On phase transition in the hard-core model on $\mathbf{Z}^d$

David Galvin
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
and
Jeff Kahn*
Department of Mathematics and RUTCOR
Rutgers University
New Brunswick, NJ 08903

Abstract

It is shown that the hard-core model on $\mathbf{Z}^d$ exhibits a phase transition at activities above some function $\lambda(d)$ which tends to zero as $d \to \infty$; that is:

Consider the usual nearest neighbor graph on $\mathbf{Z}^d$, and write $\mathcal{E}$ and $\mathcal{O}$ for the sets of even and odd vertices (defined in the obvious way). Set

$$\Lambda_M = \Lambda_M^d = \{z \in \mathbf{Z}^d : \|z\|_\infty \le M\}, \quad \partial^\star \Lambda_M = \{z \in \mathbf{Z}^d : \|z\|_\infty = M\},$$

and write $\mathcal{I}(\Lambda_M)$ for the collection of independent sets (sets of vertices spanning no edges) in Λ_M. For $\lambda > 0$ let $\mathbf{I}$ be chosen from $\mathcal{I}(\Lambda_M)$ with $\Pr(\mathbf{I} = I) \propto \lambda^{|I|}$.

Theorem There is a constant C such that if $\lambda > Cd^{-1/4}\log^{3/4} d$, then

$$\lim_{M\to\infty} \Pr(\underline{0} \in \mathbf{I} | \mathbf{I} \supseteq \partial^\star\Lambda_M \cap \mathcal{E}) \ > \lim_{M\to\infty} \Pr(\underline{0} \in \mathbf{I} | \mathbf{I} \supseteq \partial^\star\Lambda_M \cap \mathcal{O}).$$

Thus, roughly speaking, the influence of the boundary on behavior at the origin persists as the boundary recedes.

1991 Mathematics Subject Classification. Primary: 82B20, 82B26. Secondary: 05A16, 05C70.

Key words and phrases: hard-core model, phase transition, Peierls argument

*Research supported in part by NSF grant DMS-9970433.

1 Introduction

The "hard-core model" is a simple mathematical model of a gas with particles of non-negligible size. The vertices ("sites") of a graph are regarded as positions, each of which can be occupied by a particle, subject to the rule that two neighboring sites cannot both be occupied (particles cannot overlap).

We need a few definitions, but aim to be brief. For good introductions to the hard-core model see [1], [10]. See also [8] for more general background, and e.g. [2] or [5] for graph theory basics. A few conventions are mentioned at the end of this section.

Write $\mathcal{I}(\Sigma)$ for the collection of independent sets (sets of vertices spanning no edges) of graph Σ.

For Σ finite and $\lambda > 0$, the *hard-core measure* with *activity* (or *fugacity*) λ on $\mathcal{I} = \mathcal{I}(\Sigma)$ (or "on Σ") is given by

$$\mu(I) = \lambda^{|I|}/Z \qquad \text{for} \quad I \in \mathcal{I},$$

where Z is the appropriate normalizing constant (*partition function*), $Z = \sum\{\lambda^{|I'|} : I' \in \mathcal{I}\}$. (The more usual etiquette here considers probability measures on $\{0,1\}^{V(\Sigma)}$ supported on indicators of independent sets; but the present usage is convenient for us, and we adhere to it throughout.)

In particular $\lambda = 1$ gives uniform distribution. One may also assign different activities λ_v to the different vertices v and take $\mu(I)$ proportional to $\prod_{v\in I}\lambda_v$, but we will not do so here; again see [1], [10], and also e.g. [14], [11], [13] for some combinatorial applications.

For infinite Σ a measure μ on $\mathcal{I}(\Sigma)$ is hard-core with activity λ if, for $\mathbf{I}$ chosen according to μ and for each finite $W \subset V = V(\Sigma)$, the conditional distribution of $\mathbf{I} \cap W$ given $\mathbf{I} \cap (V \setminus W)$ is μ-a.s. the hard-core measure with activity λ on the independent sets of $\{w \in W : w \not\sim \mathbf{I} \cap (V \setminus W)\}$ (the vertices that can still be in $\mathbf{I}$ given $\mathbf{I} \cap (V \setminus W)$). General considerations (see [8]) imply that there is always at least one such μ; if there is more than one, the model is said to have a *phase transition*.

The canonical (and by far most studied) case of the hard-core model is that of (the usual nearest neighbor graph on) $\mathbf{Z}^d$. Here the seminal result is due to Dobrushin [6], who proved that there is a phase transition for sufficiently large λ, depending on d. (Dobrushin's result was rediscovered by Louth [18] in the context of communications networks.)

The λ required in [6] is larger than one would expect,[†] and attempted improvements have been the subject of considerable effort—if not publication—in both the statistical mechanics and discrete mathematics communities in recent years.

Even the fact that the required λ increases with d is a little strange, since one expects that as d grows phase transition should get "easier," in the sense that for a given λ, phase transition in dimension d should imply phase transition in all higher dimensions; but this remains open.

Also open is the existence of a "critical" activity, $\lambda_c(d)$, such that one has phase transition for $\lambda > \lambda_c(d)$ but not for $\lambda < \lambda_c(d)$. While this seems certain to be true for $\mathbf{Z}^d$, a cautionary note is sounded in [4], where it is shown that there are graphs (even trees) for which there is no such critical activity.

As a temporary substitute we may define $\lambda(d)$ to be the supremum of those λ for which the hard-core model with activity λ on $\mathbf{Z}^d$ does not have a phase transition.

So Dobrushin at least tells us that $\lambda(d) < \infty$, while "easier as dimension grows" would imply $\lambda(d) < O(1)$. A particular question that has received much of the attention devoted to this problem is whether $\lambda(d) \leq 1$ for large d. But in fact it has been generally believed (despite some early guesses to the contrary) that $\lambda(d)$ tends to zero as d grows; this is what we prove:

Theorem 1.1 *$\lambda(d) = O(d^{-1/4} \log^{3/4} d)$.*

The bound here is undoubtedly not best possible; $O(\log d/d)$ and $O(1/d)$ are natural guesses at the true value of $\lambda(d)$.

We assume henceforth that d is large enough to support our various assertions.

The problem of showing existence of a phase transition may be finitized as follows. Let $\Lambda = \Lambda_M = \mathbf{Z}^d \cap [-M, M]^d = \mathcal{O} \cup \mathcal{E}$ with $\mathcal{O}$ and $\mathcal{E}$ the sets of odd and even vertices (defined in the natural way: $x \in \mathbf{Z}^d$ is odd if $\sum x_i$ is odd); let μ_M be the hard-core measure with activity λ on Λ (meaning, of course, on the subgraph of $\mathbf{Z}^d$ induced by Λ); and (with $\mathbf{I}$ chosen according

[†]No explicit bound is given in [6], but several colleagues report that Dobrushin's argument works for $\lambda > C^d$ for a suitable constant C.

to μ_M) let μ_M^e be μ_M conditioned on the event $\{\mathbf{I} \supseteq \partial^\star\Lambda \cap \mathcal{E}\}$, where $\partial^\star\Lambda := [-M, M]^d \setminus [-(M-1), M-1]^d$, and define μ_M^o similarly.

In [1] it is shown (*inter alia*) that the sequences $\{\mu_M^e\}$ and $\{\mu_M^o\}$ converge to weak limits, called μ^e and μ^o, and that there is a phase transition iff these limits are different. (This is mainly based on the FKG Inequality, and applies to general bipartite graphs Σ, provided we allow $\{\Lambda_M\}$ to be an arbitrary nested sequence with $\cup\Lambda_M = V(\Sigma)$.)

Thus it is natural to try to prove phase transition by exhibiting some statistic distinguishing μ^e from μ^o. We will show $\mu^e(\underline{0} \in \mathbf{I}) \neq \mu^o(\underline{0} \in \mathbf{I})$, i.e.

$$\lim_{M\to\infty} \mu_M^e(\underline{0} \in \mathbf{I}) \neq \lim_{M\to\infty} \mu_M^o(\underline{0} \in \mathbf{I}). \tag{1}$$

(Of course we are only using the trivial direction of "phase transition iff $\mu^e \neq \mu^o$." It is not hard to show that (1), too, is equivalent to phase transition.)

To establish (1) (assuming at least $\lambda = \Omega(1/d)$, which is easily seen to be necessary for phase transition) it is in turn enough to show that for $v_0 \in \Lambda$,

$$\begin{array}{lcll} \mu_M^e(v_0) & < & o(1/d) & \text{if } v_0 \text{ is odd,} \\ \mu_M^o(v_0) & < & o(1/d) & \text{if } v_0 \text{ is even.} \end{array}$$

For then (writing N for neighborhood)

$$\begin{aligned} \mu_M^e(\underline{0} \in \mathbf{I}) &= \mu_M^e(N(\underline{0}) \cap \mathbf{I} = \emptyset)\mu_M^e(\underline{0} \in \mathbf{I} | N(\underline{0}) \cap \mathbf{I} = \emptyset) \\ &= (1 - o(1))\lambda/(1+\lambda), \end{aligned}$$

so that $\mu^e(\underline{0} \in \mathbf{I}) = (1 - o(1))\lambda/(1+\lambda)$, whereas $\mu^o(\underline{0} \in \mathbf{I}) = o(1/d)$.

So in particular the next theorem, whose proof is the main business of this paper, contains Theorem 1.1.

Theorem 1.2 *For*

$$\lambda = \omega(d^{-1/4} \log^{3/4} d), \tag{2}$$

M arbitrary, and v_0 an odd vertex of Λ_M,

$$\mu_M^e(v_0 \in \mathbf{I}) < (1+\lambda)^{-(2-o(1))d}. \tag{3}$$

The same result holds if we reverse the roles of even and odd.

Remark. It is easy to see that

$$\begin{aligned}\mu_M^e(v_0 \in \mathbf{I}) &= \mu_M^e(N(v_0) \cap \mathbf{I} = \emptyset)\mu_M^e(v_0 \in \mathbf{I} | N(v_0) \cap \mathbf{I} = \emptyset) \\ &> (1+\lambda)^{-2d}\frac{\lambda}{1+\lambda},\end{aligned}$$

so that (3) actually gives the asymptotics of $\log \mu_M^e(v_0 \in \mathbf{I})$.

Set

$$\mathcal{J} = \{I \in \mathcal{I}(\Lambda) : \partial^\star \Lambda \cap \mathcal{E} \subseteq I\}.$$

The proof of Theorem 1.2 is a sort of "Peierls argument" (see e.g. [9]): we try to associate with each $I \in \mathcal{J}$ containing v_0 a "contour"—some kind of membrane separating the outer even region from an inner odd region containing v_0—and then use this to map I to a large set of J's, also from $\mathcal{J}$ but not containing v_0, each obtained from I by some modification of the inner region.

This is no surprise: almost every attempt at settling this problem that we're aware of has attacked it more or less along these lines. (The one exception is the entropy approach of [12], which for now seems unlikely to get us to anything like what's proved here.)

