

# Enhanced noise resilience of the surface–Gottesman-Kitaev-Preskill code via designed bias

Lisa Hänggeli, Margret Heinze \*, and Robert König

*Institute for Advanced Study and Zentrum Mathematik, Technical University of Munich, Germany*



(Received 14 April 2020; accepted 4 August 2020; published 10 November 2020)

We study the code obtained by concatenating the standard single-mode Gottesman-Kitaev-Preskill (GKP) code with the surface code. We show that the noise tolerance of this surface–GKP code with respect to (Gaussian) displacement errors improves when a single-mode squeezing unitary is applied to each mode, assuming that the identification of quadratures with logical Pauli operators is suitably modified. We observe noise-tolerance thresholds of up to  $\sigma \approx 0.58$  shift-error standard deviation when the surface code is decoded without using GKP syndrome information. In contrast, prior results by 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) and C. Vuillot, H. Asasi, Y. Wang, L. P. Pryadko, and B. M. Terhal, Quantum error correction with the toric Gottesman-Kitaev-Preskill code, *Phys. Rev. A* **99**, 032344 (2019) report a threshold between  $\sigma \approx 0.54$  and  $\sigma \approx 0.55$  for the standard (toric, respectively) surface–GKP code. The modified surface–GKP code effectively renders the mode-level physical noise asymmetric, biasing the logical-level noise on the GKP qubits. The code can thus benefit from the resilience of the surface code against biased noise. We use the approximate maximum likelihood decoding algorithm of S. Bravyi, M. Suchara, and A. Vargo, Efficient algorithms for maximum likelihood decoding in the surface code, *Phys. Rev. A* **90**, 032326 (2014) to obtain our threshold estimates. Throughout, we consider an idealized scenario where measurements are noiseless and GKP states are ideal. Our paper demonstrates that Gaussian encodings of individual modes can enhance concatenated codes.

DOI: [10.1103/PhysRevA.102.052408](https://doi.org/10.1103/PhysRevA.102.052408)

## I. INTRODUCTION

We study a modified surface–Gottesman-Kitaev-Preskill (surface–GKP) code subject to classical isotropic Gaussian displacement noise. The modification includes an additional encoding of each underlying bosonic mode by means of a single-mode Gaussian unitary, as well as an adaptation of the concatenation procedure of the two codes. Specifically, we study the case where the Gaussian unitary is given by single-mode squeezing—this is not to be confused with the term squeezing in the context of nonideal, normalizable GKP states. The squeezing effectively transforms the error from isotropic to anisotropic displacement noise. At the level of the logical GKP qubits, the anisotropic displacement noise manifests itself as biased Pauli noise. To exploit this asymmetry, we concatenate with a surface code in such a way that the primary direction of the asymmetric noise is aligned with the preferred direction of biased Pauli noise for surface code decoding. We show numerically that the bias designed this way causes an improvement of the noise tolerance threshold of this concatenated asymmetric surface–GKP code. Our result shows that—contrary to a common belief—encoding bosonic modes into other bosonic modes by means of a Gaussian unitary can be beneficial—namely, when considering concatenated codes.

Without concatenation, Gaussian encodings are indeed not beneficial. At the level of a single bosonic mode, the futility of using a Gaussian unitary as an encoding map can be illustrated using the random displacement channel (also called classical

displacement noise channel),

$$\mathcal{N}_f(\rho) = \int_{\mathbb{R}^2} f(v) D(v) \rho D(v)^\dagger d^2v, \quad (1)$$

where  $f$  is a probability density function on the phase space and  $D(v)$  is the (Weyl) displacement operator. Assume that we apply a single-mode squeezing unitary  $U = U_{S_r^{-1}}$  [cf. Eq. (24) below] with squeezing parameter  $r > 1$  before the action of the noise  $\mathcal{N}_f$ . The resulting channel can be written as  $\mathcal{N}_f(U \rho U^\dagger) = U \mathcal{N}_{\tilde{f}}(\rho) U^\dagger$  where  $\mathcal{N}_{\tilde{f}}$  is another random displacement channel with  $\tilde{f}(v) = f(\text{diag}(r^{1/2}, r^{-1/2})v)$  a squeezed version of the distribution  $f$ . For example, if  $f \equiv f_{\sigma^2}$  is a centered Gaussian with covariance matrix  $\sigma^2 I_2$  associated with an isotropic Gaussian displacement channel

$$\mathcal{N}_{f_{\sigma^2}}(\rho) = \frac{1}{2\pi\sigma^2} \int_{\mathbb{R}^2} e^{-\frac{\|v\|^2}{2\sigma^2}} D(v) \rho D(v)^\dagger d^2v, \quad (2)$$

$\|v\|^2 = v \cdot v = v_1^2 + v_2^2$ , then  $\tilde{f}$  is a centered Gaussian with covariance matrix  $\text{diag}(r^{-1}\sigma^2, r\sigma^2)$ . In particular, the strength of the “worst noise”—as expressed by the maximal singular value of the covariance matrix of the noise—cannot be lowered by the introduction of the encoding unitary  $U$ . This is a fundamental limitation on the way noise can be reshaped.

This kind of obstacle to error correction by Gaussian operations (encoding) maps was formalized more generally in Ref. [1]: It was argued that Gaussian one-mode states cannot be protected by Gaussian operations (even against Gaussian errors). Such results fall in line with a number of related no-go results concerning e.g., entanglement distillation and entanglement swapping [2–4] by means of Gaussian operations.

\*Corresponding author: [margret.heinze@ma.tum.de](mailto:margret.heinze@ma.tum.de)

A related recent no-go result [5], specifically dealing with Gaussian CV-into-CV encodings, concerns  $k$  bosonic modes encoded into  $n \geq k$  modes by means of a Gaussian unitary: Consider the problem of recovering from classical isotropic Gaussian displacement noise by means of syndrome measurements of linear combinations of the quadratures followed by maximum likelihood decoding. Here the noise is described by a random variable with centered normal distribution  $Z \sim \mathbf{N}(\mathbf{0}, \sigma^2 I_{2n})$  on the phase space  $\mathbb{R}^{2n}$ . The authors of Ref. [5] show that the resulting effective logical noise is displacement noise with a centered normal distribution, whose covariance matrix has eigenvalues  $\{\sigma_P^{(j)}, \sigma_Q^{(j)}\}_{j=1}^k$  satisfying  $\sigma_P^{(j)}\sigma_Q^{(j)} = \sigma^2$  for every  $j = 1, \dots, k$ . In other words, the worst-case noise variance (per quadrature) does not improve even at the logical (encoded) level when using a Gaussian encoding map. Overall, these results appear to suggest that readily available Gaussian unitaries may not be used to boost noise tolerance levels of bosonic error-correcting codes.

We show that when bosonic degrees of freedom are used to encode logical qubits by means of concatenated codes, this apparent no-go theorem no longer applies. Indeed, we find that the introduction of additional single-mode Gaussian squeezing unitaries in surface-GKP codes increases fault-tolerance error thresholds, implying that more noise can be tolerated. Here the error-correction threshold of an infinite code family of bosonic codes is the maximum physical (single-mode) displacement noise standard deviation  $\sigma$  (for a centered normal distribution) such that the logical failure rate can be made arbitrarily small by using sufficiently large code sizes.

A scenario where additional Gaussian unitaries improve a (unconcatenated) GKP code has been discussed previously in the literature: Recall that GKP codes are constructed algebraically from certain lattices in  $\mathbb{R}^{2n}$  [6]. The chosen lattice determines the code space dimension  $K$  as well as the code's robustness against displacements. In particular, the size (norm) of the smallest uncorrectable shift depends on the lattice: Shifts inside the Voronoi cell of the dual lattice are correctable. As observed early on [6], this implies, for example, that in two dimensions the GKP code based on a hexagonal lattice beats the "standard" square lattice GKP code in this respect, being able to tolerate larger displacements. (This is related to the fact that the hexagonal lattice permits the densest sphere packing in  $\mathbb{R}^2$ , see Ref. [7].) More detailed estimates of the difference between hexagonal and square lattice GKP codes were obtained in Ref. [8] in terms of estimates of the logical error probability for a single encoded qubit in the presence of pure-loss noise. Furthermore, the characterization of correctable displacements in terms of Voronoi cells has also been used to derive bounds on the (quantum and classical) capacity of the channel (1), see Refs. [6] and [9]. Importantly, the hexagonal lattice GKP code is related to the square lattice GKP code by a Gaussian unitary (see Sec. III E). Thus the square lattice GKP code can be enhanced by a simple application of a Gaussian unitary. This improvement may appear minor for a single mode (e.g., it changes a constant prefactor from  $\pi/4$  to  $\pi/(2\sqrt{3})$ ) in the logical error probability considered in Ref. [8]). However, as we argue here, similar modifications lead to dramatic improvements in the setting of concatenated codes.

The underlying mechanism which improves fault-tolerance properties in the setting of concatenated codes is not simply

a matter of increasing the volume of a Voronoi cell. Instead, the introduction of an additional encoding map has the effect of artificially shaping the logical-level noise, yielding biased qubit noise. If the GKP code is appropriately concatenated with the surface code, this leads to improved fault-tolerance properties due to the fact that known decoders for the surface code can benefit from a noise bias. The same mechanism therefore applies to other codes resilient to biased noise. In this sense, our work on surface-GKP codes is merely a case study illustrating this principle.

In more detail, we seek to design optimized codes protecting against single-mode i.i.d. (independently and identically distributed) displacement noise of the form (2) on each mode. Starting from a standard square lattice GKP code, we apply a single-mode squeezing unitary to every mode yielding a rectangular, i.e., asymmetric lattice GKP code. As already mentioned, applying the one-mode squeezing operation with parameter  $r > 1$  [see Eq. (24) below], we are effectively dealing with anisotropic noise  $\mathcal{N}_{f_{Z_r}}$ , where  $Z_r \sim \mathbf{N}(\mathbf{0}, \Sigma_r)$  is an anisotropically distributed Gaussian random variable with covariance matrix

$$\Sigma_r = \begin{pmatrix} \sigma^2/r & 0 \\ 0 & r\sigma^2 \end{pmatrix} =: \text{diag}(\tilde{\sigma}_Q^2, \tilde{\sigma}_P^2). \quad (3)$$

In other words, this is equivalent to error correction with square lattice GKP codes subject to asymmetric noise (for  $r > 1$ ): the noise variance  $\sigma_Q^2$  in the  $Q$  quadrature is reduced, while the variance  $\sigma_P^2$  in the  $P$ -quadrature is increased. After decoding the GKP qubit, this results in biased Pauli noise: a fact that we can exploit when considering surface-GKP codes. For this to work, the concatenation of codes needs to be done in such a way as to make the primary direction of the noise asymmetry align suitably with the surface code, i.e., those (Pauli) axes where the noise bias is favorable for the surface code decoder.

The fact that surface codes perform well under certain biased noise has been established recently [18,34]; we briefly review these results in Sec. IV A. Our main contribution is a detailed discussion of how the surface-GKP code needs to be modified to exploit this property of the surface code. This is detailed in Sec. IV B. We also provide numerical evidence showing that this strategy indeed provides higher error thresholds (see Sec. V).

### A. Prior work

Among strategies to go beyond the no-go results on Gaussian encoding maps is Ref. [10]. Here Noh *et al.* show that Gaussian CV-into-CV encodings are meaningful if combined with non-Gaussian resources. By using GKP states and modular quadrature measurements, they define new families of non-Gaussian CV-into-CV encoding maps derived from (qubit) stabilizer codes. Here we take a different approach and consider DV-into-CV encoding maps obtained by concatenating surface codes with modified (asymmetric) square lattice GKP codes.

We note that the codes we study here are closely related to the surface-GKP and toric-GKP codes previously investigated in pioneering work by Fukui *et al.* [11] and Vuillot *et al.* [5], respectively, and further analyzed by Noh and Chamberland [12]. These authors study the concatenation of

the GKP code with the surface (toric) code, providing error threshold estimates. They consider a noise model which is i.i.d. Gaussian noise with isotropic distribution  $\mathbf{N}(\mathbf{0}, \sigma^2 I_2)$  on every mode. In Refs. [5] and [12], additional noise in the syndrome measurements respectively full circuit-level noise is considered. Numerically, threshold estimates are obtained in Refs. [5,11,12] by level-by-level concatenated decoding: Modular position and momentum measurements are performed to extract GKP syndrome information for every GKP qubit. By Bayesian update, an effective Pauli-error distribution is computed, either conditioned on the particular GKP syndrome information (meaning that this prior GKP information is taken into account when decoding) or averaged over the GKP measurement outcomes. Subsequently, Edmond’s maximum matching algorithm [13] (in the following referred to as the minimum weight matching decoder) is applied to decode the surface code given this effective Pauli noise: For the case where GKP information is exploited, the latter is heuristically incorporated into a weighting of the edges. Additional (heuristic) decoders are studied in Ref. [5] in the case where the GKP syndrome information (bosonic measurements) is noisy. Furthermore, the threshold error probability is related to a phase transition in a classical statistical model in this case.

Closely related to our work are approaches to exploiting asymmetry in (Gaussian) displacement noise by means of code concatenation, e.g., of cat codes with the repetition code. An additional key ingredient we use is the behavior of the surface code with respect to biased Pauli noise. To highlight the relationship to our work, we briefly discuss these prior works in Sec. IV A after introducing some more terminology.

**B. Outline**

We begin by presenting some background material to fix notation and set the stage. In Sec. II, we review the surface code, maximum likelihood decoding, and the Bravyi–Suchara–Vargo (BSV) decoder. In Sec. III, we discuss different versions of the GKP code as well as error recovery procedures with and without side information.

We then present our main findings: In Sec. IV, we introduce the (modified) surface— code and show how the introduction of single-mode Gaussian unitaries effectively leads to biased Pauli noise. In Sec. V, we present our numerical results for the noise tolerance threshold using modified surface–GKP codes. We conclude in Sec. VI.

**II. THE SURFACE CODE**

In this section, we review surface codes [14,15]. These are CSS (stabilizer) codes with geometrically local generators in two dimensions, encoding one logical qubit into  $n$  physical qubits. Subsequently, we discuss the decoding problem for these codes. In particular, we review the definition of maximum likelihood decoding for general stabilizer qubit codes. We then discuss the BSV decoder (by Bravyi *et al.* [16]), which approximates the maximum likelihood decoder for the surface code.

**A. Definition of the surface code**

Consider a  $(L - 1) \times (L - 1)$  square lattice, i.e., a rectangular grid with  $L - 1$  edges on each side. If we first identify

its vertical (left and right) boundaries, yielding a cylinder, and subsequently perform a vertical cut located in between any two vertical lines of the lattice, the result (mapped back to the plane) is the grid on which a  $L \times L$  surface code with “smooth” top and bottom boundaries and “rough” left and right boundaries is defined, cf. Fig. 1(a) below. These are the only surface codes we consider here for simplicity. Note, however, that a  $r \times s$  surface code for arbitrary  $r, s \geq 1$  can be defined similarly, and also the boundary conditions may be chosen differently, see Ref. [14] as well as Ref. [17] for the so-called rotated surface codes.

The surface code is now defined as follows: to every edge of the grid a qubit is attached, yielding a total number of  $n = L^2 + (L - 1)^2$  physical qubits. Denote by  $v$ ,  $e$ , and  $p$  a vertex, edge, and plaquette of the grid, respectively. The (co)boundaries  $\delta v$ ,  $\partial p$  consist of the edges incident on  $v$  and the edges bounding  $p$ , respectively. By the usual convention, the stabilizer generators of the surface code are on the one hand given by tensor products  $A_v := \prod_{e \in \delta v} X_e$  of Pauli- $X$  operators acting on the edges incident to a vertex  $v$ , and on the other hand by tensor products  $B_p := \prod_{e \in \partial p} Z_e$  of Pauli- $Z$  operators acting on the edges surrounding a plaquette  $p$ . Denote by  $\mathbf{V}$  the set of vertices, and by  $\mathbf{P}$  the set of plaquettes of the grid. Since  $|\mathbf{V}| = |\mathbf{P}| = L(L - 1)$ , we have  $2L(L - 1)$  independent stabilizer generators, yielding  $k = n - 2L(L - 1) = 1$  encoded qubits. The centralizer of the stabilizer group  $\mathcal{S} := \langle A_v, B_p \rangle$  with respect to the  $n$ -qubit Pauli group  $\mathcal{P}_n$ , where  $v, p$  run over the sets  $\mathbf{V}, \mathbf{P}$ , is given by

$$\mathcal{C}(\mathcal{S}) := \{P \in \mathcal{P}_n \mid PS = SP \text{ for all } S \in \mathcal{S}\}. \tag{4}$$

We can define logical Pauli- $X$  (Pauli- $Z$ ) operators, denoted by  $\bar{X}$  ( $\bar{Z}$ ), as the product of Pauli- $X$  (Pauli- $Z$ ) operators along the left (top) boundary of the lattice, cf. Fig. 1(a). Then  $\bar{X}, \bar{Z} \in \mathcal{C}(\mathcal{S}) \setminus \mathcal{S}$ , and  $\bar{X}\bar{Z} = -\bar{Z}\bar{X}$ , as required. Moreover, the weight of both  $\bar{X}$  and  $\bar{Z}$  is  $L$ . These are minimal-weight logical operators, i.e., the code distance is  $d = L$ .

**B. Error recovery for the surface code**

Recall that recovery from errors (i.e., after the action of a noise channel) in a stabilizer code proceeds by measurement of stabilizer operators (cf. Sec. III F 1) and subsequent application of a recovery map (mapping back to the code space).

A noise channel in this context is a CPTP map  $\mathcal{N} : \mathcal{D}((\mathbb{C}^2)^{\otimes n}) \rightarrow \mathcal{D}((\mathbb{C}^2)^{\otimes n})$  on the density operators  $\mathcal{D}((\mathbb{C}^2)^{\otimes n})$  on the Hilbert space  $(\mathbb{C}^2)^{\otimes n}$ . Here we consider probabilistic Pauli noise, i.e., noise channels of the form

$$\mathcal{N}_\pi(\rho) = \sum_{E \in \mathcal{P}_n} \pi(E) E \rho E^\dagger, \tag{5}$$

with  $\pi$  a probability distribution on the Pauli group  $\mathcal{P}_n$ . Note that in our applications later on, we (predominantly) consider i.i.d. channels of the form  $\mathcal{N}_{p^n}(\rho) = \mathcal{N}_{p=(p_I, p_X, p_Y, p_Z)}^{\otimes n}(\rho)$ , where the single-qubit noise channel  $\mathcal{N}_p$  represents random Pauli noise [see Eq. (47) below]. We also consider channels of the form  $\mathcal{N}_{\prod_{j=1}^n p^{(j)}} = \otimes_{j=1}^n \mathcal{N}_{p^{(j)}}$  where each qubit  $j \in \{1, \dots, n\}$  experiences random Pauli noise given by a distribution  $p^{(j)} = (p_I^{(j)}, p_X^{(j)}, p_Y^{(j)}, p_Z^{(j)})$ . However, such a restriction does not need to be imposed at this point.

