

Final Report for Period: 09/2008 - 08/2009**Submitted on:** 05/17/2011**Principal Investigator:** Randall, Dana J.**Award ID:** 0505505**Organization:** Georgia Tech Research Corp**Submitted By:**

Randall, Dana - Principal Investigator

Title:

Markov Chain Algorithms for Problems from Computer Science and Statistical Physics

Project Participants**Senior Personnel****Name:** Randall, Dana**Worked for more than 160 Hours:** Yes**Contribution to Project:****Post-doc****Graduate Student****Name:** Greenberg, Sam**Worked for more than 160 Hours:** Yes**Contribution to Project:****Name:** Streib, Amanda**Worked for more than 160 Hours:** Yes**Contribution to Project:**

During her second year in the PhD program, Amanda collaborated with me and Sam Greenberg on a couple of projects arising from nanotechnology applications of Markov chains.

Undergraduate Student**Technician, Programmer****Other Participant****Research Experience for Undergraduates****Organizational Partners****Other Collaborators or Contacts**

Collaborators over this period include: Nayantara Bhatnagar, Ivona Bezakova, Mary Cryan, Raissa D'Sousa, Martin Dyer, David Galvin, Cris Moore, Amanda Pascoe Streib, Vijay Vazirani, Eric Vigoda and Peter Winkler.

Activities and Findings**Research and Education Activities:**

We have made progress in this field by showing, for instance, that independent sets undergo a phase change as the fugacity of the system drops (SODA 2006). Moreover, we demonstrated that local chains for sampling 3-colorings become slow in sufficiently high dimensions; the dimension plays the role of the fugacity for this physical system (SODA 2007). Other results include analysis of Markov chains arising in combinatorics, statistical physics and algorithms for ad-hoc networks.

Over the last few years of the grant period I was very active in education by co-organizing a joint Georgia Tech - DIMACS special focus on 'Discrete Random Systems.' There were five workshops the first year, organized by Laci Lovasz and Benny Sudakov in October 2006, Milena Mihail, Ashish Goel, Fan Chung and Chris Wiggins in January 2007, by Gregory Sorkin and Eric Vigoda in March 2007, by Jim Fill (and a committee) in June 2007, and by Graham Brightwell, Tom Trotter and me in June 2007. In addition, we have held one working groups, organized by Eric Vigoda and me in March 2007, and a second is in Fall 2008. Later I organized a workshop with Fabio Martinelli and Prasad Tetali on "Analysis of Markov Chains on Combinatorial and Statistical Mechanics Models," at Georgia Tech in June 2009.

Findings:

We describe some of our most significant findings for this period.

1. Slow mixing of Glauber dynamics via topological obstructions:

Over the last two years, I've introduced a method for showing slow mixing based on the presence or absence of certain topological obstructions in the independent sets. The key new idea was redefining how to partition the state space into pieces to demonstrate that there is a bottleneck obstructing flow from one side to the other.

Instead of basing this on relative numbers of odd or even vertices in each independent set, as had become the norm for such proofs, the new method instead bases the partition on the presence of paths of vertices from one side of the region to the other. Then, using elementary arguments, we showed that Glauber dynamics will be slow for sampling independent sets in 2 dimensions when the fugacity $\lambda > 8.066$, improving on the best previous bound by a factor of 10. We also showed that it is slow on the torus when $\lambda > 6.1824$. (In SODA 2006.)

In work with my student Sam Greenberg, we extended the topological obstruction technique to considering an 'ocean' of sites comprising a fault line. With this new argument were able to show slow mixing of local Markov chains for independent sets on the triangular lattice, as well as showing that the so-called '8-vertex model' of statistical physics becomes slow beyond a certain fugacity (RANDOM 2007, to appear in Algorithmica). We have also applied these ideas to show the behavior of a model of DNA-based self-assembly is fast at high temperature and slow at low temperature (FNANO 2007, to appear in Theoretical Computer Science).

2. 3-colorings in high dimensional lattices:

There has been a lot of interest in understanding when local Markov chains are rapidly mixing on the set of k -colorings of a given graph. Likewise, in statistical physics there is interest in generating k -colorings on bipartite lattice graphs, in particular grid graphs. Here they arise in the context of the k -state Potts model, a fundamental model of antiferromagnetism. While we showed in previous work that local dynamics converge quickly in two dimensions, it was conjectured that they should converge slowly in sufficiently high dimensions.

