Information transmission with continuous variable quantum erasure channels

Changchun Zhong,[1, *] Changhun Oh,[1] and Liang Jiang[1]

[1]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA*
(Dated: March 6, 2023)

Quantum capacity, as the key figure of merit for a given quantum channel, upper bounds the channel's ability in transmitting quantum information. Identifying different types of channels, evaluating the corresponding quantum capacity, and finding the capacity-approaching coding scheme are the major tasks in quantum communication theory. Quantum channel in discrete variables has been discussed enormously based on various error models, while error model in the continuous variable channel has been less studied due to the infinite dimensional problem. In this paper, we investigate a general continuous variable quantum erasure channel. By defining an effective subspace of the continuous variable system, we find a continuous variable random coding model. We then derive the quantum capacity of the continuous variable erasure channel in the framework of decoupling theory. The discussion in this paper fills the gap of a quantum erasure channel in continuous variable setting and sheds light on the understanding of other types of continuous variable quantum channels.

I. INTRODUCTION

Besides storing and manipulating quantum information, efficiently transmitting them through space and time with high fidelity is essential for many practical quantum applications [1–4]. In quantum communication theory, information transmission is modeled as a quantum channel, which is defined by a completely positive and trace preserving (CPTP) map $\mathcal{N} : \rho_A \to \rho_B$, where the map $\mathcal{N}$ takes some input $\rho_A \in \mathcal{H}_A$ to the output $\rho_B \in \mathcal{H}_B$ [5]. In general, practical quantum channels are noisy, which means that the output state cannot carry the same amount of information as the input [6–8]. Given a quantum channel, an important task is to find its *quantum capacity*, which gives the ultimate limit of the achievable information transmission rate using the channel.

In practice, the exact expression of quantum capacity is generally hard to find [3, 9–11], thus a great deal of effort is devoted to improving its upper or lower bound [12–14]. Quantum erasure channel is one of the few quantum channels whose quantum capacity is known. By definition, the quantum erasure channel $\mathcal{N}_p$ erases the input state ρ with probability p while keeping the input state unchanged with chance $1 - p$ [15, 16],

$$\mathcal{N}_p(\rho) \to (1-p)\rho + p|e\rangle\langle e|, \quad (1)$$

where $|e\rangle$ is a flag state that is orthogonal to the input Hilbert space. Denote d as the dimension of the input space, the erasure channel has a quantum capacity $Q = \max\{(1 - 2p)\log_2 d, 0\}$ [16], which means positive information transmission rate is achievable if the erasure probability is less than half.

As we noticed, since the first appearance of quantum erasure channel [15], almost all discussions are based on discrete variables (DV), while less is investigated in the continuous variable (CV) setting. As CV quantum information processing shows more and more potentials
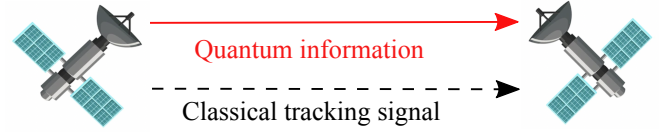
* zhong.changchun@uchicago.edu

FIG. 1. The quantum erasure channel realized by sending classical tracking signal along with the quantum information. The detection of the classical signal heralds the quantum information transmission.

[17–20], it is important to identify the CV channels or the corresponding error models. In Ref. [21], a CV model of "quantum erasure channel" is first considered, which either transmits a perfect coherent state or replaces the input by a vacuum, denoted as $\mathcal{N}(|\alpha\rangle\langle\alpha|) = (1 - p)|\alpha\rangle\langle\alpha| + p|0\rangle\langle 0|$.

Although interesting, this CV model does not faithfully capture the essence of an erasure channel since the coherent state and vacuum state are not orthogonal. In this paper, we study the CV quantum erasure channel more rigorously, where an arbitrary CV input state is either completely erased with a chance p or perfectly transmitted with a probability $1 - p$, given as

$$\mathcal{N}(\rho_{\mathrm{cv}} \otimes |\uparrow\rangle\langle\uparrow|) \to (1-p)\rho_{\mathrm{cv}} \otimes |\uparrow\rangle\langle\uparrow| + p|0\rangle\langle 0| \otimes |\downarrow\rangle\langle\downarrow|, \quad (2)$$

where the spin state is used as a flag to signal the erasure and is important to rigorously define the quantum erasure channel. Moreover, this channel is also practical which can be realized by introducing classical tracking signals along the quantum encoding. The detection of classical signal at the receiving end is used to herald the quantum information transmission, as sketched in Fig. 1. Note that the technique with classical tracking signals is already used in heralding connections between quantum satellites [22].

For a DV quantum erasure channel, it is well known that the channel capacity is in principle achievable with a random channel coding [23], as depicted in Fig. 2(a), where the quantum information carried by the input system A is unitarily scrambled with total spins N. In the

large limit N, accessing the $N(1-p)$ spins in the output can recover the scrambled quantum information, respecting the channel capacity of the DV quantum erasure channel. This recovery is indeed possible if the reference system R and the remaining Np qubits are close to a product state, which is the essence of decoupling theory [24]. This theory plays a central role in the well-known Hayden-Preskill thought experiment, where one tries to hide information by tossing it into a black hole but unfortunately the information will be quickly reflected out when the black hole is at a certain stage [25].

Naturally, one would consider investigating a similar channel coding model by replacing the spins with Bosonic modes that simulate the CV erasure channel. However, an immediate problem shows up when one notices that the CV system is infinite dimensional. The random coding scheme would involve infinite energy or noncompact group operations [26, 27], making the further discussion impossible. One might simply think to set a cutoff to get a finite dimensional Hilbert space such that the decoupling reduces to the finite dimensional situation. However, in this case, a maximally mixed state is the best encoding that achieves the optimal information transmission rate $(1-2p)\log_2 n_c$ (n_c is the cutoff on the Fock basis) [28]. With the same amount of energy carried by the maximally mixed state, a thermal state has a larger entanglement entropy, which indicates that it would be better to encode a thermal state in the CV setting without doing any cut-off, as long as we can correctly identify the unitary random scrambling.

In this paper, we will obtain the proper scrambling unitary which is essential for using the CV decoupling model to understand the CV erasure channel. By defining an effective dimension of a general Bosonic mode, we are able to get the correct Haar random unitary restricted to the effective Hilbert space. Using decoupling formalism in quantum channel theory, we show the information transmission rate with the CV erasure channel is upper bounded by $(1-2p)\mathcal{E}$, where $\mathcal{E}$ represents mode entanglement that is specifically related to the mode's effective dimension. This bound essentially gives the quantum channel capacity since random channel coding is optimal. Our discussion of the CV erasure channel naturally extends the original Hayden-Preskill thought experiment to the continuous variable case, and probably is more precise in the setting of black hole since the CV model predicts that each radiated mode is thermal, matching the property of Hawking radiation.

II. QUANTUM CHANNEL AND THE DECOUPLING THEOREM

Mathematically, a CPTP map that describes a quantum channel $\mathcal{N} : \rho_A \to \rho_B$ always has a unitary dilation given by [29]

$$\mathcal{N}(\rho_A) \equiv \mathrm{tr}_E[\mathcal{U}_{AE}(\rho_A \otimes |0\rangle \langle 0|_E)] \qquad (3)$$

where the unitary channel $\mathcal{U}_{AE}$ acts on a dilated Hilbert space H_{AE}. By convention, the subscript E is used to denote the environment degree of freedom. Since unitary evolution correlates the system and the environment, the system output usually cannot recover all the information from the input. However, it is shown that error correction for restoring all information is possible if certain decoupling condition is satisfied [24, 25, 30, 31].

