First-order phase transitions and efficient sampling algorithms

Will Perkins
(UIC)

Algorithms and phase transitions

- When are **phase transitions** barriers to efficient algorithms?
- Does the **type** of phase transition play a role?
- Can we use tools from statistical physics to design **new algorithms**?
- Can we use the **algorithmic perspective** to understand phase transitions?

Based on:

- “Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures” (joint w/ **Borgs**, **Chayes**, **Helmuth**, **Tetali**, STOC 2020)
- “Finite-size scaling and phase coexistence for the random cluster model on random graphs” (joint w/ **Helmuth**, **Jenssen** on arxiv soon!)

Outline

- Potts model and random cluster model
- What is a phase transition? What is a first-order phase transition?
- Contour representations, Pirogov-Sinai theory
- Algorithms on $\mathbb{Z}^d$ and random graphs

Potts model

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

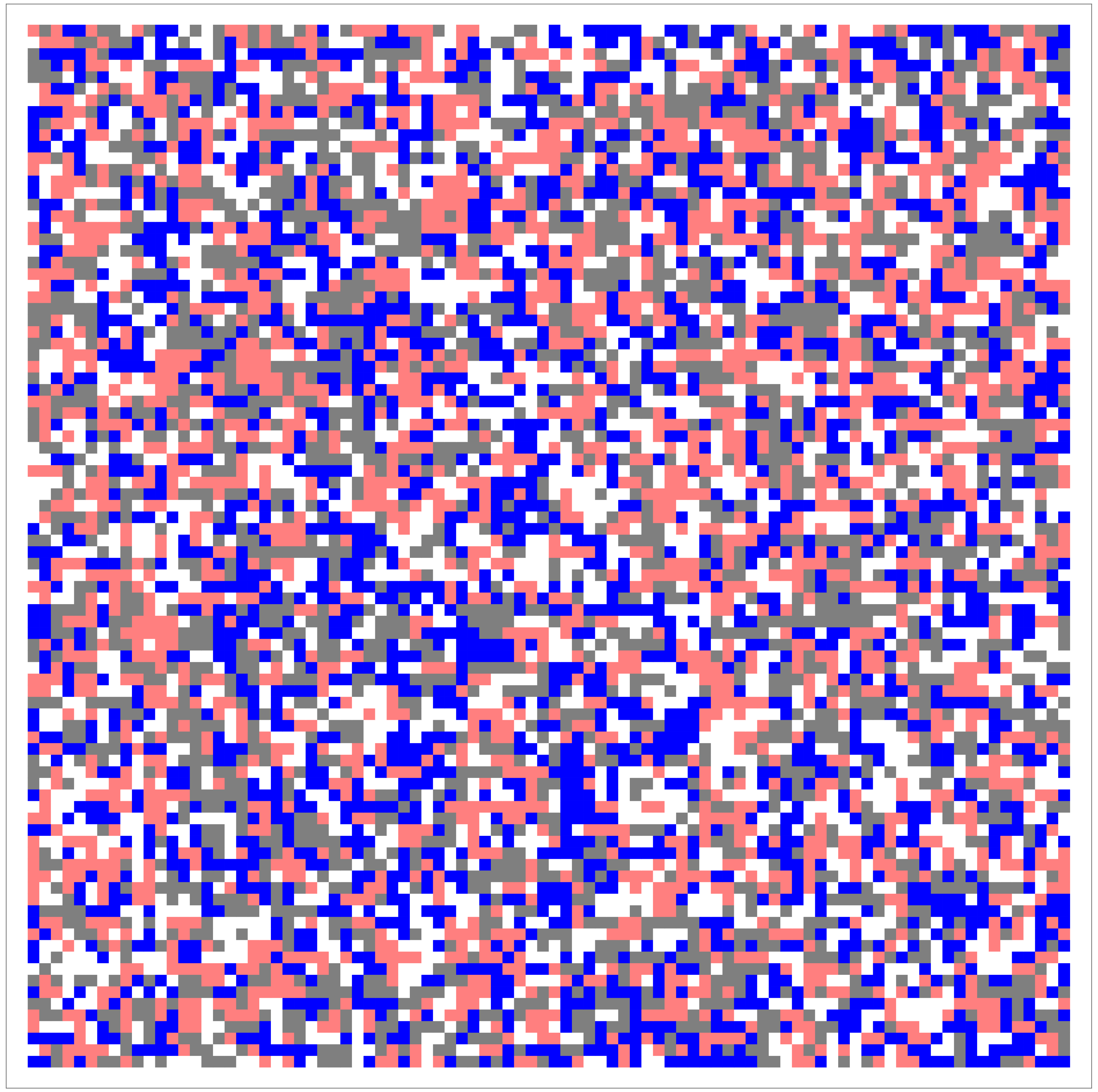
$$\mu(\sigma) = \frac{e^{\beta m(G,\sigma)}}{Z_G(\beta)}$$

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

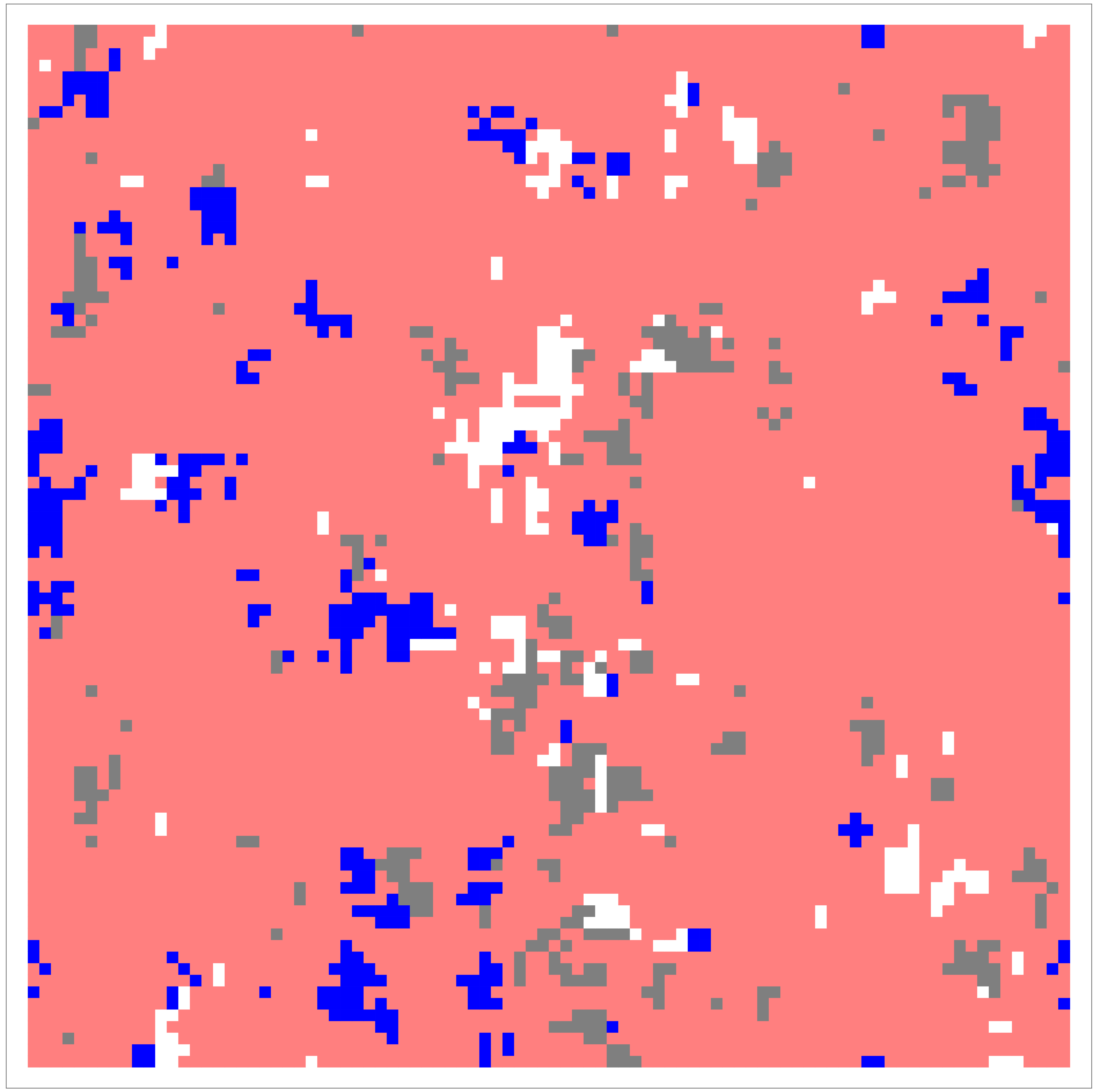
$Z_G(\beta) = \sum_{\sigma \in [q]^V} e^{\beta m(G,\sigma)}$ is the **partition function**.

