

# Quantum error correction from complexity in Brownian SYK

Vijay Balasubramanian,<sup>a,b</sup> Arjun Kar,<sup>c</sup> Cathy Li,<sup>a</sup> Onkar Parrikar<sup>d</sup> and Harshit Rajgadia<sup>d</sup>

<sup>a</sup>David Rittenhouse Laboratory, University of Pennsylvania,  
209 S. 33rd Street, Philadelphia PA 19104, U.S.A.

<sup>b</sup>Theoretische Natuurkunde, Vrije Universiteit Brussel (VUB), and  
International Solvay Institutes, Pleinlaan 2, B-1050 Brussels, Belgium

<sup>c</sup>Department of Physics and Astronomy, University of British Columbia,  
6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

<sup>d</sup>Department of Theoretical Physics, Tata Institute for Fundamental Research,  
Mumbai 400005, India

E-mail: [vijay@physics.upenn.edu](mailto:vijay@physics.upenn.edu), [arjunkar@phas.ubc.ca](mailto:arjunkar@phas.ubc.ca),  
[y1244@sas.upenn.edu](mailto:y1244@sas.upenn.edu), [parrikar@theory.tifr.res.in](mailto:parrikar@theory.tifr.res.in), [hrajgadia@gmail.com](mailto:hrajgadia@gmail.com)

**ABSTRACT:** We study the robustness of quantum error correction in a one-parameter ensemble of codes generated by the Brownian SYK model, where the parameter quantifies the encoding complexity. The robustness of error correction by a quantum code is upper bounded by the “mutual purity” of a certain entangled state between the code subspace and environment in the isometric extension of the error channel, where the mutual purity of a density matrix  $\rho_{AB}$  is the difference  $\mathcal{F}_\rho(A : B) \equiv \text{Tr} \rho_{AB}^2 - \text{Tr} \rho_A^2 \text{Tr} \rho_B^2$ . We show that when the encoding complexity is small, the mutual purity is  $O(1)$  for the erasure of a small number of qubits (i.e., the encoding is fragile). However, this quantity decays exponentially, becoming  $O(1/N)$  for  $O(\log N)$  encoding complexity. Further, at polynomial encoding complexity, the mutual purity saturates to a plateau of  $O(e^{-N})$ . We also find a hierarchy of complexity scales associated to a tower of subleading contributions to the mutual purity that quantitatively, but not qualitatively, adjust our error correction bound as encoding complexity increases. In the AdS/CFT context, our results suggest that any portion of the entanglement wedge of a general boundary subregion  $A$  with sufficiently high encoding complexity is robustly protected against low-rank errors acting on  $A$  with no prior access to the encoding map. From the bulk point of view, we expect such bulk degrees of freedom to be causally inaccessible from the region  $A$  despite being encoded in it.

**KEYWORDS:**  $1/N$  Expansion, AdS-CFT Correspondence, Field Theories in Lower Dimensions

ARXIV EPRINT: [2301.07108](https://arxiv.org/abs/2301.07108)

---

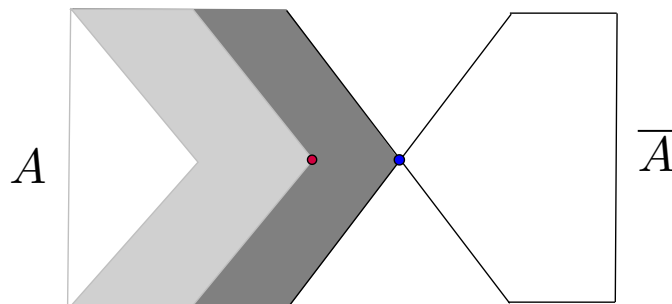
## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Setup</b>	<b>6</b>
2.1	Brief review of quantum error correction	6
2.2	An error correction bound	8
2.3	Error correction and maximum complexity encoding	10
2.4	Random circuit codes: Brownian SYK	12
2.5	Erasures errors	13
<b>3</b>	<b>Erasures in Brownian SYK codes</b>	<b>15</b>
3.1	Boundary conditions and large $N$ equations	15
3.2	Solutions: qualitative discussion	20
3.3	Disconnected solution	21
3.4	Connected solution	27
3.5	Other tunneling solutions	29
3.6	Summary	31
<b>4</b>	<b>Discussion</b>	<b>32</b>
4.1	Relation to previous work	32
4.2	Pseudorandom codes	34
4.3	Complexity and the geometry of the entanglement wedge	34
<b>A</b>	<b>Hamiltonian formalism in Brownian SYK</b>	<b>35</b>
A.1	Disconnected and connected solutions	36
A.2	One-loop determinant around disconnected solution	38
<b>B</b>	<b>Proof of the error correction bound</b>	<b>39</b>

---

## 1 Introduction

The bulk-to-boundary map in AdS/CFT has a rich structure. For any boundary subregion  $A$ , the associated Ryu-Takayanagi surface [1] singles out a certain subregion  $a$  of the bulk spacetime called the entanglement wedge of  $A$  [2]. The AdS/CFT map then satisfies subregion duality: bulk semi-classical degrees of freedom in  $a$  are encoded within  $A$  and are protected against erasures in  $\bar{A}$ . Furthermore, bulk operators within the entanglement wedge  $a$  can be reconstructed as boundary operators localized within the boundary region  $A$ , a property sometimes known as entanglement wedge reconstruction [3]. Using the language of [3–10], these properties hold because the Ryu-Takayanagi formula and its quantum generalizations imply that the bulk-to-boundary map in AdS/CFT is a quantum error



**Figure 1.** The entanglement wedge of a boundary subregion  $A$  can have a rich substructure. The outer white region on the left is the causal wedge of  $A$ , while the slightly darker grey region bounded by a non-minimal QES (shown in red) is the simple wedge. Beyond this lies the python’s lunch (dark grey).

correcting code with complementary recovery, where the entanglement wedge  $a$  of  $A$  is protected against the erasure of  $\bar{A}$ , while  $\bar{a}$  is protected against the erasure of  $A$ .

Recent progress points towards a sharper characterization of the structure of entanglement wedges that appear in holography and its generalizations [11–16]. Given a general boundary subregion, the corresponding entanglement wedge has a layered structure, i.e., it can be broken up into three regions: the causal wedge, the simple wedge [14, 15] and the python’s lunch [17]. These three regions of the entanglement wedge are defined as follows (see figure 1): the causal wedge is the region in the bulk which is causally accessible from the boundary, i.e., a boundary observer in the domain of dependence of  $A$  can send signals to and receive signals from all points in the causal wedge of  $A$ . The simple wedge is defined to be the bulk domain of dependence of the homology region between  $A$  and the outermost quantum extremal surface (QES) which need not be minimal among all the QESs associated with  $A$ . The simple wedge is generically larger than the causal wedge, and so there are points in the simple wedge which are out of causal contact with the domain of dependence of  $A$ . However, it has been argued that the simple wedge can always be brought in causal contact with the boundary by performing backwards and forwards Lorentzian time evolution with sources turned on to de-focus the causal horizons [15]. Finally, the python’s lunch region is defined as the portion of the entanglement wedge which lies between the outermost QES and the minimal QES. This region is causally inaccessible from the boundary subregion  $A$ , and furthermore since it lies behind an extremal surface it cannot be brought into causal contact with the domain of dependence of  $A$  (in contrast with the simple wedge); this follows from the fact that extremal surfaces must always lie behind causal horizons. This seems to lead to a puzzle — on the one hand, bulk operators in the python’s lunch are encoded in  $A$  and in particular one should be able to create a semi-classical bulk excitation in the python’s lunch via an operator acting on the domain of dependence of  $A$ . On the other hand, semi-classical gravity seems to forbid this!

The evaporating black hole provides a context where this apparent contradiction is particularly sharp. Beyond the Page time, a portion of the black hole interior — the island — lies in the entanglement wedge of the radiation (see [11–13, 18–26] for a partial list of articles discussing this phenomenon for black hole and cosmological horizons). But

this portion lies behind a non-minimal QES, namely the empty surface,<sup>1</sup> and therefore constitutes a python’s lunch. While one should be able to manipulate operators in the island by quantum operations on the radiation, such operations seem to blatantly violate semi-classical bulk causality. A potential way out is suggested by bounds coming from computational complexity [14, 15, 17, 27–32] — we expect that the encoding map for excitations in the python’s lunch region is extremely complex, perhaps exponentially so in the number of qubits, and so any computationally bounded observer (with access only to sub-exponential operations on the radiation) will be unable to manipulate the degrees of freedom in the island. This is how we expect that semi-classical bulk causality will be approximately respected. On the other hand, certain finely tuned, exponentially complex operations on the radiation should be able to manipulate degrees of freedom in the island, but the gravitational mechanism for this involves Euclidean wormholes.

We can get an intuition for why complexity can protect information in this way from an analogy to older results concerning the complexity of black hole microstates and the difficulty of using simple probes to extract information about them [33–35]. Consider, for example, a Schwarzschild black hole of mass  $M$  in  $\text{AdS}_5$  with a length scale  $\ell$ . A microstate of this black hole is described in the dual  $\text{SU}(N)$  Yang-Mills theory with 16 supersymmetries, by an operator  $\mathcal{O}$  of dimension  $\Delta = M\ell \sim N^2$ .  $\mathcal{O}$  is roughly a polynomial of length  $N^2$  built from the elementary fields of the Yang-Mills theory (a gauge field  $A_\mu$ , fermions  $\psi_a$  and three complex adjoint scalars  $X, Y, Z$ ) and their derivatives, with indices contracted to make the polynomial gauge and Lorentz invariant. Almost all such long polynomials are random sequences of fields and derivatives up to constraints of gauge and Lorentz invariance. A light probe of the state like the graviton corresponds to an operator of dimension  $O(1)$ , like  $P = \text{Tr}(XX)$ . The question is whether a measurement, modeled as a correlation function in the state created by  $\mathcal{O}$ ,  $\langle 0 | \mathcal{O}^\dagger P^\dagger P \mathcal{O} | 0 \rangle$ , can reveal information about the identity of  $\mathcal{O}$ . The authors of [33, 34] argue that the answer is “no” because of the universal statistics of random polynomials, which mean that almost all  $\mathcal{O}$  will lead to a similar sum of terms from contractions between the fields in the probe and the fields in  $\mathcal{O}$  in evaluating the correlator. As such, simple (i.e., low-dimension) probes cannot reveal the microstate, but an observer with prior knowledge of the state could construct a fine-tuned, complex probe to check that knowledge, by choosing these probes to match long sequences of the fields composing  $\mathcal{O}$ . One expects the situation in the python’s lunch inside an evaporating black hole to be somewhat analogous: a highly complex encoding map prevents simple operations in the radiation from affecting the black hole interior, but if the encoding map is accessible then finely tuned, complex operations affecting the interior may be performed more easily.

Kim, Preskill and Tang (KPT) have sharpened these expectations [30]. They suggested that the encoding of the black hole interior degrees of freedom in the radiation, thought of

---

<sup>1</sup>In more detail, in the toy models where these calculations are possible, the radiation is extracted into an auxiliary reservoir that is not geometrically connected to the island. Even in the absence of a geometric connection, there is still an obvious candidate extremal surface which one can consider as bounding the region dual to the radiation, namely the “empty surface”. By this, we mean that the entire black hole spacetime is taken to lie “outside” the would-be entanglement wedge. After the Page time, this surface is no longer the quantum minimal surface which computes the radiation entropy, and the true minimum QES lies in the spacetime near the black hole horizon. Nevertheless, the island is “behind the empty extremal surface” from the point of view of the radiation.

as a quantum error correcting code, has robust error correction properties against low-rank, computationally bounded errors on the radiation, or more precisely, errors which effectively see the radiation density matrix as thermal. Note that this is not the standard error correction one encounters in the context of subregion duality; in the KPT formulation, bulk degrees of freedom in the island — while being encoded in the radiation — are nevertheless approximately (up to corrections exponentially small in the black hole entropy) protected against certain errors acting on the radiation itself. KPT then argued that this approximate error correction implies the existence of “ghost logical” operators which act on the radiation to mimic bulk operators in the island and at the same time commute with computationally bounded operators on the radiation — thus realizing the approximate causality of the black hole spacetime. The language of quantum error correction thus enables one to formulate and address the question of bulk causality in a universal manner.

Recently, the novel error correction in evaporating black holes proposed by KPT was tested in a toy model for an evaporating black hole in Jackiw-Teitelboim gravity [31], and it was argued that the bulk degrees of freedom in the island are protected against a large class of low-rank error operations on the radiation which do not have access to the details of the microscopic black hole state. The low-rank criterion can be formalized as a bound on the coherent information of the error in terms of the black hole entropy. In [31], it was also conjectured that this same robust error correction should also work in the python’s lunch portions of more general entanglement wedges. The underlying reason is the high complexity of encoding in the python’s lunch. The rough picture is the same as KPT — as the encoding map becomes sufficiently complex, any generic, low-rank error operation involving “simple” operations sees only a coarse-grained<sup>2</sup> density matrix on the boundary subregion, with no sign of the encoded subspace. In other words, the encoded subspace gets lost within the exponentially large Hilbert space of the boundary subregion. This “complexity-protected error correction” makes it possible for the semi-classical degrees of freedom in the python’s lunch to be encoded in a boundary subregion and yet be causally inaccessible from it using simple probes.

The purpose of this paper is to demonstrate the above phenomenon in a toy model where the behavior of the encoding complexity is known more or less by construction. Such control is difficult to achieve directly in real holography because proving results about the complexity of the bulk-to-boundary map (without resorting to toy models like tensor networks) in different regions of the entanglement wedge is generically a very difficult task. Rather than a single code, here we consider an ensemble of quantum error correcting codes of the type relevant for entanglement wedge reconstruction in AdS/CFT. Since we want control over the complexity of encoding, our ensemble of codes is generated by picking the encoding map from an ensemble of unitaries with fixed circuit complexity.

We accomplish this by taking these unitaries to be time evolution operators  $U(T) = \mathcal{T} \exp[-i \int_0^T dt H(t)]$  in the Brownian Sachdev-Ye-Kitaev (SYK) model [37–41], a

---

<sup>2</sup>The relevant notion of coarse-graining was defined in [15, 36]: one finds the maximum-entropy state consistent with correlation functions of all simple operators, including Lorentizan time-folds with simple sources turned on. Here simple operators and sources are defined as those whose effects propagate causally in the bulk.

quantum mechanical theory of  $N$  Majorana fermions. The SYK model here is merely a trick to generate an ensemble of unitaries with fixed complexity, parametrized by the number  $T$ . When  $T$  is small, the corresponding set of unitary operators is clustered around the identity operator, but as  $T \rightarrow \infty$  this set grows [42–46] to cover (modulo global symmetries) the entire unitary group. When it covers the entire unitary group, the typical complexity of an operator in the set is exponentially large [47]. So, computing the average error correction properties of such sets gives us some insight into the behavior of a family of codes with increasing complexity.

In this paper we will consider typical, low-rank errors with no prior access to the encoding map, and acting on a small fixed fraction of the physical Hilbert space. As a particular instance of such errors, we will consider the erasure of a small fraction of the physical Hilbert space. In quantum information theory, it is standard to model an error in terms of coupling to an external environment and tracing out the environment. The error correction properties of the code can then be studied in terms of the amount of correlation generated by the error between the code subspace and the environment. Error correction works with high accuracy when these correlations are suppressed by a large parameter e.g. the dimension of the physical Hilbert space.

We will study a particular measure of correlation, namely the “mutual purity” between the code subspace and the environment. We define the mutual purity  $\mathcal{F}_\rho(A : B)$  of a density matrix  $\rho$  between Hilbert subsystems  $\mathcal{H}_A$  and  $\mathcal{H}_B$  as  $\text{Tr } \rho_{AB}^2 - \text{Tr } \rho_A^2 \text{Tr } \rho_B^2$ . The fact that this quantity is a good measure of error correction is rigorously justified in appendix B. Our main result is that for the Brownian SYK ensemble of quantum error correcting codes, there are three complexity regimes of interest.

- (i) For  $T$  smaller than a scrambling time  $T \sim \log N$  (i.e., low encoding complexity) the erasure of a small fraction of the physical qubits generate an  $O(1)$  amount of correlation that decays exponentially with  $T$  between the code subspace and the environment, and thus there is no robust quantum error correction.
- (ii) For  $T > \log N$ , the mutual purity becomes  $O(1/N)$  but keeps decaying further as the complexity  $T$  increases.
- (iii) When  $T \sim N$ , the mutual purity becomes exponentially small in  $N$ ; at this point, there is an  $O(e^{-N})$  residual correlation generated by the error which is unavoidable.

The third and final regime corresponds to an exchange of dominance between a leading saddle point and a subleading saddle point<sup>3</sup> in the Brownian SYK calculation, analogously to the exchange of dominance between a disconnected geometry and the Euclidean wormhole in gravity. This quantitative hierarchy of complexity-protected error correction, ranging from a fragile encoding at  $O(1)$  complexity, through a logarithmic complexity regime of reasonable protection, and finally an emergent robust error correction at large encoding complexity,

---

<sup>3</sup>Furthermore, there are also strictly subleading saddles controlled by a one-dimensional lattice of critical time points with the scrambling time as the lattice vector. The amount of correlation generated with the environment only changes as  $T$  passes a lattice point.

is the central result of this paper. We regard this as a step towards understanding the structure of general entanglement wedges (figure 1) from the boundary perspective in terms of quantum error correction.

Three sections follow. In section 2 we review the necessary ideas from quantum error correction. We discuss the class of errors of interest, and show that the mutual purity which is relevant for recovery from these errors can be expressed in terms of the standard purity<sup>4</sup> of a certain density matrix constructed using the encoding map. We also briefly review the Brownian SYK model. In section 3, we compute this purity in the large  $N$  limit using the Brownian SYK time evolution operator to model the encoding map. We conclude with a discussion in section 4. In appendix A, we give a Hamiltonian treatment of Brownian SYK to complement the path integral discussion in the main text and in appendix B we prove that the mutual purity provides a bound on the error correction properties of an encoding map.

## 2 Setup

### 2.1 Brief review of quantum error correction

The mathematical framework for quantum error correction involves an isometric embedding of a small “code subspace”  $\mathcal{H}_{\text{code}}$  into a larger Hilbert space  $\mathcal{H}_{\text{phys}}$ :

$$V : \mathcal{H}_{\text{code}} \rightarrow \mathcal{H}_{\text{phys}},$$

where  $V^\dagger V = \mathbb{1}$ . It is standard to model the error and recovery operations as completely positive trace-preserving linear maps, or “quantum channels”. Any such map  $\mathcal{E}$  has a representation in terms of its Kraus operators  $\{E_m\}$  [48, 49]:

$$\mathcal{E}(\rho) = \sum_m E_m \rho E_m^\dagger, \quad \sum_m E_m^\dagger E_m = \mathbb{1}. \quad (2.1)$$

The minimum number of Kraus operators needed to implement a particular channel is called the rank of the channel. These quantum channels act on physical density matrices, and the goal of error correction is to determine for a given error channel  $\mathcal{E}$  whether or not there exists a recovery channel  $\mathcal{R}$  which restores the state  $\rho_{\text{code}}$ :

$$\mathcal{R}(\mathcal{E}(V \rho_{\text{code}} V^\dagger)) = \rho_{\text{code}}. \quad (2.2)$$

On the right hand side, we have in mind that the recovery channel has eliminated redundant portions of  $\mathcal{H}_{\text{phys}}$ , leaving behind precisely the matrix  $\rho_{\text{code}}$  on the remaining subspace of  $\mathcal{H}_{\text{phys}}$ .

A second, convenient description of a quantum channel is given by its isometric extension, also known as its Stinespring dilation: we describe it as coupling the physical system via a unitary operator  $U_{\mathcal{E}}$  to an auxiliary environment with Hilbert space  $\mathcal{H}_{\text{env}}$  spanned by basis elements  $\{|e_m\rangle_{\text{env}}\}$ , initially in some fiducial state  $|e_0\rangle$ . The action of the channel  $\mathcal{E}$  on  $\rho$  is

---

<sup>4</sup>For a density matrix  $\rho$ , the purity is defined as  $\text{Tr} \rho^2$ .



then recovered by tracing out the environment:  $\mathcal{E}(\rho) = \text{Tr}_{\text{env}} \left[ U_{\mathcal{E}} (\rho \otimes |e_0\rangle\langle e_0|_{\text{env}}) U_{\mathcal{E}}^\dagger \right]$ . This implies that  $E_m = \langle e_m | U_{\mathcal{E}} | e_0 \rangle$ , or, equivalently,

$$U_{\mathcal{E}} |\psi\rangle \otimes |e_0\rangle = \sum_m E_m |\psi\rangle \otimes |e_m\rangle, \quad (2.3)$$

where  $|\psi\rangle$  is any state in the physical Hilbert space.

