Comparison of two convergence criteria for the variable-assignment Lopsided Lovász Local Lemma

David G. Harris*

Submitted: Sep 26, 2020; Accepted: Aug 5, 2022; Published: Oct 21, 2022

Abstract

The Lopsided Lovász Local Lemma (LLLL) is a cornerstone probabilistic tool for showing that it is possible to avoid a collection of "bad" events as long as their probabilities and interdependencies are sufficiently small. The strongest possible criterion in these terms is due to Shearer (1985), although it is technically difficult to apply to constructions in combinatorics.

The original formulation of the LLLL was non-constructive; a seminal algorithm of Moser & Tardos (2010) gave an efficient algorithm for nearly all its applications, including to k-SAT instances where each variable appears in a bounded number of clauses. Harris (2015) later gave an alternate criterion for this algorithm to converge; unlike the LLL criterion or its variants, this criterion depends in a fundamental way on the decomposition of bad-events into variables.

In this note, we show that the criterion given by Harris can be stronger in some cases even than Shearer's criterion. We construct k-SAT formulas with bounded variable occurrence, and show that the criterion of Harris is satisfied while the criterion of Shearer is violated. In fact, there is an exponentially growing gap between the bounds provable from any form of the LLLL and from the bound shown by Harris.

Mathematics Subject Classifications: 60C05

1 Introduction

The Lovász Local Lemma (LLL) is a general probabilistic principle for showing that, in a probability space Ω with a finite set $\mathcal{B}$ of "bad" events which are not too interdependent and are not too likely, then there is a positive probability no events in $\mathcal{B}$ occur.

*Department of Computer Science, University of Maryland, College Park, MD 20742. Research supported in part by NSF Awards CNS 1010789 and CCF 1422569. Email: davidgharris29@gmail.com

Since its introduction in [3], it has become a cornerstone of the probabilistic method of combinatorics.

There have been numerous extensions of the LLL since its original formulation. One important generalization known as the *Lopsided Lovász Local Lemma* (LLLL) [4] observes that it is not necessary for bad-events to be fully independent. If the bad-events are *positively correlated* in a certain sense, then for the purposes of the LLL this is just as good as independence. This type of correlation, which we discuss shortly, is known as *lopsidependency.*

In order to explain the LLL formally, we need to introduce a number of definitions.

For any collection of events $S \subseteq \mathcal{B}$, we define $\overline{S} = \bigcap_{B \in S} \overline{B}$; we refer to this event as *avoiding* S. A *dependency graph* is a graph G on vertex set $\mathcal{B}$ such that for any $B \in \mathcal{B}$ and any set $S \subseteq \mathcal{B} - \{B\} - N_G(B)$ (where $N_G(B)$ denotes the neighborhood of B in G), we have

$$\Pr(B \mid \overline{S}) = \Pr(B). \tag{1}$$

That is, each bad-event $B \in \mathcal{B}$ is independent of all other events in $\mathcal{B}$, except possibly those which are neighbors of B in the dependency graph. A *lopsidependency graph* is a graph G on vertex set $\mathcal{B}$, satisfying the relaxed condition that for any $B \in \mathcal{B}$ and set $S \subseteq \mathcal{B} - \{B\} - N_G(B)$,

$$\Pr(B \mid \overline{S}) \leqslant \Pr(B). \tag{2}$$

A probability space Ω and collection of bad-events $\mathcal{B}$ does not have a unique dependency graph or lopsidependency graph. Rather, we suppose that we are given $\Omega, \mathcal{B}$ and some chosen graph G which is a (lopsi-)dependency graph for them.

For such a graph G with vertex set $V = \mathcal{B}$, we say a set $S \subseteq V$ is *stable* if no elements of S are adjacent in G. For real numbers p_v, indexed by the vertices $v \in V$, we define the *stable set polynomial* of G with respect to base set $S \subseteq V$, denoted $Q(G, S, \vec{p})$, by

$$Q(G, S, \vec{p}) = \sum_{\substack{\text{stable sets } T \\ S \subseteq T \subseteq V}} (-1)^{|T|-|S|} \prod_{v \in T} p_v$$

With these definitions, we state a few formulations of the LLLL.

Theorem 1. *Suppose G is a lopsidependency graph for Ω, B. If any of the following conditions are satisfied, then* $\Pr(\overline{\mathcal{B}}) > 0$.

1. *(Symmetric LLLL) If G has maximum degree d and every $B \in \mathcal{B}$ has* $\Pr(B) \leqslant p$ *and*
$$ep(d+1) \leqslant 1$$
2. *(Asymmetric LLLL) If there is a function $x : \mathcal{B} \to (0,1)$ satisfying*
$$\forall B \in \mathcal{B} \qquad \Pr(B) \leqslant x(B) \prod_{A \in N_G(B)} (1 - x(A))$$

3. *(Cluster-expansion criterion [2]) If there is a function $\mu : \mathcal{B} \to [0, \infty)$ satisfying*

$$\forall B \in \mathcal{B} \qquad \mu(B) \geqslant \Pr(B)\Big(\mu(B) + \sum_{\substack{Y \subseteq N_G(B) \\ Y \text{ stable}}} \prod_{A \in Y} \mu(A)\Big)$$

The symmetric LLLL uses only a few crude parameters of the problem instance — namely, the maximum probability of a bad-event and the maximum degree of the lopsidependency graph. The other variants use progressively more information and take advantage of refined dependency structure. See also [14] for another criterion in this vein. In [20], Shearer derived the most powerful possible criterion in these terms.

