6.S891 Lecture 20: The Sampling Lovász Local Lemma

Kuikui Liu

November 21, 2023

In this lecture, we study the counting analog of the quintessential algorithmic problem in theoretical computer science: SAT. Recall that an instance of SAT is a Boolean formula $\Phi = (\mathcal{V}, \mathcal{C})$ in *conjunctive normal form (*CNF*)*, i.e. it is given by a collection of Boolean-valued variables $\mathcal{V}$ and an "AND-of-ORs" $\bigwedge_{C \in \mathcal{C}} C$ where each *clause* C is an OR of a subset of the variables (possibly negated). Our goal is to design an algorithm which approximates the number of solutions to a given CNF-formula in the regime where we can efficiently find a satisfying solution via the (algorithmic) Lovász Local Lemma. Building on the techniques in the previous lecture, will use a disagreement percolation argument in the analysis. Note that there is a well-known FPRAS for counting solutions to DNF-*formulas* (i.e. "OR-of-ANDs") based on rejection sampling [KLM89].

1 The Local Lemma Regime

We begin by restating the celebrated Lovász Local Lemma (LLL), which we discussed previously in the context of the cluster expansion and zero-freeness for the (multivariate) independence polynomial. We specialize the statement somewhat to make it more "compatible" with our notation for SAT. Let $X_1, \dots, X_n$ denote a sequence of independent $\{\mathsf{T}, \mathsf{F}\}$-valued random variables; write $\nu = \mathsf{Law}(X_1, \dots, X_n)$ for the associated product measure on $\{\mathsf{T}, \mathsf{F}\}^n$. For an event $B \subseteq \{\mathsf{T}, \mathsf{F}\}^n$, we write $\mathcal{V}(B) \subseteq [n]$ for the minimal subset of $\{X_i\}_{i=1}^n$ whose values completely determine whether or not B occurs. Now let $\{B_j\}_{j=1}^m$ be a collection of events. For any event A, we write $N(A) = \{j : B_j \neq A, \mathcal{V}(A) \cap \mathcal{V}(B_j) \neq \emptyset\}$ for the subset of events $\{B_j\}_{j=1}^m$ which intersect A in constituent variables. This naturally induces a graph on $[m]$ itself called a *dependency graph* for the events $\{B_j\}_{j=1}^m$; in particular, each B_i is mutually independent of $\{B_j\}_{j=1}^m \setminus N(B_i)$.

Theorem 1.1 (Asymmetric LLL)**.** *In the aforementioned setup, suppose there exists* $\boldsymbol{p} : [m] \to (0, 1)$ *such that*

$$\Pr_{\nu}[B_j] \leq \boldsymbol{p}_j \prod_{i \in N(B_j)} (1 - \boldsymbol{p}_i), \qquad \forall j \in [m].$$

Then

$$\Pr_{\nu}\left[\bigwedge_{j=1}^m \overline{B}_j\right] \geq \prod_{j=1}^m (1 - \boldsymbol{p}_j) > 0.$$

Furthemore, for every other event A*,*

$$\Pr_{\nu}\left[A \mid \bigwedge_{j=1}^m \overline{B}_j\right] \leq \Pr_{\nu}[A] \cdot \prod_{j \in N(A)} (1 - \boldsymbol{p}_j)^{-1}. \tag{1}$$

Now let us consider SAT. For simplicity, we will assume throughout this lecture that our formulas Φ are k*-uniform*, i.e. all clauses have exactly k variables, and have maximum degree d, i.e. all variables participate in at most d distinct clauses; to shorten notation, we call these *"*(k, d)*-formulas"*. In Section 4, we briefly comment on recent progress for random k-uniform CNF-formulas, which do not have bounded degrees for the variables. We also write $n = |\mathcal{V}|$ for the total number of variables, and $m = |\mathcal{C}|$ for the total number of clauses.

One can form a natural bipartite graph with variables $\mathcal{V}$ on one side, clauses $\mathcal{C}$ on the other, and an edge between a clause and each of its constituent variables; call this bipartite graph H_Φ. By

assumption, all variable vertices have degree $\leq d$, and all clause vertices have degree k. The graph $H^2_\Phi[\mathcal{V}]$ on $\mathcal{V}$ connects two variables if and only if they participate in a common clause. Similarly, the graph $H^2_\Phi[\mathcal{C}]$ on $\mathcal{C}$ connects two clauses if and only if they intersect; this is the dependency graph for the LLL defined above. Note that by assumption, both graphs have maximum degree $\leq k(d+1)$.

A classical application of the LLL (Theorem 1.1) says that every (k,d)-formula Φ has at least one satisfying assignment, so long as $e(d+1) \leq 2^k$: Take $\nu = \mathsf{Unif}\{\mathsf{T},\mathsf{F}\}^{\mathcal{V}}$, the events $\{B_C\}_{C\in\mathcal{C}}$ to correspond to clauses C *not* being satisfied, and $\boldsymbol{p}(C) = \Theta\left(\frac{1}{dk}\right)$, noting that $\Pr_\nu[B_C] = 2^{-k}$; for a more complete argument, see Theorem 2.2 and its proof. Moreover, such a satisfying assignment can be found efficiently by the following simple stochastic local search algorithm:

1. Sample $(X_1,\ldots,X_n) \sim \nu$; recall that the coordinates $X_1,\ldots,X_n$ are independent.
2. While there exists a violation, i.e. some clause C is not satisfied, resample all variables in C independently.

We will refer to this as the *Moser–Tardos algorithm*.[1] The following seminal result of Moser and Tardos gives rigorous guarantees for this algorithm [MT10].