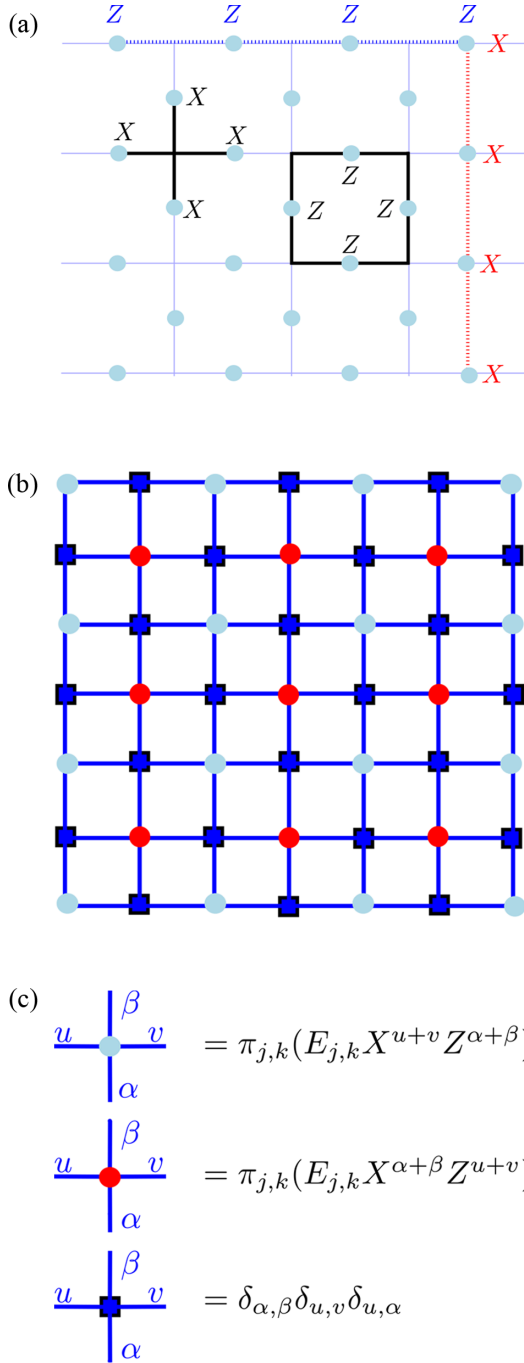


FIG. 1. Tensor network associated with the BSV decoder. Figure 1(a) shows the surface code with  $d = 4$ , logical operators  $\bar{X}$  and  $\bar{Z}$  and two stabilizer generators. Figure 1(b) gives the tensor network whose contraction gives the value  $\pi(ES)$  of a coset determined by a Pauli operator  $E = \otimes_{(j,k)} E_{j,k}$  acting with the single-qubit Pauli operator  $E_{j,k}$  on the qubit at location  $(j, k)$ . Evaluating this scalar for  $E \in \{E_s, E_s \bar{X}, E_s \bar{Y}, E_s \bar{Z}\}$ , where  $E_s$  is a representative error giving syndrome  $s$ , allows us to perform maximum likelihood decoding according to Eq. (7). Local tensors are defined as shown in Fig. 1(c). (Tensors near the boundaries are defined similarly.) *A priori* Pauli-error probabilities for a qubit located at  $(j, k)$  are denoted  $\pi_{j,k}(Q) = p_Q$ , where  $p_Q$  is the *a priori* probability of a Pauli error  $Q \in \{I, X, Y, Z\}$ . (a) Surface code with distance  $d = 4$ . (b) Tensor network for coset probability  $\pi(ES)$ . (c) Bulk tensors; the first two associated with a qubit at  $(j, k)$ .

The stabilizers we measure in the surface code are the (commuting) vertex- and plaquette-type stabilizers  $\{A_v\}_{v \in V}$ ,  $\{B_p\}_{p \in P}$ . This provides a *syndrome* (measurement outcome)  $s = (\{s_v\}_{v \in V}, \{s_p\}_{p \in P}) \in \{0, 1\}^{|V|+|P|}$  and simultaneously projects the state into the (common) eigenspaces of each  $A_v$  associated with eigenvalue  $(-1)^{s_v}$ , and the eigenspaces of each  $B_p$  associated with eigenvalue  $(-1)^{s_p}$ . A syndrome  $s$  is observed with probability

$$p(s) := \text{tr}[\Pi(s)\mathcal{N}_{q^n}(\rho)], \quad \text{where}$$

$$\Pi(s) = \prod_{p \in P} \frac{1}{2}[I + (-1)^{s_p} B_p] \prod_{v \in V} \frac{1}{2}[I + (-1)^{s_v} A_v],$$

and the postmeasurement state is given by  $p(s)^{-1}\Pi(s)\mathcal{N}_{q^n}(\rho)\Pi(s)$ . A recovery procedure mapping the postmeasurement state back to the code space is described by a unitary correction operation  $C(s)$  depending on the syndrome  $s$ . Choosing a recovery procedure (sometimes referred to as a decoder) thus amounts to the choice of a function  $s \mapsto C(s)$  associating a correction operation to a given syndrome.

The success probability of a given recovery strategy can be analyzed by considering a single  $n$ -qubit Pauli error  $E \in \mathcal{P}_n$ . The above recovery procedure [using Pauli corrections] successfully recovers from an error  $E$  if the following conditions are satisfied:

(i) First, the coset  $C(s)\mathcal{C}(S)$  of the chosen correction operation  $C(s)$  coincides with the coset  $EC(S)$  of the actual error  $E$ . Here  $\mathcal{C}(S)$  is the centralizer of  $S$  in the Pauli group, see (4). This means that  $C(s)$  causes the same syndrome as  $E$ , that is,  $C(s)$  maps the corrupted state back to the code space.

(ii) Second, the chosen correction operation  $C(s)$  belongs to the same coset of  $S$  inside  $EC(S)$ . More precisely, for every syndrome  $s$ , fix a “representative error”  $E_s$  which causes syndrome  $s$ . Then the coset  $E_s\mathcal{C}(S)$  can be partitioned into

$$E_s\mathcal{C}(S) = E_s S \cup E_s \bar{X} S \cup E_s \bar{Y} S \cup E_s \bar{Z} S, \quad (6)$$

where all elements in each of the four subsets have the same logical action. Decoding is successful if for the syndrome  $s$  caused by  $E$  [that is,  $EC(S) = E_s\mathcal{C}(S)$ ],  $E$  and the correction  $C(s)$  belong to the same subset on the right hand side of (6).

### 1. Maximum likelihood decoding

Given the available syndrome information  $s$ , the optimal choice of recovery map  $C(s)$  is determined by the maximum likelihood decoding strategy: one sets

$$C_{\text{ML}}(s) := \arg \max_{C \in \{E_s S, E_s \bar{X} S, E_s \bar{Y} S, E_s \bar{Z} S\}} \pi(C), \quad (7)$$

i.e., finds the coset with maximal weight with respect to the error distribution  $\pi$  [cf. (5)], and subsequently selects a correction  $C_{\text{ML}}(s) \in C_{\text{ML}}(s)$  belonging to this coset (arbitrarily). This guarantees that the correction satisfies (i) and simultaneously maximizes the probability that it obeys (ii). The resulting (averaged) success probability is then given by

$$P_{\text{success}} = \sum_s \pi(C_{\text{ML}}(s)).$$



**2. The Bravyi-Suchara-Vargo (BSV) decoder**

While the maximum likelihood decoder is optimal from the point of view of decoding error probability, the determination of the coset (7) is nontrivial (even for simple i.i.d. error models) as computing the coset probabilities involves sums over sets of exponential size. In other words, the maximum likelihood decoder cannot be realized efficiently.

To overcome this obstacle, Bravyi *et al.* [16] have proposed a decoding strategy  $\tilde{C}_{ML}^\chi(s)$  which approximates the maximum likelihood decoder  $C_{ML}(s)$ . The decoding strategy depends on a parameter  $\chi \in \mathbb{N}$  (the *bond dimension*) and becomes exact, i.e., maximum likelihood decoding, in the limit  $\chi \rightarrow \infty$ . In contrast to the maximum likelihood decoder, it is efficiently computable (for small, i.e., typically constant  $\chi$ ): the computation of  $\tilde{C}_{ML}^\chi(s)$  from  $s$  involves  $O(n\chi^3)$  basic arithmetic operations. The proposed decoder generally applies to “local” stochastic error models for 2D stabilizer codes with local generators, see, e.g., the Appendix of Ref. [18] for an application to 2D color codes.

To briefly summarize the construction of the BSV decoder, consider the case of Pauli noise where the distribution  $\pi$  over Pauli errors factorizes into a product of (not necessarily identical but independent) distributions over single-qubit errors. The key insight of Ref. [16] is that coset probabilities  $\pi(ES)$  can then be expressed as contractions of planar 2D tensor networks, see Fig. 1. Such tensor networks are known to be efficiently (approximately) contractible by contracting along a single axis, effectively rendering the problem one-dimensional. In more detail, a coset probability can be written as

$$\pi(ES) = \langle \Psi_1 | M_1 \cdots M_{2d-2} | \Psi_2 \rangle,$$

where  $|\Psi_1\rangle, |\Psi_2\rangle \in (\mathbb{C}^2)^{\otimes 2d-1}$  are matrix product states (MPSs) with bond dimension 2, and each  $M_j \in \mathcal{B}((\mathbb{C}^2)^{\otimes 2d-1})$  is a certain matrix product operator (MPO) of bond dimension 2. The idea then is to evaluate the right-hand side stepwise by applying the MPOs in succession, and eventually computing the inner product of two MPS. However, successive multiplication of an MPS with constant-bond dimension MPOs generically leads to an MPS with exponential bond dimension (in the number of applications of MPOs), rendering an exact evaluation inefficient. To avoid this overhead, the BSV decoder replaces intermediate MPS obtained during the computations by MPS approximations having bond dimension bounded by  $\chi$ . The latter step involves a standard truncation procedure for MPS by Murg *et al.* [19], which singles out the largest singular values across each cut.

In Ref. [16], this decoder is compared to an exact (efficiently realized) maximum likelihood decoder in the case of pure  $X$  noise, showing that moderate values of  $\chi$  suffice to achieve good accuracy. It was then used to show that the minimum weight matching decoder is typically suboptimal. The BSV decoder was subsequently applied in Refs. [18,34] to study the performance of surface codes under biased noise.

**III. GOTTESMAN-KITAEV-PRESKILL (GKP) CODES**

In this section, we review the pertinent facts about GKP codes [6]. We begin with a general discussion of bosonic

quantum systems in Sec. III A. In Sec. III B, we discuss displacement noise, a central noise model of interest in bosonic systems. We then review the general definition of a GKP code associated with a symplectically integral lattice (Sec. III C) before specializing to square lattice (Sec. III D) and asymmetric (Sec. III E) GKP codes. Finally, in Sec. III F, we discuss the decoding problem for GKP codes.

**A. Continuous-variable (CV) quantum systems**

For concreteness, we focus on a single bosonic mode ( $n = 1$ ); the generalization to  $n > 1$  modes is straightforward. A pure state of a single mode, i.e., of a particle on a line, is given by a (equivalence class with respect to a global phase of a) tempered distribution  $\psi \in \mathcal{S}'(\mathbb{R})$ . We call the state normalizable if  $\psi \in L^2(\mathbb{R})$ .

Let  $Q, P$  denote the position and momentum operators, also called quadratures, acting on  $\mathcal{S}'(\mathbb{R})$  and satisfying the canonical commutation relation [20]  $[Q, P] = iI$ . We collect them in a vector  $R := (Q \ P)^T$  such that the commutation relation takes the form  $[R_j, R_k] = iJ_{j,k}I$ . Here the matrix

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is associated with the symplectic form  $(\xi, \eta) \mapsto \xi^T J \eta$  on  $\mathbb{R}^2 \times \mathbb{R}^2$ .

For  $\xi = (\xi_1 \ \xi_2)^T \in \mathbb{R}^2$ , the *displacement operator*  $D(\xi)$  is defined as

$$D(\xi) := e^{-i\xi^T J R} = e^{i(\xi_2 Q - \xi_1 P)}. \tag{8}$$

These operators yield a representation  $(\xi, \alpha) \mapsto e^{-i\alpha} D(\xi)$  of the Weyl-Heisenberg group on  $\mathcal{S}'(\mathbb{R})$  which restricts to a unitary irreducible representation on  $L^2(\mathbb{R})$ . That is, we have

$$e^{-i\alpha} D(\xi) e^{-i\beta} D(\zeta) = e^{-i(\alpha+\beta+\frac{1}{2}\xi^T J \zeta)} D(\xi + \zeta), \tag{9}$$

for all  $\alpha, \beta \in \mathbb{R}$ , and  $\xi, \zeta \in \mathbb{R}^2$ . We note that the operator  $D(\xi)$  translates—by conjugation—the vector  $(Q \ P)^T$  of mode operators by the amount  $\xi \in \mathbb{R}^2$  in phase space.

The group  $\text{Sp}(2, \mathbb{R})$  of symplectic linear maps on  $\mathbb{R}^2$ , i.e., linear maps preserving the symplectic form, is

$$\text{Sp}(2, \mathbb{R}) := \{S \in \text{Mat}_{2 \times 2}(\mathbb{R}) \mid S^T J S = J\}.$$

The metaplectic representation  $S \mapsto U_S$  associates a Gaussian unitary  $U_S$  on  $\mathcal{S}'(\mathbb{R})$  to every  $S \in \text{Sp}(2, \mathbb{R})$ . The action of  $U_S$  (by conjugation) on the quadrature operators is linear and given by  $S$ , i.e.,

$$U_S R_j U_S^\dagger = \sum_{k=1}^2 S_{j,k} R_k, \quad j = 1, 2.$$

In particular,  $U_S$  preserves canonical commutation relations. One consequence is that for arbitrary  $S \in \text{Sp}(2, \mathbb{R})$  and  $\xi \in \mathbb{R}^2$ , we have

$$D(S\xi) = U_{S^{-1}} D(\xi) U_{S^{-1}}^\dagger. \tag{10}$$

In other words,  $D(S\xi)$  is unitarily equivalent to  $D(\xi)$  with the corresponding unitary conjugation only depending on  $S$  but not on  $\xi$ .

### B. Displacement noise in bosonic systems

GKP codes (named after Ref. [6] by Gottesman, Preskill, and Kitaev) isometrically embed a finite-dimensional quantum system  $\mathbb{C}^K$  into  $L^2(\mathbb{R})$  (and, more generally,  $L^2(\mathbb{R})^{\otimes n}$ ). Although their performance under various forms of noise has been studied in the original paper [6] and a number of subsequent papers [21,22], the codes were originally designed primarily to protect against displacement noise, i.e., noise channels of the form

$$\mathcal{N}_{f_Z}(\rho) = \int_{\mathbb{R}^2} f_Z(v) D(v) \rho D(v)^\dagger d^2v. \quad (11)$$

Here  $f_Z : \mathbb{R}^2 \rightarrow [0, 1]$  is a probability density function associated with a random variable  $Z$  on the phase space  $\mathbb{R}^2$  and  $D(v)$  is the displacement operator.

Of specific interest is the case where  $Z \sim \mathbf{N}(\mathbf{0}, \Sigma)$ , i.e., is described by a centered normal distribution with (positive semidefinite) covariance matrix  $\Sigma$ , hence

$$f_Z(v) = (2\pi)^{-1} \det(\Sigma)^{-1/2} e^{-\frac{1}{2}v^T \Sigma^{-1}v}.$$

In the isotropic case,  $\Sigma = \sigma^2 I_2$ , the channel (11) becomes the isotropic Gaussian displacement channel (2). The single-parameter family of noise channels which results by varying  $\sigma$  provides a natural test bed for assessing the noise resilience of GKP and other codes (see, e.g., Ref. [9]), as well as related capacity questions. In this context, a recent breakthrough [23] has confirmed a long-standing conjecture introduced in the seminal work [24].

### C. The GKP code $\text{GKP}(\mathcal{L})$

GKP codes are constructed algebraically from certain lattices  $\mathcal{L} \subset \mathbb{R}^{2n}$ . We first restrict our attention to  $n = 1$  which is sufficient for our purposes, i.e., we consider GKP codes for a single mode encoding a  $K$ -dimensional system. We then further restrict to  $K = 2$ , i.e., codes encoding a single qubit into a single-mode bosonic system [25]. We note that the construction discussed here was previously known [26,27] and rigorously discussed by Bouzouina and De Bièvre [28] in 1996. Its potential in terms of quantum error correcting codes was, however, only recognized by GKP in 2001 [6].

#### 1. Definition of the code space

Let us review the construction of a code space  $\text{GKP}(\mathcal{L}) \subset \mathcal{S}'$  based on a lattice  $\mathcal{L} \subset \mathbb{R}^2$  with certain properties. Specifically,  $\mathcal{L}$  needs to be *symplectically integral*: There are vectors  $\xi_1, \xi_2 \in \mathbb{R}^2$  and  $K \in \mathbb{N}$  such that

$$\xi_1^T J \xi_2 = \pm 2\pi K, \quad (12)$$

and  $\mathcal{L} := \{n_1 \xi_1 + n_2 \xi_2 \mid n_1, n_2 \in \mathbb{Z}\}$  is the (integer) span of  $\{\xi_1, \xi_2\}$ . Given such a lattice  $\mathcal{L}$ , the space  $\text{GKP}(\mathcal{L})$  is defined as the (simultaneous)  $+1$  eigenspace of all (pairwise commuting) operators  $D(\xi)$ ,  $\xi \in \mathcal{L}$ . [Condition (12) implies that  $\xi^T J \xi' \in 2\pi\mathbb{Z}$  for all  $\xi, \xi' \in \mathcal{L}$ , which, because of (9), ensures that that these operators commute.] That is, the stabilizer group  $\mathcal{S}$  of  $\text{GKP}(\mathcal{L})$  is given by

$$\mathcal{S} = \{D(\xi) \mid \xi \in \mathcal{L}\} = \langle D(\xi_1), D(\xi_2) \rangle.$$

The space  $\text{GKP}(\mathcal{L})$  can be interpreted as the Hilbert space of a quantum system with phase space  $\mathbb{R}^2/\mathcal{L}$ : states belonging to  $\text{GKP}(\mathcal{L})$  are invariant under lattice transformations realized by displacements  $D(\xi)$ ,  $\xi \in \mathcal{L}$ .

We next discuss the effect of various (error) operators on the code space  $\text{GKP}(\mathcal{L})$ . Because the displacements  $D(\zeta)$ ,  $\zeta \in \mathbb{R}^2$  form an operator basis, we restrict our attention to displacements (similar to the way qubit stabilizer codes are analyzed in terms of Pauli operators). In other words, we are interested in the effect of an arbitrary displacement  $D(\zeta)$ ,  $\zeta \in \mathbb{R}^2$ . A first question is which of these operators preserve the subspace  $\text{GKP}(\mathcal{L})$ : Ignoring irrelevant global phases, we are interested in the centralizer of the stabilizer group in the Heisenberg-Weyl group, i.e., the set

$$\begin{aligned} \mathcal{C}(\mathcal{S}) &:= \{D(\zeta) \mid \zeta \in \mathbb{R}^2 \text{ such that} \\ &D(\zeta)S = SD(\zeta) \text{ for all } S \in \mathcal{S}\}. \end{aligned}$$

The set of these operators is characterized by the set of vectors  $\zeta \in \mathbb{R}^2$  satisfying

$$\zeta^T J \xi \in 2\pi\mathbb{Z} \quad \text{for all } \xi \in \mathcal{L},$$

cf. (9). These vectors form a lattice on their own: the (symplectically) dual lattice  $\mathcal{L}^\perp$ . Because of (12), we can choose its generating vectors  $\xi_1^\perp, \xi_2^\perp \in \mathbb{R}^2$  such that

$$\xi_i^{\perp T} J \xi_j = 2\pi \delta_{ij}, \quad i, j \in \{1, 2\}. \quad (13)$$

Thus the centralizer of the stabilizer group is given by

$$\mathcal{C}(\mathcal{S}) = \{D(\zeta) \mid \zeta \in \mathcal{L}^\perp\} = \langle D(\xi_1^\perp), D(\xi_2^\perp) \rangle.$$

Operators belonging to  $\mathcal{C}(\mathcal{S})$  are logical, i.e., preserve the code space  $\text{GKP}(\mathcal{L})$ , but may or may not act nontrivially on it. Elements  $D(\zeta), D(\zeta') \in \mathcal{C}(\mathcal{S})$  have identical logical action if and only if they differ by multiplication by an element in  $\mathcal{S}$ , that is, if and only if  $\zeta - \zeta' \in \mathcal{L}$ . Thus

$$\{D(\zeta) \mid [\zeta] \in \mathcal{L}^\perp/\mathcal{L}\} \quad (14)$$

is a complete family of inequivalent logical errors indexed by the set  $\mathcal{L}^\perp/\mathcal{L} = \{[\zeta] \mid \zeta \in \mathcal{L}^\perp\}$  of cosets of  $\mathcal{L}$ , where  $[\zeta] := \zeta + \mathcal{L}$  is a coset with representative  $\zeta \in \mathcal{L}^\perp$ .

