Quantum Enigma Machines and the Locking Capacity of a Quantum Channel

Saikat Guha,[1] Patrick Hayden,[2] Hari Krovi,[1] Seth Lloyd,[3] Cosmo Lupo,[4] Jeffrey H. Shapiro,[4] Masahiro Takeoka,[5,1] and Mark M. Wilde[6]

[1]*Quantum Information Processing Group, Raytheon BBN Technologies, Cambridge, Massachusetts 02138, USA*

[2]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, California 94305-4060, USA*

[3]*Department of Mechanical Engineering, MIT, Cambridge, Massachusetts 02139, USA*

[4]*Research Laboratory of Electronics, MIT, Cambridge, Massachusetts 02139, USA*

[5]*National Institute of Information and Communications Technology, 4-2-1 Nukuikita, Koganei, Tokyo 184-8795, Japan*

[6]*Department of Physics and Astronomy, Center for Computation and Technology, Hearne Institute for Theoretical Physics, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

(Received 30 July 2013; revised manuscript received 21 November 2013; published 31 January 2014)

The locking effect is a phenomenon that is unique to quantum information theory and represents one of the strongest separations between the classical and quantum theories of information. The Fawzi-Hayden-Sen locking protocol harnesses this effect in a cryptographic context, whereby one party can encode n bits into n qubits while using only a constant-size secret key. The encoded message is then secure against any measurement that an eavesdropper could perform in an attempt to recover the message, but the protocol does not necessarily meet the composability requirements needed in quantum key distribution applications. In any case, the locking effect represents an extreme violation of Shannon's classical theorem, which states that information-theoretic security holds in the classical case if and only if the secret key is the same size as the message. Given this intriguing phenomenon, it is of practical interest to study the effect in the presence of noise, which can occur in the systems of both the legitimate receiver and the eavesdropper. This paper formally defines the *locking capacity* of a quantum channel as the maximum amount of locked information that can be reliably transmitted to a legitimate receiver by exploiting many independent uses of a quantum channel and an amount of secret key sublinear in the number of channel uses. We provide general operational bounds on the locking capacity in terms of other well-known capacities from quantum Shannon theory. We also study the important case of bosonic channels, finding limitations on these channels' locking capacity when coherent-state encodings are employed and particular locking protocols for these channels that might be physically implementable.

DOI: 10.1103/PhysRevX.4.011016 Subject Areas: Optics, Quantum Physics, Quantum Information

I. INTRODUCTION

The security of a cryptographic primitive can be assessed according to different security criteria. Most modern cryptosystems are *computationally* secure—that is, their security relies on the difficulty of breaking them in a reasonable amount of time given available technologies. This is also the case for the *enigma machines*, a family of historical polyalphabetic ciphers in use during the earlier half of the previous century—their security relied on the difficulty of uncovering patterns hidden in pseudorandom sequences [1].

A stronger security criterion requires that an encrypted message is close to being statistically independent of the corresponding unencrypted message, in which case one speaks of *information-theoretic* security. For the case of classical systems, a good measure of correlation is the mutual information between the unencrypted and the encrypted message. If the mutual information vanishes, the chance of successfully decrypting the message is exponentially small in the length of the message. Any encryption scheme with such a property cannot perform any better than one-time pad encryption, where a truly random key is used to encrypt (and decrypt) the message [2]. The one-time pad guarantees information-theoretic security as long as the key is kept secret, it has the same length as the message, and it can be used only once. However, the fact that the secret key should be the same length as the message imposes severe practical limitations on the use of the one-time pad protocol.

On the other hand, it is now known that quantum mechanics provides a way around these limitations. The *locking effect* is a phenomenon that is unique to quantum-information theory [3] and represents one of the most striking separations between the classical and

quantum theories of information. It is responsible for important revisions to security definitions for quantum key distribution (QKD) [4] and might even help to explain both how unitarity could be preserved and most of the information leaking from an evaporating black hole could be inaccessible until the final stages of evaporation [5,6]. Quantum data locking occurs when the accessible information about a classical message encoded into a quantum state decreases by an amount that is much larger than the number of qubits of a small subsystem that is discarded [3]. A device that realizes a quantum data-locking protocol is called a quantum enigma machine (QEM) [7].

Impressive locking schemes exist [6,8,9]. Suppose that a sender and a receiver share a constant number of secret key bits. Using these secret key bits, they can then encode an n-bit classical message into n qubits such that an adversary who gains access to these n qubits, but who does not know the secret key, cannot do much better than to randomly guess the message after performing an arbitrary measurement on these n qubits.

However, the cryptographic applications of quantum data locking have to be "taken with a grain of salt," as they are only applicable if the distribution of the message is completely random from the perspective of the adversary. Otherwise, the key size should increase by an amount necessary to ensure that the distribution of the message becomes uniform. Moreover, one might say that the strength of quantum data locking also exposes a weakness. Indeed, as a small key is sufficient for encrypting a long message, the leakage of a small part of the secret key may allow an adversary to uncover a disproportionate amount of information. For this reason, any cryptographic primitive based on the locking effect (called a locking scheme) does not necessarily guarantee composable security [4]. This also implies that quantum data locking cannot necessarily be used for secure key distribution. The only exception is if the adversary has no option other than to perform a collective measurement on the qubits in her possession just after she receives them.

As stated above, Shannon proved that such a locking effect is impossible classically [2]. That is, when using only classical resources, a sender and a receiver require a secret key whose size is proportional to the size of the message in order for the eavesdropper to have a negligible amount of information about the encrypted message. Thus, after Shannon's result, information scientists looked in a different direction in order to determine ways for communication systems to provide secrecy in addition to reliable transmission. In reality, all communication systems suffer from physical-layer noise, and one might be able to determine the characteristics of the noise to a legitimate receiver and to an untrusted eavesdropper. Such a model is known as the wiretap channel [10], and it is well known now that if the noise to the eavesdropper is stronger than the noise to the legitimate receiver, then it is possible to communicate error-free at a positive rate such that the eavesdropper obtains a negligible amount of information about the messages being transmitted.

II. SUMMARY OF RESULTS

In this paper, we consider the performance of locking protocols in the presence of noise, and as an important application, we consider locking protocols for bosonic channels. There are two types of noise to consider in any realistic locking protocol: that which affects the transmission to a legitimate receiver and that which affects the eavesdropper's system. Both are important to consider in any realization of a quantum enigma machine.

We begin in Sec. IV by reviewing the locking effect and a recently introduced quantum enigma machine from Ref. [7]. This QEM encodes a classical message into a single-photon state spread over a collection of discrete modes and then decodes it by direct photodetection. The encryption and decryption are realized by applying and inverting, respectively, a single multimode passive linear-optical unitary transformation, selected uniformly at random from a set of such transformations. Similar to historical enigma machines, QEMs can encrypt a long message using an exponentially shorter secret key. However, unlike historical enigma machines that were only computationally secure, quantum data locking implies security in the sense that the outcomes of any eavesdropper measurement will be essentially independent of the message.

After the review, Sec. V provides a formal definition of the locking capacity of a quantum channel. In short, a locking protocol uses a quantum channel n times (for some arbitrarily large integer n) and has three requirements.

(1) The receiver should be able to decode the transmitted message with an arbitrarily small error probability.
(2) The eavesdropper can recover only an arbitrarily small number of the message bits after performing a quantum measurement on their systems.
(3) The number of secret key bits used is no more than sublinear in the number n of channel uses (for example, logarithmic in n).

We define the locking capacity of a quantum channel to be the maximum rate at which it is possible to lock classical information according to the above requirements. Changing the systems to which the adversary has access leads to different notions of locking capacity, and we distinguish the notions by naming them the *weak locking capacity* and the *strong locking capacity*. The difference between the two is that, in the weak notion, the adversary is assumed to have access to only the channel environment, while, in the strong case, we allow access to the channel input. We emphasize that when we use the term (weak or strong) "locking capacity" without any other modifiers, we refer to the locking capacity of a quantum channel without additional resources, such as classical feedback. Most of

the results reported here correspond to such a forward-locking capacity. The locking capacity of channels with additional resources such as classical feedback remains largely open. However, at the very least, we can already say that quantum key distribution protocols provide lower bounds on the locking capacity in this setting.

We then find operational bounds on the locking capacity in terms of other well-known capacities studied in quantum Shannon theory, and we find other information-theoretic upper bounds on the locking capacity. We prove that the locking capacity of an entanglement-breaking channel is equal to zero, which demonstrates that a quantum channel should have some ability to preserve entanglement in order for it to be able to lock information according to the above requirements. We also show that any achievable locking rate is equal to zero whenever a given locking protocol has a classical simulation. Furthermore, we find a class of channels for which the weak locking capacity is equal to both the private capacity and the quantum capacity. Finally, we discuss locking protocols for some simple exemplary channels.

Section VI establishes several important upper bounds on the locking capacity of channels when restricting to coherent-state encodings. If it were possible to exploit coherent-state encodings to perform locking at high rates, this would certainly turn the locking effect from an interesting theoretical phenomenon into one with practical utility. However, we are able to show that there are fundamental limitations on the locking capacity when restricting to coherent-state encodings. In particular, we prove that the "strong" locking capacity of any channel is no larger than $\log_2(e)$ locked bits per channel use whenever the encoding consists of coherent states (where e is the base for the natural logarithm). We also prove that the "weak" locking capacity of a pure-loss bosonic channel is no larger than the sum of its private capacity and $\log_2(e)$.

In Sec. VII, we discuss an explicit protocol that uses a pulse-position modulation (PPM) encoding of coherent states. We derive bounds on the security and key efficiency of this coherent-state locking protocol and find that it has qualitative features analogous to the single-mode quantum enigma machine in the presence of linear loss.

Finally, Sec. VIII presents our conclusions, a discussion of the scaling of the required physical resources, and open questions for future research.

III. NOTATION

We briefly review some notation that we use in the rest of the paper. Let $\mathcal{B}(\mathcal{H})$ denote the algebra of bounded linear operators acting on a Hilbert space $\mathcal{H}$. The 1-norm of an operator X is defined as

$$\|X\|_1 \equiv \mathrm{Tr}\{\sqrt{X^\dagger X}\}.$$

Let $\mathcal{B}(\mathcal{H})_+$ denote the subset of positive semidefinite operators (we often simply say that an operator is "positive" if it is positive semidefinite). We also write $X \geq 0$ if $X \in \mathcal{B}(\mathcal{H})_+$. An operator ρ is in the set $\mathcal{D}(\mathcal{H})$ of density operators if $\rho \in \mathcal{B}(\mathcal{H})_+$ and $\mathrm{Tr}\{\rho\} = 1$. The tensor product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$. Given a multipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we unambiguously write $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$ for the reduced density operator on system A.

A linear map $\mathcal{N}_{A\to B}$: $\mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is positive if $\mathcal{N}_{A\to B}(\sigma_A) \in \mathcal{B}(\mathcal{H}_B)_+$, whenever $\sigma_A \in \mathcal{B}(\mathcal{H}_A)_+$. Let id_A denote the identity map acting on $\mathcal{B}(\mathcal{H}_A)$. A linear map $\mathcal{N}_{A\to B}$ is completely positive if the map $\mathrm{id}_R \otimes \mathcal{N}_{A\to B}$ is positive for a reference system R of arbitrary size. A linear map $\mathcal{N}_{A\to B}$ is trace preserving if $\mathrm{Tr}\{\mathcal{N}_{A\to B}(\tau_A)\} = \mathrm{Tr}\{\tau_A\}$ for all input operators $\tau_A \in \mathcal{B}(\mathcal{H}_A)$. If a linear map is completely positive and trace preserving, we say that it is a quantum channel or quantum operation. For simplicity, we denote a quantum channel $\mathcal{N}$: $\mathcal{B}(\mathcal{H}_A) \mapsto \mathcal{B}(\mathcal{H}_B)$ simply as $\mathcal{N}_{A\to B}$. Similarly, we denote an isometry U: $\mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_C$ simply as $U_{A\to BC}$.

The variational distance between two probability distributions $p(x)$ and $q(x)$ is defined as

$$\sum_x |p(x) - q(x)|.$$

The trace distance between two quantum states ρ and σ is defined as follows:

$$\|\rho - \sigma\|_1,$$

and it is a conventional measure used in quantum-information theory to quantify the distinguishability of two quantum states. Clearly, when the two states are commuting, the trace distance is equal to the variational distance between the two probability distributions corresponding to the eigenvalues of ρ and σ.

The von Neumann entropy of a state $\rho \in \mathcal{D}(\mathcal{H}_A)$ is given by $H(A)_\rho := -\mathrm{Tr}\{\rho \log \rho\}$. Throughout this paper, we take the logarithm base 2. For a tripartite state $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_{ABC})$, the quantum mutual information and the conditional quantum mutual information are, respectively, given by

$$I(A;B)_\rho \equiv H(A)_\rho + H(B)_\rho - H(AB)_\rho,$$
$$I(A;B|C)_\rho \equiv I(A;BC)_\rho - I(A;C)_\rho,$$

where $H(A)_\rho$ denotes the von Neumann entropy of the reduced state ρ_A, for example.

IV. REVIEW OF QUANTUM DATA LOCKING

A quantum data-locking scheme can be implemented by a set of $|\mathcal{K}|$ unitary transformations $\{U_k\}_{k\in\mathcal{K}}$ acting on a Hilbert space $\mathcal{H}_M$ of finite dimension $|\mathcal{M}|$ [3,6,8,9]. (For the moment, we restrict ourselves to finite-dimensional

Hilbert spaces, but Definitions 1 and 2, below, allow for encoding information into infinite-dimensional Hilbert spaces.) Alice encodes $|\mathcal{M}|$ equiprobable messages by means of a set of orthonormal states $\{|m\rangle\}_{m\in\mathcal{M}}$ defining a standard basis in $\mathcal{H}_M$. The encryption is then made by applying a particular unitary U_k with k chosen uniformly at random from $\mathcal{K}$, and this unitary maps a standard basis state $|m\rangle$ into a state $U_k|m\rangle$. The label k identifies the choice of the basis and plays the role of a secret key.

