Operator spreading in the memory matrix formalism

Mcculloch, Ewan; Von Keyserlingk, C W

Operator spreading in the memory matrix formalism

View the article online for updates and enhancements.

Operator spreading in the memory matrix formalism

Ewan McCulloch[*] and C W von Keyserlingk

University of Birmingham, Birmingham B15 2TT, United Kingdom

E-mail: ewan.r.mcculloch@gmail.com and c.vonkeyserlingk@bham.ac.uk

CrossMark

Abstract

The spread and scrambling of quantum information is a topic of considerable current interest. Numerous studies suggest that quantum information evolves according to hydrodynamical equations of motion, even though it is a starkly different quantity to better-known hydrodynamical variables such as charge and energy. In this work we show that the well-known memory matrix formalism for traditional hydrodynamics can be applied, with relatively little modification, to the question of operator growth in many-body quantum systems. On a conceptual level, this shores up the connection between information scrambling and hydrodynamics. At a practical level, it provides a framework for calculating quantities related to operator growth like the butterfly velocity and front diffusion constant, and for understanding how these quantities are constrained by microscopic symmetries. We apply this formalism to calculate operator-hydrodynamical coefficients perturbatively in a family of Floquet models. Our formalism allows us to identify the processes affecting information transport that arise from the spatiotemporal symmetries of the model.

Keywords: hydrodynamics, operator spreading, scrambling, thermalisation, Floquet, memory matrix

(Some figures may appear in colour only in the online journal)

[*]Author to whom any correspondence should be addressed.

Contents

1. Introduction

Under time evolution a wide class of many-body quantum systems tend towards equilibrium, where the final state is well described by a relatively small number of parameters such as temperature or pressure. At this point, information about the local conditions of the initial state is 'scrambled': it can no longer be determined by simple local measurements but is instead encoded in increasingly delicate and complicated observables. The process of scrambling has been the focus of intense study in the fields of black hole physics and holography [1–12], integrable systems [13–17], random unitary circuits [18–24], quantum field theories [25–32], and in the setting of chaotic spin-chains [33–43]. A principal reason for the flurry of interest in scrambling is the striking universality observed in the information spreading dynamics of apparently disparate models.

One universal feature of scrambling in ergodic systems is the ballistic growth of operators (with 'butterfly velocity' v_B), a feature connected to the universally observed linear growth of quantum entanglement. This result has been demonstrated for random unitary circuits [19]; in 1D a simple picture emerges where operators grow according to biased diffusion [19, 20]. This picture is altered in the presence of additional symmetries which give rise to power-law tails in the distributions of operator right end-points [21, 22]. In all of these cases, it appears that in ergodic systems quantum information obeys a sort of hydrodynamics, with an unusual conservation law, which we call 'information conservation'. The purpose of the present work is to show that the connection to hydrodynamics is not just an analogy and that a standard hydrodynamical tool—the so-called memory matrix formalism (MMF) [44]—can be applied with only a few technical (but consequential) modifications. Once we postulate a suitable slow manifold (a concept we explain), the ballistic growth of information is inevitable; in the same way that diffusion is inevitable in high temperature systems when the local conserved density is identified as the sole slow variable. Moreover, the formalism provides a framework for the perturbative calculation of information transport coefficients (e.g., the butterfly velocity) in concrete models. This is useful because operator spreading tends to require working at infinite temperature, where quantum field theoretic methods become harder to control. As such, the MMF is one of the only tools available for the analytical calculation of operator transport coefficients (although a similar effective membrane theory has been independently suggested for the related issue of entanglement growth [45]).

This work is organized as follows. In section 2 we extend the MMF to include a slow mode associated with the conservation of quantum information in the setting of translation invariant one-dimensional Hamiltonian systems and in section 3 we do the same for translation invariant one-dimensional Floquet models. Our formalism yields a succinct expression for the butterfly velocity, and shows that the biased diffusion of operator fronts observed in ergodic systems is arguably the simplest scenario consistent with the conservation laws. We also explain how microscopic symmetries constrain the butterfly velocity and front diffusion constants (D). We demonstrate the usefulness of the MMF in section 4 where we consider a translation invariant Floquet circuit with no additional symmetries and give results for the circuit averaged butterfly velocity and operator front diffusion constant in the limit of large local Hilbert space dimension

q. In section 6.2, we calculate the $\mathcal{O}(1/q^2)$ corrections to the butterfly velocity and attribute various contributions to the discrete time translation symmetry and to the spatial translation symmetry. Finally, inspired by this calculation we give predictions for v_B and D for a family of Floquet models to order $\mathcal{O}(1/q^2)$. This calculation also serves as a consistency check on our formalism, confirming that the slow manifold we have proposed is sufficiently complete to perform hydrodynamical calculations.

2. Memory matrix formalism: operator spreading as a slow mode

The MMF is a method for predicting the hydrodynamical properties of many-body systems [44, 46–48]. The input for the method is a Hamiltonian (or some other dynamics), as well as a guess as to what the likely slow modes are in the system (the local densities of conserved quantities are natural candidates). Under the assumption that all slow modes have been included and that all remaining fast modes decay sufficiently rapidly, the formalism yields predictions for the long-distance behavior of correlation functions involving the slow modes.

In this section, we adapt the formalism to include the slow mode associated with information conservation in the setting of one-dimensional quantum systems with local dynamics, i.e., local Floquet circuits, or systems with local Hamiltonians. In this paper, the 'conservation of quantum information' is equivalent to the statement that Heisenberg evolved operators have a conserved Hilbert–Schmidt norm $\mathrm{Tr}(O^\dagger(t)O(t)) = \mathrm{const.}$, a direct consequence of unitarity.

By averaging over choices of initial operator with the same right endpoint, we express the distribution of operator right endpoints (or the 'operator right density' as we refer to it in this paper), as an autocorrelation function of elements in the space of operators on two replicas of the Hilbert space. Using the MMF, we investigate the pole structure of the corresponding spectral function and give Kubo-like formula for the butterfly velocity v_B and an expression for diffusion constant D in terms of the memory matrix.

2.1. Quantifying operator spreading

We consider a one-dimensional lattice of N sites with single site Hilbert space $\mathcal{H}_{\mathrm{local}} = \mathbb{C}^q$. The space of operators on the full Hilbert space $\mathcal{H}$ is denoted $\mathcal{B}(\mathcal{H})$. A convenient basis for single site operators are the generalised Pauli matrices[1] $\{\sigma^\mu\}$, a set of unitary matrices satisfying the orthogonality relation $\mathrm{Tr}(\sigma^{\mu\dagger}\sigma^\nu)/q = \delta^{\mu,\nu}$. A time evolved operator $O(t)$ can be expressed as a linear combination of strings of generalised Pauli operators,

$$O(t) \equiv U^\dagger(t)OU(t) \equiv e^{itL}O = \sum_\mu C_\mu^O(t)\sigma^\mu. \tag{1}$$

The time evolution is generated by a Hamiltonian H, $U(t) = \exp(-itH)$. In the second equality we have introduced the Liouvillian $L(\cdot) = [H, \cdot],$[2] and in the last equality only the coefficients $C_\mu^O(t)$ depend on time. We have normalised O such that the Hilbert–Schmidt norm gives

---

[1] The generalised Pauli matrices are generated by the shift and clock matrices X and Z, $\sigma^\mu = X^{\mu^{(1)}}Z^{\mu^{(2)}}$. Where $X^q = Z^q = \mathbb{1}$ and $ZX = e^{\frac{2\pi i}{q}}XZ$.

[2] In the context of this paper, a Liouvillian L is a generator of Hamiltonian evolution (in operator space), as opposed to a generator of Markovian dynamics in open quantum systems, as the name often refers.

$\|O\|_{\text{HS}}^2 = \dim(\mathcal{H}) = q^N$, this ensures $\sum_\mu |C_\mu^O|^2 = 1$. Following the work of [19, 20], we define the *right density* $\rho_R(x, t)$ (the probability of an operator having right endpoint at position x at time t) by

$$\rho_R(x, t) \equiv \sum_\mu |C_\mu^O(t)|^2 \delta(\text{Rhs}(\mu) - x), \tag{2}$$

where $\text{Rhs}(\mu)$ denotes the rightmost site on which the Pauli string has non-trivial support. Summing over the site positions in equation (2) gives us a conserved quantity,

$$\sum_x \rho_R(x, t) = \sum_{x,\mu} |C_\mu^O(t)|^2 \delta(\text{Rhs}(\mu) - x) = \sum_\mu |C_\mu^O(t)|^2 = 1. \tag{3}$$

2.2. Operator right density as a correlation function

The starting point for the MMF is a temporal correlation function of slow variables, typically local charge operators associated to a conserved quantity such as energy or charge. Analogous to the local operators associated with energy or charge, we are able to associate pseudo-local operators $\hat{W}^x$ with the right density ρ_R. These operators, which we call the right density operators, will play the role of slow variables in the MMF. One complication is that, rather than being an operator on the original Hilbert space like charge/energy, $\hat{W}^x$ is an operator on operators—a *super-operator* (or equivalently an operator on two replicas of the original Hilbert space). Thus, in our application of MMF, we will be studying the dynamics of super-operators like W^x, eventually writing ρ_R as a temporal correlation function of 'vectorised' right density operators $|W^x\rangle$. This serves as the starting point for the MMF. $\hat{W}^x$ is implicitly defined in the equation for $\rho_R(x, t)$ below,

$$\rho_R(x, t) \equiv \langle O(t)|\hat{W}^x|O(t)\rangle, \quad \langle a|b\rangle \equiv \text{Tr}(a^\dagger b), \tag{4}$$

where the inner-product $\langle a|b\rangle$ is the infinite temperature operator inner-product suitable for $\mathcal{B}(\mathcal{H})$. An explicit definition of $\hat{W}^x$ is given by

$$\hat{W}^x \equiv \frac{1}{d_\mathcal{H}} \left(\bigotimes_{r \leq x-1} \hat{\Lambda}^+ \right) \left(\bigotimes_x \hat{\Lambda}^0 \right) \left(\bigotimes_{r \geq x+1} \hat{\Lambda}^- \right), \tag{5}$$

where $\hat{\Lambda}^+$ is the identity super-operator, $\hat{\Lambda}^-$ is the projector onto the identity operator and $\hat{\Lambda}^0$ is the projector onto the space of non-identity operators. Algebraically, $\hat{\Lambda}^\pm$ are given by

$$\hat{\Lambda}^+ \equiv \sum_\mu \frac{|\sigma^\mu\rangle\langle\sigma^\mu|}{q}, \qquad \hat{\Lambda}^- \equiv \frac{|\mathbb{1}\rangle\langle\mathbb{1}|}{q}, \qquad \hat{\Lambda}^0 \equiv \hat{\Lambda}^+ - \hat{\Lambda}^-. \tag{6}$$

By writing equtaion (4) in the language of tensor diagrams, we make the following manipulations to write the density $\rho_R(x, t)$ as an overlap of two 'vectors',

$$\langle O(t)|\hat{W}^x|O(t)\rangle \equiv \overbrace{O(t)^\dagger - \hat{W}^x - O(t)} \equiv \left. \begin{matrix} O(t)^\dagger - \hat{W}^x \\ O(t) \end{matrix} \right. \equiv \left. \begin{matrix} O(t)^\dagger \\ O(t) \end{matrix} \right| W^x . \tag{7}$$

We view the final expression as the inner-product of two vectors in the vector space $\mathcal{W} \equiv \mathcal{B}(\mathcal{H}) \boxtimes \mathcal{B}(\mathcal{H})$ of operators acting on two replicas of the Hilbert space (we have used a boxed

tensor product to represent the tensor product between two copies of the operator space). In particular, this allows us to write the density ρ_R as a temporal correlation function of the form $\langle A(t)|B\rangle_W$ as shown below,

$$\rho_R(x,t) \equiv \langle O(t)|\hat{W}^x|O(t)\rangle \equiv \langle O(t) \boxtimes O(t)^\dagger|W^x\rangle_W, \quad \langle A|B\rangle_W \equiv \mathrm{Tr}(A^\dagger B), \ A, B \in W. \quad (8)$$

Elements of W can be written as linear combinations of vectors $|A \boxtimes B\rangle$, which has the following diagrammatic representation,

$$|A \boxtimes B\rangle \equiv \begin{array}{c} \frac{1}{1} \\ \frac{2}{2} \end{array} \begin{array}{c} - A \\ - B \end{array} \Big] \Big], \quad (9)$$

where the *legs* $1, \bar{1}, 2, \bar{2}$, represent the indices of the operators A and B which act on the first and second replica respectively. An inner-product of two states has the obvious meaning of connecting legs,

$$\langle C \boxtimes D||A \boxtimes B\rangle \equiv \begin{array}{c} \lceil C^\dagger - A \rceil \\ \lfloor D^\dagger - B \rfloor \end{array} \equiv \ \mathrm{Tr}(C^\dagger A)\,\mathrm{Tr}(D^\dagger B) \equiv \langle C||A\rangle\langle D||B\rangle.$$

One also needs to include a rule for moving a symbol around a bend to ensure consistency,

$$\underline{\quad - A \ } \Big] \equiv \overline{\quad - A^T \ } \Big].$$

The manipulation of equation (7) has the following consequences for the $\hat{\Lambda}^\pm$ and $\hat{\Lambda}^0$,

$$\hat{\Lambda}^+ \rightarrow q\,|+\rangle \equiv \frac{1}{q}\sum_\mu |\sigma^\mu \boxtimes \sigma^{\mu\dagger}\rangle, \quad \hat{\Lambda}^- \rightarrow |-\rangle \equiv \frac{1}{q}\,|\mathbb{1} \boxtimes \mathbb{1}\rangle, \quad \hat{\Lambda}^0 \rightarrow q\,|0\rangle \equiv q\,|+\rangle - |-\rangle,$$

$$(10)$$

where we have chosen to normalise the vectors $|+\rangle$ and $|-\rangle$, $\langle\pm||\pm\rangle = 1$ (this introduces factors of q that differ from equation (6)). The vectors $|\pm\rangle$ take a simple diagrammatic form,

$$|+\rangle = \frac{1}{q} \begin{array}{c} \frac{1}{\bar{1}} \\ \frac{2}{\bar{2}} \end{array} \Big] \Big], \quad |-\rangle = \frac{1}{q} \begin{array}{c} \frac{1}{\bar{1}} \ \Big] \\ \frac{2}{\bar{2}} \ \Big] \end{array}. \quad (11)$$

The vectorised right density super-operators $|W^x\rangle$ are then given by

$$|W^x\rangle \equiv \frac{1}{d_{>x}} \left(\bigotimes_{r\leq x-1} |+\rangle_r \right) \left(\bigotimes_x |0\rangle_x \right) \left(\bigotimes_{r\geq x+1} |-\rangle_r \right), \quad (12)$$

where $d_{>x}$ is the dimensions of the subsystem of sites to the right of site x. A closely related super-operator is the *purity super-operator* $\hat{F}^x$, given by

$$\hat{F}^x \equiv \frac{1}{d_{\leq x}} \hat{\Lambda}^+_{\leq x} \otimes \hat{\Lambda}^-_{>x}, \quad |F^x\rangle \equiv \left(\bigotimes_{r \leq x} |+\rangle_r \right) \left(\bigotimes_{r > x} |-\rangle_r \right), \quad (13)$$

where $d_{\leq x}$ is the dimension of the subsystem of sites $r \leq x$. Using this super-operator, we are able to represent the purity $\gamma_{\leq x}(t)$ of the system (partitioned into a subsystem left (inclusive) and right of x) as a matrix element,

$$d_{\mathcal{H}} \langle \rho(t) | \hat{F}^x | \rho(t) \rangle d_{\mathcal{H}} \equiv \langle \rho(t)^\dagger \boxtimes \rho(t) | F^x \rangle_{\mathcal{W}} \equiv \mathrm{Tr}_{\leq x}(\rho_{\leq x}(t)^2) \equiv \gamma_{\leq x}(t), \quad (14)$$

where $\rho(t)$ is the density matrix for the state of the system. Up to an overall constant, the purity operators F^x and the integrated right density operators, $\sum_{y \leq x} W^y$, are equivalent. Therefore, the W^x can be written as a difference of (re-scaled) F^x's,

$$W^x = \frac{1}{d_{>x}} F^x - \frac{1}{d_{>x-1}} F^{x-1}. \quad (15)$$

Time evolution in the doubled operator space $\mathcal{W}$ is generated by a *doubled* Liouvillian $\mathcal{L} \equiv L \boxtimes \mathbb{1} + \mathbb{1} \boxtimes L$, which evolves both replicas of the operator space independently. The time derivative of $W^x(t) \equiv e^{i\mathcal{L}t}(W^x)$ is given by $\partial_t W^x(t) \equiv i\mathcal{L}(W^x(t))$. We can use this to write down a continuity equation for W^x,

$$\partial_t W^x + \Delta_x(J^x) = 0, \qquad \Delta_x(J^x) \equiv J^x - J^{x-1}, \quad (16)$$

where $J^x = -i\mathcal{L}(F^x)/d_{>x}$ is the current associated with the operator right density. The currents J^x are pseudo-local super-operators; they look, locally, like Λ^+ everywhere to the left the *cut* $\{x, x+1\}$ and Λ^- everywhere to the right. Local unitary evolution acts trivially in the $+$ and $-$ domains. We show this by considering the action of a (single-site) Liouvillian $\mathcal{L}_i$ at site i in the $+$ domain,

 $= 0.$ (17)

By moving the conjugated Hamiltonians around the 'bend' from the barred legs onto the unbarred legs, we find that the first and fourth terms cancel, as do the second and third. This can be easily generalised to l-local interactions. By swapping legs $1 \leftrightarrow 2$ in equation (17), we the find the action of $\mathcal{L}$ in the $-$ domain. The region separating the $+$ and $-$ domains can only grow via the local evolution at its edges. In this sense, one can consider equation (16) to be an equation of local conservation of operator right density.

2.3. Operator averaging

Rather than consider the evolution of the right density for a particular operator O, we average over the all operators with right endpoint x. Doing this in the basis of generalised Pauli matrices, the averaged density $\overline{\rho_R}(x, y, t)$ is given by

$$\overline{\rho_R}(x, y, t) \equiv \frac{q^2}{q^2 - 1} \frac{1}{d^2_{\leq x}} \sum_{O_x} \langle O_x^\dagger \boxtimes O_x | e^{-it\mathcal{L}} | W^y \rangle_{\mathcal{W}}, \quad (18)$$

where we have pulled the time evolution out from the operators $O(t)$. The factor before the sum is normalisation for the average (the reciprocal of the number of linearly independent operators with right endpoint x). This is in fact just a temporal-correlation function between right density super-operators,

$$\overline{\rho_{\mathrm{R}}}(x, y, t) = \frac{q^2 d_{>x}^2}{q^2 - 1} \langle W^x | \, \mathrm{e}^{-it\mathcal{L}} \, | W^y \rangle. \tag{19}$$

The W^x are orthogonal but not normalised with respect to the trace inner-product. The result of this is that the Fourier transformed right-densities W^k are not orthogonal. This lead us to an unusual inner-product $(\cdot | \cdot)$ with respect to which the W^x are orthonormal,

$$(A | B) \equiv \langle \Phi(A) | \, |B\rangle_{\mathcal{W}} = \mathrm{Tr}\left(\Phi(A)^\dagger B\right). \tag{20}$$

Where Φ is given by

$$\Phi \equiv \sum_x \frac{1}{\chi_x^2} |W^x\rangle \langle W^x| + Q, \quad \chi_x \equiv \langle W^x| \, |W^x\rangle = \frac{q^2 - 1}{q^2 d_{>x}}, \implies \Phi |W^x\rangle = \frac{1}{\chi_x} |W^x\rangle. \tag{21}$$

Q is the projector onto the *fast* subspace $\mathcal{Q}$, the orthogonal complement[3] of the *slow* subspace $\mathcal{P} \equiv \mathrm{Span}\{W^x\}$. A proof that $(\cdot | \cdot)$ satisfies the axioms of an inner-product is found in appendix A. We must be careful with expression equation (19) as $\mathcal{L}$ is not self-adjoint with respect to this inner-product. Making the operator average implicit, the right density becomes

$$\rho_{\mathrm{R}}(x, y, t) = \left(W^x | \mathrm{e}^{-it\mathcal{L}} | W^y\right). \tag{22}$$

The position dependent re-scaling of the right densities W^x by Φ reflects the strong entropic bias for operators to grow, i.e., for $y > x$ we find

$$\left(W^x | \mathrm{e}^{-it\mathcal{L}} | W^y\right) = q^{2(y-x)} \left(W^y | \mathrm{e}^{it\mathcal{L}} | W^x\right). \tag{23}$$

Operators are exponentially more likely to grow than to shrink[4].

2.4. Spectral and memory function

Having expressed the right density as a correlation function, we now investigate its late time behavior. A helpful diagnostic in the long time behaviour of a correlation function is the pole structure of the corresponding spectral function,

$$\rho_{x,y}(z) = \left(W^x \left| \frac{\mathrm{i}}{z - \mathcal{L}} \right| W^y\right) = \int_0^\infty \mathrm{e}^{\mathrm{i}zt} \left(W^x \left| \mathrm{e}^{-it\mathcal{L}} \right| W^y\right) \mathrm{d}t. \tag{24}$$

Fourier transforming in space will allow us to express the poles at small z in terms of a wave-number k. In particular this allows us to characterise the long-time and long-wavelength

---

[3] Orthogonal with respect to the inner-product $\langle \cdot | \, | \cdot \rangle$, or equivalently $(\cdot | \cdot)$.
[4] Evolution with $\mathcal{L}$ and $-\mathcal{L}$ can give rise to differing butterfly velocities (in section 2.5.2 we identify these as v_{R} and v_{L}, the right/left velocities).

behaviour of the averaged density $\rho_R(x, y, t)$. The density super-operators in k-space are given by

$$W^k = \frac{1}{\sqrt{N}} \sum_x e^{-ixk} W^x, \qquad \left(W^p | W^k \right) = \delta^{k,p}. \tag{25}$$

We choose to center the lattice at $x = 0$. In the k-space, the spectral function is given by

$$\rho_{k,p}(z) = \left(W^k \left| \frac{i}{z - \mathcal{L}} \right| W^p \right). \tag{26}$$

By taking $\mathcal{P} \equiv \mathrm{Span}\{W^k\}$ as the slow space in the MMF, we use some formal manipulations of resolvent operators [44] to express the spectral function as

$$\rho_{k,p}(z) = i \left[z\mathbb{1} - \Omega + i\Sigma(z) \right]^{-1}_{k,p}, \tag{27}$$

where $\Omega_{k,p} \equiv \left(W^k \left| \mathcal{L} \right| W^p \right)$ and $\Sigma(z)_{k,p}$ is the memory matrix,

$$\Sigma(z)_{k,p} = \left(W^k \left| \mathcal{L}Q \frac{i}{z - \mathcal{L}Q} \mathcal{L} \right| W^p \right). \tag{28}$$

By writing $[A, \cdot\,] = l_A - r_A$, where l_A (r_A) is left (right) multiplication by A, we can express the doubled Liouvillian as $\mathcal{L} = (l_H - r_H) \boxtimes \mathbb{1} + \mathbb{1} \boxtimes (l_H - r_H)$. It is simple to check using equation (13) and (11) that $\langle F^x | l_H \boxtimes \mathbb{1} | F^y \rangle = \langle F^x | r_H \boxtimes \mathbb{1} | F^y \rangle = q^{-|y-x|} \mathrm{Tr}(H)$ (and equivalently for multiplication by H in the second replica). This gives $\langle F^x | \mathcal{L} | F^y \rangle = 0$ and, by linearity, $\Omega_{k,p} = 0$.

For translationally invariant systems (in the thermodynamic limit $N \to \infty$), the memory matrix is diagonal, $\Sigma(z)_{k,p} = \Sigma(k, z)\delta^{k,p}$ and the inverse of $z\mathbb{1} + i\Sigma(k, z)$ is readily calculated,

$$\rho_{k,p}(z) = \rho(k, z)\delta^{k,p}, \qquad \rho(z, k) = \frac{i}{z + i\Sigma(z, k)}. \tag{29}$$

The long-time and long-wavelength pole structure is found by expanding the memory function $\Sigma(z, k)$, for small k and z. The k-space representation of the continuity equation equation (16) is then given by,

$$\partial_t W^k \equiv i\mathcal{L}\left(W^k \right) = -(1 - e^{-ik}) J^k, \tag{30}$$

where $J^k = \frac{1}{\sqrt{N}} \sum_x e^{-ikx} J^x$. Putting this back into the memory matrix yields

$$\Sigma(k, z) = -i(1 - e^{-ik}) \left(W^k \left| \mathcal{L}Q \frac{i}{z - \mathcal{L}Q} \right| J^k \right) \tag{31}$$

$$= -v(z)\,ik - b(z)\,k^2 + \cdots. \tag{32}$$

This will give the pole structure

$$\rho(k, z) \sim \frac{i}{z - v(z)\,k + ib(z)\,k^2 + \cdots}. \tag{33}$$

Provided the analyticity of $v(z)$ and $b(z)$ as $-iz \to 0^+$, this will be precisely the pole structure associated with a biased diffusion equation. This analyticity condition is met so long as we can assume that the fast variables J^k and $\mathcal{L}\Phi(W^k)$ have rapidly decaying correlations (faster than $1/t$).

