Counting and sampling algorithms at low temperature

New frontiers in approximate counting
STOC 2020

Will Perkins
(UIC)

Algorithms and phase transitions

- When are **phase transitions** barriers to efficient algorithms?
- What algorithmic techniques can work in the **low-temperature** regime (strong interactions)?
- Based on joint work with many coauthors: **Christian Borgs, Sarah Cannon, Jennifer Chayes, Zongchen Chen, Andreas Galanis, Leslie Goldberg, Tyler Helmuth, Matthew Jenssen, Peter Keevash, Guus Regts, James Stewart, Prasad Tetali, Eric Vigoda**

Outline

- **High** and **low temperature** regimes in the Potts and hard-core models
- What is a **phase transition**? How are algorithms and phase transitions connected?
- Some **low temperature algorithms**
- Many **open problems**!

Potts model

Probability distribution on q-colorings $\sigma\colon V(G) \to [q]$ of the vertices of G:

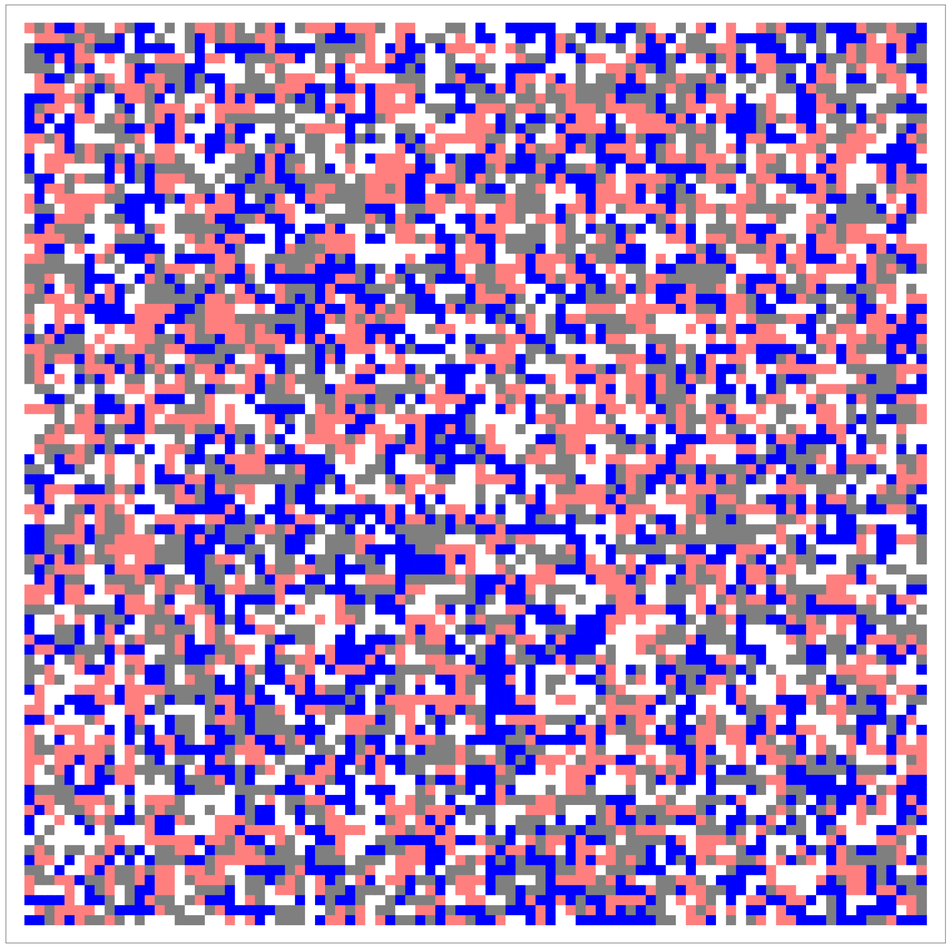
$$\mu(\sigma) = \frac{e^{\beta m(G,\sigma)}}{Z_G(\beta)}$$

