



**Title:**        **Mixing Properties of Stochastic Quantum Hamiltonians**

**Author(s):**    Onorati, E. ; Buerschaper, O.; Kliesch, M.; Brown, W.; Werner, A. H.; Eisert, J.

**Document type:**    Postprint

**Terms of Use:**     Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

**Citation:**         This is a post-peer-review, pre-copyedit version of an article published in Communications in mathematical physics. The final authenticated version is available online at:  
<http://dx.doi.org/10.1007/s00220-017-2950-6>

Onorati, E., Buerschaper, O., Kliesch, M., Brown, W., Werner, A. H., & Eisert, J. (2017). Mixing Properties of Stochastic Quantum Hamiltonians. *Communications in Mathematical Physics*, 355(3), 905–947. <https://doi.org/10.1007/s00220-017-2950-6>

# Mixing properties of stochastic quantum Hamiltonians

E. Onorati<sup>1</sup>, O. Buerschaper<sup>1</sup>, M. Kliesch<sup>1</sup>, W. Brown<sup>1</sup>, A. H. Werner<sup>1,2</sup>, J. Eisert<sup>1</sup>

<sup>1</sup> Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany

<sup>2</sup> Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

**Abstract** Random quantum processes play a central role both in the study of fundamental mixing processes in quantum mechanics related to equilibration, thermalisation and fast scrambling by black holes, as well as in quantum process design and quantum information theory. In this work, we present a framework describing the mixing properties of continuous-time unitary evolutions originating from local Hamiltonians having time-fluctuating terms, reflecting a Brownian motion on the unitary group. The induced stochastic time evolution is shown to converge to a unitary design. As a first main result, we present bounds to the mixing time. By developing tools in representation theory, we analytically derive an expression for a local  $k$ -th moment operator that is entirely independent of  $k$ , giving rise to approximate unitary  $k$ -designs and quantum tensor product expanders. As a second main result, we introduce tools for proving bounds on the rate of decoupling from an environment with random quantum processes. By tying the mathematical description closely with the more established one of random quantum circuits, we present a unified picture for analysing local random quantum and classes of Markovian dissipative processes, for which we also discuss applications.

## Contents

1. Introduction and motivation . . . . .	2
2. Preliminaries . . . . .	4
2.1 Norms . . . . .	4
2.2 Exact and approximate unitary designs . . . . .	4
2.3 Locally generated Brownian motion on the unitary group . . . . .	6
2.4 Local Hamiltonian increments . . . . .	7
3. Main results . . . . .	8
3.1 Tensor product expanders and approximate unitary designs . . . . .	8
3.2 Fast decoupling with stochastic quantum Hamiltonians . . . . .	9
4. Applications . . . . .	12
4.1 Dissipative dynamics arising from fluctuations . . . . .	12
4.2 Applications for fast scrambling . . . . .	12
4.3 Applications in quantum information processing . . . . .	14
5. Stochastic time evolution generates tensor product expanders . . . . .	15
5.1 Local gap . . . . .	19
5.2 Hamiltonian driving . . . . .	22
5.3 More general interaction graphs . . . . .	22
5.4 Example: White noise in the Pauli basis . . . . .	23
6. Decoupling with stochastic Hamiltonian time evolution . . . . .	24
6.1 Two-qubit analysis of the second moment operator . . . . .	24
6.2 Markov chain analysis on weights . . . . .	27
6.3 From the zero chain to the full distribution . . . . .	29
7. Conclusions and outlook . . . . .	31
8. Acknowledgements . . . . .	32
Appendix A. (Young diagrams) . . . . .	34
Appendix B. (Stochastic processes and Markov chains) . . . . .	35
Appendix C. (Proof of Lemma 29) . . . . .	36

## 1. Introduction and motivation

In recent years, several ramifications of quantum processes having a random component have become prominent in the literature [1–9]. These are quantum mechanical processes, but ones which have a classical random component. *Random circuits* are discrete processes of this type, so quantum circuits composed of unitary quantum gates, each quantum gate being randomly drawn according to some probability measure [3, 8, 10–13]. Continuous-time processes belong to this class too, e.g. *stochastically fluctuating local Hamiltonians* [9, 14]. For both discrete and continuous-time evolutions, similar questions arise. This is, for instance, the question how long it takes or what depth of the quantum circuit is required until suitable mixing – in a sense made precise below – is achieved, meaning that they approximate a so-called *unitary design* [15].

This recent development parallels and further develops an established body of literature on fully classical random processes: *Markov chain mixing* provides tools to capture details of the convergence of a Markov chain to its stationary distribution, giving tight bounds on relevant time scales for mixing, hitting or cover times [16]. Applications of this powerful mathematical framework range from algorithms design in computer science to the understanding of processes of equilibration and thermalisation in classical statistical mechanics. The famous cut-off phenomenon of card-shuffling is epitomic for the many intriguing insights the theory has to offer, showing that decks of 52 cards have to be shuffled seven times until the distribution is suddenly close in variation distance to the uniform mixture [17].

Such *random quantum processes*, as they will be called in this work, again have applications in algorithms design, now *quantum algorithms design* [3, 8, 18]. They are used in *quantum process tomography* and *low rank matrix recovery* [19, 20] and *benchmarking* [21], where they provide powerful tools to avoid significant overheads otherwise necessary with naive deterministic prescriptions. They play a key role in notions of *decoupling* [4], the task of approximately bringing a quantum mechanical system into a tensor product state with its environment, which constitutes a key property of quantum mixing processes. It also is an important primitive in quantum information theory [4, 22]: Indeed, it plays a central role in arguments of state merging [23], the task of conveying a subsystem from a sender to a receiver. It also is key to the proof of the quantum reverse Shannon Theorem [24] and is useful to capture quantum channel capacities [25]. *Error correcting codes*, so codes that protect quantum information against unwanted local decoherence, can be built upon such random processes [6]. It should be clear from this that the analysis of such processes constitutes a powerful proof tool in the context of quantum information theory.

Maybe most intriguingly, they are used as proxies for natural mixing processes occurring in physical systems governed by quantum mechanical laws. Clearly, random processes are reminiscent in many ways and sometimes exactly model *thermalising dynamics* of interacting quantum systems with many constituents [26]. This link has particularly prominently been explored in the context of *black hole thermalisation*. This phenomenon is connected to the still unresolved puzzle how quickly black holes release information about their microscopic state. Based on considerations from string theory and gauge-gravity correspondences [27, 28], it is increasingly becoming clear that black holes do not destroy information when evaporating. This insight raises the question on what time scales this release of information precisely happens. It has been suggested that the time scale is set by the time it takes to “*scramble*” the microscopic degrees of freedom of the black hole, in a way that initial local perturbations will be locally undetectable. Taking this idea seriously, it has been suggested in the famous “*fast scrambling conjecture*” that black holes should indeed be perfect scramblers, taking a time logarithmic in the number of degrees of freedom [9, 29]. Unfortunately, the microscopic models under consideration, most importantly the so-called *matrix models* [27], involve highly non-local interactions in interacting models that embody both bosonic and fermionic degrees of freedom, are notoriously difficult to solve, even on modern supercomputers. For this purpose, research on the fast scrambling conjecture has focused much on identifying proxies that share many similarities with the actual physical model, to get a handle on a precise quantitative understanding of the mechanisms that lead to such a fast scrambling. Classical models have been considered [30], small-dimensional quantum models [31], random processes precisely of the kind considered here, in the form of stochastically fluctuating local Hamiltonians [9], as well as random circuits [2]. It is one of the key motivations of the present work to provide tools for studies of this kind.

Indeed, stochastically fluctuating local Hamiltonians are less well understood than random circuits— and importantly, a precise understanding of the equivalences of mixing times seems urgently needed.

It should also be clear that *time-fluctuating processes* as such are ubiquitous in nature. For such processes, the dynamics is captured by a family of Hamiltonians of the form

$$H_t = H_0 + F_t, \quad (1)$$

such that both  $H_0$  and  $F_t$  are local Hamiltonians of a quantum system with many degree of freedom and  $F_t$  is randomly fluctuating in time. Any experimental setting in quantum mechanics will necessarily be interacting with a classical exterior in one way or the other. Many decoherence mechanisms can well be approximated by a classical degree of freedom fluctuating randomly in time. In fact, effects like magnetic field fluctuations are of this type, and so are Gaussian noisy processes in condensed matter physics. This type of noise is usually seen as a detrimental type of *decoherence*, deteriorating the coherence present in the quantum mechanical system. This connection to *local dissipative dynamics* will be made clear below. Again, a precise understanding of these effects and their impact seems desirable.

Again more technologically or pragmatically speaking, it should be clear that fluctuating Hamiltonians of the form (1) by no means have to reflect unwanted external noise. Quite to the contrary, in many applications in which random quantum circuits are envisioned, one can as well replace the quantum circuit by the mere time evolution under such a fluctuating Hamiltonian. In many situations this can lead to a significantly simplified prescription, compared to implementing precisely controlled quantum gates that are designed according to samples of some suitable classical probability distribution. That is to say, in a number of instances, fluctuating Hamiltonians can be seen as being vastly more feasible than random circuits that require the accurate realization of quantum gates, at least from the perspective of implementation.

Motivated by these considerations, in this work we investigate mixing properties of random quantum processes in quantum many-body systems. Specifically, we consider a family of time-fluctuating local Hamiltonians inducing a Brownian motion on the unitary group. We show that this locally generated Brownian motion gives rise to an efficient approximate unitary  $k$ -design of arbitrary order, i.e. all its moment operators converge to those of the Haar measure. Furthermore, the convergence rate is comparable to that of a random quantum circuit in discrete time. Our main technical contribution is a connection between the generator of the local diffusion and the Casimir element of the special unitary group. This allows us in turn to obtain an explicit uniform lower bound on the gap of the local generator, i.e. independent of the order  $k$ . Hence, our results also provide a class of probability measures on the unitary group, where the set of generated unitaries has the spectral gap property, albeit with an explicitly known constant [32, 33]. This might be an unexpected result, as the convergence time of the  $k$ -th moment increases with  $k$  for many processes.

We also show decoupling with almost linear scaling in the system size. We interpret the time-fluctuating Hamiltonian in the framework of a *continuous-time random walk*, relating it with the discrete random walk induced by random quantum circuits with Haar distribution. The continuous-time version has been first formalised by Montroll and Weiss [34] as a sequence of random transitions (jumps) spaced out by waiting times and been object of successive study [35], being applied to a wide range of fields of physics [36–38]. In particular, an exact correspondence between the accelerated steps of the random walk induced by random quantum circuits given in ref. [10] and the jumps of the continuous-time random walk generated by the fluctuating Hamiltonian infers a close similarity between the discrete circuit and the continuous process and can be hence used to relate results from these two settings. Much of the present work can hence be seen as providing a unifying framework to capture random quantum processes—continuous and discrete in time—under a single umbrella. By bringing notions of fluctuating Hamiltonians closer together with those of random quantum circuits, we provide a unified picture of mixing properties of random quantum processes. The results laid out here are expected to provide powerful technical tools to make further progress in those research questions for which such quantum processes having a classical random component are relevant, maybe most intriguingly the fast scrambling conjecture.

## 2. Preliminaries

In this section, we introduce basic notions and concepts that will be made use of when stating the main results. In the focus of attention will be the concept of a unitary design. A unitary design is a probability distribution over unitary matrices that mimics properties of the Haar measure, in a similar sense as a spherical design approximates the unit sphere. In order to capture approximate versions of unitary designs – which will feature strongly throughout this work – several norms will be relevant. We finish this section by introducing the concept of locally generated Brownian motion on the unitary group.

*2.1. Norms.* In this work, all Hilbert spaces and other vector spaces considered are finite dimensional, reflecting finite-dimensional quantum systems. We denote the vector space of linear operators on a Hilbert space  $\mathcal{H}$  by  $\mathcal{L}(\mathcal{H})$ . The *trace norm* of an operator  $X \in \mathcal{L}(\mathcal{H})$  is defined by

$$\|X\|_1 := \sum_i s_i(X), \quad (2)$$

where we denote by  $s_i(X)$  the  $i$ -th singular value of  $X$ .

For linear maps on operators we introduce the following two norms.

1. The *diamond norm* [39] of  $\mathcal{T} \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$  is defined to be

$$\|\mathcal{T}\|_\diamond := \sup_d \sup_{X \neq 0} \frac{\|(\mathcal{T} \otimes \mathcal{I}_d) X\|_1}{\|X\|_1}, \quad (3)$$

where  $\mathcal{I}_d$  denotes the identity element in  $\mathcal{L}(\mathcal{L}(\mathbb{C}^d))$ . The diamond norm is most meaningful to quantify how close two quantum channels are, reflecting physical processes.

2. The *infinity norm*  $\|\mathcal{T}\|_\infty$  of  $\mathcal{T}$  is given as the largest singular value of  $\mathcal{T}$ .

*2.2. Exact and approximate unitary designs.* In this work, we examine probability measures over the unitary group and their mixing properties. We refer to them as distributions on the unitary group  $\mathbb{U}(N)$ . A central role is played by the invariant distribution over all unitaries given by the *Haar measure*. In many applications, one is interested in distributions which approximate properties of the Haar measure but can be generated with limited resources (e.g. quantum circuits of a certain given depth or Hamiltonian quantum evolution generated by local Hamiltonians for a certain run-time). *Unitary  $k$ -designs* capture the ability of a distribution to mimic the properties of the Haar measure in the sense that expectation values of polynomials of a certain order  $k$  are equal to those of the Haar measure. As pointed out above, they have a wide range of applications in quantum algorithm design [3, 8, 18], in quantum state and process tomography [19, 20], and in notions of benchmarking [21] – basically as a powerful tool for *partial derandomisation*. Conceptually, they feature strongly in descriptions of equilibration, thermalisation and scrambling [2, 9, 29].

In order to make this concept of a unitary design precise, we define the  *$k$ -th moment operator*  $M_\mu^k$  on  $\mathcal{L}(\mathcal{H}^{\otimes k})$  with respect to a distribution  $\mu$  on  $\mathbb{U}(N)$  by

$$X \mapsto M_\mu^k(X) := \mathbb{E}_\mu [U^{\otimes k} X (U^\dagger)^{\otimes k}]. \quad (4)$$

Exact unitary designs as well as suitable approximate versions thereof can be defined in terms of  $M_\mu^k$ .

**Definition 1 (Unitary designs).** *Let  $\mu$  be a distribution over the unitary group  $\mathbb{U}(N)$ . Then  $\mu$  is an  $\varepsilon$ -approximate unitary  $k$ -design if*

$$\|M_\mu^k - M_{\text{Haar}}^k\|_\diamond \leq \varepsilon. \quad (5)$$

For  $\varepsilon = 0$ , the distribution  $\mu$  is also called an exact unitary  $k$ -design.

Physically implementing an exact unitary design is in general neither an obvious nor an efficient task. Fortunately, for a plethora of applications, exactness of a design is not required. Instead, we are usually interested in obtaining *approximate unitary designs*, i.e., distributions which behave similarly as the Haar measure and which can be implemented efficiently. We would also like to note that there are different formal definitions of approximate unitary designs, each of which being equipped with a different interpretation and being relevant in a different context; a close examination has been done in ref. [40].

An important method for obtaining a bound on  $\varepsilon$  is to analyse the gap of the moment operator  $M_\mu^k$ , leading to the following definition.

**Definition 2 (Tensor product expanders).** *A distribution  $\mu$  on the unitary group  $\mathbb{U}(N)$  is a quantum  $(\lambda, k)$ -tensor product expander if*

$$\|M_\mu^k - M_{\text{Haar}}^k\|_\infty \leq \lambda. \quad (6)$$

The following lemma links this definition to the one of designs.

**Lemma 3 (Criterion for being an approximate unitary design [40, Lemma 2.2.14]).** *Let  $\mu$  be a distribution on  $\mathbb{U}(N)$ . If  $\mu$  is a quantum  $(\lambda, k)$ -tensor product expander, then  $\mu$  is also an  $\varepsilon$ -approximate  $k$ -design with  $\varepsilon = N^k \lambda$ .*

The expectation value of polynomials with respect to the Haar measure over the unitary group can be understood in terms of the *Schur-Weyl duality*. This says that  $M_{\text{Haar}}^k$  is an orthogonal projection onto the span of operators representing a permutation of the  $k$  tensor copies of  $\mathcal{H}$  (see ref. [41, Proposition 2.2] for a complete description). This means that all elements of this space are eigenvectors with unit eigenvalues, while the complement space belongs to the kernel.

A similar connection between eigenvalues and eigenspaces can be made for *universal distributions* as we will see in Lemma 5 below. In order to amplify closeness of a distribution  $\mu$  on  $\mathbb{U}(N)$  to the Haar measure, one can convolute it  $\ell$  times with itself and obtain a new measure  $\mu^{*\ell}$  on  $\mathbb{U}(N)$ . One can effectively draw a unitary  $U$  from  $\mu^{*\ell}$  by drawing  $\ell$  unitaries  $U_1, U_2, \dots, U_\ell$  independently from  $\mu$  and take  $U$  as the product  $U = U_1 U_2 \dots U_\ell$ . Importantly, it holds that

$$M_{\mu^{*\ell}}^k = (M_\mu^k)^\ell. \quad (7)$$

If the support of  $\mu^{*\ell}$  becomes dense in  $\mathbb{U}(N)$  for large  $\ell$  we call  $\mu$  *universal*. More precisely, a universal distribution can be defined as follows.

**Definition 4 (Universal distribution).** *Let  $\mu$  be a distribution on  $\mathbb{U}(N)$ . Then  $\mu$  is said to be universal if for all  $V \in U(N)$  and any  $\delta > 0$  there exists a positive integer  $\ell$  such that*

$$\mu^{*\ell}(B_\delta(V)) > 0, \quad (8)$$

where  $B_\delta(V)$  is the neighbourhood of  $V$  with radius  $\delta > 0$ .

Here, the canonical way to capture the radius is in terms of the geodesic distance on  $\mathbb{U}(N)$ . It should be clear, however, that any other equivalent metric gives rise to the same definition of universality. This definition can be seen as a generalisation of a universal gate set: if  $\mu$  is the uniform distribution over finitely many unitaries then this set of unitaries is universal if and only if  $\mu$  is universal. Universal distributions induce moment operators satisfying the following property for all orders  $k$ .

**Lemma 5 (Lemma 3.7 in ref. [10]).** *Let  $\mu$  be a distribution on  $\mathbb{U}(N)$ . Then all eigenvectors of  $M_{\text{Haar}}^k$  with unit eigenvalue are eigenvectors of  $M_\mu^k$  with unit eigenvalue. Additionally, if  $\mu$  is universal then  $\mu$  is  $k$ -copy gapped for any positive integer  $k$ . This means that*

$$\|M_\mu^k - M_{\text{Haar}}^k\|_\infty < 1. \quad (9)$$

As a consequence,  $M_{\mu^{*\ell}}^k$  converges to  $M_{\text{Haar}}^k$  for  $\ell \rightarrow \infty$ . For many practical applications, however, a bound on the convergence rate is needed. Below, we will extend such a bound from quantum circuits [12] to locally generated Brownian motion on  $\mathbb{U}(N)$ .