It is helpful to consider a particular classical-quantum state when reasoning about a quantum data-locking protocol. For such a state, we have two classical systems, the first associated with Alice's message and the second associated with the secret key, and a quantum system Q of dimension $|\mathcal{M}|$ corresponding to the quantum-encoded message of Alice. This classical-quantum state is given by the following density matrix:

$$\rho_{MKQ} = \frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}|m,k\rangle\langle m,k|_{MK}\otimes(U_k|m\rangle\langle m|U_k^\dagger)_Q, \tag{1}$$

where the sets $\{|m\rangle\}$ and $\{|k\rangle\}$ are composed of orthonormal states representing the message and the secret key, respectively. The receiver, Bob, has access to the quantum system Q and the key system K. We assume that an eavesdropper, Eve, has access only to the quantum system Q (for example, before it gets passed along to the receiver Bob). The classical correlations between Alice's message M and Bob's systems K and Q can be quantified by the *accessible information* [11]. This is defined as the maximum classical mutual information that can be extracted by performing local measurements on the bipartite state:

$$I_{\mathrm{acc}}(M;KQ)_\rho = \max_{\mathcal{M}_{KQ\to Y}} I(M;Y), \tag{2}$$

where the maximization is taken over local measurement maps $\mathcal{M}_{KQ\to Y}$, and $I(X;Y) = H(X) + H(Y) - H(XY)$ is the mutual information, with $H(Z)$ denoting the Shannon entropy of the random variable Z [12].

The accessible information in Eq. (2) can never be larger than $\log_2|\mathcal{M}|$, due to the bound $I(M;Y)\le\log_2|\mathcal{M}|$, which holds for any random variable Y. A particular strategy for achieving this upper bound is for Bob to first perform the controlled unitary $\sum_k|k\rangle\langle k|_K\otimes(U_k^\dagger)_Q$, leaving the state

$$\frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}|m,k\rangle\langle m,k|_{MK}\otimes|m\rangle\langle m|_Q.$$

He then simply measures in the basis $\{|m\rangle\}$ to recover the message m perfectly, so that his accessible information is maximal, equal to $\log_2|\mathcal{M}|$.

To assess the security of the communication, let us consider the accessible information for a party, Eve, who does not have access to the secret key. We consider the following reduced state:

$$\rho_{MQ} = \frac{1}{|\mathcal{M}|}\sum_m|m\rangle\langle m|_M\otimes\frac{1}{|\mathcal{K}|}\sum_k(U_k|m\rangle\langle m|U_k^\dagger)_Q, \tag{3}$$

obtained by taking the partial trace over the key system in Eq. (1). The aim of Eve is to find an optimal positive operator-valued measure (POVM) to maximize the classical mutual information. It is sufficient to consider a POVM $\mathcal{M}_{Q\to Y}$ with rank-one measurement operators, i.e.,

$$\{\mu_y|\varphi_y\rangle\langle\varphi_y|\}, \tag{4}$$

where each $|\varphi_y\rangle$ is a normalized vector and $\mu_y>0$ (the sufficiency of rank-one POVMs follows by a data-processing argument). We then find the following expression for Eve's accessible information about Alice's message [3]:

$$I_{\mathrm{acc}}(M;Q)_\rho \le \log_2|\mathcal{M}| - \min_{\mathcal{M}_{E\to Y}}\sum_y\frac{\mu_y}{|\mathcal{M}||\mathcal{K}|}\sum_k H(q_{yk}), \tag{5}$$

where the probability distributions q_{yk} have components $q_{yk}^m = |\langle\varphi_y|U_k|m\rangle|^2$. Note that Eve's accessible information is written in terms of the minimum of the Shannon entropies $H(q_{yk}) = -\sum_m q_{yk}^m\log_2 q_{yk}^m$ averaged over y and k.

While finding Eve's optimal POVM is generally a difficult problem, one can obtain a good upper bound by a convexity argument [3]. Furthermore, one can choose the encoding unitaries uniformly at random according to the Haar measure [6,8,9,13], and if one also adjoins to the message a small ancilla system in a maximally mixed state [9], then it is possible to reduce the adversary's accessible information to become arbitrarily small. These latter results show that, for large enough $|\mathcal{M}|$, there exist data-locking schemes with $\log_2|\mathcal{K}|$ negligibly small in comparison to $\log_2|\mathcal{M}|$, and for which

$$I_{\mathrm{acc}}(M;Q)_\rho \ll I_{\mathrm{acc}}(M;KQ)_\rho.$$

That means that a relatively short secret key can be used to encrypt an exponentially longer message. To be more precise, consider the results of [9], according to which, for $|\mathcal{M}|$ large enough, there exist choices of $|\mathcal{K}|$ unitaries, with

$$\log_2|\mathcal{K}| = 4\log_2(\epsilon^{-1}) + O[\log_2\log_2(\epsilon^{-1})], \tag{6}$$

such that

$$I_{\mathrm{acc}}(M;Q)_\rho \le \epsilon\log_2|\mathcal{M}|, \tag{7}$$

for any $\epsilon > 0$. Moreover, if one randomly chooses the $|\mathcal{K}|$ unitaries according to the Haar distribution on the unitary group, the probability of picking up a set with this property approaches one exponentially fast in the limit as $|\mathcal{M}| \to \infty$.

In quantum data locking, the removal of a subsystem reduces the accessible information by an amount larger than the number of qubits removed. This is a purely quantum feature that has no classical analog. For comparison, consider a classical counterpart of the quantum data-locking setting, in which Alice has access to a message variable M, Bob has access to an output random variable Y and key variable K, while Eve has access to Y only. In the classical framework, the following inequality holds:

$$I(M;YK) - I(M;Y) = I(M;K|Y) \leq H(K) \leq \log_2|\mathcal{K}|. \tag{8}$$

This inequality shows that, in the classical framework, removal of the key variable K reduces the mutual information by no more than $\log_2|\mathcal{K}|$.

In the quantum case discussed above, this inequality can be violated by an arbitrarily large amount by replacing the classical mutual information with the accessible information. A violation of the classical inequality in Eq. (8) can be quantified in terms of the following ratios [3,8]:

$$r_1 = \frac{I_{\text{acc}}(M;Q)_\rho}{I_{\text{acc}}(M;KQ)_\rho}, \tag{9}$$

$$r_2 = \frac{\log_2|\mathcal{K}|}{I_{\text{acc}}(M;KQ)_\rho - I_{\text{acc}}(M;Q)_\rho}. \tag{10}$$

Equation (9) is the ratio of the accessible information without the secret key to that with the secret key. Equation (10) is the ratio of the key length to the amount of information that Bob can unlock by having access to the key. For a good locking scheme, both of these quantities should be small, and the quantum data-locking schemes discussed above are such that both r_1 and r_2 can be made arbitrarily small. On the other hand, the inequality in Eq. (8) implies that $r_2 \geq 1$ for any locking scheme that uses classical resources only. Note that the one-time pad protocol has $r_2 = 1$ because the number of bits in the key is equal to the amount of unlocked information for Bob.

A. Quantum enigma machine

A particular example of a QEM was proposed in Ref. [7]. This QEM implements an optical realization of quantum data locking, in which Alice exploits a pulse-position modulation encoding using single-photon states over n optical modes [7]. The message states $|m\rangle = a_m^\dagger|0\rangle$ represent the states of a single photon occupying one out of a set of n bosonic modes with canonical operators $\{a_m, a_m^\dagger\}_{m\in\{1,\ldots,n\}}$. Thus, for this case, we have $n = |\mathcal{M}|$. The unitaries $\{U_k\}_{k\in\mathcal{K}}$ are realized as passive linear-optical transformations acting on n modes. The encryption through a passive linear-optical unitary U_k transforms the message states into

$$|m\rangle_k := U_k|m\rangle = \sum_{m'=1}^{n} \tilde{U}_k^{(m,m')}|m'\rangle, \tag{11}$$

where $\tilde{U}_k$ is the corresponding $n \times n$ unitary matrix acting on the mode labels. The effect of the encryption is to spread a single photon coherently over n modes.

Let us first assume that Alice and Bob communicate via a noiseless quantum channel. Then, Bob receives the state prepared by Alice unperturbed. He decrypts the message by first applying the inverse transformation $U_k^\dagger$ and then by performing photodetection on the modes $\{a_m\}$. We assume that Eve may intercept the signal, but she does not know which unitary has been used for encryption. Then, a direct application of the results of Ref. [9] shows that Eve's accessible information can be made arbitrarily small using a preshared secret key of length logarithmic in the length of the message.

One natural application of a QEM is in synergy with standard quantum key distribution [14,15]—that is, a relatively short secret key can be first established by QKD and then used to encrypt a much (exponentially) longer message through the QEM. This combination of QKD and QEM in an all-quantum-optical cryptosystem could possibly overcome the bit-rate limitations of standard QKD, but more work is necessary to determine if this is the case.

Let us now suppose that Alice and Bob communicate through a pure-loss bosonic channel with transmissivity $\eta \in (0,1)$ and Eve makes a passive wiretap attack on the communication line, hence getting the photon lost in the channel with probability no larger than $1-\eta$. A simple feedback-assisted strategy allows for Alice and Bob to use the same scheme even for transmissivity values below 50%. Note that the only effect of the pure-loss channel is to induce a probabilistic leakage of the photon. Hence, each time Bob detects a photon (which happens with probability η), he can be sure that he has correctly decrypted Alice's message. On the other hand, if Bob's photodetectors do not produce a click (which happens with probability $1-\eta$), he can request for Alice to resend. This shows that, with the help of a classical feedback channel, Alice and Bob can attain the accessible information

$$I_{\text{acc}}(M;KQ)_\rho = \eta\log_2 n. \tag{12}$$

Although reduced, this value of the accessible information equals the maximum value achievable through a pure-loss bosonic channel with a mean value of n^{-1} photons per mode [7,16].

Lloyd argues that such a scheme should be secure, in principle [7]. However, a critical assumption for this security to hold is that Eve should attack each block that she receives independently, in which case, her accessible information is reduced by a factor $1-\eta$ when compared to the lossless case. Indeed, an important assumption for the security of any locking protocol is that the distribution of the message is uniform from the perspective of the adversary. If this is not the case (as for repeated transmission of the same message when it does not show up at the receiver's end), then the secret key needs to be large enough so that the distribution of the message becomes uniform (see Proposition 4.16 in Ref. [17]).

Concerning the key efficiency of the protocol, we can estimate the key efficiency ratio as

$$r_2 \simeq \frac{4\log_2(\epsilon^{-1})}{\eta \log_2 n}. \tag{13}$$

This expression implies that, although r_2 can be made arbitrarily small by increasing n, the number of bosonic modes needed to fulfill the key efficiency condition $r_2 < 1$ grows exponentially with decreasing r_2 and η. This feature is, first of all, a consequence of the fact that the quantum data-locking scheme in Ref. [9] (similar conclusions are also obtained using the results of [6,8]) requires a high-dimensional Hilbert space. On top of that, there is the fact that the PPM encoding, as noted above, is highly inefficient as it encodes $\log_2 n$ qubits into n optical modes.

According to Definition 1, below, this QEM is an instance of an (n, R, ϵ) weak locking protocol (assisted by classical feedback) for the pure-loss bosonic channel with transmissivity η, with a locking rate $R = \lfloor \eta \log_2 n \rfloor / n$. It is worthwhile to note that, due to the inefficiency of PPM encoding, the rate of this QEM approaches zero as n increases.

V. LOCKING CAPACITY OF A QUANTUM CHANNEL

In this section, we take a more general approach to quantum data locking than that pursued in prior work by defining the locking capacity of a quantum channel. Our goal is to understand the locking effect in the setting of quantum Shannon theory, where a sender and a receiver are given access to n independent uses of a noisy quantum channel (where n is an arbitrarily large integer). Their aim is to exploit some sublinear (in n) amount of secret key in order to lock classical messages from an adversary, in the sense that this adversary will not be able to do much better than random guessing when performing a quantum measurement to learn about the transmitted message. Also, we demand that the legitimate receiver (who knows the value of the secret key) be able to recover the classical message with an arbitrarily small probability of error. This leads us naturally to the following formal definition of a locking protocol for a noisy channel.

Definition 1: Weak locking protocol.—An (n, R, ϵ) weak locking protocol for a channel $\mathcal{N}_{A\to B}$ consists of encoding and decoding maps $\mathcal{E}_{MK\to A^n}$ and $\mathcal{D}_{B^nK\to\hat{M}}$, respectively. The encoding $\mathcal{E}_{MK\to A^n}$ acts on a message system M and a key system K and outputs the system A^n for input to n uses of the channel. The decoding map $\mathcal{D}_{B^nK\to\hat{M}}$ acts on the output systems B^n and the key system K to produce a classical system $\hat{M}$ containing the receiver's estimate of the message. Without loss of generality, the encoding consists of $|\mathcal{M}||\mathcal{K}|$ quantum states $\rho_{m,k}$, where $|\mathcal{M}|$ is the number of messages and $|\mathcal{K}|$ is the number of key values. Furthermore, the decoding consists of $|\mathcal{K}|$ POVMs $\{\Lambda_m^{(k)}\}_{m\in\mathcal{M}}$. The rate $R = \log_2|\mathcal{M}|/n$ and the parameter $\epsilon > 0$. The protocol should satisfy the following requirements.

(1) Given the key, the receiver can decode the transmitted message well on average:

$$\frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}\mathrm{Tr}\{\Lambda_m^{(k)}(\mathcal{N}_{A\to B})^{\otimes n}(\rho_{m,k})\} \geq 1-\epsilon.$$

(2) Let $\{\Gamma_y\}$ be a POVM that Eve can perform in an attempt to learn about the message M. After she performs this measurement, the joint classical-classical state of the message and her measurement outcome is as follows:

$$\frac{1}{|\mathcal{M}|}\sum_m |m\rangle\langle m|_M \otimes \sum_y \mathrm{Tr}\left\{\Gamma_y\left(\frac{1}{|\mathcal{K}|}\sum_k (\mathcal{N}_{A\to E})^{\otimes n}(\rho_{m,k})\right)\right\}|y\rangle\langle y|_Y,$$

where $\mathcal{N}_{A\to E}$ is the channel complementary to $\mathcal{N}_{A\to B}$. Equivalently, the joint probability distribution $p_{M,Y}(m,y)$ is equal to

$$p_{M,Y}(m,y) = \frac{1}{|\mathcal{M}|}\mathrm{Tr}\left\{\Gamma_y\left(\frac{1}{|\mathcal{K}|}\sum_k (\mathcal{N}_{A\to E})^{\otimes n}(\rho_{m,k})\right)\right\}.$$

Our security criterion (see also Ref. [9]) is that, for any measurement outcome y of Eve, the variational distance between the message distribution $p_M(m)$ and the distribution $p_{M|Y}(m|y)$ for the message conditioned on any particular measurement outcome should be no larger than ϵ:

$$\sum_m |p_M(m) - p_{M|Y}(m|y)| \leq \epsilon. \tag{14}$$

The interpretation here is that Eve cannot do much better than to randomly guess the message if all of the

conditional distributions $p_{M|Y}(m|y)$ are indistinguishable from the message distribution.

(3) The secret key consumption grows sublinearly in the number n of channel uses.

