Quantum Capacity Bounds of Gaussian Thermal Loss Channels and Achievable Rates With Gottesman-Kitaev-Preskill Codes

Kyungjoo Noh, Victor V. Albert, and Liang Jiang

***Abstract*— Gaussian thermal loss channels are of particular importance to quantum communication theory since they model realistic optical communication channels. Except for special cases, the quantum capacity of Gaussian thermal loss channels is not yet quantified completely. In this paper, we provide improved upper bounds of the Gaussian thermal loss channel capacity, both in energy-constrained and unconstrained scenarios. We briefly review Gottesman-Kitaev-Preskill (GKP) codes and discuss their experimental implementation. We then prove, in the energy-unconstrained case, that a family of GKP codes achieves the quantum capacity of Gaussian thermal 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 entanglement fidelity. Then, we solve the biconvex optimization heuristically by an alternating semi-definite programming method and report that, starting from Haar random initial codes, our numerical optimization yields a hexagonal GKP code as an optimal encoding in a practically relevant regime.**

***Index Terms*— Communication channels, channel capacity, error correction codes, quantum entanglement, optimization.**

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 to the classical case, practical quantum communication channels are inevitably noisy, and thus 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].

The quantum capacity of a channel quantifies the maximum number 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 classical communication theory, *regularized coherent information* of a channel characterizes that channel's quantum capacity [3]–[5]. Unlike its classical counterpart, however, coherent information might be *strictly superadditive*, which makes it hard to evaluate a channel's quantum capacity [6]–[9].

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

Evaluation of the quantum capacity, however, does not lend explicit encoding and decoding strategies that achieve capacity. In parallel with characterization of the quantum capacity of Gaussian channels, 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], and 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, Gottesman-Kitaev-Preskill (GKP) codes [28] are particularly interesting as they achieve the one-shot coherent information of Gaussian random displacement channels [41]. In our recent work, we and collaborators showed, both numerically and analytically, that GKP codes also exhibit excellent performance against bosonic pure-loss errors, outperforming many other bosonic codes [42].

Manuscript received February 7, 2018; revised July 7, 2018 and September 23, 2018; accepted September 24, 2018. Date of publication October 5, 2018; date of current version March 15, 2019. This work was supported in part by the ARL-CDQI under Grant W911NF-15-2-0067, in part by ARO under Grant W911NF-14-1-0011, Grant W911NF-14-1-0563, Grant W911NF-18-1-0020, and Grant W911NF-18-1-0212, in part by ARO MURI under Grant W911NF-16-1-0349, in part by AFOSR MURI under Grant FA9550-14-1-0052 and Grant FA9550-15-1-0015, in part by NSF under Grant EFMA-1640959, in part by the Alfred P. Sloan Foundation under Grant BR2013-049, and in part by the Packard Foundation under Grant 2013-39273. K. Noh was supported by the Korea Foundation for Advanced Studies. V. V. Albert was supported by the Walter Burke Institute for Theoretical Physics, Caltech.

K. Noh is with the Department of Physics, Yale University, New Haven, CT 06511 USA, and also with the Yale Quantum Institute, Yale University, New Haven, CT 06520 USA (e-mail: kyungjoo.noh@yale.edu).

V. V. Albert was with the Department of Physics, Yale University, New Haven, CT 06511 USA, and also with the Yale Quantum Institute, Yale University, New Haven, CT 06520 USA. He is now with the Institute for Quantum Information and Matter and the Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: valbert4@gmail.com).

L. Jiang is with the Department of Applied Physics and Physics, Yale University, New Haven, CT 06511 USA, and also with the Yale Quantum Institute, Yale University, New Haven, CT 06520 USA (e-mail: liang.jiang@yale.edu).

Communicated by M. M. Wilde, Associate Editor for Quantum Information Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2873764

In this paper, we prove that a family of GKP codes achieves the quantum capacity of Gaussian thermal 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 the quantum capacity of Gaussian thermal loss channels, 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 Eqs. (45),(47) for an improved upper bound of the Gaussian random displacement channel capacity). In section IV, we discuss experimental implementation of GKP codes and prove that the achievable quantum communication rate with 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 an energy-constrained encoding scenario and formulate a biconvex optimization problem to find an optimal encoding and decoding which maximize entanglement fidelity. We solve the biconvex optimization using an alternating semidefinite programming (SDP) method and show that a hexagonal GKP code emerges as an optimal solution. Then, we establish a lower bound for the achievable rate of the numerically optimized codes, based on an isotropic twirling and Werner state.

II. Preliminaries

In this section, we provide a summary of preliminary facts about Gaussian channels which will be referenced in later sections. For an introduction to bosonic modes, Gaussian states and Gaussian unitary operations, we refer the readers to [11] and [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 $\mathcal{D}(\mathcal{H}^{\otimes N}) \equiv \{\hat{\rho} \in \mathcal{L}(\mathcal{H}^{\otimes N}) \,|\, \hat{\rho}^\dagger = \hat{\rho} \succeq 0, \mathrm{Tr}[\hat{\rho}] = 1\}$ the set of density operators. A quantum channel $\mathcal{N} : \mathcal{D}(\mathcal{H}^{\otimes N}) \to \mathcal{D}(\mathcal{H}^{\otimes N})$ maps a quantum state $\hat{\rho} \in \mathcal{D}(\mathcal{H}^{\otimes N})$ to another state in $\mathcal{D}(\mathcal{H}^{\otimes N})$ via a completely positive and trace-preserving (CPTP) map [43]. Gaussian channels map a Gaussian state (see Eq. (82) for the definition) to another Gaussian state, and can be simulated by $\mathcal{N}(\hat{\rho}) = \mathrm{Tr}_E[\hat{U}_G(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}_G^\dagger]$, where $\hat{U}_G$ is a Gaussian unitary operation on system plus environmental modes, $\hat{\rho}_E$ is a Gaussian state, and Tr_E is the partial trace with respect to the environmental mode [44]. Let $\hat{\boldsymbol{X}}^{\boldsymbol{T}} = (\hat{\boldsymbol{x}}^{\boldsymbol{T}}, \hat{\boldsymbol{y}}^{\boldsymbol{T}})$ be a collection of quadrature operators of the system mode $\hat{\boldsymbol{x}}$ and the environmental mode $\hat{\boldsymbol{y}}$ (see the text below Eq. (78) for the definition of quadrature operators), and assume that initial system and environmental states are given by $\hat{\rho}_G(\bar{\boldsymbol{x}}, \boldsymbol{V_x})$ and $\hat{\rho}_G(\bar{\boldsymbol{y}}, \boldsymbol{V_y})$, respectively. Here, $\bar{\boldsymbol{x}}$, $\bar{\boldsymbol{y}}$ are the first moments and $\boldsymbol{V_x}$, $\boldsymbol{V_y}$ are the second moments of the system and environment–cf, Eq. (82). If the Gaussian unitary operation $\hat{U}_G$ acting on the joint system is characterized by

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S_{xx}} & \boldsymbol{S_{xy}} \\ \boldsymbol{S_{yx}} & \boldsymbol{S_{yy}} \end{pmatrix} \quad \text{and } \boldsymbol{D} = \begin{pmatrix} \boldsymbol{d_x} \\ \boldsymbol{d_y} \end{pmatrix}, \tag{1}$$

the first two moments of the system mode are transformed as $\bar{\boldsymbol{x}} \to \boldsymbol{S_{xx}}\bar{\boldsymbol{x}} + \boldsymbol{S_{xy}}\bar{\boldsymbol{y}} + \boldsymbol{d_x}$ and $\boldsymbol{V_x} \to \boldsymbol{S_{xx}V_xS_{xx}^T} + \boldsymbol{S_{xy}V_yS_{xy}^T}$, as can be derived by specializing Eq. (87) to Eq. (1). After tracing out the environment, the resulting effective Gaussian channel for the system is characterized by

$$\bar{\boldsymbol{x}} \to \boldsymbol{T}\bar{\boldsymbol{x}} + \boldsymbol{d}, \quad \boldsymbol{V_x} \to \boldsymbol{TV_xT^T} + \boldsymbol{N}, \tag{2}$$

where $\boldsymbol{T} = \boldsymbol{S_{xx}}$, $\boldsymbol{N} = \boldsymbol{S_{xy}V_yS_{xy}^T}$ and $\boldsymbol{d} = \boldsymbol{S_{xy}}\bar{\boldsymbol{y}} + \boldsymbol{d_x}$.

A major focus of this paper is the Gaussian thermal loss channel which models an excitation loss and gain, for example, in optical communication channels.

Definition 1 (Gaussian Thermal Loss Channel): Let $\hat{B}(\eta)$ be a beam splitter unitary with transmissivity $\eta \in [0, 1]$, acting on modes 1 and 2. Then, the Gaussian thermal loss channel is defined as $\mathcal{N}[\eta, \bar{n}_{\mathrm{th}}](\hat{\rho}_1) \equiv \mathrm{Tr}_2[\hat{B}(\eta)(\hat{\rho}_1 \otimes \hat{\rho}_{\bar{n}_{\mathrm{th}}})\hat{B}^\dagger(\eta)]$, where Tr_2 is the partial trace with respect to mode 2 which is initially in a thermal state $\hat{\rho}_{\bar{n}_{\mathrm{th}}}$ with an average photon number $\bar{n}_{\mathrm{th}}$. The Gaussian thermal loss channel is characterized by

$$\mathcal{N}[\eta, \bar{n}_{\mathrm{th}}] \leftrightarrow (\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{d}) = \left(\sqrt{\eta}\boldsymbol{I}_2, (1-\eta)\left(\bar{n}_{\mathrm{th}} + \frac{1}{2}\right)\boldsymbol{I}_2, 0\right), \tag{3}$$

where $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, as can be derived from the definition of beam splitter unitaries (see Eq. (89)).

The bosonic pure-loss channel is a special case of a Gaussian thermal loss channel with $\bar{n}_{\mathrm{th}} = 0$. The bosonic pure-loss channel with transmissivity $\eta \in [\frac{1}{2}, 1]$ ($\eta \in [0, \frac{1}{2}]$) is degradable (anti-degradable) [45], [46] and its quantum capacity has been evaluated. (See section III for the definitions of degradability and anti-degradability.) Except for this special case, Gaussian thermal loss channels are neither degradable nor anti-degradable [47], [48], and thus their coherent information may not be additive.

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

Definition 2 (Quantum-Limited Amplification Channel): Let $\hat{S}_2(G)$ be a two-mode squeezing unitary operation with gain $G \geq 1$, acting on modes 1 and 2. Then, the quantum-limited amplification channel is defined as $\mathcal{A}[G](\hat{\rho}_1) \equiv \mathrm{Tr}_2[\hat{S}_2(G)(\hat{\rho}_1 \otimes |0\rangle\langle 0|_2)\hat{S}_2^\dagger(G)]$, where $|0\rangle$ is the vacuum state. The quantum-limited amplification is characterized by

$$\mathcal{A}[G] \leftrightarrow (\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{d}) = \left(\sqrt{G}\boldsymbol{I}_2, \tfrac{(G-1)}{2}\boldsymbol{I}_2, 0\right), \tag{4}$$

as can be derived from the definition of two-mode squeezing unitaries (see Eq. (90)).

Note that the noise $\boldsymbol{N} = \frac{(G-1)}{2}\boldsymbol{I}_2$ is due to the variance of the ancillary vacuum state, transferred to the system via the two-mode squeezing operation. Since the vacuum state has the minimum variance allowed by the Heisenberg uncertainty

principle, the quantum-limited amplification incurs the least noise among all linear amplification channels [49].

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

Definition 3 (Gaussian Random Displacement Channel): The Gaussian random displacement channel is defined as

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

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

The Gaussian random displacement channel belongs to the class B_2 channel [47], [48] and is characterized by

$$\mathcal{N}_{B_2}[\sigma^2] \leftrightarrow (\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{d}) = (\boldsymbol{I}_2, \sigma^2\boldsymbol{I}_2, 0). \tag{6}$$

Since the convention for the Gaussian random displacement channel varies in the literature, we prove in Appendix B that $\boldsymbol{N} = \sigma^2\boldsymbol{I}_2$ holds in our convention, which is aligned with [28], [41], [42].

B. Synthesis and Decomposition of Gaussian Channels

In [42], it was shown that GKP codes outperform many other bosonic codes in protecting encoded quantum information against bosonic pure-loss errors upon the optimal decoding operation numerically obtained by semidefinite programming. This excellent performance of GKP codes can also be achieved by a sub-optimal decoding operation, which is designed based on the observation that a bosonic pure-loss channel can be converted into a Gaussian random displacement channel via quantum-limited amplification:

$$\mathcal{A}[1/\eta]\cdot\mathcal{N}[\eta, 0] = \mathcal{N}_{B_2}[\sigma^2_{\eta,0}], \tag{7}$$

where $\sigma^2_{\eta,0} \equiv \frac{1-\eta}{\eta}$ (see [42, eq. (7.21)]). This simple fact was previously noted in [10] and has been used in a key distribution scheme with GKP codes [51]. Eq. (7) implies that we can convert a bosonic pure-loss channel into a Gaussian random displacement channel and then use the conventional GKP decoding (see the text below Eq. (53)) to decode the encoded GKP states corrupted by the bosonic pure-loss channel. Here, we generalize this result and show that a general Gaussian thermal loss channel (with added thermal noise) can also be converted into a Gaussian random displacement channel. Before doing so, we state and prove the following known fact:

Theorem 4 (Gaussian Channel Synthesis; [47, eq. (26)]): Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be Gaussian channels with specification $(\boldsymbol{T}_1, \boldsymbol{N}_1, \boldsymbol{d}_1)$ and $(\boldsymbol{T}_2, \boldsymbol{N}_2, \boldsymbol{d}_2)$, respectively. Then, the synthesized channel $\mathcal{N} \equiv \mathcal{N}_2\cdot\mathcal{N}_1$ is a Gaussian channel with specification

$$\begin{aligned}\boldsymbol{T} &= \boldsymbol{T}_2\boldsymbol{T}_1,\\ \boldsymbol{N} &= \boldsymbol{T}_2\boldsymbol{N}_1\boldsymbol{T}_2^T + \boldsymbol{N}_2,\\ \boldsymbol{d} &= \boldsymbol{T}_2\boldsymbol{d}_1 + \boldsymbol{d}_2.\end{aligned} \tag{8}$$

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

$$\begin{aligned}\bar{\boldsymbol{x}}'' &= \boldsymbol{T}_2\bar{\boldsymbol{x}}' + \boldsymbol{d}_2,\\ &= \boldsymbol{T}_2\boldsymbol{T}_1\bar{\boldsymbol{x}} + \boldsymbol{T}_2\boldsymbol{d}_1 + \boldsymbol{d}_2,\\ \boldsymbol{V}'' &= \boldsymbol{T}_2\boldsymbol{V}'\boldsymbol{T}_2^T + \boldsymbol{N}_2\\ &= \boldsymbol{T}_2\boldsymbol{T}_1\boldsymbol{V}\boldsymbol{T}_1^T\boldsymbol{T}_2^T + \boldsymbol{T}_2\boldsymbol{N}_1\boldsymbol{T}_2^T + \boldsymbol{N}_2.\end{aligned} \tag{9}$$

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

Theorem 4 then leads to the following generalization of Eq. (7).

Lemma 5 (Loss + Amplification = Random Displacement): Let $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ be a general Gaussian thermal loss channel with a transmissivity $\eta \in [0, 1]$ and an average thermal photon number $\bar{n}_{\text{th}}$ in the ancillary mode. Let $\mathcal{A}[1/\eta]$ be the quantum-limited amplification with gain $G = 1/\eta$. Then,

$$\mathcal{A}[1/\eta]\cdot\mathcal{N}[\eta, \bar{n}_{\text{th}}] = \mathcal{N}_{B_2}[\sigma^2_{\eta,\bar{n}_{\text{th}}}], \tag{10}$$

where

$$\sigma^2_{\eta,\bar{n}_{\text{th}}} \equiv \Big(\frac{1-\eta}{\eta}\Big)(\bar{n}_{\text{th}} + 1). \tag{11}$$

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

$$\begin{aligned}\boldsymbol{N} &= \boldsymbol{T}_2\boldsymbol{N}_1\boldsymbol{T}_2^T + \boldsymbol{N}_2\\ &= \frac{1}{\eta}(1-\eta)\Big(\bar{n}_{\text{th}} + \frac{1}{2}\Big)\boldsymbol{I}_2 + \frac{1}{2}\Big(\frac{1-\eta}{\eta}\Big)\boldsymbol{I}_2\\ &= \Big(\frac{1-\eta}{\eta}\Big)(\bar{n}_{\text{th}} + 1)\boldsymbol{I}_2.\end{aligned} \tag{12}$$

Comparing Eq. (12) with $\boldsymbol{N} = \sigma^2\boldsymbol{I}_2$ for the Gaussian random displacement channel, the lemma follows. ∎

It is noteworthy that the noise from the Gaussian thermal loss channel (i.e., $\boldsymbol{N}_1 = (1-\eta)(\bar{n}_{\text{th}} + \frac{1}{2})\boldsymbol{I}_2$) is amplified by the quantum-limited amplification (the first term in the second line of Eq. (12)), hence increasing the variance $\sigma^2_{\eta,\bar{n}_{\text{th}}}$ of the resulting random displacement channel. We can however avoid this noise amplification simply by reversing the order of the thermal loss channel and the amplification, i.e., by amplifying signal prior to sending it through a Gaussian thermal loss channel.