*2.3. Locally generated Brownian motion on the unitary group.* In this section, we define the central objects studied in this work: Brownian motions on the unitary group. These are continuous-time stochastic processes describing the unitary evolution of a quantum system with a fluctuating Hamiltonian and whose increments satisfy specific properties. In this way, a distribution on  $\mathbb{U}(N)$  is induced, which changes over time and eventually converges to the Haar measure for arbitrary moments.

A deterministic family of Hamiltonians  $t \mapsto H_t$  depending continuously on time  $t \in \mathbb{R}$  generates a unitary time evolution (see refs. [42, 43]) via the time-ordered exponential

$$U_t = \mathcal{T} \left[ \exp \left\{ -i \int_0^t H_s ds \right\} \right]. \quad (10)$$

Conversely,  $H_t$  can be recovered from the *increments*

$$U_t^\dagger U_{t+\Delta t} = \mathbb{1} - i\Delta t \langle H_t \rangle + O(\Delta t^2) \quad (11)$$

by taking the limit  $\Delta t \rightarrow 0$ , where  $\langle H_t \rangle$  denotes the time average of  $H_s$  over the interval  $[t, t + \Delta t]$ .

In this work, we investigate how well a Brownian motion  $U_t$  that has stochastic increments of the form of a local fluctuating Hamiltonian  $H_t$  generates an approximate unitary  $k$ -design as a function of time  $t$ . First of all, we define Brownian motion as follows (c.f. refs. [44–46]):

**Definition 6 (Brownian motion on the unitary group).** *A process  $U_t$  on the unitary group  $\mathbb{U}(N)$  is called Brownian motion if the following conditions are satisfied.*

1. *For all  $0 < t_1 < t_2 < \dots < t_n$ , the (left) increments  $U_{t_1} U_0^\dagger, U_{t_2} U_{t_1}^\dagger, \dots, U_{t_n} U_{t_{n-1}}^\dagger$  are independent.*
2. *For any time  $t \geq 0$ , the increments are stationary, i.e., for any  $\Delta t > 0$  the increment  $U_{t+\Delta t} U_t^\dagger$  is equal in distribution to  $U_{\Delta t} U_0^\dagger$ .*
3. *The paths  $t \mapsto U_t$  are continuous almost surely.*

Brownian motion  $U_t$  on the Lie group  $\mathbb{U}(N)$  corresponds to Brownian motion  $W_t$  on the Lie algebra  $\mathfrak{u}(N)$  through the exponential map, which for a matrix Lie group is given by the series  $\exp(X) = \sum_{n=0}^{\infty} X^n/n!$ . More precisely, one can construct Brownian motion on  $\mathbb{U}(N)$  by *injecting the differential* of a Brownian motion from  $\mathfrak{u}(N)$  via the product integral of the exponential map,

$$U_t = \lim_{\Delta t \rightarrow 0} \prod_{\ell=t/\Delta t}^1 \exp \{ W_{\ell\Delta t} - W_{(\ell-1)\Delta t} \} U_0, \quad (12)$$

see ref. [47, Chapter 4.8] for a complete proof of existence and uniqueness of the process, and additionally ref. [48, p. 226].

The *Hamiltonian increments* are denoted by

$$H_{\ell, \Delta t} := i\Theta_{\ell, \Delta t} \quad (13)$$

with

$$\Theta_{\ell, \Delta t} := \frac{1}{\Delta t} [W_{\ell\Delta t} - W_{(\ell-1)\Delta t}] \quad (14)$$

being the increments in the Lie algebra  $\mathfrak{u}(N)$ .

*2.4. Local Hamiltonian increments.* We now turn to describing the *local* Hamiltonian increments on the physical quantum system consisting of  $n$  subsystems of dimension  $d$ , so that  $N$  becomes  $d^n$ . Those subsystems interact according to an interaction pattern captured by an interaction graph with vertex set  $V$  and edge set  $E$ . In the special case of  $d = 2$ , this is referred to the *qubit case*, and the system is an *n-qubit system*.

We assume that  $\Theta_{\ell, \Delta t}$  from eq. (14) is local with respect to an interaction graph  $(V, E)$ , where each vertex in  $V$  corresponds to a  $d$ -level subsystem. Only qudits connected by an edge  $e \in E$  may interact, i.e.,

$$\Theta_{\ell, \Delta t} = \sum_{e \in E} \theta_{\ell, \Delta t}^{(e)}, \quad (15)$$

where each local term  $\theta_{\ell, \Delta t}^{(e)}$  is supported on  $e$ . The local terms are explicitly given by

$$\theta_{\ell, \Delta t}^{(e)} = -i h_0^{(e)} + \sum_{\mu} A_{\mu}^{(e)} \xi_{\ell, \Delta t}^{(e, \mu)}, \quad (16)$$

where we specify each term in this equation in the following.  $h_0^{(e)}$  are deterministic Hermitian operators reflecting a constant drift in the evolution. Each noise operator  $A_{\mu}^{(e)}$  acts on the two vertices connected by  $e$  as  $A_{\mu}$  and as the identity elsewhere.  $\{A_{\mu}\}_{\mu}$  is a basis of the real Lie algebra

$$\mathfrak{u}(d^2) := \{X \in \mathbb{C}^{d^2 \times d^2} : X = -X^{\dagger}\}. \quad (17)$$

$\xi_k^{(e, \mu)}$  are real random variables representing the noise. We assume that the noise satisfies

$$\mathbb{E} \left[ \xi_{\ell, \Delta t}^{(e, \mu)} \right] = 0, \quad (18)$$

$$\mathbb{E} \left[ \xi_{\ell, \Delta t}^{(e, \mu)} \xi_{\ell', \Delta t}^{(e', \mu')} \right] = -\frac{a}{\Delta t} \delta_{\ell, \ell'} \delta_{e, e'} \kappa_{\mu, \mu'}^{-1}, \quad (19)$$

where  $a > 0$  is an arbitrary constant and the matrix  $\kappa$  is defined by

$$\kappa_{\mu, \nu} := -2d^2 \text{Tr}(A_{\mu}^{\dagger} A_{\nu}). \quad (20)$$

As we will explain later, this matrix is in fact the *Killing metric tensor* associated with the basis  $\{A_{\mu}\}_{\mu}$ .

*Remark 7 (Orthonormal basis).* If the basis  $\{A_{\mu}\}_{\mu}$  is orthonormal then our assumption (19) on the covariance simplifies to

$$\mathbb{E} \left[ \xi_{\ell, \Delta t}^{(e, \mu)} \xi_{\ell', \Delta t}^{(e', \mu')} \right] = \frac{a}{2d^2 \Delta t} \delta_{\ell, \ell'} \delta_{e, e'} \delta_{\mu, \mu'} \quad (21)$$

which represents *white noise*. This happens, for instance, if we choose the Pauli basis (see the example processes in Section 5.4 and Section 3.2).

*Remark 8 (Overcomplete sets of operators).* Additionally, we may consider an overcomplete set of operators  $\{A_{\mu}\} \in \mathfrak{u}(d^2)$  as long as they give rise to a negative contribution to the generator (52), since this will increase the gap of the moment operator induced by the stochastic evolution and hence make the convergence even faster.

The above described Brownian motion with Hamiltonian increments as in eq. (15) with the specified  $\theta_{\ell}^{(e)}$  induces a Brownian motion  $U_t$  on the unitary group. We denote the distribution of  $U_t$  at time  $t$  by  $\text{SLH}(t)$  and write the according expectation as  $\mathbb{E}_{\text{SLH}(t)}$ .



### 3. Main results

In this section, we present the two main results on mixing properties for local stochastic Hamiltonian evolutions. Theorem 9, together with Corollary 10, asserts that the distribution over unitaries induced by the Brownian motion as in Definition 6 is a quantum tensor product expander and hence an approximate unitary  $k$ -design after a run time scaling polynomially in  $k$  and linearly in the system size  $n$ . This means that it is suitable to efficiently reproduce certain properties of the Haar measure. Theorem 13 on decoupling says that any subsystem affected by a stochastic evolution reproducing Brownian motion becomes uncorrelated with respect to a second (possibly initially correlated) subsystem in almost linear run time in system size.

*3.1. Tensor product expanders and approximate unitary designs.* For any of the time-fluctuating local Hamiltonians fulfilling the description in Section 2.3, each moment of the generated unitary process becomes close to that of the Haar measure after a sufficiently long run time. This result can be expressed in terms of quantum tensor product expanders or approximate unitary designs as follows.

**Theorem 9 (Local Brownian motions on  $\mathbb{U}(d^n)$  are quantum  $(\lambda, k)$ -tensor product expanders).** *Let  $U_T$  be a unitary Brownian motions with the increments (15) with the interaction graph  $(V, E)$  being either a complete graph or a 1D nearest neighbour graph. Then, for any run time*

$$T \geq 850 \lceil \log_d(4k) \rceil^2 d^2 k^5 k^{3.1/\ln(d)} \frac{\ln(1/\lambda)}{a}, \quad (22)$$

*$U_T$  is a quantum  $(\lambda, k)$ -tensor product expander.*

Then, using Lemma 3, we immediately obtain the subsequent corollary.

**Corollary 10 (Approximate unitary  $k$ -designs).** *For any run time*

$$T \geq 850 \lceil \log_d(4k) \rceil^2 d^2 k^5 k^{3.1/\ln(d)} \frac{nk \log(d) + \ln(1/\varepsilon)}{a}, \quad (23)$$

*$U_T$  is an  $\varepsilon$ -approximate unitary  $k$ -design.*

Theorem 9 can be seen as a unifying statement on random quantum processes. It extends the results on random local quantum circuits, as considered in ref. [12, Corollary 7], to continuous time dynamics under fluctuating Hamiltonians. Note that the scaling of the minimal runtime required for the generating of a unitary  $k$ -design is by a factor of  $n$  smaller with respect to the circuit setting. This is due to the number of Hamiltonian interactions per time step growing linearly in the system size for a 1D graph (which is, as discussed in Lemma 22, the slowest setting among all complete graphs), while for local random quantum circuits only one gate per step is applied. If we re-scale the stochastic Hamiltonian with a pre-factor  $O(1/\sqrt{n})$  so that the  $k$ -th moment operator may be written in the same form as the one induced by a random quantum circuit, i.e. (cf. eqs. (69)-(71) for derivation),

$$M^k = \frac{1}{n} \sum_j (m^k)^{j,j+1}, \quad (24)$$

where  $(m^k)^{j,j+1}$  denotes the local moment operator applied on qubits  $j$  and  $j+1$ , then we would instead obtain the same scaling for the mixing time. Therefore, we can consider the two scenarios as perfectly compatible. As already discussed in the introduction, the time-continuous case might often be the more natural and easier implementable choice in applications, such as tomography or random benchmarking: They do not require the explicit implementation of random bipartite unitary matrices and often the natural fluctuations present in a system are already a good approximation of a locally generated Brownian motion.

The full proof of Theorem 9 is given in Section 5. Its outline is the following.

*Proof idea of Theorem 9.* We express the  $k$ -th moment operator  $M_{\text{SLH}(t)}^k$  in terms of a generator  $G_k$  so that  $M_{\text{SLH}(t)}^k = \exp(tG_k)$ . This generator has again the same locality structure as Hamiltonian increments in eq. (15), i.e.,

$$G_k = \sum_{e \in E} g_k^{(e)}. \quad (25)$$

The crucial point of our proof is to obtain a lower bound to the spectral gap for the invariant subspace of  $g_k^{(e)}$ . This gap can be recovered thanks to a simple relation between  $g_k^{(e)}$  and the so-called *Casimir element* of  $\mathfrak{su}(d^2)$  in a certain (reducible) representation denoted by  $\pi_{k,k}$ . Since the eigenvalues of the Casimir element in each irreducible representation of  $\mathfrak{su}(N)$  are well known, we can determine the spectrum in the representation  $\pi_{k,k}$  from its irreducible decomposition. In particular, using an argument based on the shape of Young diagrams, we derive a local gap,  $\Delta(g_k)$ , independently of  $k$  from which we finally deduce, applying results from ref. [12], a gap for  $G_k$ . This establishes the bound on the mixing time.  $\square$

**3.2. Fast decoupling with stochastic quantum Hamiltonians.** The second result is concerned with fast decoupling properties of the random evolutions considered in this work. In the following, we restrict our analysis to the qubit case in which  $d = 2$  and where the noise operators are given by the Pauli matrices  $\sigma_0 = \mathbb{1}, \sigma_1, \sigma_2,$  and  $\sigma_3$ . We connect to and extend the result on fast decoupling given in ref. [4] for a continuous-time evolution. Inspired by the Hamiltonian given in ref. [9], for which a reading of the *fast scrambling conjecture* has been studied, we set the increments in eqs. (15) and (16) of the Brownian motion to be

$$\Theta_{n,\ell,\Delta t} = -i H_{n,\ell,\Delta t} := -i \left( \frac{2}{n(n-1)} \right)^{1/2} \sum_{j < k} \sum_{\alpha, \beta=0}^3 \sigma_\alpha^j \otimes \sigma_\beta^k \xi_{\ell,\Delta t}^{(j,k,\alpha,\beta)}, \quad (26)$$

where  $\sigma_\alpha^j \otimes \sigma_\beta^k$  means that  $\sigma_\alpha \otimes \sigma_\beta$  is applied on qubits labeled  $j$  and  $k$ , respectively. We recall that  $\xi_{\ell,\Delta t}^{(j,k,\alpha,\beta)}$  are i.i.d. real random variables with zero mean and covariance

$$\mathbb{E} \left[ \xi_{\ell,\Delta t}^{(j,k,\alpha,\beta)} \xi_{\ell',\Delta t}^{(j',k',\alpha',\beta')} \right] = \frac{1}{\Delta t} \delta_{\ell,\ell'} \delta_{k,k'} \delta_{j,j'} \delta_{\alpha,\alpha'} \delta_{\beta,\beta'} \quad \forall j, k, \alpha, \beta \quad (27)$$

which is obtained from eq. (21) by choosing  $a$ . The pre-factor of  $(2/(n(n-1)))^{1/2}$  is chosen so that the initial rate of diffusion of a local operator scales as  $O(1/n)$ . This is to normalize the time scale for the diffusion process in order to compare it with the random quantum circuit model in refs. [4, 10], where the probability that a local operator experiences a random gate is  $2/n$  per discrete time step.

In what follows, we refer to an  $n$ -fold tensor product of Pauli operators (including the identity) on qubits as a *Pauli string* and denote it by  $\sigma_\alpha$ , with an  $n$ -dimensional index  $\alpha = \{0, 1, 2, 3\}^n$  representing the label of each sub-element.

For our results, we need to define the permutation invariance property. This condition is required to deduce a dominant probability distribution on the final Pauli coefficients when starting with an analysis of the evolution of the Pauli weights. Indeed, the random walk on Pauli weights does not distinguish among strings with same support size but different support, hence it provides the probability distribution for each set of strings with the same support size, but not on Pauli strings taken singularly.

The permutation invariance property has already been debated in the proof of ref. [10] showing that random quantum circuits with Haar measure are approximate unitary 2-designs. In ref. [49] it has been discussed that this essential condition in the proof had not been granted and an argument making use on random transpositions based on the work of Diaconis (see refs. [17, 50]) has been put forward solving this issue. Since we cannot prove that permutation invariance is achieved with sufficiently high probability by the stochastic Hamiltonian evolution itself within a run time scaling almost linearly in  $n$ , we impose it as a pre-condition for the initial state. Actually, we can relax the condition and ask for a ‘‘large portion’’

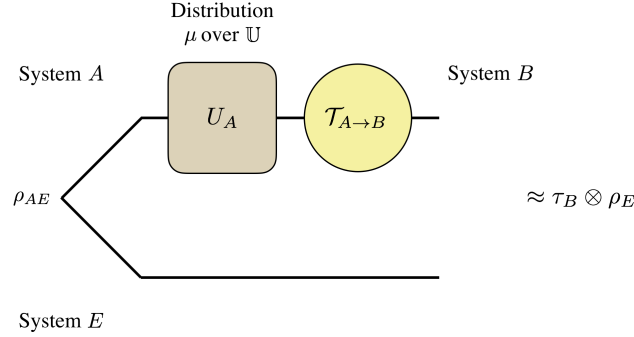


Figure 1: In the decoupling theorem, an initial bipartite state  $\rho_{AE}$  is affected by a unitary evolution  $U_A$  chosen at according to a certain distribution  $\mu$ . Then, subsystem  $A$  is mapped to another subsystem  $B$  through a completely positive map  $\mathcal{T}_{A \rightarrow B}$ . Finally, the distance between the final state and the product state  $\tau_B \otimes \rho_E$  is characterised by entropy measures.

of the qubits, but not necessarily all, to be invariant with respect to an arbitrary permutation. This allows us to apply our result to a larger family of states, for instance those whose support is very small. More formally, we define the permutation invariance property as follows.

**Definition 11 (Permutation invariance property).** Let  $0 \leq \gamma < 1$ . Let  $\sigma_{\pi(\mu)}$  denote a Pauli string whose label is given by interchanging the sub-indices of  $\mu$  according to the permutation  $\pi$ . Then, for an arbitrary quantum state  $\rho$  of a  $n$ -qubit system, we say that it satisfies the  $\gamma$ -permutation invariance property if there exists a subset of  $(1 - \gamma)n$  qubits which is invariant with respect to any permutation, i.e.,

$$\text{Tr}[\sigma_\mu \rho] = \text{Tr}[\sigma_{\pi(\mu)} \rho] \quad (28)$$

for every Pauli string  $\sigma_\mu$  and every permutation  $\pi$  on this subset of qubits.

Note that any state  $\rho$  with  $|\text{supp}(\rho)| \leq \gamma n$  is permutation invariant with respect to this definition.

For the decoupling theorem, two initially correlated subsystems, which we denote by  $A$  and  $E$ , are considered and their joint state is given by  $\rho_{AE}$ . Then,  $A$  is affected by a unitary evolution describing Brownian motion. Subsequently a completely positive map  $\mathcal{T}$  maps it into another system  $B$ . The map  $\mathcal{T}_{A \rightarrow B}$  can be equivalently described by its *Choi-Jamiolkowski isomorphism* [51], given by

$$\tau_{A'B} = (\mathcal{I}_{A'} \otimes \mathcal{T}_{A \rightarrow B}) (|\psi\rangle \langle \psi|)_{A'A} \quad (29)$$

with

$$|\psi\rangle_{A'A} = |A|^{-1/2} \sum_j |j\rangle_{A'} \otimes |j\rangle_A \quad (30)$$

being a maximally entangled state vector and  $A'$  a copy of  $A$  and  $|A|$  denoting the Hilbert space dimension of  $A$ .