In a weak locking protocol, it is assumed that the eavesdropper has access to the channel environment only. A stronger locking protocol is obtained if we allow for the eavesdropper to have access to the channel input (or, equivalently, to both the channel output and the environment).

Definition 2: Strong locking protocol.—An (n, R, ϵ) strong locking protocol is similar to a weak locking protocol, except that we allow for Eve to have access to the A^n systems, so that she can perform a measurement on the A^n systems of the following state:

$$\frac{1}{|\mathcal{M}|}\sum_m |m\rangle\langle m|_M \otimes \frac{1}{|\mathcal{K}|}\sum_k (\rho_{m,k})_{A^n}.$$

We then demand that the variational distance as in Eq. (14) can be made less than an arbitrarily small positive constant ϵ.

Remark 3.—One could, alternatively, allow for the adversary to have access to the output of the channel, but we do not explore such a possibility in this paper.

Remark 4.—The Fannes-Audenaert inequality [18,19] for continuity of entropy implies that, if Eq. (14) holds, then we get the following bound on Eve's accessible information:

$$I_{\text{acc}}(M; E^n) \leq h_2(\epsilon/2) + \epsilon n R/2, \tag{15}$$

where h_2 is the binary entropy and n and R are as in Definition 1. In more detail, recall the Fannes-Audenaert inequality for continuity of entropy,

$$T \equiv \frac{1}{2}\|\rho - \sigma\|_1 \Rightarrow |H(\rho) - H(\sigma)| \leq h_2(T) + T\log(d-1),$$

where h_2 is the binary entropy and d is the dimension of the states. Applying this to the condition in Eq. (14) gives

$$H(M) - H(M|Y=y) \leq h_2(\epsilon/2) + \frac{\epsilon}{2}\log(|\mathcal{M}|-1), \tag{16}$$

$$\leq h_2(\epsilon/2) + \epsilon n R/2. \tag{17}$$

Since the above inequality holds for any measurement of Eve, averaging it with respect to the distribution $p_Y(y)$ gives the inequality in Eq. (15).

Remark 5.—If desired, one can demand further for the secret key rate of an (n, R, ϵ) weak or strong locking protocol to be consumed at a particular sublinear rate (for example, a logarithmic number of secret key bits or perhaps $\sqrt{n}$ secret key bits for n channel uses). However, this paper establishes several upper bounds on locking capacity in an independent and identically distributed (IID) setting, and these bounds converge to the same quantity in the large n limit regardless of which sublinear rate is chosen. Also, the Fawzi-Hayden-Sen (FHS) protocol [9] is very strong, in the sense that it uses such a small amount of secret key. Thus, in light of these two observations, it seems reasonable to define locking capacity in such a coarse-grained manner. However, other characterizations of locking capacity in a finite block-length setting or in a one-shot setting might change depending on the amount of secret key allowed (so it would be necessary to specify in more detail the amount of secret key allowed).

Remark 6.—Observe that an (n, R, ϵ) strong locking protocol is also an (n, R, ϵ) weak locking protocol, but the other implication is not necessarily true.

Remark 7.—The security and key efficiency ratios become arbitrarily small for a strong locking protocol. Indeed, from the fact that $I_{\text{acc}}(M; A^n) \leq h_2(\epsilon/2) + \epsilon n R/2$, and the fact that the receiver can decode with the key, so that $I_{\text{acc}}(M; B^n K) \approx \log_2|\mathcal{M}|$, it follows that the security ratio $r_1 \leq h_2(\epsilon/2)/(nR) + \epsilon/2$. Also, since we require the key to be sublinear in the message length, it follows that the key efficiency ratio $r_2 = o(n)/O(n)$, which vanishes in the limit as $n \to \infty$.

In the following, we use the modifier "weak" or "strong" only when we need to distinguish between them.

Definition 8: Achievable rate for locking.—A rate R is achievable if $\forall\ \delta,\ \epsilon > 0$ and sufficiently large n, there exists an $(n, R-\delta, \epsilon)$ locking protocol.

Definition 9: Locking capacity.—The locking capacity $L(\mathcal{N})$ of a quantum channel is the supremum of all achievable rates:

$$L(\mathcal{N}) \equiv \sup\{R | R \text{ is achievable}\}.$$

Let $L_W(\mathcal{N})$ and $L_S(\mathcal{N})$ denote the weak and strong locking capacity, respectively.

A. Relation of the locking capacity to other capacities

Let $Q(\mathcal{N})$, $P(\mathcal{N})$, and $C(\mathcal{N})$ denote the quantum [20–26], private [26,27], and classical [28,29] capacities of a quantum channel $\mathcal{N}$, respectively. By employing operational arguments, we determine that the following bounds hold:

$$Q(\mathcal{N}) \leq P(\mathcal{N}) \leq L_W(\mathcal{N}) \leq C(\mathcal{N}). \tag{18}$$

Indeed, for any channel, its quantum capacity is less than the private classical capacity because any scheme for quantum communication can be used for private classical communication such that the classical information is protected from the environment of the channel. Furthermore, the inequality $P(\mathcal{N}) \leq L_W(\mathcal{N})$ holds because

any (n, R, ϵ) private classical communication protocol satisfies the three requirements of a weak locking protocol 26,27]]. Finally, the requirements of a weak locking protocol are more restrictive than those for classical communication, so that $L_W(\mathcal{N}) \leq C(\mathcal{N})$.

Operational arguments and the existence of the FHS locking protocol [9] also lead to the following bounds on the strong locking capacity:

$$Q(\mathcal{N}) \leq L_S(\mathcal{N}) \leq L_W(\mathcal{N}). \tag{19}$$

We first justify the bound $Q(\mathcal{N}) \leq L_S(\mathcal{N})$, already observed in some sense in Ref. [17]. The strong locking capacity of the noiseless qubit channel is equal to one, due to the existence of the FHS locking protocol (see Example 23 below). By concatenating the FHS locking protocol with a family of capacity-achieving quantum error correcting codes, we obtain a family of strong locking protocols that achieve a strong locking rate equal to the quantum capacity of the channel. The bound $L_S(\mathcal{N}) \leq L_W(\mathcal{N})$ follows because a strong locking protocol always meets the demands of a weak locking protocol (recall Remark 6). The relationship between the private capacity and the strong locking capacity is less clear. Indeed, a private communication protocol for a quantum channel protects information only from the environment of the channel (which we think of as the eavesdropper's system). For this reason, it does not meet the demands of a strong locking protocol. However, we could consider a "strong privacy" protocol in which the goal is to protect a message from both the environment and the output of the channel, under the assumption that the party controlling these systems does not have access to the shared key. In this case, the "strong private capacity" would always be equal to zero because a sublinear amount of secret key is insufficient to get any strong private capacity out of the channel. For this reason, the bounds in Eqs. (18) and (19) are the best simple ones that we can derive from operational considerations.

We can also consider the case in which a classical feedback channel is available for free from the receiver to the sender. In this case, we denote the resulting capacities with a superscript $(\leftarrow)$. By employing the same operational arguments as above, we find that the following inequalities hold:

$$Q^{(\leftarrow)}(\mathcal{N}) \leq P^{(\leftarrow)}(\mathcal{N}) \leq L_W^{(\leftarrow)}(\mathcal{N}) \leq C^{(\leftarrow)}(\mathcal{N}), \tag{20}$$

$$Q^{(\leftarrow)}(\mathcal{N}) \leq L_S^{(\leftarrow)}(\mathcal{N}) \leq L_W^{(\leftarrow)}(\mathcal{N}). \tag{21}$$

Capacities assisted by classical feedback need not be equal to the unassisted capacities. For example, it is known that the quantum and private capacities assisted by classical feedback can be strictly larger than the corresponding unassisted capacities [30], and this is true even for the classical capacity [31]. The locking capacity of quantum channels with classical feedback remains largely an open question.

B. Upper bounds on the locking capacity

Let us define the information quantity $L_W^{(u)}(\mathcal{N})$ as follows:

$$L_W^{(u)}(\mathcal{N}) \equiv \max_{\{p(x),\rho_x\}} [I(X;B) - I_{\text{acc}}(X;E)], \tag{22}$$

where the above information quantities are evaluated with respect to a state of the following form:

$$\sum_x p_X(x)|x\rangle\langle x|_X \otimes U_{A\to BE}^{\mathcal{N}}(\rho_x), \tag{23}$$

$U_{A\to BE}^{\mathcal{N}}$ is an isometric extension of the channel $\mathcal{N}$, and the superscript (u) indicates that this quantity will function as an upper bound on the locking capacity. The following theorem establishes that the regularization of $L_W^{(u)}(\mathcal{N})$ provides an upper bound on the weak locking capacity of a quantum channel. This bound is nontrivial given that the regularization of $L_W^{(u)}(\mathcal{N})$ does not depend on the secret key used in a given locking protocol.

Theorem 10.—The weak locking capacity $L_W(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is upper bounded by the regularization of $L_W^{(u)}(\mathcal{N})$:

$$L_W(\mathcal{N}) \leq \lim_{n\to\infty} \frac{1}{n} L_W^{(u)}(\mathcal{N}^{\otimes n}).$$

Proof: The proof below places an upper bound on the weak locking capacity of a quantum channel by considering the most general protocol for this task. Suppose that the task is to generate shared, locked randomness rather than to send a locked message (placing an upper bound on achievable rates for this task gives an upper bound on achievable rates for the latter task, since a protocol for the latter task can be used to accomplish the former task). The most general protocol has Alice input her share of the key K and her variable M into an encoder that outputs some systems A^n to be fed into the inputs of the channels. She then transmits these systems A^n over the channel, so that Bob receives the output systems B^n. Let the following state describe all systems at this point in the protocol:

$$\omega_{MKB^n} \equiv \frac{1}{|\mathcal{M}||\mathcal{K}|} \sum_{m,k} |m\rangle\langle m|_M \otimes |k\rangle\langle k|_K \otimes \mathcal{N}_{A\to B}^{\otimes n}(\rho_{k,m}).$$

Bob inputs his share of the key K and the systems B^n into a decoder $\mathcal{D}_{KB^n\to\hat{M}}$ to recover $\hat{M}$, which is his estimate of Alice's variable M. The final state of the protocol is given by

$$\omega'_{M\hat{M}} \equiv \frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}|m\rangle\langle m|_M \otimes \mathcal{D}_{KB^n\to\hat{M}}[|k\rangle\langle k|_K \otimes \mathcal{N}^{\otimes n}_{A\to B}(\rho_{k,m})].$$

If the protocol is any good for locking the message M, then the ideal distribution of M and $\hat{M}$ deviates from the actual distribution of these variables by no more than ϵ, in the sense that

$$\|\bar{\Phi}_{M\hat{M}} - \omega'_{M\hat{M}}\|_1 \leq \epsilon,$$

where

$$\bar{\Phi}_{M\hat{M}} \equiv \frac{1}{|\mathcal{M}|}\sum_m |m\rangle\langle m|_M \otimes |m\rangle\langle m|_{\hat{M}}.$$

The above condition is equivalent to the condition that $\Pr\{\hat{M} \neq M\} \leq \epsilon/2$, because

$$\frac{1}{2}\|\bar{\Phi}_{M\hat{M}} - \omega'_{M\hat{M}}\|_1 = \Pr\{\hat{M} \neq M\}.$$

Also, from Remark 4, Eve's accessible information $I_{\text{acc}}(M;E^n)$ about the variable M is bounded from above by $\epsilon'' n$, where $\epsilon'' \equiv h_2(\epsilon/2)/n + \epsilon R/2$, whenever Eq. (14) is satisfied. We can now proceed with bounding achievable rates for any locking protocol:

$$\begin{aligned} nR &= H(M)_{\bar{\Phi}} \\ &= I(M;\hat{M})_{\bar{\Phi}} \\ &\leq I(M;\hat{M})_{\omega'} + n\epsilon' \\ &\leq I(M;B^nK)_{\omega} + n\epsilon' \\ &= I(M;B^n)_{\omega} + I(M;K|B^n)_{\omega} + n\epsilon' \\ &\leq I(M;B^n)_{\omega} - I_{\text{acc}}(M;E^n)_{\omega} + o(n) + n\epsilon' + n\epsilon'' \\ &\leq L^{(u)}_W(\mathcal{N}^{\otimes n}) + o(n) + n\epsilon' + n\epsilon''. \end{aligned}$$

The first equality follows from the assumption that the random variable M is a uniform random variable. The second equality is an identity because $H(M|\hat{M}) = 0$ for the ideal distribution on M and $\hat{M}$. The first inequality follows from an application of the Alicki-Fannes-Audenart inequality (continuity of entropy) [19,32], where ϵ' is a function of ϵ that approaches zero as $\epsilon \to 0$. The second inequality follows from an application of quantum data processing (both B^n and K are fed into the decoder to produce $\hat{M}$). The third equality follows from an application of the chain rule for mutual information. The third inequality follows from the upper bound

$$I(M;K|B^n) \leq H(K|B^n) \leq H(K) \leq o(n)$$

(the assumption that the secret key rate is sublinear) and from the accessible information bound $I_{\text{acc}}(M;E^n) \leq \epsilon$. The final inequality follows from optimizing over all distributions, so that we have

$$R \leq \lim_{n\to\infty}\frac{1}{n}L^{(u)}_W(\mathcal{N}^{\otimes n}).$$

in the limit as n becomes large and as $\epsilon \to 0$. □

Theorem 11.—The strong locking capacity $L_S(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is upper bounded as

$$L_S(\mathcal{N}) \leq \lim_{n\to\infty}\frac{1}{n}L^{(u)}_S(\mathcal{N}^{\otimes n}),$$

where

$$L^{(u)}_S(\mathcal{N}) \equiv \max_{\{p(x),\rho_x\}}[I(X;B) - I_{\text{acc}}(X;BE)],$$

and the information quantities are with respect to the state in (23).

Proof: The proof of this theorem is nearly identical to the proof of the one above. However, we employ the bound on the accessible information $I_{\text{acc}}(M;A^n) = I_{\text{acc}}(M;B^nE^n)$ from Definition 2 and Remark 4 instead. □

Remark 12.—Observe that the bounds in the above theorem hold even if the key is allowed to be a sublinear size quantum system, as in the locking schemes discussed in Ref. [6].

It is an interesting and important open question to determine if the upper bounds given in the above theorems are achievable.

1. Entanglement-breaking channels have zero locking capacity

The above theorems and a further analysis allow us to determine that both the strong and the weak locking capacities of an entanglement-breaking channel are equal to zero.