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)

Random cluster model

The **random cluster model** is a generalization of the Potts model.

Probability distribution on **subsets of edges** of G:

$$\mu_{q,\beta}(A) = q^{c(A)}(e^{\beta} - 1)^{|A|}/Z_G(q,\beta)$$

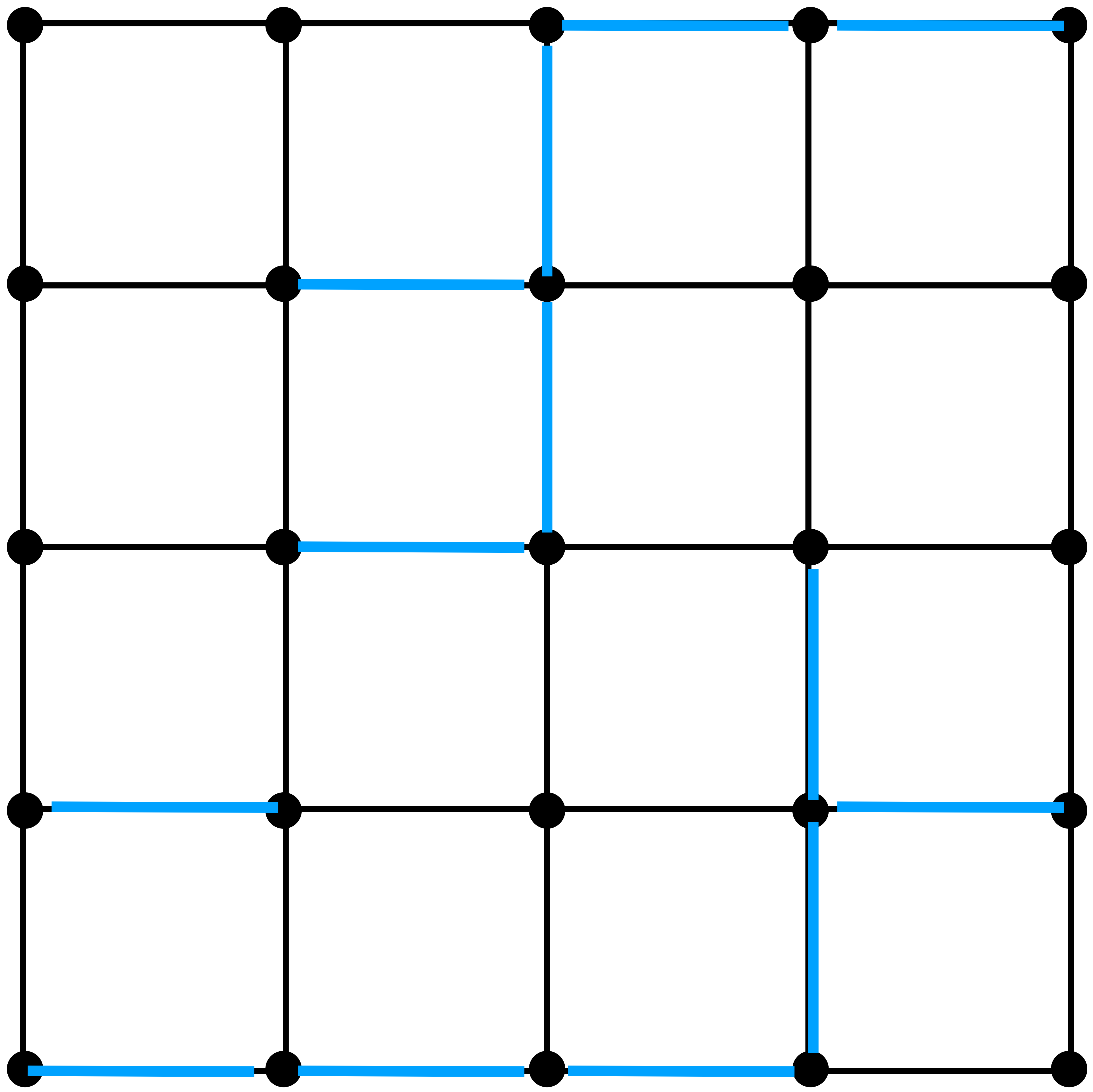
$c(A)$ is the number of connected components of (V, A).

$q > 0$ can be **non-integral**. $q = 1$ corresponds to independent edge percolation

Random cluster model

Edwards-Sokal coupling:

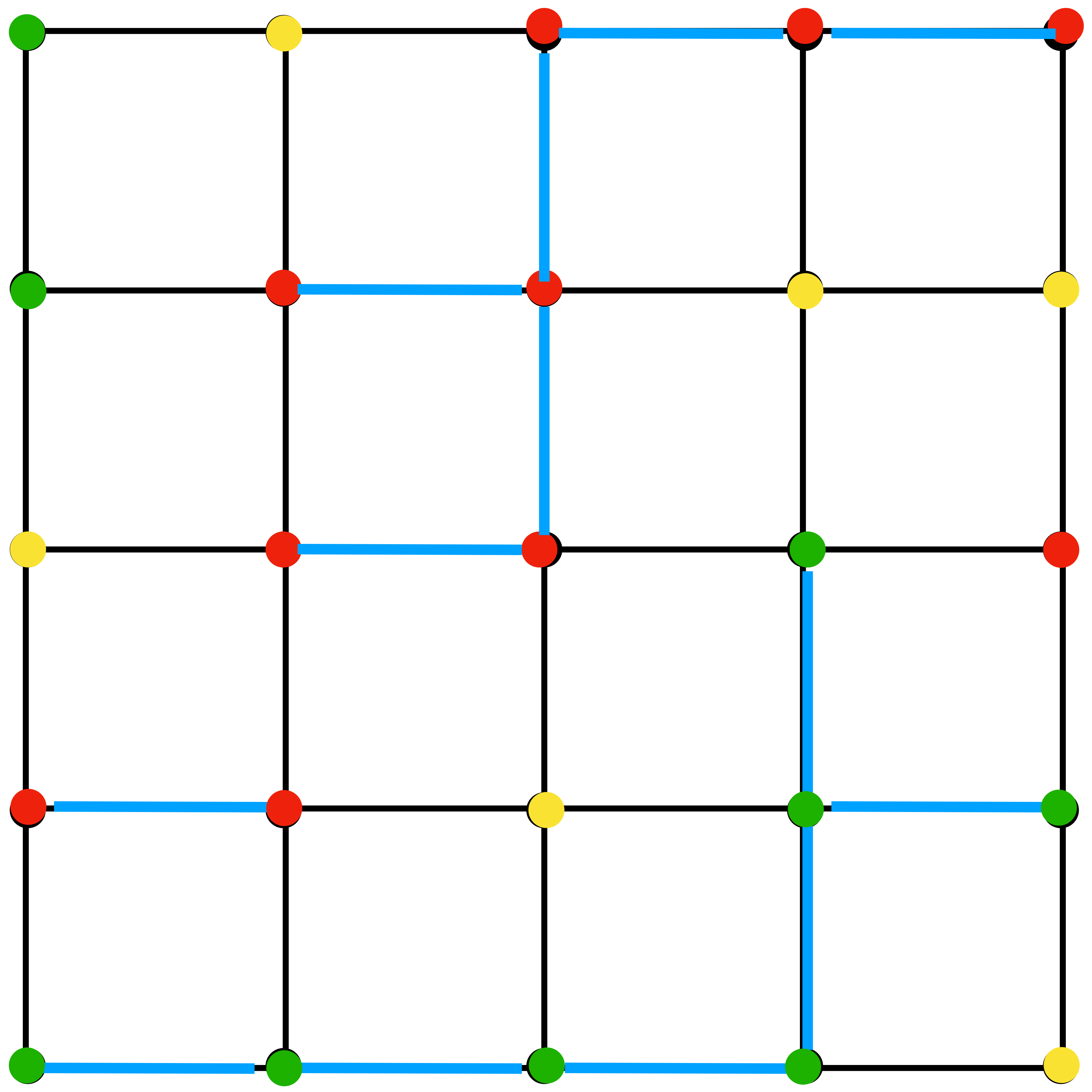
1. Pick a set of edges according to the random cluster measure
2. Determine the connected components