A decoupling theorem quantifies the distance, in terms of the 1-norm, of the final state of the above described evolution from the product state  $\tau_B \otimes \rho_E$ , considering the expectation over the unitary distribution. The bound on this expression is characterised by entropy measures of the initial state  $\rho_{AE}$  and of the Choi-Jamiolkowski representation of the map  $\mathcal{T}$ . More specifically, the entropy measure is the *conditional collision entropy of  $A$  given  $B$* , defined as (see ref. [52, Definition 2.9])

$$H_2(A|B)_\rho := \sup_{\sigma_B \in \mathcal{S}_=(\mathcal{H}_B)} -\log \text{tr} \left[ \left( (\mathbb{1}_A \otimes \sigma_B^{-1/4}) \rho_{AB} (\mathbb{1}_A \otimes \sigma_B^{-1/4}) \right)^2 \right]. \quad (31)$$

Our second main result states that, under a unitary evolution describing Brownian motion, decoupling is achieved with a run time scaling almost linear in system size. We denote the Pauli basis coefficients after a continuous-time evolution with run time  $T$  as

$$Q^T(\mu, \nu) := \frac{1}{4^n} \text{Tr} \left[ \sigma_\nu \otimes \sigma_\nu M_{n, \text{SLH}(T)}^{k=2}(\sigma_\mu \otimes \sigma_\mu) \right]. \quad (32)$$

First of all, we derive an upper bound on the distance between the distribution of these Pauli coefficients and a distribution which is close to the uniform one.

**Theorem 12 (Mixing condition for Pauli coefficients).** *For any constants  $\delta \in (0, 1/16)$ ,  $\eta \in (0, 1)$  there exist constants  $\varsigma > 0$  and  $0 < \gamma_0 \leq 1/2$  such that for a total run time  $T \geq \varsigma n \log^2 n$  and large enough  $n$*

$$\sum_{\nu \in \{0,1,2,3\}^n, \nu \neq 0} |Q^T(\mu, \nu) - p_\delta(\nu)| \leq \frac{1}{(3-\eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)} \quad (33)$$

where  $\sigma_\mu$  is an arbitrary string whose support has size  $\ell$  and has a subset of  $(1-\gamma)n$  qubits, with  $\gamma < \gamma_0$ , which is invariant with respect to any permutation, and  $p_\delta$  is a (possibly sub-normalised) distribution on Pauli strings such that:

$$p_\delta(\nu) \leq \frac{5^{\delta n}}{4^n - 1} \quad \forall \nu. \quad (34)$$

From Theorem 12 we obtain the final result on decoupling, which can be seen as a statement unifying the description of continuous and discrete processes. Specifically, it links the stochastic local Hamiltonian evolution (up to the permutation invariance property assumed for the initial state) to that of a random quantum circuit under Haar distribution given in ref. [4, Theorem 3.2], and establishes a connection between the discrete random walk induced by a random quantum circuit and the continuous-time random walk implied by the stochastic local Hamiltonian evolution. The actual proof of Theorem 13 from Theorem 12 follows by arguments analogous to the random quantum circuit case given in ref. [4].

**Theorem 13 (Fast decoupling).** *Consider a bipartite quantum state  $\rho_{AE} \in \mathcal{S}_{AE}$  of an  $n$ -qubit system  $A$  coupled with some other system  $E$ . Let then  $\rho_{AE}$  undergo a unitary evolution  $U_t$  induced by stochastic local Hamiltonian increments as in eq. (26) acting upon system  $A$ , followed by a completely positive trace preserving map  $\mathcal{T} : \mathcal{S}_A \rightarrow \mathcal{S}_B$  which maps from  $A$  to another system  $B$ . Let  $\tau_{A'B}$  denote the Choi-Jamiolkowski isomorph of  $\mathcal{T}$ . Then, for any  $\delta \in (0, 1/16)$  there exist  $\varsigma > 0$  and  $0 < \gamma_0 \leq 1/2$  such that for all  $\gamma$ -permutation invariant states with  $\gamma < \gamma_0$  and total run times  $T \geq \varsigma n \log^2 n$  and for large enough  $n$*

$$\mathbb{E}_{\text{SLH}(T)} \left\{ \left\| \mathcal{T} \left( U_T \rho_{AE} U_T^\dagger \right) - \tau_B \otimes \rho_E \right\|_1 \right\} \leq \left( \frac{1}{\text{poly}(n)} + 5^{\delta n} \cdot 2^{-H_2(A|B)_\tau - H_2(A|E)_\rho} \right)^{1/2}, \quad (35)$$

where  $\text{SLH}(T)$  denotes the distribution over the unitary group induced by the Brownian motion with run time  $T$ .

*Proof idea of Theorem 12.* As a first step, we observe the evolution of the support size of Pauli strings during the continuous-time process, looking for the probability to reach a support size within the interval  $[3/4 - \delta, 3/4 + \delta]$  when starting from an arbitrary Pauli string of weight  $\ell$ . The key point is that the jumps of the continuous-time random walk on the Pauli weights induced by the fluctuating Hamiltonian correspond exactly to the accelerated Markov chain (i.e., the chain conditioned on moving) deduced from the random quantum circuit with Haar distribution as in refs. [4, 10]. This displays an equivalence between the two settings and allows us to link results relating to the two types of processes. We then compute the contribution of each string to this probability to obtain the desired condition for Pauli coefficients. More precisely, we do so by using the permutation invariance property and the uniform randomisation of the Pauli basis  $\sigma_1, \sigma_2, \sigma_3$  – the latter achieved by the evolution itself. We can then infer that almost all Pauli strings with the same support size share the same probability.  $\square$

#### 4. Applications

We discuss in the following two interesting applications for Brownian motion on the unitary group, namely dissipative dynamics and black holes scrambling, and hint at a third one, making use of fluctuating Hamiltonian dynamics in quantum information processing.

*4.1. Dissipative dynamics arising from fluctuations.* As pointed out before, there is an intimate relationship between time-fluctuating dynamics and *Markovian dissipative evolution*, a connection that will be made manifest in this subsection. Brownian motion  $U_t$  on the unitary group yields an average dynamics given by

$$\rho(t) := \mathbb{E}[U_t \rho U_t^\dagger], \quad (36)$$

which describes a dissipative quantum Markovian evolution of the state  $\rho$ . In this sense, *time-fluctuating classical noise* is precisely a specific source of *dissipation*. As pointed out in the introduction, processes of this kind are ubiquitous in nature and in quantum systems in the laboratory: they originate whenever one does not have perfect control over the classical control parameters. At the same time, they can be used as a tool. Indeed, the use of *controlled dissipative Markovian dynamics* has received much interest in recent years [53–55].

The generator of the dynamical semi-group given by the evolution (36), and more generally of the  $k$ -th moment operator from eq. (4), is calculated below in Lemma 15. For  $k = 1$ , the generator is explicitly given in the following.

**Proposition 14 (Fluctuations as dissipative processes).** *Let  $U_t$  be a Brownian motion with increments  $\Theta_{\Delta t}$  as in eq. (14). Write  $\Theta_{\Delta t}$  as*

$$\Theta_{\Delta t} = -i H_0 + F_{\Delta t}, \quad (37)$$

where  $-i H_0$  and  $F_{\Delta t}$  are its anti-Hermitian time constant and fluctuating parts, respectively, with

$$F_{\Delta t} = \sum_{\mu} B_{\mu} \xi_{\Delta t}^{\mu}, \quad B_{\mu}^{\dagger} = -B_{\mu}, \quad \mathbb{E}[\xi_{\Delta t}^{\mu}] = 0, \quad \text{and} \quad \mathbb{E}[\xi_{\Delta t}^{\mu} \xi_{\Delta t}^{\nu}] = -\frac{a}{\Delta t} \delta_{\mu, \nu}.$$

Then  $\rho(t) = \mathbb{E}[U_t \rho U_t^\dagger]$  gives rise to a quantum dynamical semi-group and evolves according to the Lindblad equation

$$\frac{d}{dt} \rho(t) = -i [H_0, \rho(t)] - a \sum_{\mu} \left( B_{\mu} \rho B_{\mu}^{\dagger} - \frac{1}{2} (B_{\mu}^{\dagger} B_{\mu} \rho + \rho B_{\mu}^{\dagger} B_{\mu}) \right) \quad (38)$$

with  $\rho(0) = \rho$ .

This proposition is proven after Lemma 15.

*4.2. Applications for fast scrambling.* In the last decade, black holes have been considered from a quantum information perspective, providing toy models and a fresh perspective to the field. In particular it has been conjectured that they are *fast scramblers* [29, 56, 57]. A system is *scrambled* when any previous perturbation has been thoroughly spread among the degrees of freedom so that to recover information contained in the perturbation one should access simultaneously a large fraction of the entire system. The minimum time for mixing information is then called *scrambling time*. More specifically, in ref. [29] three hypotheses have been outlined: the most rapid scramblers take logarithmic time in the degrees of freedom, the bound is saturated for matrix quantum mechanics, i.e., systems whose degrees of freedom are  $n \times n$  matrices, black holes are the fastest scramblers in nature. The authors of ref. [9] brought evidence about the conjectures regarding scrambling in logarithmic time by investigating Brownian quantum circuit and Ising model on sparse random graphs. There are two related mixing conditions for unitary dynamics that satisfy the requirements for scrambling, as discussed in refs. [29, 56], or ref.

[57], respectively. The relation between our results and both of these conditions will be discussed in the following.

In ref. [56], one considers the black hole's internal system  $A$  and the radiated environment  $E$ . Furthermore, one defines an additional reference system  $R$ , initially maximally entangled with a quantum memory system  $M$  that is subsequently thrown into the black hole. As the Hawking radiation leaks out, we would like  $R$  to become maximally entangled with a subsystem of  $E$  on which we can have control, hence having recovered the initial state of  $M$ , and so interpreting the black hole as a mirror (see Fig. 2). This may be translated into a scrambling condition through a decoupling theorem. As the black hole evaporates,  $A$  shrinks into a smaller system  $B$  which decouples from  $R$ . More formally, this means that

$$\mathbb{E}_{\text{Haar}} \left\{ \left\| \text{Tr}_{A \setminus B} \left( U_A \rho_{AR} U_A^\dagger \right) - \frac{\mathbb{1}_B}{|B|} \otimes \rho_R \right\|_1 \right\} \leq 2^{-\gamma}, \quad (39)$$

where  $\rho_{AE}$  is a quantum state where subsystem  $E$  shares  $m$  Bell pair with  $A$ , and  $A$  is otherwise mixed, and  $\gamma$  is the difference between the number of qubits emitted as Hawking radiation and the number of qubits of system  $M$ . The approximate statement

$$\mathbb{E}_\omega \left\{ \left\| \text{Tr}_{A \setminus B} \left( U_A \rho_{AR} U_A^\dagger \right) - \frac{\mathbb{1}_B}{|B|} \otimes \rho_R \right\|_1 \right\} \leq \sqrt{4^{-\gamma} + 4^m \epsilon} \quad (40)$$

is satisfied in expectation for an ensemble of unitary transformations  $\omega$ , which is an approximate 2-design in the sense that the Pauli coefficients are close to the uniform distribution, i.e.

$$\sum_{\nu \neq 0} |q_\omega(\mu, \nu) - q_u(\mu, \nu)| \leq \epsilon \quad \forall \mu, \quad (41)$$

where

$$q_u(\mu, \nu) = \frac{1}{4^n - 1} \quad \forall \mu, \nu \quad \text{and} \quad q_\omega(\mu, \nu) = \frac{1}{4^n} \text{Tr} [\sigma_\nu \otimes \sigma_\nu M_\omega^{k=2}(\sigma_\mu \otimes \sigma_\mu)]. \quad (42)$$

This condition was shown in ref. [10] to be satisfied by a random quantum circuit of size  $O(n \log n)$  (when  $\epsilon = 1/\text{poly}(n)$ ) and analogously by a stochastic local Hamiltonian, according to the analysis on random walk in Section 6 and following the same reasoning as in ref. [10], with a run time  $T = O(n \log n)$ . However, in order to compare time scales with ref. [9], we take the same convention and divide the global scrambling time by the time to scramble a single subsystem ; in this case we obtain a scrambling time of  $\tau_* = O(\log n)$ . Hence, our work also provides an alternative proof for the scaling

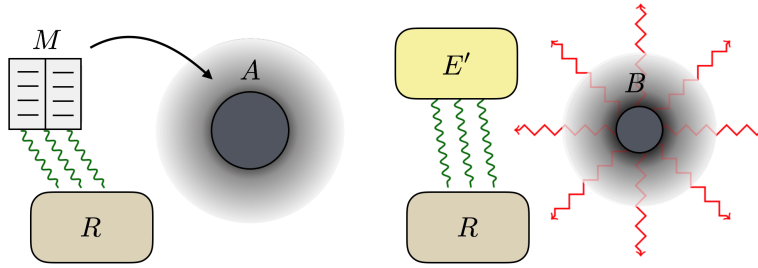


Figure 2: A quantum memory system  $M$  is initially entangled with a reference system denoted by  $R$  and is subsequently thrown into a black hole  $A$  (left picture). As the black hole leaks out Hawking radiation, it shrinks into a smaller system  $B$ . When a controlled subsystem of the irradiated environment  $E'$ , has become maximally entangled with the reference system  $R$ , the initial information  $M$  has been mirrored (right picture).

of the scrambling time in ref. [9], although our argument does not involve any intermediate conjectures, such as the final statements of ref. [9, Appendix B].

In ref. [57], a slightly different scrambling condition is required for the unitarity of black hole evaporation to hold, given postselection on the final state at the singularity inside the black hole. One considers the composite system  $\mathcal{H}_M \otimes \mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{out}}$  representing the infalling matter, the infalling negative energy Hawking radiation behind the event horizon and the outgoing positive energy Hawking radiation outside the horizon, respectively. Again, one defines a reference system  $S$  which is maximally entangled with a subsystem  $M_1 \subset M$ . After the application of a random unitary transformation  $U$  on  $\mathcal{H}_M \otimes \mathcal{H}_{\text{in}}$  and subsequently tracing out the complement subsystem of  $S$ , we have (cf. [57, eq.(3)]):

$$\mathbb{E}_{\text{Haar}} \left\{ \left\| \text{Tr}_{\bar{S}} (U \rho U^\dagger) - \frac{\mathbb{1}_S}{|S|} \right\|_1 \right\} \leq \sqrt{\frac{|\mathcal{H}_{M_1}|}{|\mathcal{H}_{\text{in}}|}}. \quad (43)$$

A relaxed version of this bound, namely

$$\mathbb{E}_\omega \left\{ \left\| \text{Tr}_{\bar{S}} (U \rho U^\dagger) - \frac{\mathbb{1}_S}{|S|} \right\|_1 \right\} \leq \sqrt{5^{\delta n} \frac{|\mathcal{H}_{M_1}|}{|\mathcal{H}_{\text{in}}|} + \frac{1}{\text{poly}(n)}}, \quad (44)$$

where  $n = \log_2 (|\mathcal{H}_{\text{in}}| |\mathcal{H}_M|)$ , follows from the condition

$$\sum_{\nu \neq 0} |q_\omega(\mu, \nu) - 4^{\delta n} q_u(\mu, \nu)| \leq \frac{1}{(3 - \eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)}, \quad (45)$$

for every Pauli string  $\sigma_\mu$  with support size  $\ell$ .

The above condition was shown to hold in ref. [4] for random quantum circuits of size  $O(n \log^2 n)$ . Applying the equivalence established Section 6, it follows from Theorem 12 that this is fulfilled by a stochastic local Hamiltonians in time  $\tau_*' = O(\log^2 n)$ , when again we take the convention of ref. [9] and divide global scrambling time by the time to scramble a single subsystem.

Note finally that recently, an interesting connection between chaos – as being captured by *out-of-time-order correlation functions* – and pseudorandomness – as formalised in the notion of a unitary design – has been established [59]. Invoking those results, the findings presented here on generating approximate unitary designs by making use of time-fluctuating dynamics can be applied to assess quantum chaos in this sense.

**4.3. Applications in quantum information processing.** We finally mentioned a third, immediate, application, which seems yet particularly important when having potential technological applications in quantum information processing in mind. It should be clear that whenever the aim is to realize an approximate unitary design, the evolution under a fluctuating Hamiltonian already constitutes a valuable option. In many domains of quantum information, specifically in notions of benchmarking, approximate unitary designs are important primitives [19–21]. It is known that with a suitable random circuit one can generate an  $\varepsilon$ -approximate  $k$ -design [12]. Such a prescription, however, requires the precise implementation of a deep quantum circuit consisting of a large number of local quantum gates, namely  $O(nk^9(nk + \log 1/\varepsilon))$ . The above results have the interesting implication: instead of implementing a quantum circuit, a suitably stochastic Hamiltonian evolution gives rise to exactly the same dynamics. In such an approach, the classically fluctuating parameters would have to be stored. This insight might have important applications in quantum information processing.

## 5. Stochastic time evolution generates tensor product expanders

In this section, we prove Theorem 9 bounding the time after which the stochastic time evolution becomes a tensor product expander. As a crucial step we investigate the gap of the local generator induced by the Hamiltonian increments as given in eq. (15).

The proof will be structured as follows: we first derive in Lemma 15 the generator of the  $k$ -th moment operator and then describe how this allow us to express it as a tensor product expander using previous results on random quantum circuits. In Subsection 5.1 we provide the central mathematical results of this work, namely a diagonalisation of the local generator by relating it to the Casimir element in the enveloping algebra of  $\mathfrak{su}(d^2)$ . Since only certain irreps are contained in the direct sum decomposition of the Casimir element, we will observe that no eigenvalue can assume a value in the interval  $(0,1)$ , giving rise to a local gap.