Definition 13: Entanglement-breaking channel Ref. [33].—A channel $\mathcal{N}_{\text{EB}}$ is entanglement breaking if the output state is separable whenever it acts on one share of an entangled state:

$$(\text{id}_R \otimes \mathcal{N}_{\text{EB}})(\rho_{RA}) = \sum_x p_X(x)\sigma^x_R \otimes \omega^x_B,$$

where $p_X(x)$ is a probability distribution, each σ^x_R is a state on the reference system R, and each ω^x_B is a state on the channel output system B.

Theorem 14.—Both the strong and the weak locking capacities of an entanglement-breaking channel $\mathcal{N}_{\text{EB}}$ are equal to zero:

$$L_W(\mathcal{N}_{\text{EB}}) = L_S(\mathcal{N}_{\text{EB}}) = 0.$$

Proof: The proof of this theorem exploits the upper bound derived in Theorem 10 and the fact that

$L_W(\mathcal{N}_{\mathrm{EB}}) \geq L_S(\mathcal{N}_{\mathrm{EB}})$. We know from Ref. [33] that any entanglement-breaking channel has a representation with rank-one Kraus operators, so that its action on an input density operator is given by

$$\mathcal{N}_{\mathrm{EB}}(\rho) = \sum_y |\varphi_y\rangle_B \langle\psi_y|_A \rho |\psi_y\rangle_A \langle\varphi_y|_B,$$

for some set of vectors $\{|\psi_y\rangle_A\}$, such that $\sum_y |\psi_y\rangle\langle\psi_y|_A = I_A$, and a set of states $\{|\varphi_y\rangle_B\}$. An isometric extension of the channel is then given by

$$U_{A\to BE}^{\mathcal{N}_{\mathrm{EB}}} \equiv \sum_y |\varphi_y\rangle_B \langle\psi_y|_A \otimes |y\rangle_E,$$

with $\{|y\rangle_E\}$ an orthonormal basis for the environment. From this representation, it is clear that the channel to the environment is of the form

$$\mathcal{N}_{\mathrm{EB}}^c(\rho) = \sum_{y,z} \langle\psi_y|\rho|\psi_z\rangle_A \langle\varphi_z|\varphi_y\rangle_B |y\rangle\langle z|_E,$$

and the environment can simulate the channel to the receiver by first performing a von Neumann measurement in the basis $\{|y\rangle\}$ followed by a preparation of the state $|\varphi_y\rangle_B$ conditioned on the measurement outcome being y.

Now, consider the information quantity $L_W^{(u)}(\mathcal{N}_{\mathrm{EB}})$ defined in Eq. (22). Theorem 10 states that the regularization of this quantity is an upper bound on the weak locking capacity. For any finite n, we can always pick the measurement to be a tensor-product von Neumann measurement of the form mentioned above, giving that

$$I_{\mathrm{acc}}(X;E^n) \geq I(X;Y^n),$$

where Y^n is the random variable corresponding to the measurement outcomes. Because of the structural relationship given above (the fact that the environment can simulate the channel to the receiver by preparing n quantum states $|\varphi_{y_1}\rangle \otimes \cdots \otimes |\varphi_{y_n}\rangle$ from the measurement outcomes y^n), we find that

$$I(X;Y^n) \geq I(X;B^n),$$

by an application of the quantum data-processing inequality. This is equivalent to $I(X;B^n) - I(X;Y^n) \leq 0$, which implies that $\lim_{n\to\infty} \frac{1}{n} L_W^{(u)}(\mathcal{N}_{\mathrm{EB}}^{\otimes n}) = 0$ and thus that the weak locking capacity vanishes for any entanglement-breaking channel. □

Remark 15.—The importance of the above theorem is the conclusion that a channel should be able to preserve entanglement between a purification of the channel input and its output in order for it to be able to lock information. If it is not able to (i.e., if it is entanglement breaking), then the locking capacity is equal to zero. Reference [34] suggested that entanglement does not play a role in quantum data locking, but this theorem shows that it does in any realistic implementation of a locking protocol.

Remark 16.—It should be possible to provide a rigorous generalization of this result to entanglement-breaking channels defined over general infinite-dimensional spaces using the techniques from Ref. [35]. For example, it is known that a lossy bosonic channel becomes entanglement breaking when the environment injects a thermal state with sufficiently high photon number [35]. However, we leave this question open for future work.

2. Protocols with classical simulations have zero strong locking rate

It is important to determine the conditions for when the locking rate of a given protocol is zero, so that we can distinguish between the classical and quantum regimes for locking. In this regard, we can exclude all protocols that have a classical simulation in the following sense.

Definition 17: Classical simulation.—We say that a locking protocol has a classical simulation if the receiver's decoding consists of performing a measurement on the output of the channel that is independent of the key K, followed by a classical postprocessing of the measurement output and the key to produce an estimate of the transmitted message.

Theorem 18.—The strong locking rate of any locking protocol with a classical simulation is equal to zero.

Proof: The fact that this theorem should hold might be obvious, but nevertheless we provide a proof. The setup for this proof is similar to that in the proof of Theorems 10 and 11, with the exception that the decoder first performs a key-independent measurement of the channel output to produce a random variable Y. The decoder then processes the random variable Y and the key K to produce an estimate $\hat{M}$ of the sender's message. We can bound the rate R of this protocol as follows:

$$\begin{aligned}
nR &= H(M)_{\bar{\Phi}} \\
&= I(M;\hat{M})_{\bar{\Phi}} \\
&\leq I(M;\hat{M})_{\omega'} + n\epsilon' \\
&\leq I(M;YK)_{\omega} + n\epsilon' \\
&= I(M;Y)_{\omega} + I(M;K|Y)_{\omega} + n\epsilon' \\
&\leq I(M;Y)_{\omega} - I_{\mathrm{acc}}(M;B^nE^n)_{\omega} + o(n) + n\epsilon' + n\epsilon'' \\
&\leq I(M;Y)_{\omega} - I(M;Y)_{\omega} + o(n) + n\epsilon' + n\epsilon'' \\
&= o(n) + n\epsilon' + n\epsilon''.
\end{aligned}$$

The first three lines above are exactly the same as those in the proof of Theorem 10. The second inequality follows from quantum data processing. The third equality is the chain rule. The third inequality follows from the condition $I_{\mathrm{acc}}(M;B^nE^n)_{\omega} \leq \epsilon' n$, with $\epsilon'' \equiv h_2(\epsilon/2)/n + \epsilon R/2$, whenever Eq. (14) is satisfied, which should hold for

any strong locking protocol. Also, it follows because $I(M;K|Y)_\omega \leq H(K) \leq o(n)$. Finally, the adversary can choose their processing of the B^nE^n systems to be a discarding of E^n followed by whatever key-independent measurement of B^n the receiver is performing to produce Y. Thus, it follows that $I_{\text{acc}}(M;B^nE^n)_\omega \geq I(M;Y)$. The statement that the strong locking rate is equal to zero follows by taking the limit as $n \to \infty$ and $\epsilon \to 0$. □

As a corollary of the above theorem, we find the following.

Corollary 19.—If a protocol does not consume any secret key at all, the strong locking rate is equal to zero.

Proof: This follows simply because the receiver's measurement on the channel outputs does not depend on a key for a scheme that does not use any key at all. □

3. The private and quantum capacity are equal to the weak locking capacity for particular Hadamard channels

In this section, we prove that, if the channel is such that the map from the input to the environment is a quantum-to-classical channel, i.e., of the following form,

$$\rho \to \sum_x \text{Tr}\{A_x \rho A_x^\dagger\}|x\rangle\langle x|, \tag{24}$$

for some orthonormal basis $\{|x\rangle\}$ and where $\sum_x A_x^\dagger A_x = I$, then the weak locking capacity of such a channel is equal to its private and quantum capacity. This result follows simply because the systems received by the environment are already classical, so that the best measurement for the adversary to perform is given by $\{|x\rangle\langle x|\}$ on each channel use. Any measurement other than this one will have mutual information with the message lower than this measurement's mutual information by a simple data-processing argument. Furthermore, since the systems given to the environment are classical, the Holevo information of the environment with the input is equal to the accessible information of the environment with the input for such channels.

For such channels, the map from the input to the output is of the following form,

$$\rho \to \sum_x A_x \rho A_x^\dagger \otimes |x\rangle\langle x|, \tag{25}$$

because the operator $\sum_x A_x(\cdot) \otimes |x\rangle \otimes |x\rangle$ is an isometric extension of the channel in Eq. (24). A notable example of such a channel is the "photon detected-jump" channel, described in Ref. [36]. Channels of the form in Eq. (25) are examples of Hadamard channels, which are generally defined as channels complementary to entanglement-breaking ones [37,38].

We state the above result as the following theorem.

Theorem 20.—The weak locking capacity of a channel of the form in Eq. (25) is equal to its private and quantum capacity and is given by the following expression,

$$\max_{\{p_X(x),\rho_x\}} [I(X;B) - I(X;E)],$$

where the information quantities are evaluated with respect to the following state,

$$\sum_x p_X(x)|x\rangle\langle x|_X \otimes U^{\mathcal{N}}_{A\to BE}(\rho_x),$$

with $U^{\mathcal{N}}_{A\to BE}$ an isometric extension of the channel $\mathcal{N}$.

Proof: A proof of this theorem follows the intuition mentioned above. In particular, we know from Refs. [26,27] that the following formula is equal to the private capacity of any channel:

$$P(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n}\{\max_{\{p_X(x),\rho_x^{(n)}\}} [I(X;B^n) - I(X;E^n)]\}.$$

Now, since we are assuming the channel to the environment to have the form given in Eq. (24), the systems given to the environment are classical so that the accessible information $I_{\text{acc}}(X;E^n)$ is equal to the Holevo information $I(X;E^n)$ for any finite n. Thus, our upper bound from Theorem 10 on the weak locking capacity of such a channel is equal to the expression given above for its private capacity. Furthermore, all Hadamard channels are degradable [39], meaning that the receiver can simulate the map from the input to the environment by acting with a degrading map on the system. Finally, it is known that the expression for the private capacity "single letterizes" to the form in the statement of the theorem for degradable channels and that the quantum capacity is equal to the private capacity for such channels [40]. (It is often said that an information theoretic formula is "single-letter" if it is function of a single instance of the channel or resource.) □

Remark 21.—Theorem 20 demonstrates that it suffices to use a private capacity achieving code for channels of the form in Eq. (25), with the benefit that these private communication codes do not require the consumption of any secret key. That is, there is no need to devise an exotic information locking protocol for such channels in order to achieve their weak locking capacity.

4. Quantum discord-based upper bound on the gap between weak locking capacity and private capacity

The quantum discord is an asymmetric measure that quantifies the quantum correlation in a bipartite quantum state [41]. For a given bipartite quantum state ρ_{AB}, the quantum mutual information $I(A;B)_\rho$ quantifies all of the bipartite correlations in ρ_{AB}, while $\max_{\Lambda_{A\to X}} I(X;B)$ is meant to capture the classical correlations in the state that are recoverable by performing a local measurement on the A system [42]. Thus, the idea behind the quantum discord $D(A,B)_\rho$ is to quantify the quantum correlations in a state by subtracting out the classical correlation from the total correlation:

$$D(A,B)_\rho \equiv I(A;B)_\rho - \max_{\Lambda_{A\to X}} I(X;B).$$

Ollivier and Zurek originally described the quantum discord as the correlations lost during a measurement process [41].

Our upper bound on the weak locking capacity from Theorem 10 appears similar to the above formula for quantum discord. Indeed, we can place an upper bound on the gap between the weak locking capacity and the private capacity of a quantum channel in terms of the discord between the environment of the channel and the classical variable sent into the channel. We can also interpret this as merely the gap between the Holevo information of the environment and its accessible information. It is clear why this gap is related to quantum discord. In a private communication protocol, the security guarantee is with respect to the Holevo information, while in a locking protocol, the guarantee is with respect to the accessible information. Thus, the gap between the two capacities should be related to the correlations lost during Eve's measurement.

Proposition 22.—The gap between the weak locking capacity and the private capacity of a quantum channel is no larger than

$$\begin{aligned} L_W(\mathcal{N}) - P(\mathcal{N}) &\leq \lim_{n\to\infty} \frac{1}{n}[\max_{\{p_X(x),\rho_x\}} I(X;E^n) - I_{\text{acc}}(X;E^n)] \\ &= \lim_{n\to\infty} \frac{1}{n}[\max_{\{p_X(x),\rho_x\}} D(E^n,X)], \end{aligned}$$

where the entropies for any finite n are with respect to a state of the following form:

$$\sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to E}^{\otimes n}(\rho_x),$$

and $\mathcal{N}_{A\to E}$ is the channel complementary to $\mathcal{N}_{A\to B} = \mathcal{N}$.

Proof: Consider that for any finite n, we have the bound

$$\begin{aligned} I(X;B^n) - I_{\text{acc}}(X;E^n) &= I(X;B^n) - I(X;E^n) \\ &\quad + I(X;E^n) - I_{\text{acc}}(X;E^n) \\ &\leq \max_{\{p_X(x),\rho_x\}} [I(X;B^n) - I(X;E^n)] \\ &\quad + \max_{\{p_X(x),\rho_x\}} [I(X;E^n) - I_{\text{acc}}(X;E^n)]. \end{aligned}$$

Then, by using the bound from Theorem 10, the inequality above, and the characterization of the private capacity as $P(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n}[\max_{\{p_X(x),\rho_x\}}[I(X;B^n) - I(X;E^n)]]$, the bound in the statement of the theorem follows. □

C. Examples

Example 23: Noiseless qudit channel.—The noiseless qudit channel trivially has weak locking capacity equal to $\log_2 d$, where d is the dimension of the input and the output for the channel. The reason for this is that an isometric extension of this channel has the following form:

$$\sum_i |i\rangle_B\langle i|_A \otimes |\varphi\rangle_E.$$

In this case, Eve's state is independent of the input, so that her accessible information is always equal to zero (even without coding in any way).

However, the noiseless qudit channel nontrivially has strong locking capacity also equal to $\log_2 d$. This follows from the results of Fawzi *et al.* [9], in which they demonstrated the existence of a locking protocol that locks n dits using $4\log_2(1/\epsilon) + O([\log_2\log_2(1/\epsilon))]$ bits of key while having the variational distance in Eq. (14) for any eavesdropper measurement no larger than ϵ, for an eavesdropper who obtains the full output of the noiseless channel. Thus, this scheme is an $(n, \log_2 d, \epsilon)$ locking protocol that consumes the secret key at a rate equal to

$$\frac{1}{n}\{4\log_2(1/\epsilon) + O[\log_2\log_2(1/\epsilon)]\}.$$

So, for any fixed $\epsilon > 0$, we can take n large so that the secret key rate vanishes in this limit, while the eavesdropper will not be able to do much better than to randomly guess the message. Thus, this construction gives a scheme to achieve the rate $\log_2 d$. Since the strong locking capacity of the noiseless qudit channel cannot be any larger than $\log_2 d$, this proves that it is equal to $\log_2 d$ for this channel.

