Mixing Times of Self Organizing Lists and Biased Permutations

(Sarah Miracle)

Joint work with Prateek Bhakta, Amanda Streib and Dana Randall

Georgia Institute of Technology

Problem

- *Sampling permutations is a fundamental problem from probability theory.*
- *We study a Markov chain $\mathcal{M}_{nn}$ which samples from the permutation group on n elements (S_n) by iteratively making nearest neighbor transitions.*

$$\sigma_t = 8\ 1\ 5\ \mathbf{3}\ \mathbf{7}\ 4\ 6\ 2$$
$$\sigma_{t+1} = 8\ 1\ 5\ \mathbf{7}\ \mathbf{3}\ 4\ 6\ 2$$

Figure: A nearest neighbor transition.

- *We are given* **P** *which is a set of preferences $0 \le p_{i,j} \le 1$ for all $i \neq j$.*

The Markov Chain $\mathcal{M}_{nn}$

Starting with an initial permutation, we

- *Pick an index $1 \le i \le n-1$ uniformly*
- *Take elements s, t at indices $i, i+1$. Place s in front of t with prob. $p_{s,t}$, and t in front of s with prob. $p_{t,s} = 1 - p_{s,t}$.*

- *$\mathcal{M}_{nn}$ converges to the distribution:*

$$\pi(\sigma) = \frac{\prod_{i<j:\sigma(i)>\sigma(j)} \frac{p_{j,i}}{p_{i,j}}}{Z}.$$

- *It is related to the "Move-Ahead-One Algorithm" for self-organizing lists.*
- *Mixing rates are only known for a few special cases of* **P**.

Previous Work

Uniform Bias:

- *Wilson [9] showed $\mathcal{M}_{nn}$ mixes in $\Theta(n^3 \log n)$ in the unbiased case $p_{i,j} = 1/2$ for all i, j.*

Constant Bias:

- *Benjamini et al. [1] showed $\mathcal{M}_{nn}$ mixes in $\Theta(n^2)$ when we have constant bias - $p_{i,j} = p$ for all i, j.*
- *Greenberg et al. [4] generalized the above to higher dimensions using a simpler proof.*

Linear extensions of a partial order:

- *Bubley and Dyer [2] showed $\mathcal{M}_{nn}$ mixes in $O(n^3 \log n)$ when $p_{i,j} = 1/2$ or 1 in the context of linear extensions of a partial order.*

Conjecture [Fill]

If **P** *satisfies a "regularity" condition, then the spectral gap is minimized when*

$$p_{i,j} = \frac{1}{2} \forall i, j.$$

- *Fill [3] proposed the above conjecture and proved it for $n \ge 4$.*
- *If $p_{i,j} \ge 1/2 \forall i < j$ we say the* **P** *is* ***positively biased***.
- *It is easy to construct slow examples when the p_{ij}'s aren't positively biased.*

Our Work

- *We show that $\mathcal{M}_{nn}$ can be slow even when* **P** *is* ***positively biased***.
- *We show fast mixing results for two large classes of* **P** *that both generalize the* ***constant bias*** *case.*

Fast Mixing Results

"Choose Your Weapon"

Theorem

If $i < j$, $p_{i,j} = r_i$ for some r_i, then $\mathcal{M}_{nn}$ is rapidly mixing

- *Here, $p_{i,j}$ only depends on the smaller element.*
- *We show that a Markov chain similar to $\mathcal{M}_{nn}$, $\mathcal{M}_{inv}$, mixes rapidly. $\mathcal{M}_{inv}$ is allowed to make moves that "hop" over smaller elements.*
- *We consider $\mathcal{M}_{inv}$'s effect on inversion tables [5, 8]. Inversion tables count the number of inversions of each element i.*

$$\sigma = 8\ 1\ 5\ 3\ 7\ 4\ 6\ 2$$
$$I(\sigma) = 1\ 7\ 2\ 3\ 1\ 2\ 1\ 0$$

Figure: The inversion table for a permutation.

