

Improved quantum capacity bounds of Gaussian loss channels and achievable rates with Gottesman-Kitaev-Preskill codes

Kyungjoo Noh,^{1,2,*} Victor V. Albert,³ and Liang Jiang^{1,2}

¹*Department of Applied Physics and Physics, Yale University, New Haven, Connecticut 06520, USA*

²*Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*

³*Walter Burke Institute for Theoretical Physics and Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

Gaussian loss channels are of particular importance since they model realistic optical communication channels. Except for special cases, quantum capacity of Gaussian loss channels is not yet known completely. In this paper, we provide improved upper bounds of Gaussian loss channel capacity, both in the energy-constrained and unconstrained scenarios. We briefly review the Gottesman-Kitaev-Preskill (GKP) codes and discuss their experimental implementation. We then prove, in the energy-unconstrained case, that the GKP codes achieve the quantum capacity of Gaussian loss channels up to at most a constant gap from the improved upper bound. In the energy-constrained case, we formulate a biconvex encoding and decoding optimization problem to maximize the entanglement fidelity. The biconvex optimization is solved by an alternating semidefinite programming (SDP) method and we report that, starting from random initial codes, our numerical optimization yields GKP codes as the optimal encoding in a practically relevant regime.

I. INTRODUCTION

Quantum communication is an important area of quantum technology wherein classical communication is enriched with non-local quantum entanglement and quantum state transmission [1]. Similarly as in the classical case, practical quantum communication channels are inevitably noisy, and thus the quantum information should be encoded and decoded in a non-trivial way such that the effects of channel noise can be reversed by quantum error correction [2].

Quantum capacity of a channel quantifies the maximum amount of quantum bits per channel use that can be transmitted faithfully (upon optimal encoding and decoding) in the limit of infinite channel uses. Analogous to mutual information in the classical communication theory, *regularized coherent information* of a channel characterizes the channel's quantum capacity [3–5]. Unlike its classical counterpart, however, coherent information might be *superadditive*, which makes it hard to evaluate the channel's quantum capacity [6–9].

Bosonic Gaussian channels [10–12] are among the most studied quantum channels due to its relevance to optical communication. The bosonic pure-loss channel is a special case of Gaussian loss channels. Since coherent information of the bosonic pure-loss channel is additive, its quantum capacity is known [13–15] (see [16] for the general formalism of energy-constrained quantum capacity, and our Theorem 9 for a pedagogical self-contained derivation of energy-constrained quantum capacity of bosonic pure-loss channels). For general Gaussian loss channels with added thermal noise, a lower bound of quantum capacity can be obtained by evaluating one-shot coherent information of the channel [13], and several upper bounds are obtained by using, e.g., data-processing inequality [13, 17–19].

Evaluation of the quantum capacity, however, does not lend explicit encoding and decoding strategies achieving the capac-

ity. In parallel with the characterization of quantum capacity, many bosonic quantum error-correcting codes have also been developed over the past two decades, using a few coherent states [20–25], position/momentum eigenstates [26–31], finite superpositions of the Fock states [32–38] of the bosonic modes. Hybrid CV-DV schemes have also been proposed [39, 40]. In the context of communication theory, the Gottesman-Kitaev-Preskill (GKP) codes [28] are particularly interesting as they can achieve one-shot coherent information of the Gaussian random displacement channel [41]. In our recent work, we showed, both numerically and analytically, that the GKP codes also exhibit excellent performance against the bosonic pure-loss errors, outperforming many the other bosonic codes [42]. It is important to understand why GKP code can perform so well, which will guide us to search for better bosonic codes.

In this paper, we prove that the GKP codes achieve quantum capacity of Gaussian loss channels up to at most a constant gap from an upper bound of the quantum capacity. In section II, we review and summarize key properties of Gaussian loss, amplification and random displacement channels, which will be used in later sections. In section III, we provide an improved upper bound of quantum capacity of the Gaussian loss channel, both in energy-constrained and unconstrained cases by introducing a slight modification of an earlier result in [18] (Theorems 10, 11 and 12; see also Eq.(42) for an improved upper bound of Gaussian random displacement channel capacity). In section IV, we discuss experimental implementation of the GKP codes and prove that the achievable quantum communication rate with the GKP codes deviates only by a constant number of qubits per channel use from our improved upper bound, assuming an energy-unconstrained scenario (Theorem 14). In section V, we address the energy-constrained encoding scenario and formulate a biconvex optimization problem to find an optimal set of encoding and decoding which maximize the entanglement fidelity. We solve the biconvex optimization using alternating semidefinite programming (SDP), and show that the GKP code emerges as an optimal solution.

* kyungjoo.noh@yale.edu

II. PRELIMINARIES

In this section, we provide a summary of preliminary facts about Gaussian channels which will be referenced in later sections. For introduction to bosonic mode, Gaussian state and Gaussian unitary operations, we refer the readers to [11, 12] (and Appendix A, which is a part of [11] relevant to this paper translated into our notation). Lemmas 6 and 8 are the key facts which will be used to derive our two main results in subsection IV D (Theorem 14) and subsection III B (Theorems 10 and 11), respectively.

A. Gaussian channels

Let $\mathcal{H}^{\otimes N}$ be a Hilbert space associated with N bosonic modes, $\mathcal{L}(\mathcal{H}^{\otimes N})$ the space of linear operators acting on $\mathcal{H}^{\otimes N}$ and $\mathcal{D}(\mathcal{H}^{\otimes N}) \equiv \{\hat{\rho} \in \mathcal{L}(\mathcal{H}^{\otimes N}) \mid \hat{\rho}^\dagger = \hat{\rho} \geq 0, \text{Tr}[\hat{\rho}] = 1\}$ the space of density operators. A quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes N})$ maps a quantum state $\hat{\rho} \in \mathcal{D}(\mathcal{H}^{\otimes N})$ to another one in $\mathcal{D}(\mathcal{H}^{\otimes N})$ via a completely positive and trace-preserving (CPTP) map [43]. Gaussian channels map a Gaussian states (see Eq.(A5) for the definition) to another Gaussian state, and can be simulated by $\mathcal{N}(\hat{\rho}) = \text{Tr}_E[\hat{U}_G(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}_G^\dagger]$, where \hat{U}_G is a Gaussian unitary operation on the system plus environmental modes, $\hat{\rho}_E$ is a Gaussian state and Tr_E is partial trace with respect to the environment. Let $\hat{\mathbf{X}}^T = (\hat{\mathbf{x}}^T, \hat{\mathbf{y}}^T)$ be a collection of quadrature operators of the system mode $\hat{\mathbf{x}}$ and the environmental mode $\hat{\mathbf{y}}$ (see the text below Eq.(A1) for the definition of quadratures), and assume that the initial system and environmental states are given by $\hat{\rho}_G(\bar{\mathbf{x}}, \mathbf{V}_x)$ and $\hat{\rho}_G(\bar{\mathbf{y}}, \mathbf{V}_y)$, respectively. Here, $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ are the first moments and $\mathbf{V}_x, \mathbf{V}_y$ are the second moments of the system and environment—cf, Eq.(A5). If the Gaussian unitary on the joint system is characterized by

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{xx} & \mathbf{S}_{xy} \\ \mathbf{S}_{yx} & \mathbf{S}_{yy} \end{pmatrix} \text{ and } \mathbf{D} = \begin{pmatrix} \mathbf{d}_x \\ \mathbf{d}_y \end{pmatrix}, \quad (1)$$

the first two moments of the system mode are transformed as $\bar{\mathbf{x}} \rightarrow \mathbf{S}_{xx}\bar{\mathbf{x}} + \mathbf{S}_{xy}\bar{\mathbf{y}} + \mathbf{d}_x$ and $\mathbf{V}_x \rightarrow \mathbf{S}_{xx}\mathbf{V}_x\mathbf{S}_{xx}^T + \mathbf{S}_{xy}\mathbf{V}_y\mathbf{S}_{xy}^T$, as can be derived by specializing Eq.(A10) to Eq.(1). After tracing out the environment, the effective Gaussian channel for the system is characterized by

$$\bar{\mathbf{x}} \rightarrow \mathbf{T}\bar{\mathbf{x}} + \mathbf{d}, \quad \mathbf{V}_x \rightarrow \mathbf{T}\mathbf{V}_x\mathbf{T}^T + \mathbf{N}, \quad (2)$$

where $\mathbf{T} = \mathbf{S}_{xx}$, $\mathbf{N} = \mathbf{S}_{xy}\mathbf{V}_y\mathbf{S}_{xy}^T$ and $\mathbf{d} = \mathbf{S}_{xy}\bar{\mathbf{y}} + \mathbf{d}_x$.

The main subject of this paper is the Gaussian channel which models excitation loss, e.g., in optical communication channels.

Definition 1 (Gaussian loss channel). Let $\hat{B}(\eta)$ be the beam splitter unitary with transmissivity $\eta \in [0, 1]$, acting on mode 1 and 2. Then, a Gaussian loss channel is defined as

$$\mathcal{N}[\eta, \bar{n}_{th}](\hat{\rho}_1) \equiv \text{Tr}_2[\hat{B}(\eta)(\hat{\rho}_1 \otimes \hat{\rho}_{\bar{n}_{th}})\hat{B}^\dagger(\eta)]. \quad (3)$$

Here, Tr_2 is the partial trace with respect to the mode 2, ini-

tially in the thermal state $\hat{\rho}_{\bar{n}_{th}}$ with average photon number \bar{n}_{th} .

Note that $\mathcal{N}[\eta, \bar{n}_{th}]$ is characterized by $\mathbf{T} = \sqrt{\eta}\mathbf{I}_2$, $\mathbf{N} = (1 - \eta)(\bar{n}_{th} + \frac{1}{2})\mathbf{I}_2$ and $\mathbf{d} = 0$, where \mathbf{I}_n is the $n \times n$ identity matrix, as can be derived from the definition of beam splitter unitary (see Eq.(A12)). Bosonic pure-loss channel is a special case of Gaussian loss channel with $\bar{n}_{th} = 0$. Bosonic pure-loss channel with transmissivity $\eta \in [\frac{1}{2}, 1]$ ($\eta \in [0, \frac{1}{2}]$) is degradable (anti-degradable) [44] and its quantum capacity is known. (See section III for the definition of degradability and anti-degradability.) Except for this special case, Gaussian loss channels are not degradable nor anti-degradable [45, 46], and thus their coherent information may not be additive.

By replacing beam splitter unitary in Definition 1 with two-mode squeezing unitary, we obtain the Gaussian amplification channel. Quantum-limited amplification channel is a special case of Gaussian amplification and is defined as follows:

Definition 2 (Quantum-limited amplification channel). Let $\hat{S}_2(G)$ be the two-mode squeezing unitary operation with gain $G \geq 1$, acting on mode 1 and 2. Then, the quantum-limited amplification channel is defined as

$$\mathcal{A}[G](\hat{\rho}_1) \equiv \text{Tr}_2[\hat{S}_2(G)(\hat{\rho}_1 \otimes |0\rangle\langle 0|_2)\hat{S}_2^\dagger(G)], \quad (4)$$

where $|0\rangle$ is the vacuum state.

$\mathcal{A}[G]$ is characterized by $\mathbf{T} = \sqrt{G}\mathbf{I}_2$, $\mathbf{N} = \frac{(G-1)}{2}\mathbf{I}_2$ and $\mathbf{d} = 0$, as can be derived from the definition of two-mode squeezing (see Eq.(A13)). Note that the noise $\mathbf{N} = \frac{(G-1)}{2}\mathbf{I}_2$ is due to the variance of the ancillary vacuum state, transferred to the system via two-mode squeezing. Since the vacuum has the minimum variance allowed by Heisenberg uncertainty principle, quantum-limited amplification incurs the least noise among all linear amplification channels [47].

Finally, we introduce the Gaussian random displacement channel, which GKP codes [28] were originally designed to protect against.

Definition 3 (Gaussian random displacement channel). Gaussian random displacement channel is defined as

$$\mathcal{N}_{B_2}[\sigma^2](\hat{\rho}) \equiv \frac{1}{\pi\sigma^2} \int d^2\alpha e^{-\frac{|\alpha|^2}{\sigma^2}} \hat{D}(\alpha)\hat{\rho}\hat{D}^\dagger(\alpha), \quad (5)$$

where $\hat{D}(\alpha)$ is the displacement operator and σ^2 is the variance of random displacement.

We note that the Gaussian random displacement channel belongs to the class B_2 channel [45, 46] and is characterized by $\mathbf{T} = \mathbf{I}_2$, $\mathbf{N} = \sigma^2\mathbf{I}_2$ and $\mathbf{d} = 0$. (Since convention for the Gaussian random displacement channel varies in literature, we prove in Appendix B that $\mathbf{N} = \sigma^2\mathbf{I}_2$ holds in our convention, which is aligned with [28, 41, 42]).

B. Synthesis and decomposition of Gaussian channels

In Ref [42], it has been shown that the GKP codes outperform many other bosonic codes in protecting encoded quan-

tum information against bosonic pure-loss errors upon optimal decoding operation numerically obtained by semidefinite programming. This excellent performance of the GKP codes can be achieved by a sub-optimal decoding operation, which is designed based on the observation that the bosonic pure-loss channel can be converted into the Gaussian displacement channel by a quantum-limited amplification:

$$\mathcal{A}[1/\eta] \cdot \mathcal{N}[\eta, 0] = \mathcal{N}_{B_2}[\sigma_{\eta, 0}^2], \quad (6)$$

where $\sigma_{\eta, 0}^2 \equiv \frac{1-\eta}{\eta}$ (see Eq.(7.21) in [42]). This simple fact was previously noted in [10] (but without proof and explicit relation between η and σ^2) and has been used in a key distribution scheme with the GKP codes [48]. Eq. (6) implies that we can convert the loss channel into the displacement channel and then use the conventional GKP decoding (see the text below Eq.(47)) to decode the encoded GKP states corrupted by bosonic pure-loss channel. Here, we generalize this result and show that general Gaussian loss channels (with added thermal noise) can also be converted into Gaussian random displacement channel. Before doing so, we state and prove the following known fact:

Theorem 4 (Gaussian channel synthesis; Eq.(26) of [45]). *Let \mathcal{N}_1 and \mathcal{N}_2 be Gaussian channels with specification $(\mathbf{T}_1, \mathbf{N}_1, \mathbf{d}_1)$ and $(\mathbf{T}_2, \mathbf{N}_2, \mathbf{d}_2)$, respectively. Then, the synthesized channel $\mathcal{N} \equiv \mathcal{N}_2 \cdot \mathcal{N}_1$ is a Gaussian channel with specification*

$$\begin{aligned} \mathbf{T} &= \mathbf{T}_2 \mathbf{T}_1, \\ \mathbf{N} &= \mathbf{T}_2 \mathbf{N}_1 \mathbf{T}_2^T + \mathbf{N}_2, \\ \mathbf{d} &= \mathbf{T}_2 \mathbf{d}_1 + \mathbf{d}_2. \end{aligned} \quad (7)$$

Proof. Let $\hat{\rho}$ be a Gaussian state, i.e., $\hat{\rho} = \hat{\rho}_G(\bar{\mathbf{x}}, \mathbf{V})$. Upon \mathcal{N}_1 , the first two moments of $\hat{\rho}$ are transformed into $\bar{\mathbf{x}}' = \mathbf{T}_1 \bar{\mathbf{x}} + \mathbf{d}_1$ and $\mathbf{V}' = \mathbf{T}_1 \mathbf{V} \mathbf{T}_1^T + \mathbf{N}_1$. The second channel \mathcal{N}_2 then transforms $\bar{\mathbf{x}}'$ and \mathbf{V}' into

$$\begin{aligned} \bar{\mathbf{x}}'' &= \mathbf{T}_2 \bar{\mathbf{x}}' + \mathbf{d}_2, \\ &= \mathbf{T}_2 \mathbf{T}_1 \bar{\mathbf{x}} + \mathbf{T}_2 \mathbf{d}_1 + \mathbf{d}_2, \\ \mathbf{V}'' &= \mathbf{T}_2 \mathbf{V}' \mathbf{T}_2^T + \mathbf{N}_2 \\ &= \mathbf{T}_2 \mathbf{T}_1 \mathbf{V} \mathbf{T}_1^T \mathbf{T}_2^T + \mathbf{T}_2 \mathbf{N}_1 \mathbf{T}_2^T + \mathbf{N}_2. \end{aligned} \quad (8)$$

The synthesized channel $\mathcal{N} = \mathcal{N}_2 \cdot \mathcal{N}_1$ is thus a Gaussian channel mapping $\hat{\rho}_G(\bar{\mathbf{x}}, \mathbf{V})$ to $\hat{\rho}_G(\bar{\mathbf{x}}'', \mathbf{V}'')$, with the specification as stated in the theorem. ■

Theorem 4 then leads to the following generalization of Eq.(6).

