Quantum error correction with the toric Gottesman-Kitaev-Preskill code

Christophe Vuillot,[1,*] Hamed Asasi,[2] Yang Wang,[1] Leonid P. Pryadko,[2] and Barbara M. Terhal[1,3]
[1]*QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*
[2]*Department of Physics, University of California, Riverside, California 92521, USA*
[3]*JARA Institute for Quantum Information, Forschungszentrum Juelich, D-52425 Juelich, Germany*

(Received 16 October 2018; revised manuscript received 3 December 2018; published 26 March 2019)

We examine the performance of the single-mode Gottesman-Kitaev-Preskill (GKP) code and its concatenation with the toric code for a noise model of Gaussian shifts, or displacement errors. We show how one can optimize the tracking of errors in repeated noisy error correction for the GKP code. We do this by examining the maximum-likelihood problem for this setting and its mapping onto a 1D Euclidean path-integral modeling a particle in a random cosine potential. We demonstrate the efficiency of a minimum-energy decoding strategy as a proxy for the path integral evaluation. In the second part of this paper, we analyze and numerically assess the concatenation of the GKP code with the toric code. When toric code measurements and GKP error correction measurements are perfect, we find that by using GKP error information the toric code threshold improves from 10% to 14%. When only the GKP error correction measurements are perfect we observe a threshold at 6%. In the more realistic setting when all error information is noisy, we show how to represent the maximum likelihood decoding problem for the toric-GKP code as a 3D compact QED model in the presence of a quenched random gauge field, an extension of the random-plaquette gauge model for the toric code. We present a decoder for this problem which shows the existence of a noise threshold at shift-error standard deviation $\sigma_0 \approx 0.243$ for toric code measurements, data errors and GKP ancilla errors. If the errors only come from having imperfect GKP states, then this corresponds to states with just four photons or more. Our last result is a no-go result for linear oscillator codes, encoding oscillators into oscillators. For the Gaussian displacement error model, we prove that encoding corresponds to squeezing the shift errors. This shows that linear oscillator codes are useless for quantum information protection against Gaussian shift errors.

DOI: 10.1103/PhysRevA.99.032344

I. INTRODUCTION

Within the framework of oscillator or continuous-variable (CV) error correcting codes, one can distinguish two classes of codes. One class generalizes qudit stabilizer codes to encode continuous degrees of freedom into a (larger) CV system [1,2]. We refer to these codes as *linear oscillator codes*. The other class, first introduced by Gottesman, Preskill, and Kitaev (GKP) in Ref. [3], and recently expanded to include many more codes [4,5], encodes a discrete (finite-dimensional) system into a CV system. Encoding and decoding for the first class of codes falls within the framework of Gaussian quantum information [6], while the second class of codes requires using non-Gaussian states.

In this paper we propose and analyze a scalable use of the GKP code [3] which encodes a single qubit into an oscillator. An example of such an oscillator is a mode in a high-Q microwave superconducting cavity coupled to superconducting qubits in a circuit-QED setup. Proposals for preparing a GKP code state in such systems exist [7]. The controlled-NOT (CNOT) gate between two GKP qubits requires about 4 dB of squeezing in both modes and a beam-splitter (see, e.g., Ref. [7]). Such a beam-splitter has been recently implemented between high-Q microwave cavity modes in Ref. [8]. Other possible physical implementations for the GKP code are the motional mode of a trapped-ion qubit [9] or atomic ensembles [10] for measurement-based CV cluster computation [11].

A bosonic code such as the GKP code or the recently implemented cat code [12] might be used to get a high-quality qubit, but the code does not provide a means to drive error rates down arbitrarily. A *scalable* fault-tolerant architecture can possibly be obtained by concatenating the GKP code with a qubit stabilizer code such as the toric or surface code. A theoretic goal is then to understand how to decode such a toric-GKP code and what is the error threshold of the architecture. Some results on using "analog" error information in concatenating the GKP code with a stabilizer code were obtained in Refs. [13–16]. A concatenation of the GKP code with the 2D surface code in the channel setting by message-passing (perfect) GKP error information to the surface code decoder was analyzed in Refs. [14,16]. In Ref. [16] the authors also consider the construction of the 3D cluster state using post-selection, with imperfect ancillas. However, these studies did not define, analyze nor decode the full fault-tolerant scenario for the 2D toric-GKP code.

Previous work has also studied the performance of the GKP code in comparison with other bosonic codes in a photon loss channel setting, not taking into account the imperfections or processing of repeated rounds of error correction [5]. Other work focused on the effect of photon loss and other sources of error on the preparation of code states [17]. Besides its good

*c.vuillot@tudelft.nl

performance compared to other bosonic codes, the GKP code is appealing since Clifford gates on the code states use only linear optical elements (including squeezing) [3].

In this work we first analyze repeated fault-tolerant quantum error correction for a single GKP qubit; see Sec. III. Our noise model in this analysis includes errors both on the GKP qubit as well as on the GKP ancilla qubit used in the error correction. We show how decoding this continuous error information in discrete time steps maps onto the evaluation of a stochastic discrete-time Euclidean path integral. We present an efficient minimum-energy decoder which chooses the path which approximately corresponds to a classical trajectory in a disordered potential.

Second, we consider the toric-GKP code in Sec. IV. Assuming that both GKP error correction and toric code correction are noiseless, we show how the use of continuous GKP error information improves the error correction for the toric code (Sec. IV B). These results are in correspondence with the previous results in Ref. [16] although our likelihood function is not identical to the one in Ref. [16].

In Sec. V we formulate the decoding problem of repeated quantum error correction with the toric-GKP code where both the GKP syndrome and the toric code syndrome contain errors. Since errors on the GKP qubits are intrinsic (getting perfect code states with infinite numbers of photons is unphysical), this is the physically relevant setting. The maximum-likelihood formulation is in terms of a 3D gauge field model with quenched randomness determined by the errors (Sec. V B). We discuss this model and its possible phase transitions in Sec. V C. Then in Sec. V D, we show how to re-express this model as a random plaquette gauge model (RPGM) with a $\mathbb{Z}_2$-field coupled to an auxiliary $U(1)$-gauge field. We then use this model to design a computationally efficient decoder and present numerical results.

Finally, in Sec. VI, we present our general no-go result for the first class of codes, namely the linear oscillator codes. This no-go result is presented as the calculation of the probability distribution of logical errors on the encoded information after perfect maximum-likelihood decoding. The result is in accordance with, but does not directly follow from, previous no-go results on Gaussian quantum information in Ref. [18]. The theorem explicitly shows that there are no linear oscillator code families of interest: there is no threshold in σ_0 below which protection of the encoded oscillators against shift errors gets better with increasing code size and the logical noise model is still Gaussian with the same σ_0, and possibly some squeezing of the logical quadratures.

The no-go result also shows that the existence of a threshold for the toric-GKP code is non-trivial. A sufficiently large departure from Gaussian quantum information is necessary to stabilize quantum information. In circuit-QED this departure comes exclusively from the use of the nonlinear Josephson junction element.

II. GENERAL CONSIDERATIONS

A. Definitions and notations

We consider n-mode oscillator codes, which are subspaces in the n-mode Hilbert space $L_2(\mathbb{R}^n)$. Such a Hilbert space can be constructed as a tensor product of n single-particle Hilbert spaces $L_2(\mathbb{R})$ of complex square-integrable functions. It supports n pairs of canonically conjugated coordinate and momentum operators, $\hat{p}_k$ and $\hat{q}_k$, such that $[\hat{q}_k, \hat{p}_l] = i\delta_{kl}$. These operators are used to define the multimode exponential shift operators,

$$\hat{U}(\boldsymbol{e}) \equiv \prod_{k=1}^{n} e^{iu_k\hat{p}_k + iv_k\hat{q}_k}, \quad \boldsymbol{e} \equiv (\boldsymbol{u}, \boldsymbol{v}), \tag{1}$$

where $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ are n-component real vectors. It is easy to check that the product of two such operators satisfies

$$\hat{U}(\boldsymbol{e})\hat{U}(\boldsymbol{e}') = \hat{U}(\boldsymbol{e} + \boldsymbol{e}')e^{i\omega(\boldsymbol{e},\boldsymbol{e}')}, \tag{2}$$

with the phase given by the symplectic product $\omega(\boldsymbol{e}, \boldsymbol{e}') = \boldsymbol{u}\cdot\boldsymbol{v}' - \boldsymbol{v}\cdot\boldsymbol{u}'$. The set $\mathcal{H}_n$ of all such operators with arbitrary phases is closed under multiplication, it forms an irreducible representation of the Heisenberg group $\mathbb{H}^n$ acting in $L_2(\mathbb{R}^n)$. Just as for the n-qubit Hilbert space and the Pauli group $\mathcal{P}_n$, any operator acting in $L_2(\mathbb{R}^n)$ can be represented as a linear combination of elements of $\mathcal{H}_n$. Furthermore, the product Eq. (2) of two exponential operators, up to a phase, can be represented in terms of the sum of the corresponding vectors, $\boldsymbol{e}'' = \boldsymbol{e} + \boldsymbol{e}'$. This map to $\mathbb{R}^{2n}$ is an analog of the symplectic representation of $\mathcal{P}_n$ used in the theory of quantum codes.

An n-mode GKP code, $\mathcal{Q}$, is a CV stabilizer code defined in terms of an Abelian stabilizer group $\mathcal{S} \subset \mathcal{H}_n$ with elements in the form Eq. (1), such that $\hat{U}(\boldsymbol{0}) \equiv \mathbb{1}$ is the only element in $\mathcal{S}$ proportional to the identity. Namely, the code $\mathcal{Q} \subseteq L_2(\mathbb{R}^n)$ is the common $+1$-eigenspace of all elements of $\mathcal{S}$,

$$\mathcal{Q} = \{|\psi\rangle \in L_2(\mathbb{R}^n) \mid \hat{S}\,|\psi\rangle = |\psi\rangle\,,\ \forall \hat{S} \in \mathcal{S}\}. \tag{3}$$

The structure of such Abelian subgroups and the implications for $\mathcal{Q}$ are described in Appendix A. In the following, we will assume the representation of such a group in terms of some number r of its members chosen as generators, $\mathcal{S} = \langle \hat{S}_1, \ldots, \hat{S}_r\rangle, \hat{S}_j \in \mathcal{H}_n$.

The formalism of qubit stabilizer codes [19,20] carries over entirely to such CV stabilizer codes, and errors from $\mathcal{H}_n$ play the special role played by Pauli errors in the qubit case. Given an error $\hat{E} \in \mathcal{H}_n$, one can compute its syndrome, $\boldsymbol{q} \equiv \boldsymbol{q}(\hat{E})$, whose components are given by the extra phases in the commutation relations with the stabilizer generators, $\hat{E}\hat{S}_j = \hat{S}_j\hat{E}e^{iq_j}$. The set of errors which commute with all elements of the stabilizer group is called the centralizer $\mathrm{C}(\mathcal{S})$; these errors have a trivial syndrome, $\boldsymbol{q} = \boldsymbol{0}$. Of these, any error that is a member of the stabilizer group acts trivially on code states, while the remaining errors $\hat{L} \in \mathrm{C}(\mathcal{S}) \setminus \mathcal{S}$ act nontrivially within the code, they are called *logical operators*. An error $\hat{E} \in \mathcal{H}_n$ that does not commute with all stabilizer generators has a nontrivial syndrome and it takes $\mathcal{Q}$ into an orthogonal subspace $\hat{E}\mathcal{Q} \equiv \{\hat{E}\,|\psi\rangle \mid |\psi\rangle \in \mathcal{Q}\}$. Two errors that differ by an element of the stabilizer group, $\hat{E}' = \hat{E}\hat{S}, \hat{S} \in \mathcal{S}$, have the same syndrome, and as such, are called mutually *degenerate*. They act identically on the code and are also called *equivalent*. Two errors that differ by a logical operator, $\hat{E}' = \hat{E}\hat{L}$, $\hat{L} \in \mathrm{C}(\mathcal{S}) \setminus \mathcal{S}$, also have the same syndrome but they act differently on the code. The set of inequivalent logical operators $\mathrm{L}(\mathcal{S})$ is formed by the cosets of $\mathcal{S}$ in the centralizer

$\mathrm{C}(\mathcal{S})$. If we ignore the phases, then the set of cosets $\mathrm{L}(\mathcal{S})$ actually forms a group, the group of logical operators.

By a slight abuse of notation, and when the global phase is irrelevant, we will often refer to an operator $\hat{U}(\boldsymbol{e}) \in \mathcal{H}_n$ directly by its symplectic vector component, $\boldsymbol{e} \in \mathbb{R}^{2n}$. For example, we can refer to a logical operator $\boldsymbol{c} \in \mathrm{L}(\mathcal{S})$ when we ought to write $\hat{U}(\boldsymbol{c}) \in \mathrm{L}(\mathcal{S})$. In accordance with classical codes terminology, we also refer to $\boldsymbol{c} \in \mathrm{L}(\mathcal{S})$ as a *codeword*. Furthermore, for two equivalent errors, $\boldsymbol{e}$ and $\boldsymbol{e}' = \boldsymbol{e} + \boldsymbol{s}$, where $\boldsymbol{s} \in \mathcal{S}$, we use $\boldsymbol{e}' \simeq \boldsymbol{e}$ to denote their equivalence and $[\boldsymbol{e}]$ to denote the entire equivalence class,

$$[\boldsymbol{e}] = \{\boldsymbol{e} + \boldsymbol{s} : \forall \boldsymbol{s} \in \mathcal{S}\}. \tag{4}$$

Throughout this work we consider the independent Gaussian displacement channel $\mathcal{N}(\rho)$ with standard deviation σ_0:

$$\mathcal{N}(\rho) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv\ \mathbb{P}_{\sigma_0}(u)\mathbb{P}_{\sigma_0}(v) e^{iu\hat{p}+iv\hat{q}} \rho\ e^{-iu\hat{p}-iv\hat{q}}, \tag{5}$$

where ρ is a single-mode density matrix and $\mathbb{P}_{\sigma_0}(x)$ is the Gaussian probability density function with mean zero and variance σ_0^2; i.e., $\mathbb{P}_{\sigma_0}(x) = (2\pi\sigma_0^2)^{-1/2} e^{-x^2/2\sigma_0^2}$. We will refer to σ_0 as the *bare* standard deviation, because we will often consider *scaled* observable, e.g., $\hat{P} = 2\alpha\hat{p}$, for which the corresponding effective rescaled standard deviation is $\sigma = 2\alpha\sigma_0$. Even though this channel may not necessarily be the one which is physically most relevant, what is important is that it covers dominant sources of imperfections stochastically. Any physically realistic GKP code state has finite photon number $\overline{n}$ and one reasonable model of such finite-photon GKP state is a coherent superposition of Gaussian displacement errors on a perfect code state; see Eqs. (40) and (41) in Ref. [3]. The quality of such an approximate GKP state can be given by an effective squeezing parameter Δ with $\overline{n} \sim \frac{1}{2\Delta^2} - \frac{1}{2}$. Assuming a coherent superposition of Gaussian displacements can be replaced by a Gaussian mixture of displacements on a perfect state we can identify $\Delta^2 = 2\sigma_0^2$. If errors are dominated by such a finite squeezing/finite photon number, then we could use $\sigma_0 \sim \sqrt{\frac{1}{2(2\overline{n}+1)}}$ to interpret our numerical data. For example, $\overline{n} = 4$ gives $\sigma_0 \approx 0.236$.

Besides this, it has been shown that photon loss with rate γ followed by an amplification or pumping step produces the Gaussian displacement channel with $\sigma_0^2 = \frac{\gamma}{1-\gamma}$ [5]. For example, the rate $\gamma = 0.02$ corresponds to $\sigma_0 = 0.14$. Such an amplification step on the GKP data qubit could be added in each step of error correction.

As a convention we use bold italic symbols, such as $\boldsymbol{u}$, to denote row vectors. We use hatted symbols, such as $\hat{P}$, $\hat{Q}$, for quantum operators and unhatted symbols for the corresponding eigenvalues, such as P and Q. We will consider modulo values for real numbers quite often where, for convention, we chose the remainder to be in a symmetrical interval around 0. For example, given $\phi \in \mathbb{R}$, writing $q = \phi \mod 2\pi$ means that $q \in [-\pi, \pi)$ and $\phi = q + 2\pi k$ for some $k \in \mathbb{N}$. In conventional notation $q := (\phi + \pi) \mod 2\pi - \pi$ and $k := \lfloor (\phi + \pi)/2\pi \rfloor$. We also denote a range of integers as $[n] \equiv \{1, \ldots, n\}$. We will refer to single-mode q-type errors as displacements of the form $\exp(i\eta\hat{q})$ for some η. Such errors induce shifts in p and are alternatively called shift-in-p errors. Similarly, for p-type errors which induce shifts in q.

B. Maximum-likelihood versus minimum-energy decoding

A (classical) binary linear code [21] of length n encoding k bits is a linear space of dimension k formed by binary strings of length n, $\mathcal{C} \subseteq \mathbb{F}_2^n$. For such a code, maximum likelihood (ML) syndrome-based decoding amounts to finding the most likely error which results in the given syndrome. Generally, there are 2^{n-k} distinct syndromes and $|\mathcal{C}| = 2^k$ codewords. It is not hard to find a vector $\boldsymbol{e}$ which produces the correct syndrome; ML decoding can then be done by comparing the probabilities of errors $\mathbb{P}(\boldsymbol{e} + \boldsymbol{c})$, where $\boldsymbol{c} \in \mathcal{C}$ goes over all the codewords. In the simplest case of the binary symmetric channel, the probabilities scale exponentially with error weight which can be thought of as the "energy" associated with the error. Thus, for linear binary codes under the binary symmetric channel, ML decoding is the same as the minimum-weight, or *minimum-energy* (ME) decoding.

Syndrome-based ML decoding for a qubit stabilizer code can be done similarly. The main difference here is the degeneracy: errors that differ by an element of the stabilizer group are equivalent, they can not and need not be distinguished. As a result, the probability $\mathbb{P}(E)$ of an n-qubit Pauli error $E \in \mathcal{P}_n$ needs to be replaced by the total probability to have any error equivalent to E. In the case of Pauli errors which are independent on different qubits, quite generally, this probability can be interpreted as a partition function of certain random-bond Ising model [22,23]. Exactly which statistical model one gets depends on the code. For a qubit square-lattice toric code with perfect stabilizer measurements the partition functions are those of 2D random-bond Ising model (RBIM). Similarly, for the toric code with repeated noisy measurements, the partition function is that of a random-plaquette gauge model (RPGM) in three dimensions [22], where the "time" dimension enumerates syndrome measurement cycles. More general models are discussed, e.g., in Refs. [23,24].

Instead of computing the partition functions proportional to the total probabilities of errors in different sectors, one could try finding a single most-likely error compatible with the syndrome. It is the latter method that is usually called the ME decoding for a quantum code. Indeed, in terms of the statistical-mechanical analogy, for ML decoding one needs to minimize the free energy, minus the logarithm of the partition function. In comparison, for ME decoding, one only looks at a minimum-energy configuration (not necessarily unique); this ignores any entropy associated with degenerate configurations. While the ME technique is strictly less accurate than ML decoding, in practice the difference may be small.

The two approaches are readily extended to GKP codes, both in the *channel model* where perfect stabilizer measurement is assumed, and in the more general *fault-tolerant* (FT) case where repeated measurements are used to offset the stabilizer measurement errors. The latter case can be interpreted in terms of a larger *space-time* code dealing with both the usual quantum errors and the measurement errors [22,24,25]. One important aspect is that the quantum errors accumulate over time, while measurement errors in different measurement rounds are independent from each other. This

leads to an extended equivalence between combined data-syndrome errors which is similar to degeneracy. The corresponding generators can be constructed, e.g., by starting with a single-oscillator error, followed by measurement errors on all adjacent stabilizer checks that result in zero syndrome, followed by the error which exactly cancels the original error. Because of this cancellation, such an invisible error has no effect and should be counted as a part of the degeneracy group of the larger space-time code.

The following discussion applies to a GKP code in either the channel model or the fault-tolerant model. In both cases we denote as $\mathcal{S} \subseteq \mathcal{H}_n$ the degeneracy group of the code. In the channel model, $\mathcal{S}$ is exactly the stabilizer group, acting on the data oscillators. In the fault-tolerant case, $\mathcal{S}$ is the degeneracy group of the space-time code, acting on both data oscillators as well as ancillary oscillators used to measure the syndrome. Consider a multi-oscillator error, $\boldsymbol{e} \in \mathbb{R}^{2n}$, see Eq. (1), and the corresponding probability density $\mathbb{P}(\boldsymbol{e})$. The probability is assumed to have a sharp [exponential or Gaussian, cf. Eq. (5)] dependence on the components of $\boldsymbol{e}$; for this reason we can also write

$$\mathbb{P}(\boldsymbol{e}) = \exp[-H(\boldsymbol{e})], \tag{6}$$

where $H(\boldsymbol{e})$ is the dimensionless *energy* associated with the error operator $\hat{U}(\boldsymbol{e})$. Syndrome-based ML decoding can be formulated as follows. The error $\boldsymbol{e}$ yields the syndrome $\boldsymbol{q}(\boldsymbol{e})$. Given a logical operator $\boldsymbol{c} \in \mathrm{L}(\mathcal{S})$, denote $\mathbb{P}([\boldsymbol{e}+\boldsymbol{c}]|\boldsymbol{q})$ as the probability for any error in the class $[\boldsymbol{e}+\boldsymbol{c}]$, see Eq. (4), conditioned on the syndrome $\boldsymbol{q}$. This probability can be written as

$$\begin{aligned}\mathbb{P}([\boldsymbol{e}+\boldsymbol{c}]|\boldsymbol{q}) &= \int_{\boldsymbol{s}\in\mathcal{S}} d\boldsymbol{s}\ \mathbb{P}(\boldsymbol{e}+\boldsymbol{c}+\boldsymbol{s}|\boldsymbol{q}) \\ &= \frac{1}{\mathbb{P}(\boldsymbol{q})}\int_{\boldsymbol{s}\in\mathcal{S}} d\boldsymbol{s}\ \mathbb{P}(\boldsymbol{e}+\boldsymbol{s}+\boldsymbol{c}) \\ &= \frac{1}{\mathbb{P}(\boldsymbol{q})}\int_{\boldsymbol{s}\in\mathcal{S}} d\boldsymbol{s}\ \mathrm{e}^{-H(\boldsymbol{e}+\boldsymbol{s}+\boldsymbol{c})} \equiv \frac{1}{\mathbb{P}(\boldsymbol{q})} Z_{\boldsymbol{c}}(\boldsymbol{e}),\end{aligned} \tag{7}$$

where an appropriate integration measure should be used, and $\mathbb{P}(\boldsymbol{q})$ is the net probability density to obtain the syndrome $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{e})$. Among all inequivalent codewords $\boldsymbol{c} \in \mathrm{L}(\mathcal{S})$, we select the most likely, i.e., with the largest $Z_{\boldsymbol{c}}(\boldsymbol{e})$. The probability of leaving a logical error $\boldsymbol{c}$ after ML decoding is the net probability of all the errors $\boldsymbol{e}$ for which the sector $[\boldsymbol{e}-\boldsymbol{c}]$ is the most likely, so

$$\mathbb{P}^{(\mathrm{ML})}(\boldsymbol{c}) = \int_{\boldsymbol{e}:\ \forall \boldsymbol{b}\not\simeq(-\boldsymbol{c}),\ Z_{-\boldsymbol{c}}(\boldsymbol{e})>Z_{\boldsymbol{b}}(\boldsymbol{e})} d\boldsymbol{e}\ \mathbb{P}(\boldsymbol{e}), \tag{8}$$

[here we disregarded the contribution from sectors $\boldsymbol{b} \not\simeq (-\boldsymbol{c})$ equiprobable with $(-\boldsymbol{c})$, $Z_{-\boldsymbol{c}}(\boldsymbol{e}) = Z_{\boldsymbol{b}}(\boldsymbol{e})$]. The probability of success of ML decoding can then be expressed as $\mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}} = \mathbb{P}^{(\mathrm{ML})}(\boldsymbol{0})$. It is easy to see that any other decoding algorithm gives success probability that is not higher than that of ML decoding. Indeed, a different algorithm would swap some errors $\boldsymbol{e}$ for $\boldsymbol{e}+\boldsymbol{c}$, which may reduce the measure in the corresponding analog of Eq. (8).

