Pseudorandom Permutations from Random Reversible Circuits

William He* Ryan O'Donnell*

April 24, 2024

Abstract

We study pseudorandomness properties of permutations on $\{0,1\}^n$ computed by random circuits made from reversible 3-bit gates (permutations on $\{0,1\}^3$). Our main result is that a random circuit of depth $n \cdot \widetilde{O}(k^2)$, with each layer consisting of $\approx n/3$ random gates in a fixed nearest-neighbor architecture, yields almost k-wise independent permutations. The main technical component is showing that the Markov chain on k-tuples of n-bit strings induced by a single random 3-bit nearest-neighbor gate has spectral gap at least $1/n \cdot \widetilde{O}(k)$. This improves on the original work of Gowers [Gow96], who showed a gap of $1/\mathrm{poly}(n,k)$ for one random gate (with non-neighboring inputs); and, on subsequent work [HMMR05, BH08] improving the gap to $\Omega(1/n^2k)$ in the same setting.

From the perspective of cryptography, our result can be seen as a particularly simple/practical block cipher construction that gives provable statistical security against attackers with access to k input-output pairs within few rounds. We also show that the Luby–Rackoff construction of pseudorandom permutations from pseudorandom functions can be implemented with reversible circuits. From this, we make progress on the complexity of the Minimum *Reversible* Circuit Size Problem (MRCSP), showing that block ciphers of fixed polynomial size are computationally secure against arbitrary polynomial-time adversaries, assuming the existence of one-way functions (OWFs).

*Computer Science Department, Carnegie Mellon University. {wrhe,odonnell}@cs.cmu.edu. Supported in part by ARO grant W911NF2110001.

1 Introduction

Motivated by questions in the analysis of practical cryptosystems (block ciphers), derandomization, and spectral graph theory, we study pseudorandomness properties of random reversible circuits.[1] That is, we study the indistinguishability of truly random permutations on $\{0,1\}^n$ versus permutations computed by small, randomly chosen reversible circuits.

Our main results concern the extent to which small random reversible circuits compute *almost k-wise independent* permutations. This corresponds to *statistical security* against adversaries that get k input-output pairs from the permutation. We also study *computational security*, showing that it can hold (assuming OWFs exist) even for adversaries running in time significantly exceeding the size of the reversible circuit.

1.1 Basic definitions

We start by defining reversible computation, which was initially studied with a motivation from physics [Lan, Lec63, Ben73, Tof80, FT82], and subsequently played an important role in the foundations of quantum computing.

Definition 1. A Boolean circuit is called *reversible* if it has n wires (hence n inputs and outputs), and each gate operates on some $c \ll n$ wires, computing a permutation on $\{0,1\}^c$. We say such a gate is of type[2] DES$[c-1]$ if (up to ordering the c bits) it is of the form $(x,b) \mapsto (x, b \oplus f(x))$ for some $f : \{0,1\}^{c-1} \to \{0,1\}$. The *size* of the circuit is the number of gates, and the *depth* is the maximum number of gates through which any wire passes.

We focus on DES[2] gates, as they are universal for reversible computation:

Theorem 2. ([CG75], cf. [EG83].) Any permutation in $\mathfrak{A}_{\{0,1\}^n}$ is computable by a reversible circuit with DES[2] gates. (Here $\mathfrak{A}_N$ denotes the alternating group of even permutations on N elements, and $\mathfrak{S}_N$ is the full permutation group.[3])

Of course, by a simple counting argument, almost all even permutations require reversible circuits of size $\widetilde{\Omega}(2^n)$

An important focus of attention in cryptography is on having super-efficient pseudorandom permutations (block ciphers) that are implementable in hardware. We introduce here some particularly simple kinds of circuits:

Definition 3. A reversible circuit has *(1D) nearest-neighbor gates* if each gate operates on c *consecutive* wires from $[n]$. The circuit is called *(3-bit) brickwork* if it has the following architecture: There are ℓ *layers* of gates, each layer having depth 3 and size (slightly less than) n. Within each layer with probability $1/2$ there are $\lfloor n/3 \rfloor$ gates on wires $\{1,2,3\}$, $\{4,5,6\}, \ldots$; with probability $1/2$ there are $\lfloor (n-1)/3 \rfloor$ gates on wires $\{2,3,4\}$, $\{5,6,7\}, \ldots$ (See Figure 1.)

Remark 4. We emphasize that a DES[2]-brickwork circuit is determined only by the parameters ℓ, n, the DES[2]-permutations used at each gate, and the "parity" of the layers.

[1] These have even been studied as toy models for black holes [HP07, Section 2].

[2] This concept has many names: Feistel, Toffoli, "width-$(c-1)$", "standard quantum oracle for f"... The terminology DES$[c-1]$ refers to the use of such permutations in the DES/AES block ciphers; we are following [Cle90]'s notation.

[3] The reader is advised that the technical distinction between $\mathfrak{A}_{\{0,1\}^n}$ and $\mathfrak{S}_{\{0,1\}^n}$ never plays an important role in this work.

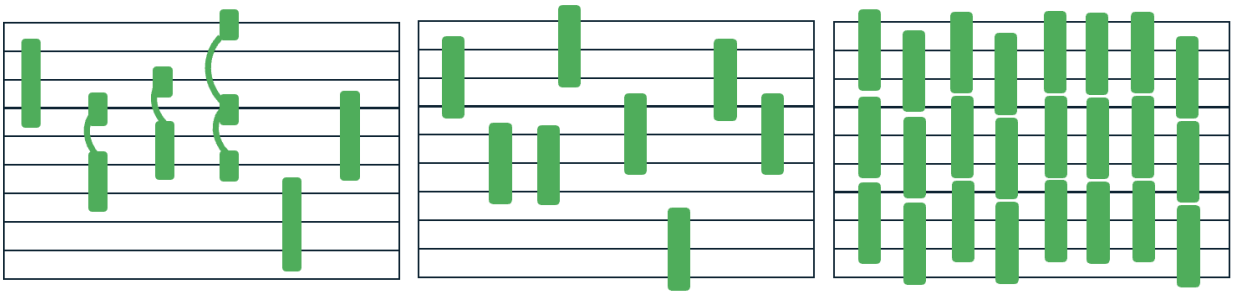

Figure 1: The first circuit is an example of a circuit with generic 3-bit gates. The second is an example of a circuit with 1D nearest-neighbor 3-bit gates. The third is an example of a brickwork circuit with 3 layers.

Brickwork circuits would be very desirable for block ciphers, as they potentially have small depth and their nearest-neighbor gates are much easier to implement. These design considerations have motivated substitution-permutation networks (SPNs) used in practice, such as AES.

We now define a notion of statistical indistinguishability for permutations:

Definition 5. Let $1 \leq k \leq N-2$, and let $\mathcal{P}$ be a probability distribution on $\mathfrak{S}_N$. We say that $\mathcal{P}$ is an *ε-approximate k-design* for $\mathfrak{S}_N$[4] if

$$\left\| \mathop{\mathbf{E}}_{\mathbf{g} \sim \mathcal{P}} \left[\mathbf{g}^{\otimes k} \right] - \mathop{\mathbf{E}}_{\mathbf{g} \sim \mathrm{Unif}(\mathfrak{S}_N)} \left[\mathbf{g}^{\otimes k} \right] \right\|_{\mathrm{op}} \leq \varepsilon.$$

Here $g^{\otimes k}$ denotes the k-fold Kronecker product of g when viewed as a permutation matrix.

Remark 6. The above definition is a rather minimal way to say that a $\mathcal{P}$-random permutation looks random when restricted to its action on size-k subsets of $[N]$. For applications, one may wish for a stricter matrix norm than operator norm; e.g., the 1-norm. As all N^k-dimensional norms are equivalent up to factors of $N^{O(k)}$, one can pass to different norms by targeting designs with $\varepsilon \ll N^{-O(k)}$, which is $2^{-O(nk)}$ in our setting of $\mathfrak{S}_{\{0,1\}^n}$.

A slight variant of this definition is popular in the pseudorandomness literature:

Definition 7. $\mathcal{P}$ is said to be *δ-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 δ from the uniform distribution on distinct k-tuples from $[N]$.

Remark 8. As there are N^k distinct k-tuples on $[N]$, the above definition is not especially strong unless $\delta \ll N^{-k}$. Also, per Remark 6, if $\mathcal{P}$ is a $\delta N^{-O(k)}$-approximate k-design, then it is δ-approximate k-wise independent. Thus the two notions are not very different in the typical error parameter setting of $N^{-O(k)} = 2^{-O(nk)}$.

Approximate k-wise independent permutation distributions $\mathcal{P}$ have many applications outside of cryptography; derandomization, for example [MOP20]. In such applications, another important parameter is the number of truly random "seed" bits needed to generate a draw from $\mathcal{P}$. By using techniques such as derandomized squaring, one can generally reduce the seed length to $O(nk)$ for any construction; see [KNR09]. This is true for the results in our paper, and we don't discuss this angle further. We are generally focused on the circuit complexity of our permutations.

[4] One may check that $k \leq N-2$ implies that it doesn't matter if we write $\mathfrak{A}_N$ here.

1.2 Our results on k-wise independence

Our main work is to demonstrate that random reversible circuits with not too many gates are ε-approximate k-designs. Perhaps our main result is the following, which shows that with the very simple/practical architecture of DES[2]-brickwork circuits, one can achieve good pseudorandomness even after just $O(n)$ random layers. Note that $\Omega(n)$ layers is obviously minimal, as otherwise the 1st and nth bits in the circuit do not interact.

Theorem 9. For $1 \leq k \leq 2^n - 2$, fix the architecture of n-bit DES[2]-brickwork reversible circuits of depth $n \cdot \widetilde{O}(k^2)$ (and hence size $n^2 \cdot \widetilde{O}(k^2)$). Then if the DES[2]-gates are chosen uniformly at random, the resulting permutations on $\{0,1\}^n$ are $2^{-O(nk)}$-approximate k-wise independent. More generally, they are ε-approximate k-wise independent for depth $(nk + \log(1/\varepsilon)) \cdot \widetilde{O}(k)$.

We are not aware of any such prior work on brickwork or nearest-neighbor gates for implementing permutations, although this has been studied in the quantum setting of random unitary circuits. The best prior result was the following theorem of Brodsky and Hoory for general (non-nearest-neighbor) reversible architecture:

Theorem 10. ([BH08].) Consider the permutation on $\{0,1\}^n$ computed by a reversible circuit of $O(n^3k^2)$ randomly chosen DES[2]-gates (meaning in particular that each gate's 3 fan-in wires are randomly chosen). This is $2^{-O(nk)}$-approximate k-wise independent. More generally, such circuits of size $O(n^2k) \cdot (nk + \log(1/\varepsilon))$ suffices for ε-approximate k-wise independence.

There is a simple reduction from general gates to nearest-neighbor gates that incurs factor-$\Omega(n)$ size blowup. Plugging this into Theorem 10 would yield $2^{-O(nk)}$-approximate k-wise independent permutations formed from nearest-neighbor reversible circuits of size $O(n^4k^2)$. Besides being non-brickwork, this is worse than our Theorem 9 by a factor of about n^2. We should note that Brodsky and Hoory also prove the following:

Theorem 11. ([BH08].) If $k \leq 2^{n/50}$, then Theorem 10 also holds for random reversible circuits of size $\widetilde{O}(n^2k^2 \log(1/\varepsilon))$.

The improved dependence on n in this theorem, namely $\widetilde{O}(n^2)$, is good, but one should caution that the dependence on $\log(1/\varepsilon)$ is *multiplicative*. Thus except in the rather weak case when $\varepsilon \gg 2^{-nk}$, this term introduces a factor of at least nk back into the bound, making it worse than Theorem 10.

All of the difficulty in our main result Theorem 9 comes from analyzing the *spectral gap* of the natural random walk on k-tuples of strings arising from picking *one* random nearest-neighbor gate. We prove:

Theorem 12. Let $\mathcal{P}$ be the distribution on $\mathfrak{S}_{\{0,1\}^n}$ given by choosing one random nearest-neighbor DES[2] gate. Then $\mathcal{P}$ is a $(1-\delta)$-approximate k-design for $\delta \geq 1/(n \cdot \widetilde{O}(k))$.

Given this result, Theorem 9 follows almost directly by using the *detectability lemma* from Hamiltonian complexity theory [AALV09], as in analogous results for unitary designs due to [BHH16]. The intermediate step is again proving the one-step spectral gap lower bound.

Theorem 13. Let $\mathcal{P}$ be the distribution on $\mathfrak{S}_{\{0,1\}^n}$ given by choosing one layer of brickwork gates. Then $\mathcal{P}$ is a $(1-\eta)$-approximate k-design for $\eta \geq 1/\widetilde{O}(k)$.

We also prove an analogous result to Theorem 12 in the case of non-nearest-neighbor gates:

Theorem 14. Let $\mathcal{P}$ be the distribution on $\mathfrak{S}_{\{0,1\}^n}$ given by choosing one random DES[2] gate. Then $\mathcal{P}$ is a $(1-\eta)$-approximate k-design for $\eta \geq \Omega(1/(nk \cdot \log k))$.

Although it may look like this result is conceptually dominated by Theorem 12, we include it as it has improved $\log k$ factors, and its proof reveals some of the ideas we use in our proof of Theorem 12.

1.3 Techniques and comparison with previous work

Prior work in this area also proceeded by bounding the spectral gap of the random walk on k-tuples induced by a single random DES[2] gate. The first such result was due to Gowers [Gow96], who established a version of our Theorem 14 with spectral gap $1/\mathrm{poly}(n,k)$. This was later improved by Hoory, Magen, Myers, and Rackoff [HMMR05] to $1/\widetilde{O}(n^2k^2)$, and by Brodsky and Hoory [BH08] to $1/O(n^2k)$. A result with gap η immediately translates into ε-approximate k-wise independence of random reversible circuits with $O(1/\eta) \cdot (nk + \log(1/\varepsilon))$ gates; e.g., the Brodsky–Hoory gap immediately gives their Theorem 10.

These earlier works bounded the spectral gap using the canonical paths method ([Gow96, HMMR05]) or multicommodity flows and the comparison method ([BH08]). We depart from these methods and use techniques from the physics literature concerned with the extent to which random *quantum* circuits are *unitary* k-designs. Specifically, for Theorem 14 we use the induction-on-n technique developed in [HHJ21], and for our main Theorem 12 we employ the more sophisticated Nachtergaele method [Nac96] as in the work of Brandão, Harrow, and Horodecki [BHH16]. These analyses use Fourier and spectral graph theory methods. In both cases, the inductive technique only begins to work for $n \geq \Theta(\log k)$, and for smaller n we need to base our argument on [BH08]; in the case of nearest-neighbor gates, this requires a further comparison-method based argument.

One can also interpret our result as the construction of a Cayley graph on $\mathfrak{A}_{2^n}$ with spectral gap $\Omega\big(1/n2^n\big)$ and degree $O(n)$. Previous work on this by Kassabov [Kas07] constructed constant-degree expanders out of Cayley graphs on $\mathfrak{A}_{2^n}$, but it is not clear if this random walk can be implemented by short circuits. Especially for the cryptographic applications of this work, it is important that our random walks have low circuit complexity.

1.4 Cryptographic results

Besides applications to derandomization, a lot of the motivation for studying ε-approximate k-wise independent permutation distributions $\mathcal{P}$ has come from cryptography. (See, e.g., recent analysis of the security properties of SPNs due to Liu, Pelecanos, Tessaro, and Vaikuntanathan [LTV21, LPTV23].) By definition, such distributions are statistically secure (up to advantage ε) against an adversary that gets access to any k input-output pairs, chosen nonadaptively. We also note that Maurer and Pietrzak [MP04] have shown that this can easily be upgraded to security against *adaptively* queries by composing two draws from the pseudorandom permutation (the second inverted). The focus in cryptography is on distributions $\mathcal{P}$ computable by size-S circuits, S a fixed polynomial in n, which are *computationally* indistinguishable from random. By this it is meant that an adversary running in time T (possibly much larger than S), having black-box query access to the permutation, cannot distinguish the pseudorandom permutation from a truly random one except with negligible advantage.

In cryptography, significant attention is paid to the designing simple and efficient block ciphers, and this motivates the question of whether reversible circuits of the type studied in this paper could be cryptographically secure pseudorandom permutations. This question was an original motivation for Gowers's work on random reversible circuits [Gow96]. Indeed, Gowers additionally asked

(cf. [Gow]) if the following general problem was hard: Suppose one is given *all* (2^n) input-output pairs (the "truth table") of a permutation on $\{0,1\}^n$. Is it computationally hard to determine if these pairs are consistent with some reversible circuit (with DES[2] gates) of size at most $S = n^5$ (say)? This is precisely the permutation analogue of the Minimum Circuit Size Problem (MCSP) (see [KC00]), and we refer to it as the *Minimum Reversible Circuit Size Problem (MRCSP)*.

Towards resolving the complexity of MRCSP, we establish it is hard (under a minimal cryptographic assumption) if the adversary can see any polynomial number of input-output pairs:

Theorem 15. Fix any constant $d \geq 3$. If one-way functions exist, then there is no deterministic poly(n)-time algorithm that — given black-box access to a permutation π on $\{0,1\}^n$ — decides whether π is computable by a reversible circuit of size at most n^d. (If subexponentially secure OWFs exist, then we can also rule out subexponential-time algorithms.)

We prove Theorem 15 as Corollary 76 in Section 7 by following the technique of Kabanets and Cai [KC00], who showed hardness of MCSP under OWFs. The key new thing we need to prove is a "reversibilized" version of the Luby–Rackoff [LR88] construction of (cryptographic) pseudorandom permutations from pseudorandom function generators. This was done much earlier by Cleve [Cle90] under the assumption that the pseudorandom function generator is computable in NC^1. We extend this to the case of any any poly-size circuit.

2 Preliminaries

2.1 The Main Characters: Some Linear Operators

We switch from $\{0,1\}$ notation to $\{\pm 1\}$ notation to facilitate later Fourier analysis. We will regard elements of $\{\pm 1\}^{nk}$ as tuples $(X^1, \ldots, X^k)$, where each $X^i \in \{\pm 1\}^n$. If $X^i \in \{\pm 1\}^n$, let X^i_S denote the vector X^i restricted to indices in S, so that if $S = \{a_1, \ldots, a_{|S|}\}$ with $a_1 < \cdots < a_{|S|}$ then $X^i_S = (X^i_{a_1}, \ldots, X^i_{a_{|S|}})$. For two vectors $x, y \in \{\pm 1\}^n$, define $\Delta(x,y) = \{a \in [n] : x_a \neq y_a\}$. Unless otherwise specified, log is $\log_2$. For positive integers $a \leq b$ denote $[a,b] = \{a, a+1, \cdots, b-1, b\}$ and let $[a] = [1,a]$. Let $\mathbf{1}$ be the all 1s vector, with length determined by context. When L is a self-adjoint operator let $\lambda_2(L)$ be its second-smallest distinct eigenvalue. If $g \in \mathfrak{S}_{\{0,1\}^\gamma}$ is a permutation and $S \subseteq [n]$ has $|S| = \gamma$ then define $g^S \in \mathfrak{S}_{\{0,1\}^n}$ by setting $(g^S x)_S = g(x_S)$, where coordinates are interpreted to be in lexicographic order, and $(g^S x)_{[n]\setminus S} = x_{[n]\setminus S}$.

When U is a set, we let $\mathbb{R}^{\mathcal{U}}$ denote the vector space of all functions $\mathcal{U} \to \mathbb{R}$. We equip this space with the inner product $\langle \cdot, \cdot \rangle$ given by

$$\langle f, g \rangle = \mathop{\mathbf{E}}_{\mathbf{x} \in \mathcal{U}} \left[f(\mathbf{x}) g(\mathbf{x}) \right]$$

for $f, g : \mathcal{U} \to \mathbb{R}$. Also define the norm by $\|f\|_2 = \sqrt{\langle f, f \rangle}$.

Definition 16. Let n, k be positive integers and let $S \subseteq [n]$.

- For every $X \in \{\pm 1\}^{nk}$ define the following distribution $\mathcal{D}^{n,S,k}_X$ on $\{\pm 1\}^{nk}$. To sample $\mathbf{Y} \sim \mathcal{D}^{n,S,k}_X$, let $\boldsymbol{\sigma} \in \mathfrak{A}_{\{\pm 1\}^{|S|}}$ be drawn uniformly at random and set $\mathbf{Y}^i = \boldsymbol{\sigma}^S X^i$ for all $i \in [k]$. Define the operator $R_{n,S,k}$ on $\mathbb{R}^{\{\pm 1\}^{nk}}$ by defining for $f : \{\pm 1\}^{nk} \to \mathbb{R}$ the new function $R_{n,m,k} f$ by

$$(R_{n,S,k} f)(X) = \mathbf{E}_{\mathbf{Y} \sim \mathcal{D}^{n,S,k}_X} \left[f(\mathbf{Y}) \right].$$

- For an integer $m \le n$ define $\mathcal{D}_X^{n,m,k}$ as follows. To sample $\mathbf{Y}$ from $\mathcal{D}_X^{n,m,k}$ sample $\mathbf{S} \in \binom{[n]}{m}$ uniformly at random and drawn $\mathbf{Y} \sim \mathcal{D}_X^{n,\mathbf{S},k}$. Define the operator $R_{n,m,k}$ on $\mathbb{R}^{\{\pm 1\}^{nk}}$ by defining for $f : \{\pm 1\}^{nk} \to \mathbb{R}$ the new function $R_{n,m,k} f$ by

$$(R_{n,m,k} f)(X) = \mathbf{E}_{\mathbf{Y} \sim \mathcal{D}_X^{n,m,k}} \left[f(\mathbf{Y}) \right].$$

- For a set $I \subseteq [n - \gamma]$ and a constant γ define $\mathcal{D}_X^{n,I,k,\gamma\text{-NN}}$ as follows. To sample $\mathbf{Y}$ from $\mathcal{D}_X^{n,I,k,\gamma\text{-NN}}$ sample $\mathbf{a} \in I$ uniformly at random and drawn $\mathbf{Y} \sim \mathcal{D}_X^{n,\{\mathbf{a},\ldots,\mathbf{a}+\gamma-1\},k}$. Define the operator $R_{n,I,k}^{\gamma\text{-NN}}$ on $\mathbb{R}^{\{\pm 1\}^{nk}}$ by defining for $f : \{\pm 1\}^{nk} \to \mathbb{R}$ the new function $R_{n,I,k}^{\gamma\text{-NN}} f$ by

$$(R_{n,I,k}^{\gamma\text{-NN}} f)(X) = \mathbf{E}_{\mathbf{Y} \sim \mathcal{D}_X^{n,I,k,\gamma\text{-NN}}} \left[f(\mathbf{Y}) \right].$$

- Define the distribution $\mathcal{D}_X^{n,k,\text{brickwork}}$ as follows. To sample $\mathbf{Y}$ from $\mathcal{D}_X^{n,k,\text{brickwork}}$ with probability $1/2$ sample $\mathbf{Y}$ by sampling $\lfloor n/3 \rfloor$ random permutations $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_{\lfloor n/3 \rfloor}$ from $\mathfrak{A}_{\{\pm 1\}^3}$ and setting $\mathbf{Y}^i = \boldsymbol{\sigma}_{\lfloor n/3 \rfloor}^{\{n\lfloor n/3 \rfloor, n\lfloor n/3 \rfloor+1, n\lfloor n/3 \rfloor+2\}} \cdots \boldsymbol{\sigma}_1^{\{1,2,3\}} X^i$ for all $i \in [k]$. Otherwise, draw $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_{\lfloor (n-1)/3 \rfloor} \in \mathfrak{A}_{\{\pm 1\}^3}$ and set $\mathbf{Y}^i = \boldsymbol{\sigma}_{\lfloor (n-1)/3 \rfloor}^{\{n\lfloor (n-1)/3 \rfloor+1, n\lfloor (n-1)/3 \rfloor+2, n\lfloor (n-1)/3 \rfloor+3\}} \cdots \boldsymbol{\sigma}_1^{\{2,3,4\}} X^i$ for all $i \in [k]$. Define the operator $R_{n,k}^{\text{brickwork}}$ on $\mathbb{R}^{\{\pm 1\}^{nk}}$ by defining for $f : \{\pm 1\}^{nk} \to \mathbb{R}$ the new function $R_{n,k}^{\text{brickwork}} f$ by

$$(R_{n,k}^{\text{brickwork}} f)(X) = \mathbf{E}_{\mathbf{Y} \sim \mathcal{D}_X^{n,k,\text{brickwork}}} \left[f(\mathbf{Y}) \right].$$

That is, the operator $R_{n,m,k}$ is essentially the transition matrix of the random walk on $\{\pm 1\}^{nk}$ given by randomizing a random subset of m bits. Similarly, the operator $R_{n,I,k}^{\gamma\text{-NN}}$ is essentially the transition matrix of the random walk on $\{\pm 1\}^{nk}$ given by randomizing a contiguous subset of γ bits. The operator $R_{n,k}^{\text{brickwork}}$ is the transition matrix of the random walk on $\{\pm 1\}^{nk}$ given by randomizing in the brickwork model.