Theorem 1. *Denote $|\psi\rangle_{RA}$ as a purification of the input ρ_A (R is generally called the reference system), we have a unitary channel $I_R \otimes \mathcal{U}_{AE}$ acting on the input $|\psi\rangle_{RA} \otimes |0\rangle_E$*

$$|\psi\rangle_{RBE} \langle\psi| = I_R \otimes \mathcal{U}_{AE}(|\psi\rangle_{RA} \langle\psi| \otimes |0\rangle_E \langle 0|). \qquad (4)$$

Perfect entanglement recovery is possible if and only if the reduced density matrix of the reference and the environment are decoupled in the following sense

$$\rho_{RE} = \rho_R \otimes \rho_E. \qquad (5)$$

This theorem is first given in Refs. [30, 31] as a condition for perfect quantum error correction. The "if part" of the theorem is interesting. Intuitively, it means that if the reference system is not correlated with the environment, its entanglement with the system must be preserved in the output. For convenience, we put the "if part" proof here and more details should resort to Ref. [30].

Proof. Suppose that ρ_R and ρ_E have the following spectral decompositions:

$$\rho_R = \sum_k \lambda_k |k^R\rangle \langle k^R|, \rho_E = \sum_l \mu_l |l^E\rangle \langle l^E|. \qquad (6)$$

Then the pure state $|\psi\rangle_{RBE}$ can be written in the form

$$|\psi\rangle_{RBE} = \sum_{kl} \sqrt{\lambda_k \mu_l} |k^R\rangle |\phi_{kl}^B\rangle |l^E\rangle. \qquad (7)$$

The states $|\phi_{kl}^B\rangle$ are orthonormal, respecting the Schmidt decomposition. Now we can design a local projection measurement on B

$$\Pi_l = \sum_k |\phi_{kl}^B\rangle \langle \phi_{kl}^B|, \qquad (8)$$

where each measurement result l occurs with probability μ_l. The post-state conditioning on l is

$$|\psi_{RBE}^l\rangle = \sum_k \sqrt{\lambda_k} |k^R\rangle |\phi_{kl}^B\rangle |l^E\rangle. \qquad (9)$$

Then a unitary transformation $U_l |\phi_{kl}^B\rangle = |k^B\rangle$ yields the state

$$U_l |\psi_{RBE}^l\rangle = \sum_k \sqrt{\lambda_k} |k^R\rangle |k^B\rangle \otimes |l^E\rangle. \qquad (10)$$

Obviously, $|\psi_{RB}\rangle = \sum_k \sqrt{\lambda_k} |k^R\rangle |k^B\rangle$ carries the exactly same amount of entanglement as the state $|\psi_{RA}\rangle$. $\qquad \square$
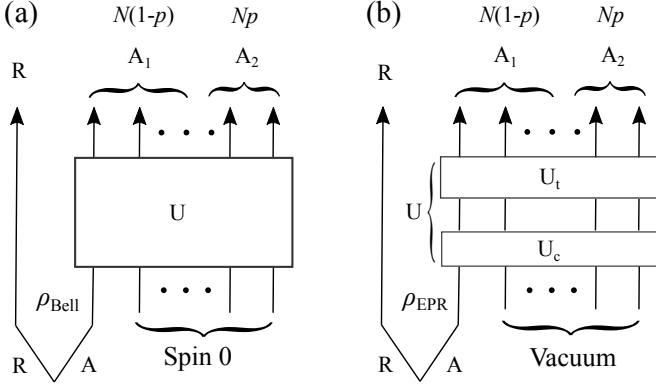
FIG. 2. The decoupling model of quantum erasure channel for: (a) discrete variable channel where quantum information is scrambled by the Haar random unitary; (b) continuous variable channels where photons are first equally distributed among all the Bosonic modes by the passive random unitary U_c, then further scrambled by the Haar random unitary U_t confining to the effective subspace.

For any quantum channel, this decoupling condition is the requirement for perfectly transmitting input states. The decoupling condition is not always satisfied. However, if the condition is approximately true, the entanglement is also approximately recoverable. Formally, it can be rigorously shown in the following theorem:

Theorem 2. *Let $\mathcal{V}_{\mathcal{N}}^{A \to BE}$ be a Stinespring extension of the channel $\mathcal{N}_{A \to B}$. Let ρ_{RA^n} be any encoding with dimension $|R| = 2^{nQ}$. The output state $\rho_{RB^n E^n} = \mathcal{V}_{\mathcal{N}}^{\otimes n}(\rho_{RA^n})$. Then there exists a decoding map $\mathcal{D}_{B^n \to \hat{R}}$ which, together with the encoding ρ_{RA^n}, comprises a (Q, n, ϵ) entanglement generation code for the channel, provided that*

$$\|\rho_{RE^n} - \Pi_R \otimes \varphi_{E^n}\|_1 \le \epsilon, \tag{11}$$

where $\Pi_R = \frac{1}{|R|}I$.

This theorem and proof can be found in Ref. [24]. Arbitrary small ϵ can be achieved for large n, indicating that the decoupling between the reference and the environment is asymptotically satisfied as one increases the number of channel usage. Thus it guarantees the existence of decoding operations on the output state to restore almost all the entanglement.

III. DECOUPLING MODEL FOR CV ERASURE CHANNEL

In the DV decoupling model, generally one party of the entangled qubit (qudit) systems is sent down to a random unitary channel, as shown in Fig. 2(a). Quantum information is fully scrambled, so that certain local access to the scrambled system cannot fully recover the information, respecting the decoupling theorem [32, 33].

Using the decoupling theorem, one can further obtain the expression of quantum capacity for quantum erasure channel [24, 25, 34].

A. The CV decoupling model

We want to extend the discussion to the CV quantum erasure channel. Naively, one would think that we just need to replace those qubits with infinite dimensional modes, e.g., Bosonic modes. For instance, as shown in Fig. 2(b), assuming two Bosonic modes are in a two-mode squeezed state ρ_{RA}^{EPR}, one of which (A) is sent to a random encoding circuit for information scrambling. The next task then is to find the maximal number of modes that can remain decoupled from the reference mode R. As in the DV case, this model is indeed simulating a CV quantum erasure channel with a random coding scheme. However, we need to identify what kind of unitary to be used, which turns out to be non-trivial for CV systems.

Unlike the DV case, the CV system is infinite dimensional, which might involve unbounded energy [27]. The unitary group for fully scrambling the system is generally noncompact [26, 35], e.g., the Gaussian unitary with squeezing forms a noncompact group. A natural guess is to use random passive unitary U_c (without squeezing, it forms a compact group) to scramble those Bosonic modes. However, a close examination could quickly give us a negative answer because the passive random unitary does not form a unitary design [36]. Actually, even including the squeezing, the Gaussian unitary is not a perfect unitary design [26], which cannot fully scramble the quantum information. It is then believed that some nontrivial (certainly non-Gaussian) unitary is required in the CV random coding, and we shall try identifying it in the following sections.

B. The effective dimension of CV system

In Ref. [26], many-mode Bosonic system scrambling is discussed in the phase space, where so-called quasi or genuine scrambling is achieved by approximated CV unitary designs. In this paper, we discuss the mode scrambling in a different approach. As mentioned, a Bosonic mode has infinite dimension that could carry infinite energy, while in practice, each mode usually has a finite amount of energy and can be effectively described as a finite dimensional system.

Theorem 3. *A single mode Bosonic state ρ with von Neumann entropy $\mathcal{E} = -tr\rho \log \rho$ has an effective dimension*

$$D = 2^{\mathcal{E}}. \tag{12}$$

Note that this theorem just means that the mode can be effective described by a D-dimensional qudit in the sense of identifying the entanglement.