Lemma 6 (Pre-Amplification Causes Less Noise): Let $\mathcal{N}[\eta, \bar{n}_{\text{th}}]$ and $\mathcal{A}[1/\eta]$ be as specified in corollary 5. Then,

$$\mathcal{N}[\eta, \bar{n}_{\text{th}}]\cdot\mathcal{A}[1/\eta] = \mathcal{N}_{B_2}[\tilde{\sigma}^2_{\eta,\bar{n}_{\text{th}}}], \tag{13}$$

where

$$\tilde{\sigma}^2_{\eta,\bar{n}_{\text{th}}} \equiv (1-\eta)(\bar{n}_{\text{th}} + 1). \tag{14}$$

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

TABLE I

CONVERSION OF A GAUSSIAN THERMAL LOSS CHANNEL INTO A GAUSSIAN RANDOM DISPLACEMENT CHANNEL (CHANNEL SYNTHESIS) AND DECOMPOSITION OF A GAUSSIAN THERMAL LOSS CHANNEL INTO A BOSONIC PURE-LOSS CHANNEL AND A QUANTUM-LIMITED AMPLIFICATION CHANNEL (CHANNEL DECOMPOSITION)

Channel synthesis	**Equation**	**Parameter**	**Application(s)**
Amp. + **Thermal loss** = Disp.	$\mathcal{A}[1/\eta]\cdot\boldsymbol{\mathcal{N}}[\boldsymbol{\eta},\bar{\boldsymbol{n}}_{\rm th}]=\mathcal{N}_{B_2}[\sigma^2_{\eta,\bar{n}_{\rm th}}]$ (See [42] and Lemma 5)	$\sigma^2_{\eta,\bar{n}_{\rm th}}=\left(\frac{1-\eta}{\eta}\right)(\bar{n}_{\rm th}+1)$	Eqs. (55),(59)
Thermal loss + Amp. = Disp.	$\boldsymbol{\mathcal{N}}[\boldsymbol{\eta},\bar{\boldsymbol{n}}_{\rm th}]\cdot\mathcal{A}[1/\eta]=\mathcal{N}_{B_2}[\tilde{\sigma}^2_{\eta,\bar{n}_{\rm th}}]$ (See Lemma 6)	$\tilde{\sigma}^2_{\eta,\bar{n}_{\rm th}}=(1-\eta)(\bar{n}_{\rm th}+1)$	Theorem 14
Channel decomposition	**Equation**	**Parameters**	**Application(s)**
Thermal loss = Amp. + Pure-loss	$\boldsymbol{\mathcal{N}}[\boldsymbol{\eta},\bar{\boldsymbol{n}}_{\rm th}]=\mathcal{A}[G']\cdot\mathcal{N}[\eta',0]$ (See [18], [47], [50] and Lemma 7)	$G'=(1-\eta)\bar{n}_{\rm th}+1=\frac{\eta}{\eta'}$	Eq. (32) (See [18])
Thermal loss = Pure-loss + Amp.	$\boldsymbol{\mathcal{N}}[\boldsymbol{\eta},\bar{\boldsymbol{n}}_{\rm th}]=\mathcal{N}[\tilde{\eta}',0]\cdot\mathcal{A}[\tilde{G}']$ (See [18], [19] and Lemma 8)	$\tilde{\eta}'=\eta-(1-\eta)\bar{n}_{\rm th}=\frac{\eta}{\tilde{G}'}$ (Not applicable if $\eta<\frac{\bar{n}_{\rm th}}{\bar{n}_{\rm th}+1}$)	Theorems 10,11 (See also [18], [19])

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

$$\begin{aligned}\boldsymbol{N}&=\boldsymbol{T}_1\boldsymbol{N}_2\boldsymbol{T}_1^T+\boldsymbol{N}_1\\&=\eta\times\frac{1}{2}\Big(\frac{1-\eta}{\eta}\Big)\boldsymbol{I}_2+(1-\eta)\Big(\bar{n}_{\rm th}+\frac{1}{2}\Big)\boldsymbol{I}_2\\&=(1-\eta)(\bar{n}_{\rm th}+1)\boldsymbol{I}_2.\end{aligned}\tag{15}$$

■

Note that in Eq. (14), the noise from the Gaussian thermal loss channel (i.e., $\boldsymbol{N}_1$) is not amplified. Moreover, the additional noise $\boldsymbol{N}_2$ from the quantum-limited amplification is now reduced by a factor of η as compared with Eq. (11), due to energy attenuation by the Gaussian thermal loss channel (see the first term in the second line of Eq. (15)).

In subsection IV-D, we combine Lemma 6 with an earlier result given in [41] to establish an achievable quantum communication rate of GKP codes for a general Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$.

Another interesting application of Gaussian channel synthesis is the composition of a bosonic pure-loss channel and a quantum-limited amplification with transmissivity and gain mismatched (i.e., $G\neq 1/\eta$). With this mismatch, one can simulate a general Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ with some $\mathcal{N}[\eta',0]$ and $\mathcal{A}[G']$ for some properly chosen parameters η', G' [47], [50]:

Lemma 7 ([18, eq. (5.1)]): A Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ can be decomposed into a bosonic pure-loss channel followed by a quantum-limited amplification

$$\mathcal{N}[\eta,\bar{n}_{\rm th}]=\mathcal{A}[G']\cdot\mathcal{N}[\eta',0],\tag{16}$$

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

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

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

$$\mathcal{N}[\eta,\bar{n}_{\rm th}]=\mathcal{N}[\tilde{\eta}',0]\cdot\mathcal{A}[\tilde{G}'],\tag{17}$$

where $\tilde{\eta}'=\eta-(1-\eta)\bar{n}_{\rm th}$ and $\tilde{G}'=\frac{\eta}{\eta-(1-\eta)\bar{n}_{\rm th}}$. Note that Eq. (17) is valid only if $\tilde{\eta}'\geq 0\leftrightarrow\eta\geq\bar{n}_{\rm th}/(\bar{n}_{\rm th}+1)$, or equivalently when $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ is not entanglement-breaking [52].

Proof: Recall that $\mathcal{A}[\tilde{G}']$ and $\mathcal{N}[\tilde{\eta}',0]$ are characterized by $(\boldsymbol{T}_1,\boldsymbol{N}_1,\boldsymbol{d}_1)=(\sqrt{\tilde{G}'}\boldsymbol{I}_2,\frac{(\tilde{G}'-1)}{2}\boldsymbol{I}_2,\boldsymbol{0})$ and $(\boldsymbol{T}_2,\boldsymbol{N}_2,\boldsymbol{d}_2)=(\sqrt{\tilde{\eta}'}\boldsymbol{I}_2,\frac{1}{2}(1-\tilde{\eta}')\boldsymbol{I}_2,\boldsymbol{0})$, respectively. Then, the synthesized channel is a Gaussian channel characterized by $\boldsymbol{T}=\boldsymbol{T}_2\boldsymbol{T}_1=\sqrt{\tilde{\eta}'\tilde{G}'}\boldsymbol{I}_2=\sqrt{\eta}\boldsymbol{I}_2$, $\boldsymbol{d}=\boldsymbol{T}_2\boldsymbol{d}_1+\boldsymbol{d}_2=\boldsymbol{0}$ and $\boldsymbol{N}=(1-\eta)(\bar{n}_{\rm th}+\frac{1}{2})\boldsymbol{I}_2$, i.e., the same specification as that of $\mathcal{N}[\eta,\bar{n}_{\rm th}]$. ■

Lemma 8 was independently discovered and stated in [18, Th. 31] and in [19, Lemma 1]. In Table I, we summarize all the Gaussian channel synthesis and decomposition properties shown in this subsection.

III. QUANTUM CAPACITY OF GAUSSIAN CHANNELS

A. Earlier Results on the 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 $\mathcal{N}(\hat{\rho})\equiv{\rm Tr}_E[\hat{U}(\hat{\rho}\otimes\hat{\rho}_E)\hat{U}^\dagger]$, where $\hat{U}$ is a unitary operator acting on $\mathcal{H}\otimes\mathcal{H}_E$, and $\mathcal{H}_E$ is the Hilbert space associated with an environment, causing the channel noise. The complementary channel $\mathcal{N}^c:\mathcal{L}(\mathcal{H})\rightarrow\mathcal{L}(\mathcal{H}_E)$ (defined with respect to the unitary operator $\hat{U}$) is the channel from the sender to the environment, i.e., $\mathcal{N}^c(\hat{\rho})\equiv{\rm Tr}_S[\hat{U}(\hat{\rho}\otimes\hat{\rho}_E)\hat{U}^\dagger]$. 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 $H(\hat{\rho})\equiv-{\rm Tr}[\hat{\rho}\log\hat{\rho}]$ is the entropy of a state $\hat{\rho}$ and log is the logarithm with base 2.

The quantum capacity of a channel $\mathcal{N}$ equals the regularized coherent information of the channel [3]–[5]:

$$Q_{\text{reg}}(\mathcal{N}) \equiv \lim_{n\to\infty} \frac{1}{n} Q(\mathcal{N}^{\otimes n}). \tag{18}$$

Coherent information might be strictly superadditive, i.e., $Q(\mathcal{N}\otimes\mathcal{N}') > Q(\mathcal{N}) + Q(\mathcal{N}')$, 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 if there exists a degrading channel $\mathcal{D} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_E)$ such that $\mathcal{N}^c = \mathcal{D}\cdot\mathcal{N}$, i.e., if the receiver can simulate a complementary channel $\mathcal{N}^c$. Coherent information of a degradable channel is additive (see, e.g., Theorem 12.5.4 in [1]), and thus the one-shot coherent information equals to the 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 $\mathcal{D} : \mathcal{L}(\mathcal{H}_E) \to \mathcal{L}(\mathcal{H})$ such that $\mathcal{N} = \mathcal{D}\cdot\mathcal{N}^c$, i.e., iff the environment can simulate the channel $\mathcal{N}$. Anti-degradable channels have zero quantum capacity and thus cannot support any reliable information transmission.

The complementary channel of a bosonic pure-loss channel with transmissivity η (with respect to the beam splitter unitary $\hat{B}(\eta)$) 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 $\eta\in[\frac{1}{2},1]$ [45], 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 $\eta\in[0,\frac{1}{2}]$ [46]).

Note that evaluation of a channel's coherent information involves optimization over all input states $\hat{\rho}$. For degradable Gaussian channels, optimization over all Gaussian states is sufficient [14]. For Gaussian thermal loss channels, optimization of the one-shot coherent information over all thermal states was performed in [13]. Related to [16, Remark 17], however, sufficiency of the thermal optimizers could be more clearly 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 $\hat{U}(\theta) \equiv e^{i\theta\hat{n}}$ is a phase rotation operator. Bosonic pure-loss channels are invariant under such a 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 $(\boldsymbol{T},\boldsymbol{N},\boldsymbol{d}) = (\boldsymbol{R}(\theta),\boldsymbol{0},\boldsymbol{0})$ and $\boldsymbol{R}(\theta)$ is given in Eq. (88). We then prove the following theorem.

Theorem 9 (Sufficiency of Thermal Optimizers for the Quantum Capacity of Bosonic Pure-Loss Channels): 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,

$$I_c(\hat{\rho},\mathcal{N}[\eta,0]) \leq I_c(\sum_{n=0}^{\infty}\rho_{nn}|n\rangle\langle n|,\mathcal{N}[\eta,0]). \tag{19}$$

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

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 function defined over $\theta\in[0,2\pi)$. Since the bosonic pure-loss channel is degradable, its coherent information $I_c(\hat{\rho},\mathcal{N}[\eta,0])$ is concave in the input state (see [1, Th. 12.5.6]):

$$\int_0^{2\pi} d\theta p(\theta) I_c(\hat{\rho}_\theta,\mathcal{N}[\eta,0]) \leq I_c\Big(\int_0^{2\pi} d\theta p(\theta)\hat{\rho}_\theta,\mathcal{N}[\eta,0]\Big). \tag{20}$$

The rotational invariance of bosonic pure-loss channels 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 quantum entropy is invariant under a unitary transformation (i.e., $H(\hat{\rho}) = H(\hat{U}\hat{\rho}\hat{U}^\dagger)$ for a unitary $\hat{U}$), 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. (20) 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

$$\begin{aligned}\int_0^{2\pi} d\theta p(\theta)\hat{\rho}_\theta &= \sum_{m,n=0}^{\infty}\frac{1}{2\pi}\int_0^{2\pi} d\theta e^{i\theta(m-n)}\rho_{mn}|m\rangle\langle n| \\ &= \sum_{n=0}^{\infty}\rho_{nn}|n\rangle\langle n|,\end{aligned} \tag{21}$$

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. (21) into the right hand side of Eq. (20), Eq. (19) 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 the energy-constrained scenario, where the average photon number is bounded from above by $\bar{n}$. ■

Note that a thermal state $\hat{\rho}_{\bar{n}}$ with average photon number $\bar{n}$ is transformed by a 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

$$g(x) \equiv (x+1)\log(x+1) - x\log x \tag{22}$$

(see Eq. (85) and the text below in Appendix A), Theorem 9 leads to the following energy-constrained quantum capacity of the bosonic pure-loss channel: ([16, eq. (1)])

$$\begin{aligned}Q_{\text{reg}}^{n\leq\bar{n}}(\mathcal{N}[\eta,0]) &= Q^{n\leq\bar{n}}(\mathcal{N}[\eta,0]) \\ &= \max[g(\eta\bar{n}) - g((1-\eta)\bar{n}),0],\end{aligned} \tag{23}$$

where $\bar{n}$ is the maximum allowed average photon number per bosonic mode and should not be confused with $\bar{n}_{\text{th}}$ in $\mathcal{N}[\eta,\bar{n}_{\text{th}}]$. When deriving the last equality of Eq. (23), we used the fact that $g(\eta\bar{n}) - g((1-\eta)\bar{n})$ increases monotonically in $\bar{n}$ (see [18, Remark 22]) and thus the optimal input state (with an 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 $\bar{n}\to\infty$, Eq. (23) reduces to

$$Q_{\text{reg}}(\mathcal{N}[\eta,0]) = Q(\mathcal{N}[\eta,0]) = \max\Big[\log\Big(\frac{\eta}{1-\eta}\Big),0\Big]. \tag{24}$$

TABLE II

LOWER AND UPPER BOUNDS OF THE QUANTUM CAPACITY OF BOSONIC PURE-LOSS CHANNELS, GAUSSIAN THERMAL LOSS CHANNELS AND GAUSSIAN RANDOM DISPLACEMENT CHANNELS, WITH AND WITHOUT AN ENERGY-CONSTRAINT $n \leq \bar{n}$

Quantum capacity	Pure-loss channel $\mathcal{N}[\eta, 0]$	Thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\rm th}]$	Displacement channel $\mathcal{N}_{B_2}[\sigma^2]$
Lower bound	$\max\left[\log\left(\frac{\eta}{1-\eta}\right), 0\right]$	$\max\left[\log\left(\frac{\eta}{1-\eta}\right) - g(\bar{n}_{\rm th}), 0\right]$	$\max\left[\log\left(\frac{1}{e\sigma^2}\right), 0\right]$
Proof	[13]	[13]	[13]
Energy-constrained	$\max\left[g(\eta\bar{n}) - g((1-\eta)\bar{n}), 0\right]$	$\max\left[I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\rm th}]), 0\right]$	$\max\left[I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}_{B_2}[\sigma^2]), 0\right]$
Proof	[13], [16]	[13], [16] (see Eq. (25))	[13], [16] (see Eq. (30))
Upper bound	$\max\left[\log\left(\frac{\eta}{1-\eta}\right), 0\right]$	$\boldsymbol{\max\left[\log\left(\frac{\eta-(1-\eta)\bar{n}_{\rm th}}{(1-\eta)(\bar{n}_{\rm th}+1)}\right), 0\right]}$	$\boldsymbol{\max\left[\log\left(\frac{1-\sigma^2}{\sigma^2}\right), 0\right]}$
Proof	[13], [46], [14], [15] and Theorem 9	**[18], [19] and Theorem 10**	**Eqs.** (28),(45)
Energy-constrained	$\max\left[g(\eta\bar{n}) - g((1-\eta)\bar{n}), 0\right]$	$\boldsymbol{Q_{\rm ODP}^{n\leq\bar{n}}(\eta, \bar{n}_{\rm th})}$ **in Theorem 12**	$\boldsymbol{\lim_{\eta\to 1} Q_{\rm ODP}^{n\leq\bar{n}}(\eta, \frac{\sigma^2}{1-\eta} - \frac{1}{2})}$
Proof	[13], [46], [14], [15], [16] and Theorem 9	**[18] and Theorem 12**	**Eqs.** (28),(47)

A general Gaussian thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\rm th}]$ (see Definition 1) with $\bar{n}_{\rm th} \neq 0$ is neither degradable nor anti-degradable [47], [48], and thus its coherent information may not be additive. A lower bound of the quantum capacity of Gaussian thermal loss channels can be obtained by evaluating the one-shot coherent information with a thermal input state [13]. In the energy-constrained case ($n \leq \bar{n}$),

$$\begin{aligned} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\rm th}]) = {} & g(\eta\bar{n} + (1-\eta)\bar{n}_{\rm th}) \\ & - g\Big(\frac{D + (1-\eta)(\bar{n} - \bar{n}_{\rm th}) - 1}{2}\Big) \\ & - g\Big(\frac{D - (1-\eta)(\bar{n} - \bar{n}_{\rm th}) - 1}{2}\Big), \end{aligned} \tag{25}$$

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

$$\lim_{\bar{n}\to\infty} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\rm th}]) = \log\Big(\frac{\eta}{1-\eta}\Big) - g(\bar{n}_{\rm th}). \tag{26}$$

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

$$\begin{aligned} Q_{\rm reg}^{n\leq\bar{n}}(\mathcal{N}[\eta, \bar{n}_{\rm th}]) &\geq \max[I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\rm th}]), 0], \\ Q_{\rm reg}(\mathcal{N}[\eta, \bar{n}_{\rm th}]) &\geq \max\Big[\lim_{\bar{n}\to\infty} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}[\eta, \bar{n}_{\rm th}]), 0\Big], \end{aligned} \tag{27}$$

in the energy-constrained and unconstrained cases, respectively.