David Galvin and I recently proved this, building on related work we are doing with Jeff Kahn and Greg Sorkin. The argument builds on sophisticated work of Galvin and Kahn showing the existence of multiple Gibbs states for independent sets in high dimensions. We used combinatorial arguments to show that the independent set comprised of the of

red points would behave similarly to a random independent set on the grid, and then adapted their previous argument. This appeared in SODA 2007.

3. The effect of boundary conditions on mixing rates of lattice models:

In the context of independent sets and colorings, we saw a change in mixing rates depending on a parameter λ that is inversely related to temperature: at high temperatures systems are well-behaved and sampling is efficient, while at low temperatures they are rigid and slow. In addition, it is believed that even at low temperature, the rate of convergence is strongly dependent on the environment in which the spin system is placed: we expect that there are boundary conditions such that one causes a local chain to mix quickly while others can cause the same chain to mix slowly. This had been verified previously only for spin systems on trees.

With my students Bhatnagar and Greenberg (RANDOM 2006), we defined a local chain for Ising configurations on the Cartesian lattice and show that with one boundary condition it will mix in polynomial time, while for another it will mix exponentially slowly. We show an analogous result for matchings on the square-octagon lattice, where, strikingly, the two boundaries differ only by the inclusion or exclusion of four vertices. These are the first rigorous proofs of this phenomenon on lattice graphs.

4. Many combinatorial and physical models on lattices have a height function representation, including several 2-dimensional tilings and 3-colorings on \mathbb{Z}^d . These height functions map a d -dimensional configuration to a surface in $(d+1)$ dimensions that has well-defined local properties. For example, lozenge tilings on the triangular lattice correspond to a monotonic surface on the 3-dimensional Cartesian lattice that looks like the upper envelope of a set of cubes, where each cube is supported by three other cubes. Previously local Markov chains have been studied to uniformly generate these objects, and it is known that they converge quickly in 2 and 3 dimensions.

Recently there has been a lot of interest in biased version of this chain where we are more likely to add unit cubes than remove them. It was previously known that the 2-dimensional chain is rapidly mixing for any constant bias, and that the 3-dimensional chain is rapidly mixing when the bias is $O(n^3)$. With my students Sam Greenberg and Amanda Pascoe we introduced a much simpler method for analyzing these biased chains, and show that for all biases in 2-dimensions (including when the bias is less than a constant), as well as any bias exceeding d^2 in d dimensions, the local chain is rapidly mixing. This paper was just submitted.

5. We have also made progress on various applications of Markov chains in computing. In work with Bhatnagar, Vazirani and Vigoda, we showed how to sample 'bichromatic perfect matchings' where the input graph has edges that are one of two colors and we specify how many edges of the perfect matching should be each color. We apply this to sampling perfect matchings on lattice graph regions, where vertical edges are one color and horizontal the other, a problem of interest in physics.

In work with Cryan and Dyer, we considered generalizations of the problem of sampling contingency tables to the 'cell-bounded' case. In the standard problem, row and column sums of a non-negative, integer valued matrix are given as input, and the goal is to count (or sample) from the set of consistent tables. Polynomial-time algorithms are only known when there are a constant number of rows, or when the row and column sums are all large. Cell-bounded tables generalize the standard model, and are interesting partly because this formulation is self-reducible (implying that counting and sampling are reducible to one another) whereas the standard formulation is not. We place the cell-bounded version on the same footing as the specialized standard model, and show that we can efficiently sample in exactly the same settings.

Training and Development:

The results achieved during this project was the major part of the Ph.D. dissertations of Nayantara Bhatnagar (graduated Summer 2007) and Sam Greenberg (graduated Summer 2008). In addition some results will be part of the dissertation of Amanda Pascoe, who has just completed her second year.

Throughout all of the Georgia Tech / DIMACS special focus activities I have been organizing, we have placed a strong emphasis on outreach to young researchers and underrepresented groups. Almost all of the workshops have included tutorials to help facilitate the inclusion of new researchers in the field.