Fact 17. All $R_{n,S,k}$ are projectors. For any n, $m \le n$, and k, we have

$$\left\| R_{n,m,k} - R_{n,n,k} \right\|_{\text{op}} \ge 0.$$

Proof. All $R_{n,S,k}$ are PSD because they are projectors. $R_{n,n,k} = R_{n,[n],k}$ is a projector to a subspace of vectors, all of which are eigenvectors of $R_{n,m,k}$ with eigenvalue 1. □

Fact 18. For any n, S, k we have that $R_{n,S,k}$ is self-adjoint.

Proof. As a matrix, $R_{n,S,k}$ is a stochastic symmetric matrix, which can be seen by evaluating transition probabilities. □

Fact 19. For any $S \subseteq [n]$ we have $R_{n,S,k}\mathbf{1} = \mathbf{1}$ ($\mathbf{1}$ is the all 1s vector). Consequently, $R_{n,m,k}\mathbf{1} = \mathbf{1}$ and $R_{n,I,k}^{\gamma\text{-NN}}\mathbf{1} = \mathbf{1}$ for all m, γ, and intervals I.

Proof. As a matrix, $R_{n,S,k}$ is doubly stochastic, as per Fact 18. □

In some contexts it is easier to work with the Laplacians of these operators:

Definition 20. For any operator R on a vector space V let $L(R)$ denote its Laplacian $\text{Id}_V - R$. In the case where $R = R_{n,m,k}$ for integers n, m, k define $L_{n,m,k} = L(R_{n,m,k})$. In the case where $R = R_{n,I,k}^{\gamma\text{-NN}}$ define $L_{n,I,k}^{\gamma\text{-NN}} = L(R_{n,I,k}^{\gamma\text{-NN}})$. In the case where $R = R_{n,k}^{\text{brickwork}}$ define $L_{n,k}^{\text{brickwork}} = L(R_{n,I,k}^{\text{brickwork}})$.

2.2 Fourier Analysis of Boolean Functions

Definition 21. Let $S_1, \dots, S_k \subseteq [n]$. Then define the function $\chi_{S_1,\dots,S_k}$ by defining for $X \in \{\pm 1\}^{nk}$,

$$\chi_{S_1,\dots,S_k}(X) = \prod_{i\in[k]} \prod_{a\in S_i} X_a^i.$$

Fact 22 ([O'D14])**.** The functions $\chi_{S_1,\dots,S_k}$ for $S_1, \dots, S_k \subseteq [n]$ form an orthonormal basis of $\mathbb{R}^{\{\pm 1\}^{nk}}$.

3 Spectral Gaps

3.1 Fully Random Gates

Throughout this section let n be a fixed positive integer. In this section we prove Theorem 14 by building on the previously-mentioned following result of Brodsky and Hoory:

Theorem 23 ([BH08], Theorem 2)**.** Given $f : \{\pm 1\}^{nk} \to \mathbb{R}$, we have for any $m \le n$ that

$$\left|\left\langle f, (R_{m,3,k} - R_{m,m,k}) f\right\rangle\right| \le 1 - \Omega\left(\frac{1}{m^2 k}\right).^{5}$$

Our main contribution is a finer analysis in the case when k is small relative to m, which results in the following theorem.

Theorem 24. Assume that $m \ge 100$ and $k \le 2^{m/3}$. Given $f : \{\pm 1\}^{mk} \to \mathbb{R}$, we have

$$\left|\left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f\right\rangle\right| \le \left(\frac{1}{m} + \frac{k^2}{2^{m/3}}\right) \langle f, f\rangle .$$

Theorem 24 is proven as Theorem 45 in Section 5.

The following lemma allows us to combine Theorem 24 with the previously-known Theorem 23 to bring the quadratic dependence on the number of wires (n) to linear.

Lemma 25 ([OSP23], Lemma 3.2)**.** Fix a positive integer $n_0 \ge 4$. For each k and $m_1 \ge m_2$ let $\tau_{m_1,m_2,k}$ be some real number such that

$$L_{m_1,m_2,k} \ge \tau_{m_1,m_2,k} L_{m_1,m_1,k}.^{6}$$

Then for any sequence $n_0 = m_0 \le m_1 \le \cdots \le m_{t-1} \le m_t = n$ we have

$$L_{n,n_0,k} \ge \left(\prod_{i\in[t-1]} \tau_{m,k}\right) L_{m_i,m_{i+1},k}.$$

Together, Theorem 23, Theorem 24, and Lemma 25 yield the following initial spectral gap.

[5] [BH08] actually proves this inequality for the operator $R^{\mathrm{DES}[2]}_{m,3,k}$ in place of $R_{m,3,k}$, which is the random walk operator induced by placing a random width-2 permutation (which acts on 3 bits). However, a standard comparison of Markov chains shows that our statement of the result easily follows. See Section 3.3.

[6] All inequalities between operators in this paper are in the PSD order.

Corollary 26. For any n and $k \le 2^{n/3}$, we have

$$L_{n,3,k} \ge \Omega\left(\frac{1}{nk \cdot \log k}\right) L_{n,n,k}.$$

Proof. Theorem 24 shows that for all $m \ge 4\log(k)$ we have $\tau_{m,k} \ge 1 - \frac{1}{m} - \frac{1}{2^{m/3}}$. By Lemma 25 applied with $n_0 = 4\log k$, we have

$$L_{n,4\log k,k} \ge \prod_{m=4\log k}^{n} \left(1 - \frac{1}{m} - \frac{1}{2^{m/3}}\right) L_{n,n,k} \ge \frac{\log k}{n} L_{n,n,k}.$$

Finally, Theorem 23 shows that

$$L_{4\log k,3,k} \ge \Omega\left(\frac{1}{k \cdot \log^2(k)}\right) L_{4\log k,4\log k,k}$$

Combining these two inequalities using Lemma 25 produces the result. □

We leverage the initial spectral gap from Corollary 26 to produce our designs by sequentially composing many copies of this pseudorandom permutation. This is akin to showing that the second largest eigenvalue of the square of a graph is quadratically smaller than that of the original graph.

Corollary 27. Let $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_t$ be drawn by drawing a random set from $\binom{[n]}{3}$ and setting each $\boldsymbol{\sigma}_i$ to be the permutation computed by a random 3-bit gate on the 3 wires corresponding to the random set drawn. Let $\boldsymbol{\pi} = \boldsymbol{\sigma}_t \ldots \boldsymbol{\sigma}_1$, and say that $\boldsymbol{\pi}$ was drawn from the distribution Ckt_t. Then

$$\left\| \mathop{\mathbf{E}}_{\boldsymbol{\pi} \sim \mathrm{Ckt}_t} \left[\rho^k(\boldsymbol{\pi})\right] - \mathop{\mathbf{E}}_{\pi \sim \mathrm{Unif}(\mathfrak{A}_N)} \left[\rho^k(\boldsymbol{\pi})\right] \right\|_{\mathrm{op}} \le \left(1 - \frac{1}{nk \cdot \log k}\right)^{\Omega(t)}$$

Proof. For any $g \in \mathfrak{A}_N$, we have $\rho^k(g)\, \mathbf{E}_{\pi \sim \mathrm{Unif}(\mathfrak{A}_N)} \left[\rho^k(\boldsymbol{\pi})\right] = \mathbf{E}_{\pi \sim \mathrm{Unif}(\mathfrak{A}_N)} \left[\rho^k(\boldsymbol{\pi})\right]$. Note also that $\mathbf{E}_{\pi \sim \mathrm{Unif}(\mathfrak{A}_N)} \left[\rho^k(\boldsymbol{\pi})\right] = R_{m,m,k}$. Therefore,

$$\left(R_{m,3,k} - R_{m,m,k}\right)^t = R_{m,3,k}^t - R_{m,m,k}^t = R_{m,3,k}^t - R_{m,m,k}.$$

Using this, we find that

$$\begin{aligned}
&\left\| \mathop{\mathbf{E}}_{\boldsymbol{\pi} \sim \mathrm{Ckt}_t} \left[\rho^k(\boldsymbol{\pi})\right] - \mathop{\mathbf{E}}_{\pi \sim \mathrm{Unif}(\mathfrak{A}_N)} \left[\rho^k(\boldsymbol{\pi})\right] \right\|_{\mathrm{op}} = \left\| R_{m,3,k}^t - R_{m,m,k} \right\|_{\mathrm{op}} \\
&= \left\| \left(R_{m,3,k} - R_{m,m,k}\right)^t \right\|_{\mathrm{op}} \le \left\| R_{m,3,k} - R_{m,m,k} \right\|_{\mathrm{op}}^t \le \left(1 - \Omega\left(\frac{1}{nk \cdot \log k}\right)\right)^t \\
&\le \left(1 - \frac{1}{nk \cdot \log k}\right)^{\Omega(t)}.
\end{aligned}$$

Here we used the bound proved in Corollary 26. □

Theorem 14 follows directly from Corollary 27 (with $t = \Theta(nk \cdot \log k)$) and Section 3.3.

3.2 Nearest-Neighbor Random Gates

3.2.1 Reduction to the Large k Case

One step in a random reversible circuit with 1D-nearest-neighbor gates is described by the operator $R^{\text{3-NN}}_{n,[n-2],k}$. Note that we can write

$$R^{\text{3-NN}}_{n,[n-2],k} = \mathop{\mathbf{E}}_{\mathbf{a}\in[n-2]}\Big[R_{n,\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\},k}\Big].$$

Because the local terms $R_{n,\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\},k}$ are projectors, to analyze such an operator we can use the following theorem of Nachtergaele.

Theorem 28 ([Nac96], Theorem 3)**.** Let $\{h_{a,a+1,a+2}\}_{a\in[n-2]}$ be projectors acting on $(\mathbb{R}^2)^{\otimes d}$ such that each $h_{\{a,a+1,a+2\}}$ only acts on the $a, a+1, a+2$th tensor factors. For $I = [a,b] \subseteq [n]$ define the subspace

$$\mathcal{G}_I = \left\{ f \in (\mathbb{R}^2)^{\otimes d} : \sum_{a'\in[a,b-2]} h_{a',a'+1,a'+2} f = 0 \right\}.$$

Let G_I be the projector to $\mathcal{G}_I$.

Now suppose there exists ℓ and n_ℓ and $\varepsilon_\ell \le \frac{1}{\sqrt{\ell}}$ such that for all $n_\ell \le m \le n$,

$$\Big\| G_{[m-\ell-1,m]}\Big(G_{[m-1]} - G_{[m]}\Big)\Big\|_{\text{op}} \le \varepsilon_\ell.$$

Then

$$\lambda_2\left(\sum_{a\in[n-2]} h_{\{a,a+1,a+2\}}\right) \ge \frac{\Big(1-\varepsilon_\ell\sqrt{\ell}\Big)^2}{\ell-2}\lambda_2\left(\sum_{a\in[\ell]} h_{\{a,a+1,a+2\}}\right).$$

Recall $\lambda_2(h)$ denotes the second-smallest distinct eigenvalue of the operator h.

Theorem 29. Fix any $m \ge 100$ and $k \le 2^m - 2$ and set $\ell = 10\log k$. Then we have

$$\Big\| R_{m,[m-\ell-1,m],k}\Big(R_{m,[m-1],k} - R_{m,[m],k}\Big)\Big\|_{\text{op}} \le \frac{1}{\ell}.$$

Theorem 29 is established in Section 6 as Theorem 61.

Corollary 30. We have for $k \ge 3$ that

$$\lambda_2\Big(L^{\text{3-NN}}_{n,[n-2],k}\Big) \ge \frac{1}{2n}\lambda_2\Big(L^{\text{3-NN}}_{10\log k+2,[10\log k-2],k}\Big)$$

Proof. Setting $h_{a,a+1,a+2} = \mathrm{Id} - R_{n,\{a,a+1,a+2\},k}$ for each $a \in [n]$, we see that the projections G (as in the statement of Theorem 28) are given by $G_{[a]} = R_{n,[\min\{n,a+2\}],k}$ for any $a \in [n]$. To see this, note first that $R_{n,[\min\{n,a+2\}],k}$ is indeed a projection. Now let f be such that $R_{n,[\min\{n,a+2\}],k} f = f$. For every $\sigma \in \mathfrak{S}_{\{\pm1\}^n}$ define f^σ by $f^\sigma(X) = f(\sigma X)$ for all $X \in \{\pm1\}^{nk}$. Then by invariance of $R_{n,[\min\{n,a+2\}],k}$ under a permutation applied to bits in $[\min\{n,a+2\}]$ we have

$$f^\sigma = R_{n,[\min\{n,a+2\}],k} f^\sigma = R_{n,[\min\{n,a+2\}],k} f = f.$$

for any $\sigma \in \mathfrak{S}_{\{\pm 1\}^{[\min\{n,a+2\}]}}$. The converse of this holds as well by a similar argument. Therefore, $f^{\sigma^{\{a',a'+1,a'+2\}}} = f$ for any $\sigma \in \mathfrak{S}_{\{0,1\}^8}$ for $a' \le a$, proving that such an f is truly in the ground space of $R_{n,\{a',a'+1,a'+2\},k}$ for any $a' \le a$. The converse of this holds as well, because the permutations of the form $\sigma^{\{a',a'+1,a'+2\}}$ with $a' \le a$ generate the group of permutations of the form $\rho^{\{a,b,c\}}$ for $\{a,b,c\} \in \binom{[\min\{n,a+2\}]}{3}$. Such an argument also proves the converse, so the R operators are serve as the projections from Theorem 28.

Therefore, by Theorem 29 we have that the hypotheses of Theorem 28 are satisfied with $\ell = 10 \log k$ and $\varepsilon_\ell = \frac{1}{\ell}$. That is, for any $m \le n$ we have

$$\left\| G_{[m-\ell-1,m]} \left(G_{[m-1]} - G_{[m]} \right) \right\|_{\text{op}} = \left\| R_{n,[m-\ell-1,m],k} \left(R_{n,[m-1],k} - R_{n,[m],k} \right) \right\|_{\text{op}}$$
$$= \left\| \left(R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) \right) \otimes \text{Id}_{[m+1,n]} \right\|_{\text{op}} \le \frac{1}{\ell}.$$

Therefore the conclusion of Theorem 28 is that

$$\lambda_2 \left((n-2) L^{\text{3-NN}}_{n,[n-2],k} \right) \ge \frac{\left(1 - \frac{1}{\sqrt{\ell}}\right)^2}{\ell - 2} \lambda_2 \left(\ell L^{\text{3-NN}}_{\ell+2,[\ell-2],k} \right) \ge \frac{1}{2} \lambda_2 \left(L^{\text{3-NN}}_{\ell+2,[\ell-2],k} \right).$$

Recalling our setting of ℓ completes the proof. □

3.2.2 Comparison Method for the Large k Case

We can use the spectral gap proved in Theorem 23 for the random walk induced by completely random 3-bit gates to show a spectral gap for the random walk induced by random 3-bit gates, where the three bits on which the gate acts on are $a, a+1, a+2$ for some $a \in [n]$. Our proof is a simple application of the comparison method applied to random walks on multigraphs.

We take the following definition of (multi)graphs. A graph is a pair of sets (V, E) such that there is a partition $E = \bigcup_{(x,y) \in V^2} E_{x,y}$. If $e \in E_{x,y}$ then we say that e *connects* the vertex x to the vertex y, and we define $u(e) = x$ and $v(e) = y$. The degree $\deg(x)$ of a vertex $x \in V$ is the number of edges originating at e, or $\sum_{y \in V} |E_{x,y}|$. We say that a graph is regular if $\deg(x) = \deg(y)$ for all $x, y \in V$.

The *random walk* on a graph (V, E) beginning at a vertex $x \in V$ consists of the Markov chain $\{\boldsymbol{x}_i\}_{i \ge 0}$ on state space V such that $\boldsymbol{x}_0 = x$ with probability 1, and to draw $\boldsymbol{x}_{i+1}$ given $\boldsymbol{x}_i$ we sample a uniform random edge $\boldsymbol{e}$ from $\bigcup_{y \in V} E_{x,y}$. We set $\boldsymbol{x}_{i+1}$ equal to the unique $y \in V$ such that $\boldsymbol{e} \in E_{x,y}$.

A *Schreier graph* is a graph with vertices V such that some group $\mathfrak{G}$ acts on V. Let $S \subseteq \mathfrak{G}$ be some subset of group elements. The edge set consists of elements of the form (X, σ) $(\sigma \in S)$, so that $(X, \sigma) \in E_{X,\sigma X}$. We call the resulting graph $\text{Sch}(V, S)$.

Definition 31. Let S and $\widetilde{S}$ be subsets of a group acting on a set V. For each $\sigma \in S$ let $\Gamma(\sigma)$ be a sequence $(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_t)$ of elements of $\widetilde{S}$ such that $\sigma v = \widetilde{\sigma}_t \ldots \widetilde{\sigma}_1 v$ for all $v \in V$, so we regard Γ as a map from S to sets of paths using edges in $\widetilde{S}$. Define the *congestion ratio* of Γ to be

$$B(\Gamma) = \max_{\widetilde{\sigma} \in \widetilde{S}} \left\{ |\widetilde{S}| \mathop{\mathbf{E}}_{\sigma \in S} \left[N(\widetilde{\sigma}, \Gamma(\sigma)) |\Gamma(\sigma)| \right] \right\},$$

where $N(\widetilde{\sigma}, \Gamma(\sigma))$ is the number of times $\widetilde{\sigma}$ appears in the sequence $\Gamma(\sigma)$.

Lemma 32. Let $G = \mathrm{Sch}(V, S)$ and $\widetilde{G} = \mathrm{Sch}(V, \widetilde{S})$ be the connected Schreier graphs of the action of a group. Let L and $\widetilde{L}$ be the Laplacian operators for the non-lazy random walks on G and $\widetilde{G}$, respectively. Suppose there exists a Γ as in Definition 31. Then

$$\lambda_2(L) \geq \left(\max_{v \in V} \frac{\pi(v)}{\widetilde{\pi}(v)}\right) B(\Gamma)\lambda_2(\widetilde{L}).$$

Here π and $\widetilde{\pi}$ are the stationary distributions for G and $\widetilde{G}$, respectively.

We prove this in Appendix A. It is essentially a reformulation of a standard result about comparisons on general Markov chains in [WLP09].

Lemma 33. For any n, k we have

$$\lambda_2\left(L^{3\text{-NN}}_{n,[n-2],k}\right) \geq \frac{1}{100000n^3}\lambda_2\left(L_{n,3,k}\right).$$

Proof. Note that $L_{n,3,k}$ and $L^{3\text{-NN}}_{n,[n-2],k}$ are simply the Laplacians of random walks on Schreier graphs with $\mathfrak{S}_{\{\pm 1\}^n}$ acting on $\{\pm 1\}^{nk}$ by $e(X^1, \ldots, X^k) = (eX^1, \ldots, eX^k)$. In the case of $L_{n,3,k}$ the edges are given by elements of the form $h^{\{a,b,c\}}$ for $h \in \mathfrak{S}_{\{\pm 1\}^3}$ and $\{a, b, c\} \in \binom{[n]}{3}$. In the case of $L^{3\text{-NN}}_{n,[n-3],k}$ the edges are given by elements of the form $g^{\{a,a+1,a+2\}}$ for $g \in \mathfrak{S}_{\{\pm 1\}^3}$ and $a \in [n-2]$. We deal with each connected component separately. Note that every connected component is isomorphic to $\{(X^1, \ldots, X^{k'}) : X^i \neq X^j \iff i \neq j\}$ for $k' \leq k$, so we bound the spectral gap for the walk on $\{(X^1, \ldots, X^k) : X^i \neq X^j \iff i \neq j\}$.

We provide a map Γ from $\{h^{\{a,b,c\}} : h \in \mathfrak{S}_{\{0,1\}^3}, a, b, c \in [n]\}$ to sequences of elements of the form $g^{\{a,a+1,a+2\}}$ for $a \in [n-2]$ such that for any $h^{\{a,b,c\}}$ the sequence $\Gamma(h^{\{a,b,c\}}) = (g_1^{\{a_1,a_1+1,a_1+2\}}, \ldots, g_t^{\{a_t,a_t+1,a_t+2\}})$ satisfies $h^{\{a,b,c\}} = g_t^{\{a_t,a_t+1,a_t+2\}} \ldots g_1^{\{a_1,a_1+1,a_1+2\}}$.

The hope is to construct Γ such that $B(\Gamma)$ is small and then to apply Lemma 32. To this end we define Γ as follows. Assume $a < b < c$. Fix $h \in \mathfrak{S}_8, a \in [n-2]$. Let $d \in [n-2]$ be arbitrary. Then write

$$h^{\{a,b,c\}} = \mathsf{Sort}^{-1} \cdot g^{\{d,d+1,d+2\}} \cdot \mathsf{Sort},$$

Here Sort sends the ath coordinate to the dth coordinate, the bth to the $d+1$th, and the cth to the $d+2$th. The permutations Sort and Sort^{-1} can each be implemented using at most $3n$ gates of the form $g^{\{a'-1,a',a'+1\}}$ and $g^{\{a'+1,a'+2,a'+3\}}$ where each g swaps coordinates; this is just by a standard partial sorting algorithm. Write $\mathsf{Sort} = g_{3n}^{\{a_{3n},a_{3n}+1,a_{3n}+2\}} \ldots g_1^{\{a_1,a_1+1,a_1+2\}}$. Then set

$$\Gamma(h^{\{a,b,c\}}) = \left(g_1^{\{a_1,a_1+1,a_1+2\}}, \ldots, g_{3n}^{\{a_{3n},a_{3n}+1,a_{3n}+2\}}, h^{(d,d+1,d+2)}, \ldots, \left(g_1^{\{a_1,a_1+1,a_1+2\}}\right)^{-1}\right).$$

We have $B(\Gamma) \leq 100000n^3$ trivially, so we have proved the result by applying Lemma 32 and the fact that the stationary distributions for both chains are uniform. □

Corollary 34. For any n, k we have

$$\lambda_2\left(L^{3\text{-NN}}_{n,[n-2],k}\right) \geq \Omega\left(\frac{1}{nk \cdot \log^5(k)}\right).$$

Proof. Lemma 33 shows that $\lambda_2\left(L^{\text{3-NN}}_{\ell+2,[\ell],k'}\right) \geq \frac{1}{\ell^3}\lambda_2\left(L_{\ell+2,3,k'}\right)$ for all ℓ and k'. Theorem 23 states that $\lambda_2\left(L_{\ell+2,3,k'}\right) \geq \frac{1}{(\ell+2)^2 k'}$ for all ℓ and k'. If we set $\ell = 10\log k$ and use Corollary 30 then we get

$$\lambda_2\left(L^{\text{3-NN}}_{n,[n-2],k}\right) \geq \frac{1}{2n} \cdot \lambda_2\left(L^{\text{3-NN}}_{10\log k+2,[10\log k],k}\right) \geq \frac{1}{2n} \cdot \frac{1}{100000k\log^5(k)}.$$

This implies the result. □

As in the case for fully random gates, Theorem 12 follows from Corollary 34 and Section 3.3.

3.3 Restricting the Gate Set

So far all of our results have dealt with random circuits with arbitrary gates acting on 3 bits. However, for practical applications we are often in further restricting the type of 3-bit gates. However, as long as the arbitrary gate set is universal on 3 bits, we lose just a constant factor in the mixing time when we restrict our random circuits to use that gate set, by a standard application of the comparison method. We prove that we can perform this conversion before proving our results about brickwork circuits in Section 3.4 because it is somewhat easier to prove for the case of single gates acting at a time.

