Quantum Pseudorandom Scramblers

Chuhan Lu[1] [0009-0000-3359-320X], Minglong Qin[2] [0009-0004-8760-5498], Fang Song[1] [0000-0002-3098-6451], Penghui Yao[2,3] [0000-0002-4104-2069], and Mingnan Zhao[2] [0009-0006-9623-9703]

[1] Computer Science Department, Portland State University, USA
{chuhan,fang.song}@pdx.edu
[2] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
mlqin@smail.nju.edu.cn {phyao1985,mingnanzh}@gmail.com
[3] Hefei National Laboratory, Hefei 230088, China

Abstract. Quantum pseudorandom state generators (PRSGs) have stimulated exciting developments in recent years. A PRSG, on a fixed initial (e.g., all-zero) state, produces an output state that is computationally indistinguishable from a Haar random state. However, pseudorandomness of the output state is not guaranteed on other initial states. In fact, known PRSG constructions *provably* fail on some initial state.
In this work, we propose and construct quantum Pseudorandom State Scramblers (PRSSs), which can produce a pseudorandom state on an *arbitrary* initial state. In the information-theoretical setting, we obtain a scrambler which maps an arbitrary initial state to a distribution of quantum states that is close to Haar random in *total variation distance*. As a result, our scrambler exhibits a *dispersing* property. Loosely, it can span an ϵ-net of the state space. This significantly strengthens what standard PRSGs can induce, as they may only concentrate on a small region of the state space as long as the average output state approximates a Haar random state in total variation distance.
Our PRSS construction develops a *parallel* extension of the famous Kac's walk, and we show that it mixes *exponentially* faster than the standard Kac's walk. This constitutes the core of our proof. We also describe a few applications of PRSSs. While our PRSS construction assumes a post-quantum one-way function, PRSSs are potentially a weaker primitive and can be separated from one-way functions in a relativized world similar to standard PRSGs.

Keywords: Quantum pseudorandom states · Kac's walk · Pseudorandom unitary operators

1 Introduction

Pseudorandomness is a fundamental concept in complexity theory and cryptography, offering efficient approximation to true randomness against computationally bounded adversaries. Recently, Ji, Liu and Song [31] introduced quantum pseudorandom state generators (PRSGs) as a family of quantum states

$\{|\phi_k\rangle\}_{k\in\mathcal{K}}$, which can be generated in polynomial time, and no computationally-bounded quantum adversary can distinguish polynomially many copies of $|\phi_k\rangle$ from polynomially many copies of a Haar random state. PRSGs can be considered as a quantum counterpart to classical pseudorandom generators, and can be constructed assuming the existence of one-way functions that are hard for efficient quantum adversaries [31,10,11,4,1]. What is surprising, PRSGs are proven weaker than one-way functions in a relativized world [36,37]. Since one-way functions are considered the *minimal* assumption in classical cryptography, this opens up the possibility of basing quantum cryptography on *weaker* assumptions. There have been exciting advances in recent years, realizing a host of cryptographic tasks based on PRSGs [6,5,40,4,19]. In addition to cryptographic interest, pseudorandom states have also inspired new developments for quantum gravity theory and string theory [8,35,12,1,47].

Another fundamental quantum pseudorandom primitive, *pseudorandom unitary operators* (PRU), was also introduced in [31] as a quantum analogue of pseudorandom functions. A PRU is a set of polynomially-time unitary operators that are computationally indistinguishable from Haar random unitaries. PRUs clearly imply PRSGs and could further enrich the toolkit in cryptography and physics [8,35,12,47,22]. Nonetheless, constructing a provably-secure PRU remains an open problem, and progress has been slow (e.g., conjectured constructions in [31], a stateful simulation in [2], and on the negative side some barriers such as impossibility of PRUs that are sparse or of real entries [27]). In fact, even basic properties that are *necessary* for PRUs have not been achieved. It is easy to see that a PRU gives a family of polynomial-sized quantum circuits which can map an *arbitrary* pure state to a family of pseudorandom states. However, a PRSG can be viewed as a family of polynomial-sized quantum circuits which map a specific initial state, typically $|0^n\rangle$, to a family of pseudorandom states. Indeed, all existing construction of PRSGs necessitates a specific initial state, and it can be shown that they *fail* to produce pseudorandom states for certain initial states. This limitation has indeed caused a variety of technical challenges in the cryptographic applications mentioned before that need to be addressed in ad hoc ways. It hence becomes imperative to understand the following question and its consequences.

> *Can we construct a family of polynomial-sized quantum circuits which map an arbitrary input (pure) state to pseudorandom states?*

1.1 Our Contributions

In this work, we answer the question affirmatively as a steady step towards bridging the gap between PRSGs and PRUs. We formally encapsulate the property of "scrambling" an arbitrary input state in a novel quantum pseudorandom primitive, termed a *quantum pseudorandom state scrambler* (PRSS), which isometrically maps an *arbitrary* pure state to a pseudorandom state. We then construct a PRSS based on any quantum-secure PRF. A central technical novelty is to design a *parallel* version of Kac's walk, which is a random walk on a unit sphere, and prove a mixing time *exponentially* faster than the standard Kac's walk [44]. Although Kac's walk was introduced by Kac in [33] more than half a century ago and has been studied by a large body of works since then, this work, to our knowledge, is the first time to employ Kac's walk to design quantum pseudorandom objects.

Our construction also exhibits a notable *dispersing* property. Loosely speaking, the output states of our scrambler constitute an ϵ-net on the sphere, and the distribution closely approximates the Haar random distribution under the strong Wasserstein distance, when sufficient randomness is supplied. Such a powerful "randomizing" capability needs not be present even in PRUs.

Overview on the construction and analysis. Our construction is inspired by Kac's walk, originally a model for a Boltzmann gas [33]. This approach differs from previous constructions for PRSGs. Let us consider an arbitrary unit-vector $v \in \mathbb{R}^N$. In one step of Kac's walk, two distinct coordinates (i, j) and an angle $\theta \in [0, 2\pi)$ are chosen uniformly at random. Then $R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is applied to rotate the two-dimensional subvector $(v_i, v_j)^T$. It is proven that it converges to the Haar measure on the unit sphere of $\mathbb{R}^N$ in $O(N \log N)$ steps [44]. However, if we view the input vector as an $n = \log N$-qubit state, then the factor N in the mixing time is prohibitive for the purpose of an efficient (polynomial in n) scrambler.

Can we *parallelize* Kac's walk in hope of shaking off a factor of N? Notice that in Kac's walk, if any two consecutive steps overlap on the random choices of coordinates, then they need to be executed in *sequence*. One might consider conditioning on the event of "collision-free" in the coordinate choices, but this occurs with negligibly small probability since we intend to compress $\Omega(N)$ steps into one.

We design a *parallel* Kac's walk that rapidly mixes in $O(\log N)$ time, an *exponential* improvement over the original walk. In each step, instead of working with an individual pair of coordinates, we randomly partition the N coordinates into $N/2$ pairs, then each pair is rotated by a random angle chosen *independently*. Although the mixing time of Kac's walk is not directly applicable, we show that the specific path-coupling proof strategy of [44] can be extended here.

We then construct a quantum circuit to implement our parallel Kac's walk. In each step, we use a random permutation to realize the coordinate partition, and employ a random function to compute a random rotation angle, under a careful discretization, for each pair of coordinates. Finally, we obtain our pseudorandom state scramblers by replacing the random permutations and functions with quantum-secure pseudorandom permutations and functions, which exist based on post-quantum one-way functions [50].

The discussion so far works with real Hilbert spaces. To construct a PRSS in a complex Hilbert space, we further develop a parallel Kac's walk on complex Hilbert spaces. The construction starts likewise by randomly partitioning N coordinates to $N/2$ pairs, and then applying random 2×2 *unitary* matrices independently to each pair. As unitary matrices have more degrees of freedom than real orthogonal rotation matricies, the analysis of the mixing time is more involved. The extension of Kac's walk to a complex Hilbert space, as well as the parallelization, has not been studied previously as far as we are aware. This may be of independent interest.

Applications. It is easy to see that PRSSs subsume standard PRSGs as well as scalable PRSGs. We also demonstrate that PRSSs can be used to achieve a black-box realization of a variant of PRSGs known as pseudorandom function-like state generators (PRFSGs), which in turn enable a host of cryptographic primitives such as IND-CPA SKE and EUF-CMA MAC [5,4]. A PRFSG takes an additional classical input x (from a poly-size domain) and produces a pseudorandom state. In the literature, a PRFSG (with logarithmic input length) can be constructed from PRSGs by measuring a part of a pseudorandom state and post-select on x. This inevitably is error-prone and consumes multiple copies, i.e., multiple invocations of a PRSG, to evaluate on a single x. Given our PRSS (with a sufficiently long key), we can simply feed $|x\rangle$ as the initial state to the PRSS, and hence only *one*, rather than polynomially-many, run of PRSS suffices.

We observe that the argument by Kretschmer [36] also implies that PRSS is *strictly* weaker than one-way functions relative to an oracle. Thus PRSSs may further enhance the new cryptographic landscape without assuming one-way functions. We demonstrate some use cases of PRSSs beyond what are already possible from PRSGs. For starters, a PRSS enables efficient encryption of quantum messages by effectively "scrambling" any initial state, and allowing *multiple copies* of the same state to be encrypted under the same key. The fact that PRSS provides a secure encryption also enables committing quantum states, thanks to a new characterization of [22]. The commitment scheme can be further made *succinct*, where the commitment message has smaller size than the size of the message to be committed. Existing constructions rely on potentially stronger assumptions than PRSSs.

Subsequent work

A follow-up work [3] gave a construction that is indistinguishable from applying the tensor product of a Haar random isometry when the input state is restricted to one of three special families: (1) $|\psi\rangle^{\otimes q}$ for a pure state $|\psi\rangle$ and polynomially-bounded q; (2) $\bigotimes_{i=1}^{q} |x_i\rangle$; and (3) $\otimes_{i=1}^{q} |\phi_i\rangle$, where every ϕ_i is Haar random. Their construction requires adding an ancilla system $|0^m\rangle$, and the security loss scales with $1/2^m$. As a result, it *necessarily* cannot preserve the input dimension and m is chosen to be a polynomial to obtain negligible security loss. This also incurs poly overheads in the applications such as quantum encryption. In other words, it only (unitarily) scrambles states $|\psi\rangle |0^m\rangle$ for a $|\psi\rangle$ chosen from one of the three families above and a polynomial m. More recently, several independent works on constructing pseudorandom unitaries which are secure against non-adaptive queries have been presented. Metger, Poremba, Sinha and Yuen [39] proposed a construction using a composition of Clifford gates, pseudorandom functions and pseudorandom permutations. Brakerski and Magrafta [9] presented a construction for real-valued unitaries that look like Haar random on any polynomial-sized set of orthogonal input states. Chen, Bouland, Brandao, Docter, Hayden, and Xu [18] achieved similar results via products of exponentiated sums of random permutations with random phases. It is not clear whether these constructions are able to generate an ϵ-net and realize the *dispersing* property (details in Appendix A), a strong randomizing property achieved by our construction.

1.2 Discussions and Open Questions

There is a rich history of studying Kac's walk in probability and mathematical physics[26,21,30,43,32,28]. Determining the total variation mixing time of Kac's walk is particularly challenging, and it is currently only known to be between the order $O(n^4 \log n)$ and $O(n^2)$ [45].

There has also been extensive efforts on approximations to Haar measures in a *statistical* setting, known as state and unitary t-designs [46,20]. For instance, a unitary t-design mimics a Haar random unitary up to the t-th moment. It is known that a unitary t-design can be constructed by a quantum circuit of size polynomially in t, composed of Haar random single or two-qubit gates [24,13,23,25,42]. It is interesting to note that a path-coupling technique in [43] for analyzing Kac's walk also plays an essential role in the proofs of these unitary design results. It is reasonable to anticipate improvements on the efficiency of the unitary designs with new advances on Kac's walk. However, it is worth stressing that another critical component in their proofs involving spectral gaps appears to inevitably incur a dependency on t, which is a serious limitation. For instance, in order for the output state to approximate a Haar random state when the number of copies can be an arbitrary polynomial, we would need to pick a superpolynomial t in the unitary design. As far as we know, our PRSS is the first to employ Kac's walk directly in the construction of a quantum pseudorandom object, and the exponential improvement on the mixing time of our parallel walk enables flipping the quantifiers, i.e., a fixed poly-size construction that is nonetheless pesudorandom against any polynomial-time distinguisher, a desired feature towards PRUs.

Kac's walks have also found applications in algorithm design. Recently, a fast and memory-optimal dimension-reduction algorithm is proposed based on Kac's walk and its discrete variants [29]. We would like to invite more exploration of Kac's walk in theoretical computer science broadly.

We describe several interesting open problems emerged from our work.

1. Is it possible to simplify the quantum circuits for these primitives? Can we replace random permutations by a sequence of parallel (pseudo) random local permutations? Can we use the same random rotation or even a fixed one (e.g., Hadamard transform) in a single iteration? Recent advances on repeated averages on graphs [41] and orthogonal repeated averaging [17,29] alludes to an affirmative answer.

2. We believe that PRSSs, potentially weaker than PRUs, are an important primitive in its own right. Can we discover more applications of PRSSs and the dispersing property, especially in cryptography as well as in quantum gravity theory? For example, We envision a form of *uncloneable knowledge tokens* from a PRSS that may enable novel quantum proof systems and *delegated* computation.
3. Is our construction of PRSS capable of scrambling polynomial quantum states? This appears to require strengthening the coupling technique in our current analysis, and it might be useful to analyze other variants of Kac's walk.
4. How far are we from a PRU? Can we get it by strengthening our parallel Kac's walk approach or can we show that our construction is already a PRU? By a simple hybrid argument, it suffices prove that our parallel Kac's walk on $\mathrm{SO}(N)$ converges within $\mathrm{polylog}(N)$ time in terms of the L^∞ Wasserstein distance. Indeed, there has been a large body of work devoting to studying the speed of the convergence with respect to different metrics [21,33,32,45]. One of the most relevant works is Oliveira's result [43] showing a tight convergence time of order $O(N^2 \log N)$ with respect to the stronger L^2 Wasserstein distance. Our parallelization achieves a quadratic speedup, which leads to an $\tilde{O}(2^n)$-time construction of PRU. Since the L^∞ Wasserstein distance is a less stringent metric than the L^2 Wasserstein distance, there is hope to obtain an improved convergence rate. To our knowledge, the speed of convergence of Kac's walk with respect to L^∞ Wasserstein distance has not been studied, and hence developing new techniques to overcome the tightness of Oliveira's L^2 result would be an exciting research direction.

Acknowledgment. We thank the anonymous reviewers' feedback. CL and FS were supported in part by the US National Science Foundation grants CCF-2042414, CCF-2054758 (CAREER) and CCF-2224131. MQ, PY and MZ were supported in part by National Natural Science Foundation of China (Grant No. 62332009, 61972191), and Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901). FS thanks Zhengfeng Ji and Yi-Kai Liu for discussions on the topic of quantum pseudorandomness.

Organization. Section 2 contains preliminary materials on basic notations and cryptographic primitives. Section 3 describes definitions and properties of our new primitives. Section 4 introduces the parallel Kac's walk. Then Section 5 constructs PRSSs via implementing the parallel Kac's walk in Section 4. Section 6 describes applications of PRSSs. In Appendix A we introduce dispersing RSS. In Appendix B, we give details on the connections between PRSSs and existing PRS variants. Some proofs are deferred to Appendix C.

2 Preliminary

2.1 Basic Notation

For $n \in \mathbb{N}$, $[n]$ denotes $\{1, \dots, n\}$. For $x \in \{0,1\}^n$, we use x_i to denote the i-th bit of x and define $\mathrm{val}(x) = \sum_{i=1}^{n} 2^{-i} x_i$. Suppose that x and y are bit strings of finite length, we denote xy to be the concatenation of x and y. For finite sets $\mathcal{X}$ and $\mathcal{Y}$, we use $\mathcal{X}^{\mathcal{Y}}$ to denote the set of all functions $\{f : \mathcal{X} \to \mathcal{Y}\}$. We use $S_{\mathcal{X}}$ to denote the *permutation group* over elements in a finite set $\mathcal{X}$. We often write S_{2^n} instead of $S_{\{0,1\}^n}$ to denote the permutation group over elements in $\{0,1\}^n$.

For any symbol x and $n \in \mathbb{N}$, $(x_i)_{i=1}^n$ represents $(x_1, \dots, x_n)$. With a slight abuse of notation, we let $(x_i)_{i=1}^n \subseteq S$ represent $x_i \in S$ for all $i \in [n]$. For $n \in \mathbb{N}$, $\mathcal{S}_{\mathbb{R}}^n$ denotes the set of all unit vectors in $\mathbb{R}^n$, $\mathcal{S}_{\mathbb{C}}^n$ denotes the set of all unit vectors in $\mathbb{C}^n$, $\mathrm{SO}(n)$ denotes the special orthogonal group of $n \times n$ real matrices, $\mathrm{SU}(n)$ denotes the special unitary group of $n \times n$ complex matrices, $\mathrm{O}(n)$ denotes the $n \times n$ orthogonal group and $\mathrm{U}(n)$ denotes the $n \times n$ unitary group. For a Hilbert space $\mathcal{H}$, we use $\mathcal{S}(\mathcal{H})$ to denote the set of pure quantum states in $\mathcal{H}$ and $\mathcal{D}(\mathcal{H})$ to denote the set of density operators on $\mathcal{H}$.

For an n-dimensional vector v and $i \in [n]$, we use $v[i]$ to denote the i-th coordinate of v. For $S \subseteq [n]$ and $v \in \mathbb{C}^n$, define

$$\|v\|_1 = \sum_{i \in [n]} |v[i]| \quad , \quad \|v\|_{1,S} = \sum_{i \in S} |v[i]| \quad , \quad \|v\|_2 = \sqrt{\sum_{i \in [n]} |v[i]|^2} \ .$$

For an $n \times n$ matrix M and $p \in \mathbb{N}$, the p-norm of M is defined to be $\|M\|_p = \left(\mathrm{Tr}\left[\left(M^\dagger M\right)^{p/2}\right]\right)^{1/p}$, and $\|M\|_\infty$ is defined to be the largest singular value of M. The following fact will be used in our paper and is easy to prove by the triangle inequality.

Fact 1. *Given $m, n \in \mathbb{N}$, $U_1, \dots, U_m, V_1, \dots, V_m \in \mathrm{O}(n)$ (or $\mathrm{U}(n)$), then*

$$\|U_1 \dots U_m - V_1 \dots V_m\|_\infty \leq \sum_{i=1}^{m} \|U_i - V_i\|_\infty \ .$$

Given two density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, the trace distance between ρ and σ is $\mathrm{TD}(\rho, \sigma) = \|\rho - \sigma\|_1$.

Let $\mathcal{V}$ be a real or complex vector space, and $\epsilon > 0$ be a positive real number. For any $\mathcal{S} \subseteq V$, a set of vectors $\mathcal{N} \subseteq \mathcal{S}$ is said to be an *ϵ-net* of $\mathcal{S}$ if, for every vector $u \in \mathcal{S}$, there exists a vector $v \in \mathcal{N}$ such that $\|u - v\|_2 \leq \epsilon$.

We adopt the standard quantum circuit model. A quantum circuit with gates drawn from a finite gate set can be encoded as a binary string. $\{Q_\lambda : \lambda \in \mathbb{N}\}$ is said to be a *polynomial-time* generated family[4] if there exists a deterministic Turing machine that, on any input $\lambda \in \mathbb{N}$, outputs an encoding of Q_λ in

[4] More precisely, each circuit should be written as Q_{1^λ}. Note that in a polynomial-time generated family, then Q_λ must have size polynomial in λ.

polynomial-time in λ. A quantum polynomial-time algorithm is identified with a polynomial-time generated circuit family. In cryptography it is conventionally to model adversaries as *non-uniform* algorithms. We model a non-uniform quantum polynomial-time algorithm as a family $\{Q_\lambda, \rho_\lambda\}_\lambda$, where $\{Q_\lambda\}$ is a polynomial-time generated circuit family, and $\{\rho_\lambda\}$ is a collection of *advice states*. Q_λ acts on ρ_λ besides the actual input state.

2.2 Probability Theory

For two probability measures ν_1 and ν_2 defined on measurable space $(\Omega, \mathcal{F})$, the *total variation distance* of ν_1 and ν_2 is defined as

$$\|\nu_1 - \nu_2\|_{\mathrm{TV}} = \sup_{A \in \mathcal{F}} |\nu_1(A) - \nu_2(A)| \quad .$$

Closeness in total variation distance is a strong promise. For example, when applied to quantum states, it implies closeness in trace distance of the average states.

Lemma 1. *Let μ and ν be two arbitrary probability measures over $\mathcal{S}_{\mathbb{R}}^{2^n}$ ($\mathcal{S}_{\mathbb{C}}^{2^n}$). Then for all $\ell \in \mathbb{N}$,*

$$\left\| \mathop{\mathbb{E}}_{|\psi\rangle \sim \mu}\left[(|\psi\rangle\langle\psi|)^{\otimes \ell}\right] - \mathop{\mathbb{E}}_{|\varphi\rangle \sim \nu}\left[(|\varphi\rangle\langle\varphi|)^{\otimes \ell}\right] \right\|_1 \leq \|\mu - \nu\|_{\mathrm{TV}} \quad .$$

We denote the distribution of a random variable X by $\mathcal{L}(X)$. If $\mathcal{L}(X) = \nu$, we write $X \sim \nu$. A coupling of two probability measures μ and ν is a joint probability measure whose marginals are μ and ν. We use $\Gamma(\mu, \nu)$ to denote the set of all couplings of μ and ν. For $p \geq 1$ The *Wasserstein p-distance* between two probability measures μ and ν is

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \mathop{\mathbb{E}}_{(x,y) \sim \gamma} [\|x - y\|_2^p] \right)^{1/p} \quad .$$

The Wasserstein ∞-distance is $W_\infty(\mu, \nu) = \lim_{p \to \infty} W_p(\mu, \nu)$.

The following lemmas about Markov chain in [38] serve as crucial tools in this work.

Lemma 2 (Coupling Lemma). *[38, Theorem 5.4] Let K be the transition kernel of a Markov chain with unique stationary distribution ν on state space Ω. Let $\{X_t\}_{t \geq 0}, \{Y_t\}_{t \geq 0}$ be two corresponding Markov chains started at $X_0 = x \in \Omega$ and $Y_0 \sim \nu$. Define the coalescence time of the chains*

$$\tau(x) = \min \{t : X_t = Y_t\} \quad .$$

Assume that a coupling of $\{X_t\}_{t \geq 0}, \{Y_t\}_{t \geq 0}$ satisfies $X_t = Y_t$ for all $t \geq \tau(x)$. Then for any $t \geq 0$,

$$\|\mathcal{L}(X_t) - \nu\|_{\mathrm{TV}} \leq \Pr[\tau(x) > t] \quad .$$

Lemma 3. *[38, Lemma 4.10, Lemma 4.11 and Equation 4.29] Let $\{X_t\}_{t\geq 0}$ be a Markov chain with unique stationary distribution ν on state space Ω. Then for integers $t_1, t_2 \in \mathbb{N}$, we have*

- *if $t_2 \geq t_1$, then*

$$\sup_{X_0\in\Omega} \|\mathcal{L}(X_{t_2}) - \nu\|_{\mathrm{TV}} \leq 2 \sup_{X_0\in\Omega} \|\mathcal{L}(X_{t_1}) - \nu\|_{\mathrm{TV}} \quad .$$

- *if $t_2 = s \cdot t_1$ for some integer $s \in \mathbb{N}$, then*

$$\sup_{X_0\in\Omega} \|\mathcal{L}(X_{t_2}) - \nu\|_{\mathrm{TV}} \leq \left(2 \cdot \sup_{X_0\in\Omega} \|\mathcal{L}(X_{t_1}) - \nu\|_{\mathrm{TV}}\right)^s \quad .$$

We also utilize the following lemmas, which present upper bounds on the probability of a coordinate in a Haar random vector being small:

Lemma 4 (Lemma 3.5 in [44]). *Let $Y \sim \mu$ where μ is the Haar measure on $\mathcal{S}^n_{\mathbb{R}}$. Then for all $1 < c < \infty$ and any $1 \leq i \leq n$,*

$$\Pr\left[Y[i]^2 \leq n^{-3c}\right] \leq 2n^{1-c} \quad .$$

Lemma 5. *Let $Y \sim \mu_{\mathbb{C}}$ where $\mu_{\mathbb{C}}$ is the Haar measure on $\mathcal{S}^n_{\mathbb{C}}$. Then for all $1 < c < \infty$ and any $1 \leq i \leq n$,*

$$\Pr\left[|Y[i]|^2 \leq (2n)^{-3c}\right] \leq 2 \cdot (2n)^{1-c} \quad .$$

Proof. Let $g_1, \ldots, g_{2n}$ be $2n$ i.i.d. real random variable with $\mathcal{N}(0,1)$ distribution. We have

$$\Pr\left[|Y[i]|^2 \leq (2n)^{-3c}\right] = \Pr\left[\frac{g_1^2 + g_2^2}{\sum_{k=1}^{2n} g_k^2} \leq (2n)^{-3c}\right]$$
$$\leq \Pr\left[\frac{g_1^2}{\sum_{k=1}^{2n} g_k^2} \leq (2n)^{-3c}\right] \leq 2 \cdot (2n)^{1-c} \quad .$$

The last inequality follows from Lemma 4. □

2.3 Cryptography

In this section, we will review various definitions and results in cryptography. Throughout this work, λ denotes a security parameter.

Pseudorandom Functions and Pseudorandom Permutations

Definition 1 (Quantum-Secure Pseudorandom Function). *Let $\mathcal{K}, \mathcal{X}$ and $\mathcal{Y}$ be the key space, the domain and range, all implicitly depending on the security parameter λ. A keyed family of functions $\{\mathsf{PRF}_k : \mathcal{X} \to \mathcal{Y}\}_{k\in\mathcal{K}}$ is a quantum-secure pseudorandom function (QPRF) if the following two conditions hold:*

1. ***Efficient generation***. PRF_k *is polynomial-time computable on a classical computer.*
2. ***Pseudorandomness***. *For any polynomial-time quantum oracle algorithm* $\mathcal{A}$*,* PRF_k *with a random* $k \leftarrow \mathcal{K}$ *is indistinguishable from a truly random function* $f \leftarrow \mathcal{Y}^{\mathcal{X}}$ *in the sense that:*

$$\left| \Pr_{k \leftarrow \mathcal{K}} [\mathcal{A}^{\mathsf{PRF}_k}(1^\lambda) = 1] - \Pr_{f \leftarrow \mathcal{Y}^{\mathcal{X}}} [\mathcal{A}^{f}(1^\lambda) = 1] \right| = \mathrm{negl}(\lambda) \quad .$$

Definition 2 (Quantum-Secure Pseudorandom Permutation). *Let* $\mathcal{K}$ *be the key space, and* $\mathcal{X}$ *be both the domain and range, implicitly depending on the security parameter* λ*. A keyed family of permutations* $\{\mathsf{PRP}_k \in S_{\mathcal{X}}\}_{k \in \mathcal{K}}$ *is a quantum-secure pseudorandom permutation (*QPRP*) if the following two conditions hold:*

1. *(**Efficient generation**).* PRP_k *and* PRP_k^{-1} *are polynomial-time computable on a classical computer.*
2. *(**Pseudorandomness**). For any polynomial-time quantum oracle algorithm* $\mathcal{A}$*,* PRP_k *with a random* $k \leftarrow \mathcal{K}$ *is indistinguishable from a truly random permutation* $\sigma \leftarrow S_{\mathcal{X}}$ *in the sense that:*

$$\left| \Pr_{k \leftarrow \mathcal{K}} \left[\mathcal{A}^{\mathsf{PRP}_k, \mathsf{PRP}_k^{-1}}(1^\lambda) = 1\right] - \Pr_{\sigma \leftarrow S_{\mathcal{X}}} \left[\mathcal{A}^{\sigma, \sigma^{-1}}(1^\lambda) = 1\right] \right| = \mathrm{negl}(\lambda) \quad .$$

We adopt the definition of a strong quantum-secure PRP in this paper. And when referring to a quantum oracle algorithm having oracle access to a permutation σ, we imply that it has oracle access to both σ and its inverse σ^{-1}.

Under the assumption that post-quantum one-way functions exist, Zhandry proved the existence of QPRFs [50]. QPRPs can be constructed from QPRFs efficiently [49].

Given two QPRFs F and G, one independently samples F_{k_1} from F and G_{k_2} from G. A standard hybrid argument shows that F_{k_1}, G_{k_2} are computationally indistinguishable from two independent random functions, as stated in the following lemma. The proof, which is deferred to Appendix C, can be readily extended to the scenario when polynomially many pseudorandom primitives (or random primitives) are given.

Lemma 6. *Let keyed families of functions* $F : \mathcal{K}_1 \times \mathcal{X}_1 \to \mathcal{Y}_1$ *and* $G : \mathcal{K}_2 \times \mathcal{X}_2 \to \mathcal{Y}_2$ *be* QPRF*s. Then we have for any polynomial-time quantum oracle algorithm* $\mathcal{A}$*,*

$$\left| \Pr_{k_1 \leftarrow \mathcal{K}_1, k_2 \leftarrow \mathcal{K}_2} [\mathcal{A}^{F_{k_1}, G_{k_2}}(1^\lambda) = 1] - \Pr_{f \leftarrow \mathcal{Y}_1^{\mathcal{X}_1}, g \leftarrow \mathcal{Y}_2^{\mathcal{X}_2}} [\mathcal{A}^{f, g}(1^\lambda) = 1] \right| = \mathrm{negl}(\lambda) \quad .$$

It also holds if $\mathcal{X}_2 = \mathcal{Y}_2$*,* G *is a family of* QPRP*s and* $g \leftarrow \mathcal{Y}_2^{\mathcal{X}_2}$ *is replaced by* $g \leftarrow S_{\mathcal{X}_2}$*.*

Quantum Pseudorandomness The concept of quantum pseudorandom state generators was originally introduced in [31].

Definition 3 (Quantum Pseudorandom State Generator). *Let $\mathcal{K}$ be a key space and $\mathcal{H}$ be a Hilbert space. $\mathcal{K}$ and $\mathcal{H}$ depend on the security parameter λ. A pair of polynomial-time quantum algorithms (K, G) is a pseudorandom state generator (*PRSG*) if the following holds:*

- ***Key Generation.*** *$K(1^\lambda)$ chooses a uniform $k \in \mathcal{K}$ and outputs it as the key.*
- ***State Generation.*** *For all $k \in \mathcal{K}$, $G(1^\lambda, k)$ outputs a quantum state $|\phi_k\rangle \in \mathcal{S}(\mathcal{H})$.*
- ***Pseudorandomness.*** *Any polynomially many copies of $|\phi_k\rangle$ with the same random k is computationally indistinguishable from the same number of copies of a Haar random state. More precisely, for any $n \in \mathbb{N}$, any efficient quantum algorithm $\mathcal{A}$ and any $\ell \in \mathrm{poly}(\lambda)$,*

$$\left| \Pr_{k \leftarrow \mathcal{K}} \left[\mathcal{A}\Big(|\phi_k\rangle^{\otimes \ell} \Big) = 1 \right] - \Pr_{|\psi\rangle \leftarrow \mu} \left[\mathcal{A}\Big(|\psi\rangle^{\otimes \ell} \Big) = 1 \right] \right| = \mathrm{negl}(\lambda) \quad ,$$

where μ is the Haar measure on $\mathcal{S}(\mathcal{H})$.

