

Optimal encoding of oscillators into more oscillators

Jing Wu¹, Anthony J. Brady², and Quntao Zhuang^{3,1,2}

¹James C. Wyant College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

²Department of Electrical and Computer Engineering, University of Arizona, Tucson, Arizona 85721, USA

³ Ming Hsieh Department of Electrical and Computer Engineering & Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA

Bosonic encoding of quantum information into harmonic oscillators is a hardware efficient approach to battle noise. In this regard, oscillator-to-oscillator codes not only provide an additional opportunity in bosonic encoding, but also extend the applicability of error correction to continuous-variable states ubiquitous in quantum sensing and communication. In this work, we derive the optimal oscillator-to-oscillator codes among the general family of Gottesman-Kitaev-Preskill (GKP)-stabilizer codes for homogeneous noise. We prove that an arbitrary GKP-stabilizer code can be reduced to a generalized GKP two-mode-squeezing (TMS) code. The optimal encoding to minimize the geometric mean error can be constructed from GKP-TMS codes with an optimized GKP lattice and TMS gains. For single-mode data and ancilla, this optimal code design problem can be efficiently solved, and we further provide numerical evidence that a hexagonal GKP lattice is optimal and strictly better than the previously adopted square lattice. For the multimode case, general GKP lattice optimization is challenging. In the two-mode data and ancilla case, we identify the D4 lattice—a 4-dimensional dense-packing lattice—to be superior to a product of lower dimensional lattices. As a by-product, the code reduction allows us to prove a universal no-threshold-theorem for arbitrary oscillators-to-oscillators codes based on Gaussian encoding, even when the ancilla are not GKP states.

1 Introduction

The power of quantum information processing comes from delicate quantum effects such as coherence, squeezing and entanglement. As noise is ubiquitous, maintaining such power relies on quantum error correction. Since the early works [1, 2], various codes and the corresponding error correction systems have been proposed, mostly focusing on discrete-variable

Jing Wu: These two authors contributed equally.

Anthony J. Brady: These two authors contributed equally.

Quntao Zhuang: qzhuang@usc.edu

physical realizations. The Gottesman-Kitaev-Preskill (GKP) code [3] provides an alternative approach to utilize the bosonic degree of freedom of an oscillator to encode a qubit. Such a qubit-to-oscillator encoding is hardware efficient, as the entire degree of freedom of a (cavity) mode is utilized and at the same time, microwave cavities have a long lifetime [4]. By this approach, the lifetime of a logical qubit has been extended beyond the error correction ‘break-even’ point for cat codes [5] and more recently for GKP codes [6].

The concatenation of bosonic codes and qubit codes has also shown great promise [7], with applications in quantum repeaters [8] and quantum computer architecture [9]. It then becomes a natural question whether an oscillator can be encoded into multiple oscillators to gain additional advantage in protecting logical data. Noh, Girvin and Jiang provide an affirmative answer by designing a GKP-stabilizer code [10, 11] to protect an oscillator by entangling it with an additional GKP ancilla. Despite that these oscillator-to-oscillator codes alone do not have a threshold theorem due to their analog nature and finite squeezing [12], an oscillator-to-oscillator encoding at the bottom-layer can substantially suppress the error in a multi-level qubit encoding [13]. Moreover, the GKP-stabilizer code can be utilized in general to protect an oscillator in an arbitrary quantum state, including a squeezed vacuum state and continuous-variable multi-partite entangled states that are widely applicable to quantum sensing [14, 15] and communication [16].

To maximally benefit from the oscillator-to-oscillator encoding for various tasks mentioned above, it is important to find the optimal GKP-stabilizer code design. In this paper, we make progress towards this endeavour for the commonly considered homogeneous noise case [11]. We prove that an arbitrary GKP-stabilizer code can be reduced to a generalized GKP two-mode-squeezing (TMS) code. Therefore, GKP-TMS codes with optimized GKP ancilla and gains can achieve the minimum geometric mean error. For decoding, we derive the minimum mean square error (MMSE) estimator, which is superior to the linear estimator [11] in terms of minimizing the residue noise on the data modes. While linear estimation leads to a break-even point (the point above which noise can no longer be reduced by error correc-

tion) of additive noise at $\sigma_{\text{lin}}^* \sim 0.558$, the MMSE estimation pushes it to $\sigma_{\text{MMSE}}^* = 0.605(5)$, which is just at the edge of the best break-even region for arbitrary GKP codes [0.607, 0.701] derived from our quantum capacity analyses.

For single-mode data and ancilla, we further show that the optimal code design problem of minimizing the geometric mean error can be efficiently solved and provide numerical evidence that a hexagonal GKP lattice is optimal and strictly better than the previously adopted square lattice. For the multimode case, solving the optimal lattice is more challenging. For two-mode data and ancilla, we identify the D4 lattice [17, 18]—a 4-dimensional dense-packing lattice—to be superior to direct products of 2-dimensional lattices. For a single data mode with two ancilla modes, we compare a few examples of lattices and find the performance to be dependent on the noise levels. Our results indicate that high-dimensional lattices have the potential of outperforming low-dimensional ones for GKP-stabilizer codes.

Besides the results on the optimal code design, the code reduction also allows us to prove a more general version of the no-threshold-theorem of Ref. [12], where the original proof relies on explicit maximum likelihood error decoder based on GKP-type syndrome information. Our proof is based on a simple, classical information-theoretical argument; moreover, our no-threshold-theorem applies to all GKP-stabilizer codes, and more generally, even when the GKP ancilla are replaced by general non-Gaussian states.

This paper is organized as the following. Section 2 introduces the general GKP-stabilizer code for encoding an oscillator into many oscillators. Our main theorem of code reduction and optimality is presented in Section 3, which then leads to the no-threshold theorem in Section 4. Section 5 introduces general GKP lattices and the MMSE estimation. The final part of the paper addresses code optimization and comparison, with single-mode case in Section 6 and multimode case in Section 7. We end the paper with discussions on the heterogeneous case in Section 8 and the imperfect GKP states in Section 9.

2 General GKP-stabilizer code

As shown in Fig. 1 (a), a general GKP-stabilizer code operates as follows. An arbitrary N -mode data system in quantum state Ψ is encoded into $K = M + N$ modes by applying a Gaussian unitary to entangle the data system with an M -mode ancilla system in a non-Gaussian state $\mathcal{L}$. In general, non-Gaussian states are required, due to the no-go theorem of Gaussian error correction [19]. Typically, we are interested in (canonical) GKP lattice states, hence $\mathcal{L}$ for lattice. The Gaussian unitary $\mathcal{U}_{S_{\text{enc}}}$ can be described by the symplectic transform S_{enc} (see Appendix B for a brief introduction of Gaussian uni-

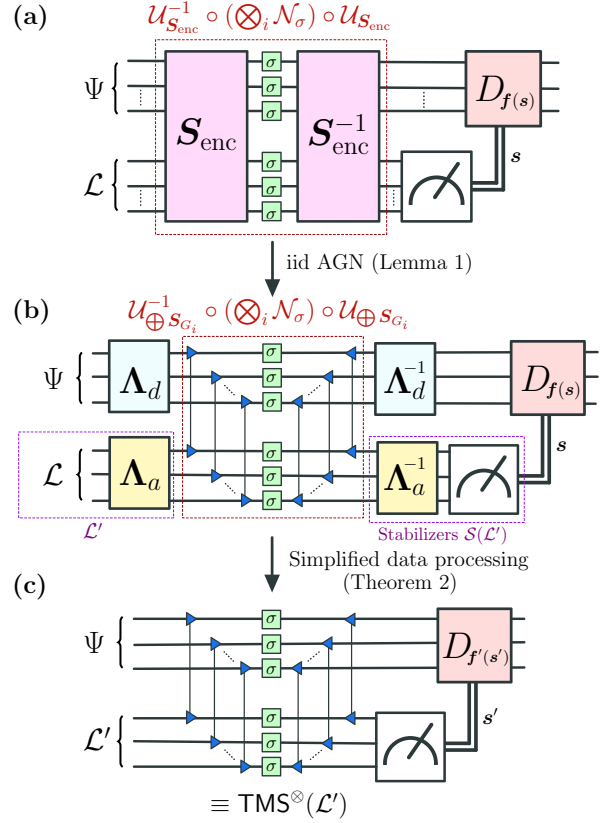


Figure 1: Illustration of GKP-stabilizer code reduction (here, $M = N$). (a) A general GKP-stabilizer code with encoding S_{enc} and an ancillary GKP lattice $\mathcal{L}$. The syndromes s are extracted from stabilizer measurements on the ancillary lattice $\mathcal{L}$ and inform the corrective displacements on the data $D_{f(s)}$. (b) Reduction of the GKP stabilizer code to a set of TMS operations between the data and ancilla modes, up to local symplectic transformations Λ_d and Λ_a . (c) Coherent data processing via Λ_d does not change the performance of the code. Acting on the initial ancillary lattice $\mathcal{L}$ by a symplectic transformation Λ_a produces another lattice $\mathcal{L}'$. Thus, a general GKP-stabilizer code reduces to a direct product of TMS codes with a GKP ancillary lattice $\mathcal{L}'$, $\text{TMS}^{\otimes}(\mathcal{L}')$.

ties). We denote the corresponding unitary channel as $\mathcal{U}_{S_{\text{enc}}}$.

At this juncture, we want to emphasize that the GKP lattice states appearing in this paper are *canonical* lattice states as opposed to *computational* lattice states. Computational lattice states allow one to encode a qubit (or qudit) into an oscillator per the original GKP approach [3]. Canonical lattices have larger spacing than their computational counterparts and, consequently, cannot support digital information. However, canonical lattices can be quite useful in multimode codes and, particularly, for oscillators-to-oscillators codes, as demonstrated in Refs [11, 20]. Therefore, when considering GKP states in this paper, we refer solely to canonical lattices. We describe the mathematical and technical details of canonical lattice states in Section 5.

As explained in the original GKP paper [3], a nat-

ural noise model for bosonic systems is random displacements. In practice, such noise arises from amplifying a lossy signal, leading to additive Gaussian noise (AGN)—i.e. the random displacements have a Gaussian distribution. As general correlated Gaussian noises can be reduced to independent noises [20], it suffices to consider a product of single-mode AGNs. Formally, we denote the single-mode AGN channel as $\mathcal{N}_\sigma$, where σ^2 is the variance of the displacement noise on a single mode.

On the decoding side, an inverse Gaussian unitary described by $\mathbf{S}_{\text{enc}}^{-1}$ is first applied to disentangle the data and the ancilla. Then one measures the ancilla system to perform error correcting displacement operations $D_{\mathbf{f}(\mathbf{s})}$ based on the measurement results (or *syndromes*) $\mathbf{s}$ ¹. Here, $\mathbf{f}$ is a vector function (i.e., $\mathbf{f}: \mathbb{R}^{2M} \rightarrow \mathbb{R}^{2N}$) that takes the syndromes $\mathbf{s}$ as input and provides an estimate for the error displacements on the data. The corrective displacements aim to cancel the additive noise on the data. Due to the analog nature of the errors, such a cancellation is never perfect, and there will be residual, random displacements ξ_d on the output data modes.

To quantify the error correction performance, we evaluate the covariance matrix $\mathbf{V}_{\text{out}}$ of the residue displacements ξ_d , despite the distribution of ξ_d being non-Gaussian after decoding. Note that, prior to encoding, one may also apply a Gaussian unitary $U_{\mathbf{S}_d}$ and then apply the inverse $U_{\mathbf{S}_d}^{-1}$ after the final decoding step. Pre- and post-processing transform the residue noise covariance matrix as $\mathbf{S}_d^{-1} \mathbf{V}_{\text{out}} \mathbf{S}_d^{-\top}$. Such an operation is considered ‘free’ and should not improve the performance of the error correction. Therefore, we consider the geometric mean (GM) error

$$\bar{\sigma}_{\text{GM}}^2 \equiv \sqrt[2N]{\det \mathbf{V}_{\text{out}}}, \quad (1)$$

as the figure of merit to benchmark code performance, which is necessarily invariant under symplectic operations on the data. Moreover, the GM error has information theoretic roots, as it relates to a lower bound on the quantum capacity for the additive non-Gaussian noise channel of a multimode GKP code. Given the geometric mean error of general additive noise $\bar{\sigma}_{\text{GM}}^2$, the quantum capacity of the N -mode additive noise channel $C_Q \geq \max \left[0, N \log_2 \left(\frac{1}{e \bar{\sigma}_{\text{GM}}^2} \right) \right]$ (see Lemma 14 of Appendix H).

As an alternative simpler metric, we also consider the root-mean-square (RMS) error,

$$\bar{\sigma}_{\text{RMS}}^2 \equiv \frac{\text{Tr}\{\mathbf{V}_{\text{out}}\}}{2N}. \quad (2)$$

The RMS error is only invariant under orthogonal symplectic (e.g., linear optical) transformations on

¹Note that the measurement and displacements can be equivalently pushed before the inverse unitary, as shown in Ref. [13].

the data. However, it is easier to evaluate and provides an upper bound for the GM error, $\bar{\sigma}_{\text{RMS}}^2 \geq \bar{\sigma}_{\text{GM}}^2$.

Before moving to our main results, we specify a few code examples. In Ref. [11], two codes are proposed based on the canonical GKP square lattice, the GKP-TMS code described by the TMS symplectic matrix

$$\mathbf{S}_G = \begin{pmatrix} \sqrt{G} \mathbf{I} & \sqrt{G-1} \mathbf{Z} \\ \sqrt{G-1} \mathbf{Z} & \sqrt{G} \mathbf{I} \end{pmatrix}, \quad (3)$$

with the gain G tunable, and the GKP-squeezed-repetition code (also see Ref. [14]), which has the following encoding matrix for $N = 2$ modes,

$$\mathbf{S}_{\text{Sq-Rep}}^{[2]} \equiv \begin{pmatrix} 1/\lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & -\lambda \\ \lambda & 0 & \lambda & 0 \\ 0 & 0 & 0 & 1/\lambda \end{pmatrix}, \quad (4)$$

with λ being a tunable parameter.

3 General reduction of encoding

We focus on the homogeneous noise model, where the displacement error on all modes are independent and identically distributed (iid), where the overall K -mode noise channel $\bigotimes_{i=1}^K \mathcal{N}_\sigma$.

Towards proving the optimal code design, we begin with the following lemma to simplify a general code.

Lemma 1. *For an iid AGN channel $\bigotimes_{i=1}^K \mathcal{N}_\sigma$ and up to local Gaussian unitaries acting on all data or all ancilla modes, a GKP-stabilizer code with any symplectic encoding matrix $\mathbf{S}_{\text{enc}}$ reduces to a direct product of N TMS operations (between the N data modes and N ancilla modes) together with an identity operation on the remaining $M - N$ ancilla modes—i.e., $\mathbf{S}_{\text{enc}} \rightarrow \bigoplus_{i=1}^N \mathbf{S}_{G_i} \oplus \mathbf{I}_{2(M-N)}$.*

In other words, we can decompose a general GKP-stabilizer code into TMS operations and local symplectic operations; see Fig. 1 (a-b) for a visual aid of the lemma for the $M = N$ case. The local Gaussian unitary U_{Λ_d} and its inverse are applied on the data modes, while the local Gaussian unitary U_{Λ_a} and its inverse are applied on the ancilla modes. Consequently, the encoding and decoding can be taken as simple product of TMS operations. The proof of this result is based on Gaussian channel synthesis and the modewise entanglement theorem [21] (Theorem 7 and Theorem 8 of Appendix B, respectively), which we present in Appendix F in full detail.

The above lemma reduces the number of parameters in code description from $2(M+N)^2$ to $2M^2 + 2N^2$ in the leading order. In fact, as we will explain later, coherent data processing (via Λ_d) only reshapes the residual noise; it does not aid in error correction. Hence a further reduction to $2M^2$ is permissible. Moreover, the generally multimode entangling operations between data and ancilla are now given by

standard TMS operations with N to-be-determined gain parameters G_i .

We now consider the GM error as a performance metric and arrive at the following main result of the work (see Fig. 1 (c)).

Theorem 2 (Sufficiency of $\text{TMS}^\otimes(\mathcal{L})$). *For an iid AGN channel, in terms of the geometric mean error on the data modes, the most general GKP-stabilizer code can be completely reduced to a direct product of TMS codes with a general (potentially multimode) ancillary lattice state $\mathcal{L}$, defined as $\text{TMS}^\otimes(\mathcal{L})$.*

Proof. Theorem 2 is a consequence of Lemma 1, together with the fact that (1) a local symplectic transformation Λ_a on the initial ancillary lattice $\mathcal{L}$ defines a new (symplectically integral) lattice $\mathcal{L}'$ (see, e.g., Appendix C) and (2) coherent pre- and post-processing of the data modes do not increase code performance. To show (2), using the geometric mean $\bar{\sigma}_{\text{GM}}^2$ as a performance metric, first observe that we can push data noise processing via Λ_d^{-1} after the corrective displacement D_f by simply redefining our error estimator $f' \equiv \Lambda_d f$, i.e., $D_f \circ \mathcal{U}_{\Lambda_d^{-1}} = \mathcal{U}_{\Lambda_d^{-1}} \circ D_{\Lambda_d f}$. Just after corrective displacements (but prior to data processing), the residual error covariance matrix on the data modes is V_{out} and transforms to $V'_{\text{out}} = \Lambda_d^{-1} V_{\text{out}} \Lambda_d^{-\top}$ after data processing. However, the geometric mean is invariant under symplectic transformations; in other words, $\det V'_{\text{out}} = \det V_{\text{out}}$. Thus, the performance of the code (as quantified by the geometric mean error) only depends on the gain parameters of the TMS operations, the ancillary lattice state, and the error estimator for corrective displacements—not on Λ_d . ■

Theorem 2 reduces code design to choosing N gain parameters for the TMS operations and finding a good ancillary lattice $\mathcal{L}$ for best code performance. [We approach this problem numerically in later sections.] Therefore, the following corollary directly follows from Theorem 2 as an observation.

Corollary 2.1 (Optimality of $\text{TMS}^\otimes(\mathcal{L})$). *The optimal GKP-stabilizer code with N data modes can be constructed from a GKP-TMS code $\text{TMS}^\otimes(\mathcal{L})$ with an optimized (potentially multimode) GKP lattice $\mathcal{L}$ and optimized TMS gain parameters $\{G_i\}_{i=1}^N$.*

As a side note, in deriving Lemma 1 and Theorem 2, observe that we do not explicitly utilize properties of the ancilla state $\mathcal{L}$. This indicates that our results hold for any Gaussian encoding with general non-Gaussian ancilla, not only for the GKP-stabilizer codes. In particular, one can likewise consider a general M -mode non-Gaussian resource state $\mathcal{L}$ and a family of such states $\mathcal{F}(\mathcal{L})$ that are related to $\mathcal{L}$ via Gaussian transformations, i.e. $\mathcal{F}(\mathcal{L}) \equiv \{\mathcal{L}' = \mathcal{U}_{\Lambda_a}(\mathcal{L}) | \Lambda_a \in \text{Sp}(2M, \mathbb{R})\}$, where $\text{Sp}(2M, \mathbb{R})$ is the set of $2M \times 2M$ real symplectic matrices. Such a class of codes represent a fairly general encoding of oscillators-to-oscillators, especially considering that Gaussian

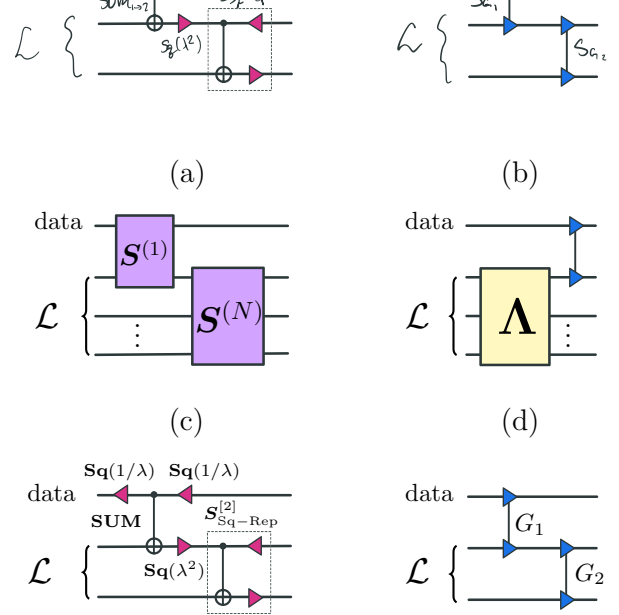


Figure 2: Circuit diagrams for concatenated codes. (a) General definition of a concatenated code. Data couples to an ancilla via $S^{(1)}$. Adding more layers leads to further error suppression. (b) Equivalent representation using TMS code reduction. Most operations in a concatenated code can be done offline on the ancillae and the ancillae can be coupled to the data at the end via TMS. (c) Example of a three mode concatenated squeezed repetition code (cf. Ref. [10]). (d) Example of a three mode concatenated TMS code.

operations supplied with non-Gaussian GKP ancillae are universal and sufficient for fault-tolerant quantum computation [22].

3.1 Concatenated codes

In general, a concatenated code applies another layer of encoding on the ancilla mode to suppress the noise level on the ancilla modes, prior to utilizing the ancilla modes to protect the data modes, as shown in Fig. 2(a). Examples of concatenated codes are the GKP-squeezed-repetition code detailed in Fig. 2(c), where repeated sum-gate and single-mode squeezing is applied in each layer of encoding, and the three-mode TMS code shown in Fig. 2(d). The purpose of concatenation is to suppress the logical output variance to higher powers of the input AGN σ , analogous to DV qubit codes. In particular, for N concatenation layers, we expect that $\sigma^{\text{out}} \sim \sigma^{N+1}$. Evidence of higher-order error suppression has been shown for squeezed repetition codes [10, 14] and TMS codes [20, 15].

We now make an interesting observation on Theorem 2 as it pertains to concatenated codes. To protect N data modes with a concatenated code, Theorem 2 states that we can prepare a multimode ancilla ($M > N$) “offline” and then couple the N data modes to the M ancilla modes with only N TMS gates. This is shown schematically in Fig. 2(b) for $N = 1$ data modes. In other words, most of the operations are pushed to entangling ancilla modes (which generally requires $\sim M^2$ elementary gates). As an example, to protect a single data mode ($N = 1$) by a concatenated code that leverages M ancilla modes, the ancilla

modes themselves need to interact by some M mode Gaussian unitary, which can be done offline, but the data only needs to couple to *one* of the ancilla modes by a single TMS operation in the last step of encoding; see Fig. 2(b) for an illustration.

3.2 Example of code reduction

We give a few example code reductions (consequences of Lemma 1) by specifying the local symplectic matrices Λ_d and Λ_a discussed in Lemma 1.

In the two-mode case, we reduce the GKP-squeezed-repetition code, $\mathbf{S}_{\text{Sq-Rep}}^{[2]}$ of Eq. (4), to a GKP-TMS code. The local symplectic transform Λ_d is a single-mode squeezer $\Lambda_d = \text{diag}(\sqrt[4]{2}\lambda, 1/\sqrt[4]{2}\lambda)$ and $\Lambda_a = \Lambda_d^{-1}$. Applying Λ_a and Λ_d , we obtain TMS with gain $G = (\sqrt{2} + 1)/2$. Hence, the two-mode squeezed repetition code is equivalent to a TMS code of fixed gain $G = (\sqrt{2} + 1)/2$ and an ancilla rectangular GKP (with dimensions specified by the squeezing parameter λ).

