More Efficient k-wise Independent Permutations from Random Reversible Circuits via log-Sobolev Inequalities

Lucas Gretta[*]   William He[†]   Angelos Pelecanos[‡]

May 30, 2024

Abstract

We prove that the permutation computed by a reversible circuit with $\widetilde{O}(nk \cdot \log(1/\varepsilon))$ random 3-bit gates is ε-approximately k-wise independent. Our bound improves on currently known bounds in the regime when the approximation error ε is not too small. We obtain our results by analyzing the log-Sobolev constants of appropriate Markov chains rather than their spectral gaps.

Contents

---

[*]UC Berkeley. Email: lucas_gretta@berkeley.edu.

[†]Carnegie Mellon University. Email: wrhe@cs.cmu.edu. Supported in part by ARO grant W911NF2110001.

[‡]UC Berkeley. Email: apelecan@berkeley.edu.

1 Introduction

We consider the extent to which small random reversible circuits compute *almost k-wise independent* permutations. The (almost) k-wise independence of permutations was first considered by Gowers [Gow96] as a proxy for pseudorandomness properties of practical cryptosystems, such as block ciphers.

Definition 1 (Approximate k-wise independent permutations)**.** A distribution $\mathcal{P}$ on the symmetric group $S_{[N]}$ is said to be ε-*approximate k-wise independent* if for all distinct $x_1, \ldots, x_k \in [N]$, the distribution of $(\boldsymbol{g}(x_1), \ldots, \boldsymbol{g}(x_k))$ for $\boldsymbol{g} \sim \mathcal{P}$ has total variation distance at most ε from the uniform distribution on distinct k-tuples over $[N]$.

A commonly studied construction of approximate k-wise independent permutations is a reversible circuit on n wires in which each gate computes a randomly chosen width-2 (see Definition 4) permutation on a random subset of 3 wires. From here on, when referring to a random reversible circuit, we mean a random circuit whose gates are drawn randomly from a set of 3-bit gates. Gowers [Gow96] introduced this construction and proved that a random reversible circuit with $\operatorname{poly}(n, k, \log(1/\varepsilon))$ gates computes an ε-approximate k-wise independent permutation of the cube $\{0, 1\}^n$ using the canonical paths technique from Markov chain mixing [Jer03]. Since then, follow-up works by Hoory et al. and Brodsky and Hoory [HMMR05, BH05] improved on the analysis of Gowers and proved that if $k \leq 2^{n/50}$, then random reversible circuits with $O(n^2 k^2 \log(1/\varepsilon))$ gates compute an ε-approximate k-wise independent permutation using the comparison method [DSC93b, DSC93a]. Finally, using quantum-inspired techniques for proving spectral gaps, He and O'Donnell [HO24] improved the number of gates needed to $\widetilde{O}(nk) \cdot (nk + \log(1/\varepsilon))$.

Random circuits have gained attention following the recent interest in random *quantum* circuits. The natural quantum analog of a (approximate) k-wise independent permutation is that of a (approximate) unitary k-design.[1] Unitary designs are widely studied in quantum computation and quantum physics as basic pseudorandom objects and models for equilibration in quantum many-body systems [BCHJ+21]. A line of work on unitary k-designs [BHH16, HHJ21] shows that for constant ε, a reversible circuit on n wires with $\widetilde{O}(n^2 \cdot \operatorname{poly}(k))$ random 3-qubit quantum gates chosen from some finite gate set (a random quantum circuit) gives a construction of an ε-approximate unitary k-design.

Recent works [MPSY24, CBB+24] obtain k-designs with size linear in k from classical k-wise independent permutations whose size is also linear in k. Even though we demonstrate that a linear-in-k number of random width-2 gates suffices to ε-approximate k-wise independence, we remark that our dependence on ε is not sufficiently tight for their k-design construction. In particular, both works employ a theorem of Alon and Lovett [AL13] which requires an exponentially small ε to translate from approximate to exact k-wise independent permutations. Plugging in such a small ε in our theorem would increase our size bound by polynomial factors in n and k.

Another line of work, motivated by the design of practical cryptosystems (such as block ciphers), studies the computational pseudorandomness properties of random reversible circuits. He and O'Donnell [HO24] consider the computational hardness of inverting the permutation computed by short reversible circuits with 3-bit gates. Another line of work by Canetti et al. [CCMR24] proposed more advanced cryptographic primitives based on the cryptographic properties of random reversible circuits. In particular, using the assumption that random reversible circuits achieve computational pseudorandomness after a modest number of rounds (much less than the super-polynomial number

---

[1]A (approximate) unitary k-design is a distribution on the unitary group that (approximately) matches the Haar distribution up to k^{th} moments.

of rounds required to reach statistical pseudorandomness), they suggest candidate obfuscation schemes along with possible ways to prove their computational security. Their approach is inspired by thermalizing processes of statistical mechanics.

In this paper, we revisit the problem of random circuits with reversible 3-bit gates and show that a random reversible circuit with $\widetilde{O}(nk \cdot \log(1/\varepsilon))$ gates gives an ε-approximate k-wise independent permutation. The following is our main theorem, which we prove in Section 6.

Theorem 2. For any n and $k \leq 2^{n/50}$, a random reversible circuit with $\widetilde{O}(nk \cdot \log(1/\varepsilon))$ width-2 gates (a subset of 3-bit gates) computes an ε-approximate k-wise independent permutation, where the $\widetilde{O}$ hides polylog(n, k) factors.

We note here that for applications of approximate k-wise independent permutation distributions $\mathcal{P}$ in derandomization, one is generally concerned with the number of truly random "seed" bits needed to generate a draw from $\mathcal{P}$. See, for example [MOP20]. By using techniques such as derandomized squaring (see [KNR09]), one can often reduce the seed length to $O(nk)$ for any construction. This is true for the results in our paper, and we don't discuss the seed length any further, as we are generally focused on the circuit complexity of our permutations.

1.1 Proof overview

We use the comparison method in a similar way as [BH05]. In particular, we bound the log-Sobolev constant of the natural Markov chain associated with the computation of a random reversible circuit, by comparing it to the log-Sobolev constant of the *k-clique 2^n-coloring* Markov chain. By working with the log-Sobolev constant rather than the spectral gap of this random walk as [BH05, HO24] do, we obtain an improved mixing time since the log-Sobolev constant gives a mixing time bound that depends doubly logarithmically on the smallest probability of the stationary distribution. In contrast, the spectral gap gives bounds that depend logarithmically on this quantity.

While it is generally more difficult to bound the log-Sobolev constant of a Markov chain, recent work of Salez [Sal20] has used the martingale method of Lee and Yau [LY98] to obtain sharp estimates for the log-Sobolev constant of a natural random walk on the multislice. Using this method, we estimate the log-Sobolev constant of a variant of k-clique 2^n-coloring chain, which we call the *uniform k-clique 2^n-coloring* chain. The log-Sobolev constant for the standard k-clique 2^n-coloring chain is then obtained via a simple application of the comparison method.

In more detail, our starting point is the work of Salez which bounds the log-Sobolev of the *multislice*. The multislice corresponds to the random walk over the set of colorings of 2^n items, where each step of the walk swaps the colors of any two items chosen uniformly at random. The colorings are comprised of $k + 1$ colors, where the first k colors appear once and the last color appears in the remaining $2^n - k$ items. The first observation is that this random walk captures the k-wise independence of a random walk with transpositions. Unfortunately, the log-Sobolev constant of this walk is too small: $(n \cdot 2^n)^{-1}$. In contrast, we would expect a random set of transpositions to mix to a k-wise independent permutation within a time that is dependent on k.

The reason that the log-Sobolev constant of the multislice chain is independent of k is because it applies a random transposition from the entire set of $\binom{2^n}{2}$ transpositions. In the case when k is much smaller than 2^n, a random transposition will most likely exchange the colors of two of the $2^n - k$ items that have color $k + 1$. Thus, with high probability, roughly $1 - \frac{k}{2^n}$, the multislice chain will not move to a new state. To avoid this artificial slowdown, we study the uniform k-clique 2^n-coloring chain, which requires that every step applies one transposition with an element that doesn't have color $k + 1$. Equivalently, one may think of the uniform k-clique 2^n-coloring chain as a random walk on the multislice that takes $\frac{2^n}{k}$ steps per time step and thus would hope that the

log-Sobolev constant scales down by a factor of $\frac{k}{2^n}$. Indeed, we employ the martingale method and prove that the log-Sobolev constant of the uniform k-clique 2^n-coloring chain is $\Omega\left(\frac{1}{nk}\right)$ as expected.

