

# Phase Transitions Workshop Program

DIMACS — Georgia Tech

19–23 March 2007

Atlanta, Georgia, USA

## Schedule of Talks (see Abstracts below)

### MONDAY, MARCH 19:

9:30 - 10:05	<b>Mark Jerrum</b>	Inapproximability of the Tutte polynomial
10:10 - 10:45	<b>Leslie Goldberg</b>	The Complexity of Weighted Boolean #CSP
Coffee break		
11:30 - 12:05	<b>Sebastien Roch</b>	Markov Models on Trees: Reconstruction and Applications
Lunch break		
2:00 - 2:35	<b>David Gamarnik</b>	Monomer-dimer model and a new deterministic approximation algorithm for computing a permanent of a 0,1 matrix
2:40 - 3:15	<b>Chandra Nair</b>	The correlation decay (CD) tree and spatial mixing in multispin systems
Coffee break		
4:00 - 4:35	<b>Juan Vera</b>	Randomly coloring planar graphs with fewer colors than the maximum degree

### TUESDAY, MARCH 20:

9:30 - 10:05	<b>Jennifer Chayes</b>	Convergent Dense Graph Sequences
10:10 - 11:10	<b>Béla Bollobás</b>	Percolation on sequences of dense graphs
Coffee break		
11:45 - 12:20	<b>Van Vu</b>	Monotonicity and Resilience of random graphs
Lunch break		
2:15 - 2:50	<b>Christian Borgs</b>	First to Market is not Everything: Phase Transitions in Preferential Attachment with Fitness
2:55 - 3:30	<b>Boris Pittel</b>	On a random graph evolving by degrees
Coffee break		
4:00 - 4:15	<b>Joel Spencer</b>	Sprinkling and the Giant Component
4:20 - 4:55	<b>Abie Flaxman</b>	Expansion and lack thereof in randomly perturbed graphs

WEDNESDAY, MARCH 21:

9:30 - 10:30	<b>Andrea Montanari</b>	Survey: Belief Propagation, Cavity Method and Pure Gibbs States in Combinatorial Problems
Coffee break		
11:00 - 11:35	<b>Thierry Mora</b>	Clustering and pairs of solutions in random Boolean problems
11:40 - 12:15	<b>Sekhar Tatikonda</b>	Mixing and clustering in message-passing schemes: A Gibbs measure view
Lunch break		
2:15 - 2:50	<b>Dimitris Achlioptas</b>	On the solution space geometry of random CSPs
2:55 - 3:30	<b>Silvio Franz</b>	Interpolation inequalities and self-averaging identities in Random k-Sat and other spin systems
Coffee break		
Free afternoon: Some excursion ideas – Aquarium, Piedmont Park and Botanical Gardens, Cyclorama		

THURSDAY, MARCH 22:

9:30 - 10:05	<b>Marek Biskup</b>	Anomalous heat kernel decay for random walk among random conductances
10:10 - 10:45	<b>Peter Winkler</b>	Coordinate Percolation
Coffee break		
11:30 - 12:05	<b>Jeffrey Steif</b>	Stochastic domination and the Ising model
Lunch break		
2:00 - 2:35	<b>Amin Coja-Oghlan</b>	Local limit theorems for the giant component
2:40 - 3:15	<b>Alan Frieze</b>	The cover time of random graphs
Coffee break		
4:00 - 4:35	<b>Yuval Peres</b>	Critical random graphs: diameter and mixing time

FRIDAY, MARCH 23:

9:30 - 10:05	<b>Santosh Vempala</b>	The Complexity of Volume Computation: Annealing vs Dispersion
10:10 - 10:45	<b>Luis Rademacher</b>	Dispersion of Mass and Lower Bounds for Randomized Algorithms
Coffee break		
11:30 - 12:05	<b>Daniel Stefankovic</b>	Adaptive Annealing: A Near-optimal Connection between Sampling and Counting
Lunch break		
1:30 - 2:05	<b>Prasad Tetali</b>	Random walks for the discrete log problem
2:10 - 2:45	<b>Martin Dyer</b>	Evolving a peer-to-peer network
Coffee break		

## Abstracts

**Dimitris Achlioptas** (UC Santa Cruz)

*On the solution space geometry of random CSPs*

For a number of random constraint satisfaction problems, such as random k-SAT and random graph coloring, we now have excellent estimates of the largest constraints-to-variables ratio for which they have solutions. These estimates imply that all known polynomial-time algorithms stop finding solutions much before solutions cease to exist. To understand the origin of this phenomenon we study the evolution of the structure of the space of solutions in random CSPs as constraints are added. In particular, we will survey both what physicists hypothesize happens in this evolution, and what has been rigorously proven so far. We will see that the rigorous results are both in agreement with the physics predictions and help demystify Survey Propagation, an extremely successful heuristic proposed by physicists for random CSPs.

