

Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes

Ilan Tzitrin,^{1, 2,*} J. Eli Bourassa,^{1, 2,*} Nicolas C. Menicucci,³ and Krishna Kumar Sabapathy^{2, †}

¹*Department of Physics, University of Toronto, Toronto, Canada*

²*Xanadu, 777 Bay Street, Toronto ON, M5G 2C8, Canada*

³*Centre for Quantum Computation and Communication Technology, School of Science, RMIT University, Melbourne, Victoria, Australia*

Encoding a qubit in the continuous degrees of freedom of an oscillator is a promising path to error-corrected quantum computation. One advantageous way to achieve this is through Gottesman-Kitaev-Preskill (GKP) grid states, whose symmetries allow for the correction of any small continuous error on the oscillator. Unfortunately, ideal grid states have infinite energy, so it is important to find finite-energy approximations that are realistic, practical, and useful for applications. In the first half of this work we investigate the impact of imperfect GKP states on computational circuits independently of the physical architecture. To this end, we analyze the behaviour of the physical and logical content of normalizable GKP states through several figures of merit, employing a recently-developed modular subsystem decomposition. By tracking the errors that enter into the computational circuit due to imperfections in the GKP states, we are able to gauge the utility of these states for NISQ (Noisy Intermediate-Scale Quantum) devices. In the second half, we focus on a state preparation approach in the photonic domain wherein photon-number-resolving measurements on some modes of Gaussian states produce non-Gaussian states in others. We produce detailed numerical results for the preparation of GKP states alongside estimating the resource requirements in practical settings and probing the quality of the resulting states with the tools we develop. Our numerical experiments indicate that we can generate any state in the GKP Bloch sphere with nearly equal resources, which has implications for magic state preparation overheads.

CONTENTS

I. Introduction	2	4. Logical subsystem Bloch sphere	20
II. Formalism of bosonic grid codes	3	D. Circuits for GKP state preparation	20
A. Ideal GKP states	3	1. Algorithm for optimal circuits	20
B. Normalizable GKP states	3	2. Circuit optimization results	21
C. Modular subsystem decomposition	5	E. Experimental imperfections	22
D. Operations on normalizable GKP states and error-tracking	6	1. Stability analysis of optical elements	22
1. Prerecovery figures of merit	8	2. Effects of photon loss	22
2. Displacements in position	10		
3. Displacements in momentum	11	IV. Conclusions	23
4. Shearing operations	11		
5. Rotations	12	Acknowledgments	24
6. Squeezing and beamsplitters	12		
7. SUM gates	12	A. Notation, nomenclature, convention, and units	25
E. Error correction with normalizable GKP states	13	B. Derivations	25
III. Photonic state preparation and characterization	16	1. Dilution for the envelope operator $E(\epsilon)$	25
A. Preparation of non-Gaussian states	16	2. Readout from the logical subsystem	26
B. Core states for GKP	16	3. Conjugation and commutation with $E(\epsilon)$	26
C. Characterization of approximate states	17	4. Measures of non-Gaussianity	27
1. Average energy	18	5. Glancy-Knill probability of no error	27
2. Orthogonality	18		
3. Glancy-Knill property	19	C. Numerical techniques	27
		1. Algorithm 1 details	28
		2. Algorithm 2 details	28
		D. Approximate GKP Z , X , and H eigenstates	29
		1. Further comments on $ 0_A\rangle$	29
		2. Results for $ 1_A\rangle$, $ +_A\rangle$, and $ H_{+A}\rangle$	30
		E. Additional characterization of approximate GKP states	31
		1. Projectors	31
		2. Quantum error correction matrix	31

* These two authors contributed equally.

† krishna@xanadu.ai

I. INTRODUCTION

Bosonic codes are quantum codes where logical qubits are encoded in states of continuous-variable systems. The most well-known among bosonic codes are those introduced by Gottesman, Kitaev and Preskill (GKP) [1], also known as **grid states**. These codes were constructed to correct displacement errors in phase space with the help of their inherent translation symmetry. An error analysis of finite energy GKP states was presented in [2]. The GKP code has garnered much interest in recent years for various applications.

Fault-tolerant quantum computation (FTQC) using GKP qubits was shown to be possible in the context of continuous-variable measurement-based quantum computation, along with a conservative upper bound of 20.5 dB on the squeezing required [3]. This result has spurred research in this direction. It has been demonstrated that the only non-Gaussian resource required for universal FTQC using GKP codes is the ability to faithfully prepare the logical zero state [4]. It has also been shown that noise in the anti-squeezing quadrature of the squeezed states is immaterial in such an implementation [5]. The authors in [5] further pointed out that practical squeezing thresholds were closer to 15 – 17 dB depending on the choice of qubit-level error-correcting code, a point previously made in [3] but rarely appreciated. The squeezing requirement was further reduced to \sim 10 dB [6] alongside methods to track the error correction [7] and perform a hybrid error correction procedure using digital and analog information [8]. Further refinements to the techniques and the incorporation of practical imperfections has led to a squeezing requirement that is less than 10 dB for a certain loss parameter range [9].

There are several other applications of the GKP code. The performance of GKP codes, including the hexagonal GKP code, was compared to other codes with respect to transmission through a bosonic loss channel in Refs. [10] and [11]. Comparison of the structure of the GKP code and the rotation-symmetric code was presented in [12]. The use of GKP qubits for quadratic noise suppression in continuous variables was considered in [13]. The surface code using GKP qubits has been considered in [6, 14–16]. The use of GKP qubits as a displacement sensor was explored in [17]. A probabilistic preparation of GKP qubits and its role in FTQC and sampling problems was presented in [18]. The sign-altered GKP states were studied as a means for decoherence suppression in [19]. Further, the fundamental mathematical structure of the GKP code states was presented in [20–22], connections among the different approximation schemes were analytically explored in [23], and aspects of the syndrome extraction was considered in [24].

Considering the extensive utility and optimality [11] of the GKP code, preparing high quality GKP states

is broadly desirable. There are several proposals and demonstrations for generating GKP qubits across various platforms, including ion traps [25–28], superconducting cavities [29], light-matter interactions [30, 31], cross-Kerr interactions [32], using phase estimation [33] (including its fault-tolerant version [34, 35]), and an analysis of practical aspects of modular quadrature measurements [36] all in circuit-QED, using cat states [37, 38], optical methods [39, 40], photon catalysis [41], a machine learning method that uses self-Kerr interactions [42], realization in time-frequency encodings [43], as eigenstates of tailored Hamiltonians in superconducting circuits [44], a plausible attempt using quantum walks [45], and a measurement-free proposal [46].

We focus on the photonic method using the minimal resources of squeezed displaced vacuum states, interferometers, and photon number-resolving detectors first introduced in [47] and subsequently developed in [48–50]. Our scheme is probabilistic in nature due to the postselection on specific photon detection outcomes. Photonic platforms have a proven track-record for implementing various forms of error correction circuits such as two-qubit encoding in polarization modes [51], topological error correction with an 8-photon cluster state [52], graph state codes [53, 54], Shor’s 9-qubit code [55], 5 wavepacket error correcting code [56], qubit codes preventing loss errors [57], and certain continuous-variable error correction [58, 59] with applications to quantum sensing [60].

In this work, we focus on the role played by physical approximations to ideal grid states. The approximations reveal themselves in various figures of merit and circuit characterizations whose understanding is central for NISQ devices. We use the recently-introduced modular subsystem decomposition [22] to monitor and mitigate errors on the logical qubit. We then provide an extensive analysis of the resource requirements for preparing these approximate states and useful benchmarks for employing them in computational circuits.

The paper is organized as follows. In Sec. II, we review the formalism of the GKP code independently of the platform for its implementation. This includes a synthesis of results for the ideal (non-normalizable) GKP states in II A and the finite-energy GKP states in II B; an overview of the modular decomposition in II C and its relevance to analyzing GKP states; a set of prescriptions for tracking, quantifying, and alleviating errors in computation induced by imperfect GKP states in II D; and a discussion of error correction and recovery with the normalizable states in II E.

Sec. III is dedicated to the preparation of approximate GKP states in the optical domain. In III A we discuss an architecture analogous to Gaussian BosonSampling devices for the generation of general non-Gaussian states; in III B we discuss a framework for approximating GKP states; in III C we characterize the approximate states using the formalism described in Sec. II; in III D we present our strategy for the preparation of these approximate

GKP states with optical circuits and provide our numerical results for the optimization of these circuits; and finally, in III E, we give comments on incorporating experimental imperfections into our scheme.

II. FORMALISM OF BOSONIC GRID CODES

Our physical landscape consists of continuous-variable systems whose Hilbert space is $L^2(\mathbb{R}^n)$, the square-integrable functions defined over the space of n real variables corresponding to n -mode systems. Examples of continuous-variable systems include modes of an electromagnetic field, harmonic oscillator chains, phonon modes in materials, continuous modes of ion traps and superconducting circuits. In what follows we work in units where $\hat{q} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger)$ and $\hat{p} = \frac{-i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger)$, so that $[\hat{q}, \hat{p}] = i$, and $\hbar = 1$. This means the measured variance of a vacuum state is $\langle \hat{q}^2 \rangle = \langle \hat{p}^2 \rangle = \frac{1}{2}$. For more details on our conventions, see App. A.

A. Ideal GKP states

The ideal GKP states [1] are defined as the simultaneous eigenstates of the continuous-variable stabilizer elements

$$S_q = e^{i(2\sqrt{\pi})(S_{11}\hat{q} + S_{21}\hat{p})} \quad (1a)$$

$$S_p = e^{i(2\sqrt{\pi})(S_{12}\hat{q} + S_{22}\hat{p})} \quad (1b)$$

such that $\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \in Sp(2, \mathbb{R})$, the real symplectic group in two dimensions. The Wigner function (see App. B 4 for the definition) of the ideal GKP states is a sum of delta functions located at the points of a lattice in phase space; the shape and spacing of this lattice is determined by \mathbf{S} . For example, rectangular GKP states are associated with the matrix

$$\mathbf{S}_{\text{rect}} = \begin{bmatrix} \sqrt{\pi}/\alpha & 0 \\ 0 & \alpha/\sqrt{\pi} \end{bmatrix}, \quad (2)$$

for $\alpha \in \mathbb{R}_{\neq 0}$, so that the unit cell of the lattice has dimensions $\frac{\sqrt{\pi}}{\alpha} \times \frac{\alpha}{\sqrt{\pi}}$ and an area of 1, while hexagonal GKP states [1, 11] have

$$\mathbf{S}_{\text{hex}} = \left(\frac{2}{\sqrt{3}} \right)^{1/2} \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}. \quad (3)$$

We be focusing on square-lattice states – that is, rectangular GKP states with $\alpha = \sqrt{\pi}$. For these states, the stabilizer elements become

$$S_q^{\text{square}} = D(i\sqrt{2\pi}) \equiv Z(2\sqrt{\pi}) \quad (4a)$$

$$S_p^{\text{square}} = D(\sqrt{2\pi}) \equiv X(2\sqrt{\pi}), \quad (4b)$$

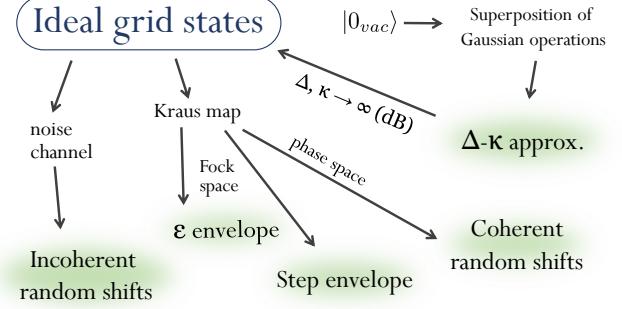


Figure 1. Various ways to obtain physical states starting from ideal grid states. From left to right, the approximations in green clouds correspond to discussions around Eq. (20), (12), (16), (15), and (8).

where $D(\beta) \equiv e^{\beta\hat{a}^\dagger - \beta^*\hat{a}}$ is the displacement operator, $X(q) \equiv D(q/\sqrt{2})$ is a displacement in position by q and $Z(p) \equiv D(ip/\sqrt{2})$ is a displacement in momentum by p . After identifying the logical Z and X operations for this code (logical operators are denoted by an overline),

$$\bar{Z} = Z(\sqrt{\pi}) \text{ and } \bar{X} = X(\sqrt{\pi}), \quad (5)$$

one can infer that the ideal encoded logical square GKP states, which we label with the subscript I , are infinite superpositions of infinitely squeezed states – delta spikes arranged like a comb – spaced by $2\sqrt{\pi}$ in position:

$$|0_I\rangle \equiv \sum_{n=-\infty}^{\infty} |2n\sqrt{\pi}\rangle_q. \quad (6)$$

The encoded logical 1 state is then just a q -displacement by $\sqrt{\pi}$ of the 0 state: $|1_I\rangle \equiv X(\sqrt{\pi})|0_I\rangle$. The periodicity of the delta spikes implies that the ideal GKP code can correct displacements of position and momentum of up to $\sqrt{\pi}/2$, i.e., those displacements that do not confuse a logical 0 for a logical 1 [1]. The \bar{X} eigenstates are then just

$$|+I\rangle \equiv \sum_{n=-\infty}^{\infty} |n\sqrt{\pi}\rangle_q, \quad (7)$$

and $|-_I\rangle \equiv Z(\sqrt{\pi})|+I\rangle$.

B. Normalizable GKP states

In an experimental setting one will be dealing with finitely squeezed states. The canonical way to generate normalizable GKP states [61], which we label with a subscript corresponding to the normalization scheme, is to replace the delta functions with Gaussians of width Δ and then introduce an overall Gaussian envelope of

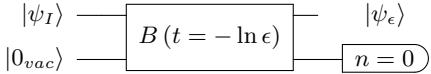


Figure 2. A schematic of how to obtain the normalizable GKP state $|\psi_\epsilon\rangle \equiv e^{-\epsilon\hat{n}} |\psi_I\rangle$ through an optical circuit, as noted in [13] and shown in App. B 1. An ideal GKP state and the vacuum state pass through the first and second mode of a beamsplitter $B(\theta, \phi)$ with transmissivity $-\ln \epsilon$ (see Eq. (62)); the second mode is measured and post-selected on a vacuum state.

width κ^{-1} :

$$|0_{\Delta,\kappa}\rangle \propto \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\kappa^2(2n\sqrt{\pi})^2} X(2n\sqrt{\pi}) |\Delta\rangle_q \quad (8)$$

$$|1_{\Delta,\kappa}\rangle \propto \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\kappa^2[(2n+1)\sqrt{\pi}]^2} X[(2n+1)\sqrt{\pi}] |\Delta\rangle_q, \quad (9)$$

where $|\Delta\rangle$ is defined so that

$$\langle q | \Delta \rangle = \left(\frac{1}{\pi\Delta^2} \right)^{\frac{1}{4}} e^{-\frac{q^2}{2\Delta^2}}. \quad (10)$$

This is just one prescription to transition from infinite to finite energy states, as shown in Fig. 1. Recently, three conventional approximations of the GKP codes were analytically shown to be equivalent in [23].

It is common to work in the regime where $\Delta = \kappa$, so that in the $\Delta \rightarrow 0$ limit the number of spikes increases while the width of each spike decreases, and the normalizable states approach the ideal ones. In this case we omit the κ from the subscript and write simply

$$|\psi_\Delta\rangle \equiv |\psi_{\Delta,\kappa=\Delta}\rangle. \quad (11)$$

A less cumbersome way than Eq. (8) of expressing $|\psi_\Delta\rangle$, as pointed out in [3, 11], is to apply a non-unitary envelope operator $E(\epsilon) \equiv e^{-\epsilon\hat{n}}$ to the ideal state, where $\hat{n} \equiv \hat{a}^\dagger \hat{a} = \frac{1}{2}(\hat{q}^2 + \hat{p}^2)$ is the number operator:

$$|\psi_\epsilon\rangle \equiv E(\epsilon) |\psi_I\rangle. \quad (12)$$

One can think of this operator as being the result of interfering an ideal GKP state and a vacuum state at a beamsplitter, measuring one of the modes, and post-selecting on the vacuum [13]. We provide an illustration of this in Fig. 2 and a derivation in App. B 1. Note that $|\psi_\epsilon\rangle \approx |\psi_\Delta\rangle$ whenever $\Delta = \kappa$ and ϵ are small. From [16], the more precise regime is whenever

$$\tanh \frac{\Delta}{2} \approx \frac{\Delta}{2}. \quad (13)$$

Since $\tanh x = x - \frac{1}{3}x^3 + \dots$, this will occur whenever $\frac{1}{24}\Delta^3$ is negligible.

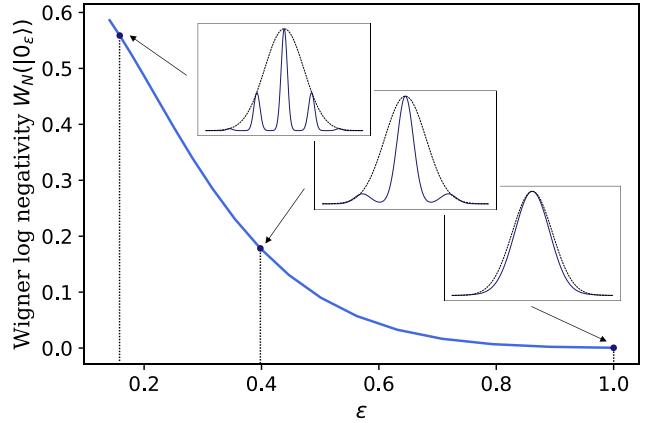


Figure 3. The Wigner logarithmic negativity W_N (defined in App. B 4) of the normalizable GKP state $|0_\epsilon\rangle$ from Eq. (12) as a function of ϵ shown as solid (blue) line. **Insets:** Highlighted from left to right are wavefunctions corresponding to ϵ values of 0.158, 0.398 and 1 (Δ values of 8, 4, and 0 dB). The overall Gaussian envelope of width $\kappa = \Delta^{-1}$ is also drawn. The Gaussian in the third inset noticeably does not perfectly envelop the $\epsilon = 1$ state because here there is an error of order 0.04 in the approximation between Δ and ϵ (as shown in Eq. (13) and Table IV). W_N decreases with increasing ϵ , i.e. as the number of peaks in our GKP state decreases, the width of each one increases, and the state becomes more Gaussian.

Since we require non-Gaussian resource states for universal quantum computation, it is important to quantify the non-Gaussianity of the normalizable GKP states. For this, we plot the Wigner logarithmic negativity W_N (see App. B 4 for the definition) of the $|0_\epsilon\rangle$ state as a function of ϵ in Fig. 3. We see that W_N decreases for increasing ϵ , as expected.

We can, in principle, come up with other ways of approximately normalizing the ideal states. In general, we have that

$$|\psi_G\rangle \equiv G |\psi_I\rangle, \quad (14)$$

for some operation G , which we can regard as an error or a single Kraus operator. Note that G will not be trace-preserving in general; even though we have written the left-hand-side of (14) as a ket, it is understood that the state will need to be normalized. For example, G can be defined as the operator

$$G = \sqrt{\frac{2}{\pi\Delta^2}} \int d^2\alpha e^{-|\alpha|^2/\Delta^2} D(\alpha), \quad (15)$$

in other words a Gaussian distribution of displacements, as in [1, 10, 23].

G can also be chosen so that the normalizable state is a weighted superposition of Fock states, as in

$$|\mu_G\rangle = \sum_{n=0}^{\infty} g(n) \langle n | \mu_I \rangle |n\rangle, \quad (16)$$

for some envelope function $g(n)$. For example, we can set g to be a step function; or, to obtain (12), we can make $g(n) = e^{-\epsilon n}$ [62]. We will explore in Sec. III which envelope functions are best to target in the state preparation scheme we consider.

There are several advantages to using Eq. (12) to denote the normalizable states. Beyond its compactness, it provides us with a convenient way to explore the effects of deviating from ideal GKP states on Gaussian operations, as required for implementing Clifford gates, and on the logical content of the states, as we will see. It is also important to point out that the displacement and envelope operators will, in general, not commute (see App. B 3 for explicit commutation and conjugation relations with $E(\epsilon)$). This means that the logical state obtained by displacing $|0_G\rangle$ by $\sqrt{\pi}$ in position will not be the same as the state obtained by applying an envelope to $|1_I\rangle$. For example,

$$X(\sqrt{\pi})|0_\epsilon\rangle = X(\sqrt{\pi})E(\epsilon)|0_I\rangle \quad (17)$$

$$= E(\epsilon)e^{\sqrt{\pi}(e^\epsilon \hat{a}^\dagger - e^{-\epsilon}\hat{a})}|0_I\rangle \quad (18)$$

$$\neq E(\epsilon)X(\sqrt{\pi})|0_I\rangle = E(\epsilon)|1_I\rangle. \quad (19)$$

To avoid ambiguity, we will use the prescription implied by Eq. (14) for our normalizable states and explicitly write, e.g., $X(\sqrt{\pi})|0_G\rangle$ where necessary. We will investigate in the following sections what impact this difference will have on practical considerations and on the logical information encoded in the states.

Note that a more general way of writing down normalizable states is through some noise channel, \mathcal{K} , acting on the ideal states:

$$\rho_\kappa = \mathcal{K}(|\mu_I\rangle\langle\mu_I|) \quad (20)$$

This approach is taken in [13] with \mathcal{K} corresponding to random Gaussian displacement errors, also known as the Gaussian classical noise channel.

C. Modular subsystem decomposition

As one deviates from the ideal GKP states, it becomes less obvious what the logical state is, where the information resides, and how to address and access it. Answers to these questions are facilitated by an important tool for analyzing states wherein a qubit is encoded periodically in a infinite-dimensional Hilbert space, the *modular subsystem decomposition*, investigated by Pantaleoni et al. [22]. Here we briefly review its formalism and discuss its application to approximate GKP states.

Given some real number α corresponding to the spacing between the logical basis states in position, established in Sec. II A, we can decompose any position eigenket $|s\rangle_q$ in an infinite-dimensional Hilbert space \mathcal{H} as

$$|s\rangle_q = |\alpha m + u\rangle \equiv |m, u\rangle, \quad (21)$$

where $m \in \mathbb{Z}$ and $u \in [-\alpha/2, \alpha/2]$. We call αm the *integer part* of $s \pmod{\alpha}$ and u the *fractional part* of $s \pmod{\alpha}$. We can subsequently decompose the physical space \mathcal{H} into $\mathcal{H} = \mathcal{L} \otimes \mathcal{G}$: a two-dimensional *logical* space \mathcal{L} corresponding to our qubit and another (virtual) infinite-dimensional *gauge* space \mathcal{G} corresponding to everything else. Position eigenkets break down as

$$|s\rangle_q = |\mu\rangle_{\mathcal{L}} \otimes |\tilde{m}, \tilde{u}\rangle_{\mathcal{G}}, \quad (22)$$

where $\mu = \text{parity}(m)$, $\tilde{m} = \frac{1}{2}(m - \mu)$, and $\tilde{u} = u$. Effectively, this decomposition amounts to stitching together the wavefunction sitting within position bins corresponding to the logical μ . In this subsystem picture, the ideal GKP states can be written

$$|\psi_I\rangle = |\bar{\psi}\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}}, \quad (23)$$

so that the logical mode is completely separable from gauge mode and we can recover the logical information perfectly through a trace over \mathcal{G} . Physically, we can access the logical information with a binned homodyne measurement, as shown in App. B 2. For the normalizable states, however, we see that

$$|\psi_\epsilon\rangle = E(\epsilon)(|\bar{\psi}\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}}), \quad (24)$$

for example. Since $E(\epsilon)$ acts on the full space \mathcal{H} , it will generally entangle the logical and gauge modes, leaving our logical qubit in a mixed state

$$\rho^{\mathcal{L}}(\epsilon) = \text{Tr}_{\mathcal{G}}[E(\epsilon)|\psi_I\rangle\langle\psi_I|E(\epsilon)]. \quad (25)$$

As ϵ grows, $\rho^{\mathcal{L}}$ will find itself somewhere inside the Bloch sphere; we plot its “trajectory” as a function of ϵ for the \bar{Z} and \bar{X} basis states in Fig. 4a. Notice that, for $\epsilon \gg 0$, the states converge at the same point: this is the vacuum state, the only state picked out by the envelope in this regime [63]. We can also see how the Bloch sphere itself changes as a function of ϵ , as depicted in Fig. 4b.

Remark. We calculate this efficiently by first noting that every unnormalized qubit operator can be decomposed as $\rho^{\mathcal{L}}(\epsilon) = \frac{1}{2}\sum_{i=0}^3 s_i \sigma_i$, where σ_i correspond to the identity and the Pauli matrices, and \mathbf{s} is the Stokes vector for the state (see for e.g. [64]). If $s_0 = 1$ then (s_1, s_2, s_3) corresponds to a Bloch vector. Thus, we can find the matrix that transforms the Stokes vectors corresponding to the ideal GKP qubits under application of the completely positive (but not trace preserving) map $E(\epsilon)$. Given the Stokes vectors corresponding to the states before renormalizing, we can now individually renormalize them by simply dividing by the first coefficient of each Stokes vector, yielding Bloch vectors in the remaining three components.