2.5. Butterfly velocity and diffusion constant

In this section we provide formal expression for the butterfly velocity v_B and diffusion constant D and sufficient conditions for the symmetry of the operator growth light-cone, $v_L = v_R$ and $D_L = D_R$.

2.5.1. Formal expressions for v_B and D. Using equation (32), the butterfly velocity v_B is given by

$$v_B = \lim_{z \to i0^+} v(z), \qquad v(z) = \lim_{k \to 0} -i\partial_k \Sigma(k,z). \tag{34}$$

We introduce the proxy $\sigma(k,z)$, defined as

$$\sigma(k,z) = \left(W^k \left| \mathcal{L} \frac{i}{z - \mathcal{L}} \mathcal{L} \right| W^k \right), \qquad \Sigma(k,z) = \frac{\sigma(k,z)}{1 + \sigma(k,z)/z}. \tag{35}$$

Using $\mathcal{L}(W^k) \sim k$, we conclude that $\lim_{k \to 0} \sigma/k = \lim_{k \to 0} \Sigma/k$, provided that we take the $k \to 0$ limit before taking $z \to i0^+$. Using this, we give a Kubo-like formula for v_B,

$$v_B = i \lim_{s \to 0} \int_0^\infty dt \ e^{-st} \left(W \left| \mathcal{L} \right| J(-t) \right), \tag{36}$$

where $J \equiv J^{k=0}$ and $W \equiv W^{k=0}$. Converting to the usual trace inner-product, we have $\left(W \left| \mathcal{L} \right| J(-t) \right) = \langle \mathcal{L}\Phi(W) | J(-t) \rangle$. Importantly, Φ and $\mathcal{L}$ do not in general commute[5]. If they were to commute, we would find $\mathcal{L}(W^{k=0}) = 0$ and hence $v_B = 0$, which would lead to the incorrect conclusion that information propagates diffusively rather than ballistically. Using the biased diffusion ansatz for the pole location,

$$z = v_B k - iDk^2 + \mathcal{O}(k^3), \tag{37}$$

the diffusion constant is given by

$$D = \lim_{z \to i0^+} \lim_{k \to 0} \left(v_B \partial_z \partial_k \Sigma + \frac{1}{2} \partial_k^2 \Sigma \right). \tag{38}$$

2.5.2. $v_R \neq v_L$ and the operator growth light-cone. Generically, the right and left butterfly velocities, v_R and v_L, are not equal [42, 49, 50]. In this section we relate v_R and v_L in a way that gives rise to a light-cone structure. We then determine sufficient conditions for a symmetric light-cone, $v_R = v_L$ and symmetric fronts $D_L = D_R$. We need to adapt our notation when talking about both left and left and right density distributions at once, to do this we denote W_R^x as the familiar right density super-operators and W_L^x to be the left density super-operator. The altered inner-product is also right/left dependent, with $(\cdot|\cdot)_R$ $((\cdot|\cdot)_L)$ corresponding to the inner-productive suitable for right (left) endpoint calculations. The right and left density distributions are given by

$$\rho_R(x,t) = \left(W_R^0 \left| e^{-it\mathcal{L}} \right| W_R^x \right)_R, \qquad \rho_L(x,t) = \left(W_L^0 \left| e^{-it\mathcal{L}} \right| W_L^x \right)_L. \tag{39}$$

The equations for v_L and D_L are found to be

$$v_L^H = \lim_{z \to i0^+} \lim_{k \to 0} i\partial_k \Sigma_L^H(k,z), \quad D_L^H = \lim_{z \to i0^+} \lim_{k \to 0} \left(-v_L \partial_z \partial_k + \frac{1}{2} \partial_k^2 \right) \Sigma_L^H(k,z). \tag{40}$$

---

[5] An example where $[\Phi, \mathcal{L}] = 0$ is when the Hamiltonian does not couple any sites, i.e., $v_B = 0$.

Where we have introduced a superscript H to label which Hamiltonian the systems is being evolved with. In order to convert between these distributions let us introduce the involution $\mathcal{I}$, the spatial inversion symmetry operation, $x \to -x$. It has the following action on the density super-operators and the Liouvillian,

$$\mathcal{I}\left|W_{\mathrm{L}}^x\right\rangle = \left|W_{\mathrm{R}}^{-x}\right\rangle, \qquad \mathcal{I}\mathcal{L}_H\mathcal{I} = \mathcal{L}_{H_{\mathrm{I}}}, \tag{41}$$

where H_{I} is the spatially inversion of the (translationally invariant) Hamiltonian H. Using this involution on ρ_{L} gives the following

$$\rho_L^H(x,t) = \left(W_{\mathrm{L}}^0\left|\mathrm{e}^{-\mathrm{i}t\mathcal{L}}\right|W_{\mathrm{L}}^x\right)_{\mathrm{L}} = \frac{1}{\chi_0}\left\langle W_{\mathrm{L}}^0\right|\mathcal{I}^2\mathrm{e}^{-\mathrm{i}t\mathcal{L}}\mathcal{I}^2\left|W_{\mathrm{L}}^x\right\rangle = \frac{1}{\chi_0}\left\langle W_{\mathrm{R}}^0\right|\mathrm{e}^{-\mathrm{i}t\mathcal{L}_{\mathrm{I}}}\left|W_{\mathrm{R}}^{-x}\right\rangle$$

$$= \left(W_{\mathrm{R}}^0\left|\mathrm{e}^{-\mathrm{i}t\mathcal{L}_{\mathrm{I}}}\right|W_{\mathrm{R}}^{-x}\right)_{\mathrm{R}} = \rho_{\mathrm{R}}^{H_{\mathrm{I}}}(-x,t), \tag{42}$$

where $\chi_0 = \left\langle W_{\mathrm{R}}^0\right|\left|W_{\mathrm{R}}^0\right\rangle = \left\langle W_{\mathrm{L}}^0\right|\left|W_{\mathrm{L}}^0\right\rangle$. We have added labels to make clear under which Hamiltonian the system is being evolved. This relation implies the following,

$$\Sigma_L^H(k,z) = \Sigma_{\mathrm{R}}^{H_{\mathrm{I}}}(-k,z). \tag{43}$$

Using this and expressions for $v_{\mathrm{R/L}}$ and $D_{\mathrm{R/L}}$ (equations (34), (38) and (40)), we find the relationship between v_{L} and v_{R} and D_{L} and D_{R} to be

$$v_{\mathrm{L}}^{(H)} \equiv v_{\mathrm{R}}^{(H_{\mathrm{I}})}, \qquad D_{\mathrm{L}}^{(H)} \equiv D_{\mathrm{R}}^{(H_{\mathrm{I}})}. \tag{44}$$

This implies a physically obvious result: in an inversion symmetric system the left and right velocities are equal. A somewhat less obvious result is the following:

Remark. For a translationally invariant Hamiltonian H, if there exists a transformation $\mathcal{R}$ that performs single site basis rotations, such that $\mathcal{R}^\dagger H\mathcal{R} = -H$ or $\mathcal{R}^\dagger H\mathcal{R} = H^*$, the operator growth light-cone is symmetric, $v_{\mathrm{L}} = v_{\mathrm{R}}$ and $D_{\mathrm{L}} = D_{\mathrm{R}}$.

The second of these sufficient conditions applies for systems with a time-reversal/anti-unitary symmetry, provided that the symmetry transformation can be achieved with single site transformations. Before we prove this result, we will apply it to several models:

- An example where both sufficient conditions are met is the spin-1/2 model $H = -\sum_{\langle i,j\rangle} JY_iZ_j - \sum_i hX_i$. The choice of $\mathcal{R} = Y^{\otimes N}$ gives $\{\mathcal{R}, H\} = 0$. Alternatively, if we had chosen $\mathcal{R}$ to be a product of Pauli X operators on every even site and $\mathbb{1}$ on every odd site, we find $\mathcal{R}^\dagger H\mathcal{R} = H^*$, (a time-reversal symmetry). Despite this model lacking inversion symmetry, it has a symmetric operator light-cone, $v_{\mathrm{L}} = v_{\mathrm{R}}$ and $D_{\mathrm{L}} = D_{\mathrm{R}}$.
- An example where the second of the sufficient conditions is met is given by another spin-1/2 model with two-body interactions and no external field coupling, $H = -\sum_{i,j}\sum_{\alpha,\beta} J_{i,j}^{\alpha,\beta}\sigma_i^\alpha\sigma_j^\beta$. By choosing $\mathcal{R} = Y^{\otimes N}$, only $\sigma_i^\alpha\sigma_j^\beta$ terms containing a single Y are sent to their negatives, every other term remains unchanged. The terms that flipped sign make up the imaginary part of H, hence $H \to H^*$, so that yet again $v_{\mathrm{L}} = v_{\mathrm{R}}$.
- Taking the integrable and non-integrable Hamiltonians of [42], in the special case $\lambda = 0$ the Hamiltonians lack inversion symmetry but satisfy both of the sufficient conditions above, giving $v_{\mathrm{L}} = v_{\mathrm{R}}$ in agreement with the results in [42].

To prove this result we make use of the symmetry operation $S = (1 \leftrightarrow 2)$, which swaps the legs 1 and 2 on every site. S has the property $S |W_R^x\rangle = |W_L^x\rangle$ and is a symmetry of $\mathcal{L}$. Therefore,

$$
\begin{aligned}
\rho_L^H(x,t) &= \left(W_L^0 \left| S^2 e^{-it\mathcal{L}_H} S^2 \right| W_L^x \right)_L = \frac{1}{\chi_0} \left\langle W_L^0 \right| S e^{-it\mathcal{L}_H} S \left| W_L^x \right\rangle \\
&= \frac{1}{\chi_0} \left\langle W_R^x \right| e^{-it\mathcal{L}_H} \left| W_R^0 \right\rangle = \frac{1}{\chi_0} \left\langle W_R^0 \right| e^{it\mathcal{L}_H} \left| W_R^{-x} \right\rangle^* \\
&= \left(W_R^0 \left| e^{it\mathcal{L}_H} \right| W_R^{-x} \right)_R^* = \rho_R^{-H}(-x,t),
\end{aligned}
\tag{45}
$$

where we have used the reality of the density. Using this gives

$$
\Sigma_L^H(k,z) = \Sigma_R^{-H}(-k,z) \implies v_L^{(H)} \equiv v_R^{(-H)}, \qquad D_L^{(H)} \equiv D_R^{(-H)}.
\tag{46}
$$

This implies the existence of a light cone structure, where the future light cone is a π rotation of the past light cone. If, in the final equality of equation (45), we had brought the complex conjugation onto each term, we would have instead found

$$
\Sigma_L^H(k,z) = \Sigma_R^{H^*}(-k,z) \implies v_L^{(H)} \equiv v_R^{(H^*)}, \qquad D_L^{(H)} \equiv D_R^{(H^*)}.
\tag{47}
$$

We can generalise this by considering any transformation which performs on-site basis rotations, such a transformation is an isometry of the density super-operators, $|W^x\rangle$. Letting $\mathcal{R} = \prod_x R_x$ be a product of unitary rotations, we can freely conjugate H by $\mathcal{R}$ at any point in equation (45) (on-site rotation is an isometry of W^x) and maintain equality. This gives

$$
v_L^H \equiv v_R^{(\mathcal{R}^\dagger H \mathcal{R})^*} \equiv v_R^{-\mathcal{R}^\dagger H \mathcal{R}}, \qquad D_L^H \equiv D_R^{(\mathcal{R}^\dagger H \mathcal{R})^*} \equiv D_R^{-\mathcal{R}^\dagger H \mathcal{R}}.
\tag{48}
$$

3. MMF for Floquet models

With only minor changes, the MMF generalises to Floquet models. In this section we repeat a number of steps taken in the continuous time case, finding Floquet analogues of the memory matrix and Ω. The main results of this section are equation (56) for the Floquet analogue of Σ and Ω and formal expressions equation (58) for v_B and D (and a Kubo-like formula for v_B in equation (61)). Readers only interested in main results should look at these equations and skip the rest of the section.

Let U be a Floquet unitary, then the Heisenberg evolution of an operator O is given by

$$
O(n) = U_\varepsilon^{-n} O U_\varepsilon^n.
\tag{49}
$$

Writing $U^{\mathrm{ad}} = l_U r_{U^{-1}}$, where l_U (r_U) is left (right) multiplication by U, allows us to rewrite the operator evolution as

$$
O(n) = (U^{\mathrm{ad}})^{-n}(O).
\tag{50}
$$

Evolution of elements in the doubled operator space $\mathcal{W}$ is given by

$$
(A \boxtimes B)(n) = (U^{\mathrm{ad}} \boxtimes U^{\mathrm{ad}})^{-n}(A \boxtimes B).
\tag{51}
$$

Denoting $\mathcal{U} = U^{\mathrm{ad}} \boxtimes U^{\mathrm{ad}}$, the evolution of an element $X \in \mathcal{W}$ is given compactly by

$$X(n) = \mathcal{U}^{-n}(X). \tag{52}$$

Using equation (52), we define the discrete time derivative $\Delta_t X(t) \equiv X(t) - X(t-1) = -\mathcal{L}X(t)$, where $\mathcal{L} = \mathcal{U} - \mathbb{1}$. The Floquet analogue of the continuous time continuity equation equation (16) is then given below, where we have analogously defined a current J,

$$\Delta_t W^k(t) = -(1 - e^{-ik})J^k(t), \qquad J^x \equiv \frac{1}{d_{>x}}\mathcal{L}(F^x), \tag{53}$$

Following the steps in section 2.1, the operator averaged right density is given by

$$\rho(x, y, n) = \left(W^x \left| \mathcal{U}^n \right| W^y \right). \tag{54}$$

For translationally invariant Floquet models, the spectral function $\rho(z, k)$ is given by equation (55), in analogy to equation (29) in the Hamiltonian case,

$$\rho(k, z) = \frac{1}{1 - e^{iz}(1 + \Omega + \Sigma)}. \tag{55}$$

Ω and the memory matrix Σ are given, in analogy to equation (28), by

$$\Omega(k) = \left(W^k \left| \mathcal{L} \right| W^k \right), \qquad \Sigma(k, z) = \left(W^k \left| \mathcal{L}Q \frac{1}{e^{-iz} - 1 - \mathcal{L}Q} \mathcal{L} \right| W^k \right). \tag{56}$$

Unlike in the Hamiltonian case, Ω does not in general vanish, as $\mathcal{L} \equiv \mathcal{U} - \mathbb{1}$ is not a (doubled) Liouvillian. The pole of the spectral function encodes the long time/distance hydrodynamical limit, and is given by equation (57),

$$z = i \log(1 + \Omega(k) + \Sigma(k, z)). \tag{57}$$

Using the biased diffusion ansatz equation (37), the butterfly velocity and operator front diffusion constants are formally given by

$$v_B = \lim_{z \to i0^+} \lim_{k \to 0} i\partial_k (\Omega + \Sigma), \qquad D = -\lim_{z \to i0^+} \lim_{k \to 0} \left(\partial_z \partial_k + \frac{1}{2}\partial_k^2 \right)(\Omega + \Sigma) - \frac{v_B^2}{2}. \tag{58}$$

As in the continuous time case, the memory function is difficult to calculate. Once again, it is convenient to define an auxiliary quantity σ, defined as

$$\sigma(k, z) \equiv \left(W^k \left| \mathcal{L}Q \frac{1}{e^{-iz} - 1 - \mathcal{L}} Q\mathcal{L} \right| W^k \right), \qquad \Sigma = \frac{\sigma}{1 + \frac{\sigma}{e^{-iz} - 1 - \Omega}}. \tag{59}$$

Using $\mathcal{L}(W^k) = i(1 - e^{-ik})J^k \sim k$, as in the continuous time case, we deduce both $\Sigma(k, z) \sim k$ and $\sigma(k, z) \sim k$ for small k. Once again giving

$$\lim_{k \to 0} \Sigma(k, z)/k = \lim_{k \to 0} \frac{\sigma/k}{1 + \frac{(\sigma/k)k}{e^{-iz} - 1 - \Omega}} = \lim_{k \to 0} \sigma(k, z)/k. \tag{60}$$

By using this equivalence and by splitting the current J it up into its slow and fast components $J_P^k = P(J^k)$ and $J_Q^k = Q(J^k)$, we find the Kubo-like formula for v_B, analogous to equation (36) found in the continuous time case,

$$v_B = -\left(W|J_P \right) - \lim_{s \to 0^+} \sum_{t=0}^{\infty} e^{-st} \left(W|\mathcal{L}Q|J_Q(-t) \right), \tag{61}$$

13
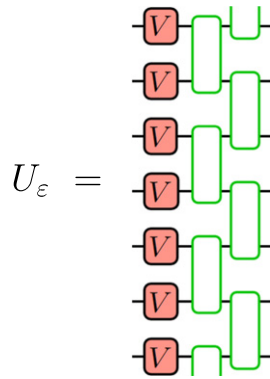
Figure 1. A Floquet layer with two-local gates followed by a layer of single-site scramblers *V*.

where $J = J^{k=0}$ and $W = W^{k=0}$. Although a slightly different object to the Liouvillian used in the continuous time case, it remains true that $\mathcal{L}$ does not commute with Φ. Leading again, inevitably, to the ballistic spreading of operators.

4. A minimal Floquet model

As a test of the formalism, we investigate a translationally invariant Floquet model with on-site random unitary scramblers. We take the single site Hilbert space to be dimension q and choose the Floquet unitary to be composed of a layer of nearest neighbour two-site unitary gates, followed by a layer of on-site Haar random scrambling unitary *V* applied to every site (figure 1), this ensures that the model has no additional conservation laws beyond those guaranteed by unitarity,

$$U_\varepsilon = V^{\otimes N} e^{-i\varepsilon H}, \qquad H = \sum_j Z_j Z_{j+1}. \tag{62}$$

The coupling unitary can be written as a product of commuting two-site unitaries $e^{-i\varepsilon H} = \prod_j e^{-i\varepsilon Z_j Z_{j+1}}$. A single gate $e^{-i\varepsilon Z_x Z_{x+1}} = \cos(\varepsilon)\mathbb{1}_x\mathbb{1}_{x+1} - i\sin(\varepsilon)Z_x Z_{x+1}$, straddling sites x and $x+1$, has diagrammatic representation

$$e^{-i\varepsilon Z_x Z_{x+1}} \equiv \quad \begin{matrix} x \\ x+1 \end{matrix} \quad = \cos(\varepsilon)\; \underline{\quad} - i\sin(\varepsilon)\; \substack{\bullet \\ \bullet}. \tag{63}$$

Where a black dot represent a Z operator,

$$\text{---}\bullet\text{---} = \text{---}Z\text{---}. \tag{64}$$

Likewise, the conjugate of the gate $e^{i\varepsilon Z_x Z_{x+1}}$ is given by

$$e^{i\varepsilon Z_x Z_{x+1}} \equiv \quad \boxed{} \quad = \cos(\varepsilon)\; \underline{\quad} + i\sin(\varepsilon)\; \substack{\bullet \\ \bullet}. \tag{65}$$

We write the Floquet unitary for the doubled operator space as $\mathcal{U} = \mathcal{U}_Z \mathcal{V}$, where $\mathcal{U}_Z = \mathrm{e}^{-\mathrm{i}\varepsilon H} \otimes \mathrm{e}^{\mathrm{i}\varepsilon H} \otimes \mathrm{e}^{-\mathrm{i}\varepsilon H} \otimes \mathrm{e}^{\mathrm{i}\varepsilon H}$ contains the two-local gates and $\mathcal{V}$ is the on-site scrambling unitary (appropriate for the four copies of state space). As in the case with a single replica, we split $\mathcal{U}_Z$ up into a product *bricks* $\mathcal{U}_{x,x+1}$, given by $\mathcal{U}_{x,x+1} = \mathrm{e}^{-\mathrm{i}\varepsilon Z_x Z_{x+1}} \otimes \mathrm{e}^{\mathrm{i}\varepsilon Z_x Z_{x+1}} \otimes \mathrm{e}^{-\mathrm{i}\varepsilon Z_x Z_{x+1}} \otimes \mathrm{e}^{\mathrm{i}\varepsilon Z_x Z_{x+1}}$. This replicated gate has the diagrammatic representation

$$\mathcal{U}_{x,x+1} \equiv \qquad\qquad\qquad\qquad \tag{66}$$

On each leg (labelled by the replica index $1, \overline{1}, 2, \overline{2}$ introduced in equation (9) and by site position), the brick has the option of carrying either a Z or a $\mathbb{1}$. If a leg is carrying a non-identity factor A, we say that the leg is *decorated* and that the factor A is the decoration. In this spirit, and using equations (63) and (65), we find the decoration expansion of the brick to be given by

$$\tag{67}$$

Multiplying $\langle F^x |$ by a layer of two-site gates yields the following,

$$\tag{68}$$

where

$$g(\varepsilon) \equiv \frac{\cos(4\varepsilon) - 1}{4}, \quad \text{and} \quad h(\varepsilon) \equiv \frac{\sin(4\varepsilon)}{4}. \tag{69}$$

We have only depicted the sites x and $x + 1$ either side of the *cut* (the domain wall between the $+$ and $-$ wiring configurations). Every brick $\mathcal{U}_{r,r+1}$ that does not straddle the cut is 'absorbed' into the state using the property

$$\tag{70}$$

To see this, we notice that the state $\langle +| \otimes \langle +|$ connects the replica 1 with $\overline{2}$ and 2 with $\overline{1}$ (see equation (11)), so that the two copies of $U_{r,r+1}$ each find a copy of $U_{r,r+1}^{\dagger}$ to yield $\langle +| \otimes \langle +| \mathcal{U}_{r,r+1} = \langle +| \otimes \langle +|$. The isometry of $\mathcal{U}_{r,r+1}$, $S = (1 \leftrightarrow 2)$, relates $\langle +|$ and $\langle -|$ through $\langle \pm| S = \langle \mp|$. Using this isometry, the first equation in equation (70) implies the second.

The calculation of $\Omega(k)$ is then straight forward. By utilising translational symmetry and definition equation (25) for the momentum space density super-operators W^k, we write

$$\Omega(k) \equiv \left(W^k \left| (\mathcal{U} - \mathbb{1}) \right| W^k \right) = \eta(k)(1 - e^{-ik}) \sum_x q^x e^{-ikx} \langle F^0 \left| (\mathcal{U} - \mathbb{1}) \right| F^x \rangle, \tag{71}$$

where $\eta(k) = \frac{1 - q^{-2} e^{ik}}{1 - q^{-2}}$. When the cuts are misaligned ($x \neq 0$), equation (70) can be used to say $\langle F^0 | (\mathcal{U} - \mathbb{1}) | F^x \rangle = \langle F^0 | (\mathcal{U}_{0,1} - \mathbb{1}) | F^x \rangle = \langle F^0 | (\mathbb{1} - \mathbb{1}) | F^x \rangle = 0$. For aligned cuts ($x = 0$), equation (68) yields $\langle F^0 | (\mathcal{U} - \mathbb{1}) | F^0 \rangle = g(\varepsilon)$. $\Omega(k)$ is then succinctly given by

$$\Omega(k) = \eta(k)(1 - e^{-ik}) g(\varepsilon). \tag{72}$$

In section 6.2 we will see that the circuit averaged[6] memory matrix is $\mathcal{O}(1/q^2)$. Therefore we can use Ω alone in equation (58) to obtain a leading order expression for v_B and D

$$v_0(\varepsilon) \equiv \lim_{q \to \infty} \langle v_B(\varepsilon) \rangle = \frac{1 - \cos(4\varepsilon)}{4}, \qquad \lim_{q \to \infty} \langle D(\varepsilon) \rangle = \frac{v_0(1 - v_0)}{2}. \tag{73}$$

A straightforward calculation shows that brick-work circuits with commuting even and odd bricks (within one Floquet layer) have a strict light-cone of $v_{LC} = 1$ (as opposed to $v_{LC} = 2$ in brick-work circuits with non-commuting even and odd layers). Notably, the $\mathcal{O}(1)$ expression equation (73) for D vanishes when there is either no ballistic spreading ($v_B = 0$) or when operators spread at the light-cone velocity ($v_B = 1$), although it should be noted that v_0 can never approach this limit in this particular model, $0 \leqslant v_0 \leqslant 1/2$. This is reassuring as an operators spreading at the geometric light-cone velocity must have a front with zero width.

The operator spreading dynamics of this model occupies a markedly different regime than that of holographic/SYK physics which exhibit a sharp front and of random unitary circuits [19, 20], where the operator front diffusion constant is strongly suppressed at large q and v_B is close to the maximum velocity allowed by any (two-local) brick-work circuit (recent work [43] investigates the crossover from holographic to random circuit behaviour of the front). In the large q limit of our Floquet circuit, the operator front diffusion constant is roughly the same size as v_B, which itself is far from the light-cone velocity.

5. Corrections from memory effects: a summary

We have so far calculated the contributions to information transport (the butterfly constant v_B and operator front diffusion constant D) coming from slow processes only. In the remainder of this paper we perturbatively calculate the corrections to the butterfly velocity coming from the fast processes packaged in the memory matrix. The parameter that controls this perturbative expansion is $1/q$, and the corrections are a sum of real-time Feynman-like diagrams. This perturbative expansion encounters technical subtleties at large times where $q \to \infty$ asymptotic

---

[6] A circuit average refers to the average of the Haar random unitary V.

methods are insufficient for circuit averaging. Fortunately, we are able to resolve this subtlety by conjecturing that processes contributing to the memory matrix possess a natural exponential decay timescale $\tau(\varepsilon) \sim 1/|g(\varepsilon)|$ [7] for the (circuit averaged) real-time memory matrix. We provide analytical evidence backing this conjecture. This allows us to argue that the problematic large-time contributions are in-fact subleading, enabling us to safely compute corrections from memory effects to $\mathcal{O}(1/q^2)$.