One can compute the log-Sobolev constant of the uniform k-clique 2^n-coloring chain by using Salez's result as a black box and viewing the multislice chain as a lazy version of the uniform k-clique 2^n-coloring chain. We instead present an alternative proof by adapting the martingale method used by Salez.

The next step is to transfer our log-Sobolev bound from the uniform k-clique 2^n-coloring chain to the k-clique 2^n-coloring chain, which has slightly different transition probabilities than its uniform counterpart. We give a randomized paths construction with only a constant amount of congestion. The comparison method implies that the log-Sobolev constant of the k-clique 2^n-coloring chain is also $\Omega\left(\frac{1}{nk}\right)$.

Finally, we obtain an estimate for the log-Sobolev constant of the random reversible circuits Markov chain by employing the comparison with the k-clique 2^n-coloring chain from [BH05]. More specifically, Brodsky and Hoory give a randomized paths construction with a comparison constant of $\Theta(n^2)$. This concludes our $\Omega\left(\frac{1}{n^3 k}\right)$ bound for the log-Sobolev constant of the reversible circuits Markov chain.

To improve our bound on the mixing time of the reversible circuits Markov chain, we use another argument from [BH05]. The observation is that after a short random walk of $\widetilde{O}(n)$ steps, the state of the reversible circuits Markov chain is very likely to be in a *generic* state. Thus it suffices to bound the mixing time of the Markov chain when restricted to generic states. We do this by bounding its log-Sobolev constant, using the log-Sobolev inequality of the clique coloring chain, which we proved earlier. This allows us to bring down the mixing time of the reversible circuits Markov chain to $O(nk \cdot \mathrm{polylog}(n, k))$.

2 Preliminaries

Notation. In this paper we will use the symbols $\gtrsim, \lesssim$ to compare two quantities in the asymptotic sense, in particular, these symbols hide constant factors. For example, $f(n) \lesssim g(n) \iff f(n) \leq O(g(n))$. When $x = (x_1, \ldots, x_k)$ is a tuple, we use the notation $\ell \in x$ whenever $\ell = x_i$ for some $i \in [k]$ and otherwise, we write $\ell \notin x$.

Definition 3 (Tuples with distinct elements). Let S be a set. We define the set of k-*tuples with distinct elements from S* as follows:

$$\Theta_{k,S} := \left\{ (x_1, \ldots, x_k) \in S^k : x_i\text{'s distinct} \right\}.$$

We frequently write $\Theta_{k,N}$ in the place of $\Theta_{k,[N]}$.

We recall the definition of width-2 simple permutations from [BH05].

Definition 4 (Width-2 simple permutations). The set of *width-2 simple permutations* is the following set of permutations on $\{0,1\}^n$

$$\Sigma := \left\{ f_{i,j_1,j_2,h} : \begin{array}{c} i, j_1, j_2 \in [n], i \neq j_1, j_2 \\ h \text{ Boolean function on } \{0,1\}^2 \end{array} \right\}.$$

The permutation $f_{i,j_1,j_2,h}$ maps $(x_1, \ldots, x_n)$ to $(x_1, \ldots, x_{i-1}, x_i \oplus h(x_{j_1}, x_{j_2}), x_{i+1}, \ldots, x_n)$.

In words, a width-2 permutation chooses 3 random indices from $[n]$: i and j_1, j_2. It further samples a random Boolean function on 2 bits. Then it XORs the value of $h(x_{j_1}, x_{j_2})$ on the ith bit of the input.

2.1 Log-Sobolev constant and mixing time

We recall some background on Markov chains from [SC97]. Let P be the transition matrix of an ergodic Markov chain over finite state space V, and let π denote its stationary distribution. We identify a Markov chain with its transition matrix, so we will often say that P is both the transition matrix for a Markov chain and also the Markov chain itself. We let p_x^t denote the probability distribution of P, starting at state x, at timestep t.

Definition 5 (Mixing time). The ε-mixing time of an ergodic Markov chain P is defined as:

$$\tau_\varepsilon(P) := \min \left\{ t \geq 0 : \max_{x \in V} \left\| p_x^t - \pi \right\|_{\mathrm{TV}} \right\}.$$

When the subscript is dropped, we mean $\tau(P) = \tau_{1/4}(P)$.

Throughout this paper, we deal only with *reversible* Markov chains.

Definition 6 (Reversible Markov chain). We say that a Markov chain P is *reversible* if for all $x, y \in V$,

$$\pi(x)P(x, y) = \pi(y)P(y, x).$$

One powerful way of bounding the mixing time of Markov chains is by functional inequalities using the Dirichlet form.

Definition 7 (Dirichlet form). For function $f : V \to \mathbb{R}_{\geq 0}$, the *Dirichlet form* of f with respect to P is

$$\mathcal{E}_P(f, f) := \frac{1}{2} \sum_{x, y \in \Omega} \big(f(x) - f(y) \big)^2 \pi(x)P(x, y).$$

Intuitively, the Dirichlet form measures the "local variation" of f with respect to the (weighted) graph underlying a Markov chain P.

Definition 8 (Entropy). For a function $f : V \to \mathbb{R}_{\geq 0}$, we define its *entropy*

$$\mathsf{Ent}_\pi[f] := \sum_{x \in V} \pi(x)f(x) \log \frac{f(x)}{\mathbb{E}_\pi[f]},$$

where $\mathbb{E}_\pi[f] = \sum_{x \in V} \pi(x)f(x)$.

The ratio of these two quantities defines the log-Sobolev constant of the Markov chain.

Definition 9 (Log-Sobolev constant of Markov chain). The *log-Sobolev constant* of P is defined by

$$\alpha(P) := \inf_{\substack{f \geq 0 \\ f \text{ non-constant}}} \frac{\mathcal{E}_P(\sqrt{f}, \sqrt{f})}{\mathsf{Ent}_\pi[f]}.$$

The log-Sobolev constant of a Markov chain bounds the mixing time of the chain according to the following theorem. Note the doubly-logarithmic dependence on $1/\pi_{\min}$, which is the conceptual advantage of using log-Sobolev inequalities over a spectral gap analysis, whenever ε is not exponentially small.

Theorem 10 ([DSC96], Theorem 3.7). Let P be the transition matrix of a reversible Markov chain whose stationary distribution is π, and $\pi_{\min}$ to be the smallest stationary probability. For $\varepsilon \leq \frac{1}{e}$, the ε-mixing time is bounded by

$$\tau_\varepsilon(P) \lesssim \frac{1}{\alpha}\left(\log\log\frac{1}{\pi_{\min}} + \log\frac{1}{\varepsilon}\right).$$

In fact, the log-Sobolev constant bounds the ℓ^∞ mixing time, which gives *pointwise* distance bounds.

Theorem 11 ([DSC96], Corollary 3.8). For reversible P, and for all $x, y \in V$

$$\left|p_x^t(y) - \pi(y)\right| \leq \varepsilon\pi(y)$$

when $t \gtrsim \frac{1}{\alpha}\left(\log\log\frac{1}{\pi_{\min}} + \log\frac{1}{\varepsilon}\right)$.

2.2 The comparison method

We bound the log-Sobolev constant of a reversible circuits Markov chain by repeated application of the comparison method [DSC93b, WLP09] which we introduce below. The comparison method is used to estimate the Dirichlet form of a *target* Markov chain with transition matrix P by relating it to the Dirichlet form of a *reference* Markov chain with transition matrix $\widetilde{P}$, for which we have previously-known estimates. This relation between Dirichlet forms can be trivially extended to an inequality between log-Sobolev constants when $\widetilde{P}$ and P are over the same state space V and have the same stationary distribution π.

The comparison is achieved by "simulating" the transition probabilities of the $\widetilde{P}$ Markov chain using paths from P. Formally, for each $(x, y) \in V^2$ we assign a random path

$$\boldsymbol{\Delta}(x, y) = \big((x, \boldsymbol{u}_1), (\boldsymbol{u}_1, \boldsymbol{u}_2), (\boldsymbol{u}_2, \boldsymbol{u}_3), \ldots, (\boldsymbol{u}_{\boldsymbol{\ell}}, y)\big),$$

where the $\boldsymbol{u}_i$'s are random elements of V that satisfy $P(x, \boldsymbol{u}_1), P(\boldsymbol{u}_{\boldsymbol{\ell}}, y) > 0$ and $P(\boldsymbol{u}_i, \boldsymbol{u}_{i+1}) > 0$. The quantity $\boldsymbol{\ell}$ is a random non-negative integer equal to the length of the path $|\boldsymbol{\Delta}(x, y)|$. The congestion of these paths (which is captured by the comparison constant $A(\boldsymbol{\Delta})$) provides a lower bound of $\mathcal{E}$ with respect to $\widetilde{\mathcal{E}}$ as shown formally in Lemma 12.