Proof. Expressing the state ρ in its diagonal basis (spectrum decomposition)

$$\rho = \sum_\lambda p_\lambda |\lambda\rangle \langle\lambda|, \qquad (13)$$

where p_λ is the probability corresponding to $|\lambda\rangle$. Imagine that we have a quantum source emitting this state and we can detect the state in this basis. In general, the emitting state is written as (assuming K copies)

$$\rho^{\otimes K} = \sum_{\lambda_1,...,\lambda_K} p_{\lambda_1}...p_{\lambda_K} |\lambda_1\rangle \langle\lambda_1| \otimes ... |\lambda_K\rangle \langle\lambda_K|. \quad (14)$$

We would be able to define a typical subspace as K becomes very large. For example, if we measure each mode on the basis, we will get a string of eigenvalues $i_1 i_2...i_j...i_K$ with $1 \le j \le K, 0 \le i_j < \infty$. Each string has the probability $p_{i_1} p_{i_2}...p_{i_j}...p_{i_K}$. In those strings, there is a set in which the strings are more probable than the strings outside the set. This set is usually called a typical set, and the probability for each string in the set is given as

$$p(\text{typical string}) = \prod_{i_j} p_{i_j}^{K p_{i_j}}. \qquad (15)$$

To identify the rank of the typical set, we take the logarithm on both sides of Eq. 15 and obtain

$$p(\text{typical string}) = 2^{K \sum_{i_j=0}^\infty p_{i_j} \log_2 p_{i_j}} \\ = 2^{-K\mathcal{E}}, \qquad (16)$$

where $\mathcal{E} = -\sum_{i_j=0}^\infty p_{i_j} \log_2 p_{i_j}$ is the von Neumann entropy of the state. (Note more rigorously, one should define the typical set by bounding small errors, e.g., $2^{-(K\mathcal{E}+\epsilon)} \le p(\text{typical}) \le 2^{-(K\mathcal{E}-\epsilon)}$, and here we take the ideal limit for demonstrating the main ideas, which is good enough for our purpose.) Obviously, each string in the typical set is equally probable, and since the probability sums to one (note that the string outside the typical set almost never happens [37, 38]), the total number of typical strings is $2^{K\mathcal{E}}$. Alternatively, we can think these K modes span a subspace with dimension $2^{K\mathcal{E}}$, thus each mode effectively contribute to the dimension with $2^{\mathcal{E}}$. □

IV. THE DECOUPLING AND QUANTUM CAPACITY

As shown in Fig. 2(b), we consider the model in which a reference R is entangled with the system A. Without losing generality, we assume that there are K Bosonic modes in R and $N \gg K$ modes in A, and they are initially entangled by K pairs of two-mode squeezed state $\rho_{\text{EPR}}^{\otimes K}$ (only one pair is shown in the figure). The rest of $N - K$ modes in A are initially set in the vacuum. We then apply a random unitary on the system A, hoping to

scramble the information in the N modes such that discarding a finite number of modes in A will not decrease the entanglement with R. Suppose that for the initial K two-mode squeezed states, each pair carries entanglement $\mathcal{E}(r)$, which is given by

$$\mathcal{E}(r) \equiv \cosh^2(r) \log_2 \cosh^2(r) - \sinh^2(r) \log_2 \sinh^2(r), \qquad (17)$$

where r denotes the strength of the squeeze. Each mode in the two-mode squeezed state contains on average $\sinh^2 r$ photons.

In the system A, we first apply passive random unitary U_c (random beam splitter and phase shifter) to redistribute $K \sinh^2 r$ photons to totally N modes, and on average each mode has $K \sinh^2 r/N$ photons. Note the passive random unitary on average does not thermalize the state in the sense (see the appendix for more details)

$$\rho'_{RA} \equiv \int dU_c U_c \rho_{RA} U_c^\dagger \ne \rho_R \otimes \rho_A^{\text{thermal}}. \qquad (18)$$

Instead, we have $\rho'_{RA} = \rho_R \otimes (\tilde\rho^{\otimes N})_A$ where the state $\tilde\rho$ for each different mode is exactly the same and is diagonal on the Fock basis

$$\tilde\rho = \sum_n p_n |n\rangle \langle n| \qquad (19)$$

since the random phase shifter washes out all correlations in that basis. The entanglement of each mode is given by $\mathcal{E}(p_n) = -\sum_n p_n \log_2 p_n$. Accordingly, the typical subspace of A has a finite dimension

$$d = 2^{N\mathcal{E}(p_n)} \qquad (20)$$

which can effectively describe the infinite dimensional Bosonic modes. Next, we apply a second random unitary U_t on the system A that is confined to this typical subspace, and we shall show below that the decoupling inequality can be recovered. Note that according to the equipartition theorem in the large mode limit [38], the typical projection measurement will have an approximately uniform distributed outcome. The second random unitary U_t essentially gives an uniform distribution of the typical state, and each mode is close to a thermal state with photon number $\sinh^2 r_0 = K \sinh^2 r/N$. The corresponding entanglement is denoted $\mathcal{E}(r_0)$, as detailed in the Appendix.

Now, we divide the system A into two parts A_1 and A_2. A_2 contains pN modes ($0 < p < 1$), which is to be discarded. The question is how many modes in A can we discard while keeping the remaining modes almost completely entangled with the reference R. The answer is summarized in the following theorem:

Theorem 4. *The average trace distance between the randomized state on RA_2 and a product state $\rho_R \otimes \prod_{A_2}$ satisfies the following inequality*

$$\int dU \|tr_{A_1}[U\rho_{RA}U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_1 \le 2^{-\frac{1}{2}(Q-\gamma)N}, \quad (21)$$

where the unitary $U = U_t U_c$ acts on the system A. $Q = (1 - 2p)\mathcal{E}(r_0)$ is the energy-constrained quantum capacity and $\gamma = K\mathcal{E}(r)/N$ is the quantum information transmission rate. The product state $\rho_R \otimes \Pi_{A_2}$ is defined as the unitary average

$$\rho_R \otimes \Pi_{A_2} \equiv tr_{A_1} \int dU_t U_t \rho'_{RA} U_t^\dagger. \tag{22}$$

In the language of the decoupling principle, to recover all initial entanglement from the system RA_1, we need the term $2^{-\frac{1}{2}(Q-\gamma)N}$ to be small, which is achievable in the large N limit if $\gamma < Q$. This clearly indicates that all the information transmission rate γ is achievable as long as it is smaller than the quantum channel capacity.

To prove the inequality, we start from the following equality: for arbitrary state ρ, the average over Haar random unitary confined to the typical subspace satisfies

$$\int dU_t U_t \rho U_t^\dagger = \frac{1}{d} I. \tag{23}$$

where d is the typical space dimension defined above and I is the corresponding identity operator. Intuitively, it is because the integral must commute with any unitary operation (the property of Haar measure), thus only identity is possible. Rigorously, we have

$$\int dU_t U_t \rho U_t^\dagger = \int dU_t \sum_{i_1 j_1, i'_1 j'_1} U^t_{i_1 j_1} \rho_{j_1 i'_1} U^{t\dagger}_{i'_1 j'_1} \tag{24}$$

$$= \sum_{i_1 j_1, i'_1 j'_1} \rho_{j_1 i'_1} \int dU_t U^t_{i_1 j_1} U^{t\dagger}_{i'_1 j'_1} \tag{25}$$

$$= \sum_{i_1 j_1, i'_1 j'_1} \rho_{j_1 i'_1} \frac{1}{d} \delta_{i_1 j'_1} \delta_{j_1 i'_1} \tag{26}$$