A Gaussian random displacement channel can be understood as a certain limit of a Gaussian thermal loss channel [53]:

$$\mathcal{N}_{B_2}[\sigma^2] = \lim_{\eta\to 1} \mathcal{N}\Big[\eta, \frac{\sigma^2}{1-\eta} - \frac{1}{2}\Big]. \tag{28}$$

Thus, a lower bound of the quantum capacity of a Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ is given by

$$\begin{aligned} Q_{\rm reg}^{n\leq\bar{n}}(\mathcal{N}_{B_2}[\sigma^2]) &\geq \max[I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}_{B_2}[\sigma^2]), 0], \\ Q_{\rm reg}(\mathcal{N}_{B_2}[\sigma^2]) &\geq \max\Big[\log\Big(\frac{1}{e\sigma^2}\Big), 0\Big], \end{aligned} \tag{29}$$

in the energy-constrained and unconstrained cases, respectively, where (given $D' \equiv \sqrt{(2\bar{n} + \sigma^2 + 1)^2 - 4\bar{n}(\bar{n}+1)}$)

$$\begin{aligned} I_c(\hat{\rho}_{\bar{n}}, \mathcal{N}_{B_2}[\sigma^2]) &= \lim_{\eta\to 1} I_c\Big(\hat{\rho}_{\bar{n}}, \mathcal{N}\Big[\eta, \frac{\sigma^2}{1-\eta} - \frac{1}{2}\Big]\Big) \\ &= g(\bar{n}) - g\Big(\frac{D' + \sigma^2 - 1}{2}\Big) \\ &\quad - g\Big(\frac{D' - \sigma^2 - 1}{2}\Big), \end{aligned} \tag{30}$$

and we used $g(x) = \log[ex] + \mathcal{O}(1/x)$ to get the second line of Eq. (29).

A general upper bound of the quantum capacity was introduced in [13] and applied to bound the quantum capacity of a Gaussian thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\rm th}]$ from above: ([13, eq. (5.7)], translated into our notation)

$$\begin{aligned} Q_{\rm reg}(\mathcal{N}[\eta, \bar{n}_{\rm th}]) &\leq Q_{\rm HW}(\eta, \bar{n}_{\rm th}) \\ &\equiv \max\Big[\log\Big(\frac{1+\eta}{(1-\eta)(2\bar{n}_{\rm th}+1)}\Big), 0\Big]. \end{aligned} \tag{31}$$

However, this upper bound is not tight: In the case of bosonic pure-loss channels, $Q_{\rm HW}(\eta, 0) = \log(\frac{1+\eta}{1-\eta}) > Q(\mathcal{N}[\eta, 0])$ for all $\eta \in [0, 1)$, where $Q(\mathcal{N}[\eta, 0])$ is given in Eq. (24). Recently, three new upper bounds were obtained in [18], where one of them is based on Lemma 7 and a data-processing argument, and the other two are based on the notion of approximate degradability of a quantum channel developed in [54]. 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: ([18, Ths. 20 and 25])

$$Q_{\rm reg}^{n\leq\bar{n}}(\mathcal{N}[\eta, \bar{n}_{\rm th}]) \leq Q_{\rm DP}^{n\leq\bar{n}}(\eta, \bar{n}_{\rm th}) \equiv Q^{n\leq\bar{n}}(\mathcal{N}[\eta', 0]), \tag{32}$$

where $Q_{\rm DP}^{n\leq\bar{n}}(\eta, \bar{n}_{\rm th})$ is the data-processing bound, $\eta' = \frac{\eta}{(1-\eta)\bar{n}_{\rm th}+1}$ and $Q^{n\leq\bar{n}}(\mathcal{N}[\eta', 0])$ is given in Eq. (23). *Proof:* [Proof of Eq. (32)] 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 a communication rate $R = Q_{\rm reg}^{n\le\bar{n}}(\mathcal{N}[\eta,\bar{n}_{\rm th}])$ for the Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$. Since $\mathcal{N}[\eta,\bar{n}_{\rm th}] = \mathcal{A}[G']\cdot\mathcal{N}[\eta',0]$ (see Lemma 7), this implies that the encoding and decoding set $\{\mathcal{E},\mathcal{D}\cdot\mathcal{A}[G']\}$ achieves a rate $R = Q_{\rm reg}^{n\le\bar{n}}(\mathcal{N}[\eta,\bar{n}_{\rm th}])$ for the bosonic pure-loss channel $\mathcal{N}[\eta',0]$. Since an achievable rate R is upper bounded by the quantum capacity $Q^{n\le\bar{n}}(\mathcal{N}[\eta',0])$, Eq. (32) follows. ∎

Note that $Q_{\rm DP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th})$ converges to $Q^{n\le\bar{n}}(\mathcal{N}[\eta,0])$ in $\bar{n}_{\rm th}\to 0$ limit, as $\eta'=\eta$ when $\bar{n}_{\rm th}=0$. Similarly, an upper bound of the quantum capacity of a Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ can be obtained by combining the data-processing argument with Eq. (7):

$$Q_{\rm reg}(\mathcal{N}_{B_2}[\sigma^2]) \le \max\Big[\log\Big(\frac{1}{\sigma^2}\Big),0\Big]. \tag{33}$$

Alternatively, this bound can also be derived from Eq. (31) by setting $\bar{n}_{\rm th}=\frac{\sigma^2}{1-\eta}-\frac{1}{2}$ and then taking the limit $\eta\to 1$.

B. Improved Upper Bound of Gaussian Loss Channel Capacity

Here, we improve the data-processing bound $Q_{\rm DP}(\eta,\bar{n}_{\rm th})$ for some parameter regimes by using Lemma 8, instead of Lemma 7, for the decomposition of Gaussian thermal loss channels, producing an improved bound $Q_{\rm IDP}(\eta,\bar{n}_{\rm th})$. We then optimize the improved upper bound further to produce an optimized upper bound $Q_{\rm ODP}(\eta,\bar{n}_{\rm th})$.

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

$$Q_{\rm reg}(\mathcal{N}[\eta,\bar{n}_{\rm th}]) \le Q_{\rm IDP}(\eta,\bar{n}_{\rm th}) \equiv Q_{\rm reg}(\mathcal{N}[\tilde{\eta}',0]), \tag{34}$$

where $\tilde{\eta}'=\eta-(1-\eta)\bar{n}_{\rm th}$ and $Q_{\rm reg}(\mathcal{N}[\tilde{\eta}',0])$ is given in Eq. (24). The improved data-processing bound $Q_{\rm IDP}(\eta,\bar{n}_{\rm th})$ simplifies to

$$Q_{\rm IDP}(\eta,\bar{n}_{\rm th}) = \max\Big[\log\Big(\frac{\eta-(1-\eta)\bar{n}_{\rm th}}{(1-\eta)(\bar{n}_{\rm th}+1)}\Big),0\Big]. \tag{35}$$

Proof: The proof goes in the same way as in the proof of Eq.(32), except that Lemma 7 is replaced by Lemma 8: $\mathcal{N}[\eta,\bar{n}_{\rm th}] = \mathcal{N}[\tilde{\eta}',0]\cdot\mathcal{A}[\tilde{G}']$ with $\tilde{G}'=\eta/(\eta-(1-\eta)\bar{n}_{\rm th}) = \eta/\tilde{\eta}'$. Then, the encoding and decoding set $\{\mathcal{A}[\tilde{G}']\cdot\mathcal{E},\mathcal{D}\}$ achieves a rate $R = Q_{\rm reg}(\mathcal{N}[\eta,\bar{n}_{\rm th}])$ for the bosonic pure-loss channel $\mathcal{N}[\tilde{\eta}',0]$. Since $R\le Q(\mathcal{N}[\tilde{\eta}',0])$, the theorem follows. ∎

Note that

$$Q_{\rm IDP}(\eta,\bar{n}_{\rm th}) < Q_{\rm DP}(\eta,\bar{n}_{\rm th}) \equiv \lim_{\bar{n}\to\infty} Q_{\rm DP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th}) \tag{36}$$

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

$$\eta' = \frac{\eta}{(1-\eta)\bar{n}_{\rm th}+1} = \eta-\frac{\eta(1-\eta)\bar{n}_{\rm th}}{(1-\eta)\bar{n}_{\rm th}+1} > \eta-(1-\eta)\bar{n}_{\rm th} = \tilde{\eta}'. \tag{37}$$

Thus, $Q_{\rm IDP}(\eta,\bar{n}_{\rm th})$ is a strictly tighter upper bound of the Gaussian thermal loss channel capacity than $Q_{\rm DP}(\eta,\bar{n}_{\rm th})$. 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 a data-processing argument, because the quantum-limited amplification in $\mathcal{E}'=\mathcal{A}[\tilde{G}']\cdot\mathcal{E}$ increases the energy of an encoded state:

$$\bar{n}' = {\rm Tr}[\mathcal{A}[\tilde{G}'](\hat{\rho})\hat{n}] = \tilde{G}'\bar{n}+(\tilde{G}'-1), \tag{38}$$

as can be derived from the symplectic transformation of a two-mode squeezing (Eq. (90) and the text above in Appendix A).

Theorem 11 (Improved Data-Processing Bound With an Energy Constraint): The energy-constrained quantum capacity of a Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ is upper bounded by

$$Q_{\rm reg}^{n\le\bar{n}}(\mathcal{N}[\eta,\bar{n}_{\rm th}]) \le Q_{\rm IDP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th}) \equiv Q^{n\le\tilde{G}'\bar{n}+(\tilde{G}'-1)}(\mathcal{N}[\tilde{\eta}',0]), \tag{39}$$

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

$$Q_{\rm IDP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th}) = \max\Big[g(\eta\bar{n}+(1-\eta)\bar{n}_{\rm th}) - g\Big(\frac{(1-\eta)(\bar{n}_{\rm th}+1)(\eta\bar{n}+(1-\eta)\bar{n}_{\rm th})}{\eta-(1-\eta)\bar{n}_{\rm th}}\Big),0\Big]. \tag{40}$$

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

Theorem 11 was independently discovered in [18] (see Theorem 46 therein). We emphasize that, unlike the energy-unconstrained case, our new upper bound $Q_{\rm IDP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th})$ is not always tighter than $Q_{\rm DP}^{n\le\bar{n}}(\eta,\bar{n}_{\rm th})$ (see the inset of Fig. 1). Physically, this is because the increased encoding energy due to the pre-amplification allows a 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 a Gaussian thermal loss channel, including both a pre-amplification and a post-amplification:

$$\mathcal{N}[\eta,\bar{n}_{\rm th}] = \mathcal{A}[G_1]\mathcal{N}[\bar{\eta},0]\mathcal{A}[G_2], \tag{41}$$

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

$$Q_{\rm reg}^{n\le\bar{n}}(\mathcal{N}[\eta,\bar{n}_{\rm th}]) \le \min_{G_1,G_2\ge 1} Q^{n\le G_2\bar{n}+(G_2-1)}(\mathcal{N}[\bar{\eta},0]). \tag{42}$$

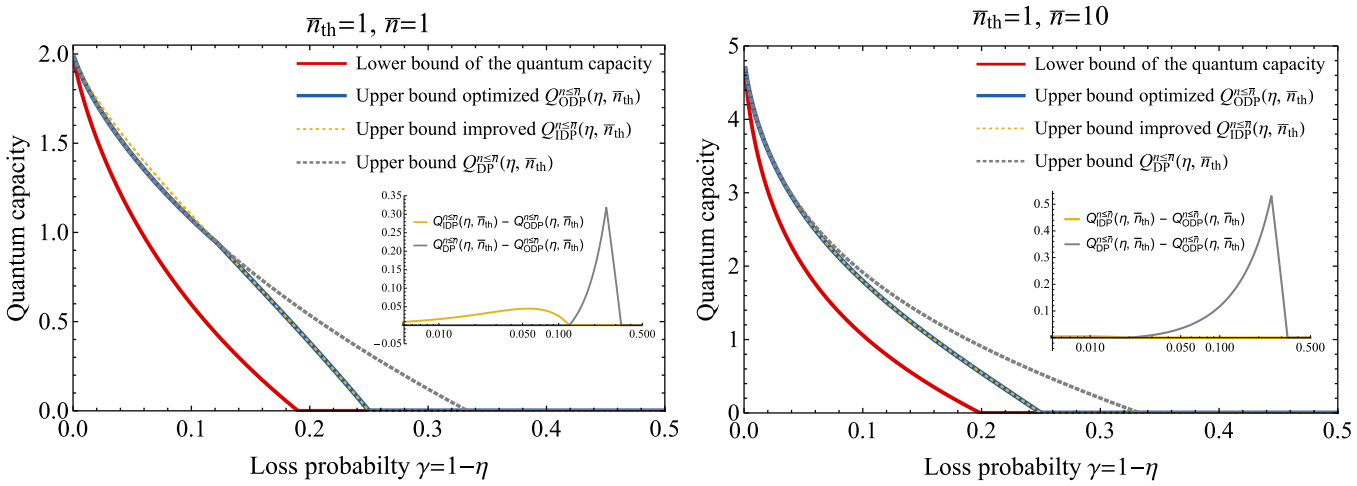

Fig. 1. Bounds of the quantum capacity of a Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ for $(\bar{n}_{\rm th},\bar{n})=(1,1)$ (left) and $(\bar{n}_{\rm th},\bar{n})=(1,10)$ (right). The lower bound (red) is obtained by evaluating the coherent information with a thermal input state (see Eq. (27)). The improved data-processing bound $Q^{n\le\bar{n}}_{\rm IDP}(\eta,\bar{n}_{\rm th})$ (dashed yellow) is identical to the optimized data-processing bound $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})$ (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 $Q^{n\le\bar{n}}_{\rm DP}(\eta,\bar{n}_{\rm th})$ (dashed grey) is optimal when $\eta\ge\eta^\star(\bar{n}_{\rm th},\bar{n})$ for some $\eta^\star(\bar{n}_{\rm th},\bar{n})$ (e.g., $\eta^\star(1,1)=0.8775\cdots$).

Theorem 12 (Optimized Data-Processing Bound With an Energy-Constraint): Energy-constrained quantum capacity of a Gaussian thermal loss channel $\mathcal{N}[\eta,\bar{n}_{\rm th}]$ is upper bounded by the optimized data-processing bound $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})$:

$$\begin{aligned}Q^{n\le\bar{n}}_{\rm reg}(\mathcal{N}[\eta,\bar{n}_{\rm th}]) &\le Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})\\ &\equiv \min_{1\le G_1\le 1+(1-\eta)\bar{n}_{\rm th}} f_{\eta,\bar{n}_{\rm th},\bar{n}}(G_1),\end{aligned}\tag{43}$$

where $\bar{n}$ is the maximum allowed average photon number in the encoding and $f_{\eta,\bar{n}_{\rm th},\bar{n}}(G_1)$ is defined as

$$\begin{aligned}f_{\eta,\bar{n}_{\rm th},\bar{n}}(G_1)\equiv\max\Big[&g\big(\bar{\eta}(G_2\bar{n}+(G_2-1))\big)\\ &-g\big((1-\bar{\eta})(G_2\bar{n}+(G_2-1))\big),0\Big].\end{aligned}\tag{44}$$

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

Proof: See appendix C for the derivation. ∎

Since $Q^{n\le\bar{n}}_{\rm DP}(\eta,\bar{n}_{\rm th})$ and $Q^{n\le\bar{n}}_{\rm IDP}(\eta,\bar{n}_{\rm th})$ 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 $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})$ is the best upper bound among all data-processing type bounds introduced above (see Fig. 1). We numerically observe, however, that either one of these two extremes is optimal: $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})=Q^{n\le\bar{n}}_{\rm DP}(\eta,\bar{n}_{\rm th})$ for $\eta\ge\eta^\star(\bar{n}_{\rm th},n)$ and $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})=Q^{n\le\bar{n}}_{\rm IDP}(\eta,\bar{n}_{\rm th})$ otherwise for some $\eta^\star(\bar{n}_{\rm th},n)$. In $\bar{n}\to\infty$ limit, $\lim_{\bar{n}\to\infty}\eta^\star(\bar{n}_{\rm th},\bar{n})=1$ and thus $Q^{n\le\bar{n}}_{\rm ODP}(\eta,\bar{n}_{\rm th})=Q^{n\le\bar{n}}_{\rm IDP}(\eta,\bar{n}_{\rm th})$ for all $\eta\in[0,1]$, which is consistent with Eq. (36). We emphasize that approximate degradability bounds [18], [54] are tighter than the data-processing bounds in the high transmissivity and high thermal noise regime and refer to [18] (e.g., Fig. 6 therein) for a more comprehensive comparison of the existing upper bounds.

