Witness Trees in the Moser–Tardos Algorithmic Lovász Local Lemma and Penrose Trees in the Hard-Core Lattice Gas

Rogério Gomes Alves · Aldo Procacci

Received: 15 March 2014 / Accepted: 10 June 2014 / Published online: 2 July 2014
© Springer Science+Business Media New York 2014

Abstract We point out a close connection between the Moser–Tardos algorithmic version of the Lovász local lemma, a central tool in probabilistic combinatorics, and the cluster expansion of the hard-core lattice gas in statistical mechanics. We show that the notion of witness trees given by Moser and Tardos is essentially coincident with that of Penrose trees in the Cluster expansion scheme of the hard-core gas. Such an identification implies that the Moser–Tardos algorithm is successful in a polynomial time if the cluster expansion converges.

Keywords Cluster expansion · Hard core lattice gas · Algorithmic Lovász local lemma

1 Introduction, State of Art, Notations and Results

1.1 The Lovász Local Lemma

The Lovász local lemma (LLL), originally formulated by Erdös and Lovász in [8], is one of the most important tools in the framework of the so called probabilistic methods in combinatorics. The philosophy of the probabilistic method (see e.g. [2] and reference therein) is to prove the existence of combinatorial objects with certain desirable properties (e.g. a proper coloring of the edges of a graph) by showing that these objects have a positive probability to occur in some suitably defined probability space. The LLL, in particular, allows to prove the existence of objects with properties occurring with exponentially small probability and it is widely used in graph theory (e.g. to obtain bounds on several problems about graph colorings, Ramsey numbers, van der Waerden function) and computer science (e.g. satisfiability problems, k-sat, set covering, latin transversal).
Given a finite set $X$, let $A = \{ A_x \}_{x \in X}$ be a collection of events (the bad events) in some probability space, each event $A_x$ with probability $\text{Prob}(A_x)$ to occur. Let $\overline{A}_x$ be the complement event of $A_x$ so that $\bigcap_{x \in X} \overline{A}_x$ is the event that none of the events $\{ A_x \}_{x \in X}$ occurs. Let $G = (X, E)$ be a graph with vertex-set $X$ and edge-set $E$ and, for each $x \in X$, let $\Gamma_G(x)$ denote the vertices of $G$ adjacent to $x$ and let $\Gamma^+_G(x) = \Gamma_G(x) \cup \{ x \}$. The graph $G$ is said to be a dependency graph for the collection of events $A$ if for each $x \in X$, $A_x$ is independent of all the events in the $\sigma$-algebra generated by $\{ A_y : y \in X \setminus \Gamma^+_G(x) \}$. The LLL gives a sufficient criterion to guarantee that the event $\bigcap_{x \in X} \overline{A}_x$ (i.e. the good event, the one with the desirable properties whose existence we would like to prove) has a strictly positive probability to occur (and hence it is non empty). In its more general form (the so-called non-symmetric version) the LLL can be stated as follows.

**Theorem 1** (Lovász Local Lemma) Let $G$ be a dependence graph for the collection of events $\{ A_x \}_{x \in X}$ with probability $\text{Prob}(A_x) = p_x$ and let $\mu = \{ \mu_x \}_{x \in X}$ be a sequence of real numbers in $[0, +\infty)$. If, for each $x \in X$,

$$p_x \leq \frac{\mu_x}{\prod_{y \in \Gamma^+_G(x)} (1 + \mu_y)}$$

then

$$\text{Prob}\left( \bigcap_{x \in X} \overline{A}_x \right) > 0$$

Shearer [27] gave an alternative formulation of this lemma which has been used as a bridge by Scott and Sokal [24,25] to point out a surprising and very interesting connection with the cluster expansion of the hard-core lattice gas on $G$ (the dependency graph). We remind rapidly below the hard-core gas setting and its state of the art.

1.2 The Self-Repulsive Hard Core Gas on a Graph $G = (X, E)$

The hard-core gas on a graph $G$ with vertex-set $X$ and edge-set $E$ is defined as follows. Suppose that each vertex $x \in X$ can be occupied by a ‘particle’ (also called sometimes, depending on the context, a ‘polymer’) or can be left empty. Moreover each particle occupying the vertex $x \in X$ carries an “activity” $w_x \in \mathbb{C}$ and we denote by $w = \{ w_x \}_{x \in X}$ the set of all activities. We further suppose that particles of this gas on $G$ interact through a self-repulsive hard-core nearest neighbor pair potential. Namely, each vertex can be occupied at most by one particle, and if a particle occupies the vertex $x \in X$, then all neighbor vertices of $x$ in $G$ must be empty. In the statistical mechanics lingo, if $x$, $y$ are vertices of the graph $G = (X, E)$ where the hard-core gas is defined such that either $\{ x, y \} \in E$ or $x = y$, it is usual to say that $x$ and $y$ are incompatible and write $x \not\sim y$ (and compatible otherwise, i.e. if $\{ x, y \} \not\in E$ and $x \neq y$, writing $x \sim y$). The grand-canonical partition function of this gas in the “volume” $X$ is then defined as

$$\Xi_X(w) = \sum_{\substack{Y \subset X \\text{independent}}} \prod_{y \in Y} w_y$$

(1.1)

where the sum in the r.h.s. is over the independent subsets of the vertex-set $X$ of $G$ (a subset $Y \subset X$ is independent in $G$ if no edge of $G$ has both endpoints in $Y$) so that $\Xi_X(z)$ coincides with the independent set (multivariable) polynomial on $G$. The “pressure” of this gas is defined via the formula (hereafter, whenever $X$ denotes a set, $|X|$ denotes its cardinality)
\[ P(w) = \frac{1}{|X|} \log \Xi_X(w) \]  

Near another key function is

\[ \Pi_{x_0}(w) = \frac{\partial}{\partial w_{x_0}} \log \Xi_X(w) = \frac{\Xi_X \Gamma_{x_0}(w)}{\Xi_X(w)} \]

The quantity \( w, \Pi_{x_0}(w) \) can be interpreted from the physical point of view (at least for positive activities \( w \geq 0 \)) as the one-point correlation function of the hard-core gas (i.e. the probability to see a particle sitting in the site \( x_0 \) regardless of where the other particles are).

It is a well-known fact that \( \log \Xi_X(w) \) (and hence \( P(w) \) and \( \Pi_{x_0}(w) \)) can be written in term of a formal series, known as the cluster expansion (CE) of the hard-core gas. Indeed, let \( G_n \) denote the set of all connected graphs with vertex-set \( I_n = \{1, 2, \ldots, n\} \) and, given an \( n \)-tuple \( (x_1, \ldots, x_n) \in X^n \), let \( g(x_1, \ldots, x_n) \) be the graph with vertex-set \( I_n \) which has the edge \( \{i, j\} \) if and only if \( x_i \not\sim x_j \) (i.e. if \( \{x_i, x_j\} \in E \) or \( x_i = x_j \)). Define, for \( n \geq 2 \)

\[ \phi^T(x_1, \ldots, x_n) = \begin{cases} 
\sum_{g \in G_n} \frac{(-1)^{|E_g|}}{n!} & \text{if } g(x_1, \ldots, x_n) \in G_n \\
0 & \text{if } g(x_1, \ldots, x_n) \not\in G_n
\end{cases} \]

Then, one can write formally (see e.g. [6, 10, 21, 22, 28])

\[ \log \Xi_X(w) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{(x_1, \ldots, x_n) \in X^n} \phi^T(x_1, \ldots, x_n) w_{x_1}, \ldots, w_{x_n} \]

whence

\[ \Pi_{x_0}(w) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(x_1, \ldots, x_n) \in X^n} \phi^T(x_0, x_1, \ldots, x_n) w_{x_1}, \ldots, w_{x_n} \]

The Eqs. (1.5) and (1.6) make sense only for those \( w \in C^{\{X\}} \) such that the formal series in the r.h.s. of (1.5) and (1.6) converge absolutely. It is again a well-known fact (see, e.g. [23] and also Proposition 1 ahead) that the number \( \phi^T(x_1, \ldots, x_n) \) defined in (1.4) has the following property

\[ \phi^T(x_1, \ldots, x_n) = (-1)^{n-1} \phi^T(x_1, \ldots, x_n) \]

We can thus consider, for \( \rho = \{\rho_x\}_{x \in X} \) with \( \rho_x \in (0, \infty) \) for all \( x \in X \), the positive term series

\[ \Pi_{x_0}(-\rho) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(x_1, x_2, \ldots, x_n) \in X^n} |\phi^T(x_0, x_1, \ldots, x_n)| \rho_{x_1}, \ldots, \rho_{x_n} \]

and, if we are able to show that \( \Pi_{x_0}(-\rho) \) converges for some (bounded) positive value \( \rho \in [0, \infty)^{|X|} \), then also \( \Pi_{x_0}(w) \) converges absolutely, whenever \( w \in \{w_x\}_{x \in X} \) is in the poly-disk \( |w| \leq \rho \) and in this poly-disk the pressure (1.2) admits the bound uniform in \( X \)

\[ |P(w)| \leq \sup_{x_0 \in X} \rho_{x_0} \Pi_{x_0}(-\rho) \]

Throughout this paper, operations and relations involving boldface symbols should be understood componentwise, for instance \( |w| \leq \rho \) is shorthand for \( \{|w_x| \leq \rho_x\}_{x \in X} \) and \(-\rho \) means \(-\rho_x\) for all \( x \in X \), and so on. The set
\[ \mathcal{R}(G) = \{ \rho \in [0, \infty)^{|X|} : \Pi_{x_0}(-\rho) < +\infty \} \]

constitutes, in the statistical mechanics lingo, the *convergence region* of the cluster expansion. Observe that, by definition, \( \mathcal{R}(G) \) is a down-set, i.e. \( \rho \in \mathcal{R}(G) \) and \( \rho^\prime \leq \rho \) implies \( \rho^\prime \in \mathcal{R}(G) \).