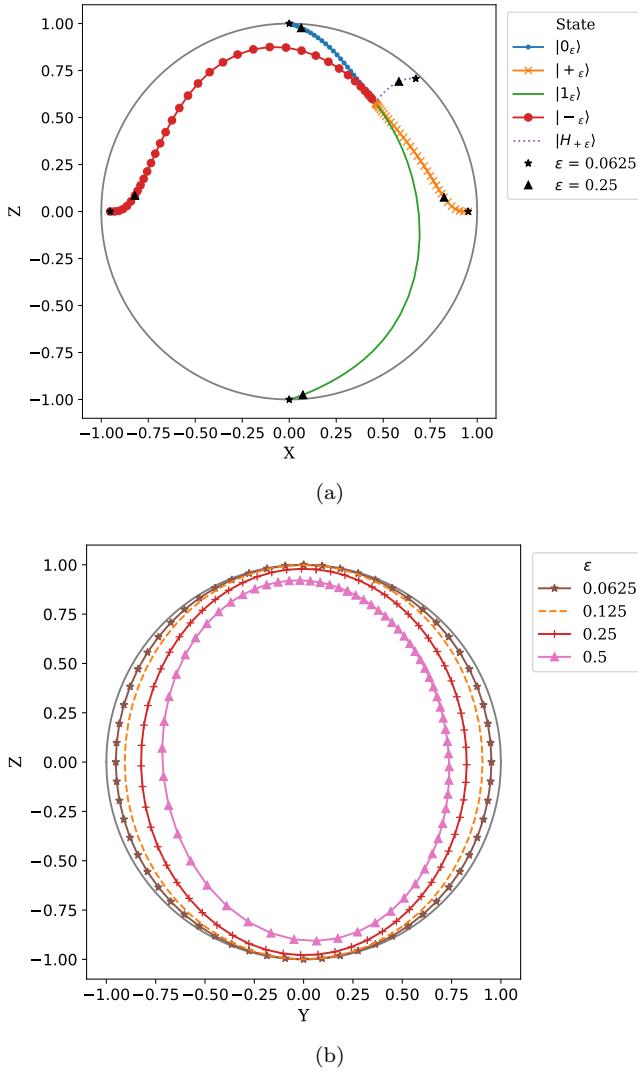


Figure 4. Effects of the normalization envelope $E(\epsilon) = e^{-\epsilon \hat{n}}$ on the logical content of GKP states. In (a), trajectories of $|0_\epsilon\rangle$, $|1_\epsilon\rangle$, $|\pm_\epsilon\rangle$, and $|H_+\rangle$ as a function of ϵ confined to the X - Z plane of the Bloch sphere (here H_+ is the +1 eigenstate of the Hadamard operator; see (31)). For large ϵ , all Fock states but $|0\rangle$ are exponentially suppressed, causing all the trajectories to converge at the logical subsystem point corresponding to the physical vacuum state. We additionally show locations for reasonable values of ϵ : $\epsilon = 0.0625 \approx \Delta = 12$ dB, and $\epsilon = 0.25 \approx \Delta = 6$ dB are denoted by a star and triangle, respectively (see Eq. 13 and App. A for conversions). We see that the required ϵ (dB) values to achieve higher purity states are lower (higher) towards the equator of the Bloch sphere. In (b), the modification of the X - Z plane of the ideal Bloch sphere as a function of ϵ . We see that points towards the equator are more distorted.

D. Operations on normalizable GKP states and error-tracking

The implementation of Clifford gates is a critical step for universal quantum computation. The Clifford group

\bar{U}	U physical (symbol)	U physical (name)
\bar{X}	$D(\sqrt{\pi/2}) = X(\sqrt{\pi})$	q displacement
\bar{Z}	$D(i\sqrt{\pi/2}) = Z(\sqrt{\pi})$	p displacement
\bar{H}	$F = R(\pi/2) = e^{i\frac{\pi}{2}\hat{n}}$	Fourier gate; $\frac{\pi}{2}$ rotation
\bar{P}	$P = e^{i\frac{1}{2}\hat{q}^2}$	Phase gate
CNOT	$SUM = e^{-i\hat{q}_1 \otimes \hat{p}_2}$	SUM gate

Table I. Conventional association between logical qubit operations and physical Gaussian transformations for ideal GKP encoding. Note that the physical gates are not unique due to fact that the stabilizers (4) and a π phase shift $e^{i\pi\hat{n}}$ act trivially on the code space. This means that any displacement by a Gaussian-integer multiple of $\sqrt{\pi/2}$ acts as a Pauli operator; F^\dagger also represents \bar{H} ; and a SUM gate of any odd-integer weight is also a CNOT. (A SUM gate of weight g is $e^{-ig\hat{q}_1 \otimes \hat{p}_2}$. The one shown in the table is weight 1.)

on n qubits, \mathcal{C}_n , is defined through its action on the Pauli group, \mathcal{P}_n , which consists of n -fold tensor products of Pauli gates. Any $U \in \mathcal{C}_n$ maps the Pauli group to itself under conjugation:

$$A \in \mathcal{P}_n \implies UAU^\dagger \in \mathcal{P}_n. \quad (26)$$

In principle, it is enough to consider a set of generators of the Clifford group, for example the Hadamard gate,

$$\bar{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (27)$$

and the phase gate,

$$\bar{P} = \sqrt{\bar{Z}} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \quad (28)$$

along with the two-qubit CNOT gate (also known as a CX or controlled- X gate),

$$\overline{\text{CNOT}} = |\bar{0}\rangle \langle \bar{0}| \otimes I + |\bar{1}\rangle \langle \bar{1}| \otimes \bar{X}. \quad (29)$$

Any Clifford element can then be obtained, in principle, from applications of the above gates in the generator set. In practice, and certainly with GKP states, one should also explicitly define other fundamental Clifford gates like X and Z rather than relying on compositions of the minimal set of generators.

In addition to these, one also requires an operation beyond the Clifford group (referred to as non-Clifford element), such as the $\pi/8$ gate or the T gate,

$$\bar{T} = \sqrt{\bar{P}} = \frac{1}{\sqrt{2}} e^{i\frac{\pi}{8}} \begin{bmatrix} e^{-i\frac{\pi}{8}} & 0 \\ 0 & e^{i\frac{\pi}{8}} \end{bmatrix}, \quad (30)$$

for fault-tolerant universal quantum computation.

Alternatively, non-Clifford gates can be effected by preparing special resource states called *magic states* that

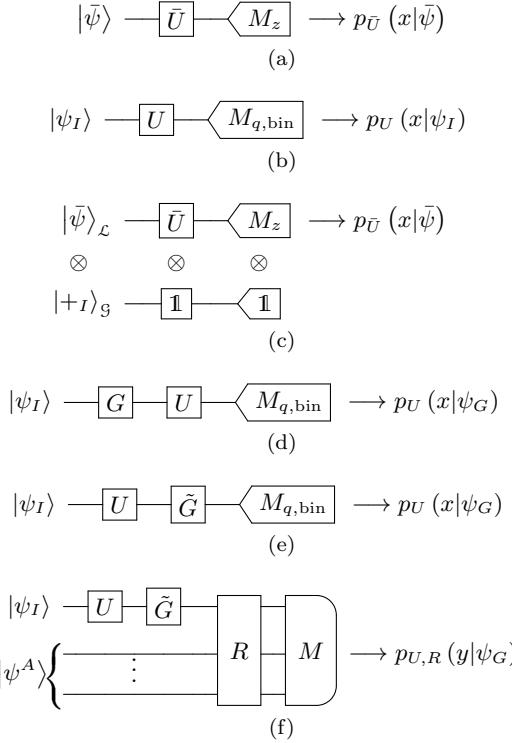


Figure 5. (a) The ideal logical circuit: a Clifford unitary, \bar{U} , is applied to a qubit, $|\bar{\psi}\rangle$, followed by Pauli Z measurements resulting in a bit, x , with probability $p_{\bar{U}}(x|\bar{\psi})$. (b) The ideal encoded physical circuit: a Gaussian unitary, U , which implements \bar{U} is applied to an ideal GKP state, followed by (binned) homodyne measurements to produce a bit with probability $p_U(x|\psi_I)$; (c) the circuit in (b) redrawn with a modular subsystem decomposition (note that we are only able to decompose the U in this way for an input ideal GKP state). The probability distributions in (a), (c), and (b) are identical: $p_{\bar{U}}(x|\bar{\psi}) = p_U(x|\psi_I)$; (d) The realistic GKP circuit, where the initial states are subject to a finite energy envelope, G . The unitary remains the same and a homodyne measurement produces a bit with probability $p_U(x|\psi_G)$; (e) the circuit in (d) but we highlight that the envelope can be conjugated with the Gaussian with appropriate modifications. (f) A general GKP circuit: the same as (e), but now we include a unitary recovery operation on the data states and ancillae $|\psi^A\rangle$. The measurement operator – which could be homodyne or PNR – is generalized to be on all modes, now producing a bit string of some composite variable, y . See the remark in the main text for a generalization of this schematic to a multimode circuit.

can be used in a type of gate teleportation circuit [65]. We can implement the \bar{T} gate, for example, with a supply of H -type magic states, which are equivalent through Clifford unitaries to the Hadamard eigenstates. Refer to Fig. 5 of [1] for an example gate teleportation circuit. The eigenstate corresponding to the $+1$ eigenvalue is

$$|\bar{H}_+\rangle \equiv \cos \frac{\pi}{8} |\bar{0}\rangle + \sin \frac{\pi}{8} |\bar{1}\rangle. \quad (31)$$

Here we will only consider Gaussian gates, because non-Clifford gates that correspond to non-Gaussian gates can

be realized via gate teleportation with magic states and Gaussian gates. This means all the gate resources remain Gaussian, and the non-Gaussian resources are imposed on the state preparation (of logical and magic GKP states), which we will discuss in Sec. III.

Suppose we would like to apply a gate $\bar{U} \in \mathcal{C}$ to a logical qubit $|\bar{\psi}\rangle$. If we encode $|\bar{\psi}\rangle$ in an ideal GKP state $|\psi_I\rangle$, \bar{U} will correspond to a Gaussian operation U on the physical space – that is, a combination of a symplectic transformation on the quadrature operators and a displacement. We recall that not all Gaussian operations correspond to Clifford gates on the ideal GKP codes, just particular non-unique ones. Writing $|\psi_I\rangle$ in its modular subsystem decomposition, we have that

$$U |\psi_I\rangle = (\bar{U} |\bar{\psi}\rangle_{\mathcal{L}}) \otimes |+I\rangle_{\mathcal{G}}. \quad (32)$$

We can see that this is true by applying U to both sides of Eq. (23): on the right-hand side we must obtain an ideal-GKP-encoded state (since U implements a Clifford unitary \bar{U}). Thus the resulting state can also be written in the form (23), and the action on the logical subsystem must be that of \bar{U} . Importantly, however, note that this does *not* mean that the physical operation can be modelled as $\bar{U}_{\mathcal{L}} \otimes I_{\mathcal{G}}$ since the decomposed operator is entangling in general [22]. The reason the gauge mode is unchanged in this particular case is that the entangling pieces of the decomposed gate do nothing when the gauge mode is exactly $|+I\rangle$. In general, there are many physical operations U corresponding to a given logical gate \bar{U} ; the standard mapping between \bar{U} and U is given in Table I.

Remark. The formalism above can be generalized to multimode states, which is necessary for generating entanglement and implementing gate teleportation. In the ideal case, a logical gate $\bar{U} \in \mathcal{C}_n$ acting on n qubits can be implemented as a Gaussian operation U on n oscillator modes. Otherwise, U will cause the logical and gauge subsystems of the different modes to interact [22]. Note further that we ought to allow for non-terminal measurements and classical feedforward within the computation. Although the unitaries will then generally depend on the result of these measurements, they will remain Gaussian.

For normalizable states, we have

$$U |\psi_\epsilon\rangle = UE(\epsilon) |\psi_I\rangle = \tilde{E}(\epsilon) U |\psi_I\rangle, \quad (33)$$

where $\tilde{E}(\epsilon) \equiv UE(\epsilon)U^\dagger$. Therefore, applying a Gaussian operation to a normalizable state can be viewed as applying this operation to an ideal state followed by a modified envelope. If we write the operation U as some function u of the creation and annihilation operators, $U = u(\hat{a}^\dagger, \hat{a})$, we can also see, as shown in App. B 3, that

$$u(\hat{a}^\dagger, \hat{a}) E(\epsilon) = E(\epsilon) u(e^\epsilon \hat{a}^\dagger, e^{-\epsilon} \hat{a}). \quad (34)$$

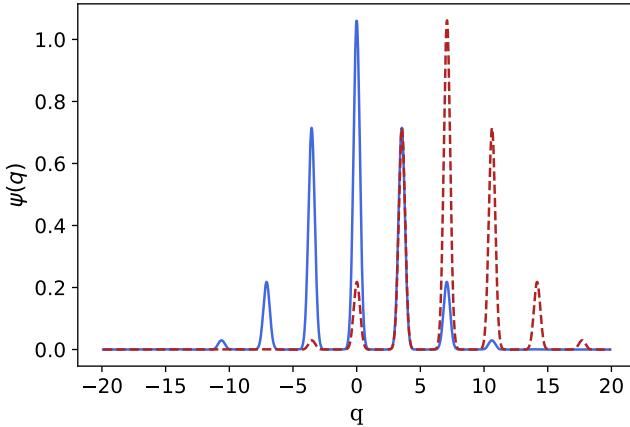


Figure 6. The normalizable GKP state $|0_\epsilon\rangle$ (blue, solid) and the same state after two applications of the $X(2\sqrt{\pi})$ gate (red, dashed) for $\epsilon = 0.063$ ($\Delta \approx 12$ dB). Because finite-energy GKP states have finite support in position space and hence lack a translational symmetry, the physical fidelity between the displaced state and the original state decreases with each application of $X(2\sqrt{\pi})$ even though this gate preserves the logical content.

Expressions (33) and (34) show that the interplay between perfect Gaussian operations and imperfect GKP states causes injury to both. The damage increases with increasing ϵ – corresponding to fewer and broader peaks in the GKP state – and with increasing powers of quadrature operators that feature in U .

In this section we explore ways to quantify the severity of this damage and approaches to mitigating it. For this, we consider models for computational circuits with increasing complexity, as in Fig. 5. First, we will overview figures of merit for normalizable GKP states before any kind of recovery operation. This will be important for ensuring that the state input to the recovery is the best possible, thereby easing the requirements on the error correction that follows. Armed with these figures of merit, we will analyze the effects of an important class of Gaussian operations on the normalizable GKP states. In Sec. II E, we will discuss the notion of error correction with approximate GKP states and revisit the metrics.

1. Prerecovery figures of merit

Here we define several metrics to probe the quality and usefulness of the normalizable GKP states. For each metric, we will use as a sanity check a toy circuit consisting of an even number, $2k$, of logical \bar{X} operations:

$$|\bar{0}\rangle \xrightarrow{} [\bar{X}] \cdots [\bar{X}] \equiv -[\bar{I}]$$

This circuit is equivalent to a logical identity. If we would like to implement it on an oscillator, we can encode $|\bar{0}\rangle$ to $|0_I\rangle$ and use the standard map $\bar{X} \rightarrow X(\sqrt{\pi})$ from Table I.

The ideal oscillator circuit is thus

$$|0_I\rangle \xrightarrow{} [X(\sqrt{\pi})] \cdots [X(\sqrt{\pi})] \equiv -[\bar{I}]$$

Thanks to the complete translational symmetry of the ideal GKP state, this circuit is also equivalent to an identity on the oscillator space (and hence the logical space), despite it effecting a net translation of $n\alpha$ in position. In a more practical setting we have

$$|0_G\rangle \xrightarrow{} [X(\sqrt{\pi})] \cdots [X(\sqrt{\pi})] \equiv -[X(n\sqrt{\pi})]$$

Because the normalized GKP states break the translational symmetry of the ideal states, the latest circuit is no longer an identity on the state space (see Fig. 6). The extent of the problems that this causes will be made clearer using the following figures of merit. Note that we always treat U as a unitary operation on the physical space that effects \bar{U} on the logical mode of an ideal GKP state.

a. Physical Fidelity. A straightforward figure of merit to consider is to compare physical states, for which we use the *physical fidelity*

$$F^P(|\psi\rangle, |\phi\rangle) \equiv |\langle\phi|\psi\rangle|^2. \quad (35)$$

We will want to compare the state before a transformation to the state after, and so we can use the shorthand notation

$$F_U^P(|\phi\rangle) \equiv F^P(U|\phi\rangle, |\phi\rangle). \quad (36)$$

If U_1 is a unitary for which $\bar{U} = \bar{1}$, then $F_{U_1}^P(|\psi_I\rangle) = 1$ for an ideal GKP state $|\psi_I\rangle$. We might therefore demand for a normalizable state that

$$F_{U_1}^P(|\psi_G\rangle) \approx 1. \quad (37)$$

However, this requirement is too stringent: to see this, we can refer back to our toy circuit, where $U_1 = X(2k\sqrt{\pi})$. In this case

$$F_{X(2k\sqrt{\pi})}^P(|\psi_\epsilon\rangle) = \langle\psi_\epsilon|X(2k\sqrt{\pi})|\psi_\epsilon\rangle \quad (38)$$

$$= \langle\psi_I|E(\epsilon)\tilde{E}(\epsilon)|\psi_I\rangle, \quad (39)$$

where $\tilde{E}(\epsilon) = e^{-\epsilon[(\hat{q}-2k\sqrt{\pi})^2 + \hat{p}^2]}$ (see Table V for more conjugation relations). For k high enough, our normalizable state can be displaced so much that the physical fidelity vanishes; however, the functional form of the wavefunction has not changed save for a rigid translation in position. This indicates that, while potentially useful for some applications, the physical fidelity does not adequately reveal the presence of information encoded in our normalizable state.

b. Logical Fidelity. Since we are interested in the logical content of the state rather than the content of the gauge mode, we might instead consider *logical fidelity*, i.e., the fidelity between the reduced logical states:

$$F^L(|\psi\rangle, |\phi\rangle) \equiv F[\text{Tr}_S(|\phi\rangle\langle\phi|), \text{Tr}_S(|\psi\rangle\langle\psi|)], \quad (40)$$

where F on the right-hand-side is the fidelity for mixed states, given by

$$F(\rho, \sigma) \equiv \text{Tr} |\sqrt{\rho}\sqrt{\sigma}|^2. \quad (41)$$

Again, we will make use of the notation

$$F_U^L(|\phi\rangle) \equiv F^L(U|\phi\rangle, |\phi\rangle), \quad (42)$$

and if U_1 is such that $\bar{U} = I$, then $F_U^L(|\psi_I\rangle) = 1$, and we might require

$$F_{U_1}^L(|\psi_G\rangle) \approx 1. \quad (43)$$

Returning to our toy circuit, if $U_1 = X(2k\sqrt{\pi})$, this amounts to a displacement of the gauge mode only:

$$X(2k\sqrt{\pi}) = \mathbb{1}_L \otimes X_S(2k\sqrt{\pi}), \quad (44)$$

so the logical fidelity is expected to be close to unity. This shows the utility of the modular subsystem decomposition in an analysis of imperfect GKP states.

c. *Distribution Distance*. Since we are ultimately interested in the result of the computation being accurate, we can also define a post-readout metric. Returning to Fig. 5, we can compare the output of the ideal circuit 5b with input state $|\psi_I\rangle$, which will be a bit string following some probability distribution $p_U(x|\psi_I)$, with the output $p_U(x|\psi_G)$ of the circuit 5d, initialized with a normalizable GKP state $|\psi_G\rangle$. For this we define the *distribution distance* through

$$\mathcal{D}_U^p(|\phi_G\rangle) = d[p_U(x|\phi_G), p_U(x|\phi_I)], \quad (45)$$

where d is a statistical distance, i.e., a generalized metric on the space of probability distributions. For a listing and comparison of probability metrics, see, for example, [66].

In general, we want that $\mathcal{D}_U^p(|\psi_G\rangle) \approx 0$ for any U . By focusing on the probability distribution, which we find after the readout, we do not require the state before the measurement in circuit 5f to be the same as the state before the measurement in circuit 5b. For example, in circuits 5d, any operation that modifies the ideal GKP states but keeps the probability distribution within the bin structure unchanged, e.g. a displacement less than $\sqrt{\pi}/2$, will still yield the same measurement statistics at readout.

With our toy circuit, if $U = X(2k\sqrt{\pi})$, we see that a binned homodyne measurement (see B 2) will extract the logical information from $|\psi_G\rangle$, implying that \mathcal{D}_U^p will be invariant with k .

d. *Mean and variant energy cost*. Given a state $|\phi\rangle$, we can characterize its energy through the quantities

$$\bar{n}(|\phi\rangle) \equiv \langle\phi|\hat{n}|\phi\rangle, \quad (46)$$

$$\sigma_{\hat{n}}^2(|\phi\rangle) \equiv \langle\phi|\hat{n}^2|\phi\rangle - \bar{n}^2(|\phi\rangle), \quad (47)$$

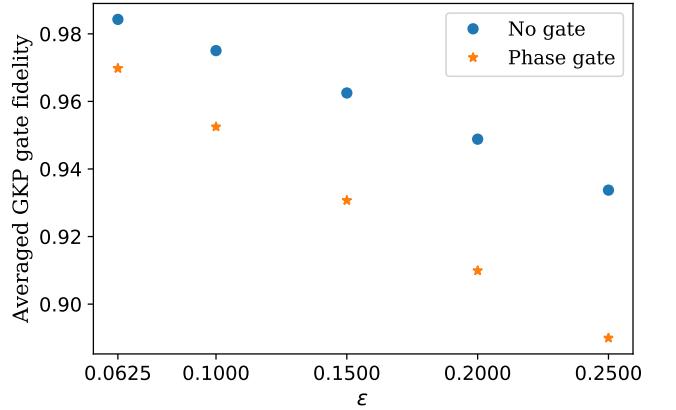


Figure 7. Averaged GKP gate fidelity $\bar{F}_{U,G}^L(\bar{U})$ from (50) evaluated for the logical and physical identity $\{\bar{U} = \mathbb{1}, U = \mathbb{1}\}$ (blue), and logical and physical phase gate $\{\bar{U} = \bar{P}, U = P\}$ (orange) assuming a normalizable GKP state with $G = E(\epsilon)$ as a function of ϵ . This plot shows that even without the application of any gate, the error envelope E creates a deviation, on average, from unity fidelity with the ideal data qubit state. The deviation is even greater in the case of a phase gate, the reasons for which are given in Sec. II D 4. In both cases the fidelity becomes worse as ϵ becomes bigger, as expected.

which correspond to the occupation number mean and variance [67]. Then we define the the mean $\mathcal{D}_U^{\bar{n}}$ and variant $\mathcal{D}_U^{\sigma^2}$ energy cost of a gate U on a state $|\phi\rangle$ to be

$$\mathcal{D}_U^{\bar{n}}(|\phi\rangle) \equiv \bar{n}(U|\phi\rangle) - \bar{n}(|\phi\rangle), \quad (48)$$

$$\mathcal{D}_U^{\sigma^2}(|\phi\rangle) \equiv \sigma_{\hat{n}}^2(U|\phi\rangle) - \sigma_{\hat{n}}^2(|\phi\rangle), \quad (49)$$

and we write \mathcal{D}_U^σ when we use the standard deviation. We would like that $|\mathcal{D}_U^{\bar{n}}|$ and $|\mathcal{D}_U^{\sigma^2}|$ are as small as possible. Note that the energy costs are physical figures of merit that translate to resource requirements. It is therefore no surprise that $\mathcal{D}_U^{\bar{n}}$ from our toy circuit with $U = X(2k\sqrt{\pi})$ increases monotonically with k as our state finds itself further and further from the origin, while the physical fidelity decreases.

e. *Extremized and averaged measures*. The figures of merit just described are single-shot measures; for a more complete picture, one can minimize, maximize, or average these measures over a set of states. Let us attempt to do so for the logical fidelity:

$$\bar{F}_{U,G}^L(\bar{U}) \equiv \int d\bar{\phi} F^L(U|\phi_I\rangle, U|\phi_G\rangle) \quad (50)$$

$$= \int d\bar{\phi} \langle\bar{\phi}|\bar{U}^\dagger \mathcal{E}_{U,G}(\bar{\phi}) \bar{U}|\bar{\phi}\rangle, \quad (51)$$

where

$$|\phi_I\rangle = |\bar{\phi}\rangle_L \otimes |+I\rangle_S \quad (52)$$

$$|\phi_G\rangle = G(|\bar{\phi}\rangle_L \otimes |+I\rangle_S), \quad (53)$$

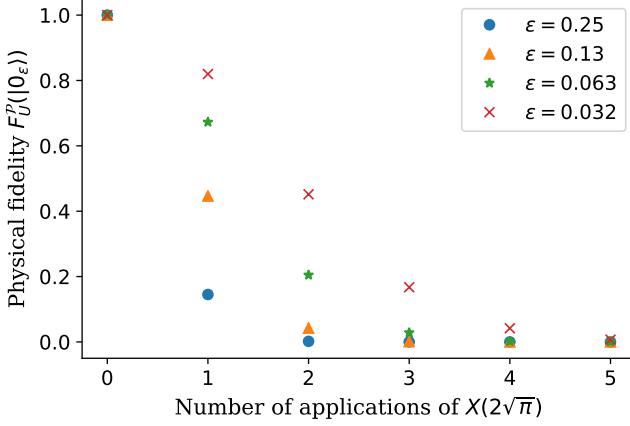


Figure 8. Physical fidelity of the normalizable GKP state $|0_\epsilon\rangle$ after sequential applications of the $X(2\sqrt{\pi})$ gate, which in the ideal case implements a logical identity. The lower the ϵ , the better the approximation to the ideal GKP state, the better the translational symmetry of the state, and the slower the physical fidelity falls as a function of the number of gate applications.

we define the channel

$$\mathcal{E}_{U,G}(\bar{\phi}) \equiv \text{Tr}_S \left[U \frac{G(|\bar{\phi}\rangle_L \langle \bar{\phi}| \otimes |+\rangle_S \langle +|) G^\dagger}{N(G, \bar{\phi})} U^\dagger \right] \quad (54)$$

with normalization factor $N(G, \bar{\phi})$, and the integral is taken over the normalized uniform Haar measure on the two-dimensional state space. We call the quantity (50) the *averaged GKP gate fidelity* (of a physical gate U); it measures the average logical fidelity of the imperfect implementation $U|\phi_G\rangle$ to the perfect implementation $U|\phi_I\rangle$ [68]. Setting $U = \mathbb{1}$ gives us an application-independent metric of the quality of our GKP states. In Fig. 7 we plot $\bar{F}_{U,G}^L(\bar{U})$ for the identity gate and phase gate with the choice of envelope $G = E(\epsilon)$. Varying ϵ produces the expected behaviour.