A standard fact in quantum error correction, sometimes called the *decoupling principle*, is that there always exists an approximate recovery channel where the error in recovery is bounded in terms of the amount of correlation the error channel generates between the code subspace and the environment. The convenient way to evaluate this correlation is to use the following procedure: (a) introduce a reference system  $\mathcal{H}_{\text{ref}}$  which is isomorphic to and maximally entangled with code Hilbert space, (b) act with the error quantum channel, (c) trace out the physical Hilbert space, and (d) evaluate the correlation between the two remaining auxiliary spaces (the environment used to represent the channel and the reference space). Thus, taking  $|i\rangle_{\text{ref}}$  and  $|i\rangle_{\text{code}}$  to be orthonormal bases for the reference space and the code subspace respectively, we construct the state

$$|\Psi'\rangle = \frac{1}{\sqrt{d_{\text{code}}}} \sum_{i=1}^{d_{\text{code}}} \sum_{m=1}^{d_{\text{env}}} |i\rangle_{\text{ref}} \otimes E_m V |i\rangle_{\text{code}} \otimes |e_m\rangle_{\text{env}}, \quad (2.4)$$

where we have defined  $d_X$  to be the dimension of a Hilbert space  $\mathcal{H}_X$ . Here the code states are embedded by  $V$  into the physical Hilbert space and maximally entangled with the reference, while the error channel acts via  $E_m$  on the physical states and thus entangles them with the environment. Then we can say that for any error channel  $\mathcal{E} = \{E_m\}$  there exists a recovery channel  $\mathcal{R}$  for which the Schatten 1-norm distance between the resulting state and the original encoded state is bounded as [50, 51]:

$$\|\mathcal{R}(\mathcal{E}(V \rho_{\text{code}} V^\dagger)) - \rho_{\text{code}}\|_1 \leq (I_{\Psi'}(\text{ref} : \text{env}))^{1/4}. \quad (2.5)$$

Here  $I_{\Psi'}(\text{ref} : \text{env})$  is the mutual information between the environment and the reference space after tracing out the physical Hilbert space. This means that the error  $\mathcal{E}$  is exactly correctable in the code  $V$  if the reference and environment do not share any correlation, hence the term “decoupling”.

In this paper, we will be interested in quantum codes with complementary recovery [52], which are the types of codes relevant for entanglement wedge reconstruction in AdS/CFT. For simplicity, consider a code subspace where we have some semi-classical bulk degrees of freedom in the entanglement wedge of a boundary subregion  $A$ , but no excitations in the entanglement wedge of the complement region  $\bar{A}$ . Let  $|i\rangle_{\text{code}}$  denote basis states for these bulk degrees of freedom. It was shown by Harlow that the Ryu-Takayanagi formula together with quantum corrections implies the following structure for the encoding map in this situation:

$$V : \mathcal{H}_{\text{code}} \rightarrow \mathcal{H}_{\text{phys}}, \quad (2.6)$$

$$V |i\rangle_{\text{code}} = (U_A \otimes \mathbb{1}_{\bar{A}}) \left( |i\rangle_{A_1} \otimes |\chi\rangle_{A_2 \bar{A}} \right), \quad (2.7)$$



where the physical Hilbert space (i.e., the Hilbert space of the dual CFT) is factorized as

$$\mathcal{H}_{\text{phys}} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}, \quad \mathcal{H}_A = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \oplus \mathcal{H}_{A_3}, \quad (2.8)$$

and  $|\chi\rangle$  is some fixed pure state in the Hilbert space  $\mathcal{H}_{A_2, \bar{A}}$ . The argument for this involves the decoupling principle applied to the erasure of  $\bar{A}$ . Let us briefly recall how this works (see [52] for details): we introduce an auxiliary system  $\mathcal{H}_{\text{aux}}$  isomorphic to the code subspace, and construct the state

$$|\Psi\rangle = \frac{1}{\sqrt{d_{\text{code}}}} \sum_{i=1}^{d_{\text{code}}} |i\rangle_{\text{aux}} \otimes |\psi_i\rangle_{A, \bar{A}}, \quad |\psi_i\rangle_{A, \bar{A}} = V|i\rangle_{\text{code}}. \quad (2.9)$$

Since the bulk degrees of freedom in the code subspace are contained in the entanglement wedge of  $A$ , one can show using the RT [1] plus FLM [53] formula that the mutual information  $I(\text{aux} : \bar{A})$  vanishes, which implies that  $\rho_{\text{aux}, \bar{A}} = \rho_{\text{aux}} \otimes \rho_{\bar{A}}$ . Therefore, viewed as a bipartite state on  $A$  and  $\text{aux} \cup \bar{A}$  the Schmidt vectors of  $\Psi$  should take a factorized form on  $\text{aux} \cup \bar{A}$ . The *canonical purification* [36, 54] of  $\rho_{\text{aux}, \bar{A}}$  will therefore also have factorized states on  $A$ . This is why the state inside the parentheses in equation (2.7) takes the factorized form between  $A_1$  and  $A_2$ ; here  $A_1$  is the canonical purifier of  $\text{aux}$  and  $A_2$  is the canonical purifier of  $\bar{A}$ . Finally, any two purifications of the same density matrix  $\rho_{\text{aux}, \bar{A}}$  should be related by a unitary on  $A$ ; this is precisely the unitary  $U_A \otimes \mathbb{1}_{\bar{A}}$  appearing in equation (2.7). It is easy to check that this code subspace is protected against the erasure of  $\bar{A}$ . We will refer to the operator  $U_A$  as the *encoding unitary*.<sup>5</sup>

There is an important caveat: the bulk-to-boundary map  $V$  need not be an exact isometry, and is often an approximate one with corrections of  $O(e^{-1/G_N})$ .<sup>6</sup> Relatedly, the quantum generalization of the Ryu-Takayanagi formula, namely the QES formula, is correct to all orders in the  $G_N$  perturbation theory for appropriate states,<sup>7</sup> but in general there are corrections of  $O(e^{-1/G_N})$ . Therefore, the bulk-to-boundary map in AdS/CFT is only approximately of the form (2.7), and has additional exponentially small corrections. As a first pass, we will focus on codes of the type (2.7) in this work. It would be interesting to incorporate the corrections mentioned above in our analysis, but we will not attempt this here.

## 2.2 An error correction bound

Putting together the considerations from above, we first introduce a reference system isomorphic to the code subspace and consider the maximally entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{d_{\text{code}}}} \sum_i |i\rangle_{\text{ref}} \otimes U_A \left( |i\rangle_{A_1} \otimes |\chi\rangle_{A_2 \bar{A}} \right) \otimes |e_0\rangle_{\text{env}}, \quad (2.10)$$

<sup>5</sup>The additional Hilbert space component  $\mathcal{H}_{A_3}$  in (2.8) is, for our purposes, a bookkeeping device for situations where the physical Hilbert space dimension is not a product of integers, also implying that a part of it does not participate in the code; so we will simply drop  $\mathcal{H}_{A_3}$  as it is not pertinent to our considerations. Henceforth, we will focus our attention on codes which have the above structure, but without  $\mathcal{H}_{A_3}$ .

<sup>6</sup>There are also more extreme situations in which the map is far from an isometry [29, 31, 32].

<sup>7</sup>See [55] for situations where there are leading order corrections.

where we included the code subspace structure in (2.7) and an auxiliary environment in some fiducial initial state  $|e_0\rangle$ . The error now acts in the form of a joint unitary operator on  $A \cup \text{env}$  (we assume the error does not act on the  $\bar{A}$  system):

$$|\Psi'\rangle = \frac{1}{\sqrt{d_{\text{code}}}} \sum_i |i\rangle_{\text{ref}} \otimes U_{\mathcal{E}} \left[ U_A \left( |i\rangle_{A_1} \otimes |\chi\rangle_{A_2 \bar{A}} \right) \otimes |e_0\rangle_{\text{env}} \right], \quad (2.11)$$

where we applied the error channel as in (2.4) to the state (2.10) in terms of a unitary operator (2.3) entangling the physical system with the environment. Next, we obtain the reduced density on the reference and environment subsystems:

$$\rho'_{\text{ref,env}} = \frac{1}{d_{\text{code}}} \sum_{i,j} |i\rangle\langle j|_{\text{ref}} \otimes \text{Tr}_A \left\{ U_{\mathcal{E}} \left[ U_A \left( |i\rangle\langle j|_{A_1} \otimes \rho_{A_2}^{\chi} \right) U_A^{\dagger} \otimes |e_0\rangle\langle e_0|_{\text{env}} \right] U_{\mathcal{E}}^{\dagger} \right\}, \quad (2.12)$$

where we have performed the trace over  $\bar{A}$  and replaced  $\chi$  with its reduced density matrix  $\rho_{A_2}^{\chi}$ . Finally, following (2.5) we can bound the error in recovery of the original state after action of the error channel in terms of the fourth root of the mutual information between the reference and the environment:

$$I_{\Psi'}(\text{ref} : \text{env}) = S(\rho'_{\text{ref}}) + S(\rho'_{\text{env}}) - S(\rho'_{\text{ref,env}}), \quad (2.13)$$

where the von Neumann entropies on the right hand side are computed from  $\rho'_{\text{ref,env}}$  and the reduced density matrices on the reference and the environment  $\rho'_{\text{ref}} = \text{Tr}_{\text{env}}(\rho'_{\text{ref,env}})$  and  $\rho'_{\text{env}} = \text{Tr}_{\text{ref}}(\rho'_{\text{ref,env}})$ .

The mutual information in (2.13) is difficult to compute directly. A standard approach is to use the replica trick to obtain the von Neumann entropies on the right hand side as analytic continuations of the Rényi entropies which are easier to compute via the relation

$$S(\rho) = -\text{Tr}(\rho \log \rho) = \lim_{n \rightarrow 1} \frac{1}{1-n} S^{(n)}(\rho), \quad (2.14)$$

where  $S^{(n)}(\rho) = \log \text{Tr}(\rho^n)$  is the  $n^{\text{th}}$  Rényi entropy. We will take a different approach. In appendix B we study a particular, well-motivated recovery channel and show that the trace distance between the recovered state under this recovery channel and the actual state satisfies

$$D(\mathcal{R} \circ \mathcal{E}(V \rho_{\text{code}} V^{\dagger}), \rho_{\text{code}}) \leq c \left( \text{Tr} \left( \rho_{\text{ref,env}}'^2 - \rho_{\text{ref}}'^2 \otimes \rho_{\text{env}}'^2 \right) \right)^{1/4}, \quad c = d_{\text{code}}^{5/2} d_{\text{env}}^{1/2}, \quad (2.15)$$

where  $D(\rho, \sigma) = \frac{1}{2} \text{Tr}(|\rho - \sigma|)$  with  $|X| = \sqrt{X^{\dagger} X}$  is the trace distance between density matrices. As above,  $d_{\text{code}}$  and  $d_{\text{env}}$  are dimensions of the code/reference subspace and the environment in the isometric extension of the error channel, respectively. In this work, (2.15) will replace the standard decoupling principle (2.5) due to the ease of evaluating the right hand side. In particular, the expression (2.15) bounds the error in recovery directly in terms of the quantity

$$\mathcal{F}_{\Psi'}(\text{ref} : \text{env}) \equiv \text{Tr} \left( \rho_{\text{ref,env}}'^2 - \rho_{\text{ref}}'^2 \otimes \rho_{\text{env}}'^2 \right), \quad (2.16)$$

which we call the *mutual purity*. If  $\mathcal{F}$  vanishes, so does the right hand side of the bound (2.15), so that perfect recovery is possible and the error can be corrected. In view of this bound, below we will compute  $\mathcal{F}$  to quantify the robustness against errors for encoding maps of increasing complexity.

### 2.3 Error correction and maximum complexity encoding

To get a more quantitative understanding of what happens when the encoding unitary becomes complex, as a first pass we can compute the Haar ensemble average with respect to  $U_A$  of  $\mathcal{F}_{\Psi'}$ . This is because we expect that the typical unitary in the Haar ensemble will be exponentially complex, and that as long as the dimension of  $\mathcal{H}_A$  is large, deviations away from the ensemble average will be exponentially suppressed in the number of qubits. For a Haar random unitary  $U$  acting on a Hilbert space  $\mathcal{H}_X$ , a standard formula for Haar integration says:

$$\begin{aligned} \langle U_{m_1 p_1} U_{q_1 n_1}^\dagger U_{m_2 p_2} U_{q_2 n_2}^\dagger \rangle_{\text{Haar}} = \\ \frac{1}{d_X^2} (\delta_{m_1, n_1} \delta_{m_2, n_2} \delta_{p_1, q_1} \delta_{p_2, q_2} + \delta_{m_1, n_2} \delta_{m_2, n_1} \delta_{p_1, q_2} \delta_{p_2, q_1}) + O\left(\frac{1}{d_X^3}\right). \end{aligned} \quad (2.17)$$

This expression has a gravitational analogue: in the Euclidean path integral computation of the radiation purity in the PSSY toy model for an evaporating black hole in JT gravity [10], the two terms displayed above are respectively analogous to the “disconnected” and “wormhole” gravitational saddles. Using the above integral we can now evaluate the Haar average of  $\text{Tr}_{\text{ref,env}} [(\rho'_{\text{ref,env}})^2]$  as follows.

$$\begin{aligned} \langle \text{Tr} \rho_{\text{ref,env}}^2 \rangle_{\text{Haar}} &= \frac{1}{d_{\text{code}}^2} \sum_{i,j,m,n} \langle \text{Tr}_A \left\{ E_m U_A \left( |i\rangle \langle j|_{A_1} \otimes \rho_{A_2}^X \right) U_A^\dagger E_n^\dagger \right\} \\ &\quad \text{Tr}_A \left\{ E_n U_A \left( |j\rangle \langle i|_{A_1} \otimes \rho_{A_2}^X \right) U_A^\dagger E_m^\dagger \right\} \rangle_{\text{Haar}} \\ &\approx \frac{1}{d_{\text{code}}^2 d_A^2} \sum_{i,j,m,n} \text{Tr}(E_m^\dagger E_n) \delta_{i,j} \text{Tr}(\rho_{A_2}^X) \text{Tr}(E_n^\dagger E_m) \delta_{j,i} \text{Tr}(\rho_{A_2}^X) \\ &\quad + \frac{1}{d_{\text{code}}^2 d_A^2} \sum_{i,j,m,n} \text{Tr}(E_m E_m^\dagger E_n E_n^\dagger) \text{Tr}(\rho_{A_2}^{X^2}) \delta_{i,i} \delta_{j,j} \\ &= \frac{1}{d_{\text{code}}^2 d_A^2} \sum_{m,n} \left( d_{\text{code}} \text{Tr}(E_m^\dagger E_n) \text{Tr}(E_n^\dagger E_m) + d_{\text{code}}^2 \text{Tr}(E_m E_m^\dagger E_n E_n^\dagger) \text{Tr}(\rho_{A_2}^{X^2}) \right) \\ &= \frac{1}{d_{\text{code}}^2 d_A^2} \left( d_{\text{code}} d_A^2 \text{Tr}(\sigma_{\text{env}}^2) + d_{\text{code}}^2 d_A^2 \text{Tr}(\sigma_A^2) \text{Tr}(\rho_{A_2}^{X^2}) \right) \\ &= \frac{\text{Tr}(\sigma_{\text{env}}^2)}{d_{\text{code}}} \left( 1 + d_{\text{code}} \frac{\text{Tr}(\sigma_A^2) \text{Tr}(\rho_{A_2}^{X^2})}{\text{Tr}(\sigma_{\text{env}}^2)} \right), \end{aligned} \quad (2.18)$$

where we have defined the density matrix  $\sigma$  on  $\mathcal{H}_A \otimes \mathcal{H}_{\text{env}}$  as

$$\sigma \equiv U_{\mathcal{E}} \left( \frac{\mathbb{1}_A}{d_A} \otimes |e_0\rangle \langle e_0|_{\text{env}} \right) U_{\mathcal{E}}^\dagger = \sum_{m,n} E_m \frac{\mathbb{1}_A}{d_A} E_n^\dagger \otimes |e_m\rangle \langle e_n|, \quad (2.19)$$

and the associated reduced density matrices

$$\sigma_A = \text{Tr}_{\text{env}} \sigma = \sum_m E_m \frac{\mathbb{1}_A}{d_A} E_m^\dagger, \quad \sigma_{\text{env}} = \text{Tr}_A \sigma = \sum_{m,n} \frac{\text{Tr}_A (E_m E_n^\dagger)}{d_A} |e_m\rangle \langle e_n|. \quad (2.20)$$

In the first step, we represented the action of  $U_{\mathcal{E}}$  in terms of the Kraus operators  $E_m = \langle e_m | U_{\mathcal{E}} | e_0 \rangle$  and traced out the reference and the environment degrees of freedom. The second step follows from equation (2.17). In the fourth step, we used equations (2.19) and (2.20). Following similar steps, we can compute the Haar average of  $\text{Tr} \rho_{\text{env}}'^2$ :

$$\langle \text{Tr} \rho_{\text{env}}'^2 \rangle_{\text{Haar}} \approx \text{Tr}(\sigma_{\text{env}}^2) \left( 1 + \frac{\text{Tr}(\sigma_A^2) \text{Tr}(\rho_{A_2}^{\chi^2})}{d_{\text{code}} \text{Tr}(\sigma_{\text{env}}^2)} \right). \quad (2.21)$$

Since  $\rho_{\text{ref}}'$  is maximally mixed, we have

$$\text{Tr} \rho_{\text{ref}}'^2 = \frac{1}{d_{\text{code}}}, \quad (2.22)$$

for any  $U_A$ , which means this expression factors out of any Haar average since it is independent of  $U_A$ . Combining the above results for  $\langle \text{Tr} \rho_{\text{ref,env}}'^2 \rangle_{\text{Haar}}$ ,  $\langle \text{Tr} \rho_{\text{env}}'^2 \rangle_{\text{Haar}}$ , and  $\text{Tr} \rho_{\text{ref}}'^2$ , the final result for the Haar averaged mutual purity is given by

$$\langle \mathcal{F}_{\Psi'}(\text{ref} : \text{env}) \rangle_{\text{Haar}} = e^{-S^{(2)}(\sigma_A) - S^{(2)}(\chi_{\bar{A}})} \left( 1 - \frac{1}{d_{\text{code}}^2} \right) + \dots, \quad (2.23)$$

where the  $\dots$  indicate exponentially small contributions that we have dropped along the way,  $S^{(2)}(\sigma_A)$  is the second Rényi entropy of the  $A$  subsystem in the mixed state  $\sigma$ , and  $S^{(2)}(\chi_{\bar{A}})$  is the second Rényi entropy of the  $\bar{A}$  subsystem in the state  $|\chi\rangle_{A_2 \bar{A}}$ . The salient feature of (2.23) is the leading exponential suppression, as we will now describe. The quantity  $1 - d_{\text{code}}^{-2}$  is simply an  $O(1)$  prefactor for a nontrivial code subspace.

Two features of (2.23) are worth highlighting. Firstly, note from the final formula that in the typical code drawn from the Haar ensemble, the error channel perceives the state on  $A$  as maximally mixed, and gains no access to the microscopic structure of the state. Consequently, as long as the error channel is low-rank, we see that  $\langle \mathcal{F}_{\Psi'}(\text{ref} : \text{env}) \rangle_{\text{Haar}}$  is exponentially suppressed by  $e^{-S^{(2)}(\sigma_A)}$ . This is a direct consequence of complexity — a general, complex encoding unitary scrambles the code subspace to a point where generic error channels do not gain any access to it. (A similar coarse-graining picture for apparent horizons and quantum extremal surfaces was advocated in [15, 36, 56].) Furthermore, there is an additional suppression factor of  $e^{-S^{(2)}(\chi_{\bar{A}})}$  in equation (2.23) coming from the shared entanglement with  $\bar{A}$ . The combination of these two effects coming from complexity and entanglement thus makes the code robust against generic, low-rank errors.

In the above analysis, we have assumed that the error channel does not have prior access to the encoding unitary  $U_A$ . This is crucial, because with prior access to the details of the encoding unitary, it is possible to construct low-rank error channels which corrupt the code subspace. For example, consider the error channel:

$$\mathcal{E}(\rho_A) = \mathcal{E}_{\text{partial SWAP}}(U_A^\dagger \rho_A U_A), \quad (2.24)$$

where the unitaries  $U_A^\dagger(\dots)U_A$  first undo the encoding, and the partial swap then swaps out the state on the first  $t$  qubits with the environment. Since the reference system in  $\Psi'$  is maximally entangled with the qubits in  $A_1$ , even if the partial SWAP acts on one of the

qubits in  $A_1$ , then it will generate an  $O(1)$  amount of mutual information between  $\mathcal{H}_{\text{ref}}$  and  $\mathcal{H}_{\text{env}}$ , and thus error correction fails. There is an analogue of this in the JT gravity model [31] — there, one assumes that the error channel does not “generate” additional asymptotic boundaries which can connect up with the bulk geometry and modify the mutual information. Of course, note that this channel is fine-tuned, in that it uses the specific unitary  $U_A$  which goes into the encoding. Nevertheless, if  $U_A$  is computationally simple, then the above error channel is also simple. On the other hand, when the encoding unitary  $U_A$  is exponentially complex, the error channel described above must be equally complex in order to first undo the encoding. Thus, if the python’s lunch has an exponentially complex encoding map, then although it will not be robust against the error channels which are constructed with prior access to the encoding unitary, the channel in question will be exponentially complex. So it will be extremely difficult to implement such errors. This is again a manifestation of the idea that semi-classical causality in the bulk is robust due to complexity.

## 2.4 Random circuit codes: Brownian SYK

Our goal in the rest of the paper is to study in more detail the dependence of the error correction against generic, low-rank errors acting on  $A$  relative to the complexity of the encoding unitary. It is convenient, for this purpose, to study the ensemble average of the mutual purity introduced above, but we would like to consider a one-parameter family of ensembles, labelled by the complexity of the typical unitary in the ensemble.

A simple way to generate such an ensemble is to consider the time evolution operators  $U_A = e^{-iTH}$ , for some ensemble of chaotic Hamiltonians. It is important that the Hamiltonians be chaotic, because for integrable Hamiltonians, the complexity of the time evolution operator is expected to saturate at a sub-exponential time-scale [46]. On the other hand, for chaotic Hamiltonians, it is widely expected that the complexity  $\mathcal{C}(e^{-iTH})$  grows linearly with time  $T$  for an exponential amount of time:  $\mathcal{C}(e^{-iTH}) \propto T$  (examples in [44, 46, 57, 58]). Thus, the parameter  $T$  is expected to be a good measure of the complexity for chaotic Hamiltonians for exponentially long times. Considering an ensemble of chaotic Hamiltonians then allows us to rely on this property holding only for the typical chaotic Hamiltonian, which is a much weaker assumption than expecting an arbitrary chaotic Hamiltonian to have linearly growing complexity.