Theorem 2 (Shearer's criterion [20])**.** *Let G be a graph on vertex set $V = \{1, \dots, n\}$ and let $p_1, \dots, p_n \in [0, 1]$.*

1. *Suppose that $Q(G, \emptyset, \vec{p}) > 0$ and $Q(G, S, \vec{p}) \geqslant 0$ for all $S \subseteq V$. Then for any probability space Ω, and any events $B_1, \dots, B_n \subseteq \Omega$ in that space such that $\Pr(B_i) = p_i$ for $i = 1, \dots, n$ and such that G is a lopsidependency graph for $\mathcal{B} = \{B_1, \dots, B_n\}$, we have $\Pr(\overline{\mathcal{B}}) \geqslant Q(G, \emptyset, \vec{p}) > 0$.*

   In this case, we say that Shearer's criterion is satisfied *by $G, \vec{p}$.*

2. *Suppose that either $Q(G, \emptyset, \vec{p}) \leqslant 0$ or there is some stable set $S \subseteq V$ with $Q(G, S, \vec{p}) < 0$. Then there is some probability space Ω and events $B_1, \dots, B_n \subseteq \Omega$ such that $\Pr_\Omega(B_i) = p_i$ for $i = 1, \dots, n$ and such that G is a dependency graph for $\mathcal{B} = \{B_1, \dots, B_n\}$ and $\Pr(\overline{\mathcal{B}}) = 0$.*

   In this case, we say that Shearer's criterion is violated *by $G, \vec{p}$.*

Having bad-events with probability 0 or 1 is not so interesting, and Theorem 2 can be simplified when we disallow these cases.

Theorem 3 ([8], Lemma 5.27)**.** *Suppose that $p_1, \dots, p_n \in (0, 1)$. Shearer's criterion is satisfied by G, p if and only if $Q(G, S, \vec{p}) > 0$ for all stable sets S.*

Thus, Shearer's criterion exactly characterizes which probability and lopsidependency structure of the bad-events guarantees a positive probability of avoiding $\mathcal{B}$. From a theoretical point of view, alternate bounds such as Theorem 1 are all weaker than, and are implied by, Shearer's criterion. However, Shearer's criterion is technically difficult to apply to constructions in combinatorics.

1.1 The variable-assignment LLLL

The LLLL has been applied to diverse probability spaces such as random permutations [16], Hamiltonian cycles [1], and perfect matchings [17]. However, by far the most common form of the LLL and LLLL concerns what we refer to as the *variable-assignment* setting. Here, the probability space Ω has m independent discrete random variables $X_1, \dots, X_m$, and the bad-events can be taken to be "monomial events"; that is, each $B \in \mathcal{B}$ can be written in the form

$$(X_{i_1} = j_1) \wedge (X_{i_2} = j_2) \wedge \cdots \wedge (X_{i_k} = j_k)$$

For such a monomial event, we define $\text{var}(B) = \{i_1, \ldots, i_k\}$. We say that two events B, B' *disagree* on variable i if B demands $X_i = j$ and B' demands $X_i = j'$ for $j \neq j'$.

Definition 4. The *canonical dependency graph* G has an edge (B, B') iff $\text{var}(B) \cap \text{var}(B') \neq \emptyset$. The *canonical lopsidependency graph* G has edge (B, B') iff B disagrees with B' on any variable $i \in \text{var}(B) \cap \text{var}(B')$.

It is immediate that the canonical dependency graph is, indeed, a dependency graph for $\Omega, \mathcal{B}$. The fact that the canonical lopsidependency graph is a lopsidependency graph follows from the FKG inequality. Most applications of the LLL use only the canonical dependency graph; some noteworthy applications of the canonical lopsidependency graph include monochromatic hypergraph coloring [18] and boolean satisfiability [6]. We will discuss the latter in much more detail later.

In [13], Kolipaka & Szegedy noted that the Shearer criterion is not tight for the variable-assignment LLL setting. Namely, they found an explicit dependency graph and vector of probabilities where the Shearer criterion is violated yet any variable-assignment realization must have a satisfying assignment. Later work [11] provided a more systematic description of which dependency graphs were satisfiable in the variable-assignment setting.

1.2 The Moser-Tardos algorithm

The LLLL ensures that $\Pr(\overline{\mathcal{B}}) > 0$, and this is usually sufficient for combinatorics where the main goal is to show existence results. However, typically $\Pr(\overline{\mathcal{B}})$ is exponentially small, and hence the LLLL does not give efficient algorithms for *constructing* such a configuration. In [19], Moser & Tardos introduced a remarkably simple algorithm for the variable-assignment LLLL setting:

Algorithm 1 The Moser-Tardos (MT) algorithm

Draw each variable independently from the distribution Ω.
while there is a true bad-event on X **do**
    Choose a true bad-event B arbitrarily.
    Resample $\text{var}(B)$ according to the distribution Ω.

They showed that when the asymmetric LLLL criterion is satisfied with respect to the canonical lopsidependency graph, then this algorithm terminates in expected polynomial time with a configuration avoiding $\mathcal{B}$. Later work [13] showed that this algorithm terminates quickly whenever the Shearer criterion is satisfied. Thus, at least for the variable-assignment LLLL setting, this gives an efficient algorithm for nearly every construction based on the LLLL.

In [7], Harris gave a different type of criterion for the Moser-Tardos algorithm. Unlike the symmetric LLLL or other similar criteria, this cannot be stated solely in terms of the dependency graph and the probabilities of the bad-events. We summarize it here (in a slightly simplified form).

Definition 5 (Orderability)**.** Given $B \in \mathcal{B}$, we say that a set of bad-events $Y \subseteq \mathcal{B}$ is *orderable* to B, if there is an ordering $Y = \{B_1, \dots, B_s\}$, such that, for each $i = 1, \dots, s$, there is a variable $z_i \in \mathrm{var}(B) \cap \mathrm{var}(B_i)$ where B disagrees with B_i on z_i but B does not disagree with $B_1, \dots, B_{i-1}$ on z_i.

