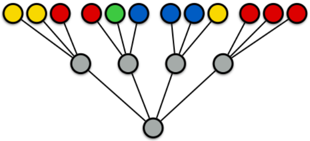
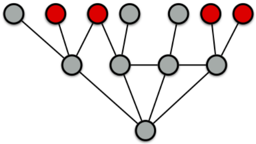
Gibbs measures in statistical physics and combinatorics

Will Perkins

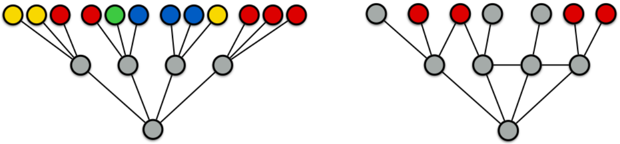
University of Birmingham

March 30, 2018





Extremal thorems for sparse graphs



Extremal theorems for bounded degree graphs

Two classic extremal theorems for regular graphs (for more see the recent survey article by Zhao):

Theorem (Kahn '01 / Zhao '10)

Let $i(G)$ be the number of independent sets of a graph G. Then for all d-regular G,

$$i(G) \leq i(K_{d,d})^{n/2d}.$$

The number of independent sets is maximized by unions of the complete d-regular bipartite graph. Proof uses the Entropy method.

Extremal theorems for bounded degree graphs

Theorem (Brégman '72 - specialized to regular graphs)

Let $m_{\mathrm{perf}}(G)$ be the number of perfect matchings of a graph G. Then for all d-regular G,

$$m_{\mathrm{perf}}(G) \leq m_{\mathrm{perf}}(K_{d,d})^{n/2d}.$$

The number of perfect matchings of any d-regular G is at most that of unions of $K_{d,d}$'s. (The extension from bipartite graphs to general graphs is due to Kahn and Lovász).

Extremal theorems for bounded degree graphs

Theorem (Csikvári '17, Lower Matching Conjecture)

For all d-regular, bipartite G,

$$m_k(G) \geq \binom{n}{k}^2 \left(\frac{d-p}{d}\right)^{n(d-p)} (dp)^{np},$$

where $p = k/n$.

Roughly, the matching polynomial is minimized by the infinite d-regular tree (or approximately by large girth graphs).

We can state these theorems in the language of statistical physics.

Independent sets in physics: the Hard-Core Model

Kahn / Zhao: for all d-regular G,

$$\frac{1}{n}\log Z_G(1) \leq \frac{1}{2d}\log Z_{K_{d,d}}(1).$$

Kahn '01; Galvin, Tetali '04: For G bipartite and d-regular, and all λ,

$$\frac{1}{n}\log Z_G(\lambda) \leq \frac{1}{2d}\log Z_{K_{d,d}}(\lambda).$$

(extended to all d-regular G by Zhao's result).

Matchings in physics: Monomer-Dimer Model

Monomer-dimer model:

$$M_G(\lambda) = \sum_{\text{Matchings } M} \lambda^{|M|}.$$

(this is an edge coloring model)

Brégman's theorem:

$$\lim_{\lambda\to\infty} \frac{1}{n} \log M_G(\lambda) - \frac{\log \lambda}{2} \leq \lim_{\lambda\to\infty} \frac{1}{2d} \log M_{K_{d,d}}(\lambda) - \frac{\log \lambda}{2}.$$

Partition functions and graph polynomials

Each of the above partition functions appears in graph theory under a different name:

- The hard-core partition function $Z_G(\lambda)$ is the independence polynomial.
- The Monomer-dimer partition function $M_G(\lambda)$ is the matching generating function (or matching polynomial).
- The Potts partition function $Z_G^q(\beta)$ is an evaluation of the Tutte polynomial and $\lim_{\beta\to\infty} Z_G^q(\beta)$ (as a function of q) is the chromatic polynomial.

How can we prove extremal bounds on these partition functions over given classes of graphs?

The occupancy fraction

The occupancy fraction of the hard-core model is the expected fraction of occupied vertices:

$$\overline{\alpha}_G(\lambda) = \frac{1}{n}\mathbb{E}_\lambda|I|.$$

It is also λ times the derivative of the free energy:

$$\overline{\alpha}_G(\lambda) = \frac{1}{n}\frac{\sum |I|\lambda^{|I|}}{Z_G(\lambda)} = \frac{\lambda Z'_G(\lambda)}{n Z_G(\lambda)}$$

$$= \lambda \cdot \frac{1}{n}(\log Z_G(\lambda))'.$$

The occupancy fraction

A small observation:

If for all $G \in \mathcal{G}$ and all $\lambda > 0$, we have $\overline{\alpha}_G(\lambda) \leq \overline{\alpha}_{G_0}(\lambda)$, then since $\overline{\alpha}_G(0) = 0$ for all G, by integrating from 0 to λ, we get a bound on the free energy:

$$\frac{1}{|V(G)|} \log Z_G(\lambda) \leq \frac{1}{|V(G_0)|} \log Z_{G_0}(\lambda).$$

The occupancy fraction

In other models, taking a derivative of the free energy yields different observables.