$$= \sum_{i_1 j_1} \rho_{j_1 j_1} \frac{1}{d} \delta_{i_1 i_1} = \sum_{i_1} \frac{\sum_{j_1} \rho_{j_1 j_1}}{d} \delta_{i_1 i_1} \tag{27}$$

$$= \frac{1}{d} \sum_{i_1} \delta_{i_1 i_1} |i_1\rangle \langle i_1|, \tag{28}$$

where the basis is omitted from the above calculation for simplicity and is recovered in the last line 28. Note in Eq. 22, the random unitary is acting on the system A, and we expect on average the state to be

$$\int dU U \rho_{RA} U^\dagger$$

$$= \int dU_t U_t \left(\int dU_c U_c \rho_{RA} U_c^\dagger \right) U_t^\dagger \tag{29}$$

$$= \rho_R \otimes \frac{1}{d} I_A = \rho_R \otimes \Pi_A,$$

where Π_A is to denote a maximally mixed state in the typical subspace. Obviously, it is also true if we confine the system to A_2

$$\text{tr}_{A_1} \int dU U \rho_{RA} U^\dagger = \rho_R \otimes \frac{1}{d} I_{A_2} = \rho_R \otimes \Pi_{A_2}. \tag{30}$$

which is the expression defined in Eq. 22. This result basically means that the correlation between R and A will be averaged out by the random unitary, and the information is completely scrambled in the sense that A is set in a maximally mixed state. Note that for each unitary realization, it is not necessarily true. In the following, we want to show that for each unitary realization, the distance between $\text{tr}_{A_1}[U \rho_{RA} U^\dagger]$ and $\rho_R \otimes \Pi_{A_2}$ is typically small in the limit of large N Bosonic modes. This is guaranteed if we can somehow bound the variance as in Eq. 21 (the average trace distance) as follows

$$\int dU \|\text{tr}_{A_1}[U \rho_{RA} U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_1 \tag{31}$$

$$\leq \int dU \sqrt{d_{RA_2}} \|\text{tr}_{A_1}[U \rho_{RA} U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_2 \tag{32}$$

$$= \int dU \sqrt{d_{RA_2}} \sqrt{\text{tr}_{RA_2}(\text{tr}_{A_1}[U \rho_{RA} U^\dagger] - \rho_R \otimes \Pi_{A_2})^2} \tag{33}$$

$$\leq \sqrt{d_{RA_2}} \sqrt{\text{tr}_{RA_2} \int dU (\text{tr}_{A_1}[U \rho_{RA} U^\dagger] - \rho_R \otimes \Pi_{A_2})^2} \tag{34}$$

$$= \sqrt{d_{RA_2} \text{tr}_{RA_2} \int dU \left[(\text{tr}_{A_1} U \rho_{RA} U^\dagger)^2 - (\rho_R \otimes \Pi_{A_2})^2 \right]} \tag{35}$$

$$= \sqrt{d_{RA_2} \left(\text{tr}_{RA_2} \int dU (\text{tr}_{A_1} U \rho_{RA} U^\dagger)^2 - \text{tr}_{RA_2}(\rho_R \otimes \Pi_{A_2})^2 \right)}. \tag{36}$$

The line 32 is obtained since the trace norm is upper bounded by the Hilbert-Schmidt norm up to the dimen-

sion factor d_{RA_2}. The line 34 is due to the Jensen's in-equality. Obviously, we only need to evaluate one term involving the random unitary, which is

$$\text{tr}_{RA_2} \int dU (\text{tr}_{A_1} U \rho_{RA} U^\dagger)^2 \tag{37}$$

$$= \int dU \text{tr}_{RAR'A'} \Big[\rho_{RA}(U) \otimes \rho_{R'A'}(U) F_{RR'} F_{A_2 A'_2} I_{A_1 A'_1} \Big] \tag{38}$$

$$= \int dU_c \text{tr}_{RAR'A'} \Big[\rho_{RA}(U_c) \otimes \rho_{R'A'}(U_c) F_{RR'} \int dU_t U_t^{\dagger \otimes 2} F_{A_2 A'_2} I_{A_1 A'_1} U_t^{\otimes 2} \Big], \tag{39}$$

where the notation $\rho(U) = U\rho U^\dagger$ is used and we use F to denote a swap operator. The line 38 is due to the swap technique

$$\text{tr}_A(\sigma_A^2) = \text{tr}_{AA'BB'}[\sigma_{AB} \otimes \sigma_{A'B'} F_{AA'} I_{BB'}]. \tag{40}$$

We see the last integral $\int dU U^\dagger \otimes U^\dagger F_{A_2 A'_2} I_{A_1 A'_1} U \otimes U$ is all we need that involves random unitary in the typical space. Now we use the double-twirling formula [39] (which can be directly obtained from the Weingarten calculus)

$$\int dU_t U_t^\dagger \otimes U_t^\dagger X U_t \otimes U_t$$
$$= \left[\text{tr}(X) - \frac{\text{tr}(X F_{AA'})}{d} \right] \frac{I_{AA'}}{d_A^2 - 1} + \left[\text{tr}(X F_{AA'}) - \frac{\text{tr}(X)}{d} \right] \frac{F_{AA'}}{d_A^2 - 1}$$
$$= \frac{1}{d_{A_2}} \frac{1 - 1/d_{A_1}^2}{1 - 1/d_A^2} I_{AA'} + \frac{1}{d_{A_1}} \frac{1 - 1/d_{A_2}^2}{1 - 1/d_A^2} F_{AA'} \tag{41}$$
$$\leq \frac{1}{d_{A_2}} I_{AA'} + \frac{1}{d_{A_1}} F_{AA'}$$

where we take $X = F_{A_2 A'_2} \otimes I_{A_1 A'_1}$. In the last equality, we used the results

$$\text{tr}(X) = d_{A_1}^2 d_{A_2}, \text{tr}(X F_{AA'}) = d_{A_1} d_{A_2}^2. \tag{42}$$

The last inequality is obtained by considering the fact that the typical dimensions for the system A and subsystems $A_{1,2}$ are much larger than one. Now we can continue to evaluate Eq. 39, which is

$$\int dU_c \text{tr}_{RAR'A'} \Big[\rho_{RA}(U_c) \otimes \rho_{R'A'}(U_c) F_{RR'} \int dU_t U_t^{\dagger \otimes 2} F_{A_2 A'_2} I_{A_1 A'_1} U_t^{\otimes 2} \Big] \tag{43}$$

$$\leq \int dU_c \text{tr}_{RAR'A'} \Big[\rho_{RA}(U_c) \otimes \rho_{R'A'}(U_c) F_{RR'} (\frac{1}{d_{A_2}} I_{AA'} + \frac{1}{d_{A_1}} F_{AA'}) \Big] \tag{44}$$

$$= \frac{1}{d_{A_2}} \text{tr}_R \rho_R^2 + \frac{1}{d_{A_1}} \text{tr}_{AR} \rho_{AR}^2. \tag{45}$$