Lots of effort have been spent during the past three decades to establish efficient upper bounds for \( \mathcal{R}(G) \) (see e.g. \[ 6, 11, 15, 22, 26 \] and references therein). These efforts can be resumed by the so called Dobrushin criterion \[ 7 \], which can be stated as follows.

**Theorem 2** (Dobrushin) *Let \( G = (X, E) \) be a graph and let \( \mathcal{R}(G) \) be the convergence region of the cluster expansion of the hard-core gas on \( G \). Let \( \mu = (\mu_x)_{x \in X} \) be a family of non negative numbers in \([0, +\infty)\). If \( \rho = (\rho_x)_{x \in X} \) is such that, for all \( x \in X \)

\[ \rho_x \leq \frac{\mu_x}{\prod_{y \in \Gamma^*_G(x)}(1 + \mu_y)} \]

then \( \rho \in \mathcal{R}(G) \) and

\[ \rho_x \Pi_x(-\rho) \leq \mu_x \]

In 2007 however the Dobrushin criterion has been improved by Fernández and Procacci \[ 10 \].

**Theorem 3** (Fernández–Procacci) *Under the same hypothesis of Theorem 2, if \( \rho = (\rho_x)_{x \in X} \) such that, for all \( x \in X \)

\[ \rho_x \leq \sum_{R \subseteq \Gamma^*_G(x) \atop R \text{ indep in } G} \frac{\mu_x}{\prod_{y \in R} \mu_y} \]

Then \( \rho \in \mathcal{R}(G) \) and

\[ \rho_x \Pi_x(-\rho) \leq \mu_x \]

The proof of Theorem 3 is based on a identity discovered by Penrose \[ 20 \] who was able to rewrite the (alternating) sum over connected graphs of the Eq. (1.4) in terms of a sum over special trees, known as Penrose trees (see Sect. 2.2, Definition 4 below). The improvement with respect to Theorem 2 is immediately recognized by noting that

\[ \prod_{y \in \Gamma^*_G(x)}(1 + \mu_y) = \sum_{R \subseteq \Gamma^*_G(x) \atop R \text{ indep in } G} \prod_{y \in R} \mu_y \geq \sum_{R \subseteq \Gamma^*_G(x) \atop R \text{ indep in } G} \prod_{y \in R} \mu_y \]

### 1.3 Connection Between the LLL and the Hard-Core Gas

As anticipated above, in 2005 Scott and Sokal \[ 24, 25 \] elucidated a surprising and beautiful connection between the repulsive hard-core gas in statistical mechanics and the LLL in probabilistic combinatorics. In particular, they pointed out that the Shearer’s formulation \[ 27 \] for the applicability of the LLL was equivalent to require the convergence of the cluster expansion of the hard-core lattice gas. As an immediate consequence, they showed that the LLL condition (Theorem 1) could be seen as a reformulation of the Dobrushin criterion (Theorem 2) for the convergence of the cluster expansion. Scott and Sokal reformulated the Shearer’s version of the LLL in terms of convergence of the cluster expansion of the hard-core gas as follows.
Theorem 4 (Scott–Sokal) Let \( G \) be a dependence graph for the family of events \( \{ A_x \}_{x \in X} \) with probability \( \text{Prob}(A_x) = p_x \). Let \( \Xi_X(w) \) be the partition function of the hard-core gas on \( G \) and let \( \mathcal{R}(G) \) the convergence region of the cluster expansion of the hard-core gas on \( G \). If \( p = \{ p_x \}_{x \in X} \in \mathcal{R}(G) \), then,

\[
\text{Prob}\left( \bigcap_{x \in X} \overline{A_x} \right) \geq \frac{\Xi_X(-p)}{\Xi_X} > 0 .
\]

Furthermore these bounds are the best possible, i.e. if \( p \notin \mathcal{R}(G) \), then there can be constructed a family of events \( \{ B_x \}_{x \in X} \) in a suitable probability space with probabilities \( \text{Prob}(B_x) = p_x \) and dependency graph \( G \), such that \( \mathbb{P}(\bigcap_{x \in X} \overline{B_x}) = 0 \).

Remark By merging Theorem 2 into Theorem 4 one obtains immediately the usual LLL, i.e. Theorem 1. On the other hand, by merging Theorem 3 into Theorem 4 we have immediately the following improved version of the LLL recently given by Bissacot et al. [5].

Theorem 5 (Bissacot–Fernández–Procacci–Scoppola) Under the same hypothesis of Theorem 1, if, for each \( x \in X \),

\[
p_x \leq \frac{\sum_{R \subseteq \Gamma_G(x)} \mu_x \prod_{x' \in R} \mu_{x'}}{\prod_{x \in R} \mu_x}
\]

Then

\[
\text{Prob}\left( \bigcap_{x \in X} \overline{A_x} \right) > 0
\]

Theorem 5 has been already used to obtain improved bounds on various graph coloring problems (see [18] and [4]).

1.4 The Algorithmic Moser–Tardos Version of the Lovász Local Lemma

The undeniable popularity of the LLL has always had to live with a recurring criticism about its inherently non-constructive character. Namely, the LLL, giving sufficient conditions for the probability that none of the undesirable events occur to be strictly positive, implies that there exists at last one configuration in the probability space of the events which realizes the occurrence of the “good” event \( \bigcap_{x \in X} \overline{A_x} \), but it does not provide any algorithm capable to find, possibly in a polynomial time, such a configuration. Efforts to devise an algorithmic version of the LLL go back to the work of Beck [3] and Alon [1], and, after various contributions (see e.g. [16] and references therein), finally culminate in a recent breakthrough paper by Moser and Tardos [17], who gave a fully algorithmic version of LLL if the events are restricted to a class which however covers basically all known applications of LLL. The Moser–Tardos scheme is as follows. Let \( \mathcal{V} \) be a finite family of mutually independent random variables and let \( \Omega \) the probability space determined by these variables. Let \( X \) be a finite set and let \( A = \{ A_x \}_{x \in X} \) be a finite family of events, each \( A_x \) depending on some subset of the random variables of the family \( \mathcal{V} \), each with probability \( \text{Prob}(A_x) = p_x \). Denote \( vbl(A_x) \), for all \( A_x \in A \), the minimal (with respect to inclusion) and unique subset of \( \mathcal{V} \) that determines \( A_x \). The dependence graph of the family \( A \) is the graph \( G = (X, E) \) with vertex-set \( X \) and edge-set \( E \) constituted by the pairs \( \{ x, x' \} \subset X \) such that \( vbl(A_x) \cap vbl(A_{x'}) \neq \emptyset \). Observe that if \( x, y \in X \) and \( vbl(A_x) \cap vbl(A_y) \neq \emptyset \) then either \( \{ x, y \} \in E \) or \( x = y \).
By analogy with the hard-core gas we denote this with the symbol \( x \not\sim x' \) and say that \( x, y \) are incompatible or overlap (so \( x, y \) compatible, denoted with \( x \sim y \) means \( \{x, y\} \notin E \), i.e. \( vbl(A_x) \cap vbl(A_{x'}) = \emptyset \)). Within this scheme Moser and Tardos defined the following algorithm.

From now on we suppose that the finite set \( X \) which indexes the family of events \( A \) (with dependency graph \( G = (X, E) \)) is ordered and indicate with \( < \) such an order.

**MT-Algorithm** As initial step choose a random evaluation of the variables \( v \in V \) (in other words choose at random a point in \( \Omega \)). If some \( A \in A \) occurs, then pick one of them (at random or according to some deterministic rule), say \( A_x \) and take a new evaluation (resampling) only of its variables, keeping unchanged all the other variables in \( V \). The algorithm stops when we reach an evaluation of the variables \( v \in V \) such that none of the events in the family \( A \) occurs. The first step of the algorithm is the initial sampling of all variables in \( V \) (the step 0 by convention) and for \( i \in \mathbb{N} \), the step \( i \) of the algorithm is the selection (according to some deterministic or random rule based on the order established in \( X \)) of an occurring bad event \( A_x \in A \) and the resampling of its variables \( vbl(A_x) \).

**Theorem 6** (Moser and Tardos) Let \( V \) be a finite set of mutually independent random variables. Let \( A = \{A_x\}_{x \in X} \) be a finite set of events determined by these variables, each with probability \( \text{Prob}(A_x) = p_x \) and with dependency graph \( G \) and let \( \mu = (\mu_x)_{x \in X} \) be a a sequence of real numbers in \( [0, +\infty) \). If, for each \( x \in X \),

\[
p_x \leq \frac{\mu_x}{\prod_{y \in \Gamma_G(x)} (1 + \mu_y)} \tag{1.11}
\]

then there exists an assignment of values to the variables \( V \) such that none of the events in \( A \) occurs. Moreover the MT-algorithm described above resamples an event \( A_x \in A \), at most an expected \( \mu_x \) times before it finds such an evaluation. Thus the expected total number of resampling steps is at most \( \sum_{x \in X} \mu_x \).