The main difficulty in all these attempts has been getting some kind of control over the set of possible "contours." Much of the inspiration for our approach to this problem was provided by the beautiful ideas of A. Sapozhenko [20], which he used to give, for example, relatively simple derivations of Korshunov's [16] description of the asymptotics for Dedekind's Problem (in [22]), and, in [21], of the asymptotics for the number of independent sets ("codes of distance 2") in the Hamming cube $\{0,1\}^n$ originally established in [17].

Some of our tools also come from [20]: Lemma 2.17 is an improved version of one of Sapozhenko's arguments, and our uses of Lemmas 2.1-2.3 are similar to his.

The rest of the paper is devoted to the proof of Theorem 1.2. Unfortunately, saying anything even mildly intelligible about the argument turns out to be awkward without some preliminaries, so we will wait: see the end of Section 2.2 and most of Section 2.6. (Section 2.2 reformulates slightly and says what we will actually prove.)

Usage

We use "bigraph" for "bipartite graph."

For a graph on vertex set V, we use $\nabla(W)$ for the set of edges having exactly one end in $W \subseteq V$ and $\nabla(U, W)$ for the set of edges having one end in U and the other in W.

The neighborhood of (i.e. set of vertices adjacent to) v is $N(v)$; $N(W) = \cup\{N(v) : v \in W\}$; and $\partial W = N(W) \setminus W$. We use $d(\cdot)$ for degree—$d(v) = |N(v)|$ and $d_W(v) = |N(v) \cap W|$—and $\mathrm{dist}(\cdot, \cdot)$ for distance.

One common abuse: we often fail to distinguish between a graph and its set of vertices, so for instance might use "component" where we should really say "set of vertices of a component."

When the difference makes no difference, we pretend that all large numbers are integers. All constants implied by the notations $O(\cdot)$, $\Omega(\cdot)$ are absolute; that is, they do not depend on d.

2 Proof of Theorem 1.2

2.1 Preliminaries

Here we collect what we will need in the way of known results.

Lemma 2.1 *In any graph with all degrees at most D, the number of connected, induced subgraphs of order n containing a fixed vertex x_0 is at most $(eD)^n$.*

This follows from the well-known fact (e.g. [15, p.396, Ex.11]) that the infinite D-branching rooted tree contains precisely $\frac{1}{(D-1)n+1}\binom{Dn}{n}$ rooted subtrees of size n.

The next lemma is a special case of a fundamental result due to Lovász [19] and Stein [23] (see also [7]). For a bigraph Σ with bipartition $X \cup Y$, say $Y' \subseteq Y$ *covers* X if each $x \in X$ has a neighbor in Y'.

Lemma 2.2 *If Σ as above satisfies $d(x) \geq a$ $\forall x \in X$ and $d(y) \leq b$ $\forall y \in Y$, then X is covered by some $Y' \subseteq Y$ of size at most $(|Y|/a)(1 + \ln b)$.*

Call a set T of vertices of a graph *c-clustered* if for any $x, y \in T$ there are vertices $x = x_0, x_1, \ldots, x_k = y$ with $\mathrm{dist}(x_{i-1}, x_i) \leq c$ for all i. The next lemma is from [20] (see Lemma 2.1); the interested reader should have no difficulty supplying a proof.

Lemma 2.3 *If Σ is a graph on V and $S, T \subseteq V$ satisfy*

(i) *S is a-clustered,*

(ii) $\mathrm{dist}(x, T) \leq b \ \forall x \in S$ *and* $\mathrm{dist}(y, S) \leq b \ \forall y \in T$,

then T is $(a + 2b)$-clustered.

Finally, we need to know something about isoperimetry in $\mathbf{Z}^d$. Write $|x|$ for the ℓ_1-norm of x, and set $B(r) = \{x \in \mathbf{Z}^d : |x| \leq r\}$, $S(r) = \{x \in \mathbf{Z}^d : |x| = r\}$, $b(r) = |B(r)|$ and $s(r) = |S(r)|$.

Lemma 2.4 *Let C be a subset of $\mathbf{Z}^d$ with*

$$|C| = b(r) + \alpha s(r+1),$$

where $0 \leq \alpha < 1$. Then

$$|\partial C| \geq (1-\alpha)s(r+1) + \alpha s(r+2).$$

This is an immediate consequence of a corresponding inequality for the torus $(\mathbf{Z}/k\mathbf{Z})^d$, given by Bollobás and Leader in [3, Cor. 5]. The case $\alpha = 0$ was proved by Wang and Wang [24].

2.2 To prove

We assume henceforth that λ satisfies (2). We prove only the first part of Theorem 1.2 ((3) for odd v_0); switching "even" and "odd" throughout the argument gives the proof of the second part.

It will be convenient to replace the box Λ_M by the discrete torus $\Gamma = \Gamma_M$ obtained from Λ_M by setting $M = -M$ and identifying vertices accordingly. Following our favorite abuse, we regard Γ as either a graph or a set of vertices as convenient.

We then use Δ for the image of $\partial^\star \Lambda_M$ under the natural projection $\Lambda_M \mapsto \Gamma$, and continue to write $\underline{0}$ for the image of $\underline{0}$ in Γ, and to use $\mathcal{O}$ and $\mathcal{E}$ for the sets of odd and even vertices of Γ.

Having done this, we replace $\partial^\star\Lambda_M$ by Δ in the definition of $\mathcal{J}$ ($\mathcal{J} = \{I \subseteq \Gamma : I \text{ independent}, \Delta \cap \mathcal{E} \subseteq I\}$), define μ_M^e, μ_M^o as before, and simply regard Theorem 1.2 as referring to Γ, a change which clearly does not affect its meaning.

We will show a bit more than (3): for $I \in \mathcal{J}$, let $Z = Z(I)$ be the component of $\Gamma - (I \cap \mathcal{O})$ containing Δ; then

$$\mu_M^e(v_0 \notin Z(\mathbf{I})) < (1+\lambda)^{-(2-o(1))d} \tag{4}$$

Let $\mathcal{J}_0 = \{I \in \mathcal{J} : v_0 \notin Z(I)\}$, and write $w(I)$ for $\lambda^{|I|}$. We prove (4) by producing a "flow" $\nu : \mathcal{J}_0 \times \mathcal{J} \to [0,1]$ satisfying

$$\sum_J \nu(I,J) = 1 \quad \forall I \in \mathcal{J}_0 \tag{5}$$

and

$$\sum_I \frac{w(I)}{w(J)} \nu(I,J) < (1+\lambda)^{-(2-o(1))d} \quad \forall J \in \mathcal{J}. \tag{6}$$

This gives (4):

$$\begin{aligned}\sum_{I \in \mathcal{J}_0} w(I) &= \sum_{I \in \mathcal{J}_0} w(I) \sum_{J \in \mathcal{J}} \nu(I,J) \\ &= \sum_{J \in \mathcal{J}} w(J) \sum_{I \in \mathcal{J}_0} \frac{w(I)}{w(J)} \nu(I,J) \\ &< (1+\lambda)^{-(2-o(1))d} \sum_{J \in \mathcal{J}} w(J).\end{aligned}$$

Throughout our discussion we fix v_0 and use I for members of $\mathcal{J}_0$ and J for general members of $\mathcal{J}$.

The definition of $\nu(I,\cdot)$ will depend on a pair $(G,A) = (G(I), A(I)) \in 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ associated with I. The construction and salient properties of the pair are given in Sections 2.3 and 2.4, but it will not be until Section 2.11 that we are able to specify ν. First steps toward this specification are taken in Section 2.5, which finally puts us in a position—in Section 2.6—to give some clue as to how the main part of the argument will proceed.

2.3 "Contours"

For a set P of vertices (in any graph) we use $\partial^\star P$ for the *internal boundary* of P:

$$\partial^\star P = \{v \in P | N(v) \not\subseteq P\}.$$

The following observation is used several times, so we record it as a lemma; its easy proof is left to the reader.

Lemma 2.5 *Let Σ be a graph, $S \subseteq V(\Sigma)$, and T (the vertex set of) some component of $\Sigma - (S \setminus \partial^\star S)$. Then $\partial^\star T \subseteq \partial^\star S$.*

Let $I \in \mathcal{J}_0$, $Z = Z(I)$ be as in Section 2.2, and set $Z_0 = \partial^\star Z$. By the definition of Z, it is clear that $Z_0 \subset \mathcal{E}$ and $Z_0 \cap I = \emptyset$. Let W' be the component of v_0 in the graph $\Gamma - (Z \setminus Z_0)$. By Lemma 2.5, $\partial^\star W' \subseteq W' \cap Z_0 \subseteq \mathcal{E}$.

Let $W'' = W' \cup \{x \in \mathcal{O} | N(x) \subseteq W'\}$. This is clearly connected, with $\partial^\star W'' \subseteq \partial^\star W'$.

Now consider $\Gamma - (W'' \setminus \partial^\star W'')$. This breaks into a number of components, one of which, C say, contains Δ. Again using Lemma 2.5, we have $\partial^\star C \subseteq C \cap \partial^\star W''$. Finally, set $W = \Gamma \setminus (C \setminus \partial^\star C)$, $G = W \cap \mathcal{E}$, $A = W \cap \mathcal{O}$, and $G_0 = \partial^\star W$.

The next proposition collects relevant properties of these objects. Once we have these properties, we will not be concerned with how G, A etc. were derived from I.

Proposition 2.6

$$v_0 \in A; \quad W \cap \Delta = \emptyset; \tag{7}$$

$$\textit{both } C \textit{ and } W \textit{ are connected;} \tag{8}$$

$$G_0 = \partial^\star C; \tag{9}$$

$$G = N(A) \quad \textit{and} \quad A = \{x \in \mathcal{O} | N(x) \subseteq G\}; \tag{10}$$

$$G_0 \cap I = \emptyset; \tag{11}$$

$$N(G_0) \cap I \subset A; \tag{12}$$

$$G_0 \subseteq N(A \cap I). \tag{13}$$

Proof. Both (7) and the connectivity of C are immediate. To see that W is connected, notice that each component of $\Gamma - (W'' \setminus \partial^\star W'')$ must meet $\partial^\star W''$ (or it would be a component of the connected graph Γ). Thus W is the union of the connected set W'' and a number of other connected sets each of which meets W'', so is itself connected. So we have (8).

For (9): $\partial^\star C \subseteq W \cap \mathcal{E}$ and the connectivity of C give

$$x \in \partial^\star C \Rightarrow \emptyset \neq N(x) \cap C \subseteq C \cap \mathcal{O} \subseteq C \setminus W \Rightarrow x \in \partial^\star W,$$

so $\partial^\star C \subseteq \partial^\star W$; and Lemma 2.5 and the connectivity of W give the reverse containment.

Connectivity of W and the fact that $G_0 \subseteq \mathcal{E}$ give $G = N(A)$. That $A \subseteq \{x \in \mathcal{O} | N(x) \subseteq G\}$ follows from $G = N(A)$ (or just $\partial^\star W \subseteq \mathcal{E}$). For the reverse containment, notice that $x \notin W \Rightarrow N(x) \cap W \subseteq G_0 \subseteq W'$, whereas $N(x) \subseteq W'$ would imply $x \in W'' \subseteq W$; so $x \notin W \Rightarrow N(x) \not\subseteq W$.

