

# Unitary circuit dynamics

Lecture notes for the course on ‘Many-body Quantum Dynamics’ at TU Dresden

December 2023

---

## Contents

<b>1 Selected bibliography</b>	<b>1</b>
<b>2 Introduction</b>	<b>2</b>
<b>3 Operator dynamics</b>	<b>6</b>
<b>4 From Pauli matrices to Pauli strings</b>	<b>11</b>
<b>5 Operator scrambling</b>	<b>12</b>
5.1 Operator density	12
5.2 Out-of-time-order correlation functions	15
5.3 Entanglement dynamics	17
<b>References</b>	<b>20</b>

---

## 1 Selected bibliography

These notes are based on the following references:

- P. W. Claeys, A. Lamacraft and J. Herzog-Arbeitman, “Absence of Superdiffusion in Certain Random Spin Models,” *Phys. Rev. Lett.* **128**, 246603 (2022).
- A. Nahum, S. Vijay and J. Haah, “Operator Spreading in Random Unitary Circuits,” *Phys. Rev. X* **8**, 021014 (2018)
- C. W. von Keyserlingk, T. Rakovszky, F. Pollmann and S. L. Sondhi, “Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws,” *Phys. Rev. X* **8**, 021013 (2018).
- For a general introduction to tensor networks and the corresponding graphical language, this review is recommended: R. Orús, “A practical introduction to tensor networks: Matrix product states and projected entangled pair states,” *Annals of Physics* **349**, 117–158 (2014).
- Slides for an introduction to unitary and dual-unitary circuit dynamics (Quantum Circuits I & II) with a significant overlap with these notes can be found on [Austen Lamacraft’s website](#).

Please send any comments/typos/errors/... to [claeys@pks.mpg.de](mailto:claeys@pks.mpg.de)

## 2 Introduction

So far we have considered dynamics generated by a (local) Hamiltonian, which could be either static or time-dependent. In these lectures, we will consider dynamics where we are no longer restricted by the Hamiltonians that are either provided by Nature or that can be smartly designed in different ways, but we will rather consider dynamics where we can construct the unitary evolution by hand. These resulting *unitary circuit dynamics* offers both theoretical and experimental advantages. On the theory side, we can either design the unitary circuits in such a way that they are amenable to exact calculations, or we can introduce randomness in a way that is not possible for Hamiltonians and consider ensemble averages. On the experimental side, these dynamics are tailormade for implementation in current quantum computation setups, since the building blocks of unitary circuits are exactly the gates underlying the circuit model of quantum computing.

Before considering unitary circuits in their full generality, let us consider a specific setup that directly relates to the dynamics governed by a local Hamiltonian. Suppose we have a Hamiltonian acting on a one-dimensional lattice

$$\hat{H} = \sum_j \hat{h}_{j,j+1}, \quad (1)$$

where the index  $j$  runs over all lattice sites and  $\hat{h}_{j,j+1}$  only acts nontrivially on the sites  $j$  and  $j+1$ . Written in this way, the Hamiltonian is manifestly local and only involves nearest-neighbor interactions. All dynamics now corresponds to a unitary transformation with

$$U(t) = \exp[-i\hat{H}t]. \quad (2)$$

This unitary evolution operator can be made more transparent by writing

$$\exp[-i\hat{H}t] = \exp\left[-i\sum_j \hat{h}_{j,j+1}t\right] = \lim_{\Delta t \rightarrow 0} \left( \exp\left[-i\sum_j \hat{h}_{j,j+1}\Delta t\right] \right)^{t/\Delta t}, \quad (3)$$

where we have simply subdivided the full time evolution over a time  $t$  in smaller time steps  $\Delta t$ . Assume that we now keep  $\Delta t$  small but nonzero, the advantage of introducing this decomposition is that we can approximate

$$\exp\left[-i\sum_j \hat{h}_{j,j+1}\Delta t\right] \approx \exp\left[-i\sum_{j \text{ even}} \hat{h}_{j,j+1}\Delta t\right] \exp\left[-i\sum_{j \text{ odd}} \hat{h}_{j,j+1}\Delta t\right] + \mathcal{O}(\Delta t^2), \quad (4)$$

where we have used the Baker-Campbell-Hausdorff expansion to approximate the exponential. The correction term arises because the two different summations do not commute with each other, but can be neglected in the limit of small  $\Delta t$ . Plugging this in the initial equation, we can write

$$\begin{aligned} U(t) &\approx \left( \exp\left[-i\sum_{j \text{ even}} \hat{h}_{j,j+1}\Delta t\right] \exp\left[-i\sum_{j \text{ odd}} \hat{h}_{j,j+1}\Delta t\right] \right)^{t/\Delta t} \\ &= \left[ \left( \prod_{j \text{ even}} e^{-i\hat{h}_{j,j+1}\Delta t} \right) \left( \prod_{j \text{ odd}} e^{-i\hat{h}_{j,j+1}\Delta t} \right) \right]^{t/\Delta t} = \left[ \left( \prod_{j \text{ even}} U_{j,j+1} \right) \left( \prod_{j \text{ odd}} U_{j,j+1} \right) \right]^{n_t}. \end{aligned} \quad (5)$$

In the first equality we have used that all the terms in the separate summations commute, since no two terms  $\hat{h}_{j,j+1}$  act on the same site, and in the last equality we have defined  $U_{j,j+1} = e^{-i\hat{h}_{j,j+1}\Delta t}$  and introduced the discrete number of time steps  $n_t = t/\Delta t$ . Crucially, the full unitary evolution operator is expressed in terms of local 2-site unitary operators – we have effectively translated the locality of the Hamiltonian to locality of the unitary evolution operator, at the cost of introducing a small error. This decomposition is known as the *Suzuki-Trotter decomposition* and in fact underlies most numerical algorithms for many-body dynamics.

Unitary operators that are constructed out of local unitary matrices, also known as *unitary gates* in this context, are known as *unitary circuits*. These circuits exhibit the two main features of many-body dynamics: unitarity and locality. In order to keep track of the different unitaries and the sites they act on it will prove convenient to introduce a graphical language for unitary circuits.

### Tensor network diagrams

Within the language of tensor networks, objects(=tensors) are represented by shapes connected by wire, also called legs. Informally, every tensor corresponds to a shape, and every index needed to specify the elements of this tensor corresponds to a wire connected to this shape. I.e. we can consider

$$\begin{aligned}
 \text{Scalar } \lambda: & \quad \boxed{\lambda} \quad \Rightarrow \quad \lambda = \boxed{\lambda} \\
 \text{Vector } v: & \quad \boxed{v} \quad \Rightarrow \quad v_b = \boxed{v}^b \\
 \text{Matrix } M: & \quad \boxed{M} \quad \Rightarrow \quad M_{ab} = \boxed{M}^a_b \\
 \text{Tensor } T: & \quad \boxed{T} \quad \Rightarrow \quad T_{ab\dots cd\dots} = \boxed{T}^{ab\dots}_{cd\dots}
 \end{aligned} \tag{6}$$

The main advantage of this notation is that it allows us to straightforwardly write down matrix multiplication and, more generally, tensor contractions. Within multiplication two indices are taken to be equal and summed over, and in the graphical language this corresponds to simply connecting wires. Every wire corresponds to a single index, with the implicit assumption that the index associated with every connecting wire in a diagram is summed over. As one example, if we multiply the matrix  $M$  with matrix elements  $M_{ab}$  with a vector  $v$  with vector elements  $v_b$ , we obtain a new vector  $Mv$  with elements

$$(Mv)_a = \sum_b M_{ab}v_b = \sum_b \boxed{M}^a_b \boxed{v}^b \quad \Rightarrow \quad Mv = \boxed{M}^a_b \boxed{v}^b \tag{7}$$

Such contractions also naturally appear when evaluating the inner product between two vec-

tors, e.g.  $v$  and  $w$ , as

$$w^T v = \sum_a w_a v_a = \sum_a \begin{array}{c} \boxed{w} \\ | \\ a \\ | \\ \boxed{v} \end{array} \Rightarrow w^T v = \begin{array}{c} \boxed{w} \\ | \\ \boxed{v} \end{array} \quad (8)$$

or when taking the trace of a matrix, since

$$\text{Tr}(M) = \sum_a M_{aa} = \sum_a \begin{array}{c} \boxed{M} \\ | \\ a \\ | \\ \boxed{M} \end{array} \Rightarrow \text{Tr}(M) = \begin{array}{c} \boxed{M} \\ | \\ \boxed{M} \end{array} \quad (9)$$

Note that both the trace and the inner product return a scalar, as evidenced by the fact that the resulting diagrams have no open wires.