Random cluster model

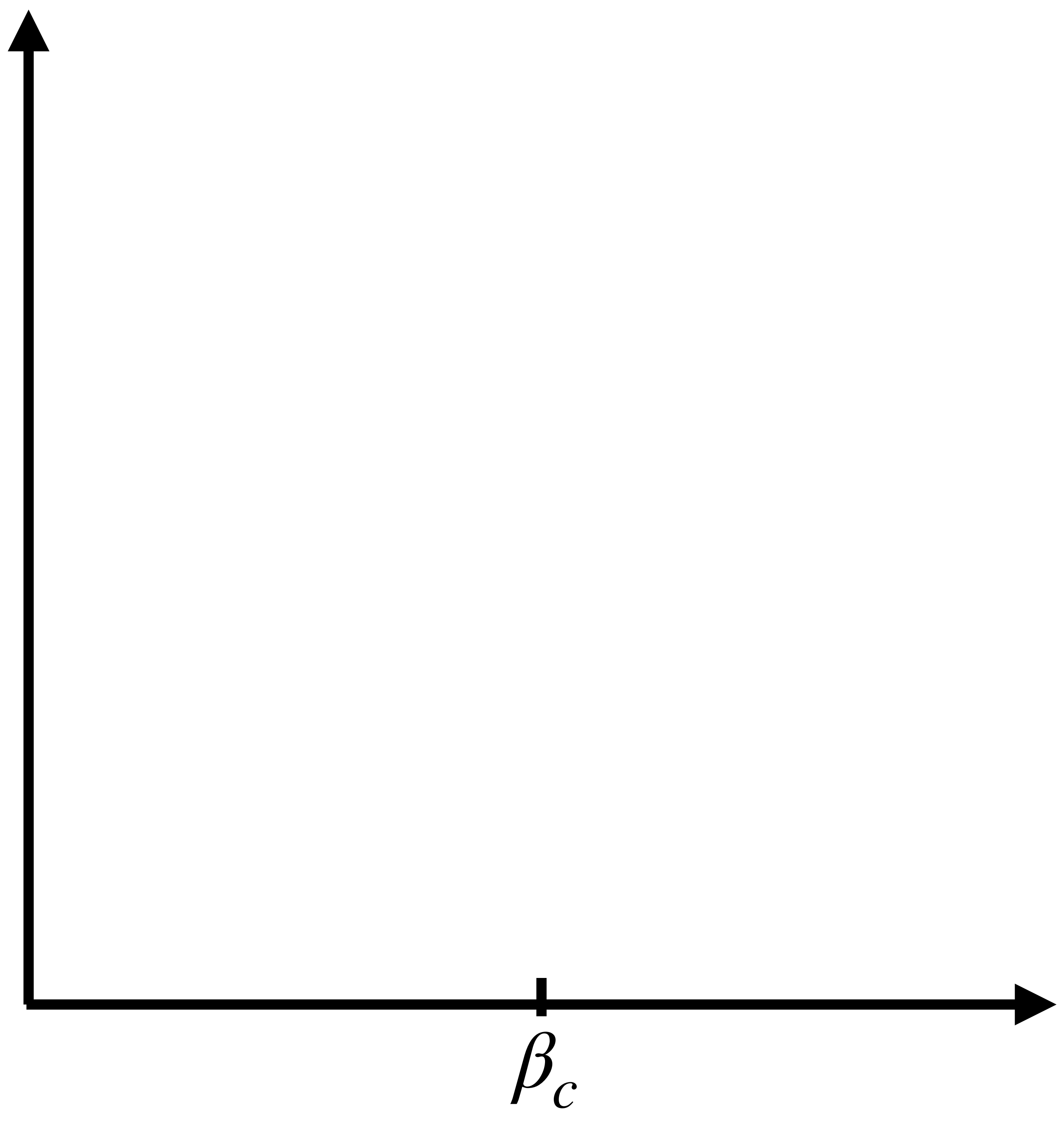
Edwards-Sokal coupling:

1. Pick a set of edges according to the random cluster measure
2. Determine the connected components
3. Assign one of the q colors uniformly and independently to each connected component

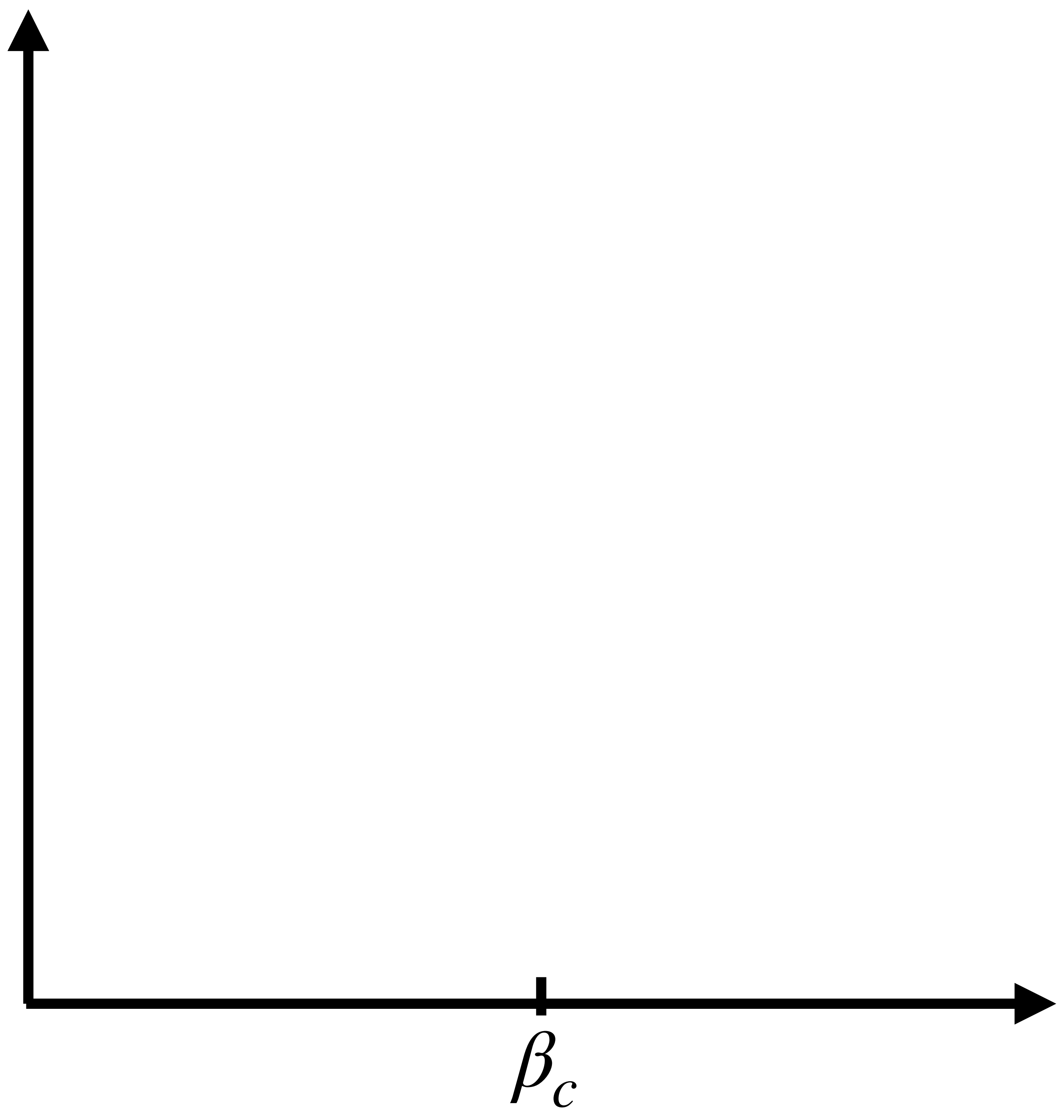


Phase transitions

- Another definition of a **phase transition**: plot the limiting value of an observable against β. A phase transition occurs at a **non-analytic** point



First-order



Second-order

Large q behavior

- For q large enough as a function of d, the random cluster model exhibits a **first-order phase transition**.
- **No middle ground**: for all β typical configurations consist of very few or very many edges (say, $\leq |E|/10$ or $\geq 9|E|/10$ for large enough q)
- Conditioned on number of edges, nice probabilistic properties, including **exponential decay of correlations**
- Proved by Laanait, Messager, Miracle-Sole, Ruiz, Shlosman, 1991 using **Pirogov-Sinai theory**

Phase transitions and 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**. All are limited by or must bypass **phase transitions** (slow mixing, long-range correlations, accumulation of zeroes on real axis)

Markov chains

- For the Potts model we have several useful Markov chains:
  - **Glauber dynamics** - pick a random vertex and update color
  - **Swendsen—Wang dynamics** - pick a cluster and update color
- **Glauber dynamics** mix slowly for $\beta \geq \beta_c$ - there is a bottleneck between mostly Red and mostly Green configurations
- **Swendsen-Wang** mixes slowly at $\beta = \beta_c$ - the middle ground is a bottleneck (Gore-Jerrum for K_n, Borgs-Chayes-Frieze-Kim-Tetali-Vigoda-Vu and Borgs-Chayes-Tetali for $\mathbb{Z}^d$, Galanis-Stefankovic-Vigoda-Yang for random graphs. Consequence of the **first-order phase transition**

Contour models

- Random cluster model has two **ground states**: the **disordered** state $A = \emptyset$ and the **ordered** state $A = E$.