Consider one data mode and two ancilla modes. We show that a downward staircase concatenation of the TMS code (as sketched in Fig. 2(d)) [20] $\mathbf{S}_{G_1, G_2}^{\text{down}} = (\mathbf{I}_2 \oplus \mathbf{S}_{G_2})(\mathbf{S}_{G_1} \oplus \mathbf{I}_2)$ is equivalent to an upward staircase concatenation $\mathbf{S}_{G_a, G_{12}}^{\text{up}} = (\mathbf{S}_{G_{12}} \oplus \mathbf{I}_2)(\mathbf{I}_2 \oplus \mathbf{S}_{G_a})$ with to be determined gains G_a and G_{12} ². The data mode is already in normal form, hence $\Lambda_d = \mathbf{I}_2$. The local symplectic transform on the ancilla modes is given by an inverse TMS transformation, $\Lambda_a = \mathbf{S}_{G_a}^{-1}$, with gain $G_a = G_2 G_1 / (1 + G_2(G_1 - 1))$. Applying Λ_d and Λ_a then reduces the staircase encoding to TMS between the first and second modes with gain $G_{12} = 1 + G_2(G_1 - 1)$.

As a final example, we reduce the three-mode squeezed-repetition code introduced in Ref. [10] (see sketch in Fig. 2(c)), which leads to cubic noise suppression for the output standard deviation. The squeezed-repetition code has encoding matrix

$$\mathbf{S}_{\text{Sq-Rep}}^{[3]} = \begin{pmatrix} 1/\lambda^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda^2 & 0 & -\lambda & 0 & 0 \\ 1 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/\lambda & 0 & -\lambda \\ \lambda^2 & 0 & \lambda^3 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\lambda \end{pmatrix}; \quad (5)$$

see Fig. 2(c) for a circuit diagram. By squeezing the data mode via $\Lambda_d = \mathbf{Sq}(\lambda(\lambda^2 + \lambda^4)^{1/4})$ and applying a two-mode ancilla transformation Λ_a , which is lengthy to show, we end up with TMS between the first and the second mode with gain $G_{12} = (\sqrt{1/\lambda^2 + 1} + 1)/2$.

In our analyses of this section, we have considered geometric mean as a performance metric. While such

²The phrase downward staircase refers to the fact that we couple the first and second modes, then the second and third modes etc., whereas an upward staircase starts from the bottom mode and goes to the top.

a choice is motivated by information-theoretical roots of the metric, some generalization of the results is possible. First, while the sufficiency and optimality of $\text{TMS}^\otimes(\mathcal{L})$ in Theorem 2 and Lemma 1 are derived with geometric mean of error as the metric, they hold for any metric that is invariant under local symplectic transformations. On the other hand, it is generally an open question to what extent some of the results here hold or approximately hold for non-Gaussian additive noise channels.

4 Code reduction implies no threshold for finite squeezing

We prove a no threshold theorem for finite squeezing that applies to general oscillator-to-oscillator codes based on Gaussian encoding. Interestingly, the no threshold theorem is a direct consequence of our general code reduction to TMS codes.

Consider iid AGN noise $\xi = (\xi_d, \xi_a)$ for the displacement errors on N data modes and M ancilla modes. After the encoding (and decoding) transformations $\mathbf{S}_{\text{enc}}$ ($\mathbf{S}_{\text{enc}}^{-1}$), the correlated noises are $(\mathbf{x}_d, \mathbf{x}_a) = \mathbf{S}_{\text{enc}}^{-1} \xi \sim \mathcal{N}(0, \mathbf{V}_x)$, where $\mathbf{V}_x = \sigma^2 \mathbf{S}_{\text{enc}}^{-1} \mathbf{S}_{\text{enc}}^{-\top}$. Since all encoding can be reduced to TMS up to local transformations via Lemma 1, we need only consider $\mathbf{V}_x$ as a direct sum of N correlated two-mode blocks (plus an identity block on the remaining $M - N$ ancilla modes), which has qq and pp correlations between data and ancilla arising from two-mode squeezing (but zero qp correlations). Therefore, to analyze general properties of the code, we can focus our attention on one data-ancilla mode pair (say, the i th mode pair) and one quadrature (say, the q quadrature) at a time; see Appendix I for further details.

Let q_{a_i} be the $(2i - 1)$ th element of $\mathbf{x}_a$ (where $i = 1, 2, \dots, N$) that is correlated with q_{d_i} of $\mathbf{x}_d$ via TMS with gain G_i , and let $\tilde{q}_{d_i} \equiv \tilde{q}_{d_i}(q_{a_i})$ be the estimation of the data noise given knowledge of the ancilla noise, which we can extract from, e.g., syndrome measurements. Although perfect knowledge of the ancilla noise is not generally available, we assume that it is in order to place an ultimate lower bound. Now from the corollary of Theorem 8.6.6 in Ref. [23], the estimation variance of a generic random variable X , given side information Y , is lower bounded by a function of the conditional differential entropy $S(X|Y)$ via $\mathbb{E}[(X - \tilde{X}(Y))^2] \geq \exp[2S(X|Y)]/2\pi e$. In our current setting, $S(q_{d_i}|q_{a_i}) = \ln\left(\frac{2\pi e \sigma^2}{2G_i - 1}\right)/2$, which is limited by the finite squeezing to correlate the noises (see Appendix I for a derivation). We point out that an equivalent relation holds for the momenta, p_{a_i} and

p_{d_i} . Therefore,

$$\mathbb{E} \left[(q_{d_i} - \tilde{q}_{d_i})^2 \right] \geq \frac{\sigma^2}{2G_i - 1}. \quad (6)$$

If the TMS has a finite squeezing level G_i , then having a larger number of ancilla modes (or an arbitrary ancilla state) will not further help error correction. This implies a universal no-threshold theorem for a wide variety of codes based on Gaussian encoding—including but not limited to GKP-stabilizer codes.

Theorem 3 (No threshold for finite squeezing). *For N data modes and arbitrary number of ancilla modes, the residual error for any oscillators-to-oscillators code using Gaussian encoding is lower bounded by*

$$\text{Tr}(\mathbf{V}_{\text{out}}) \geq \sum_{i=1}^N \frac{2\sigma^2}{2G_i - 1} \quad (7)$$

where $\mathbf{V}_{\text{out}}$ is the covariance matrix for the residual output error of the code, G_i are the two-mode squeezing gains of the encoding after code reduction [see Lemma 1 and Theorem 2], and σ^2 is the variance of the iid AGN channels.

Proof. The proof follows by summing over the individual variances (left hand side) of Eq. (6), which is less than or equal to the trace of the residual output covariance matrix, $\mathbf{V}_{\text{out}}$. The factor of two is due to the fact that the q and p quadratures of the i th mode contribute equally to the sum due to the structure of two-mode squeezing. ■

If we place a tolerance on the output error $\varepsilon \geq \text{Tr}(\mathbf{V}_{\text{out}})$, Theorem 3 implies that the (average) gain G must scale as $G \sim N\sigma^2/\varepsilon$ for the error to stay below tolerance, which is independent of the number of ancilla modes M used in the code. Thus, we cannot make ε arbitrarily small with a finite amount of squeezing even if we increase the number of ancilla modes; this is the essence of the no-threshold theorem for oscillator-to-oscillator codes.

Our proof follows from the general code reduction Theorem 2 (see also Lemma 1) and a simple, classical data-processing argument. Furthermore, our result has a broader scope than a similar no-threshold result of Ref. [12] based on GKP-stabilizer codes and maximum-likelihood decoding, as we do not require the ancilla modes to be prepared in GKP states and the decoding strategy does not enter into our proof. The only caveat here is the assumption of iid noise across all modes (likewise in [12]).

5 General GKP lattices and minimum mean square estimation

Before exploring code designs, we review general GKP lattices and derive the minimum mean square estimation (MMSE) that minimizes $\bar{\sigma}_{\text{RMS}}^2$ in Eq. (2).

5.1 GKP lattices

Consider a set of vectors $\{\boldsymbol{\lambda}_K^{\mathcal{L}}\}_{K=1}^{2N}$, where $\boldsymbol{\lambda}_K^{\mathcal{L}} \in \mathbb{R}^{2N}$, that generate a rigid phase-space lattice $\mathcal{L} \subset \mathbb{R}^{2N}$. We define the ‘stabilizers’ (formed by displacement operators) of the lattice as $S_K^{\mathcal{L}} \equiv D_{\boldsymbol{\lambda}_K^{\mathcal{L}}}$, such that

$$[S_K^{\mathcal{L}}, S_J^{\mathcal{L}}] = 0, \quad \forall K, J. \quad (8)$$

The commutator relation is equivalent to the condition $\boldsymbol{\lambda}_K^{\mathcal{L}\top} \boldsymbol{\Omega} \boldsymbol{\lambda}_J^{\mathcal{L}} = 2\pi n_{KJ}$, where $\boldsymbol{\Omega} = \mathbf{I}_N \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the N -mode symplectic form and $n_{KJ} \in \mathbb{Z}$ is an integer (see Appendix B and Appendix C). A lattice $\mathcal{L}$ with basis vectors $\{\boldsymbol{\lambda}_K^{\mathcal{L}}\}$ which satisfy this condition is known as a *symplectically integral* lattice [17]. By virtue of the commutator (8), the stabilizers form a group $\mathcal{S}_{\mathcal{L}} \equiv \langle S_1^{\mathcal{L}}, \dots, S_{2N}^{\mathcal{L}} \rangle$. We define the lattice state $|\mathcal{L}\rangle$ as the $+1$ eigenstate of all elements in $\mathcal{S}_{\mathcal{L}}$, i.e. ‘ $\mathcal{S}_{\mathcal{L}} |\mathcal{L}\rangle = |\mathcal{L}\rangle$ ’. This state is periodic in the $2N$ dimensional phase space of the modes. See Fig. 3 for an illustration of two-dimensional (symplectically integral) lattices.

From the lattice (column) vectors, one can construct a $2N \times 2N$ generator matrix,

$$\mathbf{M} = (\boldsymbol{\lambda}_1^{\mathcal{L}} \quad \boldsymbol{\lambda}_2^{\mathcal{L}} \quad \dots \quad \boldsymbol{\lambda}_{2N}^{\mathcal{L}}). \quad (9)$$

Then, the set of conditions $\boldsymbol{\lambda}_K^{\mathcal{L}\top} \boldsymbol{\Omega} \boldsymbol{\lambda}_J^{\mathcal{L}} = 2\pi n_{KJ}$ can be compactly written as

$$\mathbf{M}^{\top} \cdot \boldsymbol{\Omega} \cdot \mathbf{M} = 2\pi \mathbf{A}, \quad (10)$$

where $\mathbf{A}$ is an anti-symmetric matrix with only integer elements. We can generate the same lattice with different choices of basis vectors. For instance, given a generator matrix $\mathbf{M}$ that generates a lattice $\mathcal{L}$, one can choose a unimodular matrix $\mathbf{N}$ (i.e., integer matrix with $\det \mathbf{N} = 1$) such that $\mathbf{M}' = \mathbf{M}\mathbf{N}$ also generates $\mathcal{L}$.

We now make the distinction between computational GKP states and canonical GKP states precise. In general, if we want to encode a qudit with d -levels into a system of M oscillators, the encoded Hilbert space stabilized by $\mathcal{S}_{\mathcal{L}}$ will be related to the generator matrix $\mathbf{M}$ via $d = \sqrt{\det \mathbf{M}/(2\pi)^M}$ [3, 17, 18].³ In this paper, we are focusing on the $d = 1$ case which only supports a single code state. This is commonly referred to as canonical GKP state [11] (hence the moniker “canonical GKP lattice state” for generic lattices) or the sensor state in the square lattice case [24].

5.2 Minimum mean square error (MMSE) estimation

Here we consider MMSE for corrective displacements, which is constructed to minimize the RMS error of Eq. (2) (as RMS error is the square root of

³The factor $1/(2\pi)^M$ is due to our definition of the generator matrix $\mathbf{M}$, which differs from convention in Refs. [3, 17, 18] by a constant factor $\sqrt{2\pi}$.

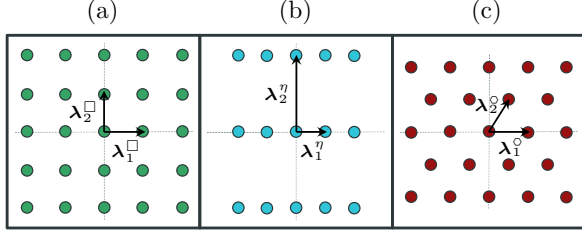


Figure 3: Illustration of two-dimensional lattices with basis vectors denoted: (a) Square, (b) Rectangular, (c) Hexagonal. Any two-dimensional symplectically integral lattice can be generated from the square lattice by rotation and squeezing.

mean square error). The encoding symplectic transform $\mathbf{S}_{\text{enc}}$ correlates the additive noise between the data and the ancilla. Let the covariance matrix of the AGN be $\mathbf{V}_\xi$. The error correlations are described by

$$\mathbf{V}_x = \mathbf{S}_{\text{enc}}^{-1} \mathbf{V}_\xi \mathbf{S}_{\text{enc}}^{-\top}. \quad (11)$$

The additive noise on the ancilla $\mathbf{x}_a$ can be extracted by measuring the stabilizers S_j^z . This leads to an error syndrome $\mathbf{s} \in [-\sqrt{\pi/2}, \sqrt{\pi/2}]^{2M}$, from which we can estimate the additive noise on the data $\mathbf{x}_d$. For a general lattice with generator matrix $\mathbf{M}$, we have $\mathbf{s} = R_{\sqrt{2\pi}}(\mathbf{M}^\top \Omega \mathbf{x}_a)$, where $R_{\sqrt{2\pi}}$ is the element-wise modulus of $\sqrt{2\pi}$ that associates the vector $\mathbf{M}^\top \Omega \mathbf{x}_a$ to the nearest lattice point within a region $[-\sqrt{\pi/2}, \sqrt{\pi/2}]^{2M}$ of that point [11]; see also Appendix D.

The joint probability density distribution (PDF) of the data and the error syndrome, $P(\mathbf{x}_d, \mathbf{s})$, is not a Gaussian distribution but a sum of Gaussian distributions. The MMSE minimizes $\bar{\sigma}_{\text{RMS}}^2$ in Eq. (2), and the estimator can be derived from the conditional distribution $P(\mathbf{x}_d | \mathbf{s})$ via $\mathbf{f}_{\text{MMSE}}(\mathbf{s}) = \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \mathbf{x}_d P(\mathbf{x}_d | \mathbf{s})$ leading to the following theorem (see Appendix G for a derivation).

Theorem 4. *For a GKP-stabilizer code with GKP lattice state $\mathcal{L}$ described by generator matrix $\mathbf{M}$, the minimum mean square estimation (MMSE) for an error syndrome $\mathbf{s}$ is given by*

$$\mathbf{f}_{\text{MMSE}}(\mathbf{s}) = - \frac{\sum_{\mathbf{n}} \mathbf{V}_d^{-1} \mathbf{V}_{da} (\mathbf{s} - \mathbf{n} \sqrt{2\pi}) e^{(\sqrt{2\pi} \mathbf{n}^\top \mathbf{V}_{d|a} \mathbf{s} - \pi \mathbf{n}^\top \mathbf{V}_{d|a} \mathbf{n})}}{\sum_{\mathbf{m}} e^{(\sqrt{2\pi} \mathbf{m}^\top \mathbf{V}_{d|a} \mathbf{s} - \pi \mathbf{m}^\top \mathbf{V}_{d|a} \mathbf{m})}}, \quad (12)$$

where $\mathbf{m}, \mathbf{n} \in \mathbb{Z}^{2M}$ sum over all integer vectors, and the matrices $\mathbf{V}_{da}$, $\mathbf{V}_d$ and $\mathbf{V}_a$ are defined through the equation

$$\begin{pmatrix} \mathbf{V}_d & \mathbf{V}_{da} \\ \mathbf{V}_{da}^\top & \mathbf{V}_a \end{pmatrix}^{-1} \equiv (\mathbf{I}_{2N} \oplus \mathbf{M}^\top \Omega) \mathbf{V}_x (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \Omega)^\top). \quad (13)$$

and $\mathbf{V}_{d|a} = \mathbf{V}_a - \mathbf{V}_{da}^\top \mathbf{V}_d^{-1} \mathbf{V}_{da}$.

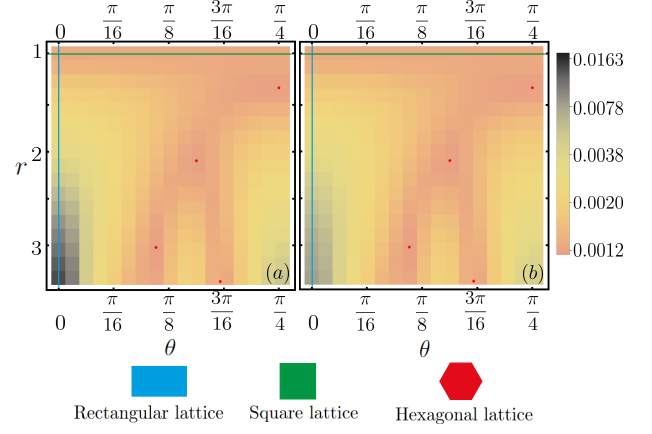


Figure 4: Output noise for a single-mode ($N = M = 1$) GKP stabilizer code. Input noise variance is $\sigma^2 = 10^{-2}$. We optimize the TMS gain G for each point (r, θ) . (a) RMS error $\bar{\sigma}_{\text{RMS}}^2$, (b) GM error $\bar{\sigma}_{\text{GM}}^2$. For the square lattice (green line), $\bar{\sigma}_{\text{RMS}}^2 = 1.25129(5) \times 10^{-3}$ and $\bar{\sigma}_{\text{GM}}^2 = 1.25129(5) \times 10^{-3}$. The four hexagonal lattice points (red dots) have the same output noises of $\bar{\sigma}_{\text{RMS}}^2 = 1.15575(5) \times 10^{-3}$ for RMS error and $\bar{\sigma}_{\text{GM}}^2 = 1.15575(5) \times 10^{-3}$ for GM error. Only the range $\theta \in [0, \pi/4]$ is considered due to symmetry; see Appendix E.

As an example, we provide the explicit application of Theorem 4 on GKP-TMS code in Appendix G.3.

Given the estimator $\mathbf{f}_{\text{MMSE}}(\mathbf{s})$ above, the residual noise covariance matrix for the data, $\mathbf{V}_{\text{out}}$, can be evaluated. In Appendix G.4, we show that the PDF of joint distribution $P(\mathbf{x}_d, \mathbf{s})$ is invariant under a change of lattice basis $\mathbf{M} \rightarrow \mathbf{M}\mathbf{N}$ where $\mathbf{N}$ is a unimodular matrix. It follows that the error correction performance with MMSE is invariant under a change of lattice basis as well.

If the noise level is much smaller than $\sqrt{2\pi}$, then $\mathbf{s} = R_{\sqrt{2\pi}}(\mathbf{M}^\top \Omega \mathbf{x}_a) \approx \mathbf{M}^\top \Omega \mathbf{x}_a$. In this case, the PDF is close to the Gaussian distribution. From Eq. (12), keeping the leading terms $\mathbf{n} = 0$ and $\mathbf{m} = 0$ for numerator and denominator, the estimator is close to linear estimation, $\mathbf{f}_{\text{Linear}}(\mathbf{s}) = -\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{s}$. However, unlike MMSE, linear estimation is not invariant under lattice basis transform. See Appendix G for technical details.

Both the linear and MMSE estimators assume knowledge of the covariance matrix of the Gaussian noise. While matrix multiplication and inversion are both required (only once) to derive the estimators, linear estimation involves no summations. On the contrary, MMSE estimation requires summation over two integer vectors of length $2M$. Fortunately, the convergence of the summation is exponential, which makes the evaluation possible, although the cost grows with the number of modes.

(θ, r)	$(\frac{\pi}{4}, \sqrt[4]{3})$	$(0.16\pi, 2.095)$	$(0.11\pi, 3.021)$	$(0.18\pi, 3.385)$
N	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$

Table 1: Equivalent representations of a hexagonal lattice. N is a unimodular matrix that relates the lattice basis vectors.

6 Optimal single-mode code

Since all single-mode canonical lattice states can be generated by local symplectic transformations on the square grid (see Appendix C), Theorem 2 and Corollary 2.1 immediately imply the following.

Theorem 5. *For single-mode data and ancilla undergoing iid AGN, the TMS code $\text{TMS}(\mathcal{L})$, with gain G and an ancillary lattice $\mathcal{L} \subset \mathbb{R}^2$ generated from the square grid by a local symplectic transformation, is the optimal GKP-stabilizer code in terms of geometric mean error.*

Therefore, to obtain the best two-mode GKP-stabilizer code, one simply needs to optimize the local Gaussian unitary and TMS gain, alongside choosing an optimal estimator $\mathbf{f}$ for the classical decoding strategy. Deriving the best estimator $\mathbf{f}$ to minimize the GM error is nontrivial. On the other hand, we can obtain an upper bound on the GM error from the RMS error since $\bar{\sigma}_{\text{RMS}}^2 \geq \bar{\sigma}_{\text{GM}}^2$. Suppose that one can optimize the RMS error (which is possible with MMSE of Theorem 4) while showing that the GM error is very close to $\bar{\sigma}_{\text{RMS}}^2$; then, this is an indication that the code is optimal for GM error as well.

For the single-mode case, any ancillary lattice $\mathcal{L}$ can be generated by applying a single-mode Gaussian unitary U_{Λ_a} on the square GKP lattice (see Fig. 3 for an illustration). The canonical square GKP lattice can be written down in the q or p quadrature bases as

$$|\square\rangle \propto \sum_{n \in \mathbb{Z}} |q = n\ell\rangle = \sum_{n \in \mathbb{Z}} |p = n\ell\rangle, \quad (14)$$

which is a translation invariant square lattice in the single-mode phase space $\mathbb{R}^2$ with period $\ell = \sqrt{2\pi}$. Moreover, any single-mode Gaussian unitary can be described by a symplectic transform, with the decomposition $\mathbf{R}(\phi)\mathbf{S}\mathbf{q}(r)\mathbf{R}(\theta)$, where $\mathbf{R}(\varphi)$ is a 2×2 rotation matrix and $\mathbf{S}\mathbf{q}(r) \equiv \text{diag}[r, 1/r]$ is single-mode squeezing. Due to the symmetry of the AGN, the effect of the last phase rotation $\mathbf{R}(\phi)$ applied on the lattice will not change the performance, therefore we parameterize the transform as $\Lambda_a = \mathbf{S}\mathbf{q}(r)\mathbf{R}(\theta)$ such that $|\mathcal{L}\rangle = U_{\Lambda_a} |\square\rangle$. As examples, a rectangular GKP state (Fig. 3b) is given by $\theta = 0$ and $r > 0$, and a hexagonal GKP state (Fig. 3c) is given by $\theta = \pi/4$ and $r_o = \sqrt[4]{3}$.

