Combinatorics, Probability and Computing (2006) 15, 253–279. © 2006 Cambridge University Press
doi:10.1017/S0963548305007182 Printed in the United Kingdom

On Dependency Graphs and the Lattice Gas

ALEXANDER D. SCOTT[1] and ALAN D. SOKAL[2]

[1]Department of Mathematics, University College London, London WC1E 6BT, England
(e-mail: scott@math.ucl.ac.uk)
[2]Department of Physics, New York University, 4 Washington Place, New York, NY 10003 USA
(e-mail: sokal@nyu.edu)

Received 8 October 2004; revised 24 June 2005

For Béla Bollobás on his 60th birthday

We elucidate the close connection between the repulsive lattice gas in equilibrium statistical mechanics and the Lovász Local Lemma in probabilistic combinatorics. We show that the conclusion of the Lovász Local Lemma holds for dependency graph G and probabilities $\{p_x\}$ if and only if the independent-set polynomial for G is nonvanishing in the polydisc of radii $\{p_x\}$. Furthermore, we show that the usual proof of the Lovász Local Lemma – which provides a sufficient condition for this to occur – corresponds to a simple inductive argument for the nonvanishing of the independent-set polynomial in a polydisc, which was discovered implicitly by Shearer [28] and explicitly by Dobrushin [12, 13]. We also present a generalization of the Lovász Local Lemma that allows for 'soft' dependencies. The paper aims to provide an accessible discussion of these results, which are drawn from a longer paper [26] that has appeared elsewhere.

1. Introduction

In probabilistic combinatorics, one is often faced with a collection of events $(A_x)_{x\in X}$ in some probability space, for which one wishes to prove that $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) > 0$, *i.e.*, that with positive probability none of the events happen. If the events are independent, then this is easily done. However, in practice there are usually some dependencies present, and so we must find some way to deal with them. One approach is to control the dependencies by a dependency graph: we say that a graph G with vertex set X is a *dependency graph* for the events $(A_x)_{x\in X}$ if, for each $x \in X$, the event A_x is independent from (the σ-algebra generated by) the collection $\{A_y : y \notin \Gamma^*(x)\}$, where $\Gamma^*(x) = \Gamma(x) \cup \{x\}$ is the closed neighbourhood of x.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Since we are interested in proving that $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) > 0$, the question is then how large the probabilities of the A_x can be while still being able to guarantee that $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) > 0$. More precisely, we have the following problem.

Problem 1.1. *Fix a graph G with vertex set X. For which sequences $\mathbf{p} = (p_x)_{x\in X} \in [0,1]^X$ is it true that, for every collection $(A_x)_{x\in X}$ of events with dependency graph G such that $\mathbb{P}(A_x) \leqslant p_x$ for all $x \in X$, we have $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) > 0$?*

We shall say that a sequence $\mathbf{p}$ with this property is *good* for the dependency graph G.

An ostensibly unrelated problem arises in statistical mechanics, in the context of the repulsive lattice gas. In its simplest form (we discuss more general versions later), the 'lattice-gas partition function' associated with the graph G (on vertex set X) is simply the multivariate generating polynomial for independent subsets of vertices, *i.e.*, the polynomial

$$Z_G(\mathbf{w}) = \sum_{\substack{X' \subseteq X \\ X' \text{ independent}}} \prod_{x\in X'} w_x, \tag{1.1}$$

where we associate a separate variable w_x (usually interpreted as a complex number) to each vertex $x \in X$. This polynomial is familiar in combinatorics, though usually in its single-variable form. (One of our contentions in this paper is that the multivariable form is quite natural, and often easier to analyse.) Since the empty set is trivially independent, we have $Z_G(\mathbf{0}) = 1$, so that Z_G is nonzero in some (complex) neighbourhood of the origin.

Mathematical physicists have devoted considerable effort to locating the zeros of the lattice-gas partition function, and in particular to finding complex polydiscs in which Z_G is nonvanishing.[1] For a sequence of radii $\mathbf{R} = (R_x)_{x\in X}$, let us define the closed polydisc $\overline{D}_{\mathbf{R}} = \{\mathbf{w} \in \mathbb{C}^X : \ |w_x| \leqslant R_x \ \forall x\}$. We then have the following problem.

Problem 1.2. *Fix a graph G with vertex set X. For which sequences $\mathbf{R} = (R_x)_{x\in X} \in [0,\infty)^X$ does the closed polydisc $\overline{D}_{\mathbf{R}}$ contain no zeros of Z_G?*

Although the two problems we have stated seem at first sight to be completely unrelated, they turn out to be closely connected. The main result that we will discuss in this paper is the following.

[1] Since the physically realizable values of w_x are *positive real*, the reader may be wondering why physicists (of all people!) would want to study the *complex* zeros. One reason is that the complex zeros of $Z_G(\mathbf{w})$ closest to the origin determine the radius of convergence of the Mayer expansion (2.14). A further reason is connected with the Yang–Lee [38] approach to studying phase transitions. Briefly: A *phase transition* is any point where one or more physical quantities depend nonanalytically on one or more control parameters. In any reasonable statistical-mechanical model (such as the polynomial (1.1)), such nonanalyticity is clearly impossible for finite graphs G; rather, phase transitions occur only in the *infinite-volume limit*. That is, we consider a countably infinite graph G_∞ (usually a regular lattice) and an increasing sequence of subgraphs $(G_n)_{n\geqslant 1}$ converging to G_∞, and study the *limiting free energy per unit volume* $f_{G_\infty}(\mathbf{w}) = \lim_{n\to\infty} |G_n|^{-1} \log Z_{G_n}(\mathbf{w})$. The nonanalyticities of $f_{G_\infty}(\mathbf{w})$ for real $\mathbf{w}$ arise from the singularities of $\log Z_{G_n}(\mathbf{w})$ for *complex* $\mathbf{w}$ that approach the real axis in the limit $n \to \infty$. But the singularities of $\log Z_{G_n}(\mathbf{w})$ are precisely the zeros of $Z_{G_n}(\mathbf{w})$. See, *e.g.*, [30, Section 1] or [31, Section 5] for a slightly more detailed explanation.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Theorem 1.3 (The equivalence theorem). *Let G be a finite graph with vertex set X, and let $\mathbf{p} = (p_x)_{x \in X} \in [0,1]^X$. Then the following two statements are equivalent:*

(a) $\mathbf{p}$ *is good for the dependency graph* G,
(b) *the closed polydisc* $\overline{D}_{\mathbf{p}}$ *contains no zeros of* Z_G.

This will follow from Theorem 3.1 below.

Physicists and mathematicians have each given sufficient conditions for a sequence of positive real numbers to have the properties specified in one (and hence both) of the problems above. For the repulsive lattice gas, Dobrushin [12, 13] gave a sufficient (but not necessary) condition on $\mathbf{R}$ for Z_G to be nonvanishing in $\overline{D}_{\mathbf{R}}$. (A slight generalization of Dobrushin's result is given below, as Theorem 4.3.) Likewise, the Lovász Local Lemma [14, 15] – which is an important tool in probabilistic combinatorics – gives a sufficient (but not necessary) condition for a sequence $\mathbf{p}$ to be good for a dependency graph G.

The equivalence between Problems 1.1 and 1.2 means that it is possible to compare the Lovász Local Lemma with Dobrushin's Theorem: both results give sufficient but not necessary conditions, and it is natural to wonder whether one of them might give stronger results than the other. Surprisingly, it turns out that the two theorems, proved in different fields and two decades apart (Dobrushin's Theorem in the 1990s and the Local Lemma in the 1970s), give *identical* criteria! Indeed, close examination of the (inductive) proofs shows that the two proofs are substantially isomorphic.

1.1. Outline of the paper

The main aim of this paper is to present the connections between dependency graphs and the lattice gas discovered in [26]. Although there are no new results in this paper, we hope that this selection of material from the (rather long) paper [26] will give a shorter and more accessible account of the topic aimed at a combinatorial audience. In order to simplify the presentation, we have omitted some of the proofs; detailed proofs of all the results in this paper, together with much more extensive discussion and many further results, can be found in [26], to which we refer the reader for further information.

In Section 2, we examine the lattice gas. After giving some basic properties and examples, we review the Mayer expansion, which is the expansion of $\log Z_G$ in a Taylor series around $\mathbf{w} = 0$. The crucial property that we shall exploit is the fact that the coefficients in the Mayer expansion have alternating signs (Proposition 2.1). We use this to prove (Theorem 2.2) that the closest zeros to the origin lie in the negative real quadrant $(-\infty, 0]^X$, along with some related facts.

In Section 3, we analyse the connection between dependency graphs and lattice gases, and prove the equivalence of Problems 1.1 and 1.2. (In fact, Theorem 3.1 asserts somewhat more than Theorem 1.3.) The proof of Theorem 3.1 has two key ingredients: first we use ideas of Shearer [28] to relate good sequences for a dependency graph G to the negative real zeros of the corresponding multivariate independent-set polynomial Z_G; then we use the results of Section 2 to relate the latter to the *complex* zeros of Z_G. These arguments can in fact be extended to the lattice gas with 'soft interactions', which are connected

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

to dependency graphs with a suitably defined notion of 'weak' dependence; we conclude Section 3 by explaining this connection.

In Section 4, we turn to results giving *sufficient* conditions for a sequence $\mathbf{p}$ or $\mathbf{R}$ to have the properties specified in Problems 1.1 and 1.2. After reviewing the Lovász Local Lemma and Dobrushin's Theorem, we prove the equivalence of the criteria given by these two results, and discuss some consequences, including a 'softened' version of the Lovász Local Lemma. We also discuss an improved bound, inspired by the work of Shearer [28].

We conclude the paper, in Section 5, with some comments on the Lovász/Dobrushin bounds.

2. The repulsive lattice gas

2.1. Definition

In statistical mechanics, a 'grand-canonical gas' is defined by a *single-particle state space* X (here a nonempty finite set), a *fugacity vector* $\mathbf{w} = \{w_x\}_{x \in X} \in \mathbb{C}^X$, and a *two-particle Boltzmann factor* $W : X \times X \to \mathbb{C}$ with $W(x, y) = W(y, x)$. The (*grand*) *partition function* $Z_W(\mathbf{w})$ is then defined to be the sum over ways of placing $n \geqslant 0$ 'particles' on 'sites' $x_1, \ldots, x_n \in X$, with each configuration assigned a 'Boltzmann weight' given by the product of the corresponding factors w_{x_i} and $W(x_i, x_j)$:

$$Z_W(\mathbf{w}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{x_1, \ldots, x_n \in X} \left(\prod_{i=1}^{n} w_{x_i} \right) \left(\prod_{1 \leqslant i < j \leqslant n} W(x_i, x_j) \right) \tag{2.1a}$$

$$= \sum_{\mathbf{n}} \left(\prod_{x \in X} \frac{w_x^{n_x} \, W(x, x)^{n_x(n_x-1)/2}}{n_x!} \right) \left(\prod_{\{x,y\} \subseteq X} W(x, y)^{n_x n_y} \right) \tag{2.1b}$$

where in (2.1b) the sum runs over all multi-indices $\mathbf{n} = \{n_x\}_{x \in X}$ of nonnegative integers, and the product runs over all two-element subsets $\{x, y\} \subseteq X$ $(x \neq y)$.[2] We shall use the notation $\mathbf{w}^{\mathbf{n}} = \prod_{x \in X} w_x^{n_x}$ and $|\mathbf{w}| = \{|w_x|\}_{x \in X}$ (although, abusing notation, we shall also write $|\mathbf{n}| = \sum_{x \in X} |n_x|$). We will also write $\mathbf{w} \geqslant \mathbf{0}$ to indicate that $\mathbf{w}$ is a vector of real numbers such that $w_x \geqslant 0$ for all $x \in X$.