Theorem 1.2 (Informal; [MT10])**.** *Suppose the conditions of Theorem 1.1 hold w.r.t. some $\boldsymbol{p} : [m] \to (0,1)$. Then there is a randomized algorithm which outputs values $X_1,\ldots,X_n$ such that $\wedge_{j=1}^m \overline{B}_j$ holds upon termination. Furthermore, if $\mathcal{V}(B_j) \leq O(1)$ for all $j \in [m]$, then the expected running time of this algorithm is upper bounded by $O\left(\sum_{j=1}^m \frac{\boldsymbol{p}_j}{1-\boldsymbol{p}_j}\right)$.*

Our goal is to use the LLL (Theorem 1.1), and our ability to find satisfying assignments Theorem 1.2, to design algorithms for (approximately) sampling uniformly random satisfying assignments. Throughout, let $\mu = \mu_\Phi$ denote the uniform measure over all satisfying assignments to the CNF-formula Φ. The main result of this lecture is the following.

Theorem 1.3 (Informal; [Moi19; Fen+21])**.** *There are constants $0 < c < 1/2$ and $c' > 0$ such that if $2^{ck} \geq c' \cdot d^5 k^5$, then there exists an FPRAS for counting solutions to any (k,d)-formula.*

Remark 1. [Moi19] gave an FPTAS for the problem, with exponent scaling polynomially in d, k. Allowing randomization, [Fen+21] gave an algorithm whose running time scales roughly as $d^2k^3n^{1+\zeta}$, where ζ can be made arbitrarily small at the expense of degrading the constant c'. We emphasize that the statement above is not the best known.

The Monotone Setting We say a CNF-formula is *monotone* if no literal is negated in any clause. Such formulas can always trivially be satisfied by setting all variables to T, regardless of what k, d are. Combinatorially, satisfying assignments to monotone CNF-formulas are equivalent to *hypergraph independent sets*, where we view $(\mathcal{V},\mathcal{C})$ as a hypergraph, and an independent set is a subset of vertices which does not fully contain any hyperedge.[2] In the monotone setting, [HSZ19] established rapid mixing of Glauber dynamics all the way up to $d \leq c \cdot 2^{k/2}$ for some constant $c > 0$, improving significantly upon previous work of [BDK08].

Hardness and Failure of Strong Spatial Mixing Even in the monotone setting, [Bez+19] proved that unless NP = RP, there is no FPRAS for counting satisfying assignments if $d \geq c' \cdot 2^{k/2}$ for some other constant $c' \approx 5$. This was achieved via a reduction to approximately computing the partition function of the hardcore model beyond uniqueness, which is NP-hard [Sly10; SS14]. This reduction was originally due to [BDK08], although at the time, the sharpest hardness results for the hardcore model were not known. An interesting aspect of this hardness is that its onset occurs *below* the uniqueness threshold for hypergraph independent sets on the k-uniform d-regular *hypertree*, which is roughly $d \leq \frac{2^k}{2k}$; note the exponential scaling in k, as opposed to $k/2$. It is an open problem to close the gap in parameters between the current best algorithmic results, and the current best NP-hardness results.

In the nonmonotone setting, additional challenges arise, as naïve Glauber dynamics can easily become disconnected. As a simple example, consider the formula

$$(x_1 \vee \neg x_2) \wedge (x_2 \vee \neg x_3) \wedge (x_3 \vee \neg x_1).$$

[1] This algorithm also sometimes goes under the name "FIX".
[2] Assigning a variable T is equivalent to pinning that vertex to be out of the independent set.

The only solutions are $(\mathsf{T},\mathsf{T},\mathsf{T})$ and $(\mathsf{F},\mathsf{F},\mathsf{F})$. Similarly, strong notions of correlation decay like strong spatial mixing fail dramatically, even in the monotone case. Indeed, regardless of how large k is, strong spatial mixing already fails for all $d \geq 6$ just by reducing hypergraph independent set in the infinite k-uniform d-regular hypertree to the usual notion of independent set in the infinite $(d-1)$-ary tree.

2 Sampling via Dynamics on Projected Spaces

Let Φ be some fixed (k,d)-formula. Recall $\mu = \mu_\Phi$ denotes the corresponding uniform measure over satisfying assignments. For any subset of variables $\mathcal{A} \subseteq \mathcal{V}$, let $\mu_\mathcal{A}$ denote the induced marginal distribution over variables in $\mathcal{A}$, i.e.

$$\mu_\mathcal{A}(\tau) \propto \sum_{\substack{\sigma:\mathcal{V}\to\{\mathsf{T},\mathsf{F}\} \\ \sigma(x)=\tau(x),\forall x\in\mathcal{A}}} \mu(\sigma).$$

Theme 2.1. *To overcome disconnectedness of Glauber dynamics, "project" the distribution onto a smaller state space. Approximately sample from this projected distribution (e.g. via Markov chains), and randomly lift the resulting sample back into the original state space.*

At a high level, for CNF-formulas, our "projection" will be the marginal distribution over a carefully chosen subset of variables $\mathcal{M} \subseteq \mathcal{V}$. We will sample a partial assignment over this subset of variables via Glauber dynamics, and then complete this partial assignment into a full satisfying assignment by sampling from the conditional measure. The latter will be achieved by leveraging sparsity of the formula and standard tools based on *shattering*. In order for this scheme to work, we need to choose $\mathcal{M}$ such that Glauber dynamics for $\mu_\mathcal{M}$ mixes quickly.

For this, we again appeal to Theorem 1.1. Notably, Eq. (1) gives us nice quantitative control on the marginals of μ. The following result will suggest a good set of variables $\mathcal{M}$.