**Marek Biskup** (UCLA)

*Anomalous heat kernel decay for random walk among random conductances*

I will consider the random walk on  $Z^d$  driven by a field of random i.i.d. conductances. The law of the conductances is bounded from above; no restriction is posed on the lower tail (at zero) except that the bonds with positive conductances percolate. The presence of very weak bonds allows the random walk in a finite box mix pretty much arbitrarily slowly. However, when we focus attention on the return probability to the starting point – i.e., the heat-kernel – it turns out that in dimensions  $d=2,3$  the decay is as for the simple random walk. On the other hand, in  $d \geq 4$  the heat-kernel at time  $2n$  may decay as slowly as  $o(1/n^2)$  and in  $d = 4$  as slowly as  $O(n^{-2} \log n)$ . These upper bounds can be matched arbitrarily closely by lower bounds in particular examples. Despite this, the random walk scales to Brownian motion under the usual diffusive scaling of space and time.

Based on joint works with N. Berger, C. Hoffman, G. Kozma and T. Prescott.

**Béla Bollobás** (Cambridge University and University of Memphis)

*Percolation on sequences of dense graphs*

**Christian Borgs** (Microsoft Research)

*First to Market is not Everything: Phase Transitions in Preferential Attachment with Fitness*

**Jennifer Chayes** (Microsoft Research)

*Convergent Dense Graph Sequences*

**Amin Coja-Oghlan** (Humboldt University)

*Local limit theorems for the giant component*

**Abraham Flaxman** (Microsoft Research)

*Expansion and lack thereof in randomly perturbed graphs*

Abstract: This talk will investigate the expansion properties of randomly perturbed graphs. These graphs are formed by, for example, adding a random 1-out or very sparse Erdős-Rényi graph to an arbitrary connected graph. When any connected  $n$ -vertex base graph is perturbed by adding a random 1-out then, with high probability, the resulting graph has expansion properties. When the perturbation is by a sparse Erdős-Rényi graph, the expansion of the perturbed graph depends on the structure of the base graph. The same techniques also apply to bound the expansion in the small worlds graphs described by Watts and Strogatz in [Nature 292 (1998), 440–442] and by Kleinberg in [Proc. of 32nd ACM Symposium on Theory of Computing (2000), 163–170]. Analysis of Kleinberg’s model reveals a phase transition: the graph stops being an expander exactly at the point where a decentralized algorithm is effective in constructing a short path. The proofs of expansion rely on a way of summing over subsets of vertices which allows an argument based on the First Moment Method to succeed.

**Alan Frieze** (CMU)

*The cover time of random graphs*

We give asymptotically precise estimates for the cover time of random graphs and digraphs drawn from various distributions: Random Regular Graphs;  $G(n, p)$ ; The Giant Component of  $G(n, p)$ ; The Preferential Attachment Graph; The Random Digraph  $D(n, p)$ .

Joint work with Colin Cooper.

**David Gamarnik** (MIT)

*Monomer-dimer model and a new deterministic approximation algorithm for computing a permanent of a 0,1 matrix.*

We construct a deterministic approximation algorithm for computing a permanent of a 0,1 matrix to within a multiplicative factor  $(1 + \text{eps})^n$ , for arbitrary positive constant  $\text{eps}$ . When the graph underlying the matrix is a constant degree expander our algorithm runs in polynomial time (PTAS). In the general case the running time of the algorithm is sub exponential  $e^{O(n^{2/3} \log^3 n)}$ . For the class of graphs which are constant degree expanders the first result is an improvement over the best known approximation factor  $e^n$ . Our results use a recent deterministic approximation algorithm for computing a partition function of a monomer-dimer model, and Jerrum-Vazirani expander decomposition method.

Joint work with Dmitriy Katz.