In this paper we shall limit attention to the *repulsive* lattice gas in which $0 \leqslant W(x, y) \leqslant 1$ for all x, y. From this assumption it follows immediately that $Z_W(\mathbf{w})$ is an entire analytic function of $\mathbf{w}$ satisfying $|Z_W(\mathbf{w})| \leqslant \exp(\sum_{x \in X} |w_x|)$.

If $W(x, x) = 0$ for all $x \in X$ – in statistical mechanics this is called a *hard-core self-repulsion* – then the only nonvanishing terms in (2.1b) have $n_x = 0$ or 1 for all x (*i.e.*, each site can be occupied by at most one particle), so that $Z_W(\mathbf{w})$ can be written as a sum over

[2] More precisely, $Z_W(\mathbf{w})$ sums $1/n!$ times the Boltzmann weight over ways of placing n *distinguishable* particles onto sites. In the case of hard-core self-repulsion (which is the case of principal interest to us), this is equivalent to summing the Boltzmann weight over ways of placing n *indistinguishable* particles onto sites. (For the two interpretations to be equivalent when multiple occupation is permitted, there would have to be an additional factor $\prod_{x \in X} n_x!$ in (2.1a) and (2.1b).)

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

subsets:

$$Z_W(\mathbf{w}) = \sum_{X' \subseteq X} \left(\prod_{x \in X'} w_x \right) \left(\prod_{\{x,y\} \subseteq X'} W(x,y) \right). \tag{2.2}$$

In this case $Z_W(\mathbf{w})$ is a multiaffine polynomial, *i.e.*, of degree 1 in each w_x separately. Combinatorially, $Z_W(\mathbf{w})$ is the generating polynomial for *induced subgraphs* of the complete graph, in which each vertex x gets weight w_x and each edge xy gets weight $W(x,y)$.

If, in addition to hard-core self-repulsion, we have $W(x,y) = 0$ or 1 for each pair $x \neq y$ – in statistical mechanics this is called a *hard-core pair interaction* – then we can define a (simple loopless) graph $G = (X, E)$ by setting $xy \in E$ whenever $W(x,y) = 0$ and $x \neq y$, so that $Z_W(\mathbf{w})$ is precisely the *independent-set polynomial* for G:

$$Z_G(\mathbf{w}) = \sum_{\substack{X' \subseteq X \\ X' \text{ independent}}} \prod_{x \in X'} w_x. \tag{2.3}$$

Traditionally the independent-set polynomial is defined as a univariate polynomial $Z_G(w)$ in which w_x is set equal to the same value w for all vertices x. But one of our main contentions in this paper is that Z_G is more naturally understood as a multivariate polynomial; this allows us, in particular, to exploit the fact that Z_G is multiaffine.

More generally, given any W satisfying $0 \leqslant W(x,y) \leqslant 1$ for all x, y, let us define a simple loopless graph $G = G_W$ (the *support graph* of W) by setting $xy \in E(G)$ if and only if $W(x,y) \neq 1$ and $x \neq y$. The partition function $Z_W(\mathbf{w})$ can be thought of as a 'soft' version of the independent-set polynomial for G, in which an edge $xy \in E(G)$ has 'strength' $1 - W(x,y) \in (0,1]$.

In the situation of hard-core self-repulsion (2.2), it is convenient to define, for each subset $\Lambda \subseteq X$, the restricted partition function

$$Z_\Lambda(\mathbf{w}) = \sum_{X' \subseteq \Lambda} \prod_{x \in X'} w_x \prod_{\{x,y\} \subseteq X'} W(x,y). \tag{2.4}$$

Of course this notation is redundant, since the same effect can be obtained by setting $w_x = 0$ for $x \in X \setminus \Lambda$, but it is useful for the purpose of inductive computations and proofs. We have, for any $x \in \Lambda$, the *fundamental identity*

$$Z_\Lambda(\mathbf{w}) = Z_{\Lambda \setminus x}(\mathbf{w}) + w_x Z_{\Lambda \setminus x}(W(x,\cdot)\mathbf{w}) \tag{2.5}$$

where

$$[W(x,\cdot)\mathbf{w}]_y = W(x,y)\, w_y; \tag{2.6}$$

here the first term on the right-hand side of (2.5) covers the summands in (2.4) with $X' \not\ni x$, while the second covers $X' \ni x$. In the special case of a hard-core interaction (= independent-set polynomial) for a graph G, (2.5) reduces to

$$Z_\Lambda(\mathbf{w}) = Z_{\Lambda \setminus x}(\mathbf{w}) + w_x Z_{\Lambda \setminus \Gamma^*(x)}(\mathbf{w}), \tag{2.7}$$

where we have used the notation $\Gamma^*(x) = \Gamma(x) \cup \{x\}$. The fundamental identity (2.5)/(2.7) plays an important role both in the inductive proof of the Lovász Local Lemma and

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

in the Dobrushin–Shearer inductive argument for the nonvanishing of Z_W in a polydisc (Section 4).

Remarks. **(1)** It is worth remarking that the hard-core lattice gas (2.3) is not merely *one* interesting statistical-mechanical model; it turns out to be the *universal* statistical-mechanical model in the sense that any statistical-mechanical model living on a vertex set V_0 can be mapped onto a gas of nonoverlapping 'polymers' on V_0, *i.e.*, a hard-core lattice gas on the intersection graph of V_0 [29, Section 5.7].[3] This construction, which is termed the 'polymer expansion' or 'cluster expansion', is an important tool in mathematical statistical mechanics [27, 7, 17, 9, 6]; it is widely employed to prove the absence of phase transition at high temperature, low temperature, large magnetic field, low density, or weak nonlinear coupling. Further information on expansion methods in statistical mechanics (and related combinatorial problems) can be found in the excellent recent survey by Borgs [6].

(2) Repeated use of (2.5) obviously gives an algorithm to compute $Z_W(\mathbf{w})$. But this algorithm takes in general exponential time. In fact, calculating $Z_G(\mathbf{w})$ for general graphs G (or even for cubic planar graphs) is NP-hard (as noted by Shearer [28]), since even calculating the *degree* of $Z_G(\mathbf{w})$ – that is, the maximum size of an independent set – is NP-hard [16, pp. 194–195]. Therefore, if P $\neq$ NP it is impossible to calculate $Z_G(\mathbf{w})$ for general graphs in polynomial time.

In this section we give some general results concerning the partition function of a lattice gas; additional related results can be found in [26]. Most of these results are valid for an arbitrary repulsive lattice gas (2.1), in which multiple occupation of a site is permitted. A few of the results are restricted to the case of a hard-core self-repulsion (2.2), in which multiple occupation of a site is forbidden.

2.2. An example: The complete r-ary rooted tree [25, 28]

Before continuing with the general theory, we pause to compute an important example. Let $T_n^{(r)}$ be the complete rooted tree with branching factor r and depth n. We limit attention to the univariate independent-set polynomial. Fix $r \geqslant 1$; and to lighten the notation, let us write Z_n as a shorthand for $Z_{T_n^{(r)}}$. Applying the fundamental identity (2.7) to the root vertex, we obtain the nonlinear recursion

$$Z_n(w) = Z_{n-1}(w)^r + wZ_{n-2}(w)^{r^2}, \tag{2.8}$$

which is valid for all $n \geqslant 0$ if we set $Z_{-1} \equiv Z_{-2} \equiv 1$. By defining

$$Y_n(w) = \frac{Z_n(w)}{Z_{n-1}(w)^r}, \tag{2.9}$$

[3] The *intersection graph* of a finite set S is the graph whose vertices are the nonempty subsets of S, and whose edges are the pairs with nonempty intersection.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

we can convert the second-order recursion (2.8) to a first-order recursion

$$Y_n(w) = 1 + \frac{w}{Y_{n-1}(w)^r} \tag{2.10}$$

with initial condition $Y_{-1} \equiv 1$. The polynomials $Z_n(w)$ can be reconstructed from the rational functions $Y_n(w)$ by

$$Z_n(w) = \prod_{k=0}^{n} Y_k(w)^{r^{n-k}}. \tag{2.11}$$

Let $w_n < 0$ be the negative real root of Z_n of smallest magnitude (set $w_n = -\infty$ if Z_n has no negative real root). Note that $w_{-1} = -\infty$ and $w_0 = -1$. Since $Z_n(0) = 1$, we have $Z_n(w) > 0$ for all $w \in (w_n, 0]$. Let us prove by induction that $w_{n-1} < w_n$ for $n \geqslant 0$. It is true for $n = 0$. For $n \geqslant 1$ we have

$$Z_n(w_{n-1}) = Z_{n-1}(w_{n-1})^r + w_{n-1} Z_{n-2}(w_{n-1})^{r^2} < 0 \tag{2.12}$$

since $Z_{n-1}(w_{n-1}) = 0$, $w_{n-1} < 0$ and $Z_{n-2}(w_{n-1}) > 0$ by the inductive hypothesis. Therefore Z_n vanishes somewhere between w_{n-1} and 0.

It follows that the w_n increase to a limit $w_\infty \leqslant 0$ as $n \to \infty$. Let us show, following Shearer [28], that

$$w_\infty = -\frac{r^r}{(r+1)^{r+1}} \tag{2.13}$$

by proving the two inequalities.

Proof of $\geqslant$. If $w \in [w_\infty, 0)$, we have $Z_n(w) > 0$ for all n and hence also $Y_n(w) > 0$ for all n. Since $Y_{-1} > Y_0$, it follows from the monotonicity of (2.10) that $\{Y_n(w)\}_{n \geqslant 0}$ is a strictly decreasing sequence of positive numbers, hence converges to a limit $y_* \geqslant 0$ satisfying the fixed-point equation $y_* = 1 + w/y_*^r$, or equivalently $w = y_*^{r+1} - y_*^r$. Elementary calculus then shows that $w \geqslant -r^r/(r+1)^{r+1}$; taking $w = w_\infty$ we obtain $w_\infty \geqslant -r^r/(r+1)^{r+1}$. □

Proof of $\leqslant$. If $-r^r/(r+1)^{r+1} \leqslant w < 0$, the equation $w = y_*^{r+1} - y_*^r$ has a unique solution $y_* \in [r/(r+1), 1)$. It then follows by induction (using (2.10) and the initial condition $Y_{-1} = 1$) that $1 = Y_{-1}(w) > Y_0(w) > \cdots > Y_{n-1}(w) > Y_n(w) > \cdots > y_*$ for all $n \geqslant 0$. In particular, $Y_n(w) > 0$ for all n, so that $w_n < w$ for all n. This shows that $w_\infty \leqslant -r^r/(r+1)^{r+1}$. □

Let us conclude by observing that (2.10) defines a degree-r rational map $R_w : y \mapsto 1 + w/y^r$ parametrized by $w \in \mathbb{C} \setminus 0$. Moreover, the zeros of $Z_n(w)$ correspond to those values w for which R_w has a (superattractive) orbit $0 \mapsto \infty \mapsto 1 \mapsto 1 + w \mapsto \cdots \mapsto 0$ of period $n + 3$ (or some divisor of $n + 3$). As $n \to \infty$, these points accumulate on a 'Mandelbrot-like' set in the complex w-plane [22].