Theorem 2.2 (Local Uniformity; [Moi19; Fen+21])**.** *Let $\Phi = (\mathcal{V},\mathcal{C})$ be a CNF-formula such that every variable participates in at most d clauses, and every clause contains between k_1 and k_2 variables. For every $s \geq k_2$, if $2^{k_1} \geq 2eds$, then Φ has at least one satisfying assignment (so that μ is well-defined), and for every variable x,*

$$\max\left\{\Pr_{\sigma\sim\mu}[\sigma(x)=\mathsf{T}], \Pr_{\sigma\sim\mu}[\sigma(x)=\mathsf{F}]\right\} \leq \frac{1}{2}\exp\left(\frac{1}{s}\right). \tag{2}$$

Remark 2. Upon inspecting the proof, it is clear that we can obtain analogous worst-case upper bounds on the joint marginal distribution of any $\mathcal{A} \subseteq \mathcal{V}$ as

$$\Pr_{\sigma\sim\mu}[\sigma\mid_\mathcal{A}=\tau] \leq \underbrace{\frac{1}{2^{|\mathcal{A}|}}}_{=\nu_\mathcal{A}(\tau)} \cdot \underbrace{\exp\left(\frac{|\mathcal{A}|}{s}\right)}_{\text{error}}, \qquad \forall \tau : \mathcal{A} \to \{\mathsf{T},\mathsf{F}\}.$$

Proof. Let ν be the uniform measure over $\{\mathsf{T},\mathsf{F}\}^\mathcal{V}$. For each clause $C \in \mathcal{C}$, define a "bad event" B_C corresponding to C not be satisfied by a random assignment $\sigma \sim \nu$. Taking $\boldsymbol{p}(B_C) = \frac{1}{2ds}$ and using the fact that any clause can have at most $k_2(d-1)$ neighboring clauses, we have

$$\begin{aligned}
\boldsymbol{p}(B_C)\prod_{C'\cap C\neq\emptyset}(1-\boldsymbol{p}(B_{C'})) &= \frac{1}{2ds}\left(1-\frac{1}{2ds}\right)^{k_2(d-1)} \\
&\geq \frac{1}{2ds}\left(1-\frac{1}{2ds}\right)^{2ds-1} && \text{(Using } s\geq k_2\text{)} \\
&\geq \frac{1}{2eds} && \text{(Using } \left(1-\tfrac{1}{z}\right)^{z-1}\geq\tfrac{1}{e} \text{ for all } z>1\text{)} \\
&\geq 2^{-k_1} && \text{(Assumption)} \\
&\geq \Pr_\nu[B_C]. && \text{(C has at least k_1 variables)}
\end{aligned}$$

For the marginal bound, we plug into Eq. (1) to obtain

$$\begin{aligned}
\Pr_{\sigma\sim\mu}[\sigma(x)=\mathsf{T}] &= \Pr_{\sigma\sim\nu}[\sigma(x)=\mathsf{T}\mid \sigma \text{ satisfies } \Phi] \\
&\le \Pr_{\sigma\sim\nu}[\sigma(x)=\mathsf{T}]\cdot\prod_{C\ni x}(1-\boldsymbol{p}(B_C))^{-1} && \text{(Eq. (1) and definition of } N(\{x\leftarrow\mathsf{T}\})) \\
&\le \frac{1}{2}\left(1-\frac{1}{2ds}\right)^{-d} && \text{(Degree} \le d) \\
&\le \frac{1}{2}\exp\left(\frac{1}{s}\right).
\end{aligned}$$

The same argument applies to $\Pr_{\sigma\sim\mu}[\sigma(x)=\mathsf{T}]$ and so we're done. □

We will also make use of the following convenient closure property for μ. Its proof is straightforward.

Definition 1 (Simplification)**.** *Let $\Phi=(\mathcal{V},\mathcal{C})$ be a* CNF*-formula. For a subset of variables $\mathcal{A}\subseteq\mathcal{V}$, and a partial assignment $\tau:\mathcal{A}\to\{\mathsf{T},\mathsf{F}\}$, define the* simplification *Φ^τ to be the formula obtained from Φ by deleting all clauses in Φ which are already satisfied by τ, and deleting all variables in $\mathcal{A}$ from the remaining clauses. Note that this definition only makes sense if τ is consistent with some global satisfying assignment $\sigma:\mathcal{V}\to\{\mathsf{T},\mathsf{F}\}$.*

Lemma 2.3 (Conditionals & Simplification)**.** *For every* CNF*-formula $\Phi=(\mathcal{V},\mathcal{C})$ and every extendable partial assignment $\tau:\mathcal{A}\to\{\mathsf{T},\mathsf{F}\}$, we have $\mu_\Phi^\tau=\mu_{\Phi^\tau}\otimes\delta_\tau$. In other words, the conditional measure μ_Φ^τ is uniform over satisfying assignments to the simplification Φ^τ, adjoined with τ.*

2.1 Moitra's Marking Scheme

Theorem 2.2 tells us that as long as every clause contains enough unpinned variables, then at least w.r.t. local statistics, the measure μ looks approximately uniform. Intuitively, this suggests something like "high-temperature behavior" or spectral independence. However, we will definitely need control over the conditionals of $\mu_\mathcal{M}$ as well. Thus, roughly speaking, we need $\mathcal{M}\subseteq\mathcal{V}$ to satisfy two competing properties:

- $\mathcal{M}$ is sufficiently small so that even if we pin $\mathcal{M}$ arbitrarily, every clause still has many unpinned variables, allowing us to invoke Theorem 2.2.
- At the same time, we want $\mathcal{M}$ to be sufficiently large so that once we pin the variables of $\mathcal{M}$ according to a random partial assignment τ drawn from $\mu_\mathcal{M}$, we shatter the formula. This ensures we can efficiently sample a full satisfying assignment from the conditional measure μ^τ.

The precise properties we need from $\mathcal{M}$ is formalized as follows.

Definition 2 (Marking; [Moi19])**.** *Let Φ be a (k,d)-formula, and let $0<\alpha<1/2$ be a constant. We call a subset of variables $\mathcal{M}\subseteq\mathcal{V}$ a* valid α-marking w.r.t. *Φ if every clause C of Φ has at least αk marked variables and at least αk unmarked variables (i.e. $|C\cap\mathcal{M}|, |C\cap(\mathcal{V}\setminus\mathcal{M})|\ge\alpha k$).*

We first show that such an α-marking can be found efficiently. Throughout, one can take $\alpha=\frac{1}{3}$ for simplicity.

