Practical Route to Entanglement-Assisted Communication Over Noisy Bosonic Channels

Haowei Shi,[1] Zheshen Zhang,[1,2] and Quntao Zhuang[1,3,*]

[1] *James C. Wyant College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA*
[2] *Department of Materials Science and Engineering, University of Arizona, Tucson, Arizona 85721, USA*
[3] *Department of Electrical and Computer Engineering, University of Arizona, Tucson, Arizona 85721, USA*

(Received 4 October 2019; revised manuscript received 16 December 2019; accepted 7 February 2020; published 11 March 2020)

Entanglement offers substantial advantages in quantum information processing, but loss and noise hinder its application in practical scenarios. Although it has been well known for decades that the classical communication capacity of lossy and noisy bosonic channels can be significantly enhanced by entanglement, no practical encoding and decoding schemes are available to realize any entanglement-enabled advantage. Here, we report structured encoding and decoding schemes for such an entanglement-assisted communication scenario. Specifically, we show that phase encoding on an entangled two-mode squeezed vacuum state saturates the entanglement-assisted classical communication capacity of a very noisy channel and overcomes the fundamental limit on covert communication that exists without the assistance of entanglement. We then construct receivers for optimum hypothesis-testing protocols with discrete phase modulation and for optimum noisy phase-estimation protocols with continuous phase modulation. Our results pave the way for entanglement-assisted communication and sensing in the radio-frequency and microwave spectral ranges.

DOI: 10.1103/PhysRevApplied.13.034029

I. INTRODUCTION

The benefit of entanglement for quantum information processing has been revealed by pioneering work in communication [1], sensing [2–4], and computation [5]. Notably, the advantage enabled by the initial entanglement even survives loss and noise in certain entanglement-breaking scenarios, as predicted [6–9] and experimentally demonstrated [10–12], in the entanglement-enhanced sensing protocol called quantum illumination.

It is also known, in theory, that preshared entanglement increases the classical communication capacity, i.e., the maximum rate of reliable communication of classical bits (cbits), of a quantum channel Φ (a completely positive trace-preserving map). In the ideal case, the superdense-coding protocol [13] allows the sending of two cbits on a single qubit, with the assistance of one entanglement bit (ebit). Formally, one characterizes the rate limit of such entanglement-assisted (EA) communication by the classical capacity with unlimited entanglement assistance [1,14–16], $C_E(\Phi)$ [17]. Compared with the classical capacity without entanglement assistance, i.e., the Holevo-Schumacher-Westmoreland capacity, $C(\Phi)$ [18–20], the improvement enabled by entanglement can be drastic even over a noisy channel Φ. In particular, it is known [1] that the ratio $C_E(\Phi)/C(\Phi)$ can diverge logarithmically with the inverse of the signal power over a noisy and lossy bosonic channel [21]. Such an EA scenario is widely applicable to radio-frequency (rf) communication, deep-space communication [22], and covert communication [23,24].

Despite the large advantage of the EA capacity, a practical EA encoding and decoding scheme that achieves *any* advantage over the classical capacity is unknown in the high-noise regime. Previous experiments [25,26] have focused on ideal scenarios with qubits. Although the EA-capacity formula for a bosonic Gaussian channel is well established [27,28], the achievability proof in Ref. [1] relies on approximating an infinite-dimensional channel as a channel with a finite but large dimension; thus a structured encoding scheme is not given for bosonic channels. In fact, simple schemes such as continuous-variable superdense coding [29–31] do not beat the classical capacity in the noisy and weak-signal regime [32], making experimental demonstrations of the EA capacity advantage elusive [33–35]. More recent encoding protocols in Refs. [36–39] use mode permutations or mode selections to encode classical information. Despite being convenient for theoretical analysis, these protocols require large quantum memories to store all quantum states and are thus difficult to implement with available technology.

The main contributions of this paper are (1) discovery of the optimum encoding scheme, (2) showing the advantage

[*] zhuangquntao@gmail.com

of EA communication in the presence of lossy entanglement distribution, and (3) the construction of practical quantum receivers for EA classical communication over lossy and noisy bosonic channels. We first prove that phase encoding on a two-mode squeezed vacuum (TMSV) is asymptotically optimal as the channel noise increases (Sec. IV A). With phase encoding, we also show that the EA advantage can still be appreciable when the entanglement distribution is lossy, thereby further reinforcing the robustness of EA communication (Sec. IV B). Next, we show that such an EA communication protocol is secure and allows one to break the square-root law of covert communication [24] by a logarithmic factor (Sec. IV C). Then, we propose practical quantum receivers, based on prior results in Refs. [8,40], that offer a constant advantage over the classical capacity $C(\Phi)$ in the weak-signal-power regime (Sec. V A). As a by-product, we show that our design in the context of continuous encoding also enables optimal phase estimation and asymptotically saturates the quantum Fisher information (QFI) upper bound [41] (Sec. V B), as the noise increases. Finally, we project the performance of a proof-of-concept experiment, based on the parameters reported in Ref. [11] (Sec. VI).

We begin our paper with a brief overview.

II. OVERVIEW

In most of our discussions, we assume that an entangled signal-idler pair is preshared before a communication, potentially through a ground-satellite and/or fiber-based quantum network. The focus of this paper is on EA communication protocols assuming that preshared entanglement is available, with the understanding that building a full-scale quantum network is a challenging task. For a less than ideal situation, we show that an EA advantage remains (see Sec. IV B) in the absence of a full-scale quantum network that completely overcomes entanglement distribution loss.

Figure 1 shows a general picture of EA communication. Bob encodes classical information on the signal mode of the preshared entanglement retrieved from a quantum memory. A quantum transducer is then employed to convert the wavelength of the signal mode to that of the information carrier, e.g., an rf field. The wavelength-converted signal is then sent to Alice through a lossy and noisy channel. After transducing the signal received from Bob, Alice jointly measures the signal mode and the entangled idler mode retrieved from another quantum memory to decode the classical information. We show that this EA communication scheme outperforms even the best classical scheme without entanglement assistance, and provides a significant advantage especially over a lossy and noisy communication channel in the weak-signal regime. Notably, such an EA communication advantage can be achieved with practical sources and receivers.

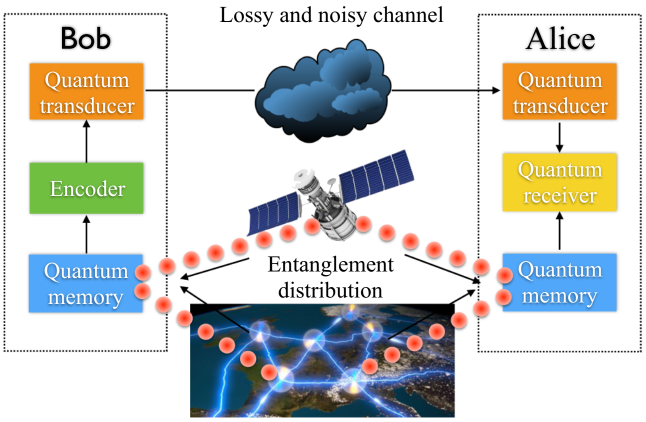

FIG. 1. Concept of entanglement-assisted communication. The entanglement is preshared through a quantum internet based on satellites and/or ground links.

Before going into technical details, we elucidate the applicable scenarios for EA communication. Noisy communication channels are commonly used in the rf domain due to black-body radiation. In optical communication, although ambient noise is not naturally present, the classical communication traffic in, e.g., the optical fibers of the internet may still be regarded as effective noise. In addition, a communication channel can also be noisy in adversarial scenarios with active jamming. There also exist multiple communication settings limited to weak signal power. For example, in deep-space communication, devices deployed on satellites or deep-space stations are likely to be power constrained. Moreover, in a scenario where Alice and Bob wish to stay undetected, known as covert communication, the signal power is minimized and embedded in a bright noise background (see Sec. IV C for details).

III. LOSSY AND NOISY BOSONIC CHANNELS: A COMPENDIUM

Communications typically involve transmitting electromagnetic waves carrying classical information through optical fibers or free space, both of which can be modeled as a bosonic thermal lossy channel $\mathcal{L}^{\kappa,N_B}$ with the following relation between the input and output modes in the Heisenberg picture:

$$\hat{a}_R = \sqrt{\kappa}\hat{a}_S + \sqrt{1-\kappa}\hat{a}_B, \tag{1}$$

as illustrated in Fig. 2. Here, the input mode is subject to an average-energy constraint $\left\langle \hat{a}_S^\dagger \hat{a}_S \right\rangle = N_S$, and the noise mode $\hat{a}_B$ is in a thermal state with mean photon number $N_B/(1-\kappa)$, where κ is the transmissivity of the bosonic channel.

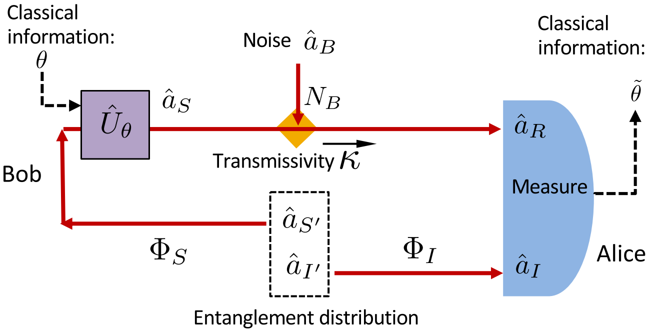

FIG. 2. Schematic illustration of the entanglement-assisted classical communication protocol. The preshared entanglement is distributed through two channels Φ_S and Φ_I. Classical information θ is encoded on the signal $\hat{a}_S$, which is sent over a noisy channel, represented by a beam splitter with transmissivity κ and noise N_B, and then jointly measured with the entangled idler $\hat{a}_I$ to decode the classical information $\tilde{\theta}$.

Without entanglement assistance, the classical capacity is known to be [42]

$$C(\mathcal{L}^{\kappa,N_B}) = g(\kappa N_S + N_B) - g(N_B), \tag{2}$$

obtained by maximizing the Holevo information [19,20,43] over the ensemble of states. Here, $g(n) = (n+1)\log_2(n+1) - n\log_2 n$ is the entropy of a thermal state with mean photon number n. The capacity is achieved by an ensemble of Gaussian-modulated coherent states in conjunction with a joint-detection receiver, which is in general difficult to build. In some special situations, however, practical receivers are known to achieve the classical capacity [22,44]. For example, in the limit of $\kappa N_S \gg 1$ and $N_B \ll 1$, an optical heterodyne receiver approaches the classical capacity. Moreover, in the large-noise case of $N_B \gg 1$, the classical capacity satisfies $C(\mathcal{L}^{\kappa,N_B}) = \kappa N_S/\ln(2)N_B + O(1/N_B^2)$ and is always saturated by a heterodyne or homodyne receiver (see Appendix A for details).

Classical communication can be enhanced by preshared entanglement. Bosonic EA classical communication operates in the following way (see Fig. 2 for an example). One starts with entangled signal-idler pairs $\hat{a}_{S'}, \hat{a}_{I'}$, which are delivered to the sender Bob and the receiver Alice through channels Φ_S and Φ_I. In this section, both of the channels Φ_S and Φ_I are assumed lossless and noiseless, i.e., perfect unlimited preshared entanglement can be shared by Alice and Bob. Entanglement preshared through a common lossy channel is considered in Sec. IV B. The encoded signal $\hat{a}_S$, with mean photon number N_S, is sent through the noisy channel. A joint measurement on the received signal-idler pairs $\hat{a}_R, \hat{a}_I$ is performed to decode information. The EA classical capacity is [1]

$$C_E(\mathcal{L}^{\kappa,N_B}) = g(N_S) + g(N_S') - g(A_+) - g(A_-), \tag{3}$$

where $A_\pm = [D - 1 \pm (N_S' - N_S)]/2$, $N_S' = \kappa N_S + N_B$, and $D = \sqrt{(N_S + N_S' + 1)^2 - 4\kappa N_S(N_S+1)}$. Various aspects of EA communication have been explored, including extensions to limited pure entanglement [45], noisy entanglement [46], trade-off capacities [36,47], and superaddivity issues [48,49].

Comparing the capacity formulas with and without entanglement assistance, one has

$$\lim_{N_B\to\infty} C_E/C = (1+N_S)\ln(1+1/N_S), \tag{4}$$

which diverges as $\ln(1/N_S)$ (see Fig. 3). Thus, in the weak-signal and strong-noise regime, entanglement assistance can offer a large capacity advantage over unassisted classical communication. Moreover, it is known that encoding on a TMSV,

$$|\psi^{N_S}\rangle_{S'I'} = \sum_{n=0}^{\infty} \sqrt{N_S^n/(N_S+1)^{n+1}}\, |n\rangle_{S'}\, |n\rangle_{I'}, \tag{5}$$

achieves the EA classical capacity over a bosonic thermal lossy channel [36–38], but the previously proposed encoding needs either large quantum memories or non-Gaussian operations without structured realizations, both of which are beyond the reach of current technology. Also, there is no known structured receiver that achieves the EA classical capacity.