More generally, we could consider unitaries  $U_A$  which are constructed from random circuits. Any unitary can be constructed as a circuit with local quantum gates — in a random circuit, we randomly choose the local gates at each instant of time from some ensemble. The resulting one-parameter family of random circuit ensembles may be thought of as a one-parameter family of measures  $d\mu(T)$  on the unitary group  $U(d_A)$ . To guarantee increasing complexity, we can choose  $d\mu(T)$  to be highly concentrated at the identity when  $T = 0$ , and as  $T$  increases we require that the support of  $d\mu(T)$  should expand outward on  $U(d_A)$  like a gas, eventually covering the entire group. If we further require that  $d\mu(T)$  approaches the Haar measure when  $T \rightarrow \infty$ , we can guarantee that the typical operator selected by averaging with  $d\mu(T)$  will have roughly increasing complexity as  $T$  increases [59].

To define such a one-parameter ensemble of random circuit codes it is convenient to pick  $U_A(T)$  to be the time evolution operator of the Brownian SYK model [37–41]. This model is constructed from  $N$  Majorana fermions  $\psi_a$ , and is defined by a set of random couplings:

$$H(t) = i^{q/2} \sum_{1 \leq a_1 < \dots < a_q \leq N} J_{a_1 \dots a_q}(t) \psi_{a_1} \dots \psi_{a_q}, \quad \{\psi_a, \psi_b\} = \delta_{ab}. \quad (2.25)$$

The coupling constants are time-dependent and are chosen to be independently Gaussian at each time point with mean zero and a fixed variance:

$$\langle J_{a_1 \dots a_q}(t) J_{b_1 \dots b_q}(t') \rangle = \delta_{a_1 b_1} \dots \delta_{a_q b_q} \frac{(q-1)!}{N^{q-1}} J^2(t, t'), \quad J^2(t, t') = J \delta(t - t'). \quad (2.26)$$

The associated encoding unitary operator is

$$U_A(T) = \mathcal{T} \exp \left( -i \int_0^T dt H(t) \right), \quad (2.27)$$

where  $\mathcal{T}$  is the time-ordering operator. Note that we are using the Brownian theory not as a model of a holographic boundary theory Hamiltonian (as has been done previously [41]), but rather as the generator of a family of holographic encoding (bulk-to-boundary) maps. Because  $H(t)$  depends on random couplings,  $U_A(T)$  is a random variable which has support on certain portions of the unitary group depending on the magnitude of  $T$ . The relevant portions are analogous to regions of space covered by a random walk of a certain fixed length.

An subtlety which we will return to later is that the SYK theory obeys certain global symmetries. The presence of these symmetries prevents the effective measure  $d\mu(T)$  from covering the entire unitary group as  $T \rightarrow \infty$ . To get around this, we will follow the strategy of [41], where a semi-classical analysis of the SYK theory gave a natural way of extracting results for SYK-like theories which do end up covering the whole unitary group.

## 2.5 Erasure errors

In order to further simplify the problem, we will consider a particular class of errors. It is important that the error channel has no prior access to the encoding unitary, i.e., we want the error to be generic and low-rank. The error channel we consider in this work will be the erasure of some subsystem  $R$ .

Let us define

$$\sigma \equiv \frac{1}{d_{\text{code}}} V V^\dagger = \frac{1}{d_{\text{code}}} \sum_i |\psi_i\rangle \langle \psi_i|, \quad (2.28)$$

and let  $\sigma_R$  be the reduced density matrix on  $R$ , and  $\sigma_L$  be the corresponding reduced density matrix on the rest of the system  $L$ . We have defined this new  $\sigma$ , which we will use for the rest of this paper, in place of the previous one in (2.19). Then, for such an erasure error,

$$\begin{aligned} \mathcal{F}_{\Psi'}(\text{ref} : \text{env}) &= \mathcal{F}_{\Psi'}(\text{ref} : R) \\ &= \text{Tr } \sigma_L^2 - \frac{1}{d_{\text{code}}} \text{Tr } \sigma_R^2, \end{aligned} \quad (2.29)$$

where recall that  $|\Psi'\rangle$  was the state which resulted from applying the error to the maximally entangled state between the reference and the code subspace. For simplicity, we will specialize to the case where  $R = A_1$  and  $L = A_2 \cup \bar{A}$ . In addition, using the fact that  $|\chi\rangle_{A_2\bar{A}}$  is maximally entangled, we then arrive at

$$\mathcal{F}_{\Psi'}(\text{ref} : R) = \left( \text{Tr} \sigma_L^2 - \frac{1}{d_{\text{code}}^2} \right). \quad (2.30)$$

Since the dimension of the environment in this case is the same as  $d_{\text{code}}$ , we find that the robustness of error correction for the above erasure error is bounded by the quantity  $d_{\text{code}} (d_{\text{code}}^2 \text{Tr} \sigma_L^2 - 1)$ .

In the next section, we will turn to the main objective of this work: computing the purity  $\text{Tr} \sigma_L^2$  for Brownian SYK codes. In particular, we are interested in the dependence of this quantity on the encoding complexity of the code, which as explained above, is linearly related to the time parameter  $T$ . From equation (2.29), we need to compute  $\text{Tr} \sigma_L^2$  as a function of  $T$ . Actually, since  $\bar{A}$  has no dynamics associated with it (i.e., there is no non-trivial time evolution operator acting on  $\bar{A}$ ), this computation boils down to a Lorentzian path integral entirely in the  $A_1 A_2$  subsystem — the relevant time contours are shown in figure 2. To arrive at figure 2, we notice that  $\text{Tr} \sigma_L^2$  involves two copies of  $U_A$  and two copies of  $U_A^\dagger$ , and so can be thought of as a matrix element of the operator  $U_A^\dagger \otimes U_A \otimes U_A^\dagger \otimes U_A$ . The matrix element in question is determined by the trace structure: since the  $R = A_1$  system is traced out first to obtain  $\sigma_L$ , the adjacent blue  $A_1$  contours are joined in figure 2, while the secondary trace over  $L$  joins the inner and outer red  $A_2$  contours.

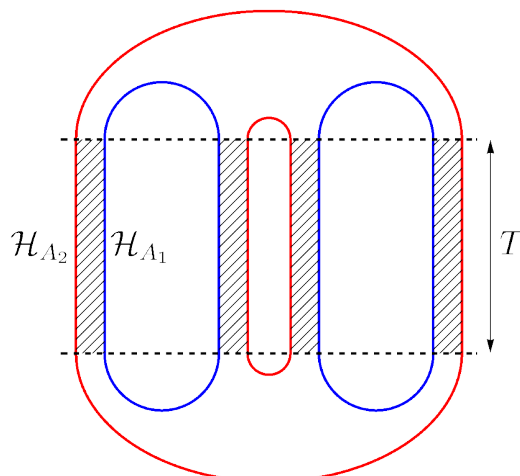
When  $T$  is small, we expect  $\text{Tr} \sigma_L^2$  to be close to 1, and so the mutual purity is non-zero. On the other hand, at very late times, we expect  $\text{Tr} \sigma_L^2$  to approach  $1/d_{\text{code}}^2$  and the mutual purity to approach zero. The intuitive argument for this is as follows: let us first purify the density matrix  $\sigma$  by including an auxiliary system  $\text{aux}$  which is isomorphic to the code subspace:

$$\begin{aligned} |\psi_\sigma\rangle &= \frac{1}{\sqrt{d_{\text{code}}}} \sum_i |i\rangle_{\text{aux}} \otimes |\psi_i\rangle \\ &= \frac{1}{\sqrt{d_{\text{code}}}} \sum_i |i\rangle_{\text{aux}} \otimes U_A(T) |i\rangle_{A_1} \otimes |\chi\rangle_{A_2, \bar{A}}. \end{aligned} \quad (2.31)$$

When  $T = 0$ , the subsystem  $A_1$  is maximally entangled with  $\text{aux}$  while  $L = A_2 \cup \bar{A}$  is in a pure state. When  $T$  becomes large (on the order of the scrambling time), we expect  $U_A(T)$  to generate nearly maximal entanglement between  $A_1$  and  $L$ . By the monogamy of entanglement, therefore,  $A_1$  cannot share much entanglement with  $\text{aux}$ . However, the unitary operator never acted on  $\text{aux}$ ; thus the reduced density matrix on  $\text{aux}$  must still be maximally mixed. We therefore conclude that both  $A_1$  and  $\text{aux}$  are close to being in a maximally entangled state with  $L$ , and so the purity of  $L$  must approach  $\frac{1}{d_{\text{code}}^2}$ . Consequently,  $\mathcal{F}_{\Psi'}(\text{ref} : R)$  should approach zero.

In what follows, we wish to understand the detailed time-dependence of the mutual purity at late times. In particular, we will demonstrate that the mutual purity becomes  $O(\frac{1}{N})$





**Figure 2.** The contour computing  $\text{Tr } \sigma_L^2$ , the quantity relevant for the mutual purity for the erasure of  $A_1$ . The red contour corresponds to the  $A_2$  fermions while the blue corresponds to the  $A_1$  fermions. The hatched regions denote an application of the time evolution operator  $U_A(T)$  or  $U_A^\dagger(T)$  of Brownian SYK, which couples the  $A_1$  and  $A_2$  systems. We have omitted the contour orientations which determine the forward and backward time evolution, but from left to right the hatched regions alternate between  $U_A^\dagger(T)$  and  $U_A(T)$ , beginning with  $U_A^\dagger(T)$ . The arcs at the top and the bottom specify the final and initial conditions respectively; in our calculation, all these arcs are actually infinitesimally small (corresponding to maximal entanglement), but they have been enlarged for visual clarity.

by the scrambling time  $T \sim \frac{1}{J} \log N$ , but then continues to decay thereafter, approaching its saturation value which is  $O(e^{-N})$  at a time of order  $T \sim \frac{1}{J} N$ . The important point is that the mutual purity keeps decaying even beyond the scrambling time, until it reaches an exponentially small plateau, which in the present model happens at polynomial time.<sup>8</sup> We will interpret this phenomenon as “complexity-protected quantum error correction”.

### 3 Erasures in Brownian SYK codes

#### 3.1 Boundary conditions and large $N$ equations

Our task now is to evaluate the path integral of Brownian SYK on the contour in figure 2. Following [41], we will use the collective-variable description of Brownian SYK. We view the path integral in figure 2 as an amplitude where we start with an “in” state, then time evolve for a time  $T$  and then take the overlap with an “out” state. The boundary conditions relevant for us are as follows. For the in boundary conditions, we have

$$\psi_{a_1}^{(1)}|\text{in}\rangle = i\psi_{a_1}^{(2)}|\text{in}\rangle, \quad \psi_{a_1}^{(3)}|\text{in}\rangle = i\psi_{a_1}^{(4)}|\text{in}\rangle, \quad (3.1)$$

$$\psi_{a_2}^{(1)}|\text{in}\rangle = i\psi_{a_2}^{(4)}|\text{in}\rangle, \quad \psi_{a_2}^{(2)}|\text{in}\rangle = i\psi_{a_2}^{(3)}|\text{in}\rangle, \quad (3.2)$$

---

<sup>8</sup>It is plausible that the time-scale at which the saturation happens is an artefact of the ensemble we have chosen, and that for other choices of ensembles, the plateau happens at exponential times.

while for the out state we have the adjoint boundary conditions:

$$\langle \text{out} | \psi_{a_1}^{(1)} = -i \langle \text{out} | \psi_{a_1}^{(2)}, \quad \langle \text{out} | \psi_{a_1}^{(3)} = -i \langle \text{out} | \psi_{a_1}^{(4)}, \quad (3.3)$$

$$\langle \text{out} | \psi_{a_2}^{(1)} = -i \langle \text{out} | \psi_{a_2}^{(4)}, \quad \langle \text{out} | \psi_{a_2}^{(2)} = -i \langle \text{out} | \psi_{a_2}^{(3)}. \quad (3.4)$$

Here  $a_1$  denotes the index of the  $N_1$  fermions corresponding to the subsystem  $A_1$ , while  $a_2$  denotes the index of the  $N_2$  fermions corresponding to the subsystem  $A_2$ . The superscript index  $(i)$  on  $\psi_a^{(i)}$  (where  $i = 1, \dots, 4$ ) labels the four contour segments corresponding to real time evolution. From left to right in figure 2, we label the contours 1, 2, 3, and 4.

In order to evaluate the path integral, it is convenient to define the two matrices:

$$g_{ij}^{(1)}(t) = \frac{1}{N_1} \sum_{a_1} \langle \psi_{a_1}^{(i)}(t) \psi_{a_1}^{(j)}(t) \rangle, \quad g_{ij}^{(2)}(t) = \frac{1}{N_2} \sum_{a_2} \langle \psi_{a_2}^{(i)}(t) \psi_{a_2}^{(j)}(t) \rangle, \quad (3.5)$$

which we can think of as the singlet part of the fermion two-point functions in the  $A_1$  and  $A_2$  sectors respectively. Here we inserted the operators on the right hand side at the specified time into the path integral in figure 2. We will soon see that these two sets of variables control the classical limit of the Brownian theory on this contour. To solve the classical equations of motion we will obtain in this limit, we require the boundary conditions that are implied by the in and out state relations above. It is also convenient to define the total two-point function (i.e., the summed two-point function of all the fermions):

$$g_{ij} = \lambda g_{ij}^{(1)} + (1 - \lambda) g_{ij}^{(2)}, \quad (3.6)$$

where we have introduced the parameter  $\lambda = \frac{N_1}{N}$ . When evaluating the path integral at large  $N$ , it will be convenient to take the double scaling limit:

$$N_1 \rightarrow \infty, \quad N_2 \rightarrow \infty, \quad \lambda = \frac{N_1}{N} \text{ fixed}. \quad (3.7)$$

In fact,  $A_1$  has the same dimension as the code subspace, so we would like to take  $N_1$  much smaller than  $N_2$ . This corresponds to taking  $\lambda \ll 1$ . Thus, we can take  $\lambda$  to be a small (but  $O(N^0)$ ) parameter and work in perturbation theory in  $\lambda$ . This makes some of the path integral calculations analytically tractable.

Note that both the in and out boundary conditions have the property that (recall that the fermions are normalized such that  $\psi^2 = 1/2$ ):

$$\psi_{a_1}^{(1)} \psi_{a_1}^{(2)} \psi_{a_1}^{(3)} \psi_{a_1}^{(4)} | \text{in} \rangle = -\frac{1}{4} | \text{in} \rangle, \quad (3.8)$$

$$\psi_{a_2}^{(1)} \psi_{a_2}^{(2)} \psi_{a_2}^{(3)} \psi_{a_2}^{(4)} | \text{in} \rangle = -\frac{1}{4} | \text{in} \rangle. \quad (3.9)$$

Since  $A_1$  fermions lie in the same parity sector as the  $A_2$  fermions, and the (effective) Hamiltonian commutes with the fermion parity operator after averaging (see appendix A, equation (A.2) and discussion below it), the above relations should hold at any time. Equations (3.8) and (3.9) imply the following symmetry properties:

$$g_{12}^{(1)} = g_{34}^{(1)}, \quad g_{14}^{(1)} = g_{23}^{(1)}, \quad g_{24}^{(1)} = -g_{13}^{(1)}, \quad (3.10)$$

$$g_{12}^{(2)} = g_{34}^{(2)}, \quad g_{14}^{(2)} = g_{23}^{(2)}, \quad g_{24}^{(2)} = -g_{13}^{(2)}. \quad (3.11)$$

These should hold at all times because time evolution preserves fermion parity flavor-wise. So the evolution reduces to the six variables  $x_\alpha = 2ig_{12}^{(\alpha)}$ ,  $y_\alpha = 2g_{13}^{(\alpha)}$  and  $z_\alpha = 2ig_{14}^{(\alpha)}$ , where  $\alpha = 1, 2$ . We can now rewrite the initial and final boundary conditions in terms of these new variables as

$$x_1(0) = 1, \quad y_1(0) = z_1(0), \quad (3.12)$$

$$x_2(0) = -y_2(0), \quad z_2(0) = 1. \quad (3.13)$$

$$x_1(T) = 1, \quad y_1(T) = -z_1(T), \quad (3.14)$$

$$x_2(T) = y_2(T), \quad z_2(T) = 1. \quad (3.15)$$

It is convenient to also introduce the total variables  $x = \lambda x_1 + (1 - \lambda)x_2$ , and similarly for  $y$  and  $z$ . The above boundary conditions imply the following constraints in terms of the  $(x, y, z)$  variables:

$$z(0) - x(0) - y(0) = (1 - 2\lambda), \quad z(T) - x(T) + y(T) = (1 - 2\lambda). \quad (3.16)$$

In order to proceed with the evaluation of the Lorentzian path integral in figure 2, recall [41] that the action for the Brownian SYK model on the contour in figure 2 is given by

$$I = \frac{1}{2} \int_0^T dt \left( \psi_a^{(j)} \partial_t \psi_a^{(j)} + i^{\frac{q}{2}} s_j J_{a_1 \dots a_q} \psi_{a_1 \dots a_q}^{(j)} \right), \quad (3.17)$$

where the flavor indices run over  $a = 1, \dots, N$  (i.e., over both  $A_1$  as well as  $A_2$  fermions), and we have introduced the notation  $\psi_{a_1 \dots a_q}^{(j)} = \psi_{a_1}^{(j)} \dots \psi_{a_q}^{(j)}$ . The quantity  $s_j$  is given by

$$s_j = \begin{cases} +i, & j \in \{2, 4\} \\ -i \cdot i^q, & j \in \{1, 3\}, \end{cases} \quad (3.18)$$

and is related to the difference between forward and backward time evolution (see [41] for details). We now wish to perform the average over the couplings. Using

$$\langle J_{a_1 \dots a_q}(t) J_{b_1 \dots b_q}(t') \rangle = \delta_{a_1 b_1} \dots \delta_{a_q b_q} \frac{(q-1)!}{N^{q-1}} J^2(t, t'), \quad (3.19)$$

the action obtained after ensemble averaging over the couplings is given by<sup>9</sup>

$$\begin{aligned} I &= \frac{1}{2} \int_0^T dt \psi_a^{(j)} \partial_t \psi_a^{(j)} - \frac{i^q (q-1)!}{2N^{q-1}} \iint_0^T dt dt' J^2(t, t') s_j s_{j'} \psi_{a_1 \dots a_q}^{(j)}(t) \psi_{a_1 \dots a_q}^{(j')}(t') \\ &= \frac{1}{2} \int_0^T dt \psi_a^{(j)} \partial_t \psi_a^{(j)} - \frac{N}{2q} \iint_0^T dt dt' J^2(t, t') s_j s_{j'} \left( \frac{1}{N} \psi_a^{(j)}(t) \psi_a^{(j')}(t') \right)^q. \end{aligned} \quad (3.20)$$

At this stage, it is convenient to introduce the collective  $(G, \Sigma)$  variables. Since we have two sets of fermions corresponding to  $A_1$  and  $A_2$ , we introduce two collective fields

$$G_{ij}^{(1)}(t, t') = \frac{1}{N_1} \sum_{a=1}^{N_1} \psi_a^{(i)}(t) \psi_a^{(j)}(t'), \quad G_{ij}^{(2)}(t, t') = \frac{1}{N_2} \sum_{a=N_1+1}^N \psi_a^{(i)}(t) \psi_a^{(j)}(t'), \quad (3.21)$$

<sup>9</sup>In the second step, we have made the same imprecise replacement of the Hamiltonian as in [41], discussed in more detail in appendix A.3 of [40].

and the corresponding Lagrange multipliers  $\Sigma_{ij}^{(1)}(t, t')$  and  $\Sigma_{ij}^{(2)}(t, t')$  to impose the constraints. We can now integrate out the fermions. The action in terms of the collective variables is

$$\begin{aligned}
 -\frac{I}{N} &= \lambda \log \text{Pf}(\partial_t - \Sigma^{(1)}) + (1 - \lambda) \log \text{Pf}(\partial_t - \Sigma^{(2)}) \\
 &\quad - \frac{1}{2} \int_0^T \int_0^T dt dt' \left[ \lambda \Sigma_{ij}^{(1)}(t, t') G_{ij}^{(1)}(t, t') + (1 - \lambda) \Sigma_{ij}^{(2)}(t, t') G_{ij}^{(2)}(t, t') \right] \\
 &\quad + \frac{1}{2q} \int_0^T \int_0^T dt dt' J^2(t, t') s_j s_{j'} G_{jj'}(t, t')^q,
 \end{aligned} \tag{3.22}$$

where Pf is the Pfaffian, and we have defined

$$G_{ij}(t, t') = \lambda G_{ij}^{(1)}(t, t') + (1 - \lambda) G_{ij}^{(2)}(t, t'). \tag{3.23}$$

In the large  $N$  limit, the path integral over the collective variables can be performed in the saddle point approximation. The equations of motion corresponding to the above action are:

$$\partial_t G_{jj'}^{(\alpha)}(t, t') - \Sigma_{jk}^{(\alpha)} \star G_{kj'}^{(\alpha)}(t, t') = \delta(t - t') \delta_{jj'}, \tag{3.24}$$

$$\Sigma_{jj'}^{(\alpha)} = s_j s_{j'} J^2(t, t') G_{jj'}^{(\alpha)}(t, t')^{q-1}, \tag{3.25}$$

where  $\alpha = 1$  corresponds to the fermions in  $A_1$  while  $\alpha = 2$  corresponds to the fermions in  $A_2$ , the repeated  $k$  index is summed, and the star product between two bi-local fields is defined as

$$(A \star B)(t, t') = \int dt'' A(t, t'') B(t'', t). \tag{3.26}$$

Using the fact that  $G^{(\alpha)}$  and  $\Sigma^{(\alpha)}$  are both anti-symmetric, we can rewrite these equations in a more convenient form:

$$(\partial_t + \partial_{t'}) G_{jj'}^{(\alpha)} = \left( \Sigma_{jk}^{(\alpha)} \star G_{kj'}^{(\alpha)} - G_{jk}^{(\alpha)} \star \Sigma_{kj'}^{(\alpha)} \right), \tag{3.27}$$

$$\Sigma_{jj'}^{(\alpha)} = s_j s_{j'} J^2(t, t') G_{jj'}^{(\alpha)q-1}. \tag{3.28}$$

Now, a simplification happens in the Brownian SYK model — recall that for Brownian SYK,  $J^2(t, t') = J\delta(t - t')$ . As a result,  $\Sigma$  is “diagonal” (in time), and only the diagonal components of all the collective variables are relevant; the off-diagonal components drop out of the equations of motion. In fact, it is easy to see from the action that for Brownian SYK, the off-diagonal modes do not have any interesting dynamics and can be integrated out of the full path integral trivially [41].