Within this language, the identity matrix can be represented as a single line, since the matrix elements of the identity are only nonzero if the two indices are equal, such that these indices corresponds to the same wire.

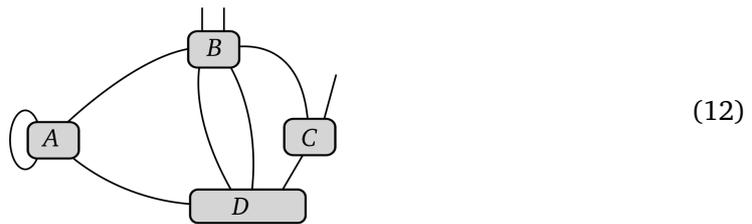
$$\mathbb{1} = \begin{array}{c} | \\ | \\ | \end{array} \Rightarrow \mathbb{1}_{ab} = \delta_{ab} = \begin{array}{c} a \\ | \\ | \\ b \end{array} \quad (10)$$

Multiplying the identity with a vector, it is clear that this action leaves the vector invariant.

As one final important example, the direct product of two matrices can simply be written by taking the diagrams of both matrices and placing them next to each other:

$$C = A \otimes B = \begin{array}{c} \boxed{C} \\ | \quad | \\ | \quad | \\ | \quad | \end{array} = \begin{array}{c} \boxed{A} \\ | \\ | \\ | \end{array} \begin{array}{c} \boxed{B} \\ | \\ | \\ | \end{array} \Rightarrow C_{ab,cd} = A_{ac} B_{bd} = \begin{array}{c} a \quad b \\ | \quad | \\ \boxed{C} \\ | \quad | \\ c \quad d \end{array} = \begin{array}{c} a \\ | \\ \boxed{A} \\ | \\ c \end{array} \begin{array}{c} b \\ | \\ \boxed{B} \\ | \\ d \end{array} \quad (11)$$

Any arbitrary tensor network can then be written as a general composition of shapes and connecting wires, e.g.



### Unitary gates

Our unitary matrices acting on two lattice sites can be represented as a shape, here chosen to be a square, with four wires: two ‘input’ wires and two ‘output’ wires. We represent the two-site gates and their matrix elements as

$$U = \begin{array}{c} \diagup \quad \diagdown \\ \boxed{U} \\ \diagdown \quad \diagup \end{array}, \quad U_{ab,cd} = \langle ab|U|cd\rangle = \begin{array}{c} a \quad b \\ \diagdown \quad \diagup \\ \boxed{U} \\ \diagup \quad \diagdown \\ c \quad d \end{array}, \quad (13)$$

where  $a, b, c, d \in \{0, 1\}$  for a local two-dimensional Hilbert space, as is the case for spin-1/2 particles or qubits. We denote  $U^\dagger$  in the same way, but with a red square instead of a blue square. The two input wires can also be thought of as representing a single index in the Hilbert space for two sites, which is equivalent to taking an individual wire to represent a single index in the Hilbert space for a single site. Here we use the two wires in order to make explicit that these operators act on two sites.

Unitarity implies that

$$UU^\dagger = U^\dagger U = \mathbb{1} \Rightarrow \begin{array}{c} \text{blue square} \\ \text{red square} \end{array} = \begin{array}{c} \text{red square} \\ \text{blue square} \end{array} = \begin{array}{|l} \hline \\ \hline \end{array} \quad (14)$$

or, expressed in matrix elements,

$$\sum_{ef} U_{ab,ef}(U^\dagger)_{ef,cd} = \sum_{ef} (U^\dagger)_{ab,ef} U_{ef,cd} = \delta_{ac,bd} \Rightarrow \begin{array}{c} a \quad b \\ \text{blue square} \\ e \quad f \\ \text{red square} \\ c \quad d \end{array} = \begin{array}{c} a \quad b \\ \text{red square} \\ e \quad f \\ \text{blue square} \\ c \quad d \end{array} = \begin{array}{|l} a \\ \hline b \\ \hline \end{array} \quad (15)$$

Within this graphical notation, the full evolution operator can be represented as a so-called 'brickwork' circuit of unitary matrices:

$$\left[ \left( \prod_{j \text{ even}} U_{j,j+1} \right) \left( \prod_{j \text{ odd}} U_{j,j+1} \right) \right]^2 = \begin{array}{cccc} \dots & \text{blue square} & \text{blue square} & \text{blue square} & \text{blue square} & \dots \\ & \diagdown & \diagup & \diagdown & \diagup & \\ \text{blue square} & & \text{blue square} & & \text{blue square} & \dots \\ & \diagup & \diagdown & \diagup & \diagdown & \\ \dots & \text{blue square} & \text{blue square} & \text{blue square} & \text{blue square} & \dots \\ & \diagdown & \diagup & \diagdown & \diagup & \\ \dots & \text{blue square} & \text{blue square} & \text{blue square} & \text{blue square} & \dots \end{array} \quad (16)$$

where the number of layers, here 4, equals  $2n_t$ . In the remainder of these notes, we will work with a discrete time  $t$  in such a way that every layer of unitary gates corresponds to a single discrete time step. If all gates are identical, this effectively realizes a Floquet protocol with period  $T = 2$ . The graphical notation significantly simplifies the bookkeeping involved when keeping track of which operators act on which sites.

The two-site unitary matrices appearing in this circuit naturally appear in quantum computing, where they form the basic building block of quantum operations on two qubits. In this context they are also referred to as 'unitary gates'. Any logical operator from (reversible) classical computing can be expressed as a permutation matrix, e.g. the CNOT gate (controlled not), acting on two bits as

$$\text{CNOT}|00\rangle = |00\rangle, \quad \text{CNOT}|01\rangle = |01\rangle, \quad \text{CNOT}|10\rangle = |11\rangle, \quad \text{CNOT}|11\rangle = |10\rangle, \quad (17)$$

which has the matrix representation

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (18)$$

in the basis  $\{00, 01, 10, 11\}$ .

Other classical gates that often appear in the context of quantum circuits are the identity gate  $\mathbb{1}$  and the swap gate  $s$ , i.e.

$$\mathbb{1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{|c|} \hline | \\ \hline \end{array}, \quad S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{|c|} \hline \diagdown \\ \diagup \\ \hline \end{array}. \quad (19)$$

Here the graphical notation directly represents the classical action, i.e.  $\mathbb{1}|ab\rangle = |ab\rangle$  and hence  $\mathbb{1}_{ab,cd} = \delta_{ac}\delta_{bd}$  and  $S|ab\rangle = |ba\rangle$  and hence  $S_{ab,cd} = \delta_{ad,bc}$ .

For completeness, we note that any arbitrary two-site unitary gate can be written as

$$U = e^{i\phi}(u_1 \otimes u_2) \exp[-i(J_x \sigma^x \otimes \sigma^x + J_y \sigma^y \otimes \sigma^y + J_z \sigma^z \otimes \sigma^z)](v_1 \otimes v_2), \quad (20)$$

where  $e^{i\phi}$  is a global phase,  $u_{1,2}$  and  $v_{1,2} \in SU(2)$  are one-site unitaries and  $J_\alpha$  are real numbers setting the entangling properties of the gate.

### 3 Operator dynamics

Unitary circuit dynamics is particularly instructive when considering the dynamics of operators rather than states, i.e. moving from the Schrödinger picture to the Heisenberg picture. In the Heisenberg picture, operators evolve as

$$O(t) = U(t)OU(t)^\dagger. \quad (21)$$

For convenience we exchange the role of  $U$  and its hermitian conjugate as opposed to the commonly used convention – in order to be technically correct the dynamics should be thought of as the dynamics of a density matrix rather than an observable, but since we can choose  $U$  freely this distinction is irrelevant in what follows.