- *There is a bijection between inversion tables and permutations.*
- *We show $\mathcal{M}_{inv}$ is fast on the space of inversion tables.*
- *Finally, we relate $\mathcal{M}_{nn}$ to $\mathcal{M}_{inv}$ using the Comparison theorem[6] to show that $\mathcal{M}_{nn}$ is fast.*

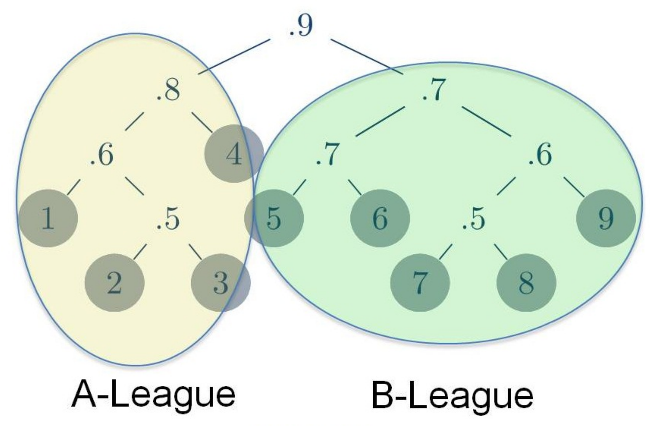
"League Hierarchies"

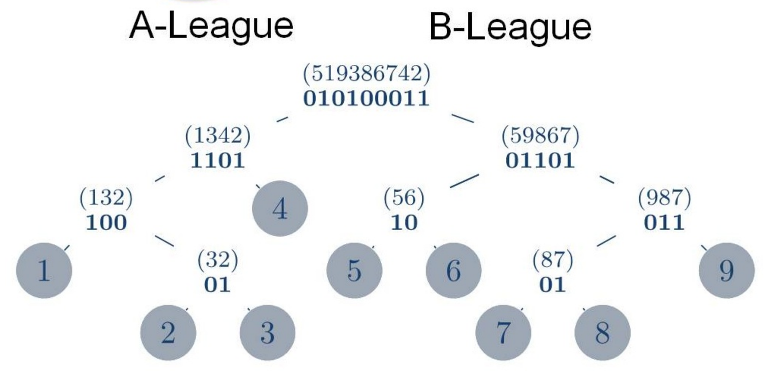
- *Say T is a binary tree with n leaf nodes. Each non-leaf node v of this tree has a value $\frac{1}{2} \le q_v \le 1$, and $p_{i,j} = q_{i \vee_T j}$. Then we say that* **P** *has* ***league structure*** *T.*
- *If $\forall i$, $p_{i,j} < p_{i,j+1}$ if $j > i$, we say that* **P** *is* ***weakly regular***.

Theorem

If **P** *has* ***league structure*** *and is* ***weakly regular***, *then $\mathcal{M}_{nn}$ is rapidly mixing*

- *League structure means that for at any node, $p_{i,j}$ is the same for any left descendant (A- League) with any right-descendant (B-League).*





Tree representation of 51938742

- *We consider a chain similar to $\mathcal{M}_{nn}$, $\mathcal{M}_{tree}(T)$, which is allowed to make moves that "hop" over elements in the same league.*
- *$\mathcal{M}_{tree}(T)$ can be shown to be fast over the space of tree representations, which are in bijection with permutations.*
- *Again, we relate $\mathcal{M}_{nn}$ to $\mathcal{M}_{tree}(T)$ using the Comparison theorem to show $\mathcal{M}_{nn}$ is fast.*

Slow Mixing Results

Biased Lattice Walk

- *We show that the* ***positive bias*** *condition is not sufficient by producing a class of* **P** *that mixes in exponential time, even though*

$$\forall i, j : p_{i,j} \ge 1/2.$$

- *We do this by reducing sampling weighted permutations to the problem of sampling biased lattice paths.*
- *Let* $p_{i,j} = \begin{cases} 1 & \text{if } i < j \in [0, n/2] \text{or} \in [n/2+1, n]; \\ 1-\delta & \text{if } i - j + 2n + 1 \ge n + M; \\ \frac{1}{2} + \frac{1}{n^2} & \text{otherwise.} \end{cases}$