Corollary 5 (Loss + amplification = displacement). *Let $\mathcal{N}[\eta, \bar{n}_{th}]$ be the general Gaussian loss channel with transmissivity $\eta \in [0, 1]$ and thermal photon \bar{n}_{th} in the ancillary mode. Let $\mathcal{A}[1/\eta]$ be the quantum-limited amplification with gain $G = 1/\eta$. Then,*

$$\mathcal{A}[1/\eta] \cdot \mathcal{N}[\eta, \bar{n}_{th}] = \mathcal{N}_{B_2}[\sigma_{\eta, \bar{n}_{th}}^2], \quad (9)$$

where

$$\sigma_{\eta, \bar{n}_{th}}^2 \equiv \left(\frac{1-\eta}{\eta} \right) (\bar{n}_{th} + 1). \quad (10)$$

Proof. Note that $\mathcal{N}[\eta, \bar{n}_{th}]$ and $\mathcal{A}(1/\eta)$ are characterized by $(\mathbf{T}_1, \mathbf{N}_1, \mathbf{d}_1) = (\sqrt{\eta} \mathbf{I}_2, (1 - \eta)(\bar{n}_{th} + \frac{1}{2}) \mathbf{I}_2, \mathbf{0})$ and $(\mathbf{T}_2, \mathbf{N}_2, \mathbf{d}_2) = (\sqrt{\frac{1}{\eta}} \mathbf{I}_2, \frac{1}{2}(\frac{1-\eta}{\eta}) \mathbf{I}_2, \mathbf{0})$, respectively. Let $(\mathbf{T}, \mathbf{N}, \mathbf{d})$ be the specification of the synthesized channel. Theorem 4 implies $\mathbf{T} = \mathbf{T}_2 \mathbf{T}_1 = \mathbf{I}_2$, $\mathbf{d} = \mathbf{T}_2 \mathbf{d}_1 + \mathbf{d}_2 = \mathbf{0}$ and

$$\begin{aligned} \mathbf{N} &= \mathbf{T}_2 \mathbf{N}_1 \mathbf{T}_2^T + \mathbf{N}_2 \\ &= \frac{1}{\eta} (1 - \eta) \left(\bar{n}_{th} + \frac{1}{2} \right) \mathbf{I}_2 + \frac{1}{2} \left(\frac{1 - \eta}{\eta} \right) \mathbf{I}_2 \\ &= \left(\frac{1 - \eta}{\eta} \right) (\bar{n}_{th} + 1) \mathbf{I}_2. \end{aligned} \quad (11)$$

Comparing Eq.(11) with $\mathbf{N} = \sigma^2 \mathbf{I}_2$ for the Gaussian random displacement channel, the statement follows. ■

It is noteworthy that the noise from the Gaussian loss channel (i.e., $\mathbf{N}_1 = (1 - \eta)(\bar{n}_{th} + \frac{1}{2}) \mathbf{I}_2$) is amplified by the quantum-limited amplification (first term in the second line of Eq.(11)), hence increasing the variance $\sigma_{\eta, \bar{n}_{th}}^2$ of resulting random displacement channel. We can however avoid the noise amplification simply by reversing the order of loss channel and amplification, i.e., amplifying the signal prior to sending it through the Gaussian loss channel.

Lemma 6 (Pre-amplification causes less noise). *Let $\mathcal{N}[\eta, \bar{n}_{th}]$ and $\mathcal{A}[1/\eta]$ be as specified in corollary 5. Then,*

$$\mathcal{N}[\eta, \bar{n}_{th}] \cdot \mathcal{A}[1/\eta] = \mathcal{N}_{B_2}[\tilde{\sigma}_{\eta, \bar{n}_{th}}^2], \quad (12)$$

where

$$\tilde{\sigma}_{\eta, \bar{n}_{th}}^2 \equiv (1 - \eta)(\bar{n}_{th} + 1). \quad (13)$$

Note that $\tilde{\sigma}_{\eta, \bar{n}_{th}}^2$ is strictly less than $\sigma_{\eta, \bar{n}_{th}}^2$ for all $0 \leq \eta < 1$.

Proof. The proof goes similarly as in corollary 5, except that the subscripts are exchanged $1 \leftrightarrow 2$:

$$\begin{aligned} \mathbf{N} &= \mathbf{T}_1 \mathbf{N}_2 \mathbf{T}_1^T + \mathbf{N}_1 \\ &= \eta \times \frac{1}{2} \left(\frac{1 - \eta}{\eta} \right) \mathbf{I}_2 + (1 - \eta) \left(\bar{n}_{th} + \frac{1}{2} \right) \mathbf{I}_2 \\ &= (1 - \eta)(\bar{n}_{th} + 1) \mathbf{I}_2. \end{aligned} \quad (14)$$

Note that in Eq.(13), noise from the Gaussian loss channel (i.e., \mathbf{N}_1) is not amplified. Moreover, additional noise \mathbf{N}_2 from the quantum-limited amplification is now reduced by the factor of η as compared with Eq.(10), due to the loss (see the first term in the second line of Eq.(14)).

In subsection IV D, we combine Lemma 6 with the earlier result given in [41] to establish the achievable quantum com-

munication rate of the GKP codes for the general Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$.

Another interesting application of the idea of Gaussian channel synthesis is the composition of bosonic pure-loss channel and quantum-limited amplification with unmatched transmissivity and gain (i.e., $G \neq 1/\eta$). With this mismatch, one can simulate the general Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ with $\mathcal{N}[\eta', 0]$ and $\mathcal{A}[G']$ for some properly chosen η', G' [45, 49]:

Lemma 7 (Eq.(5.1) of [18]). *Gaussian loss channel can be decomposed into a bosonic pure-loss channel followed by a quantum-limited amplification*

$$\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{A}[G'] \cdot \mathcal{N}[\eta', 0], \quad (15)$$

where $\eta' = \frac{\eta}{(1-\eta)\bar{n}_{\text{th}}+1}$ and $G' = (1-\eta)\bar{n}_{\text{th}}+1$.

In [18], Lemma 7 was combined with the data-processing argument to upper bound the quantum capacity of Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$. In subsection III B, we give a tighter upper bound by introducing a slight modification of this approach, i.e., reversing the order of bosonic pure-loss channel and amplification (see Theorem 10 and Theorem 11).

Lemma 8. *Reverse the order of bosonic pure-loss channel and quantum-limited amplification in Lemma 7. Then,*

$$\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{N}[\tilde{\eta}', 0] \cdot \mathcal{A}[\tilde{G}'], \quad (16)$$

where $\tilde{\eta}' = \eta - (1-\eta)\bar{n}_{\text{th}}$ and $\tilde{G}' = \frac{\eta}{\eta-(1-\eta)\bar{n}_{\text{th}}}$.

Proof. Recall that $\mathcal{A}[\tilde{G}']$ and $\mathcal{N}[\tilde{\eta}', 0]$ are characterized by $(\mathbf{T}_1, \mathbf{N}_1, \mathbf{d}_1) = (\sqrt{\tilde{G}'} \mathbf{I}_2, \frac{(\tilde{G}'-1)}{2} \mathbf{I}_2, \mathbf{0})$ and $(\mathbf{T}_2, \mathbf{N}_2, \mathbf{d}_2) = (\sqrt{\tilde{\eta}'} \mathbf{I}_2, \frac{1}{2}(1-\tilde{\eta}') \mathbf{I}_2, \mathbf{0})$, respectively. Then, the synthesized channel is a Gaussian channel characterized by $\mathbf{T} = \mathbf{T}_2 \mathbf{T}_1 = \sqrt{\tilde{\eta}' \tilde{G}'} \mathbf{I}_2 = \sqrt{\eta} \mathbf{I}_2$, $\mathbf{d} = \mathbf{T}_2 \mathbf{d}_1 + \mathbf{d}_2 = \mathbf{0}$ and

$$\begin{aligned} \mathbf{N} &= \mathbf{T}_2 \mathbf{N}_1 \mathbf{T}_2^T + \mathbf{N}_2 \\ &= \tilde{\eta}' \frac{(\tilde{G}'-1)}{2} \mathbf{I}_2 + \frac{1}{2}(1-\tilde{\eta}') \mathbf{I}_2 \\ &= \frac{1}{2}(1+\eta) \mathbf{I}_2 - \tilde{\eta}' \mathbf{I}_2 = (1-\eta) \left(\bar{n}_{\text{th}} + \frac{1}{2} \right) \mathbf{I}_2, \end{aligned} \quad (17)$$

i.e., identical specification to that of $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$. \blacksquare

We note that Lemma 8 was independently discovered and stated in Theorem 30 of [18] and in Lemma 1 of [19].

III. QUANTUM CAPACITY OF GAUSSIAN CHANNELS

A. Earlier results on Gaussian quantum channel capacity

Let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ be a noisy quantum channel from an information sender to a receiver, dilated by $\mathcal{N}(\hat{\rho}) \equiv \text{Tr}_E[\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}^\dagger]$, where \hat{U} is a unitary operator acting on $\mathcal{H} \otimes \mathcal{H}_E$, and \mathcal{H}_E is the Hilbert space

of the environment, causing the channel noise. Complementary channel $\mathcal{N}^c : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}_E)$ is the channel from the sender to the environment, i.e., $\mathcal{N}^c(\hat{\rho}) \equiv \text{Tr}_S[\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}^\dagger]$. Coherent information of a channel is defined as $Q(\mathcal{N}) \equiv \max_{\hat{\rho} \in \mathcal{D}(\mathcal{H})} I_c(\hat{\rho}, \mathcal{N})$, where $I_c(\hat{\rho}, \mathcal{N}) \equiv H(\mathcal{N}(\hat{\rho})) - H(\mathcal{N}^c(\hat{\rho}))$ and $H(\hat{\rho}) \equiv -\text{Tr}[\hat{\rho} \log \hat{\rho}]$ is the entropy of a state $\hat{\rho}$, where \log is the logarithm with base 2. Quantum capacity of a channel \mathcal{N} equals to the regularized coherent information of the channel: $Q_{\text{reg}}(\mathcal{N}) \equiv \lim_{n \rightarrow \infty} \frac{1}{n} Q(\mathcal{N}^{\otimes n})$ [3–5]. Coherent information might be superadditive, i.e., $Q(\mathcal{N} \otimes \mathcal{N}') \geq Q(\mathcal{N}) + Q(\mathcal{N}')$, and thus the one-shot expression $Q(\mathcal{N})$ only lower bounds the true quantum capacity: $Q(\mathcal{N}) \leq Q_{\text{reg}}(\mathcal{N})$.

A quantum channel \mathcal{N} is degradable iff there exists a degrading channel $\mathcal{D} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}_E)$ such that $\mathcal{N}^c = \mathcal{D} \cdot \mathcal{N}$, i.e., iff the receiver can simulate complementary channel. Coherent information of degradable channels are additive (see, e.g., Theorem 12.5.4 in [1]), and thus the one-shot coherent information fully characterizes their quantum capacity: $Q_{\text{reg}}(\mathcal{N}) = Q(\mathcal{N})$ if \mathcal{N} is degradable. A quantum channel \mathcal{N} is anti-degradable iff there exists a degrading channel $\mathcal{D} : \mathcal{L}(\mathcal{H}_E) \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathcal{N} = \mathcal{D} \cdot \mathcal{N}^c$, i.e., iff the environment can simulate the channel from sender to receiver. Anti-degradable channels have zero quantum capacity and thus are unable to support reliable information transmission.

Complementary channel of the bosonic pure-loss channel with transmissivity η is also a bosonic pure-loss channel: $\mathcal{N}^c[\eta, 0] = \mathcal{N}[1-\eta, 0]$. Since bosonic pure-loss channels are multiplicative $\mathcal{N}[\eta_1, 0] \cdot \mathcal{N}[\eta_2, 0] = \mathcal{N}[\eta_1 \eta_2, 0]$, as can be justified by applying Theorem 4, they are degradable if $\eta \in [\frac{1}{2}, 1]$, i.e., $\mathcal{N}^c[\eta, 0] = \mathcal{N}[1-\eta, 0] = \mathcal{N}[\frac{1-\eta}{\eta}, 0] \cdot \mathcal{N}[\eta, 0]$ (similarly, anti-degradable if $\eta \in [0, \frac{1}{2}]$).

Evaluation of the coherent information of a channel involves optimization over all input state $\hat{\rho}$. For degradable Gaussian channels, it was proven that optimization over all Gaussian states is sufficient [14]. For Gaussian loss channels, optimization of one-shot coherent information over thermal states was performed in [13]. As was pointed out in Remark 17 of [16], however, sufficiency of thermal optimizer is yet to be justified. Here, we justify this for bosonic pure-loss channels by using rotational invariance of bosonic pure-loss channels and concavity of coherent information for degradable channels.

Define the Gaussian unitary rotation channel $\mathcal{U}[\theta]$ as $\mathcal{U}[\theta](\hat{\rho}) \equiv \hat{U}(\theta)\hat{\rho}\hat{U}^\dagger(\theta)$, where $\hat{U}(\theta) = e^{i\theta\hat{n}}$ is the phase rotation operator. Bosonic pure-loss channels are invariant under the phase rotation: $\mathcal{U}[\theta]\mathcal{N}[\eta, 0] = \mathcal{N}[\eta, 0]\mathcal{U}[\theta]$, as can be verified by specializing Theorem 4 to $\mathcal{N}[\eta, 0]$ and $\mathcal{U}[\theta]$, where the specification of the latter channel is given by $(\mathbf{T}, \mathbf{N}, \mathbf{d}) = (\mathbf{R}(\theta), \mathbf{0}, \mathbf{0})$ and $\mathbf{R}(\theta)$ is given in Eq.(A11). We then prove the following theorem.

Theorem 9 (Sufficiency of thermal optimizer for bosonic pure-loss channel capacity). *Let $\mathcal{N}[\eta, 0]$ be a bosonic pure-loss channel with transmissivity $\eta \in [0, 1]$ (see Definition 1) and $\hat{\rho} = \sum_{m,n=0}^{\infty} \rho_{mn} |m\rangle\langle n|$ be an arbitrary bosonic state*

represented in the Fock basis. Then,

$$I_c(\hat{\rho}, \mathcal{N}[\eta, 0]) \leq I_c\left(\sum_{n=0}^{\infty} \rho_{nn}|n\rangle\langle n|, \mathcal{N}[\eta, 0]\right). \quad (18)$$

Diagonal states in the Fock basis are thus sufficient for the maximization of coherent information of bosonic pure-loss channels. Combining this with the sufficiency of Gaussian input states [14], it follows that coherent information of bosonic pure-loss channel is maximized by thermal states.

Proof. Define $\hat{\rho}_\theta \equiv \mathcal{U}[\theta]\hat{\rho} = e^{i\theta\hat{n}}\hat{\rho}e^{-i\theta\hat{n}}$ and let $p(\theta)$ be a probability density defined over $\theta \in [0, 2\pi]$. Since bosonic pure-loss channel is degradable, its coherent information $I_c(\hat{\rho}, \mathcal{N}[\eta, 0])$ is concave in the input state (see Theorem 12.5.6 in [1]):

$$\int_0^{2\pi} d\theta p(\theta) I_c(\hat{\rho}_\theta, \mathcal{N}[\eta, 0]) \leq I_c\left(\int_0^{2\pi} d\theta p(\theta) \hat{\rho}_\theta, \mathcal{N}[\eta, 0]\right), \quad (19)$$

Rotational invariance of Gaussian loss channel implies $\mathcal{N}[\eta, 0](\hat{\rho}_\theta) = \mathcal{U}[\theta]\mathcal{N}[\eta, 0](\hat{\rho})$ and similarly $\mathcal{N}^c[\eta, 0](\hat{\rho}_\theta) = \mathcal{U}[\theta]\mathcal{N}^c[\eta, 0](\hat{\rho})$. Since the quantum entropy is invariant under unitary operation, we have $I_c(\hat{\rho}_\theta, \mathcal{N}[\eta, 0]) = I_c(\hat{\rho}, \mathcal{N}[\eta, 0])$. The left hand side of Eq.(19) is then given by $I_c(\hat{\rho}, \mathcal{N}[\eta, 0])$ since $\int_0^{2\pi} p(\theta) = 1$. Choosing $p(\theta)$ to be a flat distribution $p(\theta) = 1/(2\pi)$, we find

$$\begin{aligned} \int_0^{2\pi} d\theta p(\theta) \hat{\rho}_\theta &= \sum_{m,n=0}^{\infty} \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{i\theta(m-n)} \rho_{mn}|m\rangle\langle n| \\ &= \sum_{n=0}^{\infty} \rho_{nn}|n\rangle\langle n|, \end{aligned} \quad (20)$$

where we used $\int_0^{2\pi} d\theta e^{i\theta(m-n)} = 2\pi\delta_{mn}$ to derive the last equality. Plugging Eq.(20) into the right hand side of Eq.(19), Eq.(18) follows.

We emphasize that the phase rotation $\mathcal{U}[\theta]$ does not change the average photon number of a state. Thus, the above argument also applies to an energy-constrained scenario, where average photon number is bounded from above by \bar{n} . ■

Note that the thermal state $\hat{\rho}_{\bar{n}}$ with average photon number \bar{n} is transformed by bosonic pure-loss channel into another thermal state with $\bar{n}' = \eta\bar{n}$: $\mathcal{N}[\eta, 0](\hat{\rho}_{\bar{n}}) = \hat{\rho}_{\eta\bar{n}}$. Similarly, $\mathcal{N}^c[\eta, 0](\hat{\rho}_{\bar{n}}) = \hat{\rho}_{(1-\eta)\bar{n}}$. Since $H(\hat{\rho}_{\bar{n}}) = g(\bar{n})$, where

$$g(x) \equiv (x+1)\log(x+1) - x\log x \quad (21)$$

(see Eq.(A8) and the text below in Appendix A), Theorem 9 leads to the following energy-constrained quantum capacity of bosonic pure-loss channel: (Eq.(1) in [16])

$$\begin{aligned} Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, 0]) &= Q^{n \leq \bar{n}}(\mathcal{N}[\eta, 0]) \\ &= \max[g(\eta\bar{n}) - g((1-\eta)\bar{n}), 0], \end{aligned} \quad (22)$$

where \bar{n} is the maximum allowed photon number per bosonic mode and should not be confused with \bar{n}_{th} in $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$.