Following Moser and Tardos, as the algorithm runs, resampling at each step some bad event from the family \( A \), one can define the log of the algorithm \( C = \{C(1), C(2), \ldots \} \) with \( C(i) \in X \). Namely, \( C \) lists the events as they are selected and resampled by the algorithm at each step, so that, for \( i \in \mathbb{N} \), if \( C(i) = x \) then the event \( A_x \in A \) is picked and resampled at step \( i \) of the algorithm. Note that if the algorithm stops then \( C \) is partial, i.e. there exists an \( n \in \mathbb{N} \) such that \( C : [1, n] \rightarrow X \). Using the words of Moser and Tardos, \( C \) is a random variable determined by the random choices made by the algorithm at each step.

**Rooted trees, dressed trees and witness trees.** Moser and Tardos’ proof of Theorem 6 is based on the notion of ‘witness trees’. To explain these objects we need to introduce some notations and definitions about graphs and trees.

We first recall the standard definition of graph with no loops or multiple edges as well as the notion of unlabeled graph. If \( X \) is a finite set, we denote by \( P_2(X) \) the set of all subsets of \( X \) with cardinality 2. I.e. \( P_2(X) = \{e \subset V : |e| = 2\} \).

**Definition 1** An undirected simple graph \( G \) is a pair \( (V, E) \), where \( V \) is a non-empty set and \( E \subset P_2(V) \). \( V \) is called the vertex-set of \( G \) and \( E \) is called the edge-set of \( G \). A graph \( G = (V, E) \) is said to be connected if for any pair \( B, C \) of subsets of \( V \) such that \( B \cup C = V \) and \( B \cap C = \emptyset \), there is an edge \( e \in E \) such that \( e \cap B \neq \emptyset \) and \( e \cap C \neq \emptyset \). A tree graph (or shortly tree) is a connected graph \( \tau = (V, E) \) such that \( |E| = |V| - 1 \).

**Definition 2** Two graphs \( G = (V, E) \) and \( G' = (V', E') \) are said to be isomorphic if there is a bijection \( \gamma : V \rightarrow V' \) such that \( (x, y) \in E \Rightarrow (\gamma(x), \gamma(y)) \in E' \). The isomorphism...
An undirected unlabeled simple connected graph with no cycles and such that one vertex has been designated the root is called a rooted tree. Hereafter we will use the letter \( t \) to denote a generic unlabeled rooted tree and we denote by \( \Theta \) the set of all possible unlabeled rooted trees. The vertices of a rooted tree \( t \in \Theta \) admit a natural partial order (called the tree order). Namely, given two (distinct) vertices \( u \in v \) in a rooted tree, \( v \) is said to be a descendant of \( u \) or (\( u \) is an ancestor of \( v \)), if there is a path from the root to \( v \) which contains \( u \). If \( \{ v, u \} \) is an edge of a rooted tree, then either \( v \) is a descendant of \( u \) or vice versa. So actually any edge \( \{ u, v \} \) in a rooted tree is directed (i.e. is an ordered pair) and we write \( (u, v) \) (\( v \) descendant of \( u \)) with \( u \) being the parent (or predecessor, or father) and \( v \) being the child (or successor). Note that each vertex in \( t \) different form the root has one and only one parent. The root has no predecessor and it is the extremum with respect to tree order relation.

Given a set \( X \), a \textit{\( X \)-dressed tree} is a pair \( t = (t, \sigma) \) where \( t \in \Theta \) is a rooted tree with vertex-set \( V_t \) and \( \sigma \) is function \( \sigma : V_t \rightarrow X \). Note that with this definition a dressed tree with labels in \( X \) may have distinct vertices associated to the same label in \( X \). Therefore one can bound probabilities \( \text{Prob}(\tau) \) that a witness tree \( \tau = (t, \sigma) \) with root labeled \( x \) and labels \( \{\sigma(v)\}_{v \in V_t} \) at the vertices appears in the log \( C \) of the algorithm is at most

\[
\text{Prob}(\tau) \leq \prod_{v \in V_t} \text{Prob}(A_{\sigma(v)}) = \prod_{v \in V_t} p_{\sigma(v)} \tag{1.12}
\]

Now, for \( x \in X \) let \( N_x \) be the random variable that counts how many times the event \( A_x \) is resampled during the execution of the MT-algorithm. Then \( N_x \) is, by definition, the number of occurrences of the event \( A_x \) in the log \( C \) of the algorithm and also the number of distinct proper witness trees occurring in \( C \) that have their root labeled \( x \). Therefore one can bound the expectation of \( N_x \) simply by summing the bounds (1.12) on the probabilities \( \text{Prob}(\tau) \) as \( \tau \) varies in the set \( T^x \) of the distinct witness trees with root labeled \( x \). Thus the expected value \( E(N_x) \) of \( N_x \) is bounded as

\[
E(N_x) \leq \Phi_x(p) \tag{1.13}
\]

where

\[
\Phi_x(p) = \sum_{(t, \sigma) \in T^x} \prod_{v \in V_t} p_{\sigma(v)} \tag{1.14}
\]

Moser and Tardos conclude their proof by showing, via a Galton–Watson branching process argument, that the quantity \( \Phi_x(p) \) defined in (1.14) is bounded by \( \mu_x \) if probabilities \( \{p_x\}_{x \in X} \) are such that conditions (1.11) are verified.
1.5 Results

After the work of Scott and Sokal, relating the non-constructive LLL to the statistical mechanics of hard-core gas and the improvement of the lemma obtained by Bissacot et al. exploiting this connection, it is a natural question to ask whether there can be made a similar connection between the algorithmic Lovász local lemma (ALLL) proposed by Moser and Tardos and the hard-core gas. We stress that this question is far from being trivial, since the scheme proposed by Moser and Tardos to prove their Theorem 6, based on the concept of witness trees, has, at first sight, nothing to do with the various proofs of the non-algorithmic LLL proposed in the literature.

Strong indications that a connection between the ALLL and the hard-core gas must indeed exist come form two recent works [14] and [20]. In [14] Kolipaka and Szegedy relate the Moser–Tardos algorithm to the set of Shearer’s conditions via an auxiliary algorithm (called by the authors “generalized resample”) and via a reformulation of the Moser–Tardos scheme in which the notion of witness trees is replaced by two alternative concepts (called by the authors “stable set sequences” and “stable set matrices”). However, in [14] no explicit improvement on Theorem 6 (eventually based on the equivalence of Shearer’s conditions and convergence of the cluster expansion) is presented. The improvement was later found by Pegden [20] using a completely different method. Namely, Pegden realized that, within the Moser–Tardos scheme involving witness trees, it was possible to modify the branching process argument given in [17] in order to adapt it to the Bissacot et al. condition [5] of Theorem 5.

In this paper we show that the connection between ALLL and CE of the hard-core gas is astonishingly direct, much more direct, dare we say, than the one pointed out by Scott and Sokal for the non-constructive LLL. Indeed, the connection can be obtained bypassing completely Shearer’s formulation and remaining within the original Moser–Tardos scheme involving witness trees (as the work of Pegden was implicitly suggesting). Namely, by a slight modification of the map which defines the Penrose trees in CE, we are able to show that the notion of witness tree defined in [17] is in fact coincident with that of the Penrose tree in the CE scheme of the hard-core gas. Such an identification implies that the sum over witness trees given in [17], which bounds from above the expected number of times an event \(A_x\) is resampled, happens to be exactly equal to the cluster expansion of the one point correlation function defined in (1.3) calculated at \(w = -p\) (we recall that \(p = \{p_x\}\) with \(p_x \in [0, 1]\) being the probability \(P(A_x)\) of occurrence of the event \(A_x\)). The main result of the paper can be summarized by the following theorem.

**Theorem 7** Under the same hypothesis of Theorem 6, let \(\Xi_X(w)\) be the partition function of the hard-core lattice gas on \(G\) with complex activities \(w = \{w_x\}_{x \in X}\) and let \(\mathcal{R}(G)\) the convergence region of the cluster expansion of the hard-core gas on \(G\).

If \(p = \{p_x\}_{x \in X} \in \mathcal{R}(G)\), then there exists an assignment of values to the variables \(V\) such that none of the events in \(A\) occurs. Moreover the MT-algorithm described above finds such an evaluation resampling an event \(A_x \in A\) in an expected time \(T_x\) such that

\[
T_x \leq p_x \left[ \frac{\partial \log \Xi_X(w)}{\partial w_x} \right]_{w=-p} \tag{1.15}
\]

and the expected total number of resamplings \(T\) of the variables in \(\mathcal{P}\) is at most

\[
T \leq \sum_{x \in X} p_x \left[ \frac{\partial \log \Xi_X(w)}{\partial w_x} \right]_{w=-p} \leq |X||P(w = -p)|
\]

where \(P(w)\) is the pressure of the hard-core lattice gas on \(G\) with activities \(w = \{w_x\}_{x \in X}\).
Remark 1 Theorem 7 above together with Theorem 3 immediately yields for free the following corollary, which is the result obtained by Pegden.

**Corollary 1** Under the hypothesis of Theorem 7, if \( \mu = \{\mu_x\}_{x \in X} \) is a sequence of real numbers in \([0, +\infty)\) such that, for each \( x \in X \)

\[
p_x \leq \frac{\sum \mu_x}{\prod_{y \in R} \mu_y}
\]

then the randomized algorithm resamples an event \( A_x \in A \), at most an expected \( \mu_x \) times before it finds such an evaluation. Thus the expected total number of resampling steps is at most \( \sum_{x \in X} \mu_x \).