In Fig. 4(a), we plot the contour of the RMS error $\bar{\sigma}_{\text{RMS}}^2$ for an MMSE decoder optimized over the TMS gain G for each point (r, θ) . We find four equal

minimum points for $\bar{\sigma}_{\text{RMS}}^2$, which turn out to be equivalent lattice representations of the hexagonal lattice as listed in Table 1. Meanwhile, the square lattice has $r = 1$ with θ arbitrary (represented by the green line); the rectangular lattice has $\theta = 0$ and r changes the shape of the rectangle (represented by the blue line). The square and rectangular lattices are strictly sub-optimal.

In Fig. 4(b), we plot the GM error $\bar{\sigma}_{\text{GM}}^2$ in (r, θ) parameter space for the same optimized gain values of Fig. 4(a). The two subplots are very similar, with some deviations at the left-bottom corner due to the large squeezing of a rectangular lattice which induces asymmetry between q and p quadratures. The hexagonal lattices again minimize the output noise. Moreover, for the hexagonal lattices, $\bar{\sigma}_{\text{GM}}^2 \approx \bar{\sigma}_{\text{RMS}}^2$ up to our numerical precision, which is a strong indicator that—even if we minimize the GM error instead—the hexagonal lattice is still optimal.

7 Multimode codes

It turns out that all canonical (or “self-dual”) lattices can be generated from canonical square GKP (consequence of Corollary 1 in Ref. [18]). Therefore, Theorem 5 can generalize to the multimode case. However, optimization is still challenging since $|\text{Sp}(2M, \mathbb{R})| = 2M^2 + M$ parameters need to be optimized in general. The search for an optimal GKP lattice suggested by Theorem 2 and Corollary 2.1 is therefore difficult. Nevertheless, as we will show in this section with a few examples, going to higher-dimensional lattices may indeed improve the performance of oscillators-to-oscillators codes.

Below, we present our results on multi-mode codes. We first give a lower bound on the output noise for a general multimode GKP code, then discuss break-even points. Finally, evaluate the performance of $N = M = 2$ and $N = 1, M = 2$ multimode GKP stabilizer codes for various lattice configurations (e.g., square, hexagonal, and D4) and estimation strategies (e.g., linear estimation versus MMSE).

7.1 Lower bound and AGN break-even point

By information theoretic arguments (see Appendix H), we are able to find lower bounds for the RMS and GM errors, $\bar{\sigma}_{\text{RMS}}$ and $\bar{\sigma}_{\text{GM}}$, for a general multimode GKP code, with M ancilla modes and N data modes, in terms of the variances σ_i^2 of

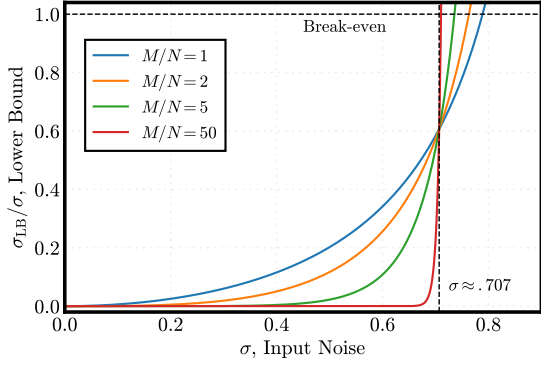


Figure 5: Lower bound σ_{LB} [Eq. (15)] on output noise versus initial noise σ for multimode GKP codes with increasing M/N . A transition occurs at $\sigma = 1/\sqrt{2} \approx .707$, which provides an upper bound for the maximal, correctable amount of AGN—i.e., $\sigma^* \leq 1/\sqrt{2}$, where σ^* is the AGN error break-even point for multimode GKP codes.

the AGN channels $\bigotimes_i \mathcal{N}_{\sigma_i}$ (Theorem 16). In particular, for iid AGN, we show that (Corollary 16.1) $\bar{\sigma}_{RMS} \geq \bar{\sigma}_{GM} \geq \sigma_{LB}$, where

$$\sigma_{LB} \equiv \frac{1}{\sqrt{e}} \left(\frac{\sigma^2}{1 - \sigma^2} \right)^{\frac{N+M}{2N}}. \quad (15)$$

For single-layer codes ($N = M$), there is at best quadratic error suppression, exactly similar to the $N = M = 1$ GKP codes discussed in Ref. [11]. Higher order error suppression can be obtained for codes with $M > N$ —with the output standard deviation scaling as $\sim \sigma^{1+\frac{M}{N}}$ per Eq. (15)—in agreement with the results on concatenated codes (for $N = 1$ and $M > 1$) found in Refs. [10, 11, 20, 15].

In Fig. 5, we plot the ratio σ_{LB}/σ versus the initial AGN noise σ , for increasing M/N ; $\sigma_{LB}/\sigma = 1$ corresponds to the break-even point. We observe a sharp transition occurring near $\sigma = 1/\sqrt{2}$ which defines an upper bound on the AGN error break-even point σ^* for general multimode GKP codes—i.e., $\sigma^* \leq 1/\sqrt{2}$. Thus, for $\sigma \gtrsim 1/\sqrt{2}$, we expect no gain to be had from such codes. This is consistent with the upper bound on the energy-unconstrained quantum capacity of an AGN channel [25] (see also Ref. [20] and Lemma 15 of Appendix H) $C_Q(\mathcal{N}_\sigma) \leq \max[0, \log_2(\frac{1-\sigma^2}{\sigma^2})]$, which vanishes as $\sigma \rightarrow 1/\sqrt{2}$. Furthermore, since a bosonic pure-loss channel with transmittance $\eta \in [0, 1]$ can be converted via pre-amplification to an AGN channel with variance $\sigma^2 = 1 - \eta$ [25], the AGN break-even point σ^* then corresponds to a pure-loss transmissivity $\eta^* = 1 - \sigma^{*2} \geq 1/2$.

The multimode GKP circuit of Fig. 1 corresponds to an (N mode) additive non-Gaussian noise channel $\tilde{\mathcal{N}}$ for the data, with output GM error $\bar{\sigma}_{GM}$. We find a lower bound for the quantum capacity of the channel $\tilde{\mathcal{N}}$ (see Lemma 14 in Appendix H), $C_Q(\tilde{\mathcal{N}}) \geq \max[0, N \log_2(\frac{1}{e\bar{\sigma}_{GM}^2})]$. Assuming break-even $\bar{\sigma}_{GM} = \sigma$, this in turn implies a lower bound for

$\mathcal{L} \backslash \sigma$	0.1	0.2	0.5
Square	0.0354	0.120	0.490
Hexagonal	0.0340	0.117	0.489
D4	0.0322	0.112	0.487

Table 2: Output geometric mean error $\bar{\sigma}_{GM}$ for a $N = M = 2$ multimode GKP stabilizer code. Various lattices $\mathcal{L}$ (Square, Hexagonal, D4) and three representative values of AGN ($\sigma = 0.1, 0.2, 0.5$) are considered. MMSE was used. See Fig. 6.

the AGN error break-even point, $\sigma^* \geq 1/\sqrt{e} \approx .607$ (thus, $\eta^* \leq 1 - 1/e \approx .632$). Hence, the break-even point σ^* (η^*) for multimode GKP codes lies within $.607 \leq \sigma^* \leq .701$ ($.5 \leq \eta^* \leq .632$). In the next section, we numerically find break-even points for multimode ($N = M = 2$) GKP stabilizer codes and MMSE estimation with AGN error break-even points near $1/\sqrt{e} \approx .607$. We remark that linear estimation strategies have a lower break-even point of .558 [11]. See Fig. 6 and below.

7.2 Two data modes and two ancilla modes ($N = M = 2$)

We compare the performance of different initial lattices $\mathcal{L}$ for a single-layer, multimode ($N = M = 2$) GKP-TMS code $\text{TMS}^{\otimes 2}(\mathcal{L})$. The encoding (decoding) is given by two TMS operations, with each TMS operation coupling one data mode to one ancilla mode; see Fig. 1. The gain values of the TMS operations are (numerically) chosen to minimize the RMS error $\bar{\sigma}_{RMS}^2 = \text{Tr}\{\mathbf{V}_{\text{out}}\}/4$. We consider MMSE estimation, which optimizes the RMS error and thus sets an upper bound on the optimal GM error. We also consider linear estimation with initial square GKP states, which was analyzed in Ref. [11].

In our analysis, we choose three initial GKP lattice states: a direct product of square GKP states, a direct product of hexagonal GKP states, and a D4 lattice which can be generated from two-square GKP states by a two-mode symplectic transformation [see Eq. (69) in Appendix C]. The D4 GKP state is necessarily entangled; hence, the D4 TMS code $\text{TMS}^{\otimes 2}(D_4)$ is a genuine multimode code. On the other hand, since the TMS encoding (decoding) operates on individual data modes and the additive noises are independent, the multimode square and hexagonal TMS codes, $\text{TMS}^{\otimes 2}(\square)$ and $\text{TMS}^{\otimes 2}(\circ)$, are simple extensions of their single-mode counterparts, i.e. $\text{TMS}^{\otimes 2}(\square) \simeq \text{TMS}(\square) \otimes \text{TMS}(\square)$ and $\text{TMS}^{\otimes 2}(\circ) \simeq \text{TMS}(\circ) \otimes \text{TMS}(\circ)$.

We report the output GM error $\bar{\sigma}_{GM} = \sqrt[4]{\det \mathbf{V}_{\text{out}}}$ of the codes in Fig 6. As shown in the figure, the D4 TMS code $\text{TMS}^{\otimes 2}(D_4)$ performs better than the square and hexagonal TMS codes $\text{TMS}^{\otimes 2}(\square)$ and $\text{TMS}^{\otimes 2}(\circ)$, at least for the few data that we gen-

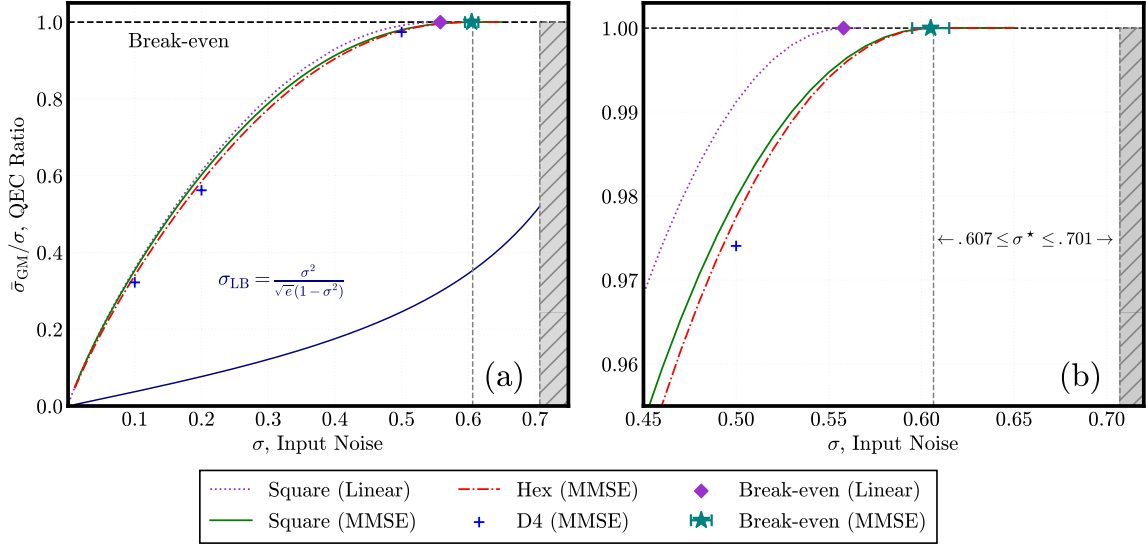


Figure 6: Quantum error correction (QEC) ratio between output noise and input noise of a single-layer, multimode ($N = M = 2$) GKP TMS code $\text{TMS}^{\otimes 2}(\mathcal{L})$ for different lattices $\mathcal{L}$ (Square, Hexagonal, D4) and estimation strategies (linear, MMSE). The square code with linear estimation (dotted purple) agrees with the original GKP TMS code presented in Ref. [11]. Grey hatched region is forbidden by information theoretic arguments. While the break-even point of the linear estimation with square lattice $\sigma \simeq 0.558$ (purple diamond), the break-even point of the MMSE estimation $\sigma = 0.605(5)$ for both square and hexagonal lattices (green star). For D4 lattice, we narrowed the break-even point to $0.60 - 0.61$, which is consistent with square and hexagonal lattices.

erated (corresponding to $\sigma = .1, .2, .5$; see also Table 2), however we expect similar findings for $\sigma \ll 1$. Indeed, using $\text{TMS}^{\otimes 2}(\square)$ as a benchmark [11], we find the relative performance of other lattices to be better for smaller initial noise σ . For instance, at $\sigma = .01$, we find that $\text{TMS}^{\otimes 2}(\circ)$ outperforms $\text{TMS}^{\otimes 2}(\square)$ by a relative difference of about 5.25%. For $\sigma = .1$, $\text{TMS}^{\otimes 2}(\circ)$ achieves a relative difference of about 3.95%, whereas $\text{TMS}^{\otimes 2}(D_4)$ achieves a relative difference of about 9.04%. For $\sigma = .2$, the relative difference is 2.5% for $\text{TMS}^{\otimes 2}(\circ)$ and 6.67% for $\text{TMS}^{\otimes 2}(D_4)$. Finally, for $\sigma = .5$ (which is near the break-even point $\sigma \approx .607$), the relative difference is .20% for $\text{TMS}^{\otimes 2}(\circ)$ and .61% for $\text{TMS}^{\otimes 2}(D_4)$. Numerical values of $\bar{\sigma}_{\text{GM}}$ for the different lattices and $\sigma = .1, .2, .5$ are displayed in Table 2.

We observe that single-layer ($N = M$) codes using MMSE—regardless of the initial lattice $\mathcal{L}$ being hexagonal or square—have a break-even point $\sigma_{\text{MMSE}}^* \approx .605(5)$ (teal star in Fig. 6), whereas linear estimation on square lattice has a break-even point of $\sigma_{\text{lin}}^* \approx .558$ [11] (purple diamond). For D4 lattice, we narrowed the break-even point to between 0.60 and 0.61, which is consistent with σ_{MMSE}^* . Indeed, the break-even point is likely universal for all lattices with MMSE decoding. The value $\sigma_{\text{MMSE}}^* \approx .605(5)$ for MMSE agrees with the lower bound on the break-even point ($.607 \leq \sigma^* \leq .701$) for general GKP codes discussed in the previous section. It is an open question whether this can be pushed further or not.

7.3 One data mode and two ancilla modes ($N = 1$ and $M = 2$)

In Sec. 3.1, we showed that general concatenation can be reduced to ancilla preparation and then presented examples of this reduction with one data mode and two ancilla modes ($N = 1, M = 2$) in Sec. 3.2. Indeed, the reduction of encoding shows that a single TMS operation between the lone data mode and only one of the ancilla modes is required (see Fig. 2(b)). The other ancilla mode—which need not directly interact with the data—is in general entangled with the ancilla that interacts with the data. Code optimization then reduces to optimizing the ancillary GKP lattice state and the TMS strength.

Due to the numerical challenge in two-mode lattice optimization, we focus on the TMS concatenation code and compare a few lattices, similar to Sec. 7.2. We examine the concatenated GKP-TMS code as it represents the best-known code for $N = 1$ and $M = 2$ (better than the squeezed-repetition code [20]). In what follows, we consider an “upward staircase” TMS concatenation, in which two ancillary modes (prepared in various lattices) are first coupled by a TMS operation of gain G_2 followed by a TMS operation of gain G_1 that couples the lone data mode to only one of the ancilla; see the inset of Fig. 7 for an illustration.⁴ Consequently, we can interpret the code as a scenario where a lone data mode is coupled to a

⁴It is worth noting that this “upward staircase” concatenation is equivalent to a “downward staircase” concatenation [Fig. 2(d)], as demonstrated in Sec. 3.2.

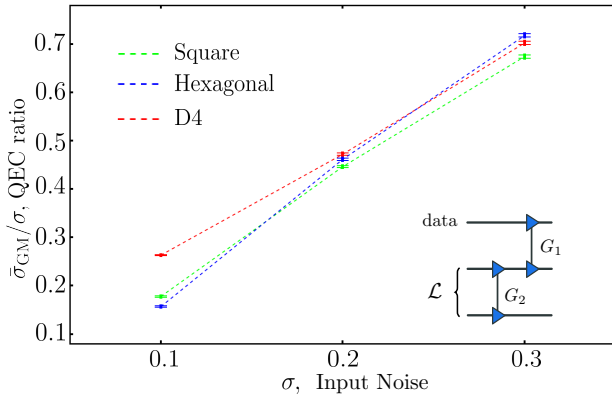


Figure 7: Quantum error correction (QEC) ratio of different lattices for a 1 data + 2 ancilla concatenated TMS code. The fences represent error bars of the local minima. Inset shows the encoding where G_1 and G_2 are the TMS gains.

σ	0.1	0.2	0.3
Square	(18.9, 1.14)	(3.43, 1.20)	(1.92, 1.20)
Hexagonal	(22.5, 2.04)	(3.20, 1.11)	(1.75, 1.11)
D4	(8.46, 1.21)	(3.21, 1.04)	(1.72, 1.01)

Table 3: Optimized two-mode squeezing gains (G_1, G_2) for different input noise σ . G_1 characterizes the two-mode squeezing between the data and ancilla, while G_2 characterizes the squeezing that correlates the two-mode lattices. See Fig. 7 inset.

correlated two-mode lattice state. To provide a more comprehensive analysis, we numerically optimize over the TMS gains G_1, G_2 for three different types of initial ancillary lattices: square, hexagonal and D4.

We consider three different levels of initial noise $\sigma = 0.1, 0.2, 0.3$ and numerically optimize the gains (G_1, G_2) to minimize $\bar{\sigma}_{\text{GM}}$. The quantum error correction (QEC) ratio $\bar{\sigma}_{\text{GM}}/\sigma$ is plotted in Fig. 7, and the corresponding optimal values of gain parameters are shown in Table 3. With the additional TMS gain $G_2 > 1$, the product of two square or hexagonal lattices becomes a four-dimensional lattice entangled between the two ancillary modes. Moreover, the four dimensional lattice generated in such way can outperform D4 lattice, especially for the case $\sigma = 0.1$ case. Although there are always caveats with numerical optimization, such results indicate that high-dimensional GKP lattice design is very much an open problem for GKP-stabilizer codes.

8 Heterogeneous noise case

We attempt to make progress towards generalizing our findings on iid AGN to more generic sources of AGN. Consider a generic N mode AGN channel $\mathcal{N}_{\mathbf{Y}}$, where $\mathbf{Y}$ is a $2N \times 2N$ noise matrix that may contain correlations. From Ref. [20], any correlated Gaussian noises can be reduced to a product of het-

erogeneous AGN channel by symplectic transformations, i.e. $\mathcal{N}_{\mathbf{Y}} \rightarrow \bigotimes_i \mathcal{N}_{\sigma_i}$, where σ_i^2 is the variance of the i th mode. It is thus sufficient to examine independent AGNs with different variances, for which we posit the following conjecture.

Conjecture. Consider N data modes and N ancilla modes undergoing heterogeneous AGN ($\bigotimes_i \mathcal{N}_{\sigma_i}$). Then, up to unitary data processing and ancilla preparation, two- (one data and one ancilla) and four-mode interactions (two data and two ancilla) are sufficient for encoding.

To promote support for this claim, consider a generic symplectic encoding matrix $\mathbf{S}_{\text{enc}}$. Observe that the Gaussian channel $\mathcal{N}_{\mathbf{Y}'} = \mathcal{U}_{\mathbf{S}_{\text{enc}}}^{-1} \circ (\bigotimes_i \mathcal{N}_{\sigma_i}) \circ \mathcal{U}_{\mathbf{S}_{\text{enc}}}$ is an AGN channel with noise matrix $\mathbf{Y}' = \mathbf{S}_{\text{enc}}^{-1} (\bigoplus_i \sigma_i \mathbf{I}_2) \mathbf{S}_{\text{enc}}^{-\top}$, which has diagonal blocks corresponding to local noises and off diagonal blocks corresponding to (generally multimode) correlated noises. By the results of Ref. [26] (see also Ref. [27]), there exists local symplectic transformations $\mathbf{\Lambda}_d, \mathbf{\Lambda}_a \in \text{Sp}(2N, \mathbb{R})$ that condense the correlation blocks to elementary two (one data and one ancilla) and four (two data and two ancilla) mode units. Since the initial noise channels are independent, the correlation units originate from the encoding, and it certainly seems plausible that two- and four-mode interactions are sufficient to generate them. However, we do not find this argument strong enough to unequivocally validate the conjecture.

Though we have inferred that a generic $2N$ mode encoding (encoded in $\mathbf{S}_{\text{enc}}$) reduces to “local” two- and four-mode interactions between data and ancillae, our reduction is not constructive, in the sense that it does not give us which particular interactions to use (such as the TMS of Theorem 2). Moreover, as stated, our reduction in the heterogeneous case is only expected to hold when the number of data modes equals the number of ancilla modes ($N = M$) and thus does not straightforwardly apply to codes where the number of ancilla modes is greater than the number of data modes (such as concatenated codes).

9 Analyses on finite-squeezing

A GKP state $|\mathcal{L}\rangle$ has support on the entire phase space of the modes and thus has infinite energy. We can regularize the state by confining the energy to a ball of radius $\sim \Delta^{-1}$ in phase space via

$$|\mathcal{L}(\Delta)\rangle \propto \exp\left(-\frac{\Delta^2 \hat{\mathbf{r}}^\top \hat{\mathbf{r}}}{2}\right) |\mathcal{L}\rangle; \quad (16)$$

see also Refs. [25, 28, 11, 29]. Note that $\hat{\mathbf{r}}^\top \hat{\mathbf{r}}/2 = \sum_{i=1}^N \hat{n}_i + N/2$. The regularizer (or envelope operator) has a nice form in the displacement operator basis,

$$\exp\left(-\frac{\Delta^2 \hat{\mathbf{r}}^\top \hat{\mathbf{r}}}{2}\right) \propto \int d\boldsymbol{\mu} \exp\left(-\frac{\boldsymbol{\mu}^2}{4 \tanh(\Delta^2/2)}\right) D_{\boldsymbol{\mu}}. \quad (17)$$

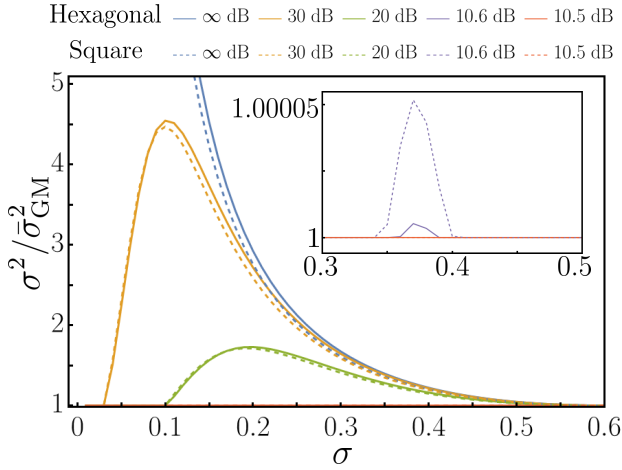


Figure 8: Quantum error correction (QEC) gain $\sigma^2/\bar{\sigma}_{\text{GM}}^2$ versus input noise σ for finite GKP squeezing in dB $s_{\text{GKP}} = (10.5, 10.6, 20, 30, \infty)$. Solid lines correspond to hexagonal lattice. Dashed lines correspond to square lattice.