*We call the keyed family of quantum states $\{\phi_k\}_{k \in \mathcal{K}}$ a pseudorandom quantum state (*PRS*) in $\mathcal{H}$.*

PRSGs exist assuming the existence of QPRFs. Given any QPRF $\mathsf{PRF} : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ (where $\mathcal{K}$ and $N = 2^n$ are implicitly functions of the security parameter λ), [31] constructed a PRS $\{\phi_k\}_{k \in \mathcal{K}}$, referred to *(pseudo)random phase states*, as follows:

$$|\phi_k\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \{0,1\}^n} \omega_N^{\mathsf{PRF}_k(x)} |x\rangle$$

for $k \in \mathcal{K}$ and $\omega_N = e^{\mathrm{i}\frac{2\pi}{N}}$. Additionally, they conjectured the variant with binary phase (i.e., replacing ω_N with -1) remains a PRS, and this was later confirmed in [10].

It is worth noting that both of these constructions rely on state generation algorithms that require a specific initial state, typically the all-zero state $|0\rangle^{\otimes n}$. If we were to use a different initial state, such as the equally weighted superposition state $|+\rangle^{\otimes n}$, their state generation algorithms would fail to produce a pseudorandom state. Therefore, the specific initial state is crucial for the success of these constructions.

3 Pseudorandom State Scramblers

We describe our new primitive *quantum Pseudorandom State Scramblers* (PRSS). A PRSS is capable of generating a pseudorandom state on an arbitrary initial state, addressing the limitation of acting on one specific initial state.

Definition 4 (Pseudorandom State Scrambler). *Let $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$ be Hilbert spaces of dimensions 2^n and 2^m respectively with $n, m \in \mathbb{N}$ and $n \leq m$. Let $\mathcal{K} = \{0,1\}^\kappa$ be a key space, and λ be a security parameter. A* pseudorandom state scrambler *(*PRSS*) is an ensemble of isometric operators*

$$\mathcal{R}^{n,m} := \{\{\mathcal{R}_k^{n,m,\lambda} : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}\}_{k \in \mathcal{K}}\}_\lambda ,$$

satisfying:

- ***Pseudorandomness***. *For any $\ell = \text{poly}(\lambda)$, any $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$, and any non-uniform poly-time quantum adversary $\mathcal{A}$,*

$$\left| \Pr_{k \leftarrow \mathcal{K}}\left[\mathcal{A}\Big(|\phi_k\rangle^{\otimes \ell}\Big) = 1\right] - \Pr_{|\psi\rangle \leftarrow \mu}\left[\mathcal{A}\Big(|\psi\rangle^{\otimes \ell}\Big) = 1\right] \right| = \text{negl}(\lambda) ,$$

  where $|\phi_k\rangle := \mathcal{R}_k^{n,m,\lambda} |\phi\rangle$ and μ is the Haar measure on $\mathcal{S}(\mathcal{H}_{\text{out}})$.
- ***Uniformity***. *$\mathcal{R}^{n,m}$ can be uniformly computed in polynomial time. That is, there is a deterministic Turing machine that, on input $(1^n, 1^m, 1^\lambda, 1^\kappa)$, outputs a quantum circuit Q in $\text{poly}(n, m, \lambda, \kappa)$ time such that for all $k \in \mathcal{K}$ and $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$*

$$Q\,|k\rangle|\phi\rangle = |k\rangle|\phi_k\rangle ,$$

  where $|\phi_k\rangle := \mathcal{R}_k^{n,m,\lambda}|\phi\rangle$.
- ***Polynomially-bounded key length***. *$\kappa = \log|\mathcal{K}| = \text{poly}(m, \lambda)$. As a result, $\mathcal{R}^{n,m}$ can be computed efficiently in time $\text{poly}(n, m, \lambda)$.*

By strengthening the pseudorandomness condition in PRSS, we define random state scramblers as follows.

Definition 5 (Random State Scrambler). *Let $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$ be Hilbert spaces of dimensions 2^n and 2^m respectively with $n, m \in \mathbb{N}$ and $n \leq m$. Let $\mathcal{K} = \{0,1\}^\kappa$ be a key space, and λ be a security parameter. A* random state scrambler *(*RSS*) is an ensemble of isometric operators $\mathcal{R}^{n,m} := \{\mathcal{R}^{n,m,\lambda}\}_\lambda$ with $\mathcal{R}^{n,m,\lambda} := \{\mathcal{R}_k^{n,m,\lambda} : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}\}_{k \in \mathcal{K}}$ satisfying:*

- ***Statistical Pseudorandomness***. *For any $\ell = \text{poly}(\lambda)$, and any $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$,*

$$\text{TD}\left(\mathop{\mathbb{E}}_{k \leftarrow \mathcal{K}}\left[|\phi_k\rangle\langle\phi_k|^{\otimes \ell}\right], \mathop{\mathbb{E}}_{|\psi\rangle \in \mu}\left[|\psi\rangle\langle\psi|^{\otimes \ell}\right] \right) = \text{negl}(\lambda) ,$$

  where $|\phi_k\rangle := \mathcal{R}_k^{n,m,\lambda} |\phi\rangle$ and μ is the Haar measure on $\mathcal{S}(\mathcal{H}_{\text{out}})$.
- ***Uniformity***. *$\mathcal{R}^{n,m}$ can be uniformly computed in polynomial time. That is, there is a deterministic Turing machine that, on input $(1^n, 1^m, 1^\lambda, 1^\kappa)$, outputs a quantum circuit Q in $\text{poly}(n, m, \lambda, \kappa)$ time such that for all $k \in \mathcal{K}$ and $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$*

$$Q\,|k\rangle|\phi\rangle = |k\rangle|\phi_k\rangle ,$$

  where $|\phi_k\rangle := \mathcal{R}_k^{n,m,\lambda}|\phi\rangle$.

Random	Pseudorandom	Main property
Haar unitary	PRU $\{U_k\}$	$\{U_k\} \approx_c$ Haar unitary
RSS	PRSS $\{\mathcal{R}_k\}$	$\forall \|\phi\rangle, \{\mathcal{R}_k \|\phi\rangle\} \approx$ Haar state (trace distance or comp. indist.)
Haar state	PRSG $\{\mathcal{R}_k\}$	for some *fixed* $\|\phi\rangle$ (e.g., $\|0\rangle$) $\{\mathcal{R}_k \|\phi\rangle\} \approx_c$ Haar state

Table 1: A collection of quantum random and pseodurandom objects.

3.1 Properties of Pseudorandom State Scramblers

We discuss basic characteristics of the new primitives, as well as their relationships with pseudorandom state generators and their siblings.

Unitary to isometry. It is sufficient to construct PRSSs from $\mathcal{H}$ to $\mathcal{H}$, since we can construct PRSSs from $\mathcal{H}_1$ to $\mathcal{H}_2$ $(n < m)$ in the following way. Let $\mathcal{R}^{m,m} := \{\mathcal{R}^{m,m,\lambda}\}_\lambda$ be a PRSS with $\mathcal{R}^{m,m,\lambda} := \{\mathcal{R}_k^{m,m,\lambda} : \mathcal{H}_2 \to \mathcal{H}_2\}$. For all $\lambda \in \mathbb{N}$ and $k \in \mathcal{K}$, we define $\mathcal{R}_k^{n,m,\lambda} = \mathcal{R}_k^{m,m,\lambda}\left(\mathbb{1} \otimes |0\rangle^{\otimes(m-n)}\right)$ where $\mathbb{1}$ is the identity of $\mathcal{H}_1$. It is not hard to verify that $\mathcal{R}^{n,m}$ is a PRSS from $\mathcal{H}_1$ to $\mathcal{H}_2$. We may write $\mathcal{R}^m$ instead of $\mathcal{R}^{m,m}$ when $m = n$.

Connections with Existing PRS *variants.* Several definitions of quantum pseudorandomness on states with slight variations have been proposed and constructed since the regular PRS has been introduced. Brakerski and Shmueli [11] introduced scalable pseudorandom states (scalable PRSs) to eliminate the dependence between the state size and the security parameter. This modification aids in assuring the security when the state size n is much smaller than the security parameter λ. Ananth, Qian and Yuen [5] introduced pseudorandom function-like states (PRFS s), which extend PRSs by augmenting with classical inputs alongside the secret key. Although the security is initially based on pre-selected classical queries to the PRFS generator, the subsequent work [4] relaxes this to allow adversaries making adaptive (classical or quantum) queries resulting in three levels of security. The following theorem states that PRSSs subsume the original PRSs and those variants. The proof is deferred to Appendix B.

Theorem 2. PRSG*s, scalable* PRSG*s, and* PRFSG*s can be constructed via invoking* PRSS*s in a black-box manner.*

Oracle Separation from OWF*s.* According to [36, Theorem 2], PRUs exist relative to a quantum oracle $\mathcal{O}$, even when $\mathsf{BQP}^{\mathcal{O}} = \mathsf{QMA}^{\mathcal{O}}$, indicating the non-existence of one-way functions. Since PRUs imply PRSSs, we obtain the same oracle separation result for PRSSs.

Theorem 3. *There exists a quantum oracle $\mathcal{O}$ relative to which* PRSS*s exist, but* $\mathsf{BQP}^{\mathcal{O}} = \mathsf{QMA}^{\mathcal{O}}$.

4 Parallel Kac's Walk

In this section, we design a *parallel version* of the standard Kac's walk on $\mathcal{S}_{\mathbb{R}}^n$ [44] and demonstrate that it mixes exponentially faster with respect to the metrics of our interest. We assume $n = 2m$ for some $m \in \mathbb{N}$ throughout this section.

4.1 Parallel Kac's Walk on Real Space

Before introducing our parallel Kac's walk, we first review the standard one. The standard Kac's walk on vectors within a real Hilbert space is a Markov process. At each discrete time t, we randomly select two coordinates (i, j) of the vector, and then apply a two-dimensional rotation to the corresponding subvector with an angle θ drawn randomly and uniformly. After a predetermined number of steps, the Markov chain converges to a Haar distribution over the unit sphere. It is proved in [44] that the mixing time of Kac's walk on $\mathcal{S}_{\mathbb{R}}^n$ with respect to the total variation distance is $\Theta(n \log n)$. The formal definition of Kac's walk is given below.

Definition 6. *Kac's walk on $\mathcal{S}_{\mathbb{R}}^n$ is a discrete-time Markov chain $\{X_t \in \mathcal{S}_{\mathbb{R}}^n\}_{t\geq 0}$. At each time t, two coordinates $i^{(t)}, j^{(t)} \in [n]$ and an angle $\theta^{(t)} \in [0, 2\pi)$ are chosen uniformly at random. X_{t+1} is obtained by the following update rules:*

$$\begin{pmatrix} X_{t+1}[i^{(t)}] \\ X_{t+1}[j^{(t)}] \end{pmatrix} = \begin{bmatrix} \cos(\theta^{(t)}) & -\sin(\theta^{(t)}) \\ \sin(\theta^{(t)}) & \cos(\theta^{(t)}) \end{bmatrix} \begin{pmatrix} X_t[i^{(t)}] \\ X_t[j^{(t)}] \end{pmatrix} ,$$

$$X_{t+1}[k] = X_t[k] \quad \text{for} \quad k \notin \left\{ i^{(t)}, j^{(t)} \right\} .$$

We denote the Kac's walk as $G : [n] \times [n] \times [0, 2\pi) \times \mathcal{S}_{\mathbb{R}}^n \to \mathcal{S}_{\mathbb{R}}^n$ such that

$$X_{t+1} = G\left(i^{(t)}, j^{(t)}, \theta^{(t)}, X_t\right) . \tag{1}$$

In our parallel Kac's walk, instead of randomly rotating one subvector, we simultaneously rotate m subvectors. Here we give its formal definition.

Definition 7. *The parallel Kac's walk is a discrete-time Markov chain $\{X_t \in \mathcal{S}_{\mathbb{R}}^n\}_{t\geq 0}$. At each step t, the parallel Kac's walk first selects a random perfect matching of the set $\{1, \ldots, n\}$, denoted by*

$$P_t = \left\{ \left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right) \right\} ,$$

where $\bigcup_{k=1}^m \left\{ i_k^{(t)}, j_k^{(t)} \right\} = \{1, \ldots, n\}$. Then m independent angles $\theta_1^{(t)}, \ldots, \theta_m^{(t)} \in [0, 2\pi)$ are chosen uniformly at random. For every pair $\left(i_k^{(t)}, j_k^{(t)}\right)$ in P_t, it sets

$$\begin{pmatrix} X_{t+1}[i_k^{(t)}] \\ X_{t+1}[j_k^{(t)}] \end{pmatrix} = \begin{bmatrix} \cos(\theta_k^{(t)}) & -\sin(\theta_k^{(t)}) \\ \sin(\theta_k^{(t)}) & \cos(\theta_k^{(t)}) \end{bmatrix} \begin{pmatrix} X_t[i_k^{(t)}] \\ X_t[j_k^{(t)}] \end{pmatrix} .$$

Let $F : ([n] \times [n])^m \times [0, 2\pi)^m \times \mathcal{S}_{\mathbb{R}}^n \to \mathcal{S}_{\mathbb{R}}^n$ *denote the map associated with the above random walk such that*

$$X_{t+1} = F\Big(P_t, \theta_1^{(t)}, \ldots, \theta_m^{(t)}, X_t\Big) \quad . \tag{2}$$

In one step of the parallel Kac's walk, we obtain m distinct coordinate pairs by randomly sampling a perfect matching P_t of set $[n]$. For each pair, a rotation angle is selected independently and uniformly at random. Recall the notation in Definition 6. Let $X_{t,1} = X_t$ and $X_{t,k+1} = G\Big(i_k^{(t)}, j_k^{(t)}, \theta_k^{(t)}, X_{t,k}\Big)$ for $1 \le k \le m$. It is evident that

$$X_{t,m+1} = X_{t+1} = F\Big(P_t, \theta_1^{(t)}, \ldots, \theta_m^{(t)}, X_t\Big) \quad .$$

We can observe that taking one step of the parallel Kac's walk can be viewed as taking $m = n/2$ steps in the original Kac's walk when there are *no collisions* in the pairing step. All the subvectors being rotated in a single step of the parallel Kac's walk are distinct, and thus not independent. Consequently, the results for the original Kac's walk cannot be directly applied. Fortunately, by enhancing the coupling technique for analyzing the mixing time of the standard Kac's walk, we are able to prove that the parallel Kac's walk rapidly mixes in time $O(\log n)$ with respect to two different metrics: (1) the Wasserstein 1-distance; and (2) the total variation distance.

In the context of the Wasserstein 1-distance, after walking T steps, the difference between the output distribution of a parallel Kac's walk and the normalized Haar measure decays exponentially as T grows, which leads to a $O(\log n)$ mixing time. Formally,

Theorem 4. *Let* $\{X_t \in \mathcal{S}_{\mathbb{R}}^n\}_{t\ge 0}$ *be a Markov chain that evolves according to the parallel Kac's walk. Then, for sufficiently large* n*,* $c > 0$*, and* $T = 10(c+1)\log n$*,*

$$\sup_{X_0 \in \mathcal{S}_{\mathbb{R}}^n} W_1(\mathcal{L}(X_T), \mu) \le \frac{1}{2^{c \log n}} \quad ,$$

where μ *is the normalized Haar measure on* $\mathcal{S}_{\mathbb{R}}^n$*.*

Furthermore, we get a stronger result regarding the total variation distance:

Theorem 5. *Let* $\{X_t \in \mathcal{S}_{\mathbb{R}}^n\}_{t\ge 0}$ *be a Markov chain that evolves according to the parallel Kac's walk. Then, for sufficiently large* n*,* $c > 515$ *and* $T = c\log n$*,*

$$\sup_{X_0 \in \mathcal{S}_{\mathbb{R}}^n} \|\mathcal{L}(X_T) - \mu\|_{\mathrm{TV}} \le \frac{1}{2^{(c/515-1)\log n - 1}} \quad ,$$

where μ *is the normalized Haar measure on* $\mathcal{S}_{\mathbb{R}}^n$*.*

Notably, while the Wasserstein 1-distance is a weaker metric compared to the total variation distance, Theorem 4 provides an adequate foundation for constructing a PRSS. Additionally, the analysis of Theorem 5 further reveals a *dispersing property* of our construction of RSS. The remainder of this section is devoted to proving Theorem 4. The proof for Theorem 5, along with an explanation of the dispersing property, is deferred to Appendix A.

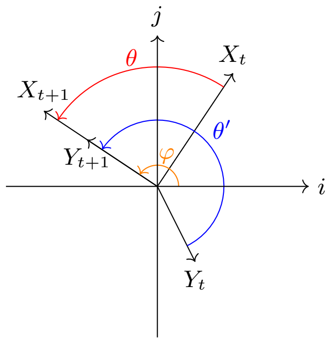

Fig. 1: Transformation of subcoordinates $X_t[i,j]$ and $Y_t[i,j]$

The Proportional Coupling. Our technique for proving the mixing time in Theorem 4 accommodates the *proportional coupling* [44] that sufficiently reduces the distance between two copies of Kac's walk. At each time t in the proportional coupling (illustrated in Figure 1), an angle θ is chosen uniformly at random from $[0, 2\pi)$ for rotating the subvector $(X_t[i], X_t[j])$, where indices i and j are picked as in Definition 6. The angle θ' is specifically selected for $(Y_t[i], Y_t[j])$ to make it collinear with $(X_t[i], X_t[j])$, i.e., they share the same argument φ. Taking into account the marginal distribution, both θ and θ' are drawn from the uniform distribution over the interval $[0, 2\pi)$, validating the proportional coupling for two Kac's walks.

Following a similar idea, we define the proportional coupling of two copies of the parallel Kac's walk, which couples each pair of indices from the randomly sampled perfect matching using the propositional coupling.

Definition 8 (Proportional Coupling for the Parallel Kac's Walk). *We define a coupling of two copies $\{X_t\}_{t\geq 0}, \{Y_t\}_{t\geq 0}$ of the parallel Kac's walk in the following way: Fix X_t, $Y_t \in \mathcal{S}^n_{\mathbb{R}}$.*

1. *Choose a perfect matching $P_t = \left\{\left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right)\right\}$ and m angles $\theta_1^{(t)}, \ldots, \theta_m^{(t)} \in [0, 2\pi)$ uniformly at random, and set*

$$X_{t+1} = F\left(P_t, \theta_1^{(t)}, \ldots, \theta_m^{(t)}, X_t\right) \quad .$$

2. *Sample m angles ${\theta'}_1^{(t)}, \ldots, {\theta'}_m^{(t)}$ in the following manner: for every $1 \leq k \leq m$,*
   (a) choose $\varphi_k \in [0, 2\pi)$ uniformly at random among all angles that satisfy

$$X_{t+1}[i_k^{(t)}] = \sqrt{X_t[i_k^{(t)}]^2 + X_t[j_k^{(t)}]^2}\cos(\varphi_k) \quad ,$$
$$X_{t+1}[j_k^{(t)}] = \sqrt{X_t[i_k^{(t)}]^2 + X_t[j_k^{(t)}]^2}\sin(\varphi_k) \quad ,$$

(b) and then choose ${\theta'}_k^{(t)} \in [0, 2\pi)$ *uniformly among the angles that satisfy*

$$\cos({\theta'}_k^{(t)}) \cdot Y_t[i_k^{(t)}] - \sin({\theta'}_k^{(t)}) \cdot Y_t[j_k^{(t)}] = \sqrt{Y_t[i_k^{(t)}]^2 + Y_t[j_k^{(t)}]^2} \cos(\varphi_k) \ ,$$
$$\sin({\theta'}_k^{(t)}) \cdot Y_t[i_k^{(t)}] + \cos({\theta'}_k^{(t)}) \cdot Y_t[j_k^{(t)}] = \sqrt{Y_t[i_k^{(t)}]^2 + Y_t[j_k^{(t)}]^2} \sin(\varphi_k) \ .$$

And set $Y_{t+1} = F\left(P_t, {\theta'}_1^{(t)}, \ldots, {\theta'}_m^{(t)}, Y_t\right)$.

In this coupling scheme, we enforce X_t and Y_t to employ an identical random matching (P_t in step 1) to generate all the m pairs of coordinates. And then we sample m rotation angles for X_t and obtain X_{t+1} by rotating the m coordinate pairs by their corresponding angles. Next, in step 2, we determine the rotation angle for each coordinate pair of Y_t. For the k-th pair, our objective is to select a suitable angle ${\theta'}_k^{(t)}$ such that the two-dimensional subvector $(Y_{t+1}[i_k^{(t)}], Y_{t+1}[j_k^{(t)}])$ aligns collinearly with $(X_{t+1}[i_k^{(t)}], X_{t+1}[j_k^{(t)}])$. To achieve this, we ensure that $(Y_{t+1}[i_k^{(t)}], Y_{t+1}[j_k^{(t)}])$ shares the same argument φ_k as $(X_{t+1}[i_k^{(t)}], X_{t+1}[j_k^{(t)}])$. Typically, the values of angles φ_k and ${\theta'}_k^{(t)}$ are uniquely determined. However, in the scenario where either $(X_t[i_k^{(t)}], X_t[j_k^{(t)}])$ or $(Y_t[i_k^{(t)}], Y_t[j_k^{(t)}])$ equals the zero vector, all angles satisfy the required conditions. In such cases, we resort to uniform random selection for determining the angles.

Remark 1. This coupling forces $X_{t+1}[i]Y_{t+1}[i] \geq 0$ for all $i \in [n]$ since the signs are determined by the same arguments.

In each step of our coupling scheme, a quarter of the distance between vectors X_t and Y_t is reduced, which is formally shown in

Lemma 7. *Let* $X_0, Y_0 \in \mathcal{S}_{\mathbb{R}}^n$*. For* $t \geq 0$*, we couple* (X_{t+1}, Y_{t+1}) *conditioned on* (X_t, Y_t) *according to the proportional coupling defined in Definition 8. We define*

$$A_t[i] = X_t[i]^2 \ , \quad B_t[i] = Y_t[i]^2 \ .$$

Then for any $l \in \mathbb{N}$*, we have*

$$\mathbb{E}\left[\sum_{i=1}^{n} (A_l[i] - B_l[i])^2\right] \leq 2 \cdot \left(1 - \frac{1}{4}\right)^l \ .$$

Proof. Fix $X_t, Y_t \in \mathcal{S}_{\mathbb{R}}^n$. Let (X_{t+1}, Y_{t+1}) obtained from (X_t, Y_t) by applying the coupling defined in Definition 8. Recall that $n = 2m$. Let $N = \frac{n!}{2^m m!}$ be the number of perfect matchings for $[n]$. To keep the notations short, the perfect matching $\left\{\left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right)\right\}$ at step t is denoted by $\left(\overrightarrow{i^{(t)}}, \overrightarrow{j^{(t)}}\right)$.

We have

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n}\left(A_{t+1}[i]-B_{t+1}[i]\right)^2\right] \\
=&\frac{1}{N}\sum_{\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)}\underbrace{\mathbb{E}\left[\sum_{i=1}^{n}\left(A_{t+1}[i]-B_{t+1}[i]\right)^2\middle| P_t=\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)\right]}_{(\star)}\quad .
\end{aligned}
\tag{3}
$$

By the definition of the parallel Kac's walk, we have

$$
\begin{aligned}
(\star)=&\sum_{k=1}^{m}\mathbb{E}\left[\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)\cos(\varphi_k)^2-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\cos(\varphi_k)^2\right)^2\right]\\
&+\sum_{k=1}^{m}\mathbb{E}\left[\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)\sin(\varphi_k)^2-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\sin(\varphi_k)^2\right)^2\right]\\
=&\frac{3}{4}\sum_{k=1}^{m}\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\right)^2\\
=&\underbrace{\frac{3}{4}\sum_{k=1}^{m}\left(\left(A_t[i_k^{(t)}]-B_t[i_k^{(t)}]\right)^2+\left(A_t[j_k^{(t)}]-B_t[j_k^{(t)}]\right)^2\right)}_{(\star\star)}\\
&+\underbrace{\frac{3}{4}\sum_{k=1}^{m}2\left(A_t[i_k^{(t)}]-B_t[i_k^{(t)}]\right)\left(A_t[j_k^{(t)}]-B_t[j_k^{(t)}]\right)}_{(\star\star\star)}\ ,
\end{aligned}
\tag{4}
$$

where the second equality is by $\mathbb{E}\left[\cos(\varphi_k)^4\right]=\mathbb{E}\left[\sin(\varphi_k)^4\right]=3/8$.

As $\left\{\left(i_1^{(t)},j_1^{(t)}\right),\ldots,\left(i_m^{(t)},j_m^{(t)}\right)\right\}$ is a perfect matching, we have

$$
(\star\star)=\frac{3}{4}\sum_{i=1}^{n}\left(A_t[i]-B_t[i]\right)^2\quad . \tag{5}
$$

Combing Eqs. (3)(4)(5), we obtain

$$
\mathbb{E}\left[\sum_{i=1}^{n}\left(A_{t+1}[i]-B_{t+1}[i]\right)^2\right]=\frac{3}{4}\sum_{i=1}^{n}\left(A_t[i]-B_t[i]\right)^2+\underbrace{\frac{1}{N}\sum_{\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)}(\star\star\star)}_{(4\star)}\quad . \tag{6}
$$

For the last term,

$$
\begin{aligned}
(4\star) =& \frac{3}{2N} \sum_{\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)} \sum_{k=1}^{m} \left(A_t[i_k^{(t)}] - B_t[i_k^{(t)}]\right)\left(A_t[j_k^{(t)}] - B_t[j_k^{(t)}]\right) \\
=& \frac{3}{2N} \cdot \frac{(n-2)!}{2^{m-1}(m-1)!} \sum_{i<j} (A_t[i] - B_t[i])\,(A_t[j] - B_t[j]) \\
=& \frac{3 \cdot m}{2n(n-1)} \left(\left(\sum_{i=1}^{n} (A_t[i] - B_t[i]) \right)^2 - \sum_{i=1}^{n} (A_t[i] - B_t[i])^2 \right) \\
=& -\frac{3}{4(n-1)} \sum_{i=1}^{n} (A_t[i] - B_t[i])^2 \quad .
\end{aligned} \tag{7}
$$

Combining Eqs. (6)(7), we have

$$
\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n} (A_l[i] - B_l[i])^2\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{n} (A_l[i] - B_l[i])^2 \,\middle|\, X_{l-1}, Y_{l-1}\right]\right] \\
&\le \frac{3}{4}\,\mathbb{E}\left[\sum_{i=1}^{n} (A_{l-1}[i] - B_{l-1}[i])^2\right] \\
&\le \left(\frac{3}{4}\right)^l \sum_{i=1}^{n} (A_0[i] - B_0[i])^2 \le 2 \cdot \left(\frac{3}{4}\right)^l \quad .
\end{aligned}
$$

□

Proof of Theorem 4. Let $T = 10(c+1)\log n$ for $c > 0$. We couple two copies $\{X_t\}_{t\ge 0}$ and $\{Y_t\}_{t\ge 0}$ of the parallel Kac's walk with starting points $X_0 = x \in \mathcal{S}_{\mathbb{R}}^n$ and $Y_0 \sim \mu$, by applying the proportional coupling. We have

$$
W_1(\mathcal{L}(X_T)\,,\mu) = W_1(\mathcal{L}(X_T)\,,\mathcal{L}(Y_T)) \le \mathbb{E}[\|X_T - Y_T\|_2] \le \left(\mathbb{E}\left[\|X_T - Y_T\|_2^4\right]\right)^{1/4} \quad .
$$

Then by Cauchy-Schwarz inequality, we have

$$
W_1(\mathcal{L}(X_T)\,,\mu) \le \left(n\,\mathbb{E}\left[\|X_T - Y_T\|_4^4\right]\right)^{1/4} \quad . \tag{8}
$$

Note that the proportional coupling forces $X_T[i]Y_T[i] \ge 0$ for all $i \in [n]$. Therefore, for all $i \in [n]$

$$
|X_T[i] - Y_T[i]| \le |X_T[i] + Y_T[i]| \quad .
$$

This gives us

$$
\|X_T - Y_T\|_4^4 = \sum_{i=1}^{n} (X_T[i] - Y_T[i])^4 \le \sum_{i=1}^{n} \left(X_T[i]^2 - Y_T[i]^2\right)^2 \quad . \tag{9}
$$

Combing Eqs. (8) and (9), we have

$$W_1(\mathcal{L}(X_T), \mu) \leq \left(n\,\mathbb{E}\left[\sum_{i=1}^{n} \left(X_T[i]^2 - Y_T[i]^2 \right)^2 \right] \right)^{1/4}$$

$$(\text{Lemma 7}) \leq \left(2n \left(\frac{3}{4} \right)^T \right)^{1/4} \leq \frac{1}{2^{c \log n}} \ .$$

□

4.2 Parallel Kac's Walk on Complex Space

In this section, we extend the parallel Kac's walk to complex vectors. In the real case, for each pair of coordinates, we uniformly select a matrix according to Haar measure on SO(2). In the complex case, we will naturally choose a matrix from SU(2) according to Haar measure on it. The Haar random unitary in SU(2) can be obtained by sampling three random angles [51]. Let

$$U(\alpha, \beta, \theta) = \begin{pmatrix} e^{\mathrm{i}\alpha}\cos(\theta) & -e^{\mathrm{i}\beta}\sin(\theta) \\ e^{-\mathrm{i}\beta}\sin(\theta) & e^{-\mathrm{i}\alpha}\cos(\theta) \end{pmatrix} \ . \tag{10}$$

If we pick $\alpha, \beta \in [0, 2\pi)$ and $\zeta \in [0, 1)$ uniformly at random and set $\theta = \arcsin\sqrt{\zeta}$, then $U(\alpha, \beta, \theta)$ is a Haar random unitary on SU(2).