Lemma 35. We have

$$\lambda_2\left(L^{\text{3-NN,DES[2]}}_{n,[n-2],k}\right) \geq \Omega\left(\lambda_2\left(L^{\text{3-NN}}_{n,[n-2],k}\right)\right)$$

Proof. We compare the Markov chains given by these two Laplacians by providing a way to write edges in the one induced by arbitrary 3-bit nearest-neighbor gates as paths in the one induced by 3-bit nearest-neighbor gates with generators $\mathcal{G}$. Again we focus on the connected component $\{(X^1,\ldots,X^k) : X^i \neq X^j \iff i \neq j\}$.

For each g^S for $g \in \mathfrak{S}_8$ and $S \in \binom{[n]}{3}$ let $\Gamma(g^S) = \left(g^S_{i_1},\ldots,g^S_{i_{8!}}\right)$ where we have fixed an arbitrary expansion of $g = g_{i_1}\ldots g_{i_{8!}}$, and each g_{i_j} is of type DES[2]. Then in the notation of Lemma 32 we have that

$$\begin{aligned}
B(\Gamma) &= \max_{g\in\mathcal{G},a\in[n-2]}\left\{|\mathcal{G}|(n-2)\mathop{\mathbf{E}}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]}\left[N\left(g^{\{a,a+1,a+2\}},\Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right)\middle|\Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right]\right\}\\
&\leq \max_{g\in\mathcal{G},a\in[n-2]}\left\{8!n\mathop{\mathbf{E}}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]}\left[8!N\left(g^{\{a,a+1,a+2\}},\Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right)\right]\right\}\\
&\leq \max_{g\in\mathcal{G},a\in[n-2]}\left\{(8!)^3 n\,\mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]}\left[g^{\{a,a+1,a+2\}} \in \Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right]\right\}\\
&\leq \max_{g\in\mathcal{G},a\in[n-2]}\left\{(8!)^3 n\,\mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]}\left[a=\mathbf{a}\right]\right\}\\
&\leq (8!)^3.
\end{aligned}$$

Applying Lemma 32 completes the proof. □

This shows that the random walk given by applying random gates on 3 bits of type DES[2] has spectral gap $\widetilde{\Omega}(1/nk)$, by combining with Corollary 34. A similar proof essentially shows the same result for the random circuit models where gates on arbitrary sets of 3 bits.

3.4 Brickwork Circuits

The spectral gap for brickwork circuits follows almost directly from the spectral gap for circuits with nearest-neighbor gates (Corollary 34), as in [BHH16]. First we show that the random walk induced by 3-bit nearest neighbor DES[2] gates, where the 3 bits on which gates act on are of the form $\{a, a+1, a+2\}$ for any $a \in [n-2]$, has approximately the same spectral gap as that in which the random gates are of the form $\{a, a+1, a+2\}$ for $a \in [n-2]$ but with the restriction that $a \neq 0 \bmod 3$. Use the notation $L^{\text{3-NN,DES[2]}}_{n,[n-2],k}$ and $L^{\text{3-NN,DES[2]}}_{n,\{a\in[n-2],a=1,2 \bmod 3\},k}$ for the Laplacians of these random walks. Assume that $n = 0 \bmod 3$; the other cases follow similarly.

Lemma 36. For any n, k we have

$$\lambda_2\left(L^{\text{3-NN,DES[2]}}_{n,\{a\in[n-2],a=1,2 \bmod 3\},k}\right) \geq \Omega\left(\lambda_2\left(L^{\text{3-NN}}_{n,[n-2],k}\right)\right).$$

Proof. By Lemma 35, it suffices to show

$$\lambda_2\left(L^{\text{3-NN,DES[2]}}_{n,\{a\in[n-2],a=1,2 \bmod 3\},k}\right) \geq \Omega\left(\lambda_2\left(L^{\text{3-NN,DES[2]}}_{n,[n-2],k}\right)\right).$$

We use the comparison method. Again we focus on the connected component $\{(X^1, \ldots, X^k) : X^i \neq X^j \iff i \neq j\}$. For each $g \in \mathfrak{S}_{\{0,1\}^3} \cong \mathfrak{S}_8$ of type DES[2] and $a \in [n-2]$ we provide a sequence $\Gamma(g^{\{a,a+1,a+2\}})$ of permutations multiplying to $g^{\{a,a+1,a+2\}}$ using only permutations of the form $h^{\{b,b+1,b+2\}}$ with $b \neq 0 \bmod 3$ such that the resulting congestion $B(\Gamma)$ is small.

We define Γ as follows. Fix $g \in \mathfrak{S}_8, a \in [n-2]$ where g is of type DES[2]. If $a \neq 0 \bmod 3$ then simply set $\Gamma(g^{\{a,a+1,a+2\}}) = (g^{\{a,a+1,a+2\}})$. Otherwise $a = 0 \bmod 3$. Then there exists a sequence of 64! permutations of the form $g_i^{\{b_i,b_i+1,b_i+2\}}$ with each $b_i \in \{a-1, a+1\}$ such that $g = g_1^{\{b_1,b_1+1,b_1+2\}} \cdots g_1^{\{b_{64!},\ldots,b_{64!}+1,\ldots,b_{64!}+2\}}$. This is because we can implement the gate $g^{\{a,a+1,a+2\}}$ as

$$g^{\{a,a+1,a+2\}} = \mathsf{Sort}^{-1} \cdot g^{\{a-1,a,a+1\}} \cdot \mathsf{Sort},$$

where Sort sends $(x_1, \ldots, x_{a-1}, x_a, x_{a+1}, x_{a+2}, \ldots, x_n) \to (x_1, \ldots, x_a, x_{a+1}, x_{a+2}, x_{a-1}, \ldots, x_n)$. The permutations Sort and Sort^{-1} can each be implemented as the product of at most 32! permutations of the form $h^{\{a-1,a,a+1\}}$ and $h^{\{a+1,a+2,a+3\}}$ where each h is of type DES[2]. This gives the implementation of $g^{\{a,a+1,a+2\}}$ as the product of at most 64! elements of the form $g^{\{a-1,a,a+1\}}$ or $g^{\{a,a+1,a+2\}}$, and this defines $\Gamma(g^{\{a,a+1,a+2\}})$ for such g and a.

In the notation of Lemma 32, the congestion of Γ is bounded:

$$\begin{aligned}
B(\Gamma) \leq & \max_{g\in\mathcal{G},a=1,2 \bmod 3} \left\{ \frac{8! \cdot 2n}{3} \mathop{\mathbf{E}}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]} \left[N\left(g^{\{a,a+1,a+2\}}, \Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right) \left| \Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}}) \right| \right] \right\} \\
\leq & 70! \max_{g\in\mathcal{G},a=1,2 \bmod 3} \left\{ n \mathop{\mathbf{E}}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]} \left[N\left(g^{\{a,a+1,a+2\}}, \Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}})\right) \right] \right\} \\
\leq & 70! \max_{g\in\mathcal{G},a=1,2 \bmod 3} \left\{ n \mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]} \left[g^{\{a,a+1,a+2\}} \in \Gamma(\mathbf{g}^{\{\mathbf{a},\mathbf{a}+1,\mathbf{a}+2\}}) \right] \right\} \\
\leq & 70! \max_{g\in\mathcal{G},a=1,2 \bmod 3} \left\{ n \mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_8,\mathbf{a}\in[n-2]} \left[a \in \{\mathbf{a}-1, \mathbf{a}, \mathbf{a}+1\} \right] \right\} \\
\leq & 71!.
\end{aligned}$$

Here we used that $\left|\Gamma(g^{\{a,a+1,a+2\}})\right| \leq 64!$ always. Applying Lemma 32 completes the proof. □

Lemma 37 ([BHH16], Section 4.A)**.** For any n, k we have

$$\lambda_2\left(L_{n,k}^{\text{brickwork,DES[2]}}\right) \geq n\Omega\left(\lambda_2\left(L_{n,\{a\in[n-2],a=1,2 \bmod 3\},k}^{\text{3-NN,DES[2]}}\right)\right).$$

The idea is to write

$$\begin{aligned} R_{n,k}^{\text{brickwork,DES[2]}} &= \frac{1}{2}P_{\text{odd}} + \frac{1}{2}P_{\text{even}} \\ &= \frac{1}{2}R_{n,\{1,2,3\},k}^{\text{DES[2]}}R_{n,\{4,5,6\},k}^{\text{DES[2]}}\cdots R_{n,\{n-2,n-1,n\},k}^{\text{DES[2]}} + \frac{1}{2}R_{n,\{2,3,4\},k}^{\text{DES[2]}}R_{n,\{5,6,7\},k}^{\text{DES[2]}}\cdots R_{n,\{n-4,n-3,n-2\},k}^{\text{DES[2]}}. \end{aligned}$$

Here the operator $R_{n,S,k}^{\text{DES[2]}}$ is the transition operator for the Markov chain that is similar to $R_{n,S,k}$ but with the restriction that each step is induced by a gate of type DES[2]. Then we note that each individual factors in each of the two products all commute. So this operator is the sum of two projections. Then we can use the inequality $\|P+Q\| \leq 1+\|PQ\|$. When we take $P = R_{n,n,k} - P_{\text{odd}}$ and $Q = R_{n,n,k} - P_{\text{even}}$ we get the inequality

$$\|P_{\text{odd}} + P_{\text{even}} - 2P_{\text{all}}\| \leq 1 + \|P_{\text{odd}}P_{\text{even}} - P_{\text{all}}\|.$$

Here P_{all} is the projector to the intersection of the images of P_{odd} and P_{even}. The quantity on the LHS is directly related to the spectral gap we are trying to bound. The detectability lemma [AALV09] bounds the quantity on the RHS.

Corollary 38. For any n, k we have

$$\lambda_2\left(L_{n,k}^{\text{brickwork,DES[2]}}\right) \geq \Omega\left(\frac{1}{k\cdot\text{polylog}(k)}\right).$$

As in the case for fully random gates and nearest-neighbor random gates, Theorem 13 follows from Corollary 38.

4 Proof Overview and Some Helpful Facts

The proof of Theorem 24 follows from directly analyzing the spectrum of the operator $R_{m,m-1,k}$. We view this operator as *noise operators* or transition matrices on the space $\mathbb{R}^{\{\pm1\}^{mk}}$. The issue is that the R matrices are not a "nice" kind of noise. In particular as transition matrices the transition probabilities aren't products of transition probabilities for each element in the k-tuple. In some sense, the Markov chain that they govern is determined by a process that is essentially resampling certain coordinates without replacement from $\{\pm1\}^m$.

Moving to a Nicer Noise Model. However, we notice that in the case where k is significantly smaller than $2^{m/2}$, the process of sampling with replacement from $\{\pm1\}^m$ is very similar to the process of sampling without replacement. Thus, we can essentially pass to this nicer noise model of sampling with replacement, represented by the Q operators we introduce in Section 5, which guarantees independence and allows the use of Fourier-analytic techniques to bound the spectral norms of Q (which are close to R). Essentially, we move to analyzing a nicer noise model. The following lemmas help us formalize this connection between closeness of Markov chains and linear-algebraic quantities we deal with.

Lemma 39. Suppose A and B are self-adjoint random walk matrices on a domain $\mathcal{U}$. That is, for every $X \in \mathcal{U}$ there are distributions $\mathcal{C}_X, \mathcal{D}_X$ such that for any $f : \mathcal{U} \to \mathbb{R}$,

$$\begin{aligned}(Af)(X) &= \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_X}\big[f(\mathbf{Y})\big],\\ (Bf)(X) &= \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{C}_X}\big[f(\mathbf{Y})\big].\end{aligned}$$

Let $\mathrm{Supp} \subseteq \mathcal{U}$. Then for any $f, g : \mathcal{Y} \to \mathbb{R}$ with $\mathrm{Supp}(f) = \mathrm{Supp}(g) = \mathrm{Supp}$ we have

$$\begin{aligned}&\Big|\big\langle f, (A-B)f\big\rangle\Big|\\ &\leq \sqrt{\sum_{X\in\mathrm{Supp}} f(X)^2 \sum_{Y\in\mathrm{Supp}} \big|p_0(X,Y) - p_1(X,Y)\big|}\sqrt{\sum_{X\in\mathrm{Supp}} g(X)^2 \sum_{Y\in\mathrm{Supp}} \big|p_0(X,Y) - p_1(X,Y)\big|}.\end{aligned}$$

where $p_0(X,Y) = \frac{1}{|\mathcal{U}|}\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X}[\mathbf{Y} = Y]$ and $p_1(X,Y) = \frac{1}{|\mathcal{U}|}\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{C}_X}[\mathbf{Y} = Y]$.

In particular, when $f = g$ we get

$$\Big|\big\langle f, (A-B)f\big\rangle\Big| \leq \sum_{X,Y\in\mathrm{Supp}(f)} f(X)^2 \sum_{Y\in\mathrm{Supp}(f)} \big|p_0(X,Y) - p_1(X,Y)\big|,$$

Proof. We directly compute

$$\begin{aligned}&\Big|\big\langle f, (A-B)g\big\rangle\Big|\\ &= \left|\mathop{\mathbf{E}}_{\mathbf{X}\in\{\pm1\}^{mk}}\Big[f(\mathbf{X})\big((Ag)(\mathbf{X}) - (Bg)(\mathbf{X})\big)\Big]\right|\\ &= \left|\mathop{\mathbf{E}}_{\mathbf{X}\in\{\pm1\}^{mk}}\left[f(\mathbf{X})\Big(\mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}}[g(\mathbf{Y})] - \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{C}_{\mathbf{X}}}[g(\mathbf{Y})]\Big)\right]\right|\\ &= \left|\mathop{\mathbf{E}}_{\mathbf{X}\in\{\pm1\}^{mk}}\left[f(\mathbf{X})\sum_{\mathbf{Y}\in\{\pm1\}^{mk}} g(\mathbf{Y})\big(\mathbf{Pr}_{\mathcal{D}_{\mathbf{X}}}[\mathbf{Y}] - \mathbf{Pr}_{\mathcal{C}_{\mathbf{X}}}[\mathbf{Y}]\big)\right]\right|\\ &= \left|\sum_{X,Y\in\{\pm1\}^{mk}} f(X)g(Y)\left(\mathbf{Pr}_{\substack{\mathbf{X}\sim\mathrm{Unif}(\{\pm1\}^{mk})\\ \mathbf{Y}\sim\mathcal{D}_X}}[\mathbf{X} = X, \mathbf{Y} = Y] - \mathbf{Pr}_{\substack{\mathbf{X}\sim\mathrm{Unif}(\{\pm1\}^{mk})\\ \mathbf{Y}\sim\mathcal{C}_X}}[\mathbf{X} = X, \mathbf{Y} = Y]\right)\right|\\ &\leq \sum_{X,Y\in\mathrm{Supp}} \big|f(X)\big|\big|g(Y)\big|\left|\mathbf{Pr}_{\substack{\mathbf{X}\sim\mathrm{Unif}(\{\pm1\}^{mk})\\ \mathbf{Y}\sim\mathcal{D}_X}}[\mathbf{X} = X, \mathbf{Y} = Y] - \mathbf{Pr}_{\substack{\mathbf{X}\sim\mathrm{Unif}(\{\pm1\}^{mk})\\ \mathbf{Y}\sim\mathcal{C}_X}}[\mathbf{X} = X, \mathbf{Y} = Y]\right|\\ &\hspace{30em}\text{(Support of } f\text{)}\\ &\leq \sum_{X,Y\in\mathrm{Supp}} \big|f(X)\big|\big|g(Y)\big|\big|p_0(X,Y) - p_1(X,Y)\big| \qquad \text{(Define } p_0, p_1 \text{ for notational convenience.)}\\ &\leq \sqrt{\sum_{X,Y\in\mathrm{Supp}} f(X)^2\big|p_0(X,Y) - p_1(X,Y)\big|}\sqrt{\sum_{X,Y\in\mathrm{Supp}} g(Y)^2\big|p_0(X,Y) - p_1(X,Y)\big|}\\ &= \sqrt{\sum_{X,Y\in\mathrm{Supp}} f(X)^2\big|p_0(X,Y) - p_1(X,Y)\big|}\sqrt{\sum_{X,Y\in\mathrm{Supp}} g(Y)^2\big|p_0(Y,X) - p_1(Y,X)\big|}\\ &\hspace{24em}\text{(Self-adjointness of } A \text{ and } B\text{)}\end{aligned}$$

$$= \sqrt{\sum_{X \in \mathrm{Supp}} f(X)^2 \sum_{Y \in \mathrm{Supp}} \left|p_0(X,Y) - p_1(X,Y)\right|} \sqrt{\sum_{X \in \mathrm{Supp}} g(X)^2 \sum_{Y \in \mathrm{Supp}} \left|p_0(X,Y) - p_1(X,Y)\right|}. \qquad \square$$

To employ this comparison bound for similar matrices, we use the following fact to formally show that when k is not too large, sampling k elements without replacement is similar to sampling k elements with replacement.

Fact 40. If $k \le \sqrt{N}$ then

$$\prod_{i=0}^{k-1} \frac{N}{N-i} \le 1 + \frac{k^2}{N}.$$

Proof. We prove by induction on i. If $i = 0$ then the result is trivially true. Now assume that $\prod_{i=0}^{k-2} \frac{N}{N-i} \le 1 + \frac{(k-1)^2}{N}$. Then

$$\prod_{i=0}^{k-1} \frac{N}{N-i} \le \left(1 + \frac{(k-1)^2}{N}\right)\left(1 + \frac{k-1}{N-k+1}\right) \le \left(1 + \frac{(k-1)^2}{N}\right)\left(1 + \frac{k}{N}\right) \le \left(1 + \frac{k^2}{N}\right).$$

The inequalities follow because $k \le \sqrt{N}$. $\square$

The Badly-Behaved Region. There is one issue: resampling some set of $m-1$ coordinates with replacement and without replacement are actually quite different when restricted to the domain of $X \in \{\pm 1\}^{mk}$ such that X^i and X^j are only Hamming distance 1 apart. We use a different tool to handle this "badly behaved" region of the domain $\{\pm 1\}^{mk}$. The intuition is in this region of the graph, we can almost directly show that expansion is good by analyzing the escape probabilities from this region, so that the contribution of functions supported on this badly-behaved region to the spectral gap of $R_{m,m-1,k}$ is already small.

Formally, the following lemma relates the escape probabilities from a region of the Markov chain to the quadratic form induced by its transition operator.

Lemma 41. Suppose A is a random walk matrix on a domain $\mathcal{U}$ such that the uniform distribution on $\mathcal{U}$ is a stationary distribution for A. That is, for every $X \in \mathcal{U}$ there is a distribution $\mathcal{D}_X$ such that for any $f : \mathcal{U} \to \mathbb{R}$,

$$(Af)(X) = \mathop{\mathbf{E}}_{\mathbf{Y} \sim \mathcal{D}_X} \left[f(\mathbf{Y})\right],$$

and $A\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the constant 1s vector.

Let $f, g : \mathcal{U} \to \mathbb{R}$ be such that for any $X \in \mathrm{Supp}(f)$,

$$\mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X} \left[\mathbf{Y} \in \mathrm{Supp}(g)\right] \le \varepsilon.$$

Then

$$\left|\langle f, Ag \rangle\right| \le \varepsilon \|f\|_2 \|g\|_2 .$$

Proof. We directly compute:

$$\begin{aligned} &\left|\langle f, Ag \rangle\right| \\ =&\left| \mathop{\mathbf{E}}_{\mathbf{X} \in \mathcal{U}, \mathbf{Y} \sim \mathcal{D}_{\mathbf{X}}} \left[f(\mathbf{X}) g(\mathbf{Y})\right] \right| \end{aligned}$$

$$
\begin{aligned}
&= \left| \mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U},\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[f(\mathbf{X})\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right] g(\mathbf{Y})\mathbf{1}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]\right]\right| \\
&= \left| \mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[f(\mathbf{X})\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right] \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})\mathbf{1}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]\right]\right]\right| \\
&\le \left| \|f\|_2 \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right]^2 \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})\mathbf{1}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]\right]^2\right]}\right| \qquad \text{(Cauchy-Schwarz)} \\
&\le \left| \|f\|_2 \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right]^2 \left(\sqrt{\mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})^2\right] \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[\mathbf{1}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]^2\right]}\right)^2\right]}\right| \\
&\qquad\qquad \text{(Cauchy-Schwarz)} \\
&= \left| \|f\|_2 \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right]^2 \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})^2\right] \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[\mathbf{1}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]^2\right]\right]}\right| \\
&= \left| \|f\|_2 \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[\mathbf{1}\left[\mathbf{X}\in \mathrm{Supp}(f)\right] \mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[\mathbf{Y}\in\mathrm{Supp}(g)\right] \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})^2\right]\right]}\right|.
\end{aligned}
$$

Now, consider each possible value that $\mathbf{X}$ may take. If $\mathbf{X}\in\mathrm{Supp}(f)$ then by the assumption, we have that $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]\le\varepsilon$. Otherwise if $\mathbf{X}\notin\mathrm{Supp}(f)$ we have that $\mathbf{1}\left[\mathbf{X}\in\mathrm{Supp}(f)\right]=0$. In either case, we have $\mathbf{1}\left[\mathbf{X}\in\mathrm{Supp}(f)\right]\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}}\left[\mathbf{Y}\in\mathrm{Supp}(g)\right]\le\varepsilon$. Continuing with our calculation, we find that

$$
\left|\langle f, Ag\rangle\right| \le \|f\|_2 \left| \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U},\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[\varepsilon \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})^2\right]\right]}\right| = \varepsilon\|f\|_2 \left| \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U},\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}} \left[g(\mathbf{Y})^2\right]}\right|.
$$

At this point, we note that sampling $\mathbf{X}$ uniformly at random and then sampling $\mathbf{Y}\sim\mathcal{D}_{\mathbf{X}}$ gives the same distribution as sampling $\mathbf{Y}$ uniformly by stationarity of the uniform distribution. This completes the proof by the following:

$$
\left|\langle f, Ag\rangle\right| \le \varepsilon\|f\|_2 \left| \sqrt{\mathop{\mathbf{E}}_{\mathbf{X}\in\mathcal{U}} \left[g(\mathbf{X})^2\right]}\right| = \varepsilon\|f\|_2\|g\|_2 . \qquad \square
$$

The proof of Theorem 29 follows a similar principle, except that we choose a different way to partition our domain into a "well-behaved" region and a "badly-behaved" region.

5 Proof of Theorem 24

Throughout this section fix $m\ge 3$. Our goal in this section is to establish Theorem 24, which states that $R_{m,m-1,k}-R_{m,m,k}$ has small spectral norm. Informally, we show that completely randomizing $m-1$ out of m wires in a reversible circuit is very similar to randomizing all m wires. Recall that $R_{m,m-1,k}$ is defined via the distributions $\mathcal{D}_X^{m,m-1,k}$ for $X\in\{\pm1\}^{mk}$, from which one samples by sampling a random set $\mathbf{S}\subseteq[m]$ with $|\mathbf{S}|=1$ (so really $\mathbf{S}=\{\mathbf{a}\}$, where $\mathbf{a}$ is a random element of $[m]$), "fixes" the entries $X^i_{\mathbf{a}}$, and applies a random permutation to the coordinates not equal to $\mathbf{a}$. We use this notation and terminology of a "fixed" coordinate $\mathbf{a}$ throughout this section.

As alluded to in Section 4, our proof will decompose the space $\mathbb{R}^{\{\pm 1\}^{mk}}$ on which these operators act into three orthogonal components, to be defined in Section 5.1. Then Section 5.1, Section 5.2, and Section 5.4 will bound the contributions from vectors lying in these orthogonal components and their cross terms.

5.1 An Orthogonal Decomposition

Definition 42. Regard elements of $\{\pm 1\}^{mk}$ as k-by-m matrices, so that the ith row of X is X^i, and the ath column of X is X_a. Define

$$B_{\geq 2} = \left\{X \in \{\pm 1\}^{mk} : \forall i \neq j \in [k], d\Big(X^i, X^j\Big) \geq 2\right\},$$
$$B_{=1} = \left\{X \in \{\pm 1\}^{mk} : \forall i \neq j \in [k], d\Big(X^i, X^j\Big) \geq 1\right\} \setminus B_{\geq 2},$$
$$B_{=0} = \left\{X \in \{\pm 1\}^{mk} : \exists i \neq j \in [k], d\Big(X^i, X^j\Big) = 0\right\}.$$

Our proof that $R_{m,m-1,k} - R_{m,m,k}$ has small spectral norm will go by induction on k. Lemma 43 helps to connect the cases of $k-1$ and k in the proof. In particular, it shows that we can pass those functions supported on $B_{=0}$ into the induction.

