BIRTHDAY INEQUALITIES, REPULSION, AND HARD SPHERES

WILL PERKINS

Abstract. We study a *birthday inequality* in random geometric graphs: the probability of the empty graph is upper bounded by the product of the probabilities that each edge is absent. We show the birthday inequality holds at low densities, but does not hold in general. We give three different applications of the birthday inequality in statistical physics and combinatorics: we prove lower bounds on the free energy of the hard sphere model and upper bounds on the number of independent sets and matchings of a given size in d-regular graphs.

The birthday inequality is implied by a *repulsion inequality*: the expected volume of the union of spheres of radius r around n randomly placed centers increases if we condition on the event that the centers are at pairwise distance greater than r. Surprisingly we show that the repulsion inequality is not true in general, and in particular that it fails in 24-dimensional Euclidean space: conditioning on the pairwise repulsion of centers of 24-dimensional spheres can *decrease* the expected volume of their union.

How many people must be in a room so that the chance at least two share a birthday is at least 1/2? This is the 'Birthday Problem', and the answer is that 23 people is enough (assuming that the birthdays are independently and identically distributed).

Our starting point is an elementary inequality, which we will call the *birthday inequality*:

Lemma 1 (The Birthday Inequality)**.** *Suppose n people have birthdays chosen independently and uniformly at random from one of m birthdays. Let E_n be the event that no two people share a birthday, and $p = 1/m$ the probability that two given people share a birthday. Then*

$$\Pr[E_n] \le (1-p)^{\binom{n}{2}}$$

Proof. Let E_k be the event that there are no shared birthdays among the first k people. Let V_k be the fraction of birthdays covered by the first k people. Then

$$\Pr[E_n] = \mathbb{E}[1 - V_{n-1} | E_{n-1}] \cdot \Pr[E_{n-1}]$$

We assume inductively that $\Pr[E_k] \le (1-p)^{\binom{k}{2}}$, and note that

$$\mathbb{E}[1 - V_k | E_k] = 1 - \frac{k}{m} \le \left(1 - \frac{1}{m}\right)^k = (1-p)^k$$

which shows that $\Pr[E_{k+1}] \le (1-p)^{\binom{k+1}{2}}$ for all $k \ge 1$. □

We are interested in geometric birthday inequalities, and in particular in settings relevant to two models from statistical physics: the hard sphere model and the hard core lattice gas model. In these models, particles are placed at random in a metric space $\mathcal{X}$ equipped with

---

Date: June 3, 2015.

2010 *Mathematics Subject Classification.* Primary: 60D05, 05C80; Secondary: 05C69, 05C70, 82B21.

Key words and phrases. Hard sphere model, random geometric graphs, hard core model, independent sets, matchings, sphere packing.

a probability measure μ (e.g. the unit cube or a subset of the d-dimensional integer lattice with uniform measure) conditioned on all pairwise distances between particles being larger than some threshold r. In this setting the birthday inequality supposes an upper bound on the probability that no two particles are within distance r when n particles are placed independently at random according to μ.

Definition 1. *Let $X_1, X_2, \dots X_n$ be independently sampled points from a space $\mathcal{X}$ according to the distribution μ. Then a birthday inequality holds if*

$$\Pr[E_n] \le (1-p)^{\binom{n}{2}} \tag{1}$$

where E_n is the event $\{\wedge_{1\le i<j\le n} d(X_i, X_j) > r\}$ and $p := \Pr[d(X_1, X_2) \le r]$.

The quantity on the right is what the probability on the left would be if all pairwise interactions were independent, and so the birthday inequality is a statement about correlations of these events.

A related inequality is the following *repulsion inequality*.

Definition 2. *In the setting above, the repulsion inequality holds if*

$$\mathbb{E}[V_k|E_k] \ge \mathbb{E}[V_k] \tag{2}$$

where E_k is the event that the centers $X_1, X_2, \dots X_k$ are at pairwise distance greater than r, V_k is the volume fraction of $\mathcal{X}$ covered by the union of the closed balls of radius r around $X_1, \dots X_k$, and the expectations are taken over choosing $X_1, \dots X_k$ independently at random in $\mathcal{X}$ according to μ.

The repulsion inequality states that conditioning on the event that the centers of randomly placed balls of radius r are at pairwise distance greater than r does not decrease the expected volume of their union (as compared to the unconditional expectation). The repulsion inequality has the flavor of a probabilistic version of the Kneser-Poulsen conjecture [16, 12]: moving a set of spheres in Euclidean space so that all pairwise distances between their centers do not decrease cannot decrease the volume of the union of the spheres. This was proved in two dimensions by Bezdek and Connelly [1], but is open in higher dimensions.

While the repulsion inequality seems intuitively obvious, we show in Corollary 2 that it is not always true; in particular it fails in dimension 24.

As in the proof of Lemma 1, the birthday inequality on n points is implied if the repulsion inequality holds for all k between 1 and $n-1$. We write $\Pr[E_n] = (1-\mathbb{E}[V_{n-1}|E_{n-1}])\cdot\Pr[E_{n-1}]$ and continue inductively. The unconditional expectation $\mathbb{E}[V_k] = 1-(1-p)^k$, and so if the repulsion inequality holds for all $1 \le k \le n-1$, the birthday inequality holds.

We will show that at sufficiently low particle densities, the repulsion inequality holds in both the hard sphere and hard core models. This leads to bounds on the free energy in both models via the birthday inequality. However, we will also show that at sufficiently high densities the birthday inequality can fail. We conclude by conjecturing that the failure of the repulsion inequality can be used to indicate the fluid/solid phase transition.

1. Hard spheres

The hard sphere model is a model of particles as non-overlapping spheres: there are no forces in the model besides the hard constraint that two spheres cannot overlap. We define the hard sphere model on $\mathcal{T}^d$, the d-dimensional unit torus.