In section 6.1, we present the tools for the computation of memory effects at $\mathcal{O}(1/q^2)$. These tools are $q \to \infty$ scaling results for the Haar average of correlators and products of correlators. For proofs of these results see [51]. In section 6.2, we use these large q results to compute $\Sigma(t)$, finding that the leading order contributions decay exponentially with a timescale $\tau(\varepsilon) \sim 1/|g(\varepsilon)|$. Although only valid for times $t \lesssim q$, this result motivates our conjecture that the exponential decay of $\Sigma(t)$ persists to arbitrarily late times. A detailed discussion of late-time correlation functions and limitation of the large q results of section 6.1 is given in section 7.1.

Before going through the calculation of memory corrections in detail, we present the result of section 6.2, in the form of a circuit averaged butterfly velocity,

$$\langle v_B \rangle = v_0(\varepsilon) + \delta v_S(\varepsilon) + \delta v_F(\varepsilon) + \mathcal{O}(1/q^3), \tag{74}$$

where the memory corrections $\delta v_S(\varepsilon)$ and $\delta v_F(\varepsilon)$ are given at $\mathcal{O}(1/q^2)$ by

$$\delta v_S(\varepsilon) = \frac{1}{q^2} \frac{1 + 5s - 4s^2}{1 - s - 3s^2}, \qquad \delta v_F(\varepsilon) = 2\frac{g^2}{q^2}(\nu(\varepsilon) - f(\varepsilon)), \tag{75}$$

with $s(\varepsilon) = \sin(\varepsilon)^2 \cos(\varepsilon)^2$ and $\nu(\varepsilon) = [4(1 - 2s)(1 - s(1 - 2s))]^{-1}$ and where $f(\varepsilon)$ is given to good approximation[8] (figure C2) by

$$f(\varepsilon) = \frac{1}{7}s(\varepsilon)(1 - 4s(\varepsilon))^2(1 + 6.8s + 16.1s^2). \tag{76}$$

It turns out that the contribution $\delta v_S(\varepsilon)$ only arises because of the spatial (S) translation symmetry of the model: the scramblers in any particular Floquet layer are identical. On the other hand, $\delta v_F(\varepsilon)$ arises due to both the spatial and Floquet (F) symmetries in the problem. To see this, we analyse variations of the model without spatial translation and Floquet symmetry (see appendix E). In figure 2 we plot the different contributions to $\langle \delta v_B \rangle \equiv \langle v_B \rangle - v_0$.

Spatial disorder is often associated with localization, and a suppression of transport. On the other hand the spatiotemporal randomness in random circuits promotes ergodicity, and the rapid growth of operators. Our results for v_B are more in agreement with the former intuition—translational symmetry enhances transport—because the butterfly velocity receives an enhancement $\delta v_S > 0$ when spatial translation symmetry is present. Even less obvious is the role of Floquet symmetry; like translational symmetry, we find an enhancement $\delta v_F > 0$ to v_B. It will be interesting to understand whether these effects are robust at higher orders in perturbation theory, and hold for more general time evolutions.

We now briefly discuss the difficulties with small ε. The $\varepsilon \to 0$ limit represents an obvious sanity check on our results, but also represents a challenging limit in a memory matrix calculation. As $\varepsilon \to 0$, the memory time must diverge, limiting our ability to truncate the memory

---

[7] $g(\varepsilon)$ is as defined in equation (69).
[8] $f(\varepsilon)$ is found by numerically evaluating a sum involving a 5×5 transfer matrix in appendix C and the form for the fitting function is motivated in section 6.3.
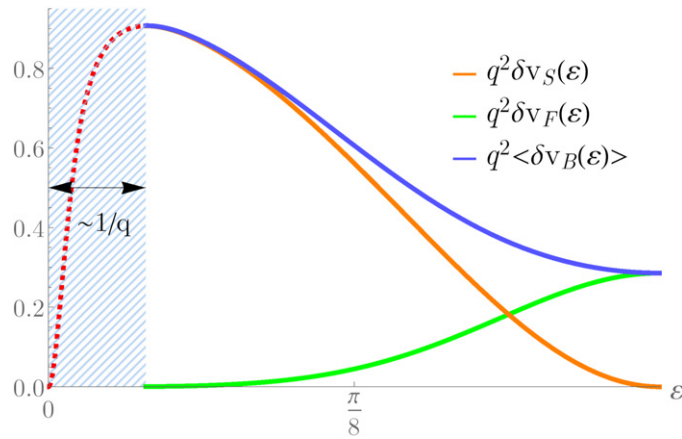
Figure 2. The contributions $\delta v_S(\varepsilon)$ and $\delta v_F(\varepsilon)$ to $\langle \delta v_B \rangle$. As $\varepsilon \to 0$, the perturbative expansion in $1/q$ breaks down (see section 7.2). δv_B must rapidly approach zero, in appendix E.1 we argue that this happen over an $\mathcal{O}(1/q)$ window.

effects. This is demonstrated in the failure of equation (74) for v_B to vanish for $\varepsilon = 0$ where sites decouple and the butterfly velocity is zero. As the memory time reaches $t \sim q$, or equivalently once $\varepsilon \sim 1/q$, our perturbative expansion in $1/q$ breaks down. Therefore, the corrections $\langle \delta v_B \rangle$ as given in (74) are valid only for $\varepsilon > \mathcal{O}(1/q)$, below which $\langle \delta v_B \rangle$ must quickly go to zero as shown in figure 2. We discuss this in detail in section 7.2.

6. Calculating corrections from memory effects

In this section we first present the large q scaling results that form the backbone of the perturbative expansion of $\Sigma(t)$, before presenting a intricate booking keeping of the $\mathcal{O}(1/q^2)$ contributions in section 6.2.

6.1. Averaging n-point correlation functions and their moments

In this section we will present several useful theorems concerning the Haar averages of correlators and products of correlators with random unitary dynamics. The reader is directed to [51] for the proofs of the theorems below. We need only consider correlation functions involving Z's:

$$\langle \mathcal{Z}(\boldsymbol{t}) \rangle \equiv \langle Z(t_1) \ldots Z(t_n) \rangle = \frac{1}{q} \operatorname{Tr}[Z(t_1) \ldots Z(t_n)], \tag{77}$$

where $\boldsymbol{t} = (t_1, \ldots, t_n)$ and $\mathcal{Z}(\boldsymbol{t}) = Z(t_1) \ldots Z(t_n)$ is a product of n 'scrambled' Pauli Z matrices $Z(t) = V^t Z V^{-t}$ (with a Haar random unitary V). We call correlators of form equation (77) arbitrary-time-ordered (ATO) n-point correlation functions as the times t_i are not forced to be ordered. A correlator is trivial if $\mathcal{Z}(\boldsymbol{t}) = \mathbb{1}$. For the purposes of this section, we assume that none of the correlators are trivial.

Before we present the theorems, let us introduce what we call the decoration delta constraint, $\delta^{A,B}$. This is zero unless the two operators $A = Z(a_1), \ldots, Z(a_n)$ and $B = Z(b_1), \ldots, Z(b_n)$

are equal, $AB^{-1} = \mathbb{1}$, for *all* scrambling unitaries V. If the strings $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_{n'})$ do not have repeated consecutive times, then the delta constraint simply checks that $\boldsymbol{a} = \boldsymbol{b}$. Otherwise, we need to introduce a concept called the *minimal form* of a operator.

Definition 1 (Minimal form). With $\mathcal{Z}(\boldsymbol{t}) = Z(t_1) \ldots Z(t_n)$ as a product of scrambled Pauli Z operators and $\mathcal{Z}(\boldsymbol{t}') = Z(t'_1) \ldots Z(t'_n)$ as the form reached after exhaustively using the property $Z(t)^2 = \mathbb{1}$, we define the minimal form of $\mathcal{Z}(\boldsymbol{t})$ as $\mathrm{Min}(\mathcal{Z}(\boldsymbol{t})) \equiv \mathcal{Z}(\boldsymbol{t}')$. We also define the minimal form of the string $\boldsymbol{t}$ as $\mathrm{Min}(\boldsymbol{t}) \equiv \boldsymbol{t}'$.

We can then give a more general definition of the delta constraint.

Definition 2 (Decoration delta constraint). The delta constraint $\delta^{A,B}$ on operators $A = Z(a_1), \ldots, Z(a_n)$ and $B = Z(b_1), \ldots, Z(b_n)$ is defined by

$$\delta^{A,B} = \begin{cases} 1, & \text{if } \mathrm{Min}(\boldsymbol{a}) = \mathrm{Min}(\boldsymbol{b}) \\ 0, & \text{otherwise.} \end{cases} \tag{78}$$

Theorem 1. *The Haar average of a product of p ATO correlators has the following scaling behaviour,*

$$\int \mathrm{d}U \langle \mathcal{Z}(\boldsymbol{t}^1) \rangle \cdots \langle \mathcal{Z}(\boldsymbol{t}^p) \rangle = \mathcal{O}(1/q^{2\lfloor p/2 \rfloor}) \quad \text{as } q \to \infty. \tag{79}$$

Theorem 2. *The Haar average of a product of two ATO correlators is given by*

$$\int \mathrm{d}U \langle \mathcal{Z}(\boldsymbol{t}) \rangle \langle \mathcal{Z}(\boldsymbol{t}') \rangle^* = \frac{S(\boldsymbol{t})}{q^2} \sum_{\tau \in \mathbb{Z}} \delta^{\mathcal{Z}(\boldsymbol{t}+\tau), \mathcal{Z}(\boldsymbol{t}')} + \mathcal{O}(1/q^3) \quad \text{as } q \to \infty, \tag{80}$$

where we have used the decoration delta constraint and where the sum over τ allows for $\boldsymbol{t}$ and $\boldsymbol{t}'$ to differ by a global shift in time. The symmetry factor $S(\boldsymbol{t})$ counts the degree of cyclic symmetry of the list of times $\boldsymbol{t}$, if there exists n cyclic permutations α such that $\alpha(\boldsymbol{t}) = \boldsymbol{t}$, then $S(\boldsymbol{t}) = n$. We will often study a special subset of ATO correlators, which we dub *physical OTOCs*. These take the form $\langle Z\Gamma_1 Z(T)\Gamma_2^\dagger Z\Gamma_2 Z(T)\Gamma_1^\dagger \rangle$, where $\Gamma_i = Z(1)^{s_1^i} Z(2)^{s_2^i} \ldots Z(T-1)^{s_{T-1}^i}$ for binary strings $\boldsymbol{s}^i = (s_1^i, \ldots, s_{T-1}^i)$.

Theorem 3. *The Haar average of a physical OTOC is given by*

$$\int \mathrm{d}V \langle Z\Gamma_1 Z(T)\Gamma_2^\dagger Z\Gamma_2 Z(T)\Gamma_1^\dagger \rangle$$
$$= \frac{1}{q^2} \left(\delta^{\Gamma_1 \Gamma_1^\dagger \Gamma_2 \Gamma_2^\dagger, \mathbb{1}} - \delta^{\Gamma_1, \Gamma_{\overline{2}}} \delta^{\Gamma_2, \Gamma_{\overline{1}}} - \delta^{\Gamma_1, \Gamma_{\overline{1}}} \delta^{\Gamma_2, \Gamma_{\overline{2}}} \right) + \mathcal{O}(1/q^3), \tag{81}$$

where we have again used the decoration delta constraint. This result relies on theorem 4 of [51] and is obtained in equation (B.6) of appendix B.

6.2. Computing Σ at $\mathcal{O}(1/q^2)$

We now compute the memory matrix at leading order in $1/q$, enabling us to calculate the butterfly velocity to $\mathcal{O}(1/q^2)$. As discussed in section 3, we calculate the proxy $\sigma(k, z)$ instead

of $\Sigma(k, z)$. We will find that the $\mathcal{O}(1/q^2)$ contributions to $\sigma(k, t)$ decays exponentially fast with a decay-rate $\gamma(\varepsilon) \approx 2|g(\varepsilon)|$ set by the interaction strength ε.

It is convenient to express $\sigma(k, z)$ as

$$\sigma(k, z) = \eta(k)(1 - e^{-ik})\mathcal{D}(k, z), \tag{82}$$

where

$$\mathcal{D}(k, z) = \sum_x q^x e^{-ikx}\mathcal{D}(x, z), \qquad \mathcal{D}(x, z) = q^x \left\langle F^0 \right| \mathcal{LQ}\frac{1}{e^{-iz} - 1 - \mathcal{L}}\mathcal{QL}\left| F^x \right\rangle. \tag{83}$$

Corrections to v_B from fast processes (i.e., processes counted by Σ) are then given by

$$\delta v = \lim_{z \to i0^+} \lim_{k \to 0} i\sigma(k, z)/k = -\mathcal{D}(k = 0, z = 0). \tag{84}$$

We will often work with the real time version of $\mathcal{D}$,

$$\mathcal{D}(x, T) = q^x \left\langle F^0 \right| \mathcal{LQU}^{T-1}\mathcal{QL}\left| F^x \right\rangle. \tag{85}$$

We separate the scrambling part of each Floquet layer from the two-local bricks as before; $\mathcal{U} = \mathcal{V}\mathcal{U}_Z$. Then a product $\mathcal{U}^n$ can be written $\mathcal{U}^n = \mathcal{U}_Z(1)\mathcal{U}_Z(2)\ldots\mathcal{U}_Z(n)\mathcal{V}^n$ where $\mathcal{U}_Z(t) = \mathcal{V}^t\mathcal{U}_Z\mathcal{V}^{-t}$. One consequence is that

$$\left\langle F^0 \right| \mathcal{L} = \left\langle F^0 \right| \mathcal{U} - \left\langle F^0 \right| \mathbb{1} = \left\langle F^0 \right| \mathcal{V}\mathcal{U}_Z - \left\langle F^0 \right| = \left\langle F^0 \right| \mathcal{L}_Z, \tag{86}$$

where $\mathcal{L}_Z = \mathcal{U}_Z - \mathbb{1}$. We can similarly show $\mathcal{L}\left| F^x \right\rangle = \mathcal{V}\mathcal{L}_Z\left| F^x \right\rangle$. Using this and $[\mathcal{V}, \mathcal{Q}] = 0$ and simplifying the notation $\mathcal{L}_Z \to L$ and $\mathcal{U}_Z \to U$, $\mathcal{D}(x, T)$ can be written as

$$\mathcal{D}(x, T) = q^x \left\langle F^0 \right| LQU(1)\ldots U(T-1)\mathcal{QL}(T)\left| F^x \right\rangle. \tag{87}$$

Using the decoration expansion for each brick equation (67), we express $\mathcal{D}(x, T)$ as a sum over decorations Γ,

$$\mathcal{D}(x, T) = \sum_\Gamma C_\Gamma \mathcal{D}_\Gamma(x, T), \qquad \mathcal{D}_\Gamma(x, T) = q^x \left\langle F^0 \right| L\mathcal{Q}\Gamma \mathcal{QL}(T)\left| F^x \right\rangle. \tag{88}$$

$\Gamma = \otimes_r \Gamma^r$ is a product of decorations on every site. The decoration on some site r is given by $\Gamma^r = \Gamma_1^r \otimes \Gamma_{\bar{1}}^{r*} \otimes \Gamma_2^r \otimes \Gamma_{\bar{2}}^{r*}$ (i.e., a product of decorations on each leg of the site) and where $\Gamma_i^r = Z(1)^{s_1^{r,i}}Z(2)^{s_2^{r,i}}\ldots Z(T-1)^{s_{T-1}^{r,i}}$ for a binary string $\boldsymbol{s}^{r,i} = (s_1^{r,i}, \ldots, s_{T-1}^{r,i})$. We will use the decoration expansion, and a carefully chosen decomposition of the initial states to express contributions as products of ATO'. We will use the theorems 1–3 to see what kinds of decorations can give rise to $O(1/q^2)$ corrections to the circuit averaged σ, and then evaluate those. It turns out that only certain values of x are relevant, $x = 0, 1, 2$ at this order, as we will see.

Using the inversion symmetry of the Floquet unitary, we find

$$\mathcal{D}(-x, T) = q^{-2x}\mathcal{D}(x, T), \quad \text{for } x > 0. \tag{89}$$

We conclude that the contributions from $x < 0$ are at least a factor $1/q^2$ smaller than the $x > 0$ contributions. In the remainder of this section we will see that the contributions for $x \geqslant 0$ are no larger than $\mathcal{O}(1/q^2)$, and that therefore $\mathcal{D}(x < 0, T) = \mathcal{O}(q^{-4})$. We will therefore only consider $x \geqslant 0$ in the following sections.

To evaluate $\mathcal{D}_\Gamma$, we first project $L(T)\,|F^x\rangle$ onto the fast space.

$$\mathcal{Q}L(T)|F^x\rangle = g\left[-\begin{array}{c}|\phi_+(T)\rangle\\|-\rangle\end{array} - \begin{array}{c}|+\rangle\\|\phi_-(T)\rangle\end{array} + \begin{array}{c}Z(T)^{\otimes 2}|+\rangle\\Z(T)^{\otimes 2}|-\rangle\end{array}\right] -\mathrm{i}h\begin{array}{cl}K(T)|+\rangle & \text{site } x\\K(T)|-\rangle & \text{site } x+1\end{array},$$

$$\underset{1}{}\quad\underset{2}{}\quad\underset{3}{}\quad\underset{4}{} \tag{90}$$

where we have suppressed sites $r < x$ $(r > x+1)$ and carefully chosen an orthogonal decomposition of $\mathcal{Q}L(T)\,|F^x\rangle$ in terms of four fast states, numbered from 1 to 4. We have used the shorthand $Z^{\otimes 2}$ and K to represent the following,

$$Z^{\otimes 2} = \begin{array}{c}-Z-\\-Z-\end{array}, \quad K = \begin{array}{c}-Z-\\===\end{array} - \begin{array}{c}===\\-Z-\end{array}. \tag{91}$$

$Z(T)^{\otimes 2}$ and $K(T)$ are defined identically but with $Z(T)$ in place of Z. Finally, $|\phi_+(T)\rangle$ and $|\phi_-(T)\rangle$ are given by

$$|\phi_+(T)\rangle = Z(T)^{\otimes 2}\,|+\rangle - \frac{1}{q-q^{-1}}\,|\perp\rangle, \qquad |\phi_-(T)\rangle = Z(T)^{\otimes 2}\,|-\rangle - \frac{1}{q-q^{-1}}\,|0\rangle, \tag{92}$$

where $|\perp\rangle = |-\rangle - \frac{1}{q}\,|+\rangle$. The initial state is easily found using $\langle F^0|\,LQ = \left(\mathcal{Q}L\,|F^0\rangle\right)^{\mathrm{T}}$.

The four states numbered in equation (90) obey a useful set of identities, which allow us to identify and discard many lower order diagrams and significantly simplify the memory matrix calculation.

6.2.1. Identities of the $\phi_\pm$ states. It will be useful to determine some properties of $|\phi_+\rangle$ and $|\phi_-\rangle$. The isometry $S = (1 \leftrightarrow 2)$ (swaps legs 1 and 2) relates the two vectors, $S\,|\phi_-\rangle = |\phi_+\rangle$. We can then investigate $|\phi_-\rangle$ only. Firstly, $|\phi_-\rangle$ has no overlap with either $|+\rangle$ or $|-\rangle$.

$$\langle -||\phi_-\rangle = \langle -|Z^{\otimes 2}|-\rangle - \frac{1}{q-q^{-1}}\langle -||0\rangle = \langle -|Z^{\otimes 2}|-\rangle = \begin{array}{c}Z\\Z\end{array} = 0$$

$$\langle +||\phi_-\rangle = \langle +|Z^{\otimes 2}|-\rangle - \frac{1}{q-q^{-1}}\langle +||0\rangle = \frac{1}{q^2}\begin{array}{c}Z\\Z\end{array} - \frac{1}{q^3-q}\left[\bigcirc - \frac{1}{q}\,\mathsf{C}\right]$$

$$= \frac{1}{q} - \frac{1}{q-q^{-1}}(1-q^{-2}) = 0.$$

Using S, we can then write

$$\langle\phi_+|\,|\pm\rangle = \langle\phi_-|\,|\pm\rangle = \langle\pm|\,|\phi_+(T)\rangle = \langle\pm|\,|\phi_-(T)\rangle = 0. \tag{93}$$

We next consider what happens when the wires of $\langle\pm|$ are decorated. Let Γ^r be a decoration on the four legs of some site r, i.e., $\Gamma^r = \Gamma_1^r \otimes \Gamma_1^{r*} \otimes \Gamma_2^r \otimes \Gamma_2^{r*}$, where each Γ_i^r take the form $\Gamma_i = Z(1)^{s_1^i}\ldots Z(T-1)^{s_{T-1}^i}$ for some bit-string s_i. This is shown graphically below.

$$
\Gamma^r = \begin{array}{c} 1 \\ \bar{1} \\ 2 \\ \bar{2} \end{array} \begin{array}{cccc} -Z(1)^{a_1^r} \ Z(2)^{a_2^r} & & Z(T-1)^{a_{T-1}^r}- \\ -Z(1)^{\bar{a}_1^{r\,*}} Z(2)^{\bar{a}_2^{r\,*}} & \cdots & Z(T-1)^{\bar{a}_{T-1}^{r}\,*}- \\ -Z(1)^{b_1^r} \ Z(2)^{b_2^r} & & Z(T-1)^{b_{T-1}^r}- \\ -Z(1)^{\bar{b}_1^{r\,*}} Z(2)^{\bar{b}_2^{r\,*}} & & Z(T-1)^{\bar{b}_{T-1}^{r}\,*}- \end{array}.
\tag{94}
$$

The overlaps between $|\phi_-(T)\rangle$ and a decorated $\langle +|$ state is then given by

$$
\langle +|\Gamma^r|\phi_-(T)\rangle = \frac{1}{q^2}\left(\begin{matrix}\Gamma_1 Z(T)\\ \Gamma_{\bar{1}}^*\\ \Gamma_2 Z(T)\\ \Gamma_{\bar{2}}^*\end{matrix}\right) - \frac{1}{q^3-q}\left[\left(\begin{matrix}\Gamma_1\\ \Gamma_{\bar{1}}^*\\ \Gamma_2\\ \Gamma_{\bar{2}}^*\end{matrix}\right) - \frac{1}{q}\left(\begin{matrix}\Gamma_1\\ \Gamma_{\bar{1}}^*\\ \Gamma_2\\ \Gamma_{\bar{2}}^*\end{matrix}\right)\right]
$$
$$
= \frac{1}{q}\langle Z(T)\Gamma_{\bar{2}}^{r\dagger}\Gamma_1^r Z(T)\Gamma_{\bar{1}}^{r\dagger}\Gamma_2^r\rangle
$$
$$
- \frac{1}{q-q^{-1}}\left(\langle\Gamma_{\bar{2}}^{r\dagger}\Gamma_1^r\rangle\langle\Gamma_{\bar{1}}^{r\dagger}\Gamma_2^r\rangle - \frac{1}{q^2}\langle\Gamma_{\bar{2}}^{r\dagger}\Gamma_1^r\Gamma_{\bar{1}}^{r\dagger}\Gamma_2^r\rangle\right).
\tag{95}
$$

Similar identities hold for the overlaps $\langle +||\phi_-(T)\rangle$ and $\langle\phi_-||\pm\rangle$. These identities can be summarised as follows

$$
\langle -|\Gamma^r|\phi_-(T)\rangle = \langle Z(T)\Gamma_{\bar{1}}^{r\dagger}\Gamma_1^r\rangle\langle Z(T)\Gamma_{\bar{2}}^{r\dagger}\Gamma_2^r\rangle
$$
$$
- \frac{1}{q^2-1}\left(\langle\Gamma_{\bar{1}}^{r\dagger}\Gamma_1^r\Gamma_{\bar{2}}^{r\dagger}\Gamma_2^r\rangle - \langle\Gamma_{\bar{1}}^{r\dagger}\Gamma_1^r\rangle\langle\Gamma_{\bar{2}}^{r\dagger}\Gamma_2^r\rangle\right),
$$
$$
\langle\phi_-|\Gamma^r|+\rangle = \frac{1}{q}\langle Z\Gamma_1^r\Gamma_{\bar{2}}^{r\dagger}Z(T)\Gamma_2^r\Gamma_{\bar{1}}^{r\dagger}\rangle
$$
$$
- \frac{1}{q-q^{-1}}\left(\langle\Gamma_1^r\Gamma_{\bar{2}}^{r\dagger}\rangle\langle\Gamma_2^r\Gamma_{\bar{1}}^{r\dagger}\rangle - \frac{1}{q^2}\langle\Gamma_1^r\Gamma_{\bar{2}}^{r\dagger}\Gamma_2^r\Gamma_{\bar{1}}^{r\dagger}\rangle\right),
$$
$$
\langle\phi_-|\Gamma^r|-\rangle = \langle Z(T)\Gamma_1^r\Gamma_{\bar{1}}^{r\dagger}\rangle\langle Z(T)\Gamma_2^r\Gamma_{\bar{2}}^{r\dagger}\rangle
$$
$$
- \frac{1}{q^2-1}\left(\langle\Gamma_1^r\Gamma_{\bar{1}}^{r\dagger}\Gamma_2^r\Gamma_{\bar{2}}^{r\dagger}\rangle - \langle\Gamma_1^r\Gamma_{\bar{1}}^{r\dagger}\rangle\langle\Gamma_2^r\Gamma_{\bar{2}}^{r\dagger}\rangle\right).
\tag{96}
$$

In general, Γ^r will insert operators on each of the four legs of the input state. However, sometimes one can utilize the wirings between each of the legs to simplify the resulting expression, an example for $\langle +|\Gamma^r$ is shown below,

$$
\left(\begin{matrix}-Z(1)-Z(2)-\\ -Z(2)^*-\\ -Z(2)-\\ -Z(1)^*-\end{matrix}\right) = \left(\begin{matrix}-Z(1)^\dagger Z(1)Z(2)-\\ Z(2)^\dagger Z(2)-\end{matrix}\right) = \left(\begin{matrix}-Z(2)-\end{matrix}\right).
\tag{97}
$$

We say Γ^r decorates the state if this simplification process cannot be used to remove all four components of Γ^r. In the example above we were able to remove all of the non-identity operators from the $(2, \overline{1})$ wiring but not from the $(1, \overline{2})$ wiring. Notice that in every case in equation (96) and in (95), if either of the wirings in the $+/-$ states carry a non-identity operator, the overlap with the $\phi_\pm$ states vanishes[9].