- In the Potts model these correspond to choosing the color of each vertex independently and choosing a single color for all vertices.

- **Contours** provide a geometric way of separating a RC configuration into ordered and disordered regions

Contour models

- Idea of Pirogov-Sinai theory is to use **contour models** to separate configurations into mostly ordered and mostly disordered
- Leads to two **new partition functions** Z_{dis} and Z_{ord}
- These can be controlled via the **cluster expansion**

Free energies

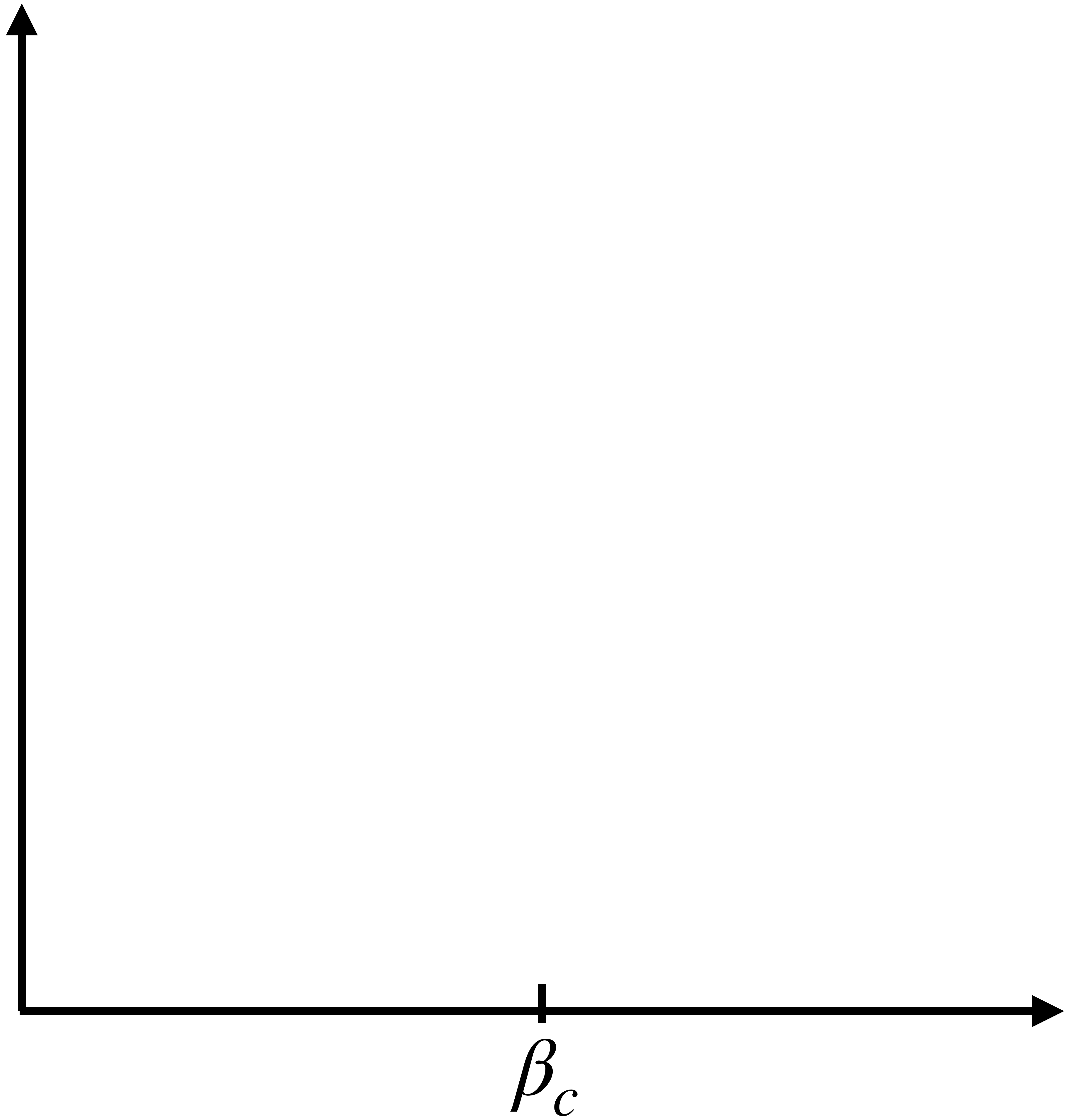
Let $f_{dis} = \lim_{n \to \infty} \frac{1}{n} \log \tilde{Z}_{dis}$ and $f_{ord} = \lim_{n \to \infty} \frac{1}{n} \log \tilde{Z}_{ord}$

By showing **no middle ground**, it follows that

$$f = \lim_{n \to \infty} \frac{1}{n} \log Z = \max\{f_{dis}, f_{ord}\}$$

β_c is the point at which $f_{dis} = f_{ord}$

Free energies



At criticality

- At $\beta = \beta_c$, Z_{dis} and Z_{ord} **match** on an exponential scale.
- On discrete torus, $Z_{ord}(\beta_c) \approx q \cdot Z_{dis}(\beta_c)$
- Smaller order corrections in n (volume of torus) are **finite-size effects** (Borgs—Kotecky—Miracle-Sole, 1991)
- $Z = Z_{ord} + Z_{dis} + \exp(-\Theta(n^{(d-1)/d}))$. Typical configurations look either **disordered** or **ordered** — no mixtures! unlike in **second-order** phase transition

