Mixing and localisation in time-periodic quantum circuits

Tom Farshi

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics

Department of Computer Science

University College London

June 15, 2022

I, Tom Farshi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This thesis introduces and analyses a new model of time-periodic (Floquet) dynamics in a quantum spin systems. This model is implemented via a time-periodic quantum circuit with local Clifford gates. All the results of this thesis are rigorous mathematical proofs, which use tools and methods from quantum information science to study problems in many-body quantum systems and condensed-matter physics. This includes proofs of a form of dynamical mixing of Pauli operators in the case of local interactions, and conditions under which the evolution operator can resemble a random unitary. The scrambling time is of critical importance to these results, and in the case of non-local interactions, a slightly larger than logarithmic scrambling time is found. Also, the model analysed in this thesis has the peculiarity that it displays a strong form of localisation in one spatial dimension and the absence of localisation in two dimensions. There is no previously known model with these features, hence, this research is important to characterise the landscape of many-body quantum physics.

Impact Statement

With the possibility of quantum simulators [1, 2], and so the artificial synthesis of quantum matter without "unphysical" Hamiltonians, a deep and thorough understanding of the landscape of possible phenomena in many-body quantum systems is essential. In this thesis, I contribute to this through an exploration of time-periodic quantum circuits with Clifford gates, which are well-known from the stabiliser formalism in quantum computation. Time-periodic dynamics, also known as Floquet dynamics, are of increasing interest in the physics community. In particular, the results of this thesis are relevant and of interest to a variety of fields within academia: quantum information theory, many-body quantum systems, condensed matter physics, and mathematical physics (random matrix theory and percolation theory). Importantly, new tools to study time-periodic random quantum circuits are developed in this thesis.

Outside of academia, the results of this thesis may be of interest to those concerned with using the sampling of pseudo-random unitaries to demonstrate the computational advantage of quantum computers (for certain tasks). The most notable example of this is the recent experiment in reference [2], which measures the output of the generated quantum circuits using Pauli measurements.

Acknowledgements

Without a doubt, I would not be where I am now without the great support of many people. I don't express my gratitude anywhere near as much as I should, but I would like to try to make up for it now.

Most of all, I would like to thank my supervisor Lluis Masanes. His help and guidance has always been invaluable, and he has been an inspirational teacher and mentor. I have learned a great deal from him, and moreover learned how to think like a scientist, with clarity, and precision. For all of this, and much else, I am incredibly grateful - thank you so much! I would also like to thank Daniele Toniolo. Our discussions taught me so much, and always gave me a deeper understanding of physics. Truly, it was an honour to have worked with you both over the past few years.

Of course, I wouldn't be here now with the love and support of my family, and particularly my parents. Thank you all! I would also like to thank all my friends, for the joyous times we have spent together and the laughter we have shared. Finally, my thanks go to Linnet for her love, support, and patience.

I owe you all a great debt, more than I could ever return. Thank you all again for all that you have done!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of time-periodic (Floquet) dynamics has provided a deeper understanding of the rich landscape of observable behaviour in many-body quantum systems, and of non-equilibrium phase transitions. This is equally true of the study of quantum system with disorder, with examples of phenomena including: single-particle (Anderson), and many-body localisation. Although despite this, the importance and role of disorder in many-body quantum systems is still far from being conclusively resolved, which is particularly the case for the definition of quantum chaos.

In this thesis, I will investigate the role of disorder in time-periodic (Floquet) quantum circuits. In chapter 2 I will present results concerning the mixing of Pauli operators (when evolving according to the Heisenberg picture), and which investigate the question: how much do local and time-periodic dynamics appear to resemble a random unitary? In chapter 3 I will present results concerning localisation. The rest of this chapter is an overview and discussion of literature relevant to the results in this thesis. I have endeavoured to make this survey broad. In each section I will draw attention to how the discussed ideas and insights from the literature serve as motivations for the results I will present in subsequent chapters, and clearly indicate the specific chapter and section. Moreover within these chapters I will discuss further how the results presented relate to previous work in the field.

1.1 Quantum chaos and random matrix theory

The distinction between chaotic and integrable quantum dynamics [3] plays a central role in many areas of physics, for example in the study of equilibration [4], thermalisation [5], and related topics such as the eigenstate thermalisation hypothesis [6, 7], quantum scars [8], and the generalised Gibbs ensemble [9]. This distinction is also important in the characterisation of many-body localisation and in arguments concerning the black-hole information paradox, both of which will be discussed later in this chapter in sections 1.5 and 1.3 respectively. In this section, I will discuss the notions of ergodicity and chaos in quantum systems, which is particularly relevant to the results in chapter 2.

Despite thorough research, the precise definitions of quantum chaos and integrability are still being debated [10–12]. In the case of classical systems, there is a defining feature of chaos; the (exponential) divergence of trajectories in phase space under a perturbation. Unfortunately however, this definition does not extend to the case of quantum systems. For all (unitary) dynamics of a quantum system, the distance or distinguishability, as quantified for example by the trace distance (1-norm) or just the overlap, of two (different) quantum states of the Hilbert space remains constant for all times.

However, through the field of random matrix theory, it is well established that the dynamics of chaotic quantum systems share important features with (typical) random unitary matrices [13]. It is worth remarking that historically the study of random matrix theory was first used to model heavy nuclei and their spectral properties, and the field itself is a highly important and influential area of physics. The most notable signature of quantum chaos is the structure of the eigenvalues of the Hamiltonian. In particular, the distribution of the spacing between (consecutive) eigenvalues, $\Delta E = E_i - E_{i+1}$, follows a distribution of the form $P(\Delta E) \sim \Delta E e^{-\Delta E^2}$, which as ΔE tends to zero gives zero probability; this is known as eigenvalue repulsion [6, 14–16]. This feature of eigenvalue repulsion and the distribution of eigenvalue spacing has been identified as the same distribution as that of a typical random unitary matrix sampled uniformly (according to the Haar measure) from the unitary group; this is

known as the circular unitary ensemble. It should be noted that the exact form of this distribution of eigenvalue spacing does depend on the underlying symmetry of the system, and this fact is also reflected by the random matrices which must satisfy some additional symmetry. For example, the case of random unitary matrices corresponds to systems with no symmetry, but for systems which exhibit time reversal symmetry then the random matrices are restricted to random orthogonal matrices. Whereas, for integrable quantum systems the distribution of eigenvalue spacing does not exhibit repulsion, and instead the signature of integrability is a Poisson distribution, $P(\Delta E) \sim \exp(-\Delta E)$, for the distribution of eigenvalue spacing [17].

There are of course many other notable signatures of quantum chaos in physically relevant systems that are commonly identified with aspects of random unitaries. Some of these include: fast decay of out-of-time order correlators [18–20], entanglement spreading [21], operator entanglement [22], entanglement spectrum [23, 24], linear growth of the spectral form factor [25–27], and Loschmidt echo [28].

1.1.1 Relevance to results in thesis

The study of quantum chaos and of ergodicity in quantum many-body systems is of particular relevance to the results I will present in chapter 2. In chapter 2, I will explore a time-periodic random quantum circuit and discuss mathematical results concerning the mixing of Pauli operators and how the dynamics of the system appear to resemble that of a random unitary.

1.2 Quantum information theory and unitary designs

In the field of quantum information, there is a prominent study of randomness as a resource for a variety of quantum information processing tasks. This itself is in direct analogy with the study of randomised algorithms in classical computing. In some cases the use of randomness is essential to the performance of the algorithm, for example in cryptography, whereas in other cases the use of randomness is instead useful for improving the efficiency of the algorithm, for example finding the

minimum-cut of a given graph. Similarly, some quantum information processing tasks which utilise randomness include: data-hiding [29], quantum state decoupling [30], entanglement distillation [31], quantum error correction [32, 33], quantum process tomography [34], and randomised benchmarking [35]. In this section, I will outline exactly what is meant by using randomness as a resource, introducing the quantum information notion of a unitary design, which is central to this work, and discuss the literature surrounding this topic. However, I will not delve into a discussion of the algorithms which make use of randomness, and in particular the discussion will focus on the construction of unitary designs.

Early in the history of the field, it was realised that to implement a (Haar) random unitary (a random unitary sampled uniformly from the unitary group [36]) on n qubits using two qubit gates was not efficient, in particular the circuit depth required is (doubly) exponential in the number of qubits, n [37–40]. This itself poses a problem, since a quantum algorithm as part of its procedure can not make use of such an inefficient operation. The solution to this comes in the form of unitary designs. In essence the idea behind unitary designs is that rather than implement a (Haar) random unitary, it suffices to instead implement a unitary which resembles a (Haar) random unitary to some degree of accuracy, in particular it resembles the first k statistical moments. (As an aside, a related notion exists in classical computation, where instead the object of interest are random permutations [41, 42].) In more operational terms, a set of unitaries $\mathcal{U} \subset \mathrm{SU}(2^n)$ forms a k-design if, despite having access to k copies of a given unitary U, we cannot discriminate between the case where U is sampled from $\mathcal{U}$ or uniformly (Haar) from the entire unitary group $\mathrm{SU}(2^n)$. Mathematically, the definition of a unitary designs [43, 44] is as follows

Definition 1.1. Unitary k-design. A distribution of unitaries $\mathcal{U} \subset \mathrm{SU}(2^n)$ is a unitary k-design if the following holds

$$\mathbb{E}_{\mathcal{U}} U^{\otimes k} \rho U^{\dagger \otimes k} = \int_{SU(2^n)} dU U^{\otimes k} \rho U^{\dagger \otimes k} \,, \tag{1.1}$$

for all density operators ρ of the k-fold tensor product Hilbert space (dimension 2^{nk}).

We can also define an approximate version of unitary designs.

Definition 1.2. Approximate unitary k-design. A distribution of unitaries $\mathcal{U} \subset \mathrm{SU}(2^n)$ is an ε-approximate unitary k-design if for the twirling maps

$$\Xi_{\mathcal{U}}^{(k)}(\rho) = \mathbb{E}_{\mathcal{U}} U^{\otimes k} \rho U^{\dagger \otimes k} \,, \tag{1.2}$$

and

$$\Xi_{\mathrm{SU}(2^n)}^{(k)}(\rho) = \int_{SU(2^n)} dU U^{\otimes k} \rho U^{\dagger \otimes k} \,, \tag{1.3}$$

the following holds

$$\left|\left| \Xi_{\mathcal{U}}^{(k)} - \Xi_{\mathrm{SU}(2^n)}^{(k)} \right|\right|_{\diamond} \leq \varepsilon \,,$$

where ρ is any density operator of the total (tensor product) Hilbert space, and where $\varepsilon > 0$. Equivalently, a distribution of unitaries $\mathcal{U} \subset \mathrm{SU}(2^n)$ is an ε-approximate unitary k-design if

$$\begin{aligned} &= \sup_{d,\rho} \left|\left| \mathbb{E}_{\mathcal{U}} \left(U \otimes \mathbb{1}_d\right)^{\otimes k} \rho \left(U \otimes \mathbb{1}_d\right)^{\dagger \otimes k} - \int_{SU(2^n)} \left(U \otimes \mathbb{1}_d\right)^{\otimes k} \rho \left(U \otimes \mathbb{1}_d\right)^{\dagger \otimes k} dU \right|\right|_1 \\ &\leq \varepsilon \end{aligned} \tag{1.4}$$

Operationally, this approximate version is equivalent to allowing some error tolerance ε in the discrimination process. For an explanation of the diamond or operator norm used in the definition of approximate unitary k-designs consult the references [45, 46]. But to briefly summarise, when discriminating between unitaries (or quantum channels) it is possible to increase the distinguishability by entangling the probe with an arbitrary ancilla.

To clearly illustrate the concept of unitary design we consider the following example. We have a collection of n-qubits and we wish to apply a unitary which resembles a Haar random unitary up to the first moment (mean), so a 1-design. To do this, we apply to each of the n qubits independently a random single qubit Pauli operator $(I, \sigma_X, \sigma_Y, \sigma_Z)$ selected with uniform probability $(1/4)$. Unfortunately,

while this construction of a unitary 1-design is straight-forward and moreover easy to implement, it appears that for most practical purposes designs of order 2 and higher are required. For example, the quantum information processing tasks I have already mentioned require: data-hiding (2-design), quantum state decoupling (2-design), randomised benchmarking (2-design), and state discrimination (4-design). Finally, it should be noted that an (exact or approximate) unitary k-design is also an (exact or approximate) unitary $k-1$-design, which can be seen by considering the state ρ to be the identity on one of the k-fold tensor product Hilbert spaces.

In the rest of this section, the discussion will focus on the construction of unitary designs, both exact and approximate. I will first present and discuss results in the literature concerning discrete time dynamics and unitary designs. The discrete time dynamics are generated by a random quantum circuit: quantum circuits in which the gates are selected at random according to some distribution. Early work focussed initially on constructing both exact and approximate unitary 2-designs. This was demonstrated in numerous works. Initially work required highly structured random quantum circuits with specific choices of gate set [39, 47], but they did indeed demonstrate (efficient) algorithms for implementing both exact and approximate unitary 2-designs. The circuits required a number of gates polynomial in the number of qubits. Later work [44, 48, 49] improved upon this by showing an approximate 2-design can be implemented by applying in series a polynomial number of independent and uniformly selected two qubit random unitaries, where the pair of qubits acted upon is also selected independent and uniformly. It is worth noting some recent work [50] that studies a notion called anti-concentration, which is a weaker but related condition than a unitary 2-design. The results of this work demonstrate that anti-concentration occurs for depth only logarithmic in the number of qubits. Another recent result concerning approximate unitary 2-designs is found in reference [51], which notably uses techniques from condensed matter and statistical mechanics, demonstrates that a local time-dependent random quantum circuit forms an approximate 2-design in a time linear in the number of qubits.

As discussed, higher than unitary 2-designs are required for many practical

applications. I will now focus on results concerning k-designs for any arbitrary k, some of which are generalisations and extensions of references already discussed. In particular, it was shown in references [33, 52] that a one-dimensional system of n qubits in which a series of independent and uniformly selected two qubit random unitaries are applied to (random) pairs of neighbouring qubits, forms an approximate k-design when $O(n^2k^{10})$ gates are applied. Moreover, later work has focussed on studying higher spatial dimensions and shown a similar result [53].

In general, less is known of the case of exact unitary designs. One particular instance of exact unitary designs is well known from references [54, 55]. In these works, the results focus upon a sub-group of unitaries known as the Clifford group. The Clifford group is the group of unitaries which transform n-qubit Pauli operators to n-qubit Pauli operators. (The Clifford group is central to all of the results of this thesis, and hence I leave a detailed presentation until later, see appendix A.) It is found that applying to a (uniformly selected) random Clifford unitary to n qubits implements an exact 3-design, and it should be noted that any Clifford unitary on n qubits can be applied using a polynomial number of single and two qubit Clifford unitaries. Also, in these works it is also shown that the Clifford group can not form any higher order exact k-design. (Although, recent work has extended the analysis of Clifford group based unitary designs to show that approximate designs of any order k can be constructed by also including a polynomial (in k) number of non-Clifford two qubit unitaries [56].)

There has also been some work investigating the construction of unitary k-designs with continuous time dynamics, so how close the time evolution operator, given by the exponential of a Hamiltonian, resembles a (Haar) random unitary. However, it would be fair to say that in this instance the discrete-time case has received more attention. One of the main contributions in this direction can be found in reference [57]. In this work, a one dimensional spin chain of n qubits with dynamics generated by time-dependent local (nearest-neighbour) stochastic Hamiltonians (Brownian quantum circuits) is studied. It is found that in this case an approximate k-design is formed after a running time that is linear in n and polynomial

in k, for any arbitrary k. Another notable work is reference [58], that presents a Hamiltonian with periodically changing interactions that forms an approximate k-design after a minimum run time.

Finally, there are some results concerning unitary designs in which there are also some added conservation laws [59, 60]. The models considered are random quantum circuits, with the conservation laws ensured by imposing that all local random unitaries commute with a fixed operator, Q. It is found that these random circuits also generate approximate k-designs in the operator space orthogonal to Q, with k increasing as time passes.

1.2.1 Relevance to results in thesis

The results I have presented from the literature concerning unitary designs focus on time-dependent dynamics. This serves as the primary motivation for the results in chapter 2 in particular section 2.1. In this section, I will present a model of a time-periodic random quantum circuit with local interactions and discuss mathematical results connected with the notion of approximate unitary designs. Additionally the Clifford group, which was discussed in the context of exact 3-designs, is central to all of the results of this thesis. (To reiterate, a detailed discussion of the Clifford group is given in appendix A.) It is also worth noting that many of the mathematical results and notation introduced in section 2.1 are also made use of in later sections (section 2.2).

1.3 Quantum information theory and black holes

The application of ideas and tools from quantum information to the study of highly chaotic quantum system, in particular investigating the physics of black holes, is particularly striking. The research along these lines has provided a useful new perspective and given further physical insight, while also presenting new research questions to investigate. In this section I will mainly focus on the results of Hayden and Preskill in the reference [61], and work which builds on this further.

Broadly speaking the question of interest is: using the tools and techniques of quantum information theory what can we learn about what happens to the (quantum)

information deposited and contained within a black hole? In particular in the work of Hayden and Preskill, the question is: using the perspective of quantum information theory what can we learn about what is the fate of the quantum information contained in a fixed number of qubits which are absorbed by a black hole? A toy model for the dynamics of a black hole is to use a (Haar) random unitary to model the dynamics, which is also a typical model used to study highly chaotic quantum dynamics. (This perspective was also taken earlier in the references [62, 63], in which the entanglement between a black hole and its environment was studied.) The results using this model are striking. Consider N qubits, which are maximally entangled with another N qubits acting as a reference system. These N qubits are then absorbed into the black hole, modelled as described as a (Haar) random unitary. If the black hole has already aged past half of its life-time then after the black hole has radiated away a further N qubits, then the total radiation from the black hole is maximally entangled with the N qubit reference system. (If the black hole is not aged past half its life time, then all that is required is to wait for this to occur.) In other words, the quantum information thrown into the black hole is almost immediately expelled and encoded within the total radiation from the black hole. (Subsequent work has found a decoding algorithm [64], which curiously has connections with the out of time order correlator studied extensively in the context of quantum chaos.)

In this work, the discussion then focuses on the implementation of this random unitary via a (random) quantum circuit, noting that a unitary 2-design is sufficient. Recalling from the previous section (1.2), it has been shown that implementing a 2-design on n qubits can be done with a (time dependent) random quantum circuit of depth polynomial in $\log n$.

This leads to what is known as the fast scrambling conjecture [65, 66]. This conjecture states the following that: the most rapid scramblers take a time logarithmic in the number of degrees of freedom, matrix quantum mechanics saturate the bound, and black holes are the fastest scramblers in nature. This conjecture, from the quantum information perspective, provides additional motivation for the continued study of unitary designs and random quantum circuits. Namely that one research

objective is to find/demonstrate a physically well-motivated random quantum circuit which implements a unitary design which has a circuit depth that is logarithmic in the number of qubits.

Progress in this direction has already been made by studying what is known as decoupling [30, 67]. In these works, it has been shown that random quantum circuits with non-local few-body (time-dependent) dynamics can in poly-logarithmic depth implement (approximately) the same process described as for the model black hole. Additionally, it is worth recalling (from section 1.2) that a related notion known as anti-concentration, which is a weaker condition than a 2-design, has been shown to occur in logarithmic depth for local (time-dependent) random quantum circuits [50]. Finally, it is briefly worth noting related studies of the growth of circuit complexity [68–70],

1.3.1 Relevance to results in thesis

The results that will be discussed in chapter 2 section 2.2 draw their primary motivation from the ideas presented in this section. That is to say, the high chaotic (quantum) dynamics of a black hole are modelled as a (Haar) random unitary. However, physically one would expect the dynamics to be time-independent. Hence, we study a time-periodic model, which share important features with time independent dynamics, and investigate the limit of the scrambling speed, by allowing interactions to be non-local although still few-body.

1.4 Single-particle (Anderson) localisation

In this section, I will discuss the phenomenon of Anderson (single-particle) localisation. This phenomenon was first studied by Anderson in 1958 [71], and was initially proposed as a model to investigate spin diffusion in materials with impurites. Anderson localisation itself is now extremely well studied [72–74], and a great deal is understood, with many elaborations, refinements, and extensions of the original model. In this section, I will give a general overview, which will provide context and motivation for the following section on many-body localisation (section 1.5) and also for our results in chapter 3, which concerns localisation in time-periodic systems.

Anderson localisation has been studied in many spatial dimensions. I will present the one-dimensional model, from which analogous higher spatial dimensional models can be inferred. Although I will still discuss and compare the phenomenology in higher dimensions, in particular since the differences in the observable behaviour between spatial dimensions is relevant and a motivation both for many-body localisation (section 1.5) and the results in chapter 3 (section 3.2).

In one spatial dimension the model consists of L sites arranged on a line, with open boundary conditions, so the lattice $\mathbb{Z}$. Using the language of second quantisation, on this chain of L sites there are (non-interacting) particles which can occupy a site or hop from one site to one of the two adjacent sites. Each site has a particle occupation energy of ε_i, and the hopping amplitude between sites is λ . Therefore, the Hamiltonian of this model is given by

$$\hat{H} = \sum_{i=1}^{L} \varepsilon_i c_i^\dagger c_i + \lambda \sum_{i=1}^{L} c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1} \ , \tag{1.5}$$

where $c_i^\dagger, c_i$ are the creation/annihilation operators for a single particle at site i, ε_i is the on site particle occupation energy. Crucially for the Anderson model, in order to model disorder, the on-site energy ε_i terms are taken to be independent random variables. Originally, each ε_i is taken to be uniformly distributed in some interval $[-W, W]$, with each ε_i identically and independently distributed (i.i.d.). In this model the particles are non-interacting, and hence the model is also referred to as single-particle localisation. The question now is: what effect does this disorder (in the on-site energy) have on the dynamics and observed behaviour of the system?

Firstly, we shall (briefly) comment on the two different extreme cases: zero disorder ($W = 0$) and infinite disorder ($W \to \infty$). In the case of zero disorder, so $W = 0$, the spin chain exhibits (discrete) translation invariance, and hence the eigenvectors are plane waves of the (discrete space) spin chain. In the case of infinite disorder, ($W \to \infty$), the hopping between sites is in effect absent and hence we have particles that are confined to a single site for all times. These two extreme cases are not particularly of interest, and naturally the main regime of focus and

interesting physics is between these two extremes (so, finite $W > 0$). Interestingly, for any non-zero disorder strength ($W > 0$) the system displays what is now known as Anderson or single-particle localisation; a particle remains fixed (localised) on its initial site or the surrounding sites for all times. By this we mean that, for a particle initially (at $t = 0$) located at site i_0 the probability that the particle is found on any site i, at any later time is given by

$$P(\text{particle at } i \text{ at time } t | \text{particle initially at } i_0) \propto \exp\left\{ -\frac{|i - i_0|}{l_{\text{loc}}} \right\}, \tag{1.6}$$

where l_{loc} is known as the localisation length scale, which depends on the disorder strength W (and its ratio to the hopping amplitude λ). Crucially, this exponential localisation occurs for all non-zero disorder strengths, $W > 0$ (although the associated length scale of the localisation will vary). Anderson localisation therefore is a classic example of disorder causing observable changes in the behaviour of system.

The landscape of the observed behaviour of this model becomes more rich and interesting (and the mathematics more complex) when considered in higher spatial dimensions. In two spatial dimensions, it is strongly believed that for any non-zero disorder ($W > 0$) the system again displays localisation (in the manner described in one dimension) [75–78]. The reason for lack of certainty is twofold: some papers claim these results are not exact for very small disorder strengths [79], although they are certainly corroborated by numerical results, while other works claim that the localisation length scale associated with very small disorder can be extremely large, and hence potentially difficult to observe in experiments [75]. In three spatial dimensions, the situation is different once again. It is well established that localisation (in the manner described in one dimension) does indeed occur, however this is after the disorder exceeds some minimum threshold, so $W > W_{\text{critical}}$. In fact generally speaking, it has been shown that in any number of spatial dimensions greater than or equal to three there exists a finite (and non-zero) threshold for the disorder strength such that the system exhibits localisation [80]. Therefore, it should be emphasised that the number of spatial dimensions is an important consideration when studying localisation phenomenon. Below this threshold, so for small disorder,

a particle is not localised and so the system then has a phase transition from a delocalised to localised phase [81].

Finally, I will briefly remark on some of the refinements and elaborations of the model. In the particular model present, the hopping term is between neighbouring sites of the lattice only. The more general case of extending the model to include more long-range hopping has been studied. In this instance long range can be for example: hopping amplitudes which decay exponentially [82] or according to some power law [83], or including hopping to next-nearest neighbours (or further) also [84–86]. Additionally, the case of more general or varied distributions for the on-site energy has been investigated, for example where the distribution is Gaussian [87], and moreover the situation where there is some correlation between the distributions for each on-site energy [88–91], so they are no longer independent. This by no means represents the entirety of the elaborations and further studies of Anderson localisation, although certainly gives some introduction to the landscape, however the articles in references [75, 78, 79, 92–94] should provide a good place to continue investigations. These references are also good resources to understand the mathematical and numerical techniques commonly used, neither of which I have discussed in detail, preferring instead to give a broad overview and introduction to the phenomenon of Anderson localisation. Finally, I should point out that I have not commented specifically on results concerning the classification of the spectrum of the Hamiltonian in terms of its pure-point/continuous and its interpretation in relation to the observation of localisation, however this is discussed within many of the references themselves.

1.4.1 Relevance to results in thesis

Of course one reason for presenting and discussing the phenomenon of single-particle localisation is to introduce the discussion in the next section (section 1.5) of many-body localisation. Importantly however, single-particle localisation is also pertinent to our results (chapter 3) and their interpretation. In particular in chapter 3 section 3.1 I will discuss results of a time-periodic model, which in one regime of dynamics exhibits a localisation effect which is in some ways reminiscent of Anderson

localisation. More over in 3 section 3.2, in part motivated by the observation of the importance of the number of spatial dimensions to single-particle localisation, I will present results from investigating an analogous time-periodic model in higher (two) spatial dimensions, which in no way exhibits the same localisation effect.

1.5 Many-body localisation

In this section, I will discuss the interesting topic of many-body localisation and give a general overview, which will provide context and motivation for the results in chapter 3, which concerns localisation in time-periodic systems. The starting point for the study of many-body localisation is, as already hinted at previously, the study of single-particle localisation (section 1.4). As discussed already, in Anderson localisation there is no form of interaction between the particles in the model, and in essence it is a single-particle system. Naturally, the extension is to consider an analogous model in which the particles interact in some manner, and hence the investigation is of a many-body quantum system. Immediately, one might expect that once interactions are included the landscape of observable behaviour becomes more rich and varied, and of course requires more techniques and tools to gain insight into what can be a more complicated system.

In one spatial dimension, one standard model used to study many-body localisation is a spin chain with dynamics given by the Heisenberg XXX model with onsite disorder, so

$$\hat{H} = \sum_{i=1}^{L} \varepsilon_i \sigma_i^Z + J \sum_{i=1}^{L} \left(\sigma_i^X \sigma_{i+1}^X + \sigma_i^Y \sigma_{i+1}^Y + \sigma_i^Z \sigma_{i+1}^Z \right) , \tag{1.7}$$

where ε_i is the onsite disorder which is uniformly distributed in the region $[-W, W]$ independently and identically for all sites, and J the interaction strength (which is commonly taken to be one), and where $\sigma_i^{X,Y,Z}$ are the Pauli X, Y, Z matrices for the site i [95–98]. Without the onsite disorder ($\forall i : \ \varepsilon_i = \varepsilon$), this model is a one dimension spin chain with (discrete) translation invariance, and is a well known example of an integrable model exactly solved using the Bethe ansatz [99, 100]. (There are many other models considered to study many-body localisation. For example, there is the

note worthy reference [101], and also references [79, 102–104].)

It is now well established, both numerically [95, 105, 106] and analytically [107] (under some reasonable assumption about the spectrum), that there exists a threshold disorder strength such that spin chain is not ergodic and instead exhibits many-body localisation. What is it that I mean when I say a system is many-body localised? Generally speaking the key aspect of many-body localisation is the absence of thermalisation, by which I mean that some aspects of an initial spin configuration remain constant throughout its time evolution for arbitrarily long times. Moreover, the distribution of eigenvalue spacing does not exhibit repulsion and instead follows a Poisson distribution. The many-body localised system can be understood as an emergent integrable system, and in this phase the Hamiltonian can be rewritten with local integrals of motion in a classical form

$$\hat{H} = \sum_i \varepsilon'_i \tau_i^Z + \sum_{i,j} J'_{i,j} \tau_i^Z \tau_j^Z + \sum_{i,j,k} J'_{i,j,k} \tau_i^Z \tau_j^Z \tau_k^Z + \dots \,, \tag{1.8}$$

where τ_i^Z is a quasi-local operator, which is a dressed version of the Pauli matrix σ_i^Z, such that $\tau_i^Z = U^\dagger \sigma_i^Z U$ with U a unitary acting in a (small) region centred upon site i of the spin chain, and where the onsite energy ε'_i is a modified onsite energy ε_i, and where the coupling terms between sites decay exponentially with distance, for example $J'_{i,j} \propto \exp(-|i-j|)$ [103, 107–109]. This conjectured effective model captures the phenomenology of the many-body localised phase, and gives a great deal of insight in to the physics.

There are many other well established features of many-body localisation. The dynamics and structure of the entanglement is a key signature of many-body localisation. When considering some region, for example the left-half, of the spin chain then it has been well established that for an initially product state the entanglement entropy grows logarithmically with time [110–112]. This slow growth of entanglement also differentiates many-body from single particle localisation. Another key signature concerns the time evolution of initially local operators; operators which are only supported on one site of the spin chain. It has been shown that under the

many-body localisation dynamics the support of initially local operators grows only logarithmically with time [113–115]. Both the slow growth of entanglement and slow growth of local operators differentiates many-body from single particle localisation.

Additionally, below this threshold disorder strength for the onset of many-body localisation, it is also well established that the many-body quantum system is ergodic [106, 116], which for example is characterised by the observation of eigenvalue repulsion. Naturally then there is also much interest and research focussing on the (non-equilibrium) phase transition between the ergodic and many-body localised phases [95, 117, 118], particularly into the entanglement structure of the eigenstates of the many-body quantum system (volume/area law transitions). However, it is fair to say that there remains to be a full understanding of the ergodic to many-body localised phase transition.

Finally, the discussion of many-body localisation in this section so far has focussed on one spatial dimension. However in the case of single particle localisation, as discussed in section 1.4, the number of spatial dimensions is an important factor determining the threshold disorder strength for localisation. Therefore, in the case of many-body localisation there is also much interest and work studying higher dimensional models [97, 119]. But, it is not yet settled either numerically or analytically whether there exists a stable many-body localised phase in spatial dimensions greater than one.

1.5.1 Relevance to results in thesis

The study of many-body localised systems, and the broad overview given in this section, is of particular relevance to the results found in chapter 3, both in terms of motivation and context. In particular in chapter 3 section 3.1 I will discuss results of a time-periodic model, which is an interacting many-body system with disorder, and which in one regime of dynamics exhibits a localisation effect which is in some ways reminiscent of many-body localisation. But, as already noted, this localisation effect is also in some ways reminiscent of single-particle localisation, and so seems to challenge the existing classification . More over in chapter 3 section 3.2, in part motivated by the fact that there is yet to be a detailed understanding or consensus

regarding many-body localisation in higher (than one) spatial dimensions, I will present results from investigating an analogous time-periodic model in two spatial dimensions, which exhibits in a strong (and precise) sense an absence of localisation.

1.6 General outline of thesis

Throughout this chapter, for all the topics covered, I have mentioned how they directly motivate and relate to the results in this thesis. I will now briefly comment on the structure and format of this thesis. In the second chapter, I will present and discuss results concerning the scrambling of quantum information and the mixing of Pauli operators in random quantum circuits. In the third chapter, I will present and discuss results concerning the localisation of quantum information in random quantum circuits. In both chapters, I will first clearly describe the models studied and then present and broadly explain the (mathematical) results obtained, initially with an emphasis on the more general side and the meaning of the results. Finally, I will then delve into a detailed discussion and derivation of the mathematical results. Then, in the final section I will discuss the results again, give possibilities for future work, and end with reflections and conclusions. In this thesis there are several appendices. The material in each appendix is additional detail that is not essential for understanding the results in the thesis, but does give a more thorough understanding, and hence is included for these reasons.

The work contained in the second and third chapters are from research projects undertaken during my doctoral studies, and I have significantly contributed to all the results. For the instances in which I cannot claim to have contributed significantly to the particular results I shall make this explicit and clearly state the attribution.

Chapter 2

Scrambling and Mixing

In the introduction of this thesis (chapter 1), I discussed the importance of random unitaries both in the study of quantum chaos and within the field of quantum information theory - for applications in a variety of quantum information processing tasks and as a toy model for black holes. In this chapter, I will present results investigating the question of scrambling and mixing (of Pauli operators) in time-periodic random quantum circuits. In the first section of this chapter (section 2.1), I will present a model of a time-periodic random quantum circuit with local interactions. I will show rigorous results that demonstrate that the dynamics can display mixing (of Pauli operators), and prove that under certain conditions the dynamics appear to resemble that of a (Haar) random unitary, with precise definitions of these statements. In the second section of this chapter (section 2.2), I will investigate an analogous time-periodic random quantum circuit model with interactions that are still few-body but are no longer local. I will show results which demonstrate that the circuit depth corresponding to the scrambling time is (at-most) poly-logarithmic in the size of the system.

In both of the sections the structure is as follows: firstly I will give a general overview of the results including a brief discussion, then I will give detailed derivations of these results, and finally I will more fully discuss the results.

2.1 Time periodic local dynamics

As discussed in the introduction (chapter 1 section 1.1), the dynamics of chaotic quantum systems share important features with random unitaries, with signatures of quantum chaos identified with aspects of random unitaries. In this section, I will take a more operational perspective and investigate instead when the time evolution operator is physically indistinguishable from a random unitary. This notion of physically indistinguishable is captured by the quantum information notion of unitary design, which was discussed in the introduction (chapter 1 section 1.2). In this section, I will present and analyse a Floquet model with disorder, characterised by a family of local, time-periodic, random quantum circuits in one spatial dimension. I will prove rigorous results demonstrating that under certain conditions: the time evolution of Pauli operators can display a form of mixing (which will be defined precisely), which is given in results 2.1, 2.2, and 2.3, and that the time evolution operator cannot be distinguished from a (Haar) random unitary (when measurements are restricted to Pauli operators only), which is given in result 2.4. In other words in this section, I will investigate the question: how much do local, disordered, time-periodic dynamics resemble a random unitary?

Finally, it is worth mentioning that the work in this section has been published as a pre-print [120].

2.1.1 Overview

Firstly, I will present the model analysed in this section. The model is a one-dimensional spin chain composed of L sites, where each site contains N qubits or modes, and with periodic boundary conditions. The dynamics of the model are time-periodic, with the first dynamical period consisting of two half-steps. In the first half-step each even site interacts with its right neighbour with a random Clifford unitary and in the second half-step each odd site interacts with its right neighbour via a random Clifford unitary. As mentioned in the introduction (chapter 1 section 1.2), the Clifford group is the group of unitaries which transform N-qubit Pauli operators to N-qubit Pauli operators. In appendix A there is a detailed description of the Clifford group [121–123]. Each of the L Clifford unitaries are independent and

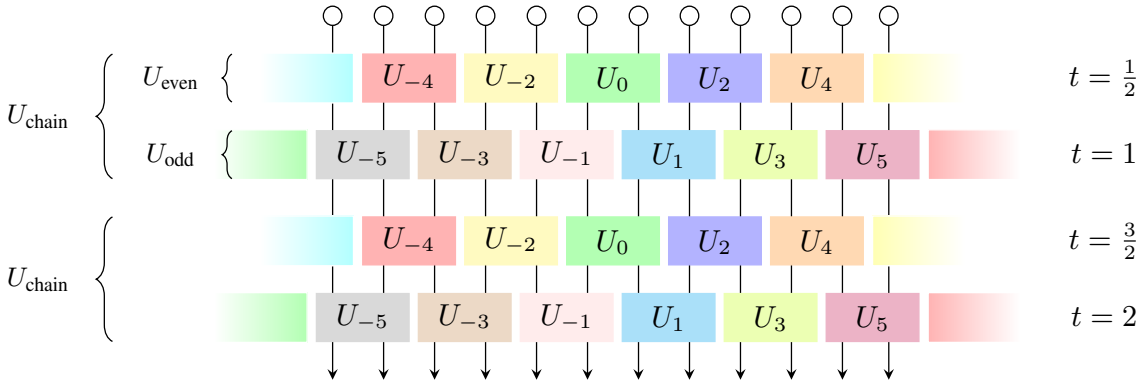

Figure 2.1: Time-periodic local dynamics. This figure illustrates the physical model analysed in this work. The circles on the top represent lattice sites, each of which consists of N qubits. Each coloured block represents a two-site unitary, with different colours representing that the unitaries are drawn independently and uniformly from the Clifford group, and hence there is spatial disorder. After the first two half time-steps, the dynamics repeats itself.

sampled uniformly from the $2N$-qubit Clifford group. Every subsequent period of the dynamics are repetitions of the first period; this is illustrated in figure 2.1. Denoting the (random) Clifford unitary which acts on sites x and $x+1$ by U_x (recalling that there are periodic boundary conditions, so that modulo L), then the evolution operator after an *integer* time t is

$$\begin{aligned} W(t) &= \left[(U_1 \otimes U_3 \otimes \cdots \otimes U_{L-1})(U_0 \otimes U_2 \otimes \cdots \otimes U_{L-2})\right]^t \\ &= (U_{\text{odd}} U_{\text{even}})^t = (U_{\text{chain}})^t \ , \end{aligned} \tag{2.1}$$

and after a *half-integer* time t is

$$W(t) = U_{\text{even}} \, (U_{\text{chain}})^{t-1/2} \ . \tag{2.2}$$

This evolution operator can also be generated by a time-periodic Hamiltonian $H(t)$ with nearest-neighbour interactions

$$W(t) = \mathcal{T} e^{-\mathrm{i} \int_0^t d\tau H(\tau)} \ , \tag{2.3}$$

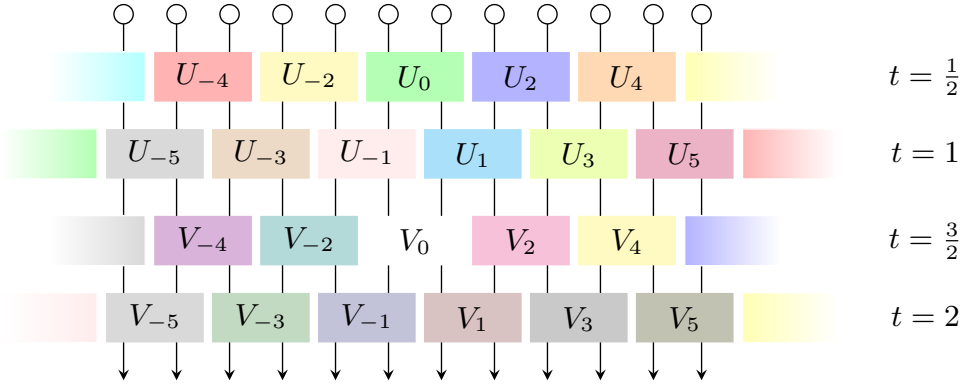

Figure 2.2: Time-dependent local dynamics. In this figure, in contrast to figure 2.1, the pictured circuit is time-dependent and not periodic in time. So, the Clifford unitaries in every time-step are sampled independently.

where $\mathcal{T}$ is the time-ordering operator. This type of dynamics is called Floquet [10, 21, 22, 26, 27, 124, 125] or quantum cellular automaton [126, 127] with disorder. It is worth emphasising that this time-periodic model is very different from the far more studied case of time-dependent random quantum circuits, an illustration of which is given in figure 2.2. Since one would expect the dynamics of physical systems to be time-independent, the case of time-periodic circuits is more relevant to physics, albeit more difficult to analyse.

One essential tool for understanding and analysing the dynamics of this time-periodic circuit is the phase space representation of the Clifford group. As I have already noted, in appendix A there is a detailed exposition of the Clifford group and its phase space representation, but I will now give a summary of this phase space representation, so that the results and following rigorous proofs can be followed. An N-qubit Pauli operator is the (N-fold) tensor product of Pauli sigma matrices, including the one-qubit identity, multiplied by a global phase $\lambda \in \{1, \mathrm{i}, -1, -\mathrm{i}\}$. Ignoring the global phase, each N-qubit Pauli operator can be represented by a binary vector $\mathbf{u} = (q_1, p_1, q_2, p_2, \ldots, q_n, p_n) \in \{0,1\}^{2N}$ as

$$\sigma_{\mathbf{u}} = \bigotimes_{i=1}^{n} (\sigma_x^{q_i} \sigma_z^{p_i}) \,. \tag{2.4}$$

By ignoring the global phase λ, the product of Pauli operators can be written as

$\sigma_{\mathbf{u}}\sigma_{\mathbf{u}'} = \lambda\,\sigma_{\mathbf{u}+\mathbf{u}'}$, where addition in the vector space $\mathbb{Z}_2^{2n}$ is modulo 2. The N-qubit Clifford group $\mathcal{C}_N \subseteq \mathrm{SU}(2^N)$ is the set of unitaries U which map each Pauli operator onto another Pauli operator $U\sigma_{\mathbf{u}}U^\dagger = \lambda\,\sigma_{\mathbf{u}'}$. Each Clifford unitary U can be represented by a $2N \times 2N$ symplectic matrix S with entries in $\mathbb{Z}_2$ such that its action on Pauli operators can be calculated in phase space

$$U\sigma_{\mathbf{u}}U^\dagger = \lambda\,\sigma_{S\mathbf{u}}\ , \tag{2.5}$$

where the matrix product $S\mathbf{u}$ is defined modulo 2. Due to the analogy with quasi free-bosons, in which dynamics are also linear and symplectic, the binary vectors $\mathbf{u} \in \{0,1\}^{2N}$ are known as the phase space representation of the N-qubit Pauli operator $\sigma_{\mathbf{u}}$. Furthermore, it is worth noting that the definition of the group of N-qubit Pauli operators can be understood as the discrete analogue of the Weyl group, which are also known in quantum optics as displacement operators.

Hence, the time-periodic model can be reinterpreted using this phase space description. The corresponding phase space of the L-site lattice with periodic boundary conditions can be written as

$$\mathcal{V}_{\text{chain}} = \bigoplus_{x\in\mathbb{Z}_L} \mathcal{V}_x\ , \tag{2.6}$$

where $\mathcal{V}_x \cong \mathbb{Z}_2^{2N}$ is the phase space of site $x \in \mathbb{Z}_L$. A local Pauli operator $\sigma_{\mathbf{u}}$ at site x, so non-identity only on site x, is represented by a phase space vector contained in the corresponding subspace $\mathbf{u} \in \mathcal{V}_x \subseteq \mathcal{V}_{\text{chain}}$, with the zero vector corresponding to the identity operator. The time evolution operator $W(t)$, defined in equations 2.1 and 2.2, has a corresponding phase space representation $S(t)$, which is a ($\mathbb{Z}_2$) symplectic matrix acting on the phase space $\mathcal{V}_{\text{chain}}$.

The time evolution operator $W(t)$ maps each Pauli operator $\sigma_{\mathbf{u}}$ to another Pauli operator via $\sigma_{\mathbf{u}'} = \lambda W(t)^\dagger \sigma_{\mathbf{u}} W(t)$. Since the Clifford unitaries $U_0, U_1, \ldots, U_{L-1}$ which define $W(t)$ are random, the mapping of Pauli operators $\mathbf{u} \to \mathbf{u}'$ is probabilistic,

and in particular the probability that $\mathbf{u}$ after a time t evolves to $\mathbf{u}'$ is

$$P_t(\mathbf{u}'|\mathbf{u}) = \mathop{\mathbb{E}}_{\{U_x\}} \left| 2^{-NL} \mathrm{tr}\Big(\sigma_{\mathbf{u}'} W(t) \sigma_{\mathbf{u}} W(t)^\dagger \Big) \right| , \tag{2.7}$$

where the orthogonality of Pauli operators $\mathrm{tr}(\sigma_{\mathbf{u}'}\sigma_{\mathbf{u}}) = 2^{NL}\delta_{\mathbf{u}'\mathbf{u}}$ is used. Since the dynamics in this model are local (as illustrated in figure 2.1), then only operators which are inside the light-cone of the initial operator $\sigma_{\mathbf{u}}$ have non zero probability. For example, if the initial operator $\sigma_{\mathbf{u}}$ is non-identity only at the origin ($x = 0$), then after a time t, the evolved operator $\sigma_{\mathbf{u}'}$ can only be non-identity inside the light cone $-2t+1 \leq x \leq 2t$. Therefore, the corresponding phase space vector $\mathbf{u}'$ is non-identity only within the light-cone with subspace given by

$$\mathbf{u}' \in \bigoplus_{x \in [-2t+1, 2t]} \mathcal{V}_x \, . \tag{2.8}$$

The scrambling time, by which I mean the minimum time t for which the light-cone of any initially local operator become the whole spin chain, is $t_{\mathrm{scr}} = L/4$. The scrambling time is a crucial quantity for all of the results in this section

Throughout this section, I will use the following definition for the mixing of Pauli operators: a dynamical system is Pauli mixing if any initial (non-identity) Pauli operator P through its time evolution $W^\dagger(t)PW(t)$ is transformed to any (non-identity) Pauli operator with uniform probability. It is worth noting that since the dynamics of the time-periodic model considered in this section is generated by Clifford unitaries, an initial Pauli operator remains a Pauli operator for all times. In the following rigorous result, I will precisely state how an initially local Pauli operator exhibits an approximate form of Pauli mixing. That is to say, the probability distribution for the time evolution of a Pauli operator, $P_t(\mathbf{u}'|\mathbf{u})$ (equation 2.7) is approximately uniformly distribution inside the light-cone.