Theorem 6 ([7])**.** *Suppose there is $\mu : \mathcal{B} \to [0, \infty)$ satisfying the condition*

$$\forall B \in \mathcal{B} \qquad \mu(B) \geqslant \Pr(B)\Big(\mu(B) + \sum_{\substack{Y \text{ orderable} \\ \text{to } B}} \prod_{A \in Y} \mu(A)\Big)$$

Then the Moser-Tardos algorithm terminates with probability 1.

Theorem 6 is superficially similar to the cluster-expansion criterion. It is strictly stronger than the asymmeric LLLL and certain simplified forms of the cluster-expansion criterion. However, its relation to the Shearer criterion is not clear. It is quite plausible, along the lines of [13, 9], that it truly takes advantage of extra information in the variable assignment LLLL. On the other hand it is quite plausible that Theorem 6 is more along the lines of [14], namely, it provides a more accurate and computationally efficient approximation to Shearer's criterion.

In this paper, we will construct a problem instance for which Theorem 6 is satisfied, yet Shearer's criterion is violated. Thus, it is impossible to deduce the fact that $\Pr(\overline{\mathcal{B}}) > 0$ based only on the probabilities and interdependency structure of the bad-events; it is necessary to take into account the decomposition of the bad-events into variables (as is provided by Theorem 6). In other words, Theorem 6 can be stronger than Shearer's criterion.

We emphasize that Shearer's criterion concerns arbitrary probability spaces; one cannot hope to provide a stronger criterion than Shearer's *for the level of generality to which the latter applies*. The strength of Theorem 6 comes from its *less general* setting (the variable assignment LLLL), which is nevertheless encompasses many applications in combinatorics.

We also remark on other related criteria for the variable-assignment LLL setting. For instance, [9, 10] derive certain convergence conditions in terms of the bipartite graph H on vertex sets $\{1, \dots, m\}$ and $\mathcal{B}$ and an edge on (i, B) when $i \in \mathrm{var}(B)$, and [11] derives conditions in terms of the probabilities that certain neighboring bad-events hold simultaneously.

2 Satisfiability with bounded variable occurrence

Consider boolean k-satisfiability instances, where we have m boolean variables $X_1, \dots, X_m$ and n clauses $C_1, \dots, C_n$ of width k, each of the form

$$C_i \equiv l_{i1} \vee l_{i2} \vee \cdots \vee l_{ik}$$

for distinct literals $l_{i1}, \dots, l_{ik}$ (i.e. expressions of the form X_j or $\neg X_j$). The goal is to produce a value for the boolean variables $X_1, \dots, X_m \in \{T, F\}^m$ such that all the clauses C_i are simultaneously true. Equivalently, we want to find a satisfying assignment of the conjuctive-normal form (CNF) formula

$$\Phi = \bigwedge_{i=1}^{n} l_{i1} \vee l_{i2} \vee \dots \vee l_{ik}$$

We are interested specifically in instances where each variable appears in a bounded number of clauses. For each $i = 1, \dots, m$, define $R_0(\Phi, i)$ and $R_1(\Phi, i)$ to be the number of clauses which contain the literal X_i (respectively $\neg X_i$), and let $R(\Phi, i) = R_0(\Phi, i) + R_1(\Phi, i)$. In [15], Kratochvíl, Savický, and Tuza defined the function $f(k)$ as the largest integer L such that whenever $R(\Phi, i) \leqslant L$ for all i, then Φ is satisfiable; they showed $f(k) \geqslant \frac{2^k}{ek}$. A series of later works [21, 12, 5, 6] showed a variety of upper and lower bounds of $f(k)$. In particular, [6] showed

$$\left\lfloor \frac{2^{k+1}}{e(k+1)} \right\rfloor \leqslant f(k) \leqslant (1 + O(k^{-1/2})) \frac{2^{k+1}}{ek},$$

The lower bound comes from the variable-assignment LLLL. Here, the probability space Ω is defined by setting each variable $X_i = T$ with a certain probability p_i given by

$$p_i = 1/2 + x \frac{R_1(\Phi, i) - R_0(\Phi, i)}{R(\Phi, i)}$$

for some carefully chosen parameter $x \geqslant 0$. Then, for each clause C_i, there is a corresponding bad-event B_i that C_i is false, namely B_i has the form

$$(X_{i1} = j_{i1}) \wedge \dots \wedge (X_{ik} = j_{ik})$$

where $j_{i1}, \dots, j_{ik} \in \{T, F\}$. Using Theorem 6 in place of the LLLL, and using a slightly different value for the probabilities p_i, Harris [7] showed a stronger bound

$$f(k) \geqslant \frac{2^{k+1}(1 - 1/k)^k}{k - 1} - \frac{2}{k} \tag{3}$$

With these constructions, we thus know the asymptotic bound

$$f(k) \sim \frac{2^{k+1}}{ek};$$

nevertheless, there are two main reasons to determine $f(k)$ as precisely as possible. First, since $f(k)$ grows exponentially in k, the asymptotic value is not as relevant for practical applications. Second, [15] showed a sudden gap in the computational complexity of k-SAT: for problem instances where variables may appear in $f(k) + 1$ clauses, it is NP-complete to determine satisfiability. On the other hand, problems instances where they appear in at most $f(k)$ clauses are always satisfiable and the problem is computationally vacuous. Thus, tiny gaps in the value of $f(k)$ can lead to huge gaps in computational hardness.

2.1 Restricting the number of occurrences of each literal

Our goal is to demonstrate that the bound in Eq. (3) cannot be shown from the Shearer criterion. If the probability space Ω is allowed to vary in a problem-specific way, then any satisfiable formula can trivially satisfy the LLL: namely, Ω puts probability mass 1 on some satisfying assignment. Thus, in order to separate the LLL and Theorem 6, we must restrict Ω to be problem-independent.