Much of the developed machinery will build upon the representation theory of the special unitary group. It will also be helpful to use the identification of maps on matrices with matrices (induced vectorisation of matrices) given by  $\text{vec}(XYZ) = (X \otimes Z^T) \text{vec}(Y)$  to express the  $k$ -th moment operator as

$$M_\mu^k = \mathbb{E}_\mu[\pi_{k,k}(U)], \quad (46)$$

where  $\pi_{k,k}(U)$  is the  $(k, k)$ -mixed tensor representation of the group element  $U \in \mathbb{S}\mathbb{U}(N)$  given by

$$\pi_{k,k}(U) := U^{\otimes k} \otimes \overline{U}^{\otimes k}. \quad (47)$$

We also make use of the corresponding representation of the Lie algebra  $\mathfrak{su}(N)$  which is also denoted by  $\pi_{k,k}$  and satisfies the following for all  $X \in \mathfrak{su}(N)$

$$\pi_{k,k}(\exp(X)) = \exp(\pi_{k,k}(X)), \quad (48)$$

with

$$\pi_{k,k}(X) = \sum_{i=1}^k X \otimes \mathbb{1}_{\bar{i}} + \sum_{i=k+1}^{2k} \overline{X} \otimes \mathbb{1}_{\bar{i}} \quad (49)$$

and

$$X \otimes \mathbb{1}_{\bar{i}} := \mathbb{1}_1 \otimes \dots \otimes \mathbb{1}_{i-1} \otimes X \otimes \mathbb{1}_{i+1} \otimes \dots \otimes \mathbb{1}_{2k}. \quad (50)$$

This representation plays a central role in our analysis of the gap of the  $k$ -th moment operator of the stochastic time evolution.

The  $k$ -th moment operator  $M_{\text{SLH}(T)}^k$  has a generator that we explicitly calculate in the following. In fact, the lemma also holds for general Brownian motions on  $\mathbb{U}(N)$ , not only the locally generated ones considered in our theorems.

**Lemma 15 (The generator of the  $k$ -th moment operator).** *Let  $M_T^k$  be the  $k$ -th moment operator of a unitary Brownian motion with increments  $\Theta_{\Delta t}$  as in eq. (14) at time  $T$ . Then*

$$M_T^k = e^{G^k T} \quad (51)$$

with

$$G^k = \lim_{t \rightarrow 0} \left( \mathbb{E}[\overline{\pi_{k,k}(\Theta_t)}] + \frac{1}{2} \mathbb{E}[\pi_{k,k}(\Theta_t)^2 t] \right). \quad (52)$$

Note that as  $\Theta_t$  is anti-Hermitian,  $G^k$  is negative semidefinite. If the Brownian motion is universal then the kernel of  $G^k$  is the invariant subspace of  $M^k$ .

Most steps in the proof of this lemma will be also used again in the proof of Theorem 9.



*Proof.* As  $M_T^k$  is a Markov process, we have

$$M_T^k = (M_{\Delta t}^k)^{T/\Delta t}. \quad (53)$$

With the mixed tensor representation (47) and the definition of the  $k$ -th moment operator (4) we obtain for a single time step

$$M_{\Delta t}^k = \mathbb{E} [\pi_{k,k}(U_{\Delta t})] \quad (54)$$

$$= \mathbb{E} [\pi_{k,k}(\exp\{\Theta_{\Delta t} \Delta t\})] \quad (55)$$

with  $U_{\Delta t} = e^{\Theta_{\Delta t} \Delta t}$  and the increments  $\Theta_{\Delta t}$  from eq. (14). Using a Taylor expansion yields

$$M_{\Delta t}^k = \mathbb{E} \left[ e^{\pi_{k,k}(\Theta_{\Delta t}) \Delta t} \right] \quad (56)$$

$$= \sum_{p=0}^{\infty} \frac{(\Delta t)^p}{p!} \mathbb{E} [\pi_{k,k}(\Theta_{\Delta t})^p] \quad (57)$$

$$= \mathbb{1}_{\dim(\pi_{k,k})} + \mathbb{E} [\pi_{k,k}(\Theta_{\Delta t})] \Delta t + \mathbb{E} [\pi_{k,k}(\Theta_{\Delta t})^2] \frac{\Delta t^2}{2} + O(\Delta t^2). \quad (58)$$

Composing the time steps as in eq. (53), we obtain

$$M_T^k = \lim_{\Delta t \rightarrow 0} \left( \mathbb{1}_{\dim(\pi_{k,k})} + \left( \mathbb{E} [\pi_{k,k}(\Theta_{\Delta t})] + \mathbb{E} \left[ \frac{1}{2} \pi_{k,k}(\Theta_{\Delta t})^2 \Delta t \right] \right) \Delta t \right)^{T/\Delta t} \quad (59)$$

and finishes the proof.  $\square$

As a direct application of this lemma and as an exercise for the proof of Theorem 9, we prove Proposition 14 in the following.

*Proof (Proposition 14).* According to Lemma 15 the evolution has a generator

$$G^1 = \lim_{t \rightarrow 0} \mathbb{E} [\Theta_t \otimes \mathbb{1} + \mathbb{1} \otimes \bar{\Theta}_t] + \frac{1}{2} \lim_{t \rightarrow 0} \mathbb{E} [(\Theta_t \otimes \mathbb{1} + \mathbb{1} \otimes \bar{\Theta}_t)^2] t, \quad (60)$$

where the mixed tensor representation  $\pi_{1,1}$  from eq. (49) is used. It remains to show that the generator is of Lindblad form. The proposition's hypothesis yields

$$G^1 = -i H_0 \otimes \mathbb{1} + i \mathbb{1} \otimes \bar{H}_0 + \frac{1}{2} \lim_{t \rightarrow 0} \mathbb{E} [(F_t^2 \otimes \mathbb{1} + \mathbb{1} \otimes \bar{F}_t^2 + 2F_t \otimes \bar{F}_t)] t \quad (61)$$

$$= -i H_0 \otimes \mathbb{1} + i \mathbb{1} \otimes H_0^T - \frac{a}{2} \sum_{\mu} (B_{\mu}^2 \otimes \mathbb{1} + \mathbb{1} \otimes B_{\mu}^{2T} - 2B_{\mu} \otimes B_{\mu}^T), \quad (62)$$

where we have used that  $\bar{B}_{\mu} = -B_{\mu}^T$ . The identification  $\text{vec}(XYZ) = (X \otimes Z^T) \text{vec}(Y)$  and  $B_{\mu}^{\dagger} = -B_{\mu}$  finish the proof.  $\square$

We note the following theorem on random quantum circuits generated by general local distributions, which is implicitly contained in ref. [12, Corollary 7].

**Lemma 16 (Relating global and local gaps).** *Let  $\mu_{loc}$  be a distribution on  $\mathbb{U}(d^2)$  and  $\mu$  be the distribution on  $\mathbb{U}(d^n)$  that applies a unitary drawn according to  $\mu_{loc}$  to a uniformly chosen edge of an interaction graph  $i, i+1$ . Then, its moment operator satisfies*

$$\|M_{\mu}^k - M_{\text{Haar}}^k\|_{\infty} \leq 1 - (1 - \|m_{\mu_{loc}}^k - m_{\text{Haar}_{loc}}^k\|_{\infty}) (1 - \|M_{RQC, \text{Haar}}^k - M_{\text{Haar}}^k\|_{\infty}), \quad (63)$$

where  $\text{Haar}_{loc}$  denotes the Haar measure on  $\mathbb{U}(d^2)$  and  $M_{RQC, \text{Haar}}^k$  the moment operator of the unitary circuit with distribution  $\mu = \text{Haar}_{loc}$ .

*Proof.* Setting  $(M_{\text{Haar}}^k)^\perp = \mathbb{1} - M_{\text{Haar}}^k$  we have the relation

$$\|M_\mu^k - M_{\text{Haar}}^k\|_\infty = \|(M_{\text{Haar}}^k)^\perp M_\mu^k (M_{\text{Haar}}^k)^\perp\|_\infty \quad (64)$$

$$= \left\| \frac{1}{|E|} \sum_{e \in E} (M_{\text{Haar}}^k)^\perp m_{\mu_{loc}}^{k,(e)} (M_{\text{Haar}}^k)^\perp \right\|_\infty. \quad (65)$$

By denoting  $\gamma := \|m_{\mu_{loc}}^k - m_{\text{Haar}_{loc}}^k\|_\infty$ , we find  $m_{\mu_{loc}}^{k,(e)} \leq (1 - \gamma) m_{\text{Haar}_{loc}}^{k,(e)} + \gamma \mathbb{1}$ , which implies the operator inequality

$$(M_{\text{Haar}}^k)^\perp \left( \sum_{e \in E} m_{\mu_{loc}}^{k,(e)} \right) (M_{\text{Haar}}^k)^\perp \leq \gamma |E| (M_{\text{Haar}}^k)^\perp \quad (66)$$

$$+ (1 - \gamma) (M_{\text{Haar}}^k)^\perp \left( \sum_{e \in E} m_{\mu_{\text{Haar}_{loc}}}^{k,(e)} \right) (M_{\text{Haar}}^k)^\perp. \quad (67)$$

Since  $(1 - \gamma)$  is positive, we can use the bound  $A \leq \|A\|_\infty$  for the second summand on the right hand side which, together with

$$\|M_{\text{RQC,Haar}}^k - M_{\text{Haar}}^k\|_\infty = \|(M_{\text{Haar}}^k)^\perp M_{\text{RQC,Haar}}^k (M_{\text{Haar}}^k)^\perp\|_\infty \quad (68)$$

finishes the proof.  $\square$

Now we present the main proof of Theorem 9. Part of it will be completed with the lemmas stated and proved subsequently.

*Proof (Proof of Theorem 9).* Thanks to Lemma 5 it is enough to bound the gap of the  $k$ -th moment operator  $M_{\text{SLH}(T)}^k$ . According to Lemma 21, the time constant part of the Hamiltonian does not affect the invariant subspace nor the gap of  $M_{\text{SLH}(T)}^k$ . Hence, we can set without loss of generality  $h_0^{(e)} = 0 \forall e$ . Additionally, in Lemma 22 we prove that the gap of an interaction graph being a complete graph is larger than the one of a 1D graph. We hence consider only the latter case in the proof.

Using the approximation (58) and expressing  $\Theta$  in terms of the local terms  $\theta^{(e)}$  (as in eq. (16)) we obtain

$$M_{\text{SLH}(\Delta t)}^k = \mathbb{1}_{\dim(\pi_{k,k})} + \sum_{e \in E} \mathbb{E} \left[ \pi_{k,k}(\theta_{\Delta t}^{(e)})^2 \right] \frac{\Delta t^2}{2} + O(\Delta t^2) \quad (69)$$

Using another Taylor approximation yields

$$M_{\text{SLH}(\Delta t)}^k = \mathbb{1}_{\dim(\pi_{k,k})} + \frac{1}{n} \sum_{e \in E} \mathbb{E} \left[ \pi_{k,k}(\sqrt{n} \theta_{\Delta t}^{(e)})^2 \right] \frac{\Delta t^2}{2} + O(\Delta t^2) \quad (70)$$

$$= \frac{1}{n} \sum_{e \in E} \mathbb{E} \left[ \left( \exp\{\sqrt{n} \theta_{\Delta t}^{(e)} \Delta t\} \right)^{\otimes k,k} \right] + O(\Delta t^2). \quad (71)$$

Next, we view  $G(\theta_{\Delta t}^{(e)}) := \exp\{\sqrt{n} \theta_{\Delta t}^{(e)} \Delta t\}$  as a random gate in a  $G$ -random quantum circuit considered in ref. [12]. The (system size independent) local  $k$ -th moment operator on edge  $e \in E$  is

$$m_{\Delta t}^{k,(e)} := \mathbb{E} \left[ \left( \exp\{\sqrt{n} \theta_{\Delta t}^{(e)} \Delta t\} \right)^{\otimes k,k} \right]. \quad (72)$$

Note that this  $k$ -th moment operator also corresponds to a Brownian motion but with a variance rescaled by a factor of  $n$ , cf. also the parameter  $a$  in eq. (19). As its gap, i.e., the difference between the largest

and second largest eigenvalue does not depend on  $e$  we simply denote the gap of  $m_{\Delta t}^{k,(e)}$  by  $\Delta(m_{\Delta t}^k)$ . Then the local gap lemma 16 yields directly

$$\left\| M_{\text{SLH}(\Delta t)}^k - M_{\text{Haar}}^k \right\|_{\infty} \leq 1 - \Delta(m_{\Delta t}^k) (1 - \|M_{\text{RQC,Haar}}^k - M_{\text{Haar}}^k\|_{\infty}), \quad (73)$$

with  $M_{\text{RQC,Haar}}^k$  being the  $k$ -th moment operator of single step of a local random quantum circuit whose gates are chosen from the Haar measure. The gap of  $M_{\text{RQC,Haar}}^k$  can be lower bounded as [12, eq. (41)]

$$(1 - \|M_{\text{RQC,Haar}}^k - M_{\text{Haar}}^k\|_{\infty}) \geq \frac{k^{-\frac{2.5}{\ln(d)} - 2.5 \frac{\ln(d^2+1)}{\ln(d)}}}{125 [\log_d(4k)]^2 e(d^2+1)} \quad (74)$$

$$\geq \frac{1}{425 n [\log_d(4t)]^2 d^2 k^5 k^{3.1/\ln(d)}}, \quad (75)$$

where we have also used the inequalities right after [12, eq. (41)]. Together with eq. (73), these results imply

$$\|M_{\text{SLH}(\Delta t)}^k - M_{\text{Haar}}^k\|_{\infty} \leq 1 - \Delta(m_{\Delta t}^k) s/n \quad (76)$$

with

$$s := \left( 425 [\log_d(4k)]^2 d^2 k^5 k^{3.1/\ln(d)} \right)^{-1}. \quad (77)$$

In order to calculate  $\Delta(m_{\Delta t}^k)$  we use Lemma 15, eq. (72) and

$$\mathbb{E} \left[ \pi_{k,k}(\theta_t^{(e)}) \right] = 0 \quad (78)$$

so that we can express  $m_{\Delta t}^{k,(e)}$  as

$$m_{\Delta t}^{k,(e)} = \exp \left( n g_k^{(e)} \Delta t \right) \quad (79)$$

$$= \mathbf{1} + n g_k^{(e)} \Delta t + O(\Delta t^2) \quad (80)$$

with

$$\begin{aligned} n g_k^{(e)} &= \frac{1}{2} \lim_{t \rightarrow 0} \mathbb{E} \left[ \pi_{k,k}(\sqrt{n} \theta_t^{(e)})^2 t \right] \\ &= \frac{n}{2} \lim_{t \rightarrow 0} \mathbb{E} \left[ \pi_{k,k}(\theta_t^{(e)})^2 t \right]. \end{aligned} \quad (81)$$

Hence,

$$\Delta(m_{\Delta t}^k) = n \Delta(g_k) \Delta t + O(\Delta t^2), \quad (82)$$

where  $\Delta(g_k)$  denotes again the spectral gap to the invariant subspace, i.e., minus the largest non-zero eigenvalue of  $g_k^{(e)}$ .

As both  $k$ -th moment operators have the same unit eigenvalue eigenspace according to Lemma 5,  $\|M_{\text{SLH}(\Delta t)}^k - M_{\text{Haar}}^k\|_{\infty}$  is the second largest eigenvalue of  $M_{\text{SLH}(\Delta t)}^k$ . Hence,

$$\|M_{\text{SLH}(T)}^k - M_{\text{Haar}}^k\|_{\infty} = \lim_{\Delta t \rightarrow 0} \left\| (M_{\text{SLH}(\Delta t)}^k)^{T/\Delta t} - M_{\text{Haar}}^k \right\|_{\infty} \quad (83)$$

$$= \lim_{\Delta t \rightarrow 0} (1 - s \Delta(g_k))^{T/\Delta t} \quad (84)$$

$$= \exp(-T s \Delta(g_k)). \quad (85)$$

Observation 18 and Lemma 19 yields that the gap is the same as the variance (19) of the noise,  $\Delta(g_k) = a/2$ , which completes the proof.  $\square$

5.1. *Local gap.* In order to calculate the local gap  $\Delta(g_k)$ , the following representations for the algebra  $\mathfrak{su}(N)$  will be used.

$$\text{Trivial rep. } \pi_1 : \mathfrak{su}(N) \rightarrow \mathfrak{gl}(1, \mathbb{C}), \quad X \mapsto 0, \quad (86)$$

$$\text{Fundamental rep. } \pi_f : \mathfrak{su}(N) \rightarrow \mathfrak{gl}(N, \mathbb{C}), \quad X \mapsto X, \quad (87)$$

$$\text{Adjoint rep. } \pi_{\text{ad}} : \mathfrak{su}(N) \rightarrow \mathfrak{gl}(\mathfrak{su}(N)), \quad X \mapsto \text{ad}_X, \quad (88)$$

where  $\text{ad}_X$  is defined by  $\text{ad}_X(Y) := [X, Y]$ .

**Observation 17 (Omitting the phase)** *From the mixed-tensor representation we note that we can restrict the analysis on the  $\mathfrak{su}(N)$  algebra instead of  $\mathfrak{u}(N)$ : the phase factor in the semi-direct product decomposition of any  $U \in \mathbb{U}(N) \simeq \mathbb{S}\mathbb{U}(N) \rtimes \mathbb{U}(1)$  is cancelled by its complex conjugate coming from  $\bar{U}$ . In this representation the two algebras are indistinguishable.*

The Killing form  $K$  in  $\mathfrak{su}(N)$  is the symmetric bilinear form defined by

$$K(X, Y) := \text{Tr}[\text{ad}_X \text{ad}_Y]. \quad (89)$$

Denoting the Hilbert-Schmidt inner product of  $X$  and  $Y$  (in the fundamental representation) by  $\langle X, Y \rangle = \text{Tr}(X^\dagger Y)$ , the Killing form of  $\mathfrak{su}(N)$  can also be written as

$$K(X, Y) = -2N \langle X, Y \rangle. \quad (90)$$

In terms of a basis  $\{X_\mu\}_{\mu=1}^{N^2-1}$  of  $\mathfrak{su}(N)$  the Killing metric tensor  $\kappa$  is defined by

$$\kappa_{\mu, \nu} := K(X_\mu, X_\nu), \quad (91)$$

as was already indicated in eq. (20). Then, the Casimir element in a matrix representation  $\pi$  is

$$C(\pi) := \sum_{\mu, \nu} \kappa_{\mu, \nu}^{-1} \pi(X_\mu) \pi(X_\nu). \quad (92)$$

According to eqs. (81) and (16), the local generator  $g_k^{(e)}$  of our unitary process with vanishing driving  $h_0^{(e)} = 0$  is given by

$$g_k^{(e)} = \frac{1}{2} \lim_{\Delta t \rightarrow 0} \mathbb{E} \left[ \pi_{k, k} \left( \sum_{\mu=1}^{N^2-1} A_\mu^{(e)} \xi^{(e, \mu)} \right)^2 \right] \Delta t = -\frac{a}{2} \sum_{\mu, \nu=1}^{N^2-1} \kappa_{\mu, \nu}^{-1} \pi_{k, k} \left( A_\mu^{(e)} \right) \pi_{k, k} \left( A_\nu^{(e)} \right) \quad (93)$$

(where  $N := d^2$ ). The second equality follows from our central assumption (19). All  $g_k^{(e)}$  are tensor copies of a local operator  $g_k$ . Therefore, we will suppress the subscripts  $e$  in this section from now on.