Remark 2 If \( p \) is outside the convergence radius of the cluster expansion, then one can say nothing about the efficiency of the algorithm since the series bounding the expected time the algorithm stops diverges. Of course in the algorithmic Moser–Tardos setting, i.e. the collection of bad events \( A = \{A_x\}_{x \in X} \) depending on a finite number of independent random variables \( \nu \in \mathcal{V} \), it is possible to think about different algorithms which could be more efficient and stop even for a set of probabilities \( p \) for which the Moser–Tardos algorithm doesn’t stop. Along these directions we would like to cite some interesting results obtained in [12] and [9].

The next two sections are devoted to the proof of Theorem 7. Specifically, in Sect. 2 we define the modified Penrose map for the CE of the hard-core gas on a graph \( G \) and write the series for the pressure and the derivative of the log of the partition function in terms of a sum over Penrose trees. In Sect. 3 we show that the witness trees of the Moser–Tardos scheme coincide with these modified Penrose trees and conclude the proof of Theorem 7.

### 2 Cluster Expansion of the Hard-Core Gas on a Graph. A Variant of the Penrose Map

In this section we will reorganize the series \( \Pi_{\lambda_0}(-\rho) \) of Eq. (1.8) in terms of a sum over certain trees using the so-called Penrose identity. To this purpose, we need to recall some definitions. In particular, a very special role in order to state the Penrose identity is played by the labeled rooted trees and plane rooted trees.

#### 2.1 Labeled Trees and Plane Rooted Tress

We will use the following notations. Given a vertex \( v \neq 0 \) in an (unlabeled) rooted tree \( t \), its depth, denoted by \( d(v) \), is the number of edges in the unique path from the root to that vertex. Given a vertex \( v \) in a rooted tree, we denote by \( v^* \) is parent (unless \( v \) is the root, since in that case \( v \) has no parent) and we denote by \( s_v \) the number of its children. Note that \( d(v^*) = d(v) - 1 \). Children of the same parent are also called siblings. Given a vertex \( v \) in a rooted tree, any vertex \( w \) such that \( d(w) = d(v) \) but \( w \) is not a sibling of \( v \) is called a cousin of \( v \) and any vertex \( w \neq v^* \) such that \( d(w) = d(v) - 1 \) is called an uncle of \( v \).

**Plane rooted trees** A plane tree is a rooted tree \( t \) for which an ordering is given for the children of each vertex. An ordering of the children in a rooted tree \( t \) is equivalent to a drawing of \( t \) in the plane, obtained, e.g., by putting parents at the left of their children which are ordered in the top-to-bottom order. Note that the number of plane rooted trees with \( n \) vertices is always greater that the number of unlabeled rooted trees with \( n \) vertices. E.g. there are 4 different
rooted trees with 4 vertices while there are 5 different plane rooted trees with 4 vertices. We denote by $T_n^0$ the set of all plane rooted trees with $n + 1$ vertices.

The set of vertices $V_t$ in a plane rooted tree $t \in T_n^0$ admits a natural total order $\prec$, which we call the plane-tree order. I.e., given two (distinct) vertices $u, v$ of $t$, we have $v \prec u$, and say that $v$ is older than $u$ or $u$ is younger than $v$, if either $d(v) < d(u)$, or $d(v) = d(u)$ but $v$ is above $u$ in the drawing of $t$.

**Labeled rooted trees** Let $I_n^0 = \{0, 1, 2, \ldots, n\}$. A rooted tree $t$ with vertex-set $I_n^0$ and root 0 is usually called a labeled rooted tree. In other words, according to the notations adopted in Sect. 1.4, a labeled rooted tree is a $I_n^0$-dressed rooted tree $\tau = (t, \sigma)$ where $t$ is a rooted tree with vertex-set $V_t$ and root $r$ and $\sigma : V_t \to I_n^0$ is a bijection (therefore $|V_t| = n + 1$) such that $\sigma(r) = 0$. Note that the number of labeled rooted trees with $n$ vertices is always greater that the number of plane rooted trees with $n$ vertices. E.g. there are 16 different labeled rooted trees with 4 vertices while there are 5 different plane rooted trees with 4 vertices. We will use the letter $\vartheta$ to denote a generic labeled tree and we denote by $T_n^0$ the set of all labeled trees with vertex-set $I_n^0$ which are rooted in 0.

There is a natural map $m : T_n^0 \to T_n^0$ which associates to each labeled rooted tree $\vartheta \in T_n^0$ a unique plane rooted tree $m(\vartheta) \in T_n^0$. This unique plane rooted tree $m(\vartheta)$ is obtained by fixing the order of the children in each vertex of $\vartheta$ according with the order of their labels in $I_n^0$. For example the plane rooted trees associated to the trees $\vartheta_1$ with edge-set $\{0, 3\}, \{1, 3\}, \{2, 3\}, \{2, 4\}, \vartheta_2$ with edge-set $\{0, 2\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \vartheta_3$ with edge-set $\{0, 2\}, \{0, 3\}, \{1, 3\}, \{3, 4\}$ and $\vartheta_4$ with edge-set $\{0, 2\}, \{0, 4\}, \{2, 3\}, \{1, 2\}$ are drawn below.

![Diagram](image-url)

Observe that $\vartheta_1$ and $\vartheta_2$, which are different labeled trees, are sent by the map $m$ into the same plane rooted tree, i.e. $m(\vartheta_1) = m(\vartheta_2)$. On the other hand $m(\vartheta_3)$ and $m(\vartheta_4)$ are different plane rooted trees (even though they correspond to the same unlabeled rooted tree).

Clearly the map $\vartheta \mapsto m(\vartheta) = t$ is many-to-one and the cardinality of the preimage $m^{-1}(t)$ of a plane rooted tree $t$ is equal to the number of ways of labeling the $n$ non-root vertices of $t$ with $n$ distinct labels from $\{1, 2, \ldots, n\}$ consistently with order of the children in each vertex, i.e.,

$$|\{\vartheta \in T_n^0 : m(\vartheta) = t\}| = \frac{n!}{\prod_{v \in V_t} s_v}$$

(2.1)

where recall that if $v \in V_t$, then $s_v$ denotes the number of the children of $v$.

There is also a natural map $\theta : T_n^0 \to T_n^0$ (an injection) which assigns to the vertices of a plane rooted tree $t$ labels in the set $I_n^0$ accordingly to the plane tree order of vertices in $t$. In other words, the root has label 0, the $s_0$ children of the root have labels 1, 2, ..., $s_0$ from top to bottom, the higher root child vertex, i.e. that with label 1, has $s_1$ children with labels $s_0 + 1 \ldots s_0 + s_1$, the root child vertex with label 1 has $s_i$ children with labels $s_0 + s_1 + \ldots s_{i-1} + 1, \ldots, s_0 + s_1 + \ldots s_{i-1} + s_i$, and so on. We call this labeling of $t$ the *natural labeling* of a plane rooted tree $t$. Given plane rooted tree $t \in T_n^0$ we will denote by $\vartheta_t$ the unique labeled tree in $T_n^0$ whose labels coincides in all vertices with the natural labels of $t$, i.e. $\vartheta_t \equiv \theta(t)$.
2.2 The Penrose Map

A pair \((\vartheta; (x_0, x_1, \ldots, x_n))\), where \(\vartheta = (t, \tilde{\sigma})\) is a labeled rooted tree (with \(t\) rooted tree and \(\tilde{\sigma} : V_t \to I_n^0\) bijection) and \((x_0, x_1, \ldots, x_n) \in X^{n+1}\) is an ordered \(n + 1\)-tuple, uniquely determines a \(X\)-dressed tree \(\tau = (t, \sigma)\) where \(t\) is a (unlabeled) rooted tree and \(\sigma = \tilde{\sigma} \circ \vartheta\) with \(\tilde{\sigma} : I_n^0 \to X\) such that \(\tilde{\sigma}(u) = x_i\). Note that \(\tilde{\sigma}\) is in general not injective, so two distinct vertices in \(\vartheta\) may carry the same label in the set \(X\). However labels of distinct vertices in the pair \((\vartheta; (x_0, x_1, \ldots, x_n))\) are always to be regarded as distinct (since distinct vertices have always at least distinct labels in \(I_n^0\)). We further suppose that vertices of the pair \((\vartheta; (x_0, x_1, \ldots, x_n))\) are ordered according the order chosen in \(X\). I.e. we order the set \(\{x_i\}_{i \in I_n^0}\) if either \(x_i < x_j\) (in the total order established in \(X\)) or \(x_i = x_j\) and \(i < j\). We call such pair \((\vartheta; (x_0, x_1, \ldots, x_n))\), with vertices equipped with the order above, a labeled \(X\)-dressed rooted tree.