When deriving the last equality of Eq.(22), we used the fact that $g(\eta\bar{n}) - g((1-\eta)\bar{n})$ increases monotonically in \bar{n} (see Remark 21 of [18]) and thus the optimal input state (with average photon number less than \bar{n}) which maximizes the coherent information is given by $\hat{\rho} = \hat{\rho}_{\bar{n}}$. In the energy-unconstrained case $\bar{n} \rightarrow \infty$, Eq.(22) reduces to

$$Q_{\text{reg}}(\mathcal{N}[\eta, 0]) = Q(\mathcal{N}[\eta, 0]) = \max\left[\log\left(\frac{\eta}{1-\eta}\right), 0\right]. \quad (23)$$

The general Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ (see Definition 1) with $\bar{n}_{\text{th}} \neq 0$ is not degradable nor anti-degradable [45],[46], and thus its coherent information may not be additive. A lower bound of quantum capacity of Gaussian loss channel can be obtained by evaluating the one-shot coherent information with a thermal input state [13]. In the energy-constrained case ($n \leq \bar{n}$),

$$\begin{aligned} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\text{th}}]) &= g(\eta\bar{n} + (1-\eta)\bar{n}_{\text{th}}) \\ &\quad - g\left(\frac{D + (1-\eta)(\bar{n} - \bar{n}_{\text{th}}) - 1}{2}\right) \\ &\quad - g\left(\frac{D - (1-\eta)(\bar{n} - \bar{n}_{\text{th}}) - 1}{2}\right), \end{aligned} \quad (24)$$

where $D \equiv \sqrt{((1+\eta)\bar{n} + (1-\eta)\bar{n}_{\text{th}} + 1)^2 - 4\eta\bar{n}(\bar{n} + 1)}$. In $\bar{n} \rightarrow \infty$ limit, this reduces to

$$\lim_{\bar{n} \rightarrow \infty} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\text{th}}]) = \log\left(\frac{\eta}{1-\eta}\right) - g(\bar{n}_{\text{th}}). \quad (25)$$

The best known lower bound of the quantum capacity of Gaussian loss channel is thus

$$\begin{aligned} Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) &\geq \max[I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\text{th}}]), 0], \\ Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) &\geq \max\left[\lim_{\bar{n} \rightarrow \infty} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\text{th}}]), 0\right], \end{aligned} \quad (26)$$

in the energy-constrained and unconstrained case, respectively.

We note that Gaussian random displacement channel can be understood as a certain limit of Gaussian loss channel $\mathcal{N}_{B_2}[\sigma^2] = \lim_{\eta \rightarrow 1} \mathcal{N}[\eta, \frac{\sigma^2}{1-\eta} - \frac{1}{2}]$. Thus, lower bound of its quantum capacity is given by

$$Q_{\text{reg}}(\mathcal{N}_{B_2}[\sigma^2]) \geq \max\left[\log\left(\frac{1}{e\sigma^2}\right), 0\right] \quad (27)$$

in the energy-unconstrained case.

General upper bound of quantum capacity was introduced in [13] and applied to Gaussian loss channels (Eq.(5.7) in [13], translated into our notation):

$$\begin{aligned} Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) &\leq Q_{\text{HW}}(\eta, \bar{n}_{\text{th}}) \\ &\equiv \max\left[\log\left(\frac{1+\eta}{(1-\eta)(2\bar{n}_{\text{th}}+1)}\right), 0\right]. \end{aligned} \quad (28)$$

However, this upper bound is not tight: In the case of bosonic pure-loss channel, $Q_{\text{HW}}(\eta, 0) = \log(\frac{1+\eta}{1-\eta}) > Q(\mathcal{N}[\eta, 0])$ for all $\eta \in [0, 1]$, where $Q(\mathcal{N}[\eta, 0])$ is given in Eq.(23). Recently,

three new upper bounds were obtained in [18], where one of them is based on Lemma 7 and data-processing argument, and the other two are based on the notion of approximate degradability of quantum channels developed in [50]. Here, we only present the data-processing bound, in the energy-constrained form, while referring to the original paper for the other two approximate degradability bounds: (Theorems 19 and 24 in [18])

$$Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) \leq Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) \equiv Q^{n \leq \bar{n}}(\mathcal{N}[\eta', 0]), \quad (29)$$

where $Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ is the data-processing bound, $\eta' = \frac{\eta}{(1-\eta)\bar{n}_{\text{th}}+1}$ and $Q^{n \leq \bar{n}}(\mathcal{N}[\eta', 0])$ is given in Eq.(22).

Proof. Since the regularized coherent information is an achievable quantum communication rate, there exists a set of encoding and decoding channels, denoted by $\{\mathcal{E}, \mathcal{D}\}$, which achieves the communication rate $R = Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ for Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$. Since $\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{A}[G'] \cdot \mathcal{N}[\eta', 0]$ (see Lemma 7), this implies that the encoding and decoding set $\{\mathcal{E}, \mathcal{D} \cdot \mathcal{A}[G']\}$ achieves the rate $R = Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ for the bosonic pure-loss channel $\mathcal{N}[\eta', 0]$. Since the achievable rate R is upper bounded by the quantum capacity $Q^{n \leq \bar{n}}(\mathcal{N}[\eta', 0])$, Eq.(29) follows. ■

Note that $Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ converges to $Q^{n \leq \bar{n}}(\mathcal{N}[\eta, 0])$ in $\bar{n}_{\text{th}} \rightarrow 0$ limit, as $\eta' = \eta$ when $\bar{n}_{\text{th}} = 0$. Similarly, an upper bound of the quantum capacity of Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ can be obtained by combining the data-processing argument and Eq.(6):

$$Q_{\text{reg}}(\mathcal{N}_{B_2}[\sigma^2]) \leq \max \left[\log \left(\frac{1}{\sigma^2} \right), 0 \right]. \quad (30)$$

Alternatively, this bound can be derived from Eq.(28) by taking the limit $\eta \rightarrow 1$ and $\bar{n}_{\text{th}} \rightarrow \frac{\sigma^2}{1-\eta} - \frac{1}{2}$.

B. Improved upper bound of Gaussian loss channel capacity

Here, we improve the data-processing bound slightly by using Lemma 8, instead of Lemma 7, for the decomposition of Gaussian loss channel.

Theorem 10 (Improved data-processing bound). *In the energy-unconstrained case, quantum capacity of Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ (see Definition 1) is upper bounded by the improved data-processing bound $Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}})$:*

$$Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) \leq Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}}) \equiv Q_{\text{reg}}(\mathcal{N}[\tilde{\eta}', 0]), \quad (31)$$

where $\tilde{\eta}' = \eta - (1-\eta)\bar{n}_{\text{th}}$ and $Q_{\text{reg}}(\mathcal{N}[\tilde{\eta}', 0])$ is given in Eq.(23). improved data-processing bound $Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}})$ simplifies to

$$Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}}) = \max \left[\log \left(\frac{\eta - (1-\eta)\bar{n}_{\text{th}}}{(1-\eta)(\bar{n}_{\text{th}}+1)} \right), 0 \right]. \quad (32)$$

Proof. The proof goes in the same way as above, except that Lemma 7 is replaced by Lemma 8: $\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{N}[\tilde{\eta}', 0] \cdot \mathcal{A}[\tilde{G}']$ with $\tilde{G}' = \eta / (\eta - (1-\eta)\bar{n}_{\text{th}}) = \eta/\tilde{\eta}'$. Then, the encoding and decoding set $\{\mathcal{A}[\tilde{G}'] \cdot \mathcal{E}, \mathcal{D}\}$ achieves the rate $R = Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ for the bosonic pure-loss channel $\mathcal{N}[\tilde{\eta}', 0]$. Since $R \leq Q(\mathcal{N}[\tilde{\eta}', 0])$, the theorem follows. ■

Note that

$$Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}}) < Q_{\text{DP}}(\eta, \bar{n}_{\text{th}}) \equiv \lim_{\bar{n} \rightarrow \infty} Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) \quad (33)$$

for all $\eta \in [0, 1)$, since

$$\begin{aligned} \tilde{\eta}' &= \frac{\eta}{(1-\eta)\bar{n}_{\text{th}}+1} \\ &= \eta - \frac{\eta(1-\eta)\bar{n}_{\text{th}}}{(1-\eta)\bar{n}_{\text{th}}+1} > \eta - (1-\eta)\bar{n}_{\text{th}} = \eta'. \end{aligned} \quad (34)$$

Thus, $Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}})$ is a strictly tighter upper bound of Gaussian loss channel capacity than $Q_{\text{DP}}(\eta, \bar{n}_{\text{th}})$. We note that Theorem 10 was independently discovered in [19] (see Eqs.(39),(40) therein).

In the energy-constrained case, more care is needed when combining Lemma 8 with the data-processing argument, because the quantum-limited amplification in $\mathcal{E}' = \mathcal{A}[\tilde{G}'] \cdot \mathcal{E}$ increases the energy of the encoded state:

$$\bar{n}' = \text{Tr}[\mathcal{A}[\tilde{G}'](\hat{\rho})\hat{n}] = \tilde{G}'\bar{n} + (\tilde{G}' - 1), \quad (35)$$

as can be derived from the symplectic transformation of the two-mode squeezing (Eq.(A13) and the text above in Appendix A).

Theorem 11 (Improved data-processing bound with energy constraint). *Energy-constrained quantum capacity of Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ is upper bounded by*

$$\begin{aligned} Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) &\leq Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) \\ &\equiv Q^{n \leq \tilde{G}'\bar{n} + (\tilde{G}' - 1)}(\mathcal{N}[\tilde{\eta}', 0]), \end{aligned} \quad (36)$$

where \bar{n} is the maximum allowed average photon number per bosonic mode in the encoded state. $Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ simplifies to

$$\begin{aligned} Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) &= \max \left[g(\eta\bar{n} + (1-\eta)\bar{n}_{\text{th}}) \right. \\ &\quad \left. - g \left(\frac{(1-\eta)(\bar{n}_{\text{th}}+1)(\eta\bar{n} + (1-\eta)\bar{n}_{\text{th}})}{\eta - (1-\eta)\bar{n}_{\text{th}}} \right), 0 \right]. \end{aligned} \quad (37)$$

Proof. Let $\{\mathcal{E}^{n \leq \bar{n}}, \mathcal{D}\}$ be the set of encoding and decoding which achieves the rate $R = Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ for the Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$. Then, the encoding and decoding set $\{\mathcal{A}[\tilde{G}'] \cdot \mathcal{E}^{n \leq \bar{n}}, \mathcal{D}\}$ achieves the rate $R = Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ for the bosonic pure-loss channel $\mathcal{N}[\tilde{\eta}', 0]$. Since the new encoding $\mathcal{E}' \equiv \mathcal{A}[\tilde{G}'] \cdot \mathcal{E}^{n \leq \bar{n}}$ has photon number $n \leq \tilde{G}'\bar{n} + (\tilde{G}' - 1)$, the rate R should be less than $Q^{n \leq \tilde{G}'\bar{n} + (\tilde{G}' - 1)}(\mathcal{N}[\tilde{\eta}', 0])$. ■

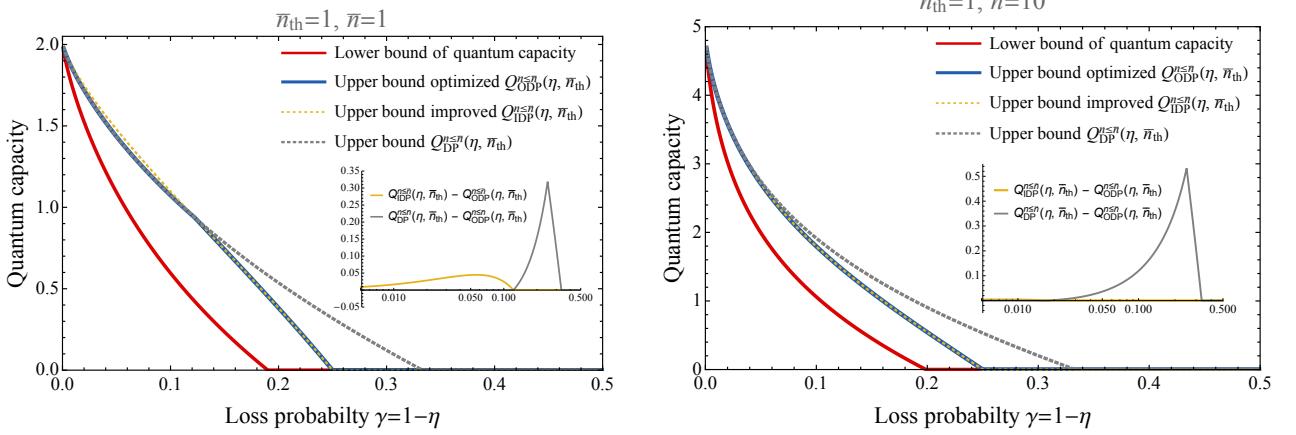


FIG. 1: Bounds of quantum capacity of Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ for $(\bar{n}_{\text{th}}, \bar{n}) = (1, 1)$ (left) and $(\bar{n}_{\text{th}}, \bar{n}) = (1, 10)$ (right). Lower bound (red) is obtained by evaluating the coherent information with a thermal input state (see Eq.(26)). The improved data-processing bound $Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ (dashed yellow) is identical to the optimized data-processing bound $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ (blue) in a wide range of parameter space, and is very close to the optimal one even when it is not optimal. The data-processing bound $Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ (dashed grey) is optimal when $\eta \geq \eta^*(\bar{n}_{\text{th}}, \bar{n})$ for some $\eta^*(\bar{n}_{\text{th}}, \bar{n})$ (e.g., $\eta^*(1, 1) = 0.8775 \dots$).

We note that Theorem 11 was independently discovered in [18] (see Theorem 45 therein). We emphasize that, unlike the energy-unconstrained case, our new upper bound $Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ is not always tighter than $Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ (see the inset of Fig. 1). Physically, this is because the increased encoding energy due to pre-amplification allows larger quantum capacity, which is crucial if the allowed average photon number in the encoding is small, i.e., $\bar{n} \ll 1$. To overcome this issue, we consider a more general decomposition of the Gaussian loss channel, including both pre-amplification and post-amplification:

$$\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{A}[G_1]\mathcal{N}[\bar{\eta}, 0]\mathcal{A}[G_2], \quad (38)$$

where $\bar{\eta} = 1 - \frac{(1-\eta)}{G_1}(\bar{n}_{\text{th}} + 1)$, $G_2 = \eta/(G_1 - (1-\eta)(\bar{n}_{\text{th}} + 1))$ and G_1 can take any value in the range $1 \leq G_1 \leq 1 + (1-\eta)\bar{n}_{\text{th}}$ (the upper bound of G_1 is imposed by $G_2 \geq 1$). In this more general setting, the upper bound of the quantum capacity of $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ is given by $Q^{n \leq G_2\bar{n} + (G_2 - 1)}(\mathcal{N}[\bar{\eta}, 0])$. This upper bound can then be optimized by choosing $G_1, G_2 \geq 1$ such that the upper bound is minimized, i.e.,

$$Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) \leq \min_{G_1, G_2 \geq 1} Q^{n \leq G_2\bar{n} + (G_2 - 1)}(\mathcal{N}[\bar{\eta}, 0]). \quad (39)$$

Theorem 12 (Optimized data-processing bound with energy-constraint). Energy-constrained quantum capacity of Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ is upper bounded by the optimized data-processing bound $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$:

$$Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) \leq Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) \equiv \min_{1 \leq G_1 \leq 1 + (1-\eta)\bar{n}_{\text{th}}} f_{\eta, \bar{n}_{\text{th}}, \bar{n}}(G_1), \quad (40)$$

where \bar{n} is the maximum allowed average photon number in

the encoding and $f_{\eta, \bar{n}_{\text{th}}, \bar{n}}(G_1)$ is defined as

$$f_{\eta, \bar{n}_{\text{th}}, \bar{n}}(G_1) \equiv \max \left[g(\bar{\eta}(G_2\bar{n} + (G_2 - 1))) - g((1 - \bar{\eta})(G_2\bar{n} + (G_2 - 1))), 0 \right]. \quad (41)$$

and $\bar{\eta} = 1 - \frac{(1-\eta)}{G_1}(\bar{n}_{\text{th}} + 1)$, $G_2 = \eta/(G_1 - (1-\eta)(\bar{n}_{\text{th}} + 1))$.

Proof. See appendix C for the derivation. ■

Since $Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ and $Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ are the two extremes with $G_2 = 1$ and $G_1 = 1$, respectively, they are greater than or equal to the optimized data-processing bound. Thus $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ is the best upper bound among all the data-processing type bounds introduced above (see Fig. 1). We numerically observe, however, that either one of these two extremes is optimal: $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) = Q_{\text{DP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ for $\eta \geq \eta^*(\bar{n}_{\text{th}}, \bar{n})$ and $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) = Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ otherwise for some $\eta^*(\bar{n}_{\text{th}}, \bar{n})$. In $\bar{n} \rightarrow \infty$ limit, $\lim_{\bar{n} \rightarrow \infty} \eta^*(\bar{n}_{\text{th}}, \bar{n}) = 1$ and thus $Q_{\text{ODP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}}) = Q_{\text{IDP}}^{n \leq \bar{n}}(\eta, \bar{n}_{\text{th}})$ for all $\eta \in [0, 1]$, consistent with Eq.(33). We refer to Ref. [18] (e.g., Fig. 6 therein) for a more comprehensive comparison of existing upper bounds, including approximate degradability bounds [50].