In both the constructions of [6] and [7], the probabilities p_i depend solely on the imbalance between $R_0(\Phi, i)$ and $R_1(\Phi, i)$. They use slightly different formulas; however, in both constructions, the extremal case is when $R_0(\Phi, i) = R_1(\Phi, i)$, in which case p_i is set to 1/2.

Accordingly, let us define $f'(k)$ to be the largest integer L such that whenever $R_0(\Phi, i) \leqslant L$ and $R_1(\Phi, i) \leqslant L$ for all i, then the formula Φ is satisfiable. Clearly $f'(k) \geqslant f(k)/2$. This function is also studied in [5], with slightly different terminology, in terms of a combinatorial object called a (k, d)-tree.

Definition 7 ([6]).[1] A (k, d)-tree is a binary tree T where every leaf has depth at least k, and every node u of T has at most d descendant leaves within distance k of u.

We quote the following two results from [5] and [6]:

Theorem 8.

- *[5, Lemma 2] If there exists a $(k-1, d)$-tree, then there is an unsatisfiable k-CNF formula where every literal occurs in at most d clauses.*
- *[6, Theorem 1.3] For any $k \geqslant 1$, there exists a (k, d) tree with $d = (2/e + O(k^{-1/2}))2^k/k$*

This immediately gives the following result:

Theorem 9. $f'(k) \leqslant (1 + O(k^{-1/2}))\frac{2^k}{ek}$

Let us use the LLL and Theorem 6 to show more precise lower bounds on $f'(k)$. We will fix a problem-independent probability space Ω to set each X_i to be T with probability $p_i = 1/2$. For each clause C_i, we have a bad-event B_i with probability $\Pr(B_i) = p = 2^{-k}$.

Theorem 10 (Follows easily from the symmetric LLLL)**.** $f'(k) \geqslant \lfloor \frac{2^k}{ek} - 1/k \rfloor$

Proof. Consider some bad-event, without loss of generality

$$B \equiv (X_1 = T) \wedge \cdots \wedge (X_k = T)$$

The neighbors of B in the canonical lopsidependency graph G are bad-events involving $X_i = F$ for some $i = 1, \ldots, k$; as each literal occurs at most L times, there are at most $d = kL$ such bad-events. The symmetric LLLL criterion $ep(d+1) \leqslant 1$ then holds if $L \leqslant \frac{2^k}{ek} - 1/k$. □

[1] The definitions of (k, d)-trees are slightly shifted in the two papers; the object referred to as a (k, d)-tree in [6] is referred to as a $(k-1, d)$-tree in [5]. To put things on a consistent footing, we have adopted the terminology of [6].

Theorem 11 (From Theorem 6)**.** *Suppose that*

$$R_0(\Phi, i), R_1(\Phi, i) \leqslant \frac{(2^k - 1)(1 - 1/k)^{k-1}}{k}$$

for all i. Then the Moser-Tardos algorithm finds a satisfying assignment of Φ in expected polynomial time. In particular,

$$f'(k) \geqslant \left\lfloor \frac{(2^k - 1)(1 - 1/k)^{k-1}}{k} \right\rfloor$$

Proof. We will set $\mu(B) = \alpha$ for all $B \in \mathcal{B}$, where $\alpha \geqslant 0$ is some parameter to be determined. Consider some bad-event, without loss of generality

$$B \equiv (X_1 = T) \wedge \cdots \wedge (X_k = T)$$

It is difficult to list all orderable sets of neighbors of B according to Definition 5. However, to apply Theorem 6, we only need to provide an *upper bound* on the sum over such orderable sets (possibly including some additional neighbor-sets Y as well). Any such orderable set will have, for each $j = 1, \dots, k$, a choice of zero or one bad-events A_j which first disagree with B on variable X_j. (That is, in Definition 5, we have $B_i = A_j$ where $z_i = X_j$). Thus, we have an upper bound:

$$\sum_{\substack{Y \text{ orderable} \\ \text{to } B}} \prod_{A \in Y} \mu(A) \leqslant \prod_{j=1}^{k} (1 + R_1(\Phi, j)\alpha) \leqslant (1 + L\alpha)^k$$

So a sufficient criterion to satisfy Theorem 6 is

$$\alpha \geqslant 2^{-k}(\alpha + (1 + L\alpha)^k) \tag{4}$$

We choose α to maximize $\alpha - 2^{-k}(\alpha + (1+L\alpha)^k)$; simple calculus gives $\alpha = \frac{\left(\frac{2^k-1}{kL}\right)^{\frac{1}{k-1}} - 1}{L}$, which is non-negative for $L \leqslant \frac{2^k-1}{k}$. With this choice of α, the condition (4) is satisfied for

$$L \leqslant \frac{(2^k - 1)(1 - 1/k)^{k-1}}{k}$$

Thus, if $L \leqslant \frac{(2^k-1)(1-1/k)^{k-1}}{k}$ and $L \leqslant \frac{2^k-1}{k}$, then Theorem 6 is satisfied. The second condition $L \leqslant \frac{2^k-1}{k}$ can be easily seen to be redundant, leading to the given bounds. □

In either case, we have $f'(k) \sim 2^k/(ek) \sim f(k)/2$. Let us define $F_{\mathrm{LLL}}(k) = \lfloor \frac{2^k}{ek} - 1/k \rfloor$ and $F_{\mathrm{MT}}(k) = \lfloor \frac{(2^k-1)(1-1/k)^{k-1}}{k} \rfloor$ to be the bounds on $f'(k)$ which are provable respectively from the symmetric LLLL (Theorem 10) and from the criterion of Theorem 11. We observe that

$$F_{\mathrm{MT}}(k) - F_{\mathrm{LLL}}(k) \geqslant \frac{2^k}{2ek^2} - 1$$

So the gap between the LLL and Theorem 6 appears to be growing exponentially in k. (The relative difference between the formulas approaches zero, however).