2.3. The Mayer expansion

Let us now return to the general case of a repulsive lattice gas (2.1). Since $Z_W(\mathbf{w})$ is an entire function of $\mathbf{w}$ satisfying $Z_W(\mathbf{0}) = 1$, its logarithm is analytic in some neighbourhood

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

of $\mathbf{w} = \mathbf{0}$, and so can be expanded in a convergent Taylor series:

$$\log Z_W(\mathbf{w}) = \sum_{\mathbf{n}} c_{\mathbf{n}}(W)\, \mathbf{w}^{\mathbf{n}}, \tag{2.14}$$

where we have used the notation $\mathbf{w}^{\mathbf{n}} = \prod_{x \in X} w_x^{n_x}$, and of course $c_{\mathbf{0}} = 0$. In statistical mechanics, (2.14) is called the *Mayer expansion* [36, 6]; there is a beautiful combinatorial formula for the Mayer coefficients $c_{\mathbf{n}}(W)$, which we shall not need here (see [36, 6, 26]).

For our purposes, the crucial property of the Mayer expansion is the following.

Proposition 2.1 (Signs of Mayer coefficients). *Suppose that the lattice gas is repulsive, i.e.,* $0 \leqslant W(x, y) \leqslant 1$ *for all* $x, y \in X$. *Then, for all* $\mathbf{n} \geqslant \mathbf{0}$, *the Mayer coefficients* $c_{\mathbf{n}}(W)$ *satisfy*

$$(-1)^{|\mathbf{n}|-1}\, c_{\mathbf{n}}(W) \geqslant 0. \tag{2.15}$$

The alternating-sign property (2.15) for the Mayer coefficients of a repulsive gas has been known in the physics literature for over 40 years: see Groeneveld [18] for a brief sketch of one proof. Our own proof [26, Section 2.2], which is based on the partitionability of a matroid complex, also controls the signs of the first two derivatives of $c_{\mathbf{n}}(W)$ with respect to W. We think that the Mayer coefficients $c_{\mathbf{n}}(W)$ merit further study from a combinatorial point of view; we would not be surprised if new identities or inequalities were waiting to be discovered.

2.4. The fundamental theorem

Let us now state the principal result of this section. We use the notation $|\mathbf{w}| = \{|w_x|\}_{x \in X}$.

Theorem 2.2 (The fundamental theorem). *Consider any repulsive lattice gas, and let* $\mathbf{R} = \{R_x\}_{x \in X} \geqslant 0$. *Then the following are equivalent.*

(a) *There exists a connected set* $C \subseteq (-\infty, 0]^X$ *that contains both* $\mathbf{0}$ *and* $-\mathbf{R}$, *such that* $Z_W(\mathbf{w}) > 0$ *for all* $\mathbf{w} \in C$. (*Equivalently,* $-\mathbf{R}$ *belongs to the connected component of* $Z_W^{-1}(0, \infty) \cap (-\infty, 0]^X$ *containing* $\mathbf{0}$.)
(b) $Z_W(\mathbf{w}) > 0$ *for all* $\mathbf{w} \in \mathbb{R}^X$ *satisfying* $-\mathbf{R} \leqslant \mathbf{w} \leqslant \mathbf{0}$.
(c) $Z_W(\mathbf{w}) \neq 0$ *for all* $\mathbf{w} \in \mathbb{C}^X$ *satisfying* $|\mathbf{w}| \leqslant \mathbf{R}$.
(d) *The Taylor series for* $\log Z_W(\mathbf{w})$ *around* $\mathbf{0}$ *is convergent at* $\mathbf{w} = -\mathbf{R}$.
(e) *The Taylor series for* $\log Z_W(\mathbf{w})$ *around* $\mathbf{0}$ *is absolutely convergent for* $|\mathbf{w}| \leqslant \mathbf{R}$.

Moreover, when these conditions hold, we have $|Z_W(\mathbf{w})| \geqslant Z_W(-\mathbf{R}) > 0$ *for all* $\mathbf{w} \in \mathbb{C}^X$ *satisfying* $|\mathbf{w}| \leqslant \mathbf{R}$.

In the case of hard-core self-repulsion, (a)–(e) *are also equivalent to the following.*

(b′) $Z_W(-\mathbf{R}\ \mathbf{1}_S) > 0$ *for all* $S \subseteq X$, *where*

$$(\mathbf{R}\, \mathbf{1}_S)_x = \begin{cases} R_x & \text{if } x \in S, \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

(f) $Z_W(-\mathbf{R}) > 0$, *and* $(-1)^{|S|} Z_W(-\mathbf{R}; S) \geqslant 0$ *for all* $S \subseteq X$, *where*

$$Z_W(\mathbf{w}; S) = \sum_{S \subseteq X' \subseteq X} \left(\prod_{x \in X'} w_x \right) \left(\prod_{\{x,y\} \subseteq X'} W(x, y) \right). \tag{2.17}$$

(g) *There exists a probability measure* P *on* 2^X *satisfying* $P(\varnothing) > 0$ *and*

$$\sum_{T \supseteq S} P(T) = \left(\prod_{x \in S} R_x \right) \left(\prod_{\{x,y\} \subseteq S} W(x, y) \right) \tag{2.18}$$

for all $S \subseteq X$. (*This probability measure is unique and is given by*

$$P(S) = (-1)^{|S|} Z_W(-\mathbf{R}; S).$$

In particular, $P(\varnothing) = Z_W(-\mathbf{R}) > 0$.)

Remarks. (1) The conditions (b′), (f) and (g) are inspired in part by Shearer [28, Theorem 1].

(2) Suppose that the univariate entire function $Z_W(w)$, defined by setting $w_x = w$ for all x, is strictly positive whenever $-R \leqslant w \leqslant 0$. Then in fact $Z_W(\mathbf{w}) > 0$ whenever $-R \leqslant w_x \leqslant 0$ for all x: this follows from (a) $\Longrightarrow$ (b) by taking C to be the segment $[-R, 0]$ of the diagonal.

Our proof of Theorem 2.2 hinges on the alternating-sign property (2.15) for the Taylor coefficients of $\log Z_W$. In preparation for this proof, let us recall the Vivanti–Pringsheim theorem in the theory of analytic functions of a single complex variable [19, Theorem 5.7.1]: if a power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with *nonnegative* coefficients has a finite nonzero radius of convergence, then the point of the circle of convergence lying on the positive real axis is a singular point of the function f. Otherwise put, if f is a function whose Taylor series at 0 has all nonnegative coefficients and f is analytic on some complex neighbourhood of the real interval $[0, R)$, then f is in fact analytic on the open disc of radius R centred at the origin and its Taylor series is absolutely convergent there. Here we will need the following multidimensional generalization [26] of the Vivanti–Pringsheim theorem.

Proposition 2.3 (Multidimensional Vivanti–Pringsheim theorem). *Let* C *be a connected subset of* $[0, \infty)^n$ *containing* $\mathbf{0}$, *let* U *be an open neighbourhood of* C *in* $\mathbb{C}^n$, *and let* f *be a function analytic on* U *whose Taylor series around* $\mathbf{0}$ *has all nonnegative coefficients. Then the Taylor series of* f *around* $\mathbf{0}$ *converges absolutely on the set* $\mathrm{hull}(C) \equiv \bigcup_{\mathbf{R} \in C} \bar{D}_{\mathbf{R}}$, *where* $\bar{D}_{\mathbf{R}}$ *denotes the closed polydisc* $\{\mathbf{w} \in \mathbb{C}^n : |w_i| \leqslant R_i \text{ for all } i\}$, *and it defines a function that is continuous on* $\mathrm{hull}(C)$ *and analytic on its interior.*

We shall also make use of the following elementary result.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Lemma 2.4. *Let F be a function on 2^X, and define*

$$F_-(S) = \sum_{X' \subseteq S} F(X'), \tag{2.19}$$

$$F_+(S) = \sum_{X' \supseteq S} F(X'). \tag{2.20}$$

Then

$$F_-(S) = \sum_{Y \subseteq S^c} (-1)^{|Y|} F_+(Y) \tag{2.21}$$

where $S^c \equiv X \setminus S$.

Finally, we shall need that fact that if $F : \mathbb{C}^X \to \mathbb{C}$ is multiaffine on a rectangle, then its value at any point inside the rectangle is a convex combination of the values at extreme points of the rectangle. For instance, if F is multiaffine on the rectangle $[0,1]^X$, then

$$F(\mathbf{z}) = \sum_{Y \subseteq X} \left(\prod_{i \in Y} z_i \right) \left(\prod_{i \in X \setminus Y} (1 - z_i) \right) F(\mathbf{1}_Y). \tag{2.22}$$

This is clear when $\mathbf{z} \in \{0,1\}^X$, and equality for all $\mathbf{z} \in [0,1]^X$ then follows because both sides are multiaffine. A similar statement holds for general rectangles.

We are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. (c) $\Longrightarrow$ (b) $\Longrightarrow$ (a) is trivial (note that $Z_W(\mathbf{0}) = 1$).

(e) $\Longrightarrow$ (d) is trivial. The alternating sign property (2.15) implies that all terms $c_{\mathbf{n}} \mathbf{w}^{\mathbf{n}}$ in the Mayer expansion (2.14) are nonpositive when $\mathbf{w} \leqslant \mathbf{0}$, and so we get (d) $\Longrightarrow$ (e).

(e) implies that the sum of the Taylor series for $\log Z_W(\mathbf{w})$ defines an analytic function on the open polydisc $D_{\mathbf{R}}$ and a continuous function on the closed polydisc $\bar{D}_{\mathbf{R}}$. Its exponential equals $Z_W(\mathbf{w})$ on $D_{\mathbf{R}}$ and hence by continuity also on $\bar{D}_{\mathbf{R}}$. Therefore (e) $\Longrightarrow$ (c).

Finally, assume (a). Since Z_W is continuous on $\mathbb{C}^X$ (and has real coefficients), we can find an open connected neighbourhood C' of C in $(-\infty, 0]^X$ on which $Z_W > 0$, and an open neighbourhood U of C' in $\mathbb{C}^X \simeq \mathbb{R}^{2|X|}$ on which $Z_W \neq 0$. It is fairly easy to show that we can find a finite polygonal path $P \subset C'$ from $\mathbf{0}$ to $-\mathbf{R}$ (consider the set of points in C' that can be reached by a finite polygonal path from $\mathbf{0}$); taking a suitably small neighbourhood of P gives a simply connected open set U' in $\mathbb{C}^X$ with $P \subset U' \subset U$ (see [26] for details). Then $\log Z_W$ is a well-defined single-valued analytic function on U', once we specify $\log Z_W(\mathbf{0}) = 0$. Applying Proposition 2.3 to $\log Z_W$ on P and U' (using the alternating-sign property (2.15)), we conclude that the Taylor series for $\log Z_W$ around $\mathbf{0}$ is absolutely convergent on $\bar{D}_{\mathbf{R}}$. Therefore (a) $\Longrightarrow$ (e).

The bound $|Z_W(\mathbf{w})| \geqslant Z_W(-\mathbf{R})$ for $|\mathbf{w}| \leqslant \mathbf{R}$, which is equivalent to $\operatorname{Re} \log Z_W(\mathbf{w}) \geqslant \log Z_W(-\mathbf{R})$, is an immediate consequence of the alternating-sign property (2.15).

Now consider the special case of a hard-core self-repulsion. (b) $\Longrightarrow$ (b′) is trivial, and (b′) $\Longrightarrow$ (b) follows from the fact that Z_W is multiaffine (*i.e.*, of degree $\leqslant 1$ in each w_x separately) because the value of Z_W at any point $\mathbf{w}$ in the rectangle $-\mathbf{R} \leqslant \mathbf{w} \leqslant \mathbf{0}$ is a

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

convex combination of the values at extreme points of the rectangle, which correspond to possible choices of $S \subset X$ in (2.16).