Furthermore, given an error $\boldsymbol{e}$, the probability $\mathbb{P}(\boldsymbol{q})$ to obtain the syndrome $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{e})$ can be written as $\mathbb{P}(\boldsymbol{q}) = \int_{\boldsymbol{b}\in\mathrm{L}(\mathcal{S})} d\boldsymbol{b}\, Z_{\boldsymbol{b}}(\boldsymbol{e})$, using the appropriate integration measure for the logical operators $\boldsymbol{b}$. For this error $\boldsymbol{e}$ we denote $\boldsymbol{c}_{\max}(\boldsymbol{e})$ as its corresponding most likely sector,

$$\boldsymbol{c}_{\max}(\boldsymbol{e}) = \underset{\boldsymbol{c}\in\mathrm{L}(\mathcal{S})}{\arg\max}\, Z_{\boldsymbol{c}}(\boldsymbol{e}). \tag{9}$$

The probability of a logical error $\boldsymbol{c}$ after ML decoding Eq. (8) can then be rewritten as an expectation by multiplying and dividing by $\mathbb{P}(\boldsymbol{q})$, changing variables, and resumming, resulting in

$$\begin{aligned}\mathbb{P}^{(\mathrm{ML})}(\boldsymbol{c}) &= \langle \mathbb{P}([\boldsymbol{e}+\boldsymbol{c}_{\max}+\boldsymbol{c}]|\boldsymbol{q}(\boldsymbol{e}))\rangle \\ &= \int d\boldsymbol{e}\ \mathbb{P}(\boldsymbol{e})\,\frac{Z_{\boldsymbol{c}_{\max}+\boldsymbol{c}}(\boldsymbol{e})}{\mathbb{P}(\boldsymbol{q})} \\ &= \int d\boldsymbol{e}\ \mathbb{P}(\boldsymbol{e})\,\frac{Z_{\boldsymbol{c}_{\max}+\boldsymbol{c}}(\boldsymbol{e})}{\int d\boldsymbol{b}\, Z_{\boldsymbol{b}}(\boldsymbol{e})}.\end{aligned} \tag{10}$$

ML decoding is successful if the most likely error is actually the one that happened, which corresponds to the trivial sector $\boldsymbol{c} \simeq \boldsymbol{0}$ being dominant over all other sectors $\boldsymbol{0} \not\simeq \boldsymbol{c} \in \mathrm{L}(\mathcal{S})$. Given the error probability distribution $\mathbb{P}(\boldsymbol{e})$, we say that a sequence of discrete GKP codes of increasing length n is in the *decodable phase* if $\mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}} \equiv \mathbb{P}^{(\mathrm{ML})}(\boldsymbol{0}) \to 1$ with $n \to \infty$.

With the definitions Eqs. (6) and (7), $Z_{\boldsymbol{c}_{\max}}(\boldsymbol{e})$ can be interpreted as a partition function of a classical model in the presence of quenched randomness determined by the actual error $\boldsymbol{e}$. The partition function $Z_{\boldsymbol{c}_{\max}+\boldsymbol{c}}(\boldsymbol{e})$ differs by an addition of a defect, e.g., a homologically nontrivial domain wall at the locations specified by nonzero components of the codeword $\boldsymbol{c}$. Having already $\boldsymbol{c}_{\max}(\boldsymbol{e}) \not\simeq \boldsymbol{0}$ means that the disorder, $\boldsymbol{e}$, energetically favors the domain wall $\boldsymbol{c}_{\max}$. In the following, we will also consider the free energy, $F_{\boldsymbol{c}}(\boldsymbol{e})$,

$$F_{\boldsymbol{c}}(\boldsymbol{e}) \equiv -\ln Z_{\boldsymbol{c}_{\max}+\boldsymbol{c}}(\boldsymbol{e}), \tag{11}$$

as well as the corresponding average $\langle F_{\boldsymbol{c}}\rangle \equiv \int d\boldsymbol{e}\, \mathbb{P}(\boldsymbol{e})\, F_{\boldsymbol{c}}(\boldsymbol{e})$. It follows from the Gibbs inequality that below the error-correction threshold for the noise parameters in $\mathbb{P}(\boldsymbol{e})$, the free-energy increment $\Delta F_{\boldsymbol{c}} \equiv F_{\boldsymbol{c}}(\boldsymbol{e}) - F_{\boldsymbol{0}}(\boldsymbol{e})$ associated with a logically distinct "incorrect" class ($\boldsymbol{c} \not\simeq \boldsymbol{0}$) necessarily diverges with n for any error, $\boldsymbol{e}$, likely to happen [23]. More precisely, if ML decoding is asymptotically successful with probability one, $\mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}} \to 1$, then the average free-energy increment, $\langle \Delta F_{\boldsymbol{c}}\rangle$ associated with any nontrivial codeword $\boldsymbol{c} \not\simeq \boldsymbol{0}$ must diverge for $n \to \infty$. Such a divergence can be seen as a signature of a phase transition in the corresponding model.

As is the case of the surface codes [22], the partition functions $Z_{\boldsymbol{c}}(\boldsymbol{e})$ are evaluated at a temperature that is not a free parameter but depends on the distribution $\mathbb{P}(\boldsymbol{e})$. For the sake of understanding the physics of the corresponding models, we could relax this, e.g., by additionally rescaling the energy $H(\boldsymbol{e}) \to \beta H(\boldsymbol{e})$, cf. Eq. (6), in the definition Eq. (7) of the partition function $Z_{\boldsymbol{c}}(\boldsymbol{e})$, while keeping the original error probability distribution in the average Eq. (10). This amounts to using ML decoder with incorrect input information, thus the corresponding success probability is not expected to increase,

$$\mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}}(\beta) \leqslant \mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}}(\beta = 1) \equiv \mathbb{P}^{(\mathrm{ML})}_{\mathrm{succ}}, \tag{12}$$

similar to decoding away from the Nishimori line in the case of qubit stabilizer code [22,23]. In particular, the limit $\beta \to \infty$ corresponds to ME decoding, where we are choosing the

codeword $\boldsymbol{c}$ to minimize the function

$$H_{\boldsymbol{c}}(\boldsymbol{e}) \equiv \min_{s\in\mathcal{S}} H(\boldsymbol{e}+\boldsymbol{c}+\boldsymbol{s}). \tag{13}$$

III. PROTECTING A SINGLE GKP QUBIT

A. Setup

The single-mode GKP code [3] is a prescription to encode a qubit—a two-dimensional Hilbert space—into the Hilbert space of an oscillator using a discrete subgroup of displacement operators $\mathcal{H}_1$ as the stabilizer group. One chooses the two commuting displacement operators, $S_p = e^{2i\alpha\hat{p}}$ and $S_q = e^{i2\pi\hat{q}/\alpha}$, where $\alpha \neq 0$ is any real number. For this encoded qubit the (logical) Pauli operators are $Z = e^{i\pi\hat{q}/\alpha}$ (with $Z^2 = S_q$) and $X = e^{i\alpha\hat{p}}$ (with $X^2 = S_p$). One can verify that $XZ = -ZX$. The oscillator observables, $\hat{P} = 2\alpha\hat{p}$ and $\hat{Q} = 2\pi\hat{q}/\alpha$, can both take any value $2\pi k$ for $k \in \mathbb{Z}$ on ideal codewords. The codeword $|\bar{0}\rangle$ (respectively, $|\bar{1}\rangle$) is distinguished by k being even (respectively, odd). The action of phase space translations S_p (respectively, S_q) on the eigenvalues of $\hat{Q}$ (respectively, $\hat{P}$) is $Q \to Q + 4\pi$ (respectively, $P \to P + 4\pi$). The action of X (respectively, Z) is $Q \to Q + 2\pi$ (respectively, $P \to P + 2\pi$).

A visual representation can be obtained by imagining the variables Q and P as a torus in phase space with both handles of circumference 2π. In this representation S_p lets Q wind around the handle exactly twice, while X lets Q go around the handle exactly once. A correctable error constitutes a shift in Q by less than half the circumference. In this convenient representation, a logical error thus occurs when the winding number is odd, and no error occurs when the winding number is even. The shifts in P corresponds to windings around the other handle of the torus.

We will assume that the oscillator undergoes noise modeled as a Gaussian displacement channel with *bare* standard deviation σ_0, see Eq. (5). The effect on the scaled observables $\hat{P}$ and $\hat{Q}$ is to map $P \to P + \epsilon_p$ and $Q \to Q + \epsilon_q$ where ϵ_p and ϵ_q are drawn from Gaussian distributions with *rescaled* variances,

$$\sigma_P^2 = 4\alpha^2\sigma_0^2 \quad \text{and} \quad \sigma_Q^2 = \frac{4\pi^2\sigma_0^2}{\alpha^2}. \tag{14}$$

For symmetry reasons, α is chosen to be $\sqrt{\pi}$ and we write $\sigma = \sigma_P = \sigma_Q$. Given perfect measurements of S_p, the error ϵ_q can be corrected if $|\epsilon_q \bmod 4\pi| < \pi$.

To measure stabilizer generators S_p and S_q we consider the fault-tolerant Steane measurement circuits [3] in Fig. 1, where encoded $|\overline{+}\rangle$ or $|\bar{0}\rangle$ ancillas, CNOTs and $\hat{q}$ or $\hat{p}$ measurements are used.

For simplicity, we consider the ancilla preparations, the CNOT and the $\hat{q}$ and $\hat{p}$ measurements to be perfect and only add Gaussian displacement channels on the data qubit and on the ancilla qubit right before its measurement. Doing this ignores the back propagation of q-type errors to the data due to an imperfect ancilla [26], but if we treat p-type and q-type error correction independently, then this back-propagation does not fundamentally alter the noise model. We will keep the freedom of choosing different standard deviations for the data and the ancilla errors and denote as σ_{M} the *scaled* standard deviation for the ancilla errors.

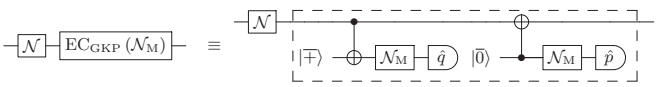

FIG. 1. A single round of fault-tolerant GKP syndrome measurement for both q and p shifts. Here $|\overline{+}\rangle$ is the $+1$ eigenstate of S_q and X, and $|\bar{0}\rangle$ is the $+1$ eigenstate of S_p and Z. The CNOT gate is the logical CNOT for the GKP code which induces the transformation $q_{\text{target}} \to q_{\text{control}} + q_{\text{target}}$ (while $p_{\text{control}} \to p_{\text{control}} - p_{\text{target}}$, $q_{\text{control}} \to q_{\text{control}}$, $p_{\text{target}} \to p_{\text{target}}$). Each measurement is a perfect homodyne measurement of $\hat{q}$ or $\hat{p}$. $\mathcal{N}$ are $\mathcal{N}_{\mathrm{M}}$ are Gaussian displacement channels in Eq. (5) which model shift errors on the encoded state in each round of error correction, respectively, shift errors in the homodyne measurement.

Since the measurement outcomes in Fig. 1 are inaccurate, they cannot be used to infer a correction which maps the state back to the code space. To perform error correction one has to measure frequently and try to use the record of measurements to stay as close as possible to the code space without incurring logical errors to preserve the codeword. Figure 2 shows this repeated measurement protocol for p-type errors (or shift-in-q errors).

We will analyze only p-type data errors with scaled standard deviation σ and measurement errors with scaled standard deviation σ_{M}, cf. Eq. (14). The analysis for q-type errors would be similar. When considering the realization of a particular shift error we will use the following notation: $\epsilon_t \in \mathbb{R}$ is the shift error occurring on the data before the tth measurement, $\delta_t \in \mathbb{R}$ is the measurement error occuring at the tth step. Furthermore, $q_t \in [-\pi, \pi)$ is the tth measurement outcome for the rescaled variable $\hat{Q}$ and $\phi_t = \epsilon_1 + \cdots + \epsilon_t$ is the cumulative shift on the data. The relations between these quantities are

$$q_t = \phi_t + \delta_t \bmod 2\pi, \quad \phi_t - \phi_{t-1} = \epsilon_t, \quad \phi_0 \equiv 0. \tag{15}$$

We consider a total of M rounds of GKP measurements indexed by $t \in [M]$. Of these, the last measurement is assumed perfect, incorporating any measurement error into the corresponding shift error. Specifically, we write $\phi_M = \phi_{M-1} + \epsilon_M$, $q_M = \phi_M \bmod 2\pi$, so that $\delta_M = 0$. This last measurement can be thought of as a destructive measurement performed directly on the data without the use of an ancilla, as one would do to retrieve the encoded information. As such, the last data error ϵ_M can equivalently be thought of as the last measurement error on the destructive measurement of the data. Having this last perfect measurement permits us to map back to the code space and easily define successful or failed

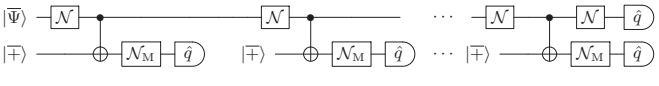

FIG. 2. Repeated rounds of error correction for the GKP code to detect and keep track of error shifts in $\hat{q}$ followed by a final destructive measurement of the data (modeled as $\mathcal{N}$ followed by a perfect measurement of $\hat{q}$). No explicit corrections based on the measured values are shown.

error correction. Specifically, we are trying to determine the parity of k_M in the relation $q_M = \phi_M + 2\pi k_M$; error correction is successful as long as we determined the parity correctly. We denote the set of M measurements as $\boldsymbol{q}$ and M cumulative shift errors as $\boldsymbol{\phi}$.

To get some intuition, imagine that we apply a single round of error correction of Fig. 2 and $\mathcal{N}_{\mathrm{M}}$ is the identity channel. The ancilla qubit $|\overline{+}\rangle$ is a uniform sum of delta functions with $Q = q = 0 \bmod 2\pi$, hence we represent the measurement outcome compactly as $q \in [-\pi, \pi)$. An incoming logical X on the data qubit is pushed (through the CNOT) onto the ancilla qubit where it translates q by a full 2π-period, hence logical information is not observed. One corrects a shift of up to π (at most half-a-logical) by shifting Q back by the least amount to make it again equal to $0 \bmod 2\pi$.

B. Decoding strategies

We start by describing the maximum-likelihood strategy. Given the measurement record, one would like to compute the conditional probabilities for different classes of errors which are distinguished by their logical action. In this case of correcting a single qubit against shift errors in q, one has to decide whether there was an X error or there was none. Knowing the details of the error model, namely σ and σ_{M}, one can write down the probability of these two classes. Formally, they are given by

$$\mathbb{P}(0|\boldsymbol{q}) = \int_{I_0} \mathbb{P}(\boldsymbol{\phi}|\boldsymbol{q}) d\boldsymbol{\phi}, \quad \mathbb{P}(1|\boldsymbol{q}) = \int_{I_1} \mathbb{P}(\boldsymbol{\phi}|\boldsymbol{q}) d\boldsymbol{\phi}, \tag{16}$$

where the integration covers all possible realizations of the shift errors described by $\boldsymbol{\phi}$, and I_0 (respectively, I_1) limits the integral to realizations leaving no X error (respectively, leaving an X error). Since the last measurement is assumed perfect, I_0 and I_1 are characterized by $\delta(\phi_M - q_M + 2\pi k_M)$, with any even k_M in I_0 and any odd k_M for I_1.

In practice, to do decoding for the given measurement history, $\boldsymbol{q}$, one needs to compare the probabilities Eq. (16). ML decoding algorithm suggests that a logical X correction is needed if $\mathbb{P}(1|\boldsymbol{q}) > \mathbb{P}(0|\boldsymbol{q})$. Of course, this does not guarantee success in each particular trial. If we take just one measurement round, $M = 1$, which corresponds to measuring the data directly, then we get

$$\begin{aligned}\mathbb{P}_{M=1}(0|q_1) &= \sum_{k_1 \text{even}} \int_{-\infty}^{\infty} d\epsilon_1 \, \mathbb{P}(\epsilon_1|q_1)\delta(\epsilon_1 - q_1 + 2k_1\pi) \\ &\propto \sum_{k_1 \text{ even}} \mathbb{P}_\sigma(q_1 - 2\pi k_1),\end{aligned} \tag{17}$$

and $\mathbb{P}_{M=1}(1|q_1)$ is given by the complementary sum over odd k_1, which makes the normalization the full sum with k_1 running over all integer values.

To compute these probabilities in general, we apply Bayes' rule:

$$\mathbb{P}(\boldsymbol{\phi}|\boldsymbol{q}) = \frac{\mathbb{P}(\boldsymbol{q}|\boldsymbol{\phi})\mathbb{P}(\boldsymbol{\phi})}{\mathbb{P}(\boldsymbol{q})}. \tag{18}$$

Then the probability for some outcome $\boldsymbol{q}$ given data errors $\boldsymbol{\phi}$ can be computed from the measurement error model and the probability for some data error $\boldsymbol{\phi}$ from the data error model. The normalization, $\mathbb{P}(\boldsymbol{q})$, can be computed by integrating the numerator over every $\boldsymbol{\phi}$. Using Eq. (15), we have from Eq. (18)

$$\mathbb{P}(\boldsymbol{\phi}|\boldsymbol{q})\mathbb{P}(\boldsymbol{q}) \propto \sum_{\boldsymbol{k}\in\mathbb{Z}^M} \left[\prod_{t=1}^{M-1} \exp\left(-\frac{(q_t - \phi_t + 2\pi k_t)^2}{2\sigma_{\mathrm{M}}^2} \right) \prod_{t=1}^{M} \exp\left(-\frac{(\phi_t - \phi_{t-1})^2}{2\sigma^2} \right) \delta(\phi_M - q_M - 2\pi k_M) \right]. \tag{19}$$

Recalling Eq. (7), we write the corresponding complementary probabilities Eq. (16) in terms of partition functions,

$$\mathbb{P}(0|\boldsymbol{q}) = \frac{Z_0(\boldsymbol{q})}{\mathbb{P}(\boldsymbol{q})}, \quad \mathbb{P}(1|\boldsymbol{q}) = \frac{Z_1(\boldsymbol{q})}{\mathbb{P}(\boldsymbol{q})}, \tag{20}$$

$$\begin{aligned}Z_c(\boldsymbol{q}) = N^{-1} \int d\boldsymbol{\phi} \sum_{\boldsymbol{k}\in\mathbb{Z}^{M-1}} \exp\left[-\sum_{t=1}^{M-1} \frac{(q_t - \phi_t + 2\pi k_t)^2}{2\sigma_{\mathrm{M}}^2} \right] \exp\left[-\sum_{t=1}^{M} \frac{(\phi_t - \phi_{t-1})^2}{2\sigma^2} \right] \\ \times \sum_{k_M\in\mathbb{Z}} \delta(\phi_M - q_M - 2\pi c - 4\pi k_M), \quad c = 0, 1,\end{aligned} \tag{21}$$

with N a normalization constant [27]. In this special case picking $\boldsymbol{e} = \boldsymbol{q}$, for a candidate error is always a valid choice, that is why we can write directly $Z_c(\boldsymbol{q})$.

The evaluation of the Gaussian integrals in Eq. (21) (see Appendix B 1) gives

$$\begin{aligned}Z_c(\boldsymbol{q}) = \frac{(2\pi)^{(M-1)/2}}{N(\det B)^{1/2}} \sum_{\boldsymbol{k}=(\tilde{\boldsymbol{k}}, c+2m):\tilde{\boldsymbol{k}}\in\mathbb{Z}^{M-1},\, m\in\mathbb{Z}} \\ \times \exp\left[-\frac{1}{2}(\boldsymbol{q} + 2\pi\boldsymbol{k})A(\boldsymbol{q} + 2\pi\boldsymbol{k})^{\mathrm{T}} \right],\end{aligned} \tag{22}$$

where B and A are symmetric positive-definite matrices given explicitly by Eqs. (B1) and (B2) in the Appendix. The sums with $c \in \{0, 1\}$ can be numerically computed by setting a cutoff K, restricting every k_t to the interval $-K \leqslant k_t \leqslant K$. The number of terms then still exponentially increases with the number of rounds. We used Eq. (22) with a cutoff $K = 2$ for up to $M = 7$ rounds for the results shown in Figs. 4 and 5. Intuitively, this cutoff corresponds to only considering events where the measurement shift errors let one wind around the torus at most twice in each round. This is pretty reasonable

since these errors follow a Gaussian distribution with small variance.

Generally, a more clever way to calculate the sum in Eq. (22) is to express it in terms of a genus-M Riemann theta function [28], and then transform the matrix so that the summation terms can be rearranged in decreasing order, stopping at a desired precision. However, this requires solving a shortest vector problem with the eigenvectors of A and is therefore also computationally difficult [29,30].

In addition to the formally exact but hard to calculate Eqs. (21) and (22) for the conditional probabilities, we would like to consider a class of approximate minimum-energy solutions of the corresponding optimization problem. To this end, we define a 2π-periodic potential, $V_\sigma(x) = V_\sigma(x+2\pi)$,

$$\exp[-V_\sigma(x)] \equiv \sum_{k\in\mathbb{Z}} e^{-(x+2\pi k)^2/2\sigma^2}. \tag{23}$$

The periodicity of the sum of the Gaussians implies that one should be able to approximate $V_\sigma(x)$ by its principal Fourier harmonic,

$$V_\sigma(x) \approx A_0 - \beta_V(\sigma)\cos x, \tag{24}$$

where β_V is Villain's effective inverse temperature parameter, and the overall shift A_0 is irrelevant. Such a simplified form is exactly the approximation used by Villain [31], but "in reverse." Indeed, for large β, one has [32]

$$e^{\beta\cos x} \approx e^{\beta} \sum_{k\in\mathbb{Z}} \exp\left(-\frac{\beta(x+2\pi k)^2}{2}\right), \tag{25}$$

which gives $\beta_V(\sigma) = 1/\sigma^2$, $\sigma \ll 1$.

With the defined periodic potential, the logarithm of the nonsingular part of Eq. (19) acquires a form of a discrete-time Euclidean action, cf. Eq. (6)

$$\begin{aligned} H(\phi;\boldsymbol{q}) &\equiv -\ln[\mathbb{P}(\boldsymbol{q}|\phi)\mathbb{P}(\phi)] \\ &= \sum_{t=1}^{M} \frac{(\phi_t-\phi_{t-1})^2}{2\sigma^2} + \sum_{t=1}^{M-1} V_{\sigma_{\mathrm{M}}}(q_t-\phi_t) + \text{const.} \end{aligned} \tag{26}$$

With the given values $\phi_0 = 0$ and ϕ_M, the corresponding extremum can be found by solving the equations

$$\phi_{t+1} - 2\phi_t + \phi_{t-1} + \sigma^2 V'_{\sigma_{\mathrm{M}}}(q_t-\phi_t) = 0, \quad t \in [M-1], \tag{27}$$

where $V'_{\sigma_{\mathrm{M}}}(x)$ denotes the derivative of the potential in Eq. (23). These equations can be readily solved one-by-one, starting with $\phi_0 = 0$ and some $\phi_1 \equiv \phi$; the boundary condition $\phi_M = q_M + 2\pi k_M$ can be satisfied by scanning over different values of ϕ_1 in a relatively small range around zero, with the global minimum subsequently found by comparing the resulting values of the sum in Eq. (26). Then, any even value of k_M corresponds to no logical error, while an odd k_M indicates an X error to be corrected. While such a minimization technique gives the exact ME solution, in practice it is rather slow. Namely, with increasing nonlinearity $\sigma^2/\sigma_{\mathrm{M}}^2$ and increasing length M of the chain, a small change in ϕ_1 may strongly affect the configuration of the entire chain. Respectively, it is easy to miss an extremum corresponding to the global minimum. This numerical complexity of minimizing Eq. (26) is a manifestation of chaotic behavior inherent in Eqs. (27).