3 Constructing the extremal formula Φ

Let us fix integers L, k. We will construct a k-SAT instance Φ with $R_0(\Phi, i), R_1(\Phi, i) \leqslant L$, in which the Shearer criterion is *violated* for the canonical lopsidependency graph corresponding to the natural space Ω where $\Pr(X_i = T) = 1/2$, and all variables X_i are independent, and with the natural collection of bad-events corresponding to the clauses. However, $L \leqslant F_{\text{MT}}(k)$; thus Theorem 6 ensures that Φ is satisfiable.

To begin the construction, start with Φ_0 containing no clauses (i.e. Φ_0 is the tautology). At stage i of the process, we modify Φ_{i-1} to produce a new formula Φ_i by adding $L-1$ clauses in which i appears positively and $L-1$ clauses in which i appears negatively. All other variables in these clauses are completely new, not appearing in any clause of Φ_{i-1}; they all appear positively in the $2L-2$ new clauses, and each of the new variables (other than variable i) appears in exactly one new clause.

Note that $\Pr(B) = p = 2^{-k}$ for all bad-events. Furthermore, since each variable i has exactly one positive occurence added in some iteration $\Phi_{i'}$ for $i' \neq i$, we have

$$R_0(\Phi_j, i) \leqslant L \qquad R_1(\Phi_j, i) \leqslant L - 1$$

for all i, j.

Define G_ℓ be the canonical lopsidependency graph corresponding to the bad-events for the formula Φ_ℓ. Although these graphs are complex, they contain a relatively simple and regular family of subgraphs H_j. We will show that Shearer's criterion is violated for these subgraphs; as shown in [20], this implies that Shearer's criterion is violated for the overall graph G_ℓ.

The graph family H_j will consist of many copies of $K_{L-1,L-1}$, the complete bipartite graph with $L-1$ vertices on each side. Each graph H_j has a special copy of $K_{L-1,L-1}$, called the *root* of H_j. We define these graphs recursively. First, H_0 is the empty graph. To form H_{j+1}, we start by taking a new copy of $K_{L-1,L-1}$ designated as the root of H_{j+1}. For each vertex v in this root, we add $k-1$ separate new copies of H_j, along with an edge connecting v to all the vertices in the right-half of the root of the corresponding H_j.

For example, H_1 consists of a single copy of $K_{L-1,L-1}$. See Figure 1.

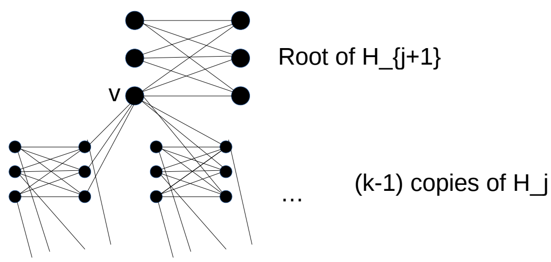

Figure 1: Construction of H_{j+1} from H_j. We have only shown here two copies of H_j corresponding to a single vertex v in the root of H_{j+1}. There are $k-1$ copies of H_j for each vertex in the root of H_{j+1} (a total of $2(L-1)(k-1)$ copies of H_j).

Proposition 12. *Any graph H_j appears as a subgraph of G_ℓ for some ℓ sufficiently large.*

Proof. Define A_i to be the collection of clauses in Φ_i but not Φ_{i-1}. We can also define a tree structure $\mathcal{T}$ on the variables of Φ: variable i is a parent of variable j if variable j appears in Φ_i but not Φ_{i-1}. For any variable i, let $\mathcal{T}_i$ denote the subtree of $\mathcal{T}$ rooted at i.

For any set of variables S, define $G_\ell[S]$ to be the subgraph of G_ℓ induced on the clauses ϕ of Φ_ℓ such that all variables in ϕ come from S. Observe that if S, S' are disjoint sets of variables then $G_\ell[S], G_\ell[S']$ are also vertex-disjoint graphs.

We will prove by induction on j a stronger claim: for any variable i, there is some integer $D(i,j)$ sufficiently large such that the induced subgraph $G_{D(i,j)}[\mathcal{T}_i]$ contains a copy of H_j, and the root of this copy of H_j corresponds to the clauses of A_i.

When $j = 0$ this is vacuously true. For the induction step, consider some variable i. Let C denote the $(2L-2)(k-1)$ variables which are children of i in $\mathcal{T}$. By inductive hypothesis, for each $i' \in C$, the graph $G_{D(i',j-1)}[\mathcal{T}_{i'}]$ contains a copy of H_{j-1} whose root corresponds to $A_{i'}$.

Let $\ell = i + \max_{i' \in C} D(i', j-1)$; we claim that the choice $D(i,j) = \ell$ satisfies the induction claim. For, in the graph $G_\ell[\mathcal{T}_i]$, the clauses of A_i in which i appears positively are lopsidependent with those clauses in which i appears negatively. Thus, it has a copy of $K_{L-1,L-1}$ corresponding to A_i; we denote this copy by J. The graph $G_\ell[\mathcal{T}_i]$ also contains the disjoint graphs $G_\ell[\mathcal{T}_{i'}]$ for each $i' \in C$. For each such $i' \in C$, let $J_{i'}$ denote the corresponding copy of H_{j-1} in $G_\ell[\mathcal{T}_{i'}]$.