We finally define a map \(\mu\) which sends each labeled \(X\)-dressed rooted tree \((\vartheta; (x_0, x_1, \ldots, x_n))\) in a unique plane rooted tree \(t = \mu(\vartheta; (x_0, x_1, \ldots, x_n)) \in \mathbb{T}_n^0\). This map \(\mu\) acts similarly as the map \(m\) previously introduced but it uses the order in the set \(\{x_i\}_{i \in I_n^0}\) established above instead of usual the order of the set \(I_n^0\). Namely, this unique plane rooted tree \(t\) associated to \((\vartheta; (x_0, x_1, \ldots, x_n))\) is obtained by fixing the order of the children in each vertex of \(\vartheta\) according with the order of their labels in the set \(\{x_i\}_{i \in I_n^0}\) ordered as described above. This map \(\mu\) automatically induces a total order, still denoted by \(<\), on the vertices of a labeled \(X\)-dressed rooted tree \((\vartheta; (x_0, x_1, \ldots, x_n))\). Indeed, given any two vertices \(u, v\) in the \(X\)-dressed rooted tree \((\vartheta; (x_0, x_1, \ldots, x_n))\) with labels \(x_i, x_j\) respectively, we say that \(u < v\) if the corresponding vertices in the plane rooted tree \(t = \mu(\vartheta; (x_0, x_1, \ldots, x_n))\), say \(\mu(u)\) and \(\mu(v)\), are such that \(\mu(u) < \mu(v)\) and we will write shortly \(x_i < x_j\).

Let’s now go back to the graph \(G = (X, E)\) on which the hard-core lattice gas has been defined. We recall that if \(x, y \in X\) are such that either \([x, y] \notin E\) or \(x = y\), we denote this with the symbol \(x \not\sim y\) (\(x\) and \(y\) are incompatible or \(x\) and \(y\) overlap), and if \([x, y] \notin E\), we denote simply \(x \sim y\) (\(x\) and \(y\) compatible). We also recall that, for fixed \((x_0, x_1, \ldots, x_n) \in X^{n+1}\), \(g(x_0, x_1, \ldots, x_n)\) is the graph with vertex-set \(I_n^0\) and edge-set \(E_{g(x_0, x_1, \ldots, x_n)} = \{(i, j) \in I_n^0 : x_i \sim x_j\}\).

**Definition 4** A labeled \(X\)-dressed rooted tree \((\vartheta; (x_0, x_1, \ldots, x_n))\) is called a Penrose tree if the following holds.

\((t0)\) \(E_{\vartheta} \subset E_{g(x_0, x_1, \ldots, x_n)}\) (i.e. if \([i, j] \notin E_{\vartheta}\) then \(x_i \sim x_j\))

\((t1)\) if two vertices \(i\) and \(j\) are such that \(d(i) = d(j)\), then \([i, j] \notin E_{g(x_0, x_1, \ldots, x_n)}\) (i.e. \(x_i \sim x_j\));

\((t2)\) if two vertices \(i\) and \(j\) are such that \(d(j) = d(i) - 1\) and \(x_{i^*} < x_j\), then \([i, j] \notin E_{g(x_0, x_1, \ldots, x_n)}\) (i.e. \(x_i \sim x_j\)).

We denote by \(P(x_0, x_1, \ldots, x_n)\) the subset of \(T_n^0\) constituted by those \(\vartheta \in T_n^0\) such that the pair \((\vartheta; (x_0, x_1, \ldots, x_n))\) is Penrose.

**Remark** Property \((t0)\) says that (labels of) children always overlap (labels of) their parents, property \((t1)\) says that siblings and/or cousins do not overlap. Finally property \((t2)\) says that children are always compatible with their uncles which are below the father in the drawing of the plane tree \(\mu(\vartheta; (x_0, x_1, \ldots, x_n))\). We want to emphasize that the map presented above is slightly different with respect to the original map given by Penrose in [19] (used also in [10, 13, 28]). However this slight alteration still allows to construct a partition scheme for the set \(G_{g(x_0, x_1, \ldots, x_n)}\) as in the original Penrose paper (see Proposition 1 below). On the other
hand, the present three conditions for a pair \((\vartheta, (x_0, x_1, \ldots, x_n))\) to be a Penrose tree have the advantage to be independent on labels in \(I^n_0\). They depend on the the underlying plane rooted tree \(t = \mu(\vartheta; (x_0, x_1, \ldots, x_n))\) associated to \(\vartheta\) and on the order established in the set \(X\) (since condition t1 guarantees that siblings always receive distinct labels in \(X\), so that they can be ordered using only the order in \(X\) which thus is sufficient to establish if \(x_{i^*} < x_j\) to check condition t2). In the original Penrose paper [19] and in works [10,13,28] the usual order in \(I^n_0\) is used which instead depends on the labels of \(\vartheta\). This subtle difference will be crucial in order to rewrite the series for (1.8) for \(\Pi_{x_0}(-\rho)\) in terms of plane rooted trees.

**Proposition 1**

\[
\phi^T(x_0, x_1, \ldots, x_n) = (-1)^n \sum_{\vartheta \in P(x_0, x_1, \ldots, x_n)} \mathbb{1}_\vartheta \in T^n_0(\vartheta) \tag{2.2}
\]

where \(\mathbb{1}_\vartheta \in P(x_0, x_1, \ldots, x_n)\) is the characteristic function of the set \(P(x_0, x_1, \ldots, x_n)\) in \(T^n_0\), i.e.

\[
\mathbb{1}_\vartheta \in P(x_0, x_1, \ldots, x_n)(\vartheta) = \begin{cases} 
1 & \text{if } \vartheta \in P(x_0, x_1, \ldots, x_n) \\
0 & \text{otherwise}
\end{cases}
\]

**Proof** Fix \((x_0, x_1, \ldots, x_n) \in X^{n+1}\). We recall that \(g(x_0, x_1, \ldots, x_n)\) denoted the graph with vertex-set \(I^n_0\) and edge-set \(E_g(x_0, x_1, \ldots, x_n) = \{[i, j] \subset I^n_0 : x_i \sim x_j\}\). Without loss of generality we may assume that \(g(x_0, x_1, \ldots, x_n)\) is connected (otherwise \(\phi^T(x_0, x_1, \ldots, x_n) = 0\) and (2.2) is trivial). We denote by \(G^n_0\) the set of all connected graphs with vertex-set \(I^n_0\) and we put

\[G_g(x_0, x_1, \ldots, x_n) = \{g \in G^n_0 : g \subset g(x_0, x_1, \ldots, x_n)\}\]

and

\[T_g(x_0, x_1, \ldots, x_n) = \{\vartheta \in T^n_0 : \vartheta \subset g(x_0, x_1, \ldots, x_n)\}\]

Note that the vertices of \(g \in G_g(x_0, x_1, \ldots, x_n)\), which can be identified as numbers in \(I^n_0\), also carry labels in \(X\), i.e. for \(i \in I^n_0\), the vertex \(i\) carries the label \(x_i \in X\). We also recall that the set \([x_i]_{i \in I^n_0}\) is ordered according the order in \(X\) if \(x_i \neq x_j\) and on the order in \(I^n_0\) if \(x_i = x_j\). We also consider the graph \(g\) as always rooted in 0, so for any \(j\) vertex of \(g\), we will denote by \(d_g(j)\) its distance from the root 0 in \(g\). Recall that \(E_g\) denotes the edge-set of \(g\).

We now define the map \(q : G_g(x_0, x_1, \ldots, x_n) \rightarrow T_g(x_0, x_1, \ldots, x_n)\) that associate to \(g \in G_g(x_0, x_1, \ldots, x_n)\) a unique labeled rooted tree \(q(g) \in T_g(x_0, x_1, \ldots, x_n)\) as follows.

**The map** \(q : G_g(x_0, x_1, \ldots, x_n) \rightarrow T_g(x_0, x_1, \ldots, x_n)\)

1. We first delete all edges \([i, j]\) in \(E_g\) with \(d_g(i) = d_g(j)\). After this operation we are left with a new connected graph \(g'\) such that \(d_{g'}(i) = d_g(i)\) for all vertices \(i = 0, 1, \ldots, n\). Moreover each edge \([i, j]\) of \(g'\) is such that \(|d_{g'}(i) - d_{g'}(j)| = 1\).
2. Let \(i_1, \ldots, i_{s_0}\) be the vertices at distance 1 from the root 0 in \(g'\) ordered according their labels in \(X\), i.e. in such way that \(i_1 < i_2 < \cdots < i_{s_0}\) if and only if \(x_{i_1} < x_{i_2} < \cdots < x_{i_{s_0}}\) according the order established in \([x_i]_{i \in I^n_0}\). Now take \(i_1\) (the smallest one) and let \(j_1^{i_1}, \ldots, j_{s_1}^{i_1}\) be the vertices connected to \(i_1\) by edges of \(E_{g'}\) (these vertices are at distance 2 from the root 0 and again are ordered according the order of their labels in \([x_i]_{i \in I^n_0}\) and delete all edges of \(g'\) connecting vertices \(j_1^{i_1}, \ldots, j_{s_1}^{i_1}\) to vertices in the set \(\{i_2, \ldots, i_{s_0}\}\). The graph so obtained \(g'_1\) is such that any of the vertices \(j_1^{i_1}, \ldots, j_{s_1}^{i_1}\) is connected only to \(i_1\) and to vertices at distance greater than 2. Then take the vertex \(i_2\) (the smallest after \(i_1\))
and let \( j_1^{i_2}, \ldots, j_{i_1}^{i_2} \) be the vertices connected to \( i_2 \) at distance 2 from the root 0 in \( g'_1 \) and delete all edges of \( g'_1 \) connecting vertices \( j_1^{i_2}, \ldots, j_{i_1}^{i_2} \) to vertices in the set \( \{i_3, \ldots, i_{i_0}\} \).

