Counting independent sets in middle two layers of Boolean lattice

Lina Li

Joint work with József Balogh and Ramon I. Garcia

Univerisity of Illinois at Urbana-Champaign

June 25, 2020

Introduction

Let Q_n be the discrete hypercube of dimension n, that is, the graph defined on $2^{[n]}$, where two sets $A \sim B$ if and only if $|A \triangle B| = 1$.

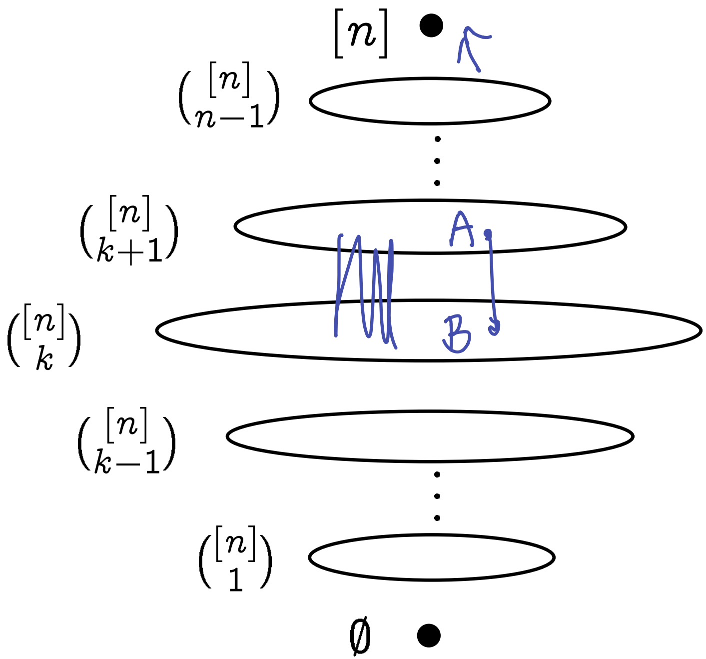


Figure: Hasse diagram of Boolean lattice

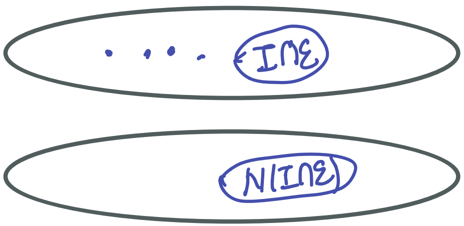
Introduction

Let $\mathcal{I}(G)$ be the set of all independent sets of the graph G.

Theorem (Korshunov and Sapozhenko, 1983)

$|\mathcal{I}(Q_n)| = 2\sqrt{e}(1+o(1))2^{2^{n-1}}$ *as* $n \to \infty$.

n-regular



even sets $|\mathcal{E}| = 2^{n-1}$

odd sets $|\mathcal{O}| = 2^{n-1}$

- trivial lower bound: $2 \cdot 2^{2^{n-1}}$;
- choose a 'small' subset K of $\mathcal{E}$ randomly, w.h.p. all its vertices have disjoint neighborhoods;
- the number of independent sets with $I \cap \mathcal{E} = K$ is $2^{2^{n-1}-n|K|}$;
- $|\mathcal{I}(Q_n)| \geq 2 \cdot \sum_{k\geq 1} \binom{2^{n-1}}{k} 2^{2^{n-1}-nk} = 2 \cdot 2^{2^{n-1}} \sum_{k\geq 1} \binom{2^{n-1}}{k} 2^{-nk}$.

$\to \sqrt{e}$

Introduction

- Sapozhenko (1989): reprove this result using the Graph Container Lemma.
- Galvin (2011): hard-core model, i.e. count the number of weighted independent sets in Q_n.

Theorem (Jenssen and Perkins 2019+)

$o(1)$

$$|\mathcal{I}(Q_n)| = 2\sqrt{e}\cdot 2^{2^{n-1}} \left(1 + \frac{3n^2 - 3n - 2}{8 \cdot 2^n} + \frac{243n^4 + \cdots}{384 \cdot 2^{2n}} + O(n^6 \cdot 2^{-3n})\right)$$

Method: the cluster expansion on abstract polymer models

Introduction

More studies on Q_n:

- Galvin (2003): the number of proper 3-colorings of Q_n is $\sim 6e2^{2^n/2}$;
- Kahn and Park (2020): the number of proper 4-colorings of Q_n is $\sim 6e2^{2^n}$;
- Jenssen and Keevash (2020+): the number of proper q-colorings of Q_n, and in general, the number of $\mathrm{Hom}(Q_n, H)$.
- Kahn and Park (2019+): the number of maximal independent sets in Q_n is $\sim 2n2^{2^n/4}$;

Introduction

Let $\mathcal{B}(n,k)$ the subgraph of Q_n induced on $\binom{[n]}{k} \cup \binom{[n]}{k-1}$.