In reality, one does not ever have access to perfectly independent uses of a quantum channel, as this is just an idealization. As such, it can be helpful to define the "one-shot" locking capacity for a single use of a quantum channel. We provide such a definition below.

Definition 24: One-shot locking capacity.—The ϵ-one-shot locking capacity of a quantum channel is the maximum number of locked bits that a sender can transmit to a receiver such that the receiver can recover the message with average error probability less than $\epsilon > 0$ and such that the total variational distance of the message distribution conditioned on the eavesdropper's measurement outcome x with the unconditioned message distribution p_M is no larger than ϵ:

$$\sum_m |p_{M|X}(m|x) - p_M(m)| \leq \epsilon.$$

We also demand that the number of secret key bits used is $O(\log_2\log_2|\mathcal{M}|)$. Similar to the IID case, we can distinguish between weak and strong locking capacities.

Example 25: Depolarizing channel.—Recall that the quantum depolarizing channel is defined as

$$\rho \to (1-p)\rho + p\frac{I}{d},$$

where $p \in [0,1]$ characterizes the noisiness of the channel and d is its dimension. For sufficiently large d, the ϵ-one-shot strong locking capacity of the depolarizing channel is equal to its ϵ-one-shot classical capacity (defined similarly as above—see Ref. [43], for example). This result follows simply because any unitary encoding commutes with the action of the depolarizing channel on the input state, and we can employ the FHS protocol combined with an ϵ-one-shot classical capacity achieving code, in order to achieve the same ϵ-one-shot strong locking capacity of the depolarizing channel.

While it is easy to prove, this example illustrates the subtle interplay between locking, entanglement, and classical communication. The fact that the depolarizing channel's one-shot strong locking and classical capacities match regardless of the strength of the noise would seem to leave little room for quantum correlations to play any role. Indeed, it seems hard to reconcile this result with the statement of Theorem 14 that entanglement-breaking channels have zero strong locking capacity, which is easily adapted to the one-shot setting. The resolution is that, for any fixed but arbitrarily large amount of noise p, the depolarizing channel eventually ceases to be entanglement breaking for some sufficiently large $d = \text{poly}(1/p)$ [44].

Our best known characterization of the locking capacity of the IID memoryless depolarizing channel is in terms of the operational inequalities given in Eqs. (18) and (19).

Example 26: Erasure channel.—Consider a d-dimensional quantum erasure channel defined as

$$\rho \to (1-p)\rho + p|e\rangle\langle e|,$$

where $|e\rangle$ is an erasure flag state that is orthogonal to the d-dimensional input state. For this channel, a unitary acting on the input commutes with the action of the channel, so that the same argument as above demonstrates that the ϵ-one-shot strong locking capacity of this channel is equal to its ϵ-one-shot classical capacity for sufficiently large d.

The feedback-assisted weak and strong locking capacities of the memoryless erasure channel are at least $(1-p)^2$ for $p \leq 1/2$ and $(1-p)/(1+2p)$ for $p \geq 1/2$. Furthemore, they are no larger than $1-p$. These results follow from the best-known lower bounds on the quantum capacity of the erasure channel assisted by classical feedback [30], the fact that the feedback-assisted classical capacity of the erasure channel cannot exceed $1-p$, and the operational inequalities in Eq. (21).

Furthermore Example 27: Parallelized locking protocols.—A simple parallelized protocol (as mentioned in Ref. [7]) is to employ the FHS protocol for each use of a memoryless depolarizing or erasure channel. However, the best-known statement regarding the parallel composition of locking protocols is given by Proposition 2.4 of Ref. [9]. That is, if one locking protocol guarantees that the total variational distance of a message distribution conditioned on the eavesdropper's measurement outcome x_1 with the unconditioned message distribution p_M is no larger than ϵ_1,

$$\sum_{m_1} |p_{M_1|X_1}(m_1|x_1) - p_{M_1}(m_1)| \leq \epsilon_1,$$

and another guarantees it is no larger than ϵ_2,

$$\sum_{m_2} |p_{M_2|X_2}(m_2|x_2) - p_{M_2}(m_2)| \leq \epsilon_2,$$

then the parallel composition of these protocols guarantees a total variational distance no larger than $\epsilon_1 + \epsilon_2$:

$$\sum_{m_1,m_2} |p_{M_1,M_2|X}(m_1,m_2|x) - p_{M_1,M_2}(m_1,m_2)| \leq \epsilon_1 + \epsilon_2.$$

Then, consider a simple parallelized protocol consisting of n uses of a d-dimensional channel, where we suppose that each channel use has a guarantee that the variational distance (as above) is no larger than $\gamma > 0$. Parallel composition of the locking protocols guarantees that the variational distance for the n channel uses is no larger than γn. By applying the Fannes-Audenaert inequality [18,19] as in Proposition 3.2 of Ref. [9], one finds the following bound on the accessible information of the adversary,

$$(\gamma n)\, \log\, d_E^n + h_2(\gamma n),$$

where d_E is the dimension of the environment for a single channel use. Thus, the number of secret key bits needed to guarantee that Eve's accessible information is no larger than $n\epsilon$ is equal to $O[n \log_2 (1/\epsilon)]$, so that the rate of the key used in this scheme grows linearly with the number of channel uses. Clearly, this approach is less desirable than simply using a one-time pad combined with a classical capacity achieving code. For this latter protocol, the rate of the key is a fixed constant independent of the number of channel uses, and the protocol guarantees perfect secrecy from an adversary with access to a quantum memory.

In information theory, results for memoryless channels usually follow straightforwardly from their one-shot counterparts. The linear key growth incurred when parallelizing locking protocols prevents us from quickly concluding that the non-one-shot strong locking capacities of the depolarizing and erasure channels match their classical capacities. Moreover, the covariance argument used to draw that conclusion does not translate directly to the setting of many channel uses. We therefore leave it as an open question to determine whether the equivalence persists beyond the one-shot setting.

VI. UPPER BOUNDS ON THE LOCKING CAPACITY WHEN RESTRICTING TO COHERENT-STATE ENCODINGS

In this section, we prove that there are fundamental limitations on the locking capacity of channels when we restrict ourselves to coherent-state encodings. In particular, we prove that the strong locking capacity of any quantum channel cannot be any larger than

$$g(N_S) - \log_2(1+N_S),$$

where $g(x) \equiv (x+1)\log_2(x+1) - x\log_2 x$, when restricting to coherent-state encodings with mean input photon number N_S. Observe that $g(N_S) - \log_2(1+N_S) \leq \log_2(e)$, and this latter bound is independent of the photon number used for the coherent-state code words. An intuitive (yet not fully rigorous) reason for why we obtain this bound is that $\log_2(1+N_S)$ is the rate of information that an adversary can recover about the message simply by performing heterodyne detection on each input to the channel, while $g(N_S)$ is an upper bound on the classical capacity of any channel with mean input photon number N_S. Thus, the difference of these two quantities should be a bound on the strong locking capacity.

We also prove that the weak locking capacity of a pure-loss bosonic channel cannot be any larger than the sum of its private capacity and

$$g[(1-\eta)N_S] - \log_2[1+(1-\eta)N_S],$$

when restricting to coherent-state encodings with mean photon number N_S, where $\eta \in [0,1]$ is the transmissivity of the channel. As before, $g[(1-\eta)N_S] - \log_2[1+(1-\eta)N_S] \leq \log_2(e)$, which is independent of the photon number.

We consider a coherent-state locking protocol in which the encrypted states $\{U_k|m\rangle\}$ are generalized to a set of n-mode coherent states $\{|\alpha^n(m,k)\rangle\}_{m\in\mathcal{M},k\in\mathcal{K}}$, where $|\alpha^n(m,k)\rangle$ is an n-fold tensor product of coherent states:

$$|\alpha^n(m,k)\rangle \equiv |\alpha_1(m,k)\rangle \otimes \cdots \otimes |\alpha_n(m,k)\rangle.$$

Definition 28: Coherent-state locking protocol.—A coherent-state locking protocol consists of coherent-state code words $\{|\alpha^n(m,k)\rangle\}_{m\in\mathcal{M},k\in\mathcal{K}}$ depending upon the message m and the key value k. These code words are then transmitted over a quantum channel to be decoded by a receiver.

Theorem 29.—The strong locking capacity of any channel when restricting to coherent-state encodings with mean photon number N_S is upper bounded by $g(N_S) - \log_2(1+N_S)$.

Proof.—As described in Definition 28, the encoder for such a scheme prepares a coherent-state code word $|\alpha^n(m,k)\rangle$ at the input of n uses of a quantum channel $\mathcal{N}$, depending upon the message m and the key value k. It is useful for us to consider the following classical-quantum state, which describes the state of the message, key, and input to many uses of the channel:

$$\rho_{MKA^n} = \frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}|m,k\rangle\langle m,k|_{MK} \otimes |\alpha^n(m,k)\rangle\langle\alpha^n(m,k)|_{A^n}. \tag{26}$$

After the isometric extension of the channel acts (it is unique up to unitaries acting on the environment), the state is then as follows:

$$\rho_{MKB^nE^n} = \frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_{m,k}|m,k\rangle\langle m,k|_{MK} \otimes U^{\mathcal{N}}_{A^n\to B^nE^n}[|\alpha^n(m,k)\rangle\langle\alpha^n(m,k)|_{A^n}],$$

where $U^{\mathcal{N}}_{A^n\to B^nE^n}$ is the isometry corresponding to n uses of the given channel. Recall from the proof of Theorem 11 that we obtain the following upper bound on the strong locking capacity of $\mathcal{N}$:

$$I(M;B^n) - I_{\text{acc}}(M;B^nE^n) + o(n) + n2\epsilon'. \tag{27}$$

[Recall that this bound holds for any (n,R,ϵ) strong locking protocol, with ϵ' a function of ϵ that vanishes as $\epsilon \to 0$.] Consider that the information quantity $I(M;B^n)$ is upper bounded as follows:

$$\begin{aligned} I(M;B^n)_\rho &\leq I(M;A^n)_\rho \\ &= I(MK;A^n)_\rho - I(K;A^n|M)_\rho, \end{aligned}$$

where the first inequality follows from quantum data processing, and the equality follows from the chain rule for quantum mutual information. We then find that

$$\begin{aligned} I(MK;A^n)_\rho &= H(A^n)_\rho - H(A^n|MK)_\rho \\ &= H(A^n)_\rho, \end{aligned} \tag{28}$$

where the second equality follows because the state on A^n is a pure coherent state when conditioned on systems M and K.

On the other hand, we obtain a lower bound on the accessible information $I_{\text{acc}}(M;B^nE^n) = I_{\text{acc}}(M;A^n)$ by having the adversary perform heterodyne detection (a particular measurement that is not necessarily the optimal one) on each of the systems A^n, giving

$$I_{\text{acc}}(M;A^n)_\rho \geq I_{\text{het}}(M;A^n)_\rho \tag{29}$$

$$= I_{\text{het}}(MK;A^n)_\rho - I_{\text{het}}(K;A^n|M)_\rho, \tag{30}$$

where in the second line we again apply the chain rule for mutual information. An ideal n-mode heterodyne

measurement is described by a POVM $\{\frac{d^{2n}\beta^n}{\pi^n}|\beta^n\rangle\langle\beta^n|\}$, where β^n is the amplitude of the n-mode coherent state $|\beta^n\rangle \equiv |\beta_1\rangle\cdots|\beta_n\rangle$ and $d^{2n}\beta^n$ denotes the Lebesgue measure on $\mathbb{C}^n$. We can then compute the heterodyne mutual information $I_{\text{het}}(MK;A^n)_\rho$ as

$$I_{\text{het}}(MK;A^n)_\rho = W(A^n)_\rho - W(A^n|MK)_\rho,$$

where

$$W(Q)_\sigma = -\int \frac{d^{2n}\beta^n}{\pi^n}\langle\beta^n|\sigma|\beta^n\rangle\log_2\langle\beta^n|\sigma|\beta^n\rangle \tag{31}$$

denotes the Wehrl entropy for a state σ defined on system Q [45] and its conditional version follows in the natural way. It is easy to see that the Wehrl entropy of an n-mode coherent state is equal to $n\log_2(e)$, so we find that

$$I_{\text{het}}(MK;A^n)_\rho = W(A^n)_\rho - n\log_2(e). \tag{32}$$

We are now in a position to derive an upper bound on Eq. (27). Observe that our development above implies that

$$\begin{aligned}
&I(M;B^n)-I_{\text{acc}}(M;B^nE^n)\\
&\quad\le I(MK;A^n)_\rho - I(K;A^n|M)_\rho\\
&\quad\quad -[I_{\text{het}}(MK;A^n)_\rho - I_{\text{het}}(K;A^n|M)_\rho]\\
&\quad\le I(MK;A^n)_\rho - I_{\text{het}}(MK;A^n)_\rho\\
&\quad\le \max_{p_X(x)}[I(X;A^n)_\omega - I_{\text{het}}(X;A^n)_\omega]\\
&\quad\le n\max_{p_X(x)}[I(X;A)_\sigma - I_{\text{het}}(X;A)_\sigma]\\
&\quad= n\{\log_2(e)+\max_{p_X(x)}[H(A)_\sigma - W(A)_\sigma]\}\\
&\quad\le n[g(N_S)-\log_2(1+N_S)].
\end{aligned} \tag{33}$$

The second inequality follows from data processing: $I(K;A^n|M)_\rho \ge I_{\text{het}}(K;A^n|M)_\rho$ (the system M is classical, and performing heterodyne detection on A^n can only reduce the mutual information). The third inequality follows by taking a maximization over all distributions $p_X(x)$, where ω_{XA^n} is a state of the following form:

$$\omega_{XA^n} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\alpha_x^n\rangle\langle\alpha_x^n|_{A^n},$$

such that the mean input photon number to the channel for each x is N_S. The fourth inequality follows by realizing that the difference between the mutual information and the heterodyne information is equal to the private information of a quantum wiretap channel in which the state $|\alpha_x^n\rangle$ is prepared for the receiver while the heterodyned version of this state (a classical variable) is prepared for the eavesdropper. Such a quantum wiretap channel has pure product input states (they are coherent states) and it is degraded. Thus, we can apply Theorem 35 from the Appendix to show that this private information is subadditive, in the sense that

$$\begin{aligned}
&\max_{p_X(x)}[I(X;A^n)_\omega - I_{\text{het}}(X;A^n)_\omega]\\
&\quad\le n\max_{p_X(x)}[I(X;A)_\sigma - I_{\text{het}}(X;A)_\sigma],
\end{aligned}$$

where we define the state σ_{XA} as follows:

$$\sigma_{XA} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\alpha_x\rangle\langle\alpha_x|_A.$$

The last equality follows from the observation in Eq. (32) and because $I(X;A)_\sigma = H(A)_\sigma - H(A|X)_\sigma = H(A)_\sigma$ (since the states are pure when conditioned on X).

