Residual and Destroyed Accessible Information after Measurements

Rui Han,[1,2,*] Gerd Leuchs,[1,2,3] and Markus Grassl[1]
[1]*Max Planck Institute for the Science of Light, 91058 Erlangen, Germany*
[2]*Institut für Optik, Information und Photonik, Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*
[3]*University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

(Received 22 November 2017; published 17 April 2018)

When quantum states are used to send classical information, the receiver performs a measurement on the signal states. The amount of information extracted is often not optimal due to the receiver's measurement scheme and experimental apparatus. For quantum nondemolition measurements, there is potentially some residual information in the postmeasurement state, while part of the information has been extracted and the rest is destroyed. Here, we propose a framework to characterize a quantum measurement by how much information it extracts and destroys, and how much information it leaves in the residual postmeasurement state. The concept is illustrated for several receivers discriminating coherent states.

DOI: 10.1103/PhysRevLett.120.160501

Introduction.—Quantum measurements are often associated with the expectation value of an observable, which corresponds to a physical quantity, such as the average energy of a system or the mean photon number. For this, one has to make repeated measurements on identically prepared copies of a quantum state (ensemble average). In the context of quantum information, on the other hand, one usually considers one-shot measurements. The result of the measurement is described as one out of M possible outcomes, and the measurement provides some classical information about the quantum state. Quantum state discrimination is a special case of this scenario. The receiver who performs the measurement knows that the state he receives is from a set of given quantum states with fixed prior probabilities, and he only needs to identify which state it is.

In the following scenario, referred to as classical-quantum (CQ) communication [1], Alice encodes her classical information using a given set of (orthogonal or nonorthogonal) quantum states and sends a particular signal state to Bob. Bob constructs a set of measurements on the signal he receives to decode the information from his measurement outcomes. In order to make the communication channel between Alice and Bob as efficient or secure as possible, Bob should not count on having identical copies of the same quantum state, but instead, make the most use of every copy he receives.

In order to characterize the communication channel between Alice and Bob, one often uses the average error probability or the mutual information. For a certain class of pure quantum signals, the average probability for Bob making an error when decoding Alice's signal is minimized by the so-called square-root measurement or by the Helstrom measurement [2–4]. From a communication perspective, the mutual information, quantifying how much information is transmitted between Alice and Bob, is the more relevant figure of merit [5]. The two concepts are not equivalent as, for example, minimizing the error probability does not necessarily result in maximal mutual information.

Bob extracts information about Alice's state through his measurement outcome, and the amount of Bob's information is upper bounded by the so-called accessible information. In general, it is very often not possible for Bob to implement an optimal measurement attaining the upper bound. When Bob performs a von Neumann measurement given by rank-one projections, the state after the measurement carries no additional information as it only depends on the measurement outcome, but no longer on the initial state. Hence, the information that has not been extracted by Bob is fully destroyed. On the other hand, if Bob performs a generalized quantum measurement—typically referred to as positive-operator valued measure (POVM) or probability operator measure—with operators of rank larger than 1, the postmeasurement state could still contain some information about the input state. That residual information can be extracted through a subsequent measurement to increase Bob's total information gain [6–9]. How much information is extracted by the measurement depends only on the POVM element. The amount of residual information left in the postmeasurement state, however, depends on the very operators used to implement the POVM measurement.

In the full realm of quantum mechanics, very often, the error probability or the gain of knowledge have been used to quantify the effectiveness of measurements, and fidelity measures have been used to quantify the disturbance of measurements on a quantum state [7–11]. In the present work, however, we fully characterize a quantum measurement using mutual information as the figure of merit, more specifically the amount of information that is extracted, how much information is destroyed, and how much is left over in the postmeasurement state. We illustrate, with practical examples, the power of this approach by looking at four different measurement schemes for binary coherent-state discrimination.

Quantum measurements.—In the general scenario of CQ communication, Alice encodes the information using an ensemble $\mathcal{E}$ of N signal states $\{\rho_j : j = 1, 2, \ldots, N\}$, which can be mixed or pure, with prior probability distribution $\{\eta_j : j = 1, 2, \ldots, N\}$. Bob, in order to identify the state he receives from Alice, can perform any von Neumann or generalized measurement on the state. The measurement can be either direct, i.e., measuring the state itself, or indirect by entangling the state to an ancilla system first and then measuring the ancilla [12]. We describe Bob's measurement by an M-element POVM on the signal state: $\Pi \equiv \{\Pi_k : k = 1, 2, \ldots, M\}$, $\sum_{k=1}^{M} \Pi_k = \mathbb{1}$. When Bob wants to establish a one-to-one correspondence between the measurement outcomes and the full set of signal states, one clearly needs $M \geq N$. For $M > N$, the most simple scheme is obtained by grouping the POVM elements; for example, Bob could associate the measurement outcomes of $\Pi_1, \ldots, \Pi_{k_1}$ with the state ρ_1.

