Error Correction of Quantum Reference Frame Information

Patrick Hayden,[1] Sepehr Nezami,[1,2] Sandu Popescu,[3] and Grant Salton[1,2,4,5,*]

[1] *Stanford Institute for Theoretical Physics, Stanford University, Stanford, California 94305, USA*
[2] *Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA*
[3] *H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*
[4] *Amazon Quantum Solutions Lab, Seattle, Washington 98170, USA*
[5] *AWS Center for Quantum Computing, Pasadena, California 91125, USA*

(Received 11 February 2019; revised 5 August 2020; accepted 12 January 2021; published 18 February 2021)

The existence of quantum error-correcting codes is one of the most counterintuitive and potentially technologically important discoveries of quantum-information theory. In this paper, we study a problem called "covariant quantum error correction", in which the encoding is required to be group covariant. This problem is intimately tied to fault-tolerant quantum computation and the well-known Eastin-Knill theorem. We show that this problem is equivalent to the problem of encoding reference-frame information. In standard quantum error correction, one seeks to protect abstract quantum information, i.e., information that is independent of the physical incarnation of the systems used for storing the information. There are, however, other forms of information that are *physical*—one of the most ubiquitous being reference-frame information. The basic question we seek to answer is whether or not error correction of physical information is possible and, if so, what limitations govern the process. The main challenge is that the systems used for transmitting physical information, in addition to any actions applied to them, must necessarily obey these limitations. Encoding and decoding operations that obey a restrictive set of limitations need not exist *a priori*. Equivalently, there may not exist covariant quantum error-correcting codes. Indeed, we prove a no-go theorem showing that no finite-dimensional, group-covariant quantum codes exist for Lie groups with an infinitesimal generator [e.g., U(1), SU(2), and SO(3)]. We then explain how one can circumvent this no-go theorem using infinite-dimensional codes, and we give an explicit example of a covariant quantum error-correcting code using continuous variables for the group U(1). Finally, we demonstrate that all finite groups have finite-dimensional codes, giving both an explicit construction and a randomized approximate construction with exponentially better parameters. Our results imply that one can, in principle, circumvent the Eastin-Knill theorem.

DOI: 10.1103/PRXQuantum.2.010326

I. INTRODUCTION

One of Shannon's original insights in the formulation of information theory was to focus on the transmission of sequences of symbols, such as strings of 0's and 1's, without regard to the semantic content of the message. This approach makes it possible to encode an enormous variety of messages, from phone numbers to photos, as long as the original information can be faithfully represented in terms of a sequence of symbols. The same situation exists in the quantum world: quantum-information theorists are primarily concerned with information that can be stored in a system of qubits (or larger quantum systems), independent of the type of information.

Here we study a situation in which the information is *physical* and cannot be represented simply as abstract qubits. Consider the following purely classical scenario [1]. Alice wishes to transmit some directional information to Bob, e.g., the axis of rotation of a gyroscope indicated by the vector $\vec{n}$, so that Bob can prepare a gyroscope rotating around the same axis as Alice's. If Alice and Bob share a reference frame, Alice can measure different components of $\vec{n}$ and describe the result *in words* to Bob, who then prepares his own gyroscope to match. However, if Alice and Bob do not share a reference frame, i.e., they do not know the relative alignment of their coordinate systems, then this task is impossible. Without a shared reference frame, Alice has no way to communicate a set of symbols to Bob

[*] gsalton@caltech.edu

indicating the axis of rotation of her gyroscope. Another simple example is clock synchronization, wherein two distant observers want to synchronize their clocks, but it is not possible to do so by sending purely symbolic messages [2].

Of course, the simple examples described above do not mean that sending physical information is impossible. For example, in the classical example, Alice can prepare and send a physical copy of her gyroscope to Bob, thereby indicating her direction. In this way, Alice and Bob can even establish a shared reference frame. Similarly, in the clock synchronization problem Alice can send a copy of her clock to Bob [3] to establish a common time standard (ignoring relativistic effects). Quantum mechanically, Alice can send direction information by sending polarized spins, while timing information can be sent using quantum clocks such as two-level atoms. As is common in the quantum case, many interesting and counterintuitive effects occur. For example, sending two antiparallel spins polarized along the desired direction is a better direction indicator than sending two parallel spins [1,4]. The problem of aligning quantum reference frames has garnered significant attention in recent years [1,3–13].

In this paper we are interested in quantum error correction of physical information. Crucially, *physical information* can only be communicated using systems that themselves have the physical property of interest. This restriction constrains the actions that can be performed on the physical systems, since we cannot, for example, destroy or change physical information arbitrarily. In particular, there may be constraints on the set of possible encoding and decoding schemes that one might have used to make the system more robust to errors, thereby limiting our ability to perform quantum error correction. In this paper, we characterize the constraints placed on quantum error correction of physical information.

In each of the examples described above, Bob's lack of knowledge about Alice's reference frame or time standard is modeled by the action of an unknown element of some group on Alice's state. For directional reference frames, Alice and Bob are related by an unknown rotation [i.e., an element of SO(3)], whereas in the example of clock synchronization their clocks are related by an unknown time translation [i.e., an element of U(1)]. In the spirit of Ref. [14] (which generalizes reference-frame information to general resource theory of asymmetry [14–19]), we study error correction of physical information that transforms under an arbitrary group G. An important reduction following from the analysis of Ref. [14] is that the existence of encoding schemes for this type of information is equivalent to the existence of ordinary, yet G-covariant, encoding schemes, which can correct the same errors.

In this paper, we first study the case in which the group G has at least one infinitesimal generator. In this first case, we find a result strikingly different from conventional, abstract quantum information: we prove a *no-go* theorem showing that it is impossible to encode physical information in any number of finite-dimensional systems such that the encoding allows for perfect correction of any erasure error. We then show that both conditions of the no-go theorem are necessary by constructing codes that circumvent the theorem when either of the conditions is violated. Specifically, we first demonstrate how one can encode physical information to protect against erasure errors when one uses continuous-variable modes (with *infinite-dimensional Hilbert spaces*). Since continuous-variable modes are used, we expect this result to be of practical interest. We then construct a perfect encoding scheme for any *finite group* G into finite-dimensional spaces, which is again robust to erasure errors. Finally, we study a family of group-covariant random codes and show that they can provide encoding schemes with better parameters than the perfect schemes for finite groups.

It is worth noting that the covariant channel formulation of the problem is closely related to other results in the literature that have very different motivations, including the *Eastin-Knill theorem* [20] and recent studies of *invariant perfect tensors* [21]. We present a more detailed comparison in the discussion.

II. REFERENCE-FRAME ERROR CORRECTION

We begin with a description of error correction of spatial reference frames, which corresponds to $G = \mathrm{SO}(3)$; the generalization to other groups is immediate. Suppose Alice and Bob share a (possibly noisy) quantum channel. Alice wants to communicate some directional quantum information (a single spin, say) to Bob, but Alice and Bob do not share a common reference frame. Specifically, their reference frames are related by an unknown rotation $R \in \mathrm{SO}(3)$. Alice and Bob will claim success if Bob receives the spin in the same direction that it was sent by Alice (i.e., the directional information is unchanged—a condition they could check at a later stage). If the task is successful, Bob can use the received spin to establish a shared reference frame, among other things.

When sending quantum information through a noisy channel, the information can be corrupted. This is also true of directional information, so we wish to error correct this type of information. We fix our error model to be *erasure* of a single spin (or mode), and our goal is to design an error-correcting code to protect the directional information from this noise.

To simplify the presentation, we focus on an encoding of one spin into three (see Fig. 1) without loss of generality. We emphasize that this choice of one into three is just for clarity—our results hold for an arbitrary one-to-many encoding. We split the process into six steps.

1. Figure 1(a). Alice starts with an unknown input state ρ_{in}, a density operator on $\mathcal{H}_{\text{in}}$, representing some directional information. Alice encodes this initial state using an encoding channel $\mathcal{E}_A$. We use the subscript A to indicate that $\mathcal{E}_A$ is the encoding map in Alice's reference frame, and to distinguish it from the map as seen in Bob's frame: $\mathcal{E}_B$, to which we return shortly. Thus, the encoded state σ_{123} on three spins is given by $\sigma_{123} = \mathcal{E}_A(\rho_{\text{in}})$.
2. Figure 1(b). Spin $j \in \{1,2,3\}$ is lost. This is an erasure error of any one of the spins, but we assume that Bob can infer which.
3. Figure 1(c). Prior to the erasure error, the encoded state as seen by Bob would be $U_1 \otimes U_2 \otimes U_3 \sigma_{123} U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger$, where $U_i = U_i(R)$ is a unitary representation of the unknown rotation R mapping Alice's coordinate system to Bob's. Bob then receives the state $\text{tr}_j\,(U_1 \otimes U_2 \otimes U_3 \sigma_{123} U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger)$.
4. Figure 1(d). Bob decodes the state with an R-independent decoding map $\mathcal{D}_j$ to obtain $\mathcal{D}_j\Big[\text{tr}_j\,(U_1 \otimes U_2 \otimes U_3 \sigma_{123} U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger)\Big]$, in Bob's reference frame. If the protocol is successful, this state should be equal to $\tilde{\rho}_{\text{in}} = U_{\text{in}} \rho_{\text{in}} U_{\text{in}}^\dagger$ in order to match Alice's original state, where $U_{\text{in}} = U_{\text{in}}(R)$ is the representation of the rotation group acting on the initial state, and the tilde signals that this is the input state as seen from Bob's rotated reference frame.
5. Figure 1(e). Bob sends the decoded state through a hypothetical perfect channel to Alice for verification.
6. Figure 1(f). Success is claimed if the received state is the same as the initial state in Alice's frame.

Using $\rho_{\text{in}} = U_{\text{in}}^\dagger \tilde{\rho}_{\text{in}}\, U_{\text{in}}$, the success condition becomes

$$\tilde{\rho}_{\text{in}} = \mathcal{D}_j\left\{\text{tr}_j\left[U_1 \otimes U_2 \otimes U_3\ \mathcal{E}_A(U_{\text{in}}^\dagger \tilde{\rho}_{\text{in}} U_{\text{in}}) U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger\right]\right\}, \tag{1}$$

for all $R \in \text{SO}(3)$, states $\tilde{\rho}_{\text{in}} \in \mathcal{H}_{\text{in}}$, and $j \in \{1,2,3\}$.