Then we are ready for the calculation of Eq. 36 using the integral we just derived

$$\sqrt{d_{RA_2}} \sqrt{\text{tr}_{RA_2} \int dU (\text{tr}_{A_1} U \rho_{RA} U^\dagger)^2 - \text{tr}_{RA_2} (\rho_R \otimes \Pi_{A_2})^2}$$

$$\leq \sqrt{d_{RA_2}} \sqrt{\frac{1}{d_{A_2}} \text{tr}_R \rho_R^2 + \frac{1}{d_{A_1}} \text{tr}_{AR} \rho_{AR}^2 - \text{tr}_{RA_2} (\rho_R \otimes \Pi_{A_2})^2}$$

$$= \sqrt{\frac{d_{RA_2}}{d_{A_1}} \text{tr}_{AR} \rho_{AR}^2}, \tag{46}$$

where the line is obtained since $\text{tr}_{RA_2}(\rho_R \otimes \Pi_{A_2})^2 = \text{tr}_R \rho_R^2 / d_{A_2}$. To summarize, we finally obtain the upper bound

$$\int dU \|\text{tr}_{A_1}[U\rho_{RA}U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_1 \leq \sqrt{\frac{d_{RA_2}}{d_{A_1}} \text{tr}_{AR}\rho_{AR}^2}, \tag{47}$$

which is the main result we have for the CV system using effective dimension. Since we are starting from a pure state ρ_{AR}, we have $tr\rho_{AR}^2 = 1$. Also the typical space dimension of A_1 and RA_2 are given by

$$d_{A_1} = 2^{(1-p)N\mathcal{E}(r_0)}, d_{RA_2} = 2^{K\mathcal{E}(r)+pN\mathcal{E}(r_0)}. \tag{48}$$

Thus the upper bound is evaluated to be

$$\int dU \|\text{tr}_{A_1}[U\rho_{RA}U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_1 \tag{49}$$

$$\leq \sqrt{\frac{d_{RA_2}}{d_{A_1}}} = 2^{\frac{K}{2}\mathcal{E}(r)-(\frac{1}{2}-p)N\mathcal{E}(r_0)}. \tag{50}$$

The above inequality can be understood in the quantum channel language: we initially have $K\mathcal{E}(r)$ ebits entanglement and try to send the entanglement through N quantum erasure channels with erasure probability p. Obviously, the entanglement transmission rate is $\gamma \equiv K\mathcal{E}(r)/N$. Then the upper bound can be rewritten in the form

$$\int dU \|\text{tr}_{A_1}[U\rho_{RA}U^\dagger] - \rho_R \otimes \Pi_{A_2}\|_1 \leq 2^{-[(\frac{1}{2}-p)\mathcal{E}(r_0)-\frac{\gamma}{2}]N}, \tag{51}$$

which is the main result of this paper. The proof to the inequality Eq. 21 is now complete.

We see if $(1/2 - p)\mathcal{E}(r_0) - \gamma/2 > 0$, the bound goes to zero in the $N \to \infty$ limit. According to the decoupling theorem, the entanglement can be recovered by decoding A_1 (A_2 discarded), which means that any entanglement transmission rate satisfying $\gamma < (1 - 2p)\mathcal{E}(r_0)$ is asymptotically achievable. Since $\mathcal{E}(r_0)$ essentially constrains the input energy for each mode, we define

$$Q \equiv (1 - 2p)\mathcal{E}(r_0) \tag{52}$$

as the *energy-constrained* quantum channel capacity for continuous variable quantum erasure channel using the random coding scheme. Interestingly, $1 - 2p$ is actually the quantum capacity of qubit erasure channel with erasure probability p, which quantitatively matches the special case when $\mathcal{E}(r_0) = 1$.

V. DISCUSSION AND OUTLOOK

The DV erasure channel was used in the Hayden-Preskill thought experiment [25], where black holes turn out to be a bad keeper of secret information. As shown in Fig. 3, Alice (A) tries to hide her secret in a black hole, assuming the black hole has a fast enough dynamics to scramble the information. Quite in contrast, if an
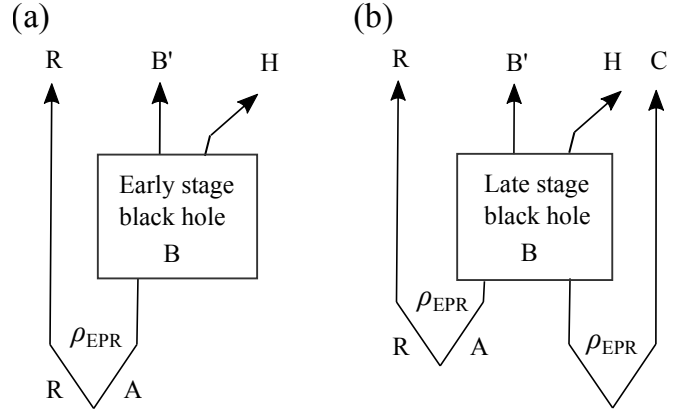

FIG. 3. Black holes as mirrors: (a) Alice's information (A) is tossed into an early stage black hole (B). Her information will be revealed as soon as the Hawking radiation (H) accumulates half the degrees of freedom of the black hole; (b) Alice's information is tossed into a late-stage black hole, where Charlie (C) has already collected a certain amount of Hawking radiation thus is entangled with the black hole. Her information will quickly revealed to Charlie when a little bit more black hole radiation is collected.

eavesdropper collects the Hawking radiation to a certain degree, i.e., a little bit more than half of the black hole degrees of freedom, he could know Alice's secret accurately. If the black hole was in its late stage, i.e., half of the black hole was evaporated, and the eavesdropper has monitored the black hole dynamics and collected all the Hawking radiation, then Alice's secret will be quickly revealed to the eavesdropper as long as a little bit more Hawking radiation is obtained. The reason is deep rooted in the quantum capacity of the DV erasure channel. With the Theorem 4 that we prove in this paper, this thought experiment naturally extends to the CV setting. More interestingly, the CV model might be more proper in describing the black hole dynamics, since the Hawking radiation is believed to be thermal. Because each Bosonic mode is in a thermal state after random scrambling, it thus correctly models the black hole radiation.

On a glance of the literature, it is easy to find that the DV-system-based theory of quantum information and computation is more advanced, while the CV counterpart is less developed. Admittedly, the recent decade sees the gradual progress of CV quantum information processing [17, 18], e.g., encoding and error correcting quantum information with harmonic oscillators [20, 40]. Although it is still hard to get the full controllability of the CV system, its infinite dimension with potential to encode larger amounts of quantum information is appealing and is usually called hardware efficient for a single system. Thus identifying the error model of CV channels and the corresponding correcting schemes are essential.

As well known, evaluating the exact capacity of quantum channels is notoriously difficult that involves optimizing two-letter functions [3, 41]. The decoupling theorem looks into the capacity in a different way, directly

calculating a random coding scheme, which is used a lot in discrete variable quantum channels. By confining the random unitary on the typical subspace of the CV systems, we derived the capacity of the CV quantum erasure channel, which avoids the problem of non-compact unitary scrambling. We expect the technique developed in the paper to shed new light on exploring other types of CV quantum channel, e.g., the long-standing capacity bound problem of the well-known Bosonic loss channels, and we leave that for future investigations.

[1] M. Hayashi, S. Ishizaka, A. Kawachi, G. Kimura, and T. Ogawa, *Introduction to quantum information science* (Springer, 2014).

[2] J. Watrous, *The theory of quantum information* (Cambridge university press, 2018).

[3] L. Gyongyosi, S. Imre, and H. V. Nguyen, A survey on quantum channel capacities, IEEE Communications Surveys & Tutorials **20**, 1149 (2018).

[4] C. H. Bennett and P. W. Shor, Quantum information theory, IEEE transactions on information theory **44**, 2724 (1998).

[5] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum measurement*, Vol. 23 (Springer, 2016).

[6] A. S. Holevo, The capacity of the quantum channel with general signal states, IEEE Transactions on Information Theory **44**, 269 (1998).

[7] H. Barnum, M. A. Nielsen, and B. Schumacher, Information transmission through a noisy quantum channel, Phys. Rev. A **57**, 4153 (1998).

[8] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A **55**, 1613 (1997).

[9] J. Eisert and M. M. Wolf, Gaussian quantum channels, arXiv preprint quant-ph/0505151 (2005).