Consider an initial observable  $O$  that only acts nontrivially on site  $j$ , e.g. a Pauli matrix  $\sigma^\alpha$  with  $\alpha \in x, y, z$ . We write

$$O(t=0) = \sigma_j^\alpha = \mathbb{1} \cdots \otimes \underbrace{\mathbb{1}}_{j-1} \otimes \underbrace{\sigma^\alpha}_j \otimes \underbrace{\mathbb{1}}_{j+1} \otimes \cdots \mathbb{1} = \begin{array}{|c|} \hline \dots \\ \hline \dots \\ \hline \end{array} \begin{array}{|c|} \hline \circlearrowleft \\ \hline \end{array} \begin{array}{|c|} \hline \dots \\ \hline \end{array} \quad (22)$$

Here the indices  $j-1, j, j+1$  label the site on which these operators act (and not the elements of these matrices). Evolving this one-site operator with the unitary brickwork circuit, we find

that after the application of a single layer we can write

$$\begin{aligned}
 O(t=1) &= \begin{array}{c} \text{[Diagram: Five pairs of blue and red squares connected by arcs, representing a one-site operator.]} \\ \text{[Diagram: The same five pairs of squares, but the middle pair is now connected to its neighbors, representing a two-site operator.]} \end{array} \\
 &= \begin{array}{c} \text{[Diagram: Five vertical lines representing sites, with a blue and red square pair on the third line.]} \end{array} \tag{23}
 \end{aligned}$$

Here we have only used the property that the unitary gates are unitary, such that every contraction of a gate with its hermitian conjugate returns the identity. We find that the time-evolved operator acts as the identity almost everywhere except on site  $j$  and  $j + 1$ : we have moved from a one-site operator to a two-site operator. Taking another time step, the time-evolved operator can be written as

$$\begin{aligned}
 O(t=2) &= \begin{array}{c} \text{[Diagram: A central blue and red square pair connected to four other blue and red square pairs, representing a four-site operator.]} \end{array} \tag{24}
 \end{aligned}$$

The operator now acts nontrivially on 4 sites. Taking an additional step, we find that

$$\begin{aligned}
 O(t=3) &= \begin{array}{c} \text{[Diagram: A central blue and red square pair connected to six other blue and red square pairs, representing a six-site operator.]} \end{array} \tag{25}
 \end{aligned}$$

now acting nontrivially on 6 sites. These examples illustrate the general principle that the support of the time-evolved operator grows *linearly* in time. The operator can only act nontrivially on a site a distance  $x$  away from the original support of the operator after  $t$  time steps: we have an effective emergent *causal light cone*. All correlations can spread no faster than ballistically with velocity  $v = 1$  simply due to the geometry of the system and the locality of the interactions. However, this velocity is only an upper bound, and in practice operator spreading is significantly slower than this upper bound.

Let us consider this operator dynamics using a particular choice of entangling two-site gate, choosing

$$U = \cos \theta \mathbb{1} + i \sin \theta S. \tag{26}$$

This gate acts as

$$U|00\rangle = \cos\theta|00\rangle + i\sin\theta|00\rangle = e^{i\theta}|00\rangle, \quad (27)$$

$$U|01\rangle = \cos\theta|01\rangle + i\sin\theta|10\rangle, \quad (28)$$

$$U|10\rangle = \cos\theta|10\rangle + i\sin\theta|01\rangle, \quad (29)$$

$$U|11\rangle = \cos\theta|11\rangle + i\sin\theta|11\rangle = e^{i\theta}|11\rangle, \quad (30)$$

effectively rotating the  $|01\rangle$  and  $|10\rangle$  states into each other. We have that

$$U = \begin{pmatrix} e^{i\theta} & 0 & 0 & 0 \\ 0 & \cos\theta & i\sin\theta & 0 \\ 0 & i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{pmatrix}. \quad (31)$$

This gate can be seen as a minimal model for gates with a symmetry, since it clearly conserves the total number of 0 and 1 states. In other words, this gate has a  $U(1)$  symmetry, conserving a total number of ‘spin excitations’,

$$[U, \sigma_1^z + \sigma_2^z] = 0. \quad (32)$$

In fact, this gate has a more general  $SU(2)$  symmetry since this symmetry also applies to the other Pauli matrices:

$$[U, \sigma_1^\alpha + \sigma_2^\alpha] = 0, \quad \forall \alpha \in \{x, y, z\}. \quad (33)$$

That this symmetry holds for all Pauli matrices can be understood by noting that the identity and swap gate are defined in a basis-agnostic way and act the same irrespective of the choice of local basis. Our initial choice of basis was the basis of eigenstates of  $\sigma^z$ , in which the conservation of  $\sigma_1^z + \sigma_2^z$  can simply be read off, but there is nothing special about this basis.

The operator dynamics generated by this gate can be understood by considering the action of the gate on a generic two-site operator  $O$ , as

$$\begin{aligned} UOU^\dagger &= [\cos\theta \mathbb{1} + i\sin\theta S]O[\cos\theta \mathbb{1} - i\sin\theta S] \\ &= \cos^2\theta O + \sin^2\theta SOS + i\sin\theta \cos\theta [S, O] \\ &= O + \sin^2(\theta)(SOS - O) + \frac{i}{2}\sin(2\theta)[S, O]. \end{aligned} \quad (34)$$

Let us consider the action of the two separate terms on the operator basis of Pauli matrices. The commutator with the swap operator can be directly calculated by noting that

$$S = \frac{1}{2} \left( \mathbb{1} \otimes \mathbb{1} + \sum_\alpha \sigma^\alpha \otimes \sigma^\alpha \right). \quad (35)$$

(Recall that the Heisenberg interaction is an ‘exchange’ interaction.)

Either  $O$  acts trivially on one site, in which case we need to consider

$$O = \sigma^\alpha \otimes \mathbb{1}: \quad S(\sigma^\alpha \otimes \mathbb{1})S = \mathbb{1} \otimes \sigma^\alpha, \quad i[S, \sigma^\alpha \otimes \mathbb{1}] = -\epsilon_{\alpha\beta\gamma} \sigma^\beta \otimes \sigma^\gamma, \quad (36)$$

$$O = \mathbb{1} \otimes \sigma^\alpha: \quad S(\mathbb{1} \otimes \sigma^\alpha)S = \sigma^\alpha \otimes \mathbb{1}, \quad i[S, \mathbb{1} \otimes \sigma^\alpha] = \epsilon_{\alpha\beta\gamma} \sigma^\beta \otimes \sigma^\gamma, \quad (37)$$

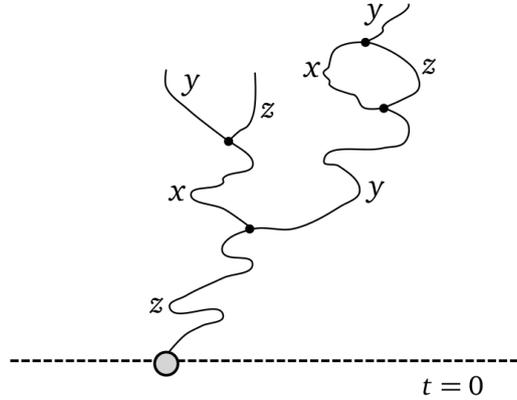
with  $\epsilon_{\alpha\beta\gamma}$  the Levi-Civita symbol, which satisfies  $\epsilon_{xyz} = 1$  and is otherwise purely antisymmetric under the exchange of indices,  $\epsilon_{\alpha\beta\gamma} = -\epsilon_{\beta\alpha\gamma} = -\epsilon_{\alpha\gamma\beta}$ , and where we use Einstein summation over  $\beta$  and  $\gamma$  in the right hand side. Note that the two commutators are necessarily antisymmetric, since we know that  $[S, \mathbb{1} \otimes \sigma^\alpha + \sigma^\alpha \otimes \mathbb{1}] = 0$ .

Alternatively,  $O$  acts nontrivially on both sites, and we now need to consider

$$O = \sigma^\alpha \otimes \sigma^\beta : \quad S(\sigma^\alpha \otimes \sigma^\beta)S = \sigma^\beta \otimes \sigma^\alpha, \quad i[S, \sigma^\alpha \otimes \sigma^\beta] = \epsilon_{\alpha\beta\gamma} (\sigma^\gamma \otimes \mathbb{1} - \mathbb{1} \otimes \sigma^\gamma). \quad (38)$$

As could be expected, the action of the swap operator just moves operators around, whereas the commutator results either in a single Pauli matrix ‘splitting’ in two other Pauli matrices or two Pauli matrices ‘merging’ in a single Pauli matrix.

These two processes are a general feature of unitary operator dynamics, which typically involves both *operator entanglement*, i.e. the generation of a linear combination of Pauli matrices, and *operator growth*, where Pauli matrices with increasing support appear. Schematically, the processes in the operator dynamics of an initial  $\sigma^z$  matrix can be illustrated as in Fig. 1



**Figure 1:** Illustration of the operator dynamics of a single initial Pauli  $z$  matrix. The swap operator can move the Pauli matrix around, and the commutator can split a single  $z$  into an  $x$  and  $y$  Pauli matrix, which can in turn split and merge again.