Similarly, for a fixed \bar{U} , U , and G we can compute an averaged or maximized distribution distance or energy cost over the normalizable states.

2. Displacements in position

A displacement in position by $\sqrt{\pi}$ – half the period of $|0_I\rangle$ – is the standard way to implement a logical \bar{X} operation on GKP-encoded qubits. However, we saw above that a physical X gate acting on a normalizable state quickly lowered the physical fidelity of the initial state while preserving the logical fidelity and the distribution distance. See Fig. 8 for the decay in physical fidelity with the number of gate applications for various envelope strengths. But there are reasons wanting to preserve the physical fidelity. For one, large displacements to a state from the origin require large amounts of energy. This can

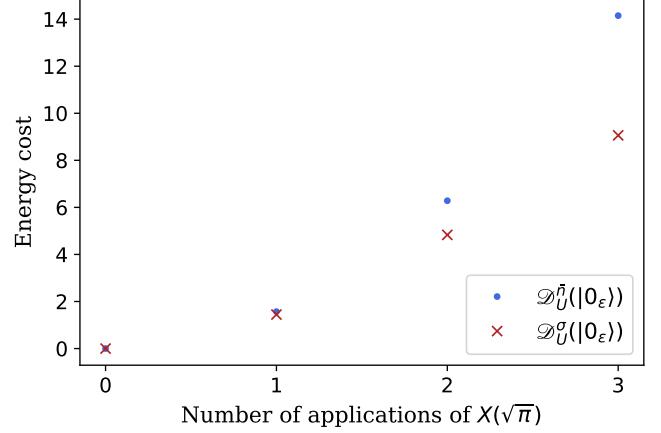


Figure 9. The energy cost mean \mathcal{D}_U^n and standard deviation \mathcal{D}_U^σ of multiple applications of the $X(\sqrt{\pi})$ gate to the state $|0_\epsilon\rangle$ for $\epsilon = 0.063$ (corresponding to a $\Delta \approx 12$ dB). Each subsequent displacement comes at an increased energy cost, and the spread increases more slowly than the mean. Note that an analytic expression for \mathcal{D}_U^n for an arbitrary state $|\psi\rangle$ yields $\frac{\pi}{2}k^2 - \langle \psi | \hat{q} | \psi \rangle k$, where k is the number of gate applications. Therefore applying X gates to zero-mean states will result in the same behaviour.

be seen through the increase in mean and standard deviation energy cost, as shown in Fig. 9. Notably, higher energy can make the states more susceptible to loss. We propose several approaches for dealing with this.

First, we might wish to monitor the changes to the mean of our wavefunction caused by the gates in the circuit. Every Gaussian operation effects the map $\bar{r} \rightarrow \mathbf{S}\bar{r} + \mathbf{d}$ on the vector of means, \bar{r} , where \mathbf{S} is a symplectic matrix and \mathbf{d} is some displacement. Finding the updated mean following an application of k Gaussian unitaries on n modes is therefore equivalent to multiplying k matrices of size $2n \times 2n$. After the circuit has been specified, we can thus classically compute the changes in the mean of our state. With this information in hand, we can then reoptimize the circuit to minimize the maximal displacement. For purely Gaussian circuits we suspect this will be an inexpensive compilation that can be done prior to the computation. The details are left to future research.

Second, we can modify the standard logical-to-physical mapping $\bar{U} \rightarrow U$; a naive approach is through

$$\bar{X} \rightarrow FPPF, \quad (55)$$

where F is the Fourier gate featured in Table I. While this new prescription has no explicit displacements, it requires a larger number of physical operations, including a shearing of position and momentum that will damage the imperfect GKP state in a worse way, as we will see shortly. A better mapping is

$$\bar{X} \rightarrow \begin{cases} X(\sqrt{\pi}) \\ X(-\sqrt{\pi}) \end{cases} \quad (56)$$

This mapping could be *probabilistic*, for example, alternating between the physical gates with probability $\frac{1}{2}$. In this case one randomizes forward and backward displacement by $\sqrt{\pi}$. This is an example of a one-dimensional simple random walk: at the end of the circuit, our state is expected to remain undisplaced with a standard deviation of $\sqrt{n\sqrt{\pi}}$. Roughly, this means that if there is a threshold displacement k after which the computation becomes practically untenable, on average a circuit of depth at most $k^2/\sqrt{\pi}$ can accommodate the computation. A better mapping is *deterministic*, that is, ensuring that a forward displacement is always followed by a backward displacement and vice versa.

As an alternative, we can recalibrate the state after the mean reaches a certain threshold, $\bar{r} = k$, that is, apply a displacement $X(-k)$ at pre-specified points in the circuit. We will discuss combining this step with error correction in Sec. II E.

We see that the new prescriptions for the physical implementation of the \bar{Z} gate maintain both the logical and physical fidelities while minimizing the energy cost. In fact, there is evidence here that demanding that energy costs be minimized generally results in a prescription for the physical implementation of a gate that also improves the other figures of merit.

3. Displacements in momentum

Normally, the logical \bar{Z} is physically realized on GKP states by a $\sqrt{\pi}$ displacement in momentum. The impact of momentum displacements $Z(\sqrt{\pi})$ on the normalizable GKP states will be equivalent to that of the X gate. This can be seen in two different ways: First, the Z gate is the Fourier transform of the X gate, $Z = FXF^\dagger$, and F commutes with the envelope operator $E(\epsilon)$ as they are both exponentials of \hat{n} . Second, the X and Z gate are both linear in the quadrature operators, and so the damage inflicted by the envelope will be on the order of e^ϵ in both cases (both points are verified in App. B 3). Thus the physical fidelity seen in Fig. 8 will be the same in both cases.

As with position, we can monitor the changes to the average momentum and recompile the circuit. We ought also to apply the mapping

$$\bar{Z} \rightarrow \begin{cases} Z(\sqrt{\pi}) \\ Z(-\sqrt{\pi}) \end{cases} \quad (57)$$

deterministically. We do not consider $\bar{Z} \rightarrow PP$, as this prescription requires two applications of the quadratic phase gate and is thus a worse solution.

4. Shearing operations

The logical phase gate \bar{P} is effected on GKP states by the physical phase gate P , which shears the wavefunction

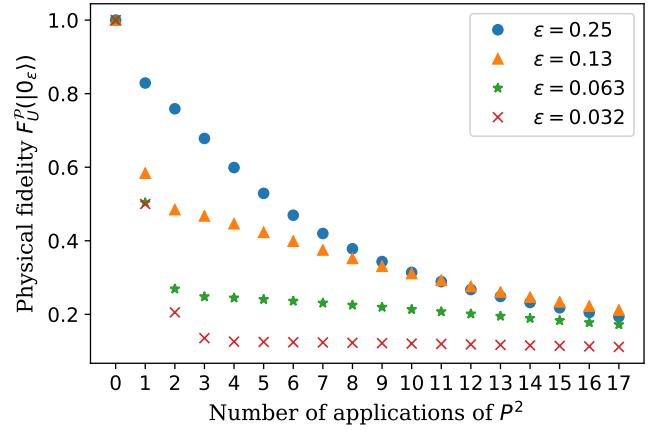


Figure 10. Physical fidelity of normalizable GKP state $|0_\epsilon\rangle$ after sequential applications of P^2 gates for various values of ϵ . In the ideal case this gate implements the logical \bar{Z} , i.e., an identity on the $|0_I\rangle$ state. The physical fidelity falls more slowly when ϵ is higher, that is, when $|0_\epsilon\rangle$ is further from the ideal state. In this regime the state has a lower squeezing and hence a wider peak at $q = 0$, where the phase gate introduces a close-to-trivial phase $e^{i\frac{1}{2}q^2}$.

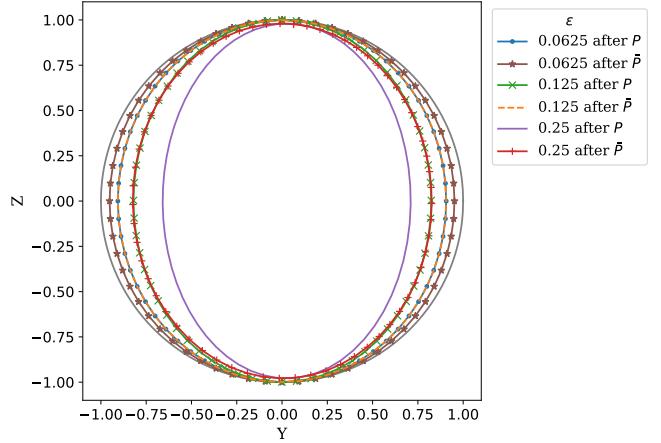


Figure 11. Warping of the logical Bloch sphere after application of the physical phase gate P to the normalizable GKP states $E(\epsilon)|0_I\rangle$, as compared to the action of a true logical phase gate \bar{P} . We show the result for various values of ϵ (for conversion to dB see App. A). We apply the gate to states in the $X - Z$ plane, so the $Y - Z$ plane is depicted since the ideal \bar{P} gate rotates the sphere about the Z axis. Whereas \bar{P} simply applies the rotation to an already distorted Bloch sphere, P additionally warps the Bloch sphere, notably near the equator where states lose more purity.

by introducing a position-dependent phase. But the state $|0_I\rangle$ only has support on $q = 2k\sqrt{\pi}$ for $k \in \mathbb{Z}$, where the phase gate introduces a trivial factor. Since $P^2 = Z(\sqrt{\pi})$, this means that an application of P^n to $|0_I\rangle$ is an identity on both the logical and physical space for any n . On the other hand, the gate P^n applied to $|0_G\rangle$ will introduce a phase for every position value, and in

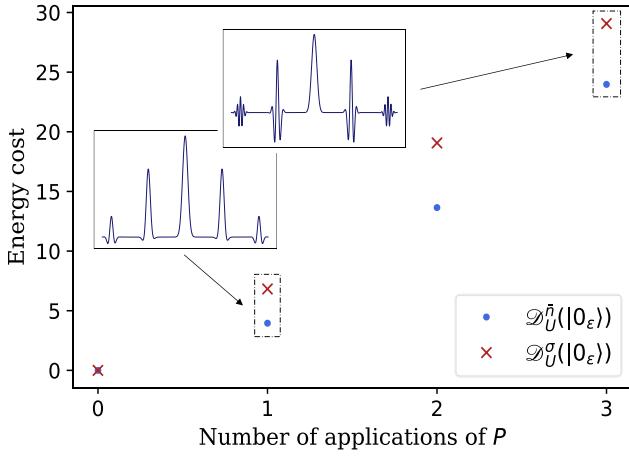


Figure 12. The energy cost mean and standard deviation of successive applications of the P gate to the state $|0_\epsilon\rangle$ for $\epsilon = 0.063$ ($\Delta \approx 12$ dB). In contrast with the X gate in Fig. 9, the spread grows faster than the mean. The analytic expression for $\mathcal{D}_U^{\bar{n}}$ after k applications of the phase gate on the state $|0_\epsilon\rangle$ is $\frac{1}{2}(k^2\langle 0_\epsilon | \hat{q}^2 | 0_\epsilon \rangle + k\langle 0_\epsilon | \{\hat{q}, \hat{p}\} | 0_\epsilon \rangle)$. **Insets:** Highlighted are the wavefunctions after one (left) and three (right) applications of the gate to visualize the damage the gate inflicts on the state in the form of shearing, that is, increasing oscillatory behaviour near the peaks.

particular a complex phase away from $q = k\sqrt{\pi}$. The damage inflicted on the state $E(\epsilon)|0_G\rangle$ will be on the order of $e^{2\epsilon}$ (App. B 3).

From Fig. 10, it looks like better states – those with higher squeezing – actually do worse under the phase gate when considering physical fidelity. This is because states with lower squeezing have a wider central peak, which is located in a region ($q \approx 0$) where the phase gate has little effect. The warping of the Bloch sphere as a result of the phase gate is shown in Fig. 11, where we see that states with higher squeezing better preserve logical information. The mean and standard deviation energy cost are plotted in Fig. 12. As with displacements, we can deal with this problem by noting the flexibility in the logical-to-physical mapping:

$$\bar{P}^k = \begin{cases} \bar{I} & k \equiv 0 \pmod{4} \\ \bar{P} & k \equiv 1 \pmod{4} \\ \bar{Z} & k \equiv 2 \pmod{4} \\ \bar{P}^{-1} & k \equiv 3 \pmod{4} \end{cases} \quad (58)$$

$$\rightarrow \begin{cases} I & k \equiv 0 \pmod{4} \\ P & k \equiv 1 \pmod{4} \\ Z(\sqrt{\pi}) & k \equiv 2 \pmod{4} \\ P^{-1} & k \equiv 3 \pmod{4} \end{cases} \quad (59)$$

Thus, in any circuit recompilation, we ought to treat even applications of \bar{P} as the identity or a \bar{Z} gate, which are easier to implement. In the latter case, we can rely on the prescription for the Z gate we have provided. Similarly, $3 \pmod{4}$ successive applications of the \bar{P} gate should

be replaced with a single application of \bar{P}^{-1} , which corresponds to the physical gate $P^{-1} \equiv e^{-i\frac{1}{2}\hat{q}^2}$.

In analogy with monitoring the position mean, we ought to keep track of how many shearing operations (positive exponents of P) and anti-shearing (negative exponents of P) we have used. For example, if the computation calls for four non-successive applications of \bar{P} , one might consider the physical pattern (P, P, P, P^{-3}) so that the net shear is 0. Another possibility is $(P, Z(\sqrt{\pi})P^{-1}, P, Z(-\sqrt{\pi})P^{-1})$, so that no more than one unit of shear is ever applied to the state. The best optimization strategy for any particular circuit is left to future research.

5. Rotations

The phase gate (a counter-clockwise rotation),

$$R(\phi) = e^{i\frac{\phi}{2}(\hat{q}^2 + \hat{p}^2)}, \quad (60)$$

which implements the logical Hadamard operation whenever $\phi = \pi/2$ (and is called in this case the Fourier gate) commutes with the envelope operator $E(\epsilon)$. Therefore repeated applications of a perfect rotation will not damage the GKP state.

6. Squeezing and beamsplitters

Although they do not effect Clifford operations per se, quadrature squeezers

$$S(z) \equiv e^{\frac{1}{2}(z^* \hat{a}^2 - z \hat{a}^\dagger)} \quad (61)$$

and beamsplitters

$$B(\theta, \phi) \equiv e^{\theta(e^{i\phi}\hat{a}_1\hat{a}_2^\dagger - e^{-i\phi}\hat{a}_1^\dagger\hat{a}_2)}, \quad (62)$$

are used in the implementation of gates, such as the SUM gate below, in the continuous-variable optical domain. For later use, note that $t = \cos\theta$ is the beamsplitter transmissivity. Since the squeezing gate has terms \hat{a}^2 and $\hat{a}^{\dagger 2}$, its impact on normalizable GKP states will be comparable to that of the phase gate. On the other hand, beamsplitter gates have only products of \hat{a} and \hat{a}^\dagger , which means they commute with the envelope operator $E(\epsilon)$ (see App. B 3) and will not be harmful to GKP states, like the rotation gates described above.

7. SUM gates

For universal quantum computation we require at least one kind of two-qubit entangling operation, for example the $\overline{\text{CNOT}}$ gate. For ideal GKP states, this can translate to the continuous-variable SUM gate,

$$\text{SUM}(g) \equiv e^{-ig\hat{q}_1 \otimes \hat{p}_2}, \quad (63)$$

with $g = 1$ being the standard weight. However, one is free to choose any odd-integer weight to effect a $\overline{\text{CNOT}}$:

$$\overline{\text{CNOT}} \rightarrow \text{SUM}(2k+1) \text{ for } k \in \mathbb{Z}. \quad (64)$$

This can be seen by noting that the SUM gate effects the following quadrature transformations:

$$\hat{q}_1 \rightarrow \hat{q}_1 \quad (65)$$

$$\hat{p}_1 \rightarrow \hat{p}_1 - g\hat{p}_2 \quad (66)$$

$$\hat{q}_2 \rightarrow \hat{q}_1 + g\hat{q}_2 \quad (67)$$

$$\hat{p}_2 \rightarrow \hat{p}_2. \quad (68)$$

When the control mode is $|0_I\rangle$, then \hat{p}_1 and \hat{q}_2 are both shifted by an even multiple of $\sqrt{\pi}$ for any integer weight, meaning neither the control nor target mode are changed, as desired. On the other hand, when the control is $|1_I\rangle$, then \hat{p}_1 and \hat{q}_2 are shifted by an odd multiple of $\sqrt{\pi}$ for odd-integer weight; since $|0_I\rangle$ and $|1_I\rangle$ are $\sqrt{\pi}$ -periodic in momentum, this implements a bit flip on the target. Therefore, like for the previous gates, one should update the weight of the SUM gate in a computational circuit depending on the weight of the previously applied SUM gate. To see how this translates to physical requirements, consider the decomposition of the SUM gate into squeezers and beamsplitters:

$$\text{SUM}(g) = B(\pi/2 + \theta, 0) (S(r, 0) \otimes S(-r, 0)) B(\theta, 0), \quad (69)$$

where $\sin(2\theta) = -\text{sech}(r)$, $\cos(2\theta) = \tanh(r)$, and $\sinh(r) = -g/2$. From the discussion above, the harmful element is not the beamsplitter but the squeezer. As expected, r is a monotonic function of g ; and positive values of g correspond to negative values of r .

One benchmark for how the normalizable states perform under the SUM gate is how entangled the modes become in the logical subsystem. Ideally, we have that $\overline{\text{CNOT}}|\bar{+}\rangle|\bar{0}\rangle$ is a maximally-entangled state. In Fig. 13, we plot the entanglement negativity of a two-qubit system in the logical subspace, a well-known measure for determining if two systems are entangled [69]. Entanglement negativity is defined to be the sum of the absolute values of the negative eigenvalues of the partial transpose of a bipartite density matrix with respect to one of the subsystems. For two-qubit systems, negativity ranges from 0 (no entanglement) to 1/2 (maximal entanglement). We initialize the two modes in the $|+\epsilon\rangle|0_\epsilon\rangle$, and then apply the CNOT [SUM(1)] gate. Interestingly, we find a threshold corresponding to ~ 4.5 dB of squeezing is required to produce entanglement in the logical subsystem.

E. Error correction with normalizable GKP states

As a continuous limit of shift-resistant qudit codes, GKP encoding allows one to correct for small displacement errors in the encoded data state. In fact, the GKP

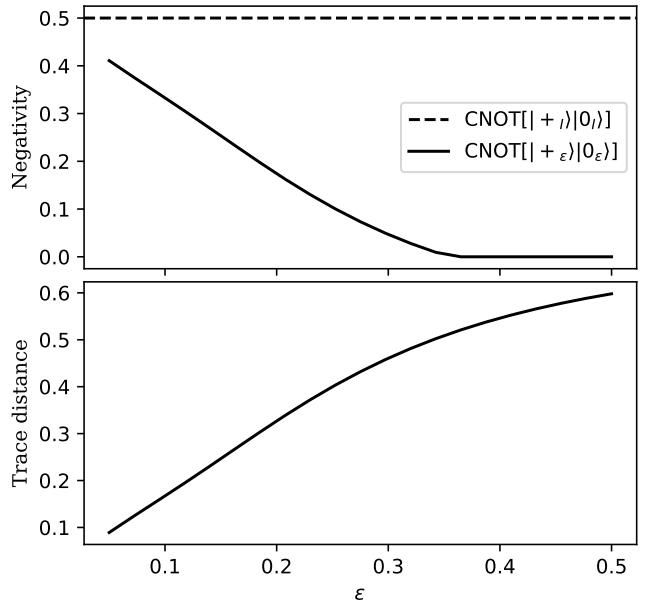


Figure 13. (a) Entanglement negativity (defined in main text) of a two-qubit system initialized in $|+\epsilon\rangle|0_\epsilon\rangle$ and then subjected to the application of a CNOT [SUM(1)] gate, as compared to the ideal (maximal) negativity for perfect qubits. We see that to generate entanglement in the logical subsystem, we require $\epsilon \approx 0.36$, which corresponds to roughly $\Delta \approx 4.5$ dB of squeezing; however, to generate maximal entanglement, significantly more stringent thresholds are required, with $\epsilon \approx 0.05$ only generating 80% of the maximal entanglement negativity. (b) The logical subsystem trace distance between $\text{CNOT}[|+\epsilon\rangle|0_\epsilon\rangle]$ and $\text{CNOT}[|+_I\rangle|0_I\rangle]$. Again, we see that to achieve the ideal distance of 0, smaller values of ϵ are required.

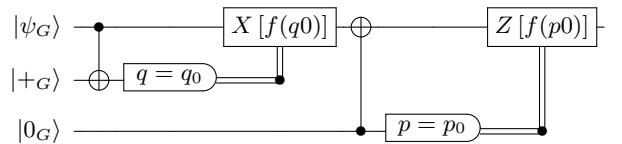


Figure 14. Error syndrome measurement with normalizable GKP states following the Steane approach. First, shifts in q are corrected: an encoded data qubit $|\psi_G\rangle$ and an ancilla $|+_G\rangle$ are sent through a SUM gate, and $|\psi_G\rangle$ is displaced according to the result of a homodyne q measurement on the ancilla. A similar procedure follows for shifts in p .

codes accommodate arbitrary errors on the oscillator, since displacements – i.e., the Weyl-Heisenberg operators $X(\alpha)$ and $Z(\beta)$ – form a complete operator basis. The error syndrome measurement in the Steane approach [70] is shown and described in Figs. 14 and 15 and in the Knill teleportation approach [71, 72] in Fig. 16. The main difference between the two pictures is that, in the former, the data qubit interacts with two separable ancillae, whereas in the latter, the data qubit interacts with

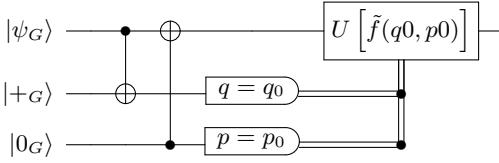


Figure 15. The circuit in Fig. 14 but with the measurements pushed to the end. Now a unitary is applied that encompasses both the correcting position and momentum shifts determined by a new function, \tilde{f} , of the homodyne measurement results q_0 and p_0 . This unitary shifts the projected state back onto the GKP grid. When the error is too big, these shifts are mistakenly performed in the wrong direction, and a logical error results.