Assuming that the decorations non-trivially decorate both wirings of the $+/-$ states, these overlaps can be summarised as follows

$$q \left\langle \pm | \Gamma^r | \phi_\mp(T) \right\rangle = \text{OTOC} - \text{Corr} \times \text{Corr}' + \mathcal{O}(1/q^2)$$

$$q \left\langle \phi_\pm | \Gamma^r | \mp \right\rangle = \text{OTOC} - \text{Corr} \times \text{Corr}' + \mathcal{O}(1/q^2)$$

$$\left\langle \pm | \Gamma^r | \phi_\pm(T) \right\rangle = \text{Corr} \times \text{Corr}' + \mathcal{O}(1/q^2)$$

$$\left\langle \phi_\pm | \Gamma^r | \pm \right\rangle = \text{Corr} \times \text{Corr}' + \mathcal{O}(1/q^2). \tag{98}$$

Where rather than give the full expressions we have simply presented the types of contributions (i.e., OTOCs, products of non-trivial correlators or terms that are manifestly $\mathcal{O}(1/q^2)$). This is often enough to identify diagrams that contribute to $\mathcal{D}(x, T)$ at $\mathcal{O}(1/q^3)$. In cases that require a more careful analysis we refer to equations (95) and (96).

These are useful identities because the diagrams contributing to the memory kernel tend to involve products of terms of this form. We will now see how, using our OTOC identities from section 5, this result allows us to pinpoint which diagrams are able to contribute at leading order in $\mathcal{O}(1/q^2)$.

6.2.2. $\mathcal{D}^{a,b}(x, T)$.

Casting our attention back to the orthogonal decomposition of the projected vector $\mathcal{Q}L(T) |F^x\rangle$ in equation (90) where labelled each of four orthogonal states from 1 to 4, we now use a short hand $\{|1, x, T\rangle, \dots, |4, x, T\rangle\}$ to denote each of these states. This is also done for $\langle F^0 | L\mathcal{Q}$. Using this, we define the following quantity,

$$\mathcal{D}^{a,b}(x, T) \equiv q^x \left\langle a, 0, 0 | U(1)U(2) \dots U(T - 1) | b, x, T \right\rangle, \tag{99}$$

and also the decoration expansion quantity,

$$\mathcal{D}^{a,b}_\Gamma(x, T) = q^x \left\langle a, 0, 0 | \Gamma | b, x, T \right\rangle. \tag{100}$$

The decoration Γ is a product of $T - 1$ *decoration layers*, $\Gamma = \Gamma(1)\Gamma(2), \dots, \Gamma(T - 1)$, one for each unitary layer $U(t)$ of equation (99).

In what follows, we examine $\mathcal{D}^{a,b}$ for all possible pairs a, b; some calculations are carried out in appendices C and D. All contributions are $\leqslant \mathcal{O}(1/q^3)$, except for the $(2, 1)$ and $(4, 4)$ terms as summarised in 6.3.

6.2.3. $(a, b) = (1, 1), (2, 2)$ are $\mathcal{O}(1/q^3)$.

The arguments used for $(a, b) = (1, 1)$ are the same as used for $(2, 2)$, for brevity we will only present them for $(1, 1)$. We study $x > 0$ and $x = 0$ separately, writing $\mathcal{D}^{1,1}_\Gamma(x, T)$ as diagram in both cases.

[9] This is easily verified in equation (95) by substituting either $\Gamma^{r\dagger}_{\overline{2}}\Gamma^r_{\overline{1}} = \mathbb{1}$ or $\Gamma^{r\dagger}_{\overline{1}}\Gamma^r_{\overline{2}} = \mathbb{1}$.

- $x > 0$:

$$\mathcal{D}_\Gamma^{1,1}(x > 0, T) = g^2 \times \begin{bmatrix} & \vdots & & \vdots \\ \text{site } 0 & \langle\phi_+| & & |+\rangle \\ \text{site } 1 & q\langle-| & & |+\rangle \\ & \vdots & \Gamma & \vdots \\ & q\langle-| & & |\phi_+(T)\rangle & \text{site } x \\ & \langle-| & & |-\rangle & \text{site } x+1 \\ & \vdots & & \vdots \end{bmatrix}$$

Using the overlap identities equations (93), (95) and (96), we see that the contribution from site 0 and x either vanish or have the form

$$\left(\text{Corr}_0 \times \text{Corr}_0' + \mathcal{O}(1/q^2)\right)\left(\text{OTOC}_x + \text{Corr}_x \times \text{Corr}_x' + \mathcal{O}(1/q^2)\right).$$

Every other site may contribute either trivial or non-trivial correlators to the product. Therefore, after Haar averaging, theorem 1 of section 6.1 gives

$$\int dV \mathcal{D}_\Gamma^{1,1}(x > 0, T) = \mathcal{O}(1/q^3). \tag{101}$$

- $x = 0$:

$$\mathcal{D}_\Gamma^{1,1}(0, T) = g^2 \times \begin{bmatrix} & \vdots & & \vdots \\ & \langle+| & & |+\rangle \\ \text{site } 0 & \langle\phi_+| & \Gamma & |\phi_+(T)\rangle \\ & \langle-| & & |-\rangle \\ & \vdots & & \vdots \end{bmatrix}$$

On site $x = 0$, we have

$$\langle\phi_+|\,\Gamma^0\,|\phi_+(T)\rangle = \langle Z\Gamma_1^0 Z(T)\Gamma_{\bar{2}}^{0\dagger}\rangle\langle Z\Gamma_2^0 Z(T)\Gamma_{\bar{1}}^{0\dagger}\rangle - \frac{1}{q^2-1}\langle\Gamma_{\bar{1}}^{0\dagger}\Gamma_1^0 Z(T)\Gamma_{\bar{2}}^{0\dagger}\Gamma_2^0 Z(T)\rangle$$

$$- \frac{1}{q^2-1}\langle Z\Gamma_1^0\Gamma_{\bar{1}}^{0\dagger}Z\Gamma_2^0\Gamma_{\bar{2}}^{0\dagger}\rangle + \frac{1}{q^2(1-q^{-2})^2}\langle\Gamma_1^0\Gamma_{\bar{1}}^{0\dagger}\rangle\langle\Gamma_2^0\Gamma_{\bar{2}}^{0\dagger}\rangle$$

$$+ \frac{1}{q^2-1}\langle Z\Gamma_1^0\Gamma_{\bar{2}}^{0\dagger}\rangle\langle Z\Gamma_2^0\Gamma_{\bar{1}}^{0\dagger}\rangle + \frac{1}{q^2-1}\langle\Gamma_{\bar{2}}^{0\dagger}\Gamma_1^0 Z(T)\rangle\langle\Gamma_{\bar{1}}^{0\dagger}\Gamma_2^0 Z(T)\rangle$$

$$- \frac{1}{(q^2-1)^2}\langle\Gamma_1^0\Gamma_{\bar{1}}^{0\dagger}\Gamma_2^0\Gamma_{\bar{2}}^{0\dagger}\rangle - \frac{1}{(q^2-1)^2}\langle\Gamma_1^0\Gamma_{\bar{2}}^{0\dagger}\Gamma_2^0\Gamma_{\bar{1}}^{0\dagger}\rangle$$

$$+ \frac{1}{(q^2-1)^2}\langle\Gamma_1^0\Gamma_{\bar{2}}^{0\dagger}\rangle\langle\Gamma_2^0\Gamma_{\bar{1}}^{0\dagger}\rangle. \tag{102}$$

The final three terms are manifestly $\mathcal{O}(1/q^4)$. The fifth and sixth terms are of the form $\text{Corr} \times \text{Corr}'/q^2$, where these correlators are non-trivial. Therefore, using theorem 1, the Haar average of these terms (possibly multiplied by additional non-trivial correlators from

other sites) is $\mathcal{O}(1/q^4)$. The second, third and fourth terms all have pre-factors of $1/q^2$; if they are to contribute at this order, the accompanying correlators must be trivial. Using decoration delta constraints, this fact (a consequence of theorem 1) is written below

$$\int dV \frac{1}{q^2} \langle \mathcal{Z}_1 \rangle \cdots \langle \mathcal{Z}_m \rangle = \frac{1}{q^2} \prod_i \delta^{\mathcal{Z}_i,\mathbb{1}} + \mathcal{O}(1/q^3). \tag{103}$$

All together, in the context of a Haar average, the following replacement is valid up to $\mathcal{O}(1/q^2)$.

$$\langle \phi_+ | \Gamma^0 | \phi_+(T) \rangle = \langle Z\Gamma_1^0 Z(T) \Gamma_2^{0\dagger} \rangle \langle Z\Gamma_2^0 Z(T) \Gamma_1^{0\dagger} \rangle - \frac{1}{q^2} \delta^{\Gamma_1^0,\Gamma_1^0} \delta^{\Gamma_2^0,\Gamma_2^0} + \mathcal{O}(1/q^3). \tag{104}$$

We say that a decoration Γ^r leaves a site r undecorated if it contributes only trivial correlators, $\langle \mathbb{1} \rangle$. In the present case keeping only $\mathcal{O}(1/q^2)$ contributions forces all sites $r \neq 0$ to be left undecorated. This allows us to take the Haar average of equation (104) directly, using theorem 2 for the Haar average of a product of two correlators. This gives,

$$\int dV \langle \phi_+ | \Gamma^0 | \phi_+(T) \rangle = \frac{1}{q^2} \delta^{\Gamma_1^0,\Gamma_1^0} \delta^{\Gamma_2^0,\Gamma_2^0} - \frac{1}{q^2} \delta^{\Gamma_1^0,\Gamma_1^0} \delta^{\Gamma_2^0,\Gamma_2^0} + \mathcal{O}(1/q^3) = \mathcal{O}(1/q^3). \tag{105}$$

The analysis of $x < 0$ would replicate that of $x > 0$, but with an additional factor of $q^{-2|x|}$. The for all x, $\int dV \mathcal{D}_\Gamma^{1,1}(x, T) = \mathcal{O}(1/q^3)$. Alternatively,

$$\int dV \mathcal{D}^{1,1}(k, T) = \mathcal{O}(1/q^3). \tag{106}$$

6.2.4. $(a, b) = (4, 4)$ is $\mathcal{O}(1/q^2)$. The (4, 4) calculation is significantly more difficult; we present the full calculation here, however; readers interested only in the final result should skip to the summary in section 6.3.

We use the decoration expansion once again to rule out contributions from $x > 0$ and to identify the relevant contributions from $x = 0$.

- $x > 0$:

$$\mathcal{D}_\Gamma^{4,4}(x > 0, T) = -h^2 \times \begin{bmatrix} & \vdots & & \vdots & \\ \text{site } 0 & \langle + | K & & | + \rangle & \\ \text{site } 1 & q\langle - | K & & | + \rangle & \\ & \vdots & \Gamma & \vdots & \\ & q\langle - | & & K(T) | + \rangle & \text{site } x \\ & \langle - | & & K(T) | - \rangle & \text{site } x+1 \\ & \vdots & & \vdots & \end{bmatrix}$$

For $x > 1$, each of the sites 0, 1, x and $x + 1$ contribute non-trivial correlators. When $x = 1$, sites 0, 1 and 2 all contribute non-trivial correlators. In either case, theorem 1 gives

$$\int dV \mathcal{D}_\Gamma^{4,4}(x > 0, T) = \mathcal{O}(1/q^3). \tag{107}$$

- $x = 0$:

$$\mathcal{D}_\Gamma^{4,4}(x = 0, T) = -h^2 \times \begin{bmatrix} \vdots & & & \vdots \\ & \langle + | & & | + \rangle \\ \text{site } 0 & \langle + | K & & K(T) | + \rangle \\ \text{site } 1 & \langle - | K & \Gamma & K(T) | - \rangle \\ & \langle - | & & | - \rangle \\ \vdots & & & \vdots \end{bmatrix}$$

Both sites 0 and 1 contribute non-trivial correlators. Keeping only the $\mathcal{O}(1/q^2)$ contributions means selecting decorations on sites $r \neq 0, 1$ that give trivial correlators only. For $r > 1$ this means selecting decorations such that $\langle - | \Gamma^r | - \rangle = \langle \Gamma_1^r \Gamma_{\bar{1}}^{r\dagger} \rangle \langle \Gamma_2^r \Gamma_{\bar{2}}^{r\dagger} \rangle = \langle \mathbb{1} \rangle^2$ (for $r < 0$ simply switch $- \leftrightarrow +$ and $1 \leftrightarrow 2$ in these equations). Choosing only Γ^r that leave a site $r < 0$ ($r > 1$) undecorated (contributing only trivial correlators) is equivalent to the decoration delta constraint $\delta^{\Gamma_1^r, \Gamma_2^r} \delta^{\Gamma_2^r, \Gamma_{\bar{1}}} (\delta^{\Gamma_1^r, \Gamma_{\bar{1}}^r} \delta^{\Gamma_2^r, \Gamma_{\bar{2}}})$. The implementation of these decoration delta constraints is discussed in appendix B. The result of which is that for sites $r < 0$ we sandwiching each decoration layer $\Gamma(t)$ by $\langle + |$ and $| + \rangle$ and by $\langle - |$ and $| - \rangle$ for sites $r < 0$.

Writing the definition of K in equation (91) as $K = Z_1 - Z_2$, where the index refers to which leg the Z decorates, we write the following,

$$\begin{array}{l} \langle + | K \Gamma^0 K(T) | + \rangle \\ \langle - | K \Gamma^1 K(T) | - \rangle \end{array} = \sum_{i,j \in \{1,2\}} \begin{array}{l} \langle + | Z_i \Gamma^0 Z(T)_i | + \rangle \\ \langle - | Z_j \Gamma^1 Z(T)_j | - \rangle \end{array} + \begin{array}{l} \text{terms with more than two} \\ \text{non-trivial correlators.} \end{array} \tag{108}$$

Because we are selecting only decorations which leave sites $r \neq 0, 1$ undecorated, we are able to take the Haar average of this expression in isolation. Terms with more than two non-trivial correlators are $\mathcal{O}(1/q^3)$ or smaller and are therefore discarded, leaving only the Haar average of the first term. Using theorem 2, this is given by

$$\int dV \sum_{i,j \in \{1,2\}} \begin{array}{l} \langle + | Z_i \Gamma^0 Z(T)_i | + \rangle \\ \langle - | Z_j \Gamma^1 Z(T)_j | - \rangle \end{array}$$
$$= \frac{1}{q^2} \delta^{\Gamma_1^0, \Gamma_{\bar{1}}^1} \delta^{\Gamma_2^0, \Gamma_{\bar{1}}^1} \delta^{\Gamma_2^0, \Gamma_{\bar{1}}^0} \delta^{\Gamma_2^1, \Gamma_{\bar{2}}^1} + \frac{1}{q^2} \delta^{\Gamma_1^0, \Gamma_{\bar{2}}^1} \delta^{\Gamma_2^0, \Gamma_{\bar{2}}^1} \delta^{\Gamma_2^0, \Gamma_{\bar{1}}^0} \delta^{\Gamma_1^1, \Gamma_{\bar{1}}^1}$$
$$+ (1 \leftrightarrow 2) + \mathcal{O}(1/q^3). \tag{109}$$

These delta constraints are implemented by sandwiching the decoration layers $\Gamma(t)$ with the appropriate wirings. The four different wirings configurations for sites 0 and 1 are

shown below.

$$\langle C_{1\bar{1}}| = \begin{pmatrix} & \end{pmatrix}, \ \langle C_{1\bar{2}}| = \begin{pmatrix} & \end{pmatrix}, \ \langle C_{2\bar{1}}| = \begin{pmatrix} & \end{pmatrix}, \ \langle C_{2\bar{2}}| = \begin{pmatrix} & \end{pmatrix}.$$

Counting only the relevant decorations Γ, $\mathcal{D}_\Gamma^{1,1}(x=0,T)$ given by

$$\int dV \mathcal{D}_\Gamma^{4,4}(x=0,T) = -\frac{h^2}{q^2} \sum_{i,j\in\{1,2\}} \left\langle C_{i\bar{j}} \middle| \Gamma(1) \middle| C_{i\bar{j}} \right\rangle \cdots \left\langle C_{i\bar{j}} \middle| \Gamma(T-1) \middle| C_{i\bar{j}} \right\rangle \begin{smallmatrix} x \\ x+1 \end{smallmatrix}$$

(110)

We now sum over all decorations Γ weighted by the coefficients C_Γ appearing in equation (88). This converts back into the picture with full unitary layers $U(t)$. Using the property equation (70) for two-local bricks contracted with the states $\langle\pm|\otimes\langle\pm|$ or $|\pm\rangle\otimes|\pm\rangle$, $\int dV \mathcal{D}^{1,1}(x=0,T)$ is given by the simplified form

$$\int dV \mathcal{D}^{4,4}(x=0,T) = -\frac{h^2}{q^2} \sum_{i,j} \left\langle C_{i\bar{j}} \middle| 1 \middle| C_{i\bar{j}} \right\rangle \cdots \left\langle C_{i\bar{j}} \middle| n \middle| C_{i\bar{j}} \right\rangle + \mathcal{O}(1/q^3)$$

$$= -\frac{h^2}{q^2} \sum_{i,j} \langle \mathcal{C}_{i,\bar{j}}| \mathcal{T} |\mathcal{C}_{i,\bar{j}}\rangle^n + \mathcal{O}(1/q^3),$$

(111)

where $\mathcal{T}$ is given by

$$\boxed{\mathcal{T}} = \begin{array}{c} \langle+| \ |+\rangle \\ \\ \langle-| \ |-\rangle \end{array} = \begin{array}{c} T_+ \\ T_- \end{array},$$

(112)

and where the tensor contractions T_+ and T_- are given algebraically by

$$T_+ = \cos\varepsilon^4 \ + \sin\varepsilon^4 \ - \frac{g}{2}\left[\ + \ \right], \quad T_- = \cos\varepsilon^4 \ + \sin\varepsilon^4 \ - \frac{g}{2}\left[\ + \ \right]$$

(113)

Due to the replica symmetry of the unitary evolution operator, $(1, \overline{1}) \leftrightarrow (2, \overline{2})$, the contributions from the $(i, j) = (1, 1)$ and $(2, 2)$ wirings are identical, as are $(1, 2)$ and $(2, 1)$ wirings. The unitary $e^{-i\varepsilon\mathcal{L}}$ also has the property $\mathrm{Sw}(1, \overline{1})\mathrm{Sw}(2, \overline{2})e^{-i\varepsilon\mathcal{L}}\mathrm{Sw}(1, \overline{1})\mathrm{Sw}(2, \overline{2}) = (e^{-i\varepsilon\mathcal{L}})^*$, where $\mathrm{Sw}(i, \overline{i})$ swaps unbarred leg i and barred leg $\overline{i}$. This transformation is a symmetry of the $\langle+|$ and $\langle-|$ wirings while exchanging the $\mathcal{C}_{1,\overline{1}}$ $(\mathcal{C}_{2,\overline{2}})$ and $\mathcal{C}_{2,\overline{1}}$ $(\mathcal{C}_{1,\overline{2}})$ wirings. Therefore,

$$\xi(\varepsilon) \equiv \langle\mathcal{C}_{1,\overline{1}}|\,\mathcal{T}\,|\mathcal{C}_{1,\overline{1}}\rangle = \langle\mathcal{C}_{2,\overline{2}}|\,\mathcal{T}\,|\mathcal{C}_{2,\overline{2}}\rangle = \langle\mathcal{C}_{1,\overline{2}}|\,\mathcal{T}\,|\mathcal{C}_{1,\overline{2}}\rangle^* = \langle\mathcal{C}_{2,\overline{1}}|\,\mathcal{T}\,|\mathcal{C}_{2,\overline{1}}\rangle^*. \tag{114}$$

Using the decoration decomposition of a brick in equation (67) and the expressions for T_+ and T_- in equation (113), $\xi(\varepsilon)$ is found to be

$$\xi(\varepsilon) = (1 + g)^2 - 2ihg. \tag{115}$$

All together we find,

$$\int \mathrm{d}V\mathcal{D}^{4,4}(x = 0, T) = -\frac{4h^2}{q^2}\mathfrak{R}\{\xi^{T-1}\} + \mathcal{O}(1/q^3). \tag{116}$$

Summing over x, we find the $\mathcal{O}(1/q^2)$ contribution to $\int \mathrm{d}V\mathcal{D}^{4,4}(k, T)$ is given by precisely the same quantity. The decay rate is given by $\gamma(\varepsilon) \equiv \ln(|\xi|^{-1})$, which is always close to $2|g(\varepsilon)|$, $2|g(\varepsilon)| \leqslant \gamma(\varepsilon) \leqslant 4\ln(2)|g(\varepsilon)|$.

6.3. Table of results and summary

We summarise the contributions $\int \mathrm{d}V\mathcal{D}^{a,b}(k = 0, T)$ in the table below, highlighting the only contributions at $\mathcal{O}(1/q^2)$.

$a\backslash b$	1	2	3	4
1	q^{-3}	q^{-3}	q^{-3}	q^{-3}
2	q^{-2}	q^{-3}	q^{-3}	q^{-3}
3	q^{-3}	q^{-3}	q^{-3}	q^{-3}
4	q^{-3}	q^{-3}	q^{-3}	q^{-2}

We calculate the $(a, b) = (2, 1)$ contribution in appendices C and in D we find that the remaining pairs (a, b) contribute at $\mathcal{O}(1/q^3)$ or smaller. Using equation (84), we are only required to know $\int \mathrm{d}V\mathcal{D}^{a,b}(k = 0, z = 0)$ to determine the butterfly velocity[10]. For $(4, 4)$ this is given below, with the re-parameterisation $s(\varepsilon) = \sin(\varepsilon)^2 \cos(\varepsilon)^2$,

$$\int \mathrm{d}V\mathcal{D}^{(4,4)}(k = 0, z = 0) = -\frac{1}{q^2}\frac{1 + 5s - 4s^2}{1 - s - 3s^2} + \mathcal{O}(1/q^3). \tag{117}$$

This re-parameterisation is motivated by the following observations about the dependence of v_B and D on ε. Firstly, under the variable shift $\varepsilon \to \pi/2 + \varepsilon$, the full coupling unitary then transforms as $e^{-i\varepsilon H} \to (-i)^{N-1}e^{-i\varepsilon H}$. The operator dynamics is blind to global phases, meaning that $\varepsilon \to \pi/2 + \varepsilon$ is a symmetry of $v_B(\varepsilon)$ and $D(\varepsilon)$. Secondly, by globally swapping leg 1 with $\overline{1}$ and 2 with $\overline{2}$, we find $\Sigma_{V,\varepsilon}(k, z) = \Sigma_{V^*,-\varepsilon}(k, z)$. and $\Omega_\varepsilon(k) = \Omega_{-\varepsilon}(k)$, where we have labelled Σ with the scrambler and coupling strength and labelled Ω with the coupling strength used (Ω is independent of V). By integrating over V, we find another symmetry of the circuit

averaged butterfly velocity and diffusion constant $\langle v_{\mathrm{B}}(\varepsilon)\rangle$ and $\langle D(\varepsilon)\rangle$, namely $\varepsilon \to -\varepsilon$. Using these symmetries, we determine that $\langle v_{\mathrm{B}}\rangle$ is a function of $s(\varepsilon)$ only.