Theorem 2.4 ([Fen+21])**.** *Let Φ be a (k,d)-formula, and let $0<\alpha<1/2$ be a constant. If $2^k\ge\Omega(dk)^{O\left(1/(1-2\alpha)^2\right)}$, then for every $\delta>0$, there is a randomized algorithm which successfully outputs a valid α-marking $\mathcal{M}\subseteq\mathcal{V}$ with probability at least $1-\delta$ and runs in time $O\left(dkn\log\frac{1}{\delta}\right)$.*

Proof Sketch. We again use the (algorithmic) LLL. For each variable, we add it to $\mathcal{M}$ independently with probability $1/2$. For each clause $C\in\mathcal{C}$, let B_C denote the complement of the event that C has at least αk marked and unmarked variables. We can control each individual $\Pr_\nu[B_C]$ via Chernoff bounds. Applying Theorem 1.1 with $\boldsymbol{p}(B_C)=\Theta\left(\frac{1}{dk}\right)$ again allows one to show that an α-marking exists. Furthermore, this α-marking can be found efficiently using the Moser–Tardos algorithm; see Theorem 1.2. □

2.2 Glauber Dynamics for $\mu_{\mathcal{M}}$

Now that we have determined a suitable subset of variables $\mathcal{M} \subseteq \mathcal{V}$, we need to draw a random partial assignment $\tau \sim \mu_{\mathcal{M}}$. Towards this, we show that Glauber dynamics mixes quickly.

Theorem 2.5 ([Fen+21]). *Let $\Phi = (\mathcal{V}, \mathcal{C})$ be a (k,d)-formula, and let $\mathcal{M} \subseteq \mathcal{V}$ be a valid α-marking for some constant $0 < \alpha < 1/2$. If $2^{\alpha k} \geq \Omega(dk)^5$, then Glauber dynamics for $\mu_{\mathcal{M}}$ mixes in $O(n \log n)$-steps.*

We sketch the proof in Section 3 by using a disagreement percolation argument to establish spectral independence for $\mu_{\mathcal{M}}$. Before we do this, however, let us emphasize that Theorem 2.5 is only part of the story. Glauber dynamics for $\mu_{\mathcal{M}}$ requires access to the conditional marginals of $\mu_{\mathcal{M}}$, which are #P-hard to compute. Hence, it is not clear that we can even implement a single step of this Glauber dynamics. It turns out one can implement an approximate version of this Glauber dynamics, using ideas from Section 2.3. The point is that with high probability, these marginals can be approximated to arbitrary accuracy, precisely for the same reason that we can draw a full satisfying assignment from the conditional measure μ^τ. The latter is actually enough for implementing Glauber dynamics; we do not need to actually approximate these conditional marginals.

2.3 Completing an Assignment: Shattering and Rejection Sampling

Now assume that we have a random partial assignment on a valid α-marking $\mathcal{M} \subseteq \mathcal{V}$. We show that with high probability, we can efficiently sample from μ^τ. This is established via two lemmas.

Definition 3 (Shattering). *Let $\Phi = (\mathcal{V}, \mathcal{C})$ be a (k,d)-formula. We say a partial assignment $\tau : \mathcal{M} \to \{\mathsf{T}, \mathsf{F}\}$ on a subset of variables $\mathcal{M} \subseteq \mathcal{V}$ (e.g. a valid α-marking) δ-*shatters Φ *if all connected components of clauses in the* simplification Φ^τ *have size $\lesssim dk \log \frac{n}{\delta}$.*

Lemma 2.6 ([Fen+21]). *Let $\Phi = (\mathcal{V}, \mathcal{C})$ be a (k,d)-formula, and let $\mathcal{M} \subseteq \mathcal{V}$ be a valid α-marking for some constant $0 < \alpha < 1/2$. If $2^{\alpha k} \geq \Omega(dk)^3$, then*

$$\Pr_{\tau \sim \mu_{\mathcal{M}}} [\tau \text{ } \delta\text{-shatters } \Phi] \geq 1 - \delta/n.$$

The proof of this lemma uses ingredients similar to the proof we will give for Theorem 3.1 below. Hence, we omit it for brevity; see [Fen+21].

Lemma 2.7 ([Fen+21]). *Let $\Phi = (\mathcal{V}, \mathcal{C})$ be a (k,d)-formula, and suppose some partial assignment $\tau : \mathcal{M} \to \{\mathsf{T}, \mathsf{F}\}$ δ-shatters Φ. If $2^{\alpha k} \geq \Omega\left(\frac{dk}{\theta}\right)$ for some (small) constant $\theta > 0$, then for every $\epsilon > 0$, we can draw a random $\sigma : \mathcal{V} \to \{\mathsf{T}, \mathsf{F}\}$ satisfying $\|\mathsf{Law}(\sigma) - \mu^\tau\|_{\mathsf{TV}} \leq \epsilon$ in time $\tilde{O}(n/\epsilon)^{1+\theta}$.*

Proof Sketch. By Lemma 2.3, suffices to draw a uniformly random satisfying assignment τ' to the simplification Φ^τ, and then take $\sigma = \tau \sqcup \tau'$. To do this, we just need to sample an assignment for each maximal connected component of Φ^τ. Since τ shatters Φ, these components all have logarithmic size, and so by the LLL (Theorem 1.1), a uniformly random $\{\mathsf{T}, \mathsf{F}\}$-assignment to the variables of each component satisfies each with probability at least $O(\delta/n)^\theta$. Hence, applying a (small) polynomial number of rejection sampling steps gives the algorithm. □

2.4 The Full Algorithm

Given what we have outlined so far, we now give a sketch of the algorithm.

1. Use Theorem 2.4 to find a valid α-marking $\mathcal{M} \subseteq \mathcal{V}$.
2. Run Glauber dynamics on $\mu_{\mathcal{M}}$ for $O(n \log n)$ iterations with initial distribution $\mathsf{Unif}\{\mathsf{T}, \mathsf{F}\}^{\mathcal{M}}$, using the algorithm in Lemma 2.7 to (approximately) implement each step. This yields a random partial configuration $\tau : \mathcal{M} \to \{\mathsf{T}, \mathsf{F}\}$ which is (approximately) distributed according to $\mu_{\mathcal{M}}$.
3. Sample $\sigma \sim \mu^\tau$ using the algorithm in Lemma 2.7, and output $\sigma : \mathcal{V} \to \{\mathsf{T}, \mathsf{F}\}$.