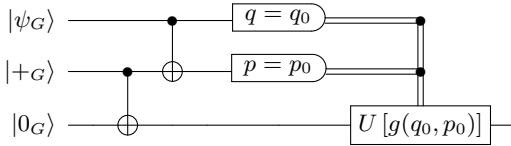


Figure 16. GKP error correction using the Knill approach. Here, the ancillary states $|+G\rangle$ and $|0_G\rangle$ are entangled and then the data state $|\psi_G\rangle$ is sent through a SUM gate with $|+G\rangle$ as the target. Homodyne position and momentum measurements are conducted on $|\psi_G\rangle$ and $|+G\rangle$, respectively; this teleports the logical data to $|0_G\rangle$ up to a unitary, applied to $|0_G\rangle$ depending on the measurement results. Notice that, unlike in the Steane approach (Fig. 15), the output state after the projection is already on the GKP grid. The purpose of this unitary, therefore, is to correct the logical-Pauli byproduct operators that result from the teleportation. When the error is too big, these byproduct operators are misidentified, and a logical error results.

only one of two entangled ancillae. If one can guarantee a supply of high-quality ancillae, the Knill approach could be advantageous, as it precludes two applications of the SUM gate to a noisy data state. There may exist other scenarios involving GKP states in which one or the other circuit is preferable; for our purposes, we focus on the Steane approach.

In the Steane error correction circuit, the SUM gate preserves the position of the data qubit and transforms the ancilla A through $q_A \rightarrow q_\psi + q_A$. This means that a measurement of the ancilla will yield $n\sqrt{\pi} + \delta q_\psi + \delta q_A$ for $n \in \mathbb{N}$, where the δq 's denote the position shift errors. Therefore the function f that ought to be applied to the measurement outcomes before the correcting shift is

$$f(r) = -\text{mod}_{\sqrt{\pi}}(r), \quad (70)$$

where $\text{mod}_t(x) \in [-\frac{t}{2}, \frac{t}{2})$. Whenever δ_ψ and $\delta_A < \frac{\sqrt{\pi}}{2}$, we have that $f(q_{\text{net}}) = -\delta_\psi - \delta_A$, and we have corrected our q displacement errors. However, we have now introduced a new shift error in momentum, since the CNOT gate also effects $p_\psi \rightarrow p_\psi + p_A$. Given a perfect ancilla,

we would be able to correct displacements of magnitude $\sqrt{\pi}/2$ in position and momentum. But the finite energy envelope introduces errors on the ancilla that restrict the range of correctable errors.

In this setting, since we accumulate three errors after a complete round of error correction, Glancy and Knill [2] found that error correction will only be successful if the magnitude of all shifts is less than $\sqrt{\pi}/6$. This ensures that the total error on the data qubit is within correctable region. For a position wavefunction $\psi(q)$ of any of the noisy states, the probability of successful error correction is thus:

$$P_{\text{no error}} = \frac{\pi}{3} \sum_{s,t} \text{sinc}\left(\frac{\pi t}{3}\right) \times \int_{\sqrt{\pi}(2s-\frac{1}{6})}^{\sqrt{\pi}(2s+\frac{1}{6})} du \psi^*(2t\sqrt{\pi} + u)\psi(u). \quad (71)$$

where we provide a short derivation of this expression in App. B 5 and explain its meaning there. Note that although we are assuming approximate GKP ancillae, the formalism just described is general enough to accommodate arbitrary ancillary states, with their usefulness quantified by Eq. (71). For error correction to succeed with high probability, $P_{\text{no error}}$ must be high; this is satisfied by close-to-ideal GKP states. Conversely, a generic state with a high $P_{\text{no error}}$ must have little modular spread in both position and momentum, implying that it approaches the form of an ideal GKP state. In this sense, the Glancy-Knill condition is both necessary and sufficient for correcting displacement errors using the standard Steane and Knill schemes described above. We explore this point further in III C 3.

If during the q (p) error correction the total magnitude of position shifts is greater than $\sqrt{\pi}/2$, the procedure results in a bit-flip (phase flip) error on the logical qubit. These additional errors can be corrected by concatenating the GKP code with standard qubit codes, for example.

Note that one does not need a physical SUM gate to implement GKP error correction; other gates that mix the physical modes can also work. For example, the gate corresponding to a beamsplitter in the optical domain will suffice, in addition to a squeezing gate and an appropriate modification of the function f . This approach is taken in [2]; its benefit is that it is less demanding to implement than the canonical optical CNOT gate, for example, which requires two beamsplitters and two squeezers.

a. Incorporating q and p calibration within error correction. GKP error correction gives information not only about the fractional part of the gauge mode – the small errors that enter haphazardly in the computation – but also the integer part, which is related to the center of the encoded state, the periodicity of the peaks, and the net displacement from the origin of a finite-energy GKP state, as discussed in II D 2 and II D 3. Therefore we might consider “feeding two birds from one hand”

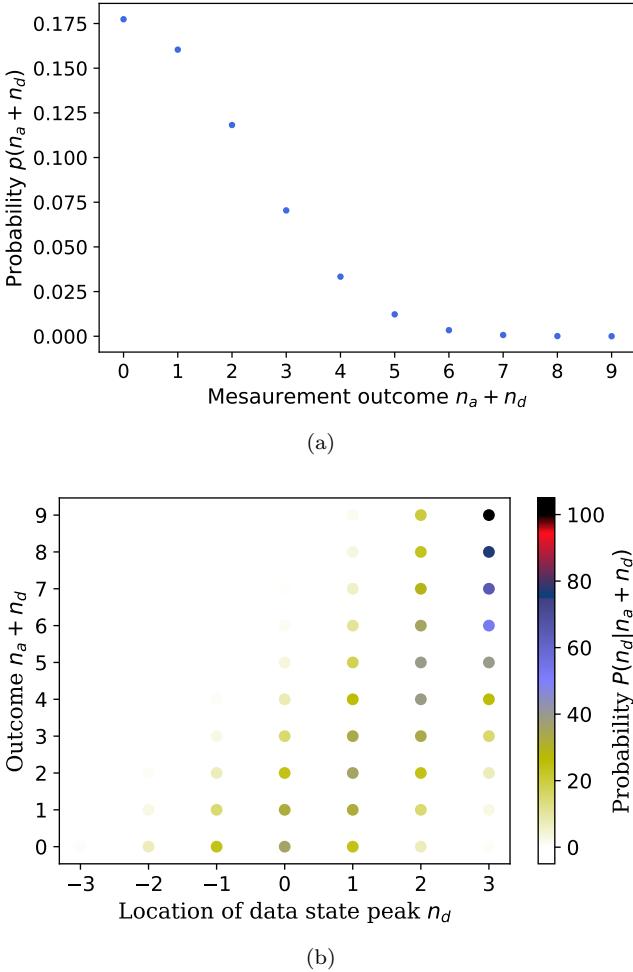


Figure 17. Homodyne measurement probabilities in a successful position error correction step with data state $|0_\epsilon\rangle$ and ancilla $|+\epsilon\rangle$ with $\epsilon = 0.063$ ($\Delta \approx 12$ dB); (a) the probability of the integer part of the measurement outcome equalling $n_a + n_d$, where $n_a\sqrt{\pi}$ and $2n_d\sqrt{\pi}$ are peak locations in the ancilla and data state, respectively, and (b) the probability of the data state having a peak at $q = 2n_d\sqrt{\pi}$ given that the integer part of the measurement outcome was $n_a + n_d$. After several rounds of error correction, we may be able to use data like this to conclude, with some probability, how far the mean of our data state is from 0 so that we may recalibrate it.

and using the outcomes of the homodyne measurements in the error correction procedure to recenter our data state. This would entail modifying the function f above so that the X and Z gates both correct and calibrate the state.

However, we run into a problem: the SUM gate mixes the quadratures of the ancillary and data qubit, limiting the information we can extract about the data state. To see this, suppose that $\delta q_\psi + \delta q_A < \sqrt{\pi}$; then the integer part of the measurement outcome above is $(n_a + n_d)\sqrt{\pi}$, corresponding to the n_d^{th} position peak of the data state and the n_a^{th} position peak of the ancillary state ($n_a, n_d \in \mathbb{Z}$). In the limiting case of an ideal GKP state, no useful

information is obtained, since n_a and n_d could have come from any peaks in the two states with equal probability.

In the normalizable case, we can say a little more, since the probability of finding oneself within a particular peak varies. Consider Fig. 17, where we assume a normalizable data state with effectively seven peaks (so that $-3 \leq n_d \leq 3$) and an ancilla with thirteen peaks ($-6 \leq n_d \leq 6$). We can see in 17b, for example, that if we find ourselves in $n_a + n_d = 9$, then $n_d = 3$ with unity probability. However, not only does this outcome happen with vanishingly small probability (17a), but it also does not necessarily mean that the center of our data state is at $2\sqrt{\pi}$. Nevertheless, one may still use Bayesian reasoning to establish the maximum-likelihood mean of the data qubit after the error-correction step and then use $2n\sqrt{\pi}$ displacements to return that mean to as close to zero as possible.

We note that the above GKP error correction cannot be expected to correct all errors. However, one can aim to construct more sophisticated codes built on the GKP qubits, such as surface codes, which can protect more general errors as considered by a few references mentioned in Sec. I [6, 14–16]. We believe that the modular decomposition picture could play an important role in developing this further.

b. Post-recovery metrics. Having outlined the GKP error correction procedure, we can revisit the metrics in Sec. IID 1 with an additional recovery step. For a unitary U_1 that effects $\bar{U} = I$ on the logical qubit, followed by a recovery operation, R , on a data state together with ancilla, we can compute the updated physical fidelity, logical fidelity, and distribution distance:

$$F_{RU_1}^P(\rho^D) = F\left[\rho^D, \text{Tr}_A\left\{RU_1\rho^{DA}U_1^\dagger R^\dagger\right\}\right] \quad (72)$$

$$F_{RU_1}^L(\rho^D) = F\left[\text{Tr}_g \rho^D, \text{Tr}_{S,A}\left\{RU_1\rho^{DA}U_1^\dagger R^\dagger\right\}\right] \quad (73)$$

$$\mathcal{D}_{RU_1}^P(\rho^D) = \|p_{RU}(x|\psi_G) - p_U(x|\psi_I)\|, \quad (74)$$

where $\rho^D = |\psi^D\rangle\langle\psi^D|$, $\rho^{DA} = \rho^D \otimes |\psi^A\rangle\langle\psi^A|$, and we refer back to Fig. 5 for the definitions of the probability distributions p_{RU} and p_U . The energy cost is a path-dependent measure: it should only increase with an additional recovery step, even if the net change in energy might be smaller. This prompts the post-recovery energy cost definitions

$$\mathcal{D}_{RU}^{\bar{n}}(\rho^D) \equiv |\bar{n}(U\rho^DU^\dagger) - \bar{n}(\rho^D)| + \quad (75)$$

$$|\bar{n}(RU\rho^{DA}U^\dagger R^\dagger) - \bar{n}(U\rho^DU^\dagger)|$$

$$\mathcal{D}_{RU}^{\sigma_n^2}(\rho^D) \equiv |\sigma_n^2(U\rho^DU^\dagger) - \sigma_n^2(\rho^D)| + \quad (76)$$

$$|\sigma_n^2(RU\rho^{DA}U^\dagger R^\dagger) - \sigma_n^2(U\rho^DU^\dagger)|,$$

where now $\bar{n}(\rho) \equiv \text{Tr}(\hat{n}\rho)$ and $\sigma_n^2(\rho) \equiv \text{Tr}(\hat{n}^2\rho) - (\bar{n}(\rho))^2$.

Having introduced various figures of merit to track and understand the way error propagates due to state preparation errors, we now set up the tools to prepare explicit optical circuits to produce the realistic grid states.

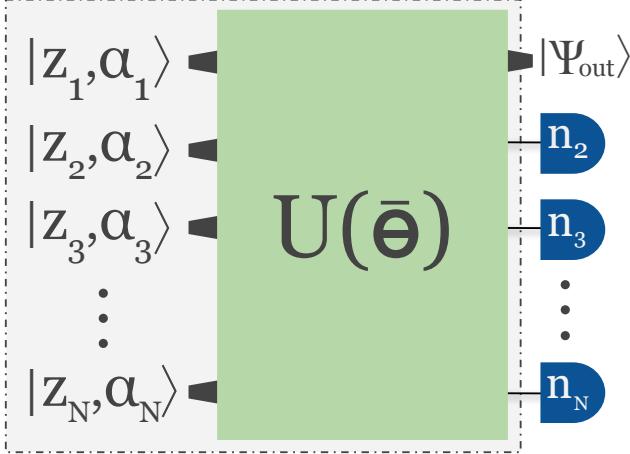


Figure 18. The model for the GBS-like device used for state preparation. Gaussian states consisting of squeezed, displaced vacuum states $|z_i, \alpha_i\rangle$ are sent through an N -mode interferometer parametrized by $U(\bar{\Theta})$, followed by a PNR measurement on all but one of the modes. Given a PNR outcome of \bar{n} , the desired output state, $|\psi_{\text{out}}\rangle$, is produced in the unmeasured mode. Our task is to optimize the circuit components (α , z , and $\bar{\Theta}$) for a given \bar{n} to produce a desired approximate GKP state. We loop over \bar{n} (subject to constraints) to find the best N -mode circuit for the task.

III. PHOTONIC STATE PREPARATION AND CHARACTERIZATION

A. Preparation of non-Gaussian states

GKP states are highly non-Gaussian (see, for example, the trend in Fig. 3), so their preparation requires non-Gaussian resources. For optical platforms, one such resource that is already experimentally accessible is the photon number-resolving (PNR) detector. A Gaussian multimode state can be prepared by applying a general interferometer, $U(\bar{\Theta})$, with N^2 independent beam splitter and phase shift parameters $\bar{\Theta}$, to a multimode input of displaced squeezed vacuum states $D(\alpha)S(z)|0\rangle$. Next, by making PNR measurements of $N - M$ of the N modes, and obtaining results other than zero photon detections across the detectors, one prepares an M -mode non-Gaussian state [47–50]. As the architecture is analogous to Gaussian BosonSampling (GBS), we will refer to the technique as state preparation with GBS devices. The circuit for generating a single-mode non-Gaussian state conditioned on measuring the remaining modes using PNR detectors is depicted in Fig. 18. This general framework encompasses other state preparation schemes, such as photon subtraction and addition (see for example, Fig. 1 of [49]).

When using GBS circuits for state preparation, the number of modes, the initial squeezing and displacement parameters, the interferometer beamsplitter angles, and the PNR measurement patterns can all be

come parameters for tailoring an output state according to a predefined cost function, such as closeness to a given target state and probability of successfully preparing the given state. Extensive analysis of this framework has been performed for the preparation of Fock, cat, NOON and weak-cubic-phase states [47–50], and proof-of-concept calculations for the technique have been made for GKP states [49]. Here, we provide a thorough analysis for preparing GKP states with GBS circuits.

There are a few important results that guided our search for optimal state preparation using GBS circuits. First, if one wishes to prepare a Fock superposition of up to N photons by measuring pure Gaussian states, then the number of photons detected in PNR measurements should sum to N [49, 73]; this allows us to significantly restrict our search over post-selection patterns. Second, Su et. al. [48, 49] conjectured that measuring $(N - 1)$ modes with PNR detectors in a GBS circuit outputs a Fock superposition with at most $(N + 2)(N - 1)/2$ independent coefficients. This implies that, by tuning α , z , and $\bar{\Theta}$ and choosing a suitable photon number post-selection pattern \bar{n} , one can always, in principle, prepare with perfect fidelity a single-mode target state of at most $(N + 2)(N - 1)/2$ photons using an N -mode circuit. For a given target state, this allows us to set an upper bound on the number of modes in the circuit for which we search for optimal α , z , $\bar{\Theta}$ and \bar{n} .

B. Core states for GKP

We now discuss a framework for approximating GKP states. In this section, we describe a formulation of GKP states that is platform-independent, so it can be applied to GKP state preparation in superconducting circuits or ion traps [25–29], for example.

An arbitrary single-mode quantum state can be constructed from a core superposition of Fock states followed by Gaussian operations: squeezing, displacement and rotation [74, 75]. This is sometimes referred to as the stellar representation of the state [75]. An approximation to a given state can be found by truncating the core state with a suitable n_{\max} :

$$|\psi\rangle \approx S(\zeta)D(\beta) \underbrace{\sum_{n=0}^{n_{\max}} \frac{c_n}{N(c, n_{\max})} |n\rangle}_{\text{truncated core state}}, \quad (77)$$

where $N(c, n_{\max})$ is the normalization constant. The exact state can be recovered by taking $n_{\max} \rightarrow \infty$.

The approximate representation (77) is particularly useful when using GBS circuits for state preparation. First, to prepare a target state with a given n_{\max} , we know how many modes and what set of PNR measurement patterns are required to guarantee production of a state with perfect fidelity to the target. Thus, we can find a circuit that optimally produces the truncated core state. Next, as the circuit consists of Gaussian

operations on Gaussian states, the additional Gaussian operations, $S(\zeta)D(\beta)$, that we apply to the core state after it is produced can simply be absorbed into the Gaussian circuit, $U(\Theta)$; moreover, the circuit can be re-decomposed, yielding new $\alpha', z', U(\bar{\Theta}')$ [76]. Thus, operations, such as the squeezing on the core state, are implemented at the start of the circuit, eliminating the need for inline squeezing that is comparatively harder to implement [77]. Finally, if n_{\max} is constrained by the available physical resources, such as the number of circuit modes, by targeting a core state from Eq. (77) rather than a truncation of the Fock expansion of $|\psi\rangle$ directly, we are generally able to attain higher fidelities between the prepared state and $|\psi\rangle$. Some of the extended support in Fock space is captured by the displacement and squeezing of the core state and may be unnecessarily discarded if one truncates the Fock expansion of $|\psi\rangle$.

Remark. The approximate representation (77) is additionally valuable if one wants to prepare GKP states with a different lattice symmetry, such as the hexagonal GKP states. As noted in II A, the hexagonal GKP is related to the square GKP via a symplectic transformation, which can be decomposed into Gaussian operations. Thus, if one has a GBS circuit which can prepare an approximate square GKP state, one can use it to prepare the hexagonal GKP state by appending the Gaussian operations to the end of circuit (or by redecoupling the circuit into a new one).

Before we examine how to prepare approximations to the GKP X , Z and Hadamard eigenstates, let us clarify some nomenclature. *Normalizable GKP states*, reviewed in Sec. II B, are ideal GKP states reduced to a finite energy state by the application of an envelope operator. Our first task is to find an approximation, in the form Eq. (77), to a choice of normalizable GKP states, i.e., for a specific choice of envelope; approximations in the form (77) will be referred to as *approximate GKP states*. Keeping with our notation, we can denote these states as $|\psi_A\rangle$. Our second task is to find a GBS circuit that can optimally prepare the core state corresponding to approximate GKP states, and then to redecouple the circuit to include the Gaussian operations applied to the core state. We call the states output by the final circuit the *circuit GKP states*. If the circuit GKP states, for which we can define a clear experimental prescription, have high enough fidelity to the approximate GKP states, and these states have high enough fidelity to the normalizable GKP states, which for a good enough choice of envelope capture the properties of the ideal GKP states, then the circuit GKP states will share the desired properties of the ideal states.

As they have been the most commonly studied form of normalizable GKP states, we choose the $|\mu_\Delta\rangle$ states with $\mu \in \{0, 1, +, H_+\}$ as the states for which we want to find approximate GKP states (such as $\Delta = \kappa$ case in

Eqs. (8), (9)). We could have also chosen to target (16) with a step function g , that is, some finite cutoff of the ideal GKP states expressed directly in the Fock basis. However, we found that, for equal cutoffs, these states had greater support than the $|\mu_\Delta\rangle$ in regions where the wavefunctions were supposed to vanish.

We note that since the wavefunctions of the above set of $|\mu_\Delta\rangle$ are all real, we do not need to apply complex squeezing to the core state, meaning $\zeta = r \in \mathbb{R}$. As the wavefunctions are also symmetric, we do not need to displace the core state, so $\alpha = 0$, and only the even Fock coefficients will contribute, meaning $c_n = 0$ for n odd. Thus, for a given n_{\max} , we want to find r and c_n (n even) such that the fidelity between the squeezed, truncated core state is maximized with $|\mu_\Delta\rangle$. We summarize our method for finding optimal parameters in Algorithm 1, with additional details provided in App. C 1. We found the `scipy basinhopping` global optimization package [78, 79] to be particularly useful.

Algorithm 1 Optimal Approximate States

```

function cost( $r, c, \Delta, \mu$ )
    initialize  $|\mu_\Delta\rangle$ 
     $|\psi\rangle \leftarrow S(r) \sum_n c_n |n\rangle$ 
    return  $|\langle \psi | \mu_\Delta \rangle|^2$ 
end function

procedure optimization( $\Delta, \mu, n_{\max}$ )
    initialize  $r$  #squeezing within desired range
    initialize  $c$  #normalized vector of dim  $n_{\max}$ 
    # basinhopping is a global search algorithm
     $r_{\text{opt}}, c_{\text{opt}} \leftarrow \text{basinhopping}[(r, c), \text{cost}, \text{args} = (\Delta, \mu)]$ 
end procedure

```

In Fig. 19, we plot the fidelity between the normalizable and approximate GKP states for $\mu = 0$ as a function of Δ from 3 to 11 dB for even values of n_{\max} from 2 to 12 photons. We provide a comment on the near identical results for $n_{\max} = 4$ and 6 in App. D 1. Our results for the other $\mu = 1, +, H_+$ states are available in App. D 2. As can be expected, for a fixed Δ , the fidelity improves monotonically with increasing n_{\max} , and for a fixed n_{\max} the fidelity worsens monotonically with increasing Δ .

C. Characterization of approximate states

We can now look at various properties of the approximate states we have considered, namely the average photon number, orthogonality relations, the Glancy-Knill error correction condition, and behaviour in the modular decomposition. In App. E, we examine even more qualitative features of the approximate states, such as the projectors and quantum error correction matrices.

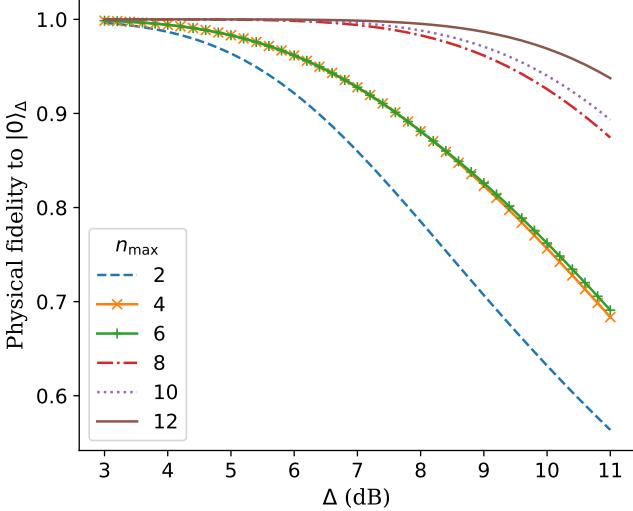


Figure 19. Physical fidelity of our approximate states, $|0_A\rangle$, to the target state, $|0_\Delta\rangle$, as Δ (dB) is varied (see Eq. (8) for the definition of $|0_\Delta\rangle$ and App. A for conversion of dB to other conventions). The $|0_A\rangle$ states are constructed by applying squeezing to a core superposition of Fock states, so different line colours and styles correspond to different values of n_{\max} for the core state. Note that, for different n_{\max} , the optimal squeezing parameter and Fock coefficients are generally different. The optimal $|0_A\rangle$ states for different n_{\max} are found by maximizing the fidelity to $|0_\Delta\rangle$.

1. Average energy

The average energy of the approximate states will have repercussions for the resources required for making the state; for circuit GKP states, this translates to a demand on the initial squeezing applied to each mode. In Fig. 20, we plot the average photon numbers of $|0_A\rangle$ for different n_{\max} as a function of Δ . We see that the average energy of the states is not too high; it increases with Δ and n_{\max} , but never exceeds five photons. We have already shown that increasing n_{\max} is required for producing higher fidelities to the target GKP states, and we know additionally that states with higher Δ values (in dB) provide better error correction and encoding properties; thus, we learn a simple trend for the resources required, even in this method of preparing approximate GKP states: better states require more energy. It is important to remember this trend in addition to the energy costs already established for gate applications (recall Figs. 9 and 12). Average photon number can also be used as a fundamental property to compare various bosonic codes; notably, when considering codes with average photon number less than five, GKP codes were shown to outperform other bosonic codes, including cat and binomial codes, for protecting against loss errors [10].

