Frontiers of Statistical Mechanics and Theoretical Computer Science

Tyler Helmuth (Durham University)
Jane Gao (University of Waterloo)
Marcus Michelen (University of Illinois, Chicago)
Will Perkins (Georgia Tech)
Gerónimo Uribe Bravo (Universidad Nacional Autónoma de México)

August 11 – 16, 2024

1 Overview

Ties between statistical mechanics and theoretical computer science date back to the earliest days of mechanical computers. A revolutionary example is the development of Markov Chain Monte Carlo methods: these were first employed to approximately determine the equation of state of a gas of hard spheres. Recently there has been intensive activity investigating fundamental connections between these two subjects. In some cases this has been driven by recent discoveries (e.g., the connection between physical and computational notions of phase transitions), while in other it is driven by mathematicians, statisticians, and computer scientists attempting to make precise non-rigorous predictions of practicing physicists (e.g., spin glass theory; perceptrons).

The interplay of these subjects has lead to fruitful exchanges of tools between different mathematical communities, and the primary goal of this meeting was to encourage further developments along these lines. Towards this the meeting brought together researchers in combinatorics, probability, statistical mechanics, and theoretical computer science to share methods and perspectives for problems of common interest. In-person meetings of extended duration are uniquely helpful for this goal, as they allow for extended interactions that can overcome some challenges (e.g., terminological barriers). One of the major highlights of the workshop was the success of three overview lectures, each delivered by a different scientist, with the aim of establishing a cooperative and collaborative base for the rest of the workshop. These lectures were all delivered on the first day of the workshop, and we summarize them below in Section 2.1.

A second major highlight of the workshop was a series of lightning talks given by (almost) all workshop participants that did not give a full-length talk. These talks were rapid fire (four minutes each), arranged in two one-hour blocks on the first day of the workshop. One reason this was a highlight was the social and collaborative benefit: everyone was quickly introduced to one another, and quickly gained an idea of common interests and topics that could be discussed. A second reason was that the rapid sequence of talks gave a sense of how the work of the participants fit together, and how past progress is driving current research. We describe some specific instances of this (without attempting to be exhaustive) in Section 2.2 below.

During the remainder of the week a selected group of participants gave lectures about recent progress they have made. The talks all concerned various probabilistic models, most of which were (or could be) motivated by physical considerations. The range of perspectives on these models, however, was broad – from

purely probabilistic to combinatorial to algorithmic to physical. Section 3 summarizes these lectures. Some concluding remarks are contained in Section 4, and Section 5 contains some comments on the workshop demographics.

2 Highlights of the Workshop

2.1 Summary of Overview Lectures

2.1.1 Matthew Jenssen: Applications of the Cluster Expansion

Jenssen gave a survey talk about recent applications of a classical tool from statistical physics, the cluster expansion, to the fields of probabilistic, extremal and enumerative combinatorics and approximate counting and sampling in computer science, thus connecting all three of the topic areas of the workshop.

The main setting Jenssen used was that of the hard-core model (or the hard-core lattice gas). For a graph G and an activity parameter $\lambda \geq 0$, the hard-core model is the probability distribution $\mu_{G,\lambda}$ on the set $\mathcal{I}(G)$ of independent sets (vertex subsets containing no edges) of G defined by

$$\mu_{G,\lambda}(I) = \frac{\lambda^{|I|}}{Z_G(\lambda)}$$

where

$$Z_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|}$$

is the partition function (or normalizing constant) of the model; in graph theory the partition function is also called the independence polynomial of G. The hard-core model generalizes to hypergraphs; in this setting independent sets are subsets of vertices that do not contain any hyperedges.

A widely studied setting in statistical physics is that in which G_n is a sequence of growing subgraphs (say, boxes) in $\mathbb{Z}^d$, $d \geq 2$. Via a limiting procedure one can then define the hard-core model on the infinite graph $\mathbb{Z}^d$. The main question is that of phase transition and phase coexistence. For a given value of λ, is there a unique limiting distribution in which vertices of the even and odd sublattices of $\mathbb{Z}^d$ are equally likely to be included in the independent set or are there multiple limiting measures possible, with one favoring independent sets supported primarily on the even sublattice and the other on the odd sublattice? This question, along with many other questions about the model, can be answered by computing (approximately) the partition function $Z_G(\lambda)$ as λ varies. For instance, following the Lee–Yang framework, phase transitions are marked by non-analyticities of the free energy,

$$f_d(\lambda) = \lim_{G_n \to \mathbb{Z}^d} \frac{1}{|V(G_n)|} \log Z_{G_n}(\lambda)\,.$$