Without loss of generality, we assume that the paths $\boldsymbol{\Delta}(x, y)$ are simple, since one can remove all loops without affecting the endpoints x, y of a path and without increasing the congestion.

Lemma 12 ([WLP09], Corollary 13.23). Let $\widetilde{P}$ and P be transition matrices for two ergodic Markov chains on the same state space V. Assume that for each $(x, y) \in V^2$ there exists a random path

$$\boldsymbol{\Delta}(x, y) = \big((x, \boldsymbol{u}_1), (\boldsymbol{u}_1, \boldsymbol{u}_2), (\boldsymbol{u}_2, \boldsymbol{u}_3), \ldots, (\boldsymbol{u}_{\boldsymbol{\ell}}, y)\big).$$

Then we have for any $f : V \to \mathbb{R}$ that

$$\widetilde{\mathcal{E}}(f, f) \leq A(\boldsymbol{\Delta}) \cdot \mathcal{E}(f, f)$$

where the comparison constant of $\boldsymbol{\Delta}$ is defined to be

$$A(\boldsymbol{\Delta}) := \max_{\substack{(a,b)\in V^2 \\ \widetilde{P}(a,b)>0}}\left\{\frac{1}{\pi(x)P(a,b)}\sum_{(x,y)\in V^2}\mathop{\mathbf{E}}_{\boldsymbol{\Delta}}\left[\mathbf{1}_{(a,b)\in\boldsymbol{\Delta}(x,y)}\cdot|\boldsymbol{\Delta}(x,y)|\right]\cdot\widetilde{\pi}(x)\cdot\widetilde{P}(x,y)\right\}.$$

Here π and $\widetilde{\pi}$ are the (unique) stationary distributions for P and $\widetilde{P}$, respectively, and $\mathbf{1}_{(a,b)\in Q}$ is the indicator variable which captures whether the edge (a, b) appears in the sequence Q.

3 The Markov chains

We now set up the Markov chains we use in the proof of Theorem 2. Throughout this section (and the rest of the paper) fix positive integers n, k, and N (which will typically be equal to 2^n). Our Markov chains all have domains isomorphic to $\Theta_{k,U}$ for some set U:

Definition 13 (Reversible circuit Markov chain). The chain $\{\boldsymbol{X}_t^{\mathsf{rev}}\}_{t \geq 0}$ on the state space of k distinct n-bit strings is given by the following distribution on $\boldsymbol{X}_{t+1}^{\mathsf{rev}} | \boldsymbol{X}_t^{\mathsf{rev}}$. Given the current state $x = (x_1, \ldots, x_k)$, to draw the next state $\boldsymbol{X}_{t+1} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$, draw a uniformly random width-2 permutation $\boldsymbol{\sigma} \in \Sigma$ and set

$$(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) = (\boldsymbol{\sigma} x_1, \ldots, \boldsymbol{\sigma} x_k).$$

Let $P_{k,n}^{\mathsf{rev}}$ be the transition matrix of this Markov chain.

This Markov chain exactly captures the evolution of k inputs to a random reversible circuit whose gates are uniformly drawn from the set of width-2 permutations Σ. Thus the statement of Theorem 2 that a random reversible circuit with s width-2 gates is an ε-approximate k-wise independent permutation is implied by the statement that $\tau_\varepsilon\left(P_{k,n}^{\mathsf{rev}}\right) \leq s$. We typically write P^{rev} and omit the parameters k and n whenever they are clear from the context or not important.

Following [BH05], we prove that this Markov chain mixes fast by comparing it to the *k-clique 2^n-coloring* Markov chain. In this paper we deal with two clique coloring chains, thus we will refer to this chain as the *standard clique coloring*, or simply the *clique coloring* chain. (Note that this chain is slightly different than the)

Definition 14 (Standard k-clique N-coloring Markov chain). Let N be the number of colors and k be the number of clique vertices. The *k-clique N-coloring* chain $\{\boldsymbol{X}_t^{\mathsf{cc}}\}_{t \geq 0}$ on the set of colorings $\Theta_{k,N}$ is given by the following distribution on $\boldsymbol{X}_{t+1}^{\mathsf{cc}} | \boldsymbol{X}_t^{\mathsf{cc}}$. To sample $\boldsymbol{X}_{t+1}^{\mathsf{cc}} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ given the current state $\boldsymbol{X}_t^{\mathsf{cc}} = x = (x_1, \ldots, x_k)$, uniformly sample $\boldsymbol{i} \in [k]$ and $\boldsymbol{\ell} \in \{\ell \in [N] : \ell \notin x\} \cup \{x_{\boldsymbol{i}}\}$ and set

$$\boldsymbol{y}_j = \begin{cases} \boldsymbol{\ell} & j = \boldsymbol{i} \\ x_j & j \neq \boldsymbol{i} \end{cases}.$$

Let $P_{k,N}^{\mathsf{cc}}$ be the transition matrix for this Markov chain.

In other words, the clique coloring chain samples a uniformly random coloring of the k-clique with N colors, by randomly choosing a vertex and randomly assigning it one of the $(N - k + 1)$ available colors (including its current color).

We directly bound the log-Sobolev constant of a related Markov chain, which we call the *uniform clique coloring* chain.

Definition 15 (Uniform k-clique N-coloring Markov chain). Let N be the number of colors and k be the number of clique vertices. The *uniform k-clique N-coloring* chain $\{\boldsymbol{X}_t^{\mathsf{ucc}}\}_{t \geq 0}$ on the set of colorings $\Theta_{k,N}$ is given by the following distribution on $\boldsymbol{X}_{t+1}^{\mathsf{ucc}} | \boldsymbol{X}_t^{\mathsf{ucc}}$. To sample $\boldsymbol{X}_{t+1}^{\mathsf{ucc}} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ given the current state $\boldsymbol{X}_t^{\mathsf{ucc}} = x = (x_1, \ldots, x_k)$ uniformly sample $\boldsymbol{i} \in [k]$ and $\boldsymbol{\ell} \in [N]$ and set

$$\boldsymbol{y}_j = \begin{cases} \boldsymbol{\ell} & j = \boldsymbol{i} \\ x_{\boldsymbol{i}} & \boldsymbol{\ell} = x_j \\ x_j & \text{otherwise} \end{cases}.$$

Let $P_{k,N}^{\mathsf{ucc}}$ be the transition matrix for this Markov chain.

We call this the uniform clique coloring chain, since at every step a random vertex $\boldsymbol{i}$ is re-colored with a uniformly random color from the entire set $[N]$. If this color is already taken by another vertex j, the two vertices swap colors. This additional symmetry allows us to obtain a bound on the log-Sobolev constant of this chain by adapting the martingale method of Lee and Yau [LY98]. Moreover, it is not hard to relate the log-Sobolev constants of the uniform and standard clique coloring chains using the comparison method.

With all of our Markov chains defined, we now state the sequence of inequalities that will allow us to conclude Theorem 2, deferring the proofs of the auxiliary results to later sections.

Theorem 16. Let $P_{k,n}^{\mathsf{rev}}$ be the transition matrix corresponding to the random walk from Definition 13. Then

$$\alpha(P_{k,n}^{\mathsf{rev}}) \geq \Omega\left(\frac{1}{n^3 k}\right).$$

Proof. We will show the following sequence of inequalities (recall that $\gtrsim$ hides constant factors):

$$\alpha(P_{k,n}^{\mathsf{rev}}) \underset{\text{Corollary 24}}{\gtrsim} \frac{1}{n^2} \cdot \alpha(P_{k,2^n}^{\mathsf{cc}}) \underset{\text{Lemma 21}}{\gtrsim} \frac{1}{n^2} \cdot \alpha(P_{k,2^n}^{\mathsf{ucc}}) \underset{\text{Lemma 18}}{\gtrsim} \frac{1}{n^3 k}. \qquad \square$$

Theorem 16 immediately gives a mixing time of $\widetilde{O}(n^3 k \cdot \log(1/\varepsilon))$ for the reversible circuits chain by Theorem 10; in Section 6 we improve the mixing time to $\widetilde{O}(nk \cdot \log(1/\varepsilon))$ by applying ideas of [BH05], thus proving Theorem 2.