III. COVARIANT ERROR CORRECTION

Covariant quantum error correction is a seemingly different problem in which the encoding map is required to commute with the action of the group. Continuing the example of mapping a single spin into three, the covariance requirement is that the encoding map satisfies

$$U_1 \otimes U_2 \otimes U_3\ \mathcal{E}(U_{\text{in}}^\dagger \rho_{\text{in}}\ U_{\text{in}}) U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger = \mathcal{E}(\rho_{\text{in}}) \tag{2}$$

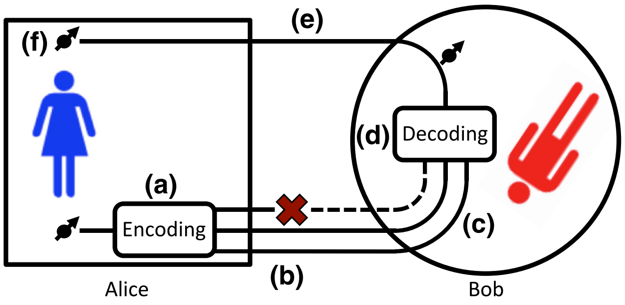

FIG. 1. Setup: Alice wants to send a spin to Bob, but Alice and Bob do not share a reference frame. (a) Alice encodes her spin into an error-correcting code. (b) The environment erases one of the spins. (c) Bob receives the encoded spins in *his* reference frame. (d) Bob then decodes the remaining spins to reveal the original state. (e) Bob sends the decoded spin using a (hypothetical) perfect channel to Alice for verification. (f) Alice confirms that the recovered state is the same as her original state.

for all $R \in \text{SO}(3)$ and initial states ρ_{in}. In this version of the problem Alice and Bob are assumed to share a single reference frame. Imposing the simple constraint, Eq. (2), on the encoding map, however, defines an error-correction problem equivalent to reference-frame erasure correction, as we now see.

Let us return to the setting of reference-frame error correction momentarily. Alice performs the encoding $\mathcal{E}_A$ in her reference frame. In Bob's reference frame, this operation is denoted by $\mathcal{E}_{B,R}$ ($\mathcal{E}_{B,R}$ is the quantum channel corresponding to the operation Alice performs as seen in Bob's reference frame). For a fixed $\mathcal{E}_A$ in Alice's reference frame, $\mathcal{E}_B$ in Bob's frame is still parametrized by the unknown rotation R, i.e., $\mathcal{E}_B = \mathcal{E}_{B,R}$. Specifically, $\mathcal{E}_{B,R}(\tilde{\rho}_{\text{in}}) = U_1 \otimes U_2 \otimes U_3\ \mathcal{E}_A(U_{\text{in}}^\dagger \tilde{\rho}_{\text{in}}\ U_{\text{in}})\ U_1^\dagger \otimes U_2^\dagger \otimes U_3^\dagger$. The success condition simplifies to

$$\mathcal{D}_j\{\text{tr}_j\,[\mathcal{E}_{B,R}(\tilde{\rho}_{\text{in}})]\} = \tilde{\rho}_{\text{in}}, \tag{3}$$

for all states $\tilde{\rho}_{\text{in}}$ and $j \in \{1,2,3\}$.

Now introduce the *average channel* $\mathcal{E} = \mathbb{E}_R[\mathcal{E}_{B,R}]$, where the average is over all rotations $R \in \text{SO}(3)$ according to the Haar measure. By the linearity, the error-correction relation Eq. (3), holds for the average channel: $\mathcal{D}_j\{\text{tr}_j\,[\mathcal{E}(\tilde{\rho}_{\text{in}})]\} = \tilde{\rho}_{\text{in}}$. Moreover, the averaged channel is clearly covariant in the sense of Eq. (2), provided we substitute $\tilde{\rho}_{\text{in}}$ for ρ_{in} in the equation.

Thus if reference-frame error correction, Eq. (1), is possible, we find a covariant erasure-correcting encoding. Moreover, it is straightforward to confirm that by choosing $\mathcal{E}$ to be $\mathcal{E}_A$, Eqs. (3) and (2) lead to Eq. (1). Therefore, reference-frame error correction and covariant error correction are equivalent.

IV. RESULTS

We now study a more general problem. Consider an encoding map $\mathcal{E}$, which encodes an initial state on $\mathcal{H}_{\text{in}}$ into n encoded systems on $\mathcal{H}_{\text{out}} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. The output Hilbert spaces are arbitrary at this point (i.e., they can be the same or different, finite or infinite dimensional, etc.). Suppose there exists a group G, and representations $U_{\text{in}}, U_1, \ldots, U_n$ acting on the different Hilbert spaces. Moreover, suppose that the channel is covariant under the action of the group:

$$\mathcal{E}(\rho_{\text{in}}) = U_1 \otimes \cdots \otimes U_n \mathcal{E}(U_{\text{in}}^\dagger \rho_{\text{in}} U_{\text{in}}) U_1^\dagger \otimes \cdots \otimes U_n^\dagger. \quad (4)$$

Our goal is to answer the following question: *is it possible to recover the original state after erasure of an arbitrary set of at most k subsystems (which we henceforth refer to as modes)?*

We study this question in different scenarios.

1. *G is a Lie group and the code is finite dimensional.* We prove a no-go theorem: no perfect covariant error-correcting scheme can be implemented in this case. This applies to the example of sending spins, as in the original reference-frame error-correction task. In fact, the *no-go* theorem applies to all groups with at least one infinitesimal generator, and it states that such generators can only act trivially on encoded states.
2. *G is a Lie group and the code is infinite dimensional.* We show that G-covariant error-correcting codes are possible when the encoding uses infinite-dimensional systems. This illustrates the existence of interesting error-correcting codes for a Lie group when the conditions of the no-go theorem above are not satisfied. We provide an explicit code for $G = \mathrm{U}(1)$ in Appendix A.
3. *G is a finite group and the code is finite dimensional.* For any finite group G, we find examples of perfect covariant error-correcting schemes. This is again consistent with our no-go theorem since finite groups do not have infinitesimal generators. We also provide a randomized construction in Appendix B to obtain approximate codes with better parameters.

V. CASE 1: *G* IS A LIE GROUP AND THE CODE IS FINITE DIMENSIONAL

Suppose that the local Hilbert space dimensions are all finite, and that the group G is a Lie group [22]. Choose one infinitesimal generator of the Lie group, without loss of generality. We denote this generator acting on the input mode by T_{in} and on the ith output mode by T_i. The generator acting on the full set of output modes is $T_{\text{out}} = T_1 + \cdots T_n$. Assume that T_{in} is nontrivial; our goal is to show that covariant quantum error correction is impossible with this assumption.

Consider an initial state ρ_{in} and a slightly rotated state $\rho_{\text{in}}(\epsilon) = e^{-i\epsilon T_{\text{in}}} \rho_{\text{in}} e^{i\epsilon T_{\text{in}}}$. These states are encoded as $\sigma_{\text{out}} = \mathcal{E}(\rho_{\text{in}})$ and $\sigma_{\text{out}}(\epsilon) = \mathcal{E}[\rho_{\text{in}}(\epsilon)]$. Using the fact that $\mathcal{E}(\rho_{\text{in}})$ is invertible on its range, we can find a set of orthogonal isometries $\{E_i\}$, $(E_i^\dagger E_j = \delta_{ij} I)$ and probabilities p_i such that

$$\mathcal{E}(\rho_{\text{in}}) = \sum_i p_i E_i \rho_{\text{in}} E_i^\dagger$$

(see, e.g., [23], Theorem 10.1 and the proof using $\mathcal{H}_{\text{in}}$ as the code space). The inverse channel $\mathcal{E}^{-1}(\sigma_{\text{out}})$ can be described by the same set of isometries on the range of $\mathcal{E}$

$$\mathcal{E}^{-1}(\rho_{\text{out}}) = \sum_i E_i^\dagger \rho_{\text{out}} E_i + \Pi_\perp \rho_{\text{out}} \Pi_\perp,$$

where $\Pi_\perp = I - \sum_i E_i E_i^\dagger$. A crucial but elementary property of $\mathcal{E}^{-1}$ is that if $\sigma_{\text{out}} = \mathcal{E}(\rho_{\text{in}})$ and A is some arbitrary operator, then $\mathcal{E}^{-1}(A\sigma_{\text{out}}) = \mathcal{E}^\dagger(A)\rho_{\text{in}}$, where $\mathcal{E}^\dagger(A) = \sum_i p_i E_i^\dagger A E_i$. Expanding the relation $\rho_{\text{in}} - \rho_{\text{in}}(\epsilon) = \mathcal{E}^{-1}[\sigma_{\text{out}} - \sigma_{\text{out}}(\epsilon)]$ to first order in ϵ we obtain

$$\begin{aligned}[T_{\text{in}}, \rho_{\text{in}}] &= \mathcal{E}^{-1}([T_{\text{out}}, \sigma_{\text{out}}]) \\ &= [\mathcal{E}^\dagger(T_{\text{out}}), \rho_{\text{in}}]. \end{aligned} \quad (5)$$

Assuming error correction succeeds, we can recover the original state from any of the $n-k$ subsets of the encoded modes. In other words, upon tracing out all output modes *except* the ith mode, the reduced state ρ_i must be independent of the initial state.

Thus, for any state ρ_{in}, we find that $\mathrm{tr}(T_i \sigma_{\text{out}}) = \alpha_i$, where α_i is independent of ρ_{in}. It is easy to see that

$$\alpha_i = \mathrm{tr}(T_i \sigma_{\text{out}}) = \mathrm{tr}[T_i \mathcal{E}(\rho_{\text{in}})] = \mathrm{tr}[\mathcal{E}^\dagger(T_i)\rho_{\text{in}}]$$

for all ρ_{in}. Hence $\mathcal{E}^\dagger(T_i) \propto I$, and consequently $\mathcal{E}^\dagger(T_{\text{out}}) \propto I$. This implies that the last term in Eq. (5) is zero, which means that $[T_{\text{in}}, \rho_{\text{in}}] = 0$ for all ρ_{in}. In order for T_{in} to commute with all ρ_{in} it must be trivial, which is a contradiction of our assumption. We conclude that perfect recoverability is impossible.