**Observation 18 (Casimir element)** *Let  $g_k$  be the generator of the local  $k$ -th moment operator in eq. (93). Then*

$$g_k = -\frac{a}{2} C(\pi_{k, k}). \quad (94)$$

More generally, an overcomplete set  $\{A_\mu\}$  can also be admitted, as already mentioned in Remark 8. The final result about the convergence rate – up to a constant  $O(1)$  – is still valid as long as the generator and the Casimir element are related by an equation of the form

$$g_k = -\frac{a'}{2} C(\pi_{k, k}) + g', \quad (95)$$

where  $a' > 0$  and  $g'$  is negative semidefinite so that it can only increase the gap.

In the following, we prove that the eigenvalues of the Casimir do not assume a value within the interval  $(0, 1)$ , for all  $k$ .

**Lemma 19 (Casimir gap).** *Let  $\mathcal{I}_k$  be the set of irreducible representations occurring in  $\pi_{k,k}$  and let  $m_k(\pi) \in \mathbb{N}$  denote the multiplicity of each such representation  $\pi$ . Then*

$$C(\pi_{k,k}) \simeq \bigoplus_{\pi \in \mathcal{I}_k} c_2(\pi) \mathbb{1}_{\dim(\pi_{k,k})} \otimes \mathbb{1}_{m_k(\pi)}, \quad (96)$$

where

$$c_2(\pi) \begin{cases} = 0 & \text{if } \pi \simeq \pi_1, \\ = 1 & \text{if } \pi \simeq \pi_{\text{ad}}, \\ > 1 & \text{otherwise.} \end{cases} \quad (97)$$

In particular, the spectral gap of  $C(\pi_{k,k})$  is independent of  $k$ .

*Proof.* Since the Casimir element is an element of the center of the universal enveloping algebra, from Schur's Lemma follows that it acts as a multiple of the identity in each irreducible representation (see ref. [60, Chapter 12]), so that (96) is immediate. Now, since the tensor product between the fundamental representation and its conjugate are isomorphic to the direct sum of the trivial and the adjoint ones, this means that the representation  $\pi_{k,k}$  is isomorphic to  $(\pi_1 \oplus \pi_{\text{ad}})^{\otimes k}$ .

The trivial representation is guaranteed to occur in the decomposition of  $\pi_{k,k}$  into irreducible representations (for example, via  $\pi_1^{\otimes k}$ ) and leads to the eigenvalue  $c_2(\pi_1) = 0$ . The adjoint representation always occurs – for example, via  $\pi_{\text{ad}} \otimes \pi_1^{\otimes(k-1)}$  and permutations thereof – too, and leads to the eigenvalue  $c_2(\pi_{\text{ad}}) = 1$ . If we can show that no other irreducible representation  $\pi$  with  $c_2(\pi) \leq 1$  occurs, the proof is complete.

One might think that this requires rather detailed knowledge about how tensor product representations of the form  $\pi_{\text{ad}}^{\otimes l}$  decompose into irreducible representations. To follow the next argument, some basic knowledge regarding Young diagrams is necessary; please refer to Appendix A. It is in fact sufficient to exploit a remarkably basic property which is shared by all the irreducible  $\mathfrak{su}(d^2)$  representations occurring in  $\pi_{k,k}$ : their Young diagrams must have a number of boxes which is divisible by  $d^2$ . This can be seen for instance by induction:  $(\pi_1$  and  $\pi_{\text{ad}})$  are two representations made of 0 and  $d^2$  boxes respectively. Now consider a representation  $\pi$  whose number of boxes is divisible by  $d^2$ ;  $\pi \otimes (\pi_1 \oplus \pi_{\text{ad}})$  is again a direct sum of representation divisible by  $d^2$ , since tensoring with the trivial one does nothing and tensoring with the adjoint adds  $d^2$  boxes to the Young diagram of  $\pi$ . According to Young calculus only  $d^2$  boxes can be cancelled at once. Hence, if the statement is true for  $(\pi_1 \oplus \pi_{\text{ad}})^{\otimes(k-1)}$ , then it holds for  $(\pi_1 \oplus \pi_{\text{ad}})^{\otimes k}$ .

Indeed, all such representations  $\pi$  other than the trivial and the adjoint one satisfy  $c_2(\pi) > 1$  as we will show in Lemma 20 below.  $\square$

Let  $\lambda := (\lambda_1, \dots, \lambda_{N-1})$  with  $\lambda_i \in \mathbb{N}_0$  denote the Dynkin label of an irreducible representation  $\pi$  of  $\mathfrak{su}(N)$ . The eigenvalue of the Casimir element in the irreducible representation  $\pi$  is

$$c_2(\pi) = \frac{1}{2N} \sum_{i,j=1}^{N-1} (\lambda_i + 2)(A^{-1})_{i,j} \lambda_j, \quad (98)$$

where  $A$  is the Cartan matrix of  $\mathfrak{su}(N)$  [61, §21.3]. The inverse Cartan matrix is directly given by

$$(A^{-1})_{i,j} = \frac{1}{N} \begin{cases} i(N-j) & \text{if } i \leq j, \\ j(N-i) & \text{if } i > j, \end{cases} \quad (99)$$

and is symmetric. We now show the following lemma.

**Lemma 20 (Young diagrams).** *Let  $N > 2$  and  $\pi$  be an irreducible representation of  $\mathfrak{su}(N)$  such that the number of boxes in its Young diagram is divisible by  $N$ . If  $\pi$  is not isomorphic to the trivial or adjoint representation, then  $c_2(\pi) > 1$ .*

*Proof.* First observe that we can immediately rule out all irreducible representations whose Young diagrams consist of a single column because the maximal column height for  $\mathfrak{su}(N)$  is  $N - 1$  (i.e. Dynkin labels having a single entry 1 and 0 everywhere else). In the following we will analyse the growth behaviour of the quadratic form (98) as we move from one irreducible representation (i.e. Dynkin label) to the next one.

It will turn out very helpful to know the column sums of the inverse Cartan matrix  $A^{-1}$ . Clearly, the sum of the first (or equally the last) column is  $(N - 1)/2$ . The sum of any other column is strictly greater than this value. Indeed, pick a column  $j$  and denote its sum by  $a_j$ . One can easily convince oneself that  $a_j = j(N - j)/2$ .

Now we compare the quadratic Casimir eigenvalues of different irreducible representations, i.e. Dynkin labels  $\lambda$ . As it turns out, adding 1 to any component of any Dynkin label  $\lambda$  always increases this eigenvalue at least by almost  $1/2$ ,

$$c_2(\lambda + e_i) - c_2(\lambda) \geq \frac{N^2 - 1}{2N^2} =: \Delta_N. \quad (100)$$

Here  $e_i$  is the  $i$ -th canonical basis vector of  $\mathbb{R}^{N-1}$ . So, starting from the trivial representation with  $c_2(0, 0, \dots, 0) = 0$  we immediately obtain the crude lower bound

$$c_2(\lambda_1, \dots, \lambda_{N-1}) \geq \Delta_N \sum_{i=1}^{N-1} \lambda_i = \Delta_N \|\lambda\|_1. \quad (101)$$

Observe that  $2\Delta_N < 1 < 3\Delta_N$ . Thus we are guaranteed to obtain a quadratic Casimir eigenvalue strictly greater than 1 whenever we add at least three arbitrary columns to the (empty!) Young diagram of the trivial representation.

This leaves us with those irreducible representations whose Young diagrams have exactly two columns, i.e. with the Dynkin labels  $(1, 1, 0, \dots, 0)$ ,  $(2, 0, 0, \dots, 0)$  and all permutations thereof. As is well known (and can be checked easily with the explicit formula below) the quadratic Casimir eigenvalue of the adjoint representation  $(1, 0, \dots, 0, 1)$  is exactly 1. We would like to show that any other placement of the two ones yields a strictly greater eigenvalue. Suppose these occur in positions  $1 \leq \alpha < \beta < N$ . Then,

$$\begin{aligned} c_2(\lambda) &= \frac{1}{2N} ((A^{-1})_{\alpha,\alpha} + 2(A^{-1})_{\alpha,\beta} + (A^{-1})_{\beta,\beta} + 2a_\alpha + 2a_\beta) \\ &\geq \frac{1}{2N} ((A^{-1})_{1,1} + 2(A^{-1})_{1,N-1} + (A^{-1})_{N-1,N-1} + 2a_1 + 2a_{N-1}) \\ &= c_2(1, 0, \dots, 0, 1) \\ &= 1. \end{aligned} \quad (102)$$

It is easy to see that this inequality turns into a strict one if either of the two ones is *not* at the first or last position. Finally consider a Dynkin label  $\lambda$  with a single non-vanishing component  $\lambda_\alpha = 2$  at position  $\alpha$  (i.e., a Young diagram with exactly two columns of height  $\alpha$ ),

$$c_2(\lambda) = \frac{2}{N} ((A^{-1})_{\alpha,\alpha} + a_\alpha) = \frac{N+2}{N^2} \alpha(N - \alpha). \quad (103)$$

From the global minimum of the quadratic function  $\alpha(N - \alpha)$  we easily obtain the lower bound

$$c_2(\lambda) \geq 1 + \frac{N-2}{N^2} \quad (104)$$

and thus  $c_2(\lambda) > 1$  for all  $N > 2$  as claimed.  $\square$

5.2. *Hamiltonian driving.* We now show that a time constant part in a stochastic Hamiltonian cannot affect the gap of the  $k$ -th moment operator.

**Lemma 21 (Hamiltonian driving).** *Let  $M_T^k$  be the  $k$ -th moment operator (4) of a universal Brownian motion with increments  $\Theta_{\Delta t}$  as in eq. (14). Write  $\Theta_{\Delta t}$  as*

$$\Theta_{\Delta t} = -iH_0 + F_{\Delta t}, \quad (105)$$

where  $-iH_0$  and  $F_{\Delta t}$  are its anti-Hermitian time constant and fluctuating parts, respectively, with

$$F_{\Delta t} = \sum_{\mu} B_{\mu} \xi_{\Delta t}^{\mu}, \quad B_{\mu}^{\dagger} = -B_{\mu}, \quad \mathbb{E}[\xi_{\Delta t}^{\mu}] = 0, \quad \text{and} \quad \mathbb{E}[\xi_{\Delta t}^{\mu} \xi_{\Delta t}^{\nu}] = -\frac{a}{\Delta t} \delta_{\mu, \nu}.$$

Let  $\tilde{M}_T^k$  be defined similarly but without driving, i.e., with  $H_0 = 0$ . Then  $\tilde{M}_T^k$  and  $M_T^k$  have the same gap, i.e.,

$$\|\tilde{M}_T^k - M_{\text{Haar}}^k\|_{\infty} = \|M_T^k - M_{\text{Haar}}^k\|_{\infty}. \quad (106)$$

*Proof.* Lemma 5 implies that the gap of  $M_{\Delta t}^k$  is  $\|M_{\Delta t}^k - M_{\text{Haar}}^k\|_{\infty}$ . Hence,

$$\|M_T^k - M_{\text{Haar}}^k\|_{\infty} = \lim_{\Delta t \rightarrow 0} \|M_{\Delta t}^k - M_{\text{Haar}}^k\|_{\infty}^{T/\Delta t} \quad (107)$$

is the gap of  $M_T^k$  and, similarly, for  $\tilde{M}_T^k$ .

Using the connection between Brownian motion and its increments (12) and a Trotter-Suzuki approximation we obtain

$$\begin{aligned} M_{\Delta t}^k &= \mathbb{E}[\exp\{\pi_{k,k}(-iH_0 + F_{\Delta t}) \Delta t\}] + O(\Delta t^2) \\ &= \mathbb{E}[\exp\{\pi_{k,k}(F_{\Delta t}) \Delta t\}] \exp\{\pi_{k,k}(-iH_0) \Delta t\} + O(\Delta t^2) \\ &= \tilde{M}_{\Delta t}^k \exp\{\pi_{k,k}(-iH_0) \Delta t\} + O(\Delta t^2). \end{aligned} \quad (108)$$

As  $\exp\{\pi_{k,k}(-iH_0) \Delta t\}$  is a fixed unitary, up to an error of order  $O(\Delta t^2)$ , the gap of  $M_{\Delta t}^k$  and  $\tilde{M}_{\Delta t}^k$  are the same. This finishes the proof.  $\square$

5.3. *More general interaction graphs.* The generator from Lemma 15 of the  $k$ -th moment operator of the unitary Brownian motion inherits the locality structure from the increments (15). Hence, it can be written as

$$G^k = \sum_{e \in E} g_k^{(e)}, \quad (109)$$

where  $G^k$  is the generator associated to  $\Theta_{\Delta t}$  and  $g_k^{(e)}$  to  $\theta_{\Delta t}^{(e)}$  according to eq. (52). Presumably, among all connected graphs, the gap of  $G^k$  could have a minimum for 1D nearest neighbour graphs. Here, we show that adding edges to this graph can only increase the gap, which can only lead to a faster mixing in Theorem 9.

In the following lemma, the spectral gap  $\Delta(G)$  of an operator  $G$  is the difference of the second smallest and smallest singular value.

**Lemma 22 (The spectral gap of the generator is concave).** *Let  $(G_i)_i$  be a finite set of negative semidefinite and Hermitian operators with common non-trivial kernel and  $p$  be a probability vector. Then*

$$\Delta\left(\sum_i p_i G_i\right) \geq \sum_i p_i \Delta(G_i). \quad (110)$$

This lemma implies that the gap of the generator (109) can only become smaller when one removes edges from  $E$ , while keeping  $E$  connected. Hence, the gap in the case of a one dimensional graph can also only be smaller as the gap in case of a complete graph.

*Proof.* Let  $K$  denote the common kernel of  $(G_i)_i$ . Then it is also the kernel of any operator in the convex hull of  $(G_i)_i$ . The gap of  $G_i$  is the smallest singular value of  $G_i$  restricted to the orthogonal complement of  $K$  and similarly for  $G := \sum_i p_i G_i$ . Hence, it is enough to show that the smallest singular value as the function

$$G \mapsto \min_{\langle x|x \rangle=1} |\langle x|G|x \rangle| \quad (111)$$

is concave. But this follows from the smallest singular value being the minimum of the linear functions  $G \mapsto \langle x|G|x \rangle$ .  $\square$

*Remark 23 (Frustration free Hamiltonians).* The same argument applies when the operators are all positive semidefinite. Hence, the gap of frustration free Hamiltonians, as considered in ref. [12], is also a concave function, i.e., can only increase under taking convex combinations.

*5.4. Example: White noise in the Pauli basis.* We conclude the discussion on approximate unitary designs with an example involving the specific setting in eqs. (26) and (27), and see that the choice of the Pauli matrices as a basis precisely matches, under the representation theoretic approach, the assumption on the covariance for the variables  $\xi$ .

Consider  $n = 2$  qubits (thus  $N = 4$ ) and the Hamiltonian increments

$$\Theta_{\Delta t} := -i \sum_{\alpha, \beta=0}^3 \sigma_\alpha \otimes \sigma_\beta \xi_{\Delta t}^{(\alpha, \beta)}, \quad (112)$$

where  $\xi_{\Delta t}^{(\alpha, \beta)}$  are i.i.d. real random variables with zero mean and covariance

$$\text{cov}[\xi_{\Delta t}^{(\alpha, \beta)} \xi_{\Delta t}^{(\alpha', \beta')}] = \delta_{\alpha, \alpha'} \delta_{\beta, \beta'} \frac{1}{\Delta t}, \quad \forall \alpha, \beta. \quad (113)$$

Leaving out the term  $\sigma_0 \otimes \sigma_0 \xi_{\Delta t}^{(0,0)}$  we can easily restrict  $\Theta_{\Delta t}$  to its traceless part

$$\Theta_{0, \Delta t} = \sum_{\mu=1}^{15} \tau_\mu \xi_{\Delta t}^\mu, \quad (114)$$

where we defined the anti-Hermitian operators  $\tau_\mu := -i \sigma_{\mu_1} \otimes \sigma_{\mu_2}$  so that  $\{\tau_1, \tau_2, \dots, \tau_{15}\} = \{\tau_{(0,1)}, \tau_{(0,2)}, \dots, \tau_{(3,3)}\}$  form a basis of the fundamental representation of  $\mathfrak{su}(4)$ . From eq. (20) we compute the Killing metric tensor (91) with respect to this basis as

$$\kappa_{\mu, \nu} = -8 \text{Tr}(\tau_\mu^\dagger \tau_\nu) = -32 \delta_{\mu, \nu}. \quad (115)$$

From eq. (19) and the assumption in eq. (113) immediately follows  $a = 32$ . Observation 18 tells us then  $g_2 = -16 C(\pi_{2,2}) =$  and hence the second moment operator  $M_{\text{SLH}(\Delta t)}^{k=2}$  has a gap of  $16\Delta t$ , matching eq. (134) in the decoupling section.



## 6. Decoupling with stochastic Hamiltonian time evolution

The section is devoted to the proof Theorem 12. To show our result, we consider a fluctuating Hamiltonian on a complete graph whose increments are given in eq. (26), in the limit of  $\Delta t \rightarrow 0$ . As already mentioned, this result implies Theorem 13 by application of the same proof technique used for the random quantum circuit case in ref. [4].

First, we analyse how the support size of an initial Pauli string evolves during the process, then we observe how the qubits are made invariant under relabelling of the Pauli elements; this, together with the permutation invariance condition, leads to the desired result. Decoupling of an arbitrary  $n$ -qubit system  $A$  is mainly described by the second moment operator induced by the evolution. The expansion coefficients in the Pauli basis are given in eq. (32). We recall that, since the Brownian motion on  $\mathbb{U}(2^n)$  is Markovian, the second moment operator at time  $T$  on  $X$  is given by concatenating  $T/\Delta t$  times the operator  $M_{n,\text{SLH}(\Delta t)}^{k=2}$ , i.e.,

$$M_{n,\text{SLH}(T)}^{k=2}(X) = \lim_{\Delta t \rightarrow 0} \underbrace{M_{n,\text{SLH}(\Delta t)}^{k=2} \circ \cdots \circ M_{n,\text{SLH}(\Delta t)}^{k=2}}_{T/\Delta t \text{ times}}(X) \quad (116)$$

$$=: \lim_{\Delta t \rightarrow 0} \bigcirc_{s=1}^{T/\Delta t} M_{n,\text{SLH}(\Delta t)}^{k=2}(X). \quad (117)$$

Note that, since the Hamiltonian in eq. (26) generating Brownian motion is dependent on system size, we must include an additional subscript.