We can thus think of the regularization procedure as coherently applying displacements (with a Gaussian envelope) on the GKP state. To simplify the analysis, we can twirl the regularized state [28] to generate an incoherent GKP state, $\bigotimes_{i=1}^N \mathcal{N}_{\sigma_{\text{GKP}}}(|\mathcal{L}\rangle\langle\mathcal{L}|)$, where we define the finite GKP noise per mode $\sigma_{\text{GKP}}^2 = \tanh(\Delta^2/2)$; the relation between finite squeezing and the variance of the GKP noise is $s_{\text{GKP}} = -10 \log_{10}(2\sigma_{\text{GKP}}^2)$. This twirling is not physical as it requires an infinite amount of energy and is only used for computational convenience. We use this noise model to approximate finite GKP noise in all of our calculations. Specifically, we assume that the GKP ancilla modes as well as the GKP states used for measurements are all noisy. We thus have two layers of finite GKP squeezing noise which enter into the analysis in slightly different ways as discussed below.

With the finite-squeezed GKP states as encoding ancilla and measurement ancilla, the final covariance matrix defined in Eq. (13) changes to

$$\begin{pmatrix} \mathbf{V}_d & \mathbf{V}_{da} \\ \mathbf{V}_{da}^T & \mathbf{V}_a \end{pmatrix}^{-1} \equiv (\mathbf{I}_{2N} \oplus \mathbf{M}^T \mathbf{\Omega}) \mathbf{V}_x (\mathbf{I}_{2N} \oplus (\mathbf{M}^T \mathbf{\Omega})^T) + \mathbf{0}_{2N} \oplus \sigma_{\text{GKP}}^2 (\mathbf{M}^T \mathbf{M} + \mathbf{I}_{2M}). \quad (18)$$

The GKP noise proportional to $\mathbf{M}^T \mathbf{M}$ originates from the GKP lattice ancilla whereas the GKP noise proportional to the identity comes from noisy measurements; see Appendix J for a derivation. We thus see that all computations from previous sections carry over (e.g., the MMSE formula Eq. (12) applies) with the updated covariance matrices to incorporate GKP noise.

In Fig. 8, we plot the QEC gain $\sigma^2/\bar{\sigma}_{\text{GM}}^2$ versus the input AGN σ for hexagonal and square GKP with finite GKP squeezing. We assume MMSE estimation.

The functional behaviors are similar to that presented in Ref. [11] for a square GKP and linear estimation. Though, an important distinction here is that we find a break-even squeezing of around $s_{\text{dB}} = 10.5\text{dB}$, below which there is no QEC gain, whereas linear estimation leads to a higher value of 11 dB [11]. Note that at each squeezing level, when $\sigma \approx \sigma_{\text{GKP}}$, the GKP finite-squeezing noise significantly contributes to the noise budget, and the square lattice is slightly better than the hexagonal lattice due to a smaller $\det(\mathbf{M}^T \mathbf{M} + \mathbf{I}_{2M})$, with values 2 for square and $2 + 4/\sqrt{3}$ for hexagonal.

10 Discussions and conclusions

In this paper, we derived the optimal form of GKP-stabilizer codes—TMS encoding on GKP ancilla with general lattices—for homogenous input noise. In the case of single-mode data and ancilla, we identified the GKP-TMS code with a two-dimensional hexagonal GKP ancilla to be optimal. For higher dimensions, we found D4 lattices to be superior to lower dimensional lattices. We were also able to prove a universal no-threshold theorem for *all* oscillators-to-oscillators codes based on Gaussian encodings and showed that the continuous errors on the data cannot be made arbitrarily small without an infinite amount of squeezing. This is not too surprising because these errors are intrinsically analog and may only be suppressed with continuous variable resources (e.g., squeezing).

Before closing, a few open questions are worth mentioning. We expect the D4 lattice to perform well and thus picked it for benchmarking, however there might be better lattices in four or higher dimensions. Searching for good lattices in higher dimension can in general be challenging, as the number of free parameters grow with the number of modes quadratically. Although we derived the minimize mean square estimator (MMSE), the optimal estimator for minimizing the geometric mean error is unknown (to our knowledge). We narrowed the range of optimal break-even point of additional noise level down to the range $.607 \leq \sigma^* \leq .701$ from quantum capacity bounds and found MMSE reaching the lower end of this range; the actual optimal break-even point is an open problem, although we expect it to be closer to the lower end. Finally, regarding heterogeneous noise sources, we have conjectured that two- and four-mode interactions are sufficient for general (Gaussian) encodings, however we do not have definitive nor constructive proof of this claim. Ideal code design for heterogeneous noise sources is thus open; dedicated numerical studies may be required to address such.

Acknowledgements

This project is supported by the Defense Advanced

Research Projects Agency (DARPA) under Young Faculty Award (YFA) Grant No. N660012014029. QZ also acknowledges support from NSF CAREER Award CCF-2142882 and National Science Foundation (NSF) Engineering Research Center for Quantum Networks Grant No. 1941583.

References

- [1] A. R. Calderbank and Peter W. Shor. “Good quantum error-correcting codes exist”. *Phys. Rev. A* **54**, 1098–1105 (1996).
- [2] Andrew Steane. “Multiple-particle interference and quantum error correction”. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **452**, 2551–2577 (1996).
- [3] Daniel Gottesman, Alexei Kitaev, and John Preskill. “Encoding a qubit in an oscillator”. *Phys. Rev. A* **64**, 012310 (2001).
- [4] A. Romanenko, R. Pilipenko, S. Zorzetti, D. Frolov, M. Awida, S. Belomestnykh, S. Posen, and A. Grassellino. “Three-dimensional superconducting resonators at $t < 20$ mk with photon lifetimes up to $\tau = 2$ s”. *Phys. Rev. Applied* **13**, 034032 (2020).
- [5] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, SM Girvin, Liang Jiang, et al. “Extending the lifetime of a quantum bit with error correction in superconducting circuits”. *Nature* **536**, 441–445 (2016).
- [6] VV Sivak, A Eickbusch, B Royer, S Singh, I Tsioutsios, S Ganjam, A Miano, BL Brock, AZ Ding, L Frunzio, et al. “Real-time quantum error correction beyond break-even” (2022). [arXiv:quant-ph/2211.09116](https://arxiv.org/abs/2211.09116).
- [7] Nithin Raveendran, Narayanan Rengaswamy, Filip Rozpędek, Ankur Raina, Liang Jiang, and Bane Vasić. “Finite rate QLDPC-GKP coding scheme that surpasses the CSS Hamming bound”. *Quantum* **6**, 767 (2022).
- [8] Filip Rozpędek, Kyungjoo Noh, Qian Xu, Saikat Guha, and Liang Jiang. “Quantum repeaters based on concatenated bosonic and discrete-variable quantum codes”. *npj Quantum Inf.* **7**, 1–12 (2021).
- [9] Christopher Chamberland, Kyungjoo Noh, Patricio Arrangoiz-Arriola, Earl T Campbell, Connor T Hann, Joseph Iverson, Harald Putterman, Thomas C Bohdanowicz, Steven T Flammia, Andrew Keller, et al. “Building a fault-tolerant quantum computer using concatenated cat codes”. *PRX Quantum* **3**, 010329 (2022).
- [10] Kyungjoo Noh, SM Girvin, and Liang Jiang. “Encoding an oscillator into many oscillators” (2019). [arXiv:1903.12615](https://arxiv.org/abs/1903.12615).
- [11] Kyungjoo Noh, S. M. Girvin, and Liang Jiang. “Encoding an Oscillator into Many Oscillators”. *Phys. Rev. Lett.* **125**, 080503 (2020).
- [12] Lisa Häggli and Robert König. “Oscillator-to-oscillator codes do not have a threshold”. *IEEE Trans. Inf. Theory* **68**, 1068–1084 (2021).
- [13] Yijia Xu, Yixu Wang, En-Jui Kuo, and Victor V Albert. “Qubit-Oscillator Concatenated Codes: Decoding Formalism and Code Comparison”. *PRX Quantum* **4**, 020342 (2023).
- [14] Quntao Zhuang, John Preskill, and Liang Jiang. “Distributed quantum sensing enhanced by continuous-variable error correction”. *New Journal of Physics* **22**, 022001 (2020).
- [15] Boyu Zhou, Anthony J. Brady, and Quntao Zhuang. “Enhancing distributed sensing with imperfect error correction”. *Phys. Rev. A* **106**, 012404 (2022).
- [16] Bo-Han Wu, Zheshen Zhang, and Quntao Zhuang. “Continuous-variable quantum repeaters based on bosonic error-correction and teleportation: architecture and applications”. *Quantum Science and Technology* **7**, 025018 (2022).
- [17] Baptiste Royer, Shraddha Singh, and S.M. Girvin. “Encoding Qubits in Multimode Grid States”. *PRX Quantum* **3**, 010335 (2022).
- [18] Jonathan Conrad, Jens Eisert, and Francesco Arzani. “Gottesman-Kitaev-Preskill codes: A lattice perspective”. *Quantum* **6**, 648 (2022).
- [19] Julien Niset, Jaromír Fiurášek, and Nicolas J. Cerf. “No-Go Theorem for Gaussian Quantum Error Correction”. *Phys. Rev. Lett.* **102**, 120501 (2009).
- [20] Jing Wu and Quntao Zhuang. “Continuous-variable error correction for general gaussian noises”. *Phys. Rev. Applied* **15**, 034073 (2021).
- [21] Alonso Botero and Benni Reznik. “Modewise entanglement of Gaussian states”. *Phys. Rev. A* **67**, 052311 (2003).
- [22] Ben Q. Baragiola, Giacomo Pantaleoni, Rafael N. Alexander, Angela Karanjai, and Nicolas C. Menicucci. “All-Gaussian Universality and Fault Tolerance with the Gottesman-Kitaev-Preskill Code”. *Phys. Rev. Lett.* **123**, 200502 (2019).
- [23] Thomas M. Cover and Joy A. Thomas. “Elements of information theory”. John Wiley & Sons. (2006). 2 edition.
- [24] Kasper Duivenvoorden, Barbara M. Terhal, and Daniel Weigand. “Single-mode displacement sensor”. *Phys. Rev. A* **95**, 012305 (2017).

- [25] Kyungjoo Noh, Victor V Albert, and Liang Jiang. “Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes”. *IEEE Transactions on Information Theory* **65**, 2563–2582 (2018).
- [26] Michael M Wolf. “Not-so-normal mode decomposition”. *Phys. Rev. Lett.* **100**, 070505 (2008).
- [27] Filippo Caruso, Jens Eisert, Vittorio Giovannetti, and Alexander S Holevo. “Multi-mode bosonic Gaussian channels”. *New J. Phys.* **10**, 083030 (2008).
- [28] Kyungjoo Noh and Christopher Chamberland. “Fault-tolerant bosonic quantum error correction with the surface-gottesman-kitaev-preskill code”. *Phys. Rev. A* **101**, 012316 (2020).
- [29] Baptiste Royer, Shraddha Singh, and S. M. Girvin. “Stabilization of Finite-Energy Gottesman-Kitaev-Preskill States”. *Phys. Rev. Lett.* **125**, 260509 (2020).
- [30] Samuel L Braunstein. “Squeezing as an irreducible resource”. *Phys. Rev. A* **71**, 055801 (2005).
- [31] Michael Reck, Anton Zeilinger, Herbert J Bernstein, and Philip Bertani. “Experimental realization of any discrete unitary operator”. *Phys. Rev. Lett.* **73**, 58 (1994).
- [32] Alessio Serafini. “Quantum Continuous Variables: A Primer of Theoretical Methods”. CRC press. (2017).
- [33] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. “Gaussian quantum information”. *Rev. Mod. Phys.* **84**, 621–669 (2012).
- [34] Alexander S Holevo. “One-mode quantum Gaussian channels: Structure and quantum capacity”. *Probl. Inf. Transm.* **43**, 1–11 (2007).
- [35] Gerardo Adesso. “Entanglement of Gaussian states” (2007). [arXiv:quant-ph/0702069](https://arxiv.org/abs/quant-ph/0702069).
- [36] Alessio Serafini, Gerardo Adesso, and Fabrizio Illuminati. “Unitarily localizable entanglement of Gaussian states”. *Phys. Rev. A* **71**, 032349 (2005).
- [37] Jim Harrington and John Preskill. “Achievable rates for the Gaussian quantum channel”. *Phys. Rev. A* **64**, 062301 (2001).
- [38] Lisa Häggli, Margret Heinze, and Robert König. “Enhanced noise resilience of the surface-Gottesman-Kitaev-Preskill code via designed bias”. *Phys. Rev. A* **102**, 052408 (2020).
- [39] Blayne W. Walshe, Ben Q. Baragiola, Rafael N. Alexander, and Nicolas C. Menicucci. “Continuous-variable gate teleportation and bosonic-code error correction”. *Phys. Rev. A* **102**, 062411 (2020).
- [40] Frank Schmidt and Peter van Loock. “Quantum error correction with higher Gottesman-Kitaev-Preskill codes: Minimal measurements and linear optics”. *Phys. Rev. A* **105**, 042427 (2022).
- [41] Benjamin Schumacher and M. A. Nielsen. “Quantum data processing and error correction”. *Phys. Rev. A* **54**, 2629–2635 (1996).
- [42] Seth Lloyd. “Capacity of the noisy quantum channel”. *Phys. Rev. A* **55**, 1613–1622 (1997).
- [43] Igor Devetak. “The private classical capacity and quantum capacity of a quantum channel”. *IEEE Transactions on Information Theory* **51**, 44–55 (2005).
- [44] Michael M. Wolf, Geza Giedke, and J. Ignacio Cirac. “Extremality of Gaussian Quantum States”. *Phys. Rev. Lett.* **96**, 080502 (2006).
- [45] A. S. Holevo and R. F. Werner. “Evaluating capacities of bosonic Gaussian channels”. *Phys. Rev. A* **63**, 032312 (2001).

Appendix

A Notation

Consider an N -mode bosonic Hilbert space $\mathcal{H}^{\otimes N}$, where $\mathcal{H}$ is the Hilbert space of a single bosonic mode. Define the vector of canonical operators for the N modes as,

$$\hat{\mathbf{r}}^\top \equiv (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_N, \hat{p}_N), \quad (19)$$

such that,

$$[\hat{\mathbf{r}}_k, \hat{\mathbf{r}}_j] = i\mathbf{\Omega}_{kj}, \quad (20)$$

where $\mathbf{\Omega}$ is the N -mode symplectic form,

$$\mathbf{\Omega} = \bigoplus_{i=1}^N \mathbf{\Omega}_1 \quad \text{with} \quad \mathbf{\Omega}_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (21)$$

The canonical operators $\hat{q}$ and $\hat{p}$ of a single mode are the real and imaginary parts, respectively, of the annihilation operator, $\hat{a}$. Throughout the appendix, we write ‘hat’ on operators only in places where operators can potentially be confused with numbers. For example, in quadrature operators; While we omit ‘hat’ for unitary operators and others that are easily recognized as operators.

In several figures, we draw quantum circuits with the circuit elements representing symplectic transformations $\mathbf{S}$ corresponding to a unitary $U_{\mathbf{S}}$. A dictionary of commonly used circuit elements is shown in Fig. 9.

B Gaussian Evolution

B.1 Gaussian unitaries

Given an N -mode symplectic transformation $\mathbf{S} \in \text{Sp}(2N, \mathbb{R})$, where $\text{Sp}(2N, \mathbb{R})$ is the set of $2N \times 2N$ real symplectic matrices (of dimension $|\text{Sp}(2N, \mathbb{R})| = 2N^2 + N$), such that $\mathbf{S}\mathbf{\Omega}\mathbf{S}^\top = \mathbf{\Omega}$ and $\det \mathbf{S} = 1$,

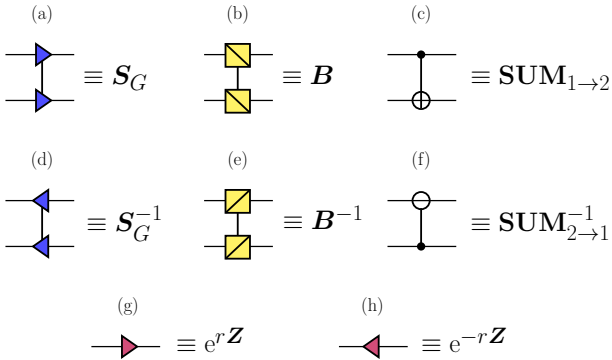


Figure 9: Dictionary of circuit elements for (a and d) two-mode squeezers, (b and e) beamsplitters, (c and f) SUM gates, and (g and h) single-mode squeezers.

one can find a unitary representation $U_{\mathbf{S}}$, which encodes the symplectic transformation $\mathbf{S}$ and acts on the canonical operators as,

$$U_{\mathbf{S}}^\dagger \hat{\mathbf{r}} U_{\mathbf{S}} = \mathbf{S} \hat{\mathbf{r}}. \quad (22)$$

A useful fact that holds for an arbitrary symplectic matrix is the so-called *Bloch-Messiah (or Euler) decomposition* [30]. The statement is that a symplectic matrix $\mathbf{S} \in \text{Sp}(2N, \mathbb{R})$ decomposes to

$$\mathbf{S} = \mathbf{B}' \cdot \left(\bigoplus_{i=1}^N e^{r_i \mathbf{Z}} \right) \cdot \mathbf{B}, \quad (23)$$

where $\mathbf{Z}$ is the Pauli-Z matrix and $e^{r_i \mathbf{Z}}$ is a single mode squeezing transformation on the i th mode with squeezing strength r_i . Here $\mathbf{B}', \mathbf{B} \in \text{Sp}(2N, \mathbb{R}) \cap \text{SO}(2N) \simeq U(N)$, where $U(N)$ is the unitary group of dimension $|U(N)| = N^2$. In other words, $\mathbf{B}'$ and $\mathbf{B}$ are passive (e.g., linear optical) transformations, which admit an even further decomposition in terms of two-mode beamsplitters and phase-shifts [31]. Since the squeezing transformation is a diagonal matrix with N free parameters, it follows that $|\text{Sp}(2N, \mathbb{R})| = 2|U(N)| + N = 2N^2 + N$, as mentioned previously.

Consider vectors $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathbb{R}^{2N}$ and define the Weyl (displacement) operator,

$$D_{\boldsymbol{\mu}} \equiv \exp(i\boldsymbol{\mu}^\top \mathbf{\Omega} \hat{\mathbf{r}}), \quad (24)$$

which form an operator basis for the space of bounded operators $\mathcal{B}(\mathcal{H}^{\otimes N})$ via $\text{Tr}(D_{\boldsymbol{\mu}} D_{-\boldsymbol{\nu}}) = (2\pi)^N \delta^{2N}(\boldsymbol{\mu} - \boldsymbol{\nu})$ [32]. Displacement operators satisfy a composition rule,

$$D_{\boldsymbol{\mu}} D_{\boldsymbol{\nu}} = e^{-i\omega(\boldsymbol{\mu}, \boldsymbol{\nu})} D_{\boldsymbol{\nu}} D_{\boldsymbol{\mu}}, \quad (25)$$

where $\omega(\boldsymbol{\mu}, \boldsymbol{\nu}) \equiv \boldsymbol{\mu}^\top \mathbf{\Omega} \boldsymbol{\nu}$. The anti-symmetric bilinear form $\omega : \mathbb{R}^{2N} \times \mathbb{R}^{2N} \rightarrow \mathbb{R}$ takes as input the real vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ and computes $\omega(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathbb{R}$, which is called the symplectic inner product between $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$. The symplectic inner product obeys $\omega(\boldsymbol{\nu}, \boldsymbol{\mu}) = -\omega(\boldsymbol{\mu}, \boldsymbol{\nu})$ and is invariant under symplectic transformations, i.e. $\omega(\mathbf{S}\boldsymbol{\mu}, \mathbf{S}\boldsymbol{\nu}) = \omega(\boldsymbol{\mu}, \boldsymbol{\nu})$. Finally, from Eq. (22) and the general conjugation formula $U^\dagger f(\hat{g})U = f(U^\dagger \hat{g} U)$, it can be shown that the displacement operators transform under symplectic transformations via

$$U_{\mathbf{S}} D_{\boldsymbol{\mu}} U_{\mathbf{S}}^\dagger = D_{\mathbf{S}\boldsymbol{\mu}}, \quad (26)$$

where $\mathbf{S}\boldsymbol{\mu} \in \mathbb{R}^{2N}$.

We now provide explicit symplectic matrices for some commonly used Gaussian unitaries in the single- and two-mode cases. For a single-mode, we only have two possible transformations, which are a single-mode squeezer and phase rotation. For two-mode operations, some common elements used in this paper and in the literature are a beamsplitter, TMS, and a SUM gate.

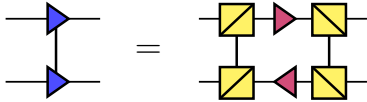


Figure 10: Bloch-Messiah decomposition for a TMS operation in terms of 50:50 beamsplitters and single mode squeezing.

A single-mode squeezer with squeezing strength r has a symplectic matrix representation,

$$\mathbf{S}_q(e^r) \equiv e^{r\mathbf{Z}}, \quad (27)$$

and a phase rotation at angle ϕ has a representation,

$$\mathbf{R}(\phi) \equiv \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}. \quad (28)$$

The two-mode beamsplitter has a symplectic representation

$$\mathbf{B} = \begin{pmatrix} \cos \theta \mathbf{I} & \sin \theta \mathbf{I} \\ -\sin \theta \mathbf{I} & \cos \theta \mathbf{I} \end{pmatrix}, \quad (29)$$

where, e.g., $\cos^2 \theta$ is the transmittance of the beamsplitter. For a 50:50 beamsplitter ($\theta = \pi/4$), we express the symplectic matrix as $\mathbf{B}_{1/2}$. [For a beamsplitter with transmittance T , we may write $\mathbf{B}_T$.]

The SUM-gate $\text{SUM}_\delta \equiv e^{-i\delta \hat{Q}_1 \otimes \hat{P}_2}$ operates on the canonical operators as,

$$\begin{aligned} \hat{Q}_1 &\rightarrow \hat{Q}_1, & \hat{P}_1 &\rightarrow \hat{P}_1 - \delta \hat{P}_2, \\ \hat{Q}_2 &\rightarrow \hat{Q}_2 + \delta \hat{Q}_1, & \hat{P}_2 &\rightarrow \hat{P}_2. \end{aligned} \quad (30)$$

where $\delta \in \mathbb{R}$ and has symplectic representation,

$$\text{SUM}_\delta = \begin{pmatrix} \mathbf{I}_2 & -\delta \mathbf{\Pi}_P \\ \delta \mathbf{\Pi}_Q & \mathbf{I}_2 \end{pmatrix}, \quad (31)$$

$$\text{SUM}_\delta^{-1} = \begin{pmatrix} \mathbf{I}_2 & \delta \mathbf{\Pi}_P \\ -\delta \mathbf{\Pi}_Q & \mathbf{I}_2 \end{pmatrix}, \quad (32)$$

where $\mathbf{\Pi}_Q = \text{diag}(1, 0)$ and $\mathbf{\Pi}_P = \text{diag}(0, 1)$ represent projections along the Q quadrature and P quadrature of the respective modes. For $\delta = 1$, we use the notation $\text{SUM} \equiv \text{SUM}_{\delta=1}$, which is the conventional definition of the SUM-gate in the literature. Observe that $\text{SUM}_{\delta_1} \cdot \text{SUM}_{\delta_2} = \text{SUM}_{\delta_1 + \delta_2}$. The SUM-gate is the CV analog to the CNOT gate for qubit-into-oscillator codes [3] and thus useful for ancilla-assisted stabilizer measurements (see Section D).