[10] I. Devetak and P. W. Shor, The capacity of a quantum channel for simultaneous transmission of classical and quantum information, Communications in Mathematical Physics **256**, 287 (2005).

[11] A. S. Holevo, Quantum systems, channels, information, in *Quantum Systems, Channels, Information* (de Gruyter, 2019).

[12] M. Rosati, A. Mari, and V. Giovannetti, Narrow bounds for the quantum capacity of thermal attenuators, Nature communications **9**, 1 (2018).

[13] 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 Journal of Physics **20**, 063025 (2018).

[14] K. Jeong, Y. Lim, J. Kim, and S. Lee, New upper bounds on the quantum capacity for general attenuator and amplifier, in *AIP Conference Proceedings*, Vol. 2241 (AIP Publishing LLC, 2020) p. 020017.

[15] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, Phys. Rev. A **56**, 33 (1997).

[16] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Capacities of quantum erasure channels, Phys. Rev. Lett. **78**, 3217 (1997).

[17] S. L. Braunstein and P. Van Loock, Quantum information with continuous variables, Reviews of modern physics **77**, 513 (2005).

[18] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Reviews of Modern Physics **84**, 621 (2012).

[19] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Physical Review A **64**, 012310 (2001).

[20] W.-L. Ma, S. Puri, R. J. Schoelkopf, M. H. Devoret, S. Girvin, and L. Jiang, Quantum control of bosonic modes with superconducting circuits, Science Bulletin **66**, 1789 (2021).

[21] J. Niset, U. L. Andersen, and N. J. Cerf, Experimentally feasible quantum erasure-correcting code for continuous variables, Phys. Rev. Lett. **101**, 130503 (2008).

[22] J. S. Sidhu, S. K. Joshi, M. Gündoğan, T. Brougham, D. Lowndes, L. Mazzarella, M. Krutzik, S. Mohapatra, D. Dequal, G. Vallone, *et al.*, Advances in space quantum communications, IET Quantum Communication **2**, 182 (2021).

[23] R. Klesse, Approximate quantum error correction, random codes, and quantum channel capacity, Phys. Rev. A **75**, 062315 (2007).

[24] P. Hayden, M. Horodecki, A. Winter, and J. Yard, A decoupling approach to the quantum capacity, Open Systems & Information Dynamics **15**, 7 (2008).

[25] P. Hayden and J. Preskill, Black holes as mirrors: quantum information in random subsystems, Journal of high energy physics **2007**, 120 (2007).

[26] Q. Zhuang, T. Schuster, B. Yoshida, and N. Y. Yao, Scrambling and complexity in phase space, Phys. Rev. A **99**, 062334 (2019).

[27] M. Fukuda and R. Koenig, Typical entanglement for gaussian states, Journal of Mathematical Physics **60**, 112203 (2019).

[28] See the appendix for a brief review of the calculations for the discrete variable decoupling with any finite dimension.

[29] V. Paulsen, *Completely bounded maps and operator algebras*, 78 (Cambridge University Press, 2002).

[30] B. Schumacher and M. A. Nielsen, Quantum data processing and error correction, Phys. Rev. A **54**, 2629 (1996).

[31] B. Schumacher and M. D. Westmoreland, Approximate quantum error correction, Quantum Information Processing **1**, 5 (2002).

[32] F. Dupuis, The decoupling approach to quantum information theory, arXiv preprint arXiv:1004.1641 (2010).

[33] M. Horodecki, J. Oppenheim, and A. Winter, Quantum state merging and negative information, Communications in Mathematical Physics **269**, 107 (2007).

[34] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, Quantum error correction in scrambling dynamics and measurement-induced phase transition, Phys. Rev. Lett. **125**, 030505 (2020).

[35] B. Zhang and Q. Zhuang, Entanglement formation in continuous-variable random quantum networks, npj Quantum Information **7**, 1 (2021).

[36] A unitary design is a subset of the unitary group where the sample averages of certain polynomials over the set match that over the whole unitary group.

[37] C. E. Shannon, A mathematical theory of communication, The Bell system technical journal **27**, 379 (1948).

[38] M. M. Wilde, *Quantum information theory* (Cambridge University Press, 2013).

[39] B. Collins and P. Śniady, Integration with respect to the haar measure on unitary, orthogonal and symplectic group, Communications in Mathematical Physics **264**, 773 (2006).

[40] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Performance and structure of single-mode bosonic codes, Phys. Rev. A **97**, 032346 (2018).

[41] K. Brádler and C. Adami, Black holes as bosonic gaussian channels, Phys. Rev. D **92**, 025030 (2015).

Appendix A: The Weingarten calculus

The Weingarten calculus is important in random matrix theory, in which an integral is repeatedly used

$$\int dU U_{i_1 j_1} U_{i_2 j_2} ... U^{\dagger}_{j'_1 i'_1} U^{\dagger}_{j'_2 i'_2} ... = \sum_{\sigma,\tau} \delta_{i_1 i'_{\sigma(1)}} \delta_{i_2 i'_{\sigma(2)}} ... \delta_{j_1 j'_{\tau(1)}} \delta_{j_2 j'_{\tau(2)}} ... W_g(\sigma^{-1}\tau), \tag{A1}$$

where W_g is called Weingarten function that depends on the permutations σ and τ. In the main text, we already use this formula to derive $\int dU U \rho U^{\dagger} = I/d$. Another formula that appears frequently in decoupling theory is the double unitary twirling formula

$$\mathcal{T}(X) = \int dU U^{\dagger} \otimes U^{\dagger} X U \otimes U$$
$$= \frac{I}{d^2-1}\left[trX - \frac{Tr(XF)}{d}\right] + \frac{F}{d^2-1}\left[tr(XF) - \frac{trX}{d}\right] \tag{A2}$$

In many literature, this result is simply stated as the consequence of Schur-Weyl duality from group representation theory, which is not friendly to beginners. The Schur-Weyl duality might be needed to derive Eq. A1, but the double twirling formula Eq. A2 can be derived straightforwardly by applying Eq. A1, as shown below:

$$\mathcal{T}(X) = \int dU U^{\dagger} \otimes U^{\dagger} X U \otimes U$$
$$= \int dU \sum_{\substack{i_1 j_1 i_2 j_2 \\ i'_1 j'_1 i'_2 j'_2}} U_{i_1 j_1} U_{i_2 j_2} X^{j_1 j_2}_{i'_1 i'_2} U^{\dagger}_{i'_1 j'_1} U^{\dagger}_{i'_2 j'_2}$$
$$= \sum_{\substack{i_1 j_1 i_2 j_2 \\ i'_1 j'_1 i'_2 j'_2}} X^{j_1 j_2}_{i'_1 i'_2} \int dU U_{i_1 j_1} U_{i_2 j_2} U^{\dagger}_{i'_1 j'_1} U^{\dagger}_{i'_2 j'_2}$$
$$= \sum_{\substack{i_1 j_1 i_2 j_2 \\ i'_1 j'_1 i'_2 j'_2}} X^{j_1 j_2}_{i'_1 i'_2} \sum_{\sigma,\tau} \delta_{i_1 j'_{\sigma(1)}} \delta_{i_2 j'_{\sigma(2)}} \delta_{j_1 i'_{\tau(1)}} \delta_{j_2 i'_{\tau(2)}} W_g(\sigma^{-1}\tau) \tag{A3}$$
$$= \sum_{\substack{i_1 j_1 i_2 j_2 \\ i'_1 j'_1 i'_2 j'_2}} X^{j_1 j_2}_{i'_1 i'_2} [\delta_{i_1 j'_1} \delta_{i_2 j'_2} \delta_{j_1 i'_1} \delta_{j_2 i'_2} W_g(\sigma_1 \tau_1) + \delta_{i_1 j'_1} \delta_{i_2 j'_2} \delta_{j_1 i'_2} \delta_{j_2 i'_1} W_g(\sigma_1 \tau_2)$$
$$+ \delta_{i_1 j'_2} \delta_{i_2 j'_1} \delta_{j_1 i'_1} \delta_{j_2 i'_2} W_g(\sigma_2 \tau_1) + \delta_{i_1 j'_2} \delta_{i_2 j'_1} \delta_{j_1 i'_2} \delta_{j_2 i'_1} W_g(\sigma_2 \tau_2)]$$
$$= trX \frac{I}{d^2-1} + tr(XF)\frac{-I}{d(d^2-1)} + trX \frac{-F}{d(d^2-1)} + tr(XF)\frac{F}{(d^2-1)}$$
$$= \frac{I}{d^2-1}\left[trX - \frac{tr(XF)}{d}\right] + \frac{F}{d^2-1}\left[tr(XF) - \frac{trX}{d}\right].$$