To show that (b) $\Longrightarrow$ (f), note that

$$Z_W(\mathbf{w}; S) = \left(\prod_{x \in S} w_x\right) \left(\prod_{\{x,y\} \subseteq S} W(x,y)\right) Z_W(W(S,\cdot)\mathbf{w}), \tag{2.23}$$

where we have defined

$$[W(S,\cdot)\mathbf{w}]_y = \left(\prod_{x \in S} W(x,y)\right) w_y \tag{2.24}$$

(note in particular that this vanishes whenever $y \in S$). Hence

$$(-1)^{|S|} Z_W(-\mathbf{R}; S) = \left(\prod_{x \in S} R_x\right) \left(\prod_{\{x,y\} \subseteq S} W(x,y)\right) Z_W(-W(S,\cdot)\mathbf{R}) \geqslant 0 \tag{2.25}$$

since $-\mathbf{R} \leqslant -W(S,\cdot)\mathbf{R} \leqslant \mathbf{0}$, with strict inequality when $|S| = 0$ or 1 (since the product over $W(x,y)$ is in that case empty).

To show that (f) $\Longrightarrow$ (b′), use Lemma 2.4 applied to the set function

$$F(S) = \left(\prod_{x \in S} -R_x\right) \left(\prod_{\{x,y\} \subseteq S} W(x,y)\right). \tag{2.26}$$

We have

$$F_-(S) = Z_W(-\mathbf{R}\,\mathbf{1}_S), \tag{2.27}$$

$$F_+(S) = Z_W(-\mathbf{R}; S), \tag{2.28}$$

so that Lemma 2.4 asserts the identity

$$Z_W(-\mathbf{R}\,\mathbf{1}_S) = \sum_{Y \subseteq S^c} (-1)^{|Y|} Z_W(-\mathbf{R}; Y). \tag{2.29}$$

By (f), the $Y = \varnothing$ term is > 0 and the other terms are $\geqslant 0$, so $Z_W(-\mathbf{R}\,\mathbf{1}_S) > 0$ for all S.

Finally, let us show that (f) $\Longleftrightarrow$ (g). By inclusion-exclusion, there are unique numbers $P(T)$ satisfying (2.18), namely $P(T) = (-1)^{|T|} Z_W(-\mathbf{R}; T)$. Moreover, taking $S = \varnothing$ in (2.18) we see that $\sum_T P(T) = 1$. Therefore, P is a probability measure if and only if $(-1)^{|T|} Z_W(-\mathbf{R}; T) \geqslant 0$ for all T; and $P(\varnothing) > 0$ if and only if $Z_W(-\mathbf{R}; \varnothing) = Z_W(-\mathbf{R}) > 0$. □

3. Dependency graphs and the lattice gas: The equivalence theorem

In this section, we begin our investigation of the relationship between dependency graphs and the lattice gas. In Section 3.1, we work with the lattice gas with hard-core pair interactions, which has partition function given by (2.3); in Section 3.2, we extend the results to the lattice gas with soft-core pair interactions, which has partition function given by (2.2).

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

3.1. Hard-core version

We begin by recalling the definition of a dependency graph. Let $(A_x)_{x\in X}$ be a finite family of events on some probability space, and let G be a graph with vertex set X. We say that G is a *dependency graph* for the family $(A_x)_{x\in X}$ if, for each $x \in X$, the event A_x is independent from the σ-algebra $\sigma(A_y : \ y \in X \setminus \Gamma^*(x))$. Note that this is much stronger than requiring merely that A_x be independent of each such A_y separately.

A family of events typically has many possible dependency graphs: for instance, if G is a dependency graph for events $(A_x)_{x\in X}$, then any graph obtained by adding edges to G is also a dependency graph. In particular, if the events A_x are independent, then any graph on X is a dependency graph. Nor must there be a unique minimal dependency graph. Consider, for instance, the set of binary strings of length n with odd digit sum (giving each such string equal probability), and let A_i be the event that the ith digit is 1. Any graph without isolated vertices is a dependency graph for this collection of events.

There is also a stronger notion of a dependency graph G for a collection of events $(A_x)_{x\in X}$, where we demand that if Y and Z are disjoint subsets of X such that G contains no edges between Y and Z, then the σ-algebras $\sigma(A_y : \ y \in Y)$ and $\sigma(A_z : \ z \in Z)$ are independent. In this case we shall refer to G as a *strong dependency graph* for the events $(A_x)_{x\in X}$. (For instance, this situation arises in any statistical-mechanical model with variables living on the set X and pair interactions only on the edges of G, where each A_x depends only on the variable at x.) Alternatively, the dependency-graph hypothesis can be replaced by a *weaker* hypothesis concerning conditional probabilities, as in the lopsided Lovász Local Lemma (Theorem 4.2). It will follow from Theorem 3.1 below that all three hypotheses lead to the same lower bound on $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x)$.

Our aim is to relate dependency graphs to lattice gases. The following result (which is a development of Shearer [28, Theorem 1]) gives the connection. For a graph G, we define

$$\mathscr{R}(G) = \{\mathbf{R} \in [0,\infty)^X : \ Z_G(\mathbf{w}) \neq 0 \ \forall \mathbf{w} \in \overline{D}_{\mathbf{R}}\}. \tag{3.1}$$

Theorem 3.1 (The equivalence theorem, hard-core case). *Let $(A_x)_{x\in X}$ be a family of events on some probability space, and let G be a graph with vertex set X. Suppose that $(p_x)_{x\in X}$ are real numbers in $[0,1]$ such that, for each x and each $Y \subseteq X \setminus \Gamma^*(x)$, we have*

$$\mathbb{P}\left(A_x \Big| \bigcap_{y\in Y} \overline{A}_y\right) \leqslant p_x. \tag{3.2}$$

(a) *If $\mathbf{p} \in \mathscr{R}(G)$, then*

$$\mathbb{P}\left(\bigcap_{x\in X} \overline{A}_x\right) \geqslant Z_G(-\mathbf{p}) > 0 \tag{3.3}$$

and more generally

$$\mathbb{P}\left(\bigcap_{x\in Y} \overline{A}_x \Big| \bigcap_{x\in Z} \overline{A}_x\right) \geqslant \frac{Z_G(-\mathbf{p}\,\mathbf{1}_{Y\cup Z})}{Z_G(-\mathbf{p}\,\mathbf{1}_Z)} > 0 \tag{3.4}$$

for any subsets $Y, Z \subseteq X$. Moreover, this lower bound is best possible in the sense that there exists a probability space on which there can be constructed a family of

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

events $(B_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B_x) = p_x$ *and strong dependency graph* G*, such that* $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x) = Z_G(-\mathbf{p})$.

(b) *If* $\mathbf{p} \notin \mathscr{R}(G)$*, then there exists a probability space on which there can be constructed:*

(i) *a family of events* $(B_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B_x) = p_x$ *and strong dependency graph* G*, satisfying* $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x) = 0$*; and*

(ii) *a family of events* $(B'_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B'_x) = p'_x \leqslant p_x$ *and strong dependency graph* G*, satisfying* $\mathbb{P}(B'_x \cap B'_y) = 0$ *for all* $xy \in E(G)$ *and* $\mathbb{P}(\bigcap_{x\in X} \overline{B}'_x) = 0$.

Remarks. **(1)** Please note that G is here an *arbitrary* graph with vertex set X; it need not be a dependency graph for the events $(A_x)_{x\in X}$. Rather, given G, we can regard $\mathbf{p}$ as *defined* by

$$p_x = \max_{Y\subseteq X\setminus\Gamma^*(x)} \mathbb{P}\left(A_x \Big| \bigcap_{y\in Y} \overline{A}_y\right) \tag{3.5}$$

(this is clearly the minimal choice). There is then a tradeoff in the choice of G: adding more edges reduces p_x (since there are fewer conditional probabilities to control) but also shrinks the set $\mathscr{R}(G)$ (see [26]).

(2) Though (3.2) is the *weak* hypothesis of the lopsided Lovász Local Lemma (Theorem 4.2), we will prove in (a) and (b) that the extremal families $(B_x)_{x\in X}$ and $(B'_x)_{x\in X}$ have G as a *strong* dependency graph. Therefore, all three dependency hypotheses lead to the same optimal lower bound on $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x)$.

(3) The proofs given here of Theorems 3.1 and 3.2 are logically independent of nearly all of Theorem 2.2. More precisely, if we were to define $\mathscr{R}(G)$ by condition (b) of Theorem 2.2, then the only part of Theorem 2.2 that is used in the proofs of Theorems 3.1 and 3.2 is the (relatively easy) implication (b) $\Longrightarrow$ (f). But we have chosen to define $\mathscr{R}(G)$ instead by condition (c), in order to emphasize the connection with the *complex* zeros of the partition function.

Proof. For $\mathbf{p} \in \mathscr{R}(G)$, we wish to define a family of events $(B_x)_{x\in X}$ (on a new probability space) such that the hypotheses of the theorem are satisfied and $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x)$ is as small as possible. An intuitively reasonable way to do this is to make the events B_x as disjoint as possible, consistent with the condition (3.2) (or with either of the two stronger notions of dependency graph). With this in mind, for $\Lambda \subseteq X$ let us define

$$\mathbb{P}\left(\bigcap_{x\in\Lambda} B_x\right) = \begin{cases} \prod_{x\in\Lambda} p_x & \text{if } \Lambda \text{ is independent in } G, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

This defines a signed measure on the σ-algebra generated by $(B_x)_{x\in X}$; indeed, inclusion-exclusion gives

$$\mathbb{P}\left(\bigcap_{x\in\Lambda} B_x \cap \bigcap_{x\notin\Lambda} \overline{B}_x\right) = \sum_{I\supseteq\Lambda} (-1)^{|I|-|\Lambda|}\, \mathbb{P}\left(\bigcap_{x\in I} B_x\right) \tag{3.7a}$$

$$= \sum_{I\supseteq\Lambda,\, I \text{ independent}} (-1)^{|I|-|\Lambda|} \prod_{x\in I} p_x \tag{3.7b}$$

$$= (-1)^{|\Lambda|}\, Z_G(-\mathbf{p};\Lambda), \tag{3.7c}$$

where $Z_G(-\mathbf{p};\Lambda)$ is defined as in (2.17). In particular, taking $\Lambda = \varnothing$, we have $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x) = Z_G(-\mathbf{p})$. Now since $\mathbf{p} \in \mathscr{R}(G)$, condition (c) (and hence all the conditions) of Theorem 2.2 is satisfied. Thus Theorem 2.2(f) implies that (3.7c) is nonnegative for all Λ, so that (3.6) defines a probability measure on $\sigma(B_x : x \in X)$. (This is the probability measure defined in Theorem 2.2(g).)

If Y and Z are disjoint subsets of X such that G contains no edges between Y and Z, it follows from (3.6) that for $Y_0 \subseteq Y$ and $Z_0 \subseteq Z$ the events $\bigcap_{x\in Y_0} B_x$ and $\bigcap_{x\in Z_0} B_x$ are independent. This implies (see, for instance, [37, Theorem 4.2] or [4, Theorem 4.2]) that $\sigma(B_x : x \in Y)$ and $\sigma(B_x : x \in Z)$ are independent, and so G is a strong dependency graph.