It may then seem that Theorem 16 is strictly weaker than Theorem 2. However, the proof of Theorem 2 does not yield a good log-Sobolev inequality for the reversible circuits Markov chain. Thus we cannot use that proof to conclude results about pointwise convergence as we can from log-Sobolev bounds using Theorem 11, such as the following result:

Corollary 17. Let p_x^t be the distribution over V after $t \gtrsim n^3 k \left(\log nk + \log \frac{1}{\varepsilon}\right)$ steps of $P_{k,n}^{\mathsf{rev}}$. For all $x, y, \in V$

$$\frac{1 - \varepsilon}{2^n (2^n - 1) \cdots (2^n - k + 1)} \leq \mathbf{Pr}[p_x^t = y] \leq \frac{1 + \varepsilon}{2^n (2^n - 1) \cdots (2^n - k + 1)}.$$

4   The Log-Sobolev Constant of the Uniform Clique Coloring Chain

The goal of this section is to lower bound the log-Sobolev constant of the uniform clique coloring Markov chain.

Recall that the uniform k-clique N-coloring Markov chain has state space $\Theta_{k,N}$ of size $N(N - 1)\ldots(N - k + 1)$. Given some $x = (x_1, \ldots, x_k) \in \Theta_{k,N}$, the action of choosing vertex $i \in [k]$ and coloring it with color $\ell \in [N]$ (where this color can already exist in the clique, as per Definition 15) will be denoted by $x^{i,\ell}$. Namely

$$x^{i,\ell} := \begin{cases} (\ldots, x_{i-1}, \ell, x_{i+1}, \ldots) & \text{if } \ell \notin x \\ (\ldots, x_{j-1}, x_i, x_{j+1} \ldots, x_{i-1}, x_j, x_{i+1}, \ldots) & \text{if } \ell = x_j. \end{cases}$$

Let $f : \Theta_{k,N} \to \mathbb{R}$ be a function on the state space of this chain. Since the stationary distribution is the uniform, the expectation of f over its state space is

$$\mathop{\mathbf{E}}_{\Theta_{k,N}}[f] := \frac{1}{|\Theta_{k,N}|} \sum_{x \in \Theta_{k,N}} f(x).$$

8

Moreover, the Dirichlet form of this chain can be written as

$$\mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}) = \frac{1}{2} \mathop{\mathbf{E}}_{x \in \Theta_{k,N}} \left[\mathop{\mathbf{E}}_{i \in [k]} \left[\mathop{\mathbf{E}}_{\ell \in [N]} \left[\left(\sqrt{f(x^{i,\ell})} - \sqrt{f(x)} \right)^2 \right] \right] \right]$$

$$= \frac{1}{2kN \cdot |\Theta_{k,N}|} \sum_{x \in \Theta_{k,N}} \sum_{i \in [k]} \sum_{\ell \in [N]} \left(\sqrt{f(x^{i,\ell})} - \sqrt{f(x)} \right)^2 .$$

With this notation in mind, we now prove that this Markov chain has a large log-Sobolev constant.

Lemma 18. The log-Sobolev constant of the uniform k-clique N-coloring Markov chain satisfies

$$\alpha(P_{k,N}^{\mathsf{ucc}}) \geq \frac{1}{12k \log N}$$

when $k \leq N/2$.

Proof. Our starting point is the recursive structure of the uniform clique coloring problem, which allows us to apply the martingale method of [LY98]. In particular, let x be uniformly distributed over the state space $\Theta_{k,N}$. Then if we condition on the ith vertex having color ℓ, the distribution of the colors of the remaining $k-1$ vertices is isomorphic to the uniform distribution over $\Theta_{k-1,N-1}$, the state space of the uniform $(k-1)$-clique $(N-1)$-coloring Markov chain.

For any vertex $i \in [k]$ and color $c \in [N]$ define the *conditional function*

$$f_{i,c} : \left\{ (x_1, \ldots, x_k) \in \Theta_{k,N} : x_i = c \right\} \to \mathbb{R}$$

to be simply the restriction of f to this domain: $f_{i,c}(x) = f(x)$ for all $x \in \Theta_{k,N}$ with $x_i = c$. Since $\left\{ (x_1, \ldots, x_k) \in \Theta_{k,N} : x_i = c \right\}$ is isomorphic to $\Theta_{k-1,N-1}$, by a slight abuse of notation we also regard $f_{i,c} : \Theta_{k-1,N-1} \to \mathbb{R}$.

Moreover, for every vertex $i \in [k]$, define the *marginal function* $F_i : [N] \to \mathbb{R}$ by defining for every color $c \in [N]$

$$F_i(c) := \mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = c}} [f(\boldsymbol{x})].$$

The *chain rule of conditional entropy* ([Sal20], Equation 13) implies that for any $i \in [k]$,

$$\mathsf{Ent}(f) = \mathop{\mathbf{E}}_{\boldsymbol{c}}[\mathsf{Ent}(f_{i,\boldsymbol{c}})] + \mathsf{Ent}(F_i). \tag{1}$$

By summing over all vertices $i \in [k]$, we get

$$k \cdot \mathsf{Ent}(f) = \sum_{i \in [k]} \mathop{\mathbf{E}}_{\boldsymbol{c}_i}[\mathsf{Ent}(f_{i,\boldsymbol{c}_i})] + \sum_{i \in [k]} \mathsf{Ent}(F_i). \tag{2}$$

We bound the two summations of the right-hand side separately in Claim 19 and Claim 20 and conclude that

$$k \cdot \mathsf{Ent}(f) \leq \frac{kN}{N-1} \cdot \alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1} \cdot \mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}) + 3k \log N \cdot \mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}).$$

$$\implies \mathsf{Ent}(f) \leq \left[\frac{N}{N-1} \cdot \alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1} + 3 \log N \right] \cdot \mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}).$$

9

This gives us a recurrence relation for the log-Sobolev constant of the uniform clique coloring chain. For every k and N, we have

$$\alpha(P_{k,N}^{\mathsf{ucc}})^{-1} \leq \frac{N}{N-1} \cdot \alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1} + 3\log N. \tag{3}$$

We proceed to solve this recurrence via induction. For fixed integers $k_{\max}$ and $N_{\max}$, we will prove that for all $1 \leq k \leq k_{\max}$,

$$\alpha(P_{k,N_{\max}-k_{\max}+k}^{\mathsf{ucc}})^{-1} \leq 6 \cdot \frac{N_{\max} - k_{\max} + k}{N_{\max} - k_{\max}} \cdot k \log N_{\max}.$$

For the base case of $k = 1$, we observe that uniform 1-clique $(N_{\max} - k_{\max} + 1)$-coloring has transition probabilities that correspond to the complete graph over $N_{\max} - k_{\max} + 1$ vertices. We use known results for the log-Sobolev constant of the complete graph ([DSC96], Corollary A.4) to deduce that

$$\alpha(P_{1,N_{\max}-k_{\max}+1}^{\mathsf{ucc}})^{-1} \leq 3\log(N_{\max} - k_{\max} + 1) \leq 6\log N_{\max}.$$

Now let $k \geq 2$ and assume that the claim holds for all $k' \leq k$. Then using Equation (3) we find

$$\alpha(P_{k,N_{\max}-k_{\max}+k}^{\mathsf{ucc}})^{-1} \leq \frac{N_{\max} - k_{\max} + k}{N_{\max} - k_{\max} + k - 1} \cdot \alpha(P_{k-1,N_{\max}-k_{\max}+k-1}^{\mathsf{ucc}})^{-1} + 3\log(N_{\max} - k_{\max} + k)$$

$$= 6 \cdot \frac{N_{\max} - k_{\max} + k}{N_{\max} - k_{\max}} \cdot (k-1)\log N_{\max} + 3\log(N_{\max} - k_{\max} + k)$$

$$\leq 6 \cdot \frac{N_{\max} - k_{\max} + k}{N_{\max} - k_{\max}} \cdot k \log N_{\max}.$$

In the above calculation, we used the fact that $k_{\max} \leq N_{\max}/2$, and that $N_{\max}$ is at least some fixed constant. This finishes the inductive proof, and by setting $k = k_{\max}$ we obtain the desired bound. $\qquad\square$

It remains to prove the two claims used in the proof of Lemma 18.