Finally, the following upper bound of energy-unconstrained quantum capacity of the Gaussian random displacement channel can be derived from the data-processing argument combined with Lemma 6 (specialized to $\bar{n}_{\text{th}} = 0$):

$$Q_{\text{reg}}(\mathcal{N}_{B_2}[\sigma^2]) \leq \max \left[\log \left(\frac{1 - \sigma^2}{\sigma^2} \right), 0 \right]. \quad (42)$$

Note that this bound is strictly tighter than the one in Eq.(30).

IV. GOTTESMAN-KITAEV-PRESKILL CODES

We now aim to find an explicit encoding and decoding strategy which achieves the quantum capacity of Gaussian loss channels. It is very important to realize that the coherent information of bosonic pure-loss channel being maximized by a Gaussian state (i.e., thermal state; Theorem 9) does not imply that Gaussian encoding and decoding is sufficient to achieve the quantum capacity of bosonic pure-loss channel. In fact, it was proven that entanglement distillation of Gaussian states with Gaussian operation is impossible [51]. Due to the close relation between entanglement distillation and quantum state transmission via the quantum teleportation protocol, this implies that Gaussian encoding and decoding can never achieve a non-vanishing quantum communication rate for Gaussian channels. In other words, quantum error correction of a Gaussian error is impossible if the states and operations are restricted to be Gaussian [52]. Therefore, non-Gaussian resources are necessary in order to achieve reliable quantum information transmission through noisy Gaussian channels.

In this section, we show that the ability to prepare a Gottesman-Kitaev-Preskill (GKP) state, which is a non-Gaussian state, is sufficient to establish non-vanishing quantum communication rate through Gaussian loss channels. In particular, we prove that the GKP codes achieve the quantum capacity of Gaussian loss channels up to at most a constant gap from the upper bound given in Theorem 10, in the energy-unconstrained scenario.

A. One-mode square lattice GKP code

The one-mode square lattice GKP code is the simplest class of the GKP codes [28]. The code space $\mathcal{C}_{\text{sq}}^{[d]} \subset \mathcal{H}$ is defined as $\mathcal{C}_{\text{sq}}^{[d]} = \{|\psi\rangle \mid \hat{S}_{\text{sq},q}^{[d]}|\psi\rangle = |\psi\rangle, \hat{S}_{\text{sq},p}^{[d]}|\psi\rangle = |\psi\rangle\}$, where

$$\begin{aligned}\hat{S}_{\text{sq},q}^{[d]} &\equiv \exp(i\hat{q}\sqrt{2\pi d}) = \hat{D}(i\sqrt{\pi d}), \\ \hat{S}_{\text{sq},p}^{[d]} &\equiv \exp(-i\hat{p}\sqrt{2\pi d}) = \hat{D}(\sqrt{\pi d}),\end{aligned}\quad (43)$$

are the stabilizers and d is the dimension of the code space, i.e., $\dim(\mathcal{C}_{\text{sq}}^{[d]}) = d$. Although simultaneous measurement of position and momentum is impossible (i.e., $[\hat{q}, \hat{p}] = i \neq 0$), they can nevertheless be measured simultaneously in modulo $\sqrt{2\pi/d}$ through the stabilizer measurement, since the two stabilizers commute with each other $\hat{S}_{\text{sq},q}^{[d]}\hat{S}_{\text{sq},p}^{[d]} = \hat{S}_{\text{sq},p}^{[d]}\hat{S}_{\text{sq},q}^{[d]}e^{i2\pi d} = \hat{S}_{\text{sq},p}^{[d]}\hat{S}_{\text{sq},q}^{[d]}$. Thus, one-mode square lattice GKP codes can correct any displacement errors in the square unit cell $|\Delta q|, |\Delta p| < \sqrt{\pi/(2d)}$.

Square lattice GKP code states are explicitly given by

$$|\mu_L^{\text{sq}}\rangle \propto \sum_{n=-\infty}^{\infty} |\hat{q} = (dn + \mu)\sqrt{2\pi/d}\rangle \quad (44)$$

in the computational basis, where $\mu = 0, \dots, d-1$ and $|\hat{q} = q_0\rangle$ is an eigenstate of the position operator \hat{q} localized at $q = q_0$, which is an infinitely squeezed state. Phase rotation by

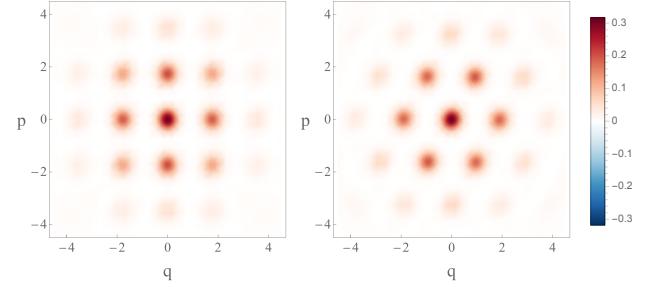


FIG. 2: Wigner function of maximally mixed code states of the one-mode square lattice (left; $\mathcal{C}_{\text{sq}}^{[d]}$) and hexagonal lattice (right; $\mathcal{C}_{\text{hex}}^{[d]}$) GKP code with $d = 2$ and average photon number $\bar{n} = 3$.

angle $\theta = -\pi/2$ implements the basis transformation from $|\mu_L^{\text{sq}}\rangle$ to

$$|\bar{\mu}_L^{\text{sq}}\rangle \propto \sum_{n=-\infty}^{\infty} |\hat{p} = (dn - \mu)\sqrt{2\pi/d}\rangle, \quad (45)$$

where $|\hat{p} = p_0\rangle$ is the eigenstate of the momentum \hat{p} at $p = p_0$. Note that $|\bar{\mu}_L^{\text{sq}}\rangle = \sqrt{1/d} \sum_{\nu=0}^{d-1} \exp(-i2\pi\mu\nu/d)|\nu_L^{\text{sq}}\rangle$.

Since GKP code states are superpositions of infinitely many infinitely squeezed states, their average photon number diverge. It is possible, however, to construct a finite energy GKP states: Replace infinitely squeezed states by finitely squeezed ones, and introduce an overall Gaussian envelop (see Eq.(31) in [28]). Our brief new observation is that this can be concisely realized by $|\mu_L^{\Delta}\rangle \propto \exp(-\Delta^2\hat{n})|\mu_L\rangle$ (which is an equivalent form of Eqs.(40),(41) in [28]). The non-unitary operation $\exp(-\Delta^2\hat{n})$ can be physically implemented by passing the input state through a beam splitter with vacuum ancilla, and then by counting photon number and post-selecting the vacuum click at the out-going idler port. As a result, large enough GKP states of any shape can be carved into smaller ones with Gaussian envelop, where the size of the resulting GKP states (i.e., Δ) can be modulated by transmissivity of the beam splitter.

In Fig. 2, we plot the Wigner function of maximally mixed code state $\hat{\rho}_C = \hat{P}_C/d$ of the square lattice and hexagonal lattice (see subsection IV B) GKP code states for $d = 2$ and $\bar{n} = 3$, where \hat{P}_C is the projection operator to the code space C . In the remainder of this section, we restrict ourselves to the properties of energy-unconstrained ideal GKP states.

Universal gate set of the square lattice GKP code is given as follows: Logical Pauli gates (i.e., Clifford gates of hierarchy 1) are given by $\hat{Z}_L^{\text{sq}} = (\hat{S}_{\text{sq},q}^{[d]})^{\frac{1}{d}} = \exp(i\hat{q}\sqrt{2\pi/d})$ and $\hat{X}_L^{\text{sq}} = (\hat{S}_{\text{sq},p}^{[d]})^{\frac{1}{d}} = \exp(-i\hat{p}\sqrt{2\pi/d})$, i.e., displacement operators, yielding $\hat{Z}_L^{\text{sq}}|\mu_L^{\text{sq}}\rangle = \exp(i2\pi\mu/d)|\mu_L^{\text{sq}}\rangle$, $\hat{X}_L^{\text{sq}}|\mu_L^{\text{sq}}\rangle = |(\mu \oplus 1)_L^{\text{sq}}\rangle$, where \oplus is the summation in modulo d . (The notion of Clifford hierarchy is due to [53].) Any logical Clifford gates of hierarchy 2 can be implemented by Gaussian unitary operations [28]. For example, logical Hadamard gate can be implemented by phase rotation $\exp(-i\pi\hat{a}^\dagger\hat{a}/2)$, as explained

above. Logical CNOT gate can be realized by the SUM gate: $\exp(-i\hat{p}_1\hat{q}_2)|\mu_L^{\text{sq}}\rangle|\nu_L^{\text{sq}}\rangle = |(\mu \oplus \nu)_L^{\text{sq}}\rangle|\nu_L^{\text{sq}}\rangle$. In $d = 2$ case, we observe that $\exp(i\hat{q}^4/(4\pi))$ is the logical T gate (a Clifford gate of hierarchy 3), i.e., $\exp(i\hat{q}^4/(4\pi))|\mu_L^{\text{sq}}\rangle = \exp(i\pi\mu/4)|\mu_L^{\text{sq}}\rangle$, completing the gate set for universal quantum computation.

Encoding of arbitrary logical states can be achieved as follows: Let $|\psi\rangle = \sum_{\mu=0}^{d-1} c_\mu |\mu\rangle$ be an arbitrary state of a physical qudit and assume that one of the GKP logical states (e.g., $|0_L^{\text{sq}}\rangle$) is available. Also, assume the gate $\sum_{\nu=0}^{d-1} (\hat{X}_L^{\text{sq}})^\nu \otimes |\nu\rangle\langle\nu|$, generated by a Hamiltonian of the form $\hat{H} \propto i(\hat{a}_1 - \hat{a}_1^\dagger) \otimes \hat{a}_2^\dagger \hat{a}_2$, can be implemented, where \hat{a}_1 is a bosonic annihilation operator associated with the encoded mode and $\hat{a}_2 \equiv \sum_{\mu=0}^{d-1} \sqrt{\mu} |\mu-1\rangle\langle\mu|$, associated with the physical qudit. Then, we get

$$\left[\sum_{\nu=0}^{d-1} (\hat{X}_L^{\text{sq}})^\nu \otimes |\nu\rangle\langle\nu| \right] (|0_L^{\text{sq}}\rangle \otimes |\psi\rangle) = \sum_{\mu=0}^{d-1} c_\mu |\mu_L^{\text{sq}}\rangle \otimes |\mu\rangle. \quad (46)$$

Upon a unitary Fourier gate on the physical qudit, $|\mu\rangle \rightarrow \sqrt{1/d} \sum_{\nu=0}^{d-1} \exp(-i2\pi\mu\nu/d) |\nu\rangle$, this state becomes $\sum_{\mu,\nu=0}^{d-1} c_\mu \exp(-i2\pi\mu\nu/d) |\mu_L^{\text{sq}}\rangle \otimes |\nu\rangle$. If one then measures the physical qudit in computational basis $\{|0\rangle, \dots, |d-1\rangle\}$ and the measurement outcome is $|\nu\rangle$, the oscillator state is collapsed into $|\psi_L^{\text{sq}}(\nu)\rangle = \sum_{\mu=0}^{d-1} c_\mu \exp(-i2\pi\mu\nu/d) |\mu_L^{\text{sq}}\rangle$. Applying $(\hat{Z}_L^{\text{sq}})^\nu$, we finally obtain the desired encoded logical state

$$(\hat{Z}_L^{\text{sq}})^\nu |\psi_L^{\text{sq}}(\nu)\rangle = \sum_{\mu=0}^{d-1} c_\mu |\mu_L^{\text{sq}}\rangle. \quad (47)$$

If the encoded GKP states are sent through the Gaussian random displacement channel, decoding can be achieved by measuring two stabilizers $\hat{S}_{\text{sq},q}^{[d]}, \hat{S}_{\text{sq},p}^{[d]}$ given in Eq.(43), providing protection against the channel noise as specified in the text below Eq.(43). Given that fresh supply of ancillary GKP state is available, such stabilizer measurements can be achieved by Gaussian operations and homodyne detection: Assume we have an ancillary oscillator mode prepared at $|\bar{0}_L^{\text{sq}}\rangle$ (i.e., $q_2 \equiv 0 \pmod{\sqrt{2\pi/d}}$). Upon the SUM gate $\exp(-i\hat{q}_1\hat{p}_2)$, the ancillary position operator \hat{q}_2 is transformed into $\hat{q}_2 + \hat{q}_1$ while the system's position operator remains at \hat{q}_1 . We can then perform homodyne measurement of \hat{q}_2 , and if the measurement outcome is $q_2^M (= q_1 + q_2)$, we can extract the value of \hat{q}_1 in modulo $\sqrt{2\pi/d}$, i.e., $q_1 \equiv q_2^M \pmod{\sqrt{2\pi/d}}$, hence measuring $\hat{S}_{\text{sq},q}^{[d]}$. The other stabilizer $\hat{S}_{\text{sq},p}^{[d]}$ can also be measured in a similar way using an initial ancillary GKP state $|0_L^{\text{sq}}\rangle$ (i.e., $p_2 \equiv 0 \pmod{\sqrt{2\pi/d}}$), SUM gate with control and target exchanged $\exp(-i\hat{p}_1\hat{q}_2)$ and homodyne measurement of \hat{p}_2 (see the text below Eq.(105) in [28]). From the extracted values of position and momentum in modulo $\sqrt{2\pi/d}$ (i.e., q_2^M, p_2^M), one can infer $\Delta q = q_2^{M*}, \Delta p = p_2^{M*}$, where q_2^{M*}, p_2^{M*} are defined such that $q_2^M \equiv q_2^{M*} \pmod{\sqrt{2\pi/d}}$, $p_2^M \equiv p_2^{M*} \pmod{\sqrt{2\pi/d}}$ and $|q_2^{M*}|, |p_2^{M*}| \leq \sqrt{\pi/(2d)}$.

One can then correct such errors by counter displacement $\exp(i\hat{p}_1 q_2^{M*})$ and $\exp(-i\hat{q}_1 p_2^{M*})$.

In both encoding and decoding, the most non-trivial task is to prepare a GKP state (e.g., $|0_L^{\text{sq}}\rangle$) from an arbitrary non-GKP state. There have been many proposals to achieve such a goal in various experimental platforms [54–60], including the one in the original paper [28]. In particular, the proposal in [58] is based on the idea of phase estimation of unitary operators [61], which allows the following representation of the GKP states (equivalent to Eq.(29) in [28])

$$|\mu_L^{\text{sq}}\rangle \propto (\hat{X}_L^{\text{sq}})^\mu \sum_{n_1, n_2=-\infty}^{\infty} (\hat{S}_{\text{sq},p}^{[d]})^{n_1} (\hat{Z}_L^{\text{sq}})^{n_2} |\phi_0\rangle \\ = \sum_{n_1, n_2=-\infty}^{\infty} e^{-i\hat{p}\sqrt{2\pi/d}(dn_1+\mu)} e^{i\hat{q}\sqrt{2\pi/d}n_2} |\phi_0\rangle. \quad (48)$$

Here, $\sum_{n_1=-\infty}^{\infty} (\hat{S}_p^{[d]})^{n_1}$ and $\sum_{n_2=-\infty}^{\infty} (\hat{Z}_L)^{n_2}$ can be understood as the projection operator associated with the phase estimation (and correction) of $\hat{S}_p^{[d]}$ and \hat{Z}_L , which enforce $\hat{S}_p^{[d]} = \hat{Z}_L = 1$, hence $|0_L^{\text{sq}}\rangle$. The last operation $(\hat{X}_L^{\text{sq}})^\mu$ then transforms $|0_L^{\text{sq}}\rangle$ into $|\mu_L^{\text{sq}}\rangle$. Note that Eq.(48) is valid for any input state $|\phi_0\rangle$ with non-zero overlap with $|0_L^{\text{sq}}\rangle$ which, for example, can be chosen to be the vacuum state. In this case, Eq.(48) can be understood as a superposition of coherent states $|\alpha\rangle$ in a 2-dimensional square lattice $\alpha = \sqrt{\frac{\pi}{d}}(dn_1 + \mu + in_2)$, where n_1, n_2 are integers.

We remark that the scheme in [54] implements Eq.(48) by post-selection, whereas [58] does so deterministically. The former protocol is within the reach of near-term technology of trapped ion systems, and some preliminary experimental progress has been made [62]. In principle, the latter protocol can be realized in circuit quantum electrodynamics systems by, e.g., extending the quantum non-demolition measurement techniques used in [63].