We now show that the maximizing distribution for $\max_{p_X(x)}[H(A)_\sigma - W(A)_\sigma]$ is given by a circularly symmetric Gaussian distribution with variance N_S, so that the optimal ensemble is a Gaussian ensemble of coherent states. Indeed, let ϱ be a single-mode quantum state with $\text{Tr}[a\varrho] = 0$ and $\text{Tr}[a^\dagger a\varrho] = N_S$, where $a^\dagger$ and a are creation and annihilation operators, respectively. The von Neumann entropy is given by $H(\varrho) = -\text{Tr}[\varrho\log_2\varrho]$. We show that

$$H(\varrho) - W(\varrho) \tag{34}$$

is maximized when ϱ is a thermal state. Our approach is based on a technique used in the Appendix of Ref. [46], which in turn is based on classical approaches to this problem [12]. Let

$$\tilde{\varrho} = \frac{1}{N_S+1}\sum_{m=0}^{\infty}\left(\frac{N_S}{N_S+1}\right)^m |m\rangle\langle m| \tag{35}$$

be a thermal state with mean photon number N_S. We will show that

$$H(\tilde{\varrho}) - W(\tilde{\varrho}) - (H(\varrho) - W(\varrho)) \ge 0 \tag{36}$$

holds for any ϱ with $\text{Tr}[a\varrho] = 0$ and $\text{Tr}[a^\dagger a\varrho] = N_S$. Putting

$$\mathcal{Q}_\varrho(\beta) = \langle\beta|\varrho|\beta\rangle, \tag{37}$$

the left-hand side of (36) is equal to

$$
\begin{aligned}
&-\mathrm{Tr}[\tilde{\varrho}\log_2\tilde{\varrho}]+\mathrm{Tr}[\varrho\log_2\varrho]+\int\frac{d^2\beta}{\pi}\mathcal{Q}_{\tilde{\varrho}}(\beta)\log_2\mathcal{Q}_{\tilde{\varrho}}(\beta)-\int\frac{d^2\beta}{\pi}\mathcal{Q}_{\varrho}(\beta)\log_2\mathcal{Q}_{\varrho}(\beta)\\
&=\mathrm{Tr}[\varrho(\log_2\varrho-\log_2\tilde{\varrho})]+\mathrm{Tr}[(\varrho-\tilde{\varrho})\log_2\tilde{\varrho}]-\left\{\int\frac{d^2\beta}{\pi}\mathcal{Q}_{\varrho}(\beta)[\log_2\mathcal{Q}_{\varrho}(\beta)-\log_2\mathcal{Q}_{\tilde{\rho}}(\beta)]+\int\frac{d^2\beta}{\pi}[\mathcal{Q}_{\varrho}(\beta)-\mathcal{Q}_{\tilde{\varrho}}(\beta)]\log_2\mathcal{Q}_{\tilde{\varrho}}(\beta)\right\}\\
&=D(\varrho||\tilde{\varrho})-D(\mathcal{Q}_{\varrho}||\mathcal{Q}_{\tilde{\varrho}})+\mathrm{Tr}[(\varrho-\tilde{\varrho})\log_2\tilde{\varrho}]-\int\frac{d^2\beta}{\pi}[\mathcal{Q}_{\varrho}(\beta)-\mathcal{Q}_{\tilde{\varrho}}(\beta)]\log_2\mathcal{Q}_{\tilde{\varrho}}(\beta),
\end{aligned}
\tag{38}
$$

where $D(\varrho||\tilde{\varrho})$ and $D(\mathcal{Q}_{\varrho}||\mathcal{Q}_{\tilde{\varrho}})$ are quantum and classical relative entropies, respectively. We can easily show that their difference is positive by the monotonicity property of the relative entropy. The third term is

$$
\begin{aligned}
\mathrm{Tr}[(\varrho-\tilde{\varrho})\log_2\tilde{\varrho}]&=\mathrm{Tr}\left\{(\varrho-\tilde{\varrho})\sum_{m=0}^{\infty}\log_2\left[\frac{1}{N_S+1}\left(\frac{N_S}{N_S+1}\right)^{a^\dagger a}\right]|m\rangle\langle m|\right\}\\
&=-\log_2(N_S+1)\mathrm{Tr}[\varrho-\tilde{\varrho}]+\log_2\left(\frac{N_S}{N_S+1}\right)\mathrm{Tr}[(\varrho-\tilde{\varrho})a^\dagger a]\\
&=0.
\end{aligned}
\tag{39}
$$

Similarly, the fourth term is

$$
\int\frac{d^2\beta}{\pi}[\mathcal{Q}_{\varrho}(\beta)-\mathcal{Q}_{\tilde{\varrho}}(\beta)]\log_2\mathcal{Q}_{\tilde{\varrho}}(\beta)=\int\frac{d^2\beta}{\pi}[\mathcal{Q}_{\varrho}(\beta)-\mathcal{Q}_{\tilde{\varrho}}(\beta)]\left[-\log_2(N_S+1)-\frac{|\beta|^2}{\ln(2)(N_S+1)}\right].
\tag{40}
$$

$$
=0.
\tag{41}
$$

Note that $\mathcal{Q}_{\tilde{\varrho}}(\beta)=\frac{1}{(N_S+1)}\exp[-\frac{|\beta|^2}{N_S+1}]$, and we used the fact that if $\mathrm{Tr}[a^\dagger a\varrho]=\mathrm{Tr}[a^\dagger a\tau]$, then

$$
\int d^2\beta\mathcal{Q}_{\varrho}(\beta)|\beta|^2=\int d^2\beta\mathcal{Q}_{\tau}(\beta)|\beta|^2.
$$

As a consequence, we have

$$
H(\tilde{\varrho})-W(\tilde{\varrho})-[H(\varrho)-W(\varrho)]=D(\varrho||\tilde{\varrho})-D(\mathcal{Q}_{\varrho}||\mathcal{Q}_{\tilde{\varrho}})\geq 0,
\tag{42}
$$

which completes the proof that $\max_{p_X(x)}[H(A)_\sigma-W(A)_\sigma]$ is optimized by a circularly symmetric complex Gaussian distribution with variance N_S.

Finally, we can rewrite $\log_2(e)+\max_{p_X(x)}[H(A)_\sigma-W(A)_\sigma]$ as $I(X;A)_\sigma-I_{\mathrm{het}}(X;A)_\sigma$ for X complex Gaussian, and these information quantities evaluate to $g(N_S)-\log_2(1+N_S)$ in such a case. By combining the bounds in Eqs. (27) and (33), we deduce the following upper bound on the rate R of any strong locking protocol that employs coherent-state code words with mean photon number N_S:

$$
R\leq g(N_S)-\log_2(1+N_S)+\frac{o(n)}{n}+2\epsilon',
$$

which converges to $g(N_S)-\log_2(1+N_S)$ in the limit as $n\to\infty$ and $\epsilon\to 0$. □

Theorem 30.—The weak locking capacity of a pure-loss bosonic channel with transmissivity $\eta\in[0,1]$ when restricting to coherent-state encodings with mean input photon number N_S is upper bounded by

$$
\begin{aligned}
&\max\{0,g(\eta N_S)-g[(1-\eta)N_S]\}\\
&\quad+\{g[(1-\eta)N_S]-\log_2[1+(1-\eta)N_S]\}.
\end{aligned}
$$

The term $\max\{0,g(\eta N_S)-g[(1-\eta)N_S]\}$ is equal to the private capacity of the pure-loss bosonic channel, while the second term is limited by the bound

$$
\{g[(1-\eta)N_S]-\log_2[1+(1-\eta)N_S]\}\leq\log_2(e).
$$

Proof: The proof of this theorem is somewhat similar to the proof of the previous theorem. Nevertheless, there are

some important differences, so we give the full proof for completeness.

In the proof of Theorem 10, we obtain the following upper bound on the weak locking capacity:

$$I(M;B^n) - I_{\text{acc}}(M;E^n) + o(n) + n2\epsilon'. \tag{43}$$

[Recall that this bound holds for any (n, R, ϵ) strong locking protocol, with ϵ' a function of ϵ that vanishes when $\epsilon \to 0$.] We begin by bounding the quantity $I(M;B^n) - I_{\text{acc}}(M;E^n)$:

$$\begin{aligned}
&I(M;B^n) - I_{\text{acc}}(M;E^n) \\
&\quad \leq I(MK;B^n) - [I_{\text{het}}(MK;E^n) - I_{\text{het}}(K;E^n|M)] \\
&\quad \leq I(MK;B^n) - I_{\text{het}}(MK;E^n) + o(n) \\
&\quad = H(B^n) - W(E^n) + n\log_2(e) + o(n) \\
&\quad = H(B^n) - H(E^n) + H(E^n) - W(E^n) \\
&\qquad + n\log_2(e) + o(n) \\
&\quad \leq n(\max\{0, g(\eta N_S) - g[(1-\eta)N_S]\}) \\
&\quad + n\{g[(1-\eta)N_S] - \log_2[1 + (1-\eta)N_S]\} + o(n).
\end{aligned}$$

The first inequality follows from data processing $I(M;B^n) \leq I(MK;B^n)$, the fact that $I_{\text{acc}}(M;E^n) \geq I_{\text{het}}(MK;E^n)$, and the identity $I_{\text{het}}(M;E^n) = I_{\text{het}}(MK;E^n) - I_{\text{het}}(K;E^n|M)$. The second inequality follows because $I_{\text{het}}(K;E^n|M) \leq H(K) \leq o(n)$. The first equality follows from the fact that $I(MK;B^n) = H(B^n)$ for the pure-loss bosonic channel and from the fact that $I_{\text{het}}(MK;E^n) = W(E^n) - n\log_2(e)$. The second equality is a simple identity. The final inequality follows because the entropy difference $H(B^n) - H(E^n)$ is equal to a coherent information of the n-use pure-loss bosonic channel. The only relevant property of the input state for which the coherent information is evaluated is that it has a mean photon number N_S, and so the coherent information is always lower than $n\max\{0, g(\eta N_S) - g[(1-\eta)N_S]\}$, which is equal to n times the quantum and private capacity of this channel [47,48]. We also employ an argument similar to that in the previous theorem to bound $H(E^n) - W(E^n) + n\log_2(e)$ from above by $n\{g[(1-\eta)N_S] - \log_2[1 + (1-\eta)N_S]\}$. Finally, by combining the above bound with the bound in Eq. (43), we deduce the following upper bound on the rate R of any weak locking protocol that employs coherent-state code words for transmission over a pure-loss bosonic channel:

$$\begin{aligned}
R \leq\ &\max\{0, g(\eta N_S) - g[(1-\eta)N_S]\} \\
&+ \{g[(1-\eta)N_S] - \log_2[1 + (1-\eta)N_S]\} + \frac{o(n)}{n} + 2\epsilon',
\end{aligned}$$

which converges to $\max\{0, g(\eta N_S) - g[(1-\eta)N_S]\} + \{g[(1-\eta)N_S] - \log_2[1 + (1-\eta)N_S]\}$ in the limit as $n \to \infty$ and $\epsilon \to 0$. □

Remark 31.—Given that the private capacity of a pure-loss bosonic channel with mean input photon number N_S is equal to $\max\{0, g(\eta N_S) - g[(1-\eta)N_S]\}$, the above theorem implies a strong limitation on the weak locking capacity of a pure-loss bosonic channel when restricting to coherent-state encodings with mean input photon number N_S. That is, the weak locking capacity when restricting to coherent-state encodings cannot be more than 1.45 bits larger than the channel's private capacity.

Remark 32.—These bounds apply, in particular, to channels that use a coherent-state locking protocol in which there is a fixed code book $\{|\alpha^n(m)\rangle\}$ and the coherent states are encrypted according to passive mode transformations U_k that transform n-mode coherent states as $|\alpha^n\rangle \to |\tilde{U}_k\alpha^n\rangle$, where $\tilde{U}_k\alpha^n$ is understood to be a label for a coherent-state vector with the following complex amplitudes:

$$\begin{bmatrix} \tilde{U}_k^{(1,1)} & \cdots & \tilde{U}_k^{(1,n)} \\ \vdots & \ddots & \vdots \\ \tilde{U}_k^{(n,1)} & \cdots & \tilde{U}_k^{(n,n)} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}. \tag{44}$$

Remark 33.—If the coherent-state locking protocol consists of passive mode transformations (as defined above) for the encryption and the receiver performs heterodyne detection to recover the message after decrypting with a passive mode transformation, then the strong locking capacity of a channel using such a scheme is equal to zero. This result follows because such a scheme has a classical simulation—passive mode transformations commute with heterodyne detection in such a way that heterodyne detection can be performed first followed by a classical postprocessing of the measurement data with a matrix multiplication as in Eq. (44). That is, the decoding in such a scheme is equivalent to first performing key-independent heterodyne detection measurements followed by classical postprocessing of the key and the measurement results. Thus, Theorem 18 applies so that the strong locking capacity of a channel using such a scheme is equal to zero. However, this theorem does not apply if the receiver performs photodetection because passive mode transformations do not commute with such a measurement.

From our upper bounds on the locking capacity of channels restricted to coherent-state encodings, it is clear that there are strong limitations on the rates that are achievable when employing bright coherent states. That is, it clearly would not be worthwhile to invest a large mean input photon number per transmission given the above limitations on locking capacity that are independent of the photon number. In spite of this result, it might be possible

to achieve interesting locking rates with weak coherent states, but we should keep in mind that the above bounds were derived by considering the information that an adversary can gain by performing heterodyne detection—the information of the adversary can only increase if they perform a better measurement. Nevertheless, we can determine values of the mean input photon number N_S such that the difference $g(N_S) - \log_2(1+N_S)$ becomes relatively large. By considering $N_S \ll 1$, we find the following expansions of $g(N_S)$ and $\log_2(1+N_S)$, respectively,

$$g(N_S) \approx \left(-N_S \ln N_S + N_S + \frac{N_S^2}{2}\right)\log_2(e), \quad (45)$$

$$\log_2(1+N_S) \approx \left(N_S - \frac{N_S^2}{2}\right)\log_2(e), \quad (46)$$

so that the difference $g(N_S) - \log_2(1+N_S) \approx [-N_S \ln N_S + N_S^2]\log_2(e)$ for $N_S \ll 1$. Indeed, in the limit as $N_S \to 0$, we find that the relative ratio of our upper bound on the strong locking capacity of a coherent-state protocol to the classical capacity $g(N_S)$ of the noiseless bosonic channel approaches one:

$$\lim_{N_S \to 0} \frac{g(N_S) - \log_2(1+N_S)}{g(N_S)} = 1,$$

so that there is some sense in which the rate at which we can lock information becomes similar to the rate at which we can insecurely communicate information if the bound $g(N_S) - \log_2(1+N_S)$ is, in fact, achievable. However, this remains an important open question.