Lemma 43. Let $f : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{=0}$. Then for any $S_1, \ldots, S_t \subseteq [m]$ and $c_1, \ldots, c_t \in \mathbb{R}$ we have

$$\left|\left\langle f, \sum_{s=1}^{t} c_s R_{m,S_s,k} f \right\rangle\right| \leq \left\|\sum_{s=1}^{t} c_s R_{m,S_s,k-1}\right\|_{\text{op}}.$$

Proof. For any map $\varphi : [k] \to [k-1]$ (viewed as a coloring of $[k]$ with $k-1$ colors) define the set

$$\mathcal{J}_\varphi = \left\{\Big(X^1, \ldots, X^k\Big) : X^i = X^j \iff \varphi(i) = \varphi(j)\right\}.$$

These sets $\mathcal{J}_\varphi$ partition $B_{=0}$. Thus, for every f supported on $B_{=0}$ we have a decomposition $f = \sum_\varphi f_\varphi$, where each f_φ is supported on $\mathcal{J}_\varphi$.

Now, for each φ define a map $\mathrm{Res}_\varphi : \left\{f : \{\pm 1\}^{mk} \to \mathbb{R} : f \text{ supported on } \mathcal{J}_\varphi\right\} \to \mathbb{R}^{\{\pm 1\}^{m(k-1)}}$ by arbitrarily choosing $i, j \in [k]$ such that $\varphi(i) = \varphi(j)$ and defining for $f_\varphi : \{\pm 1\}^{mk} \to \mathbb{R}$ supported on $\mathcal{J}_\varphi$ the new function $\mathrm{Res}_\varphi f_\varphi : \{\pm 1\}^{m(k-1)} \to \mathbb{R}$ by defining for $X' \in \{\pm 1\}^{m(k-1)}$

$$\mathrm{Res}_\varphi f_\varphi(X') = \frac{f(\mathrm{Res}^*_\varphi(X'))}{2^{m/2}}.$$

where $\mathrm{Res}^*_\varphi(X')$ is the unique element of $\mathcal{J}_\varphi$ such that $(\mathrm{Res}^*_\varphi(X'))^{[k]\setminus\{j\}} = X'$. Note this is well-defined because f_φ is supported on $\mathcal{J}_\varphi$.

Claim 44. For any $f_\varphi : \{\pm 1\}^{mk} \to \mathbb{R}$ supported on $\mathcal{J}_\varphi$ we have $\left\|f_\varphi\right\|_2 = \left\|\mathrm{Res}_\varphi f_\varphi\right\|_2$ and for any $S \subseteq [m]$,

$$\left\langle f_\varphi, R_{m,S,k} f_\varphi \right\rangle = \left\langle \mathrm{Res}_\varphi f_\varphi, R_{m,S,k} \mathrm{Res}_\varphi f_\varphi \right\rangle.$$

We prove the claim later, and for now use it to compute

$$\begin{aligned}
&\left\langle f, \sum_{s=1}^{t} c_s R_{m,S_s,k} f \right\rangle \\
=&\sum_{\varphi,\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle \\
=&\sum_{\varphi} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi} \right\rangle + \sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle \\
=&\sum_{\varphi} \left\langle \mathrm{Res}_\varphi f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} \mathrm{Res}_\varphi f_{\varphi} \right\rangle + \sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle && \text{(Claim 44)} \\
\le&\sum_{\varphi} \left\| \sum_{s=1}^{t} c_s R_{m,S_s,k-1} \right\|_{\mathrm{op}}^2 \left\| \mathrm{Res}_\varphi f_\varphi \right\|_2^2 + \sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle \\
=&\left\| \sum_{s=1}^{t} c_s R_{m,S_s,k-1} \right\|_{\mathrm{op}}^2 \sum_{\varphi} \left\| f_\varphi \right\|_2^2 + \sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle \\
=&\left\| \sum_{s=1}^{t} c_s R_{m,S_s,k-1} \right\|_{\mathrm{op}}^2 \langle f, f \rangle + \sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle .
\end{aligned}$$

The last equality follows from orthogonality of the f_φ.

To take care of the cross terms, we observe that for any $X \in \mathcal{J}_\varphi$ and $S \subseteq [n]$, we have $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,S,k}} \left[\mathbf{Y} \in \mathcal{J}_{\varphi'}\right] = \mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,S,k}} \left[\mathbf{Y} \in \mathcal{J}_{\varphi'}\right] = 0$ for any $\varphi' \neq \varphi$. Then by Lemma 41 we have

$$\sum_{\varphi\neq\varphi'} \left\langle f_\varphi, \sum_{s=1}^{t} c_s R_{m,S_s,k} f_{\varphi'} \right\rangle = 0. \qquad \square$$

Proof of Claim 44. Without loss of generality assume that $\varphi(k-1) = \varphi(k)$ so we can regard $\mathrm{Res}_\varphi f_\varphi$ as a real function on $\{\pm1\}^{m(k-1)}$. Then

$$\begin{aligned}
\left\langle f_\varphi, f_\varphi \right\rangle =& \mathop{\mathbf{E}}_{\mathbf{X}\in\{\pm1\}^{mk}} \left[f_\varphi(\mathbf{X})^2 \right] = \frac{1}{2^{mk}} \sum_{X\in\{\pm1\}^{mk}} f_\varphi(X)^2 = \frac{1}{2^{m(k-1)}} \sum_{X\in\mathcal{J}_\varphi} \left(\frac{f_\varphi(X)}{2^{m/2}} \right)^2 \\
=& \frac{1}{2^{m(k-1)}} \sum_{X'\in\{\pm1\}^{m(k-1)}} \left(\mathrm{Res}_\varphi f(X') \right)^2 = \left\langle \mathrm{Res}_\varphi f_\varphi, \mathrm{Res}_\varphi f_\varphi \right\rangle .
\end{aligned}$$

To prove the second statement, we compute

$$\begin{aligned}
&\left\langle f_\varphi, R_{m,S,k} f_\varphi \right\rangle \\
=& \mathop{\mathbf{E}}_{\mathbf{X}\in\{\pm1\}^{mk}} \left[f_\varphi(\mathbf{X}) \mathop{\mathbf{E}}_{\mathbf{Y}\sim\mathcal{D}_X^{m,S,k}} \left[f_\varphi(\mathbf{Y}) \right] \right] \\
=& \frac{1}{2^{mk}} \sum_{X\in\{\pm1\}^{mk}} f_\varphi(X) \sum_{Y\in\{\pm1\}^{mk}} f_\varphi(Y) \, \mathbf{Pr}_{\mathcal{D}_X^{m,S,k}} \left[\mathbf{Y} = Y \right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2^{mk}} \sum_{X \in \mathcal{J}_\varphi} f_\varphi(X) \sum_{Y \in \mathcal{J}_\varphi} f_\varphi(Y) \, \mathbf{Pr}_{\mathcal{D}_X^{m,S,k}} [\mathbf{Y} = Y] \\
&= \frac{1}{2^{m(k-1)}} \sum_{X \in \mathcal{J}_\varphi} f_\varphi(X) \sum_{Y \in \mathcal{J}_\varphi} f_\varphi(Y) \, \mathbf{Pr}_{\mathcal{D}_X^{m,S,k}} [\mathbf{Y} = Y] \\
&= \frac{1}{2^{m(k-1)}} \sum_{X' \in \{\pm 1\}^{m(k-1)}} \mathrm{Res}_\varphi f_\varphi(X') \sum_{Y' \in \{\pm 1\}^{m(k-1)}} \mathrm{Res}_\varphi f_\varphi(Y') \, \mathbf{Pr}_{\mathcal{D}_{X'}^{m,S,k-1}} \left[\mathbf{Y} = Y'\right] \\
&= \left\langle \mathrm{Res}_\varphi f_\varphi, R_{m,S,k} \mathrm{Res}_\varphi f_\varphi \right\rangle .
\end{aligned}
$$

The second-to-last equality follows because the corresponding X, Y are in $\mathcal{J}_\varphi$. □

We now prove Theorem 24, deferring proofs of the remaining needed auxiliary results to Section 5.2 and Section 5.4.

Theorem 45 (Theorem 24 restated)**.** Let $m \geq 100$ and assume that $k \leq 2^{m/3}$. Given any $f : \{\pm 1\}^{mk} \to \mathbb{R}$, we have

$$
\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \leq \left(\frac{1}{m} + \frac{k^2}{2^{m/3}} \right) \langle f, f \rangle .
$$

Proof. We prove by induction on k. In the base case $k = 1$ and the result holds by the following argument when we write $f = f_2$. Now assume that the result holds for real functions on $\{\pm 1\}^{m(k-1)}$.

Let $f : \{\pm 1\}^{mk} \to \mathbb{R}$. Write $f = f_0 + f_1 + f_2$ where f_0 is supported on $B_{=0}$, f_1 is supported on $B_{=1}$, and f_2 is supported on $B_{\geq 2}$. By Lemma 53 applied with both $R_{m,m-1,k}$ and $R_{m,m,k}$ and $B_{=0}$ and $\{\pm 1\}^{mk} \setminus B_{=0} = B_{=1} \cup B_{\geq 2}$, the other cross terms vanish, and we have

$$
\begin{aligned}
&\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \\
\leq &\left| \left\langle f_0, (R_{m,m-1,k} - R_{m,m,k}) f_0 \right\rangle \right| + \left| \left\langle f_1, (R_{m,m-1,k} - R_{m,m,k}) f_1 \right\rangle \right| \\
&\quad + \left| \left\langle f_2, (R_{m,m-1,k} - R_{m,m,k}) f_2 \right\rangle \right| + \left| 2 \left\langle f_1, (R_{m,m-1,k} - R_{m,m,k}) f_2 \right\rangle \right| \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Self-adjointness (Fact 18))} \\
\leq &\left(\frac{1}{m} + \frac{(k-1)^2}{2^{m/2}} \right) \langle f_0, f_0 \rangle + \left(\frac{1}{m} + \frac{k^2 m^5}{2^{m-2}} \right) \langle f_1, f_1 \rangle + \left(\frac{1}{m} + \frac{k^2}{2^{m/2}} \right) \langle f_2, f_2 \rangle + \frac{2k^2 m}{2^{m-3}} \|f_1\|_2 \|f_2\|_2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Lemma 43 + induction, Corollary 51, Proposition 60, Lemma 52, in that order)} \\
\leq &\left(\frac{1}{m} + \frac{k^2}{2^{m/2}} \right) \langle f, f \rangle + \frac{2k^2 m^5}{2^{m-3}} \|f_1\|_2 \|f_2\|_2 \\
\leq &\left(\frac{1}{m} + \frac{k^2}{2^{m/2}} \right) \langle f, f \rangle + \frac{k^2}{2^{m-20}} \langle f, f \rangle \\
= &\left(\frac{1}{m} + \frac{k^2}{2^{m/3}} \right) \langle f, f \rangle .
\end{aligned}
$$

□

5.2 f Supported on $B_{=1}$

We now use Lemma 41 to bound $\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \leq \left| \left\langle f, R_{m,m-1,k} f \right\rangle \right|$ when f is supported only on $B_{=1}$ and the cross terms contributed. We first define a partition of $B_{=1}$, and apply

Lemma 41 to these different parts. One of our main observations to bound f supported on $B_{=1}$ is the observation that when k is small, it is highly unlikely that two vectors out of any k are close to each other. Thus, a random walk beginning in $B_{=1}$ and obeying the transition probabilities given by $B_{m,m-1,k}$ will rarely remain in $B_{=1}$. Formally, Lemma 41 bounds the contributions to the spectral norm by these transition probabilities.

Definition 46. For $S \subseteq [m]$ define the set $\mathcal{I}_S \subseteq \{\pm 1\}^{mk}$ by

$$\mathcal{I}_S = \left\{ X \in B_{=1} : \forall a \in S, \exists i,j \in [k] : \Delta\left(X^i, X^j\right) = \{a\} \right\}.$$

Unfortunately, the sets $\mathcal{I}_S$ for different $S \subseteq [m]$ do not form a partition of $B_{=1}$, since there is overlap between $\mathcal{I}_S$ and $\mathcal{I}_T$ for $S \neq T$. However, we can artificially make this into a partition.

Definition 47. For $S \subseteq [m]$ with $|S| = 1$ (so $S = \{a\}$ for some $a \in [m]$) define the set $\widetilde{\mathcal{I}}_S \subseteq \{\pm 1\}^{nk}$ by

$$\widetilde{\mathcal{I}}_S = \mathcal{I}_S \setminus \bigcup_{a' \neq a} \mathcal{I}_{\{a'\}}.$$

Now place an arbitrary ordering $\preceq$ on the set $\{S \subseteq [m] : |S| = 2\}$ and define

$$\widetilde{\mathcal{I}}_S = \mathcal{I}_S \setminus \left(\bigcup_{S' \subseteq [m] : |S'| = 2, S' \preceq S} \mathcal{I}_{S'} \right).$$

Observation 48. The collection of sets $\{\widetilde{\mathcal{I}}_S : S \subseteq [m], |S| \leq 2\}$ forms a partition of $B_{=1}$.

Lemma 49. Let $S \subseteq [m]$ be such that $|S| \leq 2$. If $k \geq 2$ and $f : \{\pm 1\}^{mk} \to \mathbb{R}$ is supported on $\widetilde{\mathcal{I}}_S$ then

$$\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \leq \left(\frac{1}{m} + \frac{mk^2}{2^{m-1}} \right) \langle f, f \rangle .$$

Proof. We bound $\left| \left\langle f, R_{m,m-1,k} f \right\rangle \right| \leq \left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right|$. This inequality is true because $R_{m,m,k}$ and $R_{m,m-1,k} - R_{m,m,k}$ are PSD by Fact 17. Suppose first that $|S| = 1$ so that $S = \{a\}$ for some $a \in [m]$. Let $X \in \widetilde{\mathcal{I}}_S$. Then

$$\begin{aligned}
&\mathbf{Pr}\left[R_{m,m-1,k} X \in \widetilde{\mathcal{I}}_S \right] \\
=&\frac{1}{m} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[\mathbf{Y} \in \widetilde{\mathcal{I}}_{\{a\}} | \mathbf{a} = a \right] + \frac{m-1}{m} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[\mathbf{Y} \in \widetilde{\mathcal{I}}_{\{a\}} | \mathbf{a} \neq a \right] \\
&\qquad \text{(Recall } \mathbf{a} \text{ is the "fixed" coordinate.)} \\
\leq&\frac{1}{m} + \frac{m-1}{m} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[\mathbf{Y} \in \widetilde{\mathcal{I}}_{\{a\}} | \mathbf{a} \neq a \right] \\
\leq&\frac{1}{m} + \sum_{i,j \in [k]} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[d\left(\mathbf{Y}^i, \mathbf{Y}^j\right) \leq 1 | \mathbf{a} \neq a \right] \\
\leq&\frac{1}{m} + \binom{k}{2} \frac{m-1}{2^{m-1}-1} \\
\leq&\frac{1}{m} + \frac{mk^2}{2^{m-1}}.
\end{aligned}$$

Then applying Lemma 41, while noticing that any $R_{m,m',k}$ has the uniform distribution over $\{\pm 1\}^{mk}$ as a stationary distribution (Fact 19), completes the proof.

A similar calculation shows that if $|S| = 2$ and $X \in \widetilde{\mathcal{I}}_S$ then $\mathbf{Pr}\left[R_{m,m-1,k}X \in \widetilde{\mathcal{I}}_S\right] \le \frac{mk^2}{2^{m-1}} \le \frac{1}{m} + \frac{mk^2}{2^{m-1}}$, and again Lemma 41 completes the proof. □

Lemma 50. Let $S \neq T \subseteq [m]$ be such that $|S|, |T| \le 2$. If $k \ge 2$ and $f : \{\pm 1\}^{mk} \to \mathbb{R}$ is supported on $\widetilde{\mathcal{I}}_S$ and $g : \{\pm 1\}^{mk} \to \mathbb{R}$ is supported on $\widetilde{\mathcal{I}}_T$ then

$$\left|\left\langle f, (R_{m,m-1,k} - R_{m,m,k})g\right\rangle\right| \le \frac{mk^2}{2^{m-1}}\|f\|_2\|g\|_2 .$$

Proof. Let $X \in \widetilde{\mathcal{I}}_S$. Let $a \in T \setminus S$ be such that there does not exist $i, j \in [k]$ such that $X \in \mathcal{I}_{\{a\}}$. Then

$$\begin{aligned}
&\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m-1,k}}\left[\mathbf{Y} \in \widetilde{\mathcal{I}}_T\right] \le \mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m-1,k}}\left[\mathbf{Y} \in \widetilde{\mathcal{I}}_{\{a\}}\right] \le \mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m-1,k}}\left[\mathbf{Y} \in \mathcal{I}_{\{a\}}\right] \\
&\le \sum_{i,j\in[k]} \mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m-1,k}}\left[d\left(\mathbf{Y}^i, \mathbf{Y}^j\right) = 1\right] \le \binom{k}{2}\frac{m-1}{2^{m-1}-1}. && (X \notin \mathcal{I}_{\{a\}})
\end{aligned}$$

Then applying Lemma 41 gives

$$\left|\left\langle f, R_{m,m-1,k}g\right\rangle\right| \le \frac{mk^2}{2^{m-1}}\|f\|_2\|g\|_2 .$$

A similar calculation shows that $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m,k}}\left[\mathbf{Y} \in \widetilde{\mathcal{I}}_T\right] \le \frac{k^2}{2^m}$. Then Lemma 41 gives

$$\left|\left\langle f, R_{m,m,k}g\right\rangle\right| \le \frac{k^2}{2^m}\|f\|_2\|g\|_2 .$$

Applying the triangle inequaltiy completes the proof. □

Corollary 51. Assume $k \ge 2$. Let $f : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{=1}$. Then

$$\left|\left\langle f, (R_{m,m-1,k} - R_{m,m,k})f\right\rangle\right| \le \left(\frac{1}{m} + \frac{k^2m^5}{2^{m-1}}\right)\langle f, f\rangle .$$

Proof. Write $f = \sum_{S\subseteq[m]:|S|\le 2} f_S$ where each f_S is supported on $\mathcal{I}_S$. Then

$$\begin{aligned}
&\left|\left\langle f, (R_{m,m-1,k} - R_{m,m,k}f)\right\rangle\right| \\
&\le \left|\left\langle f, R_{m,m-1,k}f\right\rangle\right| && (R_{m,m,k} \text{ and } R_{m,m-1,k} - R_{m,m,k} \text{ both PSD Fact 17}) \\
&= \sum_{S,T\subseteq[m]:|S|,|T|\le 2}\left|\left\langle f_S, R_{m,m-1,k}f_T\right\rangle\right| \\
&\le \sum_{S\subseteq[m]:|S|\le 2}\left|\left\langle f_S, R_{m,m-1,k}f_S\right\rangle\right| + \sum_{S\neq T\subseteq[m]:|S|,|T|\le 2}\left|\left\langle f_S, R_{m,m-1,k}f_T\right\rangle\right| \\
&\le \left(\frac{1}{m} + \frac{mk^2}{2^{m-1}}\right)\sum_{S\subseteq[m]:|S|\le 2}\langle f_S, f_S\rangle + \frac{mk^2}{2^{m-1}}\sum_{S\neq T\subseteq[m]:|S|,|T|\le 2}\|f_S\|_2\|f_T\|_2 && (\text{Lemma 49, Lemma 50})
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{m} + \frac{mk^2}{2^{m-1}}\right)\langle f, f\rangle + \frac{mk^2}{2^{m-1}} \sum_{S \neq T \subseteq [m] : |S|, |T| \leq 2} \|f_S\|_2 \|f_T\|_2 \\
&\leq \left(\frac{1}{m} + \frac{mk^2}{2^{m-1}}\right)\langle f, f\rangle + \frac{mk^2}{2^{m-1}} \sum_{S \neq T \subseteq [m] : |S|, |T| \leq 2} \langle f, f\rangle \\
&\leq \left(\frac{1}{m} + \frac{mk^2}{2^{m-1}}\right)\langle f, f\rangle + \frac{m^5k^2}{2^{m-1}}\langle f, f\rangle \\
&\leq \left(\frac{1}{m} + \frac{m^5k^2}{2^{m-1}}\right)\langle f, f\rangle .
\end{aligned}
$$

Note that we can apply Lemma 49 and Lemma 50 because $k \geq 2$. □

5.3 Cross Terms

We can use the same idea to bound the contributions from the cross terms.

Lemma 52. Let $f_1 : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{=1}$ and let $f_2 : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{\geq 2}$. Then

$$
\left|\left\langle f_1, (R_{m,m-1,k} - R_{m,m,k}) f_2 \right\rangle\right| \leq \frac{k^2 m}{2^{m-2}} \|f_1\|_2 \|f_2\|_2 .
$$

Proof. For $X \in B_{\geq 2}$ we have $\mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}}[\mathbf{Y} \in B_{=1}] \leq \frac{k^2 m}{2^{m-1}}$. Apply Lemma 41 to find that

$$
\left|\left\langle f_1, R_{m,m-1,k} f_2 \right\rangle\right| \leq \frac{k^2 m}{2^{m-1}} \|f_1\|_2 \|f_2\|_2 .
$$

For $X \in B_{\geq 2}$ we have $\mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m,k}}[\mathbf{Y} \in B_{=1}] \leq \frac{k^2}{2^m}$. Apply Lemma 41 to find that

$$
\left|\left\langle f_1, R_{m,m,k} f_2 \right\rangle\right| \leq \frac{k^2}{2^m} \|f_1\|_2 \|f_2\|_2 .
$$

Applying the triangle inequality completes the proof. □

Lemma 53. Let $f_0 : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{=0}$ and let $f_1 : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{=1} \cup B_{\geq 2}$. Then

$$
\left|\left\langle f_0, (R_{m,m-1,k} - R_{m,m,k}) f_1 \right\rangle\right| \leq 0.
$$

Proof. For $X \in B_{\geq 2} \cup B_{=1}$ we have $\mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}}[\mathbf{Y} \in B_{=0}] = \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m,k}}[\mathbf{Y} \in B_{=0}] = 0$. Apply Lemma 41 to bound

$$
\left|\left\langle f_0, (R_{m,m-1,k} - R_{m,m,k}) f_1 \right\rangle\right| \leq \left|\left\langle f_0, R_{m,m-1,k} f_1 \right\rangle\right| + \left|\left\langle f_0, R_{m,m,k} f_1 \right\rangle\right| \leq 0. \qquad \square
$$

5.4 A Hybrid Argument for f supported on $B_{\geq 2}$

In this section we bound the square terms $\langle f_2, (R_{m,m-1,k} - R_{m,m,k}) f_2 \rangle$ for f_2 supported on $B_{\geq 2}$. As mentioned in Section 4, our key idea is that when k is small compared to m, the fraction of $X \in \{\pm 1\}^{mk}$ with two identical columns is so small that applying the noise by randomly permuting the rows is almost the same as randomly replacing the rows with completely random rows. That is, sampling without replacement resembles sampling with replacement closely.

This observation allows us to pass from the random walk described by $R_{m,m-1,k}$ to a different random walk described by a nicer noise model described by operators we will call $Q_{m,m-1,k}$. The key tool we use is the bound given by Lemma 39 for relating total-variation distances between Markov chain transition probabilities to a more linear-algebraic notion of closeness, stated in terms of their transition matrices. Fourier-analytic techniques will then be useful to bound the spectral norm of these Q-operators.

Definition 54. We define four operators[7] $R_{m,m-1,k}$, $Q_{m,m-1,k}$, $R_{m,m,k}$, and $Q_{m,m,k}$ on $\mathbb{R}^{\{\pm 1\}^{mk}}$.

