Harris's lopsidependency criterion can be stronger than Shearer's criterion

David G. Harris*

Abstract

The Lopsided Lovász Local Lemma (LLLL) is a probabilistic tool which is a cornerstone of the probabilistic method of combinatorics, which shows 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 that can be stated 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 constructive algorithm for nearly all applications of it, including applications to k-SAT instances with a bounded number of occurrences per variables. Harris (2015) later gave an alternate criterion for this algorithm to converge, which appeared to give stronger bounds than the standard LLL. 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.

1 Introduction

1.1 The Lovász Local Lemma

The Lovász Local Lemma (LLL) is a general probabilistic principle, first introduced in [3], for showing that, if one has a probability space Ω with a finite set $\mathcal{B}$ of "bad" events in that space, then as long as the bad-events are not interdependent and are not too likely, then there is a positive probability no events in $\mathcal{B}$ occur. This principle has become a cornerstone of the probabilistic method of combinatorics, as this establishes that a configuration avoiding $\mathcal{B}$ exists.

The definition of interdependency in the context of the LLL is somewhat technical. It is stated in terms of a *dependency graph* G, whose vertex set is $\mathcal{B}$. This graph G must satisfy the following condition: for any $B \in \mathcal{B}$ and any set $S \subseteq \mathcal{B} - \{B\} - N_G(B)$, we have

$$P(B \mid \bigcap_{A \in S} \overline{A}) = P(B) \tag{1}$$

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. (In this paper, we let $N_G(B)$ denote the neighborhood of B in the graph G).

We note that, given a probability space Ω and collection of bad-events $\mathcal{B}$, there is no unique dependency graph G. Rather, we suppose that we are given $\Omega, \mathcal{B}$ and some graph G which is a dependency graph for them.

*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

With these definitions, we can state the simplest form of the LLL, known as the symmetric LLL:

Theorem 1.1 (Symmetric LLL)**.** *Suppose Ω is a probability space, $\mathcal{B}$ is a finite set of events in Ω, and G is a dependency graph for Ω, B whose maximum degree is d, and that for each $B \in \mathcal{B}$ we have $P(B) \leq p$. Then if the criterion $ep(d+1) \leq 1$ is satisfied, then $P(\bigcap_{B \in \mathcal{B}} \overline{B}) > 0$.*

For any collection of events $\mathcal{B}$, we define the event $\overline{\mathcal{B}} = \bigcap_{B \in \mathcal{B}} \overline{B}$; we refer to this event as *avoiding $\mathcal{B}$*.

There have been numerous extensions and applications of the LLL since its original formulation. Two themes will be relevant for us here. First, there is the generalization of the LLL known as the *Lopsided Lovász Local Lemma* (LLLL). This was introduced in [4], which observed 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. More specifically, given a graph G, we say that G is a *lopsidependency graph* for Ω, B if for any $B \in \mathcal{B}$ and any set $S \subseteq \mathcal{B} - \{B\} - N_G(B)$, we have

$$P(B \mid \bigcap_{A \in S} \overline{A}) \leq P(B) \tag{2}$$

A second extension of the LLL is determining alternate criteria to ensure $P(\overline{\mathcal{B}}) > 0$. For example, the asymmetric LLL criterion can take advantage of situations in which the bad-events B have different probabilities. (The symmetric LLL uses a single quantity p to upper-bound all the bad-events' probabilities). In [19], Shearer derived the most powerful possible criterion that can be stated in terms of the probabilities of the bad-events and a lopsidependency graph G on them. We will summarize this criterion after giving a few relevant definitions.

Definition 1.2. *Let G be a graph on vertex set $[n]$ and let $p_1, \dots, p_n$ be real numbers. The* independent set polynomial *of G with respect to base set $S \subseteq V$, which we denote $Q(G, S, \vec{p})$, is defined by*

$$Q(G, S, \vec{p}) = \sum_{\substack{S \subseteq T \subseteq V \\ T \text{ independent}}} (-1)^{|T|-|S|} \prod_{i \in T} p_i$$

(In this definition, T independent means that no vertices in T are adjacent in G).

We can now state Shearer's criterion:

Theorem 1.3 (Shearer's criterion [19])**.** *Let G be a graph on vertex set $[n]$ and let $p_1, \dots, p_n \in [0, 1]$.*

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

$$P(\bigcap_{j=1}^{n} \bar{B}_j) > 0$$

   In this case, we say that Shearer's criterion is satisfied *by G, p.*

2. *Suppose that either $Q(G, \emptyset, \vec{p}) \leq 0$ or there is some independent set $S \subseteq V$ with $Q(G, S, \vec{p}) < 0$. Then there is some probability space Ω and events $B_1, \dots, B_n \subseteq \Omega$ in that probability space*

such that $P_\Omega(B_i) = p_i$ for $i = 1, \dots, n$ and such that G is a dependency graph for $\{B_1, \dots, B_n\}$ and such that

$$P(\bigcap_{j=1}^{n} \bar{B}_j) = 0$$

In this case, we say that Shearer's criterion is violated *by G, p.*

Thus, Shearer's criterion exactly characterizes what conditions on the probability of the bad-events and their lopsidependency guarantee a positive probability of avoiding $\mathcal{B}$.