#### 2. Voronoi cells and lattice modulo operation

As discussed below, logical error probabilities in the GKP error recovery process heavily depend on properties of the underlying lattices  $\mathcal{L}$  and  $\mathcal{L}^\perp$ —especially the Voronoi cell of  $\mathcal{L}^\perp$ . Let us briefly define the latter set for an arbitrary lattice  $\mathcal{L} \subset \mathbb{R}^2$ . For  $x \in \mathbb{R}^2$ , the closest lattice point in  $\mathcal{L}$  to  $x$  is

$$Q_{\mathcal{L}}(x) := \arg \min_{\xi \in \mathcal{L}} \|x - \xi\|.$$

Here the Euclidean distance is used and we assume that ties are broken in a systematic manner (e.g., such that  $Q_{\mathcal{L}}(x)_i \geq x_i$ ,  $i = 1, 2$ ). The Voronoi cell  $\mathcal{V}$  of  $\mathcal{L}$  is the set of points closest to the origin  $\mathbf{0} := (0 \ 0)^T$ , i.e.,

$$\mathcal{V} := \{x \in \mathbb{R}^2 \mid Q_{\mathcal{L}}(x) = \mathbf{0}\}.$$

For later use, we also define a lattice modulo operation by

$$\begin{aligned} (\cdot \pmod{\mathcal{L}}) : \mathbb{R}^2 &\rightarrow \mathcal{V} \\ x &\mapsto x \pmod{\mathcal{L}} := x - Q_{\mathcal{L}}(x). \end{aligned}$$

**D. Square lattice GKP codes**

In the following two sections, we list several examples of lattices giving rise to GKP codes. We begin with a discussion of the standard square lattice GKP code. We also refer to this as the symmetric GKP code below.

**1. Square lattice GKP codes of dimension  $K$**

Let  $K \in \mathbb{N}$  be arbitrary. Then the vectors

$$\begin{aligned} \xi_1 &= (\sqrt{2\pi K} \quad \mathbf{0})^T, \\ \xi_2 &= (\mathbf{0} \quad \sqrt{2\pi K})^T, \end{aligned} \tag{15}$$

satisfy the condition (12) and span a symplectically integral lattice  $\mathcal{L}$ . Its dual lattice  $\mathcal{L}^\perp$  is spanned by the vectors

$$\begin{aligned} \xi_1^\perp &= (\mathbf{0} \quad -\sqrt{2\pi/K})^T, \\ \xi_2^\perp &= (\sqrt{2\pi/K} \quad \mathbf{0})^T. \end{aligned} \tag{16}$$

In particular, the set of cosets  $\mathcal{L}^\perp/\mathcal{L}$  associated with logical errors [cf. (14)] is given by

$$\mathcal{L}^\perp/\mathcal{L} = \{[n_1\xi_1^\perp + n_2\xi_2^\perp] \mid n_1 \in \{0, \dots, K-1\}, n_2 \in \{0, \dots, K-1\}\}. \tag{17}$$

A detailed description of the action of the associated logical errors can be given by fixing a basis of  $\text{GKP}(\mathcal{L})$ . Let  $\psi \in \text{GKP}(\mathcal{L}) \subset \mathcal{S}'$ . Then the eigenvalue equation  $D(\xi_2)\psi = e^{i\sqrt{2\pi K}Q}\psi = \psi$  implies that  $\psi \in \mathcal{S}'$  is of the form

$$\psi(x) = \sum_{n \in \mathbb{N}} c_n \delta(x - n\sqrt{2\pi/K}), \tag{18}$$

with  $c_n \in \mathbb{C}$ , and with  $\delta$  denoting the Dirac- $\delta$  distribution. The eigenvalue equation  $D(\xi_1)\psi = e^{-i\sqrt{2\pi K}P}\psi = \psi$  further implies that

$$c_n = c_{n+K} \quad \text{for all } n \in \mathbb{N}. \tag{19}$$

From (18) and (19), we conclude that the dimension of  $\text{GKP}(\mathcal{L})$  is  $K$ , consistent with the fact that (17) shows the existence of  $K^2$  linearly independent logical operators. Furthermore, the set  $\{e_j\}_{j=0}^{K-1}$  defined by

$$e_j(x) = \sum_{n \in \mathbb{N}} \delta(x - n\sqrt{2\pi K} - j\sqrt{2\pi/K}) \tag{20}$$

for  $j = 0, \dots, K-1$ , defines a basis of  $\text{GKP}(\mathcal{L})$ . The action of logical error operators  $D(n_1\xi_1^\perp + n_2\xi_2^\perp) \propto D(\xi_1^\perp)^{n_1} D(\xi_2^\perp)^{n_2}$  [cf. (17)] on these basis vectors can be computed to be

$$\begin{aligned} D(\xi_1^\perp)e_j &= e^{-\frac{2\pi ij}{K}} e_j, \\ D(\xi_2^\perp)e_j &= e_{(j+1) \bmod K} \end{aligned} \tag{21}$$

for all  $j = 0, \dots, K-1$ . That is,  $(D(-\xi_1^\perp), D(\xi_2^\perp))$  are the (generalized) logical Pauli  $\bar{Z}$  and  $\bar{X}$  operators of a  $K$ -dimensional system. They generate what is sometimes referred to as a finite-dimensional Weyl system.

**2. Square lattice GKP codes encoding a qubit ( $K = 2$ )**

In the following, we specialize to the square lattice GKP code encoding a qubit, i.e., the case where  $K = 2$ . Explicitly,

the square lattice and its dual generated by the vectors (15), respectively, (16) are

$$\begin{aligned} \mathcal{L}_\square &:= \{(2\sqrt{\pi} n_1 \quad 2\sqrt{\pi} n_2)^T \mid n_1, n_2 \in \mathbb{Z}\}, \\ \mathcal{L}_\square^\perp &:= \{(\sqrt{\pi} n_1 \quad \sqrt{\pi} n_2)^T \mid n_1, n_2 \in \mathbb{Z}\}. \end{aligned}$$

We refer to the corresponding code  $\text{GKP}(\mathcal{L}_\square)$  using the expression *square lattice GKP code*. It is the most commonly used version of the GKP code and often simply referred to as *the GKP code*. Labeling the basis elements (20) as

$$|\bar{0}\rangle_\square := e_0 \quad \text{and} \quad |\bar{1}\rangle_\square := e_1, \tag{22}$$

one finds that the action (21) is that of the standard Pauli operators (with respect to the computational basis). Thus one writes

$$\begin{aligned} \bar{X}_\square &:= D((\sqrt{\pi} \quad 0)^T), \\ \bar{Z}_\square &:= D((0 \quad \sqrt{\pi})^T). \end{aligned}$$

We emphasize that (22) is simply a (commonly used) convention. Indeed, as argued below, it is essential for our purposes to choose a slightly different mapping from GKP basis states to (logical) qubit states.

It is clear from the construction above that the basis elements  $|\bar{0}\rangle_\square, |\bar{1}\rangle_\square$  are not proper quantum mechanical states in the usual sense: They constitute an infinite uniform superposition of position (equivalently: momentum) eigenstates and are therefore not normalizable. To obtain a physically meaningful description, one has to view the GKP states as a limit of approximate GKP states, in which the sum is replaced by one with weights according to a Gaussian envelope, and the delta distributions forming the position eigenstates are replaced by squeezed Gaussian states.

**E. Asymmetric GKP codes**

Applying a Gaussian unitary  $U_{S^{-1}}$  to a GKP code  $\text{GKP}(\mathcal{L})$  results in a GKP code  $\text{GKP}(\mathcal{L}') = U_{S^{-1}}\text{GKP}(\mathcal{L})$ , where the lattice  $\mathcal{L}' = S\mathcal{L}$  is obtained by applying the associated symplectic transformation  $S \in \text{Sp}(2, \mathbb{R})$  to  $\mathcal{L}$  [cf. (10)]. Note that (12) is invariant under symplectic transformations of the lattice, hence if  $\xi_1, \xi_2$  span  $\mathcal{L}$ , then  $S\xi_1, S\xi_2$  are lattice basis vectors of  $\mathcal{L}'$  satisfying (12).

This shows that the set of symplectically integral lattices  $\mathcal{L}$  [and thus GKP codes  $\text{GKP}(\mathcal{L})$ ] can be partitioned into equivalence classes of lattices that can be transformed into each other by symplectic Gaussian unitaries. Let us briefly argue that there is in fact only a single equivalence class: Every symplectic integral lattice  $\mathcal{L}$  can be obtained by applying a symplectic transformation  $S \in \text{Sp}(2, \mathbb{R})$  to the square lattice  $\mathcal{L}_\square$ . Indeed, suppose that  $\mathcal{L}$  is spanned by  $\xi_1, \xi_2 \in \mathbb{R}^2$  satisfying the integrality condition

$$\xi_1^T J \xi_2 = 4\pi, \tag{23}$$

and let

$$\begin{aligned} \xi_1^\square &:= (2\sqrt{\pi} \quad 0)^T, \\ \xi_2^\square &:= (0 \quad 2\sqrt{\pi})^T, \end{aligned}$$

be vectors generating  $\mathcal{L}_\square$ . Then it is easy to check [using the antisymmetry of the symplectic form and (23)]

that the matrix  $S := \frac{1}{2\sqrt{\pi}}(\xi_1|\xi_2)$  having normalized versions of  $\xi_1, \xi_2$  in its columns is symplectic and maps  $\mathcal{L}_\square$  to  $\mathcal{L}$ . We note that a generalization to  $N > 2$  of this argument can be obtained using a symplectic Gram-Schmidt procedure. Starting from the square lattice  $\mathcal{L}_\square$ , we can therefore obtain various deformed (asymmetric) GKP codes.

### 1. Rectangular lattice GKP codes

Of primary interest is the following example, the rectangular lattice GKP code. Let  $r \geq 1$  and consider the symplectic matrix

$$S_r := \begin{pmatrix} \sqrt{r} & 0 \\ 0 & 1/\sqrt{r} \end{pmatrix} \quad (24)$$

associated with a one-mode squeezing unitary. The lattice generated by the corresponding symplectically transformed vectors  $(S_r \xi_1^\square, S_r \xi_2^\square)$  and its dual lattice are rectangular:

$$\begin{aligned} \mathcal{L}_r &= \{(2\sqrt{\pi r} n_1 \quad 2\sqrt{\pi/r} n_2)^T \mid n_1, n_2 \in \mathbb{Z}\}, \\ \mathcal{L}_r^\perp &= \{(\sqrt{\pi r} n_1 \quad \sqrt{\pi/r} n_2)^T \mid n_1, n_2 \in \mathbb{Z}\}, \end{aligned} \quad (25)$$

and the resulting code GKP( $\mathcal{L}_r$ ) is called a *rectangular lattice GKP code*, see Fig. 2. Here the parameter  $r$  is the *ratio* between the generating vectors of the lattice. We denote the logical Pauli operators by

$$\begin{aligned} \bar{X}_r &:= D((\sqrt{\pi r} \quad 0)^T), \\ \bar{Z}_r &:= D((0 \quad \sqrt{\pi/r})^T). \end{aligned}$$

The square lattice GKP code is a rectangular lattice GKP code with ratio  $r = 1$ . For a discussion of the effect of a parameter  $r > 1$  compared to  $r = 1$  in terms of error correction and state preparation see Sec. III F 2.

We note that for a square or rectangular lattice  $\mathcal{L}$  and its symplectically dual lattice  $\mathcal{L}^\perp$ , the Voronoi cells  $\mathcal{V} \subset \mathbb{R}^2$  and  $\mathcal{V}^\perp \subset \mathbb{R}^2$ , respectively, are given by

$$\begin{aligned} \mathcal{V} &:= \{\lambda_1 \xi_1 + \lambda_2 \xi_2 \mid \lambda_1, \lambda_2 \in [-\frac{1}{2}, \frac{1}{2}]\}, \\ \mathcal{V}^\perp &:= \{\lambda_1 \xi_1^\perp + \lambda_2 \xi_2^\perp \mid \lambda_1, \lambda_2 \in [-\frac{1}{2}, \frac{1}{2}]\}, \end{aligned} \quad (26)$$

where we assume that the lattice  $\mathcal{L}$  is spanned by  $\xi_1, \xi_2$  and  $\mathcal{L}^\perp$  is spanned by  $\xi_1^\perp, \xi_2^\perp$ .

### 2. Hexagonal lattice GKP codes

We mainly focus on rectangular lattice GKP codes for concreteness and ease of illustration. Note, however, that one-mode squeezing is not the only symplectic transformation applicable to the square lattice. In particular, the transformed vectors need not be orthogonal. An example is the *hexagonal lattice GKP code* GKP( $S_\square \mathcal{L}_\square$ ) which is associated with the symplectic transformation

$$S_\square = \left(\frac{2}{\sqrt{3}}\right)^{1/2} \begin{pmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{pmatrix}.$$

The lattice  $S_\square \mathcal{L}_\square$  has an angle of  $2\pi/3$  between the two lattice basis vectors – this is the interior angle of a hexagon.

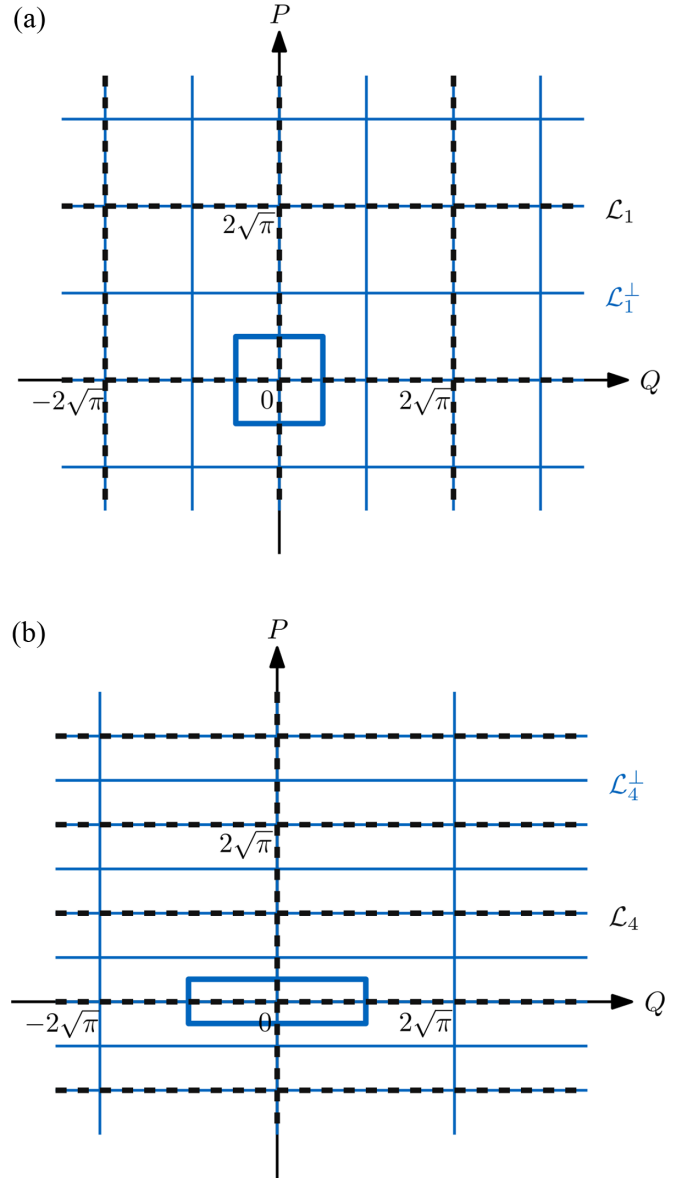


FIG. 2. The lattice  $\mathcal{L}_r$  (dashed black lines) and the dual lattice  $\mathcal{L}_r^\perp$  (solid blue lines) defining the GKP codes for two different ratio parameters  $r = 1$  (square lattice) and  $r = 4$ . The Voronoi cell of the dual lattice  $\mathcal{L}_r^\perp$  is marked with blue edges. (a)  $r = 1$ , (b)  $r = 4$ .

Since the product of symplectic matrices is a symplectic matrix, one gets an *asymmetric hexagonal lattice GKP code* by applying  $S_\square S_r$  to the square lattice code [29].

### F. Error recovery for the GKP code

Consider a logical qubit encoded in GKP( $\mathcal{L}$ ), i.e., described by a density operator  $\rho$  supported on GKP( $\mathcal{L}$ ). Assume that this state undergoes noise given by a completely positive and trace preserving (CPTP) map  $\mathcal{N} : \mathcal{B}(L^2(\mathbb{R})) \rightarrow \mathcal{B}(L^2(\mathbb{R}))$  on the set  $\mathcal{B}(L^2(\mathbb{R}))$  of bounded operators on the Hilbert space  $L^2(\mathbb{R})$ . Specifically, we are interested in random displacement channels  $\mathcal{N}_{f_z}$  [cf. (11)]. Error correction starting from the corrupted state  $\mathcal{N}_{f_z}(\rho)$  proceeds by

(i) Syndrome measurement: This measures the (eigenvalues of the) commuting stabilizer generators  $D(\xi_1), D(\xi_2)$  (with



the generating vectors  $\xi_1, \xi_2$  of the lattice  $\mathcal{L}$ . By (8) and (13), this measurement amounts to a measurement of the quadrature operators  $(Q, P)$  modulo the dual lattice  $\mathcal{L}^\perp$ . The measurement yields a *syndrome*  $s = (q \ p)^T \in \mathcal{V}^\perp$  belonging to the Voronoi cell of the dual lattice  $\mathcal{L}^\perp$ .

For example, in the square lattice case,  $Q$  and  $P$  are measured modulo  $\sqrt{\pi}$ , with outcomes  $q, p \in (-\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2})$ . This can be realized using a logical CNOT gate (realized by a beamsplitter) between the encoded GKP-qubit and an additional ancilla GKP-qubit, and subsequent homodyne measurements of  $Q$  and  $P$  of the ancilla mode (cf. Ref. [6]).

(ii) Application of a unitary correction operation  $C(s)$  depending on the syndrome  $s$ : The correction operation can be chosen to be a displacement, i.e., it is of the form  $C(s) = D(c(s))$  for a function  $c : \mathcal{V}^\perp \rightarrow \mathbb{R}^2$  determined by the chosen recovery procedure (see below).

Sequences (i) and (ii) of operations define a recovery CPTP map  $\mathcal{R}^c : \mathcal{B}(L^2(\mathbb{R})) \rightarrow \mathcal{B}(L^2(\mathbb{R}))$  in terms of the function  $c : \mathcal{V}^\perp \rightarrow \mathbb{R}^2$ . We call this function the *recovery procedure* in the following.

To analyze the effect of the recovery map  $\mathcal{R}^c$ , first consider a single (unitary) displacement error  $D(v)$  for some  $v \in \mathbb{R}^2$  applied to a density operator  $\rho$  with support on  $\text{GKP}(\mathcal{L})$ . In the recovery scheme described above, the stabilizer measurement of the corrupted state  $D(v)\rho D(v)^\dagger$  yields the syndrome

$$s(v) = v \pmod{\mathcal{L}^\perp},$$

with certainty. Combined with the subsequent correction operator  $D(c(s(v)))$ , this results in an overall (effective) displacement given by

$$D(c(s(v)))D(v) \propto D(c(s(v)) + v) \tag{27}$$

where we omit the irrelevant global phase. This overall operation acting on the initial state  $\rho$  therefore leaves the state invariant (i.e., recovery succeeds) if

(a) the operation (27) is logical, i.e., maps the code space to itself, that is,

$$c(s(v)) + v \in \mathcal{L}^\perp, \tag{28}$$

and

(b) the operation (27) has trivial action on the code space, that is, the displacement vector belongs to the trivial coset:

$$c(s(v)) + v \in \mathcal{L} = [\mathbf{0}]. \tag{29}$$

Here we have used the characterization of the effect of displacements on the code space discussed in Sec. III C. We note that condition (28) is typically guaranteed for all  $v \in \mathbb{R}^2$  for reasonable choices of the recovery procedure  $c : \mathcal{V}^\perp \rightarrow \mathbb{R}^2$  (as the ones discussed below). We call such recovery procedures *valid*. The following analysis therefore focuses on (29).

For valid recovery procedures, condition (29) guarantees that there is no residual logical error. More generally, the residual logical operator depends on the coset the vector  $c(s(v)) + v$  belongs to and can be read off from the following

table:

$c(s(v)) + v \in$	Logical Pauli operator applied
$[\mathbf{0}]$	$\bar{I}$
$[\xi_2^\perp]$	$\bar{X}$
$[\xi_2^\perp - \xi_1^\perp]$	$\bar{Y}$
$[-\xi_1^\perp]$	$\bar{Z}$

(30)

(Here we write  $\bar{I}$  for the case of a trivial operator leaving the code space invariant.)