- Define
$$(R_{m,m-1,k} f)(X) = \mathop{\mathbf{E}}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \big[f(\mathbf{Y}) \big].$$
- To define $Q_{m,m-1,k}$, for any $X \in \{\pm 1\}^{mk}$ we define the distribution $\mathcal{C}_X^{m,m-1,k}$ as follows. To sample $\mathbf{X}$ from $\mathcal{C}_X^{m,m-1,k}$, we sample $\mathbf{a} \in [m]$ uniformly randomly and set $\mathbf{X}^i_{\mathbf{a}} = X^i_{\mathbf{a}}$ for all $i \in [k]$. Then set $\mathbf{X}^i_a$ uniformly randomly for $a \neq \mathbf{a}$. Then
$$(Q_{m,m-1,k} f)(X) = \mathop{\mathbf{E}}_{\mathbf{Y} \sim \mathcal{C}_X^{m,m-1,k}} \big[f(\mathbf{Y}) \big].$$
- Define
$$(R_{m,m,k} f)(X) = \mathop{\mathbf{E}}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m,k}} \big[f(\mathbf{Y}) \big].$$
- The operator $Q_{m,m,k}$ is defined by setting for each $f : \{\pm 1\}^{mk} \to \mathbb{R}$
$$(Q_{m,m,k} f)(X) = \mathop{\mathbf{E}}_{\mathbf{Y} \sim \mathrm{Unif}\left(\{\pm 1\}^{mk}\right)} \big[f(\mathbf{Y}) \big].$$

Fact 55. The matrices $Q_{m,m-1,k}$ and $Q_{m,m,k}$ are self-adjoint and PSD for any m, k.

We intend to show that $\left| \langle f, (R_{m,m-1,k} - R_{m,m,k}) f \rangle \right|$ is small for f supported on $B_{\geq 2}$. We do so by a hybrid argument. In the following inequality, the first (and last) term on the RHS will be bounded by a simple bound on the total variation distance between the distribution $\mathcal{C}_X^{m,m-1,k}$ and $\mathcal{D}_X^{m,m-1,k}$ ($\mathcal{C}_X^{m,m,k}$ and $\mathcal{D}_X^{m,m,k}$). As mentioned before, the second term on the RHS will be bounded using Fourier analysis.

$$\begin{aligned}
&\left| \langle f, (R_{m,m-1,k} - R_{m,m,k}) f \rangle \right| \\
\leq &\left| \left\langle f, (R_{m,m-1,k} - Q_{m,m-1,k}) f \right\rangle \right| + \left| \left\langle f, (Q_{m,m,k} - Q_{m,m-1,k}) f \right\rangle \right| + \left| \left\langle f, (R_{m,m,k} - Q_{m,m,k}) f \right\rangle \right|.
\end{aligned}$$

[7] $R_{m,m-1,k}$ and $R_{m,m,k}$ have already been defined, but we define them again here for ease of comparison.

5.4.1 The First Hybrid: $R_{m,m-1,k}$ to $Q_{m,m-1,k}$

Lemma 56. Assume that $k \le 2^{m/3}$. For any $f : \{\pm 1\}^{mk} \to \mathbb{R}$ supported on $B_{\ge 2}$ and we have that

$$\left|\left\langle f, (R_{m,m-1,k} - Q_{m,m-1,k}) f \right\rangle\right| \le \frac{k^2}{2^{m-1}} \langle f, f \rangle .$$

Proof. We directly compute (using the appropriate definitions of p_0 and p_1 given in Lemma 39):

$$\begin{aligned}
&\left|\left\langle f, (R_{m,m-1,k} - Q_{m,m-1,k}) f \right\rangle\right| \\
\le &\sum_{X \in B_{\ge 2}} f(X)^2 \sum_{Y \in B_{\ge 2}} \left|p_0(X,Y) - p_1(X,Y)\right| \quad \text{(Lemma 39 + self-adjointness (Fact 18, Fact 55))} \\
\le &\frac{1}{2^{mk}} \cdot \frac{k^2}{2^{m-1}} \sum_{X \in B_{\ge 2}} f(X)^2 \quad \text{(Equation (1) below)} \\
\le &\frac{k^2}{2^{m-1}} \langle f, f \rangle .
\end{aligned}$$

It suffices to establish Equation (1). Assume that $X \in B_{\ge 2}$. Then

$$\begin{aligned}
&\sum_{Y \in B_{\ge 2}} \left|p_0(X,Y) - p_1(X,Y)\right| \\
= &\sum_{Y \in B_{\ge 2}} \left| \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} [\mathbf{Y} = Y] - \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}_X^{m,m-1,k}} [\mathbf{Y} = Y] \right| \\
= &\frac{1}{m} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}} \left| \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[\mathbf{Y} = Y | a \text{ fixed}\right] - \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}_X^{m,m-1,k}} \left[\mathbf{Y} = Y | a \text{ fixed}\right] \right| \\
= &\frac{1}{m} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}, Y_a = X_a} \left| \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m-1,k}} \left[\mathbf{Y} = Y | a \text{ fixed}\right] - \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}_X^{m,m-1,k}} \left[\mathbf{Y} = Y | a \text{ fixed}\right] \right| \\
= &\frac{1}{m 2^{mk}} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}, Y_a = X_a} \left| \prod_{i=0}^{k-1} \frac{1}{2^{m-1} - i} - \frac{1}{2^{(m-1)k}} \right| \quad (k \le 2^{m/3} \le 2^m - 2 \text{ and } X \in B_{\ge 2}) \\
\le &\frac{1}{m 2^{mk}} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}, Y_a = X_a} \left| \frac{1}{2^{(m-1)k}} \left(\prod_{i=0}^{k-1} \frac{2^{m-1}}{2^{m-1} - i} - 1 \right) \right| \\
\le &\frac{1}{m 2^{mk}} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}, Y_a = X_a} \left| \frac{1}{2^{(m-1)k}} \left(1 + \frac{k^2}{2^{m-1}} - 1 \right) \right| \quad (k \le 2^{m/3}, \text{ Fact 40}) \\
= &\frac{1}{m 2^{mk}} \sum_{a \in [m]} \sum_{Y \in B_{\ge 2}, Y_a = X_a} \frac{k^2}{2^{m-1} 2^{(m-1)k}} \\
= &\frac{1}{2^{mk}} \sum_{Y \in B_{\ge 2}, Y_1 = X_1} \frac{k^2}{2^{m-1} 2^{(m-1)k}} \\
\le &\frac{1}{2^{mk}} \cdot \frac{2^{mk}}{2^k} \cdot \frac{k^2}{2^{m-1} 2^{(m-1)k}} \\
= &\frac{k^2}{2^{m-1} 2^{mk}}. \quad (1)
\end{aligned}$$

Note that our computations for $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,m-1,k}}\left[\mathbf{Y}=Y|a \text{ fixed}\right]$ and $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{C}_X^{m,m-1,k}}\left[\mathbf{Y}=Y|a \text{ fixed}\right]$ relied on the fact that $Y \in B_{\geq 2}$. □

5.4.2 The Second Hybrid: $Q_{m,m-1,k}$ to $Q_{m,m,k}$

We use Fourier analysis to analyze the spectrum of the operator $Q_{m,m-1,k} - Q_{m,m,k}$, which will prove that $\left\|Q_{m,m-1,k} - Q_{m,m,k}\right\|_{\text{op}}$ is small. We use the Fourier characters as an eigenbasis for $Q_{m,m-1,k} - Q_{m,m,k}$.

Fact 57. Fix $S_1, \ldots, S_k \subseteq [n]$. Then

$$\left((Q_{m,m-1,k} - Q_{m,m,k})\chi_{S_1,\ldots,S_k}\right) = \begin{cases} \frac{1}{m}\chi_{S_1,\ldots,S_k} & \text{if } S_1 \cup \cdots \cup S_k = \{a\} \text{ for some } a \in [m]. \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $\left|\bigcup_i S_i\right| = 1$ then $S_1 = \cdots = S_k = \emptyset$ and it is clear that

$$Q_{m,m-1,k}\chi_{S_1,\ldots,S_k} = 1 = Q_{m,m,k}\chi_{S_1,\ldots,S_k}.$$

Now assume $\left|\bigcup_i S_i\right| \geq 2$. Then for any $X = (X^1, \ldots, X^k) \in \{\pm 1\}^{mk}$,

$$\begin{aligned}
&Q_{m,m-1,k}\chi_{S_1,\ldots,S_k}(X^1,\ldots,X^k) \\
=&\mathop{\mathbf{E}}_{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k}}\left[\chi_{S_1,\ldots,S_k}(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\right] \\
=&\frac{1}{m}\sum_{a'\in[m]}\mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k} \\ \mathbf{a}=a'}}\left[\chi_{S_1,\ldots,S_k}(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\right] \\
=&\frac{1}{m}\sum_{a'\in[m]}\mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k} \\ \mathbf{a}=a'}}\left[\prod_{i\in[k]}\prod_{a\in S_i}\mathbf{Y}_a^i\right] \\
=&\frac{1}{m}\sum_{a'\in[m]}\mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k} \\ \mathbf{a}=a'}}\left[\left(\prod_{\substack{i\in[k] \\ a'\in S_i}}\mathbf{Y}_{a'}^i\right)\left(\prod_{i\in[k]}\prod_{a\in S_i\setminus\{a'\}}\mathbf{Y}_a^i\right)\right] \\
=&\frac{1}{m}\sum_{a'\in[m]}\left(\prod_{\substack{i\in[k] \\ a'\in S_i}}\mathbf{Y}_{a'}^i\right)\mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k} \\ \mathbf{a}=a'}}\left[\prod_{i\in[k]}\prod_{a\in S_i\setminus\{a'\}}\mathbf{Y}_a^i\right] \\
=&\frac{1}{m}\sum_{a'\in[m]}\left(\prod_{\substack{i\in[k] \\ a'\in S_i}}\mathbf{Y}_{a'}^i\right)\mathop{\mathbf{E}}_{(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\in\left(\{\pm1\}^n\right)^{[k]\setminus\{a'\}}}\left[\chi_{S_1\setminus\{a'\},\ldots,S_k\setminus\{a'\}}(\mathbf{Y}^1,\ldots,\mathbf{Y}^k)\right] \\
=&0. \qquad \text{(since at least one of the } S_i \setminus \{a'\} \text{ is nonempty)}
\end{aligned}$$

It is easy to see that $Q_{m,m,k}\chi_{S_1,\ldots,S_k} = 0$, since $S_1 \cup \cdots \cup S_k \neq \emptyset$.

Now assume that $\bigcup_i S_i = \{a'\}$ for some $a' \in [m]$. Then it is clear that $Q_{m,m,k}\chi_{S_1,\ldots,S_k} = 0$, since $S_1 \cup \cdots \cup S_k \neq \emptyset$. We now compute

$$Q_{m,m-1,k}\chi_{S_1,\ldots,S_k}(X^1,\ldots,X^k)$$

$$
\begin{aligned}
&= \mathop{\mathbf{E}}_{(\mathbf{Y}^1,\dots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k}} \left[\chi_{S_1,\dots,S_k}(\mathbf{Y}^1,\dots,\mathbf{Y}^k)\right] \\
&= \frac{1}{m}\sum_{a\in[m]} \mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\dots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k}\\ \mathbf{a}=a}} \left[\prod_{\substack{i\in[k]\\ S_i=\{a'\}}} \mathbf{Y}^k_{a'}\right] \\
&= \frac{1}{m} \mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\dots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k}\\ \mathbf{a}=a'}} \left[\prod_{\substack{i\in[k]\\ S_i=\{a'\}}} \mathbf{Y}^k_{a'}\right] && \text{(If } \mathbf{a}\neq a' \text{ then the } \mathbf{Y}^i_{a'} \text{ are uniformly random.)} \\
&= \frac{1}{m} \mathop{\mathbf{E}}_{\substack{(\mathbf{Y}^1,\dots,\mathbf{Y}^k)\sim\mathcal{C}_X^{m,m-1,k}\\ \mathbf{a}=a'}} \left[\prod_{\substack{i\in[k]\\ S_i=\{a'\}}} X^k_{a'}\right] \\
&= \frac{1}{m}\chi_{S_1,\dots,S_k}(X^1,\dots,X^k).
\end{aligned}
$$

Therefore $\big(Q_{m,m-1,k} - Q_{m,m,k}\big)\chi_{S_1,\dots,S_k} = \frac{1}{m}\chi_{S_1,\dots,S_k}(X^1,\dots,X^k) - 0 = \frac{1}{m}\chi_{S_1,\dots,S_k}(X^1,\dots,X^k)$. □

Corollary 58. Let $f : \{\pm1\}^{mk} \to \mathbb{R}$. Then

$$\left|\left\langle f, (Q_{m,m-1,k} - Q_{m,m,k})f\right\rangle\right| \le \frac{1}{m}\left\langle f, f\right\rangle.$$

Proof. The $\chi_{S_1,\dots,S_k}$ form an orthonormal eigenbasis (Fact 22) for $Q_{m,m-1,k} - Q_{m,m,k}$, and by Fact 57 each basis element has eigenvalue with absolute value at most $\frac{1}{m}$. □

5.4.3 The Third Hybrid: $Q_{m,m,k}$ to $R_{m,m,k}$

Lemma 59. Assume that $k \le 2^{m/3}$. For any $f : \{\pm1\}^{mk} \to \mathbb{R}$ supported on $B_{\ge2}$ we have

$$\left|\left\langle f, (R_{m,m,k} - Q_{m,m,k})f\right\rangle\right| \le \frac{k^2}{2^m}\|f\|_2^2.$$

Proof. As in the proof of Lemma 56 we find that (with the appropriate definitions of p_0 and p_1 for use of Lemma 39,

$$
\begin{aligned}
&\left|\left\langle f, (R_{m,m,k} - Q_{m,m,k})f\right\rangle\right| \\
\le &\sum_{X\in B_{\ge2}} f(X)^2 \sum_{Y\in B_{\ge2}} \left|p_0(X,Y) - p_1(X,Y)\right| && \text{(Lemma 39 + self-adjointness (Fact 18, Fact 55))} \\
\le &\frac{1}{2^{mk}}\cdot\frac{k^2}{2^m}\sum_{X\in B_{\ge2}} f(X)^2 && \text{(Equation (2) below)} \\
\le &\frac{k^2}{2^m}\left\langle f, f\right\rangle.
\end{aligned}
$$

It remains to prove Equation (2):

$$\sum_{Y\in B_{\ge2}} \left|p_0(X,Y) - p_1(X,Y)\right|$$

$$
\begin{aligned}
&= \sum_{Y \in B_{\geq 2}} \left| \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}_X^{m,m,k}} [\mathbf{Y} = Y] - \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}_X^{m,m,k}} [\mathbf{Y} = Y] \right| \\
&= \frac{1}{2^{mk}} \sum_{Y \in B_{\geq 2}} \left| \prod_{i=0}^{k-1} \frac{1}{2^m - i} - \frac{1}{2^{mk}} \right| && (k \leq 2^{m/3} \leq 2^m - 2 \text{ and } X \in B_{\geq 2}) \\
&\leq \frac{1}{2^{mk}} \sum_{Y \in B_{\geq 2}} \left| \frac{1}{2^{mk}} \left(\prod_{i=0}^{k-1} \frac{2^m}{2^m - i} - 1 \right) \right| \\
&\leq \frac{1}{2^{mk}} \sum_{Y \in B_{\geq 2}} \left| \frac{1}{2^{mk}} \left(1 + \frac{k^2}{2^m} - 1 \right) \right| && (k \leq 2^{m/3}, \text{ Fact 40}) \\
&= \frac{1}{2^{mk}} \sum_{Y \in B_{\geq 2}} \frac{k^2}{2^m 2^{mk}} \\
&\leq \frac{1}{2^{mk}} \cdot 2^{mk} \cdot \frac{k^2}{2^m 2^{mk}} \\
&= \frac{k^2}{2^m}. && (2)
\end{aligned}
$$

Note that our computations relied on the fact that $Y \in B_{\geq 2}$. □

5.4.4 Putting Hybrids Together

Proposition 60. Assume that $k \leq 2^{m/3}$. For any $f : \{\pm 1\}^{mk} \to \mathbb{R}$ supported on $B_{\geq 2}$, we have

$$
\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \leq \left(\frac{1}{m} + \frac{k^2}{2^{m/2}} \right) \langle f, f \rangle
$$

Proof. By the triangle inequality,

$$
\begin{aligned}
&\left| \left\langle f, (R_{m,m-1,k} - R_{m,m,k}) f \right\rangle \right| \\
\leq &\left| \left\langle f, (R_{m,m-1,k} - Q_{m,m-1,k}) f \right\rangle \right| + \left| \left\langle f, (Q_{m,m,k} - Q_{m,m-1,k}) f \right\rangle \right| + \left| \left\langle f, (R_{m,m,k} - Q_{m,m,k}) f \right\rangle \right| \\
\leq &\frac{k^2}{2^{m-1}} \langle f, f \rangle + \frac{1}{m} \langle f, f \rangle + \frac{k^2}{2^m} \langle f, f \rangle && (\text{Lemma 56, Corollary 58, Lemma 59}) \\
\leq &\left(\frac{1}{m} + \frac{k^2}{2^{m-2}} \right) \langle f, f \rangle .
\end{aligned}
$$

Note we can apply Lemma 56 and Lemma 59 because f is supported on $B_{\geq 2}$. □

6 Proof of Theorem 29

In this section let m and $k \leq 2^m - 2$ be fixed positive integers. As in the proof of Theorem 24 (via its restatement Theorem 45) we will break $\mathbb{R}^{\{\pm 1\}^{mk}}$ into orthogonal components and bound the contributions of each of these cross terms to evaluation on the quadratic form given by $R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right)$.

Define the following subsets of $\{\pm 1\}^{mk}$:

- $B_{=0} = \left\{X \in \{\pm 1\}^{mk} : \exists i \neq j \in [k], X^i = X^j\right\}$.[8]
- $B_{\geq 1}^{\text{coll}} = \left\{X \in \{\pm 1\}^{mk} \setminus B_{=0} : \exists i \neq j \in [k], X^i_{[m-\ell-1,m-1]} = X^j_{[m-\ell-1,m-1]}\right\}$.
- $B_{\geq 1}^{\text{safe}} = \left\{X \in \{\pm 1\}^{mk} : \forall i \neq j \in [k], X^i_{[m-\ell-1,m-1]} \neq X^j_{[m-\ell-1,m-1]}\right\}$.

Note that these sets form a partition of $\{\pm 1\}^{mk}$.

As in the proof of Theorem 24, the contribution from the parts of functions supported on $B_{=0}$ will be bounded by induction. The component $B_{\geq 1}^{\text{safe}}$ will play the role that $B_{\geq 2}$ did: the part of the domain on which the noise model induced by random permutations (sampling without replacement) to the bits in $[m-\ell-1, m-1]$ is close to the noise model induced by completely random replacement of bits (sampling with replacement). $B_{\geq 1}^{\text{coll}}$ is the component on which these two noise models are not similar, but as in the case of $B_{=1}$, this set will already show good expansion.

Theorem 61 (Theorem 29 restated)**.** Fix any $m \geq 100$ and set $100 \leq \ell \leq m$. Suppose $k \leq 2^{\ell/10}$. Then we have for any $f : \{\pm 1\}^{mk} \to \mathbb{R}$ that

$$\left|\left\langle f, R_{m,[m-\ell-1,m],k}\left(R_{m,[m-1],k} - R_{m,[m],k}\right) f\right\rangle\right| \leq \frac{k^3}{2^{\ell/2-20}} \langle f, f\rangle \leq \frac{1}{\ell} \langle f, f\rangle .$$

Proof. We prove by induction on k. In the base case $k = 1$ and the result holds by the following argument when we write $f = f_2$. Now assume that the result holds for real functions on $\{\pm 1\}^{m(k-1)}$.

Let $f : \{\pm 1\}^{mk} \to \mathbb{R}$. Write $f = f_0 + f_1 + f_2$ where f_0 is supported on $B_{=0}$, f_1 is supported on $B_{\geq 1}^{\text{coll}}$, and f_2 is supported on $B_{\geq 1}^{\text{safe}}$. By Lemma 53 applied with $R_{m,[m-\ell-1,m],k}$, $R_{m,[m-1],k}$ and $R_{m,[m],k}$ and $B_{=0}$ and $\{\pm 1\}^{mk} \setminus B_{=0} = B_{\geq 1}^{\text{coll}} \cup B_{\geq 1}^{\text{safe}}$, the other cross terms vanish, and we have

$$\begin{aligned}
&\left|\left\langle f, R_{m,[m-\ell-1,m],k}\left(R_{m,[m-1],k} - R_{m,[m],k}\right) f\right\rangle\right| \\
=&\left|\langle f, Af\rangle\right| && \text{(Defining } A \text{ for convenience of notation)} \\
\leq&\left|\langle f_0, Af_0\rangle\right| + \left|\langle f_2, Af_2\rangle\right| + \left|\left\langle f_1, A(f_1 + f_2)\right\rangle\right| + \left|\langle f_2, Af_1\rangle\right| \\
=&\left|\langle f_0, Af_0\rangle\right| + \left|\langle f_2, Af_2\rangle\right| + \left|\left\langle f_1, A(f_1 + f_2)\right\rangle\right| + \left|\left\langle f_1, A^* f_2\right\rangle\right| \\
\leq&\frac{(k-1)^3}{2^{\ell/2-10}} \langle f_0, f_0\rangle + \frac{k^2}{2^{\ell/2-20}} \langle f_2, f_2\rangle + \frac{k^2}{2^{\ell-10}} \|f_1 + f_2\|_2 \|f_2\|_2 + \frac{k^2}{2^{\ell-10}} \|f_1\|_2 \|f_2\|_2 \\
& && \text{(Lemma 43 + Induction, Corollary 68, Corollary 64, Corollary 65, in that order)} \\
\leq&\frac{k^3}{2^{\ell/2-20}} \langle f, f\rangle.
\end{aligned}$$

In the last step we used that $k \leq 2^{\ell/10}$. □

Similar to Section 5.4 we will again argue that because ℓ is large, applying a random permutation to the same indices of all elements of a tuple of binary strings is similar to replacing those indices with random binary strings in all elements of a tuple. To model this situation, we similarly use the matrix $Q_{m,S,k}$ to denote the random walk matrix induced by the following distribution $\mathcal{C}_X^{m,S,k}$ for each $X \in \{\pm 1\}^{mk}$. To draw $\mathbf{Y} \sim \mathcal{C}_X^{m,S,k}$, set $\mathbf{Y}^i_{[m]\setminus S} = X^i_{[m]\setminus S}$ and set $\mathbf{Y}^i_S$ uniformly at random for all $i \in [k]$.

[8] $B_{=0}$ was already defined in Section 5 but we define it again here for convenience.

6.1 f Supported on $B^{\text{coll}}_{\geq 1}$ and Cross Terms

In this section we leverage that $B^{\text{coll}}_{\geq 1}$ is a set in $\{\pm 1\}^{mk}$ with very good expansion. In some sense, $B^{\text{coll}}_{\geq 1}$ plays the role that $B_{=1}$ played in Section 5. The following lemma formalizes the expansion property that we need.

Lemma 62. For any $X \in B^{\text{coll}}_{\geq 1} \cup B^{\text{safe}}_{\geq 1}$ we have that for any $I, J \subseteq [m]$ such that $I \cup J = [m]$ and $I, J \supseteq [m-\ell-1, m-1]$,

$$\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z} \in B^{\text{coll}}_{\geq 1}\right] \leq \frac{k^2}{2^{\ell-2}}.$$

In fact, this holds with $\mathcal{C}$ replacing $\mathcal{D}$ anywhere in this inequality.