The graph so obtained \( g'_2 \) is such that any of the vertices \( j_1^{i_1}, \ldots, j_{i_1}^{i_1} \) is connected only to \( i_2 \) and vertices at distance greater than 2. After \( s_0 \) steps we are left with a graph \( g'_s \) with no loops among vertices at distance \( d \leq 2 \) from the root. Continue now this procedure until all vertices of \( g \) are exhausted, always respecting the order of the labels. Namely, take \( j_1^{i_1} \) (i.e. the one with the smallest label among \( j_1^{i_1}, \ldots, j_{i_1}^{i_1} \)) and consider the vertices at distance 3 emanating from \( j_1^{i_1} \) and delete all edges linking these vertices to some vertex in the set \( \{ j_2^{i_1}, j_3^{i_1}, \ldots, j_{i_1}^{i_1}, j_1^{i_2}, \ldots, j_{i_1}^{i_2}\} \) and continue this procedure until all vertices are exhausted. The resulting graph \( g'' = q(g) \) is by construction a spanning connected subgraph of \( g(x_0, x_1, \ldots, x_n) \), i.e. \( q(g) \in G_{g(x_0, x_1, \ldots, x_n)} \), and which has no cycles, i.e. \( q(g) \in T_{g(x_0, x_1, \ldots, x_n)} \). Observe that the map \( q \) is a surjection from \( G_{g(x_0, x_1, \ldots, x_n)} \) to \( T_{g(x_0, x_1, \ldots, x_n)} \). Indeed, for any \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \) we have that \( q(\vartheta) = \vartheta \) and hence \( q^{-1}(\vartheta) \) is always non empty.

The map \( p : T_{g(x_0, x_1, \ldots, x_n)} \to G_{g(x_0, x_1, \ldots, x_n)} \)

Conversely, let \( p \) be the map that to each tree \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \) associates the graph \( p(\vartheta) \in G_{g(x_0, x_1, \ldots, x_n)} \) formed by adding to \( \vartheta \) all edges \( [i, j] \in E_{g(x_0, x_1, \ldots, x_n)} \setminus E_{\vartheta} \) such that either \( d_{\vartheta}(i) = d_{\vartheta}(j) \), or \( d_{\vartheta}(j) = d_{\vartheta}(i) - 1 \) and \( x_i < x_j \). Then by construction, for any \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \) we have that \( p(\vartheta) \in G_{g(x_0, x_1, \ldots, x_n)} \) and \( q(p(\vartheta)) = \vartheta \). Last but not least, recalling Definition 4, observe that if \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \) then \( p(\vartheta) = \vartheta \iff \vartheta \in P(x_0, x_1, \ldots, x_n) \).

Observe now that the set \( G_{g(x_0, x_1, \ldots, x_n)} \) is partially ordered by edge inclusion, namely, \( g, g' \in G_{g(x_0, x_1, \ldots, x_n)} \) and \( E_g \subseteq E_{g'} \), then \( g < g' \). Moreover if \( g, g' \in G_{g(x_0, x_1, \ldots, x_n)} \) and \( g < g' \) we denote by \( [g, g'] \) the subset of \( G_{g(x_0, x_1, \ldots, x_n)} \) formed by those \( \hat{g} \) such that \( g < \hat{g} < g' \). With these definitions it follows that if \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \) and \( g \in [\vartheta, p(\vartheta)] \), then, by construction of the map \( q \), we have that \( q(g) = \vartheta \). In other words the set of graphs \( g \in G_{g(x_0, x_1, \ldots, x_n)} \) such that \( q(g) = \vartheta \) is the Boolean interval \([\vartheta, p(\vartheta)]\) with \( \vartheta \) being minimum of the interval and \( p(\vartheta) \) being the maximal graph, with respect to the partial order relation \( < \) in \( G_{g(x_0, x_1, \ldots, x_n)} \). So \( G_{g(x_0, x_1, \ldots, x_n)} \) is partitioned into the disjoint union of the sets \([\vartheta, p(\vartheta)]\) with \( \vartheta \in T_{g(x_0, x_1, \ldots, x_n)} \). This shows that the map \( p \) provides a so-called partition scheme of the family of graphs \( G_{g(x_0, x_1, \ldots, x_n)} \).

With these definitions, recalling that \( p(\vartheta) = \vartheta \iff \vartheta \in P(x_0, x_1, \ldots, x_n) \), we have

\[
\sum_{g \in G_{g(x_0, x_1, \ldots, x_n)}} (-1)^{|E_g|} = \sum_{\vartheta \in T_{g(x_0, x_1, \ldots, x_n)}} (-1)^{|E_{\vartheta}|} \sum_{g \in G_{g(x_0, x_1, \ldots, x_n)}} (-1)^{|E_g| - |E_{\vartheta}|} \\
= (-1)^n \sum_{\vartheta \in T_{g(x_0, x_1, \ldots, x_n)}} [1 + (-1)^{|E_{p(\vartheta)}| - |E_{\vartheta}|}] \\
= (-1)^n \sum_{\vartheta \in T_{g(x_0, x_1, \ldots, x_n)}} \|_{\vartheta \in P(x_0, x_1, \ldots, x_n)} \\
= (-1)^n \sum_{\vartheta \in T_{g(x_0, x_1, \ldots, x_n)}} \|_{\vartheta \in P(x_0, x_1, \ldots, x_n)}
\]

and the proposition is proved.
Using Proposition 1 we can rewrite the formal series (1.8) as

$$\Pi_{X_0}(-\rho) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\vartheta \in T_0^n} \sum_{(x_1, \ldots, x_n) \in X^n} \mathbb{I}_{\vartheta \in P(x_0, x_1, \ldots, x_n)} \rho_{x_1} \cdots \rho_{x_n}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\vartheta \in T_0^n} \phi_{X_0}(\vartheta, \rho)$$

(2.3)

where

$$\phi_{X_0}(\vartheta, \rho) = \sum_{(x_1, \ldots, x_n) \in X^n} \mathbb{I}_{\vartheta \in P(x_0, x_1, \ldots, x_n)} \rho_{x_1} \cdots \rho_{x_n}$$

(2.4)

This equation shows that the formal series $\Pi_{X_0}(-\rho)$ can be reorganized as a sum over terms associated to labeled rooted trees. We go a little bit further, getting rid of labels in $T_0^n$, and reorganize the series in term of a sum over plane rooted trees. Indeed, the factor $\phi_{X_0}(\vartheta, \rho)$ defined in (2.4) does not depend on the labels of $\vartheta \in T_0^n$, the condition $\vartheta \in P(x_0, x_1, \ldots, x_n)$ only depends on variable $x_1, \ldots, x_n$ which are dummy variables. So, if $t = m(\vartheta)$ is the unique plane rooted associated to $\vartheta$, we can write, for any given $\vartheta \in T_0^n$ such that $m(\vartheta) = t \in T_0^n$

$$\phi_{X_0}(\vartheta, \rho) = \phi_{X_0}(\vartheta_t, \rho)$$

(2.5)

where, recall that $\vartheta_t$ denotes the labeled tree associated to the plane rooted tree $t$ (i.e. $t$ plus the natural labeling of its vertices ordered according the plane tree order). Therefore

$$\phi_{X_0}(\vartheta, \rho) = \sum_{(x_1, \ldots, x_n) \in X^n} \mathbb{I}_{\vartheta_t \in P(x_0, x_1, \ldots, x_n)} \rho_{x_1} \cdots \rho_{x_n}$$

(2.6)

and hence

$$\Pi_{X_0}(-\rho) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\vartheta_t \in T_0^n} \mathbb{I}_{\vartheta_t \in P(x_0, x_1, \ldots, x_n)} \rho_{x_1} \cdots \rho_{x_n}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{t \in T_0^n} \phi_{X_0}(\vartheta_t, \rho) \sum_{\vartheta \in T_0^n} \mathbb{I}_{\vartheta \in P(x_0, x_1, \ldots, x_n)} m(\vartheta) = t$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{t \in T_0^n} \phi_{X_0}(\vartheta_t, \rho) |m^{-1}(t)|$$

$$= \sum_{n=0}^{\infty} \sum_{t \in T_0^n} \prod_{v \in V_t} \left[ \frac{1}{S_{v_t}} \right] \sum_{(x_1, \ldots, x_n) \in X^n} \mathbb{I}_{\vartheta_t \in P(x_0, x_1, \ldots, x_n)} \prod_{i=1}^{n} \rho_{x_i}$$

As last step we can also get rid of the factorials in formula above. Indeed, given $t \in T_0^n$, let us say that an $n$-uple $(x_1, \ldots, x_n)$ is well-ordered in $t$ if labels of children (which are distinct elements of $X$) of any given vertex of $t$ are ordered according the order on $X$. Then we can write

$$\sum_{t \in T_0^n} \prod_{v \in V_t} \left[ \frac{1}{S_{v_t}} \right] \sum_{(x_1, \ldots, x_n) \in X^n} \mathbb{I}_{\vartheta_t \in P(x_0, x_1, \ldots, x_n)} \prod_{i=1}^{n} \rho_{x_i} = \sum_{t \in T_0^n} \sum_{(x_1, \ldots, x_n) \in X^n \text{ well-ordered in } t} \prod_{i=1}^{n} \rho_{x_i}$$

(2.7)

I.e. in conclusion we have obtained

$$\Pi_{X_0}(-\rho) = \sum_{n=0}^{\infty} \sum_{t \in T_0^n} \mathbb{I}_{\vartheta_t \in P(x_0, x_1, \ldots, x_n)} \prod_{i=1}^{n} \rho_{x_i}$$
3 Witness Trees are Penrose Trees

Let us now go back to the Moser–Tardos scheme illustrated in Sect. 1.4. We will make use of the concept of plane rooted tree previously introduced to redefine the witness trees in a completely deterministic way. We recall that set X which indexes the family of events A (with dependency graph $G = (X, E)$) is ordered. Let now $t \in \mathbb{T}_n^0$ be a plane rooted tree with vertex-set $V_t$ and edge-set $E_t$ and let $\sigma : V_t \rightarrow X$ be a function. We say that $\sigma$ is a good labeling of $t$ if it is such that $\{v, v'\} \in E_t \iff \sigma(v) \neq \sigma(v')$ and moreover if $v$ and $w$ are siblings and $v < w$ (according to the plane tree order) then $\sigma(v) < \sigma(w)$ (in the order introduced in $X$). Note that if an $X$-dressed tree $(t, \sigma)$ is such that $t \in \mathbb{T}_n^0$ and $\sigma$ is a good labeling, then the pair $(\vartheta_t, (x_0, x_1, \ldots, x_n))$ is such that the $n$-uple $(x_1, \ldots, x_n)$ is well-ordered in $t$.