Computing or approximating partitions functions like $Z_G(\lambda)$ arises in two other fields. In combinatorics, $Z_G(\lambda)$ is a weighted version of counting; $Z_G(1)$ counts the total number of independent sets in a graph. Many important combinatorial objects can be encoded as the set of independent sets of a graph or hypergraph; the example Jenssen used was the set of triangle-free graphs on n vertices. This can be represented as the set of independent sets of a 3-uniform hypergraph in which the $\binom{n}{2}$ vertices are the edges of the complete graph K_n and the $\binom{n}{3}$ hyperedges are the sets of three edges forming triangles. An independent set in this hypergraph is a set of edges forming no triangles, or in other words a triangle-free graph.

In algorithms, two important tasks (in machine learning, computational statistics, physics and more) are to approximate a partition function like $Z_G(\lambda)$ given G as an input and to sample (approximately) from the probability distribution $\mu_{G,\lambda}$. These tasks are generally NP-hard, but in many cases efficient algorithms exist. One of the main approaches is Markov Chain Monte Carlo and randomized algorithms (described more in the next section), but one can also ask for efficient deterministic approximate counting algorithms. The tractability of approximate counting and sampling has close connections to statistical physics phase transitions; for example, the NP-hardness threshold (in λ) for sampling from the hard-core model on the family of graphs of maximum degree Δ coincides with the phase transition on the infinite Δ-regular tree [26, 23].

Thus in all three fields, techniques to compute or approximate $Z_G(\lambda)$ are of great importance, though the kinds of graphs considered often differ considerably between these fields.

One important tool for evaluating $Z_G(\lambda)$ in statistical physics is the *cluster expansion*, dating back to the work of Mayer in the 1930s. In the setting of the hard-core model, the cluster expansion is simply the Taylor series expansion of $\log Z_G(\lambda)$ around $\lambda = 0$; it is an infinite series in powers of λ. Crucially, the coefficients of this power series can be expressed in terms of counts of connected subgraphs of the graph G. The key to applying the cluster expansion is to verify convergence. In general, it will converge for small λ (up to the radius of convergence of the power series) but diverge for larger λ. The cluster expansion has been widely used in statistical physics to understand the behavior of equilibrium lattice models.

Jenssen's overview talk focused on more recent uses of the cluster expansion in combinatorics and computer science. The cluster expansion has a strong connection to combinatorics through the Lovász Local Lemma, as discovered by Shearer [22] and elucidated by Scott and Sokal [21]. The applications Jenssen discussed are more recent and involve the use of the cluster expansion in asymptotic enumeration.

Using the cluster expansion to count at first seems surprising because when $\lambda = 1$ and $Z_G(\lambda)$ counts the number of independent sets of G, the cluster expansion cannot converge (even for an empty graph one needs $|\lambda| < 1$ for convergence). However, if the graph G has some structure – e.g. it is a lattice or a bipartite expander graph or the Hamming cube, one can rewrite $Z_G(\lambda)$ in terms of 'defects' from the largest independent sets. If these defects are rare, they will have a small activity. For example, on a bipartite graph with some structural properties one might expect most independent sets to lie predominantly in one of the partition sets. Proving the defects (deviations from entirely occupying one partition set) have small activities is possible. This perspective is not new to statistical physics; it is closely related to the use of cluster expansion within Pirogov–Sinai theory to understand low-temperature behavior of spin models. But this is a new perspective in combinatorics and algorithms, giving precise asymptotic enumeration results [14, 12] and efficient approximate counting and sampling algorithms in low-temperature settings [9, 13].

Jenssen's talk set the stage for several later talks in the workshop, including those discussing sampling algorithms for the hard-core model and those discussing graph and hypergraph container lemmas.

2.1.2 Shayan Oveis Gharan: The MCMC Method and High Dimensional Expanders

The Markov Chain Monte Carlo (MCMC) method is one of the oldest in computer science. The general setting is when one wishes to sample from a probability distribution μ that has a support that is far too large to enumerate. A very basic real-world example is when μ is the uniform distribution on permutations of 52 – i.e., the case of playing cards – and so one approximately samples from μ by shuffling many times. The MCMC method consists of devising a Markov chain whose stationary distribution is given by μ and showing that the Markov chain mixes quickly in some sense.