Having bad-events with probability 0 or 1 is not so interesting, and Theorem 1.3 can be simplified when we disallow these cases.

Theorem 1.4 ([8], Lemma 5.27)**.** *Let G be a graph on vertex set $[n]$ and let $p_1, \dots, p_n \in (0, 1)$.*

1. *Suppose that $Q(G, \emptyset, \vec{p}) > 0$ for all independent sets $S \subseteq V$. Then Shearer's criterion is satisfied by G, p.*

2. *Suppose that there is some independent set $S \subseteq V$ with $Q(G, S, \vec{p}) \leq 0$. Then Shearer's criterion is violated by G, p.*

While powerful, Shearer's criterion is technically difficult to apply to constructions in combinatorics. There has been significant research in developing forms of the LLLL which are simpler to calculate, for example [2] and [12]; these have led to improved bounds for a variety of combinatorial constructions. However, from a theoretical point of view, these are all weaker than, and are all implied by, Shearer's criterion.

1.2 The variable-assignment LLLL

The LLLL is a general principle which has been applied to diverse probability spaces such as random permutations [14], Hamiltonian cycles [1], and matchings on the complete graph [15]. 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 is defined by a series of m variables $X_1, \dots, X_m$, which take their values from the integers $\mathbf{Z}$; namely, for each $i = 1, \dots, m$ and each $z \in \mathbf{Z}$ we have $P(X_i = z) = p_{iz}$, and the random variables $X_1, \dots, X_m$ are all mutually independent. In this setting, one may assume without loss of generality that the bad-events are atomic; that is, each $B \in \mathcal{B}$ can be written

$$B \equiv X_{i_1} = z_1 \wedge X_{i_2} = z_2 \wedge \dots \wedge X_{i_k} = z_k$$

We abuse notation for such atomic events, so that B is identified with the set $\{(i_1, j_1), \dots, (i_k, j_k)\}$. Thus, for instance, when we write $(i, j) \in B$, we mean that one of the conditions of B is that $X_i = j$. For such a bad-event B we define $\text{var}(B) = \{i_1, \dots, i_k\}$.

Given any ordered pairs (i, j) and bad-event B, we say that $(i, j) \sim B$ if $(i, j') \in B$ for some $j' \neq j$. We say that B, B' *disagree* on i if there are $j \neq j'$ with $(i, j) \in B, (i, j') \in B'$. We write $B \sim B'$ if B and B' disagree on some variable $i \in [m]$.

For such a probability space Ω, there are two natural choices for the lopsidependency graph G.

Definition 1.5. *The* canonical lopsidependency graph *G has vertex set $\mathcal{B}$, and edge connecting B, B' iff $B \sim B'$. The* canonical dependency graph *G has vertex set $\mathcal{B}$, and edge connecting B, B' iff $var(B) \cap var(B') \neq \emptyset$.*

Theorem 1.6. *The canonical lopsidependency graph and canonical dependency graph are lopsidependency graph for $\Omega, \mathcal{B}$.*

Proof. The first result follows from the FKG inequality. The second result is clear since the canonical lopsidependency dependency graph is a subgraph of the canonical dependency graph. □

As we have noted, this setting from the LLLL is general enough to cover most constructions in combinatorics. Some noteworthy applications of this principle include monochromatic hypergraph coloring [16] and satisfiability [6].

In [11], Kolipaka & Szegedy noted that the Shearer criterion is not tight for the variable-assignment LLL setting. They constructed an explicit dependency graph structure and vector of probabilities where the Shearer criterion is violated yet any variable-assignment realization must have a satisfying assignment. However, this was only a small-scale example, and it was not clear how extensive this phenomenon was or whether any systematic criterion could be provided for the variable-assignment LLL setting.

Along similar lines He et al. [9] analyzed the bipartite graph H on vertex sets $[n], \mathcal{B}$, with a edge on i, B if $i \in \text{var}(B)$. The canonical dependency graph (but not the canonical lopsidependency graph) can be recovered from H (namely, B, B' have an edge in G if they have a length-two path in H). Since the graph H is only defined in the context of the variable-assignment LLL, it captures a type of dependency information which takes advantage of the variable-assignment LLL structure. They showed that, in many cases, G does not satisfy the Shearer criterion, although information about H guarantees the existence of a configuration avoiding $\mathcal{B}$. Thus, in a sense, the variable-assignment LLL can allow one to go beyond Shearer's bound.