Definition 3. *The hard sphere model $H_d(n,r)$ consists of a uniformly random configuration of n spheres of radius $r/2$ in $\mathcal{T}^d$, conditioned on the event that the centers of the n spheres are at pairwise distance greater than r.*

An important quantity in statistical physics is the partition function:

Definition 4. *The partition function, $Z_d(n,r)$, of the hard sphere model on $\mathcal{T}^d$ is defined as:*

$$Z_d(n,r) = \int_{\mathcal{T}^d} \cdots \int_{\mathcal{T}^d} \mathbf{1}_{E_n} \, dx_1 \cdots dx_n \tag{3}$$

where E_n is the event that $d(x_i, x_j) > r$ for all $1 \leq i < j \leq n$.

We define the density α of $H_d(n,r)$ as the fraction of volume of $\mathcal{T}^d$ occupied by the spheres of radius $r/2$ around the n centers, i.e., $\alpha = n(r/2)^d v_d$ where v_d is the volume of the unit ball in $\mathbb{R}^d$. As α is the density of the random sphere packing given by $H_d(n,r)$ it must lie between 0 and the maximum sphere packing density in d dimensions.

Definition 5. *The free energy of the hard sphere model at density α is:*

$$F_d(\alpha) = - \lim_{n \to \infty} \frac{1}{n} \log Z_d(n, r_n(\alpha)) \tag{4}$$

where $r_n(\alpha) = 2(\alpha/(n v_d))^{1/d}$.

Physicists believe that the hard sphere model in dimension $d \geq 2$ undergoes a fluid / solid phase transition as the density of the spheres increases: at low densities configurations show no long-range order, while after the phase transition long-range order emerges. For an introduction to the hard sphere model see [14] and the references therein. In dimension $d = 1$ there is no phase transition and the model is solved; that is, an explicit expression for the free energy is known, and it has no non-analytic points [19]. Mathematicians have proved rigorous lower bounds on the density at which a Markov chain to sample from the model mixes rapidly [11, 18, 6, 8]. See [17, 2] for a discussion of mathematical proofs of phase transitions in continuous hard core models, and in the second a proof of a phase transition in a system with zipper-like molecules.

We define the model with spheres of radius $r/2$ because it will be convenient to view the hard sphere model from the perspective of the random geometric graph $G_d(n,r)$: n points placed uniformly and independently at random in $\mathcal{T}^d$ with an edge placed between pairs of points at distance at most r.

The following proposition relates the hard sphere model to the random geometric graph, and follows immediately from Definition 3.

Proposition 1.

$$Z_d(n,r) = \Pr[G_d(n,r) \textit{ is empty}].$$

In what follows we will parameterize both the hard sphere model and the random geometric graph by $p := v_d r^d$, the probability that two uniformly random points in $\mathcal{T}^d$ are at distance at most r. Abusing notation we will write $G_d(n,p)$ for $G_d(n, r(p))$ where $v_d r(p)^d = p$. This parameterization gives some intuition for the Birthday Inequality: if $G(n,p)$ is the Erdős-Rényi random graph on n vertices (every edge present independently with probability p), then the birthday inequality is $\Pr[G_d(n,p) \text{ is empty}] \leq \Pr[G(n,p) \text{ is empty}]$. In fact, if we fix n and p and let $d \to \infty$, then $\Pr[G_d(n,p) \text{ is empty}] \to \Pr[G(n,p) \text{ is empty}]$ (Theorem 2 in [5] for the RGG defined on the surface of the d-dimensional unit sphere).

The birthday inequality holds in dimension 1 for all values of p:

Proposition 2. *For all $p \in [0,1]$,*

$$\Pr[G_1(n,p) \text{ is empty}] \le (1-p)^{\binom{n}{2}}$$

Proof. For $p > 2/n$, we are beyond the maximum packing density, and so $\Pr[G_1(n,p)$ is empty$] = 0$ and the inequality holds. For $p \le 2/n$, we can write the left-hand side explicitly: $\Pr[G_1(n,p)$ is empty$] = (1-np/2)^{n-1}$, and then it is a calculus exercise to show that $(1-np/2)^{n-1} \le (1-p)^{\binom{n}{2}}$ for $p \le 2/n$. □

Our first main result of this section is to show that in any dimension, at a low enough density the birthday inequality holds. We do this via the repulsion inequality (2).

Theorem 1. *For the the hard sphere model on $\mathcal{T}^d$, for densities $\alpha \le 2^{-2-3d}$ the repulsion inequality holds.*

Theorem 1 and the birthday inequality immediately imply a lower bound on the free energy of the hard sphere model at sufficiently low densities, which to the best of our knowledge is new.

Corollary 1. *For $\alpha \le 2^{-2-3d}$, $F_d(\alpha) \ge 2^{d-1}\alpha$.*

Proof of Theorem 1. We first define some notation. For a collection of k centers in $\mathcal{T}^d$, let V_k be the volume of points in $\mathcal{T}^d$ at distance at most r from one of the k centers, i.e. the volume of the union of balls of radius r around the centers. Let E_k be the event that the k centers are at pairwise distance greater than r. As always, we have $p = v_d r^d$, and we assume $\alpha \le 2^{-2-3d}$, i.e. $p \le 4^{-1-d}/n$. Our goal is to prove the repulsion inequality $\mathbb{E}[V_k|E_k] \ge \mathbb{E}[V_k]$ where the randomness is in placing each center uniformly and independently at random in $\mathcal{T}^d$.