Kac's walk on complex vectors We define Kac's walk on $\mathcal{S}^n_{\mathbb{C}}$ as a discrete-time Markov chain $\{X_t \in \mathcal{S}^n_{\mathbb{C}}\}_{t\geq 0}$. At each time t, two coordinates $i^{(t)}, j^{(t)} \in \{1, \ldots, n\}$ and two angles $\alpha^{(t)}, \beta^{(t)} \in [0, 2\pi)$ are chosen uniformly at random. Additionally, a real number $\zeta^{(t)} \in [0, 1)$ is selected uniformly at random and compute

$$\theta^{(t)} = \arcsin\sqrt{\zeta^{(t)}} \ .$$

X_{t+1} is obtained by the following update rules:

$$\begin{pmatrix} X_{t+1}[i^{(t)}] \\ X_{t+1}[j^{(t)}] \end{pmatrix} = U\left(\alpha^{(t)}, \beta^{(t)}, \theta^{(t)}\right) \begin{pmatrix} X_t[i^{(t)}] \\ X_t[j^{(t)}] \end{pmatrix} \ ,$$

$$X_{t+1}[k] = X_t[k] \quad \text{for} \quad k \notin \left\{ i^{(t)}, j^{(t)} \right\} \ .$$

We denote the Kac's walk on complex vectors as $G_{\mathbb{C}} : [n] \times [n] \times [0, 2\pi)^3 \times \mathcal{S}^n_{\mathbb{C}} \to \mathcal{S}^n_{\mathbb{C}}$ such that

$$X_{t+1} = G_{\mathbb{C}}\left(i^{(t)}, j^{(t)}, \alpha^{(t)}, \beta^{(t)}, \theta^{(t)}, X_t \right) \ .$$

Parallel Kac's walk on complex vectors In a parallel Kac's walk, we choose a perfect matching at each step and apply the one-step Kac's walk $G_{\mathbb{C}}$ on each pair. More specifically, the parallel Kac's walk on complex vectors is a discrete-time Markov chain $\{X_t \in \mathcal{S}_{\mathbb{C}}^n\}_{t\geq 0}$. At each step t, it first selects a random perfect matching of the set $[n]$, denoted by

$$P_t = \left\{\left(i_1^{(t)}, j_1^{(t)}\right), \dots, \left(i_m^{(t)}, j_m^{(t)}\right)\right\}$$

where $\bigcup_{k=1}^m \left\{i_k^{(t)}, j_k^{(t)}\right\} = [n]$. And then $2m$ independent angles

$$\alpha_1^{(t)}, \dots, \alpha_m^{(t)}, \beta_1^{(t)}, \dots, \beta_m^{(t)} \in [0, 2\pi)$$

are chosen uniformly at random. Additionally, m independent real numbers $\zeta_1^{(t)} \dots, \zeta_m^{(t)} \in [0,1)$ are selected uniformly at random and compute

$$\theta_k^{(t)} = \arcsin\left(\sqrt{\zeta_k^{(t)}}\right)$$

for all $k \in \{1, \dots, m\}$. Then for every pair $\left(i_k^{(t)}, j_k^{(t)}\right)$ in P_t, it sets

$$\begin{pmatrix} X_{t+1}[i_k^{(t)}] \\ X_{t+1}[j_k^{(t)}] \end{pmatrix} = U\left(\alpha_k^{(t)}, \beta_k^{(t)}, \theta_k^{(t)}\right) \begin{pmatrix} X_t[i_k^{(t)}] \\ X_t[j_k^{(t)}] \end{pmatrix} \ .$$

Let $F_{\mathbb{C}} : ([n] \times [n])^m \times [0, 2\pi)^{3m} \times \mathcal{S}_{\mathbb{C}}^n \to \mathcal{S}_{\mathbb{C}}^n$ denote the map associated with the above random walk such that

$$X_{t+1} = F_{\mathbb{C}}\left(P_t, \left\{\alpha_k^{(t)}\right\}_{k=1}^m, \left\{\beta_k^{(t)}\right\}_{k=1}^m, \left\{\theta_k^{(t)}\right\}_{k=1}^m, X_t\right) \ .$$

As the number of steps increases, the output distribution of the parallel Kac's walk on complex vectors converges exponentially fast to the Haar measure in terms of Wasserstein-1 distance and total variation distance. Formally,

Theorem 6. *Let $\{X_t \in \mathcal{S}_{\mathbb{C}}^n\}_{t\geq 0}$ be a Markov chain that evolves according to the parallel Kac's walk on complex vectors. Then, for sufficiently large n, $c > 0$ and $T = 10(c+1)\log n$,*

$$\sup_{X_0 \in \mathcal{S}_{\mathbb{C}}^n} W_1(\mathcal{L}(X_T), \mu) \leq \frac{1}{2^{c \log n}} \ ,$$

where $\mu_{\mathbb{C}}$ is the Haar measure on $\mathcal{S}_{\mathbb{C}}^n$.

Theorem 7. *Let $\{X_t \in \mathcal{S}_{\mathbb{C}}^n\}_{t\geq 0}$ be a Markov chain that evolves according to the parallel Kac's walk on complex vectors. Then, for sufficiently large n, $c > 515$ and $T = c\log n$,*

$$\sup_{X_0 \in \mathcal{S}_{\mathbb{C}}^n} \|\mathcal{L}(X_T) - \mu\|_{\mathrm{TV}} \leq \frac{1}{2^{(c/515-1)\log n - 1}} \ ,$$

where $\mu_{\mathbb{C}}$ is the Haar measure on $\mathcal{S}_{\mathbb{C}}^n$.

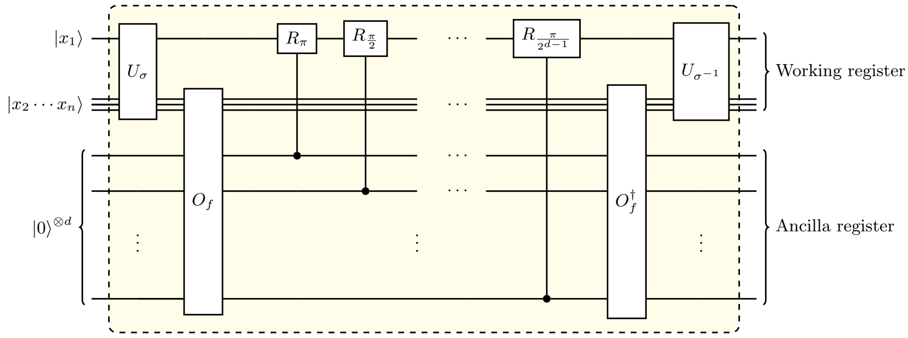

Fig. 2: Circuit diagram for the construction of the $K_{\sigma,f}$

The proofs of above theorems follow a similar line of reasoning as the proof in the real case. Therefore, to avoid redundancy, we defer the complete proof to Appendix C.

5 Constructions of RSSs and PRSSs

In this section, we present a family of circuits that implements RSSs, specifically realizing the parallel Kac's walk on the real (complex) unit sphere. To obtain circuits for PRSSs, one can simply replace the random primitives with their post-quantum secured pseudorandom counterparts.

5.1 Constructing (P)RSS over the Real Space

We begin by constructing a unitary gate that simulates a single step of the parallel Kac's walk on $\mathcal{S}_{\mathbb{R}}^{2^n}$. In every step, we denote the corresponding permutation by $\sigma \in S_{2^n}$. And we use the function $f : \{0,1\}^{n-1} \to \{0,1\}^d$ to manage the precision of the rotation angle that was originally chosen from the interval $[0, 2\pi)$, where d is the parameter controlling the precision of the rotation angle. Specifically, for every σ and f, we define a unitary gate $K_{\sigma,f} = U_{\sigma^{-1}} W_f U_\sigma$, where

$$U_\sigma = \sum_{x \in \{0,1\}^n} |\sigma(x)\rangle\langle x| \,, \quad W_f = \sum_{y \in \{0,1\}^{n-1}} \begin{pmatrix} \cos(\theta_y) & -\sin(\theta_y) \\ \sin(\theta_y) & \cos(\theta_y) \end{pmatrix} \otimes |y\rangle\langle y| \quad , \tag{11}$$

and $\theta_y = 2\pi \cdot \mathrm{val}(f(y))$ is the rotation angle for every subvector $\big(\sigma^{-1}(0y), \sigma^{-1}(1y)\big)$, $y \in \{0,1\}^{n-1}$. In Figure 2, we show a quantum circuit that realizes $K_{\sigma,f}$.

The circuit consists of:

1. Permutation: a unitary U_σ which transforms $|x\rangle$ to $|\sigma(x)\rangle$ for any $x \in \{0,1\}^n$. This unitary can be implemented via making quires to oracles O_σ and $O_{\sigma^{-1}}$,

and using n ancilla qubits: for any $x \in \{0,1\}^n$,

$$|x\rangle|0\rangle \xrightarrow{O_\sigma} |x\rangle|\sigma(x)\rangle \xrightarrow{SWAP} |\sigma(x)\rangle|x\rangle \xrightarrow{O_{\sigma^{-1}}} |\sigma(x)\rangle|0\rangle \ .$$

We omit this detail in the above figure for the sake of conciseness.

2. Implementing rotation operator W_f:
   (a) an oracle O_f which queries $f(x_2, \ldots, x_n)$ and stores the d-bit result in the ancilla qubits.
   (b) d controlled-rotation gates. The i-th ancilla qubit controls $R_{\frac{\pi}{2^{i-1}}}$ gate acting on the first qubit, where the gate R_θ denotes the rotation transformation $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.
   (c) an oracle O_f again for uncomputing the ancilla qubits.
3. Inverse permutation: a unitary $U_{\sigma^{-1}}$.

Remark. The gate $K_{\sigma,f}$ approximates one step of the parallel Kac's walk. It starts by partitioning the computational basis (indices) into 2^{n-1} pairs based on a selected permutation σ. For each pair $\left(\sigma^{-1}(0y), \sigma^{-1}(1y)\right)$ labeled by $y \in \{0,1\}^{n-1}$, the gate applies a rotation with an approximated angle θ_y indicated by f to the corresponding two dimensional subvector.

Stepwise State Evolution To gain insight into the functionality of $K_{\sigma,f}$, we assume the initial state to be a pure state

$$|\varphi\rangle = \sum_{x\in\{0,1\}^n} p_x |x\rangle \ .$$

First, to pair up the indices by applying U_σ, the initial state is transformed into

$$\begin{aligned}\sum_{x\in\{0,1\}^n} p_x |\sigma(x)\rangle \otimes \left|0^d\right\rangle &= \sum_{x'\in\{0,1\}^n} p_{\sigma^{-1}(x')} |x'\rangle \otimes \left|0^d\right\rangle \\ &= \sum_{y\in\{0,1\}^{n-1}} \left(p_{\sigma^{-1}(0y)}|0\rangle + p_{\sigma^{-1}(1y)}|1\rangle\right) \otimes |y\rangle \otimes \left|0^d\right\rangle \ .\end{aligned}$$

Then, to rotate each subvector, the oracle O_f stores $f(y)$ in the ancilla register as control qubits, resulting in the state

$$\sum_{y\in\{0,1\}^{n-1}} \left(p_{\sigma^{-1}(0y)}|0\rangle + p_{\sigma^{-1}(1y)}|1\rangle\right) \otimes |y\rangle \otimes |f(y)\rangle \ .$$

Next, a series of controlled-rotation gates are applied to the first qubit, rotating it by an angle of $\theta_y = 2\pi \cdot \mathrm{val}(f(y))$. Therefore, we have the following state:

$$\sum_{y\in\{0,1\}^{n-1}} \left(p'_{\sigma^{-1}(0y)}|0\rangle + p'_{\sigma^{-1}(1y)}|1\rangle\right) \otimes |y\rangle \otimes |f(y)\rangle$$

where

$$p'_{\sigma^{-1}(0y)} = \cos(\theta_y) \cdot p_{\sigma^{-1}(0y)} - \sin(\theta_y) \cdot p_{\sigma^{-1}(1y)} \ ,$$
$$p'_{\sigma^{-1}(1y)} = \sin(\theta_y) \cdot p_{\sigma^{-1}(1y)} + \cos(\theta_y) \cdot p_{\sigma^{-1}(1y)} \ .$$

After reverting the ancilla qubits and applying the inverse permutation, we obtain the output state

$$\sum_{y\in\{0,1\}^{n-1}} \left(p'_{\sigma^{-1}(0y)} \left|\sigma^{-1}(0y)\right\rangle + p'_{\sigma^{-1}(1y)} \left|\sigma^{-1}(1y)\right\rangle\right) \otimes \left|0^d\right\rangle = \sum_{x\in\{0,1\}^n} p'_x \left|x\right\rangle \otimes \left|0^d\right\rangle \ .$$

Constructing RSSs We first define an ensemble RSG^n of unitary operators that represents applying $K_{\sigma,f}$ for T-step with i.i.d. random selections of permutations and functions. Then, we prove that such an ensemble forms an RSS.

Definition 9. *Let $n, T, d \in \mathbb{N}$, and $\mathcal{H}$ be a real Hilbert space with dimension 2^n. An ensemble of unitary operators $\mathsf{RSG}^n := \left\{\mathsf{RSG}^{n,\lambda}\right\}_\lambda$ with*

$$\mathsf{RSG}^{n,\lambda} := \left\{\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T} : \mathcal{H} \to \mathcal{H}\right\}_{(\sigma_i)_{i=1}^T \subseteq S_{2^n}, (f_i)_{i=1}^T \subseteq \{f:\{0,1\}^{n-1}\to\{0,1\}^d\}}$$

is define as

$$\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T} = K_{\sigma_T,f_T} \cdots K_{\sigma_2,f_2} K_{\sigma_1,f_1}$$

where $K_{\sigma,f} = U_{\sigma^{-1}} W_f U_\sigma$ is defined in (11).

Theorem 8. *Let $n \in \mathbb{N}$, $d = \log^2\lambda + \log^2 n$ and $T = 10(\lambda+1)n$. The ensemble of unitary operators RSG^n defined in Definition 9 is an RSS.*

To prove Theorem 8, we define a new ensemble of (infinitely many) unitary operators $\widetilde{\mathsf{RSG}}^n := \left\{\widetilde{\mathsf{RSG}}^{n,\lambda}\right\}_\lambda$ with

$$\widetilde{\mathsf{RSG}}^{n,\lambda} := \left\{\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T} : \mathcal{H} \to \mathcal{H}\right\}_{(\sigma_i)_{i=1}^T \subseteq S_{2^n}, (\widetilde{f}_i)_{i=1}^T \subseteq \{f:\{0,1\}^{n-1}\to[0,1)\}}$$

and

$$\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T} = \widetilde{K}_{\sigma_T,\widetilde{f}_T} \cdots \widetilde{K}_{\sigma_2,\widetilde{f}_2} \widetilde{K}_{\sigma_1,\widetilde{f}_1}$$

where $\widetilde{K}_{\sigma,\widetilde{f}} = U_{\sigma^{-1}} \widetilde{W}_{\widetilde{f}} U_\sigma$ and $\widetilde{W}_{\widetilde{f}}$ is defined to be

$$\widetilde{W}_{\widetilde{f}} = \sum_{y\in\{0,1\}^{n-1}} \begin{pmatrix} \cos\left(\widetilde{\theta}_y\right) & -\sin\left(\widetilde{\theta}_y\right) \\ \sin\left(\widetilde{\theta}_y\right) & \cos\left(\widetilde{\theta}_y\right) \end{pmatrix} \otimes |y\rangle\langle y| \ , \tag{12}$$

in which $\widetilde{\theta}_y = 2\pi \cdot \widetilde{f}(y)$ for $y \in \{0,1\}^{n-1}$.

RSG^n and $\widetilde{\mathsf{RSG}}^n$ differ in the way the angles are chosen. In RSG^n, the angles are selected from the discrete set $\left\{2\pi \cdot \frac{i}{2^d} : i \in \left\{0, 1, \ldots, 2^d - 1\right\}\right\}$, while in

$\widetilde{\mathsf{RSG}}^n$, the angles are chosen from the interval $[0, 2\pi)$. For uniformly random σ and $\widetilde{f}$, applying gate $\widetilde{K}_{\sigma,\widetilde{f}}$ results in the selection of a random matching on the computational basis, with each pair in the matching being rotated by a random angle in $[0, 2\pi)$ determined by the corresponding value of $\widetilde{f}$. This is exactly one step of parallel Kac's walk described in Section 4. $\widetilde{\mathsf{RSG}}^n$ serves as an intermediate scrambler in the proof of Theorem 8. To analyse the difference between RSG^n and $\widetilde{\mathsf{RSG}}^n$, we need the following lemma.

Lemma 8. *Let $\sigma \in S_{2^n}$ and $\widetilde{f} : \{0,1\}^{n-1} \to [0,1)$. Let f be the function satisfying for any $y \in \{0,1\}^{n-1}$, $f(y)$ is the d digits after the binary point in $\widetilde{f}(y)$. Then*

$$\left\| K_{\sigma,f} - \widetilde{K}_{\sigma,\widetilde{f}} \right\|_\infty \le 2^{1-d}\pi \ ,$$

where $K_{\sigma,f} = U_{\sigma^{-1}} W_f U_\sigma$ is defined in (11) and $\widetilde{K}_{\sigma,\widetilde{f}} = U_{\sigma^{-1}} \widetilde{W}_{\widetilde{f}} U_\sigma$ is defined in (12).

Proof.

$$\begin{aligned}
\left\| K_{\sigma,f} - \widetilde{K}_{\sigma,\widetilde{f}} \right\|_\infty &= \left\| U_{\sigma^{-1}} \left(W_f - \widetilde{W}_{\widetilde{f}} \right) U_\sigma \right\|_\infty \\
&= \left\| \sum_{y\in\{0,1\}^{n-1}} \begin{pmatrix} \cos\theta_y - \cos\widetilde{\theta}_y & -\left(\sin\theta_y - \sin\widetilde{\theta}_y\right) \\ \sin\theta_y - \sin\widetilde{\theta}_y & \cos\theta_y - \cos\widetilde{\theta}_y \end{pmatrix} \otimes |y\rangle\langle y| \right\|_\infty \\
&= \max_{y\in\{0,1\}^{n-1}} \left\{ 2 \left| \sin \frac{\theta_y - \widetilde{\theta}_y}{2} \right| \left\| \begin{pmatrix} -\sin\frac{\theta_y+\widetilde{\theta}_y}{2} & -\cos\frac{\theta_y+\widetilde{\theta}_y}{2} \\ \cos\frac{\theta_y+\widetilde{\theta}_y}{2} & -\sin\frac{\theta_y+\widetilde{\theta}_y}{2} \end{pmatrix} \right\|_\infty \right\} \\
&\le \max_{y\in\{0,1\}^{n-1}} \left\{ \left| \theta_y - \widetilde{\theta}_y \right| \right\} \le 2^{1-d}\pi \ .
\end{aligned}$$

□

Proof of Theorem 8. It is easy to see that the uniformity condition is satisfied. Let κ denote the key length. Quantum circuit $\mathsf{RSG}^{n,\lambda}$ applies $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$ after reading $(\sigma_i)_{i=1}^T$ and $(f_i)_{i=1}^T$. To implement $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$, we need to realize each of the $T = 10(\lambda+1)n$ unitary gates K. Since each gate K can be implemented in $\mathrm{poly}(n,\lambda,\kappa)$ time, the total construction time for $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$ is also $\mathrm{poly}(n,\lambda,\kappa)$.

Thus, it suffices to prove the requirement of *Statistical Pseudorandomness* is satisfied. Fix $|\eta\rangle \in \mathcal{S}(\mathcal{H})$. Define three distributions:

- ν be the distribution of $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.
- $\widetilde{\nu}$ be the distribution of $\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(\widetilde{f}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$.

– μ be the Haar measure on $\mathcal{S}_{\mathbb{R}}^{2^n}$.

We first prove that the trace distance between ν and $\widetilde{\nu}$ is negligible. To this end, we construct a coupling γ_0 of ν and $\widetilde{\nu}$ by using the same permutation σ_t and letting f_t be the function satisfying $f_t(y)$ is the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$. Therefore, for any $(|\phi\rangle, |\varphi\rangle) \sim \gamma_0$, we have

$$\begin{aligned} \| |\phi\rangle - |\varphi\rangle \|_2 &= \left\| \widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T} |\eta\rangle - \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T} |\eta\rangle \right\|_2 \\ &\le \left\| \widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T} - \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T} \right\|_\infty \\ &\le 2^{1-d}\pi T = \frac{20\pi(\lambda+1)n}{\lambda^{\log\lambda} \cdot n^{\log n}} \ , \end{aligned}$$

where the last inequality is from Fact 1 and Lemma 8. Thus, for any $l \in \mathrm{poly}(\lambda, n)$

$$\begin{aligned} & \left\| \mathop{\mathbb{E}}_{|\phi\rangle\sim\nu}\left[(|\phi\rangle\langle\phi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\varphi\rangle\sim\widetilde{\nu}}\left[(|\varphi\rangle\langle\varphi|)^{\otimes l}\right] \right\|_1 \\ \le\ & \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}\left[\left\| (|\phi\rangle\langle\phi|)^{\otimes l} - (|\varphi\rangle\langle\varphi|)^{\otimes l} \right\|_1 \right] \\ \le\ & l \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0} \left[\| |\phi\rangle\langle\phi| - |\varphi\rangle\langle\varphi| \|_1 \right] \\ \le\ & l \left(\mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0} \left[\| |\phi\rangle (\langle\phi| - \langle\varphi|) \|_1 \right] + \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0} \left[\| (|\phi\rangle - |\varphi\rangle) \langle\varphi| \|_1 \right] \right) \\ \le\ & 2l \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0} \left[\| |\phi\rangle - |\varphi\rangle \|_2 \right] \le \frac{40\pi(\lambda+1)nl}{\lambda^{\log\lambda} \cdot n^{\log n}} \ . \end{aligned} \tag{13}$$

As for the trace distance between $\widetilde{\nu}$ and μ, note that $\widetilde{\nu}$ is the output distribution of T-step parallel Kac's walk. Thus by Theorem 4, we have

$$W_1(\widetilde{\nu}, \mu) \le \frac{1}{2^{\lambda n}} \ .$$

So there exists a coupling of $\widetilde{v}$ and μ, denoted by γ_1, that achieves

$$\mathop{\mathbb{E}}_{(|\varphi\rangle,|\psi\rangle)\sim\gamma_1} \left[\| |\varphi\rangle - |\psi\rangle \|_2 \right] \le \frac{3}{2^{\lambda n}} \ .$$

Therefore, similar to Eq. (13), we have for any $l \in \mathrm{poly}(\lambda, n)$

$$\left\| \mathop{\mathbb{E}}_{|\varphi\rangle\sim\widetilde{\nu}}\left[(|\varphi\rangle\langle\varphi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[(|\psi\rangle\langle\psi|)^{\otimes l}\right] \right\|_1 \le 2l \mathop{\mathbb{E}}_{(|\varphi\rangle,|\psi\rangle)\sim\gamma_1} \left[\| |\varphi\rangle - |\psi\rangle \|_2 \right] \le \frac{6l}{2^{\lambda n}} \ . \tag{14}$$

Finally, by the triangle inequality, Eqs. (13) and (14), we have

$$\left\| \mathop{\mathbb{E}}_{|\phi\rangle\sim\nu}\left[(|\phi\rangle\langle\phi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[(|\psi\rangle\langle\psi|)^{\otimes l}\right] \right\|_1 \le \frac{40\pi(\lambda+1)nl}{\lambda^{\log\lambda} \cdot n^{\log n}} + \frac{6l}{2^{\lambda n}} = \mathrm{negl}(\lambda) \ .$$

This establishes the *Statistical Pseudorandomness* property. □

Constructing PRSS We construct a PRSS by replacing the random functions and permutations used in RSS with QPRFs and QPRPs.

Definition 10. *Let $n, T \in \mathbb{N}$, $\mathcal{H}$ be a real Hilbert space with dimension 2^n, $\tau : \mathcal{K}_1 \times \{0,1\}^n \to \{0,1\}^n$ be a* QPRP *with key space $\mathcal{K}_1$ and $F : \mathcal{K}_2 \times \{0,1\}^{n-1} \to \{0,1\}^d$ be a* QPRF *with key space $\mathcal{K}_2$. An ensemble of unitary operators $\mathsf{SG}^n := \left\{\mathsf{SG}^{n,\lambda}\right\}_\lambda$ with $\mathsf{SG}^{n,\lambda} := \left\{\mathsf{SG}_k^{n,\lambda} : \mathcal{H} \to \mathcal{H}\right\}_{k \in (\mathcal{K}_1 \times \mathcal{K}_2)^T}$ is defined as*

$$\mathsf{SG}_k^{n,\lambda} = K_{\tau_{r_T}, F_{s_T}} \cdots K_{\tau_{r_2}, F_{s_2}} K_{\tau_{r_1}, F_{s_1}}$$

for $k = (r_1, s_1, r_2, s_2, \ldots, r_T, s_T) \in (\mathcal{K}_1 \times \mathcal{K}_2)^T$, where $K_{\sigma,f} = U_{\sigma^{-1}} W_f U_\sigma$ is defined in (11).

Theorem 9. *Let $n \in \mathbb{N}$, $d = \log^2\lambda + \log^2 n$ and $T = 10(\lambda+1)n$. The ensemble of unitary operators SG^n defined in Definition 10 is a* PRSS.

Proof. Due to the efficiency of τ and F, the key length is bounded by $2T \cdot \mathrm{poly}(n, d) = \mathrm{poly}(n, \lambda)$. Thus the condition of polynomial-bounded key length is satisfied. To implement $\mathsf{SG}_k^{n,\lambda}$, we need to realize each of the $T = 10(\lambda+1)n$ unitary gates K that make up $\mathsf{SG}_k^{n,\lambda}$. Since each K can be realized in $\mathrm{poly}(n, \lambda)$ time (efficiency of τ and F), the overall construction time for $\mathsf{SG}_k^{n,\lambda}$ will be $\mathrm{poly}(n, \lambda)$. Thus the uniformity is also satisfied.

We now prove the pseudorandomness property. To this end, we consider three hybrids for an arbitrary $|\phi\rangle \in \mathcal{S}(\mathcal{H})$ and $l \in \mathrm{poly}(\lambda, n)$:

H1: $|\phi_k\rangle^{\otimes l}$ for $|\phi_k\rangle = \mathsf{SG}_k^{n,\lambda}|\phi\rangle$ where $k \leftarrow (\mathcal{K}_1 \times \mathcal{K}_2)^T$ is chosen uniformly at random.

H2: $\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}\right\rangle^{\otimes l}$ for $\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}\right\rangle = \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\phi\rangle$ with independently and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$. $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$ is defined in Definition 9.

H3: $|\psi\rangle^{\otimes l}$ for $|\psi\rangle$ chosen according to the Haar measure μ on $\mathcal{S}_{\mathbb{R}}^{2^n}$.

We first prove that H1 and H2 are computationally indistinguishable. By the quantum-secure property of τ and F, we know the following two situations are computationally indistinguishable for any polynomial-time quantum oracle algorithm $\mathcal{A}$ (see Lemma 6):

- given oracle access to $\tau_{r_1}, \cdots, \tau_{r_T}$ and $F_{s_1}, \cdots, F_{s_T}$ where $(r_i)_{i=1}^T \subseteq \mathcal{K}_1$ and $(s_i)_{i=1}^T \subseteq \mathcal{K}_2$ are independently and uniformly random keys.
- given oracle access to independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.

Thus, we have for any polynomial-time quantum algorithm $\mathcal{A}$,

$$\left|\Pr\left[\mathcal{A}\left(|\phi_k\rangle^{\otimes l}\right) = 1\right] - \Pr\left[\mathcal{A}\left(\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}\right\rangle^{\otimes l}\right) = 1\right]\right| = \mathrm{negl}(\lambda) \quad .$$

For H2 and H3, they are statistically indistinguishable since RSG^n defined in Definition 9 is an RSS by Theorem 8. Finally, by the triangle inequality we establish H1 and H3 are computationally indistinguishable. This accomplishes the proof. □

5.2 Constructing (P)RSS over the Complex Space

This section provides constructions of a RSS and a PRSS over $\mathbb{C}$. Similar to the case over $\mathbb{R}$, our initial step is to create a unitary gate that can be utilized to simulate a single iteration of parallel Kac's walk within a complex Hilbert space. Fix $f, g, h : \{0,1\}^{n-1} \to \{0,1\}^d$ and $\sigma \in S_{2^n}$. Let $\widehat{L}_{\sigma,f,g,h} = U_{\sigma^{-1}} \widehat{Q}_{f,g,h} U_\sigma$, where U_σ is defined as before and $\widehat{Q}_{f,g,h}$ is

$$\sum_{y\in\{0,1\}^{n-1}} \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} \end{pmatrix} \begin{pmatrix} \cos\theta_y & -\sin\theta_y \\ \sin\theta_y & \cos\theta_y \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} \end{pmatrix} \otimes |y\rangle\langle y| , \tag{15}$$

in which

$$\theta_y = \arcsin\left(\sqrt{\mathrm{val}(f(y))}\right) \quad , \quad \alpha_y = 2\pi \cdot \mathrm{val}(g(y)) \quad , \quad \beta_y = 2\pi \cdot \mathrm{val}(h(y)) \quad .$$

Here we decompose $U(\alpha_y, \beta_y, \theta_y)$ into a product of three matrices according to the formula

$$\begin{pmatrix} e^{\mathrm{i}\alpha}\cos\theta & -e^{\mathrm{i}\beta}\sin\theta \\ e^{-\mathrm{i}\beta}\sin\theta & e^{-\mathrm{i}\alpha}\cos\theta \end{pmatrix} = \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha+\beta}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha+\beta}{2}\right)} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha-\beta}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha-\beta}{2}\right)} \end{pmatrix} . \tag{16}$$

We approximate $\widehat{Q}_{f,g,h}$ by another unitary $Q_{f,g,h}$ which can be constructed as follows applying the similar technique in Section 5:

- apply O_f, O_g, O_h and store the results $f(y), g(y), h(y)$ in ancilla qubits,
- calculate three parameters $\gamma_y^+ \approx \frac{\mathrm{val}(g(y))+\mathrm{val}(h(y))}{2}$, $\gamma_y^- \approx \frac{\mathrm{val}(g(y))-\mathrm{val}(h(y))}{2}$ and $\xi_y \approx \frac{2}{\pi}\arcsin\left(\sqrt{\mathrm{val}(f(y))}\right)$ with a precision up to d bits after the binary point,
- use γ_y^+, ξ_y and γ_y^- in the above step to construct a series of controlled gates on the first qubit which approximates the three matrices in the RHS of (16),
- uncompute all the ancilla qubits.

As a result, we construct a unitary gate $L_{\sigma,f,g,h} = U_{\sigma^{-1}} Q_{f,g,h} U_\sigma$ where $Q_{f,g,h}$ is

$$\sum_{y\in\{0,1\}^{n-1}} \begin{pmatrix} e^{\mathrm{i}2\pi\gamma_y^+} & 0 \\ 0 & e^{-\mathrm{i}2\pi\gamma_y^+} \end{pmatrix} \begin{pmatrix} \cos\left(\frac{\pi}{2}\xi_y\right) & -\sin\left(\frac{\pi}{2}\xi_y\right) \\ \sin\left(\frac{\pi}{2}\xi_y\right) & \cos\left(\frac{\pi}{2}\xi_y\right) \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}2\pi\gamma_y^-} & 0 \\ 0 & e^{-\mathrm{i}2\pi\gamma_y^-} \end{pmatrix} \otimes |y\rangle\langle y| , \tag{17}$$

and for any $y \in \{0,1\}^{n-1}$,

$$\left|\frac{\pi}{2}\xi_y - \theta_y\right| \leq 2^{-d-1}\pi \ , \quad \left|2\pi\gamma_y^+ - \frac{\alpha_y + \beta_y}{2}\right| \leq 2^{1-d}\pi \ , \quad \left|2\pi\gamma_y^- - \frac{\alpha_y - \beta_y}{2}\right| \leq 2^{1-d}\pi \ .$$

By utilizing the gate $L_{\sigma,f,g,h}$ together with random permutations and random functions, we can implement the following scheme to produce an RSS:

Definition 11. *Let $n, T, d \in \mathbb{N}$, and $\mathcal{H}$ be a complex Hilbert space with dimension 2^n. An ensemble of unitary operators $\mathsf{RSGC}^n := \left\{\mathsf{RSGC}^{n,\lambda}\right\}_\lambda$ with $\mathsf{RSGC}^{n,\lambda} :=$*

$$\left\{\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}\right\}_{(\sigma_i)_{i=1}^T \subseteq S_{2^n},(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T \subseteq \{f:\{0,1\}^{n-1}\to\{0,1\}^d\}}$$

is define as

$$\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T} = L_{\sigma_T,f_T,g_T,h_T} \cdots L_{\sigma_2,f_2,g_2,h_2} L_{\sigma_1,f_1,g_1,h_1}$$

where $L_{\sigma,f,g,h} = U_{\sigma^{-1}} Q_{f,g,h} U_\sigma$ is defined in (17).