The generation of entanglement and the large Hilbert space required for a many-body system are the two factors that limit the numerical simulation of the unitary dynamics of an initial wave function, and we see that these factors reappear here. The general numerical simulation of operator dynamics is similarly exponentially costly. However, we can now make a crucial assumption: we can introduce some degree of randomness and consider an ensemble of unitary circuits, and assume that the ensemble average of the quantities we are interested in is representative of the generic case. Typically, we choose all the individual gates in the circuit to be i.i.d. distributed, such that the dynamics is effectively random in both time and space.

For the example at hand, we can choose a random distribution of the angle  $\theta$ , e.g. we randomly choose  $\theta = \pm\sqrt{D}$  with equal probability, where  $D$  is a small number, such that we have  $\overline{\theta} = 0$  and  $\overline{\theta^2} = D$ . Let us now consider the ensemble average of the unitary transformation with  $U$ . We find that

$$\overline{UOU^\dagger} = O + \overline{\sin^2(\theta)}(SOS - O) + \frac{i}{2}\overline{\sin(2\theta)}[S, O] = O + D(SOS - O). \quad (39)$$

Introducing the ensemble averaging effectively removes the operator growth, since all two-site operators will appear with opposite signs with equal probability and hence cancel out. The ensemble averaged dynamics corresponds to the classical problem of random adjacent transpositions, where operators on two neighboring sites can switch places with probability  $D$ .

The full lattice dynamics for the ensemble averaged dynamics reads

$$\frac{d}{dt}O = D \sum_j [S_{j,j+1}OS_{j,j+1} - O], \quad (40)$$

where we have taken the continuum limit in time for convenience. We know that the swap operators do not generate any new operators, such that for an initial operator  $O(t=0) = \sigma_0^z$  localized at site  $j=0$  we can write

$$O(t) = \sum_j C_j^z(t) \sigma_j^z, \quad (41)$$

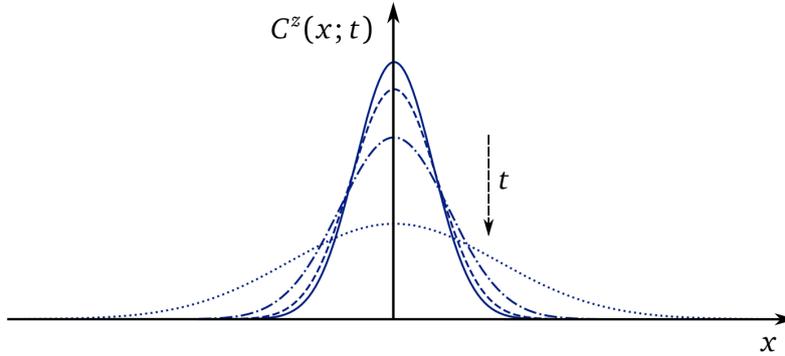
and plugging this ansatz in Eq. (40) we find that the amplitudes  $C_j^z(t)$  satisfy

$$\partial_t C_j^z = D (C_{j+1}^z + C_{j-1}^z - 2C_j^z) \equiv D \Delta_j C_j^z, \quad (42)$$

where we have identified the discrete Laplacian  $\Delta_j$ . This equation is nothing but the discrete diffusion equation with  $D$  the diffusion constant, which can be solved by

$$C_z(x; t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right], \quad (43)$$

where we have for convenience taken the continuum limit in space and introduced the continuous coordinate  $x \propto j$ . This profile is illustrated in Fig. 2.



**Figure 2:** Illustration of the Gaussian profile for  $C^z(x; t)$  in Eq. (43) with the width growing diffusively as  $\sqrt{t}$ .

These coefficients directly appear in correlation functions, since e.g.

$$\langle \sigma_0^z(t) \sigma_k^z \rangle_{\beta=0} \equiv \text{Tr}[\sigma_0^z(t) \sigma_k^z] / \mathcal{D} = \sum_j C_j^z(t) \text{Tr}[\sigma_j^z \sigma_k^z] / \mathcal{D} = C_k^z(t). \quad (44)$$

where  $\mathcal{D}$  is the dimension of the Hilbert space appearing as normalization constant,  $\mathcal{D} = \text{Tr}(\mathbb{1})$ . Here we have used the trace-orthonormality of the Pauli matrices,  $\text{Tr}[\sigma_j^z \sigma_k^z] / \mathcal{D} = \delta_{jk}$ .

The random circuit realizes a minimal model of *diffusion*. Diffusion is generically expected whenever the dynamics support a conservation law and appears naturally in any coarse-grained hydrodynamic description of quantum dynamics. Suppose that we have a global conserved charge  $Q = \sum_x q_x$ . Any local conserved charge results in a continuity equation

$$\partial_t \langle q_x \rangle = -\partial_x \langle j_x \rangle, \quad (45)$$

where  $j_x$  is a local current operator. Assuming that the system is in local equilibrium we can write an effective expansion of  $\langle j_x \rangle$  as function of the local density. Since the current needs to vanish for a constant charge density, to first order we generally obtain Fick's law,  $\langle j_x \rangle = -D \partial_x \langle q_x \rangle$ , which then results in the diffusion equation  $\partial_t \langle q_x \rangle = D \partial_x^2 \langle q_x \rangle$ . Using only minimal assumptions and no coarse-graining, we have here shown how such diffusion naturally arises in a random circuit context.

## 4 From Pauli matrices to Pauli strings

In the previous section, we showed that we could solve a minimal model of diffusion by writing the time-evolved operator as a linear combination of Pauli matrices acting on single sites. More generally, we need to consider products of Pauli matrices acting on an arbitrary number of sites, and

$$O(t) = \sum_{\alpha_j \in \{0,x,y,z\}} C_{\alpha_1, \alpha_2, \dots, \alpha_N}(t) \sigma^{\alpha_1} \otimes \sigma^{\alpha_2} \dots \otimes \sigma^{\alpha_N} \equiv \sum_{\mathcal{S}} C_{\mathcal{S}}(t) \mathcal{S}, \quad (46)$$

where we have identified  $\sigma^0 = \mathbb{1}$ . For convenience, we denote arbitrary Pauli strings as  $\mathcal{S}$ . For a system of  $N$  sites we require  $4^N$  coefficients to fully specify the operator at any arbitrary time. The Pauli strings generally form a trace-orthonormal basis, since

$$\text{Tr}[\mathcal{S}' \cdot \mathcal{S}] / \mathcal{D} = \delta_{\mathcal{S}, \mathcal{S}'}. \quad (47)$$

This identity can be directly derived by noting that all Pauli matrices are traceless and square to the identity, such that the above trace will only be nonzero if  $\mathcal{S}' \cdot \mathcal{S} = \mathbb{1}$ . As such, the expansion (46) in Pauli strings is similar to decomposing a general wave function in an orthonormal basis. In the same way that unitary dynamics preserves the norm of the wave function, unitary dynamics preserves the operator norm:

$$\text{Tr}[O(t)^\dagger O(t)] / \mathcal{D} = \sum_{\mathcal{S}, \mathcal{S}'} C_{\mathcal{S}'}^*(t) C_{\mathcal{S}}(t) \text{Tr}[\mathcal{S}' \cdot \mathcal{S}] / \mathcal{D} = \sum_{\mathcal{S}} |C_{\mathcal{S}}(t)|^2 \quad (48)$$

and the left-hand side is time-independent since

$$\text{Tr}[O(t)^\dagger O(t)] / \mathcal{D} = \text{Tr}[U O^\dagger U^\dagger U O U^\dagger] / \mathcal{D} = \text{Tr}[O^\dagger O] / \mathcal{D}. \quad (49)$$

Typically we start from an orthonormalized operator, e.g. a single Pauli matrix, for which  $\text{Tr}[O^\dagger O] / \mathcal{D} = 1$ , such that

$$\sum_{\mathcal{S}} |C_{\mathcal{S}}(t)|^2 = \sum_{\mathcal{S}} |C_{\mathcal{S}}(t=0)|^2 = 1. \quad (50)$$

Each of these coefficients has the interpretation of an infinite-temperature correlation function, since

$$C_{\mathcal{S}}(t) = \langle O(t) \cdot \mathcal{S} \rangle_{\beta=0} = \text{Tr}[O(t) \cdot \mathcal{S}] / \mathcal{D}, \quad (51)$$

recalling that the trace corresponds to the infinite-temperature correlation function, which is here the appropriate correlation function since there is no notion of energy.