In the Monomer-dimer model we can define the matching occupancy fraction $\overline{\alpha}_G^M(\lambda)$ analogously.

For the Potts model, (minus) the derivative is the internal energy, or the expected number of monochromatic edges per vertex:

$$U_G^q(\beta) = \frac{1}{n}\mathbb{E}[m(G,\sigma)] = -\frac{1}{n}(\log Z_G^q(\beta))'.$$

Our strategy: prove tight bounds on these observables to give tight bounds on the corresponding partition functions.

Maximizing the independent set occupancy fraction

Theorem (Davies, Jenssen, P., Roberts)

For all d-regular G and all $\lambda > 0$,

$$\overline{\alpha}_G(\lambda) \leq \overline{\alpha}_{K_{d,d}}(\lambda).$$

This strengthens the theorem of Kahn, Galvin-Tetali, and Zhao, as we can integrate $\overline{\alpha}_G(\lambda)/\lambda$ to get $\frac{1}{n} \log Z_G(\lambda)$.

Markov property

For any Gibbs measure μ on a graph, and sets $A, B \subset V$ separated by C, we have $\boldsymbol{\sigma}_A$ and $\boldsymbol{\sigma}_B$ are independent given $\boldsymbol{\sigma}_C$.

Definitions

We say v is occupied if $v \in I$.
We say v is uncovered if $N(v) \cap I = \emptyset$.

From the Markov property,

$$\Pr[v \text{ occupied} | v \text{ uncovered}] = \frac{\lambda}{1+\lambda}.$$

A two-part experiment

1. Choose $\boldsymbol{I}$ from the hard-core model on G at fugacity λ.
2. Choose $\boldsymbol{v} \in V(G)$ uniformly at random.

Write down local information about the independent set $\boldsymbol{I}$ (and graph structure) from the perspective of $\boldsymbol{v}$.

For example, let $\boldsymbol{Y}$ be the number of uncovered neighbors of $\boldsymbol{v}$.

Two facts

Fact 1:

$$\overline{\alpha}_G(\lambda) = \frac{\lambda}{1+\lambda}\mathbb{E}(1+\lambda)^{-\boldsymbol{Y}}$$

(for triangle-free G).

Fact 2:

$$\overline{\alpha}_G(\lambda) = \frac{\lambda}{1+\lambda}\frac{1}{d}\mathbb{E}\boldsymbol{Y}$$

Key identity

$$\frac{1}{d}\mathbb{E}\boldsymbol{Y} = \mathbb{E}(1+\lambda)^{-\boldsymbol{Y}}$$

holds for all triangle-free, d-regular G, all λ.

Relax, constrain, and optimize!

$$\overline{\alpha}_G(\lambda) \le \alpha^* = \max_{\boldsymbol{Y}} \frac{\lambda}{1+\lambda} \frac{1}{d} \mathbb{E}\boldsymbol{Y}$$
$$\text{s.t. } 0 \le \boldsymbol{Y} \le d$$
$$\frac{1}{d}\mathbb{E}\boldsymbol{Y} = \mathbb{E}(1+\lambda)^{-\boldsymbol{Y}}$$

Minimizing the occupancy for triangle-free graphs

Theorem (Davies, Jenssen, P., Roberts)

For every triangle-free graph of maximum degree d,

$$\overline{\alpha}_G(\lambda) \geq (1 + o_d(1)) \frac{\log d}{d}$$

Compare to:

Theorem (Shearer '80)

For every triangle-free graph of average degree d,

$$\alpha(G) \geq (1 + o_d(1)) \frac{\log d}{d} \cdot n.$$

Implies the best known upper bound on $R(3, k)$.

Independent sets in triangle-free graphs

Conjecture: There is a factor $4/3$ gap between the **maximum** and **average** independent set size in any triangle-free graph. There is a $(2 - o_d(1))$-gap for graphs of min degree d.

Would immediately imply an improved upper bound on $R(3, k)$.

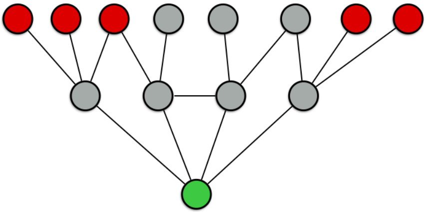
A general method

1. Pick your Gibbs measure, your favorite observable, and your class of bounded-degree graphs $\mathcal{G}$.

2. Imagine the following experiment: sample $\boldsymbol{\sigma}$ from the Gibbs distribution, and choose $\boldsymbol{v} \in G$ uniformly at random.