MCMC is a now-standard approach within theoretical computer science and dates back to a groundbreaking 1953 work of Metropolis-Rosenbluth-Rosenbluth-Teller-Teller [18]. The setting of Metroplis et al. was the so-called *canonical hard sphere gas*, i.e., the uniform distribution on sphere packings of a fixed density in two-dimensional euclidean space. The hard sphere gas is an old and well-studied model within statistical physics (a closely-related model was the subject of Daniel Hadas's talk).

One of the most classical approaches to showing a bound on the mixing time of a Markov chain is via its spectral gap: if the transition kernel is given by a matrix P with eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ then its spectral gap is $1 - \lambda_2$ and so one has that the mixing time starting from v is upper bounded by $\log(1/\pi(v))/(1 - \lambda_2)$ where π is the stationary distribution. A standard example is when one considers a random walk on a graph G that is a λ-expander, in which case one has that the mixing time is bounded above by $\log n/(1 - \lambda)$.

In many applications, the Markov chain may be viewed as a random walk in a higher dimensional setting, e.g. on subsets of a given size. The subject of Oveis Gharan's talk was how to extend the notion of expanders to higher dimensional settings, and how high-dimensional expansion (when correctly defined) leads bounds on the mixing time of an associated Markov chain. Abstractly, this is done as follows. A simplicial complex X with ground elements $[n]$ is a collection of subsets of $[n]$ that is downward closed, meaning that if $A \in X$ and $B \subset A$ then $B \subset X$. A complex is d-dimensional if $\max_{A \in X} |A| = d + 1$. Given a probability measure π on the sets in X with exactly $d + 1$ elements, one defines a Markov chain that transitions from one set $S \in X$ to another by uniformly removing an element of S and then adding a random element with

an appropriate probability. This general framework captures many Markov chains from theoretical computer science, probability theory, and statistical physics such as the clique walk and Glauber dynamics.

Oveis Gharan introduced the notion of a high-dimensional expander along with pieces of the machinery for using high-dimensional expansion: he demonstrated how high-dimensional expansion of X implies rapid mixing of the associated Markov chain described above and discussed the "trickledown theorem," a general approach for showing high-dimensional expansion. This approach is closely related to the recent development of the theory of *spectral independence* introduced by Anari, Liu and Oveis Gharan [2]. One of the major applications of spectral independence was to show fast-mixing for Glauber dynamics for the hard-core model on bounded degree graphs. The hard-core model is a central model in statistical physics, combinatorics, and computer science and was a recurring theme across multiple talks in the workshop (e.g. Jenssen's tutorial, Hadas's talk, and Davies's lightning talk). Beyond the hard-core model, high-dimensional expanders and spectral independence have found many applications throughout computer science and the analysis of models from statistical physics, including in the lightning talks of Yap and Jain on the fixed-energy Ising model and the canonical hard-core model respectively.

Oveis Gharan's talk also gave an overview of several general approaches to establishing spectral expansion and/or spectral independence including trickledown theorems, correlation decay, coupling arguments, and approaches from the geometry of polynomials. This array of approaches has led to interdisciplinary work on MCMC due to the different directions one can approach spectral expansion and mixing-time bounds more generally. The emphasis on the interdisciplinary nature of MCMC drew connections to various works discussed at the workshop. In the lightning talks for instance, we saw algorithmic bounds proven via complex-analytic techniques by Regts, coupling techniques by Carlsen, and correlation decay by Mani. Oveis Gharan's talk described how each of these techniques and approaches can combine with the framework of high dimensional expanders in order to prove robust mixing time bounds in quite general settings.

2.1.3 Mark Sellke: Algorithmic Spin Glass Theory

The field of optimization consists of solving problems of the form

$$\min_{\sigma \in \Sigma} H(\sigma) \tag{1}$$

for some function H and some set Σ. In statistical mechanics, when H is the Hamiltonian of a system, the optimizers of problem (1) are the *ground states*, i.e., configurations of minimal energy. Much of the classical theory of optimization makes stringent assumptions on H and Σ, for instance that Σ is a convex set and H is a convex function. A main feature in the convex setting is that H typically has only one local minimum, in which case there are often a host of techniques to tackle (1).