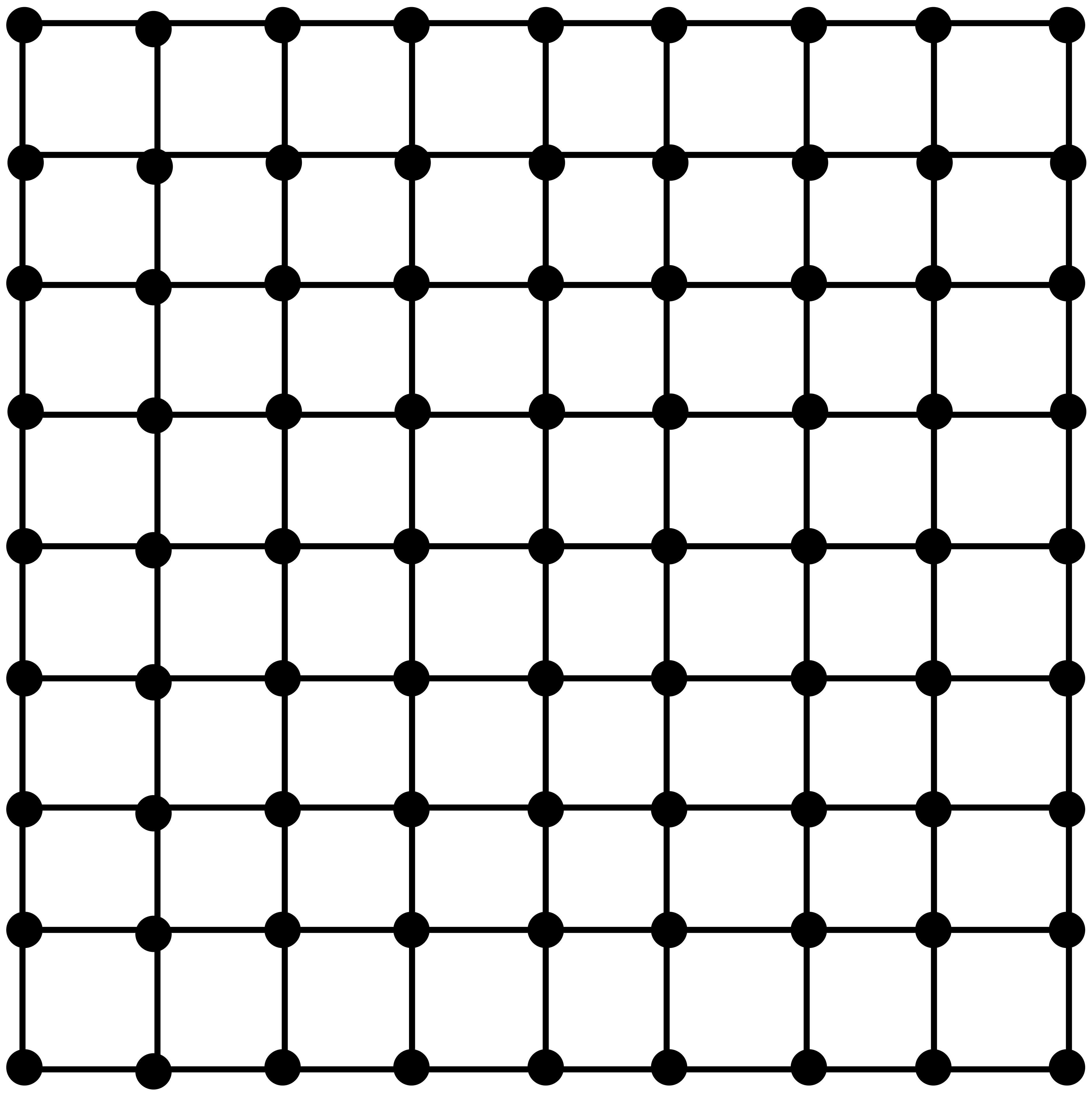
What's new

- The **cluster expansion** (and Pirogov-Sinai theory) can be made algorithmic: approximate Z_{dis} and Z_{ord} separately.
- **First-order phase transition** means that this suffices at criticality.
- Efficient sampling and counting for Potts and RC on $\mathbb{Z}^d$ at **all temperatures** (large q) [BCHTP]
- Apply the same approach to other graphs: efficient algorithms at all temperatures on **random Δ-regular graphs** (large q) [HJP].
- This gives us a detailed **phase diagram** of the random cluster model on random graphs: correlation decay, local weak convergence, precise phase coexistence etc. (see also Galanis-Stefankovic-Vigoda-Yang)

Rest of the talk

- High-level picture of **contour models** and **cluster expansions**
- How to makes these tools **algorithmic**
- What to do for **random graphs**? We lose some geometry but gain **expansion**.

Constructing Contours



Contour representation

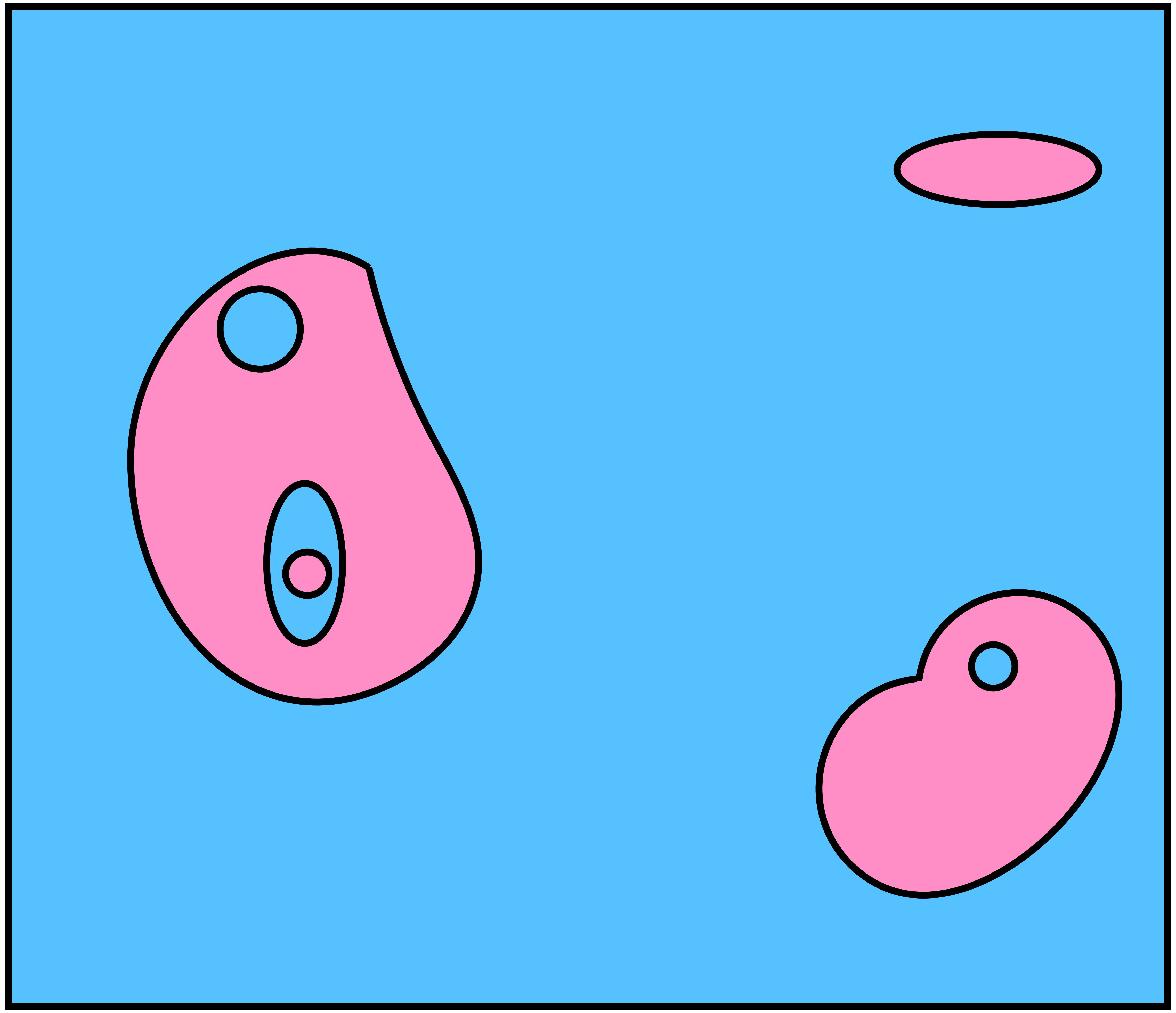
We can describe a configuration **geometrically**:

Ordered: occupied edges

Disordered: unoccupied edges

Boundaries between ordered/disordered regions are **contours**

Contours are non-intersecting, nested, with **labels** indicating the exterior and interior ground state