Theorem 10. *Let $n \in \mathbb{N}$, $d = 2\left(\log^2\lambda + \log^2 n\right)$ and $T = 10(\lambda+1)n$. The ensemble of unitary operators RSGC^n defined in Definition 11 is an RSS.*

Based on this, we obtain a PRSS by substituting the random functions and permutations utilized in RSS with their quantum-secure pseudorandom counterparts.

Definition 12. *Let $n, T \in \mathbb{N}$, $\mathcal{H}$ be a complex Hilbert space with dimension 2^n, $\tau : \mathcal{K}_1 \times \{0,1\}^n \to \{0,1\}^n$ be a QPRP with key space $\mathcal{K}_1$ and $F : \mathcal{K}_2 \times \{0,1\}^{n-1} \to \{0,1\}^d$ be a QPRF with key space $\mathcal{K}_2$. An ensemble of unitary opertors $\mathsf{SGC}^n := \left\{\mathsf{SGC}^{n,\lambda}\right\}_\lambda$ with $\mathsf{SGC}^{n,\lambda} = \left\{\mathsf{SGC}^{n,\lambda}_k : \mathcal{H} \to \mathcal{H}\right\}_{k \in (\mathcal{K}_1\times\mathcal{K}_2\times\mathcal{K}_2\times\mathcal{K}_2)^T}$ is defined as*

$$\mathsf{SGC}^{n,\lambda}_k = L_{\tau_{r_T},F_{u_T},F_{s_T},F_{t_T}} \cdots L_{\tau_{r_2},F_{u_2},F_{s_2},F_{t_2}} L_{\tau_{r_1},F_{u_1},F_{s_1},F_{t_1}}$$

for $k = (r_1, u_1, s_1, t_1, r_2, u_2, s_2, t_2, \ldots, r_T, u_T, s_T, t_T) \in (\mathcal{K}_1 \times \mathcal{K}_2 \times \mathcal{K}_2 \times \mathcal{K}_2)^T$, where $L_{\sigma,f,g,h} = U_{\sigma^{-1}} Q_{f,g,h} U_\sigma$ is defined in (17).

Theorem 11. *Let $n \in \mathbb{N}$, $d = 2\left(\log^2\lambda + \log^2 n\right)$ and $T = 10(\lambda+1)n$. The ensemble of unitary operators SGC^n defined in Definition 12 is a PRSS.*

The detailed proofs of two theorems above share similarities with the real case. For this reason, we attach them in Appendix C.

6 Applications

Since pseudorandom state scramblers subsume pseudorandom state geneators and its siblings in the literature, all applications enabled by PRSGs can also be obtained from PRSSs. This includes for instance symmetric-key encryption and commitment of classical messages as well as secure computation. In this section, we showcase a few novel applications beyond what PRSGs are capable of.

6.1 Compact Quantum Encryption

Because PRSSs map any initial state to a pseudorandom output state, we can readily employ them to encrypt quantum messages. Furthermore, it turns out that PRSS-based quantum encryption schemes offer improvements in terms of *compactness*, a point we discuss below.

We start by recalling the well-known Quantum One-Time Pad, which is the quantum analogue of one-time pad and achieves perfect secrecy. Given an n-qubit state $|\psi\rangle$, we sample a uniform $2n$-bit key $k = k_1 \| k_2$ with $k_1, k_2 \in \{0,1\}^n$ and encrypt $|\psi\rangle$ by

$$|\psi_k\rangle = \mathsf{QOTP}_k|\psi\rangle = X^{k_1} Z^{k_2}|\psi\rangle \ ,$$

where X and Z are Pauli operators applied on each qubit of $|\psi\rangle$.

We can reduce the key length by using pseudorandom keys. For instance, given a pseudorandom genrator $\mathsf{PRG} : \{0,1\}^n \to \{0,1\}^{2n}$, we can expand a uniform n-bit key under PRG and use $\mathsf{PRG}(k)$ as the key to QOTP. Namely we encrypt by

$$|\psi_k\rangle = \mathsf{QOTP}_{\mathsf{PRG}(k)}|\psi\rangle \ .$$

We refer to this scheme as prg-QOTP.

These two schemes are secure if the same key is never used more than *once.* One can extend it to multi-time security with *hybrid* encryption, using in addition a post-quantum secure encryption for *classical* bits. For concreteness, we use a post-quantum $\mathsf{PRF}_k : \{0,1\}^n \to \{0,1\}^{2n}$. To encrypt $|\psi\rangle$, we sample a uniformly random string r, and use $\mathsf{PRF}_k(r)$ as the key to QOTP, i.e., we output cipherstate $(r, |\psi_{k,r}\rangle)$ where

$$|\psi_{k,r}\rangle = \mathsf{QOTP}_{\mathsf{PRF}_k(r)}|\psi\rangle \ .$$

We call this scheme prf-QOTP.

Now suppose we have a PRSS $(\{\mathcal{R}_k^{n,m}\})$ with key space $\mathcal{K} = \{0,1\}^\kappa$, and for simplicity we assume that $n = m$ and we ignore them in the notation. We can construct three encryption schemes, analogous to each of the schemes above.

- PRSS-enc: on random key k and state $|\psi\rangle$, output $|\psi_k\rangle := \mathcal{R}_k|\psi\rangle$.
- prg-PRSS-enc: given a $\mathsf{PRG} : \{0,1\}^n \to \mathcal{K}$, on random key k and state $|\psi\rangle$, output $|\psi_k\rangle := \mathcal{R}_{\mathsf{PRG}(k)}|\psi\rangle$.
- prf-PRSS-enc: given a $\mathsf{PRF} : \{0,1\}^n \to \mathcal{K}$, the key is a random key k for the PRF. On state $|\psi\rangle$, output $(r, |\psi_{k,r}\rangle)$, where $r \leftarrow \{0,1\}^n$ and

$$|\psi_{k,r}\rangle = \mathcal{R}_{\mathsf{PRF}_k(r)}|\psi\rangle \ .$$

ℓ copies of $\|\psi\rangle$	(prg-)QOTP	(prg-)PRSS-enc
$\ell = 1$	$\|\psi_k\rangle = \mathsf{QOTP}_{k \text{ or } \mathsf{PRG}(k)}\|\psi\rangle$	$\|\psi_k\rangle = \mathsf{PRSS}_{k \text{ or } \mathsf{PRG}(k)}\|\psi\rangle$
$\ell > 1$	$(\|\psi_{k_1}\rangle, \ldots, \|\psi_{k_\ell}\rangle)$	$(\|\psi_k\rangle, \ldots, \|\psi_k\rangle)$
Comparison	Need to exchange ℓ indep. keys $k_1, \ldots, k_\ell$	Single key for any polynomial ℓ

ℓ copies of $\|\psi\rangle$	prf-QOTP	prf-PRSS-enc
$\ell = 1$	$(r, \|\psi_{k,r}\rangle = \mathsf{QOTP}_{\mathsf{PRF}_k(r)}\|\psi\rangle)$	$(r, \|\psi_{k,r}\rangle) = \mathsf{PRSS}_{\mathsf{PRF}_k(r)}\|\psi\rangle$
$\ell > 1$	$(\ldots, (r_j, \|\psi_{k,r_j}\rangle), \ldots)$	$(r, \|\psi_{k,r}\rangle, \ldots, \|\psi_{k,r}\rangle)$
Comparison	Cipher size grows by ℓ factor	Cipher size grows by $\frac{1}{2}(\ell + 1)$

Table 2: Advantages of PRSS-based encryptions: maintaining single key instead of linear number of keys or reducing the cipher size growth factor by half.

Advantages of PRSS*-based quantum encryption.* One distinct benefit of PRSS-enc over QOTP is that we can encrypt *multiple copies* of a state $|\psi\rangle$ using PRSS-enc under the *same* key k. This follows from the multi-copy indistinguishability in our PRSS definition. In contrast, QOTP needs independent keys to encrypt each copy of $|\psi\rangle$. This considerably improves *compactness*, and it holds similarly in the other two types of schemes.

A related concept called *quantum private broadcasting* has been investigated by Broadbent, Gonzàlez-Guillén and Schuknecht [14]. They employ (symmetric) t-designs to encrypt t copies of an n-qubit quantum message. While the key length in their construction scales logarithmically with t, it grows exponentially with n. Our PRSS-based scheme maintains a key size of poly(n).

We stress that this applies only to encrypting multiple copies of the *same* input state. If we want to encrypt different states, then fresh keys in (prg-)PRSS-enc or randomness in prf-PRSS-enc should be used.

6.2 Succinct Quantum State Commitment

Next we show how PRSS enables quantum commitment. Bit commitment is a fundamental primitive in cryptography. A sender Alice commits to an input bit b to a receiver Bob in the *commit* phase, which can be revealed later in the *open* phase. This naturally extends to committing bit strings. Two properties are essential.

- *Hiding.* Bob is not able to learn the message b before the open phase.
- *Binding.* Alice cannot fool Bob to accept a different message $b' \neq b$ in the open phase.

We will focus on *non-interactive* commitment schemes where both commit and open phases consist of a single message from the sender to the receiver. If the protocol involves exchanging and processing quantum information, we call it a quantum bit commitment (QBC) scheme. QBC has been extensively studied, and it is shown that QBC can be constructed based on standard PRSGs [5,40].

In a similar vein, one can also consider committing to a *quantum* input state, and this is called quantum state commitment (QSC). QSC has proven useful such as in zero-knowledge proof systems for QMA [16,15].

Recently, Gunn, Ju, Ma and Zhandry give a systematic treatment on QSC [22]. They propose a new characterization of binding termed *swap-binding*. They show a striking hiding-binding *duality* theorem for (non-interactive) quantum commitment: binding holds if the *opening* register held by the sender hides the input state. This significantly simplifies proving binding. They then construct binding commitment schemes which in addition are *succinct*, where the register containing the commitment has a *smaller* size than the message state.[5]

Succinct QSC *from* PRSS. The succinct QSC schemes by [22] are based on post-quantum one-way functions or the potentially weaker primitive of pseudorandom unitary operators (PRU). We show below the viability of building succinct commitment on PRSS. Specifically, we observed that a *succinct* PRSS implies a succinct one-time quantum encryption, and it was shown in [22] that a succinct one-time quantum encryption gives a succinct QSC scheme. Hence (succinct one-time) PRSSs offer an alternate approach of realizing succinct one-time quantum encryptions based on potentially weaker assumptions than one-way functions, and could be weaker than the instantiation via PRUs in [22]. Meanwhile, one-time quantum encryption does not seem to follow immediately from PRS or other primitives implied by PRS.

Theorem 12. *Assuming a succinct* PRSS*, i.e.,* $|\mathcal{K}| < 2^n$*, there exists a succinct* QSC*.*

Proof. This follows from a generic claim in [22]. They show that any one-time secure quantum encryption scheme with *succinct* keys, where the key is shorter than the state to be encrypted, readily gives a succinct QSC. A PRSS is a secure quantum encryption as discussed above. Succinctness translates if PRSS's key length is shorter than the size of the input state. This is stated below. We choose not to fully spelled out the syntax and definitions of the involved primitives for the sake of clarity, and refer the readers to [22]. □

Lemma 9. *Assuming a succinct* PRSS*, i.e.,* $|\mathcal{K}| < 2^n$*, there exists a succinct one-time quantum encryption scheme.*

How to instantiate a succinct PRSS? Our construction is not immediately succinct, because the key length $\Omega(\lambda \cdot n)$. We can remedy this by using a pseudorandom generator to expand a key shorter than n into pseudorandom keys for each iteration (QPRF and QPRP).

[5] Note that hiding is not required in these succinct schemes.

References

[1] Aaronson, S., Bouland, A., Fefferman, B., Ghosh, S., Vazirani, U., Zhang, C., Zhou, Z.: Quantum pseudoentanglement. In: 15th Innovations in Theoretical Computer Science Conference (ITCS 2024). Leibniz International Proceedings in Informatics (LIPIcs), vol. 287, pp. 2:1–2:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2024). https://doi.org/10.4230/LIPIcs.ITCS.2024.2 2

[2] Alagic, G., Majenz, C., Russell, A.: Efficient simulation of random states and random unitaries. In: Advances in Cryptology – EUROCRYPT 2020. pp. 759–787. Springer (2020). https://doi.org/10.1007/978-3-030-45727-3_26 2

[3] Ananth, P., Gulati, A., Kaleoglu, F., Lin, Y.T.: Pseudorandom isometries. In: Advances in Cryptology – EUROCRYPT 2024. pp. 226–254. Springer Nature Switzerland, Cham (2024) 4

[4] Ananth, P., Gulati, A., Qian, L., Yuen, H.: Pseudorandom (function-like) quantum state generators: New definitions and applications. In: 20th Theory of Cryptography Conference – TCC 2022. pp. 237–265. Springer (2022). https://doi.org/10.1007/978-3-031-22318-1_9 2, 4, 13, 52, 53

[5] Ananth, P., Qian, L., Yuen, H.: Cryptography from pseudorandom quantum states. In: Advances in Cryptology – CRYPTO 2022. pp. 208–236. Springer (2022). https://doi.org/10.1007/978-3-031-15802-5_8 2, 4, 13, 32, 52, 53

[6] Behera, A., Sattath, O.: Almost public quantum coins. arXiv preprint arXiv:2002.12438 (2020) 2

[7] Bollobás, B.: Random Graphs. Cambridge University Press, 2nd edn. (2001). https://doi.org/10.1017/CBO9780511814068 42

[8] Bouland, A., Fefferman, B., Vazirani, U.: Computational Pseudorandomness, the Wormhole Growth Paradox, and Constraints on the AdS/CFT Duality. In: 11th Innovations in Theoretical Computer Science Conference (ITCS 2020). Leibniz International Proceedings in Informatics (LIPIcs), vol. 151, pp. 63:1–63:2. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2020). https://doi.org/10.4230/LIPIcs.ITCS.2020.63 2

[9] Brakerski, Z., Magrafta, N.: Real-valued somewhat-pseudorandom unitaries (2024) 4

[10] Brakerski, Z., Shmueli, O.: (Pseudo) random quantum states with binary phase. In: 17th Theory of Cryptography Conference – TCC 2019. pp. 229–250. Springer (2019). https://doi.org/10.1007/978-3-030-36030-6_10 2, 11

[11] Brakerski, Z., Shmueli, O.: Scalable pseudorandom quantum states. In: Advances in Cryptology – CRYPTO 2020. pp. 417–440. Springer (2020). https://doi.org/10.1007/978-3-030-56880-1_15 2, 13, 51

[12] Brandão, F.G., Chemissany, W., Hunter-Jones, N., Kueng, R., Preskill, J.: Models of quantum complexity growth. PRX Quantum **2**, 030316 (Jul 2021). https://doi.org/10.1103/PRXQuantum.2.030316 2
[13] Brandão, F.G., Harrow, A.W., Horodecki, M.: Local random quantum circuits are approximate polynomial-designs. Communications in Mathematical Physics **346**, 397–434 (2016). https://doi.org/10.1007/s00220-016-2706-8 5
[14] Broadbent, A., González-Guillén, C.E., Schuknecht, C.: Quantum private broadcasting. Phys. Rev. A **105**, 022606 (Feb 2022). https://doi.org/10.1103/PhysRevA.105.022606 31
[15] Broadbent, A., Grilo, A.B.: QMA-hardness of consistency of local density matrices with applications to quantum zero-knowledge. SIAM Journal on Computing **51**(4), 1400–1450 (2022). https://doi.org/10.1137/21M140729X 32
[16] Broadbent, A., Ji, Z., Song, F., Watrous, J.: Zero-knowledge proof systems for QMA. SIAM Journal on Computing **49**(2), 245–283 (2020). https://doi.org/10.1137/18M1193530 32
[17] Chatterjee, S., Diaconis, P., Sly, A., Zhang, L.: A phase transition for repeated averages. The Annals of Probability **50**(1), 1 – 17 (2022). https://doi.org/10.1214/21-AOP1526 5
[18] Chen, C.F., Bouland, A., Brandão, F.G.S.L., Docter, J., Hayden, P., Xu, M.: Efficient unitary designs and pseudorandom unitaries from permutations (2024) 5
[19] Coladangelo, A.: Quantum trapdoor functions from classical one-way functions. arXiv preprint arXiv:2302.12821 (2023) 2
[20] Dankert, C., Cleve, R., Emerson, J., Livine, E.: Exact and approximate unitary 2-designs and their application to fidelity estimation. Physical Review A **80**(1), 012304 (2009). https://doi.org/10.1103/PhysRevA.80.012304 5
[21] Diaconis, P., Saloff-Coste, L.: Bounds for Kac's master equation. Communications in Mathematical Physics **209**, 729–755 (2000). https://doi.org/10.1007/s002200050036 5, 6
[22] Gunn, S., Ju, N., Ma, F., Zhandry, M.: Commitments to quantum states. In: Proceedings of the 55th Annual ACM Symposium on Theory of Computing. p. 1579–1588. Association for Computing Machinery (2023). https://doi.org/10.1145/3564246.3585198 2, 4, 32
[23] Haferkamp, J.: Random quantum circuits are approximate unitary t-designs in depth $O\left(nt^{5+o(1)}\right)$. Quantum **6**, 795 (2022). https://doi.org/10.22331/q-2022-09-08-795 5
[24] Harrow, A.W., Low, R.A.: Efficient quantum tensor product expanders and k-designs. In: International Workshop on Approximation Algorithms for Combinatorial Optimization – APPROX–RANDOM 2009. pp. 548–561. Springer (2009). https://doi.org/10.1007/978-3-642-03685-9_41 5
[25] Harrow, A.W., Mehraban, S.: Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates. Communications in Mathematical Physics pp. 1–96 (2023). https://doi.org/10.1007/s00220-023-04675-z 5

[26] Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**(1), 97–109 (1970). https://doi.org/10.1093/biomet/57.1.97 5

[27] Haug, T., Bharti, K., Koh, D.E.: Pseudorandom unitaries are neither real nor sparse nor noise-robust. arXiv preprint arXiv:2306.11677 (2023) 2

[28] Hough, B., Jiang, Y.: Cut-off phenomenon in the uniform plane Kac walk. The Annals of Probability **45**(4), 2248–2308 (2017), http://www.jstor.org/stable/26362254 5

[29] Jain, V., Pillai, N.S., Sah, A., Sawhney, M., Smith, A.: Fast and memory-optimal dimension reduction using Kac's walk. The Annals of Applied Probability **32**(5), 4038 – 4064 (2022). https://doi.org/10.1214/22-AAP1784 5

[30] Janvresse, É.: Bounds on semigroups of random rotations on SO(n). Theory of Probability & Its Applications **47**(3), 526–532 (2003). https://doi.org/10.1137/S0040585X97979950 5

[31] Ji, Z., Liu, Y.K., Song, F.: Pseudorandom quantum states. In: Advances in Cryptology – CRYPTO 2018. pp. 126–152. Springer (2018). https://doi.org/10.1007/978-3-319-96878-0_5 1, 2, 11, 51

[32] Jiang, Y.: Total variation bound for Kac's random walk. The Annals of Applied Probability pp. 1712–1727 (2012), https://www.jstor.org/stable/41713374 5, 6

[33] Kac, M.: Foundations of kinetic theory. In: Third Berkeley symposium on mathematical statistics and probability. vol. 3, pp. 171–197 (1956) 3, 6

[34] Kaplan, E., Naor, M., Reingold, O.: Derandomized constructions of k-wise (almost) independent permutations. Algorithmica **55**(1), 113–133 (2009). https://doi.org/10.1007/11538462_30 55

[35] Kim, I., Tang, E., Preskill, J.: The ghost in the radiation: Robust encodings of the black hole interior. Journal of High Energy Physics **2020**(6), 1–65 (2020). https://doi.org/10.1007/JHEP06(2020)031 2

[36] Kretschmer, W.: Quantum pseudorandomness and classical complexity. In: 16th Conference on the Theory of Quantum Computation, Communication and Cryptography – TQC 2021. pp. 2:1–2:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2021). https://doi.org/10.4230/LIPIcs.TQC.2021.2 2, 4, 13

[37] Kretschmer, W., Qian, L., Sinha, M., Tal, A.: Quantum cryptography in algorithmica. In: 55th Annual ACM Symposium on Theory of Computing – STOC 2023. pp. 1589–1602 (2023). https://doi.org/10.1145/3564246.3585225 2

[38] Levin, D.A., Peres, Y., Wilmer, E.: Markov chains and mixing times. American Mathematical Society (2009) 8, 9

[39] Metger, T., Poremba, A., Sinha, M., Yuen, H.: Simple constructions of linear-depth t-designs and pseudorandom unitaries (2024) 4

[40] Morimae, T., Yamakawa, T.: Quantum commitments and signatures without one-way functions. In: Advances in Cryptology – CRYPTO 2022. pp. 269–295. Springer (2022). https://doi.org/10.1007/978-3-031-15802-5_10 2, 32

[41] Movassagh, R., Szegedy, M., Wang, G.: Repeated averages on graphs. arXiv:2205.04535 (2022) 5

[42] O'Donnell, R., Servedio, R.A., Paredes, P.: Explicit orthogonal and unitary designs. In: 2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS). pp. 1240–1260. IEEE Computer Society (2023). https://doi.org/10.1109/FOCS57990.2023.00073 5

[43] Oliveira, R.I.: On the convergence to equilibrium of Kac's random walk on matrices. The Annals of Applied Probability **19**(3), 1200–1231 (2009), https://www.jstor.org/stable/30243617 5, 6

[44] Pillai, N.S., Smith, A.: Kac's walk on n-sphere mixes in $n \log n$ steps. The Annals of Applied Probability **27**(1), 631–650 (2017). https://doi.org/10.1214/16-AAP1214 3, 9, 14, 16, 38, 40, 44, 68

[45] Pillai, N.S., Smith, A.: On the mixing time of Kac's walk and other high-dimensional Gibbs samplers with constraints. The Annals of Probability **46**(4), 2345–2399 (2018), https://www.jstor.org/stable/26506605 5, 6

[46] Renes, J.M., Blume-Kohout, R., Scott, A.J., Caves, C.M.: Symmetric informationally complete quantum measurements. Journal of Mathematical Physics **45**(6), 2171–2180 (2004). https://doi.org/10.1063/1.1737053 5

[47] Yang, L., Engelhardt, N.: The complexity of learning (pseudo)random dynamics of black holes and other chaotic systems. arXiv:2302.11013 (2023), https://arxiv.org/abs/2302.11013 2

[48] Zhandry, M.: Secure identity-based encryption in the quantum random oracle model. International Journal of Quantum Information **13**(04), 1550014 (2015). https://doi.org/10.1142/S0219749915500148 55

[49] Zhandry, M.: A note on quantum-secure PRPs. Cryptology ePrint Archive, Paper 2016/1076 (2016), https://eprint.iacr.org/2016/1076 10

[50] Zhandry, M.: How to construct quantum random functions. J. ACM **68**(5) (2021). https://doi.org/10.1145/3450745 3, 10, 53

[51] Zyczkowski, K., Kus, M.: Random unitary matrices. Journal of Physics A: Mathematical and General **27**(12), 4235 (1994). https://doi.org/10.1088/0305-4470/27/12/028 20

A Dispersing RSS

As mentioned before, our parallel Kac's walk also mixes rapidly in terms of total variation distance, which endows our scramblers with a unique dispersing property. We first give the formal definitions of a random scrambler with a dispersing property in Section A.1. Then, Section A.2 analyses the total variation mixing time of the parallel Kac's walk on both real and complex Hilbert spaces. Lastly in Section A.3, we utilize this rapid mixing property to demonstrate that the ensembles of unitary operators we construct in Section 5 do exhibit a dispersing property.

A.1 Definitions

We introduce the concept of *dispersing random state scramblers* (DRSS), which ensure the approximation of Haar randomness with respect to Wasserstein distance.

Definition 13 (Dispersing Random State Scrambler). *Let $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$ be Hilbert spaces of dimensions 2^n and 2^m respectively with $n, m \in \mathbb{N}$ and $n \leq m$. Let $\mathcal{K} = \{0,1\}^\kappa$ be a key space, and λ be a security parameter. A* dispersing random state scrambler *(DRSS) is an ensemble of isometric operators $\mathcal{R}^{n,m} := \{\mathcal{R}^{n,m,\lambda}\}_\lambda$ with $\mathcal{R}^{n,m,\lambda} := \{\mathcal{R}_k^{n,m,\lambda} : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}\}_{k\in\mathcal{K}}$ satisfying:*

- ***Sphere Coverage.*** *There exist $\epsilon = \text{negl}(\lambda)$ such that for any $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$, the family of states $\{\mathcal{R}_k^{n,m}|\phi\rangle\}_{k\in\mathcal{K}}$ forms an ϵ-net of $\mathcal{S}(\mathcal{H}_{\text{out}})$.*
- ***Wasserstein Approximation of Haar randomness.*** *There exist $\delta, \delta' = \text{negl}(\lambda)$ such that for any $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$,*
  - *Let ν be the distribution of $\mathcal{R}_k^{n,m}|\phi\rangle$ with uniformly random $k \leftarrow \mathcal{K}$, and μ be the Haar measure on $\mathcal{S}(\mathcal{H}_{\text{out}})$. Then, there exists a distribution $\widetilde{\nu}$ such that*

$$\|\mu - \tilde{\nu}\|_{\text{TV}} \leq \delta, \quad \text{and} \quad W_\infty(\nu, \widetilde{\nu}) \leq \delta'.$$

- ***Uniformity.*** *$\mathcal{R}^{n,m}$ can be uniformly computed in polynomial time. That is, there is a deterministic Turing machine that, on input $(1^n, 1^m, 1^\lambda, 1^\kappa)$, outputs a quantum circuit Q in $\text{poly}(n, m, \lambda, \kappa)$ time such that for all $k \in \mathcal{K}$ and $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$*

$$Q\,|k\rangle|\phi\rangle = |k\rangle|\phi_k\rangle\ ,$$

  where $|\phi_k\rangle := \mathcal{R}_k^{n,m,\lambda}|\phi\rangle$.

In particular, small Wasserstein distance implies small trace distance between the average states drawn from the two distributions.

Proposition 1. *A DRSS is an RSS with the same parameters.*

Proof. It suffices to prove that Wasserstein approximation of Haar randomness implies statistical pseudorandomness. Let ν and μ be the distribution of the output states of a DRSS and the Haar measure, respectively. By the assumption, there exists a distribution $\widetilde{\nu}$ on $\mathcal{S}(\mathcal{H}_{\text{out}})$ and a coupling γ of ν and $\widetilde{\nu}$ such that

$$\Pr_{(|\psi\rangle,|\psi'\rangle)\sim\gamma}[\||\psi\rangle - |\psi'\rangle\|_2] \leq \text{negl}(\lambda)\,.$$

Notice that ℓ is polynomial in λ. By the triangle inequality,

$$\text{TD}\left(\mathop{\mathbb{E}}_{|\psi\rangle\sim\nu}\left[|\psi\rangle\!\langle\psi|^{\otimes\ell}\right], \mathop{\mathbb{E}}_{|\psi\rangle\sim\widetilde{\nu}}\left[|\psi\rangle\!\langle\psi|^{\otimes\ell}\right]\right) \leq \text{negl}(\lambda)\,.$$

By Lemma 1, the condition that $\|\mu - \widetilde{\nu}\|_{\text{TV}} \leq \text{negl}(\lambda)$ implies

$$\text{TD}\left(\mathop{\mathbb{E}}_{|\psi\rangle\sim\widetilde{\nu}}\left[|\psi\rangle\!\langle\psi|^{\otimes\ell}\right], \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[|\psi\rangle\!\langle\psi|^{\otimes\ell}\right]\right) \leq \text{negl}(\lambda)\,.$$

The result follows from the triangle inequality. □

Moreover, we introduce a continuous version of random state scrambler, where continuous randomness is allowed.

Definition 14 (Continuously Random State Scrambler). *Let $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$ be Hilbert spaces of dimensions 2^n and 2^m respectively with $n, m \in \mathbb{N}$ and $n \leq m$. Let $\mathcal{K}$ be a (continuous) key space, and λ be a security parameter. A* continuously random state scrambler *(*CRSS*) is an ensemble of isometric operators $\mathcal{R}^{n,m} := \{\mathcal{R}^{n,m,\lambda}\}_\lambda$ with $\mathcal{R}^{n,m,\lambda} := \{\mathcal{R}_k^{n,m,\lambda} : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}\}_{k\in\mathcal{K}}$ satisfying:*

– ***Total-Variation Approximation of Haar randomness.*** *Let $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\text{in}})$ be an arbitrary pure state. Let ν be the distribution of $\mathcal{R}_k^{n,m,\lambda}|\phi\rangle$ with uniformly random $k \leftarrow \mathcal{K}$, and μ be the Haar measure on $\mathcal{S}(\mathcal{H}_{\text{out}})$. Then there exists $\delta = \text{negl}(\lambda)$ such that the total variation distance between ν and μ is at most δ, i.e., $\|\nu - \mu\|_{\text{TV}} \leq \delta$.*

A.2 Total Variation Mixing Time of the Parallel Kac's Walk

We assume $n = 2m$ for some $m \in \mathbb{N}$ throughout this section.

Background: Two-Phase Proof Strategy in [44] It is proved in [44] that the total variation mixing time of Kac's walk on $\mathcal{S}_{\mathbb{R}}^n$ is $\Theta(n \log n)$.

Theorem 13 (Theorem 1, [44]). *Let $\{X_t \in \mathcal{S}_{\mathbb{R}}^n\}_{t\geq 0}$ be a Markov chain that evolves according to Kac's walk. Then, for sufficiently large n, and $T > 200n\log n$,*

$$\sup_{X_0\in\mathcal{S}_{\mathbb{R}}^n} \|\mathcal{L}(X_T) - \mu\|_{\text{TV}} = O\left(\frac{1}{\text{poly}(n)}\right)\,,$$

where μ is the normalized Haar measure on $\mathcal{S}_{\mathbb{R}}^n$.

We informally revisit their proof approach that utilizes the famous coupling lemma (see Lemma 2). The coupling lemma offers a practical approach to estimate the mixing time of a Markov chain by comparing the behavior of two coupled random walks. The total variation distance at time T is bounded by the probability that two coupled random walks are distinct at time T. Through a two-phase coupling of $\{X_t\}_{t\geq 0}$ and $\{Y_t\}_{t\geq 0}$, they show that the probability that two copies are not equal at time T, i.e., $\Pr[X_T \neq Y_T]$, approaches zero when n is sufficiently large and $T > 200n \log n$. Specifically, the two-step coupling consists of an initial *contracting phase* followed by a subsequent *coalescing phase*. The contracting phase aims to sufficiently reduce the distance between two copies of Kac's walk so that during the coalescing phase, they can be further fine-tuned to coalesce. These two phases are described below, with a focus on how the random angles are coupled.