There are several technical details we have swept under the rug.

- First, the iterates of Glauber dynamics are never truly distributed according to $\mu_{\mathcal{M}}$, and so we need to extend Lemma 2.6 to accommodate all of these distributions as well. This is why initializing with $\mathsf{Unif}\{\mathsf{T},\mathsf{F}\}^{\mathcal{M}}$ is required.
- Second, because we cannot implement Glauber dynamics exactly due to errors in Lemma 2.7, Theorem 2.5 does not apply. We need to transfer this mixing time bound to the approximate Glauber dynamics that we can actually implement.

For the resolution of all of these, see [Fen+21].

3 Spectral Independence for Marked Variables

In this section, we sketch the proof of Theorem 2.5, which forms the conceptual core of this lecture. The ideas here can also be used to establish Lemma 2.6 (and various analogs). More precisely, we will establish spectral independence for $\mu_{\mathcal{M}}$. The connection with Theorem 2.5 is discussed further in Remark 3.

Theorem 3.1. *Let Φ be a (k,d)-formula, and let $\mathcal{M} \subseteq \mathcal{V}$ be a valid α-marking for some constant $0 < \alpha < 1/2$. If $2^{\alpha k} \geq \Omega\left(dk\right)^5$, then $\mu_{\mathcal{M}}$ is $O(1)$-coupling independent.*

Remark 3. Theorem 3.1 by itself is *not enough* to prove Theorem 2.5, since one also needs to handle conditional distributions. Furthermore, the conditional measure $\mu_{\mathcal{M}}$ does not satisfy any conditional independence properties (even though μ does), and so $O(1)$-spectral independence for all conditionals does not immediately lead to a nearly-linear mixing time for Glauber dynamics (at least not with the technology we have developed so far).

The rough idea that gives $O(n \log n)$-mixing for Glauber dynamics on $\mu_{\mathcal{M}}$ is to extend the coupling argument in the proof of Theorem 3.1 into a *path coupling argument*, where one instead bounds a Dobrushin-type influence. This requires some additional technical ingredients, so to keep things simple, we only prove Theorem 3.1; we refer interested readers to [Fen+21] for the full proof of Theorem 2.5.

Fix any variable $x_0 \in \mathcal{V}$. We will prove that $\mathscr{W}_1\left(\mu^{x_0 \leftarrow \mathsf{T}}, \mu^{x_0 \leftarrow \mathsf{F}}\right) \leq O(1)$, which implies $O(1)$-coupling independence for μ itself, and is stronger than $O(1)$-coupling independence for $\mu_{\mathcal{M}}$. However, we stated Theorem 3.1 for $\mu_{\mathcal{M}}$ instead of μ since ultimately, this argument can only be extended to the conditionals of μ obtained by pinning up to linear-in-$|\mathcal{M}|$ many marked variables; we have little control when the pinning involves unmarked variables since we will need the local uniformity furnished by Theorem 2.2.

Our goal is to iteratively construct a randomly coupled pair of satisfying assignments $\tau \sim \mu^{x_0 \leftarrow \mathsf{T}}, \sigma \sim \mu^{x_0 \leftarrow \mathsf{F}}$ such that $\mathbb{E}[d_H(\tau, \sigma)]$ is small. We use a disagreement percolation argument due to Moitra [Moi19]. Throughout the coupling, we will track a set of "dangerous variables" $\mathcal{V}_{\mathsf{danger}}$, which will always include the set of variables with disagreeing assignments. However, in order to use local uniformity to control the propagation of disagreements, we will also include additional variables in $\mathcal{V}_{\mathsf{danger}}$ which don't necessarily disagree, but for which we will have no control over. Clearly, since $\mathcal{V}_{\mathsf{danger}}$ includes all discrepancy variables, $\mathbb{E}\left[|\mathcal{V}_{\mathsf{danger}}|\right]$ is an upper bound on $\mathscr{W}_1\left(\mu^{x_0 \leftarrow \mathsf{T}}, \mu^{x_0 \leftarrow \mathsf{F}}\right)$. We will control the growth of $\mathcal{V}_{\mathsf{danger}}$ in expectation.

3.1 The Coupling Scheme

We now define the coupling. Initialize $\mathcal{V}_{\mathsf{danger}} = \{x_0\}$, $\tau_1 = \{x_0 \mapsto \mathsf{T}\}, \sigma_1 = \{x_0 \mapsto \mathsf{F}\}$, $\mathcal{C}_{\mathsf{unprocessed}} = \mathcal{C}$, and assume we have currently built up two partial assignments τ_ℓ, σ_ℓ for some ℓ.

- **Phase 1:** Suppose there is a clause $C \in \mathcal{C}_{\mathsf{unprocessed}}$ containing at least one dangerous variable and at least one nondangerous variable. Sequentially sample the remaining unpinned *marked* variables in C using the TV-optimal coupling between their conditional marginals, ultimately yielding two new partial configurations $\tau_{\ell+1}, \sigma_{\ell+1}$. We have two cases:
  1. **Good Case:** Suppose the clause C is now satisfied by *both* $\tau_{\ell+1}$ and $\sigma_{\ell+1}$. Add to $\mathcal{V}_{\mathsf{danger}}$ all currently disagreeing variables in C.
  2. **Bad Case:** Otherwise, add *all* variables in C to $\mathcal{V}_{\mathsf{danger}}$.

  In either case, remove C from $\mathcal{C}_{\mathsf{unprocessed}}$.

- **Phase 2:** Otherwise, optimally couple all remaining variables.

Let us begin our analysis of this coupling with a few basic observations. The following claims also capture some of the key notions we will use later.