While the conservation of the norm of a wave function indicates conservation of probability, conservation of probability appears in a different way for operator dynamics as compared to

the state dynamics. Conservation of probability here fixes the single component associated with  $\mathcal{S} = \mathbb{1}$ . Denoting  $C_{\mathcal{S}=\mathbb{1}}$  as  $C_0$ , we have that

$$C_0(t) = \text{Tr}[O(t) \cdot \mathbb{1}] / \mathcal{D} = \text{Tr}[UOU^\dagger] / \mathcal{D} = \text{Tr}[O] / \mathcal{D} = C_0(t=0). \quad (52)$$

Similar in spirit to our decomposition of the dynamics of an observable in a static and dynamic part, we can here write

$$O(t) = C_0 \mathbb{1} + \sum_{\mathcal{S} \neq 0} C_{\mathcal{S}} \mathcal{S}. \quad (53)$$

The static part can now be even more explicitly related to the thermalization to an infinite-temperature reduced density matrix. If we interpret  $O$  as an initial density matrix, which we can do for an initial state with sufficiently short-range correlations, then we obtain the reduced density matrix for a subsystem  $A$  as

$$\rho_A = \text{Tr}_{\bar{A}}[O(t)] = C_0 \text{Tr}_{\bar{A}}[\mathbb{1}] + \sum_{\mathcal{S}} C_{\mathcal{S}}(t) \text{Tr}_{\bar{A}}[\mathcal{S}] = C_0 \text{Tr}_{\bar{A}}[\mathbb{1}] + \sum_{\mathcal{S} \in A} C_{\mathcal{S}}(t) \text{Tr}_{\bar{A}}[\mathcal{S}]. \quad (54)$$

In the second equality, we have restricted the summation to all operators  $\mathcal{S}$  whose support is fully contained in  $A$ : For all other terms the trace vanishes identically since the Pauli matrices are traceless. Any operator that contains a Pauli matrix that acts outside of  $A$  does not contribute to the reduced density matrix. Within the dynamics the operator spreads and the Pauli strings grow, such that the weight of the terms acting outside of  $A$  will increase, and conservation of operator norm then implies that the terms acting only within  $A$  will decay. When the full summation can be neglected, the system has thermalized to an infinite-temperature density matrix.

Effectively, as time goes on the operator becomes more complex, and an increasing fraction of the initial operator ‘leaks’ out of the subsystem  $A$  (i.e. has a support that is no longer fully contained in  $A$ ). Local information is said to ‘scramble’ nonlocally and the reduced density matrix relaxes to equilibrium since it only probes local information.

## 5 Operator scrambling

### 5.1 Operator density

We would now like to better understand and physically probe operator scrambling. In our initial random circuit calculation all terms  $C_{\mathcal{S}}$  where  $\mathcal{S}$  acted nontrivially on multiple sites cancelled out due to their random signs under averaging. We can consider a measure that is invariant under these sign changes by considering the average dynamics of  $|C_{\mathcal{S}}(\tau)|^2$ . We quantify the support of an initial operator by defining the operator density<sup>1</sup>

$$\rho_R(s; \tau) = \sum_{\mathcal{S}} |C_{\mathcal{S}}(\tau)|^2 \delta[\mathcal{S} \text{ ends on site } s]. \quad (55)$$

Here we consider the total weight of the Pauli strings that end on the right on site  $s$  (we can similarly identify an operator density for the left). E.g.  $\sigma_1^\alpha \sigma_3^\beta$  ends on site  $s = 3$ .

The operator density clearly satisfies a conservation law, since

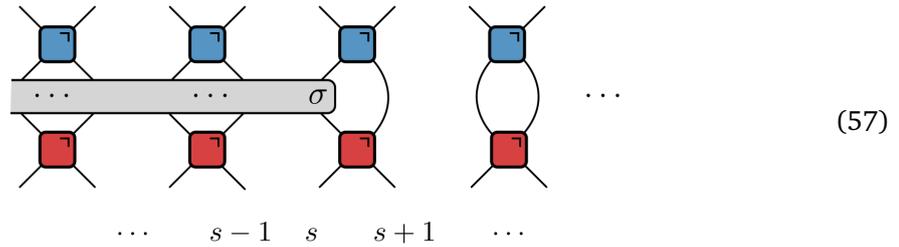
$$\sum_s \rho_R(s) = 1. \quad (56)$$

<sup>1</sup>Discrete position and time are here initially denoted by  $s$  and  $\tau$  respectively in order to be consistent with the literature. For identical gates the time evolution is periodic with period  $\tau = 2$ .

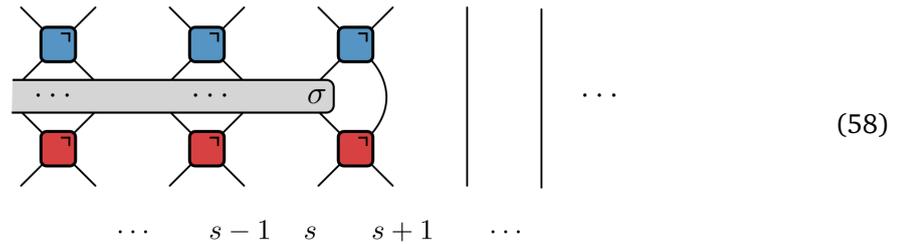
At time  $\tau = 0$  we consider an operator that is purely localized on site 0 such that the operator density is a Kronecker delta,  $\rho_R(s) = \delta_{s,0}$ . As time goes on we expect the operator to grow and the operator density to spread out in space.

Rather than consider the operator dynamics where the gates are linear combinations of the identity and swap, we consider maximally random unitary gates, i.e. gates distributed according to the *Haar distribution*. This distribution has the property that if  $U$  is Haar random distributed then both  $UV$  and  $VU$  are Haar random distributed, for any unitary  $V$ . This property uniquely fixes the distribution and is in fact is the only property that we will need.

Suppose that we have a Pauli string that ends at site  $s$ . For the unitary evolution over a single time step we can consider the update at the outermost right edge of the Pauli string as



where, since the Pauli string acts as the identity on all sites to the right of site  $s$ , we have



The change in operator density is purely determined by the action of  $U$  on the right-most Pauli matrix in the Pauli string. For the operator density there are now two options for the local update: either this unitary transformation increases the support of the Pauli string by a single site, or its support stays the same.

For Haar-random gates we need to consider the ensemble average of the local update  $U(\sigma \otimes \mathbb{1})U^\dagger$ , i.e.

$$U(\sigma \otimes \mathbb{1})U^\dagger = \text{Diagram} \quad (59)$$

Note that we here drop the superscript  $\alpha \in \{x, y, z\}$  for the Pauli operator. Since the gates are Haar-random distributed, any Pauli matrix will be equivalent under ensemble averaging. For the two-site operator basis we now have a complete set of 16 operators: the identity  $\mathbb{1} \times \mathbb{1}$ , 3 operators of the form  $\sigma \otimes \mathbb{1}$ , 3 operators of the form  $\mathbb{1} \otimes \sigma$ , and 9 operators of the form  $\sigma \otimes \sigma$ . Under Haar averaging, all Pauli operators that are not the identity are fully equivalent, such that they all get mapped to each other with equal probability<sup>2</sup>. Only the identity is 'special', since  $U(\mathbb{1} \otimes \mathbb{1})U^\dagger = \mathbb{1} \otimes \mathbb{1}$  always and

$$\mathbb{1} \otimes \mathbb{1} \quad \rightarrow \quad \mathbb{1} \otimes \mathbb{1} \quad \text{with probability 1.} \quad (60)$$

<sup>2</sup>It is easy to check that  $\mathbb{1} \otimes \sigma^\alpha$ ,  $\sigma^\alpha \otimes \mathbb{1}$  and  $\sigma^\alpha \otimes \sigma^\beta$  have the same eigenspectrum, such that they are related through a unitary transformation and hence indistinguishable w.r.t. the Haar distribution.