We will prove the following estimate for all $1 \le k \le n-1$:

$$\mathbb{E}[V_k|E_k] \ge kp - \binom{k}{2} p^2 \frac{1-4^{-d}}{(1-kp)^2} \tag{5}$$

To complete the proof of the Theorem from (5) we use inclusion/exclusion to bound $\mathbb{E}[V_k] = 1-(1-p)^k \le kp - \binom{k}{2}p^2 + \binom{k}{3}p^3$, and so

$$\mathbb{E}[V_k|E_k] - \mathbb{E}[V_k] \ge \binom{k}{2} p^2 \left(1 - \frac{1-4^{-d}}{(1-kp)^2} - \frac{k-2}{3}p\right)$$

which is non-negative when $p \le 4^{-1-d}/n$.

To prove (5) we use inclusion/exclusion and linearity of expectation to get the lower bound

$$\mathbb{E}[V_k|E_k] \ge kp - \sum_{i<j} \mathbb{E}[V(i,j)|E_k] = kp - \binom{k}{2}\mathbb{E}[V(1,2)|E_k]$$

where $V(i,j)$ is the volume covered by the balls of radius r around both centers i and j, i.e. their overlap volume.

Now let x be a fixed point in $\mathcal{T}^d$ (say the origin), $A_x^{1,2}$ the event that x is covered by the balls of radius r around both centers 1 and 2, and E_2 the event that centers 1 and 2 are at distance greater r. Then we have

$$\begin{aligned}
\mathbb{E}[V(1,2)|E_k] &= \Pr[A_x^{1,2}|E_k] \\
&= \frac{\Pr[A_x^{1,2} \cap E_k]}{\Pr[E_k]} \\
&= \frac{\Pr[A_x^{1,2} \cap E_2] \cdot \Pr[E_k|A_x^{1,2} \cap E_2]}{\Pr[E_2] \cdot \Pr[E_k|E_2]} \\
&= \Pr[A_x^{1,2}|E_2] \cdot \frac{\Pr[E_k|A_x^{1,2} \cap E_2]}{\Pr[E_k|E_2]}
\end{aligned}$$

First note that $\Pr[E_k|A_x^{1,2} \cap E_2] \leq \Pr[E_{k-2}]$. Next, write

$$\begin{aligned}
\Pr[E_k|E_2] &= \frac{\Pr[E_{k-2}] \Pr[E_k|E_{k-2}]}{\Pr[E_2]} \\
&\geq \Pr[E_{k-2}] \frac{(1-(k-2)p)(1-(k-1)p)}{1-p} \\
&\geq \Pr[E_{k-2}] \frac{(1-kp)^2}{1-p}
\end{aligned}$$

where we have used the inequalities $\Pr[E_{k-1}|E_{k-2}] \geq 1-(k-2)p$ and $\Pr[E_k|E_{k-1}] \geq 1-(k-1)p$ which follow from the union bound: the volume of the union of balls of radius r around $k-2$ centers is at most $(k-2)p$.

This gives

$$\frac{\Pr[E_k|A_x^{1,2} \cap E_2]}{\Pr[E_k|E_2]} \leq \frac{1-p}{(1-kp)^2}$$

Finally we upper bound $\Pr[A_x^{1,2}|E_2]$. We write

$$p^2 = \Pr[A_x^{1,2}] = p \Pr[A_x^{1,2}|\overline{E_2}] + (1-p) \Pr[A_x^{1,2}|E_2]$$

The probability that three given points form a triangle in the random geometric graph with connection radius r is $p \cdot \Pr[A_x^{1,2}|\overline{E_2}]$. A lower bound for the probability of forming a triangle is the probability that the first two points fall in a ball of radius $r/2$ around the third, which has probability $p^2 4^{-d}$. Putting this together we have

$$\Pr[A_x^{1,2}|E_2] \leq \frac{p^2 - p^2 4^{-d}}{1-p}$$

and

$$\mathbb{E}[V(1,2)|E_k] \leq p^2 \frac{1-4^{-d}}{(1-kp)^2}.$$

which gives (5).

□

Our next result is that the birthday inequality does not hold in general. We show this in dimension 24, using the fact that there is a sphere packing of particularly high density.

Theorem 2. *In dimension 24, the birthday inequality fails for large enough n at densities $\alpha \in ((.79)^{24} \cdot \rho, \rho)$, where $\rho = .001929$.*

This theorem implies that the repulsion inequality fails at some density in dimension 24:

Corollary 2. *For large enough n, there exists some r so that when n spheres with centers $x_1, \ldots x_n$ are placed uniformly at random in $\mathcal{T}^{24}$,*

$$\mathbb{E}\left[vol(\cup_{i=1}^n B(x_i, r))\,|d(x_i, x_j) > r \text{ for all } i \neq j\right] < \mathbb{E}\left[vol(\cup_{i=1}^n B(x_i, r))\right],$$

where $B(x_i, r)$ is the closed ball of radius r around the center x_i.

In other words, conditioning on the pairwise repulsion of the centers of the spheres can *decrease* the expected volume of their union!

Working on the torus is not essential to the result: the same holds if the centers of the spheres are chosen at random in a box in $\mathbb{R}^{24}$ large enough so that boundary effects are negligible.

Proof of Theorem 2. Consider some packing of n spheres of radius r_p in $\mathcal{T}^d$ with density ρ. Place a sphere of radius $(1-t)r_p$ for $0 < t < 1$ around each center of the packing. If we place a new set of centers within these spheres of radius $(1-t)r_p$, then the spheres of radius $tr_p =: r/2$ around them will be disjoint. The density of such a configuration is $\alpha = nv_d(tr_p)^d = t^d\rho$, or in other words, $t = (\alpha/\rho)^{1/d}$. We can lower bound the probability that n random centers of spheres of radius $r/2$ will be disjoint by the probability that each of the n centers falls into a distinct sphere of radius $(1-t)r_p$ around the centers of the packing:

$$\Pr[G_d(n, r) \text{ is empty}] \geq \frac{n!}{n^n}((1-t)^d\rho)^n = \frac{n!}{n^n}\left(1 - (\alpha/\rho)^{1/d}\right)^{dn}\rho^n$$

and

$$-\frac{1}{n}\log\Pr[G_d(n, r) \text{ is empty}] \leq 1 - d\log\left(1 - (\alpha/\rho)^{1/d}\right) - \log\rho + o(1)$$