Proof. We prove this for the case of distributions $\mathcal{D}$ in both drawing $\mathbf{Y}$ and $\mathbf{Z}$, but the proof translates easily to the case for $\mathcal{C}$. Fix any $X \in B^{\text{coll}}_{\geq 1} \cup B^{\text{safe}}_{\geq 1}$. Define the following subset of pairs of indices in $\binom{[k]}{2}$:

$$P_1 = \left\{\{i,j\} \in \binom{[k]}{2} : X^i_I = X^j_I\right\}.$$

Assume that $\{i,j\} \in P_1$. Then $\mathbf{Y}^i_I = \mathbf{Y}^j_I$ with probability 1. But now note that $X^i_J \neq X^j_J$ since otherwise $X^i = X^j$, which contradicts that $X \notin B_{=0}$. Therefore,

$$\begin{aligned}
&\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
=&\mathbf{Pr}_{\mathbf{Z}\sim\mathcal{D}_X^{m,J,k}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
=&\frac{1}{2^\ell - 1}.
\end{aligned} \tag{3}$$

Now assume that $\{i,j\} \notin P_1$. Let E be the event that $\mathbf{Y}^i_{[m-\ell-1,m-1]} = \mathbf{Y}^j_{[m-\ell-1,m-1]}$. Then

$$\begin{aligned}
&\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
=&\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}}[E]\,\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}|E\right]\\
&+\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}}\left[\overline{E}\right]\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}|\overline{E}\right]\\
\leq&\frac{1}{2^\ell} + \mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_{\mathbf{Y}}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]} = \mathbf{Z}^j_{[m-\ell-1,m-1]}|\overline{E}\right]\\
=&\frac{1}{2^\ell} + \frac{1}{2^\ell - 1}\\
\leq&\frac{1}{2^{\ell-2}}.
\end{aligned} \tag{4}$$

To complete the proof, we use a union bound and find

$$\begin{aligned}
&\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_\mathbf{Y}^{m,J,k}}}\left[\mathbf{Z}\in B_{\geq 1}^{\mathrm{coll}}\right]\\
\leq&\sum_{\{i,j\}\in\binom{[k]}{2}}\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_\mathbf{Y}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]}=\mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
\leq&\sum_{\{i,j\}\in P_1}\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_\mathbf{Y}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]}=\mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
&+\sum_{\{i,j\}\notin P_1}\mathbf{Pr}_{\substack{\mathbf{Y}\sim\mathcal{D}_X^{m,I,k}\\ \mathbf{Z}\sim\mathcal{D}_\mathbf{Y}^{m,J,k}}}\left[\mathbf{Z}^i_{[m-\ell-1,m-1]}=\mathbf{Z}^j_{[m-\ell-1,m-1]}\right]\\
\leq&\sum_{\{i,j\}\in P_1}\frac{1}{2^\ell}+\sum_{\{i,j\}\notin P_1}\frac{1}{2^{\ell-1}}\\
\leq&\sum_{\{i,j\}\in\binom{[k]}{2}}\frac{1}{2^{\ell-2}}\\
\leq&\frac{k^2}{2^{\ell-2}}.
\end{aligned}$$

□

Lemma 63. Let $k\geq 2$ and $f:\{\pm1\}^{mk}$ be supported on $B_{\geq 1}^{\mathrm{coll}}$ and $g:\{\pm1\}^{mk}$ be supported on $B_{\geq 1}^{\mathrm{safe}}\cup B_{\geq 1}^{\mathrm{coll}}$. Let I and J be such that $I\cup J=[m]$. Then

$$\left|\left\langle f, R_{m,J,k}R_{m,I,k}g\right\rangle\right|\leq\frac{k^2}{2^{\ell-2}}\|f\|_2\|g\|_2\,.$$

Moreover, this holds with Q in place of R anywhere.

Proof. The inequality directly follows from Lemma 41 and Lemma 62. We can apply Lemma 41 because the uniform distribution on $\{\pm1\}^{mk}$ is indeed a stationary distribution under $R_{m,S,k}$ for any S (Fact 19). The Q case follows from the fact that Lemma 62 applies for the distributions $\mathcal{C}$ too. □

Corollary 64. Let $k\geq 2$ and $f:\{\pm1\}^{mk}$ be supported on $B_{\geq 1}^{\mathrm{coll}}$ and $g:\{\pm1\}^{mk}$ be supported on $B_{\geq 1}^{\mathrm{safe}}\cup B_{\geq 1}^{\mathrm{coll}}$. Then

$$\left|\left\langle f, R_{m,[m-\ell-1,m],k}\Big(R_{m,[m-1],k}-R_{m,[m],k}\Big)g\right\rangle\right|\leq\frac{k^2}{2^{\ell-3}}\|f\|_2\|g\|_2\,.$$

Proof. We compute

$$\begin{aligned}
&\left|\left\langle f, R_{m,[m-\ell-1,m],k}\Big(R_{m,[m-1],k}-R_{m,[m],k}\Big)g\right\rangle\right|\\
\leq&\left|\left\langle f, R_{m,[m-\ell-1,m],k}R_{m,[m-1],k}g\right\rangle\right|+\left|\left\langle f, R_{m,[m-\ell-1,m],k}R_{m,[m],k}g\right\rangle\right|\\
=&\left|\left\langle f, R_{m,[m-\ell-1,m],k}R_{m,[m-1],k}g\right\rangle\right|+\left|\left\langle f, R_{m,[m],k}g\right\rangle\right|\\
\leq&\frac{k^2}{2^{\ell-2}}\|f\|_2\|g\|_2+\frac{k^2}{2^\ell}\|f\|_2\|g\|_2 \qquad \text{(Lemma 63)}
\end{aligned}$$

$$\leq \frac{k^2}{2^{\ell-3}} \|f\|_2 \|g\|_2 \,. \qquad \square$$

Corollary 65. Let $k \geq 2$ and $f : \{\pm 1\}^{mk}$ be supported on $B^{\text{coll}}_{\geq 1}$ and $g : \{\pm 1\}^{mk}$ be supported on $B^{\text{safe}}_{\geq 1} \cup B^{\text{coll}}_{\geq 1}$. Then

$$\left| \left\langle f, \left(R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) \right)^* g \right\rangle \right| \leq \frac{k^2}{2^{\ell-3}} \|f\|_2 \|g\|_2 \,.$$

Proof. We compute

$$\begin{aligned}
&\left| \left\langle f, \left(R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) \right)^* g \right\rangle \right| \\
=&\left| \left\langle f, \left(R_{m,[m-1],k} R_{m,[m-\ell-1,m],k} - R_{m,[m],k} \right) g \right\rangle \right| && \text{(self-adjointness (Fact 18))} \\
\leq&\left| \left\langle f, R_{m,[m-\ell-1,m],k} R_{m,[m-1],k} g \right\rangle \right| + \left| \left\langle f, R_{m,[m-\ell-1,m],k} R_{m,[m],k} g \right\rangle \right| \\
=&\left| \left\langle f, R_{m,[m-\ell-1,m],k} R_{m,[m-1],k} g \right\rangle \right| + \left| \left\langle f, R_{m,[m],k} g \right\rangle \right| \\
\leq&\frac{k^2}{2^{\ell-2}} \|f\|_2 \|g\|_2 + \frac{k^2}{2^{\ell}} \|f\|_2 \|g\|_2 && \text{(Lemma 63)} \\
\leq&\frac{k^2}{2^{\ell-3}} \|f\|_2 \|g\|_2 \,.
\end{aligned}$$

Note this is essentially the same proof as Corollary 64. $\square$

6.2 A Hybrid Argument for f Supported on $B^{\text{safe}}_{\geq 1}$

The role that $B^{\text{safe}}_{\geq 1}$ plays in this section is similar to the role played by $B_{\geq 2}$ in Section 5. It is the region of $\{\pm 1\}^{mk}$ that is "well-behaved" in the sense that the nicer noise model given by the Q operators is similar to the noise model given by the R operators on this region of $\{\pm 1\}^{mk}$.

Lemma 66. Assume that $k \leq 2^{\ell/10}$ and $f, g : \{\pm 1\}^{mk}$ be supported on $B^{\text{safe}}_{\geq 1}$. Then for any $S \supseteq [m-\ell-1, m-1]$, we have

$$\left| \left\langle f, (Q_{m,S,k} - R_{m,S,k}) g \right\rangle \right| \leq \frac{k^2}{2^{\ell-1}} \|f\|_2 \|g\|_2 \,.$$

Proof. By assumption f and g are supported on $B^{\text{safe}}_{\geq 1}$. Therefore, by Lemma 39 and self-adjointness of $Q_{m,S,k}$ and $R_{m,S,k}$ (Fact 18, Fact 55) we have

$$\begin{aligned}
&\left| \left\langle f, (R_{m,s,k} - Q_{m,S,k}) g \right\rangle \right| \\
\leq&\sqrt{\sum_{X \in B^{\text{safe}}_{\geq 1}} f(X)^2 \sum_{Y \in B^{\text{safe}}_{\geq 1}} \left| p_0(X,Y) - p_1(X,Y) \right|} \sqrt{\sum_{X \in B^{\text{safe}}_{\geq 1}} g(X)^2 \sum_{Y \in B^{\text{safe}}_{\geq 1}} \left| p_0(X,Y) - p_1(X,Y) \right|} \\
\leq&\frac{1}{2^{mk}} \cdot \frac{k^2}{2^{\ell-1}} \sqrt{\sum_{X \in B^{\text{safe}}_{\geq 1}} f(X)^2} \sqrt{\sum_{X \in B^{\text{safe}}_{\geq 1}} g(X)^2} && \text{(Equation (5) below)}
\end{aligned}$$

$$= \frac{k^2}{2^{\ell-1}} \|f\|_2 \|g\|_2 \,.$$

Here p_0 and p_1 are as used below. Now it suffices to establish Equation (5). Assume $X \in B^{\text{safe}}_{\geq 1}$. Then because $S \supseteq [m-\ell-1, m-1]$, we know that for all $i \neq j$ we have $X^i_S \neq X^j_S$.

$$\begin{aligned}
&\sum_{Y \in B^{\text{safe}}_{\geq 1}} \left| p_0(X,Y) - p_1(X,Y) \right| \\
=& \sum_{\substack{Y \in B^{\text{safe}}_{\geq 1} \\ \forall i \in [k], a \in [m] \setminus S, Y^i_a = X^i_a}} \left| \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{D}^{m,S,k}_X} [\mathbf{Y} = Y] - \frac{1}{2^{mk}} \mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}^{m,S,k}_X} [\mathbf{Y} = Y] \right| \\
=& \frac{1}{2^{mk}} \sum_{\substack{Y \in B^{\text{safe}}_{\geq 1} \\ \forall i \in [k], a \in [m] \setminus S, Y^i_a = X^i_a}} \left| \prod_{j=0}^{k-1} \frac{1}{2^{|S|} - j} - \frac{1}{2^{|S|k}} \right| && (k \leq 2^{m/3} \leq 2^m - 2 \text{ and } X, Y \in B^{\text{safe}}_{\geq 1}) \\
\leq& \frac{1}{2^{mk}} \sum_{\substack{Y \in B^{\text{safe}}_{\geq 1} \\ \forall i \in [k], a \in [m] \setminus S, Y^i_a = X^i_a}} \left| \frac{1}{2^{|S|k}} \left(\prod_{j=0}^{k-1} \frac{2^{|S|}}{2^{|S|} - j} - 1 \right) \right| \\
\leq& \frac{1}{2^{mk}} \sum_{\substack{Y \in B^{\text{safe}}_{\geq 1} \\ \forall i \in [k], a \in [m] \setminus S, Y^i_a = X^i_a}} \left| \frac{1}{2^{|S|k}} \left(1 + \frac{k^2}{2^{|S|}} - 1 \right) \right| && (k^2 \leq 2^\ell \leq 2^{|S|}, \text{ Fact 40}) \\
=& \frac{1}{2^{mk}} \sum_{\substack{Y \in B^{\text{safe}}_{\geq 1} \\ \forall i \in [k], a \in [m] \setminus S, Y^i_a = X^i_a}} \frac{k^2}{2^{|S|} 2^{|S|k}} \\
=& \frac{1}{2^{mk}} \cdot 2^{|S|k} \cdot \frac{k^2}{2^{|S|} 2^{|S|k}} \\
=& \frac{k^2}{2^{|S|} 2^{mk}} \\
\leq& \frac{k^2}{2^{mk} 2^{\ell-1}}. && (5)
\end{aligned}$$

The last inequality follows because $|S| \geq \left|[m-\ell-1, m-1]\right| = \ell - 1$. Having established Equation (5), we have completed the proof. □

At this point it may seem like we are essentially finished with the proof, since we should be able to just replace all the R operators in the expression $\left\langle f, R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle$ with the corresponding Q operators and finish the proof. However, we don't quite show an upper bound on the *operator norm* of $R - Q$. Rather, we simply show that they are close on the well-behaved region. A priori, this gives us no information about how products of these operators may behave, since the first term in the product may "rotate" vectors into the badly-behaved region. So a straightforward application of the triangle inequality fails.

However, we observe that we have already shown in Section 6.1 that random walks starting outside the badly-behaved region rarely transition into it, so that we almost can pretend as if all

of these operators are operators on $\mathbb{R}^{B^{\text{safe}}_{\geq 1}}$. We formalize this by inserting projections to the space of functions supported on $B^{\text{safe}}_{\geq 1}$, and showing that this move does very little quantitatively.

Lemma 67. We have for any $f : \{\pm 1\}^{mk} \to \mathbb{R}$ supported on $B^{\text{safe}}_{\geq 1}$ that

$$\left|\left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k}\right)\left(R_{m,[m-1],k} - R_{m,[m],k}\right) f\right\rangle\right| \leq \frac{k^2}{2^{\ell/2-20}}\langle f, f\rangle.$$

Moreover, we have

$$\left|\left\langle f, Q_{m,[m-\ell-1,m],k}\left(R_{m,[m-1],k} - Q_{m,[m-1],k} - R_{m,[m],k} + Q_{m,[m],k}\right) f\right\rangle\right| \leq \frac{k^2}{2^{\ell/2-20}}\langle f, f\rangle.$$

Proof. Let Π_{safe} be the projection to $\{h : \{\pm 1\}^{mk} \to \mathbb{R} : h \text{ supported on } B^{\text{safe}}_{\geq 1}\}$. We directly compute

$$\begin{aligned}
&\left|\left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k}\right) R_{m,[m-1],k} f\right\rangle\right| \\
\leq &\left|\left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k}\right)(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle\right| \\
&+ \left|\left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k}\right)\Pi_{\text{safe}} R_{m,[m-1],k} f\right\rangle\right| \\
\leq &\left|\left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k}\right)(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle\right| + \frac{k^2}{2^{\ell-1}}\|f\|_2 \left\|\Pi_{\text{safe}} R_{m,[m-1],k} f\right\|_2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Lemma 66)} \\
\leq &\left|\left\langle f, R_{m,[m-\ell-1,m],k}(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle\right| + \left|\left\langle f, Q_{m,[m-\ell-1,m],k}(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle\right| + \frac{k^2}{2^{\ell-1}}\|f\|_2^2 \\
\leq &\|f\|_2 \left\|(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\|_2 + \frac{k^2}{2^{\ell-1}}\|f\|_2^2 \\
\leq &\frac{k^2}{2^{\ell/2-10}}\|f\|_2^2 + \frac{k^2}{2^{\ell-1}}\|f\|_2^2. \qquad\qquad \text{(Equation (6) below)}
\end{aligned}$$

Here our application of the Lemma 66 depended on the fact that $\mathrm{Supp}(\Pi_{\text{safe}} R_{m,[m-1],k} f) \subseteq B^{\text{safe}}_{\geq 1}$ and $\mathrm{Supp}((\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f) \subseteq \{\pm 1\}^{mk} \setminus B^{\text{safe}}_{\geq 1}$.

To establish Equation (6) we compute

$$\begin{aligned}
&\left\|(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\|_2^2 \\
= &\left\langle (\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f, (\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle \\
= &\left\langle R_{m,[m-1],k} f, (\mathrm{Id} - \Pi_{\text{safe}})^2 R_{m,[m-1],k} f\right\rangle \\
= &\left\langle f, R_{m,[m-1],k}(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\rangle \qquad \text{(self-adjointness (Fact 18))} \\
\leq &\frac{k^2}{2^{\ell-1}}\|f\|_2 \left\|(\mathrm{Id} - \Pi_{\text{safe}}) R_{m,[m-1],k} f\right\|_2
\end{aligned}$$

$$\leq \frac{k^2}{2^{\ell-1}}\|f\|_2^2 . \tag{6}$$

The first inequality follows from Lemma 41 the fact that $\mathrm{Supp}(f) \subseteq B^{\mathrm{safe}}_{\geq 1}$ and for any $X \in B^{\mathrm{safe}}_{\geq 1}$, we have $\mathbf{Pr}_{\mathbf{Y}\sim\mathcal{D}_X^{m,[m-1],k}}\left[\mathbf{Y} \notin B^{\mathrm{safe}}_{\geq 1}\right] \leq \frac{k^2}{2^{\ell-1}}$ (using the same proof as Lemma 62), and that $\mathrm{Supp}\Big((\mathrm{Id} - \Pi_{\mathrm{safe}})R_{m,[m-1],k}f\Big) \subseteq \{\pm 1\}^{mk} \setminus B^{\mathrm{safe}}_{\geq 1}$.

Bounding the similar quantity but with $R_{m,[m],k}$ instead of $R_{m,[m-1],k}$ is the same and the final result follows from the triangle inequality. The second part of the lemma statement follows from the same argument and an application of the triangle inequality:

$$\begin{aligned}
&\left|\left\langle f, Q_{m,[m-\ell-1,m],k}\Big(R_{m,[m-1],k} - Q_{m,[m-1],k} - R_{m,[m],k} + Q_{m,[m],k}\Big)f\right\rangle\right| \\
=&\left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k} - R_{m,[m],k} + Q_{m,[m],k}\Big)Q_{m,[m-\ell-1,m],k}f\right\rangle\right| && \text{(Fact 18, Fact 55)} \\
\leq&\left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k}\Big)Q_{m,[m-\ell-1,m],k}f\right\rangle\right| + \left|\left\langle f, \Big(R_{m,[m],k} - Q_{m,[m],k}\Big)Q_{m,[m-\ell-1,m],k}f\right\rangle\right|.
\end{aligned} \tag{7}$$

We show how to bound the first term in this sum:

$$\begin{aligned}
&\left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k}\Big)Q_{m,[m-\ell-1,m],k}f\right\rangle\right| \\
\leq&\left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k}\Big)(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-\ell-1,m],k}f\right\rangle\right| \\
&+ \left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k}\Big)\Pi_{\mathrm{safe}}Q_{m,[m-\ell-1,m],k}f\right\rangle\right| \\
\leq&\left|\left\langle f, \Big(R_{m,[m-1],k} - Q_{m,[m-1],k}\Big)(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-\ell-1,m],k}f\right\rangle\right| + \frac{k^2}{2^{\ell-1}}\|f\|_2\left\|\Pi_{\mathrm{safe}}Q_{m,[m-\ell-1,m],k}f\right\|_2 \\
& && \text{(Lemma 66)} \\
\leq&\left|\left\langle f, R_{m,[m-1],k}(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-\ell-1,m],k}f\right\rangle\right| \\
&+ \left|\left\langle f, Q_{m,[m-1],k}(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-\ell-1,m],k}f\right\rangle\right| + \frac{k^2}{2^{\ell-1}}\|f\|_2^2 \\
\leq&\|f\|_2\left\|(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-\ell-1,m],k}\right\|_2 + \frac{k^2}{2^{\ell-1}}\|f\|_2^2 \\
\leq&\frac{k^2}{2^{\ell/2-10}}\|f\|_2^2 + \frac{k^2}{2^{\ell-1}}\|f\|_2^2 . && \text{(Equation (8) below)}
\end{aligned}$$

To establish Equation (8) we compute

$$\begin{aligned}
&\left\|(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-1],k}f\right\|_2^2 \\
=&\left\langle f, Q_{m,[m-1],k}(\mathrm{Id} - \Pi_{\mathrm{safe}})Q_{m,[m-1],k}f\right\rangle && \text{(self-adjointness (Fact 55)}
\end{aligned}$$

$$\leq \frac{k^2}{2^{\ell-1}} \left\| Q_{m,[m-1],k} f \right\|_2 \left\| R_{m,[m-1],k} f \right\|_2$$
$$\leq \frac{k^2}{2^{\ell-1}} \|f\|_2^2 . \tag{8}$$

The first inequality follows from Lemma 41 the fact that $\mathrm{Supp}(f) \subseteq B_{\geq 1}^{\mathrm{safe}}$ and for any $X \in B_{\geq 1}^{\mathrm{safe}}$, we have $\mathbf{Pr}_{\mathbf{Y} \sim \mathcal{C}_X^{m,[m-\ell-1,m],k}} \left[\mathbf{Y} \notin B_{\geq 1}^{\mathrm{safe}} \right] \leq \frac{k^2}{2^{\ell-1}}$ (using the same proof as Lemma 62). The second term in the sum Equation (7) follows from the same argument. □

Corollary 68. Let $k \geq 2$ and $f : \{\pm 1\}^{mk} \to \mathbb{R}$ be supported on $B_{\geq 1}^{\mathrm{safe}}$. Then

$$\left| \left\langle f, R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle \right| \leq \frac{k^2}{2^{\ell/2-50}} \langle f, f \rangle .$$

Proof. We directly compute

$$\begin{aligned}
&\left| \left\langle f, R_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle \right| \\
\leq &\left| \left\langle f, Q_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle \right| + \\
&\left| \left\langle f, \left(R_{m,[m-\ell-1,m],k} - Q_{m,[m-\ell-1,m],k} \right) \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle \right| \\
\leq &\left| \left\langle f, Q_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - R_{m,[m],k} \right) f \right\rangle \right| + \frac{k^2}{2^{\ell/2-20}} \langle f, f \rangle \qquad \text{(Lemma 67, first part)} \\
\leq &\frac{k^2}{2^{\ell/2-20}} \langle f, f \rangle + \left| \left\langle f, Q_{m,[m-\ell-1,m],k} \left(Q_{m,[m-1],k} - Q_{m,[m],k} \right) f \right\rangle \right| \\
&+ \left| \left\langle f, Q_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - Q_{m,[m-1],k} - R_{m,[m],k} + Q_{m,[m],k} \right) f \right\rangle \right| \\
= &\frac{k^2}{2^{\ell/2-20}} \langle f, f \rangle + \left| \left\langle f, Q_{m,[m-\ell-1,m],k} \left(R_{m,[m-1],k} - Q_{m,[m-1],k} - R_{m,[m],k} + Q_{m,[m],k} \right) f \right\rangle \right| \\
= &\frac{k^2}{2^{\ell/2-20}} \langle f, f \rangle + \frac{k^2}{2^{\ell/2-20}} \langle f, f \rangle . \qquad \text{(Lemma 67, second part)}
\end{aligned}$$

For the second-to-last equality we used that $Q_{m,[m-\ell-1,m],k} \left(Q_{m,[m-1],k} - Q_{m,[m],k} \right) = 0$ because $Q_{m,S,k} Q_{m,T,k} = Q_{m,S \cup T,k}$ for any $S, T \subseteq [m]$. □

7 Computational Hardness of Inverting Block Ciphers

In this section we show a connection between the existence of one-way functions (OWFs) and the hardness of inverting classes of block ciphers. We view the problem of inverting block ciphers as equivalent to the problem of learning a reversible circuit that computes a permutation using query access to the permutation, which frames the problem as a minimum (reversible) circuit size problem (MRCSP).

The idea is essentially the same as that of Kabanets and Cai [KC00], who showed that the analogous result about the minimum circuit size problem (MCSP). In that paper, it is observed that pseudorandom function ensembles computed by small circuits are exactly the type of input distributions to the MCSP problem that "fool" algorithms for MCSP running in polynomial time. So the existence of OWFs (which by [GGM86, HILL99] implies the existence of such pseudorandom function ensembles) precludes the existence of polynomial time algorithms for MCSP.

Thus, to prove the result for MRCSP, we would like to construct *pseudorandom permutations* implementable by small reversible circuits. It is known how to construct pseudorandom permutations implementable by small circuits by the Feistel cipher construction Luby and Rackoff [LR88], but these circuits are not necessarily reversible. We build on this construction by simulating these circuits using reversible circuits, as one often does in the context of quantum computing, for example. The problem with this is that such simulation requires ancilla bits (let's say n^d total bits are required), which may reveal information about the computation, so naively simulating the Feistel cipher does not yield a pseudorandom permutation. At this point we are viewing our ciphers as acting on $\{0,1\}^{n^d}$. Our solution is to simply hide away the ancilla bits with another round of the Feistel cipher, this time applied to the ancilla bits. We iterate this a polynomial number of times, and find that this is sufficient for pseudorandomness.

We note here that the construction of Luby-Rackoff can be made reversible when the underlying pseudorandom function is computable in NL with no extra ancillas by a construction of Cleve [Cle90].

7.1 Pseudorandom Permutations Implementable Using Reversible Circuits

Throughout this section we use $\{0,1\}$-notation and interpret $\{0,1\}$ as the field $\mathbb{F}_2$ with two elements.