Contracting Phase (from $t = 0$ to T_0). The contracting phase starts from time 0 and continues until time T_0. In this phase, $\{X_t\}$ and $\{Y_t\}$ undergo the *proportional coupling.* which aims at reducing the distance between two copies of Kac's walk. The proportional coupling is introduced in Section 4.

Coalescing Phase (from $t = T_0$ to $T_0 + T_1$). This phase employs a non-Markovian coupling[6] starting from a close-by pair X_{T_0} and Y_{T_0}. Let T_1 be determined later. Initially, the coupling independently and identically samples T_1 pairs of coordinates $\{(i_t, j_t)\}_{t=T_0}^{T_0+T_1-1}$ all at once, which are subsequently used to generate T_1+1 partitions of $[n]$, denoted by $\{\mathcal{P}_t\}_{t=T_0}^{T_0+T_1}$. The construction of these partitions is done inductively in reverse order. The last partition is enforced to be $\mathcal{P}_{T_0+T_1} = \{\{1\}, \{2\}, \ldots, \{n\}\}$. Starting from $t = T_1 + T_0 - 1$ and decrementing down to T_0, the construction of partition $\mathcal{P}_t$ uses the chosen coordinate pair (i_t, j_t) as a guide. Specifically, it is generated by merging two sets in $\mathcal{P}_{t+1}$: one set includes i_t, and the other includes j_t; while leaving other sets untouched. Then, the value of T_1 is determined such that $\mathcal{P}_{T_0} = \{[n]\}$ with high probability.

The aim of this phase is to ultimately coalesce X and Y. To see how to achieve this, we introduct the event $\mathcal{A}_t$ in which, at time t,

$$\sum_{i\in S} X_t[i]^2 = \sum_{i\in S} Y_t[i]^2, \quad \forall S \in \mathcal{P}_t$$

and

$$X_t[k]Y_t[k] \geq 0, \quad k \in \{i_{t-1}, j_{t-1}\}\,.$$

Intuitively, event $\mathcal{A}_t$ states that if we partition X_t and Y_t based on $\mathcal{P}_t$, then both X_t and Y_t carry equal significance within each segment S at time t; and meanwhile the corresponding updated subvectors share the same sign at time $t-1$. Conditioning on $\mathcal{P}_{T_0} = \{\{1, \cdots, n\}\}$, which holds with high probability, it is not hard to verify that $\mathcal{A}_{T_0}$ occurs and $\cap_{t=T_0}^{T_0+T_1} \mathcal{A}_t$ implies that the corresponding entries in vectors $X_{T_0+T_1}$ and $Y_{T_0+T_1}$ are equal. Thus, to prove that X and Y

[6] The non-Markovian coupling refers to a situation where the transition between states depend not only on the current state but also on the future states, violating the memoryless property of a standard Markov process.

are identical by the end of this phase, it suffices to prove that all events occur with a high probability during the process.

So, the non-Markovian coupling aims to ensure that $\mathcal{A}_{t+1}$ takes place with a high probability, conditioned on all previous events occur. This is achieved by sampling θ and θ' from a "good" joint distribution, which makes sure that both marginal distributions are uniformly distributed on $[0, 2\pi)$. Such a desirable distribution is made possible by the entry-wise closeness achieved during the first phase.

Intuitively, the parallel Kac's walk is expected to have a mixing time that is only on the order of $O(\log n)$, saving factor of n in the mixing time of the original Kac's walk. However, the mixing time for the parallel Kac's walk cannot be derived from the mixing time for the original Kac's walk, directly. Fortunately, through careful modifications to the two-phase coupling approach above, we have discovered a logarithmic mixing time for the parallel Kac's walk, resulting in exponential speedup compared to the original random walk.

Real Case As we introduced in Section 4, the total variation distance between the output distribution of a parallel Kac's walk after T steps and the normalized Haar measure on $\mathcal{S}^n_{\mathbb{R}}$ decays exponentially as T grows. We restate the theorem here.

Theorem 5. *Let $\{X_t \in \mathcal{S}^n_{\mathbb{R}}\}_{t\geq 0}$ be a Markov chain that evolves according to the parallel Kac's walk. Then, for sufficiently large n, $c > 515$ and $T = c\log n$,*

$$\sup_{X_0\in\mathcal{S}^n_{\mathbb{R}}} \|\mathcal{L}(X_T) - \mu\|_{\mathrm{TV}} \leq \frac{1}{2^{(c/515-1)\log n - 1}} ,$$

where μ is the normalized Haar measure on $\mathcal{S}^n_{\mathbb{R}}$.

To prove Theorem 5, we will use the coupling lemma (see Lemma 2) and extend the two-phase coupling method described in [44] to accommodate parallel Kac's walks. We have already extended the proportional coupling used in the contracting phase in Section 4.1. We now introduce the non-Markovian coupling employed in the coalescing phase to ensure that X and Y converge to an identical state, and integrate these two couplings into a comprehensive two-phase coupling to establish the mixing time of the parallel Kac's walk.

Coalescing Phase: the Non-Markovian Coupling The non-markovian coupling is defined in Definition 15. As we will see later, if the initial vectors of two parallel Kac's walks, namely X_{T_0} and Y_{T_0}, are close, this coupling guarantees a high probability of collision between X_T and Y_T, when T is sufficiently large.

Definition 15 (Non-Markovian Coupling). *Fix $T_0 \leq T \in \mathbb{N}$. We couple $\{X_t\}_{T_0\leq t\leq T}, \{Y_t\}_{T_0\leq t\leq T}$ in the following way:*

1. *For each $T_0 \leq t < T$, choose a perfect matching*

$$P_t = \left\{ \left(i_1^{(t)}, j_1^{(t)}\right), \dots, \left(i_m^{(t)}, j_m^{(t)}\right) \right\}$$

uniformly at random.

2. *Set* $\mathcal{P}_{T,1} = \{\{1\}, \ldots, \{n\}\}$*, and define a sequence of partitions*

$$\{\mathcal{P}_{t,k}\}_{T_0 \le t < T,\ 1 \le k \le m+1}$$

of $[n]$ *inductively by the process:*
(a) *If* $k = m+1$*, let* $\mathcal{P}_{t,k} = \mathcal{P}_{t+1,1}$*.*
(b) *If* $1 \le k \le m$*, write* $\mathcal{P}_{t,k+1} = \left\{S_1(t,k+1), \ldots, S_{l_{t,k+1}}(t,k+1)\right\}$ *with* $S_r(t,k+1) \subseteq [n]$ *for* $1 \le r \le l_{t,k+1}$*. Let* $u_{t,k},\ v_{t,k}$ *be the indices such that*

$$i_k^{(t)} \in S_{u_{t,k}}(t,k+1) \quad \text{and} \quad j_k^{(t)} \in S_{v_{t,k}}(t,k+1) \ .$$

   i. *If* $u_{t,k} = v_{t,k}$*, set* $\mathcal{P}_{t,k} = \mathcal{P}_{t,k+1}$*.*
   ii. *If* $u_{t,k} \ne v_{t,k}$*, construct* $\mathcal{P}_{t,k}$ *by merging* $S_{u_{t,k}}(t,k+1)$ *and* $S_{v_{t,k}}(t,k+1)$ *in* $\mathcal{P}_{t,k+1}$*.*

3. *If* $\mathcal{P}_{T_0,1} = \{[n]\}$*, we couple* $\{X_t\}_{T_0 \le t \le T}, \{Y_t\}_{T_0 \le t \le T}$ *in the following way:*
   – *Define the set*

$$H = \{(t,k) : T_0 \le t < T,\ 1 \le k \le m,\ \mathcal{P}_{t,k} \ne \mathcal{P}_{t,k+1}\} \ . \tag{18}$$

   – *Fix* $T_0 \le t < T$*,* X_t *and* Y_t*, and we couple* X_{t+1} *and* Y_{t+1} *in the following way:*
   (a) *Set* $X_{t,1} = X_t$ *and* $Y_{t,1} = Y_t$*.*
   (b) *For* $1 \le k \le m$*,*
      i. *If* $(t,k) \notin H$*, uniformly choose* $\theta_k^{(t)} \in [0, 2\pi)$*. Let*

$$X_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, \theta_k^{(t)}, X_{t,k}) \text{ and } Y_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, {\theta'}_k^{(t)}, Y_{t,k})$$

where $G(\cdot)$ *is defined in Eq. (1) and* ${\theta'}_k^{(t)}$ *is obtained in the same way as the proportional coupling defined in Definition 8.*
      ii. *If* $(t,k) \in H$*, let* θ_0 *be the angle satisfies*

$$X_{t,k}[i_k^{(t)}] = \sqrt{X_{t,k}[i_k^{(t)}]^2 + X_{t,k}[j_k^{(t)}]^2}\cos(\theta_0) \ ,$$

$$X_{t,k}[j_k^{(t)}] = \sqrt{X_{t,k}[i_k^{(t)}]^2 + X_{t,k}[j_k^{(t)}]^2}\sin(\theta_0) \ ,$$

θ_0' *be the angle satisfies*

$$Y_{t,k}[i_k^{(t)}] = \sqrt{Y_{t,k}[i_k^{(t)}]^2 + Y_{t,k}[j_k^{(t)}]^2}\cos(\theta_0') \ ,$$

$$Y_{t,k}[j_k^{(t)}] = \sqrt{Y_{t,k}[i_k^{(t)}]^2 + Y_{t,k}[j_k^{(t)}]^2}\sin(\theta_0') \ ,$$

and then choose the best distribution ν *among all joint distributions on* $[0,2\pi) \times [0,2\pi)$ *with both marginal distributions uniformly distributed on* $[0,2\pi)$ *which maximizes the probability of the following events when* $(\theta, \theta') \sim \nu$*:*

$$\sum_{i \in S_r(t,k+1)} X_{t,k+1}[i]^2 = \sum_{i \in S_r(t,k+1)} Y_{t,k+1}[i]^2 \ , \ 1 \le r \le l_{t,k+1}$$

$$X_{t,k+1}[i] \cdot Y_{t,k+1}[i] \ge 0 \ , \ i \in \left\{i_k^{(t)}, j_k^{(t)}\right\} \ ,$$

where

$$X_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, \theta - \theta_0, X_{t,k}) \text{ and } Y_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, \theta' - \theta'_0, Y_{t,k}) \ .$$

Then choose $(\theta_k^{(t)}, {\theta'}_k^{(t)}) \sim \nu$, *and set*

$$X_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, \theta_k^{(t)} - \theta_0, X_{t,k}) \text{ and } Y_{t,k+1} = G(i_k^{(t)}, j_k^{(t)}, {\theta'}_k^{(t)} - \theta'_0, Y_{t,k}) \ .$$

(c) *Set* $X_{t+1} = X_{t,m+1}$ *and* $Y_{t+1} = Y_{t,m+1}$.

4. *If* $\mathcal{P}_{T_0,1} \neq \{[n]\}$, *for* $T_0 \leq t \leq T$, *we couple* X_{t+1} *and* Y_{t+1} *in the following way: choose* m *independent angles* $\theta_1^{(t)}, \ldots, \theta_m^{(t)} \in [0, 2\pi)$ *uniformly at random and set*

$$X_{t+1} = F\left(P_t, \theta_1^{(t)}, \ldots, \theta_m^{(t)}, X_t\right) \quad \text{and} \quad Y_{t+1} = F\left(P_t, \theta_1^{(t)}, \ldots, \theta_m^{(t)}, Y_t\right) \ ,$$

where $F(\cdot)$ *is given in Eq.* (2).

Step 1 samples $T - T_0$ matchings, generating all coordinate pairs that will be updated in the succeeding process. Step 2 utilizes this matchings to construct a series of partitions of $[n]$ in a back propagation manner. Starting from $\mathcal{P}_{T,1} = \{\{1\}, \ldots, \{n\}\}$, it sequentially construct

$$\mathcal{P}_{T-1,m+1} \quad \mathcal{P}_{T-1,m} \quad \cdots \quad \mathcal{P}_{T-1,1} \quad \mathcal{P}_{T-2,m+1} \quad \cdots \quad \mathcal{P}_{T-2,1} \quad \cdots \quad \mathcal{P}_{T_0,1} \ .$$

$\mathcal{P}_{t,m+1}$ is set equal to $\mathcal{P}_{t+1,1}$ directly. For $1 \leq k \leq m$, $\mathcal{P}_{t,k}$ is obtained based on $\mathcal{P}_{t,k+1}$ and the k-th pair of coordinates in matching P_t. If two coordinates of the k-th pair belong to different components in partition $\mathcal{P}_{t,k+1}$, we merge these two components. Otherwise, $\mathcal{P}_{t,k}$ is set equal to $\mathcal{P}_{t,k+1}$. This series of partitions thus consists of random partitions of set $[n]$ and with high probability the first partition $\mathcal{P}_{T_0,1}$ is $\{[n]\}$ (see Lemma 10). This follows from the argument for bounding the probability of the connectivity of Erdös-Rényi graphs [7, Theorem 7.3]. The proof is deferred to Appendix C.

Lemma 10. *Fix* $c > 0$ *and* $T_0 \in \mathbb{N}$. *Let* $l = 5(1+c)\log n$ *and* $T = T_0 + l$. *Then we have for* n *sufficiently large,*

$$\Pr[\mathcal{P}_{T_0,1} \neq \{[n]\}] \leq 2n^{-c} \ ,$$

where $\mathcal{P}_{T_0,1}$ *is defined in Definition 15.*

If $\mathcal{P}_{T_0,1} = \{[n]\}$, step 3 serves as the crucial step of this coupling. To provide a clearer explanation of how this coupling technique works, Definition 16 introduces a series of events $\{\mathcal{A}(t,k)\}$. Intuitively, $\mathcal{A}(t,k)$ indicates that the $(k-1)$-th updated coordinates in both vectors have the same signs at time t, and both vectors have the same weight within each component of the partition $\mathcal{P}_{t,k}$.

Definition 16. *Let $\mathcal{A}(T_0, 1)$ denote the event*

$$\sum_{i \in S_r(T_0,1)} X_{T_0,1}[i]^2 = \sum_{i \in S_r(T_0,1)} Y_{T_0,1}[i]^2, 1 \le r \le l_{T_0,1} \ .$$

For other $T_0 \le t \le T$ and $1 \le k \le m+1$, we define the event $\mathcal{A}(t,k)$ as [7]

$$\sum_{i \in S_r(t,k)} X_{t,k}[i]^2 = \sum_{i \in S_r(t,k)} Y_{t,k}[i]^2 \ \ , \ \ 1 \le r \le l_{t,k} \tag{19}$$

$$X_{t,k}[i] \cdot Y_{t,k}[i] \ge 0 \ \ , \ \ i \in \left\{ i_{k-1}^{(t)}, j_{k-1}^{(t)} \right\} \ . \tag{20}$$

It is worthy noting that under the assumption $\mathcal{P}_{T_0,1} = \{[n]\}$, $\mathcal{A}(T_0, 1)$ occurs since the sum of squares of the components of a unit vector is equal to 1. And if $\{\mathcal{A}(t,k)\}$ take place in the entire process, we can conclude that $X_T = Y_T$ because $\mathcal{A}(T,1)$ guarantees the corresponding coordinates have the same absolute values and Remark 1 together with (20) guarantees that they have the same signs as well. So we want to couple the rotation angles to ensure that $\mathcal{A}(t, k+1)$ occurs with a high probability, given that all previous events have already taken place. The coupling of the rotation angles are divided into two cases: whether the coordinate pair $\left(i_k^{(t)}, j_k^{(t)}\right)$ is a "merge point" in the construction process of partitions. If not, the rotation angles are coupled using the proportional coupling (see step 3.b.i). For non-"merge point", the proportional coupling does not affect the partition and also maintain the weight within each component, and thus $\mathcal{A}(t, k+1)$ must occur conditioned on $\mathcal{A}(t,k)$. In the other case, we sample the rotation angles from a "good" joint distribution which maximizes the probability that $\mathcal{A}(t, k+1)$ occurs (see step 3.b.ii).

We next show why such a "good" joint distribution exists. Before that, we set some notations. For $T_0 \le t \le T$ and $1 \le k \le m+1$, define

$$A_{t,k}[i] = X_{t,k}[i]^2 \ \ , \quad B_{t,k}[i] = Y_{t,k}[i]^2 \ . \tag{21}$$

For $T_0 \le t_1, t_2 \le T$ and $1 \le k_1, k_2 \le m+1$, we define a partial order $\sqsubseteq$ as

$$(t_1, k_1) \sqsubseteq (t_2, k_2) \quad \text{iff} \quad (t_1 < t_2) \vee (t_1 = t_2 \wedge k_1 \le k_2) \ .$$

Note that if $\left(i_k^{(t)}, j_k^{(t)}\right)$ is a "merge point", $\mathcal{P}_{t,k+1}$ differs from $\mathcal{P}_{t,k}$ only in the components in which $i_k^{(t)}$ and $j_k^{(t)}$ are, namely $S_{u_{t,k}}(t, k+1)$ and $S_{v_{t,k}}(t, k+1)$. To see whether $\mathcal{A}(t, k+1)$ occurs given $\mathcal{A}(t,k)$, we only need to check whether (19) holds for $r = u_{t,k}$, that is,

$$\sum_{i \in S_{u_{t,k}}(t,k+1)} X_{t,k}[i]^2 = \sum_{i \in S_{u_{t,k}}(t,k+1)} Y_{t,k}[i]^2 \ .$$

[7] In Eq. (20), if $k = 1$, then $i \in \left\{ i_m^{(t-1)}, j_m^{(t-1)} \right\}$.

Rewrite

$$A = \sum_{i \in S_{u_{t,k}}(t,k+1) \setminus i_k^{(t)}} X_{t,k}[i]^2 \ , \qquad B = X_{t,k}[i_k^{(t)}]^2 + X_{t,k}[j_k^{(t)}]^2 \ ,$$

$$C = \sum_{i \in S_{u_{t,k}}(t,k+1) \setminus i_k^{(t)}} Y_{t,k}[i]^2 \ , \qquad D = Y_{t,k}[i_k^{(t)}]^2 + Y_{t,k}[j_k^{(t)}]^2 \ .$$

We will have

$$\sum_{i \in S_{u_{t,k}}(t,k+1)} X_{t,k}[i]^2 = A + B\cos(\theta_k^{(t)})^2 \ , \sum_{i \in S_{u_{t,k}}(t,k+1)} Y_{t,k}[i]^2 = C + D\cos({\theta'}_k^{(t)})^2 \ .$$

The following lemma states that $|A-C|$ and $|B-D|$ are bounded by the initial distance of two vectors. Its proof is the same as Lemma 4.4 in [44].

Lemma 11. *Fix $T_0 < T$, and couple two chains $\{X_t\}_{T_0 \le t \le T}, \{Y_t\}_{T_0 \le t \le T}$ using the non-Markovian coupling defined in Definition 15. Fix $T_0 \le t_0 \le T$ and $1 \le k_0 \le m+1$. Then, on the event $\bigcap_{(t,k) \sqsubseteq (t_0,k_0)} \mathcal{A}(t,k) \cap \{\mathcal{P}_{T_0,1} = \{[n]\}\}$, we have*

$$\|A_{t,k} - B_{t,k}\|_{1,S} \le \|A_{T_0,1} - B_{T_0,1}\|_1$$

for all $(t,k) \sqsubseteq (t_0,k_0)$ and $S \in \mathcal{P}_{t,k}$. Moreover, for all $(t,k) \sqsubseteq (t_0,k_0)$,

$$\|A_{t,k} - B_{t,k}\|_1 \le n \|A_{T_0,1} - B_{T_0,1}\|_1 \ .$$

Knowing that A, B and C, D are close, the following lemma states the existence of a good distribution ν for $\theta_k^{(t)}$ and ${\theta'}_k^{(t)}$ such that $\sum_{i \in S_{u_{t,k}}(t,k+1)} X_{t,k}[i]^2$ agrees with $\sum_{i \in S_{u_{t,k}}(t,k+1)} Y_{t,k}[i]^2$ with high probability.

Lemma 12 (Lemma 4.6 in [44]). *Fix positive reals $1 < p < q' < q/2$. Let $\theta, \theta' \sim \mathrm{Unif}[0, 2\pi)$ and let*

$$S = A + B\cos(\theta)^2 \quad \textit{and} \quad S' = C + D\cos(\theta')^2$$

for some $0 \le A, B, C, D \le 1$ that satisfy

$$|A-C|, |B-D| \le n^{-q} \quad \textit{and} \quad B, D \ge n^{-p} \ .$$

Then for sufficiently large n, there exists a coupling of θ, θ' so that

$$\Pr[S = S'] \ge 1 - 6 \times 10^3 n^{-c}$$

and

$$\cos(\theta)\cos(\theta') \ge 0 \quad \textit{and} \quad \sin(\theta)\sin(\theta') \ge 0$$

where $c = \min\left(\frac{q'}{2}, q - 2q'\right) > 0$.

Proof of the Total Variation Mixing Time

Proof of Theorem 5. Let $a = 66,\ b = 24,\ T_0 = 500\log n,\ T_1 = 15\log n,\ T = T_0 + T_1 = 515\log n$. We construct a coupling of two copies $\{X_t\}_{t\geq 0}$ and $\{Y_t\}_{t\geq 0}$ of the parallel Kac's walk with starting points $X_0 = x \in \mathcal{S}^n_{\mathbb{R}}$ and $Y_0 \sim \mu$. The coupling is as follows:

1. couple $\{X_t\}_{0\leq t\leq T_0}, \{Y_t\}_{0\leq t\leq T_0}$ by using the proportional coupling defined in Definition 8,
2. couple $\{X_t\}_{T_0\leq t\leq T}, \{Y_t\}_{T_0\leq t\leq T}$ by using the non-Markovian coupling defined in Definition 15.

Define the events

$$\begin{aligned}\mathcal{E}_1 &= \left\{\|A_{T_0} - B_{T_0}\|_1 \geq n^{-a}\right\} \ , \\ \mathcal{E}_2 &= \{\mathcal{P}_{T_0,1} \neq \{\{1,\dots,n\}\}\} \ , \\ \mathcal{E}_3 &= \{X_T \neq Y_T\} \ .\end{aligned}$$

By Lemma 2,

$$\sup_{X_0\in\mathcal{S}^n_{\mathbb{R}}} \|\mathcal{L}(X_T) - \mu\|_{\mathrm{TV}} \leq \sup_{X_0\in\mathcal{S}^n_{\mathbb{R}}} \Pr[\mathcal{E}_3] \leq \sup_{X_0\in\mathcal{S}^n_{\mathbb{R}}} (\Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2] + \Pr[\mathcal{E}_3\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c]) \ . \tag{22}$$

By Markov's inequality, we have

$$\begin{aligned}\Pr[\mathcal{E}_1] &= \Pr\left[\|A_{T_0} - B_{T_0}\|_1 \geq n^{-a}\right] \\ &\leq \Pr\left[\|A_{T_0} - B_{T_0}\|_2 \geq n^{-a-1/2}\right] \\ (\text{Lemma 7}) &\leq n^{2a+1}\cdot 2\cdot\left(\frac{3}{4}\right)^{T_0} \leq \frac{1}{n^2} \ .\end{aligned} \tag{23}$$

Moreover, by Lemma 10, we have

$$\Pr[\mathcal{E}_2] \leq 2n^{-2} \ . \tag{24}$$

In order to bound $\Pr[\mathcal{E}_3\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c]$, recall the definition of $\mathcal{A}(t,k)$ in Definition 16. It is evident that $\bigcap_{(t,k)\sqsubseteq(T,1)}\mathcal{A}(t,k)$ implies $\mathcal{E}_3^c$. Thus, $\mathcal{E}_3\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c$ implies $\bigcup_{(t,k)\sqsubseteq(T,1)}\mathcal{A}(t,k)^c\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c$. So, we have

$$\begin{aligned}\Pr[\mathcal{E}_3\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c] &\leq \Pr\left[\bigcup_{(t,k)\sqsubseteq(T,1)}\mathcal{A}(t,k)^c\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right] \\ &\leq \sum_{t=T_0}^{T-1}\sum_{k=1}^{m}\Pr\left[\mathcal{A}(t,k+1)^c\cap\left(\bigcap_{(t',k')\sqsubseteq(t,k)}\mathcal{A}(t',k')\right)\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right] \\ &\quad + \Pr[\mathcal{A}(T_0,1)^c\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c] \ .\end{aligned} \tag{25}$$

Notice that if $\mathcal{E}_2^c$ happens, we have

$$\sum_{i\in[n]} X_{T_0,1}[i]^2 = \sum_{i\in[n]} Y_{T_0,1}[i]^2 = 1 \ ,$$

and the proportional coupling forces $X_{T_0,1}[i] \cdot Y_{T_0,1}[i] = X_{T_0}[i] \cdot Y_{T_0}[i] \geq 0$ for all $1 \leq i \leq n$. Therefore, $A(T_0, 1)$ must occur, i.e.,

$$\Pr[\mathcal{A}(T_0,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] = 0 \ . \tag{26}$$

Combining (25) and (26), we have

$$\Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \leq \sum_{t=T_0}^{T-1}\sum_{k=1}^{m} \Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \ . \tag{27}$$

We are now left to find a upper bound for

$$\Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right]$$

when $T_0 \leq t \leq T-1$ and $1 \leq k \leq m$. To this end, we define

$$\mathcal{B}(t,k) = \left\{ \min_{(t',k')\sqsubseteq(t,k):t'\leq t} \min_{1\leq i\leq n} Y_{t',k'}[i]^2 \geq n^{-b} \right\} \ .$$

Note that

$$\begin{aligned} &\Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \\ &\leq \Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] + \Pr[\mathcal{B}(t,k)^c] \ . \end{aligned} \tag{28}$$

By Lemma 4 and a union bound over all (t',k') such that $(t',k') \sqsubseteq (t,k)$ and $t' \leq t$, we have for sufficiently large n,

$$\Pr[\mathcal{B}(t,k)^c] \leq 15n^{3-\frac{b}{3}}\log(n) \ . \tag{29}$$

Next, we consider two cases of the term

$$\Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right]$$

in (28): $(t,k) \notin H$ and $(t,k) \in H$, where H is defined in Definition 15. In the case that $(t,k) \notin H$, we have $\mathcal{P}_{t,k} = \mathcal{P}_{t,k+1}$ and we apply the proportional coupling. Thus $\mathcal{A}(t,k)$ implies $\mathcal{A}(t,k+1)$ which means

$$\Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] = 0 \ . \tag{30}$$

In the other case that $(t,k) \in H$, let

$$A = \sum_{i \in S_{u_{t,k}}(t,k+1)\setminus i_k^{(t)}} X_{t,k}[i]^2 \ , \qquad B = X_{t,k}[i_k^{(t)}]^2 + X_{t,k}[j_k^{(t)}]^2 \ ,$$

$$C = \sum_{i \in S_{u_{t,k}}(t,k+1)\setminus i_k^{(t)}} Y_{t,k}[i]^2 \ , \qquad D = Y_{t,k}[i_k^{(t)}]^2 + Y_{t,k}[j_k^{(t)}]^2 \ ,$$

$$S = A + B\cos(\theta_k^{(t)})^2 \ , \qquad S' = C + D\cos({\theta'}_k^{(t)})^2 \ .$$

On the event $\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k') \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c$, we have by Lemma 11

$$|A - C| \le \|A_{t,k} - B_{t,k}\|_1 \le n\,\|A_{T_0} - B_{T_0}\|_1 \le n^{1-a} \ .$$

Similarly,

$$|B - D| \le \|A_{t,k} - B_{t,k}\|_1 \le n\,\|A_{T_0} - B_{T_0}\|_1 \le n^{1-a} \ .$$

Moreover, $D \ge n^{-b}$ and $B \ge D - |B-D| \ge n^{-b}$ for sufficiently large n. Then apply Lemma 12 with $p = b,\ q = a-1$ and $q' = \frac{2(a-1)}{5}$, we know there exists a distribution ν_0 such that when $(\theta_k^{(t)}, {\theta'}_k^{(t)}) \sim \nu_0$, we have

$$\cos(\theta_k^{(t)})\cos({\theta'}_k^{(t)}) \ge 0 \quad \text{and} \quad \sin(\theta_k^{(t)})\sin({\theta'}_k^{(t)}) \ge 0$$

$$\Pr_{(\theta_k^{(t)}, {\theta'}_k^{(t)})\sim\nu_0}\left[S \ne S' \middle| \bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k') \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \le 6\times 10^3 n^{-\frac{a-1}{5}} \ .$$

We choose the best distribution which maximizes the probability of event described in (19) and (20), so

$$\begin{aligned}
&\Pr\left[\mathcal{A}(t,k+1)^c \cap \left(\bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\right) \cap \mathcal{B}(t,k)\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right]\\
&\leq \Pr\left[\mathcal{A}(t,k+1)^c | \bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\cap\mathcal{B}(t,k)\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right]\\
&\leq \Pr_{(\theta_k^{(t)},\theta'^{(t)}_k)\sim\nu_0}\left[\mathcal{A}(t,k+1)^c | \bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\cap\mathcal{B}(t,k)\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right]\\
&= \Pr_{(\theta_k^{(t)},\theta'^{(t)}_k)\sim\nu_0}\left[S\neq S' | \bigcap_{(t',k')\sqsubseteq(t,k)} \mathcal{A}(t',k')\cap\mathcal{B}(t,k)\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c\right] \leq 6\times 10^3 n^{-\frac{a-1}{5}} \ .
\end{aligned} \tag{31}$$

Combining (27), (28), (29), (30) and (31), we have for sufficiently large n

$$\Pr[\mathcal{E}_3\cap\mathcal{E}_1^c\cap\mathcal{E}_2^c] \leq \sum_{t=T_0}^{T-1}\sum_{k=1}^{m} 6\times 10^3 n^{-\frac{a-1}{5}} + 15 n^{3-\frac{b}{3}}\log(n) \leq \frac{1}{n^2} \ . \tag{32}$$

By (22), (23), (24) and (32), we have for sufficiently large n and $T = 515\log n$

$$\sup_{X_0\in\mathcal{S}_{\mathbb{R}}^n} \|\mathcal{L}(X_T)-\mu\|_{\mathrm{TV}} \leq \sup_{X_0\in\mathcal{S}_{\mathbb{R}}^n} \Pr[\mathcal{E}_3] \leq \frac{1}{2n} \ .$$

As for $T = c\log n$ where $c > 515$, by Lemma 3 we have

$$\sup_{X_0\in\mathcal{S}_{\mathbb{R}}^n} \|\mathcal{L}(X_T)-\mu\|_{\mathrm{TV}} \leq 2\left(\frac{1}{n}\right)^{\left\lfloor\frac{T}{515\log n}\right\rfloor} \leq \frac{1}{2^{(c/515-1)\log n-1}} \ .$$

□

Complex Case The output distribution of the parallel Kac's walk on complex vectors after T steps approaches the Haar measure on the complex unit sphere of $\mathbb{C}^n$ exponentially fast as T grows. We restate the theorem here.

Theorem 7. *Let $\{X_t\in\mathcal{S}_{\mathbb{C}}^n\}_{t\geq 0}$ be a Markov chain that evolves according to the parallel Kac's walk on complex vectors. Then, for sufficiently large n, $c > 515$ and $T = c\log n$,*

$$\sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \|\mathcal{L}(X_T)-\mu\|_{\mathrm{TV}} \leq \frac{1}{2^{(c/515-1)\log n-1}} \ ,$$

where $\mu_{\mathbb{C}}$ is the Haar measure on $\mathcal{S}_{\mathbb{C}}^n$.