We next show that $(B_x)_{x\in X}$ is a family minimizing $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x)$. For $\Lambda \subseteq X$, we define

$$P_\Lambda = \mathbb{P}\left(\bigcap_{x\in\Lambda} \overline{A}_x\right), \tag{3.8}$$

$$Q_\Lambda = \mathbb{P}\left(\bigcap_{x\in\Lambda} \overline{B}_x\right). \tag{3.9}$$

Let us now prove by induction on $|\Lambda|$ that P_Λ/Q_Λ is monotone increasing in Λ. Note first that by inclusion-exclusion,

$$Q_\Lambda = \sum_{I\subseteq\Lambda} (-1)^{|I|}\, \mathbb{P}\left(\bigcap_{x\in I} B_x\right) \tag{3.10a}$$

$$= \sum_{I\subseteq\Lambda,\, I \text{ independent}} (-1)^{|I|} \prod_{x\in I} p_x \tag{3.10b}$$

$$= Z_G(-\mathbf{p}\,\mathbf{1}_\Lambda). \tag{3.10c}$$

Thus $Q_\Lambda > 0$ for all Λ, since $\mathbf{p} \in \mathscr{R}(G)$ and $\mathscr{R}(G)$ is a down-set. Furthermore, for $y \notin \Lambda$,

$$Q_{\Lambda\cup\{y\}} = \sum_{I\subseteq\Lambda\cup\{y\},\, I \text{ independent}} (-1)^{|I|} \prod_{x\in I} p_x \tag{3.11a}$$

$$= Q_\Lambda - p_y \sum_{I\subseteq\Lambda\setminus\Gamma(y),\, I \text{ independent}} (-1)^{|I|} \prod_{x\in I} p_x \tag{3.11b}$$

$$= Q_\Lambda - p_y Q_{\Lambda\setminus\Gamma(y)}. \tag{3.11c}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

(Note that this is just the fundamental identity (2.7) applied to $Z_G(-\mathbf{p}\,\mathbf{1}_\Lambda)$.) On the other hand,

$$P_{\Lambda\cup\{y\}} = P_\Lambda - \mathbb{P}\left(A_y \cap \bigcap_{x\in\Lambda} \overline{A}_x\right) \tag{3.12a}$$

$$\geqslant P_\Lambda - \mathbb{P}\left(A_y \cap \bigcap_{x\in\Lambda\setminus\Gamma(y)} \overline{A}_x\right) \tag{3.12b}$$

$$\geqslant P_\Lambda - p_y P_{\Lambda\setminus\Gamma(y)} \tag{3.12c}$$

by the hypothesis (3.2). Now we want to show that $P_{\Lambda\cup\{y\}}/Q_{\Lambda\cup\{y\}} \geqslant P_\Lambda/Q_\Lambda$, or equivalently that $P_{\Lambda\cup\{y\}}Q_\Lambda - Q_{\Lambda\cup\{y\}}P_\Lambda \geqslant 0$. By (3.11) and (3.12) we have

$$P_{\Lambda\cup\{y\}}Q_\Lambda - Q_{\Lambda\cup\{y\}}P_\Lambda \geqslant [P_\Lambda - p_y P_{\Lambda\setminus\Gamma(y)}]Q_\Lambda - [Q_\Lambda - p_y Q_{\Lambda\setminus\Gamma(y)}]P_\Lambda \tag{3.13a}$$

$$= p_y\,[P_\Lambda Q_{\Lambda\setminus\Gamma(y)} - Q_\Lambda P_{\Lambda\setminus\Gamma(y)}] \tag{3.13b}$$

$$\geqslant 0 \tag{3.13c}$$

since

$$\frac{P_\Lambda}{Q_\Lambda} \geqslant \frac{P_{\Lambda\setminus\Gamma(y)}}{Q_{\Lambda\setminus\Gamma(y)}} \tag{3.14}$$

by the inductive hypothesis.

Since P_Λ/Q_Λ is monotone increasing in Λ, we have $P_X/Q_X \geqslant P_\varnothing/Q_\varnothing = 1$, which proves (3.3). More generally, for any subsets $Y, Z \subseteq X$, we have $P_{Y\cup Z}/Q_{Y\cup Z} \geqslant P_Z/Q_Z$ and hence $P_{Y\cup Z}/P_Z \geqslant Q_{Y\cup Z}/Q_Z$, which gives (3.4).

For $\mathbf{p} \notin \mathscr{R}(G)$, choose a minimal vector $\mathbf{p}' \leqslant \mathbf{p}$ such that $\mathbf{p}' \geqslant 0$ and $Z_G(-\mathbf{p}') = 0$ (such a $\mathbf{p}'$ is in general non-unique). Then the family of events $(B'_x)_{x\in X}$ defined by (3.6) with p_x replaced by p'_x satisfies $\mathbb{P}(\bigcap_{x\in X} \overline{B}'_x) = Z_G(-\mathbf{p}') = 0$ (by (3.7c) with $\Lambda = \varnothing$). Since $\mathbf{p}'$ is in the closure of $\mathscr{R}(G)$, it follows by the minimality of $\mathbf{p}'$ and the continuity of Z_G that this is a well-defined probability measure; note that if x and y are adjacent then $\mathbb{P}(B'_x \cap B'_y) = 0$ by (3.6). Thus we have constructed a collection of events satisfying part (b)(ii) of the theorem.

To construct a collection of events satisfying part (b)(i), let $(C_x)_{x\in X}$ be an (independent) collection of independent events satisfying

$$[1 - \mathbb{P}(B'_x)]\,[1 - \mathbb{P}(C_x)] = 1 - p_x. \tag{3.15}$$

Then the events $B_x = B'_x \cup C_x$ satisfy $\mathbb{P}(B_x) = p_x$ and $\mathbb{P}(\bigcap \overline{B}_x) \leqslant \mathbb{P}(\bigcap \overline{B}'_x) = 0$. □

Remarks. **(1)** If $(A_x)_{x\in X}$ is a family of events satisfying (3.3) with equality, then we have $P_X = Q_X$ in the foregoing proof; and since $P_\varnothing = Q_\varnothing = 1$, the monotonicity of P_Λ/Q_Λ implies that we have $P_\Lambda = Q_\Lambda$ for every $\Lambda \subseteq X$. Thus, if $(A_x)_{x\in X}$ is an extremal family, the probabilities of all events in $\sigma(A_x\colon\ x \in X)$ are completely determined and are given by (3.6)/(3.7).

(2) More generally, dependencies between events can be expressed in terms of a dependency *digraph*: each event A_x is independent from the σ-algebra $\sigma(A_y\colon\ y \in X \setminus \Gamma^*_+(x))$,

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

where $\Gamma_+^*(x) = \Gamma_+(x) \cup \{x\}$ and $\Gamma_+(x)$ is the out-neighbourhood of x. See, *e.g.*, [2, Lemma 5.1.1] or [5, Theorem 1.17]. It would be interesting to have a digraph analogue of Theorem 3.1, but we do not know how to do this.

3.2. Soft-core version

Let us now consider how to extend Theorem 3.1 to the more general case of a soft-core pair interaction, *i.e.*, to allow 'soft edges' xy of strength $1 - W(x,y) \in [0,1]$. The first step here is to replace the hard-core dependency condition (3.2) by an appropriate soft-core version.

Let $W : X \times X \to [0,1]$ be symmetric and satisfy $W(x,x) = 0$ for all $x \in X$; and let $(A_x)_{x\in X}$ be a collection of events in some probability space. For each $x \in X$, let S_x be a random subset of X, independent of the σ-algebra $\sigma(A_x : x \in X)$, defined by the probabilities

$$\mathbb{P}(y \in S_x) = W(x,y) \tag{3.16}$$

independently for each $y \in X$. (Thus in the case of a hard-core pair interaction, we have $S_x = X \setminus \Gamma^*(x)$ with probability 1.) Let $(p_x)_{x\in X}$ be real numbers in $[0,1]$. We say that $(A_x)_{x\in X}$ satisfies the *weak dependency conditions with interaction W and probabilities $(p_x)_{x\in X}$* if, for each $x \in X$ and each $Y \subseteq X \setminus x$ we have

$$\mathbb{E}\left(\mathbb{P}\left(A_x \cap \bigcap_{y\in Y\cap S_x} \overline{A}_y\right)\right) \leqslant p_x \mathbb{E}\left(\mathbb{P}\left(\bigcap_{y\in Y\cap S_x} \overline{A}_y\right)\right), \tag{3.17}$$

where the expectations are taken over the random choice of subset S_x. (Note that in the special case of a hard-core pair interaction, we have $Y \cap S_x = Y \setminus \Gamma^*(x)$ with probability 1, so that (3.17) reduces to (3.2).) Of course, the reference here to a random subset S_x can be replaced by an explicit expression for the probabilities $\mathbb{P}(Y \cap S_x = Y')$, so that (3.17) is equivalent to

$$\begin{aligned}\sum_{Y'\subseteq Y} \left(\prod_{y\in Y'} W(x,y)\right)\left(\prod_{y\in Y\setminus Y'} [1 - W(x,y)]\right)\mathbb{P}\left(A_x \cap \bigcap_{y\in Y'} \overline{A}_y\right) \leqslant \\ p_x \sum_{Y'\subseteq Y} \left(\prod_{y\in Y'} W(x,y)\right)\left(\prod_{y\in Y\setminus Y'} [1 - W(x,y)]\right)\mathbb{P}\left(\bigcap_{y\in Y'} \overline{A}_y\right).\end{aligned} \tag{3.18}$$

We also replace (3.1) by the definition

$$\mathscr{R}(W) = \{\mathbf{R} \in [0,\infty)^X : Z_W(\mathbf{w}) \neq 0 \,\forall \mathbf{w} \in \overline{D}_{\mathbf{R}}\}. \tag{3.19}$$

We can now state a soft-core version of Theorem 3.1.

Theorem 3.2 (The equivalence theorem, soft-core case). *Let $(A_x)_{x\in X}$ be a family of events in some probability space, and let $W : X \times X \to [0,1]$ be symmetric and satisfy $W(x,x) = 0$ for all $x \in X$. Suppose that $(A_x)_{x\in X}$ satisfies the weak dependency conditions* (3.17)/(3.18) *with interaction W and probabilities $(p_x)_{x\in X}$.*

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

(a) *If* $\mathbf{p} \in \mathscr{R}(W)$, *then*

$$\mathbb{P}\left(\bigcap_{x\in X} \overline{A}_x\right) \geqslant Z_W(-\mathbf{p}) > 0 \tag{3.20}$$

and more generally

$$\mathbb{P}\left(\bigcap_{x\in Y} \overline{A}_x \Big| \bigcap_{x\in Z} \overline{A}_x\right) \geqslant \frac{Z_W(-\mathbf{p}\,\mathbf{1}_{Y\cup Z})}{Z_W(-\mathbf{p}\,\mathbf{1}_Z)} > 0 \tag{3.21}$$

for any subsets $Y, Z \subseteq X$. *Furthermore, this bound is best possible in the sense that there exists a family* $(B_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B_x) = p_x$ *that satisfies the weak dependency conditions* (3.17)/(3.18) *with interaction* W *and probabilities* $(p_x)_{x\in X}$, *has strong dependency graph* G_W, *and has* $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x) = Z_W(-\mathbf{p})$.

(b) *If* $\mathbf{p} \notin \mathscr{R}(W)$, *then there exists a probability space on which there can be constructed:*

(i) *a family of events* $(B_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B_x) = p_x$ *and satisfying the weak dependency conditions* (3.17)/(3.18) *with interaction* W, *such that* $\mathbb{P}(\bigcap_{x\in X} \overline{B}_x) = 0$; *and*

(ii) *a family of events* $(B'_x)_{x\in X}$ *with probabilities* $\mathbb{P}(B'_x) = p'_x \leqslant p_x$ *and satisfying the weak dependency conditions* (3.17)/(3.18) *with interaction* W, *such that* $\mathbb{P}(B'_x \cap B'_y) = W(x,y)\mathbb{P}(B'_x)\mathbb{P}(B'_y)$ *for all* x, y *and* $\mathbb{P}(\bigcap_{x\in X} \overline{B}'_x) = 0$.