3. Record the graph structure up to depth t from $\boldsymbol{v}$ and the boundary condition induced by $\boldsymbol{\sigma}$: call this the (random) local view.

Figure: A local view of the hard-core model on a 4-regular graph at depth $t = 2$



A general method

4. Every graph induces a probability distribution over local views

$$p_1, p_2, \dots p_T$$

5. The expectation of your chosen observable can be computed as an expectation over this probability distribution (using the spatial Markov property).

$$\beta_G(\lambda) = \sum_{j=1}^{T} \beta_j p_j$$

6. Now relax the optimization problem to the set of all probability distributions over local views, and add constraints that all graph distributions must satisfy.

7. Solve the resulting Linear Program! Integrate to get a bound on the free energy.

Convergence and optimization

A key step is representing a graph by a probability distribution on a finite set via sampling.

A similar experiment and representation is the basis of Benjamini-Schramm convergence or local weak convergence of bounded-degree graphs

The metric is the weighted sum of total variation distances of these probability distributions from $r = 1, \ldots \infty$.

In a sense, we are taking a finite subset of these (decorated) Benjamini-Schramm numbers and optimizing over them.

Matchings in regular graphs

What about matchings? While Brégman gives the maximizer in the $\lambda \to \infty$ limit, neither the partition function statement or counting statement was known.

Theorem (Davies, Jenssen, P., Roberts)

For all d-regular G and all $\lambda > 0$,

$$\alpha_G^M(\lambda) \le \alpha_{K_{d,d}}^M(\lambda),$$

and so

$$\frac{1}{|V(G)|} \log M_G(\lambda) \le \frac{1}{2d} \log M_{K_{d,d}}(\lambda).$$

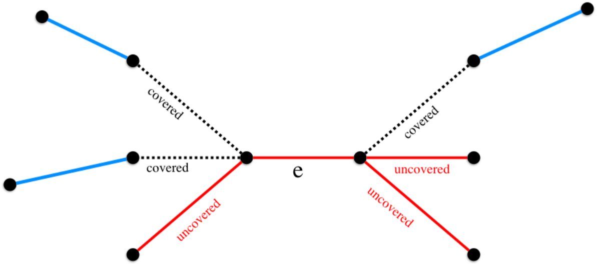
Matchings in regular graphs

This theorem generalizes Brégman's Theorem on perfect matchings in regular graphs to the matching polynomial at all values of λ.

In particular, a union of $K_{d,d}$'s maximizes the total number of matchings of any d-regular graph on the same number of vertices.

Proof idea

We can do the same two-part experiment: pick a random matching M from the Monomer-Dimer model and pick a uniformly random edge e from G.
Now look at the left and right adjoining edges of e and ask which are externally uncovered by the matching M. This defines a configuration C.



A constraint

For any d-regular G, by double counting we must have:

$$\begin{aligned}\mathbb{E}_C[\alpha_e(C)] &:= \mathbb{E}_C[\mathbb{P}[e \in M|C]] \\ &= \frac{1}{2(d-1)}\mathbb{E}_C[\mathbb{E}[|M \cap N(e)||C]] \\ &=: \mathbb{E}_C[\alpha_N(C)].\end{aligned}$$

This puts a constraint on the probability distribution a graph G induces on configurations.

We can relax the optimization problem from d-regular graphs to probability distributions on configurations:

$$\begin{aligned}\text{maximize } & \mathbb{E}_C[\alpha_e(C)] \\ \text{subject to } & \mathbb{E}_C[\alpha_e(C)] = \mathbb{E}_C[\alpha_N(C)].\end{aligned}$$

A linear program

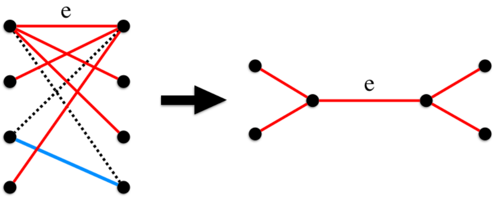
We can pose this optimization problem as a linear program with variables p_C, the probability of seeing a configuration C.

$$\begin{aligned} \text{maximize} \quad & \sum_C p_C \cdot \alpha_e(C) \\ \text{subject to} \quad & \sum p_C = 1,\ p_C \geq 0\ \forall C \\ & \sum_C p_C \cdot [\alpha_e(C) - \alpha_N(C)] = 0 \end{aligned}$$

Notice that **n** has disappeared. But the numbers of variables grows with **d**.