The Birthday Inequality, however, asserts that

$$-\frac{1}{n}\log\Pr[G_d(n, r) \text{ is empty}] \geq 2^{d-1}\alpha$$

For $d = 24$, there is a sphere packing of $\mathbb{R}^{24}$ of density $\frac{\pi^{12}}{12!}$ via the Leech lattice [13, 4]. For any $\epsilon > 0$, and all large enough n, we can find a packing of $\mathcal{T}^{24}$ with n disjoint spheres which has density at least $\frac{\pi^{12}}{12!} - \epsilon$. Choosing $\rho = .001929 \approx \frac{\pi^{12}}{12!} - 5 \cdot 10^{-7}$, we can compare the birthday inequality lower bound on the free energy $BI(\rho, t)$ to the cell model upper bound $CM(\rho, t)$.

$$F(t) := BI(\rho, t) - CM(\rho, t) = \frac{\rho}{2}(2t)^{24} - 1 + 24\log(1-t) + \log(\rho)$$

A calculation gives $F(.79) > 0$ and $F'(t) > 0$ for $t \in (.79, 1)$. This proves Theorem 2. Corollary 2 follows immediately since the sequence of repulsion inequalities implies the birthday inequality. ∎

The *hard square model* is the hard sphere model under l_∞ distance: configurations of disjoint d-dimensional axis-parallel cubes. Cubes pack particularly nicely, with a maximum packing density of 1. We use this to show that the birthday inequality fails in all sufficiently high dimensions.

Theorem 3. *The birthday inequality fails in the hard square model for some range of densities in dimension $d \geq 6$. For $d = 6$, the inequality fails for $\alpha \in (.4, .95)$. For $d > 6$, the inequality fails for $\alpha \in (\underline{\eta}_d, \overline{\eta}_d)$, where $\underline{\eta}_d \sim 2^{1-d} \log(2) \cdot d$ as $d \to \infty$, $\overline{\eta}_d \to 1$ as $d \to \infty$.*

Proof. Here we ask for a given d if there is some $\alpha \in (0, 1)$ so that

$$1 - d \log(1 - \alpha^{1/d}) > 2^{d-1} \alpha$$

A numerical calculation for $d = 6$ and some calculus give the theorem. □

2. The hard core model

Assume n is such that $n^{1/d}$ is an even integer. Let $\mathbb{Z}_d(n)$ be the d-dimensional discrete torus of sidelength $n^{1/d}$ (for a total of n sites). Assume α is such that αn is an integer. We define a fixed-density hard core model as follows:

Definition 6. *The fixed-density hard core model $HC_d(n, \alpha)$ for $\alpha \in [0, 1/2]$ consists of a uniformly chosen random independent set of $k = \alpha n$ sites in $\mathbb{Z}_d(n)$.*

This model is a natural discretization of the hard sphere model. It is closely related to the hard core model with an activity parameter λ: an independent set $I \subset \mathbb{Z}_d(n)$ chosen with probability proportional to $\lambda^{|I|}$. By conditioning on $|I| = \alpha n$ we obtain the fixed-density hard core model defined above. In the terminology of statistical physics, the fixed-density model is the canonical ensemble, while the activity parameter model is the grand canonical ensemble.

We can define the partition function and free energy of the hard core model:

Definition 7. *The partition function, $Z_d(n, k)$, of the hard core model is defined as:*

$$Z_d(n, k) = IS(k) := \#\text{ of independent sets of size } k \text{ in } \mathbb{Z}_d(n) \tag{6}$$

We can write $Z_d(n, k) = \frac{n^k}{k!} \Pr[X_k \text{ independent}]$ where X_k is a (multi)-set of k independent and uniformly chosen sites from $\mathbb{Z}_d(n)$.

Definition 8. *The free energy of the hard core model at density α is:*

$$F_d(\alpha) = \lim_{n \to \infty} \frac{1}{n} \log Z_d(n, \alpha n) \tag{7}$$

Note that we do not take the negative of the log partition function here, and so while we obtained lower bounds on the free energy of the hard sphere model, here we will obtain upper bounds.

We can write the free energy in terms of the probability X_k is an independent set:

$$F_d(\alpha) = \alpha - \alpha \log \alpha + \lim_{n \to \infty} \frac{1}{n} \log \Pr[X_{\alpha n} \text{ independent}] \tag{8}$$

We can also define the fixed-size hard core model on the d-dimensional Hamming cube Q_d, with vertex set $\{0,1\}^d$ and edges between vectors that differ in exactly one coordinate. The partition function and free energy are defined as in the hard core model on $\mathbb{Z}_d(n)$.

Our first theorem of this section is that the repulsion inequality (and thus the Birthday Inequality) holds in the hard core model at a sufficiently low density for any d-regular graph. We consider a d-regular graph G on n vertices, and select a set of k vertices X_k uniformly at random with replacement. We consider two vertices in X_k to form an edge if they are neighbors in G or if they are identical; so if X_k has no edges, it is an independent set of size k in G. Thus we have $p = \frac{d+1}{n}$, the probability that two randomly chosen vertices form an edge.

Theorem 4. *For the the hard core model on any d-regular graph G on n vertices, at densities $\alpha \leq (d+1)^{-2}$, the repulsion inequality (2) holds.*

As a corollary via the birthday inequality we get an improved upper bound on the number of independent sets of size αn in all d-regular graphs, for $\alpha \leq (d+1)^{-2}$.