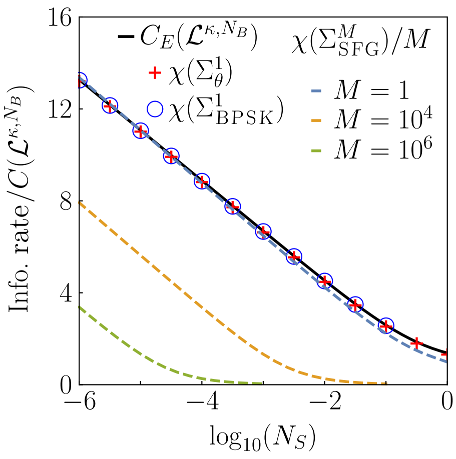

FIG. 3. Information (Info.) rate divided by the unassisted classical capacity $C(\mathcal{L}^{\kappa,N_B})$ vs the transmitted power N_S. The channel transmissivity is $\kappa = 0.1$ and the noise is $N_B = 10$. The entanglement-assisted classical capacity $C_E(\mathcal{L}^{\kappa,N_B})$ (black solid line) has a large advantage when the power is low, and the modewise phase encodings Σ_θ^1 (red crosses) and $\Sigma_{\rm BPSK}^1$ (blue circles) have Holevo information $\chi\left(\Sigma_\theta^1\right)$ and $\chi\left(\Sigma_{\rm BPSK}^1\right)$, achieving the entanglement-assisted classical capacity (black solid line). (BPSK, binary phase-shift keying.) For phase modulation in M-mode blocks, the sum-frequency-generation (SFG) process gives an estimate of the Holevo information per mode $\chi(\Sigma_{\rm SFG}^M)/M$ for various repetition-encoding block sizes M (dashed lines).

It is worth mentioning that both the coherent states for classical communication and the TMSV for EA classical communication belong to the class of Gaussian states [50], whose Wigner functions have a Gaussian shape. Gaussian states are important for quantum information processing, because they enable nonclassical resources such as squeezing and entanglement and, moreover, they often allow analytical solutions of various problems. An n-mode Gaussian state $\hat{\rho}$ comprising modes $\hat{a}_k$, $1 \leq k \leq n$, is fully characterized by the mean and the covariances of real quadrature field operators $\hat{q}_k = \hat{a}_k + \hat{a}_k^\dagger, \hat{p}_k = i\left(\hat{a}_k^\dagger - \hat{a}_k\right)$. Formally, we can define a real $2n$-dimensional vector of operators $\hat{\mathbf{x}} = \left(\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_n, \hat{p}_n\right)$, and then the mean $\bar{\mathbf{x}} = \langle\hat{\mathbf{x}}\rangle_{\hat{\rho}}$ and the elements of the $2n$-by-$2n$ covariance matrix are given by

$$\mathbf{\Lambda}_{ij} = \frac{1}{2}\left\langle\{\hat{x}_i - \bar{x}_i, \hat{x}_j - \bar{x}_j\}\right\rangle_{\hat{\rho}}, \tag{6}$$

where $\{,\}$ is the anticommutator and $\left\langle\hat{A}\right\rangle_{\hat{\rho}} = \mathrm{Tr}\left(\hat{A}\hat{\rho}\right)$. As an example, from the wave function in Eq. (5), we can obtain the covariance matrix of a TMSV as

$$\mathbf{\Lambda}_{\mathrm{TMSV}} = \begin{pmatrix} (2N_S+1)\mathbf{I} & 2C_0\mathbf{Z} \\ 2C_0\mathbf{Z} & (2N_S+1)\mathbf{I} \end{pmatrix}, \tag{7}$$

where $\mathbf{I}$, $\mathbf{Z}$ are 2×2 Pauli matrices, and $C_0 = \sqrt{N_S\left(N_S+1\right)}$ is the amplitude of the phase-sensitive cross-correlation.

IV. OPTIMAL ENCODING—PHASE MODULATION

A. Channel capacity with perfect preshared entanglement

In this section, we show that a set of states produced by phase modulation on an ideal TMSV is an asymptotically optimal encoding scheme, in that it achieves $C_E\left(\mathcal{L}^{\kappa,N_B}\right)$ for $N_B \gg 1$. Mathematically, phase modulation is described by the unitary operator $\hat{U}_\theta = \exp\left(i\theta\hat{a}^\dagger\hat{a}\right)$ [50], which maps $\hat{a} \to e^{i\theta}\hat{a}$, where $\hat{a}$ is the annihilation operator of the incoming field. Under phase encoding (see Fig. 2), the joint state of the returned signal $\hat{a}_R$ and the retained idler $\hat{a}_I$ at the receiver is

$$\hat{\rho}_{RI}^{\theta} \equiv \mathcal{L}^{\kappa,N_B}\left[(\hat{U}_\theta\otimes\hat{I})\hat{\psi}_{S'I'}^{N_S}(\hat{U}_\theta^\dagger\otimes\hat{I})\right], \tag{8}$$

where $\hat{\psi}_{S'I'}^{N_S}$ is the density operator of the TMSV. From the input-output relation in Eq. (1) and the covariance matrix of a TMSV in Eq. (7), we can obtain the covariance matrix of the zero-mean Gaussian state $\hat{\rho}_{RI}^{\theta}$ as

$$\mathbf{\Lambda}_\theta = \begin{pmatrix} (2\left(N_B+\kappa N_S\right)+1)\mathbf{I} & 2C_p\mathbf{R}_\theta \\ 2C_p\mathbf{R}_\theta & (2N_S+1)\mathbf{I} \end{pmatrix}, \tag{9}$$

where $\mathbf{R}_\theta = \mathrm{Re}\left[\exp\left(i\theta\right)\left(\mathbf{Z}-i\mathbf{X}\right)\right]$. The amplitude of the cross-correlation for each mode pair is $C_p = \sqrt{\kappa}C_0$.

Thus, the set of states at the receiver is given by $\Sigma_\theta^1 \equiv \{\hat{\rho}_{RI}^{\theta}, \theta \sim U[0,2\pi)\}$, where the phase θ is uniformly distributed. Under optimal decoding, the accessible information after the channel can be obtained from $\chi\left(\Sigma_\theta^1\right)$, where

$$\chi\left(\{\hat{\rho}_x, p(x)\}\right) = S\left(\int_x p(x)\hat{\rho}_x\right) - \int_x p(x)S\left(\hat{\rho}_x\right) \tag{10}$$

is the Holevo information and $S(\cdot)$ is the von Neumann entropy.

The conditional entropy $S(\hat{\rho}_{RI}^{\theta})$ can be straightforwardly calculated because the state is Gaussian [50]. The calculation of the unconditional entropy is, however, more involved, as detailed in Appendix B; numerical results are shown in Fig. 3 by red crosses. By asymptotic expansion in the limit $N_B \gg 1$, we can show (details are given in Appendix B) that

$$\begin{aligned}\chi\left(\Sigma_\theta^1\right) &= C_E(\mathcal{L}^{\kappa,N_B}) + O(1/N_B^2) \\ &= \frac{1}{N_B}\kappa N_S(1+N_S)\log_2\left(1+\frac{1}{N_S}\right) + O\left(\frac{1}{N_B^2}\right).\end{aligned} \tag{11}$$

Because phase encoding achieves the EA capacity $C_E(\mathcal{L}^{\kappa,N_B})$, it is the optimal encoding over a lossy and noisy bosonic channel in the asymptotic limit of $N_B \gg 1$.

While continuous phase encoding is asymptotically optimal, encoding with a set of discrete phases is more practical in real-world operations. As an example, Sec. V A demonstrates binary phase-shift keying as a handy implementation that overcomes the limit on the classical capacity. In BPSK, the ensemble of the quantum states at the receiver is $\Sigma_{\mathrm{BPSK}}^1 = \{\hat{\rho}_{RI}^{\theta}, \theta \sim U\{0,\pi\}$. Similarly, the Holevo information for Σ_{BPSK}^1 is calculated and is depicted in Fig. 3, showing the asymptotic optimality of BPSK encoding in the large-noise limit.

In many protocols, such as quantum illumination and floodlight quantum key distribution [51], repetition coding of the same θ on M signal-idler mode pairs, i.e., $\Sigma_\theta^M \equiv \{\otimes_{k=1}^M\hat{\rho}_{R_kI_k}^{\theta}, \theta \sim U[0,2\pi)\}$, is used to obtain sufficiently large mutual information per encoding so that efficient error correction codes can be employed. Let M mode pairs be a phase modulation block. The derivation of the Holevo information per mode, $\chi\left(\Sigma_\theta^M\right)/M$, is computationally challenging when $M \gg 1$. However, we obtain a precise estimate of the information per mode $\chi\left(\Sigma_{\mathrm{SFG}}^M\right)/M$

based on the results in Sec. VII, where a SFG process [8] on the modes within each phase-modulation block is devised. Figure 3 shows good agreement between the estimate $\chi(\Sigma_{\rm SFG}^{M})/M$ (blue dashed line) and the exact result $\chi(\Sigma_{\theta}^{1})$ (red crosses) for $M=1$.

B. Channel capacity with imperfect preshared entanglement distributed via a lossy channel

The previous analysis assumes perfect preshared entanglement distributed through lossless and noiseless channels Φ_S and Φ_I. It shows that phase encoding on perfect preshared TMSV states leads to a $\ln(1/N_S)$ capacity advantage in the weak-signal and high-noise regime. Although perfect preshared entanglement could be constructed in a full-scale quantum network in the future, current technology allows only nonideal distribution of entanglement. As such, imperfections in Φ_S and Φ_I need to be accounted for in a practical scenario.

Suppose the entangled TMSV signal-idler pairs are generated by Alice, and so the idler distribution channel remains perfect, i.e., $\Phi_I = I$, while the signal is distributed to Bob through a noiseless lossy channel Φ_S, i.e., a pure-loss bosonic channel $\mathcal{L}^{\kappa_0,0}$. Since the phase encoding $\hat{U}_\theta$ commutes with $\mathcal{L}^{\kappa_0,0}$, as can be verified by using the input-output relation in Eq. (1) of a general thermal loss channel $\mathcal{L}^{\kappa,N_B}$ and the fact that a thermal state is invariant under phase rotation, we may consider an equivalent protocol in which the initial signal modes $\hat{a}_{S'}$ are first phase encoded, subsequently go through the channels $\Phi_S = \mathcal{L}^{\kappa_0,0}$ and $\mathcal{L}^{\kappa,N_B}$ consecutively, and are finally received by Alice. In the equivalent protocol, the overall noisy channel is

$$\Phi_{\rm All} = \mathcal{L}^{\kappa,N_B} \circ \mathcal{L}^{\kappa_0,0} = \mathcal{L}^{\kappa_0\kappa,N_B}. \tag{12}$$

To match the mean photon number going through the channel $\mathcal{L}^{\kappa,N_B}$ with the classical case, the mean photon number $N_{S'}$ of $\hat{a}_{S'}$ is constrained by $\kappa_0 N_{S'} = N_S$. Because the overall channel is again a lossy thermal channel, in the $\kappa \ll 1, N_B \gg 1$ limit, the accessible information at Alice's receiver reads

$$\begin{aligned}\chi\left(\Sigma^1_{\theta,\kappa_0}\right) &= C_E(\mathcal{L}^{\kappa_0\kappa,N_B}, N_S/\kappa_0) + O(1/N_B^2)\\ &= \frac{1}{N_B}\kappa N_S\left(1+\frac{N_S}{\kappa_0}\right)\log_2\left(1+\frac{\kappa_0}{N_S}\right)\\ &\quad + O\left(\frac{1}{N_B^2}\right),\end{aligned} \tag{13}$$

where the input power N_S/κ_0 is made explicit. If $N_S/\kappa_0 \ll 1$, the accessible information $\chi\left(\Sigma^1_{\theta,\kappa_0}\right)$ remains a factor of $\ln(\kappa_0/N_S)$ larger than the unassisted capacity $C(\mathcal{L}^{\kappa,N_B}, N_S)$.

Note that prior to EA communication, preshared entanglement always needs to be constructed via an entanglement-distribution channel. It thus behooves us to consider using the pure-loss entanglement distribution channel for classical communication without entanglement assistance, which is anticipated to outperform EA communication over a lossy and noisy channel. However, various scenarios can preclude direct utilization of the entanglement-distribution channel for classical communication; we enumerate two examples in the following. First, the performance of rf communications can be improved by preshared entanglement distributed via optical links. In this case, the rf communication link is well modeled by a lossy thermal channel, while the optical entanglement-distribution link is a pure-loss channel. Second, a pure-loss channel for entanglement establishment may not always be available. With long-lived quantum memories, entanglement can be distributed in the presence of an entanglement-distribution channel and be subsequently stored with high fidelity until it is retrieved on demand for EA communication.

C. Entanglement-assisted covert communication

An additional benefit of the EA communication protocol is its covertness and security [23,24]. Covert communication refers to the scenario in which two parties are able to communicate, while the presence of the communication signals cannot be easily detected by any passive adversary. This is possible when unavoidable environmental noise hides weak signals. In the EA communication protocol, suppose that a passive adversary endeavors to detect Alice and Bob's communication attempt by monitoring the mode lost to the environment, but does not have access to the idler $\hat{a}_I$, since the entanglement is preshared prior to communication. In the presence of EA communication with N mode pairs, the reduced state of the modes lost to the environment is a product of N thermal states $\hat{\rho}_1$, each with mean photon number $n_1 = \kappa N_B/(1-\kappa) + (1-\kappa)N_S \simeq \kappa N_B + N_S$, irrespective of the message being transmitted. In the absence of communication, the state $\hat{\rho}_0$ remains thermal, with a mean photon number $n_0 = \kappa N_B/(1-\kappa) \simeq \kappa N_B$. If $\kappa N_B \gg 1$, the difference between $\hat{\rho}_0$ and $\hat{\rho}_1$ is so small that communication covertness can be guaranteed. The Helstrom bound on the probability of an error in distinguishing $\hat{\rho}_0^{\otimes N}$ from $\hat{\rho}_1^{\otimes N}$ can be numerically calculated, as both states are diagonal in the number basis. Here, we use the quantum Chernoff bound [52,53] to estimate the error probability of the adversary,

$$P_E \sim \exp\left[-NN_S^2/(8\kappa^2 N_B^2)\right]. \tag{14}$$

Under the condition $P_E \sim 1/2$, we can still communicate with $N \sim \kappa^2 N_B^2/N_S^2$ modes, which is large when $\kappa N_B \gg 1$.