The relevant setting for Sellke's talk is in the case when H is random and far from convex. In the setting of *glassy models*, the function H often is indexed by a parameter N and the objective function H_N has exponentially many local minima. Here, there is often no hope of solving (1) in the worst-case setting for H as the problem can be NP-hard. However, there has recently been much development of algorithms for finding the optimizers to (1) in the random (typical) case.

The specific model that is the subject of Sellke's talk is when one has either $\Sigma = \{-1, 1\}^N$ or $\Sigma = \mathbb{S}^{N-1}(\sqrt{N})$—i.e. the $N-1$ dimensional sphere of radius $\sqrt{N}$—and

$$H_N(\sigma) = \sum_{p=2}^{P} \frac{\gamma_p}{N^{(p-1)/2}} \langle \Gamma^{(p)}, \sigma^{\otimes p} \rangle \tag{2}$$

where $\gamma_j \geq 0$ are constants and $\Gamma^{(j)}$ are rank-j tensors of independent standard Gaussian random variables. This is a *mixed p-spin glass*. The maximum value in (1) turns out to be described by *the Parisi formula* which was predicted by Parisi [20] and proved mathematically by Guerra [7] and Talagrand [25]. For questions of sampling and statistical physics, one looks at the *Gibbs measure* where one assigns probability proportional to $\exp(-\beta H_N(\sigma))$ to each state σ. As β increases to infinity, this assigns all probability to the optimizers of (1), while at $\beta = 0$ the Hamiltonian H_N plays no role and the Gibbs measure is the same as a uniformly random choice of $\sigma \in \Sigma$. This already presents a connection between the two seemingly-disparate fields: in statistical mechanics, the Gibbs measure is the equilibrium measure for a given energy function at a given temperature; in computer science, it provides a softening of an optimization problem.

Sellke's talk concerned the problem of sampling the Gibbs measure for mixed p-spin glasses and finding configurations $\sigma \in \Sigma$ with a given energy level. Work on this area tends to first begin with getting a heuristic or rigorous understanding of the geometry of the solution space: depending on β, what do typical samples from the Gibbs measure look like? Similarly, for which x can one find a solution to $H_N(\sigma) \leq x$? When β is small or x is large, the solution space is connected and one expects that one can sample with an efficient algorithm. As β increases or x decreases, the solution set shatters into exponentially many clusters and one expects that sampling should not be possible.

A breakthrough work of Subag [24] introduced an algorithm to the spin glass literature known as *Hessian ascent* which tackles the former regime: when things are sufficiently random-like (e.g. when x is large enough), one can find configurations with $H_N(\sigma) \leq x$. Recent work of Huang and Sellke [10] showed a complementary *hardness* result. Subag's work finds an algorithm up to an energy level x_{ALG}, and Huang and Sellke showed that Lipschitz algorithms – i.e., algorithms whose outputs only have small changes under small changes to the input Γ in (2) – *cannot* find a configuration σ with $H_N(\sigma) \leq x_{\mathrm{ALG}} - \varepsilon$ for any $\varepsilon > 0$. This is done by showing that the solution space $\{\sigma \in \Sigma : H_N(\sigma) \leq x_{\mathrm{ALG}} - \varepsilon\}$ exhibits a geometric property called the *branching overlap gap property* and subsequently showing that any problem with such solution space cannot be solved with a Lipschitz algorithm. Similarly, a class of algorithms for approximate sampling of the Gibbs measure was ruled out by El-Alaoui, Montanari and Sellke [5] for the analagous sampling problem in a certain regime of β.

While much of the content of Sellke's talk focused on the case of spin glasses, a similar story is expected to hold for various other glassy models which appear in probability theory, combinatorics, computer science, statistics and statistical physics. Some key examples are spin glasses (as covered in this talk as well as the talk of Reza Gheissari); the random perceptron model (see also the talks of Brice Huang and Shuangping Li as well as the lightning talk of Dylan Altschuler); tensor principal component analysis (which appeared in the talk of Alex Wein); random graph matching (the focus of Gerandy Brito's lightning talk); independent sets in random graphs (an algorithm for which was provided in Shayan Oveis-Gharan's tutorial lecture); and many other models.

Sellke's tutorial lecture provided a description not only of the state-of-the-art for the specific case of mixed p-spin glasses but also a road-map for the development of a work on glassy models. A sociological feature common to many of these models is that often early work on the model begins with heuristics and predictions by physicists which guide more mathematically rigorous work on the models for decades to come. Due to the wide class of applicable problems, often certain problems are more popular among computer scientists or statisticians or combinatorialists despite the techniques having commonality. Sellke's tutorial provided a unified bird's-eye-view of algorithmic work on glassy models by using his work on spin glasses as an extended example.