1.3 The Moser-Tardos algorithm

The LLLL ensures that $P(\overline{\mathcal{B}}) > 0$. This is typically sufficient for the probabilistic method of combinatorics, in which the main goal is show that a configuration avoiding $\mathcal{B}$ exists. However, usually $P(\overline{\mathcal{B}})$ is exponentially small, and hence the LLLL does not give efficient algorithms for *constructing* such a configuration.

In [17], Moser & Tardos introduced a remarkably simple algorithm to find configurations for the variable-assignment LLLL setting:

1. Draw each variable independently from the distribution Ω.

2. While there is some true bad-event:

   2a. Choose a true bad-event B arbitrarily.

   2b. Resample all the variables involved in 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 such as [18], [11] 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 termination of the Moser-Tardos algorithm. Unlike the symmetric LLLL or other similar criteria, this cannot be stated solely in terms of the dependency graph for G and the probabilities of the bad-events. Rather, it depends critically on the variable-assignment LLLL setting and the decomposition of bad-events into conjunction of atomic terms. We summarize this criterion, which we refer to as the *Lopsided MT criterion*, here.

Definition 1.7 (Orderability)**.** *Given an event $B \in \mathcal{B}$, we say that a set of bad-events $Y \subseteq \mathcal{B}$ is* orderable *to B, if either of the conditions hold:*

(O1) $Y = \{B\}$, *or*

(O2) there is some ordering $Y = \{B_1, \dots, B_s\}$, *with the following property. For each* $i = 1, \dots, s$, *there is some* $z_i \in B$ *such that* $z_i \sim B_i, z_i \not\sim B_1, \dots, z_i \not\sim B_{i-1}$.

Theorem 1.8 ([7])**.** *In the variable-assignment setting, suppose there is* $\mu : \mathcal{B} \to [0, \infty)$ *satisfying the following condition:*

$$\forall B \in \mathcal{B}, \mu(B) \geq P(B) \sum_{\substack{Y \text{ orderable} \\ \text{to } B}} \prod_{A \in Y} \mu(A)$$

Then the Moser-Tardos algorithm terminates with probability 1.

We note that Theorem 1.8 is superficially similar to the cluster-expansion criterion of [2]; the difference is that [2] requires Y is an independent set of neighbors of B (which is typically, although not always, a weaker condition than Y being orderable to B).

In [7], Harris also discussed a variety of combinatorial constructions based on Theorem 1.8, which appeared to lead to stronger bounds than had been shown using the standard LLLL or its variants. Harris also showed that Theorem 1.8 was always stronger than certain commonly-used variants of the asymmetric LLLL and cluster-expansion criterion.

It is not clear from [7], whether Theorem 1.8 could truly be stronger than Shearer's criterion. Is is quite plausible, along the lines of [11] and [9], that it more accurately carves out the satisfying assignment for the variable-assignment setting. On the other hand it is quite plausible that the improvement given by Theorem 1.8 is more along the lines of [12], namely, it provides a more accurate and computationally efficient approximation to Shearer's criterion.

We emphasize that Shearer's criterion is a general result concerning 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 1.8 comes from the fact that it applies to a *less general* setting (the variable assignment LLLL), which is nevertheless sufficiently general to encompass many applications in combinatorics.

In this paper, we will construct a problem instance for which Theorem 1.8 is satisfied, yet Shearer's criterion is violated. In other words, it is impossible to deduce the fact that $P(\overline{\mathcal{B}}) > 0$ 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 1.8). We can summarize this situation as showing that the lopsidependent Moser-Tardos criterion can be stronger than Shearer's criterion.

2 Satisfiability with bounded variable occurrence

Let us consider boolean k-satisfiability instances, in which the number of occurrences of each variable is bounded. Specifically, in such a problem instance we have m boolean variables and n clauses $C_1, \dots, C_n$ each of the form

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

where $l_{i1}, \dots, l_{ik}$ are distinct literals (i.e. an expression of the form X_j or $\overline{X}_j$). The goal is to produce a value for the boolean variables $X_1, \dots, X_m$ such that all the clauses C_i are simultaneously true on the assignment X. For each $i = 1, \dots, m$, we 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 we define $R(\Phi, i) = R_0(\Phi, i) + R_1(\Phi, i)$.

Such a problem instance can be be viewed as equivalent to determining the satisfiability of the formula

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

The formula Φ is in conjunctive-normal form, a conjunction of m disjunctions of k literals.