Let us denote the diagonal components of the collective variables as

$$G_{ij}^{(\alpha)}(t, t) = g_{ij}^{(\alpha)}(t), \quad \Sigma_{ij}^{(\alpha)}(t, t') = \delta(t - t') \sigma_{ij}^{(\alpha)}(t), \tag{3.29}$$

The resulting equations of motion for the  $(g, \sigma)$  variables are

$$\frac{dg^{(\alpha)}}{dt} = \left[ \sigma^{(\alpha)}(t), g^{(\alpha)}(t) \right], \quad \sigma_{ij}^{(\alpha)}(t) = \sigma_{ij}(t) \equiv \begin{cases} Js_i s_j (g_{ij}(t))^{q-1}, & i \neq j \\ 0, & i = j \end{cases} \tag{3.30}$$

where the equation on the left is written for the  $g$  and  $\sigma$  matrices and

$$g(t) = \lambda g^{(1)}(t) + (1 - \lambda)g^{(2)}(t). \quad (3.31)$$

In terms of the  $(x_\alpha, y_\alpha, z_\alpha)$  variables, we get the following equations of motion:

$$\begin{aligned} \dot{x}_\alpha &= \frac{J}{2^{q-2}}(y^{q-1}z_\alpha - z^{q-1}y_\alpha) \\ \dot{y}_\alpha &= \frac{J}{2^{q-2}}(x^{q-1}z_\alpha - z^{q-1}x_\alpha) \\ \dot{z}_\alpha &= \frac{J}{2^{q-2}}(x^{q-1}y_\alpha - y^{q-1}x_\alpha), \end{aligned} \quad (3.32)$$

where recall that  $x = \lambda x_1 + (1 - \lambda)x_2$ , and so on. These relations imply the following equations of motion for the total variables:

$$\begin{aligned} \dot{x} &= \frac{J}{2^{q-2}}(y^{q-1}z - z^{q-1}y) \\ \dot{y} &= \frac{J}{2^{q-2}}(x^{q-1}z - z^{q-1}x) \\ \dot{z} &= \frac{J}{2^{q-2}}(x^{q-1}y - y^{q-1}x). \end{aligned} \quad (3.33)$$

As an aside, the above equations of motion have a Hamiltonian structure. To see this, let us denote  $\mathbf{x}_\alpha = (x_\alpha, y_\alpha, z_\alpha)$  and  $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2 = (x, y, z)$ . Further, we define  $h_2(\mathbf{x}) = \frac{1}{2}(x^2 - y^2 + z^2)$  and  $h_q(\mathbf{x}) = \frac{1}{q}(x^q - y^q + z^q)$ . Then, equations (3.32) take the succinct form

$$\dot{\mathbf{x}}_\alpha^I = \omega_\alpha^{IJ} \frac{\partial h_q(\mathbf{x})}{\partial \mathbf{x}_\alpha^J}, \quad \omega_\alpha^{IJ} = -p_\alpha \epsilon^{IJK} \frac{\partial h_2(\mathbf{x}_\alpha)}{\partial \mathbf{x}_\alpha^K}, \quad (3.34)$$

where  $p_\alpha = (\frac{1}{\lambda}, \frac{1}{1-\lambda})$ . Similarly, the equations for the total variables take the form

$$\dot{\mathbf{x}}^I = \omega^{IJ} \frac{\partial h_q(\mathbf{x})}{\partial \mathbf{x}^J}, \quad \omega^{IJ} = -\epsilon^{IJK} \frac{\partial h_2(\mathbf{x})}{\partial \mathbf{x}^K}. \quad (3.35)$$

Thus, these equations take the form of Hamilton's equations of motion — the underlying phase space is that of two copies, labelled by  $\alpha$ , of a co-adjoint orbit of  $\mathfrak{sl}(2, \mathbb{R})$  specified by a constant value of the conserved quantity  $h_2(\mathbf{x})$ .<sup>10</sup> The Hamiltonian  $h_q(\mathbf{x})$  couples the two copies, with the effective coupling constant being  $\lambda$ . It may seem unusual that we have an odd number of variables (e.g.  $(x, y, z)$ ) in Hamiltonian mechanics, but this is simply because we have parametrized the two-dimensional dynamics on the hypersurface  $h_2(\mathbf{x}) = \text{const.}$  in terms of coordinates in the ambient  $\mathbb{R}^3$ .

---

<sup>10</sup>The solutions we will find turn out to have  $2h_2(\mathbf{x}) = 1$ , so for the total variables  $\mathbf{x}$  the orbits in question are related to the “continuous series” of unitary  $\mathfrak{sl}(2, \mathbb{R})$  representations by geometric quantization. Understanding the significance of this structure is an interesting problem in its own right, but we will not address it further in this work.

**Normalization.** The fermionic path integral depicted in figure 2 is the result of writing  $\text{Tr } \sigma_L^2$  in a Hilbert space form and replacing the maximally entangled state projectors with connections in the contour. However, translating a Hilbert space expression into a fermionic path integral comes with a standard normalization issue since Majorana fermions have a somewhat unusual Hilbert space interpretation (when they admit one at all). So, we must relate the path integral  $Z(T)$  to  $\text{Tr } \sigma_L^2(T)$  with an overall normalization that ensures  $\text{Tr } \sigma_L^2(T=0) = 1$ .

When  $T = 0$ , the contour in figure 2 consists of four disconnected circles which are not coupled by any time evolution. The result of the path integral for a single free Majorana fermion on a circular contour of length  $T$  with antiperiodic boundary conditions is actually equal to  $\sqrt{2}$ , independent of  $T$ , so in the limit  $T \rightarrow 0$  we still have  $\sqrt{2}$ . As such, when  $T = 0$ , the contour in figure 2 yields  $Z(T=0) = 2^N$  since there are  $N_1$  Majorana fermions on the two  $A_1$  circles and  $N_2$  Majorana fermions on the two  $A_2$  circles. So we must relate the path integral  $Z(T)$  to  $\text{Tr } \sigma_L^2(T)$  by the formula

$$\text{Tr } \sigma_L^2(T) = 2^{-N} Z(T). \quad (3.36)$$

### 3.2 Solutions: qualitative discussion

We will first qualitatively discuss what the solutions to the equations of motion (3.32) should look like, leaving a quantitative treatment for sections 3.3, 3.4, and 3.5. When  $N_1 = 0$  (i.e.,  $\lambda = 0$ ), then the  $\mathbf{x}_2$  equations are easy to solve. In this case, the boundary conditions imply that the solution stays at the fixed point  $\mathbf{x}_2^* = (0, 0, 1)$ . When  $\lambda$  is small but non-zero, we expect that this saddle point remains, but with small corrections. In particular, the  $\mathbf{x}_2$  variables will stay close to their original fixed point values. The corrections to the  $\mathbf{x}_2$  solutions can be obtained in perturbation theory in  $\lambda$ , and we describe them in detail in section 3.3. (Recall from the discussion under (3.7) that  $\lambda \ll 1$ .) This resulting solution is the dominant saddle point at small times, and is the analogue of the “disconnected” contribution in equation (2.17), or the disconnected geometry in JT gravity [31].

When  $\lambda$  is small, to zeroth order, the solution for  $\mathbf{x}_2$  variables will be unaffected by the  $\mathbf{x}_1$  variables, and in particular will correspond to a fixed point of the Hamiltonian picked out by the boundary conditions as we have just described. However, the initial backreaction on the  $\mathbf{x}_1$  variables will be large. The source of this strong backreaction is the mismatch between the boundary conditions of the  $\mathbf{x}_1$  and  $\mathbf{x}_2$  variables. As the  $A_1$  system is small compared to  $A_2$ , the Brownian dynamics quickly thermalizes the  $A_1$  system so that the correlation between contours (in figure 2) can be measured with any subset of all  $N$  fermions; at the level of the solutions, this means that we will have  $\mathbf{x}_1 \approx \mathbf{x}_2$  at all times except in small neighborhoods around  $t = 0$  and  $t = T$  where we expect large transient behaviors for  $\mathbf{x}_1$  to arrive at their “thermalized” values. These transient behaviors can be computed analytically in perturbation theory for the disconnected solution and must be treated numerically otherwise. These qualitative properties of the  $\mathbf{x}_1$  solutions hold both for the disconnected solution in section 3.3 as well as the other solutions we now describe.

In addition to generating nontrivial time-dependence for the disconnected solution, turning on a small  $\lambda$  has another important effect — it gives rise to new “tunneling” solutions

(i.e., instantons) which are absent at  $\lambda = 0$ . The tunneling allows the  $\mathbf{x}_2$  variables to jump between different fixed points; the leading tunneling solution jumps from  $\mathbf{x}_2(0) \approx (1, -1, 1)$  to  $\mathbf{x}_2(T) \approx (1, 1, 1)$ . The  $\mathbf{x}_1$  variables again have large transient signals near  $t = 0$  and  $t = T$ , but this time both transients are different than the ones which occur for the disconnected solution due to the different fixed points approached by the  $\mathbf{x}_2$  variables, and in fact these are the only other two types of transient behavior which can occur. This saddle point, which we describe in section 3.4, is the analogue of the “connected” saddle point in equation (2.17), or the “wormhole” in JT gravity [31]. While it is suppressed by a factor of  $e^{-(1-2\lambda)N}$  relative to the leading, disconnected solution, the contribution of the disconnected saddle point decays exponentially in time. So, at a time  $t_* \sim O(N)$ , there is an exchange of dominance between these two saddle points.

There are also other tunneling saddle points, described in section 3.5, where the  $\mathbf{x}_2$  variables tunnel back and forth multiple times between the two possible initial fixed points  $\mathbf{x}_2 \approx (0, 0, 1)$  and  $\mathbf{x}_2 \approx (1, -1, 1)$  and the two possible final fixed points  $\mathbf{x}_2 \approx (0, 0, 1)$  and  $\mathbf{x}_2 \approx (1, 1, 1)$ ; these are even more subleading in powers of  $e^{-N}$ , and occur with all possible combinations of the previously described types of transient behaviors for the  $\mathbf{x}_1$  variables. Explicitly, there are two possible behaviors at  $t = 0$  and two at  $t = T$  corresponding to the possible initial and final fixed points for  $\mathbf{x}_2$ , and all four combinations of initial and final transient behaviors occur in the multiply tunneling solutions. These multiply tunneling solutions show interesting behavior as a function of  $T$ . We will see that they become genuine solutions of the equations (3.32) only after certain critical values of  $T$ , related to the scrambling time. Before these critical times, these configurations are actually off-shell. Configurations which tunnel more times take longer to become solutions. As the presence of these contributions is important for unitarity of the overall evolution [41], it is intriguing that they can be invisible on-shell for a parametrically (though not polynomially) long time in  $N$ .

In summary, we began with the goal of studying the error correction dynamics of a family of unitary operators with increasing average circuit complexity. The specific family which we chose for convenience was the set of time evolution operators in the Brownian SYK model, a family of time-dependent Hamiltonians which are essentially a continuous random circuit. We found that the error correction dynamics are governed by the quantity  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$ , and this quantity in turn depends on a Brownian SYK path integral (figure 2). What we have just discussed are the saddle point solutions to that path integral. Evaluating the effective action of these solutions will allow us to draw conclusions about the error correction behavior of the family of unitary operators with increasing complexity.

### 3.3 Disconnected solution

We will first solve for the disconnected solution at small, non-zero  $\lambda$ , and evaluate its on-shell action together with the one-loop determinant. We begin by expanding our variables in a power series expansion in  $\lambda$ :

$$\mathbf{x}_\alpha = \sum_{n=0}^{\infty} \lambda^n \mathbf{x}_\alpha^{(n)}, \quad \mathbf{x} = \sum_{n=0}^{\infty} \lambda^n \mathbf{x}^{(n)}. \quad (3.37)$$



At  $O(\lambda^0)$ , we must have  $\mathbf{x}^{(0)} = \mathbf{x}_2^{(0)}$ . Therefore, the boundary conditions, equations (3.14) and (3.15), imply that at leading order these variables sit at a fixed point of the Hamiltonian:

$$\mathbf{x}^{(0)}(t) = \mathbf{x}_2^{(0)}(t) = (0, 0, 1). \quad (3.38)$$

After substituting these solutions in (3.32), we get the following equations for  $\mathbf{x}_1^{(0)}$ :

$$\begin{aligned} \dot{x}_1^{(0)} &= -\frac{J}{2^{q-2}} y_1^{(0)}, \\ \dot{y}_1^{(0)} &= -\frac{J}{2^{q-2}} x_1^{(0)}, \\ \dot{z}_1^{(0)} &= 0. \end{aligned} \quad (3.39)$$

We need to solve these equations with the boundary conditions (3.12) and (3.14). The solution is

$$\begin{aligned} x_1^{(0)}(t) &= \frac{\cosh\left(\frac{J}{2^{q-2}}\left(t - \frac{T}{2}\right)\right)}{\cosh\left(\frac{JT}{2^{q-1}}\right)}, \\ y_1^{(0)}(t) &= -\frac{\sinh\left(\frac{J}{2^{q-2}}\left(t - \frac{T}{2}\right)\right)}{\cosh\left(\frac{JT}{2^{q-1}}\right)}, \\ z_1^{(0)}(t) &= \tanh\left(\frac{JT}{2^{q-1}}\right). \end{aligned} \quad (3.40)$$

We can think of this solution as the backreaction of the  $\mathbf{x}_2^{(0)}$  variables on the  $\mathbf{x}_1^{(0)}$  variables. For instance,  $x_1$  starts off at one at  $t = 0$ , but after a brief transient behavior, it thermalizes to a small value of about  $e^{-\frac{JT}{2^{q-1}}}$  owing to its coupling to the  $\mathbf{x}_2$  variables, which act like a bath and dynamically force  $\mathbf{x}_1 \approx \mathbf{x}_2$ . There is another transient near  $t = T$ , where  $\mathbf{x}_1$  again deviates from  $\mathbf{x}_2$  significantly to reach the final boundary condition.

At order  $\lambda$ , we note from  $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$ , that

$$\mathbf{x}^{(1)} = \mathbf{x}_2^{(1)} + \mathbf{x}_1^{(0)} - \mathbf{x}_2^{(0)}. \quad (3.41)$$

Now we use the  $O(\lambda^0)$  solutions to find the following boundary conditions up to  $O(\lambda)$ :

$$\begin{aligned} x_1^{(1)}(0) &= 0, & y_1^{(1)}(0) &= z_1^{(1)}(0), \\ x_1^{(1)}(T) &= 0, & y_1^{(1)}(T) &= -z_1^{(1)}(T), \end{aligned} \quad (3.42)$$

$$\begin{aligned} x_2^{(1)}(0) &= -y_2^{(1)}(0), & z_2^{(1)}(0) &= 0, \\ x_2^{(1)}(T) &= y_2^{(1)}(T), & z_2^{(1)}(T) &= 0, \end{aligned} \quad (3.43)$$

$$\begin{aligned} \mathbf{x}^{(1)}(0) &= \left(1 + x_2^{(1)}(0), \tanh\left(\frac{JT}{2^{q-1}}\right) + y_2^{(1)}(0), \tanh\left(\frac{JT}{2^{q-1}}\right) - 1\right), \\ \mathbf{x}^{(1)}(T) &= \left(1 + x_2^{(1)}(T), -\tanh\left(\frac{JT}{2^{q-1}}\right) + y_2^{(1)}(T), \tanh\left(\frac{JT}{2^{q-1}}\right) - 1\right). \end{aligned} \quad (3.44)$$

With these boundary conditions in hand, we can in principle solve for all the variables at  $O(\lambda)$ . However, in what follows, we will only need  $\mathbf{x}^{(1)}$  in order to evaluate the on-shell action up to  $O(\lambda)$ . The corresponding differential equations are given by

$$\begin{aligned}\dot{x}^{(1)} &= -\frac{J}{2^{q-1}} y^{(1)}, \\ \dot{y}^{(1)} &= -\frac{J}{2^{q-1}} x^{(1)}, \\ \dot{z}^{(1)} &= 0.\end{aligned}\tag{3.45}$$

The solution is:

$$\begin{aligned}x^{(1)}(t) &= \left(1 + \tanh\left(\frac{JT}{2^{q-1}}\right)\right) \cosh\left(\frac{J}{2^{q-2}}\left(t - \frac{T}{2}\right)\right) e^{-\frac{JT}{2^{q-1}}}, \\ y^{(1)}(t) &= -\left(1 + \tanh\left(\frac{JT}{2^{q-1}}\right)\right) \sinh\left(\frac{J}{2^{q-2}}\left(t - \frac{T}{2}\right)\right) e^{-\frac{JT}{2^{q-1}}}, \\ z^{(1)}(t) &= \tanh\left(\frac{JT}{2^{q-1}}\right) - 1.\end{aligned}\tag{3.46}$$

**Classical on-shell action.** Having obtained the classical solutions, we can evaluate the action in (3.22) at leading order in  $\lambda$ . We first compute the Pfaffian for the  $A_1$  fermions:

$$\text{Pf}(\partial_t - \sigma^{(1)}) = \int_{\psi^{(1)}=i\psi^{(2)}, \psi^{(3)}=i\psi^{(4)}}^{\psi^{(1)}=-i\psi^{(2)}, \psi^{(3)}=-i\psi^{(4)}} \mathcal{D}\psi^{(1)} \dots \mathcal{D}\psi^{(4)} \exp\left(-\frac{1}{2} \int_0^T dt \left[\psi^{(j)} \partial_t \psi^{(j)} - \sigma_{jj'}(t) \psi^{(j)} \psi^{(j')}\right]\right).\tag{3.47}$$

On the right hand side of the above expression, we have used the fact that on-shell,  $\sigma_{ij}^{(1)}(t) = \sigma_{ij}(t)$ ; see equation (3.30). Following [41], we can write the Pfaffian in a Hilbert space representation in terms of a single qubit:<sup>11</sup>

$$\text{Pf}(\partial_t - \sigma^{(1)}) = 2 \times \langle + | \mathcal{T} \exp\left(-\int_0^T dt h(t)\right) | + \rangle,\tag{3.48}$$

where  $h(t)$  is the qubit Hamiltonian:<sup>12</sup>

$$h(t) = \frac{J}{2^{q-1}} \left(-x^{q-1}(t)X + iy^{q-1}(t)Y - z^{q-1}(t)Z\right),\tag{3.49}$$

and  $X, Y, Z$  are the Pauli matrices. Further,  $|+\rangle$  is an eigenstate of the Pauli  $X$  operator with the eigenvalue  $+1$ :  $X|+\rangle = |+\rangle$ . The initial and final states are fixed by the boundary conditions of the path integral. Since the leading contribution to  $x^{q-1}(t), y^{q-1}(t)$  is at  $O(\lambda^{q-1})$  we can ignore them in the evaluation of the Pfaffian (assuming  $q \geq 4$ ). The Pfaffian is therefore given by

$$\begin{aligned}\text{Pf}(\partial_t - \sigma^{(1)}) &= 2 \times \langle + | \exp\left(\int_0^T dt \frac{J}{2^{q-1}} z^{q-1}(t)Z\right) | + \rangle + O(\lambda^{q-1}) \\ &= 2 \cosh\left(\frac{JT}{2^{q-1}} z^{q-1}(0)\right) + O(\lambda^2).\end{aligned}\tag{3.50}$$

<sup>11</sup>Naively, we would need two qubits given that there are four Majorana fermions. However, since time evolution preserves fermion number, we need only use one qubit.

<sup>12</sup>The factor of 2 appearing in (3.48) ensures that, when  $T = 0$ , the Pfaffian gives 2, as this is the result for a single Majorana fermion path integral on two disjoint circles of any length with antiperiodic boundary conditions.

In the second step, we used the fact that  $z(t)$  is a constant at  $O(\lambda)$ . Similarly, we can evaluate the Pfaffian for the  $A_2$  fermions:

$$\begin{aligned}
 \text{Pf}(\partial_t - \sigma^{(2)}) &= \int_{\psi^{(1)}=i\psi^{(4)}, \psi^{(2)}=i\psi^{(3)}}^{\psi^{(1)}=-i\psi^{(4)}, \psi^{(2)}=-i\psi^{(3)}} \mathcal{D}\psi^{(1)} \dots \mathcal{D}\psi^{(4)} \exp \left( -\frac{1}{2} \int_0^T dt \left[ \psi^{(j)} \partial_t \psi^{(j)} - \sigma_{jj'}(t) \psi^{(j)} \psi^{(j')} \right] \right) \\
 &= 2 \times \langle 0 | \exp \left( \frac{JT}{2^{q-1}} z^{q-1}(0) Z \right) | 0 \rangle + O(\lambda^{q-1}) \\
 &= 2 \exp \left( \frac{JT}{2^{q-1}} z^{q-1}(0) \right) + O(\lambda^{q-1}).
 \end{aligned} \tag{3.51}$$

Here,  $|0\rangle$  is an eigenstate of the Pauli  $Z$  operator with eigenvalue  $+1$ :  $Z|0\rangle = |0\rangle$ .