Contour representation

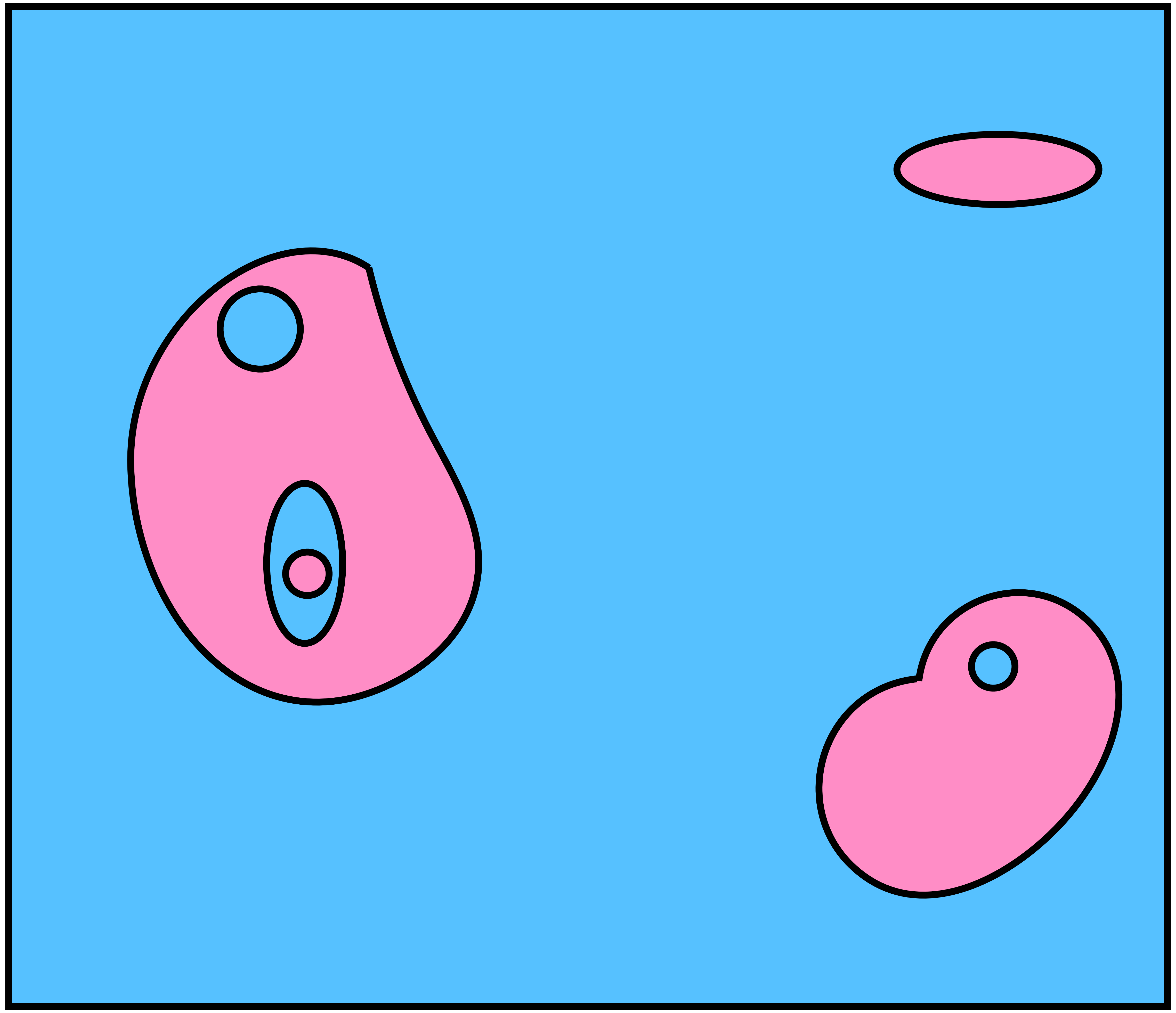
We can describe a configuration **geometrically**:

Ordered: occupied edges

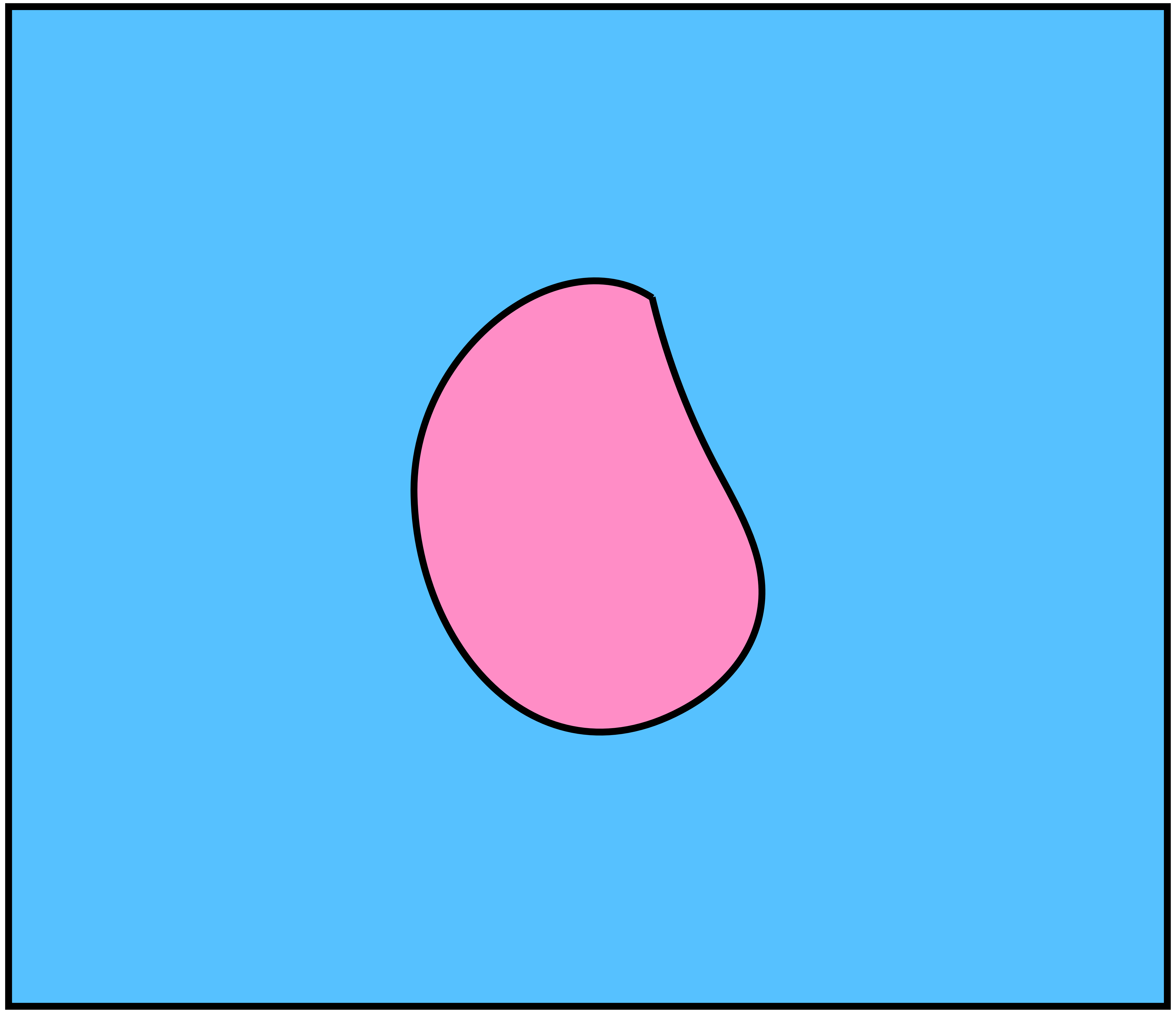
Disordered: unoccupied edges

A configuration is (**dis**)**ordered** if the only region that wraps around the torus (winding number >0) is (**dis**)**ordered**

If any contour wraps around, the configuration is part of the **middle ground**.



Contour representation



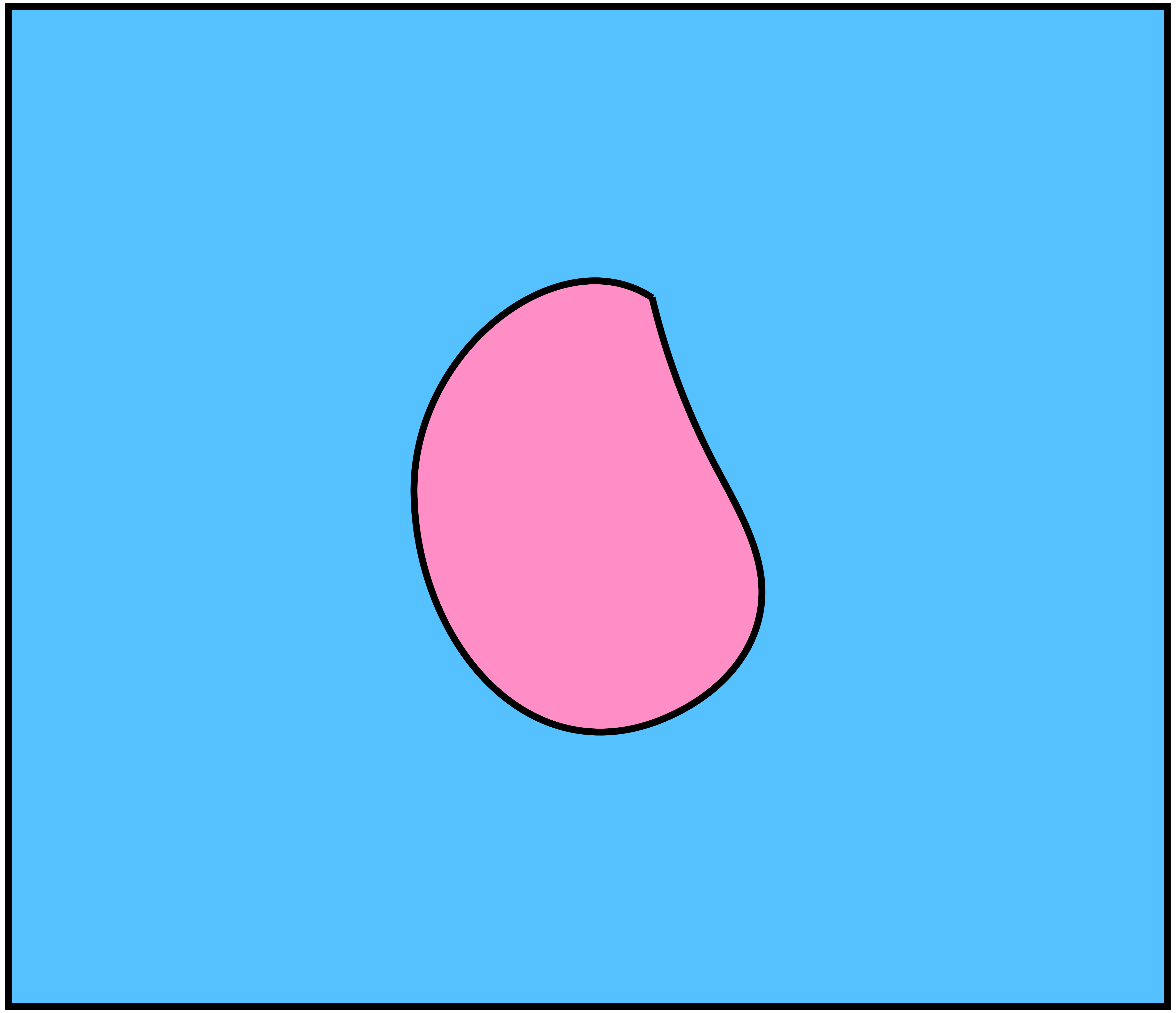
We can express the RC weight of a configuration in terms of its contours.