We defer the complete proof to Appendix C, as it is similar to the proof of Theorem 5.

A.3 Construction of DRSS

Real Case The following theorem proves that the RSS we construct over real space is also a DRSS.

Theorem 14. *Let $n \in \mathbb{N}$, $d = \log^2\lambda + \log^2 n$ and $T = 515(\lambda+1)n$. The ensemble of unitary operators RSG^n defined in Definition 9 is a* DRSS.

To prove Theorem 14, recall the ensemble of unitary operators $\widetilde{\mathsf{RSG}}^n := \left\{\widetilde{\mathsf{RSG}}^{n,\lambda}\right\}_\lambda$ we define in Section 5.1. We have the following proposition for $\widetilde{\mathsf{RSG}}^n$.

Proposition 2. *For $T = 515(\lambda+1)n$, the ensemble of unitary operator $\widetilde{\mathsf{RSG}}^n$ is a* CRSS.

Proof. Note that a uniformly random $\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\tilde{f}_i)_{i=1}^T}$ corresponds to a T-step parallel Kac's walk on $\mathcal{S}_{\mathbb{R}}^{2^n}$. The proposition then follows from Theorem 5 and the definition of the CRSS. □

Let $|\eta\rangle \in \mathcal{S}(\mathcal{H})$ be an arbitrary real state. Set

$$\mathcal{N} = \left\{\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\eta\rangle\right\} \quad \text{and} \quad \widetilde{\mathcal{N}} = \left\{\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\tilde{f}_i)_{i=1}^T}|\eta\rangle\right\} \quad . \tag{33}$$

We need the following two lemmas. Both proofs are deferred to Appendix C.

Lemma 13. $\widetilde{\mathcal{N}} = \mathcal{S}_{\mathbb{R}}^{2^n}$.

Lemma 14. *There exists an $\epsilon = \mathrm{negl}(\lambda)$ such that $\mathcal{N}$ is an ϵ-net for real vectors in $\mathcal{S}(\mathcal{H})$, where $\mathcal{N}$ is defined in Eq. (33).*

Proof of Theorem 14. It is easy to see that the uniformity condition is satisfied. Let κ denote the key length. Quantum circuit $\mathsf{RSG}^{n,\lambda}$ applies $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$ after reading $(\sigma_i)_{i=1}^T$ and $(f_i)_{i=1}^T$. To implement $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$, we need to realize each of the $T = 515(\lambda+1)n$ unitary gates K. Since each gate K can be implemented in $\mathrm{poly}(n,\lambda,\kappa)$ time, the total construction time for $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}$ is also $\mathrm{poly}(n,\lambda,\kappa)$.

Combining with Lemma 14, it suffices to prove that there exists a good distribution $\tilde{\nu}$ satisfying the requirement in Definition 13. Fix $|\eta\rangle \in \mathcal{S}(\mathcal{H})$. Define three distributions:

- ν be the distribution of $\mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.
- $\tilde{\nu}$ be the distribution of $\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\tilde{f}_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(\tilde{f}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$.

– μ be the Haar measure on $\mathcal{S}_{\mathbb{R}}^{2^n}$.

Note that $\widetilde{\nu}$ is the output distribution of T-step parallel Kac's walk. Thus by Theorem 5, we have

$$\|\widetilde{\nu} - \mu\|_{\text{TV}} \leq \frac{1}{2^{\lambda n-1}} = \text{negl}(\lambda) \quad . \tag{34}$$

We are left to show the Wasserstein ∞-distance between ν and $\widetilde{\nu}$ is negligible. To this end, we construct a coupling γ_0 of ν and $\widetilde{\nu}$ by using the same permutation σ_t and letting f_t be the function satisfying $f_t(y)$ is the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$. Therefore

$$\begin{aligned} & W_\infty(\nu, \widetilde{\nu}) \\ = & \lim_{p\to\infty} \left(\inf_{\gamma\in\Gamma(\nu,\widetilde{\nu})} \mathop{\mathbb{E}}_{(|v\rangle,|u\rangle)\sim\gamma} [\||v\rangle - |u\rangle\|_2^p] \right)^{1/p} \\ \leq & \lim_{p\to\infty} \left(\mathop{\mathbb{E}}_{(|v\rangle,|u\rangle)\sim\gamma_0} [\||v\rangle - |u\rangle\|_2^p] \right)^{1/p} \overset{(\text{Eq. }(62))}{\leq} \frac{1030\pi(\lambda+1)n}{\lambda^{\log\lambda} n^{\log n}} = \text{negl}(\lambda) \quad . \end{aligned} \tag{35}$$

This completes the proof. □

Complex Case Likewise, the ensemble of unitary operators built over complex space in Section 5.2 turns out to be a DRSS. The proof is provided in Appendix C.

Theorem 15. *Let $n \in \mathbb{N}$, $d = 2\left(\log^2\lambda + \log^2 n\right)$ and $T = 515(\lambda+1)n$. The ensemble of unitary operators* RSGC^n *defined in Definition 11 is a* DRSS.

B Connections with Existing PRS variants

In this section, we formally demonstrate that the existing PRS and its variants can all be constructed from PRSS in a black-box manner.

(Scalable) Pseudorandom States. Originally, the definition of PRS in [31] identifies the number of qubits n as the security parameter. Consequently, the security of PRS is not guaranteed when n is small. This issue was addressed by [11] through defining *scalable* PRSG*s*, in which n and λ are treated separately. This allows for the tuning of security when $n < \lambda$. We rephrase the scalable definition in a style that is congruent with PRSS.

Definition 17 ((Scalable) PRSG). *Let $\mathcal{K} = \{0,1\}^\kappa$ be a key space, $\mathcal{H}$ be a Hilbert space of dimension 2^n with $n \in \mathbb{N}$, λ be a security parameter. A* (scalable) pseudorandom state generator (PRSG) *is an ensemble of unitaries*

$$\mathcal{G}^n := \{\{\mathcal{G}_k^{n,\lambda} : \mathcal{H} \to \mathcal{H}\}_{k\in\mathcal{K}}\}_\lambda$$

satisfying

- ***Pseudorandomness.*** *Any polynomially many ℓ copies of $|\phi_k\rangle$ with the same random k is computationally indistinguishable from the same number of copies of a Haar random state. More precisely, for any $n \in \mathbb{N}$, any efficient quantum algorithm $\mathcal{A}$ and any $\ell \in \mathrm{poly}(\lambda)$,*

$$\left|\Pr_{k\leftarrow\mathcal{K}}\left[\mathcal{A}\Big(|\phi_k\rangle^{\otimes\ell}\Big) = 1\right] - \Pr_{|\psi\rangle\leftarrow\mu}\left[\mathcal{A}\Big(|\psi\rangle^{\otimes\ell}\Big) = 1\right]\right| = \mathrm{negl}(\lambda) \quad ,$$

  where $|\phi_k\rangle := \mathcal{G}_\kappa^{n,\lambda}|0^n\rangle$ and μ is the Haar measure on $\mathcal{S}(\mathcal{H})$.
- ***Uniformity.*** *$\mathcal{G}^n$ can be uniformly computed in polynomial time. That is, there is a deterministic Turing machine that, on input $(1^n, 1^\lambda, 1^\kappa)$, outputs a quantum circuit Q in $\mathrm{poly}(n, \lambda, \kappa)$ time such that for all $k \in \mathcal{K}$ and $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\mathrm{in}})$*

$$Q\,|k\rangle|\phi\rangle = |k\rangle|\phi_k\rangle \ ,$$

  where $|\phi_k\rangle := \mathcal{G}_k^{n,\lambda}|\phi\rangle$.
- ***Polynomially-bounded key length.*** *$\kappa = \log|\mathcal{K}| = \mathrm{poly}(m, \lambda)$. As a result, $\mathcal{G}^n$ can be computed efficiently in time $\mathrm{poly}(n, \lambda)$.*

We informally call the keyed family of quantum states $\{|\phi_k\rangle\}_{k\in\mathcal{K}}$ a (scalable) pseudorandom quantum state *in $\mathcal{H}$.*

The existence of PRSSs implies the existence of (scalable) PRSGs [11,31] straightforwardly by definition.

Lemma 15. *If $\mathcal{R}^{n,m}$ is a* PRSS *from $\mathcal{H}_{\mathrm{in}}$ to $\mathcal{H}_{\mathrm{out}}$ over a key space $\mathcal{K}$ with security parameter λ, $\left\{\mathcal{R}_k^{n,m,\lambda}\,|\phi\rangle\right\}_{k\in\mathcal{K}}$ is a (scalable)* PRS *in $\mathcal{H}_{\mathrm{out}}$ for any $|\phi\rangle \in \mathcal{S}(\mathcal{H}_{\mathrm{in}})$.*

Pseudorandom Function-like States. We recall the definition of the pseudorandom function-like states generator with three levels of security:

Definition 18 (PRFSG [4,5]). *Let $\mathcal{K} = \{0,1\}^\lambda$ be a key space. Let $\mathcal{C} = \{0,1\}^n$ be a classical input space and $\mathcal{H}$ be a Hilbert space of dimension 2^m. A pair of* $\mathrm{poly}(\lambda, m)$*-time quantum algorithm (K, G) is a* pseudorandom function-like state *generator if the following holds:*

- ***Key Generation.*** *For all $\lambda \in \mathbb{N}$, $K(1^\lambda)$ outputs a uniform key $k \in \mathcal{K}$.*
- ***State Generation.*** *For all $k \in \mathcal{K}$ and $x \in \mathcal{C}$, $G(1^\lambda, k, x)$ computes $|\phi_{x,k}\rangle \in \mathcal{S}(\mathcal{H})$.*
- ***Pseudorandomness.*** *The pseudorandomness can be defined in three different levels (from weaker to stronger):*
  - ***Selective security.*** *For any $\lambda \in \mathbb{N}$, $s \in \mathrm{poly}(\lambda)$, $\ell \in \mathrm{poly}(\lambda)$, any efficient quantum algorithm (non-uniform) $\mathcal{A}$ and a set of pre-declared input $\{x_1, \ldots, x_s\} \subseteq \mathcal{C}$,*

    $$\Bigg| \Pr_{k \leftarrow \mathcal{K}}\left[\mathcal{A}(x_1, \ldots, x_s, |\phi_1\rangle^{\otimes \ell}, \ldots, |\phi_s\rangle^{\otimes \ell}) = 1\right] - \Pr_{|\psi_1\rangle, \ldots, |\psi_s\rangle \leftarrow \mu}\left[\mathcal{A}(x_1, \ldots, x_s, |\psi_1\rangle^{\otimes \ell}, \ldots, |\psi_s\rangle^{\otimes \ell}) = 1\right] \Bigg| \leq \mathrm{negl}(\lambda),$$

    where, for each i, $|\phi_i\rangle$ denotes $|\phi_{x_i,k}\rangle$ generated by G; and $|\psi_i\rangle$ is a Haar random state.
  - ***Adaptive security.*** *Given* classical-access *to either the* PRFS *oracle $\mathcal{O}_{\mathsf{PRFS}}$ or the Haar-random oracle $\mathcal{O}_{\mathsf{Haar}}$, for any $\lambda \in \mathbb{N}$, any efficient quantum algorithm (non-uniform) $\mathcal{A}$ with polynomial length quantum advice ρ_λ,*

    $$\left| \Pr_{k \leftarrow \mathcal{K}}\left[\mathcal{A}^{\mathcal{O}_{\mathsf{PRFS}}(k,\cdot)}(\rho_\lambda) = 1\right] - \Pr_{\mathcal{O}_{\mathsf{Haar}}}\left[\mathcal{A}^{\mathcal{O}_{\mathsf{Haar}}(\cdot)}(\rho_\lambda) = 1\right] \right| \leq \mathrm{negl}(\lambda),$$

    where on input $x \in \mathcal{C}$, $\mathcal{O}_{\mathsf{PRFS}}(k, \cdot)$ outputs $G(1^\lambda, k, x)$; and $\mathcal{O}_{\mathsf{Haar}}(\cdot)$ outputs a Haar random $|\psi_x\rangle$.
  - ***Quantum-accessible adaptive security.*** *Given* quantum-access *to a* PRFS *or Haar-random oracle, for any $\lambda \in \mathbb{N}$, for any efficient quantum algorithm (non-uniform) $\mathcal{A}$ with polynomial length quantum advice ρ_λ,*

    $$\left| \Pr_{k \leftarrow \mathcal{K}}\left[\mathcal{A}^{|\mathcal{O}_{\mathsf{PRFS}}(k,\cdot)\rangle}(\rho_\lambda) = 1\right] - \Pr_{\mathcal{O}_{\mathsf{Haar}}}\left[\mathcal{A}^{|\mathcal{O}_{\mathsf{Haar}}(\cdot)\rangle}(\rho_\lambda) = 1\right] \right| \leq \mathrm{negl}(\lambda),$$

    where on input a n-qubit register X, $\mathcal{O}_{\mathsf{PRFS}}(k, \cdot)$ applies a channel that controlled on the register X containing x, and stores $G(1^\lambda, k, x)$ in the register Y, then output (X, Y); instead, $\mathcal{O}_{\mathsf{Haar}}(\cdot)$ stores a Haar random $|\psi_x\rangle\langle\psi_x|$ on the register Y, then output (X, Y).

As previously mentioned, the security of our PRSS is maintained when applied to any arbitrary initial pure state, making it straightforward to derive a PRFSG with selective security.

Lemma 16. *Assume $\mathcal{R}^{n,m}$ is a* PRSS *from $\mathcal{H}_{\text{in}}$ to $\mathcal{H}_{\text{out}}$ over a key space $\mathcal{K}$ with security parameter λ s.t. $n = O(\log \lambda)$. Construct $(\hat{K}, \hat{G})$ in the following way:*

- *(Key generation.) For all $\lambda \in \mathbb{N}$, $\hat{K}(1^\lambda)$ sets a large enough $s \in \text{poly}(\lambda)$ and generates a key $k = \{k_1, \ldots, k_s\}$ such that for $i \in [s]$, k_i is chosen uniformly and independently from $\mathcal{K}$;*
- *(State generation.) For all k and classical queries $\{x_i \in \mathcal{C}\}_{i=1}^s$, $\hat{G}(1^\lambda, k, x_i)$ computes $|\phi_i\rangle = \mathcal{R}_{k_i}^{n,m,\lambda} |x_i\rangle$.*

Then, $(\hat{K}, \hat{G})$ is a PRFSG *with selective security.*

Proof. The algorithm $(\hat{K}, \hat{G})$ is efficient as it essentially performs PRSS a polynomial number of times. Meanwhile, the indistinguishability holds because all output states are obtained via independent keys. □

When we consider log-inputs by restricting $n = O(\log \lambda)$ and setting $s = O(2^n) \in \text{poly}(\lambda)$, the construction in Lemma 16 produces sufficient key segments to ensure that every $x \in \mathcal{C}$ has its own independent key, without assuming the number of queries from the adversary. As a consequence, log-input PRFSGs with adaptive security can be achieved through the use of PRSSs. Furthermore, as demonstrated by the results in [50], quantum superposition queries over the input domain do not provide any additional advantages when the output state is known for every input string, which results in quantum-accessible adaptive security.

Lemma 17. *If $\mathcal{R}^{n,m}$ is a* PRSS *from $\mathcal{H}_{\text{in}}$ to $\mathcal{H}_{\text{out}}$ over a key space $\mathcal{K}$ with security parameter λ s.t. $n = O(\log \lambda)$, then $(\hat{K}, \hat{G})$ is a* PRFSG *satisfying adaptive security and quantum-accessible adaptive security.*

Prior works [4,5] have demonstrated that log-input PRFSGs can be constructed from a PRS. However, this approach requires a *test procedure* involving a post-selection among the classical input domain, which introduces errors and incurs computation overhead. Our PRSS, with its flexibility in choosing initial states, provides a cleaner method. It is worth noting that PRFSGs on long inputs (i.e., exponentially large domain) may appear stronger and might not be achievable from either PRSs or PRSSs in a black-box manner.

C Deferred Proofs

C.1 Proof of Lemma 1

Proof of Lemma 1. We demonstrate that the real case and the complex case can be established using the same approach.

Let $\zeta^+ - \zeta^-$ be the Hahn decomposition of the signed measure $\mu - \nu$. Then

$$\begin{aligned}
&\left\| \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[(|\psi\rangle\langle\psi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\varphi\rangle\sim\nu}\left[(|\varphi\rangle\langle\varphi|)^{\otimes l}\right] \right\|_1 \\
= &\left\| \int_{\mathcal{S}_{\mathbb{R}}^{2^n}} (|u\rangle\langle u|)^{\otimes l} (\mu - \nu)(\mathrm{d}\,|u\rangle) \right\|_1 \\
= &\left\| \int_{\mathcal{S}_{\mathbb{R}}^{2^n}} (|u\rangle\langle u|)^{\otimes l} \left(\zeta^+ - \zeta^-\right)(\mathrm{d}\,|u\rangle) \right\|_1 \\
\leq &\left\| \int_{\mathcal{S}_{\mathbb{R}}^{2^n}} (|u\rangle\langle u|)^{\otimes l} \zeta^+(\mathrm{d}\,|u\rangle) \right\|_1 + \left\| \int_{\mathcal{S}_{\mathbb{R}}^{2^n}} (|u\rangle\langle u|)^{\otimes l} \zeta^-(\mathrm{d}\,|u\rangle) \right\|_1 \\
\leq &\int_{\mathcal{S}_{\mathbb{R}}^{2^n}} \left\|(|u\rangle\langle u|)^{\otimes l}\right\|_1 \zeta^+(\mathrm{d}\,|u\rangle) + \int_{\mathcal{S}_{\mathbb{R}}^{2^n}} \left\|(|u\rangle\langle u|)^{\otimes l}\right\|_1 \zeta^-(\mathrm{d}\,|u\rangle) \\
= &\ \zeta^-\left(\mathcal{S}_{\mathbb{R}}^{2^n}\right) + \zeta^-\left(\mathcal{S}_{\mathbb{R}}^{2^n}\right) \\
= &\ \|\mu - \nu\|_{\mathrm{TV}} \quad .
\end{aligned}$$

□

C.2 Proof of Lemma 6

Proof of Lemma 6. We first prove the case that $F : \mathcal{K}_1 \times \mathcal{X}_1 \to \mathcal{Y}_1$ and $G : \mathcal{K}_2 \times \mathcal{X}_2 \to \mathcal{Y}_2$ are both QPRFs. We prove the lemma by contradiction. Suppose there exists a polynomial-time quantum oracle algorithm $\mathcal{A}$ who queries q times such that

$$\left| \Pr_{k_1\leftarrow\mathcal{K}_1, k_2\leftarrow\mathcal{K}_1}\left[\mathcal{A}^{F_{k_1}, G_{k_2}}\left(1^\lambda\right) = 1\right] - \Pr_{f\leftarrow\mathcal{Y}_1^{\mathcal{X}_1}, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{f,g}\left(1^\lambda\right) = 1\right] \right| = \frac{1}{\mathrm{poly}(\lambda)} \quad .$$

By the triangle inequality,

$$\begin{aligned}
&\left| \Pr_{k_1\leftarrow\mathcal{K}_1, k_2\leftarrow\mathcal{K}_1}\left[\mathcal{A}^{F_{k_1}, G_{k_2}}\left(1^\lambda\right) = 1\right] - \Pr_{k_1\leftarrow\mathcal{K}_1, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{F_{k_1}, g}\left(1^\lambda\right) = 1\right] \right| \\
&\quad + \left| \Pr_{k_1\leftarrow\mathcal{K}_1, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{F_{k_1}, g}\left(1^\lambda\right) = 1\right] - \Pr_{f\leftarrow\mathcal{Y}_1^{\mathcal{X}_1}, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{f,g}\left(1^\lambda\right) = 1\right] \right| \geq \frac{1}{\mathrm{poly}(\lambda)} \quad .
\end{aligned}$$

Without loss of generality, we assume

$$\left| \Pr_{k_1\leftarrow\mathcal{K}_1, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{F_{k_1}, g}\left(1^\lambda\right) = 1\right] - \Pr_{f\leftarrow\mathcal{Y}_1^{\mathcal{X}_1}, g\leftarrow\mathcal{Y}_2^{\mathcal{X}_2}}\left[\mathcal{A}^{f,g}\left(1^\lambda\right) = 1\right] \right| \geq \frac{1}{2\cdot\mathrm{poly}(\lambda)} \quad .$$

Thus we can define a polynomial-time quantum oracle algorithm $\mathcal{A}'$ that is able to distinguish F_{k_1} from a random function f. Once $\mathcal{A}'$ gets an oracle access to some function $h \in \{F_{k_1}, f\}$, it simulates the execution of $\mathcal{A}$ with oracle access to h and a random function g. Since $\mathcal{A}$ makes at most q queries, $\mathcal{A}'$ can efficiently simulate a random oracle using $2q$-wise independent function (see [48, Theorem 6.1]). This contradicts the quantum-security property of F.

To prove the case that $F : \mathcal{K}_1 \times \mathcal{X}_1 \to \mathcal{Y}_1$ is a QPRF and $G : \mathcal{K}_2 \times \mathcal{X}_2 \to \mathcal{X}_2$ is a QPRP, we may assume as above that there exists a polynomial-time quantum oracle algorithm $\mathcal{A}$ who queries q times such that

$$\left| \Pr_{k_1 \leftarrow \mathcal{K}_1, g \leftarrow S_{\mathcal{X}_2}} \left[\mathcal{A}^{F_{k_1}, g}(1^\lambda) = 1\right] - \Pr_{f \leftarrow \mathcal{Y}_1^{\mathcal{X}_1}, g \leftarrow S_{\mathcal{X}_2}} \left[\mathcal{A}^{f, g}(1^\lambda) = 1\right] \right| \geq \frac{1}{2 \cdot \mathrm{poly}(\lambda)} .$$

Then we define the following efficient algorithm $\mathcal{A}''$ to distinguish a QPRF from a random function:

1. Given $h \in \{F_{k_1}, f\}$, it chooses a permutation g uniformly at random from a $2q$-wise almost independent family of permutations to simulate a random permutation oracle. This sampling procedure can be done in polynomial time (see [34, Theorem 5.9]).
2. It simulates $\mathcal{A}$ with oracle access to h and g, and outputs what $\mathcal{A}$ returns.

This breaks the quantum-security property of QPRFs.

□

C.3 Proof of Theorem 6

To prove Theorem 6, we extend the the proportional coupling introduced in real case to complex case. In the proportional coupling, the real case lets $(X_t[i], X_t[j])$ be collinear with $(Y_t[i], Y_t[j])$ so that the distance from X_t to Y_t is reduced by a constant factor in each step. However in complex case, $(X_t[i], X_t[j])$ is actually a two dimensional complex vector and this makes it unsuitable to adopt the previous approach directly. To deal with this, in the complex case, we let $(|X_t[i]|, |X_t[j]|)$ align collinearly with $(|Y_t[i]|, |Y_t[j]|)$ in real two dimensional real plane, and make $X_t[i]$ and $Y_t[i]$, as well as $X_t[j]$ and $Y_t[j]$, share the same argument in complex plane in the meantime. We assume $n = 2m$.

Proportional Coupling

Definition 19 (Proportional Coupling). *We define a coupling of two copies $\{X_t\}_{t\geq 0}, \{Y_t\}_{t\geq 0}$ of parallel Kac's walk on complex vectors in the following way: Fix X_t, $Y_t \in \mathbb{C}^n$.*

1. *Choose a perfect matching of $[n]$, denoted by $P_t = \left\{ \left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right) \right\}$, uniformly at random.*
2. *Let $X_{t,1} = X_t$ and $Y_{t,1} = Y_t$. For every $1 \leq k \leq m$:*

(a) let $l_k^{(t)} = \sqrt{\left|X_{t,k}[i_k^{(t)}]\right|^2 + \left|X_{t,k}[j_k^{(t)}]\right|^2}$ *and* $l'^{(t)}_k = \sqrt{\left|Y_{t,k}[i_k^{(t)}]\right|^2 + \left|Y_{t,k}[j_k^{(t)}]\right|^2}$. *Let* U_0 *and* U_0' *be the unitary operators in* $\mathrm{SU}(2)$ *which satisfy*

$$U_0 \begin{pmatrix} X_{t,k}[i_k^{(t)}] \\ X_{t,k}[j_k^{(t)}] \end{pmatrix} = \begin{pmatrix} l_k^{(t)} \\ 0 \end{pmatrix} \quad \text{and} \quad U_0' \begin{pmatrix} Y_{t,k}[i_k^{(t)}] \\ Y_{t,k}[j_k^{(t)}] \end{pmatrix} = \begin{pmatrix} l'^{(t)}_k \\ 0 \end{pmatrix} \ .$$

(b) pick $\alpha_k^{(t)}, \beta_k^{(t)} \in [0, 2\pi)$ *and* $\zeta_k^{(t)} \in [0,1)$ *uniformly at random and set*

$$\theta_k^{(t)} = \arcsin \sqrt{\zeta_k^{(t)}} \ . \tag{36}$$

(c) set

$$X_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha_k^{(t)}, \beta_k^{(t)}, \theta_k^{(t)}, U_0 X_{t,k}\right) \ ,$$
$$Y_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha_k^{(t)}, \beta_k^{(t)}, \theta_k^{(t)}, U_0' Y_{t,k}\right) \ .$$

3. Finally, set $X_{t+1} = X_{t,m+1}$ *and* $Y_{t+1} = Y_{t,m+1}$.

Remark 2. In step 2(a), if $l_k^{(t)} \neq 0$, $X_{t,k}[i_k^{(t)}] = r_1 e^{\mathrm{i}\gamma}$ and $X_{t,k}[j_k^{(t)}] = r_2 e^{\mathrm{i}\delta}$ with $r_1, r_2 \in [0,1]$ and $\gamma, \delta \in [0, 2\pi)$, then U_0 is $U(\alpha_0, \beta_0, \theta_0)$ where

$$\alpha_0 = -\gamma \ , \quad \beta_0 = \pi - \delta \ , \quad \theta_0 = \arccos \frac{r_1}{\sqrt{r_1^2 + r_2^2}} \ .$$

If $l_k^{(t)} = 0$, U_0 can be arbitrary matrix in SU(2). The same applies to U_0'.

Due to the unitary invariance of Haar measures, UU_0 is Haar distributed on SU(2) for a random random U sampled according to Haar measure on SU(2). This guarantees that the proportional coupling is indeed a valid coupling.

Remark 3. The proportional coupling forces $X_{t+1}[i], Y_{t+1}[i]$ and the original point to be collinear in the complex plane. In other words, $X_{t+1}[i]$ and $Y_{t+1}[i]$ have the same argument. Specifically, we can write

$$X_{t+1}[i] = e^{i\phi} l \quad \text{and} \quad Y_{t+1}[i] = e^{i\phi} l'$$

for some $l, l' \in [0,1], \phi \in [0, 2\pi)$.

Through proportional coupling , X and Y approach each other at an exponential rate. Formally, we have

Lemma 18. *Let* $X_0, Y_0 \in \mathcal{S}_{\mathbb{C}}^n$. *For* $t \geq 0$, *we couple* (X_{t+1}, Y_{t+1}) *conditional on* (X_t, Y_t) *according to the proportional coupling defined in Definition 19. We define*

$$A_t[i] = |X_t[i]|^2 \ , \quad B_t[i] = |Y_t[i]|^2 \ .$$

Then for any $l \in \mathbb{N}$, *we have*

$$\mathbb{E}\left[\sum_{i=1}^{n} (A_l[i] - B_l[i])^2\right] \leq 2 \cdot \left(\frac{2}{3}\right)^l \ .$$

Proof of Lemma 18. Fix $X_t, Y_t \in \mathcal{S}^n_{\mathbb{C}}$. Let (X_{t+1}, Y_{t+1}) obtained from (X_t, Y_t) by applying the coupling defined in Definition 19. Recall that $n = 2m$. Let $N = \frac{n!}{2^m m!}$ be the number of perfect matchings for $[n]$. A perfect matching $\left\{\left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right)\right\}$ of $[n]$ at step t is denoted by $\left(\overrightarrow{i^{(t)}}, \overrightarrow{j^{(t)}}\right)$.

We have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n}(A_{t+1}[i]-B_{t+1}[i])^2\right]\\
=&\frac{1}{N}\sum_{\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)}\underbrace{\mathbb{E}\left[\sum_{i=1}^{n}(A_{t+1}[i]-B_{t+1}[i])^2\,\middle|\,P_t=\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)\right]}_{(\star)}\ .
\end{aligned} \tag{37}$$

By the definition of the parallel Kac's walk on complex vectors, we have

$$\begin{aligned}
(\star)=&\sum_{k=1}^{m}\mathbb{E}\left[\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)\cos(\theta_k^{(t)})^2-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\cos(\theta_k^{(t)})^2\right)^2\right]\\
&+\sum_{k=1}^{m}\mathbb{E}\left[\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)\sin(\theta_k^{(t)})^2-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\sin(\theta_k^{(t)})^2\right)^2\right]\\
=&\frac{2}{3}\sum_{k=1}^{m}\left(\left(A_t[i_k^{(t)}]+A_t[j_k^{(t)}]\right)-\left(B_t[i_k^{(t)}]+B_t[j_k^{(t)}]\right)\right)^2\\
=&\underbrace{\frac{2}{3}\sum_{k=1}^{m}\left(\left(A_t[i_k^{(t)}]-B_t[i_k^{(t)}]\right)^2+\left(A_t[j_k^{(t)}]-B_t[j_k^{(t)}]\right)^2\right)}_{(\star\star)}\\
&+\underbrace{\frac{2}{3}\sum_{k=1}^{m}2\left(A_t[i_k^{(t)}]-B_t[i_k^{(t)}]\right)\left(A_t[j_k^{(t)}]-B_t[j_k^{(t)}]\right)}_{(\star\star\star)}\ ,
\end{aligned} \tag{38}$$

where the second equality is by $\mathbb{E}\left[\cos(\theta_k^{(t)})^4\right]=\mathbb{E}\left[\sin(\theta_k^{(t)})^4\right]=1/3$.