A TMS transformation with gain G has symplectic representation,

$$\mathbf{S}_G = \begin{pmatrix} \sqrt{G} \mathbf{I} & \sqrt{G-1} \mathbf{Z} \\ \sqrt{G-1} \mathbf{Z} & \sqrt{G} \mathbf{I} \end{pmatrix}. \quad (33)$$

Parameterizing the gain G in terms of the squeezing strength r via $G = \cosh^2 r$, one can provide a Bloch-Messiah decomposition for TMS in terms of two-mode beamsplitters and single-mode squeezers,

$$\mathbf{S}_G = \mathbf{B}_{1/2} \cdot \left(\mathbf{S}_q(e^r) \oplus \mathbf{S}_q(e^{-r}) \right) \cdot \mathbf{B}_{1/2}^\top. \quad (34)$$

See Fig. 10 for an illustration.

B.2 Gaussian channels

Consider an initial separable quantum state $\Psi \otimes \rho_E$, where $\Psi \in \mathcal{H}^{\otimes N}$ is a N mode quantum state of the system (not necessarily Gaussian) and $\rho_E \in \mathcal{H}^{\otimes M}$ is a M mode Gaussian quantum state of the environment. Given a symplectic transformation $\mathbf{S}$ with unitary representation $U_{\mathbf{S}}$, we define a Gaussian quantum channel $\mathcal{G} : \mathcal{H}^{\otimes N} \rightarrow \mathcal{H}^{\otimes N}$ which acts on the system modes via

$$\mathcal{G}(\Psi) \equiv \text{Tr}_E \left[U_{\mathbf{S}} (\Psi \otimes \rho_E) U_{\mathbf{S}}^\dagger \right]. \quad (35)$$

Let the symplectic matrix $\mathbf{S}$ be written in the following form,

$$\mathbf{S} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \quad (36)$$

where $\mathbf{A}$ is a $2N \times 2N$ matrix that dictates internal evolution of the system and $\mathbf{B}$ is a $2N \times 2M$ rectangular matrix which encodes the interaction between environment and system. [For a system of N modes evolving under the Gaussian channel $\mathcal{G}$, it is sufficient to choose $M \leq 2N$ to fully characterize the channel [33].] From thus, one can prove the following theorem,

Theorem 6 (Gaussian channel characterization). *Given a Gaussian environment state ρ_E —characterized by the first and second moments $\boldsymbol{\mu}_E$ and $\boldsymbol{\sigma}_E$ —a Gaussian quantum channel $\mathcal{G} : \mathcal{H}^{\otimes N} \rightarrow \mathcal{H}^{\otimes N}$ is completely determined by a displacement-noise vector $\mathbf{d}$, a scaling matrix $\mathbf{X}$, and a noise matrix $\mathbf{Y}$ such that*

$$\mathbf{d} = \mathbf{B} \boldsymbol{\mu}_E, \quad (37)$$

$$\mathbf{X} = \mathbf{A}, \quad (38)$$

$$\mathbf{Y} = \mathbf{B} \boldsymbol{\sigma}_E \mathbf{B}^\top, \quad (39)$$

where $\mathbf{A}$ and $\mathbf{B}$ are sub-matrices of the symplectic matrix $\mathbf{S}$ [Eq. (36)] that couples the system to the environment.

The displacement-noise vector $\mathbf{d}$ can always be subsumed into a unitary displacement occurring on the system state; thus of primary interest to the dynamics are the scaling and noise matrices, $\mathbf{X}$ and $\mathbf{Y}$, respectively.

Another useful fact to know about Gaussian channels is how their noise vectors, scaling matrices, and noise matrices combine when we concatenate Gaussian channels.

Theorem 7 (Gaussian channel synthesis). *Consider two Gaussian channels $\mathcal{G}_1$ and $\mathcal{G}_2$ with characterizations $(\mathbf{d}_1, \mathbf{X}_1, \mathbf{Y}_1)$ and $(\mathbf{d}_2, \mathbf{X}_2, \mathbf{Y}_2)$, respectively. Then the composite channel $\mathcal{G}_{12} = \mathcal{G}_2 \circ \mathcal{G}_1$, which is also a Gaussian channel, has a characteri-*

ization $(\mathbf{d}_{12}, \mathbf{X}_{12}, \mathbf{Y}_{12})$ that take the form [34, 25]

$$\mathbf{d}_{12} = \mathbf{d}_2 + \mathbf{X}_2 \mathbf{d}_1, \quad (40)$$

$$\mathbf{X}_{12} = \mathbf{X}_2 \mathbf{X}_1, \quad (41)$$

$$\mathbf{Y}_{12} = \mathbf{X}_2 \mathbf{Y}_1 \mathbf{X}_2^\top + \mathbf{Y}_2. \quad (42)$$

B.3 Condensing correlations

We discuss the condensation of correlations—from multimode correlations to elementary pair-like correlations—for general Gaussian transformations.

B.3.1 Phase-space Schmidt decomposition

As highlighted in the main text (Theorem 2), for the special case of iid AGN, we can significantly reduce the freedom in a generic GKP-stabilizer code by effectively reducing the code to a direct product of TMS codes and a general ancillary GKP lattice $\mathcal{L}$. The workhorse in proving Theorem 2 is the *modewise entanglement theorem* just below.

Theorem 8 (Modewise entanglement [21]). *Consider a subsystem A of N modes and a subsystem B of M modes, such that the joint system AB consists of $K = M + N$ modes, and define the positive definite matrix $\varrho_{AB} = \varrho \mathbf{S} \mathbf{S}^\top$, where $\varrho \in (0, \infty)$ and $\mathbf{S} \in \text{Sp}(2K, \mathbb{R})$. Then, there exists local symplectic matrices $\mathbf{\Lambda}_A \in \text{Sp}(2N, \mathbb{R})$ and $\mathbf{\Lambda}_B \in \text{Sp}(2M, \mathbb{R})$ such that,*

$$(\mathbf{\Lambda}_A \oplus \mathbf{\Lambda}_B) \varrho_{AB} (\mathbf{\Lambda}_A^\top \oplus \mathbf{\Lambda}_B^\top) = \varrho \left(\bigoplus_{i=1}^N \mathbf{S}_{G_i} \mathbf{S}_{G_i}^\top \right) \oplus \mathbf{I}_{K-2N}, \quad (43)$$

where $\mathbf{S}_{G_i}$ is a TMS squeezing operation of gain G_i [see, e.g., Eq. (33)] between the i th mode in A and the $(N+i)$ th mode in B .

Theorem 8 is sometimes referred to as the *phase-space Schmidt decomposition* for Gaussian states [35]; see Refs. [21] and [32] for proofs and Ref. [36, 35] for more discussion and a generalization to ‘bisymmetric’ noise.

As an illustrative example, let ϱ_{AB} be the covariance matrix for an isotropically mixed K mode Gaussian state Ψ_{AB} ($\varrho \geq 1$). Then, by Lemma 8, the state Ψ_{AB} is locally equivalent (up to a Gaussian unitary $U_{\mathbf{\Lambda}_A}^\dagger \otimes U_{\mathbf{\Lambda}_B}^\dagger$) to a direct product of N (noisy) TMS vacuum states and $K-2N$ uncorrelated thermal states. In other words, Ψ_{AB} is modewise entangled—i.e., each mode in A is entangled with only one corresponding mode in B .

B.3.2 Not-so-normal mode decomposition

We briefly discuss a generalization of the normal mode decomposition (see, e.g., Refs. [33, 32]) to a

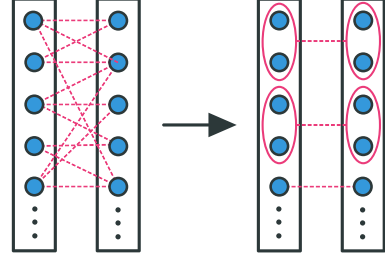


Figure 11: Multimode correlations condense to elementary two- and four-mode units (see Theorem 9 below and cf. Ref. [26]).

“not-so-normal” mode decomposition [26, 27]. We utilize this decomposition in our argument for the reduction of GKP codes against heterogeneous AGN. It is easier to state the results of Ref. [26] by using the following order of canonical operators $\hat{\mathbf{r}} = (\hat{q}_1, \hat{q}_2, \dots, \hat{p}_1, \hat{p}_2, \dots)^\top$.

Theorem 9 (Not-so-normal mode decomposition [26]). *Consider a $2N \times 2N$ invertible, non-defective matrix $\mathbf{X}$. Then, there exists symplectic transformations $\mathbf{S}_A, \mathbf{S}_B \in \text{Sp}(2N, \mathbb{R})$ such that*

$$\mathbf{S}_A \mathbf{X} \mathbf{S}_B = \begin{pmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathbb{R}} \end{pmatrix}, \quad (44)$$

where $\mathbf{J}_{\mathbb{R}} = \bigoplus_{k=1}^N \mathbf{J}(\lambda_k)$ is a $N \times N$ block-diagonal matrix consisting of the (two-fold degenerate) eigenvalues $\{\lambda_k\}_{k=1}^N$ of the matrix $\mathbf{X} \mathbf{\Omega}^\top \mathbf{X}^\top \mathbf{\Omega}$. For complex $\lambda_k = a_k + ib_k$, the diagonal blocks are 2×2 matrices $\mathbf{J}_k(\lambda_k) = \begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}$. For real λ_k , $\mathbf{J}_k(\lambda_k) = \lambda_k$.

As a technical aside, for defective $\mathbf{X}$, the above theorem still holds, however in that case, $\mathbf{J}_{\mathbb{R}}$ is an $N \times N$ matrix in real Jordan form (hence the subscript), with Jordan blocks containing the degenerate eigenvalues of $\mathbf{X}$.

As an example, if $\mathbf{X}$ is a positive definite real matrix, then the above is effectively equal to the normal mode decomposition. Recall from the normal mode decomposition that there exists $\mathbf{S}'_A, \mathbf{S}'_B \in \text{Sp}(2N, \mathbb{R})$ such that $\mathbf{S}'_A \mathbf{X} \mathbf{S}'_B = \nu \oplus \nu$ where $\nu = \text{diag}(\nu_1, \nu_2, \dots, \nu_N)$ and ν_k are the symplectic eigenvalues of $\mathbf{X}$ [33, 32]. The eigenvalues λ_k of $\mathbf{X} \mathbf{\Omega}^\top \mathbf{X}^\top \mathbf{\Omega}$ are real in this case and are related to the symplectic eigenvalues ν_k of $\mathbf{X}$ via $\lambda_k = \nu_k^2$. Indeed, one can show that the decomposition in Eq. (44) and the normal mode decomposition just above are equivalent up to local squeezing transformations $\bigoplus_k \mathbf{S}_q(e^{r_k})$, with squeezing strengths $e^{r_k} = \nu_k$.

As another example, consider a $4N \times 4N$ covariance matrix σ for an $2N$ mode Gaussian quantum state. Let us go back to $(q_1, p_1, q_2, p_2, \dots)$ ordering; this will allow us to write the covariance matrix in block form as below. Then, consider a bipartite cut which partitions a subsystems A of N modes from a subsystem

B of N modes. We write the covariance matrix in blocks as

$$\sigma = \begin{pmatrix} \sigma_A & \sigma_{AB} \\ \sigma_{AB}^\top & \sigma_B \end{pmatrix}, \quad (45)$$

where σ_A, σ_B are local covariance matrices and σ_{AB} is the correlation matrix. Consider local symplectic transformations $\mathbf{S}_A, \mathbf{S}_B \in \text{Sp}(2N, \mathbb{R})$ such that $\sigma \rightarrow (\mathbf{S}_A \oplus \mathbf{S}_B) \sigma (\mathbf{S}_A^\top \oplus \mathbf{S}_B^\top)$. The correlation matrix transforms as $\sigma_{AB} \rightarrow \mathbf{S}_A \sigma_{AB} \mathbf{S}_B^\top$. By Theorem 9, we can choose $\mathbf{S}_A$ and $\mathbf{S}_B$ to reduce σ_{AB} into “not-so-normal” form, completely determined by Jordan blocks $\mathbf{J}(\lambda_k)$. This transformation condenses the correlations into two- and four-mode units. In particular, each $\lambda_k \in \mathbb{R}$ implies two-mode correlations, whereas $\lambda_k \in \mathbb{C}$ implies four-mode correlations; see Fig. 11 for an illustration.

C GKP lattice states

We present some details regarding lattice GKP states for N modes and, for concreteness, provide explicit examples for one and two modes; see also Refs. [17, 18].

C.1 Modular quadratures

Consider a GKP lattice state $|\mathcal{L}\rangle$ and the stabilizer group $\mathcal{S}_{\mathcal{L}} = \langle S_1^{\mathcal{L}}, \dots, S_{2N}^{\mathcal{L}} \rangle$, such that $\mathcal{S}_{\mathcal{L}} |\mathcal{L}\rangle = |\mathcal{L}\rangle$. Since $S_j^{\mathcal{L}} = \exp(i\hat{g}_j^{\mathcal{L}})$, where $\hat{g}_j^{\mathcal{L}} \equiv \lambda_j^{\mathcal{L}\top} \Omega \hat{\mathbf{r}}$, we can equivalently say that $|\mathcal{L}\rangle$ is an eigenstate of the modular quadratures $\{\hat{g}_j^{\mathcal{L}}\}$ (which are the generators of translations in the lattice) with eigenvalue 0 mod 2π . Given a generator matrix $\mathbf{M}$ of $\mathcal{L}$, we can package all the modular quadratures quite nicely into a vector operator,

$$\hat{\mathbf{g}}^{\mathcal{L}} = \mathbf{M}^\top \Omega \hat{\mathbf{r}} = \begin{pmatrix} \hat{g}_1^{\mathcal{L}} \\ \hat{g}_2^{\mathcal{L}} \\ \vdots \end{pmatrix}. \quad (46)$$

If the lattice is a symplectic self-dual lattice [37] (see below), such that $\mathbf{M}^\top \Omega \mathbf{M} = 2\pi \Omega$, then $\hat{\mathbf{g}}^{\mathcal{L}} = \Omega \mathbf{M}^{-1} \hat{\mathbf{r}}$ and $\hat{\mathbf{g}}^{\mathcal{L}}$ defines a new set of quadrature operators, up to a factor $\ell = \sqrt{2\pi}$.

C.2 Canonical GKP state

The prototypical example of a lattice state is the canonical (square) GKP state $|\square\rangle$, which one can define via the stabilizers,

$$S_1^\square \equiv \exp(i\lambda_1^{\square\top} \Omega_1 \hat{\mathbf{r}}) = e^{i\ell \hat{q}}, \quad (47)$$

$$S_2^\square \equiv \exp(i\lambda_2^{\square\top} \Omega_1 \hat{\mathbf{r}}) = e^{i\ell \hat{p}}, \quad (48)$$

where $\ell \equiv \sqrt{2\pi}$, such that $S_K^\square |\square\rangle = |\square\rangle \forall K \in \{1, 2\}$. The vectors $\lambda_K^\square$ are given explicitly as

$$\lambda_1^\square \equiv \ell(0, -1)^\top, \quad (49)$$

$$\lambda_2^\square \equiv \ell(1, 0)^\top. \quad (50)$$

Observe that $[S_1^\square, S_2^\square] = 0$ since $\omega(\lambda_1^\square, \lambda_2^\square) = \ell^2 = 2\pi$. For reference, a generator matrix built from these vectors is

$$\mathbf{M}^\square = \begin{pmatrix} \lambda_1^\square & \lambda_2^\square \end{pmatrix} = \ell \Omega_1, \quad (51)$$

which is just the symplectic form in $\mathbb{R}^2$. The modular quadratures $\hat{\mathbf{g}}^\square$ are thus equivalent (up to a factor ℓ) to the canonical operators $\hat{q}$ and $\hat{p}$, as explicitly seen above for the stabilizers (47)-(48). One may therefore refer to the quadrature basis for the square GKP state as the ‘canonical’ or ‘standard’ basis. In terms of the basis states for $\hat{q}$ and $\hat{p}$, the canonical GKP state can be written as

$$|\square\rangle \propto \sum_{n \in \mathbb{Z}} |q = n\ell\rangle = \sum_{n \in \mathbb{Z}} |p = n\ell\rangle, \quad (52)$$

which is a translation invariant square lattice in the single-mode phase space with period $\ell = \sqrt{2\pi}$.

We denote N canonical GKP states as $|\square^N\rangle \equiv |\square\rangle^{\otimes N}$, corresponding to a $2N$ dimensional hypercube. The multimode lattice can be described by the vectors $\lambda_J^\square$, where $J = \{1, \dots, 2N\}$, such that, e.g.,

$$\lambda_1^\square = (0, -\ell)^\top \oplus \mathbf{0}_{2(N-1)}^\top,$$

$$\lambda_2^\square = (\ell, 0)^\top \oplus \mathbf{0}_{2(N-1)}^\top,$$

$$\vdots$$

$$\lambda_{2N-1}^\square = \mathbf{0}_{2(N-1)}^\top \oplus (0, -\ell)^\top,$$

$$\lambda_{2N}^\square = \mathbf{0}_{2(N-1)}^\top \oplus (\ell, 0)^\top,$$

where $\mathbf{0}_{2(N-1)}$ is a $2(N-1)$ dimensional zero vector. The generator matrix is the symplectic form on N modes, which is simply a direct sum of the generator matrices for the individual modes,

$$\mathbf{M}^{\square^N} = (\lambda_1^\square \dots \lambda_{2N}^\square) = \ell \Omega, \quad (53)$$

such that $\mathcal{L}_{\square^N} \simeq \ell \mathbb{Z}^{2N}$. This represents a canonical hypercube in $2N$ dimensions, which is a (symplectic) self-dual lattice.

C.3 Self-dual GKP states

One can generate other all other (symplectic) self-dual lattices via symplectic transformations on the canonical hypercube (a consequence of Corollary 1 in Ref. [18]). The corresponding lattice states are ‘symplectically equivalent’ to N canonical GKP states but can nonetheless be useful resources for practical tasks.

Consider a symplectic transformation $\Lambda \in \text{Sp}(2N, \mathbb{R})$ with unitary representation U_Λ and define new lattice vectors $\lambda_J^\Lambda \equiv \Lambda \lambda_J^\square$. Since the symplectic inner product is invariant under symplectic transformations, it follows that $\omega(\lambda_J^\Lambda, \lambda_K^\Lambda) = \omega(\lambda_J^\square, \lambda_K^\square)$, and thus one can naturally define a symplectically integral lattice $\mathcal{L}_\Lambda$ via $\{\lambda_J^\Lambda\}$. Furthermore, we can

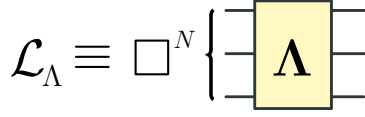


Figure 12: Self-dual GKP lattice state $\mathcal{L}_\Lambda$ generated from N canonical square GKP by a symplectic transformation Λ .

associate these vectors with the lattice stabilizers $S_J^\Lambda \equiv D_{\lambda_J^\Lambda}$, which generate the stabilizer group $\mathcal{S}_{\mathcal{L}_\Lambda} = \langle S_1^\Lambda, S_2^\Lambda, \dots, S_{2N}^\Lambda \rangle$. By relation (26), the stabilizers for $\mathcal{L}_\Lambda$ are related to canonical stabilizers via conjugation,

$$S_J^\Lambda = U_\Lambda S_J^\square U_\Lambda^\dagger. \quad (54)$$

It follows immediately that the state $|\mathcal{L}_\Lambda\rangle \equiv U_\Lambda |\square^N\rangle$ is a $+1$ eigenstate of any element in $\mathcal{S}_{\mathcal{L}_\Lambda}$, i.e. $\mathcal{S}_{\mathcal{L}_\Lambda} |\mathcal{L}_\Lambda\rangle = |\mathcal{L}_\Lambda\rangle$.

Consider a generator matrix, M^Λ , built from the lattice vectors,

$$M^\Lambda \equiv (\lambda_1^\Lambda \quad \lambda_2^\Lambda \quad \dots \quad \lambda_{2N}^\Lambda). \quad (55)$$

A generator matrix and the symplectic transformation Λ are related via

$$M^\Lambda = \ell \Lambda \Omega = \Lambda M^{\square^N}, \quad (56)$$

i.e., a generator matrix for the new lattice M^Λ can be found by transforming a generator matrix of the canonical lattice $M^{\square^N}$ by a symplectic transformation Λ . It immediately follows that $M^{\Lambda^\top} \Omega M^\Lambda = 2\pi \Omega$.

Using Eq. (46), one can show that the modular quadratures $\hat{g}^\Lambda$ which generate discrete translations on the lattice are explicitly given as

$$\hat{g}^\Lambda = \ell \Lambda^{-1} \hat{r}. \quad (57)$$

Since $\Lambda \in \text{Sp}(2N, \mathbb{R})$, $\hat{g}^\Lambda$ defines a good set of canonical operators which obey the canonical commutation relations up to a factor $\ell^2 = 2\pi$.

We provide some examples of self-dual lattices below.

Example: Hexagonal GKP. An example of a single-mode ($N = 1$) symplectic lattice built from the canonical lattice is the hexagonal lattice $\mathcal{L}_\square$, which admits the densest sphere-packing in $\mathbb{R}^2$. Consider the following symplectic transformation,

$$\Lambda_\square \equiv \frac{\ell_\square}{\ell} \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}, \quad (58)$$

where $\ell_\square/\ell = \sqrt{2}/3^{1/4} \approx 1.07$. The hexagonal lattice vectors are defined from the square lattice vectors via $\lambda_K^\square = \Lambda_\square \lambda_K^\square$ and written explicitly as

$$\lambda_1^\square \equiv \ell_\square \left(1/2, -\sqrt{3}/2 \right)^\top, \quad (59)$$

$$\lambda_2^\square \equiv \ell_\square (1, 0)^\top. \quad (60)$$

Note that $\|\lambda_K^\square\| = \|\lambda_2^\square - \lambda_1^\square\| = \ell_\square \approx 1.07$, where $\|\cdot\|$ is the Euclidean norm; this is also the minimal distance between lattice points in $\mathcal{L}_\square$. [As an aside, the Euclidean norm of a vector μ can be defined in symplectic geometry via the symplectic inner product $\omega(\cdot, \cdot)$ without needing to refer to Euclidean geometry. Indeed, consider the dual vector $\tilde{\mu} \equiv \Omega \mu$, then $\omega(\tilde{\mu}, \mu) = \|\mu\|$.]