Theorem

If **P** *is defined as above, then the Markov chain $\mathcal{M}_{nn}$ takes exponential time to converge.*

- *The first constraint causes the permutation to behave like a lattice walk. The remaining $p_{i,j}$ are the biases for each cell.*

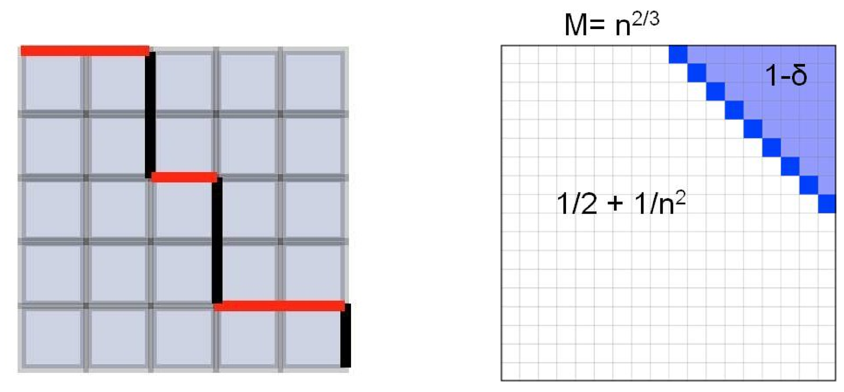
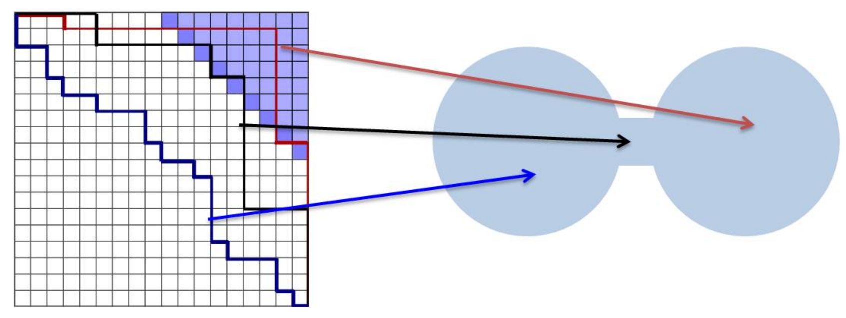


Figure: This Markov chain behaves like a walk with n/2 0's and n/2 1's. The bias in each cell varies as in the chart.

- *We show a bottleneck in the state space.*
- *There are* ***many states with low probability***, *and a* ***few states with high probability***. *In between, there is an exponentially small* ***cut*** *with few states of low probability.*



This cut in the state space shows us that the Markov chain is slow. [7]

DOE Office of Science Research Programs

- *Advanced Scientific Computing Research (ASCR): Applied Mathematics*
- *Advanced Scientific Computing Research (ASCR): Computer Science*

References

- *Itai Benjamini, Noam Berger, Christopher Hoffman, and Elchanan Mossel.* *Mixing times of the biased card shuffling and the asymmetric exclusion process.* Trans. Amer. Math. Soc, *2005.*
- *Russ Bubley and Martin Dyer.* *Faster random generation of linear extensions.* *In* Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms, *SODA '98, 1998.*
- *Jim Fill.* *Background on the gap problem.* Unpublished manuscript, *2003.*
- *Sam Greenberg, Amanda Pascoe, and Dana Randall.* *Sampling biased lattice configurations using exponential metrics.* *In* Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, *SODA '09, 2009.*
- *Donald E. Knuth.* The Art of Computer Programming, *volume 3: Sorting and Searching.* *Addison Wesley, 1973.*
- *Dana Randall and Prasad Tetali.* *Analyzing glauber dynamics by comparison of Markov chains.* Journal of Mathematical Physics, *41:1598–1615, 2000.*
- *Alistair Sinclair.* Algorithms for random generation and counting. *Progress in theoretical computer science. 1993.*
- *Silvio Turrini.* *Optimization in permutation spaces.* Western Research Laboratory Research Report, *1996.*
- *David Wilson.* *Mixing times of lozenge tiling and card shuffling markov chains.* The Annals of Applied Probability, *1:274–325, 2004.*