As $\left\{\left(i_1^{(t)}, j_1^{(t)}\right), \ldots, \left(i_m^{(t)}, j_m^{(t)}\right)\right\}$ is a perfect matching, we have

$$(\star\star)=\frac{2}{3}\sum_{i=1}^{n}(A_t[i]-B_t[i])^2\ . \tag{39}$$

Combing Eqs. (37)(38)(39), we obtain

$$\mathbb{E}\left[\sum_{i=1}^{n}(A_{t+1}[i]-B_{t+1}[i])^2\right]=\frac{2}{3}\sum_{i=1}^{n}(A_t[i]-B_t[i])^2+\underbrace{\frac{1}{N}\sum_{\left(\overrightarrow{i^{(t)}},\overrightarrow{j^{(t)}}\right)}(\star\star\star)}_{(4\star)}\ . \tag{40}$$

Using the same calculation in Eq. (7), we have

$$(4\star)=-\frac{2}{3(n-1)}\sum_{i=1}^{n}(A_t[i]-B_t[i])^2\ . \tag{41}$$

Combining Eqs. (40)(41), we have

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}(A_l[i]-B_l[i])^2\right]&=\mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{n}(A_l[i]-B_l[i])^2\,\middle|\,X_{l-1},Y_{l-1}\right]\right]\\&\le\frac{2}{3}\,\mathbb{E}\left[\sum_{i=1}^{n}(A_{l-1}[i]-B_{l-1}[i])^2\right]\\&\le\left(\frac{2}{3}\right)^l\sum_{i=1}^{n}(A_0[i]-B_0[i])^2\le 2\cdot\left(\frac{2}{3}\right)^l\ .\end{aligned}$$

□

Proof of the Mixing Time

Proof of Theorem 6. Let $T=10(c+1)\log n$ for $c>0$. We couple two copies $\{X_t\}_{t\ge 0}$ and $\{Y_t\}_{t\ge 0}$ of the parallel Kac's walk with starting points $X_0=x\in\mathcal{S}^n_{\mathbb{C}}$ and $Y_0\sim\mu$, by applying the proportional coupling. We have

$$W_1(\mathcal{L}(X_T),\mu)=W_1(\mathcal{L}(X_T),\mathcal{L}(Y_T))\le\mathbb{E}[\|X_T-Y_T\|_2]\le\left(\mathbb{E}\left[\|X_T-Y_T\|_2^4\right]\right)^{1/4}\ .$$

Then by Cauthy-Schwarz inequality, we have

$$W_1(\mathcal{L}(X_T),\mu)\le\left(n\,\mathbb{E}\left[\|X_T-Y_T\|_4^4\right]\right)^{1/4}\ . \tag{42}$$

Note that the proportional coupling forces $X_T[i]$ and $Y_T[i]$ share the same argument for all $i\in[n]$. Therefore, for all $i\in[n]$

$$|X_T[i]-Y_T[i]|=||X_T[i]|-|Y_T[i]||\le|X_T[i]|+|Y_T[i]|\ .$$

This gives us

$$\|X_T-Y_T\|_4^4=\sum_{i=1}^{n}|X_T[i]-Y_T[i]|^4\le\sum_{i=1}^{n}\left(|X_T[i]|^2-|Y_T[i]|^2\right)^2\ . \tag{43}$$

Combing Eqs. (42) and (43), we have

$$W_1(\mathcal{L}(X_T),\mu) \le \left(n\,\mathbb{E}\left[\sum_{i=1}^{n}\left(|X_T[i]|^2 - |Y_T[i]|^2\right)^2\right]\right)^{1/4}$$
$$(\text{Lemma 18}) \le \left(2n\left(\frac{2}{3}\right)^T\right)^{1/4} \le \frac{1}{2^{c\log n}} \ .$$

□

C.4 Proof of Theorem 10

To prove Theorem 10, we first introduce a new ensemble of (infinitely many) unitary operators $\widetilde{\mathsf{RSGC}}^n := \left\{\widetilde{\mathsf{RSGC}}^{n,\lambda}\right\}_\lambda$ with $\widetilde{\mathsf{RSGC}}^{n,\lambda} :=$

$$\left\{\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}\right\}_{(\sigma_i)_{i=1}^T\subseteq S_{2^n},(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T\subseteq\{f:\{0,1\}^{n-1}\to[0,1)\}}$$

and

$$\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T} = \widetilde{L}_{\tau_T,\widetilde{f}_T,\widetilde{g}_T,\widetilde{h}_T}\cdots\widetilde{L}_{\tau_2,\widetilde{f}_2,\widetilde{g}_2,\widetilde{h}_2}\widetilde{L}_{\tau_1,\widetilde{f}_1,\widetilde{g}_1,\widetilde{h}_1}$$

where $\widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}} = U_{\sigma^{-1}}\widetilde{Q}_{\widetilde{f},\widetilde{g},\widetilde{h}}U_\sigma$ and $\widetilde{Q}_{\widetilde{f},\widetilde{g},\widetilde{h}}$ is defined to be

$$\widetilde{Q}_{\widetilde{f},\widetilde{g},\widetilde{h}} = \sum_{y\in\{0,1\}^{n-1}} U\left(\widetilde{\alpha}_y,\widetilde{\beta}_y,\widetilde{\theta}_y\right)\otimes|y\rangle\langle y| \ , \tag{44}$$

in which $U(\alpha,\beta,\theta)$ is defined in (10) and for any $y\in\{0,1\}^{n-1}$

$$\widetilde{\theta}_y = \arcsin\left(\sqrt{\widetilde{f}(y)}\right) \ , \ \widetilde{\alpha}_y = 2\pi\cdot\widetilde{g}(y) \ , \ \widetilde{\beta}_y = 2\pi\cdot\widetilde{h}(y) \ .$$

Similar to the real case, $\widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}}$ represents one step of parallel Kac's walk in complex space for independently and uniformly random $\sigma,\widetilde{f},\widetilde{g}$ and $\widetilde{h}$.

Lemma 19. *Let $\sigma\in S_{2^n}$ and $\widetilde{f},\widetilde{g},\widetilde{h}:\{0,1\}^{n-1}\to[0,1)$. Let $f:\{0,1\}^{n-1}\to\{0,1\}^d$ be the function satisfying for any $y\in\{0,1\}^{n-1}$, $f(y)$ is the d digits after the binary point in $\widetilde{f}(y)$. The same applies to g and h. Then*

$$\left\|L_{\sigma,f,g,h} - \widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}}\right\|_\infty \le 2^{6-\frac{d}{2}} \ ,$$

where $L_{\sigma,f,g,h} = U_{\sigma^{-1}}Q_{f,g,h}U_\sigma$ is defined in (17) and $\widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}} = U_{\sigma^{-1}}\widetilde{Q}_{\widetilde{f},\widetilde{g},\widetilde{h}}U_\sigma$ is defined in (44).

Proof of Lemma 19. Recall the unitary $\widehat{L}_{\sigma,f,g,h} = U_{\sigma^{-1}}\widehat{Q}_{f,g,h}U_\sigma$ defined in (15). We will prove

– $\left\| L_{\sigma,f,g,h} - \widehat{L}_{\sigma,f,g,h} \right\|_\infty \le 2^{3-d}\pi$.
– $\left\| \widehat{L}_{\sigma,f,g,h} - \widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}} \right\|_\infty \le 2^{4-\frac{d}{2}}$.

The claim then follows from the triangle inequality.

Proof of the first bound Fix a $y \in \{0,1\}^{n-1}$, we have

$$\left\| \begin{pmatrix} \cos(\theta_y) & -\sin(\theta_y) \\ \sin(\theta_y) & \cos(\theta_y) \end{pmatrix} - \begin{pmatrix} \cos\left(\frac{\pi}{2}\xi_y\right) & -\sin\left(\frac{\pi}{2}\xi_y\right) \\ \sin\left(\frac{\pi}{2}\xi_y\right) & \cos\left(\frac{\pi}{2}\xi_y\right) \end{pmatrix} \right\|_\infty \le 2^{-d-1}\pi \ ,$$
$$\left\| \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} \end{pmatrix} - \begin{pmatrix} e^{\mathrm{i}2\pi\gamma_y^+} & 0 \\ 0 & e^{-\mathrm{i}2\pi\gamma_y^+} \end{pmatrix} \right\|_\infty \le 2^{1-d}\pi \ ,$$
$$\left\| \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} \end{pmatrix} - \begin{pmatrix} e^{\mathrm{i}2\pi\gamma_y^-} & 0 \\ 0 & e^{-\mathrm{i}2\pi\gamma_y^-} \end{pmatrix} \right\|_\infty \le 2^{1-d}\pi \ .$$

Thus, by the triangle inequality and the decomposition for matrix $U(\alpha,\beta,\theta)$, we have for any $y \in \{0,1\}^{n-1}$

$$\left\| U(\alpha_y,\beta_y,\theta_y) - U\Big(2\pi(\gamma_y^+ + \gamma_y^-), 2\pi(\gamma_y^+ - \gamma_y^-), \frac{\pi}{2}\xi_y\Big) \right\|_\infty \le 2^{3-d}\pi \ .$$

Therefore we have

$$\begin{aligned} &\left\| L_{\sigma,f,g,h} - \widehat{L}_{\sigma,f,g,h} \right\|_\infty = \left\| Q_{f,g,h} - \widehat{Q}_{f,g,h} \right\|_\infty \\ &= \max_{y\in\{0,1\}^{n-1}} \left\| U\Big(2\pi(\gamma_y^+ + \gamma_y^-), 2\pi(\gamma_y^+ - \gamma_y^-), \frac{\pi}{2}\xi_y\Big) - U(\alpha_y,\beta_y,\theta_y) \right\|_\infty \\ &\le 2^{3-d}\pi \ . \end{aligned}$$

Proof of the second bound Fix a $y \in \{0,1\}^{n-1}$, we have $|\alpha_y - \widetilde{\alpha}_y| \le 2^{1-d}\pi$, and $\left|\beta_y - \widetilde{\beta}_y\right| \le 2^{1-d}\pi$. Therefore,

$$\left\| \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y+\beta_y}{2}\right)} \end{pmatrix} - \begin{pmatrix} e^{\mathrm{i}\left(\frac{\widetilde{\alpha}_y+\widetilde{\beta}_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\widetilde{\alpha}_y+\widetilde{\beta}_y}{2}\right)} \end{pmatrix} \right\|_\infty \le 2^{1-d}\pi \ , \quad (45)$$
$$\left\| \begin{pmatrix} e^{\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\alpha_y-\beta_y}{2}\right)} \end{pmatrix} - \begin{pmatrix} e^{\mathrm{i}\left(\frac{\widetilde{\alpha}_y-\widetilde{\beta}_y}{2}\right)} & 0 \\ 0 & e^{-\mathrm{i}\left(\frac{\widetilde{\alpha}_y-\widetilde{\beta}_y}{2}\right)} \end{pmatrix} \right\|_\infty \le 2^{1-d}\pi \ . \quad (46)$$

Moreover, we have $\left|\mathrm{val}(f(y)) - \widetilde{f}(y)\right| \le 2^{-d}$ and thus

$$\begin{aligned}
&\left\| \begin{pmatrix} \cos(\theta_y) & -\sin(\theta_y) \\ \sin(\theta_y) & \cos(\theta_y) \end{pmatrix} - \begin{pmatrix} \cos\left(\widetilde{\theta}_y\right) & -\sin\left(\widetilde{\theta}_y\right) \\ \sin\left(\widetilde{\theta}_y\right) & \cos\left(\widetilde{\theta}_y\right) \end{pmatrix} \right\|_\infty \\
= &\left\| \begin{pmatrix} \sqrt{1-\mathrm{val}(f(y))} - \sqrt{1-\widetilde{f}(y)} & -\sqrt{\mathrm{val}(f(y))} + \sqrt{\widetilde{f}(y)} \\ \sqrt{\mathrm{val}(f(y))} - \sqrt{\widetilde{f}(y)} & \sqrt{1-\mathrm{val}(f(y))} - \sqrt{1-\widetilde{f}(y)} \end{pmatrix} \right\|_\infty \\
\le &\left\| \begin{pmatrix} \sqrt{1-\mathrm{val}(f(y))} - \sqrt{1-\widetilde{f}(y)} & -\sqrt{\mathrm{val}(f(y))} + \sqrt{\widetilde{f}(y)} \\ \sqrt{\mathrm{val}(f(y))} - \sqrt{\widetilde{f}(y)} & \sqrt{1-\mathrm{val}(f(y))} - \sqrt{1-\widetilde{f}(y)} \end{pmatrix} \right\|_2 \\
= &\sqrt{2\left(\sqrt{1-\mathrm{val}(f(y))} - \sqrt{1-\widetilde{f}(y)}\right)^2 + 2\left(\sqrt{\mathrm{val}(f(y))} - \sqrt{\widetilde{f}(y)}\right)^2} \le 2^{-\frac{d}{2}+\frac{3}{2}} \ ,
\end{aligned} \tag{47}$$

where the last inequality is by the following fact.

Fact 16. *For $a, b \in [0,1]$ and $d \in \mathbb{N}$, if $|a-b| \le 2^{-d}$, then $\left|\sqrt{a} - \sqrt{b}\right| \le 2^{-\frac{d}{2}+\frac{1}{2}}$.*

Proof. Let $\delta = 2^{-d+1}$. If $a \le \delta$ and $b \le \delta$, we have $\left|\sqrt{a}-\sqrt{b}\right| \le \max\left\{\sqrt{a}, \sqrt{b}\right\} \le \sqrt{\delta} = 2^{-\frac{d}{2}+\frac{1}{2}}$. If $a > \delta$ or $b > \delta$, we will have $a > \delta/2$ and $b > \delta/2$, and therefore

$$\left|\sqrt{a}-\sqrt{b}\right| \le \frac{1}{\sqrt{2\delta}} \cdot |a-b| \le 2^{-\frac{d}{2}-1} \ .$$

□

Hence, we have by (45), (46) and (47), for any $y \in \{0,1\}^{n-1}$,

$$\left\| U(\alpha_y, \beta_y, \theta_y) - U\left(\widetilde{\alpha}_y, \widetilde{\beta}_y, \widetilde{\theta}_y\right) \right\|_\infty \le 2^{4-\frac{d}{2}} \ .$$

Therefore,

$$\begin{aligned}
\left\| \widehat{L}_{\sigma,f,g,h} - \widetilde{L}_{\sigma,\widetilde{f},\widetilde{g},\widetilde{h}} \right\|_\infty &= \left\| \widehat{Q}_{f,g,h} - \widetilde{Q}_{\widetilde{f},\widetilde{g},\widetilde{h}} \right\|_\infty \\
&= \max_{y \in \{0,1\}^{n-1}} \left\| U(\alpha_y, \beta_y, \theta_y) - U\left(\widetilde{\alpha}_y, \widetilde{\beta}_y, \widetilde{\theta}_y\right) \right\|_\infty \\
&\le 2^{4-\frac{d}{2}} \ .
\end{aligned}$$

□

Proof of Theorem 10. It is easy to see that the uniformity condition is satisfied. Let κ denote the key length. Quantum circuit $\mathsf{RSGC}^{n,\lambda}$ applies the operator $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$ after reading $(\sigma_i)_{i=1}^T$, $(f_i)_{i=1}^T$, $(g_i)_{i=1}^T$ and $(h_i)_{i=1}^T$. To implement $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$, we need to realize each of the

$T = 10(\lambda+1)n$ unitary gates L. Since each gate L can be implemented in $\text{poly}(n,\lambda,\kappa)$ time, the total construction time for $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$ is also $\text{poly}(n,\lambda,\kappa)$.

Thus, it suffices to prove the requirement of *Statistical Pseudorandomness* is satisfied. Fix $|\eta\rangle \in \mathcal{S}(\mathcal{H})$. Define three distributions:

- ν be the distribution of $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(f_i)_{i=1}^T, (g_i)_{i=1}^T, (h_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.
- $\widetilde{\nu}$ be the distribution of $\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(\widetilde{f}_i)_{i=1}^T, (\widetilde{g}_i)_{i=1}^T, (\widetilde{h}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$.
- μ be the Haar measure on $\mathcal{S}(\mathcal{H})$.

We first proof the trace distance between ν and $\widetilde{\nu}$ is negligible. To this end, we construct a coupling γ_0 of ν and $\widetilde{\nu}$ by using the same permutation σ_t and letting f_t be the function satisfying $f_t(y)$ is the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$ (The same applies to g_t and h_t). Therefore, for any $(|\phi\rangle, |\varphi\rangle) \sim \gamma_0$, we have

$$\begin{aligned}
&\||\phi\rangle - |\varphi\rangle\|_2 \\
= &\left\|\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle - \widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle\right\|_2 \\
\le &\left\|\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T} - \widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}\right\|_\infty \\
\le &\ 2^{6-\frac{d}{2}}T = \frac{640(\lambda+1)n}{\lambda^{\log\lambda}\cdot n^{\log n}} \ ,
\end{aligned}$$

where the last inequality is from Fact 1 and Lemma 19. Thus, for any $l \in \text{poly}(\lambda, n)$

$$\begin{aligned}
&\left\|\mathop{\mathbb{E}}_{|\phi\rangle\sim\nu}\left[(|\phi\rangle\langle\phi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\varphi\rangle\sim\widetilde{\nu}}\left[(|\varphi\rangle\langle\varphi|)^{\otimes l}\right]\right\|_1 \\
\le &\mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}\left[\left\|(|\phi\rangle\langle\phi|)^{\otimes l} - (|\varphi\rangle\langle\varphi|)^{\otimes l}\right\|_1\right] \\
\le &\ l \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}[\||\phi\rangle\langle\phi| - |\varphi\rangle\langle\varphi|\|_1] \\
\le &\ l\left(\mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}[\||\phi\rangle(\langle\phi| - \langle\varphi|)\|_1] + \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}[\|(|\phi\rangle - |\varphi\rangle)\langle\varphi|\|_1]\right) \\
\le &\ 2l \mathop{\mathbb{E}}_{(|\phi\rangle,|\varphi\rangle)\sim\gamma_0}[\||\phi\rangle - |\varphi\rangle\|_2] \le \frac{1280(\lambda+1)nl}{\lambda^{\log\lambda}\cdot n^{\log n}} \ . \qquad (48)
\end{aligned}$$

As for the trace distance between $\widetilde{\nu}$ and μ, note that $\widetilde{\nu}$ is the output distribution of T-step parallel Kac's walk. Thus by Theorem 6, we have

$$W_1(\widetilde{\nu}, \mu) \le \frac{1}{2^{\lambda n}} \ .$$

So there exists a coupling of $\widetilde{\upsilon}$ and μ, denoted by γ_1, that achieves

$$\mathop{\mathbb{E}}_{(|\varphi\rangle,|\psi\rangle)\sim\gamma_1}[\||\varphi\rangle - |\psi\rangle\|_2] \leq \frac{3}{2^{\lambda n}} \ .$$

Therefore, similar to Eq. (48), we have for any $l \in \mathrm{poly}(\lambda, n)$

$$\left\|\mathop{\mathbb{E}}_{|\varphi\rangle\sim\widetilde{\upsilon}}\left[(|\varphi\rangle\!\langle\varphi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[(|\psi\rangle\!\langle\psi|)^{\otimes l}\right]\right\|_1 \leq 2l \mathop{\mathbb{E}}_{(|\varphi\rangle,|\psi\rangle)\sim\gamma_1}[\||\varphi\rangle - |\psi\rangle\|_2] \leq \frac{6l}{2^{\lambda n}} \ . \tag{49}$$

Finally, by the triangle inequality, Eqs. (48) and (49), we have

$$\left\|\mathop{\mathbb{E}}_{|\phi\rangle\sim\nu}\left[(|\phi\rangle\!\langle\phi|)^{\otimes l}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[(|\psi\rangle\!\langle\psi|)^{\otimes l}\right]\right\|_1 \leq \frac{1280(\lambda+1)nl}{\lambda^{\log\lambda}\cdot n^{\log n}} + \frac{6l}{2^{\lambda n}} = \mathrm{negl}(\lambda) \ .$$

This establishes the *Statistical Pseudorandomness* property.

□

C.5 Proof of Theorem 11

Proof of Theorem 11. The key length is bounded by $4T \cdot \mathrm{poly}(n, d) = \mathrm{poly}(n, \lambda)$ since τ and F are efficient. Thus the condition of polynomial-bounded key length is satisfied. To implement $\mathsf{SGC}_k^{n,\lambda}$, we need to realize each of the $T = 10(\lambda+1)n$ unitary gates L that compose $\mathsf{SGC}_k^{n,\lambda}$. Since each gate L can be implemented in $\mathrm{poly}(n, \lambda)$ time due to the efficiency of τ and F, the total construction time for $\mathsf{SGC}_k^{n,\lambda}$ is also $\mathrm{poly}(n, \lambda)$. Thus the uniformity is also satisfied.

We now prove the pseudorandomness property. To this end, we consider three hybrids for an arbitrary $|\phi\rangle \in \mathcal{S}(\mathcal{H})$ and $l \in \mathrm{poly}(\lambda, n)$:

H1: $|\phi_k\rangle^{\otimes l}$ for $|\phi_k\rangle = \mathsf{SGC}_k^{n,\lambda}|\phi\rangle$ where $k \leftarrow (\mathcal{K}_1 \times \mathcal{K}_2 \times \mathcal{K}_2 \times \mathcal{K}_2)^T$ is chosen uniformly at random.

H2: $\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}\right\rangle^{\otimes l}$ for

$$\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}\right\rangle = \mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\phi\rangle$$

with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T, (g_i)_{i=1}^T, (h_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$. Here the unitary $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$ is defined in Definition 11.

H3: $|\psi\rangle^{\otimes l}$ for $|\psi\rangle$ chosen according to the Haar measure μ on $\mathcal{S}(\mathcal{H})$.

We first prove that H1 and H2 are computationally indistinguishable. By the quantum-secure property of τ and F, we know the following two situations are computationally indistinguishable for any polynomial-time quantum oracle algorithm $\mathcal{A}$ (see Lemma 6):

- given oracle access to $\tau_{r_1}, \cdots, \tau_{r_T}$ and $F_{u_1}, \cdots, F_{u_T}, F_{s_1}, \cdots, F_{s_T}, F_{t_1}, \cdots, F_{t_T}$ where $(r_i)_{i=1}^T \subseteq \mathcal{K}_1$ and $(u_i)_{i=1}^T, (s_i)_{i=1}^T, (t_i)_{i=1}^T \subseteq \mathcal{K}_2$ are independent and uniformly random keys.
- given oracle access to independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and random functions $(f_i)_{i=1}^T, (g_i)_{i=1}^T, (h_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.

Thus, we have for any polynomial-time quantum algorithm $\mathcal{A}$,

$$\left|\Pr\left[\mathcal{A}\left(|\phi_k\rangle^{\otimes l}\right) = 1\right] - \Pr\left[\mathcal{A}\left(\left|\varphi_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}\right\rangle^{\otimes l}\right) = 1\right]\right| = \mathrm{negl}(\lambda) \quad .$$

For H2 and H3, they are statistically indistinguishable since RSGC^n defined in Definition 11 is an RSS by Theorem 10. Finally, by the triangle inequality we establish H1 and H3 are computationally indistinguishable. This accomplishes the proof. □

C.6 Proof of Lemma 10

Proof of Lemma 10. Let graph $G_0 = (V = [n], E_0 = \emptyset)$. We recursively define $G_1, \ldots, G_l$ as follows: given $G_i = ([n], E_i)$, choose a perfect matching M_i of $[n]$ uniformly at random, and set

$$G_{i+1} = ([n], E_{i+1} = E_i \cup M_i) \quad .$$

Then

$$
\begin{aligned}
&\Pr[\mathcal{P}_{T_0,1} \neq \{[n]\}] \\
&= \Pr[G_l \text{ is disconnected}] \\
&= \Pr[\exists\ S \subseteq [n] \text{ such that there is no edge between } S \text{ and } [n]\backslash S \text{ in } G_l] \\
&\leq \sum_{\substack{i\in[m] \\ i \text{ is even}}} \sum_{S\in\binom{[n]}{i}} \Pr[\text{There is no edge between } S \text{ and } [n]\backslash S \text{ in } G_l] \\
&\leq \sum_{\substack{i\in[m] \\ i \text{ is even}}} \sum_{S\in\binom{[n]}{i}} \left(\frac{\frac{i!}{2^{i/2}(i/2)!} \cdot \frac{(n-i)!}{2^{(n-i)/2}((n-i)/2)!}}{\frac{n!}{2^{n/2}(n/2)!}} \right)^l \\
&= \sum_{\substack{i\in[m] \\ i \text{ is even}}} \sum_{S\in\binom{[n]}{i}} \left(\frac{(n-i)(n-i-1)\cdots((n-i)/2+1)}{n(n-1)\cdots((n+i)/2+1)} \cdot \frac{i(i-1)\cdots(i/2+1)}{((n+i)/2)\cdots(n/2+1)} \right)^l \\
&\leq \sum_{i\in[m],\ i \text{ is even}} \sum_{S\in\binom{[n]}{i}} \left(\frac{i(i-1)\cdots(i/2+1)}{((n+i)/2)\cdots(n/2+1)} \right)^l \\
&\leq \sum_{\substack{i\in[m] \\ i \text{ is even}}} \sum_{S\in\binom{[n]}{i}} \left(\frac{2i}{n+i} \right)^{il/2} \\
&\leq \sum_{\substack{i\in[m] \\ i \text{ is even}}} \binom{n}{i} \left(\frac{2}{3}\right)^{il/2} \leq \left(1+\left(\frac{2}{3}\right)^{l/2}\right)^n - 1 \leq n\left(\frac{2}{3}\right)^{l/2} \left(1+\left(\frac{2}{3}\right)^{l/2}\right)^{n-1} .
\end{aligned}
$$

When $l = 5(1+c)\log n$ and n is sufficiently large,

$$
n\left(\frac{2}{3}\right)^{l/2} \leq n^{-c} \quad \text{and} \quad \left(1+\left(\frac{2}{3}\right)^{l/2}\right)^{n-1} \leq 2 .
$$

□

C.7 Proof of Theorem 7

To prove Theorem 7, we extend the two-stage coupling introduced in real case to complex case. We have introduced the proportional coupling of complex space in Definition 19. We now extend the the non-Markovian coupling and then prove the mixing time. We assume $n = 2m$.

Non-Markovian Coupling

Definition 20 (Non-Markovian Coupling). *Fix $T_0 \leq T \in \mathbb{N}$. We couple $\{X_t\}_{T_0 \leq t \leq T}, \{Y_t\}_{T_0 \leq t \leq T}$ in the following way:*

1. *For each* $T_0 \le t < T$*, choose a perfect matching*

$$P_t = \left\{ \left(i_1^{(t)}, j_1^{(t)}\right), \dots, \left(i_m^{(t)}, j_m^{(t)}\right) \right\}$$

uniformly at random.

2. *Set* $\mathcal{P}_{T,1} = \{\{1\}, \dots, \{n\}\}$*, and define a sequence of partitions*

$$\{\mathcal{P}_{t,k}\}_{T_0 \le t < T,\ 1 \le k \le m+1}$$

of $[n]$ *in the same way as Definition 15.*

3. *If* $\mathcal{P}_{T_0,1} = \{\{1, \dots, n\}\}$*, we couple* $\{X_t\}_{T_0 \le t \le T}$*,* $\{Y_t\}_{T_0 \le t \le T}$ *in the following way:*
   - *Define the set*

$$H = \{(t,k) : T_0 \le t < T,\ 1 \le k \le m,\ \mathcal{P}_{t,k} \ne \mathcal{P}_{t,k+1}\} \quad .$$

   - *Fix* $T_0 \le t < T$*,* X_t *and* Y_t*, and we couple* X_{t+1} *and* Y_{t+1} *in the following way:*
     (a) *Set* $X_{t,1} = X_t$ *and* $Y_{t,1} = Y_t$*.*
     (b) *For* $1 \le k \le m$*,*
        i. *If* $(t,k) \notin H$*, we obtained* $X_{t,k+1}$ *and* $Y_{t,k+1}$ *in the same way as the proportional coupling defined in Definition 19.*
        ii. *If* $(t,k) \in H$*, let*

$$l_k^{(t)} = \sqrt{\left|X_{t,k}[i_k^{(t)}]\right|^2 + \left|X_{t,k}[j_k^{(t)}]\right|^2}$$

and

$$l'^{(t)}_k = \sqrt{\left|Y_{t,k}[i_k^{(t)}]\right|^2 + \left|Y_{t,k}[j_k^{(t)}]\right|^2}.$$

Let U_0 *and* U_0' *be the unitary operators which satisfy*

$$U_0 \begin{pmatrix} X_{t,k}[i_k^{(t)}] \\ X_{t,k}[j_k^{(t)}] \end{pmatrix} = \begin{pmatrix} l_k^{(t)} \\ 0 \end{pmatrix} \quad \text{and} \quad U_0' \begin{pmatrix} Y_{t,k}[i_k^{(t)}] \\ Y_{t,k}[j_k^{(t)}] \end{pmatrix} = \begin{pmatrix} l'^{(t)}_k \\ 0 \end{pmatrix} \quad .$$

Then we choose the best distribution ν *among all joint distributions on* $[0,1) \times [0,1)$ *with both marginal distributions uniformly distributed on* $[0,1)$ *which maximizes the probability of the following events when* $(\zeta, \zeta') \sim \nu$ *and* α, β *are uniformly sample from* $[0, 2\pi)$*:*

$$\sum_{i \in S_r(t,k+1)} |X_{t,k+1}[i]|^2 = \sum_{i \in S_r(t,k+1)} |Y_{t,k+1}[i]|^2, \quad 1 \le r \le l_{t,k+1}$$

where

$$X_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha, \beta, \arcsin\sqrt{\zeta}, U_0 X_{t,k}\right) \quad ,$$
$$Y_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha, \beta, \arcsin\sqrt{\zeta'}, U_0' Y_{t,k}\right) \quad .$$

Then choose $(\zeta_k^{(t)}, {\zeta'}_k^{(t)}) \sim \nu$ *and* $\alpha_k^{(t)}, \beta_k^{(t)}$ *uniformly from* $[0, 2\pi)$*, and set*

$$X_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha_k^{(t)}, \beta_k^{(t)}, \arcsin\sqrt{\zeta_k^{(t)}}, U_0 X_{t,k}\right) \quad ,$$

$$Y_{t,k+1} = G_{\mathbb{C}}\left(i_k^{(t)}, j_k^{(t)}, \alpha_k^{(t)}, \beta_k^{(t)}, \arcsin\sqrt{{\zeta'}_k^{(t)}}, U_0' Y_{t,k}\right) \quad .$$

(c) *Set* $X_{t+1} = X_{t,m+1}$ *and* $Y_{t+1} = Y_{t,m+1}$.

4. *If* $\mathcal{P}_{T_0,1} \neq \{\{1,\dots,n\}\}$*, for* $T_0 \leq t \leq T$*, we couple* X_{t+1} *and* Y_{t+1} *in the following way: choose* $2m$ *independent angles*

$$\alpha_1^{(t)}, \dots, \alpha_m^{(t)}, \beta_1^{(t)}, \dots, \beta_m^{(t)} \in [0, 2\pi)$$

uniformly at random. Additionally, m *independent real numbers* $\zeta_1^{(t)} \dots, \zeta_m^{(t)} \in [0,1)$ *are selected uniformly at random and compute*

$$\theta_k^{(t)} = \arcsin\left(\sqrt{\zeta_k^{(t)}}\right)$$

for all $k \in \{1,\dots,m\}$*. We set*

$$X_{t+1} = F_{\mathbb{C}}\Big(P_t, \left\{\alpha_k^{(t)}\right\}_{k=1}^m, \left\{\beta_k^{(t)}\right\}_{k=1}^m, \left\{\theta_k^{(t)}\right\}_{k=1}^m, X_t\Big) \quad ,$$

$$Y_{t+1} = F_{\mathbb{C}}\Big(P_t, \left\{\alpha_k^{(t)}\right\}_{k=1}^m, \left\{\beta_k^{(t)}\right\}_{k=1}^m, \left\{\theta_k^{(t)}\right\}_{k=1}^m, Y_t\Big) \quad .$$

For $T_0 \leq t \leq T$, $1 \leq k \leq m+1$ and $1 \leq i \leq n$ we define

$$A_{t,k}[i] = |X_{t,k}[i]|^2 \quad , \quad B_{t,k}[i] = |Y_{t,k}[i]|^2 \quad ,$$

and define the event $\mathcal{A}(t,k)$ by

$$\mathcal{A}(t,k) = \{\text{Eq. (50) are satisfied for all } (t',k') \sqsubseteq (t,k) \text{ such that } T_0 \leq t' \leq t\} \,.$$

$$\sum_{i \in S_r(t',k')} A_{t',k'}[i] = \sum_{i \in S_r(t',k')} B_{t',k'}[i] \quad , \quad 1 \leq r \leq l_{t',k'} \quad . \tag{50}$$

Similarly, we have the following two lemmas as Lemma 11 and Lemma 12 before.

