UNIVERSITY OF CALIFORNIA
SANTA CRUZ

LOCAL BOUNDS FOR THE INDEPENDENT SET POLYNOMIAL AND THE PROBABILISTIC METHOD

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER SCIENCE

by

Konstantinos Zampetakis

September 2022

The Dissertation of Konstantinos Zampetakis is approved:

Prof. Seshadhri Comandur, Chair

Prof. John Musacchio

Prof. Daniel Fremont

Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

ABSTRACT

LOCAL BOUNDS FOR THE INDEPENDENT SET POLYNOMIAL AND THE PROBABILISTIC METHOD

by

Konstantinos Zampetakis

The independent set polynomial plays a prominent role in combinatorics, statistical mechanics, and the probabilistic method. Specifically, in combinatorics the independent set polynomial appears as the generating function of the independent sets of a graph, in statistical mechanics as the partition function of the hard-core model, while in probabilistic method, due to the seminal work of Shearer [1], it expresses the full range of applicability of the celebrated Lovász Local Lemma (LLL).

The first three chapters of this thesis are concerned with the study the computability and approximability of the independent set polynomial, with emphasis on its applications in the probabilistic method.

In particular, in Chapter 1, we briefly review previous work, and present our first results: (i) an exact computation procedure of the independent set polynomial in linear time, for all arguments (ii) an improved version of the asymmetric LLL, (iii) two novel local lemmata inspired from non-backtracking walks, which we use to improve the rigorous bound of the "negative fugacity singularity of hard core model" for the triangular lattice, a central problem in statistical physics of lattice gases.

In Chapter 2 we show a rigorous correspondence between walks and local lemmata, which we use to develop a hierarchy of increasingly powerful, increasingly non-local lemmata. To demonstrate the power of our hierarchy, we prove new rigorous lower bounds for aforementioned negative fugacity singularity of the hard core model on several lattices, matching their conjectured values up to three decimal digits.

In Chapter 3 we prove that Shearer's connection between the probabilistic method and the independent set polynomial holds for *arbitrary supermodular* functions, not just probability measures. This means that all LLL machinery can be employed to bound from below an arbitrary supermodular function, based only on information regarding its value at singleton sets and partial information regarding their interactions. We show that our lemma readily implies both the "Quantum LLL" of Ambainis, Kempe, and Sattath [41], and the "Quantum Shearer" criterion of Sattath, Morampudi, Laumann, and Moessner [43].

Finally, in Chapter 4 we turn on the Bethe approximation for partition functions of general graphical models. While, a priori, there is no connection between the (analytically defined) Bethe approximation and the independent set polynomial, we use a recent combinatorial characterization of the Bethe approximation by Vontobel [44] to suggest precisely such a connection, by relating typical random k-lifts of graphs where $k \to \infty$, with the aforementioned tree of non-backtracking walks. In particular, we revisit a recent result of Ruozzi showing that the Bethe partition function is a lower bound for the true partition function, for every graphical model whose constituent factors are log-supermodular. We give a significantly shorter proof of this result. More importantly, we give a new much shorter proof of the celebrated four functions theorem.

To my family, and my advisor Sesh for their constant and generous support.

Acknowledgements

I am grateful to my co-author, and former advisor, Dimitris Achlioptas, for introducing me to the beautiful world of probabilistic combinatorics, for the countless hours he spent patiently working and discussing with me. His high quality research and teaching, have been a constant source of inspiration to me.

I also want to offer my sincere thanks:

To Seshadhri Comandur, for advising me the last three years of my graduate career. His generous support and advice, both in research and in life, greatly helped me to successfully complete my PhD.

To Phokion Kolaitis, Alistair Sinclair, and Abhradeep Guha Thakurta, for serving in my advancement committee, and for the valuable advice they gave me about my future steps.

To John Musacchio, and Daniel Fremont, for serving in my dissertation defense committee, and for their insightful comments and advice.

To my family for being always there for me.

Chapter 1

The Independent Set Polynomial

The independent set polynomial of a graph has one variable per vertex and one monomial per independent set, each monomial being the product of the corresponding variables. When the variables take positive real values, the value of the polynomial coincides with the partition function of the *hard-core* model on the graph, wherein atoms of a gas are distributed upon the vertices so that each vertex accommodates at most one atom, no two adjacent vertices are simultaneously occupied, and where the positive real number associated with each vertex, called fugacity, expresses its attractiveness. Estimating the partition function of the hard-core model for different graphs and fugacities is a central problem of statistical mechanics. Somewhat surprisingly, this estimation is also interesting when the variables take negative values, due to a connection with probabilistic combinatorics and, in particular, the Lovász Local Lemma. Going further, the Lee-Yang theory of phase transitions motivates the estimation of the independent set polynomial for complex arguments, as the absence of zeros in regions of the complex plane implies the absence of phase transitions (non-analyticities) for corresponding real arguments.

We start by revisiting known methods for bounding the independent set polynomial and then by devising new ones. We develop two distinct approaches, corresponding to two different ways of dealing with the exponentially large computation tree that results when the fundamental recurrence obeyed by the independent set polynomial is unfolded. In the first, more combinatorial, approach we develop a map enjoying crisp monotonicity properties from subgraphs of the input graph to truncations of the graph's tree of self-avoiding walks. This allows us to develop several new results, including results about exact evaluation of the polynomial, a first local upper bound for it, and several new local lower bounds. The second, more analytic, approach stays closest to existing methods and treats the computation tree as a circuit. Efficiency in computation is now achieved by showing that, under certain conditions, retaining a logarithmic number of layers closest to the root suffices, as the rest of the circuit has very little influence. This view allows us to recover some seminal results with much simpler proofs and, crucially, without any appeal to notions and results from statistical physics. We hope that this last fact will enable the development of new results.

1.1 Preliminaries and Notation

Let $[n]$ denote the set $\{1, 2, \ldots, n\}$. A (simple, undirected) *graph* G is a pair (V, E), where V is a set whose elements are called *vertices*, and E is a set of paired vertices, whose elements are called *edges*. The graph $G' = (V', E')$ is a *subgraph* of G, if $V' \subseteq V$, and $E' \subseteq E \cap (V' \times V')$. If $E' = E \cap (V' \times V')$, i.e., G' is formed by a subset V' of V and all of the edges in E connecting pairs of V', then we say G' is the subgraph of G *induced* by V', and we write $G\langle V' \rangle := G'$. An *independent set* of a graph G is a subset S of V, such that $G\langle S \rangle$ is devoid of edges, in other words, no edge has both of its endpoints within S. We write $\mathrm{Ind}(G)$ to denote the set of all independent sets of the graph G (note that $\emptyset \in \mathrm{Ind}(G)$, as it vacuously satisfies the definition above). Unless stated otherwise, from now on we assume that G is a graph on $[n]$, i.e., $V = [n]$.

1.2 The Independent Set Polynomial

The independent set polynomial of a graph G, is nothing but the (multivariate) generating function of the independent sets of G.

Definition 1. *For variables $x_1, \ldots, x_n \in \mathbb{C}$, the (multivariate)* independent set polynomial *of G is*

$$Z(\boldsymbol{x}; G) = Z([n]) := \sum_{I \in \mathrm{Ind}(G)} \prod_{i \in I} x_i \ . \tag{1.1}$$

Remark 2. *Unless stated otherwise, we will think of $\boldsymbol{x}$ as fixed and, thus, of Z as a function on $2^{[n]}$ (induced subgraphs of G). Also, we will often refer to the components of the vector $\boldsymbol{x}$ as* activities*.*

The complexity of computing and approximating the independent set polynomial is an extensively studied subject, as there are meaningful instantiations of the polynomial when the activities are positive reals, negative reals, and even complex numbers.

1.2.1 Positive Reals: The Hard Core Model

In a plethora of natural computational problems in combinatorics, probability, and statistical physics we are given as input a graph G that defines a set $\Omega = \Omega(G)$ of objects (configurations) of interest (matchings, spanning trees, etc.), A weight function $w : \Omega \to (0, +\infty)$ assigns a positive weight to each element $\sigma \in \Omega$, giving rise to a probability distribution $\pi(\sigma) = w(\sigma)/Z$, where the normalizing factor $Z := \sum_{\sigma \in \Omega} w(\sigma)$ is called the *partition function*.

When $\Omega = \mathrm{Ind}(G)$ and each $I \in \mathrm{Ind}(G)$ has weight $w(I) = \prod_{i \in I} x_i$, for some reals $\boldsymbol{x} \in (0, +\infty)^n$, the distribution is known as the *hard-core* model of statistical physics, and the independent set polynomial of G equals the partition function. Observe that in the *univariate* case where all vertex activities equal $x > 0$, i.e., $\boldsymbol{x} = x\mathbf{1}$, as $x \to \infty$ the polynomial is increasingly dominated by the contribution of the largest independent sets, which, due to the hardness of the maximum independent set problem, suggests that evaluating the polynomial for arbitrarily large values of x should be intractable. A celebrated achievement in this area is the characterization of the computational tractability of approximating the partition function. For a graph G, let $\Delta = \Delta(G) \geq 3$ denote its maximum degree and let

$$x_c = x_c(\Delta) := \frac{(\Delta - 1)^{\Delta - 1}}{(\Delta - 2)^{\Delta}} \searrow \frac{\mathrm{e}}{\Delta} \ . \tag{1.2}$$

Theorem 3 ([2])**.** *There exists an algorithm which given a graph G and (activities) $\boldsymbol{x} \in (0, x_c)^n$, for any $\varepsilon > 0$, returns $Y \in (1 \pm \varepsilon)Z$ in time $O(n/\varepsilon)^{\log \Delta/(1-\delta)}$, for some $\delta = \delta(\boldsymbol{x}, \Delta) < 1$.*

Theorem 3 is best possible in terms of maximum degree. Specifically, Sly and Sun [3] proved that approximating the partition function is NP-hard even when $\boldsymbol{x} = x\mathbf{1}$, and $x > x_c$.

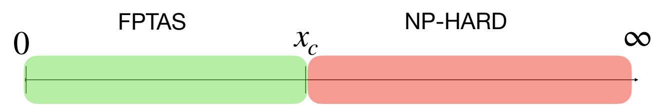

Figure 1.1: Characterization of approximability of Z_G for positive activities

1.2.2 Complex Numbers: Phase Transitions

The study of partition functions when the arguments of the corresponding polynomial are complex numbers dates back to the 1952 work of Lee and Yang [4] who established a connection between the location of zeros of the partition function on the complex plane and the presence of phase transitions on the real axis. The high-level idea is that since we identify phase transitions as discontinuities in the derivatives of free energy, i.e., $\log Z$, such a transition can only occur at a point of the complex plane if there is at least one nearby zero of the partition function. Specifically, in the follow-up paper [5], Lee and Yang instantiated this connection for the ferromagnetic Ising model by proving that the zeros of the partition function always lie on the unit circle in the complex plane, and using this fact to conclude that the ferromagnetic Ising model can have at most one phase transition. The Lee-Yang approach has since become a cornerstone of the study of phase transitions, and has been used extensively in the statistical physics literature: see, e.g., [6–9] for specific examples, and Ruelle's book [10] for background. There have also been attempts to relate the Lee-Yang program to the Riemann hypothesis [11].

Zeros of partition functions when the variables take complex values have also been studied in a purely combinatorial setting without reference to the physical interpretation: see, for example, Choe et al. [12]. Another important example is the work of

Chudnovsky and Seymour [13], who show that the zeros of the univariate independent set polynomial of claw-free graphs lie on the real line. Also, Barvinok [14] initiated a line of research leading to *quasi*-polynomial time approximation algorithms for several types of partition functions and graph polynomials. This approach, known as interpolation method, is based on Taylor approximations of the $\log Z$ allowing approximation in regions of the complex plane where Z is non-vanishing. Later, Patel and Regts [15] showed that the interpolation method, in fact, yields *polynomial* time approximation algorithms when restricted to bounded degree graphs. Finally, in a seminal work, Scott and Sokal [16] proved that the independent set polynomial is non-zero in the polydisk of $\boldsymbol{p} \in [0,1)^n$, if and only if $Z_G(-\lambda \boldsymbol{p}) > 0$ for every $\lambda \in [0,1]$.

1.2.3 The Probabilistic Method and the Lovász Local Lemma

The *Probabilistic Method* [17] amounts to establishing the existence of mathematical objects with a certain property of interest by demonstrating a probability distribution under which they have strictly positive probability. The power of the method stems from the fact that underestimating the probability of the objects by *any* factor (multiplicative approximation) suffices to establish existence. Typically, the property of interest, $\mathcal{P}$, is the intersection of the complements of several simpler properties, each property expressing some particular "flaw", so that $\mathcal{P}$ coincides with flawlessness. Thus, if we endow a universe of candidate objects Ω, where sets $\{F_i\}_{i=1}^n \subseteq \Omega$ correspond to different flaws, with a probability measure μ, we want to prove that the *avoidance probability*, $\mu(\bigcap_{i \in [n]} \overline{F}_i)$, is strictly positive.

Now, if we only know the marginals $p_i := \mu(F_i)$, the best lower bound we can give for the avoidance probability is $1 - \sum_i p_i$, as, for all we know, the flaws could be disjoint. To improve upon the union bound, we need to constrain the flaw overlaps. A natural (and extremely successful) way to express this is as a graph G on $[n]$. Concretely, let $\Gamma_i(G) = \Gamma_i$ denote the neighborhood of vertex i in G and let $\Gamma_i^+ = \Gamma_i \cup \{i\}$. We say that G is a *dependency graph* for $\{F_i\}_{i=1}^n$ with respect to μ if for every $i \in [n]$ and every set $\{j_1, j_2, \ldots\} \subseteq [n] \setminus \Gamma_i^+$,

$$\mu(F_i \mid \overline{F_{j_1}} \cap \overline{F_{j_2}} \cap \cdots) = \mu(F_i) = p_i \quad . \tag{1.3}$$

Note that the presence of an edge in G does not prescribe any specific kind of dependency between the corresponding events, only a lack of constraint thereof. Thus, a complete dependency graph (clique) conveys no information at all about how the n events overlap, while an empty dependency graph implies that the n events are mutually independent.

Say that a measure μ on $\{F_1, \ldots, F_n\}$ is *compatible* with $\boldsymbol{p}, G$, if $p_i = \mu(F_i)$, and G is a dependency graph for μ. For a measure μ it is typically easy to compute $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ and a dependency graph G. As a result, it is natural to seek a condition which given $\boldsymbol{p}, G$ determines if the avoidance probability is strictly positive for *every* measure compatible with $\boldsymbol{p}, G$, or not. This task was wonderfully tackled by Shearer in [1], who showed that to minimize the avoidance probability, given $\boldsymbol{p}, G$, one should try to realize the (unique) measure μ^* under which events adjacent in G are disjoint. Explicitly, μ^* is defined by demanding $\mu^*(\cap_{i \in S} F_i) = \prod_{i \in S} p_i$, if S is an independent set of G, and $\mu^*(\cap_{i \in S} F_i) = 0$, otherwise. It is not hard to see that if $Z(-\boldsymbol{p}; S) < 0$, for some $S \subseteq [n]$ (or, equivalently, if $Z(-\lambda \boldsymbol{p}; [n]) < 0$ for some $\lambda \in [0, 1]$), then μ^* can not be realized, and one can easily construct a probability measure compatible with $G, \boldsymbol{p}$, for which the avoidance probability is zero; otherwise, inclusion-exclusion implies that $\mu^*(\bigcap_{i \in S} \overline{F}_i) = Z(-\boldsymbol{p}; S) \geq 0$ and, therefore, that the avoidance probability is at least $Z(-\boldsymbol{p}; [n])$, i.e., the value of the independent set polynomial of G evaluated at $-\boldsymbol{p}$. Unfortunately, performing this evaluation is generally intractable, as it involves a summation over all independent sets of the dependency graph G.

The Lovász Local Lemma is a *sufficient* condition for $Z(-\lambda \boldsymbol{p}) > 0$ for all $\lambda \in [0, 1]$, along with a lower bound for $Z(-\boldsymbol{p})$. Below is a general formulation (the so-called *asymmetric*).

Theorem 4 (Lovász [18])**.** *Let μ be a probability measure on set Ω and let G be a dependency graph for $\{F_i\}_{i\in[n]} \subseteq \Omega$. If there exist $r_1, r_2, \ldots, r_n \in [0,1)$ such that for every $i \in [n]$,*

$$p_i \leq r_i \prod_{j\in\Gamma_i} (1 - r_j) \ , \tag{1.4}$$

then $\mu\left(\bigcap_{i=1}^{n} \overline{F_i}\right) \geq \prod_{i\in[n]} (1 - r_i) > 0$.

Remark 5. *Theorem 4 holds (and is known as the "Lopsided LLL" of Erdős and Spencer [19]) if condition* (1.3) *holds with "$\leq$" instead of "$=$". All our results also hold in that setting but we stick to "$=$" for simplicity of presentation.*

1.3 Our Results

1.3.1 An Improved General / Asymmetric LLL

We strictly improve the asymmetric LLL, as follows, and thus all its (several hundred) applications.

Theorem 6. *Theorem 4 holds if* (1.4) *is replaced by*

$$p_i \leq r_i \prod_{j \in \Gamma_i} \frac{1 - r_j}{1 - r_i r_j} \ . \tag{1.5}$$

Even though Theorem 6 retains all the flexibility of the asymmetric LLL to adjust to events with different degrees and probabilities, it is sharp enough to recover the *optimal* bound in terms of the maximum degree $\Delta(G)$, attained as a limit by Δ-regular trees as depth goes to infinity. Specifically, if every event has probability at most p and is mutually independent of all but $\Delta \geq 2$ other events, Theorem 6 recovers the optimal condition $p < \frac{(\Delta-1)^{(\Delta-1)}}{\Delta^\Delta}$ originally proven by Shearer [1], whereas the asymmetric LLL requires $p < \frac{\Delta^\Delta}{(\Delta+1)^{(\Delta+1)}}$.

A fairly recent improvement of the asymmetric LLL is the so-called *cluster expansion* LLL by Bissacot et al. [20], wherein the presence of edges in the neighborhood of a vertex, i.e., the presence of triangles in G, relaxes the condition corresponding to that vertex. In general, our Theorem 6 is incomparable with the cluster expansion LLL, but the overall trend is that the former is better when the neighborhoods are sparse, while the latter is better when they are dense.

In Section 1.3.4, we will see two significant improvements of Theorem 6. The *weaker* of these is already *exact* on *arbitrary* trees (uniform trees being the worst case for given Δ).

1.3.2 An Upper Bound for the Partition Function on the Negative Reals

Recall that given a vector $\boldsymbol{p} \in [0,1)^n$, the central problem is determining whether $Z_G \neq 0$ on the *polydisk* of $\boldsymbol{p}$, i.e., for every $\boldsymbol{x} \in \mathbb{C}^n$ such that $|x_i| \leq p_i$ for all $i \in [n]$.

Recall also that Scott and Sokal [16] showed that "the worst case" occurs when all arguments are negative real, so that the independent set does not vanish anywhere in the polydisk iff $Z(-\lambda\boldsymbol{p}) > 0$ for every $\lambda \in [0,1]$. Concomitantly, when this occurs, the magnitude of Z_G over the polydisk of $\boldsymbol{p}$ is minimized at $-\boldsymbol{p}$.

Definition 7. *Given a graph G, let*

$$\mathbb{S}(G) = \{\boldsymbol{p} \in [0,1)^n : Z_G(-\lambda\boldsymbol{p}) > 0, \text{ for all } \lambda \in [0,1]\} \ .$$

Now, recall that by the work of Shearer [1], if G is a dependency graph and $\boldsymbol{p} \in \mathbb{S}(G)$, then the avoidance probability of every measure compatible with $\boldsymbol{p}, G$ is at least $Z(-\boldsymbol{p}) > 0$, whereas if $\boldsymbol{p} \notin \mathbb{S}(G)$, there exists a measure compatible with $\boldsymbol{p}, G$ for which the avoidance probability is zero. Finally, recall that the LLL is a local sufficient condition for $\boldsymbol{p} \in \mathbb{S}(G)$, providing a (positive) lower bound for $Z(-\boldsymbol{p})$ (and, thus, for $|Z|$ on the polydisk of $\boldsymbol{p}$). We show that if $\boldsymbol{p} \in \mathbb{S}(G)$, then $|Z_G|$ can also be bounded *from above* on the entire polydisk of $\boldsymbol{p}$.

Definition 8. *Given a permutation π of $[n]$, let $\overleftarrow{\Gamma_i} = \overleftarrow{\Gamma_i}(\pi) = \Gamma_i \cap \{j \in [n] : \pi(j) < \pi(i)\}$ and let $\overrightarrow{\Gamma_i} = \overrightarrow{\Gamma_i}(\pi) = \Gamma_i \cap \{j \in [n] : \pi(j) > \pi(i)\}$.*

Theorem 9 (Upper Bound)**.** *Given $\boldsymbol{p}, G$ and any permutation π of $[n]$, define $\boldsymbol{r} = \boldsymbol{r}(\pi)$ by*

$$p_i = r_i \prod_{j \in \overleftarrow{\Gamma_i}(\pi)} (1 - r_j) \ , \quad \text{for every } i \in [n] \ . \tag{1.6}$$

(Note that $\boldsymbol{r}$ is well-defined as $r_1 = p_1$, while r_i is determined by $p_i, r_1, \ldots, r_{i-1}$ for $i > 1$.)

If $\boldsymbol{p} \in \mathbb{S}(G)$, then $Z(-\boldsymbol{p}; S) \le \prod_{j \in S}(1 - r_j)$, for every $S \subseteq [n]$.

Remark 10. *If $\boldsymbol{r'}, \boldsymbol{r}$ satisfy* (1.4), (1.6)*, respectively, then $r'_i \ge r_i$ for every $i \in [n]$.*

1.3.3 Exact Computation for Chordal Graphs

Recall that a graph is *chordal* if all its induced cycles have length three. We prove that the independent set polynomial of a chordal graph can be evaluated *anywhere* on the complex plane in *linear* time. We conjecture that chordality is closely related to the exact solvability of the hard-core model for certain highly transitive graphs, e.g., triangular lattice (hard hexagons model [21]), and that the hard-core model is not the only statistical mechanics model for which chordality relates to exact solvability.

Fact 1. *A graph G on $[n]$ is chordal iff there exists a permutation π of $[n]$ such that $\overrightarrow{\Gamma}_i(\pi)$ is a clique for every $i \in [n]$, in which case we say that π is a* chordal presentation *of G.*

Theorem 11. *If π is a chordal presentation of G, then $Z_G(\boldsymbol{x}) = \prod_{i\in[n]}(1+r_i)$, where*

$$x_i = r_i \prod_{j \in \overleftarrow{\Gamma}_i(\pi)} (1+r_j) \ , \quad \textit{for every } i \in [n] \ . \tag{1.7}$$

(Note that $\boldsymbol{r}$ is well-defined as $r_1 = x_1$, while for $i > 1$, r_i is given by $x_i, r_1, \ldots, r_{i-1}$.)

Corollary 12. *The independent set polynomial of a chordal graph can be evaluated anywhere on the complex plane in linear time. A perfect sample from the hard-core distribution on a chordal graph can be obtained in polynomial time.*

Proof. A chordal presentation of chordal graph $G = (V, E)$ can be found in time $O(|V| + |E|)$. Computing each r_i given $r_1, \ldots, r_{i-1}$ requires $O(|\Gamma_i|)$ steps. By evaluating $Z_G(\boldsymbol{x}, [n])$ and $Z_G((x_1, \ldots, x_{n-1}, 0), [n])$ we can sample the correct distribution for vertex n. Having done so, i.e., whether n is absent or present in our sample, we proceed to sample an independent set from the chordal graph induced by $[n-1]$ or $[n] \setminus \Gamma_i^+$, respectively. ∎

The previous best result on the independent set polynomial of chordal graphs is due to Okamoto, Uno, and Uehara [22] who showed that it can be evaluated exactly in linear time at $\boldsymbol{x} = \mathbf{1}$, i.e., that the number of independent sets can be counted. It is not clear that the approach of [22] can be extended to handle even the case $\boldsymbol{x} = x\mathbf{1}$ for real $x > 0$, let alone the multivariate, or (especially) the negative real case. A very recent related work by Heinrich and Müller [23] showed that the independent set polynomial can be evaluated exactly for $\boldsymbol{x} \in \mathbb{R}^n$, when G is strongly orderable. Stronly orderable graphs form a subclass of weakly chordal graphs that contains chordal bipartite graphs. Finally, in terms of (arbitrarily good, randomized) approximate evaluation of the independent set polynomial, Bezakova and Sun [24] showed that a natural Markov chain for the hard core model with positive fugacities, i.e., for the case $\boldsymbol{x} \in \mathbb{R}^n$, mixes in polynomial time on chordal graphs with separators of *bounded* size.