$m(G, \sigma)$ is the number of monochromatic edges of G under σ

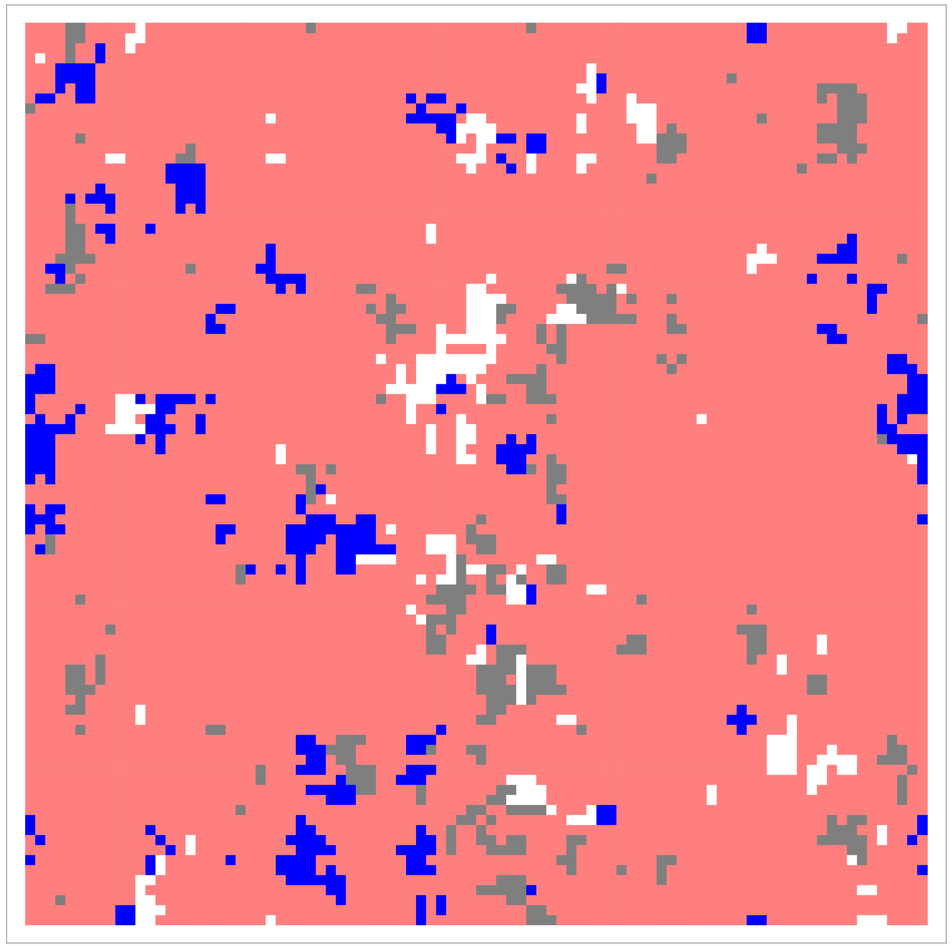
$Z_G(\beta) = \sum_{\sigma \in [q]^V} e^{\beta m(G,\sigma)}$ is the **partition function**.

β is the inverse temperature. $\beta \geq 0$ is the **ferromagnetic** case: same color preferred

Potts model



High temperature (β small)



Low temperature (β large)

Phase transitions

- On $\mathbb{Z}^d$ the Potts model undergoes a **phase transition** as β increases
- For small β **influence of boundary conditions** diminishes as volume grows; for large β influence of boundary conditions persists in infinite volume
- For small β, **correlations decay** exponentially fast, configurations are disordered (on, say, the discrete torus)
- For large β, we have **long range order** (and a dominant color in a typical configuration)
- For small β, Glauber dynamics **mix rapidly**; for large β **mix slowly**

Hard-core model

The **hard-core model** is a simple model of a gas.