Ordered regions B have a volume factor $q(e^{\beta} - 1)^{d|B|}$

Disordered regions B have a volume factor $q^{|B|}$

Contours have a penalty factor exponentially small in their size (number of edges crossing the contour) $e^{-\kappa\|\gamma\|}$

Contour representation

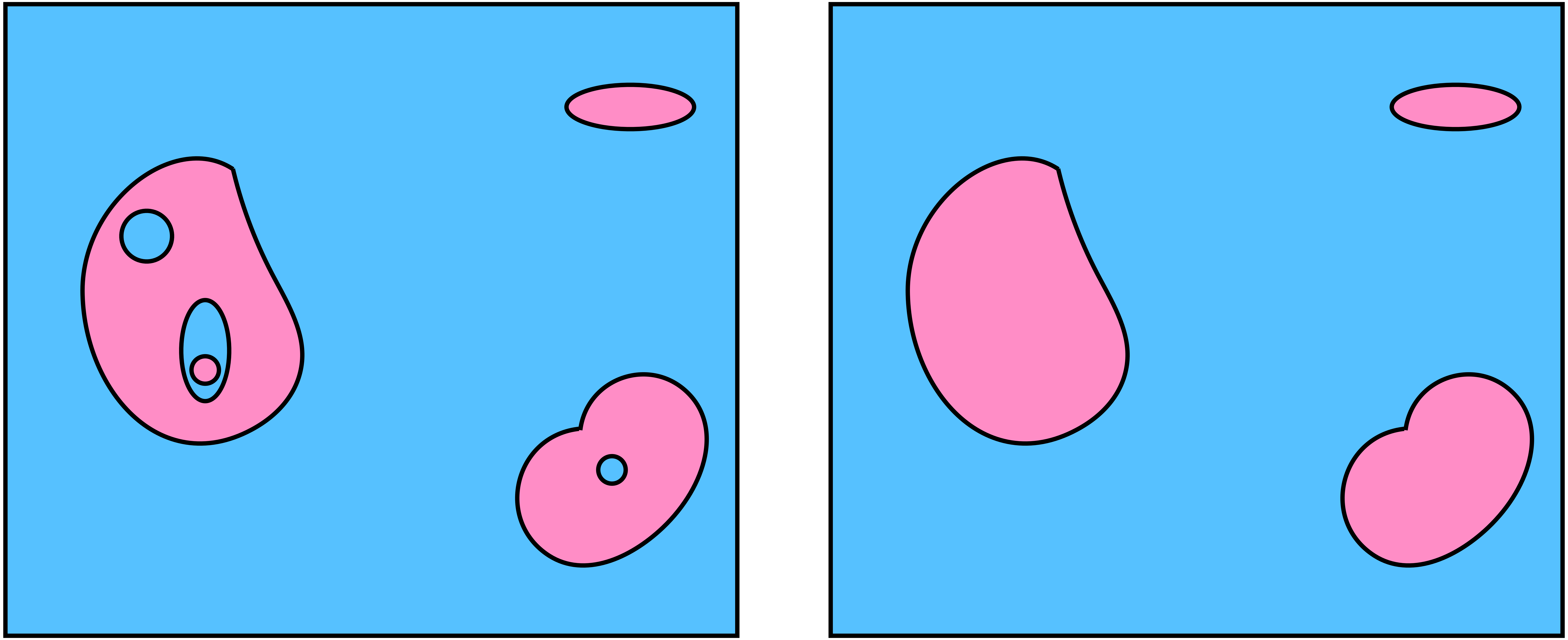


Weight of this configuration:

$$(e^{\beta} - 1)^{nd/2} \cdot e^{-\kappa \|\gamma\|} \cdot \left(\frac{q}{(e^{\beta} - 1)^{d}} \right)^{|Int(\gamma)|}$$

κ is increasing in q. $\beta_c \approx \log q/d$

Contour representation



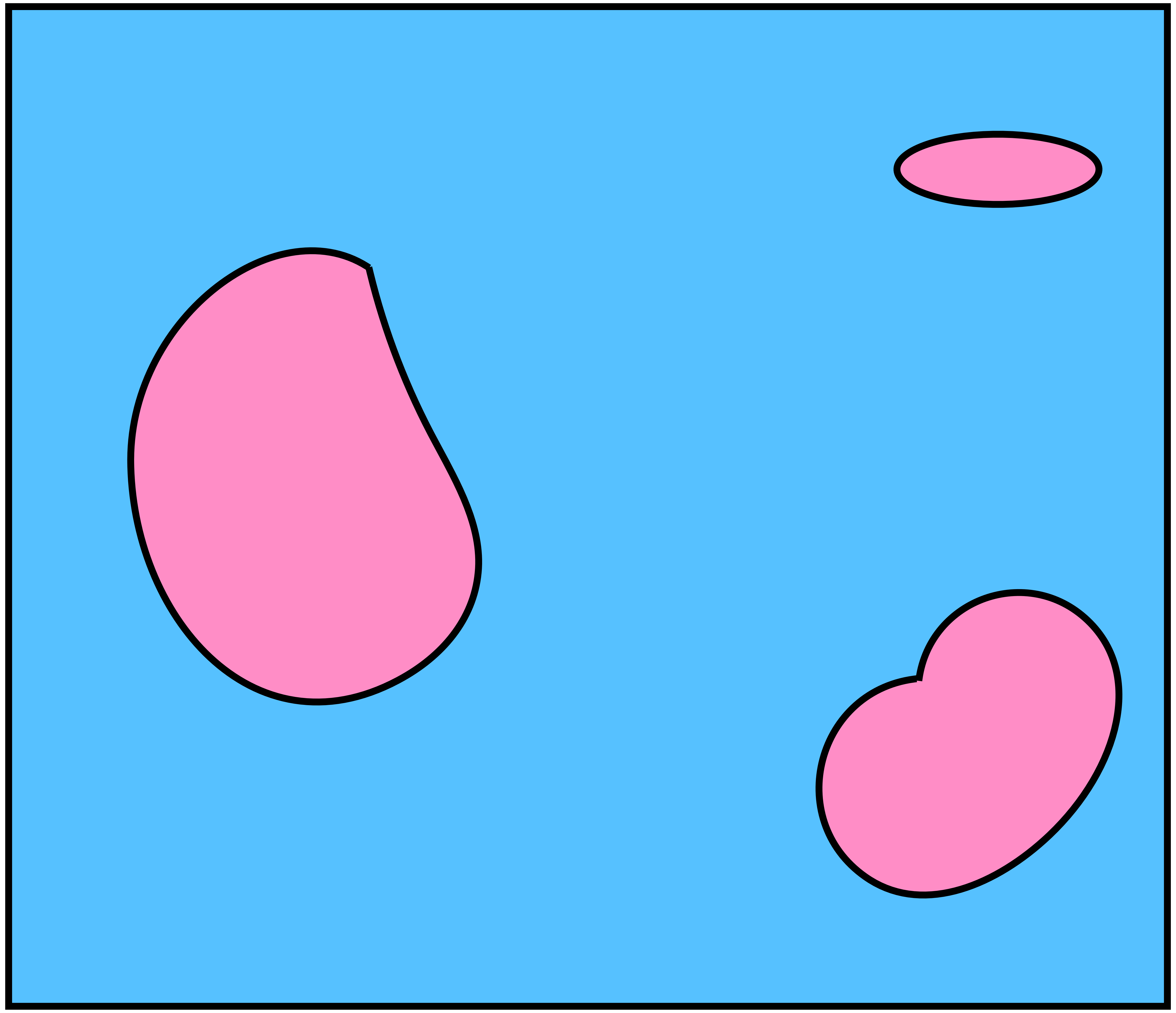
Outer contours

Contour representation

Weight of **all** configurations with these outer contours:

$$(e^{\beta}-1)^{nd}\prod_{\gamma\in\Gamma} e^{-\kappa\|\gamma\|}Z_{dis}(Int(\gamma))(e^{\beta}-1)^{-|Int(\gamma)|}$$

where $Z_{dis}(\Lambda)$ is the partition function with disordered boundary conditions