Outreach Activities:

The Georgia Tech / DIMACS activities that are part of the special focus on 'Discrete Random Systems' I am organizing have encouraged participation from underrepresented groups and from young researchers. There have been many tutorials included in these workshops to help graduate students as well as to facilitate interdisciplinary interactions. I have given a keynote talk at FNANO 2007 (Foundations of NanoScience) giving an overview of Markov chain Monte Carlo methods and their relation to self-assembly based nanosystems.

I published a survey article on the Markov chain Monte Carlo method in Computing in Science and Engineering to help engineers and physicists understand the mathematical requirements of designing a provably efficient algorithm for sampling and related applications.

Journal Publications

Russell Martin and Dana Randall, "Disjoint decomposition of Markov chains and sampling circuits in Cayley graphs.", *Combinatorics, Probability and Computing*, p. 411, vol. 15, (2006). Published,

Dana Randall, "Rapidly mixing Markov chains with applications in computer science and physics.", *Computing in Science and Engineering*, p. 30, vol. 8, (2006). Published,

Mary Cryan, Martin Dyer and Dana Randall, "Approximately counting integer flows and cell-bounded contingency tables.", *SIAM Journal on Computing*, p. , vol. , (). Submitted,

Nayantara Bhatnagar, Dana Randall, Vijay Vazirani, Eric Vigoda, "Random birchomatic matchings", *Algorithmica*, p. 418, vol. 50, (2008). Published,

Sam Greenberg and Dana Randall, "Convergence Rates of Markov Chains for Some Self-Assembly and Non-Saturated Ising Models", *Theoretical Computer Science*, p. , vol. , (2008). Accepted,

Sam Greenberg and Dana Randall, "Slow Mixing of Markov Chains Using Fault Lines and Fat Contours", *Algorithmica*, p. , vol. , (2008). Accepted,

Books or Other One-time Publications

D. Randall and P. Winkler, "Mixing Points on an Interval", (2005). Book, Published
Bibliography: Second Workshop on Analytic Algorithmics and Combinatorics

M. Cryan, M. Dyer, and D. Randall, "Approximately counting integral flows and cell-bounded contingency tables", (2005). Proceedings, Published
Bibliography: 37th ACM Symposium on Theory of Computing

D. Randall and P. Winkler, "Mixing Points on a Circle", (2005). Proceedings, Published
 Bibliography: Proceedings of the 9th International
 Workshop on Randomization and Approximation Techniques in Computer
 Science

D. Randall, "Slow Mixing of Glauber Dynamics via
 Topological Obstructions", (2006). Proceedings, Published
 Bibliography: Proceedings on the 17th Symposium of Discrete Algorithms

N. Bhatnagar, D. Randall, V. Vazirani, E.
 Vigoda, "Random Bichromatic Matchings", (2006). Proceedings, Published
 Bibliography: 11th Latin American Theoretical Informatics

R. D'Sousa, D. Galvin, C. Moore and D.
 Randall, "Global Connectivity from Local Geometric Constraints in Sensor Networks for
 many Wireless Footprints", (2006). Proceedings, Published
 Bibliography: 5th International Conference on Information Processing in Sensor Networks

N. Bhatnagar, S. Greenberg, D. Randall, "The Effects of Boundary Conditions on Mixing Rates of Markov Chains", (2006). Proceedings,
 Published
 Bibliography: Proceedings of the 10th International
 Workshop on Randomization and Approximation Techniques in Computer
 Science

D. Galvin and D. Randall, "Torpid Mixing of Local Markov Chains on 3-Colorings of the Discrete Torus", (2007). Book, Published
 Bibliography: Proceedings of the 18th Symposium on Discrete Algorithms

D. Randall, "Markov Chain Convergence and the Efficiency of Some Self-Assembly Models", (2007). Proceedings, Published
 Bibliography: Proceedings of the 4th Annual Conference on Nanoscience

S. Greenberg and D. Randall, "Slow Mixing of Markov Chains using Fault Lines and Fat Contours", (2007). Proceedings, Published
 Bibliography: Proceedings of the 11th International
 Workshop on Randomization and Approximation Techniques in Computer
 Science