2.2 Summary and Highlights of Lightning Talks

Every participant that was not giving a full-length talk was invited to give a four-minute lightning talk on the first day. Over two one-hour sessions the talks given were:

- Corrine Yap: Fast and slow mixing of Kawasaki dynamics
- Dylan Altshuler: Zero-One laws for random feasibility problems
- Ewen Davies: Algorithms for independent sets of fixed size
- Gourab Ray: Structure of Gibbs measures of integer Lipschitz functions on trees
- Nitya Mani: Strong spatial mixing for colorings on trees and applications to sampling
- Vishesh Jain: Optimal mixing of the down up walk for independent sets of fixed size
- Laura Eslava: Random recursive trees
- Elizabeth Collins-Woodfin: Stochastic gradient descent in high dimensions
- Gerandy Brito: Random graph matching

- Guus Regts: Approximately counting proper colorings
- Candida Bowtell: Transversals in random latin squares
- Holden Lee: Sampling from the continuous random energy model in TV distance
- Joseph Chen: Entropic repulsion in the SOS model
- Sarah Cannon: Pirogov-Sinai theory beyond lattices
- Edward Zeng: The integer-valued Gaussian free field in three dimensions
- Charlie Carlson: Sampling colorings (with Markov Chains)
- Raimundo Briceño: Approximability of sofic entropy
- Thuy-Duong (June) Vuong: Fast algorithms for sampling
- Sara Hernndez-Torres: 3D loop-erased random walks

While the talk titles give some sense of the scope of the workshop, they leave many connections hidden. We highlight a few of these connections here to give a sense of how progress has been achieved, and to emphasize the common themes and tools of the mathematical communities involved in the meeting.

1. The present workshop developed out of two online meetings organized by T. Helmuth and W. Perkins in 2020 and 2021. Following discussions at the first of these meetings, an understanding of computational hardness of approximate counting for independent sets of fixed size was developed by Davies and Perkins [4]. The lightning talks of Jain and Yap reported on further developments in this area [11, 16]. Later in the meeting we heard (in the talk of Randall) about other problems in which similar conceptual and technical issues concerning canonical distributions (instead of grand-canonical distributions) arise.

2. Ray's talk about Lipschitz functions on trees highlighted phenomena familiar to those who have studied the hard-core model: the presence of different Gibbs measures reflecting the parity at which the tree is truncated. This was an excellent prelude to Park's talk for participants more familiar with the hard-core model/algorithms. Park's talk in turn made use of the cluster expansion, an approach detailed in Jenssen's tutorial.

3. The lightning talks presented various perspectives on closely related questions, particularly concerning graph colorings. The talks of Carlson and Mani concerned randomized approximation algorithms for these problems, while the talk of Regts concerned deterministic approximation algorithms. These methods exploit different different statistical-mechanical properties of colourings: randomized algorithms use as input some type of correlation decay, while deterministic algorithms utilize zero-freeness of the partition function. The equivalence of these notions is not fully understood and this leads to an interesting link with an important question in computing (the power of randomness).

4. Bowtell's talk presented a wholly new (and seemingly challenging) problem, namely understanding the typical structure of latin squares. Despite this, connections to other talks were present: the absorber approach used by Bowtell is a descendant of the "Rödl nibble" (called "sprinkling" in percolation theory). Sprinkling played an important role in Kotecký's talk, and originated in work of Ajtai-Komlós-Szemerédi on Sidon sequences.

3 Scientific Progress Made

To give an indication of the scientific progress made at the meeting, we summarize the main developments that were formally discussed. Further progress occurred in informal discussions – both via the formation of new collaborations and via the continuation of ongoing collaborations.

3.1 Probability and Combinatorics

3.1.1 Omer Angel: The Long-Range Frog Model

Consider the frog model on either $\mathbb{Z}^d$ or the torus. At each site there are initially a Poisson number of sleeping frogs. When an active frog visits a site, all frogs there wake up, perform a random walk for some time, and die. In this work, our random walks are heavy tailed. We connect this model with long range percolation, and derive estimates for the hitting time of a vertex and the cover time of the torus.

Based on joint work with Jonathan Hermon and Yuliang Shi.