Claim 19. For any $f : \Theta_{k,N} \to \mathbb{R}$ we have

$$\sum_{i\in[k]} \mathop{\mathbf{E}}_{\mathbf{c}_i} \left[\mathsf{Ent}(f_{i,\mathbf{c}_i})\right] \leq \frac{kN}{N-1} \cdot \alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1} \cdot \mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}).$$

Proof. Recall that when we condition f on vertex i having color c_i, its domain is isomorphic to the state space of the uniform $(k-1)$-clique $(N-1)$-coloring chain. The log-Sobolev constant of this smaller restricted chain implies that

$$\mathsf{Ent}(f_{i,c_i}) \leq \alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1} \cdot \mathcal{E}_{P_{k-1,N-1}^{\mathsf{ucc}}}\left(\sqrt{f_{i,c_i}}, \sqrt{f_{i,c_i}}\right).$$

Our goal is to relate the Dirichlet form of $P_{k-1,N-1}^{\mathsf{ucc}}$ to the Dirichlet form of $P_{k,N}^{\mathsf{ucc}}$. We start by expanding the right-hand side while keeping in mind that f_{i,c_i} has fixed the color of vertex i to c_i.

$$\mathsf{Ent}(f_{i,c_i}) \leq \frac{\alpha(P_{k-1,N-1}^{\mathsf{ucc}})^{-1}}{2(N-1)(k-1)|\Theta_{k-1,N-1}|} \sum_{\substack{x\in\Theta_{k,N} \\ x_i=c_i}} \sum_{\substack{j\in[k] \\ j\neq i}} \sum_{\substack{\ell\in[N] \\ \ell\neq c_i}} \left(\sqrt{f(x^{j,\ell})} - \sqrt{f(x)}\right)^2$$

10

Let us take the expectation now over all N values of c_i. We note that the log-Sobolev of $P^{\text{ucc}}_{k-1,N-1}$ is not dependent on the value of c_i due to symmetry, thus we factor it outside the summation.

$$\mathop{\mathbf{E}}_{c_i}\left[\text{Ent}(f_{i,c_i})\right] \leq \frac{\alpha(P^{\text{ucc}}_{k-1,N-1})^{-1}}{2N(N-1)(k-1)|\Theta_{k-1,N-1}|} \sum_{c_i\in[N]} \sum_{\substack{x\in\Theta_{k,N}\\ x_i=c_i}} \sum_{\substack{j\in[k]\\ j\neq i}} \sum_{\substack{\ell\in[N]\\ \ell\neq c_i}} \left(\sqrt{f(x^{j,\ell})}-\sqrt{f(x)}\right)^2.$$

Summing over all $i\in[k]$ yields the following

$$\sum_{i\in[k]}\mathop{\mathbf{E}}_{c_i}\left[\text{Ent}(f_{i,c_i})\right] \leq \frac{\alpha(P^{\text{ucc}}_{k-1,N-1})^{-1}}{2N(N-1)(k-1)|\Theta_{k-1,N-1}|} \sum_{i\in[k]}\sum_{c_i\in[N]} \sum_{\substack{x\in\Theta_{k,N}\\ x_i=c_i}} \sum_{\substack{j\in[k]\\ j\neq i}} \sum_{\substack{\ell\in[N]\\ \ell\neq c_i}} \left(\sqrt{f(x^{j,\ell})}-\sqrt{f(x)}\right)^2.$$

Notice that each tuple x is counted k times in the summation of the right-hand side, one time for each (i,c_i) that satisfies $c_i=x_i$. Then each $\left(\sqrt{f(x^{j',\ell})}-\sqrt{f(x)}\right)^2$ term appears at most $(k-1)$ times, since out of the k times that x appears, one of them satisfies $j'=i$, and thus it does not contribute to the sum.

This implies that the sum above is at most $(k-1)$ times the summation that corresponds to the Dirichlet form of $\mathcal{E}_{P^{\text{ucc}}_{k,N}}$.

$$\sum_{i\in[k]}\mathop{\mathbf{E}}_{c_i}\left[\text{Ent}(f_{i,c_i})\right] \leq \frac{\alpha(P^{\text{ucc}}_{k-1,N-1})^{-1}}{2N(N-1)(k-1)|\Theta_{k-1,N-1}|} \cdot (k-1)\cdot 2kN\cdot|\Theta_{k,N}|\cdot \mathcal{E}_{P^{\text{ucc}}_{k,N}}(\sqrt{f},\sqrt{f})$$

$$= \frac{kN}{N-1}\cdot\alpha(P^{\text{ucc}}_{k-1,N-1})^{-1}\cdot\mathcal{E}_{P^{\text{ucc}}_{k,N}}(\sqrt{f},\sqrt{f}).$$

$\square$

Claim 20. Let $f:\Theta_{k,N}\to\mathbb{R}$ be a function, and for all $i\in[k]$, $F_i:[N]\to\mathbb{R}$ is the ith marginal function of f that maps color c to $F_i(c) := \mathbf{E}_{x\in\Theta_{k,N},x_i=c}[f(x)]$. Then it holds that

$$\sum_{i=1}^{k}\text{Ent}(F_i)\leq k\log N\cdot\mathcal{E}_{P^{\text{ucc}}_{k,N}}(\sqrt{f},\sqrt{f}).$$

Proof. Consider the random walk on the set $[N]$ of colors where at every step we move to a uniformly random color (including the color we are currently in). The transition matrix of this walk is the complete graph over N vertices and we denote it by P^{compl}_N. Let us apply the log-Sobolev inequality of P^{compl}_N to the function F_i:

$$\text{Ent}\,(F_i)\leq\alpha(P^{\text{compl}}_N)^{-1}\cdot\mathcal{E}_{P^{\text{compl}}_N}\left(\sqrt{F_i},\sqrt{F_i}\right)$$

$$= \frac{\alpha(P^{\text{compl}}_N)^{-1}}{2N^2}\cdot\sum_{\ell\in[N]}\sum_{\ell'\in[N]}\left(\sqrt{F_i(\ell')}-\sqrt{F_i(\ell)}\right)^2. \tag{4}$$

We would like to rewrite the Dirichlet form of P^{compl}_N in terms of $P^{\text{ucc}}_{k,N}$. We start by expanding the definition of F_i

$$\left(\sqrt{F_i(\ell')}-\sqrt{F_i(\ell)}\right)^2 = \left(\sqrt{\mathop{\mathbf{E}}_{\substack{x\in\Theta_{k,N}\\ x_i=\ell'}}[f(x)]}-\sqrt{\mathop{\mathbf{E}}_{\substack{x\in\Theta_{k,N}\\ x_i=\ell}}[f(x)]}\right)^2.$$

11

Observe that sampling a random $\boldsymbol{x} \in \Theta_{k,N}$ such that $\boldsymbol{x}_i = \ell'$, is equivalent to sampling a random $\boldsymbol{x}$ with $\boldsymbol{x}_i = \ell$, and then outputting $\boldsymbol{x}^{i,\ell'}$:

$$\left(\sqrt{F_i(\ell')} - \sqrt{F_i(\ell)}\right)^2 = \left(\sqrt{\mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = \ell}}\left[f(\boldsymbol{x}^{i,\ell'})\right]} - \sqrt{\mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = \ell}}\left[f(\boldsymbol{x})\right]}\right)^2 .$$

Since the function on the right-hand side is convex, Jensen's inequality implies that

$$\left(\sqrt{F_i(\ell')} - \sqrt{F_i(\ell)}\right)^2 \leq \mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = \ell}}\left[\left(\sqrt{f(\boldsymbol{x}^{i,\ell'})} - \sqrt{f(\boldsymbol{x})}\right)^2\right].$$

Plugging in the above inequality to Equation (4) we get

$$\mathsf{Ent}(F_i) \leq \frac{\alpha(P_N^{\mathsf{compl}})^{-1}}{2N^2} \sum_{\ell \in [N]} \sum_{\ell' \in [N]} \mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = \ell}}\left[\left(\sqrt{f(\boldsymbol{x}^{i,\ell'})} - \sqrt{f(\boldsymbol{x})}\right)^2\right].$$

We sum over all $i \in [k]$ to get

$$\sum_{i=1}^{k}\mathsf{Ent}(F_i) \leq \frac{\alpha(P_N^{\mathsf{compl}})^{-1}}{2N^2} \sum_{\ell \in [N]} \sum_{\ell' \in [N]} \sum_{i \in [k]} \mathop{\mathbf{E}}_{\substack{\boldsymbol{x} \in \Theta_{k,N} \\ \boldsymbol{x}_i = \ell}}\left[\left(\sqrt{f(\boldsymbol{x}^{i,\ell'})} - \sqrt{f(\boldsymbol{x})}\right)^2\right]$$

$$= \frac{\alpha(P_N^{\mathsf{compl}})^{-1}}{2N|\Theta_{k,N}|} \sum_{\ell \in [N]} \sum_{\ell' \in [N]} \sum_{i \in [k]} \sum_{\substack{x \in \Theta_{k,N} \\ x_i = \ell}}\left[\left(\sqrt{f(x^{i,\ell'})} - \sqrt{f(x)}\right)^2\right].$$

The right-hand side now contains all $\left(\sqrt{f(x^{i,\ell'})} - \sqrt{f(x)}\right)^2$ terms that appear in $\mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}$ exactly once. Thus we can substitute this Dirichlet form (and adjust its scaling). Moreover, the log-Sobolev constant of the complete graph over N vertices is well-studied and satisfies $\alpha(P_N^{\mathsf{compl}})^{-1} \leq 3 \cdot \log N$ ([DSC96], Corollary A.4). We conclude that

$$\sum_{i=1}^{k}\mathsf{Ent}(F_i) \leq 3k \log N \cdot \mathcal{E}_{P_{k,N}^{\mathsf{ucc}}}(\sqrt{f}, \sqrt{f}). \qquad \square$$

5 The Log-Sobolev Constant of the Standard Clique Coloring Chain

The goal of this section is to translate the log-Sobolev bound from the uniform clique coloring chain Lemma 18 to the standard clique coloring chain. Since the two chains are very similar, applying the comparison method is a natural approach.