Result 2.1 (Approximate Pauli mixing)**.** If the initial Pauli operator $\sigma_{\mathbf{u}}$ is only non-identity at site $x = 0$ then the probability distribution for its evolution $\sigma_{\mathbf{u}'}$, $P_t(\mathbf{u}'|\mathbf{u})$

(equation 2.7), is close to uniform inside the light cone

$$\sum_{\mathbf{u}'} \left| P_t(\mathbf{u}'|\mathbf{u}) - Q_t(\mathbf{u}') \right| \;\leq\; 130 \times t^2\, 2^{-N} \,, \tag{2.9}$$

for any integer or half-integer time $t \in [1/2, 2t_{\text{scr}}]$, where $Q_t(\mathbf{u}')$ denotes the uniform distribution over all non-zero vectors $\mathbf{u}'$ within the light-cone (equation 2.8), and where it is understood that after the scrambling time ($t \geq t_{\text{scr}}$) the light-cone is the total phase space of the spin chain, $\mathcal{V}_{\text{chain}}$. An analogous statement holds for any other initial location $x \neq 0$.

This result, as well as all other results in this section, is meaningful in the limit of large N ($N \gg \log L$). I will investigate the opposite regime, where $N \ll \log L$, in chapter 3 section 3.1. In this result, the mixing of Pauli operators appears to be weaker as time progresses, which is primarily due to temporal correlations and dynamical recurrences. These dynamical recurrences occur periodically due to the fact that the time-evolution operator is a Clifford unitary, and since the Clifford group is finite, there exists a recurrence time such that $W(t) = \mathbb{1}$. Although, the time required for these recurrences is far larger than twice the scrambling time, $2t_{\text{scr}}$, which is the time after which this result (and others in this section) does not hold. This is in contrast to the observed behaviour in time-dependent quantum circuits, as discussed in chapter 1 section 1.2, in which the error or difference capturing the proximity to uniformity decreases with time.

The approximate Pauli mixing result (2.1) applies to local operators only. In the next result, I will present a more general result which applies to a large class of non-local initial operators. Unfortunately, I have only been able to analyse initial operators which are non-identity inside a (continuous) region $\mathcal{R} = \{1, \ldots, L_{\text{s}}\} \subset \mathbb{Z}_L$ only.

Result 2.2. Consider an initial vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$ which is non-identity only on all lattice sites $x \in \mathbb{Z}_L$. Denote by $\mathbf{u}^t_{\mathcal{R}}$ the projection onto the subspace $\mathcal{V}_{\mathcal{R}} = \bigoplus_{x \in \mathcal{R}} \mathcal{V}_x$ of the evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ in the region $\mathcal{R} = \{1, \ldots, L_{\text{s}}\}$ with even size L_{s}.

When averaging over circuits, this projection is approximately uniform,

$$\sum_{\mathbf{v}\in\mathcal{V}_{\mathcal{R}}}\left|\text{prob}\{\mathbf{v}=\mathbf{u}^t_{\mathcal{R}}\}-\frac{1}{|\mathcal{V}_{\mathcal{R}}|}\right|\leq 34\times t\,2^{c|\mathcal{R}|-N}\,, \tag{2.10}$$

at times $t\in[1,\frac{L-|\mathcal{R}|}{4}]$, where we define the constant $c=\log_2\sqrt{3}$.

A stronger set of results can be proven for the case of half-integer times. This is due to the fact that the evolution operator of the model $W(t)$ at half-integer time has an extra symmetry. I will discuss this feature in detail in due course. Hence, the following results apply more generally to any initial Pauli operator.

Result 2.3. Let $\sigma_{\mathbf{u}'}=\lambda W(t)\sigma_{\mathbf{u}}W(t)^\dagger$ be the evolution of any initial Pauli operator $\sigma_{\mathbf{u}}\neq\mathbb{1}$. At any half-integer time t larger than the scrambling time, in the interval $t\in[t_{\text{scr}},2t_{\text{scr}}]$ the probability distribution (equation 2.7) for the evolved operator $\sigma_{\mathbf{u}'}$ is close to uniform

$$\sum_{\mathbf{u}'}\left|P_t(\mathbf{u}'|\mathbf{u})-Q_t(\mathbf{u}')\right|\;\leq\;33\times tL2^{-N}\,. \tag{2.11}$$

It is worth noting that the fact that mixing is more prominent at half-integer times is not restricted to Clifford dynamics, since it applies to a large class of periodic random quantum circuit or Floquet dynamics with disorder. Specifically, this fact holds for any quantum circuit in which the two-site interactions U_x also include single site random gates that are 1-designs.

As a consequence of the previous result, it is possible to prove that at half-integer times a Haar-random unitary is hard to distinguish from the time evolution operator $W(t)$, when the only measurements available to distinguish between the two are Pauli operators. More precisely, imagine that one is given a unitary transformation V which has been sampled from either the set of evolution operators $\{W(t)\}$ or the full unitary group $\text{SU}(2^{NL})$. The task is to choose the optimal state ρ, process the state according to the given transformation $\rho\mapsto V\rho V^\dagger$, and then measure the transformed state with a Pauli operator $\sigma_{\mathbf{u}}$, and finally use this measurement to guess whether the sampled V belongs to the set of evolution operators $\{W(t)\}$ or the full

unitary group $\mathrm{SU}(2^{NL})$. In order to enhance this discrimination procedure, two uses of the transformation V are allowed, which therefore allows one to input half of an entangled state ρ into each of the transformations. In the following result, it is shown that in the large-N limit the optimal guessing probability to distinguish the transformations is almost as good as a random guess.

Result 2.4. Discriminating between two copies of $W(t)$ and two copies of a Haar-random unitary can be done with success probability

$$\begin{aligned} p_{\text{guess}} &= \frac{1}{2} + \frac{1}{4} \max_{\rho,\mathbf{u},\mathbf{v}} \operatorname{tr}\left(\sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{v}} \left[\mathop{\mathbb{E}}_{W(t)} W(t)^{\otimes 2} \rho\, W(t)^{\otimes 2\dagger} - \int_{\mathrm{SU}(d)} dU\, U^{\otimes 2} \rho\, U^{\otimes 2\dagger} \right] \right) \\ &\leq \frac{1}{2} + 8tL2^{-N}\,, \end{aligned} \tag{2.12}$$

provided measurements are restricted to Pauli operators, for (half-integer) times $t \in [t_{\text{scr}}, 2t_{\text{scr}}]$.

If in this result the measurements were not restricted to Pauli operators, then $W(t)$ would be an $(8tL2^{-N})$-approximate unitary 2-design. Hence, the indistinguishability quantified in result 2.4 is a weaker variant notion of approximate unitary 2-design. Moreover, I should point out that the precise definition of approximate 2-design includes the use of an ancillary system in the discrimination process [47], which has not been included in result 2.4 as it does not provide any advantage.

2.1.2 Details and derivations

In this section, I will give detailed proofs for the stated results. Since the series of rigorous results which culminate in the proofs for the stated results are quite involved and require many stages, I will first give a broad outline of the general approach and methods used, and in this way provide an index for the proofs in this section.

As already mentioned, the main tool used in all of our results is the phase space description of the dynamics; a detailed description is given in appendix A. The phase space description allows the dynamics generated by the time evolution operator $W(t)$ to be studied instead by investigating linear dynamics generated by the ($\mathbb{Z}_2$) symplectic matrix, $S(t)$. I will first give a description of phase space

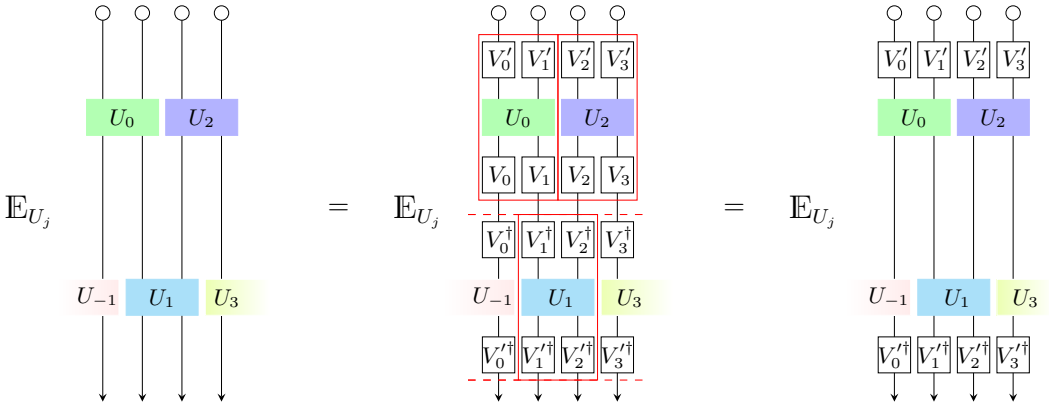

Figure 2.3: Twirling technique. This figure illustrates the fact that the statistical properties of the evolution operator are invariant under local transformations. In the image on the left there is a section of the circuit shown in figure 2.1. The middle image follows from noting that for any any pair of local (single-site) Clifford unitaries V_x, V_{x+1}, the random two-site Clifford unitary $(V_x \otimes V_{x+1})U_x(V_x' \otimes V_{x+1}')$ has the same probability distribution than U_x. Finally, in the image on the right, all of the local unitaries except for the those at the initial and final times cancel. This invariance property of $W(t)$ is different for integer and half-integer times t, which is specified in equations 2.14 and 2.16 respectively. This figure was produced by a collaborator, Daniele Toniolo.

description of the model, and then in lemmas 2.1 - 2.6 I will prove some properties of random symplectic matrices. The study of random symplectic matrices is an essential ingredient in the proofs of this section. In particular, for a random symplectic matrix S which is decomposed in block form as

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{2.13}$$

where A, B, C, D are all square matrices of equal size. I will characterise the probability distribution for the rank of a "quarter" submatrix C (in lemma 2.4), and generalise this distribution to the rank of a product $C_r \cdots C_2 C_1$ of independently sampled C-submatrices (in lemma 2.5).

Additionally, as already mentioned, the (random) time evolution operator $W(t)$ has an extra symmetry (lemma 2.7). This symmetry has been used previously in the reference [125] to obtain analytical results on localisation. This symmetry, which is

illustrated in figure 2.3, means that the time evolution operator $W(t)$ at any integer time t is invariant under the transformation

$$W(t) \;\mapsto\; \left(\bigotimes_x V'_x\right)^\dagger W(t) \left(\bigotimes_x V'_x\right) \;, \tag{2.14}$$

for any choice of local (single-site) Clifford unitaries $V'_1, \ldots, V'_L \in \mathcal{C}_N$. Consequently, for the probability distribution $P_t(\mathbf{u}'|\mathbf{u})$, which describes the probability that after a time t an initial Pauli operator $\mathbf{u}$ evolves to the Pauli operator $\mathbf{u}'$, this symmetry means that

$$P_t\left([\bigoplus_x S_x^{-1}]\mathbf{u}' \middle| [\bigoplus_x S_x]\mathbf{u}\right) = P_t(\mathbf{u}'|\mathbf{u}) \;, \tag{2.15}$$

for any choice of local (singe-site) symplectic matrices $S_1, \ldots, S_L$. Moreover, at half-integer times t the evolution operator displays a still higher degree of symmetry. In particular, the time evolution operator $W(t)$ for half-integer times is invariant under the transformation

$$W(t) \;\mapsto\; \left(\bigotimes_x V_x\right) W(t) \left(\bigotimes_x V'_x\right) \;, \tag{2.16}$$

for any choice of local (single-site) Clifford unitaries $V_1, V'_1, \ldots, V_L, V'_L \in \mathcal{C}_N$. Hence, analogously the probability distribution is such that

$$P_t\left([\bigoplus_x S'_x]\mathbf{u}' \middle| [\bigoplus_x S_x]\mathbf{u}\right) = P_t(\mathbf{u}'|\mathbf{u}) \;, \tag{2.17}$$

for any choice of local (singe-site) symplectic matrices $S_1, S'_1, \ldots, S_L, S'_L$.

I will then use and combine the proof of the additional symmetries of the evolution operator, and its implications for the probability distribution $P_t(\mathbf{u}'|\mathbf{u})$, with the proven results concerning probability distribution of the rank of the product of independently sampled C-submatrices. In particular, I will proceed to prove a series of results (lemmas 2.8 and 2.9), which are then used (in lemma 2.6) to calculate a bound for the probability that any initial non-identity Pauli operator $\mathbf{u}$ evolves to a Pauli operator $\mathbf{u}'$ which is non-identity at every site of the spin chain.

With this I will give the proofs which justify both half-integer time results

(results 2.3 and 2.4). I will first prove (lemma 2.11) that conditioned on the evolved Pauli operator being non-identity at every site of the spin chain, then the evolved Pauli operator is approximately uniformly distributed over all non-identity Pauli operators. Using this, I will then finally justify result 2.3 with a rigorous proof (lemma 2.12) and after justify result 2.4 also with a rigorous proof (lemma 2.13).

After this, I will then focus on the proofs which justify the results which apply to both integer and half-integer times (result 2.1 and 2.2). I will first prove a very general result, which describes the form of the probability distribution of the evolved Pauli operator (lemma 2.14). Then, after proving another pair of results (lemma 2.15 and 2.16), I will (in lemma 2.17) give the rigorous proof which justifies result 2.1. Finally, I will prove a rather involved result (lemma 2.15) which itself is a generalisation of another (lemma 2.15), for the case of the initial Pauli operator being non-identity only a continuous sub-set of sites of the spin chain. Then, I will (using lemma 2.18) give the proof which justifies result 2.2.

Also, I would like to reemphasise that all of the results in this section are meaningful only in the regime in which N, the number of qubits per site of the spin chain, is large ($N \gg \log L$).

Phase space description and random symplectic matrices

It is worth briefly describing again with more detail the phase space description of the model. An N-qubit Pauli operator is the (N-fold) tensor product of Pauli sigma matrices, including the one-qubit identity, multiplied by a global phase $\lambda \in \{1, \mathrm{i}, -1, -\mathrm{i}\}$. Ignoring the global phase, each N-qubit Pauli operator can be represented by a binary vector $\mathbf{u} = (q_1, p_1, q_2, p_2, \ldots, q_N, p_N) \in \{0,1\}^{2N}$ as

$$\sigma_{\mathbf{u}} = \bigotimes_{i=1}^{N} (\sigma_x^{q_i} \sigma_z^{p_i}) \tag{2.18}$$

and the product of Pauli operators can be written as $\sigma_{\mathbf{u}} \sigma_{\mathbf{u}'} = \lambda\, \sigma_{\mathbf{u}+\mathbf{u}'}$, where addition in the vector space $\mathbb{Z}_2^{2N}$ is modulo 2. The N-qubit Clifford group $\mathcal{C}_N \subseteq \mathrm{SU}(2^N)$ is the set of unitaries U which map each Pauli operator onto another Pauli operator $U \sigma_{\mathbf{u}} U^\dagger = \lambda\, \sigma_{\mathbf{u}'}$. Each Clifford unitary U can be represented by a $2N \times 2N$ symplectic

matrix S with entries in $\mathbb{Z}_2$ such that its action on Pauli operators can be calculated as a linear transformation of the binary vector, so

$$U\sigma_{\mathbf{u}}U^{\dagger} = \lambda\,\sigma_{S\mathbf{u}}\ , \tag{2.19}$$

where the matrix product $S\mathbf{u}$ is defined modulo 2. Due to the analogy with quasi free-bosons, in which dynamics are also linear and symplectic, the binary vectors $\mathbf{u} \in \{0,1\}^{2N}$ are known as the phase space representation of the N-qubit Pauli operator $\sigma_{\mathbf{u}}$, and similarly the symplectic matrices S are known as the phase space representation of the N-qubit Clifford unitary U.

Since any Pauli operator for the whole chain, $\sigma_{\mathbf{u}}$, can be represented by a binary vector $\mathbf{u}$, the whole chain of L sites with N-qubits per site has the phase space

$$\mathcal{V}_{\text{chain}} = \bigoplus_x \mathcal{V}_x \cong \mathbb{Z}_2^{2NL}\ , \tag{2.20}$$

where $\mathcal{V}_x \cong \mathbb{Z}_2^{2N}$ is the phase space of the single site x. It is worth emphasising that the tensor product of Hilbert spaces becomes a direct sum in phase space. Additionally, each Clifford unitary U_x, which acts on two sites of the spin chain, has the phase space representation given by the symplectic matrix $S_x \in \mathcal{S}_{2N}$, which acts on two sites of the phase space given by the subspace $\mathcal{V}_x \oplus \mathcal{V}_{x+1}$, and which has elements in $\mathbb{Z}_2$. Moreover, by using this direct sum decomposition, each symplectic matrix, S_x, can be written in the form

$$S_x = \begin{pmatrix} A_x & B_x \\ C_x & D_x \end{pmatrix}\ , \tag{2.21}$$

where A_x, B_x, C_x, D_x are $2N \times 2N$ matrices, with $A_x : \mathcal{V}_x \to \mathcal{V}_x$, $B_x : \mathcal{V}_{x+1} \to \mathcal{V}_x$, $C_x : \mathcal{V}_x \to \mathcal{V}_{x+1}$, $D_x : \mathcal{V}_{x+1} \to \mathcal{V}_{x+1}$. Therefore, since every Clifford unitary U_x has the phase space representation S_x, the (Clifford) unitary which acts on the whole chain, U_{chain}, has the phase space representation

$$S_{\text{chain}} = \left(\bigoplus_{x\,\text{odd}} S_x\right)\left(\bigoplus_{x\,\text{even}} S_x\right)\ , \tag{2.22}$$

which act on the entire phase space $\mathcal{V}_{\text{chain}}$. Hence, the time evolution operator $W(t)$ has the corresponding phase space representation, $S(t)$, which for integer and half-integer times is given by

$$S(t) = \begin{cases} (S_{\text{chain}})^t & \text{integer } t \\ (\bigoplus_{x\,\text{even}} S_x)(S_{\text{chain}})^{t-1/2} & \text{half-integer } t \end{cases} . \tag{2.23}$$

Before proceeding to the mathematical proofs concerning uniformly distributed symplectic matrices, I shall briefly note some features of symplectic matrices (for greater detail see appendix A). The group of symplectic matrices, $\mathcal{S}_N$, is the set of matrices $S : \mathbb{Z}_2^{2N} \to \mathbb{Z}_2^{2N}$ such that

$$\langle S\mathbf{u}, S\mathbf{u}' \rangle = \langle \mathbf{u}, \mathbf{u}' \rangle = \mathbf{u}^T J \mathbf{u}' = \mathbf{u}^T \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{u}' \tag{2.24}$$

for all $\mathbf{u}, \mathbf{u}' \in \mathbb{Z}_2^{2N}$, and where addition is defined modulo 2. This condition is equivalent to the condition $S^T J S = J \bmod 2$. Another equivalent way of writing these conditions is in terms of the columns of the matrix. So the columns of the matrix $S = (\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2, \ldots, \mathbf{u}_N, \mathbf{v}_N)$ satisfy

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 , \tag{2.25}$$

$$\langle \mathbf{u}_i, \mathbf{v}_j \rangle = \delta_{ij} . \tag{2.26}$$

This perspective and set of conditions on the columns of the symplectic matrix is important for all of the proofs in this section. With this in mind, in the following lemma I will prove an algorithm for uniformly sampling a symplectic matrix (from the group $\mathcal{S}_N$). In the lemma after I will upper-bound and lower-bound the order of the ($\mathbb{Z}_2$) symplectic group, and then in the lemma after that prove that a uniformly random symplectic matrix acting on a binary vector has a random output.

Lemma 2.1. The following algorithm allows to uniformly sample from the symplectic group $\mathcal{S}_N$.

1. Generate $\mathbf{u}_1$ by picking any of the $(2^{2N} - 1)$ non-zero vectors in $\mathbb{Z}_2^{2N}$.

2. Generate $\mathbf{v}_1$ by picking any of the 2^{2N-1} vectors satisfying $\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = 1$.

3. Generate $\mathbf{u}_2$ by picking any of the $(2^{2N-2}-1)$ non-zero vectors satisfying $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{v}_1, \mathbf{u}_2 \rangle = 0$.

4. Generate $\mathbf{v}_2$ by picking any of the 2^{2N-3} vectors satisfying $\langle \mathbf{u}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ and $\langle \mathbf{u}_2, \mathbf{v}_2 \rangle = 1$.

5. Continue generating $\mathbf{u}_3, \mathbf{v}_3, \ldots, \mathbf{u}_N, \mathbf{v}_N$ in analogous fashion, completing the matrix $S = (\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2, \ldots, \mathbf{u}_N, \mathbf{v}_N)$.

Proof. We first look at the number of vectors $(\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2, \ldots, \mathbf{u}_N, \mathbf{v}_N)$, as stated above, that ensures S symplectic.

The first column, $\mathbf{u}_1$, can be any non-zero vector in $\mathbb{Z}_2^{2N}$, of which there are $2^{2N}-1$ possible choices. The second column, $\mathbf{v}_1$ must have $\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = 1$. This imposes one constraint on the $2N$ components of $\mathbf{v}_1$, and hence the number of independent components is $2N-1$, each of which belongs to $\mathbb{Z}_2$. Therefore the number of vectors $\mathbf{v}_1$ equals 2^{2N-1}, where it is worth noting that we need not account for $\mathbf{v}_1$ being the zero vector, since this is captured by the condition $\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = 1$. This completes the argument for the first pair of columns $\mathbf{u}_1, \mathbf{v}_1$, and the proof proceeds in a similar manner. The column $\mathbf{u}_2$ must now satisfy two constraints, being linearly independent of the previous columns. Hence, $\mathbf{u}_2$ has $2N-2$ independent components, and hence neglecting the zero vector, therefore there are 2^{2N-2} possible choices of $\mathbf{u}_2$. The next column $\mathbf{v}_2$ must satisfy $\langle \mathbf{u}_2, \mathbf{v}_2 \rangle = 1$ and moreover be linearly independent of the previous column $\mathbf{u}_1, \mathbf{v}_1$, for which there are a total of 2^{2N-3} possible choices. Hence, this argument proceeds in a similar manner again for the next pair of columns.

The argument made in this proof does not depend on the particular choice of any column. Therefore, if we select each column uniformly from all the possible choices satisfying the constraints explained, then the resulting matrix from the $2N$ choices of column is also uniformly distributed.

□

Lemma 2.2. The order of the symplectic group is

$$|\mathcal{S}_N| = (2^{2N}-1)2^{2N-1}(2^{2N-2}-1)2^{2N-3}\cdots(2^2-1)2^1\,, \tag{2.27}$$

and it satisfies

$$a(N)\,2^{2N^2+N} \;\le\; |\mathcal{S}_N| \;\le\; b(N)\,2^{2N^2+N} \tag{2.28}$$

with $0.64 < a(N) < b(N) < 0.78$.

Proof. We start considering $\ln|\mathcal{S}_N|$.

$$\begin{aligned}
\ln|\mathcal{S}_N| &= \ln\left[\prod_{i=1}^{N}(2^{2i}-1)\prod_{j=1}^{N}2^{2j-1}\right] \\
&= \sum_{i=1}^{N}\ln\left(2^{2i}-1\right)+\sum_{j=1}^{N}\ln 2^{2j-1} \\
&= \sum_{i=1}^{N}\ln\left[2^{2i}(1-2^{-2i})\right]+\sum_{j=1}^{N}(2j-1)\ln 2 \\
&= \sum_{i=1}^{N}2i\ln 2+\sum_{i=1}^{N}\ln\left(1-2^{-2i}\right)+\sum_{j=1}^{N}(2j-1)\ln 2 \\
&= N(N+1)\ln 2+\sum_{i=1}^{N}\ln\left(1-2^{-2i}\right)+N^2\ln 2 \\
&= N(2N+1)\ln 2+\sum_{i=1}^{N}\ln\left(1-2^{-2i}\right).
\end{aligned} \tag{2.29}$$

We use $\frac{x}{1+x} < \ln(1+x) < x$, with $x \neq 0$ and $x > -1$, to upper and lower bound the logarithm in equation 2.29. The corresponding bounds on $|\mathcal{S}_N|$ are obtained after exponentiating

$$\sum_{i=1}^{N}\ln\left(1-2^{-2i}\right) < -\sum_{i=1}^{N}2^{-2i} = -\sum_{i=1}^{N}\frac{1}{4^i} = -\frac{1}{3}\left(1-\frac{1}{4^N}\right) \tag{2.30}$$

$b(N)$ is defined to be $b(N) \equiv e^{-\frac{1}{3}(1-\frac{1}{4^N})}$, moreover $b(N) < b(1) = e^{-\frac{1}{4}} < 0.78$.

To obtain the lower bound of $|S_N|$, from 2.29 we have:

$$\sum_{i=1}^{N} \ln\left(1-2^{-2i}\right) > -\sum_{i=1}^{N} \frac{2^{-2i}}{1-2^{-2i}}$$

and $a(N) \equiv e^{-\sum_{i=1}^{N} \frac{2^{-2i}}{1-2^{-2i}}} < b(N)$. We observe that:

$$a(N) \equiv e^{-\sum_{i=1}^{N} \frac{1}{2^{2i}-1}} \geq e^{-\frac{1}{3}\sum_{i=0}^{N-1} \frac{1}{2^{2i}}} > e^{-\frac{4}{9}} > 0.64 \tag{2.31}$$

□

In the next lemma, I will show that uniformly distributed symplectic matrix have random outputs.

Lemma 2.3 (Uniform output)**.** If $S \in \mathcal{S}_N$ is uniformly distributed, then for any pair of non-zero vectors $\mathbf{u}, \mathbf{u}' \in \mathbb{Z}_2^{2N}$ we have

$$\text{prob}\{\mathbf{u}' = S\mathbf{u}\} = (2^{2N}-1)^{-1} \ . \tag{2.32}$$

Proof. Let us first consider the case $\mathbf{u} = (1,0,\ldots,0)^T$. If we follow the algorithm of Lemma 2.1, then the image of $(1,0,\ldots,0)^T$ is uniformly distributed over the $(2^{2N}-1)$ non-zero vectors, and hence, it follows the probability distribution given in equation 2.32. To show this distribution holds for any given $\mathbf{u}$, take any $S_0 \in \mathcal{S}_N$ such that $S_0\mathbf{u} = (1,0,\ldots,0)^T$, and note that, if S is uniformly distributed then so is SS_0. □

With these results concerning random symplectic matrices, and in particular this approach based on columns of symplectic matrices, I will now prove a result concerning the probability distribution of the rank of a submatrix of a symplectic matrix (equation 2.21). Also, I would like to point out that in the phase space description of the model each random symplectic matrix, which represents a Clifford unitary, acts on two subspaces and hence is a $4N \times 4N$ symplectic matrix ($S_x \in \mathcal{S}_{2N}$). Hence, for the rest of this section I will consider $4N \times 4N$ symplectic matrix (with elements in $\mathbb{Z}_2$) only.

Lemma 2.4. Any given $S \in \mathcal{S}_{2N}$ can be written in block form

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{2.33}$$

according to the local decomposition $\mathbb{Z}_2^{4N} = \mathbb{Z}_2^{2N} \oplus \mathbb{Z}_2^{2N}$. If S is uniformly distributed this then induces a distribution on the sub-matrices A, B, C, D. For each of them ($E = A, B, C, D$) the induced distribution satisfies

$$\text{prob}\big\{\text{rank}\, E \leq 2N-k\big\} \;\leq\; \min\{2^k, 4\} \frac{2^{-k^2}}{(1-2^{-2N})^k} \approx 4 \times 2^{-k^2} \;. \tag{2.34}$$

Proof. We proceed by studying the rank of C and later generalizing the results to A, B, D. In the following, we will also analyse the case where $kge1$, since the result in equation 2.34 is trivial for $k = 0$. Let us start by counting the number of matrices $S \in \mathcal{S}_{2N}$ with a sub-matrix C satisfying $C\mathbf{u} = \mathbf{0}$ for a given (arbitrary) non-zero vector $\mathbf{u} \in \mathbb{Z}_2^{2N}$. Let r denote the position of the last "1" in $\mathbf{u}$, so that it can be written as:

$$\mathbf{u} = \big(\underbrace{\mathbf{u}^1, \ldots, \mathbf{u}^{r-1}}_{r-1}, 1, \underbrace{0, \ldots, 0}_{2N-r}\big)^T \;, \tag{2.35}$$

where $\mathbf{u}^1, \ldots, \mathbf{u}^{r-1} \in \{0, 1\}$. Then, the constraint $C\mathbf{u} = \mathbf{0}$ can be written as

$$\begin{cases} C_{i,1} = 0, \quad \text{if } r = 1, \quad \text{with } 1 \leq i \leq 2N \\ C_{i,r} = \sum_{j=1}^{r-1} C_{i,j}\, \mathbf{u}^j, \quad \text{if } r > 1, \quad \text{with } 1 \leq i \leq 2N \end{cases} \tag{2.36}$$

where $C_{i,j}$ are the components of C. The statement in equation 2.36 is a constraint on the r-th column of the matrix C.

Next, we follow the algorithm introduced in lemma 2.1 for generating a matrix $S \in \mathcal{S}_{2N}$ column by column (from left to right) and in addition to the symplectic constraints we include equation 2.36. The conditions in equation 2.36 can be imposed by ignoring the constraints during the generation of columns $1, \ldots, r-1$, completely fixing the rows $2N < i \leq 4N$ of the r column that corresponds to the r-th column of

the matrix C, and ignoring the constraints again during the generation of columns $r+1,\ldots,4N$. By counting as in lemma 2.2, we obtain that the number of matrices $S \in \mathcal{S}_{2N}$ satisfying $C\mathbf{u} = \mathbf{0}$ follows

$$\begin{aligned} &|\{S \in \mathcal{S}_{2N} : C\mathbf{u} = \mathbf{0}\}| \qquad (2.37)\\ &\leq (2^{4N}-1)(2^{4N-1})\cdots(2^{4N-(r-2)} - |\alpha - 1|)(2^{2N-(r-1)} - \alpha)(2^{4N-r} - |\alpha-1|)\cdots 2^1 \ , \end{aligned}$$

where $\alpha = 1$ if r is odd and $C_{i,r} = 0$ for all i; and $\alpha = 0$ otherwise. In the case of $r = 1$, equation 2.37is the following equality

$$|\{S \in \mathcal{S}_{2N} : C\mathbf{u} = \mathbf{0}\}| = (2^{2N}-1)(2^{4N-1})(2^{4N-2}-1)\cdots 2^1 \ .$$

For some values of the first $r-1$ columns of S and the r-th column of C, it is impossible to complete the r-th column of A satisfying the symplectic constraints (equations 2.25 and 2.26), and hence equation 2.37 is an inequality.

The probability that a random S satisfies $C\mathbf{u} = \mathbf{0}$ is

$$\text{prob}\{C\mathbf{u} = \mathbf{0}\} = \frac{|\{S \in \mathcal{S}_{2N} : C\mathbf{u} = \mathbf{0}\}|}{|\mathcal{S}_{2N}|} \ . \qquad (2.38)$$

By noting that all factors in equation 2.37 are the same as in equation 2.27 except for the factor at position r, we obtain

$$\begin{aligned} \text{prob}\{C\mathbf{u} = \mathbf{0}\} &\leq \frac{2^{2N-(r-1)} - \alpha}{2^{4N-(r-1)} - \alpha'} \leq \frac{2^{2N-(r-1)}}{2^{4N-(r-1)} - 1} \\ &= \frac{2^{-2N}}{1 - 2^{(r-1)-4N}} \leq \frac{2^{-2N}}{1-2^{-2N}} \ , \qquad (2.39) \end{aligned}$$

where $\alpha' = 1$ if r is odd and $\alpha' = 0$ otherwise. The last inequality above follows from $r \leq 2N$. The bound in equation 2.39 is correct also for $r = 1$. The fact that bound in equation 2.39 is independent of r is crucial for the rest of the proof.

Next, we generalise bound in equation 2.39 to the case where $C\mathbf{u}_i = 0$ for k given linearly-independent vectors $\mathbf{u}_i \in \{\mathbf{u}_1,\ldots,\mathbf{u}_k\}$. To do this, we take the $2n \times k$ matrix $[\mathbf{u}_1,\ldots,\mathbf{u}_k]$ and perform Gauss-Jordan elimination, operating on the columns,

to obtain a matrix $[\mathbf{v}_1,\dots,\mathbf{v}_k]$ having column-echelon form. $\{C\mathbf{u}_1=\mathbf{0},\dots,C\mathbf{u}_k=\mathbf{0}\}$ is equivalent to $\{C\mathbf{v}_1=\mathbf{0},\dots,C\mathbf{v}_k=\mathbf{0}\}$, in fact only two operations are performed on the set $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$ to obtain the set $\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$: changing the order of the vectors $\{\mathbf{u}_1,\dots,\mathbf{u}_k\}$, replacing a vector $\mathbf{u}_j$ with the sum of $\mathbf{u}_j$ with another vector $\mathbf{u}_l$. If we denote by r_i the position of the last "1" of $\mathbf{v}_i$, then column-echelon form amounts to $r_1<r_2<\cdots<r_k$. Now we proceed as above to generate each column of S satisfying the symplectic and the $C\mathbf{v}_i=\mathbf{0}$ constraints. This gives

$$\text{prob}\{C\mathbf{u}_1=\mathbf{0},\dots,C\mathbf{u}_k=\mathbf{0}\}\leq\frac{2^{2N-(r_1-1)}-\alpha_1}{2^{4N-(r_1-1)}-\alpha_1'}\frac{2^{2N-(r_2-1)}-\alpha_2}{2^{4N-(r_2-1)}-\alpha_2'}\cdots\frac{2^{2N-(r_k-1)}-\alpha_k}{2^{4N-(r_k-1)}-\alpha_k'}\,, \tag{2.40}$$

where $\alpha_i,\alpha_i'\in\{0,1\}$. Similarly as in equation 2.39, we obtain

$$\text{prob}\{C\mathbf{u}_1=\mathbf{0},\dots,C\mathbf{u}_k=\mathbf{0}\}\leq\frac{2^{-2Nk}}{(1-2^{-2N})^k}\,. \tag{2.41}$$

Multiplying the above bound by the number $\mathcal{N}_k^{2N}$ of k-dimensional subspaces of $\mathbb{Z}_2^{2N}$ (see appendix B), then we obtain

$$\begin{aligned}\text{prob}\{\text{rank}(C)\leq 2N-k\}&=\mathcal{N}_k^{2N}\,\text{prob}\{C\mathbf{u}_1=\mathbf{0},\dots,C\mathbf{u}_k=\mathbf{0}\}\\&\leq\min\{2^k,4\}\,\frac{2^{2Nk}}{2^{k^2}}\,\frac{2^{-2Nk}}{(1-2^{-2N})^k}\,,\\&=\min\{2^k,4\}\,\frac{2^{-k^2}}{(1-2^{-2N})^k}\,,\end{aligned} \tag{2.42}$$

where in the last inequality we used the upper-bound proven in lemma B.2 of appendix B.

I will now demonstrate that this argument applies to any of the four sub-matrices A,B,C,D. Consider the symplectic matrix

$$M=\begin{pmatrix}0&\mathbb{1}\\\mathbb{1}&0\end{pmatrix}\,, \tag{2.43}$$

and its left and right action on the symplectic matrix S

$$SM = \begin{pmatrix} B & A \\ D & C \end{pmatrix} , \quad MS = \begin{pmatrix} C & D \\ A & B \end{pmatrix} , \tag{2.44}$$

so M permutes the block matrix S by swapping the rows and columns. Since, the product of symplectic matrices is also a symplectic matrix, we have that the argument above for the case of $C = 0$ applies to any of the four sub-matrices A, B, C, D. With this the proof of equation 2.34 is complete. □

I will now generalise this result to include the product of multiple independent symplectic matrices, and in particular to the probability distribution for the rank of a product $C_r \cdots C_2 C_1$ of (r) independently sampled C-submatrices.

Lemma 2.5. Let the random matrices $S_1, S_2, \ldots, S_r \in \mathcal{S}_{2N}$ be independent and uniformly distributed, which induces a distribution for the sub-matrices

$$S_i = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix} . \tag{2.45}$$

For any choice $E_i \in \{A_i, B_i, C_i, D_i\}$ for each $i \in \{1, \ldots, r\}$, we have

$$\text{prob}\big\{\text{rank}(E_r \cdots E_1) \leq 2N-k\big\} \leq \frac{2^k}{(1-2^{-2N})^k} \binom{k+r-1}{k} 2^{-\frac{1}{2}k^2} . \tag{2.46}$$

Proof. Before analysing the rank of the product of r independent random matrices $C_r \cdots C_1$, we start by studying a simpler problem. Analysing the rank of the product CF where C follows the usual C-distribution and F is a fixed $2N \times 2N$ matrix with $\text{rank}(F) = 2N - k_1$. Noting that the input space of C has dimension $2N - k_1$, using the result in lemma 2.4, with $k_2 \equiv k - k_1 \geq 0$, we obtain

$$\begin{aligned} \text{prob}\big\{\text{rank}(CF) \leq 2N-k\big\} &\leq \mathcal{N}_{k_2}^{2N-k_1} \text{prob}\{C\mathbf{u}_1 = \mathbf{0}, \ldots, C\mathbf{u}_{k_2} = \mathbf{0}\} \\ &\leq \min\{2^{k_2}, 4\} \frac{2^{(2N-k_1)k_2}}{2^{k_2^2}} \frac{2^{-2Nk_2}}{(1-2^{-2N})^{k_2}} \\ &\leq \frac{2^{k_2 - k_1 k_2 - k_2^2}}{(1-2^{-2N})^{k_2}} , \end{aligned} \tag{2.47}$$

Proceeding in a similar fashion, we can analyse the product of two independent C-matrices. To do so, we multiply two factors of equation 2.47 and sum over all possible intermediate kernel sizes k_1, obtaining

$$\begin{aligned}\text{prob}\big\{\text{rank}(C_2C_1) \leq 2N-k\big\} &\leq \sum_{k_1=0}^{k} \frac{2^{k_2-k_2k_1-k_2^2}}{(1-2^{-2N})^{k_2}} \frac{2^{k_1-k_1^2}}{(1-2^{-2N})^{k_1}} \\ &= \sum_{k_1=0}^{k} \frac{2^{k-k_2k_1-k_1^2-k_2^2}}{(1-2^{-2N})^{k}} \,, \end{aligned} \tag{2.48}$$

where again $k_2 = k-k_1$. Analogously, we can bound the rank of a product of r independent random C-matrices as

$$\begin{aligned}\text{prob}\big\{\text{rank}(C_r\cdots C_1) \leq 2N-k\big\} &\leq \sum_{\{k_i\}}\prod_{i=1}^{r} \frac{2^{k_i-k_i\sum_{j=1}^{i}k_j}}{(1-2^{-2N})^{k_i}} \\ &= \frac{2^k}{(1-2^{-2N})^{k}} \sum_{\{k_i\}} 2^{-\sum_{i=1}^{r}k_i\sum_{j=1}^{i}k_j} \,, \end{aligned} \tag{2.49}$$

where the sum $\sum_{\{k_i\}}$ runs over all sets of r non-negative integers $\{k_1,\ldots,k_r\}$ such that $\sum_{i=1}^{r}k_i = k$. In other words, this is all the possible ways of sharing k units among r distinguishable parts. The number of all these sets can be bounded (see lemma B.3 in appendix B) as

$$\sum_{\{k_i\}} 1 = \binom{k+r-1}{r-1} = \binom{k+r-1}{k} \leq (1+r)^k \leq (2r)^k \,. \tag{2.50}$$

Finally, for any set $\{k_1,\ldots,k_r\}$ we have

$$\begin{aligned} k^2 = \sum_{i=1}^{r}\sum_{j=1}^{r} k_ik_j &\leq \sum_{i=1}^{r}\sum_{j=1}^{i} k_ik_j + \sum_{i=1}^{r}\sum_{j=i}^{r} k_ik_j \\ &= 2\sum_{i=1}^{r}\sum_{j=1}^{i} k_ik_j \,. \end{aligned} \tag{2.51}$$

Substituting equations 2.50 and 2.51 back in to equation 2.49 we obtain

$$\text{prob}\big\{\text{rank}(C_r\cdots C_1) \leq 2N-k\big\} \;\leq\; \frac{2^k}{(1-2^{-2N})^{k}} \binom{k+r-1}{k} 2^{-\frac{1}{2}k^2} \,. \tag{2.52}$$

Using the same argument as in the proof of lemma 2.4, this proof applies to all products of submatrix $E \in \{A,B,C,D\}$. □

As I have discussed already, in the phase space description of the dynamics the Clifford unitaries are represented by symplectic matrices, which act on and map binary vectors, which represent Pauli operators. Hence, to study the probabilistic transformation of Pauli operators by the time evolution operator, in the next lemma I will, using lemma 2.5, prove an upper-bound to the probability that a random binary vector $\mathbf{u}$ is mapped to the zero vector by the product $C_r \cdots C_2 C_1$ of (r) independently sampled C-submatrices. This next lemma is an essential component in the proof of all of the main results (results 2.1,2.2,2.3, and 2.4).

Lemma 2.6. If the random variables $S_1, S_2, \ldots, S_r \in \mathcal{S}_{2N}$ and $\mathbf{u} \in \mathbb{Z}_2^{2N}$ are independent and uniformly distributed it follows that

$$\text{prob}\big\{E_r \cdots E_1 \mathbf{u} = \mathbf{0}\big\} \leq 8\, r\, 2^{-N} \,. \tag{2.53}$$

being $E_j \in \{A_j, B_j, C_j, D_j\}$ the subblocks of the symplectic matrices $S_1, ..., S_r$.

Proof. If M is a fixed $2N \times 2N$ matrix with rank $M = 2N - k$ and $\mathbf{u} \in \mathbb{Z}_2^{2N}$ is uniformly distributed, then

$$\text{prob}\big\{M\mathbf{u} = \mathbf{0}\big\} = \frac{2^k}{2^{2N}} \,. \tag{2.54}$$

Also, if rank $M > 2N - k$ then

$$\text{prob}\big\{M\mathbf{u} = \mathbf{0}\big\} \leq \frac{2^{k-1}}{2^{2N}} \,. \tag{2.55}$$

This inequality is useful for the following bound

$$\begin{aligned}
&\text{prob}\big\{C_r \cdots C_1 \mathbf{u} = \mathbf{0}\big\} \\
&\quad = \text{prob}\big\{C_r \cdots C_1 \mathbf{u} = \mathbf{0} \text{ and } \text{rank}(C_r \cdots C_1) > 2N - k\big\} \\
&\quad + \text{prob}\big\{C_r \cdots C_1 \mathbf{u} = \mathbf{0} \text{ and } \text{rank}(C_r \cdots C_1) \leq 2N - k\big\} \\
&\quad \leq \text{prob}\big\{C_r \cdots C_1 \mathbf{u} = \mathbf{0} \,\big|\, \text{rank}(C_r \cdots C_1) > 2N - k\big\}
\end{aligned}$$

$$
\begin{aligned}
&+\operatorname{prob}\big\{\operatorname{rank}(C_r\cdots C_1)\leq 2N-k\big\}\\
&\leq 2^{k-1-2N}+\frac{2^k}{(1-2^{-2N})^k}(1+r)^k 2^{-\frac{1}{2}k^2}\ , \qquad (2.56)
\end{aligned}
$$

where the last inequality uses equation 2.55 and lemma 2.5 (and additional lemma B.3 in the appendix B).

Using

$$
\begin{aligned}
\frac{1}{(1-2^{-2N})^k}\leq\frac{1}{(1-2^{-2N})^{2N}}&=\left(1+\frac{1}{2^{2N}-1}\right)^{2N}=\left(1+\frac{1}{2\left(2^{2N-1}-\frac{1}{2}\right)}\right)^{2N}\\
&\leq\left(1+\frac{1}{4N}\right)^{2N}\leq\sqrt{e}<2\ , \qquad (2.57)
\end{aligned}
$$

we obtain

$$
\operatorname{prob}\big\{C_r\cdots C_1\mathbf{u}=\mathbf{0}\big\}\leq 2^{k-2N}+2\,(4r)^k 2^{-\frac{1}{2}k^2}=\varepsilon\ , \qquad (2.58)
$$

where the last equality defines ε. Note that the left-hand side above is independent of k. Hence, for each value of k we have a different upper bound. We are interested in the tightest one of them. Therefore, we need to find a value of $k\in[1,2N]$ that makes the upper bound in equation 2.58 have a small enough value. This can be done by equating each of the two terms to $\varepsilon/2$ as

$$
2^{k-2N}=2\,(4r)^k 2^{-\frac{1}{2}k^2}=\frac{\varepsilon}{2}\ . \qquad (2.59)
$$

Isolating k from the first and second terms gives

$$
k=2N-\log_2\frac{2}{\varepsilon}\ , \qquad (2.60)
$$

$$
k=\log_2 4r+\sqrt{\log_2^2 4r+\log_2\frac{2}{\varepsilon}+1}\ , \qquad (2.61)
$$

where we only keep the positive solution. Equating the above two identities for k we

obtain

$$\begin{aligned} N &= \frac{1}{2}\left(\log_2 4r + \log_2\frac{2}{\varepsilon} + \sqrt{\log_2^2 4r + \log_2\frac{2}{\varepsilon}+1}\right) \\ &\leq \frac{1}{2}\left(\log_2 4r + \log_2\frac{2}{\varepsilon} + \log_2 4r + \sqrt{\log_2\frac{2}{\varepsilon}+1}\right) \\ &\leq \log_2 4r + \log_2\frac{2}{\varepsilon}\,, \end{aligned} \tag{2.62}$$

which implies

$$\varepsilon \leq 8\,r\,2^{-N}\,. \tag{2.63}$$

Substituting this into equation 2.58 we finish the proof of this lemma. □

Additional symmetry

I will now demonstrate another key tool in the proof of the main results of this section. I will now present what is referred to as the twirling technique in the work [125] and discuss how it applies to the random Clifford circuit model we consider. Before stating the lemma itself, it is worthwhile recalling that the definition of the evolution operator after an integer time t is:

$$\begin{aligned} W(t) &\equiv \left[(U_1\otimes U_3\otimes\cdots\otimes U_{L-1})(U_0\otimes U_2\otimes\cdots\otimes U_{L-2})\right]^t \\ &= (U_{\text{odd}}U_{\text{even}})^t = (U_{\text{chain}})^t\,, \end{aligned}$$

and after a half-integer time t is:

$$W(t) \equiv U_{\text{even}}\,(U_{\text{chain}})^{t-1/2}\,.$$

Lemma 2.7. Consider a set of $2L$ single-site Clifford unitaries $V_x, V_x' \in \mathcal{C}_N$, these unitaries are fixed. At integer time t, the random evolution operator $W(t)$, as defined above, has identical statistical properties to

$$\left(\bigotimes_{x=0}^{L-1} V_x'^{\dagger}\right) W(t) \left(\bigotimes_{x=0}^{L-1} V_x'\right)\,. \tag{2.64}$$

Similarly, at half-integer time t the evolution operator $W(t)$ has identical statistical properties to

$$\left(\bigotimes_{x=0}^{L-1} V_x\right) W(t) \left(\bigotimes_{x=0}^{L-1} V'_x\right) . \tag{2.65}$$

Proof. First, we note that any uniformly distributed two-site Clifford unitary $U_x \in \mathcal{C}_{2N}$ has identical statistical properties to the unitary $(V_x \otimes V_{x+1})U_x(V'_x \otimes V'_{x+1})$ for any arbitrary choice of $V_x, V_{x+1}, V'_x, V'_{x+1} \in \mathcal{C}_N$; this is known as single-site Haar invariance. Hence, we introduce the primed notation for the random two-site Clifford unitary U_x

$$U'_x \equiv (V_x \otimes V_{x+1})U_x(V'_x \otimes V'_{x+1}) \text{ for even } x \in \mathbb{Z}_L , \tag{2.66}$$

$$U'_x \equiv (V'_x \otimes V'_{x+1})^{-1}U_x(V_x \otimes V_{x+1})^{-1} \text{for odd } x \in \mathbb{Z}_L , \tag{2.67}$$

where $V_x, V_{x+1}, V'_x, V'_{x+1} \in \mathcal{C}_N$ are any arbitrary choice of single-site Clifford unitary. Consequently, the primed version of the global dynamics for integer t becomes

$$W'(t) = \left(\bigotimes_{x=0}^{L-1} V_x'^{\dagger}\right) W(t) \left(\bigotimes_{x=0}^{L-1} V'_x\right) , \tag{2.68}$$

and for half-integer t

$$W'(t) = \left(\bigotimes_{x=0}^{L-1} V_x\right) W(t) \left(\bigotimes_{x=0}^{L-1} V'_x\right) . \tag{2.69}$$

The single-site Haar invariance of the probability distributions of the primed and not-primed evolution operators are identical, this proves the result. □

Binary vector time evolution

With the two mathematical results proven in lemma 2.6 and lemma 2.7, I will now prove a series of results which culminate in bounding the probability that any (non-zero) initial binary vector of the total phase space (of the spin chain) $\mathbf{u} \in \mathbb{Z}_2^{2NL}$ is through its time evolution mapped to a binary vector which is non-zero at every site of the phase space $\mathcal{V}_x$. In other words, I will bound the probability that the initial Pauli operator after a time t evolves to another Pauli operator which is non-identity

at every site of the spin chain.