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

$$Q_{\rm reg}(\mathcal{N}_{B_2}[\sigma^2])\le\max\Big[\log\Big(\frac{1-\sigma^2}{\sigma^2}\Big),0\Big].\tag{45}$$

(Alternatively, this bound can also be obtained by evaluating $\lim_{\eta\to1}Q_{\rm IDP}(\eta,\frac{\sigma^2}{1-\eta}-\frac{1}{2})$.) Note that this bound is strictly tighter than the one in Eq. (33). In the energy-constrained case, the upper bound of the quantum capacity of the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ is given by

$$Q^{n\le\bar{n}}_{\rm reg}(\mathcal{N}_{B_2}[\sigma^2])\le Q^{n\le\bar{n}}_{\rm ODP}(\mathcal{N}_{B_2}[\sigma^2]),\tag{46}$$

where

$$\begin{aligned}Q^{n\le\bar{n}}_{\rm ODP}(\mathcal{N}_{B_2}[\sigma^2])&\equiv\lim_{\eta\to1}Q^{n\le\bar{n}}_{\rm ODP}\Big(\mathcal{N}\Big[\eta,\frac{\sigma^2}{1-\eta}-\frac{1}{2}\Big]\Big)\\ &=\min_{1\le G_1\le1+\sigma^2}\tilde{f}_{\sigma^2,\bar{n}}(G_1),\end{aligned}\tag{47}$$

and $\tilde{f}_{\sigma^2,\bar{n}}\equiv\lim_{\eta\to1}f_{\eta,\frac{\sigma^2}{1-\eta}-\frac{1}{2},\bar{n}}(G_1)$ is given by

$$\begin{aligned}\tilde{f}_{\sigma^2,\bar{n}}(G_1)=\max\Big[&g\Big(\frac{\bar{n}+\sigma^2-G_1+1}{G_1}\Big)\\ &-g\Big(\frac{\sigma^2(\bar{n}+\sigma^2-G_1+1)}{G_1(G_1-\sigma^2)}\Big),0\Big].\end{aligned}\tag{48}$$

IV. Gottesman-Kitaev-Preskill Codes

We now aim to find an explicit encoding and decoding strategy that achieves the quantum capacity of Gaussian thermal loss channels. It is very important to realize that the coherent information of bosonic pure-loss channels 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 a bosonic pure-loss channel. In fact, it was proven that entanglement distillation of a Gaussian state with only Gaussian operations is impossible [55]. Due to

the close relation between entanglement distillation and quantum state transmission via quantum teleportation [56], [57], this implies that Gaussian encoding and decoding can never achieve a non-vanishing quantum communication rate for a Gaussian channel. In other words, quantum error correction of a Gaussian error is impossible if the states and operations are restricted to be Gaussian [58]. Therefore, non-Gaussian resources are necessary in order to achieve a reliable quantum information transmission through a noisy Gaussian channel.

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 a non-vanishing quantum communication rate through a Gaussian thermal loss channel. In particular, we prove that a family of GKP codes achieves the quantum capacity of Gaussian thermal 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: Introduction, Preparation, and Gates

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

$$\begin{aligned} \hat{S}_{{\rm sq},q}^{[d]} &\equiv \exp(i\hat{q}\sqrt{2\pi d}) = \hat{D}(i\sqrt{\pi d}), \\ \hat{S}_{{\rm sq},p}^{[d]} &\equiv \exp(-i\hat{p}\sqrt{2\pi d}) = \hat{D}(\sqrt{\pi d}), \end{aligned} \tag{49}$$

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

The square lattice GKP code states are explicitly given by

$$|\mu_L^{\rm sq}\rangle \propto \sum_{n=-\infty}^{\infty} |\hat{q} = (dn+\mu)\sqrt{2\pi/d}\rangle \tag{50}$$

in the computational basis, where $\mu = 0, \cdots, d-1$ 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. Note that the phase rotation by angle $\theta = -\pi/2$ implements a basis transformation from $|\mu_L^{\rm sq}\rangle$ to

$$|\bar{\mu}_L^{\rm sq}\rangle \propto \sum_{n=-\infty}^{\infty} |\hat{p} = (dn-\mu)\sqrt{2\pi/d}\rangle, \tag{51}$$

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

Since the GKP code states are superpositions of infinitely many infinitely squeezed states, they have infinite average photon number. It is possible, however, to construct a GKP state with a finite energy: Replace the infinitely squeezed states by finitely squeezed ones, and introduce an overall Gaussian envelope (see [28, eq. (31)]). Our brief new observation is that this can be concisely realized by

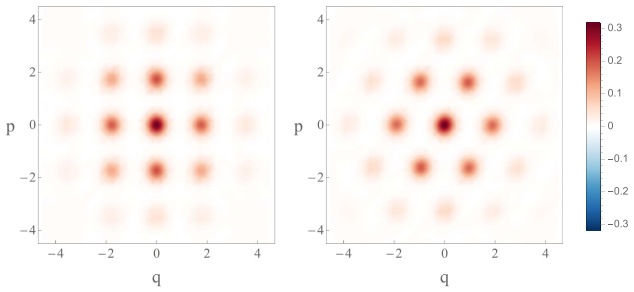

Fig. 2. Wigner function of the maximally mixed code states of the one-mode square lattice (left; $\mathcal{C}_{\rm sq}^{[d]}$) and hexagonal lattice (right; $\mathcal{C}_{\rm hex}^{[d]}$) GKP code with $d = 2$ and an average photon number $\bar{n} = 3$.

$|\mu_L^{\Delta}\rangle \propto \exp(-\Delta^2\hat{n})|\mu_L\rangle$ (which is an equivalent form [28, eqs. (40) and (41)]), where $\Delta \geq 0$. The non-unitary operation $\exp(-\Delta^2\hat{n})$ can be physically implemented by mixing an input state with an ancilla vacuum state by a beam splitter interaction, and then by counting the photon number of an output idler mode and post-selecting on no clicks. As a result, a large enough GKP state of any shape can be carved into a smaller one with a Gaussian envelope, where the size of the resulting GKP state (i.e., Δ) can be modulated by changing transmissivity $\bar{\eta}$ of the beam splitter interaction, i.e., $\bar{\eta} = e^{-2\Delta^2}$.

In Fig. 2 (a), we plot the Wigner function of the maximally mixed code state $\hat{\rho}_{\mathcal{C}_{\rm sq}^{[d]}} = \hat{P}_{\mathcal{C}_{\rm sq}^{[d]}}/d$ of the square lattice GKP code states for $d = 2$ and $\bar{n} = 3$, where $\hat{P}_{\mathcal{C}_{\rm sq}^{[d]}}$ is the projection operator to the code space $\mathcal{C}_{\rm sq}^{[d]}$. In the remainder of this section, we restrict ourselves to the properties of the energy-unconstrained ideal GKP states.

A universal gate set of the square lattice GKP code is given as follows: The logical Pauli gates (i.e., Clifford gates of hierarchy 1) can be implemented by $\hat{Z}_L^{\rm sq} = (\hat{S}_{{\rm sq},q}^{[d]})^{\frac{1}{d}} = \exp(i\hat{q}\sqrt{2\pi/d})$ and $\hat{X}_L^{\rm sq} = (\hat{S}_{{\rm sq},p}^{[d]})^{\frac{1}{d}} = \exp(-i\hat{p}\sqrt{2\pi/d})$, i.e., displacement operators, yielding $\hat{Z}_L^{\rm sq}|\mu_L^{\rm sq}\rangle = \exp(i2\pi\mu/d)|\mu_L^{\rm sq}\rangle$, $\hat{X}_L^{\rm sq}|\mu_L^{\rm sq}\rangle = |(\mu\oplus 1)_L^{\rm sq}\rangle$, where $\oplus$ is the summation in modulo d (see [59] for the notion of Clifford hierarchy). Also, any logical Clifford gate of hierarchy 2 can be implemented by a Gaussian unitary operation [28]. For example, the logical Hadamard gate can be implemented by a phase rotation $\exp(-i\pi\hat{a}^\dagger\hat{a}/2)$, as explained above. The logical CNOT gate can be realized by the SUM gate: $\exp(-i\hat{p}_1\hat{q}_2)|\mu_L^{\rm sq}\rangle|\nu_L^{\rm sq}\rangle = |(\mu\oplus\nu)_L^{\rm sq}\rangle|\nu_L^{\rm sq}\rangle$. In $d = 2$ case, we observe that $\exp(i\hat{q}^4/(4\pi))$ implements the logical T gate (a Clifford gate of hierarchy 3), i.e., $\exp(i\hat{q}^4/(4\pi))|\mu_L^{\rm sq}\rangle = \exp(i\pi\mu/4)|\mu_L^{\rm sq}\rangle$, completing the gate set for universal quantum computation.

Encoding of an arbitrary logical state 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., $|0_L^{\rm sq}\rangle$) is available. Also, assume the gate $\sum_{\nu=0}^{d-1}(\hat{X}_L^{\rm sq})^\nu \otimes |\nu\rangle\langle\nu|$, 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

$$\Big[\sum_{\nu=0}^{d-1}(\hat{X}_L^{\rm sq})^\nu\otimes|\nu\rangle\langle\nu|\Big](|0_L^{\rm sq}\rangle\otimes|\psi\rangle)=\sum_{\mu=0}^{d-1}c_\mu|\mu_L^{\rm sq}\rangle\otimes|\mu\rangle. \tag{52}$$

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

$$(\hat{Z}_L^{\rm sq})^\nu|\psi_L^{\rm sq}(\nu)\rangle=\sum_{\mu=0}^{d-1}c_\mu|\mu_L^{\rm sq}\rangle. \tag{53}$$

If the encoded GKP states are sent though a Gaussian random displacement channel, decoding can be achieved by measuring the two stabilizers $\hat{S}_{{\rm sq},q}^{[d]}$, $\hat{S}_{{\rm sq},p}^{[d]}$, providing protection against the channel noise as specified in the text below Eq. (49). Given that a fresh supply of ancillary GKP states is available, such a stabilizer measurement can be achieved by a Gaussian operation and homodyne detection: Assume that we have an ancillary oscillator mode prepared at $|\bar{0}_L^{\rm sq}\rangle$ (i.e., $q_2\equiv 0\ \mathrm{mod}\sqrt{2\pi/d}$). Upon performing 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 a 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., $q_1\equiv q_2^M\ \mathrm{mod}\sqrt{2\pi/d}$, hence measuring $\hat{S}_{{\rm sq},q}^{[d]}$. The other stabilizer $\hat{S}_{{\rm sq},p}^{[d]}$ can also be measured in a similar way using an initial ancillary GKP state $|0_L^{\rm sq}\rangle$ (i.e., $p_2\equiv 0\ \mathrm{mod}\sqrt{2\pi/d}$), a SUM gate with control and target exchanged $\exp(-i\hat{p}_1\hat{q}_2)$ and a homodyne measurement of $\hat{p}_2$ (see the text below [28, eq. (105)]). From the extracted values of the position and momentum operator in modulo $\sqrt{2\pi/d}$ (i.e., q_2^M, p_2^M), one can infer $\Delta q=q_2^{M\star}$, $\Delta p=p_2^{M\star}$, where $q_2^{M\star}$, $p_2^{M\star}$ are defined such that $q_2^M\equiv q_2^{M\star}\ \mathrm{mod}\sqrt{2\pi/d}$, $p_2^M\equiv p_2^{M\star}\ \mathrm{mod}\sqrt{2\pi/d}$ and $|q_2^{M\star}|,|p_2^{M\star}|\leq\sqrt{\pi/(2d)}$. One can then correct such errors by counter displacements $\exp(i\hat{p}_1q_2^{M\star})$ and $\exp(-i\hat{q}_1p_2^{M\star})$.

In both the encoding and decoding procedures, the most non-trivial task is to prepare a GKP state (e.g., $|0_L^{\rm sq}\rangle$) from an arbitrary non-GKP state. There have been many proposals to achieve such a goal in various experimental platforms [60]–[66], including one in the original paper [28]. In particular, the proposal in [64] is based on the idea of phase estimation of a unitary operator [67], which allows the following representation of the GKP states (equivalent to [28, eq. (29)])

$$\begin{aligned}|\mu_L^{\rm sq}\rangle&\propto(\hat{X}_L^{\rm sq})^\mu\sum_{n_1,n_2=-\infty}^{\infty}(\hat{S}_{{\rm sq},p}^{[d]})^{n_1}(\hat{Z}_L^{\rm sq})^{n_2}|\phi_0\rangle\\&=\sum_{n_1,n_2=-\infty}^{\infty}e^{-i\hat{p}\sqrt{\frac{2\pi}{d}}(dn_1+\mu)}e^{i\hat{q}\sqrt{\frac{2\pi}{d}}n_2}|\phi_0\rangle.\end{aligned} \tag{54}$$

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 (an unnormalized) 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 $|0_L^{\rm sq}\rangle$. The last operation $(\hat{X}_L^{\rm sq})^\mu$ then transforms $|0_L^{\rm sq}\rangle$ into $|\mu_L^{\rm sq}\rangle$. Note that Eq. (54) is valid for any input state $|\phi_0\rangle$ with a non-zero overlap with $|0_L^{\rm sq}\rangle$ which, for example, can be chosen to be the vacuum state. In this case, Eq. (54) can be understood as a superposition of coherent states $|\alpha\rangle$ on the 2-dimensional square lattice sites $\alpha=\sqrt{\frac{\pi}{d}}(dn_1+\mu+in_2)$, where n_1,n_2 are integers.

We remark that the scheme in [60] implements Eq. (54) by a post-selection, whereas the scheme in [64] does so deterministically. The former protocol was recently realized experimentally in a trapped ion system [68], [69]. In principle, the latter protocol can be realized in a circuit quantum electrodynamics system by, e.g., extending the quantum non-demolition measurement technique used in [70].

Based on a numerically optimized decoding operation, we earlier showed that GKP codes offer an excellent protection against a bosonic pure-loss channel $\mathcal{N}[\eta,0]$, despite not being specifically designed for such a purpose [42]. In addition, by using a sub-optimal decoding (i.e., a 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

$$\epsilon_L^{\rm sq}\sim\exp\Big[-\frac{\pi}{4d}\Big(\frac{\eta}{1-\eta}\Big)\Big], \tag{55}$$

(see [42, eq. (7.24)]) 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 a hexagonal lattice to define the GKP code states, instead of a square lattice.