Corollary 3. *For $\alpha \leq (d+1)^{-2}$, the number of independent sets of size αn in any d-regular graph G satisfies:*

$$IS(\alpha n) \leq \frac{n^{\alpha n}}{(\alpha n)!} \left(1 - \frac{d+1}{n}\right)^{\binom{\alpha n}{2}}$$

On the scale of the free energy, this gives:

$$\frac{1}{n} \log IS(\alpha n) \leq \alpha - \alpha \log \alpha - \alpha^2 \frac{d+1}{2}. \tag{9}$$

For $\alpha \leq (d+1)^{-2}$, Corollary 3 improves the bound for d-regular bipartite graphs given by Carroll, Galvin, and Tetali [3][1]. Corollary 3 holds for all d-regular graphs. Specializing to $\mathbb{Z}_d(n)$ and Q_d we get upper bounds of $\alpha\,(1 - \log \alpha - \alpha(2d+1)/2)$ and $\alpha\,(1 - \log \alpha - \alpha(d+1)/2)$ respectively on the normalized log of number of independent sets of size αn. As far as we know these are the best bounds known on the number of independent sets of a given size in $\mathbb{Z}_d(n)$ and Q_d at these densities.

Proof of Theorem 4. The proof is essentially the same as the proof of Theorem 1, and this is one of the motivations of this work: to find methods for analyzing the hard core model that generalize to the hard sphere model.

Let V_k be the fraction of vertices in G at distance at most 1 to a set of k randomly chosen vertices. Let E_k be the event that the set of k random vertices is at pairwise distance at least 2 in G. We can assume that $k \geq 2$ (and thus $n \geq 2(d+1)^2$) since the case $k = 1$ is immediate. We let $p = \frac{d+1}{n}$, the probability that two random chosen vertices in G coincide or are neighbors. We will prove the following estimate:

$$\mathbb{E}[V_k | E_k] \geq \frac{k(d+1)}{n} - \binom{k}{2} \frac{d(d-1)}{n^2(1-kp)^2} \tag{10}$$

[1]For this range of α, the best upper bound in [3] on $IS(\alpha n)$ is the third bound given in Theorem 1.6, $2^{\alpha n}\binom{n/2}{\alpha n}$. On the scale of the free energy this is $-\alpha \log(\alpha) - (1/2 - \alpha)\log(1 - 2\alpha)$. Some calculus shows that the bound in (9) is lower for $\alpha \in (0, (d+1)^{-2})$.

where the randomness is in selecting the k vertices uniformly and independently at random.

By inclusion/exclusion we have

$$\mathbb{E}[V_k] = 1-(1-p)^k \leq kp - \binom{k}{2}p^2 + \binom{k}{3}p^3 = \frac{k(d+1)}{n} - \binom{k}{2}\frac{(d+1)^2}{n^2}\left(1 - \frac{(k-2)(d+1)}{3n}\right)$$

and using (10) we get

$$\begin{aligned}\mathbb{E}[V_k|E_k] - \mathbb{E}[V_k] &\geq \binom{k}{2}p^2\left(1 - \frac{d(d-1)}{(d+1)^2}\cdot\frac{1}{(1-kp)^2} - \frac{(k-2)(d+1)}{3n}\right)\\ &\geq \binom{k}{2}p^2\left(1 - \frac{d(d-1)}{(d+1)^2}\cdot\frac{1}{(1-kp)^2} - \frac{kp}{3}\right)\end{aligned}$$

This is non-negative when $\alpha = k/n \leq \frac{1}{(d+1)^2}$: the RHS is decreasing in k, and so it is enough to prove when $k = n/(d+1)^2$. This follows from an elementary calculation and proves Theorem 4, modulo the estimate (10).

To prove (10), we use inclusion/exclusion again to bound

$$\mathbb{E}[V_k|E_k] \geq kp - \binom{k}{2}\mathbb{E}[V(1,2)|E_k],$$

where $V(1,2)$ is the fraction of vertices in G at distance 0 or 1 of the first and second of the k randomly selected vertices. Let $A_v^{1,2}$ be the event that vertex v neighbors both of the first two selected vertices. We write

$$\begin{aligned}\mathbb{E}[V(1,2)|E_k] &= \frac{\mathbb{E}[V(1,2)\cdot\mathbf{1}_{E_k}]}{\Pr[E_k]}\\ &= \frac{1}{n}\sum_{v\in G}\frac{\Pr[A_v^{1,2}\cap E_k]}{\Pr[E_k]}\\ &= \frac{1}{n}\sum_{v\in G}\frac{\Pr[A_v^{1,2}\cap E_2]\cdot\Pr[E_k|A_v^{1,2}\cap E_2]}{\Pr[E_2]\cdot\Pr[E_k|E_2]}\\ &= \frac{1}{n}\sum_{v\in G}\Pr[A_v^{1,2}|E_2]\cdot\frac{\Pr[E_k|A_v^{1,2}\cap E_2]}{\Pr[E_k|E_2]}\end{aligned}$$

If the neighbors of v form a clique then that term in the sum is 0. We assume from here that there is at least one edge missing from the subgraph of v's neighbors.

Consider $\frac{\Pr[E_k|A_v^{1,2}\cap E_2]}{\Pr[E_k|E_2]}$. Again we have $\Pr[E_k|A_v^{1,2}\cap E_2] \leq \Pr[E_{k-2}]$, and

$$\begin{aligned}\Pr[E_k|E_2] &\geq \frac{\Pr[E_{k-2}]\Pr[E_k|E_{k-2}]}{\Pr[E_2]}\\ &\geq \frac{\Pr[E_{k-2}](1-kp)^2}{1-p}\end{aligned}$$

which gives

$$\frac{\Pr[E_k|A_v^{1,2}\cap E_2]}{\Pr[E_k|E_2]} \leq \frac{1-p}{(1-kp)^2} \tag{11}$$

Next note that for any d-regular graph G,

$$\frac{1}{n}\sum_{v\in G}\Pr[A_v^{1,2}|E_2] = \frac{1}{n}\frac{n\binom{d}{2} - 3\cdot \#C_3's \text{ in } G}{\binom{n}{2} - dn/2} \le \frac{d}{n}\cdot\frac{d-1}{n(1-p)} \tag{12}$$

Inequalities (11) and (12) give

$$\mathbb{E}[V(1,2)|E_k] \le \frac{d(d-1)}{n^2(1-kp)^2}$$

and thus (10).