Returning to our error model (11), an error  $D(v)$  occurs with probability  $f_Z(v)d^2v$ . Table (30) allows us to conclude that the combined CPTP map  $\mathcal{R}^c \circ \mathcal{N}_f$  obtained by applying the recovery operation after the noise—when restricted to states supported on  $\text{GKP}(\mathcal{L})$ —is given by

$$\bar{\mathcal{N}}_{p_{\bar{I}}, p_{\bar{X}}, p_{\bar{Y}}, p_{\bar{Z}}}(\rho) = p_{\bar{I}}\rho + p_{\bar{X}}\bar{X}\rho\bar{X}^\dagger + p_{\bar{Y}}\bar{Y}\rho\bar{Y}^\dagger + p_{\bar{Z}}\bar{Z}\rho\bar{Z}^\dagger, \tag{31}$$

where [cf. (30)]

$$p_{\bar{P}} = \Pr_v[c(s(v)) + v \in [\xi_{\bar{P}}^\perp]] = \int_{v: c(s(v))+v \in [\xi_{\bar{P}}^\perp]} f_Z(v)d^2v, \tag{32}$$

with  $\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$  and

$$(\xi_{\bar{I}}^\perp, \xi_{\bar{X}}^\perp, \xi_{\bar{Y}}^\perp, \xi_{\bar{Z}}^\perp) := (\mathbf{0}, \xi_2^\perp, \xi_2^\perp - \xi_1^\perp, -\xi_1^\perp). \tag{33}$$

In particular, the success probability of a valid recovery procedure  $c : \mathcal{V}^\perp \rightarrow \mathbb{R}^2$  is thus given by  $P_{\text{success}}^c = p_{\bar{I}}$ .

We note that the effective logical error channel (31) is the result of averaging over displacement errors and syndrome measurement outcomes. The latter average is incorporated in the definition of the recovery map  $\mathcal{R}^c$ . In fact, a more detailed description of this process treats the sequence (i) and (ii) as an instrument (rather than a CPTP map): this generates both syndrome information  $s \in \mathcal{V}^\perp$  and a corresponding state  $\rho(s)$  (obtained by applying the correction to the post-measurement state). Such a description is needed if the syndrome information  $s$  is used, e.g., in concatenated coding to update Bayesian priors. The distribution over syndromes  $s$  is

$$P_s(s_0) = \Pr_v[s(v) = s_0] \quad \text{for } s_0 \in \mathcal{V}^\perp. \tag{34}$$

Conditioned on the syndrome being  $s_0 \in \mathcal{V}^\perp$ , the conditional probability  $p_{\bar{P}}^{s_0}$  of finding an error  $\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$  is

$$p_{\bar{P}}^{s_0} = \Pr_v[c(s(v)) + v \in [\xi_{\bar{P}}^\perp] | s(v) = s_0] = \frac{\Pr_v[s(v) = s_0 \text{ and } c(s(v)) + v \in [\xi_{\bar{P}}^\perp]]}{P_s(s_0)}, \tag{35}$$

by Bayes' rule, i.e.,

$$p_{\bar{P}} = \int_{\mathcal{V}^\perp} P_s(s_0) p_{\bar{P}}^{s_0} d^2s_0 = \int_{\mathcal{V}^\perp} \Pr_v[s(v) = s_0 \text{ and } c(s(v)) + v \in [\xi_{\bar{P}}^\perp]] d^2s_0. \tag{36}$$

In particular, the state after applying the correction operation is  $\rho(s_0) = \bar{\mathcal{N}}_{p_{\bar{I}}^{s_0}, p_{\bar{X}}^{s_0}, p_{\bar{Y}}^{s_0}, p_{\bar{Z}}^{s_0}}(\rho)$ . In other words, for a fixed syndrome measurement outcome  $s_0$ , the residual error model is a

(logical) Pauli-noise channel with syndrome-dependent error probabilities  $(p_{\bar{I}}^{s_0}, p_{\bar{X}}^{s_0}, p_{\bar{Y}}^{s_0}, p_{\bar{Z}}^{s_0})$ .

### 1. Nearest lattice point recovery procedure

Let us now consider a specific simple choice of recovery procedure  $c : \mathcal{V}^\perp \rightarrow \mathbb{R}^2$ , namely,

$$c(s) = -s \quad \text{for } s \in \mathcal{V}^\perp.$$

We call this the nearest lattice point recovery procedure: it attempts to correct a given shift error by mapping to the nearest lattice point in  $\mathcal{L}^\perp$ . While it is not equivalent to maximum likelihood decoding for general error distributions  $f_Z$ , it is expected to work well for distributions concentrated around the origin (i.e., predominantly small displacement errors). Because of its simplicity, it is commonly used in the context of GKP codes. Figure 3 illustrates this recovery procedure, as well as the phase space regions of displacements associated with errors that are successfully recovered from, respectively lead to a logical error.

To analyze its behavior, first observe that the recovery procedure  $c$  is valid, i.e., satisfies (28) for all  $v \in \mathbb{R}^2$ : indeed, we have

$$\begin{aligned} c(s(v)) + v &= -s(v) + v \\ &= -v \pmod{\mathcal{L}^\perp} + v \in \mathcal{L}^\perp \end{aligned}$$

by definition. Thus we can use the expression (32) for the residual (i.e., postcorrection) logical error (coset) probabilities without syndrome information [cf. (33) and Appendix A1],

$$p_{\bar{P}} = \sum_{\xi \in \mathcal{L}} \int_{\mathcal{V}^\perp} f_Z(v + \xi + \xi_{\bar{P}}^\perp) d^2v, \quad (37)$$

for  $\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$ . Similarly, using the syndrome information  $s_0$ , Eq. (35) yields

$$p_{\bar{P}}^{s_0} = \frac{\sum_{\xi \in \mathcal{L}} f_Z(s_0 + \xi + \xi_{\bar{P}}^\perp)}{\sum_{\xi^\perp \in \mathcal{L}^\perp} f_Z(s_0 + \xi^\perp)}. \quad (38)$$

### 2. Biased logical noise from isotropic physical noise in $\text{GKP}(\mathcal{L}_r)$

As an example, consider the rectangular lattice GKP code  $\text{GKP}(\mathcal{L}_r)$  with the physical mode subject to the isotropic Gaussian displacement channel  $\mathcal{N}_{f_{\sigma^2}}$  [cf. (2)]. Inserting the associated probability density  $f_{\sigma^2}$  into (37) and (38) yields (see Appendix A2)

$$\begin{aligned} p_{\bar{I}} &= (1 - q_{\bar{X}})(1 - q_{\bar{Z}}), \\ p_{\bar{X}} &= q_{\bar{X}}(1 - q_{\bar{Z}}), \\ p_{\bar{Y}} &= q_{\bar{X}}q_{\bar{Z}}, \\ p_{\bar{Z}} &= (1 - q_{\bar{X}})q_{\bar{Z}}, \end{aligned} \quad (39)$$

where

$$q_{\bar{P}} := \int_{\mathcal{V}^\perp} n_{\bar{P}}(s_0) q_{\bar{P}}^{s_0} d^2s_0 \quad (40)$$

for  $\bar{P} \in \{\bar{X}, \bar{Z}\}$  with

$$\begin{aligned} q_{\bar{X}}^{s_0} &:= 1 - n_{\bar{X}}(s_0)^{-1} e(4r, x), \\ q_{\bar{Z}}^{s_0} &:= 1 - n_{\bar{Z}}(s_0)^{-1} e(4/r, z), \end{aligned} \quad (41)$$

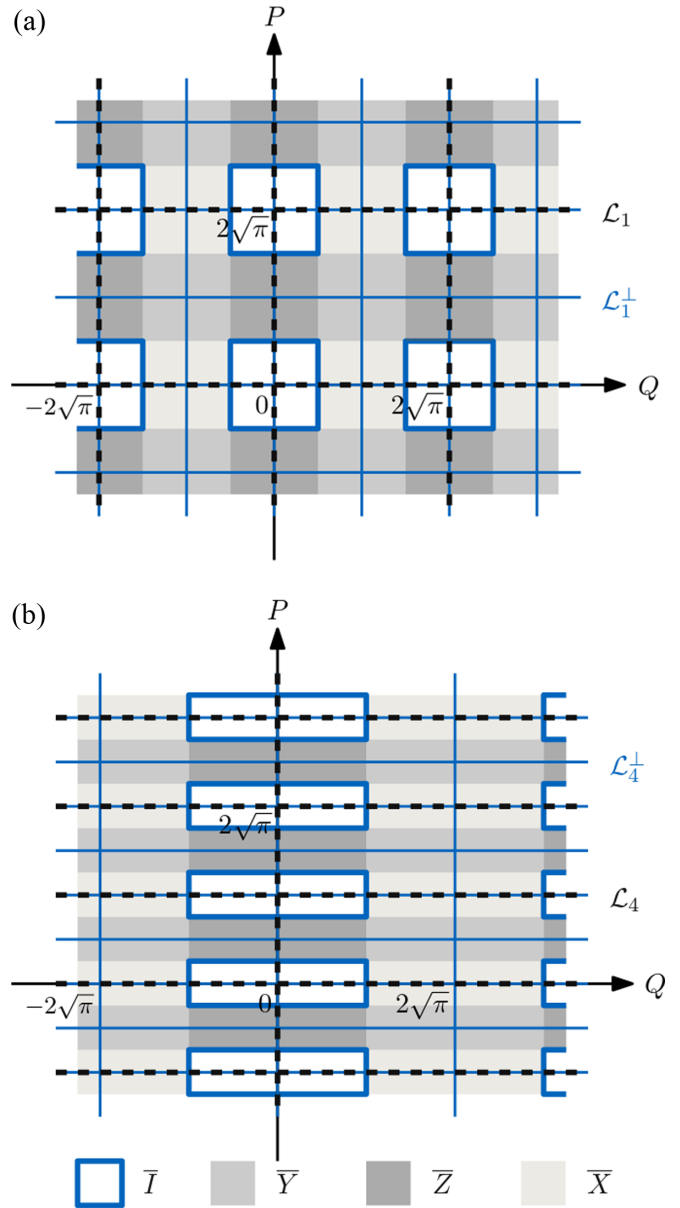


FIG. 3. Lattices for  $\text{GKP}(\mathcal{L}_r)$  for the ratio parameters  $r = 1, 4$ . In the phase space, the lattice  $\mathcal{L}_r$  (dashed black lines) and its dual  $\mathcal{L}_r^\perp$  (solid blue lines) are depicted, respectively. For the square lattice ( $r = 1$ ) GKP code undergoing a displacement  $D(v)$ ,  $v \in \mathbb{R}^2$ , measuring the stabilizer generators  $e^{i2\sqrt{\pi}Q}$  and  $e^{i2\sqrt{\pi}P}$  corresponds to measuring the values  $\eta = v \pmod{\mathcal{L}^\perp}$ , with  $\eta = (\eta_1 \ \eta_2)^T$ ,  $\eta_1, \eta_2 \in [-\frac{\sqrt{\pi}}{2}, \frac{\sqrt{\pi}}{2}]$ . Applying the correction  $D(\eta)$  corresponds to a mapping to the nearest lattice point in  $\mathcal{L}^\perp$ . The correction is successful if the actual error displacement  $v = (v_1 \ v_2)^T$  satisfies  $v_1 = 2\sqrt{\pi}n_1 + \eta_1$ ,  $v_2 = 2\sqrt{\pi}n_2 + \eta_2$ , with  $n_1, n_2 \in \mathbb{Z}$ . In other words, the correction is successful if the error vector  $v$  lies in the areas marked by the blue edges. This correspond to shifts of the dual lattice Voronoi cell by a lattice vector  $\xi \in \mathcal{L}$ . If the error vector  $v$  lies in one of the grey shaded areas, the correction results in a logical error  $\bar{X}, \bar{Y}, \bar{Z}$  [cf. table in (30)]. (a)  $r = 1$ , (b)  $r = 4$ .

and

$$\begin{aligned} n_{\bar{X}}(s_0) &:= e(r, x), \\ n_{\bar{Z}}(s_0) &:= e(1/r, z). \end{aligned} \quad (42)$$

Here we write  $s_0 = (x \ z)^T$  for the syndrome and make use of the function

$$e(u, w) := \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{n \in \mathbb{Z}} e^{-\frac{(\sqrt{\pi}un+w)^2}{2\sigma^2}} \quad (43)$$

for  $u, w \in \mathbb{R}$ ,  $u \geq 0$ . More explicitly, (40) can be written as

$$1 - q_{\bar{X}} = \frac{1}{2} \sum_{n \in \mathbb{Z}} \operatorname{erf} \left( \sqrt{\frac{2\pi r}{\sigma^2}} \left( n + \frac{1}{4} \right) \right) - \operatorname{erf} \left( \sqrt{\frac{2\pi r}{\sigma^2}} \left( n - \frac{1}{4} \right) \right), \quad (44)$$

$$1 - q_{\bar{Z}} = \frac{1}{2} \sum_{n \in \mathbb{Z}} \operatorname{erf} \left( \sqrt{\frac{2\pi}{r\sigma^2}} \left( n + \frac{1}{4} \right) \right) - \operatorname{erf} \left( \sqrt{\frac{2\pi}{r\sigma^2}} \left( n - \frac{1}{4} \right) \right), \quad (45)$$

where  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\tau^2} d\tau$ . Note that (44) is a monotonically increasing function in  $r$ , whereas (45) is monotonically decreasing in  $r$ .

Similarly, we obtain the following expressions for the conditional logical error probabilities given syndrome  $s_0$ :

$$\begin{aligned} p_I^{s_0} &= (1 - q_{\bar{X}}^{s_0})(1 - q_{\bar{Z}}^{s_0}), \\ p_{\bar{X}}^{s_0} &= q_{\bar{X}}^{s_0}(1 - q_{\bar{Z}}^{s_0}), \\ p_{\bar{Y}}^{s_0} &= q_{\bar{X}}^{s_0}q_{\bar{Z}}^{s_0}, \\ p_{\bar{Z}}^{s_0} &= (1 - q_{\bar{X}}^{s_0})q_{\bar{Z}}^{s_0}. \end{aligned} \quad (46)$$

As discussed, the rectangular lattice GKP code  $\text{GKP}(\mathcal{L}_r)$  with squeezing parameter  $r > 1$  is the result of applying the one-mode squeezing operation with parameter  $r > 1$  to the square lattice GKP code. Thus, using identity (10) and a variable substitution, we get  $\mathcal{N}_{f_{r,2}}(U_{S_r^{-1}} \rho U_{S_r^{-1}}^\dagger) = U_{S_r^{-1}} \mathcal{N}_{f_2}(\rho) U_{S_r^{-1}}^\dagger$  where  $\tilde{Z} \sim \mathbf{N}(\mathbf{0}, \Sigma_r)$ , with the diagonal matrix  $\Sigma_r$  defined in (3). That is, the noise is effectively transformed to one where the two quadratures are displaced independently with variances  $(\tilde{\sigma}_Q^2, \tilde{\sigma}_P^2) := (\sigma^2/r, r\sigma^2)$  [cf. (3)]. Note that  $\tilde{\sigma}_Q^2 < \tilde{\sigma}_P^2$  for  $r > 1$ . This results in biased noise on the level of GKP qubits as can be seen in (39): For  $r > 1$ , we have  $p_{\bar{X}} < p_{\bar{Z}}$ .

Since we use the code  $\text{GKP}(\mathcal{L}_r)$  extensively below, let us briefly discuss the degree of squeezing necessary to produce associated code states from standard square lattice GKP code states, i.e., the parameter  $r$  in (24). As common in quantum optics, we quantify squeezing in terms of the *squeezing factor*  $s$ , which is given in units of dB and defined as

$$s(\sigma_Q) := -10 \log_{10} \left( \frac{\sigma_Q^2}{\sigma_0^2} \right),$$

where  $\sigma_0^2 = 1/2$  is the variance in the  $Q$  quadrature in the vacuum state. As just derived, the effective variance in the  $Q$  quadrature of a rectangular ( $r > 1$ ) lattice code is  $\tilde{\sigma}_Q^2 = \sigma^2/r$ , where  $\sigma^2$  is the corresponding variance of a square ( $r = 1$ ) lattice code. We therefore have  $s(\tilde{\sigma}_Q) = s(\sigma^2) + 10 \log_{10}(r)$ , i.e., the introduction of asymmetry ( $r > 1$ ) yields an increase of the squeezing factor by an additive term  $10 \log_{10}(r)$ .

We note that for more general Gaussian encodings corresponding to some Gaussian unitary  $U_S$ , the resulting modified density  $f_{\tilde{Z}}$  over displacements may not be a product distribution with respect to the quadratures. This generally results in correlations between the  $\bar{X}$  and  $\bar{Z}$  errors. One such example is the asymmetric hexagonal lattice GKP code (cf. Sec. III E 2). Expressions for the associated logical error probabilities under the noise channel (2) are given in Appendix A3.

#### IV. THE (MODIFIED) SURFACE-GKP CODE

In this section, we introduce our main idea, namely, the effective generation of biased noise in modified surface-GKP codes by the introduction of asymmetry. The modified surface-GKP codes we study are obtained by concatenating an asymmetric GKP code at the base (inner) level with a qubit (surface) code at the top (outer) level. They benefit from the fact that surface codes are resilient to certain forms of biased noise.

In Sec. IV A we briefly review prior uses of biased noise in the discrete- and continuous-variable (CV) settings. In Sec. IV B, we give the details of how the concatenation has to be done in order to yield an improvement over standard surface-GKP codes. Then, in Sec. IV C, we explain how to use the BSV decoder to decode asymmetric surface-GKP codes with and without GKP side information.

##### A. Prior work on biased noise in codes

For certain physical systems, physically relevant processes naturally lead to biased noise. There is a long history of considering such scenarios in quantum error correction: It was shown that biased noise is typically less detrimental than more general noise if suitable encodings are used. For example, early work [30] on qubit codes considered the extreme case of pure dephasing noise. Fault-tolerance schemes for asymmetric (i.e., predominantly dephasing) qubit noise models were introduced and analyzed in Refs. [31–33], giving improved estimates for error rates and error thresholds for biased noise.

More recently, and directly relevant to our work, a significant degree of resilience to biased noise of surface codes was observed numerically by Tuckett *et al.* [34] when using the approximate maximum likelihood decoder of Bravyi *et al.* [16]. That is, recovery from Pauli noise,

$$\mathcal{N}_{(p_I, p_X, p_Y, p_Z)}(\rho) = p_I \rho + p_X X \rho X^\dagger + p_Y Y \rho Y^\dagger + p_Z Z \rho Z^\dagger, \quad (47)$$

(i.i.d. on each qubit) is successful (in the limit of large code sizes) even for values of  $p_Y$  close to  $1/2$ , in the case where  $p_X = p_Z$  and the bias

$$\eta := p_Y / (p_X + p_Z) \quad (48)$$

is sufficiently large. This is in sharp contrast to the case of independent  $X$  and  $Z$  noise, where

$$\begin{aligned} p_X &= q_X (1 - q_Z), \\ p_Y &= q_X q_Z, \\ p_Z &= (1 - q_X) q_Z, \end{aligned}$$

and where the tolerable noise thresholds for  $q_X$  and  $q_Z$  are smaller than  $\approx 11\%$  [35].

Subsequent work by Tuckett *et al.* [18] extended these findings significantly in several directions: on the one hand, new numerical results for (rotated), nonsquare, i.e.,  $r \times s$ -surface codes were established. On the other hand, additional analytical arguments were given to support the numerical evidence: the minimum weight of a  $Y$ -type logical operator was computed (as a function of  $r, s$ ), and in the limiting case of pure  $Y$ -noise, a threshold of 50% error probability was established by analyzing a corresponding decoder.

Note, however, that not all qubit codes show such improvements: In fact, it was shown numerically in Ref. [18] that the threshold of color codes [36] actually decreases with stronger bias. It should also be emphasized that improvements can only be obtained by suitably aligning the code's stabilizers with the asymmetry in the noise as discussed below.