For $(a, b) = (2, 1)$, we find

$$\int \mathrm{d}V \mathcal{D}^{(2,1)}(k = 0, z = 0) = \frac{2g^2}{q^2}(f(\varepsilon) - \nu(\varepsilon)) + \mathcal{O}(1/q^3), \tag{118}$$

where $\nu(\varepsilon) = [4(1 - 2s)(1 - s(1 - 2s))]^{-1}$ and where $f(\varepsilon)$ is found by numerically evaluating a sum involving a 5×5 transfer matrix in appendix C and given to good approximation (figure C2) by

$$f(\varepsilon) = \frac{1}{7}s(\varepsilon)(1 - 4s(\varepsilon))^2(1 + as + bs^2). \tag{119}$$

where $a = 6.8$ and $b = 16.1$. The factors $\frac{1}{7}s(1 - 4s)^2$ is obtained analytically by diagonalising the transfer matrix at small s and around the point $s = 1/4$. Altogether, we then find

$$\int \mathrm{d}V \mathcal{D}(k = 0, z = 0) = 2\frac{g^2}{q^2}(f(\varepsilon) - \nu(\varepsilon)) - \frac{1}{q^2}\frac{1 + 5s - 4s^2}{1 - s - 3s^2} + \mathcal{O}(1/q^3), \tag{120}$$

so that the correction to the circuit averaged butterfly velocity is

$$\langle \delta v_{\mathrm{B}}(\varepsilon)\rangle = -\int \mathrm{d}V \mathcal{D}(k = 0, z = 0) = \delta v_F(\varepsilon) + \delta v_S(\varepsilon) + \mathcal{O}(1/q^3), \tag{121}$$

where

$$\delta v_F(\varepsilon) = 2\frac{g^2}{q^2}(\nu(\varepsilon) - f(\varepsilon)), \qquad \delta v_S(\varepsilon) = \frac{1}{q^2}\frac{1 + 5s - 4s^2}{1 - s - 3s^2}. \tag{122}$$

As discussed in section 5, the correction $\delta v_S(\varepsilon)$ only arises because of the spatial translation symmetry of the model. Whereas $\delta v_F(\varepsilon)$ arises due to both the spatial and Floquet symmetries. To illustrate this point, we carry out a similar calculation for a version of the model with independently distributed scramblers $V_{x,t}$ for each site x and at each layer of unitaries U_t. This is done in appendix E where we find the corrections to the circuit averaged butterfly velocity $\langle \delta v\rangle \sim \mathcal{O}(1/q^3)$, i.e., $v_{S,F}$ both vanish. Turning to the variant of the model with spatial translation symmetry but no time translation symmetry, we find $\langle \delta v_{\mathrm{B}}\rangle = \delta v_S(\varepsilon) + \mathcal{O}(1/q^3)$. Comparing these result to equation (74) allows us to identify the corrections to v_{B} with the symmetries of the model.

Away from the weak coupling limit, the averaged butterfly velocity is given by equation (74). The weak coupling limit is discussed in section 7.2.

7. Discussion

In this section we discuss the late time behaviour of the memory function, addressing the limitations of the Haar averaged ATO' results of section 6.1. We also discuss the break down of the perturbative scheme used in section 6.2 as $\varepsilon \to 0$.

7.1. Late times

We have presented a Kubo-like formula for $v_{\rm B}$ in equation (61), involving the time integral of a correlator of fast operators, analogous to the *J–J* correlation functions in Kubo formulae for more conventional transport. For $\varepsilon > 0$ (more precisely $g(\varepsilon) \neq 0$), we find that the leading order correlation functions decay exponentially in time at leading order in $1/q$. Given that the correlation functions are between fast variables, it is plausible that this exponential decay is not simply an artifact of the large q limit, i.e., the correlation functions decay over some time scale which is bounded below by a q-independent quantity $\tau(\varepsilon)$, which is positive except at exceptional point where the sites decouple i.e., when $\varepsilon = n\pi/2$ (we discuss the decoupling limit later). Assuming that $t_\varepsilon(q)$ is a (cutoff) time up to which we can safely use the large q correlator theorems of section 6.1 to identify the $\mathcal{O}(1/q^2)$ contributions to Σ, the error incurred by truncating the infinite time evolution is $\mathcal{O}(\exp[-t_\varepsilon(q)/\tau(\varepsilon)])$. We will now argue that $t_\varepsilon(q)$ increases linearly with q. As a result, the error due to truncation decays exponentially in q, which is much smaller than the $\mathcal{O}(1/q^2)$ corrections we calculate. This demonstrates the validity of our expansion in q^{-1}.

We now give some justification for the existence of a cutoff $t_\varepsilon(q) = cq$ for a q-independent constant c. We do this in two parts: (1) we argue that the expressions in section 6.1 for the Haar average of few correlators are valid for times $t < cq$; (2) we argue that this is enough to guarantee that contributions involving many correlators (more than two) are $\mathcal{O}(1/q^3)$ (for times $t < cq$).

In order to do this, we must go beyond the $q \to \infty$ scaling results of section 6.1. In the absence of an analytical understanding of the corrections at finite q, we took a numerically Haar average of an assortment of ATO correlators and products of correlators, finding good evidence that the expression for the Haar average of a product of two correlators (theorem 2) and of OTOCs (theorem 3) experience only $\mathcal{O}(1/q^3)$ errors for times $t < cq$. It turns out that this, along with the use of a Holder's inequality and the monotonicity of the *p*-norm, is enough to bound the Haar average of arbitrarily many correlators as $\mathcal{O}(1/q^3)$.

To show this, we start by bound the expectation of a product of many (more than two) non-trivial correlators as shown below,

$$|\int dV \prod_i \langle \mathcal{Z}(t_i)\rangle| \leqslant \int dV \prod_i |\langle \mathcal{Z}(t_i)\rangle|. \tag{123}$$

Then, using the fact $|\langle \mathcal{Z}(t_i)\rangle| \leqslant 1$, we can bound the average of a product of many correlators by an average over only a few. We choose to highlight three correlators

$$|\int dV \prod_i \langle \mathcal{Z}(t_i)\rangle| \leqslant \int dV |\langle \mathcal{Z}(t_1)\rangle||\langle \mathcal{Z}(t_2)\rangle||\langle \mathcal{Z}(t_3)\rangle|. \tag{124}$$

Using a generalised Holder's inequality and monotonicity of the *p*-norm, we further bound equation (124) by

$$|\int dV \prod_i \langle \mathcal{Z}(t_i)\rangle| \leqslant \sqrt{\int dV |\langle \mathcal{Z}(t_1)\rangle|^2 \int dV |\langle \mathcal{Z}(t_2)\rangle|^2 \int dV |\langle \mathcal{Z}(t_3)\rangle|^2}. \tag{125}$$

Then, for times $t < cq$ and using theorem 2 for the second moment of a correlator, this bound takes the form

$$| \int dV \prod_i \langle \mathcal{Z}(t_i) \rangle | \leqslant C \frac{S(t_1)S(t_2)S(t_3)}{q^3}, \qquad (126)$$

for an $\mathcal{O}(1)$ constant C. The correlators contributing to Σ have $1 \leqslant S(t) \leqslant 2$, so that the right hand-side of equation (126) then simplifies to C'/q^3 for an $\mathcal{O}(1)$ constant C'. For times $t < cq$, any contribution to Σ with three of more non-trivial correlators is $\mathcal{O}(1/q^3)$, and the $\mathcal{O}(1/q^2)$ contributions are counted precisely as we have done in section 6.2 and the appendix C.

7.2. The $\varepsilon \to 0$ limit

In the $\varepsilon \to 0$ limit, the Floquet unitary is given by a product of single-site unitaries, so that there is no operator growth dynamics at all ($v_B = 0$). However, the $\varepsilon \to 0$ limit of the expression equation (74) yields $v_B = 1/q^2$. The reason for this failure to predict the correct operator dynamics as $\varepsilon \to 0$ lies in the fact that our $1/q$ perturbative scheme breaks down once ε is as small as $\varepsilon \sim 1/q$. Diagrams that we previously dismissed as $\mathcal{O}(1/q^3)$ at strong coupling, can in fact contribute at $\mathcal{O}(1/q^2)$ due to appearance of factors of $\varepsilon \sim 1/q$ in the denominator (after summing over time).

We do not have access to the exact expressions for the Haar average of ATO correlators in the Floquet model, instead relying on $\mathcal{O}(1/q^2)$ results. However, in appendix E, we study a variant of the model with independently distributed scramblers V_t between time-step for which exact Haar averaged results are possible for certain diagrams. If we were to naively identify the $1/q^2$ contributions before taking a sum over time, as we did in section 6.2, we would find the same, incorrect $(a, b) = (4, 4)$ contribution and incorrect behaviour of v_B as $\varepsilon \to 0$. In appendix E.1, we consider the same family of diagrams, those involving only two non-trivial correlators (whose contribution we dub $\mathcal{D}_2^{4,4}$), but now take the Haar average exactly. In doing so, we find that the previously troublesome (4, 4) contribution now vanishes as $\varepsilon \to 0$ equation (E.23),

$$\langle \mathcal{D}_2^{4,4}(k = 0, z = 0) \rangle \approx -\frac{8\varepsilon^2}{1 + 8q^2\varepsilon^2}. \qquad (127)$$

We suggest then, that there is a region of width $\mathcal{O}(1/q)$ in figure 2, where v_B rapidly approaches zero.

8. Conclusion

In this paper, we re-purposed a hydrodynamic formalism (the MMF) for information transport calculations by identifying the conservation of the right density of a Heisenberg time evolved operator as a pseudo-local conservation law. A number of modifications to the existing MMF are necessary: in particular, we are led to use an unusual inner product on our space of observables (equation (20)), which leads to the prediction of ballistic operator growth assuming we have identified a sufficiently complete space of slow operators. We use this new formalism to produce a Kubo formula for the butterfly velocity (equation (36)) and also find symmetry constraints on the operator growth light-cone (equation (48)).

In section 4 we used this formalism to investigate a family of translationally invariant Floquet models, finding leading order expressions for the circuit averaged butterfly velocity v_B

and operator front diffusion constant D. By leveraging large q random unitary dynamics [51], we found that a simple hierarchy of contributions to the averaged memory matrix $\langle \Sigma \rangle$ emerges, organised by the number of non-trivial correlators contributed. This enabled us to select only the $\mathcal{O}(1/q^2)$ contributions, associated with processes that explore a manageable sub-region of $\mathcal{Q}$ in which only the sites directly either side of the cut (the $+/-$ domain wall) are decorated by non-identity operators. We have then counted all these processes and found that at $\mathcal{O}(1/q^2)$, the memory matrix decays exponentially fast, with an $\mathcal{O}(1)$ decay rate. We used this to calculate corrections to v_{B}, and were able to distinguish the effects of spatiotemporal symmetry on information transport by identifying which processes arise as a consequence of spatial translation and Floquet symmetry.

9. Further work/outlook

More exotic information hydrodynamics than biased diffusion is possible in the presence of additional conserved charges. With a $U(1)$ charge, the diffusive conserved components acts as a source of non-conserved operators, giving rise to power-law tails in the spatial distributions of operator weight [21, 22]. It will be interesting to incorporate additional symmetries, such as a $U(1)$ or fracton symmetry, into the MMF and perform a mode coupling analysis to confirm and perhaps extend existing results. Even in the presence of conservation laws, the diffusive broadening of the operator front appears ubiquitous in chaotic and interacting integrable systems in 1D [15, 52]. What makes this diffusive broadening so universal? Perhaps, by examining MMF expressions for the front diffusion constant, we can say something about the nature or number of additional slows required to find a hydrodynamical equation other than biased diffusion. Other avenues to explore include the formulation of an information mode MMF at finite temperature/chemical potential and, with only minor modifications, calculating purity. Another potentially fruitful application of formalism is in the setting of perturbed dual unitary circuits, which may serve as a testing ground for perturbative MMF calculations.

Acknowledgments

Data availability statement

No new data were created or analysed in this study.

Appendix A. An unusual inner-product

In this appendix we prove that the inner-product $(\cdot|\cdot)$ defined in equation (20) satisfies the necessary axioms. We repeat the definition below,

$$(A|B) \equiv \langle \Phi(A)| \, |B\rangle_{\mathcal{W}} = \mathrm{Tr}\left(\Phi(A)^\dagger B\right), \tag{A.1}$$

where the super-operator Φ is given by

$$\Phi \equiv \sum_x \frac{1}{\chi_x^2} |W^x\rangle \langle W^x| + Q, \qquad \chi_x \equiv \langle W^x| |W^x\rangle, \ \Rightarrow \ \Phi |W^x\rangle = \frac{1}{\chi_x} |W^x\rangle.$$

(A.2)

Q is the Hermitian projector onto $\mathcal{Q}$. We must show that this constitutes a bona fide inner product by checking each of the inner product axioms.

(a) Conjugate symmetry: $(A|B)^* = (B|A)$,

Exploiting the conjugate symmetry of the inner product $\langle \cdot| |\cdot\rangle$, we write

$$(A|B) = \langle \Phi(A)||B\rangle = \langle A||\Phi(B)\rangle = \langle \Phi(B)||A\rangle^* = (B|A)^*,$$

(A.3)

where we have used the fact that Φ is Hermitian, this is because χ_x is real and the projector $\hat{Q}$ is Hermitian.

(b) Linearity in second argument: $(A|\beta B + \gamma C) = \beta (A|B) + \gamma (A|C)$ for scalars β, γ,

$$(A|\beta B + \gamma C) = \langle \Phi(A)||\beta B + \gamma C\rangle = \beta\langle \Phi(A)||B\rangle + \gamma\langle \Phi(A)||C\rangle$$
$$= \beta(A|B) + \gamma(A|C).$$

(A.4)

where we have used linearity in the second argument of the inner product $\langle \cdot| |\cdot\rangle$.

(c) Positive definiteness: $(A|A) > 0$ for $A \neq 0$,

$$(A|A) = \langle \Phi(A)| |A\rangle = \langle A| \Phi |A\rangle > 0$$

(A.5)

where we have used the fact that Φ is a positive definite matrix. This is easily seen by noticing that all χ_x^2 are real and positive.

This confirms that $(A|B)$ is indeed an inner-product and allows us to consider a new inner-product space in which the weight operators are orthonormal.

Appendix B. Imposing the decorations delta constraints

In this section we revisit the Haar average identities for ATO correlators and their moments, specifically theorems 2 and 3. The final results (at leading order in $1/q$) involved 'delta constraints' (as defined in definition 2) which are zero/one depending on whether or not the decorations were equal (for all scramblers V). We have also seen these delta constraints whenever we have demanded that a correlator be trivial (see equation (104)). In this section, we show how these delta constraints can be imposed by placing the decorations on a contour and inserting projectors at every time step; the insertion of projectors has an appealingly simple graphical interpretation, which facilitates our calculation of σ in the main text.

B.1. The constraint $\delta^{\Gamma_1, \Gamma_{\bar{1}}}$

The simplest example of a decoration delta constraint to consider is $\delta^{\Gamma_1, \Gamma_{\bar{1}}}$, where the two decorations are time-ordered, i.e., $\Gamma_1 = Z(1)^{a_1} Z(2)^{a_2} \ldots Z(n)^{a_n}$ and $\Gamma_{\bar{1}} = Z(1)^{b_1} Z(2)^{b_2} \ldots Z(n)^{b_n}$ for binary strings $\boldsymbol{a}$ and $\boldsymbol{b}$. In this case, the delta constraint checks that $a_i = b_i$ for all i. This is equivalent to putting Γ_1 and $\Gamma_{\bar{1}}$ on a wiring with a single forward and backward contour

$$\Gamma \;=\; \begin{matrix} -\Gamma_1- \\ -\Gamma_{\overline{1}}^{*}- \\ -\Gamma_2- \\ -\Gamma_{\overline{2}}^{*}- \end{matrix} \;=\; \begin{matrix} -Z(1)^{a_1}- \\ -Z(1)^{\overline{a}_1\,*}- \\ -Z(1)^{b_1}- \\ -Z(1)^{\overline{b}_1\,*}- \end{matrix} \begin{matrix} -Z(2)^{a_2}- \\ -Z(2)^{\overline{a}_2\,*}- \\ -Z(2)^{b_2}- \\ -Z(2)^{\overline{b}_2\,*}- \end{matrix} \cdots \begin{matrix} Z(n)^{a_{T-1}}- \\ Z(n)^{\overline{a}_{T-1}\,*}- \\ Z(n)^{b_{T-1}}- \\ Z(n)^{\overline{b}_{T-1}\,*}- \end{matrix}$$
$$\qquad\qquad\qquad \Gamma(1) \qquad \Gamma(2) \qquad\quad \Gamma(n)$$
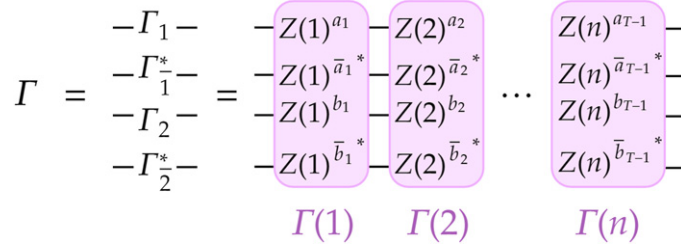
Figure B1. A decoration Γ on the four legs $1, \overline{1}, 2$ and $\overline{2}$ with n decoration layers.

(labelled 1 and $\overline{1}$) and then placing a projector between each of the decoration layers are shown below.

$$\delta^{\Gamma_1,\Gamma_{\overline{1}}} = \frac{1}{q^n} \left(\begin{matrix} Z(1)^{a_1} \\ Z(1)^{b_1*} \end{matrix} \supset\subset \begin{matrix} Z(2)^{a_2} \\ Z(2)^{b_2*} \end{matrix} \supset\subset \cdots \supset\subset \begin{matrix} Z(n)^{a_n} \\ Z(n)^{b_n*} \end{matrix} \right). \qquad (B.1)$$

The right-hand side checks that at each decoration layer t, $a_t = b_t$, this is precisely the same as the decoration delta constraint. A site $r \leqslant 0$ with decoration Γ^r contributes $\langle+|\,\Gamma^r\,|+\rangle$ in the decoration expansion, suppose that we demanded that this contribution contained only trivial correlators ($\langle+|\,\Gamma^r\,|+\rangle = \langle\mathbb{1}\rangle^2$). The decorations that meet this condition are those that satisfy the constraint $\delta^{\Gamma_1^r,\Gamma_{\overline{1}}^r}\delta^{\Gamma_2^r,\Gamma_{\overline{2}}^r}$. The projector insertion technique described above selects precisely these relevant decorations as follows

$$\delta^{\Gamma_1^r,\Gamma_{\overline{1}}^r}\delta^{\Gamma_2^r,\Gamma_{\overline{2}}^r} = \langle+|\,\Gamma^r(1)\,|+\rangle \,\langle+|\,\Gamma^r(2)\,|+\rangle \ldots \langle+|\,\Gamma^r(n)\,|+\rangle. \qquad (B.2)$$

For sites $r > x + 1$, where the contributions take the form $\langle-|\,\Gamma^r\,|-\rangle$, the decorations that contribute trivial correlators are identified in the same way but by projecting with $|-\rangle\,\langle-|$.

B.2. The constraint $\delta^{\Gamma_1\Gamma_{\overline{2}}^\dagger\Gamma_2\Gamma_{\overline{1}}^\dagger,\mathbb{1}}$

Consider next, a decoration $\Gamma = \Gamma_1 \otimes \Gamma_{\overline{1}}^* \otimes \Gamma_2 \otimes \Gamma_{\overline{2}}^*$ with n decoration layers ($\Gamma = \Gamma(1),\ldots,\Gamma(n)$) given in figure B1, with $\Gamma_1 = Z(1)^{a_1}\ldots Z(n)^{a_n}$, $\Gamma_{\overline{1}} = Z(1)^{\overline{a}_1}\ldots Z(n)^{\overline{a}_n}$, $\Gamma_2 = Z(1)^{b_1}\ldots Z(n)^{b_n}$ and $\Gamma_{\overline{2}} = Z(1)^{\overline{b}_1}\ldots Z(n)^{\overline{b}_n}$ for binary strings $\boldsymbol{a}, \overline{\boldsymbol{a}}, \boldsymbol{b}$ and $\overline{\boldsymbol{b}}$.

As in the previous case, we will obtain a prescription for rewiring the legs of a contour at every time step, this prescription will identify decorations Γ that satisfy the constraint. The delta constraint $\delta^{\Gamma_1^r\Gamma_{\overline{2}}^{r\dagger}\Gamma_2^r\Gamma_{\overline{1}}^{r\dagger},\mathbb{1}}$ appears whenever we demand that the decoration Γ^r on some site r, with $1 < r < x$, contributes only a trivial correlator, $q\,\langle-|\,\Gamma^r\,|+\rangle = \langle\mathbb{1}\rangle = 1$. Finding the Γ which contribute non-trivial correlators is equivalent to finding Γ which satisfy the delta constraint.

Start then, with $\langle-|\,\Gamma\,|+\rangle$. Assume that, working in from the left, at least one of the decoration layers non-trivially decorates the $\langle-|$ wiring (so that the either one or both of the $(i,\overline{i})$ wirings in $\langle-|$ carry non-identity operators). This excludes the case where $\langle-|\,\Gamma = \langle-|$, which

---

[10] We sum over times up to the cutoff $t_\varepsilon(q)$ as discussed in section 5, incurring only an error exponentially small in q.
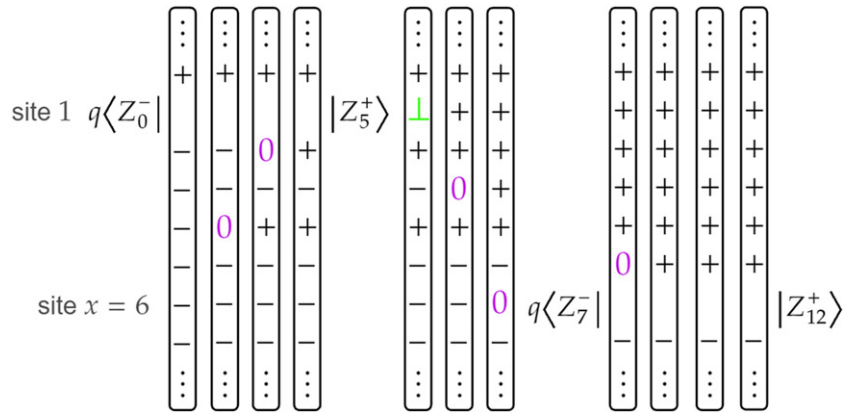
Figure C1. The contribution $(t_1, t_2, t_3, t_4, t_5, t_6) = (5, 3, 6, 2, 8, 7)$ to $\int dV \mathcal{D}^{(2,1)}(x = 6, T = 12)$.
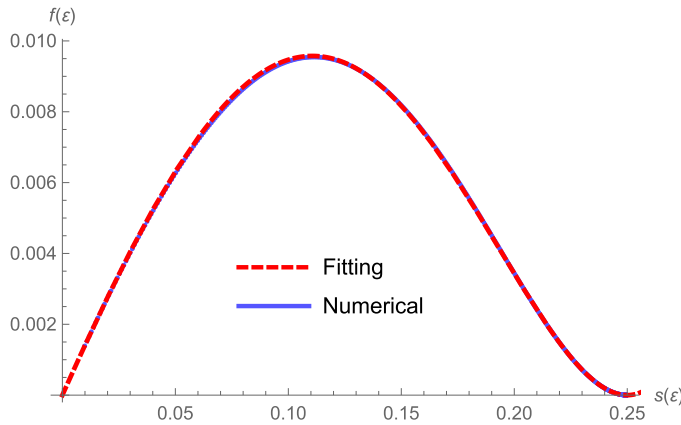
Figure C2. The contribution to v_B from processes with overlapping OTOCs is given by $2g(\varepsilon)^2 f(\varepsilon)/q^2$, $f(\varepsilon)$ is by numerically and plotted in blue. A fitting function is plotted in red.

we will examine last. Let t_i, $0 < t_i \leqslant n$, be the first decoration layer in from the left that non-trivially decorates the $\langle -|$ wiring. Likewise, let t_f be the first decoration layer in from the right that non-trivially decorates the $|+\rangle$ wiring. If $t_i \neq t_f$, the resulting correlator is certainly non-trivial and therefore $X \neq \mathbb{1}$ (the delta constraint is not satisfied). Otherwise, if $t_i = t_f$, we have $\langle -| \Gamma |+\rangle = \langle -| \Gamma(t_i) |+\rangle$. In order for this to be a trivial correlator, $\Gamma(t_i)$ must decorate $\langle -|$ such that $\langle -| \Gamma(t_i) = \langle -| Z(t_i)^{\otimes 2}$ (see equation (91) for definition of $Z^{\otimes 2}$). We can select this case by sandwiching every decoration layer $t < t_i$ by $\langle -|$ and $|-\rangle$, every layer $t > t_i$ by $\langle +|$ and $|+\rangle$ and the layer t_i by $\langle -|$ on the left and on right by $q\,|0\rangle$. Finally, the cases where no decoration layer decorates the $\langle -|$ (an obvious example where $X = \mathbb{1}$) wiring can be selected

by sandwiching every layer with $\langle -|$ and $|-\rangle$. Therefore, the decoration delta constraint can be rewritten as

$$\delta^{\Gamma_1^r \Gamma_2^{r\dagger} \Gamma_2^r \Gamma_1^{r\dagger}, \mathbb{1}} = \sum_{m=1}^{n} \left(\prod_{t=1}^{m-1} \langle -| \Gamma^r(t) |-\rangle \right) q \langle -| \Gamma^r(m) |0\rangle$$
$$\times \left(\prod_{t=m+1}^{n} \langle +| \Gamma^r(t) |+\rangle \right) + \prod_{t=1}^{n} \langle -| \Gamma^r(t) |-\rangle . \quad (B.3)$$

Or diagrammatically as

$$(B.4)$$

B.3. Connection to the OTOC Haar average

In this section we use the results of [51] to re-express the Haar average of a physical OTOC in terms of decoration delta constraints, arriving at the form of this theorem presented in theorem 3 of section 6.1. The Haar average of an OTOC given in [51] is quoted below

$$\int dV \langle Z\Gamma_1 Z(T)\Gamma_1^\dagger Z\Gamma_2 Z(T)\Gamma_2^\dagger \rangle = \frac{1}{q^2} \sum_{m=1}^{n} \left(\prod_{t=1}^{m-1} \langle -| \Gamma(t) |-\rangle \right)$$
$$\times q \langle -| \Gamma(m) |0\rangle \left(\prod_{t=m+1}^{n} \langle +| \Gamma(t) |+\rangle \right)$$
$$- \frac{1}{q^2} \prod_{t=1}^{n} \langle -| \Gamma(t) |-\rangle + \mathcal{O}(1/q^3). \quad (B.5)$$

Comparing this to equation (B.3), we have the following,

$$\int dV \langle Z\Gamma_1 Z(T)\Gamma_1^\dagger Z\Gamma_2 Z(T)\Gamma_2^\dagger \rangle = \frac{1}{q^2} \left(\delta^{\Gamma_1 \Gamma_2^\dagger \Gamma_2 \Gamma_1^\dagger, \mathbb{1}} - \delta^{\Gamma_1, \Gamma_{\bar{1}}} \delta^{\Gamma_2, \Gamma_{\bar{2}}} - \delta^{\Gamma_1, \Gamma_{\bar{2}}} \delta^{\Gamma_2, \Gamma_{\bar{1}}} \right) + \mathcal{O}(1/q^3).$$
$$(B.6)$$

Appendix C. $(a, b) = (2, 1)$

In this appendix, we evaluate the Haar average of the $(a, b) = (2, 1)$ contribution to $\mathcal{D}^{a,b}(x > 0, T)$.