Furthermore, the hexagonal stabilizer group $\mathcal{S}_\square$ can be generated from the following stabilizers,

$$S_1^\square \equiv \exp\left(i\lambda_1^{\square^\top} \Omega_1 \hat{r}\right) = \exp\left[i\ell_\square \left(\frac{\sqrt{3}}{2} \hat{q} + \frac{1}{2} \hat{p}\right)\right], \quad (61)$$

$$S_2^\square \equiv \exp\left(i\lambda_2^{\square^\top} \Omega_1 \hat{r}\right) = \exp[i\ell_\square \hat{p}], \quad (62)$$

which can also be found via conjugation of the square stabilizers via $S_K^\square = U_\square S_K^\square U_\square^\dagger$, where $U_\square$ is a unitary representation of $\Lambda_\square$. As in the general case, the hexagonal GKP state is formally given by $|\square\rangle = U_\square |\square\rangle$, which is necessarily a $+1$ eigenstate of the stabilizer group elements, i.e. $\mathcal{S}_\square |\square\rangle = |\square\rangle$.

Example: Rectangular GKP. Consider the symplectic squeezing matrix with squeezing strength $\eta \equiv e^r$,

$$\mathbf{Sq}(\eta) = \begin{pmatrix} \eta & 0 \\ 0 & 1/\eta \end{pmatrix}. \quad (63)$$

We define the new lattice vectors $\lambda_K^\eta \equiv \mathbf{Sq}(\eta) \lambda_K^\square$ from which we construct the stabilizers,

$$S_1^\eta \equiv \exp\left(i\lambda_1^{\eta^\top} \Omega_1 \hat{r}\right) = e^{i\hat{q}\ell/\eta}, \quad (64)$$

$$S_2^\eta \equiv \exp\left(i\lambda_2^{\eta^\top} \Omega_1 \hat{r}\right) = e^{i\hat{p}\ell\eta}. \quad (65)$$

The state $|\eta\rangle \equiv U_{\mathbf{Sq}(\eta)} |\square\rangle$ is a $+1$ eigenstate of these stabilizers, and defines a rectangular GKP state—i.e., squeezed along one quadrature and stretched along the other. This state can be written in the position and momentum bases as

$$|\eta\rangle \propto \sum_{k \in \mathbb{Z}} |q = k\ell\eta\rangle = \sum_{k \in \mathbb{Z}} |p = k\ell/\eta\rangle. \quad (66)$$

The rectangular GKP state has proven useful for bias-enhanced QEC with the GKP surface code [38].

Example: GKP Bell state. An interesting class of two-mode lattice states are entangled GKP states generated via two-mode interactions. A prominent example is a GKP Bell state Φ often used in DV quantum information processing with bosonic qubits [39, 40]. A GKP Bell state can be formed by interacting two canonical GKP states on a 50:50 beamsplitter [39], i.e.

$$|\Phi\rangle \equiv U_{B_{1/2}} |\square^2\rangle. \quad (67)$$

The lattice vectors $\{\lambda_j^\Phi\}$ are given by the column vectors of the generator matrix M^Φ , which can be directly found via Eq. (56); explicitly,

$$M^\Phi = \ell B_{1/2} \Omega = \frac{\ell}{\sqrt{2}} \begin{pmatrix} \Omega_1 & \Omega_1 \\ -\Omega_1 & \Omega_1 \end{pmatrix}, \quad (68)$$

where Ω_1 is the symplectic form on $\mathbb{R}^2$. The GKP Bell state has the same lattice spacing as the canonical GKP state $|\square^2\rangle$, since the former is simply a version of the latter in $\mathbb{R}^4$ and are thus not useful for QEC. On the other hand, GKP Bell states are useful in other quantum information processing tasks [40], such as quantum teleportation.

Example: D_4 GKP. The D_4 lattice admits the densest sphere packing in $\mathbb{R}^4$ [17]. One can generate a D_4 GKP state from a direct product of two canonical GKP states by the following symplectic transformation,

$$\Lambda_{D_4}^\top = \sqrt[4]{2} \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad (69)$$

such that $|D_4\rangle \equiv U_{\Lambda_{D_4}} |\square^2\rangle$. Since $\Lambda_{D_4} \Lambda_{D_4}^\top \neq I$, there is squeezing involved.

D Syndrome measurements

We discuss how to extract error syndromes from a general GKP lattice.

D.1 Error syndrome

Consider a lattice state $\mathcal{L}$ and a displacement error e such that

$$|\mathcal{L}; e\rangle \equiv D_e |\mathcal{L}\rangle. \quad (70)$$

Then observe the following set of equalities,

$$\begin{aligned} S_K^\mathcal{L} |\mathcal{L}; e\rangle &= S_K^\mathcal{L} D_e |\mathcal{L}\rangle \\ &= e^{-i\omega(\lambda_K^\mathcal{L}, e)} D_e S_K^\mathcal{L} |\mathcal{L}\rangle \\ &= e^{-i\omega(\lambda_K^\mathcal{L}, e)} D_e |\mathcal{L}\rangle \\ &= e^{-i\omega(\lambda_K^\mathcal{L}, e)} |\mathcal{L}; e\rangle, \end{aligned}$$

where $\lambda_K^\mathcal{L}$ are the lattice vectors that satisfy $\omega(\lambda_K^\mathcal{L}, \lambda_J^\mathcal{L}) = 2\pi n_{KJ}$ with $n_{KJ} \in \mathbb{Z}$, and we have used the composition rule (25) to go from the first equality to the second equality. In other words, the error state $|\mathcal{L}; e\rangle$ is an eigenstate of the stabilizer $S_K^\mathcal{L}$ with eigenvalue $\exp[-i\omega(\lambda_K^\mathcal{L}, e)]$.

The set of quantities $\{\omega(\lambda_K^\mathcal{L}, e) \bmod 2\pi\}_{K=1}^{2N}$ are the error syndromes extracted from stabilizer measurements on the GKP lattice $\mathcal{L}$. We can package them nicely into an *error syndrome* vector s with the help of a generator matrix M ,

$$s \equiv s(e) = \ell^{-1} M^\top \Omega e \bmod \sqrt{2\pi}, \quad (71)$$

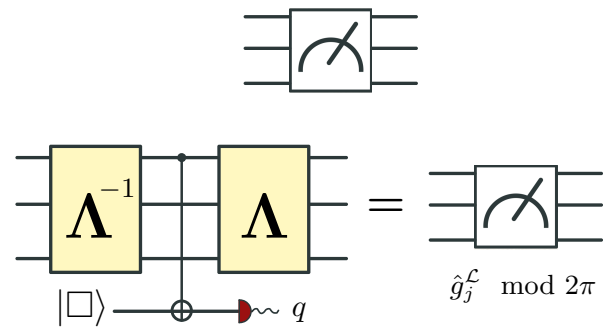


Figure 13: GKP-assisted stabilizer measurement circuit for $S_K^\mathcal{L}$. Equivalent to measuring the modular quadrature $\hat{g}_j^\mathcal{L} \bmod 2\pi$.

where the modulo operation acts elementwise and we have factored out the canonical spacing ℓ . If $\mathcal{L}$ is a symplectic self-dual lattice such that $M = \Lambda \Omega$ where $\Lambda \in \text{Sp}(2N, \mathbb{R})$ [see Eq. (56)], then we may write the syndrome vector as

$$s = \Lambda^{-1} e \bmod \sqrt{2\pi}. \quad (72)$$

D.2 Homodyne and ancilla-assisted measurements

One way to measure displacements on a lattice state is to perform homodyne measurements along the modular quadratures ($\hat{g} = M^\top \Omega \hat{r} \bmod 2\pi$). However, homodyne measurements are destructive, hence an ancilla-assisted measurement scheme is warranted.

For lattices that can be generated from the canonical lattice via symplectic transformations (i.e., self-dual lattices), we introduce a measurement circuit that is related to the canonical measurement circuit (consisting of standard SUM gates and square GKP [3]) via unitary conjugation; see Fig. 13 for an illustration of the stabilizer measure circuit. Our measurement scheme relies on the SUM gate, however we remark that SUM-gate measurement strategies necessarily require inline squeezing, as the SUM-gate is not an orthogonal transformation. An ancilla-assisted measurement strategy which moves squeezing offline has been proposed in Ref. [40].

There are $2N$ stabilizers that need to be measured, two for each mode of the GKP lattice $\mathcal{L}$. However, if the GKP lattice $\mathcal{L}$ is an ancillary state in a QEC circuit and can thus be discarded after use, then only N additional GKP ancillae (on top of the N mode GKP lattice $\mathcal{L}$) are required for measurements. This reduces the GKP resources by half. In more detail, we can first perform N nondestructive stabilizer (e.g., q quadrature) measurements via ancilla-assisted measurements, which consumes N GKP measurement ancillae. Following these nondestructive measurements, we subsequently perform N *destructive* homodyne (e.g., p quadrature) measurements on the GKP lattice $\mathcal{L}$. Moreover, this measurement strategy can be performed in parallel on the individual modes.

E Symmetry of single-mode lattices

We can have the same lattice $\mathcal{L}$ with different choices of bases. In particular, the generator matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ generate the same lattice if there exist a unimodular matrix $\mathbf{N}$ (i.e. a matrix with integer entries and $\det \mathbf{N} = 1$) such that

$$\mathbf{M}_1^\top = \mathbf{N} \mathbf{M}_2^\top. \quad (73)$$

Suppose that a single-mode symplectic self-dual lattice is given by (we drop the $\ell = \sqrt{2\pi}$ factor in front of $\mathbf{M}$ here for brevity)

$$\mathbf{M} = \mathbf{R}(\phi) \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta), \quad (74)$$

where $\mathbf{R}(\theta)$ and $\mathbf{R}(\phi)$ are 2×2 rotation matrices and $\mathbf{S} \mathbf{q}(r)$ is the single-mode squeezing. Fixing r , it is easy to see that two bases

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta), \\ \mathbf{M}_2 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta + \pi/2), \end{aligned}$$

are the same lattice since $\mathbf{M}_2^\top = \mathbf{R}(\frac{\pi}{2}) \mathbf{M}_1^\top$ and $\mathbf{R}(\frac{\pi}{2})$ is a unimodular matrix. Furthermore, it is easy to show that

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta), \\ \mathbf{M}_2 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(-\theta), \end{aligned}$$

are the same lattice under a reflection about the x -axis. Combining the above symmetries, we conclude that two bases

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta), \\ \mathbf{M}_2 &= \mathbf{S} \mathbf{q}(r) \mathbf{R}(\pi/2 - \theta), \end{aligned}$$

correspond to the same lattice.

E.1 Hexagonal lattice

A generator matrix for the hexagonal lattice can be written as

$$\mathbf{M}_\square = \frac{\ell_\square}{\ell} \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}, \quad (75)$$

where $\ell_\square/\ell = \sqrt{2}/3^{1/4} \approx 1.07$; see Eq. (58). By the Bloch-Messiah decomposition,

$$\mathbf{M}_\square = \mathbf{R}\left(-\frac{\pi}{6}\right) \mathbf{S} \mathbf{q}(3^{1/4}) \mathbf{R}\left(\frac{\pi}{4}\right), \quad (76)$$

where $\mathbf{S} \mathbf{q}(3^{1/4}) = \text{diag}(3^{1/4}, 3^{-1/4})$. Due to rotation symmetry, we have the same lattice if

$$\begin{aligned} \mathbf{M}'_\square &\equiv \mathbf{S} \mathbf{q}(r) \mathbf{R}(\theta) \\ &= \mathbf{R}(\phi) \mathbf{M}_\square \mathbf{N}^\top, \end{aligned} \quad (77)$$

where $\mathbf{N}$ is a unimodular matrix. For $\mathbf{N} \neq \mathbf{I}$, the squeezing value $r \geq 3^{1/4}$; see Table 1 of the main text.

F Proof of Lemma 1

We prove Lemma 1 of the main text. The primary techniques that we use are Gaussian channel synthesis and the modewise entanglement theorem [21] (Theorem 7 and Theorem 8 of Appendix B, respectively). Below, we sometimes write $\mathcal{N}_\sigma$ for the K mode iid AGN channel, as opposed to $\bigotimes_{i=1}^K \mathcal{N}_\sigma$, for the sake of brevity.

Denoting $\mathbf{S} = \mathbf{S}_{\text{enc}}^{-1}$ for simplicity, it is sufficient to show that,

$$\begin{aligned} \mathcal{U}_\mathbf{S} \circ \mathcal{N}_\sigma \circ \mathcal{U}_\mathbf{S}^{-1} &\simeq \\ \left(\mathcal{U}_{\bigoplus \mathbf{S}_{G_i}} \otimes \text{id}_{M-N} \right) \circ \mathcal{N}_\sigma \circ &\left(\mathcal{U}_{\bigoplus \mathbf{S}_{G_i}}^{-1} \otimes \text{id}_{M-N} \right), \end{aligned} \quad (78)$$

where $\mathbf{S} \in \text{Sp}(2K, \mathbb{R})$ is an arbitrary symplectic matrix for $K = M + N$ modes, ‘ $\simeq$ ’ means ‘equivalent up to local unitaries’, id_{M-N} is an identity superoperator on $M - N$ modes, and $\bigoplus \mathbf{S}_{G_i}$ is shorthand for $\bigoplus_{i=1}^N \mathbf{S}_{G_i}$. The channel $\mathcal{G}_1 \equiv \mathcal{U}_\mathbf{S} \circ \mathcal{N}_\sigma \circ \mathcal{U}_\mathbf{S}^{-1}$ is a Gaussian channel characterized by a displacement-noise vector $\mathbf{d}_1 = \mathbf{0}$, a scaling matrix $\mathbf{X}_1 = \mathbf{I}$, and a noise matrix $\mathbf{Y}_1 = \sigma^2 \mathbf{S} \mathbf{S}^\top$ (see Appendix B for details about characterizing general Gaussian channels). By the modewise entanglement theorem [21], there exists local symplectic matrices $\mathbf{\Lambda}_d \in \text{Sp}(2N, \mathbb{R})$ and $\mathbf{\Lambda}_a \in \text{Sp}(2M, \mathbb{R})$ such that,

$$\begin{aligned} (\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a) \mathbf{Y}_1 (\mathbf{\Lambda}_d^\top \oplus \mathbf{\Lambda}_a^\top) &= \\ \sigma^2 \left(\bigoplus_{i=1}^N \mathbf{S}_{G_i} \mathbf{S}_{G_i}^\top \right) \oplus \mathbf{I}_{2(M-N)}, \end{aligned} \quad (79)$$

where $\mathbf{S}_{G_i}$ is a TMS operation between the i th data mode and the $(N+i)$ th ancilla mode. The transformation $\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a$ corresponds to a direct product of local, unitary Gaussian channels $\mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a} = \mathcal{U}_{\mathbf{\Lambda}_d} \otimes \mathcal{U}_{\mathbf{\Lambda}_a}$. We can pre- and post-process with $\mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a}$ and $\mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a}^{-1}$ to generate a new Gaussian channel,

$$\begin{aligned} \mathcal{G}_2 &= \mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a} \circ \mathcal{G}_1 \circ \mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a}^{-1} \\ &= \mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a} \circ (\mathcal{U}_\mathbf{S} \circ \mathcal{N}_\sigma \circ \mathcal{U}_\mathbf{S}^{-1}) \circ \mathcal{U}_{\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a}^{-1}. \end{aligned} \quad (80)$$

Then, by Gaussian channel synthesis, starting from $\mathbf{d}_1$, $\mathbf{X}_1$ and $\mathbf{Y}_1$, $\mathcal{G}_2$ has a characterization $\mathbf{d}_2 = \mathbf{0}$, $\mathbf{X}_2 = \mathbf{I}$, and

$$\begin{aligned} \mathbf{Y}_2 &= (\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a) \mathbf{Y}_1 (\mathbf{\Lambda}_d^\top \oplus \mathbf{\Lambda}_a^\top) \\ &= \sigma^2 \left(\bigoplus_{i=1}^N \mathbf{S}_{G_i} \mathbf{S}_{G_i}^\top \right) \oplus \mathbf{I}_{2(M-N)}, \end{aligned} \quad (81)$$

where Eq. (79) was used for the second equality. Defining the unitary Gaussian channel for the (modewise) TMS operations $\mathcal{U}_{\bigoplus \mathbf{S}_{G_i}} \equiv \bigotimes_{i=1}^N \mathcal{U}_{\mathbf{S}_{G_i}}$, it follows that,

$$\mathcal{G}_2 = \left(\mathcal{U}_{\bigoplus \mathbf{S}_{G_i}} \otimes \text{id}_{M-N} \right) \circ \mathcal{N}_\sigma \circ \left(\mathcal{U}_{\bigoplus \mathbf{S}_{G_i}}^{-1} \otimes \text{id}_{M-N} \right). \quad (82)$$

Therefore, up to pre- and post-processing with local

Gaussian unitaries, we have that $\mathcal{G}_2 \simeq \mathcal{G}_1$, which was to be proved.

G Error estimation for general GKP lattices

Here we derive properties of error estimation—such as the joint PDF $P(\mathbf{x}_d, \mathbf{s})$ after a QEC protocol and the minimum mean square estimation (MMSE) $\hat{\mathbf{x}}_d = \mathbf{f}(\mathbf{s})$ for error syndrome $\mathbf{s}$ —given an encoding $\mathbf{S}_{\text{enc}}$ and ancilla GKP state with generator matrix $\mathbf{M}$. To simplify our derivations, we denote the probability density function (PDF) of the $2n$ -dimensional multivariate Gaussian distribution as

$$g(\mathbf{\Sigma}, \mathbf{x} - \boldsymbol{\mu}) = \frac{1}{(2\pi)^n \sqrt{\det \mathbf{\Sigma}}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (83)$$

where $\mathbf{x} = (x_1, \dots, x_{2n})^\top$, $\mathbf{\Sigma}$ is the covariance matrix and $\boldsymbol{\mu}$ is the mean. The dimension of the distribution is implicitly given by the dimensions of $\mathbf{\Sigma}$ and $\boldsymbol{\mu}$.

G.1 Proof of Theorem 4

In the QEC protocol, the original noise covariance matrix $\mathbf{V}_\xi = \bigoplus_{i=1}^{N+M} \sigma_i^2 \mathbf{I}_2$ is transformed into

$$\mathbf{V}_x^{-1} = \mathbf{S}_{\text{enc}}^\top \mathbf{V}_\xi^{-1} \mathbf{S}_{\text{enc}}, \quad (84)$$

via the encoding symplectic transform $\mathbf{S}_{\text{enc}}$ and decoding symplectic transform $\mathbf{S}_{\text{enc}}^{-1}$. Let the additive noise on the data and ancilla be

$$\mathbf{x}_d = (x_1^{(q)}, x_1^{(p)}, \dots, x_N^{(q)}, x_N^{(p)})^\top, \quad (85)$$

$$\mathbf{x}_a = (x_{N+1}^{(q)}, x_{N+1}^{(p)}, \dots, x_{N+M}^{(q)}, x_{N+M}^{(p)})^\top, \quad (86)$$

which are random variables following the joint distribution

$$P(\mathbf{x}_d, \mathbf{x}_a) = g[\mathbf{V}_x, (\mathbf{x}_d, \mathbf{x}_a)]. \quad (87)$$

We define the interval $\mathcal{J} \equiv [-\sqrt{\pi/2}, \sqrt{\pi/2}]$, and the error syndrome

$$\mathbf{s} = R_{\sqrt{2\pi}}(\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) \in \mathcal{J}^{2M}. \quad (88)$$

To get the joint distribution of $\mathbf{x}_d$ and $\mathbf{s}$, we first rewrite Eq. (87) as

$$P(\mathbf{x}_d, \mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) = g \left[\begin{pmatrix} \mathbf{V}_d & \mathbf{V}_{da} \\ \mathbf{V}_{da}^\top & \mathbf{V}_a \end{pmatrix}^{-1}, (\mathbf{x}_d, \mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) \right], \quad (89)$$

where we have defined the covariance matrix in the block form

$$\begin{pmatrix} \mathbf{V}_d & \mathbf{V}_{da} \\ \mathbf{V}_{da}^\top & \mathbf{V}_a \end{pmatrix}^{-1} \equiv (\mathbf{I}_{2N} \oplus \mathbf{M}^\top \mathbf{\Omega}) \mathbf{V}_x (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top). \quad (90)$$

and used the property $g(\mathbf{\Sigma}, \mathbf{x}) = g(\mathbf{L} \mathbf{\Sigma} \mathbf{L}^\top, \mathbf{L} \mathbf{x})$ for some invertible matrix $\mathbf{L}$. From here, the distribution of the error and syndrome can be solved as

$$P(\mathbf{x}_d, \mathbf{s}) = \int_{\mathbb{R}^{2M}} d(\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) P(\mathbf{x}_d, \mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) \sum_{\mathbf{k}} \delta(\mathbf{s} - \mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a - \mathbf{k} \sqrt{2\pi}) \quad (91)$$

$$= \sum_{\mathbf{k}} g \left((\mathbf{I}_{2N} \oplus \mathbf{M}^\top \mathbf{\Omega}) \mathbf{V}_x (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{s} - \mathbf{k} \sqrt{2\pi}) \right) \quad (92)$$

$$= \sum_{\mathbf{k}} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da} (\mathbf{s} - \mathbf{k} \sqrt{2\pi})) g(\mathbf{V}_{da|a}^{-1}, \mathbf{s} - \mathbf{k} \sqrt{2\pi}). \quad (93)$$

where we sum over all vector of integers $\mathbf{k} \in \mathbb{Z}^{2M}$, $\delta(\cdot)$ is Dirac delta distribution and $\mathbf{V}_{da|a} = \mathbf{V}_a - \mathbf{V}_{da}^\top \mathbf{V}_d^{-1} \mathbf{V}_{da}$. From (91) to (92), we integrate over $\mathbf{x}_a$. From (92) to (93), we adopt the block form of Eq. (90) and separate the joint distribution into two parts.

The PDF of the syndrome measurement result can be obtained by integrating over $\mathbf{x}_d$,

$$P(\mathbf{s}) = \int_{\mathbb{R}^{2N}} d\mathbf{x}_d P(\mathbf{x}_d, \mathbf{s}) = \sum_{\mathbf{k}} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}). \quad (94)$$

The PDF of the conditional distribution is therefore

$$\begin{aligned} P(\mathbf{x}_d|\mathbf{s}) &= P(\mathbf{x}_d, \mathbf{s})/P(\mathbf{s}) \\ &= \sum_{\mathbf{n}} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{n}\sqrt{2\pi})) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{n}\sqrt{2\pi}) / \left(\sum_{\mathbf{m}} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{m}\sqrt{2\pi}) \right). \end{aligned} \quad (95)$$

The MMSE estimator is obtained by evaluating the mean of the conditional distribution

$$\begin{aligned} \tilde{\mathbf{x}}_d(\mathbf{s}) &\equiv \mathbf{f}_{\text{MMSE}}(\mathbf{s}) = \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \mathbf{x}_d P(\mathbf{x}_d|\mathbf{s}) \\ &= - \sum_{\mathbf{n}} \mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{n}\sqrt{2\pi}) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{n}\sqrt{2\pi}) / \left(\sum_{\mathbf{m}} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{m}\sqrt{2\pi}) \right) \\ &= \frac{- \sum_{\mathbf{n}} \mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{n}\sqrt{2\pi}) \exp\{\sqrt{2\pi}\mathbf{n}^\top \mathbf{V}_{d|a}\mathbf{s} - \pi\mathbf{n}^\top \mathbf{V}_{d|a}\mathbf{n}\}}{\sum_{\mathbf{m}} \exp\{\sqrt{2\pi}\mathbf{m}^\top \mathbf{V}_{d|a}\mathbf{s} - \pi\mathbf{m}^\top \mathbf{V}_{d|a}\mathbf{m}\}}, \end{aligned} \quad (96)$$

where $\mathbf{n}, \mathbf{m} \in \mathbb{Z}^{2M}$. We have thus proven Eq. (12) of Theorem 4.