Lemma 21. The log-Sobolev constant of the k-clique N-coloring Markov chain satisfies

$$\alpha(P_{k,N}^{\mathsf{cc}}) \geq \frac{1}{19} \cdot \alpha(P_{k,N}^{\mathsf{ucc}}).$$

Proof. Define the following (randomized) map $\boldsymbol{\Delta}$ that maps edges of $P_{k,N}^{\mathsf{ucc}}$ to paths in $P_{k,N}^{\mathsf{cc}}$. Each edge of $P_{k,N}^{\mathsf{ucc}}$ that connects x and $x^{i,\ell}$ is determined by a vertex $x \in \Theta_{k,N}$ and the pair $(i, \ell) \in$

$[k] \times [N]$. We assign to this edge a path in $P^{\mathsf{cc}}_{k,N}$ drawn according to the following distribution:

$$\boldsymbol{\Delta}(x, x^{i,\ell}) = \begin{cases} (x, x^{i,\ell}) & \ell \notin x \setminus \{x_i\}, \\ (x, \underbrace{x^{i,\boldsymbol{\ell}'}}_{y}) \,||\, (y, \underbrace{y^{j,x_i}}_{z}) \,||\, (z, z^{i,x_j}) & \ell = x_j \text{ for } j \neq i, \ \ \boldsymbol{\ell}' \sim [N] \setminus x. \end{cases}$$

Here the symbol "$||$" denotes the concatenation of edges to make a path. Intuitively, the path assigned to edge $(x, x^{i,\ell})$ is either itself (whenever $(x, x^{i,\ell}$ is also an edge of $P^{\mathsf{cc}}_{k,N}$), or a sequence of three edges that swap the colors x_i and x_j by using a random unused color $\boldsymbol{\ell}'$.

Now we bound the comparison constant $A(\boldsymbol{\Delta})$.

$$A(\boldsymbol{\Delta}) = \max_{(a,b) \in E^{\mathsf{cc}}} \left\{ \frac{1}{\pi^{\mathsf{cc}}(x) P^{\mathsf{cc}}(a,b)} \sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \cdot \pi^{\mathsf{ucc}}(x) \cdot P^{\mathsf{ucc}}(x,y) \right\}$$

The stationary distributions of both chains are the uniform over $\Theta_{k,N}$, and thus the stationary probabilities cancel.

$$A(\boldsymbol{\Delta}) = \max_{(a,b) \in E^{\mathsf{cc}}} \left\{ k(N-k+1) \sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \cdot \frac{1}{kN} \right\}$$

$$= \max_{(a,b) \in E^{\mathsf{cc}}} \left\{ \frac{N-k+1}{N} \sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \right\}.$$

Our goal will be to bound the sum of expectations. First, let us partition the paths into the ones with length 1 and length 3. To do that, we observe that the length of each path $\boldsymbol{\Delta}(x,y)$ is deterministic and only depends on x and y.

$$\sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] = \sum_{\substack{(x,y) \in E^{\mathsf{ucc}} \\ |\boldsymbol{\Delta}(x,y)|=1}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \right] + 3 \sum_{\substack{(x,y) \in E^{\mathsf{ucc}} \\ |\boldsymbol{\Delta}(x,y)|=3}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \right].$$

We can now easily bound the first term. For a path with a single edge to include (a,b), it must hold that $(x,y) = (a,b)$. Thus the first term is at most 1. To bound the second term, we consider the location $t \in \{1,2,3\}$ where (a,b) appears in $\boldsymbol{\Delta}(x,y)$. We write $(a,b) = \boldsymbol{\Delta}(x,y)_t$ if (a,b) appears as the tth edge of the path. Formally,

$$\sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \leq 1 + 3 \sum_{t \in \{1,2,3\}} \sum_{\substack{(x,y) \in E^{\mathsf{ucc}} \\ |\boldsymbol{\Delta}(x,y)|=3}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) = \boldsymbol{\Delta}(x,y)_t} \right].$$

Observe now that once we fix the tth edge to be (a,b), there are only $k-1$ possible 3-edge paths. This is because our map $\boldsymbol{\Delta}$ performs three transpositions between the elements $x_i, x_j, \boldsymbol{\ell}'$. The edge (a,b) specifies two of the elements, and the third element is one of the remaining $k-1$ elements of the tuples at the endpoints of (a,b). Once this third element is specified, the edge (x,y) and its respective path $\boldsymbol{\Delta}(x,y)$ is fully determined.

Each 3-edge path has a probability of $\frac{1}{N-k}$ to appear, since it depends on the random choice of $\boldsymbol{\ell}'$ from the set $[N] \setminus x$. Thus we bound the expectation above to be at most

$$\sum_{(x,y) \in E^{\mathsf{ucc}}} \mathbf{E}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \leq 1 + \frac{9(k-1)}{N-k}.$$

13

We conclude that the comparison constant of $\boldsymbol{\Delta}$ is

$$A(\boldsymbol{\Delta}) = \max_{(a,b) \in E^{\mathsf{cc}}} \left\{ \frac{N-k+1}{N} \sum_{(x,y) \in E^{\mathsf{ucc}}} \mathop{\mathbf{E}}_{\boldsymbol{\Delta}} \left[\mathbf{1}_{(a,b) \in \boldsymbol{\Delta}(x,y)} \cdot |\boldsymbol{\Delta}(x,y)| \right] \right\}$$

$$\leq \frac{N-k+1}{N} \left(1 + \frac{9(k-1)}{N-k} \right)$$

$$= \frac{N-k+1}{N} + \frac{9(k-1)}{N} \cdot \frac{N-k+1}{N-k}$$

$$\leq 1 + 9 \cdot 2 = 19. \qquad \square$$

Our log-Sobolev bound for the standard clique-coloring chain now follows directly from Lemma 18 and Lemma 21.

Corollary 22. The log-Sobolev constant of the k-clique N-coloring Markov chain satisfies

$$\alpha(P_{k,N}^{\mathsf{cc}}) \geq \Omega\left(\frac{1}{k \log N} \right).$$

5.1 Clique-Coloring Walk to Random Circuits Walk

We would like to transfer our log-Sobolev constant bound of the k-clique N-coloring Markov chain from Corollary 22, to the random circuits Markov chain. This is done via the randomized paths construction of Brodsky and Hoory to compare this walk to clique coloring.

Lemma 23 ([BH05]). When $k \leq 2^n/3$ there exists a randomized map $\boldsymbol{\Phi}$ that takes as input an edge (x,y) of $P_{k,2^n}^{\mathsf{cc}}$ and outputs a sequence of edges in $P_{k,n}^{\mathsf{rev}}$ connecting x and y such that the comparison constant satisfies

$$A(\boldsymbol{\Phi}) = O(n^2).$$

Corollary 24. If $k \leq 2^n/3$ then

$$\alpha(P_{\mathsf{rev}}) \gtrsim \frac{1}{n^2} \cdot \alpha(P_{\mathsf{cc}}).$$

Proof. This follows immediately from Lemma 23 and Lemma 12. $\qquad \square$

6 Even Faster Mixing of the Random Circuits Walk via Generic States

We can improve the dependence on n of the mixing time of the random reversible circuits Markov chain $P_{k,n}^{\mathsf{rev}}$ from cubic to linear using an idea of [BH05]. The main observation is that after $n \cdot \mathrm{polylog}\,(n,k)$ steps of $P_{k,n}^{\mathsf{rev}}$, the chain is very likely to be in a *generic* state, that is a state where no two of the bit-strings agree on many bits. Generic states happen with good probability and are nicer to work with, thus when we restrict our Markov chain $P_{k,n}^{\mathsf{rev}}$ to generic states we apply the comparison theorem with a better (logarithmic) comparison constant.