C.1. x = 0

$$
\mathcal{D}_\Gamma^{2,1}(x=0,T) = g^2 \times \begin{bmatrix} & \vdots & \vdots \\ \text{site 0} & \langle +| & \boxed{\Gamma} & |\phi_+(T)\rangle \\ \text{site 1} & \langle \phi_-| & & |-\rangle \\ & \vdots & \vdots \end{bmatrix}
\tag{C.1}
$$

Sites 0 and 1 each contribute a product of non-trivial correlation functions plus terms of size $1/q^2$. The Haar average of this is $\mathcal{O}(1/q^4)$.

C.2. x = 1

$$
\mathcal{D}_\Gamma^{2,1}(x=1,T) = g^2 \times \begin{bmatrix} & \vdots & & \vdots \\ & \langle +| & & |+\rangle \\ \text{site 1} & q\langle \phi_-| & \boxed{\Gamma} & |\phi_+(T)\rangle \\ & \langle -| & & |-\rangle \\ & \vdots & & \vdots \end{bmatrix}
\tag{C.2}
$$

Splitting the decoration site by site, we find

$$
\mathcal{D}_\Gamma^{2,1}(x=1,T) = qg^2 \langle \phi_-| \Gamma^1 |\phi_+(T)\rangle \left(\prod_{r\leqslant 0} \langle +| \Gamma^r |+\rangle \right) \left(\prod_{r>1} \langle -| \Gamma^r |-\rangle \right).
\tag{C.3}
$$

The contribution from site 1 is given in full below,

$$
\begin{aligned}
q\langle \phi_-| &\, \Gamma^1 |\phi_+(T)\rangle \\
&= \langle Z\Gamma_1^1 Z(T)\Gamma_2^{1\dagger} Z\Gamma_2^1 Z(T)\Gamma_1^{1\dagger}\rangle - \frac{1}{1-q^{-2}} \langle Z\Gamma_1^1 \Gamma_1^{1\dagger}\rangle \langle Z\Gamma_2^1 \Gamma_2^{1\dagger}\rangle \\
&\quad - \frac{1}{1-q^{-2}} \langle \Gamma_2^{1\dagger}\Gamma_1^1 Z(T)\rangle \langle \Gamma_1^{1\dagger}\Gamma_2^1 Z(T)\rangle + \frac{1}{q^2-1} \langle Z\Gamma_1^1 \Gamma_2^{1\dagger} Z\Gamma_2^1 \Gamma_1^{1\dagger}\rangle \\
&\quad + \frac{1}{q^2-1} \langle \Gamma_1^{1\dagger}\Gamma_1^1 Z(T)\Gamma_2^{1\dagger}\Gamma_2^1 Z(T)\rangle + \frac{1}{q^2(1-q^{-2})^2} \langle \Gamma_1^1 \Gamma_1^{1\dagger}\Gamma_2^1 \Gamma_2^{1\dagger}\rangle \\
&\quad - \frac{1}{q^2(1-q^{-2})^2} \langle \Gamma_1^1 \Gamma_2^{1\dagger}\rangle \langle \Gamma_2^1 \Gamma_1^{1\dagger}\rangle - \frac{1}{q^2(1-q^{-2})^2} \langle \Gamma_1^1 \Gamma_1^{1\dagger}\rangle \langle \Gamma_2^1 \Gamma_2^{1\dagger}\rangle \\
&\quad + \frac{1}{(q^2-1)^2} \langle \Gamma_1^1 \Gamma_2^{1\dagger}\Gamma_2^1 \Gamma_1^{1\dagger}\rangle.
\end{aligned}
\tag{C.4}
$$

Every term is either an OTOC, a product of two non-trivial correlators, or is manifestly $\mathcal{O}(1/q^2)$. The decorations on each site $r \neq 1$ may result in contributions that are either: (1) a trivial correlator; (2) a single non-trivial correlator; (3) a product of two non-trivial correlators. Note that none of these non-trivial correlators are OTOCs because they live on a contour with only a single forward and backward segments. Therefore, if any decoration on sites $r \neq 1$ does anything other than contribute trivial correlators, we have $\int dV \mathcal{D}_\Gamma^{2,1}(x = 1, T) = \mathcal{O}(1/q^3)$. Keeping only $\mathcal{O}(1/q^2)$ contributions forces every site $r \neq 1$ to contribute trivial correlators only. This allows us to take the Haar average of equation (C.4) in isolation. To do this we find it useful to write the follows results (consequences of theorem 1),

$$\int dU \langle \Gamma \rangle \langle \Gamma'^\dagger \rangle = \delta^{\Gamma,\mathbb{1}} \delta^{\Gamma',\mathbb{1}} + \mathcal{O}(1/q^2),$$

$$\int dU \langle \Gamma \rangle = \delta^{\Gamma,\mathbb{1}} + \mathcal{O}(1/q^2),$$

$$\int dU \langle Z\Gamma Z\Gamma' \rangle = \delta^{\Gamma,\mathbb{1}} \delta^{\Gamma',\mathbb{1}} + \mathcal{O}(1/q^2).$$

Where all Γ are products $Z(1)^{\alpha_1} \ldots Z(T-1)^{\alpha_{T-1}}$ for some binary string $\alpha = (\alpha_1^i, \ldots, \alpha_{T-1}^i)$. Using theorem 2 we find the useful result

$$\int dU \langle Z\Gamma \rangle \langle Z\Gamma'^\dagger \rangle = \frac{1}{q^2} \delta^{\Gamma,\Gamma'}(1 - \delta^{\Gamma,\mathbb{1}}) + \mathcal{O}(1/q^4). \tag{C.5}$$

Using these results and theorem 3 for the Haar average of a physical OTOC, we find that at $\mathcal{O}(1/q^2)$, every term in (C.4) cancels,

$$
\begin{aligned}
q \langle \phi_- | \, \Gamma^1 \, | \phi_+(T) \rangle_{\text{Haar}} = \; & \frac{1}{q^2} \Big[\delta^{\Gamma_1^1 \Gamma_{\bar{1}}^{1\dagger} \Gamma_2^1 \Gamma_{\bar{2}}^{1\dagger}, \mathbb{1}} - \delta^{\Gamma_1, \Gamma_{\bar{1}}} \delta^{\Gamma_2, \Gamma_{\bar{2}}} - \delta^{\Gamma_1, \Gamma_{\bar{2}}} \delta^{\Gamma_2, \Gamma_{\bar{1}}} \\
& - \delta^{\Gamma_1^1 \Gamma_{\bar{1}}^{1\dagger}, \Gamma_2^1 \Gamma_{\bar{2}}^{1\dagger}} (1 - \delta^{\Gamma_1^1, \Gamma_{\bar{1}}^1}) \\
& - \delta^{\Gamma_1^1 \Gamma_{\bar{1}}^{1\dagger}, \Gamma_2^1 \Gamma_{\bar{2}}^{1\dagger}} (1 - \delta^{\Gamma_1^1, \Gamma_{\bar{2}}^1}) \\
& + \delta^{\Gamma_1^1, \Gamma_{\bar{2}}^1} \delta^{\Gamma_2^1, \Gamma_{\bar{1}}^1} + \delta^{\Gamma_1^1, \Gamma_{\bar{1}}^1} \delta^{\Gamma_2^1, \Gamma_{\bar{2}}^1} \\
& + \delta^{\Gamma_1^1 \Gamma_{\bar{1}}^{1\dagger} \Gamma_2^1 \Gamma_{\bar{2}}^{1\dagger}, \mathbb{1}} - \delta^{\Gamma_1^1, \Gamma_{\bar{2}}^1} \delta^{\Gamma_2^1, \Gamma_{\bar{1}}^1} \\
& - \delta^{\Gamma_1^1, \Gamma_{\bar{1}}^1} \delta^{\Gamma_2^1, \Gamma_{\bar{2}}^1} \Big] + \mathcal{O}(1/q^3) \\
= \; & \mathcal{O}(1/q^3). \tag{C.6}
\end{aligned}
$$

and we find $\int dV \mathcal{D}^{2,1}(x = 1, T) = \mathcal{O}(1/q^3)$.

C.3. $x \geqslant 2$

$$
\mathcal{D}_\Gamma^{2,1}(x > 2, T) = g^2 \times \begin{bmatrix} \vdots & & \vdots \\ \text{site } 0 & \langle + | & & | + \rangle \\ \text{site } 1 & q\langle \phi_- | & & | + \rangle \\ & \vdots & \Gamma & \vdots \\ & q\langle - | & & | \phi_+(T) \rangle & \text{site } x \\ & \langle - | & & | - \rangle & \text{site } x+1 \\ \vdots & & \vdots \end{bmatrix}
\tag{C.7}
$$

Sites 1 and x each contribute factors of form $\text{OTOC} + \text{Corr} \times \text{Corr}' + \mathcal{O}(1/q^2)$ (see equation (98)). Our theorem for the Haar average of a product of correlators (theorem 1) implies that if any other site contributes a non-trivial correlator, the Haar average of the total contribution will be $\mathcal{O}(1/q^3)$ or smaller. Thus, working to $\mathcal{O}(1/q^2)$, we will look for contributions where sites $r \neq 1, x$ give only trivial correlators. Moreover, theorem 1 also implies that the leading order contribution comes from the $\text{OTOC}_1 \text{OTOC}_x$ cross term

$$
\int dU (\text{OTOC}_1 + \text{Corr}_1 \times \text{Corr}_1' + \mathcal{O}(1/q^2))
$$
$$
\times (\text{OTOC}_x + \text{Corr}_x \times \text{Corr}_x' + \mathcal{O}(1/q^2))
$$
$$
= \int dU\, \text{OTOC}_1 \times \text{OTOC}_x + \mathcal{O}(1/q^3).
\tag{C.8}
$$

In summary, in evaluating the contributions for $x \geqslant 2$, we need only consider those terms in the decoration expansion corresponding to OTOCs on site $1, x$, and trivial correlators on all other sites. As in the $(a, b) = (4, 4)$ calculation, we select decorations that leave the contours on sites $r > x$ ($r < 1$) undecorated by inserting the projector $|+\rangle\langle+|$ ($|-\rangle\langle-|$) between every Floquet layer. For sites $1 < r < x$ the non-decoration condition is more delicate. For these sites, the input wiring configuration is of $-$ type and the output wiring configuration is of $+$ type giving an OTO type contour. The requirement that the OTO contour is undecorated (i.e., a trivial correlator) is equivalent to the decoration delta constraint $\delta^{\Gamma_1^r \Gamma_2^{r\dagger} \Gamma_2^r \Gamma_1^{r\dagger}, \mathbb{1}}$. In appendix B we show that this decoration delta constraint can be rewritten as

$$
\delta^{\Gamma_1^r \Gamma_2^{r\dagger} \Gamma_2^r \Gamma_1^{r\dagger}, \mathbb{1}} = \sum_{m=1}^{n} \left(\prod_{t=1}^{m-1} \langle - | \Gamma^r(t) | - \rangle \right) q \langle - | \Gamma^r(m) | 0 \rangle
$$
$$
\times \left(\prod_{t=m+1}^{n} \langle + | \Gamma^r(t) | + \rangle \right) + \prod_{t=1}^{n} \langle - | \Gamma^r(t) | - \rangle.
\tag{C.9}
$$

The final term in equation (C.9) selects the decorations Γ^r that never decorate the initial state $\langle -|$, so that at each time step the $\langle -|$ wirings never carry any non-identity operators. With these $|-\rangle \langle -|$ projectors in place, let us sum over all decorations Γ with the coefficients C_Γ of equation (88), in doing so we replace each of the decoration layers with the full unitary layers $U(t)$. This is pictured below, where we have highlighted site x, with its terminating state $|\phi\rangle_+(T)$.

$$\text{(C.10)}$$

Where, crucially, the brick property of equation (70) can be used to remove every brick to the right of site r (below site r in the diagram above). This yields the right hand-side of the equation above. A consequence of which is that the terminating states $|\phi_+(T)\rangle$ on site x is contracted directly with $\langle -|$. Using equation (93), we see that this diagram vanishes. Therefore, the decorations selected by the final term of equation (C.9) cannot contribute to $\mathcal{D}^{2,1}(x > 2, T)$. In what follows, we will consider only the decorations selected by the sum in equation (C.9).

As previously noted, only the $\text{OTOC}_1 \times \text{OTOC}_x$ terms can contribute at $\mathcal{O}(1/q^2)$. This allows us to drop all but the $\langle -| Z^{\otimes 2}$ term in the $\langle \phi_-|$ state (see equation (92)) of site 1 and the $Z(T)^{\otimes 2} |+\rangle$ term in the state $|\phi_+(T)\rangle$ of site x. The fact that OTOC_1 begins with $\langle -| Z^{\otimes 2}$ and OTOC_x ends with $Z(T)^{\otimes 2} |+\rangle$ and the condition that these OTOCs are complex conjugates of each other (using theorem 2) forces both OTOCs to be physical OTOCs of the same length (the length of physical OTOCs is the difference between the latest and earliest time appearing in the OTOC). We will sum over all possible OTOC lengths τ, $1 \leqslant \tau \leqslant T-1$.

We now describe how we select only those decorations that produce a physical OTOC with length τ. For OTOC_1, we must ensure that every unitary layer $t > \tau$ does not decorate the $|+\rangle$ wirings on site 1. We do this by inserting the projector $|+\rangle \langle +|$ to the left of each of these layers. Requiring then that the τth unitary layer decorates the $|+\rangle$ wiring by leaving $Z(\tau)^{\otimes 2} |+\rangle$ is achieved by sandwiching the layer with $q \langle \perp|$ and $|+\rangle$. Making all of these selections, the contribution from site 1 is takes the form shown below,

$$q \langle -| Z^{\otimes 2} \Gamma^1(1) \ldots \Gamma^1(\tau - 1) Z(\tau)^{\otimes 2} |+\rangle \, q \langle \perp| \Gamma^1(\tau) |+\rangle$$
$$\times \langle +| \Gamma^1(\tau + 1) |+\rangle \ldots \langle +| \Gamma^1(T-1) |+\rangle. \quad \text{(C.11)}$$

We use the same strategy to select decorations that contribute physical OTOCs of length τ on site x as well. The resulting contribution takes the form shown below, where we have defined $\tau' = T - \tau$,

$$\langle -| \Gamma^x(1) |-\rangle \ldots \langle -| \Gamma^x(\tau' - 1) |-\rangle \, q \langle -| \Gamma^x(\tau') |0\rangle$$
$$\times q \langle -| Z(\tau')^{\otimes 2} \Gamma^x(\tau' + 1) \ldots \Gamma^x(T-1) Z(T)^{\otimes 2} |+\rangle. \quad \text{(C.12)}$$

Rather than focus on a single decoration Γ, we are able to select every $\mathcal{O}(1/q^2)$ to $\mathcal{D}^{2,1}$ ($x = 2, T$) simultaneous by summing over decorations Γ with the appropriate coefficients C_Γ (as introduced in the decoration expansion in equation (88)),

$$\mathcal{D}^{2,1}(x = 2, T) = \sum_\Gamma C_\Gamma \mathcal{D}_\Gamma^{2,1}(x = 2, T). \tag{C.13}$$

Each of these layers is contracted by various combinations of the $+, -, \perp$ and 0 states. We introduce a short-hand for each of these contractions, this is given below,

$$\langle +| \begin{bmatrix} \vdots \\ \\ \vdots \end{bmatrix} |+\rangle = \begin{bmatrix} \vdots \\ + \\ \vdots \end{bmatrix}, \qquad \langle -| \begin{bmatrix} \vdots \\ \\ \vdots \end{bmatrix} |-\rangle = \begin{bmatrix} \vdots \\ - \\ \vdots \end{bmatrix},$$

$$q\langle -| \begin{bmatrix} \vdots \\ \\ \vdots \end{bmatrix} |0\rangle = \begin{bmatrix} \vdots \\ 0 \\ \vdots \end{bmatrix}, \qquad q\langle \perp| \begin{bmatrix} \vdots \\ \\ \vdots \end{bmatrix} |+\rangle = \begin{bmatrix} \vdots \\ \perp \\ \vdots \end{bmatrix}. \tag{C.14}$$

One further short-hand we use is $\langle Z_t^- | \equiv \langle -| Z(t)^{\otimes 2}$ and $|Z_t^+\rangle \equiv Z(t)^{\otimes 2} \langle +|$. The contractions (of the unitary layers) at site 1 now take the more readable form,

$$\text{site 1} \quad q\langle Z_0^-| \begin{bmatrix} \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \\ \end{bmatrix} |Z_\tau^+\rangle \begin{bmatrix} \perp \end{bmatrix} \begin{bmatrix} + \end{bmatrix} \cdots \begin{bmatrix} + \end{bmatrix}. \tag{C.15}$$
$$\quad\quad\quad\quad\quad 1 \quad \tau-1 \quad\quad \tau \quad \tau+1 \quad T-1$$

The short-hand version for site x is found similarly, but with $T - \tau - 1$ '$-$' contractions, followed by a '0' contraction on layer $T - \tau$, followed by the OTOC. We now also apply this short-hand to the contributions on site r, $1 < r < x$, in particular, this yields

$$\sum_{t_r=1}^{T-1} \text{site } r \begin{bmatrix} - \end{bmatrix} \cdots \begin{bmatrix} - \end{bmatrix} \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} + \end{bmatrix} \cdots \begin{bmatrix} + \end{bmatrix}, \tag{C.16}$$
$$\quad\quad\quad 1 \quad t_r-1 \quad t_r \quad t_r+1 \quad T-1$$

where we have discarded the decoration that never decorates the initial state, as previously discussed.

By keeping only the $\mathcal{O}(1/q^2)$ contributions to $\mathcal{D}^{2,1}(x \geqslant 2, T)$, have found a set of diagrams labelled by: τ, the length of each of the physical OTOCs; t_r for each site $r \in \{2, \ldots, x-1\}$, the positions the '0' contraction on site r. Setting $t_1 = \tau$ and $t_x = T - \tau$, we label each diagram by a sequence $(t_1, t_2, \ldots, t_{x-1}, t_x)$. An example diagram, labelled $(t_1, t_2, t_3, t_4, t_5, t_6) = (5, 3, 6, 2, 8, 7)$, that contributes to $\mathcal{D}^{2,1}(x = 6, T = 12)$ is given below

In fact, for $x > 2$, only contributions where $t_1 \leqslant t_2 < t_3 < \ldots < t_x$ are non-zero. One of the following motifs must appear in any diagram not satisfying this property, each of which is

zero,

$$
\begin{vmatrix} \vdots \\ - \\ 0 \\ - \\ \vdots \end{vmatrix} = q \begin{matrix} \langle -| \\ \langle -| \\ \langle -| \end{matrix} \begin{vmatrix} \vdots \\ \\ \\ \vdots \end{vmatrix} \begin{matrix} |-\rangle \\ |0\rangle \\ |-\rangle \end{matrix} = \begin{matrix} \langle -| \\ q\langle -| \\ \langle -| \end{matrix} \begin{bmatrix} \vdots \\ \\ \vdots \end{bmatrix} \begin{matrix} |-\rangle \\ |0\rangle \\ |-\rangle \end{matrix} = \frac{\langle -| \Box |-\rangle}{\langle -| \Box |-\rangle} = 0 , \qquad \begin{vmatrix} \vdots \\ + \\ \bot \\ + \\ \vdots \end{vmatrix} = q \begin{matrix} \langle +| \\ \langle \bot| \\ \langle +| \end{matrix} \begin{vmatrix} \vdots \\ \\ \\ \vdots \end{vmatrix} \begin{matrix} |+\rangle \\ |+\rangle \\ |+\rangle \end{matrix} = \frac{\langle +| \Box |+\rangle}{\langle +| \Box |+\rangle} = 0 ,
$$

$$
\begin{vmatrix} \vdots \\ 0 \\ 0 \\ - \\ \vdots \end{vmatrix} = \begin{matrix} q\langle -| \\ q\langle -| \\ \langle -| \end{matrix} \begin{vmatrix} \vdots \\ \\ \\ \vdots \end{vmatrix} \begin{matrix} |0\rangle \\ |0\rangle \\ |-\rangle \end{matrix} = \frac{q\langle -| \Box |0\rangle}{\langle -| \Box |-\rangle} = 0 .
$$

(C.17)

In the non-zero diagrams (i.e., those for which $t_1 \leqslant t_2 < t_3 < \ldots < t_x$) the contracted unitary layers collapse into contractions of only a short portion of the full layer this follows from the brick property equation (70). We demonstrate this process by collapsing a semi-infinite domain of '+' contractions below,

$$
\begin{vmatrix} \vdots \\ + \\ + \\ + \\ \vdots \end{vmatrix} = \begin{matrix} \langle +| \Box |+\rangle \\ \langle +| \Box |+\rangle \\ \langle +| \Box |+\rangle \end{matrix} = \langle +| \Box |+\rangle = \begin{vmatrix} + \\ \vdots \end{vmatrix} .
$$

(C.18)

We collapse the '−' domain in the same way. Ultimately, every layer reduces to one of the motifs below,

$$
\boxed{\begin{matrix} + \\ - \end{matrix}} = 1 + g, \qquad \boxed{\begin{matrix} + \\ 0 \\ - \end{matrix}} = \boxed{\begin{matrix} + \\ \bot \\ - \end{matrix}} = -g, \qquad \boxed{\begin{matrix} + \\ \bot \\ 0 \\ - \end{matrix}} = g, \qquad M(t) \equiv \boxed{\begin{matrix} + \\ - \end{matrix}}_t ,
$$

(C.19)

where $M(t)$ is a single site operator. Each diagram as a whole decomposes into a products of the motifs below, which we use to introduce a compact diagrammatic notation. We also give the numerical value of each motif.

$$
\bullet \;=\; \boxed{\begin{matrix} + \\ 0 \\ - \end{matrix}} = -g, \qquad \circ \;=\; \boxed{\begin{matrix} + \\ - \end{matrix}} = 1 + g,
$$

$$
\underset{t}{\bullet}\!\!\sim\!\!\!\sim\!\!\!\sim\!\!\underset{t'}{\bullet} \;=\; q\langle Z_t^-| \boxed{\begin{matrix}+\\-\end{matrix}} \cdots \boxed{\begin{matrix}+\\-\end{matrix}} |Z_{t'}^+\rangle \boxed{\begin{matrix}+\\\bot\\-\end{matrix}} = \boxed{\begin{matrix}+\\0\\-\end{matrix}} q\langle Z_t^-| \boxed{\begin{matrix}+\\-\end{matrix}} \cdots \boxed{\begin{matrix}+\\-\end{matrix}} |Z_{t'}^+\rangle = -g\lambda(t'-t),
$$

$$
\underset{t}{\bullet}\!\!\sim\!\!\!\sim\!\!\!\sim\!\!\underset{t'}{\bullet} \;=\; q\langle Z_t^-| \boxed{\begin{matrix}+\\-\end{matrix}} \cdots \boxed{\begin{matrix}+\\-\end{matrix}} |Z_{t'}^+\rangle \boxed{\begin{matrix}+\\\bot\\0\\-\end{matrix}} = g\lambda(t'-t) = -\; \underset{t}{\bullet}\!\!\sim\!\!\!\sim\!\!\!\sim\!\!\underset{t'}{\bullet} ,
$$

(C.20)

where we have defined $\lambda(t - t') = q \langle Z_t^- | M(t+1) \ldots M(t'-1) | Z_{t'}^+ \rangle$. To give an example of the correspondence between the contracted unitary layer diagrams and this new diagrammatic notation, consider the diagram corresponding to the sequence $(t_1, \ldots, t_4) = (3, 3, 6, 7)$ below

$$(C.21)$$

For $x > 2$, the diagrams come in two qualitatively different types: (1) $t_1 < t_2$ and (2) $t_1 = t_2$. All type 1 diagrams have the following diagrammatic form

$$(C.22)$$

where $T - t_x = t_1$. Whereas, type 2 diagrams have the form

$$(C.23)$$

For $x = 2$, we will find diagrams with similar motifs. We had found that for $x \geqslant 2$, the relevant diagrams are labelled by a sequence $(t_1, t_2, \ldots, t_x)$, with $t_x = T - t_1$ and where $t_1 = \tau$ is the length of the OTOCs. For $x = 2$ the diagrams are simply labelled by $(\tau, T - \tau)$. A complication for $x = 2$ is the fact that the OTOCs may overlap in time. This is because no matter where we position the '$\perp$' contraction of site 1 and '0' contraction of site 2, we can never encounter any of the vanishing motifs of equation (C.17), which in the case of $x > 2$ force $t_1 \leqslant t_2$.