Definition 69. Fix $\kappa : \mathbb{N} \to \mathbb{N}$ and $\varepsilon : \mathbb{N} \to [0,1]$. A family (indexed by n) of maps map $f^n_{(\cdot)} : \{0,1\}^{\kappa(n)} \to \left\{f : \{0,1\}^n \to \{0,1\}^n\right\}$ is an *ε-pseudorandom function ensemble against a class of adversaries $\mathcal{C}$*[9] *with key length κ* if for all large enough n and $\mathcal{A} \in \mathcal{C}_n$, we have

$$\left|\mathbf{Pr}_{\mathsf{key}\in\{0,1\}^{\kappa(n)}}\left[\mathcal{A}^{f^n_{\mathsf{key}}} = 1\right] - \mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_{\{0,1\}^n}}\left[\mathcal{A}^{\mathbf{g}} = 1\right]\right| \leq \varepsilon(n).$$

Here (and throughout this section) $\mathcal{A}^f$ denotes the circuit $\mathcal{A}$ with *nonadaptive* oracle gates computing f (all oracles are at the bottom of the circuit).

Given the existence of OWFs, we can construct efficiently-implementable ensembles of pseudorandom functions by constructing pseudorandom generators.

Theorem 70 ([HILL99] + [GGM86])**.** If OWFs exist, then there exists constants $c, d > 0$ such that for any large enough n there exists a $\delta > 0$ and 2^{-n^δ}-pseudorandom function ensemble against P/poly implementable by a family of Boolean circuits with size at most n^d.[10]

We first show how to get pseudorandom permutation ensembles implemented by short reversible circuits from pseudorandom function ensembles implemented by small Boolean circuits. Throughout this section reversible circuits are those with arbitrary 3-bit gates, and Boolean circuits have 2-bit gates computing AND and OR and 1-bit gates computing NOT. The specific gate set is not important, since we can simulate universal gate sets with other universal gate sets using constant overhead.

[9] Here $\mathcal{C}$ actually defines a set $\mathcal{C}_n$ of adversaries for each n. We will commonly take $\mathcal{C}$ to be the class of circuits with size restricted by some function $r : \mathbb{N} \to \mathbb{N}$, so that $\mathcal{C}_n$ consists of the set of all circuits of size at most $r(n)$.

[10] The family $f_{(\cdot)}$ is implementable by circuits of size at most s if every f^n_{key} is computed using Boolean circuits of size at most $s(n)$. By Boolean circuits we mean circuits with fan-in at most 2 and $(\mathsf{AND}, \mathsf{OR}, \mathsf{NOT})$ gates.

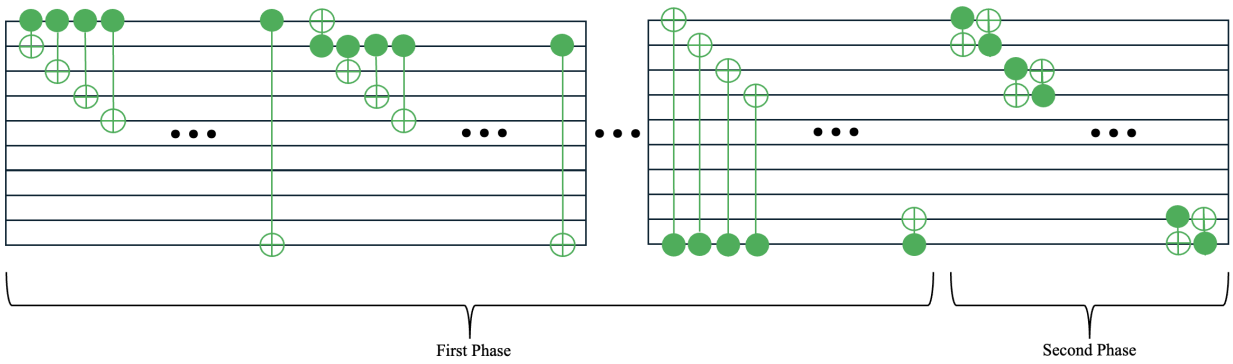

Figure 2: An illustration of $p^N_{\left(f_{(1,2)},\ldots,f^2_{s(n)}\right)}$. Here each wire carries a single n-bit string, and there are $s(n)$ wires. The first layer applies the operation $\mathsf{Ctr}_{f_{(1,2)}}$, the second layer applies $\mathsf{Ctr}_{f_{(1,3)}}$, etc. We apply the operations in two phases. The first phase consists of the operations $\mathsf{Ctr}_{f_{(i,j)}}$ for $i \neq j \in [s(n)]$. The second phase applies the operations $\mathsf{Ctr}^{(i,i+1)}_{f^1_i}$ and $\mathsf{Ctr}^{(i+1,i)}_{f^2_i}$.

Implementing Luby-Rackoff Reversibly. Fix any positive integer N, and assume that $N = ns$ for some s, so that we regard strings in $\{0,1\}^N$ as s-tuples of n-bit strings. That is, we write $x \in \{0,1\}^N$ as $x = (x_1, \ldots, x_s)$ where each $x_i \in \{0,1\}^n$.

Definition 71. For each $a, b \in [s]$ and a function $f : \{0,1\}^n \to \{0,1\}^n$ define the following permutation $\mathsf{Ctr}^{(i,j)}_f$ on $\{0,1\}^N$:

$$\mathsf{Ctr}^{(a,b)}_f(x_1, \ldots, x_s) = \big(x_1, \ldots, x_{a-1}, x_a, x_{a+1}, \ldots, x_{b-1}, x_j + f(x_a), x_{b+1}, \ldots, x_s\big).$$

The notation Ctr refers to "controlled", so $\mathsf{Ctr}^{(a,b)}_f$ is a "controlled f" with control string a and target string b, since such functions can be thought of as generalization of CNOTs.

We now describe our pseudorandom ensemble of functions $\left\{p^N_{(\cdot)}\right\}_{N\in\mathbb{N}}$ implemented using reversible circuits. Let $\left\{f^n_{(\cdot)}\right\}_{n\in\mathbb{N}}$ be a ε-pseudorandom function ensemble with key length κ implementable using Boolean circuits of size at most $s(n)$. We will prove that the ensemble resulting from our construction, when instantiated with f[11], is pseudorandom against P/poly.

To begin, define for functions $f_{(1,2)}, f_{(1,3)}, \ldots, f_{(s(n),s(n)-1)}, f^1_1, f^2_1, \ldots, f^1_{s(n)}, f^2_{s(n)} : \{0,1\}^n \to \{0,1\}^n$ the permutation

$$p^N_{\left(f_{(1,2)},\ldots,f^2_{s(n)}\right)} = \prod_{i\in[s(n)],\, i \text{ odd}} \left(\mathsf{Ctr}^{(a+1,a)}_{f_a}\mathsf{Ctr}^{(a,a+1)}_{f_a}\right) \cdot \prod_{a\in[s(n)]} \left(\prod_{b\in[s(n)]\setminus\{a\}} \mathsf{Ctr}^{(a,b)}_{f_{(a,b)}}\right).$$

Here the order of factors in the product is given by $\prod_{a\in[s]} x_a = x_s x_{s-1} \ldots x_1$. Here we can think of the permutation as acting in two phases: the first phase consists of $s(n)(s(n)-1)$ many controlled f permutation from each $a \in [s(n)]$ to each $b \in [s(n)] \setminus \{a\}$. Then the second phase consists of a sequence of permutations from each a to one of its adjacent bits. We will show that applying the first phase to any $x \in \{0,1\}^{Nq(n)}$ will make all of its matrix entries pairwise distinct with high

[11] f here really denotes the ensemble $\left\{f^n_{(\cdot)}\right\}_{n\in\mathbb{N}}$. We employ this abuse of notation throughout this section.

probability. Then the second phase will leverage this distinctness to make all of the entries close to i.i.d. uniform on $\{0,1\}^n$.

To construct our ensemble, we construct p^N_{key} for each $\mathsf{key} \in \{0,1\}^{\kappa'(N)} = \{0,1\}^{s(n)^2\kappa(n)}$ (recalling that $N = s(n)n$). Given key, interpret it as

$$\mathsf{key} = \left(\left(\mathsf{key}_{(a,b)} \right)_{a \neq b \in [s(n)]}, \left(\mathsf{key}^1_a, \mathsf{key}^2_a \right)_{a \in [s(n)]} \right),$$

where each $\mathsf{key}_{i,1}, \mathsf{key}_{i,2}, \mathsf{key}_{(i,j)} \in \{0,1\}^{\kappa(n)}$. Then define

$$p^N_{\mathsf{key}} = p^N_{\left(f_{\mathsf{key}_{(1,2)}}, \ldots, f_{\mathsf{key}_{s(n),2}} \right)}.$$

Here the order of factors in the product is given by $\prod_{i \in [t]} x_i = x_t x_{t-1} \ldots x_1$.

Proposition 72. Let $c > 0$ be fixed. Assume $\left\{ f^n_{(\cdot)} \right\}_{n \in \mathbb{N}}$ is ε-pseudorandom against circuits of size at most $100n^{c+1}s(n)^3$. Fix $n \in \mathbb{N}$ large enough and let $s(n)$ be as above. Let $N = s(n)n$. Then for any adversary circuit $\mathcal{A}$ of size at most n^c that makes $q(n)$ black box nonadaptive queries, we have

$$\left| \mathbf{Pr}_{\mathbf{g} \in \mathfrak{S}_{\{0,1\}^N}} \left[\mathcal{A}^{\mathbf{g}} = 1 \right] - \mathbf{Pr}_{\mathsf{key} \in \{0,1\}^{s(n)^2\kappa(n)}} \left[\mathcal{A}^{p^N_{\mathsf{key}}} \right] = 1 \right| \leq s(n)^2 \varepsilon(n) + \frac{q(n)^2 s(n)^2}{2^{n-10}}.$$

Proof. We proceed by a hybrid argument. The first step is to replace the pseudorandom function ensemble given by the f's with completely random permutations. By the triangle inequality,

$$\begin{aligned}
&\left| \mathbf{Pr}_{\mathbf{g}:\{0,1\}^n \to \{0,1\}^n} \left[\mathcal{A}^{\mathbf{g}} = 1 \right] - \mathbf{Pr}_{\mathsf{key} \in \{0,1\}^{s(n)^2\kappa(n)}} \left[\mathcal{A}^{p^N_{\mathsf{key}}} \right] = 1 \right| \\
\leq &\left| \mathbf{Pr}_{\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2}:\{0,1\}^n \to \{0,1\}^n} \left[\mathcal{A}^{p^N_{(\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2})}} = 1 \right] - \mathbf{Pr}_{\mathsf{key}} \left[\mathcal{A}^{p^N_{\mathsf{key}}} \right] = 1 \right| \\
&+ \left| \mathbf{Pr}_{\mathbf{g} \in \mathfrak{S}_{\{0,1\}^N}} \left[\mathcal{A}^{\mathbf{g}} = 1 \right] - \mathbf{Pr}_{\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2}:\{0,1\}^n \to \{0,1\}^n} \left[\mathcal{A}^{p^N_{(\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2})}} = 1 \right] \right| \\
\leq &s(n)^2 \varepsilon(n) + \left| \mathbf{Pr}_{\mathbf{g} \in \mathfrak{S}_{\{0,1\}^N}} \left[\mathcal{A}^{\mathbf{g}} = 1 \right] - \mathbf{Pr}_{\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2}:\{0,1\}^n \to \{0,1\}^n} \left[\mathcal{A}^{p^N_{(\mathbf{g}_{(1,2)}, \ldots, \mathbf{g}_{s(n),2})}} = 1 \right] \right|.
\end{aligned}$$

This last inequality follows from a standard hybrid argument in Claim 74 and the ε-pseudorandomness of f against adversaries with size at most $100n^{c+1}s(n)^3$.

We now work to bound the difference between these two probabilities. We do so by proving that the total variation distance between the distributions on $\{0,1\}^{Nq(n)}$ induced by applying a fully random permutation of $\{0,1\}^N$ to all the elements in a fixed $q(n)$-tuple $x \in (\{0,1\}^N)^{q(n)}$ and by applying the product of the $s(n)^2 + 1$ permutations as in the second experiment to the same tuple is small. Without loss of generality we may assume that $x_i \neq x_j$ for all $i \neq j$.

We view elements $x \in \{0,1\}^{Nq(n)} \cong \left(\left(\{0,1\}^n \right)^{s(n)} \right)^{q(n)}$ as $q(n)$-by-$s(n)$ matrices of elements of $\{0,1\}^n$:

$$x = \begin{bmatrix} x^1_1 & \ldots & x^1_{s(n)} \\ & \vdots & \\ x^{q(n)}_1 & \ldots & x^{q(n)}_{s(n)} \end{bmatrix}.$$

Here each x_a^i is an element of $\{0,1\}^n$.

Claim 73. Fix any $x \in \{0,1\}^{Nq(n)}$ such that $x^i \neq x^j$ if $i \neq j$. Except with probability at most $\frac{q(n)^2 s(n)^2}{2^{n-1}}$, we have that

$$y = \left(\prod_{a \in [s(n)]} \prod_{b \in [s(n)] \setminus \{a\}} \mathsf{Ctr}_{\mathbf{g}_{(a,b)}}^{(a,b)} \right)(x)$$

has $y_a^i \neq y_b^j$ if $(i,a) \neq (j,b)$, where each $\mathbf{g}_{(a,b)}$ is a uniformly random function $\{0,1\}^n \to \{0,1\}^n$.

Proof. We bound $\mathbf{Pr}\left[y_a^i = y_b^j\right]$ for each $(i,a) \neq (j,b)$. First assume that $a \neq b$. Then (addition and subtraction are the same because operations are over $\mathbb{F}_2$) we have

$$\begin{aligned}
&y_a^i - y_b^j \\
=&x_a^i + x_b^j + \sum_{a' \neq a} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \sum_{b' \neq b} \mathbf{g}_{(b',b)}\left(F_{b'}\left(x_1^j, \ldots, x_{b'-1}^j\right) + x_{b'}^j\right) \\
=&x_a^i + x_b^j + \mathbf{g}_{(b,a)}\left(F_b\left(x_1^i, \ldots, x_{b-1}^i\right) + x_b^i\right) + \mathbf{g}_{(a,b)}\left(F_a\left(x_1^j, \ldots, x_{a-1}^j\right) + x_a^j\right) \\
&+ x_a^i + x_b^j + \sum_{a' \neq a,b} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \sum_{b' \neq a,b} \mathbf{g}_{(b',b)}\left(F_{b'}\left(x_1^j, \ldots, x_{b'-1}^j\right) + x_{b'}^j\right).
\end{aligned}$$

The $F_{a'}$ are arbitrary functions that depend on the $\mathbf{g}_{(a'',a')}$ for $a'' \leq a'$. The right-hand side is of this is therefore the sum of a uniformly random element $\{0,1\}^n$, given by $\mathbf{g}_{(b,a)}\left(F_b\left(x_1^i, \ldots, x_{s(n)}^i\right)\right)$, with independent random elements of $\{0,1\}^n$, which is a uniformly random element of $\{0,1\}^n$. As a result we have that $\mathbf{Pr}\left[y_a^i = y_b^j\right] = 2^{-n}$.

Now assume that $a = b$; this means that $i \neq j$. Because $x^i \neq x^j$, it must be that $x_{a'}^i \neq x_{a'}^j$ for some $a^* \in [s(n)]$. Let a^* be the least such index. Then if $a^* \neq a$, we have

$$\begin{aligned}
&y_a^i - y_b^j \\
=&x_a^i + x_a^j + \sum_{a' \neq a} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \sum_{a' \neq a} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^j, \ldots, x_{a'-1}^j\right) + x_{a'}^j\right) \\
=&x_a^i + x_a^j + \mathbf{g}_{(a^*,a)}\left(F_{a^*}\left(x_1^i, \ldots, x_{a^*-1}^i\right) + x_{a^*}^i\right) + \mathbf{g}_{(a^*,a)}\left(F_{a^*}\left(x_1^j, \ldots, x_{a^*-1}^j\right) + x_{a^*}^j\right) \\
&+ x_a^i + x_a^j + \sum_{a' \neq a,a^*} \left(\mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^j, \ldots, x_{a'-1}^j\right) + x_{a'}^j\right)\right).
\end{aligned}$$

Again the $F_{a'}$ are arbitrary functions that depend on the $\mathbf{g}_{(a'',a')}$ for $a'' \leq a'$. Since $x_{a^*}^i \neq x_{a^*}^j$ and $x_{a''}^i \neq x_{a''}^j$ for all $a'' \leq a^*$, the above random variable is the sum of the images under a uniformly random permutation of two distinct elements, which is a uniformly random nonzero element of $\{0,1\}^n$, so we have $\mathbf{Pr}\left[y_a^i = y_b^j\right] \leq 2^{-n+1}$.

If $a^* = a$ then

$$y_a^i - y_b^j$$

$$
\begin{aligned}
&= x_a^i + x_a^j + \sum_{a' \neq a} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \sum_{a' \neq a} \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^j, \ldots, x_{a'-1}^j\right) + x_{a'}^j\right) \\
&= x_a^i + x_a^j + \mathbf{g}_{(a+1,a)}\left(F_{a+1}\left(x_1^i, \ldots, x_a^i\right) + x_{a+1}^i\right) + \mathbf{g}_{(a+1,a)}\left(F_{a+1}\left(x_1^j, \ldots, x_{a^*}^j\right) + x_{a+1}^j\right) \\
&\quad + x_a^i + x_a^j + \sum_{a' \neq a, a+1} \left(\mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^i, \ldots, x_{a'-1}^i\right) + x_{a'}^i\right) + \mathbf{g}_{(a',a)}\left(F_{a'}\left(x_1^j, \ldots, x_{a'-1}^j\right) + x_{a'}^j\right)\right).
\end{aligned}
$$

We know that $F_{a+1}\left(x_1^i, \ldots, x_a^i\right) = \mathbf{g}_{(a,a+1)}\left(F_a\left(x_1^i, \ldots, x_{a-1}^i\right) + x_a^i\right)$ for some function F_a depending on $\mathbf{g}_{(a'',a)}$ for $a'' \leq a$. Because $x_a^i \neq x_a^j$, we know $F_{a+1}\left(x_1^i, \ldots, x_a^i\right)$ and $F_{a+1}\left(x_1^j, \ldots, x_a^j\right)$ are two uniform elements sampled without replacement from $\{0,1\}^n$. With probability at least $1 - \frac{1}{2^n - 1}$ we have

$$
F_{a+1}\left(x_1^i, \ldots, x_a^i\right) + x_{a+1}^i \neq F_{a+1}\left(x_1^j, \ldots, x_a^j\right) + x_{a+1}^j.
$$

If this occurs, then $y_a^i - y_b^j = y_a^i - y_a^j$ is the sum of a uniformly random element of $\{0,1\}^n$ with an arbitrary independent element of $\{0,1\}^n$, which is uniformly random. Thus,

$$
\mathbf{Pr}\left[y_a^i = y_b^j\right] \leq \frac{1}{2^n - 1} + \frac{1}{2^n} \leq \frac{1}{2^{n-1}}.
$$

To complete the proof let the event that $y_a^i \neq y_b^j$ for all $(i,a) \neq (j,b)$ be denoted by $\mathcal{E}$. Then

$$
\mathbf{Pr}\left[\overline{\mathcal{E}}\right] \leq \sum_{(i,a) \neq (j,b)} \mathbf{Pr}\left[y_a^i = y_b^j\right] \leq \frac{q(n)^2 s(n)^2}{2^{n-1}}. \qquad \square
$$

Conditioned on the event $\mathcal{E}$ (so that $y_a^i \neq y_b^j$ for $(i,a) \neq (j,b)$), the distribution

$$
\left(\prod_{a \in [s(n)], a \text{ odd}} \mathsf{Ctr}_{\mathbf{g}_a^1}^{(a,a+1)} \mathsf{Ctr}_{\mathbf{g}_a^2}^{(a+1,a)}\right)(y)
$$

$\frac{q(n)^2 s(n)^2}{2^{n-1}}$-close to uniform on $\{0,1\}^{Nq(n)}$. To see this, we prove that every two columns of the form $\{a, a+1\}$ for a odd ends up close to uniform on $\{0,1\}^{2nq(n)}$. WLOG assume that $a = 1$. Then because $y_2^i \neq y_2^j$ for all $i \neq j$, we have that $\left(\mathbf{g}(y_1^i)\right)_{i \in q(n)}$ is uniformly distributed on $\{0,1\}^{nq(n)}$, independently of the n-bit strings occupying the entries not in the second column of $\mathsf{Ctr}_{\mathbf{g}_1^2}^{(2,1)}(y)$. By a union bound, with probability at least $1 - \frac{q(n)^2}{2^n}$ all of the strings in the second column are distinct, and if this occurs then the first column of $\mathsf{Ctr}_{\mathbf{g}_1^1}^{(1,2)} \mathsf{Ctr}_{\mathbf{g}_1^2}^{(2,1)}(y)$ is uniformly distributed on $\{0,1\}^{nq(n)}$ independently of the strings in the entries not in the first column. Repeating this argument for $s(n)/2$ other pairs of columns and using a union bound gives that the resulting distribution is $\frac{q(n)^2 s(n)}{2^{n-2}}$-close to uniform on $\{0,1\}^{Nq(n)}$.

The uniform distribution on $\{0,1\}^{Nq(n)}$ has distance at most $\frac{q(n)^2}{2^n}$ from the uniform distribution on $B_{\geq 1} = \left\{z \in \{0,1\}^{Nq(n)} : z^i \neq z^j \iff i \neq j\right\}$, so if we let $\mathcal{D}$ denote the distribution induced by applying the local permutations given above to a matrix $x \in \{0,1\}^{Nq(n)}$, we have by the previous paragraph that

$$
\left\|\mathcal{D} - \mathrm{Unif}\left(B_{\geq 1}\right)\right\|_1 \leq \left\|\mathcal{D} - \mathrm{Unif}\left(\{0,1\}^{Nq(n)}\right)\right\|_1 + \left\|\mathrm{Unif}\left(B_{\geq 1}\right) - \mathrm{Unif}\left(\{0,1\}^{Nq(n)}\right)\right\|_1
$$

$$\leq \left(\frac{q(n)^2 s(n)}{2^{n-2}} + \frac{q(n)^2 s(n)^2}{2^{n-1}} \right) + \frac{q(n)^2}{2^n} \leq \frac{q(n)^2 s(n)^2}{2^{n-3}}.$$

Since $\mathrm{Unif}(B_{\geq 1})$ is exactly the distribution induced by applying a uniformly random permutation from $\mathfrak{S}_{\{0,1\}^N}$ to x, this completes our proof. □

Claim 74. Let $c > 0$ be fixed. Assume $\left\{ f^n_{(\cdot)} \right\}_{n \in \mathbb{N}}$ is ε-pseudorandom against circuits of size at most $100n^{c+1}s(n)^3$ and implementable by circuits of size at most $s(n)$. Then for any adversary circuit $\mathcal{A}$ of size at most n^c, we have

$$\left| \mathbf{Pr}_{\mathbf{g}_{(1,2)},\ldots,\mathbf{g}_{s(n),2}:\{0,1\}^n \to \{0,1\}^n} \left[\mathcal{A}^{p^N_{(\mathbf{g}_{(1,2)},\ldots,\mathbf{g}_{s(n),2})}} = 1 \right] - \mathbf{Pr}_{\mathsf{key}} \left[\mathcal{A}^{p^N_{\mathsf{key}}} \right] = 1 \right| \leq s(n)^2 \varepsilon(n).$$

Proof. Let $\mathcal{A}$ be a circuit of size at most n^c. We replace the pseudorandom functions in $p^N_{(\cdot)}$ with completely random functions one-by-one, so that the quantity we intend to bound can be written as $\left| \mathbf{Pr}^{s(n)^2} - \mathbf{Pr}^0 \right|$, where for each $\ell \in [s(n)^2]$ we define

$$\mathbf{Pr}^\ell = \mathbf{Pr}_{\substack{\mathsf{key}_{\ell+1},\ldots,\mathsf{key}_{s(n)^2} \\ \mathbf{g}_1,\ldots,\mathbf{g}_\ell:\{0,1\}^n \to \{0,1\}^n \\ \mathbf{g}_{t+1} = f^n_{\mathsf{key}_t} \forall t \in [\ell+1, s(n)^2]}} \left[\mathcal{A}^{p^N_{(\mathbf{g}_1,\ldots,\mathbf{g}_{s(n)^2})}} = 1 \right].$$

It suffices to show that $\left| \mathbf{Pr}^\ell - \mathbf{Pr}^{\ell-1} \right| \leq \varepsilon(n)$ for all $\ell \in [s(n)^2 - 1]$. Assume otherwise, so that this does not hold for a particular value of ℓ. Then we can build an adversary $\mathcal{A}'$ breaking pseudorandomness of f by simulating oracle to access to p using oracle access to f. To do this, we build a randomized adversary $\mathcal{R}$ that is the same circuit as $\mathcal{A}$, but each oracle gate $\mathcal{A}$ has is replaced with the following random circuit $\mathcal{P}$. The circuit $\mathcal{P}$ randomly draws $\mathsf{key}_1, \ldots, \mathsf{key}_{\ell-1}$ uniformly at random. It then implements the function

$$p^N_{f^n_{\mathsf{key}_1},\ldots,f^n_{\mathsf{key}_{\ell-1}},\mathcal{O},\mathbf{g}_{\ell+1},\ldots,\mathbf{g}_{s(n)^2}},$$

where $\mathcal{O} : \{0,1\}^n \to \{0,1\}^n$ is the oracle gate (which is either a pseudorandom $f^n_{\mathsf{key}_\ell}$ or a completely random function $\mathbf{g}_\ell$) and each $\mathbf{g}_t : \{0,1\}^n \to \{0,1\}^n$ is a completely random function. To implement the completely random functions, $\mathcal{P}$ just draws uniformly random elements each time it is asked for the output of a random function, and remembers which input it has been queried on.