In Taylor approximation, up to an error  $O(\Delta t^2)$ ,  $M_{n,\text{SLH}(\Delta t)}^{k=2}$  results from the sum of two-qubit moment operators acting on any possible qubit pair  $j, k$ , i.e.

$$M_{n,\text{SLH}(\Delta t)}^{k=2} = \frac{2}{n(n-1)} \sum_{j < k} \left( M_{2,\text{SLH}(\Delta t)}^{k=2} \right)^{j,k} + O(\Delta t^2). \quad (118)$$

This can be seen through calculations analogous to the ones from eqs. (69)-(71). We can hence interpret this process as a qubit pair being uniformly randomly chosen at every time step  $(\ell-1)\Delta t$  and a two-qubit unitary  $U_{2,\ell,\Delta t} := \exp\{-i H_{2,\ell,\Delta t} \Delta t\}$  being applied. Therefore, in the following section we first consider the restricted two-qubit case, which provides useful results and insights to be used for the investigation of the general case with  $n$  qubits.

*6.1. Two-qubit analysis of the second moment operator.* Considering a two-qubit system, here we would like to understand the evolution of  $M_{2,\text{SLH}(T)}^{k=2}$  through  $M_{2,\text{SLH}(\Delta t)}^{k=2}$  and show the following lemma, which is compatible to the analysis of the local gap discussed in the previous section (as showed in Example 5.4).

**Lemma 24 (Two-qubit case).** *Then the local second moment operator associated to the Hamiltonian increments (26) converges exponentially to the second moment operator of the uniform distribution, i.e.*

$$\|M_{2,\text{SLH}(T)}^{k=2} - M_{2,\text{Haar}}^{k=2}\|_{\infty} \leq e^{-16T}. \quad (119)$$

*Proof.* To prove the convergence rate, we want to express  $M_{n,\text{SLH}(\Delta t)}^{k=2}$  in terms of the Pauli basis and compute the gap. We can see directly that the identity on 4 qubits is an eigenvector with unit eigenvalue

$$M_{2,\text{SLH}(\Delta t)}^{k=2}(\mathbb{1}_4) = \mathbb{E} \left[ U_{2,\ell,\Delta t}^{\otimes 2} \mathbb{1}_4 (U_{2,\ell,\Delta t}^{\dagger})^{\otimes 2} \right] = \mathbb{1}_4. \quad (120)$$

We then observe the unitary evolution acting on a Pauli element  $\sigma_{\mu} \otimes \sigma_{\nu}$ , with  $\mu, \nu \in \{0, 1, 2, 3\}^2$  and calculate its expectation with a Taylor expansion for the unitary, taking into account terms with leading

order in  $\Delta t$  (and omitting subscripts for  $H$ ),

$$\begin{aligned} M_{2,\text{SLH}(\Delta t)}^{k=2}(\sigma_\mu \otimes \sigma_\nu) &= \mathbb{E} \left[ U_{2,\ell,\Delta t}(\sigma_{\mu_1} \otimes \sigma_{\mu_2}) U_{2,\ell,\Delta t}^\dagger \otimes U_{2,\ell,\Delta t}(\sigma_{\nu_1} \otimes \sigma_{\nu_2}) U_{2,\ell,\Delta t}^\dagger \right] \\ &= \mathbb{E} \left[ \left( \mathbb{1}_2 - iH\Delta t - \frac{1}{2}H^2\Delta t^2 \right) (\sigma_{\mu_1} \otimes \sigma_{\mu_2}) \left( \mathbb{1}_2 + iH\Delta t - \frac{1}{2}H^2\Delta t^2 \right) \right. \\ &\quad \left. \otimes \left( \mathbb{1}_2 - iH\Delta t - \frac{1}{2}H^2\Delta t^2 \right) (\sigma_{\nu_1} \otimes \sigma_{\nu_2}) \left( \mathbb{1}_2 + iH\Delta t - \frac{1}{2}H^2\Delta t^2 \right) \right] \\ &\quad + O(\Delta t^2). \end{aligned} \quad (121)$$

We now recall that the  $\xi$  white noise variables are i.i.d. with zero mean and covariance as in eq. (27). Considering only the non-vanishing linear terms in  $\Delta t$  in the expectation, we have

$$\begin{aligned} M_{2,\text{SLH}(\Delta t)}^{k=2}(\sigma_\mu \otimes \sigma_\nu) &= \sigma_\mu \otimes \sigma_\nu + \Delta t^2 \mathbb{E} [H\sigma_\mu H \otimes \sigma_\nu + \sigma_\mu \otimes H\sigma_\nu H] \\ &\quad - \frac{\Delta t^2}{2} \mathbb{E} [H^2\sigma_\mu \otimes \sigma_\nu + \sigma_\mu H^2 \otimes \sigma_\nu + \sigma_\mu \otimes H^2\sigma_\nu + \sigma_\mu \otimes \sigma_\nu H^2] \\ &\quad - \Delta t^2 \mathbb{E} [[H, \sigma_\mu] \otimes [H, \sigma_\nu]] + O(\Delta t^2). \end{aligned} \quad (122)$$

Let us consider the second term, in particular

$$\mathbb{E} [H\sigma_\mu H \otimes \sigma_\nu] = \frac{1}{\Delta t} \left( \sum_{\alpha,\beta} (\sigma_\alpha \otimes \sigma_\beta) (\sigma_{\mu_1} \otimes \sigma_{\mu_2}) (\sigma_\alpha \otimes \sigma_\beta) \right) \otimes (\sigma_{\nu_1} \otimes \sigma_{\nu_2}). \quad (123)$$

If  $\mu = 0$ , then

$$\mathbb{E} [H \mathbb{1}_2 H \otimes \sigma_\nu] = \mathbb{E} [H^2 \otimes \sigma_\nu] = \frac{16}{\Delta t} \mathbb{1}_2 \otimes \sigma_\nu. \quad (124)$$

Otherwise, for  $\mu \neq 0$ , at least one among  $\mu_1$  and  $\mu_2$  is not 0. Let us assume  $\mu_1 \neq 0$ . Then,  $\forall \beta$ ,  $\sigma_\alpha \sigma_{\mu_1} \sigma_\alpha \otimes \sigma_\beta \sigma_{\mu_2} \sigma_\beta$  equals  $\sigma_{\mu_1} \otimes \sigma_\beta \sigma_{\mu_2} \sigma_\beta$  for  $\alpha = 0, \mu_1$  and  $-\sigma_{\mu_1} \otimes \sigma_\beta \sigma_{\mu_2} \sigma_\beta$  for the other two indices of  $\alpha$ . Thus, summing over  $\alpha$  gives 0. The same applies for  $\mu_1$  arbitrary,  $\mu_2 \neq 0$ . We conclude that the second term in the expression for  $M_{2,\text{SLH}(\Delta t)}^{k=2}$  vanishes if both  $\mu$  and  $\nu$  are different from  $\{0, 0\}$ .

Now we look at the first part of the third term and we get that

$$\mathbb{E} [H^2\sigma_\mu \otimes \sigma_\nu] = \frac{1}{\Delta t} \sum_{\alpha,\beta} (\sigma_\alpha \otimes \sigma_\beta)^2 \sigma_\mu \otimes \sigma_\nu = \frac{16}{\Delta t} \sigma_\mu \otimes \sigma_\nu. \quad (125)$$

Hence, keeping terms to leading order in  $\Delta t$  we have

$$M_{2,\text{SLH}(\Delta t)}^{k=2}(\sigma_\mu \otimes \sigma_\nu) = (1 - 32\Delta t)\sigma_\mu \otimes \sigma_\nu - \Delta t^2 \mathbb{E} [[H, \sigma_\mu] \otimes [H, \sigma_\nu]] \quad (126)$$

$$= (1 - 32\Delta t)\sigma_\mu \otimes \sigma_\nu - \Delta t \sum_{\alpha,\beta} [\sigma_\alpha \otimes \sigma_\beta, \sigma_{\mu_1} \otimes \sigma_{\mu_2}] \otimes [\sigma_\alpha \otimes \sigma_\beta, \sigma_{\nu_1} \otimes \sigma_{\nu_2}], \quad (127)$$

when both  $\mu$  and  $\nu$  are different from  $\{0, 0\}$ , and conversely

$$M_{2,\text{SLH}(\Delta t)}^{k=2}(\mathbb{1}_2 \otimes \sigma_\nu) = (1 - 16\Delta t)\mathbb{1}_2 \otimes \sigma_\nu, \quad (128)$$

$$M_{2,\text{SLH}(\Delta t)}^{k=2}(\sigma_\mu \otimes \mathbb{1}_2) = (1 - 16\Delta t)\sigma_\mu \otimes \mathbb{1}_2. \quad (129)$$

We now divide the set of all possible strings  $\sigma_\mu \otimes \sigma_\nu$  in three parts: the identity  $\mathbb{1}_4$ , the set of strings of the form  $\sigma_\mu \otimes \sigma_\mu$ , and all remaining strings of the form  $\sigma_\mu \otimes \sigma_\nu$  with  $\mu \neq \nu$ . We can then make use of the matrix representation of the operator  $M_{2,\text{SLH}(\Delta t)}^{k=2}$  as a matrix with respect to Pauli basis, which gives

$$M_{2,\text{SLH}(\Delta t)}^{k=2} = \begin{pmatrix} 1 & & & \\ & A & & \\ & & & \\ & & & B \end{pmatrix}, \quad (130)$$

where  $A$  is a  $15 \times 15$  matrix related to the set of  $\sigma_\mu \otimes \sigma_\mu$  elements (without the identity  $\mathbb{1}_{16}$ ) and  $B$  is a  $240 \times 240$  matrix for  $\sigma_\mu \otimes \sigma_\nu$  elements. The detailed proof of this finding is laid out in the separate subsequent Lemma 25.

We now consider the matrix  $A$ ; we compute the action of  $M_{2,\text{SLH}}^{k=2}(\Delta t)$  over all possible  $\sigma_\mu \otimes \sigma_\mu$  and look for eigenvalues. We obtain a non-degenerate eigenvalue 1 whose eigenvector is the uniform sum over all non-identity Pauli matrices

$$\mathbb{F} = \frac{1}{15} \sum_{\gamma \neq 0} \sigma_\gamma \otimes \sigma_\gamma. \quad (131)$$

We then have a 9-fold degenerate eigenvalue  $1 - 40\Delta t$  and a 5-fold degenerate eigenvalue  $1 - 24\Delta t$ . We are free to bound all these eigenvalues with  $1 - 16\Delta t$ . We now deal with the action of the second moment operator on terms of the form  $\sigma_\mu \otimes \sigma_\nu$  with  $\mu, \nu \neq 0$  and  $\mu \neq \nu$ . Only four choices of  $\sigma_\alpha \otimes \sigma_\beta$  do not commute for a given pair  $\mu, \nu$ , i.e.:

$$M_{2,\text{SLH}}^{k=2}(\sigma_\mu \otimes \sigma_\nu) = (1 - 32\Delta t)\sigma_\mu \otimes \sigma_\nu - 4\Delta t \{ \pm \sigma_{\gamma_1} \otimes \sigma_{d_1} \pm \sigma_{\gamma_2} \otimes \sigma_{d_2} \pm \sigma_{\gamma_3} \otimes \sigma_{d_3} \pm \sigma_{\gamma_4} \otimes \sigma_{d_4} \} \quad (132)$$

with  $\gamma_i \neq d_i$ , for each  $\sigma_\mu \otimes \sigma_\nu$ . This means that each column of the matrix  $B$  has one entry  $(1 - 32\Delta t)$  (in the diagonal element) and four entries  $\pm 4\Delta t$ , and 0 otherwise. Hence,

$$\|B\|_1 = \max_j \sum_i |a_{i,j}| = 1 - 16\Delta t. \quad (133)$$

By the Gershgorin circle theorem, and taking also into account (128) and (129), we can upper bound the highest eigenvalue of  $B$  with  $1 - 16\Delta t$ . For a single time step, the two-qubit second moment operator can be upper bounded by the following diagonal matrix

$$M_{2,\text{SLH}}^{k=2} \leq \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 - 16\Delta t & & \\ & & & \ddots & \\ & & & & 1 - 16\Delta t \end{pmatrix}, \quad (134)$$

where we recall that the 2-fold degenerate eigenvalue 1 corresponds to the identity and  $\omega$ .  $\square$

**Lemma 25 (Local second moment operator).**  $M_{2,\text{SLH}}^{k=2}(\Delta t)$  is Hermitian, maps elements of the set of strings of the form  $\sigma_\mu \otimes \sigma_\mu$  to a linear combination of elements of the same set and elements of the set of strings of the form  $\sigma_\mu \otimes \sigma_\nu$  with  $\mu \neq \nu$  again to a linear combination of elements of the same set, such that there is no mixing between the two sets. Hence, we can represent the operator  $M_{2,\text{SLH}}^{k=2}(\Delta t)$  as a matrix with respect to Pauli basis in the following form

$$M_{2,\text{SLH}}^{k=2}(\Delta t) = \begin{pmatrix} 1 & \\ & A \\ & & B \end{pmatrix}, \quad (135)$$

where  $A$  is a  $15 \times 15$  matrix related to the set of  $\sigma_\mu \otimes \sigma_\mu$  elements (without the identity  $\mathbb{1}_{16}$ ) and  $B$  is a  $240 \times 240$  matrix for  $\sigma_\mu \otimes \sigma_\nu$  elements.

*Proof.* From eq. (126) and (127), follows directly that  $M_{2,\text{SLH}}^{k=2}(\Delta t)$  is Hermitian. Moreover we see, again from eq. (127), that elements of the set  $\sigma_\mu \otimes \sigma_\mu$  are mapped to a linear combination of elements of the same set. This, in addition to the fact that  $M_{2,\text{SLH}}^{k=2}(\Delta t)$  is Hermitian, implies that elements of the set  $\sigma_\mu \otimes \sigma_\nu$  with  $\mu \neq \nu$  are mapped again to a linear combination of elements of the same set.  $\square$

Next, we make use of this analysis to understand how  $n$ -qubit Pauli strings evolve during the continuous-time process. In Appendix B we collect the most relevant mathematical tools used in the second part of this section. As already mentioned, the continuous-time random walk induced by the Hamiltonian increments can be interpreted as a sequence of jumps defining a discrete random walk spaced out by i.i.d. waiting times.

**6.2. Markov chain analysis on weights.** The proof strategy for Lemma 12 begins with the analysis of the evolution of the coefficients: we observe how the support size behaves during the process, inferring a probability that, for a given initial string  $\sigma_\mu$  with support size  $\ell$ , after run time  $T$  the string has support size  $k$ . Conditioned on some specific event  $E_W$  that we will discuss later, this probability can be upper bounded as

$$\mathbb{P}(\{T, \ell, k\} \mid E_W) := \sum_{|\nu|=k} Q_{E_W}^T(\mu, \nu) \leq \binom{n}{k} 3^k \frac{4^{\delta n}}{4^n - 1}. \quad (136)$$

Having a total of  $\binom{n}{k} 3^k$  strings with support size  $k$ , we then show that almost all of them have the same probability.

Considering the analysis in the previous section on the two-qubit case and that, the local structure of  $M_{n, \text{SLH}(\Delta t)}^{k=2}$  given in eq. (118) we introduce a Markov chain over the weights of the string similarly to ref. [10] (where this projected chain is called *zero chain*). The chain runs over the state space  $\Omega = \{1, 2, \dots, n\}$  and the transition probability from  $\ell$  at time  $t$  to  $k$  at time  $t + \Delta t$  is described by the matrix element

$$P(\ell, k) := \sum_{\nu: |\nu|=k} \frac{1}{4^n} \text{Tr} \left[ \sigma_\nu \otimes \sigma_\nu M_{n, \text{SLH}(\Delta t)}^{k=2}(\sigma_\mu \otimes \sigma_\mu) \right] \quad (137)$$

for any choice of  $\mu$  with support size  $\ell$ .

**Lemma 26 (Transition matrix of the zero chain).** *The zero chain has transition matrix  $P$  on state space  $\Omega = \{1, 2, \dots, n\}$ ,*

$$P(\ell, k) = \begin{cases} 1 - \frac{16\ell(3n-2\ell-1)}{n(n-1)} \Delta t & k = \ell \\ \frac{16\ell(\ell-1)}{n(n-1)} \Delta t & k = \ell - 1 \\ \frac{48\ell(n-\ell)}{n(n-1)} \Delta t & k = \ell + 1 \\ 0 & \text{otherwise} \end{cases} \quad (138)$$

for  $1 \leq x, y \leq n$ .

*Proof.* We consider the analysis of the two-qubit second moment operator in Section 6.1. It is straightforward to note that, after application of  $M_{n, \text{SLH}(\Delta t)}^{k=2}$ , the weight of the string can only vary by 1 or stay the same. The weight decreases if a pair of two non-identity terms  $\sigma \otimes \sigma$  is chosen and is transformed in a pair with one identity element (namely  $\sigma \otimes \mathbb{1}$  or  $\mathbb{1} \otimes \sigma$ ); there are in total four choices for  $\sigma_\alpha \otimes \sigma_\beta$  which produce such a transition. According to the two-qubit case, the probability that one of these Pauli operators is chosen is  $4 \cdot 4\Delta t = 16\Delta t$  and since the probability of choosing a pair with weight 2 is  $\ell(\ell-1)/(n(n-1))$ , we have

$$P(\ell, \ell - 1) = \frac{16\ell(\ell-1)}{n(n-1)} \Delta t. \quad (139)$$

The weight of the string can be increased if an identity term paired with a non-identity term is chosen (i.e.,  $\sigma \otimes \mathbb{1}$  or  $\mathbb{1} \otimes \sigma$ ) and transformed into a pair of two non-identity terms  $\sigma \otimes \sigma$ . The probability of obtaining such a result (conditioned on choosing such a pair) after application of the two-qubit second moment operator is  $24\Delta t$ , since there are in total 6 choices for  $\sigma_\alpha \otimes \sigma_\beta$  to produce such a transition. Furthermore, the probability of choosing an identity and non-identity pair is given by  $2\ell(n-\ell)/(n(n-1))$ ; hence

$$P(\ell, \ell + 1) = \frac{48\ell(n-\ell)}{n(n-1)} \Delta t. \quad (140)$$

Finally, the probability of staying at the same weight is obtained by simply requiring the total probability to sum to unity.  $\square$

It is therefore possible to reach each state of the chain, meaning that it is *irreducible*. Moreover, the chain contains self loops, being hence *aperiodic*. From these two properties follows that the chain is also *ergodic*, thus converging to a unique stationary distribution.