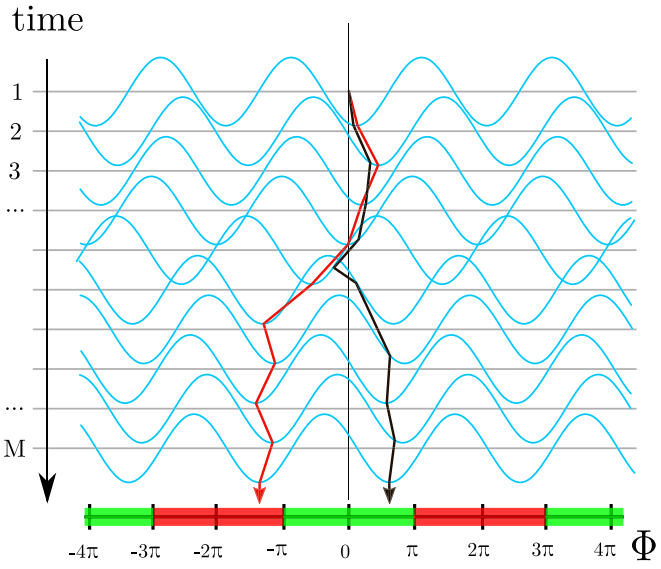

FIG. 3. Sketch of GKP decoding problem and its solution strategy. For each round of error correction $t \in [M]$ one has a Villain potential $V_{\sigma_{\mathrm{M}}}(q_t-\phi_t)$, depicted in blue, with the minimum centered at the measured value q_t. Green and red intervals along the horizontal axis denote the sets I_0 and I_1 in Eq. (16), they correspond to realizations leaving no error or an X error, respectively. In a memoryless decoding strategy one decides how to shift the code state after each measurement based on the value of q_t. In a maximum-likelihood decoding procedure one evaluates the path integral in Eq. (22). In a minimum-energy decoder one determines an optimal path for the sequence ϕ_t, $t \in [M]$, given the random potential and the quadratic "kinetic energy" term proportional to $(\phi_t - \phi_{t-1})^2$, see Eq. (26). Red and black lines show two decoding trajectories which start at the same point but have different winding numbers. Upon a final decoding step, choosing the black trajectory leads to deciding that no logical X error has taken place, since the final value ϕ_M lies in the green region. Choosing the red trajectory leads to deciding that a logical X has taken place since the final value ϕ_M lies in the red region. Examples of actual trajectories are shown in Fig. 13 in the Appendix.

Indeed, the problem of minimizing the energy Eq. (26) can be interpreted as a disordered version of a generalized Frenkel-Kontorova (FK) model [33], where a chain of masses coupled by springs lies in a periodic potential. In our setup, random shifts q_t can be traded for randomness in the initial (unstretched) lengths of the springs, $q_t - q_{t-1}$. The original FK model, with $V_\sigma(x)$ replaced by a harmonic function, is obtained if one uses the Villain approximation "in reverse," see Eq. (24). Even in the absence of disorder, the FK model is an example of a minimization problem with multiple competing minima which can be extremely close in energy. The corresponding Eqs. (27), viewed as a two-dimensional map $(\phi_{t-1},\phi_t) \to (\phi_t,\phi_{t+1})$, are a version of the Chirikov-Taylor area-preserving map from a square of size 2π to itself [34], one of the canonical examples of emergent chaos.

For this reason, and also in an attempt to come up with a numerically efficient decoding algorithm, we have

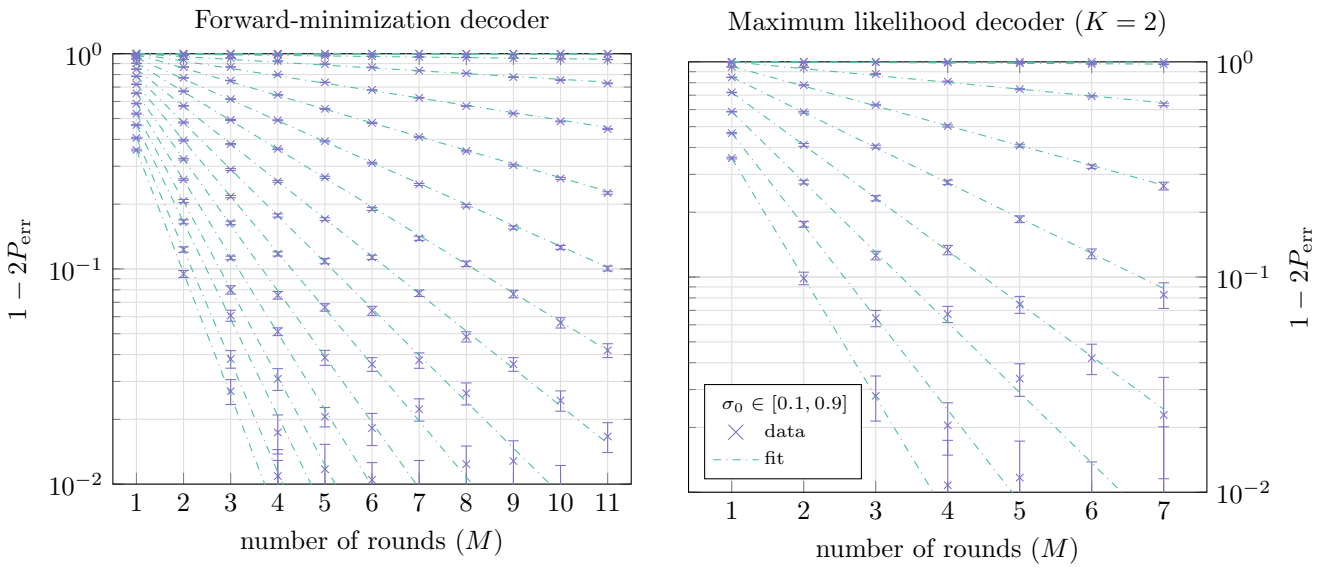

FIG. 4. Some of the numerical results for the forward-minimization and maximum-likelihood decoders (with cutoff $K = 2$). To observe the exponential decay towards $1/2$ we plot $1 - 2P_{\text{err}}$ on a log scale for different number of rounds and different bare standard deviation σ_0. Low hundred thousands of trials are performed for each data point and the confidence intervals at 95% are shown. For each σ_0 we fit an exponential decay, the slope gives us an effective logical error rate per round. On both plots the value for σ_0 varies between 0.1 and 0.9, on the left by 0.05 increments, on the right by 0.1 increments. All effective logical error rates are plotted in Fig. 5.

designed the following approximate *forward-minimization* technique. For each $t < M$, starting from $t = 1$, given the present value ϕ_{t-1}, one determines the next value ϕ_t such that $\frac{\partial H}{\partial \phi_t}|_{\phi_{t+1}=\phi_t} = 0$. Given the syndrome q_t, this implies

$$\phi_t = \arg\min_{\phi} \left[\frac{(\phi - \phi_{t-1})^2}{2} - \sigma^2 V_{\sigma_M}(q_t - \phi) \right]$$
$$\Rightarrow \phi_t = \sigma^2 V'_{\sigma_M}(q_t - \phi_t) + \phi_{t-1}. \quad (28)$$

At the end, after one obtains ϕ_{M-1}, one chooses a k_M such that $q_M + 2\pi k_M$ is the closest to ϕ_{M-1}. The parity of thus chosen k_M then tells if a logical error happened. This strategy is illustrated in Fig. 3.

These equations are certainly different from the exact extremum Eqs. (27), and the configuration found by this forward minimization technique necessarily has the energy higher than the exact minimum. However, empirically, the corresponding energy difference is typically small, much smaller what one gets, if the correct minimum is missed by the formally exact technique based on Eqs. (27). Even though this technique is only an approximation, it is fast and is accurate enough in practice. To ensure that the approximation does not hurt

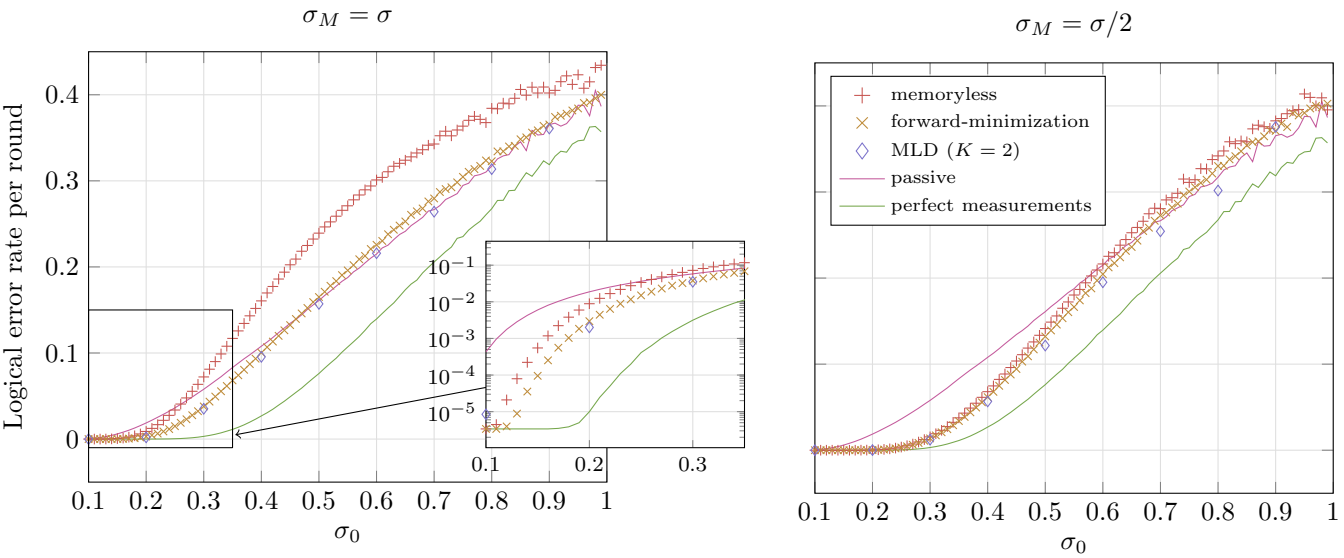

FIG. 5. Plots of the observed effective logical error rate per round for different decoding techniques. In the left plot we have taken the same standard deviation for measurement and data errors. On the right we have taken the measurement standard deviation equal to half that of the data errors. The horizontal axis shows the bare standard deviation σ_0.

the performance of our decoder, we have compared it to a rigorous dynamic programming approach; see Appendix B 2. This comparison shows that the dynamic programming approach has very little advantage while it is substantially slower in its execution.

A very simple decoding strategy that one might also try is to trust every measurement outcome and immediately correct each round. This does not require any memory so we refer to it as the *memoryless* decoder. Intuitively, this method is risky as every round transfers the measurement errors to the data, increasing the variance of the effective error model acting on the data.

C. Numerical results

We have numerically simulated these different decoders: maximum likelihood with a cutoff, forward-minimization, and memoryless decoding. We have also compared them to the scenario where the measurements are perfect, as well as the completely *passive* decoder where one lets shift errors happen without performing error correction measurements. For each scenario we considered up to 11 rounds of measurements ($M = 7$ for the maximum likelihood decoder), sampled errors, applied the decoder and gathered statistics of success or failure of the procedure for different bare standard deviation for the errors $\sigma_0 \in [0.1, 1]$.

In every scenario we observe an exponential decay toward $1/2$ of the probability of logical error as seen in Fig. 4. Thus, as expected, there is an eventual loss of logical information for any values of σ and σ_M. The decay can be fitted to extract an effective logical error rate per round which we then plotted in Fig. 5.

One striking observation is that above a certain bare standard deviation, the measurement outcomes are simply not reliable enough, so that one cannot do substantially better than throwing away the measurements and passively letting errors accumulate. Roughly speaking, this occurs for $\sigma_0 \gtrsim 0.5$ when $\sigma_M = \sigma$ and for $\sigma_0 \gtrsim 0.7$ when $\sigma_M = \sigma/2$. Another observation for $\sigma_M = \sigma$ is that the memoryless technique actually quickly does more harm than the passive approach which forgoes error correction altogether. Finally, we observe that in the range of parameters studied, the forward-minimization technique performs almost as well as the maximum-likelihood decoding, while having the advantage of being much simpler computationally.

IV. CONCATENATION: TORIC-GKP CODE

A. Setup

We consider the following setup shown in Fig. 6. We have a 2D lattice of oscillators such that each oscillator encodes a single GKP qubit. To error correct these GKP qubits by the repeated application of the circuits in Fig. 1, a GKP ancilla qubit oscillator is placed next to each data oscillator, allowing for the execution of these circuits. After each step of GKP error correction, we measure the checks of a surface or toric code: a single error correction cycle for one of the toric-code checks is shown in Fig. 7. Note first that we omit GKP error correction after each gate in the circuit in Fig. 7: the reason is that we assume that these components are noiseless in this setup so nothing would be gained by adding this. Second, the check operators of the toric code are those of the continuous-variable toric code [35] which are commuting operators on the whole oscillator space; see Appendix F. The reason for using these checks is that for the displacements X and Z of a GKP qubit, it only holds that $X = X^{-1}$ and $Z = Z^{-1}$ on the code space. Expressed as displacement operators on two oscillators 1 and 2, it holds that $[X_1X_2, Z_1Z_2^{-1}] = 0$. The upshot is that one has to use some inverse CNOTs in the circuit in Fig. 7.

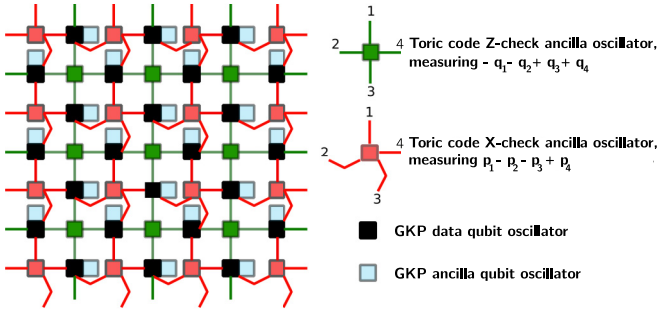

FIG. 6. The two-dimensional layout of oscillators, e.g., high-Q cavities, for the toric-GKP code. Shown is a fragment of a surface or toric code lattice. The different $\pm$ signs are defined by the orientations as explained in the main text.

In the following, we will denote vertices of the square lattice with letters i, j, k, l, and the directed edges with the corresponding vertex pairs, e.g., $e = (ij)$. Quantities defined on the edges will be considered as vector quantities, e.g., $\hat{p}_{ij} = -\hat{p}_{ji}$ for the momentum operator. The *preferred orientation* is given by the direction of the coordinate axis. For such a lattice vector field, say some field $\boldsymbol{f}$, defined on edges, with components $f_{ij} = -f_{ji}$, we will denote the sum of the vectors from a vertex j as

$$(\nabla \cdot f)_j \equiv \sum_{k\sim j} f_{jk}, \tag{29}$$

where the summation is over all vertices k that are neighboring with j. Note that this choice, recovers the $\pm$ signs for a X-check (red) in Fig. 6, when the quantities f_{jk} are written in their preferred orientation. The circulation of a vector around

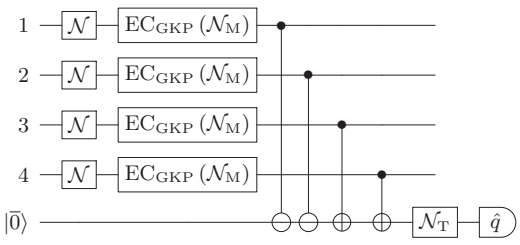

FIG. 7. A single round of error correction for a Z-check for the toric-GKP code in Fig. 6, on oscillators numbered 1 to 4. The GKP error correction unit is given in Fig. 1, $|\bar{0}\rangle$ is a $+1$ eigenstate of Z and S_p. The inverse CNOT which induces the transformation $q_{\text{target}} \to q_{\text{target}} - q_{\text{control}}$ (while $p_{\text{control}} \to p_{\text{control}} + p_{\text{target}}$) is denoted using a $\ominus$ at the target qubit instead of a $\oplus$. A parallel execution of the CNOTs for the X-checks is possible in the toric code.

a (square) plaquette $p \equiv (ijkl)$ is denoted as

$$(\nabla \times f)_p \equiv f_{ij} + f_{jk} + f_{kl} + f_{li}. \tag{30}$$

The preferred orientation for a plaquette, $p = (ijkl)$, is the one for which the closed path $i \to j \to k \to l \to i$ turns counter-clockwise. With this choice Eq. (30) recovers $\pm$ signs for a Z-check (green) in Fig. 6, when the quantities f_{jk} are written in their preferred orientation. With these notations, the vertex $\hat{A}_j$ (X-type) and the plaquette $\hat{B}_p$ (Z-type) operators of the toric code in Fig. 6 can be denoted as,

$$\hat{A}_j \equiv e^{i\sqrt{\pi}(\nabla \cdot \hat{p})_j} \quad \text{and} \quad \hat{B}_p \equiv e^{i\sqrt{\pi}(\nabla \times \hat{q})_p}. \tag{31}$$

B. Noiseless measurements and numerical results

We first examine the operation of the toric-GKP code in the channel setting. This simplified error model is based on the assumption that there are no measurement errors, i.e., $\mathcal{N}_{\mathrm{M}} = \mathbb{1}$ and $\mathcal{N}_{\mathrm{T}} = \mathbb{1}$ in Fig. 7, or equivalently, $\sigma_{\mathrm{M}} = 0$ and $\sigma_{\mathrm{T}} = 0$. In other words, the assumption is that both the GKP syndrome and the toric code syndrome are measured perfectly at every round.

Generally, when two codes are concatenated, it is possible to pass error information of the lower-level code (in this case, the GKP code) to the decoder for the top-level code (here the toric code). The information that can be passed on is an estimation of the error rate on the underlying GKP qubits based on the outcome of the GKP error correction measurement. Intuitively, if the GKP measurement gives a $q \in [-\pi, \pi]$ which lies at the boundaries of the interval, say, beyond $-\pi/2$ or $\pi/2$, we are less sure that we have corrected this shift correctly [36]. In other words, the logical error rate depends on the measured value of the GKP syndrome, and this conditional error rate can be used in the standard minimum-weight matching decoder for the toric code. If the conditional single-qubit error rate fluctuates throughout the lattice, then one can expect that using this information will substantially benefit the toric code decoder.

We numerically demonstrate that this is the case for the toric-GKP code, reproducing some of the results in Ref. [16]. The threshold of the toric code without measurement error is about 11% [22]. If we are not using any GKP error information in the toric code decoding, then the threshold for σ_0 is set by the value for which $\mathbb{P}(X) = 11\%$ with $\mathbb{P}(X)$ shown as the green line in Fig. 5. We can run a standard minimum-weight matching toric code decoder where qubit X-errors are generated by sampling Gaussian noise with standard deviation σ followed by perfect GKP error correction on every GKP qubit. The left plot in Fig. 8 presents our numerical data in this scenario, showing a crossing point at $\sigma_0^c \approx 0.54$.

To use the GKP error information, we use Eq. (17) for the probability of an X error conditioned on the outcome $q \in [-\pi, \pi)$. Including normalization, this probability reads [37]

$$\mathbb{P}(1|q) = \frac{\sum_{k\in\mathbb{Z}} \mathbb{P}_\sigma(-2\pi + q + 4\pi k)}{\sum_{k\in\mathbb{Z}} \mathbb{P}_\sigma(q + 2\pi k)}. \tag{32}$$

To use these expressions we replace $k \in \mathbb{Z}$ by the corresponding sum with a cutoff K, restricting the summation to the interval $-K \leqslant k \leqslant K$. This is warranted since the Gaussian weight for large k is small. In the numerics we used the cutoff $K = 3$.

We numerically simulate the following process. For each toric code qubit, (ij), we first generate a shift error ϵ_{ij} according to the Gaussian distribution which leads to a GKP syndrome value $q_{ij} \in [-\pi, \pi)$. Given q_{ij}, we infer a correction which may give rise to an X error on qubit (ij). We evaluate the Z-checks of the toric code given this collection of errors and perform a minimum-weight matching algorithm to pair up the toric code defects. Logical failure is determined when the toric decoder makes a logical X error on any of the two logical qubits of the toric code. To use the information about the logical error rates $\mathbb{P}(1|q_{ij})$, for each qubit (ij), we define a weight:

$$w_{ij} = \log\left[\frac{1 - \mathbb{P}(1|q_{ij})}{\mathbb{P}(1|q_{ij})}\right]. \tag{33}$$

Then, we define a new weighted graph $G = (V, E)$, whose vertices, $p \in V$, are plaquette defects from the toric code graph and whose edges constitute the complete graph. Given an edge, $(p, p') \in E$, its weight $\omega_{p,p'}$ is the minimum weight of a path on the dual of the toric code graph connecting the defect plaquettes p and p'. Here, the path weight, $\omega_{p,p'}$, is the sum of the weights, w_{ij}, of all edges crossed by the path. Minimum-weight-matching (Blossom) algorithm is then run on this ω-weighted graph G, leading to a matching of defects and thus an inferred X error.

Specifically, we used Dijkstra's algorithm for finding a minimum-weight path in a weighted graph as provided by the Python library `Graph-tools` [38], and the minimum-weight matching algorithm from the C++ library `BlossomV` [39]. The process of sampling from shift errors is repeated many times; the logical error rate plotted in Fig. 8 is given by the fraction of runs which result in logical failure over the total number of runs.

V. NOISY MEASUREMENTS: 3D SPACE-TIME DECODING

A. Error model

In this section we consider how to use both GKP and toric code error information when both error correction steps are noisy, using repeated syndrome measurements. This is the full error model in Fig. 7, which represents one complete QEC cycle. We only consider p-type shift errors (inducing shifts in q), the initial state at $t = 0$ is assumed to be perfect, and the last of M rounds of measurements noise-free, both for the GKP and the toric code ancillas. We will address the question of whether or not there is a *decodable* phase in the space of parameters, such that by increasing the size of the code and the number of measurement cycles, the probability of a logical error can be made arbitrarily small.

To visualize errors of different origin, for the M-times repeated measurement of the toric code on an $L \times L$ square lattice, it is convenient to consider a three-dimensional cubic lattice, with periodic boundary conditions along x and y directions, separated into horizontal layers. Each layer corresponds to a measurement round $t \in [M]$; see Fig. 9. In each time layer t, we use the same notations and conventions of directed edges and vector quantities as in Sec. IV A. Hence we associate a

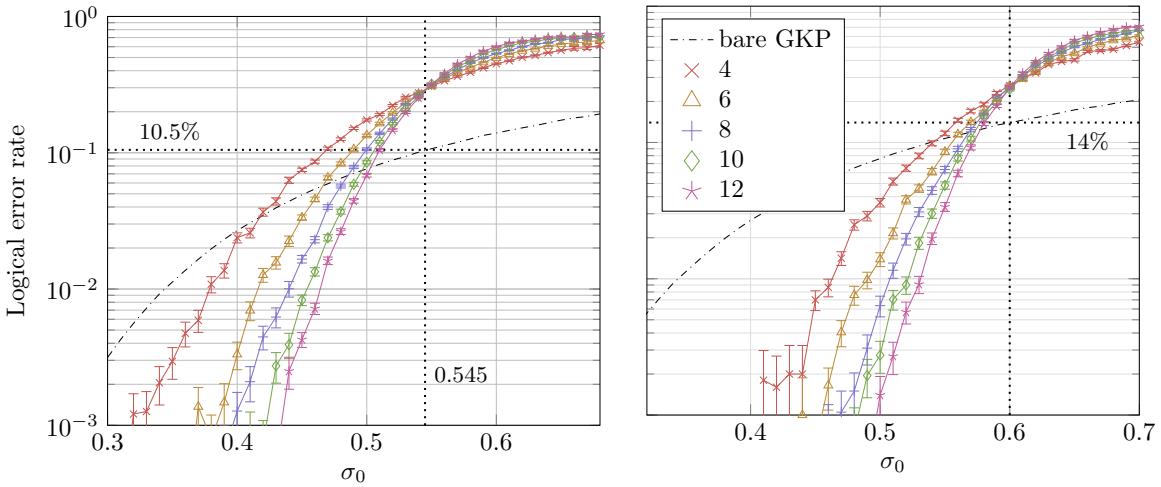

FIG. 8. Threshold comparison between decoding with or without GKP error information. On the left, the simulation only takes the average error rate into account and one obtains a threshold between $\sigma_0 \approx 0.54$ and $\sigma_0 \approx 0.55$ corresponding to $P(X) \approx 10\%$ and $P(X) \approx 10.7\%$, respectively. On the right, the simulation takes the GKP error information into account. In this case, the crossing point is around $\sigma_0^c \approx 0.6$ corresponding to $P(X) \approx 14\%$. The data are labeled by the distance of the toric code. "Bare GKP" is the logical error rate for a single GKP qubit whose errors are processed perfectly without measurement errors (green line in Fig. 5).