Based on the numerically optimized decoding operation, we earlier showed that the GKP codes offer excellent protection against bosonic pure-loss channel $\mathcal{N}[\eta, 0]$ as well, despite not being specifically designed for such a purpose [42]. In addition, by using sub-optical decoding (i.e., quantum-limited amplification $\mathcal{A}[1/\eta]$ followed by the conventional GKP decoding as described above), we can prove that the logical error probability scales as

$$\epsilon_L^{\text{sq}} \sim \exp \left[-\frac{\pi}{4d} \left(\frac{\eta}{1-\eta} \right) \right], \quad (49)$$

(cf. Eq.(7.24) in [42]) which vanishes non-analytically in the limit of perfect transmission $\eta \rightarrow 1$. In the following subsection, we will explain that it is possible to improve the constant prefactor from $\frac{\pi}{4}$ to $\frac{\pi}{2\sqrt{3}}$ by using hexagonal lattice GKP code instead of square lattice.

B. One-mode hexagonal lattice GKP code

Code space of one-mode hexagonal lattice GKP code is defined by $\mathcal{C}_{\text{hex}}^{[d]} \equiv \{|\psi\rangle \mid \hat{S}_{\text{hex},q}^{[d]}|\psi\rangle = |\psi\rangle, \hat{S}_{\text{hex},p}^{[d]}|\psi\rangle = |\psi\rangle\}$, where the stabilizers are given by

$$\begin{aligned}\hat{S}_{\text{hex},q}^{[d]} &\equiv \exp[i\sqrt{2\pi d}(S_{11}\hat{q} + S_{21}\hat{p})], \\ \hat{S}_{\text{hex},p}^{[d]} &\equiv \exp[-i\sqrt{2\pi d}(S_{12}\hat{q} + S_{22}\hat{p})],\end{aligned}\quad (50)$$

and S_{ij} are matrix elements the symplectic matrix

$$\mathbf{S}_{\text{hex}} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \left(\frac{2}{\sqrt{3}}\right)^{\frac{1}{2}} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}. \quad (51)$$

Note that stabilizers of the square lattice GKP code are Eq.(50) with $\mathbf{S}_{\text{sq}} = \mathbf{I}_2$, where \mathbf{I}_2 is the 2×2 identity matrix. Thus, hexagonal lattice GKP code states can be generated by applying Gaussian unitary operation, with corresponding symplectic transformation $\mathbf{S}_{\text{hex}}^{-1}$, to the square lattice GKP code states. Similarly as in the case of square lattice GKP code, logical Pauli operators are given by $\hat{Z}_L^{\text{hex}} = (\hat{S}_{\text{hex},q}^{[d]})^{\frac{1}{d}}$ and $\hat{X}_L^{\text{hex}} = (\hat{S}_{\text{hex},p}^{[d]})^{\frac{1}{d}}$. Following the same reasoning to obtain Eq.(48), logical states of the hexagonal GKP code are given by

$$\begin{aligned}|\mu_L^{\text{hex}}\rangle &\propto (\hat{X}_L^{\text{hex}})^\mu \sum_{n_1, n_2=-\infty}^{\infty} (\hat{S}_{\text{hex},p}^{[d]})^{n_1} (\hat{Z}_L^{\text{hex}})^{n_2} |\phi_0\rangle \\ &= \sum_{n_1, n_2=-\infty}^{\infty} e^{-i(\frac{1}{2}\hat{q} + \frac{\sqrt{3}}{2}\hat{p})\sqrt{\frac{4\pi}{\sqrt{3}d}}(dn_1 + \mu)} e^{i\hat{q}\sqrt{\frac{4\pi}{\sqrt{3}d}}n_2} |\phi_0\rangle,\end{aligned}\quad (52)$$

in the computational basis, where $|\phi_0\rangle$ is an arbitrary state with non-zero overlap with $|0_L^{\text{hex}}\rangle$. If $|\phi_0\rangle$ is chosen to be the vacuum state, $|\mu_L^{\text{hex}}\rangle$ is a superposition of coherent states $|\alpha\rangle$ in a 2-dimensional hexagonal lattice $\alpha = \sqrt{\frac{2\pi}{\sqrt{3}d}}((\frac{\sqrt{3}}{2} - \frac{i}{2})(dn_1 + \mu) + in_2)$, where n_1, n_2 are integers. For visualization of the one-mode hexagonal lattice GKP code space, see the right panel of Fig. 2.

Note that since \mathbf{S}_{sq} and \mathbf{S}_{hex} are symplectic matrices (see Eq.(A9)), the lattice generated by each of them has a unit cell with unit area: $\det(\mathbf{S}_{\text{sq}}) = \det(\mathbf{S}_{\text{hex}}) = 1$. For the square lattice, the minimum distance between different lattice points is given by $d_{\min}^{\text{sq}} = 1$, and for hexagonal lattice, $d_{\min}^{\text{hex}} = (2/\sqrt{3})^{1/2} > d_{\min}^{\text{sq}}$. Thus, square lattice GKP code can correct displacement errors within the radius $r \equiv \sqrt{|\Delta q|^2 + |\Delta p|^2} \leq \sqrt{\frac{\pi}{2d}}$, whereas hexagonal lattice GKP code does so for $r \leq \sqrt{\frac{\pi}{\sqrt{3}d}}$. This leads to the following logical error probability

$$\epsilon_L^{\text{hex}} \sim \exp\left[-\frac{\pi}{2\sqrt{3}d}\left(\frac{\eta}{1-\eta}\right)\right], \quad (53)$$

for the hexagonal lattice GKP code against bosonic pure-loss channel. Note that the hexagonal lattice allows the densest

sphere packing in 2-dimensional Euclidean space [64].

C. Multi-mode symplectic dual lattice GKP codes

Generalizing Eq.(50), one can consider the following stabilizers for N -mode GKP code space, encoding d logical states per mode:

$$\hat{S}_k^{[d]} \equiv \exp[i\sqrt{2\pi d}(-1)^{k+1}(\hat{\mathbf{x}}^T \mathbf{S})_k], \quad (54)$$

for $k \in \{1, \dots, 2N\}$. Here, $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_{2N})^T = (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_N, \hat{p}_N)^T$ are the quadrature operators of N oscillator modes and \mathbf{S} is a $2N \times 2N$ symplectic matrix (see also [28, 41]). This more general type of GKP code states can be generated by applying Gaussian operation with corresponding symplectic transformation \mathbf{S}^{-1} to the N copies of one-mode square lattice GKP states, i.e., $|\vec{\mu}_L^{\text{sq}}\rangle = \hat{U}_{\mathbf{S}^{-1}}|\vec{\mu}_L^{\text{sq}}\rangle$, where $\vec{\mu} = (\mu^1, \dots, \mu^N)$, $|\vec{\mu}_L^{\text{sq}}\rangle = |\mu_L^1\rangle \otimes \dots \otimes |\mu_L^N\rangle$ and $\mu^k \in \{0, \dots, d-1\}$.

D. Achievable quantum communication rate of the GKP codes

Similar to the one-mode case, it is possible to increase the correctable radius of displacement in N -mode case, by choosing a $2N$ -dimensional symplectic lattice allowing more efficient sphere packing than the square lattice. It is known that there exists a $2N$ -dimensional lattice in Euclidean space allowing $d_{\min} \geq \sqrt{N/(\pi e)}$ [65] and a stronger statement was proven in [66] that the same holds also for symplectic lattices. Choosing such a lattice to define the GKP code, one can correct all displacement error within radius $r \leq \sqrt{N/(2ed)}$. In the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$, the probability of displacement with radius larger than $\sqrt{2N}\sigma$ occurring vanishes in the limit of infinitely many modes $N \rightarrow \infty$. Thus, if $\sqrt{N/(2ed)} \geq \sqrt{2N}\sigma$ is satisfied, i.e.,

$$d \leq d_\sigma \equiv \frac{1}{4e\sigma^2}, \quad (55)$$

encoded information can be transmitted faithfully with vanishing decoding error probability. Then, it follows that the communication rate $R = \log[d_\sigma] = \log\lfloor\frac{1}{4e\sigma^2}\rfloor$ can be achieved for the Gaussian random displacement channel (see Eq.(55) in [41]; floor function is due to the fact that d can only be an integer).

Note that the above estimation is overly conservative since we did not take into account the correctable displacement outside the correctable sphere. With an improved estimation of the decoding error probability, the following statement was ultimately established in [41]:

Lemma 13 (Eq.(66) in [41]). Let $\mathcal{N}_{B_2}[\sigma^2]$ be a Gaussian random displacement channel as defined in Definition 3. Also, let $\mathcal{C}_{\mathbf{S}}^{[d]}$ be the N -mode GKP code space defined by the stabilizers in Eq.(54). Then, there exists a symplectic lattice generated by \mathbf{S} such that the GKP code achieves the following rate for the

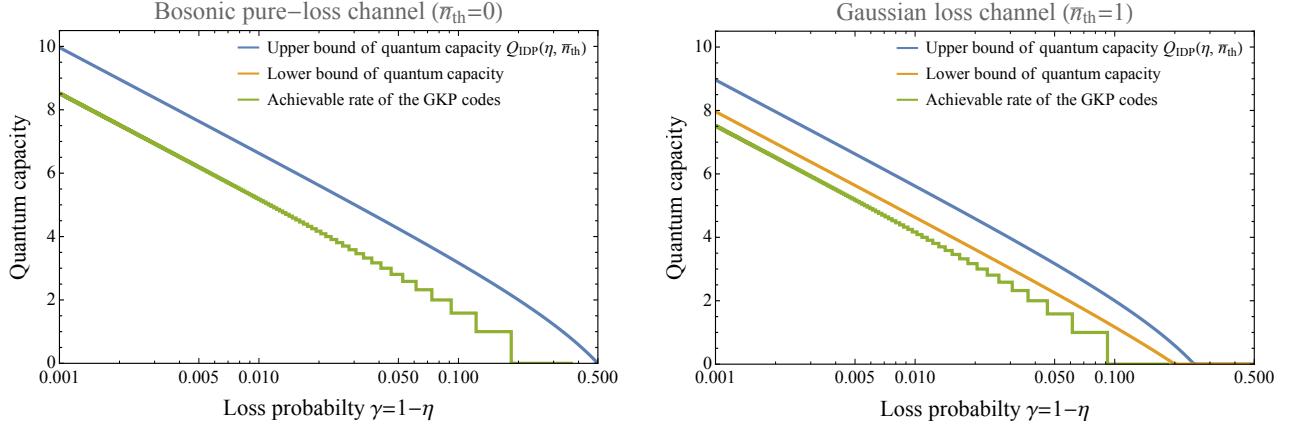


FIG. 3: Achievable quantum communication rate of the GKP codes

(green) compared with lower (orange) and upper (blue) bounds of quantum capacity of bosonic pure-loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}} = 0]$ (left) and Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}} = 1]$ (right). Orange line in the left panel is hidden behind the blue line, since for the bosonic pure-loss channel, lower and upper bounds coincide with each other and the quantum capacity is completely known.

Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ in $N \rightarrow \infty$ limit.

$$R = \max \left(\log \left\lfloor \frac{1}{e\sigma^2} \right\rfloor, 0 \right). \quad (56)$$

Here, $\lfloor x \rfloor$ is the floor function, due to the fact that d can only be an integer.

Thus, the GKP codes achieve one-shot coherent information of the Gaussian random displacement channel (cf. Eq.(27)) in modulo floor function. Based on this, we now establish the following main result:

Theorem 14 (Achievable quantum communication rate of the GKP codes for Gaussian loss channel). *Let $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ be a Gaussian loss channel as defined in Definition 1. Upon optimal choice of $2N \times 2N$ symplectic matrix $\mathbf{S} = \mathbf{S}^*$, the N -mode GKP code as defined in Eq.(54) achieves the following rate for the Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ in $N \rightarrow \infty$ limit:*

$$R = \max \left(\log \left\lfloor \frac{1}{e(1-\eta)(\bar{n}_{\text{th}}+1)} \right\rfloor, 0 \right), \quad (57)$$

where $\lfloor x \rfloor$ is the floor function.

Proof. Lemma 6 states that the Gaussian loss channel $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ can be converted into the Gaussian random displacement channel $\mathcal{N}_{B_2}[\tilde{\sigma}_{\eta, \bar{n}_{\text{th}}}^2]$ by quantum-limited amplification $\mathcal{A}[1/\eta]$, where $\tilde{\sigma}_{\eta, \bar{n}_{\text{th}}}^2 = (1-\eta)(\bar{n}_{\text{th}}+1)$. Combining this with Lemma 13, Eq.(57) follows. ■

Recall Theorem 10 and note that $Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) \leq Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}})$ where

$$Q_{\text{IDP}}(\eta, \bar{n}_{\text{th}}) < \max \left[\log \left(\frac{1}{(1-\eta)(\bar{n}_{\text{th}}+1)} \right), 0 \right], \quad (58)$$

as derived from Eq.(32). Comparing this with Eq.(57), we get $Q_{\text{reg}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}]) - R \lesssim \log e = 1.44269 \dots$, where \sim is due to the floor function. Thus, GKP codes defined over optimal symplectic lattice achieve quantum capacity of Gaussian loss channels up to at most a constant gap from the upper bound of quantum capacity (see Fig. 3 for illustration).

We note that the established rate in Theorem 14 relies on the existence of a symplectic lattice in higher dimensions satisfying a certain desired condition (see Eqs.(56),(57) in [41]). In this regard, we remark that E_8 lattice and Leech lattice Λ_{24} (both symplectic; see appendix of [66]) were recently shown to support the densest sphere packing in 8 and 24 dimensional Euclidean space, respectively [67, 68].

V. BICONVEX ENCODING AND DECODING OPTIMIZATION

In this subsection, we consider the energy-constrained scenario by imposing maximum allowed photon number in each mode. Instead of attempting to establish a theoretically rigorous statement as we did in the energy-unconstrained case, we tackle the problem by numerical optimization. Since decoding error probability cannot be suppressed arbitrarily close to zero with only finite number of modes, we aim to find a set of optimal encoding and decoding maps which maximize the *entanglement fidelity*. The main reason we chose entanglement fidelity (instead of other measures, e.g., diamond norm) as the fidelity measure is to make the optimization problem more tractable, as maximization of entanglement fidelity can be formulated as a biconvex optimization, which can be tackled by alternating semidefinite programming method.

A. Choi matrix and superoperator of a quantum channel

A quantum channel $\mathcal{A} : \mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2)$ is a completely positive trace preserving (CPTP) map [43], which has one-to-one correspondence with the Choi matrix $\hat{X}_{\mathcal{A}} \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Matrix elements of the Choi matrix are defined as $(\hat{X}_{\mathcal{A}})_{[ij],[i'j']} = \langle j_{\mathcal{H}_2} | \mathcal{A}(|i_{\mathcal{H}_1}\rangle\langle i'_{\mathcal{H}_1}|) | j'_{\mathcal{H}_2} \rangle$, where $|i_{\mathcal{H}_1}\rangle, |i'_{\mathcal{H}_1}\rangle$ and $|j_{\mathcal{H}_2}\rangle, |j'_{\mathcal{H}_2}\rangle$ are orthonormal basis of \mathcal{H}_1 and \mathcal{H}_2 , respectively. Choi matrix $\hat{X}_{\mathcal{A}}$ is hermitian positive semidefinite by definition of complete positivity (CP) of \mathcal{A} [43]. Also, trace preserving (TP) condition imposes the affine constraint $\text{Tr}_{\mathcal{H}_2} \hat{X}_{\mathcal{A}} \equiv \sum_{j=0}^{\dim(\mathcal{H}_2)-1} (X_{\mathcal{A}})_{[ij],[i'j]} |i_{\mathcal{H}_1}\rangle\langle i'_{\mathcal{H}_1}| = \hat{I}_{\mathcal{H}_1}$.

Let $\mathcal{B} : \mathcal{L}(\mathcal{H}_2) \rightarrow \mathcal{L}(\mathcal{H}_3)$ be another quantum channel (i.e., a CPTP map) and consider the composite channel $\mathcal{B} \cdot \mathcal{A} : \mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_3)$. To analyze the composite channel, it is convenient to define superoperator $\hat{T}_{\mathcal{A}}$ of a channel \mathcal{A} whose matrix elements are given by $(\hat{T}_{\mathcal{A}})_{jj',ii'} \equiv \langle j_{\mathcal{H}_2} | \mathcal{A}(|i_{\mathcal{H}_1}\rangle\langle i'_{\mathcal{H}_1}|) | j'_{\mathcal{H}_2} \rangle = (\hat{X}_{\mathcal{A}})_{[ij],[i'j']}$ [69]: Superoperator of a composite channel is then given by the matrix multiplication of the superoperators of its constituting channels, i.e., $\hat{T}_{\mathcal{B} \cdot \mathcal{A}} = \hat{T}_{\mathcal{B}} \hat{T}_{\mathcal{A}}$.