For (11) recall that $G_0 = \partial^\star C \subseteq \partial^\star W'' \subseteq \partial^\star W' \subseteq Z_0$ and $Z_0 \cap I = \emptyset$.

That $N(G_0) \cap I \subseteq A$ follows from $G_0 \subseteq \partial^\star W'$, since $N(\partial^\star W') \cap I$ is clearly contained in A.

Finally, $v \in G_0 \Rightarrow v \in Z_0 \Rightarrow v \sim I$, so (13) follows from (12).

∎

2.4 Topology

The purpose of this section is to prove, for any $I \in \mathcal{J}_0$ and W, G etc. produced from I as in Section 2.3,

$$G_0 \text{ is 2-clustered} \tag{14}$$

Our proof of this, which is considerably longer than we would wish and unrelated to the methods in the rest of the paper, might profitably be skipped on a first reading.

Though (14) turns out to follow from the connectivity of W and C (see (8)), we could not see a simple combinatorial proof of the implication, and our argument requires a little topological detour, based on

Lemma 2.7 *If U, V are connected subsets of $X = \mathbf{R}^n$ or S^n, $n > 1$, with $U \cup V = X$, U closed and V compact, then $U \cap V$ is connected.*

(As usual, S^n is the unit sphere $\{x \in \mathbf{R}^{n+1} : \sum x_i^2 = 1\}$. We also write B^{n+1} for the corresponding unit ball.)

The (presumably well-known) proof of Lemma 2.7 is given at the end of this section.

It will be convenient here to write Ω for the nearest neighbor graph on $\mathbf{Z}^d$. As usual, $\Omega[S]$ is the subgraph induced by S. We will prove (14) in the following more general form.

Proposition 2.8 *Let $R \cup B$ be a decomposition of $V(\Omega)$ ($= \mathbf{Z}^d$), with both $\Omega[R]$ and $\Omega[B]$ connected and R finite. Suppose $G := R \cap B$ is contained in $\mathcal{E}$ and is the internal boundary of each of R, B. Then G is 2-clustered.*

Remark. We will actually show that G is 2-clustered in each of R and B.

Proof With Ω embedded in $\mathbf{R}^d$ in the natural way, we extend R and B to closed connected subsets R^* and B^* of $\mathbf{R}^d$ so that $R^* \cup B^* = \mathbf{R}^d$ and $G^* := R^* \cap B^*$ is path-connected. We then derive the 2-clusteredness of G from the path-connectedness of G^*.

We view $\mathbf{R}^d$ as the union of $\mathbf{Z}^d$-translates of $[0,1]^d$ (the *cells* of $\mathbf{R}^d$), and define R^* and B^* cell by cell. Within a cell we proceed by dimension, first defining the extensions for 0-dimensional faces (the vertices of Ω), 1-dimensional faces (the edges of Ω), and 2-dimensional faces, and then continuing inductively. (As usual a *face* of a cell is the intersection of the cell with some supporting hyperplane. Henceforth we use "k-face" for "k-dimensional face.") For the inductive step, we need a topological lemma (Lemma 2.11), for the statement of which it's convenient to introduce two local definitions. Let us say that a subset of a topological space is *civilized* if it is closed, has only finitely many components, and each of its components is path-connected.

Definition 2.9 *A decomposition $X = R \cup B$ of a topological space X, with $R \cap B = G$, is* nice *if it satisfies:*

(i) $G = \partial R = \partial B$;

(ii) *each of R, B, G is civilized; and*

(iii) *each of R, B—and so each component of R and B—is the closure of the union of finitely many open, path-connected sets.*

If $X = R \cup B$ is a nice decomposition, and R', B' are obtained from R, B by adding finitely many points, then we also call the decomposition $X = R' \cup B'$ nice.

(Of course there is some redundancy in conditions (i)-(iii).)

We say that two nice decompositions $X_1 = R_1 \cup B_1$ and $X_2 = R_2 \cup B_2$ are *compatible* if $R_1 \cap X_1 \cap X_2 = R_2 \cap X_1 \cap X_2$ and $B_1 \cap X_1 \cap X_2 = B_2 \cap X_1 \cap X_2$. It's straightforward to check that nice decompositions of different spaces can be combined if they are compatible:

Lemma 2.10 *Suppose $X = X_1 \cup \cdots \cup X_m$ with each X_i closed. If $X_i = R_i \cup B_i$ are pairwise compatible, nice decompositions, then $(\cup R_i) \cup (\cup B_i)$ is a nice decomposition of X.*

We now state the topological lemma alluded to above, deferring its proof until after the derivation of Proposition 2.8. (Recall B^{n+1} and S^n are the unit ball and sphere in $\mathbf{R}^{n+1}$.)

Lemma 2.11 *Assume $n > 1$. If $R \cup B$ is a nice decomposition of S^n, then there is a nice decomposition $R^* \cup B^*$ of B^{n+1}, with $R^* \cap S^n = R$, $B^* \cap S^n = B$, and such that if C is any component of R^* (resp. B^*, G^*), then $C \cap S^n$ is a component of R (resp. B, G).*

(This is easily seen to fail for $n = 1$. It may be worth pointing out that for R and B, condition (iii) of Definition 2.9 refers to sets that are open *in* S^n; similarly ∂R and ∂B are boundaries relative to S^n, while ∂R^* and ∂B^* are boundaries relative to B^{n+1}.)

Of course Lemma 2.11 still applies if we replace the B^{n+1} by any of its homeomorphic images (and S^n by the corresponding homeomorphic copy); in our case the relevant image will be $[0,1]^d$.

We now fix a cell, and begin defining our extensions. For vertices and edges we do the natural things: $R^* \cap V(\Omega) = R$, $B^* \cap V(\Omega) = B$; and we put (the interior of) an edge in R^* (resp. B^*) iff both its ends are in R^* (resp. B^*), noting that exactly one of these possibilities occurs, since $\nabla(G, G) = \emptyset$.

Next, we deal with 2-dimensional faces. If the vertices of such a face are all in R (resp. B), then put the interior of the face in R^* (resp. B^*). Otherwise, the face has two opposite corner vertices (v_1, v_3, say) in G, with one of its remaining two vertices (v_2) in $R \setminus B$ and the other (v_4) in $B \setminus R$. Put the interior of the convex hull of v_1, v_2, v_3 in R^*, the interior of the convex hull of v_1, v_3, v_4 in B^*, and the interior of the diagonal joining v_1 and v_3 in $R^* \cap B^*$. It is easy to check that these (R^*, B^*)-decompositions of the 2-dimensional faces are nice. (It may be worth observing that a 2-dimensional face contained in R^* may still have one or two of its vertices in B^*, and vice versa.)

We now proceed by induction, assuming the decomposition has been defined on faces of dimension less than $k \in \{3, \ldots, d\}$. Each k-face F is homeomorphic to B^k, and is bounded by the union of finitely many $(k-1)$-dimensional faces. The decomposition of each of these bounding faces is nice, and the decompositions on any two faces are compatible (since we are defining the decomposition from lower dimensions up). So, by Lemma 2.10, we have a nice decomposition of the boundary of F. We now apply Lemma 2.11 to extend to a nice decomposition of the entire face. Once we have a nice decomposition of each cell, we get the full decomposition $\mathbf{R}^d = R^* \cup B^*$ by combining the decompositions of the cells, again appealing to Lemma 2.10 for "nice." (For formal applicability of the lemma, we can use a single $X_i = B_i$ for the union of all cells not meeting R.)

It is clear from the construction that R^* and B^* are closed, R^* is bounded, and $R^* \cup B^* = \mathbf{R}^d$. To see that R^* is connected, notice that by construction, any component of R^* contains an edge of $\Omega[R]$, and that every edge of $\Omega[R]$ is contained in a component of R^*; connectivity of R^* then follows from connectivity of $\Omega[R]$. The same argument shows that B^* is connected.

Lemma 2.7 now shows that G^* is connected, which, since G^* is also civilized (since $R^* \cup B^*$ is nice), implies that it is actually path-connected.

It remains to show that path-connectedness of G^* implies 2-clusteredness of G. It is enough to show that for each pair of vertices $u, v \in G$, there is a path connecting them in G^* which is supported entirely on the 2-dimensional faces of $\mathbf{R}^d$; for, by the construction of R^* and B^*, such a path is supported on diagonals (of 2-dimensional faces) connecting pairs of vertices from G, and such diagonals correspond to steps of length 2 in Ω. (This also justifies the remark following Proposition 2.8.)

So, consider a (u, v)-path P in G^* given by the continuous function $f : [0, 1] \to \mathbf{R}^d$. If P is supported on 2-dimensional faces of $\mathbf{R}^d$, then we are done. Otherwise, let $k > 2$ be the maximum dimension of a face whose interior meets P. It's enough to show that we can replace P by a path meeting the interiors of fewer k-faces than P and no faces of dimension more than k.

To do this, choose a k-face F and component C of $G^* \cap F$ with $C \cap F^0 \cap P \neq \emptyset$ (where F^0 is the interior of F). Let $p = \inf\{x \in [0, 1] : f(x) \in C \cap F^0\}$ and $q = \sup\{x \in [0, 1] : f(x) \in C \cap F^0\}$. Then $f(p), f(q) \in C \cap \partial F$, which, by construction, is path-connected. So we may replace $f([p, q])$ in P by a path contained in ∂F.

∎

Proof of Lemma 2.11

To avoid confusion, we now write ∂X, $\partial' X$ and $\partial'' X$ for the boundaries of X relative to, respectively, $\mathbf{R}^{n+1}$, B^{n+1} and S^n.

We may assume neither R nor B contains isolated points: otherwise we can simply delete such points, produce R^* and B^* for the resulting "reduced" R and B, and then add the deleted points of R (B) to R^* (B^*).

We use (R, B)-*component* to mean a component of either R or B, and proceed by induction on the number of (R,B)-components in the decomposition of S^n.

If there is exactly one such component (a component of R, say), then $R = S^n$, and $B = \emptyset$. Setting $R^* = B^{n+1}$ and $B^* = \emptyset$, we get a nice decomposition of B^{n+1} which satisfies the conditions of the lemma.

Otherwise, there must be at least one (R,B)-component T for which $S^n \setminus T^0$ is connected. For suppose $S^n \setminus T^0$ is disconnected for every (R,B)-component T. Choose an (R,B)-component T_0 ($\subseteq R$, say) such that one of the components of $S^n \setminus T_0^0$, C say, contains as few (R, B)-components as possible, and let T_1 be an (R, B)-component of C (i.e. contained in C, noting that each (R, B)-component other than T_0 is either contained in or disjoint from C). Now $S^n \setminus C^0$ is connected in $S^n \setminus T_1^0$, so $S^n \setminus T_1^0$ (which by assumption is not connected) contains a component whose (R,B)-components form a proper subset of the (R,B)-components of C, contradicting the choice of T_0.

Let T, then, be an (R,B)-component with $S^n \setminus T^0$ connected. We may assume that T is a component of R. Applying Lemma 2.7 with $X = S^n$,

$U = T$ and $V = S^n \setminus T^0$, we find that $\partial'' T$ is connected, so that T meets exactly one component, say C, of B (and $C \supseteq \partial'' T$).