A more involved calculation, similar to that in Ref. [24], shows that under the condition $P_E \geq 1/2 - \delta$, the relative entropy satisfies $D(\hat{\rho}_0^{\otimes N} \| \hat{\rho}_1^{\otimes N}) \leq 2\delta^2/\ln(2)$. Using the additivity of relative entropy and the properties of thermal states (or Theorem 7 in Ref. [54]),

$$D(\hat{\rho}_0^{\otimes N}\|\hat{\rho}_1^{\otimes N}) = ND(\hat{\rho}_0\|\hat{\rho}_1)$$
$$= N\left\{\log_2\left[\frac{n_1+1}{n_0+1}\right]\right.$$
$$\left.+n_0\log_2\left[\frac{n_0(n_1+1)}{n_1(n_0+1)}\right]\right\}. \tag{15}$$

Therefore, one has $N \le N_\delta \equiv 4\delta^2\kappa N_B(\kappa N_B+1)/N_S^2 + O(N_S^3)$. With a large κN_B, and based on the capacity formula in Eqs. (2) and (3), we expect the information transmitted in classical communication without entanglement assistance to be $N_\delta C(\mathcal{L}^{\kappa,N_B}) \simeq 4\kappa^3\delta^2 N_B/[N_S\ln(2)] \sim \sqrt{N_\delta}\delta/\kappa^2$, which is often referred to as the square-root law for covert communication [24]. The EA communication, however, allows a factor of $\ln(1/N_S) \sim \ln(N_\delta)$ more bits of information to be transmitted while maintaining the same level of covertness. The $\sqrt{N_\delta}\ln(N_\delta)$ scaling for EA covert communication thus breaks the square-root law for classical covert communication by a logarithmic factor, as illustrated by the example in Fig. 4.

Moreover, because the quantum states accessible to the adversary are identical for any encoded message, the adversary is unable to learn any information about the message. As such, the protocol is secure, as long as the preshared entanglement is perfect and the idler is retained securely in Alice's laboratory. Note that the security is conditioned only on the entanglement being perfect; it does not rely on a specific noise model for the communication channel. Presharing perfect entanglement, however, is challenging in practice. Nonetheless, the quantum internet is a technology under active development, and the challenges associated with entanglement distribution can in principle be overcome by quantum repeaters.

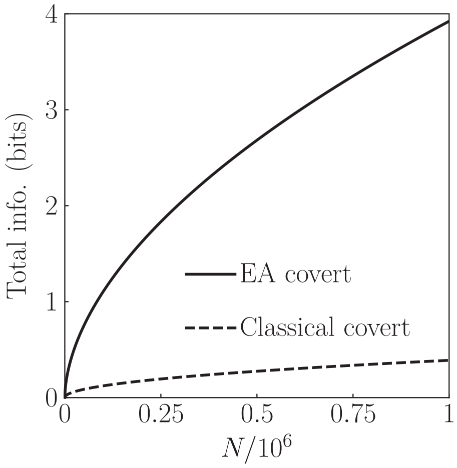

FIG. 4. Ultimate total covert information, in bits transferred in N modes by entanglement-assisted covert communication and classical covert communication without entanglement assistance. Channel transmissivity $\kappa = 0.1$ and noise $N_B = 10$, covertness $\delta = 0.01$.

V. PRACTICAL RECEIVER STRUCTURES

A. Quantum receivers for discrete modulation and optimum hypothesis testing

Section IV A demonstrates the optimality of phase encoding in EA communication without specifying a structured receiver that approaches the channel capacity. In this section, we focus on practical receiver designs. To allow efficient error correction codes, we consider repetition coding with the BPSK-modulated state ensemble $\Sigma_{\rm BPSK}^M = \{\otimes_{k=1}^M \hat{\rho}^\theta_{R_k I_k}, \theta = 0, \pi\}$. Formally, the decoding of BPSK may be viewed as a binary hypothesis-testing task that discriminates between two modulation phases $\theta = 0, \pi$. Such a hypothesis-testing task is similar to that of quantum illumination. It is known that the optical-parametric-amplifier (OPA) receiver and phase-conjugate receiver (PCR) in quantum illumination [40] both offer a 3-dB advantage in the error-probability exponent over classical illumination, while the optimum quantum receiver offers a 6-dB error-probability exponent advantage. The advantage enabled by the OPA receiver has been demonstrated in a quantum-illumination experiment [11]. A more recent publication [8] presented the optimum receiver, based on sum-frequency generation and feedforward (FF), for obtaining the full advantage of quantum illumination over the optimum classical scheme. Let the error probability of the symmetric hypothesis testing be P_E; the per-mode communication rate is given by

$$R_{P_E} = \frac{1}{M}\left[1 + P_E\log_2 P_E + (1-P_E)\log_2(1-P_E)\right]. \tag{16}$$

The per-mode communication rates for the OPA receiver, the PCR, and the FF-SFG receiver in EA communication are evaluated and are plotted in Fig. 5.

1. OPA receiver

We first elaborate on the OPA receiver (Fig. 6). The OPA receiver applies parametric amplification across all returned-signal and retained-idler mode pairs $\{\hat{a}_R^{(m)}, \hat{a}_I^{(m)}\}$ to transform the cross-correlations between the input modes to photon-number differences. The two-mode squeezing in the amplification produces the modes $\hat{c}^{(m)} = \sqrt{G}\hat{a}_I^{(m)} + \sqrt{G-1}\hat{a}_R^{\dagger(m)}$, with mean photon number

$$\overline{N}(\theta) \equiv \left\langle \hat{c}^{\dagger(m)}\hat{c}^{(m)}\right\rangle$$
$$= GN_S + (G-1)(\kappa N_S + N_B + 1)$$
$$+ 2\sqrt{G(G-1)}\cos\theta C_p \tag{17}$$

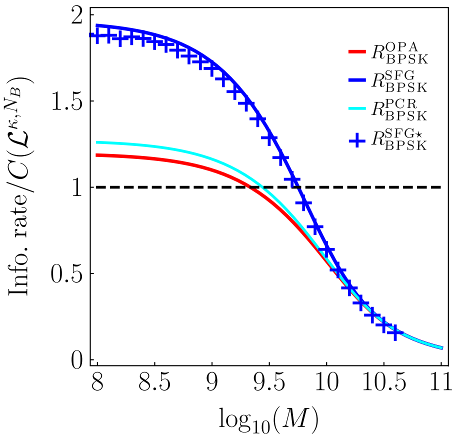

FIG. 5. Entanglement-assisted communication in comparison with classical communication. Information rate per mode normalized by the unassisted classical capacity $C(\mathcal{L}^{\kappa,N_B})$. The unassisted information rate $C(\mathcal{L}^{\kappa,N_B})$ is plotted as a black dashed line as a benchmark. The plots are drawn in blue for a sum-frequency-generation-based receiver, cyan for a phase-conjugate receiver, and red for an optical-parametric-amplifier receiver. The blue crosses show the numerical results of Monte Carlo simulations of an FF-SFG receiver with 50 adaptive measurement cycles and 8×10^5 samples. Parameters: $N_s = 10^{-3}, N_B = 10^4, \kappa = 10^{-3}$.

for an encoded phase θ. The distribution of the total photon number across the M modes can be obtained as

$$P_{\mathrm{OPA}}(n|\theta; M) = \binom{n+M-1}{n}\left(\frac{\overline{N}(\theta)}{1+\overline{N}(\theta)}\right)^n \times \left(\frac{1}{1+\overline{N}(\theta)}\right)^M. \tag{18}$$

Given the conditional probability distribution, we evaluate the performance of maximum-likelihood decision in Appendix C 1, with the error probability shown in Fig. 9 and the communication rate plotted in Fig. 5. An ideal OPA receiver applied to a BPSK-encoded TMSV source (red line) beats the classical capacity by approximately 18.6% at $M = 10^8$ and 10.0% at $M = 10^9$. As the number of modes M in the repetition block increases, the rate per mode decreases as expected. Note that the normalization $C(\mathcal{L}^{\kappa,N_B})$ does not change with M.

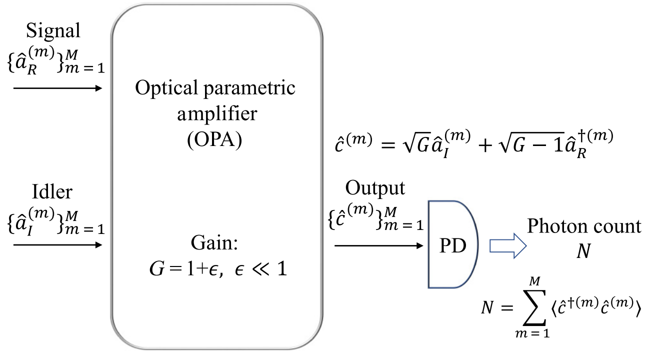
Signal
$\{\hat{a}_R^{(m)}\}_{m=1}^M$
Idler
$\{\hat{a}_I^{(m)}\}_{m=1}^M$
Optical parametric amplifier (OPA)
Gain:
$G = 1{+}\epsilon,\ \epsilon \ll 1$
$\hat{c}^{(m)} = \sqrt{G}\hat{a}_I^{(m)} + \sqrt{G-1}\hat{a}_R^{\dagger(m)}$
Output
$\{\hat{c}^{(m)}\}_{m=1}^M$
PD
Photon count
N
$N = \sum_{m=1}^{M}\langle \hat{c}^{\dagger(m)}\hat{c}^{(m)}\rangle$

FIG. 6. Setup for optical-parametric-amplifier receiver. PD, photodetector. The returned signal $\hat{a}_R$ and idler $\hat{a}_I$ travel through the OPA, ejecting amplified beams at two output ports, with the amplifier gain $G \simeq 1$. We collect photons at the port where the idler is amplified.

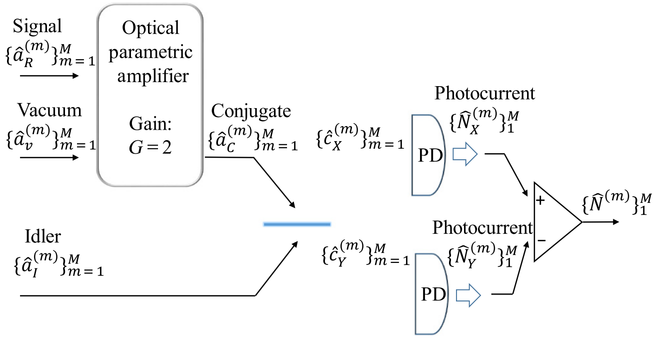
Signal
$\{\hat{a}_R^{(m)}\}_{m=1}^M$
Vacuum
$\{\hat{a}_v^{(m)}\}_{m=1}^M$
Optical parametric amplifier
Gain:
$G = 2$
Conjugate
$\{\hat{a}_C^{(m)}\}_{m=1}^M$
Idler
$\{\hat{a}_I^{(m)}\}_{m=1}^M$
$\{\hat{c}_X^{(m)}\}_{m=1}^M$
$\{\hat{c}_Y^{(m)}\}_{m=1}^M$
PD
PD
Photocurrent
$\{\hat{N}_X^{(m)}\}_1^M$
Photocurrent
$\{\hat{N}_Y^{(m)}\}_1^M$
+
−
$\{\hat{N}^{(m)}\}_1^M$

FIG. 7. Setup for a phase-conjugate receiver. The returned signal $\hat{a}_R$ travels through an OPA with an amplifier gain $G = 2$. The phase-conjugated field appears at the empty port, interfering with the idler $\hat{a}_I$ through a 50:50 beam splitter. The photon counts in the two arms of the interferometer are collected to derive the differential photon counts, based on which the message is decoded.

2. Phase-conjugate receiver

The PCR (Fig. 7), a variant of the OPA receiver, reaches the same asymptotic error exponent for $N_S \ll 1, N_B \gg 1$ but yields a slight advantage for nonzero N_S (see Fig. 5). The PCR conjugates the M input modes $\hat{a}_R^{(m)}$ while amplifying the vacuum $\hat{a}_v^{(m)}$ at the empty port, i.e.,

$$\hat{a}_C^{(m)} = \sqrt{2}\hat{a}_v^{(m)} + \hat{a}_R^{\dagger(m)}. \tag{19}$$

Then, the conjugated signal along with the idler is detected by a balanced difference detector from the photon count $\hat{N}^{(m)} = \hat{N}_X^{(m)} - \hat{N}_Y^{(m)}$, where $\hat{N}_X^{(m)}, \hat{N}_Y^{(m)}$ are the photon counts of the two outputs of the 50:50 beam splitter: $\hat{c}_X^{(m)} = (\hat{a}_C^{(m)} + \hat{a}_I^{(m)})/\sqrt{2}$, $\hat{c}_Y^{(m)} = (\hat{a}_C^{(m)} - \hat{a}_I^{(m)})/\sqrt{2}$. In analogy to the OPA receiver, the decision is made according to the total photon count across the M modes.