In [13], Kratochvíl, Savický, and Tuza defined the function $f(k)$ as the largest integer L such that whenever $R(\Phi, i) \leq L$ for all $i = 1, \dots, m$, then the formula Φ is satisfiable. They showed the bound $f(k) \geq \frac{2^k}{ek}$. A series of later works [21, 10, 5] showed a variety of upper and lower bounds of $f(k)$; most recently [6] showed that

$$\Big\lfloor \frac{2^{k+1}}{e(k+1)} \Big\rfloor \leq f(k) \leq (2/e + O(1/\sqrt{k})) \frac{2^k}{k},$$

and in particular

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

The lower bound comes from the variable-assignment LLLL, and is perhaps its most important and exemplary application. The probability space Ω is defined by setting each variable $X_i = T$ with probability p_i, and $X_i = F$ with probability $1 - p_i$. For each clause C_i, we have a bad-event that C_i is false. Each such bad-event B_i can be written

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

where $j_{i1}, \dots, j_{ik} \in \{T, F\}$.

The probabilities p_i are selected in a delicate and problem-specific way. Specifically, [6] sets

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

where $x \geq 0$ is a carefully chosen parameter. Determining the correct formula for p_i was one of the major technical innovations of [6]. This is counter-intuitive because one would expect that if $R_0(i)$ is larger than $R_1(i)$, then one should be more likely to set $X_i = T$ (as this will satisfy more clauses, on average); the formula in [6] does precisely the opposite.

While the bound of [6] is asymptotically tight, it is still open to determine the exact value of $f(k)$. One might argue that for algorithmic applications, the asymptotic value of $f(k)$ is not too relevant since $f(k)$ grows exponentially in k. Harris showed a slightly stronger bound:

Theorem 2.1 ([7]). *Suppose that*

$$R(\Phi, i) \leq \frac{2^{k+1}(1 - 1/k)^k}{k - 1} - \frac{2}{k}$$

for all $i = 1, \dots, m$. Then Φ is satisfiable, and the Moser-Tardos algorithm finds a satisfying assignment in expected polynomial time.

In particular,

$$f(k) \geq \frac{2^{k+1}(1 - 1/k)^k}{k - 1} - \frac{2}{k}$$

The construction of [7] is quite similar to that of [6], except that it uses Theorem 1.8 instead of the LLLL.

2.1 Restricting the number of occurrences of each literal

It is possible a priori that the gap between the bound of [7] and that of [6] is due to the latter not taking advantage of the strongest available form of the LLL. Our goal in this paper is to show that this is not the case: the bound of Theorem 2.1 cannot be shown even using the full Shearer criterion.

We note that if the probability space Ω is allowed to vary in a problem-specific way, then for any satisfiable problem instance we can always satisfy the LLL criterion trivially: namely, Ω puts probability mass 1 on some satisfying assignment. Thus, in order to achieve a separation between the LLL and MT criteria, we must restrict Ω to be problem-independent.

In both the LLL constructions of [6] and [7], the probabilities p_i depend solely on the imbalance between $R_0(\Phi, i)$ and $R_1(\Phi, i)$. The formulas for p_i are slightly different in these two constructions. It turns out that the extremal case for both constructions is when $R_0(\Phi, i) = R_1(\Phi, i)$, and in those cases we set $p_i = 1/2$.

We are thus lead to define the function $f'(k)$ as the largest integer L such that whenever $R_0(\Phi, i) \leq L$ and $R_1(\Phi, i) \leq L$ for all $i = 1, \dots, m$, then the formula Φ is satisfiable. We clearly have $f'(k) \geq f(k)/2$. Using results of [6], we can show that $f'(k)$ is indeed within a $(1 + o(1))$ factor of $f(k)/2$.

Theorem 2.2. *We have*

$$f'(k) \leq (1/e + O(k^{-1/2}))\frac{2^k}{k}$$

Proof. This is based on a construction introduced in [6] referred as a (k, d) tree. The key result of [6] is the existence of (k, d) trees for every integer $k \geq 1$ with

$$d \leq (2/e + O(k^{-1/2}))\frac{2^k}{k}$$

Furthermore, [6] shows that the existence of $(k-1, d)$ tree leads to the existence of an unsatisfiable k-SAT formula in which each literal appears at most d times. (This result appears in the proof of Theorem 1.4 of that paper). This implies the existence of an unsatisfiable formula in which every literal appears at most

$$(2/e + O((k-1)^{-1/2}))\frac{2^{k-1}}{k-1} = (1/e + O(k^{-1/2}))\frac{2^k}{k}$$

times.

Please see [6] for more information about this construction. □

This shows that indeed $f'(k) \sim \frac{2^k}{ek} \sim f(k)/2$. Our goal is to use the LLL to show more precise lower bounds on $f'(k)$ (that are asymptotically tight beyond the first-order). We will restrict our probability space Ω to set $X_i = T$ or $X_i = F$ with probability 1/2. For this restricted case, we can also show simpler and slightly stronger bounds than were shown in [7], [6].