The proof of Theorem 3.2 is similar to that of Theorem 3.1; details can be found in [26].

4. Dependency graphs and the lattice gas: Sufficient conditions

In this section we shall consider sufficient conditions on a set of radii $\mathbf{R} = \{R_x\}_{x\in X}$ so that the partition function $Z_W(\mathbf{w})$ is nonvanishing in the closed polydisc $|\mathbf{w}| \leqslant \mathbf{R}$. Our main tool will be the fundamental identity (2.5), applied inductively.

4.1. The Lovász Local Lemma

Let G be a graph with vertex set X. Recall that G is a *dependency graph* for the family $(A_x)_{x\in X}$ if, for each $x \in X$, the event A_x is independent from the σ-algebra generated by the events $\{A_y : y \in X \setminus \Gamma^*(x)\}$. Erdős and Lovász [14] proved the following fundamental result.

Theorem 4.1 (Lovász Local Lemma). *Let* G *be a dependency graph for the family of events* $(A_x)_{x\in X}$, *and suppose that* $(r_x)_{x\in X}$ *are real numbers in* $[0,1)$ *such that, for each* x,

$$\mathbb{P}(A_x) \leqslant r_x \prod_{y\in\Gamma(x)} (1 - r_y). \tag{4.1}$$

Then $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) \geqslant \prod_{x\in X}(1 - r_x) > 0$.

Erdős and Spencer [15] (see also [2, 23]) later noted that the same conclusion holds even if A_x and $\sigma(A_y : y \in X \setminus \Gamma^*(x))$ are not independent, provided that the 'harmful'

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

conditional probabilities are suitably bounded. More precisely:

Theorem 4.2 (Lopsided Lovász Local Lemma). *Let $(A_x)_{x\in X}$ be a family of events on some probability space, and let G be a graph with vertex set X. Suppose that $(r_x)_{x\in X}$ are real numbers in $[0,1)$ such that, for each x and each $Y \subseteq X \setminus \Gamma^*(x)$, we have*

$$\mathbb{P}\left(A_x \Big| \bigcap_{y\in Y} \overline{A}_y\right) \leqslant r_x \prod_{y\in\Gamma(x)} (1-r_y). \tag{4.2}$$

Then $\mathbb{P}(\bigcap_{x\in X} \overline{A}_x) \geqslant \prod_{x\in X}(1-r_x) > 0$.

In fact, the arguments of [14, 15] (see also [32, 33]) show that in Theorems 4.1 and 4.2 a slightly stronger conclusion holds: for all pairs Y, Z of subsets of X we have

$$\mathbb{P}\left(\bigcap_{x\in Y} \overline{A}_x \Big| \bigcap_{x\in Z} \overline{A}_x\right) \geqslant \prod_{x\in Y\setminus Z} (1-r_x). \tag{4.3}$$

The Lovász Local Lemma has proved incredibly useful in probabilistic combinatorics. However, one limitation of the result is that it does not take into account the 'strength' of dependence. Our aim in this section is to relate the Lovász Local Lemma to Dobrushin's Theorem (presented in Section 4.2) and to discuss some consequences. In particular, we present an extension of the Lovász Local Lemma to the context of 'weak' dependence (Theorem 4.6). Here the precise definition of 'weak dependence' is essentially forced upon us by Theorem 3.2, and it may be a little difficult to use in practice. It would be very interesting to see some concrete applications of Theorem 4.6.

4.2. Basic bound

In this section, we will provide some sufficient conditions for the nonvanishing of Z_W in a closed polydisc $\bar{D}_{\mathbf{R}}$, based on 'local' properties of the interaction W (or of the graph G). Results of this type have traditionally been proved [24, 11, 27, 7, 8, 10, 29, 9] by explicitly bounding the terms in the Mayer expansion (2.14); this requires some rather nontrivial combinatorics (for example, facts about partitionability together with the counting of trees). Once this is done, an immediate consequence is that Z_W is nonvanishing in any polydisc where the series for $\log Z_W$ is convergent. Dobrushin's brilliant idea [12, 13] was to prove these two results in the opposite order. First one proves, by an elementary induction on the cardinality of the state space, that Z_W is nonvanishing in a suitable polydisc (Theorem 4.3); it then follows immediately that $\log Z_W$ is analytic in that polydisc, and hence that its Taylor series (2.14) is convergent there. Let us remark that the Dobrushin–Shearer inductive method employed in Section 4 is limited, at present, to models with hard-core self-repulsion (2.2), for which Z_W is a multiaffine polynomial. It is an interesting open question to know whether this approach can be made to work without the assumption of hard-core self-repulsion.[4]

[4] See also Kotecký and Preiss [21] for a third approach to proving the convergence of the Mayer expansion.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Our first (and most basic) bound is due to Dobrushin [12, 13] in the case of a hard-core interaction; the generalization to a soft repulsive interaction was proved a few years ago by one of us [30]. The method of proof is, however, already implicit (in more powerful form) in Shearer [28, Theorem 2].

Theorem 4.3 (Dobrushin [12, 13], Sokal [30]). *Let X be a finite set, and let W satisfy*

(a) $0 \leqslant W(x, y) \leqslant 1$ *for all* $x, y \in X$,
(b) $W(x, x) = 0$ *for all* $x \in X$.

Let $\mathbf{R} = \{R_x\}_{x \in X} \geqslant 0$. *Suppose that there exist constants* $\{K_x\}_{x \in X}$ *satisfying* $0 \leqslant K_x < 1/R_x$ *and*

$$K_x \geqslant \prod_{y \neq x} \frac{1 - W(x, y) K_y R_y}{1 - K_y R_y} \tag{4.4}$$

for all $x \in X$. *Then, for each subset* $\Lambda \subseteq X$, $Z_\Lambda(\mathbf{w})$ *is nonvanishing in the closed polydisc* $\bar{D}_{\mathbf{R}} = \{\mathbf{w} \in \mathbb{C}^X : |w_x| \leqslant R_x \text{ for all } x\}$ *and satisfies there*

$$\left| \frac{\partial \log Z_\Lambda(\mathbf{w})}{\partial w_x} \right| \leqslant \begin{cases} \dfrac{K_x}{1 - K_x |w_x|} & \text{for all } x \in \Lambda, \\ 0 & \text{for all } x \in X \setminus \Lambda. \end{cases} \tag{4.5}$$

Moreover, if $\mathbf{w}, \mathbf{w}' \in \bar{D}_{\mathbf{R}}$ *and* $w'_x / w_x \in [0, +\infty]$ *for each* $x \in \Lambda$, *then*

$$\left| \log \frac{Z_\Lambda(\mathbf{w}')}{Z_\Lambda(\mathbf{w})} \right| \leqslant \sum_{x \in \Lambda} \left| \log \frac{1 - K_x |w'_x|}{1 - K_x |w_x|} \right|, \tag{4.6}$$

where on the left-hand side we take the standard branch of the log, i.e., $|\operatorname{Im} \log \cdots| \leqslant \pi$.

Remark 1. It follows from (4.4) that $K_x \geqslant 1$ and hence that $R_x < 1$.

It is convenient to rewrite Theorem 4.3 in terms of the new variables $r_x = K_x R_x$.

Corollary 4.4. *Let X be a finite set, and let W satisfy* $0 \leqslant W(x, y) \leqslant 1$ *for all* $x, y \in X$ *and* $W(x, x) = 0$ *for all* $x \in X$. *Let* $\mathbf{R} = \{R_x\}_{x \in X} \geqslant 0$. *Suppose that there exist constants* $0 \leqslant r_x < 1$ *satisfying*

$$R_x \leqslant r_x \prod_{y \neq x} \frac{1 - r_y}{1 - W(x, y) r_y} \tag{4.7}$$

for all $x \in X$. *Then, for all* $\mathbf{w}$ *satisfying* $|\mathbf{w}| \leqslant \mathbf{R}$, *the partition function* Z_W *satisfies*

$$|Z_W(\mathbf{w})| \geqslant Z_W(-\mathbf{R}) \geqslant \prod_{x \in X} (1 - r_x) > 0 \tag{4.8}$$

and more generally

$$\left| \frac{Z_W(\mathbf{w} \mathbf{1}_{Y \cup Z})}{Z_W(\mathbf{w} \mathbf{1}_Z)} \right| \geqslant \prod_{x \in Y} (1 - r_x) > 0. \tag{4.9}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

In particular, if we define the 'maximum weighted degree'

$$\Delta_W = \max_{x \in X} \sum_{y \neq x} [1 - W(x,y)] \tag{4.10}$$

and write

$$F(\Delta_W) = \frac{2 + \Delta_W - \sqrt{\Delta_W^2 + 4\Delta_W}}{2}, \tag{4.11}$$

$$R(\Delta_W) = F(\Delta_W)\, e^{-[1-F(\Delta_W)]} \tag{4.12}$$

we have

$$|Z_W(\mathbf{w})| \geqslant [1 - F(\Delta_W)]^{|X|} > 0 \tag{4.13}$$

whenever $|w_x| \leqslant R(\Delta_W)$ *for all* $x \in X$.

Proof. Setting $r_x = K_x R_x$, we find that (4.4) becomes (4.7), and (4.6) with $\Lambda = X$ and $\mathbf{w}' = 0$ becomes (4.8).

To obtain the last claim, note first that

$$\frac{1-r}{1-Wr} = \frac{1-r}{1-r+(1-W)r} = \frac{1}{1+(1-W)\frac{r}{1-r}} \geqslant e^{-(1-W)r/(1-r)} \tag{4.14}$$

whenever $0 \leqslant W \leqslant 1$ and $0 \leqslant r \leqslant 1$. Therefore, if we set $r_x = r$ for all $x \in X$, we have

$$r_x \prod_{y \neq x} \frac{1 - r_y}{1 - W(x,y) r_y} \geqslant r e^{-\Delta_W r/(1-r)}. \tag{4.15}$$

We then choose r to maximize the right-hand side of (4.15); simple calculus yields $\Delta_W r = (1-r)^2$ and $r = F(\Delta_W)$, so that the right-hand side of (4.15) is bounded below by $R(\Delta_W)$. It follows that if we define $R_x = R(\Delta_W)$ and $r_x = F(\Delta_W)$ for all $x \in X$ then (4.7) and so (4.8) are satisfied. □

Remarks. **(1)** The radius $R(\Delta_W)$ behaves as

$$R(\Delta_W) = \begin{cases} 1 - 2\Delta_W^{1/2} + \frac{5}{2}\Delta_W + O(\Delta_W^{3/2}) & \text{as } \Delta_W \to 0, \\ \frac{1}{e\Delta_W}\left[1 - \frac{1}{\Delta_W} + \frac{3}{2\Delta_W} + O(\Delta_W^{-3})\right] & \text{as } \Delta_W \to \infty. \end{cases} \tag{4.16}$$

Example 4.6 (the r-ary rooted tree) shows that this bound is sharp (to leading order) as $\Delta_W \to \infty$. At the other extreme, the $1 - \text{const} \times \Delta_W^{1/2}$ behaviour at small Δ_W is also best possible, since the two-site lattice gas with $W(x,x) = W(y,y) = 0$ and $W(x,y) = 1 - \epsilon$ has $Z_W(w) = 1 + 2w + (1-\epsilon)w^2$ and hence has a root at $w = -1/(1 + \sqrt{\epsilon})$. (However, the coefficient 2 rather than 1 in the $\Delta_W^{1/2}$ term of (4.16) may not be best possible.)