1.3.4 Local Lemmata from Non-Backtracking Walks

We develop a framework that allows us to derive local lemmata from different sets of walks on the graph G. Concretely, in each set the walks share the property that the next vertex in each step avoids the previous q, for some $q \geq 0$. In particular, Theorems 4 , 13, and 14 correspond to taking $\mathcal{T}_i$ to be the (infinite) tree of walks starting at i, that are arbitrary ($q = 0$), non-backtracking ($q = 1$), and non-2-backtracking, respectively. These trees preserve the local structure of G while being tractable, as they enjoy highly recursive structure (decreasingly so as q is increased).

A Local Lemma from Non-Backtracking Walks

Theorem 13. *Given $G = ([n], E)$ and $\boldsymbol{p} \in [0,1)^n$, assume that for every $i \in [n]$, and every ordered pair (i,j) where $\{i,j\} \in E$, there exist $r_i, r_{i,j}, \in [0,1)$, respectively, such that*

$$p_i \leq r_i \prod_{j \in \Gamma_i} (1 - r_{j,i}) \tag{1.8}$$

$$p_i \leq r_{i,j} \prod_{k \in \Gamma_i \setminus \{j\}} (1 - r_{k,i}) \ . \tag{1.9}$$

Then, $Z(-\boldsymbol{p}) \geq \prod_{i \in [n]} (1 - r_i)$.

Theorem 13 implies Theorem 6 and is implied by Theorem 14, below.

A Lower Bound from Non-2-Backtracking Walks

Theorem 14. *Given* $G = ([n], E)$ *and* $\boldsymbol{p} \in [0,1)^n$*, assume that for every* $i \in [n]$*, every ordered pair* (i,j) *where* $\{i,j\} \in E$*, and every ordered triple* (i,j,k) *where* $\{i,j,k\}$ *induce a triangle in* G*, there exist* $r_i, r_{i,j}, r_{i,j,k} \in [0,1)$*, respectively, such that*

$$p_i \le r_i \prod_{j \in \Gamma_i} (1 - r_{j,i}) \tag{1.10}$$

$$p_i \le r_{i,j} \prod_{k \in \Gamma_i \setminus \Gamma_j^+} (1 - r_{k,i}) \prod_{k \in \Gamma_i \cap \Gamma_j} (1 - r_{k,i,j}) \tag{1.11}$$

$$p_i \le r_{i,j,k} \prod_{\ell \in \Gamma_i \setminus \Gamma_j^+} (1 - r_{\ell,i}) \prod_{\ell \in \Gamma_i \cap \Gamma_j \setminus \{k\}} (1 - r_{\ell,i,j}) \ . \tag{1.12}$$

Then, $Z(-\boldsymbol{p}) \ge \prod_{i \in [n]} (1 - r_i)$.

1.3.5 Application: Shearer Region of the Triangular Lattice

As mentioned earlier, determining the set $\mathbb{S}(G)$ of activities for which the partition function is non-vanishing in the corresponding polydisk is a central problem in statistical mechanics, motivated by the Lee-Yang [4] approach to studying phase transitions. Since phase transitions can occur only in infinite-size systems, to study them on locally-finite countable graph G_∞ (typically a regular lattice), we consider an increasing sequence of subgraphs $(G_n)_{n\geq 1}$ converging to G_∞ and study the limiting free energy per vertex $f_{G_\infty} = \lim_{n\to\infty} n^{-1} \log Z_{G_n}(\boldsymbol{x})$. Nonanalyticities of f_{G_∞} for real $\boldsymbol{x}$, arise from singularities of $\log Z_{G_n}(\boldsymbol{x})$ for complex $\boldsymbol{x}$ that approach the real axis in the limit $n \to \infty$. But the singularities of $\log Z_{G_n}(\boldsymbol{x})$ are precisely the zeros of $Z_{G_n}(\boldsymbol{x})$, hence the desire to determine $\mathcal{S}(G)$. Of particular interest is the so-called *uniform* case $\boldsymbol{p} = p\mathbf{1}$, where all the activities are the same.

To benchmark our methods, we consider one of the very few exactly solved cases of the hard-core model, namely the case where G_∞ is the triangular lattice. This is known as the "hard hegaxons" model, since its valid configurations amount to placements of (centers of) hexagons in a triangular lattice so that no two hexagons overlap, i.e., to selecting an independent set of the triangular lattice (serving as the centers of the hexagons). For this model it is known that the critical value $p_c = \frac{5\sqrt{5}-11}{2} = 0.09016...$

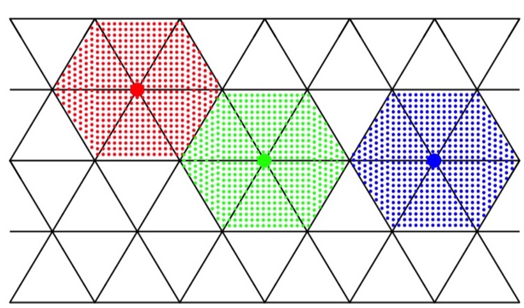

Figure 1.2: The "hard hegaxons" model: its valid configurations amount to placements of (centers of) hexagons in a triangular lattice so that no two hexagons overlap

Applying the asymmetric LLL, which only exploits that $\Delta(G_n) = 6$, implies $p_c \geq$

$6^6/7^7 = 0.0566$. Our improved asymmetric LLL (Theorem 6), refining the dependence on Δ, yields $p_c \geq 5^5/6^6 = 0.0669$. Kolipaka, Szegedy, and Xu [25], introduced a family of sufficient conditions for the avoidance probability to be positive that range between the asymmetric LLL and the exact result of Shearer [1]. To apply their so-called "clique LLL" to the triangular lattice we color the triangular faces in a chess board pattern and decompose it using the white triangles as the parts of the clique-decomposition. Optimizing the resulting parameters yields $p_c \geq 0.07407$. Finally, the cluster expansion LLL [20], exploiting the presence of 6 triangles in the neighborhood of each vertex, yields $p_c \geq 0.0776$.

Applying Theorem 14 with $r_i = 0.81614$, $r_{i,j} = 0.2$, and $r_{i,j,k} = 0.2$ yields $p_c \geq 0.08192$.

1.3.6 Recovering Known Approximability Results with Simpler Proofs

Recall that for positive real activities, Weitz [2] showed that the problem of evaluating the independent set polynomial admits a FPTAS when all activities are in the interval $[0, x_c)$, where $x_c \searrow \frac{\mathrm{e}}{\Delta}$ is defined in (1.2). Let

$$x_c^* = \frac{(\Delta-1)^{(\Delta-1)}}{\Delta^\Delta} \searrow \frac{1}{\mathrm{e}\Delta} \ .$$

For *complex* arguments, Patel and Regts [15], as well as, Harvey, Srivastava and Vondrák [26] showed that there is an FPTAS for every G of maximum degree Δ for activities inside the disk of modulus x^* centered at the origin. Later, Peters and Regts [27] extended the regime of approximability to a small strip surrounding the interval $[0, x_c)$, using the polynomial interpolation approach of Barvinok [14], confirming a conjecture of Scott and Sokal.

As we will see, for the independent set polynomial with (uniform) activities $x\mathbf{1}$ a key function is $f_x : \mathbb{C} \mapsto \mathbb{C}$ with $f_x(w) = x/(1+w)^{\Delta-1}$. In their paper [27], Peters and Regts defined a cardioid region on the complex plane as: the locus of x (activities) for which f_x has at least one attractive fixed point. They asked whether this cardioid

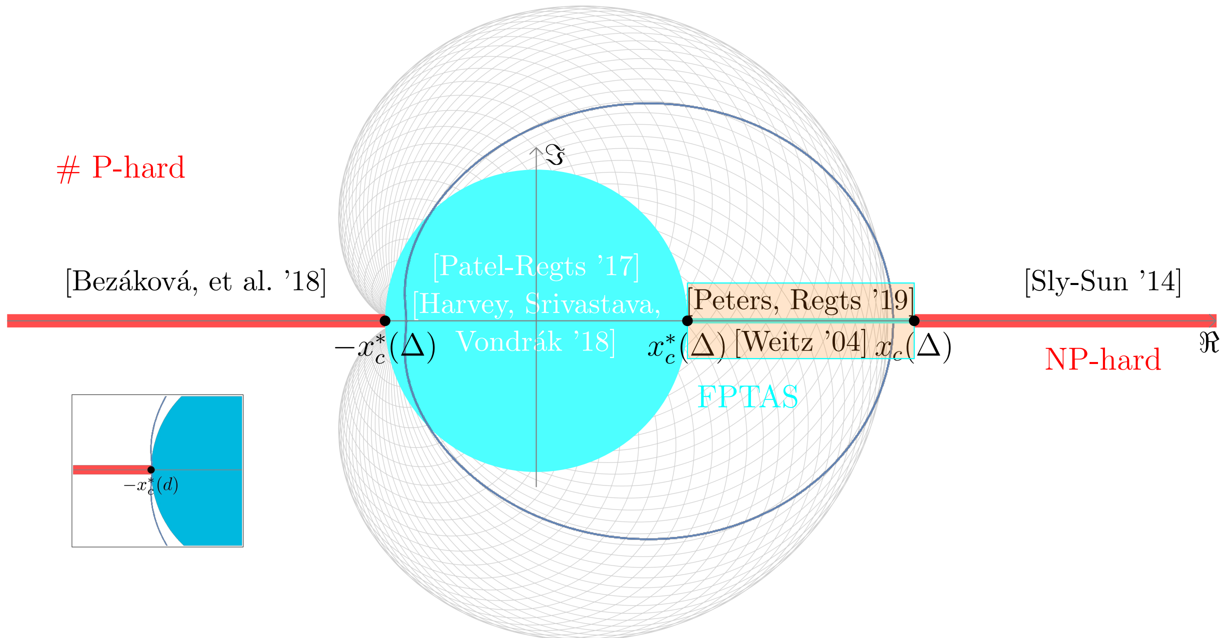

Figure 1.3: The locus of activities for which f_x has at least one attractive fixed point is the cardioid. The locus of activities for which $\widehat{f}_x$ has at least one attractive fixed point is the egg-shaped region.

region corresponds precisely to the region where efficient approximability is possible. In [28], Bezáková, Galanis, Goldberg, and Štefankovič showed that everywhere *outside* the cardioid region lies intractability: approximating the independent set polynomial is#P-hard, except on the real, positive axis where it is NP-hard.

We introduce analytic machinery that allows us to reproduce both the result of [2] and of [15, 26]. A positive aspect of our machinery is that the resulting proofs are *significantly shorter, entirely self-contained, and elementary*, i.e., do not appeal to notions from statistical physics such as correlation decay, or (sophisticated) polynomial interpolation.

An important feature of our machinery is that it operates in the *log-domain*, i.e., instead of the function f_x mentioned above, we analyze $\widehat{f}_x = \log \circ f_x \circ \exp$. Thus, for the result in [15, 26], the disk of radius x_c^* centered at the origin is transformed to the half-plane comprised of complex numbers whose real part is bounded above by $\log x_c^*$. Similarly, the cardioid locus of activities for which $\widehat{f}_x$ has at least one attractive fixed point becomes the egg-shaped region in Figure 1.3 that is nearly *convex* (see the relevant inset around $-x_c$.). We hope to exploit this near-convexity to establish efficient approximability for the independent set polynomial inside (most of) the egg-shaped region.

1.4 Relating the Independent Set Polynomial to Walk Trees

In this section we exploit the so called self-reducibility of the independent set polynomial, being the starting point of both [1] and [16]. Simply put, one can show that the ratios of Z, satisfy a simple recurrence, whose computation tree can be seen as a tree of appropriately defined walks on G. However, the emerging computation tree is in general exponential in the size of G, and hence no direct access to it is possible.

1.4.1 Main Recurrence and the Computation Tree

For $i \in [n]$, and $S \subseteq [n] \setminus \{i\}$, given input $\boldsymbol{x}$, we define

$$Z(\boldsymbol{x}; i \mid S) := \frac{Z(\boldsymbol{x}; S \cup \{i\})}{Z(\boldsymbol{x}; S)} = Z(i \mid S) \ .$$

Trivially, $Z = Z([n]) = \prod_{i\in[n]} Z(i \mid [i-1])$, since $Z(i, \emptyset) = 1$. To estimate $Z(i \mid S)$ observe that the contribution to $Z(S \cup \{i\})$ of the sets including vertex i equals x_i times the contribution of the sets not including $\Gamma^+(i)$. Therefore,

$$Z(S \cup \{i\}) = Z(S) + x_i Z(S \setminus \Gamma_i) \ . \tag{1.13}$$

With the above in mind, let $\{j_1, \ldots, j_d\}$ be an ordering of $\Gamma_i \cap S$, and write $S_\ell = S \setminus \{j_1, \ldots, j_\ell\}$. Dividing (1.13) by $Z(S)$ and writing the ratio $Z(S \setminus \Gamma_i)/Z(S)$ in telescopic form yields

$$Z(i \mid S) = 1 + x_i \frac{1}{\frac{Z(S)}{Z(S\setminus\Gamma_i)}} = 1 + x_i \frac{1}{\prod_{\ell=1}^{d} \frac{Z(S\setminus\{j_1,\ldots,j_{\ell-1}\})}{Z(S\setminus\{j_1,\ldots,j_\ell\})}} = 1 + x_i \prod_{\ell=1}^{d} \frac{1}{Z(j_\ell \mid S_\ell)} \ . \tag{1.14}$$

It is convenient to introdude the quantity $\text{ratio}_G(\boldsymbol{x}; (i, S)) := Z(i \mid S) - 1 = \text{ratio}(i, S)$ and rewrite (1.14) as

$$\text{ratio}(i, S) = x_i \prod_{\ell=1}^{d} \frac{1}{1 + \text{ratio}(j_\ell, S_\ell)} \ . \tag{1.15}$$

Thus,

$$Z = Z([n]) = \prod_{i \in [n]} Z(i \mid [i-1]) = \prod_{i \in [n]} (1 + \text{ratio}(i, [i-1])) \ . \tag{1.16}$$

To compute $\text{ratio}(i, [i-1])$ for each $i \in [n]$, we observe that since all sets on the right hand side of (1.15) are strictly smaller than S, the recursion (1.15) for $\text{ratio}(i, [i-1])$ unfolds to a computation tree of depth at most $i-1$, with leaves of the form $\text{ratio}(j, \emptyset) = x_j$.

1.4.2 From Arbitrary Graphs to Trees

A *walk* on G starting at vertex i is a sequence of vertices $(v_0, v_1, \ldots, v_\ell)$, such that $v_0 = i$ and v_{k-1} is adjacent to v_k for all $k \in [\ell]$.

Definition 15. *A walk $(v_0, v_1, \ldots)$ is q-non-backtracking if $v_j \notin \{v_{j-1}, v_{j-2}, \ldots, v_{j-q}\}$. A walk is* self-avoiding *if its vertices are* distinct*, i.e., it is $(|V|-1)$-non-backtracking. A walk is* descending *if $v_{k-1} > v_k$ for all $k \in [\ell]$.*

Definition 16. *Given a walk $w = (v_0, v_1, \ldots, v_\ell)$, let $\mathcal{F}_w(v_0) = \emptyset$, while for $k \in [\ell]$ let*

$$\mathcal{F}_w(v_k) = \mathcal{F}_w(v_{k-1}) \cup \{u \in \Gamma_{v_{k-1}} : u \geq v_{k-1}\} \ . \tag{1.17}$$

If $v_{k+1} \notin \mathcal{F}_w(v_k)$ for every $k \in [\ell]$, we say that the walk w is leftward*.*

Remark 17. *A descending walk is both leftward and self-avoiding.*

Given a graph G, we write $\overleftarrow{G}_i$ for the graph induced by $[i]$.

Definition 18. *Let $\mathcal{W}$ be a non-empty set of walks on G all starting at i, such that $w \in \mathcal{W}$ implies $w' \in \mathcal{W}$ for every prefix w' of w. The* tree *corresponding to set $\mathcal{W}$ has as its root the walk (i), while the children of each vertex (walk) are its extensions by one step. The* activity *of each vertex $(i, v_1, \ldots, v_\ell)$ of the tree is x_{v_ℓ}.*

We use $\mathcal{L}_i := \mathcal{L}_i(G)$ to denote the tree of leftward self-avoiding walks on $\overleftarrow{G}_i$ starting at i.

Definition 19. *If G is a rooted tree, we use* $\text{ratio}_G(i)$ *to denote the quantity* $\text{ratio}_G(i, T(i))$*, where $T(i)$ is the set of vertices in the subtree rooted at i other than i.*

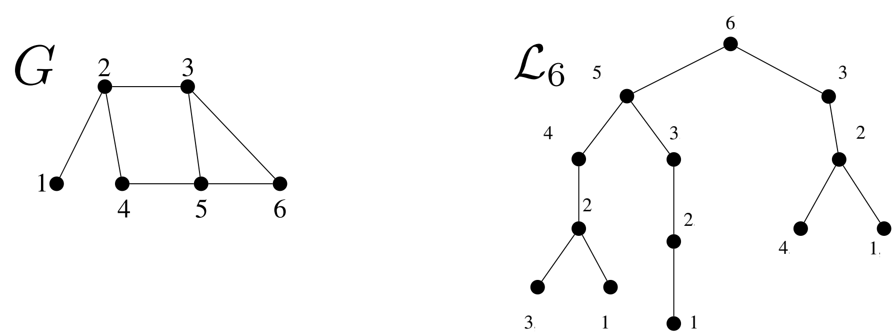

Figure 1.4: A graph G, and its leftward self-avoiding tree rooted at vertex 6

The following theorem shows that the independent set polynomial of any graph can be expressed as the independent set polynomial of some, possibly exponentially larger, tree.

Theorem 20. *For every graph G on $[n]$, $i \in [n]$, and walk $w = (v_0, v_1, \ldots, v_\ell)$ in $\mathcal{L}_i$,*

$$\text{ratio}_{\mathcal{L}_i}(w) = \text{ratio}_G\left(v_\ell, [i-1] \setminus \mathcal{F}_w(v_\ell)\right) \quad . \tag{1.18}$$

In particular, $\text{ratio}_{\mathcal{L}_i}\left((i)\right) = \text{ratio}_G\left(i, [i-1]\right)$.

Proof. We proceed by induction on the size of the subtree rooted at w.

If w is a leaf in $\mathcal{L}_i$ then, trivially, $\text{ratio}_{\mathcal{L}_i}(w) = z_{v_\ell}$. Moreover, $\mathcal{F}_w(v_\ell) \supseteq \Gamma_{v_\ell}$ (otherwise w could be extended), and thus, $\text{ratio}_G\left(v_\ell, [i-1] \setminus \mathcal{F}_w(v_\ell)\right) = \text{ratio}_G(v_\ell, \emptyset) = z_{v_\ell}$.

If w is not a leaf in $\mathcal{L}_i$, assume the theorem holds for its descendants. Let $\{j_1, \ldots, j_d\} = \Gamma_i \setminus \mathcal{F}_w(v_\ell)$, with $j_1 \geq \ldots \geq j_d$, and write $w_t = (v_0, v_1, \ldots, v_\ell, j_t)$, for $t \in [\ell]$

$$\operatorname{ratio}_{\mathcal{L}_i}(w) = \prod_{t\in[d]} \frac{1}{1+\operatorname{ratio}_{\mathcal{L}_i}(w_t)} \tag{1.19}$$

$$= \prod_{t\in[d]} \frac{1}{1+\operatorname{ratio}_G\left(j_t, [i-1] \setminus \mathcal{F}_{w_t}(j_t)\right)} \tag{1.20}$$

$$= \prod_{t\in[d]} \frac{1}{1+\operatorname{ratio}_G\left(j_t, ([i-1] \setminus \mathcal{F}_{w}(v_\ell)) \setminus \{j_1,\ldots,j_t\}\right)} \tag{1.21}$$

$$= \operatorname{ratio}_G\left(v_\ell, [i-1] \setminus \mathcal{F}_{w}(v_\ell)\right) \quad . \tag{1.22}$$

■

The following lemma establishes a strong monotonicity of ratios with respect to the activity vector in the negative axis.

Lemma 21. *Let* $i \in [n]$, $S \subseteq [n]\backslash\{i\}$, *and* $\boldsymbol{p} \in [0,1)^n$, *be such that* $\operatorname{ratio}(-\boldsymbol{p}; (j, S')) > -1$ *for every* $j \in S$ *and every* $S' \subseteq S \setminus \{j\}$. *Let* $P = \{(x_1,\ldots,x_n) : 0 \le x_i \le p_i\}$. *The function* $f : \boldsymbol{x} \mapsto \operatorname{ratio}(-\boldsymbol{x}; (i,S))$ *is smooth and strictly decreasing in* P.

Proof. We use induction on the size of S. If $S = \emptyset$ then, $\operatorname{ratio}(-\boldsymbol{x}; (i,\emptyset)) = -x_i$, satisfying the claim trivially. Assume now that the lemma holds for every $j \in S$ and every proper subset of S not containing j. Let $\{j_1,\ldots,j_d\}$ be an ordering of $\Gamma_i \cap S$, and write $S_\ell = S \setminus \{j_1,\ldots,j_\ell\}$; then (1.15) gives

$$\operatorname{ratio}(-\boldsymbol{x}; (i,S)) = -x_i \prod_{\ell=1}^{d} \frac{1}{1+\operatorname{ratio}(-\boldsymbol{x}; (j_\ell, S_\ell))} \quad . \tag{1.23}$$

Since each set S_ℓ that appears in (1.23) is a proper subset of S that does not contain j, the inductive hypothesis implies that $\operatorname{ratio}(-\boldsymbol{x}; (j_\ell, S_\ell))$ is decreasing inside P. Therefore, the assumed fact that $\operatorname{ratio}(-\boldsymbol{p}; (j_\ell, S_\ell)) > -1$ implies that $\operatorname{ratio}(-\boldsymbol{x}; (j_\ell, S_\ell)) > -1$. Since the function $1/(1+x)$ is smooth and decreasing for $x > -1$, the claim follows by the smoothness and monotonicity of the d factors in (1.23), afforded by the inductive hypothesis. ■

Next, we provide a simple characterization of $\mathbb{S}(G)$ with respect to ratios of $\mathcal{L}_i$.