GKP data qubit oscillator with each edge (ij) of the square lattice.

We thus denote the shift occurring on a data oscillator just before the measurement at time t as the shift $\epsilon_{ij}(t)$ (induced by channel $\mathcal{N}$). Since the shifts accumulate, the net shift on the oscillator at bond ij just before the measurement at time t is [cf. Eq. (15)]

$$\phi_{ij}(t) \equiv \sum_{t'=1}^{t} \epsilon_{ij}(t'). \tag{34}$$

Furthermore, we denote the GKP measurement error of this oscillator at time t as $\delta_{ij}(t)$. This is the shift error on the ancilla inside the corresponding $\mathrm{EC}_{\mathrm{GKP}}$ unit (induced by channel $\mathcal{N}_{\mathrm{M}}$). With these notations, we can write for the GKP syndrome $q_{ij}^{\mathrm{GKP}}(t)$ at time t,

$$q_{ij}^{\mathrm{GKP}}(t) = \delta_{ij}(t) + \phi_{ij}(t) \bmod 2\pi. \tag{35}$$

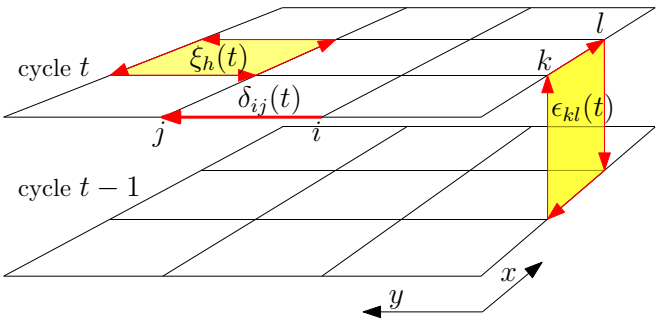

FIG. 9. Notations for repeated toric-GKP errors. Data oscillators are located on the bonds of the square lattice. A GKP measurement error for the bond $b \equiv (ij,t)$ in the measurement cycle t is denoted as $\delta_{ij}(t) \equiv \delta_b$, while the corresponding measurement error for the toric code generator on horizontal plaquette $p \equiv (h,t)$ is denoted as $\xi_h(t) \equiv \xi_p$. A data qubit error $\epsilon_{kl}(t) \equiv \epsilon_p$ is associated with the vertical plaquette p directly below the bond kl in the layer t.

In addition, we have the toric code syndrome. Specifically, we consider the toric code plaquette operators; see Fig. 6 and Eq. (31). The result of the toric code syndrome measurement on the plaquette $h \equiv (ijkl)$ at time t is

$$q_h^{\mathrm{tor}}(t) = \xi_h(t) + (\nabla \times \phi)_h(t) \bmod 4\pi, \tag{36}$$

where the bond vectors $\phi_{ij}(t)$ are the accumulated errors in Eq. (34), and $\xi_h(t)$ is the plaquette measurement error (induced by channel $\mathcal{N}_{\mathrm{T}}$). Note that, unlike for the GKP measurements, the syndrome $\boldsymbol{q}^{\mathrm{tor}}$ is measured modulo 4π, since the ancilla starts in the state $|\bar{0}\rangle$.

Since we assume that the measurement errors in the last layer, $t = M$, are absent, we have $\xi_h(M) = \delta_{ij}(M) = 0$. Writing the product of the corresponding probability densities, we obtain an analog of Eq. (26) for the effective energy

$$\begin{aligned} H(\boldsymbol{\phi};\boldsymbol{q}) = &\sum_{t=1}^{M} \sum_{\langle ij \rangle} \frac{[\phi_{ij}(t) - \phi_{ij}(t-1)]^2}{2\sigma^2} \\ &+ \sum_{t=1}^{M-1} \sum_{\langle ij \rangle} V_{\sigma_{\mathrm{M}}}\left[q_{ij}^{\mathrm{GKP}}(t) - \phi_{ij}(t)\right] \\ &+ \sum_{t=1}^{M-1} \sum_{h} V_{\sigma_{\mathrm{T}}/2}\left[\frac{q_h^{\mathrm{tor}}(t) - (\nabla \times \boldsymbol{\phi})_h(t)}{2}\right], \end{aligned} \tag{37}$$

which depends on the accumulated field $\boldsymbol{\phi}$ with components $\phi_{ij}(t)$ and on the total measured syndrome $\boldsymbol{q} \equiv \{\boldsymbol{q}^{\mathrm{GKP}}, \boldsymbol{q}^{\mathrm{tor}}\}$. Here, $\sum_{\langle ij \rangle}$ indicates a summation over all bonds of the square lattice, the summation over h runs over all square-lattice faces (horizontal), and the structure of the last term accounts for the 4π-periodicity of toric syndrome measurements; see Eq. (36).

The energy in Eq. (37) defines the conditional probability $\mathbb{P}(\boldsymbol{\phi}|\boldsymbol{q}) \propto \exp[-H(\boldsymbol{\phi};\boldsymbol{q})]$, up to a normalization factor. The measurements in the last time-layer, $t = M$, constrain the values $\phi_{ij}(M)$ as follows. From the GKP syndrome we have,

as in Sec. III,

$$\phi_{ij}(M) = q_{ij}^{\mathrm{GKP}}(M) - 2\pi k_{ij}, \tag{38}$$

while the toric syndrome for each square-lattice face $h = (ijkl)$ gives

$$\begin{aligned}(\nabla \times \phi)_h(M) &\equiv \phi_{ij}(M) + \phi_{jk}(M) + \phi_{kl}(M) + \phi_{li}(M) \\ &= q_h^{\mathrm{tor}}(M) - 4\pi k_h. \end{aligned} \tag{39}$$

These equations can be solved to find a last-layer binary candidate error vector $\boldsymbol{b} \in \mathbb{F}_2^{2L^2}$, whose components $b_{ij} \equiv k_{ij} \bmod 2$ give the parity of the integer shifts k_{ij} in Eq. (38). Just as for the usual toric code, Eqs. (38) and (39) determine $\boldsymbol{b}$ up to arbitrary cycles on the dual lattice, i.e., X-stabilizers (homologically trivial) and X logical errors (homologically nontrivial) of the toric code. Adding trivial cycles to $\boldsymbol{b}$ gives another equivalent last-layer candidate which should be summed as part of the same sector. A trivial cycle can be written as the gradient of a binary field, $\boldsymbol{b}_{ij}^{\mathrm{triv}} = n_j - n_i$, with $n_i \in \mathbb{F}_2$. To turn $\boldsymbol{b}$ into an inequivalent last-layer candidate error, one should add a homologically nontrivial cycle, $\boldsymbol{c} \in \mathbb{F}_2^{2L^2}$.

We can now write explicitly the partition function $Z_{\boldsymbol{c}}(\boldsymbol{b}|\boldsymbol{q})$, equivalent to Eq. (7), which determines the conditional probability, given the measurement outcomes $\boldsymbol{q}$ of the equivalence class of last-layer candidate error $[\boldsymbol{b} + \boldsymbol{c}]$,

$$\begin{aligned} Z_{\boldsymbol{c}}(\boldsymbol{b}|\boldsymbol{q}) = N'^{-1} \int d\phi\, \mathrm{e}^{-H(\phi;\boldsymbol{q})} \prod_{\langle ij \rangle} \sum_{\{n_i \in \mathbb{F}_2\}} \sum_{m_{ij} \in \mathbb{Z}} \\ \times \delta\big[\phi_{ij}(M) - q_{ij}^{\mathrm{GKP}}(M) + 2\pi(b_{ij} + c_{ij} - n_j \\ + n_i + 2m_{ij})\big]. \end{aligned} \tag{40}$$

For ML decoding, given a $\boldsymbol{b}$ which satisfies Eqs. (38) and (39), one needs to compare $Z_{\boldsymbol{c}}(\boldsymbol{b}|\boldsymbol{q})$ for different $\boldsymbol{c} \in \mathbb{F}_2^{2L^2}$ which are inequivalent binary codewords of the toric code, i.e., the three homologically nontrivial domain walls on the square lattice or the trivial vector. ML decoding then prescribes that we choose the error $\boldsymbol{b} + \boldsymbol{c}$ as the correction where $\boldsymbol{c}$ has the largest partition function $Z_{\boldsymbol{c}}(\boldsymbol{b}|\boldsymbol{q})$.

B. Equivalent formulation with $U(1)$ symmetry

The partition function in Eq. (40) with the Hamiltonian in Eq. (37), as a statistical-mechanical model, is not so convenient to analyze, since the components of the syndrome are not independent of each other. So, we will consider an equivalent form of the partition function, that explicitly depends on the data errors $\epsilon_{ij}(t) \equiv \epsilon_p$, the measurement errors $\delta_{ij}(t) \equiv \delta_b$, and the toric code measurement errors $\xi_h(t) \equiv \xi_p$. We group all these errors into one error record $\boldsymbol{e} = \{\boldsymbol{\epsilon}, \boldsymbol{\delta}, \boldsymbol{\xi}\}$. Any error $\boldsymbol{e}'$ that is equivalent to $\boldsymbol{e}$ can be obtained by adding, so to say, stabilizer generators of the space-time code. This can be expressed using a 2π-periodic vector field $\boldsymbol{A}$ whose components are real-valued on horizontal bonds in layers $t \in [M-1]$, and $\{0, \pi\}$-valued on the vertical bonds connecting layers $t-1$ and t for all $t \in [M]$. For horizontal bonds b in layers $t = 0$ and $t = M$, $A_b = 0$. With these notations, we can express the partition function in Eq. (7) as

$$\begin{aligned} H(\boldsymbol{A};\boldsymbol{e}) = \sum_{b \| xy} V_{\sigma_{\mathrm{M}}}(\delta_b - 2A_b) + \sum_{p \| xy} V_{\sigma_{\mathrm{T}}/2}\left[\frac{\xi_p}{2} - (\nabla \times A)_p\right] \\ + \sum_{p \perp xy} V_{\sigma/2}\left[\frac{\epsilon_p}{2} + (\nabla \times A)_p\right], \end{aligned} \tag{41}$$

$$Z_{\boldsymbol{0}}(\boldsymbol{e}) = N''^{-1} \sum_{A_b \in \{0,\pi\}: b \perp xy} \prod_{b \| xy} \int_{-\pi}^{\pi} dA_b\ e^{-H(\boldsymbol{A};\boldsymbol{e})}, \tag{42}$$

with some normalization N''. To derive these equations, we determine the shift errors that leave the syndrome record $\{\boldsymbol{q}^{\mathrm{GKP}}, \boldsymbol{q}^{\mathrm{tor}}\}$ unchanged without inducing a logical error. These are called the gauge degrees of freedom and form the stabilizer group for the space-time code. In our case there are five types of gauge degrees of freedom, four discrete and one continuous. The discrete ones are genuine symmetries of the quantum states involved in the code or measurement circuits. Namely, the input state of an ancilla in the GKP measurement circuit in Fig. 2 is stabilized by an X operator, whose action is equivalent to a 2π-shift of the corresponding GKP measurement error $\delta_{ij}(t)$. Similarly, application of an $X^2 = S_p$ GKP stabilizer generator to a data qubit or a toric-code ancilla in Fig. 7 is equivalent to a 4π-shift of the corresponding error, $\epsilon_{ij}(t)$ or $\xi_h(t)$, respectively. We also have toric code vertex operators $\hat{A}_j$ [Eq. (31)] whose action corresponds to simultaneous 2π-shifts on the four adjacent qubits, $\epsilon_{ij}(t) \to \epsilon_{ij}(t) + 2\pi$. This discrete gauge freedom will be captured by the two-valued field A_b on the vertical bonds.

The only continuous degree of freedom is a space-time one: it corresponds to adding a continuous shift a on a data oscillator at some time step and then canceling it at the next time step, while hiding the shift from the adjacent GKP and toric syndrome measurements by adding the shifts $\pm a$ as necessary on the corresponding ancillas.

When applied to the Gaussian distribution, the discrete local shifts are responsible for forming the Villain potentials Eq. (23) in the effective Hamiltonian Eq. (41), where additional rescaling in the last two terms was necessary to account for the 4π-periodicity. The remaining two degrees of freedom are represented by the vertical (discrete) and horizontal (continuous) components of the doubled vector potential $2\boldsymbol{A}$. The scale of the vector potential $\boldsymbol{A}$ was chosen to make easier contact with previous literature on related models. With this choice, adding a π-shift to a component of $\boldsymbol{A}$ correspond to a GKP X-logical, which was a 2π-shift in the previous sections.

Equations (41) and (42) have to be supplemented with the appropriate boundary conditions to be used in decoding. Given the toric-code codeword $\boldsymbol{c}$, to calculate $Z_{\boldsymbol{c}}(\boldsymbol{e})$, corresponding to the sector $[\boldsymbol{e} + \boldsymbol{c}]$, one can add $2\pi\boldsymbol{c}$ to the data error in the top layer, $t = M$. This can be achieved by introducing a fixed nonzero vector potential in this layer, namely, $A_b = \pi c_{ij}$ for all top-layer horizontal bonds $b = (ij, t = M)$, instead of zero for the trivial sector.

Equations (41) and (42) look very different from the equivalent form that we first derived, i.e., Eqs. (37) and (40). A map between these two formulations is given in Appendix C.

C. Anisotropic charge-two $U(1)$ gauge model with flux disorder

We would like to get some intuition about the constructed $U(1)$-symmetric model and its features that are relevant for decoding. To this end, we are going to relax the constraint $A_b \in \{0, \pi\}$ for vertical bonds and consider the following anisotropic charge-two Villain $U(1)$ model in three dimensions, with quenched uncorrelated gauge and flux disorder,

$$H = \sum_{b\parallel xy} V_{\sigma_{\rm M}}(2A_b - \delta_b) + \sum_{b\perp xy} V_{\eta}(2A_b - \delta_b) + \sum_{p\parallel xy} V_{\sigma_{\rm T}/2}[(\nabla \times A)_p - \xi_p/2] + \sum_{p\perp xy} V_{\sigma/2}[(\nabla \times A)_p + \epsilon_p/2]. \tag{43}$$

The model Eqs. (41) and (42) are recovered with the help of the symmetry of the Villain potential, $V_\sigma(x) = V_\sigma(-x)$, by setting $\delta_b = 0$ for vertical bonds, $b \perp xy$, and taking the limit $\eta \to +0$, in which case the field A_b along the vertical bonds be only allowed to take the values 0 or π.

In addition to making all components of the vector field $\boldsymbol{A}$ continuous, we will also relax the constraint on the parameters of the quenched disorder. Specifically, instead of using the components of the fields $\{\boldsymbol{\epsilon}, \boldsymbol{\delta}, \boldsymbol{\xi}\}$ as normally distributed with specific root-mean-square deviations σ, $\sigma_{\rm M}$, and $\sigma_{\rm T}$, respectively, we are going to treat the parameters of the disorder as independent from the parameters in the Hamiltonian Eq. (43). This is similar to the trick used originally for qubit-based surface codes [22], where only the Nishimori line on the phase diagram of the disordered random-bond Ising model corresponds to ML decoding, while points away from that line correspond to a decoder given an incorrect input information; see Eq. (12).

We first examine the parent model Eq. (43) in the absence of background fields, by setting all $\delta_b = \epsilon_p = \xi_p = 0$. The case without anisotropy is relatively well studied, that is $\eta = \sigma_{\rm M}$ and $\sigma = \sigma_{\rm T}$. The Wilson Hamiltonian of the compact charge-q $U(1)$ lattice gauge model reads

$$H = -\kappa \sum_{\langle uv\rangle} \cos(\theta_u - \theta_v + qA_{uv}) - \lambda \sum_p \cos(\nabla \times A)_p, \qquad \kappa = \frac{1}{\sigma_{\rm M}^2}, \quad \lambda = \frac{4}{\sigma^2}, \tag{44}$$

where we restored local gauge symmetry $\theta_v \to \theta_v - q\chi_v$, $A_{uv} \to A_{uv} + \chi_u - \chi_v$ by adding a scalar *matter field*, $U(1)$ phases θ_v on the vertices of the lattice. Notice that the phases θ_v can be always suppressed by a gauge transformation with $\chi_v = q^{-1}\theta_v$. For $U(1)$ symmetry, the charge q in Eq. (44) must be an integer; our original model with the Hamiltonian Eq. (43) corresponds to $q = 2$. In application to this model, the boundary conditions for the sectors with nontrivial codewords $\boldsymbol{c} \not\simeq \boldsymbol{0}$ [see discussion below Eq. (42)] are equivalent to an externally applied uniform magnetic field $B = (\nabla \times A)_p$ (flux per plaquette), with the total flux of π piercing the system along x, y, or both directions. Quite generally, when the couplings κ and λ are sufficiently small, the net magnetic field remains uniform on average, with the total free-energy cost, see Eq. (D5) in the Appendix, proportional to the volume times B^2. For a system with the volume $V = L^2M$ and $B = \pi/(LM)$, this gives the free-energy cost $\Delta_B F \propto 1/M$, vanishing in the large-system limit. The situation is different in the *Meissner phase*, analogous in properties to that in type-II superconductors, where the magnetic field is expelled from the bulk, and is forced into *vortices* (vortex lines) which can carry a flux quantized in the units of $2\pi/q$. Such a vortex is a topological excitation, meaning that it cannot disappear without moving to the system boundary or annihilating with another vortex that carries the opposite flux, and it has a nonzero line tension (finite or logarithmically divergent with the system size), which gives a free-energy cost proportional to the system size, $\Delta_B F \propto L$.

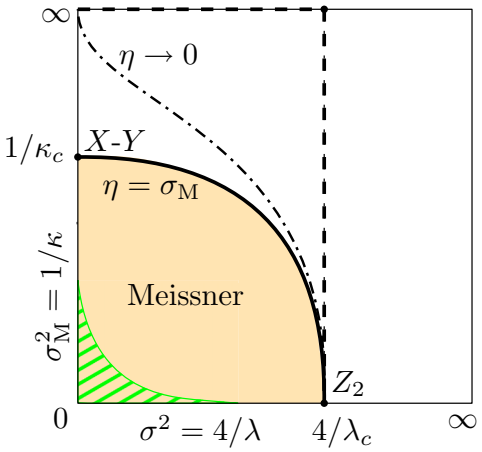

FIG. 10. Schematic phase diagram of the 3D anisotropic $q = 2$ Villain gauge model Eq. (43) with $\sigma_{\rm T} = \sigma$. Solid line bounds the Meissner phase (light orange shading) in the clean isotropic limit, $\eta = \sigma_{\rm M}$, which corresponds to the Villain version of the model Eq. (44). With $\eta \to 0$, decodability condition for the toric-GKP code with $\sigma = \sigma_{\rm T}$ can not be satisfied in the region to the right from the dashed line, $\sigma^2 > 4/\lambda_c(\mathbb{Z}_2)$, and along the upper boundary, $\sigma_{\rm M} \to \infty$, and it is satisfied along the left boundary, $\sigma \to 0$. We therefore expect the boundary of the Meissner phase in the clean $\eta \to 0$ limit as shown with dash-dotted line; this region includes the entire Meissner phase of the isotropic model. The green-hatched region represents the expected location of the ML decodable phase for the toric-GKP code; the sign of the curvature matches the bound in Ref. [24].

The 3D lattice model Eq. (44) (along with related non-Abelian models) has been first discussed by Fradkin and Shenker [40] as a toy model for quark confinement. Subsequently, both the model Eq. (44) and its Villain version have been studied analytically and numerically in a number of papers, e.g., Refs. [41–47]. The conclusion is that the model in 3D has only two phases; see Fig. 10. The weak-coupling phase is characterized by the area law in the Wilson loop correlator and the absence of the Meissner effect. In comparison, the strong-coupling phase, which requires both λ and κ sufficiently large, $\lambda > \lambda_c(q) > 0$, $\kappa > \kappa_c > 0$, is characterized by the presence of both the perimeter law in the Wilson loop and the Meissner effect. Here, $\lambda_c(q)$ corresponds to the limit of $\kappa \to \infty$, which forces the gauge field A_{uv} to take values in $2\pi/q$ times an element of $\mathbb{Z}_q$. In the case $q = 2$, the corresponding critical point [48] $\lambda_c(2) \approx 0.7613$ is that of the

three-dimensional $\mathbb{Z}_2$ lattice gauge theory [49]. Similarly, in the limit $\lambda \to \infty$ the fluxes are all frozen to zero, the charge q is irrelevant and the remaining degrees of freedom are the on-site phases θ_v. This model is known as the X-Y model, and its critical point in 3D is at $\kappa_c \approx 0.453$ (or $\kappa_c^{(\text{Villain})} \approx 0.4542$ for the corresponding Villain model [50]). Such a phase diagram shape with no reentrance as a function of either variable is consistent with the monotonicity of the correlation functions and free-energy increments which follow from generalized GKS inequalities; see Appendix D.

We should notice that the perimeter law in the Wilson loop and the Meissner effect do not necessarily come together; examples are given by compact gauge models similar to Eq. (44) in $D \geqslant 4$ which also have small-κ large-λ phases characterized by the perimeter law but no Meissner effect [40,51]. Of course, it is the Meissner phase that is associated with the formation of magnetic vortices with a nonzero line tension.

What do these results tell us about the anisotropic model Eq. (43) of interest, in particular, about the singular limit $\eta \to +0$? To answer these questions, we notice that, in the absence of disorder, both the correlation functions and the response to external magnetic field (existence of the Meissner effect) are monotonically nondecreasing with respect to any coupling. This follows from general correlation inequalities which are briefly discussed in Appendix D. Moreover, these inequalities also predict an upper bound, see Eq. (D6) in the Appendix, on the ML decoding probability in terms of a similar quantity defined in the absence of disorder. It follows that finite vortex line tension in the clean (Meissner effect) limit is a necessary condition for perfect decoding.

Since decreasing η corresponds to increasing some of the couplings, the entire strong-coupling (Meissner) phase of the 3D lattice gauge model with $q = 2$ should be inside of the corresponding phase of the model Eq. (42) with $\sigma_{\text{T}} \leqslant \sigma$ and the values of $\sigma_{\text{M}}, \sigma$ given by the map in Eq. (44). Second, this phase cannot exist for $\sigma^2 > 4/\lambda_c(2)$, the limit $\kappa \to \infty$ which corresponds to taking *both* η and σ_{M} to zero.

Furthermore, if we started with the model Eq. (37) in the limit of unusable GKP syndrome, $\sigma_{\text{M}} \to \infty$, the first term in Eq. (43) would be absent. In this case the continuous gauge symmetry $A_{uv} \to A_{uv} + \chi_u - \chi_v$ is not broken, which is sufficient to recover a continuous field A_b along the vertical bonds $b \perp xy$; with $\sigma = \sigma_{\text{T}}$, the resulting model is the Villain version of Eq. (44) with $\kappa = 0$. According to Polyakov's argument [51], only one phase is expected in this limit; we expect no Meissner effect, and no decoding threshold.

However, the large-λ (small-σ) limit of the model Eq. (44) corresponds to all fluxes $(\nabla \times A)_p$ frozen in the minimum-energy configuration. The remaining degrees of freedom are the phases in the first term, which gives an X-Y model. However, if we look at the model Eq. (43) with $\sigma = \sigma_{\text{T}} \to 0$, in the singular anisotropic limit $\eta \to 0$, the phases $\theta_v \equiv \theta_{i,t}$ at the same square lattice position i are forced to fluctuate together, which gives arbitrarily large effective X-Y coupling $\kappa_{\text{eff}} = M/\sigma_{\text{M}}^2$ as $M \to \infty$. Assuming this argument also holds for $\sigma = \sigma_{\text{T}}$ small but finite, we expect the phase line as shown in Fig. 10 with a dot-dashed line, with the region below it in the Meissner phase.