Set $T^* = \{\lambda x : x \in T, \lambda \in [1/2, 1]\}$. This will be one component of R^*. It is easy to see that T^* is closed and path-connected (so civilized), as is $\partial' T^*$, and that $T^* \cap S^n = T$, a component of R.

Now let $(T^*)^0$ be the relative interior of T^* with respect to B^{n+1} (namely, $(T^*)^0 = \{\lambda x : x \in T^0, \lambda \in (1/2, 1]\}$), $P = \partial(B^{n+1} \backslash (T^*)^0)$ $(= (S^n \backslash T^0) \cup \partial' T^*)$, and $Q = B^{n+1} \setminus (T^*)^0$. Then (Q, P) is (easily seen to be) homeomorphic to (B^{n+1}, S^n).

Let, further, $R_1 = R \setminus T$, $B_1 = B \cup \partial' T^*$, and $C_1 = C \cup \partial' T^*$. Then

(i) the components of R_1 are precisely the components of R other than T,

(ii) the components of B_1 are C_1 and the components of B other than C,

and it is easy (if tedious) to deduce that $R_1 \cup B_1$ is a nice decomposition of P.

Our inductive hypothesis thus gives a nice decomposition $R_1^* \cup B_1^*$ of Q, and we obtain the desired decomposition, $R^* \cup B^*$, of B^{n+1} by setting $B^* = B_1$ and $R^* = R_1 \cup T^*$ (again an easy verification using (i) and (ii)).

∎

Proof of Lemma 2.7

We first establish a corresponding statement for open sets: *if U, V are connected, open subsets of $X = \mathbf{R}^n$ or S^n, $n > 1$, with $U \cup V = X$, then $U \cap V$ is connected.*

Proof. We use the Mayer-Vietoris sequence. If X is a topological space, and U and V are open subsets of X whose union is X, then this is a long exact sequence of group homomorphisms ending with

$$\cdots \to H_1(X) \to H_0(U \cap V) \to H_0(U) \oplus H_0(V) \to H_0(X) \to 0,$$

where H_m is the m^{th} homology group. We apply this with $X = \mathbf{R}^n$ or S^n. Using the facts that $H_m(\mathbf{R}^n) = 0$ whenever $m \geq 1$ and that if O is an open subset of $\mathbf{R}^n$ or S^n, then $H_0(O) \cong \mathbf{Z}$ iff O is connected, this long exact sequence becomes

$$0 \to H_0(U \cap V) \to \mathbf{Z} \oplus \mathbf{Z} \to \mathbf{Z} \to 0.$$

From the exactness of this sequence, it follows that $H_0(U \cap V) \cong \mathbf{Z}$, so that $U \cap V$ is connected.

∎

Now let U, V be as in the lemma, and for each $\varepsilon > 0$, set $U_\varepsilon = \{x \in X : d(x, U) < \varepsilon\}$ and $V_\varepsilon = \{x \in X : d(x, V) < \varepsilon\}$. These are open, connected sets whose union is X, so by the preceding result, $U_\varepsilon \cap V_\varepsilon$ is connected. Thus $\overline{U_\varepsilon \cap V_\varepsilon}$ is connected; it is also closed and bounded, so compact. So $U \cap V = \cap_{\varepsilon > 0} \overline{U_\varepsilon \cap V_\varepsilon}$ is the intersection of a nested sequence of compact, connected sets, so is itself connected.

∎

2.5 Shifts and φ_j

We again fix $I \in \mathcal{J}_0$ and take W, G, A etc. to be as in Section 2.3.

For $j \in \{\pm 1, \ldots, \pm d\}$, define σ_j, the *shift in direction* j, by

$$\sigma_j(v) = v + e_j,$$

where e_j is the j^{th} standard basis vector if $j > 0$ and $e_j = -e_{-j}$ if $j < 0$, and set

$$G_0^j = \{v \in G_0 : \sigma_j^{-1}(v) \notin A\} = G_0 \cap \sigma_j(\mathcal{O} \setminus A).$$

Proposition 2.12 *For each j, the sets $I \setminus W$, $\sigma_j(I \cap W)$ and G_0^j are pairwise disjoint, and their union is an independent set.*

Proof. Trivially, $\sigma_j(I) \cap I = \emptyset$, so in particular $(I \setminus W) \cap \sigma_j(I \cap W) = \emptyset$; $(I \setminus W) \cap G_0^j = \emptyset$ is trivial (because $G_0^j \subseteq W$); and $\sigma_j(I \cap W) \cap G_0^j = \emptyset$ follows from the definiton of G_0^j. So the union is disjoint.

Clearly $(I \setminus W)$, $\sigma_j(I \cap W)$ and G_0^j are all independent sets. To show independence of the union, we must show that there are no edges between any two of them. Since $\nabla(I \setminus W, W) = \emptyset$ (by (12)) and $\sigma_j(I \cap W) \subseteq W$ (by (11)), we have $\nabla((I \setminus W), (\sigma_j(I \cap W) \cup G_0^j)) = \emptyset$.

This leaves $\nabla(\sigma_j(I \cap W), G_0^j)$. Suppose, for a contradiction, that $y \in G_0^j$ and $\sigma_k(y) \in \sigma_j(I \cap W)$ for some k. Then $z := \sigma_j^{-1}(\sigma_k(y)) \in I \cap W \cap \mathcal{E} \subset G \setminus G_0$ (by (11)), implying $\sigma_j^{-1}(y) = \sigma_k^{-1}(z) \in A$, contrary to the assumption $y \in G_0^j$. So $\nabla(\sigma_j(I \cap W), G_0^j) = \emptyset$.

∎

Define $\sigma_j^*(I) = (I \setminus W) \cup \sigma_j(I \cap W)$ and

$$\varphi_j(I) = \{J : \sigma_j^*(I) \subseteq J \subseteq \sigma_j^*(I) \cup G_0^j\}.$$

Then Proposition 2.12 implies

$$\varphi_j(I) \subseteq \mathcal{J}.$$

Notice also that we recover I from j, J ($\in \varphi_j(I)$) and (G, A); namely, if we are given (G, A), j, and $J \in \varphi_j(I)$, then

$$I = (J \setminus W) \cup \sigma_j^{-1}(J \cap (W \setminus G_0^j)). \tag{15}$$

2.6 Conventions and preview

Conventions

In much of what remains we can ignore I and concentrate on pairs from

$$\mathcal{G} := \{(G, A) \in 2^{\mathcal{E}} \times 2^{\mathcal{O}} : (G, A) \text{ satisfies (10) }\}.$$

Notice that under (10) each of G, A determines the other.

If (G, A) is produced from I as in Section 2.3 then we write $(G(I), A(I))$, noting that a given (G, A) may correspond to more than one I.

We will always take $W = G \cup A$ and $G_0 = \partial^\star W$ (a subset of $\mathcal{E}$ because of (10)).

Set $\ell = 2d$; so Γ is an ℓ-regular bigraph. (We tend to think in terms of d and use ℓ sparingly, for instance usually preferring $O(d)$ to the equivalent $O(\ell)$.) Though we usually work in Γ, we sometimes—especially in Section 2.9—consider more general graphs Σ, always assumed to satisfy

$$\Sigma \text{ is an } \ell\text{-regular bigraph with bipartition } V = \mathcal{O} \cup \mathcal{E}. \tag{16}$$

We always take $|G| = g$ and $|A| = a = (1-\delta)g$, and for given g, δ set

$$\mathcal{G}(g,\delta) = \{(G,A) \in \mathcal{G} : |G| = g, |A| = (1-\delta)g\},$$

$$\mathcal{J}(g,\delta) = \{I \in \mathcal{J}_0 : (G(I), A(I)) \in \mathcal{G}(g,\delta)\}.$$

(It's generally best to think of δ as small, though it will not always be so.)

As will appear, the quantity that really matters is almost always δg ($= |G| - |A|$), and it will be convenient to take, for any t,

$$\mathcal{G}(t) = \{(G,A) \in \mathcal{G} : |G| - |A| = t\}.$$

Notice that for $(G,A) \in \mathcal{G}(t)$,

$$|\nabla(W, V \setminus W)| \quad (= |\nabla(G_0, \mathcal{O} \setminus A)|) \quad = t\ell. \tag{17}$$

Though we don't really need t, we use it to emphasize a certain *duality*: if $(G,A) \in \mathcal{G}(t)$ in some graph Σ satisfying (16), then $(\mathcal{O} \setminus A, \mathcal{E} \setminus G)$ belongs to the analogue of $\mathcal{G}(t)$ obtained by reversing the roles of $\mathcal{O}$ and $\mathcal{E}$ in Σ—but of course g and δ, unlike t, are not usually preserved by this switch.

Preview

Our tasks are to define ν, for which (5) will turn out to be obvious, and establish (6).

We will eventually associate with each (G,A) a particular index $j = j(G,A)$, and set $j(I) = j(G(I), A(I))$. (This is basically a j for which $|G_0^j| = \log_2 |\varphi_j(I)|$ is large, though there are some additional considerations.) We then define $\varphi(I) = \varphi_{j(I)}(I)$ and require

$$J \notin \varphi(I) \Rightarrow \nu(I,J) = 0. \tag{18}$$

Let us call I *small* if $|G(I)| \leq d^3$ (we could get by with $d^{9/4}$; see (68)), and *large* otherwise.

For small I—an easy case, as we will see in Section 2.13—we simply choose $j = j(I)$ to maximize $|G_0^j|$ (where $G = G(I)$), so that, since

$$\sum_j |G_0^j| = |\nabla(G, \mathcal{O} \setminus A)| = \delta g \ell, \tag{19}$$

we have

$$|G_0^j| \geq \delta g. \tag{20}$$

We then set

$$\nu(I,J)=\lambda^{|J|-|I|}(1+\lambda)^{-|G_0^j|}\quad \forall J\in\varphi(I). \tag{21}$$

(Note this satisfies (5). The separate treatment of small I is unnecessary if we only want the phase transition, but is needed for the "correct" bound in (3).)

Most of our work (including everything in Sections 2.4 and 2.8-2.12) is geared to large I (though often valid in general). For most of our discussion we fix (g,δ), and aim to bound the contribution of $\mathcal{J}(g,\delta)$ to (6). Of course these contributions must eventually be summed, but this turns out not to add anything significant.

Before beginning in earnest, we pause in Section 2.7 to adapt the isoperimetric Lemma 2.4 to our situation (Lemma 2.13). This is needed especially in Section 2.13, but will also make an appearance in Section 2.8.

In Sections 2.8-2.10 we associate with each relevant (G,A) some $(F,S)\in 2^{\mathcal{E}}\times 2^{\mathcal{O}}$ which "approximates" (G,A) in an appropriate sense. The definitions of $j(I)$ and $\nu(I,\cdot)$ (in Section 2.11) are then based on our approximation to $(G(I),A(I))$. The main points are: (i) the set of possible approximations is small (Lemma 2.18); and (ii) for a given J, I's for which $(G(I),A(I))$ is approximated by a particular (F,S) don't contribute too much in (6) (see (53)), construction of a ν achieving this being made possible by the accuracy of our approximations.