Remark. In App. D 2, we see that the average photon numbers for the $|1_A\rangle$, $|+_A\rangle$ and $|H_{+A}\rangle$ states are all

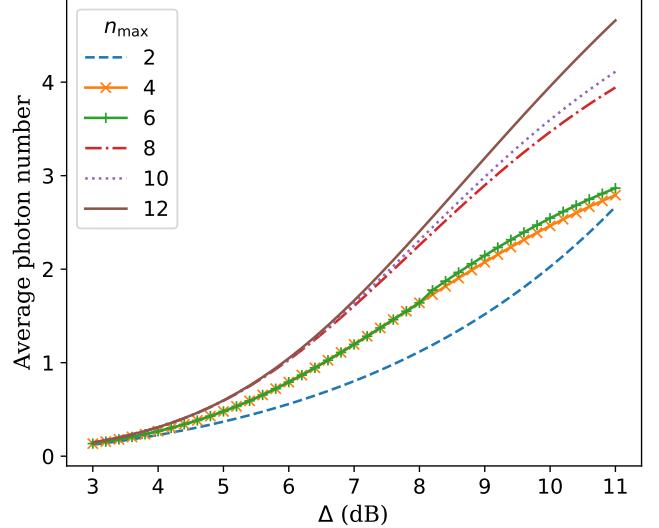


Figure 20. Average photon number of the approximate states $|0_A\rangle$ as a function of Δ (dB), the parameter that characterizes the target state $|0_\Delta\rangle$ that $|0_A\rangle$ is designed to approximate. Line colours/styles reflect different values of n_{\max} for the core state used to construct $|0_A\rangle$. As expected, higher quality states – those with larger Δ in dB – require more energy.

on the same order and follow similar trends. Therefore, given our approach for preparing states with optical circuits, we expect that this will mean preparing different approximate GKP states on the Bloch sphere will require comparable resources, i.e., the same required order of magnitude for squeezing, and the same number of interferometer elements, and PNR detectors (see also Ref. [80]). This means one has the options when designing a computation of only preparing $|0_A\rangle$ states and $|H_{+A}\rangle$ states for non-Clifford gates, or of preparing a collection of states on the Bloch sphere from the outset: the resource requirements will be similar, and the number of gate applications in the circuit will be reduced. In other words, there might be a tradeoff between the number of different state preparation devices and the number of gate elements in the computational circuit.

2. Orthogonality

For the ideal GKP states, the logical 0 and 1 are orthogonal; however, for physically realizable GKP states, $|\psi_G\rangle$, they have nonzero overlap due to the tails of the wavefunction existing outside the bins. We can check the orthogonality of the approximate 0 and 1 GKP states to see how close it is to the orthogonality between the normalizable 0 and 1 states. This is valuable even when the approximate states do not have high fidelity to the normalizable states; if their overlap is small, they can still be used to encode a qubit, although they may not preserve the error-correcting properties of the GKP states nor the simplicity of the canonical gate implementation.

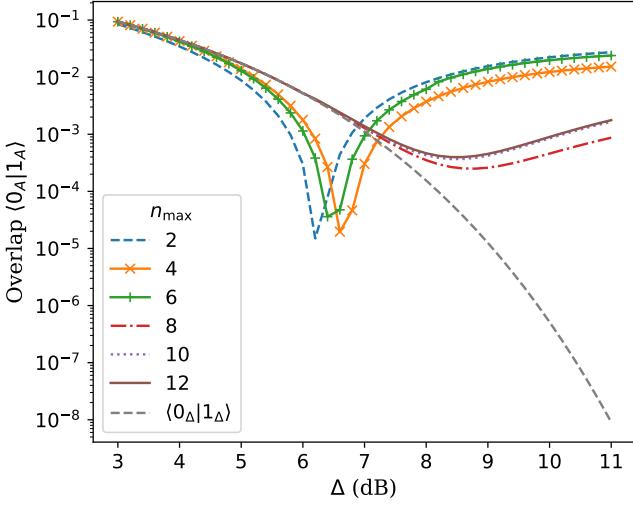


Figure 21. Overlap $\langle 1_A | 0_A \rangle$ between the approximate logical states as a function of Δ (dB), the parameter that characterizes the GKP states that the approximate states are targeting. Line colours/styles reflect different values of n_{\max} for the core state used to construct $|0_A\rangle$. For comparison, we also plot the overlap $\langle 1_\Delta | 0_\Delta \rangle$ of the target states (grey/dashed). A small overlap is a minimal condition for being able to distinguish logical states.

A small overlap is also the most basic necessary condition for error correction, as we will require a low probability of mistaking a 0 for a 1.

In Fig. 21 we plot the overlap $\langle 1_A | 0_A \rangle$ as a function of Δ , for different n_{\max} . We see that all the states roughly follow the trend of the target overlap for low dB values of Δ , but as the squeezing of the target states increases, the approximate states become less orthogonal. This is because as Δ increases, the fidelity gets worse for both $|0_A\rangle$ and $|1_A\rangle$ relative to their target states; specifically, while for larger Δ the $|\mu_\Delta\rangle$ states consist of smooth, narrow peaks, the truncated core of the $|\mu_A\rangle$ states mean that their wavefunctions have additional oscillation between the peaks, where they should instead be close to zero (see Fig. 23 for an example wavefunction). This now increases the overlap $\langle 1_A | 0_A \rangle$ because they gain non-zero contributions at q values between the peak locations.

3. Glancy-Knill property

Using the Glancy-Knill condition in Eq. (71), we compute $P_{\text{no err}}$ for $|0_A\rangle$ with results depicted in Fig. 22. Even though, for example, $|0_A\rangle$ with $n_{\max} = 12$ has higher than 93% fidelity for all Δ considered, this can still translate to a drop in $P_{\text{no err}}$ from a target value of 74% to 58%. One must therefore be cautious in how one constructs approximate GKP states, as their error correcting properties may differ significantly. While physical fidelity depends on the specific choice of targeted finite-energy GKP state, there may be some global properties,

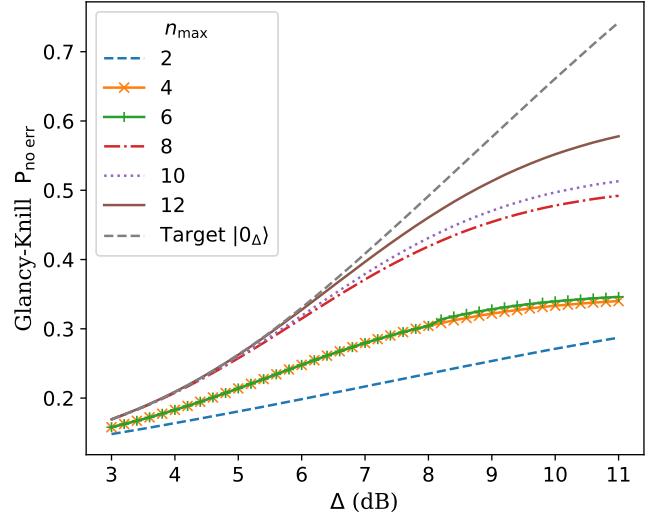


Figure 22. The Glancy-Knill property, $P_{\text{no err}}$, which characterizes the probability a state would not yield a logical error if used as an ancilla for Steane error correction with GKP states (see Sec. II E for definition and details). The line colours/styles reflect $P_{\text{no err}}$ for the various $|0_A\rangle$, each constructed with a different n_{\max} in the core state. We vary $|0_A\rangle$ with Δ , which parametrizes $|0_\Delta\rangle$, the state $|0_A\rangle$ approximates. Additionally, we provide $P_{\text{no err}}$ for $|0_\Delta\rangle$ (grey/dashed). For $\Delta = 11$ dB and $n_{\max} = 12$, we see that, even though $|0_A\rangle$ can have 93% fidelity to $|0_\Delta\rangle$ (see Fig. 19), $P_{\text{no err}}$ drops by 16%.

such as the Glancy-Knill condition or logical fidelity, that are more valuable than the choice of representation.

This motivates examining an additional question: given an n_{\max} , what is the best $P_{\text{no err}}$ one can achieve? This question does not require defining a target state, since the cost function in Algorithm 1 is simply replaced with $P_{\text{no err}}$. We found that, using a core state with $n_{\max} = 12$ and a squeezing of $r \approx 1.87$ dB, we could modestly increase $P_{\text{no err}}$ from 57% to 61%. In Fig. 23, we plot three wavefunctions corresponding to three different values of $P_{\text{no err}}$: $|0_\Delta\rangle$ with $\Delta = 11$ dB, which yields $P_{\text{no err}} = 74\%$; $|0_A\rangle$ with $n_{\max} = 12$ that achieved $P_{\text{no err}} = 57\%$ by maximizing fidelity to $|0_\Delta\rangle$ with $\Delta = 11$ dB; and the state obtained from optimizing directly with respect to $P_{\text{no err}}$ and employing core states with $n_{\max} = 12$, giving $P_{\text{no err}} = 61\%$. One interesting feature of the last wavefunction is that it has more support outside of the logical bins than the $|0_A\rangle$ state we plotted. This means that, although $|0_A\rangle$ is worse for error correction, it is better at encoding information in the logical subsystem of the modular decomposition. An additional issue with only using $P_{\text{no err}}$ as the cost function is that $P_{\text{no err}}$ can be perfectly satisfied by ideal GKP states, so the optimization procedure may be overly demanding and push towards the ideal states which we know to be non-normalizable, while we know that there exist finite energy GKP states that are suitable for computation [3].

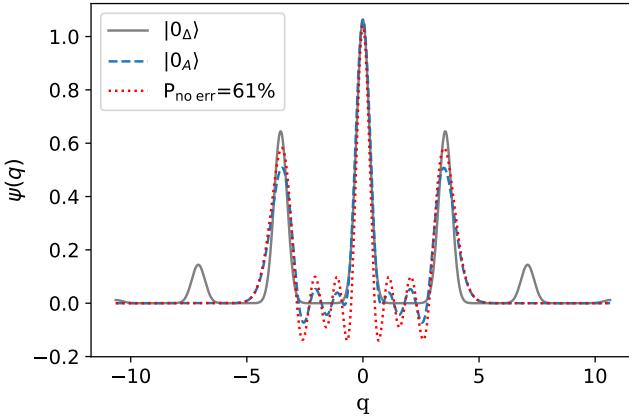


Figure 23. Examples of different normalizable and approximate GKP wavefunctions and their performance under the Glancy-Knill property (71): $|0_\Delta\rangle$ with $\Delta = 11$ dB, which yields $P_{\text{no err}} = 74\%$ (grey/solid); $|0_A\rangle$ with $n_{\text{max}} = 12$ that achieves $P_{\text{no err}} = 57\%$ by maximizing fidelity to $|0_\Delta\rangle$ with $\Delta = 11$ dB (blue/dashed); and the state obtained from optimizing directly with respect to $P_{\text{no err}}$ with core states of $n_{\text{max}} = 12$, giving $P_{\text{no err}} = 61\%$ (red/dotted). Although the red/dotted wavefunction has higher $P_{\text{no err}}$, the blue/dashed wavefunction is better confined to the bin structure of the GKP states, meaning it has better logical encoding when using the modular subsystem decomposition which only depends on the bin structure of the wavefunction.

It should also be noted that the Glancy-Knill condition is specifically a benchmark for using GKP states in error correction in quantum computation. It does not, for example, address the utility of GKP in quantum communication, where the values of Δ required for useful states are much lower. In [10] the authors showed that GKP states worked best for correcting noise resulting from a loss channel. For states with an average photon number of less than two photons, square lattice GKP states with $\Delta = 6.4$ dB were shown to be better than codes designed to correct errors due to loss. For such a Δ , with $n_{\text{max}} = 12$, we found the fidelity of $|0_A\rangle$ to $|0_\Delta\rangle$ to be greater than 99.9%.

4. Logical subsystem Bloch sphere

While the Glancy-Knill condition quantifies the error-correcting capability of the approximate GKP states, examining the logical subsystem in the modular decomposition provides a benchmark for the encoding properties of the states. These are two distinct characteristics: for example, if one takes an ideal GKP state and blurs it such that the delta peaks now become distributions over position, as long as those distributions are confined to the original modular bins of the delta functions of width $\sqrt{\pi}$, then the logical information has not been disturbed, since the results of a binned homodyne measurement will be the same. However, the states might be inadequate

for error correction, since the blurring could easily extend beyond the $\sqrt{\pi}/6$ threshold.

We obtain the logical subsystem state from Eq. (25), except instead of $E(\epsilon)$ the error operators correspond to the envelopes associated with each approximate state. In Fig. 24, we plot the trajectories of the logical subsystems of the $|\mu_A\rangle$ states on the Bloch sphere as the target Δ is varied for different n_{max} . We see that, in some cases, the logical information can be relatively close to the target position on the Bloch sphere even when the fidelity is quite low. For instance, with $n_{\text{max}} = 2$, for $|0_A\rangle$ and $|+_A\rangle$, the approximate states are basically squeezed states in q and p . While this may not compromise the logical information in those states, we know that applications of Gaussian operations will keep the states almost Gaussian, and so we would not expect them to be suitable for universal computation. Combining the Bloch sphere picture with the Glancy-Knill results, we find that the better approximate states are provided, unsurprisingly, by increasing n_{max} and the target Δ .

D. Circuits for GKP state preparation

1. Algorithm for optimal circuits

Given the representation and characterization of approximate GKP states as single-mode Gaussian operations applied to truncated core states, we can now design GBS devices for producing the approximate states. As described in Sec. III A, our state preparation framework is to apply an interferometer, $U(\bar{\Theta})$, to N modes of displaced, squeezed vacuum states, $D(\alpha)S(z)|0\rangle$, perform a PNR measurement on $N - 1$ modes, and postselect on a specific photodetection pattern, \bar{n} , to obtain the target state in the N^{th} mode.

As we summarize in Algorithm 2 (see complete details in App. C2), for a given truncated core state, we train $\alpha, z, \bar{\Theta}$ and \bar{n} for fixed numbers of modes using machine learning algorithms, the strawberryfields [81] and the walrus simulators [82, 83]. Even once we find the optimal circuit for producing the truncated core state, we still have to include the squeezing, $S(r)$; to avoid the need for inline squeezing, we can take the Gaussian unitary $\tilde{U} = S(r)U(\bar{\Theta})D(\alpha)S(z)$ applied to vacuum states on all the modes and express it according to the Euler or Bloch-Messiah [76] decomposition as an equivalent Gaussian unitary of the form $U(\bar{\Theta}')D(\alpha')S(z')$, where now all the squeezing is relegated to the source. This framework for training the circuits is analogous to the process described in [47].

Since the truncated core states only have even Fock coefficients, we do not need to implement the initial displacement on each of the modes. We use the rectangular decomposition from strawberryfields to decompose $U(\bar{\Theta})$ as an N -mode interferometer with $N(N - 1)$ independent beamsplitter and phase-shifter parameters. While a completely general interferometer will also include rotations

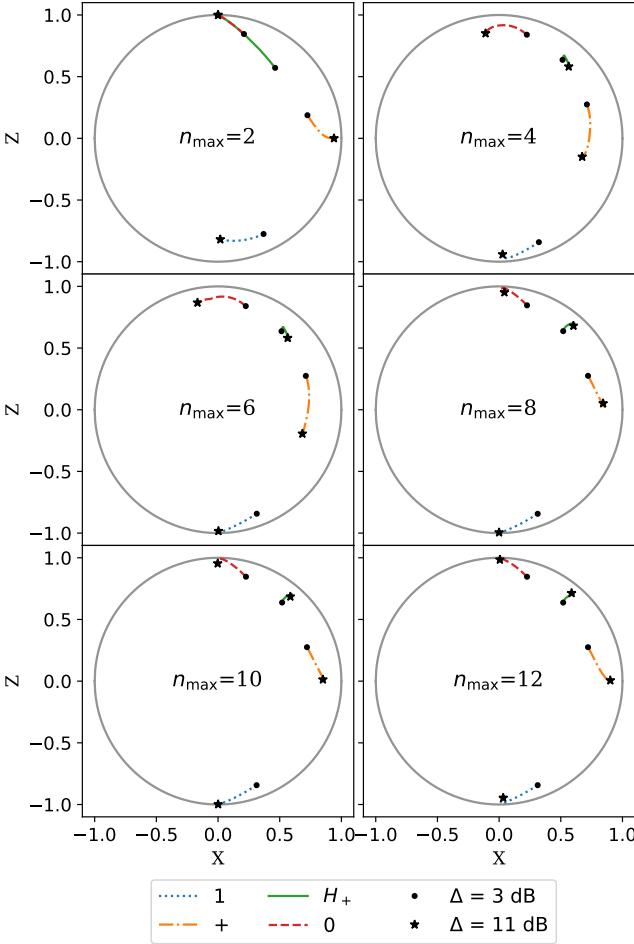


Figure 24. Location of the logical subsystem of the approximate states $|\mu_A\rangle$, where $\mu = 0, 1, +, H_+$, on the X-Z slice of the Bloch sphere. The trajectories are formed as the $|\mu_A\rangle$ vary with Δ , which parametrizes the states $|\mu_\Delta\rangle$ that the $|\mu_A\rangle$ approximate. We denote where the trajectories begin and end using a point for $\Delta = 3$ dB and a star for $\Delta = 11$ dB. We repeat the plot for different choices of n_{\max} for the core states of $|\mu_A\rangle$. While in certain cases $|0_A\rangle$ and $|+_A\rangle$ can be mapped correctly on the Bloch sphere even with $n_{\max} = 2$, if we want all μ to be well-mapped, we need core states with higher n_{\max} .

on each of the modes after the application of all the beam-splitters, since we are performing PNR measurements on $N - 1$ of the modes, we neglect these rotations.

Algorithm 2 Optimal Circuits for Approximate States

```

function cost( $c, \mathbf{x}, \bar{\mathbf{n}}, \text{handle}$ )
     $|\psi_{\text{target}}\rangle \leftarrow \sum_n c_n |n\rangle$ 
    #  $\mathbf{x} = (\mathbf{z}, \Theta)$ 
     $|\psi_{\text{out}}\rangle \leftarrow \langle \bar{\mathbf{n}} | U(\bar{\Theta}) S(\mathbf{z}) | 0\rangle$ 
    prob  $\leftarrow \langle \psi_{\text{out}} | \psi_{\text{out}} \rangle$ 
    fid  $\leftarrow |\langle \psi_{\text{out}} | \psi_{\text{target}} \rangle|^2 / \text{prob}$ 
    if handle = 'global' then
        return fid + 0.1  $\times$  prob
    else if handle = 'local' then
        return fid + prob
    end if
end function

procedure optimization( $c, \bar{\mathbf{n}}, N$ )
    # for an  $N$ -mode circuit
    # random squeezing and BS angles within ranges
    initialize  $\mathbf{x}_0 = (\zeta, \bar{\Theta})$ 
    # basinhopping is a global search algorithm
     $\mathbf{x}_1 \leftarrow \text{basinhopping}[\mathbf{x}_0, \text{cost}, \text{args} = (\bar{\mathbf{n}}, \text{'global'})]$ 
    # local_search is a local maximization algorithm
     $\mathbf{x}_2 \leftarrow \text{local\_search}[\mathbf{x}_1, \text{cost}, \text{args} = (\bar{\mathbf{n}}, \text{'local'})]$ 
end procedure

procedure redecompose_circuit( $r, \mathbf{x}_2$ )
    #  $|\psi_{\text{approx}}\rangle = S(r) \sum_n c_n |n\rangle$ 
    # squeezing only applied to mode with  $|\psi_{\text{out}}\rangle$ 
     $\mathbf{r} \leftarrow (r, 0, \dots, 0)$ 
     $\tilde{U} \leftarrow S(\mathbf{r}) U(\bar{\Theta}) S(\mathbf{z})$ 
    # given a Gaussian unitary applied to vacuum, euler
    returns circuit parameters in the form  $U(\bar{\Theta}') S(\zeta')$ 
     $(\mathbf{z}', \bar{\Theta}') \leftarrow \text{euler}(\tilde{U})$ 
end procedure

```

2. Circuit optimization results

We now employ Algorithm 2 to find the circuits to produce $|0_A\rangle$ for $\Delta = 10$ dB, for $n_{\max} = 4, 8$ and 12 photons. The minimum number of circuit modes we examined was 2, while for $n_{\max} = 4, 8$, and 12 , we examined circuits with up to 3, 4, and 5 modes, respectively. For each circuit, we checked all PNR measurement patterns $\bar{\mathbf{n}}$, such that the number of photodetections summed to n_{\max} . We restricted our squeezing parameter search to, at most, 12 dB of squeezing, which is within the state of the art [84]. In nearly all cases, we found that the optimal fidelities were achieved by saturating the squeezing in at least one of the modes to $z \sim \pm 12$ dB. We can increase the search over larger squeezing values in a straight-forward manner.

In Table II, we provide our results for the best fidelities found using Algorithm 2 for different circuit sizes and values of n_{\max} . Additionally, we provide some other numerical results that still returned fidelities above 99%, but with modestly higher probabilities. Even though increasing N yields more independent parameters to tune to the target state, we see from our results that the increase in fidelity gained beyond three modes is marginal, with the corresponding probability of success vanishing.

In Fig. 25, we plot the Wigner function of $|0_\Delta\rangle$ with $\Delta = 10$ dB, as well as the Wigner functions of the op-

(a) $n_{\max} = 4$			
N	1-(Fidelity to $ 0_A\rangle$)	Probability	\bar{n}
2	0.33*	6.8%	(4)
3	$1 \times 10^{-5}*$	2.1%	(1,3)
	3×10^{-4}	2.2%	(2,2)

(b) $n_{\max} = 8$			
N	1-(Fidelity to $ 0_A\rangle$)	Probability	\bar{n}
2	0.34*	4.7%	(8)
3	$1 \times 10^{-3}*$	0.41%	(4,4)
4	$2 \times 10^{-6}*$	0.14%	(2,2,4)
	5×10^{-6}	0.19%	(1,3,4)

(c) $n_{\max} = 12$			
N	1-(Fidelity to $ 0_A\rangle$)	Probability	\bar{n}
2	0.35*	2.3%	(12)
3	$3 \times 10^{-3}*$	0.11%	(5,7)
4	$4 \times 10^{-8}*$	5.5×10^{-5}	(3,3,6)
	2×10^{-5}	2.3×10^{-4}	(2,4,6)
5	$7 \times 10^{-9}*$	6.5×10^{-5}	(1,1,3,7)
	7×10^{-8}	7.2×10^{-5}	(1,2,3,6)

Table II. Results for the GBS circuits optimized to produce an approximate GKP state $|0_A\rangle$ constructed to approach $|0_\Delta\rangle$ with $\Delta = 10$ dB. We present, as a function of number of circuit modes N , the best fidelities (starred) along with other points of comparably high fidelity and probability found using Algorithm 2, with corresponding probabilities and PNR measurement patterns \bar{n} . We examine $|0_A\rangle$ with core states of n_{\max} = (a) 4, (b) 8, and (c) 12 photons.

timal states (highest fidelity) output by the three-mode circuits designed to produce $|0_A\rangle$ with $n_{\max} = 4, 8$ and 12 photons. These correspond to the starred results for $N = 3$ in Table II. We see that, with increasing n_{\max} – that is, as the core state resource improves – the number and sharpness of peaks approaches to that of $|0_\Delta\rangle$. The difference is smallest near the origin in phase space.

E. Experimental imperfections

1. Stability analysis of optical elements

The stability of the numerically computed optimal circuit parameters is important for experiment since, for example, one might have a given uncertainty in tuning the initial squeezing parameters and beamsplitter angles. As a benchmark of solution stability, we can find the worst-case fidelity within a small region in parameter space

about the optimal solution. To illustrate this, let us take from Table II as an initial guess the optimal solution for a three-mode circuit designed to produce the approximate state $|0_A\rangle$ with a core state of $n_{\max} = 12$. We can then modify Algorithm 2 to minimize the fidelity and to only search within a region set by stability parameters. That is, given the optimal squeezing and beamsplitter parameters, $r_{\text{opt}}, \theta_{\text{opt}}, \phi_{\text{opt}}$, we search for the worst fidelity in a region $r_{\text{opt}} \pm \delta r/r_{\text{opt}}, \theta_{\text{opt}} \pm \delta\theta/\theta_{\text{opt}}, \phi_{\text{opt}} \pm \delta\phi/\phi_{\text{opt}}$. In Fig. 26 we show how much fidelity could change as a function of squeezing stability, $\delta r/r_{\text{opt}}$, and beamsplitter stability, $\delta\theta/\theta_{\text{opt}} = \delta\phi/\phi_{\text{opt}}$. We see that a much larger instability in the squeezing parameters can be tolerated compared to the beamsplitter parameters.

2. Effects of photon loss

In an experimental implementation of a GBS circuit device for state preparation, there will inevitably be loss in the optical components. Here we examine how such loss affects the results of the optimal circuit GKP states. We employ a simple loss model for our circuit: after each squeezer and each beamsplitter (with a complex transmissivity) of the interferometer in the rectangular decomposition, we apply a loss channel with loss parameter η . As the squeezers act on vacuum, they represent the only source of input light, so we capture the effect of lossy sources. Additionally, at the end of each mode, we apply a circuit out-coupling loss, η ; this is followed by loss, also of magnitude η , before each PNR detector to account for detector inefficiency. The loss channel is modelled by coupling a beamsplitter of transmissivity $\sqrt{\eta}$ to an ancillary mode and then tracing out the mode. In Fig. 27, we plot how the fidelity, probability, and Wigner log negativity of the optimal $N = 3$ mode solutions from Table II change as a function of the single loss parameter η . Notably, we see that with increasing n_{\max} , loss becomes increasingly detrimental to fidelity, as it affects higher-photon number components. The probability remains relatively stable, while the Wigner log negativity also decreases with higher loss.