More constraints

For matchings this isn't enough: the distribution induced by $K_{d,d}$ has support size d, so we need at least $d - 1$ constraints (in addition to one insisting on a probability distribution). $K_{d,d}$ has support $(L, R) = (t, t)$ for $t = 0, \ldots d - 1$.

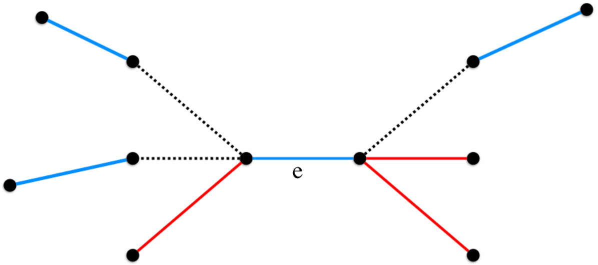


More constraints

We introduce the additional constraints that for $t = 0, \ldots d-1$,

$$\mathbb{P}[e \text{ has } t \text{ unc. left neighbors}] = \mathbb{P}[f \text{ has } t \text{ unc. right neighbors}]$$

where f is a uniformly chosen left incident edge to e (and again e is a uniform edge in the graph.



The linear program

This gives a tighter linear program:

$$\begin{aligned} \text{maximize} \quad & \sum_C p_C \cdot \alpha_e(C) \\ \text{subject to} \quad & \sum p_C = 1,\ \ p_C \geq 0\ \forall C \\ & \sum_C p_C \cdot [\gamma_e^t(C)] - \gamma_f^t(C)] = 0 \text{ for } t = 0, \ldots d-1. \end{aligned}$$

Duality

How do we prove the distribution induced by $K_{d,d}$ is the optimizer?

We look at the dual linear program:

$$\text{minimize } \Lambda_p$$

$$\text{subject to } \Lambda_p + \sum_{t=0}^{d-1} \Lambda_t \cdot [\gamma_e^t(C)] - \gamma_f^t(C)] \geq \alpha_e(C) \ \forall C$$

Beyond $K_{d,d}$ and K_{d+1}?

The extremal theory of dense graphs is rich with many different extremal constructions based on the quantities being optimized and the class of graphs considered.

Is there something comparable for sparse graphs? Can we find extremal graphs that are not $K_{d,d}$, K_{d+1}?

One answer is the infinite tree! (see recent work of P. Csikvári)

Another way is to optimize free energies subject to constraints on the local structure of the graphs.

Independent sets under girth constraints

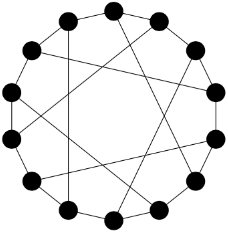
Which 3-regular graph of girth at least 5 has the **most** independent sets?

Theorem (Perarnau, P.)

For all 3-regular, graphs G of girth ≥ 5, and all $\lambda > 0$,

$$\frac{1}{n} \log Z_G(\lambda) \leq \frac{1}{14} \log Z_{H_{3,6}}(\lambda),$$

where $H_{3,6}$ is the Heawood Graph.



Heawood Graph $H_{3,6}$

Independent sets under girth constraints

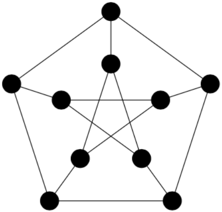
Which 3-regular, triangle-free graph has the **fewest** independent sets?

Theorem (Perarnau, P.)

For all 3-regular, triangle-free graphs G,

$$\frac{1}{n} \log i(G) \geq \frac{1}{10} \log i(P_{5,2}),$$

where $P_{5,2}$ is the Petersen Graph.



Petersen Graph $P_{5,2}$

Independent sets under girth constraints

Conjecture: All **Moore graphs** maximize or minimize the number of independent sets for regular graphs of a given minimum girth.

Examples of Moore graphs include $K_{d,d}, K_{d+1}$, Petersen graph, Heawood graph, even cycles, and odd cycles, all of which are extremal for the number of independent sets under a minimum girth constraint.

Some graphs to try next: the Tutte-Coxeter graph; the Wong graph (4,6 cage graph).

More open problems

1. Prove that $K_{d,d}$ maximizes the number of q-colorings of a d-regular graph.
   Known for bipartite graphs by Galvin and Tetali. DJPR proved $d = 3$. Davies proved $d = 4, q \geq 5$.
2. Prove that K_{d+1} minimizes the chromatic polynomial for real $q \geq d + 1$ (known for integer q).
3. For a given local view determine the maximal set of constraints all graphs must satify.
4. Incorporate global information into the constraints.