VII. ONE-SHOT PPM COHERENT-STATE LOCKING PROTOCOL

In spite of the limitations on the locking capacity of coherent-state protocols in the previous section, we still think it is interesting to explore what kind of locking protocols are possible using coherent-state encodings. To this end, we now discuss a one-shot strong locking protocol (in the sense of Definition 24) that employs a coherent-state encoding. For simplicity, we consider the case of a noiseless channel, with the generalization to the pure-loss bosonic channel and weak locking being straightforward.

An explicit scheme for locking using weak coherent states can be obtained by analogy with the PPM encryption presented in Ref. [7] and reviewed in Sec. IVA. Similar to the single-photon scheme, to encode a message m, Alice prepares an n-mode coherent state $|\alpha_m\rangle$, which is a tensor product of a single-mode coherent state of amplitude α on the mth mode and the vacuum on the remaining $n-1$ modes:

$$|\alpha_m\rangle \equiv |0\rangle_1 \ldots |0\rangle_{m-1} |\alpha\rangle_m |0\rangle_{m+1} \ldots |0\rangle_n.$$

(Note that, as in the single-photon case, the PPM encoding is highly inefficient in terms of the number of modes, as it encodes only $\log_2 n$ bits into n bosonic modes.) Let us fix $N_{\text{tot}} = |\alpha|^2$ to be the total mean number of photons involved in the protocol and

$$N_S \equiv \frac{N_{\text{tot}}}{n} \quad (47)$$

to be the mean photon number per mode. Before sending anything to Bob, Alice encrypts a message by applying a unitary selected uniformly at random (according to the shared secret key) from a set of $|\mathcal{K}|$ n-mode linear-optical passive transformations. If the unitary U_k is used, then the final state is the n-mode coherent state

$$U_k|\alpha_m\rangle = \bigotimes_{m'=1}^{n} |\tilde{U}_k^{(m',m)}\alpha\rangle. \quad (48)$$

Bob, who knows which unitary has been chosen by Alice, applies the inverse transformation and performs photodetection on the received modes. He will detect (one or more) photons only in the mth mode, hence successfully decrypting the message in case of a detection.

Different from the single-photon architecture of Ref. [7], there is a nonzero probability that Bob's detector does not click. Analogous to the case of the single-photon locking protocol in the presence of loss, if no photon is detected, Bob may use a public classical communication channel to ask Alice to resend, yielding

$$I_{\text{acc}}(M;KQ)_\rho = N_{\text{tot}} \log_2 n. \quad (49)$$

However, the same observations from Sec. IVA apply here. That is, locking is only known to be secure when the message distribution is uniform, and this is certainly not the case for a feedback-assisted scheme unless Eve attacks each PPM block independently. If she attacks collectively, then it is necessary for Alice and Bob to exploit an amount of key necessary to ensure that the message distribution is uniform.

Assuming that Eve independently attacks each block that she receives, we have to evaluate her accessible information with respect to the following state:

$$\rho_{MKQ} = \frac{1}{n|\mathcal{K}|} \sum_{m,k} |m,k\rangle\langle m,k|_{MK} \otimes (U_k|\alpha_m\rangle\langle\alpha_m|U_k^\dagger)_Q. \quad (50)$$

If the set of n-mode unitaries is selected uniformly at random according to the Haar measure, one might expect that such a set of unitaries scrambles phase information of α so that it is not accessible to Eve. Thus, a presumably clever strategy for Eve is to perform a measurement that commutes with the total number of photons. Such a POVM has the elements

$$\{|0\rangle\langle 0|, \{\mu_y^{(1)}|\varphi_y^{(1)}\rangle\langle\varphi_y^{(1)}|\}_y, \{\mu_y^{(2)}|\varphi_y^{(2)}\rangle\langle\varphi_y^{(2)}|\}_y, \ldots\}, \quad (51)$$

where $|0\rangle$ is the n-mode vacuum, and for any $k \geq 1$, each vector $|\varphi_y^{(k)}\rangle$ belongs to the k-photon subspace. This suboptimal measurement allows Eve to achieve a mutual information $I_{\text{num}}(M;Q)_\rho$ such that

$$I_{\text{num}}(M;Q)_\rho \leq I_{\text{acc}}(M;Q)_\rho. \quad (52)$$

For small values of $N_{\text{tot}} \ll 1$, the probability of having more than one photon is of order N_{tot}^2. We can hence argue that, for $N_{\text{tot}}^2 \ll 1$, the main contribution to $I_{\text{num}}(M;Q)_\rho$ comes from POVM elements in Eq. (51), with $k = 0, 1$, and that the contribution of those with $k \geq 2$ is, in the worst case, of order $N_{\text{tot}}^2 \log_2 n$. Noting that the projection of the coherent state $U_k|\alpha_m\rangle$ in the subspace spanned by the vacuum and the single-photon subspace is

$$\begin{aligned} &e^{-|\alpha|^2/2}(|0\rangle + \alpha U_k|m\rangle) \\ &\quad = e^{-|\alpha|^2/2}\left(|0\rangle + \alpha\sum_{m'}\tilde{U}_k^{(m',m)}|m'\rangle\right), \end{aligned} \quad (53)$$

where $|m'\rangle$ is the state of a single photon on the m'th mode, a straightforward calculation leads to the following expression for the lower bound $I_{\text{num}}(M;Q)_\rho$:

$$\begin{aligned} I_{\text{num}}(M;Q)_\rho \leq N_{\text{tot}}&\left[\log_2 n - \min_{\mathcal{M}_E^{(1)}}\sum_y \frac{\mu_y^{(1)}}{n|\mathcal{K}|}\sum_k H(q_{yk})\right] \\ &+ O(N_{\text{tot}}^2 \log_2 n), \end{aligned} \quad (54)$$

where the optimization is over the POVM $\mathcal{M}_E^{(1)}$ defined on the single-photon subspace, with elements $\{\mu_y^{(1)}|\varphi_y^{(1)}\rangle\langle\varphi_y^{(1)}|\}_i$, and $q_{yk}^m = |\langle\varphi_y^{(1)}|U_k|m\rangle|^2$.

The expression in square brackets in Eq. (54) is formally the same as that in Eq. (5). We can hence bound $I_{\text{num}}(M;Q)_\rho$ using the results of Ref. [9]. It follows that there exist choices of $|\mathcal{K}|$ n-mode passive linear-optical unitaries with

$$\log_2|\mathcal{K}| = 4\log_2(\epsilon^{-1}) + O[\log_2\log_2(\epsilon^{-1})], \quad (55)$$

such that

$$I_{\text{num}}(M;Q)_\rho \leq \epsilon N_{\text{tot}}\log_2 n + O(N_{\text{tot}}^2\log_2 n), \quad (56)$$

with ϵ arbitrarily small if n is large enough.

Clearly, the security condition $r_1 \ll 1$ can be satisfied only if $I_{\text{num}}(M;Q)_\rho \ll I_{\text{acc}}(M;QK)_\rho$. A necessary condition for $r_1 \ll 1$ to hold is

$$\epsilon + O(N_{\text{tot}}) \ll 1, \quad (57)$$

which can be fulfilled in the case of weak coherent states, where $N_{\text{tot}} \ll 1$. In this regime, the key-efficiency condition $r_2 < 1$ can be satisfied only if

$$4\log_2(\epsilon^{-1}) < N_{\text{tot}}\log_2 n. \quad (58)$$

This implies that the value of N_{tot} has to be in the range

$$1 \gg N_{\text{tot}} > \frac{4\log_2(\epsilon^{-1})}{\log_2 n}. \quad (59)$$

In conclusion, this weak coherent state PPM protocol is analogous to the single-photon one in the presence of linear loss. In principle, the condition in Eq. (59) can always be fulfilled for n large enough, yet the minimum value of n increases exponentially with decreasing key-efficiency ratio r_2 and with decreasing N_{tot}.

VIII. CONCLUSION

In this paper, we formally define the locking capacity of a quantum channel in order to establish a framework for understanding the locking effect in the presence of noise. We distinguish between a weak locking capacity and a strong one, the difference being whether the adversary has access to the environment of the channel or to its input. We relate these locking capacities to other well-known capacities from quantum Shannon theory, such as the quantum, private, and classical capacity. The existence of the FHS locking protocol [9] establishes that both the weak and the strong capacity locking capacities are not smaller than the quantum capacity, while the weak locking capacity is not smaller than the private capacity because a private communication protocol always satisfies the demands of a weak locking protocol. Furthermore, the classical capacity is a trivial upper bound on both locking capacities. We also prove that the strong locking capacity is equal to zero whenever a locking protocol has a classical simulation and that both locking capacities are equal to zero for an entanglement-breaking channel. This latter result demonstrates that a channel should have some ability to preserve entanglement in order for nonzero locking rates to be achievable. Moreover, we find a class of channels for which the weak locking capacity is equal to both the private capacity and the quantum capacity.

As an important application, we consider the case of the pure-loss bosonic channel and the locking capacities for channels restricted to coherent-state encodings. We note that a particular example of such a protocol is the $\alpha\eta$ protocol (also known as $Y00$) [49]. We find limitations of the locking capacity for these coherent-state schemes: the strong locking capacity of any channel is not larger than $\log_2(e)$ locked bits per channel use, while the weak locking capacity of the pure-loss bosonic channel is not larger than the sum of its private capacity and $\log_2(e)$ locked bits per channel use. If the scheme exploits passive mode

transformations and the receiver uses heterodyne detection, the restrictions are as severe as they can be: the strong locking capacity is equal to zero because there is a classical simulation of such a protocol.

As a final contribution, we discuss locking schemes that exploit weak coherent states and that might be physically implementable. They are similar to the single-photon QEM of Ref. [7], with the exception that information is encoded by PPM of a coherent state of a given amplitude α, with $|\alpha|^2 \ll 1$, over n modes. The necessary conditions for security and key efficiency of this scheme are qualitatively equivalent to that of the single-photon QEM.

The realization of a proof-of-principle demonstration of a quantum enigma machine is a tremendous experimental challenge. The main difficulty to overcome concerns the scaling of the physical resources required for key efficient encryption. Note that, for single-photon PPM encoding, n optical modes are needed to encode $\log_2 n$ bits, while the required secret key has length of the order of $\log_2 \log_2 n$. As a consequence, the number of modes increases very quickly if one requires small values of the key efficiency ratio r_2, as defined in Eq. (10). Although keeping the same scaling law, the resources required for a key efficient single-photon QEM become even more demanding when one introduces loss in the single-photon scheme and when one moves from the single-photon PPM to the weak coherent-state PPM. On the other hand, in the case of a weak coherent-state QEM, one can trade off the increase in the number of modes (and/or the reduction in the key efficiency level) with the fact that coherent states can be prepared deterministically and are much easier to handle than single-photon states. However, it seems that one has to go beyond PPM encoding to overcome the key efficiency limitations.

There are many open questions to consider going forward. Perhaps the most pressing question is to determine a formula that serves as a good lower bound on the locking capacities (that is, one would need to demonstrate locking protocols with nontrivial achievable rates according to the requirements stated in Definitions 1, 2, and 8). One might suspect that the formulas given in Theorems 10 and 11 are in fact achievable, but it is not clear to us if this is true. Furthermore, it is important to determine if there is an example of a channel (perhaps many?) for which its weak locking capacity is strictly larger than its private classical capacity, and similarly, if there is a channel for which its strong locking capacity is strictly larger than its quantum capacity. We also suspect that the locking capacity is nonadditive, as is the case for other capacities in quantum Shannon theory [31,50,51]. If it is the case that the 50% quantum erasure channel has a weak locking capacity equal to zero, then it immediately follows from the results of Ref. [31] and our operational bounds in Eqs. (18) and (19) that both the weak and the strong locking capacities are nonadditive.

Another intriguing question is the relationship between the strong locking and quantum identification capacities [52] of a quantum channel. Both seem to involve a weak form of coherent data transmission from a sender to a receiver. FHS even established an explicit connection between locking using unitary encodings and quantum identification over a channel built out of the inverses of those unitaries [9]. It is tempting to speculate that the single-letter formula for the amortized quantum identification capacity found in Ref. [53] could thereby be recruited as a tool to study the locking capacity.

ACKNOWLEDGMENTS

We are grateful to Omar Fawzi, Graeme Smith, and Andreas Winter for helpful discussions. This research was supported by the DARPA Quiness Program through U.S. Army Research Office Grant No. W31P4Q-12-1-0019. P. H. was supported by the Canada Research Chairs program, CIFAR, NSERC, and ONR through Grant No. N000140811249.

APPENDIX: PRIVATE CAPACITY OF DEGRADED QUANTUM WIRETAP CHANNELS

This Appendix contains a proof that the private capacity of a degraded quantum wiretap channel when restricted to product-state encodings is single letter. Also, we show by an appeal to Hastings's counterexample to the additivity conjecture [54] that there exist two quantum wiretap channels that are degraded but nevertheless have nonadditive private information. This latter result provides a simple answer to a question that has been open since the introduction of weakly degradable channels [55].