□

We now show that the birthday inequality fails in general for d-regular, bipartite graphs with $d \ge 6$.

Theorem 5. *For $d \ge 6$, there exists constants $\alpha_d^l \in (0, 1/2)$ so that for n large enough, the birthday inequality fails for the hard core model on any d-regular, bipartite graph G on n vertices at densities $\alpha \in [\alpha_d^l, 1/2]$. Asymptotically, $\alpha_d^l \sim 2\log 2/d$ as $d \to \infty$.*

Proof. For a lower bound on the number of independent sets of size αn in G, we use the *parity* lower bound: any subset of one side of the bipartition is an independent set, and so

$$\mathrm{IS}(\alpha n) \ge \binom{n/2}{\alpha n}$$

The corresponding bound on the free energy is

$$F_d(\alpha) \ge -\alpha\log(2\alpha) - (1/2-\alpha)\log(1-2\alpha) + o(1)$$

The birthday inequality asserts the upper bound:

$$F_d(\alpha) \le \alpha\left(1 - \log\alpha - \alpha\frac{d+1}{2}\right) + o(1)$$

Some calculus shows that these bounds cross for $d \ge 6$ (see Figure 1), and that asymptotically as $d \to \infty$, the crossing point is $\alpha_d^l \sim 2\log 2/d$.

□

3. Matchings

In this section we use the repulsion inequality to give bounds on the number of matchings of size k in a d-regular graph G on n vertices. Such a graph has $nd/2$ edges, and each edge shares a vertex with $2d-2$ other edges. We let $p = \frac{2d-1}{nd/2}$, the probability that two uniformly random edges (with replacement) coincide or intersect at a vertex. Then the birthday inequality asserts that $\Pr[E_k] \le (1-p)^{\binom{k}{2}}$, where E_k is the event that k edges chosen uniformly at random from G form a matching of size k. The repulsion inequality states that $\mathbb{E}[V_k|E_k] \ge \mathbb{E}[V_k]$, where V_k is the fraction of edges covered by a set of k edges: the fraction of edges that are contained in or intersect the set. Since a matching in G is an independent set in the line graph of G, we could apply Theorem 4 to get a bound, but we can do somewhat better directly, since for $d \ge 3$, the line graph of a d-regular graph contains many triangles. Let $M(k)$ be the number of matchings consisting of k edges in G. We show:

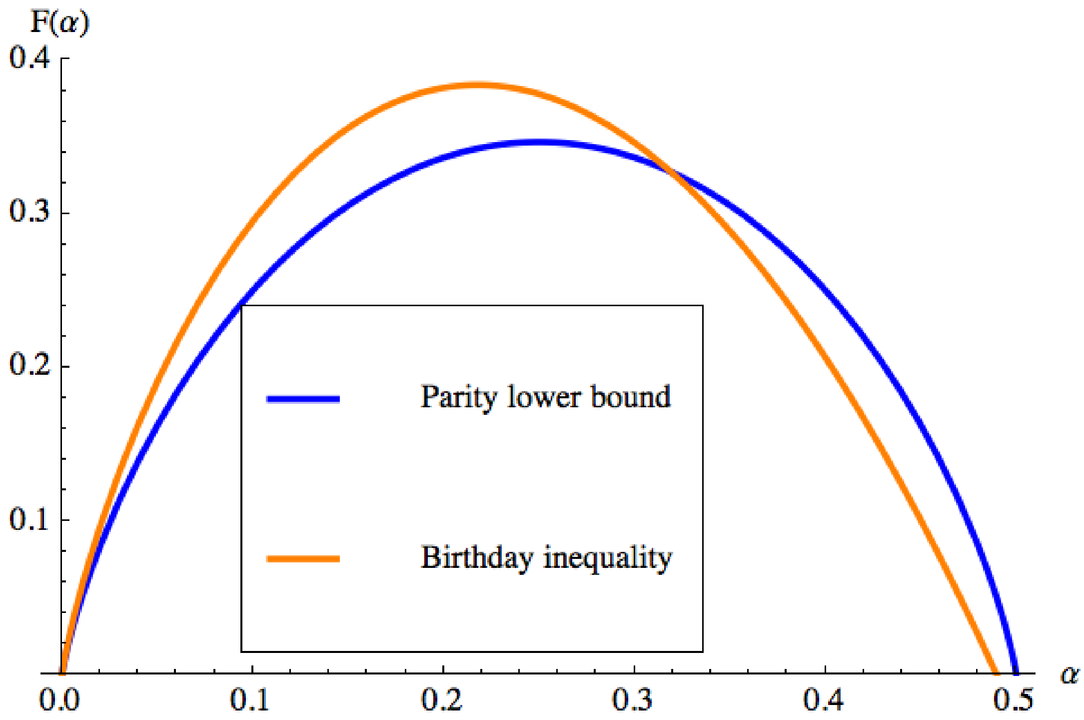

FIGURE 1. Free Energy bounds for the hard core model on 6-regular bipartite graphs

Theorem 6. *For $\alpha \leq \frac{3}{28}$, the repulsion inequality holds for matchings of size $\alpha\frac{n}{2}$ in a d-regular graph on n vertices, and as a consequence*

$$M(\alpha n/2) \leq \frac{(nd/2)^{\alpha n/2}}{(\alpha n/2)!}\left(1 - \frac{2d-1}{nd/2}\right)^{\binom{\alpha n/2}{2}}.$$

On a logarithmic scale, this gives

$$\frac{2}{n}\log M(\alpha n/2) \leq \alpha \log d - \alpha \log \alpha + \alpha - \frac{\alpha^2}{2}\frac{2d-1}{d}. \tag{13}$$

For $\alpha = O(d^{-1/3})$, Theorem 6 improves the bound given by Ilinca and Kahn in [10][2]. Together Theorem 6 and [10] show that the birthday inequality holds for matchings of all sizes in d-regular graphs.