Before this, I will first define some notation. The time evolution of an initial vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$ at time t is denoted by $\mathbf{u}^t = S(t)\mathbf{u}^0$. If the initial vector is supported only at the origin $\mathbf{u}^0 \in \mathcal{V}_0$ then, as time t increases, the evolved vector $\mathbf{u}^t$ is supported on the light-cone

$$x \in \{-(2t-1), -(2t-2), \ldots, 2t\} \subseteq \mathbb{Z}_L \ . \tag{2.70}$$

It is worth recalling that the scrambling time, the minimum time such that any initially local operator spreads throughout the entire spin chain, in this model is $t_{\text{scr}} = \frac{L}{4}$. Finally, we denote the projection of $\mathbf{u}$ on the local subspace $\mathcal{V}_x$ by $\mathbf{u}_x$.

In the next lemma, I will bound the probability distribution of the evolved binary vector $\mathbf{u}^t$ being zero or non-zero on the boundary of the light-cone for an initial an initial vector which is non-zero only on a single-site.

Lemma 2.8. Consider a non-zero vector located at the origin $\mathbf{u}^0 \in \mathcal{V}_0$ and its time evolution $\mathbf{u}^t$ for any $t \in \{1/2, 1, 3/2, \ldots\}$. The projection of $\mathbf{u}^t$ at the rightmost site of the lightcone $x = 2t$ follows the probability distribution

$$P(\mathbf{u}^t_{2t}) = \begin{cases} \frac{1-q_t}{2^{2N}-1} & \text{if } \mathbf{u}^t_{2t} \neq \mathbf{0} \\ q_t & \text{if } \mathbf{u}^t_{2t} = \mathbf{0} \end{cases} \ , \tag{2.71}$$

where $q_t \leq 2t\, 2^{-2N}$. The projection onto the second rightmost site $\mathbf{u}^t_{2t-1}$ also obeys distribution given in equation 2.71.

Proof. After half a time step the evolved vector $\mathbf{u}^{1/2}$ is supported on sites $x \in \{0, 1\}$ and it is determined by

$$\mathbf{u}^{1/2}_0 \oplus \mathbf{u}^{1/2}_1 = S_0(\mathbf{u}^0_0 \oplus \mathbf{0}) \ . \tag{2.72}$$

Lemma 2.3 tells us that the vector $\mathbf{u}^{1/2}_0 \oplus \mathbf{u}^{1/2}_1$ is uniformly distributed over all non-zero vectors in $\mathcal{V}_0 \oplus \mathcal{V}_1$. This implies that the vectors $\mathbf{u}^{1/2}_0$ and $\mathbf{u}^{1/2}_1$ satisfy

$$\text{prob}\{\mathbf{u}^{1/2}_x = \mathbf{0}\} = \frac{2^{2N}-1}{2^{4N}-1} \leq 2^{-2N} \ , \tag{2.73}$$

and have probability distribution of the form given in equation 2.71 with $t = 1/2$. In the next time step we have

$$\mathbf{u}_1^1 \oplus \mathbf{u}_2^1 = S_1(\mathbf{u}_1^{1/2} \oplus \mathbf{0}) \,. \tag{2.74}$$

Hence, if $\mathbf{u}_1^{1/2} = \mathbf{0}$ then $\mathbf{u}_1^1 = \mathbf{u}_2^1 = \mathbf{0}$. Also, applying again lemma 2.3 we see that, if $\mathbf{u}_1^{1/2} \neq \mathbf{0}$, then $\mathbf{u}_1^1 \oplus \mathbf{u}_2^1$ is uniformly distributed over all non-zero values. Putting these things together we conclude that $\mathbf{u}_1^1$ and $\mathbf{u}_2^1$ satisfy

$$\begin{aligned} \text{prob}\{\mathbf{u}_x^1 = \mathbf{0}\} &= \text{prob}\{\mathbf{u}_1^{1/2} = \mathbf{0}\} + \text{prob}\{\mathbf{u}_1^{1/2} \neq \mathbf{0}\}\,\text{prob}\{\mathbf{u}_x^1 = \mathbf{0} | \mathbf{u}_1^{1/2} \neq \mathbf{0}\} \\ &\leq \text{prob}\{\mathbf{u}_1^{1/2} = \mathbf{0}\} + \text{prob}\{\mathbf{u}_1^{1/2} \neq \mathbf{0}\}\, 2^{-2N} \\ &\leq 2 \times 2^{-2N} \,, \end{aligned} \tag{2.75}$$

and have probability distribution of the form given in equation 2.71 with $t = 1$.

We can proceed as above, applying lemma 2.3 to each evolution step

$$\mathbf{u}_{2t-1}^t \oplus \mathbf{u}_{2t}^t = S_{2t-1}(\mathbf{u}_{2t-1}^{t-1/2} \oplus \mathbf{0}) \,, \tag{2.76}$$

for $t = 1/2, 1, 3/2, 2, \ldots$ This gives us the recursive equation

$$\begin{aligned} \text{prob}\{\mathbf{u}_{2t}^t = \mathbf{0}\} &= \text{prob}\{\mathbf{u}_{2t-1}^{t-1/2} = \mathbf{0}\} + \text{prob}\{\mathbf{u}_{2t-1}^{t-1/2} \neq \mathbf{0}\}\,\text{prob}\{\mathbf{u}_{2t}^t = \mathbf{0} | \mathbf{u}_{2t-1}^{t-1/2} \neq \mathbf{0}\} \\ &\leq 2t \times 2^{-2N} \,. \end{aligned} \tag{2.77}$$

And the same for $\mathbf{u}_{2t-1}^t$. Also, lemma 2.3 implies that $\mathbf{u}_{2t-1}^t$ and $\mathbf{u}_{2t}^t$ follow the probability distribution in equation 2.71 for all $t = 1/2, 1, 3/2, 2, \ldots$ □

With this proof, I will in the following lemma prove a similar result for the case where the initial binary vector $\mathbf{u}^0$ is non-zero at every site $\mathcal{V}_x$. This proof, along with the previous lemma, will then be used to generalise further to any arbitrary (non-zero) initial vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$.

Lemma 2.9. If the initial vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$ is non-zero on all lattice sites ($\mathbf{u}_x^0 \neq \mathbf{0}$

for all x) then the projection of its evolution $\mathbf{u}^t$ onto any site $x \in \mathbb{Z}_L$ satisfies

$$\text{prob}\{\mathbf{u}_x^t \neq \mathbf{0}\} \geq 1 - 16t\,2^{-N}, \tag{2.78}$$

for all $t \in \{1/2, 1, 3/2, \ldots\}$.

Proof. To prove this lemma we proceed similarly as in lemma 2.8. However, here, the recursive equation 2.76 need not have a $\mathbf{0}$-input in the right system

$$\mathbf{u}_{2t-1}^t \oplus \mathbf{u}_{2t}^t = S_{2t-1}(\mathbf{u}_{2t-1}^{t-1/2} \oplus \mathbf{u}_{2t}^{t-1/2}) \,. \tag{2.79}$$

This difference in the premises does not change the conclusion in equation 2.73, due to the fact that the bound in equation 2.32 is independent of $\mathbf{u}_1^0$ being zero or not. This gives equation 2.71 for $t = 1/2$. Also, using

$$\begin{aligned}
\text{prob}\{\mathbf{u}_x^1 = \mathbf{0}\} &= \text{prob}\{\mathbf{u}_1^{1/2} \oplus \mathbf{u}_2^{1/2} = \mathbf{0}\} + \text{prob}\{\mathbf{u}_1^{1/2} \oplus \mathbf{u}_2^{1/2} \neq \mathbf{0} \text{ and } \mathbf{u}_x^1 = \mathbf{0}\} \\
&\leq \text{prob}\{\mathbf{u}_1^{1/2} = \mathbf{0}\} + \text{prob}\{\mathbf{u}_x^1 = \mathbf{0} \,|\, \mathbf{u}_1^{1/2} \oplus \mathbf{u}_2^{1/2} \neq \mathbf{0}\} \\
&\leq 2 \times 2^{-2N} \,,
\end{aligned} \tag{2.80}$$

we obtain equation 2.71 for $t = 1$. However, here there is a very delicate point. As can be seen in figure 2.4, the vector $\mathbf{u}_2^1$ is partly determined by S_2, and hence, it is not independent. Crucially, the bound given in equation 2.75 for $\mathbf{u}_2^1$ holds regardless of the right input $\mathbf{u}_2^{1/2}$, and hence, it is independent of S_2. This fact can be summarized with the following bound

$$P(\mathbf{u}_2^1 | S_2) = \begin{cases} \frac{1-q_1}{2^{2N}-1} & \text{if } \mathbf{u}_2^1 \neq \mathbf{0} \\ q_1 & \text{if } \mathbf{u}_2^1 = \mathbf{0} \end{cases} \,, \tag{2.81}$$

for any S_2, where $q_1 \leq 2\,2^{-2N}$. That is, the correlation between $\mathbf{u}_2^1$ and S_2 can only happen through small variations of q_1.

For $t > 1$, the inputs in equation 2.79 are not independent of the matrix S_{2t-1}, as illustrated in figure 2.4, and hence, lemma 2.3 cannot be applied. If we restrict

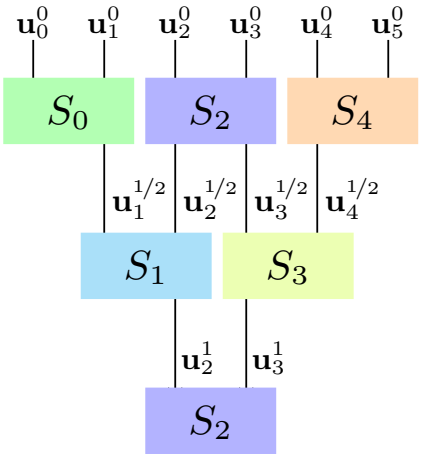

Figure 2.4: This figure shows that the causal past of $\mathbf{u}_2^1$ is partly determined by S_2. Hence, at $t = 3/2$, the input $\mathbf{u}_2^1$ of S_2 is not independent of S_2. This makes the exact probability distribution of $\mathbf{u}^t$ very complicated. However, we prove that $\mathbf{u}_2^1$ is approximately independent of S_2. This figure was produced by a collaborator, Daniele Toniolo.

equation 2.79 to the rightmost output ($x = 2t$) then we obtain

$$\begin{aligned}\mathbf{u}_{2t}^t &= C_{2t-1}\mathbf{u}_{2t-1}^{t-1/2} + D_{2t-1}\mathbf{u}_{2t}^{t-1/2} \\ &= C_{2t-1}\mathbf{u}_{2t-1}^{t-1/2} + \mathbf{v}^{t-1/2} \,, \end{aligned} \tag{2.82}$$

where the vector $\mathbf{v}^{t-1/2} = D_{t-1/2}\mathbf{u}_{2t}^{t-1/2} \in \mathbb{Z}_2^{2N}$ is not independent of $C_{t-1/2}$. Expanding this recursive relation we obtain

$$\begin{aligned}\mathbf{u}_{2t}^t &= C_{2t-1}C_{2t-2}\mathbf{u}_{2t-2}^{t-1} + C_{2t-1}\mathbf{v}^{t-1} + \mathbf{v}^{t-1/2} \\ &= C_{2t-1}\cdots C_2\mathbf{u}_2^1 + \mathbf{w}^t \,, \end{aligned} \tag{2.83}$$

where the random vector

$$\mathbf{w}^t = C_{2t-1}\cdots C_3\mathbf{v}^1 + \cdots + C_{2t-1}C_{2t-2}\mathbf{v}^{t-3/2} + C_{2t-1}\mathbf{v}^{t-1} + \mathbf{v}^{t-1/2} \tag{2.84}$$

is not independent of the matrices $C_{2t-1}, \ldots, C_2$. Crucially, the bound in equation 2.81 for the distribution of $\mathbf{u}_2^1$ is independent of all these matrices. This bound can

be equivalently stated as follows: with probability $q = q_1 - \frac{1-q_1}{2^{2N}-1}$ we have $\mathbf{u}_2^1 = \mathbf{0}$, and with probability $1-q$ the random variable $\mathbf{u}_2^1$ is uniformly distributed over all vectors in $\mathbb{Z}_2^{2N}$ including zero. This allows to write the following identity

$$\text{prob}\{\mathbf{u}_{2t}^t = \mathbf{0}\} = q + (1-q)\,\text{prob}\{C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t\}\ , \tag{2.85}$$

where the random variable $\mathbf{u} \in \mathbb{Z}_2^{2N}$ is uniform (including zero) and independent of $\mathbf{w}^t$ and C_i for all $i \in \{2t-1,\ldots,2\}$. Clearly, we can write

$$\text{prob}\{C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t\} = \mathop{\mathbb{E}}_{C_i,\mathbf{w}^t}\mathop{\mathbb{E}}_{\mathbf{u}}\delta\left[C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t\right]\ . \tag{2.86}$$

Now consider the average $\mathbb{E}_{\mathbf{u}}\,\delta[C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t]$ for a fixed value of the variables $\mathbf{w}^t$ and C_i. If the vector $\mathbf{w}^t$ is not in the range of the matrix $(C_{2t-1}\cdots C_2)$ then the average is zero. If the vector $\mathbf{w}^t$ is in the range of the matrix $(C_{2t-1}\cdots C_2)$ then there is a vector $\tilde{\mathbf{w}}$ such that $\mathbf{w}^t = (C_{2t-1}\cdots C_2)\tilde{\mathbf{w}}$. Then we can write the average as

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{u}}\delta\left[C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t\right] &= \mathop{\mathbb{E}}_{\mathbf{u}}\delta[C_{2t-1}\cdots C_2(\mathbf{u}+\tilde{\mathbf{w}}) = \mathbf{0}] &(2.87)\\ &= \mathop{\mathbb{E}}_{\mathbf{u}}\delta[C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{0}]\ , &(2.88)\end{aligned}$$

where the last equality follows from the fact that the random variable $\mathbf{u}+\tilde{\mathbf{w}}$ is uniform and independent of C_i, likewise $\mathbf{u}$. Combining together the two cases for $\mathbf{w}^t$ we can write

$$\begin{aligned}\text{prob}\{C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{w}^t\} &\le \mathop{\mathbb{E}}_{C_i}\mathop{\mathbb{E}}_{\mathbf{u}}\delta[C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{0}]\\ &= \text{prob}\{C_{2t-1}\cdots C_2\mathbf{u} = \mathbf{0}\}\\ &\le 8(2t-2)2^{-N}\ , &(2.89)\end{aligned}$$

where the last step follows from lemma 2.6. Substituting this back into equation 2.85 and using $q \le (2^{2N}-1)^{-1}$ we obtain

$$\text{prob}\{\mathbf{u}_{2t}^t = \mathbf{0}\} \le q + (1-q)16\,(t-1)2^{-N} \le 16t\,2^{-N}\ . \tag{2.90}$$

It is worth noting that lemma 2.6 implies that the same bound also holds for $\mathbf{u}^t_{2t-1}$. Also, since the premises of this lemma are invariant under translations in the chain $\mathbb{Z}_L$, then the conclusions hold for all $x \in \mathbb{Z}_L$. □

The previous two proofs analyse two specific cases of the initial vector $\mathbf{u}^0$. In the following lemma, I will prove a similar result for the more general case in which the initial vector $\mathbf{u}^0 \in \mathbb{Z}_2^{2NL}$ is any non-zero vector. Also, I would like to comment that while I did contribute significantly to the proof of the following lemma, the proof in its current form I primarily credit to my supervisor, Lluis Masanes, and collaborator, Daniele Toniolo.

Lemma 2.10. After the scrambling time $t \in [t_{\text{scr}}, 2t_{\text{scr}}]$, with t integer or half-integer, the evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ is non-zero at each lattice site with probability

$$\text{prob}\{\mathbf{u}^t_x \neq \mathbf{0}, \forall x \in \mathbb{Z}_L\} \geq 1 - 16tL2^{-N}, \tag{2.91}$$

for any initial non-zero vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$.

Proof. Let $\mathcal{F}(\mathbf{u}^0) \subseteq \mathbb{Z}_L \times \mathbb{N}$ be the set of spacetime points consisting of the causal future of the sites $x' \in \mathbb{Z}_L$ where the initial vector $\mathbf{u}^0$ has support ($\mathbf{u}^0_{x'} \neq \mathbf{0}$). For example, if the initial vector is supported in the origin of the chain $\mathbf{u}^0 \in \mathcal{V}_0$ then the causal future is given by the light cone (equation 2.70).

The main objective in this proof is to bound the probability of $\mathbf{u}^t_x \neq \mathbf{0}$ for any fixed site $x \in \mathbb{Z}_L$ and time $t \in [t_{\text{scr}}, 2t_{\text{scr}}]$. For the sake of simplicity, let us start by considering the case of x odd and t integer. In this case, the left-most spacetime points in the causal past of (x,t) that are also contained in $\mathcal{F}(\mathbf{u}^0)$ are

$$(x-1, t-1/2), \ldots, (x-n, t-n/2), \ldots, (x_{\text{e}}, t_{\text{e}}) \,. \tag{2.92}$$

We have that either $t_{\text{e}} = 0$ or $t_{\text{e}} > 0$. In the first case ($t_{\text{e}} = 0$) we have that $\mathbf{u}^0$ has support on x_{e} or $x_{\text{e}} + 1$. And we can prove

$$\text{prob}\{\mathbf{u}^t_x = \mathbf{0}\} \leq 16t\, 2^{-N} \,, \tag{2.93}$$

by applying the same procedure as in lemma 2.9. It is worth noting that the possibility that $\mathbf{u}^0_{x'} = \mathbf{0}$ for $x' > x_\mathrm{e} + 1$ does not affect the argument.

In the second case ($t_\mathrm{e} > 0$), the sequence 2.92 can be continued by including the following points from $\mathcal{F}(\mathbf{u}^0)$,

$$(x_\mathrm{e}, t_\mathrm{e} - 1/2), \ldots, (x_\mathrm{e}+n, t_\mathrm{e} - 1/2 - n/2), \ldots, (x_0 - 1, 1/2), \begin{cases} (x_0 - 1, 0) \\ (x_0, 0) \end{cases} , \quad (2.94)$$

where the last element is chosen so that it belongs to $\mathcal{F}(\mathbf{u}^0)$. If $\mathbf{u}^0$ has support on both $(x_0 - 1, 0)$ and $(x_0, 0)$ then the choice is arbitrary. Here, for the sake of concreteness, we assume that $\mathbf{u}^0_{x_0} \neq \mathbf{0}$ and take $(x_0, 0)$ as the last point of the sequence. The sub-index "e" stands for "elbow", because it labels the point where the sequence 2.92 changes direction to the sequence 2.94 (see figure 2.5).

Now we can write our chosen vector $\mathbf{u}^t_x$ as

$$\mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} = B_{x_\mathrm{e}} \cdots B_{x_0-2} B_{x_0-1} \mathbf{u}^0_{x_0} , \quad (2.95)$$

$$\mathbf{u}^{t_\mathrm{e}}_{x_\mathrm{e}} = D_{x_\mathrm{e}-1} \mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} , \quad (2.96)$$

$$\mathbf{u}^t_x = C_{x-1} \cdots C_{x_\mathrm{e}+1} C_{x_\mathrm{e}} \mathbf{u}^{t_\mathrm{e}}_{x_\mathrm{e}} + \mathbf{w} , \quad (2.97)$$

where the random vector $\mathbf{w}$ is correlated with $B_{x_\mathrm{e}}, \ldots, B_{x_0-1}$ and $C_{x-1}, \ldots, C_{x_\mathrm{e}}$ but not with $D_{x_\mathrm{e}-1}$. Vector $\mathbf{w}$ is analogous to $\mathbf{w}^t$, defined in equation 2.84. Note also that the random matrices $B_{x_\mathrm{e}}, \ldots, B_{x_0-1}$ are not independent from $C_{x-1}, \ldots, C_{x_\mathrm{e}}$, but that $D_{x_\mathrm{e}-1}$ is independent from all the rest. (Figure 2.5 contains an example where the gates associated to $B_{x_\mathrm{e}}, \ldots, B_{x_0-1}$ are in blue, those of $C_{x-1}, \ldots, C_{x_\mathrm{e}}$ in red, and that of $D_{x_\mathrm{e}-1}$ in yellow.)

Now we can start constructing our bound as

$$\begin{aligned} \mathrm{prob}\{\mathbf{u}^t_x = \mathbf{0}\} &= \mathrm{prob}\{\mathbf{u}^t_x = \mathbf{0} \text{ and } \mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} = \mathbf{0}\} + \mathrm{prob}\{\mathbf{u}^t_x = \mathbf{0} \text{ and } \mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} \neq \mathbf{0}\} \\ &\leq \mathrm{prob}\{\mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} = \mathbf{0}\} + \mathrm{prob}\{\mathbf{u}^t_x = \mathbf{0} \text{ and } \mathbf{u}^{t_\mathrm{e}-1/2}_{x_\mathrm{e}} \neq \mathbf{0}\} \quad (2.98) \end{aligned}$$

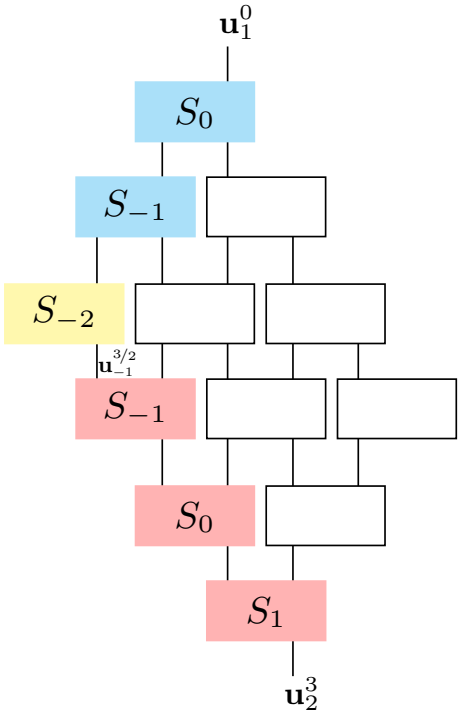

Figure 2.5: This figure represents the evolution of an initially local operator $\mathbf{u}^0 \in \mathcal{V}_1$ at site $x = 1$. The figure only displays gates S_x that are in the intersection of the causal future of the initial location $x = 1$ and the causal past of the chosen point $\mathbf{u}_2^3$. The probability of $\mathbf{u}_2^3 = \mathbf{0}$ is bounded by analysing the sequence of coloured gates, which has an "elbow" at location $(x_e, t_e) = (-1, 3/2)$. The analysis of blue gates uses lemma 2.8, and that of red gates uses lemma 2.9. The key feature of the bound is that the yellow gate S_{-2} only appears once. This figure was produced by a collaborator, Daniele Toniolo.

Using lemma 2.8, the first term can be upper-bounded by

$$\text{prob}\{\mathbf{u}_{x_e}^{t_e-1/2} = \mathbf{0}\} \leq 2(t_e - 1/2)2^{-2N} \,. \tag{2.99}$$

The second term can be bounded by using the independence of D_{x_e-1}, the fact that $\mathbf{u}_{x_e}^{t_e-1/2}$ is not zero, and proceeding in a manner similar to lemma 2.9. Therefore, we again introduce the uniformly distributed random vector $\mathbf{u} \in \mathbb{Z}_2^{2N}$, which is independent of all gates $S_{x_e}, S_{x_e+1}, \ldots, S_{x-1}$.

The statistical distance between $\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}}$ and $\mathbf{u}$, conditioned on $\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}-1/2} \neq \mathbf{0}$, is

$$\begin{aligned}
\mathrm{d}(\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}}, \mathbf{u}) &= \sum_{\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}}} \left| P(\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}} | \mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}-1/2} \neq \mathbf{0}) - 2^{-2N} \right| \\
&= \left(2^{2N}-1\right) \left| \frac{2^{2N}}{2^{4N}-1} - 2^{-2N} \right| + \left| \frac{2^{2N}-1}{2^{4N}-1} - 2^{-2N} \right| \\
&\leq 2^{1-4N} \,, \qquad (2.100)
\end{aligned}$$

where we have used lemma 2.3. Once again, continuing using the proof technique of lemma 2.9 (in particular equation 2.85) we find that

$$\begin{aligned}
&\mathrm{prob}\{\mathbf{u}_x^t = \mathbf{0} \text{ and } \mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}-1/2} \neq \mathbf{0}\} \\
&= \mathrm{prob}\{C_{x-1} \cdots C_{x_\mathrm{e}} \mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}} = \mathbf{w} \text{ and } \mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}-1/2} \neq \mathbf{0}\} \\
&\leq \mathrm{d}(\mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}}, \mathbf{u}) + \mathrm{prob}\{C_{x-1} \cdots C_{x_\mathrm{e}} \mathbf{u} = \mathbf{w} \text{ and } \mathbf{u}_{x_\mathrm{e}}^{t_\mathrm{e}-1/2} \neq \mathbf{0}\} \\
&\leq 2^{1-4N} + \mathrm{prob}\{C_{x-1} \cdots C_{x_\mathrm{e}} \mathbf{u} = \mathbf{w}\} \,. \qquad (2.101)
\end{aligned}$$

This upper-bound makes use of the fact that that $\mathbf{w}$ and $\mathbf{u}$ are independent, giving

$$\mathrm{prob}\{C_{x-1} \cdots C_{x_\mathrm{e}} \mathbf{u} = \mathbf{w}\} \leq 8\,(x - x_\mathrm{e})\, 2^{-N} = 8\,(t - t_\mathrm{e})\, 2^{-N} \,. \qquad (2.102)$$

Putting all things together we obtain

$$\mathrm{prob}\{\mathbf{u}_x^t = \mathbf{0}\} \leq 2(t_\mathrm{e} - 1/2) 2^{-2N} + 2^{1-4N} + 8\,(t - t_\mathrm{e})\, 2^{-N} \leq 8\, t\, 2^{-N} \,. \qquad (2.103)$$

Finally, we use the union bound to conclude that

$$\mathrm{prob}\{\exists x \in \mathbb{Z}_L : \mathbf{u}_x^t = \mathbf{0}\} \leq 8\, t\, L\, 2^{-N} \,, \qquad (2.104)$$

which is equivalent to the statement in equation 2.91. □

The result in lemma 2.10 demonstrates that in the regime $N \gg \log L$ an initial non-zero vector $\mathbf{u}^0$ is mapped by the time evolution operator to an evolved vector $\mathbf{u}^t$ which is non-zero at every site of the lattice (with probability close to one). This

lemma is key for the proof of the main results of this section.

Half-integer time results: proof of results 2.3 and 2.4

With the mathematical tools developed in this section so far, I will now show the rigorous proofs for results 2.3 and 2.4. These results concern half-integer times only. I am focussing on the proof of the half-integer time results first, because they require fewer steps than the integer time results, but use the same approach. Therefore, the proof of the half-integer time results can serve as an informative guide to the more involved results for integer times, which require a few extra lemmas and proofs.

In the next lemma, I will prove that for any evolved vector $\mathbf{u}^t$ which is non-zero at every site the probability distribution is uniformly distributed over all possible vectors, for any half-integer time after the scrambling time.

Lemma 2.11. At half-integer $t \geq t_{\text{scr}}$ the probability distribution of the evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ conditioned on the evolved vector being non-zero at every site is uniform:

$$\text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}^t_x \neq \mathbf{0}, \forall x \in \mathbb{Z}_L\} = \frac{1}{(2^{2N}-1)^L}\ , \tag{2.105}$$

for all vectors $\mathbf{v}$ that are non-zero at every site $\mathbf{v}_x \neq \mathbf{0}, \forall x \in \mathbb{Z}_L$.

Proof. The proof of this lemma follows from the twirling technique discussed and proven in lemma 2.7. The probability distribution of the evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ has identical statistical properties to

$$\mathbf{u}^t = \left(\bigoplus_{x=0}^{L-1} X_x\right) S(t) \left(\bigoplus_{x=0}^{L-1} Y_x\right) \mathbf{u}^0, \tag{2.106}$$

where $X_x, Y_x \in \mathcal{S}_N$ are arbitrary single-site symplectic matrices. Hence, since the choice of each X_x is arbitrary, each X_x is independent and uniformly distributed over all single-site symplectic matrices. Therefore, imposing the condition that the evolved vector is non-zero on every site, then, since the twirling matrices X_x are independent and uniform, the probability distribution of the evolved vector at each site is independent and uniformly distributed over all non-zero vectors. Applying lemma 2.3 the stated conditional probability in equation 2.105 is obtained. □

With this uniformly distributed conditional probability, I will now give the proof of result 2.3, which states that at any half-integer time in the interval $t \in [t_{\text{scr}}, 2t_{\text{scr}}]$ the probability distribution of the evolved operator is close to uniform, when $N \gg \log L$.

Lemma 2.12. For any initial non-zero vector $\mathbf{u}^0 \in \mathcal{V}_{\text{chain}}$, the probability distribution of the time evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$, at any half-integer time after the scrambling time, is approximately uniformly distributed over all non-zero vectors of the total system, and bounded by

$$\sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v}\} - \frac{1}{2^{2NL}-1} \right| \leq 32tL2^{-N} + L2^{-2N} \,. \tag{2.107}$$

Proof. Defining $q \equiv \text{prob}\{\mathbf{u}^t_x \neq \mathbf{0}, \forall x \in \mathbb{Z}_L\}$, the probability $\text{prob}\{\mathbf{u}^t = \mathbf{v}\}$ can be rewritten in the following way

$$\begin{aligned}\text{prob}\{\mathbf{u}^t = \mathbf{v}\} = q\,\text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}^t_x \neq \mathbf{0}, \forall x \in \mathbb{Z}_L\} \\ + (1-q)(\text{prob}\{\mathbf{u}^t = \mathbf{v} | \exists y \in \mathbb{Z}_L : \mathbf{u}^t_y = \mathbf{0}\}).\end{aligned}$$

Adding and subtracting $\frac{q}{2^{2NL}-1}$ in the sum and then applying the triangle inequality, we find that:

$$\begin{aligned}\sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v}\} - \frac{1}{2^{2NL}-1} \right| \leq \sum_{\mathbf{v}} q \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}^t_x \neq \mathbf{0}\, \forall x\} - \frac{1}{2^{2NL}-1} \right| \\ + (1-q) \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \exists y \in \mathbb{Z}_L : \mathbf{u}^t_y = \mathbf{0}\} - \frac{1}{2^{2NL}-1} \right| \,.\end{aligned}$$

We can upper bound the first term with $q \leq 1$ and apply lemma 2.11 to find that

$$q \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}^t_x \neq \mathbf{0}\, \forall x\} - \frac{1}{2^{2NL}-1} \right| \leq L2^{-2N} \,. \tag{2.108}$$

To bound the second term, notice that the maximum value of the sum is 2, so

$$\begin{aligned}&(1-q) \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \exists y \in \mathbb{Z}_L : \mathbf{u}^t_y = \mathbf{0}\} - \frac{1}{2^{2NL}-1} \right| \\ &\leq (1-q) \sum_{\mathbf{v}} \left(\text{prob}\{\mathbf{u}^t = \mathbf{v} | \exists y \in \mathbb{Z}_L : \mathbf{u}^t_y = \mathbf{0}\} + \frac{1}{2^{2NL}-1} \right) = 2(1-q)\end{aligned}$$

and then using the result of lemma 2.10 to find that

$$(1-q) \leq 16tL2^{-N} , \tag{2.109}$$

gives the stated result. □

This concludes the proof of result 2.3. In the next lemma, I will give the proof of result 2.4, which shows that the time evolution operator $W(t)$ is an approximate 2-design in the weak sense that I defined - when measurements are are restricted to Pauli operators only.

Lemma 2.13. At half integer times after the scrambling time, $t \geq t_{scr}$, the dynamics $W(t)$ give rise to an approximate 2-design with respect to two-Pauli measurements. That is to say, for any state ρ the following holds

$$\mathrm{tr}\left(\sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{v}} \left[\mathop{\mathbb{E}}_{W(t)} W(t)^{\otimes 2} \rho\, W(t)^{\otimes 2\dagger} - \int_{\mathrm{SU}(d)} dU\, U^{\otimes 2} \rho\, U^{\otimes 2\dagger} \right] \right) \leq 33tL2^{-N} \delta_{\mathbf{u},\mathbf{v}} . \tag{2.110}$$

Proof. Let us consider a general state describing two copies of the system

$$\rho = \sum_{\mathbf{u},\mathbf{v}} \alpha_{\mathbf{u},\mathbf{v}}\, \sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{v}} , \tag{2.111}$$

where $\alpha_{0,0} = 2^{-2NL}$ by normalisation, and the coefficients $\alpha_{\mathbf{u},\mathbf{v}}$ must satisfy the following

$$\alpha_{\mathbf{u},\mathbf{v}}\, 2^{2NL} = \mathrm{tr}(\rho\, \sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{v}}) \in [-1,1] . \tag{2.112}$$

Applying the average dynamics to ρ we obtain

$$\mathop{\mathbb{E}}_{W(t)} W(t)^{\otimes 2} \rho\, W(t)^{\otimes 2\dagger} = 2^{-2NL} \mathbb{1} \otimes \mathbb{1} + \sum_{\mathbf{u},\mathbf{v} \neq 0} \alpha_{\mathbf{v},\mathbf{v}} \mathrm{prob}\{\mathbf{v} = S(t)\mathbf{u}\} \sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{u}} . \tag{2.113}$$

The fact that terms $\alpha_{\mathbf{u},\mathbf{u}'}$ and $\sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{u}'}$ with $\mathbf{u} \neq \mathbf{u}'$ are not present in the above expression follows from the fact that $W(t)$ has an additional symmetry at half-integer times (referred to as the twirling technique), which is described and proven in lemma

2.7. Recall that at half-integer t we have the time-reversal symmetry

$$\text{prob}\{\mathbf{v} = S(t)\mathbf{u}\} = \text{prob}\{\mathbf{u} = S(t)\mathbf{v}\} \ . \tag{2.114}$$

Applying the Haar twirling on ρ we obtain

$$\int_{\text{SU}(d)} dU\, U^{\otimes 2} \rho\, U^{\otimes 2\dagger} = 2^{-2NL} \mathbb{1} \otimes \mathbb{1} + \sum_{\mathbf{u},\mathbf{v}\neq 0} \alpha_{\mathbf{v},\mathbf{v}}\, \gamma\, \sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{u}} \ , \tag{2.115}$$

where $\gamma = (2^{2NL} - 1)^{-1}$ is the uniform distribution over non-zero vectors in $\mathcal{V}_{\text{chain}}$. Substituting equations 2.113 and 2.115 into equation 2.110 we obtain

$$\text{tr}\left(\sigma_{\mathbf{u}} \otimes \sigma_{\mathbf{v}} \left[\mathop{\mathbb{E}}_{W(t)} W(t)^{\otimes 2} \rho\, W(t)^{\otimes 2\dagger} - \int_{\text{SU}(d)} dU\, U^{\otimes 2} \rho\, U^{\otimes 2\dagger} \right] \right) \tag{2.116}$$

$$= \delta_{\mathbf{u},\mathbf{v}} \sum_{\mathbf{w}\neq 0} \alpha_{\mathbf{w},\mathbf{w}} \left(\text{prob}\{\mathbf{u} = S(t)\mathbf{w}\} - \gamma\right) 2^{2NL} \tag{2.117}$$

$$\leq \delta_{\mathbf{u},\mathbf{v}} \sum_{\mathbf{w}\neq 0} \left|\text{prob}\{\mathbf{u} = S(t)\mathbf{w}\} - \gamma\right| \leq 33 t L 2^{-N} \delta_{\mathbf{u},\mathbf{v}} \ , \tag{2.118}$$

where the last two inequalities follow from equations 2.112 and 2.114, and lemma 2.12. □

Integer and half-integer time results: proof of results 2.1 and 2.2

With the mathematical tools proven previously in this section, I will now show the rigorous proofs for results 2.1 and 2.2. These results apply to both integer and half-integer times. The proof of these two results make use of the same tools and argument as the proofs for the two results that apply to half-integer times only. However, a few extra and more involved proofs will be required, since at integer times the time evolution operator has a more restricted additional symmetry than in the half-integer case - as proven in lemma 2.7.

First, I will prove a general result (lemma 2.14) , which is not specific to integer times (or quantum circuits), but is primarily used in the all proofs related to integer time results. Then, I will give a pair of proofs (lemma 2.15 and 2.16), which will conclude with the proof justifying result 2.1. After, I will then proceed to the proof

justifying result 2.2 (lemma 2.19), which first requires an involved derivation (lemma 2.18).

Lemma 2.14. Let $\mathbf{u}$ be a fixed non-zero element of $\mathbb{Z}_2^{2N}$. Let the probability distribution $P(\mathbf{v})$ over $\mathbf{v} \in \mathbb{Z}_2^{2N}$ have the property that $P(S\mathbf{v}) = P(\mathbf{v})$ for any $S \in \mathcal{S}_N$ such that $S\mathbf{u} = \mathbf{u}$. Then it must be of the form

$$P(\mathbf{v}) = \begin{cases} q_1 & \text{if } \mathbf{v} = \mathbf{0} \\ q_2 & \text{if } \mathbf{v} = \mathbf{u} \\ q_3 & \text{if } \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ and } \mathbf{v} \neq \mathbf{0}, \mathbf{u} \\ q_4 & \text{if } \langle \mathbf{v}, \mathbf{u} \rangle = 1 \end{cases} , \tag{2.119}$$

where the positive numbers q_i are constrained by the normalization of $P(\mathbf{v})$.

Proof. We initially consider that $\mathbf{u} = (1, 0, \ldots, 0)^T$, and the subgroup of $\mathcal{S}_n$ that leaves $\mathbf{u}$ unchanged. If $\mathbf{v} = \mathbf{0}$ or $\mathbf{u}$, then the action of this subgroup has no effect, and hence we require a parameter for each in the distribution, q_1 and q_2 respectively. This is not the case for all other choices of $\mathbf{v}$, since the action of the subgroup will transform $\mathbf{v}$ into some other vector in $\mathbb{Z}_2^{2N}$. This transformation is constrained by the symplectic form:

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle S\mathbf{v}, S\mathbf{u} \rangle = \langle S\mathbf{v}, \mathbf{u} \rangle , \tag{2.120}$$

and hence the subgroup is composed of two subgroups, which transform $\mathbf{v}$ into another vector in $\mathbb{Z}_2^{2N}$ that has the same value for the symplectic form. Furthermore, the two subgroups are such they can map any vector to any other vector with the same value for the symplectic form.

This can be seen by considering the case where $\mathbf{v} = (0, 1, \ldots, 0)^T$, so $\langle \mathbf{u}, \mathbf{v} \rangle = 1$. The subgroup that keeps $\mathbf{u} = (1, 0, ..., 0)^T$ unchanged consists of all the elements of $\mathcal{S}_N$ with $\mathbf{u}$ as the first column of the matrix. Hence, by lemma 2.1, we can select the second column of the matrix to be any vector which has symplectic form of 1 with the first column, which is $\mathbf{u}$. Thus, we can map $\mathbf{v}$ to any other vector with symplectic form one with $\mathbf{u}$, which is also unchanged. Then, by noting that the product of symplectic matrices is a symplectic matrix, the subgroup can map any vector with

symplectic form of one with $\mathbf{u}$ to any other. Similarly, this argument applies to the other case where the symplectic form has a value of zero.

Then since $P(S\mathbf{v}) = P(\mathbf{v})$, all vectors that give the same value for $\langle \mathbf{v}, \mathbf{u} \rangle$ have the same probability. Thus, we get the probability distribution in equation 2.119.

Finally, since via a symplectic transformation $\mathbf{u}$ can be mapped to any other vector in $\mathbb{Z}_2^{2N}$, and that the product of two symplectic matrices is symplectic, this result applies for any $\mathbf{u} \in \mathbb{Z}_2^{2N}$. □

This rather general result demonstrates that at integer times the symplectic form of the initial vector at one particular site $\mathbf{u}_x^0$ with the evolved vector at the same particular site $\mathbf{u}_x^t$ is a crucial factor in the form of the probability distribution. In the next lemma, I will address the question of bounding the symplectic form for the case of an initially local operator $\mathbf{u}^0$ and its time evolution $\mathbf{u}^t$.

Lemma 2.15. For an initial vector $\mathbf{u}^0 \in \mathcal{V}_0$ which is non-zero only at site $x = 0$, the probability that the value of the symplectic form between the evolved vector $\mathbf{u}^t = S_{\text{chain}}^t \mathbf{u}^0$, with integer t, and the initial vector, $\langle \mathbf{u}^t, \mathbf{u}^0 \rangle = \langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle$ is equal to s has an s-independent upper bound given by:

$$\text{prob}\{\langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle = s\} \leq 8t2^{-N} + \frac{1}{2} \,. \tag{2.121}$$

Furthermore, this result is independent of the initial site at which the initial vector $\mathbf{u}^0$ in non-zero.