Probability distribution on **independent sets** of G:

$$\mu_\lambda(I) = \lambda^{|I|}/Z_G(\lambda)$$

where $Z_G(\lambda) = \sum_I \lambda^{|I|}$ is the partition function (independence polynomial)

$\lambda \geq 0$ is the **fugacity**. Larger λ means stronger interaction

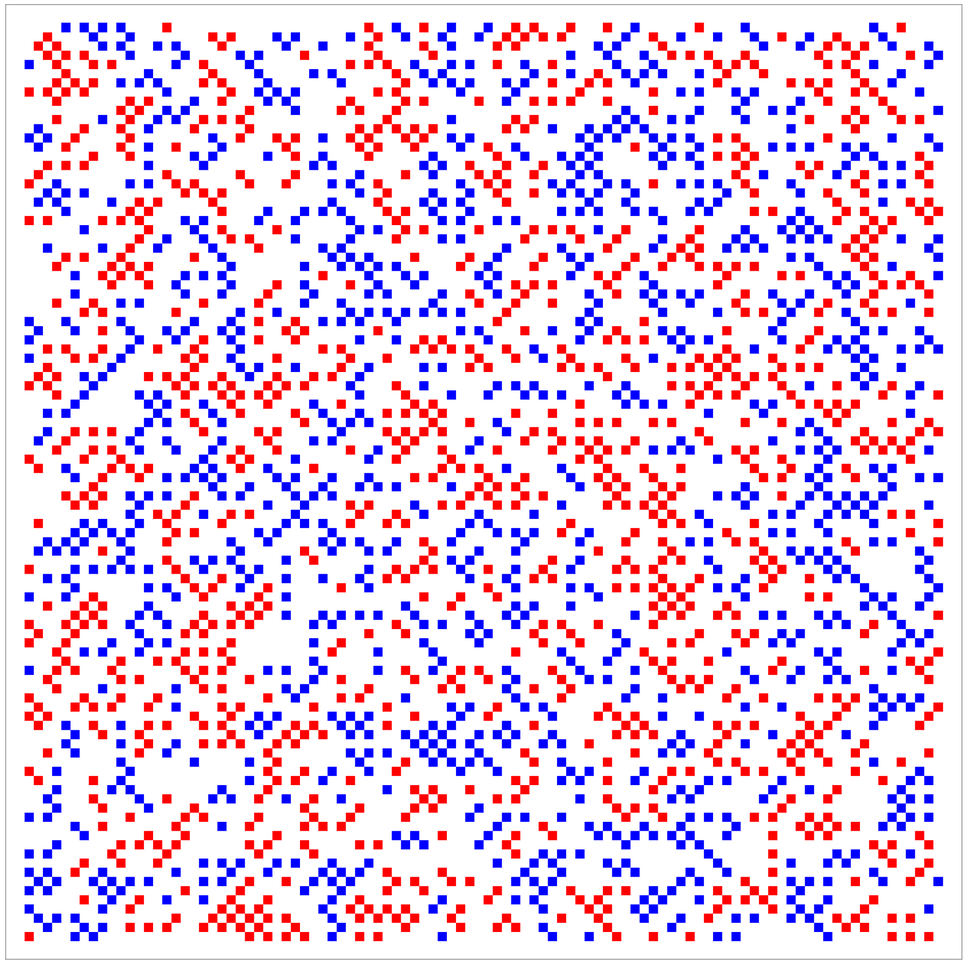
Hard-core model

On $\mathbb{Z}^d$ the hard-core model exhibits a **phase transition** as λ changes