While these results apply to qubit codes, more recent work has shown that biasedness of noise can naturally emerge when using CV systems to encode logical qubits. Specifically, Guillaud and Mirrahimi [37] and Puri *et al.* [38] argued that bosonic cat codes exhibit logical-level noise biased towards dephasing (under certain bosonic noise processes affecting the physical modes), where the bias is tunable by adjusting the “cat size” parametrizing the code. They show how to exploit this by concatenating the cat code with the (qubit) repetition code, which has some resilience to biased noise. In addition, methods for achieving bias-preserving logical gates are proposed and analyzed in detail.

### B. Exploiting engineered bias in surface–GKP codes

To tap into the potential of the surface code to correct biased noise, we require the following modifications of what are typically considered to be surface–GKP codes:

(i) First, we use a Gaussian unitary  $U_S$  applied to each mode to turn isotropic Gaussian noise into biased noise at the GKP-qubit level, that is, we work with modified GKP code states  $U_S|\bar{0}\rangle_{\square}$  and  $U_S|\bar{1}\rangle_{\square}$ , where  $|\bar{0}\rangle_{\square}$  and  $|\bar{1}\rangle_{\square}$  are the standard square lattice GKP basis states.

(ii) Second, we change the way the GKP code is concatenated with the surface code. More precisely, we map Pauli- $Y$  (instead of Pauli- $Z$ ) eigenstates of a (GKP) qubit to the modified GKP basis states  $U_S|\bar{0}\rangle_{\square}$  and  $U_S|\bar{1}\rangle_{\square}$  according to

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \mapsto U_S|\bar{0}\rangle_{\square}, \\ |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \mapsto U_S|\bar{1}\rangle_{\square}, \end{aligned} \quad (49)$$

and linearly extend this to a definition of a (modified) isometric GKP-qubit encoding  $\mathbb{C}^2 \rightarrow L^2(\mathbb{R})$ .

We note that the definition (49) ensures that qubit operators are identified with logical GKP operators according to

$$\begin{aligned} Y &\leftrightarrow U_S \bar{Z}_{\square} U_S^{\dagger}, \\ Z &\leftrightarrow U_S \bar{X}_{\square} U_S^{\dagger}. \end{aligned} \quad (50)$$

In the case of a rectangular lattice GKP code, the correspondence (50) leads to effective (qubit level) noise channels with

independent  $Y$  and  $Z$  noise, i.e.,

$$\begin{aligned} p_X &= q_Z q_Y, \\ p_Y &= (1 - q_Z) q_Y, \\ p_Z &= q_Z (1 - q_Y), \end{aligned} \quad (51)$$

in the resulting qubit noise channel (47), where

$$(q_Y, q_Z) = (q_{\bar{Z}}, q_{\bar{X}}), \quad (52)$$

with  $q_{\bar{X}}, q_{\bar{Z}}$  given in (40). In particular,  $p_Z$  is biased compared to  $p_Y$  for  $r$  large enough (i.e.,  $p_Z \ll p_Y$ ), since  $q_Z$  is monotonically decreasing in  $r$  whereas  $q_Y$  is monotonically increasing in  $r$  by the relation (52) and Eqs. (44), (45). We note that the triple  $(p_X, p_Y, p_Z)$  thus is a two-parameter family parameterized by  $(q_Y, q_Z)$ , and the set of such triples differs from the one studied in [18,34], which consists of tuples of the form  $(p_X, p_Y, p_Z) = (p_X, p_Y, p_X)$ , cf. (48). In particular, this means that the threshold estimates obtained in Refs. [18,34] cannot directly be lifted to our setting. Rather, we need to separately study the threshold behavior of biased noise of the form (51).

We find (see Sec. V) that the surface code is well-equipped against displacement noise specified by (51) if  $p_Z \ll p_Y$ . Specifically, we find threshold standard deviations  $\sigma$  of the noise corresponding to – in a standard (non-modified) surface–GKP code—Pauli- $X$  (equivalently: Pauli- $Z$ ) error probabilities strictly above the  $\approx 11\%$  of the plain surface code, similar to the results in Refs. [18,34].

### C. Decoding modified surface–GKP codes with the BSV decoder

In the surface–GKP code, each of the  $n$  qubits of the surface code is encoded in a GKP-qubit (i.e., in a single mode). As with any concatenated code, the surface–GKP code provides natural families of decoders obtained by combining a decoder for the GKP code with a decoder for the surface code. Such a decoding procedure first decodes the GKP code (the inner code), and subsequently the surface code (the outer code). Here we primarily use the nearest lattice point recovery procedure discussed in Sec. III F 1 to decode the GKP code, and the BSV decoder discussed in Sec. IIB 2 for the surface code.

Recall that the GKP decoder produces syndrome information  $s^{j,k}$  for every GKP-encoded qubit  $(j, k)$  (see Sec. III F). Here we use  $j$  and  $k$  to index the row and column in the square lattice geometry [see Fig. 1(a)]. This syndrome information can either be ignored or passed on to the outer decoder. The difference is illustrated in Fig. 4. In the former case, i.e., decoding without side information, the prior distribution  $\pi_{j,k} = (p_{\bar{X}}^{j,k}, p_{\bar{Y}}^{j,k}, p_{\bar{Z}}^{j,k})$  over logical Pauli-error probabilities (seen by the outer decoder) are identical for each qubit  $(j, k)$ , i.e.,  $\pi_{j,k} = \pi$ , and equal to the averaged GKP-qubit probabilities  $\pi = (p_{\bar{X}}, p_{\bar{Y}}, p_{\bar{Z}})$  given in expression (32) [that is, (39) for rectangular lattice GKP-qubits subject to isotropic Gaussian noise]. Thus the BSV decoder is provided with the i.i.d. distribution  $\pi^n$ , that is, the tensors in Fig. 1(c) are qubit independent.

Alternatively, one may use the GKP syndrome information in the surface code decoding step. In this case, the prior probabilities  $\pi_{j,k}$  passed to the surface code decoder are po-



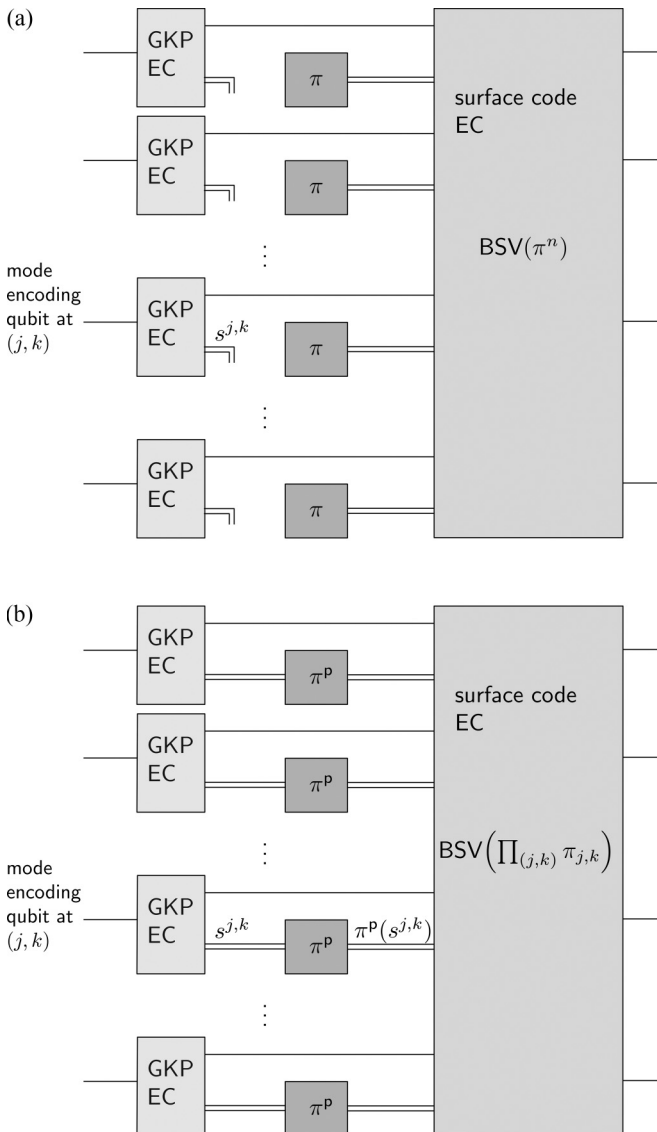


FIG. 4. Two ways of decoding the surface-GKP code. In both cases the GKP error correction circuit produces syndrome information  $s^{j,k}$ . When decoding without side information [Fig. 4(a)], this information is ignored and the same prior distribution  $\pi$  over logical Pauli errors  $\{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$  is used for every site. The surface-code BSV decoder (which also involves syndrome measurement but whose syndromes are not shown here) is run with the associated product (i.i.d.) distribution  $\pi^n$ . Figure 4(b) shows how to exploit the GKP side information: For each qubit  $(j, k)$  a prior distribution  $\pi_{j,k} = \pi^{P(s^{j,k})}$  over logical Pauli errors is computed from the GKP syndrome  $s^{j,k}$ , and this is then used in the BSV decoder as the prior distribution. (a) Decoding without GKP side information. (b) Decoding with GKP side information.

tentially different for each GKP-encoded qubit  $(j, k)$ . They are computed from the associated GKP syndrome  $s^{j,k}$  according to expression (35), which involves the distribution over the bosonic displacement errors, and we write  $\pi_{j,k} = \pi^{P(s^{j,k})}$ . The BSV decoder can then be run with the (nonidentical) product distribution  $\prod_{(j,k)} \pi_{j,k}$ , where each qubit experiences independent noise with a distribution depending on the outcome of the associated GKP measurement.

We note that decoding with and without side information has been studied previously in Refs. [5,11,12], for the surface-GKP code and the toric-GKP code respectively, based on square lattices. These authors used a minimum weight matching decoder for the toric (surface) code. Syndrome information from the GKP measurements was translated into different edge weights (where the three papers use different heuristic expressions) to be passed to the minimum weight matching decoder. In contrast, we consider nonsquare lattice GKP codes, and use the BSV decoder: here the way syndrome information is passed is completely determined. Indeed, for bond dimension  $\chi \rightarrow \infty$ , this results in maximum likelihood decoding based on the available syndrome information.

## V. THRESHOLD ESTIMATES FOR THE MODIFIED SURFACE-GKP CODE

In this section, we present numerical results obtained by applying the decoding procedure discussed in Sec. IV C to study the effect of asymmetry in surface-GKP codes. We primarily study if asymmetry increases thresholds, even when no GKP side information is used, in rectangular lattice GKP codes. Additionally, we provide numerical data for rectangular and asymmetric hexagonal lattice GKP codes using GKP side information.

### A. Simulation method

The results are based on Monte Carlo simulation of the error correction processes illustrated in Fig. 4. A pseudocode of the corresponding routine in the case where the GKP side information is ignored is presented in Fig. 9 in Appendix B, whereas Fig. 10 in Appendix B gives pseudocode in the case where the side information is used in the BSV decoder. Given a standard deviation  $\sigma$ , both routines simulate the process of applying a displacement error distributed according to  $\mathbf{N}(\mathbf{0}, \sigma^2 I_2)$  independently of each GKP qubit of a distance- $d$  surface code, subsequent GKP syndrome extraction and error correction, and surface-code error correction with or without this syndrome information. The output of both routines is the residual logical Pauli error that this process yields on the logical surface-GKP qubit. Thus the routines permit us to empirically estimate the averaged logical error channel. More coarsely, we use these procedures to study the logical error probability  $P_{\text{err}} = P_{\text{err}}(\sigma, d, r)$  as a function of the standard deviation  $\sigma$  of the noise [cf. (2)], the code distance  $d$ , and the asymmetry ratio  $r$  defining the GKP code lattice.

For a fixed ratio  $r$ , we are interested in the *threshold*, which is the critical noise level, i.e., the critical value of  $\sigma$ , below which the error probability  $P_{\text{err}}$  can be made arbitrarily small by choosing a sufficiently large code distance  $d$ . Following standard reasoning, an estimate  $\sigma_c$  for this quantity can be obtained by studying the intersection points of a family  $\{\sigma \mapsto P_{\text{err}}(\sigma, d, r)\}_{d \in D}$  of curves parameterized by a (finite) set of distances  $D \subset \mathbb{N}$ . More precisely, we use the critical exponent method of Ref. [39]: define a correlation length  $\xi = (\sigma - \sigma_c)^{-\mu}$  for some critical exponent  $\mu$ , and assume that for  $d \gg \xi$ , the failure probability only depends on the dimensionless ratio  $d/\xi$ , i.e., it is a function of the variable  $x := (\sigma - \sigma_c)d^{1/\mu}$ . Our thresholds are obtained by numeri-

cally fitting the data to a power series up to quadratic terms in  $x$ , i.e., using the same fitting formula as in Ref. [34].

For our simulations, we choose physical parameters as follows. We consider code sizes  $d$  satisfying  $9 \leq d \leq 21$ , Gaussian displacement noise standard deviations  $\sigma$  in the interval  $[0.4, 0.7]$ , and asymmetry ratios  $r \in [1, 4]$ . The parameters for the simulation are as follows: The number of Monte Carlo samples for an estimate of  $P_{\text{err}}$  for fixed physical parameters  $(\sigma, d, r)$  is chosen between 10 000 and 50 000, depending on the distance  $d$ . For the BSV decoder, we use bond dimensions  $\chi$  in the interval  $[48, 100]$ . We provide a more detailed description and justification of these choices in Appendix D.

## B. Numerical results for surface-GKP( $\mathcal{L}_r$ )

We primarily focus on decoding surface-GKP codes without making use of GKP side information (other scenarios will be discussed below). We present (selected) simulation results for this scenario: in Figs. 5(a)–5(c), the curves  $\{\sigma \mapsto P_{\text{err}}(\sigma, d, r)\}_{d \in \{9, 13, 17, 21\}}$  are shown for asymmetry ratios  $r \in \{1, 2, 3\}$ , respectively. These curves are used to extract a threshold value  $\sigma_c$  as described above. Figure 6 shows the curves  $\{\sigma \mapsto P_{\text{err}}(\sigma, d, r)\}_{r \in \{1, 1.5, 2, 2.5, 3, 3.5, 4\}}$  indicating how different asymmetry ratios  $r$  change the logical error probability for a fixed distance  $d \in \{9, 13, 17, 21\}$ .

### 1. Standard (symmetric) surface-GKP code

The standard (symmetric, i.e.,  $r = 1$ ) GKP code corresponding to the square lattice provides our reference point for comparison when studying the effect of asymmetry. For this reason, we work with a relatively high bond dimension of  $\chi = 100$  for the BSV decoder with the goal of obtaining accurate estimates of the maximum likelihood decoding probability (see Appendix D for a detailed discussion). We obtain a threshold value  $\sigma_c \approx 0.540$ , see Fig. 5(a).

We note that this value is comparable to the threshold estimate for  $\sigma_c$  between 0.54 and 0.55 obtained in Ref. [5] for independent  $X$  and  $Z$  noise using the minimum weight matching decoder. We emphasize, however, that while these results deal with the same physical error model, there is no *a priori* reason the obtained thresholds should coincide. This is because the mapping to qubit-level noise used here [cf. Sec. IV B, in particular modification (ii)] differs from that used in Ref. [5].

Note also that the data presented in Ref. [16] suggests that the difference between using the minimum weight matching decoder and the BSV decoder may play a minor role for the symmetric case  $r = 1$ . In fact, for pure Pauli- $X$  (equivalently, Pauli- $Z$ ) errors even the discrepancy between the threshold values obtained by using the minimum weight matching decoder [39] and the maximum likelihood decoder (for which a threshold can be derived from the numerical estimates for the random-bond Ising model obtained in Ref. [40]) is known to be less than 0.7%, see Ref. [16]. This is within the accuracy regime of the results found in Ref. [5].

### 2. Asymmetric surface-GKP codes

Consider now the asymmetric case, i.e.,  $r > 1$ . A subset of corresponding simulation results is shown in Figs. 5(b) and

5(c). These curves are used to extract threshold estimates. We note that the choice of bond dimension (discussed in Appendix D) is less crucial here as our goal is mainly to demonstrate the advantage of asymmetry. Indeed, for any chosen value of  $\chi$ , the associated curve in Fig. 5 represents the actual error probability of some (albeit approximate and hence possibly nonoptimal) decoder.

It is also instructive and experimentally relevant in the near future to consider the error probability for a fixed code distance  $d$ . Corresponding curves for a collection of asymmetry ratios are shown in Fig. 6. Figures 6(a)–6(d) give the curves  $\{\sigma \mapsto P_{\text{err}}(\sigma, d, r)\}_{r \in \{1, 2, 2.5, 3, 3.5, 4\}}$  for different ratios in separate graphs for the same distance  $d \in \{9, 13, 17, 21\}$ .

We observe a qualitative difference between the regime of small distances and small asymmetry ratios, and the regime of large distances and large asymmetry ratios. More precisely, we may ask where the error probability decreases monotonically with increasing asymmetry ratio  $r$ , i.e.,  $P_{\text{err}}(\sigma, d, r') < P_{\text{err}}(\sigma, d, r)$  for  $r' > r$  (for all  $\sigma$ ). We find that this property holds for all  $r, r' \leq 2.0$  for all distances  $d$ . For higher distances, e.g.,  $d = 17$  and  $d = 21$ , this property extends up to all asymmetry ratios  $r, r' \leq 2.5$ . In other words, the monotonicity in  $r$  appears to be a function of the distance  $d$ : For example, for  $d = 9$  and ratios  $r' = 3.5$  and  $r = 1$ , there are values  $\sigma$  such that  $P_{\text{err}}(\sigma, 9, 3.5) > P_{\text{err}}(\sigma, 9, 1)$ , whereas for  $d = 21$  we find that  $P_{\text{err}}(\sigma, 21, 3.5) < P_{\text{err}}(\sigma, 21, 1)$  for all  $\sigma$ .

It is conceivable that this nonmonotonic behavior of  $P_{\text{err}}$  is merely an artefact of the finite size, i.e., the limited code distances  $d$  that can be explored by simulation. For large distances  $d$ , the logical error probability may in fact decrease monotonically for increasing asymmetry ratios  $r$ . Our current numerical data does not permit us to draw a definite conclusion in this regard.

### 3. Thresholds

Figure 7 summarizes our main numerical findings for the effect of asymmetry on the threshold. It gives the observed threshold values  $\sigma_c$  for all considered asymmetry ratios  $r$ . For comparison, the vertical axis on the right hand side gives the value of  $q_{\bar{X}}$  (cf. (44)) for  $\sigma = \sigma_c$  and  $r = 1$ , i.e., the Pauli- $X$  (equivalently, Pauli- $Z$ ) error probability of the individual qubits of the surface code. This probability corresponds to the scenario where—under the same displacement error noise model—the standard GKP code and GKP error correction without syndrome information is used to encode each qubit.

The numerical results gathered in Fig. 7 show that every asymmetry ratio  $r > 1$  considered here yields an improved threshold compared to the symmetric ( $r = 1$ ) case: for example, the threshold (tolerable error probability) improves to  $\sigma_c = 0.581$  (0.127) for  $r = 3$  from  $\sigma_c = 0.540$  (0.101) for the symmetric case  $r = 1$ . That is, even a moderate, experimentally achievable amount (cf. Appendix D3) of squeezing results in more noise resilience.