B. Entanglement fidelity and Choi matrix

Let \mathcal{H}_n be a Hilbert space of dimension n and $\mathcal{N} : \mathcal{L}(\mathcal{H}_n) \rightarrow \mathcal{L}(\mathcal{H}_n)$ be a CPTP map describing a noisy quantum channel. Consider d -dimensional ($d \leq n$) Hilbert spaces \mathcal{H}_0 and \mathcal{H}'_0 and assume the information sender prepared a maximally entangled state in a local noiseless memory:

$$|\Phi^+\rangle \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i_{\mathcal{H}_0}\rangle |i'_{\mathcal{H}'_0}\rangle. \quad (59)$$

The sender then encodes half of the entangled state in \mathcal{H}_0 to the channel input by a CPTP encoding map $\mathcal{E} : \mathcal{L}(\mathcal{H}_0) \rightarrow \mathcal{L}(\mathcal{H}_n)$, and then send it through the noisy channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_n) \rightarrow \mathcal{L}(\mathcal{H}_n)$. The receiver then maps the received state at the channel output to the local memory by a CPTP decoding map $\mathcal{D} : \mathcal{L}(\mathcal{H}_n) \rightarrow \mathcal{L}(\mathcal{H}_0)$. As a result, both parties obtain an approximate entangled state

$$\hat{\rho} \equiv (\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E} \otimes \text{id}_{\mathcal{H}'_0})(|\Phi^+\rangle\langle\Phi^+|), \quad (60)$$

where $\text{id}_{\mathcal{H}'_0} : \mathcal{L}(\mathcal{H}'_0) \rightarrow \mathcal{L}(\mathcal{H}'_0)$ is the identity map associated with the noiseless local memory at the sender's side. Note that $\hat{\rho}$ in Eq.(60) is explicitly given by

$$\hat{\rho} = \frac{1}{d} \sum_{i,i'=0}^{d-1} (\hat{X}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}})_{[ij],[i'j']} |j_{\mathcal{H}_0}\rangle\langle j'_{\mathcal{H}_0}| \otimes |i_{\mathcal{H}'_0}\rangle\langle i'_{\mathcal{H}'_0}|, \quad (61)$$

where $\hat{X}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}} \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_0)$ is the Choi matrix of the composite channel $\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}$. Thus, $\hat{\rho}$ is isomorphic to the Choi matrix $\hat{X}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}}$.

Quality of the approximate non-local entangled state $\hat{\rho}$ can be measured by the entanglement fidelity $F_e(\hat{\rho}) \equiv \langle \Phi^+ | \hat{\rho} | \Phi^+ \rangle$ [70], where $F_e(\hat{\rho}) = 1$ iff $\hat{\rho}$ is a noiseless maximally entangled state $|\Phi^+\rangle\langle\Phi^+|$. Isomorphism between $\hat{\rho}$ and $\hat{X}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}}$ then yields

$$F_e(\hat{\rho}) = (1/d^2) \sum_{i,i'=0}^{d-1} (\hat{X}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}})_{[ii],[i'i']} = \frac{1}{d^2} \text{Tr}[\hat{T}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}}], \quad (62)$$

where $\hat{T}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}}$ is the superoperator of $\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}$. Note that $\hat{T}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}} = \hat{T}_{\mathcal{D}} \hat{T}_{\mathcal{N}} \hat{T}_{\mathcal{E}}$ can be decomposed as

$$\begin{aligned} & (\hat{T}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}})_{jj',ii'} \\ &= \sum_{k,k',l,l'=0}^{n-1} (\hat{X}_{\mathcal{D}})_{[lj],[l'j']} (\hat{X}_{\mathcal{N}})_{[kl],[k'l']} (\hat{X}_{\mathcal{E}})_{[ik],[i'k']}, \end{aligned} \quad (63)$$

where $\hat{X}_{\mathcal{D}} \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_n)$, $\hat{X}_{\mathcal{N}} \in \mathcal{L}(\mathcal{H}_n \otimes \mathcal{H}_n)$ and $\hat{X}_{\mathcal{E}} \in \mathcal{L}(\mathcal{H}_n \otimes \mathcal{H}_0)$ are the Choi matrices of decoding, noisy channel, and encoding, respectively, and $i, i', j, j' \in \{0, \dots, d-1\}$. We note that the entanglement fidelity $F_e(\hat{\rho})$ is a bi-linear function of $\hat{X}_{\mathcal{E}}$ and $\hat{X}_{\mathcal{E}}$, since $\hat{T}_{\mathcal{D} \cdot \mathcal{N} \cdot \mathcal{E}}$ is bi-linear in $\hat{X}_{\mathcal{E}}$ and $\hat{X}_{\mathcal{E}}$, as can be seen from Eq.(63).

To make the bi-linearity more evident, we define a linear map $f_{\mathcal{N}} : \mathcal{L}(\mathcal{H}_n \otimes \mathcal{H}_0) \rightarrow \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_n)$ such that

$$(f_{\mathcal{N}}(\hat{X}))_{[l'l'],[li]} \equiv \sum_{k,k'=0}^{n-1} (\hat{X}_{\mathcal{N}})_{[kl],[k'l']} (\hat{X})_{[ik],[i'k']}, \quad (64)$$

where $l, l' \in \{0, \dots, n-1\}$. Entanglement fidelity $F_e(\hat{\rho})$ is then given by

$$F_e(\hat{\rho}) = \frac{1}{d^2} \text{Tr}[\hat{X}_{\mathcal{D}} f_{\mathcal{N}}(\hat{X}_{\mathcal{E}})], \quad (65)$$

which is apparently bi-linear in $\hat{X}_{\mathcal{E}}$ and $\hat{X}_{\mathcal{D}}$.

C. Maximization of entanglement fidelity

It is known that finding optimal decoding operation which maximizes the entanglement fidelity is a semidefinite programming, if the encoding map is fixed to $\mathcal{E} = \bar{\mathcal{E}}$ [71, 72]:

$$\begin{aligned} & \max_{\hat{X}_{\mathcal{D}}} \text{Tr}[\hat{X}_{\mathcal{D}}(f_{\mathcal{N}}(\hat{X}_{\mathcal{E}}))] \\ & \text{s.t. } \hat{X}_{\mathcal{D}} = \hat{X}_{\mathcal{D}}^\dagger \succeq 0, \text{ Tr}_{\mathcal{H}_0} \hat{X}_{\mathcal{D}} = \hat{I}_{\mathcal{H}_n}, \end{aligned} \quad (66)$$

where the constraints are due to the CPTP nature of \mathcal{D} . Similarly, optimizing encoding while fixing decoding is also a semidefinite programming, and thus the entire problem is a biconvex optimization [73]. Here, we further impose an energy constraint to the encoding map while still preserving biconvexity of the problem. Let $\hat{E} \in \mathcal{L}(\mathcal{H}_n)$ be an energy observable. We define $\text{Tr}_{\mathcal{H}_n}[\hat{E} \hat{\rho}_{\mathcal{E}}]$ to be the average energy

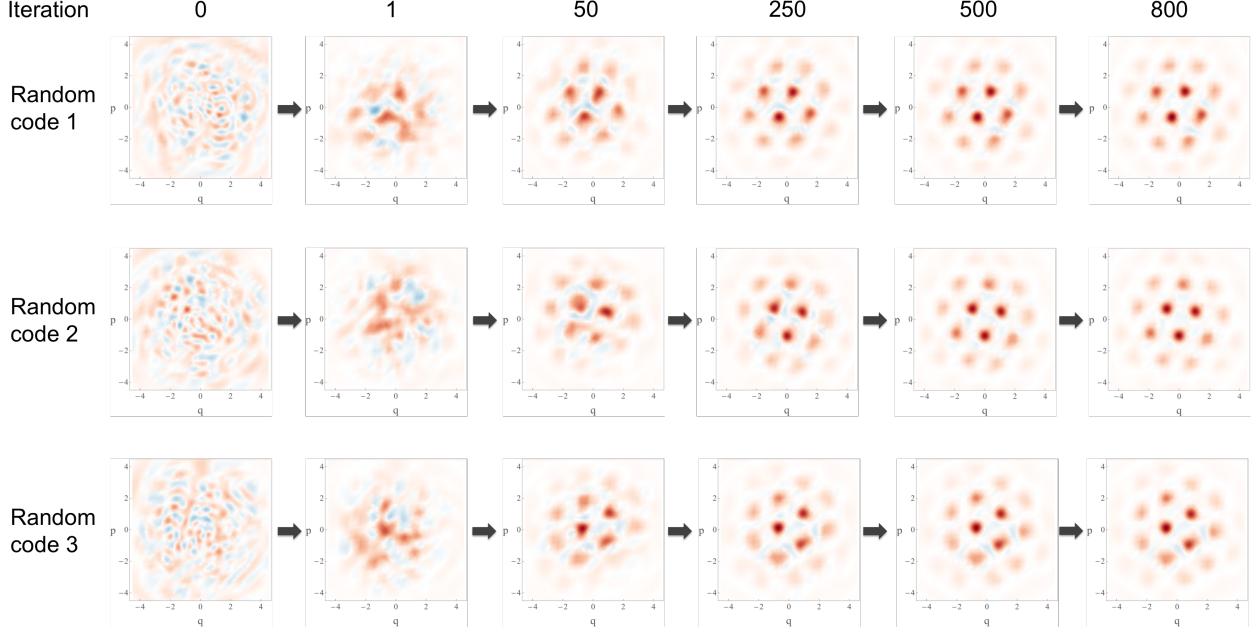


FIG. 4: Biconvex optimization of the encoding and

decoding maps for bosonic pure-loss channel $\mathcal{N}[\eta, 0]$ with $\eta = 0.9$. We chose $n = 20$ and $d = 2$ and imposed photon number constraint $\text{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}] \leq 3$, where $\hat{\rho}_{\mathcal{E}} = (1/d)\text{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{E}}$ is the maximally mixed code state. First column of each row is the Wigner function of $\hat{\rho}_{\mathcal{E}}$ for a randomly generated encoding map \mathcal{E} . From the second to the sixth columns represent the updated code spaces after 1, 50, 250, 500 and 800 iterations of alternating semidefinite programming. Color scale is the same as in Fig. 2.

in the encoded state, where $\hat{\rho}_{\mathcal{E}} \equiv \mathcal{E}(\frac{1}{d} \sum_{i=0}^{d-1} |i_{\mathcal{H}_0}\rangle\langle i_{\mathcal{H}_0}|) = \frac{1}{d}\text{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{E}}$ is the maximally mixed encoded state. Then, the energy constraint reads $\text{Tr}[(\hat{E} \otimes \hat{I}_{\mathcal{H}_0})\hat{X}_{\mathcal{E}}] \leq \bar{E}d$ and we obtain the following biconvex encoding and decoding optimization with energy constraint:

$$\begin{aligned} & \max_{\hat{X}_{\mathcal{E}}, \hat{X}_{\mathcal{D}}} \text{Tr}[\hat{X}_{\mathcal{D}}^\dagger f_{\mathcal{N}}(\hat{X}_{\mathcal{E}})], \\ & \text{s.t. } \hat{X}_{\mathcal{D}} = \hat{X}_{\mathcal{D}}^\dagger \succeq 0, \quad \text{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{D}} = \hat{I}_{\mathcal{H}_n}, \\ & \quad \hat{X}_{\mathcal{E}} = \hat{X}_{\mathcal{E}}^\dagger \succeq 0, \quad \text{Tr}_{\mathcal{H}_n}\hat{X}_{\mathcal{E}} = \hat{I}_{\mathcal{H}_0}, \\ & \quad \text{and } \text{Tr}[(\hat{E} \otimes \hat{I}_{\mathcal{H}_0})\hat{X}_{\mathcal{E}}] \leq \bar{E}d, \end{aligned} \quad (67)$$

where the second line in the constraint is due to the CPTP nature of the encoding map \mathcal{E} , and the third line is due to the encoding energy constraint.

We apply Eq.(67) to find an optimal qubit-into-oscillator encoding (i.e., $\dim(\mathcal{H}_0) = d = 2$) against the bosonic pure-loss channel $\mathcal{N}[\eta, 0]$ subject to the photon number constraint $\text{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}] \leq \bar{n}$. For practical implementation, since the photon number cannot increase under the bosonic pure-loss channel, we confine the bosonic Hilbert space to a truncated subspace $\mathcal{H}_n \equiv \text{span}\{|0\rangle, \dots, |n-1\rangle\}$, yielding $\mathcal{N}(\hat{\rho}) = \mathcal{N}[\eta, 0](\hat{\rho}) = \sum_{\ell=0}^{n-1} \hat{E}_\ell \hat{\rho} \hat{E}_\ell^\dagger$ where $\hat{E}_\ell = \sqrt{(1-\eta)^\ell / \ell!} \eta^{\frac{n}{2}} \hat{a}^\ell$. We choose $n \gg \bar{n}$ to avoid artifacts caused by truncation and solve Eq.(67) heuristically by alternating encoding and decoding optimization, each solved by semidefinite programming, starting from a random initial encoding map. We generate random initial codes by taking the first d columns of a Haar random

$n \times n$ unitary matrix. To solve each SDP sub-problem we used CVX, a package for specifying and solving convex programs [74, 75].

In Fig. 4, we took $\eta = 0.9$, $n = 20$, $d = 2$ and $\bar{n} = 3$ and plot the Wigner function of maximally mixed code states of the numerically optimized codes (last column), starting from three different random initial encoding maps (first column). In all instances, the obtained codes are similar to the hexagonal GKP code (in Fig. 2), except an overall displacement. We note that the optimized code in second row exhibits the best performance (i.e., $1 - F_N^* = 0.002092$) among all. Although we do not attempt to make a rigorous claim, this might indicate that the GKP codes defined over the symplectic lattice allowing the most efficient sphere packing may be the optimal encoding for the Gaussian loss channels. In particular, the constant gap in Theorem 14 may be closed if we assume the optimal decoding, instead of the sub-optimal one involving (noisy) quantum-limited amplification, which we assumed to prove Eq.(57).

We emphasize that the biconvex optimization in Eq.(67) explores the most general form of encoding maps, including the mixed state encoding. From the numerical optimization, however, we only obtained pure-state encoding (i.e., $\mathcal{E}(\hat{\rho}_0) = \hat{V}\hat{\rho}_0\hat{V}^\dagger$, where $\hat{V} : \mathcal{H}_0 \rightarrow \mathcal{H}_n$ is an isometry $\hat{V}^\dagger\hat{V} = \hat{I}_{\mathcal{H}_0}$) as an optimal solution at all iterations of SDP sub-problem. We also stress that the alternating semidefinite programming method is not guaranteed to yield a global optimal solution, although the obtained hexagonal GKP code in Fig. 4 may be a global optimum. In principle, global op-

timal solution of Eq.(67) can be deterministically found by global optimization algorithm outlined in [76]. To implement the algorithm, however, one should in general solve exponentially many convex sub-problems in the number of complicating variables (responsible for non-convexity of the problem; see [76] for details), which is intractable in our application.

VI. CONCLUSION

Exploiting various synthesis and decomposition of Gaussian channels, we provided improved upper bounds on the Gaussian thermal loss channel, both in the energy constrained and unconstrained cases, and the Gaussian displacement channel (Theorems 10, 11, 12 and Eq.(42)). We also established an achievable rate of the GKP codes for the Gaussian loss channels (Theorem 14). In the energy-unconstrained case, in particular, we showed that the GKP codes achieve the upper bound of Gaussian loss channel capacity upto at most a constant gap $\simeq \log e = 1.44269 \dots$ in the unit of qubits per channel use. In the energy-constrained case, we formulated a biconvex encoding and decoding optimization problem and solved it via alternating semidefinite programming method. We demonstrated in the one-mode case that hexagonal GKP code emerges as an optimal encoding from Haar random initial codes.

As was proven in [77], entanglement infidelity of the Petz decoding [78, 79] (or, transpose decoding) is at most twice as large as the optimal entanglement infidelity. Thus, communication rate achievable by Petz decoding equals to the optimal achievable rate. We leave finding the optimal communication rate of the GKP codes for Gaussian loss channels as an open problem.

Note added: While preparing the manuscript, we became aware of a related work [19] in which Lemma 8 and Theorem 10 were independently discovered. We also realized our Theorem 11 was independently discovered in [18]. We thank the authors of [18] for pointing out a mistake we made in Eq.(35) in an earlier version of our manuscript.

ACKNOWLEDGMENT

We would like to thank Steven M. Girvin, Barbara Terhal, John Preskill, Steven T. Flammia, Sekhar Tatikonda, Richard Kueng, Linshu Li and Mengzhen Zhang for fruitful discussions. We also thank Mark M. Wilde and Matteo Rosati for useful comments on our manuscript. We acknowledge support from the ARL-CDQI (W911NF-15-2-0067), ARO (W911NF-14-1-0011, W911NF-14-1-0563), ARO MURI (W911NF-16-1-0349), AFOSR MURI (FA9550-14-1-0052, FA9550-15-1-0015), NSF (EFMA-1640959), the Alfred P. Sloan Foundation (BR2013-049), and the Packard Foundation (2013-39273). KN acknowledges support through the Korea Foundation for Advanced Studies. VAA acknowledges support from the Walter Burke Institute for Theoretical Physics at Caltech.