Consider some clause $\phi \in A_i$, corresponding to a vertex of J, and some variable $i' \neq i$ in this clause. The root of $J_{i'}$ corresponds to the clauses $A_{i'}$. Note that ϕ is the only clause of A_i in which i' appears, and it appears positively in ϕ. Variable i' also appears negatively in exactly $L-1$ clauses of $A_{i'}$, which correspond to the right-half of $J_{i'}$. Thus, there are edges from ϕ in J to all the right-vertices in $k-1$ copies of H_{j-1}. As this is true for every $\phi \in J$, the resulting graph is precisely H_j. This completes the induction. □

4 Computing the Shearer criterion for H_j

We now compute the Shearer criterion for the family of graphs H_j. For our intermediate calculations, we also need to work with another closely-related family of graphs. For each $j \geqslant 0$, define a graph H'_j by taking a single vertex v along with $k-1$ new copies of H_j. We include an edge from v to all the vertices in the right-half of the roots of H_j. See Figure 2.

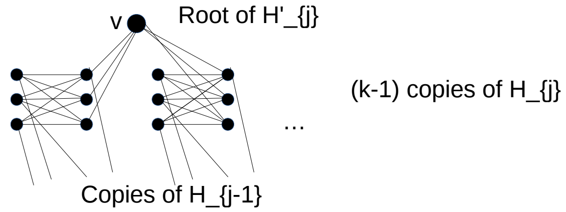

Figure 2: The construction of H'_j from H_j.

We will make use of two computational tricks for stable set polynomials; the proofs of these are elementary and are omitted here.

Proposition 13. *If vertex set V is partitioned into connected-components as $V = V_1 \sqcup V_2$, then*

$$Q(G, \emptyset, \vec{p}) = Q(G[V_1], \emptyset, \vec{p}) Q(G[V_2], \emptyset, \vec{p})$$

Proposition 14. *Suppose $X \subseteq V$. Then*

$$Q(G, \emptyset, \vec{p}) = \sum_{\text{stable set } U \subseteq X} Q(G[V - X - N(U)], \emptyset, \vec{p}) \prod_{i \in U} (-p_i)$$

We now begin the calculation.

Proposition 15. *Let us define*

$$r_j = Q(H'_j, \emptyset, \vec{p}) \qquad s_j = Q(H_j, \emptyset, \vec{p})$$

Then $r_0 = 1 - p, s_0 = 1$, and r, s satisfy the mutual recurrence relations for $j \geqslant 1$:

$$\begin{aligned} r_j &= s_j^{(k-1)} - p r_{j-1}^{(k-1)(L-1)} s_{j-1}^{(k-1)^2(L-1)} \\ s_j &= 2 r_{j-1}^{(L-1)} s_{j-1}^{(k-1)(L-1)} - s_{j-1}^{(k-1)(2L-2)} \end{aligned}$$

Proof. The base cases are clear, since H_0 is empty and H'_0 is a single node. We first show the bound on s_j for $j \geqslant 1$. In any stable set U of H_j, either U contains zero vertices from the left half of the root of H_j, or zero vertices from the right-half of the root of H_j, or both. In the first two cases, when we remove the vertices in the left (respectively right) half of H_j, then we are left with $L - 1$ copies of H'_{j-1} and $(k-1)(L-1)$ copies of H_{j-1}. In the third case, we are left with $(k-1)(2L-2)$ copies of H_{j-1}. We can sum the first two contributions and subtract the third, as it is double-counted: this gives

$$s_j = 2 r_{j-1}^{(L-1)} s_{j-1}^{(k-1)(L-1)} - s_{j-1}^{(k-1)(2L-2)}$$

Next consider the bound for r_j. Let v denote the root node of H'_j and let $J_1, \dots, J_{k-1}$ be the copies of H_j to which it is connected, and let P_i denote the root of each J_i. We apply Proposition 14 with $X = \{v\}$, and so either $U = \emptyset$ or $U = \{v\}$. For $U = \emptyset$, the graph $H'_j[V - X - N(U)]$ consists of $k - 1$ independent copies of H_j. For $U = \{v\}$, consider the graph $H'_j[V - X - N(U)]$: the vertices in the left half of P_i now yield $L - 1$ disconnected copies of H'_{j-1} and each vertex u in the right half of P_i now yields $k - 1$ disconnected copies of H_{j-1}. Over all $k - 1$ choices of i and all $(k-1)(L-1)$ choices for u in each P_i, we see that $H'_j[V - v - N(v)]$ consists of $(k-1)(L-1)$ copies of H'_{j-1} and $(k-1)^2(L-1)$ copies of H_{j-1}. See Figure 3.

Summing the contributions of these two terms according to Proposition 14 gives

$$r_j = Q(H'_j, \emptyset, \vec{p}) = s_j^{(k-1)} - p r_{j-1}^{(k-1)(L-1)} s_{j-1}^{(k-1)^2(L-1)}. \qquad \square$$

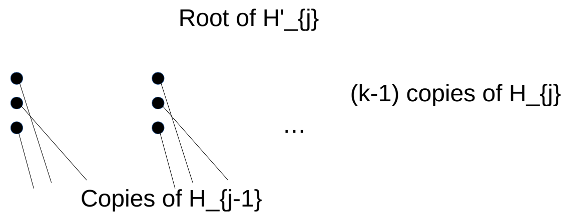

Figure 3: Removing the root node from H'_j

Proposition 16. *Suppose that G_ℓ satisfies the Shearer condition for all $\ell \geqslant 0$. Then, if we define the function $g : [0,1] \to \mathbb{R}$ by*

$$g(a) = 1 - \frac{p}{(2 - a^{-(L-1)})^{k-1}},$$

there is some $a \in (2^{\frac{-2}{2L-2}}, 1]$ satisfying $g(a) = a$.