A quantum wiretap channel is defined as a completely positive trace-preserving map $\mathcal{N}_{A\to BE}$ from an input system A to a legitimate receiver's system B and an eavesdropper's system E. Such a map has an isometric extension $U_{A\to BEF}$, with the property that

$$\mathcal{N}_{A\to BE}(\rho) = \mathrm{Tr}_F\{U_{A\to BEF}\rho U^\dagger_{A\to BEF}\}.$$

The private capacity of a quantum wiretap channel is given by [26,27]

$$\lim_{n\to\infty}\frac{1}{n}P(\mathcal{N}^{\otimes n}_{A\to BE}),$$

where $P(\mathcal{N}_{A\to BE})$ is the private information, defined as

$$P(\mathcal{N}_{A\to BE}) \equiv \max_{\{p_X(x),\rho_x\}} I(X;B)_\rho - I(X;E)_\rho, \tag{A1}$$

with the entropies taken with respect to the following classical-quantum state:

$$\rho_{XBE} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to BE}(\rho_x). \tag{A2}$$

Such a wiretap channel is degraded if there exists a degrading map $\mathcal{D}_{B\to E}$ such that

$$\mathcal{D}_{B\to E}\circ\mathcal{N}_{A\to B} = \mathcal{N}_{A\to E}.$$

Theorem 34.—The private information formula in Eq. (A1) for two degraded quantum wiretap channels is generally nonadditive. That is, there exist degraded quantum wiretap channels $\mathcal{N}_1$ and $\mathcal{N}_2$ such that

$$P(\mathcal{N}_1 \otimes \mathcal{N}_2) > P(\mathcal{N}_1) + P(\mathcal{N}_2).$$

Proof: This result follows by exploiting the counterexample of Hastings [54] for the Holevo information formula. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the channels from Hastings's counterexample, i.e., they satisfy

$$\chi(\mathcal{M}_1 \otimes \mathcal{M}_2) > \chi(\mathcal{M}_1) + \chi(\mathcal{M}_2), \tag{A3}$$

where $\chi(\mathcal{N})$ is the Holevo information of a channel $\mathcal{N}$. We then construct our quantum wiretap channels $\mathcal{N}_1$ and $\mathcal{N}_2$ as $\mathcal{N}_1(\rho) = \mathcal{M}_1(\rho) \otimes \sigma_E$ and $\mathcal{N}_2(\rho) = \mathcal{M}_2(\rho) \otimes \sigma_E$. Both channels are obviously degraded wiretap channels because the channel to the environment simply prepares a constant state σ_E. Also, there is no dependence of the environment's output on the input state, so that the private information of these channels reduces to Holevo information:

$$P(\mathcal{N}_1 \otimes \mathcal{N}_2) = \chi(\mathcal{M}_1 \otimes \mathcal{M}_2),$$
$$P(\mathcal{N}_1) = \chi(\mathcal{M}_1),$$
$$P(\mathcal{N}_2) = \chi(\mathcal{M}_2).$$

Thus, the inequality in the statement of the theorem follows from Eq. (A3)). □

Theorem 35.—The private information of a degraded quantum wiretap channel is additive when restricted to product state encodings.

Proof: First, consider that we can always restrict the optimization in the private information formula to be taken over pure input states whenever the quantum wiretap channel $\mathcal{N}_{A\to BE}$ is degraded. Indeed, consider the extension state

$$\rho_{XYBE} \equiv \sum_{x,y} p_X(x)p_{Y|X}(y|x)|x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \otimes \mathcal{N}_{A\to BE}(\psi_{x,y}), \tag{A4}$$

where we are using a spectral decomposition for each state ρ_x:

$$\rho_x = \sum_y p_{Y|X}(y|x)\psi_{x,y}.$$

Thus, the state in Eq. (A2)) is a reduction of the state in Eq. (A4). Now, consider that

$$\begin{aligned} I(X;B)_\rho - I(X;E)_\rho &= I(XY;B) - I(XY;E) - [I(Y;B|X) - I(Y;E|X)] \\ &\le I(XY;B) - I(XY;E) \\ &\le P(\mathcal{N}_{A\to BE}). \end{aligned}$$

The first equality is from the chain rule for mutual information. The first inequality follows by exploiting the degrading condition and from the fact that X is classical. The final inequality follows by considering XY as a joint classical system, so that the private information of the channel can only be larger than $I(XY;B) - I(XY;E)$.

Now, consider an isometric extension $U_{A\to BEF}$ of a quantum wiretap channel $\mathcal{N}_{A\to BE}$. By using the fact that the private information is optimized for pure state ensembles, we can always rewrite it as

$$\begin{aligned} I(X;B) - I(X;E) &= H(B) - H(E) - H(B|X) + H(E|X) \\ &= H(B) - H(E) - H(B|X) + H(BF|X) \\ &= H(B) - H(E) + H(F|BX), \end{aligned} \tag{A5}$$

where in the second line we used the fact that $H(E|X) = H(BF|X)$ for pure-state ensembles.

Now we show the additivity property for product-state ensembles. Consider the following state on which we evaluate information quantities:

$$\sigma_{XB_1E_1F_1B_2E_2F_2} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes U_{A_1\to B_1E_1F_1}(\varphi_x) \otimes U_{A_2\to B_2E_2F_2}(\psi_x),$$

where we are restricting the signaling states to be product, and without loss of generality, we can take them to be pure, as shown above. Consider that

$$
\begin{aligned}
&I(X;B_1B_2)_\sigma - I(X;E_1E_2)_\sigma \\
&\quad = H(B_1B_2)_\sigma - H(E_1E_2)_\sigma + H(F_1F_2|B_1B_2X) \\
&\quad = H(B_1)_\sigma + H(B_2)_\sigma - H(E_1)_\sigma - H(E_2)_\sigma - [I(B_1;B_2)_\sigma - I(E_1;E_2)_\sigma] + H(F_1F_2|B_1B_2X) \\
&\quad \leq H(B_1)_\sigma + H(B_2)_\sigma - H(E_1)_\sigma - H(E_2)_\sigma + H(F_1|B_1X) + H(F_2|B_2X) \\
&\quad = [I(X;B_1) - I(X;E_1)] + [I(X;B_2) - I(X;E_2)].
\end{aligned}
$$

The first equality follows from the identity in Eq. (A5). The second equality follows from entropy identities. The first inequality follows from the degraded wiretap channel assumption, so that $I(B_1;B_2)_\sigma - I(E_1;E_2)_\sigma \geq 0$ and by applying strong subadditivity of entropy [56] 3 times to get that $H(F_1F_2|B_1B_2X) \leq H(F_1|B_1X) + H(F_2|B_2X)$. The last equality follows from the identity in Eq. (A5) and the fact that we are restricting to product-state signaling ensembles. □

[1] A. Bruen and M. A. Forcinito, *Cryptography, Information Theory, and Error-Correction: A Handbook for the 21st Century* (Wiley-Interscience, New York, 2004).

[2] C. Shannon, *Communication Theory of Secrecy Systems*, Bell Syst. Tech. J. **28**, 656 (1949).

[3] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, *Locking Classical Correlations in Quantum States*, Phys. Rev. Lett. **92**, 067902 (2004).

[4] R. König, R. Renner, A. Bariska, and U. Maurer, *Small Accessible Quantum Information Does Not Imply Security*, Phys. Rev. Lett. **98**, 140502 (2007).

[5] J. A. Smolin and J. Oppenheim, *Locking Information in Black Holes*, Phys. Rev. Lett. **96**, 081302 (2006).

[6] F. Dupuis, J. Florjanczyk, P. Hayden, and D. Leung, Locking Classical Information, arXiv:1011.1612.

[7] S. Lloyd, Quantum Engima Machines, arXiv:1307.0380.

[8] P. Hayden, D. Leung, P. W. Shor, and A. Winter, *Randomizing Quantum States: Constructions and Applications*, Commun. Math. Phys. **250**, 371 (2004).

[9] O. Fawzi, P. Hayden, and P. Sen, *From Low-Distortion Norm Embeddings to Explicit Uncertainty Relations and Efficient Information Locking*, J. ACM **60**, 44 (2013).

[10] A. Wyner, *The Wire-Tap Channel*, Bell Syst. Tech. J. **54**, 1355 (1975).

[11] B. Schumacher, in *Complexity, Entropy, and the Physics of Information*, Santa Fe Institute Studies in the Sciences of Complexity Vol. VIII (Addison-Wesley, Redwood City, CA, 1990), pp. 29–37.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 2006).

[13] H. Buhrman, M. Christandl, P. Hayden, H.-K. Lo, and S. Wehner, *Possibility, Impossibility, and Cheat Sensitivity of Quantum-Bit String Commitment*, Phys. Rev. A **78**, 022316 (2008).

[14] C. H. Bennett and G. Brassard, Quantum Cryptography: Public Key Distribution and Coin Tossing, in Proceedings of the IEEE International Conference on Computer Systems and Signal Processing, Bangalore, India 1984 (IEEE, New York, 1984), pp. 175–179.

[15] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, *The Security of Practical Quantum Key Distribution*, Rev. Mod. Phys. **81**, 1301 (2009).

[16] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, *Classical Capacity of the Lossy Bosonic Channel: The Exact Solution*, Phys. Rev. Lett. **92**, 027902 (2004).

[17] O. Fawzi, Ph.D. thesis, McGill University, 2012, arXiv:1208.5918.

[18] M. Fannes, *A Continuity Property of the Entropy Density for Spin Lattices*, Commun. Math. Phys. **31**, 291 (1973).

[19] K. M. R. Audenaert, *A Sharp Continuity Estimate for the von Neumann Entropy*, J. Phys. A **40**, 8127 (2007).

[20] B. Schumacher, *Sending Entanglement through Noisy Quantum Channels*, Phys. Rev. A **54**, 2614 (1996).

[21] B. Schumacher and M. A. Nielsen, *Quantum Data Processing and Error Correction*, Phys. Rev. A **54**, 2629 (1996).

[22] H. Barnum, M. A. Nielsen, and B. Schumacher, *Information Transmission through a Noisy Quantum Channel*, Phys. Rev. A **57**, 4153 (1998).

[23] H. Barnum, E. Knill, and M. A. Nielsen, *On Quantum Fidelities and Channel Capacities*, IEEE Trans. Inf. Theory **46**, 1317 (2000).

[24] S. Lloyd, *Capacity of the Noisy Quantum Channel*, Phys. Rev. A **55**, 1613 (1997).

[25] P. W. Shor, The Quantum Channel Capacity and Coherent Information, in Proceedings of the MSRI Workshop on Quantum Computation, 2002.

[26] I. Devetak, *The Private Classical Capacity and Quantum Capacity of a Quantum Channel*, IEEE Trans. Inf. Theory **51**, 44 (2005).

[27] N. Cai, A. Winter, and R. W. Yeung, *Quantum Privacy and Quantum Wiretap Channels*, Probl. Inf. Transm. **40**, 318 (2004).

[28] A. S. Holevo, *The Capacity of the Quantum Channel with General Signal States*, IEEE Trans. Inf. Theory **44**, 269 (1998).

[29] B. Schumacher and M. D. Westmoreland, *Sending Classical Information via Noisy Quantum Channels,* Phys. Rev. A **56**, 131 (1997).

[30] D. Leung, J. Lim, and P. Shor, *Capacity of Quantum Erasure Channel Assisted by Backwards Classical Communication*, Phys. Rev. Lett. **103**, 240505 (2009).

[31] G. Smith and J. A. Smolin, *Extensive Nonadditivity of Privacy*, Phys. Rev. Lett. **103**, 120503 (2009).
[32] R. Alicki and M. Fannes, *Continuity of Quantum Conditional Information*, J. Phys. A **37**, L55 (2004).
[33] M. Horodecki, P. W. Shor, and M. B. Ruskai, *Entanglement Breaking Channels*, Rev. Math. Phys. **15**, 629 (2003).
[34] S. Boixo, L. Aolita, D. Cavalcanti, K. Modi, and A. Winter, *Quantum Locking of Classical Correlations and Quantum Discord of Classical-Quantum States*, Int. J. Quantum. Inform. **09**, 1643 (2011).
[35] A. S. Holevo, *Entanglement-Breaking Channels in Infinite Dimensions*, Probl. Inf. Transm. **44**, 171 (2008).
[36] M. Grassl, Z. Ji, Z. Wei, and B. Zeng, *Quantum-Capacity-Approaching Codes for the Detected-Jump Channel*, Phys. Rev. A **82**, 062324 (2010).
[37] C. King, An Application of the Lieb-Thirring Inequality in Quantum Information Theory, arXiv:quant-ph/0412046.
[38] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, *Properties of Conjugate Channels with Applications to Additivity and Multiplicativity*, Markov Proc. Relat. Fields **13**, 391 (2007).
[39] K. Brádler, P. Hayden, D. Touchette, and M. M. Wilde, *Trade-off Capacities of the Quantum Hadamard Channels*, Phys. Rev. A **81**, 062312 (2010).
[40] G. Smith, *Private Classical Capacity with a Symmetric Side Channel and Its Application to Quantum Cryptography*, Phys. Rev. A **78**, 022306 (2008).
[41] H. Ollivier and W. H. Zurek, *Quantum Discord: A Measure of the Quantumness of Correlations*, Phys. Rev. Lett. **88**, 017901 (2001).
[42] L. Henderson and V. Vedral, *Classical, Quantum and Total Correlations*, J. Phys. A **34**, 6899 (2001).
[43] L. Wang and R. Renner, *One-Shot Classical-Quantum Capacity and Hypothesis Testing*, Phys. Rev. Lett. **108**, 200501 (2012).
[44] L. Gurvits and H. Barnum, *Largest Separable Balls around the Maximally Mixed Bipartite Quantum State*, Phys. Rev. A **66**, 062311 (2002).
[45] A. Wehrl, *General Properties of Entropy*, Rev. Mod. Phys. **50**, 221 (1978).
[46] A. S. Holevo, M. Sohma, and O. Hirota, *Capacity of Quantum Gaussian Channels*, Phys. Rev. A **59**, 1820 (1999).
[47] M. M. Wolf, D. Perez-Garcia, and G. Giedke, *Quantum Capacities of Bosonic Channels*, Phys. Rev. Lett. **98**, 130501 (2007).
[48] M. M. Wilde, P. Hayden, and S. Guha, *Quantum Trade-off Coding for Bosonic Communication*, Phys. Rev. A **86**, 062306 (2012).
[49] H. P. Yuen, KCQ: A New Approach to Quantum Cryptography I. General Principles and Key Generation, arXiv: quant-ph/0311061.
[50] G. Smith and J. Yard, *Quantum Communication with Zero-Capacity Channels*, Science **321**, 1812 (2008).
[51] K. Li, A. Winter, X. Zou, and G.-C. Guo, *Private Capacity of Quantum Channels Is Not Additive*, Phys. Rev. Lett. **103**, 120501 (2009).
[52] A. Winter, *Identification via Quantum Channels*, Lect. Notes Comput. Sci. **7777**, 217 (2013).
[53] P. Hayden and A. Winter, *Weak Decoupling Duality and Quantum Identification*, IEEE Trans. Inf. Theory **58**, 4914 (2012).
[54] M. B. Hastings, *Superadditivity of Communication Capacity Using Entangled Inputs*, Nat. Phys. **5**, 255 (2009).
[55] F. Caruso and V. Giovannetti, *Degradability of Bosonic Gaussian Channels*, Phys. Rev. A **74**, 062307 (2006).
[56] E. H. Lieb and M. B. Ruskai, *Proof of the Strong Subadditivity of Quantum-Mechanical Entropy*, J. Math. Phys. (N.Y.) **14**, 1938 (1973).