Theoretically it is also interesting to examine the limit of large asymmetry  $r$ . Figure 7 shows that the estimated threshold value  $\sigma_c$  obtained by simulation is nonmonotonic, exhibiting a maximum around  $r = 3$ . As explained above, this numerical finding may be due to finite-size effects, i.e., the limited system sizes considered in our simulations. We

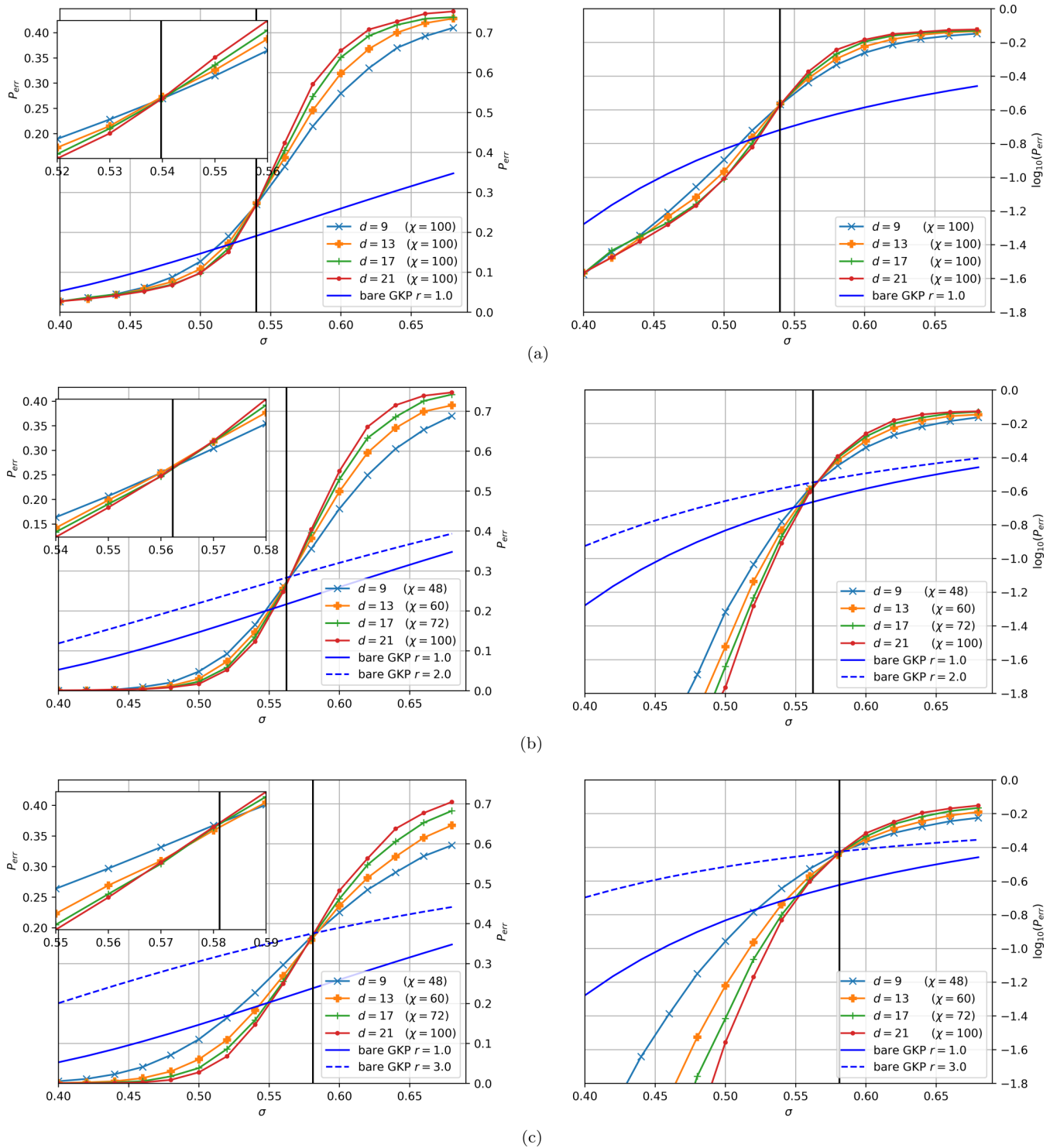


FIG. 5. The error probability  $P_{\text{err}}(\sigma, d, r)$  (left) for decoding without GKP side information for different asymmetry ratios  $r \in \{1.0, 2.0, 3.0\}$ . Insets give higher-resolution data around the observed threshold estimate  $\sigma_c$  for the critical noise variance. The latter is indicated by a vertical line. The right-hand side shows a log-plot of the error probability  $P_{\text{err}}(\sigma, d, r)$ . For comparison, the error probability of the bare GKP code, i.e., without concatenation with the surface code, is shown in blue with the solid line corresponding to ratio  $r = 1.0$  and the dashed lines corresponding to the respective ratio of the figure. (a) Ratio  $r = 1.0$ , threshold  $\sigma_c \approx 0.540$ . (b) Ratio  $r = 2.0$ , threshold  $\sigma_c \approx 0.562$ . (c) Ratio  $r = 3.0$ , threshold  $\sigma_c \approx 0.581$ .

cannot conclusively deduce that the actual threshold increases monotonically with  $r$  (for large distances  $d$ ). We note that the limiting effective surface-code qubit-level noise for  $r \rightarrow \infty$  is pure  $Y$  noise with 50% probability. For the latter, an analytical

threshold result is known [18]. However, this result does not allow us to draw conclusions about the threshold for any fixed asymmetry ratio  $r$  because of the different order of limits with respect to  $d$  and  $r$ , respectively.

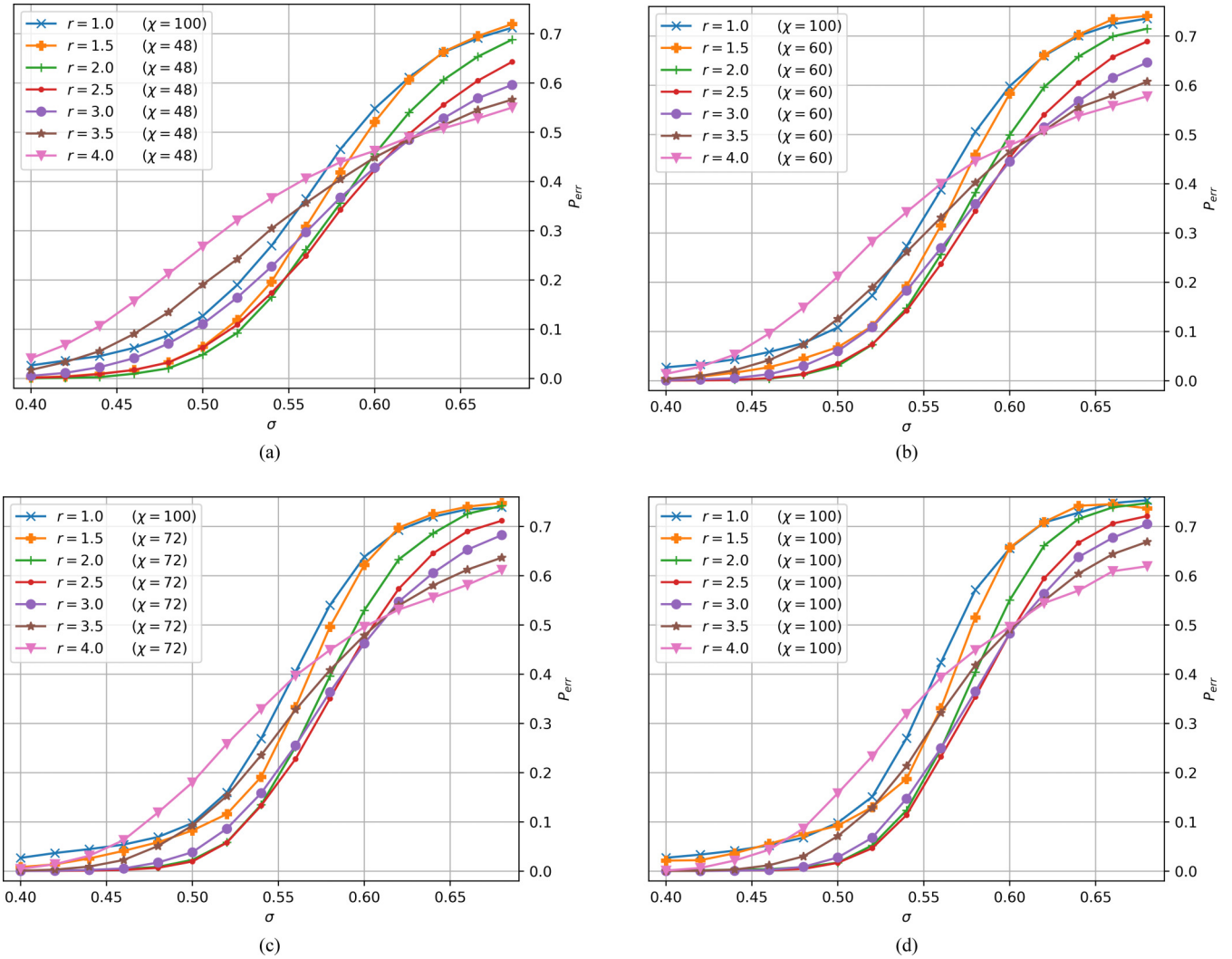


FIG. 6. The dependence of the error probability  $P_{\text{err}}(\sigma, d, r)$  on the noise variance  $\sigma$  for decoding without GKP side information. In each figure, the curves  $\sigma \mapsto P_{\text{err}}(\sigma, d, r)$  for asymmetry ratios  $r$  between 1.0 and 4.0 are given. The individual figures treat code distances  $d \in \{9, 13, 17, 21\}$ . (a) Distance  $d = 9$ . (b) Distance  $d = 13$ . (c) Distance  $d = 17$ . (d) Distance  $d = 21$ .

Note that the effective qubit error distribution resulting from the use of asymmetric GKP codes depends on both the asymmetry ratio  $r$  (a parameter which may in principle be chosen arbitrarily subject to experimental capabilities) and the physical error strength  $\sigma$ . In particular, the parameter  $r$  determines the relative weight of  $X$ -,  $Y$ -, and  $Z$ -Pauli errors similar to the bias parameter  $\eta$  in the work of Refs. [18,34]. We emphasize however that our effective noise model does not have the form of the error distributions considered in those works. In particular, their bias parameter  $\eta$  is not in one-to-one correspondence with our asymmetry ratio  $r$ .

### C. Numerical results for further scenarios

Up to this point, we have not considered all possible optimizations over different choices of codes and decoders in our simulations. For example, we restricted our attention to rectangular GKP codes and did not attempt to optimize the underlying lattice by varying its axial angle. It can be expected that such an optimization yields additional benefits. Indeed, as already pointed out in the seminal paper [6], using the

hexagonal lattice  $\mathcal{L}_{\square}$  instead of  $\mathcal{L}_{\square}$  results in improved error correcting properties because of the difference in volume of the associated Voronoi cells.

A further significant improvement should result from modifying the decoder: So far, we did not use the GKP syndrome information in the surface code decoding step. For the standard (symmetric) surface-GKP code, it was shown in Refs. [5,11,12] that using this side information improves threshold estimates.

To examine if such modifications lead to improvements as expected, we numerically estimate logical error probabilities for surface-GKP codes based on the rectangular lattice  $\mathcal{L}_r$  with  $r = 2$ , decoded without and with side information [Figs. 8(a) and 8(b), respectively]. Analogously to the results of Refs. [5,11] for the symmetric case where a threshold of  $\sigma_c \approx 0.61$  was observed, we find that the use of side information significantly increases the threshold in the asymmetric case also.

As a paradigmatic test case, we additionally consider the use of the asymmetric hexagonal lattice surface GKP code



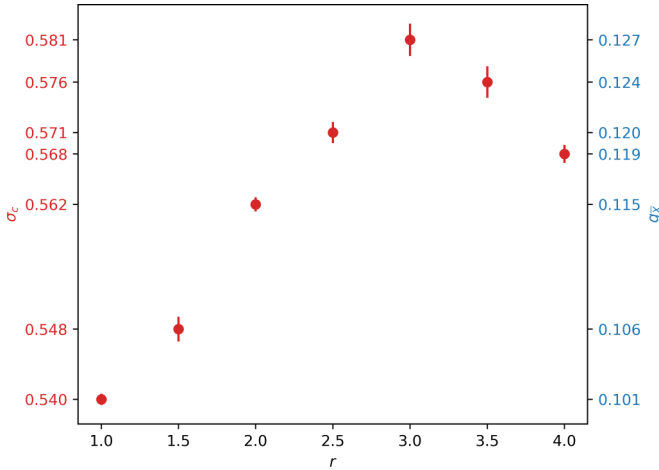


FIG. 7. Empirically computed thresholds for different asymmetry ratios  $r$ . The error bars depict the standard deviation of the fitted threshold value  $\sigma_c$  (left hand side axis) and range from 0.0006 (for  $r = 1$ ) to 0.0019 (for  $r = 3$ ). The values on the right-hand side axis correspond to  $q_{\bar{X}}$  [Eq. (46)] for the parameters  $\sigma = \sigma_c$  and  $r = 1$ . These values are given for comparison with existing estimates for (qubit) surface codes: For example, it is known [35] that the logical  $X$ -error probability  $q_{\bar{X}}$  for minimum matching decoding for independent  $X$  and  $Z$  noise is around 10.1%. We recover this value for ratio  $r = 1$ . The overall logical error probability is given by  $p = q_{\bar{X}}^2 + 2q_{\bar{X}}(1 - q_{\bar{X}})$ .

GKP( $S_{\square}, S_r, \mathcal{L}_{\square}$ ) introduced in Sec. III E 2 for asymmetry ratio  $r = 2$ , see Fig. 8(c). Again, this simulation data is for a decoder using side information. Here the observed threshold is comparable to that obtained for the square lattice surface GKP code.

These numerical results indicate that such additional modifications aimed to improve fault-tolerance properties do not negatively interfere with the use of asymmetry. In particular, combining these strategies and additionally optimizing over lattices (rather than simply the hexagonal and rectangular ones) has the potential to yield further increases of the threshold estimates.

## VI. CONCLUSIONS

Our results show that no-go theorems for Gaussian CV-into-CV encodings (see Sec. I) no longer apply when considering concatenated codes: Our numerical results for the (modified) surface-GKP code indicate that suitably chosen Gaussian CV-into-CV encodings (such as the one presented in Sec. IV B) effectively lead to a deformation of the effective error distribution for the GKP-qubits which known decoders for the surface code can benefit from. We expect that similar strategies may improve other CV codes constructed by concatenation. Due to the minimal additional experimental requirements, this kind of modification of bosonic fault-tolerance schemes may be particularly attractive once the basic components of an experimental setup are in place.

Our paper provides a proof of principle and shows that artificially engineered asymmetry not only benefits surface-GKP codes, but is also compatible with other improvements to the

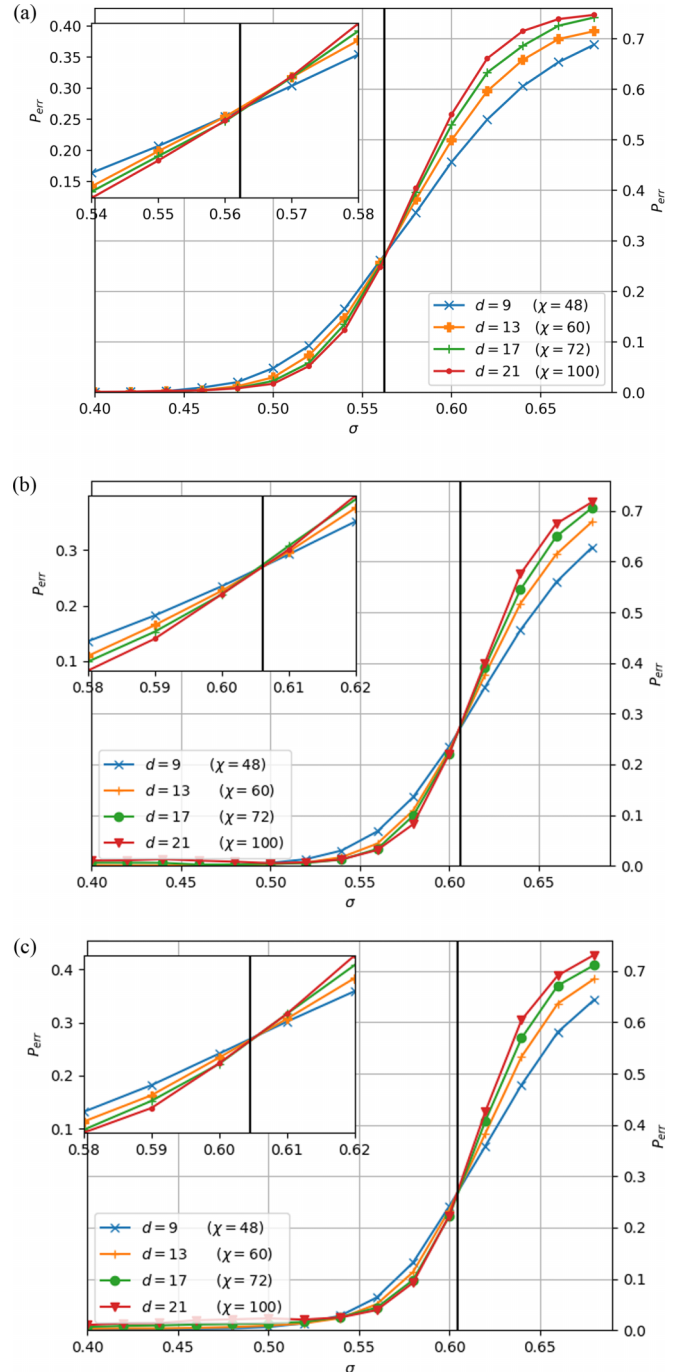


FIG. 8. Other improvements for ratio  $r = 2.0$ : Here the plots show the error probability  $P_{\text{err}}(\sigma, d, r)$  as a function of  $\sigma$  for decoding without side information [Fig. 8(a)] and with side information [Fig. 8(b)] for the rectangular lattice GKP code. Figure 8(c) shows the error probability for the asymmetric hexagonal lattice GKP code decoded with side information. Observed threshold estimates  $\sigma_c$  are indicated with vertical lines. (a) Rectangular lattice GKP without side information, threshold  $\sigma_c \approx 0.562$ . (b) Rectangular lattice GKP with side information, threshold  $\sigma_c \approx 0.6062$  (with standard deviation 0.0007). (c) Asymmetric hexagonal lattice GKP with side information, threshold  $\sigma_c \approx 0.6045$  (with standard deviation 0.0009).

GKP encoding: This includes, e.g., the use of nonrectangular lattices having larger Voronoi cells, and the use of GKP side information at the surface-code decoding stage.

Future work may seek to establish monotonicity of the logical error probability as a function of the degree of asymmetry. Furthermore, the effect of nonideal (finitely squeezed) GKP states, as well as noise in the syndrome information, and eventually full circuit-level noise should be examined. Finally, methods for fault-tolerant computation with asymmetric surface–GKP codes, e.g., using bias-preserving gates should be further developed.

### ACKNOWLEDGMENTS

R.K. acknowledges support by the Technical University of Munich Institute of Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under Grant Agreement No. 291763 and by the DFG cluster of excellence 2111 (Munich Center for Quantum Science and Technology). R.K. and L.H. acknowledge support by the German Federal Ministry of Education through the program Photonics Research Germany, Contract No. 13N14776 (QCDA-QuantERA). M.H. is supported by the International Max Planck Research School for Quantum Science and Technology at the Max-Planck-Institut für Quantenoptik. The authors gratefully acknowledge computing resources of the Leibniz-Rechenzentrum.

### APPENDIX A: COMPUTATION OF COSET PROBABILITIES

In this Appendix, we derive the logical error (coset) probabilities for nearest lattice point decoding (NLPD, cf. Sec. III F 1) given a classical isotropic Gaussian displacement noise channel  $\mathcal{N}_{f_{\sigma^2}}$  [cf. (2)]; the latter channel is characterized by its density

$$f_{\sigma^2}(v) = \frac{1}{2\pi\sigma^2} e^{-\frac{\|v\|^2}{2\sigma^2}}, \quad (\text{A1})$$

where  $\|v\|^2 = v^T v = v_1^2 + v_2^2$ . We start by proving the formulas (37), (38) for the coset probabilities without and with GKP side information for NLPD for a general random displacement error channel (11) in Appendix A 1. Subsequently, the corresponding explicit expressions (39), (46) for a rectangular lattice and the error channel corresponding to (A1) are derived in Sec. A 2. Finally, in Appendix A 3, we give an explicit expression for the coset probabilities with GKP side information in the case of an asymmetric hexagonal lattice and density (A1). The latter expression is not presented in the main text but added here for completeness as it is used in the simulations for the asymmetric hexagonal case.