Proof. For $j \geqslant 0$ define

$$a_j = \frac{r_j}{s_j^{(k-1)}}.$$

We will show a recurrence relation for a_j. Using Proposition 15, we calculate for $j \geqslant 1$:

$$a_j = \frac{s_j^{(k-1)} - p r_{j-1}^{(k-1)(L-1)} s_{j-1}^{(k-1)^2(L-1)}}{s_j^{(k-1)}} = 1 - \frac{p r_{j-1}^{(k-1)(L-1)}}{s_{j-1}^{(k-1)^2(L-1)}} \cdot \frac{s_{j-1}^{(k-1)^2(2L-2)}}{s_j^{(k-1)}} = 1 - \frac{p a_{j-1}^{(k-1)(L-1)}}{\left(\frac{s_j}{s_{j-1}^{(k-1)(2L-2)}}\right)^{k-1}}$$

Here again using Proposition 15, we get

$$\frac{s_j}{s_{j-1}^{(k-1)(L-2)}} = \frac{2 r_{j-1}^{(L-1)} s_{j-1}^{(k-1)(L-1)} - s_{j-1}^{(k-1)(2L-2)}}{s_{j-1}^{(k-1)(2L-2)}} = \frac{2 r_{j-1}^{(L-1)}}{s_{j-1}^{(k-1)(L-1)}} - 1 = 2 a_{j-1}^{(L-1)} - 1 \qquad (5)$$

and, substituting this into the equation for a_j, this implies:

$$a_j = 1 - \frac{p a_{j-1}^{(k-1)(L-1)}}{(2 a_{j-1}^{(L-1)} - 1)^{k-1}} = g(a_{j-1}). \qquad (6)$$

We must have $a_j > 2^{\frac{-2}{2L-2}}$ for all $j \geqslant 1$. For, if not, then Eq. (5) would otherwise imply that $\frac{s_j}{s_{j-1}^{(k-1)(2L-2)}} \leqslant 0$; thus, either $s_j \leqslant 0$ or $s_{j-1} \leqslant 0$. Thus, either H_j or H_{j-1} violates the Shearer condition, and so would some G_ℓ; this contradicts our hypothesis.

Now suppose $g(a) < a$ for all $a \in (2^{-\frac{-2}{2L-2}}, 1]$, so from Eq. (6) the sequence $a_1, a_2, \ldots$ decreases monotonically. Because of the lower bound $a_j \geqslant 2^{\frac{-2}{2L-2}}$, it converges to some limit point $a \geqslant 2^{\frac{-2}{2L-2}}$. By continuity, this must be a fixed point, i.e. $g(a) = a$, as desired.

Furthermore, since $g(a)$ diverges to infinity at $a = 2^{\frac{-2}{2L-2}}$, we must indeed have $a > 2^{-\frac{2}{2L-2}}$ strictly.

Otherwise, suppose that $g(a) \geqslant a$ for some $a \in (2^{-\frac{-2}{2L-2}}, 1]$. Observe that $g(1) = 1 - p < 1$. Hence, the function $g(a) - a$ changes sign on the interval $(2^{-\frac{-2}{2L-2}}, 1]$. This implies there must be a fixed point $g(a) = a$ on this interval. □

Proposition 17. *Suppose*

$$L > 1 - \frac{\ln(2-t)}{\ln\left(1 - 2^{-k} t^{1-k}\right)}$$

for all $t \in (2^{k/(k-1)}, 2)$. Then the Shearer condition is violated on G_ℓ, for ℓ sufficiently large.

Proof. Suppose not; by Proposition 16, the function g then has a fixed point $a \in (2^{-\frac{-2}{2L-2}}, 1]$. So

$$a = 1 - \frac{2^{-k}}{(2 - a^{-(L-1)})^{k-1}}$$

Solving for L, we thus obtain:

$$L = 1 - \frac{\ln\left(2 - 2^{\frac{k}{1-k}}(1-a)^{\frac{1}{1-k}}\right)}{\ln a} \qquad \text{for } t = 2^{k/(1-k)}(1-a)^{1/(1-k)} \tag{7}$$

where here $t \in (2^{k/(k-1)}, 2)$. This contradicts our hypothesis. □

For any $k \geqslant 1$, let us define the quantity $\tilde{F}_{\text{Shearer}}(k)$ by:

$$\tilde{F}_{\text{Shearer}}(k) = \left\lfloor \max_{t \in (2^{k/(k-1)}, 2)} 1 - \frac{\ln(2-t)}{\ln\left(1 - 2^{-k} t^{1-k}\right)} \right\rfloor$$

In light of Proposition 17, this is an upper bound on the value of $f'(k)$ that can be shown using the LLL or any variant of it. We observe that $\tilde{F}_{\text{Shearer}}(k) \geqslant F_{\text{LLL}}(k)$ for all values of k — this must be the case, since the bound F_{LLL} was indeed derived using the LLL and this is always weaker than Shearer's criterion. To illustrate, we list $F_{\text{LLL}}, \tilde{F}_{\text{Shearer}}$, and F_{MT} for a few small values of k in Table 1.

The gap between $\tilde{F}_{\text{Shearer}}$ and F_{LLL} is very small, suggesting that there is little to no improvement possible in the bound for $f'(k)$ from a more advanced more of the LLL.

We next derive an asymptotic approximation to $\tilde{F}_{\text{Shearer}}$.

Proposition 18. $\tilde{F}_{\text{Shearer}} = \frac{2^k}{ek} + \Theta(\frac{2^k}{k^3})$

Proof. We can show the lower bound by taking $t = 1 - 1/k$, i.e.

$$\tilde{F}_{\text{Shearer}} \geqslant \left\lfloor 1 - \frac{\ln(2-t)}{\ln(1 - 2^{-k} t^{1-k})} \right\rfloor \geqslant -\frac{\ln(2-t)}{\ln(1 - 2^{-k} t^{1-k})} = \frac{2^k}{ek} + \Omega(\frac{2^k}{k^3}).$$

k	F_{LLL}	$\tilde{F}_{\mathrm{Shearer}}$	F_{MT}
9	20	21	22
10	37	38	39
11	68	69	71
12	125	126	131
13	231	233	241
14	430	432	446
15	803	806	831
16	1506	1510	1555
17	2836	2842	2922
18	5357	5366	5511
19	10151	10165	10426
20	19287	19311	19784

Table 1: $F_{\mathrm{LLL}}, \tilde{F}_{\mathrm{Shearer}}$, and F_{MT} for a few small values of k.