As a proof-of-concept, we also reoptimize some of the circuits in the presence of loss. For the three-mode circuit designed to produce the $n_{\max} = 12$ core state for $\Delta = 10$ dB, we re-ran a modified version of Algorithm 2 to find optimal circuits in the presence of loss. One change we made was to include Wigner log negativity in the cost function to ensure the states had non-Gaussian properties. Additionally, we skip the step of redecoding the circuit, which we discuss in the next paragraph. In Fig. 27, we plot the reoptimized results for $n_{\max} = 12$ for three values of η . We stop at $\eta = 0.06$ dB because by that point we have dropped below the fidelity that can be achieved with only a Gaussian state, a threshold which we indicate with a black dashed line. Although the state still has nonzero Wigner log negativity, being in the regime where Gaussian states are approximating

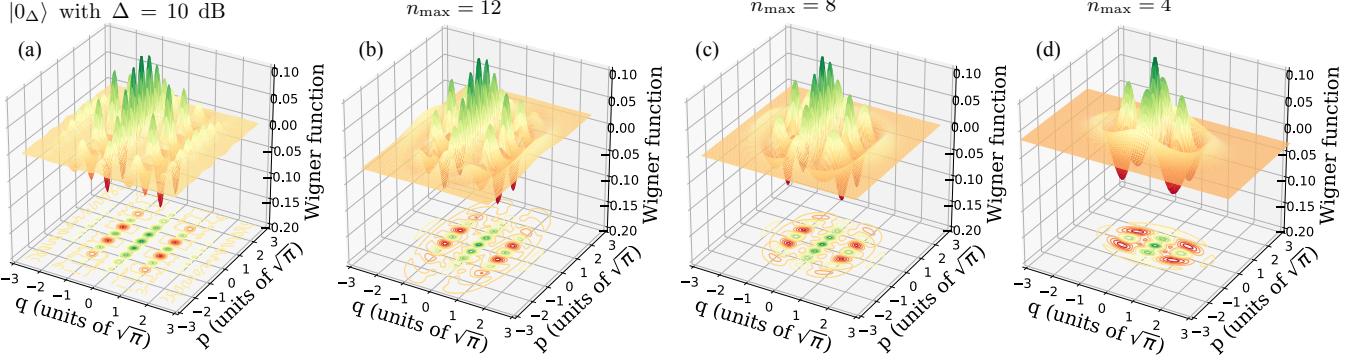


Figure 25. Wigner functions for (a) $|0_{\Delta}\rangle$ with $\Delta = 10$ dB, as well as for the optimal states output by the three-mode GBS devices designed to produce $|0_A\rangle$ with $n_{\max} =$ (b) 12, (c) 8, and (d) 4 photons. These correspond to the starred results for $N = 3$ in Table II. We see the peak structure gets better with increasing n_{\max} , but differences to $|0_{\Delta}\rangle$ are still apparent further from the origin in phase space.

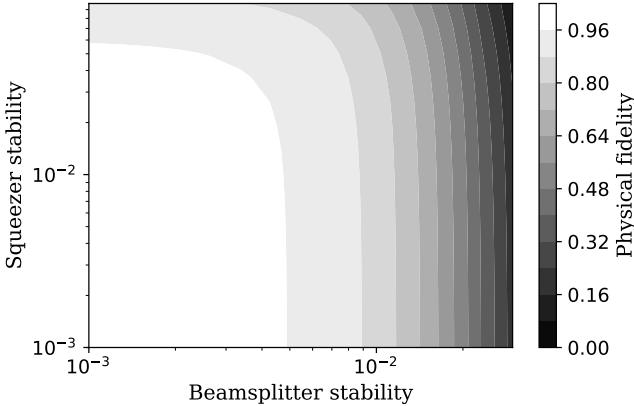


Figure 26. Stability analysis for the optimal 3-mode circuit designed to produce $|0_A\rangle$ with a core state of $n_{\max} = 12$ targeting $|0_{\Delta}\rangle$ with $\Delta = 10$ dB. We find the worst case fidelity between the circuit output and $|0_A\rangle$ as a function of the circuit beamsplitter and squeezer stability parameters, which are defined as the relative error to the ideal parameters.

the state just as well as non-Gaussian ones significantly hampered our search for optimal states.

There are several complications to our strategy for constructing GKP states with GBS devices that arise in the presence of loss. First, as already alluded to, there is the question of how to include the squeezing, r , applied to the core state, since the Euler decomposition is only valid for circuits without loss, so it cannot be applied perfectly in this situation. One option is to employ inline squeezing after the core state is produced by the lossy circuit, but this can be difficult experimentally [85]. Another (suboptimal) strategy is to find the optimal circuit for preparing the core state in the presence of loss, remove the loss channels, add the squeezing r , redecompose the circuit with the Euler decomposition, and then finally reinsert the loss channels. Since the r values applied to the core states of $|0_A\rangle$ with $n_{\max} = 4, 8$, and 12 are rel-

atively small (less than 3 dB), we do not expect a large drop in fidelity using this strategy. For example, we used this strategy for the reoptimized result in Fig. 27 with $\eta = 0.02$ dB, and found that the fidelity only decreased from 82% to 81%.

A second complication to our strategy is the need to revisit the restriction that n_{\max} photodetections ought to be detected to prepare a Fock superposition up to n_{\max} with high fidelity. In the presence of loss, we might expect fewer photodetections are required as some photons could have been lost in the circuit. This leads to many more detection patterns to check, which we leave open to future study.

IV. CONCLUSIONS

In this paper we have investigated, in considerable detail, the methodology for tracking abstract qubit computation encoded into grid states, which are among the earliest and most prominent bosonic codes. A bird's-eye view of our analysis reveals an attempt to understand how realistic and practical considerations affect the quality of the encoded qubit and thereby the information that it carries. Any general qubit computation can be broken down into state preparation, gate implementation, and finally a measurement readout. In the ideal case of a GKP or grid code, the state preparation is the only non-Gaussian resource required. The operations on these ideal states are specific Gaussian gates that implement Clifford unitaries on the encoded qubits, and the Pauli measurements can be realized through homodyne measurements. Our focus is solely on state preparation errors; we assume that our gates and measurements are error-free. We are able to track how an imperfect state preparation manifests itself in the computation and provide ideas around compilation and error mitigation for the physical implementation of the circuits.

To design good error correction protocols, we first need

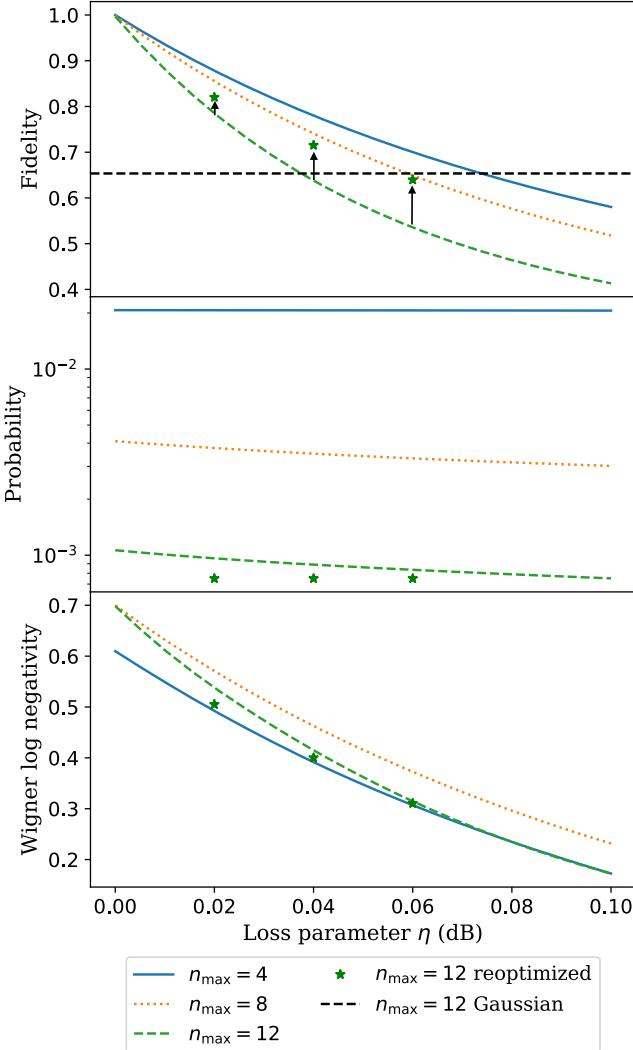


Figure 27. The effect of lossy optical components on the optimal GBS devices for GKP state preparation. We examine how the optimal circuits for $N = 3$ modes from Table II for core states of $n_{\max} = 4, 8, 12$ perform as a function of a single loss parameter, η . Our loss model consists of applying a loss channel parametrized by η to each optical component, as well as to the mode outcoupling and PNR detectors. We plot how the fidelity, probability and Wigner log negativity change as loss increases. Additionally, for a few values of loss, we plot reoptimized results for lossy circuits designed to produce $n_{\max} = 12$ (green stars with arrows indicating which values were reoptimized). We see an increase in fidelity as a result. We stop reoptimization once fidelity is on the same order as that of the Gaussian state with highest fidelity to the target, a threshold we plot with a dashed black line.

to adequately characterize and monitor the errors. To this end we deploy the modular subsystem decomposition, trace out the gauge mode, and follow a logical qubit encoded in a grid state through a computation. The qubit’s journey begins with the ideal, infinite-energy states, which become physical following the application

of an error map; we present several ways this can be done. The qubits then continue their travels through single and two-mode gates. We note that there are many physical operations that lead to the same logical operation on the encoded qubit, meaning a modified real-time mapping or a circuit compilation is necessary to find the best physical gates for the implementation of an abstract qubit circuit. What is meant by “best” is facilitated by the various figures of merit that we introduce in both the physical and logical space, the latter giving us a direct indicator of the information content. Finally, we offer a closer look at the standard Steane-type error correction on GKP qubits, including a discussion of recalibrating the gauge mode alongside an error-correction step.

In the second half of the paper, we focus on obtaining the explicit resources required to produce a class of approximate grid states using a photonic platform. More specifically, we find the fidelity-probability trade-off of producing these single-mode approximate GKP qubits for different circuit sizes, the circuits being measured in all but one of the modes of multimode Gaussian states using PNR detectors. We provide detailed optimization tools for producing the target states and various ways the approximations manifest, such as in the non-orthogonality of the qubits. The stability analysis of these numerical optimization schemes suggest that the states output by the circuit are more robust to uncertainty in squeezing than beamsplitter angles. We furthermore investigate how loss decreases the quality of our approximate GKP states, and show how circuits can be reoptimized in the presence of loss. In addition to looking at the normalizable states through the figures of merit introduced above, we give a few more metrics that could further inform the suitability of these noisy qubits for computation.

In the lead-up to the implementation of useful algorithms on near-term noisy devices, we believe that the analytical and numerical toolboxes that we have developed are valuable in the context of bosonic codes. Since bosonic codes are viewed as promising candidates, especially in the context of lossy circuits, our analysis informs the construction of practical error correction protocols, which, when concatenated with more elaborate codes, would lead ultimately to fault-tolerant computation.

ACKNOWLEDGMENTS

We are grateful to Hoi-Kwong Lo, Saikat Guha, Arne L. Grimsmo, Victor V. Albert, and colleagues at Xanadu for helpful discussions. This work is supported by Mitacs through the Mitacs Accelerate program grant. I. T. is supported by an Ontario Graduate Scholarship. J. E. B. is supported by an Ontario Graduate Scholarship and the Lachlan Gilchrist Fellowship. N. C. M. is supported by the Australian Research Council Centre of Excellence for Quantum Computation and Communication Technology (Project No. CE170100012). All the numerical codes and

Category	Symbol	Description	Eq.
Ideal	$ \psi_I\rangle$	Infinite energy square-lattice GKP states	(6)
Normalized	$ \psi_G\rangle$	$G \psi_I\rangle$; ideal GKP states normalized with operator G	(14)
	$ \psi_\epsilon\rangle$	$ \psi_G\rangle$ with $G = E(\epsilon) \equiv e^{-\epsilon\hat{n}}$	(12)
	$ \psi_{\Delta,\kappa}\rangle$	Gaussians of width Δ enveloped by a Gaussian of width κ^{-1}	(8)
Approximate	$ \psi_\Delta\rangle$	$ \psi_{\Delta,\kappa}\rangle$ with $\Delta = \kappa$; equals $ \psi_\epsilon\rangle$ for small Δ	(11)
	$ \psi_A\rangle$	Approximations to a given choice of normalizable states	(77)

Table III. Notation for various kinds of GKP states referred to in the paper.

Δ_{dB}	ϵ	$\frac{\Delta^3}{24}$
0	1	2×10^{-2}
2	0.631	2×10^{-2}
4	0.398	1×10^{-2}
6	0.251	5×10^{-3}
8	0.158	3×10^{-3}
10	0.100	1×10^{-3}
12	0.0630	7×10^{-4}
15	0.0316	2×10^{-4}
20	0.0100	4×10^{-5}

Table IV. Conversion between selected squeezing values expressed in decibels, corresponding to the width of the peaks in normalizable GKP states $|\psi_\Delta\rangle$ and the epsilon parameter in $|\psi_\epsilon\rangle$ from Eq. (12). The final column contains the error in the approximation between Δ and ϵ , expressed as the magnitude of the third-order term in the expansion of $\tanh \frac{\Delta}{2}$ (see Eq. (13)).

data are available in Ref. [86]. For a distilled discussion of the topics in this paper, see the blog post [87].

Appendix A: Notation, nomenclature, convention, and units

We use this section to disambiguate some of the notation and conventions we use in the paper.

a. *Grid states.* See Table III for the notation we use for the different GKP states we refer to in the paper.

b. *Logical and physical gates.* Gates with a bar, as in \bar{U} , always refer to logical gates acting on qubits, and gates without a bar, as in U , are always physical gates acting on the oscillator space.

c. *Squeezing.* The convention for the squeezing parameter r we use is such that the effect of the squeezing gate on the quadrature operators is

$$S^\dagger(r)\hat{q}S(r) = e^{-r}\hat{q}, \quad S^\dagger(r)\hat{p}S(r) = e^r\hat{p}, \quad (\text{A1})$$

meaning that, in the position representation,

$$S(r)\psi(q) = e^{r/2}\psi(e^r q). \quad (\text{A2})$$

Squeezing values are often expressed in decibels in this paper. The conversion we use for the parameter in the squeezing gate is

$$r = \frac{\ln 10}{20}r_{\text{dB}}, \quad (\text{A3})$$

and for the width Δ of the normalizable GKP state peaks, it is

$$\Delta = 10^{-\frac{\Delta_{\text{dB}}}{20}}. \quad (\text{A4})$$

Although we often work in the regime $\Delta = \kappa$, it is κ^{-1} that is the width of the overall envelope in the $|\psi_{\Delta,\kappa}\rangle$ states. Therefore we are interested in low values of both Δ and κ (large positive in dB).

Lastly, we often use the envelope operator $E(\epsilon)$ from Eq. (12) to express our normalizable states. In the regimes we consider, it is true that

$$\epsilon \approx \Delta^2, \quad (\text{A5})$$

meaning

$$\epsilon \approx 10^{-\frac{\Delta_{\text{dB}}}{10}}. \quad (\text{A6})$$

In Eq. IV we list the conversion between a few squeezing values expressed in Δ_{dB} and ϵ , as well as the error in the approximation.

Appendix B: Derivations

1. Dilation for the envelope operator $E(\epsilon)$

As stated in Sec. II B, a convenient method for constructing normalizable GKP states is by applying the operator $e^{-\epsilon\hat{n}}$ to the ideal GKP states. According to [13], this can be constructed as follows: first pass an ideal GKP state through a beamsplitter of transmissivity t with a vacuum state:

$$|\psi_I\rangle|vac\rangle = \int \frac{d^2\beta}{\pi} \langle \beta|\psi_I\rangle |\beta\rangle|vac\rangle \xrightarrow{\text{BS}} \int \frac{d^2\beta}{\pi} \langle \beta|\psi_I\rangle |t\beta\rangle|r\beta\rangle \quad (\text{B1})$$

Then, post-select on measuring the second mode in the vacuum state to obtain

$$\begin{aligned}
& \int \frac{d^2\beta}{\pi} e^{-|r\beta|^2/2} \langle \beta | \psi_I \rangle |t\beta\rangle \\
&= \int \frac{d^2\beta}{\pi} \langle \beta | \psi_I \rangle \sum_n e^{-|\beta|^2/2} \frac{\beta^n t^n}{n!} |n\rangle \quad (\text{B2}) \\
&= t^{\hat{n}} |\psi_I\rangle,
\end{aligned}$$

Finally, make the association $e^{-\epsilon} \rightarrow t$.

2. Readout from the logical subsystem

A measurement in the computational basis of a qubit can be effected with a binned homodyne measurement of the q quadrature of the corresponding GKP state. Here we show that this translates easily to the subsystem picture, meaning binned homodyne measurements allow direct access to the logical subsystem.

Just as with a measurement in the computational basis, a binned homodyne measurement must yield a binary output. Naturally, for the square lattice, the union of bins of width $\sqrt{\pi}$ centered at $2n\sqrt{\pi}$ ($(2n+1)\sqrt{\pi}$) corresponds to binary output 0 (1). We note that the POVMs corresponding to these bins can naturally be written in terms of displacements of the GKP states:

$$M_\mu = \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} d\beta X(\beta) |\mu_I\rangle \langle \mu_I| X^\dagger(\beta); \quad \mu = 0, 1; \beta \in \mathbb{R}. \quad (\text{B3})$$

Recall the decomposition from Eq. (23), and note that

$$X(\beta) |\mu_I\rangle = \begin{cases} |\mu\rangle_{\mathcal{L}} \otimes X(\beta) |+_I\rangle_{\mathcal{G}}, & \beta \in 0 \text{ bins}, \\ \bar{X} |\mu\rangle_{\mathcal{L}} \otimes X(\beta) |+_I\rangle_{\mathcal{G}}, & \beta \in 1 \text{ bins}, \end{cases} \quad (\text{B4})$$

taking care to realize that this is not a decomposition of $X(\beta)$ but only a statement about its action on ideal GKP states. Here, the 0 (1) bins refer to the q quadrature region of $[(2n - \frac{1}{2})\sqrt{\pi}, (2n + \frac{1}{2})\sqrt{\pi}]$ ($[(2n + \frac{1}{2})\sqrt{\pi}, (2n + \frac{3}{2})\sqrt{\pi}]$).

Thus, Eq. (B3) can be rewritten as:

$$\begin{aligned}
M_\mu &= |\mu\rangle \langle \mu|_{\mathcal{L}} \otimes \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} d\beta X(\beta) |+_I\rangle \langle +_I|_{\mathcal{G}} X^\dagger(\beta) \quad (\text{B5}) \\
&= |\mu\rangle \langle \mu|_{\mathcal{L}} \otimes \mathbb{1}_{\mathcal{G}},
\end{aligned}$$

which is exactly a measurement on the logical subsystem in the computational basis.

3. Conjugation and commutation with $E(\epsilon)$

In Table V, we display how the envelope operator $E(\epsilon)$ from (12) transforms under conjugation with Gaussian operations through $E(\epsilon) \rightarrow UE(\epsilon)U^\dagger$. In addition to this we might wish to see how the operations themselves

U	$\tilde{E}(\epsilon) = UE(\epsilon)U^\dagger$
$X(\alpha)$	$e^{-\frac{\epsilon}{2}[(\hat{q}-\alpha)^2+\hat{p}^2]}$
$Z(\alpha)$	$e^{-\frac{\epsilon}{2}[\hat{q}^2+(\hat{p}-\alpha)^2]}$
$P(s)$	$e^{-\frac{\epsilon}{2}[\hat{q}^2+(\hat{p}-s\hat{q})^2]}$
$R(\phi)$	$E(\epsilon)$
SUM(g)	$e^{-\frac{\epsilon}{2}[\hat{q}_1^2+(\hat{p}_1-g\hat{p}_2)^2]}e^{-\frac{\epsilon}{2}[(g\hat{q}_1+\hat{q}_2)^2+\hat{p}_2^2]}$
$S(r)$	$e^{-\frac{\epsilon}{2}[e^{2r}\hat{q}^2+e^{-2r}\hat{p}^2]}$
$B(\theta, \phi)$	$e^{-\epsilon\hat{n}_1}e^{-\epsilon\hat{n}_2}$

Table V. Physical operators U and conjugated envelope operator $\tilde{E}(\epsilon) = UE(\epsilon)U^\dagger$ introduced in Eq. 12.

change: $U \rightarrow E(-\epsilon)UE(\epsilon)$. Note that there are some mathematical difficulties in working with $E(\epsilon)$ that one should be wary of, since it is an exponential of an unbounded operator. For a discussion and a rigorous treatment of some the issues that arise, see, for example, [88].

To derive the commutation relations, we first show that, for any $k \in \mathbb{N}$,

$$[\hat{n}, \hat{a}^k] = -k\hat{a}^k \quad (\text{B6})$$

$$[\hat{n}, \hat{a}^{\dagger k}] = k\hat{a}^{\dagger k}. \quad (\text{B7})$$

For this, proceed by induction. First, using the identity $[AB, C] = A[B, C] + [A, C]B$, we can see that

$$[\hat{n}, \hat{a}] = [\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger \overbrace{[\hat{a}, \hat{a}]}^0 + \overbrace{[\hat{a}^\dagger, \hat{a}]}^{-1} \hat{a} = -\hat{a}. \quad (\text{B8})$$

Now choose $j \in \mathbb{N}$, and assume $[\hat{n}, \hat{a}^{j-1}] = -a^{j-1}$. In light of the related identity $[A, BC] = [A, B]C + B[A, C]$, we have

$$[\hat{n}, \hat{a}^j] = [\hat{n}, \hat{a}^{j-1}\hat{a}] \quad (\text{B9})$$

$$= [\hat{n}, \hat{a}^{j-1}] \hat{a} + \hat{a}^{j-1} [\hat{n}, \hat{a}] \quad (\text{B10})$$

$$= -(j-1) \hat{a}^{j-1}\hat{a} + \hat{a}^{j-1} (-\hat{a}) \quad (\text{B11})$$

$$= -j\hat{a}^j. \quad (\text{B12})$$

To show $[\hat{n}, \hat{a}^{\dagger k}] = \hat{a}^{\dagger k}$, simply take the Hermitian conjugate of both sides of (B6).

Let us now rewrite these commutation relations in a more helpful form:

$$[r\hat{n}, s\hat{a}^k] = (-kr)(s\hat{a}^k) \quad (\text{B13})$$

$$[r\hat{n}, t\hat{a}^{\dagger k}] = (kr)(t\hat{a}^{\dagger k}). \quad (\text{B14})$$

For relations that look like this, that is $[X, Y] = bY$, there is a braiding identity

$$e^X e^Y = e^{e^b Y} e^X. \quad (\text{B15})$$

Thus we may write

$$e^{r\hat{n}} e^{s\hat{a}^k} = e^{e^{-kr}s\hat{a}^k} e^{r\hat{n}} \quad (\text{B16})$$

$$\implies e^{r\hat{n}} e^{s\hat{a}^k} e^{-r\hat{n}} = e^{s(e^{-r}\hat{a})^k} \quad (\text{B17})$$

$$e^{r\hat{n}} e^{t\hat{a}^{\dagger k}} = e^{e^{kr}t\hat{a}^{\dagger k}} e^{r\hat{n}} \quad (\text{B18})$$

$$\implies e^{r\hat{n}} e^{t\hat{a}^{\dagger k}} e^{-r\hat{n}} = e^{t(e^{r}\hat{a}^{\dagger})^k}. \quad (\text{B19})$$

Now consider an arbitrary single-mode operator of the form

$$U = e^{p_k(\hat{a}^{\dagger}, \hat{a})},$$

where p_k is a k -th degree polynomial. We can insert $E(-\epsilon)E(\epsilon) = \mathbb{1}$ between any two operators in this polynomial, implying that

$$E(-\epsilon)p_k(\hat{a}^{\dagger}, \hat{a})E(\epsilon) \quad (\text{B20})$$

$$= p_k [E(-\epsilon)\hat{a}^{\dagger}E(\epsilon), E(-\epsilon)\hat{a}E(\epsilon)] \quad (\text{B21})$$

$$= p_k (e^{\epsilon}\hat{a}, e^{-\epsilon}\hat{a}^{\dagger}). \quad (\text{B22})$$

Again using the fact $E(-\epsilon)E(\epsilon) = \mathbb{1}$, we see that

$$E(-\epsilon)e^{p_k(\hat{a}^{\dagger}, \hat{a})}E(\epsilon) = e^{E(-\epsilon)p_k(\hat{a}^{\dagger}, \hat{a})E(\epsilon)} \quad (\text{B23})$$

$$= e^{p_k(e^{\epsilon}\hat{a}^{\dagger}, e^{-\epsilon}\hat{a})}. \quad (\text{B24})$$

Therefore we conclude that the envelope conjugates across any operator in the form (B24) – an exponential of an arbitrary polynomial of the creation and annihilation operators – at the expense of changing the inputs via

$$\hat{a} \rightarrow e^{-\epsilon}\hat{a} \quad (\text{B25})$$

$$\hat{a}^{\dagger} \rightarrow e^{\epsilon}\hat{a}^{\dagger}. \quad (\text{B26})$$

4. Measures of non-Gaussianity

One way to characterize genuine non-Gaussianity of a quantum state is through the negativity of the Wigner function [89, 90]. The Wigner function is a phase space quasiprobability distribution defined through

$$W_{\rho}(q, p) = \frac{1}{\pi} \int_{-\infty}^{\infty} \langle q + x | \rho | q - x \rangle e^{-2ipx} dx, \quad (\text{B27})$$

and the Wigner negativity is the area of the negative part of the Wigner function:

$$W(\rho) = \int dq dp |W_{\rho}(q, p)| - 1. \quad (\text{B28})$$

The Wigner logarithmic negativity is then

$$W_N(\rho) = \log [W(\rho) + 1]. \quad (\text{B29})$$

Although pure states with a vanishing Wigner negativity must be Gaussian, there exist non-Gaussian mixed states with a positive Wigner function [91].