To address this complication, we split $x = 2$ in three types of diagram: (1) $T - \tau > \tau$, the OTOCs do not overlap—the treatment of these diagrams is exactly the same the type-1 diagrams discussed for $x > 2$; (2) $\tau = T - \tau$, the OTOC's 'touch'—this is similar to the type-2 diagrams in $x > 2$; (3) $\tau > T - \tau$, the OTOCs overlap.

The touching OTOC contributions ($\tau = T - \tau$) have the following form/motif,

$$(C.24)$$

We give the overlapping OTOC contributions ($\tau > T - \tau$) the following compact notation,

$$ \tag{C.25} $$

We study these contributions in detail in the next section, appendix C.4.

We are interested in computing $\int dV \sum_x \mathcal{D}^{(2,1)}(x, T)$, we have seen earlier in this appendix that the $x = 0, 1$ contributions are $\mathcal{O}(1/q^3)$ (as are the $x < 0$ contributions, equation (89)). Therefore, at $\mathcal{O}(1/q^2)$, we need only sum over $x \geqslant 2$. Fortunately, there is an abundance of cancellation between these diagrams. We will cover some examples and then give the general result. Starting with the simplest, we compute $\int dV \sum_x \mathcal{D}^{(2,1)}(x, T)$ for $T = 2, 3$ and 4 explicitly (where the Haar average is implied, but not written below). The additional factor of g^2 in equation (C.7) has been divided through in the equations below.

We have used the third rule of equation (C.20) to cancel the two terms (associated with $x = 2$ and $x = 3$) for $T = 2$ and to cancel the second and fourth terms and the fifth and sixth terms for $T = 4$. Notice that the terms that remain after cancellation are all connected diagrams (i.e., the OTOCs either touch or overlap), all the diagrams with 'gaps' (i.e., where a vertical line can be drawn through them without intersecting a wobbly line, representing an OTOC) have conspired to cancel. This is no coincidence, it is a consequence of the fact that processes that contribute to Σ (and hence the corrections to v_B) must explore only the fast space, this is due to the Q projectors that project out all slow components at every time-step in Σ. Diagrams with a gap represent processes that take a detour to the slow space. We can see this by returning to the contracted unitary layer picture; take, for example, the $T = 3$ diagram associated with $(t_1 = 1, t_2 = 2)$, this is the first diagram in the $T = 3$ sum above. It is equivalently given by

$$ \tag{C.26} $$

Between the two unitary layers we have vectors that are clearly in $\mathcal{P}$. The only diagrams that have no gaps are those for $x = 1$ (which we have seen all vanish at $\mathcal{O}(1/q^2)$) and the overlapping or touching OTOC diagrams for $x = 2$. We will next evaluate the overlapping OTOC diagram contributions, before finally calculating the touching OTOC diagram contributions.

C.4. Overlapping OTOC diagrams

In this section we investigate the contributions to $\mathcal{D}^{(2,1)}(x = 2, T)$ that take the form of overlapping OTOCs. We name this contribution $\mathcal{D}_O^{(2,1)}(x = 2, T)$. These OTOC overlap diagrams are given in detail below for OTOC length τ and total diagram length T.

$$\int dV \mathcal{D}_O^{(2,1)}(x = 2, T) \approx g^2 \int dV \sum_{\tau=1}^{T-1} \; q\langle Z_0^- | \cdots | \, q\langle Z_{T-\tau}^- | \cdots | Z_\tau^+ \rangle \cdots | Z_T^+ \rangle$$

$$\tag{C.27}$$

It will be useful to define the single site operators $M_+(\tau)$ and $M_-(\tau')$ and repeat the definition for $M(t)$ seen in equation (C.19).

$$M_-(\tau) \equiv \boxed{\begin{matrix}+\\1\\-\end{matrix}}_{\tau} \;, \quad M_+(\tau') \equiv \boxed{\begin{matrix}+\\0\\-\end{matrix}}_{\tau'} \;, \quad M(t) \equiv \boxed{\begin{matrix}+\\-\end{matrix}}_{t} \;.$$

$$\tag{C.28}$$

The Floquet layers $T - \tau + 1$ through to $\tau - 1$ have been reduced to two site operators which has already been introduced in equation (112) and named $\mathcal{T}(t)$ for Floquet layer t. With these definitions we have

$$\int dV \mathcal{D}_O^{(2,1)}(2, T) \approx g^2 \int dV \sum_{\tau=1}^{T-1} \langle -|Z^{\otimes 2} - \bigcirc \cdots \bigcirc - \bigcirc \cdots - Z(\tau)^{\otimes 2}|+\rangle \langle -|Z(\tau')^{\otimes 2} \cdots - Z(T)^{\otimes 2}|+\rangle,$$

$$\tag{C.29}$$

where we have denoted the following,

$$-\bigcirc- = M(t), \quad -\bigcirc- = M_+(t), \quad -\bigcirc- = M_-(t), \quad \boxed{} = \mathcal{T}(t)$$

Using theorem 2 of section 6.1, the Haar average of a product of two physical OTOCs is found to $\mathcal{O}(1/q^2)$ to be

$$
\int dV \begin{matrix} \langle Z\Gamma_1^1 Z(T)\Gamma_2^{1\dagger} Z\Gamma_2^1 Z(T)\Gamma_1^{1\dagger} \rangle \\ \langle Z\Gamma_1^2 Z(T)\Gamma_2^{2\dagger} Z\Gamma_2^2 Z(T)\Gamma_1^{2\dagger} \rangle \end{matrix} = \frac{1}{q^2} \left(\delta^{\Gamma_1^1,\Gamma_{\bar{1}}^2} \delta^{\Gamma_{\bar{2}}^1,\Gamma_2^2} \delta^{\Gamma_2^1,\Gamma_{\bar{2}}^2} \delta^{\Gamma_{\bar{1}}^1,\Gamma_1^2} \right.
$$
$$
\left. + \delta^{\Gamma_1^1,\Gamma_{\bar{2}}^2} \delta^{\Gamma_{\bar{2}}^1,\Gamma_{\bar{1}}^2} \delta^{\Gamma_2^1,\Gamma_{\bar{1}}^2} \delta^{\Gamma_{\bar{1}}^1,\Gamma_2^2} \right)
$$
$$
+ \mathcal{O}(1/q^3). \tag{C.30}
$$

Each of the decoration delta constraints can be implemented as described in appendix B. In doing so, we sandwich each decoration layer with a wiring configuration labelled A for the first term in equation (C.30) and by a configuration labelled B for the second term. We then use the shorthand below.

$$
\int dV \; \begin{matrix} q\langle -|Z^{\otimes 2}\text{-}\square\text{-}\square\text{-} \cdots \text{-}\square\text{-}Z(\tau)^{\otimes 2}|+\rangle \\ q\langle -|Z^{\otimes 2}\text{-}\square\text{-}\square\text{-} \cdots \text{-}\square\text{-}Z(\tau)^{\otimes 2}|+\rangle \end{matrix} = \frac{1}{q^2} \sum_{a=A,B} {}_a\!\left(\!\!\square\!\!\right)_a \; {}_a\!\left(\!\!\square\!\!\right)_a \cdots {}_a\!\left(\!\!\square\!\!\right)_a, \tag{C.31}
$$

where the grey boxes are placeholders for the possible decorations at each layer. A grey box (labeled t) may decorated each of the incoming legs 1 and 2 with $Z(t)$ and each of the legs $\bar{1}$ and $\bar{2}$ with $Z(t)^*$. On the right-hand side, every super-leg carries a label a which labels one of the permutations A, B of the legs $1, \bar{1}, 2, \bar{2}$ given in equation (C.32) below,

$$
\frac{\quad a \quad}{} = -R_a\!-, \quad R_A = \asymp\!\!\asymp, \quad R_B = \nparallel \tag{C.32}
$$

Each super-leg is implicitly carrying a factor $1/q^2$. The contraction between decorations within a column is given more explicitly below

$$
{}_a\!\left(\begin{matrix} \Gamma' \\ \Gamma^* \end{matrix}\right)_a = \left(\begin{matrix} \Gamma' \\ R_a\Gamma^* R_a \end{matrix}\right) = \mathrm{Tr}(R^a\Gamma^\dagger R^a\Gamma')/q^4 = \langle R^a\Gamma R^a || \Gamma' \rangle. \tag{C.33}
$$

All of the onsite scrambling evolution can be dropped on the right-hand-side of equation (C.31) as the leg contractions are between decorations at the same time. In the case of equation (C.29), the OTOCs are off-set by $\tau' = T - \tau$ Floquet time-steps. These OTOCs can be brought into alignment by globally shift the time arguments in the OTOCs, then we can use equation (C.31). After Haar averaging, and shifting the OTOCs back to their original positions, the leg contractions will stretch over τ' steps, as shown below.

$$
\frac{g^2}{q^2} \sum_{s\geq 0} \sum_{t\geq 0} \delta^{s+2t+3,T} \sum_{a=A,B} {}_a\!\left(\overbrace{\bigcirc \cdots \bigcirc}^{t+1} \overbrace{\cdots}^{s} \cdots \underbrace{\bigcirc \cdots \bigcirc}_{t+1} \right)_a, \tag{C.34}
$$

where all scrambling evolution time arguments have been dropped. The new variables are related to those in equation (C.29) by $t = T - \tau - 1 \geqslant 0$ and $s = 2\tau - T - 1 \geqslant 0$. As an

example, we show the $t = 1$, $s = 2$ case,

$$(C.35)$$

where on the right-hand side we have introduced a shorthand which makes obvious the decomposition into two chains of tensor contractions. In general, the Haar average decomposes into $t + 1$ chains. By summing over $T \geqslant 1$, we are able to drop the delta constraint above, doing this sum is equivalent to calculating the Laplace transformed $\int dV \mathcal{D}^{2,1}(x = 2, T)$ at $z = 0$, i.e. $\int dV \mathcal{D}^{2,1}(x = 2, z = 0)$. This is sufficient for calculating v_B. Before we determine the decomposition for a general $s \geqslant 0$ and $t \geqslant 0$, we first define the following chains,

$$(C.36)$$

For any (s, t) There are two qualitatively distinct types of contribution to equation (C.34): (1) $t + 1 \equiv 0 \pmod{s + 1}$, a product of $t C^a$ chains and one C_{+-}^a chain; (2) otherwise, a product of $t - 1 C^a$ chains, one C_+^a chain and one C_-^a chain. For case 1, the C_{+-}^a chain is $n = \frac{s+1}{t+1} - 1 \geqslant 0$ transfer matrices long, i.e., $C_{+-}^a(n)$, while all t of C^a chains are $n + 1$ matrices long, $C^a(n + 1)$. For case 2 with $s + 1 = n(t + 1) + k$ for $1 \leqslant k \leqslant t$ and starting from the left, the first $k - 1$ M's are on chains of length $n + 1$ and terminate on an M. The kth M sits on a chain of length n and terminates on M_-, the following $t - k$ M's sit on chains length n that terminate on an M. Finally, the M_+ sits on a chain of length n and terminates on an M. All together, the contribution is $C_+^a(n) C_-^a(n) C^a(n + 1)^{k-1} C^a(n)^{t-k}$. This is summarised below,

$$\frac{g^2}{q^2} \sum_{a=A,B} \sum_{n \geqslant 0} \left[\sum_{t \geqslant 0} C_{+-}^a(n) C^a(n+1)^t + \sum_{t \geqslant 1} \right.$$

$$\left. \times \sum_{k=1}^{t} C_+^a(n) C_-^a(n) C^a(n+1)^{k-1} C^a(n)^{t-k} \right]. \tag{C.37}$$

In both cases, all but the n sums can be evaluated to give,

$$\int dV \mathcal{D}_O^{(2,1)}(x = 2, z = 0)$$

$$= \frac{g^2}{q^2} \sum_{a=A,B} \sum_{n \geqslant 0} \left[\frac{C_{+-}^a(n)}{1 - C^a(n+1)} + \frac{C_+^a(n) C_-^a(n)}{(1 - C^a(n))(1 - C^a(n+1))} \right]. \tag{C.38}$$

It remains to calculate the different chains. To do this we approach the problem as a transfer matrix problem, where $\mathcal{T}$ is the transfer matrix. Using the definition of $\mathcal{T}$ in equation (112),

the decoration decomposition of a two-site brick in equation (67) and definitions of M and $M^{\pm}$ in equation (C.28), $\mathcal{T}$ and each of M, M_+ and M_- have the following properties,

$$
\begin{array}{cc}
{}^{-R_a}\bigcap^{R_a^-} = \overline{\underset{-R_a^-\bigcup^{R_a^-}}{}} \,, & {}^{-}R_a^-M_{(\pm)}^-R_a^- = {}^-M_{(\pm)}^- \,.
\end{array}
\tag{C.39}
$$

Then, using the definition of the labelled super legs in equation (C.32), the chain $C^a(n)$ simplifies to the expression below,

$$
C^a(n) = \overbrace{\bigcup^{M}}\ \underset{n}{\underset{\longleftarrow \quad \longrightarrow}{\bigcup\bigcup}} \cdots \bigcup\overbrace{}^{M} \,.
\tag{C.40}
$$

Notice that the dependence on the label a vanished. This is true of all chains. Algebraically, these chains (now without leg labels) are equivalently given by

$$
\begin{aligned}
C(n) &= \langle M | \, \mathcal{T}^n \, | M \rangle, & C_+(n) &= \langle M_+ | \, \mathcal{T}^n \, | M \rangle \,, \\
C_+(n) &= \langle M | \, \mathcal{T}^n \, | M_- \rangle, & C_{+-}(n) &= \langle M_+ | \, \mathcal{T}^n \, | M_- \rangle \,,
\end{aligned}
\tag{C.41}
$$

where these angles braces reflect the trace inner product for tensors with input and output super legs $l = (1, \overline{1}, 2, \overline{2})$.

$$
\langle B || A \rangle = \mathrm{Tr}(B^\dagger A)/q^4 = \binom{A}{B^*} \,.
\tag{C.42}
$$

Equipped with this, we can now write, for the contribution due to overlapping OTOC diagrams, the following

$$
\int \mathrm{d}V \mathcal{D}_O^{(2,1)}(x = 2, z = 0) = \frac{2g^2}{q^2} f(\varepsilon),
\tag{C.43}
$$

where $f(\varepsilon)$ is given by

$$
f(\varepsilon) = \sum_{n \geqslant 0} \left[\frac{C_{+-}(n)}{1 - C(n+1)} + \frac{C_+(n) C_-(n)}{(1 - C(n))(1 - C(n+1))} \right],
\tag{C.44}
$$

where the chains are implicitly dependent on ε. To evaluate this sum, we must understand the space that the transfer matrix acts on. Each of the legs 1, $\overline{1}$, 2 and $\overline{2}$ may be either undecorated or carry a Z decoration. This means that the state our state space is dimension 2^4 and $\mathcal{T}$ is a 16×16 matrix. Because all the $M_{(\pm)}$ are even in the number of Z decorations, and $\mathcal{T}$ preserves decoration parity, we are able to reduce the state space to those states with an even number of Z decorations only, i.e., 8 states. It will be useful to use the basis below,

$$
\begin{aligned}
|\mathbb{1}\rangle &= \overset{=}{=} & |S_{1\overline{2}}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{-z-}{=}} + \overset{=}{\underset{-z-}{=}}\right] & |S_{1\overline{1}}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{=}{=}} + \overset{=}{\underset{-z-}{=}}\right] & |S_{12}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{=}{=}} + \overset{=}{\underset{-z-}{=}}\right] \\
|4\rangle &= \overset{-z-}{\underset{-z-}{=}} & |A_{1\overline{2}}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{=}{=}} - \overset{=}{\underset{-z-}{=}}\right] & |A_{1\overline{1}}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{=}{=}} - \overset{=}{\underset{-z-}{=}}\right] & |A_{12}\rangle &= \frac{1}{\sqrt{2}}\left[\overset{-z-}{\underset{=}{=}} - \overset{=}{\underset{-z-}{=}}\right] \,.
\end{aligned}
\tag{C.45}
$$

Let $\mathcal{S}$ be the space spanning $\{|1\rangle, |S_{1\bar{1}}\rangle, |S_{1\bar{2}}\rangle, |S_{12}\rangle, |4\rangle\}$ and $\mathcal{A}$ be the space spanning $\{|A_{1\bar{1}}\rangle, |A_{1\bar{2}}\rangle, |A_{12}\rangle\}$ One can easily check that $|M\rangle$, $|M_+\rangle$ and $|M_-\rangle$ are all in $\mathcal{S}$. Explicitly, with $u(\varepsilon) = \sin(\varepsilon)^2$

$$|M\rangle = (u^4 + (1-u)^4)\,|1\rangle + 2u^2(1-u)^2\,|4\rangle$$

$$- (1+g)\frac{g}{\sqrt{2}}\left(|S_{1\bar{1}}\rangle + |S_{1\bar{2}}\rangle\right) + \frac{g^2}{\sqrt{2}}\,|S_{12}\rangle, \tag{C.46}$$

$$|M_+\rangle = -\frac{g(1+g)}{\sqrt{2}}\left(|S_{12}\rangle - |S_{1\bar{2}}\rangle\right) + \frac{g^2}{2}\left(\sqrt{2}\,|S_{1\bar{2}}\rangle - |1\rangle - |4\rangle\right), \tag{C.47}$$

$$|M_-\rangle = -\frac{g(1+g)}{\sqrt{2}}\left(|S_{12}\rangle - |S_{1\bar{1}}\rangle\right) + \frac{g^2}{2}\left(\sqrt{2}\,|S_{1\bar{1}}\rangle - |1\rangle - |4\rangle\right). \tag{C.48}$$

$\mathcal{T}$ is block diagonal in the subspaces $\mathcal{S}$ and $\mathcal{A}$, and since all $|M_{(\pm)}\rangle$ lie in $\mathcal{S}$, we are able to restrict our considerations to this five dimensional space only. In this restricted space with basis order $(1, 4, S_{1,\bar{1}}, S_{1,\bar{2}}, S_{1,2})$, $\mathcal{T}$ is given by the product $\mathcal{T} = \tilde{T}_-\tilde{U}\tilde{T}_+$, where

$$\tilde{U} = \begin{pmatrix} (1-u)^2 & 0 & 0 & 0 & 0 \\ 0 & u^2 & 0 & 0 & 0 \\ 0 & 0 & -\frac{g}{2} & 0 & 0 \\ 0 & 0 & 0 & -\frac{g}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{g}{2} \end{pmatrix} \tag{C.49}$$

$$\tilde{T}_- = \begin{pmatrix} (1-u)^2 & u^2 & -\frac{g}{\sqrt{2}} & 0 & 0 \\ u^2 & (1-u)^2 & -\frac{g}{\sqrt{2}} & 0 & 0 \\ -\frac{g}{\sqrt{2}} & -\frac{g}{\sqrt{2}} & g+1 & 0 & 0 \\ 0 & 0 & 0 & g+1 & -g \\ 0 & 0 & 0 & -g & g+1 \end{pmatrix} \tag{C.50}$$

$$\tilde{T}_+ = \begin{pmatrix} (1-u)^2 & u^2 & 0 & -\frac{g}{\sqrt{2}} & 0 \\ u^2 & (1-u)^2 & 0 & -\frac{g}{\sqrt{2}} & 0 \\ 0 & 0 & g+1 & 0 & -g \\ -\frac{g}{\sqrt{2}} & -\frac{g}{\sqrt{2}} & 0 & g+1 & 0 \\ 0 & 0 & -g & 0 & g+1 \end{pmatrix}. \tag{C.51}$$

Four of the eigenvalues of $\mathcal{T}$ are bounded by $|\lambda_{1,2,3,4}| \leqslant |g(\varepsilon)|(1 - 2|g(\varepsilon)|)/2$. The largest eigenvalue (for all ε) is bounded by $|\lambda_5| \leqslant (1 - |g(\varepsilon)|)^3$.

In section 4 we showed that $\langle v_{\mathrm{B}}(\varepsilon)\rangle$ must have the symmetries $\varepsilon \to -\varepsilon$ and $\varepsilon \to \pi/2 + \varepsilon$. All analytically computed $\mathcal{O}(1/q^2)$ contributions (the $(a, b) = (4, 4)$ contribution and the touching OTOC contribution) are found to respect this symmetry, as does the $\mathcal{O}(1)$ contribution from Ω. We therefore conclude that $f(\varepsilon)$ must also have this symmetry. Additionally, we know that $f(\varepsilon)$ is a function of $\sin(\varepsilon)^2$ only (the transfer matrix and the initial and final vectors in our transfer matrix calculation are functions of $\sin(\varepsilon)^2$ only). Together with the symmetry requirements, this means that $f(\varepsilon)$ is in fact a function of $s(\varepsilon) = \sin(\varepsilon)^2 \cos(\varepsilon)^2$.

We evaluate the sum in equation (C.44) analytically for small ε, finding $f(\varepsilon) \approx \varepsilon^2/7$. Knowing that f is a function of $s(\varepsilon)$ only, we use this to factorise $f(\varepsilon) = \frac{1}{7}s(\varepsilon)w(s(\varepsilon))$, where $w(\varepsilon)$ is a function that approaches 1 as $\varepsilon \to 0$. The transfer matrix dramatically simplifies at the point $s(\varepsilon) = 1/4$. Where all but one eigenvalue is zero. We find analytically that $f(\varepsilon) \propto (1 - 4s)^2$ around this point. This suggests a further factorisation $f(\varepsilon) = \frac{1}{7}s(\varepsilon)(1 - 4s)^2 p(s(\varepsilon))$. We find a very good quadratic polynomial approximation for $p(\varepsilon) = 1 + as + bs^2$, with $a = 6.8$ and $b = 16.1$.

$$f(\varepsilon) = \frac{1}{7}s(\varepsilon)(1 - 4s(\varepsilon))^2(1 + as(\varepsilon) + bs(\varepsilon)^2). \tag{C.52}$$

We also find numerically that $\int dV \mathcal{D}_O^{2,1}(x = 2, T)$ decays exponentially quickly, with a decay rate bounded below by $\gamma_O(\varepsilon) \geqslant 6s(\varepsilon)$.

C.5. Touching OTOC diagrams

We will now sum up all of the touching OTOC diagrams. These are given in equation (C.24) with $T = 2\tau$ so that the two OTOC have the same length. We name this contribution $\mathcal{D}_{\text{Touch}}^{(2,1)}(x = 2, T)$.

We are only interested in the $z \to i0^+$ limit, this is given by

$$\int dV \mathcal{D}_{\text{Touch}}^{(2,1)}(k = 0, z = 0) = g^2 \int dV \sum_{t=1}^{\infty} g\lambda(t)^2 + \mathcal{O}(1/q^3). \tag{C.53}$$

The factor $g\lambda(t)^2$ is the contribution from the diagrams equation (C.24). The Haar average of a product of OTOCs given in equation (C.31) is deployed again to say that

$$\int dV \lambda(t)^2 = \frac{2}{q^2} \langle M| |M\rangle^{t-1}. \tag{C.54}$$

This gives

$$\int dV \sum_{t=1}^{\infty} g^3 \lambda(t)^2 = \frac{2g^3}{q^2} \frac{1}{1 - \langle M| |M\rangle} = -\frac{2g^2}{q^2}\nu(\varepsilon), \tag{C.55}$$

where $\langle M| |M\rangle = 1 - 8s(1 - 2s)(1 - s(1 - 2s))$, $s(\varepsilon) = \sin(\varepsilon)^2 \cos(\varepsilon)^2$ and $\nu(\varepsilon) = -g/(1 - \langle M| |M\rangle)$. Touching OTOC diagrams decay exponentially with a decay rate $\gamma_{\text{Touch}}(\varepsilon) = \log(\langle M| |M\rangle^{-1}) \geqslant 8s(\varepsilon)$.