Appendix A: Bosonic mode, Gaussian states and Gaussian unitary operations

Let \mathcal{H} denote an infinite-dimensional Hilbert space. Quantum states of N bosonic modes are in the tensor product of N such Hilbert spaces $\mathcal{H}^{\otimes N}$. Each bosonic mode is associated with the annihilation and creation operators \hat{a}_k and \hat{a}_k^\dagger , satisfying the bosonic communication relation $[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0$, $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$, where δ_{ij} is the Kronecker delta function and $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. Hilbert space \mathcal{H} is spanned by the eigenstates of the number operator $\hat{n} \equiv \hat{a}^\dagger \hat{a}$, i.e., $\mathcal{H} = \text{span}\{\lvert n \rangle\}_{n=0}^\infty$ where $\hat{n}\lvert n \rangle = n\lvert n \rangle$. In the number basis (or Fock basis), annihilation and creation operators are given by $\hat{a} = \sum_{n=1}^\infty \sqrt{n} \lvert n-1 \rangle \langle n \rvert$, $\hat{a}^\dagger = \sum_{n=0}^\infty \sqrt{n+1} \lvert n+1 \rangle \langle n \rvert$. Coherent state $\lvert \alpha \rangle$ is an eigenstate of the annihilation operator \hat{a} with eigenvalue α , i.e., $\hat{a}\lvert \alpha \rangle = \alpha\lvert \alpha \rangle$. In the Fock basis, $\lvert \alpha \rangle$ is given by $\lvert \alpha \rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^\infty \frac{\alpha^n}{\sqrt{n!}} \lvert n \rangle$. Note that the vacuum state $\lvert 0 \rangle$ is a special case of coherent state with $\alpha = 0$. Displacement operator $\hat{D}(\alpha)$ is defined as $\hat{D}(\alpha) \equiv \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$. Coherent state $\lvert \alpha \rangle$ is a displaced vacuum state:

$$\lvert \alpha \rangle = \hat{D}(\alpha)\lvert 0 \rangle. \quad (\text{A1})$$

Quadrature operators are defined as $\hat{q}_k \equiv \frac{1}{\sqrt{2}}(\hat{a}_k + \hat{a}_k^\dagger)$, $\hat{p}_k \equiv \frac{i}{\sqrt{2}}(\hat{a}_k^\dagger - \hat{a}_k)$ and are called position and momentum operator, respectively. We follow the same convention as used for the GKP codes [28, 42] which differs from Ref. [11] by a factor of $\sqrt{2}$ in the definition of \hat{q}_k and \hat{p}_k . Define $\hat{\mathbf{x}} \equiv (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_N, \hat{p}_N)^T$. Then, the bosonic commutation relation reads $[\hat{x}_i, \hat{x}_j] = i\Omega_{ij}$, where Ω is defined as

$$\Omega \equiv \begin{pmatrix} \omega & & & \\ & \ddots & & \\ & & \omega & \\ & & & \omega \end{pmatrix} \text{ and } \omega \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (\text{A2})$$

Eigenvalue spectrum of quadrature operators are continuous, $\hat{q}\lvert q \rangle = q\lvert q \rangle$, $\hat{p}\lvert p \rangle = p\lvert p \rangle$, where $q, p \in (-\infty, \infty)$ and the eigenstates are normalized by the Dirac delta function: $\langle q \lvert q' \rangle = \delta(q - q')$ and $\langle p \lvert p' \rangle = \delta(p - p')$. Also, $\lvert q \rangle$ and $\lvert p \rangle$ are related by Fourier transformation $\lvert q \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp e^{-iqp} \lvert p \rangle$, $\lvert p \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dq e^{ipq} \lvert q \rangle$. Upon displacement,

$$\begin{aligned} \hat{D}(\xi_1/\sqrt{2})\lvert q \rangle &= e^{-i\xi_1\hat{p}}\lvert q \rangle = \lvert q + \xi_1 \rangle, \\ \hat{D}(i\xi_2/\sqrt{2})\lvert p \rangle &= e^{i\xi_2\hat{q}}\lvert p \rangle = \lvert p + \xi_2 \rangle. \end{aligned} \quad (\text{A3})$$

Let $\mathcal{L}(\mathcal{H})$ be the space of linear operators on the Hilbert space \mathcal{H} . General quantum states (including mixed states) are described by density operator $\hat{\rho} \in \mathcal{D}(\mathcal{H})$, where $\mathcal{D}(\mathcal{H}) \equiv \{\hat{\rho} \in \mathcal{L}(\mathcal{H}) \mid \hat{\rho}^\dagger = \hat{\rho} \succeq 0, \text{Tr}[\hat{\rho}] = 1\}$. Expectation value of an observable \hat{E} of the state $\hat{\rho}$ is given by $\langle \hat{E} \rangle = \text{Tr}[\hat{\rho}\hat{E}]$. Wigner characteristic function $\chi(\xi)$ is defined as $\chi(\xi) \equiv \text{Tr}[\hat{\rho} \exp(i\hat{\mathbf{x}}^T \Omega \xi)]$, where $\hat{\rho} \in \mathcal{D}(\mathcal{H}^{\otimes N})$ and $\xi = (\xi_1, \dots, \xi_{2N})^T$. Weyl operator $\exp(i\hat{\mathbf{x}}^T \Omega \xi)$ is of the form

of displacement operator and satisfies the orthogonality relation $\text{Tr}[\exp(-i\hat{\mathbf{x}}^T \Omega \xi) \exp(i\hat{\mathbf{x}}^T \Omega \xi')] = (2\pi)^N \delta(\xi - \xi')$. Wigner characteristic function $\chi(\xi)$ is in one-to-one correspondence with the state $\hat{\rho}$ and the inverse function is given by $\hat{\rho} = \frac{1}{(2\pi)^N} \int d^{2N} \xi \chi(\xi) \exp(-i\hat{\mathbf{x}}^T \Omega \xi)$. Wigner function $W(\mathbf{x})$ is Fourier transformation of $\chi(\xi)$, i.e.,

$$W(\mathbf{x}) = \frac{1}{(2\pi)^N} \int d^{2N} \xi \chi(\xi) \exp(-i\mathbf{x}^T \Omega \xi), \quad (\text{A4})$$

where $\mathbf{x} = (x_1, \dots, x_N)^T$ is the eigenvalue of the quadrature operator $\hat{\mathbf{x}}$.

A quantum state $\hat{\rho}$ is Gaussian iff its Wigner characteristic function and Wigner function are Gaussian [11]:

$$\begin{aligned} \chi(\xi) &= \exp \left[-\frac{1}{2} \xi^T (\Omega \mathbf{V} \Omega^T) \xi - i(\Omega \bar{\mathbf{x}}) \xi \right], \\ W(\mathbf{x}) &= \frac{\exp \left[-\frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{V}^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \right]}{(2\pi)^N \sqrt{\det \mathbf{V}}}, \end{aligned} \quad (\text{A5})$$

where $\bar{\mathbf{x}}, \mathbf{V}$ are first and second moments of the state $\hat{\rho}$.

$$\bar{\mathbf{x}} \equiv \langle \hat{\mathbf{x}} \rangle = \text{Tr}[\hat{\rho} \hat{\mathbf{x}}], \quad V_{ij} \equiv \frac{1}{2} \langle \{ \hat{x}_i - \bar{x}_i, \hat{x}_j - \bar{x}_j \} \rangle, \quad (\text{A6})$$

where $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$. Thus, Gaussian states are fully characterized by the first two moments, i.e., $\hat{\rho} = \hat{\rho}_G(\bar{\mathbf{x}}, \mathbf{V})$. Heisenberg uncertainty relation reads $\mathbf{V} + \frac{i}{2} \Omega \succeq 0$ and implies $V(\hat{q}_k)V(\hat{p}_k) \geq \frac{1}{4}$ for all $k \in \{1, \dots, N\}$, where $V(\hat{x}_i) \equiv \mathbf{V}_{ii}$.

Vacuum state $|0\rangle\langle 0|$ is the simplest example of one-mode Gaussian states with $\bar{\mathbf{x}} = \mathbf{0}$ and $\mathbf{V} = \frac{\mathbf{I}_2}{2}$, where \mathbf{I}_n is defined as the $n \times n$ identity matrix. Coherent state $|\alpha\rangle\langle\alpha|$ is also Gaussian: $|\alpha\rangle\langle\alpha| = \hat{\rho}_G(\bar{\mathbf{x}}_\alpha, \frac{\mathbf{I}_2}{2})$ with $\bar{\mathbf{x}}_\alpha \equiv \sqrt{2}(\alpha_R, \alpha_I)^T$ and $\alpha = \alpha_R + i\alpha_I$. Coherent states (including the vacuum state) saturate the uncertainty inequality and thus have the minimum uncertainty. Thermal state is an example of Gaussian mixed states and is given by

$$\hat{\rho}_{\bar{n}_{\text{th}}} \equiv \sum_{n=0}^{\infty} \frac{(\bar{n}_{\text{th}})^n}{(\bar{n}_{\text{th}} + 1)^{n+1}} |n\rangle\langle n| = \hat{\rho}_G(\mathbf{0}, (\bar{n}_{\text{th}} + \frac{1}{2}) \mathbf{I}_2), \quad (\text{A7})$$

in the Fock basis, where \bar{n}_{th} is the average photon number, i.e., $\bar{n}_{\text{th}} = \text{Tr}[\hat{\rho}_{\bar{n}_{\text{th}}} \hat{n}]$. Quantum entropy of a state $\hat{\rho}$ is defined as $H(\hat{\rho}) \equiv -\text{Tr}[\hat{\rho} \log \hat{\rho}]$, where \log is the logarithm with base 2. Entropy of a thermal state $\hat{\rho}_{\bar{n}_{\text{th}}}$ is given by

$$H(\hat{\rho}_{\bar{n}_{\text{th}}}) = g(\bar{n}_{\text{th}}), \quad (\text{A8})$$

where $g(x) \equiv (x + 1) \log(x + 1) - x \log x$. Since thermal state is a mixed state, $H(\hat{\rho}_{\bar{n}_{\text{th}}}) \geq 0$ where the equality holds only when the state is vacuum, i.e., $\bar{n}_{\text{th}} = 0$.

Unitary operations that map a Gaussian state to another Gaussian state are called Gaussian unitary operations. Gaussian unitaries are generated by second order polynomials of $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_N)^T$ and $\hat{\mathbf{a}}^\dagger = (\hat{a}_1^\dagger, \dots, \hat{a}_N^\dagger)^T$, i.e., $\hat{U}_G = \exp(-i\hat{H})$ with $\hat{H} = i(\alpha^T \hat{\mathbf{a}}^\dagger + \hat{\mathbf{a}}^\dagger \mathbf{F} \hat{\mathbf{a}} + \hat{\mathbf{a}}^\dagger \mathbf{G} \hat{\mathbf{a}}^\dagger) + \text{h.c.}$, where $\alpha^T = (\alpha_1, \dots, \alpha_N)$ and \mathbf{F}, \mathbf{G} are $N \times N$ complex matrices.

In the Heisenberg picture, annihilation operator $\hat{\mathbf{a}}$ is transformed into $\hat{U}_G^\dagger \hat{\mathbf{a}} \hat{U}_G = \mathbf{A} \hat{\mathbf{a}} + \mathbf{B} \hat{\mathbf{a}}^\dagger + \alpha$, where $N \times N$ complex matrices \mathbf{A}, \mathbf{B} (determined by \mathbf{F}, \mathbf{G}) satisfy $\mathbf{AB}^T = \mathbf{BA}^T$ and $\mathbf{AA}^\dagger = \mathbf{BB}^\dagger + \mathbf{I}_N$. In terms of quadrature operators, the transformation reads $\mathbf{x} \rightarrow \hat{U}_G^\dagger \mathbf{x} \hat{U}_G = \mathbf{S} \mathbf{x} + \mathbf{d}$, where $\mathbf{d} = (d_1, \dots, d_{2N})^T = \sqrt{2}(\alpha_1^R, \alpha_1^I, \dots, \alpha_N^R, \alpha_N^I)^T$ and $2N \times 2N$ matrix \mathbf{S} is symplectic:

$$\mathbf{S} \Omega \mathbf{S}^T = \Omega. \quad (\text{A9})$$

Gaussian unitaries are thus fully characterized by \mathbf{S}, \mathbf{d} , and under $\hat{U}_{\mathbf{S}, \mathbf{d}}$ the first two moments are transformed as

$$\bar{\mathbf{x}} \rightarrow \mathbf{S} \bar{\mathbf{x}} + \mathbf{d}, \quad \mathbf{V} \rightarrow \mathbf{S} \mathbf{V} \mathbf{S}^T. \quad (\text{A10})$$

Displacement operator $\hat{D}(\alpha)$ is a one-mode Gaussian unitary operation with α and $\mathbf{F} = \mathbf{G} = \mathbf{0}$, yielding $\mathbf{A} = \mathbf{1}, \mathbf{B} = \mathbf{0}$ and $\mathbf{S} = \mathbf{I}_2, \mathbf{d} = \sqrt{2}(\alpha_R, \alpha_I)^T$. Squeeze operator $\hat{S}(r) \equiv \exp(\frac{r}{2}(\hat{a}^2 - \hat{a}^\dagger 2))$ is another example of one-mode Gaussian unitaries and transforms quadrature operators by $\hat{q} \rightarrow e^{-r} \hat{q}$ and $\hat{p} \rightarrow e^r \hat{p}$, i.e., $\mathbf{S} = \text{diag}(e^{-r}, e^r)$. Quadrature eigenstate can thus be understood as an infinitely squeezed state: For example, $|\hat{q} = 0\rangle \propto \lim_{r \rightarrow +\infty} \hat{S}(r)|0\rangle$ and $|\hat{p} = 0\rangle \propto \lim_{r \rightarrow -\infty} \hat{S}(r)|0\rangle$, where $|0\rangle$ is the vacuum state. Phase rotation operator is defined as $\hat{U}(\theta) \equiv \exp(i\theta \hat{a}^\dagger \hat{a})$. Under the phase rotation, quadrature operators are transformed as

$$\mathbf{x} \rightarrow \mathbf{R}(\theta)\mathbf{x} \text{ and } \mathbf{R}(\theta) \equiv \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (\text{A11})$$

yielding, e.g., $\hat{U}(\theta)|\alpha\rangle = |\alpha e^{i\theta}\rangle$.

Beam splitter is a two-mode Gaussian unitary operation generated by the Hamiltonian of the form $\hat{H} \propto i(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger)$. Beam splitter unitary $\hat{B}(\eta)$ transforms the annihilation operators by $\hat{a}_1 \rightarrow \sqrt{\eta} \hat{a}_1 + \sqrt{1-\eta} \hat{a}_2$ and $\hat{a}_2 \rightarrow -\sqrt{1-\eta} \hat{a}_1 + \sqrt{\eta} \hat{a}_2$, where $\eta = \cos^2 \theta \in [0, 1]$ is called the transmissivity. In terms of the quadrature operator $\hat{\mathbf{x}} = (\hat{q}_1, \hat{p}_1, \hat{q}_2, \hat{p}_2)^T$, the transformation reads

$$\hat{\mathbf{x}} \rightarrow \mathbf{B}(\eta)\hat{\mathbf{x}} \text{ and } \mathbf{B}(\eta) \equiv \begin{pmatrix} \sqrt{\eta} \mathbf{I}_2 & \sqrt{1-\eta} \mathbf{I}_2 \\ -\sqrt{1-\eta} \mathbf{I}_2 & \sqrt{\eta} \mathbf{I}_2 \end{pmatrix}. \quad (\text{A12})$$

Another example of two-mode Gaussian operation is the two-mode squeezing generated by $\hat{H} \propto i(\hat{a}_1 \hat{a}_2 - \hat{a}_1^\dagger \hat{a}_2^\dagger)$. Under the two-mode squeezing $\hat{S}_2(G)$, annihilation operators are transformed as $\hat{a}_1 \rightarrow \sqrt{G} \hat{a}_1 + \sqrt{G-1} \hat{a}_2^\dagger$ and $\hat{a}_2 \rightarrow \sqrt{G-1} \hat{a}_1^\dagger + \sqrt{G} \hat{a}_2$, where $G \geq 1$ is the gain. Under $\hat{S}_2(G)$, quadrature operators are transformed as

$$\hat{\mathbf{x}} \rightarrow \mathbf{S}_2(G)\hat{\mathbf{x}} \text{ and } \mathbf{S}_2(G) \equiv \begin{pmatrix} \sqrt{G} \mathbf{I}_2 & \sqrt{G-1} \mathbf{Z}_2 \\ \sqrt{G-1} \mathbf{Z}_2 & \sqrt{G} \mathbf{I}_2 \end{pmatrix}, \quad (\text{A13})$$

where $\mathbf{Z}_2 \equiv \text{diag}(1, -1)$.

Appendix B: Characterization of Gaussian random displacement channel

Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2](\hat{\rho})$ as defined in Definition 3 is characterized by $\mathbf{T} = \mathbf{I}_2$, $\mathbf{N} = \sigma^2 \mathbf{I}_2$ and $\mathbf{d} = 0$.