Having evaluated the Pfaffians to  $O(\lambda)$ , we now evaluate the rest of the terms in the action following [41]:

$$\begin{aligned}
 & -\frac{1}{2} \iint_0^T dt dt' \left[ \lambda \Sigma_{ij}^{(1)}(t, t') G_{ij}^{(1)}(t, t') + (1 - \lambda) \Sigma_{ij}^{(2)}(t, t') G_{ij}^{(2)}(t, t') \right] \\
 & + \frac{1}{2q} \iint_0^T dt dt' J^2(t, t') s_j s_{j'} G_{jj'}(t, t')^q \\
 & = -\frac{JT}{2^{q-1}q} - \frac{1}{2} \int_0^T dt \left[ \sigma_{ij}(t) g_{ij}(t) - \frac{1}{q} s_j s_{j'} g_{jj'}(t, t')^q \right] \\
 & = -\frac{JT}{2^{q-1}q} - \frac{JT}{2^{q-1}} \frac{q-1}{q} r \\
 & = \frac{JT}{2^{q-1}q} (r-1) - \frac{JT}{2^{q-1}} r,
 \end{aligned} \tag{3.52}$$

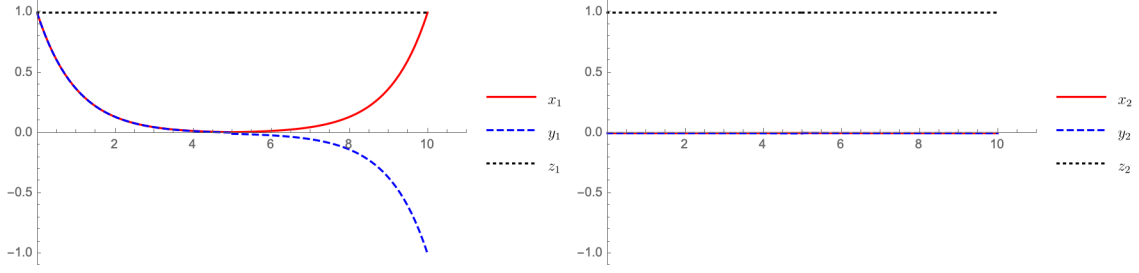
where  $r = qh_q(\mathbf{x}) = x^q(t) - y^q(t) + z^q(t)$  is a constant of motion, namely the total energy. Note that this form of the bulk contribution to the action is independent of the particular solution we are considering. Now, we can combine the above terms with equations (3.50) and (3.51) to obtain the full on-shell action for the disconnected saddle point:

$$\begin{aligned}
 \frac{-I}{N} &= \log 2 + \lambda \log \cosh \left( \frac{JT}{2^{q-1}} z^{q-1}(0) \right) + (1 - \lambda) \frac{JT}{2^{q-1}} z^{q-1}(0) + \frac{JT}{2^{q-1}q} (r-1) - \frac{JT}{2^{q-1}} r \\
 &= \log 2 + \lambda \log \cosh \left( \frac{JT}{2^{q-1}} \right) + \frac{JT}{2^{q-1}} \left( (1 - \lambda) z^{q-1}(0) + \frac{r-1}{q} - r \right) \\
 &= \log 2 + \lambda \left( \log \cosh \left( \frac{JT}{2^{q-1}} \right) - \frac{JT}{2^{q-1}} \right) + O(\lambda^2).
 \end{aligned} \tag{3.53}$$

As the normalization relation (3.36) cancels the leading  $\log 2$  in the effective action, the contribution of the disconnected saddle point in the large  $N$  limit is given by

$$\text{Tr } \sigma_L^2 \Big|_{\text{disc}} \approx \left( \frac{1 + \exp \left( -\frac{JT}{2^{q-2}} \right)}{2} \right)^{N\lambda}. \tag{3.54}$$

This formula is consistent with physical expectations; see the discussion around equation (2.31). At small times  $\frac{JT}{2^{q-1}} \ll 1$ , we find that  $\text{Tr } \sigma_L^2 \rightarrow 1$ . This is expected, since



**Figure 3.** The disconnected solution for the  $\mathbf{x}_1$  (left) and  $\mathbf{x}_2$  (right) variables with  $J = q = 4$ ,  $r = 1 - 10^{-8}$ , and  $\lambda = 10^{-4}$ . There is a long region where  $\mathbf{x}_1 \approx \mathbf{x}_2$  around the fixed point  $(0, 0, 1)$ , while there are large transient behaviors where  $\mathbf{x}_1 \neq \mathbf{x}_2$  near  $t = 0$  and  $t = T \approx 10$  due to the difference between the  $\mathbf{x}_1$  and  $\mathbf{x}_2$  boundary conditions. The difference between the exact numerical  $\mathbf{x}_1$  solution shown here and the perturbative solution (3.40) is essentially invisible.

in this limit,  $U_A(T)$  does not introduce much entanglement between  $L$  and  $A_1$ . On the other hand, at late times  $\frac{JT}{2^{q-1}} \gg 1$ ,  $\text{Tr } \sigma_L^2 \rightarrow 2^{-N\lambda} = \frac{1}{d_{\text{code}}^2}$ , as we anticipated based on monogamy of entanglement. As a further check, we also reproduce the above formula from a “Hamiltonian” point of view in appendix A. For now, we proceed to evaluate the one-loop determinant around the disconnected solution. But, before doing so, we display the numerical solutions for  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in figure 3, where we can clearly see the leading order nontrivial time-dependence of  $\mathbf{x}_1$  takes the rough hyperbolic forms we found for  $\mathbf{x}_1^{(0)}(t)$  in (3.40).

**One-loop determinant.** To compute the one-loop determinant in the path integral formalism, we need to expand the action around the saddle point and integrate over small fluctuations. We will follow the notations and conventions of appendix B in [41]. Recall that the action is given by

$$\begin{aligned} \frac{-I}{N} &= \lambda \log \text{Pf}(\partial_t - \sigma^{(1)}) + (1 - \lambda) \log \text{Pf}(\partial_t - \sigma^{(2)}) \\ &\quad - \frac{1}{2} \int_0^T dt \left[ \lambda \sigma_{ij}^{(1)}(t) g_{ij}^{(1)}(t) + (1 - \lambda) \sigma_{ij}^{(2)}(t) g_{ij}^{(2)}(t) \right] \\ &\quad + \frac{J}{2q} \int_0^T dt s_j s_{j'} g_{jj'}(t)^q. \end{aligned} \quad (3.55)$$

We can write the Pfaffian in the Hilbert space representation, as in equations (3.50) and (3.51):

$$\begin{aligned} -\frac{I}{N} &= \log 2 + \lambda \log \left[ \langle + | \exp \left( \int_0^T dt \left( -\sigma_x^{(1)}(t) X + i \sigma_y^{(1)}(t) Y - \sigma_z^{(1)}(t) Z \right) \right) | + \rangle \right] \\ &\quad + (1 - \lambda) \log \left[ \langle 0 | \exp \left( \int_0^T dt \left( -\sigma_x^{(2)}(t) X + i \sigma_y^{(2)}(t) Y - \sigma_z^{(2)}(t) Z \right) \right) | 0 \rangle \right] \\ &\quad + \lambda \int_0^T dt \left[ \sigma_x^{(1)}(t) x^{(1)}(t) - \sigma_y^{(1)}(t) y^{(1)}(t) + \sigma_z^{(1)}(t) z^{(1)}(t) \right] \\ &\quad + (1 - \lambda) \int_0^T dt \left[ \sigma_x^{(2)}(t) x^{(2)}(t) - \sigma_y^{(2)}(t) y^{(2)}(t) + \sigma_z^{(2)}(t) z^{(2)}(t) \right] \\ &\quad - \frac{J}{2^{q-1}q} \int_0^T dt (1 - x^q(t) + y^q(t) - z^q(t)). \end{aligned} \quad (3.56)$$

We now expand around the saddle point solution  $\mathbf{x}_*^{(\alpha)}$  found in section 3.3. We will use hatted variables to denote fluctuations:

$$\mathbf{x}^{(\alpha)} = \mathbf{x}_*^{(\alpha)} + \sqrt{\hbar^{(\alpha)}} \hat{\mathbf{g}}^{(\alpha)}, \quad \boldsymbol{\sigma}^{(\alpha)} = \boldsymbol{\sigma}_*^{(\alpha)} + \sqrt{\hbar^{(\alpha)}} \hat{\boldsymbol{\sigma}}^{(\alpha)}, \quad (3.57)$$

$$\hat{\boldsymbol{\sigma}}^{(\alpha)} = (\hat{\sigma}_x^{(\alpha)}, \hat{\sigma}_y^{(\alpha)}, \hat{\sigma}_z^{(\alpha)}), \quad \hat{\mathbf{g}}^{(\alpha)} = (\hat{x}^{(\alpha)}, \hat{y}^{(k)}, \hat{z}^{(\alpha)}), \quad (3.58)$$

where  $\hbar^{(\alpha)} = \left(\frac{1}{N\lambda}, \frac{1}{N(1-\lambda)}\right)$ . The quadratic action for the fluctuations has the following form:

$$-\hat{I} = \begin{pmatrix} \hat{\sigma}^{(1)} \\ \hat{g}^{(1)} \\ \hat{\sigma}^{(2)} \\ \hat{g}^{(2)} \end{pmatrix} \cdot M \cdot \begin{pmatrix} \hat{\sigma}^{(1)} \\ \hat{g}^{(1)} \\ \hat{\sigma}^{(2)} \\ \hat{g}^{(2)} \end{pmatrix}, \quad (3.59)$$

where

$$M = \begin{pmatrix} K_1 & S & 0 & 0 \\ S & \lambda \tilde{S} & 0 & \sqrt{\lambda(1-\lambda)} \tilde{S} \\ 0 & 0 & K_2 & S \\ 0 & \sqrt{\lambda(1-\lambda)} \tilde{S} & S & (1-\lambda) \tilde{S} \end{pmatrix}. \quad (3.60)$$

The matrices  $K_1$  and  $K_2$  can be derived by variation of the Pfaffian terms at quadratic order.  $S$  and  $\tilde{S}$  are defined as

$$S = \delta(t_{12}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{S} = \frac{J(q-1)}{2^{q-2}} \delta(t_{12}) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 + O(\lambda) \end{pmatrix}. \quad (3.61)$$

To compute the determinant at leading order in  $\lambda$ , it turns out to be sufficient to note that  $K_2$  has non-zero matrix elements only in  $\hat{\sigma}_x^{(2)}$  and  $\hat{\sigma}_y^{(2)}$ . Therefore,  $K_2$  satisfies the relation

$$\tilde{S} K_2 = K_2 \tilde{S} = 0. \quad (3.62)$$

Moreover, we will only need the  $(\sigma_z^{(1)}, \sigma_z^{(1)})$  component of  $K_1$  which is

$$\iint_0^T dt_1 dt_2 \sigma_z^{(1)}(t_1) K_1^{zz}(t_1, t_2) \sigma_z^{(1)}(t_2) = \frac{1}{2} \text{sech}^2\left(\frac{JT}{2^{q-2}}\right) \left(\int_0^T dt \sigma_z^{(1)}(t)\right)^2. \quad (3.63)$$

We can now compute the determinant of  $M$  to leading order in  $\lambda$ :

$$\det M = \det A \det C \left[1 - \lambda \text{Tr} \left(A^{-1} B C^{-1} B\right)\right] + O(\lambda^2), \quad (3.64)$$

where

$$A = \begin{pmatrix} K_1 & S \\ S & \lambda \tilde{S} \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{S} \end{pmatrix}, \quad C = \begin{pmatrix} K_2 & S \\ S & (1-\lambda) \tilde{S} \end{pmatrix}. \quad (3.65)$$

We first compute the trace term in the determinant

$$\begin{aligned} \text{Tr} \left(A^{-1} B C^{-1} B\right) &= \text{Tr} \left[A_{22}^{-1} \tilde{S} C_{22}^{-1} \tilde{S}\right] \\ &= \text{Tr} \left[SK_1 S \tilde{S} SK_2 (S - \tilde{S} K_2)^{-1} \tilde{S}\right] + O(\lambda) \\ &\approx \text{Tr} \left[SK_1 \tilde{S} K_2 \tilde{S}\right] \\ &= 0. \end{aligned} \quad (3.66)$$

In the first step, we used the fact that  $B_{22}$  is the only non-zero entry in  $B$ . In the second step, we inserted the  $A_{22}^{-1}$  and  $C_{22}^{-1}$  components upto  $O(\lambda)$  corrections. The third and fourth step follow from the relation (3.62). Using relation (3.62) once again, we conclude that  $\det C = 1$ . We are left with the evaluation of  $\det A$ .

$$\begin{aligned}\det A &= \det A|_{\lambda=0} \left[ 1 - \lambda \text{Tr} \left( K_1 \tilde{S} \right) \right] \\ &= 1 - \lambda \frac{J(q-1)}{2^{q-2}} \iint_0^T dt_1 dt_2 \delta(t_{12}) K^{zz}(t_1, t_2) \\ &= 1 - \lambda \frac{JT(q-1)}{2^{q-1}} \text{sech}^2 \left( \frac{JT}{2^{q-1}} \right).\end{aligned}\tag{3.67}$$

Thus,

$$\det M = 1 - \lambda \frac{JT(q-1)}{2^{q-1}} \text{sech}^2 \left( \frac{JT}{2^{q-1}} \right) + O(\lambda^2).\tag{3.68}$$

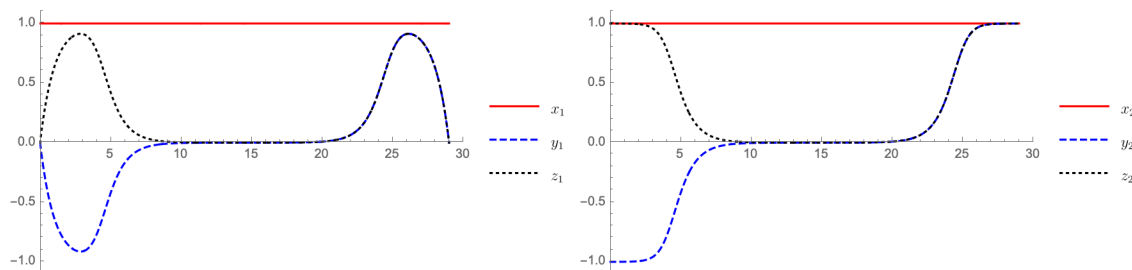
So, the one-loop determinant does not significantly modify the  $T$  dependence of  $\text{Tr} \sigma_L^2$  at leading order. The coefficient of the  $O(\lambda)$  term above is bounded by an  $O(q)$  number, and  $q\lambda$  is always small in our regime of interest.

### 3.4 Connected solution

When  $\lambda$  is slightly non-zero,  $\mathbf{x}_2 \approx (0, 0, 1)$  is not the only fixed point for the  $\mathbf{x}_2$  variables which enters the analysis. The leading solution involving more than one fixed point is the tunneling solution between the  $\mathbf{x}_2 \approx (1, -1, 1)$  and  $\mathbf{x}_2 \approx (1, 1, 1)$  fixed points. This solution has nontrivial time-dependence for the  $\mathbf{x}_2$  variables which involves an initial region where  $\mathbf{x}_2(t) \approx (1, -1, 1)$ , then a transition to the  $\mathbf{x}_2(t) \approx (1, 0, 0)$  fixed point where the solution remains for a long period, and then a final transition to the  $\mathbf{x}_2(t) \approx (1, 1, 1)$  fixed point. The  $\mathbf{x}_1$  solution has large transient behaviors in the initial and final fixed point regions, but matches very closely with  $\mathbf{x}_2$  in the long region where  $\mathbf{x}_2 \approx \mathbf{x}_1 \approx (1, 0, 0)$ .

Because this solution is non-perturbative in  $\lambda$ , we cannot hope to use perturbation theory to evaluate the effective action. We will instead follow the approximate analysis of [41]. Unlike the disconnected solution we described in section 3.3, the connected solution is suppressed exponentially in  $N$ , and the main aim of our approximate analysis will be to demonstrate this suppression quantitatively. The numerical connected solution is shown in figure 4. We will give an approximate analytical computation of the action for this solution. As argued in section 3.3, for any solution of the equations of motion the bulk terms in the effective action contribute to the total path integral  $\exp(-JT/2^{q-1})$  for  $r \approx 1$ . So what remains is to evaluate the two Pfaffian contributions, again using the qubit Hamiltonian approach.

We see in figure 4 that there is a long region with  $\mathbf{x}_1 \approx \mathbf{x}_2 \approx (1, 0, 0)$ . In this region, we may approximate the time-ordered exponential expressions as projectors  $|+\rangle\langle+|$ , the lowest energy state of the Hamiltonian  $-JT X/2^{q-1}$  generating the time evolution in that



**Figure 4.** The connected solution for the  $\mathbf{x}_1$  (left) and  $\mathbf{x}_2$  (right) variables with  $J = q = 4$ ,  $r = 1 - 10^{-8}$ , and  $\lambda = 10^{-4}$ . There is a long region where  $\mathbf{x}_1 \approx \mathbf{x}_2$  around the fixed point  $(1, 0, 0)$ , while there are large transient behaviors where  $\mathbf{x}_1 \neq \mathbf{x}_2$  near  $t = 0$  and  $t = T \approx 29$  due to the difference between the  $\mathbf{x}_1$  and  $\mathbf{x}_2$  boundary conditions. The regions where  $\mathbf{x}_1$  displays transient behavior are comparable in size to a transition region where the  $\mathbf{x}_2$  variables tunnel between fixed points.

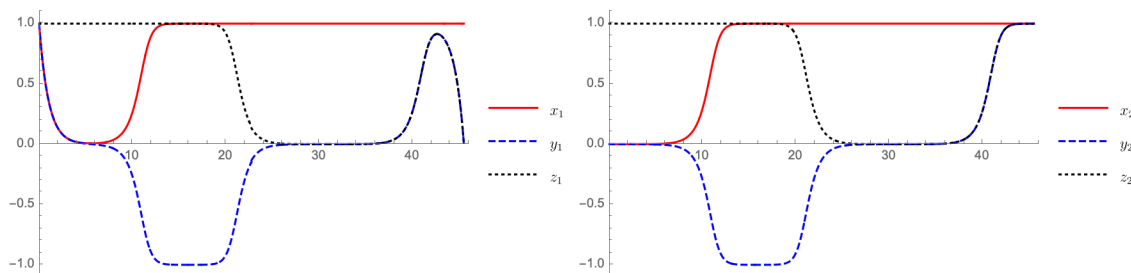
region.<sup>13</sup> The energy contribution from this ground state exactly cancels the bulk term, so the result of the long region for both  $\mathbf{x}_1$  and  $\mathbf{x}_2$  is a projector  $|+\rangle\langle+|$ . At this point, the analysis splits between  $\mathbf{x}_1$  and  $\mathbf{x}_2$ .

The  $\mathbf{x}_2$  variables include an initial region around the  $(1, -1, 1)$  fixed point and a final region around  $(1, 1, 1)$ . Both of these regions share an important property with their adjacent transition regions, namely that the first has  $y_2 \approx -z_2$  and the second has  $y_2 \approx z_2$ . Because these regions are adjacent to the long middle region that yields a projector  $|+\rangle\langle+|$ , we may use the null state relations  $\langle+|(iY + Z) = 0$  and  $(iY - Z)|+\rangle = 0$  to conclude that the Pfaffian does not depend on the precise details of these transition regions nor on the initial and final fixed point regions, and the remaining  $X$  term in the Hamiltonian simply cancels against the bulk contribution as was the case in the long middle fixed point region. So, the overall Pfaffian for the  $\mathbf{x}_2$  variables is determined by the overlap of the initial state  $|0\rangle$  and final state  $\langle 0|$  with the projector from the middle region:  $2\langle 0|+\rangle\langle+|0\rangle = 1$ .

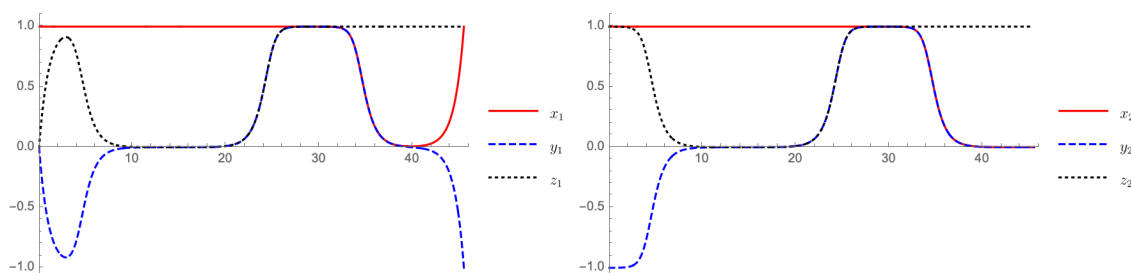
We may analyze the  $\mathbf{x}_1$  variables similarly. Because the long region where  $\mathbf{x}_1 \approx \mathbf{x}_2$  again gives a projector  $|+\rangle\langle+|$ , and because the transient regions satisfy the same relations between the variables as the transition regions from the  $\mathbf{x}_2$  analysis, the same arguments we made about the transition regions for  $\mathbf{x}_2$  goes through for the transient behaviors of  $\mathbf{x}_1$ , and the Pfaffian does not depend on the precise form of the transient behaviors. The remaining Hamiltonian contribution from  $X$  again cancels the bulk term in the transient regions. There are no initial or final fixed point regions for  $\mathbf{x}_1$ , so the total contribution is from another projector overlap with the relevant initial and final states:  $2\langle+|+\rangle\langle+|+\rangle = 2$ . It may seem redundant to analyze the  $\mathbf{x}_1$  variables separately as we have done here, since the  $A_1$

<sup>13</sup>When  $T$  is on the order of  $(1/J)\log N$  and not much larger, there are exponentially suppressed  $T$  dependent corrections to this projector which lead to  $O(1)$  factors in the wormhole contribution to the Rényi mutual information. While these corrections could be addressed in the path integral formalism we employ here, it is easier to study them in the Hamiltonian picture (appendix A). We will continue to approximate the long region as a projector because these corrections are highly subleading by the time the wormhole dominates at  $T \sim N/J$  and are therefore unimportant for the qualitative error correction properties of the Brownian circuit.