G.2 Output covariance matrix

Given an error syndrome $\mathbf{s}$ extracted from a general GKP ancilla in a QEC protocol, the element at the i th row and j th column of the output covariance matrix $\mathbf{V}_{\text{out}}$ for N data modes is given as

$$\begin{aligned} [\mathbf{V}_{\text{out}}]_{ij} &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d\mathbf{s} (\mathbf{x}_d - \tilde{\mathbf{x}}_d)_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \\ &= \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} \int_{\mathbb{R}^{2N}} d\mathbf{x}_d (\mathbf{x}_d - \tilde{\mathbf{x}}_d)_{i \times j} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \\ &= \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} \left([\mathbf{V}_d^{-1}]_{ij} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) + g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \left\{ \tilde{\mathbf{x}}_d + [\mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})] \right\}_{i \times j} \right) \\ &= [\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \left\{ \tilde{\mathbf{x}}_d + [\mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})] \right\}_{i \times j} \\ &= [\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} 2\pi \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \left[\frac{\sum_{\mathbf{n}} \mathbf{V}_d^{-1}\mathbf{V}_{da}(\mathbf{n} - \mathbf{k}) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{n}\sqrt{2\pi})}{\sum_{\mathbf{m}} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{m}\sqrt{2\pi})} \right]_{i \times j}. \end{aligned} \quad (97)$$

where we use the notation $\mathbf{x}_{i \times j} = \mathbf{x}_i \mathbf{x}_j$, as the product vector components i and j and have expanded and simplified by using the fact that

$$\int_{\mathbb{R}^{2M}} d\mathbf{x}_d (\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_j) \frac{\det(\mathbf{V}_d)^{\frac{1}{2}}}{(2\pi)^N} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\tau})^\top \mathbf{V}_d(\mathbf{x} - \boldsymbol{\tau})\right\} = [\mathbf{V}_d^{-1}]_{ij} + (\boldsymbol{\mu}_i - \boldsymbol{\tau}_i)(\boldsymbol{\mu}_j - \boldsymbol{\tau}_j). \quad (98)$$

G.3 Explicit analyses of the decoding strategies for GKP-TMS code

Here we show explicit calculation to obtain the estimators for GKP-TMS codes with GKP square lattice. According to Eq. (11) and the encoding operation of the GKP-TMS code in Eq. (3), the random displacement $\mathbf{z} = (z_q^{(1)}, z_p^{(1)}, z_q^{(2)}, z_p^{(2)})$ of the output AGN channel now has the covariance matrix

$$\mathbf{V}_x = \mathbf{S}_{\text{enc}}^{-1} \mathbf{V}_\xi \mathbf{S}_{\text{enc}}^{-\top} = \begin{pmatrix} [G\sigma_1^2 + (G-1)\sigma_2^2]\mathbf{I} & -\sqrt{G(G-1)}(\sigma_1^2 + \sigma_2^2)\mathbf{Z} \\ -\sqrt{G(G-1)}(\sigma_1^2 + \sigma_2^2)\mathbf{Z} & [G\sigma_2^2 + (G-1)\sigma_1^2]\mathbf{I} \end{pmatrix}. \quad (99)$$

The GKP square lattice has generator matrix $\mathbf{M} = \mathbf{I}$. From eq. (13), we have

$$\mathbf{V}_d = \frac{(G-1)\sigma_1^2 + G\sigma_2^2}{\sigma_1^2\sigma_2^2}\mathbf{I}, \quad \mathbf{V}_{da} = \frac{\sqrt{G(G-1)}(\sigma_1^2 + \sigma_2^2)}{\sigma_1^2\sigma_2^2}\boldsymbol{\Omega}\mathbf{Z}, \quad \mathbf{V}_{d|a} = \frac{1}{(G-1)\sigma_1^2 + G\sigma_2^2}\mathbf{I}. \quad (100)$$

The linear estimation becomes

$$\mathbf{f}_{\text{Linear}}(\mathbf{s}) = -\mathbf{V}_d^{-1}\mathbf{V}_{da}\mathbf{s} = \tilde{\mu} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{s}, \quad (101)$$

where

$$\tilde{\mu} = \frac{\sqrt{G(G-1)}(\sigma_1^2 + \sigma_2^2)}{(G-1)\sigma_1^2 + G\sigma_2^2}. \quad (102)$$

Further calculations on the output variance can be found in Refs. [11, 20].

The MMSE estimator of Eq. (12) becomes

$$\mathbf{f}_{\text{MMSE}}(\mathbf{s}) = \frac{\sum_{\mathbf{n}} \tilde{\mu} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} (\mathbf{s} - \mathbf{n}\sqrt{2\pi}) g(\sigma_G^2 \mathbf{I}, \mathbf{s} - \mathbf{n}\sqrt{2\pi})}{\sum_{\mathbf{m}} g(\sigma_G^2 \mathbf{I}, \mathbf{s} - \mathbf{m}\sqrt{2\pi})}. \quad (103)$$

The output covariance matrix of Eq. (97) leads to

$$[\mathbf{V}_{\text{out}}]_{ij} = \frac{\sigma_1^2 \sigma_2^2}{(G-1)\sigma_1^2 + G\sigma_2^2} [\mathbf{I}]_{ij} + \sum_{\mathbf{k}} 2\pi \tilde{\mu}^2 \int_{\mathcal{I}^2} d\mathbf{s} \ g(\sigma_G^2 \mathbf{I}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \left[\frac{\sum_{\mathbf{n}} (\mathbf{n} - \mathbf{k}) g(\sigma_G^2 \mathbf{I}, \mathbf{s} - \mathbf{n}\sqrt{2\pi})}{\sum_{\mathbf{m}} g(\sigma_G^2 \mathbf{I}, \mathbf{s} - \mathbf{m}\sqrt{2\pi})} \right]_{i \times j}. \quad (104)$$

G.4 Lattice basis transformations and estimators

We show that the output joint PDF of the QEC protocol is invariant under a change of lattice basis. We then show that the MMSE estimator of Theorem 4 is likewise invariant but that linear estimation depends upon the choice of the lattice basis.

Theorem 10 (Invariance of the joint PDF). *Consider a generator matrix $\mathbf{M}$ and a change of lattice basis by a unimodular matrix $\mathbf{N}$, which defines another generator matrix $\mathbf{M}' = \mathbf{M}\mathbf{N}^\top$. Let $\mathbf{s} = R_{\sqrt{2\pi}}(\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a)$ be the error syndrome and consider the joint PDF $P(\mathbf{x}_d, \mathbf{s})$. Likewise, let $\mathbf{s}' = R_{\sqrt{2\pi}}(\mathbf{M}'^\top \mathbf{\Omega} \mathbf{x}_a) = R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a)$ be the error syndrome in the new basis and the corresponding joint PDF $P'(\mathbf{x}_d, \mathbf{s}')$. Then, $P'(\mathbf{x}_d, \mathbf{s}') = P(\mathbf{x}_d, \mathbf{s})$.*

Proof. From properties of modulo operations, we have that $R_{\sqrt{2\pi}}(\mathbf{N}^{-1}\mathbf{s}') = R_{\sqrt{2\pi}}(\mathbf{N}^{-1}\mathbf{N}\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a) = \mathbf{s}$, as both $\mathbf{N}$ and $\mathbf{N}^{-1}$ are a matrices of integers. Therefore $\mathbf{N}^{-1}\mathbf{s}' = \mathbf{s} + \ell_{\mathbf{N}^{-1}\mathbf{s}'}\sqrt{2\pi}$ for some vector of integers $\ell_{\mathbf{N}^{-1}\mathbf{s}'}$ determined by $\mathbf{N}^{-1}\mathbf{s}'$. The joint PDF for the data $\mathbf{x}_d$ and the error syndrome $\mathbf{s}'$ in the new basis is then

$$P'(\mathbf{x}_d, \mathbf{s}') = \sum_{\mathbf{k}} g\left((\mathbf{I}_{2N} \oplus (\mathbf{N}\mathbf{M}^\top \mathbf{\Omega}))\mathbf{V}_{\mathbf{x}}(\mathbf{I}_{2N} \oplus (\mathbf{N}\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{s}' - \mathbf{k}\sqrt{2\pi})\right) \quad (105)$$

$$= \sum_{\mathbf{k}} g\left((\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega}))\mathbf{V}_{\mathbf{x}}(\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{N}^{-1}(\mathbf{s}' - \mathbf{k}\sqrt{2\pi}))\right) \quad (106)$$

$$= \sum_{\mathbf{k}'} g\left((\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega}))\mathbf{V}_{\mathbf{x}}(\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{N}^{-1}\mathbf{s}' - \mathbf{k}'\sqrt{2\pi})\right) \quad (107)$$

$$= \sum_{\mathbf{k}'} g\left((\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega}))\mathbf{V}_{\mathbf{x}}(\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{s} - (\mathbf{k}' - \ell_{\mathbf{N}^{-1}\mathbf{s}'})\sqrt{2\pi})\right) \quad (108)$$

$$= \sum_{\mathbf{k}''} g\left((\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega}))\mathbf{V}_{\mathbf{x}}(\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \mathbf{\Omega})^\top), (\mathbf{x}_d, \mathbf{s} - \mathbf{k}''\sqrt{2\pi})\right) \quad (109)$$

$$= P(\mathbf{x}_d, \mathbf{s}). \quad (110)$$

(105) follows directly from the Eq. (93). To move from (106) to (107), we use the fact that $\mathbf{N}^{-1}$ is a unimodular matrix and acting with $\mathbf{N}^{-1}$ on an integer vector simply changes the summation index, $\mathbf{N}^{-1}\mathbf{k} \rightarrow \mathbf{k}'$. In (108), we use $\mathbf{N}^{-1}\mathbf{s}' = \mathbf{s} + \ell_{\mathbf{N}^{-1}\mathbf{s}'}\sqrt{2\pi}$. And in (109), we change the summation index, which is over all integer vectors in $\mathbb{Z}^{2M}$. ■

We next show that the MMSE estimator $\tilde{\mathbf{x}}_d(\mathbf{s}) = \mathbf{f}_{\text{MMSE}}(\mathbf{s})$ of Eq. (96) is invariant under a basis transformation, but before doing so, we prove a useful lemma.

Lemma 11. Given a vector $\mathbf{s} \in \mathbb{R}^{2M}$, a unimodular matrix $\mathbf{N}$, and a function $J(\mathbf{s})$, it follows that

$$\int_{\mathfrak{J}^{2M}} J(\mathbf{s}) d(R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})) = \int_{\mathfrak{J}^{2M}} J(\mathbf{s}) d\mathbf{s}. \quad (111)$$

Proof. For brevity, we use the shorthand $R(\cdot)$ for the modulo function $R_{\sqrt{2\pi}}(\cdot)$. Given a unimodular matrix $\mathbf{N}$, the function $f : \mathbf{s} \rightarrow R(\mathbf{N}\mathbf{s})$ is a bijection from $\mathfrak{J}^{2M}$ to $\mathfrak{J}^{2M}$, as we can find the inverse function $f^{-1} : R(\mathbf{N}\mathbf{s}) \rightarrow \mathbf{s}$ by $\mathbf{s} = R(\mathbf{N}^{-1}R(\mathbf{N}\mathbf{s}))$. We can then divide the region $\mathfrak{J}^{2M}$ into sub regions $A_{\mathbf{k}} = \{\mathbf{x} \in \mathfrak{J}^{2M} | \mathbf{N}\mathbf{x} - \mathbf{k}\sqrt{2\pi} \in \mathfrak{J}^{2M}\}$. There are finite number of these regions since $\mathbf{N}$ and $\mathbf{x}$ are finite. Let $B_{\mathbf{k}} = \{f(\mathbf{x}) | \mathbf{x} \in A_{\mathbf{k}}\}$ be the image of $A_{\mathbf{k}}$. Then $\mathbf{x} \in A_{\mathbf{k}}$ is equivalent to $f(\mathbf{x}) = \mathbf{N}\mathbf{x} - \mathbf{k}\sqrt{2\pi} \in B_{\mathbf{k}}$ by the map f and thus

$$\int_{\mathfrak{J}^{2M}} J(\mathbf{s}) d(f(\mathbf{s})) = \sum_{\mathbf{k}} \int_{B_{\mathbf{k}}} J(\mathbf{s}) d(f(\mathbf{s})) \quad (112)$$

$$= \sum_{\mathbf{k}} \int_{B_{\mathbf{k}}} J(\mathbf{s}) d(\mathbf{N}\mathbf{s} - \mathbf{k}\sqrt{2\pi}) \quad (113)$$

$$= \sum_{\mathbf{k}} \int_{A_{\mathbf{k}}} J(\mathbf{s}) d\mathbf{s} \quad (114)$$

$$= \int_{\mathfrak{J}^{2M}} J(\mathbf{s}) d\mathbf{s}, \quad (115)$$

where we use $|\mathbf{N}| = 1$ from (113) to (114). ■

Theorem 12 (Invariance of MMSE). Consider a generator matrix $\mathbf{M}$ and change of basis by a unimodular matrix $\mathbf{N}$, which defines another generator matrix $\mathbf{M}' = \mathbf{M}\mathbf{N}^\top$. Let $\mathbf{s} = R_{\sqrt{2\pi}}(\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a)$ be the error syndrome and consider the MMSE estimator $\tilde{\mathbf{x}}_d(\mathbf{s}) = \mathbf{f}_{\text{MMSE}}(\mathbf{s})$. Likewise, let $\mathbf{s}' = R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{M}^\top \mathbf{\Omega} \mathbf{x}_a)$ be the error syndrome in the new basis and the corresponding MMSE estimator $\tilde{\mathbf{x}}'_d(\mathbf{s}') = \mathbf{f}'_{\text{MMSE}}(\mathbf{s}')$. Then, $\tilde{\mathbf{x}}'_d(\mathbf{s}') = \tilde{\mathbf{x}}_d(\mathbf{s})$.

Proof. From Theorem 10, we have that $P(\mathbf{x}_d, \mathbf{s}) = P'(\mathbf{x}_d, \mathbf{s}')$ and thus $P(\mathbf{x}_d | \mathbf{s}) = P'(\mathbf{x}_d | \mathbf{s}')$ for the conditional distribution, such that

$$\begin{aligned} \tilde{\mathbf{x}}'_d(\mathbf{s}') &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \mathbf{x}_d P'(\mathbf{x}_d | \mathbf{s}') \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \mathbf{x}_d P(\mathbf{x}_d | R_{\sqrt{2\pi}}\mathbf{s}) \\ &= \tilde{\mathbf{x}}_d(\mathbf{s}), \end{aligned}$$

which was to be proved. ■

Corollary 12.1. The output covariance matrix $\mathbf{V}_{\text{out}}$ describing the noise on the data modes (i.e., the second moments of data error) is invariant under a basis transformation on the ancillary lattice.

This corollary follows immediately from Theorem 12, but we explicitly derive such for clarity.

Proof. Consider the output covariance matrix in the new basis $\mathbf{V}'_{\text{out}}$ [defined in a similar fashion as Eq. (97)] and the following set of equalities,

$$\begin{aligned} [\mathbf{V}'_{\text{out}}]_{ij} &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathfrak{J}^{2M}} d\mathbf{s}' [\mathbf{x}_d - \tilde{\mathbf{x}}'_d(\mathbf{s}')]_{i \times j} P'(\mathbf{x}_d, \mathbf{s}') \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathfrak{J}^{2M}} d\mathbf{s}' [\mathbf{x}_d - \tilde{\mathbf{x}}_d(\mathbf{s})]_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathfrak{J}^{2M}} d(R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})) (\mathbf{x}_d - \tilde{\mathbf{x}}_d(\mathbf{s}))_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathfrak{J}^{2M}} d\mathbf{s} (\mathbf{x}_d - \tilde{\mathbf{x}}_d(\mathbf{s}))_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \\ &= [\mathbf{V}_{\text{out}}]_{ij}. \end{aligned}$$

We have used Theorems 10 and 12 to go from the first equality to the second equality and used the relation $\mathbf{s}' = R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})$ and Lemma 11 to go from the third equality to the fourth equality. The final equality follows by definition of the covariance matrix. ■

We now discuss linear estimation and show that such is not invariant under a lattice basis transformation.

Theorem 13. *Linear estimation is not invariant under a lattice basis transformation.*

Proof. Considering only the $\mathbf{n} = \mathbf{0}$ element in the sum of Eq. (96), we get the linear estimator

$$\mathbf{f}_{\text{Linear}}(\mathbf{s}) = -\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{s}, \quad (116)$$

which is obviously not invariant under a basis transformation. Indeed, consider a basis transformation such that $\mathbf{s}' = R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})$. The estimator in Eq. (116) then changes to

$$\begin{aligned} \mathbf{f}_{\text{Linear}}(\mathbf{s}') &= -\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{s}' \\ &= -\mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}). \end{aligned} \quad (117)$$

and $\mathbf{f}_{\text{Linear}}(\mathbf{s}) \neq \mathbf{f}_{\text{Linear}}(\mathbf{s}')$. ■

As a consequence, the output covariance matrix for linear estimation also changes under a change of lattice basis. To show this explicitly, let us first write the output covariance matrix $\mathbf{V}_{\text{out}}$ in full,

$$\begin{aligned} [\mathbf{V}_{\text{out}}]_{ij} &= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d\mathbf{s} (\mathbf{x}_d - \mathbf{f}_{\text{Linear}}(\mathbf{s}))_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \\ &= \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} \int_{\mathbb{R}^{2N}} d\mathbf{x}_d (\mathbf{x}_d - \mathbf{f}_{\text{Linear}}(\mathbf{s}))_{i \times j} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \\ &= \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} [\mathbf{V}_d^{-1}]_{ij} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) + \left[-\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{s} + \mathbf{V}_d^{-1} \mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi}) \right]_{i \times j} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \\ &= [\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} 2\pi [\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{k}]_{i \times j} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}), \end{aligned} \quad (118)$$

where we have used the expression of $P(\mathbf{x}_d, \mathbf{s})$ in (93) and (98) in the first and second steps, respectively. Under change of basis, the new elements of the covariance matrix $\mathbf{V}'_{\text{out}}$ are

$$[\mathbf{V}'_{\text{out}}]_{ij} = \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d\mathbf{s}' [\mathbf{x}_d - \mathbf{f}_{\text{Linear}}(\mathbf{s}')]_{i \times j} P'(\mathbf{x}_d, \mathbf{s}') \quad (119)$$

$$= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d(R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})) [\mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})]_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \quad (120)$$

$$= \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d\mathbf{s} [\mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})]_{i \times j} P(\mathbf{x}_d, \mathbf{s}) \quad (121)$$

$$= \sum_{\mathbf{k}} \int_{\mathbb{R}^{2N}} d\mathbf{x}_d \int_{\mathcal{J}^{2M}} d\mathbf{s} [\mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})]_{i \times j} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \quad (122)$$

$$= \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) \left\{ [\mathbf{V}_d^{-1}]_{ij} + [\mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{V}_d^{-1} \mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})]_{i \times j} \right\} \quad (123)$$

$$= [\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) [\mathbf{V}_d^{-1} \mathbf{V}_{da} R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{V}_d^{-1} \mathbf{V}_{da}(\mathbf{s} - \mathbf{k}\sqrt{2\pi})]_{i \times j} \quad (124)$$

$$= [\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) [\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{k}\sqrt{2\pi} + \mathbf{V}_d^{-1} \mathbf{V}_{da} (R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{s})]_{i \times j} \quad (125)$$

$$\begin{aligned} &= \underbrace{[\mathbf{V}_d^{-1}]_{ij} + \sum_{\mathbf{k}} 2\pi [\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{k}]_{i \times j} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi})}_{=[\mathbf{V}_{\text{out}}]_{ij}} \\ &\quad + \sum_{\mathbf{k}} \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) [\mathbf{V}_d^{-1} \mathbf{V}_{da} (R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{s})]_{i \times j} \\ &\quad + \sum_{\mathbf{k}} \sqrt{2\pi} [\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{k}]_i \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) [\mathbf{V}_d^{-1} \mathbf{V}_{da} (R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{s})]_j \\ &\quad + \sum_{\mathbf{k}} \sqrt{2\pi} [\mathbf{V}_d^{-1} \mathbf{V}_{da} \mathbf{k}]_j \int_{\mathcal{J}^{2M}} d\mathbf{s} g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}) [\mathbf{V}_d^{-1} \mathbf{V}_{da} (R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s}) - \mathbf{s})]_i. \end{aligned} \quad (126)$$

From (119) to (120), we use $\mathbf{s}' = R_{\sqrt{2\pi}}(\mathbf{N}\mathbf{s})$, $P'(\mathbf{x}_d, \mathbf{s}') = P(\mathbf{x}_d, \mathbf{s})$ and the expression in (117). From (120) to (121), we apply Lemma 11. To obtain (122), we input the explicit equation of (93). The integral of $\mathbf{x}_d$ is simplified with (98) from (122) to (123). From (123) to (125), we first recognize that the $\mathbf{V}_d^{-1}$ part does not depend on $\mathbf{s}$ and then reorganized the integrand. To obtain (125), we make use of the definition $\mathbf{x}_{i \times j} = \mathbf{x}_i \mathbf{x}_j$, where $\mathbf{x}_i$ is the i th component of $\mathbf{x}$, so that $(\mathbf{a} + \mathbf{b})_{i \times j} = (\mathbf{a} + \mathbf{b})_i (\mathbf{a} + \mathbf{b})_j = \mathbf{a}_{i \times j} + \mathbf{b}_{i \times j} + \mathbf{a}_i \mathbf{b}_j + \mathbf{a}_j \mathbf{b}_i$. Comparing equations (118) and (126), we see that the linear estimation is not invariant under a lattice basis transformation.

H Lower bound on output variance for multimode codes

We find a lower bound on the output variance for multimode GKP codes (with N data modes and $M \geq N$ GKP ancilla modes), in a similar vein to the single-mode lower bound found in Ref. [11]. We first provide an effective reduction from multimode additive noise to independent single-mode additive noises which we utilize to prove the result.

Given a N mode input data state ρ , the output data state ρ' after a round of error correction with a multimode GKP stabilizer code (which protects against AGN) is related to the input by an additive non-Gaussian noise channel $\tilde{\mathcal{N}}$ via

$$\rho' = \tilde{\mathcal{N}}(\rho) = \int_{\mathbb{R}^{2N}} d\mathbf{e}_d P(\mathbf{e}_d) D_{\mathbf{e}_d} \rho D_{\mathbf{e}_d}^\dagger, \quad (127)$$

where $P(\mathbf{e}_d)$ is a multivariate non-Gaussian pdf that describes the residual data noise $\mathbf{e}_d$ —e.g., $\langle \mathbf{e}_d \rangle = 0$ and $\langle \{(\mathbf{e}_d)_i, (\mathbf{e}_d)_j\} \rangle = (\mathbf{V}_{\text{out}})_{ij}$; for brevity in what follows, we let $\mathbf{V} \equiv \mathbf{V}_{\text{out}}$. Consider the Gaussian approximations ρ_G and ρ'_G for the input and output states ρ and ρ' , which agree with ρ and ρ' at the level of first and second moments. We have that

$$\rho'_G = \mathcal{N}_{\mathbf{V}}(\rho_G) = \int_{\mathbb{R}^{2N}} d\mathbf{e}_d g(\mathbf{V}, \mathbf{e}_d) D_{\mathbf{e}_d} \rho_G D_{\mathbf{e}_d}^\dagger, \quad (128)$$

where $g(\mathbf{V}, \mathbf{e}_d)$ is a Gaussian distribution according to Eq. (83) with the same first and second moments as the non-Gaussian pdf $P(\mathbf{e}_d)$ and $\mathcal{N}_{\mathbf{V}}$ is a multimode AGN channel with noise matrix $\mathbf{V}$.