Proof. The approach of this proof is similar to the proof of lemma 2.9. That is to say, we consider a sequence of gates which are in the causal past of $\mathbf{u}_0^t$, as shown for example in figure 2.5. Mathematically, we write $\mathbf{u}_0^t$ as

$$\mathbf{u}_{1-t}^{t/2} = B_{1-t} \cdots B_{-2} B_{-1} A_0 \mathbf{u}_0^0 \,, \tag{2.122}$$

$$\mathbf{u}_{1-t}^{t/2+1/2} = D_{-t} \mathbf{u}_{1-t}^{t/2} \,, \tag{2.123}$$

$$\mathbf{u}_0^t = C_{-1} \cdots C_{2-t} C_{1-t} \mathbf{u}_{1-t}^{t/2+1/2} + \mathbf{w} \,, \tag{2.124}$$

where, crucially, the random vector $\mathbf{w}$ is independent of the random matrix D_{-t}. This

vector $\mathbf{w}$ is defined in a way similar to equation 2.84 (of lemma 2.9). To proceed further, we can follow similar steps to those in equations 2.98 to 2.103, so

$$\begin{aligned}
&\text{prob}\big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s\big\} \\
&= \text{prob}\Big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s \text{ and } \mathbf{u}_{1-t}^{t/2} = \mathbf{0}\Big\} + \text{prob}\Big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s \text{ and } \mathbf{u}_{1-t}^{t/2} \neq \mathbf{0}\Big\} \\
&\leq \text{prob}\Big\{\mathbf{u}_{1-t}^{t/2} = \mathbf{0}\Big\} + \text{prob}\Big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s \text{ and } \mathbf{u}_{1-t}^{t/2} \neq \mathbf{0}\Big\} \ . \qquad (2.125)
\end{aligned}$$

The first term can be upper-bounded using lemma 2.8, so

$$\text{prob}\Big\{\mathbf{u}_{1-t}^{t/2} = \mathbf{0}\Big\} \leq 2t\, 2^{-2N} \ . \qquad (2.126)$$

To upper-bound the second term, following the method of lemma 2.9, we introduce two uniformly distributed random vectors $\mathbf{u}, \mathbf{u}' \in \mathbb{Z}_2^{2N}$, which importantly are independent of the gates $S_{1-t}, S_{2-t}, \ldots, S_{-2}$. Hence, this probability term can be written as

$$\begin{aligned}
&\text{prob}\Big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s \text{ and } \mathbf{u}_{1-t}^{t/2} \neq \mathbf{0}\Big\} \\
&= \text{prob}\Big\{\mathbf{u}_0^{0T} J C_{-1} \cdots C_{1-t} \mathbf{u}_{1-t}^{t/2+1/2} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \text{ and } \mathbf{u}_{1-t}^{t/2} \neq \mathbf{0}\Big\} \\
&\leq \text{d}\Big(\mathbf{u}_{1-t}^{t/2+1/2}, \mathbf{u}\Big) + \text{prob}\Big\{\mathbf{u}_0^{0T} J C_{-1} \cdots C_{1-t} \mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \text{ and } \mathbf{u}_{1-t}^{t/2} \neq \mathbf{0}\Big\} \\
&\leq \text{d}\Big(\mathbf{u}_{1-t}^{t/2+1/2}, \mathbf{u}\Big) + \text{prob}\big\{\mathbf{u}_0^{0T} J C_{-1} \cdots C_{1-t} \mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle\big\} \\
&\leq \text{d}\Big(\mathbf{u}_{1-t}^{t/2+1/2}, \mathbf{u}\Big) + \text{d}\big(J C_{-1}^T J \mathbf{u}_0^0, \mathbf{u}'\big) + \text{prob}\big\{\mathbf{u}'^T J C_{-2} \cdots C_{1-t} \mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle\big\} \ , \\
&\qquad (2.127)
\end{aligned}$$

where $\text{d}(\cdot,\cdot)$ denotes the statistical distance between the two probability distribution. These two statistical distance terms can be bounded in the same way as equation 2.100, so

$$\text{d}\Big(\mathbf{u}_{1-t}^{t/2+1/2}, \mathbf{u}\Big) \leq 2^{1-4N} \ , \qquad (2.128)$$

$$\text{d}\big(J C_{-1}^T J \mathbf{u}_0^0, \mathbf{u}'\big) \leq 2^{1-4N} \ . \qquad (2.129)$$

To bound the third term of equation 2.127, note that for any non-zero $\mathbf{a} \in \mathbb{Z}_2^{2N}$ then $\text{prob}\{\mathbf{u}'^T\mathbf{a} = s\} = 1/2$, for both $s = 0, 1$, and hence therefore

$$\begin{aligned}
&\text{prob}\big\{\mathbf{u}'^T JC_{-2}\cdots C_{1-t}\mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle\big\} \\
&= \text{prob}\big\{\mathbf{u}'^T JC_{-2}\cdots C_{1-t}\mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \text{ and } C_{-2}\cdots C_{1-t}\mathbf{u} \neq \mathbf{0}\big\} \\
&+ \text{prob}\big\{\mathbf{u}'^T JC_{-2}\cdots C_{1-t}\mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \text{ and } C_{-2}\cdots C_{1-t}\mathbf{u} = \mathbf{0}\big\} \ .
\end{aligned}$$

Next we bound the first term by using the fact that the uniformly distributed vector $\mathbf{u}'$ is independent of $\mathbf{a} := JC_{-2}\cdots C_{1-t}\mathbf{u}$ and $\langle \mathbf{u}_0^0, \mathbf{w}\rangle$, as

$$\text{prob}\big\{\mathbf{u}'^T\mathbf{a} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \text{ and } \mathbf{a} \neq \mathbf{0}\big\} \leq \text{prob}\big\{\mathbf{u}'^T\mathbf{a} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle \big| \, \mathbf{a} \neq \mathbf{0}\big\} = \frac{1}{2} \ .$$

The second term can be upper-bounded by applying lemma 2.6 to find Combining these bounds, it is found that

$$\text{prob}\big\{\mathbf{u}'^T JC_{-2}\cdots C_{1-t}\mathbf{u} = s + \langle \mathbf{u}_0^0, \mathbf{w}\rangle\big\} \ \leq \ \frac{1}{2} + 8(t-2)2^{-N} \ . \tag{2.130}$$

Finally, with all of the terms accounted for and upper-bounded, we arrive at

$$\begin{aligned}
\text{prob}\big\{\langle \mathbf{u}_0^0, \mathbf{u}_0^t\rangle = s\big\} &\leq 2t\,2^{-2N} + 4\,2^{-4N} + \frac{1}{2} + 8(t-2)2^{-N} \\
&\leq \frac{1}{2} + 8t\,2^{-N},
\end{aligned} \tag{2.131}$$

which is the stated result. □

By combining the previous two results (lemmas 2.14 and 2.15), I will now prove a similar statement to lemma 2.11. That is to say, for an initially local operator $\mathbf{u}^0$ and an evolved vector $\mathbf{u}^t$, which is non-zero at every site, the probability distribution of the evolved vector $\mathbf{u}^t$ is uniformly distributed over all possible vectors (within the light-cone).

Lemma 2.16. For an initial vector $\mathbf{u}^0$ which is non-zero only on site $x = 0$ the probability distribution of the evolved vector $\mathbf{u}^t = S_{\text{chain}}^t\mathbf{u}^0$ at integer times, conditioned on the evolved vector being non-zero at every site and different from the initial

single-site non-zero vector, $\mathbf{u}_x^t \neq 0 \; \forall x \in \mathbb{Z}_L$ and $\mathbf{u}_0^t \neq \mathbf{u}_0^0$, after the scrambling time t_{scr} is of the form

$$\text{prob}\{\mathbf{u}^t | \mathbf{u}_x^t \neq 0 \; \forall x \in \mathbb{Z}_L, \mathbf{u}_0^t \neq \mathbf{u}_0^0\} \leq \frac{1}{(2^{2N}-1)^{L-1}} \begin{cases} \frac{8t2^{-N}+1/2}{2^{2N-1}-2} & \text{if } \langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle = 0 \\ \frac{8t2^{-N}+1/2}{2^{2N-1}} & \text{if } \langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle = 1 \end{cases},$$

and before the scrambling time for all sites within the causal light-cone the probability distribution is of the form

$$\begin{aligned} &\text{prob}\{\mathbf{u}^t | \mathbf{u}_x^t \neq 0 \; \forall x \in [-2t+1, 2t], \mathbf{u}_0^t \neq \mathbf{u}_0^0\} \\ &\qquad \leq \frac{1}{(2^{2N}-1)^{4t-1}} \begin{cases} \frac{8t2^{-N}+1/2}{2^{2N-1}-2} & \text{if } \langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle = 0 \\ \frac{8t2^{-N}+1/2}{2^{2N-1}} & \text{if } \langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle = 1 \end{cases}. \end{aligned} \tag{2.132}$$

Furthermore, this result is independent of the initial site which the initial vector $\mathbf{u}^0$ in non-zero.

Proof. The proof of this lemma uses the twirling technique discussed in lemma 2.7. The probability distribution of the evolved vector $\mathbf{u}^t = (S_{\text{chain}})^t \mathbf{u}^0$ at integer times has identical statistical properties to

$$\begin{aligned} \mathbf{u}^t &= \left(\bigoplus_{x=0}^{L-1} X_x\right) S_{\text{chain}}^t \left(\bigoplus_{x=0}^{L-1} X_x^{-1}\right) \mathbf{u}^0 \\ &= \left(\bigoplus_{x=0}^{L-1} X_x\right) S_{\text{chain}}^t \left(X_0^{-1}\mathbf{u}_0^0 \bigoplus_{x=1}^{L-1} \mathbf{0}\right) \end{aligned} \tag{2.133}$$

where $X_x \in \mathcal{S}_N$ are arbitrary single-site symplectic matrices. Equation 2.133 follows from the fact that it has been assumed that $\mathbf{u}^0$ is non-zero at $x = 0$, and therefore $\left(\bigoplus_{x=0}^{L-1} X_x^{-1}\right) \mathbf{u}^0$ is supported at $x = 0$ as well. If we restrict X_0 to the elements of $\mathcal{S}_N$ that satisfy $X_0 \mathbf{u}_0^0 = \mathbf{u}_0^0$, then the probability distribution of $\mathbf{u}^t$ is identical to $\left(\bigoplus_{x=0}^{L-1} X_x\right) \mathbf{u}^t$. Since the choice of symplectic matrices $\bigoplus_{x=0}^{L-1} X_x$ with which to twirl is arbitrary, we can take each single-site matrix to be independent and uniformly distributed over all single-site symplectic matrices, except for X_0 which is uniformly distributed over the restricted set satisfying $X_0 \mathbf{u}_0^0 = \mathbf{u}_0^0$. Then, we condition on the evolved vector being non-zero at all sites x and different for the initial single-site

non-zero vector, $\mathbf{u}_x^t \neq 0\ \forall x \in \mathbb{Z}_L$ and $\mathbf{u}_0^t \neq \mathbf{u}_0^0$. Therefore under this condition, the evolved vector at each site is independent and uniformly distributed over all non-zero vectors (lemma 2.3) apart from the initial vector $\mathbf{u}_0^0$. For the vector space $\mathcal{V}_0$ we invoke lemma 2.14, and hence the evolved vector $\mathbf{u}_0^t$ ($\neq \mathbf{0}, \mathbf{u}_0^0$) at $x = 0$ is uniformly distributed over all the vectors with the same symplectic form with $\mathbf{u}_0^0$, $\langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle$. Hence, using lemma 2.15, which gives an upper bound for the probability of $\langle \mathbf{u}_0^t, \mathbf{u}_0^0 \rangle \in \{0,1\}$, we get the stated result. □

With this uniformly distributed conditional probability, I will now give the proof of result 2.1. This result states that at any integer or half-integer time in the interval $t \in [1/2, 2t_{\text{scr}}]$ the probability distribution of the evolved operator $\mathbf{u}^t$ is close to uniform, when $N \gg \log L$, when the initial vector $\mathbf{u}^0$ is non-zero on a single site only.

Lemma 2.17. For an initial vector which is non-zero at $x = 0$ only, the evolved vector $\mathbf{u}^t = S_{\text{chain}}^t \mathbf{u}^0$, at integer times, is approximately uniformly distributed over all non-zero vectors within the light-cone. With $t \leq t_{scr}$ within the causal light-cone $x \in [-2t+1, 2t]$ we have:

$$\sum_{\mathbf{v} \in \mathbb{Z}_2^{8Nt}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v}\} - \frac{1}{2^{8Nt} - 1} \right| \leq\ 32t(4t+1)2^{-N} + 4t2^{-2N} \tag{2.134}$$

With $t \geq t_{scr}$ it holds:

$$\sum_{\mathbf{v} \in \mathbb{Z}_2^{2NL}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v}\} - \frac{1}{2^{2NL} - 1} \right| \leq 32t(L+1)2^{-N} + L2^{-2N} \tag{2.135}$$

Additionally, we note that this result holds for any choice of single site at which the initial vector is non-zero.

Proof. The proof of this result is similar to that of lemma 2.12 . I will first consider the case $t \geq t_{scr}$ first. Defining $q \equiv \text{prob}\{\mathbf{u}_x^t \neq \mathbf{0} \forall x \in \mathbb{Z}_L \wedge \mathbf{u}_0^t \neq \mathbf{u}_0^0\}$, the $\text{prob}\{\mathbf{u}^t = \mathbf{v}\}$

can be rewritten in the following way

$$\begin{aligned}\text{prob}\{\mathbf{u}^t = \mathbf{v}\} = q\,&\text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}_x^t \neq \mathbf{0}\,\forall x \in \mathbb{Z}_L \wedge \mathbf{u}_0^t \neq \mathbf{u}_0^0\} \\ &+ (1-q)(\text{prob}\{\mathbf{u}^t = \mathbf{v} |\, \exists x \in \mathbb{Z}_L \,\text{such that}\, \mathbf{u}_x^t = \mathbf{0} \vee \mathbf{u}_0^t = \mathbf{u}_0^0\})\,,\end{aligned}$$

Adding and subtracting the term $\frac{q}{2^{2NL}-1}$ into the sum over $\mathbf{v}$ and using the triangle inequality we find that

$$\begin{aligned}&\sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v}\} - \frac{1}{2^{2NL}-1} \right| \\ &\leq q \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}_x^t \neq \mathbf{0}\,\forall x \in \mathbb{Z}_L, \mathbf{u}_0^t \neq \mathbf{u}_0^0\} - \frac{1}{2^{2NL}-1} \right| \\ &+ (1-q) \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} |\, \exists x \in \mathbb{Z}_L \,\text{such that}\, \mathbf{u}_x^t = \mathbf{0} \vee \mathbf{u}_0^t = \mathbf{u}_0^0\} - \frac{1}{2^{2NL}-1} \right|\,.\end{aligned}$$

We can bound the first term using $q \leq 1$ and apply lemma 2.16 to find that

$$q \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} | \mathbf{u}_x^t \neq \mathbf{0}\,\forall x \in \mathbb{Z}_L, \mathbf{u}_0^t \neq \mathbf{u}_0^0\} - \frac{1}{2^{2NL}-1} \right| \leq 16t2^{-N} + (L+1)2^{-2N}\,.$$

To evaluate the second term above, we upper bound the sum with its maximum value of 2 and use the result of lemma 2.10 to find that

$$(1-q) \sum_{\mathbf{v}} \left| \text{prob}\{\mathbf{u}^t = \mathbf{v} |\, \exists x \in \mathbb{Z}_L \,\text{such that}\, \mathbf{u}_x^t = \mathbf{0} \vee \mathbf{u}_0^t = \mathbf{u}_0^0\} - \frac{1}{2^{2NL}-1} \right| \leq 32t(L+1)2^{-N}\,.$$

Combining, this gives the stated result for integer times after the scrambling time.

To derive the results for integer times before the scrambling time, note that the derivation is identical with the substitution $L \to 4t$, which agree when $t = t_{\text{scr}}$ (and after this time). □

This concludes the proof of result 2.1, which finally leaves the proof which justifies result 2.2. In this result, I will consider a subsystem of the spin chain comprising L_{s} consecutive sites, where L_{s} is even. Without loss of generality, this subsystem is chosen to be $\{1,2,\ldots,L_{\text{s}}\} \subseteq \mathbb{Z}_L$. I will analyse the state of this

subsystem at times

$$t \leq \frac{L - L_{\rm s}}{4} \, , \tag{2.136}$$

which ensures that the left boundary of the light-cone of $\mathbf{u}_1^t$ and the right boundary of the light-cone of $\mathbf{u}_{L_{\rm s}}^t$ do not collide and intersect. In the next lemma, I will prove a generalisation of the result in lemma 2.15. Specifically, I will prove that joint probability of the symplectic form between the initial vector $\mathbf{u}^0$ and the evolved vector $\mathbf{u}^t$ evaluated between each and every site of the L_s region is approximately uniformly distributed (once again, in the regime of large N).

Lemma 2.18. Consider an initial vector $\mathbf{u}^0 \in \mathcal{V}_{\rm chain}$ supported on all lattice sites ($\mathbf{u}_x^0 \neq \mathbf{0}$ for all $x \in \mathbb{Z}_L$), and its evolution at time t, $\mathbf{u}^t$. Define the random variable $s_x = \langle \mathbf{u}_x^t, \mathbf{u}_x^0 \rangle$ at each site of the region $x \in \{1, \ldots, L_{\rm s}\} \subseteq \mathbb{Z}_L$, where $L_{\rm s}$ is even. Then we have

$$P(s_1, \ldots, s_{L_{\rm s}}) \leq 2^{-L_s} + 32 t \, 3^{\frac{L_s}{2}+1} \, 2^{-N} \, , \tag{2.137}$$

as long as $t \leq (L - L_{\rm s})/4$.

Proof. The value of the random vectors $\mathbf{u}_1^t, \ldots, \mathbf{u}_{L_{\rm s}}^t$ is only determined by the random matrices $S_{2-2t}, \ldots, S_{L_{\rm s}+2t-2}$. The rest of matrices S_x are not contained in the causal past of the region under consideration $\{1, 2, \ldots, L_{\rm s}\}$. In order to simplify this proof, we will replace $S_{2-2t}, \ldots, S_{L_{\rm s}+2t-2}$ by a new set of random variables defined in what follows.

Let us label by $y \in \{1, \ldots, L_{\rm s}/2\}$ the pair of neighbouring sites $\{2y-1, 2y\} \subseteq \{1, \ldots, L_{\rm s}\}$. For each pair y we consider a given non-zero vector $\mathbf{a}_y \in \mathbb{Z}_2^{4N}$ and define the random variables

$$\mathbf{b}_y = S_{2y-1}^{-1} \mathbf{a}_y \, , \tag{2.138}$$

$$h_y = \left\langle \mathbf{a}_y, \mathbf{u}_{2y-1}^t \oplus \mathbf{u}_{2y}^t \right\rangle = \left\langle \mathbf{b}_y, \mathbf{u}_{2y-1}^{t-1/2} \oplus \mathbf{u}_{2y}^{t-1/2} \right\rangle \, . \tag{2.139}$$

The left-most random contribution to h_y is the matrix S_{2y-2t}, or equivalently the

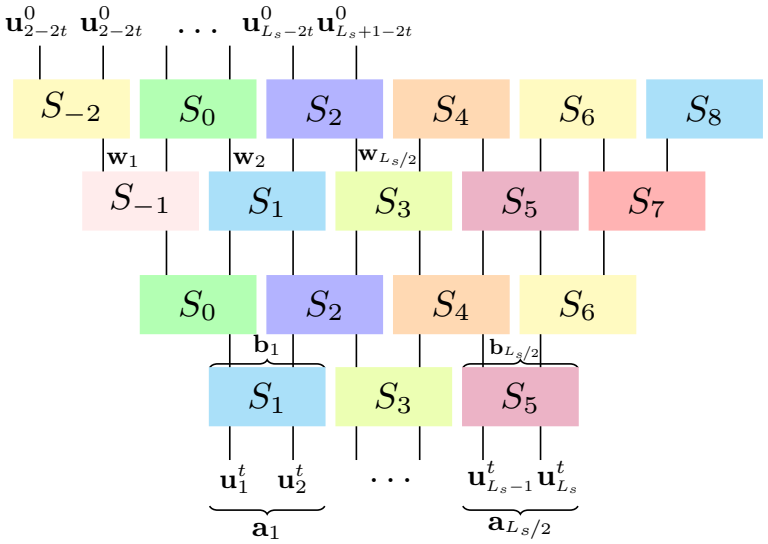

Figure 2.6: This figure represents the region $\{1,2,\ldots,6\}$ at time $t=2$, and its causal past back to $t=0$. (Hence $L_s=6$.) All the random matrices $S_{-2},\ldots,S_8$ contribute to the value of the vectors $\mathbf{u}^t_1,\ldots,\mathbf{u}^t_6$. The left-most contribution to the vector $\mathbf{u}^t_1$ is the matrix S_{-2}, or equivalently the vector $\mathbf{w}_1$. The given vector $\mathbf{a}_y$ associated to the pair of neighbouring sites y, and its 1/2-step backwards time translations $\mathbf{b}_y$, are also represented. This figure was produced by a collaborator, Daniele Toniolo.

vector $\mathbf{w}_y$, defined through

$$\tilde{\mathbf{w}}_y \oplus \mathbf{w}_y = S_{2y-2t}(\mathbf{u}^0_{2y-2t} \oplus \mathbf{u}^0_{2y-2t+1}) \ . \tag{2.140}$$

We note that $\mathbf{w}_y \in \mathcal{V}_{2y-2t+1}$. This contribution and others are illustrated in figure 2.6. The contribution of the vector $\mathbf{w}_y$ to h_y (and $\mathbf{u}^{t-1/2}_{2y-1}$) is "transmitted through" the matrices S_{2y-2}, S_{2y-3}, …, $S_{2y-2t+2}$, $S_{2y-2t+1}$. More precisely, $\mathbf{w}_y$ is mapped via the matrix product

$$F_y = C_{2y-2}C_{2y-3}\cdots C_{2y-2t+2}C_{2y-2t+1} \ , \tag{2.141}$$

where we have used decomposition in equation 2.21. We denote by $\mathbf{v}_y$ all contributions to $\mathbf{u}^{t-1/2}_{2y-1}$ that are not $F_y\mathbf{w}_y$,

$$\mathbf{v}_y = (\mathbf{u}^{t-1/2}_{2y-1} + F_y\mathbf{w}_y) \oplus \mathbf{u}^{t-1/2}_{2y}. \tag{2.142}$$

We remark that $\mathbf{v}_y \in \mathcal{V}_{2y-1} \oplus \mathcal{V}_{2y}$. The last random variable that we need to define is $g_y = \langle \mathbf{b}_y, \mathbf{v}_y \rangle$, which together with equation 2.139 allows us to write

$$h_y = \left\langle \mathbf{b}_y, F_y \mathbf{w}_y + \mathbf{v}_y \right\rangle = \left\langle \mathbf{b}_y, F_y \mathbf{w}_y \right\rangle + g_y \ . \tag{2.143}$$

Note the slight abuse of notation in that we write $F_y \mathbf{w}_y$ instead of $F_y \mathbf{w}_y \oplus \mathbf{0}$.

In summary, we have replaced the variables $S_{2-2t}, \dots, S_{L_s+2t-2}$ by the variables $\mathbf{w}_y, \mathbf{b}_y, F_y, g_y$ for $y = 1, \dots, L_s/2$. (We are not using $\mathbf{v}_y, \tilde{\mathbf{w}}_y$ any more.) These variables are not all independent, but they satisfy the following independence relations:

- $\mathbf{w}_1, \mathbf{b}_1, \dots, \mathbf{w}_{L_s/2}, \mathbf{b}_{L_s/2}$ are independent.
- $\mathbf{w}_y$ is independent of $g_{y'}$ for all $y' \geq y$.
- F_y is independent of $\mathbf{w}_{y'}$ and $\mathbf{b}_{y''}$ for all $y' \leq y$ and $y'' \geq y$.

To continue with the proof it is convenient to introduce the following notation:

$$\mathbf{u}_{\geq y} = (\mathbf{u}_y, \mathbf{u}_{y+1}, \dots, \mathbf{u}_{L_s/2}) \ , \tag{2.144}$$

$$\mathbf{u}_{\leq y} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_y) \ , \tag{2.145}$$

and analogously for $>, <$ and the rest of variables $\mathbf{b}_y, F_y, g_y$. This allows us to write the joint probability distribution of $h_1, \dots, h_{L_s/2}$ as

$$P(h_{\geq 1}) = \sum P(\mathbf{w}_{\geq 1}, \mathbf{b}_{\geq 1}, F_{\geq 1}, g_{\geq 1}) \prod_y \delta\left(h_y, \left\langle \mathbf{b}_y, F_y \mathbf{w}_y \right\rangle + g_y\right) \ . \tag{2.146}$$

Note that we can write the above distribution $P(\mathbf{w}_{\geq 1}, \mathbf{b}_{\geq 1}, F_{\geq 1}, g_{\geq 1})$ as

$$\begin{aligned} P(\mathbf{w}_{\geq 1}, \mathbf{b}_{\geq 1}, F_{\geq 1}, g_{\geq 1}) = & \sum_{S_0, S_1, \dots, S_{L-1}} P(S_0) P(S_1) \cdots P(S_{L-1}) \times \\ & \times \prod_{y=1}^{L_s/2} \delta\left(\mathbf{w}_y - S_{2y-2t}[\mathbf{u}^0_{2y-2t} \oplus \mathbf{u}^0_{2y-2t+1}]\right) \times \delta\left(\mathbf{b}_y - S^{-1}_{2y-1} \mathbf{a}_y\right) \times \\ & \times \delta\left(F_y - C_{2y-2} \cdots C_{2y-2t+1}\right) \times \delta\left(g_y - \left\langle \mathbf{b}_y, (\mathbf{u}^{t-1/2}_{2y-1} + F_y \mathbf{w}_y) \oplus \mathbf{u}^{t-1/2}_{2y} \right\rangle\right) . \end{aligned} \tag{2.147}$$

The following sum-rule is repeatedly exploited below

$$\sum_{\mathbf{w}_1} P(\mathbf{w}_1)\delta\left(h_1, \langle \mathbf{b}_1, F_1\mathbf{w}_1\rangle + g_1\right) = \begin{cases} \delta\left(h_1, g_1\right) & \text{if } (F_1 \oplus \mathbf{0}_{2N})^T J\mathbf{b}_1 = \mathbf{0}\,, \\ 1/2 & \text{otherwise.} \end{cases} \tag{2.148}$$

Using $\delta(h,h') \leq 1$ for all h,h' we can write

$$\begin{aligned} P(h_{\geq 1}) &= \sum P(\mathbf{w}_1) P(\mathbf{w}_{\geq 2}, \mathbf{b}_{\geq 1}, F_{\geq 1}, g_{\geq 1}) \prod_y \delta\left(h_y, \left\langle \mathbf{b}_y, F_y\mathbf{w}_y\right\rangle + g_y\right) \\ &\leq \text{prob}\{(F_1 \oplus \mathbf{0}_{2N})^T J\mathbf{b}_1 = \mathbf{0}\} + \frac{1}{2}\sum P(\mathbf{w}_{\geq 2}, \mathbf{b}_{\geq 2}, F_{\geq 2}, g_{\geq 2}) \prod_{y\geq 2} \delta\left(h_y, \left\langle \mathbf{b}_y, F_y\mathbf{w}_y\right\rangle + g_y\right) \end{aligned}$$

where in the last term we extended the sum over $\mathbf{b}_1, F_1$ from the values satisfying $(F_1 \oplus \mathbf{0}_{2N})^T J\mathbf{b}_1 \neq \mathbf{0}$ to all values. Since the variables $\mathbf{b}_1, F_1, g_1$ do not appear in any of the remaining δ-functions, we can trace them out. Subsequently we repeat the above process by summing over $\mathbf{w}_2$, using the analog of equation 2.148 for $y = 2$, and summing over $\mathbf{w}_2, F_2, g_2$, obtaining

$$P(h_{\geq 1}) = \varepsilon + \frac{1}{2}\left(\varepsilon + \frac{1}{2}\sum P(\mathbf{w}_{\geq 3}, \mathbf{b}_{\geq 3}, F_{\geq 3}, g_{\geq 3}) \prod_{y\geq 3} \delta\left(h_y, \left\langle \mathbf{b}_y, F_y\mathbf{w}_y\right\rangle + g_y\right)\right)\,, \tag{2.149}$$

where we define $\varepsilon = \text{prob}\{(F_1 \oplus \mathbf{0}_{2N})^T J\mathbf{b}_1 = \mathbf{0}\}$. Continuing in this fashion yields

$$\begin{aligned} P(h_1, \ldots, h_{L_s/2}) &= \varepsilon \sum_{k=0}^{L_s/2-1} 2^{-k} + 2^{-L_s/2}\,, \\ &\leq 2\varepsilon + 2^{-L_s/2}\,. \end{aligned} \tag{2.150}$$

We now wish to turn this bound from a distribution of h_y to the distribution of $s_x \equiv \langle \mathbf{u}_x^t, \mathbf{u}_x^0\rangle$ (recalling that $x \in \{1,2,\ldots L_s\}$ and $y \in \{1,\ldots,L_s/2\}$), that is to say we want $P(s_1, s_2, \ldots, s_{L_s})$. I would like to point out here that the rest of the proof of this lemma I attribute to a collaborator, Daniele Toniolo. The bound in equation 2.150 extends to the case where rather than the values of $h_y \equiv s_{2y-1} + s_{2y}$ are fixed, the values of certain h_y and of certain s_x are fixed. For concreteness, let us first consider

the case $L_s = 4$, in which equation 2.150 entails 36 (not necessarily independent) bounds corresponding to the categories:

h_1 fixed, h_2 fixed

h_1 fixed, s_3 or s_4 fixed

h_2 fixed, s_1 or s_2 fixed

s_1 or s_2 fixed, s_3 or s_4 fixed .

For example there are four inequalities arising from the category h_1 fixed, h_2 fixed. Setting $h_1 = 0, h_2 = 0$ we obtain the following inequality in terms of the distribution $P(s_1, s_2, s_3, s_4)$

$$\frac{1}{4} - 6\varepsilon \leq P(0,0,0,0) + P(1,1,0,0) + P(0,0,1,1) + P(1,1,1,1) \leq \frac{1}{4} + 2\varepsilon \ , \quad (2.151)$$

similarly setting $h_1 = 1, h_2 = 0$ we obtain

$$\frac{1}{4} - 6\varepsilon \leq P(1,0,0,0) + P(0,1,0,0) + P(1,0,1,1) + P(0,1,1,1) \leq \frac{1}{4} + 2\varepsilon \ , \quad (2.152)$$

with $h_1 = 0, h_2 = 1$ we obtain

$$\frac{1}{4} - 6\varepsilon \leq P(0,0,1,0) + P(0,0,0,1) + P(1,1,1,0) + P(1,1,0,1) \leq \frac{1}{4} + 2\varepsilon \ , \quad (2.153)$$

with $h_1 = 1, h_2 = 1$

$$\frac{1}{4} - 6\varepsilon \leq P(1,0,1,0) + P(1,0,0,1) + P(0,1,1,0) + P(0,1,0,1) \leq \frac{1}{4} + 2\varepsilon \ , \quad (2.154)$$

where the lower bounds follow from normalisation. To make proceeding further easier, we make the upper and lower bounds similar by replacing 2ε with 6ε in the upper bound. The overall idea is that through specific linear combinations of bounds such as those in equations 2.151-2.154 one can obtain lower and upper bounds on the distribution $P(s_1, s_2, s_3, s_4)$. The tightest set of bounds for $P(s_1, s_2, s_3, s_4)$ in the case $L_s = 4$ (and all the higher orders $L_s \geq 2$) can be obtained by tensor products of

the matrix

$$A \equiv \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \leftrightarrow \begin{pmatrix} h_1 = 0 \\ h_1 = 1 \\ s_1 = 0 \\ s_1 = 1 \\ s_2 = 0 \\ s_2 = 1 \end{pmatrix} . \tag{2.155}$$

The reason for this is that each row of A comes with a label given by the inequalities for the probabilities that it generates. So, the tensor product $A \otimes A$ has rows associated with a couple of labels. For example the first row of the tensor product $A \otimes A$, that has label $(h_1 = 0, h_2 = 0)$ will give rise to the inequality in equation 2.151. Since A is a full rank matrix and the rank of a tensor product is the product of the ranks, then $A \otimes A$ is full rank.

We will now show that by taking linear combinations of specific bounds, such as those in equation 2.151, which is equivalent to linear combinations of rows of $A \otimes A$, we can obtain bounds on the distribution $P(s_1, s_2, s_3, s_4)$. Since the bounds for the probabilities $P(s_1, s_2, s_3, s_4)$ correspond to rows of the tensor product $A \otimes A$ this implies that:

$$\frac{1}{4}(\mathbf{a}_3 + \mathbf{a}_5 - \mathbf{a}_2) \otimes (\mathbf{a}_3 + \mathbf{a}_5 - \mathbf{a}_2) \cdot \mathbf{P} = P(0,0,0,0) , \tag{2.156}$$

where $\mathbf{a_j}$ denotes row j of matrix A, and the tensor product of two rows is meant to be the corresponding row of the tensor product $A \otimes A$, and where $\mathbf{P}$ denotes the vector of all possible choice of $P(s_1, s_2, s_3, s_4)$. The equation 2.156 involves nine bounds, since there are nine terms in the tensor product $(\mathbf{a}_3 + \mathbf{a}_5 - \mathbf{a}_2) \otimes (\mathbf{a}_3 + \mathbf{a}_5 - \mathbf{a}_2)$, and so

$$\frac{1}{16} - 56\varepsilon \leq P(0,0,0,0) \leq \frac{1}{16} + 56\varepsilon . \tag{2.157}$$

Note that the error 56ε arises as the product of the error associated with each bound in equation 2.151 and the number of inequalities that is 9. Any choice of $P(s_1, s_2, s_3, s_4)$ can be found through a linear combination of three inequalities for each pair of

sites, and hence the bound applies in general to $P(s_1,s_2,s_3,s_4)$. To generalise this to arbitrary L_s, we just consider further tensor products of A, and hence

$$2^{-L_s} - 2\,3^{\frac{L_s}{2}+1}\varepsilon \leq P(s_1,...,s_{L_s}) \leq 2^{-L_s} + 2\,3^{\frac{L_s}{2}+1}\varepsilon\ . \tag{2.158}$$

□

Using this proof, in the following lemma I will show the proof which justifies result 2.2.

Lemma 2.19. Consider an initial vector $\mathbf{u}^0 \in \mathcal{V}_\mathrm{chain}$ with non-zero support in all lattice sites ($\mathbf{u}^0_x \neq \mathbf{0}$ for all $x \in \mathbb{Z}_L$). Consider the evolved vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ inside a region $x \in \{1,\ldots,L_\mathrm{s}\} \subseteq \mathbb{Z}_L$ where L_s is even and the time is $t \leq \frac{L-L_\mathrm{s}}{4}$. If $\mathbf{u}^t_{[1,Ls]}$ is the projection of $\mathbf{u}^t$ in the subspace $\bigoplus_{x=1}^{L_\mathrm{s}} \mathcal{V}_x$ then

$$\sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_\mathrm{s}}} \left|\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t_{[1,Ls]}\} - \frac{1}{2^{2NL_\mathrm{s}}}\right| \leq 32t\,2^{-N}(2L_\mathrm{s}+3^{\frac{L_\mathrm{s}}{2}+1}) + 4L2^{-2N}\ . \tag{2.159}$$

Proof. First, we re-state $\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t\}$ in the following way

$$\begin{aligned}\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t\} &= q\,\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t|\mathbf{u}^t_x\neq\mathbf{0},\mathbf{u}^0_x\forall x\in\mathbb{Z}_{L_\mathrm{s}}\}\\ &\quad+(1-q)(1-\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t|\mathbf{u}^t_x\neq\mathbf{0},\mathbf{u}^0_x\forall x\in\mathbb{Z}_{L_\mathrm{s}}\})\ ,\end{aligned}$$

where $x \in \{1,\ldots,L_\mathrm{s}\} \subseteq \mathbb{Z}_L$ with L_s is even, q is the probability of distribution $\mathrm{prob}\{\mathbf{u}^t_x \neq \mathbf{0}, \mathbf{u}^t_x \forall x \in \mathbb{Z}_{L_\mathrm{s}}\}$, and similarly with the complement. Then using convexity we find that

$$\begin{aligned}\sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_\mathrm{s}}} \left|\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t\} - \frac{1}{2^{2NL_\mathrm{s}}}\right| &\leq q \sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_\mathrm{s}}} \left|\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t|\mathbf{u}^t_x\neq\mathbf{0},\mathbf{u}^0_x\forall x\} - \frac{1}{2^{2NL_\mathrm{s}}}\right|\\ &\quad+(1-q)\sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_\mathrm{s}}} \left|1-\mathrm{prob}\{\mathbf{v}=\mathbf{u}^t|\mathbf{u}^t_x\neq\mathbf{0},\mathbf{u}^0_x\forall x\} - \frac{1}{2^{2NL_\mathrm{s}}}\right|\ .\end{aligned}$$

We can evaluate the first term using the upper bound $q \leq 1$ and use lemma 2.14

combined with lemma 2.18 to find that

$$q \sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_s}} \left| \text{prob}\{\mathbf{v}=\mathbf{u}^t | \mathbf{u}_x^t \neq \mathbf{0}, \mathbf{u}_x^0 \forall x \in \mathbb{Z}_{L_s}\} - \frac{1}{2^{2NL_s}} \right| \leq 32t3^{\frac{L_s}{2}+1}2^{-N} + L2^{2-2N} . \tag{2.160}$$

To evaluate the second term, we can upper bound the sum by its maximum value, 2, and use the result of lemma 2.10 to upper bound $(1-q)$ to find that

$$(1-q) \sum_{\mathbf{v}\in\mathbb{Z}_2^{2NL_s}} \left| 1 - \text{prob}\{\mathbf{v}=\mathbf{u}^t | \mathbf{u}_x^t \neq \mathbf{0}, \mathbf{u}_x^0 \forall x \in \mathbb{Z}_{L_s}\} - \frac{1}{2^{2NL_s}} \right| \leq 64L_s t 2^{-N} . \tag{2.161}$$

Combining these two terms we get the stated result. □

This concludes the proof which justifies result 2.2, and moreover the proof of all results concerning the time-periodic random quantum circuit model I have presented.

2.1.3 Discussion

To summarise, I have shown and proven a selection of results concerning the mixing of Pauli operators in the time-periodic random quantum circuit model in the regime of large N ($N \gg \log L$). Moreover, the indistinguishability result (result 2.4) is a weaker, variant notion of approximate 2-design, which requires measurements to be restricted to Pauli operators only. The results in this section contrast with the case of time-dependent random quantum circuits (discussed in chapter 1 section 1.2), in which, for example, the resemblance to a (Haar) random unitary increases with the time evolution. The question of whether local and time-independent or time-periodic dynamics can generate a 2 (or higher) approximate unitary design remains open. However, results in references [10, 21, 22, 26, 27] suggest that further results in this direction are possible.

All of the results I have proven in this are relevant only in the regime of large N ($N \gg \log L$). Naturally, one might wonder about the opposite regime of small N, when there is only one or a few qubits per site of the spin chain. I will discuss and prove results in this regime in chapter 3 section 3.1, in which a strong form of localisation is observed. In the final chapter of this thesis (chapter 4), I will discuss

more generally Clifford dynamics and why they are an interesting model with which to study the landscape of quantum many-body phenomena.

Also, I would like to point out that there is a different interpretation of the model in this section, which is a spin chain of L sites, with N qubits per site, and nearest-neighbour interactions. Instead, equivalently the model could be presented as a spin chain with LN sites, with a single qubit per site, and interactions of range $2N$. This modification uses the fact that any Clifford gate acting on $2N$ qubits can be rewritten as a circuit of depth $\mathcal{O}(N^2/\log N)$ [122, 123], and so the dynamical period of the LN-site spin chain decomposes into a circuit with $\mathcal{O}(N^2/\log N)$ elementary time steps.

The scrambling time t_{scr} has been of great importance to the results I have presented in this section. This is because the (strict) causality imposed by local dynamics means that time evolution of at least the scrambling time is required for an initially local operator to evolve to a completely non-local operator. Additionally, all of the results of presented in this section hold only up until a maximum time of $2t_{\text{scr}}$, which is a consequence of the proof technique. Moreover, in all of the result I have discussed the proximity to the uniform distribution and resemblance to a random unitary decrease with time. However, this is also the case for a (Haar) random unitary. For example, the powers U^t of a (Haar) random unitary $U \in \text{SU}(d)$ lose their resemblance to a (Haar) random unitary as t increases, as quantified by the spectral form factor. In the model I have considered, the time evolution operator is composed of, and is itself, a Clifford unitary. So, since the Clifford group is finite, this means that there is a recurrence time, t_{rec}, such that the time evolution operator is trivial, $W(t_{\text{rec}}) = \mathbb{1}$. Therefore, it is reasonable to expect that the time at which the time evolution operator $W(t)$ resembles as closely as possible a random unitary, is the smallest time such that recurrences are avoided, but also larger than the scrambling time.

Additionally, I would like to briefly comment on the role of the boundary conditions of the model. For instance, the periodic boundary conditions mean that the minimum time for the light-cone of any initially local operator to contain the

whole spin chain does not depend on the site at which the initially local operator is non-identity, and so the scrambling time is $t_{\text{scr}} = L/4$. However, if instead one considered a modified version of the model with closed boundary conditions this would not be the case. The proofs of all the results in this section make use of the fact that the model has periodic boundary conditions, specifically from lemma 2.8 onwards. Nonetheless, I expect that the proof method used in this section can be adapted and applied to versions of the model with different boundary conditions, and believe that similar mixing results can be obtained. Although, the results with modified boundary conditions may not hold over the same timescales, and may depend more strongly on the initial operator, and which particular sites of the spin chain the operator is initially non-identity.

2.2 Non-local few-body dynamics

The scrambling time is an important timescale when studying mixing (of Pauli operators) and scrambling. As discussed in the introduction (chapter 1 section 1.3), this is particularly the case for studying the dynamics of highly chaotic quantum systems, such as black holes which are often modelled with (Haar) random unitaries. This model is often justified by the fact that local random quantum circuits generate unitary designs. However, as I have noted in the introduction, these circuits are time dependent, while physically one would expect the dynamics to be time-independent [66]. In this section, I will present a time-periodic model, which shares important features with time independent dynamics, and in which the interactions are non-local but few-body. I will investigate the circuit depth corresponding to the scrambling time. I will show that this model approaches the circuit depth corresponding to the scrambling time postulated in the fast scrambling conjecture, logarithmic in the number of qubits, but does not reach it. Unlike in the previous section (2.1), I will not present results about how closely the dynamics resemble a (Haar) random unitary, or any other results concerning Pauli mixing. I leave this important question for future work.

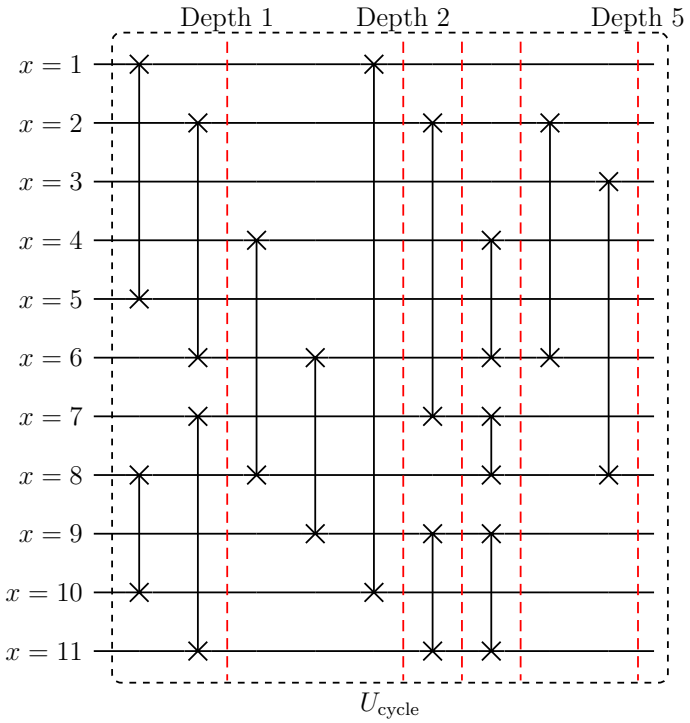

Figure 2.7: This is an example of the evolution operator U_{cycle}, as enclosed by the dashed-box, acting on eleven sites ($L = 11$), with N qubits per site. The circuit itself is of depth 5, with increases denoted by vertical dashed lines. Each interaction of two sites, signified by two crosses connected by a (vertical) line, is a Clifford unitary sampled from the Clifford group of $2N$ qubits according to the uniform distribution. It is worth noting that while the depicted quantum circuit is one-dimensional, the interactions are not local.

2.2.1 Overview

I will first describe the physical model analysed in this section. The model considered is a spin system composed of L sites. Each site is labeled by $x \in \mathbb{Z}_L$ and situated on each site there are N qubits. Hence, the (local) Hilbert space of each site has dimension 2^N and the total Hilbert space for the entire system has dimension 2^{NL}. The dynamics of the spin system is discrete in time, and so is characterised by a unitary U_{cycle}. Since, we consider time-periodic dynamics, each evolution time cycle is characterised by the same unitary U_{cycle}, and hence $|\psi(t)\rangle = \left(U_{\text{cycle}}\right)^t |\psi(0)\rangle$ for all $t \in \mathbb{Z}$.

The unitary U_{cycle}, which is implemented via a quantum circuit, is generated

through the following random process:

1. A random pair of (different) sites (x, y) are selected uniformly from the spin system, $x, y \in \mathbb{Z}_L$.

2. Then, for this selected pair of sites, a random Clifford unitary is applied to the $2N$ qubits.

3. Repeat from step 1 again independently M times.

I will denote this sequence, which generates the quantum circuit, by S_{cycle}. The variable M is a free parameter which is chosen such that, with high probability, the random quantum circuit has gates applied to every site of the L-site spin chain and is a connected graph. By connected graph, I mean that every site of the spin system is (via the interactions) causally connected to all other sites of the spin chain.

Hence, the sequence S_{cycle} produced by this random process, and so the quantum circuit which implements the random unitary U_{cycle}, has two sources of randomness: gates and location. Figure 2.7 illustrates one particular instance of the random unitary U_{cycle}. Therefore, the random unitary U_{cycle} can be written as

$$U_{\text{cycle}} = \prod_{i=1}^{M} U_{(x_i, y_i)} \ , \tag{2.162}$$

where (x_i, y_i) is the random pair of (different) sites selected in trial i (of total M in the sequence S_{cycle}), and $U_{(x_i, y_i)}$ denotes the random Clifford unitary applied to the $2N$ qubits of this pair of different sites, where it is understood that this Clifford unitary is in tensor product with the identity operator on all other sites of the spin system.

Since in this model the dynamics are not generated by nearest-neighbour interactions, there is no sense of locality. However, the interactions, while not local, are still between two-sites only. In this section, I will investigate the scrambling time, and the corresponding circuit depth, of the model I have described.

Definition 2.20. In this model, I will call the scrambling time the (minimum) number of repetitions of the quantum circuit of U_{cycle}, as generated by the sequence S_{cycle} of gates, such that any initially local operator (non-identity on a single site at $t = 0$),

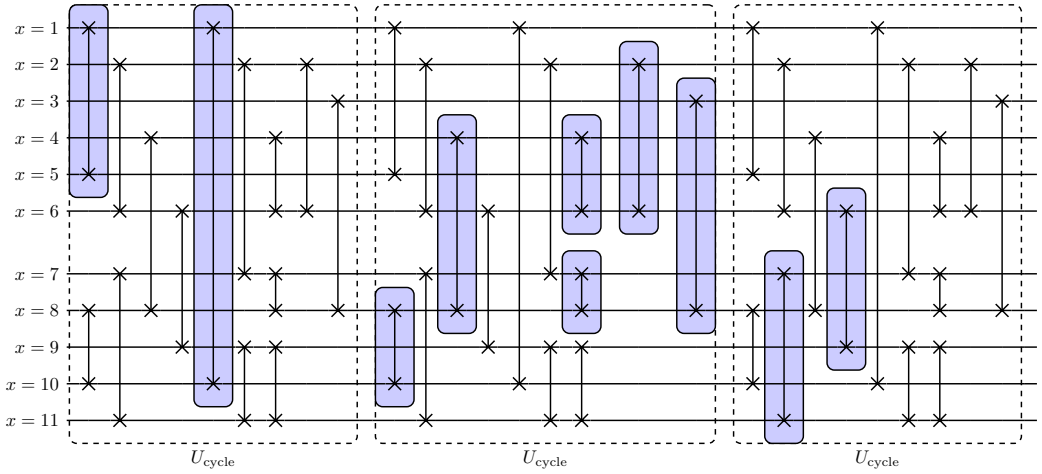

Figure 2.8: In this figure, there is an illustration of an example quantum circuit to clarify the definition of scrambling time. The quantum circuits in this example is the same as illustrated in figure 2.7, and is repeated three times, so U_{cycle}^3. In this example, I consider the initial operator to be non-identity on site 1 only. Each blue box, surrounding an interacting pair of sites, highlights when the operator evolving via the quantum circuit spreads and its support grows to a new site. After three repetitions of the driving unitary, the initial operator has spread through the entire system. The scrambling time is the number of repetitions such that this occurs for any initial operator which is non-identity on any single site.

through its time evolution, to grow and spread throughout the entire system. To clarify this, there is an illustration of an example in figure 2.8.