- Duffus, Frankl and Rödl (2011): initiate the study of $\text{mis}(\mathcal{B}(n,k))$, the number of maximal independent sets.
- Ilinca and Kahn (2013): show that

$$\log \text{mis}(\mathcal{B}(n,k)) = (1+o(1))\binom{n-1}{k-1},$$

and also conjecture that $\text{mis}(\mathcal{B}(n,k)) = (1+o(1))n2^{\binom{n-1}{k-1}}$.

- Balogh, Treglown and Wagner (2016): for $|k-n/2| \le \sqrt{n}$,

$$\text{mis}(\mathcal{B}(n,k)) \ge 2^{\binom{n-1}{k-1}+Cn^{3/2}}.$$

Main result

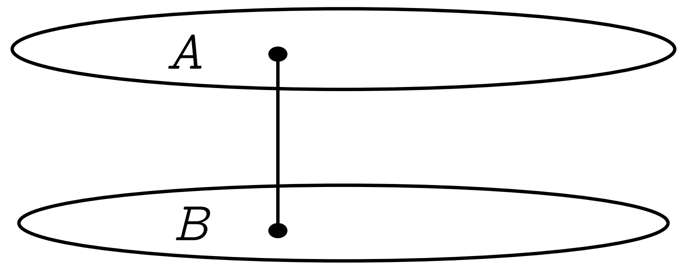
Question

What is the number of independent sets in $\mathcal{B}(n,k)$?

In particular, we focus on the case when $n = 2d-1$, $k = d$.

$$\mathcal{L}_d = \binom{[n]}{d}$$

$$\mathcal{L}_{d-1} = \binom{[n]}{d-1}$$



d-regular

$$A \sim B \text{ iff } A \supseteq B$$

Main result

For simplicity, we let $N = \binom{n}{d}$.

- Trivial lower bound: $2 \cdot 2^N$.
- Let $k = N2^{-d}$ and take a random k-set from $\mathcal{L}_d$:

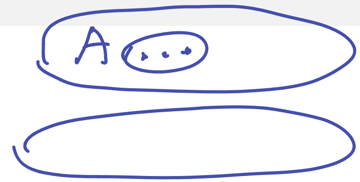
  w.h.p most of vertices have disjoint neighborhood, but there are $\Theta(d^{3/2})$ pairs of vertices, which are at distance 2 from each other.

Proposition

The number of independent sets I with $|I \cap \mathcal{L}_d| = k$ is at least

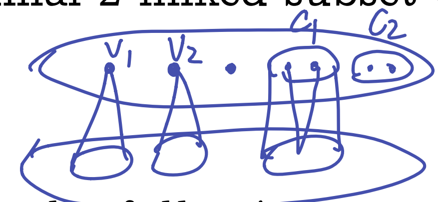
$$(1+o(1))2^N \exp\left(\overset{\sim\ 2^d/\sqrt{d}}{N2^{-d}} + (\ln 2 + o(1))\overset{d^{3/2}}{\binom{d}{2} N2^{-2d}}\right).$$

Main result



For a bipartite graph G with parts X and Y, a set $A \subseteq X$ is 2-linked if A is connected in G^2, where G^2 is a simple graph defined on $V(G)$, in which $v \sim u$ if $d_G(u, v) \leq 2$.

A 2-linked component of a set $A \subseteq X$ is a maximal 2-linked subset of A.



Theorem (Balogh, Garcia and L., 2020+)

Amost all independent sets I in $\mathcal{B}(n, d)$ have the following property. There exists $k \in \{d-1,\ d\}$ such that every 2-linked component of $I \cap \mathcal{L}_k$ is either of size 1 or 2.

Proof idea

The proof uses a variant of Sapozhenko's graph container lemma for $\mathcal{B}(n, d)$.

Let $[A] = \{v \in \mathcal{L}_d : N(v) \subseteq N(A)\}$ be the closure of A.