Let $\mathbf{S}$ be the symplectic transformation that diagonalizes $\mathbf{V}$ such that $\mathbf{S}\mathbf{V}\mathbf{S}^\top = \bigoplus_{i=1}^N \nu_i \mathbf{I}_2$, where $\nu_i \geq 0$ are the symplectic eigenvalues (single-mode variances) of $\mathbf{V}$. We append Gaussian unitaries $U_{\mathbf{S}}^\dagger$ and $U_{\mathbf{S}}$ before and after error correction, respectively, and define the resulting channel $\tilde{\mathcal{N}}^{(\mathbf{S})} \equiv \mathcal{U}_{\mathbf{S}} \circ \tilde{\mathcal{N}} \circ \mathcal{U}_{\mathbf{S}}^{-1}$, where $\mathcal{U}_{\mathbf{S}}$ is the unitary channel of $U_{\mathbf{S}}$. Then,

$$\begin{aligned} \rho'(\mathbf{S}) &= \tilde{\mathcal{N}}^{(\mathbf{S})}(\rho) \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{e}_d P(\mathbf{e}_d) \left(U_{\mathbf{S}} D_{\mathbf{e}_d} U_{\mathbf{S}}^\dagger \right) \rho \left(U_{\mathbf{S}} D_{\mathbf{e}_d}^\dagger U_{\mathbf{S}} \right) \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{e}_d P(\mathbf{e}_d) D_{\mathbf{S}\mathbf{e}_d} \rho D_{\mathbf{S}\mathbf{e}_d}^\dagger \\ &= \int_{\mathbb{R}^{2N}} d\mathbf{e}_d P(\mathbf{S}^{-1}\mathbf{e}_d) D_{\mathbf{e}_d} \rho D_{\mathbf{e}_d}^\dagger, \end{aligned} \quad (129)$$

where Eq. (26) was used in the second line and a change of variables was used in the third line, along

with the fact that $d(\mathbf{S}^{-1}\mathbf{e}_d) = d\mathbf{e}_d$. The Gaussian approximation for $\rho'(\mathbf{S})$ is then simply,

$$\rho'_G(\mathbf{S}) = \mathcal{N}_{\mathbf{S}\mathbf{V}\mathbf{S}^\top}(\rho_G) = \bigotimes_{i=1}^N \mathcal{N}_{\nu_i}(\rho_G). \quad (130)$$

This is due to the fact that displacements are always local, $D_{\mathbf{e}_d} = \bigotimes_i D_{\mathbf{e}_{d_i}}$, and that $P(\mathbf{S}^{-1}\mathbf{e}_d)$ has a corresponding Gaussian distribution satisfying

$$\begin{aligned} g(\mathbf{V}, \mathbf{S}^{-1}\mathbf{e}_d) &= g(\mathbf{S}\mathbf{V}\mathbf{S}^\top, \mathbf{e}_d) \\ &= g\left(\bigoplus_i \nu_i \mathbf{I}_2, \mathbf{e}_d\right) \\ &= \prod_i g(\nu_i \mathbf{I}_2, \mathbf{e}_{d_i}), \end{aligned} \quad (131)$$

where $\mathbf{e}_{d_i} \in \mathbb{R}^2$. In other words, at the level of first and second moments, we have decorrelated the residual noises on the data modes via $\mathbf{S}$ and reduced the channel to a direct product of independent AGN channels. We will use this simplification in what follows to prove a lower bound on the quantum capacity of a general non-Gaussian additive noise channel $\tilde{\mathcal{N}}$, which includes the single-mode case proven in Ref. [11].

Lemma 14. *Consider a N mode additive non-Gaussian noise channel $\tilde{\mathcal{N}}$ [see Eq. (127)]. Let the geometric mean error of the channel be defined as $\sigma_{\text{GM}}^2 = \sqrt[2N]{\det \mathbf{V}}$, where $\mathbf{V}$ is the $2N \times 2N$ covariance matrix of $\tilde{\mathcal{N}}$. Then the quantum capacity C_Q of $\tilde{\mathcal{N}}$ has the following lower bound,*

$$C_Q(\tilde{\mathcal{N}}) \geq \max \left[0, N \log_2 \left(\frac{1}{e \sigma_{\text{GM}}^2} \right) \right]. \quad (132)$$

Proof. To prove this result, we first consider the channel $\tilde{\mathcal{N}}^{(\mathbf{S})}$ of Eq. (129). Since $\tilde{\mathcal{N}}^{(\mathbf{S})}$ is related to $\tilde{\mathcal{N}}$ by unitary pre- and post-processing, the quantum capacities of the two channels are equivalent, $C_Q(\tilde{\mathcal{N}}) = C_Q(\tilde{\mathcal{N}}^{(\mathbf{S})})$. We focus on the channel $\tilde{\mathcal{N}}^{(\mathbf{S})}$ from hereon.

Observe that the single-shot coherent information places a lower bound on the quantum capacity [41, 42, 43]

$$C_Q(\tilde{\mathcal{N}}^{(\mathbf{S})}) \geq \max_{\rho} I_c(\rho, \tilde{\mathcal{N}}^{(\mathbf{S})}) \quad (133)$$

where $I_c(\rho, \tilde{\mathcal{N}}^{(\mathbf{S})}) = S(\tilde{\mathcal{N}}^{(\mathbf{S})}(\rho)) - S(\tilde{\mathcal{N}}^{(\mathbf{S})c}(\rho))$ is the coherent information, S is the von Neumann entropy and $\tilde{\mathcal{N}}^{(\mathbf{S})c}$ is the complementary channel of $\tilde{\mathcal{N}}^{(\mathbf{S})}$. By Gaussian extremality [44], one can show that $I_c(\rho, \tilde{\mathcal{N}}^{(\mathbf{S})}) \geq S(\tilde{\mathcal{N}}^{(\mathbf{S})}(\rho)_G) - S(\tilde{\mathcal{N}}^{(\mathbf{S})c}(\rho)_G)$ where,

e.g., $\tilde{\mathcal{N}}^{(\mathcal{S})}(\rho)_G$ is a Gaussian state with the same first and second moments as $\tilde{\mathcal{N}}^{(\mathcal{S})}(\rho)$. In fact, $\tilde{\mathcal{N}}^{(\mathcal{S})}(\rho)_G = \bigotimes_{i=1}^N \mathcal{N}_{\nu_i}(\rho_G)$ from Eq. (130), which is just a direct product of AGN channels acting on an input Gaussian state ρ_G ; similarly, $\tilde{\mathcal{N}}^{(\mathcal{S})^c}(\rho)_G = \bigotimes_{i=1}^N \mathcal{N}_{\nu_i}^c(\rho_G)$. Therefore, $I_c(\rho, \tilde{\mathcal{N}}^{(\mathcal{S})}) \geq I_c(\rho_G, \bigotimes_{i=1}^N \mathcal{N}_{\nu_i})$. In other words, we have reduced the problem to analyzing a set of independent AGN channels with Gaussian input states; this has been done in previous works for a single mode with an input thermal state $\Theta_{\bar{n}}$ [45, 25]. In the energy unconstrained setting ($\bar{n} \rightarrow \infty$), the single mode result is $\max_{\rho_G} I_c(\rho_G, \mathcal{N}_{\nu_i}) = \log(1/e\nu_i)$. Hence, for the multimode channel $\bigotimes \mathcal{N}_{\nu_i}$,

$$\max_{\rho_G} I_c \left(\rho_G, \bigotimes_{i=1}^N \mathcal{N}_{\nu_i} \right) \geq \sum_{i=1}^N \log \left(\frac{1}{e\nu_i} \right) \quad (134)$$

$$= \log \left(\frac{1}{e^N \prod_{i=1}^N \nu_i} \right). \quad (135)$$

Since ν_i are the symplectic eigenvalues of the channel covariance matrix $\mathbf{V}$, $\det \mathbf{V} = \prod_{i=1}^N \nu_i^2$, and the result is proved. ■

Now we consider the upper bound of the quantum capacity. We begin by quoting a lemma in Refs. [25, 20].

Lemma 15. *Given an $N + M$ mode AGN channel, $\mathcal{N}_{\mathbf{Y}}$, with noise covariance matrix $\mathbf{Y} \geq 0$, the quantum capacity C_Q then has the following upper bound [25, 20],*

$$C_Q(\mathcal{N}_{\mathbf{Y}}) \leq \sum_{i=1}^{N+M} \log_2 \left(\frac{1 - \sigma_i^2}{\sigma_i^2} \right), \quad (136)$$

where σ_i^2 are the symplectic eigenvalues of $\mathbf{Y}$ (i.e., the single-mode variances).

From the above results of lower and upper bounds, we can have the following theorem.

Theorem 16 (Lower bound for output noise of multimode codes). *Consider a multimode GKP stabilizer code which consumes $M \geq N$ ancillary GKP modes to protect N data modes against independent AGN. The joint noise channel across all modes is given by $\bigotimes_{i=1}^{N+M} \mathcal{N}_{\sigma_i^2}$ with variances σ_i^2 . Let the $2N \times 2N$ output noise matrix of the data be $\mathbf{V}_{\text{out}}$, with RMS and GM errors given as $\bar{\sigma}_{\text{RMS}}^2 = \text{Tr}\{\mathbf{V}_{\text{out}}\}/2N$ and*

$\bar{\sigma}_{\text{GM}}^2 = \sqrt[2N]{\det \mathbf{V}_{\text{out}}}$, respectively. Then,

$$\bar{\sigma}_{\text{RMS}} \geq \bar{\sigma}_{\text{GM}} \geq \frac{1}{\sqrt{e}} \sqrt[2N]{\left(\prod_{i=1}^{N+M} \frac{\sigma_i^2}{1 - \sigma_i^2} \right)}. \quad (137)$$

Proof. The proof for the geometric mean error follows by concurrently considering the lower bound and upper bound of Lemmas 14 and 15, respectively. In detail, we can use a multimode GKP code—and thus the resultant additive non-Gaussian channel $\tilde{\mathcal{N}}$ —to transmit information across the AGN channels $\bigotimes_{i=1}^{N+M} \mathcal{N}_{\sigma_i^2}$. The transmission rate is bounded from below by the quantum capacity $C_Q(\tilde{\mathcal{N}})$ since $C_Q(\tilde{\mathcal{N}})$ is achievable, which in turn is bounded from below by Lemma 14. Likewise, the transmission rate is bounded from above by the quantum capacity of the original channel, $C_Q(\bigotimes_{i=1}^{N+M} \mathcal{N}_{\sigma_i^2})$, which in turn is bounded from above by Lemma 15. This establishes the second inequality in Eq. (137). The bound on the RMS error [the first inequality in Eq. (137)] follows from the fact that the arithmetic mean is an upper bound on the geometric mean. ■

Corollary 16.1. *For iid AGN, such that $\sigma_i = \sigma \forall i \in \{1, \dots, M + N\}$, $\bar{\sigma}_{\text{RMS}} \geq \bar{\sigma}_{\text{GM}} \geq \frac{1}{\sqrt{e}} \left(\frac{\sigma^2}{1 - \sigma^2} \right)^{\frac{N+M}{2N}}$.*

Thus, for general multimode codes, if the number of ancilla modes is equal to the number of data modes ($M = N$), then error suppression is at most quadratic in the initial noise σ . The performance can be further enhanced with concatenated codes ($M > N$), i.e. $\bar{\sigma}_{\text{GM}} \sim \sigma^{1 + \frac{M}{N}}$. We find an upper bound on the break-even point, $\sigma^* \leq 1/\sqrt{2}$; see Fig. 5 of the main text. For $\sigma > 1/\sqrt{2}$, there is no gain to be had from QEC. This agrees with the fact that the upper bound for the quantum capacity of the AGN channel [Eq. (136)] vanishes as $\sigma \rightarrow 1/\sqrt{2}$.

I Classical entropies for the iid Gaussian channel

The differential entropy of multivariate normal distribution is [23]

$$S(\mathbf{x}) = \frac{1}{2} \ln [(2\pi e)^K \det \mathbf{V}], \quad (138)$$

where K is the dimension of the random variable $\mathbf{x}$. In the case of N data modes and M ancilla modes [$K = 2(M+N)$], the correlated data and ancilla noises (just prior to corrective displacements on the data) has a covariance matrix

$$\begin{aligned} \mathbf{V}_{\mathbf{x}} &= \mathbf{S}_{\text{enc}}^{-1} \mathbf{V}_{\xi} \mathbf{S}_{\text{enc}}^{-\top} \\ &= (\mathbf{\Lambda}_d^{-1} \oplus \mathbf{\Lambda}_a^{-1}) \left[\sigma^2 \left(\bigoplus_{i=1}^N \mathbf{S}_{G_i} \mathbf{S}_{G_i}^{\top} \right) \oplus \mathbf{I}_{2(M-N)} \right] (\mathbf{\Lambda}_d^{-\top} \oplus \mathbf{\Lambda}_a^{-\top}), \end{aligned} \quad (139)$$

where $\mathbf{\Lambda}_d$ and $\mathbf{\Lambda}_a$ are local symplectic transformations on the data and ancilla, respectively. The general form in terms of two-mode squeezing blocks, for an arbitrary Gaussian encoding $\mathbf{S}_{\text{enc}}$, is a consequence of Lemma 1.

Since the local symplectic transforms do not change the determinants of the subsystems, we can easily calculate all entropies of interest,

$$S(\mathbf{x}_d) = \sum_{i=1}^N \ln [2\pi e(2G_i - 1)\sigma^2], \quad (140)$$

$$S(\mathbf{x}_a) = \sum_{i=1}^N \ln [2\pi e(2G_i - 1)\sigma^2] + (M - N) \ln(2\pi e\sigma^2), \quad (141)$$

$$S(\mathbf{x}_d, \mathbf{x}_a) = (N + M) \ln(2\pi e\sigma^2). \quad (142)$$

The mutual information of the joint distribution is then

$$\begin{aligned} I(\mathbf{x}_d; \mathbf{x}_a) &= S(\mathbf{x}_d) + S(\mathbf{x}_a) - S(\mathbf{x}_d, \mathbf{x}_a) \\ &= 2 \sum_{i=1}^N \ln(2G_i - 1), \end{aligned} \quad (143)$$

and the conditional differential entropy is

$$\begin{aligned} S(\mathbf{x}_d | \mathbf{x}_a) &= S(\mathbf{x}_d, \mathbf{x}_a) - S(\mathbf{x}_a) \\ &= \sum_{i=1}^N \ln \left(\frac{2\pi e\sigma^2}{2G_i - 1} \right). \end{aligned} \quad (144)$$

In obtaining these results, we have utilized the fact that $\mathbf{\Lambda}_d \oplus \mathbf{\Lambda}_a$ does not change any of the above quantities. So in the calculations, we can actually consider the random variables $\mathbf{\Lambda}_d \mathbf{x}_d$ and $\mathbf{\Lambda}_a \mathbf{x}_a$ for simplicity. In terms of these variables, the data modes (and ancilla modes) are uncorrelated amongst themselves, and there are only two-mode correlations for each data-ancilla mode pair. Moreover, due to the structure of two-mode squeezing, there are only qq and pp correlations and no qp correlations. Therefore, the total entropy, mutual information, and conditional entropy originate from: (1) two-mode pairs (one data and one ancilla) due to two-mode squeezing and (2) individual quadratures (qq or pp correlations but no qp correlations). As an example, $S(\mathbf{x}_d | \mathbf{x}_a) = S(\mathbf{\Lambda}_d \mathbf{x}_d | \mathbf{\Lambda}_a \mathbf{x}_a) = \sum_{i=1}^{2N} S(\mathbf{\Lambda}_d \mathbf{x}_d | \mathbf{\Lambda}_a \mathbf{x}_a)_i$, where $S(\mathbf{\Lambda}_d \mathbf{x}_d | \mathbf{\Lambda}_a \mathbf{x}_a)_i = \ln \left(\frac{2\pi e\sigma^2}{2G_i - 1} \right) / 2$ is the conditional entropy for the qq (pp) quadrature correlations of the i th data-ancilla mode pair. Everything considered, we see that we can treat each data quadrature independently and sum their individual contributions to calculate global properties.

J Error from approximated states

Given the noise model for finite GKP squeezing discussed in Section 9, it is not hard to show that the error syndrome changes to

$$\mathbf{s} = R_{\sqrt{2\pi}} \left(\mathbf{M}^{\top} \mathbf{\Omega} (\boldsymbol{\xi}_{\text{GKP}}^{(1)} + \mathbf{x}_a) + \boldsymbol{\xi}_{\text{GKP}}^{(2)} \right) \quad (145)$$

$$= R_{\sqrt{2\pi}} \left(\mathbf{M}^{\top} \mathbf{\Omega} \mathbf{x}_a + \boldsymbol{\delta} \right), \quad (146)$$

where $\boldsymbol{\xi}_{\text{GKP}}^{(1)} \sim \mathcal{N}(0, \sigma_{\text{GKP}}^2)$ is the GKP noise from the ancilla modes, $\boldsymbol{\xi}_{\text{GKP}}^{(2)} \sim \mathcal{N}(0, \sigma_{\text{GKP}}^2)$ is the GKP noise from the noisy stabilizer measurements, and $\boldsymbol{\delta} = \mathbf{M}^{\top} \mathbf{\Omega} \boldsymbol{\xi}_{\text{GKP}}^{(1)} + \boldsymbol{\xi}_{\text{GKP}}^{(2)}$. The PDF of $\boldsymbol{\delta}$ is

$$P(\boldsymbol{\delta}) = g(\sigma_{\text{GKP}}^2 (\mathbf{M}^{\top} \mathbf{M} + \mathbf{I}_{2M}), \boldsymbol{\delta}). \quad (147)$$

We can show that the joint distribution of $\mathbf{x}_d$ and the error syndrome $\mathbf{s}$ is

$$P(\mathbf{x}_d, \mathbf{s}) = \int_{\mathbb{R}^{4M}} d(\mathbf{M}^\top \boldsymbol{\Omega} \mathbf{x}_a) d\boldsymbol{\delta} P(\mathbf{x}_d, \mathbf{M}^\top \boldsymbol{\Omega} \mathbf{x}_a) P(\boldsymbol{\delta}) \sum_{\mathbf{k}} \delta(\mathbf{s} - \mathbf{M}^\top \boldsymbol{\Omega} \mathbf{x}_a - \boldsymbol{\delta} - \mathbf{k}\sqrt{2\pi}) \quad (148)$$

$$= \sum_{\mathbf{k}} \int_{\mathbb{R}^{2M}} d\boldsymbol{\delta} g\left((\mathbf{I}_{2N} \oplus \mathbf{M}^\top \boldsymbol{\Omega}) \mathbf{V}_{\mathbf{x}} (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \boldsymbol{\Omega})^\top), (\mathbf{x}_d, \mathbf{s} - \boldsymbol{\delta} - \mathbf{k}\sqrt{2\pi})\right) g(\sigma_{\text{GKP}}^2 (\mathbf{M}^\top \mathbf{M} + \mathbf{I}_{2M}), \boldsymbol{\delta}) \quad (149)$$

$$= \sum_{\mathbf{k}} g\left((\mathbf{I}_{2N} \oplus \mathbf{M}^\top \boldsymbol{\Omega}) \mathbf{V}_{\mathbf{x}} (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \boldsymbol{\Omega})^\top + (\mathbf{0}_{2N} \oplus \sigma_{\text{GKP}}^2 (\mathbf{M}^\top \mathbf{M} + \mathbf{I}_{2M}))), (\mathbf{x}_d, \mathbf{s} - \mathbf{k}\sqrt{2\pi})\right) \quad (150)$$

$$\equiv \sum_{\mathbf{k}} g(\mathbf{V}_d^{-1}, \mathbf{x}_d + \mathbf{V}_d^{-1} \mathbf{V}_{da} (\mathbf{s} - \mathbf{k}\sqrt{2\pi})) g(\mathbf{V}_{d|a}^{-1}, \mathbf{s} - \mathbf{k}\sqrt{2\pi}), \quad (151)$$

where we have redefined the notation $\mathbf{V}_d, \mathbf{V}_{da}, \mathbf{V}_{d|a}$ to incorporate finite-squeezing GKP noise.

We use a moment-generating function to prove the identities (149) and (150). Let $\mathbf{X}$ and $\mathbf{Y}$ be the vectors of independent random Gaussian variables, then

$$P(\mathbf{X} = (\mathbf{x}_d, \mathbf{s} - \mathbf{k}\sqrt{2\pi})) = g\left((\mathbf{I}_{2N} \oplus \mathbf{M}^\top \boldsymbol{\Omega}) \mathbf{V}_{\mathbf{x}} (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \boldsymbol{\Omega})^\top), (\mathbf{x}_d, \mathbf{s} - \mathbf{k}\sqrt{2\pi})\right), \quad (152)$$

$$P(\mathbf{Y} = (\mathbf{0}_{2N}, \boldsymbol{\delta})) = g((\mathbf{0}_{2N} \oplus \sigma_{\text{GKP}}^2 (\mathbf{M}^\top \mathbf{M} + \mathbf{I}_{2M})), (\mathbf{0}_{2N}, \boldsymbol{\delta})). \quad (153)$$

The corresponding moment generating functions are

$$M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{-\frac{1}{2} \mathbf{t}^\top (\mathbf{I}_{2N} \oplus \mathbf{M}^\top \boldsymbol{\Omega}) \mathbf{V}_{\mathbf{x}} (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \boldsymbol{\Omega})^\top) \mathbf{t}\right\}, \quad (154)$$

$$M_{\mathbf{Y}}(\mathbf{t}) = \exp\left\{-\frac{1}{2} \mathbf{t}^\top (\mathbf{0}_{2N} \oplus \sigma_{\text{GKP}}^2 (\mathbf{M}^\top \mathbf{M} + \mathbf{I}_{2M})) \mathbf{t}\right\}. \quad (155)$$

Let $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$. Since $\mathbf{X}$ and $\mathbf{Y}$ are independent, then

$$M_{\mathbf{Z}}(\mathbf{t}) = M_{\mathbf{X}}(\mathbf{t}) M_{\mathbf{Y}}(\mathbf{t}) \quad (156)$$

$$= \exp\left\{-\frac{1}{2} \mathbf{t}^\top \left[(\mathbf{I}_{2N} \oplus \mathbf{M}^\top \boldsymbol{\Omega}) \mathbf{V}_{\mathbf{x}} (\mathbf{I}_{2N} \oplus (\mathbf{M}^\top \boldsymbol{\Omega})^\top) + (\mathbf{0}_{2N} \oplus \sigma_{\text{GKP}}^2 (\mathbf{M}^\top \mathbf{M} + \mathbf{I}_{2M}))\right] \mathbf{t}\right\}, \quad (157)$$

and we have

$$(92) = P(\mathbf{Z} = \mathbf{X} + \mathbf{Y}) = (150). \quad (158)$$

This complete the proof.