algebraic curve or where data classified differently is close together).

Markov Chain Dynamics

Pick a tile at random, select one of its edges, and **flip** neighboring tiles if you stay in Ω_k.



- This MC is ergodic with stationary distribution uniform. (Check detailed balance equations!)
- MCMC question: How long must chain be sampled to obtain satisfactory uniform sample from uniform distribution on tilings?
- N.B.: If you really want to sample, there are arguably better ways than MCMC! Recursive constructions are almost uniform.

Mixing, Relaxation, and all that

Ergodic Markov chain $\{X_t\}$ with stationary distribution π:

- **Mixing Time**:

$$t_{\text{mix}}(\epsilon) = \min\{t : \max_x \|P_x(X_t \in \cdot) - \pi\|_{\text{TV}} < \epsilon\}.$$

- **Spectral Gap**: Order eigenvalues $1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ of transition matrix. Then $\gamma = 1 - \lambda_2$ is the spectral gap. [Assume all eigenvalues are non-negative, which is satisfied, e.g., when the chain has some laziness.] The gap measures the asymptotic rate of convergence.
- The **relaxation time** is $t_{\text{rel}} = \frac{1}{\gamma}$.
- t_{rel} is the time scale on which correlations decay.

- Relation between mixing and relaxation:

$$(t_{\text{rel}} - 1)\log\left(\frac{1}{2\varepsilon}\right) \le t_{\text{mix}}(\varepsilon) \le \log\left(\frac{1}{\varepsilon\pi_{min}}\right) t_{\text{rel}}.$$

- The game is: how does $t_{\text{mix}}(1/4)$ scales with instance size n; i.e. is it polynomial, or exponential?

Previous results

- Weighted tilings. $\pi(x) \propto \lambda^{|x|}$, where $|x|$ is the sum of the length of edges in x.
- If $\lambda > 1$, then $t_{\text{mix}} = \Omega(e^{cn^2})$ [Cannon, Miracle, Randall]. (There is an obvious bottleneck.)
- If $\lambda < 1$, then $t_{\text{mix}} = O(n^2 \log n)$ and $t_{\text{rel}} = O(n)$. [Cannon, Miracle, Randall]
- Cannon et al: What happens at criticality??? ($\lambda = 1$).

Some motivations

- Weighted version introduces "temperature" or weight parameter λ: Distribution on tilings is $\pi(x) \propto \lambda^{|x|}$, where $|x|$ is the sum of the length of edges in x.
- For many models with Boltzmann/Gibbs-type distributions

$$\pi(x) = Ze^{-H(x)/t} \quad H(x) = \text{energy},$$

$x \in \Omega = S^n$, where functions in S^n may be constrained in some way. [E.g. $\Omega = \{-1, 1\}^V$, V are the vertices of a graph.] Let n be the number of *sites*, so

$$n = \log(\text{state-space}),$$

(For dyadic-tilings, above enumeration shows $n = 2^k$ is order log of the state-space.)

The **ansantz** is:

- High temperature is fast mixing ($n \log n$) with a sharp **cut-off** ($t_{\text{mix}}(\epsilon)/t_{\text{mix}}(1/4) \to 1$),
- Low temperature is slow mixing (exponential in n),
- Critical temperature is higher degree polynomial mixing than high temperature.

- Complete picture confirmed for special cases of Ising model. (L.-Luczak-Peres for complete graph, Lubetzky-Sly confirm this in a series of remarkable papers for lattices.)
- Potts on complete graph (Cuff et al), Lattices (Ghessari-Lubetzky)
- We can confirm the main features of this picture for dyadic tilings, but unlike the examples above, mixing times are not sharp.

Our Main Results

Theorem

The relaxation time of the edge-flip Markov chain for dyadic tilings of size n is at most $O(n^{\log 17})$. As a consequence, the mixing time of this chain is at most $O(n^{1+\log 17})$.

Theorem

The relaxation time of the edge-flip Markov chain for dyadic tilings of size n is at least $\Omega(n^{2\log_2 \phi})$, where $\phi = \frac{\sqrt{5}+1}{2}$ is the golden ratio.

Proof methods

- The upper bound uses a rescaling/block dynamics idea exploited in spin systems by, e.g.m F. Martinelli and coauthors. The recursive nature of the tilings make this argument natural here!
- The lower bound is via Cheeger's inequality, identifying a bottleneck in the state-space.

Lower Bound via Conductance (Cheeger Inequality)

- The **conductance**

$$Q(A, B) = \sum_{x\in A, y\in B} \pi(x) P(x, y)$$

  measures the (stationary) flow from A to B.

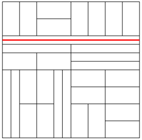
- **Cheeger's inequality**. For any S with $\pi(S) \leq 1/2$, the spectral gap γ satisifies

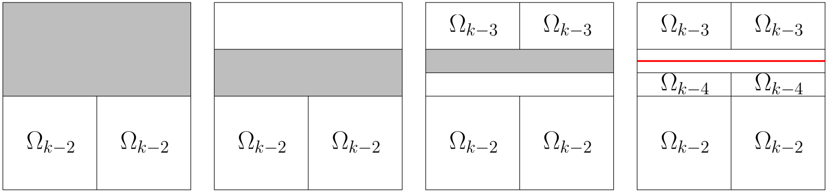
$$\gamma \leq 2\frac{Q(S, S^c)}{\pi(S)}.$$

- Bottlenecks make the right-hand side small, slowing mixing.