Specializing Corollary 4.4 to the case of a hard-core pair interaction for a graph G,

$$W(x,y) = \begin{cases} 0 & \text{if } x = y \text{ or } xy \in E(G), \\ 1 & \text{if } x \neq y \text{ and } xy \notin E(G), \end{cases} \tag{4.17}$$

we have the following result.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Corollary 4.5. *Let G be a finite graph with vertex set X, and let $\mathbf{R} = \{R_x\}_{x \in X} \geqslant 0$. Suppose that there exist constants $0 \leqslant r_x < 1$ satisfying*

$$R_x \leqslant r_x \prod_{y \in \Gamma(x)} (1 - r_y) \tag{4.18}$$

for all $x \in X$. Then, for all $\mathbf{w}$ satisfying $|\mathbf{w}| \leqslant \mathbf{R}$, the independent-set polynomial Z_G satisfies

$$|Z_G(\mathbf{w})| \geqslant Z_G(-\mathbf{R}) \geqslant \prod_{x \in X} (1 - r_x) > 0 \tag{4.19}$$

and more generally

$$\left| \frac{Z_G(\mathbf{w}\, \mathbf{1}_{Y \cup Z})}{Z_G(\mathbf{w}\, \mathbf{1}_Z)} \right| \geqslant \prod_{x \in Y} (1 - r_x) > 0. \tag{4.20}$$

In particular, if G has maximum degree Δ, then $|Z_G(\mathbf{w})| \geqslant [\Delta/(\Delta+1)]^{|X|} > 0$ whenever $|w_x| \leqslant \Delta^\Delta/(\Delta+1)^{\Delta+1}$ for all $x \in X$.

Proof. The last claim is obtained by setting $r_x = 1/(\Delta+1)$ for all $x \in X$. □

Remark. The radius $\Delta^\Delta/(\Delta+1)^{\Delta+1}$ behaves for large Δ as

$$\frac{\Delta^\Delta}{(\Delta+1)^{\Delta+1}} = \frac{1}{e\Delta}\left[1 - \frac{1}{2\Delta} + \frac{7}{24\Delta^2} - \frac{3}{16\Delta^3} + O(\Delta^{-4})\right], \tag{4.21}$$

which agrees with (4.16) to leading order in $1/\Delta$ but is slightly larger (hence better) at order $1/\Delta^2$.

Combining Corollary 4.5 with Theorem 3.1, we immediately obtain the lopsided Lovász Local Lemma (Theorem 4.2). It is equally possible to go in the opposite direction, and deduce the Dobrushin bounds from the Lovász Local Lemma.

Let us remark that we have been able to relate the Lovász Local Lemma to a combinatorial polynomial (namely, the independent-set polynomial) only in the case of an *undirected* dependency graph G. Although the Local Lemma can be formulated quite naturally for a dependency *digraph* [2, 5, 23], we do not know whether the digraph Lovász problem can be related to any combinatorial polynomial. (Clearly the independent-set polynomial cannot be the right object in the digraph context, since exclusion of simultaneous occupation is manifestly a symmetric condition.)

The results also allow us to deduce a 'soft-core' version of the lopsided Lovász Local Lemma. Combining Corollary 4.4 with Theorem 3.2, we obtain the following result.

Theorem 4.6. *Let $(A_x)_{x \in X}$ be a family of events in some probability space, and let $W : X \times X \to [0,1]$ be symmetric and satisfy $W(x,x) = 0$ for all $x \in X$. Suppose that $(A_x)_{x \in X}$ satisfies the weak dependency conditions (3.17)/(3.18) with interaction W and probabilities $(p_x)_{x \in X}$. Suppose further that $(r_x)_{x \in X}$ are real numbers in $[0,1)$ satisfying*

$$p_x \leqslant r_x \prod_{y \in \Gamma(x)} (1 - r_y). \tag{4.22}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

Then

$$\mathbb{P}\left(\bigcap_{x\in X}\overline{A}_x\right) \geqslant \prod_{x\in X}(1-r_x) > 0, \tag{4.23}$$

and more generally for sets $Y, Z \subseteq X$, *we have*

$$\mathbb{P}\left(\bigcap_{x\in Y}\overline{A}_x \Big| \bigcap_{x\in Z}\overline{A}_x\right) \geqslant \prod_{x\in Y\setminus Z}(1-r_x) > 0. \tag{4.24}$$

Defining the weighted degree Δ_W as in (4.10), we obtain the following.

Lemma 4.7. *Let* $(A_x)_{x\in X}$ *satisfy the weak dependency conditions* (3.17)/(3.18) *with interaction* W *and probabilities* $(p_x)_{x\in X}$. *If* $p_x < \Delta_W^{\Delta_W}/(\Delta_W+1)^{\Delta_W+1}$ *for every* $x \in X$, *then* $\mathbb{P}(\bigcap_{x\in X}\overline{A}_x) > 0$.

Proof. As in the proof of Corollary 4.5, set $r_x = r \equiv 1/(\Delta_W+1)$ for all $x \in X$. Then check (4.7):

$$\begin{aligned} r_x \prod_{y\neq x}\frac{1-r_y}{1-W(x,y)r_y} &\leqslant r_x \prod_{y\neq x}(1-r_y)^{1-W(x,y)} \\ &\leqslant r(1-r)^{\Delta_W} \\ &\leqslant \frac{\Delta_W^{\Delta_W}}{(\Delta_W+1)^{\Delta_W+1}}. \end{aligned} \tag{4.25}$$

In the first inequality we have used the fact that $1 - W(x,y)r_y \leqslant (1-r_y)^{W(x,y)}$ for $0 \leqslant W(x,y) \leqslant 1$. □

As noted above, it would be interesting to see applications of Theorem 4.6 and Lemma 4.7.

4.3. Improved bound

Finally in this section, we note that Theorem 4.3 can be slightly sharpened. Note, first of all, that we need not insist that the bound (4.5) hold with the *same* constant K_x for all $\Lambda \ni x$; rather, we can use constants $K_{x,\Lambda}$ that depend on Λ.

Let us define the constants $K_{x,\Lambda} \in [0, +\infty]$ as a function of the family $\{R_x\}$ by the recursion

$$K_{x,\Lambda} = \prod_{\substack{y\in\Lambda\setminus x \\ W(x,y)\neq 1 \\ R_y>0}} \frac{1 - W(x,y)K_{y,\Lambda\setminus x}R_y}{1 - K_{y,\Lambda\setminus x}R_y} \tag{4.26}$$

if $K_{y,\Lambda\setminus x}R_y < 1$ for all terms in the product, and $K_{x,\Lambda} = +\infty$ otherwise.

We define a graph G with vertex set $V = \{x \in X : R_x > 0\}$ and edge set $E = \{x, y \in V : W(x,y) \neq 1\}$; and for each $\Lambda \subseteq X$, let G_Λ be the subgraph of G induced by $\Lambda \cap V$. Then only the connected component of G_Λ containing x plays any role in the definition

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

of $K_{x,\Lambda}$: that is, if G_Λ has several connected components with vertex sets $\Lambda_1, \dots, \Lambda_k$ and $x \in \Lambda_i$, then $K_{x,\Lambda} = K_{x,\Lambda_i}$.

Let us now call a pair (x, Λ) 'good' if $K_{x,\Lambda} < \infty$ and $K_{x,\Lambda} R_x < 1$. It is easily shown that if (x, Λ) is good, then $(y, \Lambda \setminus x)$ is also good whenever $y \in \Lambda \setminus x$ with $W(x, y) \neq 1$ and $R_y > 0$, *i.e.*, whenever y is a neighbour of x in G_Λ. (Indeed, this follows under the weaker hypothesis that $K_{x,\Lambda} < \infty$.)

The following result can then be proved (see [26] for a proof).

Theorem 4.8 (Improved Dobrushin–Shearer bound). *Let X be a finite set, and let W satisfy*

(a) $0 \leqslant W(x, y) \leqslant 1$ *for all* $x, y \in X$,
(b) $W(x, x) = 0$ *for all* $x \in X$.

Let $\mathbf{R} = \{R_x\}_{x \in X} \geqslant 0$. *Define the constants* $K_{x,\Lambda} \in [0, +\infty]$ *as above. Suppose that in each connected component of* G_Λ *there exists at least one vertex* x *for which the pair* (x, Λ) *is good. Then* $Z_\Lambda(\mathbf{w})$ *is nonvanishing in the closed polydisc* $\bar{D}_{\mathbf{R}}$; *and for every good pair* (x, Λ) *and every* $\mathbf{w} \in \bar{D}_{\mathbf{R}}$, *we have*

$$\left| \frac{\partial \log Z_\Lambda(\mathbf{w})}{\partial w_x} \right| \leqslant \frac{K_{x,\Lambda}}{1 - K_{x,\Lambda} |w_x|}. \tag{4.27}$$

Moreover, if $\mathbf{w}, \mathbf{w}' \in \bar{D}_{\mathbf{R}}$ *and* $w'_x / w_x \in [0, +\infty]$ *for each* $x \in \Lambda$, *and in addition the pair* (x, Λ) *is good whenever* $w'_x \neq w_x$, *then*

$$\left| \log \frac{Z_\Lambda(\mathbf{w}')}{Z_\Lambda(\mathbf{w})} \right| \leqslant \sum_{\substack{x \in \Lambda \\ w'_x \neq w_x}} \left| \log \frac{1 - K_{x,\Lambda} |w'_x|}{1 - K_{x,\Lambda} |w_x|} \right|, \tag{4.28}$$

where on the left-hand side we take the standard branch of the log, i.e., $|\operatorname{Im} \log \cdots| \leqslant \pi$.

As a corollary of Theorem 4.8, we can deduce a bound due originally (in the Lovász context) to Shearer [28, Theorem 2], which improves the last sentence of Corollary 4.5 by replacing Δ by $\Delta - 1$. Indeed, we can very slightly improve Shearer's bound by allowing *one* vertex x_0 to have a larger radius R_{x_0}.

Corollary 4.9. *Let* $G = (X, E)$ *be a finite graph of maximum degree* $\Delta \geqslant 2$, *and fix one vertex* $x_0 \in X$. *Suppose that* $|w_{x_0}| \leqslant (\Delta - 1)^\Delta / \Delta^\Delta$ *and that* $|w_x| \leqslant (\Delta - 1)^{\Delta - 1} / \Delta^\Delta$ *for all* $x \neq x_0$. *Then* $Z_G(\mathbf{w}) \neq 0$.

Proof. Since Z_G factorizes over connected components, we can assume without loss of generality that G is connected. (Indeed, if G is disconnected, then we can allow one 'x_0-like' vertex in *each* connected component.) Set $R_{x_0} = (\Delta - 1)^\Delta / \Delta^\Delta$ and $R_x = (\Delta - 1)^{\Delta - 1} / \Delta^\Delta$ for all $x \neq x_0$.