5. Glancy-Knill probability of no error

In [2], the authors find that GKP error correction can be performed perfectly with ideal GKP states displaced by less than $\sqrt{\pi}/6$. Thus, to determine the probability of successful error correction with normalizable GKP states, they first expand an arbitrary oscillator state in a basis of displaced ideal GKP states:

$$|\psi\rangle = \int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} dv \langle u, v | \psi \rangle |u, v\rangle \quad (\text{B30})$$

where $|u, v\rangle = \pi^{-1/4} e^{-iup} e^{-iv\hat{x}} |0_I\rangle$ is reminiscent of the modular basis later explored in [21] and $\pi^{-1/4}$ is a normalization factor that ensures completeness:

$$\int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} dv |u, v\rangle \langle u, v| = \mathbb{1}. \quad (\text{B31})$$

Then, the probability of finding $|\psi\rangle$ within a displaced region of less than $\sqrt{\pi}/6$, so that the errors introduced by the ancilla are small enough for the error correction to go through, is given by

$$P_{\text{no error}} = \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv |\langle u, v | \psi \rangle|^2. \quad (\text{B32})$$

Expanding, we find

$$P_{\text{no error}} = \pi^{-1/2} \sum_{s,t=-\infty}^{\infty} \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv e^{2iv(s-t)\sqrt{\pi}} \times \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \psi^*(2t\sqrt{\pi} + u) \psi(2s\sqrt{\pi} + u). \quad (\text{B33})$$

Performing the integral over v and taking $t \rightarrow t+s$ and $u \rightarrow u - 2s\sqrt{\pi}$, we recover Eq. (71). Eq. (B33) can be interpreted as follows: For $s = t$, one gets the probability that $\psi(x)$ lies within $\frac{\sqrt{\pi}}{6}$ of all integer multiples of $2\sqrt{\pi}$ in position space, that is, where $|0_I\rangle$ has support. For $s \neq t$, one gets the probability that $\psi(x)$ lies within $\frac{\sqrt{\pi}}{6}$ of integer multiples of $2\sqrt{\pi}$ in momentum, without overcounting the correctable position regions.

It is straightforward to extend the formula to error correction with arbitrary mixed states:

$$P_{\text{no error}} = \frac{\pi}{3} \sum_{s,t} \text{sinc}\left(\frac{\pi t}{3}\right) \times \int_{\sqrt{\pi}(2s-\frac{1}{6})}^{\sqrt{\pi}(2s+\frac{1}{6})} du \rho(u, 2t\sqrt{\pi} + u). \quad (\text{B34})$$

where $\rho(x, x')$ is the density matrix in the position basis.

Appendix C: Numerical techniques

In addition to the `strawberryfields` and the `walrus` libraries, we employed packages from `scipy` libraries for

special functions, numerical integration, and optimization algorithm implementations. Here we present details of our implementation of the algorithms.

1. Algorithm 1 details

To initialize the normalizable state $|\mu_\Delta\rangle$, we build a numerical wavefunction on a discretized position space as follows:

1. To be safe, we take 7 standard deviations of $1/\Delta$ as the range in q space so that the heights of the peaks at the edge of the range will be negligible.
2. We solve for the integer number of peaks separated by $\alpha = \sqrt{\pi}$ that fit into that range, and choose the number of points in the discretization of the range of q space to be 100 times the number of peaks, so that each peak is well-resolved.
3. With the array of q values in hand, we build the $|\mu_\Delta\rangle$ wavefunction by summing together Gaussians of width Δ centred at each of the $n\sqrt{\pi}$ within the range of q , and weighted according to the value of $\mu \in \{0, 1, +, H_+\}$.
4. We numerically integrate the function using the `numpy trapz` function to determine the normalization factor.

We define the `cost` function to be the fidelity between the target state $|\mu_\Delta\rangle$ and the approximate state of the form of a squeezed finite superposition of Fock states. We construct a numerical wavefunction for $|\psi\rangle = S(r) \sum_n c_n |n\rangle$. Using the same q array as the $|\mu_\Delta\rangle$ wavefunction, we sum the weighted q space wavefunctions for the Fock states (built using the `scipy eval_hermite` function). Then, we apply the squeezing using the convention from Eq. (A2). The fidelity, $|\langle \mu_\Delta | \psi \rangle|^2$, is evaluated as a numerical integration with the `numpy trapz` function.

In the optimization algorithm, we employ the `basinhopping` algorithm available from the `scipy optimize` library. An overview of this algorithm is available in [47], with a more detailed description available in the library documentation and references therein [78, 79]. Broadly, `basinhopping` consists of two phases:

1. A local minimization (we minimize the negative of the `cost` function) over the optimization parameters (r, \mathbf{c}) given an initial guess, performed using the sequential least-squares programming method. We impose the constraint that \mathbf{c} needs to be normalized and only has even components up to n_{\max} . This step produces a candidate for the global optimum (not a true global optimum).
2. A stochastic “hop” is performed in parameter space once a local minimum is found. After the hop is made, the local minimization phase is repeated with the new point in parameter space as the initial

guess. The candidate global optimum is updated based on an acceptance test. We found 40 hops to be a suitable number.

When the algorithm terminates we are provided with an optimal squeezing parameter, r_{opt} , and a vector of coefficients, \mathbf{c}_{opt} , consistent with our choice of n_{\max} .

Importantly, we find that `basinhopping` is helpful: in some cases the optimal squeezing parameter and the truncated coefficients of the core state do not vary smoothly as we change Δ , even though the fidelity appears to be changing smoothly. That is, two or more different regions of parameter space can yield comparable fidelities, so if one region becomes better than another there is an apparent jump in parameter space. The global properties of the approximate states, such as the fidelity and the average photon number, still vary smoothly because the interplay between the squeezing and the core state coefficients combine to yield similar final states even if the optimal squeezing parameters and coefficients are undergoing discontinuous jumps. Put differently, even though the stellar representation [75] of a state is unique, when we apply a truncation to the core state, there can be many choices for states that meet a fidelity threshold to the target state. Thus, when one chooses the state with the highest fidelity among a collection of states that meet a fidelity threshold, the states can come from very different regions of parameter space even if Δ is only changing slightly.

To speed up the search for optimal parameters, for a given μ and n_{\max} , we pass the optimal results from the previous value of Δ as the initial guess for the next value. Since the `basinhopping` global search algorithm employs random jumps in parameter space, the initial guess does not exclude finding better local optima in other regions of parameter space.

2. Algorithm 2 details

To run Algorithm 2 for finding the optimal circuit-produced state relative to an approximate state $|\mu_A\rangle$, we supply the squeezing parameter, r ; the core state Fock coefficients, \mathbf{c} , of $|\mu_A\rangle$; the number of modes for the GBS device; and a post-selection pattern, $\bar{\mathbf{n}}$, consistent with the need for the number of photodetections in $\bar{\mathbf{n}}$ to total the n_{\max} of the core state. We loop Algorithm 2 over all choices of $\bar{\mathbf{n}}$ up to permutations of modes, as permutations can be absorbed into the interferometer.

There are three essential parts of Algorithm 2. First, we define a `cost` function which is called in the `optimization` procedure. Second, the `optimization` procedure returns the optimal circuit parameters (relative to the `cost` function) for producing the core state of $|\mu_A\rangle$, \mathbf{c} . Third, the `recompose_circuit` procedure uses the optimal circuit parameters from `optimization` as well as the squeezing parameter, r , associated with $|\mu_A\rangle$ to find new GBS device parameters, where r is offloaded to the beginning of the circuit instead of having

to be applied in-line after the interferometer to the core state. The algorithm outputs are the parameters for a GBS device that produces a state given a post-selection pattern, $\bar{\mathbf{n}}$, along with the fidelity of that state to $|\mu_A\rangle$ and the success probability of obtaining that photon detection pattern.

The optimization procedure runs in two steps:

1. After initializing a random set of circuit parameters, \mathbf{x}_0 , consistent with the number of modes (constraining the initial squeezed light source to within 12 dB to maintain realistically achievable values), we employ the `basinhopping` algorithm already described in App. C1 to find a global candidate solution that maximizes the `cost` function labelled by the “global” handle. This global `cost` function is the fidelity between the core state of $|\mu_A\rangle$, \mathbf{c} , and the state output by the circuit given the post-selection pattern, $\bar{\mathbf{n}}$. Additionally, we add to the cost function the probability of the postselection pattern weighted by a factor of 0.1; if we solely optimized fidelity, we sometimes found solutions with probabilities smaller than 10^{-6} . We found the L-BFGS-B algorithm within `basinhopping` to be the most efficient local minimization algorithm in our case. We employed 50 hops to search the parameter space.
2. The optimal circuit parameters, \mathbf{x}_1 , from `basinhopping` are passed to `local_search`, which employs the `scipy minimize` function to find the optimal solution to the `cost` function with the “local” handle. This local `cost` function is almost the same as the global `cost` function, but now the weight for the probability of the state is increased to 1 so that a solution in the neighbourhood of \mathbf{x}_1 with comparable fidelity but higher probability is returned. The `scipy minimize` function stops running when the first local minimum is found instead of the iterative process of random hops implemented in `basinhopping`. This change ensures only the neighbourhood of \mathbf{x}_1 is searched instead of the entire parameter space.

We now move on to how we calculate the `cost` function, including how to calculate the output of the GBS circuit. The `cost` function is given the $|\mu_A\rangle$ core state coefficients, \mathbf{c} , initialized as a vector in the Fock basis that will be used to compute fidelity; the squeezing and interferometer parameters; and the postselection pattern, $\bar{\mathbf{n}}$, for the GBS device. The algorithm proceeds as follows:

1. Using a `strawberryfields` engine we apply the squeezing to each mode followed by the interferometer in the rectangular decomposition. As the state of the modes at this point is still Gaussian, we use the `gaussian` backend for the engine, which returns the mean and covariance matrix for the resulting state.
2. The mean and covariance matrix, along with $\bar{\mathbf{n}}$, is passed to the `state_vector` function from the `walrus`

library to get $|\psi_{\text{out}}\rangle$, the coefficients of the output state in the Fock basis (up to the dimension of \mathbf{c}), and before they have been normalized by the square root of the probability of obtaining $\bar{\mathbf{n}}$.

3. The probability is computed in the following way:

- (a) The mean and covariance matrix of all the modes before PNR measurements, as well as the indices of the modes we intend to measure, are passed to the `strawberryfields reduced_gaussian` function, which returns the mean and covariance matrix of ρ_{n-1} , the reduced $(n-1)$ -mode state of all but the output mode. ρ_{n-1} is still Gaussian, as it is the partial trace of a Gaussian state.
- (b) Using the mean and covariance matrix of ρ_{n-1} , we can use the `density_matrix_element` function of the `walrus` and find the diagonal element $\langle \bar{\mathbf{n}} | \rho_{n-1} | \bar{\mathbf{n}} \rangle$ which is exactly the probability of finding the post-selection pattern $\bar{\mathbf{n}}$.
4. We normalize $|\psi_{\text{out}}\rangle$ with the probability obtained in the previous step.
5. We take the inner product between $|\psi_{\text{out}}\rangle$ and \mathbf{c} in the Fock basis to compute the fidelity, and use the already calculated probability to compute the relevant `cost` function being called.

Finally, to implement `recompose_circuit`, we find the Gaussian unitary associated with $S(r)U(\Theta)D(\alpha)S(z)$ and use the `strawberryfields GaussianTransform` function to find $U(\Theta')D(\alpha')S(z')$.

Appendix D: Approximate GKP Z , X , and H eigenstates

1. Further comments on $|0_A\rangle$

Our results for constructing the $|0_A\rangle$ state using $n_{\max} = 4$ and 6 were identical until $\Delta \approx 8$ dB, meaning the $n = 6$ component was not required; after that point, there was a jump in parameter space and the $n = 6$ state was used to attain better fidelity than $n_{\max} = 4$, albeit only slightly. Here we explain why it makes sense that the results for $n_{\max} = 4$ and 6 are so close, and why we see a small jump in parameter space for $n_{\max} = 6$.

First, we can use the $|0_\epsilon\rangle$ state to understand the distribution of the Fock coefficients for $|0_\Delta\rangle$, since the states are almost indistinguishable for $\epsilon = \Delta^2$. In Fig. 28, we plot the probabilities of detecting the even Fock states in $|0_\epsilon\rangle$ up to $n = 12$ as a function of Δ , for $\epsilon = \Delta^2$. We see that, up to $\Delta \approx 8$ dB, the $n = 2$ and 4 states are the most probable states after the vacuum, and beyond $\Delta \approx 8$ dB, the $n = 8$ state becomes more probable. The probabilities for $n = 6, 10$, and 12 remain significantly smaller over the whole range. Thus, if one wants to increase the $n = 8$

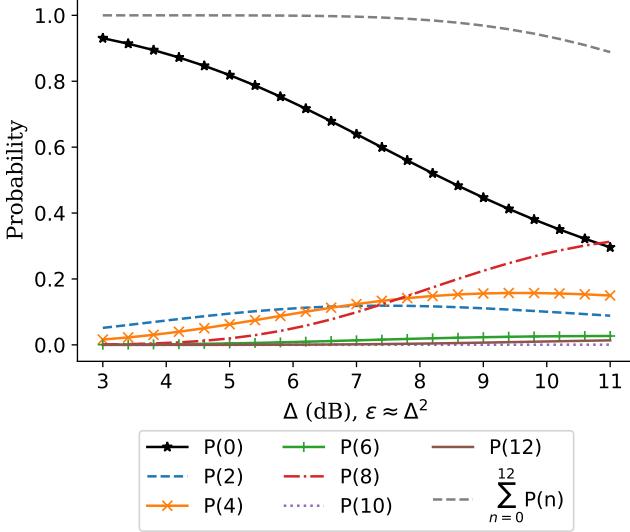


Figure 28. Probabilities of different even Fock states in $|0_\epsilon\rangle$ up to $n = 12$ as a function of $\epsilon \approx \Delta^2$. We see that the probability of measuring 8 photons is higher than 6; thus, when constructing the GKP states from squeezed core states, increasing n_{\max} of the core state from 4 to 6 provides little to no advantage compared to increasing n_{\max} to 8.

component using only squeezing applied to a core state of $n_{\max} = 6$, one will also end up increasing the $n = 6$ component in the distribution. If one wants to keep the $n = 6$ component small, one option is to truncate the core state at $n = 4$; as it turns out, this is the optimal option found numerically up to $\Delta \approx 8$ dB, leading to similar results between $n_{\max} = 4$ and 6 in that regime.

We can also use the wavefunctions of the Fock states to understand why the $n = 2, 4$, and 8 components are weighted higher than $n = 6$. In Fig. 29, we plot the wavefunction produced by the sum of the Fock state components in $|0_\Delta\rangle$ up to $n = 8$, as well as the wavefunctions for the weighted Fock states in the superposition. We start to see the peak structure of the GKP wavefunctions – a central peak and two outer peaks at the correct locations – in addition to some wiggles in between due to the truncation. Importantly, we see that the outer set of peaks is created almost exclusively by the $n = 8$ Fock state, with the $n = 2$ and 4 states working to narrow the central peak. Thus, were one to construct $|0_\Delta\rangle$ by only using a core state up to $n = 6$ and then stretching the Fock wavefunctions through squeezing, building the outer peaks of the $|0_\Delta\rangle$ state would require stretching the $n = 6$ or the $n = 4$ state to align with the GKP grid. But using the $n = 6$ state and weighting it heavily would shrink the central peak, since the $n = 6$ state has a dip at $q = 0$. Thus, the better option for matching the symmetry of the $|0_\Delta\rangle$ state is to stretch the $n = 4$ state, which has outer peaks and a central peak. This is further evidence for why we see no difference between the results for $|0_A\rangle$ using $n_{\max} = 4$ and 6 for $\Delta < 8$ dB.

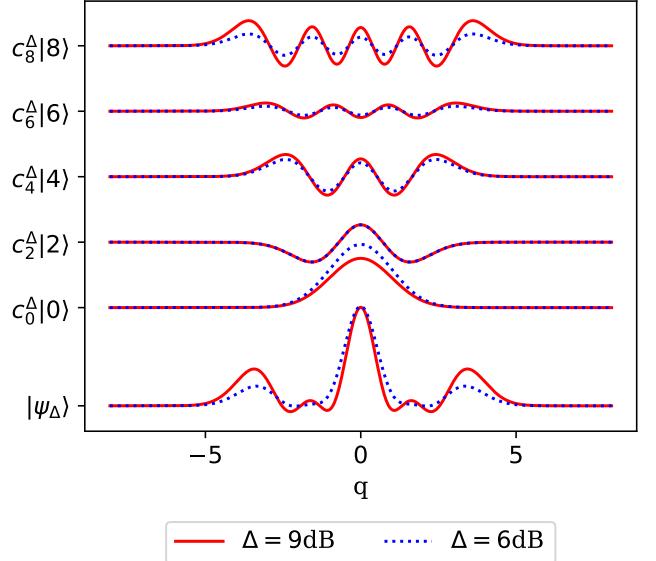


Figure 29. Wavefunctions produced from summing the first eight Fock states in the distribution for $|0_\epsilon\rangle$ (using $\epsilon \approx \Delta^2$), as well as the wavefunctions for the weighted Fock states in the superposition. We see the important contribution of the 8 photon state to the outer set of peaks for the normalizable GKP state. Thus, without the 8-photon state in the core of $|0_A\rangle$, the $n = 4$ or 6 states are needed to build the outer peaks, leading to worse fidelities.

For $\Delta > 8$ dB, we see a small jump in the (r, c) parameter space for the $n_{\max} = 6$ state, although the resulting wavefunction does not change much, leading to only slight jumps in the fidelity and average photon numbers. In that regime, all coefficients in the core state are now nonzero, and the squeezing applied to the core state as compared to $\Delta < 8$ dB is slightly greater. We understand this jump to come from the fact that the achievable fidelity to the target state is already relatively low ($\approx 90\%$) for this range of Δ ; in other words, since the core state resource available from n_{\max} is not too high, we are simply trying to find the best fidelity among several poorly contending regions of parameter space. There are two basins in parameter space that provide very similar fidelities when $\Delta = 8$ dB; as Δ increases, one basin overtakes the other to provide marginally higher fidelity, leading to a jump from the first basin to the second.

2. Results for $|1_A\rangle$, $|+_A\rangle$, and $|H_{+A}\rangle$

In Fig. 30, we provide the numerical results for the construction of the $|1_A\rangle$, $|+_A\rangle$, and $|H_{+A}\rangle$ states using Algorithm 1. All the trends for fidelity to the target states are comparable to those of the $|0_A\rangle$ states. The $|+_A\rangle$ state is effectively as resource-intensive to create as the $|0_A\rangle$ state given their relation by a Fourier transform. Finally, the $|H_{+A}\rangle$ state has squeezing and average pho-

ton numbers on the same order as $|0_A\rangle$ and $|+_A\rangle$, so it should also be comparatively demanding to prepare.

Notice that the results for the $|0_A\rangle$ in the main text and $|+_A\rangle$ states are nearly indistinguishable. Since the states are closely relatable by a Fourier transform [92], r for the $|0_A\rangle$ state is approximately $-r$ for the $|+_A\rangle$ state: one is squeezed in q and the other in p . Additionally, since the state $|n\rangle$ is an eigenfunction of the Fourier transform with eigenvalue $(-i)^n$, we expect the coefficients of the Fock states in the core state superposition to be related by phases of the form $(-i)^n$. In fact, we do see in our results for $|0_A\rangle$ and $|+_A\rangle$ that the core Fock states with $n \bmod 4 = 0$, for which $(-i)^n = 1$, have almost the same coefficients, while all the states with $n \bmod 4 = 2$, for which $(-i)^n = -1$, have coefficients that differ effectively by a phase of -1 . As a result of the states having such similar properties, their global properties end up being very close.

Appendix E: Additional characterization of approximate GKP states

1. Projectors

Projectors provide a valuable understanding as well as some visual intuition for how the code subspace behaves. Here we consider projectors defined by

1. The normalizable states $|\psi_\Delta\rangle$: $|0_\Delta\rangle\langle 0_\Delta| + |1_\Delta\rangle\langle 1_\Delta|$;
2. The approximate states $|\psi_A\rangle$: $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$;
3. $|0_A\rangle$ and displaced $|0_A\rangle$:

$$|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi}).$$

We consider the third kind of projector since one may want to prepare just one GKP Z eigenstate and generate the other through gate application. In Fig. 31, we provide some examples of these projectors plotted in the q basis. We take $\Delta = 10$ dB, and choose the approximate states that are meant to target the $\Delta = 10$ dB states with core states of $n_{\max} = 12$. Note that the projector $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$ maintains the symmetry of the target state projector, while the displacement in $|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi})$ breaks it. As we know from the modular subsystem decomposition, a loss of such symmetry does not affect the logical encoding. Of course, because the approximate states only have $\sim 95\%$ fidelity to the target states, the sharp peaks get blurrier further away from the origin in q space.

2. Quantum error correction matrix

One object that encompasses the error-correcting properties of general codes is the quantum error correction (QEC) matrix [10, 93]. For an error map with Kraus

operators $\{E_k\}$ and code states $|0_G\rangle$ and $|1_G\rangle$, the QEC matrix elements are defined by

$$\gamma_{kk'} \equiv P_G E_k^\dagger E_{k'} P_G, \quad (\text{E1})$$

where P_G is the projector constructed from the code states, as defined in App. E.

The 2×2 matrix γ can then be expanded in a basis of Pauli matrices:

$$\gamma_{kk'} = \varepsilon_0 \mathbb{1} + \varepsilon_x \sigma_x + \varepsilon_y \sigma_y + \varepsilon_z \sigma_z. \quad (\text{E2})$$

If γ is proportional to the identity, then there exists a recovery which can perfectly correct the error [93]. When this is not the case, errors can only be approximately corrected. However, the magnitude of the coefficients ε_i offer a convenient interpretation of the error's effects; namely, ε_x , ε_y , and ε_z are the probability of bit, bit-phase, and phase flip errors [10].

Ideal square-lattice GKP states were designed to correct displacement errors in phase space of up to $\sqrt{\pi}/2$. Thus, we examine the QEC matrix for displacement errors applied to the normalizable and approximate states. In Fig. 32, we calculate $P_G D(\beta) P_G$ as a function of displacement, β , expand it in terms of the Pauli matrices, and then plot the magnitudes of ε_i as functions of β . We do this for four choices of states (that is, four different P_G): the normalizable GKP states with $\Delta = 10$ dB and the approximate GKP states for the same Δ but with $n_{\max} = 4, 8$, and 12 .

The first row of plots in Fig. 32 corresponds to the normalizable states $|0_\Delta\rangle$ with $\Delta = 10$ dB. We see that, for small displacements, the largest magnitude is ε_0 , meaning the matrix is basically proportional to the identity, meaning these displacement are correctable. A displacement by $\sqrt{\pi}$ in q (p) leads to the dominant terms becoming ε_x (ε_z), corresponding to a bit (phase) flip. Further away from the origin, the blobs become more smeared, indicating that a displacement of that magnitude is more likely to lead to an uncorrectable error.

As we progress down the column, we see how the approximate states perform. The radii of the blobs become bigger as n_{\max} decreases, meaning the probability of error increases. Moreover, the blobs further from the origin become quite smeared in phase space, meaning only the first multiples of $\sqrt{\pi}/2$ will be correctable.