VI. CASE 2: *G* IS A LIE GROUP AND THE CODE IS INFINITE DIMENSIONAL

If we allow Alice to use infinite-dimensional Hilbert spaces (violating one of the hypotheses of our no-go theorem), then even a naïve solution to the problem exists. Intuitively, a simple way to achieve the task is for Alice to append a classical gyroscope to the encoded state that she sends to Bob [24]. Bob can then infer information about Alice's reference frame by measuring the state of the gyroscope, thereby establishing a common reference frame. Indeed, this is one strategy we outline below. Since the full state is sent through the noisy channel, Alice must actually send *two* gyroscopes in order to safeguard against loss of one of the encoded shares [25].

In the reference-frame error-correction paradigm, Alice chooses her favorite (noncovariant) erasure code, encodes, and then appends two redundant ancilla (the classical gyroscopes) indicating her reference frame to the encoded state. The ancilla must necessarily be states in infinite-dimensional Hilbert spaces so that the no-go theorem does not apply (and in this protocol this is also necessary so that Alice can specify her reference frame with perfect precision) [26]. If any shares of the erasure code are lost, Bob can first measure the gyroscopes to learn Alice's reference frame, and then use the standard decoding on the remaining shares in the aligned frame. Since Alice sent two ancilla, one can freely be lost without failure.

Let us now study this problem in the covariant quantum error-correction paradigm. Let $\mathcal{H}_G = \text{span}\{|g\rangle\}$, where $g \in G$, and the set $\{|g\rangle\}$ forms an orthonormal basis for $\mathcal{H}_G$. The group acts via $U(g)\,|h\rangle = |gh\rangle$ [27]. To encode her state, Alice chooses her favorite, noncovariant erasure-correcting code (denoted by $\mathcal{E}_0$) [e.g., the $\mathbb{C}^3 \to (\mathbb{C}^3)^{\otimes 3}$ qutrit code] without loss of generality. As before, define the rotated encoding map (i.e., the map in Bob's frame) by

$$\mathcal{E}_g(\Psi) = U(g)^{\otimes 3}\mathcal{E}_0\left[U^\dagger(g)\Psi U(g)\right]U^{\dagger\otimes 3}(g). \qquad (6)$$

To complete the encoding, Alice appends two ancilla in the state $|e\rangle\langle e|$ (where $e \in G$ is the identity element) for a full encoded state $\mathcal{E}_0(\Psi) \otimes |e\rangle\langle e|^{\otimes 2}$ as seen in her frame. The two $|e\rangle\langle e|$ registers represent the classical gyroscopes above. The encoding is made *covariant* by averaging over the group G. Thus, the full encoding is defined by symmetrizing the channel and ancilla together:

$$\mathcal{E}(\Psi) = \int_{g\in G} dg \quad \mathcal{E}_g(\Psi) \otimes |g\rangle\langle g|^{\otimes 2},$$

which is clearly covariant.

Decoding is then fairly simple: one need only measure any surviving ancilla, which collapses the state to one corresponding to the *measured* group element. Bob can then recover the encoded state from the surviving shares of the code.

The procedure described above is not the only method one can use in this case. In Appendix A we describe an explicit, group-covariant, continuous-variable quantum erasure code for the example of $G = \text{U}(1)$. An input continuous variable mode is mapped into three physical modes via the encoding

$$E_{U(1)} = \sum_{x,y\in\mathbb{Z}} |-3y, -x+y, 2(y+x)\rangle_{123}\,\langle x|_{\text{in}}.$$

We leave all relevant details to Appendix A.

VII. CASE 3: *G* IS A FINITE GROUP AND THE CODE IS FINITE DIMENSIONAL

Consider a *finite* group G. Here we show that there exist G-covariant channels that encode the input Hilbert space into finite-dimensional Hilbert spaces, while satisfying the erasure-correction conditions.

Suppose the group G acts on some set A. By definition, the action of G permutes the elements of A. Our goal is to construct an error-correction scheme for which the action of the group commutes with the process of encoding, erasure, and decoding. To achieve our goal, we first start with a noncovariant code. We then consider a tensor product of many copies of this noncovariant code, one tensor factor for each element of A. This new code is already a covariant code! To see this, note that the encoding acts as a tensor product over the factors, while the group action simply permutes the factors. Therefore, the encoding map and the group action commute, which implies that the encoding is G covariant.

To be more precise, consider a channel $\mathcal{E}_0 : S(\mathcal{H}_{\text{in}}) \to S(\mathcal{H}_{\text{out}} := \mathcal{H}^{\otimes n})$, where $S(\mathcal{H})$ denotes the space of density matrices on the Hilbert space $\mathcal{H}$. Suppose that $\mathcal{E}_0$ is an encoding map that allows for recovery after erasure of an arbitrary set of k of the n output modes. We make no assumptions about the covariance of $\mathcal{E}_0$—it is an arbitrary erasure-correcting map. We now introduce a new encoding

$$\mathcal{E} = \bigotimes_{a\in A}\mathcal{E}_0 = \mathcal{E}_0^{\otimes|A|}, \qquad \mathcal{E} : S(\mathcal{H}_{\text{in}}^{\otimes|A|}) \to S(\mathcal{H}_{\text{out}}^{\otimes|A|}),$$

where we use $\bigotimes_{a\in A}\mathcal{E}_0$ to indicate that the different tensor copies are labeled by elements of A. For each $g \in G$ the action of the representation on $\mathcal{H}^{\otimes|A|}$ is defined by

$$U(g)\,|\phi_{a_1}\rangle\,|\phi_{a_2}\rangle\cdots|\phi_{a_{|A|}}\rangle = |\phi_{g^{-1}a_1}\rangle\,|\phi_{g^{-1}a_2}\rangle\cdots|\phi_{g^{-1}a_{|A|}}\rangle.$$

Here $a_1 \cdots a_{|A|}$ is a list of the elements of A. The covariance of $\mathcal{E}$ follows from the definition, and the error-correction properties of $\mathcal{E}$ are directly inherited from those of $\mathcal{E}_0$. Therefore, we succeed in finding a perfect G-covariant channel. Figure 2 shows an example in which $G = S_3$ (the permutation group on three elements) and $A = \{1, 2, 3\}$.

While our construction can be formally extended to infinite groups with their associated infinite-dimensional representations, we have not determined which additional conditions need to be imposed in order for the argument to remain mathematically rigorous.

The construction presented in this section provides codes in which the Hilbert spaces can be exponentially large in $|G|$. However, it is known that in many cases random codes give near optimal error-correcting schemes with good parameters [28–32]. In Appendix B, we show

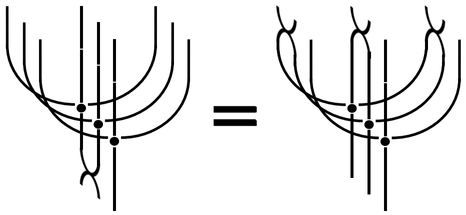

FIG. 2. Permutation covariance for the group S_3 acting on S_3 (i.e., $G = A = S_3$). Each fork represents a code that maps one qudit into three, and can correct an erasure error on any one output qudit. $\pi_{12} \in G$ is the transposition that swaps systems 1 and 2. Left: the map $\mathcal{E}[U_{\text{in}}(\pi_{12})\rho_{\text{in}}U_{\text{in}}(\pi_{12})^\dagger]$. The group action permutes the inputs to the channel. Right: the map $U_{\text{out}}(\pi_{12})\mathcal{E}(\rho_{\text{in}})U_{\text{out}}(\pi_{12})^\dagger$. As it is evident from the wiring of the forks, these two maps are equivalent.

that choosing a random, covariant isometry yields approximate error-correcting codes for which the dimension of each mode is $|G|$. For these codes, the worst-case fidelity of recovery, F_{worst}, behaves well with high probability. Specifically, $P(F_{\text{worst}} < 1 - \epsilon)$ decays exponentially in $|G|$. For example, we show in the Appendix E

$$P\left(F_{\text{worst}} < 1 - |G|^{\frac{9-2n}{8}}\right) \leq \exp\left\{-\frac{|G|^2}{216}\left[|G|^{\frac{2n-8}{4}} - 432\log\left(30|G|^{\frac{7+2n}{8}}\right)\right]\right\}. \tag{7}$$

It is clear that for $n \geq 5$ and $|G|$ sufficiently large, the exponent on the right-hand side becomes arbitrarily negative, indicating that the worst-case fidelity of recovery is close to 1 with high probability.

VIII. DISCUSSION

We showed that perfect error correction of physical information against erasure is a process that depends on the details of the symmetry group and dimensions of the code. For example, covariant quantum error correction is impossible when the symmetry group is a Lie group and the code is finite dimensional. This is connected to the following *no-go* theorems in the literature:

(a) *Eastin-Knill theorem* [33]. Eastin and Knill proved [20] that it is not possible to encode information in an error-detecting code such that a set of universal gates can be implemented transversally. We can reproduce the main thrust of the Eastin-Knill theorem [34] using an instance of our no-go theorem in which the input space is the set of N logical qudits, the output consists of physical qudits, and letting the group be $G = U(N)$. Moreover, our continuous-variable code construction provides a demonstration that *the Eastin-Knill theorem can be circumvented in principle*, although our explicit examples do not immediately appear to be useful for fault-tolerant quantum computation.

(b) *Invariant perfect tensors*. A quantum state on the tensor product of a number of Hilbert spaces is a *perfect tensor* if, for any bipartition of the Hilbert space into two collections of constituent factors, it forms an isometry from the smaller space to the larger [35]. Motivated by the construction of physical states in the Hilbert space of loop quantum gravity, the authors in Ref. [21] defined the notion of *invariant perfect tensors* as those perfect tensors that are invariant with respect to the action of SU(2). The authors proved that there are no invariant perfect tensors with four tensor factors. This can be seen as a direct consequence of our no-go theorem for $G = \text{SU}(2)$, by considering a four-partite invariant perfect tensor as a one-mode to three-mode isometry. Such an invariant perfect tensor with four tensor factors would define an SU(2)-covariant erasure-correcting code, which is prohibited by our no-go theorem. Furthermore, our no-go theorem states that there are no invariant perfect tensors with higher numbers of tensor factors, thereby solving an open question in Ref. [21].

One might hope to find a more quantitative relation between some measure of the size of the group and the dimension of the code when error correction is possible. For example, a condition of the form $|G| \leq \dim(\text{code})$ (i.e., dimension of the physical Hilbert space) is consistent with our no-go theorem and the examples in cases 1 and 2.