Note for simplicity, we left out the basis during the calculation and recover that in the end. F is the swap operator acting on the tensor product space

$$F = \sum_{i_1, j'_1, i_2, j'_2} |i_1\rangle \langle j'_1| \otimes |i_2\rangle \langle j'_2| \delta_{i_1 j'_2} \delta_{i_2 j'_1}. \tag{A4}$$

Note the permutation group on a string with length two has only two elements $\{\sigma_1(1 \to 1, 2 \to 2), \sigma_2(1 \to 2, 2 \to 1)\}$, and similar for τ. We also used the Weingarten formula

$$W_g(\sigma_1\tau_1) = W_g(\sigma_2\tau_2) = \frac{1}{d^2 - 1}, \tag{A5}$$

$$W_g(\sigma_1\tau_2) = W_g(\sigma_2\tau_1) = \frac{-1}{d(d^2 - 1)}. \tag{A6}$$

Appendix B: Random passive operations does not thermalize

Let us consider a state

$$\rho_{\text{in}} = \rho_{\text{T}}(\bar{n}) \otimes |0\rangle\langle0| \otimes |0\rangle\langle0| \otimes \ldots |0\rangle\langle0| \otimes |0\rangle\langle0|, \tag{B1}$$

and we apply random passive unitary operations to the state, where $\rho_{\text{T}}(\bar{n})$ is a thermal state with mean photon number $\bar{n}$. We claim that after random passive unitary operations, it does not transform to a product of thermal states,

$$\int dU_c U_c \rho_{\text{in}} U_c^\dagger \neq \rho_{\text{T}}(\bar{n}/N) \otimes \cdots \otimes \rho_{\text{T}}(\bar{n}/N), \tag{B2}$$

where N is the number of Bosonic modes. To see this, let us compare the state projected on the single-photon subspace. First, the lhs is

$$\int dU_c U_c p_1(|1\rangle\langle1| \otimes |0\rangle\langle0| \otimes \cdots \otimes |0\rangle\langle0|)U_c^\dagger$$
$$= \frac{p_1}{N}(|1\rangle\langle1| \otimes |0\rangle\langle0| \otimes \cdots \otimes |0\rangle\langle0| + |0\rangle\langle0| \otimes |1\rangle\langle1| \otimes \cdots \otimes |0\rangle\langle0| + \cdots + |0\rangle\langle0| \otimes |0\rangle\langle0| \otimes \cdots \otimes |1\rangle\langle1|), \tag{B3}$$

where $p_1 = \frac{\bar{n}}{(\bar{n}+1)^2}$. On the other hand, the rhs is

$$q_1(|1\rangle\langle1| \otimes |0\rangle\langle0| \otimes \cdots \otimes |0\rangle\langle0| + |0\rangle\langle0| \otimes |1\rangle\langle1| \otimes \cdots \otimes |0\rangle\langle0| + \cdots + |0\rangle\langle0| \otimes |0\rangle\langle0| \otimes \cdots \otimes |1\rangle\langle1|), \tag{B4}$$

where $q_1 = \frac{\bar{n}/N}{(\bar{n}/N+1)^2}$. Since $p_1/N \neq q_1$, they are not equal.

In the limit of large number of modes, the state of N_1 ($N_1 \ll N_2 \equiv N - N_1$) modes after random average will be typically close to the thermal state in the following sense

$$\int dU_c \|\text{tr}_{N_2}[U_c \rho_{\text{in}} U_c^\dagger] - \rho_{\text{T}}^{\otimes N_1}\|_1 \ll \epsilon, \tag{B5}$$

which can be straightforwardly shown in the phase space and we refer the reader to the nice paper Ref. [27] for more details.

Each mode will be in diagonal form in the Fock basis, since the random phase shift in each mode gives

$$\frac{1}{2\pi} \int d\theta e^{i\theta\hat{a}^\dagger\hat{a}} \rho e^{-i\theta\hat{a}^\dagger\hat{a}}$$
$$= \frac{1}{2\pi} \int d\theta e^{i\theta\hat{a}^\dagger\hat{a}} \sum_{m,n} \rho_{mn} |m\rangle\langle n| e^{-i\theta\hat{a}^\dagger\hat{a}}$$
$$= \frac{1}{2\pi} \sum_{m,n} \int d\theta e^{i\theta(m-n)} \rho_{mn} |m\rangle\langle n|$$
$$= \sum_n \rho_{nn} |n\rangle\langle n|. \tag{B6}$$

Appendix C: Decoupling with truncated Fock space doesn't recover the CV erasure channel capacity

The typical subspace in the main text gives an efficient way of avoiding the infinite dimension problem in the CV random scrambling unitary. Naively, we might think we can achieve the same goal by simply truncating each mode in the Fock basis. In this appendix, we show this is not true.

Denote in each mode the truncation as n_c in the Fock basis. Thus, the EPR state can be approximated as

$$\rho_{EPR}^{n_c} = \sum_{n=0}^{n_c} \frac{\tanh^n(r)}{\cosh(r)} |n\rangle |n\rangle. \tag{C1}$$

The entanglement entropy of this state is denoted as $\mathcal{E}_c(r)$ which is in general smaller than $\mathcal{E}(r)$ (the true entanglement of the EPR state), but they will be very close as long as the cutoff number n_c is chosen to be large enough. Now in the same notation, we want to show the decoupling inequality in the discrete variable setting in the hope of revealing the CV erasure quantum capacity

$$\int dU \|\mathrm{tr}_{A_1}(U\rho_{RA}U^\dagger) - \rho_R \otimes \Pi_{A_2}\|_1 \le \epsilon. \tag{C2}$$