Claim 3.2. *1. Throughout the coupling, all variables in $\mathcal{V}_{\mathsf{danger}}$ are dangerous due to at least one of the following types of error:*

- ***Type I Error:*** *The variable has disagreeing assignments under τ_ℓ, σ_ℓ.*
- ***Type II Error:*** *The variable participates in a fully processed clause C which is not satisfied w.r.t. both τ_ℓ and σ_ℓ.*

Furthermore, every variable of disagreement is contained in $\mathcal{V}_{\mathsf{danger}}$.

2. The set of dangerous variables $\mathcal{V}_{\mathsf{danger}}$ is connected in $H^2_\Phi[\mathcal{V}]$.

3. Every clause $C \in \mathcal{C}$ is visited at most once in Phase 1.

Proof. The first claim is obvious. The second claim follows inductively from the fact that every clause selected for consideration in Phase 1 is required to contain a variable in $\mathcal{V}_{\mathsf{danger}}$. For the third claim, if a clause C is visited in Phase 1, then it is removed from $\mathcal{C}_{\mathsf{unprocessed}}$ by the end of that iteration, and so it will never be considered again in Phase 1. □

Claim 3.3. *The moment we enter Phase 2, all clauses $C \in \mathcal{C}$ come in three types.*

- ***Dangerous Clauses*** $\mathcal{C}_{\mathsf{danger}}$***:*** *$C \subseteq \mathcal{V}_{\mathsf{danger}}$ entirely, i.e. all variables in C are dangerous.*
- ***Satisfied Clauses*** $\mathcal{C}_{\mathsf{sat}}$***:*** *$C$ contains a dangerous variable, but is satisfied in both τ_ℓ and σ_ℓ.*
- ***Safe Clauses*** $\mathcal{C}_{\mathsf{safe}}$***:*** *$C \subseteq \mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}$, i.e. all variables in C are not dangerous.*

Moreover, we can perfectly couple all remaining variables not in $\mathcal{V}_{\mathsf{danger}}$.

Proof. Every clause is either in $\mathcal{C}_{\mathsf{unprocessed}}$ or $\mathcal{C} \setminus \mathcal{C}_{\mathsf{unprocessed}}$. The execution of Phase 1 enforces that throughout the coupling, all clauses in $\mathcal{C} \setminus \mathcal{C}_{\mathsf{unprocessed}}$ are either satisfied or completely contained in $\mathcal{V}_{\mathsf{danger}}$. At the termination of Phase 1, all clauses in $\mathcal{C}_{\mathsf{unprocessed}}$ are either fully contained in $\mathcal{V}_{\mathsf{danger}}$ or fully contained in $\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}$. The classification of clauses immediately follows.

Observe that in the graph $H^2_\Phi[\mathcal{C}]$, any path from $\mathcal{C}_{\mathsf{danger}}$ to $\mathcal{C}_{\mathsf{good}}$ must go through a satisfied clause in $\mathcal{C}_{\mathsf{sat}}$. It follows that μ^{τ_ℓ} (resp. μ^{σ_ℓ}) is the uniform measure over satisfying assignments to the formula with clause set $\mathcal{C} \setminus \mathcal{C}_{\mathsf{sat}}$ which is consistent with τ_ℓ (resp. σ_ℓ). Since this formula factorizes into two formulas, one supported on $\mathcal{V}_{\mathsf{danger}}$ and one supported on $\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}$, and τ_ℓ agrees with σ_ℓ on the visited variables in $\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}$, it follows that $\mu^{\tau_\ell}_{\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}} = \mu^{\sigma_\ell}_{\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}}$. In particular, all variables in $\mathcal{V} \setminus \mathcal{V}_{\mathsf{danger}}$ can be coupled perfectly in Phase 2. □

3.2 Controlling the Growth of $\mathcal{V}_{\mathsf{danger}}$

Now let us analyze the expected number of dangerous variables. Connectivity of $\mathcal{V}_{\mathsf{danger}}$ gives us bounds on the worst-case growth of $|\mathcal{V}_{\mathsf{danger}}|$ due to our assumptions on the uniformity and degree of the input formula Φ. Type I errors can be controlled with no problem, but it is not clear how to iteratively bound the probability of Type II errors, since this heavily depends upon which clauses share the bulk of their marked variables.

Obstacle 1. *If two clauses C, C' share almost all of their marked variables, then knowing C is dangerous drastically increases the probability that C' is also dangerous.*

Indeed, if C is dangerous, then C is still not satisfied even though all of its marked variables have assignments. Since C' shares almost all of its marked variables with C, this means C' has almost no free marked variables left, leaving it with "few chances to become satisfied". As an extremal case, one can imagine a *"sunflower"* type of structure for the clauses, with the core entirely consisting of marked variables. At the same time, there is hope, because even if C, C' only share at most a constant fraction of their marked variables, then there are still "enough chances" for C' to become satisfied. Furthermore, any such *"sunflower"* structure cannot have more than $\approx d$ petals due to our degree assumption for the variables, so this extremal case isn't a genuine barrier.

To control $|\mathcal{V}_{\mathsf{danger}}|$, we will look for a "large" collection *disjoint* clauses in $\mathcal{C}_{\mathsf{danger}}$ which we will charge for the dangerous variables. Disjointness will be key, since it will ensure each such clause has a good probability of becoming satisfied. Following [Moi19], we formalize this as follows.

Definition 4 (3-Tree)**.** *Let $G = (V, E)$ be a graph. We call a subset of vertices $\mathcal{T} \subseteq V$ a 3-tree if the following two properties hold:*

1. ***"Independence":*** *Every pair of distinct vertices $u, v \in \mathcal{T}$ satisfies $\mathrm{dist}_G(u, v) \geq 3$. In other words, $\mathcal{T}$ is an independent set in G^2.*
2. ***"Connectivity":*** *If we add an edge between every pair of distinct vertices $u, v \in \mathcal{T}$ satisfying $\mathrm{dist}_G(u, v) = 3$, then $\mathcal{T}$ becomes connected. In other words, $G^3[\mathcal{T}]$ is a connected graph.*

We will use 3-trees in $H_\Phi^2[\mathcal{V}]$ to argue that $|\mathcal{V}_{\mathsf{danger}}|$ is small in expectation. The independence property of 3-trees will allow us to reason about the probability of propagation, while the connectivity property of 3-trees will allow us to control the number of such 3-trees. We have the following claims, which follow straightforwardly from Definition 4.