Corollary 4. *The birthday inequality holds for matchings of size k, for all k, in every d-regular graph on n vertices.*

In would be nice to prove that in fact the repulsion inequality holds for matchings of any size in a d-regular graph.

[2]The bound in [10], translated to natural logarithms, is $\frac{2}{n}\log M(\alpha n/2) \leq \alpha\log d - \alpha\log\alpha - 2(1-\alpha)\log(1-\alpha) - \alpha + (\log d)/(d-1)$. Subtracting the first two matching terms then power expanding around $\alpha = 0$ gives $\alpha - \alpha^2 + (\log d)/(d-1) - \alpha^3/3 - \ldots$ for the Ilinca-Kahn bound and $\alpha - \alpha^2 + \alpha^2/(2d)$ for the birthday inequality bound (13). These cross when $\alpha = \Theta(((\log d)/d)^{1/3})$. In particular, for all larger α, the Ilinca-Kahn bound is stronger than the birthday inequality, giving Corollary 4.

Proof of Theorem 6. Let $m = nd/2$ be the number of edges of G, and $p = \frac{2d-1}{m}$. We want to show that $\mathbb{E}[V_k|E_k] \geq \mathbb{E}[V_k]$. Let L_2 be the number of edges of G that are covered by two edges of a matching of size k but are not part of the matching themselves. Then

$$\mathbb{E}[V_k|E_k] = k + (2d-2)k - \mathbb{E}[L_2|E_k] = m\left(pk - \mathbb{E}[L_2/m|E_k]\right)$$

By inclusion/exclusion we have

$$\begin{aligned}\mathbb{E}[V_k] = m(1-(1-p)^k) &\leq mpk - mp^2\binom{k}{2} + mp^3\binom{k}{3} \\ &\leq m\left(pk - \binom{k}{2}p^2\left(1 - \frac{kp}{3}\right)\right)\end{aligned}$$

So it is enough to show that

$$\mathbb{E}[L_2|E_k] \leq m\binom{k}{2}p^2\left(1 - \frac{kp}{3}\right)$$

We can write

$$\mathbb{E}[L_2|E_k] = \sum_{e \in G}\binom{k}{2}\Pr[A_2^{1,2}|E_k]$$

where $A_e^{1,2}$ is the event that edge e is covered by edges 1 and 2. Now it is enough to show that $\Pr[A_2^{1,2}|E_k] \leq p^2(1-kp/3)$. As in the proofs above we write

$$\begin{aligned}\Pr[A_2^{1,2}|E_k] &= \Pr[A_e^{1,2}|E_2] \cdot \frac{\Pr[E_k|A_e^{1,2} \cap E_2]}{\Pr[E_k|E_2]} \\ &\leq \Pr[A_e^{1,2}|E_2] \cdot \frac{1-p}{(1-kp)^2}\end{aligned}$$

We calculate

$$\Pr[A_e^{1,2}|E_2] = \frac{2(d-1)^2}{m^2(1-p)}$$

which gives

$$\Pr[A_2^{1,2}|E_k] \leq \frac{2(d-1)^2}{m^2(1-kp)^2}$$

Our assumption that $\alpha \leq 3/28$ implies that $kp \leq \frac{3}{14}$, and so

$$\Pr[A_2^{1,2}|E_k] \leq \frac{2(d-1)^2}{m^2(1-kp)^2} \leq \frac{(2d-1)^2}{m^2}\left(1 - \frac{kp}{3}\right) = p^2(1-kp/3)$$

which shows that the repulsion inequality holds. $\square$

4. Conclusions and conjectures

We conjecture that the lower bounds on the density at which the birthday inequality holds in Theorem 1 and Theorem 4 can be extended to the entire fluid phase of the hard sphere and hard core models.

We describe two notions of the fluid phase of the hard sphere and fixed-size hard core model. The first is decay of correlations:

Definition 9. *Let x_0, x_0', x_t be positions in $\mathcal{T}^d$ or lattice sites on $\mathbb{Z}_d(n)$. Let A_0 (resp. A_0', A_t) be the event that the position x_0 (x_0', x_t) is covered by a sphere in the hard sphere model or occupied by a particle in the hard core model. Then the model exhibits decay of correlations at density α if there is some constant $c_\alpha > 0$ so that*

$$\left|\Pr[A_t|A_0] - \Pr[A_t|A_0']\right| \leq g(c_\alpha d_t/r)$$

where $d_t = \min\{d(x_0, x_t), d(x_0', x_t)\}$ and $g(s)$ is some function so that $\lim_{s\to\infty} g(s) = 0$. (For the hard core model, we take $r = 1$). The model exhibits exponentially fast decay of correlations if we can take $g(s) \leq e^{-cs}$ for some $c > 0$.

The second notion is the rapid mixing of a specific Markov chain with the hard sphere or hard core distribution as its stationary distribution. One such chain is the single-particle, global-move dynamics (see e.g.[8]). A single move of the Markov chain consists of selecting one center of a sphere or one particle on the lattice uniformly at random, then selecting a position or a site uniformly at random from $\mathcal{T}^d$ or $\mathbb{Z}_d(n)$ and moving the center or the particle to the new position as long as it does not violate the hard constraints of the model. We say the chain mixes rapidly if the mixing time is a polynomial in n.