For the lower bound, let $L = \tilde{F}_{\mathrm{Shearer}}(k)$, so that

$$L \leqslant 1 - \frac{\ln(2-t)}{\ln\left(1 - 2^{-k} t^{1-k}\right)}$$

for some $t \in (2^{k/(k-1)}, 2)$. Using the bound $-\ln(1-x) \geqslant x$ for $x \geqslant 0$, we have:

$$L \leqslant 1 + t^{k-1} 2^k \ln(2-t) \tag{8}$$

Since $\ln(2-t)$ is a concave-down function of t, we have the bound

$$\ln(2-t) \leqslant \ln(2-t_0) + \frac{t_0 - t}{2 - t_0}$$

for any chosen value $t_0 \in (0, 2)$. Substituting this bound into (8), and differentiating with respect to t to maximize the resulting value, we get

$$L \leqslant 1 + \frac{\Big(2(1-1/k)(t_0 + (2-t_0)\ln(2-t_0))\Big)^k}{(2-t_0)(k-1)} \tag{9}$$

If we set $t_0 = 1 - 1/k$ in Eq. (9), then straightforward analysis gives:

$$L \leqslant \frac{2^k}{ek} + O(\frac{2^k}{k^3}) \qquad \square$$

On the other hand, one can easily verify that $F_{\mathrm{MT}}(k) \geqslant \frac{2^k}{ek} + \Omega(\frac{2^k}{k^2})$; thus, there is a large and growing gap between F_{MT} and $\tilde{F}_{\mathrm{Shearer}}$.

Acknowledgements

Thanks to Aravind Srinivasan and anonymous journal referees for many helpful comments and suggestions.

References

[1] Albert, M., Frieze, A., Reed, B. Multicoloured Hamilton Cycles. *The Electronic Journal of Combinatorics* 2:#R10, 1995.

[2] Bissacot, R., Fernandez, R., Procacci, A., Scoppola, B. An improvement of the Lovász Local Lemma via cluster expansion. *Combinatorics, Probability and Computing* 20(5): 709–719, 2011.

[3] Erdős, P., Lovász, L. Problems and results on 3-chromatic hypergraphs and some related questions. In A. Hajnal, R. Rado, and V. T. Sos, eds. Infinite and Finite Sets II, pages 607–726, 1975.

[4] Erdős, P., Spencer, J. Lopsided Lovász Local Lemma and Latin transversals. *Discrete Applied Math* 30: 151–154, 1990.

[5] Gebauer, H. Disproof of the neighborhood conjecture with implications to SAT *Combinatorica* 32(5):573–587, 2012.

[6] Gebauer, H., Szabó, T., Tardos, G. The local lemma is asymptotically tight for SAT. *Journal of the ACM* 63(5), Article #43, 2016.

[7] Harris, D. Lopsidependency in the Moser-Tardos framework: beyond the Lopsided Lovász Local Lemma. *ACM Transactions on Algorithms* 13(1), Article #17, 2016.

[8] Harvey, N., Vondrák, J. An algorithmic proof of the Lovász local lemma via resampling oracles. Proc. 56th annual IEEE Sympsium on Foundations of Computer Science (FOCS), pages 1327–1346, 2015.

[9] He, K., Li, L., Liu, X., Wang, Y., Xia, M. Variable version Lovász local lemma: beyond Shearer's bound. Proc. 58th annual IEEE Symposium on Foundations of Computer Science (FOCS), pages 451–462, 2017.

[10] He., K., Li, Q., Sun, X., Zhang, J. Quantum Lovász local lemma: Shearer's bound is tight. Proc. 51st annual ACM SIGACT Symposium on Theory of Computing (STOC), pages 461–472, 2019.

[11] He, K., Li, Q., Sun, X. Moser-Tardos algorithm: beyond Shearer's bound. arXiv:2111.06527, 2021.

[12] Hoory, S., Szeider, S. Computing unsatisfiability k-SAT instances with few occurrences per variable. *Theoretical Computer Science* 337(1-3):347–359, 2005.

[13] Kolipaka, K., Szegedy, M. Moser and Tardos meet Lovász. Proc. 43rd annual ACM Symposium on Theory of Computing (STOC), pages 235–244, 2011.

[14] Kolipaka, K., Szegedy, M., Xu, Y. A sharper local lemma with improved applications. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques* LNCS 7408, pages 603–614, 2012.

[15] Kratochvíl, J., Savický, P., Tuza, Z. One more occurrence of variables makes satisfiability jump from trivial to NP-complete. *SIAM Journal of computing* 22(1):203–210, 1993.

[16] Lu, L., Székély, L. Using Lovász Local Lemma in the space of random injections. *The Electronic Journal of Combinatorics* 13:#R63, 2007.

[17] Lu, L., Székély, L. A new asymptotic enumeration technique: the Lovász local lemma. arXiv:0905.3983v3, 2011.

[18] McDiarmid, C. Hypergraph coloring and the Lovász Local Lemma. *Journal of Discrete Mathematics* 167/168:481–486, 1995.

[19] Moser, R., Tardos, G.: A constructive proof of the general Lovász Local Lemma. *Journal of the ACM* 57(2), Article #11, 2010

[20] Shearer, J. B. On a problem of Spencer. *Combinatorica* 5:241–245, 1985.

[21] Savický, P., Sgall, J. DNF tautologies with a limited number of occurrences of every variable. *Theoretical Computer Science* 238(1-2):495–498, 2000.