**Leslie Ann Goldberg** (University of Liverpool)

*The Complexity of Weighted Boolean #CSP*

This work gives a dichotomy theorem for the complexity of exactly computing the partition function of a weighted Boolean constraint satisfaction instance. The problem is parameterised by a fixed "constraint language" which is the set of weighted relations that may be used to constrain an instance. We show that the partition function can be computed in polynomial time if either

- (1) every weighted relation in the constraint language is "of product type", or
- (2) every weighted relation in the constraint language is "pure affine".

For every other fixed constraint language, computing the partition function is #P-hard.

Joint work with Martin Dyer and Mark Jerrum.

**Mark Jerrum** (Queen Mary)

*Inapproximability of the Tutte polynomial*

The (classical) Tutte polynomial of a graph  $G$  is a two-variable polynomial  $T(G; x, y)$  that encodes much information about  $G$ . The number of spanning trees in  $G$ , the number of acyclic orientations of  $G$  and the partition function of the  $q$ -state Potts model are all specialisations of the Tutte polynomial. From a complexity-theoretic point of view, "mapping the Tutte plane" amounts to determining the computational complexity of evaluating  $T(G; x, y)$ , given  $G$ , for each rational pair  $(x, y)$ . For exact computation, the mapping was done in detail by Jaeger, Vertigan and Welsh. For approximate computation (within specified relative error) much less was known. I'll report on recent work with Leslie Goldberg (Liverpool) that makes some progress in this direction.

**Andrea Montanari** (Stanford)

*Belief Propagation, Cavity Method and Pure Gibbs States in Combinatorial Problems*

Consider the uniform measure over satisfying assignments of a  $k$ -satisfiability formula, or the size-weighted measure over independent sets of a graph. Belief propagation and the cavity method provide heuristic estimates for the marginals of such distribution. They both are based on a remarkable assumption on the structure of the corresponding measure, prescribing a precise form of (approximate) factorization according to the underlying graph. I will discuss conditions for the assumption to hold. This will bring up the connection with pure states, and 'clusters' of solutions in constraint satisfaction problems.

**Thierry Mora** (Universit Paris-Sud)

*Clustering and pairs of solutions in random Boolean problems*

One of the most striking predictions of the statistical physics analysis of random constraint satisfactions problems is the existence of a clustering phenomenon, by which the space of solutions divides itself into many connected components far from each other. In order to investigate this geometrical property, we introduce the notion of  $x$ -satisfiability: a problem is  $x$ -satisfiable if there exist two solutions separated by a relative distance  $x$ , i.e. differing by a proportion  $x$  of variables. Using first and second moment methods, I will show how the existence of a gap in the distance spectrum can be proved rigorously in the  $k$ -satisfiability problem, for  $k$  large enough. This means that one can find pairs of solutions close to each other, and around  $x = 1/2$ , but that a whole range of intermediate distances is inaccessible. I will then show how to yield an arguably exact value of the  $x$ -satisfiability threshold in the simpler case of random  $k$ -XORSAT using message-passing techniques.

**Chandra Nair** (Microsoft Research)

*The correlation decay (CD) tree and spatial mixing in multispin systems*

**Yuval Peres** (Microsoft Research and UC Berkeley)

*Critical random graphs: diameter and mixing time*

Let  $C_1$  denote the largest connected component of the critical Erdős-Rényi random graph  $G(n, 1/n)$ . We show that, typically, the diameter of  $C_1$  is of order  $n^{1/3}$  and the mixing time of the lazy simple random walk on  $C_1$  is of order  $n$ . The latter answers a question of Benjamini, Kozma and Wormald. These results extend to clusters of size  $n^{2/3}$  of  $p$ -bond percolation on any  $d$ -regular  $n$ -vertex graph where such clusters exist, provided that  $p(d-1) \leq 1$ .

Joint work with Asaf Nachmias.

**Boris Pittel** (Ohio State University)

*On a random graph evolving by degrees.*

We consider a random (multi)graph process  $\{G_a(n, m)\}$  on a vertex set  $[n]$ , which is the special case of a more general graph process put forth by Laci Lovász a few years ago.  $G_a(n, 0)$  is empty, and  $G_a(n, m + 1)$  is obtained from  $G_a(n, m)$  by inserting a new edge  $e$  at random: given  $G_a(n, m)$ ,  $e$  joins two currently disjoint vertices,  $i$  and  $j$ , with probability proportional to  $(d(i) + a)(d(j) + a)$ , where  $d(i)$ ,  $d(j)$  are the degrees of  $i$ ,  $j$ , and  $a > 0$  is a fixed parameter. (The limiting case  $a = \infty$  is the Erdős-Rényi process.) We show that whp  $G_a(m, n)$  contains a unique giant component iff  $c = 2m/n$  exceeds  $c(a) = a/(a + 1)$ , and the size of the giant is asymptotic to  $n\{1 - [(a + c^*)/(a + c)]^a\}$ , where  $c^* < c(a)$  satisfies  $c/(a + c)^{2+a} = c^*/(a + c^*)^{2+a}$ . The multigraph  $MG_a(n, m)$  version becomes whp connected once  $m \gg n^{1+1/a}$ .