Conjecture 1. *If the hard sphere or hard core model is in the fluid phase (say it exhibits exponentially fast decay of correlations or the Markov chain above mixes rapidly), then the repulsion inequality (and thus the Birthday Inequality) holds.*

The intuition behind Conjecture 1 is that at a sufficiently low density, conditioning on particles being repulsed from each other should have an essentially local effect, and locally, conditioning on repulsion increases the volume covered by the union of balls around the particles. However, beyond the fluid/solid phase transition, long-range correlations come into play, and conditioning on the repulsion of particles can force them into global lattice-like configurations with holes, and thus the volume covered may actually decrease. Note that a model of random matchings of a given size on the d-dimensional lattice, the monomer-dimer model, does not exhibit a phase transition [9], and Corollary 4 shows the the birthday inequality holds at all densities.

Conjecture 1 has several consequences. First, it would give a mathematical proof that the hard sphere model in dimension 24 undergoes a phase transition, which to the best of our knowledge has not been proved yet in any dimension. The density at which the birthday inequality fails would mark an upper bound on the critical density for the model, as exponential decay of correlations (or fast mixing) could not hold.

Second, it would imply that the critical density in the fixed-size hard core model is upper bounded by α_d^l from Theorem 5, in particular showing that a phase transition occurs at densities $O(1/d)$ in $\mathbb{Z}_d(n)$ or Q_d. The best known analogous bounds in the hard core model with fugacity parameter λ are $\lambda_c = \tilde{O}(d^{-1/3})$ given by Peled and Samotij [15] improving the bound of $\tilde{O}(d^{-1/4})$ from Galvin and Kahn [7]. Proving Conjecture 1 would give the optimal bound up to a constant factor $\lambda_c = O(d^{-1})$ as the typical particle density α in the hard core model with fugacity λ is bounded below by $\frac{\lambda}{6(1+\lambda)}$.

We also make the bold predictions that the critical density on $\mathbb{Z}_d$ satisfies $\alpha_d^* \sim \log 2/d$ as $d \to \infty$, and on Q_d, $\alpha_d^* \sim 2\log 2/d$. For the hard square model in continuous space we conjecture that the phase transition occurs at a critical density $\alpha_d^* \sim \frac{2\log(2)d}{2^d}$.

Acknowledgments

Thanks to Alfredo Hubard for many fruitful discussions on this topic. Thanks to Peter Winkler and Tyler Helmuth for their careful reading of a draft of this paper.

References

[1] K Bezdek and R Connelly. Pushing disks apart. *J. Reine Angew. Math*, 553:221–236, 2003.
[2] Lewis Bowen, Russell Lyons, Charles Radin, and Peter Winkler. Fluid-solid transition, in a hard-core system. *Physical review letters*, 96(2):025701, 2006.
[3] Teena Carroll, David Galvin, and Prasad Tetali. Matchings and independent sets of a fixed size in regular graphs. *Journal of Combinatorial Theory, Series A*, 116(7):1219–1227, 2009.
[4] Henry Cohn and Abhinav Kumar. The densest lattice in twenty-four dimensions. *Electronic Research Announcements of the American Mathematical Society*, 10(7):58–67, 2004.
[5] Luc Devroye, András György, Gábor Lugosi, and Frederic Udina. High-dimensional random geometric graphs and their clique number. *Electronic Journal of Probability*, 16:2481–2508, 2011.
[6] Persi Diaconis, Gilles Lebeau, and Laurent Michel. Geometric analysis for the metropolis algorithm on lipschitz domains. *Inventiones mathematicae*, 185(2):239–281, 2011.
[7] David Galvin and Jeff Kahn. On phase transition in the hard-core model on zˆ d. *Combinatorics, Probability and Computing*, 13(02):137–164, 2004.
[8] Thomas P Hayes and Cristopher Moore. Lower bounds on the critical density in the hard disk model via optimized metrics. *arXiv preprint arXiv:1407.1930*, 2014.
[9] Ole J Heilmann and Elliott H Lieb. Theory of monomer-dimer systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972.
[10] Liviu Ilinca and Jeff Kahn. Asymptotics of the upper matching conjecture. *Journal of Combinatorial Theory, Series A*, 120(5):976–983, 2013.
[11] Ravi Kannan, Michael W Mahoney, and Ravi Montenegro. Rapid mixing of several markov chains for a hard-core model. In *Algorithms and computation*, pages 663–675. Springer, 2003.
[12] Martin Kneser. Einige bemerkungen über das minkowskische flächenmass. *Archiv der Mathematik*, 6(5):382–390, 1955.
[13] John Leech. Notes on sphere packings. *Canad. J. Math*, 19(251):267, 1967.
[14] Hartmut Löwen. Fun with hard spheres. In *Statistical physics and spatial statistics*, pages 295–331. Springer, 2000.
[15] Ron Peled and Wojciech Samotij. Odd cutsets and the hard-core model on zˆ d. *Annales de l'Institut Henri Poincare, Probabilites et Statistiques*, 50:975–998, 2014.
[16] E Thue Poulsen. Problem 10. *Mathematica Scandinavica*, 2:346–346, 1954.
[17] Charles Radin. Low temperature and the origin of crystalline symmetry. *International Journal of Modern Physics B*, 1(05n06):1157–1191, 1987.
[18] Dana Randall and Peter Winkler. Mixing points on a circle. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 426–435. Springer, 2005.
[19] Lewi Tonks. The complete equation of state of one, two and three-dimensional gases of hard elastic spheres. *Physical Review*, 50(10):955, 1936.

School of Mathematics, University of Birmingham, UK, and the IMA, Minneapolis, MN, USA.
E-mail: w.f.perkins@bham.ac.uk.