A different version of this argument can be obtained by examining the Hamiltonian in the form of Eq. (37), with $\boldsymbol{q} = \boldsymbol{0}$. With σ small, the fields $\phi_{ij}(t)$ in the neighboring t-layers are forced to move together, which is equivalent to increasing the couplings for the remaining terms. The resulting model is a two-dimensional version of the gauge model Eq. (44), with in-plane vector potential $\boldsymbol{A}$. Just like its 3D counterpart, this model is in a disordered phase except when the fluxes are suppressed in the limit $\sigma_{\text{T}} \to 0$, which gives a 2D X-Y model. With σ sufficiently small, the effective in-plane coupling $\kappa_{\text{eff}}^{(\text{2D})}$ can be made arbitrarily large, driving the model below the BKT transition.

We have discussed the expected phase boundary of the model Eqs. (41) and (42) in the absence of disorder, $\boldsymbol{e} = \boldsymbol{0}$. We expect a Meissner phase to survive with $\boldsymbol{e} \neq \boldsymbol{0}$. The basic effect of a weak disorder on a topological excitation like a vortex line is to force its random displacements. The displacements can be accounted in a simple linear approximation up to certain distance scale called the Larkin length [52]. Even though the displacements become nonlinear beyond this scale, signifying the onset of glassiness, with weak enough disorder, topological excitations are not expected to be generated; this is called Bragg glass phase of an elastic solid [53]. The absence of topological excitations would indicate a divergent free-energy cost for a π vortex excitation in the Meissner phase of the model Eqs. (41) and (42), with sufficiently weak disorder.

We thus expect a decodable phase for a toric-GKP code to exist in a finite region at sufficiently small σ, σ_{M}, and σ_{T}. In the limit $\sigma_{\text{M}} \to 0$, this phase should go continuously into a decodable phase of the regular (qubit) toric code [22]. These expectations are confirmed by our numerical results with two (suboptimal) decoders presented in the next section.

D. Decoder and numerical results

Maximum likelihood decoding can be done by comparing conditional probabilities in different sectors; see Eq. (7). Just as in the case of a single GKP qubit, Gaussian integrations in the relevant partition functions, $Z_{\boldsymbol{c}}(\boldsymbol{b}|\boldsymbol{q})$, in Eq. (40), or its equivalent form $Z_{\boldsymbol{c}}(\boldsymbol{e})$, in Eq. (42), can be carried through exactly. This would leave expressions similar to Eq. (22), with an additional summation over $2L^2$ binary spins. In principle, such expressions can be evaluated using Monte Carlo sampling techniques. In practice, the complexity of such a calculation is expected to be high, because the corresponding coupling matrix is not sparse, just as the matrix A in Eq. (22) is not sparse; see Appendix B. For this reason, we have not attempted ML decoding for toric-GKP codes.

We have constructed several decoders which approximate ME decoding. The idea is to find a configuration of the field $\boldsymbol{A}$ minimizing the Wilson version of the Hamiltonian Eq. (41), i.e., with Villain potentials replaced by cosines, by decomposing it into a continuous part (to be guessed or found using a local minimization algorithm), and a binary field which represents frustration, to be found using minimum weight matching. This decomposition relies on the analysis of the Hamiltonian Eq. (41) in the limit of perfect GKP measurements, $\sigma_{\text{M}} \to 0$.

1. ME decoding in the limit of perfect GKP measurements

Let us consider what happens with our model Eqs. (41) and (42) in the limit $\sigma_{\mathrm{M}} \to 0$. First, in this limit all GKP measurement errors δ_b vanish with probability one. Second, the first term in Eq. (41) forces $A_b \in \{0, \pi\}$ for all horizontal bonds, the same as we already had for vertical bonds. This forces all plaquette fluxes to take integer values times π. We show here that in this limit we recover a version of the random-plaquette gauge model (RPGM) associated with decoding the usual qubit toric code in the presence of (toric) syndrome measurement errors [54].

Since the limit $\sigma_{\mathrm{M}} \to 0$ makes the vector potential $A_b \in \{0, \pi\}$, and in turn the plaquette flux $B_p \equiv (\nabla \times A)_p$, discrete, then one can interpret it as a spin degree of freedom, using the fact that the Villain potential is 2π-periodic as well as even. Indeed, considering for example a vertical plaquette, $p \perp xy$, in Eq. (41), one can write

$$V_{\sigma/2}\left(\frac{\epsilon_p}{2} + B_p\right) = -\tau_p(\boldsymbol{e})\mathrm{e}^{iB_p} + \text{const},$$

$$p \perp xy, \quad \tau_p(\boldsymbol{e}) \equiv \frac{1}{2}\left[V_{\sigma/2}\left(\frac{\epsilon_p}{2} - \pi\right) - V_{\sigma/2}\left(\frac{\epsilon_p}{2}\right)\right] \approx \frac{4}{\sigma^2}\cos\left(\frac{\epsilon_p}{2}\right). \tag{45}$$

The additive constant has no effect and can be ignored. Similarly for horizontal plaquettes, $p \parallel xy$, where one obtains the weights

$$p \parallel xy, \quad \tau_p(\boldsymbol{e}) \equiv \frac{1}{2}\left[V_{\sigma_{\mathrm{T}}/2}\left(\frac{\xi_p}{2} - \pi\right) - V_{\sigma_{\mathrm{T}}/2}\left(\frac{\xi_p}{2}\right)\right] \approx \frac{4}{\sigma_{\mathrm{T}}^2}\cos\left(\frac{\xi_p}{2}\right). \tag{46}$$

Then, if one defines from $\boldsymbol{A}$ some Ising spins, $\boldsymbol{s}$, using for each bond, $s_b \equiv \mathrm{e}^{iA_b} \in \{-1, 1\}$, one obtains, in place of Eq. (41), a RPGM very similar to that in Ref. [54],

$$H(\boldsymbol{s}; \boldsymbol{e}) = -\sum_p \tau_p(\boldsymbol{e}) u_p, \quad u_p \equiv \prod_{b \in p} s_b. \tag{47}$$

Unlike in the usual RBGM obtained for the qubit toric code [54] where plaquette weights can take only two values, $\tau_p = \pm J$, here quenched randomness leads to a continuous distribution of the weights. This model is similar to the 2D random-bond Ising model constructed in Sec. IV B for decoding a toric-GKP code in the channel setting, also in the limit $\sigma_{\mathrm{M}} = 0$. In fact, the weights concerning data errors in Eq. (45) (without the cosine approximation), are equivalent to those given by Eqs. (32) and (33). Similar to its counterpart with sign disorder, the phase-diagram of the RPGM with the Hamiltonian Eq. (47) will show a transition from an ordered to disordered phase as the temperature β^{-1} or the strength of the quenched disorder is increased. If we increase the disorder σ along a line with any fixed ratio $r = \sigma_{\mathrm{T}}/\sigma$, then a version of the standard argument [55] shows that no ordered phase can exist beyond the critical value of σ reached along the "Nishimori line," $\beta = 1$ [cf. Eq. (12)]. We expect that it is this critical value of σ that is associated with the memory phase transition.

An important quantity that governs the structure of the minimum of the RPGM Hamiltonian Eq. (47) is *frustration*. We call a cube frustrated when it has an odd number of its boundary plaquettes p with $\tau_p(\boldsymbol{e}) < 0$. Since every spin, s_b, affects the sign of two plaquettes in a cube, for a frustrated cube, no spin configuration on the edges can simultaneously satisfy all the plaquette terms. For the usual toric codes, frustrated cubes can be readily identified by stabilizer defects without referring to a candidate error. For toric-GKP codes one has to be more careful. In the limit $\sigma_{\mathrm{M}} \to 0$, the frustration cannot be read directly from the value of the toric code syndromes but all candidate errors exhibit the same frustration. So picking any candidate error, $\boldsymbol{e}$, finds it. When $\sigma_{\mathrm{M}} \neq 0$, this is no longer the case, the frustration can change between different candidate errors.

We examine the problem of finding the optimum spin configuration, $\boldsymbol{s}$, which minimizes the RPGM Hamiltonian Eq. (47), given the weights $\tau_p(\boldsymbol{e})$. For a given $\boldsymbol{s}$, we call a plaquette term, $\tau_p(\boldsymbol{e})u_p$, *satisfied* when $\tau_p(\boldsymbol{e})u_p > 0$ and *unsatisfied* when $\tau_p(\boldsymbol{e})u_p < 0$. Necessarily, a frustrated cube is incident to an odd number of unsatisfied plaquettes, in this sense frustrated cubes are the source of unsatisfied plaquettes. Let S be any set of plaquettes such that the frustrated cubes are incident to an odd number of plaquettes in S and the unfrustrated cubes are incident to an even number of plaquettes in S. One has

$$\min_{\boldsymbol{s}} H(\boldsymbol{s}; \boldsymbol{e}) = -\sum_p |\tau_p(\boldsymbol{e})| + 2 \min_S \sum_{p \in S} |\tau_p(\boldsymbol{e})|.$$

Hence, minimum-weight matching on a 3D lattice with vertices representing the frustrated cubes determines the optimal set of unsatisfied plaquettes $S_{\min}$. Given a candidate error $\boldsymbol{e}$ let S_{cand} be the set of plaquettes with $\tau_p(\boldsymbol{e}) < 0$. The candidate error $\boldsymbol{e}$ is now modified using the minimum-weight matching by adding 2π for all plaquettes in $S_{\min}$ as well as adding 2π on all plaquettes in S_{cand} (remember that 4π-shifts are elements of the stabilizer group). This corresponds to an addition of a stabilizer or logical operator to the candidate error $\boldsymbol{e}$, leaving the syndrome unchanged. This modified candidate error is the proposed correction for this decoder.

We implemented the described decoder to minimize the energy Eq. (47), with Villain potentials replaced with cosines; see Eqs. (45) and (46). A very simple estimate of the expected performance of the toric-GKP code in this setting is the following. It is known that the threshold for the toric code is about 2.9% under phenomenological noise with independent X and Z errors [54]. If we assume that all errors are due to the logical error on the underlying GKP qubits, see Fig. 5, then one can ask what σ_0 leads to a probability for an X error (or equivalently Z) equal to 2.9%. This of course depends on the measurement error σ_{M} as well as the decoding method for the GKP qubit. In case of no measurement error ($\sigma_{\mathrm{M}} = 0$), the error probability can be found by averaging $\mathbb{P}_{M=1}(1|q_1)$, see Eq. (17), over the Gaussian error distribution, it is plotted as the green line in Fig. 5. An error probability of 2.9% corresponds to $\sigma_0 \approx 0.41$. Our numerical results are shown in Fig. 11. One can see a crossing point around $\sigma_0 = 0.47$ which can be converted to around 6% error rate per round for a GKP qubit with $\sigma_{\mathrm{M}} = 0$; see Fig. 5. We can conclude that, similarly to the results in Sec. IV B, the continuous-valued

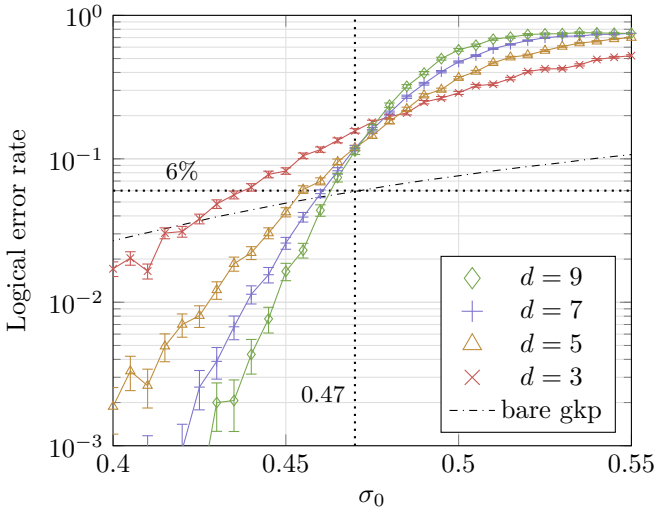

FIG. 11. Numerical results for perfect GKP measurements, with $\sigma = \sigma_T$ and $\sigma_M = 0$. The logical error rate as a function of the bare standard deviation σ_0, see Eq. (14), is shown for toric code distances $d = L$ as indicated in the caption. The vertical dotted line indicates the position of the crossing point in the data, a threshold at $\sigma_0 \approx 0.47$, which corresponds to a logical error probability $p = 6\%$ for single-qubit GKP code with $\sigma_M = 0$. The latter value is recovered as the vertical position on the dash-dotted line which reproduces the green line $\sigma_M = 0$ from Fig. 5 (left). The observed threshold is well above $\sigma_0 = 0.41$ which can be superficially expected from the $p = 2.9\%$ threshold for the 2D toric code under phenomenological error model.

weights $\tau_p(\boldsymbol{e})$, constitute valuable information to the decoder about the likelihood of errors and permit to surpass decoders which do not have access to such information.

2. *Dealing with multiple competing minima*

A difficulty in minimizing the Hamiltonian Eq. (41) is the existence of a large number of competing local minima. This was already the case for a single GKP qubit, which we considered in Sec. III. We saw in the previous section that in the case of perfect GKP measurements, the solution can be found efficiently because the problem is equivalent to a minimum weight matching on a graph. Our approach to minimizing Eq. (41) in the general case will be to decompose the vector potential $A_b \in [-\pi, \pi)$ for horizontal bonds $b = (ij, t)$ into a discrete field $A_b^{(0)} \in \{0, \pi\}$, and an auxiliary continuous field $a_b \in [-\frac{\pi}{2}, \frac{\pi}{2})$,

$$A_b = A_b^{(0)} + a_b + \delta_b/2. \tag{48}$$

For vertical bonds there is no need of such a substitution since the corresponding field is already discrete, see Eq. (42), so we set $a_b = 0$ for all $b \perp xy$. The discrete part of the field, $\boldsymbol{A}^{(0)}$, can be used to define Ising spins, $\boldsymbol{s}$, similarily as before, $s_b \equiv e^{iA_b^{(0)}}$. The definition of the weigths, $\tau_p(\boldsymbol{e})$, has to be extended since now they depend also on $\delta_b \neq 0$ and $a_b \neq 0$, or more precisely the residual fluxes $b_p \equiv (\nabla \times a)_p$. Adding these dependencies, the new weights, $\tau_p(\boldsymbol{a}; \boldsymbol{e})$, given only in their cosine approximation, read

$$\tau_p(\boldsymbol{a}; \boldsymbol{e}) = \frac{4}{\sigma_T^2} \cos\left(\frac{\xi_p}{2} - \frac{1}{2}(\nabla \times \delta)_p - b_p\right), \quad p \| xy, \tag{49}$$

$$\tau_p(\boldsymbol{a}; \boldsymbol{e}) = \frac{4}{\sigma^2} \cos\left(\frac{\epsilon_p}{2} - \frac{1}{2}(\delta_{ij}(t) - \delta_{ij}(t-1)) + b_p\right), \quad p \equiv (ij, t) \perp xy. \tag{50}$$

Rewriting the Hamiltonian Eq. (41), using the substitution Eq. (48) and using the cosine approximation of the Villain potentials gives

$$H(\boldsymbol{a}, \boldsymbol{s}; \boldsymbol{e}) = -\frac{1}{\sigma_M^2} \sum_{b \| xy} \cos(2a_b) - \sum_p \tau_p(\boldsymbol{a}; \boldsymbol{e}) u_p. \tag{51}$$

Then, the minimization over the gauge field $\boldsymbol{A}$, or equivalently over $\{\boldsymbol{a}, \boldsymbol{s}\}$, factors out into minimizing the RPGM similar to Eq. (47), and a minimization over the continuous field $\boldsymbol{a}$.

In addition to the candidate error $\boldsymbol{e}$, the RPGM weights now also depend on $\boldsymbol{a}$ via the the flux field $\boldsymbol{b}$. Moreover, even with the restriction on the auxiliary vector potentials, $|a_b| \leqslant \pi/2$, the corresponding fluxes are not so restricted; in particular, both for horizontal and vertical plaquettes one may have $|b_p| = \pi$, sufficient to flip the sign of the RPGM weight, $\tau_p(\boldsymbol{a}; \boldsymbol{e})$, and flip the frustration of the two adjacent cubes.

Nevertheless, even though frustration depends on configuration of both the spins and the residual fluxes b_p, increasing the number of variables by the substitution Eq. (48) does simplify the minimization problem. First, unlike in the isotropic model Eq. (44), the fields a_b are uniquely defined by the fluxes b_p. Indeed, since the gauge fields a_b are only nonzero on the horizontal bonds, and are zero at the bottom layer, $t = 0$, the gauge is fixed. Furthermore, the GKP terms tend to suppress order-reducing fluctuations by favoring small $|a_b|$. Thus, with σ_M small compared to σ and σ_T, we can hope to find a reasonably good solution just by setting $a_b = 0$.

3. *Actual decoder algorithms and their performance*

To design a syndrome-based decoder with the starting point $a_b = 0$, we first need to come up with a candidate error $\boldsymbol{e} \equiv \boldsymbol{e}(\boldsymbol{q})$. Since we are not doing the full minimization of the corresponding energy, the method to find $\boldsymbol{e}$ will necessarily affect the performance of the resulting decoder.

ALGORITHM 1:

(1) For each plaquette h, starting from $t = M - 1$ down to $t = 1$, set the toric code measurement error $\xi_h(t) = q_h^{\text{tor}}(t) - q_h^{\text{tor}}(t+1)$, to suppress the increments of the toric syndrome. This leaves nonzero toric code syndromes only in the first layer, $q_h^{\text{tor}}(t = 1)$.

(2) Set data errors in the first layer, $\epsilon_{ij}(t = 1)$, to move nonzero toric code syndromes to the left (with $ij \| y$), then down along the leftmost column. Due to the boundary conditions, the sum of all toric code syndromes is 0, meaning that after this procedure, all toric code syndromes are removed including the one in the left-bottom plaquette.

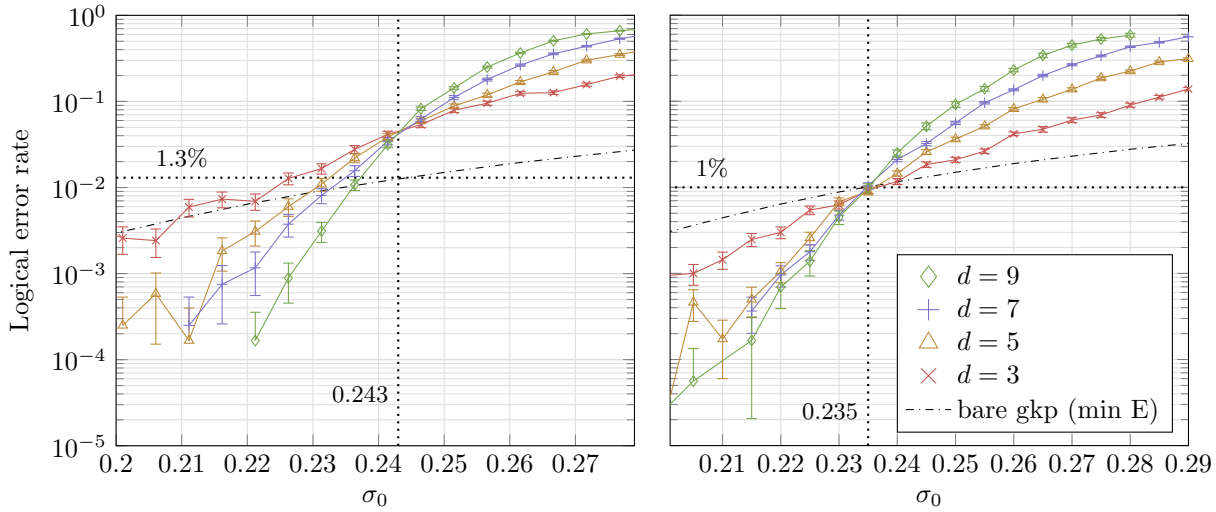

FIG. 12. For both plots the dash-dotted line shows the logical error rate per round for a single GKP qubit corrected with the forward-minimization decoder, with $\sigma = \sigma_T = \sigma_M$, as a function of the bare standard deviation $\sigma_0 = \sigma/2\sqrt{\pi}$. (Left) Numerical results for ALGORITHM 1 without preprocessing of the GKP syndrome information (see text). We observe a crossing point at $\sigma_0 \approx 0.243$ which can be translated to $p = 1.3\%$. (Right) Numerical results for the ALGORITHM 2 with preprocessing of GKP syndrome information (see text). We observe that this decoder improves the logical error rates. However, the improvement being greater for smaller distances, the crossing point moves left to $\sigma_0 \approx 0.235$ which corresponds to an error rate $p = 1\%$.

(3) For each square lattice bond (ij), starting with $t = 1$ up to $t = M - 1$, use Eq. (35) to set GKP errors $\delta_{ij}(t)$ to suppress the GKP syndromes $q_{ij}^{\mathrm{GKP}}(t)$ [without changing $\phi_{ij}(t)$]. This leaves nonzero GKP syndromes only in the top layer, $t = M$, with toric syndromes all zero.

(4) Make a gauge transformation on the top layer data errors, $\epsilon_{ij}(t = M) \to \epsilon_{ij}(t = M) + \chi_j - \chi_i$ to move nonzero GKP syndromes to the left (to $x = 0$) and then down (to $y = 0$). The last step works because Eqs. (38) and (39) are satisfied for the updated syndrome; with $q_h^{\mathrm{tor}}(t = M) = 0$ these guarantee that the GKP syndrome is a gradient.

Having determined the candidate error $\boldsymbol{e}$, we calculate the RPGM weights Eqs. (49) and (50) with $b_p = 0$, and continue with minimum-weight matching decoding described in Sec. V D 1.

The numerical results obtained with the described decoder at $\sigma = \sigma_M = \sigma_T$ are shown in Fig. 12 (left). Despite the fact that this decoder does not make a particularly good use of the GKP syndrome information, $\boldsymbol{q}^{\mathrm{GKP}}$, and does not even try to find a good candidate error, the logical error rate rapidly goes down with increasing code distance and decreasing σ_0 below the crossing point at $\sigma_0 \approx 0.243$. With the forward-minimization decoder on a single GKP qubit with $\sigma = \sigma_M$, this would correspond to a logical error rate of $p \approx 1.3\%$. If the errors only come from having imperfect GKP states, then this also can be translated to having states with at least four photons.

It seems possible that, by making a better use of the GKP syndrome, one should be able to improve this decoder while preserving its computational efficiency. To this end, we tried a *preprocessing* algorithm. The basic idea is, given the syndrome $\boldsymbol{q}^{\mathrm{GKP}}$, to find an initial approximation, $\boldsymbol{e}_0$, for the data errors which would bring back the GKP qubits closer to their code space. Given $\boldsymbol{e}_0$, Algorithm 1 can be used to find an error $\boldsymbol{e}_1$ matching the updated syndrome, after which the RPGM weights can be computed using the full candidate error $\boldsymbol{e}_0 + \boldsymbol{e}_1$. The hope is then that the candidate error found is one for which the first term of the Hamiltonian in Eq. (51) does not need to be minimized anymore. In particular, we tried using our single-oscillator forward-minimization decoder from Sec. III as the preprocessing step. Using it directly produced a degradation of performance, seemingly resulting from the fact that the minimization of the RPGM Hamiltonian also tries to optimize the GKP measurement errors δ_b. Our solution was to drop the measurement errors from the decomposition of the field in Eq. (48), which results in RPGM weights identical to those in Sec. V D 1. Our second decoder can then be summarized as follows.

ALGORITHM 2:

(1) For each bond (ij), use the forward-minimization decoder from $t = 1$ up to $M - 1$ to calculate the accumulated data error $\phi_{ij}(t)$, calculate the corresponding $\phi_{ij}(t = M)$, and then go back from $t = M$ to $t = 1$ with a version of the same algorithm, but using previously found values for a more accurate minimization. This gives the data errors $\boldsymbol{\epsilon}_{ij}(t) = \phi_{ij}(t) - \phi_{ij}(t - 1)$, which we use to define the error vector $\boldsymbol{e}_0 = \{\mathbf{0}, \boldsymbol{\epsilon}, \mathbf{0}\}$.

(2) Calculate the residual syndrome, and run the entire ALGORITHM 1, to calculate the corresponding error $\boldsymbol{e}_1$.

(3) Calculate RPGM weights $\tau_p(\boldsymbol{e})$ using Eqs. (49) and (50) with the combined error $\boldsymbol{e} = \boldsymbol{e}_0 + \boldsymbol{e}_1$ and $\boldsymbol{\delta}$ set to zero.