ACKNOWLEDGMENTS

We thank Dawei Ding, Iman Marvian, Michael Walter, and Beni Yoshida for helpful discussions. S.N. acknowledges support from Stanford Graduate Fellowship. G.S. acknowledges support from a NSERC postgraduate scholarship. This work is supported by the CIFAR and the Simons Foundation. The work of G.S. was performed before joining Amazon Web Services.

Note added.—Recently, there has been follow-up work on covariant quantum error correction (e.g., Refs. [36–40]). In Ref. [36], approximate versions of the no-go theorem above and of the Eastin-Knill theorem were developed, in which the recovery fidelity is bounded above by a function of the symmetry group G and the code dimension. These approximate theorems address our earlier questions about a quantitative relation between group size and code dimension. Moreover, several of the papers listed above attempt to address the question of practical fault-tolerant quantum computation, either using approximate quantum error correction, or approximate computation.

APPENDIX A: *G* = U(1) AND THE CODE IS CONTINUOUS VARIABLE

Here we provide an explicit U(1)-covariant $1 \to 3$ encoding. The construction presented in this section does not violate the no-go theorem stated in *case 1* above as the local systems are infinite dimensional. Since the symmetry group in question is U(1), this code could be implemented in optical modes, and it is arguably more natural than the construction presented in *case 2*.

We take the Hilbert space to be the space of functions on a circle using the position basis $\{|\phi\rangle\}_{\phi\in[0,2\pi)}$. U(1) acts on this space via the regular representation: if $g = e^{i\theta} \in U(1)$, then the action of the regular representation is defined by $U(g)\,|\alpha\rangle = |\alpha+\theta\rangle$. It is convenient to work in the Fourier basis where the Hilbert space is described by the conjugate momentum basis $\{|n\rangle\}_{n\in\mathbb{Z}}$ and the group acts by $U(g)\,|n\rangle = e^{in\theta}\,|n\rangle$. We define the isometry to be the following operator expressed in the conjugate momentum basis

$$E_{\mathrm{U}(1)} = \sum_{x,y\in\mathbb{Z}} |-3y, -x+y, 2(y+x)\rangle_{123}\,\langle x|_{\mathrm{in}}.$$

More explicitly, the isometry maps the state $\sum_x \phi(x)\,|x\rangle_{\mathrm{in}}$ to $|\Psi\rangle_{123} = \sum_{x,y}\phi(x)\,|-3y,-x+y,2(y+x)\rangle_{123}$. It is easy to see that this isometry is U(1) covariant:

$$\begin{aligned} U(g)^{\otimes 3}E_{\mathrm{U}(1)}U(g)^{\dagger} &= e^{i[-3y-x+y+2(y+x)]}E_{\mathrm{U}(1)}e^{-ix} \\ &= E_{\mathrm{U}(1)}. \end{aligned}$$

Here we show, step by step, that this mapping can correct an erasure error. Consider the encoded density matrix

$$\begin{aligned}\Psi_{123} = &\sum_{x_1,y_1,x_2,y_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|-3y_1,-x_1+y_1,2(y_1+x_1)\rangle \\ &\langle -3y_2,-x_2+y_2,2(y_2+x_2)|_{123}\,.\end{aligned}$$

We study the loss of modes 1, 2, and 3, in turn.

1. *Loss of the first mode*. The resulting density matrix is

   $$\begin{aligned}\Psi_{23} = &\sum_{x_1,x_2,y\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|-x_1+y,2(y+x_1)\rangle \\ &\langle -x_2+y,2(y+x_2)|_{23}\,.\end{aligned}$$

   Decoding starts with the linear map $|a,b\rangle \to |a,b-2a\rangle$, yielding

   $$\sum_{x_1,x_2,y\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|x_1+y,4x_1\rangle\,\langle x_2+y,4x_2|_{23}.$$

   We then use an isometry, which maps the states of the form $|a,4b\rangle$ to $|a,b\rangle$

   $$\sum_{x_1,x_2,y\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|x_1+y,x_1\rangle\,\langle x_2+y,x_2|_{23}.$$

   Finally, by $|a,b\rangle \to |a-b,b\rangle$, we obtain

   $$\sum_{x_1,x_2,y\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|y,x_1\rangle\,\langle y,x_2|_{23}.$$

   Therefore, tracing out mode 2 reveals the original state.
2. *Loss of the second mode*. The resulting density matrix is

   $$\begin{aligned}\Psi_{13} = &\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|-3y,2(y+x_1)\rangle \\ &\langle -3(-x_1+y+x_2),2(-x_1+y+2x_2)|_{13}\,,\end{aligned}$$

   or, equivalently by the change of variable $y \to y + x_1$,

   $$\begin{aligned}\Psi_{13} = &\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|-3(y+x_1),2(y+2x_1)\rangle \\ &\langle -3(y+x_2),2(y+2x_2)|_{13}\,.\end{aligned}$$

   We now use an isometry, which maps states of the form $|3a,2b\rangle$ to $|a,b\rangle$

   $$\begin{aligned}&\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|-(y+x_1),(y+2x_1)\rangle \\ &\langle -(y+x_2),(y+2x_2)|_{13}\,.\end{aligned}$$

   By $|a,b\rangle \to |a,2a+b\rangle$, we have

   $$\begin{aligned}&\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \\ &|-(y+x_1),-y\rangle\,\langle -(y+x_2),y|_{13}\,.\end{aligned}$$

   We now use $|a,b\rangle \to |-(a+b),b\rangle$ to obtain

   $$\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^*\,|x_1,-y\rangle\,\langle x_2,y|_{13}.$$

   Tracing out mode 3 reveals the original state.

3. *Loss of the third mode*. Again, the resulting density matrix is

$$\Psi_{12} = \sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \left|-3y,-x_1+y\right\rangle \left\langle -3(y+x_1-x_2), -2x_2+y+x_1\right|_{12}.$$

Using the change of variable $y \to y + x_1$ we have

$$\Psi_{12} = \sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \left|-3(y-x_1),-2x_1+y\right\rangle \left\langle -3(y-x_2), -2x_2+y\right|_{12}.$$

Applying an isometry that maps $|3a,b\rangle$ to $|a,b\rangle$ yields

$$\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \left|-(y-x_1),-2x_1+y\right\rangle \left\langle -(y-x_2), -2x_2+y\right|_{12}.$$

Using $|a,b\rangle \to |a,a+b\rangle$,

$$\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \left|-(y-x_1),-x_1\right\rangle \left\langle -(y-x_2), -x_2\right|_{12}.$$

Finally, the isometry $|a,b\rangle \to |a+b,-a\rangle$ turns the state to

$$\sum_{x_1,y,x_2\in\mathbb{Z}} \phi(x_1)\phi(x_2)^* \left|-y,x_1\right\rangle \left\langle -y,x_2\right|_{12}.$$

Thus we can recover the state on mode 2.

APPENDIX B: *G* IS A FINITE GROUP AND THE CODE IS A RANDOM *G*-COVARIANT ISOMETRY

In the construction presented for *case 3*, the local Hilbert space dimension can grow exponentially with $|G|$. In this section we present an alternative, approximate method for error correction in which the local Hilbert space dimensions are equal to $|G|$. Our goal is to prove Eq. (7) of the main text.

Consider a $1 \to n$ encoding. We look for isometries that map $\mathcal{H}_G \to \mathcal{H}_G^{\otimes n}$, where $\mathcal{H}_G$ denotes the Hilbert space associated to the regular representation of G with the basis $\{|g\rangle\}_{g\in G}$. Thus $\dim \mathcal{H}_G = |G| = d$. We represent the action of the regular representation of $g \in G$ on $\mathcal{H}_G$ by $U(g)$.

To construct a random covariant map, we start with a random invariant state $|\Psi\rangle \in \mathcal{H}_G^{\otimes(n+1)}$. For our purposes, a random state is one that is chosen randomly with respect to the unitary invariant measure; random unitaries are unitaries chosen randomly with respect to the Haar measure; and a state is invariant if $U(g)^{\otimes(n+1)}|\Psi\rangle = |\Psi\rangle$ for all $g \in G$. By projecting our chosen state onto an un-normalized, maximally entangled state $|\phi^+\rangle_{AB} = \sum |i\rangle_A |i\rangle_B$ we obtain a map E (which is close to an isometry with high probability) from $\mathcal{H}_{\text{in}} \to \mathcal{H}^{\otimes n}$,

$$E_{\text{in},1\cdots n} = \sqrt{d}\,\langle\phi^+|_{\text{in},0}\,|\Psi\rangle_{0\cdots n}.$$

Note that the covariance of E defined by $U(g)^{\otimes n}E = EU(g)$, which follows from the invariance of $|\Psi\rangle$. From E we can define the exact isometry T as

$$T := E(E^\dagger E)^{-1/2}.$$

One can verify that T is also a covariant map, since $[E^\dagger E, U(g)]$ for all $g \in G$. Our encoding is then defined by

$$\mathcal{E}(\rho_{\text{in}}) = T\rho_{\text{in}}T^\dagger.$$

With the covariant encoding in hand, we now turn our attention to the decoding. Before diving in, let us first define two notational conventions that are used frequently henceforth. Firstly, we use $\text{tr}_{\hat{x}}$ to indicate tracing out all subsystems *except* the set x. Secondly, if there are two isomorphic Hilbert spaces $\mathcal{H}_\alpha$ and $\mathcal{H}_\beta$ with the same preferred basis, and if the operator X_α acts on $\mathcal{H}_\alpha$, then by $(X_\alpha)_\beta$ we mean the operator X_α acting on $\mathcal{H}_\beta$ (in the sense that the matrix corresponding to X_α is simply applied to $\mathcal{H}_\beta$). One can think of $(X_\alpha)_\beta$ as overriding the Hilbert space indices. When it is clear to do so, we use X_β instead of $(X_\alpha)_\beta$ for brevity.