A detailed analysis of the communication performance in Appendix C 2 shows that the PCR has a slight edge over the OPA receiver in that its signal-to-noise ratio is better in the higher-order terms. Illustrated by the cyan line in Fig. 5, an ideal PCR with a BPSK-encoded TMSV source exceeds the classical capacity by approximately 26.0% at $M = 10^8$ and 16.3% at $M = 10^9$.

3. FF-SFG receiver

The FF-SFG receiver (Fig. 8) improves on the performance of the OPA receiver and is the optimum for

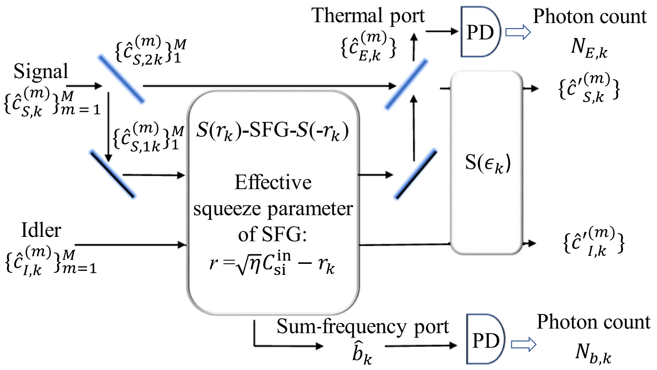
Thermal port
$\{\hat{c}_{E,k}^{(m)}\}$
PD
Photon count
$N_{E,k}$
Signal
$\{\hat{c}_{S,k}^{(m)}\}_{m=1}^{M}$
$\{\hat{c}_{S,2k}^{(m)}\}_1^M$
$\{\hat{c}_{S,1k}^{(m)}\}_1^M$
$S(r_k)$-SFG-$S(-r_k)$
Effective squeeze parameter of SFG:
$r=\sqrt{\eta}C_{\rm si}^{\rm in}-r_k$
$S(\epsilon_k)$
$\{\hat{c}'^{(m)}_{S,k}\}$
$\{\hat{c}'^{(m)}_{I,k}\}$
Idler
$\{\hat{c}_{I,k}^{(m)}\}_{m=1}^{M}$
Sum-frequency port
$\hat{b}_k$
PD
Photon count
$N_{b,k}$

FIG. 8. Setup for a single cycle of the feedforward sum-frequency-generation receiver. The signal in the kth cycle, $\hat{c}_{S,k}$, is first divided into a bright main stream $\hat{c}_{S,2k}$ and a weak slice $\hat{c}_{S,1k}$ by a highly transmissive beam splitter with reflectivity $\sqrt{\eta} \ll 1$. The weak slice goes through a FF-SFG module containing three processes in sequence, $S(r_k)$, SFG, and $S(-r_k)$, with the phase of the squeeze parameter r_k adaptively tuned. Eventually, the processed weak slice $\hat{c}'_{S,1k}$ is merged back into the main stream by a second highly transmissive beam splitter. We collect the photon counts at the sum-frequency port of the SFG module and the thermal port of the second beam splitter.

quantum illumination in the strong-noise and weak-signal limit. Through an SFG process, the FF-SFG receiver converts the cross-correlations between the signal-idler pairs and produces quantum states with photon-number statistics approximating a coherent state. Thus, by analogy with the Dolinar receiver, the optimum receiver for binary coherent-state discrimination, the FF-SFG receiver asymptotically achieves the quantum Chernoff bound for quantum illumination. The principle of the FF-SFG receiver is briefly introduced below (more details are given in the Supplemental Material of Ref. [8]).

The FF-SFG receiver consists of a sequence of multiple cycles of adaptive detection. The measurement results from all previous cycles are combined through a Bayesian strategy that produces a posterior distribution of different hypotheses. In the kth cycle, the prior probabilities $P_0^{(k)}, P_1^{(k)}$ for the hypotheses $\theta_0=0, \theta_1=\pi$ are used to design the measurements, whose results are used to obtain the posteriors (and also the priors of the $(k+1)$th cycle) $P_0^{(k+1)}, P_1^{(k+1)}$ through a Bayesian formula. We denote the maximum-likelihood decision before the cycle as $\tilde{h}=\arg\max_\ell P_\ell^{(k)}$, while the true hypothesis is h. As shown in Fig. 8, the FF-SFG slices off a proportion $\eta \ll 1$ of the strong returned-signal modes $\hat{c}_{S,k}^{(m)}$ to interact with the weak idler modes $\hat{c}_{I,k}^{(m)}$ through a SFG process to produce a sum mode $\hat{b}_k$ for detection. We denote the cross-correlation between $\hat{c}_{S,k}^{(m)}$ and $\hat{c}_{I,k}^{(m)}$ by $C_{{\rm si},k}^{\rm in}$. The interaction consists of (1) two adaptively tuned two-mode squeezing modules $\hat{S}(r_k)$ and $\hat{S}(-r_k)$ that change the cross-correlation, adopting the same feedforward strategy as in the Dolinar receiver, and (2) a SFG process that converts the cross-correlation into a sum-frequency mode $\hat{b}_k$. The sum-frequency mode is approximately in a coherent state $|e^{i\theta_h}\sqrt{M}r\rangle$, with $r=\sqrt{\eta}C_{{\rm si},k}^{\rm in}-r_k$, plus thermal noise $\eta N_S N_B$, which is to be measured by photon counting. This shows an analogy to the Dolinar receiver [55], which, based on the maximum-likelihood decision $\tilde{h}$, chooses an r_k to displace the coherent state to a near-vacuum state, i.e., $r \sim 0$, for optimum state discrimination at the Helstrom limit (more details are given in Appendix C 3).

Subsequently, the sliced signal modes are recombined with the other part of the signal modes, forming an interferometer structure. An M-mode thermal state of $\hat{c}_{E,k}^{(m)}$'s is generated with the same mean photon number $M|r|^2$ at the dim port, which is therefore denoted as the *thermal port*. The total number of photons at the thermal port is measured as well.

Finally, the bright output goes through an additional two-mode squeezing $\hat{S}(\epsilon_k)$ that wipes out the r_k dependence in the evolution of the cross-correlation. The evolution is terminated when the cross-correlation has almost been used up, i.e., when the residual cross-correlation is only a proportion $\epsilon \ll 1$ of the initial cross-correlation.

Similarly to the results for target detection [8], the FF-SFG receiver also demonstrates its optimality for phase discrimination. Monte Carlo simulations of the FF-SFG receiver are performed with various parameters for EA communication as shown in Fig. 9, in which the Helstrom limit under a uniform prior is estimated. From the

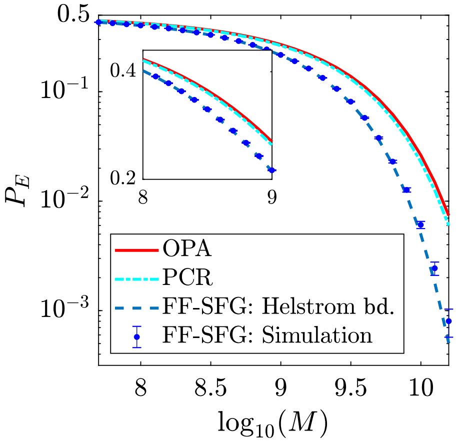

FIG. 9. Error probability of hypothesis testing between two encoded phases for optical-parametric-amplifier receiver (red), phase-conjugate receiver (cyan), and feedforward sum-frequency-generation receiver (dashed line, theoretical bound; dots, numerical simulation with error bars). The dots are from Monte Carlo simulations with 8×10^5 samples, which saturates the Helstrom bound. The inset is an enlargement around an error probability of one half, with the error probability on a linear scale, showing the synchronicity better. Parameters: $N_S=10^{-3}, N_B=10^4, \kappa=10^{-3}, \eta=4\times10^{-6}$.

error probabilities, the communication rate can be evaluated from Eq. (16). As indicated by the blue stars in Fig. 5, a FF-SFG receiver with a BPSK-encoded TMSV source exceeds the classical capacity by an advantage of approximately 90% for $M = 10^8$ and 71% for $M = 10^9$.

B. Quantum receivers for continuous encoding and noisy phase estimation

Although BPSK encoding is handy for practical communications, its capacity is intrinsically bounded at one bit per symbol. This rapidly undermines the advantage of EA communication as the number of modes M in a repetition block increases, as shown in Fig. 5. An immediate solution is to increase the alphabet size in the phase modulation. Continuous phase encoding is the limiting case when the alphabet size approaches infinity. With continuous phase encoding, decoding becomes a parameter-estimation problem, in which one endeavors to acquire an estimate $\tilde{\theta}$ of the encoded phase θ based on the received state in the ensemble Σ_θ^M. The conditional distribution $P(\tilde{\theta}|\theta)$ describes the measurement statistics. Since the encoding θ is uniformly distributed in $[0, 2\pi)$, the per-mode communication rate reads

$$R_{P(\cdot|\cdot)} = \frac{1}{M} \left(\log_2(2\pi) + \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^{2\pi} d\tilde{\theta} P(\tilde{\theta}|\theta) \log_2 P(\tilde{\theta}|\theta)\right). \tag{20}$$

The decoding of a continuous phase requires a design for phase estimation in the presence of large amount of noise with $N_B \gg 1$. To this end, we first show that the TMSV is the asymptotically optimal input state for noisy phase estimation in the limit of strong noise and weak signal, as it maximizes the QFI among all states (see the details in Appendix D 1). Moreover, we derive an adaptive version of the OPA receiver which is the asymptotically optimal receiver for phase estimation with the TMSV state, as it saturates the QFI in the limit of strong noise and weak signal. By combining these results, a noisy phase-estimation protocol that is asymptotically optimal in the limit of a large number of copies is devised.

Unfortunately, noisy phase estimation operating in the large-number-of-copies limit ($M \to \infty$) cannot be used as a decoding strategy, because it leads to a zero per-mode rate. Therefore, it is crucial to optimize the phase-estimation performance with a finite or even small number of modes. In Appendix D 2, two adaptive receiver designs are presented. The basic idea is to introduce a sequence of measurements, each performed on a subset of M_l modes. The setting for each measurement is determined by a prior probability distribution, updated based on previous measurement results through an approach based on the maximum Fisher information or the maximum Van Trees information [56–58]. We find that the Van Trees approach gives much better performance.

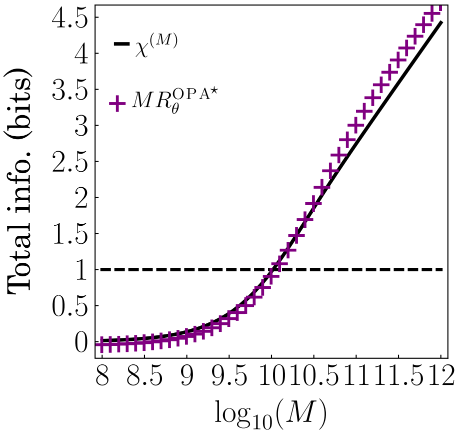

FIG. 10. Performance of continuous phase encoding with adaptive optical-parametric-amplifier receiver. Purple crosses, M-mode information rate MR_θ^{OPA} of OPA; black line, M-mode unassisted phase-encoding information rate $\chi^{(M)}$, repetition-coded with phase encoding. The 1-bit-information upper bound of BPSK codes is plotted as a dashed line as a benchmark. Based on Monte Carlo simulations with 50 adaptive measurement cycles and 8×10^5 samples. Parameters: $N_s = 10^{-3}, N_B = 10^4, \kappa = 10^{-3}$.

Using the Van Trees approach, the total information rate is calculated using Eq. (20), and the results are depicted in Fig. 10. In the context of repetition coding, we take the M-mode unassisted phase-encoded Holevo information $\chi^{(M)}$ as a benchmark, assuming repetition coding of identical phase-encoded coherent states in M-mode blocks and no entanglement assistance (see Appendix E for details). Overall, in the region where BPSK saturates the one-bit bound, an extended practical EA advantage enabled by continuous encoding is observed over the performance of repetition-phase-encoded classical communication $\chi^{(M)}$. Although the current numerical-simulation result for the adaptive OPA receiver shows no EA advantage over the unassisted classical capacity without repetition coding, a systematic optimization of the finite-copy phase-estimation protocol may further improve the EA performance.