Definition 25 (Generic states, [BH05]). Let $w = \lceil 10 \cdot (\log k + \log n) \rceil, p = \lceil \frac{n}{2w} \rceil$. Let $C_1, \cdots C_p, C$ be a partition of $[n]$ such that $|C_t| = w$ for $t \in [p]$, and $|C| = n - pw$. A state $(x_1, \cdots, x_k)$ is *generic* if for $i \neq i'$, x_i and $x_{i'}$ are distinct when restricted to a part C_t (but not C). Let $\mathsf{Generic}_{k,n}$ denote the set of generic states.

In other words, we divide the n bits of the input into two subsets $\bigcup_{t \in [p]} C_t$ and C of roughly equal size. Then we further divide the first subset into p equal-length blocks that hold a logarithmic number of bits. A state is generic if no two distinct elements $x_i, x_{i'}$ are equal in any of the C_t parts. Since we now deal with n-bit strings, we will extend our notation and write $x_{i,j}$ to denote the jth bit of the ith element of the state x.

We define below the *generic state reversible circuit* Markov chain $P^{\mathsf{grev}}_{k,n}$ to be the restriction of $P^{\mathsf{rev}}_{k,n}$ to generic states.

Definition 26 (Generic state reversible circuit Markov chain)**.** The matrix P^{grev} is the transition matrix of the Markov chain on $\mathsf{Generic}_{k,n}$ such that for any $x, y \in \mathsf{Generic}_{k,n}$,

$$P^{\mathsf{grev}}(x,y) = \frac{P^{\mathsf{rev}}(x,y)}{\sum_{z \in \mathsf{Generic}_{k,n}} P^{\mathsf{rev}}(x,z)}.$$

Lemma 27 ([BH05], Equation (3))**.** There exists a constant $\varepsilon > 0$ such that if $\tau_\varepsilon(P^{\mathsf{grev}}) \leq O(n^3 k^3)$, and $k \leq 2^{n/50}$, then

$$\tau(P^{\mathsf{rev}}) \leq \tau_\varepsilon(P^{\mathsf{grev}}) + O(n \cdot \mathrm{polylog}(n,k)).$$

We bound the mixing time of the P^{grev} Markov chain by bounding its log-Sobolev constant. We use the comparison of [BH05] as stated in Lemma 32 to relate its log-Sobolev constant to the log-Sobolev constant of a related product chain on generic states, $\widetilde{P}^{\mathsf{grev}}$. We get our final estimate by bounding the log-Sobolev constant of the $\widetilde{P}^{\mathsf{grev}}$ Markov chain in Lemma 31 using results for product chains from [DSC96].

Below we introduce the $\widetilde{P}^{\mathsf{grev}}$ Markov chain.

Definition 28 (Product chain on generic states)**.** Let $\widetilde{P}^{\mathsf{grev}}$ be the Markov chain on state space $\mathsf{Generic}_{k,n}$, where to sample the next state $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ given the current state $x = (x_1, \ldots, x_k) \in \mathsf{Generic}_{k,n}$ we do the following:

- With probability $\frac{1}{2}$, toss a fair coin.

  - If the coin has landed heads, set $\boldsymbol{y} = x$.
  - Else, sample uniformly at random $\boldsymbol{c} \sim C, \boldsymbol{r} \sim [k]$ and set for all $i \in [k]$ and $j \in [n]$

$$\boldsymbol{y}_{i,j} = \begin{cases} x_{i,j} & \text{if } i \neq \boldsymbol{r} \text{ or } j \neq \boldsymbol{c} \\ 1 - x_{i,j} & \text{if } i = \boldsymbol{r} \text{ and } j = \boldsymbol{c}. \end{cases}$$

- With probability $\frac{1}{2}$, sample uniformly at random $\boldsymbol{\ell} \sim [p], \boldsymbol{r} \sim [k]$ and a random string $\boldsymbol{u} \in \{0,1\}^w$ such that $\boldsymbol{u} \neq x_{i,C_\ell}$ for any $i \neq \boldsymbol{r}$. Set

$$\boldsymbol{y}_{i,j} = \begin{cases} x_{i,C_\ell} & \text{if } i \neq \boldsymbol{r} \text{ or } \ell \neq \boldsymbol{\ell} \\ \boldsymbol{u} & \text{if } i = \boldsymbol{r} \text{ and } \ell = \boldsymbol{\ell}. \end{cases}$$

Informally, given the current state x, one step of this Markov chain performs a change in exactly one of the two subsets of bits (C or $\bigcup_{i \in [p]} C_i$) with equal probability. In the first case, it either flips the $\boldsymbol{c}^{th}$ bit from the subset C of a random element $\boldsymbol{r}$ with probability $\frac{1}{2}$, or it does nothing. In the second case, it samples a uniformly random subset of bits C_ℓ and replaces that subset with a new bit string $\boldsymbol{u}$ for a random element $\boldsymbol{r}$. All of the operations above are performed such that the resulting state remains generic.

It is not hard to observe that $\widetilde{P}^{\mathsf{grev}}$ is a *product chain*, that is it acts "independently" on different parts of its state space. This means that we can compute its log-Sobolev constant by breaking it down into smaller chains.

Definition 29 (Product Markov chain)**.** Consider t Markov chains $\{P_i\}_{i\in[t]}$ with state spaces $\{V_i\}_{i\in[t]}$ respectively. We define the *product Markov chain* $\prod\left(\{P_i\}_{i\in[t]}\right)$ over the state space $\prod_{i\in[t]} V_i$ to be the Markov chain with transition matrix

$$\frac{1}{t}\sum_{i\in[t]} I \otimes \cdots \otimes P_i \otimes \cdots \otimes I.$$

We will refer to the P_i's as the *factors* of $\prod\left(\{P_i\}_{i\in[t]}\right)$.

Lemma 30 (Log-Sobolev constant of product chain, Lemma 3.2 of [DSC96])**.** The log-Sobolev constant of the product chain $\prod\left(\{P_i\}_{i\in[t]}\right)$ is related to the log-Sobolev constant of its factors as follows:

$$\alpha\left(\prod\left(\{P_i\}_{i\in[t]}\right)\right) = \frac{1}{t}\min_{i\in[t]}\alpha(P_i).$$

Using Lemma 30 we obtain the following bound by decomposing $\widetilde{P}^{\mathsf{grev}}$ into factor chains whose log-Sobolev constants are known.

Lemma 31. The following bound on the log-Sobolev constant of $\widetilde{P}^{\mathsf{grev}}$ holds:

$$\alpha\left(\widetilde{P}^{\mathsf{grev}}\right) \geq \Omega\left(\frac{1}{nk}\right).$$

Proof. We first write the state space $\mathsf{Generic}_{k,n}$ in the form of a product

$$\mathsf{Generic}_{k,n} = \left(\prod_{i\in[p]} \Theta_{k,\{0,1\}^w}\right) \times \left(\{0,1\}^{k(n-wp)}\right).^2$$

Then decompose $\widetilde{P}^{\mathsf{grev}}$ as the product of two Markov chains $\prod\left(\left\{\widetilde{P}_1,\widetilde{P}_2\right\}\right)$. The first chain $\widetilde{P}_1$ corresponds to performing a change in the $\bigcup_{i\in[p]} C_i$ subset of the bits, and the second chain $\widetilde{P}_2$ corresponds to operating in the C subset of the bits.

The chain $\widetilde{P}_1$. The state space of this chain is $\prod_{i\in[p]} \Theta_{k,\{0,1\}^w}$. We further decompose[3] this chain as $\widetilde{P}_1 = \prod\left(\left\{\widetilde{P}_{1,\ell}\right\}_{\ell\in[p]}\right)$, where $\tilde{P}_{1,\ell}$ corresponds to performing an operation on the C_ℓ subset of the bits. Thus the chain $\widetilde{P}_{1,\ell}$ has state space $\Theta_{k,\{0,1\}^w}$, since it corresponds to the size-w subset C_ℓ. To sample the next state $\boldsymbol{y} = (\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k)$ from the current state $\boldsymbol{x} = (x_1,\ldots,x_k)$, we choose a random $\boldsymbol{i}\in[k]$ and a random $\boldsymbol{z}\in\{z\in\{0,1\}^w \mid z\notin x\}\cup\{x_{\boldsymbol{i}}\}$ and set for each $j\in[k]$

$$\boldsymbol{y}_j = \begin{cases} x_j & \text{if } j\neq\boldsymbol{i}. \\ \boldsymbol{z} & \text{if } j=\boldsymbol{i}. \end{cases}$$

---

[2]Recall from Definition 3 that $\Theta_{k,\{0,1\}^w}$ denotes the set of k-tuples of distinct elements of $\{0,1\}^w$.

[3]We don't directly decompose $\widetilde{P}^{\mathsf{grev}}$ into all of its $t+1$ factors because to use Lemma 30 we need each factor of the product chain to have equal weight.