Graph Container Lemma

For integers $a, b \geq 1$, let

$$\mathcal{G}(a, b) = \{A \subseteq \mathcal{L}_d : A \text{ 2-linked}, |[A]| = a, |N(A)| = b\}.$$

Then $|\mathcal{G}(a, b)| \leq 2^b \binom{n}{d} \exp\left(-\Theta\left(\frac{(b-a)\ln d}{d^{2/3}}\right)\right)$ *for all* $a \leq \frac{1}{2}\binom{n}{d}$.

Proof idea

- For each $I \in \mathcal{I}$, let

$$\mathcal{LC}(I) := \{B \subseteq I \cap \mathcal{L}_d \mid B \text{ is a 2-linked component, and } |B| \geq 3\},$$

and $m(I) := \sum_{B \in \mathcal{LC}(I)} |N(B)|$.

- Let $\mathcal{U}_i$ be the collection of $I \in \mathcal{I}$ with $m(I) = i$.
- $\mathcal{I} = \mathcal{U}_0 \cup \bigcup_{i \geq 3d-3} \mathcal{U}_i$ and it is enough to prove for every $i \geq 3d-3$ we have $|\mathcal{U}_i| = o(|\mathcal{U}_0|/N)$.
- Define a bipartite graph G_i with parts $\mathcal{U}_0$ and $\mathcal{U}_i$: for $I \in \mathcal{U}_i$ and $J \in \mathcal{U}_0$,

$$I \sim J \text{ if } J = I - \bigcup_{B \in \mathcal{LC}(I)} B + K,$$

where $K \subseteq \bigcup_{B \in \mathcal{LC}(I)} N(B)$.

- Note that for $I \in \mathcal{U}_i$ we have $d_{G_i}(I) = 2^i$.

Proof idea

sets with no large comps.

- It is enough to show that $d_{G_i}(J) = o(2^i/N)$ for every $J \in \mathcal{U}_0$.
- $d_{G_i}(J) =$ number of ways to add large 2-linked components, whose neighborhood is of size i.
- First, we specify the number of components k and a decomposition $i = i_1 + \cdots i_k$.
- For 'small' i_ℓ, it is relatively easy to show the number of 2-linked sets A with $N(A) = i$ is small; for 'large' i_ℓ, we use graph container lemma.

Main result

For many combinatorial problems, getting the typical structure is harder than the corresponding enumeration problems.

Here, even though we have the typical structure, counting the typical independent sets is not a easy task.

That is why we need a new technique, the polymer method, which uses polymer models and the cluster expansion from statistical physics.

The polymer method

Let $H_{\mathcal{P}}$ be a graph defined on the finite set $\mathcal{P}$, in which every vertex has a loop edge and there is no multiple edge.

The vertices $S \in \mathcal{P}$ are called polymers.

We equip each polymer S with a complex-valued weight $w(S)$. Such a weighted graph $(H_{\mathcal{P}}, w)$ is referred as the polymer model.

Let $\Omega_{\mathcal{P}}$ be the collection of independent sets, where loops are allowed, of $H_{\mathcal{P}}$, including the empty set.

The polymer method

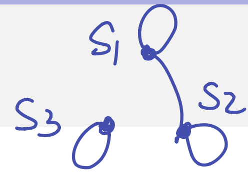
The polymer model partition function

$$\Xi(\mathcal{P}, w) = \sum_{\Lambda \in \Omega_{\mathcal{P}}} \prod_{S \in \Lambda} w(S) \tag{1}$$

is essentially a weighted independent polynomial of the polymer model $(H_{\mathcal{P}}, w)$.

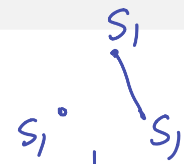
The polymer method

e.g.



$\mathcal{P} = (S_1, S_3, S_1)$

$H_{\mathcal{P}}(\mathcal{P})$



not a cluster

Let $\Gamma = (S_1, S_2, \ldots, S_k)$ be an ordered tuple of polymers, where repetitions are allowed.

Let $H_{\mathcal{P}}(\Gamma)$ be the simple graph defined on the multiset $\{S_1, S_2, \ldots, S_k\}$ with $E = \{S_i S_j : S_i \sim S_j \text{ in } H_{\mathcal{P}}\}$.

We say such a tuple Γ is a cluster if the graph $H_{\mathcal{P}}(\Gamma)$ is connected.

The weight function of a cluster Γ is defined as follows:

$$w(\Gamma) := \phi(H_{\mathcal{P}}(\Gamma)) \prod_{S \in \Gamma} w(S), \tag{2}$$

where $\phi(G)$ is the Ursell function of G.