To decode after loss of one of the modes, say mode 1 without loss of generality, Bob first replaces the lost mode by a maximally mixed state τ_1 and then decodes the state $\tau_1 \otimes \text{tr}_1[\mathcal{E}(\rho_{\text{in}})]$. The decoding map is given by

$$\begin{aligned}\sigma_{\text{out}} &= \mathcal{D}_1(\rho_{12\cdots n}) \\ &= \{\text{tr}_{\hat{2}}\left[(U_{23}^T V_{23\cdots n})\rho_{12\cdots n}(U_{23}^T V_{23\cdots n})^\dagger\right]\}_{\text{out}},\end{aligned}$$

where U_{01}, and $V_{23\cdots n}$ are unitaries that transform $|\Psi\rangle_{0\cdots n}$ into its Schmidt form:

$$U_{01}\otimes V_{2\cdots n}\,|\Psi\rangle_{0\cdots n} = \sum_{i,j}\sqrt{\lambda_{ij}}\,|ij\rangle_{01}\otimes|ij\,0\cdots0\rangle_{23\cdots n},$$

and $U_{23} = (U_{01})_{23}$ is the same operator as U_{01} but acting on the Hilbert spaces indexed by 2 and 3. In other words, $U_{01} = (U_{23})_{01}$.

With the decoding above, our task is now to prove Eq. (7) of the main text. Our first step is bounding the worst-case fidelity of recovery F_{worst} in terms of the distance between Ψ_{01} (the reduced density matrix of the invariant state $|\Psi\rangle$) and the maximally mixed state.

Lemma 1. *For* $0 \le \epsilon \le 1$, *if* $\|\Psi_{01} - \tau_{01}\|_\infty \le (\epsilon/3d^2)$, *then* $1 - \epsilon \le F_{\text{worst}}$.

Proof. We prove this in three steps.

(a) Step 1. We first simplify the expression for the recovered state and show that

$$\begin{aligned}&\mathcal{D}_1[\tau_1 \otimes \mathcal{E}(\rho_{\mathrm{in}})]\\ &\quad = \mathrm{tr}_1\left[\Psi_{01}^{T\ 1/2}\Psi_0^{T-1/2}(\rho_{\mathrm{in}})_0\Psi_0^{T-1/2}\Psi_{01}^{T\ 1/2}\right].\end{aligned}$$

(b) Step 2. We then use joint concavity of the fidelity, and properties of the Schatten norm to bound the worst-case fidelity

$$F_{\mathrm{worst}} \geq \min_{|\kappa\rangle}\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right|.$$

From the above equation, it is already clear that if Ψ_0 and Ψ_{01} are close to the maximally mixed state, then the worst-case fidelity will be close to 1. We quantify this in the last step.

(c) Step 3. We show that for $0 \leq \epsilon \leq 1$, if $\|\Psi_{01} - \tau_{01}\|_\infty \leq (\epsilon/3d^2)$, then $1-\epsilon \leq F_{\mathrm{worst}}$.

■

APPENDIX C: STEP 1

We begin with the expression for the recovered state,

$$\mathcal{D}_1[\tau_1 \otimes \mathcal{E}(\rho_{\mathrm{in}})] = \mathrm{tr}_{\hat{2}}\left[U_{23}^T V_{23\cdots n}\,\mathrm{tr}_1(T\rho_{\mathrm{in}}T^\dagger)V_{23\cdots n}^\dagger U_{23}^*\right]. \tag{C1}$$

Using the fact that $E^\dagger E = d(\Psi_0^T)_{\mathrm{in}}$, and the definition $\tilde{\rho}_{\mathrm{in}} = 1/d\left(\Psi_0^{T-1/2}\right)_{\mathrm{in}}\rho_{\mathrm{in}}\left(\Psi_0^{T-1/2}\right)_{\mathrm{in}}$, we have that $T\rho_{\mathrm{in}}T^\dagger = E\tilde{\rho}_{\mathrm{in}}E^\dagger$. From the definition of E we can simplify the formula for the encoding map:

$$\mathcal{E}(\rho_{\mathrm{in}}) = E\tilde{\rho}_{\mathrm{in}}E^\dagger = d\,\mathrm{tr}_0\left(|\Psi\rangle\langle\Psi|_{0\cdots n}\,\tilde{\rho}_0^T\right).$$

Therefore,

$$\mathcal{D}_1[\tau_1 \otimes \mathcal{E}(\rho_{\mathrm{in}})] = d\,\mathrm{tr}_{\hat{2}}\left(U_{23}^T V_{2\cdots n}|\Psi\rangle\langle\Psi|_{0\cdots n}V_{2\cdots n}^\dagger U_{23}^*\tilde{\rho}_0^T\right). \tag{C2}$$

However, recall that $U_{01} \otimes V_{2\cdots n}|\Psi\rangle_{0\cdots n} = \sum_{i,j}\sqrt{\lambda_{ij}}\,|ij\rangle_{01} \otimes |ij\,0\cdots 0\rangle_{23\cdots n}$, and that $U_{01}\Psi_{01}^{1/2}U_{01}^\dagger = \sum_{ij}\sqrt{\lambda_{ij}}\,|ij\rangle\langle ij|$. Thus we obtain

$$V_{2\cdots n}|\Psi\rangle_{0\cdots n} = \Psi_{01}^{1/2}U_{01}^\dagger\,|\phi\rangle_{02}^+|\phi\rangle_{13}^+|0\cdots 0\rangle_{4\cdots n} = U_{23}^*\left(\Psi_{01}^{T\ 1/2}\right)_{23}|\phi^+\rangle_{02}|\phi_{13}^+\rangle|0\cdots 0\rangle_{4\cdots n}.$$

Using Eq. (C2), we find

$$\begin{aligned}&\mathcal{D}_1[\tau_1 \otimes \mathcal{E}(\rho_{\mathrm{in}})]\\ &= d\,\mathrm{tr}_{\hat{2}}\left[\left(\Psi_{01}^{1/2^T}\right)_{23}|\phi^+\rangle_{02}|\phi^+\rangle_{13}\langle\phi^+|_{02}\langle\phi^+|_{13}\left(\Psi_{01}^{1/2^T}\right)_{23}\tilde{\rho}_0^T\right]\\ &= d\,\mathrm{tr}_3\left[\left(\Psi_{01}^{1/2^T}\right)_{23}(\tilde{\rho}_0)_2\left(\Psi_{01}^{1/2^T}\right)_{23}\right]\\ &= \mathrm{tr}_1\left[\Psi_{01}^{T\ 1/2}\Psi_0^{T-1/2}\rho_0\Psi_0^{T-1/2}\Psi_{01}^{T\ 1/2}\right].\end{aligned}$$

Therefore, we have achieved goal of step 1.

APPENDIX D: STEP 2

Our goal now is to lower bound the fidelity of recovery. Since the fidelity is jointly concave, we know that the minimum fidelity of recovery for the channel is achieved with a pure input state, say $\rho_0 = (|\kappa\rangle\langle\kappa|)^T$, where we add the transpose to simplify the expressions. In this case, the recovered state takes the following form:

$$\mathcal{D}_1[\tau_1 \otimes \mathcal{E}(\rho_{\text{in}})] = \text{tr}_1 \left(\Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle_0 \langle\kappa|_0 \Psi_0^{-1/2} \Psi_{01}^{1/2} \right)^T,$$

so that the minimum fidelity is

$$F_{\min} = \min_{|\kappa\rangle} \sqrt{\text{tr}\left(\langle\kappa|_0 \Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle_0 \langle\kappa|_0 \Psi_0^{-1/2} \Psi_{01}^{1/2} |\kappa\rangle_0 \right)} = \min_{|\kappa\rangle} \left(\left\| \langle\kappa|_0 \Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle_0 \right\|_2 \right).$$

To proceed, we use the following basic property of the Schatten norm: for $(1/p) + (1/q) = 1$, $\|Y\|_p \geq |\text{tr}(XY^\dagger)|$ if $\|X\|_q = 1$. Applying this inequality when $X = I_1/\sqrt{d}$ and $p = q = 2$ we find

$$\begin{aligned} \left\| \langle\kappa| \Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle \right\|_2 &= \max \left\{ \left| \text{tr}\left(X_1 \langle\kappa|_0 \Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle_0 \right) \right| \; \middle| \; \|X\|_2 = 1 \right\} \\ &\geq \frac{1}{\sqrt{d}} \left| \text{tr}\left(\langle\kappa|_0 \Psi_{01}^{1/2} \Psi_0^{-1/2} |\kappa\rangle_0 \right) \right| \\ &= \left| \langle\kappa|_0 \, \text{tr}_1 \left(\Psi_{01}^{1/2} \right) \left(\frac{\Psi_0^{-1/2}}{\sqrt{d}} \right) |\kappa\rangle_0 \right|. \end{aligned} \tag{D1}$$

This concludes step 2.

APPENDIX E: STEP 3

We ultimately want to lower bound the worst-case fidelity using concentration of measure techniques for Ψ_{01} and Ψ_0. We start by upper bounding $\left\| \text{tr}_1 \left(\Psi_{01}^{1/2} \right) - I_0 \right\|_\infty$ and $\left\| \left(\Psi_0^{-1/2}/\sqrt{d} \right) - I_0 \right\|_\infty$, assuming that $\|\Psi_0 - \tau_0\|_\infty \leq (1/2d)$.