Unoccupied

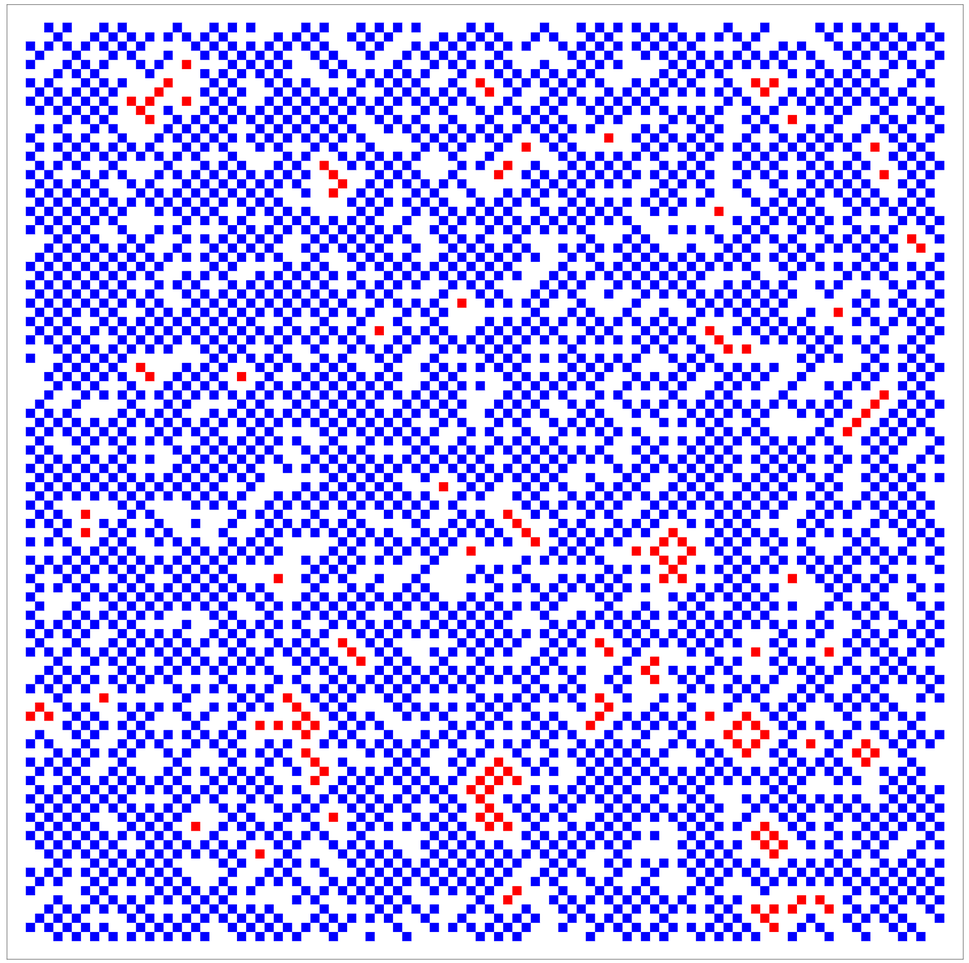
Even occupied

Odd occupied



Low fugacity

High temperature



High fugacity

Low temperature

Ground states

The **ground states** (maximum weight configurations) of the ferromagnetic Potts model are simple: they are the q **monochromatic configurations**.

The ground states of the hard-core model on $\mathbb{Z}^d$ are also simple: the **all even** and **all odd** occupied configurations.

Algorithms

- Two main **computational problems** associated to a statistical physics model: approximate the partition function (**counting**) and output an approximate sample from the model (**sampling**)

- Many different approaches including **Markov chains**, **correlation decay** method, **polynomial interpolation**.