(4) Use random-weight matching to minimize the RPGM Hamiltonian, and update the error $\boldsymbol{e}$, with the result being the output of the decoder.

The results are shown in Fig. 12 (right). One can observe that for each distance, there is an improvement in the encoded logical error rate, compared to the results from Algorithm 1 on the left of Fig. 12. One can see, for each curve, a higher *pseudothreshold*, i.e., the point below which the logical error rate becomes smaller than the physical one, determined using

the logical error rate for a single GKP qubit given by the forward-minimization decoder. However, this improvement is greater for smaller distances, so the overall crossing point is shifted to the left, indicating a smaller threshold. Specifically, the crossing point is at $\sigma_0 \approx 0.235$ which corresponds to $p = 1\%$ for the single GKP qubit with the forward-minimization decoder.

We should note that even though both simple decoders we tried result in finite thresholds, with substantial reduction of the logical error rates with increasing distance below the crossing points, one could expect better performance. Indeed, the toric code with a phenomenological error model shows a threshold at $p = 2.9\%$. The forward-minimization decoder with $\sigma = \sigma_M$ in Sec. III reaches this error rate at $\sigma_0 \approx 0.28$.

Of course, a direct comparison is technically incorrect: forward-minimization or any other single-oscillator GKP decoder would return highly correlated errors and one cannot expect that the toric code would achieve the same performance. Nevertheless, we expect that adding a minimization step for the continuous part of the potential $\boldsymbol{a}$ in Eq. (48) would significantly improve the performance.

VI. NO-GO RESULT FOR LINEAR OSCILLATOR CODES

We turn to the class of codes defined by continuous subgroups of displacement operators. One can think of linear combinations of position and momentum operators as nullifiers for the code space, i.e., any code state is annihilated by these nullifiers; see Appendix A.

It is known that one cannot distill entanglement from Gaussian states by means of purely Gaussian local operations and classical communication [56,57]. In addition, the authors of Ref. [18] defined a quantity, "entanglement degradation," for any single-mode Gaussian channel, such as the Gaussian displacement channel, and showed that it cannot decrease under Gaussian encoding and decoding. In the setting here we consider any input state which is perfectly encoded into a linear oscillator code. This encoding map, $\mathcal{E}$, is a Gaussian operation as it is a linear transformation of the $\hat{p}$ and $\hat{q}$ variables. After this encoding, the modes go through the Gaussian displacement channel $\mathcal{N}$, see Eq. (5). After this, linear combinations of $\hat{p}$ and $\hat{q}$ are measured to give rise to a syndrome $\boldsymbol{q}$. Again, this operation is Gaussian. Our results thus follow the model considered in Ref. [18]. However, our results do not require the definition of a new quantity but rather give a description of the logical noise model of Eq. (10). Namely, we show in the following Theorem, that it can only lead to an effective squeezing of the original Gaussian displacement channel. Hence, whatever protection is gained in one quadrature, is lost in the other quadrature. This is a property of all linear oscillator codes, CSS or non-CSS, i.e., mixing $\hat{p}$ and $\hat{q}$ quadratures or not, with respect to the Gaussian displacement channel. Since the result is a detailed expression of the logical Gaussian displacement channel, it does not immediately follow from earlier no-go results.

Theorem (No-Go). Let $\mathcal{C}$ be a linear oscillator code on n physical oscillators defined by a set of $n-k$ independent nullifiers, thus encoding k logical oscillators. Let this code undergo independent Gaussian shifts in $\hat{p}$ and $\hat{q}$ of variance σ_0^2 on each of its physical oscillators, followed by a perfect (maximum-likelihood) decoding step. Then the remaining logical displacement noise model, Eq. (10), for logical shift errors $\boldsymbol{\epsilon} \in \mathbb{R}^{2k}$ is

$$\mathbb{P}^{(\mathrm{ML})}(\boldsymbol{\epsilon}) = \left(2\pi\sigma_0^2\right)^{-2k} \exp\left(-\frac{1}{2\sigma_0^2}\boldsymbol{\epsilon}\,\Sigma^{-1}\boldsymbol{\epsilon}^T\right),$$

and the eigenvalues of the covariance matrix, Σ^{-1}, are paired by conjugated logical operators, $(\lambda_j^p, \lambda_j^q)_{j\in[k]}$, such that

$$\forall j \in [k], \quad \lambda_j^p \lambda_j^q = 1.$$

In particular one has $\det \Sigma^{-1} = 1$.

The proof of this Theorem, which is rather lengthy, can be found in Appendix E. The remaining logical displacement noise is obtained by working out Eq. (10), using general properties of linear oscillator codes. The theorem says that for each logical mode, in the basis given by the eigenvectors of Σ, the only effect of the encoding is to squeeze the displacement noise model between the conjugated operators of the mode. The amount of squeezing can depend on code size but it is impossible to reduce the noise in both quadratures.

As a concrete example, we consider the linear oscillator code version of the 2D toric code in Appendix F and show that the squeezing depends simply on the ratio L_x/L_y for a $L_x \times L_y$ toric lattice.

VII. DISCUSSION

In this paper we have made the first strides in tackling the problem of error correction and decoding for the toric-GKP code. Interestingly, the decoding problem maps onto a class of physical continuous-variable models with quenched-disorder that goes beyond the random plaquette gauge model corresponding to toric code decoding [54]. We have presented an efficient minimum-energy decoding method for a single GKP oscillator and an efficient decoder for the toric-GKP code. We have also presented a combination of these two decoders for the toric-GKP code which improves the achieved logical error rates but in greater proportion for the small distances, hence achieving better pseudothresholds but a slightly worse threshold. It would be interesting to design a better decoder to improve this threshold.

An interesting open problem is to study the phase diagram of the toric-GKP code numerically. A particularly interesting question, even in the absence of disorder, is whether the Meissner phase is extended all the way to $\sigma_M \to \infty$ as we conjectured, see Fig. 10, or terminates at a point along the vertical axis. This would indicate the singularity of the limit $\sigma \to 0$. An example of such a numerical study for 3D color codes is Ref. [58].

Future studies could look at the question of decoding coherent errors and/or correcting both p and q-shifts simultaneously. Another question is whether it is possible to handle more realistic noise models, e.g., consider a model of repeated photon loss [5] for a single GKP oscillator. All such different error models will have a particular path integral representation and the idea of choosing an energy-minimizing path can be examined.

A variant of the toric-GKP code is the toric-rotor code. This is the concatenation of a rotor space with integer $\hat{n}$ and

2π-periodic $\hat{\varphi}$ with a rotor toric code whose nullifiers are linear combinations of $\hat{n}$s and $\hat{\varphi}$s of four rotors [59]. When one uses a two-dimensional rotor subspace such as a cat code [12] or just a transmon qubit, one could still express the proper toric code checks in the entire rotor space and examine when a memory phase is present. Note that the difference in such analysis versus the usual toric code analysis is that in this model the effect of leakage errors is automatically included.

A possible realization of a surface code variant of the toric-GKP code is an array of superconducting 2D or 3D resonators. For a code such as Surface-17, this would require $2 \times 17 = 34$ (coplanar) microwave resonators. One can compare this to the transmon + resonators layout for the regular surface code [60,61], in which each CNOT gate is mediated via a coupling bus, hence one uses $8 \times 4 = 32$ such bus resonators for Surface-17.

ACKNOWLEDGMENTS

B.M.T., Y.W., and C.V. acknowledge support by the European Research Council (EQEC, ERC Consolidator Grant No. 682726) as well as a QuantERA grant for the QCDA consortium. L.P.P. acknowledges support from NSF Division of Physics via Grants No. 1416578 and No. 1820939.

APPENDIX A: GENERAL MULTIMODE GKP CODES

Here we provide a mathematical summary of the formalism of continuous-variable stabilizer codes, as introduced by Gottesman, Kitaev, and Preskill (GKP) [3]. We show that these codes contain two main classes, proper GKP codes, encoding discrete information and linear oscillator codes, encoding continuous-variable information. Hybrids between the two classes are also possible but do not seem particularly useful.

Recall that we write a displacement operator on a n-oscillator Hilbert space as follows:

$$\hat{U}(\boldsymbol{e}) \equiv \prod_{k=1}^{n} \mathrm{e}^{iu_k\hat{p}_k + iv_k\hat{q}_k}, \quad \boldsymbol{e} \equiv (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{R}^{2n}. \tag{A1}$$

We also denote by $\hat{O}(\boldsymbol{e})$ the operator in the exponent of $\hat{U}(\boldsymbol{e})$,

$$\hat{O}(\boldsymbol{e}) \equiv \boldsymbol{u}\cdot\hat{\boldsymbol{p}} + \boldsymbol{v}\cdot\hat{\boldsymbol{q}} = \sum_{k=1}^{n} u_k\hat{p}_k + \sum_{k=1}^{n} v_k\hat{q}_k,$$

$$\hat{U}(\boldsymbol{e}) = \exp[i\hat{O}(\boldsymbol{e})]. \tag{A2}$$

These operators, $\hat{U}(\boldsymbol{e})$, follow the product rule

$$\hat{U}(\boldsymbol{e})\hat{U}(\boldsymbol{e}') = \hat{U}(\boldsymbol{e}+\boldsymbol{e}')\mathrm{e}^{i\omega(\boldsymbol{e},\boldsymbol{e}')}, \tag{A3}$$

with the standard symplectic form $\omega(\boldsymbol{e},\boldsymbol{e}') = \boldsymbol{u}\cdot\boldsymbol{v}' - \boldsymbol{v}\cdot\boldsymbol{u}'$.

Take $\mathcal{S} \subset \mathcal{H}_n$ an Abelian subgroup of the displacement operators with no element proportional to the identity with a nontrivial phase, stabilizing some code space, $\mathcal{Q}$. One can simply characterise the structure of such a subgroup. Each element $\hat{S} \in \mathcal{S}$ can be uniquely written as

$$\hat{S} = \mathrm{e}^{i\theta}\hat{U}(\boldsymbol{e}), \tag{A4}$$

for some real vector $\boldsymbol{e} \in \mathbb{R}^{2n}$, and some phase $\theta \in [0, 2\pi)$. Each vector $\boldsymbol{e}$ can only appear with a unique phase θ in $\mathcal{S}$; otherwise, an element proportional to the identity with a nontrivial phase would also be in $\mathcal{S}$. Hence, we have an isomorphism between $\mathcal{S}$ and the additive subgroup $G \subset \mathbb{R}^{2n}$, given by all the vectors, $\boldsymbol{e}$, obtained from the decomposition in Eq. (A4). We can then use the following theorem to characterize the structure of $\mathcal{S}$.

Theorem ([62]). Let G be a closed additive subgroup of $\mathbb{R}^{2n}$, G can be decomposed in the direct sum of a linear subspace, $G_0 \subset \mathbb{R}^{2n}$, and a discrete lattice, L, generated by some vectors orthogonal to G_0, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \in G_0^{\perp}$:

$$G = G_0 \oplus L, \quad L = \left\{ \sum_{j=1}^{m} k_j\boldsymbol{u}_j \;\middle|\; k_1, \ldots, k_m \in \mathbb{Z} \right\}.$$

The linear subspace G_0 is the largest linear subspace contained in G.

The condition for G to be closed is without consequences in our case as an open set and its closure would stabilize the same space. If G is open, then we can replace it by its closure and appropriately complete $\mathcal{S}$. The limiting case when $G = G_0$ corresponds to the continuous case whereas $G = L$ corresponds to the discrete case. Denote ℓ as the dimension of G_0, one can choose some basis vectors, $G_0 = \langle \boldsymbol{g}_1, \ldots, \boldsymbol{g}_\ell \rangle$. Consider one of the generators of G_0, w.l.o.g. $\boldsymbol{g}_1$, for this generator, given a scalar factor $\lambda \in \mathbb{R}$, there is some angle function $\theta_1(\lambda)$ such that

$$\mathrm{e}^{i\theta_1(\lambda)}\hat{U}(\lambda\boldsymbol{g}_1) \in \mathcal{S}.$$

It is easy to check that θ_1 obeys Cauchy's functional equation and as such θ_1 is automatically $\mathbb{Q}$-linear:

$$\forall(\lambda,\mu) \in \mathbb{R}^2, \quad \theta_1(\lambda+\mu) = \theta_1(\lambda) + \theta_1(\mu) \;\Rightarrow\; \forall r \in \mathbb{Q},$$
$$\theta_1(r) = r\theta_1(1) \equiv r\theta_1.$$

Hence, for any code state, $|\overline{\Psi}\rangle \in \mathcal{Q}$, and rational $r \in \mathbb{Q}$, we can write down

$$\mathrm{e}^{ir\theta_1}\hat{U}(r\boldsymbol{g}_1)\,|\overline{\Psi}\rangle = \exp[ir(\theta_1 + \hat{O}(\boldsymbol{g}_1))]\,|\overline{\Psi}\rangle = |\overline{\Psi}\rangle\,. \tag{A5}$$

The previous equation means that code states are eigenstates of the operator $\hat{O}(\boldsymbol{g}_1)$ with eigenvalue $O(\boldsymbol{g}_1)$, which satisfies

$$\forall r \in Q, \quad O(\boldsymbol{g}_1) + \theta_1 = 0 \mod 2\pi/r,$$
$$\Leftrightarrow O(\boldsymbol{g}_1) = -\theta_1.$$

Usually, $\mathcal{S}$ will be chosen such that $\theta_1 = 0$, and $\hat{O}(\boldsymbol{g}_1)$ will be called a *nullifier* as it only takes eigenvalue 0 on the code space. Choosing some nontrivial θ_1 just corresponds to shifting the whole code space by $\hat{U}(\theta_1\boldsymbol{d}_1)$, with $\boldsymbol{d}_1$ describing the conjugated pair to $\boldsymbol{g}_1$, i.e., such that $[\hat{O}(\boldsymbol{g}_1), \hat{O}(\boldsymbol{d}_1)] = i$. Similarly, each generator $\boldsymbol{g}_j$ is a nullifier on the code space [up to some possible shift $\hat{U}(\theta_j\boldsymbol{d}_j)$].

At this point, if $G = G_0$ ($L = \emptyset$), then we have described a linear oscillator code. It is defined by the ℓ nullifiers, $\hat{O}(\boldsymbol{g}_j)$, which each remove a single continuous degree of freedom from the system, leaving $k \equiv n - \ell$ logical oscillator modes. The logical operators can be found by completing the $\boldsymbol{g}_j$ into a full symplectic basis. For the stabilized code space to be nontrivial we therefore require that $\ell < n$.

Consider now that L is nontrivial, then it will constrain the code space as described in Ref. [3, Sec.VI] on the remaining $k = n - \ell$ modes available. Take one of the generators of L, w.l.o.g. $\boldsymbol{u}_1$. The difference with elements of G_0 is that it occurs

in $\mathcal{S}$ only with integer multiples. Similarly as previously, given a code state, $|\overline{\Psi}\rangle \in \mathcal{Q}$, and an integer $k \in \mathbb{Z}$ there will be some angle, ϑ_1, such that

$$e^{ik\vartheta_1}\hat{U}(k\boldsymbol{u}_1)\,|\overline{\Psi}\rangle = \exp[ik(\vartheta_1 + \hat{O}(\boldsymbol{u}_1))]\,|\overline{\Psi}\rangle = |\overline{\Psi}\rangle\,. \quad \text{(A6)}$$

This means that on the code states, the operator $\hat{O}(\boldsymbol{u}_1)$ can take now several values, given by

$$\forall k \in \mathbb{Z}, \quad O(\boldsymbol{u}_1) + \vartheta_1 = 0 \mod 2\pi/k$$
$$\Leftrightarrow\ O(\boldsymbol{u}_1) = -\vartheta_1 \mod 2\pi\,.$$

The effect is to discretize this mode. As explained in Ref. [3], one then needs $m = 2k$ lattice generators to fully discretize the remaining k modes.

Summing up, the case where $G = G_0$ and $\ell < n$ ($m = 0$) corresponds to linear oscillator codes defined by ℓ nullifiers, encoding $k = n - \ell$ oscillators. However, the case $G = L$ and $m = 2n$ ($\ell = 0$) correspond to proper GKP codes described in Ref. [3]. Finally, the hybrid case where $G = G_0 \oplus L$, with $n = \ell + m/2 + k'$ correspond to a case where ℓ nullifiers leave $n - \ell$ modes in the code space, among which $n - \ell - k'$ are discretized into a qudit (depending on the characteristics of L) and k' remain as logical oscillators. As we show in Sec. VI, under Gaussian noise, the logical oscillators defined by the nullifiers have essentially the same noise model as the physical oscillator modes, so there is little interest in going beyond proper GKP codes. Note that linear oscillator codes can nevertheless correct erasure errors [1,2].

APPENDIX B: DETAILS ABOUT THE DECODERS FOR THE GKP QUBIT

1. Maximum likelihood decoder

Maximum likelihood decoding for a single GKP qubit requires the calculation of the partition functions $Z_c(\boldsymbol{q})$, $c \in \{0, 1\}$ given by Eq. (21). Here we do the Gaussian integration. Denote as B the symmetric matrix with the components

$$B_{ij} \equiv \frac{1}{2}\frac{\partial^2}{\partial\phi_i\partial\phi_j}R(\boldsymbol{\phi}, \boldsymbol{0}), \quad i, j \in [M-1], \quad \text{(B1)}$$

associated with the first $M - 1$ variables ϕ_t in the quadratic form

$$R(\boldsymbol{\phi}, \boldsymbol{q}) = \frac{1}{\sigma_{\mathrm{M}}^2}\sum_{t=1}^{M-1}(q_t - \phi_t)^2 + \frac{1}{\sigma^2}\sum_{t=1}^{M}(\phi_t - \phi_{t-1})^2\Big|_{\phi_M = q_M}$$

in the exponent in the integrand of Eq. (21). Collecting the remaining terms and completing the square we obtain an M-variable quadratic form $\boldsymbol{q}A\boldsymbol{q}^{\mathrm{T}}$ with the block matrix

$$A = \begin{pmatrix} \tilde{A} & \boldsymbol{c}^{\mathrm{T}} \\ \boldsymbol{c} & b \end{pmatrix}, \quad \text{(B2)}$$

expressed in terms of the $(M-1)\times(M-1)$ matrix $\tilde{A}$, row vector $\boldsymbol{c}$, and a scalar b:

$$\tilde{A} = \frac{1}{\sigma_{\mathrm{M}}^2}\mathbb{1} - \frac{1}{\sigma_{\mathrm{M}}^4}B^{-1},$$
$$\boldsymbol{c}_i = -\frac{1}{\sigma^2\sigma_{\mathrm{M}}^2}[B^{-1}]_{M-1,i}, \quad \text{(B3)}$$
$$b = \frac{1}{\sigma^2} - \frac{1}{\sigma_{\mathrm{M}}^4}[B^{-1}]_{M-1,M-1}.$$

2. Dynamic programming for the minimum-energy decoder

In this section we express the minimum-energy decoder for one GKP qubit as a dynamic programming problem. This allows us to check how well the forward-minimization technique performs. The goal is to minimize the energy from Eq. (26), which we write using the Villain approximation in reverse,

$$H[\phi_1, \ldots, \phi_M] = \sum_{t=1}^{M}\frac{(\phi_t - \phi_{t-1})^2}{2\sigma^2} - \sum_{t=1}^{M-1}\frac{\cos(q_t - \phi_t)}{\sigma_{\mathrm{M}}^2},$$
$$\phi_0 \equiv 0.$$

We define the partial energy, $H_k[\phi_1, \ldots, \phi_k]$ as the contribution from the first k terms from each sum,

$$H_k[\phi_1, \ldots, \phi_k] = \sum_{t=1}^{k}\frac{(\phi_t - \phi_{t-1})^2}{2\sigma^2} - \frac{\cos(q_t - \phi_t)}{\sigma_{\mathrm{M}}^2},$$

and a single-variable function of ϕ_k as

$$M_k(\phi_k) = \min_{\phi_1,\ldots,\phi_{k-1}}\{H_k[\phi_1, \ldots, \phi_k]\}.$$

Then M_k, $k < M$, can be defined recursively by

$$M_1(\phi_1) = \frac{\phi_1^2}{2\sigma^2} - \frac{\cos(q_1 - \phi_1)}{\sigma_{\mathrm{M}}^2},$$
$$M_k(\phi_k) = \min_{\phi}\left[M_{k-1}(\phi) + \frac{(\phi_k - \phi)^2}{2\sigma^2}\right] - \frac{\cos(q_k - \phi_k)}{\sigma_{\mathrm{M}}^2}.$$

If one discretizes the values of ϕ_k to a desired precision and restricts them to lie in a reasonable interval, then the minimization with M time steps amounts to computing M lists of values of discrete functions $M_k(\phi_k)$. Unlike the minimization technique based on solving Eq. (27), there is no accuracy loss at larger M, and much less danger of missing the desired minimum with the present dynamic programming method.

We have compared our forward minimization technique with dynamic programming with a discretization of 200 points per period in a four-periods window around the last measurement result. The statistics of success or failure of both decoders agree pretty closely while the forward minimization is much faster. In Fig. 13 we show two of the obtained realizations for illustration purposes.

APPENDIX C: ALTERNATIVE DERIVATION OF THE $U(1)$-SYMMETRIC MODEL

Here we derive Eqs. (41) and (42) directly from Eqs. (37) and (40). First, note that the integration over components of the field ϕ in Eq. (40) is done in an infinite interval, while the two last terms in the bulk Hamiltonian (37) are 2π- and 4π-periodic, respectively. In addition, the background fluxes $(\nabla \times \phi)_h(t)$ in the last term of Eq. (37) are explicitly symmetric with respect to the gauge transformation

$$\phi_{ij}(t) \to \phi_{ij}(t) + \alpha_j(t) - \alpha_i(t), \quad \text{(C1)}$$

where $\alpha_j(t) = \alpha_v$ is a real-valued scalar field associated with the vertices $v = (i, t)$ of the cubic lattice. While these are *not* the symmetries of the full Hamiltonian Eq. (37), we can

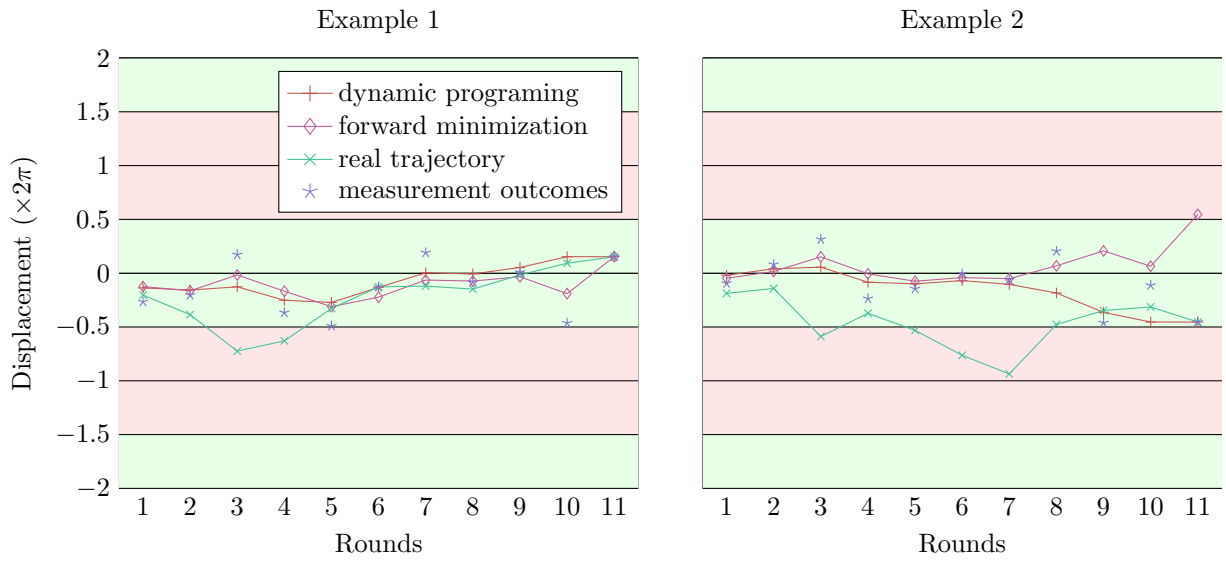

FIG. 13. Two examples of 11 rounds of error correction with data and measurement errors sampled from Gaussian distributions with $\sigma_0 = 0.4$. In the example on the left, forward-minimization and dynamic programming reach the same conclusion, whereas on the right forward-minimization reaches a wrong conclusion. One can compare this data with the sketched trajectories in Fig. 3.

now render it in a more familiar $U(1)$-symmetric form with a simple change of variables.