In this model, the scrambling time itself is also a random variable, which can vary between particular instances of the random quantum circuit. Hence, statements made about the scrambling time in this model are probabilistic.

The quantum circuit of the time evolution operator $(U_{\text{cycle}})^t$ generated by the sequence S_{cycle} has two contributions towards the circuit depth corresponding to the scrambling time; one contribution comes from the depth of the quantum circuit of the random unitary (U_{cycle}), which is made up of M independent unitary gates, and the other from the number of repetitions of this driving unitary, in other words the scrambling time t_{scr}. Generally speaking, two consecutive Clifford unitaries that make up the time evolution operator U_{cycle} will not act upon any of the same sites, and hence the two unitaries commute.

In the following result, I will bound the probability that the circuit depth

corresponding to the scrambling time is at most $\sim (\log L)^2$, which is given by $t_{\text{scr}} \sim \log L$ time cycles of the quantum circuit, which itself is composed of $\sim L \log L$ independently selected random gates, that has depth $\sim \log L$.

Result 2.5. The quantum circuit of the random unitary U_{cycle} generated by the sequence S_{cycle}, which is composed of $M = 6L\log L$ independent gates, where $L \geq 70$, has with probability close to one a circuit depth of at most $\sim \log L$ and scrambling time of at most $\sim \log L$. More precisely,

$$\text{prob}\left\{\text{depth}(S_{\text{cycle}}) \leq 24 \log L\right\} > 1 - \frac{L^{-5}}{4} , \tag{2.163}$$

and

$$\text{prob}\left\{t_{\text{scr}}(S_{\text{cycle}}) \leq \frac{2\log(L-1)}{\log\log L}\right\} > 1 - L^{-6\log L} - 12L^{-1} . \tag{2.164}$$

where $\text{depth}(S_{\text{cycle}})$ denotes the depth of the quantum circuit generated by the sequence S_{cycle}, and $t_{\text{scr}}(S_{\text{cycle}})$ the scrambling time.

2.2.2 Details and derivation

I will now give the detailed proofs which justify result 2.5. Firstly, I will define some notation (definitions 2.21 - 2.25), to make the mathematical proofs more concise and understandable. Then, I will (in lemmas 2.26 - 2.28) bound the probability for the minimum number of gates, M, such that the unitary U_{cycle} acts upon every site at least once. Using these results as a guide, (in lemmas 2.30 - 2.32) I will bound the probability for the number of gates, M, such that the random unitary U_{cycle} is a connected graph. Then, I will prove (lemmas 2.33 - 2.35) that the random unitary U_{cycle} with the number of gates M is a quantum circuit of a certain depth, with high probability.

To prove the number of required time cycles t_{scr} of the random unitary U_{cycle} and hence justify result 2.5, I will use the perspective and tools of random graph theory. First, I will precisely state I mean by the scrambling time in the context of the model considered (definitions 2.36 - 2.39). Then, I will introduce and define

two related models of random graphs, which are of particular relevance (definitions 2.41 and 2.42), and reproduce two well-known proofs of their relationship (lemmas 2.43 and 2.45). In order to apply well-known results of random graph theory, I will (in lemma 2.46) bound the probability that the number of gates M in the quantum circuit has a certain number of repeated (independent) choices of pairs of sites. Then, I will (in lemma 2.50) reproduce and expand upon a well known result concerning the diameter of sparse random graphs. Finally, in lemma 2.51 I will combine these proofs and hence justify result 2.5.

Circuit connectivity

First, I will define some notation, for the purposes of making the proofs slightly more concise. Then, I will present results which concern the random process of picking pairs of sites of the spin chain, to form the unitary U_{cycle}, and in particular bound the probability that the random unitary is a connected graph after M gate selections.

Definition 2.21. The (uniformly sampled) random pair of (different) sites selected in trial i (of M) is denoted by $\alpha_i = (x_i, y_i)$, where x, y are different elements of $\mathbb{Z}_L$.

Definition 2.22. A sequence of M independent choices of random pairs of sites is denoted by $S_M = (\alpha_1, \ldots, \alpha_M)$.

Definition 2.23. The set of points of a given sequence $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)\}$ is denoted by $\text{set}(S_M) = \{x_1, y_1, x_2, y_2, \ldots, x_k, y_k\} \subseteq \mathbb{Z}_L$, and its cardinality by $|\text{set}(S_M)|$.

Definition 2.24. The depth of the quantum circuit formed from a sequence of gates on random pairs S_M is denoted by $\text{depth}(S_M)$. (A sequence of disjoint pairs has unit depth.)

Definition 2.25. For any given sequence S_M we denote by $\text{cl}(S_M) \subseteq S_M$ the subsequence of random pairs which "grow" from the initial element α_1 - referred to as a cluster. That is:

1. $\alpha_1 \in \text{cl}(S_M)$

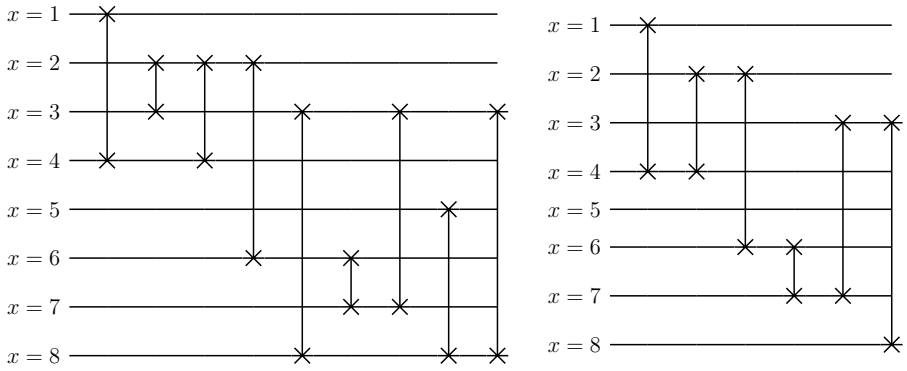

Figure 2.9: This figure is to illustrate the concept of a cluster introduced in definition 2.25. The left hand image contains an example of a sequence S_9, for a system consisting of a total of eight sites. The right hand image shows the cluster $\mathrm{cl}(S_9)$, which is the sequence S_9 with the second, fifth, and eighth elements (so $\alpha_3, \alpha_5, \alpha_8$) removed.

2. $\alpha_i = (x_i, y_i) \in \mathrm{cl}(S_M)$ if and only if either x_i or y_i belongs to $\mathrm{set}(\mathrm{cluster}(S_{i-1}))$ but not both, for $i = 2, \ldots, M$.

Figure 2.9 shows an illustrative example of the definition of a cluster (definition 2.25).

With these definitions, I will now commence with the proofs, which culminate in bounding the probability for the random unitary U_{cycle} being fully connected after M selections of pairs of sites. In the following, I will find the probability for the random choice of a pair of sites to be a particular pair, which is important for all of the proofs in this section.

Lemma 2.26. The probability of picking a particular pair of sites $x, y \in \mathbb{Z}_L$ is

$$\mathrm{prob}\{\alpha_i = (x, y)\} = \binom{L}{2}^{-1}. \tag{2.165}$$

Proof. The total number of distinct choices of pairs of sites of the spin chain is $\binom{L}{2}$. Hence, since we are selecting uniformly from this set of choices, we get the stated probability. □

I will now prove an upper-bound to the probability that a single particular site is never selected in any of the M independent trials of pairs of sites.

Lemma 2.27. The probability that a single site $x \in \mathbb{Z}_L$ is not selected after M random choices of pairs is given by

$$\text{prob}\{x \notin \text{set}(S_M)\} = \left(1 - \frac{2}{L}\right)^M \leq e^{\frac{-2M}{L}} \,. \tag{2.166}$$

Proof. Using the result from lemma 2.26, the probability that a random pair α_i does not include a specific site $x \in \mathbb{Z}_L$ is given by

$$\text{prob}\{x \notin \text{set}(\alpha_i)\} = \frac{\binom{L-1}{2}}{\binom{L}{2}} = 1 - \frac{2}{L} \,. \tag{2.167}$$

Hence, since each choice of pairs of sites of the spin chain α_i in the sequence S_M is independent and uniformly distributed, we get the stated (exact) probability.

The upper bound follows from the inequality

$$(1+k)^c \leq e^{ck} \,, \tag{2.168}$$

for $k, c \in \mathbb{R}$ and $1 + k > 0$, $c > 0$. □

With this result, I will now prove a bound on the probability that every site has been selected at least once. This does not imply that the quantum circuit of the unitary U_{cycle} is a connected graph, however it does serve as a useful guide and perspective for later results.

Lemma 2.28. The probability that, after M random pairs, all sites have been selected at least once is given by

$$\text{prob}\{\text{set}(S_M) = \mathbb{Z}_L\} \geq 1 - L\left(1 - \frac{2}{L}\right)^M \geq 1 - Le^{\frac{-2M}{L}} \,. \tag{2.169}$$

Proof. To prove this lemma, we study the complementary event

$$\text{prob}\{\text{set}(S_M) \subset \mathbb{Z}_L\} = 1 - \text{prob}\{\text{set}(S_M) = \mathbb{Z}_L\} \,. \tag{2.170}$$

Using the union bound, and the result from lemma 2.27, we have that

$$\text{prob}\{\text{set}(S_M) \subset \mathbb{Z}_L\} \leq \sum_{x \in \mathbb{Z}_L} \text{prob}\{x \notin \text{set}(S_M)\} \,, \tag{2.171}$$

$$\leq L\left(1-\frac{2}{L}\right)^M \leq L e^{\frac{-2M}{L}} \,. \tag{2.172}$$

This gives the result stated in the lemma. □

Broadly speaking, this proof implies that for $M \sim L\log L$ every single site has been selected, and therefore a random Clifford unitary has been applied, at least once, with probability close to one. I will now consider the more involved question of the quantum circuit being a connected graph. To do this, I will make use of the notion of a cluster, of which there is a precise definition in definition 2.25. This notion of cluster is useful as if the cluster contains every single site then the quantum circuit is a connected graph. Before proceeding to any proofs, I will first introduce another random variable.

Definition 2.29. Consider a sequence S_M with $|\text{cl}(S_M)| = i$. We define the random variable μ_i as the number of pairs generated until the cluster increases by one: $|\text{cl}(S_M)| = |\text{cl}(S_{M+1})| = \cdots = |\text{cl}(S_{M+\mu_i-1})| = i$ and $|\text{cl}(S_{M+\mu_i})| = i+1$.

In the next lemma, I will prove the form of the probability distribution of this newly defined random variable μ_i. In particular, the fact that the probability distribution is a geometric distribution, is of particular importance.

Lemma 2.30. The probability distribution of the random variable μ_i is given by

$$\text{prob}\{\mu_i = k\} = (1-\gamma_i)^{k-1}\gamma_i = \left(1-\frac{i(L-i)}{\binom{L}{2}}\right)^{k-1} \frac{i(L-i)}{\binom{L}{2}} \,, \tag{2.173}$$

where $i \in [2, L-1]$ and $\gamma_i = \frac{i(L-i)}{\binom{L}{2}}$. This is a geometric distribution with $p = \gamma_i$.

Proof. First, we note that the probability that the size of the cluster increases from i

to $i+1$ after a single choice of a pair of sites, α_j, is given by

$$\text{prob}\{\mu_i = 1\} = \text{prob}\{|\text{cl}(S_{M+1})| = i+1 \,\big|\, |\text{cl}(S_M)| = i\} \,, \tag{2.174}$$

$$= \frac{i(L-i)}{\binom{L}{2}} = \gamma_i \,. \tag{2.175}$$

To extend this to $k > 1$, we note that all choices of pairs of sites apart from the final choice must not increase the size of the cluster. Hence, since all choices of pairs of sites are independent, then we get the stated result. □

I will use this proof in the following two lemmas. First, I will find the expected number of choices M such that the cluster contains every site. Then, I will bound the probability that the cluster contains every site after a (minimum) number of choices $M \sim L\log L$.

Lemma 2.31. The expected number of independent choices $\mathbb{E}(M)$ such that the cluster reaches all of the spin chain, that is to say $\text{set}(\text{cl}(S_M)) = \mathbb{Z}_L$, is

$$\mathbb{E}(M) = 1 + \frac{L-1}{2}\left(H_{L-1} + H_{L-2} - 1\right) \,, \tag{2.176}$$

where H_n is the n-th harmonic number, which can be bounded as

$$0.5L\log(L) \leq \mathbb{E}(M) \leq 2L\log(L) \,. \tag{2.177}$$

Proof. The expected number of independent choices such that $\text{set}(\text{cl}(S_M)) = L$ is given by

$$\mathbb{E}\left(\sum_{i=2}^{L-1} \mu_i\right) + 1 = 1 + \sum_{i=2}^{L-1} \mathbb{E}(\mu_i) \,, \tag{2.178}$$

where the equality follows from the linearity of the expected value, and the $+1$ term follows from the fact that the very first choice of pairs of sites, α_1, initially defines the cluster. Since, the probability distribution $\text{prob}\{\mu_i = k\}$ is a geometric distribution, the expected value is given by

$$\mathbb{E}(\mu_i) = \frac{1}{\gamma_i} = \frac{\binom{L}{2}}{i(L-i)} \,. \tag{2.179}$$

So,

$$\mathbb{E}\left(\sum_{i=2}^{L-1}\mu_i\right)+1 = 1+\sum_{i=2}^{L-1}\frac{\binom{L}{2}}{i(L-i)} = 1+\binom{L}{2}\sum_{i=2}^{L-1}\frac{1}{i(L-i)} \tag{2.180}$$

$$= 1+\binom{L}{2}\sum_{i=2}^{L-1}\left(\frac{1}{Li}+\frac{1}{L(L-i)}\right) \tag{2.181}$$

$$= 1+\frac{L-1}{2}\left(H_{L-1}+H_{L-2}-1\right)\ , \tag{2.182}$$

where H_n is the n-th harmonic number. This gives the exact result, which can be bounded to a region, using the bounds $\log(n)+n^{-1} \leq H_n \leq \log(n)+1$,

$$L\log(L)-\log(L)-\frac{L}{2}+1 \leq \mathbb{E}\left(\sum_{i=2}^{L-1}\mu_i\right)+1 \leq L\log(L)+\frac{L}{2}+1-\log(L)\ . \tag{2.183}$$

□

Lemma 2.32. The probability that the cluster contains every site of the spin chain, that is to say $\mathrm{set}(\mathrm{cl}(S_M)) = L$, when the number of independent choices $M \geq (1+\varepsilon)L\log L$, is bounded by

$$\mathrm{prob}\{\mathrm{set}(\mathrm{cl}(S_{1+M})) = L\} \leq L^{-\varepsilon}\ . \tag{2.184}$$

Proof. In this lemma, I will move from the expected value, as found in lemma 2.31, to a bound on the probability. To do this, we first note that the random variables μ_i are all independent, and hence we can apply the Chernoff bound (see appendix B), which states that

$$\mathrm{prob}\left\{\sum_{i=2}^{L-1}\mu_i \geq a\right\} \leq e^{-ta}\prod_{i=2}^{L-1}\mathbb{E}\left(e^{t\,\mu_i}\right)\ , \tag{2.185}$$

for any $t>0$. The quantity $\mathbb{E}\left(e^{t\,\mu_i}\right)$ is the generating function, which in the case of a geometric distribution is given by

$$\mathbb{E}\left(e^{t\ \mu_i}\right) = \frac{\gamma_i e^t}{1-(1-\gamma_i)e^t}\ . \tag{2.186}$$

We now will evaluate the product

$$\prod_{i=2}^{L-1} \mathbb{E}\left(e^{t\,\mu_i}\right) = \prod_{i=2}^{L-1} \frac{\gamma_i e^t}{1-(1-\gamma_i)e^t} = \prod_{i=2}^{L-1} \frac{\gamma_i}{e^{-t}-(1-\gamma_i)} \,. \tag{2.187}$$

Let $t = L^{-1}$ and hence $\exp\left(-L^{-1}\right) \geq 1 - L^{-1}$, then the product can be upper bound as

$$\prod_{i=2}^{L-1} \frac{\gamma_i e^t}{1-(1-\gamma_i)e^t} \leq \prod_{i=2}^{L-1} \frac{\gamma_i L}{\gamma_i L - 1} = \prod_{i=2}^{L-1} \frac{2i(L-i)}{2i(L-i)-(L-1)} \tag{2.188}$$

$$= 2^{L-2}(L-1)!(L-2)! \prod_{i=2}^{L-1} \frac{1}{2i(L-i)-(L-1)} \tag{2.189}$$

$$= (-1)^{L-2}(L-1)!(L-2)! \prod_{i=2}^{L-1} \frac{1}{(i^2 - Li + 0.5(L-1))} \,. \tag{2.190}$$

The term inside the product can be factorised as

$$\frac{1}{(i^2 - Li + 0.5(L-1))} = \frac{1}{(i - \frac{L}{2} - \frac{1}{2}\sqrt{L^2-2L+2})(i - \frac{L}{2} + \frac{1}{2}\sqrt{L^2-2L+2})} \,, \tag{2.191}$$

which, using the bounds $L - 1 \leq \sqrt{L^2-2L+2} \leq L$, allows us to upper bound the product as

$$\prod_{i=2}^{L-1} \frac{\gamma_i e^t}{1-(1-\gamma_i)e^t} \leq (-1)^{L-2}(L-1)!(L-2)! \prod_{i=2}^{L-1} \frac{1}{i-L} \frac{1}{i-\frac{1}{2}} \tag{2.192}$$

$$\leq (L-1)!(L-2)! \prod_{i=2}^{L-1} \frac{1}{L-i} \frac{1}{i-1} \tag{2.193}$$

$$= (L-1)!(L-2)! \frac{1}{(L-2)!(L-2)!} \leq L \tag{2.194}$$

Therefore,

$$\text{prob}\left\{ \sum_{i=2}^{L-1} \mu_i \geq a \right\} \leq L \exp\left(-\frac{a}{L}\right) \,, \tag{2.195}$$

and so if $a = (1+\varepsilon) L \log L$, then, we get the stated upper bound. For the same reasons as in lemma 2.31, we must include an additional $+1$ term to the number of independent random choices. □

In other words, I have shown that for $M \sim L\log L$ the quantum circuit of the random unitary U_{cycle} is a connected graph with probability close to one. In the rest of this section, I will focus only on this scenario.

Circuit depth

I will now prove a series of results, which culminate in demonstrating that the quantum circuit of the random unitary U_{cycle} made up of $M \sim L\log L$ independently selected and applied Clifford gates has a depth $\sim \log L$. In the following lemma, I will give a proof for the probability that for $M < L$ choices of pairs of sites, the quantum circuit is of maximum depth, so circuit depth of M.

Lemma 2.33. The probability that $M < L$ random choices of pairs of sites forms a quantum circuit of depth M is given exactly by

$$\text{prob}\{\text{depth}(S_M) = M\} = \left(\frac{4}{L} - \binom{L}{2}^{-1}\right)^{M-1} = \left(\frac{4}{L} - \frac{2}{L(L-1)}\right)^{M-1} . \quad (2.196)$$

Proof. We start by noting that the probability $\text{prob}\{\text{depth}(S_M) = M\}$ is equal to

$$\begin{aligned}\text{prob}\{\text{depth}(S_M) = M\} = (&\text{prob}\{\text{depth}(S_{M-1}) = M-1\}\times \\ &\text{prob}\{\text{depth}(S_M) = M|\text{depth}(S_{M-1}) = M-1\}) .\end{aligned} \quad (2.197)$$

This equality also applies in a similar fashion to the term $\text{prob}\{\text{depth}(S_{M-1}) = M-1\}$, and hence we will now analyse the term $\text{prob}\{\text{depth}(S_M) = M|\text{depth}(S_{M-1}) = M-1\}$. We note that when adding another pair of sites, α_M to the sequence S_{M-1}, which is of depth $M-1$, the only way to increase the depth (by one) is if α_M shares a site (or both) with the previous choice α_{M-1}. Therefore, we have that

$$\text{prob}\{\text{depth}(S_M) = M|\text{depth}(S_{M-1}) = M-1\} = \frac{2(L-1)-1}{\binom{L}{2}} , \quad (2.198)$$

where the negative term arises from avoiding double counting for the case $\alpha_M =$

α_{M-1}. So, we have that

$$\text{prob}\{\text{depth}(S_M) = M\} = \frac{2(L-1)-1}{\binom{L}{2}}\text{prob}\{\text{depth}(S_{M-1}) = M-1\} \tag{2.199}$$

$$= \left(\frac{4}{L} - \binom{L}{2}^{-1}\right)\text{prob}\{\text{depth}(S_{M-1}) = M-1\}\ . \tag{2.200}$$

Therefore, by making the same argument for the term $\text{prob}\{\text{depth}(S_{M-1}) = M-1\}$, we find that

$$\text{prob}\{\text{depth}(S_M) = M\} = \left(\frac{4}{L} - \binom{L}{2}^{-1}\right)^{M-1}\ , \tag{2.201}$$

which is the stated result. □

Using this, I will now prove a general result upper-bounding the probability that the depth is at least D after M random (independent) choices of pairs of site. Finally, I will make use of this proof and upper-bound the probability further for the particular case of $M = (1+\varepsilon)L\log L$, which as was demonstrated in lemma 2.32 is the region of M for which the quantum circuit is a connected graph.

Lemma 2.34. The probability that M random choices of pairs of sites forms a quantum circuit with a depth of at least D is bounded by

$$\text{prob}\{\text{depth}(S_M) \geq D\} \leq \binom{M}{D}\left(\frac{4}{L} - \binom{L}{2}^{-1}\right)^{D-1}\ . \tag{2.202}$$

Proof. The proof of this result is as follows

$$\text{prob}\{\text{depth}(S_M) \geq D\} = \text{prob}\{\exists S_D \subseteq S_M : \text{depth}(S_D) = D\}\ , \tag{2.203}$$

where by $S_D \subseteq S_M$ we mean a selection of D (different) elements from the sequence S_M. By using the union bound becomes

$$\text{prob}\{\text{depth}(S_M) \geq D\} \leq \binom{M}{D}\text{prob}\{\text{depth}(S_D) = D\}\ , \tag{2.204}$$

and hence using the result in lemma 2.33 we get the stated result. □

Lemma 2.35. The probability that for $M = (1+\varepsilon)L\log L$ the depth of the circuit exceeds $D = e(1+\delta)(1+\varepsilon)\log L$, is bounded by

$$\text{prob}\{\text{depth}(S_M) \geq e(1+\delta)(1+\varepsilon)\log L\} \leq \frac{1}{4}L^{1-e(1+\varepsilon)(1+\delta)\log(1+\delta)}\,, \quad (2.205)$$

where $\delta > 0$.

Proof. In this lemma, we are essentially upper-bounding the result of lemma 2.34, for the case of $M = (1+\varepsilon)L\log L$. This is done in the following way

$$\begin{aligned}\text{prob}\{\text{depth}(S_M) \geq D\} &\leq \binom{M}{D}\left(\frac{4}{L} - \binom{L}{2}^{-1}\right)^{D-1}\,, && (2.206)\\ &\leq \left(\frac{e(1+\varepsilon)L\log L}{D}\right)^{D}\left(\frac{4}{L}\right)^{D-1}\,, && (2.207)\\ &\leq \frac{L}{4}\left(\frac{e(1+\varepsilon)\log L}{D}\right)^{D}\,, && (2.208)\end{aligned}$$

where the second line follows from using the inequality $\binom{N}{k} \leq (\frac{eN}{k})^k$. Therefore, if $D = e(1+\delta)(1+\varepsilon)\log L$, where $\delta > 0$, then we find that

$$\begin{aligned}\text{prob}\{\text{depth}(S_M) \geq e(1+\delta)(1+\varepsilon)\log L\} &\leq \frac{L}{4}(1+\delta)^{-e(1+\delta)(1+\varepsilon)\log L}\,, && (2.209)\\ &\leq \frac{1}{4}L^{1-e(1+\varepsilon)(1+\delta)\log(1+\delta)}\,, && (2.210)\end{aligned}$$

which gives the stated result. □

Therefore, In general, the quantum circuit of the random unitary U_{cycle} for $M = (1+\varepsilon)L\log L$ gates has a depth of at most $\sim \log L$, with probability close to one.

Random graphs and the scrambling time

To summarise, so far I have demonstrated that the quantum circuit of the random unitary U_{cycle} after $M = (1+\varepsilon)L\log L$ random gates applied to independently selected pairs of sites is a connected graph, and has a depth of at most $\sim \log L$, with probability close to one. I will now investigate how many time cycles, t, of the time

evolution operator, so U^t_{cycle}, are required such that any initially local operator or perturbation spreads throughout the entire system - in other words the scrambling time t_{scr}. Before this, I will define some notation and rigorously state what the scrambling time, t_{scr}, is in the context of this model.

Definition 2.36. For the sequence S_M, a path from site x to site y, path(x,y), is a sub-set of S_M such that the pairs of selected sites α_i can be arranged in a (new) sequence of the form $((x,v_1),(v_1,v_2),(v_2,v_3),\ldots,(v_l,y))$, where v_i are all different choices of sites. It is worth noting that path(x,y) is one possible choice of path from the set of all possible paths.

Definition 2.37. We define the length of a path in the following manner:

- the first element, α_i, of the sequence path(x,y) increases the length of path(x,y) (from 0) by $+1$,
- If the next element of (the sequence) path(x,y) occurs chronologically (as defined by S_M) after the previous element then the length remains the same, otherwise the length of the path increases again by $+1$,
- this calculation proceeds in the same way for the next element again.

For an example of this notion of length of a path, which will hopefully make the definition more clear, please refer to figure 2.10. It is worth noting that the definition of length of a path is not (necessarily) symmetric (under exchange of sites x and y), which is also illustrated in figure 2.10.

Definition 2.38. For the sequence S_M, the distance from site x to site y, $\mathrm{d}(x,y)$, is the shortest length path, path(x,y). If the sequence S_M is such that sites x and y are disjoint, in other words no path connects them, then we say $\mathrm{d}(x,y) = \mathrm{d}(y,x) = \infty$.

Definition 2.39. The scrambling time, t_{scr}, is the maximum distance between pairs of sites, $t_{\text{scr}} = \max_{x,y\in\mathbb{Z}_L} \mathrm{d}(x,y)$.

In order to clarify these definitions further, in figure 2.11, there is an illustration of two examples of different paths from the same initial to the same final site.

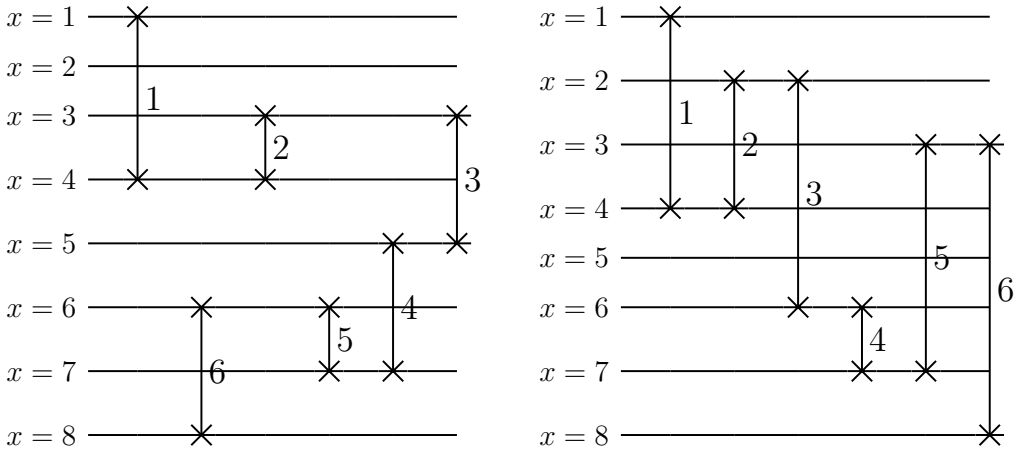

Figure 2.10: This figure is to help illustrate the notion of path length. The two images illustrate two different possible choices of the sequence S_6, a sequence of six interaction pairs in this case on eight sites. In the left-hand image, the length of the path path$(1,8)$ is 4, whereas in the right-hand image, the length of the path path$(1,8)$ is 1. In both images, the numbers next to each pair indicate the order of the pair in the sequence path$(1,8)$. The length of a path is not symmetric, which can be seen in both images. In the left-hand image the length of the path path$(8,1)$ is 3, and in the right-hand image the length of the path path$(8,1)$ is 6.

Since the process which generates the quantum circuit of the random unitary U_{cycle} is itself random, the scrambling time is also a random variable, which can vary between instances of the quantum circuit. In the following lemma, I prove a fairly straight-forward result that the scrambling time cannot be increased by adding more gates into the unitary U^t_{cycle}.

Lemma 2.40. The scrambling time is a non-increasing function of M, and hence

$$\text{prob}\{S_{M'} : t_{\text{scr}} \geq b\} \leq pr\{S_M : t_{\text{scr}} \geq b\} \ , \tag{2.211}$$

for $M \leq M'$.

Proof. If the random quantum circuit, S_M, has some scrambling time t_{scr}, then if we select another pair of sites independently (so go from S_M to S_{M+1}) this additional selection can not increase the distance between any sites at all. The scrambling time can either stay the same or decrease if an additional selection of pairs of sites is made. Since each selection of sites is independent, this argument applies to any number of

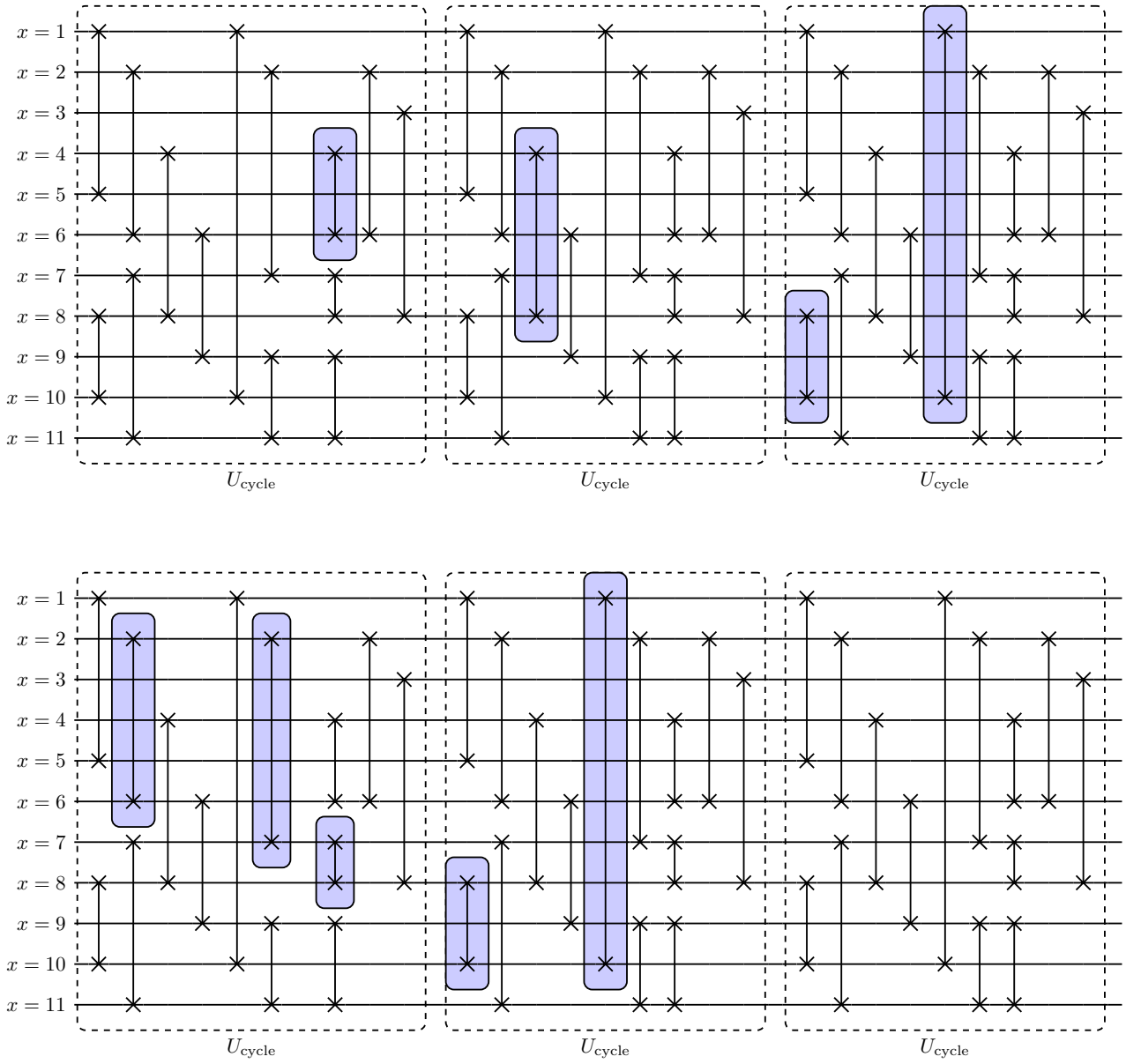

Figure 2.11: In this figure, there is an illustration of two examples of different paths from site 6 to site 1, so path$(6,1)$. The evolution operator, U_{cycle}, in these examples is the same as illustrated in figure 2.7, and is repeated three times, so U_{cycle}^3 . The blue squares surrounding a pair of sites highlights each element in the sequence path$(6,1)$. In the top image, path$(6,1)$ is made up of four elements, or choices of pairs of sites, and has a path length of 3. In the bottom image, path$(6,1)$ is made up of more (five) elements, but has a path length of 2. Hence, it is found that $\mathrm{d}(6,1) = 2$.

additional selections of pairs of sites. Thus, we get the stated result. □

In order to prove statements about the scrambling time in this model, I will use the perspective and results from random graph theory (see references [128–130]). In the study of random graphs there are two well studied models which are of particular relevance: the $G(n,p)$ model and the $G(n,M)$. These are defined in the following way.

Definition 2.41. A $G(n,p)$ random graph is a graph with n vertices, in which every possible edge occurs independently with probability p.

Definition 2.42. A $G(n,M)$ random graph is a graph chosen uniformly at random from the set of graphs with n vertices and M edges.

These models are related, and proving properties of one model can be translated into proofs for the other. In the following two lemmas, I will reproduce two well-known results in the field of random graph theory, which can be found in references such as [128, 129]. (In the case of the following lemma see, for example, chapter 2 theorem 2.2 in the reference [129]).

Lemma 2.43. For any graph property $\mathcal{P}$, and $p = \frac{M}{\binom{n}{2}} \in [0,1]$, where $p(1-p)\binom{n}{2} \to \infty$ when $n \to \infty$, then

$$\text{prob}\{G(n,M) \text{ has } \mathcal{P}\} \leq \sqrt{2\pi M}\text{prob}\{G(n,p) \text{ has } \mathcal{P}\} \ . \tag{2.212}$$

Proof. We first note that we can rewrite the probability, $\text{prob}\{G(n,p) \text{ has } \mathcal{P}\}$, in the following way

$$\text{prob}\{G(n,p) \text{ has } \mathcal{P}\} = \sum_{k=0}^{\binom{n}{2}} \text{prob}\{G(n,p) \text{ has } \mathcal{P} | E_{n,p} = k\}\text{prob}\{E_{n,p} = k\} \ , \tag{2.213}$$

where $E_{n,p}$ is a random variable denoting the number of edges in the $G(n,p)$ random graph. Hence, since $\text{prob}\{G(n,p) \text{ has } \mathcal{P} | E_{n,p} = k\} = \text{prob}\{G(n,k) \text{ has } \mathcal{P}\}$, where $G(n,k)$ is a $G(n,M)$ random graph with k edges, and since in the $G(n,p)$ random

graph each edge is independent and identically distributed (with probability p) we find that

$$\text{prob}\{G(n,p) \text{ has } \mathcal{P}\} = \sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} \text{prob}\{G(n,k) \text{ has } \mathcal{P}\} \ . \quad (2.214)$$

This sum can be lower bounded by considering a single term only, and hence we find

$$\text{prob}\{G(n,p) \text{ has } \mathcal{P}\} \geq \binom{\binom{n}{2}}{M} p^M (1-p)^{\binom{n}{2}-M} \text{prob}\{G(n,M) \text{ has } \mathcal{P}\} \ . \quad (2.215)$$

Using the two inequalities

$$\sqrt{2\pi x}\left(\frac{x}{e}\right)^x < x! < \sqrt{2\pi x}\left(\frac{x}{e}\right)^x e^{1/(12x)} \ , \quad (2.216)$$

to lower bound the binomial coefficient, and that $p = \frac{M}{\binom{n}{2}}$, we find that

$$\text{prob}\{G(n,p) \text{ has } \mathcal{P}\} \geq \frac{1}{\sqrt{2\pi M}} \text{prob}\{G(n,M) \text{ has } \mathcal{P}\} \ , \quad (2.217)$$

which is the stated result. □

With this lemma, one can use the the $G(n,p)$ model to prove results for the $G(n,M)$ model. An improvement of this result can be made for what is known as monotone graph properties. First, I will define what a monotone increasing graph property is and then reproduce a well-known proof (for example see chapter 1, lemma 1.3 of reference [128]).

Definition 2.44. A monotone increasing graph property is a property which is preserved when an additional edge is added to the graph. **Example**: Connectivity and existence of a Hamiltonian path.

Lemma 2.45. For any monotone increasing graph property $\mathcal{P}_{\text{m.i.}}$, and $p = \frac{M}{\binom{n}{2}} \in [0,1/2]$ such that $\lim_{n\to\infty} p\binom{n}{2} \to \infty$, $\lim_{n\to\infty}(1-p)\binom{n}{2}/\sqrt{p\binom{n}{2}} \to \infty$, we have that

$$\text{prob}\{G(n,M) \text{ has } \mathcal{P}_{\text{m.i.}}\} \leq 8 \ \text{prob}\{G(n,p) \text{ has } \mathcal{P}_{\text{m.i.}}\} \ . \quad (2.218)$$

Proof. The proof of this result follows on from the proof of lemma 2.43, and uses the fact that for a monotone graph increasing property $\mathcal{P}_{\text{m.i.}}$

$$\text{prob}\{G(n,M) \text{ has } \mathcal{P}_{\text{m.i.}}\} \leq \text{prob}\{G(n,M') \text{ has } \mathcal{P}_{\text{m.i.}}\} \,, \tag{2.219}$$

for $M < M'$. Therefore, for a monotone graph increasing property $\mathcal{P}_{\text{m.i.}}$

$$\text{prob}\{G(n,p) \text{ has } \mathcal{P}_{\text{m.i.}}\} \geq \text{prob}\{G(n,M) \text{ has } \mathcal{P}_{\text{m.i.}}\} \sum_{k=M}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} \,. \tag{2.220}$$

To proceed we shall now lower bound this sum.

First, we will denote the summand by $a_k = \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k}$ and from lemma 2.43 we have that $a_M > 1/\sqrt{2\pi M}$. We can now lower bound the ratio between consecutive terms in the series by

$$\frac{a_{M+t+1}}{a_{M+t}} = \frac{p}{1-p} \frac{\binom{\binom{n}{2}}{M+t+1}}{\binom{\binom{n}{2}}{M+t}} = \frac{1-\frac{t}{\binom{n}{2}-M}}{1+\frac{t+1}{M}} \,, \tag{2.221}$$

$$\geq \exp\left(-\frac{t+1}{M} - \frac{t}{\binom{n}{2}-M-t}\right) \,, \tag{2.222}$$

where $t \geq 0$, for the second equality we have used that $p = \frac{M}{\binom{n}{2}}$, and for the inequality we have used that $\forall x \; 1+x \leq e^x$ and for $-1 < x < 0$ then $1-x \geq \exp\left(\frac{-x}{1-x}\right)$. Using this inequality on ratio of consecutive terms, we can lower bound the summation by

$$\sum_{k=M}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} > \frac{1}{\sqrt{2\pi M}} \left[1 + \sum_{k=0}^{\binom{n}{2}-M-1} \exp\left(\sum_{t=0}^{k} -\frac{t+1}{M} - \frac{t}{\binom{n}{2}-M-t}\right)\right] , \tag{2.223}$$

$$> \frac{1}{\sqrt{2\pi M}} \sum_{k=0}^{\binom{n}{2}-M-1} \exp\left(-\frac{(k+1)(k+2)}{2M} - \sum_{t=0}^{k} \frac{t}{\binom{n}{2}-M-t}\right) \tag{2.224}$$

If we impose that $M < \frac{1}{2}\binom{n}{2}$ (so $p \leq 1/2$), and then lower-bound the sum further by decreasing the range of k from $\left[0, \binom{n}{2}-M\right]$ to $\left[0, \sqrt{M}\right]$, then we can lower bound

the sum further as

$$\sum_{k=M}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} > \frac{1}{\sqrt{2\pi M}} \sum_{k=0}^{\sqrt{M}} \exp\left(-\frac{(k+1)(k+2)}{2M} - 1\right) . \quad (2.225)$$

Then, since the sum is a decreasing function, we can lower bound the sum using an integral, which itself is like a Gaussian integral in the region $\left[0, \sqrt{M}\right]$. Hence, we find that,

$$\sum_{k=M}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} > \frac{1}{8} . \quad (2.226)$$

Hence, with the lower bound of this sum, we have the stated result. □

An analogous definition and result can be found in the case of monotonically decreasing graph properties. With these definitions of two random graph models, and some proofs which show how they are connected, I will now discuss how they relate to the problem of finding the scrambling time of the random quantum circuit.

The random quantum circuit of the unitary U_{cycle} can be viewed as a graph by representing the sequence S_M as a graph - in other words removing the ordering, and projecting the circuit into the plane. That is to say, the sites $\mathbb{Z}_L$ form the vertices of the graph, which are connected by the edges $\alpha_i \in S_M$. The random graph corresponding to S_M is similar to the $G(n,M)$ model. However, in the $G(n,M)$ model there are no repeated edges within the graph, whereas in the sequence S_M repetitions of selections are allowed, and hence in the random graph corresponding to S_M there can be repeated edges. Hence, before making use of the tools and results of random graphs, I will first bound the probability of the number of repeated choices in the sequence S_M, which is given in the following proof.

Lemma 2.46. The probability that after $M = (1+\varepsilon)L\log L$ independent selections of pairs of sites there are at least R repetitions is upper-bounded by

$$\text{prob}\{S_M : N_{\text{repetitions}} \geq R\} \leq \binom{(1+\varepsilon)L\log L}{R} \left(\frac{(1+\varepsilon)L\log L - R}{\binom{L}{2}}\right)^R , \quad (2.227)$$

where $N_{\text{repetitions}}$ denotes the total number of repeated choices of pairs of sites. Hence, we find that

$$\text{prob}\{S_M : N_{\text{repetitions}} \geq 2e(1+\varepsilon)^2(1+\omega)(\log L)^2\} \leq L^{-2e(1+\varepsilon)^2\omega\log L}\,, \quad (2.228)$$

for $\omega > 0$.

Proof. We will first derive the bound on the probability, and then we will upper-bound this derived quantity further for some specific choice of number of repetitions.

First, we note that the number of choices $M = (1+\varepsilon)L\log L$ is made up D distinct choices and R repeated choices of pairs of sites: so, $M = D + R$. The probability that a single selection α_i is a repetition depends on the number of distinct selections up until this point. However, we can upper-bound the probability that any selection is a repetition with

$$\text{prob}\{S_M : \alpha_i = \alpha_j \text{ for } j < i\} \leq \frac{D}{\binom{L}{2}}\,. \quad (2.229)$$

Hence, using the fact that there are $\binom{M}{R}$ ways of the M selections having R repetitions, then we can upper-bound the probability that the number of repetitions is at least R by

$$\text{prob}\{S_M : N_{\text{repetitions}} \geq R\} \leq \binom{M}{R}\left(\frac{R}{\binom{L}{2}}\right)^R\,, \quad (2.230)$$

To evaluate this bound further, we first note that $\binom{N}{k} \leq (eN/k)^k$, and so

$$\binom{(1+\varepsilon)L\log L}{R}\left(\frac{(1+\varepsilon)L\log L - R}{\binom{L}{2}}\right)^R \leq \left(\frac{2e(1+\varepsilon)\log L}{R(L-1)}((1+\varepsilon)L\log L - R)\right)^R\,. \quad (2.231)$$

We now wish to find the minimum number of repetitions, R, such that the upper bound is small (and tends to zero for large L). For this to be the case, we require that the term within the bracket is smaller than 1, and this is satisfied when $R = 2e(1+\varepsilon)^2(1+\omega)(\log L)^2$ for $\omega > 0$. Evaluating the upper-bound for $R = 2e(1+\varepsilon)^2(1+\omega)(\log L)^2$,

we find that

$$\binom{(1+\varepsilon)L\log L}{R}\left(\frac{(1+\varepsilon)L\log L-R}{\binom{L}{2}}\right)^{R} \leq \left(\frac{L-2e(1+\varepsilon)(1+\omega)}{(1+\omega)(L-1)}\right)^{2e(1+\varepsilon)^2(1+\omega)(\log L)^2}$$

$$\leq (1+\omega)^{-2e(1+\varepsilon)^2(1+\omega)(\log L)^2} \quad (2.232)$$

$$\leq L^{-2e(1+\varepsilon)^2(1+\omega)\log L\log(1+\omega)}, \quad (2.233)$$

which by using the inequality $\log(1+x) > \frac{x}{1+x}$ we get the stated result.

□

To briefly summarise, I have demonstrated how the quantum circuit is related to the $G(n,M)$ random graph model, which also is in turn related to the $G(n,p)$ random graph model. In the following lemma, I will reproduce a well-known proof of the diameter of sparse random graphs [131, 132] (and chapter 10 theorem 10.17 of reference [129], also references [128, 133]). The derived bound on the diameter of the (sparse) random graph $G(n,p)$, for a specific range of values of p, will then be used to bound the scrambling time of the random circuit model. First, I will introduce the following notation.

Definition 2.47. We denote by $\Gamma_k(v)$ as the set of sites at distance k from a specific site v, so $\Gamma_k(v) = \{w \in \mathbb{Z}_L : d(v,w) = k\}$.

Definition 2.48. We call the neighbourhood $N_k(v)$ as the set of sites within a distance k of a specific site v, so $N_k(v) = \bigcup_{j=1}^{j=k} \Gamma_j(v)$.