Algorithms

- These algorithmic approaches work in great generality at high temperatures (weak interactions) but are **limited by phase transitions**
- Local Markov chains **mix slowly** at low temperatures
- **Long-range correlations** emerge on trees and graphs at low temperatures
- **Complex zeroes** of partition functions accumulate at a phase transition point

Algorithms

- How to **circumvent** these barriers?
- Design Markov chains on a different state space or with different transitions to avoid bottlenecks: **Jerrum-Sinclair** algorithm for the Ising model; **Swendsen-Wang** dynamics for the Potts model
- Today's talk: use **structure of the phase transition** to design efficient algorithms

Algorithms

- Phase transitions come in **many different varieties**!
- Compare hard-core model on **random regular** graphs to the hard-core model on **random regular bipartite** graphs (replica symmetry breaking vs replica symmetric)
- Ferro Potts and hard-core on bipartite graphs: **easy to find a ground state**. Does this mean it is easy to count and sample?
- Models like these are distinctive for both **phase transitions** and **algorithms**

#BIS

- No known FPTAS/FPRAS or NP-hardness for counting the number of **independent sets in a bipartite graph** G.

- **Dyer, Goldberg, Greenhill, Jerrum**: defined a class of problems as hard to approximate as BIS.

- Many natural approximate counting problems are **#BIS-hard** (counting stable matchings, ferromagnetic Potts model, counting colorings in bipartite graphs, etc..)

- #BIS-hardness even on graphs of maximum degree $\Delta \geq 3$

#BIS

- #BIS plays a role in approximate counting similar to that of **Unique Games** in optimization - not known to be hard or easy and captures complexity of many interesting problems

- Caveat / open problem: many problems are known to be **#BIS-hard** (like ferro Potts) but not known to be **#BIS equivalent**

Algorithms for #BIS-hard problems

- We can exploit the structure of instances to design **efficient algorithms** for models like Potts and hard-core at **low temperatures**
- Results for subgraphs of $\mathbb{Z}^d$, random regular graphs, expander graphs
- Uses techniques from statistical physics and computer science used to understand **phase transitions** and prove **slow mixing** results for Markov chains

Algorithms for #BIS-hard problems

- First step is to **separate the state space** into pieces dominated by a single ground state (e.g. mostly red, mostly green, mostly blue configurations for Potts; mostly even and mostly odd occupied for hard-core)

- Prove that contributions from intermediate configurations is **exponentially small** (a bottleneck!)

- Control each piece by showing that **deviations from the ground state** behave like a new **high-temperature spin model**

Unbalanced bipartite graphs

- Example: hard-core model on **unbalanced bipartite graphs** (different degrees or fugacities for left/right vertices (paper w/ S. Cannon)
- Setting: G is a biregular, bipartite graph with degrees Δ_L, Δ_R fugacity $\lambda = 1.$
- Condition: $\Delta_R \geq 10\Delta_L \log(\Delta_L)$
- This includes regimes with **non-uniqueness** on the infinite biregular tree and **slow mixing** in random graphs
- We obtain an **FPTAS** and point-to-point **correlation decay** on all graphs

Unbalanced bipartite graphs

- We expect to see **many left occupied** vertices and **few right occupied** vertices in a typical independent set

- We think of the `ground state' as the collection of **independent sets with no right occupied vertices**: these contribute $(1+\lambda)^{|L|}$ to the partition function

- **Deviations** from this ground state are occupied right vertices

Unbalanced bipartite graphs

- A **polymer** γ is a 2-linked set of vertices from R
- The **weight** of a polymer is $w_\gamma = \dfrac{\lambda^{|\gamma|}}{(1+\lambda)^{|N(\gamma)|}}$
- Two polymers are **compatible** if their union is not 2-linked

$$Z_G(\lambda) = (1+\lambda)^{|L|} \sum_{\Gamma} \prod_{\gamma \in \Gamma} w_\gamma$$

where the sum is over collections of compatible polymers

Unbalanced bipartite graphs

- How to **analyze** this new model?
- $|N(\gamma)| \geq \frac{\Delta_R}{\Delta_L}|\gamma|$ so $w_\gamma \leq 2^{-\frac{\Delta_R}{\Delta_L}|\gamma|}$ at $\lambda = 1$
- **Exponentially decaying weights** when $\Delta_R > \Delta_L$
- We have **switched** from strong interactions to weak interactions! **Low temperature** to **high temperature**

Cluster expansion

- The **cluster expansion** is a tool from mathematical physics for analyzing probability laws on ‘dilute’ collections of geometric objects.