Under the unitary transformation  $\sigma \otimes \mathbb{1}$  hence gets mapped to an operator of the form  $\sigma \otimes \mathbb{1}$  with probability  $p = 3/15 = 1/5$  and the operator support does not increase. With a probability  $1 - p = 12/15 = 4/5$  it gets mapped to an operator of the form  $\mathbb{1} \otimes \sigma$  or  $\sigma \otimes \sigma$  and the operator support increases by 1. We can extend this argument to an arbitrary  $q$ -dimensional local Hilbert space, where the number of ‘Pauli matrices’ (Hermitian, orthonormal and traceless) equals  $q^2 - 1$ , such that

$$p = \frac{q^2 - 1}{q^4 - 1} = \frac{1}{q^2 + 1}. \quad (61)$$

Taking this together, under a Haar random unitary transformations  $\mathbb{1} \otimes \sigma$  maps to

$$\sigma \otimes \sigma \rightarrow \begin{cases} \mathbb{1} \otimes \mathbb{1} & \text{with probability } 0 \\ \sigma \otimes \mathbb{1} & \text{with probability } p \\ \mathbb{1} \otimes \sigma \text{ or } \sigma \otimes \sigma & \text{with probability } 1 - p \end{cases} \quad (62)$$

Similarly, we find that under a Haar random unitary transformation  $\sigma \otimes \sigma$  maps to

$$\sigma \otimes \sigma \rightarrow \begin{cases} \mathbb{1} \otimes \mathbb{1} & \text{with probability } 0 \\ \sigma \otimes \mathbb{1} & \text{with probability } p \\ \mathbb{1} \otimes \sigma \text{ or } \sigma \otimes \sigma & \text{with probability } 1 - p \end{cases} \quad (63)$$

Note that we average over Haar random unitaries acting on two sites, so it does not matter whether the initial Pauli has the identity on the left or on the right site, only an identity on both sites is important. We hence find that

$$\rho_R(s, \tau + 1) = p [\rho_R(s, \tau) + \rho_R(s + 1, \tau)] \quad (64)$$

$$\rho_R(s + 1, \tau + 1) = (1 - p) [\rho_R(s, \tau) + \rho_R(s + 1, \tau)] \quad (65)$$

In order to account for the two-layer update and the parity effect of the gates, we can define an averaged operator density for a rescaled time

$$\rho_R(x, t) = \rho_R(s = 2x - 1, \tau = 2t) + \rho_R(s = 2x, \tau = 2t). \quad (66)$$

The full update equation then reads

$$\rho_R(x, t + 1) = 2p(1 - p)\rho_R(x, t) + (1 - p)^2\rho_R(x - 1, t) + p^2\rho_R(x + 1, t). \quad (67)$$

The right edge can either move to the right with probability  $(1 - p)^2$ , it can move to the left with probability  $p^2$ , or it can stay put with probability  $2p(1 - p)$ . Since  $p < 1/2$  this corresponds to a biased random walk to the right: the operator tends to grow in time.

We can define a so-called butterfly velocity  $v_B$  and a diffusion constant  $D$  as

$$v_B = (1 - p)^2 - p^2 = \frac{q^2 - 1}{q^2 + 1}, \quad D = \frac{(1 - p)^2 + p^2}{2} = 1 - \frac{2q^2}{(q^2 + 1)^2}, \quad (68)$$

and rewrite the discrete update equation as

$$\begin{aligned} \rho_R(x, t + 1) - \rho_R(x, t) = & -v_B \left[ \frac{\rho_R(x + 1, t) - \rho_R(x - 1, t)}{2} \right] \\ & + D [\rho_R(x + 1, t) + \rho_R(x - 1, t) - 2\rho_R(x, t)]. \end{aligned} \quad (69)$$

Taking the appropriate continuum limits, we recover the *Fokker-Planck equation*

$$\partial_t \rho_R = -v_B \partial_x \rho_R + D \partial_x^2 \rho_R. \quad (70)$$

This equation doesn't just describe a biased random walk but appears in many different contexts, e.g. the evolution of a particle's velocity when undergoing Brownian motion and experiencing drag forces and random forces. Note again that all hydrodynamic equations reflect some underlying (local) conservation law, which is here the conservation of the operator norm, which is also referred to as conservation of quantum information.

In our specific setup, we can reduce the problem to a combinatoric one due to its origin as a biased random walk. Suppose that the boundary has taken  $u$  moves to the right and  $v$  moves to the left. Then the total number of time steps equals  $u + v = 2t$  and the boundary is located at position  $u - v = 2x$ . The probability of this process occurring follows as

$$P(u, v) = \binom{u+v}{u} (1-p)^u p^v = \binom{2t}{t+x} \frac{q^{2(t+x)}}{(q^2+1)^{2t}} = \rho_R(x, t). \quad (71)$$

The conservation of operator norm at time  $t$  can be directly checked since

$$\sum_{u+v=2t} P(u, v) = \sum_{u=0}^{2t} \binom{2t}{u} (1-p)^u p^{2t-u} = [(1-p) + p]^{2t} = 1. \quad (72)$$

We can obtain a closed-form expression for the profile of the operator density by applying Stirling's approximation to Eq. (71) for  $x = v_B t + \mathcal{O}(\sqrt{t})$ . To very good approximation, we then find that

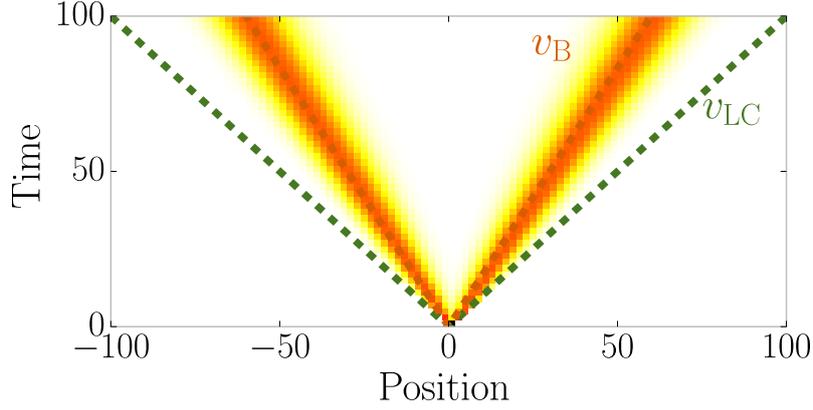
$$\rho_R(x, t) \approx \frac{1}{\sqrt{\pi(1-v_B^2)t}} \exp \left[ -\frac{(x - v_B t)^2}{(1-v_B^2)t} \right]. \quad (73)$$

We find that the operator density exhibits a Gaussian profile and is concentrated on a light ray  $x = v_B t$ , such that the operator can be said to grow linearly in time. This butterfly velocity  $v_B = (q^2 - 1)/(q^2 + 1)$  is smaller than the maximally allowed light-cone velocity  $v_{LC} = 1$  imposed by the geometry. Only in the limit  $q \rightarrow \infty$  do these two velocities coincide, since then the probability  $p$  of obtaining a single-site Pauli after a unitary transformation goes to zero and the Pauli edge hops to the right at each time step. This ballistic propagation of the operator is accompanied by a diffusive broadening of the front, where the width of the operator density around the front grows as  $\sqrt{t}$ . In the limit  $q \rightarrow \infty$  the profile becomes sharp. The general operator density profile is illustrated in Fig. 3. Note that this Figure illustrates the outer edges of the Pauli strings: these Pauli strings, constituting the operator dynamics, have a full support ranging from the left edge to the right edge and hence grow linearly in time, indicating operator scrambling.

## 5.2 Out-of-time-order correlation functions

The operator density is not directly experimentally accessible. In order to measure the support of a time-evolved operator  $\sigma_0^\alpha(t)$  acting as  $\sigma^\alpha$  on site  $x = 0$  at time  $t = 0$ , we can consider the commutator between the time-evolved operator and an operator  $\sigma_y^\beta$  acting as  $\sigma^\beta$  on site  $y$ , which we choose to be to the right of  $x = 0$ . For two initial operators that are sufficiently far apart we expect these to commute, and the commutator will only be nontrivial once the support of the initial operator has grown sufficiently. We have that

$$[\sigma_0^\alpha(t), \sigma_y^\beta] = \sum_S C_S(t) [\mathcal{S}, \sigma_y^\beta] = \sum_{\mathcal{S} \text{ ends at } y' \geq y} C_S(t) [\mathcal{S}, \sigma_y^\beta], \quad (74)$$



**Figure 3:** Illustration of the renormalized operator density  $(\rho_R(x, t) + \rho_L(x, t))\sqrt{t}$  as a function of position  $x$  and time  $t$ . Reproduced from Ref. [1].

where we have expanded  $\sigma_0^\alpha(t) = \sum_S C_S(t) S$  and have made use of the fact that the only non-commuting Pauli strings necessarily end at  $y' \geq y$ . In order to remove any sign-dependence and recover the operator density we can take the operator norm of the commutator, i.e.

$$\text{Tr}\left([\sigma_0^\alpha(t), \sigma_y^\beta]^\dagger [\sigma_0^\alpha(t), \sigma_y^\beta]\right) \propto \sum_{\substack{S \text{ ends at } y' \geq y}} |C_S(t)|^2 = \sum_{y' \geq y} \rho_R(y', t), \quad (75)$$

where we used that the commutator of a Pauli string with a Pauli matrix again returns a Pauli string and that under Haar random evolution all Pauli matrices are equally likely, with the commutator nonzero whenever the Pauli matrix on site  $y$  is different from  $\sigma^\beta$ .