VI. EXPERIMENTAL DESIGN

A proof-of-concept experiment using the adaptive OPA receiver to beat the Holevo classical capacity can be readily built with off-the-shelf components, as conceptually illustrated in Fig. 11. Similarly to the quantum-illumination experiment [11], broadband entanglement from spontaneous parametric down-conversion (SPDC) can be generated and employed as the signal and the idler. A loosely focused pump is needed to achieve a collection

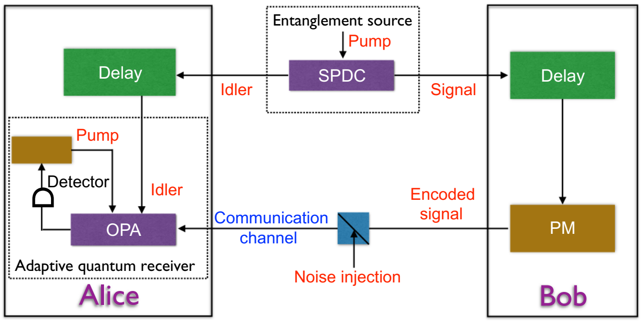

FIG. 11. Experimental setup for EA communication with an adaptive optical-parametric-amplifier receiver. Broadband entangled signal and idler pairs are generated via spontaneous parametric down-conversion and distributed to Bob and Alice. Bob employs phase modulation (PM) to encode on the signal and then sends the encoded photons to Alice. Alice applies an adaptive OPA receiver to her retained idler and received signal to decode Bob's message. In the adaptive OPA receiver, the phase of the pump is adjusted based on prior measurement outcomes to achieve optimal performance.

efficiency greater than 99% for the entanglement source. The idler photons can be stored in a spool of optical fibers with an efficiency in excess of 95%. Other experimental imperfections include free-space-to-fiber coupling loss (less than 5%), detector loss ($1-\eta_D < 2\%$), and filter losses (less than 10%), which contribute to an overall excess loss $1-\kappa_S \sim 15\%$ in the signal and $1-\kappa_I \sim 15\%$ in the idler (combining the storage loss and filter loss). The noisy and lossy channel is usually induced by an adversary in a contested environment, which can be emulated by a beam splitter and a power-tunable amplified spontaneous emission source, e.g., an erbium-doped fiber amplifier, to deliver an N_B up to 500×10^3.

The adaptive OPA receiver can be realized with a field-programmable gate array that processes real-time detector output with a bandwidth greater than 100 MHz, capable of generating a feedforward signal within approximately 100 ns. In conjunction with a 20-GHz electro-optic phase modulator that controls the pump phase, the response time of the adaptive OPA receiver is sufficient to cope with a 1-kbit/s communication rate, corresponding to $M = 2 \times 10^9$. This experimental platform also allows the demonstration of the optimal noisy phase-estimation protocol described in Appendix D 2.

To analyze the key rate of the communication, we include the extra losses $1-\kappa_I$, $1-\kappa_S$, and the detector inefficiency $1-\eta_D$ in the theoretical analysis. We focus on BPSK, which is easier to implement. The analysis is parallel to that in Sec. V A1. With the imperfections, the mean photon count in Eq. (17) changes to $\overline{N}'(\theta) = \eta_D[G\kappa_I N_S + (G-1)(\kappa_S\kappa N_S + \kappa_S N_B + 1) + 2\sqrt{G(G-1)}\cos\theta\sqrt{\kappa_I\kappa_S}C_p]$. As a result, the optimum gain shifts to $G' = 1 + \sqrt{\kappa_I N_S/\kappa_S N_B}$. The distribution of the total photon number across M modes is still given by Eq. (18), with the new mean $\overline{N}'(\theta)$. With some algebra, we find that the variable inside the error function in Eq. (C1) is a factor of $\sqrt{\kappa_I\eta_D}$ smaller than in the ideal case. It is independent of the excess signal loss because of the large noise background N_B. As an example, for $M = 10^9$, and using the same parameters as in Fig. 5, $P_E^{\mathrm{OPA}'} = \mathrm{erfc}(0.43\sqrt{\kappa_I\eta_D})/2$, where erfc denotes the complementary error function. To beat the classical capacity $C(\mathcal{L}^{\kappa,N_B})$, the efficiencies need to satisfy $\kappa_I\eta_D \gtrsim 90\%$. To reach this threshold, the efficiencies need to be improved over the ones in the previous experiment [11]. In particular, if we replace the filter with a free-space filter, the filter loss can be reduced to less than 1%, thus leading to $1-\kappa_I \sim 5\%$ and $1-\eta_D \sim 2\%$. In this case, the communication rate can have an advantage of 3% over the ultimate unassisted classical capacity. When it comes to $M = 10^8$, $P_E^{\mathrm{OPA}'} = \mathrm{erfc}(0.14\sqrt{\kappa_I\eta_D})/2$. With these parameters, the required efficiencies are subject to $\kappa_I\eta_D \gtrsim 84\%$. With the same loss $1-\kappa_I \sim 5\%$, $1-\eta_D \sim 2\%$, the remaining advantage rises to approximately 10%.

VII. BLUEPRINTS FOR JOINT RECEIVERS

Before concluding, we point out some future directions for the design of joint receivers, via combining a FF-SFG receiver and other receivers, for EA communication. As shown in Ref. [8], conditioned on the encoded phase θ and in the limit $N_S \to 0$, one is effectively dealing with a displaced thermal state $\hat{\rho}^{\theta}_{\lambda,n_e}$ with mean $\lambda = e^{i\theta}\sqrt{\kappa\,(1-\epsilon)\,MN_S(N_S+1)/(N_B+1)}$ and thermal noise $n_e \simeq N_S \ln(1/\epsilon)/2$ at the two output ports of the FF-SFG receiver. As explained in Sec. IV, the overall Holevo information is difficult to calculate for the repetition-encoded ensemble Σ^M_θ. As an estimate, the Holevo information of the ensemble $\Sigma^M_{\mathrm{SFG}} = \{\hat{\rho}^{\theta}_{\lambda,n_e}, \theta \sim U[0,2\pi)\}$ is calculated. Although this is not the exact Holevo information of Σ^M_θ, since the equivalence of the quantum states in the ensemble is effective only for the evaluation of the performance of the SFG receiver, one can still obtain interesting observations from this estimate.

The Holevo information of Σ^M_{SFG} can be efficiently calculated (details are given in Appendix E). The result for $\epsilon = 0.05$ is shown in Fig. 3. For $M = 1$, this estimate agrees well with the exact result $\chi(\Sigma^1_\theta)$ and also reaches the EA capacity $C_E(\mathcal{L}^{\kappa,N_B})$. As M increases, the per-mode Holevo information decreases, as expected. Nonetheless, the advantage over the classical capacity survives even for $M > 10^5$.

This analogy inspires us to consider concatenation of the FF-SFG receiver with Holevo-capacity-achieving receivers for classical communication, such as the joint-detection receivers designed in Refs. [59,60]. While the FF-SFG receiver transforms the detection in EA communication into a coherent-state detection problem, the joint receiver optimally extracts information from the coherent states. The complete design of such a receiver will be the subject of future work.

VIII. CONCLUSION

In conclusion, we propose structured encoding and decoding devices to achieve the advantages of entanglement assistance in communication over noisy bosonic channels. We show that phase encoding on a TMSV is asymptotically optimal as the noise increases. In particular, a simple BPSK encoding approaches the optimum Holevo information. In addition to offering higher-than-classical communication rates, the EA communication protocol is secure when the preshared entanglement is perfect, and it beats the fundamental limit on covert communication without entanglement assistance.

We also show that practical repetition coding, e.g., on frequency modes, maintains a $\ln(1/N_S)$ rate advantage, even though the per-mode communication rate decreases. Moreover, with only lossy preshared entanglement available at the current stage of technology, with no feasible quantum network, we show that a slightly smaller advantage remains. For repetitive BPSK encoding, we analyze practical receivers that offer a constant advantage over the classical capacity in the low-signal-power regime. For continuous phase encoding, we show that the use of a TMSV with practical receivers is asymptotically optimum for noisy phase estimation, in the high-noise and large-number-of-copies region. To optimize its parameter-estimation performance with a finite number of copies of states, we develop adaptive Bayesian Van Trees phase-estimation schemes, with fast convergence to the quantum Cramér-Rao bound. However, the effect of the finite number of states prevents any quantum advantage, and optimization is still an open question. Nevertheless, the results on repetition coding provide a straightforward way to implement communications with a practically correctable error rate.

ACKNOWLEDGMENTS

This research is sponsored by the Army Research Office and is accomplished under Grant No. W911NF-19-1-0418. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation herein. Z.Z. acknowledges the National Science Foundation, Grant No. CCF-1907918. Q.Z. and Z.Z. also acknowledge the University of Arizona for support. Q.Z. acknowledges Boulat Bash, Saikat Guha, Jeffrey Shapiro, and Nicolas Cerf for discussions. The authors are also grateful for the comments from the anonymous referee, which greatly improved the paper.

APPENDIX A: CLASSICAL COMMUNICATION WITH COHERENT STATES

Classical communication protocols transmit coherent states $|\alpha\rangle$, encoding information on their real and imaginary quadratures with a bounded average photon number $\overline{|\alpha|^2} = N_S$. Then the transmitted states suffer thermal noise modeled by a thermal-loss channel $\mathcal{L}^{\kappa,N_B}$ with transmissivity κ and noise N_B. Finally, at the receiver the quadratures of the noisy coherent states are measured by homodyne or heterodyne detection, producing Gaussian measurement statistics. The information rate can be obtained through the Shannon capacity. For homodyne detection, the information is encoded on a single quadrature, and thereby the average signal power is $4\kappa N_S$ and the white noise is $2N_B + 1$; for heterodyne detection, both of the quadratures are encoded, and so there are effectively two white-noise channels, each with average signal power κN_S (with a factor of one half from dividing the encoding power into two quadratures and a factor of two from the heterodyne splitting) and noise $(2N_B + 1)/2 + 1/2 = N_B + 1$. Then the information rate is given by the following equation (which can also be found in Ref. [22]):

$$
\begin{aligned}
C_{\text{hom}} &= \frac{1}{2}\log_2\left(1 + \frac{4\kappa N_S}{1 + 2N_B}\right),\\
C_{\text{het}} &= \log_2\left(1 + \frac{\kappa N_S}{1 + N_B}\right).
\end{aligned}
\tag{A1}
$$

In the asymptotic limit of $\kappa N_S \gg 1$, we have

$$
C_{\text{het}}/C(\mathcal{L}^{\kappa,N_B}) = 1 + O[1/\ln(\kappa N_S), N_B]. \tag{A2}
$$

In the asymptotic limit $\kappa N_S \ll N_B$, we have

$$
C_{\text{hom}} \simeq C_{\text{het}} \simeq C(\mathcal{L}^{\kappa,N_B}) \simeq \frac{\kappa N_S}{N_B \ln 2}. \tag{A3}
$$

APPENDIX B: ACCESSIBLE INFORMATION FOR TMSV WITH CONTINUOUS PHASE ENCODING

The accessible (Holevo) information after the channel can be obtained from

$$
\chi\left(\Sigma_\theta^1\right) = S\left(\int_0^{2\pi} d\theta\, \hat{\rho}_{RI}^\theta/2\pi\right) - \int_0^{2\pi} d\theta\, S(\hat{\rho}_{RI}^\theta)/2\pi. \tag{B1}
$$

The conditional entropy $S(\hat{\rho}_{RI}^\theta)$ can be straightforwardly calculated because the state is Gaussian [50]. Note also that $\hat{\rho}_{RI}^\theta = (\hat{U}_\theta \otimes \hat{I})\hat{\rho}_{RI}(\hat{U}_\theta^\dagger \otimes \hat{I})$, where $\hat{\rho}_{RI} \equiv \mathcal{L}^{\kappa,N_B}\left[\hat{\psi}_{SI}^{N_S}\right]$ has the covariance matrix $\mathbf{\Lambda}_{\theta=0}$ in Eq. (9).

Thus the unconditional term is

$$\int_0^{2\pi} d\theta\, S(\hat{\rho}_{RI}^{\theta})/2\pi = S(\hat{\rho}_{RI}) \\ = g[(\mu_+ - 1)/2] + g[(\mu_- - 1)/2]. \tag{B2}$$

Here the symplectic eigenvalues of the covariance matrix $\mathbf{\Lambda}_{\theta=0}$ in Eq. (9) are $\mu_\pm = \frac{1}{2}\left[\pm(S-A) + \sqrt{(A+S)^2 - 4C^2}\right]$, with $A = 2\,(N_B + \kappa N_S) + 1, C = 2C_p, S = 2N_S + 1$. In the limit of $N_B \gg 1, N_S \ll 1$, one can obtain

$$S(\hat{\rho}_{RI}^{\theta=0}) = \log_2(N_B) - N_S \log_2(N_S) + \frac{1+N_S}{\ln(2)} \\ + \frac{\kappa N_S[1/\ln(2) + \log_2(N_S)] + 1/2\ln(2)}{N_B} \\ + O(N_S^2, 1/N_B^2). \tag{B3}$$

The number-basis matrix element of the unconditional state $\int_0^{2\pi} d\theta\, \hat{\rho}_{RI}^{\theta}/2\pi$ can be obtained analytically. We first obtain the number-basis density matrix and then integrate over the unitary $\hat{U}_\theta \otimes \hat{I}$.

From the covariance matrix $\mathbf{\Lambda}_{\theta=0}$ of the Gaussian state $\hat{\rho}_{RI}$, we can obtain the result that the density matrix in the number basis $\langle n_1, n_2|\hat{\rho}_{RI}|n_1', n_2{\prime}\rangle$ is only nonzero when $n_1 - n_1' = n_2 - n_2'$, and the nonzero terms equal

$$\sqrt{\frac{n_1!n_2!}{n_1'!n_2'!}}(-1)^{1+n_2+n_2'}2^{2+n_2-n_2'}C^{n_2-n_2'} \\ \times \frac{(-1+C^2+E+S-ES)^{1+n_1'+n_2}}{X^{1+n_1}Y^{1+n_2}} \\ \times F_R\left(1+n_1, 1+n_2, 1+n_2-n_2', \frac{4C^2}{XY}\right), \tag{B4}$$

where $F_R(a,b,c,z)$ is the regularized hypergeometric function, $X = [1 + C^2 + E - (1+E)S]$, and $Y = [C^2 - (E-1)(S+1)]$, with $C = 2C_p, E = 1 + 2(N_B + \kappa N_S), S = (1 + 2N_S)$.