For a fixed error $\boldsymbol{e} \equiv \{\boldsymbol{\epsilon}, \boldsymbol{\delta}, \boldsymbol{\xi}\}$ which corresponds to the syndrome $\boldsymbol{q}$, let us denote the corresponding accumulated error, see Eq. (34), as $\phi^{(0)}$. We can make a change of integration variables in Eq. (42), for all bonds (ij) and layers $1 \leqslant t \leqslant M$, following

$$\begin{aligned} \phi_{ij}(t) &= \phi^{(0)}_{ij}(t) + 2A_{ij}(t) + 4\pi m_{ij}(t) \\ &\quad \pm 2\pi \sum_{t'=1}^{t} [s_j(t') - s_i(t')], \\ &\quad -\pi < A_{ij}(t) \leqslant \pi, \end{aligned} \tag{C2}$$

where the field $A_{ij}(t)$ is continuous, $m_{ij}(t) \in \mathbb{Z}$ is integer-valued, and an additional 2π-shift is proportional to the lattice gradient of the binary field $s_i(t) \in \mathbb{F}_2$ accumulated over time. Strictly speaking, the last term is unnecessary as it causes some double counting in the measure. However, the corresponding factor is a constant that is finite on a finite lattice, so it does not cause any trouble. However, these binary charges simplify the boundary conditions, since we can simply write $s_i(t) + \cdots + s_i(M) = n_i$, see Eq. (40), and thus trade the summation over the boundary field n_i for the binary field $s_i(t)$ in the bulk. To complete the derivation, also define $A_{ij}(t = 0) = 0$, and take $A_b = \pi s_i(t)$ for the vertical bond b at the square-lattice vertex i, between layers $t - 1$ and t for all $t \in [M]$. This gives for the vertical plaquette p at the bond (ij), between the same two layers,

$$\begin{aligned} &\phi_{ij}(t) - \phi_{ij}(t-1) \to \phi^{(0)}_{ij}(t) - \phi^{(0)}_{ij}(t-1) + 2(\nabla \times A)_p \\ &\quad \equiv \epsilon_p + 2(\nabla \times A)_p, \end{aligned}$$

where the appropriate sign in Eq. (C2) needs to be chosen to recover the part of the flux that is missing in the first term of Eq. (37), and we absorbed the summation over the integer-valued $m_{ij}(t)$ into the definition of the Villain potential $V_{\sigma/2}[\epsilon_p/2 + (\nabla \times A)_p]$, an even 2π-periodic function of the argument; see Eq. (23). In the remaining two terms we use Eqs. (35) and (36), to recover Eq. (41) exactly [63]. Comparing with Eqs. (38) and (39), it is also easy to check that for this particular error $\boldsymbol{e}$, one can simply take the vector potential $A_b = 0$ for the top-layer in-plane bonds, $b = (ij, M)$. Similarly, for the three sectors where the error in the top layer is shifted by a nontrivial toric codeword $\mathbf{0} \not\simeq \boldsymbol{c} \in \mathbb{F}_2^{2L^2}$, we can take $A_{ij}(M) = \pi c_{ij}$.

APPENDIX D: DETAILS ABOUT THE PHASE-DIAGRAM DERIVATION

The model in Eq. (43) is in a general class of compact $U(1)$ models whose partition functions can be written (in a conventional *Wilson*, or non-Villain form) as

$$\begin{aligned} Z(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P) &= \int d\boldsymbol{\theta}\, e^{-H}, \\ H &\equiv -\sum_b \kappa_b \cos\left(\varphi_b + \sum_i \theta_i P_{ib}\right), \end{aligned} \tag{D1}$$

where P is an $r \times n$ coupling matrix with integer components P_{ib} that determines the structure of the model, θ_i, $i \in [r]$ are the $U(1)$ variables, $\theta_i \in [0, 2\pi)$, $\kappa_b \geqslant 0$, $b \in [r]$, are the coupling constants, and the additional phases φ_b can represent quenched disorder and/or a uniform background field. The only requirement on the integration measure for any component θ_j in Eq. (D1) is 2π-periodicity. For example, one can have the usual integration over the period, or a summation over discrete phases $\theta_j = 2\pi m_j/q_j$, $m_j \in \mathbb{Z}_q$, with some integer $q_j \geqslant 2$ that may differ for different "spins" j. According to the Fourier theorem, the most general correlation function of the variables θ_j can be written as the average

$$\mathcal{C}_m \equiv C_m(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P) = \langle e^{im\theta} \rangle, \tag{D2}$$

where the vector $\boldsymbol{m}$ has integer components. The physics of the model Eq. (D1) is characterized by the dependence of the free energy $F = -\ln Z$ and the correlation functions C_m on the parameters, the coupling constants κ_b and the phases φ_b.

This dependence is restricted by several constraints. Two of them, the first and the second generalized Griffiths-Kelly-Sherman (GKS) inequalities, concern the correlation functions in the absence of background phases, $\boldsymbol{\varphi} = \mathbf{0}$ (in this case the averages Eq. (D2) are real-valued),

$$\langle f \rangle \geqslant 0, \tag{D3}$$

$$\langle fg \rangle - \langle f \rangle \langle g \rangle \geqslant 0, \tag{D4}$$

where f and g can be any nonnegative combination of products of $\cos(\boldsymbol{m} \cdot \boldsymbol{\theta})$, with various integer vectors $\boldsymbol{m} \in \mathbb{Z}^n$. In the case of the model Eq. (D1) in Wilson form these are called Ginibre inequalities [64]. It is also easy to check that the average $C_m(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P)$ can only be nonzero if $\boldsymbol{m}$ is a linear combination of the rows of the matrix P. Furthermore, the left-hand side of the second inequality Eq. (D4), with $g = \cos(\boldsymbol{m}' \cdot \boldsymbol{\theta})$, equals the derivative of $\langle f \rangle$ with respect to the coupling constant corresponding to the term $\cos(\boldsymbol{m}' \cdot \boldsymbol{\theta})$. This implies a monotonic nondecreasing dependence of any correlation function $\langle f \rangle$, including $C_m(\boldsymbol{\kappa}, \mathbf{0}; P)$, on any coupling constant $\kappa_b \geqslant 0$. The Villain version of the same model has similar properties, since the potential $V_\sigma(\varphi)$ can be approximated to an arbitrary precision with a chain of phases with pairwise Wilson couplings [65]. In particular, generalized GKS inequalities apply for the averages $C_m(\boldsymbol{\kappa}, \mathbf{0}; P)$ in the Villain form of the model, which are also monotonically nondecreasing as a function of any coupling.

The second type of constraints concerns the free energy $F(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P) \equiv -\ln Z(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P)$, or, more precisely, the free-energy cost associated with the background phases φ,

$$\Delta_\varphi F(\boldsymbol{\kappa}; P) \equiv F(\boldsymbol{\kappa}, \boldsymbol{\varphi}; P) - F(\boldsymbol{\kappa}, \mathbf{0}; P). \tag{D5}$$

Again, the free-energy cost is nonnegative, $\Delta_\varphi F(\boldsymbol{\kappa}; P) \geqslant 0$, and it is a nondecreasing function of the coupling constants κ_b. This can be obtained from the generalized GKS inequalities for the dual form of the model, where the free-energy cost $\Delta_\varphi F$ is mapped into the logarithm of a correlation function parameterized by the phases $\boldsymbol{\varphi}$, up to a sign, with the coupling constants inverted.

An important (albeit nearly self-evident) consequence of the monotonicity of the free-energy cost Eq. (D5) is a generalization of the inequalities Eqs. (17) and (18) from Ref. [66] to all GKP codes. These inequalities imply an upper bound for the success probability of ML decoding, see Eq. (8), in terms of the partition functions in the absence of quenched disorder. Namely, in the case of CSS-like codes, the bound reads

$$\mathbb{P}_{\text{succ}}^{(\text{ML})} \leqslant \frac{Z_{\mathbf{0}}(\mathbf{0})}{\sum_{\boldsymbol{c}} Z_{\boldsymbol{c}}(\mathbf{0})}, \tag{D6}$$

where the summation is over all 2^k inequivalent binary codewords that correspond to the entire group of X-type CSS logical operators. When applied to the case of repeated syndrome measurement for the toric-GKP code, the numerator in Eq. (D6) is the partition function (42) with zero argument, while the denominator is the sum of the same partition function $Z_{\mathbf{0}}(\mathbf{0})$ with those for the three remaining nontrivial sectors, $Z_{\boldsymbol{c}}(\mathbf{0})$. For the existence of an ML-decodable region it is necessary that these contributions vanish in the large-system limit, namely, when both the distance of the toric code L and the number of layers M diverge.

Typically examined are the Wilson loops, they are correlation functions in the form of Eq. (D2),

$$\mathcal{W}_\Omega \equiv \langle e^{i \sum_{p \in \Omega} (\nabla \times A)_p} \rangle \equiv \langle e^{i \sum_{b \in \partial\Omega} A_b} \rangle, \tag{D7}$$

where Ω is some oriented surface, a set of plaquettes, with the boundary $\partial\Omega$, and the gauge-invariant two-point correlation function

$$\mathcal{C}_\Pi \equiv \langle e^{i(\theta_u - \theta_v) - iq \sum_{b \in \Pi} A_b} \rangle, \tag{D8}$$

where $\Pi \equiv \Pi(u, v)$ is a directed path (sequence of bonds) on the cubic lattice connecting vertices u and v. Quite generally, in a high-temperature phase, when coupling constants κ and λ in Eq. (44) are sufficiently small, the correlation function Eqs. (D7) and (D8) are characterized by the *area law*: $\mathcal{W}_\Omega$ has an upper bound that decays exponentially with the area of Ω, the minimum number of plaquettes that are needed to form the boundary $\partial\Omega$, while $\mathcal{C}_\Pi$ has an upper bound that decays exponentially with the distance between the points u and v. A low-temperature, or strong-coupling asymptotic form is qualitatively different and is characterized by the *perimeter law*, where $\mathcal{W}_\Omega$ scales exponentially with the length $|\partial\Omega|$ of the perimeter, while $\mathcal{C}_\Pi$ becomes a constant or falls as a power law of the distance; the latter is the case below the Berezinskii-Kosterlitz-Thouless (BKT) transition in two dimensions.

APPENDIX E: PROOF OF THE NO-GO THEOREM IN SECTION VI

For convenience, we rename the $\hat{p}_j$ and $\hat{q}_j$ operators of the oscillators as $\hat{r}_k$, i.e.,

$$\forall k \in [2n], \quad \hat{r}_k = \begin{cases} \hat{p}_k & \text{when } j \leqslant n, \\ \hat{q}_{k-n} & \text{when } j > n. \end{cases}$$

Let $\boldsymbol{g}_j$ be the real vector corresponding to the jth nullifier $\hat{O}(\boldsymbol{g}_j) = \boldsymbol{g}_j \cdot \hat{\boldsymbol{r}}$ of the code; see Appendix A. One can extend the set of nullifiers to a full canonical linear transformation given by a real $2n \times 2n$ matrix A defining new variables $\hat{R}_k$, as follows:

$$\hat{\boldsymbol{R}} = A\hat{\boldsymbol{r}}^{\mathrm{T}}, \quad \text{i.e., } \hat{R}_k = \sum_{j=1}^{2n} A_{kj} \hat{r}_j.$$

To preserve the commutation relation the condition on the matrix A is that it preserves the symplectic form S. This can be expressed in two ways:

$$ASA^{\mathrm{T}} = S \quad \text{or} \quad A^{\mathrm{T}}SA = S \quad \text{with}$$
$$S = \begin{pmatrix} 0 & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & 0 \end{pmatrix}. \tag{E1}$$

The matrix A can be decomposed in blocks

$$A = \begin{array}{c} \\ n-k\updownarrow \\ k\updownarrow \\ n-k\updownarrow \\ k\updownarrow \end{array} \begin{array}{c} \overleftarrow{\ 2n\ }\rightarrow \\ \begin{pmatrix} G \\ P \\ D \\ Q \end{pmatrix} \end{array} = \begin{array}{cc} \overleftarrow{\ n\ }\rightarrow & \overleftarrow{\ n\ }\rightarrow \\ \begin{pmatrix} G_p & G_q \\ P_p & P_q \\ D_p & D_q \\ Q_p & Q_q \end{pmatrix} \end{array},$$

where the rows of G are the nullifiers. The rows of D represent the corresponding conjugated variables which will be called *pure errors* (they are sometimes referred to as destabilizers). These pure errors give a convenient basis for expressing an error which is compatible with a given syndrome which will be used below. The rows of P (respectively, Q) represent the logical $\hat{p}$ (respectively, $\hat{q}$) operators of the code as linear combinations of the original $\hat{p}_i$ and $\hat{q}_j$. The subscript p (respectively, q) indicates the $\hat{p}$ part (respectively, $\hat{q}$ part) of the operators. Inside the code space the operators $\hat{O}(\boldsymbol{g}_j)$ only take the value 0. Let us assume that a displacement happens along a pure error direction $\boldsymbol{d}_j$, say $\hat{U}(-\lambda \boldsymbol{d}_j)$. Then, measuring $\hat{O}(\boldsymbol{g}_j)$ [equivalently, $\hat{U}(\eta \boldsymbol{g}_j)$ for all $\eta \in \mathbb{R}$] would give outcome $\lambda \in \mathbb{R}$ called the syndrome, since

$$\begin{aligned} \hat{U}(\eta \boldsymbol{g}_j)[\hat{U}(-\lambda \boldsymbol{d}_j)|\Psi\rangle] &= \mathrm{e}^{i\lambda\eta}\hat{U}(-\lambda \boldsymbol{d}_j)\hat{U}(\eta \boldsymbol{g}_j)|\Psi\rangle \\ &= \mathrm{e}^{i\lambda\eta}[\hat{U}(-\lambda \boldsymbol{d}_j)|\Psi\rangle]. \end{aligned}$$

Note that the logical operators $\hat{O}(\boldsymbol{p}_k)$ or $\hat{O}(\boldsymbol{q}_\ell)$, or nullifiers, act on the code space without affecting the measurement of $\hat{O}(\boldsymbol{g}_j)$ since they commute with it.

We can express the constraints on the matrix A in Eq. (E1) in terms of the matrices G, D, P, and Q. The blocks which should be equal to zero are shown in blue and the blocks proportional to the identity are shown in green. The first condition gives

$$\begin{aligned} ASA^{\mathrm{T}} &= \begin{pmatrix} GSG^{\mathrm{T}} & GSP^{\mathrm{T}} & GSD^{\mathrm{T}} & GSQ^{\mathrm{T}} \\ PSG^{\mathrm{T}} & PSP^{\mathrm{T}} & PSD^{\mathrm{T}} & PSQ^{\mathrm{T}} \\ DSG^{\mathrm{T}} & DSP^{\mathrm{T}} & DSD^{\mathrm{T}} & DSQ^{\mathrm{T}} \\ QSG^{\mathrm{T}} & QSP^{\mathrm{T}} & QSD^{\mathrm{T}} & QSQ^{\mathrm{T}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \mathbb{1} & 0 \\ 0 & 0 & 0 & \mathbb{1} \\ -\mathbb{1} & 0 & 0 & 0 \\ 0 & -\mathbb{1} & 0 & 0 \end{pmatrix} = S, \end{aligned} \tag{E2}$$

while the second implies

$$\mathbb{1} = A^{\mathrm{T}}SAS^{T} = \begin{pmatrix} D_q^{\mathrm{T}}G_p + Q_q^{\mathrm{T}}P_p - G_q^{\mathrm{T}}D_p - P_q^{\mathrm{T}}Q_p & D_q^{\mathrm{T}}G_q + Q_q^{\mathrm{T}}P_q - G_q^{\mathrm{T}}D_q - P_q^{\mathrm{T}}Q_q \\ -D_p^{\mathrm{T}}G_p - Q_p^{\mathrm{T}}P_p + G_p^{\mathrm{T}}D_p + P_p^{\mathrm{T}}Q_p & -D_p^{\mathrm{T}}G_q - Q_p^{\mathrm{T}}P_q + G_p^{\mathrm{T}}D_q + P_p^{\mathrm{T}}Q_q. \end{pmatrix} \tag{E3}$$

1. Logical error model under displacement errors

We consider the Gaussian displacement noise model on every oscillator as in Eq. (5). Given a realization of the displacement error as a vector of real amplitudes, $\boldsymbol{e}' \in \mathbb{R}^{2n}$, one can compute the vector of syndromes $\boldsymbol{q} \in \mathbb{R}^{n-k}$, and given $\boldsymbol{q}$, a candidate error with the same syndrome, $\boldsymbol{e} \in \mathbb{R}^{2n}$,

$$\boldsymbol{q} = \boldsymbol{e}' S G^{\mathrm{T}}, \quad \boldsymbol{e} = -\boldsymbol{q} D.$$

In addition, when any two errors $\boldsymbol{e}$ and $\boldsymbol{e}'$ have the same syndrome $\boldsymbol{q}$, then they can only differ by a stabilizer and a logical operator and one has

$$\begin{aligned} &\exists \boldsymbol{u}_c, \boldsymbol{v}_c \in \mathbb{R}^k, \quad \exists \boldsymbol{a} \in \mathbb{R}^{n-k}, \\ &\boldsymbol{e}' = \boldsymbol{e} + \boldsymbol{u}_c P + \boldsymbol{v}_c Q + \boldsymbol{a} G = \boldsymbol{e} + \boldsymbol{c} C + \boldsymbol{a} G, \end{aligned}$$

where we have used the notation

$$C := \begin{pmatrix} P \\ Q \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} \boldsymbol{u}_c & \boldsymbol{v}_c \end{pmatrix}.$$

We can compute the associated partition function, $Z_{\boldsymbol{c}}(\boldsymbol{e})$ from Eq. (7), for the error $\boldsymbol{e}$ and the sector equivalent to $\boldsymbol{c}$,

$$\begin{aligned} Z_{\boldsymbol{c}}(\boldsymbol{e}) = &\left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n \int \prod_{j=0}^{n-k-1} da_j \\ &\times \exp\left[-\frac{(\boldsymbol{c}C + \boldsymbol{a}G + \boldsymbol{e})(\boldsymbol{c}C + \boldsymbol{a}G + \boldsymbol{e})^{\mathrm{T}}}{2\sigma_0^2}\right]. \end{aligned} \tag{E4}$$

This integral can be evaluated since it is Gaussian, resulting, after some manipulations, in

$$Z_{\boldsymbol{c}}(\boldsymbol{e}) = \mathcal{C}(\boldsymbol{e}) \exp\left[-\frac{1}{2\sigma_0^2}(\boldsymbol{c} - \boldsymbol{\mu}(\boldsymbol{e}))\Sigma^{-1}(\boldsymbol{c} - \boldsymbol{\mu}(\boldsymbol{e}))^{\mathrm{T}}\right], \tag{E5}$$

where $\mathcal{C}(\boldsymbol{e})$ only depends on $\boldsymbol{e}$. The covariance matrix Σ and the off-set vector, $\boldsymbol{\mu}(\boldsymbol{e})$, are defined as

$$\Sigma^{-1} = C'C'^{\mathrm{T}}, \quad \text{and} \quad \boldsymbol{\mu}(\boldsymbol{e}) = -\Sigma C' \boldsymbol{e}^{\mathrm{T}}, \tag{E6}$$

where

$$C' = C\Pi_{G^\perp}, \quad \text{and} \quad \Pi_{G^\perp} = \mathbb{1} - G^T(GG^T)^{-1}G. \tag{E7}$$

Remark that $\Pi_{G^\perp}$ is the projector onto $\ker(G)$ along $\mathrm{im}(G^{\mathrm{T}})$ [67]. Indeed it is easy to check that $\Pi_{G^\perp}^2 = \Pi_{G^\perp}$, $\ker(G) \subset \mathrm{im}(\Pi_{G^\perp})$, $\mathrm{im}(G^{\mathrm{T}}) \subset \ker(\Pi_{G^\perp})$, and the dimensions coincide.

One can see that Eq. (E5) describes a multivariate Gaussian distribution over the logical variable $\boldsymbol{c}$, hence its maximum is readily given by the mean value, $\boldsymbol{\mu}(\boldsymbol{e})$; see Eq. (E6). This means that given the error $\boldsymbol{e}$, one can directly express its most likely error class: It is $[\boldsymbol{e} + \boldsymbol{\mu}(\boldsymbol{e})]$. Using this, one can directly compute the probability density of a remaining logical error after ML decoding as given by Eq. (10):

$$\mathbb{P}^{(\mathrm{ML})}(\boldsymbol{c}) = \int d\boldsymbol{e}\, \mathbb{P}(\boldsymbol{e}) \frac{Z_{\boldsymbol{\mu}(\boldsymbol{e})+\boldsymbol{c}}(\boldsymbol{e})}{\int d\boldsymbol{b}\, Z_{\boldsymbol{b}}(\boldsymbol{e})} = \frac{1}{N} \exp\left(-\frac{1}{2\sigma_0^2}\boldsymbol{c}\Sigma^{-1}\boldsymbol{c}^{\mathrm{T}}\right), \tag{E8}$$

where N is a normalization constant. All possible dependence on the choice and size of the code is contained in the

covariance matrix Σ, leading to some rescaled displacement noise model. Recall that the logical variable vector $\boldsymbol{c} \in \mathbb{R}^{2k}$ represents all k pairs of conjugated logical operators. We will prove in Appendix E 2 that for any linear oscillator code, the eigenvalues of Σ^{-1} can be paired between corresponding conjugated logical operators, denoted $(\lambda_j^p, \lambda_j^q)_{j\in[n]}$, and are such that

$$\forall j \in [n], \quad \lambda_j^p \lambda_j^q = 1. \tag{E9}$$

This means that the noise remaining after ML decoding on the logical variables is identical to the original physical noise except for a possible squeezing between each logical operator and its conjugated pair.

2. Eigenvalues of the covariance matrix

First we remark that C', appearing in $\Sigma^{-1} = C'C'^{\mathrm{T}}$ in Eq. (E6), correspond to a valid basis for the logical operators—we call this the spread-out logical basis. This basis is obtained by adding linear combination of stabilizer generators to C. Such addition can be summarized, using Eq. (E7), by the matrix equation,

$$C' = C - CG^{\mathrm{T}}(GG^{\mathrm{T}})^{-1}G = C + \Lambda G,$$

where Λ is a $2k \times (n-k)$ matrix defining the linear combination of stabilizer generators added to the logical operators. One can verify that one can replace C by C' and still satisfy the constraints of Eq. (E1), if one appropriately redefines the pure errors to be

$$D' = D + \Lambda^{\mathrm{T}} S_{2k}^{\mathrm{T}} C,$$

where S_{2k} is the symplectic form of size $2k \times 2k$. One can also check that this choice of basis is the only choice which enforces the following constraint:

$$C'G^{\mathrm{T}} = (C + \Lambda G)G^{\mathrm{T}} = 0. \tag{E10}$$

Indeed, solving for Λ, using the fact that G is full rank and therefore GG^{T} is invertible, gives

$$C' = C - CG^{T}(GG^{\mathrm{T}})^{-1}G. \tag{E11}$$

Now we use this spread-out logical basis to prove Eq. (E9). We thus consider that we already have chosen the spread-out logical basis, so $C' = C$, and want to get information about the eigenvalues of $\Sigma^{-1} = CC^{\mathrm{T}}$.