3.1.2 Louigi Addario-Berry: Root reconstruction for SI trees

Fix a graph G and a starting vertex r (the root), and let T be a spanning tree of G grown as follows. Initially, vertex r is infected and all other vertices are susceptible. At constant rate, each infected node infects each neighbouring susceptible node. Eventually, all vertices are infected; the edge set of T is the set of edges along which the infection spread. How easy is it to recover the identity of the initially infected individual r if one is given access to T?

I will discuss recent work on this problem when G is a random graph with a fixed degree sequence. Along the way, it is also useful to consider the case when G is an infinite tree.

Based on joint work with Sasha Bell, Tasmin Chu, Théodore Conrad-Frenkiel, Catherine Fontaine, Robin Khanfir, and Simone Têtu.

3.1.3 Jinyoung Park: Lipschitz functions on expanders

We will discuss the typical behavior of M-Lipschitz functions on d-regular expander graphs, where an M-Lipschitz function means any two adjacent vertices admit integer values differ by at most M. While it is easy to see that the maximum possible height of an M-Lipschitz function on an n-vertex expander graph is about $CM \log n$, where C depends (only) on d and the expansion of the given graph, it was shown by Peled, Samotij, and Yehudayoff (2012) that a uniformly chosen random M-Lipschitz function has height at most $C'M \log\log n$ with high probability, showing that the typical height of an M-Lipschitz function is much smaller than the extreme case. Peled-Samotij-Yehudayoff's result holds under the condition that, roughly, subsets of the expander graph expand by the rate of about $M \log(dM)$. We will show that the same result holds under a much weaker condition assuming that d is large enough.

Based on joint work with Robert Krueger and Lina Li [15].

3.1.4 Wojciech Samotij: The hypergraph container lemma revisited

The hypergraph container lemma is a powerful tool in probabilistic combinatorics that has found many applications since it was first proved a decade ago. Roughly speaking, it asserts that the family of independent sets of every uniform hypergraph can be covered by a small number of almost-independent sets, called containers.

In this talk, I will present two new versions of the hypergraph container lemma that utilise alternative notions of almost-independence, in place of the usual 'balanced supersaturation' property. The two lemmas display a number of other attractive features and have surprising connections to several other well-studied topics in probabilistic combinatorics. Based on joint work with Marcelo Campos (Cambridge) [3].

3.2 Statistical Mechanics

3.2.1 Roman Kotecký: Emergence of a giant component for random cluster model on the n-cube

Sudden emergence of a giant component discovered by Erdős and Rényi concerns random subgraphs defined in terms of Bernoulli percolation on the complete graph K_n. Similar results were obtained in other instances of percolation on various finite graphs. In particular, the case of percolative random subgraphs of n-cube Q_n was treated by Ajtai, Komls, and Szemerdi.

Surprisingly, the only result concerning emergence of a giant component when percolation is replaced by a random cluster measure has been obtained in the case of the complete graph by Bollobás, Grimmett, and Janson.

While formulating our results and some hints of proofs, I will touch several related topics: Potts model and its mean field limit, Edwards-Sokal coupling, sprinkling for random cluster measures, and motivations from statistical physics via weighted random interchange on Q_n.

Based on joint work with Darion May.

3.2.2 Daniel Hadas: Random high-density packing of disks on a lattice

Consider random configurations of disjoint disks of euclidean diameter D centered on integer coordinates in a bounded domain, where the probability for the appearance of a configuration is proportional to λ^n, with n being the number of tiles in the configuration. This is a special case of the Hard-core model in statistical physics. We are interested in characterizing the periodic Gibbs measures for sufficiently large values of λ. For all but a finitely many cases for D, this was solved recently by Mazel, Stuhl and Suhov. The remaining cases seem to display a variety of interesting behaviours. We solved the case $D = 2$, using Reflection Positivity. I will present ideas from our proof, and some initial thoughts (and MCMC simulations) for the remaining open cases. Based on joint work with Ron Peled [8].

3.3 Perceptrons and Spin Glasses

3.3.1 Reza Gheissari: Metastability in heavy-tailed spin glass dynamics

Many low-temperature dynamics in complex landscapes are expected to exhibit a sharp form of metastability, where the state space can be partitioned into wells, such that the equilibration time within each well is much faster than the transit time between wells, and the process tracking which well the Markov chain belongs to, itself is asymptotically Markovian. We overview this predicted picture for mean-field spin glass dynamics, its relation to aging predictions, and then describe recent results with Curtis Grant proving this for Glauber dynamics for mean-field heavy-tailed spin glasses. Joint work with Curtis Grant [6].