Global information

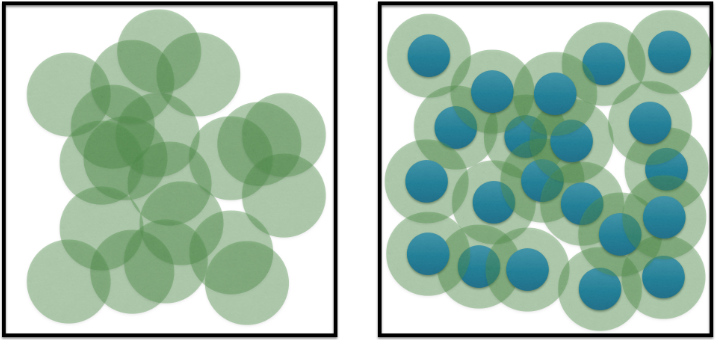
For some models, bipartiteness gives extra information: occupancy probabilities for independent sets are positively correlated on either side of the bipartition.

Theorem (DJPR)

For every d-regular, vertex-transitive, bipartite G,

$$\overline{\alpha}_G(\lambda) > \alpha_{\mathbb{T}_d}(\lambda) \,.$$

Two experiments



Experiment 1:
n spheres placed uniformly and independently at random in a large box

Experiment 2:
n spheres placed uniformly at random in a large box *conditioned* on the pairwise repulsion of their centers

Under which experiment is the expected covered volume larger?

Two experiments

Let V_k be the fraction of volume covered by k spheres of radius r placed at random.
Let E_k be the event that the centers of the k spheres are at pairwise distance at least r.

Theorem (P. '16)

At low enough densities,

$$\mathbb{E}[V_k | E_k] \geq \mathbb{E}[V_k]$$

for all $1 \leq k \leq n$.

Theorem (P. '16)

In dimension 24, there is some density so that

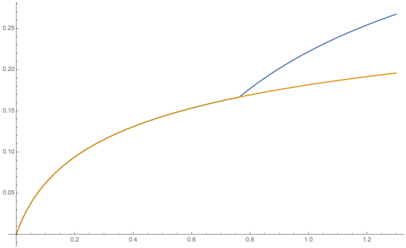
$$\mathbb{E}[V_k | E_k] < \mathbb{E}[V_k]$$

A heuristic for the phase transition

For trees or random bipartite graphs, we can say that the phase transition is the point at which the graph notices that it is bipartite.

We can plot the expected fraction of occupied vertices as a function of λ:

Figure: Random 6-regular graph vs. random bipartite 6-regular graph



A repulsion inequality

The occupancy fraction should reveal when the model notices the graph has a very large independent set.

Find a strong upper bound on $\alpha_G(\lambda)$ in the disordered phase and prove something weak at high densities to contradict this.

The repulsion inequality:

$$\alpha/\lambda \leq e^{-(d+1)\alpha}.$$

A conjecture

Conjecture

The repulsion inequality holds in the hard sphere model and the hard-core model on $\mathbb{Z}^D$ in the uniqueness phase.

$$\alpha/\lambda \leq e^{-(2D+1)\alpha}$$

for $\lambda < \lambda_c(\mathbb{Z}^D)$.

Some evidence:

- Failure of the repulsion inequality coincides asymptotically with phase transition on the d-regular infinite tree.
- In agreement with numerical simulation on $\mathbb{Z}_D$ for D up to 6.
- Repulsion inequality **always holds** in some models where we know there is no phase transition (Hard-core model in 1-D; Monomer-dimer model).

Regularity lemmas and pinning of probability measures on the cube

Notation

For $\mu \in \mathcal{P}(\Omega^n)$, let μ_i be the marginal distribution of coordinate i
Let $\mu_{i,j}$ be the joint marginal of coordinates i and j.

For $S \subseteq \Omega^n$, define the conditional measure $\mu[\cdot|S]$ by

$$\mu[\sigma|S] = \frac{\mathbf{1}_{\sigma \in S} \cdot \mu(\sigma)}{\mu(S)}.$$

Replica symmetry

We say a measure μ is ϵ-symmetric if

$$\frac{1}{n^2} \sum_{i,j \in [n]} \|\mu_{i,j} - \mu_i \otimes \mu_j\|_{TV} < \epsilon.$$

For a given model and choice of parameters, we say the model is replica symmetric if μ_G is ϵ-symmetric for $\epsilon(n) \to 0$.

If this does not hold, then the model exhibits replica symmetry breaking.

Regularity decompositions

Theorem (Bapst, Coja-Oglan '15)

For every Ω, ϵ, there exists T so that for every $\mu \in \mathcal{P}(\Omega^n)$, there exists a decomposition $S_1 \cup \cdots \cup S_T$ of Ω^n so that $\mu[\cdot|S_j]$ is ϵ-symmetric for each j.