We first claim that if $x_0 \notin \Lambda$, and $x \in \Lambda$ is a vertex with at least one neighbour in $X \setminus \Lambda$, then

$$K_{x,\Lambda} < \left(\frac{\Delta}{\Delta - 1} \right)^{\Delta - 1} \tag{4.29}$$

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

(note the strict inequality). The proof is by induction on $|\Lambda|$, using the definition (4.26). It certainly holds if $\Lambda = \{x\}$. For general Λ, note first that since every y appearing in the product on the right-hand side of (4.26) has at least one neighbour outside of $\Lambda \setminus x$ (namely, x itself), $K_{y,\Lambda\setminus x}$ satisfies (4.29) by the inductive hypothesis and so $K_{y,\Lambda\setminus x} R_y < 1/\Delta$. Also, since x has at least one neighbour outside Λ, there are at most $\Delta - 1$ factors in the product. Thus

$$K_{x,\Lambda} < \left(\frac{1}{1-1/\Delta}\right)^{\Delta-1} = \left(\frac{\Delta}{\Delta-1}\right)^{\Delta-1}. \tag{4.30}$$

It then follows that

$$K_{x_0,X} < \left(\frac{\Delta}{\Delta-1}\right)^{\Delta}, \tag{4.31}$$

since the bound (4.29) applies to all the terms $K_{y,\Lambda\setminus x_0}$ appearing on the right-hand side of (4.26). We therefore have $K_{x_0,X} R_{x_0} < 1$, and so the pair (x_0, X) is good. The claim then follows from Theorem 4.8. □

Replacing $\Delta^\Delta/(\Delta+1)^{\Delta+1}$ by $(\Delta-1)^{\Delta-1}/\Delta^\Delta$ may seem to be a negligible improvement, since both quantities have the same leading behaviour $\approx 1/(e\Delta)$ as $\Delta \to \infty$, and differ only at higher order:

$$\frac{(\Delta-1)^{\Delta-1}}{\Delta^\Delta} = \frac{1}{e\Delta}\left[1 + \frac{1}{2\Delta} + \frac{7}{24\Delta^2} + \frac{3}{16\Delta^3} + O(\Delta^{-4})\right] \tag{4.32}$$

(*cf.* (4.21)).[5] But Shearer's bound $(\Delta-1)^{\Delta-1}/\Delta^\Delta$ has the great merit of being *best possible*: for, as he showed [28], if G is the complete rooted tree with branching factor $r = \Delta - 1$ and depth n, then $Z_G(w)$ has negative real zeros that tend to $w = -(\Delta-1)^{\Delta-1}/\Delta^\Delta$ as $n \to \infty$ (see Section 2.2 above).

We remark that Corollary 4.9 does not appear to extend naturally to the soft-core case (note that having one neighbour outside Λ in the argument around (4.29) need not reduce the weighted degree of a vertex in Λ by 1).

For additional detail and further results, as well as a discussion of the *optimal* bounds, we refer the reader to [26].

5. Conclusion

How good are the bounds given by the results in Section 4? We shall consider the diagonal case, where all radii are the same. Let us define, for a finite graph G,

$$\lambda_c(G) = \sup\{\lambda : \lambda \mathbf{1} \in \mathscr{R}(G)\}. \tag{5.1}$$

For a countably infinite graph G, we define $\lambda_c(G)$ to be the infimum of $\lambda_c(H)$ over finite induced subgraphs H of G.

For graphs G with maximum degree Δ, we know from Corollary 4.9 that $\lambda_c(G) \leqslant (\Delta-1)^{\Delta-1}/\Delta^\Delta$, and by Shearer's result this is optimal for the infinite Δ-regular tree.

[5] The amusing similarity of (4.21) and (4.32) arises from the fact that $-(-\Delta)^{-\Delta}/(-\Delta+1)^{-\Delta+1} = (\Delta-1)^{\Delta-1}/\Delta^\Delta$.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

The bound is close to optimal for Δ-regular graphs of large girth (as can be seen from monotonicity results from [26]). However, the value of λ_c is less easy to determine for graphs with short cycles. For instance, consider the square lattice $\mathbb{Z}^2$. This is 4-regular, and so $\lambda_c \geqslant 3^3/4^4 = 0.105$, but the correct value of λ_c is not so clear. Fortunately, this problem has also been considered by physicists. Indeed, Todo [35] has given the extraordinarily precise (but nonrigorous) numerical estimate

$$\lambda_c(\mathbb{Z}^2) = 0.119\,338\,881\,88(1), \tag{5.2}$$

obtained by using transfer matrices and the phenomenological-renormalization method (a variant of finite-size scaling). It would be interesting to gain good rigorous estimates (see [26] for further discussion).

There are other probabilistic inequalities that are expressed in terms of a dependency graph (see for instance Suen [34] or Janson [20]); it would be interesting to know if any of these can be related to a combinatorial polynomial (= statistical-mechanical partition function) in a manner analogous to Theorem 3.1. However, even without such a result, there may be scope for proving further inequalities in the presence of weak dependency conditions of the form discussed in Section 3.2 above.

Finally, we note that combinatorics and statistical physics have seen a very extensive and fruitful interaction in recent years; we hope that the results in this paper indicate that much remains to be discovered.

Acknowledgements

We wish to thank Keith Ball for informing us of the important work of Shearer [28]; Keith Ball, Pierre Leroux and Joel Spencer for helpful discussions; and Synge Todo for communicating to us some of his unpublished numerical results. We would also like to thank the referees for their helpful comments.

This research was supported in part by US National Science Foundation grants PHY–9900769 and PHY–0099393 and UK Engineering and Physical Sciences Research Council grant GR/S26323/01.

References

[1] Alon, N. (1991) A parallel algorithmic version of the local lemma. *Random Struct. Alg.* **2** 367–378.
[2] Alon, N. and Spencer, J. (2000) *The Probabilistic Method*, 2nd edn, Wiley, New York.
[3] Beck, J. (1991) An algorithmic approach to the Lovász local lemma I. *Random Struct. Alg.* **2** 343–365.
[4] Billingsley, P. (1986) *Probability and Measure*, 2nd edn, Wiley, New York.
[5] Bollobás, B. (2001) *Random Graphs*, 2nd edn, Cambridge University Press, Cambridge.
[6] Borgs, C. (2003) *Expansion Methods in Combinatorics*. Preprint: to be published in the *Conference Board of the Mathematical Sciences* book series.
[7] Brydges, D. C. (1986) A short course on cluster expansions. In *Phénomènes Critiques, Systèmes Aléatoires, Théories de Jauge / Critical Phenomena, Random Systems, Gauge Theories*, Les Houches summer school, Session XLIII, 1984 (K. Osterwalder and R. Stora, eds), Elsevier/North-Holland, Amsterdam, pp. 129–183.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

[8] Brydges, D. C. and Kennedy, T. (1987) Mayer expansions and the Hamilton–Jacobi equation. *J. Statist. Phys.* **48** 19–49.
[9] Brydges, D. C. and Martin, Ph. A. (1999) Coulomb systems at low density: A review. *J. Statist. Phys.* **96** 1163–1330, `cond-mat/9904122` at `arXiv.org`.
[10] Brydges, D. C. and Wright, J. D. (1988) Mayer expansions and the Hamilton–Jacobi equation II: Fermions, dimensional reduction formulas. *J. Statist. Phys.* **51** 435–456. Erratum **97** (1999) 1027.
[11] Cammarota, C. (1982) Decay of correlations for infinite range interactions in unbounded spin systems. *Commun. Math. Phys.* **85** 517–528.
[12] Dobrushin, R. L. (1996) Estimates of semi-invariants for the Ising model at low temperatures. In *Topics in Statistical and Theoretical Physics*, Vol. 177 of *American Mathematical Society Translations*, Ser. 2, pp. 59–81.
[13] Dobrushin, R. L. (1996) Perturbation methods of the theory of Gibbsian fields. In *Lectures on Probability Theory and Statistics* Ecole d'Eté de Probabilités de Saint-Flour XXIV, 1994 (P. Bernard, ed.), Vol. 1648 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 1–66.
[14] Erdős, P. and Lovász, L. (1975) Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and Finite Sets*, Vol. II, *Colloq. Math. Soc. Janos Bolyai*, Vol. 10, North-Holland, Amsterdam, pp. 609–627.
[15] Erdős, P. and Spencer, J. (1991) Lopsided Lovász local lemma and Latin transversals. *Discrete Appl. Math.* **30** 151–154.
[16] Garey, M. R. and Johnson, D. S. (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco.
[17] Glimm, J. and Jaffe, A. (1987) *Quantum Physics: A Functional Integral Point of View*, 2nd edn, Springer, New York.
[18] Groeneveld, J. (1962) Two theorems on classical many-particle systems. *Phys. Lett.* **3** 50–51.
[19] Hille, E. (1973) *Analytic Function Theory*, 2nd edn, Chelsea, New York.
[20] Janson, S. (1998) New versions of Suen's correlation inequality. *Random Struct. Alg.* **13** 467–483.
[21] Kotecký, R. and Preiss, D. (1986) Cluster expansion for abstract polymer models. *Commun. Math. Phys.* **103** 491–498.
[22] Milnor, J. (2000) On rational maps with two critical points. *Experiment. Math.* **9** 481–522.
[23] Molloy, M. and Reed, B. (2002) *Graph Colouring and the Probabilistic Method*, Springer, Berlin/New York.
[24] Penrose, O. (1967) Convergence of fugacity expansions for classical systems. In *Statistical Mechanics: Foundations and Applications* (T. A. Bak, ed.), Benjamin, New York/Amsterdam, pp. 101–109.
[25] Runnels, L. K. (1967) Phase transition of a Bethe lattice gas of hard molecules. *J. Math. Phys.* **10** 2081–2087.
[26] Scott, A. D. and Sokal, A. D. (2005) The repulsive lattice gas, the independent-set polynomial, and the Lovász Local Lemma. *J. Statist. Phys.* **118** 1151–1261, `cond-mat/0309352` at `arXiv.org`.
[27] Seiler, E. G. (1982)) *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics*, Vol. 159 of *Lecture Notes in Physics*, Springer, Berlin/New York.
[28] Shearer, J. B. (1985) On a problem of Spencer. *Combinatorica* **5** 241–245.
[29] Simon, B. (1993) *The Statistical Mechanics of Lattice Gases*, Princeton University Press, Princeton, NJ.
[30] Sokal, A. D. (2001) Bounds on the complex zeros of (di)chromatic polynomials and Potts-model partition functions. *Combin. Probab. Comput.* **10** 41–77, `cond-mat/9904146` at `arXiv.org`.
[31] Sokal, A. D. (2005) The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in Combinatorics, 2005*, (Bridget S. Webb, ed.), Cambridge University Press, pp. 173–226.
[32] Spencer, J. (1975) Ramsey's Theorem: a new lower bound. *J. Combin. Theory Ser. A* **18** 108–115.
[33] Spencer, J. (1977/78) Asymptotic lower bounds for Ramsey functions. *Discrete Math.* **20** 69–76.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press

[34] Suen, W. C. (1990) A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph. *Random Struct. Alg.* **1** 231–242.
[35] Todo, S. (1999) Transfer-matrix study of negative-fugacity singularity of hard-core lattice gas. *Int. J. Mod. Phys. C* **10** 517–529. `cond-mat/9703176` at `arXiv.org`.
[36] Uhlenbeck, G. E. and Ford, G. W. (1962) The theory of linear graphs with applications to the theory of the virial development of the properties of gases. In *Studies in Statistical Mechanics*, Vol. I (J. de Boer and G. E. Uhlenbeck, eds), North-Holland, Amsterdam, pp. 119–211.
[37] Williams, D. (1991) *Probability with Martingales*, Cambridge University Press, Cambridge.
[38] Yang, C. N. and Lee, T. D. (1952) Statistical theory of equations of state and phase transitions I: Theory of condensation. *Phys. Rev.* **87** 404–409.

https://doi.org/10.1017/S0963548305007182 Published online by Cambridge University Press