**Figure 5.** The DW solution for the  $\mathbf{x}_1$  (left) and  $\mathbf{x}_2$  (right) variables with  $J = q = 4$ ,  $r = 1 - 10^{-8}$ ,  $\lambda = 10^{-4}$ , and  $T \approx 45.5$ . There are two long regions with  $\mathbf{x}_2 \approx (0, 0, 1)$  and  $(1, 0, 0)$ , although the  $(0, 0, 1)$  region is a little smaller for this value of  $\lambda$ . The initial transient behavior for  $\mathbf{x}_1$  matches the disconnected solution in figure 3 while the final transient matches the connected solution in figure 4.



**Figure 6.** The WD solution for the  $\mathbf{x}_1$  (left) and  $\mathbf{x}_2$  (right) variables with  $J = q = 4$ ,  $r = 1 - 10^{-8}$ ,  $\lambda = 10^{-4}$ , and  $T \approx 45.5$ . There are two long regions with  $\mathbf{x}_2 \approx (1, 0, 0)$  and  $(0, 0, 1)$ , though the  $(0, 0, 1)$  region is a little smaller for this value of  $\lambda$ . The initial transient behavior for  $\mathbf{x}_1$  matches the connected solution in figure 4 while the final transient matches the disconnected solution in figure 3.

Pfaffian term involves the total  $\mathbf{x}$  variables like the  $A_2$  Pfaffian. However, it was important here to conclude that the transient behaviors do not contribute any time-dependence at  $O(\lambda)$ , and we actually obtained a constant result that is independent of  $T$ .

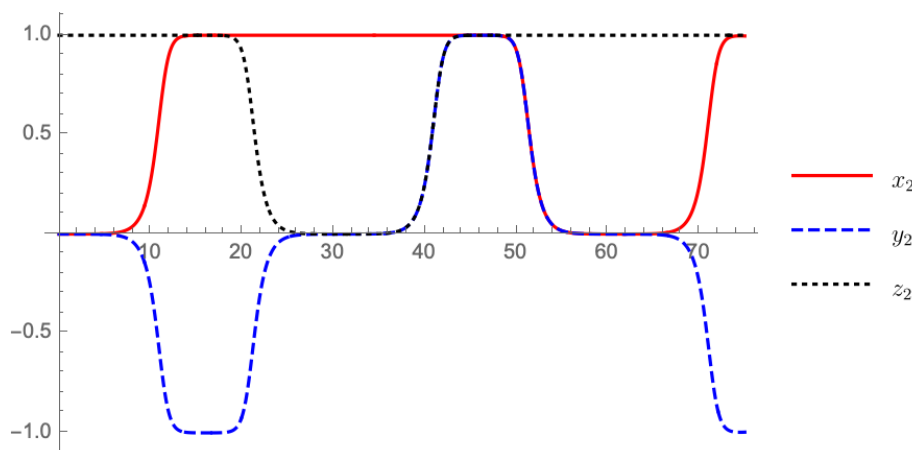
Thus, again using the normalization (3.36), from the connected solution we have a contribution

$$\text{Tr } \sigma_L^2 \Big|_{\text{conn}} = 2^{-N(1-\lambda)}. \quad (3.69)$$

### 3.5 Other tunneling solutions

There are also solutions which tunnel from  $\mathbf{x}_2 \approx (0, 0, 1)$  to  $\mathbf{x}_2 \approx (1, 1, 1)$  and from  $\mathbf{x}_2 \approx (1, -1, 1)$  to  $(0, 0, 1)$ . We name these the “DW” and “WD” solutions, respectively, after the order of transient behavior which occurs for the  $\mathbf{x}_1$  variable: the first has “Disk” initial transient behavior and “Wormhole” final transient behavior, while the second has the opposite ordering. The DW solution is shown in figure 5 while the WD solution is shown in figure 6. The contribution of these solutions to  $\text{Tr } \sigma_L^2$  can be evaluated in the same approximate manner as section 3.4.

We begin with the DW solution in figure 5. The  $\mathbf{x}_2$  variables has the same long region with  $\mathbf{x}_2 \approx (1, 0, 0)$  which appears in the connected solution (figure 4), and by the same reasoning as in section 3.4 we conclude that this region yields for the path integral the



**Figure 7.** A full period of the  $\mathbf{x}_2$  variables with  $J = q = 4$ ,  $r = 1 - 10^{-8}$ ,  $\lambda = 10^{-4}$ . This periodic segment can be inserted into the disconnected, connected, DW, or WD solutions however many times we like to produce new solutions (for different values of  $T$ ) that are suppressed with powers of  $2^{-kN}$ , where  $k$  is the number of inserted periods.

projector  $|+\rangle\langle+|$ . Similarly, the long region with  $\mathbf{x}_2 \approx (0, 0, 1)$  gives a projector  $|0\rangle\langle 0|$ . These two projectors cancel the transition regions and the other constant regions associated with other fixed points, and we get the overlap  $2\langle 0|0\rangle\langle 0|+\rangle\langle +|0\rangle = 1$ . The  $\mathbf{x}_1$  variables have a leading contribution determined by simply changing the initial and final states:  $2\langle +|0\rangle\langle 0|+\rangle\langle +|+\rangle = 1$ . Thus, including the normalization (3.36), we have the additional suppression  $2^{-N}$  for the DW solution.<sup>14</sup>

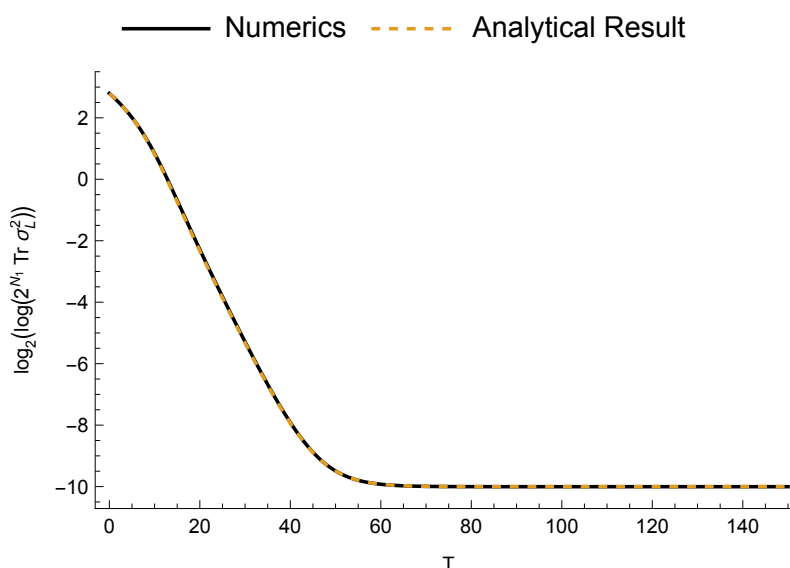
$$\text{Tr } \sigma_L^2 \Big|_{\text{DW}} = 2^{-N}. \quad (3.70)$$

We will not bother to compute the  $O(\lambda)$  contribution from the transient behaviors of  $\mathbf{x}_1$  in the  $A_1$  Pfaffian (which could lead to nontrivial  $T$  dependence), since this solution is already highly suppressed compared to the connected one in section 3.4.

The WD solution can be analyzed similarly and also has a  $2^{-N}$  leading suppression. Thus, both the DW and WD solutions are subleading compared to the connected solution from section 3.4. There are also even more highly suppressed solutions which can be formed by inserting additional periods into any of the four solutions we have discussed up to this point. A full period of the  $\mathbf{x}_2$  variables is shown in figure 7. By the same approximate reasoning, inserting a full period in the solution will suppress the contribution to the path integral by an additional  $2^{-N}$ .

Interestingly, the long regions of the solution have a minimum length which scales like the scrambling time  $T_s \sim (1/J) \log N$ . What this means is that they are actually not solutions for all values of  $T$ . For instance, the connected solution in section 3.4 is only a solution for  $T > (1/J) \log N$ . A configuration with  $k$  long regions will not appear

<sup>14</sup>We are neglecting the one-loop determinant here. As shown in [41], this determinant can lead to an overall minus sign for some of these subleading solutions. Because they are subleading anyway, we will omit this effect, which does not affect the disconnected or connected saddle points.



**Figure 8.** A comparison between numerical evaluation of  $\log_2[\log(2^{N_1} \text{Tr} \sigma_L^2)]$  with the approximate analytical result of equation (3.71) for  $N = 30$  and  $\lambda = 1/3$ . At late times, the quantity saturates to  $-(N_2 - N_1) = -(1 - 2\lambda)N = -10$ .

as a solution until  $T > (k/J) \log N$ . This lattice of critical times is interesting from a unitarity perspective. These subleading saddles are necessary to ensure the total Brownian evolution is unitary, so an experimentalist with access to only on-shell configurations will discover that it is impossible to verify unitarity with accuracy better than  $2^{-kN}$  until at least  $T > (k/J) \log N$ .

### 3.6 Summary

We have shown that the leading  $T$  dependence of the purity which controls the mutual purity in Brownian SYK is<sup>15</sup>

$$\text{Tr} \sigma_L^2 = \left( \frac{1 + e^{-JT/2^{q-1}}}{2} \right)^{\lambda N} + \frac{\Theta(JT - \log N)}{2^{(1-\lambda)N}} + \dots, \quad (3.71)$$

where the first term comes from the disconnected saddle point, the second term from the leading connected saddle point, and the dots represent further subleading solutions that are suppressed in powers of  $2^{-N}$ . We present a comparison of this saddle point analysis with an exact numerical computation of  $\text{Tr} \sigma_L^2$  in figure 8. The form of (3.71) means  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$  is initially  $O(1)$  and subsequently decays for a polynomial  $T \sim N/J$  amount

<sup>15</sup>In this analysis, we have purposefully ignored the presence of discrete symmetries. The presence of such symmetries generically prevents the time evolution from covering the entire unitary group. Following [41], we can adapt the analysis of Brownian SYK so that the time evolution covers the entire unitary group by only including the saddle points we have discussed. Incorporating the discrete symmetries of the SYK model requires additional saddle points [41]. This means our results are effectively valid for an SYK-like model with no discrete symmetries which does end up covering the whole unitary group.

of time. When  $T > N/J$ , the connected solution begins to dominate and leads to an exponentially small mutual purity.<sup>16</sup> Thus, when the encoding complexity is sufficiently large, the code is robustly protected from the erasure of  $A_1$ . At multiples of the scrambling time  $T_s \sim (1/J) \log N$ , subleading contributions become genuine on-shell solutions of the equations of motion, though these contributions never dominate  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$ .

## 4 Discussion

We have studied the error correction properties of Brownian SYK quantum codes against the erasure of a small number of qubits, but we expect our results to be valid more generally for generic, low-rank errors with no prior access to the encoding map. As a measure of quantum error correction, we computed the mutual purity  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$ , which is related to the purity  $\text{Tr } \sigma_L^2$ , where  $\sigma = VV^\dagger/d_{\text{code}}$  is the density matrix built from the encoding map  $V$ ,  $\sigma_L = \text{Tr}_R \sigma$ , and  $R$  is a small fraction of the physical Hilbert space which is being erased. In codes defined using Brownian SYK time evolution, which have a linearly growing encoding complexity — mimicking the expected behavior of the bulk-to-boundary map for an infalling observer in AdS/CFT — this purity is related to a four-contour Lorentzian (Schwinger-Keldysh) path integral. We found two special saddle point solutions to the large  $N$  equations of motion in Brownian SYK — analogous to the disconnected disks and connected wormhole geometries in JT gravity — which dominate this path integral. At early times  $T \ll N/J$ , the disconnected solution gives an exponentially decaying value for the mutual purity, while at late times the connected solution dominates and gives a constant, exponentially small mutual purity. Thus, when the encoding complexity is sufficiently large, we find emergent, “complexity-protected” quantum error correction against generic, low-rank errors with no prior access to the encoding map. We should emphasize that it is important that the error does not have access to the encoding map — with prior access, it is possible to violate the above conclusions.

### 4.1 Relation to previous work

Understanding how the complexity of an encoding operator affects certain error correction properties of the code is a problem that has been explored previously from a variety of viewpoints. The most common method of studying codes with increasing complexity is to employ the randomization trick as we have done, where one instead considers a one-parameter family of ensembles of codes with increasing complexity and studies ensemble-averaged properties.

For instance, [60, 61] argued that  $n$ -qubit random quantum circuits with  $O(n \log^2 n)$  two-qubit gates and  $O(\log^3 n)$  depth can encode  $k$  qubits into  $n$  while correcting erasure errors on  $d$  qubits where

$$\frac{k}{n} < 1 - \frac{d}{n} \log_2 3 - h(d/n), \quad (4.1)$$

---

<sup>16</sup>Recall that  $\text{Tr } \sigma_L^2$  enters in  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$  along with a subtraction of a baseline value, and so the contribution which dominates  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$  is not necessarily the one which makes the largest contribution to  $\text{Tr } \sigma_L^2$ .

with  $h(x)$  being the binary entropy function

$$h(x) \equiv -x \log x - (1-x) \log(1-x). \quad (4.2)$$

In our analysis, we studied a random error with  $k/n = d/n = \lambda$ , and with these replacements the inequality (4.1) is true for roughly  $\lambda \leq 1/5$ . It would be interesting to understand whether our analytics can be extended to this rather large value of  $\lambda$  without the need for novel techniques, although the numerical results in figure 8 suggest we may have an accurate picture of the Brownian theory even when  $\lambda = 1/3$ . At any rate, it appears that the Brownian codes we have studied in this work are able to *approximately* (with error of order  $1/n$ ) correct errors on a fraction  $\lambda$  of the physical qubits with a depth  $T \sim (1/J) \log n$ . This polynomial improvement in depth, if true, is likely due to differences in how the random two-qubit quantum circuit theory of [61] and the Brownian SYK theory scramble quantum information.

More recently, [62] studied low depth random circuits with spatial connectivity restrictions in various spatial dimensions  $D$  as stabilizer codes. They discovered that such circuits can correct fairly large erasure errors (converging to both the optimal threshold and zero failure probability at large  $n$ ) with a depth of just  $O(\log n)$  for  $D \geq 2$ . These results are similar to ours, although we have no restriction on spatial connectivity, but rather a restriction on the number of fermions which can couple in the Hamiltonian. It would be interesting to understand if there is a relation between the universality for  $D \geq 2$  found in [62] and the expected universality of our results for  $q \geq 4$ . A significant difference of our analysis compared with [62] is that we do not restrict ourselves to stabilizer codes, though we also have not studied the decoding problem in any detail.

Beyond questions of depth, we may also consider the total gate complexity of efficient quantum codes. Several bounds on this complexity exist for stabilizer codes [63–65] and their generalizations [65, 66]. In particular, for a generic stabilizer code encoding  $k$  qubits into  $n$ , [65] showed that  $O(n(n-k)/\log n)$  gates are sufficient. Entanglement-assisted stabilizer codes were also studied in [65] and were shown to have gate complexity linear in the number of additional entangled qubits  $c$ , with  $O(n(n-k+c)/\log n)$  gates. As we have not restricted ourselves to stabilizer circuits, our gate complexity is not expected to have such small polynomial asymptotic behavior.<sup>17</sup> However, if we used a sparse SYK model instead [69], we may achieve equal or better gate complexity compared to stabilizer circuits. This issue deserves further study as it would represent an interesting development in efficient random code design.

Our work is also closely related to measurement-induced phase transitions which have recently been studied extensively in the condensed matter community (see for instance [70, 71]). In these studies, the quantum circuit usually consists of local unitary gates with some quenched disorder and forms a brickwall pattern. These local unitaries are interspersed with local measurements, which are viewed as “errors”. The long range entanglement generated

---

<sup>17</sup>Efficient Hamiltonian simulation of Brownian SYK would likely involve discretization of the contact correlation  $\delta(t-t')$  in the variance, along with a sparse query model like the one studied in [67, 68]. Because the sparsity of the full SYK Hamiltonian scales with  $N^q$ , we do not expect simulation to be efficient compared to stabilizer circuits.

by the random unitary gates is identified as the volume-law phase, suitable for quantum error correction. However, a transition to a short-range entanglement phase can occur when measurement rate is high, i.e. when the error rate is high enough to disentangle different subsystems. The volume-law to area-law transition is identified as a transition in quantum error correction, when the error rate exceeds a critical value [70, 71]. In essence, the size of the Hilbert space of the principal quantum system needs to be large enough (spatial) and the time for the unitary gates need be long enough (temporal) to scramble the information so that the entanglement is robust against local disturbances. It would be interesting to compare these results with those presented here.

In another direction, ensembles of encoding maps that satisfy some global symmetry have also been explored [72, 73]. The general idea is that there is a tension between the existence of a continuous symmetry leaving the encoding map invariant and strong protection against erasure errors. However, approximate error correction can be achieved in certain circumstances [74]. In Brownian SYK, there are discrete global symmetries (which we did not include in the analysis since we were interested in covering the entire unitary group) but no continuous symmetries, allowing us to avoid these no-go arguments. However, it is easy to implement continuous symmetries in analogues of the SYK model; for instance, SYK with complex fermions satisfies a  $U(1)$  global symmetry [75]. It would be interesting to understand the error correction behavior of a complex analogue of Brownian SYK to further elucidate the tension between codes with continuous symmetries and erasure error correction.

## 4.2 Pseudorandom codes

Our results seem to suggest that after a polynomial time, a random quantum circuit, which likely has polynomial circuit complexity, has powerful error correction properties that are essentially as good as a Haar random unitary code, which likely has exponential complexity. One explanation for why this is possible may be that the majority of unitary operators with polynomial complexity are in fact pseudorandom unitary operators, and a simple test of error correction properties cannot distinguish polynomially complex pseudorandom unitary operators from unitary operators of exponential complexity.

A pseudorandom unitary operator is, roughly speaking, an operator which has polynomial complexity but which cannot be distinguished from one with exponential complexity by any sort of simple test which can be implemented efficiently. The transition between disconnected and connected solutions that we found, hints at a sharp transition point where most random circuits with complexity less than some polynomial ( $N$ , for the purity transition) are not pseudorandom, while the typical circuit and perhaps the majority of circuits above that critical complexity are in fact pseudorandom, at least for the purposes of error correction. It would be very interesting to understand in more detail what properties of Haar random circuits can be reproduced by such low complexity Brownian circuits.

## 4.3 Complexity and the geometry of the entanglement wedge

From the AdS/CFT point of view, it would be very interesting to understand the bulk significance of our results; indeed, one of our main motivations in this paper was to

understand the geometry of figure 1 in terms of quantum error correction. Following [30, 31], we expect that this error correction is a sign of “causal inaccessibility” from the boundary subregion. By this, we mean that including backreaction from turning on simple sources in the asymptotic boundary does not render the relevant degrees of freedom causally accessible from the boundary; the mechanism behind this is that the relevant bulk degrees of freedom lie behind a non-minimal quantum extremal surface. In our calculation, we encountered two significant complexity scales, i.e., the mutual purity becomes  $O(1/N)$  at  $T \sim \log N$ , and the mutual purity saturates to an exponentially small plateau at a much larger time-scale. It is tempting to speculate that these thresholds have natural bulk interpretations: the  $\log N$  time-scale could correspond to the bulk degrees of freedom crossing the causal horizon, while the plateau could correspond to the bulk degrees of freedom crossing over to the python’s lunch. In a similar vein, the lattice of subleading solutions we found may also have a geometric meaning, although it is less clear because they do not dominate the calculation of the crucial quantity  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$ .

## Acknowledgments

We thank Bartek Czech, Abhijit Gadde, Issac Kim, Gautam Mandal, Shiraz Minwalla, Pranab Sen and Sandip Trivedi for helpful discussions. VB is supported in part by the Department of Energy through grant DE-SC0013528 and grant QuantISED DE-SC0020360, as well as the Simons Foundation through the It From Qubit Collaboration (Grant No. 38559). AK is supported by the Simons Foundation through the It from Qubit Collaboration. CL is supported by the Department of Energy through QuantISED grant DE-SC0020360. OP and HR are supported by the Department of Atomic Energy, Government of India, under project identification number RTI 4002.