**Proposition 2** Let $X$ be the vertex-set of a graph $G$. An $X$-dressed tree $\tau = (t, \sigma)$ is a proper witness tree if and only if is a finite plane rooted tree $t = (V_t, E_t) \in \mathbb{T}_0^0$ with a good labeling $\sigma : V_t \rightarrow X$.

**Proof** Let us first show that a finite plane rooted tree $t = (V_t, E_t) \in \mathbb{T}_0^0$ with a good labeling $\sigma : V_t \rightarrow X$ is a proper witness tree according to the Moser and Tardos Definition 3 given in Sect. 1.4. Indeed it is obvious that $\{v, v'\} \in E_t \iff \sigma(v) \neq \sigma(v')$ is the same as requiring that the children of a vertex $u \in V_t$ receive labels from $\Gamma^*_G(\sigma(u))$. Moreover, since labels of siblings must respect their order in the plane tree, these labels must be necessarily distinct. It is finally simple to construct a one-to-one correspondence between rooted trees $t$ whose vertices are labeled by a function $\sigma : V_t \rightarrow X$ according to the rule that children always overlap their parents and always receive distinct labels and plane rooted trees $t$ whose vertices are well-labeled with labels from $X$. Indeed, any plane rooted tree $t$ equipped with a good-labeling with labels from $X$ can also be viewed as a rooted tree $t$ whose vertices are labeled with labels from $X$. Conversely, to any rooted tree $t$ whose vertices are labeled with labels from $X$ according to the rule that the labels of the children always overlap the labels of their parents and always receive distinct labels we can associate a unique plane rooted tree $t$ equipped with a good labeling with labels from $X$: just order the children of the rooted tree $t$ according to the order of their labels in $X$ obtaining in this way a (unique) plane rooted tree $t$ whose vertices are automatically equipped with a good labeling with labels from $X$. \square

**Definition 5** A proper witness tree $\tau = (t, \sigma)$ is called a Penrose tree if the following occurs:

1. if two vertices $v$ and $v'$ are such that $d(v) = d(v')$, then $\sigma(v) \sim \sigma(v')$;
2. if two vertices $v$ and $v'$ are such that $d(v') = d(v) - 1$ and $v^* < v'$ (i.e. $v'$ is an uncle of $v$ which is below the father $v^*$ of $v$), then $\sigma(v) \sim \sigma(v')$

We denote by $S^+_X$ the set of all Penrose trees $\tau = (t, \sigma)$ with root label $x$.

**Remark** Note that this definition coincides, *mutatis mutandis*, with Definition 4 given in Sect. 2 in the sense that we can construct a one-to-one correspondence between Penrose trees $\tau = (t, \sigma)$ according to Definition 5 and Penrose tree $(\vartheta_t, (x_0, x_1, \ldots, x_n))$ according to Definition 4. Indeed if $\tau = (t, \sigma)$ is a Penrose tree according to Definition 5 then $t$, being a plane rooted tree, defines uniquely the labeled rooted tree $\vartheta_t \in \mathbb{T}_n^0$ previously seen. Moreover the function $\sigma$ defines uniquely a $n + 1$-tuple $(x_0, x_1, \ldots, x_n)$ such that $\sigma(i) = x_i$ for each $i \in I_n^0$ (we are identifying vertices of $V_t$ with numbers in $I_n^0$ through the bijection $t \mapsto \vartheta_t$). Then, since $\sigma$ is a good labeling, we have that $\mu(\vartheta_t; (x_0, x_1, \ldots, x_n)) = t$ so that the pair $(\vartheta_t; (x_0, x_1, \ldots, x_n))$ automatically fulfills condition 2 of Definition 4 and
thus \( \vartheta_t \in P(x_0, x_1, \ldots, x_n) \) according to Definition 4. Conversely, if \((\vartheta; (x_0, x_1, \ldots, x_n)) \) is Penrose tree according to Definition 4, let \( t = \mu(\vartheta; (x_0, x_1, \ldots, x_n)) \) be the plane tree associated to \((\vartheta; (x_0, x_1, \ldots, x_n)) \) by the map \( \mu \) and let \( \pi \) be the permutation of \([1, \ldots, n] \) such that \( x_{\pi(1)} < x_{\pi(2)} < \cdots < x_{\pi(n)}. \) Again identifying vertices of \( V_t \) with numbers in \( T_n^0 \) through the bijection \( t \mapsto \vartheta_t, \) we define \( \sigma(0) = x_0 \) and \( \sigma(i) = x_{\pi(i)}. \) Then \((t, \sigma)\) is is a Penrose tree according to Definition 5 since labels (in the set \( X \)) of the children at each vertex respect the order of the plane tree \( t = \mu(\vartheta; (x_0, x_1, \ldots, x_n)) \in T_n^0 \) by definition of the map \( \mu. \)

Now we are in the position to explain how Moser and Tardos associate to each step \( s \) of the algorithm, with log C, a witness tree \( \tau(s) \in T_X^C(s) \) (with vertex labels chosen in the set \( X \) and root with label \( C(s) \)). The tree \( \tau(s) \) is obtained by constructing a sequence \( \tau_1(s), \tau_2(s), \ldots, \tau_1(s) \) of witness trees and then posing \( \tau(s) = \tau_1(s). \) Let \( \tau_1(s) \) be the witness tree formed only by a single vertex (i.e. the root) with label \( C(s) \). For \( i - 1 \in [1, \ldots, s - 1] \), \( \tau_{i-1}(s) \) is obtained from \( \tau_1(s) \) by attaching a new vertex to \( \tau_1(s) \) with label \( C(i - 1) \) in the following way. Let \( W_i \) be constituted by all vertices of \( \tau_{i-1}(s) \) whose labels (which, recall, are elements of \( X \)) are incompatible with the event \( C(i - 1). \) If \( W_i \) is empty (i.e. all vertices in \( \tau_1(s) \) have labels in \( X \) which are compatible with \( C(i - 1) \)) do nothing, i.e. put \( \tau_{i-1}(s) = \tau_1(s) \) and skip to the next step. If \( W_i \) is not empty, then, quoting Moser and Tardos, “choose among all such vertices the one (say \( v \)) having the maximum distance from the root and attach a new child vertex \( u \) to \( v \) with label \( C(i - 1) \), thereby obtaining the tree \( \tau_{i-1}(s)’ \).” Of course, even if not stated explicitly, if there is more than one vertex in \( W_i \) at maximum distance from the root, then one has to choose one, at random or according to some deterministic rule. Moser and Tardos does not give any deterministic rule to choose the vertex \( v \) when a choice is necessary. That is, if \( v \) is non unique, they choose it at random. We instead give a deterministic rule (that’s why we need to order the set \( X \) and to work with plane rooted trees in place of simple rooted trees). Let \( \tilde{W}_i \) be the subset of \( W_i \) formed by those vertices of \( W_i \) which are at the maximal distance from the root\(^1\) and attach the new vertex \( u \) with label \( C(i - 1) \) to the lowest (a.k.a. younger) vertex of the set \( \tilde{W}_i \), say \( v \), according to the tree order described before, forming in this way the tree \( \tau_{i-1}(s) \), in which \( u \) is a child of \( v \). Of course, in order to obtain a good labeling of \( \tau_{i-1}(s) \), if the vertex \( v \) had already children in \( \tau_1(s) \) (so that \( u \) becomes a new sibling of these children of \( v \)) attach the new vertex \( u \) with label \( C(i - 1) \) respecting the order of the children of \( v \).

According to Moser and Tardos we say that a witness tree \( \tau \) occurs in the log \( C \) of the algorithm if there exists \( s \in \mathbb{N} \) such that \( \tau(s) = \tau. \) We now prove a generalization of Lemma 2.1 in [17]. Recall that \( T_X^C \) denotes the set of all distinct witness tree which can be generated by the algorithm according to the procedure described above and with root labeled \( x \in X. \)

**Proposition 3** Let \( \tau = (t, \sigma) \) be a proper witness tree and let \( C \) be the (random) log produced by the algorithm. If \( \tau \) occurs in the log \( C \), then \( \tau \) is a Penrose tree.