T only depends on Ω, ϵ not on n. Proof is the same as that of Szemeredi's regularity lemma.

Regularity decompositions

- How can we get a handle on such a decomposition in general?
- Is such a decomposition unique in any sense?

A pinning lemma

For $I \subset [n], \tau \in \Omega^n$, let

$$\mu^{I,\tau} = \mu[\cdot|\{\sigma : \sigma_I = \tau_I\}].$$

Theorem (Coja-Oghlan, Krzakala, P., Zdeborova)

For every Ω, ϵ, there exists $T = O(\epsilon^{-3})$ so that for every $\mu \in \mathcal{P}(\Omega^n)$, if we choose τ according to μ, $I \subset [n]$ of size T uniformly at random, then with probability at least $1 - \epsilon$, $\mu^{I,\tau}$ is ϵ-symmetric.

Proof follows work of Montanari, see also Raghavendra-Tan.

A pinning lemma

In other words, a very small (random) perturbation to μ results in an ϵ-symmetric measure!

Think of the ferromagnetic Potts model on $\mathbb{Z}^d$:

- In the non-uniqueness regime, pinning has the effect of selecting a preferred spin.
- In the uniqueness regime, the measure is already ϵ-symmetric (correlation decay) and pinning has no effect.

Reconciling the two results

The two results have been very useful in making parts of the cavity method rigorous. We wanted to reconcile the two points of view.

- Is a tower function really required for number of parts of the regularity decomposition?
- How are the pinned measures $\mu^{I,\tau}$ related to the measures $\mu[\cdot|S_j]$?
- If we fix I, and vary τ over Ω^I, do we 'hit' all parts of the regularity decomposition with the pinned measures or do we miss some?
- Is the regularity decomposition unique in any sense?

Reconciling the two results

To measure uniqueness, we use a variant of the cut metric of Frieze and Kannan adapted to probability measures:

For $\mu, \nu \in \mathcal{P}(\Omega^n)$,

$$\Delta(\mu,\nu) = \frac{1}{n} \inf_{\gamma \in \Gamma(\mu,\nu)} \max_{\substack{I \subseteq [n] \\ B \subseteq \Omega^n \times \Omega^n \\ \omega \in \Omega}} \left| \sum_{i \in I} \sum_{(\sigma,\tau) \in B} \gamma(\sigma,\tau)(\mathbf{1}_{\sigma_i = \omega} - \mathbf{1}_{\tau_i = \omega}) \right|$$

Under the best alignment of the two measures, what is the worst possible discrepancy?

Reconciling the two results

Theorem (Coja-Oghlan, P.)

1. With probability $1-\epsilon$ over the choice of I, τ,

$$\Delta\left(\mu^{I,\tau}, \bigotimes_{i=1}^{n} \mu_i^{I,\tau}\right) < \epsilon.$$

2. With probability $1-\epsilon$ over the choice of I,

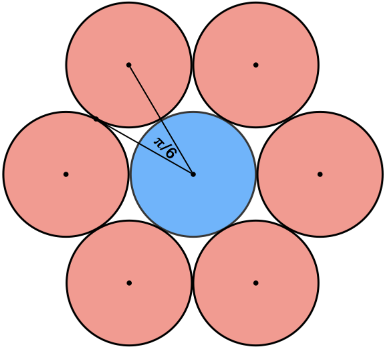
$$\Delta\left(\mu, \sum_{\tau \in \Omega^I} \mu_I(\tau)\mu^{I,\tau}\right) < \epsilon\,.$$

Reconciling the two results

We now have some answers:

- A tower function is too pessimistic: exponential in poly$(1/\epsilon)$ is enough.
- We can identify the measures $\mu^{I,\tau}$ with the measures $\mu[\cdot|S_j]$.
- The pinning procedure, with high probability over the choice of I, does hit all pieces of the measure.
- There is a canonical and simple way to decompose any $\mu \in \mathcal{P}(\Omega^n)$ into ϵ-symmetric parts.

Kissing numbers and spherical codes in high dimensions



Kissing numbers

Definition

The kissing number in dimension d, $K(d)$, is the largest number of non-overlapping unit spheres in d-dimensions than can simultaneously touch a single sphere.

Determining $K(d)$ is a very old problem in mathematics, and the answer is known only in a small number of dimensions:

$$
\begin{aligned}
K(2) &= 6 \\
K(3) &= 12 \\
K(4) &= 24 \\
K(8) &= 240 \\
K(24) &= 196560 \,.
\end{aligned}
$$

Kissing numbers

In high dimensions, very little is known.