The proof that ν behaves as desired (that is, of (53)) is given in Section 2.12, and Section 2.13 is a mopping up operation, combining what we already know for large I's with the easy analysis for small I's and the isoperimetric information from Lemma 2.13, to finally establish (6).

More conventions

For whatever G,A,F,S we have under discussion, we set $H=\mathcal{E}\setminus G$, $B=\mathcal{O}\setminus A$, $E=\mathcal{E}\setminus F$, $T=\mathcal{O}\setminus S$, $B_0=B\cap N(G)$, $S_0=S\cap N(E)$, and $E_0=E\cap N(S)$.

From now until Section 2.13 we fix g,δ and always take $I\in\mathcal{J}(g,\delta)$ and $(G,A)\in\mathcal{G}(g,\delta)$. (We will not see I again until Section 2.11.)

2.7 Isoperimetry

Before continuing, we need to work out what Lemma 2.4 implies in the way of a lower bound on δ for given g.

Lemma 2.13 *Suppose* $(G, A) \in \mathcal{G}(g, \delta)$ *satisfies*

$$(G \cup A) \cap \Delta = \emptyset. \tag{22}$$

Then

$$\delta = \begin{cases} \Omega(g^{-1/d}/d) & \text{for all } g \\ 1 - O(1/d) & \text{if } g < d^{O(1)}. \end{cases}$$

(For the (G, A)'s of interest to us, (22) is given by (7).)

Proof. In view of (22), the lemma does not change if we replace the torus Γ by the box Λ.

For the first part of the lemma, the main thing we have to show is

Proposition 2.14 $\quad s(r) = \Omega(b(r)^{1-1/d})$

(where $B(r)$, $S(r)$, $b(r)$, $s(r)$ are as defined before Lemma 2.4). Notice that this, combined with Lemma 2.4, implies that for any $C \subset \mathbf{Z}^d$,

$$|\partial C| = \Omega(|C|^{(d-1)/d}). \tag{23}$$

Proposition 2.14 is again something for which one would hope to just give a reference; but we could not find one, or even give the short proof that seems called for.

For the proof, we'll be interested in the average number of nonzero entries in an element of $S(q)$,

$$t(q) := s(q)^{-1} \sum_{x \in S(q)} |supp(x)|.$$

This is useful because, setting

$$N(q) = |\{(x, y) \in S(q) \times S(q+1) : x \sim y\}|,$$

we have

$$s(q)(2d - t(q)) = N(q) \leq s(q+1) \min\{q+1, d\},$$

implying

$$\frac{s(q)}{s(q+1)} \leq \frac{\min\{q+1, d\}}{2d - t(q)}. \tag{24}$$

This already implies Proposition 2.14 for, say, $r \le .9d$, since in this case we have

$$b(r) \le s(r) \sum_{i=0}^{r} \frac{(r)_i}{(2d-r+i)_i} \le s(r) \sum_{i \ge 0} \left(\frac{r}{2d-r} \right)^i = O(s(r)).$$

For larger r we will have to work harder. Here we first show, for $q = \beta d$ with $\beta > .9$,

$$t(q) < (1 - 1/(20\beta))d. \tag{25}$$

Let

$$S(q,t) = \{x \in S(q) : |supp(x)| = t\},$$

$s(q,t) = |S(q,t)|$, and define $B(q,t)$ and $b(q,t)$ similarly. Then

$$f(q,t) := \frac{s(q,t+1)}{s(q,t)} = 2\frac{(d-t)(q-t)}{(t+1)t}.$$

Set $t_0 = t_0(q) = \lceil (1 - 1/(4\beta))d \rceil$. Then $t \ge t_0$ implies

$$\begin{aligned} f(q,t) &\le 2\frac{(1/(4\beta))(\beta - 1 + 1/(4\beta))}{(1-1/(4\beta))^2} \\ &= 2\left(\frac{2\beta-1}{4\beta-1}\right)^2 < \frac{1}{2}. \end{aligned}$$

Thus

$$\begin{aligned} t(q) &= s(q)^{-1} \sum_{t \le q} t s(q,t) \\ &< t_0 + \sum_{i \ge 1} i 2^{-i} = t_0 + 2. \end{aligned}$$

This gives (25) provided $\beta \le d/15$. For larger β we just use

$$\frac{s(q,d-1)}{s(q,d)} = \frac{d(d-1)}{2(\beta d - d + 1)} > \frac{d-1}{2\beta},$$

whence

$$\begin{aligned} d - t(q) &= s(q)^{-1} \sum (d-i) s(q,i) \ge \sum_{i<d} s(q,i) / (\sum_{i \le d} s(q,i)) \\ &\ge s(q,d-1)/(s(q,d-1) + s(q,d)) \ge (d-1)/(2\beta + d - 1), \end{aligned}$$

which again gives (25).

Now let $r = \gamma d \geq .9d$. By (25) and (24) we have, for $r - i \geq .9d$,

$$s(r-i) \leq s(r) \prod_{j=1}^{i} \frac{d}{d + d^2/(20(r-j))} < s(r)(1 - \Omega(1/\gamma))^i,$$

so

$$b(r) \leq s(r) \sum_{i=0}^{r-.9d} (1 - \Omega(1/\gamma))^i + b(.9d) = O(\gamma s(r)) \tag{26}$$

(since we know $b(.9d) = O(s(.9d)) = O(s(r))$).

On the other hand, with $t_0 = t_0(r)$, we have

$$b(r) > b(r, t_0) = 2^{t_0} \binom{d}{t_0} \binom{r}{t_0} > \exp[t_0 \log(r/t_0)],$$

and $b(r)^{1/d} > \exp[(1 - 1/(4\gamma)) \log(r/t_0)] = \Omega(\gamma)$; and this with (26) gives Proposition 2.14.

∎

Now for the first part of Lemma 2.13, we consider the possibilities $|G_0| > |A|$ and $|G_0| \leq |A|$ separately, in both cases using the fact that $|G_0| \leq \delta g d$ (since $|G_0| \leq |\nabla(G, \mathcal{O} \setminus A)| = \delta g d$).

If $|G_0| > |A|$, then $\delta > 1/(d+1)$, so certainly $\delta = \Omega(g^{-1/d}/d)$. If, on the other hand, $|G_0| \leq |A|$, then we have (using (23) and the fact that $\partial((G \setminus G_0) \cup A) = G_0$)

$$\begin{aligned}
\delta &\geq |G_0|/(dg) \\
&= \Omega(|(G \setminus G_0) \cup A|^{(d-1)/d}/(dg)) \\
&= \Omega(|G|^{(d-1)/d}/(dg)) \\
&= \Omega(g^{-1/d}/d).
\end{aligned}$$

For small g notice that for $r < O(1)$,

$$s(r) = 2^r d^r / r! + O(d^{r-1}),$$

which in view of Lemma 2.4 implies that for $C \subseteq \mathbf{Z}^d$ with $|C| < d^{O(1)}$,

$$|\partial C| = \Omega(|C| d).$$

Applying this with $C = W \setminus G_0$ gives $|G_0| = (1 - O(1/d))g$. But then $|\nabla(G_0, A)| \le \ell|A| = O(|G_0|)$ implies

$$\delta g \ell = |\nabla(G_0, \mathcal{O} \setminus A)| \ge (\ell - O(1))|G_0| = \ell(1 - O(1/d))g.$$

∎

2.8 First approximation: covering the boundary

Say a set $C \subseteq \Gamma$ *separates* $P, Q \subseteq \Gamma$ if any path meeting both P and Q also meets C.

In this section we begin the process of approximation by showing that there is a "small" collection of subsets of Γ, at least one of which separates W ($= G \cup A$) and $\Gamma \setminus W$ for each relevant (G, A). We then use these separations to show that there is a small $\mathcal{S} \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ such that each of our (G, A)'s is approximated by some $(F, S) \in \mathcal{S}$ in the sense that

$$S \supseteq A, \ F \subseteq G \tag{27}$$

and

$$|S \setminus A|, |G \setminus F| < O(\delta g \sqrt{d \log d}\ ). \tag{28}$$

This is stated formally in Lemma 2.16 at the end of the section.

Our argument applies to pairs from

$$\mathcal{G}^\star := \{(G, A) \in \mathcal{G}(g, \delta) : (G, A) \text{ satisfies (7) and (14)}\},$$

though the main point, Lemma 2.15, is valid for all of $\mathcal{G}(t)$.

In this section (unlike in the next) we make substantial use of properties particular to Γ, specifically the isoperimetric properties given by Lemma 2.4 and

$$\forall\ w \sim v \text{ and } L \subseteq N(v),\ |N(w) \cap N(L)| \ge |L| \tag{29}$$

(which follows from the fact that for vertices $v \sim w$, $\Gamma[(N(v) \cup N(w)) \setminus \{v, w\}]$ is a matching of all but one vertex of $N(v)$ and all but one vertex of $N(w)$).

Let

$$G_0' = \{v \in G : d_A(v) \le \ell/2\} \quad (\subseteq G_0),$$
$$B_0' = \{v \in B : d_H(v) \le \ell/2\} \quad (\subseteq B_0),$$

$G_0'' = G_0 \setminus G_0'$ and $B_0'' = B_0 \setminus B_0'$. Then

$$\nabla(G_0'', B_0'') = \emptyset. \tag{30}$$

(The more general statement here is: if $v \in G_0$, $w \in B_0$ and $v \sim w$, then (by (29) with $L = N(v) \cap A$) $d_G(w) \geq d_A(v)$ $(= \ell - d_B(v))$, implying $d_B(v) + d_G(w) \geq \ell$.)

Notice that (30) implies

$$G_0' \cup B_0' \text{ separates } W \text{ and } \Gamma \setminus W \tag{31}$$

(equivalently, $\nabla(W, \Gamma \setminus W) \subseteq \nabla(G_0') \cup \nabla(B_0')$).

Lemma 2.15 *In any graph satisfying* (16) *and* (29)*, for any* $(G, A) \in \mathcal{G}(t)$*, there exists* $U \subseteq N(G_0' \cup B_0')$ *satisfying*

$$N(U) \supseteq G_0' \cup B_0' \tag{32}$$

and

$$|U| < O(t\sqrt{\log \ell / \ell}\ ). \tag{33}$$

Before proving this, we observe that it does accomplish the first goal stated at the beginning of this section (existence of a small set of separations). For (G, A) and U as in Lemma 2.15, we have

$$N(U) \text{ separates } W \text{ and } \Gamma \setminus W \tag{34}$$

(by (31) and (32)). So we just need to limit the number of possibilities for U when $(G, A) \in \mathcal{G}^\star$.

To do so, notice that

$$U \text{ is 6-clustered.} \tag{35}$$

This follows from Lemma 2.3 and (14), once we observe that $\text{dist}(u, G_0) \leq 2\ \forall u \in U$ (since $U \subseteq N(G_0' \cup B_0')$), and that (32) and (30) imply $\text{dist}(v, U) \leq 2\ \forall v \in G_0$.