Theorem 22. *$\boldsymbol{p} \in \mathbb{S}(G)$ iff* $\text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p}; w) > -1$*, for every $i \in [n]$, $w \in \mathcal{L}_i$.*

Proof. If $\text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p}; w) > -1$, for every $i \in [n]$ and $w \in \mathcal{L}_i$, then, by Lemma 21, $\text{ratio}_{\mathcal{L}_i}(-\lambda\boldsymbol{p}; (i)) \geq \text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p}; (i))) > -1$, for all $\lambda \in [0, 1]$. Thus, using telescoping and Theorem 20 to write the first equality, we conclude that $\boldsymbol{p} \in \mathbb{S}(G)$ since for all $\lambda \in [0, 1]$,

$$Z(-\lambda\boldsymbol{p}; [n]) = \prod_{i \in [n]} (1 + \text{ratio}_{\mathcal{L}_i}(-\lambda\boldsymbol{p}; (i))) \geq \prod_{i \in [n]} (1 + \text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p}; (i))) > 0 \ .$$

For the converse, we show that if $\text{ratio}\,(-\boldsymbol{p}; (i, S)) \leq -1$, for some $i \in [n]$, $S \subseteq [n] \setminus \{i\}$, then, there is $\lambda^* \in (0, 1]$ with $Z(-\lambda^*\boldsymbol{p}; [n]) = 0$, implying $\boldsymbol{p} \notin \mathbb{S}(G)$ (this is stronger than the contrapositive of what we need to prove as it does not require the set S to amount to the vertices of a subtree). Indeed, let S be any minimal set satisfying $\text{ratio}\,(-\boldsymbol{p}; (i, S)) \leq -1$; by renaming the vertices of G, we can assume that this happens for $S = [i - 1]$, i.e., that (a) $\text{ratio}\,(-\boldsymbol{p}; (i, [i - 1])) \leq -1$, while (b) $\text{ratio}\,(-\boldsymbol{p}; (j, T)) > -1$, for every $j \in [i-1]$ and $T \subseteq [i-1] \setminus \{j\}$. Per Lemma 21, fact (b) implies that $\text{ratio}\,(-\lambda\boldsymbol{p}; (i, [i - 1]))$ is a smooth and strictly decreasing function for $\lambda \in [0, 1]$. Hence, there exists $\lambda^* \in (0, 1]$ such that $\text{ratio}\,(-\lambda^*\boldsymbol{p}; (i, [i - 1])) = -1$, which per (1.16) yields $Z(-\lambda^*\boldsymbol{p}; [n]) = 0$. ■

Definition 23. *If T is a tree rooted at vertex i and we delete zero or more vertices of T other than i, then the remaining tree rooted at i is called a* prefix *of T.*

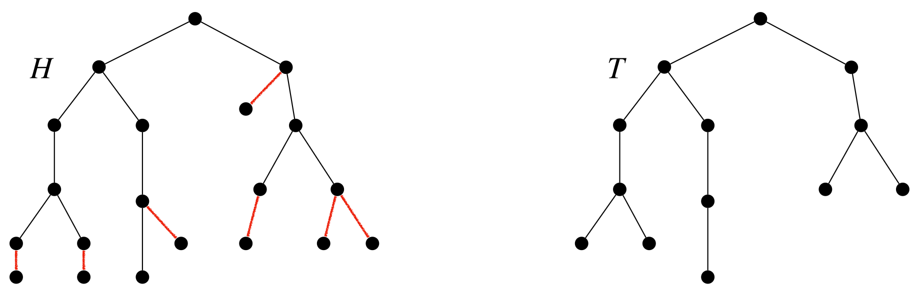

Figure 1.5: A prefix tree T obtained from the tree H after deleting its red edges

Corollary 24. *Let T_i be a tree on $[n]$ rooted at i and assume that $\boldsymbol{p} \in \mathbb{S}(T_i)$. If T_i' is any prefix of T_i, then $\text{ratio}_{T_i}(-\boldsymbol{p}; i) \leq \text{ratio}_{T_i'}(-\boldsymbol{p}; i)$.*

Proof. Zeroing a vertex activity p_j effectively deletes j. As $\boldsymbol{p} \in \mathbb{S}(T_i)$, Lemma 21 applies. ∎

1.5 Proofs of the Main Results

Using the tools we established above, we are now ready to prove our main results.

1.5.1 Upper Bound

Let $\mathcal{D}_i := \mathcal{D}_i(G)$ denote the tree of descending walks on G starting at i.

Theorem 25. *Given a graph G on $[n]$ and $\boldsymbol{x} \in \mathbb{C}^n$, for $i = 1, 2, \ldots, n$ let*

$$r_i = x_i \prod_{j \in \overleftarrow{\Gamma}_i} \frac{1}{1 + r_j} \ . \tag{1.24}$$

Then, $\text{ratio}_{\mathcal{D}_i}(\boldsymbol{x}; (i)) = r_i$ *for every $i \in [n]$.*

Proof. We use induction on i. For $i = 1$, trivially, $\text{ratio}_{\mathcal{D}_1}(\boldsymbol{x}; (1)) = x_1 = r_1$. Assume now that (1.24) holds for all $i < k$. Clearly, the root walk (k) can only be extended by taking a step to a neighbor smaller than k. If $\{j_1, \ldots, j_d\}$ is an ordering of $\overleftarrow{\Gamma}_k$, then (1.15) yields (1.25), while the inductive hypothesis yields the second equality in (1.27),

$$\text{ratio}_{\mathcal{D}_k}(\boldsymbol{x}; (k)) = x_k \prod_{\ell=1}^{d} \frac{1}{1 + \text{ratio}_{\mathcal{D}_k}(\boldsymbol{x}; (j_\ell, \mathcal{D}_k \setminus \{(k), (k, j_1), \ldots, (k, j_\ell)\}))} \tag{1.25}$$

$$= x_k \prod_{\ell=1}^{d} \frac{1}{1 + \text{ratio}_{\mathcal{D}_k}(\boldsymbol{x}; (j_\ell, \mathcal{D}_\ell \setminus \{(j_\ell)\}))} \tag{1.26}$$

$$= x_k \prod_{\ell=1}^{d} \frac{1}{1 + \text{ratio}_{\mathcal{D}_\ell}(\boldsymbol{x}; (j_\ell))} = x_i \prod_{j \in \overleftarrow{\Gamma}_i} \frac{1}{1 + r_j} = r_k \ . \tag{1.27}$$

■

Corollary 26. *If $\boldsymbol{p} \in \mathcal{S}(G)$, and $\boldsymbol{r} \in [0, 1)^n$ satisfies $r_i \prod_{j \in \overleftarrow{\Gamma}_i}(1 - r_j) = p_i$ for every $i \in [n]$, then $Z(-\boldsymbol{p}; S) \leq \prod_{i \in S}(1 - r_i)$.*

Proof. Equation (1.16) yields (1.28) and (1.29) follows rom Theorem 20. Recalling that a descending walk is both leftward and self-avoiding implies that $\mathcal{D}_i$ is a prefix of $\mathcal{L}_i$ and, thus, Corollary 24, implies $\text{ratio}_{\mathcal{D}_i}(-\boldsymbol{p}; (i)) \geq \text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p}; (i))$, yielding (1.30).

Finally, it is easy to see that our hypothesis is equivalent to $-r_i$ satisfying (1.24) for $\boldsymbol{x} = -\boldsymbol{p}$ so that claim 25, implies $\text{ratio}_{\mathcal{D}_i}(-\boldsymbol{p};(i)) = -r_i$ and, thus, (1.31). Putting everything together we get

$$
\begin{aligned}
Z(-\boldsymbol{p};S) &= \prod_{i\in[n]} (1 + \text{ratio}_G(-\boldsymbol{p};(i,[i-1]))) && (1.28)\\
&= \prod_{i\in[n]} (1 + \text{ratio}_{\mathcal{L}_i}(-\boldsymbol{p};(i)) && (1.29)\\
&\leq \prod_{i\in[n]} (1 + \text{ratio}_{\mathcal{D}_i}(-\boldsymbol{p};(i))) && (1.30)\\
&= (1 - r_i)\ . && (1.31)
\end{aligned}
$$

■

1.5.2 Proof of Theorem 6 from Theorem 13

We first show how Theorem 13 readily implies Theorem 6.

Proof. Given $\{r'_i\}_{i\in[n]}$, let $r_i = r'_i$ and $r_{i,j} = r'_i \frac{1-r'_j}{1-r'_i r'_j}$. We show that if (1.5) is satisfied, then (1.8) and (1.9) are satisfied. Indeed,

$$r_i \prod_{j\in\Gamma_i} (1 - r_{j,i}) = r'_i \prod_{j\in\Gamma_i} \left(1 - r'_j \frac{1-r'_i}{1-r'_j r'_i}\right) = r'_i \prod_{j\in\Gamma_i} \left(\frac{1-r'_j}{1-r'_j r'_i}\right) \geq p_i \ ,$$

and

$$r_{i,j} \prod_{k\in\Gamma_i\setminus\{j\}} (1 - r_{k,i}) = r'_i \frac{1-r'_j}{1-r'_i r'_j} \prod_{k\in\Gamma_i\setminus\{j\}} \left(1 - r'_k \frac{1-r'_i}{1-r'_k r'_i}\right) = r'_i \prod_{j\in\Gamma_i} \left(\frac{1-r'_j}{1-r'_j r'_i}\right) \geq p_i.$$

■

1.5.3 Proof of Theorems 4, 13, 14

Theorems 4 , 13, and 14 can be derived in a unified framework using walk trees. This is because, per Theorem 20 and Lemma 24, showing that $\boldsymbol{p} \in \mathbb{S}(\mathcal{T}_i)$ for an enlargement $\mathcal{T}_i$ of $\mathcal{L}_i$, implies $\boldsymbol{p} \in \mathbb{S}(G)$. Indeed, each of these theorems is the result of considering the (infinite) tree of all walks starting at i, which additionally:

- In the case of Theorem 4, are non-0-backtracking, i.e., have no restriction.
- In the case of Theorem 13, are non-1-backtracking, i.e., non-backtracking.
- In the case of Theorem 14, are non-2-backtracking.

The value of considering these tress is that, being infinite, they have highly recursive structure which can be characterized in terms of fixed-point equations. Specifically, as q is increased and more and more restrictions are placed on the walks, the amount of the graph's local structure preserved is increased, but so is the complexity of the recursive characterization (and, corresponding, fixed-point equations). In the interest of compactness, instead of writing our proofs via this intuitive tree approach (which motivated Theorems 13 and 14), we show how Theorem 14 yields Theorem 13 at the algebraic level and, similarly, for Theorem 13 yielding Theorem 6. For Theorem 14 itself, instead of the longer tree proof we give a shorter but rather less transparent inductive proof.

1.5.4 Proof of Theorem 13 from Theorem 14

Proof. Given $\{r'_i\}_{i\in[n]} \in [0,1)$ and $\{r'_{i,j}\}_{\{i,j\}\in E} \in [0,1)$, let

$$r_i = r'_i, r_{i,j} = r'_{i,j}, \text{ and } r_{i,j,k} = r'_{i,j}\frac{r'_{j,k}r'_{j,i} - r'_{k,i} + 1}{r'_{i,j}r'_{j,k}r'_{k,i} + 1} \ . \tag{1.32}$$

We claim that if (1.8) and (1.9) are satisfied, then (1.10), (1.11), and (1.12), are satisfied, implying

$$Z(-\boldsymbol{p}) \geq \prod_{i\in[n]}(1-r_i) = \prod_{i\in[n]}(1-r'_i) \ .$$

Indeed, $r_i \prod_{j\in\Gamma_i}(1-r_{j,i}) = r'_i \prod_{j\in\Gamma_i}\left(1-r'_{j,i}\right) \geq p_i$.

Using that $\{r'_{i,j}\}_{\{i,j\}\in E} \in [0,1)$ we see that $r_{i,j,k} \leq r'_{i,j}$, for all (i,j,k). Thus,

$$\begin{aligned}
& r_{i,j} \prod_{k\in\Gamma_i\setminus\Gamma_j^+}(1-r_{k,i}) \prod_{k\in\Gamma_i\cap\Gamma_j}(1-r_{k,i,j}) \\
\geq\ & r'_{i,j} \prod_{k\in\Gamma_i\setminus\Gamma_j^+}(1-r'_{k,i}) \prod_{k\in\Gamma_i\cap\Gamma_j}(1-r'_{k,i}) \\
=\ & r'_{i,j} \prod_{k\in\Gamma_i\setminus\{j\}}(1-r'_{k,i}) \\
\geq\ & p_i \ .
\end{aligned}$$

Similarly,

$$\begin{aligned}
& r_{i,j} \prod_{\ell\in\Gamma_i\setminus\Gamma_j^+}(1-r_{\ell,i}) \prod_{\ell\in\Gamma_i\cap\Gamma_j\setminus\{k\}}(1-r_{\ell,i,j}) \\
\geq\ & r'_{i,j} \prod_{\ell\in\Gamma_i\setminus\Gamma_j^+}(1-r'_{\ell,i}) \prod_{\ell\in\Gamma_i\cap\Gamma_j\setminus\{k\}}(1-r'_{\ell,i}) \\
=\ & r'_{i,j} \prod_{\ell\in\Gamma_i\setminus\{j,k\}}(1-r'_{\ell,i}) \\
\geq\ & r'_{i,j} \prod_{\ell\in\Gamma_i\setminus\{j\}}(1-r'_{\ell,i}) \\
\geq\ & p_i \ .
\end{aligned}$$

■

1.5.5 Proof of Theorem 14

Proof. We will show the following for every $i \in [n]$ and $S \subseteq [n] \setminus \{i\}$, by induction on $|S|$:

(a) $\text{ratio}(-\boldsymbol{p}; (i, S)) \geq -r_{i,j,k}$, for every ordered triple (i, j, k) where $\{i, j, k\}$ induce a triangle in G and $S \subseteq [n] \setminus \{i, j, k\}$.

(b) $\text{ratio}(-\boldsymbol{p}; (i, S)) \geq -r_{i,j}$, for every ordered pair (i, j) where $\{i, j\} \in E$ and $S \subseteq [n] \setminus \{i, j\}$.

(c) $\text{ratio}(-\boldsymbol{p}; (i, S)) \geq -r_i$, for every $i \in [n]$ and $S \subseteq [n] \setminus \{i\}$.

For $S = \emptyset$, $\text{ratio}(i, S) = -p_i \geq \max\{-r_{i,j,k}, -r_{i,j}, -r_i\}$, for all $\{i, j, k\} \subseteq [n]$. For the inductive step, let $S \subseteq [n] \setminus \{i\}$, and assume (a), (b), (c) hold for all proper subsets of S.

(a) Let $\{j_1, \ldots, j_q\} = S \cap \Gamma_i$, and write $S_\ell = S \setminus \{j_1, \ldots, j_\ell\}$, for $\ell \in [q]$. Per (1.15)

$$\text{ratio}(i, S) = -p_i \prod_{\ell=1}^{q} \frac{1}{1 + \text{ratio}(j_\ell, S_\ell)} \geq -p_i \prod_{\ell=1}^{q} \frac{1}{1 - r_{j_\ell}} \geq p_i \prod_{j \in \Gamma_i} \frac{1}{1 - r_j} \geq r_i \ .$$

In the following, for $(i, j) \in [n] \times [n]$ where $\{i, j\} \in E$, and $S \subseteq [n] \setminus \{i, j\}$, let $\{j_1, \ldots, j_q\} = S \cap (\Gamma_i \setminus \Gamma_j)$, $\{j_{q+1}, \ldots, j_d\} = S \cap (\Gamma_i \cap \Gamma_j)$, and write $S_\ell = S \setminus \{j_1, \ldots, j_\ell\}$, for $\ell \in [d]$.

(b) Let $S \subseteq [n] \setminus \{i, j\}$, and recall that (a), (b), (c) hold for all proper subsets of S. Per (1.15),

$$\text{ratio}(i, S) = -p_i \prod_{\ell=1}^{d} \frac{1}{1 + \text{ratio}(j_\ell, S_\ell)} \tag{1.33}$$

$$= -p_i \prod_{\ell=1}^{q} \left(\frac{1}{1 + \text{ratio}(j_\ell, S_\ell)} \right) \prod_{\ell=q+1}^{d} \left(\frac{1}{1 + \text{ratio}(j_\ell, S_\ell)} \right) \tag{1.34}$$

$$\geq -p_i \prod_{\ell=1}^{q} \left(\frac{1}{1 - r_{\ell,i}} \right) \prod_{\ell=q+1}^{d} \left(\frac{1}{1 - r_{\ell,i,j}} \right) \tag{1.35}$$

$$\geq -p_i \prod_{\Gamma_i \setminus \Gamma_j^+} \left(\frac{1}{1 - r_{\ell,i}} \right) \prod_{\Gamma_i \cap \Gamma_j} \left(\frac{1}{1 - r_{\ell,i,j}} \right) \geq r_{i,j} \ . \tag{1.36}$$

(c) Let $S \subseteq [n] \setminus \{i, j, k\}$, and recall that (a), (b), (c) hold for all proper subsets of S. Per (1.15),

$$\begin{aligned}
\mathrm{ratio}(i, S) &= -p_i \prod_{\ell=1}^{d} \frac{1}{1 + \mathrm{ratio}(j_\ell, S_\ell)} && (1.37)\\
&= -p_i \prod_{\ell=1}^{q} \left(\frac{1}{1 + \mathrm{ratio}(j_\ell, S_\ell)} \right) \prod_{\ell=q+1}^{d} \left(\frac{1}{1 + \mathrm{ratio}(j_\ell, S_\ell)} \right) && (1.38)\\
&\geq -p_i \prod_{\ell=1}^{q} \left(\frac{1}{1 - r_{\ell,i}} \right) \prod_{\ell=q+1}^{d} \left(\frac{1}{1 - r_{\ell,i,j}} \right) && (1.39)\\
&\geq -p_i \prod_{\Gamma_i \setminus \Gamma_j^+} \left(\frac{1}{1 - r_{\ell,i}} \right) \prod_{\Gamma_i \cap \Gamma_j \setminus \{k\}} \left(\frac{1}{1 - r_{\ell,i,j}} \right) \geq r_{i,j,k} \ . && (1.40)
\end{aligned}$$

■

1.5.6 Chordal Graphs

We claim that if G is chordally presented, then the set of leftward self-avoiding walks coincides with the set of decreasing walks. Theorem 11 then follows from Theorem 25. To prove the claim, as decreasing walks are leftward and self-avoiding, the following (more than) suffices.

Theorem 27. *G is chordally presented iff for every vertex i, every leftward self-avoiding walk on $\overleftarrow{G}_i$ starting at i is descending.*

Proof. We prove the contrapositive statement of both directions. Let $(i =: v_0, v_1, \ldots, v_\ell)$ be a leftward self-avoiding walk on $\overleftarrow{G}_i$ and let $2 \leq k \leq \ell$ be the minimum index such that $v_{k-1} < v_k$. The minimality of k implies $v_{k-1} < v_{k-2}$ and, hence, that $v_k, v_{k-2} \in \overrightarrow{\Gamma}_{v_{k-1}}$. Leftwardness implies that $v_k \notin \Gamma_{v_{k-2}}$, i.e., that there is no edge between v_{k-2} and v_k, contradicting that G is chordally presented.

If G is not chordally presented, then there exist vertices $a < b < c$ such that a is connected to b and c, but b is not connected to c. Clearly, the walk (c, a, b) on $\overleftarrow{G}_c$ is leftward and self-avoiding but not descending. ■

1.6 Taming the Computation Tree

In this section we devise analytical material for controlling the computation tree that results by unfolding the fundamental recurrence by treating it as a circuit. Thus, the the inputs to each gate come from its children in the tree, while its output travels upwards to its parent.

1.6.1 Circuits, Contraction, and Truncation

For arbitrary $D \subseteq \mathbb{C}^n$, write $\operatorname{diam}(D) := \sup\{\|\boldsymbol{w} - \boldsymbol{w}'\|_\infty : \boldsymbol{w}, \boldsymbol{w}' \in D\}$.

Definition 28. *A function F is a δ-*contraction on a set D *if* $|F(\boldsymbol{w}) - F(\boldsymbol{w}')| \leq \delta\|\boldsymbol{w} - \boldsymbol{w}'\|_\infty$*, for all $\boldsymbol{w}, \boldsymbol{w}' \in D$.*

Definition 29. *Let $D \subseteq \mathbb{C}$ and let $\mathcal{H}$ be a set of functions, each $f \in \mathcal{H}$ mapping D^d to D for some $d = d(f)$. We say that each $F \in \mathcal{H}$ is a* circuit of depth 1 on $\mathcal{H}$*, or* gate*. For $k \geq 1$, a function that results by replacing one or more inputs of a function $f \in \mathcal{H}$ with the outputs of depth-k circuits is a circuit of depth $k+1$ on $\mathcal{H}$.*

Lemma 30. *Let C be a depth-k circuit on $\mathcal{H}$ and let $\boldsymbol{w}, \boldsymbol{w}'$ be inputs to C that only differ in coordinates at depth k. If every gate in $\mathcal{H}$ is a δ-contraction, then $|C(\boldsymbol{w}) - C(\boldsymbol{w}')| \leq \delta^k \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty$.*

Proof. We use induction on k. For $k = 1$ the proposition is a tautology. Assume that our claim holds for some $k \geq 1$, and let C be a depth-$(k+1)$ circuit $C(\boldsymbol{z}) = F\left(C_1(\boldsymbol{z}_1), \ldots, C_d(\boldsymbol{z}_d)\right)$. Since $C_1, \ldots, C_d$ are depth-k circuits whose inputs differ only in coordinates at depth k,

$$\begin{aligned}
|C(\boldsymbol{z}) - C(\boldsymbol{z}')| &= \left|F\left(C_1(\boldsymbol{z}_1), \ldots, C_d(\boldsymbol{z}_d)\right) - F\left(C_1(\boldsymbol{z}'_1), \ldots, C_d(\boldsymbol{z}'_d)\right)\right| \\
&\leq \delta \left\|\left(C_1(\boldsymbol{z}_1), \ldots, C_d(\boldsymbol{z}_d)\right) - \left(C_1(\boldsymbol{z}'_1), \ldots, C_d(\boldsymbol{z}'_d)\right)\right\|_\infty \\
&= \delta \max\{\left|C_1(\boldsymbol{z}_1) - C_1(\boldsymbol{z}'_1)\right|, \ldots, \left|C_d(\boldsymbol{z}_d) - C_d(\boldsymbol{z}'_d)\right|\} \\
&\leq \delta \max\{\delta^k \left\|\boldsymbol{z}_1 - \boldsymbol{z}'_1\right\|_\infty, \ldots, \delta^k \left\|\boldsymbol{z}_d - \boldsymbol{z}'_d\right\|_\infty\} \\
&= \delta^{k+1} \|\boldsymbol{z} - \boldsymbol{z}'\|_\infty \ .
\end{aligned}$$

■

Theorem 31. *Assume that every gate in $\mathcal{H}$ is a δ-contraction for some $\delta < 1$. Given any circuit C on $\mathcal{H}$ and any $\varepsilon > 0$, let T be the truncation of C at depth $\ell = \lceil \log_{1/\delta}(\mathrm{diam}(D)/\varepsilon) \rceil$. If $\boldsymbol{w}$ is any input to C and $\boldsymbol{r}$ is any input to T that agrees with $\boldsymbol{w}$ in all coordinates at depth strictly less than k, then $|C(\boldsymbol{w}) - T(\boldsymbol{r})| < \varepsilon$.*

Proof. Trivially, there is an input $\boldsymbol{r}'$ to T that only differs from $\boldsymbol{r}$ in coordinates at depth k for which $C(\boldsymbol{w}) = T(\boldsymbol{r}')$. By Lemma 30, $|T(\boldsymbol{r}) - T(\boldsymbol{r}')| \leq \delta^\ell \cdot \|\boldsymbol{r} - \boldsymbol{r}'\|_\infty \leq \delta^\ell \cdot \mathrm{diam}(A)$. ■

1.6.2 Tools for Establishing Contraction

The following basic observations are useful for establishing contraction. Their proofs are given in Appendix **??**.

Bounded derivative implies contraction

Lemma 32. *Let $F : \mathbb{C}^n \mapsto \mathbb{C}$. A holomorphic function F is a δ-contraction on set D iff $\sup_{\boldsymbol{w} \in D} |\nabla F(\boldsymbol{w})|_1 \leq \delta$.*

Proof.

From the mean-value theorem for multiple variables, for any points $\boldsymbol{w}, \boldsymbol{w'} \in S \subseteq F^n$ there exist $c \in [0,1]$ such that

$$F(\boldsymbol{w}) - F(\boldsymbol{w}') = \nabla F\left((1-c)\boldsymbol{w} + c\boldsymbol{w}'\right) \cdot (\boldsymbol{w} - \boldsymbol{w}') \ .$$

Hence if F is not δ-contractive on a pair $\boldsymbol{w}, \boldsymbol{w'}$, there exists $\boldsymbol{w''}$ in the segment connecting $\boldsymbol{w}$ and $\boldsymbol{w'}$ with $\nabla F(\boldsymbol{w''}) \geq \delta$. For the other direction, Hölder's inequality gives the first inequality below

$$\|F(\boldsymbol{w}) - F(\boldsymbol{w}')\| \leq \|\nabla F(\boldsymbol{w})\|_1 \cdot \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \leq \sup_{\boldsymbol{w} \in S} \|\nabla F(\boldsymbol{w})\|_1 \cdot \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \leq \delta \cdot \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \ .$$

■

Let $f : A \subseteq \mathbb{C} \mapsto \mathbb{C}$, and fix $d \geq 1$. For $\zeta \in \mathbb{C}$ and $\boldsymbol{a} \in \mathbb{C}^d$ such that $|\boldsymbol{a}|_1 \leq 1$, let $F_\zeta : A^d \mapsto \mathbb{C}$ with

$$F_\zeta(\boldsymbol{w}) = \zeta + a_1 f(w_1) + \cdots + a_d f(w_d) \ .$$

Lemma 33. $\sup_{\boldsymbol{w} \in A^d} |\nabla F_\zeta(\boldsymbol{w})|_1 \leq \sup_{w \in A} |f'(w)|$.