1. Upper bound for $\left\| \text{tr}_1 \left(\Psi_{01}^{1/2} \right) - I_0 \right\|_\infty$:

$$\begin{aligned} \left\| \text{tr}_1 \left(\Psi_{01}^{1/2} \right) - I_0 \right\|_\infty &= \left\| \text{tr}_1 \left(\Psi_{01}^{1/2} - \frac{I_{01}}{d} \right) \right\|_\infty = \max_{|\alpha\rangle} \left| \langle\alpha|_0 \, \text{tr}_1 \left(\Psi_{01}^{1/2} - \frac{I_{01}}{d} \right) |\alpha\rangle_0 \right| \\ &\leq \max_{|\alpha\rangle} \sum_{g \in G} \left| \langle\alpha|_0 \langle g|_1 \, \text{tr}_1 \left(\Psi_{01}^{1/2} - \frac{I_{01}}{d} \right) |\alpha\rangle_0 |g\rangle_1 \right| \\ &\leq d \left\| \Psi_{01}^{1/2} - \frac{I_{01}}{d} \right\|_\infty, \end{aligned}$$

where $|g\rangle$, $g \in G$ form a basis for evaluating the trace, the first inequality is the triangle inequality, and the second inequality comes from the fact that the infinite Schatten norm of a Hermitian operator is equal to its maximum eigenvalue. Now, one can check that for any $\lambda \geq 0$, $|\lambda^{1/2} - 1/d| \leq d|\lambda - 1/d^2|$. Taking $\{\lambda_i\}$ to be the set of eigenvalues of $\Psi_{01}^{1/2}$, and using the aforementioned inequality, we obtain

$$\left\| \Psi_{01}^{1/2} - \frac{I_{01}}{d} \right\|_\infty = \max_i \left| \lambda_i^{1/2} - \frac{1}{d} \right| \leq d \max_i \left| \lambda_i - \frac{1}{d^2} \right| \leq d \, \|\Psi_{01} - \tau_{01}\|_\infty. \tag{E1}$$

Thus

$$\left\| \text{tr}_1 \left(\Psi_{01}^{1/2} \right) - I_0 \right\|_\infty \leq d^2 \, \|\Psi_{01} - \tau_{01}\|_\infty.$$

2. Upper bound for $\left\| \left(\Psi_0^{-1/2}/\sqrt{d} \right) - I_0 \right\|_\infty$:one can simply check that for any real number λ such that $|\lambda - 1/d| \leq 1/2d$, then $\left| \lambda^{-1/2}/\sqrt{d} - 1 \right| \leq d \, |\lambda - 1/d|$. In particular, since this inequality holds for all of the eigenvalues of Ψ_0, we can derive the following bound for the operator norm:

$$\left\| \left(\frac{\Psi_0^{-1/2}}{\sqrt{d}} \right) - I_0 \right\|_\infty \leq d \, \|\Psi_0 - \tau_0\|_\infty. \tag{E2}$$

To proceed, we *assume* that $\|\Psi_{01}-\tau_{01}\|_\infty \le (\epsilon/3d^2)$ for $0\le\epsilon\le 1$. Combining this assumption with Eq. (D1), we have

$$\begin{aligned}&\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right|\\ &=\left|1+\langle\kappa|_0\left[\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right]|\kappa\rangle_0+\langle\kappa|_0\left[\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right]|\kappa\rangle_0+\langle\kappa|_0\left[\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right]\left[\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right]|\kappa\rangle_0\right|\\ &\ge 1-\left|\langle\kappa|_0\left[\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right]|\kappa\rangle_0\right|-\left|\langle\kappa|_0\left[\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right]|\kappa\rangle_0\right|-\left|\langle\kappa|_0\left[\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right]\left[\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right]|\kappa\rangle_0\right|,\end{aligned}$$

where the inequality in the last line is the triangle inequality. Now, since $\langle\kappa|X|\kappa\rangle\le\|X\|_\infty$ for any matrix X, we have

$$\begin{aligned}\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right| &\ge 1-\left\|\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right\|_\infty-\left\|\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right\|_\infty-\left\|\left[\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right]\left[\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right]\right\|_\infty\\ &\ge 1-\left\|\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right\|_\infty-\left\|\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)-I_0\right\|_\infty-\left\|\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)-I_0\right\|_\infty\left\|\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)\right.\\ &\quad\left.-I_0\right\|_\infty,\end{aligned}$$

where the second inequality follows from the fact that $\|XY\|_\infty\le\|X\|_\infty\|Y\|_\infty$ for any pair of matrices X and Y. Using Eqs. (E1) and (E2) above,

$$\begin{aligned}\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right| &\ge 1-d^2\|\Psi_{01}-\tau_{01}\|_\infty-d\|\Psi_0-\tau_0\|_\infty\\ &\quad-\left(d^2\|\Psi_{01}-\tau_{01}\|_\infty\right)\left(d\|\Psi_0-\tau_0\|_\infty\right).\end{aligned}$$

Note that the condition $\|\Psi_0-\tau_0\|_\infty\le(1/2d)$ is satisfied, since $\|\Psi_0-\tau_0\|_\infty\le d\|\Psi_{01}-\tau_{01}\|_\infty$ and $\|\Psi_{01}-\tau_{01}\|_\infty\le(\epsilon/3d^2)$. Finally, since $\|\Psi_0-\tau_0\|_\infty\le d\|\Psi_{01}-\tau_{01}\|_\infty$, we have that

$$\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right|\ge 1-2d^2\|\Psi_{01}-\tau_{01}\|_\infty-\left(d^2\|\Psi_{01}-\tau_{01}\|_\infty\right)^2,$$

and we therefore conclude that

$$F_{\text{worst}}\ge\left|\langle\kappa|_0\,\mathrm{tr}_1\left(\Psi_{01}^{1/2}\right)\left(\frac{\Psi_0^{-1/2}}{\sqrt{d}}\right)|\kappa\rangle_0\right|\ge 1-2d^2\|\Psi_{01}-\tau_{01}\|_\infty-\left(d^2\|\Psi_{01}-\tau_{01}\|_\infty\right)^2\ge 1-\epsilon,$$

which proves the lemma.

To complete the proof, it remains to be shown that our assumption is valid. Specifically, in order to show that the worst-case fidelity is close to 1, it suffices to prove that the reduced density matrix of random invariant states, Ψ_{01}, is very close to the maximally mixed state in operator norm (i.e., $\|\Psi_{01}-\tau_{01}\|_\infty$ is small) with high probability. Since

$$\|\Psi_{01}-\tau_{01}\|_\infty=\max_{\sigma_{01}}|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]|,$$

where the maximization is done over all possible density matrices σ, we can instead study the quantity on the right-hand side. To show that this is small, we follow the techniques used in Refs. [41–44].

Before stating the proof in its full glory, let us first gain an imprecise, high-level overview of the strategy. We first define an ϵ-net on the set of density matrices on $\mathcal{H}_0\otimes\mathcal{H}_1$, i.e., a finite set of density matrices $\tilde{\sigma}_{01}$ such that any other density matrix σ_{01} is close to one of the elements of the net in the trace norm. If we can then show that $|\mathrm{tr}\,[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]|$ is small for every $\tilde{\sigma}$ in the net, then it must be small for *all* density matrices σ_{01}. Using large deviation methods, we then prove that for any fixed density matrix σ_{01} (including the elements of the net), $|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]|$ is small with very high probability. Since the number of elements in the net is finite (with a known upper bound), we can then use a union bound to show that $|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]|$ is small for all elements in the net with high probability. Therefore, we can bound $|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]|$, arriving at our desired conclusion.

We now give a detailed proof of Eq. (7) of the main text. Let $\mathcal{P}_{\delta,\sigma_{01}}$ be the probability that, for a fixed σ_{01}, $|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]|\geq\delta/d^2$, and let $\mathcal{P}_\delta=\max_{\sigma_{01}}\mathcal{P}_{\delta,\sigma_{01}}$, where the maximum is over all density matrices on $\mathcal{H}_0\otimes\mathcal{H}_1$. The following lemma relates $P\left[\|\Psi_{01}-\tau_{01}\|_\infty\leq(\epsilon/3d^2)\right]$ to $\mathcal{P}_\delta$.

Lemma 2. *For* $0\leq\alpha\leq\epsilon$, *we have*

$$P\left(\|\Psi_{01}-\tau_{01}\|_\infty\leq\frac{\epsilon}{3d^2}\right)\geq 1-\mathcal{P}_{\frac{\epsilon-\alpha}{3}}\left[\frac{15d^2}{\alpha}\right]^{2d^2}.$$

Proof. Consider an $\alpha/3d^2$-trace distance net $\mathcal{M}$ of pure states in $\mathcal{H}_0\otimes\mathcal{H}_1$, with $\alpha\leq\epsilon$. For every pure state σ_{01}, there exists a pure state $\tilde{\sigma}_{01}$ such that

$$\|\sigma_{01}-\tilde{\sigma}_{01}\|_1\leq\frac{\alpha}{3d^2}, \tag{E3}$$

by definition. It is known that we can choose $\mathcal{M}$ such that $|\mathcal{M}|\leq\left(15d^2/\alpha\right)^{2d^2}$ [43, Lemma II.4]. Now if $|\mathrm{tr}[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]|\leq(\epsilon-\alpha/3d^2)$, then from Eq. (E3) it follows that

$$\begin{aligned}\left|\mathrm{tr}[\sigma_{01}(\Psi_{01}-\tau_{01})]\right|&\leq\left|\mathrm{tr}[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]\right|+\left|\mathrm{tr}[(\sigma_{01}-\tilde{\sigma}_{01})(\Psi_{01}-\tau_{01})]\right|\\&\leq\frac{\epsilon-\alpha}{3d^2}+\|\sigma_{01}-\tilde{\sigma}_{01}\|_1\|\Psi_{01}-\tau_{01}\|_\infty\\&\leq\frac{\epsilon-\alpha}{3d^2}+\|\sigma_{01}-\tilde{\sigma}_{01}\|_1\\&\leq\frac{\epsilon}{3d^2}.\end{aligned}$$

Therefore,

$$\begin{aligned}P\left(\|\Psi_{01}-\tau_{01}\|_\infty\leq\frac{\epsilon}{3d^2}\right)&=P\left\{\forall\sigma_{01}:\left|\mathrm{tr}\,[\sigma_{01}(\Psi_{01}-\tau_{01})]\right|\leq\frac{\epsilon}{3d^2}\right\}\\&\geq P\left\{\forall\tilde{\sigma}_{01}\in\mathcal{M}:\left|\mathrm{tr}\,[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]\right|\leq\frac{\epsilon-\alpha}{3d^2}\right\}\\&=1-P\left\{\exists\tilde{\sigma}_{01}\in\mathcal{M}:\left|\mathrm{tr}\,[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]\right|\geq\frac{\epsilon-\alpha}{3d^2}\right\}.\end{aligned} \tag{E4}$$

We can simplify Eq. (E4) using a union bound:

$$P\left\{\exists\tilde{\sigma}_{01}\in\mathcal{M}:\left|\mathrm{tr}\,[\tilde{\sigma}_{01}(\Psi_{01}-\tau_{01})]\right|\geq\frac{\epsilon-\alpha}{3d^2}\right\}\leq\sum_{\tilde{\sigma}_{01}\in\mathcal{M}}\mathcal{P}_{\frac{\epsilon-\alpha}{3},\tilde{\sigma}_{01}}\leq\mathcal{P}_{\frac{\epsilon-\alpha}{3}}|\mathcal{M}|.$$

This, along with Eq. (E4), conclude the proof of the lemma. ■

In the next section, we use large deviation techniques to show that

$$\mathcal{P}_\delta \leq \exp\left(-d^{n-2}\delta^2/6\right) \quad \text{for} \quad 0 \leq \delta \leq 1. \tag{E5}$$

We defer the proof to the next section, but we use the result immediately. Combining Lemma 1, Lemma 2, and Eq. (E5) we have

$$\begin{aligned} P\left(F_{\text{worst}} \leq 1-\epsilon\right) &\leq P\left(\|\Psi_{01}-\tau_{01}\|_\infty \geq \frac{\epsilon}{3d^2}\right) \leq \min_{0\leq\alpha\leq\epsilon} \mathcal{P}_{\frac{\epsilon-\alpha}{3}} \times \left[\frac{15d^2}{\alpha}\right]^{2d^2} \\ &\leq \min_{0\leq\alpha\leq\epsilon} \exp\left[-d^{n-2}(\epsilon-\alpha)^2/54 + 2d^2\log\left(\frac{15d^2}{\alpha}\right)\right]. \end{aligned}$$

One convenient choice of ϵ and α is $\epsilon = d^{(9-2n/8)}$ and $\alpha = \epsilon/2$. With this choice we find

$$P\left(F_{\text{worst}} \geq 1-\epsilon\right) \leq \exp\left\{-\frac{d^2}{216}\left[d^{\frac{2n-8}{4}} - 432\log\left(30d^{\frac{7+2n}{8}}\right)\right]\right\},$$

which reduces to Eq. (7) of the main text after substituting $|G|$ for d.