In the continuum limit, we can replace the summation by an integral and express the profile as

$$\int_y^\infty dy' \rho_R(y', t) = 1 - \int_{-\infty}^y dy' \rho_R(y', t) \propto 1 - \text{erf}\left(\frac{x - v_B t}{(1 - v_B^2)t}\right). \quad (76)$$

The operator norm of the commutator inherits the ballistic spreading and diffusive front of the operator density. This operator norm can be recast as a so-called out-of-time-order correlation function (OTOC) by making use of the fact that the Pauli matrices are unitary and hermitian. Expanding the squared commutator from the operator norm, we find that

$$\begin{aligned} [\sigma_0^\alpha(t), \sigma_y^\beta]^\dagger [\sigma_0^\alpha(t), \sigma_y^\beta] &= (\sigma_y^\beta \sigma_0^\alpha(t) - \sigma_0^\alpha(t) \sigma_y^\beta) (\sigma_0^\alpha(t) \sigma_y^\beta - \sigma_y^\beta \sigma_0^\alpha(t)) \\ &= 2 \times \mathbb{1} - \sigma_y^\beta \sigma_0^\alpha(t) \sigma_y^\beta \sigma_0^\alpha(t) - \sigma_0^\alpha(t) \sigma_y^\beta \sigma_0^\alpha(t) \sigma_y^\beta, \end{aligned} \quad (77)$$

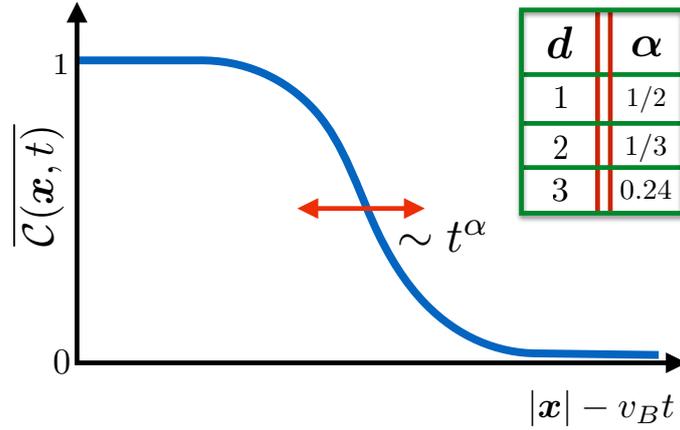
such that the trace norm returns

$$\text{Tr}\left([\sigma_0^\alpha(t), \sigma_y^\beta]^\dagger [\sigma_0^\alpha(t), \sigma_y^\beta]\right) / \mathcal{D} \propto 2 - \frac{2}{\mathcal{D}} \text{Tr}\left[\sigma_0^\alpha(t) \sigma_y^\beta \sigma_0^\alpha(t) \sigma_y^\beta\right]. \quad (78)$$

The nontrivial term in the right-hand side is known as the OTOC, and can be explicitly written out as

$$\text{Tr}\left[\sigma_0^\alpha(t) \sigma_y^\beta \sigma_0^\alpha(t) \sigma_y^\beta\right] = \text{Tr}\left[U(t) \sigma_0^\alpha U(t)^\dagger \sigma_y^\beta U(t) \sigma_0^\alpha U(t)^\dagger \sigma_y^\beta\right]. \quad (79)$$

The appearance of alternating terms  $U(t)$  and  $U(t)^\dagger$  results in a forward and a backward time evolution, leading to the term ‘out-of-time-order’ in out-of-time-order correlations. The OTOC profile is illustrated in Fig. 4. While the OTOC was initially proposed in a different context as



**Figure 4:** Illustration of the OTOC profile in different dimensions  $d$ , where the ballistic propagation with a velocity  $v_B$  is accompanied by a front widening as  $t^\alpha$ , where the exponent  $\alpha$  depends on the dimension and reproduces diffusive spreading with  $\alpha = 1/2$  in a one-dimensional lattice ( $d = 1$ ). Reproduced from Ref. [2].

a probe of quantum chaos, it has since been realized that the OTOC should be thought of as quantifying operator scrambling. The OTOC can be considered for both Hamiltonian dynamics and unitary circuit dynamics. However, while the forward and backward time evolution is difficult (but not necessarily impossible) to implement in Hamiltonian dynamics, in unitary circuit dynamics both circuits can be directly implemented, since these only corresponds to different choices of unitary gates. The OTOC and operator spreading have been directly probed in current quantum computing setups. As one example, in X. Mi *et al.*, “Information scrambling in quantum circuits”, *Science* **374**, 1479–1483 (2021), the OTOC was directly measured for different classes of unitary circuit dynamics in Google’s Sycamore quantum processor, observing the butterfly velocity and broadening of the operator front.

### 5.3 Entanglement dynamics

So far our discussion was focused on the dynamics of operators, but the dynamics of initial states can similarly be considered in a straightforward manner. Operator spreading here results in e.g. thermalization and the growth of entanglement as encoded in the reduced density matrix. Consider an initial wave function  $|\psi(t=0)\rangle = |1\rangle \otimes |1\rangle \otimes \dots \otimes |1\rangle$  and dynamics generated by Haar random circuits. The initial choice of product wave function is immaterial due to the Haar random dynamics, it only matters that the initial wave function is an untangled product wave function. The reduced density matrix follows as

$$\rho_A(t) = \text{Tr}_{\bar{A}}[\rho(t)], \quad \rho(t) = U(t)|\psi(t=0)\rangle \langle\psi(t=0)| U(t)^\dagger. \quad (80)$$

(Note that in this subsection  $\rho_A(t)$  denotes a reduced density matrix and not the operator density.)

We can use our results on operator spreading by writing

$$\rho(t=0) = |1\rangle \langle 1| \otimes |1\rangle \langle 1| \otimes \dots \otimes |1\rangle \langle 1| \quad \text{and} \quad |1\rangle \langle 1| = \frac{1}{2}(\mathbb{1} + \sigma^z). \quad (81)$$

As such, the initial density matrix for a system of  $N$  sites can be written as

$$\rho(t=0) = \frac{1}{2^N} \sum_{\alpha \in \{0,z\}} \sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \dots \sigma_N^{\alpha_N} \equiv \frac{1}{2^N} \sum_{\nu \in z\text{-strings}} S^\nu, \quad (82)$$

and hence

$$\rho(t) = \frac{1}{2^N} \sum_{\alpha \in \{0, z\}} \sigma_1^{\alpha_1}(t) \sigma_2^{\alpha_2}(t) \cdots \sigma_N^{\alpha_N}(t) \equiv \frac{1}{2^N} \sum_{\nu \in z\text{-strings}} \mathcal{S}^\nu(t). \quad (83)$$

The reduced density matrix for a subsystem  $A$  follows by tracing out the complement of  $A$ , such that we have

$$\rho_A(t) = \frac{1}{2^N} \sum_{\nu \in z\text{-strings}} \text{Tr}_{\bar{A}}[\mathcal{S}^\nu(t)]. \quad (84)$$

In order to quantify the entanglement, we could consider the second Rényi entropy,

$$S^{(2)}(t) = -\log \text{Tr}[\rho_A(t)^2]. \quad (85)$$

However, in order to use our results on operator dynamics, it is more straightforward to consider the exponential of the second Rényi entropy, i.e. the purity  $\text{Tr}[\rho_A(t)^2]$ . We can then consider the logarithm of the averaged purity as a proxy for the averaged entanglement.