**Lemma 27 (Stationary distribution of zero chain).** *The stationary distribution of the zero chain is*

$$\omega_0(k) = \frac{3^k \binom{n}{k}}{4^n - 1}. \quad (141)$$

*Proof.* This follows from straightforward calculation.  $\square$

The stationary distribution is actually analogous to the one of the chain induced by a random quantum circuit under the Haar measure (see ref. [10, Lemma 5.3]). Another crucial analogy is the exact equivalence of the *accelerated chain* (i.e., the chain conditioned on moving) of the two different settings. This means that, when moving, the random walk on weights is identically biased for both random quantum circuits under Haar distribution and the stochastic Hamiltonian process. From the description of Montroll and Weiss, the jumps of the random quantum circuit are contained in the fluctuating Hamiltonian evolution, spaced out by i.i.d. waiting times. Concretely, the accelerated chain is given by

$$P_{\text{accel}}(\ell, k) = \begin{cases} 0 & k = \ell \\ \frac{\ell-1}{3^{n-2\ell-1}} & k = \ell - 1 \\ \frac{3^{(n-\ell)}}{3^{n-2\ell-1}} & k = \ell + 1 \\ 0 & \text{otherwise.} \end{cases} \quad (142)$$

With these analogies, we can prove the next theorem using results from the proof of ref. [4, Theorem 4.2]. We should take care of the parts of the proof involving the waiting time, because it is where the two walks differ. We will also deal with the permutation invariance property in a more precise and explicit way. Now, we reformulate the result for the continuous-time case.

**Lemma 28 (Mixing condition on support size).** *Let  $P$  be the Markov chain transition matrix defined in Lemma 26. For any constants  $\delta \in (0, 1/16)$ ,  $\eta \in (0, 1)$  there exists a constant  $\varsigma > 0$  such that for  $T \geq \varsigma n \log^2 n$  and all integers  $1 \leq \ell \leq n$  and  $1 \leq k \leq n$ , we have for large enough  $n$*

$$\mathbb{P}(\{T, \ell, k\}) = \sum_{\nu: |\nu|=k} Q^T(\ell, k) \leq \binom{n}{k} 3^k \frac{4^{\delta n}}{4^n - 1} + \frac{1}{(3 - \eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)}, \quad (143)$$

where  $\{T, \ell, k\}$  is the event that an initial Pauli string with support size  $\ell$ , after a run time  $T$ , has weight equal to  $k$ .

*Proof.* We start by defining the following points,

$$r_- := \left(\frac{3}{4} - \delta\right) n \quad \text{and} \quad r_+ := \left(\frac{3}{4} + \delta\right) n. \quad (144)$$

Then, considering ref. [4, eq. (20)], it follows that for an initial weight of  $\ell \in [r_-, r_+]$

$$\mathbb{P}(\{T, \ell, k\}) \leq \binom{n}{k} 3^k \frac{4^{\delta n}}{4^n - 1} \quad (145)$$

for any  $T > 0$ .

To deal with the case  $\ell \in [1, r_-)$ , for random quantum circuits it has been shown that the probability that the interval  $[r_-, r_+]$  of the state space has been reached is very high for a number of gates  $O(n \log^2 n)$ . Here we prove the same scaling result for the run time of the continuous-time process, that is, the total waiting time between the jumps can be bound with the following lemma.

**Lemma 29 (Waiting time).**

$$\mathbb{P}(E_W^c) := \mathbb{P}(W_{r_-} > \varsigma n \log^2 n) \leq \frac{1}{(3-\eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)} \quad (146)$$

for some sufficiently large  $\varsigma$ .

The proof of the lemma is postponed to Appendix C to help readability. The case that remains to be discussed is the one of an initial Pauli string with support size  $\ell \in (r_+, n]$  to reach  $[r_-, r_+]$ ; again the analysis is divided on accelerated steps and waiting times. Regarding the former, the probability of going backward is larger than the one of moving forward starting from point  $z$  with

$$P(z, z+1) \stackrel{!}{=} P(z, z-1), \quad (147)$$

$$\frac{z-1}{3n-2z-1} = 3 \frac{n-z}{3n-2z-1}, \quad (148)$$

from which follows that

$$z = \frac{3}{4}n + \frac{1}{4}. \quad (149)$$

This means that for any  $n > 1/(4\delta)$  the probability of moving backward at each site of region  $(r_+, n]$  is at least  $1/2 + \epsilon$  for some  $\epsilon > 0$ , and again using the argument for the case with  $\ell < r_-$  the probability of not reaching  $r_+$  in  $S \leq s$  steps is upper bounded by an exponential decreasing function for  $s \geq \phi' n$  for sufficiently large  $\phi'$ . In this instance, all waiting times are stochastically dominated by parameter  $p(3n/4) = 12$ , hence there is no necessity to define an event equivalent to  $H$ . For  $S \leq s$  accelerated steps, using again a Markov's inequality, the bound on the total waiting time is exponentially decreasing in  $s$  for a run time  $W_{r_+} > (\log 2/6)s$ . The proof of Lemma 28 is then complete.  $\square$

**6.3. From the zero chain to the full distribution.** Once the weight distribution has reached an equilibrium such that the condition in eq. (136) is fulfilled, we need to show that all Pauli strings sharing the same weight have a similar probability. To prove this, we need to show that almost all Pauli strings with the same support but different Pauli labels  $\{1, 2, 3\}$  are equivalent in probability. This, together with the permutation invariance property assumed for the initial state, which is conserved during the whole stochastic Hamiltonian process, will bring us to the desired result.

Let  $M$  be the Markov chain on the first  $n$ -qubits induced by  $M_{n,\Delta t}^2$ , and define an accelerated version as

$$A := \frac{1}{36\Delta t} (M - (1 - 36\Delta t)\mathcal{I}). \quad (150)$$

If we define an operator

$$R = \frac{2}{n(n-1)} \sum_{j < k} R_{j,k}, \quad (151)$$

where  $R_{j,k}$  randomises one qubit site in the following way,

$$R_{j,k}(\sigma_\mu^j \otimes \sigma_\nu^k) = \begin{cases} \frac{1}{3} \sum_{\alpha=1,2,3} \sigma_\alpha^j \otimes \mathbb{1}^k & \text{if } \mu \neq 0, \nu = 0, \\ \frac{1}{3} \sum_{\alpha=1,2,3} \mathbb{1}^j \otimes \sigma_\alpha^k & \text{if } \mu = 0, \nu \neq 0, \\ \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_\alpha^j \otimes \sigma_\nu^k + \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_\mu^j \otimes \sigma_\alpha^k & \text{if } \mu \neq 0, \nu \neq 0, \\ \mathbb{1}^j \otimes \mathbb{1}^k & \text{if } \mu = \nu = 0, \end{cases} \quad (152)$$

then according to Section 6.1, the accelerated chain can be written as

$$A = \frac{1}{3}R + \frac{2}{3}L, \quad (153)$$

where

$$L = \frac{2}{n(n-1)} \sum_{j < k} L_{j,k} \quad (154)$$

and

$$L_{j,k}(\sigma_\mu^j \otimes \sigma_\nu^k) = \begin{cases} \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_{\mu+1}^j \otimes \sigma_\alpha^k + \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_{\mu+2}^j \otimes \sigma_\alpha^k & \text{if } \mu \neq 0, \nu = 0, \\ \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_\mu^j \otimes \sigma_{\nu+1}^k + \frac{1}{6} \sum_{\alpha=1,2,3} \sigma_\mu^j \otimes \sigma_{\nu+2}^k & \text{if } \mu = 0, \nu \neq 0, \\ \frac{1}{12} \left( \sigma_{\mu+1}^j \otimes \sigma_\nu^k + \sigma_{\mu+2}^j \otimes \sigma_\nu^k + \sigma_\mu^j \otimes \sigma_{\nu+1}^k + \sigma_\mu^j \otimes \sigma_{\nu+2}^k \right) \\ + \frac{1}{6} \left( \sigma_{\mu+1}^j \otimes \mathbb{1}^k + \sigma_{\mu+2}^j \otimes \mathbb{1}^k + \mathbb{1}^j \otimes \sigma_{\nu+1}^k + \mathbb{1}^j \otimes \sigma_{\nu+2}^k \right) & \text{if } \mu \neq 0, \nu \neq 0, \\ \mathbb{1}^j \otimes \mathbb{1}^k & \text{if } \mu = \nu = 0, \end{cases} \quad (155)$$

with the notation  $\sigma_{3+1} = \sigma_{2+2} = \sigma_1$  and  $\sigma_{3+2} = \sigma_2$ . Note that  $R$  does not produce any change in the weight or transpositions between identities and non-identity elements, it solely performs a local randomisation of the Pauli labels. This means that only the chain  $L$  is responsible for the random walk on the weights.

We would like to upper bound the probability that more than  $\beta n$  sites have not been randomised after  $s$  steps of chain  $R$  (we denote the complement of this event as  $E_R$ ). Knowing that there are  $\binom{n}{\beta n}$  such regions, this is given by union bound

$$\mathbb{P}(E_R^c) \leq \binom{n}{\beta n} (1 - \beta)^s \leq 2^{h(\beta)n} e^{-\beta s}, \quad (156)$$

where  $h : [0, 1] \rightarrow [0, 1]$  is the binary entropy function. This probability can then be upper bounded by an arbitrary exponentially decreasing function in  $n$  for some  $s = O(n)$ . Hence, to ensure that  $s$  randomisations have been performed to fulfill the event  $E_R$  with sufficiently large probability, given eq. (153) and by application of an Hoeffding's inequality follows that it is again sufficient to apply  $O(n)$  steps of the accelerated chain  $A$ . Since the waiting time is dominated by an exponential distribution with parameter 36, the bound on the probability for the waiting time of this process to exceed  $W_R = \varsigma_R n$  can be bounded by an arbitrarily exponentially decreasing function in  $n$  for a sufficiently large  $\varsigma_R$  with the same argument used for the random walk on weights when starting from  $\ell > r_+$ .

In conclusion, assuming that event  $E_W$  and  $E_R$  have been satisfied, we have for  $\gamma < \gamma_0 \leq 1/2$ :

1. For strings  $\nu$  with support size  $k \leq \gamma_0 n$ ,

$$Q^T(\mu, \nu) \leq \sum_{|\nu|=k} Q^T(\mu, \nu) \leq \binom{n}{\gamma_0 n} 3^{\gamma_0 n} \frac{4^{\delta n}}{4^n - 1} \leq 2^{n h(\gamma_0)} 3^{\gamma_0 n} \frac{4^{\delta n}}{4^n - 1}. \quad (157)$$

2. For strings  $\nu$  with support size  $k \geq (1 - \gamma_0)n$ , given event  $E_R$  at least  $(1 - \beta)n$  sites of the support have been uniformly randomised, hence

$$Q^T(\mu, \nu) \leq \frac{1}{3^{k-\beta n}} \sum_{|\nu|=k} Q^T(\mu, \nu) \leq 2^{n h(\gamma_0)} 3^{\beta n} \frac{4^{\delta n}}{4^n - 1}. \quad (158)$$

3. For strings  $\nu$  with support size  $\gamma_0 n < k = \kappa n < (1 - \gamma_0)n$  such that  $\kappa - \gamma_0 = O(1)$  (otherwise, we can apply slightly modified versions of the bounds in the two previous cases), given event  $E_R$  at least  $(1 - \beta)n$  sites of the support have been uniformly randomised. In addition, if we assume the  $\gamma$ -permutation invariance property for the initial string  $\sigma_\mu$ , we obtain

$$Q^T(\mu, \nu) \leq \frac{1}{3^{k-\beta n}} \frac{1}{\binom{(1-\gamma)n}{k-\gamma n}} \sum_{|\nu|=k} Q^T(\mu, \nu) \leq 3^{\beta n} \left[ \frac{1}{\kappa - \gamma_0} \right]^{\gamma_0 n} \frac{4^{\delta n}}{4^n - 1}. \quad (159)$$

Now, for an appropriate choice of  $\beta$  and  $\gamma_0$ ,

$$Q^T(\mu, \nu) \leq \frac{5^{\delta n}}{4^n - 1} \quad (160)$$

for all  $\mu$  and  $\nu$ .

Also, having proven that there exists  $\varsigma$  such that, for all  $T \geq \varsigma n \log^2 n$ ,  $\mathbb{P}(E_R^c)$  is bounded by an exponentially decreasing function in  $n$  and

$$\mathbb{P}(E_W^c) \leq \frac{1}{(3 - \eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)}. \quad (161)$$

Having proven that, if both event have been satisfied and the permutation invariance property is assumed, we have

$$Q^T(\mu, \nu) \leq \frac{5^{\delta n}}{4^n - 1} \quad (162)$$

for all  $\mu$  and  $\nu$ , we conclude the proof for the main Lemma 12.

As mentioned in the main result section, the decoupling Theorem holds for all states which are invariant with respect to any permutation on  $(1 - \gamma)n$  qubits, in the sense of Definition 11, and not only for Pauli strings taken singularly. Consider a set of  $\min \left\{ \binom{n-\gamma n}{\ell-\gamma n}, \binom{n-\gamma n}{\ell} \right\} \leq b_{\ell, \gamma} \leq \max \left\{ \binom{n-\gamma n}{\ell-\gamma n}, \binom{n-\gamma n}{\ell} \right\}$  Pauli strings  $\{\sigma_\mu\}_\mu$  with support size  $\ell$  which is invariant with respect to any of such permutations. Assuming that the above events have been satisfied, at least the same number of qubits in the final Pauli strings  $\{\sigma_\nu\}_\nu$  is invariant with respect to permutations since the stochastic evolution preserves this property. Hence, for the argument from the previous subsection, we have:

$$\sum_{\mu} Q^T(\mu, \nu) \leq b_{\ell, \gamma} \frac{5^{\delta n}}{4^n - 1} \quad (163)$$

This, together with the fact that  $\text{Tr}[\sigma_\mu \rho]$  is the same for all strings related by these permutations, allows to apply the proof in ref. [4] for the decoupling Theorem for all density states  $\rho$  composed by permutation invariant sets of Pauli strings.

## 7. Conclusions and outlook

In this work, we have investigated mixing properties of fluctuating local Hamiltonian evolutions, establishing a connection with random quantum circuits. The two settings differ on the distribution over unitary group: in the random quantum circuits considered in other works two-qudit gates are chosen from the Haar measure or a fixed distribution. The discretised stochastic Hamiltonian is described by local terms weighted random coefficients also generating a gate set. However, the gate set depends on the discretisation which required an involved analysis of the gap of the local moment operators. We show that scaling in the system size in order to obtain an approximate unitary  $k$ -design are compatible in the two settings: the total run time of the diffusion process provides a faster mixing time, by a factor of  $n$ , in comparison of a local random quantum circuit due to the larger number of interactions per time step, but



the two scenarios display the same scaling when they interact with two qubits only at each step. In this way, we provide a unifying framework of random quantum processes.

In order to bound the gap of the local moment operator, we have made use of and further developed tools from representation theory, significantly going beyond uses of representation theory in related contexts [7, 62]. With this, we analyse how quickly the diffusion on the unitary group induced by the local stochastic Hamiltonian mixes, where the local gap characterises the speed of the diffusion. The gap can be lower bounded by an expression which is entirely independent of the number of copies  $k$  of the system, which constitutes a possibly surprising result in its own right.

In the framework of a continuous-time random walk on weights induced by the stochastic Hamiltonian evolution, we prove a decoupling theorem with almost linear scaling in  $n$ , already shown to be valid for random quantum circuits. The exact correspondence between the accelerated steps of the walk derived from the random circuit and the jumps of the continuous-time random walk originating from the stochastic Hamiltonian is a strong element of similarity: we can consider the steps of the circuit as if they were dispersed within the continuous-time process and spaced out by i.i.d. waiting times. Again, a unifying picture is hence provided.

All these results allow us to unify in one single mathematical framework random quantum processes in the form of quantum circuits and continuous-time phenomena governed by time-fluctuating Hamiltonians. This is of interest for both a pragmatic and application-oriented [3, 4, 8, 10–13, 22] as well as a conceptual point of view [2, 9, 29], indeed giving guidance on how fast time-fluctuating processes lead to mixing or “fast scrambling”.

Given the close connection of fluctuating processes with classes of local dissipative processes, we also gain new insights into the impact of dissipation to quantum many-body dynamics. Turning the logic of approximating the Haar measure upside down, this work shows how dynamics can deviate from the uniform measure without affecting its mixing properties. It is the hope that the present work stimulates further research on random quantum processes, both as far as the mathematical development and the exploration of its implications are concerned.

## 8. Acknowledgements

We thank M. Horodecki, D. Gross, H. Wojewódka and I. Roth for fruitful discussions and acknowledge support from the EU (RAQUEL, AQUeS), the DFG (CRC 183, EI 519/7-1), the Templeton Foundation (RQ-35601), the ERC (TAQ), and the BMBF (Q.com).

## References

1. W. Brown and L. Viola. Convergence rates for arbitrary statistical moments of random quantum circuits, 2009. arXiv:0910.0913.
2. W. Brown and O. Fawzi. Scrambling speed of random quantum circuits, 2012. arXiv:1210.6644.
3. F. G. S. L. Brandao and M. Horodecki. Exponential quantum speed-ups are generic. *Q. Inf. Comp.*, 13:0901, 2013.
4. W. Brown and O. Fawzi. Decoupling with random quantum circuits, 2013. arXiv:1307.0632.
5. R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio. Efficient generation of generic entanglement. *Phys. Rev. Lett.*, 98:130502, 2007.
6. W. Brown and O. Fawzi. Short random circuits define good quantum error correcting codes. *Proc. ISIT*, page 346, 2013.
7. F. G. S. L. Brandao, P. Cwiklinski, M. Horodecki, P. Horodecki, J. Korbicz, and M. Mozrymas. Convergence to equilibrium under a random Hamiltonian. *Phys. Rev. E*, 86:031101, 2012.
8. S. Hallgren and A. W. Harrow. Superpolynomial speedups based on almost any quantum circuit. *Proc. of the 35th Int. Coll. Aut. Lang. Prog. LNCS*, 5125:782, 2008.
9. N. Lashkari, D. Stanford, M. Hastings, T. J. Osborne, and P. Hayden. Towards the fast scrambling conjecture. *JHEP*, 2013: 22, 2013.
10. A. W. Harrow and R. A. Low. Random quantum circuits are approximate 2-designs. *Comm. Math. Phys.*, 291:257, 2009.
11. L. Bouten and R. van Handel. Discrete approximation of quantum stochastic models. *J. Math. Phys.*, 49:102109, 2008.
12. F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki. Local random quantum circuits are approximate polynomial-designs, 2012. arXiv:1208.0692.