Lemma 20. *Fix* $T_0 < T$ *and two chains* $\{X_t\}_{T_0 \leq t \leq T}, \{Y_t\}_{T_0 \leq t \leq T}$ *are coupled using the non-Markovian coupling defined in Definition 20. Fix* $T_0 \leq t \leq T$ *and* $1 \leq k \leq m+1$*. Then, on the event* $\mathcal{A}(t,k) \cap \{\mathcal{P}_{T_0,1} = \{1,\dots,n\}\}$*, we have*

$$\|A_{t',k'} - B_{t',k'}\|_{1,S} \leq \|A_{T_0} - B_{T_0}\|_1$$

for all $(t',k') \sqsubseteq (t,k)$ *such that* $T_0 \leq t' \leq t$ *and* $S \in \mathcal{P}_{t',k'}$*. Moreover, for all* $(t',k') \sqsubseteq (t,k)$ *such that* $T_0 \leq t' \leq t$*,*

$$\|A_{t',k'} - B_{t',k'}\|_1 \leq n \, \|A_{T_0} - B_{T_0}\|_1 \quad .$$

The proof of above lemma is the same as Lemma 4.4 in [44].

Lemma 21. *Fix positive reals* $1 < p < q$. *Let* $\zeta, \zeta' \sim \mathrm{Unif}[0,1)$ *and let*

$$S = A + B\zeta \quad \text{and} \quad S' = C + D\zeta'$$

for some $0 \le A, B, C, D \le 1$ *that satisfy*

$$|A - C|, |B - D| \le n^{-q} \quad \text{and} \quad B, D \ge n^{-p} \ .$$

Then for sufficiently large n, *there exists a coupling of* ζ, ζ' *so that*

$$\Pr[S = S'] \ge 1 - 3n^{-(q-p)} \ .$$

Proof. Without loss of generality, we assume $B \ge D$. The total variation distance between S and S' is

$$\begin{aligned} \|S - S'\|_{\mathrm{TV}} &\le 1 - \int_{(A,A+B)\cap(C,C+D)} \frac{1}{B}\,\mathrm{d}x \\ &\le 1 - \int_{A+n^{-q}}^{A+B-2n^{-q}} \frac{1}{B}\,\mathrm{d}x \\ &= 1 - \left(B - 3n^{-q}\right)\frac{1}{B} \\ &= \frac{3n^{-q}}{B} \le 3n^{-(q-p)} \ . \end{aligned}$$

This implicitly defines a coupling of ζ, ζ' that satisfies $\Pr[S = S'] \ge 1 - 3n^{-(q-p)}$. □

Proof of Theorem 7 Let $a = 30,\ b = 24,\ T_0 = 500\log n,\ T_1 = 15\log n,\ T = T_0 + T_1 = 515\log n$. We construct a coupling of two copies $\{X_t\}_{t\ge 0}, \{Y_t\}_{t\ge 0}$ with starting point $X_0 \in \mathcal{S}_{\mathbb{C}}^n$ and $Y_0 \sim \mu_{\mathbb{C}}$. The coupling is as follows:

1. Couple $\{X_t\}_{0\le t\le T_0}, \{Y_t\}_{0\le t\le T_0}$ by using the proportional coupling defined in Definition 19.
2. Couple $\{X_t\}_{T_0\le t\le T}, \{Y_t\}_{T_0\le t\le T}$ by using the non-markovian coupling defined in Definition 20.

Define the event

$$\begin{aligned} \mathcal{E}_1 &= \left\{\|A_{T_0} - B_{T_0}\|_1 \ge n^{-a}\right\} \ , \\ \mathcal{E}_2 &= \{\mathcal{P}_{T_0,1} \ne \{\{1,\ldots,n\}\}\} \ , \\ \mathcal{E}_3 &= \{X_T \ne Y_T\} \ . \end{aligned}$$

By Lemma 2,

$$\begin{aligned} \sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \|\mathcal{L}(X_T) - \mu_{\mathbb{C}}\|_{\mathrm{TV}} &\le \sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \Pr[\mathcal{E}_3] \\ &\le \sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \left(\Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2] + \Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c]\right) \ . \end{aligned} \tag{51}$$

By Lemma 18 and Markov's inequality, we have

$$\begin{aligned}\Pr[\mathcal{E}_1] &= \Pr\left[\|A_{T_0} - B_{T_0}\|_1 \geq n^{-a}\right] \\ &\leq \Pr\left[\|A_{T_0} - B_{T_0}\|_2 \geq n^{-a-1/2}\right] \\ &\leq n^{2a+1} \cdot 2 \cdot \left(\frac{2}{3}\right)^{T_0} \leq \frac{1}{n^2} \ . \end{aligned} \tag{52}$$

Moreover, by Lemma 10, we have

$$\Pr[\mathcal{E}_2] \leq 2n^{-2} \ . \tag{53}$$

In order to bound $\Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c]$, recall the definition of $\mathcal{A}(t,k)$ in (50). If $\mathcal{A}(T,1)$ occurs, we have $|X_t[i]| = |Y_t[i]|$ for all $i \in \{1,\ldots,n\}$. Meanwhile, our coupling ensures $X_t[i]$ and $Y_t[i]$ have the same argument. This means $\mathcal{A}(T,1)$ implies $\mathcal{E}_3^c$. As a result, $\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c$ implies $\mathcal{A}(T,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c$. So, we have

$$\begin{aligned} &\Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\ \leq\ &\Pr[\mathcal{A}(T,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\ \leq\ &\sum_{t=T_0}^{T-1} \Pr[\mathcal{A}(t+1,1)^c \cap \mathcal{A}(t,1) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] + \Pr[\mathcal{A}(T_0,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\ =\ &\sum_{t=T_0}^{T-1} \Pr[\mathcal{A}(t,m+1)^c \cap \mathcal{A}(t,1) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] + \Pr[\mathcal{A}(T_0,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\ \leq\ &\sum_{t=T_0}^{T-1} \sum_{k=1}^{m} \Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] + \Pr[\mathcal{A}(T_0,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \ . \end{aligned} \tag{54}$$

Notice that if $\mathcal{E}_2^c$ happens, we have

$$\sum_{i \in \{1,\ldots,n\}} |X_{T_0,1}[i]|^2 = \sum_{i \in \{1,\ldots,n\}} |Y_{T_0,1}[i]|^2 = 1 \ .$$

Therefore,

$$\Pr[\mathcal{A}(T_0,1)^c \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] = 0 \ . \tag{55}$$

Combining (54) and (55), we have

$$\Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \leq \sum_{t=T_0}^{T-1} \sum_{k=1}^{m} \Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \ . \tag{56}$$

We are now left to find a upper bound for $\Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c]$ when $T_0 \leq t \leq T-1$ and $1 \leq k \leq m$. To this end, we define

$$\mathcal{B}(t,k) = \left\{ \min_{(t',k') \sqsubseteq (t,k): t' \leq t} \ \min_{1 \leq i \leq n} |Y_{t',k'}[i]|^2 \geq (2n)^{-b} \right\} \ .$$

Note that

$$\Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\ \leq \Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] + \Pr[\mathcal{B}(t,k)^c] \ . \tag{57}$$

By Lemma 5 and a union bound over all (t',k') such that $(t',k') \sqsubseteq (t,k)$ and $t' \leq t$, we have for sufficiently large n,

$$\Pr[\mathcal{B}(t,k)^c] \leq 15 \cdot 2^{1-\frac{b}{3}} \cdot n^{3-\frac{b}{3}} \log(n) \leq \frac{1}{n^4} \ . \tag{58}$$

Next, we consider two cases of the term

$$\Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c]$$

in (57): $(t,k) \notin H$ and $(t,k) \in H$, where H is defined in Definition 20. In the case that $(t,k) \notin H$, we have $\mathcal{P}_{t,k} = \mathcal{P}_{t,k+1}$ and we apply the proportional coupling. Thus $\mathcal{A}(t,k)$ implies $\mathcal{A}(t,k+1)$ which means

$$\Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] = 0 \ . \tag{59}$$

In the other case that $(t,k) \in H$, let

$$A = \sum_{i \in S_{v_{t,k}}(t,k+1) \setminus j_k^{(t)}} |X_{t,k}[i]|^2 \ , \qquad B = \left|X_{t,k}[i_k^{(t)}]\right|^2 + \left|X_{t,k}[j_k^{(t)}]\right|^2 \ ,$$

$$C = \sum_{i \in S_{v_{t,k}}(t,k+1) \setminus j_k^{(t)}} |Y_{t,k}[i]|^2 \ , \qquad D = \left|Y_{t,k}[i_k^{(t)}]\right|^2 + \left|Y_{t,k}[j_k^{(t)}]\right|^2 \ ,$$

$$S = A + B\zeta_k^{(t)} \ , \qquad S' = C + D{\zeta'}_k^{(t)} \ .$$

On the event $\mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c$, we have by Lemma 11

$$|A - C| \leq \|A_{t,k} - B_{t,k}\|_1 \leq n \|A_{T_0} - B_{T_0}\|_1 \leq n^{1-a} \ .$$

Similarly,

$$|B - D| \leq \|A_{t,k} - B_{t,k}\|_1 \leq n \|A_{T_0} - B_{T_0}\|_1 \leq n^{1-a} \ .$$

Moreover, $D \geq n^{-(b+1)}$ and $B \geq D - |B - D| \geq n^{-(b+1)}$ for sufficiently large n. Then apply Lemma 21 with $p = b+1$, $q = a-1$, we know there exists a distribution ν_0 such that when $\left(\zeta_k^{(t)}, {\zeta'}_k^{(t)}\right) \sim \nu_0$, we have

$$\Pr_{\left(\zeta_k^{(t)}, {\zeta'}_k^{(t)}\right) \sim \nu_0} [S \neq S' | \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \leq 3n^{-(a-b-2)} \ .$$

We choose the best distribution which maximizes the probability of event described in (50) , so

$$\begin{aligned}
&\Pr[\mathcal{A}(t,k+1)^c \cap \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \\
\leq &\Pr\left[\mathcal{A}(t,k+1)^c \middle| \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \\
\leq &\Pr_{\left(\varsigma_k^{(t)},\varsigma'^{(t)}_k\right)\sim\nu_0}\left[\mathcal{A}(t,k+1)^c \middle| \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \\
= &\Pr_{\left(\varsigma_k^{(t)},\varsigma'^{(t)}_k\right)\sim\nu_0}\left[S \neq S' \middle| \mathcal{A}(t,k) \cap \mathcal{B}(t,k) \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c\right] \leq 3n^{-(a-b-2)} \ . \qquad (60)
\end{aligned}$$

Combining (56), (57), (59), (60) and (58), we have for sufficiently large n

$$\Pr[\mathcal{E}_3 \cap \mathcal{E}_1^c \cap \mathcal{E}_2^c] \leq \sum_{t=T_0}^{T-1}\sum_{k=1}^{m} 3n^{-(a-b-2)} + n^{-4} \leq \frac{1}{n^2} \ . \qquad (61)$$

By (51), (52), (53) and (61), we have for sufficiently large n and $T = 515\log n$

$$\sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \|\mathcal{L}(X_T) - \mu_{\mathbb{C}}\|_{\mathrm{TV}} \leq \frac{1}{2n} \ .$$

As for $T \geq 515\log n$, by Lemma 3 we have

$$\sup_{X_0\in\mathcal{S}_{\mathbb{C}}^n} \|\mathcal{L}(X_T) - \mu_{\mathbb{C}}\|_{\mathrm{TV}} \leq 2\left(\frac{1}{n}\right)^{\left\lfloor\frac{T}{515\log n}\right\rfloor} \leq \frac{1}{2^{(c/515-1)\log n - 1}} \ .$$

C.8 Proof of Lemma 13

Proof of Lemma 13. To see why $\widetilde{\mathcal{N}} = \mathcal{S}_{\mathbb{R}}^{2^n}$, we prove $\mathcal{S}_{\mathbb{R}}^{2^n} \subseteq \widetilde{\mathcal{N}}$ since $\widetilde{\mathcal{N}} \subseteq \mathcal{S}_{\mathbb{R}}^{2^n}$ is trivial. Given $|\xi\rangle \in \mathcal{S}_{\mathbb{R}}^{2^n}$, we prove $|\xi\rangle \in \widetilde{\mathcal{N}}$ by constructing a series of $(\sigma_t, \widetilde{f}_t)$ for $t \leq n$ such that

$$\widetilde{K}_{\sigma_n,\widetilde{f}_n}\cdots\widetilde{K}_{\sigma_1,\widetilde{f}_1}|\eta\rangle = |\xi\rangle$$

(we let $\widetilde{K}_{\sigma_t,\widetilde{f}_t} = \mathbb{1}$ for $t > n$).

The idea is to bisect $\widetilde{K}_{\sigma_t,\widetilde{f}_t}\cdots\widetilde{K}_{\sigma_1,\widetilde{f}_1}|\eta\rangle$ and $|\xi\rangle$ accordingly and recursively, always keeping the 2-norm of each part of the two vectors equal.

To illustrate the process in detail, we begin with indexing the entries of a vector in $\mathcal{S}_{\mathbb{R}}^{2^n}$ by a bit string of length n. In step $t \in [n]$, we divide the index set $\{0,1\}^n$ into 2^t sets based on the first t bits. For $y \in \{0,1\}^t$, we define S_y to be the set of all elements in $\{0,1\}^n$ with prefix y. We then split $|\eta\rangle$ and $|\xi\rangle$ into 2^t sub-vectors, according to S_y where $y \in \{0,1\}^t$. Let L_y be the function from $\mathcal{S}_{\mathbb{R}}^{2^n}$ to $\mathbb{R}$ that gives the length of the sub-vector corresponding to S_y. Let

$$|\eta_0\rangle = |\eta\rangle\,, \quad |\eta_t\rangle = \widetilde{K}_{\sigma_t,\widetilde{f}_t}|\eta_{t-1}\rangle \quad \text{for } t \in [n] \ .$$

Our goal is to construct $(\sigma_t, \widetilde{f}_t)$ at each step $t \in [n]$, such that $L_y(|\eta_t\rangle) = L_y(|\xi\rangle)$ for every $t \in [n-1], y \in \{0,1\}^t$ and that $|\eta_n\rangle = |\xi\rangle$.

To accomplish this, we first define $\sigma_t \in S_{2^n}$ to be

$$\sigma_t(x) = x_t x_1 \dots x_{t-1} x_{t+1} \dots x_n \quad \text{for all } x \in \{0,1\}^n \ .$$

For any $y \in \{0,1\}^{t-1}$, σ_t matches every index in S_{y0} with another index in S_{y1} that shares a common suffix of length $n-t$.

Next, we move to construct $\widetilde{f}_t$. Define α_y for each $y \in \{0,1\}^{t-1}$ as

$$\alpha_y = \begin{cases} \arccos \frac{L_{y0}(|\xi\rangle)}{L_y(|\xi\rangle)} & \text{if } L_y(|\xi\rangle) \neq 0 \text{ and } t < n \ , \\ 0 & \text{if } L_y(|\xi\rangle) = 0 \text{ and } t < n \ . \end{cases}$$

When $t = n$, we define α_y for $y \in \{0,1\}^{n-1}$ to be any angle satisfying

$$(|\xi\rangle)_{y0} = L_y(|\xi\rangle) \cos \alpha_y \ ,$$
$$(|\xi\rangle)_{y1} = L_y(|\xi\rangle) \sin \alpha_y \ .$$

We want to design $\widetilde{f}_t$ which controls the rotation of each index pair to let each pair of indices between S_{y0} and S_{y1} (induced by σ_t) form an angle α_y with the x-axis in a two-dimensional Cartesian coordinate system. To this end, for each $y \in \{0,1\}^{t-1}$ and $z \in \{0,1\}^{n-t}$, we define $\beta_y(z)$ to be any angle satisfying

$$(|\eta_{t-1}\rangle)_{y0z} = \sqrt{(|\eta_{t-1}\rangle)_{y0z}^2 + (|\eta_{t-1}\rangle)_{y1z}^2} \cos \beta_y(z),$$
$$(|\eta_{t-1}\rangle)_{y1z} = \sqrt{(|\eta_{t-1}\rangle)_{y0z}^2 + (|\eta_{t-1}\rangle)_{y1z}^2} \sin \beta_y(z),$$

and we define $\widetilde{f}_t$ to be

$$\widetilde{f}_t(yz) = (\alpha_y - \beta_y(z)) / (2\pi)$$

for all $y \in \{0,1\}^{t-1}$ and $z \in \{0,1\}^{n-t}$. It can be easily verified that

$$L_y(|\eta_t\rangle) = L_y(|\xi\rangle) \quad \text{for} \quad t \in [n-1] \ \text{ and } \ y \in \{0,1\}^t,$$

and $|\eta_n\rangle = |\xi\rangle$. □

C.9 Proof of Lemma 14

Proof of Lemma 14. Let $|u\rangle = \widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T}|\eta\rangle \in \widetilde{\mathcal{N}}$ for some $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and $(\widetilde{f}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$. For every $t \in [T]$, we define f_t by letting $f_t(y)$ be the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$. It is evident that $|v\rangle = \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\eta\rangle \in \mathcal{N}$. And

$$\begin{aligned} \||u\rangle - |v\rangle\|_2 &= \left\|\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T}|\eta\rangle - \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}|\eta\rangle\right\|_2 \\ &\le \left\|\widetilde{\mathsf{RSG}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T} - \mathsf{RSG}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T}\right\|_\infty \le 2^{1-d}\pi T = \frac{1030\pi(\lambda+1)n}{\lambda^{\log\lambda} n^{\log n}} \ , \end{aligned} \tag{62}$$

where the last inequality is from Fact 1 and Lemma 8. This proves that $\mathcal{N}$ is indeed an ϵ-net for $\widetilde{\mathcal{N}}$ where $\epsilon = \frac{1030\pi(\lambda+1)n}{\lambda^{\log\lambda} n^{\log n}} = \mathrm{negl}(\lambda)$. Combining with Lemma 13, we conclude the result. □

C.10 Proof of Theorem 15

To prove Theorem 15, recall the ensemble of (infinitely many) unitary operators $\widetilde{\mathsf{RSGC}}^n := \left\{\widetilde{\mathsf{RSGC}}^{n,\lambda}\right\}_\lambda$ defined in Section C.4.

Proposition 3. *For $T = 515(\lambda+1)n$, the ensemble of unitary operator $\widetilde{\mathsf{RSGC}}^n$ is a CRSS.*

Proof. Note that a uniformly random $\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}$ corresponds to a T-step parallel Kac's walk on $\mathcal{S}_{\mathbb{C}}^{2^n}$. The proposition then follows from Theorem 7 and the definition of CRSS. □

Let $\eta \in \mathcal{S}(H)$ be an arbitrary state. Denote

$$\mathcal{N} = \left\{\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle\right\} \text{ and } \widetilde{\mathcal{N}} = \left\{\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle\right\} .$$

We prove the following two lemmas.

Lemma 22. $\widetilde{\mathcal{N}} = \mathcal{S}(\mathcal{H})$.

Proof. We prove $\mathcal{S}_{\mathbb{C}}^{2^n} \subseteq \widetilde{\mathcal{N}}$ since $\widetilde{\mathcal{N}} \subseteq \mathcal{S}_{\mathbb{C}}^{2^n}$ is trivial. Given $|\xi\rangle \in \mathcal{S}_{\mathbb{C}}^{2^n}$, we prove $|\xi\rangle \in \widetilde{\mathcal{N}}$ by constructing a series of $(\sigma_t, \widetilde{f}_t, \widetilde{g}_t, \widetilde{h}_t)$ for $t \le n+1$ such that

$$\widetilde{L}_{\sigma_n,\widetilde{f}_{n+1},\widetilde{g}_{n+1},\widetilde{h}_{n+1}} \cdots \widetilde{L}_{\sigma_1,\widetilde{f}_1,\widetilde{g}_1,\widetilde{h}_1} |\eta\rangle = |\xi\rangle$$

(we let $\widetilde{L}_{\sigma_t,\widetilde{f}_t,\widetilde{g}_t,\widetilde{h}_t} = \mathbb{1}$ for $t > n+1$).

The proof idea is similar to Appendix C.8. For any $t \in [n]$ and $y \in \{0,1\}^t$, we use the definition of S_y, L_y (change the domain to $\mathcal{S}_{\mathbb{C}}^{2^n}$), and $|\eta_t\rangle$ (change $\widetilde{K}_{\sigma_t,\widetilde{f}_t}$ to $\widetilde{L}_{\sigma_t,\widetilde{f}_t,\widetilde{g}_t,\widetilde{h}_t}$) there. Our goal is to construct $(\sigma_t, \widetilde{f}_t, \widetilde{g}_t, \widetilde{h}_t)$ at each step $t \in [n+1]$, such that $L_y(|\eta_t\rangle) = L_y(|\xi\rangle)$ for every $t \in [n-1], y \in \{0,1\}^t$ and that $|\eta_{n+1}\rangle = |\xi\rangle$. That is, after $n-1$ steps, for every $y \in \{0,1\}^{n-1}$, the two-dimensional sub-vectors of $|\eta_n\rangle$ and $|\xi\rangle$ induced by S_y have the same length. In the final two steps, we adjust the two sub-vectors to be equal.

For any $t \in [n]$, let σ_t be defined as in Appendix C.8. We now construct $\widetilde{f}_t, \widetilde{g}_t, \widetilde{h}_t$.

For any $t \in [n], y \in \{0,1\}^{t-1}, z \in \{0,1\}^{n-t}$, suppose that

$$(|\eta_{t-1}\rangle)_{y0z} = e^{\mathrm{i}\theta^\eta_{y,0,z}} r^\eta_{y,0,z} \quad \text{and} \quad (|\eta_{t-1}\rangle)_{y1z} = e^{\mathrm{i}\theta^\eta_{y,1,z}} r^\eta_{y,1,z} \ ,$$

$$(|\xi\rangle)_{y0z} = e^{\mathrm{i}\theta^\xi_{y,0,z}} r^\xi_{y,0,z} \quad \text{and} \quad (|\xi\rangle)_{y1z} = e^{\mathrm{i}\theta^\xi_{y,1,z}} r^\xi_{y,1,z} \ ,$$

where $\theta^\eta_{y,0,z}, \theta^\eta_{y,1,z}, \theta^\xi_{y,0,z}, \theta^\xi_{y,1,z} \in [0, 2\pi), r^\eta_{y,0,z}, r^\eta_{y,1,z}, r^\xi_{y,0,z}, r^\xi_{y,1,z} \in [0,1]$.

For $t \in [n-1]$, define

$$\alpha_y = \begin{cases} \arccos \frac{L_{y0}(|\xi\rangle)}{L_y(|\xi\rangle)} & \text{if } L_y(|\xi\rangle) \neq 0 \ , \\ 0 & \text{otherwise.} \end{cases}$$

and

$$\rho_y(z) = \begin{cases} \arccos \frac{r^\eta_{y,0,z}}{\sqrt{\left(r^\eta_{y,0,z}\right)^2 + \left(r^\eta_{y,1,z}\right)^2}} & \text{if } \left(r^\eta_{y,0,z}\right)^2 + \left(r^\eta_{y,1,z}\right)^2 > 0 \ , \\ 0 & \text{otherwise.} \end{cases}$$

We then define $\widetilde{f}_t, \widetilde{g}_t, \widetilde{h}_t$ to be

$$\widetilde{f}_t(yz) = \sin^2\left(\alpha_y - \rho_y(z)\right),$$
$$\widetilde{g}_t(yz) = \theta^\eta_{y,1,z}/(2\pi),$$
$$\widetilde{h}_t(yz) = \theta^\eta_{y,0,z}/(2\pi),$$

for all $y \in \{0,1\}^{t-1}$ and $z \in \{0,1\}^{n-t}$. It can be easily verified that

$$L_y\left(|\eta_t\rangle\right) = L_y\left(|\xi\rangle\right) \quad \text{for} \quad t \in [n-1] \text{ and } y \in \{0,1\}^t.$$

For $t = n$, we set the second entry of the sub-vector of $|\eta_{n-1}\rangle$ induced by S_y to zero for all $y \in \{0,1\}^{n-1}$. That is, define

$$\widetilde{f}_n(y) = \begin{cases} \frac{\left(r^\eta_{y,1}\right)^2}{\left(r^\eta_{y,0}\right)^2 + \left(r^\eta_{y,1}\right)^2} & \text{if } \left(r^\eta_{y,0}\right)^2 + \left(r^\eta_{y,1}\right)^2 > 0, \\ 0 & \text{otherwise,} \end{cases}$$
$$\widetilde{g}_n(y) = -\theta^\eta_{y,0}/(2\pi),$$
$$\widetilde{h}_n(y) = \left(\pi - \theta^\eta_{y,1}\right)/(2\pi),$$

for all $y \in \{0,1\}^{n-1}$. We then have

$$L_y\left(|\eta_n\rangle\right) = L_y\left(|\eta_{n-1}\rangle\right) = L_y\left(|\xi\rangle\right) \text{ for all } y \in \{0,1\}^{n-1}.$$

For the final step, we let $\sigma_{n+1} = \sigma_n$, and define

$$\widetilde{f}_{n+1}(y) = \begin{cases} \frac{\left(r^\xi_{y,1}\right)^2}{\left(r^\xi_{y,0}\right)^2 + \left(r^\xi_{y,1}\right)^2} & \text{if } \left(r^\xi_{y,0}\right)^2 + \left(r^\xi_{y,1}\right)^2 > 0, \\ 0 & \text{otherwise,} \end{cases}$$
$$\widetilde{g}_{n+1}(y) = \theta^\xi_{y,0}/(2\pi),$$
$$\widetilde{h}_{n+1}(y) = -\theta^\xi_{y,1}/(2\pi),$$

for all $y \in \{0,1\}^{n-1}$. It can be easily verified that $|\eta_{n+1}\rangle = |\xi\rangle$. □

Lemma 23. *There exists an* $\epsilon = \mathrm{negl}(\lambda)$ *such that* $\mathcal{N}$ *is an* ϵ*-net for* $\mathcal{S}(\mathcal{H})$.

Proof. By Lemma 22, it suffices to prove that there exists an $\epsilon = \mathrm{negl}(\lambda)$ such that $\mathcal{N}$ is an ϵ-net for $\widetilde{\mathcal{N}}$.

Let $|u\rangle = \widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle \in \widetilde{\mathcal{N}}$ for some $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$ and $(\widetilde{f}_i)_{i=1}^T$, $(\widetilde{g}_i)_{i=1}^T$, $(\widetilde{h}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$. For every $t \in [T]$, we define f_t by letting $f_t(y)$ be the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$. We define g_t and h_t for $t \in [T]$ in the same way. It is evident that $|v\rangle = \mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle \in \mathcal{N}$. And

$$\begin{aligned}
& \||u\rangle - |v\rangle\|_2 \\
= & \left\| \mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle - \widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle \right\|_2 \\
\le & \left\| \mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T} - \widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T} \right\|_\infty \\
\le & \; 2^{6-\frac{d}{2}} T = \frac{32960(\lambda+1)n}{\lambda^{\log\lambda} n^{\log n}} \quad ,
\end{aligned} \tag{63}$$

where the last inequality is from Fact 1 and Lemma 19. This proves that $\mathcal{N}$ is indeed an ϵ-net for $\widetilde{\mathcal{N}}$ where $\epsilon = \frac{32960(\lambda+1)n}{\lambda^{\log\lambda} n^{\log n}} = \mathrm{negl}(\lambda)$. □

Proof of Theorem 15. It is easy to see that the uniformity condition is satisfied. Let κ denote the key length. Quantum circuit $\mathsf{RSGC}^{n,\lambda}$ applies $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$ after reading $(\sigma_i)_{i=1}^T, (f_i)_{i=1}^T, (g_i)_{i=1}^T$ and $(h_i)_{i=1}^T$. To implement $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$, we need to realize each of the $T = 515(\lambda+1)n$ unitary gates L. Since each gate L can be implemented in $\mathrm{poly}(n,\lambda,\kappa)$ time, the total construction time for $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}$ is also $\mathrm{poly}(n,\lambda,\kappa)$.

In conjunction with Lemma 23, it suffices to prove the existence of a good distribution $\widetilde{\nu}$ meeting the requirement in Definition 13. Fix $|\eta\rangle \in \mathcal{S}(\mathcal{H})$. Define three distributions:

- ν be the distribution of $\mathsf{RSGC}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(f_i)_{i=1}^T,(g_i)_{i=1}^T,(h_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(f_i)_{i=1}^T, (g_i)_{i=1}^T, (h_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to \{0,1\}^d\}$.
- $\widetilde{\nu}$ be the distribution of $\widetilde{\mathsf{RSGC}}^{n,\lambda}_{(\sigma_i)_{i=1}^T,(\widetilde{f}_i)_{i=1}^T,(\widetilde{g}_i)_{i=1}^T,(\widetilde{h}_i)_{i=1}^T}|\eta\rangle$ with independent and uniformly random permutations $(\sigma_i)_{i=1}^T \subseteq S_{2^n}$, and random functions $(\widetilde{f}_i)_{i=1}^T, (\widetilde{g}_i)_{i=1}^T, (\widetilde{h}_i)_{i=1}^T \subseteq \{f : \{0,1\}^{n-1} \to [0,1)\}$.
- μ be the Haar measure on $\mathcal{S}(\mathcal{H})$.

Note that $\widetilde{\nu}$ is the output distribution of T-step parallel Kac's walk in $\mathcal{S}_{\mathbb{C}}^{2^n}$. Thus by Theorem 7, we have

$$\|\widetilde{\nu} - \mu\|_{\mathrm{TV}} \le \frac{1}{2^{\lambda n - 1}} = \mathrm{negl}(\lambda) \quad . \tag{64}$$

We are left to show the Wasserstein ∞-distance between ν and $\widetilde{\nu}$ is negligible. To this end, we construct a coupling γ_0 of ν and $\widetilde{\nu}$ by using the same permutation σ_t and letting f_t be the function satisfying $f_t(y)$ is the d digits after the binary point in $\widetilde{f}_t(y)$ for all $y \in \{0,1\}^{n-1}$ (The same applies to g_t and h_t). Therefore

$$\begin{aligned} W_\infty(\nu, \widetilde{\nu}) = & \lim_{p \to \infty} \left(\inf_{\gamma \in \Gamma(\nu, \widetilde{\nu})} \mathop{\mathbb{E}}_{(|v\rangle, |u\rangle) \sim \gamma} \left[\| |v\rangle - |u\rangle \|_2^p \right] \right)^{1/p} \\ \leq & \lim_{p \to \infty} \left(\mathop{\mathbb{E}}_{(|v\rangle, |u\rangle) \sim \gamma_0} \left[\| |v\rangle - |u\rangle \|_2^p \right] \right)^{1/p} \overset{(\text{Eq. } (63))}{\leq} \frac{32960(\lambda+1)n}{\lambda^{\log \lambda} n^{\log n}} = \text{negl}(\lambda) \quad . \end{aligned} \tag{65}$$

Thus we conclude the result. □