Proof. Let $\hat{\rho}' \equiv \mathcal{N}_{B_2}[\sigma^2](\hat{\rho})$. Note that, in the case of one-mode, $\exp(i\hat{x}^T \Omega \xi) = \hat{D}(\xi/\sqrt{2})$ where $\xi = \xi_1 + i\xi_2$ (see Appendix A for the definition of Ω and \hat{x}). The Wigner characteristic function of $\hat{\rho}'$ is then given by

$$\begin{aligned}\chi'(\xi) &= \text{Tr}\left[\mathcal{N}_{B_2}[\sigma^2](\hat{\rho})\hat{D}\left(\frac{\xi}{\sqrt{2}}\right)\right] \\ &= \text{Tr}\left[\frac{1}{\pi\sigma^2} \int d^2\alpha e^{-\frac{|\alpha|^2}{\sigma^2}} \hat{\rho}\hat{D}^\dagger(\alpha)\hat{D}\left(\frac{\xi}{\sqrt{2}}\right)\hat{D}(\alpha)\right] \\ &= \frac{1}{\pi\sigma^2} \int d^2\alpha e^{-\frac{|\alpha|^2}{\sigma^2}} e^{\frac{1}{\sqrt{2}}(\alpha^*\xi - \alpha\xi^*)} \text{Tr}\left[\hat{\rho}\hat{D}\left(\frac{\xi}{\sqrt{2}}\right)\right] \\ &= e^{-\frac{\sigma^2}{2}|\xi|^2} \chi(\xi),\end{aligned}\quad (\text{B1})$$

where $\chi(\xi)$ is the Wigner characteristic function of the input state $\hat{\rho}$ —cf. text above Eq.(A4). We used cyclic property of trace to derive the second equality, and applied Baker-Campbell-Hausdorff formula to obtain $\hat{D}^\dagger(\alpha)\hat{D}(\beta)\hat{D}(\alpha) = \exp(\alpha^*\beta - \alpha\beta^*)\hat{D}(\beta)$ and the third equality. The last equality follows from the evaluation of Gaussian integral and the definition of $\chi(\xi)$. If the input state is Gaussian, i.e., $\hat{\rho} = \hat{\rho}_G(\bar{x}, \mathbf{V})$, Wigner characteristic function of the output state is given by

$$\chi'(\xi) = \exp\left[-\frac{1}{2}\xi^T(\Omega(\mathbf{V} + \sigma^2 \mathbf{I}_2)\Omega^T)\xi - i(\Omega\bar{x})\xi\right], \quad (\text{B2})$$

and thus $\hat{\rho}' = \hat{\rho}_G(\bar{x}, \mathbf{V} + \sigma^2 \mathbf{I}_2)$. ■

Appendix C: Derivation of the optimized data-processing bound

Proof. Assume that a Gaussian loss channel is decomposed as

$$\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{A}[G_1]\mathcal{N}[\bar{\eta}, 0]\mathcal{A}[G_2], \quad (\text{C1})$$

cf. Eq.(38). Then, the channel on the right hand side is characterized by $\mathbf{T} = \sqrt{G_1\bar{\eta}G_2}\mathbf{I}_2$ and

$$\begin{aligned}\mathbf{N} &= \left[G_1\left(\bar{\eta} \times \frac{1}{2}(G_2 - 1) + \frac{1}{2}(1 - \bar{\eta})\right) + \frac{1}{2}(G_1 - 1)\right]\mathbf{I}_2 \\ &= \frac{1}{2}\left[G_1\bar{\eta}G_2 - 2G_1\bar{\eta} + 2G_1 - 1\right]\mathbf{I}_2.\end{aligned}\quad (\text{C2})$$

Imposing $\mathbf{T} = \sqrt{\eta}\mathbf{I}_2$ and $\mathbf{N} = (1 - \eta)(\bar{n}_{\text{th}} + \frac{1}{2})\mathbf{I}_2$, we get $(1 - \bar{\eta})G_1 = (1 - \eta)(\bar{n}_{\text{th}} + 1)$ and thus $\bar{\eta} = 1 - \frac{1}{G_1}(1 - \eta)(\bar{n}_{\text{th}} + 1)$, $G_2 = \frac{\eta}{G_1\bar{\eta}} = \frac{\eta}{G_1 - (1 - \eta)(\bar{n}_{\text{th}} + 1)}$. Then, $Q_{\text{reg}}^{n \leq \bar{n}}(\mathcal{N}[\eta, \bar{n}_{\text{th}}])$ is upper bounded by

$$\begin{aligned}Q^{n \leq G_2\bar{n} + (G_2 - 1)}(\mathcal{N}[\bar{\eta}, 0]) \\ \equiv \max \left[g(\bar{\eta}(G_2\bar{n} + (G_2 - 1))) \right. \\ \left. - g((1 - \bar{\eta})(G_2\bar{n} + (G_2 - 1))), 0 \right].\end{aligned}\quad (\text{C3})$$

hence Eq.(41). To ensure $G_1 \geq 1$ and $G_2 = \frac{\eta}{G_1 - (1 - \eta)(\bar{n}_{\text{th}} + 1)} \geq 1$, G_1 should be in the range $1 \leq G_1 \leq 1 + (1 - \eta)\bar{n}_{\text{th}}$. Thus, Eq.(40) follows. ■

-
- [1] Mark M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
 - [2] Peter W. Shor, “Scheme for reducing decoherence in quantum computer memory,” Phys. Rev. A **52**, R2493–R2496 (1995).
 - [3] Benjamin Schumacher and M. A. Nielsen, “Quantum data processing and error correction,” Phys. Rev. A **54**, 2629–2635 (1996).
 - [4] Seth Lloyd, “Capacity of the noisy quantum channel,” Phys. Rev. A **55**, 1613–1622 (1997).
 - [5] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel,” IEEE Transactions on Information Theory **51**, 44–55 (2005).
 - [6] David P. DiVincenzo, Peter W. Shor, and John A. Smolin, “Quantum-channel capacity of very noisy channels,” Phys. Rev. A **57**, 830–839 (1998).
 - [7] Graeme Smith and Jon Yard, “Quantum communication with zero-capacity channels,” Science **321**, 1812–1815 (2008), <http://science.sciencemag.org/content/321/5897/1812.full.pdf>.
 - [8] Graeme Smith, John A. Smolin, and Jon Yard, “Quantum communication with gaussian channels of zero quantum capacity,” Nature Photonics **5**, 624 EP – (2011).
 - [9] Toby Cubitt, David Elkouss, William Matthews, Maris Ozols, David Pérez-García, and Sergii Strelchuk, “Unbounded number of channel uses may be required to detect quantum capacity,” Nature Communications **6**, 6739 EP – (2015).
 - [10] J. Eisert and M. M. Wolf, “Gaussian quantum channels,” eprint arXiv:quant-ph/0505151 (2005), quant-ph/0505151.
 - [11] 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).
 - [12] Alessio Serafini, *Quantum Continuous Variables* (CRC Press, 2017).
 - [13] A. S. Holevo and R. F. Werner, “Evaluating capacities of bosonic gaussian channels,” Phys. Rev. A **63**, 032312 (2001).
 - [14] Michael M. Wolf, David Pérez-García, and Geza Giedke, “Quantum capacities of bosonic channels,” Phys. Rev. Lett. **98**, 130501 (2007).
 - [15] Mark M. Wilde, Patrick Hayden, and Saikat Guha, “Quantum trade-off coding for bosonic communication,” Phys. Rev. A **86**, 062306 (2012).
 - [16] M. M. Wilde and H. Qi, “Energy-constrained private and quantum capacities of quantum channels,” ArXiv e-prints (2016), arXiv:1609.01997 [quant-ph].
 - [17] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi, “Fundamental limits of repeaterless quantum communications,” Nature Communications **8**, 15043 EP – (2017).

- [18] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, “Bounding the energy-constrained quantum and private capacities of phase-insensitive Gaussian channels,” ArXiv e-prints (2017), arXiv:1708.07257v2 [quant-ph].
- [19] M. Rosati, A. Mari, and V. Giovannetti, “Narrow Bounds for the Quantum Capacity of Thermal Attenuators,” ArXiv e-prints (2018), arXiv:1801.04731v2 [quant-ph].
- [20] P. T. Cochrane, G. J. Milburn, and W. J. Munro, “Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping,” Phys. Rev. A **59**, 2631–2634 (1999).
- [21] J. Niset, U. L. Andersen, and N. J. Cerf, “Experimentally Feasible Quantum Erasure-Correcting Code for Continuous Variables,” Phys. Rev. Lett. **101**, 130503 (2008).
- [22] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J. Schoelkopf, Michel H. Devoret, and Mazyar Mirrahimi, “Hardware-efficient autonomous quantum memory protection,” Phys. Rev. Lett. **111**, 120501 (2013).
- [23] Mazyar Mirrahimi, Zaki Leghtas, Victor V Albert, Steven Touzard, Robert J Schoelkopf, Liang Jiang, and Michel H Devoret, “Dynamically protected cat-qubits: a new paradigm for universal quantum computation,” New Journal of Physics **16**, 045014 (2014).
- [24] Felipe Lacerda, Joseph M. Renes, and Volkher B. Scholz, “Coherent state constellations for Bosonic Gaussian channels,” in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2016) pp. 2499–2503.
- [25] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang, “Multimode cat codes,” ArXiv e-prints (2018), arXiv:1801.05897 [quant-ph].
- [26] Seth Lloyd and Jean-Jacques E. Slotine, “Analog Quantum Error Correction,” Phys. Rev. Lett. **80**, 4088–4091 (1998).
- [27] Samuel L. Braunstein, “Error correction for continuous quantum variables,” Phys. Rev. Lett. **80**, 4084–4087 (1998).
- [28] Daniel Gottesman, Alexei Kitaev, and John Preskill, “Encoding a qubit in an oscillator,” Phys. Rev. A **64**, 012310 (2001).
- [29] Nicolas C. Menicucci, “Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States,” Phys. Rev. Lett. **112**, 120504 (2014).
- [30] Patrick Hayden, Sepehr Nezami, Grant Salton, and Barry C Sanders, “Spacetime replication of continuous variable quantum information,” New J. Phys. **18**, 083043 (2016).
- [31] A. Ketterer, A. Keller, S. P. Walborn, T. Coudreau, and P. Milman, “Quantum information processing in phase space: A modular variables approach,” Phys. Rev. A **94**, 022325 (2016).
- [32] Isaac L. Chuang, Debbie W. Leung, and Yoshihisa Yamamoto, “Bosonic quantum codes for amplitude damping,” Phys. Rev. A **56**, 1114–1125 (1997).
- [33] E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics,” Nature **409**, 46–52 (2001).
- [34] T. C. Ralph, A. J. F. Hayes, and Alexei Gilchrist, “Loss-Tolerant Optical Qubits,” Phys. Rev. Lett. **95**, 100501 (2005).
- [35] Wojciech Wasilewski and Konrad Banaszek, “Protecting an optical qubit against photon loss,” Phys. Rev. A **75**, 042316 (2007).
- [36] Marcel Bergmann and Peter van Loock, “Quantum error correction against photon loss using NOON states,” Phys. Rev. A **94**, 012311 (2016).
- [37] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin, “New class of quantum error-correcting codes for a bosonic mode,” Phys. Rev. X **6**, 031006 (2016).
- [38] M. Yuezhen Niu, I. L. Chuang, and J. H. Shapiro, “Hardware-Efficient Bosonic Quantum Error-Correcting Codes Based on Symmetry Operators,” ArXiv e-prints (2017), arXiv:1709.05302 [quant-ph].
- [39] Seung-Woo Lee and Hyunseok Jeong, “Near-deterministic quantum teleportation and resource-efficient quantum computation using linear optics and hybrid qubits,” Phys. Rev. A **87**, 022326 (2013).
- [40] Eliot Kapit, “Hardware-Efficient and Fully Autonomous Quantum Error Correction in Superconducting Circuits,” Phys. Rev. Lett. **116**, 150501 (2016).
- [41] Jim Harrington and John Preskill, “Achievable rates for the gaussian quantum channel,” Phys. Rev. A **64**, 062301 (2001).
- [42] V. V. Albert, K. Noh, K. Duivenvoorden, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, “Performance and structure of bosonic codes,” ArXiv e-prints (2017), arXiv:1708.05010 [quant-ph].
- [43] Man-Duen Choi, “Completely positive linear maps on complex matrices,” Linear Algebra and its Applications **10**, 285 – 290 (1975).
- [44] Filippo Caruso and Vittorio Giovannetti, “Degradability of bosonic gaussian channels,” Phys. Rev. A **74**, 062307 (2006).
- [45] F Caruso, V Giovannetti, and A S Holevo, “One-mode bosonic gaussian channels: a full weak-degradability classification,” New Journal of Physics **8**, 310 (2006).
- [46] A. S. Holevo, “One-mode quantum gaussian channels: Structure and quantum capacity,” Problems of Information Transmission **43**, 1–11 (2007).
- [47] H. Heffner, “The fundamental noise limit of linear amplifiers,” Proceedings of the IRE **50**, 1604–1608 (1962).
- [48] Daniel Gottesman and John Preskill, “Secure quantum key distribution using squeezed states,” Phys. Rev. A **63**, 022309 (2001).
- [49] Raúl García-Patrón, Carlos Navarrete-Benlloch, Seth Lloyd, Jeffrey H. Shapiro, and Nicolas J. Cerf, “Majorization theory approach to the gaussian channel minimum entropy conjecture,” Phys. Rev. Lett. **108**, 110505 (2012).
- [50] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, “Approximate degradable quantum channels,” IEEE Transactions on Information Theory **63**, 7832–7844 (2017).
- [51] J. Eisert, S. Scheel, and M. B. Plenio, “Distilling gaussian states with gaussian operations is impossible,” Phys. Rev. Lett. **89**, 137903 (2002).
- [52] Julien Niset, Jaromír Fiurášek, and Nicolas J. Cerf, “No-go theorem for gaussian quantum error correction,” Phys. Rev. Lett. **102**, 120501 (2009).
- [53] Daniel Gottesman and Isaac L. Chuang, “Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations,” Nature **402**, 390 EP – (1999).
- [54] B. C. Travaglione and G. J. Milburn, “Preparing encoded states in an oscillator,” Phys. Rev. A **66**, 052322 (2002).
- [55] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, “Constructing finite-dimensional codes with optical continuous variables,” EPL (Europhysics Letters) **68**, 323 (2004).
- [56] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, “Generating continuous variable quantum codewords in the near-field atomic lithography,” Journal of Physics B Atomic Molecular Physics **39**, 997–1009 (2006), quant-ph/0510053.
- [57] H. M. Vasconcelos, L. Sanz, and S. Glancy, “All-optical generation of states for “encoding a qubit in an oscillator”,” Opt. Lett. **35**, 3261–3263 (2010).
- [58] B. M. Terhal and D. Weigand, “Encoding a qubit into a cavity mode in circuit qed using phase estimation,” Phys. Rev. A **93**, 012315 (2016).
- [59] Keith R. Motes, Ben Q. Baragiola, Alexei Gilchrist, and Nicolas C. Menicucci, “Encoding qubits into oscillators with

- atomic ensembles and squeezed light,” Phys. Rev. A **95**, 053819 (2017).
- [60] D. J. Weigand and B. M. Terhal, “Breeding Grid States From Schrödinger Cat States without Post-Selection,” ArXiv e-prints (2017), arXiv:1709.08580 [quant-ph].
- [61] Alexei Kitaev, “Quantum measurements and the abelian stabilizer problem,” *Electronic Colloquium on Computational Complexity (ECCC)*, **3** (1996).
- [62] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, “Sequential modular position and momentum measurements of a trapped ion mechanical oscillator,” ArXiv e-prints (2017), arXiv:1709.10469 [quant-ph].
- [63] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. M. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, L. Frunzio, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, “Tracking photon jumps with repeated quantum non-demolition parity measurements,” Nature **511**, 444–448 (2014).
- [64] L. Fejes, “Über die dichteste kugellagerung,” Mathematische Zeitschrift **48**, 676–684 (1942).
- [65] Edmund Hlawka, “Zur geometrie der zahlen,” Mathematische Zeitschrift **49**, 285–312 (1943).
- [66] P. Buser and P. Sarnak, “On the period matrix of a Riemann surface of large genus (with an Appendix by J. H. Conway and N. J. A. Sloane).” Inventiones Mathematicae **117**, 27 (1994).
- [67] Maryna S Viazovska, “The sphere packing problem in dimension 8,” Annals of Mathematics **185**, 991–1015 (2017).
- [68] Henry Cohn, Abhinav Kumar, Stephen D Miller, Danylo Radchenko, and Maryna Viazovska, “The sphere packing problem in dimension 24,” Annals of Mathematics **185**, 1017–1033 (2017).
- [69] John Preskill, “Lecture note for quantum computation, chapter 3. foundations of quantum theory ii: Measurement and evolution,” .
- [70] Benjamin Schumacher, “Sending entanglement through noisy quantum channels,” Phys. Rev. A **54**, 2614–2628 (1996).
- [71] M. Reimpell and R. F. Werner, “Iterative optimization of quantum error correcting codes,” Phys. Rev. Lett. **94**, 080501 (2005).
- [72] Andrew S. Fletcher, Peter W. Shor, and Moe Z. Win, “Optimum quantum error recovery using semidefinite programming,” Phys. Rev. A **75**, 012338 (2007).
- [73] Robert L. Kosut and Daniel A. Lidar, “Quantum error correction via convex optimization,” Quantum Information Processing **8**, 443–459 (2009).
- [74] Inc. CVX Research, “CVX: Matlab software for disciplined convex programming, version 2.0,” (2012).
- [75] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110.
- [76] C. A. Floudas and V. Visweswaran, “A global optimization algorithm (gop) for certain classes of nonconvex nlps—i. theory,” Computers & Chemical Engineering **14**, 1397–1417 (1990).
- [77] H. Barnum and E. Knill, “Reversing quantum dynamics with near-optimal quantum and classical fidelity,” Journal of Mathematical Physics **43**, 2097–2106 (2002), <http://aip.scitation.org/doi/pdf/10.1063/1.1459754>.
- [78] Dénes Petz, “Sufficient subalgebras and the relative entropy of states of a von neumann algebra,” Communications in Mathematical Physics **105**, 123–131 (1986).
- [79] Dénes Petz, “Sufficiency of channels over von neumann algebras,” The Quarterly Journal of Mathematics **39**, 97–108 (1988).