Definition 2.49. The diameter of a graph G is such that diam$G \leq d$ if and only if $N_d(v) = V(G)$ for every vertex v, with $V(G)$ denoting the set of all vertices of the graph G. This implies that diam$G \geq d$ if and only if it exists a vertex w such that $N_{d-1}(w) \neq V(G)$.

Once again, I would like to emphasise that the following result and rigorous proof is well established [128, 129, 131–133]. The primary modifications and extensions I have made are to include more details and steps, such that the proof is easier to follow, and to state the probability bound explicitly.

Lemma 2.50. For the random graph $G(n,p)$, when $p = \frac{c\log n}{n-1}$ for some constant $c > 4$, the probability that the diameter of the random graph, $\mathrm{diam}\,(G(n,p))$, is $o(\log n)$, in the sense specified below, for large n is upper bounded by

$$\mathrm{prob}\left\{\mathrm{diam}\,(G(n,p)) > \frac{2\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)} + 1\right\} \leq 1.5n^{2-0.5c}\,, \tag{2.234}$$

provided that $c\log n \geq 32$.

Proof. The initial stage of the proof of this result is to consider a fixed vertex v and bound the probability distribution of $|\Gamma_k(v)|$. Firstly, we will bound $|\Gamma_1(v)|$ and then expand from there. The probability distribution of the random variable $|\Gamma_1(v)|$ is a binomial distribution with probability p and $n-1$ total trials. Using the Chernoff bounds (appendix B result B.2), we can bound this distribution to find that

$$\mathrm{prob}\{|\Gamma_1(v)| \geq \gamma c\log n\} \leq n^{\frac{c}{\ln 2}(\gamma - 1 - \gamma\ln\gamma)}\,, \tag{2.235}$$

where $\gamma > 1$, and

$$\mathrm{prob}\{|\Gamma_1(v)| \leq \nu c\log n\} \leq n^{\frac{c}{\ln 2}(\nu - 1 - \nu\ln\nu)}\,, \tag{2.236}$$

with $\nu < 1$. Hence, setting $\gamma = 2$ and $\nu = 1/4$, we have that

$$\mathrm{prob}\{|\Gamma_1(v)| \geq 2\log n\} \leq n^{-0.5c}\,, \tag{2.237}$$

$$\mathrm{prob}\left\{|\Gamma_1(v)| \leq \frac{1}{4}c\log n\right\} \leq n^{-0.58c} \leq n^{-0.5c}\,. \tag{2.238}$$

To bound the probability distribution of $|\Gamma_k(v)|$ we note that conditional on $|\Gamma_j(v)|$ for $j \in [1,k-1]$ the distribution is binomial with probability $p' = 1-(1-$

$p)^{|\Gamma_{k-1}(v)|}$ and number of trials $n-1-|N_{k-1}(v)|$, so

$$\begin{aligned}&\text{prob}\{|\Gamma_k(v)| = i_k||\Gamma_{k-1}(v)| = i_{k-1},\ldots,|\Gamma_1(v)| = i_1\}\\&= \binom{n-1-|N_{k-1}(v)|}{i_k}\left(1-(1-p)^{|\Gamma_{k-1}(v)|}\right)^{i_k}\left((1-p)^{|\Gamma_{k-1}(v)|}\right)^{(n-1-|N_{k-1}(v)|-i_k)}.\end{aligned} \tag{2.239}$$

Once again we will use the Chernoff bounds to bound this probability distribution. We will first focus on the upper-tail, and demonstrate that

$$\text{prob}\{|\Gamma_k(v)| \geq (2c\log n)^k | \bigcap_{j=1}^{j=k-1} l_j \leq |\Gamma_j(v)| \leq (2c\log n)^j\} \leq n^{-0.5c(2c\log n)^{k-1}}, \tag{2.240}$$

where l_j denotes the lower tail of $|\Gamma_k(v)|$, which we will investigate and determine next. Applying the Chernoff bound to the probability distribution of $|\Gamma_k(v)|$, recalling that with the conditioning the distribution is binomial (equation 2.239), we find that

$$\begin{aligned}&\text{prob}\{|\Gamma_k(v)| \geq (2c\log n)^k | \bigcap_{j=1}^{j=k-1} l_j \leq |\Gamma_j(v)| \leq (2c\log n)^j\}\\&\leq \min_{t>0}\exp\left(-t\,(2c\log n)^k + (n-1-|N_{k-1}(v)|)(e^t-1)(1-(1-p)^{|\Gamma_{k-1}(v)|})\right), \qquad (2.241)\\&\leq \min_{t>0}\exp\left(-t\,(2c\log n)^k + (n-1)(e^t-1)p|\Gamma_{k-1}(v)|\right), \qquad (2.242)\\&\leq \min_{t>0}\exp\left(-t\,(2c\log n)^k + (e^t-1)c\log n(2c\log n)^{k-1}\right), \qquad (2.243)\\&\leq \exp\left(2c\log n(2c\log n)^{k-1}(-\ln 2+1/2)\right) \leq n^{-0.5c(2c\log n)^{k-1}}, \qquad (2.244)\end{aligned}$$

which gives that stated result of equation 2.240, and moreover agrees with the calculation for the case $k=1$ in equation 2.237. We shall now investigate the lower

tail (denoted above by l_k), and demonstrate a similar result

$$\text{prob}\left\{|\Gamma_k(v)| \leq \left(\frac{c}{4}\log n\right)^k \middle| \bigcap_{j=1}^{j=k-1}\left(\frac{c}{4}\log n\right)^j \leq |\Gamma_j(v)| \leq (2c\log n)^j\right\}$$
$$\leq n^{-0.5c\left(\frac{c}{4}\log n\right)^{k-1}}, \tag{2.245}$$

with the condition that k is such that

$$1-\frac{2(2c\log n)^{k-1}}{n-1}\left(1+c\log n\frac{((c/4)\log n)^{k-1}}{n-1}\right)^{-1} \geq 0.95\,, \tag{2.246}$$

which essentially imposes a maximum value upon k, and roughly means that $|\Gamma_k(v)|$ is not in the region of $\sim n-1$. The proof of this again makes use of the Chernoff bound to bound the $|\Gamma_k(v)|$, recalling that with the conditioning the distribution is binomial (equation 2.239), we find that

$$\text{prob}\left\{|\Gamma_k(v)| \leq \left(\frac{c}{4}\log n\right)^k \middle| \bigcap_{j=1}^{j=k-1}\left(\frac{c}{4}\log n\right)^j \leq |\Gamma_j(v)| \leq (2c\log n)^j\right\}$$
$$\leq \min_{t>0}\exp\left(t\left((c/4)\log n\right)^k - (n-1-|N_{k-1}(v)|)(1-e^{-t})(1-(1-p)^{|\Gamma_{k-1}(v)|})\right)\,, \tag{2.247}$$

using $|N_{k-1}(v)| \leq 2(2c\log n)^{k-1}$, $|\Gamma_{k-1}(v)| \geq \left(\frac{c}{4}\log n\right)^{k-1}$, and the fact that $(1-(1-p)^{|\Gamma_{k-1}(v)|}) \geq p|\Gamma_{k-1}(v)|(1+p|\Gamma_{k-1}(v)|)^{-1}$, then we can upper bound further to find

$$\leq \min_{t>0}\exp\left(\left((c/4)\log n\right)^{k-1} c\log n\left(\frac{t}{4}-(1-\exp(-t))\frac{1-\frac{2(2c\log n)^{k-1}}{n-1}}{1+c\log n\frac{((c/4)\log n)^{k-1}}{n-1}}\right)\right)\,. \tag{2.248}$$

Imposing the condition given in equation 2.246, and taking the minimum over $t>0$, we find that the upper bound gives

$$\leq n^{-0.52c\left(\frac{c}{4}\log n\right)^{k-1}}\,, \tag{2.249}$$

and hence we have the stated result given in equation 2.245.

The condition given in equation 2.246 imposes an upper bound on k such that the probability bound in equation 2.245 holds. Provided that $c\log n \geq 32$ this condition (equation 2.246) is satisfied for

$$k \leq \frac{\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)} \ . \tag{2.250}$$

However, we should note that this upper-bound on k is more restricted than the condition in equation 2.246.

The result in equation 2.240, which hold for all k that satisfy the condition in equation 2.250, and the result in equation 2.245, which holds without any such condition on k, are conditional probabilities. We will now upper-bound the unconditioned probabilities using the equations themselves, and then use this result to address the main question of this proof (the diameter of $G(n,p)$). Using the abbreviated notation $u = 2c\log n$ and $l = \frac{c}{4}\log n$ to make the equations more compact, we upper-bound in the following manner

$$\begin{aligned}
&\text{prob}\left\{|\Gamma_k(v)| \geq u^k \bigcup |\Gamma_k(v)| \leq l^k\right\} \\
&\leq \text{prob}\left\{|\Gamma_k(v)| \geq u^k \bigcup |\Gamma_k(v)| \leq l^k \middle| \bigcap_{j=1}^{j=k-1} l^j \leq |\Gamma_j(v)| \leq u^j \right\} \\
&+ \left(1 - \text{prob}\left\{\bigcap_{j=1}^{j=k-1} l^j \leq |\Gamma_j(v)| \leq u^j\right\}\right) \ .
\end{aligned} \tag{2.251}$$

We can upper-bound the second term as follows

$$\begin{aligned}
&\left(1-\text{prob}\left\{\bigcap_{j=1}^{j=k-1} l^j \leq |\Gamma_j(v)| \leq u^j\right\}\right) \\
&\leq \left(1-\text{prob}\left\{\bigcap_{j=1}^{j=k-2} l^j \leq |\Gamma_j(v)| \leq u^j\right\}\right) \\
&+\text{prob}\left\{|\Gamma_{k-1}(v)| \geq u^{k-1} \bigcup |\Gamma_{k-1}(v)| \leq l^{k-1} \middle| \bigcap_{j=1}^{j=k-2} l^j \leq |\Gamma_j(v)| \leq u^j\right\}.
\end{aligned} \tag{2.252}$$

Therefore by performing this procedure repeatedly on the $1-\text{prob}\{\cdots\}$ term, we find that the unconditioned probability can be upper-bounded by

$$\begin{aligned}
&\text{prob}\left\{|\Gamma_k(v)| \geq u^k \bigcup |\Gamma_k(v)| \leq l^k\right\} \\
&\leq \sum_{j=1}^{j=k} \text{prob}\left\{|\Gamma_j(v)| \geq u^j \bigcup |\Gamma_j(v)| \leq l^j \middle| \bigcap_{i=1}^{j-1} l^i \leq |\Gamma_i(v)| \leq u^i\right\},
\end{aligned} \tag{2.253}$$

which is just the sum of the conditional probabilities we have upper-bounded previously (equations 2.240 and 2.245). Evaluating this sum, we find that

$$\text{prob}\left\{|\Gamma_k(v)| \geq u^k \bigcup |\Gamma_k(v)| \leq l^k\right\} \leq \frac{2}{1-n^{-0.5c}} n^{-0.5c} \tag{2.254}$$

recalling that $u = 2c\log n$ and $l = \frac{c}{4}\log n$.

At last, we can address the main question. Firstly, we note that when k takes the maximum value allowed by the condition in equation 2.250, then

$$\left|\Gamma_{\frac{\log(0.8(n-1))}{\log(2c\log n)}}(v)\right| \geq \left(\frac{c}{4}\log n\right)^{\frac{\log(0.8(n-1))}{\log(2c\log n)}} \geq \sqrt{0.8(n-1)}, \tag{2.255}$$

once again recalling that $c\log n > 32$. We now consider two fixed different vertices v and w, and consider their neighbourhoods up to this maximum value of k (equation 2.250), $N_{k_{\max}}(v) = N_{\frac{\log(0.8(n-1))}{\log(2c\log n)}}(v)$ and $N_{k_{\max}}(w) = N_{\frac{\log(0.8(n-1))}{\log(2c\log n)}}(w)$. Clearly, if $N_{k_{\max}}(v) \wedge N_{k_{\max}}(w) = \emptyset$ then $\mathrm{d}(v,w) > 2k_{\max}$. Moreover, if there are no edges

connecting any vertices of $\Gamma_{k_{\max}}(v)$ and $\Gamma_{k_{\max}}(w)$, then $\mathrm{d}(v,w) > 2k_{\max}+1$. Hence, we can upper-bound the probability that $\mathrm{d}(v,w) > 2k_{\max}+1$ by

$$\begin{aligned}
&\mathrm{prob}\left\{\mathrm{d}(v,w) > \frac{2\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)}+1\right\} \\
&\leq \mathrm{prob}\{\Gamma_{k_{\max}}(v) \not\leftrightarrow \Gamma_{k_{\max}}(w) \big| l^{k_{\max}} \leq |\Gamma_{k_{\max}}(v)| \leq u^{k_{\max}}, l^{k_{\max}} \leq |\Gamma_{k_{\max}}(w)| \leq u^{k_{\max}}\} \\
&\quad + 2(1-\mathrm{prob}\{l^{k_{\max}} \leq |\Gamma_{k_{\max}}(v)| \leq u^{k_{\max}}\}) \qquad (2.256)
\end{aligned}$$

where the notation $\Gamma_{k_{\max}}(v) \not\leftrightarrow \Gamma_{k_{\max}}(w)$ means that there are no edges connecting any vertices of $\Gamma_{k_{\max}}(v)$ and $\Gamma_{k_{\max}}(w)$. Using results found earlier in this proof (equations 2.254 and 2.255), we can evaluate this upper bound

$$\begin{aligned}
&\mathrm{prob}\left\{\mathrm{d}(v,w) > \frac{2\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)}+1\right\} \\
&\leq (1-p)^{|\Gamma_{k_{\max}}|^2} + 2\mathrm{prob}\left\{|\Gamma_k(v)| \geq u^k \bigcup |\Gamma_k(v)| \leq l^k\right\} \qquad (2.257) \\
&\leq \left(1-\frac{c\log n}{n-1}\right)^{\left(\frac{\log(0.8(n-1))}{\log(2c\log n)}\right)^2} + \frac{2}{1-n^{-0.5c}} n^{-0.5c}\,, \qquad (2.258) \\
&\leq \exp\left(-\frac{c\log n}{n-1}\left(\frac{\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)}\right)^2\right) + \frac{2}{1-n^{-0.5c}} n^{-0.5c}\,, \qquad (2.259) \\
&\leq n^{-0.8c} + \frac{2}{1-n^{-0.5c}} n^{-0.5c} \leq 3n^{-0.5c}\,. \qquad (2.260)
\end{aligned}$$

Finally, the proof so far has only considered two fixed different sites, and the diameter is the maximum distance between two sites of the random graph. Therefore we use the union bound (appendix B result B.1) and sum over $n(n-1)/2$ possible pairs of sites, and hence get the stated result. □

With this proof, I am now in a position rigorously prove result 2.5. In the following lemma, I will use and combine the previous proofs in order to bound the scrambling time, and the depth of the quantum circuit of U_{cycle}.

Lemma 2.51. The quantum circuit of the random unitary, U_{cycle}, as described by the sequence S_M, with $M = (1+\varepsilon)L\log L$, has circuit depth corresponding to the

scrambling time $\sim (\log L)^2$, with high probability. In particular,

$$\begin{aligned}&\text{prob}\{S_M : \text{depth}(S_M) < (e+1)(1+\varepsilon)\log L \text{ and } t_{\text{scr}} \leq \frac{2\log\left(0.8(L-1)\right)}{\log\left(2(1+\varepsilon)\log L\right)} + 1\} \\ &\geq 1 - \frac{1}{4}L^{-\varepsilon} - L^{-(1+\varepsilon)\log L} - 12L^{2-0.5(1+\varepsilon)} \end{aligned} \tag{2.261}$$

with $M = (1+\varepsilon)L\log L$, and with the conditions

$$(1+\varepsilon)\log L \geq 32 \quad \text{and} \quad \frac{\log L}{L} \leq \frac{1+\varepsilon}{4e(1+\varepsilon)+2}\,, \tag{2.262}$$

and the implied condition that $\varepsilon > 3$.

Proof. In the first part of this proof, I will first proof an upper-bound to the scrambling time of the random quantum circuit. In lemma 2.50 there is a proof bounding the probability that the diameter of a $G(n,p)$ random graph is at least a certain value, for the case of sparse random graphs $p \sim (\log n)(n-1)^{-1}$. This property is a monotonically decreasing graph property. By combining lemma 2.50, with the analogous version of lemma 2.45 for the case of monotonically decreasing graph properties, the statement made for the $G(n,p)$ random graph model can be applied to the $G(n,M)$ random graph model. Hence,

$$\text{prob}\left\{\text{diam}\left(G(n,M')\right) > \frac{2\log\left(0.8(n-1)\right)}{\log\left(2c\log n\right)} + 1\right\} \leq 12n^{2-0.5c}\,, \tag{2.263}$$

where $M' = 0.5cn\log n$, where we recall that c is a constant, and there is the requirement that $c\log n \geq 32$.

We now wish to use this result for the diameter of the $G(n,M)$ random graph model to understand the scrambling time of the random circuit, given by the sequence S_M. We can relate the two in the following way, using lemma 2.46. First, we can rewrite the probability $\text{prob}\{S_M : t_{\text{scr}} > T\}$, where T is an arbitrary time, in the

following way

$$\begin{aligned}
&= \text{prob}\{S_M : t_{\text{scr}} > T\} \\
&= \text{prob}\{S_M : t_{\text{scr}} > T | N_{\text{repetitions}} \geq R\} \text{prob}\{S_M : N_{\text{repetitions}} \geq R\} \\
&\quad + \text{prob}\{S_M : t_{\text{scr}} > T | N_{\text{repetitions}} < R\} \text{prob}\{S_M : N_{\text{repetitions}} < R\} \ , \qquad (2.264)
\end{aligned}$$

where $N_{\text{repetitions}}$ and R are the same as in lemma 2.46. This can be upper-bounded by

$$\text{prob}\{S_M : t_{\text{scr}} > T\} \leq \text{prob}\{S_M : N_{\text{repetitions}} \geq R\} + \text{prob}\{S_M : t_{\text{scr}} > T | N_{\text{repetitions}} < R\} \ . \qquad (2.265)$$

The region of M in which we are interested in is $M = (1+\varepsilon) L \log L$, applying the proof in lemma 2.46, and using $\omega = (2e(1+\varepsilon))^{-1} > 0$, then this probability can be upper-bounded further by

$$\begin{aligned}
&\text{prob}\{S_M : t_{\text{scr}} > T\} \qquad (2.266) \\
&\leq L^{-(1+\varepsilon)\log L} + \text{prob}\{S_M : t_{\text{scr}} > T | N_{\text{repetitions}} < (2e(1+\varepsilon)+1)(\log L)^2\} \ .
\end{aligned}$$

To upper-bound further, we note that the scrambling time is a non-increasing function of M (see lemma 2.40), and hence

$$\begin{aligned}
&\text{prob}\{S_M : t_{\text{scr}} > T | N_{\text{repetitions}} < (2e(1+\varepsilon)+1)(\log L)^2\} \\
&\leq \text{prob}\{S^*_{M-(2e(1+\varepsilon)+1)(\log L)^2} : t_{\text{scr}} > T\} \ , \qquad (2.267)
\end{aligned}$$

where $S^*_{M-(2e(1+\varepsilon)+1)(\log L)^2}$ denotes the sequence S_M in which all of the $(2e(1+\varepsilon)+1)(\log L)^2$ repetitions have been omitted. Now that we have this new sequence, in which there are no repetitions, we can use the $G(n,M)$ random graph model to understand the scrambling time. In particular, we can understand the sequence $S^*_{(1+\varepsilon)L\log L-(2e(1+\varepsilon)+1)(\log L)^2}$ as a graph, by using each element of the sequence (α_i) as an edge drawn on a graph of L vertices - in other words by ignoring the ordering

of each element of the sequence. The diameter of this (random) graph is larger than (or equal to) the scrambling time of the sequence, which follows from the definitions 2.36 - 2.39. Therefore,

$$\begin{aligned} &\text{prob}\{S^*_{M-(2e(1+\varepsilon)+1)(\log L)^2} : t_{\text{scr}} > T\} \qquad (2.268)\\ &\leq \text{prob}\left\{\text{diam}\left(G(L,(1+\varepsilon)L\log L-(2e(1+\varepsilon)+1)(\log L)^2)\right) > T\right\}\ , \end{aligned}$$

which we can upper-bound further, using again the fact that the diameter is monotonically decreasing graph property, by

$$\text{prob}\{S^*_{M-(2e(1+\varepsilon)+1)(\log L)^2} : t_{\text{scr}} > T\} \leq \text{prob}\left\{\text{diam}\left(G(L,0.5(1+\varepsilon)L\log L\right) > T\right\}\ , \qquad (2.269)$$

with the added condition that

$$\frac{\log L}{L} \leq \frac{1+\varepsilon}{4e(1+\varepsilon)+2}\ . \qquad (2.270)$$

Hence, using bound on the diameter of the $G(n,M)$ random graph in equation 2.263, and $c=(1+\varepsilon)$, we find that

$$\begin{aligned} &\text{prob}\{S_M : t_{\text{scr}} > \frac{2\log\left(0.8(L-1)\right)}{\log\left(2(1+\varepsilon)\log L\right)}+1\} \qquad (2.271)\\ &\leq L^{-(1+\varepsilon)\log L} + \text{prob}\{S^*_{M-(2e(1+\varepsilon)+1)(\log L)^2} : t_{\text{scr}} > \frac{2\log\left(0.8(L-1)\right)}{\log\left(2(1+\varepsilon)\log L\right)}+1\}\ ,\\ &\leq L^{-(1+\varepsilon)\log L} + 12L^{2-0.5(1+\varepsilon)}\ , \qquad (2.272) \end{aligned}$$

with the conditions

$$(1+\varepsilon)\log L \geq 32 \quad \text{and} \quad \frac{\log L}{L} \leq \frac{1+\varepsilon}{4e(1+\varepsilon)+2}\ . \qquad (2.273)$$

Finally, using the value $M=(1+\varepsilon)L\log L$ and the result for the depth of the

quantum circuit in lemma 2.35, with $\delta = 1/e > 0$, we find that

$$\text{prob}\{\text{depth}(S_M) \geq (e+1)(1+\varepsilon)\log L\} \leq \frac{1}{4}L^{-\varepsilon}\,, \tag{2.274}$$

and applying the union bound, the stated result is obtained. □

This concludes all of the proofs required to justify result 2.5.

2.2.3 Discussion

In summary, in this section I have presented a model with time-periodic dynamics and non-local but few-body interactions. I have demonstrated with rigorous proofs that the circuit depth corresponding to the scrambling time in this model is $\sim (\log L)^2$, where L is the number of sites of the system. However, unlike in the previous section (section 2.1), I have not analysed the question of how closely the dynamics appear to be Pauli mixing, or if under some circumstances can resemble a random unitary. This is an important question that requires answering. Despite the model I presented having interactions given by random Clifford unitaries, none of the proofs in this section are specific to or require this feature. However, to make use of the techniques and proofs from the previous section (section 2.1), the fact that the interactions are random Clifford unitaries will be essential. In fact, this aspect of the model is only relevant for results investigating how closely the dynamics can exhibit Pauli mixing. Hence, result 2.5 applies more generally. Additionally, I would like to emphasise that while I consider a time-periodic model with Clifford dynamics in order to make use of the mathematical results in the previous section, Clifford dynamics will not fully capture the highly chaotic dynamics of black holes. As discussed in the introduction, the Clifford group can at most form a 3-design, and does not therefore exhibit all features of highly chaotic quantum systems. For instance the exponential decay of out-of-time order correlators [19, 134], which is a well-studied diagnostic feature of quantum chaos, is not observed in the case of Clifford dynamics In fact, for any Clifford unitary W and two Pauli operators $\sigma_{\mathbf{u}}, \sigma_{\mathbf{v}}$ the out-of-time order correlator at infinite temperature takes the maximum (absolute) value $|\frac{1}{d}\,\text{tr}\left(\sigma_{\mathbf{u}} W \sigma_{\mathbf{v}} W^\dagger \sigma_{\mathbf{u}} W \sigma_{\mathbf{v}} W^\dagger\right)| = 1$.

As discussed in the introduction (chapter 1 section 1.3), the main motivation for the work in this section is drawn from the study of black holes and the fast scrambling conjecture. Highly chaotic quantum systems are often modelled with (Haar) random unitaries, which is justified by the fact that local random quantum circuits generate unitary designs. However, these circuits are time dependent, while physically one would expect the dynamics to be time-independent. In the fast scrambling conjecture, the main object of interest is finding a random quantum circuit which scrambles quantum information in a depth that is logarithmic in the number of qubits. It is worth considering then whether the model in this section may, with further work and improved proof techniques, approach this limit. The proofs in this section, in particular bounding the scrambling time, make use of well-known results from the study of random graphs. In fact, further results from the study of random graphs prove that with probability close to one the scrambling time is upper-bounded and lower-bounded by $\sim \log L$ for the case where $M \sim L \log L$. Moreover, to achieve a faster scrambling time of, for example, a constant independent of the system size L, would require the driving unitary U_{cycle} to have a much larger number of constituent gates (M), and hence a larger circuit depth. For this reason, I believe that to achieve the desired result of a scrambling time reached in circuit depth logarithmic in the number of qubits, a different (random) circuit generation model is necessary.

Finally, I would also like to draw attention again to results in the references [30, 67], in which it is proven that a random quantum circuit acting on n qubits of depth $\sim (\log n)^3$ can implement a decoupling process. The model I have presented and analysed in this section is in essence a time-periodic model analogous to the model studied in the references [30, 67].

Chapter 3

Localisation

In the introduction of this thesis (chapter 1), I discussed the important role that disorder can have in preventing the onset of thermalisation in quantum systems, in particular in sections 1.4 and 1.5 which discussed single-particle and many body localisation respectively. In this chapter, I will present results investigating the question of localisation in time-periodic random quantum circuits. The main feature I will use to identify the phenomenon of localisation is the time evolution of local operators, which was one of a variety of signatures of localisation discussed in the introduction. In the previous chapter (chapter 2 section 2.1), I presented a model of a time-periodic random quantum circuit in one spatial dimension, and demonstrated that the model can exhibit the mixing of Pauli operators in one regime of the dynamics. In the first section of this chapter (3.1) I will present results for the opposite regime of dynamics, which exhibits a novel type of localisation. In the second section, motivated both by the importance of the number of spatial dimensions in single-particle localisation and by the fact that there is yet to be a detailed understanding or consensus regarding many-body localisation in two or more spatial dimensions, I will present work studying an analogous two dimensional model. I will present a result which shows that a model of a time-periodic quantum circuit in two spatial dimensions does not exhibit localisation, unlike in one dimension, in a strong and precise sense.

In both of the sections the structure is as follows: firstly I will give a general overview of the results including a brief discussion, then I will give detailed

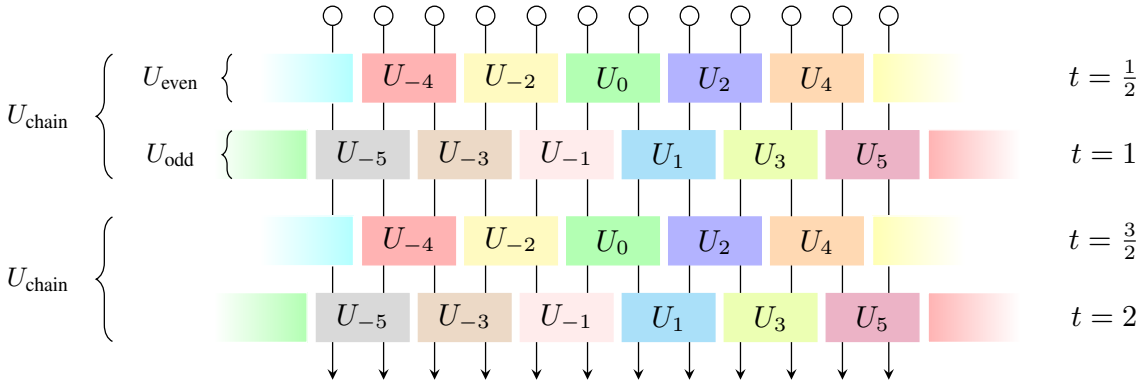

Figure 3.1: Time-periodic local dynamics. This figure illustrates the physical model analysed in this work. The circles on top represent lattice sites, each consisting of N qubits. Coloured blocks represent two-site unitaries, and different colours stand for independently and identically distributed Clifford unitaries, representing the spatial disorder. After the first two half time-steps the dynamics repeats.

derivations of these results, and finally I will more fully discuss the results.

3.1 Strong localisation in one dimension

In chapter 2 section 2.1 I presented results demonstrating that time-periodic random quantum circuits with local interactions, given by random Clifford unitaries, in one regime exhibits Pauli mixing and can in some situations resemble a (Haar) random unitary. This regime of dynamics required that the local spin dimension of each site was larger than the logarithm of the number of sites of the spin chain. In this section, we will investigate the opposite regime when the local spin dimension is small, so the spin chain contains one or only a few qubits per site of the spin chain. I will show that in this regime the time-periodic dynamics results in a novel and strong form of localisation, which means that local operators are strictly confined to a fixed region of the spin-chain at all times. This observed localisation is characterised in terms of one-sided "walls", across which the spreading of perturbations is prevented.

3.1.1 Overview

Firstly, I will describe the model (discussed in chapter 2) once more, so that this section can be followed in a self-contained manner without reference to the previous chapter. The system is a spin chain with L sites and periodic boundary conditions, where each site contains N modes or qubits. The first dynamical period consists of

two half-steps. In the first half-step each even site interacts with its right neighbour with a random Clifford unitary (for the definition and exposition of the Clifford group see appendix A) and in the second half-step each odd site interacts with its right neighbour with a random Clifford unitary. These L Clifford unitaries are independent and uniformly sampled from the $2N$-qubit Clifford group. The subsequent periods of the dynamics are repetitions of the first period, as illustrated in figure 3.1. If we denote by U_x the above-mentioned unitary action on sites x and $x+1$ (modulo L due to periodic boundary conditions) then the evolution operator after an *integer* time t is

$$\begin{aligned} W(t) &= \left[(U_1 \otimes U_3 \otimes \cdots \otimes U_{L-1})(U_0 \otimes U_2 \otimes \cdots \otimes U_{L-2})\right]^t \\ &= (U_{\text{odd}} U_{\text{even}})^t = (U_{\text{chain}})^t \ . \end{aligned} \tag{3.1}$$

As discussed in the previous chapter, this model exhibits the mixing of Pauli operators in the regime $N \gg \log L$, and hence in this section the focus is instead on the opposite regime of $N \ll \log L$, which in particular includes the case of a single qubit per site $N = 1$. In this section, the feature used to study and identify the phenomenon of localisation is the time evolution of initially local operators, $A(t) = W(t)^\dagger A W(t)$, whereby initially local operator I mean an operator which at $t = 0$ is supported (non-identity) on a single site of the spin chain. This model has the property that specific combinations of unitaries acting on consecutive sites, $U_x, U_{x+1}, \ldots, U_{x+l}$, generate left or right sided "semi-permeable" walls. By "semi-permeable" wall I mean the following: a right-sided "semi-permeable" wall at site x prevents the growth of the support of any operator from the left of x towards the right of x, but it does not necessarily stop the growth of any operator from the right to the left of x. An analogous definition applies for left-sided walls. This effect, and how it produces localisation, is well demonstrated in figure 3.2. In this figure, the time evolution of the initial operator σ_z at site $x = 1$ is shown to initially spread linearly until colliding with both left and right sided walls, which prevents further propagation for all times.

The one-sided walls have some penetration length l into a forbidden region. To illustrate this, consider a specific realisation of $W(t)$ with a right-sided wall at site

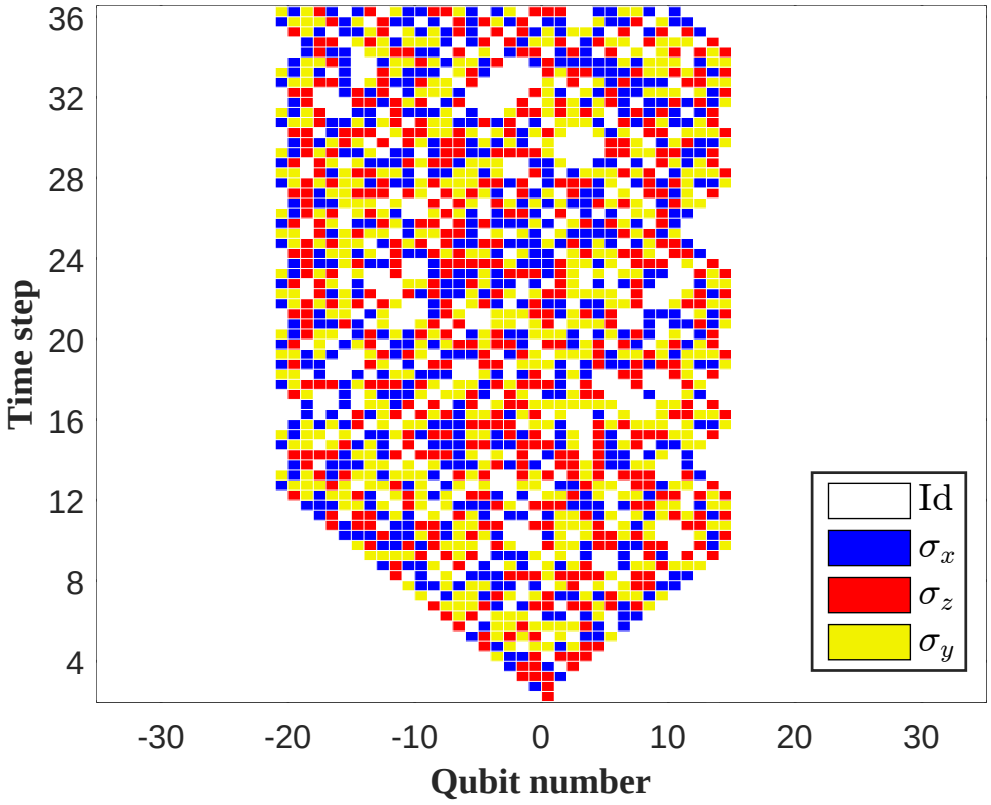

Figure 3.2: This figure displays the Heisenberg evolution of the initial operator σ_z at site $x = 1$. Each lattice site consists of one qubit ($N = 1$) with first-neighbour interactions. After a phase of linear growth the lateral wings collide with left- and right-sided walls with penetration length $l = 1$, that confine the evolution for all times. This confinement affects all (not necessarily local or Pauli) operators between the two walls. Inside the confined region, the evolution seems to mix Pauli operators. This figure was produced by a collaborator, Álvaro M. Alhambra, who also performed the required numerics.

$x = 0$ with penetration length l. Any operator which is supported (non-identity) only on sites $x \leq 0$ evolves in time, mapped by $W(t)$, to an operator supported only on sites $x \leq l$, and not supported (so non-identity) on sites $x > l$ for all $t \geq 1$. Whereas, it should be noted that, if the initial operator was supported on the region of sites given by the interval $x \in [1, l]$ it is possible that the time evolution allowed the operator to be mapped to an operator with support on sites $x > l$.

This property of combinations of Clifford unitaries generating walls can be most clearly seen in the case of a single unitary U_x, which forms a wall with zero penetration length $l = 0$ and hence is in fact a two-sided wall. These walls occur

when a single unitary U_x is of product form $U_x = V_x \otimes V_{x+1}$, wich is equivalent to a non-interacting unitary. I will refer to this case as trivial localisation, as in essence is equivalent to the spin chain being split at site x into two completely independent parts. The following result gives the probability of these trivial ($l = 0$) walls.

Result 3.1. The probability that a Clifford unitary $U \in \mathcal{C}_{2N}$ is of product form is

$$\frac{1}{2}2^{-4N^2} \leq \text{prob}\{U \text{ is product }\} \leq 2^{-4N^2} \ . \tag{3.2}$$

The more interesting case is considering combinations of more than one unitary, which through one-sided walls produce non-trivial localisation. I shall now characterise pairs of gate $U_x,\ U_{x+1}$, which act on the pair of sites $\{x, x+1\}$ and $\{x+1, x+2\}$ respectively, that produce a right-sided wall at site x with penetration length $l \leq 1$. I shall again use the phase-space representation of Clifford unitaries, which has been discussed in chapter 2 and in more detail in appendix A. The unitaries $U_x,\ U_{x+1}$ have phase-space representations $S_x,\ S_{x+1}$, for which subsystems are decomposed with direct sums rather than tensor products. Hence, $S_x,\ S_{x+1}$ (symplectic matrices with entries in $\mathbb{Z}_2$) decompose in to $2N$-dimensions blocks

$$S_x = \begin{pmatrix} A_x & B_x \\ C_x & D_x \end{pmatrix} \ . \tag{3.3}$$

The growth of the support of an operator caused by S_x, which can also be referred to as the flow of information, can be understood by considering the action of S_x on a ($\mathbb{Z}_2$) vector $(\mathbf{u}_x, \mathbf{u}_{x+1})^T$

$$\begin{pmatrix} A_x & B_x \\ C_x & D_x \end{pmatrix} \begin{pmatrix} \mathbf{u}_x \\ \mathbf{u}_{x+1} \end{pmatrix} = \begin{pmatrix} A_x\mathbf{u}_x + B_x\mathbf{u}_{x+1} \\ C_x\mathbf{u}_x + D_x\mathbf{u}_{x+1} \end{pmatrix} \ . \tag{3.4}$$

So, the blocks A_x and D_x represent local dynamics at sites x and $x+1$ respectively. Whereas, C_x (and analogously for B_x) represents interactions between sites x and $x+1$, and in particular the propagation of the operator from site x to $x+1$. This idea is illustrated in figure 3.3.

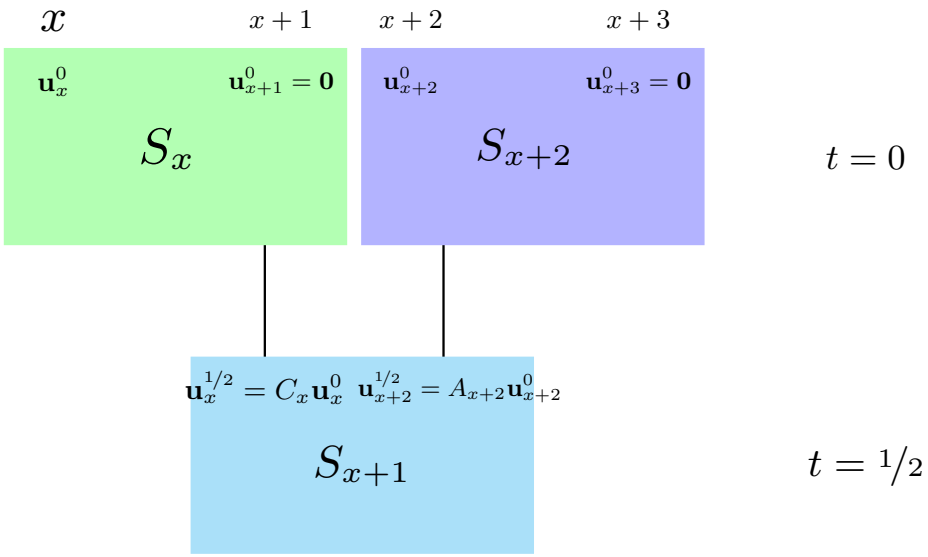

Figure 3.3: Information flow. The flow of information in phase space according to eq. (3.4) is illustrated in the particular case $\mathbf{u}_x^0 = \mathbf{u}_{x+2}^0 = 0$, where $\mathbf{u}_x^t$ denotes the projection of the phase-space vector $\mathbf{u}^t = S(t)\mathbf{u}^0$ onto the subspace $\mathcal{V}_x$ of site x. For graphical convenience we write the inputs $(\mathbf{u}_x, \mathbf{u}_{x+1})$ to each symplectic block matrix S_x inside the block. Note that the input at time $t = \frac{1}{2}$ at site $x+1$ $(x+2)$ is equal to the output given by S_x (S_{x+2}) at the site $x+1$ $(x+2)$. This figure is based on another figure which was produced by a collaborator, Daniele Toniolo.

For the pair of unitaries, U_x, U_{x+1} to form a right-sided wall it must be imposed that no operator can spread to the site $x+2$ (from site x) at any time. For the first time-step ($t = 1$), this is equivalent to the following condition on the phase-space representation: $C_{x+1}C_x = 0$, and more generally for any integer number of time-steps $t > 0$ this requires imposing the conditions

$$C_1(D_0A_1)^t C_0 = 0\,. \tag{3.5}$$

For the simplest case of $N = 1$, a spin chain of qubits, this infinite family of conditions in equation 3.5 reduces to only two conditions ($t = 0$ and $t = 1$). Whereas for the case $N > 1$ this infinite family of conditions is instead implied by the cases $t = 0, 1, \ldots, (4N^2 - 1)$. The conditions in equation 3.5 can be understood as characterising a pattern of destructive interference due to disorder which causes localisation. To illustrate this more clearly, I present the following example for the case of $N = 1$

for a pair of Clifford unitaries U_x, U_{x+1}

$$U_x = \frac{1}{\sqrt{2}} \begin{pmatrix} i & 0 & 0 & -i \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad U_{x+1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}, \tag{3.6}$$

which have phase-space representations

$$S_0 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.7}$$

which as a pair of matrices do satisfy the conditions in equation 3.5, and hence produce a wall.

The probability of this type of one-sided wall, those with penetration length $l \leq 1$ is bounded in the following result. The inverse of this probability is comparable to the average distance between walls, and so can be understood as the localisation length scale.

Result 3.2. The probability that a pair of gates $U_x, U_{x+1} \in \mathcal{C}_{2N}$ generates a right-sided wall with penetration length $l \leq 1$ is

$$\text{prob}\{\forall t \geq 0, C_{x+1} (D_x A_{x+1})^t C_x = 0\} \leq 4N 2^{-2N(N-1)} . \tag{3.8}$$

For $N = 1$ the exact probability is

$$\text{prob}\{C_{x+1} D_x A_{x+1} C_x = 0 \text{ and } C_{x+1} C_x = 0\} = 0.12 . \tag{3.9}$$

By symmetry, left-sided walls have the same probabilities.

Both results demonstrate an exponential dependence on N, and hence suggests that the distance between walls increases rapidly with increasing N. Unfortunately,

I have only been able to characterise and calculate the probability for walls with penetration length $l \leq 1$. I expect that walls with penetration length $l = 0$, so trivial localisation, are less likely than walls with penetration length $l \geq 1$, which would allow for a regime of L and N in which the system does not display trivial localisation. However, this is a conjecture. Although, one could instead consider an equivalent model in which the Clifford unitaries are sampled uniformly, excluding those which are of product form.

Finally, it is worth emphasising that since these walls prevent the propagation of all Pauli operators, then they also prevent the propagation of all operators, because the Pauli operators form a basis. Hence, all initially local operators are confined to a fixed and finite region of the spin chain.

3.1.2 Details and derivation

In this section, I will give in detail the mathematical derivations for both result 3.1 and result 3.2. The proof for result 3.1 is given in lemma 3.1. The set of proofs which give result 3.2 are given in lemmas 3.2 and 3.3.

The case of trivial localisation, described in result 3.1, occurs when a Clifford unitary is of product form, and so the two sites upon which the unitary acts do not interact. Since the quantum circuit is time-periodic this non-interacting Clifford unitary is repeated periodically, and so the two sites never interact at all, and hence there is a wall. The product form of the unitary, U, has an equivalent phase-space description, in which the corresponding symplectic matrix S is written in the form of a direct sum, so $S = T \oplus T'$, which when S is written in block matrix form is equivalent to $B = C = 0$. The following lemma bounds the probability of a uniformly sampled ($\mathbb{Z}_2$) symplectic matrix being in direct sum form.

Lemma 3.1. Any given $S \in \mathcal{S}_{2N}$ can be written in block form

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3.10}$$

according to the decomposition $\mathbb{Z}_2^{4N} = \mathbb{Z}_2^{2N} \oplus \mathbb{Z}_2^{2N}$, and if S is uniformly distributed

then this induces a distribution on the sub-matrices A,B,C,D. For each of the sub-matrices ($E = A,B,C,D$) the induced distribution satisfies

$$\frac{2^{-4N^2}}{2} \leq \text{prob}\{E = 0\} = \frac{|\mathcal{S}_N|^2}{|\mathcal{S}_{2N}|} \leq 2^{-4N^2} , \tag{3.11}$$

with the implied additional property that $\text{prob}\{A = 0|D = 0\} = \text{prob}\{D = 0|A = 0\} = \text{prob}\{B = 0|C = 0\} = \text{prob}\{C = 0|B = 0\} = 1$.

Proof. First consider the case where $C = 0$. Following the algorithm for generating a symplectic matrix in lemma 2.1, we see that A must be $(2N \times 2N)$ symplectic matrix. Hence, any choice for the columns of B will have symplectic form of one with at least one column of the matrix A. Therefore, to fulfil the symplectic constraints for the entire matrix S, the corresponding column of D must have symplectic form of one with a column of C. However, this is not possible since $C = 0$, therefore $B = 0$ and both A and D are $(2N \times 2N)$ symplectic matrices. Again using the counting algorithm in lemma 2.1, the number of choices of S with $C = 0$ is given exactly by

$$|\{S \in \mathcal{S}_{2N} : C = 0\}| = |\mathcal{S}_N||\mathcal{S}_N| = |\mathcal{S}_N|^2 . \tag{3.12}$$

Finally, dividing by the total number of choices for S gives the probability.

I will now demonstrate that this argument applies to any of the four sub-matrices A,B,C,D. Consider the symplectic matrix

$$M = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} , \tag{3.13}$$

and its left and right action on the symplectic matrix S

$$SM = \begin{pmatrix} B & A \\ D & C \end{pmatrix} , \quad MS = \begin{pmatrix} C & D \\ A & B \end{pmatrix} , \tag{3.14}$$

so M permutes the block matrix S by swapping the rows and columns. Since, the product of symplectic matrices is also a symplectic matrix, we have that the argument

above for the case of $C = 0$ applies to any of the four sub-matrices A, B, C, D.