Claim 3.4. *$H_\Phi^2[\mathcal{V}_{\mathsf{danger}}]$ admits a 3-tree $\mathcal{T}$ of size at least $\Omega\left(\frac{1}{d^2k^2}\right) \cdot |\mathcal{V}_{\mathsf{danger}}|$ variables.*

Proof Sketch. Use the same greedy algorithm as the one for constructing maximal independent sets in bounded-degree graphs. Maximality of the output will certify the connectivity property. □

Claim 3.5.

- ***Disjointness:*** *For any 3-tree $\mathcal{T}$ in $H_\Phi^2[\mathcal{V}]$ and any pair of distinct variables $x, x' \in \mathcal{T}$, if x, x' both participate in dangerous clauses C, C' respectively (i.e. they suffered from Type II errors), then $C \cap C' = \emptyset$. (This only requires Item 1.)*
- ***Bounded Counts:*** *The number of 3-trees of size-t in $H_\Phi^2[\mathcal{V}]$ containing x_0 is at most $O(d^3k^3)^t$. (This only requires Item 2.)*

Proof Sketch. For the first item, if $C \cap C' \neq \emptyset$, then any variable in $C \cap C'$ certifies that the distance between x and x' is at most 2 in $H_\Phi^2[\mathcal{V}]$, contradicting the "independence" property of 3-trees. The second item just follows from the standard bound on the number of connected induced subgraphs containing a particular vertex in a sparse graph [Bor+13]; note that $\mathcal{T}$ is connected in $H_\Phi^2[\mathcal{V}]$, which has maximum degree $O(d^3k^3)$. □

Proof of Theorem 3.1. For any $t \in \mathbb{N}$, we have that

$$\begin{aligned}\Pr\left[|\mathcal{V}_{\mathsf{danger}}| \geq Cd^2k^2t\right] &\leq \Pr\left[\exists \text{ 3-tree } \mathcal{T} \text{ in } H_\Phi^2[\mathcal{V}] \text{ s.t. } |\mathcal{T}| = t, \mathcal{T} \subseteq \mathcal{V}_{\mathsf{danger}}\right] && \text{(Claim 3.4)}\\ &\leq \sum_{\substack{\text{3-tree } \mathcal{T} \text{ in } H_\Phi^2[\mathcal{V}] \\ \mathcal{T} \ni x_0, |\mathcal{T}| = t}} \Pr[\mathcal{T} \subseteq \mathcal{V}_{\mathsf{danger}}]\end{aligned}$$

Our goal is to upper bound this by something exponentially small in t, which is clearly sufficient (e.g. by the "layered cake representation" of an expectation). Fix some 3-tree $\mathcal{T}$ in $H_\Phi^2[\mathcal{V}]$ with $\mathcal{T} \ni x_0$ and $|\mathcal{T}| = t$. We bound the probability that all variables in $\mathcal{T}$ become dangerous after the above coupling process. Arbitrarily order the variables of $\mathcal{T}$ as $x_0, x_1, \ldots, x_{t-1}$. We have

$$\begin{aligned}\Pr[\mathcal{T} \subseteq \mathcal{V}_{\mathsf{danger}}] &= \prod_{j=0}^{t-1} \Pr[x_i \text{ dangerous} \mid x_0, \ldots, x_{j-1} \text{ dangerous}]\\ &\leq \prod_{j=0}^{t-1} \Bigg(\Pr\left[x_i \text{ Type I} \mid x_0, \ldots, x_{j-1} \in \mathcal{V}_{\mathsf{danger}}\right]\\ &\qquad + \Pr\left[x_i \text{ Type II} \mid x_0, \ldots, x_{j-1} \in \mathcal{V}_{\mathsf{danger}}\right] \Bigg). && \text{(Union Bound)}\end{aligned}$$

By Claim 3.5, these variables do not participate in clauses which intersect, and since our coupling is faithful,

$$\begin{aligned}\Pr\left[x_i \text{ Type I} \mid x_0, \ldots, x_{j-1} \in \mathcal{V}_{\mathsf{danger}}\right] &\leq 1 - \exp\left(\frac{1}{s}\right) && \text{(Theorem 2.2)}\\ &\leq \frac{2}{s},\end{aligned}$$

and

$$\Pr\left[x_i \text{ Type II} \mid x_0, \dots, x_{j-1} \in \mathcal{V}_{\mathsf{danger}}\right] \le d \cdot \left(\frac{1}{2}\exp\left(\frac{1}{s}\right)\right)^{|C\cap\mathcal{M}|} \quad \text{(Theorem 2.2 and } \deg(x_i) \le d\text{)}$$
$$\le d \cdot \left(\frac{1}{2}+\frac{1}{s}\right)^{\alpha k} \quad \text{(Disjointness and } \alpha\text{-marking guarantee)}$$

It follows that for such $\mathcal{T}$,

$$\Pr[\mathcal{T} \subseteq \mathcal{V}_{\mathsf{danger}}] \le \left(\frac{2}{s} + d \cdot \left(\frac{1}{2}+\frac{1}{s}\right)^{\alpha k}\right)^t,$$

whence

$$\Pr\left[|\mathcal{V}_{\mathsf{danger}}| \ge Cd^2k^2t\right] \le \#\left\{\substack{\text{3-tree } \mathcal{T} \text{ in } H^2_\Phi[\mathcal{V}] \\ \mathcal{T}\ni x_0, |\mathcal{T}|=t}\right\} \cdot \left(\frac{2}{s} + d \cdot \left(\frac{1}{2}+\frac{1}{s}\right)^{\alpha k}\right)^t$$
$$\lesssim \left(\frac{O(d^3k^3)}{s} + O(d^4k^3) \cdot \left(\frac{1}{2}+\frac{1}{s}\right)^{\alpha k}\right)^t \quad \text{(Claim 3.5)}$$
$$\le \frac{1}{2^t}. \quad \text{(Assuming } 0<\alpha<1/2 \text{ constant and } s \ge \Omega\left(d^4k^5\right)\text{)}$$