**Proof** If \( \tau \) occurs in the log \( C \), then, by definition, there exists \( s \in \mathbb{N} \) such that \( \tau(s) = \tau = (t, \sigma). \) By construction the plane rooted tree \( t \) associated to \( \tau \) is such that in any given vertex \( v \in V_t \), the label \( \sigma(v) \) is compatible with all labels at the same distance from the root. Indeed, suppose (by contradiction) that \( v \) and \( v’ \) are two vertices of \( \tau(s) = \tau \) at the same distance from the root, i.e. \( d(v) = d(v’) \) and that the label of \( v \) is incompatible with the label of \( v’ \). Suppose, without loss of generality, that \( v’ \) has been attached after \( v. \) But then, since the

---

\(^1\) Be careful! Here ‘maximal’ means maximal in \( W_i \), so that \( \tau^{(i)}(s) \) can have vertices with depth greater than those in \( \tilde{W}_i \) as long as the labels in these vertices are all compatible with \( C(i - 1) \).
the hypothesis that \( v \) is an uncle of \( v \) who is younger than the father \( v^* \) of \( v \) (i.e. an uncle of \( v \) who is below the father \( v^* \) of \( v \) in the drawing of \( \tau \)). Then we have to show that \( \sigma(v) \sim \sigma(v') \). Indeed, suppose by absurd that \( v' \) is a younger uncle of \( v \) and that \( \sigma(v) \not\sim \sigma(v') \). We have to consider two cases. First we suppose that \( v \) has been added after \( v' \) to form \( \tau(s) = \tau \). Since \( \sigma(v) \not\sim \sigma(v') \) and \( v' \) is below \( v^* \), then, according to the deterministic rule described above, \( v \) cannot be attached to \( v^* \); it must be attached to \( v' \) or to another uncle below \( v' \), contrary to the hypothesis that \( v \) is attached to \( v^* \). Secondly, suppose that \( v' \) has been added after \( v \). But then \( d(v') \geq d(v) + 1 \), contrary to the hypothesis that \( v' \) is uncle of \( v \) (and hence \( d(v') = d(v) - 1 \)).

Recalling the definition of \( T^X_x \) given in Sect. 1.4 and the definition of the set \( S^X_x \), we can summarize Proposition 3 by saying that \( T^X_x \subseteq S^X_x \). Therefore recalling formulas (1.13) and (1.14) we have that the expected number of times an event \( A \in A \) is resampled by the MT-algorithm is bounded by

\[
E(N_x) \leq \Phi_x(p) \tag{3.1}
\]

where

\[
\Phi_x(p) = \sum_{\tau = (\ell, \sigma) \in T^X_x} \prod_{\tau \in S^X_x} p_{\sigma(v)} \tag{3.2}
\]

Now note that the sum on the r.h.s. of (3.2) can also be rewritten as follows. Let us denote by \( T^X_x \) the set of witness trees with \( n + 1 \) vertices (the root plus \( n \) vertices) with root \( r \) carrying the label \( x \in X \). Then

\[
\Phi_x(p) = p_x \sum_{n=0}^{\infty} \sum_{\tau \in T^X_x, n} \prod_{\tau \in S^X_x} p_{\sigma(v)} \tag{3.2}
\]

Now, recalling the remark after Definition 5, we observe the following.

1. To say that \( \tau \in T^X_x \), i.e. \( \tau \) is a pair \( \tau = (i, \sigma) \) where \( i \) is a plane rooted tree and \( \sigma \) is a good labeling of the vertices of \( i \), is equivalent to say (see remark after Definition 5) that the pair \( (\vartheta_i, (x_0, x_1, \ldots, x_n)) \) with \( \sigma(i) = x_i \) is such that the \( n \)-uple \( (x_1, \ldots, x_n) \) is well-ordered in \( i \).

2. To say that \( \tau \) is also such that \( \tau \in S^X_x \) (i.e. \( \tau \) is Penrose according to Definition 5) is equivalent to say that the pair \( (\vartheta_i, (x_0, x_1, \ldots, x_n)) \) is such \( \vartheta_i \in P(x_0, x_1, \ldots, x_n) \) (i.e. is Penrose according to Definition 4).

Therefore we can write

\[
\sum_{\tau \in T^X_x, n} \prod_{\tau \in S^X_x} p_{\sigma(v)} = \sum_{\tau \in T^X_x, n} \prod_{(x_1, \ldots, x_n) \in X^n} \prod_{i=1}^{n} p_{x_i}
\]

and so recalling (2.7), we get

\[
\sum_{n=0}^{\infty} \sum_{\tau \in T^X_x, n} \prod_{\tau \in S^X_x} p_{\sigma(v)} = \sum_{n=0}^{\infty} \sum_{\tau \in T^X_x, n} \prod_{\tau \in S^X_x} p_{\sigma(v)} = \prod_{x_0}(-p) \]

\( \square \)
Hence, recalling Definition (3.2) and bound (3.1), we conclude that

\[ E(N_x) \leq \tilde{\Phi}_x(p) = p_x \Pi_x(-p) \]

which concludes the proof of Theorem 7.

Acknowledgments Aldo Procacci has been partially supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do estado de Minas Gerais (FAPEMIG—Programa de Pesquisador Mineiro).

References

1. Alon, N.: A parallel algorithmic version of the local lemma. Random Struct. Algorithms 2(4), 367–378 (1991)
2. Alon, N., Spencer, J.: The Probabilistic Method, 3rd edn. Wiley-Interscience, New York (2008)
3. Beck, J.: An algorithmic approach to the Lovász local lemma. Random Struct. Algorithms 2(4), 343–365 (1991)
4. Böttcher, J., Kohayakawa, Y., Procacci, A.: Properly coloured copies and rainbow copies of large graphs with small maximum degree. Random Struct. Algorithms 40(4), 425–436 (2012)
5. Bissacot, R., Fernández, R., Procacci, A., Scoppola, B.: An improvement of the Lovász local lemma via cluster expansion. Comb. Probab. Comput. 20(5), 709–719 (2011)
6. Cammarota, C.: Decay of correlations for infinite range interactions in unbounded spin systems. Comm. Math. Phys. 85, 517–28 (1982)
7. Dobrushin, R.L.: Perturbation methods of the theory of Gibbsian fields. In: Bernard, P. (ed.) Lectures on Probability Theory and Statistics. Lecture Notes in Mathematics vol. 1648, pp. 1–66. Springer-Verlag, Berlin (1996)
8. Erdős, P., Lovász, L.: Problems and results on 3-chromatic hypergraphs and some related questions. In Infinite and Finite Sets, vol. 10, pp. 609–627. Colloq. Math. Soc. Janos Bolyai, Amsterdam (1975)
9. Esperet, L., Parreau, A.: Acyclic edge-coloring using entropy compression. Eur. J. Comb. 34(6), 1019–1027 (2013)
10. Fernández, R., Procacci, A.: Cluster expansion for abstract polymer models. New bounds from an old approach. Commun. Math. Phys. 274(1), 123–140 (2007)
11. Gruber, C.: General properties of polymer systems. Commun. Math. Phys. 22, 133–161 (1971)
12. Grytczuk, J., Kozik, K., Micek, P.: New approach to nonrepetitive sequences. Random Struct. Algorithms 42(2), 214–225 (2013)
13. Jackson, B., Procacci, A., Sokal, A.D.: Complex zero-free regions at large \(|q|\) for multivariate Tutte polynomials (alias Potts-model partition functions) with general complex edge weights. J. Comb. Theory, Series B 103, 21–45 (2013)
14. Kolipaka, K.B.R., Szegedy, M.: Moser and Tardos meet Lovász. In: Proceedings of the 43rd annual ACM symposium on Theory of computing, pp. 235–244. ACM New York, NY (2011)
15. Kotecký, R., Preiss, D.: Cluster expansion for abstract polymer models. Comm. Math. Phys. 103, 491–498 (1986)
16. Moser, R.A.: A constructive proof of the Lovász local lemma. In: Proceedings of the 41st Annual ACM Symposium on the Theory of Computing (STOC). ACM, New York (2009)
17. Moser, R., Tardos, G.: A constructive proof of the general Lovász local Lemma. JACM 57, 11 (2010)
18. Ndreca, S., Procacci, A., Scoppola, B.: Improved bounds on coloring of graphs. Eur. J. Comb. 33(4), 592–609 (2012)
19. Penrose, O.: Convergence of fugacity expansions for classical systems. In: Bak, A. (ed.) Statistical Mechanics: Foundations and Applications. Benjamin, New York (1967)
20. Pegden, W.: An extension of the Moser–Tardos algorithmic local lemma, SIAM J. Discrete Math. (2013). To appear http://arxiv.org/abs/1102.2853
21. Procacci, A., de Lima, B.N.B., Scoppola, B.: A remark on high temperature polymer expansion for lattice systems with infinite range pair interactions. Lett. Math. Phys. 45, 303–322 (1998)
22. Procacci, A., Scoppola, B.: Polymer gas approach to N-body lattice systems. J. Statist. Phys. 96, 49–68 (1999)
23. Ruelle, D.: Statistical Mechanics: Rigorous Results. W. A. Benjamin Inc, New York-Amsterdam (1969)
24. Scott, A., Sokal, A.D.: The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. J. Stat. Phys. 118(5–6), 1151–1261 (2005)
25. Scott, A., Sokal, A.D.: On dependency graphs and the lattice gas. Comb. Probab. Comput. 15, 253–279 (2006)
26. Seiler, E.: Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics. Lecture Notes in Physics, vol. 159. Springer-Verlag, Berlin (1982)
27. Shearer, J.B.: On a problem of spencer. Combinatorica 5, 241–245 (1985)
28. Sokal, A.D.: Bounds on the complex zeros of (Di)chromatic polynomials and Potts-model partition functions. Comb. Probab. Comput. 10(1), 41–77 (2001)