## A Hamiltonian formalism in Brownian SYK

In this appendix, we will reproduce some of the results of section 3 from a different point of view. Recall that after averaging over the couplings  $J_{a_1 \dots a_q}(t)$  in the Brownian SYK model, we obtain the following effective action:

$$I = \frac{1}{2} \int_0^T dt \psi_a^{(j)} \partial_t \psi_a^{(j)} - \frac{i^q J(q-1)!}{2N^{q-1}} \int_0^T dt s_j s_{j'} \psi_{a_1 \dots a_q}^{(j)}(t) \psi_{a_1 \dots a_q}^{(j')}(t). \quad (\text{A.1})$$

From the action, we can read off an “effective Hamiltonian”:

$$H_{\text{eff}} = -i^q \frac{J(q-1)!}{2N^{q-1}} \sum_{j,k} \sum_{a_1 < a_2 < \dots < a_q} s_j s_k \psi_{a_1 \dots a_q}^{(j)} \psi_{a_1 \dots a_q}^{(k)}, \quad (\text{A.2})$$

where,  $\psi_{a_1 \dots a_q}^{(j)} = \prod_{i=1}^q \psi_{a_i}^{(j)}$ . We note that  $H_{\text{eff}}$  commutes with the fermion parity operator defined in equation (3.8) and (3.9). Therefore, we can write  $H_{\text{eff}}$  in terms of the Pauli matrices defined as follows:

$$\psi_a^{(1)} \psi_a^{(2)} = -\frac{i}{2} X_a, \quad \psi_a^{(1)} \psi_a^{(3)} = \frac{i}{2} Y_a, \quad \psi_a^{(1)} \psi_a^{(4)} = -\frac{i}{2} Z_a. \quad (\text{A.3})$$



After substituting the above relations in (A.2), the effective Hamiltonian can be written as

$$H_{\text{eff}} = -\frac{J(q-1)!}{(2N)^{q-1}} \sum_{a_1 < a_2 \dots < a_q} (X_{a_1} \dots X_{a_q} - i^q Y_{a_1} \dots Y_{a_q} + Z_{a_1} \dots Z_{a_q} - 1), \quad (\text{A.4})$$

where the last term (proportional to the identity) comes from the  $j = k$  terms.  $\text{Tr } \sigma_L^2$  can now be written as the Euclidean transition amplitude:

$$\text{Tr } \sigma_L^2 = \langle \psi | \exp(-H_{\text{eff}} T) | \psi \rangle, \quad (\text{A.5})$$

where the initial and final states are dictated by the boundary conditions in the path integral, and are given by

$$|\psi\rangle = |+\rangle^{N_1} |0\rangle^{N_2}. \quad (\text{A.6})$$

### A.1 Disconnected and connected solutions

In the above expression, we can separate contributions from the ground states and the excited states as follows:

$$\text{Tr } \sigma_L^2 = \sum_n |\langle \psi | g_n \rangle|^2 + \sum_k \exp(-E_k T) |\langle \psi | e_k \rangle|^2. \quad (\text{A.7})$$

Here,  $g_n$  are all the ground states of  $H_{\text{eff}}$  (which all have zero energy) and  $e_k$  are the excited states with energies  $E_k$ . We first look at the contribution from the ground states. The set of ground states depends on whether we choose  $q = 4k$  or  $q = 4k + 2$  but the following two ground states contribute to the leading order independently of  $q$ :

$$\begin{aligned} |g_1\rangle &= |0\rangle^N, \\ |g_2\rangle &= |+\rangle^N. \end{aligned} \quad (\text{A.8})$$

Therefore, the contribution from the ground states is

$$\begin{aligned} \sum_n |\langle \psi | g_n \rangle|^2 &\approx \langle \psi | g_1 \rangle|^2 + \langle \psi | g_2 \rangle|^2 \\ &= \frac{1}{2^{N_1}} + \frac{1}{2^{N_2}}, \end{aligned} \quad (\text{A.9})$$

which reproduces the two leading order terms (i.e., the disconnected and the connected contributions) in the Haar ensemble. The contribution from excitations near the ground states can be approximated in the following manner. The Hamiltonian can be written in terms of the ladder operator as

$$H = -\frac{2JN}{q} \left[ \left( \frac{S_x}{N} \right)^q - \left( \frac{iS_y}{N} \right)^q + \left( \frac{S_z}{N} \right)^q - \frac{1}{2^q} \right] + O\left( \frac{1}{N} \right), \quad (\text{A.10})$$

where  $S_x = \sum_a X_a/2$ ,  $S_y = \sum_a Y_a/2$ , and  $S_z = \sum_a Z_a/2$ . The matrix elements of the first two terms in the Hamiltonian with excited states near  $|g_1\rangle$  are suppressed by a factor of  $1/N^{q/2}$  and can be ignored for  $q \geq 4$  at leading order in  $1/N$ . Thus, the Hamiltonian up to  $O(1/N)$  corrections is

$$H = -\frac{2JN}{q} \left[ \left( \frac{S_z}{N} \right)^q - \frac{1}{2^q} \right]. \quad (\text{A.11})$$

The contribution from the states near  $|g_1\rangle$  is:

$$\begin{aligned}\mathrm{Tr}_{g_1} \sigma_L^2 &= \sum_k \exp(-E_k T) |\langle \psi | e_k^{g_1} \rangle|^2 \\ &\approx \sum_{k=0}^{N_1} \frac{1}{2^{N_1}} \binom{N_1}{k} \exp\left(-\frac{JkT}{2^{q-2}}\right) \\ &= \left(\frac{1 + \exp\left(-\frac{JT}{2^{q-2}}\right)}{2}\right)^{N_1},\end{aligned}\tag{A.12}$$

where  $|e_k^{g_1}\rangle$  denotes the  $k^{th}$  excited state near  $|g_1\rangle$ , explicitly given by a choice of  $k$  qubits which are flipped to  $|1\rangle$  from  $|0\rangle$ . These  $k$  must come from the first  $N_1$  qubits to give a nonvanishing overlap  $\langle \psi | e_k^{g_1} \rangle$ .

Similarly, we can compute the correction due to the excited states  $|e_k^{g_2}\rangle$  near  $|g_2\rangle$ . The perturbative Hamiltonian is now

$$H = -\frac{2JN}{q} \left[ \left(\frac{S_x}{N}\right)^q - \frac{1}{2^q} \right],\tag{A.13}$$

and the excited states are formed by flipping  $k$  qubits to  $|-\rangle$  from  $|+\rangle$ , where these must come from the last  $N_2$  qubits to give a nonvanishing overlap with the boundary state. We get

$$\begin{aligned}\mathrm{Tr}_{g_2} \sigma_L^2 &= \sum_k \exp(-E_k T) |\langle \psi | e_k^{g_2} \rangle|^2 \\ &\approx \frac{1}{2^{N_2}} \left( 1 + N_2 \exp\left(-\frac{JT}{2^{q-2}}\right) + \dots \right).\end{aligned}\tag{A.14}$$

Note that the contribution from the second and higher excited states (denoted here by ellipsis) is not negligible. Moreover, unlike the case of the disconnected saddle where  $N_1/N = \lambda$  was a small parameter, we cannot resum all the contributions from higher excited states near  $|g_2\rangle$ . Since  $N_2/N \sim 1$ , one must also take the quantum corrections into account. Nevertheless, the above expression is sufficient to infer that the ground state contribution dominates when  $T > \frac{2^{q-2}}{J} \log(N_2)$ .

Summing up these contributions we have the following result for  $\mathrm{Tr} \sigma_L^2$ :

$$\begin{aligned}\mathrm{Tr} \sigma_L^2 &\approx \mathrm{Tr}_{g_1} \sigma_L^2 + \mathrm{Tr}_{g_2} \sigma_L^2 \\ &\approx \left(\frac{1 + \exp\left(-\frac{JT}{2^{q-2}}\right)}{2}\right)^{N_1} + \frac{1}{2^{N_2}} \left( 1 + N_2 \exp\left(-\frac{JT}{2^{q-2}}\right) + \dots \right).\end{aligned}\tag{A.15}$$

The  $T$  dependent term proportional to  $N_2/2^{N_2}$  is a contribution from corrections to the projector approximation to the long region we made in section 3.4. We could have incorporated such terms in the path integral saddle point approximation of section 3.4 by writing the long region as a projector  $|+\rangle\langle+|$  plus an exponentially suppressed correction  $e^{-JT}|-\rangle\langle-|$ . However, the evaluation of the saddle point including this correction is difficult because the transient regions no longer cancel against the  $|-\rangle\langle-|$  operator, so this term

induces large corrections which depend sensitively on the transient shape. This is the path integral analogue of the Hamiltonian picture difficulty we described under (A.14). Of course, these corrections are only important before  $T < (1/J) \log N$ , when the connected configuration is not actually a solution of the equations. By the time the connected configuration becomes a genuine saddle point, this  $T$  dependence is subleading and the constant  $2^{-N_2}$  term dominates up to possible  $O(1)$  factors just as  $T$  crosses  $(1/J) \log N$ . Furthermore, by the time the connected solution actually dominates the mutual purity, these terms are suppressed by an even stronger factor of  $e^{-N}$  compared to the constant  $2^{-N_2}$  term.

## A.2 One-loop determinant around disconnected solution

In the Hamiltonian picture, the one-loop determinant is related to corrections in the energy eigenstates and the corresponding eigenvalues near the ground state  $|g_1\rangle$ . From equation (A.10), we see that the energy eigenstates  $|e_k^{g_1}\rangle$  gain corrections from the first two terms related to the ladder operators. However, since they are suppressed by a factor of  $1/\sqrt{N^q}$  we can ignore these corrections. The correction to energy eigenvalues can be computed by expanding the  $S_z^q$  term to  $O(1/N^2)$ :

$$E_k = \frac{J}{2^{q-2}} \left( k - (q-1) \frac{k^2}{N} \right) + O(N^{-2}). \quad (\text{A.16})$$

Thus, the contribution to  $\text{Tr} \sigma_L^2$  from the first saddle including corrections at  $O(1/N)$  is

$$\text{Tr}_{g_1}^{(1)} \sigma_L^2 = \sum_{k=1}^{N_1} \binom{N_1}{k} \exp \left( -\frac{JT N_1}{2^{q-2}} \left( \frac{k}{N_1} - (q-1) \lambda \frac{k^2}{N_1^2} \right) \right). \quad (\text{A.17})$$

To extract the one-loop determinant from the above expression we divide it by the classical saddle point result in equation (A.12) and take the large  $N_1$  limit keeping  $\lambda = \frac{N_1}{N}$  fixed. Define  $F(T)$  as

$$\begin{aligned} F(T) &\equiv \frac{\text{Tr}_{g_1}^{(1)} \sigma_L^2}{\text{Tr}_{g_1}^{(0)} \sigma_L^2} \\ &= \frac{\sum_{k=1}^{N_1} \binom{N_1}{k} \exp \left( -\frac{JT N_1}{2^{q-2}} \left( \frac{k}{N_1} - (q-1) \lambda \frac{k^2}{N_1^2} \right) \right)}{\sum_{m=1}^{N_1} \binom{N_1}{m} \exp \left( -\frac{J N_1 T}{2^{q-2}} \frac{m}{N_1} \right)}. \end{aligned} \quad (\text{A.18})$$

In the large  $N_1$  limit, we use Stirling's approximation for the factorial terms and replace the sum over  $k$  by an integral over  $x \equiv k/N_1$  to write  $F(T)$  as follows:

$$F(T) \approx \frac{\int_0^1 dx \frac{1}{\sqrt{x(1-x)}} \exp[-N_1 f(x)]}{\int_0^1 dx \frac{1}{\sqrt{x(1-x)}} \exp[-N_1 g(x)]}, \quad (\text{A.19})$$

where we have defined the functions

$$\begin{aligned} g(x) &= x \log x + (1-x) \log(1-x) + \frac{JT}{2^{q-2}} x, \\ f(x) &= g(x) - \lambda \frac{JT(q-1)}{2^{q-2}} x^2. \end{aligned} \quad (\text{A.20})$$

The integrals can be evaluated in the saddle point approximation and we get the following result:

$$F(T) \approx \sqrt{\frac{g''(x_g)x_g(1-x_g)}{f''(x_f)x_f(1-x_f)}} \exp[-N_1(f(x_f) - g(x_g))]. \quad (\text{A.21})$$

Here,  $x_f$  and  $x_g$  are saddle points of  $f(x)$  and  $g(x)$  respectively. Since  $f(x)$  and  $g(x)$  differ by a term proportional to  $\lambda$ , we can evaluate  $F(T)$  perturbatively in  $\lambda$ . We have the following equations:

$$\begin{aligned} g'(x_g) = 0 &\implies \frac{x_g}{1-x_g} = \exp\left(-\frac{JT}{2^{q-2}}\right), \\ f'(x_f) = 0 &\implies g'(x_f) = \lambda \frac{JT(q-1)}{2^{q-3}} x_f \\ &\implies (x_f - x_g)g''(x_g) = \lambda \frac{JT(q-1)}{2^{q-3}} x_g + O(\lambda^2). \end{aligned} \quad (\text{A.22})$$

Another useful relation is

$$g'''(x) = -g''(x)h(x), \quad (\text{A.23})$$

where the function  $h(x)$  is

$$h(x) = \frac{1}{x} - \frac{1}{1-x}. \quad (\text{A.24})$$

Using the above relations, we first evaluate the term in the square root.

$$\begin{aligned} \frac{f''(x_f)x_f(1-x_f)}{g''(x_g)x_g(1-x_g)} &= 1 + \frac{g'''(x_g)(x_f - x_g) - \lambda \frac{JT(q-1)}{2^{q-3}}}{g''(x_g)} + (x_f - x_g)h(x_g) + O(\lambda^2) \\ &\approx 1 - \frac{\lambda}{g''(x_g)} \frac{JT(q-1)}{2^{q-3}} \\ &= 1 - \lambda \frac{JT(q-1)}{2^{q-1}} \operatorname{sech}^2\left(\frac{JT}{2^{q-1}}\right). \end{aligned} \quad (\text{A.25})$$

In a similar manner we can evaluate the expression in the exponential. Finally, we get:

$$F(T) \approx \left(1 + \lambda \frac{JT(q-1)}{2^q} \operatorname{sech}^2\left(\frac{JT}{2^{q-1}}\right)\right) \exp\left[\frac{N_1 \lambda JT(q-1)}{2^q} \exp\left(-\frac{JT}{2^{q-2}}\right) \operatorname{sech}^2\frac{JT}{2^q}\right]. \quad (\text{A.26})$$

The term in the exponential turns out to be equal to the  $O(\lambda^2)$  contribution from the classical action while the factor multiplying the exponential piece is the contribution from the one-loop determinant that we computed in (3.68).

## B Proof of the error correction bound

We use the two different measures of distance in the proof [49]: the trace distance and fidelity. The trace distance between two states  $\rho$  and  $\sigma$  is:

$$\begin{aligned} D(\rho, \sigma) &= \frac{1}{2} \operatorname{Tr}(|\rho - \sigma|) \\ &= \max_Q \operatorname{Tr}(Q(\rho - \sigma)), \end{aligned} \quad (\text{B.1})$$

where  $|A| = \sqrt{A^\dagger A}$ . In the second expression, we maximize over all possible projectors  $Q$ . The fidelity between two states  $\rho$  and  $\sigma$  is defined as:

$$\begin{aligned} F(\rho, \sigma) &= \text{Tr} \left( \sqrt{\sqrt{\sigma} \rho \sqrt{\sigma}} \right) \\ &= \max_{|\psi_\sigma\rangle} |\langle \psi_\rho | \psi_\sigma \rangle|, \end{aligned} \quad (\text{B.2})$$

where  $|\psi_\rho\rangle$  and  $|\psi_\sigma\rangle$  are purification of  $\rho$  and  $\sigma$  respectively.

Consider a maximally entangled state  $|\Psi\rangle$  between the encoded code subspace and a reference system isomorphic to the code subspace:

$$|\Psi\rangle = \sum_i \frac{1}{\sqrt{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes |\psi_i\rangle_{\text{phys}}. \quad (\text{B.3})$$

The physical system interacts with the environment initially in some pure state  $|0\rangle_{\text{env}}$ . This interaction is described by a joint evolution of the physical system and the environment by a unitary  $U_{\mathcal{E}}$  leading to the following final state:

$$|\Psi'\rangle = \sum_i \frac{1}{\sqrt{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes U_{\mathcal{E}}(|\psi_i\rangle_{\text{phys}} \otimes |0\rangle_{\text{env}}). \quad (\text{B.4})$$

Consider now a fictitious state

$$\tilde{\rho}_{\text{ref,env}} = \rho'_{\text{ref}} \otimes \rho'_{\text{env}}, \quad (\text{B.5})$$

where the reduced states are

$$\begin{aligned} \rho'_{\text{ref}} &= \text{Tr}_{\text{env,phys}}(|\Psi'\rangle\langle\Psi'|), \\ \rho'_{\text{env}} &= \text{Tr}_{\text{ref,phys}}(|\Psi'\rangle\langle\Psi'|). \end{aligned} \quad (\text{B.6})$$

Note that  $\tilde{\rho}_{\text{ref,env}}$  is *not* the state  $\rho'_{\text{ref,env}}$ , but is instead a factorized state between  $\mathcal{H}_{\text{ref}}$  and  $\mathcal{H}_{\text{env}}$  that is built from its reduced states. Consider a purification  $|\tilde{\Psi}\rangle$  of  $\tilde{\rho}_{\text{ref,env}}$  such that its trace distance with  $|\Psi'\rangle$ , the quantity  $D(|\tilde{\Psi}\rangle, |\Psi'\rangle)$ , is minimum. By the Schmidt decomposition of pure states, any purification of  $\tilde{\rho}_{\text{ref,env}}$  may be written in the following form:

$$|\tilde{\Psi}\rangle = \sum_{i,j} \sqrt{\frac{\alpha_j}{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes |\phi_{ij}\rangle_{\text{phys}} \otimes |j\rangle_{\text{env}} \quad (\text{B.7})$$

where the set  $\{|\phi_{ij}\rangle\}$  form an orthonormal basis of the physical Hilbert space. The Schmidt coefficients  $\alpha_j$  depend only on the environment index because  $\rho'_{\text{ref}}$  is maximally mixed which restricts the form of the Schmidt coefficients in this manner. Indeed, the state  $\rho'_{\text{env}}$  determines the real non-negative coefficients  $\sqrt{\alpha_j}$  completely. The condition that  $|\tilde{\Psi}\rangle$  should be a purification with minimal  $D(|\tilde{\Psi}\rangle, |\Psi'\rangle)$  is hidden in the basis vectors  $|\phi_{ij}\rangle_{\text{phys}}$ . Define projection operators  $\Pi_j$  as:

$$\Pi_j = \sum_i |\phi_{ij}\rangle\langle\phi_{ij}|. \quad (\text{B.8})$$

These projectors satisfy the following relation:

$$\Pi_j \Pi_k = \delta_{jk} \Pi_k. \quad (\text{B.9})$$

Moreover, every subspace corresponding to the projector  $\Pi_j$  is isomorphic to the code subspace i.e. for each  $\Pi_j$ , there is a unitary operator  $U_j$  such that  $U_j \Pi_j U_j^\dagger = \Pi_{\text{code}}$ , where  $\Pi_{\text{code}}$  is a projector onto the code subspace.

Following [76], we construct a recovery channel  $\tilde{\mathcal{R}}$  which consists of the following two operations:

1. Measurement with some projection operator  $\Pi_j$ , and
2. Rotation of the resulting state by the unitary operator  $U_j$ .

Consider acting with  $\tilde{\mathcal{R}}$  on  $|\tilde{\Psi}\rangle$ . Measurement of  $|\tilde{\Psi}\rangle$  with  $\Pi_j$  projects the state  $|\tilde{\Psi}\rangle$  to the following state with probability  $\alpha_j$ :

$$|\tilde{\Psi}\rangle_j = \sum_i \frac{1}{\sqrt{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes |\phi_{ij}\rangle_{\text{phys}} \otimes |j\rangle_{\text{env}}. \quad (\text{B.10})$$

The unitary transformation  $U_j$  acts on  $|\tilde{\Psi}_j\rangle$  as

$$\begin{aligned} U_j |\tilde{\Psi}\rangle_j &= \sum_i \frac{1}{\sqrt{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes U_j |\phi_{ij}\rangle_{\text{phys}} \otimes |j\rangle_{\text{env}} \\ &= \sum_i \frac{1}{\sqrt{d_{\text{code}}}} |i\rangle_{\text{ref}} \otimes |\psi_i\rangle_{\text{phys}} \otimes |j\rangle_{\text{env}} \\ &= |\Psi\rangle \otimes |j\rangle_{\text{env}}. \end{aligned} \quad (\text{B.11})$$

Thus,  $\tilde{\mathcal{R}}$  acts on  $|\tilde{\Psi}\rangle$  to give back the original state  $|\Psi\rangle$  because the unitary  $U_j$  acts to precisely rotate the basis  $|\phi_{ij}\rangle$  via  $U_j |\phi_{ij}\rangle = |\psi_i\rangle$ . However, we are interested in recovery from the state  $|\Psi'\rangle$ , after the action of the error channel. We will now rephrase the condition for approximate recovery, derived in [76] in terms of the trace distance between the recovered state and the initial state, using the mutual purity between the reference and the environment. We have the following bound on the trace distance between the state obtained by action of the recovery channel  $\tilde{\mathcal{R}}$  on the actual state  $|\Psi'\rangle$  and the initial state  $|\Psi\rangle$ .

$$\begin{aligned} D\left(\tilde{\mathcal{R}}(|\Psi'\rangle\langle\Psi'|), |\Psi\rangle\langle\Psi|\right) &= D\left(\tilde{\mathcal{R}}(|\Psi'\rangle\langle\Psi'|), \tilde{\mathcal{R}}(|\tilde{\Psi}\rangle\langle\tilde{\Psi}|)\right) \\ &\leq D\left(|\Psi'\rangle, |\tilde{\Psi}\rangle\right) \\ &= \sqrt{1 - |\langle\Psi'| \tilde{\Psi}\rangle|^2} \\ &= \sqrt{1 - F^2(\rho'_{\text{ref,env}}, \rho'_{\text{ref}} \otimes \rho'_{\text{env}})} \\ &\leq \sqrt{2 - 2F(\rho'_{\text{ref,env}}, \rho'_{\text{ref}} \otimes \rho'_{\text{env}})} \\ &\leq \sqrt{2D(\rho'_{\text{ref,env}}, \rho'_{\text{ref}} \otimes \rho'_{\text{env}})} \\ &\leq \sqrt{d_{\text{ref}} d_{\text{env}}} \lambda_{\text{max}} \\ &\leq \sqrt{d_{\text{ref}} d_{\text{env}}} \left( \text{Tr} \left( \rho'_{\text{ref,env}} - \rho'_{\text{ref}} \otimes \rho'_{\text{env}} \right)^2 \right)^{1/4} \\ &= \sqrt{d_{\text{ref}} d_{\text{env}}} \left( \text{Tr} \left( \rho_{\text{ref,env}}^2 - \rho_{\text{ref}}^2 \otimes \rho_{\text{env}}^2 \right) \right)^{1/4}. \end{aligned} \quad (\text{B.12})$$

In the second step, we used the monotonicity property of trace distance with respect to the action of a channel (see chapter 9 of [49]). The fourth step follows from the definition of

fidelity and the fact that  $|\tilde{\Psi}\rangle$  is a purification of  $\tilde{\rho}_{\text{ref,env}}$  that minimizes its trace distance with  $|\Psi'\rangle$ . The sixth step is a standard inequality between fidelity and trace distance [49]. As in the main text,  $d_{\text{ref}}$  and  $d_{\text{env}}$  are respective dimensions of the reference and the environment Hilbert spaces. In the seventh step,  $\lambda_{\text{max}}$  is the maximum eigenvalue of  $|\rho'_{\text{ref,env}} - \rho'_{\text{ref}} \otimes \rho'_{\text{env}}|$ . Since  $\lambda_{\text{max}}^2 < \text{Tr}(\rho'_{\text{ref,env}} - \rho'_{\text{ref}} \otimes \rho'_{\text{env}})^2$ , the eighth step follows. The final step is true because  $\rho'_{\text{ref}}$  is maximally mixed.