Finally, the upper and lower bounds for the probability are found using lemma 2.2. □

This lemma includes the statement of result 3.1 for the case where $C = 0$. As I have already mentioned, this trivial form of localisation is equivalent to having the spin chain split into independent parts. In the rest of this section, I will investigate other conditions for localisation which are not trivial, and occur due to interacting dynamics. The following two lemmas together constitute the proof for result 3.2. The first lemma gives a full justification for the infinite set of conditions which result in a pair of Clifford unitaries (U_x, U_{x+1}) forming a wall of penetration length $l \leq 1$. In this lemma, it is also found that the infinite family of conditions can in fact be reduced to a set of $4N^2$ conditions.

Lemma 3.2. The conditions

$$C_{x+1} \left(D_x A_{x+1}\right)^k C_x = 0, \tag{3.15}$$

for all $k \in \mathbb{N}$, are sufficient for preventing all right-wards propagation past position x at any time. The probability of this is upper-bounded by

$$\text{prob}\{C_{x+1} \left(D_x A_{x+1}\right)^k C_x = 0, \forall k \in \mathbb{N}\} \leq \frac{2N+1}{(1-2^{-2N})^{2N}} 2^{2N-2N^2}, \tag{3.16}$$

Proof. This proof is clearer with reference to figure 3.1. The condition $C_{x+1}C_x = 0$ prevents right-wards propagation for a single time-step, however (unless $C_x = 0$) then $A_{x+1}C_x \neq 0$ and hence in subsequent time-steps there could be right-wards propagation. In the next time-step, the only way for possible right-ward propagation to occur that would not be blocked by the condition $C_{x+1}C_x = 0$ is $C_{x+1}D_xA_{x+1}C_x$, and so the additional requirement $C_{x+1}D_xA_{x+1}C_x = 0$ prevents right-ward propagation. Once again the same argument applies for subsequent time-steps, and hence we require that $C_{x+1} \left(D_x A_{x+1}\right)^k C_x = 0$ for $k \geq 2$ ($k \in \mathbb{N}$).

Finally, to obtain the upper-bound for the probability, first note that

$$\text{prob}\{C_{x+1}\left(D_xA_{x+1}\right)^k C_x = 0, \forall k \in \mathbb{N}\} \leq \text{prob}\{C_{x+1}C_x = 0\}\ , \tag{3.17}$$

and hence by applying the result in lemma 2.5 we get the stated result. □

I would like to point out that this lemma gives a sufficient condition, and there are other potential conditions and mechanisms by which right-wards propagation can be prevented. For example, one could instead consider conditions such that three or more Clifford unitaries prevent right-wards propagation ($l \geq 2$ penetration length).

This lemma applies for any general choice of N. In the following lemma I will consider the particular case of $N = 1$, a spin chain with a single qubit per site, and demonstrate that the family of conditions actually reduces to just two conditions, and give exactly the probability of these two conditions being satisfied. The following lemma, combined with the previous lemma, are the justification for result 3.2.

Lemma 3.3. For $N = 1$ the conditions

$$C_{x+1}\left(D_xA_{x+1}\right)^k C_x = 0, \tag{3.18}$$

for $k = 0, 1, 2, \ldots$ are implied by the two conditions

$$C_{x+1}C_x = 0 \quad \text{and} \quad C_{x+1}D_xA_{x+1}C_x = 0\ . \tag{3.19}$$

Furthermore the probability of this is given exactly by

$$\text{prob}\{C_{x+1}C_x = 0, C_{x+1}D_xA_{x+1}C_x = 0\} = 0.12\ \ , \tag{3.20}$$

which includes trivial localisation.

Proof. In this result we are concerned only in the case $N = 1$, for which $S_x,\ S_{x+1}$ are 4×4 symplectic matrices and the sub blocks A, B, C, D are 2×2 matrices. We first note that if $C_x = 0$ and/or $C_{x+1} = 0$, which is trivial localisation, then it is clear that the conditions for all k are satisfied. Hence, we now focus only on the cases

where $C_x \neq 0$ and $C_{x+1} \neq 0$. Moreover, we note that we will only focus on the cases where $\text{Rank}(C_x) = \text{Rank}(C_{x+1}) = 1$, since if either of C_x or C_{x+1} are full rank then to satisfy $C_{x+1}C_x = 0$ the other of the C matrices must be the zero matrix.

When $\text{Rank}(C_{x+1}) = 1$ the columns of C_{x+1} have symplectic form of zero with themselves, or in other words $C_{x+1}^T J C_{x+1} = 0$. Therefore, by the symplectic conditions, this implies that A_{x+1} is a 2×2 symplectic matrix. This argument also applies to C_x, and so D_x is also a 2×2 symplectic matrix.

Therefore, since the product of symplectic matrices is also a symplectic matrix, for $N = 1$ neglecting the cases of trivial localisation ($C_x = 0$ and/or $C_{x+1} = 0$) the conditions become

$$C_{x+1} S^k C_x = 0, \tag{3.21}$$

where S is a generic 2×2 symplectic matrix. For 2×2 symplectic matrices, there exists $\alpha,\ \beta \in \mathbb{Z}_2$ such that the following holds

$$S^2 = \alpha \mathbb{I} + \beta S\ , \tag{3.22}$$

which can be verified by a direct check. So, if $C_{x+1}C_x = 0$ and $C_{x+1}SC_x = 0$ both hold then $C_{x+1}S^kC_x = 0$ for all $k > 1$. Hence, we have the stated result for $N = 1$.

The exact result for the probability for the case of $N = 1$ follows from directly counting the number of symplectic matrices that satisfy the two conditions. □

Unfortunately, it has not been possible to prove for the cases $N > 1$ a similar reduction of the family of conditions to some smaller number. Indeed, one can construct many counter examples for the case $N = 2$ which shows that more than two conditions ($k = 0, 1$) are required.

3.1.3 Discussion

To summarise, I have shown that the time-periodic quantum circuit considered in the regime of small N ($N \ll \log L$), one or a few qubits per site of the spin chain, exhibits localisation. This novel form of localisation, which I have characterised in terms of "semi-permeable" (one-sided) walls, is demonstrated by observing that all initially

local (only non-identity on a single site) operators over their time evolution are only supported on the same fixed finite region of the spin chain. This finite region, in terms of one-sided walls, is the space between a left-sided wall and a right-sided wall.

Firstly, I will clarify in what sense this form of localisation is novel. In the introduction, I discussed both Anderson localisation (chapter 1 section 1.4) and many-body localisation (chapter 1 section 1.5). In many-body localised systems, the support of initially local operators grows logarithmically with time. Whereas in Anderson localisation, operators are confined to a finite region for all times, with exponentially decaying tails beyond the region. The form of localisation I have described and characterised is novel in that operators are localised to the same fixed and finite region of the spin chain. This strict localisation is reminiscent of Anderson localisation, however, the model considered is an interacting many-body quantum system which cannot be understood as a system of free or weakly interacting particles, and so is unlike Anderson localisation. Additionally, the localisation does not behave in the same manner as many-body localisation, and hence seems to challenge the existing classification.

Naturally, the results in this section give impetus to further research. In particular, a full characterisation of this novel type of localisation is required. Moreover, it is not clear if this phenomenon is robust against perturbations, as one would expect for a phase of matter. By this I mean: does the observed phenomenon persist when one considers the more general case by including non-Clifford unitaries in the time-periodic dynamics?

In the next section, I will investigate a related question: in two spatial dimensions does an analogous model still demonstrate this novel type of localisation?

3.2 Absence of localisation in two dimensions

The number of spatial dimensions is of critical importance to both single particle and many-body localisation. This was discussed in chapter 1 sections 1.4 and 1.5. Importantly, for many-body localisation it is not yet fully resolved whether the

	Quasi-free systems	Clifford dynamics	Chaotic many-body
Phase-space rep.	✓	✓	×
1D localisation	✓	✓	✓
2D localisation	High disorder	×	?
Order of design	0	3	∞

Table 3.1: This table broadly illustrates the comparison between observable physical behaviour in quasi-free systems (bosons/fermions), systems with Clifford dynamics, and chaotic many-body quantum systems.

phenomenon persists in greater than one spatial dimension. Whereas, in the case of Anderson localisation the role of the number of spatial dimensions in determining the observed behaviour is well understood. In the previous section (section 3.1), I presented results demonstrating a novel type of localisation for a one-dimensional spin chain with dynamics generated via a time-periodic random quantum circuit. In this section, I will present results from investigating an analogous model in two spatial dimensions. I will show that in fact there is a strong and precise sense in which there is an absence of localisation. In table 3.1, I have briefly summarised a comparison between the observed physical behaviour of systems with quasi-free bosons or fermions and systems with Clifford dynamics. This comparison will be discussed in greater detail in the final chapter.

3.2.1 Overview

Initially, I will define precisely what I mean by the absence of localisation, where again the perspective taken is the time evolution of initially local operators. Then, I will present a model, which is a two-dimensional analogue of the model previously considered (chapter 2 section 2.1 and chapter 3 section 3.1) with $N = 1$ and four site interactions, and show a result demonstrating the absence of localisation.

The restricted growth of initially local operators with the evolution of time is a well established feature indicating localisation. That is to say, if an initially local operators grows for some time before eventually remaining confined to some finite region, then there is localisation. Additionally, if an initially local operator grows indefinitely but at a very slow rate (such as logarithmically), then again I would identify this as localisation. Whereas, if an initially local operator grows in some

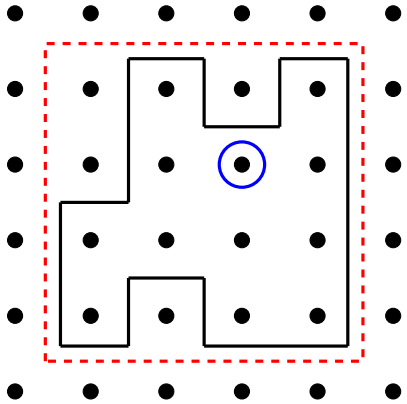

Figure 3.4: This figure gives an illustration of an example of light-speed operator growth, which implies the absence of localisation, in two spatial dimensions, where each black circle represents a qubit. Consider an operator which initially (at $t = 0$) is only non-identity on the qubit highlighted by the (blue) circle. At time $t = 1$ according to the circuit described in figure 3.5, this initially local operator has some light-cone represented by the dashed red line. If, for example, at this specific time, t, the operator is non-identity on all the sites within the region indicated by the continuous black line, then this operator displays light-speed growth given in definition 3.4, which implies the absence of localisation

direction indefinitely at linear rate, there is no localisation. With this in mind, I propose the following definition, which will be used throughout this section.

Definition 3.4. Light-speed operator growth: Consider an initially ($t = 0$) local operator (non-identity only on a single site). If for all times $t \in \mathbb{Z}$, the time evolution of this initially local operator has non-trivial support (so non-identity) on the boundary of the light-cone then the operator has light-speed growth in (at least) one direction.

In figure 3.4, there is an example to illustrate this definition. I would like to emphasise that this definition implies an absence of localisation, and indeed light-speed operator growth is a sufficient condition for a system to display the absence of localisation. This definition while extreme, has the benefit of being uncontroversial. However, I should note that since this definition only concerns what happens on the boundary of the causal light-cone, if the condition in the definition is violated it does not imply that the system displays localisation. For example, if at all times the time-evolved version of the initially local operator is supported on all sites except

those on the boundary of the light-cone, then the system can not be said to display localisation. In other words, some initially local operators can grow in some direction indefinitely at a linear, but not maximal, rate. However, I will not investigate or prove anything about operators that grow in this way. In this section I will present results showing that in both models considered, there is a light-speed operator growth, which implies the absence of localisation, with some constant minimum probability, which can also be interpreted as a lower bound on the fraction of operators which grow in some direction indefinitely at a linear rate.

The model consider in this section is an $L \times L$ square lattice of qubits. Each site of the lattice is labelled by (x,y) where $x,y \in \mathbb{Z}_L$. Hence, the total Hilbert space for the entire system has dimension 2^{L^2}. The dynamics of the spin system is discrete in time, and so is characterised by a unitary W_{cycle} (and not a Hamiltonian). Since the dynamics are time-periodic, each evolution time-step is characterised by the same unitary W_{cycle}, and hence $|\psi(t)\rangle = \left(W_{\text{cycle}}\right)^t |\psi(0)\rangle$ for all $t \in \mathbb{Z}$. Locality is imposed by the fact that the unitary W_{cycle} is generated via four-qubit interactions in the following way

$$W_{\text{cycle}} = \left(\bigotimes_{x \text{ odd}, y \text{ odd}} W_{(x,y)} \right) \left(\bigotimes_{x \text{ even}, y \text{ even}} W_{(x,y)} \right) , \tag{3.23}$$

where the unitary $W_{(x,y)}$ acts on the sites $(x,y),(x+1,y),(x,y+1),$ and $(x+1,y+1)$. This expression indicates that each time cycle decomposes into two half time-steps, which is illustrated in figure 3.5. Therefore, note that alternatively I will refer to the number of layers in the quantum circuit, l, as the duration of the time evolution, so $l = 2t$. Furthermore, each unitary $W_{(x,y)}$ is selected independently from the uniform distribution over the Clifford group (acting on four qubits).

Result 3.3. Consider a Pauli operator which is non-identity on only one site at $t = 0$. With probability at least 0.4, the time evolution of this initial operator for the model with time-periodic dynamics as described in equation 3.23 is non-identity on some qubits on the boundary of the light-cone at all times $t \in \{1/2, 1, 3/2, 2, 5/2, \ldots\}$. Hence, this model displays light-speed operator growth, which implies the absence

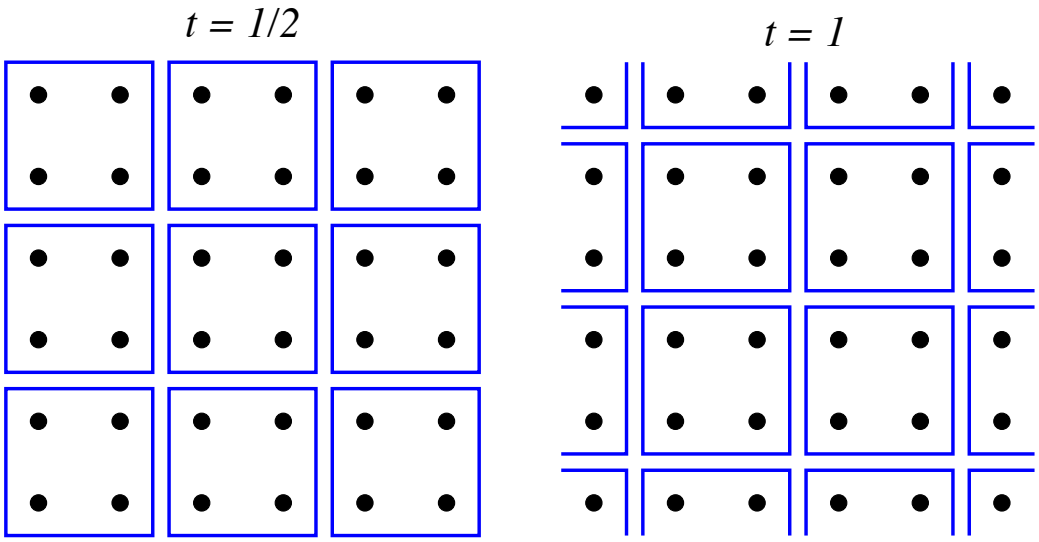

Figure 3.5: This figure illustrates the model we are considering, in which the driving circuit is depth 2; it is composed of two half time-steps. All unitaries, which are represented by blue squares, act upon 4 qubits, which are represented by dots.

of localisation.

In summary, in this model there is an absence of localisation, which is implied by the light-speed operator growth established in definition 3.4; a significant fraction of initially local operators grow at the maximum linear rate in some direction indefinitely. Crucially, this lower bound for the probability is independent of the size of the system and the time. Although, I should point out that for finite systems these results only hold for the initial spreading of the local operator, so $t \sim L$, since after this time the causal light-cone of the initially local operator is the entire system. Nonetheless despite this, it is fair to conclude that there is an absence of localisation. Strikingly, this result holds in the thermodynamic limit, $L \to \infty$, and so the majority of local operators grow indefinitely.

3.2.2 Details and derivation

I will now give the arguments and mathematical details which justify result 3.3, which is given in lemmas 3.5, 3.7, 3.8, and 3.9 .

Firstly, I will describe how the spreading of an initially local operator due to the time-periodic dynamics, as described in equation 3.23 and illustrated in figure 3.5, can be interpreted as a directed graph, when considering only the boundary of

the light-cone. For an initially local operator (supported on a single site at $t = 0$), the shape of the light-cone boundary at any time $t \in \{1/2, 1, 3/2, 2, \ldots\}$ is a square centred upon the initial site, with $4(4t-1)$ qubits on this boundary (perimeter). The state of the operator on the light-cone boundary at any time t depends only on the state of the operator on the light-cone boundary at time $t - 1/2$, and importantly does not on the operator inside the "bulk" of the light-cone. Hence, the main reason for focussing on the boundary of the light-cone is that at each time-step the random Clifford unitaries are statistically independent from all previous random unitaries, which makes it mathematically tractable.

The evolution of the light-cone boundary from one half time-step to the next proceeds in the following general way, an illustrative example of which is depicted in figure 3.6. The four qubits situated on the four corners of the square light-cone boundary at time t each interact with their three nearest neighbours outside of the light-cone via a random Clifford unitary. Each set of three qubits outside of the light-cone boundary at time t are qubits on the light-cone boundary at time $t + 1/2$. One qubit in each of the four sets of three qubits form the corners of the square light-cone boundary at time $t + 1/2$, and the other two qubits in each of the sets are qubits on the edge of the light-cone boundary at time $t + 1/2$. The qubits situated on the edges of the square light-cone boundary at time t each interact with one of their nearest neighbours within the light-cone boundary at time t, and the two nearest qubits outside the light-cone boundary, via a random Clifford unitary. The two interacting qubits on the outside of the light-cone boundary become qubits on the light-cone boundary at time $t + 1/2$. This interaction pattern for the light-cone boundary can be translated into a directed graph in the following way. Each random Clifford unitary is represented by a vertex, and each qubit which becomes a qubit on the boundary of the light-cone in the following half time-step is represented as an outwards arrow from the corresponding vertex. So, unitaries which act on the corners of the light-cone boundary become a vertex with one inwards arrow and three outwards arrows, and similarly unitaries which act on the edges of the light-cone boundary become a vertex with two inwards arrows and two outwards arrows. All of

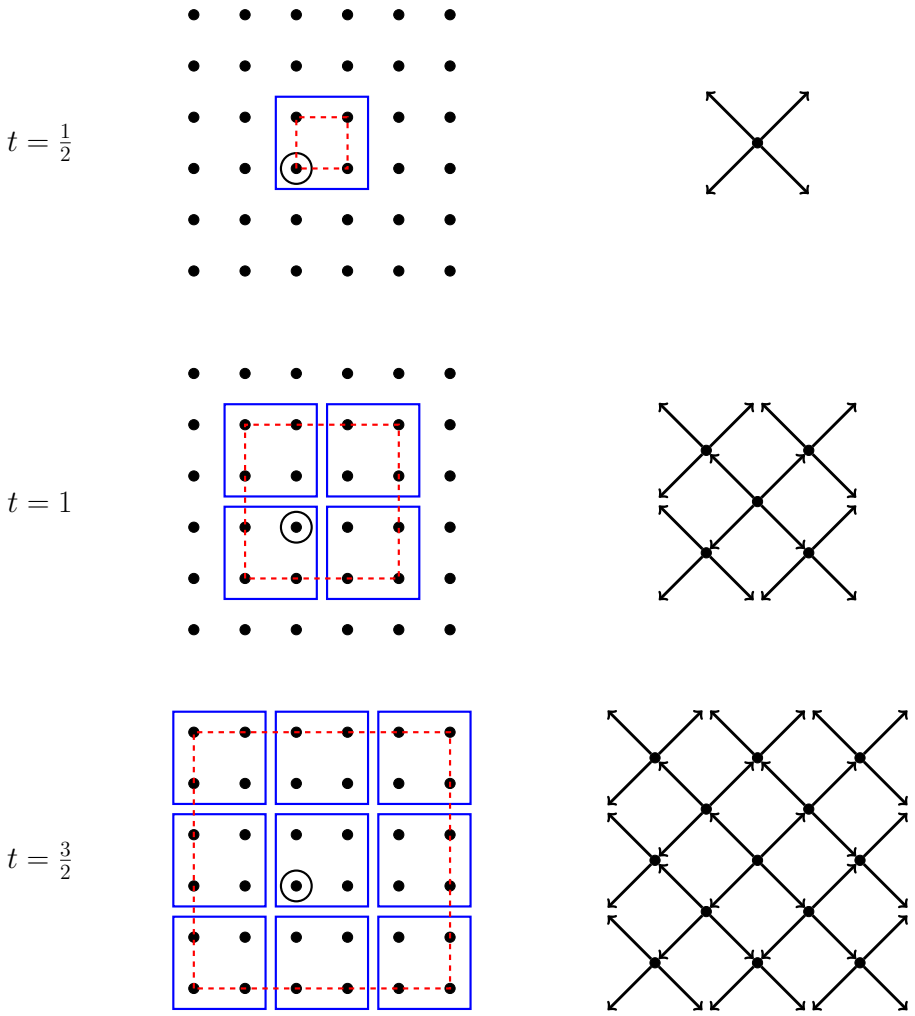

Figure 3.6: This figure illustrates how the light-cone boundary grows up to $t = 3/2$ for the model described in equation 3.23. The images in the column on the left show the quantum circuit at each time-step, with black dots representing qubits, where the qubit highlighted with a black circle is the qubit upon which the initially local operator is non-identity, and the blue rectangles represent the random Clifford unitaries. The red dashed line represents the rectangular region corresponding to the light-cone; qubits on the edges of the rectangle are on the boundary of the light-cone. The images in the column on the right show the directed graph corresponding to the light-cone boundary. Arrows in the directed graph represent qubits on the boundary of the light cone at a particular time, where as the vertices represent the random Clifford unitaries which act on four qubits and cause the growth of the initially local operator.

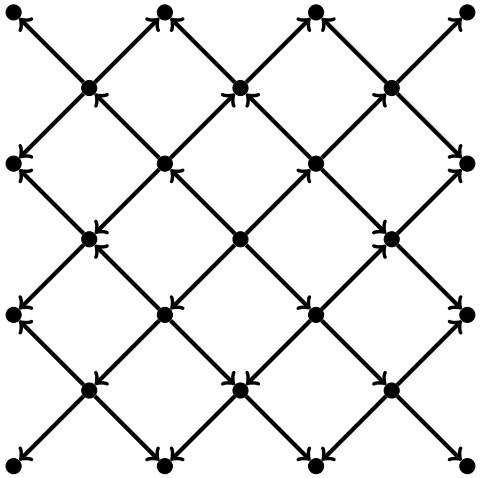

Figure 3.7: This figure illustrates the directed graph which corresponds to the boundary of the causal light-cone up to $t = 3/2$ for the model shown in figure 3.5 and described in equation 3.23.

these outwards arrows themselves then point to another vertex, which corresponds to a Clifford unitary acting on the light-cone boundary at the next half-time step. The growth of the light-cone boundary, and its interpretation as a directed graph, is illustrated in figure 3.6.

In summary, we can study the growth of an initially local operator and investigate if it is non-identity on the light-cone boundary at any time by using a directed graph. Translating the problem in this manner to studying a directed graph is in itself noteworthy, and I shall now explain the benefit of studying the problem in this way. I should reiterate that we are only interested in whether an operator is non-identity on a qubit situated on the boundary of the light-cone or not. Consequently, this can be represented by the arrows from a vertex being present, indicating the operator has support, or absent, indicating that the operator is not supported on this site (qubit). Since each vertex is a random Clifford unitary, the outwards arrows have a joint probability distribution. Hence, if there exists from the origin a directed path along the arrows to the square boundary centred on the origin, representing the light-cone boundary at a certain time, then we can conclude that the operator is supported on the light-cone boundary at this time. This event has some probability, and I will demonstrate that for any time t this probability is (lower-bounded by) a non-zero

constant. In other words, using the random directed graph I will prove a lower bound for the probability

$$\text{prob}\{\text{light-speed operator growth at time } t\} = \text{prob}\{G \text{ has } \omega_{2t}\}\ , \tag{3.24}$$

where ω_{2t} is a path from the origin along the arrows in the directed graph of length $2t$, and where G denotes the random directed graph representing the light-cone boundary that we have discussed and shown examples of in figures 3.6and 3.7. I would like to emphasise that G is a random directed graph, in which some of the arrows are absent and others are present according to a probability distribution, the form of which will be proven in lemma 3.5, and hence is one particular instance of G_{rand}, the set of all directed graphs representing the light-cone boundary. So for example, the directed graph shown in figure 3.7 is one instance of the random directed graph G (of G_{rand}), in which all arrows are present.

In the rest of this section, I will prove a series of rigorous results which lower bound prob{light-speed operator growth at time t} and hence give the statement in result 3.3.

There is one minor point to consider. Firstly, if a vertex has no inwards arrows, which corresponds to the evolved operator having no support on the sites which a random Clifford unitary acts, then strictly speaking it should not have any outwards arrows also. However, since I am considering a directed paths from the origin, this is not a problem. That is to say, if there are no inwards arrows to a vertex, the presence or absence of outwards arrows is irrelevant since a directed path from the origin using these erroneous arrows could not be made anyway. Therefore, in the directed graph picture, when we are only interested in paths from the origin, we can consider the presence of outwards arrows from a vertex independently of whether inwards arrows are present or not.

In the next lemma, I will give the probability distribution for the presence/absence of outwards arrows for the three types of vertex in this directed graph.

Lemma 3.5. For the vertex with four outwards arrows, the probability distribution

for the outwards arrows is given by

$$P(x_1,x_2,x_3,x_4) = \begin{cases} 0 \text{ if } x_1 = x_2 = x_3 = x_4 = 0, \\ \frac{3^{x_1+x_2+x_3+x_4}}{255} \text{ otherwise }, \end{cases} \tag{3.25}$$

where $x_i = 0$ indicates that the i^{th} arrow is absent, and $x_i = 1$ indicates that the i^{th} arrow is present. For the vertices with three outwards arrows, the probability distribution for the outwards arrows is given by

$$P(x_1,x_2,x_3) = \frac{1}{255} \begin{cases} 3 \text{ if } x_1 = x_2 = x_3 = 0, \\ 4 \times 3^{x_1+x_2+x_3} \text{ otherwise }, \end{cases} \tag{3.26}$$

where x_i are defined in the same manner. For the vertices with two outwards arrows, the probability distribution for the outwards arrows is given by

$$P(x_1,x_2) = \frac{1}{255} \begin{cases} 15 \text{ if } x_1 = x_2 = 0, \\ 2^4 \times 3^{x_1+x_2} \text{ if otherwise }, \end{cases} \tag{3.27}$$

where x_i are again defined in the same manner.

Proof. We recall that presence of an arrow represents that the operator is supported on the qubit site, and that each vertex is an independent and uniformly distributed random Clifford unitary. We will first look at the case of a vertex with four outwards arrows. In essence, we are interested in the probability distribution of the transformation

$$U^\dagger (P \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1}) U \,, \tag{3.28}$$

where U is a (uniformly distributed) random Clifford unitary. Since the random Clifford unitary is uniformly distributed, then the transformed Pauli operator is uniformly distributed over all (non-identity) Pauli operators on four qubits. Immediately, we can see from this that the probability that all arrows are absent, so the operator on all four qubit sites is the identity, is zero. For each present arrow, the operator on the qubit site is one of the three non-identity Pauli operators $(\sigma_X, \sigma_Y, \sigma_Z)$. Hence, since the

distribution is uniform over all non-identity Pauli operators, for each present arrow we must include a factor of 3, and dividing by the total number of non-identity Pauli operators on four qubits, $4^4 - 1 = 255$, we get the stated probability distribution.

Using the distribution for four outwards arrows, we can find the probability distribution for three and two outwards arrows as marginal distributions, in effect summing over a single or two outwards arrows. So, to obtain the distribution of three outwards arrows, we sum over the fourth arrow (x_4), and to obtain the distribution for two outwards arrows, we sum over the third and forth arrow (x_3, x_4). Hence, we get all three stated probability distributions. □

I will now use this directed graph picture to lower bound the probability

$$\text{prob}\{\text{light-speed operator growth at time } t\} = \text{prob}\{G \text{ has } \omega_{2t}\} \,, \tag{3.29}$$

where, to reiterate, G is a random directed graph, in which some of the arrows are absent and others are present according to the probability distribution proven in lemma 3.5. To do this, I will introduce the following definition, which concerns paths in the directed graph.

Definition 3.6. In the (random) directed graph G, an l-path is a path starting at the origin along l arrows, and so of total length l. (In figure 3.8 there is an illustration of an example.)

Using these definitions, we can rewrite prob{light-speed operator growth at time t}, in the following way

$$\begin{aligned} &\text{prob}\{\text{light-speed operator growth at time } t\} \\ &= 1 - \text{prob}\{G \text{ has no } 2l - \text{path}\} \,. \end{aligned} \tag{3.30}$$

Hence, in order to establish a lower-bound to prob{light-speed operator growth at time t}, I will now investigate the upper-bound of prob{G has no $2l-$path}.

In the next lemma, I will demonstrate that one can upper bound the probability prob{G has no $2l-$path} by considering one quadrant of the directed graph, which

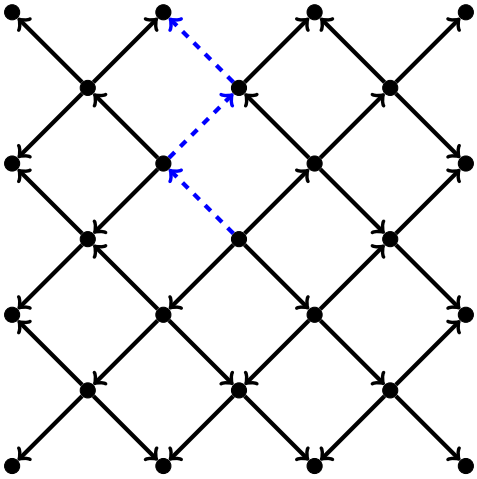

Figure 3.8: This figure illustrates an example of an l-path, a description of which is given in definition 3.6. The directed graph corresponding to the outer boundary of the causal light-cone up to $t = 3/2$ is shown in the figure, with one example of an l-path, with $l = 3$, highlighted with blued dashed arrows.

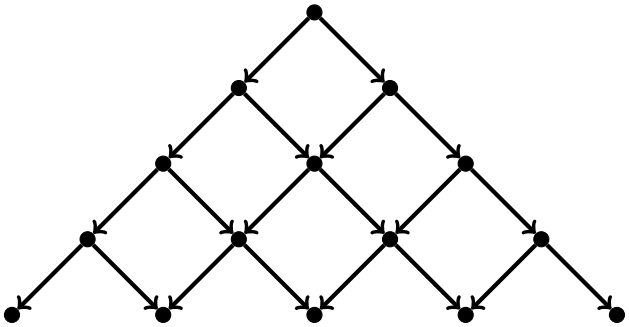

Figure 3.9: This figure illustrates one quadrant of the directed graph, which is shown in figure 3.7.

is illustrated in figure 3.9.

Lemma 3.7. The probability that in the directed graph G there is no l-path is upper bounded by

$$\text{prob}\{G \text{ has no } l-\text{path}\} \leq \text{prob}\{G^{\triangle} \text{ has no } l-\text{path}\}\ , \tag{3.31}$$

where $G^{\triangle}$ is one quadrant of the (random) directed graph G, an illustration of which is given in figure 3.9

Proof. The proof of this lemma is quite direct. First, we note that for the random directed graph G, we can upper-bound prob$\{G$ has no $l-$path$\}$ by neglecting l-paths which terminate at any choice of vertex, or set of vertices. Hence, we can upper-bound prob$\{G$ has no $l-$path$\}$ by considering only one quadrant of the (random) directed graph, $G^{\triangle}$. This single quadrant of the directed graph (figure 3.7) is illustrated in figure 3.9. □

One might wonder: why just study one quadrant of the directed graph? As I have already mentioned, the probability distribution for arrows emerging from the same vertex, as given in lemma 3.5, is not an independent distribution. In the following lemma, I will show that by studying one quadrant of the directed graph one can upper-bound further and replace the probability distribution for arrows, as given in lemma 3.5, by an independent and identical distribution.

Lemma 3.8. The probability that for one quadrant of the directed graph, $G^{\triangle}$, there is no l-path can be upper-bounded by

$$\text{prob}_p\{G^{\triangle} \text{ has no } l-\text{path}\} \leq \text{prob}_{p'}\{G^{\triangle} \text{ has no } l-\text{path}\} \ , \tag{3.32}$$

where $\text{prob}_p\{\cdot\}$ indicates that the probability distribution for the two outwards arrows from each vertices in the quadrant is given by equation 3.27 (lemma 3.5), and where $\text{prob}_{p'}\{\cdot\}$ indicates that the probability distribution for each outwards arrow is independent and given by

$$P(x_i) = \begin{cases} \frac{21}{85} + \varepsilon \text{ for } x_i = 0, \\ \frac{64}{85} - \varepsilon \text{ for } x_i = 1 \ , \end{cases} \ , \tag{3.33}$$

where x_i indicates if the i^{th} arrow is absent ($x_i = 0$) or present ($x_i = 1$), and $\varepsilon = \frac{1}{170}(43 - \sqrt{1785}) \approx 0.00441612$.

Proof. Firstly, we note that the probability distribution of outwards arrows from different vertices are independent, and hence we focus on only a single vertex with two outwards arrows. We can rewrite the probability distribution for the vertices

with two outwards arrows, given in equation 3.27 (lemma 3.5), in a matrix form as

$$\begin{pmatrix} \text{prob}\{0,0\} & \text{prob}\{0,1\} \\ \text{prob}\{1,0\} & \text{prob}\{1,1\} \end{pmatrix} = \begin{pmatrix} \frac{1}{17} & \frac{16}{85} \\ \frac{16}{85} & \frac{48}{85} \end{pmatrix} . \tag{3.34}$$

We can modify the distribution to make it more less likely that a path exists from the origin, by making more probable that the two outwards arrows are absent and less probable that the two outwards arrows are present, so

$$\begin{pmatrix} \text{prob}\{0,0\}+\varepsilon & \text{prob}\{0,1\} \\ \text{prob}\{1,0\} & \text{prob}\{1,1\}-\varepsilon \end{pmatrix} , \tag{3.35}$$

where $\varepsilon \geq 0$, and where we note that this modification preserves normalisation. Hence, we get an upper bound to the probability

$$\text{prob}_p\{G^{\triangle} \text{ has no } l-\text{path}\} . \tag{3.36}$$

Since we are interested in this distribution being an independent distribution, which is equivalent to selecting an ε such that

$$\begin{vmatrix} \text{prob}\{0,0\}+\varepsilon & \text{prob}\{0,1\} \\ \text{prob}\{1,0\} & \text{prob}\{1,1\}-\varepsilon \end{vmatrix} = 0 . \tag{3.37}$$

This relation can be solved to find that $\varepsilon = \frac{1}{170}(43-\sqrt{1785}) \leq 0.00441612$. Hence taking marginals of this new independent distribution we get the stated result. □

Finally, the following lemma establishes an upper bound to the probability $\text{prob}\{G^{\triangle} \text{ has no } 2l-\text{path}\}$ and hence provides the lower bound which justifies the statement given in result 3.3 .

Lemma 3.9. The probability that for one quadrant of the directed graph, $G^{\triangle}$, there is no l-path is be upper-bounded by

$$\text{prob}\{G^{\triangle} \text{ has no } l-\text{path}\} \leq 0.6 . \tag{3.38}$$

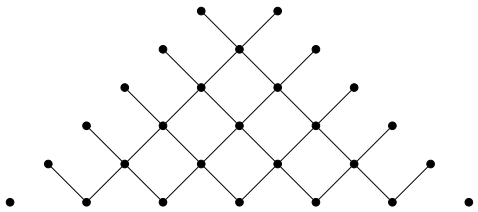

Figure 3.10: This figure illustrates the dual of a single quadrant of the directed graph shown in figure 3.9. It is worth noting that the dual graph is not directed, since the edges in the dual correspond to absence of arrows in the directed graph.

Proof. Firstly, we note the following inequality

$$\text{prob}\{G^{\triangle} \text{ has no } l-\text{path}\} \leq \text{prob}\{G^{\triangle} \text{ has no } (l+1)-\text{path}\} \ , \tag{3.39}$$

in other words the probability is increasing with increasing path length.

The proof of this result makes use of the notion of the dual graph of the directed graph, $G^{\star}$, which I will now introduce. The vertices of the dual graph corresponding to the directed graph quadrant ($G^{\triangle}$) are located within the faces of the squares in the directed graph. The edges in the dual graph are not directed. The presence of an edge in the dual graph corresponds to the absence of an arrow it intersects in the regular directed graph. In other words, an edge in the dual graph severs the corresponding arrow it crosses in the directed graph. Therefore, the edges in the dual and the arrows in the regular graph are related and complementary, so if an arrow in the directed graph is present with probability p then the corresponding edge in the dual is absent with probability $1-p$. Figure 3.10 illustrates an example of a dual graph of the directed graph quadrant shown in figure 3.9.

In analogy with the definition of l-paths (definition 3.6), a d-wall in the dual graph $G^{\star}$ is a set of d edges (and hence has length d) in the dual graph which creates a path connecting the left side of the quadrant with the right side of the quadrant. The presence of a d-wall in the dual graph means that in the directed graph $G^{\triangle}$ there is a boundary which no outwards arrows cross, and hence also a boundary which no l-paths cross.

Therefore, using the union bound and the upper-bound given in lemma 3.8 it is

d	$N(d-\text{walls})$
2	1
3	2
4	3
5	6
6	18

Table 3.2: This table shows the exact number of possible d-walls in the dual graph, $G^\star$, for $d = 2,3,4,5,6$.

found that

$$\text{prob}\{G^{\triangle} \text{ has no } l-\text{path}\} \leq \sum_{d=2}^{\infty} N(d-\text{walls}) \left(\frac{21}{85}+\varepsilon\right)^d , \tag{3.40}$$

where $N(d-\text{walls})$ denotes the total number of d-walls.

We can be upper-bounded the number of d-walls, $N(d-\text{walls})$, by

$$N(d-\text{walls}) \leq \left(3^{d-2}(d-3)+2\right) \tag{3.41}$$

However, for shorter paths ($d = 2,3,4,5,6$) we are able to count the number of possible d-walls, $N(d-\text{walls})$, exactly - the values for which are shown in table 3.2. A better upper-bound is obtained by incorporating these terms exactly and using the upper bound for $N(d-\text{walls})$ for $d > 6$. The upper bound for $N(d-\text{walls})$ in equation 3.41 can be obtained in the following way.

This upper-bound can be seen in the following way, which is most clear with reference to the illustration in figure 3.10. Consider a d-wall starting at a specific vertex position on the left side of the dual graph $G^\star$. For the choice of the first edge in this path there is only one possible choice, for the next $d-2$ there are at most 3 possible choices, and for the final edge in the path there is a single choice. Hence, we can upper-bound the number of d-walls starting at a specific choice of vertex on the left side of the dual graph $G^\star$ with 3^{d-2}. It is worth noting that this upper-bound includes paths which do not connect the left and right sides of the quadrant of the dual graph, and hence do not actually form d-walls. Now, for a d-wall the total number of vertices which can be the initial specific vertex on the left side of the

quadrant is $d-1$. Hence, we can upper-bound $N(d-\text{walls})$ by the product of the total number of starting positions of the d-wall with the upper-bound for the total number of d-walls starting at a specific vertex. Moreover, for any d-wall when the starting vertex on the left side is either the 1^{st} or $(d-1)^{\text{th}}$ from the top, then there is only one possible choice of d-wall. So, to incorporate this correction, we can change the factor accounting for the total number of starting vertices from $d-1$ to $d-3$ and instead add $+2$, resulting in the stated upper-bound.

The stated result follows from evaluating this sum. □

3.2.3 Discussion

To summarise, I have shown that for the time-periodic quantum circuit model considered there is a non-zero constant minimum probability of an initially local Pauli operator being supported on the boundary of its light-cone at all times. In this sense, there is an absence of localisation in this models. This result is notable since in the previous section 3.1 the one-dimensional version of the models considered displayed strong localisation. Also, I expect that this result holds more generally for any non-fine-tuned initial operator.

It is worth noting a few points. Firstly, since the definition of absence of localisation used only concerns the boundary of light-cone, as do all of the mathematical derivations, then all of the results also apply to a time-dependent version of the model analysed. However, this is not the case for the interior of the light-cone, which has not been studied. Indeed, it would be interesting to study the interior of the light-cone, but this would not be amenable using the same mathematical tools. This is because the proofs rely on the fact that when studying only the light-cone boundary all of the random Clifford unitaries are independent, whereas if we consider the interior of the light-cone this is not the case, since the dynamics are time-periodic. Additionally, in the model we have considered there is only a single qubit situated on each lattice site, but one could also have considered the case where the number of qubits per lattice site is larger, $N > 1$. When the number of qubits per site is N the Hilbert space of the $L \times L$ square lattice has dimension 2^{NL^2} and the local random unitaries $U_{(x,y)}$ are selected independently from the uniform distribution over the

Clifford group of $4N$ qubits, $\mathcal{C}_{4N}$. In this case, the same mathematical approach could be used, and I expect that the lower bound for the probability is larger; the probability that an operator localises is smaller.

In the model I have considered the random unitaries are sampled uniformly from the Clifford group, which is a sub-group of the unitary group. It would be interesting to consider instead the case where the random circuit is composed of unitaries sampled from the unitary group. In the case of time-dependent circuits this has been well studied in references such as [53, 59, 135], and discussed in the introduction of this thesis (chapter 1). Whereas, as mentioned in chapter 1, the case of time-periodic quantum circuits in one spatial dimension has been studied in references such as [22, 26, 27, 125].

Finally, I would like to comment on connections between the mathematical approach in this section and directed percolation theory [136]. Both cases analyse the presence of infinitely long paths which start at the origin in random directed graphs. But in the case of the model in this section the arrows that emerge from the same vertex are not statistically independently, while in directed percolation theory they are. In this sense, the results here study a variant of (directed) percolation theory.

Chapter 4

General Conclusions

In summary, in this thesis I have analysed time-periodic quantum circuits and proven results concerning mixing and localisation. Initially, in the first chapter, I discussed a selection of topics which are relevant to the results of this thesis. In the first section of the second chapter, I analysed a time-periodic quantum circuit with local interactions given by random Clifford unitaries, and demonstrated with mathematical proofs that the dynamics can, under certain conditions, display (approximately) the mixing of Pauli operators and appear to resemble a (Haar) random unitary. I have stated precisely the conditions under which these results hold - a large ($N \gg \log L$) local spin dimension and at times in the region of the scrambling time (t_{scr}). Then, motivated by these results, in the second section I analysed an analogous version of the same model in which the interactions are no longer local, but remain two-body. I demonstrated that the time-periodic random quantum circuit has a poly-logarithmic ($\sim (\log L)^2$) circuit depth corresponding to the scrambling time. For now, I have left the important question of whether or how closely the dynamics of this model appear to resemble a random unitary, or exhibit Pauli mixing, unanswered. This is an important avenue for further work, and one which is of great importance for the study of highly chaotic quantum systems, and in particular the fast scrambling conjecture. The results in both of these sections are markedly different from the more studied case of time-dependent random quantum circuits.

In the third chapter I presented a series of results concerning localisation. In the first section, I analysed the same time-periodic quantum circuit with local

interactions given by Clifford unitaries as in the second chapter. I demonstrated that in the opposite case of a small local spin dimension, so one or a few qubits per site of the spin chain, the system exhibits a novel form of localisation. I characterised this form of localisation, in which initially local operators are confined to fixed finite regions of the spin chain for all time. Then, motivated by these results, in the second section I presented (two) analogous time-periodic quantum circuits in two spatial dimensions and investigated the growth of initially local (Pauli) operators, in order to gain an understanding of localisation. Remarkably, this problem can be recast in terms of a directed graph, and the question of localisation studied as a form of percolation. Hence, using these tools, I gave rigorous proofs that the system displays an absence of localisation. The meaning of this was stated precisely; with some minimum constant probability, independent of the system-size, an initially local operator grows at the maximum (linear) rate in some directions. Moreover, since these results are independent of the system size, they hold in the thermodynamic limit, $L \to \infty$.

The dynamics of all of the models analysed in this thesis arise from interactions given by Clifford unitaries. It is worthwhile summarising now some interesting features of Clifford dynamics, which demonstrate that they are a worthwhile lens with which to study many-body quantum systems. One major tool I have made use of throughout this thesis is the phase space description of the dynamics. Clifford unitaries can be represented as symplectic transformations in phase space, which crucially has an exponentially smaller dimension than the Hilbert space. One major consequence of this fact is that it is efficient to simulate the evolution (under Clifford dynamics) of Pauli operators on a classical computer [122]. Equally however, due to the symplectic matrices having entries in $\mathbb{Z}_2$, there are no eigen-modes and one cannot use many tools from quasi-free bosons/fermions. Another feature is that Clifford dynamics with disorder can, as I have discussed, display a strong form of localisation. This strong form of localisation is both reminiscent of Anderson localisation, but also not understandable in terms free particles. Indeed, typical Clifford dynamics cannot be understood in terms of free or interacting particles. Although some specific

cases of Clifford dynamics do have gliders, which are understood as discrete time analogues of free particles, in the case of translation-invariant Clifford dynamics the time evolution of operators have fractal patterns [127, 137]. In fact, there is an even more extreme case of fully non-local integrals of motion [138]; operators which commute with the time evolution operator act on an extensive number of sites of the system. However, on the other hand, we have seen that Clifford dynamics do display signatures of chaos also. Indeed, as I mentioned in the introduction (chapter 1 section 1.2), Clifford dynamics can (in the non-local case) generate (at most) a unitary 3-design.

Finally, I will now comment further on the scope for future work. Of course, one question that is still unanswered is: can local and time-periodic or time-independent dynamics generate a unitary 2-design? In terms of the result I have presented in this thesis, this would mean removing the restriction of the measurements in the discrimination process. This is of particular importance to the study of high chaotic quantum systems, in which the dynamics are modelled with a (Haar) random unitary. Since one would expect that the dynamics of physical systems are time-independent, the results in this thesis (chapter 2 section 2.1) have provided some justification for the use of (Haar) random unitaries as a model for chaotic quantum systems. On the localisation side, one important question that is worth investigating further is whether the strong localisation in one spatial dimension I have analysed (chapter 3 section 3.1) is robust against perturbations. Moreover, it is worth considering the results in this thesis from the perspective that it is now possible to implement quantum circuits on small quantum devices. One notable example is the recent experiment in reference [139], which investigates a quantum many-body system with Floquet dynamics. Both the mixing results in chapter 2 section 2.1 and the localisation results of chapter 3 section 3.1 could be suitable for study on current quantum devices. In particular, investigating the question of strong localisation and whether it is robust to perturbations seems suitable, since it can be investigated using only a short quantum circuit on a small number of qubits, and would provide insight into the landscape of many-body phenomena.