Proof. Linearity of derivative gives

$$\begin{aligned}
\sup_{\boldsymbol{w} \in A^d} |\nabla F_\zeta(\boldsymbol{w})|_1 &= \sup_{\boldsymbol{w} \in A^d} \sum_{k=1}^{d} \left| \frac{\partial F_\zeta}{\partial w_i} \right| \\
&= \sup_{\boldsymbol{w} \in A^d} \sum_{k=1}^{d} |a_k| \cdot |f'(w_i)| \\
&= \sum_{k=1}^{d} |a_k| \cdot \sup_{w \in A} |f'(w)| \\
&\leq \sup_{w \in A} |f'(w)| \ .
\end{aligned}$$

■

Contraction of the second iteration of a function that is decreasing and concave

Let $f : A \subseteq \mathbb{R} \mapsto \mathbb{R}$, and fix $d \geq 1$. For $\lambda \in \mathbb{R}$, $\boldsymbol{a} \in \mathbb{R}^d_{\geq 0}$ with $\sum_{i=1}^{d} a_i \leq 1$, let $F_\lambda : A^d \mapsto \mathbb{R}$ with

$$F_\lambda(\boldsymbol{w}) = \lambda + a_1 f(w_1) + \cdots + a_d f(w_d) \ .$$

Write $\mathcal{F}_\gamma = \{F_\lambda : \lambda < \gamma\}$.

Lemma 34. *If $f', f'' < 0$ and $F^{(2)}_\gamma$ is a δ-contraction on set D, then so is every depth-2 circuit on $\mathcal{F}_\gamma$.*

Proof. If C^2 is any depth-2 circuit on $\mathcal{F}_\gamma$, then there exist $\gamma_0, \gamma_1, \ldots, \gamma_d < \gamma$, such that

$$C^2(\boldsymbol{x}) = \gamma_0 + \sum_{i=1}^{d} a_i g\left(\gamma_i + \sum_{j=1}^{d} a_j g(x_{ij})\right) \ .$$

Given $\boldsymbol{x}, \boldsymbol{x}' \in E$, let

$$\boldsymbol{s} = (\min\{x_{11}, x'_{11}\}, \ldots, \min\{x_{dd}, x'_{dd}\}) \quad ,$$

and

$$\boldsymbol{t} = (\max\{x_{11}, x'_{11}\}, \ldots, \max\{x_{dd}, x'_{dd}\}) \quad .$$

Since $g' < 0$, we see that C^2 is increasing in every coordinate, implying

$$\left|C^2(\boldsymbol{x}) - C^2(\boldsymbol{x}')\right| \leq C^2(\boldsymbol{t}) - C^2(\boldsymbol{s}) \quad .$$

Moreover, letting $C^2(\boldsymbol{t}) - C^2(\boldsymbol{s}) := H(\gamma_1, \ldots, \gamma_d)$ we see that, since $g'' < 0$, for all $i \in [d]$,

$$\frac{\partial H}{\partial \gamma_i} = \sum_{i=1}^{d} a_i \left[g'\left(\gamma_i + \sum_{j=1}^{d} a_j g(t_{ij})\right) - g'\left(\gamma_i + \sum_{j=1}^{d} a_j g(s_{ij})\right)\right] \geq 0 \quad .$$

Therefore,

$$\begin{aligned}
\left|C^2(\boldsymbol{x}) - C^2(\boldsymbol{x}')\right| &\leq C^2(\boldsymbol{t}) - C^2(\boldsymbol{s}) \\
&\leq F_\gamma^{(2)}(\boldsymbol{t}) - F_\gamma^{(2)}(\boldsymbol{s}) \\
&\leq \delta\|\boldsymbol{t} - \boldsymbol{s}\|_\infty \\
&= \delta\|\boldsymbol{x} - \boldsymbol{x}'\|_\infty \quad .
\end{aligned}$$

■

1.6.3 Passing to the Log-domain and Hardcoding the Activities

Since our main tool, Theorem 31, gives a bound on the *difference* between circuit outputs corresponding to different inputs (assuming contraction), in order to use it for the computation tree of the independent set polynomial we need to switch to the log-domain. Recall that $\mathrm{ratio}(\boldsymbol{x}; (i, S)) := Z(i \mid S) - 1$. Since we will need to consider iterates of our functions, to normalize the range of the function with the domain, we define the quantity

$$\mathrm{des}(\boldsymbol{x}; (i, S)) := \log\left[\mathrm{ratio}(\exp(\boldsymbol{x}); (i, S))\right] = \mathrm{des}(i, S)\ \ . \tag{1.41}$$

Expressed in terms of $\mathrm{des}(i, S)$, the fundamental recurrence,

$$\mathrm{ratio}(i, S) = x_i \prod_{\ell=1}^{d} (1 + \mathrm{ratio}(j_\ell, S_\ell))^{-1}\ \ ,$$

is

$$\mathrm{des}(i, S) = x_i - \sum_{i=1}^{d} \log\left(1 + \exp\left(\mathrm{des}(j_\ell, S_\ell)\right)\right)\ \ . \tag{1.42}$$

Exactly as (1.15), the recursion (1.42) for each $\mathrm{des}(i, [i-1])$, where $i \in [n]$, unfolds to a computation tree of depth at most $i-1$ with leaves of the form $\mathrm{des}(j, \emptyset) = x_j$, as all sets on the right hand side of (1.42) are strictly smaller than S. Besides working in the log-domain, i.e., with (1.42) instead of (1.15), it will also be convenient to *hard-code* the argument $\boldsymbol{x}$ of the independent set polynomial by making its components part of the specification of the different gates. Thus, different arguments cause the specification of the gates to change, while the structure of the computation tree (mandated by the graph G) stays the same. Specifically, let $D = \{(2k+1)\pi i : k \in \mathbb{Z}\}$ and recall that $\exp(w) = -1$ iff $w \in D$. For each $i \in [n]$ we define $f_i : (\mathbb{C} \setminus D)^{|\Gamma_i|} \mapsto \mathbb{C}$ with

$$f_i(\boldsymbol{w}) := x_i - \sum_{j \in \Gamma_i} \log\left(1 + \exp(w_j)\right)\ \ . \tag{1.43}$$

Letting $\mathcal{H} = \{f_1, \ldots, f_n\}$, we see that $\mathrm{des}(i, [i-1])$ can be realized as a depth-i circuit on $\mathcal{H}$, where every occurrence of an input x_i as a leaf of the computation tree is realized via a gate f_i having all its inputs set to $-\infty$ (causing the sum in the definition of f_i to vanish and, thus, the output to equal x_i). In other words, to compute $\mathrm{des}(i, [i-1])$ we evaluate a circuit all whose inputs equal $-\infty$.

1.6.4 Truncating the Tree

To additively approximate $\log Z(\exp(\boldsymbol{x}))$ within $\zeta > 0$, it suffices to approximate $\log(1 + \exp(\mathrm{des}(i, [i-1])))$ for each $i \in [n]$ within ζ/n. Observing that $|\log(1 + e^{x+\eta}) - \log(1 + e^{x})| < \eta$, it follows that it suffices to approximate each circuit computing $\mathrm{des}(i, [i-1])$ within an additive error of ζ/n. If each gate of the circuit is a δ-contraction, then by Theorem 31, this approximation can be achieved by truncating each such circuit at depth $\lceil \log_{1/\delta}(\mathrm{diam}(D)\,(n/\zeta)) \rceil$. If the maximum degree in G is Δ, then each truncated circuit will have $O(n \cdot \mathrm{diam}(D)/\zeta)^{(\log \Delta)/(1-\delta)}$ gates.

1.7 Positive Real Activities

We will recover the celebrated, optimal result of Weitz [2], below, without appealing to any (correlation decay) result from statistical physics, making the proof entirely self-contained and significantly shorter than the original. The key idea in our proof is that while the individual gates of the computation tree do not contract on the positive reals, treating pairs of successive layers of gates as a single layer of macro-gates yields contraction. In other words, we show that working with the *second iterate* of the recursion suffices. Given an integer $\Delta \geq 3$, let

$$\widehat{x}_c(\Delta) = \log(x_c(\Delta)) = \log\left(\frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}\right) .$$

In the log-domain, the approximability result of Weitz [2], i.e., Theorem 3, becomes the following.

Theorem 35 ([2]). *There exists an algorithm which given a graph of maximum degree Δ and activities $\boldsymbol{x} \in (-\infty, \widehat{x}_c(\Delta))^n$, returns $Y \in [\log Z - \varepsilon, \log Z + \varepsilon]$ in time $O(n/\varepsilon)^{\log \Delta/(1-\delta)}$, for some $\delta = \delta(\boldsymbol{x}, \Delta) < 1$.*

Let $d = \Delta - 1$. Let $f(w) = -d \log(1 + \exp(w))$. For $x \in \mathbb{C}$, let

$$F_x(\boldsymbol{w}) := x + \sum_{i=1}^{d} \frac{1}{d} f(w_i) = x - \sum_{i=1}^{d} \log(1 + \exp(w_i)) \quad . \tag{1.44}$$

Theorem 31 allows us to derive Theorem 35 immediately from the following analytical fact.

Lemma 36. *If $x < \widehat{x}_c$, then $F_x^{(2)}$ is a contraction on $\mathbb{R}$.*

Proof of Theorem 35. Without loss of generality we can assume that the computation tree (circuit) is regular, full, and has an even number of layers, since if that's not the case we can always add dummy "$-\infty$" gates as needed to make it so (with "$-\infty$" inputs when such gates are added as leaves). The key point is to then "redraw" the circuit, by grouping together as single gates the gates in pairs of successive layers.

Since, clearly, $f', f'' < 0$, if x is such that $F_x^{(2)}$ is a contraction, Lemma 34 implies that any depth-2 circuit with gates from $\{F_\mu : \mu < x\}$ is a contraction as well. Thus, Lemma 36 and Theorem 31 yield the result. ■

To prove Lemma 36, we observe that the permutation symmetry of F_x with respect to its arguments implies that $F_x^{(2)}$ is a contraction on $\mathbb{R}^d$ iff $f_x^{(2)}$ is a contraction on $\mathbb{R}$, where $f_x(w) = x - d \log(1 + \exp(w))$. Thus, by Lemma 32, it suffices to establish $|f_x^{(2)\prime}(w)| < 1$ for all $w \in \mathbb{R}$ and $x < \widehat{x}_c$. This is readily achieved by Lemmata 37, 38 below, whose proof appears in Appendix 35.

Lemma 37. *$f_x^{(2)\prime}(w)$ is strictly increasing in x.*

Proof.

$$f_z^{(2)\prime}(w) = \frac{d^2 \exp(w) \exp(z - d \log(1 + \exp(w)))}{(1 + \exp(w))(1 + \exp(z - d \log(1 + \exp(w))))} \quad .$$

■

Lemma 38. *$f_{\widehat{x}_c}^{(2)\prime}(w) \leq 1$, for all $w \in \mathbb{R}$.*

Proof. We start by establishing that for every $x \in \mathbb{R}$, $f_x(w)$ has a unique fixed point $\alpha := \alpha(x)$ and that this fixed point α satisfies the equation $x = \alpha + d\log(1 + \exp(\alpha))$. To see this we first observe that since f_x is strictly decreasing, it has at most one fixed point. At the same time, $w - f_x(w)$ is continuous and changes sign in $(-\infty, +\infty)$, implying that f_x has at least one fixed point. Hence, f_x has a unique fixed point, call it α. Requiring $\alpha = f_x(\alpha) = x - d\log(1 + \exp(\alpha))$ is equivalent to $x = \alpha + d\log(1 + \exp(\alpha))$.

Next we prove that $|f'_x(\alpha)| < 1$ iff $x < \widehat{x}_c$. To see this observe that $|f'_x(\alpha)| < 1$ is equivalent to $d\frac{e^\alpha}{1+e^\alpha} < 1$, which is equivalent to $e^\alpha < 1/(d-1)$. Thus, at fixed point,

$$x = \alpha + d\log(1+\exp(\alpha)) < \log\left(\frac{1}{d-1}\right) + d\log\left(1 + \frac{1}{d-1}\right) = \log\left(\frac{d^d}{(d-1)^{(d+1)}}\right) \ .$$

Finally, let $h(w) := h_x(w) = \exp(x - d\log(1 + e^w)) > 0$. The second derivative of $f_x^{(2)}$ is

$$\frac{d^2 e^w h(w)(1 + h(w) - de^w)}{(1 + e^w)^2(1 + h(w))^2} \ ,$$

so that its sign is the sign of $g(w) := g_x(w) = 1 + h(w) - de^w$. It is easy to check that g is strictly decreasing and that $g > 0$ for $w < \alpha$. Since $g(\alpha) = 0$ for $x = \widehat{x}_c$, we see that $f_{\widehat{x}_c}^{(2)\prime}$ attains its global maximum at $w = \alpha$. Since $f'_{\widehat{x}_c}(\alpha) = f'_{\widehat{x}_c}(\log(1/(d-1))) = -1$, we see that

$$f_{\widehat{x}_c}^{(2)\prime}(\alpha) = f'_{\widehat{x}_c}(f_{\widehat{x}_c}(\alpha))f'_{\widehat{x}_c}(\alpha) = f'_{\widehat{x}_c}(\alpha)f'_{\widehat{x}_c}(\alpha) = (-1)^2 = 1 \ .$$

■

Remark 39. *We will actually prove that the condition in Lemma 36 is also necessary for $F_x^{(2)}$ to be a contraction. This is makes the origin of the approximability barrier transparent and may lead to a simpler proof of the hardness result by Sly and Sun [3]. We leave this as further work.*

1.8 Complex Activities of Bounded Modulus

For $\Delta \geq 3$, let

$$\widehat{x}^*(\Delta) = \log(x^*(\Delta)) = \log\left(\frac{(\Delta-1)^{\Delta-1}}{\Delta^\Delta}\right) \quad .$$

Patel and Regts [15] and Harvey, Srivastava and Vondrák [26] proved that the approximability of the independent set polynomial can be extended to all complex numbers of modulus at most x^*, which in the logarithmic domain corresponds to the complex half-plane comprising numbers whose real part is less than $\widehat{x}^*$. In the following, $\Re(w)$ denotes the real part of $w \in \mathbb{C}$.

Theorem 40 ([15], [26])**.** *For every $\varepsilon > 0$, there exists an efficient algorithm which given a graph of maximum degree Δ and activities $\boldsymbol{x} \in \{x \in \mathbb{C} : \Re(x) < \widehat{x}^*(\Delta)\}^n$, returns $Y \in [\log Z - \varepsilon, \log Z + \varepsilon]$.*

Let $d = \Delta - 1$. Recall the definition (1.44) of function F_x and for arbitrary $\theta > 0$ define the set of gates

$$\mathcal{H}_\theta = \{F_x(\boldsymbol{w}) : \Re(x) \leq \widehat{x}^*(\Delta) - \theta\} \quad .$$

We prove a significantly more general result than Theorem 40.

Theorem 41. *The output of* **any** *circuit on $\mathcal{H}_\theta$ can be additively ε-approximated by truncating the circuit at depth proportional to $\log_{1/\delta} 1/\varepsilon$, where $\delta = \delta(\theta) < 1$.*

Thus, Theorem 40 is the special case of Theorem 41 when the circuit on $\mathcal{H}_\theta$ is the one computing the independent set polynomial. We leave the exploration other models for which Theorem 41 can be applied as future work. To prove Theorem 41 we let $A = \{w \in \mathbb{C} : \Re(w) < -\log(d+1))\}$, and establish that (a) F_x is a δ-contraction on A^d for some $\delta < 1$, and that (b) $F_x(A^d) \subseteq A$ for all $\{x \in \mathbb{C} : \Re(x) < \widehat{x}^*(\Delta)\}$. We now prove Theorem 41.

Proof. We study the functions (gates) in $\mathcal{H}_\theta$. Specifically, observe that

$$|f'(w)| = d\left|\frac{\exp(w)}{1+\exp(w)}\right| \leq d\frac{|\exp(w)|}{1-|\exp(w)|} = d\frac{\exp(\Re(w))}{1-\exp(\Re(w))} \quad ,$$

so that $|f'(w)| < 1$ whenever $\Re(w) < -\log(d+1)$. Letting $A = \{w \in F : \Re(w) < -\log(d+1))\}$, Lemma 33 implies $|\nabla F_x(\boldsymbol{w})| < 1$ for all $x \in F$ and $\boldsymbol{w} \in A^d$ and, therefore, by Lemma 32, for all such $x, \boldsymbol{w}$ there exists $\delta = \delta(x, \boldsymbol{w}) < 1$ such that F_x is a δ-contraction on A^d. We are thus left to identify a set (of activities) $G = \{x \in F : F_x(A^d) \subseteq A\}$ so that we can apply Theorem 31. To determine such a set G we observe that

$$F_x(A^d) \subseteq A \iff \sup_{\boldsymbol{w} \in A^d} \Re\left(F_x(\boldsymbol{w})\right) < -\log(d+1) \tag{1.45}$$

and that

$$\begin{aligned}
\sup_{\boldsymbol{w} \in A^d} \Re\left(F_x(\boldsymbol{w})\right) &= \Re\left(x\right) - \sum_{i=1}^{d} \inf_{w_i \in A} \Re\left(\log\left(1 + \exp(w_i)\right)\right) \\
&= \Re\left(x\right) - \sum_{i=1}^{d} \inf_{w_i \in A} \log\left(\left|1 + \exp(w_i)\right|\right) \\
&= \Re\left(x\right) - d\log\left(1 - \inf_{w \in A} |\exp(w)|\right) \\
&= \Re\left(x\right) - d\log\left(1 - \inf_{w \in A} \exp(\Re(w))\right) \\
&= \Re\left(x\right) - d\log\left(1 - \frac{1}{d+1}\right) \quad .
\end{aligned}$$

Thus, we can take G to comprise the disk $\{x \in F : \Re\left(x\right) < \widehat{x}^*(\Delta)\}$. ■

1.9 Future Work

- As we will see later, our local lemma corresponding to non-1-backtracking walks, is intimately related to the Bethe approximation (which we will also see later). More generally, the hierarchy of local lemmata mentioned above, corresponds to a hierarchy of Kikuchi approximations. Making this connection explicit will (i) connect the LLL with variational inference, and (ii) extend notions such as the Bethe/Kikuchi approximations to graphical models that take *negative* values.

- Passing to negative values suggests a vast generalization of the Probabilistic Method. Specifically, observe that to prove that objects with certain properties exist, the Probabilistic Method assigns to each candidate object a positive number (probability) and the goal is to prove that the sum of the numbers assigned to the desired objects is positive. But another tack could have been to assign each candidate an *arbitrary* number, i.e., positive or negative, and then prove that the sum of the numbers assigned to the desired objects is *non-zero*.

- We believe that our approach to the independent set polynomial through walk-trees establishes a new line of attack for *exactly solvable* models in statistical mechanics. Our first target is to recover the exact results for the hard-hexagon on the triangular lattice (which, while being a triangulation of the place, is *not* a chordal graph).

- Recall that chordal graphs have no induced cycles of length greater than 3. Call a graph q-cyclic if it has no induced cycles of length greater than q. The methodology we have developed for chordal graphs, causing the number of distinct subproblems in the fundamental recurrence to collapse, appears to be extendable to q-cyclic graphs, where the size of the collapsed set is exponential in q.

- Our “circuit” proof of Weitz’s theorem, is using only few properties of hard-core model. We believe it can, perhaps be adapted to other models.

- As mentioned, we hope to exploit the near-convexity of the egg-shaped (in the logarithmic domain) region in Figure 1.3 to establish the approximability of the independent set polynomial within it.

Chapter 2

A Hierarchy of Local Lemmata

In this chapter we provide a novel hierarchy of sufficient conditions for membership in $\mathbb{S}(G)$ that interpolates smoothly between the LLL and a new exact criterion, expressed in terms of walks on the graph. The walks viewpoint unifies all known local lemmas and yields new, stronger ones with minimal effort. For example, as soon as we take into account the subgraph induced by each inclusive neighborhood $\Gamma_i \cup \{i\}$, we get a local lemma that dominates every known *local* condition, e.g., the asymmetric LLL [29], the cluster-expansion LLL [30], and the non-backtracking LLL [31]. More generally, our hierarchy makes it possible to account for arbitrary short cycles, improving the pessimistic tree-like bound of the LLL.

To demonstrate the power of our hierarchy, we improve the lower bound for the negative fugacity singularity of the hard-core model on several lattices, a central problem of statistical physics [32–34]. Our bound of $\lambda_c \geq 0.1191$ for $\mathbb{Z}^2$ improves upon the previous best rigorous lower bound of 0.113 by Kolipaka, Szegedy, and Xu [25] and matches the *conjectured* value 0.11933888188..., to three decimal digits. We achieve a similar level of accuracy relative to the conjectured values for other lattices. As we will see, each bound is derived by selecting an integer q and enumerating all walks starting at the origin of the lattice that have length at most q and satisfy some additional conditions. The results presented here come from relatively small values of q, corresponding to computational effort in the order of minutes on a laptop computer.

Regarding algorithmic implications, we note that Kolipaka and Szegedy [35] already proved that if $\boldsymbol{p} \in \mathbb{S}(G)$, then the resampling algorithm of Moser and Tardos [36] will terminate in polynomial time. As a result, any algorithmic improvements must lie outside the dependency graph setting. That said, by providing new *tractable* sufficient conditions for membership in $\mathbb{S}(G)$ our work extends the range of problems that can be *provably* made constructive.

2.1 Formal Statement of Results

Throughout, we will be referring to a graph G with vertex set $[n]$ and edge set E. Write Γ_i for the neighborhood of i in G. A *walk* on G is a sequence of vertices where successive elements are adjacent in G. We denote the terminal (last) vertex of a walk w by $\mathrm{term}(w)$. For a set of walks $\mathcal{W}$ and a walk $w \in \mathcal{W}$, we define the continuations of w as $\mathrm{Cont}(w) = \{y : wy \in \mathcal{W}\}$.

Self-Bounding Walks. *A walk is* self-bounding *if in each step: (i) it proceeds from the current vertex i to a non-forbidden vertex $j \in \Gamma_i$ (initially no vertex is forbidden), and (ii) declares i and all neighbors of i greater than j forbidden. The set of self-bounding walks,* $\mathrm{SB}(G)$, *is, clearly, a subset of self-avoiding walks.*

Walk Equivalence. *Given a set of walks $\mathcal{W}$, we declare $w, w' \in \mathcal{W}$* equivalent *(in the context of $\mathcal{W}$) if* $\mathrm{term}(w) = \mathrm{term}(w')$ *and* $\mathrm{Cont}(w) = \mathrm{Cont}(w')$. *We denote the set of equivalence classes of a set of walks $\mathcal{W}$ by $\mathcal{C}(\mathcal{W})$. Since the function* term *is constant within each equivalence class, we write* $\mathrm{term}(c)$, *for $c \in \mathcal{C}(\mathcal{W})$.*

Causality. *For classes $c_1, c_2 \in \mathcal{C}(\mathcal{W})$, we write $c_1 \to c_2$ if there exist $w \in c_1$ and $z \in [n]$ such that $wz \in c_2$.*

For example, if $\mathcal{W}$ is the set of all walks on G, then two walks are equivalent iff they have the same last vertex. Thus, the equivalence classes are indexed by $[n]$ and

$c_i \to c_j$ iff $j \in \Gamma_i$. As another example, if $\mathcal{W}$ is the set of *non-backtracking* walks, i.e., walks without contiguous (i, j, i) triplets, then two walks are equivalent iff they either amount to the same single vertex, or they have the same last two vertices. Thus, there is an equivalent class per vertex and edge orientation, while $c_{(i,j)} \to c_j$ and $c_{(i,j)} \to c_{(j,k)}$, where $k \neq i$.

Our fundamental technical result is the following alternative characterization of Shearer's region.

Theorem 42 (Alternative Criterion for Membership in $\mathbb{S}(G)$)**.** $\boldsymbol{p} \in \mathbb{S}(G)$ *if and only if for each* $c_i \in \mathcal{C}(\mathrm{SB}(G))$ *there exists* $0 \le r_i < 1$*, such that* $p_{\mathrm{term}(c_i)} = r_i \prod_{c_j \leftarrow c_i}(1 - r_j)$*, for all* $c_i \in \mathcal{C}(\mathrm{SB}(G))$*.*

Since the number of self-bounding walks, $|\mathrm{SB}(G)|$, can be exponential, the criterion in Theorem 42 is, in general, intractable, like Shearer's. (Chordal graphs being a notable exception, as we discuss in Section 2.4). Unlike Shearer's criterion, though, our criterion in terms of self-bounding walks is extraordinarily amenable to relaxation: just turn the equality to an *in*equality and take *any* superset of $\mathrm{SB}(G)$. We emphasize that Theorem 43, below, is also a necessary and sufficient condition for membership in Shearer's region.