Theorem 2.3 (Follows easily from the symmetric LLLL)**.** *We have* $f'(k) \geq \lfloor \frac{2^k}{ek} - 1/k \rfloor$

Proof. We have a bad-event for each clause. Consider some bad-event, without loss of generality

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

This event has probability $P(B) = p = 2^{-k}$. Also, the neighbors of B in the canonical lopsidependency graph G are bad-events involving $X_i = F$ for some $i = 1, \dots, k$; as each literal occurs at most L times, there are at most $d = kL$ such bad-events.

Thus, by the symmetric LLLL, it is possible to avoid all such bad-events (and in particular Φ is satisfiable), if $ep(d+1) \leq 1$, which occurs iff

$$L \leq \frac{2^k}{ek} - 1/k$$

$\square$

Theorem 2.4 (From the lopsidependent Moser-Tardos criterion)**.** *Suppose that*

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

for all $i = 1, \dots, m$. Then Φ is satisfiable, and the Moser-Tardos algorithm finds a satisfying assignment in expected polynomial time.

In particular,

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

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

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

This event has probability $P(B) = p = 2^{-k}$. We may form an orderable set Y of neighbors of B as follows: first, we may set $Y = \{B\}$. Second, for each $i = 1, \dots, k$, we may select zero or one bad-events B_i which disagree with B on variable i. Thus, we have that

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

Thus, a sufficient criterion to satisfy Theorem 1.8 is to have

$$\alpha \geq 2^{-k}(\alpha + (1 + L\alpha)^k) \tag{3}$$

We choose α to maximize $\alpha - 2^{-k}(\alpha + (1 + L\alpha)^k)$; simple calculus shows that this occurs at

$$\alpha = \frac{\left(\frac{2^k-1}{kL}\right)^{\frac{1}{k-1}} - 1}{L};$$

and that $\alpha \geq 0$ for $L \leq \frac{2^k-1}{k}$.

With this choice of α, then simple algebraic manipulations show that (3) is satisfied for

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

Thus, if $L \leq \frac{(2^k-1)(1-1/k)^{k-1}}{k}$ and $L \leq \frac{2^k-1}{k}$, then Theorem 1.8 is satisfied. The second condition $L \leq \frac{2^k-1}{k}$ can be easily seen to be redundant, leading to the given bounds. $\square$

Let us define $F_{\text{LLL}}(k) = \lfloor \frac{2^k}{ek} - 1/k \rfloor$ and $F_{\text{MT}} = \lfloor \frac{(2^k-1)(1-1/k)^{k-1}}{k} \rfloor$ to be the bounds on $f'(k)$ which are provable from the symmetric LLLL (Theorem 2.3) and the Moser-Tardos lopsidependency criterion (Theorem 2.4). We observe that $F_{\text{MT}}(k) \geq F_{\text{LLL}}(k)$ for all integers $k \geq 1$. Furthermore,

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

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

3 Constructing the extremal formula Φ

For any k-SAT instance, we have the natural problem space Ω which assigns $X_i = T$ with probability $1/2$, independently for all variables i, and we have the natural collection of bad-events corresponding to the clauses. For this probability space, we have $P(B) = p = 2^{-k}$ for all bad-events.

We will next construct a k-SAT problem instance in which $R_0(\Phi, i), R_1(\Phi, i) \leq L$, in which the Shearer criterion is *violated* for this B, Ω. However, this value of L will be less than $F_{\mathrm{MT}}(k)$, implying that the lopsidependent MT criterion can be stronger than Shearer's criterion.

For a given integer $L \geq 1$, we construct Φ recursively. Initially, Φ_0 contains no clauses. At stage i of the process, we modify Φ_{i-1} to produce a new formula Φ_i, by adding $2L-2$ clauses containing variable i; exactly $L-1$ clauses in which i appears positively and exactly $L-1$ clauses in which i appears negatively. All the 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. When we form Φ_i, each of the new variables (other than variable i) appears in exactly one new clause. We refer to the process of adding $2L-2$ clauses containing variable i as *expanding* variable i. We define A_i to be the collection of new clauses added during the expansion of i (that is, the clauses in Φ_i but not Φ_{i-1})

Proposition 3.1. *For any $r \in \mathbf{N}$ and every $i \in \mathbf{N}$, we have*

$$R_0(\Phi_r, i) \leq L \qquad R_1(\Phi_r, i) \leq L-1$$

Proof. There is only one case in which variable $j < i$ is expanded and variable i appears, and in that case it may produce at most one positive occurrence of variable i. Otherwise, the only occurrences of variable i appear when expanding variable i; this adds $L-1$ positive occurrences of i and $L-1$ negative occurrences of i. $\square$

Let us define G_r be the canonical lopsidependency graph corresponding to the bad-events for the formula Φ_r. Although the graphs G_r are complicated, we will show that they contains a relatively simple and regular type of subgraph. We will actually show that Shearer's criterion is violated for this subgraph; as shown in [19], this implies that Shearer's criterion is violated for the overall graph G_r.