B. One-Mode Hexagonal Lattice GKP Code

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

$$\begin{aligned}\hat{S}_{{\rm hex},q}^{[d]}&\equiv\exp[i\sqrt{2\pi d}(S_{11}\hat{q}+S_{21}\hat{p})],\\\hat{S}_{{\rm hex},p}^{[d]}&\equiv\exp[-i\sqrt{2\pi d}(S_{12}\hat{q}+S_{22}\hat{p})],\end{aligned} \tag{56}$$

and S_{ij} are the matrix elements of a symplectic matrix

$$\boldsymbol{S}_{\rm hex}=\begin{pmatrix}S_{11}&S_{12}\\S_{21}&S_{22}\end{pmatrix}=\Big(\frac{2}{\sqrt{3}}\Big)^{\frac{1}{2}}\begin{pmatrix}1&\frac{1}{2}\\0&\frac{\sqrt{3}}{2}\end{pmatrix}. \tag{57}$$

Note that the stabilizers of the square lattice GKP code are given by the same form as in Eq. (56) but with $\boldsymbol{S}_{\rm sq}$ the 2×2 identity matrix. Thus, a hexagonal lattice GKP code state can be generated by applying a Gaussian unitary operation (with a corresponding symplectic transformation $\boldsymbol{S}_{\rm hex}^{-1}$) to a square lattice GKP code state [28], [41]. Similarly as in the case of the square lattice GKP code, the logical Pauli operators are given by $\hat{Z}_L^{\rm hex}=(\hat{S}_{{\rm hex},q}^{[d]})^{\frac{1}{d}}$ and $\hat{X}_L^{\rm hex}=(\hat{S}_{{\rm hex},p}^{[d]})^{\frac{1}{d}}$. Following the same reasoning to obtain Eq. (54), we provide an explicit

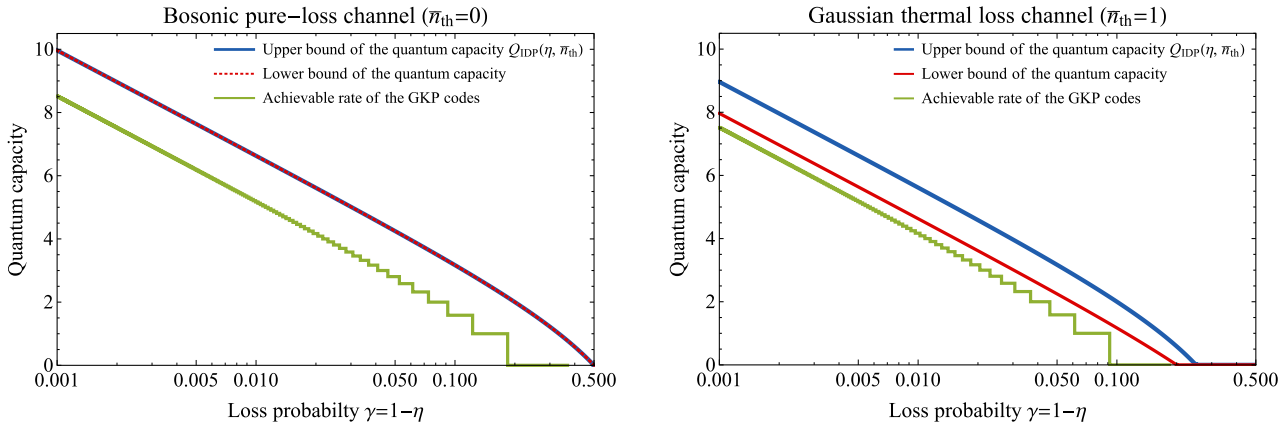

Fig. 3. Achievable quantum communication rate of GKP codes (green) compared with a lower (red) and an upper (blue) bound of the quantum capacity of bosonic pure-loss channels $\mathcal{N}[\eta, \bar{n}_{\rm th} = 0]$ (left) and Gaussian thermal loss channels $\mathcal{N}[\eta, \bar{n}_{\rm th} = 1]$ (right). The red line in the left panel overlaps with the blue line, since for bosonic pure-loss channels, the lower and upper bounds coincide with each other and the quantum capacity is completely known.

expression of the hexagonal GKP code states

$$|\mu_L^{\rm hex}\rangle \propto (\hat{X}_L^{\rm hex})^{\mu} \sum_{n_1,n_2=-\infty}^{\infty} (\hat{S}_{{\rm hex},p}^{[d]})^{n_1} (\hat{Z}_L^{\rm hex})^{n_2} |\phi_0\rangle$$
$$= \sum_{n_1,n_2=-\infty}^{\infty} e^{-i(\frac{1}{2}\hat{q}+\frac{\sqrt{3}}{2}\hat{p})\sqrt{\frac{4\pi}{\sqrt{3}d}}(dn_1+\mu)} e^{i\hat{q}\sqrt{\frac{4\pi}{\sqrt{3}d}}n_2} |\phi_0\rangle, \tag{58}$$

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

Note that since $\boldsymbol{S}_{\rm sq}$ and $\boldsymbol{S}_{\rm hex}$ are symplectic matrices (see Eq. (86)), the lattice generated by each of them has a unit cell with area one: $\det(\boldsymbol{S}_{\rm sq}) = \det(\boldsymbol{S}_{\rm hex}) = 1$. For the square lattice, the minimum distance between different lattice points is given by $d_{\rm min}^{\rm sq} = 1$, and for the hexagonal lattice, $d_{\rm min}^{\rm hex} = (2/\sqrt{3})^{1/2} > d_{\rm min}^{\rm sq}$. Thus, the 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 the hexagonal lattice GKP code does so for $r \leq \sqrt{\frac{\pi}{\sqrt{3}d}}$. This leads to the following logical error probability

$$\epsilon_L^{\rm hex} \sim \exp\Big[-\frac{\pi}{2\sqrt{3}d}\Big(\frac{\eta}{1-\eta}\Big)\Big], \tag{59}$$

for the hexagonal lattice GKP code against a bosonic pure-loss channel $\mathcal{N}[\eta, 0]$. Note that the hexagonal lattice allows the densest sphere packing in the 2-dimensional Euclidean space [71].

C. Multi-Mode Symplectic Dual Lattice GKP Codes

Generalizing Eq. (56), one can consider the following stabilizers for an N-mode GKP code space, encoding d logical states per mode:

$$\hat{S}_k^{[d]} \equiv \exp[i\sqrt{2\pi d}(-1)^{k+1}(\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{S})_k], \tag{60}$$

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

D. Achievable Quantum Communication Rate of GKP Codes

Similar to the one-mode case, it is possible to increase the correctable radius of a random displacement in the 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 the Euclidean space allowing $d_{\rm min} \geq \sqrt{N/(\pi e)}$ [72] and a stronger statement was proven in [73] that the same holds also for symplectic lattices. Choosing such a lattice to define the GKP code, one can correct all random displacement errors within the radius $r \leq \sqrt{N/(2\,ed)}$. For the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$, the probability of a displacement with radius larger than $\sqrt{2N}\sigma$ occurring vanishes in the limit of infinitely many modes $N \to \infty$. Thus, if $\sqrt{N/(2\,ed)} \geq \sqrt{2N}\sigma$ is satisfied, i.e.,

$$d \leq d_\sigma \equiv \frac{1}{4e\sigma^2}, \tag{61}$$

encoded information can be transmitted faithfully with an asymptotically vanishing decoding error probability as $N \to \infty$. Then, it follows that a communication rate $R = \log\lfloor d_\sigma \rfloor = \log\lfloor \frac{1}{4e\sigma^2} \rfloor$ can be achieved for the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ (see [41, eq. (55)]; the 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 correctable displacements 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 the Gaussian random displacement channel as defined in Definition 3. Also, let $\mathcal{C}_{\boldsymbol{S}}^{[d]}$ be an N-mode GKP code space defined by the stabilizers in Eq. (60). Then, there exists a symplectic lattice generated by $\boldsymbol{S}$ such that the corresponding GKP code achieves the following rate for the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2]$ in $N \to \infty$ limit.

$$R = \max\left(\log\left\lfloor \frac{1}{e\sigma^2}\right\rfloor, 0\right). \tag{62}$$

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

Thus, a family of GKP codes achieves the one-shot coherent information of Gaussian random displacement channels (see Eq. (29)) modulo the floor function. Based on this, we now establish the following main result:

Theorem 14 (Achievable Quantum Communication Rate of GKP Codes for Gaussian Thermal Loss Channels): Let $\mathcal{N}[\eta, \bar{n}_{\rm th}]$ be the Gaussian thermal loss channel as defined in Definition 1. Upon the optimal choice of a $2N \times 2N$ symplectic matrix $\boldsymbol{S} = \boldsymbol{S}^\star$, the N-mode GKP code as defined in Eq. (60) achieves the following rate for the Gaussian thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\rm th}]$ in $N \to \infty$ limit:

$$R = \max\left(\log\left\lfloor \frac{1}{e(1-\eta)(\bar{n}_{\rm th}+1)}\right\rfloor, 0\right). \tag{63}$$

Proof: Lemma 6 states that a Gaussian thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\rm th}]$ can be converted into a Gaussian random displacement channel $\mathcal{N}_{B_2}[\tilde{\sigma}^2_{\eta,\bar{n}_{\rm th}}]$ by a quantum-limited amplification $\mathcal{A}[1/\eta]$, where $\tilde{\sigma}^2_{\eta,\bar{n}_{\rm th}} = (1-\eta)(\bar{n}_{\rm th}+1)$. Combining this with Lemma 13, Eq. (63) follows. ∎

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

$$Q_{\rm IDP}(\eta, \bar{n}_{\rm th}) < \max\left[\log\left(\frac{1}{(1-\eta)(\bar{n}_{\rm th}+1)}\right), 0\right], \tag{64}$$

as can be derived from Eq. (35). Comparing this with Eq. (63), we get $Q_{\rm reg}(\mathcal{N}[\eta, \bar{n}_{\rm th}]) - R \lesssim \log e = 1.44269\cdots$, where $\sim$ is due to the floor function. Thus, a family of the GKP code defined over an optimal symplectic lattice achieves the quantum capacity of Gaussian thermal loss channels up to at most a constant gap from an upper bound of the quantum capacity (see Fig. 3 for an illustration).

The established rate in Theorem 14 relies on the existence of a symplectic lattice in higher dimensions satisfying a certain desired condition (see [41, eqs. (56) and (57)]). In this regard, we remark that the E_8 lattice and the Leech lattice Λ_{24} (both symplectic; see appendix of [73]) were recently shown to support the densest sphere packing in 8 and 24 dimensional Euclidean spaces, respectively [74], [75], and can be used to define a 4-mode and a 12-mode GKP code, respectively.

V. Biconvex Encoding and Decoding Optimization

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

A. Choi Matrix and Superoperator of a Quantum Channel

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

Let $\mathcal{B} : \mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_3)$ be another quantum channel (i.e., a CPTP map) and consider a composite channel $\mathcal{B}\cdot\mathcal{A} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_3)$. To analyze the composite channel, it is convenient to define a superoperator $\hat{T}_{\mathcal{A}}$ of a channel $\mathcal{A}$ whose matrix elements are defined as $(\hat{T}_{\mathcal{A}})_{jj',ii'} \equiv \langle j_{\mathcal{H}_2}|\mathcal{A}(|i_{\mathcal{H}_1}\rangle\langle i'_{\mathcal{H}_1}|)|j'_{\mathcal{H}_2}\rangle = (\hat{X}_{\mathcal{A}})_{[ij],[i'j']}$ [76]: Then, the superoperator of a composite channel is given by 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 $\mathcal{N} : \mathcal{L}(\mathcal{H}_n) \to \mathcal{L}(\mathcal{H}_n)$ 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 an information sender prepared a maximally entangled state in a local noiseless memory:

$$|\Phi^+\rangle \equiv \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|i_{\mathcal{H}_0}\rangle|i_{\mathcal{H}'_0}\rangle. \tag{65}$$

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

$$\hat{\rho} \equiv (\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}\otimes \mathrm{id}_{\mathcal{H}'_0})(|\Phi^+\rangle\langle\Phi^+|), \tag{66}$$

where $\mathrm{id}_{\mathcal{H}'_0} : \mathcal{L}(\mathcal{H}'_0) \to \mathcal{L}(\mathcal{H}'_0)$ is an identity map associated with the noiseless local memory at the sender's side. Note

that $\hat{\rho}$ in Eq. (66) is explicitly given by

$$\hat{\rho} = \frac{1}{d}\sum_{i,i'=0}^{d-1} (\hat{X}_{\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}})_{[ij],[i'j']} |j_{\mathcal{H}_0}\rangle\langle j'_{\mathcal{H}_0}| \otimes |i_{\mathcal{H}'_0}\rangle\langle i'_{\mathcal{H}'_0}|, \tag{67}$$

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 an approximate non-local entangled state $\hat{\rho}$ can be measured by entanglement fidelity $F_e(\hat{\rho}) \equiv \langle\Phi^+|\hat{\rho}|\Phi^+\rangle$ [77], where $F_e(\hat{\rho}) = 1$ iff $\hat{\rho}$ is the noiseless maximally entangled state $|\Phi^+\rangle\langle\Phi^+|$. The isomorphism between $\hat{\rho}$ and $\hat{X}_{\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}}$ then yields

$$F_e(\hat{\rho}) = (1/d^2)\sum_{i,i'=0}^{d-1} (\hat{X}_{\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}})_{[ii],[i'i']} = \frac{1}{d^2}\mathrm{Tr}[\hat{T}_{\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}}], \tag{68}$$

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

$$(\hat{T}_{\mathcal{D}\cdot\mathcal{N}\cdot\mathcal{E}})_{jj',ii'} = \sum_{k,k',l,l'=0}^{n-1} (\hat{X}_{\mathcal{D}})_{[lj],[l'j']} (\hat{X}_{\mathcal{N}})_{[kl],[k'l']} (\hat{X}_{\mathcal{E}})_{[ik],[i'k']}, \tag{69}$$

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 the decoding $\mathcal{D}$, the noisy channel $\mathcal{N}$, and the encoding $\mathcal{E}$, respectively, and $i, i', j, j' \in \{0, \cdots, d-1\}$. 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. (69).

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

$$\big(f_{\mathcal{N}}(\hat{X})\big)_{[l'i'],[li]} \equiv \sum_{k,k'=0}^{n-1} (\hat{X}_{\mathcal{N}})_{[kl],[k'l']} (\hat{X})_{[ik],[i'k']}, \tag{70}$$

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

$$F_e(\hat{\rho}) = \frac{1}{d^2}\mathrm{Tr}\Big[\hat{X}_{\mathcal{D}} f_{\mathcal{N}}(\hat{X}_{\mathcal{E}})\Big], \tag{71}$$

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 an optimal decoding operation which maximizes entanglement fidelity is a semidefinite program, if the encoding map is fixed to $\mathcal{E} = \bar{\mathcal{E}}$ [78], [79]:

$$\begin{aligned} &\max_{\hat{X}_{\mathcal{D}}} \ \mathrm{Tr}[\hat{X}_{\mathcal{D}}(f_{\mathcal{N}}(\hat{X}_{\bar{\mathcal{E}}}))] \\ &\text{s.t. } \hat{X}_{\mathcal{D}} = \hat{X}_{\mathcal{D}}^\dagger \succeq 0, \quad \mathrm{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{D}} = \hat{I}_{\mathcal{H}_n}, \end{aligned} \tag{72}$$

where the constraints are due to the CPTP nature of $\mathcal{D}$. Similarly, optimizing encoding while fixing decoding is also a semidefinite program, and thus the entire problem is a biconvex optimization [80]. Here, we further impose an energy constraint to the encoding map while still preserving the bi-convexity of the problem. Let $\hat{E} \in \mathcal{L}(\mathcal{H}_n)$ be an energy observable. Let $\mathrm{Tr}_{\mathcal{H}_n}[\hat{E}\hat{\rho}_{\mathcal{E}}]$ be the average energy in the encoded state, where $\hat{\rho}_{\mathcal{E}} \equiv \mathcal{E}(\frac{1}{d}\sum_{i=0}^{d-1} |i_{\mathcal{H}_0}\rangle\langle i_{\mathcal{H}_0}|) = \frac{1}{d}\mathrm{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{E}}$ is the state resulting from $\mathcal{E}$ being applied to the maximally mixed encoded state. Then, the energy constraint reads $\mathrm{Tr}[(\hat{E} \otimes \hat{I}_{\mathcal{H}_0})\hat{X}_{\mathcal{E}}] \le \bar{E}d$ and we obtain the following energy-constrained biconvex encoding and decoding optimization:

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

and

$$\mathrm{Tr}[(\hat{E} \otimes \hat{I}_{\mathcal{H}_0})\hat{X}_{\mathcal{E}}] \le \bar{E}d, \tag{73}$$

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

In principle, a global optimal solution of Eq. (73) can be deterministically found by a global optimization algorithm outlined in [81]. 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 [81] for details), which is intractable in our application below. Thus, we instead solve Eq.(73) heuristically by alternating between encoding and decoding optimization (i.e., SDP sub-problems) starting from a random initial encoding map. We generate the random initial code by taking the first d columns of an $n \times n$ Haar random unitary matrix. To solve each SDP sub-problem we used CVX, a package for specifying and solving convex programs [82], [83].

D. Optimized Qudit-Into-an-Oscillator Codes

First, we specialize Eq. (73) to $\mathcal{N} = \mathcal{N}[\eta, 0]$ and $d = 2$ to find an optimal qubit-into-an-oscillator code for a bosonic pure-loss channel, subject to an average photon number constraint $\mathrm{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}] \le \bar{n}$. To make the optimization tractable, we confine the bosonic Hilbert space to a truncated subspace $\mathcal{H}_n \equiv \mathrm{span}\{|0\rangle, \cdots, |n-1\rangle\}$ and choose $n \gg \bar{n}$ to avoid any artifacts caused by truncation. In particular, we use the Kraus representation of a bosonic pure-loss channel $\mathcal{N}[\eta, 0](\hat{\rho}) = \sum_{\ell=0}^{n-1} \hat{N}_\ell \hat{\rho} \hat{N}_\ell^\dagger$, where the Kraus operators $\hat{N}_\ell$ are given by [32], [84], [85]

$$\hat{N}_\ell = \sqrt{\frac{(1-\eta)^\ell}{\ell!}}\eta^{\frac{\hat{n}}{2}}\hat{a}^\ell. \tag{74}$$

In Fig. 4, we took $\eta = 0.9$, $n = 20$, $d = 2$ and $\bar{n} = 3$ and plot the Wigner function of the maximally mixed code states of the numerically optimized codes (last column), starting from three different random Haar initial codes (first column). In all instances, the obtained codes are given by a hexagonal GKP code (see Fig. 2), up to an overall displacement. The optimized code in the second row exhibits the best performance

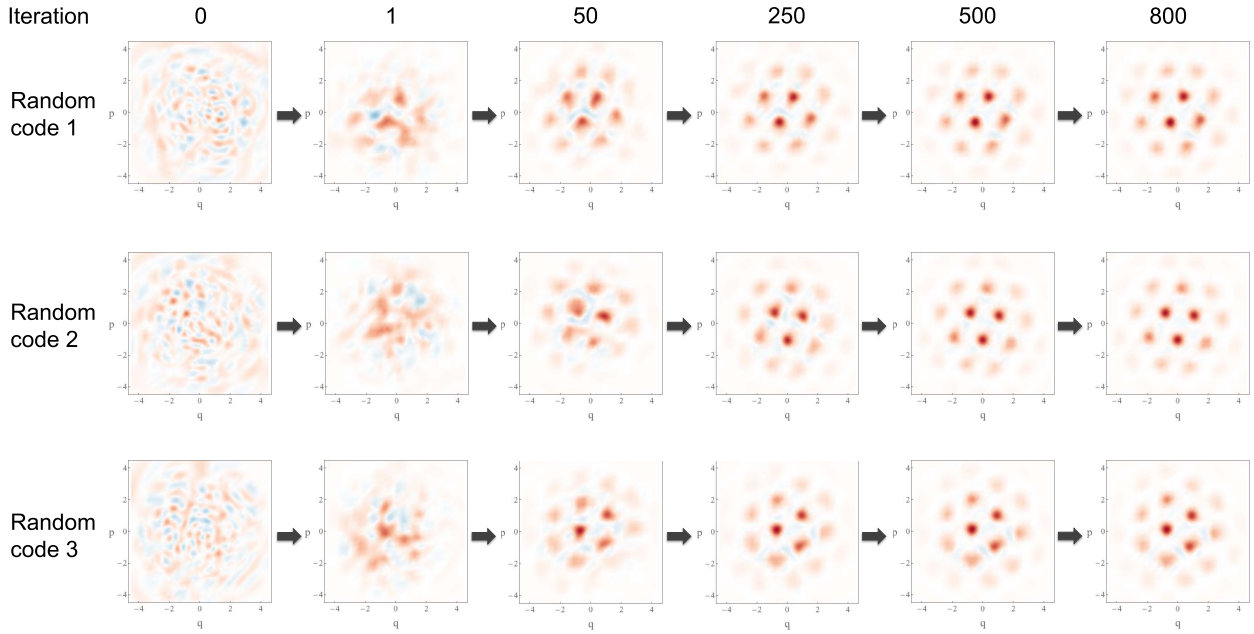

Fig. 4. Biconvex optimization of encoding and decoding maps for a bosonic pure-loss channel $\mathcal{N}[\eta, 0]$ with $\eta = 0.9$. We chose $n = 20$ and $d = 2$ and imposed an average photon number constraint $\mathrm{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}] \leq 3$, where $\hat{\rho}_{\mathcal{E}} = (1/d)\mathrm{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{E}}$ is the maximally mixed code state. The first column of each row represents 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.