3.3.2 Brice Huang: Algorithmic threshold for the random perceptron

We consider the problem of efficiently optimizing the random (spherical or Ising) perceptron model with general bounded Lipschitz activation. We focus on a class of algorithms with Lipschitz dependence on the disorder, which includes constant-order methods such as gradient descent, Langevin dynamics, and AMP on dimension-free time-scales. Our main result exactly characterizes the optimal value ALG such algorithms can attain in terms of a 1-dimensional stochastic control problem. Qualitatively, ALG is the largest value whose level set contains a certain "dense solution cluster." Quantitatively, this characterization yields both improved algorithms and hardness results for a variety of asymptotic regimes, which are sharp up to absolute constant factors.

Based on joint work with Mark Sellke and Nike Sun.

3.3.3 Shuangping Li: Discrepancy algorithms for the binary perceptron

The binary perceptron problem asks us to find a sign vector in the intersection of independently chosen random halfspaces with a fixed intercept. The computational landscape of the binary perceptron is not yet well-understood. In some regimes there may be an information-computation gap, but there is much room for improvement on both the algorithms and lower-bounds side. In this talk I will discuss a forthcoming work in which we analyze the performance of canonical discrepancy minimization algorithms for the binary perceptron problem, obtaining new algorithmic results with simple off-the-shelf algorithms and relatively simple analysis. In some settings, we complement these algorithmic results with (close, but non-matching) overlap-gap lower bounds.

Based on joint work with Tselil Schramm and Kangjie Zhou [17].

3.4 Theoretical Computer Science

3.4.1 Nima Anari: Parallel Sampling via Counting

I will discuss how to speed up sampling from an arbitrary distribution on a product space $[q]^n$, given oracle access to conditional marginals, as in any-order autoregressive models. The algorithm takes $n^{2/3}\, polylog(n, q)$ parallel time, the first sublinear-in-n bound for arbitrary distributions. We also show a lower bound of $n^{1/3}$ on the parallel runtime of any algorithm, putting the complexity firmly in the sublinear but polynomially large territory.

Based on joint work with Aviad Rubinstein and Ruiquan Gao [1].

3.4.2 Dana Randall: Emergent phenomena in equilibrium and nonequilibrium collectives

Programmable matter explores how collections of computationally limited agents acting locally and asynchronously can achieve useful coordinated behaviors. We take a stochastic approach using techniques from randomized algorithms and equilibrium statistical physics to develop distributed algorithms for emergent collective behaviors that give guarantees and are robust to failures. We will also introduce some new tools that may prove fruitful in nonequilibrium settings as well.

3.4.3 Alex Wein: Optimality of AMP Among Low-Degree Polynomials

The approximate message passing (AMP) framework has been widely successful at producing algorithms with provable guarantees for a variety of high-dimensional statistical inference tasks. In some settings, AMP is conjectured to be optimal in the sense that no computationally efficient estimator can achieve a better mean squared error (MSE). In a simple "signal plus noise" model (spiked Wigner), we prove a variant of this conjecture by showing that AMP has the best possible MSE within a larger class of algorithms, namely those that can be described as constant-degree polynomials.

Based on joint work with Andrea Montanari [19].

4 Outcome of the Meeting

The meeting was a success, exemplifying the positive effects of having researchers spend an extended period of time together. While many participants arrived knowing only a few (or no) others, everyone was quickly introduced via the lightning talks, which were a big success that contributed to lively discussions in breaks, over meals, and in the moments of unstructured time (hikes; evenings). The connectedness of the subjects of the meeting became apparent to all participants as the days progressed. While initially some speakers mentioned that their talk may not fit the theme of the conference, as the week progressed we stopped hearing this. Many threads connecting the subjects of the talks became apparent. The speakers at the workshop have to be thanked for their efforts in preparing accessible and inspiring talks, which enables these connections to be drawn without undue difficulty.

The broad scope of the workshop was gratifying for both senior and junior researchers. For more senior participants, the workshop gave them an opportunity to meet new people and hear about new topics, in areas where their experience already gives them a foothold. Moreover, it provided an opportunity for them to be brought to the forefront of knowledge in specific topics (e.g., spin glass algorithms) or to be caught up on developments in areas where they have an interest, but are not currently active participants. For junior participants the workshop was an opportunity for them to situate their work into a broader picture and to see how it can impact other research communities. For both junior and senior participants a recurrent theme was that the broad set of participants shed light on definitions, conjectures, and research directions from adjacent fields – i.e., from the ability to discuss what is typically not explicated in research papers.