Theorem (Chabauty '53, Shannon '59, Wyner '65)

$$K(d) \geq (1 + o(1))\sqrt{\frac{3\pi d}{8}}\left(\frac{2}{\sqrt{3}}\right)^d .$$

The proof is very simple: the covering bound (Gilbert-Varshamov bound). The bound is the reciprocal of the fraction of the surface of the unit sphere covered by a spherical cap of angular radius $\pi/3$.

Kissing numbers

An equally simple upper bound is a volume bound: caps of radius $\pi/6$ must form a packing, and so $K(d) \leq 2^d$.

Rankin ('55) obtained the bound $K(d) \leq 2^{d/2}$.

Theorem (Kabatyanskii and Levenshtein '78)

$$K(d) \leq 2^{.401\ldots \cdot d}.$$

This leaves an exponential gap between the covering lower bound and the KL upper bound - no progress has been made since!

A new lower bound

Theorem (Jenssen, Joos, P. 18+)

$$K(d) \geq (1 + o(1))\sqrt{\frac{3\pi}{8}} \log \frac{3}{2\sqrt{2}} \cdot d^{3/2} \left(\frac{2}{\sqrt{3}}\right)^d$$

as $d \to \infty$.

- Improves the Chabauty-Shannon-Wyner lower bound by a linear factor in d.
- The constant is $\approx .0639$.
- Proof uses a hard cap model and the same two-part experiment I've shown you before.

Spherical codes

Definition

A spherical code of angle θ in d dimensions is a set of unit vectors $x_1, \ldots, x_k \in \mathbb{R}^d$ so that $\langle x_i, x_j \rangle \leq \cos\theta$ for $i \neq j$.

Let $A(d, \theta)$ be the maximal size of a spherical code of angle θ in d dimensions.

Spherical codes

$A(d, \theta)$ known exactly (Rankin) for $\theta \geq \pi/2$, so we'll take $\theta \in (0, \pi/2)$ fixed as $d \to \infty$.

The spherical cap of angle θ around $x \in S_{d-1}$ is $C_\theta(x) = \{y \in S_{s-1} : \langle x, y \rangle \geq \cos\theta\}$.

Let $s_d(\theta)$ be the normalized surface area of $C_\theta(x)$.

The covering bound gives $A(d, \theta) \geq \frac{1}{s_d(\theta)}$.

Kabatyanskii and Levenshtein obtain $A(d, \theta) \leq e^{\phi(\theta)d}$ for a certain $\phi(\theta) > -\log\sin\theta$.

Our result: $A(d, \theta) \geq (1 + o(1))\frac{c_\theta d}{s_d(\theta)}$.

Sphere packings

Compare to the state of affairs for $\theta(d)$, the maximum sphere packing density in $\mathbb{R}^d$.

- Covering lower bound of 2^{-d}.
- Rogers (1947) improves to $\Omega(d \cdot 2^{-d})$ using a random lattice packing.
- Subsequent improvements to the leading constant by Davenport-Rogers '47, Ball '92, Vance '11, Venkatesh '13.
- Best upper bound of $2^{-.599..d}$ follows from the KL spherical code bound by a geometric argument. Constant factor improvement by Cohn-Zhao '14.

Lattice kissing number is very different than lattice packing density. Only very recently proved to be exponentially large in d (Vlăduț 18+).

A hard cap model

Let $\boldsymbol{X}$ be a Poisson process of intensity λ on S_{d-1} (with normalized surface area as the underlying measure) conditioned on the event that $\boldsymbol{X}$ forms a spherical code of angle θ.

The partition function is

$$Z_d^\theta(\lambda) = \sum_{k \geq 0} \lambda^k \hat{Z}_d^\theta(k)$$

where $\hat{Z}_d^\theta(0) = 1$ and

$$\hat{Z}_d^\theta(k) = \frac{1}{k!} \int_{S_{d-1}^k} \mathbf{1}_{D_\theta(x_1, \ldots x_k)} dx_1 \cdot dx_k \,.$$

A hard cap model

Let $\alpha_d^\theta(\lambda) = \mathbb{E}_\lambda |\boldsymbol{X}|$ be the expected size of the random spherical code.

Let $q(\theta)$ be the angular radius of the smallest spherical cap that contains the intersection of two spherical caps of angular radius θ whose centers are at angle θ:

$$q(\theta) = \arcsin\left(\frac{\sqrt{(\cos\theta - 1)^2(1 + 2\cos\theta)}}{\sin\theta}\right) .$$

Theorem

For $\lambda > \frac{1}{d s_d(q(\theta))}$,

$$\alpha_d^\theta(\lambda) \geq (1 + o(1))\frac{c_\theta d}{s_d(\theta)} ,$$

where $c_\theta = \log(\sin\theta / \sin q(\theta))$.