Again moving to an arbitrary  $q$ -dimensional local Hilbert space for generality<sup>3</sup>, we can write

$$\rho(t) = \frac{1}{q^N} \sum_{\nu \in z\text{-strings}} \mathcal{S}^\nu(t) = \frac{1}{q^N} \sum_{\nu \in z\text{-strings}} \sum_{\mu} C_\mu^\nu(t) S^\mu, \quad (86)$$

where we have expanded  $\mathcal{S}^\nu(t) = \sum_{\mu} C_\mu^\nu(t) S^\mu$  in Pauli strings in the same way as done in the previous subsection. Taking the trace over  $\bar{A}$  then returns a nonzero value for the Pauli strings  $S^\mu$  that only act on  $A$ , i.e.

$$\rho_A(t) = \frac{1}{q^N} \sum_{\nu \in z\text{-strings}} \sum_{\mu \in A} C_\mu^\nu(t) \text{Tr}_{\bar{A}}[S^\mu]. \quad (87)$$

The purity follows as

$$e^{-S^{(2)}(t)} = \text{Tr}[\rho_A(t)^2] = \frac{1}{q^{N_A}} \sum_{\nu, \nu' \in z\text{-strings}} \sum_{\mu \in A} C_\mu^\nu(t) C_\mu^{\nu'*}(t). \quad (88)$$

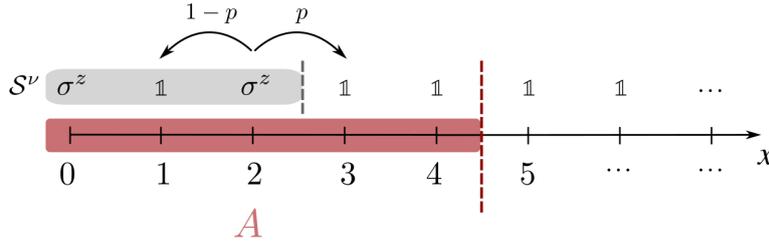
Introducing the Haar averaging leads to a vanishing of all off-diagonal terms, and the diagonal terms return the average purity as

$$\overline{e^{-S^{(2)}(t)}} = \frac{1}{q^{N_A}} \sum_{\nu \in z\text{-strings}} \sum_{\mu \in A} |C_\mu^\nu(t)|^2. \quad (89)$$

This expression can now be explicitly evaluated using our results for the averaged operator density. The problem again reduces to a classical combinatorics problem: consider an initial Pauli  $z$ -string  $\nu$  that is contained only within  $A$ . As time goes on, this operator spreads out in a superposition of Pauli strings  $\mu$ , and all Pauli strings that fall outside the support of  $A$  no longer contribute since they are traced out. For concreteness, we consider a subsystem of  $N_A$  consecutive sites ending at the left boundary of the lattice (see also Fig. 5). Using the random walk approach, we find that

$$\overline{e^{-S^{(2)}(t)}} = \frac{1}{q^{N_A}} + \frac{1}{q^{N_A}} \sum_{x=-t}^{N_A/2-1} \binom{2t}{t+x} \frac{q^{2t+2x}}{(q^2+1)^{2t}} \sum_{y=1}^{N_A/2-x} q^{2(y-1)}(q^2-1). \quad (90)$$

<sup>3</sup>Instead of  $\alpha \in \{0, z\}$ , the index  $\alpha$  can now take  $q$  different values labelling the diagonal generalized Pauli matrices in the  $z$ -string.



**Figure 5:** Illustration of how operator spreading relates to entanglement dynamics. Pauli  $z$ -strings  $S^\nu$  that initially act only within a subsystem  $A$  and hence contribute to the reduced density matrix  $\rho_A$  grow in time until they act nontrivially outside  $A$  and no longer contribute to the  $\rho_A$  since they are traced out. Once all Pauli strings that are not the identity have spread outside of  $A$  only the constant part  $\propto \mathbb{1}_A$  remains and the system is thermalized and maximally entangled.

The two summations have a direct interpretation: the first summation runs over the number of steps  $x$  to the right that the edge of the Pauli string has taken, where the summand is the corresponding probability, and the second summation counts all initial Pauli  $z$ -strings that end at site  $y \leq L_A/2 - x$ . Extending the second summation to start from  $y = 1 - x$  rather than  $y = 1$  induces an exponentially small error and allows this expression to be simplified to

$$\overline{e^{-S^{(2)}(t)}} \approx \frac{1}{q^{N_A}} + \left[ 1 - \frac{1}{q^{N_A}} \right] \frac{q^{2t}}{(1+q^2)^{2t}} \sum_{x=-t}^{N_A/2-1} \binom{2t}{t+x}. \quad (91)$$

While a closed form expression exists for the sum over binomial coefficients, it is not particularly transparent. However, two limits are already clear from this final expression, as also illustrated in Fig. 6. At time  $t = 0$  the purity equals one and the entropy vanishes, consistent with the initial unentangled product state. At late times the second summation vanishes and only the first term survives, and we have that  $S^{(2)} \approx N_A \log(q)$ . This result indicates a maximally entangled state with entanglement scaling with the size of  $A$ , leading to so-called volume law entanglement. In this limit all Pauli strings have effectively ‘leaked’ out of the subsystem, and we recover the infinite-temperature thermal reduced density matrix  $\rho_A \approx \mathbb{1}_A/q^{N_A}$ .

At short times the entanglement grows linearly, since the full summation over binomial coefficients can be evaluated to return  $2^{2t}$  and hence

$$\overline{e^{-S^{(2)}(t)}} \approx \left( \frac{2q}{1+q^2} \right)^{2t}, \quad (92)$$

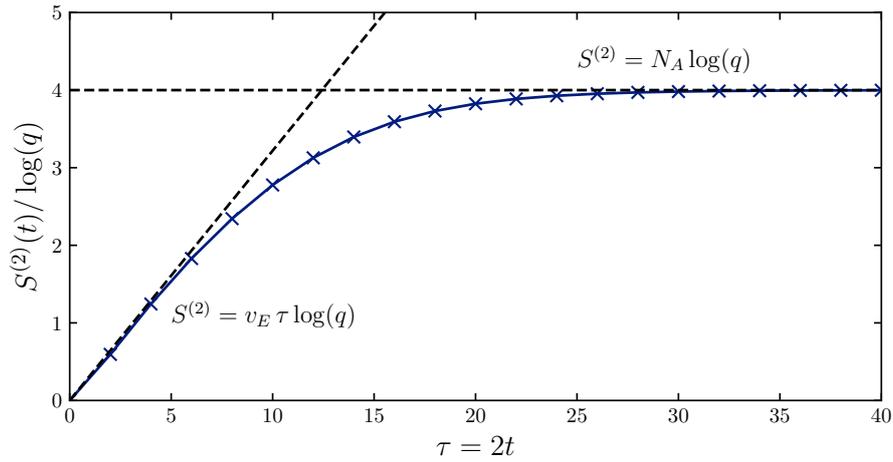
leading to a linear growth of the entanglement as

$$S^{(2)}(t) \approx 2t \log \left( \frac{q+q^{-1}}{2} \right). \quad (93)$$

The slope of the entanglement growth re-expressed in the original discrete time,  $\tau = 2t$ , and normalized to  $\log(q)$ , i.e. the unit of entanglement, is also known as the entanglement velocity  $v_E$ :

$$v_E = \log \left( \frac{q+q^{-1}}{2} \right) / \log(q). \quad (94)$$

The entanglement velocity has a general interpretation as the rate with which a maximally entangled region would need to grow in order to have entanglement  $S^{(2)}(t)$ . Note that this entanglement velocity approaches 1 in the limit of an infinitely large local Hilbert space  $q \rightarrow \infty$ ,



**Figure 6:** Entanglement dynamics for a subsystem of size  $N_A = 4$  under Haar-random unitary circuit dynamics and  $S^{(2)}(t)$  approximated from the averaged purity (91).

but does so logarithmically slowly. The entanglement velocity generically satisfies  $v_E < v_B$ , with the difference arising from the diffusive broadening of the operator front leading to a slower entanglement growth.

The two observed behaviors are universal: starting from an initial unentangled product state, the subsystem entanglement is expected to grow linearly with a (non-universal) entanglement velocity, before saturating at later times to a steady-state value exhibiting volume-law entanglement.

## References

- [1] C. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. Sondhi, “Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws,” *Phys. Rev. X* **8**, 021013 (2018).
- [2] A. Nahum, S. Vijay, and J. Haah, “Operator Spreading in Random Unitary Circuits,” *Phys. Rev. X* **8**, 021014 (2018).