Theorem 43 (Main result)**.** $\boldsymbol{p} \in \mathbb{S}(G)$ *if and only if there exists a set of walks* $\mathcal{W} \supseteq \mathrm{SB}(G)$*, and a real number* $0 \le r_i < 1$ *for each class* $c_i \in \mathcal{C}(\mathcal{W})$*, such that for every class* $c_i \in \mathcal{C}(\mathcal{W})$*,*

$$p_{\mathrm{term}(c_i)} \le r_i \prod_{c_j \leftarrow c_i} (1 - r_j) \ . \tag{2.1}$$

The power of Theorem 43 is that it allows *arbitrary* sets of walks $\mathcal{W} \supset \mathrm{SB}(G)$ to certify membership in $\mathbb{S}(G)$, thus defining an infinite hierarchy of sufficient conditions, with the case $\mathcal{W} = \mathrm{SB}(G)$ at the top of the hierarchy. At the bottom of the hierarchy, naturally, $\mathcal{W}$ is simply the set of *all* walks on G (an infinite set). Rather delightfully, this bottom case amounts precisely to the asymmetric LLL.

Corollary 44 ([29]). *$\boldsymbol{p} \in \mathbb{S}(G)$ if there exist $\{r_i\}_{i\in[n]}$ such that $p_i \leq r_i \prod_{j\in\Gamma_i}(1-r_j)$ for all $i \in [n]$.*

Proof. If $\mathcal{W}$ is the set of all walks on G, then the equivalence classes are indexed by $[n]$ and $c_i \to c_j$ iff $j \in \Gamma_i$. Thus, (2.1) reads $p_i \leq r_i \prod_{j\in\Gamma_i}(1-r_j)$. ∎

Between the top and the bottom of the hierarchy, we are completely free. For example, if we take $\mathcal{W}$ to be the set of non-backtracking walks, we readily recover the local lemma of [31], a strict improvement of the asymmetric LLL. And by pursuing the same spirit, we can take $\mathcal{W}$ to be the set of 2-non-backtracking walks, i.e., walks that avoid the immediately previous vertex and the one before it, 3-non-backtracking walks, etc., and get a hierarchy of increasingly sharper sufficient conditions. As we discuss next, though, there is a much better alternative, yielding both far greater flexibility and much sharper bounds.

2.2 A Focal Lemma and Infinite Graphs

The idea is to chose certain subsets of vertices $S_1, \ldots, S_q$ and focus on them in the following sense: each set S_i "accepts" a walk w iff every time w enters S_i, the subwalk it performs within S_i until it re-exits belongs in $\mathrm{SB}(G[S_i])$. We take $\mathcal{W} = \mathcal{W}(S_1, \ldots, S_q)$ to be the set of walks accepted by *all* filters. More formally:

Definition 45. *Given a set of vertices S and a walk w, the S-*restriction *of w is the multiset of maximal contiguous subsequences of w whose elements are all in S. For example, if $w = (2, 1, 2, 5, 3, 5, 2)$, then its $\{2, 5\}$–restriction is $\{(2), (2, 5), (5, 2)\}$. Given a family $\mathbf{S} = \{S_1, \ldots, S_q\}$ of subsets of $[n]$, a walk is* $\mathbf{S}$-self-bounding *if for every $S_i \in \mathbf{S}$, every element of the S_i-restriction of w belongs in $\mathrm{SB}(G[S_i])$. For a family of vertex subsets $\mathbf{S}$, we denote the set of* $\mathbf{S}$-self-bounding *walks by $\mathcal{W}(\mathbf{S})$.*

Below are some examples of the flexibility and power of this viewpoint. We will see more later.

(i) If $\mathbf{S} = \{\{1\}, \{2\}, \ldots, \{n\}\}$, then $\mathcal{W}(\mathbf{S})$ is the set of all walks, yielding the asymmetric LLL.

(ii) If $\mathbf{S} = \{\{i, j\} : \{i, j\} \in E\}$, then $\mathcal{W}(\mathbf{S})$ is the set of non-backtracking walks, yielding the LLL of [31].

(iii) If $\mathbf{S} = \{[n]\}$, then $\mathcal{W}(\mathbf{S}) = \mathrm{SB}(G)$ and we recover the exact criterion of Theorem 42.

Theorem 46 addresses the two immediate concerns one may have for the set $\mathcal{W}(\mathbf{S})$ of $\mathbf{S}$-self-bounding walks. Namely, whether it always contains $\mathrm{SB}(G)$ and whether it has a finite number of equivalence classes.

2.2.1 Infinite Graphs

When all variables of the independent set polynomial, i.e., all vertices (sites) of the graph (lattice), have the same value (fugacity), λ, the polynomial becomes univariate.

For a graph G on $[n]$, we are interested in

$$\lambda_c(G) = \sup\{\lambda : Z_G(\boldsymbol{x}) > 0 \text{ for all } \boldsymbol{x} \in \mathbb{C}^n \text{ with } |x_i| \leq \lambda\} \ .$$

For countably infinite G, we take $\lambda_c(G)$ as the infimum of $\lambda_c(H)$ over finite induced subgraphs H of G. As discussed in Section 2.3, this number, λ_c, known as "negative-fugacity singularity of the hard-core lattice gas," is a central object in statistical physics.

Naturally, to compute λ_c for infinite graphs we need to impose some conditions, as otherwise even the problem description is infinite. To start, we restrict to countably infinite, locally-finite graphs. Next, to avoid enumerating their vertices, we require that they are *arc-labeled*, i.e., that each edge carries a label ℓ for each direction, from some alphabet Σ, so that the arcs leaving a vertex carry distinct labels. This way, any walk on G can be specified by an initial vertex and a sequence in Σ^*. Finally, to be able to define self-bounding walks, we require Σ to be totally ordered. Thus, a walk is *self-bounding* if initially no vertices are forbidden and every time we move from a vertex i to a non-forbidden vertex $j \in \Gamma_i$, we add i to the set of forbidden vertices along with all $k \in \Gamma_i$ with $\ell(i,k) > \ell(i,j)$. (We could have defined self-bounding walks in this manner also for finite graphs, but we did not want to introduce this level of complexity unneccessarily.)

Our main result (Theorem 43) applies to infinite arc-labelled graphs, yielding that $\lambda \leq \lambda_c(G)$ if and only if there exists a set of walks $\mathcal{W} \supseteq \mathrm{SB}(G)$, and a real number $0 \leq r_i < 1$ for each class $c_i \in \mathcal{C}(\mathcal{W})$, such that $r_i \prod_{c_j \leftarrow c_i}(1 - r_j) \geq \lambda$, for every class $c_i \in \mathcal{C}(\mathcal{W})$. But, of course, this condition is impossible to verify, since, without further assumptions, $\mathcal{C}(\mathcal{W})$ is infinite and unstructured.

To generate tractable sets of walks we require symmetry from $G(V,E)$. Recall that an automorphism is a bijection $\phi : V \mapsto V$ such that $\{u,v\} \in E$ iff $\{\phi(u), \phi(v)\} \in E$. An automorphism ϕ of an arc-labeled graph is *label-preserving* iff $\ell(x,y) = \ell(\phi(x), \phi(y))$, for every arc (x,y). An infinite graph is *quasi-transitive* if its vertices can be partitioned into a *finite* number of equivalence classes, called *orbits*, such that if u, v are in the same orbit, there is an automorphism mapping u to v. A quasi-transitive graph is *arc-label-quasi-transitive*, if for all u, v in the same orbit there is a label-preserving automorphism mapping u to v.

To get a sense for arc-label-quasi-transitivity, observe that as soon as $\phi(u) = v$ has been specified, aligning the labels of the arcs leaving u with the labels of the arcs leaving v, forces the image under ϕ of every neighbor of u. And as soon as this image has been specified, by the same logic, the image of the second neighbors of u is forced, etc. As a result, we see that in a quasi-transitive, arc-labeled graph, for any two vertices u, v in the same orbit there can be *at most* one label-preserving automorphism mapping u to v; yet, we require such an automorphism to exist for every pair of vertices in the same orbit. This makes arc-label transitivity a very strict requirement. Nevertheless, it is a requirement satisfied by many infinite graphs of interest, including lattices. For example, it is easy to see that it is satisfied by $\mathbb{Z}^d$ if all parallel arcs with the same direction receive the same label.

Theorem 46. *Let $G(V, E)$ be any arc-label-quasi-transitive graph with orbits $\mathcal{O}_1, \ldots, \mathcal{O}_t$. For each $i \in [t]$, select $x_i \in \mathcal{O}_i$ and $\mathbf{S}_{x_i} = \{S_1, \ldots, S_{q_i}\}$ arbitrarily, subject only to $x_i \in S_j$ and $|S_j| < \infty$, for all $j \in [q_i]$. For each $v \in \mathcal{O}_i$, let $\mathbf{S}_v$ be the image of $\mathbf{S}_{x_i}$ under the unique automorphism mapping x_i to v. Let $\mathbf{S} = \cup_{v \in V} \mathbf{S}_v$.*

- *$\mathcal{W}(\mathbf{S}) \supseteq \mathrm{SB}(G)$ and $\mathcal{C}(\mathcal{W}(\mathbf{S}))$ is finite.*
- *If for each $c_i \in \mathcal{C}(\mathcal{W}(\mathbf{S}))$ there exists $0 \leq r_i < 1$, such that for every $c_i \in \mathcal{C}(\mathcal{W}(\mathbf{S}))$,*

$$\lambda \leq r_i \prod_{c_j \leftarrow c_i} (1 - r_j) \ , \tag{2.2}$$

  then $\lambda_c(G) \geq \lambda$.

In our application of Theorem 46 to $\mathbb{Z}_2$, $\mathbb{Z}_3$, the triangular, and the hexagonal lattice, we label the arcs so that parallel arcs with the same direction receive the same label. For $\mathbb{Z}_2$, $\mathbb{Z}_3$, and the triangular lattice, which only have one orbit, we take x to be the lattice origin. For the hexagonal lattice, which has two orbits, we take x_1 to be the origin of the lattice and x_2 to be one of the vertices adjacent to the origin. For $\mathbb{Z}_2$, $\mathbb{Z}_3$, and the hexagonal lattice, we take each $\mathbf{S}_x = \mathbf{S}_x(q)$ to be such that a walk starting at x is accepted iff its q-prefix is self-bounding, for some integer $q \in [13, 15]$, depending on the lattice. For the triangular lattice, we take $\mathbf{S}_x$ so that only 2-non-backtracking walks are accepted.

2.3 Lower Bounds for Negative-Fugacity Singularity of Hard-Core Lattice Gas

In statistical physics, λ_c is known as the "negative-fugacity singularity of the hard-core lattice gas," and has been extensively studied for many lattices. Its importance is primarily motivated by the Lee-Yang [4] theory of phase transitions which implies that for $\lambda \in [0, \lambda_c)$ the hard-core model, i.e., the probability distribution on independent sets where each independent set I has probability proportional to $\lambda^{|I|}$, does not exhibit any phase transition. Very recently, Regts [37] strengthened the Lee-Yang conclusion, by showing that for $\lambda \in [0, \lambda_c)$, the hard-core model exhibits decay of correlations in the form of *strong spatial mixing*, a very helpful property for designing efficient algorithms for approximating the independent set polynomial.

We apply our Theorem 46 to four important lattices: the square lattice ($\mathbb{Z}^2$), the cubic lattice ($\mathbb{Z}^3$), the triangular lattice, and the hexagonal planar lattice. In each case, we label the arcs so that parallel arcs with the same direction receive the same label. For $\mathbb{Z}_2$, $\mathbb{Z}_3$, and the triangular lattice, which only have one orbit, we take x to be the lattice origin. For the hexagonal lattice, which has two orbits, we take x_1 to be the origin of the lattice and x_2 to be one of the vertices adjacent to the origin. For $\mathbb{Z}_2$, $\mathbb{Z}_3$, and the hexagonal lattice, we take each $\mathbf{S}_x = \mathbf{S}_x(q)$ to be such that a walk starting at x is accepted iff all of its subwalks of length q are self-bounding, for some integer $q \in \{13, 14, 15\}$, the choice depending on the lattice. In each case, the corresponding set of walks has approximately a million elements.

As can be seen from Table 2.1, our method dominates all other known methods, matching the numerical estimate derived using non-rigorous methods from statistical physics to the third decimal digit. (Of the four lattices considered, the exact value of λ_c is actually known for the triangular lattice, per the celebrated result of Baxter [21] and equals $\frac{5\sqrt{5}-11}{2} = 0.09016...$) There are two further observations worth making.

The first is that for the triangular lattice, the bound delivered even by just the non-backtracking lemma beats the bound offered by the cluster expansion LLL, even though the triangular lattice is ideal for that LLL, as a significant number of triangles occupies the neighborhood of every vertex. The second is that the bound given by the non-backtracking LLL for the square lattice is equal to the bound given by the asymmetric LLL for the hexagonal lattice, as the astute reader may have noticed. This is not a coincidence. In the eyes of the non-backtracking LLL, each vertex of the square lattice (other than the origin) has effective degree three, since no backtracking is allowed. In the eyes of the asymmetric LLL, each vertex of the hexagonal lattice has effective degree equal to its actual degree, i.e., three, as the corresponding set of walks is completely unrestricted. Thus, in the eyes of the two methods, the two lattices are indistinguishable.

	Asymmetric [29]	Cluster [30]	Non-back [31]	Decomposition [25]	**Our Work**	Numerical [33]
Square ($\mathbb{Z}^2$)	0.0819	0.0896	0.1054	0.1130	0.1191	0.1193
Cubic ($\mathbb{Z}^3$)	0.0566	0.0601	0.0669	0.0702	0.0741	0.0744
Triangular	0.0566	0.0776	0.0811	0.0811	0.0899	0.0902
Hexagonal	0.1054	0.1190	0.1481	0.1481	0.1542	0.1547

Table 2.1: Lower bounds for λ_c for different lattices and methods.

2.4 Related Work

2.4.1 Zero-Free Regions and Evaluation of the Independent Set Polynomial

Sinclair, Srivastava, Štefankovič, and Ying [38] consider the hard core model on graphs with a bounded connective constant, such as the lattices we consider here. Harvey, Srivastava, and Vondrák [26] consider the problem of evaluating the independent set polynomial inside $\mathbb{S}(G)$. Bezáková, Galanis, Goldberg, and Štefankovič [39] consider the inapproximability of the independent set polynomial outside $\mathbb{S}(G)$. Finally, Vera, Vigoda, and Yang [40] consider the independent set polynomial for $\mathbb{Z}^2$.

self-bounding walks, the technical foundation for our work, have been related before to the independent set polynomial in the seminal work of Scott and Sokal [16] relating the polynomial to the LLL. They also appear, in a somewhat disguised form, in the seminal work of Weitz [2], on evaluating the hard-core model partition function. The disguise amounts to the fact that in [2] the tree of self-bounding walks is presented as a truncation of the tree of self-avoiding walks (SAW), enforced by imposing certain boundary conditions on the vertices. We believe that working instead with our explicit combinatorial representation, brings out the crucial monotonicity property of the tree for negative real arguments.

The first ever strict improvement to the asymmetric LLL was given in [31], showing that (under the exact same hypothesis) one can replace $\prod_{j\in\Gamma_i}(1-r_j)$ with $\prod_{j\in\Gamma_i}(1-r_j)/(1-r_ir_j)$. That was a corollary to the main result of [31], a local lemma requiring a condition on both the vertices and the edges of G. While that local lemma was proven by an ad hoc inductive argument, we show that it follows readily from Theorem 43 by taking $\mathcal{W}$ to be the set of non-backtracking walks.

The cluster expansion LLL of Bissacot, Fernández, Procacci, and Scoppola [20], improves upon the asymmetric LLL, when the neighborhood of each vertex has one or more triangles (for triangle-free graphs it is equivalent to the asymmetric LLL).

Even for graphs such as the triangular lattice, though, taking as a filter $S_i = \{i\} \cup \Gamma_i$ for each vertex $i \in [n]$, already gives better results.

2.4.2 The Decomposition Local Lemma of Kolipaka, Szegedy, and Xu

To the best of our knowledge, the only work prior to ours that exploits features of the dependency graph beyond vertex neighborhoods is the *Decomposition Local Lemma* of Kolipaka, Szegedy, and Xu [25]. While [25] did not draw any connection to walks, we can establish such a connection and show that for any given family of sets $\mathbf{S}$ (and thus two conditions of similar complexity), our method always dominates the Decomposition LLL of [25]. Specifically, for a given collection of sets $S_1, \ldots, S_q$, a walk is "accepted" by [25] if it can be partitioned into subwalks each of which lies entirely and is self-bounding within some set S_i. While the two approaches are *identical* when each edge of G belongs in exactly one filter, as soon as any edge is included in multiple filters, our method acquires an advantage, as it is more selective in the walks it accepts. Since in typical applications the filters correspond, roughly, to depth-d vertex neighborhoods, the amount of filter overlap (and the corresponding gap in performance) grows rapidly in d.

The bigger difference between the two approaches, though, is that while our method enjoys monotonicity, i.e., enlarging filters and/or adding new ones, can only shrink the set of accepted walks and, thus, improve the resulting lower bound, that is *not* the case for the decomposition LLL of [25]. As a result, besides performing worse for any given set of filters, designing a good decomposition (collection of sets) for the decomposition LLL is far from obvious.

2.4.3 Benchmarking on Lattices

While for the asymmetric LLL, the cluster expansion LLL, and the non-backtracking LLL, deriving the lower bound implied for λ_c on each lattice is straightforward, for the decomposition LLL of [25], more work is needed, as the bound depends on the decomposition. Specifically, since [25] reports a lower bound only for the square lattice, in order to make the comparison more comprehensive, we derived suitable decompositions ourselves, as follows. For the cubic lattice imagine the unit cubes of $\mathbb{Z}^3$ colored black and white in a 3D-chessboard pattern. Taking each set (filter) in the decomposition to be the set of vertices of a white unit cube, so that every edge is included in exactly one filter, gave 0.0695. A more complex decomposition, yielded 0.0702. For the hexagonal lattice, the number we report corresponds to the decomposition where each filter amounts to a single edge of the lattice, making the method equivalent to the non-backtracking LLL. We did this because several natural choices of larger filters only made things worse, manifesting the non-monotonicity of the decomposition LLL. For example, taking the filters to be the hexagonal faces gives only 0.067, much less than taking the filters to be individual edges which gives 0.1481.

2.5 Proof of Theorems 42 and 43

2.5.1 Proof of Basic Tools and Theorem 42

Let $\mathcal{W}$ be any *finite* set of walks on G, and let $\boldsymbol{s} \in \mathbb{R}^{\mathcal{W}}$. For two walks $w, z \in \mathcal{W}$, we write $\{z\} \leftarrow \{w\}$ to denote that z is a one-step continuation of w, i.e., that $z = wv$ for some $v \in [n]$.

Given $\boldsymbol{s}, \mathcal{W}$, we define the *message vector* $\mathbf{m}(\boldsymbol{s}, \mathcal{W}) = \mathbf{m} \in \mathbb{R}^{\mathcal{W}}$ to be the unique vector satisfying

$$\mathrm{m}_w = s_w \prod_{\{z\} \leftarrow \{w\}} \frac{1}{1 - \mathrm{m}_z} \ . \tag{2.3}$$

To see that $\mathbf{m}$ is unique, observe that the graph having one vertex per walk in $\mathcal{W}$ and an edge between every $w, z \in \mathcal{W}$ such that $\{z\} \leftarrow \{w\}$, is a forest. Rooting each tree of this forest by the vertex corresponding to the tree's shortest walk, we can compute the coefficients of $\mathbf{m}$ in a bottom-up fashion in each tree.

Definition 47. *We say that* $\boldsymbol{s} \in \mathbb{R}^{\mathcal{W}}$ *is* valid *for* $\mathcal{W}$ *if* $\mathbf{m}(\boldsymbol{s}, \mathcal{W}) \in [0,1)^{\mathcal{W}}$.

We next show that the set of valid messages is downward-closed with respect to $\boldsymbol{s}$

Lemma 48. *If* $\boldsymbol{s}$ *is valid for* $\mathcal{W}$ *and* $\boldsymbol{q} \le \boldsymbol{s}$, *then* $\mathbf{m}(\boldsymbol{q}, \mathcal{W}) \le \mathbf{m}(\boldsymbol{s}, \mathcal{W})$.

Proof. Recall that $\mathcal{W}$ is finite and let $\ell(w) = |\mathrm{Cont}(w)| < \infty$. We will prove that $\mathrm{m}_w(\boldsymbol{q}, \mathcal{W}) \le \mathrm{m}_w(\boldsymbol{s}, \mathcal{W})$ by induction on $\ell(w)$. Trivially, if $\ell(w) = 0$, then per (2.3), $\mathrm{m}_w(\boldsymbol{q}, \mathcal{W}) = q_w \le s_w = \mathrm{m}_w(\boldsymbol{s}, \mathcal{W})$.

Assume now that $\mathrm{m}_z(\boldsymbol{q}, \mathcal{W}) \le \mathrm{m}_z(\boldsymbol{s}, \mathcal{W})$ for all $z \in \mathcal{W}$ with $\ell(z) < k$, and let $w \in \mathcal{W}$ with $\ell(w) = k$. The two equalities in (2.4) follow from (2.3). For the inequality, $\boldsymbol{q} \le \boldsymbol{s}$ yields $q_w \le s_w$, while the fact that $\boldsymbol{s}$ is valid for $\mathcal{W}$ implies $1 - \mathrm{m}_z(\boldsymbol{q}, \mathcal{W}) > 0$. Combining the latter fact with the inductive hypothesis that $\mathrm{m}_z(\boldsymbol{q}, \mathcal{W}) \le \mathrm{m}_z(\boldsymbol{s}, \mathcal{W})$ the inequality follows, yielding

$$\mathrm{m}_w(\boldsymbol{q}, \mathcal{W}) = q_w \prod_{\{z\} \leftarrow \{w\}} \frac{1}{1 - \mathrm{m}_z(\boldsymbol{q}, \mathcal{W})} \tag{2.4}$$

$$\le s_w \prod_{\{z\} \leftarrow \{w\}} \frac{1}{1 - \mathrm{m}_z(\boldsymbol{s}, \mathcal{W})} = \mathrm{m}_w(\boldsymbol{s}, \mathcal{W}) \ . \tag{2.5}$$

■

Recall that a walk is self-bounding if in each step: (i) it proceeds from the current vertex i to a non-forbidden vertex $j \in \Gamma_i$ (initially no vertex is forbidden), and (ii) declares i and all neighbors of i greater than j forbidden. Recall that we denote the set of self-bounding walks on a graph G by $\mathrm{SB}(G)$. For $w \in \mathrm{SB}(G)$, let $\mathrm{Forb}(w)$ denote the set of forbidden vertices of w when the walk reaches $\mathrm{term}(w)$.

Definition 49. *Given* $\boldsymbol{p} \in \mathbb{R}^n$ *and a set of walks* $\mathcal{W}$, *let* $\tilde{\boldsymbol{p}} \in \mathbb{R}^{\mathcal{W}}$, *where* $\tilde{p}_w := p_{\mathrm{term}(w)}$, *for each* $w \in \mathcal{W}$.

Lemma 50. *For every* $\boldsymbol{p} \in \mathbb{R}^n$ *and walk* $w \in \mathrm{SB}(G)$,

$$\mathrm{m}_w(\tilde{\boldsymbol{p}}, \mathrm{SB}(G)) = \mathrm{ratio}\left(\boldsymbol{p}; (\mathrm{term}(w), ([n] - \mathrm{term}(w)) \setminus \mathrm{Forb}(w))\right). \tag{2.6}$$

Proof. We will show that the rhs of (2.6) satisfies equation (2.3) for $\tilde{\boldsymbol{p}}, \mathcal{W}$. The uniqueness of $\mathbf{m}$, yields the result. Let $w \in \mathrm{SB}(G)$ be arbitrary, and let $j_1 > \cdots > j_q$ be a descending ordering of $\Gamma_{\mathrm{term}(w)} \setminus \mathrm{Forb}(w)$. Applying recurrence (1.15) on ratio $(\boldsymbol{p}; (\mathrm{term}(w), ([n] - \mathrm{term}(w)) \setminus \mathrm{Forb}(w)))$ gives the first equality below. The second, and key, equality follows from the definition of Forb.