The reasonable amount of unstructured time during the workshop (coffee breaks; mealtimes; evenings) lead to many discussions. Participants have reported progress in both ongoing and new collaborations. We are excited to see the progress that this leads to, hopefully at a future iteration of this event.

5 Demographics

In total there were thirty-nine in-person participants, of which thirteen were women. A mix of career stages were represented, with five PhD students, six postdocs, and twenty-eighth tenure-track or tenured faculty present. Two participants hold positions in which teaching is the primary job role, as opposed to the more traditional R1 mix of teaching and research. Participants were primarily from institutions in the United States (23) and Canada (6), but several other countries were represented (Chile (1), Czechia (1), Israel (2), Netherlands (1), Mexico (2), United Kingdom (3)). Total international representation was somewhat diminished due to substantial delays in the Canadian visa system.

References

[1] Nima Anari, Ruiquan Gao, and Aviad Rubinstein. Parallel sampling via counting. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 537–548, 2024.

[2] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, (0):FOCS20–1, 2021.

[3] Marcelo Campos and Wojciech Samotij. Towards an optimal hypergraph container lemma. *arXiv preprint arXiv:2408.06617*, 2024.

[4] Ewan Davies and Will Perkins. Approximately counting independent sets of a given size in bounded-degree graphs. *SIAM Journal on Computing*, 52(2):618–640, 2023.

[5] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic stochastic localization. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 323–334. IEEE, 2022.

[6] Reza Gheissari and Curtis Grant. Metastability in Glauber dynamics for heavy-tailed spin glasses. *arXiv preprint arXiv:2407.11874*, 2024.

[7] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in mathematical physics*, 233:1–12, 2003.

[8] Daniel Hadas and Ron Peled. Columnar order in random packings of 2×2 squares on the square lattice. *arXiv preprint arXiv:2206.01276*, 2022.

[9] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov–Sinai theory. *Probability Theory and Related Fields*, 176(3):851–895, 2020.

[10] Brice Huang and Mark Sellke. Tight lipschitz hardness for optimizing mean field spin glasses. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 312–322. IEEE, 2022.

[11] Vishesh Jain and Clayton Mizgerd. Rapid mixing of the down-up walk on matchings of a fixed size. *arXiv preprint arXiv:2408.03466*, 2024.

[12] Matthew Jenssen and Peter Keevash. Homomorphisms from the torus. *Advances in Mathematics*, 430:109212, 2023.

[13] Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM Journal on Computing*, 49(4):681–710, 2020.

[14] Matthew Jenssen and Will Perkins. Independent sets in the hypercube revisited. *Journal of the London Mathematical Society*, 102(2):645–669, 2020.

[15] Robert A. Krueger, Lina Li, and Jinyoung Park. Lipschitz functions on weak expanders, 2024. *arXiv preprint arXiv:2408.14702*, 2024.

[16] Aiya Kuchukova, Marcus Pappik, Will Perkins, and Corrine Yap. Fast and slow mixing of the Kawasaki dynamics on bounded-degree graphs. *arXiv preprint arXiv:2405.06209*, 2024.

[17] Shuangping Li, Tselil Schramm, and Kangjie Zhou. Discrepancy algorithms for the binary perceptron. *arXiv preprint arXiv:2408.00796*, 2024.

[18] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[19] Andrea Montanari and Alexander S Wein. Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation. *arXiv preprint arXiv:2212.06996*, 2022.

[20] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Physical Review Letters*, 43(23):1754, 1979.

[21] Alexander D Scott and Alan D Sokal. The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *Journal of Statistical Physics*, 118:1151–1261, 2005.

[22] James B. Shearer. On a problem of Spencer. *Combinatorica*, 5:241–245, 1985.

[23] Allan Sly. Computational transition at the uniqueness threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 287–296. IEEE, 2010.

[24] Eliran Subag. Following the ground states of full-RSB spherical spin glasses. *Communications on Pure and Applied Mathematics*, 74(5):1021–1044, 2021.

[25] Michel Talagrand. The Parisi formula. *Annals of mathematics*, pages 221–263, 2006.

[26] Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 140–149, 2006.