Appendix A

Clifford dynamics and discrete phase space

In this section, we first define the Pauli and Clifford groups and then present the phase-space description of Clifford dynamics. This description is known from previous works [121–123] and we include it here for completeness and clarity.

The Pauli sigma matrices together with the identity $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ form a basis of the space of operators of one qubit $\mathbb{C}^2$. Also, the sixteen matrices obtained by multiplying $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ times the coefficients $\{1, \mathrm{i}, -1, -\mathrm{i}\}$ form a group. This is called the Pauli group of one qubit and it is denoted by $\mathcal{P}_1$. The generalization to n qubits is the following.

Definition A.1. The **Pauli group** of n qubits $\mathcal{P}_n$ is the set of matrices $\mathrm{i}^u \sigma_{\mathbf{u}}$ where

$$\sigma_{\mathbf{u}} = \bigotimes_{i=1}^{n} (\sigma_x^{q_i} \sigma_z^{p_i}) \in \mathrm{U}(2^n) \ , \tag{A.1}$$

for all phases $u \in \mathbb{Z}_4$ and vectors $\mathbf{u} = (q_1, p_1, q_2, p_2, \ldots, q_n, p_n) \in \mathbb{Z}_2^{2n}$. We also define $\bar{\mathcal{P}}_n = \mathcal{P}_n / \{1, \mathrm{i}, -1, -\mathrm{i}\}$ which satisfies $\bar{\mathcal{P}}_n \cong \mathbb{Z}_2^{2n}$.

Here $\mathbb{Z}_2^{2n}$ stands for a $2n$-dimensional vector space with addition and multiplication operations defined modulo 2. Using the identity $\sigma_z \sigma_x = -\sigma_x \sigma_z$ and the definition

$\beta(\mathbf{u},\mathbf{u}') = \sum_{i=1}^{n} p_i q'_i$ we obtain the multiplication and inverse rules

$$\sigma_{\mathbf{u}}\sigma_{\mathbf{u}'} = (-1)^{\beta(\mathbf{u},\mathbf{u}')}\sigma_{\mathbf{u}+\mathbf{u}'} \ , \tag{A.2}$$

$$\sigma_{\mathbf{u}}^{-1} = (-1)^{\beta(\mathbf{u},\mathbf{u})}\sigma_{\mathbf{u}} \ . \tag{A.3}$$

The Pauli group A.1 is the discrete version of the Weyl group, or the *displacement operators* used in quantum optics. Concretely, if $\hat{Q}$ and $\hat{P}$ are quadrature operators (satisfying the canonical commutation relations $[\hat{Q},\hat{P}] = \mathrm{i}$) then we can write the analogy as

$$\sigma_x^q \sigma_z^p \ \longleftrightarrow \ e^{\mathrm{i}\hat{P}q} e^{\mathrm{i}\hat{Q}p} \ , \tag{A.4}$$

where the phase space variables (q,p) take values in $\mathbb{Z}_2^2$ on the left of A.4, and in $\mathbb{R}^2$ on the right. This analogy also extends to the set of transformations that preserve the phase space structure. Before characterizing these transformations let us define the phase space associated to the Pauli group.

Definition A.2. The **discrete phase space** of n qubits $\mathbb{Z}_2^{2n}$ is the $2n$-dimensional vector space over the field $\mathbb{Z}_2$, endowed with the symplectic (antisymmetric) bilinear form

$$\langle \mathbf{u},\mathbf{u}' \rangle = \mathbf{u}^T J \mathbf{u}' \ , \text{ where } \ J = \bigoplus_{i=1}^{n} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \tag{A.5}$$

for all $\mathbf{u},\mathbf{u}' \in \mathbb{Z}_2^{2n}$. Note that the form is indeed antisymmetric $\langle \mathbf{u},\mathbf{u}' \rangle = \langle \mathbf{u}',\mathbf{u} \rangle = -\langle \mathbf{u}',\mathbf{u} \rangle$ mod 2, which implies $\langle \mathbf{u},\mathbf{u} \rangle = 0$.

Using the symplectic form A.5 and the rules (A.2-A.3) we can write the commutation relations of the Pauli group as

$$\sigma_{\mathbf{u}}\,\sigma_{\mathbf{u}'}\,\sigma_{\mathbf{u}}^{-1}\,\sigma_{\mathbf{u}'}^{-1} = (-1)^{\langle \mathbf{u},\mathbf{u}' \rangle} \ . \tag{A.6}$$

In analogy with the continuous (bosonic) phase space, in the following two definitions we introduce the transformations that preserve the symplectic form A.5 and the Pauli group, respectively.

Definition A.3. The **s**ymplectic group $\mathcal{S}_n$ is the set of matrices $S : \mathbb{Z}_2^{2n} \to \mathbb{Z}_2^{2n}$ such that

$$\langle S\mathbf{u}, S\mathbf{u}' \rangle = \langle \mathbf{u}, \mathbf{u}' \rangle \ , \tag{A.7}$$

for all $\mathbf{u}, \mathbf{u}' \in \mathbb{Z}_2^{2n}$. This is equivalent to the condition $S^T J S = J \bmod 2$.

Definition A.4. The **Clifford group** of n qubits $\mathcal{C}_n$ is the subset of unitaries $U \in \mathrm{U}(2^n)$ which map the Pauli group onto itself

$$U \sigma_{\mathbf{u}} U^\dagger \in \mathcal{P}_n \ \text{ for all } \mathbf{u}. \tag{A.8}$$

Since adding a global phase $e^{\mathrm{i}\theta} U$ does not change the map A.8, we identify all unitaries $\{e^{\mathrm{i}\theta} U : \forall \theta \in \mathbb{R}\}$ with the same element of $\mathcal{C}_n$. In other words, $\mathcal{C}_n$ is the quotient of the centralizer of $\mathcal{P}_n$ by the group $\mathrm{U}(1)$.

Lemma A.5. [Structure of $\mathcal{C}_n$] Each Clifford transformation $U \in \mathcal{C}_n$ is characterized by a symplectic matrix $S \in \mathcal{S}_n$ and a vector $\mathbf{s} \in \mathbb{Z}_2^{2n}$ so that

$$U \sigma_{\mathbf{u}} U^\dagger = \mathrm{i}^{\alpha[S,\mathbf{u}]} (-1)^{\langle \mathbf{s}, \mathbf{u} \rangle} \sigma_{S\mathbf{u}} \ , \tag{A.9}$$

where the function α takes values in $\mathbb{Z}_4$. More precisely we have $\mathcal{C}_n \cong \bar{\mathcal{P}}_n \ltimes \mathcal{S}_n$.

Proof. In this work the function α does not play any role, hence, we do not provide a characterization. Moving to the proof, for each $U \in \mathcal{C}_n$ there are two functions

$$s : \mathbb{Z}_2^{2n} \to \mathbb{Z}_4 \ , \tag{A.10}$$

$$S : \mathbb{Z}_2^{2n} \to \mathbb{Z}_2^{2n} \ , \tag{A.11}$$

such that

$$U \sigma_{\mathbf{u}} U^\dagger = \mathrm{i}^{s[\mathbf{u}]} \, \sigma_{S[\mathbf{u}]} \ . \tag{A.12}$$

Note that, at this point, we do not make any assumption about these functions, such

as linearity. Using A.2 we obtain the equality between the following two expressions

$$\begin{aligned} U\sigma_{\mathbf{u}}\sigma_{\mathbf{u}'}U^{\dagger} &= (-1)^{\beta(\mathbf{u},\mathbf{u}')}\, U\sigma_{\mathbf{u}+\mathbf{u}'}U^{\dagger} \\ &= (-1)^{\beta(\mathbf{u},\mathbf{u}')}\, \mathrm{i}^{s[\mathbf{u}+\mathbf{u}']}\, \sigma_{S[\mathbf{u}+\mathbf{u}']}\ , \qquad \text{(A.13)} \\ U\sigma_{\mathbf{u}}U^{\dagger}U\sigma_{\mathbf{u}'}U^{\dagger} &= (\mathrm{i}^{s[\mathbf{u}]}\,\sigma_{S[\mathbf{u}]})\,(\mathrm{i}^{s[\mathbf{u}']}\,\sigma_{S[\mathbf{u}']}) \\ &= (-1)^{\beta(S\mathbf{u},S\mathbf{u}')}\, \mathrm{i}^{s[\mathbf{u}]+s[\mathbf{u}']}\, \sigma_{S[\mathbf{u}]+S[\mathbf{u}']}\ , \qquad \text{(A.14)} \end{aligned}$$

which implies the $\mathbb{Z}_2$-linearity of the S function. Hence, from now on, we write its action as a matrix $S[\mathbf{u}] = S\mathbf{u}$. Next, if we impose the commutation relations of the Pauli group A.6 as follows

$$\begin{aligned} (-1)^{\langle \mathbf{u},\mathbf{u}'\rangle} &= U\sigma_{\mathbf{u}}\,\sigma_{\mathbf{u}'}\,\sigma_{\mathbf{u}}^{-1}\,\sigma_{\mathbf{u}'}^{-1}\,U^{-1} \\ &= (\mathrm{i}^{s[\mathbf{u}]}\sigma_{S\mathbf{u}})\,(\mathrm{i}^{s[\mathbf{u}']}\sigma_{S\mathbf{u}'})\,(\mathrm{i}^{s[\mathbf{u}]}\sigma_{S\mathbf{u}})^{-1}\,(\mathrm{i}^{s[\mathbf{u}']}\sigma_{S\mathbf{u}'})^{-1} \\ &= (-1)^{\langle S\mathbf{u},S\mathbf{u}'\rangle}\ , \qquad \text{(A.15)} \end{aligned}$$

we find that the matrices S are symplectic. Conversely, it has been proven [121, 140–142] that for each symplectic matrix $S \in \mathcal{S}_n$ there is $U \in \mathcal{C}_n$ such that $U\sigma_{\mathbf{u}}U^{\dagger} \propto \sigma_{S\mathbf{u}}$ for all $\mathbf{u}$.

Now, let us obtain the set of pairs (S,s) associated to the subgroup $\bar{\mathcal{P}}_n \subseteq \mathcal{C}_n$. Using A.6 we see that the Clifford transformation $\sigma_{\mathbf{v}} \in \bar{\mathcal{P}}_n$ has $S = \mathbb{1}$ and $s[\mathbf{u}] = 2\langle \mathbf{v},\mathbf{u}\rangle$, for any $\mathbf{v} \in \mathbb{Z}_2^{2n}$. Next, let us prove the converse. By equating A.13 and A.14 with $S = \mathbb{1}$, we see that any Clifford transformation U with $S = \mathbb{1}$ has a phase function s satisfying

$$s[\mathbf{u}+\mathbf{u}'] = s[\mathbf{u}] + s[\mathbf{u}']\ , \qquad \text{(A.16)}$$

for all pairs $\mathbf{u},\mathbf{u}'$. Also, since the map $\sigma_{\mathbf{u}} \to U\sigma_{\mathbf{u}}U^{\dagger}$ preserves the Hermiticity or anti-Hermiticity of $\sigma_{\mathbf{u}}$, the phase function in $U\sigma_{\mathbf{u}}U^{\dagger} = \mathrm{i}^{s[\mathbf{u}]}\sigma_{\mathbf{u}}$ has to satisfy $s[\mathbf{u}] \in \{0,2\}$ for all $\mathbf{u}$. Combining this with A.16 we deduce that, if $S = \mathbb{1}$ then $s[\mathbf{u}] = 2\langle \mathbf{v},\mathbf{u}\rangle$ for some vector $\mathbf{v} \in \mathbb{Z}_2^{2n}$. In summary, an element of the Clifford group belongs to the Pauli group if, and only if, there is a vector $\mathbf{v} \in \mathbb{Z}_2^{2n}$ such that $S = \mathbb{1}$ and $s[\mathbf{u}] = 2\langle \mathbf{v},\mathbf{u}\rangle$.

Now let us show that $\mathcal{C}_n/\bar{\mathcal{P}}_n \cong \mathcal{S}_n$. By definition, any Clifford element $U\bar{\mathcal{P}}_n U^\dagger \subseteq \bar{\mathcal{P}}_n$ satisfies $U\bar{\mathcal{P}}_n = \bar{\mathcal{P}}_n U$, hence $\bar{\mathcal{P}}_n \subseteq \mathcal{C}_n$ is a normal subgroup. This allows us to allocate each element $U \in \mathcal{C}_n$ into an equivalence class $U\bar{\mathcal{P}}_n \subseteq \mathcal{C}_n$, and define a group operation between classes. In order to prove the isomorphism $\mathcal{C}_n/\bar{\mathcal{P}}_n \cong \mathcal{S}_n$, we need to check that two transformations U, U' are in the same equivalence class ($\exists\, \mathbf{v} : U = U'\sigma_{\mathbf{v}}$) if and only if they have the same symplectic matrix $S = S'$. Identity A.6 tells us that $U = U'\sigma_{\mathbf{v}}$ implies $S = S'$. To prove the converse, let us assume that U, U' have symplectic matrices $S = S'$. Due to the fact U^{-1} has symplectic matrix S^{-1}, the product $U^{-1}U'$ has symplectic matrix $S^{-1}S = \mathbb{1}$. As proven above, this implies that $U^{-1}U' \in \bar{\mathcal{P}}_n$, and therefore both are in the same class.

Finally, for each symplectic matrix S we define $\alpha[S, \mathbf{u}] = s[\mathbf{u}]$ where s is the phase function of an arbitrarily chosen element in the equivalence class defined by S. The phase function of the other elements in the class S is $s[\mathbf{u}] = \alpha[S, \mathbf{u}] + 2\langle \mathbf{v}, \mathbf{u} \rangle$ for all $\mathbf{v} \in \mathbb{Z}_2^{2n}$. □

Appendix B

Miscellaneous mathematical results

In this section, we collate a variety of mathematical results which are used within the proofs written in the main body of the thesis. Some of these results are well-known, such as the Chernoff bound, others are minor and particular, such as a specific upper-bound for the binomial coefficient, but all are collected in this appendix so that the proofs written in the main body are understandable and at the same time concise. The mathematical results contained in this section which are well-known are provided without proof, since these are easily found, whereas the results which are more particular are given with proofs.

Result B.1. The **union bound** states that for a (countable) set of events $A_1, \ldots, A_n$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) \ . \tag{B.1}$$

Result B.2. The **Chernoff bounds** for a random variable X state that

$$\text{prob}\{X \geq a\} \leq \min_{t>0} e^{-ta}\, \mathbb{E}\left(e^{tX}\right) \ , \tag{B.2}$$

and

$$\text{prob}\{X \leq a\} \leq \min_{t>0} e^{ta}\, \mathbb{E}\left(e^{-tX}\right) \ . \tag{B.3}$$

Lemma B.1. The number of k-dimensional subspaces of $\mathbb{Z}_2^n$ is

$$\mathcal{N}_k^n = \prod_{i=0}^{k-1} \frac{2^n - 2^i}{2^k - 2^i} \ . \tag{B.4}$$

Proof. Let us start by counting how many lists of k linearly independent vectors $(\mathbf{u}_1,\ldots,\mathbf{u}_k)$ are in $\mathbb{Z}_2^n$. The first vector $\mathbf{u}_1$ can be any element of $\mathbb{Z}_2^n$ except the zero vector $\mathbf{0}$, giving a total of (2^n-1) possibilities. Following that, $\mathbf{u}_2$ can be any element of $\mathbb{Z}_2^n$ that is not contained in the subspace generated by $\mathbf{u}_1$, which is $\{\mathbf{0},\mathbf{u}_1\}$, giving (2^n-2) possibilities. Analogously, $\mathbf{u}_3$ can be any element of $\mathbb{Z}_2^n$ that is not contained in the subspace generated by $\{\mathbf{u}_1,\mathbf{u}_2\}$, which is $\{\mathbf{0},\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_1+\mathbf{u}_2\}$, giving (2^n-2^2) possibilities. Following in this fashion we arrive at the following conclusion. The number of lists of k linearly independent vectors is

$$\mathcal{L}_k^n = (2^n-2^0)(2^n-2^1)(2^n-2^2)\cdots(2^n-2^{k-1})\ . \tag{B.5}$$

It is important to note that many lists $(\mathbf{u}_1,\ldots,\mathbf{u}_k)$ generate the same subspace. So, in order to obtain $\mathcal{N}_k^n$, we have to divide $\mathcal{L}_k^n$ by the number of lists which generate that same subspace.

First, we note that a list $(\mathbf{u}_1,\ldots,\mathbf{u}_n)$ is a basis of $\mathbb{Z}_2^n$ with its vectors in a particular order. Hence, $\mathcal{L}_n^n$ is the number of basis (in particular order) of $\mathbb{Z}_2^n$. Second, we use the fact that the subspace of $\mathbb{Z}_2^n$ generated by the list $(\mathbf{u}_1,\ldots,\mathbf{u}_k)$ is isomorphic to $\mathbb{Z}_2^k$, so that, the number of basis (in a particular order) generating that subspace is $\mathcal{L}_k^k$. Putting things together, we obtain $\mathcal{N}_k^n = \mathcal{L}_k^n/\mathcal{L}_k^k$, as in equation B.4 □

Lemma B.2. Let $\mathcal{N}_k^n$ be the number of k-dimensional subspaces of $\mathbb{Z}_2^n$; then we have

$$2^{(n-k)k}(1-2^{k-n})^k \ \leq\ \mathcal{N}_k^n \ \leq\ 2^{(n-k)k}\min\{2^k,4\}\ . \tag{B.6}$$

Proof. Taking lemma B.1 and neglecting the negative terms in the numerator gives

$$\mathcal{N}_k^n = \prod_{i=0}^{k-1}\frac{2^n-2^i}{2^k-2^i} \leq \prod_{i=0}^{k-1}\frac{2^n}{2^k-2^i} \tag{B.7}$$

$$= \frac{2^{nk}}{2^{k^2}}\prod_{i=0}^{k-1}\frac{1}{1-2^{i-k}} = 2^{(n-k)k}\prod_{j=1}^{k}\frac{1}{1-2^{-j}} \tag{B.8}$$

$$\leq 2^{(n-k)k}\prod_{j=1}^{\infty}\frac{1}{1-2^{-j}}\ , \tag{B.9}$$

where in the last inequality we have extended the product to infinity. It turns out that this infinite product is the inverse of Euler's function ϕ evaluated at $1/2$, which has the value

$$\phi(1/2) = \prod_{j=1}^{\infty}(1-2^{-j}) \approx .28 \geq \frac{1}{4} \,. \tag{B.10}$$

Combining the two above inequalities we obtain

$$\mathcal{N}_k^n \leq 2^{(n-k)k}4 \,. \tag{B.11}$$

For the cases where $k = 0, 1$, we can improve this bound. When $k = 0$ the coefficient is 1 by definition, and when $k = 1$ the product $\prod_{i=0}^{k-1}(1-2^{i-k})^{-1}$ evaluates to 2. Hence, for $k = 0, 1$ we can replace 4 by 2^k, and therefore this improvement is captured concisely by changing 4 to $\min\{2^k, 4\}$.

We obtain the lower bound by instead neglecting the negative terms in the denominator

$$\mathcal{N}_k^n \geq \prod_{i=0}^{k-1}\frac{2^n - 2^i}{2^k} = \frac{2^{nk}}{2^{k^2}}\prod_{i=0}^{k-1}(1-2^{i-n}) \,. \tag{B.12}$$

The remaining product can be bounded using by

$$\prod_{i=0}^{k-1}(1-2^{i-n}) \geq \prod_{i=0}^{k-1}(1-2^{k-n}) \geq (1-2^{k-n})^k \,, \tag{B.13}$$

since $n \geq k > i$, and hence we get the final lower bound. □

Lemma B.3. The binomial coefficient can be bounded by

$$\binom{k+r-1}{k} \leq (1+r)^k \leq (2r)^k \,. \tag{B.14}$$

Proof. We start with the bound

$$\binom{k+r-1}{k} = \prod_{i=1}^{k}\frac{r+k-i}{i} \leq \prod_{i=1}^{k}1+\frac{r}{i} \,. \tag{B.15}$$

Using the AM-GM inequality, we find that

$$\binom{k+r-1}{k} \leq \prod_{i=1}^{k} 1+\frac{r}{i} \leq \left(1+\frac{rH_k}{k}\right)^k \leq (1+r)^k \leq (2r)^k \,. \tag{B.16}$$

□

Bibliography

[1] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbène. Quantum simulations with ultracold quantum gases. *Nature Physics*, 8(4):267–276, April 2012.

[2] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, October 2019.

[3] M. J. Giannoni (Editor), A. Voros (Editor), and J. Zinn-Justin (Editor). *Chaos and Quantum Physics: Proceedings of the Les Houches Summer School 1989*. Elsevier, 1991.

[4] Anthony J Short and Terence C Farrelly. Quantum equilibration in finite time. *New Journal of Physics*, 14(1):013063, January 2012.

[5] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, 452(7189):854–858, 2008.

[6] Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, 2016.

[7] J. M. Deutsch, Haibin Li, and Auditya Sharma. Microscopic origin of thermodynamic entropy in isolated systems. *Physical Review E*, 87(4):042135, April 2013.

[8] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papic. Weak ergodicity breaking from quantum many-body scars. *Nature Physics*, 14(7):745–749, 2018.

[9] Fabian H L Essler and Maurizio Fagotti. Quench dynamics and relaxation in isolated integrable quantum spin chains. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(6):064002, June 2016.

[10] Tomaz Prosen. Chaos and complexity of quantum motion. *Journal of Physics A: Mathematical and Theoretical*, 40(28):7881–7918, June 2007.

[11] Jean-Sébastien Caux and Jorn Mossel. Remarks on the notion of quantum integrability. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(02):P02023, February 2011.

[12] Jasen A. Scaramazza, B. Sriram Shastry, and Emil A. Yuzbashyan. Integrable matrix theory: Level statistics. *Physical Review E*, 94(3):032106, September 2016.

[13] Fritz Haake. *Quantum Signatures of Chaos*. Springer, 3rd edition, 2010.

[14] Freeman J. Dyson. Statistical Theory of the Energy Levels of Complex Systems. I. *Journal of Mathematical Physics*, 3(1):140–156, January 1962.

[15] Freeman J. Dyson. Statistical Theory of the Energy Levels of Complex Systems. III. *Journal of Mathematical Physics*, 3(1):166–175, January 1962.

[16] Freeman J. Dyson. Statistical Theory of the Energy Levels of Complex Systems. II. *Journal of Mathematical Physics*, 3(1):157–165, January 1962.

[17] M.L. Mehta. *Random Matrices*. ISSN. Elsevier Science, 2004.

[18] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. A bound on chaos. *Journal of High Energy Physics*, 2016(8):106, 2016.

[19] Yingfei Gu and Alexei Kitaev. On the relation between the magnitude and exponent of OTOCs. *Journal of High Energy Physics*, 2019(2):75, 2019.

[20] Amos Chan, Andrea De Luca, and J. T. Chalker. Solution of a minimal model for many-body quantum chaos. *Physical Review X*, 8(4):041019, November 2018.

[21] Bruno Bertini, Pavel Kos, and Tomaz Prosen. Entanglement spreading in a minimal model of maximal many-body quantum chaos. *Physical Review X*, 9(2):021033, May 2019.

[22] Bruno Bertini, Pavel Kos, and Tomaz Prosen. Operator Entanglement in Local Quantum Circuits I: Maximally Chaotic Dual-Unitary Circuits. *arXiv:1909.07407*, September 2019.

[23] Claudio Chamon, Alioscia Hamma, and Eduardo R. Mucciolo. Emergent irreversibility and entanglement spectrum statistics. *Physical Review Letters*, 112(24):240501, June 2014.

[24] Shiyu Zhou, Zhi-Cheng Yang, Alioscia Hamma, and Claudio Chamon. Single T gate in a Clifford circuit drives transition to universal entanglement spectrum statistics. *arXiv:1906.01079v2 [cond-mat.stat-mech]*, September 2019.

[25] Amos Chan, Andrea De Luca, and J. T. Chalker. Spectral statistics in spatially extended chaotic quantum many-body systems. *Physical Review Letters*, 121(6):060601, August 2018.

[26] Pavel Kos, Marko Ljubotina, and Tomaz Prosen. Many-body quantum chaos: Analytic connection to random matrix theory. *Physical Review X*, 8(2):021062, June 2018.

[27] Bruno Bertini, Pavel Kos, and Tomaz Prosen. Exact spectral form factor in a minimal model of many-body quantum chaos. *Physical Review Letters*, 121(26):264101, December 2018.

[28] Bin Yan, Lukasz Cincio, and Wojciech H. Zurek. Information Scrambling and Loschmidt Echo. *arXiv:1903.02651 [quant-ph]*, September 2019.

[29] D.P. DiVincenzo, D.W. Leung, and B.M. Terhal. Quantum data hiding. *IEEE Transactions on Information Theory*, 48(3):580–598, March 2002.

[30] Winton Brown and Omar Fawzi. Decoupling with Random Quantum Circuits. *Communications in Mathematical Physics*, 340(3):867–900, December 2015.

[31] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824–3851, November 1996.

[32] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. The mother of all protocols: Restructuring quantum information's family tree.

Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 465(2108):2537–2563, August 2009.

[33] Fernando G. S. L. Brandão, Aram W. Harrow, and Michał Horodecki. Efficient Quantum Pseudorandomness. *Physical Review Letters*, 116(17):170502, April 2016.

[34] A. J. Scott. Optimizing quantum process tomography with unitary2-designs. *Journal of Physics A: Mathematical and Theoretical*, 41(5):055308, January 2008.

[35] Easwar Magesan, J. M. Gambetta, and Joseph Emerson. Scalable and Robust Randomized Benchmarking of Quantum Processes. *Physical Review Letters*, 106(18):180504, May 2011.

[36] Barry Simon. *Representations of Finite and Compact Groups*. AMS, 1996.

[37] Joseph Emerson, Yaakov S. Weinstein, Marcos Saraceno, Seth Lloyd, and David G. Cory. Pseudo-Random Unitary Operators for Quantum Information Processing. *Science*, 302(5653):2098–2100, December 2003.

[38] Joseph Emerson, Etera Livine, and Seth Lloyd. Convergence conditions for random quantum circuits. *Physical Review A*, 72(6):060302, December 2005.

[39] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-Designs and their application to fidelity estimation. *Physical Review A*, 80(1):012304, July 2009.

[40] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio. Generic Entanglement Can Be Generated Efficiently. *Physical Review Letters*, 98(13):130502, March 2007.

[41] Shlomo Hoory, Avner Magen, Steven Myers, and Charles Rackoff. Simple Permutations Mix Well. In Josep Díaz, Juhani Karhumäki, Arto Lepistö,

and Donald Sannella, editors, *Automata, Languages and Programming*, Lecture Notes in Computer Science, pages 770–781, Berlin, Heidelberg, 2004. Springer.

[42] Alex Brodsky and Shlomo Hoory. Simple permutations mix even better. *Random Structures & Algorithms*, 32(3):274–289, 2008.

[43] Aidan Roy and A. J. Scott. Unitary designs and codes. *Designs, Codes and Cryptography*, 53(1):13–31, October 2009.

[44] Richard A. Low. Pseudo-randomness and Learning in Quantum Computation. *arXiv:1006.5227 [quant-ph]*, June 2010.

[45] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 20–30, New York, NY, USA, 1998. Association for Computing Machinery.

[46] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, Cambridge, 2018.

[47] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *Journal of Mathematical Physics*, 48(5):052104, 2007.

[48] Aram W. Harrow and Richard A. Low. Random Quantum Circuits are Approximate 2-Designs. *Communications in Mathematical Physics*, 291(1):257–302, October 2009.

[49] Igor Tuche Diniz and Daniel Jonathan. Comment on "Random Quantum Circuits are Approximate 2-designs" by A.W. Harrow and R.A. Low (Commun. Math. Phys. 291, 257–302 (2009)). *Communications in Mathematical Physics*, 304(1):281–293, May 2011.

[50] Alexander M. Dalzell, Nicholas Hunter-Jones, and Fernando G. S. L. Brandão. Random quantum circuits anti-concentrate in log depth. *arXiv:2011.12277 [cond-mat, physics:quant-ph]*, November 2020.

[51] Nicholas Hunter-Jones. Unitary designs from statistical mechanics in random quantum circuits. *arXiv:1905.12053 [cond-mat, physics:hep-th, physics:quant-ph]*, May 2019.

[52] Fernando G. S. L. Brandão, Aram W. Harrow, and Michał Horodecki. Local Random Quantum Circuits are Approximate Polynomial-Designs. *Communications in Mathematical Physics*, 346(2):397–434, September 2016.

[53] Aram Harrow and Saeed Mehraban. Approximate unitary t-Designs by short random quantum circuits using nearest-neighbor and long-range gates. *arXiv:1809.06957 [quant-ph]*, September 2018.

[54] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross. The Clifford group fails gracefully to be a unitary 4-Design. *arXiv:1609.08172 [quant-ph]*, September 2016.

[55] Zak Webb. The Clifford group forms a unitary 3-design. *Quantum Information & Computation*, 16(15&16):1379–1400, 2016.

[56] Jonas Haferkamp, Felipe Montealegre-Mora, Markus Heinrich, Jens Eisert, David Gross, and Ingo Roth. Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates. *arXiv:2002.09524 [math-ph, physics:quant-ph]*, June 2020.

[57] E. Onorati, O. Buerschaper, M. Kliesch, W. Brown, A. H. Werner, and J. Eisert. Mixing Properties of Stochastic Quantum Hamiltonians. *Communications in Mathematical Physics*, 355(3):905–947, November 2017.

[58] Yoshifumi Nakata, Christoph Hirche, Masato Koashi, and Andreas Winter. Efficient Quantum Pseudorandomness with Nearly Time-Independent Hamiltonian Dynamics. *Physical Review X*, 7(2):021006, April 2017.

[59] Vedika Khemani, Ashvin Vishwanath, and David A. Huse. Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws. *Physical Review X*, 8(3):031057, September 2018.

[60] Nicholas Hunter-Jones. Operator growth in random quantum circuits with symmetry. *arXiv:1812.08219 [cond-mat, physics:hep-th, physics:quant-ph]*, December 2018.

[61] Patrick Hayden and John Preskill. Black holes as mirrors: Quantum information in random subsystems. *Journal of High Energy Physics*, 2007(09):120–120, September 2007.

[62] Don N. Page. Average entropy of a subsystem. *Physical Review Letters*, 71(9):1291–1294, August 1993.

[63] Don N. Page. Information in black hole radiation. *Physical Review Letters*, 71(23):3743–3746, December 1993.

[64] Beni Yoshida and Alexei Kitaev. Efficient decoding for the Hayden-Preskill protocol. *arXiv:1710.03363 [hep-th, physics:quant-ph]*, October 2017.

[65] Yasuhiro Sekino and L. Susskind. Fast scramblers. *Journal of High Energy Physics*, 2008(10):065–065, October 2008.

[66] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the fast scrambling conjecture. *Journal of High Energy Physics*, 2013(4):22, April 2013.

[67] Winton Brown and Omar Fawzi. Scrambling speed of random quantum circuits. *arXiv:1210.6644 [hep-th, physics:quant-ph]*, July 2013.

[68] Fernando G. S. L. Brandão, Wissam Chemissany, Nicholas Hunter-Jones, Richard Kueng, and John Preskill. Models of quantum complexity growth. *arXiv:1912.04297 [cond-mat, physics:hep-th, physics:quant-ph]*, December 2019.

[69] Daniel A. Roberts and Beni Yoshida. Chaos and complexity by design. *Journal of High Energy Physics*, 2017(4):121, April 2017.

[70] Jonas Haferkamp, Philippe Faist, Naga B. T. Kothakonda, Jens Eisert, and Nicole Yunger Halpern. Linear growth of quantum circuit complexity. *arXiv:2106.05305 [hep-th, physics:math-ph, physics:quant-ph]*, June 2021.

[71] P. W. Anderson. Absence of Diffusion in Certain Random Lattices. *Physical Review*, 109(5):1492–1505, March 1958.

[72] R Abou-Chacra, D J Thouless, and P W Anderson. A selfconsistent theory of localization. *Journal of Physics C: Solid State Physics*, 6(10):1734–1752, May 1973.

[73] D. J. Thouless. Anderson's theory of localized states. *Journal of Physics C: Solid State Physics*, 3(7):1559–1566, July 1970.

[74] D C Herbert and R Jones. Localized states in disordered systems. *Journal of Physics C: Solid State Physics*, 4(10):1145–1161, July 1971.

[75] Cord A. Müller and Dominique Delande. Disorder and interference: Localization phenomena. *arXiv:1005.0915 [cond-mat, physics:quant-ph]*, June 2016.

[76] P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher. New method for a scaling theory of localization. *Physical Review B*, 22(8):3519–3526, October 1980.

[77] Patrick A. Lee and Daniel S. Fisher. Anderson Localization in Two Dimensions. *Physical Review Letters*, 47(12):882–885, September 1981.

[78] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan. Scaling Theory of Localization: Absence of Quantum Diffusion in Two Dimensions. *Physical Review Letters*, 42(10):673–676, March 1979.

[79] Antonello Scardicchio and Thimothée Thiery. Perturbation theory approaches to Anderson and Many-Body Localization: Some lecture notes. *arXiv:1710.01234 [cond-mat]*, October 2017.

[80] Dirk Hundertmark. A short introduction to Anderson localization. *Analysis and stochastics of growth processes and interface models*, 1:194–219, 2008.

[81] Ferdinand Evers and Alexander D. Mirlin. Anderson transitions. *Reviews of Modern Physics*, 80(4):1355–1417, October 2008.

[82] Abel Klein. Localization in the Anderson model with long range hopping. *BrazilianJournal of Physics*, 23:363–371, 1993.

[83] Alexander D. Mirlin, Yan V. Fyodorov, Frank-Michael Dittes, Javier Quezada, and Thomas H. Seligman. Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices. *Physical Review E*, 54(4):3221–3230, October 1996.

[84] Yan V. Fyodorov, Alexander D. Mirlin, and Hans-Jürgen Sommers. A novel field theoretical approach to the Anderson localization : Sparse random hopping model. *Journal de Physique I*, 2(8):1571–1605, August 1992.

[85] J. Biddle, D. J. Priour, B. Wang, and S. Das Sarma. Localization in one-dimensional lattices with non-nearest-neighbor hopping: Generalized Anderson and Aubry-Andr\'e models. *Physical Review B*, 83(7):075105, February 2011.

[86] G. L. Celardo, R. Kaiser, and F. Borgonovi. Shielding and localization in presence of long range hopping. *Physical Review B*, 94(14):144206, October 2016.

[87] Henrique von Dreifus and Abel Klein. A new proof of localization in the Anderson tight binding model. *Communications in Mathematical Physics*, 124(2):285–299, June 1989.

[88] Francisco A. B. F. de Moura and Marcelo L. Lyra. Delocalization in the 1D Anderson Model with Long-Range Correlated Disorder. *Physical Review Letters*, 81(17):3735–3738, October 1998.

[89] René Carmona, Abel Klein, and Fabio Martinelli. Anderson localization for Bernoulli and other singular potentials. *Communications in Mathematical Physics*, 108(1):41–66, 1987.

[90] Alexander Croy, Philipp Cain, and Michael Schreiber. Anderson Localization in 1D Systems with Correlated Disorder. *The European Physical Journal B*, 82(2):107–112, July 2011.

[91] Henrique von Dreifus and Abel Klein. Localization for random Schrödinger operators with correlated potentials. *Communications in Mathematical Physics*, 140(1):133–147, August 1991.

[92] Ad Lagendijk, Bart van Tiggelen, and Diederik S. Wiersma. Fifty years of Anderson localization. *Physics Today*, 62(8):24–29, August 2009.

[93] Marcel Filoche and Svitlana Mayboroda. Universal mechanism for Anderson and weak localization. *Proceedings of the National Academy of Sciences*, 109(37):14761–14766, September 2012.

[94] Sophie S. Shamailov, Dylan J. Brown, Thomas A. Haase, and Maarten D. Hoogerland. Anderson localisation in two dimensions: Insights from Localisation Landscape Theory, exact diagonalisation, and time-dependent simulations. *arXiv:2003.00149 [cond-mat]*, March 2021.

[95] Arijeet Pal and David A. Huse. Many-body localization phase transition. *Physical Review B*, 82(17):174411, November 2010.

[96] Rahul Nandkishore and David A. Huse. Many-Body Localization and Thermalization in Quantum Statistical Mechanics. *Annual Review of Condensed Matter Physics*, 6(1):15–38, March 2015.

[97] Fabien Alet and Nicolas Laflorencie. Many-body localization: An introduction and selected topics. *Comptes Rendus Physique*, 19(6):498–525, September 2018.

[98] Ehud Altman. Many-body localization and quantum thermalization. *Nature Physics*, 14(10):979–983, October 2018.

[99] Rodney J Baxter. *Exactly Solved Models in Statistical Mechanics*. Elsevier, 2016.

[100] L. D. Faddeev. How Algebraic Bethe Ansatz works for integrable model. *arXiv:hep-th/9605187*, May 1996.

[101] D. M. Basko, I. L. Aleiner, and B. L. Altshuler. Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states. *Annals of Physics*, 321(5):1126–1205, May 2006.

[102] Dmitry A. Abanin and Zlatko Papić. Recent progress in many-body localization. *Annalen der Physik*, 529(7):1700169, 2017.

[103] Dmitry A. Abanin, Ehud Altman, Immanuel Bloch, and Maksym Serbyn. Colloquium: Many-body localization, thermalization, and entanglement. *Reviews of Modern Physics*, 91(2):021001, May 2019.

[104] Marko Znidaric, Tomaz Prosen, and Peter Prelovsek. Many-body localization in the Heisenberg XXZ magnet in a random field. *Physical Review B*, 77(6):064426, February 2008.

[105] John Z. Imbrie, Valentina Ros, and Antonello Scardicchio. Local integrals of motion in many-body localized systems. *Annalen der Physik*, 529(7):1600278, 2017.

[106] David J. Luitz, Nicolas Laflorencie, and Fabien Alet. Many-body localization edge in the random-field Heisenberg chain. *Physical Review B*, 91(8):081103, February 2015.

[107] John Z. Imbrie. On many-body localization for quantum spin chains. *Journal of Statistical Physics*, 163(5):998–1048, 2016.

[108] David A. Huse, Rahul Nandkishore, and Vadim Oganesyan. Phenomenology of fully many-body-localized systems. *Physical Review B*, 90(17):174202, November 2014.

[109] Maksym Serbyn, Z. Papić, and Dmitry A. Abanin. Local Conservation Laws and the Structure of the Many-Body Localized States. *Physical Review Letters*, 111(12):127201, September 2013.

[110] Maksym Serbyn, Z. Papić, and Dmitry A. Abanin. Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems. *Physical Review Letters*, 110(26):260601, June 2013.

[111] Dong-Ling Deng, Xiaopeng Li, J. H. Pixley, Yang-Le Wu, and S. Das Sarma. Logarithmic entanglement lightcone in many-body localized systems. *Physical Review B*, 95(2):024202, January 2017.

[112] Fabio Anza, Francesca Pietracaprina, and John Goold. Logarithmic growth of local entropy and total correlations in many-body localized dynamics. *Quantum*, 4:250, April 2020.

[113] Christian K. Burrell and Tobias J. Osborne. Bounds on the Speed of Information Propagation in Disordered Quantum Spin Chains. *Physical Review Letters*, 99(16):167201, October 2007.

[114] Brian Swingle and Debanjan Chowdhury. Slow scrambling in disordered quantum systems. *Physical Review B*, 95(6):060201, February 2017.

[115] Yichen Huang, Yong-Liang Zhang, and Xie Chen. Out-of-time-ordered correlators in many-body localized systems. *Annalen der Physik*, 529(7):1600318, 2017.

[116] David J. Luitz and Yevgeny Bar Lev. The ergodic side of the many-body localization transition. *Annalen der Physik*, 529(7):1600350, July 2017.

[117] Maksym Serbyn and Joel E. Moore. Spectral statistics across the many-body localization transition. *Physical Review B*, 93(4):041424, January 2016.

[118] Subhayan Sahu, Shenglong Xu, and Brian Swingle. Scrambling Dynamics across a Thermalization-Localization Quantum Phase Transition. *Physical Review Letters*, 123(16):165902, October 2019.

[119] Thorsten B. Wahl, Arijeet Pal, and Steven H. Simon. Signatures of the many-body localized regime in two dimensions. *Nature Physics*, 15(2):164–169, February 2019.

[120] Tom Farshi, Daniele Toniolo, Carlos E. González-Guillén, Álvaro M. Alhambra, and Lluis Masanes. Time-periodic dynamics generates pseudo-random unitaries. *arXiv:2007.03339 [cond-mat, physics:math-ph, physics:quant-ph]*, July 2020.

[121] Robert Koenig and John A. Smolin. How to efficiently select an arbitrary Clifford group element. *Journal of Mathematical Physics*, 55(12):122202, December 2014.

[122] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70(5):052328, November 2004.

[123] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *arXiv:quant-ph/9807006*, July 1998.

[124] Marin Bukov, Luca D'Alessio, and Anatoli Polkovnikov. Universal high-frequency behavior of periodically driven systems: From dynamical stabilization to Floquet engineering. *Advances in Physics*, 64(2):139–226, 2015.

[125] Christoph Sünderhauf, David Pérez-García, David A. Huse, Norbert Schuch, and J. Ignacio Cirac. Localization with random time-periodic quantum circuits. *Physical Review B*, 98(13):134204, October 2018.

[126] Terry Farrelly. A review of Quantum Cellular Automata. *arXiv:1904.13318 [quant-ph]*, April 2019.

[127] Dirk-M. Schlingemann, Holger Vogts, and Reinhard F. Werner. On the structure of Clifford quantum cellular automata. *Journal of Mathematical Physics*, 49(11):112104, November 2008.

[128] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, 2015.

[129] Béla Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition, 2001.

[130] Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[131] Fan Chung and Linyuan Lu. The Diameter of Sparse Random Graphs. *Advances in Applied Mathematics*, 26(4):257–279, May 2001.

[132] Béla Bollobás. The Diameter of Random Graphs. *Transactions of the American Mathematical Society*, 267(1):41–52, 1981.

[133] Oliver Riordan and Nicholas Wormald. The diameter of sparse random graphs. *Combinatorics, Probability and Computing*, 19(5-6):835–926, November 2010.

[134] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. A bound on chaos. *Journal of High Energy Physics*, 2016(8):106, 2016.

[135] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator Spreading in Random Unitary Circuits. *Physical Review X*, 8(2):021014, April 2018.

[136] Richard Durrett. Oriented Percolation in Two Dimensions. *The Annals of Probability*, 12(4):999–1040, November 1984.

[137] Johannes Gütschow, Sonja Uphoff, Reinhard F. Werner, and Zoltán Zimborás. Time asymptotics and entanglement generation of Clifford quantum cellular automata. *Journal of Mathematical Physics*, 51(1):015203, 2010.

[138] Zoltán Zimborás, Terry Farrelly, Szilárd Farkas, and Lluis Masanes. Does causal dynamics imply local interactions? *arXiv:2006.10707 [cond-mat, physics:quant-ph]*, June 2020.

[139] Xiao Mi, Matteo Ippoliti, Chris Quintana, Ami Greene, Zijun Chen, Jonathan Gross, Frank Arute, Kunal Arya, Juan Atalaya, Ryan Babbush, Joseph C. Bardin, Joao Basso, Andreas Bengtsson, Alexander Bilmes, Alexandre Bourassa, Leon Brill, Michael Broughton, Bob B. Buckley, David A. Buell, Brian Burkett, Nicholas Bushnell, Benjamin Chiaro, Roberto Collins, William Courtney, Dripto Debroy, Sean Demura, Alan R. Derk, Andrew Dunsworth, Daniel Eppens, Catherine Erickson, Edward Farhi, Austin G. Fowler, Brooks Foxen, Craig Gidney, Marissa Giustina, Matthew P. Harrigan, Sean D. Harrington, Jeremy Hilton, Alan Ho, Sabrina Hong, Trent Huang, Ashley Huff, William J. Huggins, L. B. Ioffe, Sergei V. Isakov, Justin Iveland, Evan Jeffrey, Zhang Jiang, Cody Jones, Dvir Kafri, Tanuj Khattar, Seon Kim, Alexei Kitaev, Paul V. Klimov, Alexander N. Korotkov, Fedor Kostritsa, David Landhuis, Pavel Laptev, Joonho Lee, Kenny Lee, Aditya Locharla, Erik Lucero, Orion Martin, Jarrod R. McClean, Trevor McCourt, Matt McEwen, Kevin C. Miao, Masoud Mohseni, Shirin Montazeri, Wojciech Mruczkiewicz, Ofer Naaman, Matthew Neeley, Charles Neill, Michael Newman, Murphy Yuezhen Niu, Thomas E. O'Brien, Alex Opremcak, Eric Ostby, Balint Pato, Andre Petukhov, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vladimir Shvarts, Yuan Su, Doug Strain, Marco Szalay, Matthew D. Trevithick, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Juhwan Yoo, Adam Zalcman, Hartmut Neven, Sergio Boixo, Vadim Smelyanskiy, Anthony Megrant, Julian Kelly, Yu Chen, S. L. Sondhi, Roderich Moessner, Kostyantyn Kechedzhi, Vedika Khemani, and Pedram Roushan. Time-crystalline eigenstate order on a quantum processor. *Nature*, 601(7894):531–536, January 2022.

[140] David Gross, Sepehr Nezami, and Michael Walter. Schur-Weyl Duality for the Clifford Group with Applications: Property Testing, a Robust Hud-

son Theorem, and de Finetti Representations. *arXiv:1712.08628 [math-ph, physics:quant-ph]*, June 2019.

[141] A.R. Calderbank, E.M. Rains, P.M. Shor, and N.J.A. Sloane. Quantum error correction via codes over GF(4). *IEEE Transactions on Information Theory*, 44(4):1369–1387, July 1998.

[142] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, December 2006.