(i.e., $1 - F_{\mathcal{N}}^{\star} = 0.002092$). However, we stress that the alternating semidefinite programming method is not guaranteed to yield a global optimal solution.

Although we do not attempt to make a rigorous claim, this optimization result might indicate that the family of GKP codes defined over an optimal symplectic lattice allowing the densest sphere packing may be the optimal encoding for Gaussian thermal loss channels (see below for more numerical evidence). In particular, the constant gap in Theorem 14 may be reduced (or even closed) if we use the optimal decoding, instead of the sub-optimal decoding involving a (noisy) quantum-limited amplification, which we assumed to prove Eq. (63). In this regard, we remark that, for a bosonic pure-loss channel with $\eta = 0.9$, the sub-optimal decoding yields a logical error probability $\epsilon_L^{\mathrm{hex}} \simeq 0.0169$ (obtained by specializing Eq.(59) to $d = 2$ and $\eta = 0.9$) even for the ideal hexagonal GKP code with an infinite photon number, whereas the numerically optimized decoding yields a much smaller infidelity $1 - F_e^{\star} = 0.002092$ for a finite-sized hexagonal GKP code with an average photon number $\bar{n} = 3$.

We also emphasize that the biconvex optimization in Eq. (73) explores the most general form of CPTP encoding maps, including mixed state encoding maps. However, from the numerical optimization, we only obtained a pure-state encoding (i.e., $\mathcal{E}(\hat{\rho}_0) = \hat{V}\hat{\rho}_0\hat{V}^{\dagger}$, where $\hat{V} : \mathcal{H}_0 \to \mathcal{H}_n$ is an isometry $\hat{V}^{\dagger}\hat{V} = \hat{I}_{\mathcal{H}_0}$) as an optimal solution at all iterations of SDP sub-problems.

Now, we further apply the biconvex optimization method to find an optimal qudit-into-an-oscillator code with $2 \leq d \leq 5$ for a bosonic pure-loss channel $\mathcal{N}[\eta, \bar{n}_{\mathrm{th}}] = \mathcal{N}[0.9, 0]$ and also for a Gaussian thermal loss channel $\mathcal{N}[\eta, \bar{n}_{\mathrm{th}}] = \mathcal{N}[0.9, 1]$. To represent the Gaussian thermal loss channel, we use the decomposition in Lemma 7, i.e., $\mathcal{N}[\eta, \bar{n}_{\mathrm{th}}] = \mathcal{A}[G']\mathcal{N}[\eta', 0]$ where $G' = (1 - \eta)\bar{n}_{\mathrm{th}} + 1 = \eta/\eta'$, and the Kraus representation of the quantum-limited amplification $\mathcal{A}[G](\hat{\rho}) = \sum_{g=0}^{\infty}\hat{A}_g\hat{\rho}\hat{A}_g^{\dagger}$ where [86]

$$\hat{A}_g = \sqrt{\frac{1}{g!}\Big(1 - \frac{1}{G}\Big)^g}(\hat{a}^{\dagger})^g\Big(\frac{1}{G}\Big)^{\frac{\hat{n}+1}{2}}. \tag{75}$$

Note that the photon gain parameter g can take any non-negative integer values even when we consider a truncated bosonic space $\mathcal{H}_n$ and thus we should also truncate g at some sufficiently large $g_{\max}$.

In Fig. 5 (a), we took $\eta = 0.9$, $\bar{n}_{\mathrm{th}} \in \{0, 1\}$, $n = 30$, $d \in \{2, 3, 4, 5\}$, $g_{\max} = 15$ (for $\bar{n}_{\mathrm{th}} = 1$) and plot the Wigner function of the maximally mixed code state $\hat{\rho}_{\mathcal{E}}$ of the optimized qudit-into-an-oscillator codes, subject to an average photon number constraint $\mathrm{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}] \leq \bar{n} = 3$. Similarly as in the qubit-into-an-oscillator case, the alternating SDP method (starting from a Haar random initial code) yields a hexagonal GKP code (up to an overall displacement) as an optimal solution of the biconvex optimization for all values of $d \in \{2, 3, 4, 5\}$ and $\bar{n}_{\mathrm{th}} \in \{0, 1\}$. We remark that the lattice spacing of the optimized codes in Fig. 5 (a) scales as $1/\sqrt{d}$, which is consistent with Eq. (56).

Finally, in order to characterize an achievable rate of the numerically optimized codes, we take the obtained code (optimized at $\eta = 0.9$) for each $(d, \bar{n}_{\mathrm{th}})$ and compute the optimal entanglement fidelity $F_e^{\star}(\eta)$ by optimizing the decoding map for bosonic pure-loss channels $\mathcal{N}[\eta, 0]$ and for Gaussian thermal loss channels $\mathcal{N}[\eta, 1]$ for $0.5 \leq \eta \leq 1$ via a semidefinite program (see Eq. (72)). Note that any bipartite

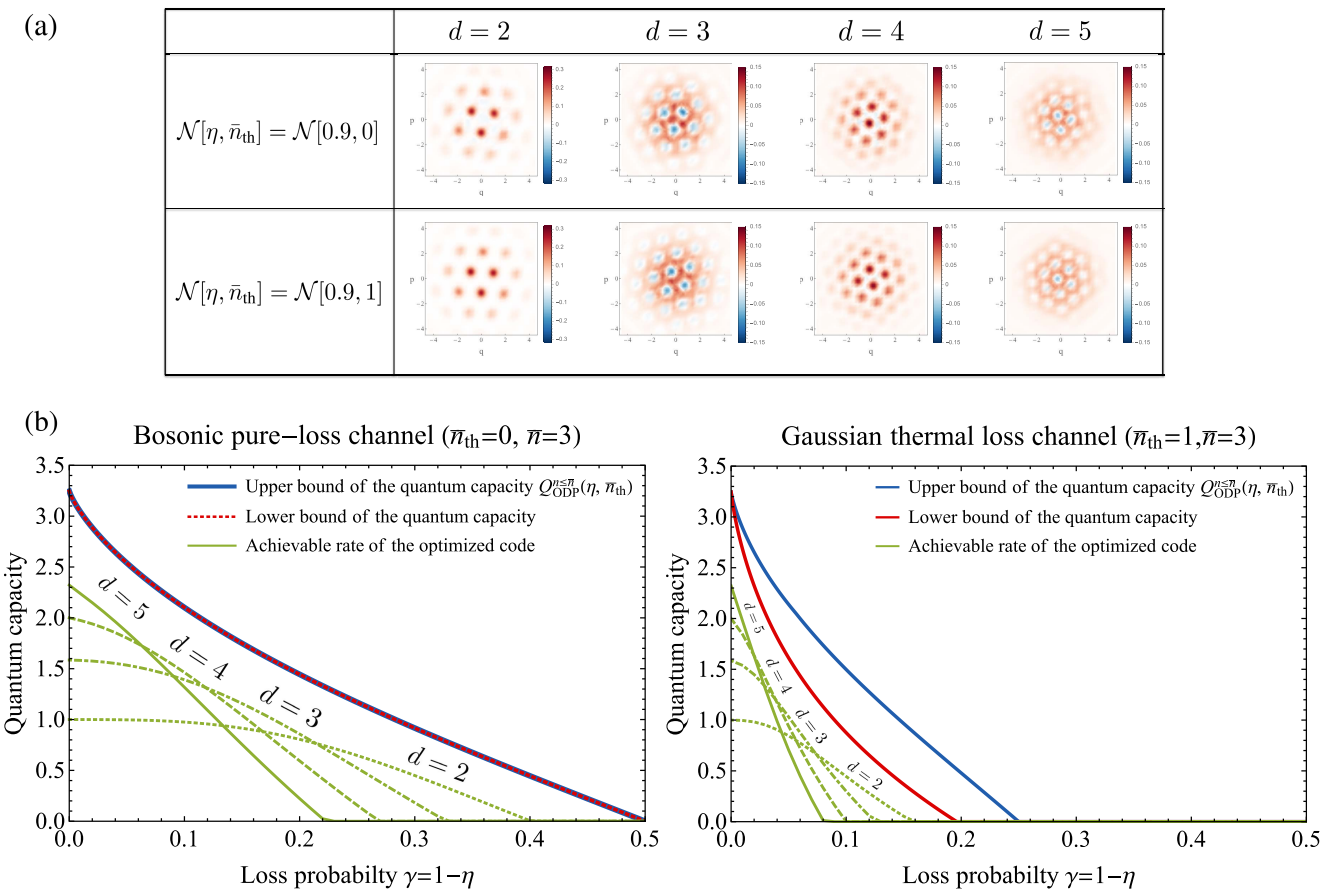

Fig. 5. (a) Wigner function of the maximally mixed code state $\hat{\rho}_{\mathcal{E}}=(1/d)\mathrm{Tr}_{\mathcal{H}_0}\hat{X}_{\mathcal{E}}$ of an optimized qudit-into-an-oscillator code ($2\le d\le 5$) for a bosonic pure-loss channel $\mathcal{N}[0.9,0]$ (top) and a Gaussian thermal loss channel $\mathcal{N}[0.9,1]$ (bottom), subject to an average photon number constraint $\mathrm{Tr}[\hat{n}\hat{\rho}_{\mathcal{E}}]\le\bar{n}=3$, obtained by an alternating SDP method. (b) Lower bound of an achievable rate of the numerically optimized one-mode qudit-into-an-oscillator codes (at $\eta=0.9$) for bosonic pure-loss channels $\mathcal{N}[\eta,0]$ (left) and Gaussian thermal loss channels $\mathcal{N}[\eta,1]$ (right) for each $2\le d\le 5$, compared with lower and upper bounds of the quantum capacity.

state $\hat{\rho}\in\mathcal{D}(\mathcal{H}_0\otimes\mathcal{H}_0)$ ($\dim(\mathcal{H}_0)=d$) with an entanglement fidelity F_e can be converted into an Werner state,

$$\hat{W}(F_e,d)\equiv F_e|\Phi^+\rangle\langle\Phi^+|+\frac{(1-F_e)}{d^2-1}(\hat{I}-|\Phi^+\rangle\langle\Phi^+|), \quad (76)$$

by an LOCC protocol based on the isotropic twirling $\int d\hat{U}\hat{U}\otimes\hat{U}^*\hat{\rho}(\hat{U}\otimes\hat{U}^*)^\dagger$, where $\hat{U}^*$ is the complex conjugate of $\hat{U}$ [87]. Thus, the coherent information of the Werner state

$$R(F_e,d)=\log d+F_e\log F_e+(1-F_e)\log\left(\frac{1-F_e}{d^2-1}\right) \quad (77)$$

is a lower bound of the achievable rate of the state $\hat{\rho}$ with $\langle\Phi^+|\hat{\rho}|\Phi^+\rangle=F_e$. We remark that this is another way to extract an achievable rate from the entanglement fidelity than the methods used in [42] and [88].

In Fig. 5 (b), we plot the lower bound of the achievable rate of the numerically optimized qudit-into-an-oscillator codes (obtained by evaluating $R(F_e^\star,d)$ for the optimized entanglement fidelity $F_e^\star$) for bosonic pure-loss channels $\mathcal{N}[\eta,0]$ and for Gaussian thermal loss channels $\mathcal{N}[\eta,1]$ and compare it with the lower and upper bounds of the energy constrained quantum capacity (with $\bar{n}=3$). The achievable rate deviates from an upper bound of the quantum capacity at most by 0.923 and 1.141 qubits per channel use in the case of $\mathcal{N}[\eta,0]$ and $\mathcal{N}[\eta,1]$ respectively, which are smaller than the gap $\log e=1.44269\cdots$ we found in subsection IV-D. This is because the decoding we considered in subsection IV-D is not the optimal decoding. We remark that the gap between the achievable rate and the upper bound in Fig. 5 (b) may be further reduced if we use a qudit-into-N-oscillators encoding (such as the GKP code defined over an optimal $2N$-dimensional symplectic lattice for $N\ge 2$), instead of qudit-into-an-oscillator encoding.

Finally, we remark that one can find the optimal achievable rate of GKP codes by using Petz decoding [89], [90] (or, transpose decoding), since entanglement infidelity of Petz decoding is at most twice as large as the optimal entanglement infidelity [91]. We leave finding such an optimal achievable rate of GKP codes for Gaussian thermal loss channels as an open problem.

VI. Conclusion

Exploiting the various ways to synthesize and decompose Gaussian channels, we provided improved upper bounds of Gaussian thermal loss channels and Gaussian displacement channels, both in the energy constrained and unconstrained

cases (Theorems 10, 11, 12 and Eqs. (45),(47)). We also established an achievable rate of GKP codes for Gaussian thermal loss channels (Theorem 14), implying (when combined with the previous work [41]) that GKP codes yield an achievable rate for all types of Gaussian noise channels. In the energy-unconstrained case, in particular, we showed that a family of GKP codes achieves an upper bound of the quantum capacity of Gaussian thermal loss channels up to at most a constant gap $\simeq \log e = 1.44269\cdots$ 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 an alternating semidefinite programming method. We demonstrated that, in the one-mode case, a hexagonal GKP code emerges as an optimal encoding from a Haar random initial code.

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]. All of these results are now concisely summarized in Tables I and II. We thank Sharma *et al.* [18] for pointing out a mistake we made in Eq. (38) in an earlier version of our manuscript.