- It applies to a very general **weighted independent set model** — on a graph with inhomogeneous weights and unbounded vertex degrees. Each vertex represents a geometric object, neighboring objects overlap.

$$Z = \sum_{\Gamma} \prod_{\gamma \in \Gamma} w_{\gamma}$$

Cluster expansion

- The **cluster expansion** says that, under some conditions,

$$\log Z = \sum_{\Gamma_c} \Phi(\Gamma_c) \prod_{\gamma \in \Gamma_c} w_\gamma$$

- The sum is over **connected** collections of polymers. Informally, the conditions say that the **weights are exponentially small** in the size of the contours.

- The algorithm is to **truncate the cluster expansion** (like **Barvinok's** algorithm of truncating the Taylor series)

Algorithms

Making the cluster expansion **algorithmic** requires:

Enumerating polymers of size $O(\log n)$: essentially enumerating connected subgraphs in a bounded degree graph

Computing polymer weights

Sampling is done via self-reducibility on the level of polymers

Markov chains

- The results and techniques suggest a simpler and **faster sampling algorithm**: start with the all left occupied independent set and run Glauber dynamics.

- This chain may mix slowly from a worst-case start but converge close to stationarity from a **good start**.

- More generally for models with multiple dominant ground states, start chains from each. **Fast mixing** within a state

- How to **prove** that this works?

Markov chains

- Some progress w/ Chen, Galanis, Goldberg, Stewart, Vigoda: define a Markov chain on polymer configurations, adding or removing a single polymer at a time

- Under **weaker conditions** than cluster expansion convergence, this chain **mixes rapidly**

- Need stronger than cluster expansion conditions to **implement a single step** efficiently

- **Comparison techniques** give polynomial-time mixing within a state (with rejection) but not O(n log n) as we'd expect

Perturbative Approach

- The cluster expansion is a **perturbative tool** in statistical physics: needs some parameter to get large to ensure sufficient exponential decay
- In general we can't expect the techniques to work in a **sharp** range of parameters
- Semi-exception is large q Potts and random cluster models: can get efficient algorithms at **all temperatures** (on $\mathbb{Z}^d$ w/ Borgs-Chayes-Helmuth-Tetali; on random graphs w/ Helmuth-Jenssen)
- Can we sample from the hard-core model on random bipartite graphs for **all fugacities** λ?

Summary

- The connection between **phase transitions** and **algorithms** is fascinating and complex
- **#BIS** captures a class of counting problems in which ground states are easy to find but complexity of approximate counting is unknown
- On structured instances probabilistic tools can be made **algorithmic at low temperatures**
- Two tools: **polymer models** and the **cluster expansion**

Open Questions

- **More algorithms for #BIS** - more classes of graphs, better running times (subexponential?) see Goldberg-Lapinskas-Richerby for exponential-time algorithms

- **Markov chains beyond mixing times** - using well chosen starting configurations to sample efficiently despite slow mixing

- Deeper understanding of the relationship between **phase transitions and algorithms**: explanation for ‘coincidence’ of Lee-Yang and Heilmann-Lieb theorems and efficient algorithms for ferro Ising and matchings

- Make **non-perturbative** tools algorithmic

Open Questions

- **More algorithms for #BIS** - more classes of graphs, better running times (subexponential?) see Goldberg-Lapinskas-Richerby for exponential-time algorithms

- **Markov chains beyond mixing times** - using well chosen starting configurations to sample efficiently despite slow mixing

- Deeper understanding of the relationship between **phase transitions and algorithms**: explanation for ‘coincidence’ of Lee-Yang and Heilmann-Lieb theorems and efficient algorithms for ferro Ising and matchings

- Make **non-perturbative** tools algorithmic

Thank you!