In view of (33) (with $t = \delta g$), Lemma 2.1 then gives, for example, a bound

$$O(gd^2)(Cd^6)^{O(\delta g\sqrt{\log d/d}\ )} = \exp[O(\delta g d^{-1/2} \log^{3/2} d)] \tag{36}$$

on the number of possibilities for U. Here we used Lemma 2.13 for the equality in (36). The initial $O(gd^2)$ corresponds to a choice of x_0 in Lemma 2.1:

in view of (7), there must be some $j \in [-d,d] \setminus \{0\}$ and $k \le g/(2d)$ for which $y_0 := v_0 + (2k-1)e_j \in G_0$; there are at most g possibilities for this y_0, so at most $O(gd^2)$ possibilities for a vertex x_0 with $d(x_0, y_0) \le 2$; and by (32) and (30) U must contain such an x_0.

Proof of Lemma 2.15.

By "duality" (see Section 2.6) it's enough to show the existence of $S \subseteq N(G_0')$ with

$$N(S) \supseteq G_0' \tag{37}$$

and

$$|S| < O(t\sqrt{\log \ell/\ell}\,). \tag{38}$$

Define $Q = \{v \in G_0 : d_A(v) \le \sqrt{\ell \log \ell}\}$, $K = G_0 \setminus Q$, and $P = N(Q) \cap A$. By (29),

$$d_{G_0}(v) \ge \ell - \sqrt{\ell \log \ell} \quad \forall v \in P. \tag{39}$$

Let $P' = \{v \in P : d_K(v) \ge \ell/2\}$, $P'' = P \setminus P'$, $Q' = Q \cap N(P')$, $Q'' = Q \setminus Q'$ and $R = \{v \in B_0 \cap N(G_0') : d_{G_0}(v) > \sqrt{\ell \log \ell}\}$.

Now P'' is a cover of Q'' of size $O(t\sqrt{\log \ell/\ell}\,)$, the size bound following from $|Q| \le t\ell/(\ell - \sqrt{\ell \log \ell}\,) = O(t)$ (using (17)), $d_{P''}(v) \le d_A(v) \le \sqrt{\ell \log \ell}$ $\forall v \in Q$, and $d_Q(v) > \ell/2 - \sqrt{\ell \log \ell}$ $\forall v \in P''$ (using (39) and the definition of P'').

On the other hand, we can cover $G_0' \setminus Q''$ by a similarly small subset of R, as follows. From (29) we have $N(K) \cap N(G_0') \cap B_0 \subseteq R$. This gives $d_R(v) > \ell/2$ for $v \in G_0' \setminus Q$, while for $v \in Q'$,

$$d_R(v) \ge |N(v) \cap N(K)| - |N(v) \cap A| \ge \ell/2 - \sqrt{\ell \log \ell}$$

(the second inequality following from (29) and the definitions of Q' and Q). So, noting that $|R| < t\sqrt{\ell/\log \ell}$ (again using (17)), Lemma 2.2 says that we can cover $G_0' \setminus Q''$ by some $T \subseteq R$ of size at most $|R|(1+\log \ell)/(\ell/2 - \sqrt{\ell \log \ell}\,) < O(t\sqrt{\log \ell/\ell}\,)$. (And note $P \subseteq N(G_0')$ since $Q \subseteq G_0'$, and $R \subseteq N(G_0')$ by definition, so $S := P'' \cup T \subseteq N(G_0')$.)

∎

We now return to Γ. Given U as above, let us temporarily set $L = N(U)$. Then $|L| = O(\delta g\sqrt{d \log d}\,)$.

Say a component C of $\Gamma - L$ is *large* if $|C| > d$ and *small* otherwise. Lemma 2.4 implies

$$|\nabla(C, L)| = |\nabla(C)| \geq |\partial C| = \Omega(|C|d)$$

for small C (actually also for considerably larger C), and

$$|\nabla(C, L)| = \Omega(d^2)$$

for large C. But $|\nabla(L)| \leq 2d|L| = O(\delta g d^{3/2}\sqrt{\log d}\,)$, so

$$\text{the number of large components is } O(\delta g d^{-1/2}\sqrt{\log d}\,), \tag{40}$$

and the number of vertices in small components is $O(\delta g\sqrt{d\log d}\,)$.

It follows that if (G, A) is any pair satisfying (10) for which L separates W and $\Gamma \setminus W$, then we satisfy (27) and (28) with

$$F = P \cap \mathcal{E} \quad \text{and} \quad S = (P \cup Q \cup L) \cap \mathcal{O}, \tag{41}$$

where P is the union of those large components of $\Gamma - L$ that meet (equivalently, are contained in) W, and Q is the union of (all) the small components. In particular this is true if (G, A) is any pair from $\mathcal{G}^\star$ for which Lemma 2.15 applied to (G, A) produces U.

By (40) the number of possibilities (given L) for (F, S) as in (41) is at most $\exp[O(\delta g d^{-1/2}\sqrt{\log d}\,)]$, and combining this with the bound (36) on the number of U's we have

Lemma 2.16 *There exist* $\mathcal{S} \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ *with*

$$|\mathcal{S}| < \exp[O(\delta g d^{-1/2}\log^{3/2} d)] \tag{42}$$

and a map $\pi_1 : \mathcal{G}^\star \to \mathcal{S}$ *such that* (27) *and* (28) *hold for each* $(G, A) \in \mathcal{G}^\star$ *and* $(F, S) = \pi_1(G, A)$.

2.9 Second approximation

The discussion in this section is valid for any graph Σ satisfying (16). It may be worth reiterating that we follow the conventions given at the end of Section 2.6.

Given $(F^*, S^*) \in 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ and a positive x, write $\mathcal{G}' = \mathcal{G}'(F^*, S^*, x)$ for the set of (G, A)'s in $\mathcal{G}(t)$ satisfying (27) (with (F^*, S^*) in place of (F, S)) and

$$|S^* \setminus A|, |G \setminus F^*| < x. \tag{43}$$

Lemma 2.17 *With notation as above, for any $0 < \psi < \ell$, there exist $\mathcal{T} \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$,*

$$|\mathcal{T}| < \exp[O((x/\ell) + (t/\psi)) \log \ell], \tag{44}$$

and a map $\pi_2 : \mathcal{G}' \to \mathcal{T}$ such that for each $(G, A) \in \mathcal{G}'$ and $(F, S) = \pi_2(G, A)$ we have (27) *and*

$$v \in S \Rightarrow d_F(v) > \ell - \psi, \quad v \in E \Rightarrow d_T(v) > \ell - \psi \tag{45}$$

(where as usual $E = \mathcal{E} \setminus F$ and $T = \mathcal{O} \setminus S$).

Remarks. We only need Lemma 2.17 when $(F^*, S^*) \in \mathcal{S}$ (with $\mathcal{S}$ as in Lemma 2.16), in which case we take $t = \delta g$ and $x = O(\delta g \sqrt{d \log d}\ )$ (with an appropriate constant), so that $\mathcal{G}' \supseteq \pi_1^{-1}(F^*, S^*)$; but the extra generality costs us nothing. The pairs we produce will satisfy $S \subseteq S^*$ and $F \supseteq F^*$, but we don't need this in what follows.

Proof of Lemma 2.17

We would like to exhibit a procedure which, for a given $(G, A) \in \mathcal{G}'$, outputs a pair (F, S) satisfying (27) and (45), and show that the set $\mathcal{T}$ of pairs produced in this way is small.

We produce (F, S) via a sequence of modifications, initializing at $(F, S) = (F^*, S^*)$. Note that whenever we update (F, S), we also automatically update E, T, etc.

One preliminary observation:

$$|S_0^*|, |E_0^*| < x + \ell x \tag{46}$$

(since $S_0^* \subseteq (S^* \setminus A) \cup N(G \setminus F^*)$, and similarly for E_0^*; recall $S_0^* = S^* \cap N(E^*)$ and $E_0^* = E^* \cap N(S^*)$, where $E^* = \mathcal{E} \setminus F^*$).

Stage 1A Set $\xi = \ell/2$.

(A.1) Repeat for as long as possible: choose $w \in H$ with $d_S(w) \geq \xi$ and do $S \leftarrow S \setminus N(w)$.

(A.2) When no longer possible, do $F \leftarrow F \cup \{w \in \mathcal{E} : d_S(w) \geq \xi\}$.

Stage 1B Do the same thing in the dual; that is,

(B.1) for as long as possible, choose $w \in A$ with $d_E(w) \geq \xi$ and do $F \leftarrow F \cup N(w)$, and

(B.2) when no longer possible, do $S \leftarrow S \setminus \{w \in \mathcal{O} : d_E(w) \geq \xi\}$.

Notice—a crucial idea—that (F, S) produced by Stage 1 does satisfy (27).

Analysis:

The output (F, S) of Stage 1 is determined by the sets of w's used in (A.1) and (B.1).

Since each iteration in (A.1) shrinks $|S|$ by at least ξ while maintaining $A \subseteq S$, the number of iterations is less than $x/\xi = 2x/\ell$. Moreover, each w used in (A.1) lies in $N(S_0^*)$. So the number of possibilities for the set of w's used in (A.1) is less than $\sum_{i \leq x/\xi} \binom{\ell |S_0^*|}{i} < \exp[O((x/\ell) \log \ell)]$ (using (46)).

At the end of (A.2) we have $w \in G \setminus F \Rightarrow d_T(w) > \ell - \xi = \ell/2$, which, since $|\nabla(G, T)| \leq t\ell$ (see (17)), gives $|G \setminus F| < 2t$.

Similarly, the number of choices for the set of w's used in Stage 1B is at most $\exp[O((x/\ell) \log \ell)]$ (note Stage 1A does not increase E_0^*), and at the end of this stage we have $|S \setminus A| < 2t$.

Stage 2 now repeats Stage 1, starting with the revised (F, S), using ψ in place of ξ, and replacing (43) and (46) by

$$|S \setminus A|, |G \setminus F| < 2t$$

and

$$|S_0|, |E_0| < 2t(1 + \ell).$$

This clearly produces an (F, S) satisfying (27) and (45). Moreover, repeating the analysis above, we find that the number of possible outputs of Stage 2, for a given output of Stage 1, is at most $\exp[O((t/\psi) \log \ell)]$. So the number of possible outputs of the entire procedure is no more than $\exp[O((x/\ell) + (t/\psi)) \log \ell]$.

∎

2.10 Status

We now specify $t = \delta g$ and $x = O(\delta g\sqrt{d \log d}\ )$ (the bound in (28)), and $\psi = \sqrt{d}$ (any $\psi \in (\Omega(\sqrt{d/\log d}), O(\sqrt{d \log d}))$ would do; see the remark following (62).) Specializing to these values and combining Lemmas 2.16 and 2.17, we have

Lemma 2.18 *There exist* $\mathcal{U} \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$*,*

$$|\mathcal{U}| < \exp[O(\delta g d^{-1/2} \log^{3/2} d)], \tag{47}$$

and $\pi : \mathcal{G}^\star \to \mathcal{U}$ *such that* (27) *and* (45) *hold for each* (G, A) *and* $(F, S) = \pi(G, A)$*.*

(The expression in the exponent in (47) is the maximum of the corresponding expressions from (42) and (44).)