13. Y. S. Weinstein, W. G. Brown, and L. Viola. Parameters of pseudo-random quantum circuits. *Phys. Rev. A*, 78:052332, 2008.
14. A. C. R. Belton, M. Gnacik, and J. M. Lindsay. The convergence of unitary quantum random walks, 2014.
15. D. Gross, K. M. R. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J. Math. Phys.*, 48:052104, 2007.
16. D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2008.
17. D. Aldous and P. Diaconis. Shuffling cards and stopping time. *Am. Math. Soc. Mon.*, 93(5):333–348, 1986.
18. V. Dunjko and H. J. Briegel. Quantum mixing of Markov chains for special distributions. arXiv:1502.05511.
19. M. Kabanava, R. Kueng, H. Rauhut, and U. Terstiege. Stable low-rank matrix recovery via null space properties, 2015. arXiv:1507.07184.
20. M. Ohliger, V. Nesme, and J. Eisert. Efficient and feasible state tomography of quantum many-body systems. *New J. Phys.*, 15:015024, 2013.
21. E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:12307, 2008.
22. O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner. Decoupling with unitary approximate two-design. *New J. Phys.*, 15:053022, 2013.
23. M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. *Commun. Math. Phys.*, 269:107, 2007.
24. C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter. Quantum reverse Shannon theorem, 2009. arXiv:0912.5537.
25. F. Buscemi. Private quantum decoupling and secure disposal of information. *New J. Phys.*, 11:123002, 2009.
26. C. Gogolin and J. Eisert. Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems. *Rep. Prog. Phys.*, 79:56001, 2016.
27. T. Banks, W. Fischler, S. Shenker, and L. Susskind.  $m$  theory as a matrix model: A conjecture. *Phys. Rev. D*, 55:5112, 1997.
28. J. Maldacena. The large  $n$  limit of super-conformal field theories and supergravity. *Adv. Th. Math. Phys.*, 2:213, 1998.
29. Y. Sekino and L. Susskind. Fast scramblers. *JHEP*, 10:65, 2008.
30. C. T. Asplund, D. Berenstein, and D. Trancanelli. Evidence for fast thermalization in the plane-wave matrix model. *Phys. Rev. Lett.*, 107:171602, 2011.
31. R. Hübener, Y. Sekino, and J. Eisert. Equilibration in low-dimensional quantum matrix models. *JHEP*, 2015:166, 2015.
32. J. Bourgain and A. Gamburd. A spectral gap theorem in  $su(d)$ . *Journal of the European Mathematical Society*, 14(5):1455–1511, 2012.
33. Y. Benoist and N. de Saxcé. A spectral gap theorem in simple lie groups. *Inventiones mathematicae*, pages 1–25.
34. E. W. Montroll and G. H. Weiss. Random walks on lattices ii. *J. Math. Phys.*, 6:167–181, 1965.
35. G. H. Weiss. Aspects and applications of the random walk. *J. Stat. Phys.*, 79(1):497–500, 1995.
36. V. Zaburdaev, S. Denisov, and P. Hanggi. Perturbation spreading in many-particle systems: A random walk approach. *Phys. Rev. Lett.*, 106, 2011.
37. J. H. P. Schulz and E. Barkai. Fluctuations around equilibrium laws in ergodic continuous-time random walks. *Phys. Rev. E*, 91, 2015.
38. P. Chaudhuri, Y. Gao, L. Berthier, M. Kilfoil, and W. Kob. A random walk description of the heterogeneous glassy dynamics of attracting colloids. *J. Phys. Cond. Mat.*, 20, 2008.
39. J. Watrous. Semidefinite programs for completely bounded norms. *Th. Comp.*, 5(11), 2009.
40. R. Low. *Pseudo-randomness and learning in quantum computation*. PhD thesis, university of Bristol, 2010.
41. B. Collins and P. Sniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Commun. Math. Phys.*, 264:773, 2006.
42. F. J. Dyson. The radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.*, 75, 1949.
43. J. D. Dollard and C. N. Friedman. Product integrals and the Schrödinger equation. *J. Math. Phys.*, 18:1598, 1977.
44. S. Ito. Brownian motions in a topological group and in its covering group. *Rend. Circ. Mat. Palermo*, 1:40–48, 1952.
45. B. Tsirelson. Unitary Brownian motions are linearisable, 1988. arXiv:math/9806112.
46. M. Liao. *Lévy processes in Lie groups*, volume 162. Cambridge university press, 2004.
47. H. P. McKean. *Stochastic integrals*. Academic Press, 1969.
48. L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales*, volume 2. Cambridge Mathematical Library, 2 edition, 2000.
49. I. T. Diniz and D. Jonathan. Comment on the paper "Random quantum circuits are approximate 2-designs". *Comm. Math. Phys.*, 304, 2011.
50. P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Probability theory and related fields*, 57(2):159–179, 1981.
51. M. Choi. Completely positive linear maps on complex matrices. *Lin. Alg. App.*, pages 285–290, 1975.
52. F. Dupuis, M. Berta, J. Wullschleger, and R. Renner. One-shot decoupling. *Commun. Math. Phys.*, 328, 2014.

53. S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Buechler, and P. Zoller. Quantum states and phases in driven open quantum systems with cold atoms. *Nature Phys.*, 4:878, 2008.
54. F. Verstraete, M. M. Wolf, and J. I. Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Phys.*, 5(9):633, 2009.
55. M. J. Kastoryano, M. M. Wolf, and J. Eisert. Precisely timing dissipative quantum information processing. *Phys. Rev. Lett.*, 110:110501, 2013.
56. P. Hayden and J. Preskill. Black holes as mirrors: quantum information in random subsystems. *JHEP*, 2007:120, 2007.
57. S. Lloyd and J. Preskill. Unitarity of black hole evaporation in final-state projection models. *JHEP*, 2014:1, 2014.
58. H. Georgi. *Lie algebras in particle physics*. Westview Press, 2 edition, 1999.
59. D. A. Roberts and B. Yoshida. Chaos and complexity by design. arXiv:1610.04903.
60. D. H. Sattinger and O. L. Weaver. *Lie groups and algebras with applications to physics, geometry and mechanics*. Springer-Verlag Berlin Heidelberg, 1986.
61. W. Fulton and J. Harris. *Representation theory: A first course*. Springer, Heidelberg, 1991.
62. J. Eisert, T. Felbinger, P. Papadopoulos, M. B. Plenio, and M. Wilkens. Classical information and distillable entanglement. *Phys. Rev. Lett.*, 84:1611, 2000.

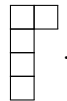
## Appendix A (Young diagrams)

In order to study the decomposition of the mixed tensor representation  $\pi_{k,k}$ , we make use of *Young diagrams* for  $\mathfrak{su}(N)$ . These are arrays of boxes arranged in  $N - 1$  left-justified rows whose length is non-increasing from top to bottom, each of them connected to an irreducible representation, e.g.,

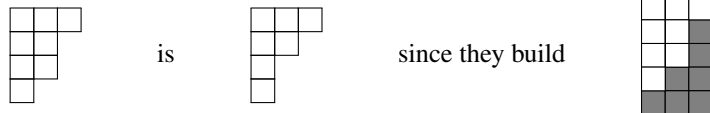


In particular, the following holds true.

- The Young diagram of the fundamental representation is given by one single box  $\square$ .
- The trivial representation does not have any box; we can denote it by  $\emptyset$ .
- The adjoint representation is given by a column of  $N - 1$  boxes and a second column made of a single box. For example, the adjoint representation of  $\mathfrak{su}(5)$  is given by



- The conjugate representation of a Young diagram whose first row contains  $\ell$  boxes is given by the complementary diagram (rotated by 180 degrees) shaping the rectangle of  $N$  rows and  $\ell$  columns. For example, for  $\mathfrak{su}(5)$  the conjugate representation of



Note that the conjugate diagram of the fundamental representation is given by a single column of  $N - 1$  boxes, while the adjoint representation is self-conjugate.

Young diagrams are particularly helpful when decomposing the tensor product of two representations into a direct sum of irreducible representations. Here, one follows two steps: first, one combines the boxes of the two diagrams by adding, one at a time, all boxes in the first row of the second diagram to the first one, respecting the condition of non-increasing length from top to bottom for the rows of

the newly created diagrams and remembering that each of them can have at most  $N$  rows. One repeats the procedure for all rows in the second diagram. As a second step, one discards all diagrams which do not satisfy specific rules that we are not going to mention here; for a full description, see ref. [58]. Furthermore, for the algebra  $\mathfrak{su}(N)$  all columns with  $N$  boxes occurring in a diagram can be deleted.

Recalling that the tensor product of the fundamental representation and its conjugate can be decomposed as a direct sum of the trivial and the adjoint representation and taking again  $\mathfrak{su}(5)$  as an example, we have

$$\bar{U} \otimes U = \bar{\pi}_f \otimes \pi_f = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array} \otimes \square = \emptyset \oplus \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} = \pi_1 \oplus \pi_{\text{ad}}, \quad (164)$$

since a diagram with a column of  $N = 5$  boxes is equivalent to the trivial representation.

An alternative way to express an irreducible representation of  $\mathfrak{su}(n)$  is to associate a *Dynkin label*  $(\lambda_1, \lambda_2, \dots, \lambda_{N-1})$ , where  $\lambda_b$  gives the number of columns made of  $b$  boxes. For instance, the fundamental representation is given by the label  $(1, 0, \dots, 0)$  and the adjoint representation by  $(1, 0, \dots, 0, 1)$ .

## Appendix B (Stochastic processes and Markov chains)

A (discrete) stochastic process with a sequence of random variables  $X_1, X_2, \dots$  whose next step depends solely on the current state is called a *Markov chain*. We consider a countable set of values  $\Lambda = \{\lambda_1, \lambda_2, \dots\}$  which the variables  $X_j$  can assume during the process and denote it as *state space*. For the variable  $X_j$  we can then assign a *probability distribution*  $\omega_j = (\omega_j^1, \omega_j^2, \dots)$  where  $\omega_j^k = \mathbb{P}(X_j = \lambda_k)$ . If the state space is finite, the transition from  $X_j$  to  $X_{j+1}$  can be described by a *transition matrix*  $P_j$  with entries

$$p_{a,b} = \mathbb{P}(X_{j+1} = \lambda_b | X_j = \lambda_a) \quad (165)$$

such that we have

$$\omega_{j+1} = \omega_j P_j. \quad (166)$$

If the process is homogeneous, then each transition is governed by the same transition matrix  $P$ , and

$$\omega_n = \omega_0 P^n. \quad (167)$$

The *stationary distribution* of the process  $\omega$  satisfies

$$\omega = \omega P \quad (168)$$

and can hence be regarded as a fixed point of the chain.

For an ergodic chain, we refer as the *mixing time* of the chain to the number of steps required to reach closeness to the stationary distribution. For two arbitrary distributions  $\omega$  and  $\eta$ , the *total variation distance* is given by

$$\|\omega - \eta\|_{TV} = \frac{1}{2} \|\omega - \eta\|_1 = \frac{1}{2} \sum_j |(\omega)_j - (\eta)_j|. \quad (169)$$

Then the *mixing time* is defined as

$$\tau(\varepsilon) := \max_{\omega_0} \min_{t \geq 0} \{t : \|\omega_0 P^t - \omega\|_{TV} \leq \varepsilon\}, \quad (170)$$

where  $\omega_0$  is the initial probability distribution and  $\omega$  the stationary distribution.

The function  $\{N(t) : t \geq 0\}$  counting the number of jumps occurred up to the positive time  $t$  defines a *Poisson process* if the following properties are satisfied.

1.  $N(0) = 0$ .
2. The increments are independent and stationary.
3. Each increment  $N(t + \Delta t) - N(t)$  is distributed as a Poisson random variable with parameter (mean)  $\lambda t$ .

The last condition implies that  $\mathbb{E}[N(t)] = \lambda t$  and, in particular, the probability that two or more jumps occur in the time interval is negligible when it is small.

The waiting time  $W$  between two consecutive jumps is then described by an *exponential distribution*, having for  $\lambda > 0$  a *cumulative distribution function*

$$\mathbb{P}(W \leq t) = 1 - e^{-\lambda t} \quad (171)$$

and a *probability density function*

$$f(t) = \lambda e^{-\lambda t}. \quad (172)$$

### Appendix C (Proof of Lemma 29)

To prove this result on the waiting time, we first assume that we reach the region  $[r_-, r_+]$  within  $S \leq s$  accelerated steps for some  $s = O(n)$  and we bound the probability that the waiting time exceeds  $\varsigma n \log^2 n$ . We will deal with the case of  $\mathbb{P}(S > s)$  afterwards. Now, let  $M$  be the smallest site visited during the walk, and let  $\{y_i\}_{i=1}^S$  be a sequence of accelerated steps where  $S \leq s$ , with waiting times  $\{W_i\}_{i=1}^S$  respectively, satisfying the event

$$H = \bigcap_{j=1}^n \left[ \sum_{k=1}^S \mathbb{I}(X_k \leq j) \leq zj/\mu \right], \quad (173)$$

where  $\mathbb{I}$  is the indicator function and  $X_k$  is the random variable assuming values in  $\Omega = \{1, 2, \dots, n\}$  describing the state of the chain at step  $k$  and  $z$  chosen as  $O(\log n)$ . In words, this means that, if  $H$  occurs, then no site has been visited “too often”. This is a useful event, since the smaller is the value of the current state of the chain, the smaller is the parameter of the exponential distribution dominating the waiting time. Namely, we have

$$1 - P(k, k) = \frac{16k(3n - 2k - 1)}{n(n-1)} dt \geq \frac{16k}{n} \Delta t. \quad (174)$$

So, dealing with three events, we consider the bound

$$\begin{aligned} \mathbb{P}(W > t) &= \mathbb{P}(W > t \cap H \cap S \leq s) + \mathbb{P}(W > t \cap H \cap S > s) \\ &\quad + \mathbb{P}(W > t \cap H^c \cap S \leq s) + \mathbb{P}(W > t \cap H^c \cap S > s) \\ &\leq \mathbb{P}(W > t \mid H \cap S \leq s) + \mathbb{P}(H \cap S > s) \\ &\quad + \mathbb{P}(H^c \cap S \leq s) + \mathbb{P}(H^c \cap S > s) \\ &\leq \mathbb{P}(W > t \mid H \cap S \leq s) + \mathbb{P}(H^c \mid S \leq s) + \mathbb{P}(S > s). \end{aligned} \quad (175)$$

Conditioning on the two previous event and setting  $M = m$  for arbitrary  $m \in \{1, \dots, \ell\}$ , we have to find an upper bound for the waiting time being too large; more precisely for a given run time  $t$ , we show:

#### Lemma 30 (Waiting time conditioning on event $H$ ).

$$\max_{\{y_i\}} \mathbb{P}(W(y_1) + \dots + W(y_S) \geq t \mid M = m, H) \leq e^{-\frac{8k}{n}t} 2^{zm/\mu} e^{zm/(2\mu) \log n}. \quad (176)$$

*Proof.* (Proof of Lemma 30) We recall that this is the exactly the sequence visiting  $m$  for  $zm/\mu$  (for simplicity, we assume it to be an integer) times and all other  $j > m$  sites for  $z/\mu$  times, hence

$$W(y_1) + \dots + W(y_S) \leq \sum_{i=1}^{zm/\mu} E_{m,i} + \sum_{i=1}^{z/\mu} \sum_{k=m+1}^r E_{k,i}, \quad (177)$$

where  $E_{k,i}$  are i.i.d. exponential distributions with parameter  $p(k) = 16k/n$ . Now applying Markov's inequality we obtain

$$\begin{aligned} \mathbb{P} \left( \sum_{i=1}^{zm/\mu} E_{m,i} + \sum_{i=1}^{z/\mu} \sum_{k=m+1}^r E_{k,i} > t \right) &\leq \frac{\mathbb{E} \left[ \exp \left\{ \alpha \left( \sum_{i=1}^{zm/\mu} E_{m,i} + \sum_{i=1}^{z/\mu} \sum_{k=m+1}^r E_{k,i} \right) \right\} \right]}{e^{\alpha t}} \\ &= e^{-\alpha t} \left( \frac{p(m)}{p(m) - \alpha} \right)^{zm/\mu} \prod_{k=m+1}^r \left( \frac{p(k)}{p(k) - \alpha} \right)^{z/\mu} \end{aligned} \quad (178)$$

for  $\alpha < p(m)$ . Let us choose  $\alpha = p(m)/2$ , then we have

$$\begin{aligned} \mathbb{P} \left( \sum_{i=1}^{zm/\mu} E_{m,i} + \sum_{i=1}^{z/\mu} \sum_{k=m+1}^r E_{k,i} > t \right) &\leq e^{-\frac{8m}{n}t} 2^{zm/\mu} \left( \prod_{k=m+1}^r \frac{2k}{2k-m} \right)^{z/\mu} \\ &\leq e^{-\frac{8m}{n}t} 2^{zm/\mu} e^{zm/(2\mu) \log n}. \end{aligned} \quad (179)$$

□

With this lemma we obtain an equivalent result for the waiting time as in ref. [4] up to the prefactor of  $t$ . Hence, for  $t > \varsigma n \log^2 n$  with  $\varsigma$  sufficiently large, applying the bounds on the probabilities  $\mathbb{P}(M = m)$  for each value of  $m \in \{1, \dots, \ell\}$  proved for the random quantum circuit case, we have

$$\begin{aligned} \mathbb{P}(W_{r_-} > t \mid H \cap S \leq s) &= \sum_{m=1}^{\ell} (M = m) \max_{\{y_i\}} \mathbb{P}(W(y_1) + \dots + W(y_s) \geq t \mid M = m) \\ &\leq \frac{1}{(3-\eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)}. \end{aligned} \quad (180)$$

The last two probability terms in eq. (175) depend only on the path of the accelerated random walk before reaching the interval  $[r_-, r_+]$ . Looking at the accelerated chain and considering  $\ell$  being in the region  $[1, (3/4 - \delta)n]$ , we have  $3(n - \ell)/(3n - 2\ell - 1) \geq 1/2 + \delta$  for any  $n$ . So, constructing a random walk  $X'_k$  starting at the origin moving forward with probability  $1/2 + \delta$  and backward with  $1/2 - \delta$ , it follows

$$\begin{aligned} \mathbb{P}(S > s) &\leq \mathbb{P}(X'_s < r_- - \ell) \\ &= \mathbb{P}(X'_s < 2\delta s - (2\delta s + \ell - r_-)) \\ &\leq \exp \left( -\frac{(2\delta s + \ell - r_-)^2}{2s} \right), \end{aligned} \quad (181)$$

where in the last inequality we have used the Chernoff bound in ref. [10, Lemma A.3] assuming  $2\delta s + \ell - r_- > 0$ . We conclude that the probability for the waiting time to be larger than  $s \geq \phi n$  is exponentially decreasing in  $n$  for large enough  $\phi$ . The last remaining term in eq. (175) can instead be bounded by (see ref. [4])

$$\mathbb{P}(H^c \mid S \leq s) \leq \frac{1}{(3-\eta)^\ell \binom{n}{\ell}} \frac{1}{\text{poly}(n)} \quad (182)$$

so that the proof of Lemma 29 is now complete. □