To see it, recall that transitioning from $\mathrm{term}(w)$ to j_ℓ, we declare $\mathrm{term}(w)$ and all neighbors of $\mathrm{term}(w)$ greater than v_j, i.e., $\{j_1, \ldots, j_\ell\}$, forbidden, which explains why we can rewrite $\mathrm{Forb}(w) \cup \{\mathrm{term}(w), j_1, \ldots, j_\ell\}$ as $\mathrm{Forb}(wj_\ell) \cup \{j_\ell\}$. For the next equality, observe that every vertex in $\Gamma_{\mathrm{term}(w)} \setminus \mathrm{Forb}(w)$, is by the definition of self-bounding walks, a valid next step for w, which explains why the new indexing of the product and the use of v instead of j_ℓ. The last equality is simply a rewriting of 1-step continuations in terms of the "$\leftarrow$" notation, using $\mathrm{term}(z)$ instead of v.

$$
\begin{aligned}
&\mathrm{ratio}(\mathrm{term}(w), ([n] - \mathrm{term}(w)) \setminus \mathrm{Forb}(w)) \\
&\quad = p_{\mathrm{term}(w)} \prod_{\ell=1}^{d} \frac{1}{1 - \mathrm{ratio}\left(j_\ell, (([n] - \mathrm{term}(w)) \setminus \mathrm{Forb}(w)) - \{j_1, \ldots, j_\ell\}\right)} && (2.7) \\
&\quad = p_{\mathrm{term}(w)} \prod_{\ell=1}^{d} \frac{1}{1 - \mathrm{ratio}\left(j_\ell, [n] \setminus (\mathrm{Forb}(wj_\ell) \cup \{j_\ell\})\right)} && (2.8) \\
&\quad = p_{\mathrm{term}(w)} \prod_{v:\{wv\} \leftarrow \{w\}} \frac{1}{1 - \mathrm{ratio}\left(v, [n] \setminus (\mathrm{Forb}(wv) \cup \{v\})\right)} && (2.9) \\
&\quad = \tilde{p}_w \prod_{\{z\} \leftarrow \{w\}} \frac{1}{1 - \mathrm{ratio}\left(\mathrm{term}(z), ([n] - \mathrm{term}(z)) \setminus \mathrm{Forb}(z)\right)} \quad . && (2.10)
\end{aligned}
$$

■

Lemma 51. $\boldsymbol{p} \in [0,1)^n$ *is good for* G *if and only if* $\tilde{\boldsymbol{p}} \in \mathbb{R}^{\mathrm{SB}(G)}$ *is valid for* $\mathrm{SB}(G)$.

Proof.

($\Longrightarrow$) We claim that if $\boldsymbol{p}$ is good for G, then ratio $(\boldsymbol{p}; (i, S)) \in [0,1)$, for all $i \in [n]$ and $S \subseteq [n] \setminus \{i\}$. Per Lemma 50, this implies $\mathrm{m}\left(\tilde{\boldsymbol{p}}; \mathrm{SB}(G)\right) \in [0,1)^{\mathcal{W}}$, i.e., that $\tilde{\boldsymbol{p}}$ is valid for $\mathrm{SB}(G)$.

For the sake of contradiction, let $S = \{j_1, \ldots, j_q\}$ be a minimal subset of $[n]$, such that $\mathrm{ratio}(i, S) \notin [0, 1)$. Notice that it cannot be $S = \emptyset$, since $\mathrm{ratio}(i, \emptyset) = p_i \in [0, 1)$, and that the minimality of S and (1.15) imply that it must be $\mathrm{ratio}(i, S) \geq 1$. For $\ell \in [q]$, write $S_\ell = S - \{j_1, \ldots, j_\ell\}$, with $S_0 = S$, and note that

$$
\begin{aligned}
Z(S \cup \{i\}) &= \frac{Z(S \cup \{i\})}{Z(S)} \prod_{\ell \in [q]} \frac{Z(S_{\ell-1})}{Z(S_\ell)} \\
&= Z(i \mid S) \prod_{\ell \in [q]} Z(j_\ell \mid S_\ell) \\
&= (1 - \mathrm{ratio}(i, S)) \prod_{\ell \in [q]} (1 - \mathrm{ratio}(j_\ell, S_\ell)) \ .
\end{aligned}
$$

Theorem **??** implies $Z(S \cup \{i\}) > 0$ since $\boldsymbol{p}$ is good. The minimality of S implies $\prod_{\ell \in [q]}(1 - \mathrm{ratio}(j_\ell, S_\ell)) > 0$. Therefore, it must be $(1 - \mathrm{ratio}(i, S)) > 0$, a contradiction to our initial assumption.

($\Longleftarrow$) By Theorem **??** it suffices to prove that if $\tilde{\boldsymbol{p}}$ is valid for $\mathrm{SB}(G)$, then $Z(S) > 0$ for every $S \subseteq [n]$.

Let $S = \{j_1, \ldots, j_q\} \subseteq [n]$. As before, for $\ell \in [q]$, write $S_\ell = S - \{j_1, \ldots, j_\ell\}$, with $S_0 = S$. By telescoping,

$$
Z(S) = \prod_{\ell \in [q]} \frac{Z(S_{\ell-1})}{Z(S_\ell)} = \prod_{\ell \in [q]} Z(j_\ell \mid S_\ell) = \prod_{\ell \in [q]} (1 - \mathrm{ratio}(\boldsymbol{p}; j_\ell, S_\ell)) \ . \tag{2.11}
$$

Write H_ℓ for the subgraph of G induced by $\{j_1, \ldots, j_\ell\}$ and (i) for the walk consisting solely of vertex i. Lemma 50 implies $\mathrm{ratio}(\boldsymbol{p}; j_\ell, S_\ell)) = \mathrm{m}_{(j_\ell)}(\tilde{\boldsymbol{p}}, \mathrm{SB}(H_\ell))$. Moreover, $\mathrm{m}_{(j_\ell)}(\tilde{\boldsymbol{p}}, \mathrm{SB}(H_\ell)) = \mathrm{m}_{(j_\ell)}(\tilde{\boldsymbol{p}}_{\boldsymbol{\ell}}, \mathrm{SB}(G))$, where $\tilde{\boldsymbol{p}}_{\boldsymbol{\ell}}$ is the mutation of $\tilde{\boldsymbol{p}}$ that results by setting to 0 all coordinates corresponding to walks outside $\mathrm{SB}(H_\ell)$. Since $\tilde{\boldsymbol{p}}$ is valid for $\mathrm{SB}(G)$, we see that $1 > \mathrm{m}_{(j_\ell)}(\tilde{\boldsymbol{p}}, \mathrm{SB}(G)) \geq \mathrm{m}_{(j_\ell)}(\tilde{\boldsymbol{p}}_{\boldsymbol{\ell}}, \mathrm{SB}(G))$, the second inequality due to Lemma 48. Therefore, every factor in the rhs of (2.11) is positive, concluding the proof. ■

Lemma 52. *For every finite set of walks $\mathcal{W}$, and all $w, y \in \mathcal{W}$, with $w \sim y$, we have* $\mathrm{m}_w(\boldsymbol{q}, \mathcal{W}) = \mathrm{m}_y(\boldsymbol{q}, \mathcal{W})$.

Proof. The result follows from the observation that equation (2.3) depends only on term(w), and Cont(w). ■

Combining now Lemma 51 and Lemma 52 for $\mathcal{W} = \mathrm{SB}(G)$, gives Theorem 42.

2.5.2 Proof of Theorem 43

We first prove the following Theorem 43 for the special case where each inequality in (2.1) holds as equality.

Proof of Theorem 43 with equalities.
($\Longleftarrow$) We will construct $\boldsymbol{s} \geq \tilde{\boldsymbol{p}}$ valid for SB(G). Lemma 48 then implies that $\tilde{\boldsymbol{p}}$ is valid for SB(G) and Lemma 51 then implies that $\boldsymbol{p}$ is good for G, i.e., $\boldsymbol{p} \in \mathbb{S}(G)$.

We start by noting that for every class $c_i \in C(\mathcal{W})$, and every $w \in c_i$,

$$\{\mathcal{C}(z) : \{z\} \leftarrow \{w\}\} = \{c_j : c_j \leftarrow c_i\} \ .$$

This is because, by the definition of causality, $\{c_j : c_j \leftarrow c_i\} = \cup_{w \in c_i}\{\mathcal{C}(z) : \{z\} \leftarrow \{w\}\}$, while by the definition of walk equivalence, $\{\mathcal{C}(z) : \{z\} \leftarrow \{w\}\} = \{\mathcal{C}(z) : \{z\} \leftarrow \{y\}\}$, for any two equivalent $w, y \in \mathcal{W}$. Thus, for every $c_i \in \mathcal{C}(\mathcal{W})$, we can select $w \in c_i$ arbitrarily and rewrite (2.2) as

$$p_{\mathrm{term}(w)} = r_{\mathcal{C}(w)} \prod_{\{z\} \leftarrow \{w\}} (1 - r_{\mathcal{C}(z)}) \ . \tag{2.12}$$

For every $w \in \mathcal{W}$, by rearranging factors in (2.12) we derive (2.13), while we define s_w via (2.14),

$$\begin{aligned} r_{\mathcal{C}(w)} &= p_{\mathrm{term}(w)} \prod_{\substack{\{z\} \leftarrow \{w\} \\ z \notin \mathrm{SB}(G)}} \frac{1}{1 - r_{\mathcal{C}(z)}} \prod_{\substack{\{z\} \leftarrow \{w\} \\ z \in \mathrm{SB}(G)}} \frac{1}{1 - r_{\mathcal{C}(z)}} \qquad (2.13) \\ &:= s_w \prod_{\substack{\{z\} \leftarrow \{w\} \\ z \in \mathrm{SB}(G)}} \frac{1}{1 - r_{\mathcal{C}(z)}} \ . \qquad (2.14) \end{aligned}$$

Since $0 \leq r_i < 1$ for all $c_i \in C(\mathcal{W})$, the left product in (2.13) is at least 1, implying $\boldsymbol{s} \geq \tilde{\boldsymbol{p}}$, as desired.

If we now let $\boldsymbol{r} \in \mathbb{R}^{\mathrm{SB}(G)}$ where $r_w = r_{\mathcal{C}(w)}$, we see that $\boldsymbol{r}, \boldsymbol{s}$ satisfy equation (2.3) for SB(G). The uniqueness of the solution of (2.3), implies that $\boldsymbol{r} = \mathbf{m}(\boldsymbol{s}, \mathrm{SB}(G))$. Thus, the hypothesis that $0 \leq r_i < 1$, for all $c_i \in C(\mathcal{W})$, implies that $\mathbf{m}(\boldsymbol{s}, \mathrm{SB}(G)) \in [0,1)^{\mathrm{SB}(G)}$, i.e., that $\boldsymbol{s}$ is valid for SB(G).

($\Longrightarrow$) By Lemma 51, if $\boldsymbol{p}$ is good for G, then $\tilde{\boldsymbol{p}}$ is valid for SB(G). Note now that for any finite set of walks $\mathcal{W}$, if w, z are equivalent, then $\mathrm{m}_w = \mathrm{m}_z$, as equation (2.3) depends only on term(w), and Cont(w). Thus, the projection of $\mathbf{m}(\tilde{\boldsymbol{p}}, \mathrm{SB}(G))$ to $\mathcal{C}(\mathrm{SB}(G))$ belongs in $[0,1)^{\mathrm{SB}(G)}$ and satisfies (2.2). ■

For the general case, we prove the following Lemma, asserting that whenever (2.1) holds as inequality it also holds as equality. We use $\boldsymbol{v} \geq \boldsymbol{u}$ for two real vectors $\boldsymbol{v}$ and $\boldsymbol{u}$, to denote that $\boldsymbol{v}$ is coordinate-wise greater than $\boldsymbol{u}$.

Lemma 53. *Let $q \geq 1$, $\boldsymbol{p} \in [0,1)^q$, and let $S_1, \ldots, S_q \subseteq [q]$. Assume there exist a $\boldsymbol{y} \in [0,1)^q$ such that $y_i \prod_{j \in S_i}(1 - y_j) \geq p_i$, then, there exists a $\boldsymbol{y}^* \in [0,1)^q$ such that $y_i^* \prod_{j \in S_i}(1 - y_j^*) = p_i$.*

Proof. For $\boldsymbol{x} \in \mathbb{R}^q$, and $i \in [q]$, define $g_i(\boldsymbol{x}) = p_i / \prod_{j \in S_i}(1 - x_j)$, and $g : (\mathbb{R} - \{1\})^q \mapsto \mathbb{R}^q$, with $g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_q(\boldsymbol{x}))$. Let $\boldsymbol{x}^{(1)} = \mathbf{0}$, and $\boldsymbol{x}^{(k+1)} = g(\boldsymbol{x}^{(k)})$, for $k \geq 1$. We and we claim that for every k, $\boldsymbol{x}^{(k+1)} \geq \boldsymbol{x}^{(k)}$, and that $\boldsymbol{x}^{(k)} \leq \boldsymbol{y}$, i.e., each coordinate of $\boldsymbol{x}^{(k)}$ consist an increasing and bounded sequence in $[0,1)$. Hence, $\boldsymbol{x}^{(k)}$ must converge to a limit $\mathbf{0} \leq \boldsymbol{y}^* < \mathbf{1}$, which must be a fixed point of g, i.e., $g(\boldsymbol{y}^*) = \boldsymbol{y}^*$, yielding $y_i^* \prod_{j \in S_i}(1 - y_j^*) = p_i$, as desired.

To prove $\boldsymbol{x}^{(k)}$ is monotone and bounded we proceed by induction on k. Clearly, $\boldsymbol{x}^{(1)} = \mathbf{0} \leq \boldsymbol{p} = \boldsymbol{x}^{(2)}$, and $\boldsymbol{x}^{(2)} = \boldsymbol{p} \leq \boldsymbol{y}$. Now assume the result holds for all $2 \leq k' \leq k$, we will show it for $k+1$. Indeed, for any $i \in [q]$, we have

$$x_i^{(k+1)} = g_i\left(\boldsymbol{x}^{(k)}\right) = \frac{p_i}{\prod_{j \in S_i}\left(1 - x_j^{(k)}\right)} \leq \frac{p_i}{\prod_{j \in S_i}\left(1 - x_j^{(k-1)}\right)} = g_i\left(\boldsymbol{x}^{(k-1)}\right) = x_i^{(k)} \leq y_i \ ,$$

where the two inequalities follow from the inductive hypotheses that $\boldsymbol{x}^{(k)} \geq \boldsymbol{x}^{(k-1)}$, and that $\boldsymbol{x}^{(k-1)} \leq \boldsymbol{y} \leq \mathbf{1}$.

■

2.5.3 Proof of Theorem 46

Proof of Theorem 46.

- The fact that $\mathcal{W}(\mathbf{S}) \supseteq \mathrm{SB}(G)$ is immediate from Definition 45. For the finiteness of $C(\mathcal{W}(\mathbf{S}))$, we may assume that every filter of $\mathbf{S}$ covering $u \in V$ belongs to $\mathbf{S}_u$ (as we can always achieve this by reorganizing filters into the sets $\mathbf{S}_v$).

  Consider now the following procedure for generating $\mathcal{W}(\mathbf{S})$. Start with $A = V$. For a walk $w \in A$ and $v \in V$, we add wv to A iff for every $S \in \mathbf{S}_v$, the greatest suffix of wv fully contained in S, belongs to is self-bounding. It is clear that after k steps this process generates all walks of $\mathcal{W}(\mathbf{S})$ of size at most k.

  Let $\delta = \max_{i\in[t]}\{\sum_{S\in\mathbf{S}_{x_i}} |S|\}$; since each $\mathbf{S}_{x_i}$ is a finite family of finite filters, $\delta < \infty$.

  Since a self-bounding walk visit a vertex at most one, the greatest suffix of wv fully contained in a filter $S \in \mathbf{S}_v$, has length at most δ. Therefore, w is equivalent with its δ-suffix.

  Let now $\mathcal{W}_x(\mathbf{S})$ be the set of all walks in $\mathcal{W}(\mathbf{S})$ that end up in vertex x. Since the number of walks in $\mathcal{W}_x(\mathbf{S})$ of length at most δ is bounded above by δ^δ, $C(\mathcal{W}_x(\mathbf{S})) \le \delta^\delta$. Finally, per the quasi-transitivity of G, we have $C(\mathcal{K}) = t|C(\mathcal{K}_x)| < t\delta^\delta$.

- Since $\mathcal{W}(\mathbf{S}) \supseteq \mathrm{SB}(G)$, we have that $\mathcal{W}(\mathbf{S}) \supseteq \mathrm{SB}(H)$ for every finite sugraph H,of G. Hence, per Theorem 43, equation (2.2) implies that $\lambda\mathbf{1}$ is good for H. The result follows from the definition of λ_c.

■

2.6 Explaining Our Lattice Numbers

In practice, it is convenient to choose the filters $\mathbf{S}_{x_i}$ in a uniform manner, so that $\mathcal{W}(\mathbf{S})$ has a simple description. One such choice is to fix an integer q (the larger the better), and take as filters $\mathbf{S}_{x_i} = \mathbf{S}_{x_i}(q)$ the set of all connected subgraphs of G of size q that contain x_i. Notice that these filters give rise to the set $\mathcal{W}(\mathbf{S})$ comprised by all walks on G, whose q-prefix is self-bounding.

Since the continuations of a walk are completely determined by its q-prefix and the orbit of its terminal vertex, two walks agreeing on the orbit of their terminal vertex, and on their q-prefix are equivalent. We identify a class of walks with a pair (w, i), where w is a word of length at most q on the label-alphabet of G, and $i \in [t]$, such that the walk performed on G that moves according to w and terminates on x_i is self-bounding. We call such a pair SB-*realizable*. Specializing Theorem 46 for the above case gives

Corollary 54. *Assume that for every* SB-realizable *pair* (w, i) *there exist number* $0 \leq r_{w,i} < 1$*, such that*

$$\lambda \leq r_{w,i} \prod_{(z,j) \leftarrow (w,i)} (1 - r_{z,j}) \ . \tag{2.15}$$

Then, $\lambda \leq \lambda_c$.

To derive the numbers presented in Table 2.1, we apply Corollary 54. In particular, in all three lattices we consider, we label their arcs so that geometrically parallel arcs pointing to the same direction receive the same label. For $\mathbb{Z}_2$, and $\mathbb{Z}_3$ we take $x = (0, 0)$, and $x = (0, 0, 0)$, respectively, to be the origin of the lattice. Then, fixing the integer q (the greatest value of q we were able to run in a reasonable time was 15), we generate all self-bounding walks of length up to q keeping track of the causality relationship between them in a list.

The situation is pretty similar for the hexagonal lattice, except from the fact that label-preserving automorphisms partition the set of vertices in two orbits. Indeed, think of painting the vertices of the lattice black and white in a chessboard manner. It is easy to observe that for any two vertices of the same color there is a label-preserving translation mapping one to the other, while the labels of the outgoing arcs of any two vertices of different color are disjoint, and thus, no label-preserving map can send one to the other. To account for this, we distinguish one white vertex representative x, say $(0, 0)$, and one black vertex representative y, say $(1, 0)$, and consider the set of walks ending at x or y, that are q-prefix-self-bounding. Two walks would now be equivalent if the color of their terminal vertex is the same and they have the same q-prefix. To check wether an activity λ is valid for the above set of walks, we have to check wether the (huge but) finite system (2.2) has a solution. Here we use the line of thinking of Lemma 53: starting from all variables of (2.2) equal to zero, we iteratively update the value of each variable to satisfy the equation it appears (outside of the product). We either reach a (approximate) fixed point satisfying each equation, or on the k-th iteration some r_i becomes greater than 1, in which case we declare λ is not valid . Doing binary search on the values λ we find quickly a good estimation of the maximum valid λ.

Chapter 3

Submodularity is All that Matters

We show that the independent set polynomial with negative arguments is useful in a far more general setting than that of the Lovász Local Lemma and the Probabilistic Method. Specifically, we prove that the independent set polynomial provides a lower bound for general supermodular functions that are also “mostly” log-supermodular. This is a very natural class of functions corresponding to the (shrinking/decreasing) volume of sets as they are subjected to a sequence of “mostly symmetric” restrictions. As a fist application we recover the (standard) Quantum Lovász Local Lemma effortlessly. The fact that we do so through the independent set polynomial, immediately makes available all the improvements and specializations of the lemma that we have derived, which are new in the quantum setting.

3.1 A generalization of Shearer’s result

Given a graph G on $[n]$, Shearer [1] characterized the probability vectors $\boldsymbol{p} = \boldsymbol{p}(G) \in [0,1]^n$, such that under *every* measure defined on n events, where the i-th event occurs with probability at most p_i, and is independent from its non-neighbors in G, the probability that none of the events occurs is strictly positive.

Theorem 55 (Shearer [1]). *Let Ω be an arbitrary set and let $F_1, \ldots, F_n$ be arbitrary subsets of Ω. Let G be a graph on $[n]$, and $\boldsymbol{p} \in [0,1]^n$. The following statements are equivalent:*

(i) $\mu\left(\bigcap_{i\in[n]} \overline{F}_n\right) > 0$*, for every measure* μ *on* Ω *such that* $\mu(F_i) = p_i$ *and* G *is a dependency graph for* $\{F_i\}_{i=1}^n$*.*

(ii) $Z_G(-\boldsymbol{p}; S) > 0$*, for every* $S \subseteq [n]$*.*

Moreover, whenever (ii) holds, $\mu\left(\cap_{i\in[n]}\overline{F}_n\right) \geq Z_G(-\boldsymbol{p}) > 0$*.*

In this section, we give a new, simple proof of Shearer's classic result, demonstrating that the *only* property of the function μ needed for the proof to go through is supermodularity. In particular, μ does not need to be a probability measure. This simple observation greatly broadens the scope of Theorem 55, giving as immediate corollaries (i) the celebrated "Quantum LLL", a geometric variant of the classic LLL [41], and (ii) a tight, universal, lower bound on the partition function of arbitrary graphical models under minimal assumptions. Moreover, the fact that this lower bound comes from an evaluation of the independent set polynomial with negative real inputs makes *every* local lower bound, such as the LLL and our improvements in Part I, applicable, when a tractable (but lossy) lower bound is desirable.

3.1.1 Our Setting

Let $f : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$ be a non-negative real function on the binary cube. We will find it useful to also think of f as a function on the subsets of $[n]$, by thinking of each $x \in \{0,1\}^n$ as the characteristic vector of a set $S \in 2^{[n]}$. For $i \in [n]$, and $S \subseteq [n] \setminus \{i\}$, let

$$\Delta_i f(S) := f(S \cup \{i\}) - f(S), \text{ and } f(i \mid S) := \frac{f(S \cup \{i\})}{f(S)} \ .$$

Thus, $\Delta_i f(S)$ is the *marginal difference* of adding i to S, while $f(i \mid S)$ is the *marginal ratio* of adding i to S. The former is also known as the *discrete derivative* of f with respect to i at S, as it measures the *additive* change of f when moving along the direction of i, and plays a fundamental role in the analysis of functions on the Boolean cube. The latter also measures the change of f along the direction of i, but in a *multiplicative* sense, and is far less studied in the literature. The closest, very well-studied but far coarser, notion is that of *curvature*, defined as $f(S \cup \{i\})/f(\{i\})$ and introduced in [42]. The following simple equation relates the two quantities:

$$f(i \mid S) = 1 + \frac{\Delta_i f(S)}{f(S)} \ . \tag{3.1}$$

We say f is *increasing* if $f(S) \leq f(T)$ for all $S \subseteq T$, and *decreasing* if the reverse inequality holds. We call f *modular* if $\Delta_i f$ is independent of S, for all $i \in [n]$. Also, f is *supermodular* if $\Delta_i f$ is increasing, and *submodular* if $\Delta_i f$ is decreasing. Finally, f is *log-modular*, *log-supermodular*, and *log-submodular*, if for all $i \in [n]$, the quantity $f(i \mid S)$ is constant, increasing, and decreasing with respect to S, respectively.