A hard cap model

We can define the hard cap model $\boldsymbol{X}_A$ on $A \subset S_{d-1}$ in the obvious way, with partition function $Z_A^\theta(lam)$.

The free area is

$$F_A^\theta(\lambda) = \mathbb{E}_\lambda \left[s \left(\{ y \in A : \langle y, x \rangle \leq \cos\theta \ \forall x \in \mathbf{X}_A \} \right) \right] .$$

The two-part experiment

1. Sample $\boldsymbol{X}_A$ from the hard cap model on A at fugacity λ.
2. Choose $\boldsymbol{v} \in A$ uniformly at random.

The local view is the random set

$$\mathbf{T}_A = \{x \in C_\theta(\mathbf{v}) \cap A : \langle x, y \rangle \leq \cos\theta \ \forall y \in \mathbf{X}_A \cap C_\theta(\boldsymbol{v})^c\};$$

that is, $\mathbf{T}_A$ is the set of all externally uncovered points of A in the spherical cap of angular radius θ around $\mathbf{v}$.

The two-part experiment

Lemma

Let $A \subseteq S_{d-1}$ be measurable and suppose $s(A) > 0$. Then the following hold.

1. $\alpha_A^\theta(\lambda)$ *is strictly increasing in* λ.
2. $\log Z_A^\theta(\lambda) \leq \lambda \cdot s(A)$.
3. $\alpha_d^\theta(\lambda) = \frac{1}{s_d(\theta)} \cdot \mathbb{E}\left[\alpha_{\mathbf{T}}^\theta(\lambda)\right]$.
4. $\alpha_A^\theta(\lambda) = \lambda \cdot \mathrm{F}_A^\theta(\lambda)$.
5. $\alpha_A^\theta(\lambda) = \lambda \cdot s(A) \cdot \mathbb{E}\left[\frac{1}{Z_{\mathbf{T}_A}^\theta(\lambda)}\right]$.
6. $\alpha_A^\theta(\lambda) \geq \lambda \cdot s(A) \cdot e^{-\lambda \cdot \mathbb{E}[s(\mathbf{T}_A)]}$.

A geometric lemma

Lemma

Let $x \in S_{d-1}$ *and* $A \subseteq C_\theta(x)$ *be measurable with* $s(A) > 0$*. Let* $\mathbf{u}$ *be a uniformly chosen point in* A*. Then*

$$\mathbb{E}[s(C_\theta(\mathbf{u}) \cap A)] \leq 2 \cdot s_d(q(\theta)) \,.$$

In particular,

$$\alpha_A^\theta(\lambda) \geq \lambda \cdot s(A) \cdot e^{-2\lambda \cdot s_d(q(\theta))} \,.$$

Key facts

1. $\alpha_A^\theta(\lambda) = \lambda \cdot s(A) \cdot \mathbb{E}\left[\frac{1}{Z_{\mathbf{T}_A}^\theta(\lambda)}\right]$.
2. $\alpha_d^\theta(\lambda) = \frac{1}{s_d(\theta)} \cdot \mathbb{E}\left[\alpha_{\mathbf{T}}^\theta(\lambda)\right]$.
3. $\log Z_A^\theta(\lambda) \leq \lambda \cdot s(A)$.
4. $\alpha_A^\theta(\lambda) \geq \lambda \cdot s(A) \cdot e^{-2\lambda \cdot s_d(q(\theta))}$.

The bound

$$\alpha \geq \lambda e^{-z^*}$$

where

$$z^* = W\left(\lambda s_d(\theta) e^{2\lambda s_d(q(\theta))}\right) .$$

$$W(x) = \log x - \log\log x + o(1) \text{ as } x \to \infty .$$

$$z^* = \log\lambda + \log s_d(\theta) - \log d - \log(c_\theta) + o(1) .$$

The bound

$$\alpha \geq (1 + o(1)) \frac{c_\theta d}{s_d(\theta)} = \Theta\left(\frac{d^{3/2}}{\sin^d \theta}\right)$$

Remarks

- The bound on λ doesn't tell us anything about the size of spherical codes, but does give a non-trivial bound on the number of spherical codes (or kissing configurations)
- The exact same proof works for sphere packings! We get a lower bound of $\Theta(d \cdot 2^{-d})$ and a lower bound on the entropy of sphere packings.

Open problems

- Improve the lower bound, even by a constant factor.
- Prove a lower bound on the minimum covering density of spherical caps of angle θ. $\Omega(d)$?
- Other applications of this method? Certainly in coding theory, anywhere the Gilbert-Varshamov bound is used.
- What are dense sphere packings like in high dimensions? Ordered? Disordered? No one knows!

Thank you all for your attention, and thanks especially to Jan Hladky for organizing a great workshop!