The polymer method

Let $\mathcal{C}$ be the set of all clusters.

The cluster expansion is the formal power series of the logarithm of the partition function $\Xi(\mathcal{P}, w)$, which takes the form

$$\ln \Xi(\mathcal{P}, w) = \sum_{\Gamma \in \mathcal{C}} w(\Gamma) = L_1 + L_2 + \cdots , \tag{3}$$

where $L_k = \sum_{\Gamma \in \mathcal{C},\ |\Gamma| = k} w(\Gamma)$. $|\Gamma| = \sum_{S \in \Gamma} |S|$

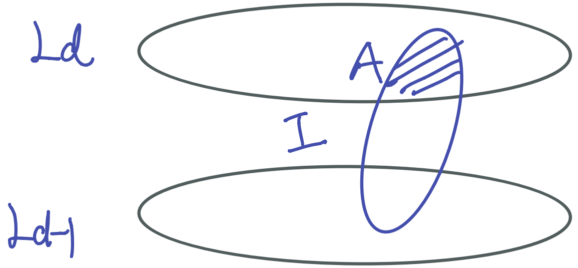
To apply the cluster expansion, we require this infinite series converges.

To prove the convergence condition, we need Sapozhenko's graph container lemma.

How to build a proper polymer model?

We use the idea of container method:

- every independent set has a fingerprint, which uniquely determines a container;
- instead of counting independent sets, we can count the number of containers and the number of independent sets in each container.



$I \cap L_d = A$

A fingerprint

Container: $A \cup (L_{d-1} \setminus N(A))$

How to build a proper polymer model?

For a given fingerprint, that is, a collection of independent 2-linked components $\{S_1, S_2, \ldots, S_k\}$, the number of independent sets is exactly

$$= \; 2^{N-\sum |N(S_i)|}.$$

Therefore, we have

$$\mathcal{I}(\mathcal{B}(n,d)) = \sum_{\{S_1,\ldots,S_k\}} 2^{N-\sum |N(S_i)|},$$

where the sum is over all fingerprints $\{S_1, \ldots, S_k\}$.

How to build a proper polymer model?

- Let $\mathcal{P}$ be the collection of 2-linked set of $\mathcal{L}_d$ (and similarly for $\mathcal{L}_{d-1}$); for $S_1, S_2 \in \mathcal{P}$, $S_1 \sim S_2$ iff $S_1 \cup S_2$ is also 2-linked;
- Let $\Omega_{\mathcal{P}}$ be the collection of independent sets of $\mathcal{P}$; each element in $\Omega_{\mathcal{P}}$ is a fingerprint!
- Let $w(S) = 2^{-|N(S)|}$. Then

$$\mathcal{I}(\mathcal{B}(n,d)) = 2 \cdot \sum_{\Lambda \in \Omega_{\mathcal{P}}} 2^{N - \sum |N(S)|} = 2 \cdot 2^N \boxed{\sum_{\Lambda \in \Omega_{\mathcal{P}}} \prod_{S \in \Lambda} w(S)}$$

$$= 2 \cdot 2^N \cdot \boxed{\Xi(P, w)} = 2 \cdot 2^N \cdot \exp\left(L_1 + L_2 + \dots\right)$$

$L_1 = \sum_{P \in C} w(P)$

$|P| = 1$

$|P| = \sum_{S \in P} |S|$

1 type :

3 types

① ② ③

v u

Main results

Theorem (Balogh, Garcia, and L., 2020+)

As $d \to \infty$, the number of independent sets in $\mathcal{B}(n,d)$ is

$$|\mathcal{I}(\mathcal{B}(n,d))| = 2(1+o(1))2^N \exp\left(N2^{-d} + \binom{d}{2} N2^{-2d}\right).$$

Proof sketch:

- Check the convergence of the above polymer model;
- Compute that $L_1 = N2^{-d}$, $L_2 = \binom{d}{2} N2^{-2d}$, and note that $L_3 = o(1)$.

- Recall that we have a lower bound

$$(1+o(1))2^N \exp\left(N2^{-d} + (\ln 2 + o(1))\binom{d}{2}N2^{-2d}\right).$$

- The behavior of clusters also 'implies' the typical structure, that is, $I \cap \mathcal{L}_d$ or $I \cap \mathcal{L}_{d-1}$ only have 2-linked components of size 1 or 2.
- Indeed, the polymer method can be used to get detailed probabilistic information about the typical structure of weighted independent sets, such as the distribution of 2-linked components of fixed size.

Thank you!