**Luis Rademacher** (MIT)

*Dispersion of Mass and Lower Bounds for Randomized Algorithms*

How much can randomness help computation? Motivated by this general question and by volume computation, one of the few instances where randomness provably helps, we analyze a notion of dispersion and connect it to asymptotic convex geometry. We obtain a nearly quadratic lower bound on the complexity of randomized volume algorithms for convex bodies in  $R^n$  (the current best algorithm has complexity roughly  $n^4$ , conjectured to be  $n^3$ ). Our main tools, dispersion of random determinants and dispersion of the length of a random point from a convex body, are of independent interest and applicable more generally; in particular, the latter is closely related to the variance hypothesis from convex geometry. This geometric dispersion also leads to lower bounds for matrix problems and property testing.

**Sebastien Roch** (UC Berkeley)

*Markov Models on Trees: Reconstruction and Applications*

Markov models on trees arise naturally in many fields, notably in molecular biology - as models of evolution; in statistical physics - as models of spin systems; and in networking - as models of broadcasting. In this talk, I will discuss various inference problems motivated especially by applications in statistical phylogenetics, i.e. the reconstruction of evolutionary histories of organisms from their molecular sequences. In particular, I will consider the "root reconstruction" problem: how accurately can one guess the value at the root of the tree, given the state at the leaves? I will focus on recent work establishing new conditions for the impossibility of such reconstruction. I will also discuss the related "phylogenetic reconstruction" problem: given enough samples at the leaves, can one reconstruct the tree that generated this data and, if so, how efficiently? I will present a recent result on a sharp transition in the number of samples required to recover the tree topology, using a connection to the root reconstruction problem above. Time permitting, I will describe briefly connections to computational learning theory and network tomography as well.

Joint work with S. Bhamidi, C. Borgs, J. Chayes, C. Daskalakis, E. Mossel, and R. Rajagopal.

**Joel Spencer** (NYU)

*Sprinkling and the Giant Component*

The uniqueness of the Erdős-Rényi giant component should be obvious as two giant components would be highly unstable. Despite a plethora of proofs, none seems to be easy. I still don't know an easy one but I'll discuss one that takes advantage of the high instability by using sprinkling.

**Daniel Stefankovic** (University of Rochester)

*Adaptive Annealing: A Near-optimal Connection between Sampling and Counting*

We present a near-optimal reduction from approximately counting the cardinality of a discrete set to approximately sampling elements of the set. Our result is set in the framework of approximating the partition function  $Z$  of a discrete system, such as the Ising model, the matchings, or colorings of a graph. The typical approach to estimating the partition function  $Z(\beta)$  at some desired (inverse) temperature  $\beta^*$  is to define a sequence, which we call a *cooling schedule*,  $\beta_0 = 0 < \beta_1 < \dots < \beta_\ell = \beta^*$  where  $Z(0)$  is trivial to compute and the ratios  $Z(\beta_{i+1})/Z(\beta_i)$  are easy to estimate by sampling from the distribution corresponding to  $Z(\beta_i)$ . Previous approaches required a cooling schedule of length  $O^*(\ln A)$  where  $A = Z(0)$ , thereby ensuring that each ratio  $Z(\beta_{i+1})/Z(\beta_i)$  is bounded. We present a cooling schedule of length  $O^*(\sqrt{\ln A})$ . For well-studied problems such as estimating the partition function of the Ising model, or approximating the number of  $k$ -colorings or matchings of a graph, our cooling schedule is of length  $O^*(\sqrt{n})$ , which yields an overall savings of  $O^*(n)$  in the running time of the approximate counting algorithm compared to previous results.

Joint work with Santosh Vempala and Eric Vigoda.

**Jeff Steif** (Chalmers)

*Stochastic domination and the Ising model*

(1). We show that the plus and minus states for the Ising model on  $Z^d$  dominate the same set of product measures. We show that this latter fact however completely fails on the homogenous 3-ary tree.