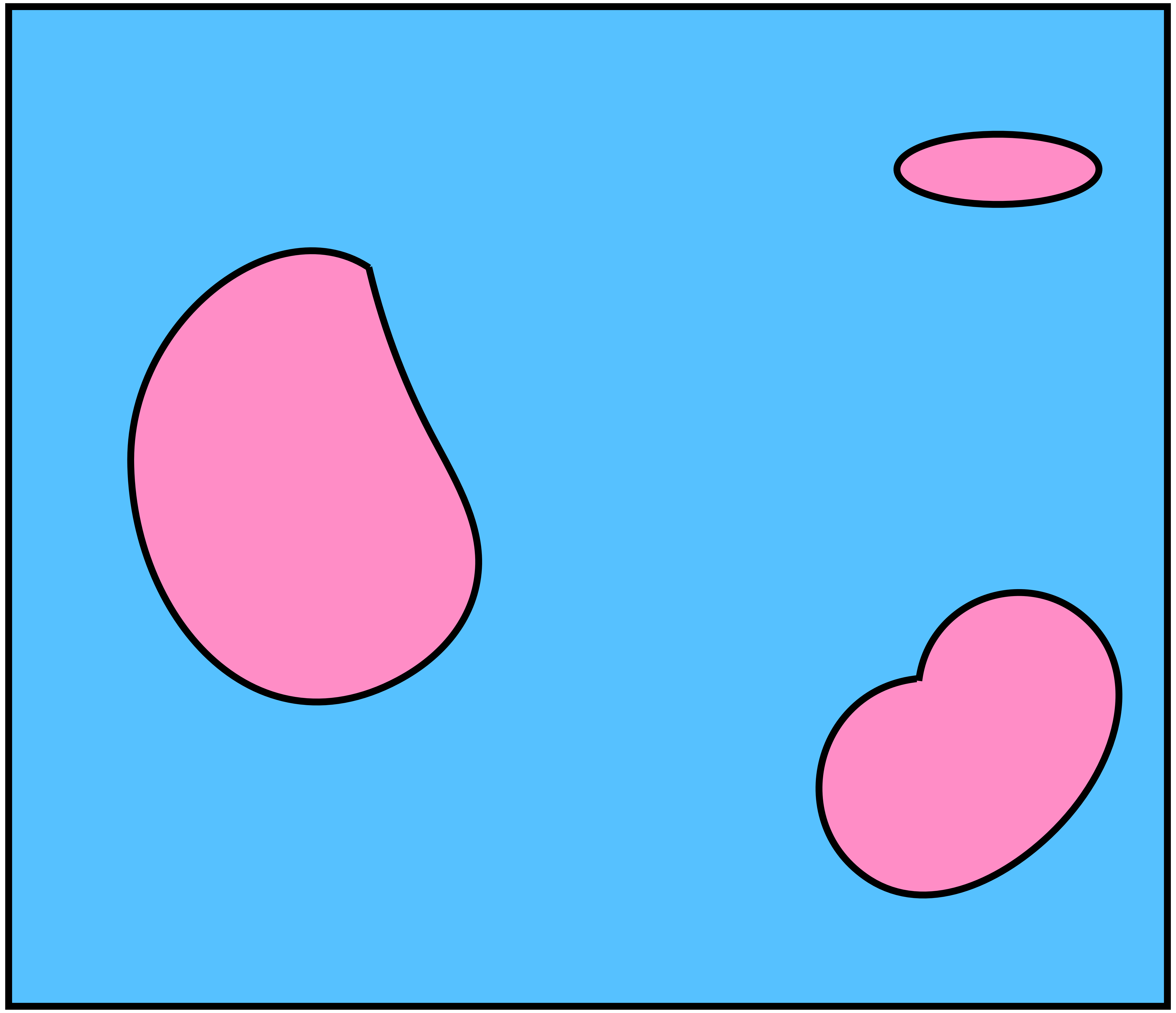
Outer contours

Contour representation

Express Z_{ord} as a sum over collections of compatible outer contours:

$$Z_{ord} = q(e^{\beta} - 1)^{nd} \sum_{\Gamma} \prod_{\gamma \in \Gamma} e^{-\kappa \|\gamma\|} Z_{dis}(Int(\gamma))(e^{\beta} - 1)^{-|Int(\gamma)|}$$

This looks like a generalized **hard-core model**: sum over `independent set', product of weights



Outer contours

Contour representation

With some additional manipulations (standard in Pirogov-Sinai theory) we have (sums are over collections of non-intersecting contours):

$$Z_{ord} = q(e^{\beta} - 1)^{nd} \sum_{\Gamma} \prod_{\gamma \in \Gamma} K_{ord}(\gamma)$$

$$Z_{dis} = q^{n} \sum_{\Gamma} \prod_{\gamma \in \Gamma} K_{dis}(\gamma)$$

For q large, either $K_{ord}(\gamma)$ or $K_{dis}(\gamma)$ is exponentially small in $\|\gamma\|$ ($\beta \geq \beta_c$ and $\beta \leq \beta_c$ respectively)

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 objects. Informally, the conditions say that the **weights are exponentially small** in the size of the objects.

Algorithms

Making the cluster expansion **algorithmic** requires:

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

Computing contour weights $K_{dis}(\gamma), K_{ord}(\gamma)$: tricky because weights involve ratios of partition functions, but this can be done inductively using the cluster expansion

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

Random graphs

We can apply a similar approach to the random cluster model on Δ-regular random graphs.

Thm (Helmuth, Jenssen, P.) For $\Delta \geq 5, q = q(\Delta)$ large enough:

There are efficient approximate counting and sampling algorithms for the Potts and random cluster models on random Δ-regular graphs at **all temperatures**.

Random graphs

Thm (Helmuth, Jenssen, P.) For $\Delta \geq 5$, $q = q(\Delta)$ large enough:

- Determine distribution of Z_{dis}/Z_{ord} at β_c
- Exponential decay of correlations for $\beta \neq \beta_c$
- Local convergence of RC measure to free or wired RC measure on infinite Δ-regular tree

Compare to Galanis-Stefankovic-Vigoda; Montanari-Mossel-Sly; Dembo-Montanari-Sly-Sun; Sly-Sun; very different techniques based on verifying the predictions of the **cavity method**; these results are generally on the level of the free energy

Random graphs

Random graphs are very good **expanders** and are locally **tree-like**.

We can define Z_{dis} and Z_{ord} via **polymer models** representing deviations from the disordered and ordered ground state.

We don't have to define weights inductively since the boundary of polymers is proportional to volume. But we lose the geometry of $\mathbb{R}^d$ and have to argue indirectly about the non-local RC interaction via expansion.

Summary

- The **first-order phase transition** in the Potts and Random Cluster model is a barrier to Markov chains like Swendsen-Wang

- But it facilitates a **different type of algorithm** based on approximating / sampling from ordered and disordered configurations separately

- This allows us to find efficient algorithms at **all temperatures**, including critical

- We can follow this framework for **random graphs** as well, obtaining new algorithms and new probabilistic results

Open questions

- Are there provably fast **Markov chain algorithms** to sample from these models at all temperatures?
- Can we deal with boundary conditions and **interfaces** algorithmically?
- How can we deal with **second-order** phase transitions algorithmically?

Thank you!