To summarize, we have shown that there exists a set of projection operators  $\Pi_j$ , the measurement of which followed by a unitary transformation with  $U_j$  approximately recovers the maximally entangled state between the reference and the physical system. The accuracy of this recovery in terms of trace distance is bounded by the combination we have found, which is the mutual purity  $\mathcal{F}_{\Psi'}(\text{ref} : \text{env})$  from the main text.

The inequality (B.12) was derived for a specific recovery channel  $\tilde{\mathcal{R}}$ . However, there may exist a better recovery channel  $\mathcal{R}$  which must also satisfy the inequality:

$$D(\mathcal{R}(|\Psi'\rangle\langle\Psi'|), |\Psi\rangle\langle\Psi|) \leq \sqrt{d_{\text{ref}} d_{\text{env}}} \left( \text{Tr} \left( \rho_{\text{ref,env}}'^2 - \rho_{\text{ref}}'^2 \otimes \rho_{\text{env}}'^2 \right) \right)^{1/4}. \quad (\text{B.13})$$

We can use the above inequality to compute a bound on recovery of arbitrary states in the code subspace after the action of the error channel  $\mathcal{E}$ . We will use the channel-state isomorphism of [77] as follows: consider a state  $\sigma = \sum_{m,n} \sigma_{mn} |\psi_m\rangle\langle\psi_n|$  in the code subspace and let  $\sigma' = \mathcal{R} \circ \mathcal{E}(\sigma)$ . We can write  $\sigma'$  in terms of  $\sigma_{\text{ref}} = \sum_{m,n} \sigma_{mn} (|m\rangle\langle n|)_{\text{ref}}$  and  $\omega = \mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|)$  as

$$\begin{aligned} \sigma' &= \sum_{m,n} \sigma_{mn} \mathcal{R} \circ \mathcal{E}(|\psi_m\rangle\langle\psi_n|) \\ &= d_{\text{ref}} \sum_{k,l} \text{Tr}_{\text{ref}} \left( \frac{1}{d_{\text{ref}}} |k\rangle\langle l| \sigma_{\text{ref}}^T \right) \mathcal{R} \circ \mathcal{E}(|\psi_k\rangle\langle\psi_l|) \\ &= d_{\text{ref}} \text{Tr}_{\text{ref}} \left( \sigma_{\text{ref}}^T \mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|) \right). \end{aligned} \quad (\text{B.14})$$

Here  $\sigma_{\text{ref}}^T$  is the transpose of  $\sigma_{\text{ref}}$ . We have a similar expression for  $\sigma$ :

$$\sigma = d_{\text{ref}} \text{Tr}_{\text{ref}} \left( \sigma_{\text{ref}}^T |\Psi\rangle\langle\Psi| \right). \quad (\text{B.15})$$

We can derive a bound on the trace distance between  $\sigma$  and  $\sigma'$  as follows:

$$\begin{aligned} D(\sigma', \sigma) &= \frac{1}{2} \text{Tr}_{\text{phys}} |\sigma' - \sigma| \\ &= \frac{1}{2} d_{\text{ref}} \text{Tr}_{\text{phys}} \left| \text{Tr}_{\text{ref}} \left( \sigma_{\text{ref}}^T (\mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|) - |\Psi\rangle\langle\Psi|) \right) \right| \\ &\leq \frac{1}{2} d_{\text{ref}} \text{Tr} (|\sigma_{\text{ref}}^T (\mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|) - |\Psi\rangle\langle\Psi|)|) \\ &= \frac{1}{2} d_{\text{ref}} \text{Tr} (|\sigma_{\text{ref}}^T \otimes \Pi_{\text{code}} (\mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|) - |\Psi\rangle\langle\Psi|)|) \\ &\leq d_{\text{ref}} \text{Tr} \left( \sigma_{\text{ref}}^T \otimes \Pi_{\text{code}} \right) \frac{1}{2} \text{Tr} (|\mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|) - |\Psi\rangle\langle\Psi|)|) \\ &= d_{\text{ref}} \text{Tr} \left( \sigma_{\text{ref}}^T \otimes \Pi_{\text{code}} \right) D(\mathcal{R} \circ \mathcal{E}(|\Psi\rangle\langle\Psi|), |\Psi\rangle\langle\Psi|) \\ &\leq d_{\text{ref}}^{5/2} d_{\text{env}}^{1/2} \left( \text{Tr} \left( \rho_{\text{ref,env}}'^2 - \rho_{\text{ref}}'^2 \otimes \rho_{\text{env}}'^2 \right) \right)^{1/4}. \end{aligned} \quad (\text{B.16})$$

In the fifth step,  $\Pi_{\text{code}}$  is the projector on the code subspace. In the final step, we used the inequality in (B.13). The result above is precisely the one quoted in (2.15).

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [[hep-th/0603001](#)] [[INSPIRE](#)].
- [2] M. Headrick, V.E. Hubeny, A. Lawrence and M. Rangamani, *Causality & holographic entanglement entropy*, *JHEP* **12** (2014) 162 [[arXiv:1408.6300](#)] [[INSPIRE](#)].
- [3] X. Dong, D. Harlow and A.C. Wall, *Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality*, *Phys. Rev. Lett.* **117** (2016) 021601 [[arXiv:1601.05416](#)] [[INSPIRE](#)].
- [4] E. Verlinde and H. Verlinde, *Black Hole Entanglement and Quantum Error Correction*, *JHEP* **10** (2013) 107 [[arXiv:1211.6913](#)] [[INSPIRE](#)].
- [5] K. Papadodimas and S. Raju, *State-Dependent Bulk-Boundary Maps and Black Hole Complementarity*, *Phys. Rev. D* **89** (2014) 086010 [[arXiv:1310.6335](#)] [[INSPIRE](#)].
- [6] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090 [[arXiv:1304.4926](#)] [[INSPIRE](#)].
- [7] D.L. Jafferis, A. Lewkowycz, J. Maldacena and S.J. Suh, *Relative entropy equals bulk relative entropy*, *JHEP* **06** (2016) 004 [[arXiv:1512.06431](#)] [[INSPIRE](#)].
- [8] A. Almheiri, X. Dong and D. Harlow, *Bulk Locality and Quantum Error Correction in AdS/CFT*, *JHEP* **04** (2015) 163 [[arXiv:1411.7041](#)] [[INSPIRE](#)].
- [9] N. Engelhardt and A.C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [[arXiv:1408.3203](#)] [[INSPIRE](#)].
- [10] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, *JHEP* **03** (2022) 205 [[arXiv:1911.11977](#)] [[INSPIRE](#)].
- [11] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [[arXiv:1905.08762](#)] [[INSPIRE](#)].
- [12] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, *JHEP* **09** (2020) 002 [[arXiv:1905.08255](#)] [[INSPIRE](#)].
- [13] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [[arXiv:1908.10996](#)] [[INSPIRE](#)].
- [14] N. Engelhardt, G. Penington and A. Shahbazi-Moghaddam, *Finding pythons in unexpected places*, *Class. Quant. Grav.* **39** (2022) 094002 [[arXiv:2105.09316](#)] [[INSPIRE](#)].
- [15] N. Engelhardt, G. Penington and A. Shahbazi-Moghaddam, *A world without pythons would be so simple*, *Class. Quant. Grav.* **38** (2021) 234001 [[arXiv:2102.07774](#)] [[INSPIRE](#)].
- [16] C. Akers and G. Penington, *Quantum minimal surfaces from quantum error correction*, *SciPost Phys.* **12** (2022) 157 [[arXiv:2109.14618](#)] [[INSPIRE](#)].
- [17] A.R. Brown, H. Gharibyan, G. Penington and L. Susskind, *The Python's Lunch: geometric obstructions to decoding Hawking radiation*, *JHEP* **08** (2020) 121 [[arXiv:1912.00228](#)] [[INSPIRE](#)].



- [18] A. Almheiri et al., *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013 [[arXiv:1911.12333](#)] [[INSPIRE](#)].
- [19] V. Balasubramanian et al., *Geometric secret sharing in a model of Hawking radiation*, *JHEP* **01** (2021) 177 [[arXiv:2003.05448](#)] [[INSPIRE](#)].
- [20] H. Geng et al., *Inconsistency of islands in theories with long-range gravity*, *JHEP* **01** (2022) 182 [[arXiv:2107.03390](#)] [[INSPIRE](#)].
- [21] H. Geng, Y. Nomura and H.-Y. Sun, *Information paradox and its resolution in de Sitter holography*, *Phys. Rev. D* **103** (2021) 126004 [[arXiv:2103.07477](#)] [[INSPIRE](#)].
- [22] H. Geng et al., *Information Transfer with a Gravitating Bath*, *SciPost Phys.* **10** (2021) 103 [[arXiv:2012.04671](#)] [[INSPIRE](#)].
- [23] V. Balasubramanian, A. Kar and T. Ugajin, *Entanglement between two disjoint universes*, *JHEP* **02** (2021) 136 [[arXiv:2008.05274](#)] [[INSPIRE](#)].
- [24] T. Hartman, Y. Jiang and E. Shaghoulian, *Islands in cosmology*, *JHEP* **11** (2020) 111 [[arXiv:2008.01022](#)] [[INSPIRE](#)].
- [25] L. Anderson, O. Parrikar and R.M. Soni, *Islands with gravitating baths: towards  $ER = EPR$* , *JHEP* **21** (2020) 226 [[arXiv:2103.14746](#)] [[INSPIRE](#)].
- [26] V. Balasubramanian, A. Kar and T. Ugajin, *Entanglement between two gravitating universes*, *Class. Quant. Grav.* **39** (2022) 174001 [[arXiv:2104.13383](#)] [[INSPIRE](#)].
- [27] D. Harlow and P. Hayden, *Quantum Computation vs. Firewalls*, *JHEP* **06** (2013) 085 [[arXiv:1301.4504](#)] [[INSPIRE](#)].
- [28] Y. Zhao, *Petz map and Python's lunch*, *JHEP* **11** (2020) 038 [[arXiv:2003.03406](#)] [[INSPIRE](#)].
- [29] C. Akers et al., *The black hole interior from non-isometric codes and complexity*, [arXiv:2207.06536](#) [[INSPIRE](#)].
- [30] I. Kim, E. Tang and J. Preskill, *The ghost in the radiation: Robust encodings of the black hole interior*, *JHEP* **06** (2020) 031 [[arXiv:2003.05451](#)] [[INSPIRE](#)].
- [31] V. Balasubramanian, A. Kar, C. Li and O. Parrikar, *Quantum error correction in the black hole interior*, *JHEP* **07** (2023) 189 [[arXiv:2203.01961](#)] [[INSPIRE](#)].
- [32] A. Kar, *Non-isometric quantum error correction in gravity*, *JHEP* **02** (2023) 195 [[arXiv:2210.13476](#)] [[INSPIRE](#)].
- [33] V. Balasubramanian, V. Jejjala and J. Simon, *The Library of Babel*, *Int. J. Mod. Phys. D* **14** (2005) 2181 [[hep-th/0505123](#)] [[INSPIRE](#)].
- [34] V. Balasubramanian, J. de Boer, V. Jejjala and J. Simon, *The Library of Babel: On the origin of gravitational thermodynamics*, *JHEP* **12** (2005) 006 [[hep-th/0508023](#)] [[INSPIRE](#)].
- [35] V. Balasubramanian, P. Kraus and M. Shigemori, *Massless black holes and black rings as effective geometries of the  $D1$ - $D5$  system*, *Class. Quant. Grav.* **22** (2005) 4803 [[hep-th/0508110](#)] [[INSPIRE](#)].
- [36] N. Engelhardt and A.C. Wall, *Coarse Graining Holographic Black Holes*, *JHEP* **05** (2019) 160 [[arXiv:1806.01281](#)] [[INSPIRE](#)].
- [37] A. Kitaev, *A simple model of quantum holography (part 1)*, talk at KITP, April 7, 2015, <http://online.kitp.ucsb.edu/online/entangled15/kitaev/>.
- [38] A. Kitaev, *A simple model of quantum holography (part 2)*, talk at KITP, May 27, 2015, <http://online.kitp.ucsb.edu/online/entangled15/kitaev2/>.

- [39] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, *Phys. Rev. D* **94** (2016) 106002 [[arXiv:1604.07818](#)] [[INSPIRE](#)].
- [40] P. Saad, S.H. Shenker and D. Stanford, *A semiclassical ramp in SYK and in gravity*, [arXiv:1806.06840](#) [[INSPIRE](#)].
- [41] D. Stanford, Z. Yang and S. Yao, *Subleading Weingartens*, *JHEP* **02** (2022) 200 [[arXiv:2107.10252](#)] [[INSPIRE](#)].
- [42] A.R. Brown and L. Susskind, *Second law of quantum complexity*, *Phys. Rev. D* **97** (2018) 086015 [[arXiv:1701.01107](#)] [[INSPIRE](#)].
- [43] F.G.S.L. Brandão et al., *Models of Quantum Complexity Growth*, *PRX Quantum* **2** (2021) 030316 [[arXiv:1912.04297](#)] [[INSPIRE](#)].
- [44] V. Balasubramanian, M. Decross, A. Kar and O. Parrikar, *Quantum Complexity of Time Evolution with Chaotic Hamiltonians*, *JHEP* **01** (2020) 134 [[arXiv:1905.05765](#)] [[INSPIRE](#)].
- [45] J. Haferkamp et al., *Linear growth of quantum circuit complexity*, *Nature Phys.* **18** (2022) 528 [[arXiv:2106.05305](#)] [[INSPIRE](#)].
- [46] V. Balasubramanian et al., *Complexity growth in integrable and chaotic models*, *JHEP* **07** (2021) 011 [[arXiv:2101.02209](#)] [[INSPIRE](#)].
- [47] L. Susskind, *Three Lectures on Complexity and Black Holes*, Springer Cham (2018) [[DOI:10.1007/978-3-030-45109-7](#)] [[arXiv:1810.11563](#)] [[INSPIRE](#)].
- [48] M.-D. Choi, *Completely positive linear maps on complex matrices*, *Linear Algebra Appl.* **10** (1975) 285.
- [49] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, U.S.A. (2011) [[DOI:10.1017/cbo9780511976667](#)].
- [50] B. Schumacher and M.A. Nielsen, *Quantum data processing and error correction*, *Phys. Rev. A* **54** (1996) 2629 [[quant-ph/9604022](#)] [[INSPIRE](#)].
- [51] C. Bény and O. Oreshkov, *General Conditions for Approximate Quantum Error Correction and Near-Optimal Recovery Channels*, *Phys. Rev. Lett.* **104** (2010) 120501.
- [52] D. Harlow, *The Ryu-Takayanagi Formula from Quantum Error Correction*, *Commun. Math. Phys.* **354** (2017) 865 [[arXiv:1607.03901](#)] [[INSPIRE](#)].
- [53] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [[arXiv:1307.2892](#)] [[INSPIRE](#)].
- [54] S. Dutta and T. Faulkner, *A canonical purification for the entanglement wedge cross-section*, *JHEP* **03** (2021) 178 [[arXiv:1905.00577](#)] [[INSPIRE](#)].
- [55] C. Akers and G. Penington, *Leading order corrections to the quantum extremal surface prescription*, *JHEP* **04** (2021) 062 [[arXiv:2008.03319](#)] [[INSPIRE](#)].
- [56] J. Chandra and T. Hartman, *Coarse graining pure states in AdS/CFT*, [arXiv:2206.03414](#) [[INSPIRE](#)].
- [57] V. Balasubramanian, P. Caputa, J.M. Magan and Q. Wu, *Quantum chaos and the complexity of spread of states*, *Phys. Rev. D* **106** (2022) 046007 [[arXiv:2202.06957](#)] [[INSPIRE](#)].
- [58] V. Balasubramanian, J.M. Magan and Q. Wu, *Tridiagonalizing random matrices*, *Phys. Rev. D* **107** (2023) 126001 [[arXiv:2208.08452](#)] [[INSPIRE](#)].

- [59] S.-K. Jian, G. Bentsen and B. Swingle, *Linear Growth of Circuit Complexity from Brownian Dynamics*, [arXiv:2206.14205](#) [[INSPIRE](#)].
- [60] W. Brown and O. Fawzi, *Decoupling with Random Quantum Circuits*, *Commun. Math. Phys.* **340** (2015) 867 [[arXiv:1307.0632](#)].
- [61] W. Brown and O. Fawzi, *Short random circuits define good quantum error correcting codes*, in the proceedings of the *2013 IEEE International Symposium on Information Theory*, Istanbul Turkey, July 07–12 (2013) [[DOI:10.1109/isit.2013.6620245](#)].
- [62] M.J. Gullans et al., *Quantum Coding with Low-Depth Random Circuits*, *Phys. Rev. X* **11** (2021) 031066 [[arXiv:2010.09775](#)] [[INSPIRE](#)].
- [63] R. Cleve and D. Gottesman, *Efficient computations of encodings for quantum error correction*, *Phys. Rev. A* **56** (1997) 76 [[quant-ph/9607030](#)] [[INSPIRE](#)].
- [64] S. Aaronson and D. Gottesman, *Improved simulation of stabilizer circuits*, *Phys. Rev. A* **70** (2004) 052328 [[quant-ph/0406196](#)].
- [65] K.-Y. Kuo and C.-Y. Lai, *The Encoding and Decoding Complexities of Entanglement-Assisted Quantum Stabilizer Codes*, [arXiv:1903.10013](#) [[DOI:10.1109/isit.2019.8849380](#)].
- [66] T. Brun, I. Devetak and M.-H. Hsieh, *Correcting Quantum Errors with Entanglement*, *Science* **314** (2006) 436 [[quant-ph/0610092](#)].
- [67] D.W. Berry, A.M. Childs and R. Kothari, *Hamiltonian simulation with nearly optimal dependence on all parameters*, [arXiv:1501.01715](#) [[DOI:10.1109/FOCS.2015.54](#)].
- [68] D.W. Berry et al., *Time-dependent Hamiltonian simulation with  $L^1$ -norm scaling*, [arXiv:1906.07115](#) [[DOI:10.22331/q-2020-04-20-254](#)].
- [69] S. Xu, L. Susskind, Y. Su and B. Swingle, *A Sparse Model of Quantum Holography*, [arXiv:2008.02303](#) [[INSPIRE](#)].
- [70] S. Choi, Y. Bao, X.-L. Qi and E. Altman, *Quantum Error Correction in Scrambling Dynamics and Measurement-Induced Phase Transition*, *Phys. Rev. Lett.* **125** (2020) 030505 [[arXiv:1903.05124](#)] [[INSPIRE](#)].
- [71] M.J. Gullans and D.A. Huse, *Dynamical Purification Phase Transition Induced by Quantum Measurements*, *Phys. Rev. X* **10** (2020) 041020 [[arXiv:1905.05195](#)] [[INSPIRE](#)].
- [72] B. Eastin and E. Knill, *Restrictions on Transversal Encoded Quantum Gate Sets*, *Phys. Rev. Lett.* **102** (2009) 110502.
- [73] P. Faist et al., *Continuous symmetries and approximate quantum error correction*, *Phys. Rev. X* **10** (2020) 041018 [[arXiv:1902.07714](#)] [[INSPIRE](#)].
- [74] L. Kong and Z.-W. Liu, *Near-Optimal Covariant Quantum Error-Correcting Codes from Random Unitaries with Symmetries*, *PRX Quantum* **3** (2022) 020314 [[arXiv:2112.01498](#)] [[INSPIRE](#)].
- [75] Y. Gu, A. Kitaev, S. Sachdev and G. Tarnopolsky, *Notes on the complex Sachdev-Ye-Kitaev model*, *JHEP* **02** (2020) 157 [[arXiv:1910.14099](#)] [[INSPIRE](#)].
- [76] B. Schumacher and M.D. Westmoreland, *Approximate quantum error correction*, [quant-ph/0112106](#).
- [77] M. Horodecki, P. Horodecki and R. Horodecki, *General teleportation channel, singlet fraction, and quasidistillation*, *Phys. Rev. A* **60** (1999) 1888.