APPENDIX F: PROOF OF EQ. (E5)

The goal of this appendix is to prove Eq. (E5). The discussion is split into two parts: we first explain the random invariant state construction, and then we prove the desired bound.

APPENDIX G: CONSTRUCTION OF RANDOM INVARIANT STATES

Consider the invariant subspace of $\mathcal{H}^{\otimes(n+1)}$—it is easy to see that the invariant subspace is spanned by states of the form

$$\frac{1}{\sqrt{d}}\sum_{g\in G}|gh_1, gh_2, \ldots, gh_n, g\rangle_{0\cdots n}.$$

We now introduce an isometry M from $\mathcal{H}^{\otimes n}$ to the invariant subspace of $\mathcal{H}^{\otimes(n+1)}$,

$$M = \frac{1}{\sqrt{d}}\sum_{g,h_1,\ldots,h_n}|gh_1, gh_2, \ldots, gh_n, g\rangle_{0\cdots n}\langle h_1, h_2, \ldots, h_n|_{0\cdots n-1}.$$

The projector onto the invariant subspace is defined as $\Pi_{0\cdots n} = MM^\dagger$. $\Pi_{0\cdots n}$ has the important property that, upon tracing out any one of the subsystems, it becomes the identity operator on the remaining subsystems. That is

$$\operatorname{tr}_i \Pi_{0\cdots n} = I_{0\cdots\hat{i}\cdots n}. \tag{G1}$$

A random invariant state $|\Psi\rangle_{0\cdots n}$ is constructed by choosing a random state $|\phi\rangle_{0\cdots n-1}$ in $\mathcal{H}^{\otimes n}$ from the unitary invariant measure, and then mapping $|\phi\rangle$ to $\mathcal{H}^{\otimes(n+1)}$ using the isometry M, $|\Psi\rangle_{0\cdots n} = M\,|\phi\rangle_{0\cdots n-1}$.

APPENDIX H: PROOF OF EQ. (E4)

To begin, we upper bound the moment generating function, $\mathbb{E}_\Psi \exp\left[t\operatorname{tr}\left(\sigma_{01}\Psi_{01}\right)\right]$, for an arbitrary density matrix σ_{01}, where $\Psi_{01} = \operatorname{tr}_{\hat{0}\hat{1}}\left(\Psi_{0\cdots n}\right)$ and the average is over random invariant states $|\Psi\rangle_{0\cdots n}$. Note that $\operatorname{tr}\left(\sigma_{01}\Psi_{01}\right) = \operatorname{tr}\left(\sigma_{01}\Psi_{0\cdots n}\right) =$

$\operatorname{tr}\left(\sigma_{01}M\phi_{0\cdots n-1}M^\dagger\right) = \operatorname{tr}\left(M^\dagger\sigma_{01}M\phi_{0\cdots n-1}\right)$. One can easily check that $M^\dagger\sigma_{01}M = \sigma_{01}^G \otimes I_{2\cdots n-1}$, where

$$\sigma_{01}^G = \frac{1}{d}\sum_{g,h_1,h_2,h_1',h_2'} |h_1,h_2\rangle\langle gh_1,gh_2|\,\sigma_{01}\,|gh_1',gh_2'\rangle\langle h_1',h_2'|\,.$$

One can also check that σ_{01}^G is a density matrix, specifically a version of σ_{01} symmetrized by the group G. Therefore,

$$\operatorname{tr}\left(\sigma_{01}\Psi_{0\cdots n}\right) = \langle\phi|\,\sigma_{01}^G\,|\phi\rangle\,,$$

where $|\phi\rangle = |\phi\rangle_{0\cdots n-1}$ is a state on $\mathcal{H}^{\otimes n}$ chosen from the unitary invariant measure (see the first subsection of this appendix).

We now choose a Gaussian state $|g\rangle_{0\cdots n-1}$ in which the coefficients of the wave function are chosen independent and identically distributed from a complex Gaussian distribution centered at zero with variance d^{-n}. Thus $\mathbb{E}_{|g\rangle}\|g\|_2^2 = 1$. Therefore, we have

$$\begin{aligned}\mathbb{E}_{|g\rangle}\exp\left(t\,\langle g|\,\sigma_{01}^G\,|g\rangle\right) &= \mathbb{E}_{|\phi\rangle}\mathbb{E}_{\|g\|_2}\exp\left(t\,\|g\|_2^2\,\langle\phi|\,\sigma_{01}^G\,|\phi\rangle\right)\\ &\geq \mathbb{E}_{|\phi\rangle}\exp\left[t\left(\mathbb{E}_{\|g\|_2}\|g\|_2^2\right)\langle\phi|\,\sigma_{01}^G\,|\phi\rangle\right]\\ &= \mathbb{E}_{|\phi\rangle}\exp\left(t\,\langle\phi|\,\sigma_{01}^G\,|\phi\rangle\right)\\ &= \mathbb{E}_{\Psi}\exp\left[t\operatorname{tr}\left(\sigma_{01}\Psi_{0\cdots n}\right)\right],\end{aligned}$$

where the inequality follows from the convexity of the exponential function.

Now suppose that the eigenvalues of σ_{01}^G are p_{i_0,i_1}. Since the Gaussian states are unitarily invariant, we can evaluate $\mathbb{E}_{|g\rangle}\exp\left(t\,\langle g|\,\sigma_{01}^G\,|g\rangle\right)$ in a basis in which $\sigma_{01}^G\otimes I_{2\cdots n-1}$ is diagonal. In that basis,

$$\mathbb{E}_{|g\rangle}\exp\left(t\,\langle g|\,\sigma_{01}^G\,|g\rangle\right) = \mathbb{E}_{|g\rangle}\exp\left(t\sum_{i_0\cdots i_{n-1}} p_{i_0,i_1}|g_{i_0\cdots i_{n-1}}|^2\right) = \prod_{i_0\cdots i_{n-1}}\mathbb{E}_{g_{i_0\cdots i_{n-1}}}\exp\left(t\,p_{i_0,i_1}|g_{i_0\cdots i_{n-1}}|^2\right).$$

However, the radial probability density for each coefficient is $p\left(|g_{i_0\cdots i_{n-1}}|\right) = 2d^n|g_{i_0\cdots i_{n-1}}|\exp\left(-d^n|g_{i_0\cdots i_{n-1}}|^2\right)$. Using the probability density formula, we find

$$\mathbb{E}_{g_{i_0\cdots i_{n-1}}}\exp\left(tp_{i_0,i_1}|g_{i_0\cdots i_{n-1}}|^2\right) = \frac{1}{1-\frac{t\,p_{i_0,i_1}}{d^n}}\quad\text{for}\quad t\leq d^n/p_{i_0,i_1}.$$

Assuming $t \leq d^n$, we have

$$\mathbb{E}_{|g\rangle}\exp\left(t\,\langle g|\,\sigma_{01}^G\,|g\rangle\right) = \prod_{i_0,i_1}\left(1-\frac{t\,p_{i_0,i_1}}{d^n}\right)^{-d^{n-2}}.$$

Ultimately, we fix the value of t to prove the bound in Eq. (E4), but we need to distinguish the cases in which t is positive or negative to bound the fluctuations of $\operatorname{tr}\left(\sigma_{01}\Psi_{01}\right)$ above or below $1/d^2$. Therefore, we discuss these two different ranges for t separately.

1. Positive t:
   We use the assumption that t is positive to limit the fluctuations of $\operatorname{tr}\left(\sigma_{01}\Psi_{01}\right)$ *above* $1/d^2$. Let $0 < s < 1$ be a fixed number, and restrict t to $0 \leq t \leq sd^n$. Under these conditions, we have,

$$\left(1-\frac{t\,p_{i_0,i_1}}{d^n}\right)^{-1} \leq \left(1+\frac{1}{1-s}\frac{t\,p_{i_0,i_1}}{d^n}\right).$$

   Therefore,

$$\left(1-\frac{t\,p_{i_0,i_1}}{d^n}\right)^{-d^{n-2}} \leq \left(1+\frac{1}{1-s}\frac{t\,p_{i_0,i_1}}{d^n}\right)^{d^{n-2}} \leq \exp\left(\frac{1}{1-s}t\,p_{i_0,i_1}d^{-2}\right).$$

Combining with Eq. (H1), we have

$$\mathbb{E}_{|g\rangle}\exp\left(t\langle g|\,\sigma_{01}^{G}\,|g\rangle\right)\leq\mathbb{E}_{|g\rangle}\exp\left(t\langle g|\,\sigma_{01}^{G}\,|g\rangle\right)=\prod_{i_0,i_1}\left(1-\frac{p_{i_0,i_1}t}{d^n}\right)^{-d^{n-2}}\leq\exp\left(\frac{1}{1-s}td^{-2}\right).\qquad\text{(H1)}$$

To bound the probabilities, we use Bernstein’s trick:

$$\begin{aligned}P\left[\operatorname{tr}(\sigma_{01}\Psi_{01})\geq\frac{1}{d^2}+\frac{\delta}{d^2}\right]&=P\left\{\exp\left[t\operatorname{tr}(\sigma_{01}\Psi_{01})\right]\geq\exp\left(t\frac{1+\delta}{d^2}\right)\right\}\\&\leq\left\{\mathbb{E}_{\Psi}\exp\left[t\operatorname{tr}(\sigma_{01}\Psi_{01})\right]\right\}\exp\left(-t\frac{1+\delta}{d^2}\right)\\&\leq\exp\left[-td^{-2}\left(1+\delta-\frac{1}{1-s}\right)\right],\end{aligned}$$

where we use Markov’s inequality for the exponentials and Eq. (H1). To obtain the best result, we now set $t=sd^n$ and $s=1-(1+\delta)^{-1/2}$. With this substitution,

$$P\left[\operatorname{tr}(\sigma_{01}\Psi_{01})\geq\frac{1}{d^2}+\frac{\delta}{d^2}\right]\leq\exp\left[-d^{n-2}(\sqrt{1+\delta}-1)^2\right]\leq\exp\left(-d^{n-2}\delta^2/6\right),$$

where the last inequality is valid for $0\leq\delta\leq1$.