#### 1. Nearest lattice point decoding (NLPD)

Recall that for NLPD (cf. Sec. III F 1) we have  $c(s(v)) = -s(v) \in \mathcal{V}^\perp$  and thus

$$\{v \in \mathbb{R}^2 \mid c(s(v)) + v \in [\xi_{\bar{P}}^\perp]\}$$

$$\begin{aligned} &= \{v \in \mathbb{R}^2 \mid v - s(v) \in \xi_{\bar{P}}^\perp + \mathcal{L}\} \\ &= \{v \in \mathbb{R}^2 \mid v = \omega + \xi + \xi_{\bar{P}}^\perp, \xi \in \mathcal{L} \text{ and } \omega \in \mathcal{V}^\perp\}. \end{aligned}$$

Hence by (32) and variable substitution, we have

$$\begin{aligned} p_{\bar{P}} &= \Pr_v [c(s(v)) + v \in [\xi_{\bar{P}}^\perp]] \\ &= \sum_{\xi \in \mathcal{L}} \int_{\mathcal{V}^\perp} f_Z(v + \xi + \xi_{\bar{P}}^\perp) \quad d^2v, \end{aligned}$$

proving (37). Furthermore, since  $s(v) = v \pmod{\mathcal{L}^\perp}$ , we have

$$\{v \in \mathbb{R}^2 \mid s(v) = s_0\} = \{v \in \mathbb{R}^2 \mid v = \xi^\perp + s_0, \xi^\perp \in \mathcal{L}^\perp\}.$$

By (34) and variable substitution, it thus follows that

$$P_s(s_0) = \Pr_v [s(v) = s_0] = \sum_{\xi^\perp \in \mathcal{L}^\perp} f_Z(\xi^\perp + s_0). \quad (\text{A2})$$

This together with

$$\begin{aligned} &\{v \in \mathbb{R}^2 \mid c(s_0) + v \in [\xi_{\bar{P}}^\perp]\} \\ &= \{v \in \mathbb{R}^2 \mid -s_0 + v \in [\xi_{\bar{P}}^\perp]\} \\ &= \{v \in \mathbb{R}^2 \mid v = s_0 + \xi + \xi_{\bar{P}}^\perp, \xi \in \mathcal{L}\} \end{aligned}$$

and an additional variable substitution implies, by (35), that

$$\begin{aligned} p_{\bar{P}}^{s_0} &= \Pr_v [c(s(v)) + v \in [\xi_{\bar{P}}^\perp] \mid s(v) = s_0] \\ &= P_s(s_0)^{-1} \Pr_v [s(v) = s_0 \text{ and } c(s(v)) + v \in [\xi_{\bar{P}}^\perp]] \\ &= \left( \sum_{\xi^\perp \in \mathcal{L}^\perp} f_Z(s_0 + \xi^\perp) \right)^{-1} \sum_{\xi \in \mathcal{L}} f_Z(s_0 + \xi + \xi_{\bar{P}}^\perp). \quad (\text{A3}) \end{aligned}$$

This yields (38).

#### 2. NLPD with Gaussian noise channel: rectangular lattice

Let now  $\mathcal{L} \equiv \mathcal{L}_r$  and  $f_Z \equiv f_{\sigma^2}$  be given by (A1). Furthermore, let us write  $s_0 = (x \ z)^T$ . Then by the definition of the dual lattice  $\mathcal{L}_r^\perp$  [cf. (25)] we may write Eq. (A2) as

$$\begin{aligned} P_s(s_0) &= \frac{1}{2\pi\sigma^2} \sum_{n_1, n_2 \in \mathbb{Z}} e^{-\frac{(\sqrt{\pi}n_1+x)^2}{2\sigma^2}} e^{-\frac{(\sqrt{\pi}n_2+z)^2}{2\sigma^2}} \\ &= e(r, x) e(1/r, z), \end{aligned}$$

cf. (43). Recall that  $(\xi_{\bar{I}}^\perp, \xi_{\bar{X}}^\perp, \xi_{\bar{Y}}^\perp, \xi_{\bar{Z}}^\perp) := (0, \xi_2^\perp, \xi_1^\perp, -\xi_1^\perp)$ . If we interpret  $\xi_{\bar{P}}^\perp$  as a function of  $\bar{P}$ , we may thus write

$$\xi_{\bar{P}}^\perp = (\sqrt{\pi}r \cdot 1_{\bar{P} \in \{\bar{X}, \bar{Y}\}} \sqrt{\pi/r} \cdot 1_{\bar{P} \in \{\bar{Z}, \bar{Y}\}})^T,$$

where  $1_{\bar{P} \in \{\bar{X}, \bar{Y}\}}, 1_{\bar{P} \in \{\bar{Z}, \bar{Y}\}}$  are indicator functions. Then, by the definition of the lattice  $\mathcal{L}_r$  [cf. (25)],

$$\begin{aligned} \sum_{\xi \in \mathcal{L}_r} f_Z(s_0 + \xi + \xi_{\bar{P}}^\perp) &= \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{n_1 \in \mathbb{Z}} e^{-\frac{(\sqrt{\pi}r(2n_1+1)_{\bar{P} \in \{\bar{X}, \bar{Y}\}}+x)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{n_2 \in \mathbb{Z}} e^{-\frac{(\sqrt{\pi}/r(2n_2+1)_{\bar{P} \in \{\bar{Z}, \bar{Y}\}}+z)^2}{2\sigma^2}} \\ &= e(4r, x + \sqrt{\pi}r \cdot 1_{\bar{P} \in \{\bar{X}, \bar{Y}\}}) e(4/r, z + \sqrt{\pi}/r \cdot 1_{\bar{P} \in \{\bar{Z}, \bar{Y}\}}), \end{aligned}$$

whence the coset probabilities with GKP information (A3) can be written as

$$\begin{aligned} p_{\bar{P}}^{s_0} &= P_s(s_0)^{-1} \sum_{\xi \in \mathcal{L}_r} f_Z(s_0 + \xi + \xi_{\bar{P}}^\perp) \\ &= \frac{e(4r, x + \sqrt{\pi r} \mathbf{1}_{\bar{P} \in \{\bar{X}, \bar{Y}\}})}{e(r, x)} \frac{e(4/r, z + \sqrt{\pi/r} \mathbf{1}_{\bar{P} \in \{\bar{Z}, \bar{Y}\}})}{e(1/r, z)}. \end{aligned}$$

This together with the identities

$$e(r, x) - e(4r, x + \sqrt{\pi r}) = e(4r, x), \quad e(1/r, z) - e(4/r, z + \sqrt{\pi/r}) = e(4/r, z)$$

yields (41), (42), and (46).

By (36), the coset probabilities (39) without GKP information follow from (46) by integration, cf. (40). Finally, (44) is verified as

$$\begin{aligned} (1 - q_{\bar{X}}) &= \int_{\mathcal{V}_r^\perp} n_{\bar{X}}(s_0) (1 - q_{\bar{X}}^{s_0}) d^2 s_0 \\ &= \int_{\{x = \lambda_1 \sqrt{\pi r} \mid \lambda_1 \in [-1/2, 1/2]\}} e(4r, x) dx \\ &= \sum_{n_1 \in \mathbb{Z}} \frac{1}{\sqrt{\pi}} \int_{\sqrt{\frac{2\pi r}{\sigma^2}}(n_1 - 1/4)}^{\sqrt{\frac{2\pi r}{\sigma^2}}(n_1 + 1/4)} e^{-\tau_1^2} d\tau_1 \\ &= \frac{1}{2} \sum_{n \in \mathbb{Z}} \operatorname{erf}\left(\sqrt{\frac{2\pi r}{\sigma^2}}\left(n + \frac{1}{4}\right)\right) - \operatorname{erf}\left(\sqrt{\frac{2\pi r}{\sigma^2}}\left(n - \frac{1}{4}\right)\right), \end{aligned}$$

where we used (in this order) the identity

$$\int_{\mathcal{V}_r^\perp} n_{\bar{P}}(s_0) d^2 s_0 = 1,$$

the definition of the dual Voronoi cell  $\mathcal{V}_r^\perp$  [cf. (26)], variable substitution, and the definition of the error function. The identity (45) is verified analogously.

### 3. NLPD with Gaussian noise channel: hexagonal lattice

Here we derive the coset probabilities with GKP side information for an asymmetric hexagonal lattice  $\mathcal{L}_{\square, r}$  (cf. Sec. III E 2), with  $f_Z \equiv f_{\sigma^2}$  given in (A1). We have

$$\begin{aligned} \mathcal{L}_{\square, r} &= \{c(2\sqrt{\pi r}n_1 + \sqrt{\pi/r}n_2 \quad \sqrt{3\pi/r}n_2)^T \mid n_1, n_2 \in \mathbb{Z}\}, \\ \mathcal{L}_{\square, r}^\perp &= \left\{ \frac{c}{2}(2\sqrt{\pi r}n_1 + \sqrt{\pi/r}n_2 \quad \sqrt{3\pi/r}n_2)^T \mid n_1, n_2 \in \mathbb{Z} \right\}, \end{aligned}$$

and

$$\mathcal{V}_{\square, r}^\perp = \left\{ \frac{c}{2}(2\sqrt{\pi r}\lambda_1 + \sqrt{\pi/r}\lambda_2 \quad \sqrt{3\pi/r}\lambda_2)^T \mid \lambda_1, \lambda_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right] \right\},$$

where  $c := (2/\sqrt{3})^{1/2}$ . For  $u, w_q, w_p \in \mathbb{R}$ ,  $u \geq 0$ , define

$$e_{\square}(u, w_q, w_p) := \frac{1}{2\pi\sigma^2} \sum_{n_1, n_2 \in \mathbb{Z}} e^{-\frac{(c\sqrt{\pi}u_1 + (c/2)\sqrt{\pi/ru_2} + w_q)^2 + ((c/2)\sqrt{3\pi/ru_2} + w_p)^2}{2\sigma^2}}.$$

Then, with  $s_0 = (x \quad z)^T$ , Eq. (A2) can be written as

$$\begin{aligned} P_s(s_0) &= \sum_{\xi^\perp \in \mathcal{L}_{\square, r}^\perp} f_Z(\xi^\perp + s_0) = \sum_{n_1, n_2 \in \mathbb{Z}} \frac{e^{-\frac{((c/2)(2\sqrt{\pi r}n_1 + \sqrt{\pi/r}n_2) + x)^2 + ((c/2)\sqrt{3\pi/r}n_2 + z)^2}{2\sigma^2}}}{2\pi\sigma^2} \\ &= e_{\square}(r, x, z). \end{aligned}$$

Furthermore, writing

$$\xi_{\bar{P}}^\perp = \frac{c}{2}(2\sqrt{\pi r} \cdot \mathbf{1}_{\bar{P} \in \{\bar{X}, \bar{Y}\}} + \sqrt{\pi/r} \cdot \mathbf{1}_{\bar{P} \in \{\bar{Z}, \bar{Y}\}} \quad \sqrt{3\pi/r} \cdot \mathbf{1}_{\bar{P} \in \{\bar{Z}, \bar{Y}\}})^T$$

(cf. Appendix A 2), we get

$$\begin{aligned} \sum_{\xi \in \mathcal{L}_{\square, r}} f_Z(s_0 + \xi + \xi_{\bar{P}}^{\perp}) &= \sum_{n_1, n_2 \in \mathbb{Z}} \frac{e^{-\frac{(2c\sqrt{\pi r}(n_1 + \frac{1}{2}\bar{P}_{\in\{\bar{X}, \bar{Y}\}}) + c\sqrt{\pi r}(n_2 + \frac{1}{2}\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}) + x)^2 + (c\sqrt{3\pi/r}(n_2 + \frac{1}{2}\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}) + z)^2}{2\sigma^2}}}{2\pi\sigma^2} \\ &= e_{\square} \left( 4r, x + c\sqrt{\pi r} 1_{\bar{P}_{\in\{\bar{X}, \bar{Y}\}}} + \frac{c}{2}\sqrt{\pi/r} 1_{\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}}, z + \frac{c}{2}\sqrt{3\pi/r} 1_{\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}} \right), \end{aligned}$$

whence the coset probabilities with GKP information (A3) can be written:

$$\begin{aligned} p_{\bar{P}}^{s_0} &= P_s(s_0)^{-1} \sum_{\xi \in \mathcal{L}_{\square, r}} f_Z(s_0 + \xi + \xi_{\bar{P}}^{\perp}) \\ &= \frac{e_{\square} \left( 4r, x + c\sqrt{\pi r} \cdot 1_{\bar{P}_{\in\{\bar{X}, \bar{Y}\}}} + \frac{c}{2}\sqrt{\pi/r} \cdot 1_{\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}}, z + \frac{c}{2}\sqrt{3\pi/r} \cdot 1_{\bar{P}_{\in\{\bar{Z}, \bar{Y}\}}} \right)}{e_{\square}(r, x, z)}. \end{aligned} \quad (\text{A4})$$

## APPENDIX B: PSEUDOCODE FOR THE MONTE CARLO SIMULATION ALGORITHMS

Here we give pseudocode for the algorithms used to obtain our numerical results in Sec. V. The routine MCWITHOUTSIDEINFO shown in Fig. 9 simulates one Monte Carlo step when decoding without side information [41]. In this description, the errors  $E_{j,k}$  are sampled exactly from the distribution  $\pi$  and the BSV decoder is given the exact probabilities  $\pi$ .

We note that Fig. 9 does not yet constitute a practical algorithm: Because expression (39) for  $\pi$  involves infinite sums, an approximation has to be made in our implementation, i.e., we work with a suitably chosen approximation  $\tilde{\pi}$  to  $\pi$ . We argue in Appendix C that using this approximation  $\tilde{\pi}$  does not impact the validity of our conclusions about the effect of asymmetry.

In Fig. 10, we give pseudocode for the simulation of error correction when using GKP side information. Again, in our implementation, Step 10 in this algorithm is replaced by an approximate computation described and analyzed in Appendix C.

## APPENDIX C: CUTOFF ANALYSIS

For the computation of the distributions  $\pi$  (cf. Step 10 in Fig. 9) and  $\pi_{j,k}$  (cf. Step 10 in Fig. 10), infinite sums are approximated by finite ones by introducing a cutoff  $\kappa$ .

More precisely, the probability distribution  $\pi$  is computed using the expressions (44) and (45) with a cutoff  $\kappa = 10$ , meaning that all summands with indices of absolute value  $\geq 10$  are neglected. Similarly, for the conditional distribution  $\pi_{j,k}$ , a cutoff  $\kappa = 15$  is chosen when evaluating expressions (46) respectively (A4). Let us denote the corresponding approximate distributions by  $\tilde{\pi}$  and  $\tilde{\pi}_{j,k}$ , respectively. Here we discuss the impact of using these approximations instead of the actual distributions  $\pi$ , respectively,  $\pi_{j,k}$  on the accuracy of our simulation.

We first observe that these approximations do not enter in the sampling procedure generating the errors  $E_{j,k}$  when following the description of Figs. 9 or 10. That is, the simulation algorithm produces the correct distribution over errors [42]. This means that only the effect of using  $\tilde{\pi}$  instead of the exact distribution  $\pi$  (respectively,  $\tilde{\pi}_{j,k}$  instead of  $\pi_{j,k}$ ) as input to the BSV decoder has to be considered.

If the BSV decoder receives an approximation of the actual *a priori* probabilities  $\pi$  (or  $\pi_{j,k}$ ), its action is not that of a maximum likelihood decoder even in the limit of infinite bond dimension  $\chi$ . Nevertheless, it still acts as a (hopefully decent) decoder: The fact that the bond dimension is finite and the input probabilities are approximate as a consequence of a finite cutoff  $\kappa$  can only possibly affect the performance of the decoder. In other words, the simulation algorithms shown in Figs. 9 and 10 allow us to compute estimates of logical

---

```

1: function MCWITHOUTSIDEINFO( $\sigma^2, d, r$ )
2: _____
3: Input: Variance  $\sigma^2$ , code distance  $d$ , asymmetry ratio  $r$ 
4: Output: Residual logical Pauli error  $\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$ 
5: _____
6:   for each qubit  $(j, k)$  do
7:     Sample displacement error vector  $\nu_{j,k} \in \mathbb{R}^2$  according to  $N(\mathbf{0}, \sigma^2 I_2)$ .
8:     Compute the (logical) Pauli error  $E_{j,k} \in \{I, X, Y, Z\}$  from  $\nu_{j,k}$  according to (38) and (31) (for  $\text{GKP}(\mathcal{L}_r)$ ).
9:     Set  $E_s := \otimes_{j,k} E_{j,k}$ .
10:    Compute the distribution  $\pi = (p_{\bar{I}}, p_{\bar{X}}, p_{\bar{Y}}, p_{\bar{Z}})$  according to (41) (for  $\text{GKP}(\mathcal{L}_r)$ ).
11:    Use the BSV decoder with input  $\pi^n$  to compute an  $n$ -qubit Pauli correction  $C$ .
12:    Compute  $\bar{P} = \arg \max_{\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}} 1_{E_s C \in \bar{P}S}$  (i.e., decide which coset  $E_s C$  belongs to).

```

---

FIG. 9. Subroutine MCWITHOUTSIDEINFO simulates one instance of the error-recovery process when GKP side information is ignored. It returns the residual logical error.



```

1: function MCWITHSIDEINFO( $\sigma^2, d, r$ )
2:
3: Input: Variance  $\sigma^2$ , code distance  $d$ , asymmetry ratio  $r$ 
4: Output: Residual logical Pauli error  $\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}$ 
5:
6:   for each qubit  $(j, k)$  do
7:     Sample displacement error vector  $\nu_{j,k} \in \mathbb{R}^2$  according to  $\mathbf{N}(\mathbf{0}, \sigma^2 I_2)$ .
8:     Compute the (logical) Pauli error  $E_{j,k} \in \{I, X, Y, Z\}$  from  $\nu_{j,k}$  according to (38) and (31) (for  $\text{GKP}(\mathcal{L}_r)$ 
       respectively  $\text{GKP}(\mathcal{L}_{\mathcal{O},r})$ ).
9:     Compute the syndrome  $s_0 = s(\nu_{j,k}) = (z \ y)^T$  according to (27).
10:    Compute the conditional distribution  $\pi_{j,k} = (p_I^{s_0}, p_X^{s_0}, p_Y^{s_0}, p_Z^{s_0})$  according to (48) (for  $\text{GKP}(\mathcal{L}_r)$ ) respec-
       tively (A4) (for  $\text{GKP}(\mathcal{L}_{\mathcal{O},r})$ ).
11:    Set  $E_s := \otimes_{j,k} E_{j,k}$ .
12:    Use the BSV decoder with input  $\prod_{j,k} \pi_{j,k}$  to compute an  $n$ -qubit Pauli correction  $C$ .
13:    Compute  $\bar{P} = \arg \max_{\bar{P} \in \{\bar{I}, \bar{X}, \bar{Y}, \bar{Z}\}} \mathbb{1}_{E_s C \in \bar{P}S}$  (i.e., decide which coset  $E_s C$  belongs to).
    
```

FIG. 10. Subroutine MCWITHSIDEINFO Monte-Carlo-simulates one instance of the error-recovery process when GKP side information is used. It returns the residual logical error.

error probabilities of *some* decoder (with and without side information).

Running these simulation algorithms hence establishes lower bounds on achievable error thresholds obtained by using nonideal decoders. Since our general goal is to show that asymmetry is beneficial, this means that we do not need to worry about the choice of cutoff  $\kappa$  as long as we observe an improvement over symmetric codes (i.e., asymmetry ratio  $r = 1$ ).

In this comparison, it is of course important to have a reliable estimate of the threshold for the symmetric case  $r = 1$  since this is our reference point. We make three observations concerning this point:

First, from an operational viewpoint, it is natural to optimize the decoding success probability while restricting attention to efficient decoders. We use the BSV decoder with a particularly high bond dimension (see Appendix D 5). This should guarantee that the decoder provides a good approximation to maximum likelihood decoding while still being efficient.

Second, our threshold estimate for the symmetric case is comparable to previously obtained threshold values (cf. Sec. V B) in this well-studied scenario, increasing confidence in their reliability.

Third, we argue here that the approximate probability distributions  $\tilde{\pi}$  are—for the considered parameter regime and cutoff  $\kappa$ —very close to the actual probability distributions  $\pi$ . That is, the input to the BSV decoder only differs by a small amount from the ideal input. While we do not make any claims about the continuity properties of the BSV decoder, this is at least some indication that using these approximate distributions is justified.