Note each of the N mode in system A is truncated to have dimension n_c and the random unitary U is acting on a Hilbert space with dimension n_c^N. Within this finite dimensional Hilbert space, we have

$$\int dU \|\mathrm{tr}_{A_1}(U\rho_{RA}U^\dagger) - \rho_R \otimes \Pi_{A_2}\|_1 \tag{C3}$$

$$\le \int dU \sqrt{d_{RA_2}} \|\mathrm{tr}_{A_1}(U\rho_{RA}U^\dagger) - \rho_R \otimes \Pi_{A_2}\|_2 \tag{C4}$$

$$= \sqrt{d_{RA_2}} \int dU \sqrt{\mathrm{tr}_{RA_2}[(\mathrm{tr}_{A_1}(U\rho_{RA}U^\dagger) - \rho_R \otimes \Pi_{A_2})^2]} \tag{C5}$$

$$\le \sqrt{d_{RA_2}} \sqrt{\int dU \mathrm{tr}_{RA_2}\left[(\mathrm{tr}_{A_1}[U\rho_{RA}U^\dagger])^2\right] - \mathrm{tr}_{RA_2}[(\rho_R \otimes \Pi_{A_2})^2]} \tag{C6}$$

$$\le \sqrt{d_{RA_2}} \sqrt{\frac{1}{d_{A_2}}\mathrm{tr}_R\rho_R^2 + \frac{1}{d_{A_1}}\mathrm{tr}_{RA}\rho_{RA}^2 - \mathrm{tr}_{RA_2}[(\rho_R \otimes \Pi_{A_2})^2]} \tag{C7}$$

$$= \sqrt{d_{RA_2}} \sqrt{\frac{1}{d_{A_1}}\mathrm{tr}_{RA}\rho_{RA}^2}, \tag{C8}$$

where we used the result $\int dU \mathrm{tr}_{RA_2}[(\mathrm{tr}_{A_1}[U\rho_{RA}U^\dagger])^2] \le \frac{1}{d_{A_2}}\mathrm{tr}_R\rho_R^2 + \frac{1}{d_{A_1}}\mathrm{tr}_{AR}\rho_{AR}^2$ in line C7. Taking the pure initial condition $\mathrm{tr}_{RA}\rho_{RA}^2 = 1$, we have

$$\int dU \|\mathrm{tr}_{A_1}(U\rho_{RA}U^\dagger) - \rho_R \otimes \Pi_{A_2}\|_1 \le \sqrt{\frac{d_{RA_2}}{d_{A_1}}}. \tag{C9}$$

The dimension factors on the rhs is determined by the truncation, which has nothing to do with the encoded CV state entanglement. While in the main text, the effective dimension is directly related to the encoded CV state entanglement, enabling us to obtain the CV erasure channel capacity. Actually, the rhs gives

$$\sqrt{\frac{d_{RA_2}}{d_{A_1}}} = 2^{-N(Q-\gamma)}, \tag{C10}$$

where $\gamma = \frac{K}{N}\log_2 n_c$ and $Q = (1-2p)\log_2 n_c$, indicating the n_c-dimensional Bell state is the optimal encoding in the decoupling scheme. Imposing energy constraint in the encoded state, we know thermal state would has the maximal entanglement entropy, which justify the CV decoupling scheme shown in the main text is better.

Appendix D: Each mode is a thermal state on average

As stated in the main text, after the random unitary each mode of A is on average in a thermal state with thermal photon $K\sinh^2 r/N$. We have

$$\int dU U\rho_{RA}U^\dagger = \int dU_t U_t \left(\int dU_c U_c \rho_{RA}U_c^\dagger\right) U_t^\dagger = \rho_R \otimes \frac{1}{d}I_A \equiv \rho_R \otimes \rho_A'. \tag{D1}$$

$\rho_A' = \frac{1}{d}I_A$ describes the state of A after the random unitary, which essentially is a maximally mixed state in the typical subspace. Note each typical vector in the typical subspace represents a state with an average photon number $K\sinh^2(r)$. Thus on average the state ρ_A' has a total photon number $K\sinh^2(r)$ and each mode in A has $K\sinh^2(r)/N$ photons.

Now we will further show the reduced density operator for each single mode in A is in a thermal state. Since ρ'_A is a maximally mixed state in the typical subspace, each typical state is equally probable, which essentially defines a micro-canonical ensemble with fixed energy, thus any small subsystem will looks like thermal. To show that, we express the state as

$$
\begin{aligned}
\rho'_A &= \sum_n P_n \, |\Phi_n\rangle \langle \Phi_n| \\
&= \sum_n P_n \sum_{i\alpha,j\beta} C^n_{i\alpha} C^{n*}_{j\beta} \, |\phi_i \psi_\alpha\rangle \langle \phi_j \psi_\beta| ,
\end{aligned}
\tag{D2}
$$

where $|\Phi_n\rangle$ denotes the typical state (energy eigen-states) and $P_n = 1/d$ is the probability (density of state independent of n). In order to get the reduced density matrix of a single mode, we separate it from the rest $N-1$ modes, and we take $|\Phi_n\rangle = \sum_{i\alpha} C^n_{i\alpha} |\phi_i \psi_\alpha\rangle$ where $|\phi_i\rangle$ and $|\psi_\alpha\rangle$ are the orthonormal basis for the single mode and the remaining $N-1$ modes, respectively. $C^n_{i\alpha}$ is the coefficient satisfying $\sum_{i,\alpha} |C^n_{i\alpha}|^2 = 1$. The reduced density state of the single mode can be obtained by tracing out the rest $N-1$ modes

$$
\rho'_1 = \mathrm{tr}_{N-1}[\rho'_A] \tag{D3}
$$

$$
= \sum_\gamma \langle \psi_\gamma | \sum_n P_n \sum_{i\alpha,j\beta} C^n_{i\alpha} C^{n*}_{j\beta} |\phi_i \psi_\alpha\rangle \langle \phi_j \psi_\beta| \psi_\gamma\rangle \tag{D4}
$$

$$
= \sum_n P_n \sum_{i\alpha,j\beta,\gamma} C^n_{i\alpha} C^{n*}_{j\beta} \delta_{\alpha\gamma} \delta_{\beta\gamma} |\phi_i\rangle \langle \phi_j| \tag{D5}
$$

$$
= \sum_n P_n \sum_{i,j,\gamma} C^n_{i\gamma} C^{n*}_{j\gamma} |\phi_i\rangle \langle \phi_j| \tag{D6}
$$

$$
= \sum_n P_n \sum_{i,j,\gamma} \langle \phi_i \psi_\gamma | \Phi_n\rangle \langle \Phi_n | \phi_j \psi_\gamma\rangle |\phi_i\rangle \langle \phi_j| \tag{D7}
$$

$$
= \frac{1}{d(E)} \sum_{i,j,\gamma} \delta_{ij} \delta_{\gamma,\gamma} |\phi_i\rangle \langle \phi_j| = \frac{d_{N-1}(E - E_i)}{d(E)} \sum_i |\phi_i\rangle \langle \phi_i| , \tag{D8}
$$

where E is the total system eigen-energy and E_i is the eigen-energy of the single mode corresponding to $|\phi_i\rangle$. The above state Eq. D8 is indeed thermal. To show that, note $d(E)$ and $d(E - E_i)$ are basically the number of quantum states for the total system and the $N-1$ modes, respectively. We have

$$
d(E) = e^{S(E)/k_B}, \quad d_{N-1}(E - E_i) = e^{S(E-E_i)/k_B}. \tag{D9}
$$

where the symbol S denotes the entropy. Plug them into Eq. D8, we get

$$
\rho'_1 = \sum_i e^{\frac{S(E-E_i)-S(E)}{k_B}} |\phi_i\rangle \langle \phi_i| . \tag{D10}
$$

If we do the expansion

$$
S(E - E_i) = S(E) - \left(\frac{\partial S}{\partial E}\right)_{E_i=0} E_i + ..., \tag{D11}
$$

we obtain

$$
\begin{aligned}
\rho'_1 &\simeq \sum_i e^{-\frac{1}{k_B}\left(\frac{\partial S}{\partial E}\right)_{E_i=0} E_i} |\phi_i\rangle \langle \phi_i| \\
&= \sum_i e^{-\beta E_i} |\phi_i\rangle \langle \phi_i| ,
\end{aligned}
\tag{D12}
$$

which is indeed a thermal state. Following the main text, we know each mode has energy (photon number) $K \sinh^2(r)/N$. If this thermal state purified into a two-mode squeezed state, it will have a squeezing factor r_0 satisfying $\sinh^2(r_0) = K \sinh^2(r)/N$. The entanglement is obviously $\mathcal{E}(r_0)$, as used in the main text.