If we run $\mathcal{A}$ with its oracle calls replaced by these circuits $\mathcal{P}$, we get a distinguisher $\mathcal{R}$ with oracle access to a function $f : \{0,1\}^n \to \{0,1\}^n$ that distinguishes the case where f is chosen completely randomly and the case where f is chosen according to $\left\{ f^n_{(\cdot)} \right\}$ with advantage equal to $\left| \mathbf{Pr}^\ell - \mathbf{Pr}^{\ell-1} \right|$.

The circuit $\mathcal{P}$ has size at most $s(n)^2 \cdot s(n) + 100s(n) \cdot n$ (the $100s(n) \cdot n$ comes from the size required for the memorization of queried strings). There are a fixed set of random strings so that $\mathcal{R}$ has the largest distinguishing advantage on that set of strings, so we just let $\mathcal{A}'$ be $\mathcal{R}$ equipped with those random strings, and $\mathcal{A}'$ has size at most $\mathrm{size}(\mathcal{A}) + \mathrm{size}(\mathcal{A}) + \mathrm{size}(\mathcal{P}) \leq n^c + 100n \cdot s(n)^3$. Thus, we have a contradiction, since $\mathcal{A}'$ has distinguishing advantage at least $\left| \mathbf{Pr}^\ell - \mathbf{Pr}^{\ell-1} \right| > \varepsilon(n)$ against f. □

When we instantiate this construction with the pseudorandom f ensembles given by [HILL99] and [GGM86], we get pseudorandom permutations implementable by small reversible circuits.

Corollary 75. If OWFs exist then for some $\delta > 0$ there exists an 2^{-n^δ}-pseudorandom function ensemble $\left\{p^n_{(\cdot)}\right\}$ against P/poly implementable by reversible circuits of size at most n^3.

Proof. If OWFs exist then by Theorem 70, for some constants δ and d and any $c > 0$, there exists a 2^{-n^δ}-pseudorandom function ensemble $\left\{f^n_{(\cdot)}\right\}$ against circuits of size at most n^c implementable using Boolean circuits of size n^d. Proposition 72 shows that the $\left\{p^N_{(\cdot)}\right\}$ built from $\left\{f^n_{(\cdot)}\right\}$ as above is ε-pseudorandom against circuits of size at most $n^{c/(d+1)}/n^{3d+1}100 \leq n^{c/d-4d}$, where (recalling that $N = ns(n) = n^{d+1}$)

$$\varepsilon(N) \leq \frac{n^{2c}}{2^{n^{\delta'}}} + \frac{n^{2c/d-8d}n^{2d}}{2^{n-10}} \leq 2^{-N^\delta},$$

where we set $\delta = 0.001\delta'/d$ and assume n to be large enough. This applies for any c, so the ensemble is pseudorandom against P/poly.[12]

Moreover, each $\mathsf{Ctr}^{(i,j)}_{f_{\mathsf{key}}}$ is implementable using a reversible circuit of size $O(n^d) \leq O(N)$ by standard simulation of Boolean circuits with reversible circuits. There are at most $O(N^2)$ many applications of such a permutation, so there are reversible circuits with at most $O(N^3)$ gates capable of implementing this ensemble. □

7.2 Hardness of MRCSP for Poly-Time Adversaries

As in [KC00], the capability of small reversible circuits to implement pseudorandom functions implies that their minimization problems are hard.

Corollary 76 (Theorem 15, restated)**.** Fix any constant $d \geq 3$. If one-way functions exist, then there is no deterministic $\mathrm{poly}(n)$-time algorithm that — given black-box access to a permutation π on $\{0,1\}^n$ — decides whether π is computable by a reversible circuit of size at most n^d.

Proof. Given that OWFs exist, by Corollary 75 there exists $\delta > 0$ such that there exists a 2^{-n^δ}-pseudorandom function ensemble $\left\{p^n_{(\cdot)}\right\}_{n\in\mathbb{N}}$ against P/poly implemented by reversible circuits with at most n^3 gates for some constant $\delta > 0$. Now assume that there exists such a deterministic $\mathcal{A}$. Then, by definition we have that for n large enough,

$$\left|\mathbf{Pr}_{\mathsf{key}\in\{0,1\}^{\kappa(n)}}\left[\mathcal{A}^{p^n_{\mathsf{key}}} = 1\right] - \mathbf{Pr}_{\mathbf{g}\in\mathfrak{S}_{\{0,1\}^n}}\left[\mathcal{A}^{\mathbf{g}} = 1\right]\right| \geq \left|1 - \frac{n^{4n^d}}{2^{n2^n}}\right| > 2^{-n^\delta}.$$

This is because for large enough n, there are at most n^{4n^d} permutations of $\{0,1\}^n$ computable by reversible circuits with at most n^d gates. This contradicts the pseudorandomness of $\left\{p^n_{(\cdot)}\right\}_{n\in\mathbb{N}}$. (When p is pseudorandom against subexponential algorithms, then this argument gives subexponential-time hardness.) □

[12] For subexponential-time adversaries, $q(n)$ is subexponential, and this is still exponentially small when f is subexponentially secure.

This shows that given the existence of OWFs, it is impossible for polynomial-time adversaries to invert our block ciphers. In fact, the result applies when one replaces reversible circuits with reversible circuits with nearest-neighbor gates, with some cost coming from the simulation of arbitrary gates with nearest-neighbor gates. In fact, our result applies to any restricted set of reversible circuits that can simulate arbitrary reversible circuits with 3-bit gates with polynomial overhead.

We leave a few directions open:

- Provide more efficient implementations using reversible circuits of the pseudorandom permutation ensemble of Proposition 72 (as in polynomial-size circuits with small exponent).
- Show average-case hardness of inverting block ciphers (an average-case version of Corollary 76) under a well-founded computational hardness assumption.

7.3 Hardness of MRCSP for General Adversaries

In Section 7.2 we showed that it is hard for circuits/algorithms running in time bounded in $\mathrm{poly}(n)$ (or subexponential in n) with query-access to a function to distinguish between functions of large reversible circuit complexity and those of small reversible circuit complexity. Part of our proof relied on the fact that such adversaries can query a limited number of inputs to the function that it receives. However, often we are interested in the case where the adversary can actually query the entire truth table and run in time $\mathrm{poly}(2^n)$. To show that such adversaries still struggle, we provide a direct reduction from a gapped version of the MCSP problem, which is known to be hard under the existence of OWFs. For this reduction we need a less naive simulation of Boolean circuits using reversible circuits than the one used in the proof of Proposition 72; we get a better simulation from [Wil10].

Definition 77. Let $a, b : \mathbb{N} \to \mathbb{N}$. In the promise problem $\mathsf{gapMCSP}_{a,b}$, the input is a truth-table of a permutation $f : \{0,1\}^n \to \{0,1\}^n$ such that there either exists a circuit of size at most $a(n)$ computing f or there is no circuit of size less than $b(n)$ computing f. The objective is to decide whether f falls into the former case or the latter case.

Lemma 78. If OWFs exist then there exists d such that the are no polynomial-size circuits for $\mathsf{gapMCSP}_{n^d,n^{100d}}$.

Proof. If OWFs exist then by [HILL99], [GGM86], and [LR88] there exists a 2^{-n^δ}-pseudorandom function ensemble against polynomial-time adversaries computable with circuits with at most n^d gates. The existence of such an ensemble precludes the existence of a polynomial-time algorithm for $\mathsf{gapMCSP}_{n^d,n^{100d}}$. □

Definition 79. Let $a, b : \mathbb{N} \to \mathbb{N}$. In the promise problem MRCSP_a, the input is a truth-table of a permutation $f : \{0,1\}^n \to \{0,1\}^n$ such that there either exists a reversible circuit with at most $a(n)$ gates computing f or there is no circuit with less than $a(n)$ gates computing f. The objective is to decide whether f falls into the former case or the latter case.

Proposition 80. Let $d < d'/10$. If there exist circuits for $\mathsf{MRCSP}_{n^{5d}}$ with size at most $T(2^n)$ then there exist circuits for $\mathsf{gapMCSP}_{n^d,n^{d'}}$ with size at most $T(2^{2n}) + O(2^{4n})$.

Proof. Assume the existence of such circuits for MRCSP_{n^d,n^d}. We present distinguisher circuits $\mathcal{D}_n$ for $\mathsf{gapMCSP}_{n^{5d},n^{d'}}$ satisfying the size bound. The circuit $\mathcal{D}_n$ proceeds as follows. Given the input to $\mathsf{gapMCSP}_{n^d,n^{d'}}$, which is a truth table

$$\mathsf{tt} = (x_i, y_i)_{i \in [2^n]},$$

where each x_i and y_i are elements of $\{0,1\}^n$. Now form the truth table tt' which is given by

$$\mathsf{tt}' = (x_i \circ z, y_i \circ z)_{i \in [2^n], z \in \{0,1\}^n},$$

where $\circ$ denotes concatenation of strings, so that tt' is the truth table of a permutation on $\{0,1\}^{2n}$. Then $\mathcal{D}_n$ runs the size $T(2^{2n})$ circuit $\mathcal{C}'_{2n}$ on tt' and returns 1 if $\mathcal{C}'$ returns 1 and 0 otherwise.

First assume that tt is computable by circuits of size at most n^d. Then by Claim 81 tt' is computable by reversible circuits with at most $n^{5d} \leq (2n)^{5d}$ gates, so that our $\mathcal{D}_n$ is correct on such an input.

Claim 81 ([Wil10], Section 2.2.1)**.** If $f : \{0,1\}^n \to \{0,1\}^n$ is computable using circuits of size n^d, then there exists circuits of size at most n^{5d} computing the function $f \circ \mathrm{Id} : \{0,1\}^{2n} \to \{0,1\}^{2n}$ given by $(f \circ \mathrm{Id})(x \circ z) = f(x) \circ z$.

Otherwise, tt (the truth table for the function f) has no circuits of size at most $n^{d'}$, and therefore $f \circ \mathrm{Id}$ has no reversible circuits of size at most $n^{d'/2}$. This implies that $f \circ \mathrm{Id}$ has no reversible circuits of size at most $n^d < n^{d'/2}$, so that the internal circuit $\mathcal{C}'_{2n}$ returns 0. Therefore $\mathcal{D}_n$ returns 0 and is again correct. The size of $\mathcal{D}_n$ is at most $T(2^{2n}) + O(2^{4n})$ because it simply builds the size 2^{4n} truth table tt' and runs a size at most $T(2^{2n})$ circuit on it. □

Combining Lemma 78 and Proposition 80 results in the conditional (based on the existence of OWFs) hardness of MRCSP.

Corollary 82. If OWFs exist then there exists $d > 0$ such that $\mathsf{MRCSP}_{n^d} \notin \mathsf{P/poly}$.

Acknowledgements

W.H. thanks Angelos Pelecanos and Lucas Gretta for helpful discussions on the comparison method, and for allowing him to reuse results from their discussion and in-progress work in Appendix A.

References

[AALV09] Dorit Aharonov, Itai Arad, Zeph Landau, and Umesh Vazirani. The Detectability Lemma and Quantum Gap Amplification. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, pages 417–426, 2009.

[Ben73] Charles H. Bennett. Logical Reversibility of Computation. *IBM Journal of Research and Development*, 17(6):525–532, 1973.

[BH08] Alex Brodsky and Shlomo Hoory. Simple Permutations Mix Even Better. *Random Structures & Algorithms*, 32(3):274–289, 2008.

[BHH16] Fernando G.S.L. Brandao, Aram W. Harrow, and Michał Horodecki. Local Random Quantum Circuits are Approximate Polynomial-Designs. *Communications in Mathematical Physics*, 346:397–434, 2016.

[CG75] Don Coppersmith and Edna Grossman. Generators for Certain Alternating Groups with Applications to Cryptography. *SIAM Journal on Applied Mathematics*, 29(4):624–627, 1975.

[Cle90] Richard Cleve. Complexity Theoretic Issues Concerning Block Ciphers Related to DES. In *Conference on the Theory and Application of Cryptography*, pages 531–544. Springer, 1990.

[EG83] Shimon Even and Oded Goldreich. DES-like Functions can Generate the Alternating group. *Transactions on Information Theory*, 29(6):863–865, 1983.

[FT82] Edward Fredkin and Tommaso Toffoli. Conservative Logic. *International Journal of Theoretical Physics*, 21(3):219–253, 1982.

[GGM86] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to Construct Random Functions. *Journal of the ACM (JACM)*, 33(4):792–807, 1986.

[Gow] W. T. Gowers. Are There Very Strongly Pseudorandom Permutations? MathOverflow. URL:https://mathoverflow.net/q/134060 (version: 2013-06-27).

[Gow96] W.T. 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.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A Pseudorandom Generator from any One-Way Function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[HMMR05] Shlomo Hoory, Avner Magen, Steven Myers, and Charles Rackoff. Simple Permutations Mix Well. *Theoretical Computer Science*, 348(2-3):251–261, 2005.

[HP07] Patrick Hayden and John Preskill. Black Holes as Mirrors: Quantum Information in Random Subsystems. *Journal of High Energy Physics*, 2007(09):120, 2007.

[Kas07] Martin Kassabov. Symmetric Groups and Expander Graphs. *Inventiones Mathematicae*, 170(2):327–354, 2007.

[KC00] Valentine Kabanets and Jin-Yi Cai. Circuit Minimization Problem. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pages 73–79, 2000.

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

[Lan] R. Landauer. Dissipation and Heat Generation in the Computing Process. *IBM Journal of Research and Development*, 5.

[Lec63] Yves Lecerf. Machines de Turing réversibles. Récursive insolubilité en $n \in N$ de l'équation $u = \theta^n u$, où θ est un "isomorphisme de codes". *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 257:2597–2600, 1963.

[LPTV23] Tianren Liu, Angelos Pelecanos, Stefano Tessaro, and Vinod Vaikuntanathan. Layout Graphs, Random Walks and the t-Wise Independence of SPN Block Ciphers. In *Annual International Cryptology Conference*, pages 694–726. Springer, 2023.

[LR88] Michael Luby and Charles Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM Journal on Computing*, 17(2):373–386, 1988.

[LTV21] Tianren Liu, Stefano Tessaro, and Vinod Vaikuntanathan. The t-wise independence of substitution-permutation networks. In *Advances in Cryptology–CRYPTO 2021: 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16–20, 2021, Proceedings, Part IV 41*, pages 454–483. Springer, 2021.

[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.

[MP04] Ueli Maurer and Krzysztof Pietrzak. Composition of Random Systems: When Two Weak Make One Strong. In *Theory of Cryptography Conference*, pages 410–427. Springer, 2004.

[Nac96] Bruno Nachtergaele. The Spectral Gap for Some Spin Chains with Discrete Symmetry Breaking. *Communications in Mathematical Physics*, 175:565–606, 1996.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[OSP23] Ryan O'Donnell, Rocco A. Servedio, and Pedro Paredes. Explicit Orthogonal and Unitary Designs. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1240–1260. IEEE, 2023.

[Tof80] T. Toffoli. Reversible Computing/Automata, Languages and Programming, Volume 85 of Lecture Notes in Computer Science, 1980.

[Wil10] Colin P Williams. *Explorations in quantum computing*. Springer Science & Business Media, 2010.

[WLP09] E.L. Wilmer, David A. Levin, and Yuval Peres. Markov Chains and Mixing Times. *American Mathematical Soc., Providence*, 2009.

A Comparison Method

Theorem 83 ([WLP09], Theorem 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

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

Then we have that

$$\lambda_2(L) \geq \left(\max_{v \in V} \frac{\pi(v)}{\widetilde{\pi}(v)}\right) A(\boldsymbol{\Delta}) \lambda_2(\widetilde{L}).$$

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

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

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

Lemma 32 is a direct consequence of the following Corollary 84, with symmetry applied.

Corollary 84. Let $\widetilde{P}$ and P be transition matrices for random walks on undirected (multi)graphs $\widetilde{G} = (V, \widetilde{E})$ and $G = (V, E)$, respectively. The graph G is d-regular, and $\widetilde{G}$ is $\widetilde{d}$-regular. Assume that for each $e \in E$ there exists a random path

$$\mathbf{\Gamma}(e) = (\widetilde{\boldsymbol{e}}_1, \ldots, \widetilde{\boldsymbol{e}}_{T_e}),$$

where $(\widetilde{\boldsymbol{e}}_1, \ldots, \widetilde{\boldsymbol{e}}_{T_e})$[13] is drawn from a distribution on sequences of edges connecting the endpoints of E. Then we have for any $f : V \to \mathbb{R}$ that

$$\lambda_2(L) \geq \left(\max_{v\in V} \frac{\pi(v)}{\widetilde{\pi}(v)}\right) B(\mathbf{\Gamma})\lambda_2(\widetilde{L}).$$

where the *multigraph comparison constant* is defined as the maximum congestion over all edges $\tilde{e} \in \widetilde{E}$ connecting vertices $a, b \in V$,

$$B(\mathbf{\Gamma}) \coloneqq \max_{\widetilde{e}\in\widetilde{E}} \left\{ \frac{\widetilde{d}}{d} \sum_{e\in E} \mathop{\mathbf{E}}_{\mathbf{\Gamma}} \left[\mathbf{1}_{\widetilde{e}\in\mathbf{\Gamma}(e)} \cdot |\mathbf{\Gamma}(e)|\right] \right\}.$$[14]

Here π and $\widetilde{\pi}$ are the (unique) stationary distributions for P and $\widetilde{P}$, respectively.

Proof. We construct a (randomized) map $\mathbf{\Delta}$ as in Lemma 32 from the map $\mathbf{\Gamma}$ as follows. For each $(x, y) \in V^2$, if x and y are not connected by an edge in E then set $\mathbf{\Delta}(x, y) = ()$ (the sequence of length 0). Otherwise select a random edge $\boldsymbol{e}$ from $(E)_{x,y}$ and let $(\widetilde{\boldsymbol{e}}_1, \ldots, \widetilde{\boldsymbol{e}}_{T_e})$ be the random path $\mathbf{\Gamma}(\boldsymbol{e})$. For each $i \in [T_e]$ let $(\boldsymbol{u}_i, \boldsymbol{v}_i)$ be the vertices connected by $\widetilde{\boldsymbol{e}}_i$. Then set

$$\mathbf{\Delta}(x, y) = \big((\boldsymbol{u}_1, \boldsymbol{v}_1), \ldots, (\boldsymbol{u}_\ell, \boldsymbol{v}_{T_e})\big) = \big((x, \boldsymbol{v}_1), \ldots, (\boldsymbol{u}_{T_e}, y)\big).$$

The comparison constant of $\mathbf{\Delta}$ is

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

Since both Markov chains have the same stationary distribution over the same state space, the π terms cancel, so the above is equal to

$$\begin{aligned} &\max_{(a,b)\in V^2} \left\{ \frac{1}{\widetilde{P}(a,b)} \sum_{(x,y)\in V^2} \mathop{\mathbf{E}}_{\mathbf{\Delta}} \left[\mathbf{1}_{(a,b)\in\mathbf{\Delta}(x,y)} \cdot |\mathbf{\Delta}(x, y)|\right] \cdot P(x, y) \right\} \\ = &\max_{(a,b)\in V^2} \left\{ \frac{\widetilde{d}}{\left|\widetilde{E}_{a,b}\right|} \sum_{(x,y)\in V^2} \mathop{\mathbf{E}}_{\mathbf{\Delta}} \left[\mathbf{1}_{(a,b)\in\mathbf{\Delta}(x,y)} \cdot |\mathbf{\Delta}(x, y)|\right] \cdot \frac{|E_{x,y}|}{d} \right\}. \end{aligned}$$

[13] Here T_e is a (deterministic) quantity determined by the edge e.

[14] If $\mathcal{P}$ is a sequence then $|\mathcal{P}|$ is its length.

Let us now start translating from pairs of vertices to edges of the multigraph. Each pair of vertices (x, y) has $\left|E_{x,y}\right|$ edges connecting them, we can change the summation from pairs of vertices to edges $e \in E$. Recall that $u(e), v(e)$ are the endpoints of edge e. Continuing the calculation, the above is equal to

$$\begin{aligned}
&\max_{(a,b)\in V^2} \left\{ \frac{\widetilde{d}}{d\left|\widetilde{E}_{a,b}\right|} \sum_{e\in E} \mathop{\mathbf{E}}_{\mathbf{\Delta}} \left[\mathbf{1}_{(a,b)\in \mathbf{\Delta}(u(e),v(e))} \cdot |\mathbf{\Delta}(u(e), v(e))| \right] \right\} \\
&= \max_{(a,b)\in V^2} \left\{ \frac{\widetilde{d}}{d} \sum_{e\in E} \frac{1}{\left|\widetilde{E}_{a,b}\right|} \mathop{\mathbf{E}}_{\mathbf{\Delta}} \left[\mathbf{1}_{(a,b)\in \mathbf{\Delta}(u(e),v(e))} \cdot |\mathbf{\Delta}(u(e), v(e))| \right] \right\} \\
&= \max_{(a,b)\in V^2} \left\{ \frac{\widetilde{d}}{d} \sum_{e\in E} \frac{1}{\left|\widetilde{E}_{a,b}\right|} |\mathbf{\Gamma}(e)| \mathop{\mathbf{Pr}}_{\mathbf{\Delta}} \left[\mathbf{1}_{(a,b)\in \mathbf{\Delta}(u(e),v(e))} \right] \right\}.
\end{aligned}$$

The last equality is because $|\mathbf{\Delta}(u(e), v(e))| = |\mathbf{\Gamma}(e)|$ with certainty, and $|\mathbf{\Gamma}(e)|$ is a deterministic quantity that only depends on e.

The sum of probabilities that $\widetilde{e} \in (\widetilde{E})_{a,b}$ appears in the sequence $\mathbf{\Gamma}(e)$, over all such $\widetilde{e}$, is equal to the probability that (a, b) appears in $\mathbf{\Delta}(u(e), v(e))$. By averaging, we have that the probability that $\widetilde{e} \in \widetilde{E}_{a,b}$ appears in $\mathbf{\Gamma}(e)$, where $\widetilde{e}$ maximizes this quantity, is at least $\frac{1}{|\widetilde{E}_{a,b}|}$ times the appearance probability of (a, b). This results in

$$A(\mathbf{\Delta}) \leq \max_{\widetilde{e}\in\widetilde{E}} \left\{ \frac{\widetilde{d}}{d} \sum_{e\in E} \mathop{\mathbf{E}}_{\mathbf{\Gamma}} \left[\mathbf{1}_{\widetilde{e}\in\mathbf{\Gamma}(e)} \cdot |\mathbf{\Gamma}(e)| \right] \right\} = B(\mathbf{\Gamma}).$$

Applying Theorem 83 with this new map $\mathbf{\Delta}$ completes the proof. □