For this analysis, define the quantities

$$\alpha_+ := \max_{r,\sigma} \left\{ \sqrt{\frac{2\pi r}{\sigma^2}} \sqrt{\frac{2\pi}{r\sigma^2}} \right\} \quad \text{and}$$

$$\alpha_- := \min_{r,\sigma} \left\{ \sqrt{\frac{2\pi r}{\sigma^2}}, \sqrt{\frac{2\pi}{r\sigma^2}} \right\},$$

where the optimizations are over the parameter values  $(r, \sigma)$  considered in our simulations. The logical error probabili-

ties without GKP side information (39) are computed via the probabilities  $q_{\bar{X}}, q_{\bar{Z}}$ , i.e., the expressions (44) and (45). As the slope of the error function  $\text{erf}(\tau)$  is decreasing with increasing  $|\tau|$ , the contribution of the summands in (44) and (45) decreases with  $n \in \mathbb{Z}$ . For  $n \neq 0$ , the function

$$\text{inc}(n) := \frac{1}{2} \left( \text{erf}\left(\alpha_- n + \frac{\alpha_+}{4}\right) - \text{erf}\left(\alpha_- n - \frac{\alpha_+}{4}\right) \right)$$

is thus an upper bound on the increment caused by the  $n$ th summand in both cases (44) and (45). Note that the error function is an odd function and thus  $\text{inc}(n) = \text{inc}(-n)$ . We can therefore bound the error introduced by a cutoff  $\kappa \geq 1$  by

$$\begin{aligned} & \max_{\bar{P} \in \{\bar{X}, \bar{Z}\}} (1 - q_{\bar{P}}) - (1 - \tilde{q}_{\bar{P}}) \\ & \leq 2 \sum_{n=\kappa}^{\infty} \text{inc}(n) \\ & = \frac{2}{\sqrt{\pi}} \sum_{n=\kappa}^{\infty} \int_{\alpha_- n - \alpha_+/4}^{\alpha_- n + \alpha_+/4} e^{-\tau^2} d\tau \\ & \leq \frac{2}{\sqrt{\pi}} \int_{\alpha_- \kappa - \alpha_+/4}^{\infty} e^{-\tau^2} d\tau \\ & \leq \frac{2}{\sqrt{\pi}} \int_{\alpha_- \kappa - \alpha_+/4}^{\infty} \frac{\tau}{\alpha_- \kappa - \alpha_+/4} e^{-\tau^2} d\tau \\ & = \frac{1}{(\alpha_- \kappa - \alpha_+/4)\sqrt{\pi}} e^{-(\alpha_- \kappa - \alpha_+/4)^2}. \end{aligned} \quad (\text{C1})$$

Here  $\tilde{q}_{\bar{X}}$  and  $\tilde{q}_{\bar{Z}}$  are given by the expressions (44) and (45) with the sum over  $n \in \mathbb{Z}$  replaced by a sum from  $-\kappa + 1$  to  $\kappa - 1$ . On the other hand,  $1 - q_{\bar{X}}$  and  $1 - q_{\bar{Z}}$  as given in (44) and (45) can be bounded from below by their summand  $n = 0$ , respectively. The latter in turn is in both cases greater than or equal to  $\text{erf}(\alpha_-/4)$ . Thus the ratio of the approximation error in the calculation of (44) and (45) and the true value is

$$\begin{aligned} & \max_{\bar{P} \in \{\bar{X}, \bar{Z}\}} \frac{(1 - q_{\bar{P}}) - (1 - \tilde{q}_{\bar{P}})}{1 - q_{\bar{P}}} \\ & \leq (\sqrt{\pi}(\alpha_- \kappa - \alpha_+/4) \text{erf}(\alpha_-/4) e^{(\alpha_- \kappa - \alpha_+/4)^2})^{-1}. \end{aligned} \quad (\text{C2})$$

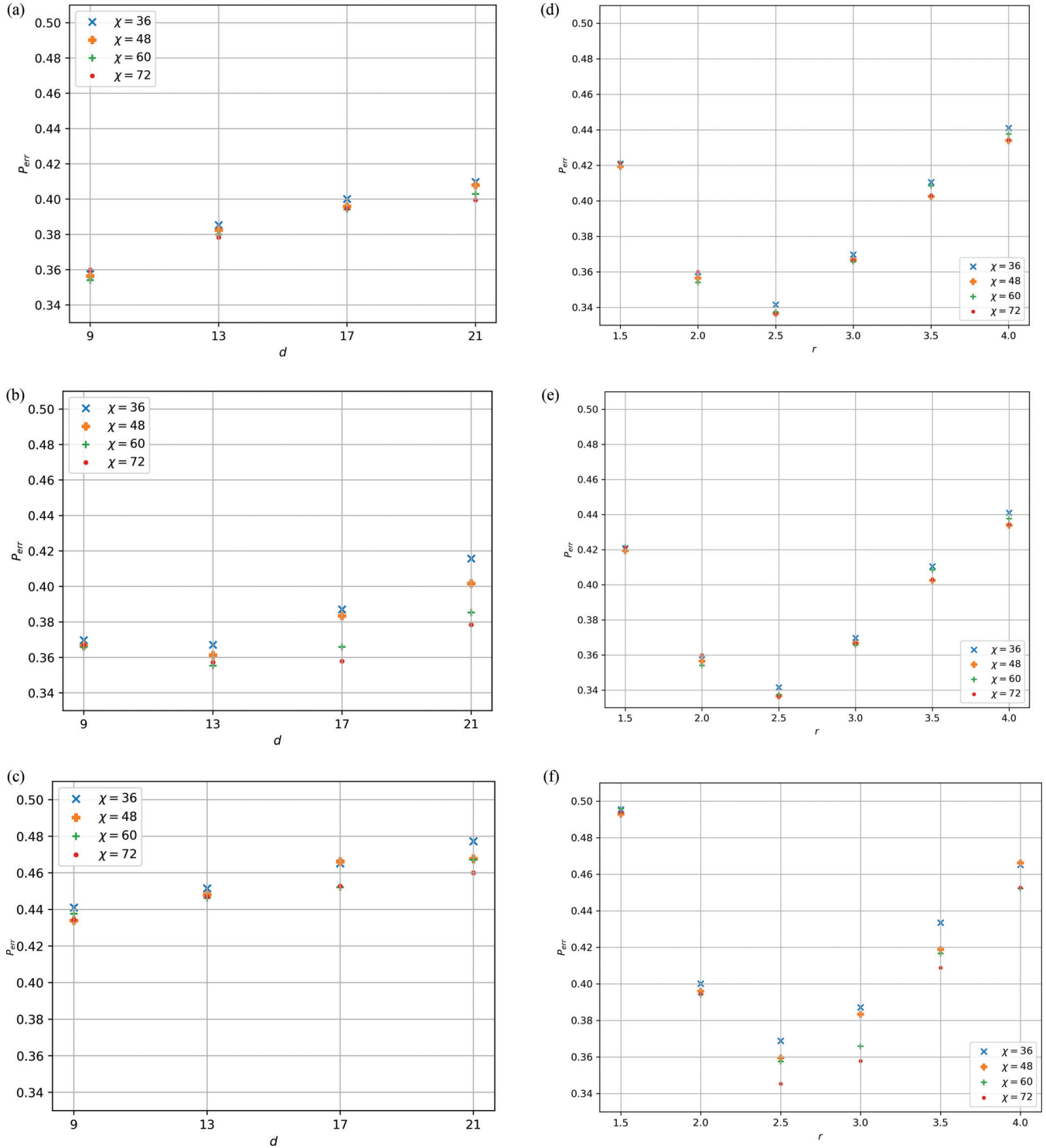


FIG. 11. Test of the necessary bond dimension for the GKP code (without side information): Error probabilities are estimated for different bond dimensions  $\chi$ . Heuristically, a bond dimension  $\chi$  suffices if  $P_{\text{err}}$  does not decrease significantly when the bond dimension is increased further. We observe that the required bond dimension  $\chi$  appears to be nonmonotonic as a function of the asymmetry ratio  $r$ , being maximal in some intermediate regime of  $r$  (left hand side). The required bond dimension increases with increasing distance  $d$  (right-hand side). (a) Ratio  $r = 2.0$ . (b) Ratio  $r = 3.0$ . (c) Ratio  $r = 4.0$ . (d) Distance  $d = 9$ . (e) Distance  $d = 13$ . (f) Distance  $d = 17$ .

For the given parameter values  $r \in [1, 4]$  and  $\sigma \in [4/10, 7/10]$ , the parameters  $\alpha_+$  and  $\alpha_-$  evaluate to  $(\alpha_+, \alpha_-) = 5\sqrt{2\pi}(1, 1/7)$ . Inserting this into (C2) shows that with a cutoff  $\kappa = 10$ , the approximate values  $1 - \tilde{q}_{\bar{X}}, 1 - \tilde{q}_{\bar{Z}}$

are within  $10^{-93}\%$  of the correct values  $1 - q_{\bar{X}}, 1 - q_{\bar{Z}}$ . Since the probabilities of interest are given by degree-2 polynomials in  $(q_{\bar{X}}, q_{\bar{Z}})$ , see Eq. (51), this shows that the introduction of a cutoff leads to a negligible error in the computation of  $\pi$ .

## APPENDIX D: CHOICE OF PARAMETERS

In this Appendix, we discuss the choice of parameters used in our simulation. This includes physical parameters such as the code size  $d$  (Appendix D 1), the noise strength as quantified by the variance  $\sigma$  (Appendix D 2), and the asymmetry ratio  $r$  (Appendix D 3).

We also discuss parameters related to the simulation, including the number of Monte Carlo iterations (Appendix D 4), as well as the necessary bond dimension  $\chi$  in the BSV decoder (Appendix D 5).

### 1. Code sizes

We consider code sizes  $d \in \{9, 13, 17, 21\}$ , which are comparable to what has been used in [18,34] to study the performance of the surface code under biased noise using the BSV decoder.

### 2. Noise strengths (variance)

To identify noise levels of interest, we use previously known results: In Ref. [5], it was shown that (standard) square lattice toric-GKP codes can tolerate isotropic Gaussian displacement noise with standard deviations up to values between  $\sigma \approx 0.54$  and  $\sigma \approx 0.55$  when a certain weighted minimum-matching decoding is used without GKP side information, and around  $\sigma \approx 0.61$  when GKP side information is employed. Based on this, we simulate noise levels given by standard deviations  $\sigma \in [0.4, 0.7]$  in increments of 0.02 (a higher resolution of 0.01 is used in the vicinity of empirical thresholds).

### 3. Asymmetry ratios

As discussed in Sec. III F 2, a lattice asymmetry ratio  $r$  corresponds to an increase of  $10 \log_{10}(r)$  in squeezing with respect to the square lattice GKP code (we consider  $r \geq 1$ ). The relevant quantity to decide whether an amount of squeezing is physically reasonable is the *total* squeezing resulting from the isotropic noise variance  $\sigma^2$  and the ratio  $r$ . The physically reasonable range of  $r$  therefore depends on the value of  $\sigma$  under consideration. Here we consider asymmetry ratios  $r \in [1, 4]$  with increments of 0.5, which translate to a squeezing  $< 11$  dB

for the considered range of variances  $\sigma \in [0.4, 0.7]$ . Note that this is around the amount of squeezing within reach of near term experimental setups, see, e.g. the discussion in Ref. [11]. Ratios up to  $r = 4$  turned out to be sufficient for the numerical demonstration of the beneficial effect of asymmetry.

### 4. Number of Monte Carlo iterations

The number of simulations completed for every tuple  $(\sigma, d, r)$  depends on the distance  $d$ : the distances 9, 13, 17, and 21 are simulated approximately 50 000, 30 000, 30 000, and 10 000 times, respectively; for higher distances the fluctuations decrease and hence the empirical estimates converge faster. For the bond dimension tests (see Appendix D 5) the number of simulations for all distances is approximately 30 000.

### 5. Necessary bond dimension

For given choices of  $\sigma$ ,  $r$ , and  $d$ , the bond dimension  $\chi$  has to be chosen suitably large in order for the BSV decoder to be sufficiently accurate.

To determine appropriate bond dimensions for our model, we conducted test simulations for distances  $d \in \{9, 13, 17, 21\}$  and ratios  $r \in [1.5, 4]$  (in increments of 0.5), for a *fixed* standard deviation  $\sigma = 0.58$  (i.e., in the vicinity of the expected threshold estimate) and for bond dimensions  $\chi \in \{36, 48, 60, 72\}$ . Note that  $\chi = 36$  and  $\chi = 48$  are the bond dimensions used in Refs. [18] and [34], respectively. Examples are depicted in Fig. 11. We find a similar behavior as observed in Ref. [34]: intermediate ratios (biases) require higher bond dimensions. Furthermore, the necessary bond dimension increases with the distance  $d$ . Our tests show that for distances  $d = 9, 13, 17$ , the bond dimensions  $\chi = 48, 60, 72$  are sufficient, respectively, for all considered ratios. For distance  $d = 21$  we choose  $\chi = 100$ , since the tests indicate that  $\chi = 72$  is insufficient. Also, the values for ratio  $r = 1$ , which serve as benchmarks, are simulated with bond dimension  $\chi = 100$  to prevent any bias toward our interpretation. As explained in the main text, for the asymmetric ( $r > 1$ ) cases choosing lower bond dimensions  $\chi$  suffices. This is because any choice of  $\chi$  provides estimates for the logical error probability of some decoder. This suffices to demonstrate improvements over the symmetric ( $r = 1$ ) case.

- 
- [1] J. Niset, J. Fiurášek, and N. J. Cerf, No-Go Theorem for Gaussian Quantum Error Correction, *Phys. Rev. Lett.* **102**, 120501 (2009).
  - [2] J. Eisert, S. Scheel, and M. B. Plenio, Distilling Gaussian States with Gaussian Operations is Impossible, *Phys. Rev. Lett.* **89**, 137903 (2002).
  - [3] J. Fiurášek, Gaussian Transformations and Distillation of Entangled Gaussian States, *Phys. Rev. Lett.* **89**, 137904 (2002).
  - [4] G. Giedke and J. Ignacio Cirac, Characterization of Gaussian operations and distillation of Gaussian states, *Phys. Rev. A* **66**, 032316 (2002).
  - [5] C. Vuillot, H. Asasi, Y. Wang, L. P. Pryadko, and B. M. Terhal, Quantum error correction with the toric Gottesman-Kitaev-Preskill code, *Phys. Rev. A* **99**, 032344 (2019).
  - [6] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, *Phys. Rev. A* **64**, 012310 (2001).
  - [7] L. Fejes, Über die dichteste Kugellagerung, *Math. Z.* **48**, 676 (1942).
  - [8] K. Noh, V. V. Albert, and L. Jiang, Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes, *IEEE Trans. Inf. Theory* **65**, 2563 (2019).
  - [9] J. Harrington and J. Preskill, Achievable rates for the Gaussian quantum channel, *Phys. Rev. A* **64**, 062301 (2001).
  - [10] K. Noh, S. M. Girvin, and L. Jiang, Encoding an Oscillator into Many Oscillators, *Phys. Rev. Lett.* **125**, 080503 (2020).
  - [11] 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).

- [12] K. Noh and C. Chamberland, Fault-tolerant bosonic quantum error correction with the surface–Gottesman-Kitaev-Preskill code, *Phys. Rev. A* **101**, 012316 (2020).
- [13] J. Edmonds, Paths, trees, and flowers, *Can. J. Math.* **17**, 449 (1965).
- [14] S. B. Bravyi and A. Yu. Kitaev, Quantum codes on a lattice with boundary, [arXiv:quant-ph/9811052](https://arxiv.org/abs/quant-ph/9811052).
- [15] A. Yu. Kitaev, Fault-tolerant quantum computation by anyons, *Ann. Physics* **303**, 2 (2003).
- [16] S. Bravyi, M. Suchara, and A. Vargo, Efficient algorithms for maximum likelihood decoding in the surface code, *Phys. Rev. A* **90**, 032326 (2014).
- [17] H. Bombin and M. A. Martin-Delgado, Optimal resources for topological two-dimensional stabilizer codes: Comparative study, *Phys. Rev. A* **76**, 012305 (2007).
- [18] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Tailoring Surface Codes for Highly Biased Noise, *Phys. Rev. X* **9**, 041031 (2019).
- [19] V. Murg, F. Verstraete, and J. I. Cirac, Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states, *Phys. Rev. A* **75**, 033605 (2007).
- [20] We work in units  $\hbar = 1$ .
- [21] S. Glancy and E. Knill, Error analysis for encoding a qubit in an oscillator, *Phys. Rev. A* **73**, 012325 (2006).
- [22] D. J. Weigand and B. M. Terhal, Generating grid states from Schrödinger-cat states without postselection, *Phys. Rev. A* **97**, 022341 (2018).
- [23] V. Giovannetti, A. S. Holevo, and R. García-Patrón, A solution of Gaussian optimizer conjecture for quantum channels, *Commun. Math. Phys.* **334**, 1553 (2015).
- [24] A. S. Holevo and R. F. Werner, Evaluating capacities of bosonic Gaussian channels, *Phys. Rev. A* **63**, 032312 (2001).
- [25] Although we only treat the case  $n = 1$  here, the discussion in this section generalizes to arbitrary  $n \in \mathbb{N}$ . Note that for  $n > 1$ ,  $\text{Sp}(2n, \mathbb{R})$  is a proper subgroup of  $\text{SL}(2n, \mathbb{R})$ , however for  $n = 1$  we have  $\text{Sp}(2, \mathbb{R}) = \text{SL}(2, \mathbb{R})$ .
- [26] N. L. Balazs and A. Voros, The quantized Baker’s transformation, *Ann. Phys.* **190**, 1 (1989).
- [27] J. H. Hannay and M. V. Berry, Quantization of linear maps on a torus-fresnel diffraction by a periodic grating, *Phys. D: Nonlinear Phenom.* **1**, 267 (1980).
- [28] A. Bouzouina and S. De Bièvre, Equipartition of the eigenfunctions of quantized ergodic maps on the torus, *Commun. Math. Phys.* **178**, 83 (1996).
- [29] Note that  $S_{\square} S_r \neq S_r S_{\square}$  in general. Here we use the convention that the former order of transformations is associated with an asymmetric hexagonal lattice.
- [30] I. Gourlay and J. F. Snowdon, Concatenated coding in the presence of dephasing, *Phys. Rev. A* **62**, 022308 (2000).
- [31] P. Aliferis and J. Preskill, Fault-tolerant quantum computation against biased noise, *Phys. Rev. A* **78**, 052331 (2008).
- [32] Z. W. E. Evans, A. M. Stephens, J. H. Cole, and L. C. L. Hollenberg, Error correction optimisation in the presence of X/Y asymmetry, [arXiv:0709.3875](https://arxiv.org/abs/0709.3875).
- [33] A. M. Stephens, Z. W. E. Evans, S. J. Devitt, and L. C. L. Hollenberg, Asymmetric quantum error correction via code conversion, *Phys. Rev. A* **77**, 062335 (2008).
- [34] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh Error Threshold for Surface Codes with Biased Noise, *Phys. Rev. Lett.* **120**, 050505 (2018).
- [35] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *J. Math. Phys.* **43**, 4452 (2002).
- [36] H. Bombin and M. A. Martin-Delgado, Topological Quantum Distillation, *Phys. Rev. Lett.* **97**, 180501 (2006).
- [37] J. Guillaud and M. Mirrahimi, Repetition Cat Qubits for Fault-Tolerant Quantum Computation, *Phys. Rev. X* **9**, 041053 (2019).
- [38] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, S. T. Flammia, and S. M. Girvin, Bias-preserving gates with stabilized cat qubits, [arXiv:1905.00450](https://arxiv.org/abs/1905.00450).
- [39] 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).
- [40] F. Merz and J. T. Chalker, Two-dimensional random-bond Ising model, free fermions, and the network model, *Phys. Rev. B* **65**, 054425 (2002).
- [41] Our actual implementation differs slightly from the pseudocode in Fig. 9 for ease of implementation: We directly use the distribution  $\pi$  to sample the errors  $E_{j,k}$ . That is, instead of sampling a displacement error vector  $v$  from a normal distribution and then computing the corresponding Pauli error  $E_{j,k}$  as in steps 7 and 8, we first compute the distribution  $\pi$  as in Step 10 and then sample each  $E_{j,k}$  independently and identically from  $\pi$ . These two descriptions are of course equivalent because  $\pi$  is the induced distribution over (single-qubit) errors.
- [42] For the modified algorithm described in Footnote 4, this is only approximately the case. But the difference is negligible as  $\tilde{\pi}$  is a good approximation to  $\pi$  as argued here.