Because $\hat{U}_\theta \otimes \hat{I}\,|n_1\rangle_R\,|n_2\rangle_I\,\langle n_1'|_R\,\langle n_2'|_I\,\hat{U}_\theta^\dagger \otimes \hat{I} = e^{i\theta(n_1-n_1')}\,|n_1\rangle_R\,|n_2\rangle_I\,\langle n_1'|_R\,\langle n_2'|_I$, the integration will lead to $n_1 = n_1'$. Combined with the fact that $n_1 - n_1' = n_2 - n_2'$, we see that the density matrix of $\int_0^{2\pi} d\theta\,\hat{\rho}_{RI}^{\theta}/2\pi$ is diagonal in the number basis, with $p(n_1,n_2) \equiv \langle n_1, n_2|\hat{\rho}_{RI}|n_1,n_2\rangle$ given by

$$p(n_1,n_2) = -4F_R\left(1+n_1, 1+n_2, 1, \frac{4C^2}{XY}\right) \\ \frac{(-1+C^2+E+S-ES)^{1+n_1+n_2}}{X^{1+n_1}Y^{1+n_2}}. \tag{B5}$$

Thus the unconditional entropy is

$$S\left(\int_0^{2\pi} d\theta\,\hat{\rho}_{RI}^{\theta}/2\pi\right) = -\sum_{n_1,n_2=0}^{\infty} p(n_1,n_2)\log_2[p(n_1,n_2)]. \tag{B6}$$

The rest of the analysis is to asymptotically expand the result. In the limit of $N_B \gg 1$, we have $4C^2/XY = \kappa/[N_B(1-\kappa) + N_B^2] \ll 1$, and thus we can expand using $F_R(1+n_1, 1+n_2, 1, x) = 1 + (1 + n_1 + n_2 + n_1 n_2)x + O(x^2)$. With the above expansion, we denote the first-order result for $p(n_1,n_2)$ as $p_1(n_1,n_2)$, which is too long to display here. This expansion can be justified by checking the normalization $\sum_{n_1,n_2=0}^{\infty} p_1(n_1,n_2) = 1 - \kappa^2(1+N_S)^2/N_B^2 + O(1/N_B^3)$, which is accurate to high order. Further expansion and summation lead to

$$S\left(\int_0^{2\pi} d\theta\,\hat{\rho}_{RI}^{\theta}/2\pi\right) = \log_2(N_B) - N_S\log_2(N_S) + \frac{1+N_S}{\ln(2)} \\ + \frac{2\kappa N_S/2\ln(2) + 1/2\ln(2)}{N_B} \\ + O(N_S^2, 1/N_B^2). \tag{B7}$$

Overall, combining Eqs. (B1), (B3), and (B7) and noticing that all higher-order terms in N_S cancel, we have

$$\chi\left(\Sigma_\theta^1\right) = \frac{\kappa N_S(1+N_S)\log_2(1+1/N_S)}{N_B} + O(1/N_B^2). \tag{B8}$$

In the limit of $N_B \gg 1$, by comparing this with the EA classical capacity [Eq. (3)], we have Eq. (11).

APPENDIX C: INFORMATION RATES OF RECEIVERS WITH BINARY PHASE SHIFT KEYING

1. Optical parametric receiver

We are interested in the limit $M \gg 1$, while $M\kappa N_S/N_B \ll 1$ still holds, such that $P_{\rm OPA}(n|\theta=0;M)$ and $P_{\rm OPA}(n|\theta=\pi;M)$ have approximately the same Gaussian distribution. In this regime, the optimum binary encoding yields an approximately symmetric Gaussian channel. For equal priors, the maximum-likelihood decision rule gives the threshold $N_{\rm th} = M[\sigma(\pi)\overline{N}(0) + \sigma(0)\overline{N}(\pi)]/[\sigma(0) + \sigma(\pi)]$, and the error probability

$$P_E^{\rm OPA} = \frac{1}{2}{\rm erfc}\left(\sqrt{\frac{M\mu_{\rm OPA}^2}{2\sigma_{\rm OPA}^2}}\right), \tag{C1}$$

where $\mu_{\rm OPA} = |\overline{N}(0) - \overline{N}(\pi)|$ and $\sigma_{\rm OPA}^2 = [\sigma(0) + \sigma(\pi)]^2 \simeq 4\overline{N}(\pi/2)[1 + \overline{N}(\pi/2)]$, with $N_B \gg 1$ and the optimal gain $G = 1 + \sqrt{N_S}/N_B$, giving the leading-order signal-to-noise ratio

$$\mu^2_{\rm OPA}/\sigma^2_{\rm OPA} = 4\kappa N_S(1+N_S)/\left[N_B(1+2\sqrt{N_S}+2N_S)\right]. \tag{C2}$$

Here, $\mathrm{erfc}(x) = 1 - 2\int_0^x dt\, e^{-t^2}/\sqrt{\pi}$ is the complementary error function. With Eq. (16), we obtain the information rate by inputting the error probability given in Eq. (C1).

In the limit $N_S \ll 1, N_B \gg 1$, with Eq. (17) we may simplify the variable in the error function to $\sqrt{M\mu^2_{\rm OPA}/2\sigma^2_{\rm OPA}} \simeq \sqrt{2M\kappa N_S/N_B}$. Expanding around $P_E = 1/2$, we obtain

$$R^{\rm OPA}_{\rm BPSK} \simeq 1.27\frac{\kappa N_S}{N_B \ln 2}. \tag{C3}$$

Compared with Eq. (A3), the OPA receiver with BPSK theoretically offers an EA advantage of approximately 27% at best over the unassisted case. However, in the region in which the error rates are small enough for practical error correction, the advantage is smaller.

2. Phase-conjugate receiver

For large M, the photon statistics of the PCR are also approximately Gaussian and symmetric. With the maximum-likelihood decision rule, we have

$$P^{\rm PCR}_E = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{M\mu^2_{\rm PCR}}{2\sigma^2_{\rm PCR}}}\right), \tag{C4}$$

where $\mu_{\rm PCR} = |N_+ - N_-|$ and $\sigma^2_{\rm PCR} = (\sigma_+ + \sigma_-)^2$. Here the means and variances, depending on the phase encoding $\theta \in \{0, \pi\}$, are given by $N_\pm = \pm C_p$ and $\sigma^2_\pm = (1 + N_{X,\pm})N_{X,\pm} + (1 + N_{Y,\pm})N_{Y,\pm} - (N_C - N_I)^2/2$, where the X arm contributes $N_{X,\pm} = (N_C + N_I)/2 \pm C_p$ and the Y arm yields $N_{Y,\pm} = (N_C + N_I)/2 \mp C_p$. Note that the photon number of the idler, $N_I = N_S$, and that of the conjugated signal, $N_C = \kappa N_S + N_B + 1$, are independent of the phase, and the variances are symmetric, i.e., $\sigma^2_+ = \sigma^2_-$. Finally, we have the signal-to-noise ratio

$$\mu^2_{\rm PCR}/\sigma^2_{\rm PCR} = 4\kappa N_S(1+N_S)/\left[N_B(1+2N_S)\right], \tag{C5}$$

in the limit of $N_B \gg 1$.

Note that the PCR has the same signal-to-noise ratio to leading order as the OPA and thus the same asymptotic advantage. However, in the practical error-correctable region, we see that the higher-order term in the denominator is smaller for the PCR than for the OPA, which means that the performance is enhanced, especially when the influence of the higher-order terms is comparable to the EA advantage.

3. Sum-frequency-generation receiver

Based on an analogy with the Dolinar receiver, the choice of r_k is

$$r_{k,\tilde{h}_k} = \sqrt{\eta}|C^{\rm in}_{{\rm si},k}|\left(\frac{(-1)^{\tilde{h}_k}}{\sqrt{1-\exp\left[-2M\left(\sum_{\ell=0}^{k}\lambda^2_\ell - \lambda^2_k/2\right)\right]}}\right), \tag{C6}$$

where $\lambda^2_k = 4\eta|C^{\rm in}_{{\rm si},k}|^2$. The intuition behind this is that when one guesses $\tilde{h}_k = h$ correctly, with the information sufficiently extracted, i.e., $M\sum_{\ell=0}^{k}\lambda^2_\ell \gg 1$, this condition reduces to $r_{k,\tilde{h}_k} \simeq \sqrt{\eta}C^{\rm in}_{{\rm si},k}$, leaving the sum-frequency mode $\hat{b}_k$ close to the vacuum. In this case, any click of the photon detector implies, with high likelihood, that a wrong hypothesis has been made. When this happens, nearly unambiguous information is obtained that can be used to improve the performance.

Similarly to the Dolinar receiver, the minimum error probability for discriminating $\Sigma^M_{\rm BPSK}$ on the SFG receiver, determined by the Helstrom bound, can be estimated based on the discrimination between noisy coherent states with mean $e^{i\theta_h}\sqrt{(1-\epsilon)M\kappa N_S/N_B}$ and noise $-N_S\ln(\epsilon)/2$, with the residual correlation $\epsilon \ll 1$. The numerical results are plotted in Fig. 9.

In the limit where $N_S \ll 1$, the noisy coherent state approximates to a pure coherent state. Its Helstrom bound yields $P_H = \frac{1}{2}\left[1 - \sqrt{1 - \exp\left(-4M\kappa N_S/N_B\right)}\right]$. With $M\kappa N_S/N_B \ll 1$, we have $P_H \simeq 1/2 - \sqrt{M\kappa N_S/N_B}$. The Taylor expansion of Eq. (16) around $P_E = 1/2$ yields

$$R^{\rm SFG}_{\rm BPSK} = \frac{2}{M\ln 2}\left(P_H - \frac{1}{2}\right)^2 = 2\frac{\kappa N_S}{N_B\ln 2}, \tag{C7}$$

which produces an EA advantage of 3 dB.

APPENDIX D: ADAPTIVE NOISY PHASE ESTIMATION

1. Precision limit of noisy phase estimation

The precision limit of the root-mean-square (rms) error in estimating a parameter θ on $M \gg 1$ input states $\hat{\rho}_\theta$ is given by the quantum Cramér-Rao lower bound (CRLB) $\delta\theta \geq 1/\sqrt{M\mathcal{J}_\theta}$ [64–66], where the single-parameter QFI [67–69],

$$\mathcal{J}_\theta = \lim_{d\theta\to 0} 8\frac{1-\sqrt{\mathcal{F}\left(\hat{\rho}_\theta, \hat{\rho}_{\theta+d\theta}\right)}}{d\theta^2}, \tag{D1}$$

is obtained from the Uhlmann fidelity $\mathcal{F}\left(\hat{\rho},\hat{\sigma}\right) = \mathrm{tr}\left(\sqrt{\sqrt{\hat{\rho}}\hat{\sigma}\sqrt{\hat{\rho}}}\right)^2$.

Although the well-known NOON state [70,71] is the optimum for phase estimation in the absence of noise for a fixed photon number, it quickly becomes useless as the noise and loss rise. While the optimum quantum state for noisy phase estimation remains unknown, an upper bound on the QFI has been found [41]. It is straightforward to show that the maximum of the upper bound is achieved in the limit of large photon-number variance limit, i.e., $\Delta^2_{N_S} \to \infty$ [72] and

$$\mathcal{J}_\theta^{\rm UB} = \frac{4\kappa N_S\left[\kappa N_S + (1-\kappa)N_B + 1\right]}{(1-\kappa)\left[\kappa N_S(2N_B+1) - \kappa N_B(N_B+1) + (N_B+1)^2\right]}. \tag{D2}$$

In the limit of $\kappa \ll 1, \kappa N_S \ll N_B, N_B \gg 1$, one has $\mathcal{J}_\theta^{\rm UB} \simeq 4\kappa N_S/N_B$. Since the rms error of the phase estimation is bounded by the period 2π, this QFI holds only in an asymptotic limit, in which the $1/\sqrt{M}$ factor decreases the rms error to $\delta\theta \ll 2\pi$.

With a TMSV source (TMSS), the joint state $\hat{\rho}^\theta_{RI}$ at the receiver in the EA communication protocol is Gaussian, and thus the fidelity and the QFI can be analytically obtained [73,74]:

$$\mathcal{J}_\theta^{\rm TMSS} = \frac{4\kappa N_S(N_S+1)}{1 + N_B(1+2N_S) + N_S(1-\kappa)}. \tag{D3}$$

As a comparison, suppose one uses the coherent state $|\sqrt{N_S}\rangle$, in lieu of the TMSV; the returned state $\mathcal{L}_\theta^{\kappa,N_B}\left(|\sqrt{N_S}\rangle\langle\sqrt{N_S}|\right)$ is then a displaced thermal state with mean $e^{i\theta}\sqrt{\kappa N_S}$ and thermal noise N_B. It is straightforward to derive the fidelity [75], and thus the QFI, under these circumstances: $\mathcal{J}_\theta^{\rm coh} = 4\kappa N_S/(1+2N_B)$. In the limit of $N_B \gg 1$, $\kappa \ll 1$, and $N_S \ll 1$, one has $\mathcal{J}_\theta^{\rm UB} \simeq \mathcal{J}_\theta^{\rm TMSS} \simeq 2\mathcal{J}_\theta^{\rm coh}$. Note that the QFI, in this limit, is related only to the mean of the displacement. As such, the coherent state is anticipated to also be the optimum state in the absence of entanglement assistance. With entanglement assistance, a 3-dB advantage can be achieved. In fact, the EA protocol presented above based on a TMSS is asymptotically optimal in the limit of strong noise and weak signal. In the following, we describe the optimum receiver that saturates the maximum QFI.