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 a tensor product of N such Hilbert spaces $\mathcal{H}^{\otimes N}$. Each bosonic mode is associated with an annihilation and a creation operator $\hat{a}_k$ 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 $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. The Hilbert space $\mathcal{H}$ is spanned by eigenstates of the number operator $\hat{n} \equiv \hat{a}^\dagger\hat{a}$, i.e., $\mathcal{H} = \mathrm{span}\{|n\rangle\}_{n=0}^\infty$ where $\hat{n}|n\rangle = n|n\rangle$. In the number basis (or the Fock basis), the annihilation and creation operators are given by $\hat{a} = \sum_{n=1}^\infty \sqrt{n}|n-1\rangle\langle n|$, $\hat{a}^\dagger = \sum_{n=0}^\infty \sqrt{n+1}|n+1\rangle\langle n|$. A coherent state $|\alpha\rangle$ is an eigenstate of the annihilation operator $\hat{a}$ with a complex eigenvalue α, i.e., $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$. In the Fock basis, $|\alpha\rangle$ is given by $|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^\infty \frac{\alpha^n}{\sqrt{n!}}|n\rangle$. Note that the vacuum state $|0\rangle$ is a special case of coherent states with $\alpha = 0$. The displacement operator $\hat{D}(\alpha)$ is defined as $\hat{D}(\alpha) \equiv \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$, and a coherent state $|\alpha\rangle$ can be understood as a displaced vacuum state:

$$|\alpha\rangle = \hat{D}(\alpha)|0\rangle. \tag{78}$$

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 GKP codes [28], [42] which differs from [11] by a factor of $\sqrt{2}$ in the definition of $\hat{q}_k$ and $\hat{p}_k$. Define $\hat{\boldsymbol{x}} \equiv (\hat{q}_1, \hat{p}_1, \cdots, \hat{q}_N, \hat{p}_N)^T$. Then, the bosonic commutation relation reads $[\hat{x}_i, \hat{x}_j] = i\Omega_{ij}$, where $\boldsymbol{\Omega}$ is defined as

$$\boldsymbol{\Omega} \equiv \begin{pmatrix} \boldsymbol{\omega} & & \\ & \ddots & \\ & & \boldsymbol{\omega} \end{pmatrix} \quad \text{and } \boldsymbol{\omega} \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{79}$$

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

$$\begin{aligned} \hat{D}(\xi_1/\sqrt{2})|q\rangle &= e^{-i\xi_1\hat{p}}|q\rangle = |q+\xi_1\rangle, \\ \hat{D}(i\xi_2/\sqrt{2})|p\rangle &= e^{i\xi_2\hat{q}}|p\rangle = |p+\xi_2\rangle. \end{aligned} \tag{80}$$

Let $\mathcal{L}(\mathcal{H})$ be the space of linear operators on the Hilbert space $\mathcal{H}$. A general quantum state (pure or mixed) is described by a density operator $\hat{\rho} \in \mathcal{D}(\mathcal{H})$, where $\mathcal{D}(\mathcal{H}) \equiv \{\hat{\rho} \in \mathcal{L}(\mathcal{H}) \,|\, \hat{\rho}^\dagger = \hat{\rho} \succeq 0, \mathrm{Tr}[\hat{\rho}] = 1\}$. The expectation value of an observable $\hat{E}$ of a state $\hat{\rho}$ is given by $\langle\hat{E}\rangle = \mathrm{Tr}[\hat{\rho}\hat{E}]$. The Wigner characteristic function $\chi(\boldsymbol{\xi})$ is defined as $\chi(\boldsymbol{\xi}) \equiv \mathrm{Tr}[\hat{\rho}\exp(i\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi})]$, where $\hat{\rho} \in \mathcal{D}(\mathcal{H}^{\otimes N})$ and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{2N})^T$. The Weyl operator $\exp(i\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi})$ is of the form of a displacement operator and satisfies the orthogonality relation $\mathrm{Tr}[\exp(-i\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi})\exp(i\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi}')] = (2\pi)^N\delta(\boldsymbol{\xi}-\boldsymbol{\xi}')$. The Wigner characteristic function $\chi(\boldsymbol{\xi})$ is in one-to-one correspondence with a state $\hat{\rho}$ and the inverse function is given by $\hat{\rho} = \frac{1}{(2\pi)^N}\int d^{2N}\boldsymbol{\xi}\chi(\boldsymbol{\xi})\exp(-i\hat{\boldsymbol{x}}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi})$. The Wigner function $W(\boldsymbol{x})$ is the Fourier transformation of $\chi(\boldsymbol{\xi})$, i.e.,

$$W(\boldsymbol{x}) = \frac{1}{(2\pi)^N}\int d^{2N}\boldsymbol{\xi}\chi(\boldsymbol{\xi})\exp(-i\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Omega}\boldsymbol{\xi}), \tag{81}$$

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

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

$$\begin{aligned} \chi(\boldsymbol{\xi}) &= \exp\big[-\frac{1}{2}\boldsymbol{\xi}^{\boldsymbol{T}}(\boldsymbol{\Omega}\boldsymbol{V}\boldsymbol{\Omega}^{\boldsymbol{T}})\boldsymbol{\xi} - i(\boldsymbol{\Omega}\bar{\boldsymbol{x}})\boldsymbol{\xi}\big], \\ W(\boldsymbol{x}) &= \frac{\exp\big[-\frac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^T\boldsymbol{V}^{-\boldsymbol{1}}(\boldsymbol{x}-\bar{\boldsymbol{x}})\big]}{(2\pi)^N\sqrt{\det\boldsymbol{V}}}, \end{aligned} \tag{82}$$

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

$$\bar{\boldsymbol{x}} \equiv \langle\hat{\boldsymbol{x}}\rangle = \mathrm{Tr}[\hat{\rho}\hat{\boldsymbol{x}}], \quad V_{ij} \equiv \frac{1}{2}\langle\{\hat{x}_i - \bar{x}_i, \hat{x}_j - \bar{x}_j\}\rangle, \tag{83}$$

where $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$. Thus, a Gaussian state is fully characterized by its first two moments, i.e., $\hat{\rho} = \hat{\rho}_G(\bar{\boldsymbol{x}}, \boldsymbol{V})$. The Heisenberg uncertainty relation reads $\boldsymbol{V} + \frac{i}{2}\boldsymbol{\Omega} \succeq 0$ and implies $V(\hat{q}_k)V(\hat{p}_k) \geq \frac{1}{4}$ for all $k \in \{1, \cdots, N\}$, where $V(\hat{x}_i) \equiv V_{ii}$.

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

$$\hat{\rho}_{\bar{n}_{\mathrm{th}}} \equiv \sum_{n=0}^\infty \frac{(\bar{n}_{\mathrm{th}})^n}{(\bar{n}_{\mathrm{th}}+1)^{n+1}}|n\rangle\langle n| = \hat{\rho}_G\Big(\boldsymbol{0}, \Big(\bar{n}_{\mathrm{th}} + \frac{1}{2}\Big)\boldsymbol{I}_2\Big), \tag{84}$$

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

$$H(\hat{\rho}_{\bar{n}_{\rm th}}) = g(\bar{n}_{\rm th}), \tag{85}$$

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

A unitary operation that maps a Gaussian state to another Gaussian state is called a Gaussian unitary operation. A Gaussian unitary operation is generated by a second order polynomial of $\hat{\boldsymbol{a}} = (\hat{a}_1, \cdots, \hat{a}_N)^T$ and $\hat{\boldsymbol{a}}^\dagger = (\hat{a}_1^\dagger, \cdots, \hat{a}_N^\dagger)^T$, i.e., $\hat{U}_G = \exp(-i\hat{H})$ with $\hat{H} = i(\boldsymbol{\alpha}^{\boldsymbol{T}}\hat{\boldsymbol{a}}^\dagger + \hat{\boldsymbol{a}}^\dagger \boldsymbol{F}\hat{\boldsymbol{a}} + \hat{\boldsymbol{a}}^\dagger \boldsymbol{G}\hat{\boldsymbol{a}}^\dagger) + {\rm h.c.}$, where $\boldsymbol{\alpha}^{\boldsymbol{T}} = (\alpha_1, \cdots, \alpha_N)$ and $\boldsymbol{F}, \boldsymbol{G}$ are $N \times N$ complex matrices. In the Heisenberg picture, the annihilation operator $\hat{\boldsymbol{a}}$ is transformed into $\hat{U}_G^\dagger \hat{\boldsymbol{a}} \hat{U}_G = \boldsymbol{A}\hat{\boldsymbol{a}} + \boldsymbol{B}\hat{\boldsymbol{a}}^\dagger + \boldsymbol{\alpha}$, where $N \times N$ complex matrices $\boldsymbol{A}, \boldsymbol{B}$ (determined by $\boldsymbol{F}, \boldsymbol{G}$) satisfy $\boldsymbol{A}\boldsymbol{B}^{\boldsymbol{T}} = \boldsymbol{B}\boldsymbol{A}^{\boldsymbol{T}}$ and $\boldsymbol{A}\boldsymbol{A}^\dagger = \boldsymbol{B}\boldsymbol{B}^\dagger + \boldsymbol{I}_N$. In terms of the quadrature operators, the transformation reads $\boldsymbol{x} \to \hat{U}_G^\dagger \boldsymbol{x} \hat{U}_G = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{d}$, where $\boldsymbol{d} = (d_1, \cdots, d_{2N})^T = \sqrt{2}(\alpha_1^R, \alpha_1^I, \cdots, \alpha_N^R, \alpha_N^I)^T$ and the $2N \times 2N$ matrix $\boldsymbol{S}$ is symplectic:

$$\boldsymbol{S}\boldsymbol{\Omega}\boldsymbol{S}^{\boldsymbol{T}} = \boldsymbol{\Omega}. \tag{86}$$

A Gaussian unitary operation is thus fully characterized by $\boldsymbol{S}, \boldsymbol{d}$, and under $\hat{U}_{\boldsymbol{S},\boldsymbol{d}}$ the first two moments of a Gaussian state $\hat{\rho}_G(\bar{\boldsymbol{x}}, \boldsymbol{V})$ are transformed as

$$\bar{\boldsymbol{x}} \to \boldsymbol{S}\bar{\boldsymbol{x}} + \boldsymbol{d}, \quad \boldsymbol{V} \to \boldsymbol{S}\boldsymbol{V}\boldsymbol{S}^{\boldsymbol{T}}. \tag{87}$$

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

$$\hat{\boldsymbol{x}} \to \boldsymbol{R}(\theta)\hat{\boldsymbol{x}} \text{ and } \boldsymbol{R}(\theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{88}$$

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

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

$$\hat{\boldsymbol{x}} \to \boldsymbol{B}(\eta)\hat{\boldsymbol{x}} \text{ and } \boldsymbol{B}(\eta) \equiv \begin{pmatrix} \sqrt{\eta}\boldsymbol{I}_2 & \sqrt{1-\eta}\boldsymbol{I}_2 \\ -\sqrt{1-\eta}\boldsymbol{I}_2 & \sqrt{\eta}\boldsymbol{I}_2 \end{pmatrix}. \tag{89}$$

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

$$\hat{\boldsymbol{x}} \to \boldsymbol{S_2}(G)\hat{\boldsymbol{x}} \text{ and } \boldsymbol{S_2}(G) \equiv \begin{pmatrix} \sqrt{G}\boldsymbol{I}_2 & \sqrt{G-1}\boldsymbol{Z}_2 \\ \sqrt{G-1}\boldsymbol{Z}_2 & \sqrt{G}\boldsymbol{I}_2 \end{pmatrix}, \tag{90}$$

where $\boldsymbol{Z}_2 \equiv {\rm diag}(1, -1)$.

Appendix B
Characterization of Gaussian Random Displacement Channel

Here, we prove that the Gaussian random displacement channel $\mathcal{N}_{B_2}[\sigma^2](\hat{\rho})$, as defined in Definition 3, is characterized by $\boldsymbol{T} = \boldsymbol{I}_2$, $\boldsymbol{N} = \sigma^2\boldsymbol{I}_2$ and $\boldsymbol{d} = 0$. Although this is a standard calculation [12], we give this proof for the convenience of the readers.

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

$$\begin{aligned}
\chi'(\boldsymbol{\xi}) &= {\rm Tr}\Big[\mathcal{N}_{B_2}[\sigma^2](\hat{\rho})\hat{D}\Big(\frac{\xi}{\sqrt{2}}\Big)\Big] \\
&= {\rm Tr}\Big[\frac{1}{\pi\sigma^2}\int d^2\alpha e^{-\frac{|\alpha|^2}{\sigma^2}}\hat{\rho}\hat{D}^\dagger(\alpha)\hat{D}\Big(\frac{\xi}{\sqrt{2}}\Big)\hat{D}(\alpha)\Big] \\
&= \frac{1}{\pi\sigma^2}\int d^2\alpha e^{-\frac{|\alpha|^2}{\sigma^2}} e^{\frac{1}{\sqrt{2}}(\alpha^*\xi - \alpha\xi^*)}{\rm Tr}\Big[\hat{\rho}\hat{D}\Big(\frac{\xi}{\sqrt{2}}\Big)\Big] \\
&= e^{-\frac{\sigma^2}{2}|\xi|^2}\chi(\boldsymbol{\xi}),
\end{aligned} \tag{91}$$

where $\chi(\boldsymbol{\xi})$ is the Wigner characteristic function of an input state $\hat{\rho}$ (see the text above Eq. (81)). We used the cyclic property of the trace to derive the second equality, and applied the 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 evaluation of a Gaussian integral and the definition of $\chi(\xi)$. If the input state is Gaussian, i.e., $\hat{\rho} = \hat{\rho}_G(\bar{\boldsymbol{x}}, \boldsymbol{V})$, the Wigner characteristic function of the output state is given by

$$\chi'(\boldsymbol{\xi}) = \exp\Big[-\frac{1}{2}\boldsymbol{\xi}^{\boldsymbol{T}}(\boldsymbol{\Omega}(\boldsymbol{V} + \sigma^2\boldsymbol{I}_2)\boldsymbol{\Omega}^{\boldsymbol{T}}\boldsymbol{\xi}) - i(\boldsymbol{\Omega}\bar{\boldsymbol{x}})\boldsymbol{\xi}\Big], \tag{92}$$

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

Appendix C
Derivation of the Optimized Data-Processing Bound

Proof: Assume that a Gaussian thermal loss channel is decomposed as

$$\mathcal{N}[\eta, \bar{n}_{\rm th}] = \mathcal{A}[G_1]\mathcal{N}[\bar{\eta}, 0]\mathcal{A}[G_2], \tag{93}$$

(see Eq. (41)). Then, the channel on the right hand side is characterized by $\boldsymbol{T} = \sqrt{G_1\bar{\eta}G_2}\boldsymbol{I}_2$ and

$$\begin{aligned}
\boldsymbol{N} &= \Big[G_1\Big(\bar{\eta} \times \frac{1}{2}(G_2 - 1) + \frac{1}{2}(1 - \bar{\eta})\Big) + \frac{1}{2}(G_1 - 1)\Big]\boldsymbol{I_2} \\
&= \frac{1}{2}\Big[G_1\bar{\eta}G_2 - 2G_1\bar{\eta} + 2G_1 - 1\Big]\boldsymbol{I_2}.
\end{aligned} \tag{94}$$

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

$$\begin{aligned} &Q^{n\leq G_2\bar{n}+(G_2-1)}(\mathcal{N}[\bar{\eta},0]) \\ &\quad\equiv \max\Big[g\big(\bar{\eta}(G_2\bar{n}+(G_2-1))\big) \\ &\qquad - g\big((1-\bar{\eta})(G_2\bar{n}+(G_2-1))\big),0\Big]. \end{aligned} \tag{95}$$

hence Eq. (44). To ensure $G_1 \geq 1$ and $G_2 = \frac{\eta}{G_1-(1-\eta)(\bar{n}_{\rm th}+1)} \geq 1$, G_1 should be in the range $1 \leq G_1 \leq 1+(1-\eta)\bar{n}_{\rm th}$. Thus, Eq. (43) follows. ∎

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.

REFERENCES

[1] M. M. Wilde, *Quantum Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[2] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A, Gen. Phys.*, vol. 52, pp. R2493–R2496, Oct. 1995, doi: 10.1103/PhysRevA.52.R2493.

[3] B. Schumacher and M. A. Nielsen, "Quantum data processing and error correction," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 4, pp. 2629–2635, 1996, doi: 10.1103/PhysRevA.54.2629.

[4] S. Lloyd, "Capacity of the noisy quantum channel," *Phys. Rev. A, Gen. Phys.*, vol. 55, no. 3, pp. 1613–1622, 1997, doi: 10.1103/PhysRevA.55.1613.

[5] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, Jan. 2005.

[6] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, "Quantum-channel capacity of very noisy channels," *Phys. Rev. A, Gen. Phys.*, vol. 57, no. 2, pp. 830–839, Feb. 1998, doi: 10.1103/PhysRevA.57.830.

[7] G. Smith and J. T. Yard, "Quantum communication with zero-capacity channels," *Science*, vol. 321, pp. 1812–1815, Sep. 2008. [Online]. Available: http://science.sciencemag.org/content/321/5897/1812

[8] G. Smith, J. A. Smolin, and J. Yard, "Quantum communication with Gaussian channels of zero quantum capacity," *Nature Photon.*, vol. 5, pp. 624–627, Aug. 2011, doi: 10.1038/nphoton.2011.203.

[9] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, "Unbounded number of channel uses may be required to detect quantum capacity," *Nature Commun.*, vol. 6, p. 6739, Mar. 2015, doi: 10.1038/ncomms7739.

[10] J. Eisert and M. M. Wolf, "Gaussian quantum channels," in *Quantum Information With Continuous Variables of Atoms and Light*, 2007, pp. 23–42.

[11] C. Weedbrook *et al.*, "Gaussian quantum information," *Rev. Mod. Phys.*, vol. 84, pp. 621–669, May 2012, doi: 10.1103/RevModPhys.84.621.

[12] A. Serafini, *Quantum Continuous Variables*. Boca Raton, FL, USA: CRC Press, Jul. 2017, doi: 10.1201%2F9781315118727.

[13] A. S. Holevo and R. F. Werner, "Evaluating capacities of bosonic Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 63, p. 032312, Feb. 2001, doi: 10.1103/PhysRevA.63.032312.

[14] M. M. Wolf, D. Pérez-García, and G. Giedke, "Quantum capacities of bosonic channels," *Phys. Rev. Lett.*, vol. 98, p. 130501, Mar. 2007, doi: 10.1103/PhysRevLett.98.130501.

[15] M. M. Wilde, P. Hayden, and S. Guha, "Quantum trade-off coding for bosonic communication," *Phys. Rev. A, Gen. Phys.*, vol. 86, no. 6, p. 062306, Dec. 2012, doi: 10.1103/PhysRevA.86.062306.

[16] M. M. Wilde and H. Qi. (Sep. 2016). "Energy-constrained private and quantum capacities of quantum channels." [Online]. Available: https://arxiv.org/abs/1609.01997

[17] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, "Fundamental limits of repeaterless quantum communications," *Nature Commun.*, vol. 8, p. 15043, Apr. 2017, doi: 10.1038/ncomms15043.