as desired. □

4 Random k-SAT

Since this line of research was initiated, recent works have turned to *random* k-CNF *formulas*, where for each $j = 1, \dots, m$, we add a clause on a uniformly random k-subset of $\mathcal{V}$, with each variable negated independently with probability 1/2. This setting is more technically challenging, since the maximum degree of a variable now increases with n. Nonetheless, it is known that if the *average degree* $\frac{m}{n}$ is at most $2^{O(k)}$, efficient sampling and counting algorithms for satisfying assignments still exist. Very roughly speaking, one sampling method is to replace Glauber dynamics for the set of marked variables $\mathcal{M} \subseteq \mathcal{V}$ with an appropriate *block dynamics* which can still be implemented efficiently via shattering-type lemmas, and then establish $O(\log n)$-spectral independence for $\mu_{\mathcal{M}}$ [Gal+22; CMM23]. Many other ingenious algorithmic techniques have also been developed [Moi19; GJL19; Gal+21; HWY23]. See also the recent breakthrough on the *satisfiability threshold* [DSS22]. For extensions of these ideas to other constraint satisfaction problems, see e.g. [JPV21; HWY22].

References

[BDK08] Magnus Bordewich, Martin Dyer, and Marek Karpinski. "Path coupling using stopping times and counting independent sets and colorings in hypergraphs". In: *Random Structures & Algorithms* 32.3 (2008), pp. 375–399 (cit. on p. 2).

[Bez+19] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Štefankovič. "Approximation via Correlation Decay When Strong Spatial Mixing Fails". In: *SIAM Journal on Computing* 48.2 (2019), pp. 279–349. DOI: 10.1137/16M1083906 (cit. on p. 2).

[Bor+13] Christian Borgs, Jennifer Chayes, Jeff Kahn, and László Lovász. "Left and right convergence of graphs with bounded degree". In: *Random Structures & Algorithms* 42.1 (2013), pp. 1–28 (cit. on p. 8).

[CMM23] Zongchen Chen, Nitya Mani, and Ankur Moitra. "From Algorithms to Connectivity and Back: Finding a Giant Component in Random k-SAT". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2023, pp. 3437–3470. DOI: 10.1137/1.9781611977554.ch132. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611977554.ch132 (cit. on p. 9).

[DSS22] Jian Ding, Allan Sly, and Nike Sun. "Proof of the satisfiability conjecture for large k". In: *The Annals of Mathematics* 196 (1 2022), pp. 1–388 (cit. on p. 9).

[Fen+21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. "Fast Sampling and Counting K-SAT Solutions in the Local Lemma Regime". In: *J. ACM* 68.6 (Oct. 2021). ISSN: 0004-5411. DOI: 10.1145/3469832 (cit. on pp. 2–6).

[Gal+21] Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Kuan Yang. "Counting Solutions to Random CNF Formulas". In: *SIAM Journal on Computing* 50.6 (2021), pp. 1701–1738. DOI: 10.1137/20M1351527 (cit. on p. 9).

[Gal+22] Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Andrés Herrera-Poyatos. "Fast sampling of satisfying assignments from random k-SAT". In: *arXiv preprint arXiv:2206.15308* (2022) (cit. on p. 9).

[GJL19] Heng Guo, Mark Jerrum, and Jingcheng Liu. "Uniform Sampling Through the Lovász Local Lemma". In: *J. ACM* 66.3 (Apr. 2019). ISSN: 0004-5411. DOI: 10.1145/3310131 (cit. on p. 9).

[HSZ19] Jonathan Hermon, Allan Sly, and Yumeng Zhang. "Rapid mixing of hypergraph independent sets". In: *Random Structures & Algorithms* 54.4 (2019), pp. 730–767 (cit. on p. 2).

[HWY22] Kun He, Chunyang Wang, and Yitong Yin. "Sampling Lovász local lemma for general constraint satisfaction solutions in near-linear time". In: *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022, pp. 147–158. DOI: 10.1109/FOCS54457.2022.00021 (cit. on p. 9).

[HWY23] Kun He, Kewen Wu, and Kuan Yang. "Improved Bounds for Sampling Solutions of Random CNF Formulas". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2023, pp. 3330–3361. DOI: 10.1137/1.9781611977554.ch128. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611977554.ch128 (cit. on p. 9).

[JPV21] Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. "On the sampling Lovász Local Lemma for atomic constraint satisfaction problems". In: *arXiv preprint arXiv:2102.08342* (2021) (cit. on p. 9).

[KLM89] Richard M Karp, Michael Luby, and Neal Madras. "Monte-Carlo approximation algorithms for enumeration problems". In: *Journal of Algorithms* 10.3 (1989), pp. 429–448. ISSN: 0196-6774. DOI: https://doi.org/10.1016/0196-6774(89)90038-2 (cit. on p. 1).

[Moi19] Ankur Moitra. "Approximate Counting, the Lovász Local Lemma, and Inference in Graphical Models". In: *J. ACM* 66.2 (Apr. 2019). ISSN: 0004-5411. DOI: 10.1145/3268930 (cit. on pp. 2–4, 6, 7, 9).

[MT10] Robin A. Moser and Gábor Tardos. "A Constructive Proof of the General Lovász Local Lemma". In: *J. ACM* 57.2 (Feb. 2010). ISSN: 0004-5411. DOI: 10.1145/1667053.1667060 (cit. on p. 2).

[Sly10] Allan Sly. "Computational Transition at the Uniqueness Threshold". In: *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. 2010, pp. 287–296. DOI: 10.1109/FOCS.2010.34 (cit. on p. 2).

[SS14] Allan Sly and Nike Sun. "The Computational Hardness of Counting in Two-Spin Models on d-Regular Graphs". In: *The Annals of Probability* 42.6 (2014), pp. 2383–2416 (cit. on p. 2).