2. Optimum receiver for noisy phase estimation—adaptive OPA receiver

As set out in Eq. (18), the OPA receiver's photon-number counting statistics are $P_{\rm OPA}(n|\theta;M)$, conditioned on the encoded phase θ. The corresponding classical Fisher information, $\mathcal{J}_\theta^{\rm OPA} = \sum_{n=0}^\infty [\partial_\theta \log P_{\rm OPA}(n|\theta;M)]^2 P_{\rm OPA}(n|\theta;M)$, can be analytically solved for:

$$\mathcal{J}_\theta^{\rm OPA} = \frac{4(G-1)GM\kappa N_S(1+N_S)\sin^2\theta}{\overline{N}(1+\overline{N})}. \tag{D4}$$

For $N_B \gg 1$ and $G = 1 + \sqrt{N_S}/N_B$, this becomes $\mathcal{J}_\theta^{\rm OPA} \simeq M\sin^2\theta\,\mathcal{J}_\theta^{\rm TMSS}$.

The factor $\sin^2\theta$ indicates that the QFI $\mathcal{J}_\theta^{\rm OPA}$ is phase dependent and is maximized only at $\theta = \pi/2$. Thus, a single-shot phase estimation of a random phase does not usually achieve the maximum QFI. However, with multiple copies of the joint signal-idler state available, viz., $M \gg 1$, this phase-dependent factor can be asymptotically eliminated through a FF mechanism, as utilized in the achievability proof of the single-parameter CRLB [76–78]. A simple FF approach involves first performing an OPA operation on $\sqrt{M}$ modes to obtain an initial estimate $\tilde{\theta} = \theta^\star + O(1/M^{1/4})$ of the true value $\theta^\star$, followed by a phase shift of $\Delta\theta = \pi/2 - \tilde{\theta}$ to set the phases to $\theta^\star + \Delta\theta = \pi/2 + O(1/M^{1/4})$ so that a near-maximum QFI can be attained. A subsequent OPA operation on $M - \sqrt{M}$ modes gives a QFI of $\left(M - \sqrt{M}\right)\left[1 - O(1/\sqrt{M})\right]\mathcal{J}_\theta^{\rm TMSS}$, which, to first order, achieves $M\mathcal{J}_\theta^{\rm TMSS}$.

In EA communication, however, the rate of convergence to the maximum QFI is important. Thus, a systematic Bayesian FF approach is adopted (shown schematically in Fig. 12). The entire set of M mode pairs are measured in K cycles, with each cycle consuming M_k modes such that $\sum_{k=1}^K M_k = M$. In this process, an adaptive strategy $\mathbb{S}_{\mathbf{M}}$ specified by the parameters $\mathbf{M} = \{M_k, 1 \le k \le K\}$ is executed as follows. Initially, the prior probability $p^{(0)}_{\theta^\star}(\theta)$ is set to be uniformly distributed in $[0, 2\pi)$, because the phase encoding is uniform. In the $2 \le k \le K$th cycle, the prior probability distribution $p^{(k-1)}_{\theta^\star|\{n_{k-1}\}}(\cdot|\{n_{k-1}\})$ equals the posterior in the $(k-1)$th cycle, based on all previous measurement results $\{n_{k-1}\} \equiv \{n_1, \ldots, n_{k-1}\}$. Prior to the measurement, a phase shift $\hat{U}_{\Delta\theta_k}$ with $\Delta\theta_k = f\left[p^{(k-1)}_{\theta^\star|\{n_{k-1}\}}\right]$ is applied. The phase shift is a functional of the Bayesian posterior probability of the last cycle, which will be specified later.

After the measurement, the posterior probability is updated, based on the measured photon number n_k and the prior probability using the Bayesian formula

$$p^{(k)}_{\theta^\star|\{n_k\}}(\theta|\{n_k\}) \propto P_{\rm OPA}(n_k|\theta;M_k)\,p^{(k-1)}_{\theta^\star|\{n_{k-1}\}}(\theta|\{n_{k-1}\}). \tag{D5}$$

From this, one can construct the estimator $\tilde{\theta}_k = \arg\max p^{(k)}_{\theta^\star|\{n_k\}}(\theta|\{n_k\})$. After all cycles are executed, the output from the last cycle is chosen as the final estimate.

Approaches based on the maximum Fisher information and the maximum Van Trees information [56–58] are used to determine the phase shift $\Delta\theta_k = f\left[p^{(k-1)}_{\theta^\star|\{n_{k-1}\}}\right]$. The Fisher-information approach simply maximizes the Fisher information by taking $\Delta\theta_k = \arg\max_{\Delta\theta'_k}\mathcal{J}^{\rm OPA}_{\tilde{\theta}_{k-1}+\Delta\theta'_k} = \arg\max_{\Delta\theta'_k}\sin^2(\tilde{\theta}_{k-1}+\Delta\theta'_k)$ based on the current estimator, giving $\Delta\theta_k = \pi/2 - \tilde{\theta}_{k-1}$. The Van Trees approach

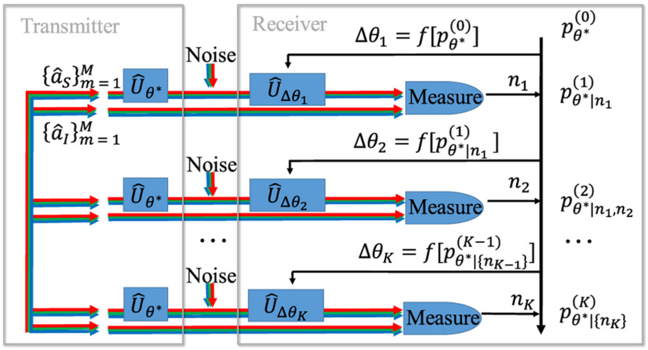
Transmitter
Receiver
$\{\hat{a}_S\}_{m=1}^M$
$\{\hat{a}_I\}_{m=1}^M$
$\hat{U}_{\theta^*}$
Noise
$\hat{U}_{\Delta\theta_1}$
$\hat{U}_{\Delta\theta_2}$
$\hat{U}_{\Delta\theta_K}$
Measure
$\Delta\theta_1 = f[p_{\theta^*}^{(0)}]$
$\Delta\theta_2 = f[p_{\theta^*|n_1}^{(1)}]$
$\Delta\theta_K = f[p_{\theta^*|\{n_{K-1}\}}^{(K-1)}]$
n_1
n_2
n_K
$p_{\theta^*}^{(0)}$
$p_{\theta^*|n_1}^{(1)}$
$p_{\theta^*|n_1,n_2}^{(2)}$
$p_{\theta^*|\{n_K\}}^{(K)}$
...

FIG. 12. Feedforward setup for adaptive schemes. On the transmitter side, the unitary phase encoding $\hat{U}_{\theta^\star}$ encodes identical information on multiple signal modes. On the receiver side, a phase compensation $\hat{U}_{\Delta\theta_k}$ is applied to the signal before the measurement. The compensation angle $\Delta\theta_k$ is determined from the posterior distribution $p_{\theta^\star|\{n_{k-1}\}}^{(k-1)}$.

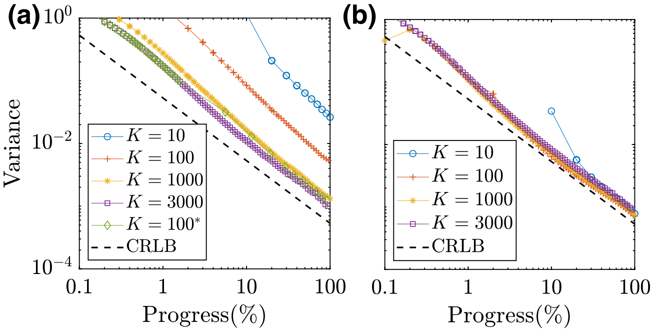

FIG. 13. Evolution of variance for Bayesian phase estimation using the OPA receiver. (a) Maximum-Fisher-information approach. (b) Maximum-Van Trees-information approach. Parameters: $M = 5 \times 10^{12}, N_S = 10^{-3}, N_B = 10^4, \kappa = 10^{-3}$. * The line marked with diamonds is obtained by distributing resources heterogeneously to optimize the performance.

maximizes the average Fisher information, also known as the Van Trees information:

$$\Delta\theta_k = \arg\max_{\Delta\theta_k'} \int d\theta_0\, p_{\theta^\star|\{n_{k-1}\}}^{(k-1)}(\theta_0|\{n_{k-1}\})\mathcal{J}_{\theta_0+\Delta\theta_k'}^{\mathrm{OPA}}. \tag{D6}$$

Because the Van Trees approach makes use of the entire posterior distribution, it yields a performance superior to that of the maximum-Fisher-information approach when the posterior probability has multiple peaks with similar heights.

Seeking an analytical solution for the ultimate posterior probability is challenging. We thus resort to a Monte Carlo simulation to evaluate the performance. We simulate the parameter-estimation process with 8×10^5 samples and record the evolution of the variance evaluated from the posterior probability $p_{\theta^\star|\{n_k\}}^{(k)}(\theta|\{n_k\})$ in each estimation cycle. In Fig. 13, the variance in the kth cycle is plotted against the progress, i.e., the proportion of the modes that have been utilized up to the current cycle, $\sum_{\ell=1}^{k} M_\ell/M$. To benchmark the convergence, the CRLB in Eq. (D3) for each number of modes $\sum_{\ell=1}^{k} M_\ell$ is shown. First, an equal slicing of $M_k = M/K$ is considered. In this case, the Fisher-information approach gives a variance converging to the CRLB as the number of cycles K increases [Fig. 13(a)]. However, the Van Trees approach converges to the CRLB much faster. With $K = 10$ slices, the variance is already close to the CRLB [Fig. 13(b)].

In practice, the implementation of the FF process can be challenging, and so the number of cycles K needs to be minimized. Hence, the Van Trees approach is favorable. One can reduce the number of cycles in the maximum-Fisher-information approach by heterogeneously slicing M into larger segments M_k as we progress to a small-variance region. As an example, the line marked with diamonds is obtained by using $K = 100$ estimation cycles with heterogeneously distributed resources. The first 50 cycles are assigned a small M_k equivalent to that for $K = 3000$, whereas the last 50 cycles are sliced wider, with M_k comparable to that for uniform slices with $K = 100$ (red crosses). A large advantage from the optimization of $\mathbf{M} = \{M_k, 1 \leq k \leq K\}$ is observed. The systematic optimization of the parameter $\mathbf{M}$ is in general a dynamical programming problem, which will be the subject of future work.

APPENDIX E: PHOTON STATISTICS OF THE DISPLACED THERMAL STATE

A displaced thermal state (DTS) $\hat{\rho}_{\lambda,n_e}^{\theta}$ with mean $\lambda = e^{i\theta}|\lambda|$ and thermal noise n_e has the form of the Glauber-Sudarshan P function $P(\alpha) = \exp\left[-|\alpha - \lambda|^2/(2\sigma_P^2)\right]/(2\pi\sigma_P^2)$, where $\sigma_P^2 = n_e/2$. We immediately obtain the density matrix $\hat{\rho}_{\lambda,n_e}^{\theta}$ in the Fock basis,

$$\begin{aligned}\langle n|\hat{\rho}_{\lambda,n_e}^{\theta}|m\rangle &= \langle n| \int d\alpha\, P(\alpha)\, |\alpha\rangle\, \langle\alpha|\, m\rangle \\ &= \frac{e^{-|\lambda|^2/n_e} e^{i(m-n)\theta} n_e^n\, |\lambda|^{m-n}\sqrt{m!}}{(1+n_e)^{m+1}\sqrt{n!}} \\ &\quad \times {}_1\tilde{F}_1\left[m+1, m-n+1, \frac{|\lambda|^2}{n_e(1+n_e)}\right],\end{aligned} \tag{E1}$$

where ${}_1\tilde{F}_1$ is the regularized confluent hypergeometric function [79]. Also, we can obtain the photon-number distribution $P_{\mathrm{DTS}}(n; \lambda, n_e)$ by letting $n = m$, which leads to Laguerre statistics.

Now we calculate the Holevo information of an ensemble of uniformly phase-encoded displaced thermal states, i.e., $\Sigma_\theta^{M,C} = \{(\hat{\rho}_{\lambda,n_e}^{\theta})^{\otimes M}, \theta \sim U[0, 2\pi)\}$. Here M is the number of repetition encodings, and C implies classical states (cf. the TMSV ensemble Σ_θ^M). First, we can use a balanced beam-splitter array to transform each state $(\hat{\rho}^\theta_{\lambda,n_e})^{\otimes M}$ to $\hat{\rho}^\theta_{\sqrt{M}\lambda,n_e} \otimes (\hat{\rho}_{0,n_e})^{\otimes M}$. Because the Holevo information is unchanged under unitary and appending constant states, effectively we can consider the ensemble $\{\hat{\rho}^\theta_{\sqrt{M}\lambda,n_e}, \theta \sim U[0,2\pi)\}$.