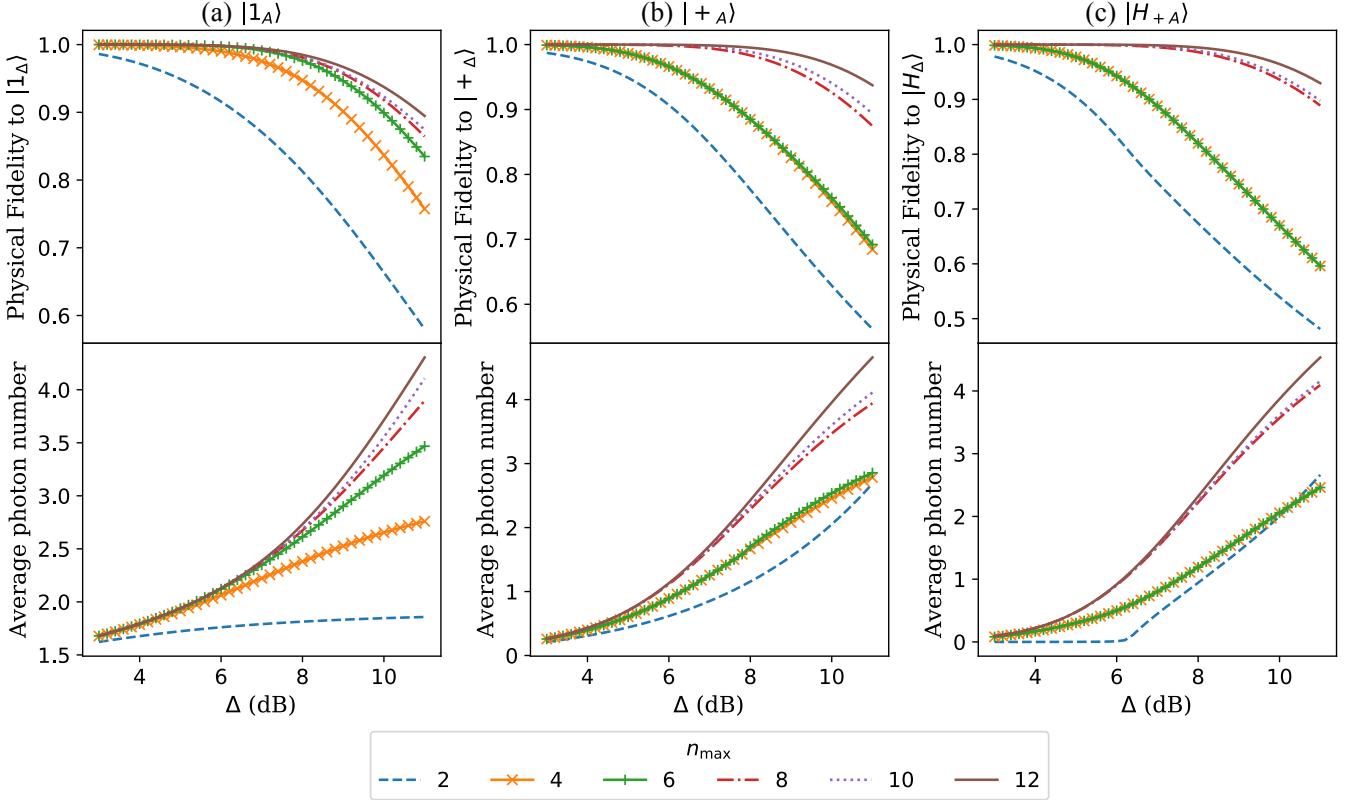


Figure 30. Fidelity of approximate states $|\mu_A\rangle$ to normalizable states $|\mu_\Delta\rangle$; and photon number vs. Δ for (by column) $\mu =$ (a) 1, (b) +, and (c) H_+ . Line colours/styles in each plot represent the results for different n_{\max} in the core states of $|\mu_A\rangle$. The trend in all cases is that a greater n_{\max} leads to better fidelity to the target state and higher energy; if the quality of the target state is improved by increasing the dB value of Δ , for a fixed n_{\max} , the fidelity gets worse, and the required energy gets higher. We expect comparable circuit resources (see Sec. III D) are required to construct approximate states for any point on the Bloch sphere.

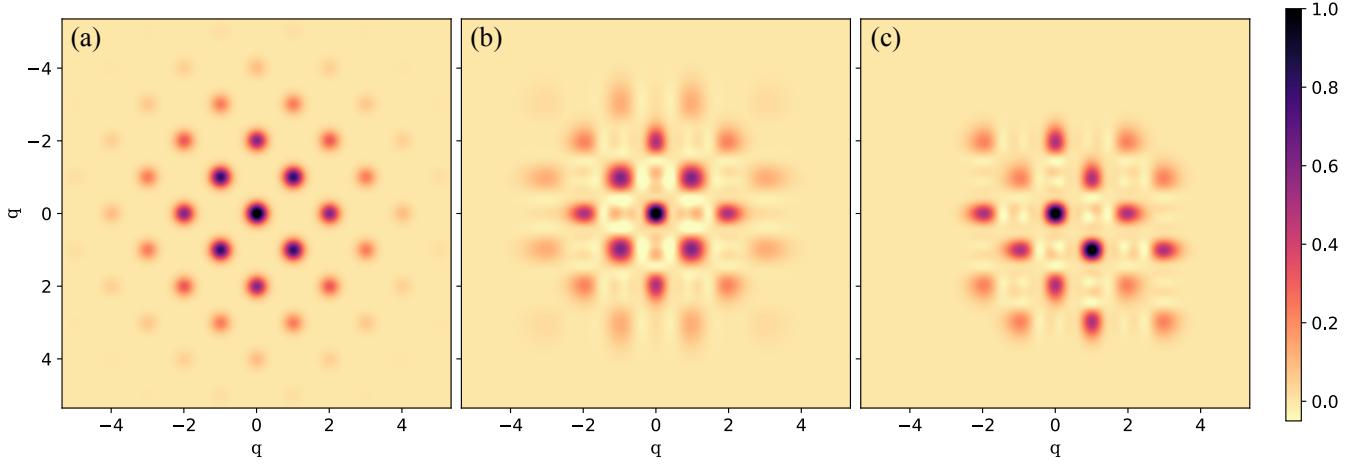


Figure 31. Single-qubit codespace projectors $P(q, q')$ in q basis (in units of $\sqrt{\pi}$) for three cases: (a) The normalizable GKP states $|0_\Delta\rangle\langle 0_\Delta| + |1_\Delta\rangle\langle 1_\Delta|$, with $\Delta = 10$ dB. (b) Our approximate GKP states $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$, designed to target the normalizable states with $\Delta = 10$ dB, constructed using squeezing applied to a core state with $n_{\max} = 12$. Here we assume we prepare both $|0_A\rangle$ and $|1_A\rangle$. (c) Our approximate GKP states $|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi})$ with the same parameters as (b), except here we assume we prepare only $|0_A\rangle$ and create the logical 1 state by displacing $|0_A\rangle$. We see that the symmetries of the target projector are best preserved when preparing both $|0_A\rangle$ and $|1_A\rangle$, although that symmetry may not be required for encoding.

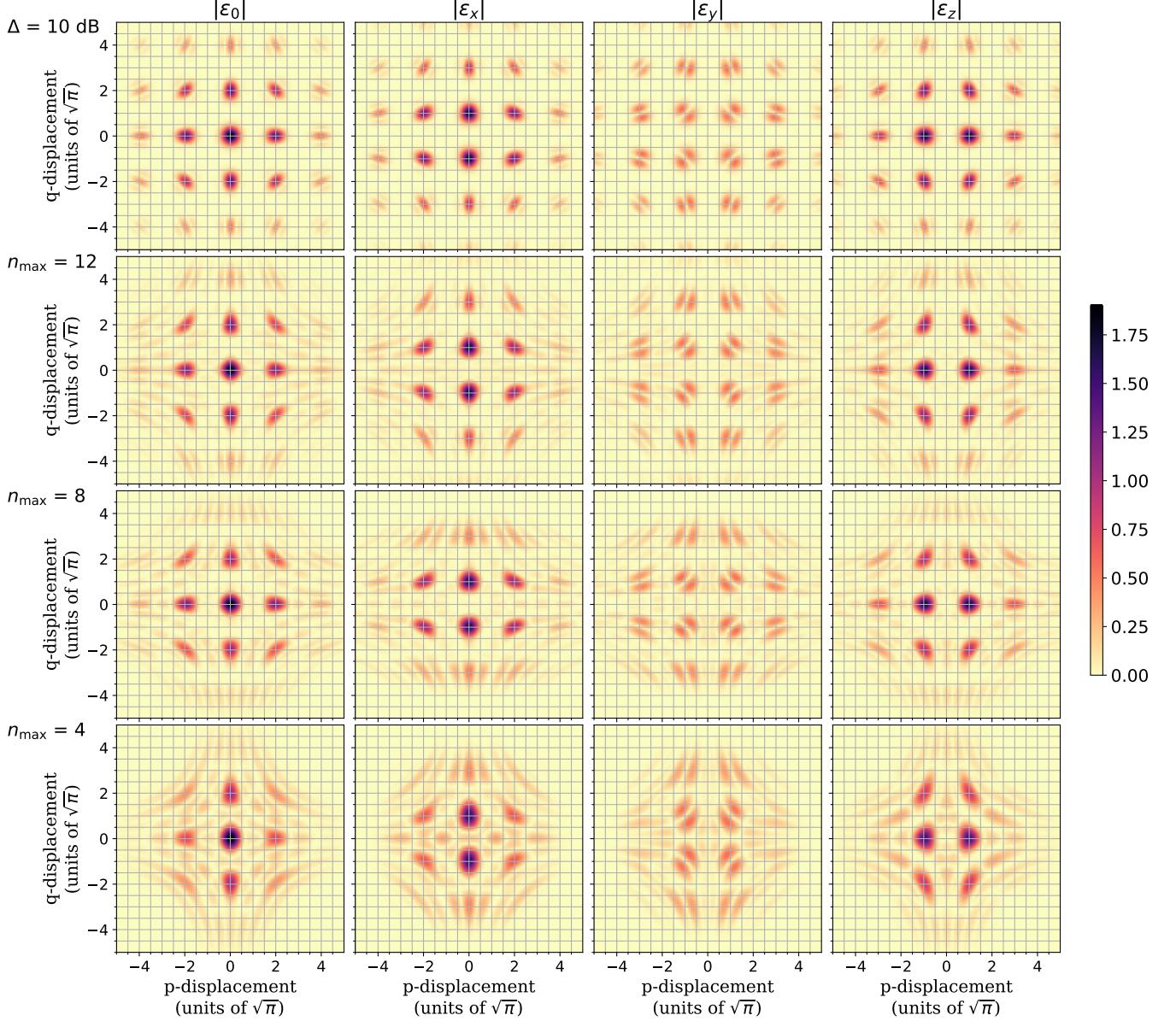


Figure 32. Visualization of the quantum error correction (QEC) matrices (Eq. (E1)) of normalizable and approximate GKP states, $|\psi_\Delta\rangle$ and $|\psi_A\rangle$, subject to displacement errors. Each row of subfigures corresponds to a different choice of code states. In each subfigure, the y (x) axis corresponds to a value of displacement along q (p) in phase space, so that a single point corresponds to a net displacement error. The QEC matrix associated with this displacement error can then be decomposed in terms of the identity and Pauli operators. Each column corresponds to the magnitude of the coefficients in the decomposition, which can be interpreted as no error, bit flip, bit-phase flip, and phase flip probabilities. The first row corresponds to $|\psi_\Delta\rangle$ with $\Delta = 10$ dB. The second through fourth rows correspond to $|\psi_A\rangle$ meant to approximate $|\psi_\Delta\rangle$ with $\Delta = 10$ dB with core states of $n_{\max} = 12, 8$, and 4 photons.

-
- [1] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310 (2001).
- [2] S. Glancy and E. Knill, Phys. Rev. A **73**, 012325 (2006).
- [3] N. C. Menicucci, Phys. Rev. Lett. **112**, 120504 (2014).
- [4] B. Q. Baragiola, G. Pantaleoni, R. N. Alexander, A. Karanjai, and N. C. Menicucci, Phys. Rev. Lett. **123**, 200502 (2019), arXiv:1903.00012.
- [5] B. W. Walshe, L. J. Mensen, B. Q. Baragiola, and N. C. Menicucci, Phys. Rev. A **100**, 010301 (2019).
- [6] K. Fukui, A. Tomita, A. Okamoto, and K. Fujii, Phys. Rev. X **8**, 021054 (2018).
- [7] K. Fukui, A. Tomita, and A. Okamoto, Phys. Rev. A **98**, 022326 (2018).
- [8] K. Fukui, A. Tomita, and A. Okamoto, Phys. Rev. Lett. **119**, 180507 (2017).
- [9] K. Fukui, “High-threshold fault-tolerant quantum computation with the gkp qubit and realistically noisy devices,” (2019), arXiv:1906.09767.
- [10] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. Girvin, *et al.*, Phys. Rev. A **97**, 032346 (2018).
- [11] K. Noh, V. V. Albert, and L. Jiang, IEEE Transactions on Information Theory **65**, 2563 (2019).
- [12] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, Phys. Rev. X **10**, 011058 (2020), arXiv:1901.08071.
- [13] K. Noh, S. M. Girvin, and L. Jiang, “Encoding an oscillator into many oscillators,” (2019), arXiv:1903.12615.
- [14] C. Vuillot, H. Asasi, Y. Wang, L. P. Pryadko, and B. M. Terhal, Phys. Rev. A **99**, 032344 (2019).
- [15] Y. Wang, “Quantum error correction with the gkp code and concatenation with stabilizer codes,” (2019), arXiv:1908.00147.
- [16] K. Noh and C. Chamberland, Phys. Rev. A **101**, 012316 (2020), arXiv:1908.03579.
- [17] K. Duivenvoorden, B. M. Terhal, and D. Weigand, Phys. Rev. A **95**, 012305 (2017).
- [18] T. Douce, D. Markham, E. Kashefi, P. Van Loock, and G. Ferrini, Phys. Rev. A **99**, 012344 (2019).
- [19] L. Li, D. J. Young, V. V. Albert, K. Noh, C.-L. Zou, and L. Jiang, “Designing good bosonic quantum codes via creating destructive interference,” (2019), arXiv:1901.05358.
- [20] B.-G. Englert, K. L. Lee, A. Mann, and M. Revzen, Journal of Physics A: Mathematical and General **39**, 1669 (2006).
- [21] A. Ketterer, A. Keller, S. Walborn, T. Coudreau, and P. Milman, Phys. Rev. A **94**, 022325 (2016).
- [22] G. Pantaleoni, B. Q. Baragiola, and N. C. Menicucci, “Modular bosonic subsystem codes,” (2019), arXiv:1907.08210.
- [23] T. Matsuurra, H. Yamasaki, and M. Koashi, “On the equivalence of approximate gottesman-kitaev-preskill codes,” (2019), arXiv:1910.08301.
- [24] K. H. Wan, A. Neville, and W. S. Kolthammer, “A memory-assisted decoder for approximate gottesman-kitaev-preskill codes,” (2019), arXiv:1912.00829.
- [25] B. Travaglione and G. J. Milburn, Phys. Rev. A **66**, 052322 (2002).
- [26] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics **37**, 283 (2006).
- [27] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. Home, Nature **566**, 513 (2019).
- [28] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, Phys. Rev. X **8**, 021001 (2018).
- [29] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, “A stabilized logical quantum bit encoded in grid states of a superconducting cavity,” (2019), arXiv:1907.12487.
- [30] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, Journal of Physics B: Atomic, Molecular and Optical Physics **39**, 997 (2006).
- [31] K. R. Motes, B. Q. Baragiola, A. Gilchrist, and N. C. Menicucci, Phys. Rev. A **95**, 053819 (2017).
- [32] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, EPL (Europhysics Letters) **68**, 323 (2004).
- [33] B. Terhal and D. Weigand, Phys. Rev. A **93**, 012315 (2016).
- [34] Y. Shi, C. Chamberland, and A. W. Cross, New J. Phys. **21**, 093007 (2019), arXiv:1905.00903.
- [35] Y. Shi, C. Chamberland, A. Cross, and F. Chong, “Building fault-tolerant qubits using more physical states,” (2019).
- [36] D. J. Weigand and B. M. Terhal, “Realizing modular quadrature measurements via a tunable photon-pressure coupling in circuit-qed,” (2019), arXiv:1909.10075.
- [37] D. J. Weigand and B. M. Terhal, Phys. Rev. A **97**, 022341 (2018).
- [38] S. Puri, A. Grimm, P. Campagne-Ibarcq, A. Eickbusch, K. Noh, G. Roberts, L. Jiang, M. Mirrahimi, M. H. Devoret, and S. M. Girvin, Phys. Rev. X **9**, 041009 (2019).
- [39] H. M. Vasconcelos, L. Sanz, and S. Glancy, Optics letters **35**, 3261 (2010).
- [40] J. Etesse, R. Blandino, B. Kanseri, and R. Tualle-Brouri, New J. Phys. **16**, 053001 (2014).
- [41] M. Eaton, R. Nehra, and O. Pfister, arXiv:1903.01925 (2019).
- [42] J. M. Arrazola, T. R. Bromley, J. Izaac, C. R. Myers, K. Brádler, and N. Killoran, Quantum Science and Technology **4**, 024004 (2019).
- [43] N. Fabre, G. Maltese, F. Appas, S. Felicetti, A. Ketterer, A. Keller, T. Coudreau, F. Baboux, M. Amanti, S. Ducci, and P. Milman, “Continuous variables error correction with integrated biphoton frequency combs,” (2019), arXiv:1904.01351.
- [44] D. T. Le, A. Grimsmo, C. Mueller, and T. M. Stace, Phys. Rev. A **100**, 062321 (2019), arXiv:1904.01843.
- [45] W.-C. Su, C.-Y. Lin, and S.-T. Wu, “Encoding qubits into harmonic-oscillator modes via quantum walks in phase space,” (2018), arXiv:1808.08722.
- [46] J. Hastrup, K. Park, J. B. Brask, R. Filip, and U. L. Andersen, “Measurement-free preparation of grid states,” (2019), arXiv:1912.12645.
- [47] K. K. Sabapathy, H. Qi, J. Izaac, and C. Weedbrook, Phys. Rev. A **100**, 012326 (2019).
- [48] D. Su, C. R. Myers, and K. K. Sabapathy, “Generation of photonic non-gaussian states by measuring multimode gaussian states,” (2019), arXiv:1902.02331.
- [49] D. Su, C. R. Myers, and K. K. Sabapathy, Phys. Rev. A **100**, 052301 (2019), arXiv:1902.02323.

- [50] N. Quesada, L. G. Helt, J. Izaac, J. M. Arrazola, R. Shahrokhshahi, C. R. Myers, and K. K. Sabapathy, *Phys. Rev. A* **100**, 022341 (2019).
- [51] T. B. Pittman, B. C. Jacobs, and J. D. Franson, *Phys. Rev. A* **71**, 052332 (2005).
- [52] X.-C. Yao, T.-X. Wang, H.-Z. Chen, W.-B. Gao, A. G. Fowler, R. Raussendorf, Z.-B. Chen, N.-L. Liu, C.-Y. Lu, Y.-J. Deng, *et al.*, *Nature* **482**, 489 (2012).
- [53] B. Bell, D. Herrera-Martí, M. Tame, D. Markham, W. Wadsworth, and J. Rarity, *Nat. Commun.* **5**, 3658 (2014).
- [54] J. C. Adcock, C. Vigliar, R. Santagati, J. W. Silverstone, and M. G. Thompson, *Nat. Commun.* **10**, 3528 (2019), arXiv:1811.03023.
- [55] T. Aoki, G. Takahashi, T. Kajiya, J.-i. Yoshikawa, S. L. Braunstein, P. Van Loock, and A. Furusawa, *Nature Physics* **5**, 541 (2009).
- [56] S. Hao, X. Su, C. Tian, C. Xie, and K. Peng, *Scientific reports* **5**, 15462 (2015).
- [57] C.-Y. Lu, W.-B. Gao, J. Zhang, X.-Q. Zhou, T. Yang, and J.-W. Pan, *Proceedings of the National Academy of Sciences* **105**, 11050 (2008).
- [58] M. Lassen, A. Berni, L. S. Madsen, R. Filip, and U. L. Andersen, *Phys. Rev. Lett.* **111**, 180502 (2013).
- [59] M. Lassen, M. Sabuncu, A. Huck, J. Niset, G. Leuchs, N. J. Cerf, and U. L. Andersen, *Nature Photonics* **4**, 700 (2010).
- [60] Q. Zhuang, J. Preskill, and L. Jiang, *New J. Phys.* **22**, 022001 (2020), arXiv:1910.14156.
- [61] What we call normalizable is generally referred to as approximate in the literature. We prefer this nomenclature because there are several different approximations to ideal GKP states we will be considering.
- [62] Note that we avoid ambiguity in notation because the subscript of an ideal state $|\psi_I\rangle$ can be interpreted as a normalizable state with an identity envelope.
- [63] Note that in Pantaleoni et al., Fig. 2, the Bloch sphere trajectory for the $|+_G\rangle$ state differs from ours. This is because a different approximation is used there, effectively corresponding to a different envelope operator.
- [64] K. K. Sabapathy, “Some aspects of the interplay between bipartite correlations and quantum channels,” (2016), arXiv:1604.08017.
- [65] S. Bravyi and A. Kitaev, *Phys. Rev. A* **71**, 022316 (2005).
- [66] A. L. Gibbs and F. E. Su, *Int. Stat. Rev.* **70**, 419 (2002), arXiv:math/0209021.
- [67] In [10], the authors show that the GKP code (i.e. the codespace spanned by $|0_\Delta\rangle$ and $|1_\Delta\rangle$) is geometrically distributed in Fock space and has a mean occupation number of $\bar{n}(|0_\Delta\rangle + |1_\Delta\rangle) \approx \frac{1}{\Delta^2} - 1$.
- [68] In [94] there is a simplified, easy-to-compute expression for averaged gate fidelity. However, it assumes that the error channel is trace-preserving, so we can only use it in our setting with appropriate modifications.
- [69] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [70] A. M. Steane, *Phys. Rev. Lett.* **77**, 793 (1996).
- [71] E. Knill, *Phys. Rev. A* **71**, 042322 (2005).
- [72] E. Knill, *Nature* **434**, 39 (2005).
- [73] J. Fiurášek, R. García-Patrón, and N. J. Cerf, *Phys. Rev. A* **72**, 033822 (2005).
- [74] D. Menzies and R. Filip, *Phys. Rev. A* **79**, 012313 (2009).
- [75] U. Chabaud, D. Markham, and F. Grosshans, *Phys. Rev. Lett.* **124**, 063605 (2020), arXiv:1907.11009.
- [76] B. Dutta, N. Mukunda, R. Simon, *et al.*, *Pramana* **45**, 471 (1995).
- [77] Y. Miwa, J.-i. Yoshikawa, N. Iwata, M. Endo, P. Marek, R. Filip, P. van Loock, and A. Furusawa, *Phys. Rev. Lett.* **113**, 013601 (2014).
- [78] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, *et al.*, “SciPy 1.0—fundamental algorithms for scientific computing in Python,” arXiv:1907.10121 [cs.MS].
- [79] D. J. Wales and J. P. K. Doye, *The Journal of Physical Chemistry A* **101**, 5111 (1997).
- [80] H. Yamasaki, T. Matsuura, and M. Koashi, “Cost-reduced all-gaussian universality with the gottesmann-kitaev-preskill code,” (2019), arXiv:1911.11141.
- [81] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, *Quantum* **3**, 129 (2019).
- [82] A. Björklund, B. Gupt, and N. Quesada, *ACM Journal of Experimental Algorithms* **24**, 17 (2019), arXiv:1805.12498.
- [83] B. Gupt, J. Izaac, and N. Quesada, *Journal of Open Source Software* **4**, 1705 (2019).
- [84] H. Vahlbruch, M. Mehmet, K. Danzmann, and R. Schnabel, *Phys. Rev. Lett.* **117**, 110801 (2016).
- [85] K. Miyata, H. Ogawa, P. Marek, R. Filip, H. Yonezawa, J.-i. Yoshikawa, and A. Furusawa, *Phys. Rev. A* **90**, 060302 (2014).
- [86] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, “Approximate GKP preparation repository,” <https://github.com/XanaduAI/approximate-GKP-prep> (2020).
- [87] I. Tzitrin, J. E. Bourassa, and K. K. Sabapathy, “Riding bosonic qubits towards fault-tolerant quantum computation,” (2019).
- [88] S. Biagi, A. Bonfiglioli, and M. Matone, *Linear and Multilinear Algebra* **0**, 1 (2018), arXiv:1805.10089.
- [89] A. Kenfack and K. Zyczkowski, *Journal of Optics B: Quantum and Semiclassical Optics* **6**, 396 (2004), arXiv:0406015 [quant-ph].
- [90] F. Albarelli, M. G. Genoni, M. G. Paris, and A. Ferraro, *Phys. Rev. A* **98**, 052350 (2018).
- [91] T. Bröcker and R. Werner, *Journal of Mathematical Physics* **36**, 62 (1995).
- [92] They would be exactly relatable if we chose the $E(\epsilon)$ envelope since $E(\epsilon)$ commutes with phase space rotations.
- [93] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, New York, NY, USA, 2011).
- [94] M. A. Nielsen, *Phys. Lett. Sect. A Gen. At. Solid State Phys.* **303**, 249 (2002).