Notice that the transition matrix of this chain is equal to the transition matrix $P_{k,\{0,1\}^w}^{\mathsf{cc}}$ of the standard k-clique 2^w-coloring chain. Therefore, by Corollary 22, we have for all $\ell \in [p]$ that

$$\alpha\left(\widetilde{P}_{1,\ell}\right) \gtrsim \frac{1}{k \log |\{0,1\}^w|} = \frac{1}{kw}.$$

Applying Lemma 30, we have

$$\alpha\left(\widetilde{P}_1\right) = \frac{1}{p} \min_{\ell \in [p]} \alpha\left(\widetilde{P}_{1,\ell}\right) \gtrsim \frac{1}{p} \cdot \frac{1}{kw} \gtrsim \frac{1}{nk}. \tag{5}$$

The chain $\widetilde{P}_2$. We will "flatten" the bits from the subset C of the k elements into a sequence of $k(n-wp)$ bits. Then the $\widetilde{P}_2$ Markov chain corresponds to the random walk on the hypercube $\{0,1\}^{k(n-wp)}$ where to sample the next state $\boldsymbol{y}$ from the current state x we sample $\boldsymbol{i} \in [k(n-wp)]$ uniformly at random and flip the $\boldsymbol{i}^{th}$ bit with probability $\frac{1}{2}$. This chain is the product chain of $k(n-wp)$ chains on the space $\{0,1\}$ with transition probabilities $\frac{1}{2}$ to each state. We can write the transition matrix of $\widetilde{P}_2$ as the product

$$\prod\left(\left\{\widetilde{P}_{2,\ell}\right\}_{\ell \in [k(n-wp)]}\right),$$

where each $\widetilde{P}_{2,\ell}$ is the 2×2 matrix with $\frac{1}{2}$'s. Equivalently, it corresponds to the transition matrix of the complete graph on two states. It is easy to see (e.g. [DSC96], Corollary A.4) that $\alpha(\widetilde{P}_{2,\ell}) \geq \frac{1}{3}$ for all ℓ. Therefore, by Lemma 30 we have

$$\alpha\left(\widetilde{P}_2\right) = \frac{1}{k(n-wp)} \min_{\ell \in [k(n-wp)]} \alpha\left(\widetilde{P}_{2,\ell}\right) \gtrsim \frac{1}{k(n-wp)} \gtrsim \frac{1}{nk}. \tag{6}$$

Applying Lemma 30 with Equation (5) and Equation (6) yields

$$\alpha\left(\widetilde{P}^{\mathsf{grev}}\right) = \frac{1}{2} \min\left\{\alpha\left(\widetilde{P}_1\right), \alpha\left(\widetilde{P}_2\right)\right\} = \Omega\left(\frac{1}{nk}\right). \qquad \square$$

Armed with the log-Sobolev constant of $\widetilde{P}^{\mathsf{grev}}$, we employ the comparison method of [BH05] to bound the log-Sobolev constant of $\widetilde{P}^{\mathsf{rev}}$.

Lemma 32 ([BH05], Lemma 16)**.** There exists a randomized map $\boldsymbol{\Psi}$ that takes as input an edge (x,y) of $\widetilde{P}^{\mathsf{grev}}$ and outputs a sequence of edges in P^{grev} connecting x and y with congestion $A(\boldsymbol{\Psi}) = \mathrm{polylog}(n,k)$. Consequently,

$$\alpha(P^{\mathsf{grev}}) \geq \frac{\alpha\left(\widetilde{P}^{\mathsf{grev}}\right)}{\mathrm{polylog}(n,k)}.$$

Corollary 33. It holds that

$$\alpha(P_{\mathsf{grev}}) \gtrsim \frac{1}{nk \cdot \mathrm{polylog}(n,k)}$$

Using now the well-known relation between the log-Sobolev constant and the mixing time of a Markov chain in total variation distance, we conclude:

Theorem 2. For any n and $k \leq 2^{n/50}$, a random reversible circuit with $\widetilde{O}(nk \cdot \log(1/\varepsilon))$ width-2 gates (a subset of 3-bit gates) computes an ε-approximate k-wise independent permutation, where the $\widetilde{O}$ hides $\mathrm{polylog}(n,k)$ factors.

Proof. Combining Lemma 31 and Lemma 32 we find that $\alpha(P^{\mathsf{grev}}) \geq \Omega\left(\frac{1}{nk}\right)$. This implies that for the constant $\varepsilon' > 0$ referenced in Lemma 27, we have $\tau_{\varepsilon'}(P^{\mathsf{grev}}) \leq O(nk \cdot \mathrm{polylog}(n,k))$. Then applying Lemma 27 we have

$$\tau(P^{\mathsf{rev}}) \leq \tau_{\varepsilon'}(P^{\mathsf{grev}}) + O(nk \cdot \mathrm{polylog}(n,k)) \leq O(nk \cdot \mathrm{polylog}(n,k)).$$

Finally, we can decrease the total variation distance down to an arbitrary $\varepsilon > 0$ by increasing the length of the walk by a multiplicative factor of $O(\log(1/\varepsilon))$, and the statement follows. $\qquad\square$

Acknowledgments

References

[AL13]   Noga Alon and Shachar Lovett. Almost k-wise vs. k-wise independent permutations, and uniformity for general group actions. *Theory of Computing*, 9(15):559–577, 2013.

[BCHJ+21] Fernando G.S.L. Brandão, Wissam Chemissany, Nicholas Hunter-Jones, Richard Kueng, and John Preskill. Models of quantum complexity growth. *PRX Quantum*, 2(3), July 2021.

[BH05]   Alex Brodsky and Shlomo Hoory. Simple permutations mix even better, 2005.

[BHH16]   Fernando GSL Brandao, Aram W Harrow, and Michał Horodecki. Local random quantum circuits are approximate polynomial-designs. *Communications in Mathematical Physics*, 346:397–434, 2016.

[CBB+24]   Chi-Fang Chen, Adam Bouland, Fernando G. S. L. Brandão, Jordan Docter, Patrick Hayden, and Michelle Xu. Efficient unitary designs and pseudorandom unitaries from permutations, 2024.

[CCMR24]   Ran Canetti, Claudio Chamon, Eduardo Mucciolo, and Andrei Ruckenstein. Towards general-purpose program obfuscation via local mixing. Cryptology ePrint Archive, Paper 2024/006, 2024. https://eprint.iacr.org/2024/006.

[DSC93a]   Persi Diaconis and Laurent Saloff-Coste. Comparison Techniques for Random Walk on Finite Groups. *The Annals of Probability*, 21(4):2131 – 2156, 1993.

[DSC93b]   Persi Diaconis and Laurent Saloff-Coste. Comparison Theorems for Reversible Markov Chains. *The Annals of Applied Probability*, 3(3):696 – 730, 1993.

[DSC96]   P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695 – 750, 1996.

[Gow96]   W Timothy Gowers. An almost m-wise independent random permutation of the cube. *Combinatorics, Probability and Computing*, 5(2):119–130, 1996.

[HHJ21]   Jonas Haferkamp and Nicholas Hunter-Jones. Improved spectral gaps for random quantum circuits: large local dimensions and all-to-all interactions. *Physical Review A*, 104(2):022417, 2021.

[HMMR05] Shlomo Hoory, Avner Magen, Steven Myers, and Charles Rackoff. Simple permutations mix well. *Theoretical Computer Science*, 348(2):251–261, 2005. Automata, Languages and Programming: Algorithms and Complexity (ICALP-A 2004).

[HO24]   William He and Ryan O'Donnell. Pseudorandom permutations from random reversible circuits. *arXiv preprint arXiv:2404.14648*, 2024.

[Jer03]   Mark Jerrum. *Counting, sampling and integrating: algorithms and complexity.* Springer Science & Business Media, 2003.

[KNR09]   Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. *Algorithmica*, 55(1):113–133, 2009.

[LY98]   Tzong-Yow Lee and Horng-Tzer Yau. Logarithmic sobolev inequality for some models of random walks. *The Annals of Probability*, 26(4):1855–1873, 1998.

[MOP20]   Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 510–523, 2020.

[MPSY24]   Tony Metger, Alexander Poremba, Makrand Sinha, and Henry Yuen. Simple constructions of linear-depth t-designs and pseudorandom unitaries, 2024.

[Sal20]   Justin Salez. A sharp log-sobolev inequality for the multislice, 2020.

[SC97]   Laurent Saloff-Coste. *Lectures on finite Markov chains*, pages 301–413. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.

[WLP09]   EL Wilmer, David A Levin, and Yuval Peres. Markov chains and mixing times. *American Mathematical Soc., Providence*, 2009.