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 this $K_{L-1,L-1}$, which is labeled as the *root* of H_j. We define the graph family H_j recursively. First, H_0 is the null graph (the graph on 0 vertices). To form H_{j+1}, we start by taking a new copy of $K_{L-1,L-1}$, which we will designate as the root of H_{j+1}. Then, 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.

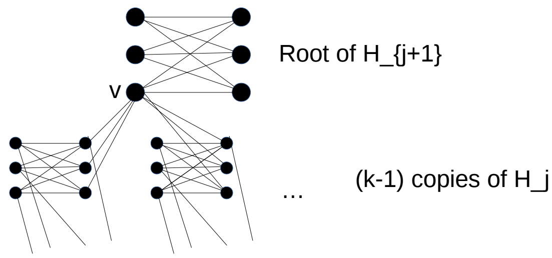

Figure 1: The construction of H_{j+1} from H_j. We have only shown here two copies of H_j corresponding to a single vertex v in the root of H_{j+1}. There are $k-1$ copies of H_j for each vertex in the root of H_{j+1} (a total of $2(L-1)(k-1)$ copies of H_j).

Proposition 3.2. *Let $j > 0$ be any fixed integer. There is some r sufficiently large (which may depend on j) such that G_r contains a copy of H_j.*

Proof. We define a tree structure $\mathcal{T}$ on the variables of Φ_r: variable i is a parent of variable j if variable j appears in Φ_i but not Φ_{i-1}, that is, variable j was introduced during the expansion of variable i. For any variable i, let $\mathcal{T}_i$ denote the subtree of $\mathcal{T}$ rooted at i.

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

When $j = 0$ this is vacuously true. To show the inductive step, consider some variable i. Variable i has $(2L-2)(k-1)$ children in $\mathcal{T}$, which we denote by C.

By inductive hypothesis, for each $i' \in C$, $G_{R(i',j-1)}[\mathcal{T}_{i'}]$ contains a copy of H_{j-1} where the root of this copy of H_{j-1} corresponds to $A_{i'}$.

We now claim that the choice $r = R(i, j) = i + \max_{i' \in C} R(i', j-1)$ satisfies the stated conditions. Let $\hat{G} = G_T[\mathcal{T}_i]$. The graph $\hat{G}$ contains the disjoint graphs $G_T[\mathcal{T}_{i'}]$ for each $i' \in C$. For each such $i' \in C$, let $J_{i'}$ denote this copy of H_{j-1} in $G_r[\mathcal{T}_{i'}] \subseteq \hat{G}$. Since the graphs $G_r[\mathcal{T}_{i'}]$ are disjoint, so are the graphs $J_{i'}$

In the graph $\hat{G}$, the clauses of A_i in which i appears positively are lopsidependent with those clauses in which i appears negatively. Thus, there is a copy in $\hat{G}$ of $K_{L-1,L-1}$ corresponding to A_i; we denote this copy by J.

Consider some clause $\phi \in C$ (which corresponds to a vertex of J). This clause contains the variable i and $k-1$ other variables in C. Consider one such variable i'. The root of $J_{i'}$ corresponds to the clauses $A_{i'}$. Note that ϕ is the only clause of C 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 $\phi \in J$, we see that the resulting graph is precisely H_j.

Furthermore, the root of this H_j is J, whose vertices correspond to the clauses in A_i. Finally, this graph structure all appears in $\hat{G} = G_r[\mathcal{T}_i]$. The induction thus holds. $\square$

4 Computing the Shearer criterion for H_j

We now discuss how to compute the Shearer criterion for the family of graphs H_j. We will show that, for j sufficiently large, $Q(H_j, \emptyset, \vec{p}) \leq 0$.

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

Proposition 4.1. *If V are 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 4.2. *Suppose $X \subseteq V$. Then*

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

We will need to work also with a family of graphs H'_j, which are slight variant of the graphs H_j. We define a graph H'_{j+1} 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.

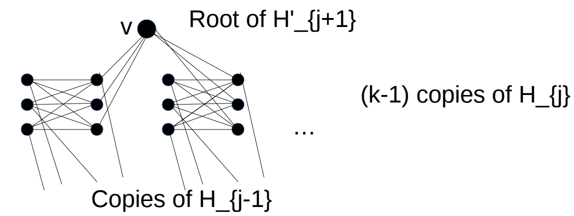

Figure 2: The construction of H'_{j+1} from H_j.