We will be interested in decreasing, supermodular functions, i.e., f whose decrease per additional element in the input is a negative increasing function, thus capturing the familiar notion of "diminishing returns," just as in the more familiar setting of increasing submodular, i.e., *polymatroid*, functions, which characterize matroids (since f is decreasing, $\Delta_i f$ is negative and, thus, $\Delta_i f$ being increasing (supermodularity) makes $|\Delta_i f|$ decreasing (diminishing returns)). Since we will seek to bound f from below, naturally, we need some control over its *rate* of decrease. This will come in the form of lower bounds for *some* of its marginal ratios. Specifically, we require that the marginal ratios information takes the form of a graph G on $[n]$ and a vector $\boldsymbol{p} \in [0,1)^n$, such that for each $i \in [n]$, and each set $S \subseteq \Phi_i$, where $\Phi_i = [n] - \Gamma_i$ comprising only *non*-neighbors of i in G,

$$f(i \mid S) \geq 1 - p_i \ . \tag{3.2}$$

In other words, the neighbors of i in G are the only "uncontrolled" dimensions for dimension i: if S contains no neighbors of i, introducing i to form $i \cup S$ can cause f to shrink by a factor of at most p_i. When (3.2) holds, we say that f *factorizes according to* $G, \boldsymbol{p}$.

To see that our framework includes the probabilistic setting of Theorem 55, let μ be a probability measure and G a dependency graph for $\{F_1, \ldots, F_n\}$, where $\mu(F_i) =$

p_i for all $i \in [n]$, and define $f(S) := \mu\left(\cap_{i\in S}\overline{F_i}\right)$. Clearly, f is non-negative and decreasing. To see that it is also supermodular, observe that for every $i \in [n]$, and $S \subseteq T \subseteq [n] \setminus \{i\}$,

$$\Delta_i f(S) = \mu\left(\overline{F_i} \cap \bigcap_{j\in S} \overline{F_j}\right) - \mu\left(\bigcap_{j\in S} \overline{F_j}\right) \tag{3.3}$$

$$= -\mu\left(F_i \cap \bigcap_{j\in S} \overline{F_j}\right) \tag{3.4}$$

$$\leq -\mu\left(F_i \cap \bigcap_{j\in T} \overline{F_j}\right) = \Delta_i f(T) \ . \tag{3.5}$$

Finally, it is easy to check that G is a dependency graph of μ iff $f(i \mid S) = 1 - p_i$, for all $i \in [n]$, and $S \subseteq \Phi_i$.

3.1.2 Our Result

Theorem 56. *Let G be a graph on $[n]$, and $\boldsymbol{p} \in [0,1)^n$. The following statements are equivalent:*

(i) Every $f : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$ that is supermodular and factorizes according to $G, \boldsymbol{p}$ is strictly positive.

(ii) $Z_G(-\boldsymbol{p}; S) > 0$, for every $S \subseteq [n]$.

Moreover, whenever (ii) holds, $f(\boldsymbol{x}) \geq f(\boldsymbol{0}) \cdot Z_G(-p_1x_1, -p_2x_2, \ldots, -p_nx_n)$.

Proof of Theorem 56.

To establish that (ii) implies (i) it suffices to prove that $f(i \mid S) \geq Z_G\left(-\boldsymbol{p}; (i \mid S)\right) =: Z\left(i \mid S\right)$, for every $i \in [n]$, and $S \subseteq [n] \setminus \{i\}$. We prove this by induction on $|S|$. Assume the claim holds for every $i \in [n]$ and every proper subset of S. Let $\{j_1, \ldots, j_d\}$ be an ordering of $S \setminus \Phi_i$ and write $S_\ell = S \setminus \{j_1, \ldots, j_\ell\}$. Then, the first inequality in (3.6) follows from the supermodularity of f, while the second follows from the fact that f factorizes according to $G, \boldsymbol{p}$.

$$f(i \mid S) := 1 + \frac{\Delta_i f(S)}{f(S)} \tag{3.6}$$

$$\geq 1 + \frac{\Delta_i f(S \cap \Phi_i)}{f(S)} \tag{3.7}$$

$$= 1 + \left(f\left(i \mid S \cap \Phi_i\right) - 1\right) \frac{f(S \cap \Phi_i)}{f(S)} \tag{3.8}$$

$$\geq 1 - p_i \frac{f(S \cap \Phi_i)}{f(S)} \quad . \tag{3.9}$$

By telescoping, the r.h.s. of (3.6) equals the l.h.s. of (3.10) below, the inequality in (3.10) follows from the inductive hypothesis, while the remaining two equalities follow from the fundamental recurrence (1.15) for the independent set polynomial and the definition of $Z(i \mid S)$.

$$1 - p_i \prod_{\ell=1}^{d} \frac{1}{f(j_\ell \mid S_\ell)} \geq 1 - p_i \prod_{\ell=1}^{d} \frac{1}{Z(j_\ell \mid S_\ell)} = 1 - p_i \frac{Z(-\boldsymbol{p}; S \cap \Phi_i)}{Z(-\boldsymbol{p}; S)} = Z(i \mid S) \ . \tag{3.10}$$

To prove that (i) implies (ii) we prove the contrapositive, i.e., that if $Z_G(-\boldsymbol{p}; S) \leq 0$ for some $S \subseteq [n]$, then we can find $f : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$ that is supermodular and factorizes according to $G, \boldsymbol{p}$, such that $f([n]) = 0$. To do this, let $\lambda = \min\{\theta : \exists S \subseteq [n]$ such that $Z_G(-\theta\boldsymbol{p}) \leq 0\}$ and take $f = Z_G(-\lambda\boldsymbol{p})$. Thus, f inherits factorization according to $G, \boldsymbol{p}$ from Z_G. To prove supermodularity we invoke the fact that Z_G is supermodular if $Z_G \geq 0$ and observe that, by the continuity of the independent set polynomial, if $Z_G(-\lambda\boldsymbol{p}; S) < 0$ for some $S \subseteq [n]$, then λ would not be minimal.

Finally, we observe that if (ii) holds, then we can use telescoping to bound $f(\boldsymbol{x})$ from below as

$$f(\boldsymbol{x}) = f(\mathbf{0}) \prod_{i:x_i=1} f(i \mid \{j < i : x_j = 1\}) \tag{3.11}$$

$$\leq f(\mathbf{0}) \prod_{i:x_i=1} Z(i \mid \{j < i : x_j = 1\}) = f(\mathbf{0}) Z_G(-p_1 x_1, \ldots, -p_n x_n) \ . \tag{3.12}$$

■

3.1.3 Quantum Local Lemma

Let V be a vector space over some field. For a subspace $X \subseteq V$, write $R(X) := \frac{\dim(X)}{\dim(V)}$, for the relative dimension of X with respect to V.

Definition 57. *Say that X is* mutually independent *from the subspaces $Y_1, \ldots, Y_m$, if for all $S \subseteq [m]$, $R(X \cap_{j\in S} Y_j) = R(X)R(\cap_{j\in S} Y_j)$.*

Definition 58. *The graph G on $[m]$ is a R-*dependency graph *for $X_1, \ldots, X_m \subseteq V$ if X_i is mutually independent from $\{X_j\}_{j\in S}$, for any $S \subseteq \Phi_i$.*

Let $X_1, \ldots, X_m$ be subspaces of V, and define $f(\emptyset) = 1$, and $f(S) = R\left(\cap_{j\in S} X_j\right)$, for $S \subseteq [n] \setminus \{i\}$. Then f is non-negative, decreasing, and supermodular. Indded, clearly f is non-negative and decreasing.

To prove supermodularity, we need to show that $\Delta_i f(S) \le \Delta_i f(T)$, for every $i \in [n]$ and $S \subseteq T \subseteq [n] - \{i\}$. Write $A := \bigcap_{j\in S} X_j$, $A_i := A \cap X_i$, and $B := \bigcap_{j\in T} X_j$. Observe now that $\Delta_i f(S) = R\left(A \cap X_i\right) - R\left(A\right) = \left(\dim\left(A_i\right) - \dim\left(A\right)\right)/\dim(V)$ and $\Delta_i f(T) = R\left(B \cap X_i\right) - R\left(B\right) = \left(\dim\left(B \cap X_i\right) - \dim\left(B\right)\right)/\dim(V)$. Since A_i, B are subspaces of finite-dimensional vector space A, it follows that the dimension of vector space $A_i + B = \{u + v : u \in A_i, v \in B\}$ is trivially upper bounded by $\dim(A)$ and equals $\dim(A_i) + \dim(B) - \dim(A_i \cap B)$ (see, e.g., p.47 of [65]). Thus, $\dim\left(A_i\right) - \dim\left(A\right) \le \dim\left(A_i\right) - \dim\left(A_i + B\right) = \dim\left(A_i \cap B\right) - \dim\left(B\right) = \dim\left(X_i \cap B\right) - \dim\left(B\right)$, implying $\Delta_i f(S) \le \Delta_i f(T)$, as required.

If additionally $R(X_i) = 1 - p_i$ for all $i \in [n]$, and G is a R-dependency graph for $\{X_1, \ldots, X_m\}$, then f is compatible with $G, \boldsymbol{p}$. Indeed, for $i \in [m]$, and $S \subseteq \Phi_i$

$$f(i \mid S) = \frac{R\left(X_i \cap \bigcap_{j\in S} X_j\right)}{R\left(\bigcap_{j\in S} X_j\right)} = \frac{R(X_i)R(\cap_{j\in S} X_j)}{R\left(\bigcap_{j\in S} X_j\right)} = R(X_i) = 1 - p_i \ . \tag{3.13}$$

Applying Theorem 56 to f gives the following result of Sattath et al. [43]

Theorem 59. *Let* $X_1, \ldots, X_m$ *be subspaces of* V *with* $R(X_i) = 1 - p_i$, *and* R-*dependency graph* G. *Then* $R\left(\cap_{j\in[m]} X_j\right) > 0$ *iff* $Z_G(-\boldsymbol{p}; S) > 0$ *for every* $S \subseteq [m]$.

Combining with Theorem 4 from Part I, we get the Quantum Local Lemma of Ambainis et al. [41].

Theorem 60 (Quantum LLL)**.** *Let* $X_1, \ldots, X_m$ *be subspaces of* V *with dependency graph* G, *and* $R(X_i) = 1 - p_i$. *If there exist* $\boldsymbol{r} \in [0,1)^m$ *such that* $p_i \le r_i \prod_{j\in\Gamma_i}(1 - r_i)$, *then,* $R(\cap_{i\in[n]} X_i) \ge \prod_{i\in[n]}(1 - r_i)$.

3.2 Future Work

- Devise an *algorithmic* version of Theorem 56 with potential applications in submodular optimization.

- Given a function $f : D^n \mapsto \mathbb{R}_{\geq 0}$ where $f(\boldsymbol{x}) = \prod_{i=1}^{m} f_i(\{x\}_i)$, form a graph G on $[m]$ were i, j are adjacent if $\{x\}_i \cap \{x\}_j \neq \emptyset$. Use Theorem 56 to derive lower bounds for $Z(f) = \sum_{\boldsymbol{x} \in D^n} f(\boldsymbol{x})$ when G is sparse.

- Formulate and apply the new quantum LLL that result by incorporating the results of Chapter **??** to the quantum settting.

- In a certain sense, the so-called quantum LLL is "not really quantum" but rather geometric, as no quantum notions enter our Theorem 56, which readily yields all known quantum LLLs. We would like to devise a truly quantum LLL, by relaxing the supermodularity assumption our abstract framework.

Chapter 4

Random Lifts and the Bethe Approximation

We now turn on the Bethe approximation for partition functions of general graphical models. While, a priori, there is no connection between the (analytically defined) Bethe approximation and the independent set polynomial, we use a recent combinatorial characterization of the Bethe approximation by Vontobel [44] to suggest precisely such a connection, by relating typical random k-lifts of graphs where $k \to \infty$, with the aforementioned tree of non-backtracking walks. In particular, we revisit a recent result of Ruozzi showing that the Bethe partition function is a lower bound for the true partition function, for every graphical model whose constituent factors are log-supermodular. We give a much shorter proof of this result. More importantly, we give a new much shorter proof of the celebrated four functions theorem.

4.1 Random Lifts of Graphs

Imagine representing a graph $G = (V, E)$ using $|V|$ coins and $|E|$ pieces of string, by spreading the coins on a table and tying together each pair of coins that correspond to an edge in E. To form a k-lift of G for some $k \geq 2$, we take another $(k-1)|V|$ coins and $(k-1)|E|$ pieces of string, as follows. First, we place $k-1$ new coins below each existing coin on the table, so that the graph G is lifted off-the table. Then, for each tied pair of top-tier coins corresponding to some edge $e = \{u, v\} \in E$ we: untie the two coins; pick a k-permutation π_e; connect the i-th coin in the u-stack to the $\pi(i)$-th coin in the v-stack. That is, each edge of G is now mapped to a k-matching. Clearly, if all permutations chosen are identical, we simply have k copies of G. If, on the other hand, the k-permutations are chosen uniformly and independently, we have a (uniformly) random k-lift of G. (We give a formal definition shortly.)

Random lifts of graphs form a relatively new, reasonably well-studied area, where the primary focus has been to establish different combinatorial properties of lifted graphs, primarily expansion. What makes random lifts most interesting to us is the following fact: lifting a graph does not introduce any new cycles, i.e., every cycle of the lifted graph corresponds to a unique cycle of G; moreover, the length of each cycle is typically increases by a factor of $\Theta(k)$. Thus, as $k \to \infty$, we expect that a typical random k-lift is (nearly) devoid of cycles. It is not hard to see (and is widely known), that the only tree that can (and does) serve as the limiting object for random k-lifts of a graph G is the graph's so-called *universal cover*, which is nothing but the tree of non-1-backtraking walks on G that we met in Part I.

Another fact that makes random lifts fascinating is that they can be refined so as to introduce the notion of a *vantage point* of a graph. Specifically, as seen from the ceiling, a lift is identical to the underlying graph G. But as seen from the side of the table, a far more complex picture emerges. Things get particularly interesting if one limits the choice of permutations. Specifically, imagine that instead of allowing all $k!$ permutations between two stacks of coins, we chose some small number $q \geq 0$ and only allowed permutations such that $|i - \pi(i)| \leq q$, for every $i \in [k]$ (treating coins 1 and k as adjacent). Such lifts, known as *spatially coupled* random lifts, only allow permutations where each piece of string tying two coins is near-parallel to the ground (as q is fixed while k grows). As it turns out, even for q as small as 2 or 3, the resulting lifts inherit most of the desirable properties of random lifts, while having truly remarkable *computational* properties. These properties have so far been used to achieve landmark results in coding theory [45, 46] and to establish lower bounds for random Constraint Satisfaction Problems [47, 48].

Our long-term goal is to introduce random lifts, and ideally spatially coupled random lifts, to machine learning, via graphical models. Our starting step in this direction, has been to revisit the only application of random lifts to machine learning (that we are aware of) and greatly simplify its proof. In doing so, we also greatly simplified the proof of the celebrated *four functions theorem* of Ahlswede and Daykin [49].

4.2 Binary graphical models and the Bethe approximation

A graphical model, $\mathcal{G}$, on n binary variables is a function that maps $\boldsymbol{x} = (x^1, \ldots, x^n) \in \{0,1\}^n$ to $\prod_{\alpha \in \mathcal{A}} \psi_\alpha(\boldsymbol{x}^\alpha)$, where $\mathcal{A} \subseteq 2^{[n]}$ is arbitrary and where for each variable subset $\alpha \in \mathcal{A}$ the factor/function ψ_α maps $\{0,1\}^{|\alpha|}$ to $\mathbb{R}_{\geq 0}$. Graphical models are often represented as bipartite graphs called "factor graphs", with "variable" and "factor" vertices, where for each $\alpha \in \mathcal{A}$ the corresponding factor vertex is connected to the vertices corresponding to the variables in α.

Computing the partition function, $Z(\mathcal{G})$, of a graphical model, i.e., the sum of $\mathcal{G}(\boldsymbol{x})$ over all $\boldsymbol{x} \in \{0,1\}^n$, is a fundamental problem in combinatorics, physics, and machine learning. Exact solutions may be obtained via variable elimination or the junction tree method, but unless the tree-width of the factor graph of $\mathcal{G}$ is bounded, this can take exponential time. As a result, several approximate methods have been developed. Of particular note is the so called *Bethe approximation*, motivated by the work of Nobel laureate Hans Bethe. For the purposes of our discussion there is no need to define the Bethe approximation to the partition function, $Z_B(\mathcal{G})$, other than to the say that it is the minimum of a (highly non-convex) function over $2|E|$ variables. (Indeed, these variables are morally analogous to the $2|E|$ variables that appear in our Theorem 13. See the book of Wainright and Jordan[50] for a detailed treatment of Bethe approximation). While the Bethe approximation is often computationally tractable and the results can be quite accurate, in general it comes with no approximation guarantees, or even knowledge of whether $Z_B(\mathcal{G})$ is a lower or upper bound for $Z(\mathcal{G})$. One of very few rigorous results about the Bethe approximation, is due to Ruozzi [51] and forms the only application we are aware of of random lifts to machine learning. Recall that $f : \{0,1\}^k \mapsto \mathbb{R}_{\geq 0}$ is log-*supermodular* if $f(\boldsymbol{x})f(\boldsymbol{y}) \leq f(\boldsymbol{x} \wedge \boldsymbol{y})f(\boldsymbol{x} \vee \boldsymbol{y})$, for every $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^k$.

Theorem 61. *If all factors of a binary graphical model $\mathcal{G}$ are* log-*supermodular, then $Z(\mathcal{G}) \geq Z_{\mathrm{B}}(\mathcal{G})$.*

To prove Theorem 61, Ruozzi relied heavily on (i) a beautiful combinatorial characterization of the Bethe partition function of arbitrary (non-negative) graphical models in terms of random lifts by Vontobel [44], which we discuss next, and (ii) the *$2k$-functions theorem* of Rinott and Saks [52] and Aharoni and Keich [53], a generalization of the celebrated *four functions theorem* of Ahlswede and Daykin [49], which we discuss in Section 4.1.

4.2.1 The Bethe partition function in term of random lifts

For $A \in \{0,1\}^{k \times n}$ and $\alpha \subseteq [n]$, write A^α for the matrix obtained by dropping the columns of A outside α. For $i \in [k], j \in [n]$ write A_{i*}, A_{*j} for the i-th row and j-th column of A, respectively. For an n-tuple of k-permutations $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n) \in (S_k)^n$, write $\boldsymbol{\sigma} A$ for the matrix obtained by permuting the elements of A_{*j} according to σ_j, for each $j \in [n]$.

Definition 62. *For $k \in \mathbb{N}$, the k-*fold product *of ψ_α is the function $\Psi_\alpha := \Psi_\alpha^{(k)}$ mapping $A \in \{0,1\}^{k \times |\alpha|}$ to $\prod_{i \in [k]} \psi_\alpha(A_{i*})$. For a graphical model $\mathcal{G}$, if $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}^\alpha\}_{\alpha \in \mathcal{A}}$ comprises an $|\alpha|$-tuple of k-permutations for each $\alpha \in \mathcal{A}$, then the k-*lift of $\mathcal{G}$ corresponding to $\boldsymbol{\sigma}$ *is the function $\mathcal{G}_{\boldsymbol{\sigma}}$ mapping $A \in \{0,1\}^{k \times n}$ to $\prod_{\alpha \in \mathcal{A}} \Psi_\alpha(\boldsymbol{\sigma}^\alpha A^\alpha)$.*

Definition 63. *We denote the set of all k-lifts of a graphical model $\mathcal{G}$ by $\mathcal{S}^k(\mathcal{G})$. We say that $\mathcal{G}$ is* lift-decreasing *if $Z(\mathcal{G}_{\boldsymbol{\sigma}})$ is maximum when every k-permutation appearing in $\boldsymbol{\sigma}$ is the identity (correspondingly for* lift-increasing*).*

While, in general, $Z_{\mathrm{B}}(G)$ can be more or less than $Z(G)$, Vontobel [44] proved the following.

Theorem 64.

$$Z_{\mathrm{B}}(\mathcal{G}) = \limsup_{k\to\infty} \left(\underset{\mathcal{G}_{\boldsymbol{\sigma}}\in\mathcal{S}^k(\mathcal{G})}{\mathrm{Avg}} Z\left(\mathcal{G}_{\boldsymbol{\sigma}}\right) \right)^{1/k} .$$

To prove Theorem 61, Ruozzi invoked Theorem 64, but did not have to actually perform the associated averaging, as he showed the following very strong proposition.

Theorem 65. *If all factors of a binary graphical model are* log-*supermodular, then it is lift-decreasing.*

4.3 The $2k$-functions theorem and our contribution

As mentioned, besides exploiting Vontobel's combinatorial characterization of the Bethe partition function, the other main step in Ruozzi's proof is establishing a variation of the $2k$-functions theorem [53], used to establish Theorem 65. To state the $2k$-functions theorem, we need to introduce the following notation.

Definition 66. *For an arbitrary finite-valued function f on a finite domain D we denote $Z(f) = \sum_{x\in D} f(x)$.*

Definition 67. *For a 0-1 matrix A, let $\uparrow\! A$ be the matrix obtained by independently sorting each column of A so that the 1s are on top and $\overleftarrow{A}$ be the matrix obtained by independently sorting each row of A so that the 1s are on the left.*

It is worth pointing out that instantiating the $2k$-functions theorem, below, for $k = 2$, yields the celebrated *four functions theorem* of Ahlswede and Daykin [49], which, in turn, readily implies the famous FKG inequality [54], the Holley inequality [55], and the Fishburn-Shepp [56, 57] inequality.

Theorem 68 (2k-functions)**.** *If $f_1, \ldots, f_k, g_1, \ldots, g_k : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$, are such that for all $X \in \{0,1\}^{k \times n}$,*

$$\prod_{i \in [k]} g_i\left(X_{i*}\right) \leq \prod_{i \in [k]} f_i\left((\uparrow X)_{i*}\right) \quad , \tag{4.1}$$

then $\prod_{i \in [k]} Z(g_i) \leq \prod_{i \in [k]} Z(f_i)$.

Ruozzi proved the following variation of the 2k-functions theorem.

Theorem 69. *Let $f_1, \ldots, f_k : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$ and $g : \{0,1\}^{k \times n} \mapsto \mathbb{R}_{\geq 0}$, be such that for all $X \in \{0,1\}^{k \times n}$,*

$$g\left(X\right) \leq \prod_{i \in [k]} f_i\left((\uparrow X)_{i*}\right) \quad . \tag{4.2}$$

If g is log-*supermodular, then, $Z(g) \leq \prod_{i \in [k]} Z(f_i)$.*

In other words, while the right hand side of (4.2) is the same as the right hand side of (4.1), on the left hand side Ruozzi's variant replaces the product of k functions on $\{0,1\}^n$ with a single log-supermodular function on $\{0,1\}^{kn}$.

4.3.1 Our Contribution

We make two contributions. The first is to dramatically shorten Ruozzi's proof of Theorem 65 given Theorem 69. To state the proof we need to introduce the following notation.

Definition 70. *A function $g : \{0,1\}^{k \times n} \mapsto \mathbb{R}_{\geq 0}$ is a k-*product*, if $g\left(X\right) = \prod_{i \in [k]} g_i\left(X_{i*}\right)$, where $g_1, \ldots, g_k$ are arbitrary positive functions on $\{0,1\}^n$.*

Definition 71. *For a 0-1 matrix A, let $\uparrow A$ be the matrix obtained by independently sorting each column of A so that the 1s are on top and $\overleftarrow{A}$ be the matrix obtained by independently sorting each row of A so that the 1s are on the left.*

Proof of Theorem 65. Let $\mathcal{G}^k$ and $\mathcal{G}_{\boldsymbol{\sigma}}$, be the trivial and an arbitrary k-lift of $\mathcal{G}$, respectively. We apply Theorem 69 with $f = \mathcal{G}^k$ and $g = \mathcal{G}_\sigma$. Specifically, $\mathcal{G}^k$ is a k-product (corresponding to the trivial k-lift), $\mathcal{G}_{\boldsymbol{\sigma}}$ is log-supermodular since it is the product of log-supermodular factors, and $\mathcal{G}^k, \mathcal{G}_{\boldsymbol{\sigma}}$ satisfy condition (4.2), as for all $A \in \{0,1\}^{k \times n}$,

$$\mathcal{G}_{\boldsymbol{\sigma}}(A) = \prod_{\alpha \in \mathcal{A}} \Psi_\alpha\left(\boldsymbol{\sigma}^\alpha A^\alpha\right) \leq \prod_{\alpha \in \mathcal{A}} \Psi_\alpha\left(\uparrow(\boldsymbol{\sigma}^\alpha A^\alpha)\right) = \prod_{\alpha \in \mathcal{A}} \Psi_\alpha\left(\uparrow(A^\alpha)\right) = \prod_{\alpha \in \mathcal{A}} \Psi_\alpha\left((\uparrow A)^\alpha\right) = \mathcal{G}^k(\uparrow A) \ . \tag{4.3}$$

■

Our second contribution is to give a substantially simpler proof of both Ruozzi's variant of the four functions theorem and of the theorem itself by proving the following, trivially implying Theorem 68 and Theorem 69.