2. Negative t:
We now use the constraint on t to limit the fluctuations of $\operatorname{tr}(\sigma_{01}\Psi_{01})$ *below* $1/d^2$. Assuming that $s>0$ and $-sd^n\leq t\leq0$, one can show that

$$\left(1-\frac{tp_{i_0,i_1}}{d^n}\right)^{-d^{n-2}}\leq\exp\left[t\frac{\log(1+s)}{s}p_{i_0,i_1}d^{-2}\right].$$

Therefore,

$$\begin{aligned}\mathbb{E}_{|g\rangle}\exp\left(t\langle g|\,\sigma_{01}^{G}\,|g\rangle\right)=\prod_{i_0,i_1}\left(1-\frac{p_{i_0,i_1}t}{d^n}\right)^{-d^{n-2}}&\leq\prod_{i_0,i_1}\exp\left[t\frac{\log(1+s)}{s}p_{i_0,i_1}d^{-2}\right]\\&\leq\exp\left[\frac{\log(1+s)}{s}td^{-2}\right].\end{aligned}$$

Thus,

$$\begin{aligned}P\left[\operatorname{tr}(\sigma_{01}\Psi_{01})\leq\frac{1}{d^2}-\frac{\delta}{d^2}\right]&=P\left[t\operatorname{tr}(\sigma_{01}\Psi_{01})\geq t\left(\frac{1}{d^2}-\frac{\delta}{d^2}\right)\right]\\&=P\left\{\exp\left[t\operatorname{tr}(\sigma_{01}\Psi_{01})\right]\geq\exp\left(t\frac{1-\delta}{d^2}\right)\right\}\\&\leq\left\{\mathbb{E}_{\Psi}\exp\left[t\operatorname{tr}(\sigma_{01}\Psi_{01})\right]\right\}\exp\left[-td^{-2}(1-\delta)\right]\\&\leq\exp\left\{-td^{-2}\left[1-\delta-\frac{\log(1+s)}{s}\right]\right\}.\end{aligned}$$

We now fix $t=-sd^n$ and $s=\delta/(1-\delta)$ to get

$$\begin{aligned}P\left[\operatorname{tr}(\sigma_{01}\Psi_{01})\leq\frac{1}{d^2}-\frac{\delta}{d^2}\right]&\leq\exp\left\{d^{n-2}\left[\delta+\log(1-\delta)\right]\right\}\\&\leq\exp\left(-d^{n-2}\delta^2/2\right)\leq\exp\left(-d^{n-2}\delta^2/6\right).\end{aligned}$$

This concludes the proof of Eq. (E5).

[1] N. Gisin and S. Popescu, Spin Flips and Quantum Information for Antiparallel Spins, Phys. Rev. Lett. **83**, 432 (1999).

[2] If the transmission time for each message is predetermined, that provides a resource that could itself be used for clock synchronization. To avoid this loophole, symbolic messages should not be received at predetermined times.

[3] J. Preskill, Quantum clock synchronization and quantum error correction, arXiv preprint quant-ph/0010098 (2000).

[4] S. Massar and S. Popescu, Optimal Extraction of Information from Finite Quantum Ensembles, Phys. Rev. Lett. **74**, 1259 (1995).

[5] E. Bagan, M. Baig, A. Brey, R. Munoz-Tapia, and R. Tarrach, Optimal encoding and decoding of a spin direction, Phys. Rev. A **63**, 052309 (2001).

[6] R. Jozsa, D. S. Abrams, J. P. Dowling, and C. P. Williams, Quantum Clock Synchronization Based on Shared Prior Entanglement, Phys. Rev. Lett. **85**, 2010 (2000).

[7] C. Souza, C. Borges, A. Khoury, J. Huguenin, L. Aolita, and S. Walborn, Quantum key distribution without a shared reference frame, Phys. Rev. A **77**, 032345 (2008).

[8] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, Degradation of a quantum reference frame, New J. Phys. **8**, 58 (2006).

[9] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Classical and Quantum Communication Without a Shared Reference Frame, Phys. Rev. Lett. **91**, 027901 (2003).

[10] A. Peres and P. F. Scudo, Entangled Quantum States as Direction Indicators, Phys. Rev. Lett. **86**, 4160 (2001).

[11] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reference frames, superselection rules, and quantum information, Rev. Mod. Phys. **79**, 555 (2007).

[12] G. Gour and R. W. Spekkens, The resource theory of quantum reference frames: Manipulations and monotones, New J. Phys. **10**, 033023 (2008).

[13] I. Marvian and R. W. Spekkens, Modes of asymmetry: The application of harmonic analysis to symmetric quantum dynamics and quantum reference frames, Phys. Rev. A **90**, 062110 (2014).

[14] I. Marvian and R. W. Spekkens, The theory of manipulations of pure state asymmetry: I. basic tools, equivalence classes and single copy transformations, New J. Phys. **15**, 033001 (2013).

[15] I. Marvian and R. W. Spekkens, Asymmetry properties of pure quantum states, Phys. Rev. A **90**, 014102 (2014).

[16] I. Marvian and R. W. Spekkens, Extending noether's theorem by quantifying the asymmetry of quantum states, Nat. Commun. **5**, 1 (2014).

[17] F. G. Brandao, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource Theory of Quantum States out of Thermal Equilibrium, Phys. Rev. Lett. **111**, 250404 (2013).

[18] V. Veitch, S. H. Mousavian, D. Gottesman, and J. Emerson, The resource theory of stabilizer quantum computation, New J. Phys. **16**, 013009 (2014).

[19] I. Devetak, A. W. Harrow, and A. J. Winter, A resource framework for quantum shannon theory, IEEE Trans. Inf. Theory **54**, 4587 (2008).

[20] B. Eastin and E. Knill, Restrictions on Transversal Encoded Quantum Gate Sets, Phys. Rev. Lett. **102**, 110502 (2009).

[21] Y. Li, M. Han, M. Grassl, and B. Zeng, Invariant perfect tensors, arXiv preprint arXiv:1612.04504 (2016).

[22] We exclude the case of zero-dimensional Lie groups. Also, G does not actually need to be a Lie group, but it must have at least one infinitesimal generator.

[23] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (2002).

[24] To be precise, each classical gyroscope determines one axis. In order to send a classical reference frame we need at least two gyroscopes for the x and y axes. By "gyroscope" we mean a complete indicator of the reference frame.

[25] Any reader disappointed by the construction's use of effectively classical gyroscopes should be heartened to know that the one-into-three encoding described in Appendix A achieves covariant error correction without them.

[26] Some readers might take issue with calling this an erasure code, since such codes are usually constructed such that the Hilbert spaces of each share are the same. However, one can use infinite-dimensional Hilbert spaces for each share and simply embed finite dimensional spaces such that the group acts on these subspaces according to the associated finite-dimensional representation and trivially on the rest.

[27] In fact, it is not necessary to assume that the basis is indexed by group elements—they can be indexed by any set on which the group acts faithfully.

[28] P. W. Shor, in *lecture notes, MSRI Workshop on Quantum Computation* (2002), https://www.msri.org/workshops/203/schedules/1181.

[29] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Trans. Inf. Theory **51**, 44 (2005).

[30] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A **55**, 1613 (1997).

[31] P. Hayden, M. Horodecki, A. Winter, and J. Yard, A decoupling approach to the quantum capacity, Open Syst. Inf. Dyn. **15**, 7 (2008).

[32] M. Hamada, Information rates achievable with algebraic codes on quantum discrete memoryless channels, IEEE Trans. Inf. Theory **51**, 4263 (2005).

[33] We thank Beni Yoshida for pointing out the connection to the Eastin-Knill theorem.

[34] The Eastin-Knill theorem also discusses the possibility of encoding information in the disconnected components of the Lie group, a point that is absent in our work. Furthermore, the full Eastin-Knill theorem makes reference to universal gates. In order to fully reproduce the theorem, we would need additional arguments concerning continuity of the channel and error detection.

[35] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence, arXiv preprint arXiv:1503.06237 (2015).

[36] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, Continuous symmetries and approximate quantum error correction, arXiv preprint arXiv:1902.07714 (2019).

[37] M. P. Woods and Á. M. Alhambra, Continuous groups of transversal gates for quantum error correcting codes from finite clock reference frames, arXiv preprint arXiv:1902.07725 (2019).

[38] S. Zhou, Z.-W. Liu, and L. Jiang, New perspectives on covariant quantum error correction, arXiv preprint arXiv:2005.11918 (2020).

[39] D.-S. Wang, G. Zhu, C. Okay, and R. Laflamme, Quasi-exact quantum computation, Phys. Rev. Res. **2**, 033116 (2020).

[40] Y. Yang, Y. Mo, J. M. Renes, G. Chiribella, and M. P. Woods, Covariant quantum error correcting codes via reference frames, arXiv preprint arXiv:2007.09154 (2020).

[41] P. Hayden, D. W. Leung, and A. Winter, Aspects of generic entanglement, Commun. Math. Phys. **265**, 95 (2006).

[42] A. Harrow, P. Hayden, and D. Leung, Superdense Coding of Quantum States, Phys. Rev. Lett. **92**, 187901 (2004).

[43] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, Commun. Math. Phys. **250**, 371 (2004).

[44] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, Remote preparation of quantum states, IEEE Trans. Inf. Theory **51**, 56 (2005).