We use the diagonal block of Eq. (E3) as well as the block only about logical operators in Eq. (E2):

$$-D_p^{\mathrm{T}}G_q - Q_p^{\mathrm{T}}P_q + G_p^{\mathrm{T}}D_q + P_p^{\mathrm{T}}Q_q = \mathbb{1}, \tag{E12}$$

$$D_q^{\mathrm{T}}G_p + Q_q^{\mathrm{T}}P_p - G_q^{\mathrm{T}}D_p - P_q^{\mathrm{T}}Q_p = \mathbb{1}, \tag{E13}$$

$$PSQ^{\mathrm{T}} = P_pQ_q^{\mathrm{T}} - P_qQ_p^{\mathrm{T}} = \mathbb{1}. \tag{E14}$$

We multiply Eq. (E12) on the left by P_p and on the right by Q_q^T. We also multiply Eq. (E13) on the left by P_q and on the right by Q_p^T. Then we take the difference, i.e., we consider

$$P_p(E12)Q_q^{\mathrm{T}} - P_q(E13)Q_p^{\mathrm{T}}.$$

By Eq. (E14), the right-hand side is still the identity, so we have

$$\begin{aligned}\mathbb{1} = {} & P_pG_p^{\mathrm{T}}D_qQ_q^{\mathrm{T}} + P_qG_q^{\mathrm{T}}D_pQ_p^{\mathrm{T}} \\ & - P_qD_q^{\mathrm{T}}G_pQ_p^{\mathrm{T}} - P_pD_p^{\mathrm{T}}G_qQ_q^{\mathrm{T}} \\ & + P_pP_p^{\mathrm{T}}Q_qQ_q^{\mathrm{T}} + P_qP_q^{\mathrm{T}}Q_pQ_p^{\mathrm{T}} \\ & - P_pQ_p^{\mathrm{T}}P_qQ_q^{\mathrm{T}} - P_qQ_q^{\mathrm{T}}P_pQ_p^{\mathrm{T}}.\end{aligned}$$

Since the logical operators are in their spread-out basis we have the corresponding equations

$$P_pG_p^{\mathrm{T}} + P_qG_q^{\mathrm{T}} = 0, \quad G_pQ_p^{\mathrm{T}} + G_qQ_q^{\mathrm{T}} = 0,$$

and can write

$$\begin{aligned}\mathbb{1} = {} & P_q\big(D_q^{\mathrm{T}}G_q - G_q^{\mathrm{T}}D_q\big)Q_q^{\mathrm{T}} \\ & + P_p\big(D_p^{\mathrm{T}}G_p - G_p^{\mathrm{T}}D_p\big)Q_p^{\mathrm{T}} \\ & + P_p\big(P_p^{\mathrm{T}}Q_q - Q_p^{\mathrm{T}}P_q\big)Q_q^{\mathrm{T}} \\ & + P_q\big(P_q^{\mathrm{T}}Q_p - Q_q^{\mathrm{T}}P_p\big)Q_p^{\mathrm{T}}.\end{aligned}$$

One can now recognize that the off-diagonal terms in Eq. (E3) can be used to make a equation only about logical variables, using

$$\begin{aligned}D_q^{\mathrm{T}}G_q - G_q^{\mathrm{T}}D_q &= P_q^{\mathrm{T}}Q_q - Q_q^{\mathrm{T}}P_q, \\ D_p^{\mathrm{T}}G_p - G_p^{\mathrm{T}}D_p &= P_p^{\mathrm{T}}Q_p - Q_p^{\mathrm{T}}P_p.\end{aligned}$$

Hence, we have the following matrix identity:

$$\mathbb{1} = PP^{\mathrm{T}}QQ^{\mathrm{T}} - PQ^{\mathrm{T}}PQ^{\mathrm{T}}. \tag{E15}$$

In Appendix E 3, we show that there always exists a logical basis which is both spread out as well as orthogonal. Hence, for this new basis, Eq. (E10) is satisfied, as well as

$$\tilde{P}\tilde{P}^{\mathrm{T}} = \mathrm{Diag}\big(\lambda_j^p\big), \quad \tilde{Q}\tilde{Q}^{\mathrm{T}} = \mathrm{Diag}\big(\lambda_j^q\big), \quad \text{and} \quad \tilde{P}\tilde{Q}^{\mathrm{T}} = 0. \tag{E16}$$

Therefore, deriving Eq. (E15) in this basis yields

$$\mathbb{1} = \tilde{P}\tilde{P}^{\mathrm{T}}\tilde{Q}\tilde{Q}^{\mathrm{T}} = \mathrm{Diag}\big(\lambda_j^p\lambda_j^q\big), \tag{E17}$$

and therefore

$$\forall j \in [k], \quad \lambda_j^p\lambda_j^q = 1. \tag{E18}$$

Last, recall that starting with this choice of basis, one has

$$\Sigma^{-1} = \tilde{C}\tilde{C}^{\mathrm{T}} = \begin{pmatrix} \tilde{P}\tilde{P}^{\mathrm{T}} & \tilde{P}\tilde{Q}^{\mathrm{T}} \\ \tilde{Q}\tilde{P}^{\mathrm{T}} & \tilde{Q}\tilde{Q}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \mathrm{Diag}\big(\lambda_j^p\big) & 0 \\ 0 & \mathrm{Diag}\big(\lambda_j^q\big) \end{pmatrix}. \tag{E19}$$

■

3. Existence of a spread-out and orthogonal logical operator basis

We start from a given stabilizer matrix G and denote the rows by $\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-k}\}$. We showed above that one can always find a logical basis, P and Q with rows $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_k\}$ and $\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_k\}$, which is orthogonal to the stabilizer generators, i.e.,

$$\begin{aligned}\mathcal{L}_k &= \mathrm{Span}\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_k, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_k\} \perp \mathcal{G} \\ &= \mathrm{Span}\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-k}\}.\end{aligned}$$

We want to construct a new symplectic and orthogonal basis, $\{\tilde{\boldsymbol{p}}_1, \ldots, \tilde{\boldsymbol{p}}_k, \tilde{\boldsymbol{q}}_1, \ldots, \tilde{\boldsymbol{q}}_k\}$, for the $2k$-dimensional space $\mathcal{L}_k$. One can first find an orthogonal basis for $\mathcal{L}_k$ by the Gram-Schmidt process for the regular inner product, denote it $\{\boldsymbol{c}_1^k, \ldots, \boldsymbol{c}_{2k}^k\}$. We will now construct a new basis pair by pair, decreasing at each step the dimension of the space $\mathcal{L}_k$ by 2. We can choose

$$\tilde{\boldsymbol{p}}_1 = \boldsymbol{c}_1^k.$$

There has to exist a conjugated pair, $\tilde{\boldsymbol{q}}_1$, in the space spanned by $\{\boldsymbol{c}_2^k, \ldots, \boldsymbol{c}_{2k}^k\}$. Indeed, all stabilizer generators, pure errors as well as $\boldsymbol{c}_1^k$ have a trivial symplectic product with $\tilde{\boldsymbol{p}}_1$, hence any conjugated $\tilde{\boldsymbol{q}}_1 \in \mathrm{Span}\{\boldsymbol{c}_2^k, \ldots, \boldsymbol{c}_{2k}^k\}$. So the following equation,

$$\boldsymbol{\lambda} O_k S \tilde{\boldsymbol{p}}_1^{\mathrm{T}} = -1,$$

where O_k are the basis vectors $\{\boldsymbol{c}_2^k, \ldots, \boldsymbol{c}_{2k}^k\}$ stacked in rows, has a solution for $\boldsymbol{\lambda}$. Then we can choose

$$\tilde{\boldsymbol{q}}_1 = \boldsymbol{\lambda} O_k.$$

Note that the created pair is indeed conjugated for the symplectic inner product as well has orthogonal. Moreover, they are composed only of linear combinations of the original $\boldsymbol{p}_i$s and $\boldsymbol{q}_j$s, so they do have trivial symplectic product with stabilizer generators and pure errors and they are orthogonal to the stabilizer generators. Now we define the $2(k-1)$-dimensional subspace $\mathcal{L}_{k-1}$, as

$$\begin{aligned}\mathcal{L}_{k-1} &= \mathrm{Span}\{\boldsymbol{c}_2^k(\mathbb{1} - \Pi_{\tilde{\boldsymbol{q}}_1}), \ldots, \boldsymbol{c}_{2k}^k(\mathbb{1} - \Pi_{\tilde{\boldsymbol{q}}_1})\},\\ &= \mathrm{Span}\{\boldsymbol{c}_1^{k-1}, \ldots, \boldsymbol{c}_{2(k-1)}^{k-1}\},\end{aligned}$$

where $\Pi_{\tilde{\boldsymbol{q}}_1}$ is the orthogonal projector onto $\tilde{\boldsymbol{q}}_1$, and the $\boldsymbol{c}_j^{k-1}$ form a new orthogonal basis for this space. We can then repeat the procedure until we reach $\mathcal{L}_0 = \{0\}$. To summarize we have constructed a new symplectic and orthogonal basis $\tilde{C} = \binom{\tilde{P}}{\tilde{Q}}$ obeying Eq. (E16).

APPENDIX F: CONTINUOUS-VARIABLE TORIC-CODE

In this section we examine the continuous-variable toric code [35,68] as an example of a continuous-variable topological code. Since the toric code is a homological code, it is easy to convert it from a $\mathbb{Z}_2$-code to a $\mathbb{R}$-code using orientation to add appropriate minus signs to the stabilizer checks.

The stabilizer checks are shown on the left in Fig. 14. Since the toric code is a CSS code we will always have the orthogonality condition $PQ^{\mathrm{T}} = 0$. The vectors $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the two rows of P while the vectors $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are the two rows of Q. On the left, one sees the usual string-like $\boldsymbol{p}_1$ and $\boldsymbol{q}_1$. The spread-out basis for the code can be computed and is shown on the right of Fig. 14. The name *spread-out basis* comes from the fact that these logical operators have support over the full lattice: The continuous-variable stabilizer checks have been used to spread out or distribute their support over the lattice.

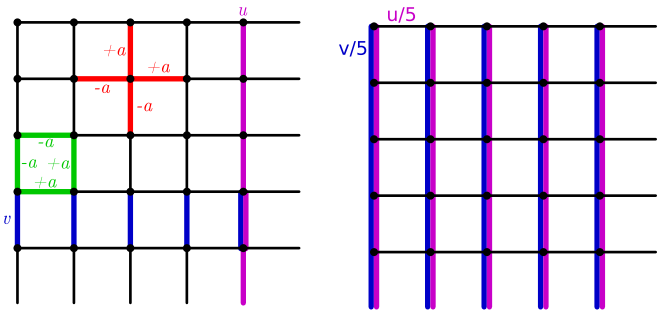

FIG. 14. In both figures periodic boundary conditions are assumed. (Left) Example of stabilizer checks and logical operators for the distance-5 continuous-variable toric-code. The stabilizer checks are also shown in Fig. 7. The support of the logical $\boldsymbol{p}_1$ (blue/dark gray) and $\boldsymbol{q}_1$ (purple/light gray) of one of the encoded oscillators is depicted. Shifts of strength v on the support of $\boldsymbol{p}_1$ is a logical $\boldsymbol{p}$-shift of strength v of the first oscillator. (Right) The support of the spread-out version of the logical operators $\boldsymbol{p}_1$ and $\boldsymbol{q}_1$. Both operators have the same support (blue/dark gray and purple/light gray) and one can verify that these logical operators are orthogonal to the stabilizer checks as well as commuting with them. A logical $\boldsymbol{p}$-shift of strength v on the first oscillator is now realized by applying $v/5$ on the support of the spread-out $\boldsymbol{p}_1$. In general, if the torus had dimension $L_x \times L_y$, the spread-out $\boldsymbol{p}_1$ would have a shift rescaling of $\frac{1}{L_x}$ over the whole lattice while $\boldsymbol{q}_1$ would a shift rescaling of $\frac{1}{L_y}$. At the same time $\boldsymbol{p}_2$ (respectively, $\boldsymbol{q}_2$) would have rescaling $\frac{1}{L_y}$ (respectively, $\frac{1}{L_x}$).

Computing CC^T, Eq. (E19), with C being the spread-out logical operators in the general case of different dimensions L_x and L_y gives directly a diagonal matrix:

$$CC^T = \mathrm{diag}\big(\lambda_1^p, \lambda_2^p, \lambda_1^q, \lambda_2^q\big) = \mathrm{diag}\big(\boldsymbol{p}_1^2, \boldsymbol{p}_2^2, \boldsymbol{q}_1^2, \boldsymbol{q}_2^2\big), \quad \text{(F1)}$$

with

$$\begin{aligned}\lambda_1^p = \boldsymbol{p}_1^2 &= \sum_{e_x} \frac{1}{L_y^2} = \frac{L_x L_y}{L_y^2} = \frac{L_x}{L_y},\\ \lambda_2^p = \boldsymbol{p}_2^2 &= \sum_{e_y} \frac{1}{L_x^2} = \frac{L_x L_y}{L_x^2} = \frac{L_y}{L_x},\\ \lambda_1^q = \boldsymbol{q}_1^2 &= \sum_{e_x} \frac{1}{L_x^2} = \frac{L_x L_y}{L_x^2} = \frac{L_y}{L_x},\\ \lambda_2^q = \boldsymbol{q}_2^2 &= \sum_{e_y} \frac{1}{L_y^2} = \frac{L_x L_y}{L_y^2} = \frac{L_x}{L_y}.\end{aligned}$$

One sees that it is possible to choose the squeezing amount by choosing the ratio of the dimensions L_x/L_y. In particular, with Eq. (E8), one sees that the first encoded oscillator experiences a Gaussian displacement noise model with variance $\sigma_p^2 = \frac{L_y}{L_x}\sigma_0^2$ for $\hat{p}$ and variance $\sigma_q^2 = \frac{L_x}{L_y}\sigma_0^2$ for $\hat{q}$. For the second oscillator the quadrature squeezing goes in the other direction.

[1] S. L. Braunstein, Error Correction for Continuous Quantum Variables, Phys. Rev. Lett. **80**, 4084 (1998).

[2] S. Lloyd and J.-J. E. Slotine, Analog Quantum Error Correction, Phys. Rev. Lett. **80**, 4088 (1998).

[3] D. Gottesman, A. Yu. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[4] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New Class of Quantum

Error-Correcting Codes for a Bosonic Mode, Phys. Rev. X **6**, 031006 (2016).
[5] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Performance and structure of single-mode bosonic codes, Phys. Rev. A **97**, 032346 (2018).
[6] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).
[7] B. M. Terhal and D. Weigand, Encoding a qubit into a cavity mode in circuit QED using phase estimation, Phys. Rev. A **93**, 012315 (2016).
[8] Y. Y. Gao, B. J. Lester, Y. Zhang, C. Wang, S. Rosenblum, L. Frunzio, L. Jiang, S. M. Girvin, and R. J. Schoelkopf, Programmable Interference Between Two Microwave Quantum Memories, Phys. Rev. X **8**, 021073 (2018).
[9] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, Sequential Modular Position and Momentum Measurements of a Trapped Ion Mechanical Oscillator, Phys. Rev. X **8**, 021001 (2018).
[10] K. R. Motes, B. Q. Baragiola, A. Gilchrist, and N. C. Menicucci, Encoding qubits into oscillators with atomic ensembles and squeezed light, Phys. Rev. A **95**, 053819 (2017).
[11] N. C. Menicucci, Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States, Phys. Rev. Lett. **112**, 120504 (2014).
[12] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Extending the lifetime of a quantum bit with error correction in superconducting circuits, Nature **536**, 441 (2016).
[13] K. Fukui, A. Tomita, and A. Okamoto, Analog Quantum Error Correction with Encoding a Qubit into an Oscillator, Phys. Rev. Lett. **119**, 180507 (2017).
[14] Y. Wang, Quantum error correction with the GKP code and concatenation with stabilizer codes, Master's thesis, RWTH Aachen University, 2017.
[15] L. P. Pryadko, H. Asasi, M. Mulligan, and A. A. Kovalev, Maximum likelihood decoding threshold for oscillator quantum error correcting codes, APS Meeting Abstracts (2017) p. V46.007.
[16] K. Fukui, A. Tomita, A. Okamoto, and K. Fujii, High-Threshold Fault-Tolerant Quantum Computation with Analog Quantum Error Correction, Phys. Rev. X **8**, 021054 (2018).
[17] K. Duivenvoorden, B. M. Terhal, and D. Weigand, Single-mode displacement sensor, Phys. Rev. A **95**, 012305 (2017).
[18] J. Niset, J. Fiurášek, and N. J. Cerf, No-Go Theorem for Gaussian Quantum Error Correction, Phys. Rev. Lett. **102**, 120501 (2009).
[19] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, Caltech, 1997.
[20] E. Hostens, J. Dehaene, and B. De Moor, Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic, Phys. Rev. A **71**, 042315 (2005).
[21] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1981).
[22] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, J. Math. Phys. **43**, 4452 (2002).
[23] A. A. Kovalev and L. P. Pryadko, Spin glass reflection of the decoding transition for quantum error correcting codes, Quantum Info. Comput. **15**, 825 (2015).
[24] I. Dumer, A. A. Kovalev, and L. P. Pryadko, Thresholds for Correcting Errors, Erasures, and Faulty Syndrome Measurements in Degenerate Quantum Codes, Phys. Rev. Lett. **115**, 050502 (2015).
[25] A. J. Landahl, J. T. Anderson, and P. R. Rice, Fault-tolerant quantum computing with color codes, in *Proceedings of the QIP 2012*, Retrieved from arXiv:1108.5738.
[26] S. Glancy and E. Knill, Error analysis for encoding a qubit in an oscillator, Phys. Rev. A **73**, 012325 (2006).
[27] Strictly speaking the normalization constant, N, diverges due to $\mathbb{P}(\boldsymbol{q})$ being an infinite sum of delta peaks. In practice we always compute the quantities $Z_c(\boldsymbol{q})$ with a cutoff and can adjust N such that $1 = Z_0(\boldsymbol{q}) + Z_1(\boldsymbol{q})$.
[28] B. Deconinck, Multidimensional Theta functions, NIST digital library of mathematical functions, https://dlmf.nist.gov/21.
[29] B. Deconinck, M. Heil, A. Bobenko, M. van Hoeij, and M. Schmies, Computing Riemann theta functions, Math. Comp. **73**, 1417 (2004).
[30] J. Frauendiener, C. Jaber, and C. Klein, Efficient computation of multidimensional theta functions, arXiv:1701.07486.
[31] J. Villain, Theory of one- and two-dimensional magnets with an easy magnetization plane. II. The planar, classical, two-dimensional magnet, J. Phys. **36**, 581 (1975).
[32] W. Janke and H. Kleinert, How good is the Villain approximation? Nucl. Phys. B **270**, 135 (1986).
[33] O. M. Braun and Y. S. Kivshar, Nonlinear dynamics of the Frenkel-Kontorova model, Phys. Rep. **306**, 1 (1998).
[34] Y. S. Kivshar, H. Benner, and O. M. Braun, Nonlinear models for the dynamics of topological defects in solids, *Nonlinear Science at the Dawn of the 21st Century*, Lecture Notes in Physics Vol. 542 (Springer, Berlin, 2008).
[35] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).
[36] In fact, when one has approximate GKP states with finite photon number for ancilla and data oscillator, there is slightly more information in q than in $q \bmod 2\pi$ but we neglect this extra information here.
[37] Note that this conditional probability distribution is not identical to the likelihood function used in Ref. [16]. This might be why Ref. [16] does not observe a single cross-over point in their threshold plots.
[38] T. P. Peixoto, The graph-tool Python library, http://figshare.com/articles/graph_tool/1164194 (2014).
[39] V. Kolmogorov, Blossom V: A new implementation of a minimum cost perfect matching algorithm, Math. Program. Comput. **1**, 43 (2009).
[40] E. Fradkin and S. H. Shenker, Phase diagrams of lattice gauge theories with Higgs fields, Phys. Rev. D **19**, 3682 (1979).
[41] G. Bhanot and B. A. Freedman, Finite-size scaling for the 3D Abelian Higgs model, Nucl. Phys. B **190**, 357 (1981).
[42] C. Borgs and F. Nill, The phase diagram of the Abelian lattice Higgs model: A review of rigorous results, J. Stat. Phys. **47**, 877 (1987).
[43] M. N. Chernodub, E.-M. Ilgenfritz, and A. Schiller, Lattice study of 3D compact QED at finite temperature, Phys. Rev. D **64**, 054507 (2001).
[44] M. N. Chernodub, E.-M. Ilgenfritz, and A. Schiller, Photon Propagator, Monopoles, and the Thermal Phase Transition in Three-Dimensional Compact QED, Phys. Rev. Lett. **88**, 231601 (2002).

[45] M. N. Chernodub, R. Feldmann, E.-M. Ilgenfritz, and A. Schiller, Monopole chains in the compact Abelian Higgs model with doubly charged matter field, Phys. Lett. B **605**, 161 (2005).

[46] J. Smiseth, E. Smørgrav, F. S. Nogueira, J. Hove, and A. Sudbø, Phase structure of $(2+1)$-dimensional compact lattice gauge theories and the transition from Mott insulator to fractionalized insulator, Phys. Rev. B **67**, 205104 (2003).

[47] O. Borisenko, R. Fiore, M. Gravina, and A. Papa, Critical behavior of the compact 3D $u(1)$ gauge theory on isotropic lattices, J. Stat. Mech. (2010) P04015.

[48] R. Balian, J. M. Drouffe, and C. Itzykson, Gauge fields on a lattice. II. Gauge-invariant Ising model, Phys. Rev. D **11**, 2098 (1975).

[49] F. Wegner, Duality in generalized Ising models and phase transitions without local order parameters, J. Math. Phys. **12**, 2259 (1971).

[50] A. P. Gottlob, M. Hasenbusch, and S. Meyer, Critical behavior of the 3D XY-model: A Monte Carlo study, Nucl. Phys. B, Proc. Suppl. **30**, 838 (1993).

[51] A. M. Polyakov, Compact gauge fields and the infrared catastrophe, Phys. Lett. **59B**, 82 (1975).

[52] A. I. Larkin, Effect of inhomogeneities on the structure of the mixed state of superconductors, Zh. Exp. Teor. Fiz. **58**, 1466 (1970) [Sov.-Phys. JETP **31**, 784 (1970)].

[53] T. Giamarchi and P. Le Doussal, Statics and dynamics of disordered elastic systems, in *Spin Glasses and Random Fields*, edited by A. P. Young (World Scientific, Singapore, 1997), chap. 6.

[54] C. Wang, J. Harrington, and J. Preskill, Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory, Ann. Phys. **303**, 31 (2003).

[55] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Clarendon Press, Oxford, 2001).

[56] G. Giedke and J. Ignacio Cirac, Characterization of Gaussian operations and distillation of Gaussian states, Phys. Rev. A **66**, 032316 (2002).

[57] J. Eisert, S. Scheel, and M. B. Plenio, Distilling Gaussian States with Gaussian Operations is Impossible, Phys. Rev. Lett. **89**, 137903 (2002).

[58] A. Kubica, M. E. Beverland, F. Brandão, J. Preskill, and K. M. Svore, Three-Dimensional Color Code Thresholds via Statistical-Mechanical Mapping, Phys. Rev. Lett. **120**, 180501 (2018).

[59] V. V. Albert, S. Pascazio, and M. H. Devoret, General phase spaces: From discrete variables to rotor and continuum limits, J. Phys. A: Math. Theor. **50**, 504002 (2017).

[60] D. P. DiVincenzo, Fault-tolerant architectures for superconducting qubits, Phys. Scr. **T137**, 014020 (2009).

[61] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, Scalable Quantum Circuit and Control for a Superconducting Surface Code, Phys. Rev. Appl. **8**, 034021 (2017).

[62] N. Bourbaki, *Éléments de mathématique. III. Topologie générale. Groupes additifs de R^n* (Hermann, Paris, 1960).

[63] Strictly speaking we get $-(\nabla \times A)_p$ for a vertical p along y. But we can push this sign to the disorder component using the parity of $V_{\sigma/2}$ and use the symmetry of our noise model to ignore this sign.

[64] J. Ginibre, General formulation of Griffiths' inequalities, Comm. Math. Phys. **16**, 310 (1970).

[65] J. Fröhlich and T. Spencer, The Kosterlitz-Thouless transition in two-dimensional Abelian spin systems and the Coulomb gas, Commun. Math. Phys. **81**, 527 (1981).

[66] A. A. Kovalev, S. Prabhakar, I. Dumer, and L. P. Pryadko, Numerical and analytical bounds on threshold error rates for hypergraph-product codes, Phys. Rev. A **97**, 062320 (2018).

[67] This is not an orthogonal projection as $\ker(G)$ and $\mathrm{im}(G^{\mathrm{T}})$ are not orthogonal to one another.

[68] J. Zhang, C. Xie, K. Peng, and P. van Loock, Anyon statistics with continuous variables, Phys. Rev. A **78**, 052121 (2008).