Now notice that the conditional entropy $S(\hat{\rho}^\theta_{\sqrt{M}\lambda,n_e})$ is simply $g(n_e)$ due to the invariance of entropy under a unitary transform. Furthermore, the unconditional single-mode state is diagonal in the photon-number basis due to an average over a uniform phase modulation $\theta \sim U[0,2\pi)$. As such, one needs only the Shannon entropy of the photon-number distribution $P_{\rm DTS}(\cdot;\sqrt{M}\lambda,n_e)$ of the displaced thermal state. The final result is

$$\chi(\Sigma_\theta^{M,C}) = H\left[P_{\rm DTS}(\cdot;\sqrt{M}\lambda,n_e)\right] - g(n_e), \tag{E2}$$

which can be efficiently evaluated.

[1] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem, IEEE Trans. Inf. Theory **48,** 2637 (2002).

[2] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nat. Photonics **12,** 724 (2018).

[3] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306,** 1330 (2004).

[4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5,** 222 (2011).

[5] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Rev. **41,** 303 (1999).

[6] S. Lloyd, Enhanced sensitivity of photodetection via quantum illumination, Science **321,** 1463 (2008).

[7] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum Illumination with Gaussian States, Phys. Rev. Lett. **101,** 253601 (2008).

[8] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Optimum Mixed-State Discrimination for Noisy Entanglement-Enhanced Sensing, Phys. Rev. Lett. **118,** 040801 (2017).

[9] S. Barzanjeh, S. Guha, C. Weedbrook, D. Vitali, J. H. Shapiro, and S. Pirandola, Microwave Quantum Illumination, Phys. Rev. Lett. **114,** 080503 (2015).

[10] Z. Zhang, M. Tengner, T. Zhong, F. N. C. Wong, and J. H. Shapiro, Entanglement's Benefit Survives an Entanglement-Breaking Channel, Phys. Rev. Lett. **111,** 010501 (2013).

[11] Z. Zhang, S. Mouradian, F. N. C. Wong, and J. H. Shapiro, Entanglement-Enhanced Sensing in a Lossy and Noisy Environment, Phys. Rev. Lett. **114,** 110506 (2015).

[12] S. Barzanjeh, S. Pirandola, D. Vitali, and J. Fink, Experimental microwave quantum illumination, arXiv:1908.03058 (2019).

[13] C. H. Bennett and S. J. Wiesner, Communication via One- and Two-Particle Operators on Einstein-Podolsky-Rosen States, Phys. Rev. Lett. **69,** 2881 (1992).

[14] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Entanglement-Assisted Classical Capacity of Noisy Quantum Channels, Phys. Rev. Lett. **83,** 3081 (1999).

[15] A. S. Holevo, On entanglement-assisted classical capacity, J. Math. Phys. **43,** 4326 (2002).

[16] M.-H. Hsieh, I. Devetak, and A. Winter, Entanglement-assisted capacity of quantum multiple-access channels, IEEE Trans. Inf. Theory **54,** 3078 (2008).

[17] The initial proof was for finite-dimensional systems, and Refs. [61,62] proved the EA capacity formula for an infinite-dimensional channel with more rigor.

[18] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Classical information capacity of a quantum channel, Phys. Rev. A **54,** 1869 (1996).

[19] B. Schumacher and M. D. Westmoreland, Sending classical information via noisy quantum channels, Phys. Rev. A **56,** 131 (1997).

[20] A. S. Holevo, The capacity of the quantum channel with general signal states, IEEE Trans. Inf. Theory **44,** 269 (1998).

[21] A similar large improvement can also happen in large-dimension depolarizing channels [15,63].

[22] K. Banaszek, L. Kunz, M. Jarzyna, and M. Jachura, in *Proc. SPIE 10910, Free-Space Laser Communications* (SPIE, 2019), Vol. XXXI, p. 109100A.

[23] B. A. Bash, A. H. Gheorghe, M. Patel, J. L. Habif, D. Goeckel, D. Towsley, and S. Guha, Quantum-secure covert communication on bosonic channels, Nat. Commun. **6,** 8626 (2015).

[24] M. S. Bullock, C. N. Gagatsos, S. Guha, and B. A. Bash, Fundamental limits of quantum-secure covert communication over bosonic channels, arXiv:1907.04228 (2019).

[25] R. Prevedel, Y. Lu, W. Matthews, R. Kaltenbaek, and K. J. Resch, Entanglement-Enhanced Classical Communication Over a Noisy Classical Channel, Phys. Rev. Lett. **106,** 110505 (2011).

[26] A. Chiuri, S. Giacomini, C. Macchiavello, and P. Mataloni, Experimental achievement of the entanglement-assisted capacity for the depolarizing channel, Phys. Rev. A **87,** 022333 (2013).

[27] A. S. Holevo and R. F. Werner, Evaluating capacities of bosonic gaussian channels, Phys. Rev. A **63,** 032312 (2001).

[28] G. De Palma, D. Trevisan, and V. Giovannetti, Gaussian States Minimize the Output Entropy of One-Mode Quantum Gaussian Channels, Phys. Rev. Lett. **118,** 160503 (2017).

[29] M. Ban, Quantum dense coding via a two-mode squeezed-vacuum state, J. Opt. B: Quantum Semiclassical Opt. **1,** L9 (1999).

[30] S. L. Braunstein and H. J. Kimble, Dense coding for continuous variables, Phys. Rev. A **61,** 042302 (2000).

[31] M. Ban, Quantum dense coding of continuous variables in a noisy quantum channel, J. Opt. B: Quantum Semiclassical Opt. **2,** 786 (2000).

[32] M. Sohma and O. Hirota, Capacity of a channel assisted by two-mode squeezed states, Phys. Rev. A **68,** 022303 (2003).

[33] J. Mizuno, K. Wakui, A. Furusawa, and M. Sasaki, Experimental demonstration of entanglement-assisted coding using a two-mode squeezed vacuum state, Phys. Rev. A **71,** 012304 (2005).

[34] S. Barzanjeh, S. Pirandola, and C. Weedbrook, Continuous-variable dense coding by optomechanical cavities, Phys. Rev. A **88,** 042331 (2013).

[35] X. Li, Q. Pan, J. Jing, J. Zhang, C. Xie, and K. Peng, Quantum Dense Coding Exploiting a Bright Einstein-Podolsky-Rosen Beam, Phys. Rev. Lett. **88,** 047904 (2002).

[36] M. M. Wilde, P. Hayden, and S. Guha, Information Trade-Offs for Optical Quantum Communication, Phys. Rev. Lett. **108,** 140501 (2012).

[37] A. Anshu, R. Jain, and N. A. Warsi, Building blocks for communication over noisy quantum networks, IEEE Trans. Inf. Theory **65,** 1287 (2019).

[38] H. Qi, Q. Wang, and M. M. Wilde, Applications of position-based coding to classical communication over quantum channels, J. Phys. A: Math. Theor. **51,** 444002 (2018).

[39] S. Khabbazi Oskouei, S. Mancini, and M. M. Wilde, Union bound for quantum information processing, Proc. R. Soc. London, Ser. A **475,** 20180612 (2019).

[40] S. Guha and B. I. Erkmen, Gaussian-state quantum-illumination receivers for target detection, Phys. Rev. A **80,** 052310 (2009).

[41] C. N. Gagatsos, B. A. Bash, S. Guha, and A. Datta, Bounding the quantum limits of precision for phase estimation with loss and thermal noise, Phys. Rev. A **96,** 062306 (2017).

[42] V. Giovannetti, R. García-Patrón, N. J. Cerf, and A. S. Holevo, Ultimate classical communication rates of quantum optical channels, Nat. Photonics **8,** 796 (2014).

[43] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Problemy Peredachi Informatsii **9**, 3 (1973).

[44] J. H. Shapiro, The quantum theory of optical communications, IEEE J. Sel. Top. Quantum Electron. **15,** 1547 (2009).

[45] P. W. Shor, The classical capacity achievable by a quantum channel assisted by limited entanglement, arXiv quant-ph/0402129 (2004).

[46] Q. Zhuang, E. Y. Zhu, and P. W. Shor, Additive Classical Capacity of Quantum Channels Assisted by Noisy Entanglement, Phys. Rev. Lett. **118,** 200503 (2017).

[47] M. M. Wilde and M.-H. Hsieh, The quantum dynamic capacity formula of a quantum channel, Quantum Inf. Process. **11,** 1431 (2012).

[48] E. Y. Zhu, Q. Zhuang, and P. W. Shor, Superadditivity of the Classical Capacity with Limited Entanglement Assistance, Phys. Rev. Lett. **119,** 040503 (2017).

[49] E. Y. Zhu, Q. Zhuang, M.-H. Hsieh, and P. W. Shor, Superadditivity in trade-off capacities of quantum channels, IEEE Trans. Inf. Theory **65,** 3973 (2018).

[50] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84,** 621 (2012).

[51] Q. Zhuang, Z. Zhang, J. Dove, F. N. Wong, and J. H. Shapiro, Floodlight quantum key distribution: A practical route to gigabit-per-second secret-key rates, Phys. Rev. A **94,** 012322 (2016).

[52] S. Pirandola and S. Lloyd, Computable bounds for the discrimination of gaussian states, Phys. Rev. A **78,** 012331 (2008).

[53] K. M. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Discriminating States: The Quantum Chernoff Bound, Phys. Rev. Lett. **98,** 160501 (2007).

[54] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Fundamental limits of repeaterless quantum communications, Nat. Commun. **8,** 15043 (2017).

[55] S. J. Dolinar, “Processing and Transmission of Information,” Technical Report (Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT), 1973).

[56] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* (John Wiley & Sons, New York, 2004).

[57] E. Martínez-Vargas, C. Pineda, F. Leyvraz, and P. Barberis-Blostein, Quantum estimation of unknown parameters, Phys. Rev. A **95,** 012136 (2017).

[58] M. G. Paris, Quantum estimation for quantum technology, Int. J. Quantum Inf. **7,** 125 (2009).

[59] S. Guha, Structured Optical Receivers to Attain Superadditive Capacity and the Holevo Limit, Phys. Rev. Lett. **106,** 240502 (2011).

[60] M. M. Wilde, S. Guha, S.-H. Tan, and S. Lloyd, in *2012 IEEE International Symposium on Information Theory Proceedings* (IEEE, Cambridge, MA, USA, 2012), p. 551.

[61] A. S. Holevo, in *First International Symposium on Quantum Informatics* (International Society for Optics and Photonics, Lipki, Russian Federation, 2003), Vol. 5128, p. 62.

[62] A. S. Holevo and M. E. Shirokov, On classical capacities of infinite-dimensional quantum channels, Probl. Inf. Transm. **49,** 15 (2013).

[63] A. Holevo, Information capacities of quantum measurement channels, Phys. Scr. **2013,** 014034 (2013).

[64] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[65] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Mechanics* (North-Holland, Amsterdam, 1982).

[66] H. Yuen and M. Lax, Multiple-parameter quantum estimation and measurement of nonselfadjoint observables, IEEE Trans. Inf. Theory **19,** 740 (1973).

[67] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Quantum-enhanced measurements without entanglement, Rev. Mod. Phys. **90,** 035006 (2018).

[68] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, Phys. Rev. Lett. **72,** 3439 (1994).

[69] M. Jarzyna and R. Demkowicz-Dobrzański, True precision limits in quantum metrology, New J. Phys. **17,** 013010 (2015).

[70] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. Heinzen, Optimal frequency measurements with maximally correlated states, Phys. Rev. A **54,** R4649 (1996).

[71] U. Dorner, R. Demkowicz-Dobrzanski, B. Smith, J. Lundeen, W. Wasilewski, K. Banaszek, and I. Walmsley, Optimal Quantum Phase Estimation, Phys. Rev. Lett. **102,** 040403 (2009).

[72] The photon-number variance can be unbounded; e.g., $(1-p)\,|0\rangle\,\langle 0| + p\,|N\rangle\,\langle N|$, with mean $pN = N_S$, has a variance diverging in proportion to N.

[73] P. Marian and T. A. Marian, Quantum fisher information on two manifolds of two-mode gaussian states, Phys. Rev. A **93**, 052330 (2016).

[74] L. Banchi, S. L. Braunstein, and S. Pirandola, Quantum Fidelity for Arbitrary Gaussian States, Phys. Rev. Lett. **115,** 260501 (2015).

[75] H. Scutaru, Fidelity for displaced squeezed thermal states and the oscillator semigroup, J. Phys. A: Math. Gen. **31,** 3659 (1998).

[76] A. Fujiwara, Strong consistency and asymptotic efficiency for adaptive quantum estimation problems, J. Phys. A: Math. Gen. **39**, 12489 (2006).

[77] R. D. Gill and S. Massar, in *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers* (World Scientific Publishing, Singapore, 2005), p. 178.

[78] M. Hayashi, *Quantum Information Theory* (Springer-Verlag, Berlin, Heidelberg, 2017).

[79] G. Lachs, Theoretical aspects of mixtures of thermal and coherent radiation, Phys. Rev. **138**, B1012 (1965).