Now consider some $(F, S) \in \mathcal{U}$. Notice that, for any $(G, A) \in \pi^{-1}(F, S)$, $Q := S_0 \cup E_0$ contains all vertices whose locations in the partition $\Gamma = G \cup H \cup A \cup B$ are as yet unknown; namely, we have

$$F \subseteq G, \quad T \subseteq B, \quad S \setminus S_0 \subseteq A, \quad E \setminus E_0 \subseteq H$$

(the first two containments are just (27); $S \setminus S_0 \subseteq A$ follows from $F \subseteq G$, (10) and the definition of S_0, and $E \setminus E_0 \subseteq H$ is similar).

By convention, whenever we are given an (F, S)*, we take* Q *to be as defined in the preceding paragraph, and write* Γ_Q *for the subgraph induced by* Q*.*

2.11 Flow

Here, finally, we define ν (for large I; for small I, see Section 2.6).

Throughout the section we fix $(F, S) \in \mathcal{U}$. It is now convenient to write $G \sim (F, S)$ if $\pi(G, A) = (F, S)$ and $I \sim (F, S)$ if $G(I) \sim (F, S)$.

To define $\nu(I, \cdot)$ for $I \sim (F, S)$, we first need to choose a direction $j = j(I)$. Fix such an I and let $G = G(I)$, $A = A(I)$, etc. The choice of j will depend only on (G, A). Observe that (using (45))

$$\sum_j |\sigma_j(S_0 \cap A) \cap E_0| = |\nabla(S_0 \cap A, G \cap E_0)| < |G \cap E_0|\psi,$$

$$\sum_j |\sigma_j^{-1}(E_0) \cap (S_0 \setminus A)| = |\nabla(E_0, S_0 \setminus A)| < |S_0 \setminus A|\psi.$$

But (45) and (17) imply $|G \cap E_0| + |S_0 \setminus A| < \delta g\ell/(\ell - \psi)$, so that

$$\begin{aligned}\sum_j |\sigma_j(S_0) \cap E_0| &= \sum_j (|\sigma_j(S_0 \cap A) \cap E_0| + |\sigma_j^{-1}(E_0) \cap (S_0 \setminus A)|) \\ &< \delta g\ell\psi/(\ell - \psi). \qquad (48)\end{aligned}$$

We assert that we can choose j so that

$$|G_0^j| > .8\delta g \qquad (49)$$

and

$$|\sigma_j(S_0) \cap E_0| < 10|G_0^j|\psi/\ell. \qquad (50)$$

To see this, let

$$P = \{j \in [-d, d] \setminus \{0\} : |\sigma_j(S_0) \cap E_0| \geq 10|G_0^j|\psi/\ell\}.$$

Then (48) gives

$$\sum_{j \in P} |G_0^j| \leq \frac{\ell}{10\psi} \sum |\sigma_j(S_0) \cap E_0| < \delta g \frac{\ell^2}{10(\ell - \psi)},$$

so (using (19))

$$\sum_{j \notin P} |G_0^j| > (1 - \ell/(10(\ell - \psi)))\delta g\ell.$$

So there exists $j \notin P$ with (say) $|G_0^j| > .8\delta g$, which is what we want.

∎

Having chosen j satisfying (49) and (50), we turn to defining $\nu(I, \cdot)$. Let

$$C = C^j(I) = G_0^j \cap F \cap \sigma_j(S_0) \ \ (= \sigma_j(S_0 \setminus A) \cap F),$$

$$D = D^j(I) \ = \ G_0^j \cap (\sigma_j(T) \cup (\sigma_j(S_0) \cap E_0)).$$

Then

$$C \cup D \text{ is a partition of } G_0^j. \qquad (51)$$

Setting $\alpha = \alpha(\lambda) = \lambda/(1+\lambda)^2$ and $\beta = \beta(\lambda) = 1-\alpha\lambda = (1+2\lambda)/(1+\lambda)^2$, define

$$\nu(I,J) = \begin{cases} (\alpha\lambda)^{|C\cap J|}\beta^{|C\setminus J|}(\lambda/(1+\lambda))^{|D\cap J|}(1+\lambda)^{-|D\setminus J|} & \\ \qquad = \frac{w(J)}{w(I)}\alpha^{|C\cap J|}\beta^{|C\setminus J|}(1+\lambda)^{-|D|} & \text{if } j \in \varphi_j(I) \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\sum_J \nu(I,J) = 1 \qquad \forall I \tag{52}$$

(because of (51)). On the other hand we will show, for any J,

$$\sum_{I\sim(F,S)} \frac{w(I)}{w(J)}\nu(I,J) \leq \ell\beta^{\delta g/2}. \tag{53}$$

2.12 Proof of (53)

We need one easy lemma. Given a bigraph Σ on $P\cup R$ and $U \subseteq R$, say that a (vertex) cover $K\cup L\cup M$ of Σ with $K\subseteq P$, $L\subseteq U$ and $M\subseteq R\setminus U$ is *legal* (with respect to U) if it is a minimal cover and

$$K = N(U\setminus L).$$

(Note minimality implies $K = N(R\setminus(L\cup M))$.)

Lemma 2.19 *With notation as above, let $K\cup L\cup M$ be a legal cover with $|K\cup L|$ as small as possible. Then*

(a) $\forall K'\subseteq K \quad |N(K')\cap(U\setminus L)| \geq |K'|$,

(b) $\forall L'\subseteq L \quad |N(L')\setminus K| \geq |L'|$.

Proof. (a) Given $K'\subseteq K$, let $S = N(K')\cap(U\setminus L)$,

$$K'' = \{v\in K : N(v)\cap U \subseteq S\cup L\} \;\; (\supseteq K'),$$

and $T = N(K'')\cap(R\setminus U)$. Then

(i) $(K\setminus K'')\cup(L\cup S)\cup(M\cup T)$ is a minimal cover

(a straightforward verification using the fact that each vertex of $K\setminus K''$ has a neighbor in $U\setminus(L\cup S)$), and

(ii) $K \setminus K'' = N(U \setminus (L \cup S))$.

Minimality of $|K \cup L|$ thus implies $|K \setminus K''| + |L \cup S| \geq |K| + |L|$, so $|S| \geq |K''| \geq |K'|$.

(b) This is similar. Given $L' \subseteq L$, let $W = N(L') \setminus K$ and

$$L'' = \{u \in L \cup M : N(u) \subseteq K \cup W\} \quad (\supseteq L').$$

Then

(i) $K \cup W \cup ((L \cup M) \setminus L'')$ is a minimal cover, and

(ii) $K \cup W = N(U \setminus (L \setminus L''))$.

Minimality of $|K \cup L|$ thus implies $|K \cup W| + |L \setminus L''| \geq |K| + |L|$, and $|W| \geq |L''| \geq |L'|$.

∎

Proof of (53).

Given (F, S), J and j, set

$$\mathcal{I}^\star = \mathcal{I}^\star(F, S, J, j) = \{I \sim (F, S) : j(I) = j, J \in \varphi_j(I)\}.$$

We will show

$$\sum_{I \in \mathcal{I}^\star} \frac{w(I)}{w(J)} \nu(I, J) < \beta^{\delta g/2},$$

which of course gives (53).

Set $U = \sigma_j^{-1}(J) \cap S_0$. Suppose $I \in \mathcal{I}^\star$, and set $G = G(I)$, $A = A(I)$, and

$$K = K(I) = G \cap E_0, \quad L = L(I) = U \setminus A, \quad M = M(I) = (S_0 \setminus U) \setminus A.$$

Then $K \cup L \cup M$ $(= (G \cup B) \cap Q)$ is a minimal cover of Γ_Q. (That it is a cover follows from (10); for minimality, notice (e.g.) that each $v \in G \cap E_0$ has a neighbor in A, which must be in S_0 (using $A \subseteq S$ and the definition of S_0).) Moreover, we assert,

$$K = N_{\Gamma_Q}(U \setminus L). \tag{54}$$

Proof. We show that each side of (54) contains the other. The obvious direction is

$$N_{\Gamma_Q}(U \setminus L) = N_{\Gamma_Q}(U \cap A) \subseteq N(A) \cap E_0 = G \cap E_0 = K.$$

For the reverse containment, suppose $v \in K$. Since $K \subseteq G_0$, (13) says that v has a neighbor $u \in A \cap I$. Then $u \in S_0$ (because $v \in E_0 \not\sim S \setminus S_0$), implying $u \in U$ (since $u \in A \cap I \Rightarrow \sigma_j(u) \in J$). And of course $u \notin L$ (since $u \in A$).

∎

Thus $K \cup L \cup M$ is a legal cover of Γ_Q with respect to U in the sense of Lemma 2.19.

Now fix $K_0 \cup L_0 \cup M_0$, a legal cover of Γ_Q with respect to U with $|K_0 \cup L_0|$ as small as possible.

Given $I \in \mathcal{I}^\star$, let $K = K(I)$ etc. be as above and set $K' = K_0 \setminus K$, $L' = L_0 \setminus L$. Then by Lemma 2.19,

$$|L| \geq |K'| + |L_0 \setminus L'|, \quad |K| \geq |L'| + |K_0 \setminus K'|. \tag{55}$$

Furthermore, we assert,

$$K = (K_0 \setminus K') \cup N_{\Gamma_Q}(L'). \tag{56}$$

The point of this is that it says that (K', L') determines G (so also A), and therefore $I \in \mathcal{I}^\star$ (because of (15)).

To see (56), just observe that the only point requiring proof is $K \setminus K_0 \subseteq N_{\Gamma_Q}(L_0 \setminus L)$, and that this follows from (54) once we notice that $\nabla(K \setminus K_0, U \setminus (L_0 \cup L)) = \emptyset$ (since $K_0 \cup L_0$ covers $\nabla(E_0, U)$).