Theorem 72. *Let $f_1, \ldots, f_k : \{0,1\}^n \mapsto \mathbb{R}_{\geq 0}$ be arbitrary and for $X \in \{0,1\}^{k \times n}$ write $f(X) := \prod_{i \in [k]} f_i\left(X_{i*}\right)$. Let $g : \{0,1\}^{k \times n} \mapsto \mathbb{R}_{\geq 0}$ be such that for every $X \in \{0,1\}^{k \times n}$,*

$$g\left(X\right) \leq f(\uparrow X) \ . \tag{4.4}$$

If g is log*-supermodular or a k-product, then $Z(g) \leq Z(f)$.*

4.4 Proof of Theorem 72

Definition 73. *If $g(\boldsymbol{x}) \leq f\left(\uparrow \boldsymbol{x}\right)$ for every $\boldsymbol{x} \in \{0,1\}^k$, we say that f dominates g.*

We start by proving the following powerful lemma, which will form the base case for our later induction.

Lemma 74. *Let $f, g : \{0,1\}^k \mapsto \mathbb{R}_{\geq 0}$. If g is* log*-supermodular, f is* log*-submodular, and f dominates g, then $Z(g) \leq Z(f)$.*

Proof. For this proof only, it will be convenient to define the t-fold product of f by organizing its t inputs in columns instead of rows, i.e., $F = F^{(t)} : \{0,1\}^{k \times t} \mapsto \mathbb{R}_{\geq 0}$, where $F(A) = \prod_{i \in [t]} f(A_{*i})$. Let us say that g is *(weakly)* log*-majorized* by f and write $g \prec_{\log} f$, if for every $A \in \{0,1\}^{k \times t}$ with distinct columns, there exists $B \in \{0,1\}^{k \times t}$ with distinct columns, such that $G(A) \leq F(B)$. It is well-known (and easy to see) that if $g \prec_{\log} f$, then $Z(g) \leq Z(f)$.

To prove $g \prec_{\log} f$, fix any $t \geq 1$. Let $A \in \{0,1\}^{k\times t}$ have distinct rows and let π be any permutation of the rows of A in order of weight (number of ones). Write $B = \pi(A)$. Observe that $\uparrow\left(\overleftarrow{A}\right) = \pi\left(\overleftarrow{A}\right) = \overleftarrow{\pi(A)} = \overleftarrow{B}$ and that B has distinct columns as it is the result of permuting rows of a matrix with distinct columns. Therefore, $g \prec_{\log} f$, since

$$G\left(A\right) \leq G\left(\overleftarrow{A}\right) \leq F\left(\uparrow\left(\overleftarrow{A}\right)\right) = F\left(\overleftarrow{B}\right) \leq F\left(B\right) \ ,$$

where the second inequality follows from the fact that f dominates g, while the first and last inequalities follow from the log-supermodularity of g and the log-submodularity of f, respectively. ■

Armed with Lemma 74, we prove Theorem 72 by induction on n. Specifically, let $P(d)$ be the proposition that Theorem 72 holds for $n = d$. For $n = 1$, note that hypothesis (4.4) states that f dominates g, and that f is log-modular, and therefore, log-submodular. Similarly, g is log-supermodular under either of the two assumptions for it. Hence, Lemma 74 applies, yielding $Z(g) \geq Z(f)$, i.e., $P(1)$.

For $n \geq 2$, let $\overline{g}, \overline{f}$ be the the average of g, f over the n-th column, respectively. Note that if g is log-supermodular, or k-product, then $\overline{g}$ being a marginal of g, remains log-supermodular, or k-product, respectively. We will prove, using $P(1)$, that $\overline{g}, \overline{f}$ satisfy (4.4), so that $P(n-1)$ implies $Z(\overline{g}) \leq Z(\overline{f})$, i.e., $P(n)$ (since $Z(\overline{g}) = Z(g)$ and $Z(\overline{f}) = Z(f)$).

To prove that $\overline{g}, \overline{f}$ satisfy (4.4), fix any $Y \in \{0,1\}^{k\times(n-1)}$. For $\boldsymbol{t} \in \{0,1\}^{k\times 1}$, write $Y' = \left(Y \mid \boldsymbol{t}\right) \in \{0,1\}^{n\times k}$, and define $g_* : \{0,1\}^k \mapsto \mathbb{R}_{\geq 0}$, as $g_*(\boldsymbol{t}) = g\left(Y'\right)$, and $\widehat{f} : \{0,1\}^k \mapsto \mathbb{R}_{\geq 0}$ as $\widehat{f}(\boldsymbol{t}) = f\left(\uparrow Y \mid \boldsymbol{t}\right)$. The fact that g, f satisfy (4.4), implies that g_* is dominated by $\widehat{f}$, since $g_*(\boldsymbol{t}) = g\left(Y'\right) \leq f\left(\uparrow Y'\right) = \widehat{f}\left(\uparrow \boldsymbol{t}_i\right)$. Thus, invoking $P(1)$ for g_* and $\widehat{f}$, we conclude that $\overline{g}, \overline{f}$ satisfy (4.4), since $\overline{g}(Y) = Z\left(g_*(\boldsymbol{t})\right) \leq Z\left(\widehat{f}(\boldsymbol{t})\right) = \overline{f}\left(\uparrow Y\right)$.

Bibliography

[1] J. B. Shearer, "On a problem of Spencer," *Combinatorica*, vol. 5, no. 3, pp. 241–245, 1985. DOI: 10.1007/BF02579368. [Online]. Available: https://doi.org/10.1007/BF02579368.

[2] D. Weitz, "Counting independent sets up to the tree threshold," in *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, J. M. Kleinberg, Ed., ACM, 2006, pp. 140–149. DOI: 10.1145/1132516.1132538. [Online]. Available: https://doi.org/10.1145/1132516.1132538.

[3] A. Sly and N. Sun, "The computational hardness of counting in two-spin models on d-regular graphs," in *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, IEEE Computer Society, 2012, pp. 361–369. DOI: 10.1109/FOCS.2012.56. [Online]. Available: https://doi.org/10.1109/FOCS.2012.56.

[4] C. N. Yang and T. D. Lee, "Statistical theory of equations of state and phase transitions. i. theory of condensation," *Phys. Rev.*, vol. 87, pp. 404–409, 3 1952. DOI: 10.1103/PhysRev.87.404. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.87.404.

[5] T. D. Lee and C. N. Yang, "Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model," *Phys. Rev.*, vol. 87, pp. 410–419, 3 1952. DOI: 10.1103/PhysRev.87.410. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.87.410.

[6] T. Asano, "Lee-Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet," *Phys. Rev. Lett.*, vol. 24, pp. 1409–1411, 25 1970. DOI: 10.1103/PhysRevLett.24.1409. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.24.1409.

[7] O. J. Heilmann and E. H. Lieb, "Theory of monomer-dimer systems," *Comm. Math. Phys.*, vol. 25, no. 3, pp. 190–232, 1972. [Online]. Available: https://projecteuclid.org:443/euclid.cmp/1103857921.

[8] C. M. Newman, "Zeros of the partition function for generalized Ising systems," *Communications on Pure and Applied Mathematics*, vol. 27, no. 2, pp. 143–159, 1974. DOI: 10.1002/cpa.3160270203. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160270203. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160270203.

[9] M. Suzuki and M. E. Fisher, “Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models,” *Journal of Mathematical Physics*, vol. 12, no. 2, pp. 235–246, 1971. DOI: 10.1063/1.1665583. eprint: https://doi.org/10.1063/1.1665583. [Online]. Available: https://doi.org/10.1063/1.1665583.

[10] D. Ruelle, *Statistical mechanics: rigorous results*. Singapore: World Scientific, 1999. [Online]. Available: http://cds.cern.ch/record/392442.

[11] C. M. Newman, “The GHS inequality and the Riemann hypothesis,” *Constructive Approximation*, vol. 7, no. 1, pp. 389–399, 1991. DOI: 10.1007/BF01888165. [Online]. Available: https://doi.org/10.1007/BF01888165.

[12] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner, “Homogeneous multivariate polynomials with the half-plane property,” *Advances in Applied Mathematics*, vol. 32, no. 1, pp. 88 –187, 2004, Special Issue on the Tutte Polynomial, ISSN: 0196-8858. DOI: https://doi.org/10.1016/S0196-8858(03)00078-2. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196885803000782.

[13] M. Chudnovsky and P. Seymour, “The roots of the independence polynomial of a claw-free graph,” *J. Comb. Theory, Ser. B*, vol. 97, pp. 350–357, 2007.

[14] A. I. Barvinok, *Combinatorics and Complexity of Partition Functions* (Algorithms and combinatorics). Springer, 2016, vol. 30, ISBN: 978-3-319-51828-2. DOI: 10.1007/978-3-319-51829-9. [Online]. Available: https://doi.org/10.1007/978-3-319-51829-9.

[15] V. Patel and G. Regts, “Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials,” *SIAM J. Comput.*, vol. 46, no. 6, pp. 1893–1919, 2017. DOI: 10.1137/16M1101003. [Online]. Available: https://doi.org/10.1137/16M1101003.

[16] A. D. Scott and A. D. Sokal, “The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma,” *Journal of Statistical Physics*, vol. 118, pp. 1151–1261, 2005.

[17] N. Alon and J. H. Spencer, *The Probabilistic Method*, 4th. Wiley Publishing, 2016, ISBN: 1119061954.

[18] P. Erdős and L. Lovász, “Problems and results on 3-chromatic hypergraphs and some related questions,” *Coll Math Soc J Bolyai*, vol. 10, Jan. 1974.

[19] P. Erdös and J. Spencer, “Lopsided lovász local lemma and latin transversals,” *Discret. Appl. Math.*, vol. 30, no. 2-3, pp. 151–154, 1991. DOI: 10.1016/0166-218X(91)90040-4. [Online]. Available: https://doi.org/10.1016/0166-218X(91)90040-4.

[20] R. Bissacot, R. FernÁndez, A. Procacci, and B. Scoppola, “An improvement of the lovász local lemma via cluster expansion,” *Comb. Probab. Comput.*, vol. 20, no. 5, 709?719, Sep. 2011, ISSN: 0963-5483. DOI: 10.1017/S0963548311000253. [Online]. Available: https://doi.org/10.1017/S0963548311000253.

[21] R. J. Baxter, “Hard hexagons: Exact solution,” *Journal of Physics A: Mathematical and General*, vol. 13, no. 3, pp. L61–L70, 1980. DOI: 10.1088/0305-4470/13/3/007.

[22] Y. Okamoto, T. Uno, and R. Uehara, “Counting the number of independent sets in chordal graphs,” *Journal of Discrete Algorithms*, vol. 6, no. 2, pp. 229–242, 2008, Selected papers from CompBioNets 2004, ISSN: 1570-8667. DOI: https://doi.org/10.1016/j.jda.2006.07.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570866707000330.

[23] M. Heinrich and H. M?ller, *Counting independent sets in strongly orderable graphs*, 2021. arXiv: 2101.01997 [cs.DM].

[24] I. Bezáková and W. Sun, “Mixing of markov chains for independent sets on chordal graphs with bounded separators,” in *Computing and Combinatorics - 26th International Conference, COCOON 2020, Atlanta, GA, USA, August 29-31, 2020, Proceedings*, D. Kim, R. N. Uma, Z. Cai, and D. H. Lee, Eds., ser. Lecture Notes in Computer Science, vol. 12273, Springer, 2020, pp. 664–676. DOI: 10.1007/978-3-030-58150-3_54. [Online]. Available: https://doi.org/10.1007/978-3-030-58150-3_54.

[25] K. B. R. Kolipaka, M. Szegedy, and Y. Xu, “A sharper local lemma with improved applications,” in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012. Proceedings*, A. Gupta, K. Jansen, J. D. P. Rolim, and R. A. Servedio, Eds., ser. Lecture Notes in Computer Science, vol. 7408, Springer, 2012, pp. 603–614. DOI: 10.1007/978-3-642-32512-0_51. [Online]. Available: https://doi.org/10.1007/978-3-642-32512-0_51.

[26] N. J. A. Harvey, P. Srivastava, and J. Vondrák, “Computing the independence polynomial: From the tree threshold down to the roots,” in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, A. Czumaj, Ed., SIAM, 2018, pp. 1557–1576. DOI: 10.1137/1.9781611975031.102. [Online]. Available: https://doi.org/10.1137/1.9781611975031.102.

[27] H. Peters and G. Regts, “On a conjecture of sokal concerning roots of the independence polynomial,” *Michigan Math. J.*, vol. 68, no. 1, pp. 33–55, Apr. 2019. DOI: 10.1307/mmj/1541667626. [Online]. Available: https://doi.org/10.1307/mmj/1541667626.

[28] I. Bezáková, A. Galanis, L. A. Goldberg, and D. Stefankovic, “Inapproximability of the independent set polynomial in the complex plane,” in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, I. Diakonikolas, D. Kempe, and M. Henzinger, Eds., ACM, 2018, pp. 1234–1240. DOI: 10.1145/3188745.3188788. [Online]. Available: https://doi.org/10.1145/3188745.3188788.

[29] P. Erdős and J. Spencer, "Lopsided Lovász local lemma and latin transversals," *Discrete Applied Mathematics*, vol. 30, no. 2-3, pp. 151–154, 1991.

[30] R. Bissacot, R. Fernández, A. Procacci, and B. Scoppola, "An improvement of the lovász local lemma via cluster expansion," *Comb. Probab. Comput.*, vol. 20, no. 5, pp. 709–719, 2011. DOI: 10.1017/S0963548311000253. [Online]. Available: https://doi.org/10.1017/S0963548311000253.

[31] D. Achlioptas and K. Zampetakis, "Local approximations of the independent set polynomial," in *ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, ser. LIPIcs, vol. 198, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 8:1–8:16. DOI: 10.4230/LIPIcs.ICALP.2021.8. [Online]. Available: https://doi.org/10.4230/LIPIcs.ICALP.2021.8.

[32] D. S. Gaunt and M. E. Fisher, "Hard-Sphere Lattice Gases. I. Plane-Square Lattice," *Journal of Chemical Physics*, vol. 43, pp. 2840–2863, 1965.

[33] S. Todo, "Transfer-matrix study of negative-fugacity singularity of hard-core lattice gas," *International Journal of Modern Physics C*, vol. 10, no. 04, pp. 517–529, 1999.

[34] Y.-b. Chan, "Series expansions from the corner transfer matrix renormalization group method: The hard-squares model," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 8, p. 085 001, 2012.

[35] K. B. R. Kolipaka and M. Szegedy, "Moser and Tardos meet Lovász," in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, ser. STOC '11, San Jose, California, USA: Association for Computing Machinery, 2011, pp. 235–244, ISBN: 9781450306911. DOI: 10.1145/1993636.1993669. [Online]. Available: https://doi.org/10.1145/1993636.1993669.

[36] R. A. Moser and G. Tardos, "A constructive proof of the general lovász local lemma," *J. ACM*, vol. 57, no. 2, 2010, ISSN: 0004-5411. DOI: 10.1145/1667053.1667060. [Online]. Available: https://doi.org/10.1145/1667053.1667060.

[37] G. Regts, *Absence of zeros implies strong spatial mixing*, 2021. arXiv: 2111.04809 [math.PR].

[38] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin, "Spatial mixing and the connective constant: Optimal bounds," in *Proceedings of the 2015 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1549–1563. DOI: 10.1137/1.9781611973730.102. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611973730.102. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611973730.102.

[39] I. Bezáková, A. Galanis, L. A. Goldberg, and D. Štefankovič, "Inapproximability of the independent set polynomial in the complex plane," *SIAM Journal on Computing*, vol. 49, no. 5, STOC18–395–STOC18–448, 2020. DOI: 10.1137/18M1184485. eprint: https://doi.org/10.1137/18M1184485. [Online]. Available: https://doi.org/10.1137/18M1184485.

[40] J. C. Vera, E. Vigoda, and L. Yang, "Improved bounds on the phase transition for the hard-core model in 2 dimensions," *SIAM Journal on Discrete Mathematics*, vol. 29, no. 4, pp. 1895–1915, 2015. DOI: 10.1137/140976923. eprint: https://doi.org/10.1137/140976923. [Online]. Available: https://doi.org/10.1137/140976923.

[41] A. Ambainis, J. Kempe, and O. Sattath, "A quantum Lovász local lemma," *J. ACM*, vol. 59, no. 5, 24:1–24:24, 2012. DOI: 10.1145/2371656.2371659. [Online]. Available: https://doi.org/10.1145/2371656.2371659.

[42] M. Conforti and G. Cornuéjols, "Submodular set functions, matroids and the greedy algorithm: Tight worst-case bounds and some generalizations of the rado-edmonds theorem," *Discret. Appl. Math.*, vol. 7, no. 3, pp. 251–274, 1984. DOI: 10.1016/0166-218X(84)90003-9. [Online]. Available: https://doi.org/10.1016/0166-218X(84)90003-9.

[43] O. Sattath, S. C. Morampudi, C. R. Laumann, and R. Moessner, "When a local Hamiltonian must be frustration-free," *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6433–6437, 2016, ISSN: 0027-8424. DOI: 10.1073/pnas.1519833113. eprint: https://www.pnas.org/content/113/23/6433.full.pdf. [Online]. Available: https://www.pnas.org/content/113/23/6433.

[44] P. O. Vontobel, "Counting in graph covers: A combinatorial characterization of the bethe entropy function," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 6018–6048, 2013. DOI: 10.1109/TIT.2013.2264715. [Online]. Available: https://doi.org/10.1109/TIT.2013.2264715.

[45] S. Kumar, A. J. Young, N. Macris, and H. D. Pfister, "Threshold saturation for spatially coupled ldpc and ldgm codes on bms channels," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7389–7415, 2014.

[46] D. G. Mitchell, M. Lentmaier, and D. J. Costello, "Spatially coupled ldpc codes constructed from protographs," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4866–4889, 2015.

[47] D. Achlioptas, S. H. Hassani, N. Macris, and R. Urbanke, "Bounds for random constraint satisfaction problems via spatial coupling," in *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, SIAM, 2016, pp. 469–479.

[48] S Hamed Hassani, N. Macris, and R. Urbanke, "Threshold saturation in spatially coupled constraint satisfaction problems," *Journal of Statistical Physics*, vol. 150, no. 5, pp. 807–850, 2013.

[49] R. Ahlswede and D. E. Daykin, "An inequality for the weights of two families of sets, their unions and intersections," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 43, no. 3, pp. 183–185, 1978. DOI: 10.1007/BF00536201. [Online]. Available: https://doi.org/10.1007/BF00536201.

[50] M. J. Wainwright, M. I. Jordan, *et al.*, "Graphical models, exponential families, and variational inference," *Foundations and Trends® in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.

[51] N. Ruozzi, "The bethe partition function of log-supermodular graphical models," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., vol. 25, Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/03afdbd66e7929b125f8597834fa83a4-Paper.pdf.

[52] Y. Rinott and M. Saks, "Correlation inequalities and a conjecture for permanents," *Combinatorica*, vol. 13, no. 3, pp. 269–277, 1993. DOI: 10.1007/BF01202353. [Online]. Available: https://doi.org/10.1007/BF01202353.

[53] R. Aharoni and U. Keich, "A generalization of the Ahlswede-Daykin inequality," *Discrete Mathematics*, vol. 152, no. 1, pp. 1–12, 1996, ISSN: 0012-365X. DOI: https://doi.org/10.1016/0012-365X(94)00294-S. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0012365X9400294S.

[54] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre, "Correlation inequalities on some partially ordered sets," *Communications in Mathematical Physics*, vol. 22, no. 2, pp. 89–103, 1971. DOI: 10.1007/BF01651330. [Online]. Available: https://doi.org/10.1007/BF01651330.

[55] R. Holley, "Remarks on the fkg inequalities," *Communications in Mathematical Physics*, vol. 36, no. 3, pp. 227–231, 1974. DOI: 10.1007/BF01645980. [Online]. Available: https://doi.org/10.1007/BF01645980.

[56] P. C. Fishburn, "A correlational inequality for linear extensions of a poset," *Order*, vol. 1, no. 2, pp. 127–137, 1984. DOI: 10.1007/BF00565648. [Online]. Available: https://doi.org/10.1007/BF00565648.

[57] L. A. Shepp, "The XYZ Conjecture and the FKG Inequality," *The Annals of Probability*, vol. 10, no. 3, pp. 824 –827, 1982. DOI: 10.1214/aop/1176993791. [Online]. Available: https://doi.org/10.1214/aop/1176993791.

[58] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*, Second. Springer, 2011, vol. 143. DOI: 10.1007/978-0-387-68276-1.

[59] D. Straszak and N. K. Vishnoi, "Belief propagation, bethe approximation and polynomials," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4353–4363, 2019. DOI: 10.1109/TIT.2019.2901854.

[60] M. Jerrum, A. Sinclair, and E. Vigoda, "A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries," *J. ACM*, vol. 51, no. 4, 671–697, Jul. 2004, ISSN: 0004-5411. DOI: 10.1145/1008731.1008738. [Online]. Available: https://doi.org/10.1145/1008731.1008738.

[61] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities. 2nd ed.* English, Cambridge, Engl.: At the University Press. XII, 324 p. (1952). 1952.

[62] N. Ruozzi, "Beyond log-supermodularity: Lower bounds and the bethe partition function," in *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence, UAI 2013, Bellevue, WA, USA, August 11-15, 2013*, A. E. Nicholson and P. Smyth, Eds., AUAI Press, 2013. [Online]. Available: https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1\&smnu=2\&article_id=2416\&proceeding_id=29.

[63] H. A. Bethe, "Statistical Theory of Superlattices," *Proceedings of the Royal Society of London Series A*, vol. 150, no. 871, pp. 552–575, Jul. 1935. DOI: 10.1098/rspa.1935.0122.

[64] A. Sinclair and P. Srivastava, "Lee-yang theorems and the complexity of computing averages," in *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, ser. STOC '13, Palo Alto, California, USA: Association for Computing Machinery, 2013, 625–634, ISBN: 9781450320290. DOI: 10.1145/2488608.2488686. [Online]. Available: https://doi.org/10.1145/2488608.2488686.

[65] S. J. Axler, *Linear Algebra Done Right* (Undergraduate Texts in Mathematics). New York: Springer, 1997, ISBN: 0387982582. [Online]. Available: http://linear.axler.net/.

[66] R. L. Dobrushin, "Estimates of semi-invariants for the Ising model at low temperatures," *Translations of the American Mathematical Society-Series 2*, vol. 177, pp. 59–82, 1996.

[67] R. L. Dobrushin, "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. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 1–66, ISBN: 978-3-540-49635-9. DOI: 10.1007/BFb0095674. [Online]. Available: https://doi.org/10.1007/BFb0095674.

[68] P. Erdős, "Some remarks on the theory of graphs," *Bulletin of the American Mathematical Society*, vol. 53, no. 4, pp. 292 –294, 1947. DOI: bams/1183510596. [Online]. Available: https://doi.org/.

[69] N. J. A. Harvey and J. Vondrák, *Short proofs for generalizations of the Lovász local lemma: Shearer's condition and cluster expansion*, 2017. arXiv: 1711.06797 [math.CO].

[70] W. Pegden, "The lefthanded local lemma characterizes chordal dependency graphs," *Random Struct. Algorithms*, vol. 41, no. 4, pp. 546–556, 2012. DOI: 10.1002/rsa.20439. [Online]. Available: https://doi.org/10.1002/rsa.20439.

[71] P. Erdős and L. Lovász, "Problems and results on 3-chromatic hypergraphs and some related questions," *Coll Math Soc J Bolyai*, vol. 10, Jan. 1974.

[72] J. Spencer, "Asymptotic lower bounds for ramsey functions," *Discrete Mathematics*, vol. 20, pp. 69–76, 1977, ISSN: 0012-365X. DOI: https://doi.org/10.1016/0012-365X(77)90044-9. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0012365X77900449.