[18] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, "Bounding the energy-constrained quantum and private capacities of phase-insensitive bosonic Gaussian channels," *New J. Phys.*, vol. 20, p. 063025, Jun. 2018. [Online]. Available: http://stacks.iop.org/1367-2630/20/i=6/a=063025

[19] M. Rosati, A. Mari, and V. Giovannetti. (Jan. 2018). "Narrow bounds for the quantum capacity of thermal attenuators." [Online]. Available: https://arxiv.org/abs/1801.04731

[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, Gen. Phys.*, vol. 59, pp. 2631–2634, Apr. 1999, doi: 10.1103/PhysRevA.59.2631.

[21] J. Niset, U. L. Andersen, and N. J. Cerf, "Experimentally feasible quantum erasure-correcting code for continuous variables," *Phys. Rev. Lett.*, vol. 101, no. 13, p. 130503, Sep. 2008, doi: 10.1103/PhysRevLett.101.130503.

[22] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, "Hardware-efficient autonomous quantum memory protection," *Phys. Rev. Lett.*, vol. 111, p. 120501, Sep. 2013, doi: 10.1103/PhysRevLett.111.120501.

[23] M. Mirrahimi *et al.*, "Dynamically protected cat-qubits: A new paradigm for universal quantum computation," *New J. Phys.*, vol. 16, no. 4, p. 045014, 2014. [Online]. Available: http://stacks.iop.org/1367-2630/16/i=4/a=045014

[24] F. Lacerda, J. M. Renes, and V. B. Scholz, "Coherent-state constellations and polar codes for thermal Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 95, p. 062343, Jun. 2017, doi: 10.1103/PhysRevA.95.062343.

[25] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang. (Jan. 2018). "Pair-cat codes: Autonomous error-correction with low-order nonlinearity." [Online]. Available: https://arxiv.org/abs/1801.05897

[26] S. Lloyd and J.-J. E. Slotine, "Analog quantum error correction," *Phys. Rev. Lett.*, vol. 80, no. 18, pp. 4088–4091, May 1998, doi: 10.1103/PhysRevLett.80.4088.

[27] S. L. Braunstein, "Error correction for continuous quantum variables," *Phys. Rev. Lett.*, vol. 80, pp. 4084–4087, May 1998, doi: 10.1103/PhysRevLett.80.4084.

[28] D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator," *Phys. Rev. A, Gen. Phys.*, vol. 64, p. 012310, Jun. 2001, doi: 10.1103/PhysRevA.64.012310.

[29] N. C. Menicucci, "Fault-tolerant measurement-based quantum computing with continuous-variable cluster states," *Phys. Rev. Lett.*, vol. 112, no. 12, p. 120504, Mar. 2014, doi: 10.1103/PhysRevLett.112.120504.

[30] P. Hayden, S. Nezami, G. Salton, and B. C. Sanders, "Spacetime replication of continuous variable quantum information," *New J. Phys.*, vol. 18, no. 8, p. 083043, Aug. 2016. [Online]. Available: http://stacks.iop.org/1367-2630/18/i=8/a=083043?key=crossref.8f7fb8717a56e53e6d6ba36ed20c74aa

[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, Gen. Phys.*, vol. 94, no. 2, p. 022325, Aug. 2016, doi: 10.1103/PhysRevA.94.022325.

[32] I. L. Chuang, D. W. Leung, and Y. Yamamoto, "Bosonic quantum codes for amplitude damping," *Phys. Rev. A, Gen. Phys.*, vol. 56, pp. 1114–1125, Aug. 1997, doi: 10.1103/PhysRevA.56.1114.

[33] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, no. 6816, pp. 46–52, Jan. 2001. [Online]. Available: http://www.nature.com/doifinder/10.1038/35051009

[34] T. C. Ralph, A. J. F. Hayes, and A. Gilchrist, "Loss-tolerant optical qubits," *Phys. Rev. Lett.*, vol. 95, no. 10, p. 100501, Aug. 2005, doi: 10.1103/PhysRevLett.95.100501.

[35] W. Wasilewski and K. Banaszek, "Protecting an optical qubit against photon loss," *Phys. Rev. A, Gen. Phys.*, vol. 75, no. 4, p. 042316, Apr. 2007, doi: 10.1103/PhysRevA.75.042316.

[36] M. Bergmann and P. van Loock, "Quantum error correction against photon loss using NOON states," *Phys. Rev. A, Gen. Phys.*, vol. 94, no. 1, p. 012311, Jul. 2016, doi: 10.1103/PhysRevA.94.012311.

[37] M. H. Michael *et al.*, "New class of quantum error-correcting codes for a bosonic mode," *Phys. Rev. X*, vol. 6, p. 031006, Jul. 2016, doi: 10.1103/PhysRevX.6.031006.

[38] M. Y. Niu, I. L. Chuang, and J. H. Shapiro, "Hardware-efficient bosonic quantum error-correcting codes based on symmetry operators," *Phys. Rev. A, Gen. Phys.*, vol. 97, p. 032323, Mar. 2018, doi: 10.1103/PhysRevA.97.032323.

[39] S.-W. Lee and H. Jeong, "Near-deterministic quantum teleportation and resource-efficient quantum computation using linear optics and hybrid qubits," *Phys. Rev. A, Gen. Phys.*, vol. 87, no. 2, p. 022326, Feb. 2013, doi: 10.1103/PhysRevA.87.022326.

[40] E. Kapit, "Hardware-efficient and fully autonomous quantum error correction in superconducting circuits," *Phys. Rev. Lett.*, vol. 116, no. 15, p. 150501, Apr. 2016, doi: 10.1103/PhysRevLett.116.150501.

[41] J. Harrington and J. Preskill, "Achievable rates for the Gaussian quantum channel," *Phys. Rev. A, Gen. Phys.*, vol. 64, p. 062301, Nov. 2001, doi: 10.1103/PhysRevA.64.062301.

[42] V. V. Albert *et al.*, "Performance and structure of single-mode bosonic codes," *Phys. Rev. A, Gen. Phys.*, vol. 97, p. 032346, Mar. 2018, doi: 10.1103/PhysRevA.97.032346.

[43] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra Appl.*, vol. 10, no. 3, pp. 285–290, 1975. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0024379575900750

[44] F. Caruso, J. Eisert, V. Giovannetti, and A. S. Holevo, "Multi-mode bosonic Gaussian channels," *New J. Phys.*, vol. 10, no. 8, p. 083030, Aug. 2008. [Online]. Available: http://stacks.iop.org/1367-2630/10/i=8/a=083030

[45] I. Devetak and P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," *Commun. Math. Phys.*, vol. 256, no. 2, pp. 287–303, Jun. 2005, doi: 10.1007/s00220-005-1317-6.

[46] F. Caruso and V. Giovannetti, "Degradability of bosonic Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 74, no. 6, p. 062307, Dec. 2006, doi: 10.1103/PhysRevA.74.062307.

[47] F. Caruso, V. Giovannetti, and A. S. Holevo, "One-mode bosonic Gaussian channels: A full weak-degradability classification," *New J. Phys.*, vol. 8, no. 12, p. 310, Dec. 2006. [Online]. Available: http://stacks.iop.org/1367-2630/8/i=12/a=310

[48] A. S. Holevo, "One-mode quantum Gaussian channels: Structure and quantum capacity," *Problems Inf. Transmiss.*, vol. 43, no. 1, pp. 1–11, Mar. 2007, doi: 10.1134/S0032946007010012.

[49] H. Heffner, "The fundamental noise limit of linear amplifiers," *Proc. IRE*, vol. 50, no. 7, pp. 1604–1608, Jul. 1962.

[50] R. García-Patrón, C. Navarrete-Benlloch, S. Lloyd, J. H. Shapiro, and N. J. Cerf, "Majorization theory approach to the Gaussian channel minimum entropy conjecture," *Phys. Rev. Lett.*, vol. 108, p. 110505, Mar. 2012, doi: 10.1103/PhysRevLett.108.110505.

[51] D. Gottesman and J. Preskill, "Secure quantum key distribution using squeezed states," *Phys. Rev. A, Gen. Phys.*, vol. 63, p. 022309, Jan. 2001, doi: 10.1103/PhysRevA.63.022309.

[52] A. S. Holevo, "Entanglement-breaking channels in infinite dimensions," *Problems Inf. Transmiss.*, vol. 44, no. 3, pp. 171–184, Sep. 2008, doi: 10.1134/S0032946008030010.

[53] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro, "Minimum output entropy of bosonic channels: A conjecture," *Phys. Rev. A, Gen. Phys.*, vol. 70, p. 032315, Sep. 2004, doi: 10.1103/PhysRevA.70.032315.

[54] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, "Approximate degradable quantum channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 12, pp. 7832–7844, Dec. 2017.

[55] J. Eisert, S. Scheel, and M. B. Plenio, "Distilling Gaussian states with Gaussian operations is impossible," *Phys. Rev. Lett.*, vol. 89, p. 137903, Sep. 2002, doi: 10.1103/PhysRevLett.89.137903.

[56] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar. 1993, doi: 10.1103/PhysRevLett.70.1895.

[57] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 5, pp. 3824–3851, Nov. 1996.

[58] J. Niset, J. Fiurášek, and N. J. Cerf, "No-go theorem for Gaussian quantum error correction," *Phys. Rev. Lett.*, vol. 102, p. 120501, Mar. 2009, doi: 10.1103/PhysRevLett.102.120501.

[59] D. Gottesman and I. L. Chuang, "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations," *Nature*, vol. 402, no. 6760, pp. 390–393, Nov. 1999, doi: 10.1038/46503.

[60] B. C. Travaglione and G. J. Milburn, "Preparing encoded states in an oscillator," *Phys. Rev. A, Gen. Phys.*, vol. 66, p. 052322, Nov. 2002, doi: 10.1103/PhysRevA.66.052322.

[61] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, "Constructing finite-dimensional codes with optical continuous variables," *Europhys. Lett.*, vol. 68, no. 3, p. 323, 2004. [Online]. Available: http://stacks.iop.org/0295-5075/68/i=3/a=323

[62] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, "Generating continuous variable quantum codewords in the near-field atomic lithography," *J. Phys. B, At., Mol. Opt. Phys.*, vol. 39, no. 4, p. 997, 2006. [Online]. Available: http://stacks.iop.org/0953-4075/39/i=4/a=023

[63] H. M. Vasconcelos, L. Sanz, and S. Glancy, "All-optical generation of states for 'Encoding a qubit in an oscillator,"' *Opt. Lett.*, vol. 35, no. 19, pp. 3261–3263, Oct. 2010. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-35-19-3261

[64] B. M. Terhal and D. Weigand, "Encoding a qubit into a cavity mode in circuit QED using phase estimation," *Phys. Rev. A, Gen. Phys.*, vol. 93, p. 012315, Jan. 2016, doi: 10.1103/PhysRevA.93.012315.

[65] K. R. Motes, B. Q. Baragiola, A. Gilchrist, and N. C. Menicucci, "Encoding qubits into oscillators with atomic ensembles and squeezed light," *Phys. Rev. A, Gen. Phys.*, vol. 95, p. 053819, May 2017, doi: 10.1103/PhysRevA.95.053819.

[66] D. J. Weigand and B. M. Terhal, "Generating grid states from Schrödinger-cat states without postselection," *Phys. Rev. A, Gen. Phys.*, vol. 97, p. 022341, Feb. 2018, doi: 10.1103/PhysRevA.97.022341.

[67] A. Kitaev, "Quantum measurements and the Abelian stabilizer problem," *Electron. Colloq. Comput. Complex.*, vol. 3, 1996.

[68] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, "Sequential modular position and momentum measurements of a trapped ion mechanical oscillator," *Phys. Rev. X*, vol. 8, p. 021001, Apr. 2018, doi: 10.1103/PhysRevX.8.021001.

[69] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. Home. (Jul. 2018). "Encoding a qubit in a trapped-ion mechanical oscillator." [Online]. Available: https://arxiv.org/abs/1807.01033

[70] L. Sun *et al.*, "Tracking photon jumps with repeated quantum non-demolition parity measurements," *Nature*, vol. 511, pp. 444–448, Jul. 2014, doi: 10.1038/nature13436.

[71] L. Fejes, "Über die dichteste Kugellagerung," *Math. Zeitschrift*, vol. 48, no. 1, pp. 676–684, Dec. 1942, doi: 10.1007/BF01180035.

[72] E. Hlawka, "Zur Geometrie der Zahlen," *Math. Zeitschrift*, vol. 49, no. 1, pp. 285–312, Dec. 1943, doi: 10.1007/BF01174201.

[73] 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 Math.*, vol. 117, no. 1, pp. 27–56, Dec. 1994, doi: 10.1007/BF01232233.

[74] M. S. Viazovska, "The sphere packing problem in dimension 8," *Ann. Math.*, vol. 185, no. 3, pp. 991–1015, 2017.

[75] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, "The sphere packing problem in dimension 24," *Ann. Math.*, vol. 185, no. 3, pp. 1017–1033, 2017.

[76] J. Preskill. *Lecture Note for Quantum Computation, Chapter 3. Foundations of Quantum Theory II: Measurement and Evolution*. [Online]. Available: http://www.theory.caltech.edu/people/preskill/ph229/notes/chap3.pdf

[77] B. Schumacher, "Sending entanglement through noisy quantum channels," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 4, pp. 2614–2628, Oct. 1996, doi: 10.1103/PhysRevA.54.2614.

[78] M. Reimpell and R. F. Werner, "Iterative optimization of quantum error correcting codes," *Phys. Rev. Lett.*, vol. 94, p. 080501, Mar. 2005, doi: 10.1103/PhysRevLett.94.080501.

[79] A. S. Fletcher, P. W. Shor, and M. Z. Win, "Optimum quantum error recovery using semidefinite programming," *Phys. Rev. A, Gen. Phys.*, vol. 75, p. 012338, Jan. 2007, doi: 10.1103/PhysRevA.75.012338.

[80] R. L. Kosut and D. A. Lidar, "Quantum error correction via convex optimization," *Quantum Inf. Process.*, vol. 8, no. 5, pp. 443–459, Oct. 2009, doi: 10.1007/s11128-009-0120-2.

[81] C. A. Floudas and V. Visweswaran, "A global optimization algorithm (GOP) for certain classes of nonconvex NLPs—I. Theory," *Comput. Chem. Eng.*, vol. 14, no. 12, pp. 1397–1417, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/pii/009813549080020C

[82] CVX Research. (Aug. 2012). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.0*. [Online]. Available: http://cvxr.com/cvx

[83] 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), V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag, 2008, pp. 95–110.
[84] M. Ueda, "Probability-density-functional description of quantum photodetection processes," *Quantum Opt., J. Eur. Opt. Soc. B*, vol. 1, no. 2, p. 131, 1989. [Online]. Available: http://stacks.iop.org/0954-8998/1/i=2/a=005
[85] C. T. Lee, "Superoperators and their implications in the hybrid model for photodetection," *Phys. Rev. A, Gen. Phys.*, vol. 49, pp. 4888–4894, Jun. 1994, doi: 10.1103/PhysRevA.49.4888.
[86] J. S. Ivan, K. K. Sabapathy, and R. Simon, "Operator-sum representation for bosonic Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 84, p. 042311, Oct. 2011, doi: 10.1103/PhysRevA.84.042311.
[87] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Phys. Rev. A, Gen. Phys.*, vol. 40, pp. 4277–4281, Oct. 1989, doi: 10.1103/PhysRevA.40.4277.
[88] L. Li, C.-L. Zou, V. V. Albert, S. Muralidharan, S. M. Girvin, and L. Jiang, "Cat codes with optimal decoherence suppression for a lossy bosonic channel," *Phys. Rev. Lett.*, vol. 119, p. 030502, Jul. 2017, doi: 10.1103/PhysRevLett.119.030502.
[89] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra," *Commun. Math. Phys.*, vol. 105, no. 1, pp. 123–131, Mar. 1986, doi: 10.1007/BF01212345.
[90] D. Petz, "Sufficiency of channels over von Neumann algebras," *Quart. J. Math.*, vol. 39, no. 1, pp. 97–108, Mar. 1988, doi: 10.1093/qmath/39.1.97.
[91] H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," *J. Math. Phys.*, vol. 43, no. 5, pp. 2097–2106, 2002. [Online]. Available: http://aip.scitation.org/doi/abs/10.1063/1.1459754

Kyungjoo Noh is a graduate student in physics at Yale University. He received his B.S. in physics from Seoul National University in 2014. He is a fellow of the Korea Foundation for Advanced Studies. His research interests include quantum Shannon theory, quantum error-correcting codes, open quantum systems, and experimental realization of quantum computation.

Victor V. Albert received his Bachelor's degree in physics and mathematics from the University of Florida in 2010 and his Ph.D. in physics from Yale University in 2017. He is currently a Lee A. DuBridge Postdoctoral Scholar in Physics at the California Institute of Technology. He pursues an interdisciplinary line of research in quantum science, including open quantum systems, error-correction, experimental realizations, and topological band theory.

Liang Jiang is an associate professor of applied physics at Yale University. He obtained his B.S. from Caltech in 2004 and Ph.D. from Harvard in 2009. After working as a Sherman Fairchild postdoctoral fellow at Caltech, he joined Yale University in 2012. Liang's research interests are quantum optics, condensed matter physics, nanotechnology, and quantum information science.