N. Bhatnagar, S. Greenberg and D.
 Randall, "Sampling Stable Marriages: Why Spouse-Swapping Won't Work", (2008). Proceedings, Published
 Bibliography: Proceedings of the 19th Symposium on Discrete Algorithms

S. Greenberg, A. Pascoe and D. Randall, "Sampling biased lattice configurations using exponential metrics", (2009). Proceedings, Published
 Bibliography: Proceedings of the 20th Symposium on Discrete Algorithms

A. Pascoe and D. Randall, "Self-assembly and Convergence Rates of Heterogenous Reversible Growth
 Processes", (2009). Proceedings, Published
 Bibliography: Proceedings of the 6th Annual Foundations of Nanoscience Conference

I. Bezaković, N. Bhatnagar and D.
 Randall, "On the Diaconis-Gangolli Markov Chain for sampling contingency tables with
 cell-bounded entries", (2009). Proceedings, Published
 Bibliography: Proceedings of the 15th International Computing and Combinatorics Conference

Web/Internet Site

Other Specific Products

Contributions

Contributions within Discipline:

The research has produced new methods for showing the inefficiency of certain local Markov chain algorithms, including 'fault lines' and 'fat faults.' This will help researchers produce similar results in the future, and it also helps guide the design of efficient algorithms by characterizing obstacles in fast mixing for certain classes of algorithms.

We have also made contributions that help expand the set of tools for showing fast mixing of Markov chains. In our work on sampling biased routings, we use a geometrically increasing distance function in order to argue that there is a good coupling. This approach is fine when the space is known to be contracting (i.e., that the expected change in distance between any pair of configurations is negative and bounded away from zero), but the standard coupling theorems are insufficient when the expected change in distance is very small or zero. This is because the maximal distance enters as a factor of the running time, and we require it to be exponentially large. We introduced a modification of the theorem that adds additional conditions under which the chain can still be proven to be rapidly mixing.

Contributions to Other Disciplines:

I presented a keynote address at FNANO 2007 (Foundations of Nanoscience) on the Markov chain Monte Carlo method, its mathematical foundations, and its connection to self-assembly based nanotechnologies.

I wrote a survey article on the design and analysis of Markov chain based algorithms for Computing in Science and Engineering intended for their Engineering and Physics readers.

I delivered the Wetterhahn Science Research Symposium at Dartmouth College May 2006. This is a lecture to all undergraduate students given each year by one woman in science.

I participated in several interdisciplinary National Academy of Sciences Annual Japanese-American Frontiers of Science Symposia from 2004-6 in Shonan Village, Japan and Irvine, CA. In 2005 I was on the planning committee and in 2006 I co-chaired the committee for the subsequent meeting.

Most of the research pursued in this period is motivated by problems in statistical physics, algorithms, and network systems. Our contributions have implications in each of these domains as well as computer science and mathematics.

Contributions to Human Resource Development:

I have worked closely with three Ph.D. students during this period. Nayantara Bhatnagar graduated Summer 2007 and is currently a visiting professor at U.C. Berkeley. Sam Greenberg graduated Summer 2008, and Amanda Pascoe has just completed her second year. They are all very promising students who have already made significant contributions to the discipline.

I actively participated in activities to increase the participation on women in computing and mathematics. In 2005 I delivered a couple of Invited lectures on Markov Chains at the Mathematics REU Program for undergraduates at Clemson University. In 2007 I participated as a panelist and speaker at Panelist for 'OurCS: Opportunities for Undergraduate Research in Computer Science' at CMU, Pittsburgh, PA. This program was to help encourage undergraduate women to continue with graduate studies.

In 2005 I wrote a book review 'Summing up a Set of Inspirational Female Figures,' for the Times Higher Education Supplement (THES) on

the book, 'Complexities: Women in Mathematics.'

Contributions to Resources for Research and Education:

The Georgia Tech / DIMACS special focus on 'Discrete Random Systems' has been designed to facilitate interactions between researchers across disciplines, and to encourage participation of young researchers and underrepresented groups.

Contributions Beyond Science and Engineering:

Conference Proceedings

Categories for which nothing is reported:

Organizational Partners

Any Web/Internet Site

Any Product

Contributions: To Any Beyond Science and Engineering

Any Conference