(2). While it is known that the plus states for different temperatures on  $Z^d$  are never stochastically ordered, on the homogenous 3-ary tree, almost the complete opposite is the case.

(3). On  $Z^d$ , the set of product measures which the plus state for the Ising model dominates is strictly decreasing in the interaction parameter.

(4). An FKG exchangeable finite process dominates a product measure if and only the relevant inequality holds for the event of all 0's.

(5). Extending these results to an amenable/nonamenable dichotomy would be interesting.

This is based on joint work with Tom Liggett.

**Sekhar Tatikonda** (Yale)

*Mixing and clustering in message-passing schemes - A Gibbs measure view*

The belief propagation algorithm and its variants have recently been shown to work quite well on hard constraint satisfaction problems. Survey propagation is one such algorithm that has been developed for solving the  $k$ -SAT problem. This algorithm is based on two hypotheses deriving from the replica symmetry breaking theory developed for the statistical mechanics of spin glasses. The first is that the solutions of constraint satisfaction problems near threshold organize into disjoint clusters. The second is that within each cluster there is spatial mixing. In this paper we show that both hypotheses hold for a large class of constraint satisfaction problems including the  $k$ -SAT problem. Our approach is based on examining an appropriate infinite site Gibbs measure. Each cluster corresponds to a different extremal measure. Thus the so-called greedy threshold is equivalent to the Gibbs measure uniqueness threshold. Below this threshold there is one cluster and above it there are many.

**Prasad Tetali** (Georgia Tech)

*Random walks for the discrete log problem*

We analyze the classical Pollard's Rho algorithm for finding the discrete logarithms in a cyclic group  $G$ . We prove that, with high probability, a collision occurs (and the discrete logarithm found) in  $O(\sqrt{|G|} \log |G| \log \log |G|)$  steps, by analyzing an appropriate (nonreversible, nonlazy) random walk on a discrete cycle of odd length.

This is joint work with Jeong Han Kim and Ravi Montenegro

**Juan C. Vera** (Georgia Tech)

*Randomly coloring planar graphs with fewer colors than the maximum degree*

We study Markov chains for randomly sampling  $k$ -colorings of a graph with maximum degree  $D$ . We prove the first results, for a general class of graphs, showing fast convergence of such chains when the number of colors  $k \ll D$ . For planar graphs, we prove polynomial mixing time of the Glauber dynamics when  $k > D/\log D$ . A challenging aspect of the proof is when  $D$  is constant and  $k < D$ , even in a random coloring, a constant fraction of the vertices are "frozen" with a unique color not present among its neighbors.

Joint work with Thomas P. Hayes and Eric Vigoda.

**Santosh Vempala** (Georgia Tech)

*The Complexity of Volume Computation: Annealing vs Dispersion*

Computing the volume of a convex body in  $n$ -dimensional space is a basic problem that has led to many nice techniques. In this talk, we survey the algorithmic and geometric ideas behind the most recent developments in sampling, integration and optimization over convex bodies and more generally, logconcave functions. In particular, we will discuss the analysis of the random walk called "hit-and-run", and simulated annealing in the continuous setting which together lead to the current best algorithms for both integration and optimization. We'll briefly touch upon possible future improvements given recent dispersion-based lower bounds.

Parts of the talk are joint work with Laci Lovász and with Luis Rademacher.

**Van Vu** (Rutgers)

*Monotonicity and Resilience of random graphs*

(1) Does it make sense to talk about thresholds for random regular graphs? In other words, is the random regular graphs model monotone? (If one tells us that a random  $n^{1/2}$ -regular graph has a.s. a monotone property  $P$ , can we conclude that a random  $n/2$ -regular graph has a.s. the same property?)

(2) The resilience of a graph towards a specific property. Assume that  $G$  has a property  $P$ , how much should we change  $G$  in order to destroy  $P$ ? Different notions of resilience: Global and Local resilience. The resilience of a random graphs with respect to the most basic properties.

(3) The relation between (1) and (2).

(4) The resilience of other random structures. Open questions.

Based on joined work with J. H. Kim and B. Sudakov.

**Peter Winkler** (Dartmouth)

*Coordinate Percolation*

In coordinate percolation, the life or death of a site is determined by events associated with the lines or planes through the point. We will review some examples and show some results (with Lizz Moseman) on a particularly tractable form.