Proposition 4.3. *Let us define*

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

Then s, r satisfy the mutual recurrence relations

$$\begin{aligned} r_j &= s_{j-1}^{k-1} - p r_{j-1}^{(k-1)(L-1)} s_{j-2}^{(k-1)^2(L-1)} \\ s_j &= 2 r_j^{(L-1)} s_{j-1}^{(k-1)(L-1)} - s_{j-1}^{(k-1)(2L-2)} \\ s_0 &= 1, r_0 = 1 \end{aligned}$$

Proof. We will first show the bound s_j. In any independent 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 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

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

Next, let us consider the bound for r_j. We apply Proposition 4.2, taking X as the singleton root node. In this case, U is either the empty set, or U is the root node. In the former case, the residual graph $G[V - U - N(U)]$ consists of $k-1$ independent copies of H_{j-1}.

In the latter case, we have removed the root node v of H'_j and its neighbors; let J denote of the copies of H_{j-1} to which v was connected (there are $k-1$ such copies). In J, all the vertices in the left half of the root are now disconnected and isolated, leaving $L-1$ disconnected copies of H'_{j-1}. In addition, all the vertices in the right-half of the root of J are removed; each of those was connected to $k-1$ copies of H_{j-2}, which now become isolated copies. In total, J consists of $L-1$ copies of H'_{j-1} and $(k-1)(L-1)$ copies of H_{j-2}. As there are $k-1$ isomorphic copies of J, then $H'_j - v - N(v)$ consists of $(k-1)(L-1)$ copies of H'_{j-1} and $(k-1)^2(L-1)$ copies of H_{j-2}. See Figure 3.

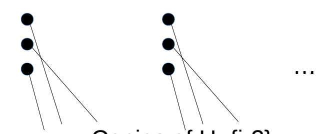

Figure 3: Removing the root node from H'_j

Summing the contributions of these two terms according to Proposition 4.2 gives

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

□

Proposition 4.4. *Define the function* $g : [0,1] \to \mathbf{R}$ *by*

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

Suppose that the Shearer condition is satisfied for all $G_r, r \geq 0$. *Then there is some* $a \in (2^{\frac{-2}{2L-2}}, 1]$ *satisfying* $g(a) = a$.

Proof. Define

$$a_j = \frac{r_j}{s_{j-1}^{k-1}}, b_j = \frac{s_j}{s_{j-1}^{(k-1)(2L-2)}}$$

With this definition, we have

$$b_j = \frac{2 r_j^{(L-1)} s_{j-1}^{(k-1)(L-1)} - s_{j-1}^{(k-1)(2L-2)}}{s_{j-1}^{(k-1)(2L-2)}} = \frac{2 r_j^{(L-1)}}{s_{j-1}^{(k-1)(L-1)}} - 1 = 2 a_j^{(L-1)} - 1$$

and we can obtain a pure first-order recurrence on the sequence a_j:

$$a_j = \frac{s_{j-1}^{k-1} - p r_{j-1}^{(k-1)(L-1)} s_{j-2}^{(k-1)^2(L-1)}}{s_{j-1}^{k-1}} = 1 - p \frac{r_{j-1}^{(k-1)(L-1)}}{s_{j-2}^{(k-1)^2(L-1)}} \times \frac{s_{j-2}^{(k-1)^2(2L-2)}}{s_{j-1}^{k-1}} = 1 - \frac{p a_{j-1}^{(k-1)(L-1)}}{b_{j-1}^{k-1}}$$

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

One may verify also that $a_0 = 1$.

Now suppose that for some $j \geq 1$ we have $a_j \leq 2^{\frac{-2}{2L-2}}$. In this case, we have $b_j \leq 0$ and hence $\frac{s_j}{s_{j-1}^{(k-1)(2L-2)}} \leq 0$. This implies that either $s_j \leq 0$ or $s_{j-1} \leq 0$. Since every bad-event has probability $p = 2^{-k} > 0$, by Theorem 1.4 so the Shearer condition is violated for H_j or H_{j-1}. This implies that the Shearer condition is violated for G_r for r sufficiently large.

Next, suppose that it holds that $g(a) < a$ for all $a \in (2^{-\frac{-2}{2L-2}}, 1]$. This implies that the sequence a_j is decreasing for $j \in \mathbf{N}$. As the sequence a_j also satisfies $a_j \geq 2^{\frac{-2}{2L-2}}$, it must converge to some limit point a. But, by continuity, this limit point must be a fixed point of the functional iteration, i.e. $g(a) = a$, which is a contradiction.

So we know that $g(a) \geq a$ for some $a \in (2^{-\frac{-2}{2L-2}}, 1]$. But also note 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. $\square$

Proposition 4.5. *Suppose*

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

for all $t \in [0, 2]$. Then the Shearer condition is violated on G_r, for r sufficiently large.