The cut

- $\Omega^{|}$ are all tilings without a vertical bisector.
- $\pi((\Omega^{|})^c) < 1/2$.
- $$Q((\Omega^{|})^c, \Omega^{|}) = \frac{|\partial(\Omega^{|})^c|}{|\Omega_k| 2n}$$
- Example of configuration in $\partial(\Omega^{|})^c$, the set of configurations one move away from $\Omega^{|}$





(a) (b) (c) (d)

$$|\partial\Omega_k^{|}| = \prod_{i=2}^{k}\left(2|\Omega_{k-i}|^2\right) = 2^{k-1}\prod_{i=0}^{k-2}|\Omega_i|^2 = \frac{n}{2}\prod_{i=0}^{k-2}|\Omega_i|^2.$$

Hence,

$$Q = \frac{1}{4|\Omega_k|}\prod_{i=0}^{k-2}|\Omega_i|^2 \leq \frac{1}{4}\phi^{-2k+2},$$

where the last inequality follows from another combinatorial argument.

Therefore, there exists a constant $c > 0$ such that

$$\gamma_k \le c\phi^{-2k}.$$

This implies that the relaxation time and mixing time satisfy

$$t_{\mathrm{rel}}, t_{\mathrm{mix}} \ge \frac{1}{c}\phi^{2k} = \frac{1}{c}\phi^{2\log_2 n} = \frac{1}{c}n^{2\log_2 \phi} = \Omega(n^{2\log_2 \phi}).$$

Note $2\log_2 \phi \approx 1.39 > 1$.

Compare with Cannon et al, $t_{\mathrm{rel}} = O(n)$ when $\lambda < 1$.

This suggests but does not prove the mixing time should be larger at criticality than at high temperature. (Cannon et al *upper bound* is $O(n^2 \log n)$, while our *lower bound* is $\Omega(n^{1.39})$.

Block dynamics

For $k \geq 2$, the block dynamics Markov chain on the state space Ω_k of all dyadic tilings of size 2^k is given by the following rules.

Beginning at any dyadic tiling σ_0, repeat:

- Uniformly at random choose a tiling $\rho \in \Omega_{k-1}$.
- Uniformly at random choose *Left*, *Right*, *Top*, or *Bottom*.
- To obtain σ_{i+1}:
  - If *Left* was chosen and σ has a vertical bisector, retile σ's left half with ρ, under the mapping $x \to x/2$.
  - If *Right* was chosen and σ has a vertical bisector, retile σ's right half with ρ, under the mapping $x \to (x+1)/2$.
  - If *Bottom* was chosen and σ has a horizontal bisector, retile σ's bottom half with ρ, under the mapping $y \to y/2$.
  - If *Top* was chosen and σ has a horizontal bisector, retile σ's top half with ρ, under the mapping $y \to (y+1)/2$.
- Else, set $\sigma_{i+1} = \sigma_i$.

Upper bound

- Uses “block updates”
- Perfect set-up for this: the system “looks the same at different scales”!
- $\gamma_k \geq \gamma_{k,BLOCK} \cdot \gamma_{k-1}$
- $\gamma_{k,BLOCK} \geq \frac{1}{17}$
- $\gamma_k > 17^{-(k-k_0)} \gamma_{k_0}$.

Variational characterization

- **Dirichlet form**

$$\mathcal{E}(f) = \sum_{x,y} \pi(x) P(x,y) [f(x) - f(y)]^2$$

- The spectral gap satisfied, via its variational characterization,

$$\gamma = \min_f \frac{\mathcal{E}(f)}{\mathrm{Var}(f)}$$

- Relate $\mathcal{E}_k(f)$ to $\mathcal{E}_{k,BLOCK}(f)$ and $\mathcal{E}_{k-1}$

- Just need to wait a constant expectation time to couple any two configurations.
- Spectral gap is bounded below by a constant.

some details

It can be shown that

$$\mathcal{E}_{k,block}(f) \leq \frac{\mathcal{E}_{k,edge}(f)}{\gamma_{k-1}}.$$

Note this implies that for any f,

$$\mathrm{Var}_k(f) \leq \frac{\mathcal{E}_{k,block}(f)}{\gamma_{k,block}} \leq \frac{\mathcal{E}_{k,edge}(f)}{\gamma_{k,block} \cdot \gamma_{k-1}}.$$

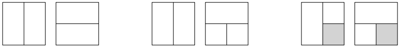
Let f be chosen to be the function achieving equality in

$$\mathrm{Var}_k(f) \leq \frac{\mathcal{E}_{k,edge}(f)}{\gamma_k}.$$

We conclude

$$\gamma_k = \frac{\mathcal{E}_{k,edge}(f)}{\mathrm{Var}_k(f)} \geq \gamma_{k,block} \cdot \gamma_{k-1}.$$

Contractive coupling shows Block chain is an expander



$d(x,y) =$ $4b+4$ $\quad$ $3b+4$ $\quad$ $2b+4-i$

(a) (b) (c)

- $$d(x,y) = b\#\text{ non-shared half-bisectors} + \#\text{ disagreeing quadrants}$$
- Case analysis shows that
$$E_{x,y}[d(X_1,Y_1)] \le \left(1-\frac{1}{17}\right)d(x,y)$$
- Result of M.-F. Chen shows that $\gamma \ge 1/17$

- Probability of a left-bisector is $A_{k-2}^2 / A_{k-1} \geq 1/2$
- A fixed sequence of moves will bring two configurations into agreement.
- $t_{\text{rel}} \leq t_{\text{mix}} \leq C$.

Loose ends/open questions

- What are the actual mixing times?
- Is there cut-off for small λ? Is the mixing time $n \log n$? (I think we can get $n \log^2 n$.)