Now with $C = C^j(I)$, $D = D^j(I)$ as in the discussion preceding (51), observe that

$$C \cap J = \sigma_j(L \setminus \sigma_j^{-1}(E_0)) \quad \text{and} \quad C \setminus J = \sigma_j(M \setminus \sigma_j^{-1}(E_0)),$$

and that we may partition D as

$$D = (\sigma_j(T) \cap F) \cup (K \setminus \sigma_j(S_0 \setminus (L \cup M))).$$

Thus, with inequalities justified below,

$$\begin{aligned}
\frac{w(I)}{w(J)}\nu(I,J) &= \alpha^{|\sigma_j(L\setminus\sigma_j^{-1}(E_0))|}\beta^{|\sigma_j(M\setminus\sigma_j^{-1}(E_0))|} \\
&\qquad \cdot (1+\lambda)^{-(|\sigma_j(T)\cap F|+|K\setminus\sigma_j(S_0\setminus(L\cup M))|)} \\
&\le \alpha^{|L|}\beta^{|M|}(1+\lambda)^{-(|K|+|\sigma_j(T)\cap F|)} \\
&\qquad \cdot \alpha^{-(|\sigma_j(S_0\cap A)\cap K|+|\sigma_j^{-1}(E_0)\cap(S_0\setminus A)|)} &(57)\\
&\le \alpha^{|L|}(1+\lambda)^{-|K|}\beta^{|G_0^j|-(|K|+|L|)}\alpha^{-O(|G_0^j|\psi/\ell)} &(58)\\
&\le \beta^{\delta g/2}\alpha^{|L|}(1+\lambda)^{-|K|}\beta^{-(|K|+|L|)} &(59)\\
&= \beta^{\delta g/2}\left(\frac{1+\lambda}{1+2\lambda}\right)^{|K|}\left(\frac{\lambda}{1+2\lambda}\right)^{|L|} \\
&\le \beta^{\delta g/2}\left(\frac{1+\lambda}{1+2\lambda}\right)^{|L'|+|K_0\setminus K'|}\left(\frac{\lambda}{1+2\lambda}\right)^{|K'|+|L_0\setminus L'|} &(60)\\
&= \beta^{\delta g/2}\left(\frac{1+\lambda}{1+2\lambda}\right)^{|K_0|}\left(\frac{\lambda}{1+2\lambda}\right)^{|L_0|}\left(\frac{\lambda}{1+\lambda}\right)^{|K'|-|L'|}.
\end{aligned}$$

(In (57) we used $\alpha^{-1} = \max\{\alpha^{-1}, \beta^{-1}, 1+\lambda\}$; in (58) we used $G_0^j \subseteq \sigma_j(L \cup M) \cup K \cup (\sigma_j(T) \cap F)$, $(1+\lambda)^{-1} < \beta$ and (50); (59) is from (49), using $(\psi/\ell)\log(1/\alpha) = o(\log(1/\beta))$, which is a consequence of

$$\lambda^2 = \omega((\psi/\ell)\log(1/\lambda)) \tag{61}$$

for small λ, and easily verified when λ is larger; and (60) comes from (55).)

Thus, recalling—see the remark following (56)—that each (K', L') corresponds to at most one $I \in \mathcal{I}^\star$,

$$\begin{aligned}
\sum_{I\in\mathcal{I}^\star}\frac{w(I)}{w(J)}\nu(I,J) &\le \beta^{\delta g/2}\left(\frac{1+\lambda}{1+2\lambda}\right)^{|K_0|}\left(\frac{\lambda}{1+2\lambda}\right)^{|L_0|}\sum_{K'\subseteq K_0}\sum_{L'\subseteq L_0}\left(\frac{\lambda}{1+\lambda}\right)^{|K'|-|L'|} \\
&= \beta^{\delta g/2}.
\end{aligned}$$

As noted earlier this gives (53).

2.13 Finally

Now fixing $J \in \mathcal{J}$, we are ready to verify (6) (thus completing the proofs of Theorems 1.2 and 1.1).

Note first of all (referring to (47)) that for $\lambda \le 2$ (say) (53) implies

$$\begin{aligned}
\sum_{I \in \mathcal{J}(g,\delta)} \frac{w(I)}{w(J)} \nu(I,J) &= \sum_{(F,S)\in\mathcal{U}} \sum_{I\sim(F,S)} \frac{w(I)}{w(J)}\nu(I,J) \\
&\le |\mathcal{U}|\ell\beta^{\delta g/2} \\
&< \ell\exp[\{O(d^{-1/2}\log^{3/2} d) - \Omega(\lambda^2)\}\delta g] \\
&< \exp[-\Omega(\lambda^2\delta g)], \qquad (62)
\end{aligned}$$

while for larger λ,

$$\sum_{I\in\mathcal{J}(g,\delta)} \frac{w(I)}{w(J)}\nu(I,J) < \lambda^{-\Omega(\delta g)}. \qquad (63)$$

Remark. Our choice of ψ was constrained by the demands of (61) and (62) (the latter since $\psi = o(\sqrt{d/\log d})$ would give—via (44)—a larger bound in (47)).

We first deal with large I's (recall I is large if $|G(I)| > d^3$). Here we have already done the work: Assuming first that $\lambda \le 2$, and with justifications to follow, we have

$$\begin{aligned}
\sum_{I \text{ large}} \frac{w(I)}{w(J)}\nu(I,J) &= \sum_{g>d^3}\sum_{\delta}\sum_{I\in I(g,\delta)} \frac{w(I)}{w(J)}\nu(I,J) \\
&= \sum_{g>d^3}\sum_{\delta} \exp[-\Omega(\lambda^2\delta g)] \qquad (64)\\
&\le \sum_{g>d^3}\sum\{\exp[-\Omega(\lambda^2 i)] : i \ge \Omega(d^{-1}g^{1-1/d})\} \qquad (65)\\
&\le \sum_{g>d^3} \exp[-\Omega(\lambda^2(d^{-1}g^{1-1/d}))] \qquad (66)\\
&< \exp[-\Omega(\lambda^2 d^{3(1-1/d)-1})] \qquad (67)\\
&< \exp[-\omega(\lambda d)]. \qquad (68)
\end{aligned}$$

Of course sums involving δ, are restricted to δ for which δg is an integer. The main inequality (64) is just (62), and (65) comes from Lemma 2.13. In (66) we have absorbed a factor λ^{-2} in the exponent. One way (probably not the most natural) to see the inequality in (67) is to use

$$(1-\varepsilon)^{g^{1-\delta}} < (1-\varepsilon)^{iK^{1-\delta}} \quad \text{for} \quad i^{1/(1-\delta)}K < g \leq (i+1)^{1/(1-\delta)}K$$

with $K = d^3$, $\delta = 1/d$ and $1-\varepsilon = \exp[-\Omega(\lambda^2 d^{-1})]$.

For $\lambda > 2$ a similar analysis (using (63)) gives

$$\sum_{I \text{ large}} \frac{w(I)}{w(J)} \nu(I,J) \leq \lambda^{-\Omega(d^2)}. \tag{69}$$

Finally we turn to the easy case of small I. Here we abuse our notation slightly and set

$$\mathcal{J}(g,a) = \{I \in \mathcal{J}_0 : |G(I)| = g, |A(I)| = a\}.$$

For a (nonempty) $\mathcal{J}(g,a)$ with $g < d^3$, Lemma 2.13 gives $a = O(g/d)$, so that, since each $A(I)$ is 2-clustered and contains v_0, Lemma 2.1 bounds the number of possibilities for $A(I)$ with $I \in \mathcal{J}(g,a)$ by $\exp[O((g/d)\log d)]$.

But we also know (see (15)) that, given J and j, $I \in \varphi_j^{-1}(J)$ is determined by $G(I)$ (or $A(I)$), and that (by (21), (20), and again Lemma 2.13)

$$\begin{aligned} \frac{w(I)}{w(J)}\nu(I,J) &= (1+\lambda)^{-|G_0^j(I)|} \\ &\leq (1+\lambda)^{-\delta g} \\ &= (1+\lambda)^{-(1-O(1/d))g}. \end{aligned}$$

So finally, noting that $A(I) \neq \emptyset$ implies $|G(I)| \geq \ell$, we have

$$\begin{aligned} \sum_{I \in \mathcal{J}(g,a)} \frac{w(I)}{w(J)}\nu(I,J) &< \ell \exp[O((g/d)\log d)](1+\lambda)^{-(1-O(1/d))g} \\ &< (1+\lambda)^{-(1-o(1))g} \end{aligned}$$

and

$$
\begin{aligned}
\sum_{I \text{ small}} \frac{w(I)}{w(J)}\nu(I,J) &= \sum_{\ell\le g\le d^3}\sum_{a\le g}\sum_{I\in\mathcal{J}(g,a)} \frac{w(I)}{w(J)}\nu(I,J) \\
&< \sum_{\ell\le g<d^3} g(1+\lambda)^{-(1-o(1))g} \\
&\le (1+\lambda)^{-(1-o(1))\ell};
\end{aligned}
$$

and combining this with (68) or (69) gives (6).

Acknowledgments We are very grateful to Vladimir Gurvich for translating parts of [20]. Thanks also to Chuck Weibel for help with the proof of Lemma 2.8.

References

[1] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, *Stoch. Proc. Appl.* **49** (1994), 179–197.

[2] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998.

[3] B. Bollobás and I. Leader, An isoperimetric inequality on the discrete torus, *SIAM J. Disc. Math.* **3** (1990), 32–37.

[4] G.R. Brightwell, O. Häggström and P. Winkler, Nonmonotonic behavior in the hard-core and Widom-Rowlinson models, *J. Stat. Phys.* **94** (1999), 415–435.

[5] R. Diestel, *Graph Theory*, Springer, New York, 2000.

[6] R.L. Dobrushin, The problem of uniqueness of a Gibbs random field and the problem of phase transition, *Functional Anal. Appl.* **2** (1968), 302–312.

[7] Z. Füredi, Matchings and covers in hypergraphs, *Graphs and Comb.* **4** (1988), 115–206.

[8] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin, 1988.

[9] G. Grimmett, *Percolation*, Springer-Verlag, Berlin, 1999.

[10] O. Häggström, Ergodicity of the hard-core model on $\mathbf{Z}^2$ with parity-dependent activities, *Ark. Mat.* **35** (1997), 171–184.

[11] J. Kahn, Asymptotics of the list-chromatic index for multigraphs, *Random Structures & Algorithms* **17** (2000), 117–156.

[12] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Probab. Comput.* **10** (2001), 219–237.

[13] J. Kahn, Entropy, independent sets and antichains: a new approach to Dedekind's Problem, *Proc. Amer. Math. Soc.* **130** (2001), 371–378.

[14] J. Kahn and P.M. Kayll, Fractional vs. integer covers in hypergraphs of bounded edge size, *J. Combinatorial Th. (A)* **78** (1997), 199–235.

[15] D. Knuth, *The Art of Computer Programming*, Vol. I, Addison Wesley, London, 1969.

[16] A.D. Korshunov, The number of monotone Boolean functions, *Problemy Kibernet.*, **38** (1980), 5–108. (Russian)

[17] A.D. Korshunov and A.A. Sapozhenko, The number of binary codes with distance 2, *Problemy Kibernet.* **40** (1983), 111–130. (Russian)

[18] G.M. Louth, *Stochastic networks: complexity, dependence and routing*, thesis, Cambridge University, 1990.

[19] L. Lovász, On the ratio of optimal integral and fractional covers, *Discrete Math.* **13** (1975) 383–390.

[20] A.A. Sapozhenko, On the number of connected subsets with given cardinality of the boundary in bipartite graphs, *Metody Diskret. Analiz.* **45** (1987), 42–70. (Russian)

[21] A.A. Sapozhenko, The number of antichains in ranked partially ordered sets, *Diskret. Mat.* **1** (1989), 74–93. (Russian; translation in Discrete Math. Appl. 1 (1991), no. 1, 35–58

[22] A.A. Sapozhenko, The number of antichains in multilevelled ranked sets, *Diskret. Mat.* **1** (1989), 110–128. (Russian; translation in Discrete Math. Appl. 1 (1991), no. 2, 149–169.)

[23] S.K. Stein, Two combinatorial covering theorems, *J. Combinatorial Th. (A)* **16** (1974), 391–397.

[24] D. Wang and P. Wang, Discrete isoperimetric problems, *SIAM J. Appl. Math.* **32** (1977), 860–870.