Proof. Suppose that the Shearer condition is satisfied for G_r for all integers $r \geq 0$. By Proposition 4.4, the function g 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:

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

Setting $t = 2^{k/(1-k)}(1-a)^{1/(1-k)}$, we see that $t \in [0, 2]$ and

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

This contradicts our hypothesis. $\square$

For any $k \geq 1$, let us define $\tilde{F}_{\text{Shearer}}(k)$ to be the largest integer L with the property that $L \leq 1 - \frac{\ln(2-t)}{\ln\left(1-2^{-k}t^{1-k}\right)}$ for some $t \in [0, 2]$. Equivalently,

$$\tilde{F}_{\text{Shearer}}(k) = \left\lfloor \max_{t \in [0,2]} 1 - \frac{\ln(2-t)}{\ln\left(1 - 2^{-k} t^{1-k}\right)} \right\rfloor$$

Thus $\tilde{F}_{\text{Shearer}}$ is an upper bound on the value of $f'(k)$ that can be shown using the LLL or any variant of it. To illustrate, we list the values of $F_{\text{LLL}}, \tilde{F}_{\text{Shearer}}$, and F_{MT} for a few small values of k.

k	F_{LLL}	$\tilde{F}_{\text{Shearer}}$	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

We observe that $\tilde{F}_{\text{Shearer}} \geq F_{\text{LLL}}$ 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. 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 4.6. $\tilde{F}_{\mathit{Shearer}} = \frac{2^k}{ek} + \Theta(\frac{2^k}{k^3})$

Proof. We first show the upper bound. Let $L = \tilde{F}_{\text{Shearer}}(k)$, so that

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

for some $t \in [0, 2]$.

Using the bound $-\ln(1-x) \geq x$ for $x \geq 0$, we have:

$$L \leq 1 + t^{k-1} 2^k \ln(2-t) \tag{4}$$

Now observe that $\ln(2-t)$ is a concave-down function of t for $t \in [0, 2]$. Hence, for any $t_0 \in [0, 2]$ we have the bound

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

for all $t \in [0, 2]$.

Substituting this bound into (4) gives

$$L \leq 1 + \frac{\Big(2(1-1/k)(t_0 + (2-t_0)\ln(2-t_0))\Big)^k}{(2-t_0)(k-1)} \tag{5}$$

for any $t_0 \in [0, 2]$.

Set $t_0 = 1 - 1/k$ and after some simple calculus we obtain:

$$L \leq \frac{2^k}{ek} + O(\frac{2^k}{k^3})$$

Next, we observe that $\tilde{F}_{\text{Shearer}} \geq \lfloor 1 - \frac{\ln(2-t_0)}{\ln\left(1-2^{-k} t_0^{1-k}\right)} \rfloor \geq -\frac{\ln(2-t_0)}{\ln\left(1-2^{-k} t_0^{1-k}\right)}$. Again, with $t_0 = 1 - 1/k$, simple calculus show that this is at least $\frac{2^k}{ek} + \Omega(\frac{2^k}{k^3})$. □

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

5 Acknowledgments

Thanks to Aravind Srinivasan for many helpful comments and suggestions.

References

[1] Albert, M., Frieze, A., Reed, B.: Multicoloured Hamilton Cycles. The Electronic Journal of Combinatorics 2-1, 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, pp. 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, pp. 607-726 (1975)

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

[5] Gebauer, H.: Disproof of the neighborhood conjecture with implications to SAT. Combinatorica 32-5, pp. 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 of the Lovász local lemma via resampling oracles. Proceedings of the 56th annual IEEE Sympsoium on Foundations of Computer Science (FOCS) (2015).

[9] He, K., Li, L., Liu, X., Wang, Y., Xia, M.: Variable version Lovász local lemma: beyond Shearer's bound. arXiv:1709.05143 (2017)

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

[11] Kolipaka, K., Szegedy, M.: Moser and Tardos meet Lovász. Symposium on Theory of Computing, pp. 235-244 (2011)

[12] Kolipaka, K., Szegedy, M., Xu, Y.: A sharper local lemma with improved applications. In "Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques" LNCS 7408, pp. 603-614 (2012)

[13] 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, pp. 203-210 (1993)

[14] Lu, L., Székély, L.: Using Lovász Local Lemma in the space of random injections. The Electronic Journal of Combinatorics 13-R63 (2007)

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

[16] McDiarmid, C.: Hypergraph coloring and the Lovász Local Lemma. Journal of Discrete Mathematics 167/168, pp. 481-486 (1995)

[17] Moser, R., Tardos, G.: A constructive proof of the general Lovász Local Lemma. Journal of the ACM 57-2, pp. 11:1-11:15 (2010)

[18] Pegden, W.: An extension of the Moser-Tardos algorithmic local lemma. SIAM Journal of Discrete Math 28-2, pp. 911-917 (2014)

[19] Shearer, J. B.: On a problem of Spencer. Combinatorica 5, pp. 241-245 (1985)

[20] Spencer, J.: Asymptotic lower bounds for Ramsey functions. Discrete Mathematics 20, pp. 69-76 (1977)

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