C.6. Summary

Recalling equation (89), the contributions from $x < 0$ are $\mathcal{O}(1/q^4)$ or smaller. Piecing together the $\mathcal{O}(1/q^2)$ corrections, we find $\int dV \mathcal{D}^{(2,1)}(k = 0, z = 0)$ is given by

$$\int dV \mathcal{D}^{(2,1)}(k = 0, z = 0) = \frac{2g^2}{q^2}(f(\varepsilon) - \nu(\varepsilon)) + \mathcal{O}(1/q^3). \tag{C.56}$$

Appendix D. Remaining (*a*, *b*)

D.1. (1, 2)

$$
\mathcal{D}_\Gamma^{1,2}(x \geq 0, T) = g^2 \times
\begin{bmatrix}
 & \vdots & & \vdots & \\
\text{site } 0 & \langle \phi_+| & & |+\rangle & \\
\text{site } 1 & q\langle -| & & |+\rangle & \\
 & \vdots & \Gamma & \vdots & \\
 & q\langle -| & & |+\rangle & \text{site } x \\
 & \langle -| & & |\phi_-(T)\rangle & \text{site } x+1 \\
 & \vdots & & \vdots &
\end{bmatrix}
\tag{D.1}
$$

Here, we can simply use equation (98) to see that both sites 0 and $x + 1$ contribute a product of non-trivial correlation functions plus terms of size $1/q^2$. The Haar average of this is $\mathcal{O}(1/q^4)$.

$$
\int \mathrm{d}V \mathcal{D}_\Gamma^{1,2}(x \geqslant 0, T) = \mathcal{O}(1/q^4).
\tag{D.2}
$$

D.2. (3, 3)

- $x = 0$:

$$
\mathcal{D}_\Gamma^{3,3}(x = 0, T) = g^2 \times
\begin{bmatrix}
\text{site } 0 & \langle +|Z^{\otimes 2} & Z(T)^{\otimes 2}|+\rangle \\
\text{site } 1 & \langle -|Z^{\otimes 2} & Z(T)^{\otimes 2}|-\rangle
\end{bmatrix} \Gamma
\left(\prod_{r<0} \langle +|\Gamma^r|+\rangle \right) \left(\prod_{r>1} \langle -|\Gamma^r|-\rangle \right)
\tag{D.3}
$$

In this case, both contours on site 0 and on site 1 are non-trivially decorated, contributing a total of four non-trivial correlators. The Haar average is then $\mathcal{O}(1/q^4)$.

- $x \geqslant 1$: in this case, each site 0 and $x + 1$ contribute a product of non-trivial correlators and therefore, as above, are $\mathcal{O}(1/q^4)$ or smaller.

All together

$$
\int \mathrm{d}V \mathcal{D}_\Gamma^{1,2}(x \geqslant 0, T) = \mathcal{O}(1/q^4).
\tag{D.4}
$$

D.3. (3, 4) *and* (4, 3)

- $x = 0$:

$$
\mathcal{D}_\Gamma^{3,4}(0, T) = -ihg \times
\begin{bmatrix}
\text{site } 0 & \langle +|Z^{\otimes 2}\Gamma^0 K(T)|+\rangle \\
\text{site } 1 & \langle -|Z^{\otimes 2}\Gamma^1 K(T)|-\rangle
\end{bmatrix}
\left(\prod_{r<0} \langle +|\Gamma^r|+\rangle \right) \left(\prod_{r>1} \langle -|\Gamma^r|-\rangle \right)
\tag{D.5}
$$

Both site 0 and site 1 contribute two non-trivial correlators. The Haar average of this term is $\mathcal{O}(1/q^4)$ or smaller.

- $x \geqslant 1$: each of the sites 0 and $x + 1$ contribute a product of two non-trivial correlators. Each of the sites 1 and x contribute a non-trivial correlator. If $x = 1$, then this is only five non-trivial correlators, otherwise it is six. The Haar average is then $\mathcal{O}(1/q^5)$.

All together,

$$\int \mathrm{d}V \mathcal{D}_\Gamma^{4,4}(x \geqslant 0, T) = \mathcal{O}(1/q^4). \tag{D.6}$$

D.4. (a, b), where either $a \in 1, 2$ and $b \in 3, 4$ or the converse

In these case, equation (98) is again enough to show that $\int \mathrm{d}V \mathcal{D}_\Gamma^{a,b}(x, T) = \mathcal{O}(1/q^3)$.

Appendix E. Model variation: independent scramblers $V_{x,t}$

We can carry out a similar calculation for a version of the model with independently distributed scramblers $V_{x,t}$ for each site and at each layer of unitaries U_t (analogous to the Floquet layer in the Floquet model), i.e., we break both spatial and temporal translation symmetry. We start with, for large t, the fact $\sum_x x\rho(x, t) = v_\mathrm{B} t$ to write

$$
\begin{aligned}
v_\mathrm{B} &= \lim_{t \to \infty} \sum_x x \left[\rho(x, t) - \rho(x, t - 1)\right] \\
&= \lim_{t \to \infty} \lim_{k \to 0} \sum_x \mathrm{i}\partial_k \mathrm{e}^{-\mathrm{i}kx} \left[\rho(x, t) - \rho(x, t - 1)\right] \\
&= \lim_{t \to \infty} \lim_{k \to 0} \mathrm{i}\partial_k \left[\rho(k, t) - \rho(k, t - 1)\right].
\end{aligned} \tag{E.1}
$$

Using the definition of $\rho(k, n)$ we write the following

$$
\begin{aligned}
\rho(k, n) &= \left(W^k | U_0(P + Q)U_1 \ldots U_n | W^k\right) \\
&= \left(W^k | U_0 | W^k\right) \rho(k, n - 1) + H(k, n) \\
&= (1 + \Omega(k))\rho(k, n - 1) + H(k, n),
\end{aligned} \tag{E.2}
$$

where $H(k, n)$ is given by

$$
\begin{aligned}
H(k, n) &= \left(W^k | U_0 Q U_1 \ldots U_n | W^k\right) \\
&= \left(W^k | U_0 Q U_1 \ldots U_{n-1}(P + Q)U_n | W^k\right) \\
&= (1 + \Omega(k))H(k, n - 1) + \sigma(k, n),
\end{aligned} \tag{E.3}
$$

where $\sigma(k, n)$ is as previously defined, but with the independently distributed scramblers.

$$\sigma(k, n) = \left(W^k | L_0 Q U_1 \ldots U_{n-1} Q L_n | W^k\right). \tag{E.4}$$

Using equation (E.3) inductively and the fact that $H(k, n = 0) = 0$ we find

$$\rho(k, n) - \rho(k, n - 1) = \Omega(k)\rho(k, n - 1) + \sum_{m=1}^{n} (1 + \Omega(k))^{n-m-1}\sigma(k, m). \tag{E.5}$$

Using this in equation (E.1), we find

$$
v_{\mathrm{B}} = \lim_{n\to\infty} \lim_{k\to 0} \mathrm{i}\partial_k \left[\Omega(k)\rho(k, n-1) + \sum_{m=1}^{n} (1 + \Omega(k))^{n-m-1}\sigma(k, m) \right]. \tag{E.6}
$$

Then writing $\Omega(k) = v_0(1 - \mathrm{e}^{-\mathrm{i}k})$ where $v_0 = |g(\varepsilon)|$ is the $q \to \infty$ butterfly velocity, and noticing that $\lim_{k\to 0}\partial_k (\Omega(k)\rho(k, n)) = \mathrm{i}v_0$, we find

$$
\delta v_{\mathrm{B}} \equiv v_{\mathrm{B}} - v_0 = \lim_{n\to\infty} \lim_{k\to 0} \mathrm{i}\partial_k \sum_{m=1}^{n} (1 + \Omega(k))^{n-m-1}\sigma(k, m). \tag{E.7}
$$

Analogous to the definition in equation (85) for the Floquet model, we define the quantity,

$$
\mathcal{D}(x \to y, n) = q^{y-x} \langle F^x | L_0 Q U_1 \dots U_{n-1} Q L_n | F^y \rangle \tag{E.8}
$$

then $\sigma(k, n)$ is given by

$$
\sigma(k, n) = \frac{1}{L(1 - q^{-2})}\sum_{x,y}\mathrm{e}^{\mathrm{i}k(x-y)}[\mathcal{D}(x \to y, n) - \mathcal{D}(x \to y - 1, n)
$$

$$
- \frac{1}{q^2}(\mathcal{D}(x - 1 \to y, n) - \mathcal{D}(x - 1 \to y - 1, n))]. \tag{E.9}
$$

The circuit averaged $\mathcal{D}(x \to y, n)$ is translationally invariant $\langle \mathcal{D}(x \to y, n) \rangle = \langle \mathcal{D}(y - x, n) \rangle$ where $\mathcal{D}(x, n) = \mathcal{D}(0 \to x, n)$. The circuit average of $\sigma(k, n)$ is given by

$$
\langle \sigma(k, n) \rangle = \eta(k)(1 - \mathrm{e}^{-\mathrm{i}k})\sum_{x}\mathrm{e}^{-\mathrm{i}kx}\langle \mathcal{D}(x, n) \rangle = \eta(k)(1 - \mathrm{e}^{-\mathrm{i}k})\langle \mathcal{D}(k, n) \rangle. \tag{E.10}
$$

The circuit averaging of equation (E.7) is given by

$$
\langle \delta v_{\mathrm{B}} \rangle = \lim_{n\to\infty} \lim_{k\to 0} \mathrm{i}\partial_k \sum_{m=1}^{n} (1 + \Omega(k))^{n-m-1}\eta(k)(1 - \mathrm{e}^{-\mathrm{i}k})\langle \mathcal{D}(k, m) \rangle
$$

$$
= -\sum_{m=1}^{\infty} \langle \mathcal{D}(k = 0, m) \rangle = -\langle \mathcal{D}(k = 0, z = 0) \rangle. \tag{E.11}
$$

This is same expression for the circuit averaged correction to v_{B} that we found using the MMF in equation (121). We calculate $\langle \mathcal{D}(k = 0, z = 0) \rangle$ in the same way as in section 6.2. However, the two $\mathcal{O}(1/q^2)$ contributions found for the Floquet model (i.e., the so-called (2, 1) and (4, 4) terms) relied on correlations between different sites and on the discrete time translation symmetry. With independently distributed scramblers, such terms are smaller than $\mathcal{O}(1/q^2)$. This gives the result $\langle \delta v_{\mathrm{B}} \rangle = \mathcal{O}(1/q^3)$ which leads to

$$
\langle v_{\mathrm{B}} \rangle = \frac{1 - \cos(4\varepsilon)}{4} + \mathcal{O}(1/q^3). \tag{E.12}
$$

We can employ the same calculation above for the variant of the model with spatial translation symmetry but no time translation symmetry (i.e., independently random scramblers between time-steps) in order to find $\langle \delta v_{\mathrm{B}} \rangle = \delta v_S(\varepsilon) + \mathcal{O}(1/q^3)$ where $\delta v_S(\varepsilon)$ is as given in equation (122).

$$\langle v_\mathrm{B} \rangle = \frac{1 - \cos(4\varepsilon)}{4} + \delta v_S(\varepsilon) + \mathcal{O}(1/q^3), \qquad \delta v_S(\varepsilon) = \frac{1}{q^2} \frac{1 + 5s - 4s^2}{1 - s - 3s^2}. \tag{E.13}$$

Comparing these result to the result of equation (121) allows us to see that the $\mathcal{O}(1/q^2)$ corrections only arise when there is spatial translation symmetry. Moreover, this enables us to identify the δv_F contribution (see equation (122)) with the presence of time translation symmetry.

E.1. The $\varepsilon \to 0$ limit in the absence of time-translation symmetry

In section 6.2.4, we identified all the contributions to $\sigma(x = 0, t)$ that contained the product of two non-trivial correlators and used theorem 2 to evaluate these contributions to $\mathcal{O}(1/q^2)$. We were unable to extract the correct $\varepsilon \to 0$ behaviour of these contributions due to the appearance of ε in the denominator (after summing over the time t). Luckily, for the model variant with independently random scramblers at each time-step, we are able to accurately sum over all 'two-correlator' contributions, after circuit averaging. While this is ignores correlations between Floquet layers, it demonstrated the failure of naively expanding the memory matrix in $1/q$ before taking the $\varepsilon \to 0$ limit and the resolution by summing over all contributions.

We identify the contributions that have the form of two non-trivial correlators in the same way as done in section 6.2.4. We repeat this below for ease of reading.

$$\mathcal{D}^{4,4}(x = 0, T) = -h^2 \sum_{i,j \in \{1,2\}} \frac{\langle +|Z_i}{\langle -|Z_j} \boxed{1}\boxed{2} \cdots \boxed{T-1} \frac{Z(T)_i|+\rangle}{Z(T)_j|-\rangle} + \cdots, \tag{E.14}$$

where, so as to ensure that only sites 0 and 1 contribute non-trivial correlators, we have sandwiched each layer by $\langle +|$ ($\langle -|$) and $|+\rangle$ ($|-\rangle$) for sites $x < 0$ ($x > 1$) to leave the two site operators $\mathcal{T}(t)$ introduced in equation (112) and represented above in equation (E.14) by the rectangle blocks. The second ellipsis represents terms that contain at least three non-trivial correlators. Using the same argument as was used in section 6.2.4, we see that the $(i, j) = (1, 1)$ and $(2, 2)$ choices give identical contributions and the $(1, 2)$ and $(2, 1)$ contributions are the complex conjugate of the (i, i) contributions, equation (114). Therefore, we are free to study only the $(i, j) = (1, 1)$ case. We insert projectors ensure that the $2, \bar{1}$ contour of site 0 and the $2, \bar{2}$ contour of site 1 remain undecorated (so as to ensure that there are only two non-trivial correlators). The contribution is given below,

$$\tag{E.15}$$
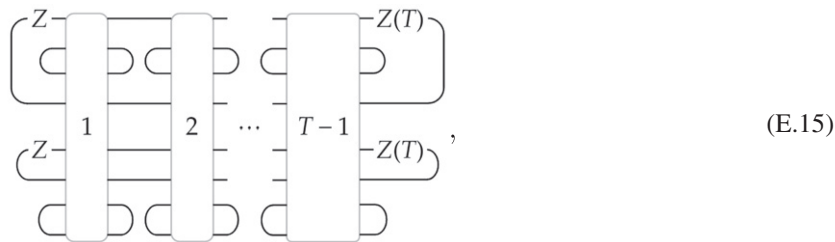
where each closed loops is associated with a normalising factor of $1/q$ that is suppressed in the diagram. The contracted single time-step unitaries are now just operators on four indices. We denote the contribution to $\mathcal{D}^{4,4}(x = 0, T)$ that includes only two non-trivial contours as

$\mathcal{D}_2^{4,4}(x = 0, T)$. Separating out the scrambling part of the unitaries we find

$$\mathcal{D}_2^{4,4}(x = 0, T) = -\frac{2h^2}{q^2} \quad \cdots \quad + \text{c.c,}$$

(E.16)

where each of the pink bricks is given by the following contraction of $\mathcal{T}$.

$$Y = \quad = \quad \times \frac{1}{q^2}.$$

(E.17)

Circuit averaging this tensor contraction is easy as each unitary only appears twice. We use the following result for the Haar average of two V's and two $V^\dagger$'s.

$$\int dV \quad = \frac{1}{q^2 - 1} \left[\quad + \quad \right] - \frac{1}{q^3 - q} \left[\quad + \quad \right] = K,$$

(E.18)

where K is a projector. Using the definitions of $|+\rangle$, $|-\rangle$, $|0\rangle$ and $|\perp\rangle$, K is given algebraically as

$$K \equiv \frac{|-\rangle \langle \perp|}{1 - q^{-2}} + \frac{|+\rangle \langle 0|}{1 - q^{-2}} = \frac{|\perp\rangle \langle -|}{1 - q^{-2}} + \frac{|0\rangle \langle +|}{1 - q^{-2}}.$$

(E.19)

Using this, we write

$$\langle \mathcal{D}_2^{4,4}(x = 0, T) \rangle = -2h^2 \langle -| Z^{\otimes 2} (KYK)^{T-1} Z^{\otimes 2} |-\rangle + \text{c.c}$$

$$= -\frac{2h^2}{q^2} \langle 0| (KYK)^{T-1} |0\rangle + \text{c.c.}$$

(E.20)

The matrix KYK is a 2×2 matrix acting on the space of wirings $\text{Span}\{|+\rangle, |-\rangle\}$. Choosing the orthormal basis $|S, A\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)$. KYK is given by

$$KYK = \frac{\xi}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1+g}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{1+g+ih}{q} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(E.21)

By diagonalising this, we find

$$\sum_{n=0}^{\infty} \frac{1}{q^2} \langle 0| (KYK)^n |0\rangle = \frac{2}{1 + 8q^2 \varepsilon^2 + \mathcal{O}(q^2 \varepsilon^3, q\varepsilon^2)}.$$

(E.22)

Using $h(\varepsilon) \approx \varepsilon$ for small ε, $\langle \mathcal{D}_2^{4,4}(x=0, z=0) \rangle$ is then given by

$$\langle \mathcal{D}_2^{4,4}(x=0, z=0) \rangle = -\frac{8\varepsilon^2}{1 + 8q^2\varepsilon^2 + \mathcal{O}(q^2\varepsilon^3, q\varepsilon^2)}. \tag{E.23}$$

Importantly, this contribution, which keeps account of all orders in q, vanishes as $\varepsilon \to 0$, unlike the $\mathcal{O}(1/q^2)$ contribution (naively) calculated in section 6.2.4, which (incorrectly) predicts $\lim_{\varepsilon \to 0} \langle \mathcal{D}^{4,4}(x=0, z=0) \rangle = -1/q^2$. However, this is still far from a complete account of contributions to $\langle \mathcal{D}(k=0, z=0) \rangle$ (for instance, contributions involving more than two correlators), but it does demonstrate that the apparent pathological behaviour as $\varepsilon \to 0$ is the result of a naive (and incorrect at small ε) counting of the $\mathcal{O}(1/q^2)$ contributions. Notice that for $q\varepsilon \gg 1$, the small ε results of section 6.2.4 (which predict $\langle \mathcal{D}^{4,4}(x=0, z=0) \rangle \approx -1/q^2$) are in agreement with the analysis presented here.

ORCID iDs

Ewan McCulloch ⓘ https://orcid.org/0000-0003-4521-7628

References

[1] Hayden P and Preskill J 2007 Black holes as mirrors: quantum information in random subsystems *J. High Energy Phys.* JHEP09(2007)120
[2] Sekino Y and Susskind L 2008 Fast scramblers *J. High Energy Phys.* JHEP10(2008)065
[3] Lashkari N, Stanford D, Hastings M, Osborne T and Hayden P 2013 Towards the fast scrambling conjecture *J. High Energy Phys.* JHEP04(2013)022
[4] Shenker S H and Stanford D 2014 Black holes and the butterfly effect *J. High Energy Phys.* JHEP03(2014)067
[5] Shenker S H and Stanford D 2014 Multiple shocks *J. High Energy Phys.* JHEP12(2014)046
[6] Shenker S H and Stanford D 2015 Stringy effects in scrambling *J. High Energy Phys.* JHEP05(2015)132
[7] Maldacena J, Shenker S H and Stanford D 2016 A bound on chaos *J. High Energy Phys.* JHEP08(2016)106
[8] Hartman T and Maldacena J 2013 Time evolution of entanglement entropy from black hole interiors *J. High Energy Phys.* JHEP05(2013)014
[9] Liu H and Suh S J 2014 Entanglement tsunami: universal scaling in holographic thermalization *Phys. Rev. Lett.* **112** 011601
[10] Liu H and Suh S J 2014 Entanglement growth during thermalization in holographic systems *Phys. Rev. D* **89** 066012
[11] Mezei M and Stanford D 2017 On entanglement spreading in chaotic systems *J. High Energy Phys.* JHEP05(2017)065
[12] Blake M 2016 Universal charge diffusion and the butterfly effect in holographic theories *Phys. Rev. Lett.* **117** 091601
[13] Dóra B and Moessner R 2017 Out-of-time-ordered density correlators in luttinger liquids *Phys. Rev. Lett.* **119** 026802
[14] Fagotti M and Calabrese P 2008 Evolution of entanglement entropy following a quantum quench: analytic results for thexychain in a transverse magnetic field *Phys. Rev. A* **78** 010306
[15] Gopalakrishnan S, Huse D A, Khemani V and Vasseur R 2018 Hydrodynamics of operator spreading and quasiparticle diffusion in interacting integrable systems *Phys. Rev. B* **98** 220303
[16] Lin C-J and Motrunich O I 2018 Out-of-time-ordered correlators in a quantum ising chain *Phys. Rev. B* **97** 144304
[17] Prosen T and Pižorn I 2007 Operator space entanglement entropy in a transverse Ising chain *Phys. Rev. A* **76** 032316

[18] Nahum A, Ruhman J, Vijay S and Haah J 2017 Quantum entanglement growth under random unitary dynamics *Phys. Rev.* X **7** 031016

[19] Nahum A, Vijay S and Haah J 2018 Operator spreading in random unitary circuits *Phys. Rev.* X **8** 021014

[20] von Keyserlingk C W, Rakovszky T, Pollmann F and Sondhi S L 2018 Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws *Phys. Rev.* X **8** 021013

[21] Khemani V, Vishwanath A and Huse D A 2018 Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws *Phys. Rev.* X **8** 031057

[22] Rakovszky T, Pollmann F and von Keyserlingk C W 2018 Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation *Phys. Rev.* X **8** 031058

[23] Brown W and Fawzi O 2012 Scrambling speed of random quantum circuits (arXiv:1210.6644)

[24] Chan A, De Luca A and Chalker J T 2018 Solution of a minimal model for many-body quantum chaos *Phys. Rev.* X **8** 041019

[25] Stanford D 2016 Many-body chaos at weak coupling *J. High Energy Phys.* JHEP10(2016)009

[26] Asplund C T, Bernamonti A, Galli F and Hartman T 2015 Entanglement scrambling in 2D conformal field theory *J. High Energy Phys.* JHEP09(2015)110

[27] Banerjee S and Altman E 2017 Solvable model for a dynamical quantum phase transition from fast to slow scrambling *Phys. Rev.* B **95** 134302

[28] Roberts D A, Stanford D and Streicher A 2018 Operator growth in the syk model *J. High Energy Phys.* JHEP06(2018)122

[29] Roberts D A and Swingle B 2016 Lieb–Robinson bound and the butterfly effect in quantum field theories *Phys. Rev. Lett.* **117** 091602

[30] Chowdhury D and Swingle B 2017 Onset of many-body chaos in the $O(N)$ model (arXiv:1703.02545)

[31] Aleiner I L, Faoro L and Ioffe L B 2016 Microscopic model of quantum butterfly effect: out-of-time-order correlators and traveling combustion waves *Ann. Phys., NY* **375** 378–406

[32] Calabrese P and Cardy J 2005 Evolution of entanglement entropy in one-dimensional systems *J. Stat. Mech.* P04010

[33] Kim H and Huse D A 2013 Ballistic spreading of entanglement in a diffusive nonintegrable system *Phys. Rev. Lett.* **111** 127205

[34] Bohrdt A, Mendl C B, Endres M and Knap M 2017 Scrambling and thermalization in a diffusive quantum many-body system *New J. Phys.* **19** 063001

[35] Kukuljan I, Grozdanov S and Prosen T 2017 Weak quantum chaos (arXiv:1701.09147)

[36] Chiara G D, Montangero S, Calabrese P and Fazio R 2006 Entanglement entropy dynamics of heisenberg chains *J. Stat. Mech.* P03001

[37] Han X and Hartnoll S 2019 Quantum scrambling and state dependence of the butterfly velocity *SciPost Phys.* **7** 045

[38] Lieb E H and Robinson D W 1972 The finite group velocity of quantum spin systems *Commun. Math. Phys.* **28** 251–7

[39] Ho W W and Abanin D A 2017 Entanglement dynamics in quantum many-body systems *Phys. Rev.* B **95** 094302

[40] Luitz D J and Bar Lev Y 2017 Information propagation in isolated quantum systems *Phys. Rev.* B **96** 020406

[41] Bertini B, Kos P and Prosen T 2019 Entanglement spreading in a minimal model of maximal many-body quantum chaos *Phys. Rev.* X **9** 021033

[42] Zhang Y-L and Khemani V 2020 Asymmetric butterfly velocities in two-local Hamiltonians *SciPost Phys.* **9**

[43] Xu S and Swingle B 2019 Locality, quantum fluctuations, and scrambling *Phys. Rev.* X **9** 031048

[44] Forster D 2018 *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Boca Raton, FL: CRC Press)

[45] Zhou T and Nahum A 2020 Entanglement membrane in chaotic many-body systems *Phys. Rev.* X **10** 031066

[46] Lucas A and Sachdev S 2015 Memory matrix theory of magnetotransport in strange metals *Phys. Rev.* B **91** 195122

[47] Davison R A, Delacrétaz L V, Goutéraux B and Hartnoll S A 2016 Hydrodynamic theory of quantum fluctuating superconductivity *Phys. Rev.* B **94** 054502

[48] Bentsen G, Gu Y and Lucas A 2019 Fast scrambling on sparse graphs *Proc. Natl Acad. Sci. USA* **116** 6689–94

[49] Liu F, Garrison J R, Deng D-L, Gong Z-X and Gorshkov A V 2018 Asymmetric particle transport and light-cone dynamics induced by anyonic statistics *Phys. Rev. Lett.* **121** 250404
[50] Stahl C, Khemani V and Huse D A 2018 Asymmetric butterfly velocities in Hamiltonian and circuit models (arXiv:1812.05589)
[51] McCulloch E 2021 Haar averaged moments of correlation functions and OTOCs in Floquet systems (arXiv:2110.15151)
[52] Gopalakrishnan S 2018 Operator growth and eigenstate entanglement in an interacting integrable Floquet system *Phys. Rev.* B **98** 060302