The initial knowledge of the signal state can be represented by the statistical operator $\rho = \sum_{j=1}^{N} \eta_j \rho_j$, where $\mathrm{tr}(\rho) = 1$. The joint probability that the state ρ_j is sent and that the measurement outcome is Π_k is given by $P(\rho_j, \Pi_k) = \eta_j \mathrm{tr}(\Pi_k \rho_j)$. The marginal over the label j of the input states,

$$P_{\Pi_k} = \sum_{j=1}^{N} \eta_j \mathrm{tr}(\Pi_k \rho_j) = \mathrm{tr}(\Pi_k \rho), \tag{1}$$

gives the total probability of having measurement outcome Π_k, and the marginal over the label k of the measurement outcomes, $\sum_{k=1}^{M} P(\rho_j, \Pi_k) = \eta_j$, is just the prior probability of the state ρ_j. The mutual information,

$$I(\mathcal{E} : \Pi) = H(\mathcal{E}) - \sum_{k=1}^{M} P_{\Pi_k} H(\mathcal{E}|\Pi_k), \tag{2}$$

quantifies how much information is shared between Alice and Bob through Bob's POVM measurement Π. The Shannon entropy of Alice's signal is given by $H(\mathcal{E}) = -\sum_{j=1}^{N} \eta_j \log_2 \eta_j$. The conditional entropy $H(\mathcal{E}|\Pi_k)$ quantifies Bob's remaining ignorance about the signal state given the measurement outcome Π_k. The accessible information of the ensemble $\mathcal{E}$ is defined as the maximal mutual information attainable over all possible POVMs,

$$I_{\mathrm{acc}}(\mathcal{E}) = H(\mathcal{E}) - \min_{\text{all } \Pi} \sum_{k=1}^{M} P_{\Pi_k} H(\mathcal{E}|\Pi_k). \tag{3}$$

The accessible information and the set of optimal measurements is known in closed form only for very few special cases, namely, for a communication channel with pure binary states or with real-symmetric trine states [13–15]. In general, the accessible information is usually obtained using numerical optimization methods [16–19]. Holevo's theorem provides an upper bound on the accessible information in terms of the so-called Holevo quantity. Although the Holevo bound is asymptotically achievable when collective measurements on a large number of signals are allowed, it is very often not tight when only single-copy measurements are allowed [13,20–22].

When Bob's POVM does not extract all possible information, he could, at least in principle, perform a subsequent measurement on the postmeasurement state to proceed further. Each POVM element corresponds to a general quantum operation with Kraus operator A_k, where $\Pi_k = A_k^\dagger A_k$ [23]. When the measurement outcome for Π_k is obtained, the normalized postmeasurement state $\rho_j^{(k)}$ corresponding to Alice's state ρ_j and the new prior probabilities are [24]

$$\rho_j^{(k)} = \frac{A_k \rho_j A_k^\dagger}{\mathrm{tr}(\Pi_k \rho_j)} \quad \text{and} \quad \eta_j^{(k)} = \frac{\eta_j \mathrm{tr}(\Pi_k \rho_j)}{\mathrm{tr}(\Pi_k \rho)}. \tag{4}$$

They form the ensemble of postmeasurement states $\mathcal{E}^{(k)}$, conditioned on a particular measurement outcome Π_k. Note that we can also express the conditional Shannon entropy $H(\mathcal{E}|\Pi_k)$ in Eqs. (2) and (3) as the Shannon entropy $H(\mathcal{E}^{(k)})$. To discriminate the postmeasurement states, Bob then can perform any subsequent POVM. For an optimal subsequent measurement on $\mathcal{E}^{(k)}$, the remaining ignorance about the ensemble $\mathcal{E}$ is reduced to $H(\mathcal{E}^{(k)}) - I_{\mathrm{acc}}(\mathcal{E}^{(k)})$. Then the maximal mutual information between the ensemble $\mathcal{E}$ of signal states and the outcomes of optimal subsequent measurements is given by

$$I'_{\max}(\mathcal{E}, \Pi) = H(\mathcal{E}) - \sum_{k=1}^{M} P_{\Pi_k} [H(\mathcal{E}^{(k)}) - I_{\mathrm{acc}}(\mathcal{E}^{(k)})], \tag{5}$$

which only depends on $\mathcal{E}$ and Bob's first measurement Π. Note that Bob's final message solely depends on the outcome of the subsequent measurement, because the result of the first measurement is incorporated in the updated new prior probabilities $\{\eta_j^{(k)}\}$ for the discrimination of $\{\rho_j^{(k)}\}$. Therefore, $I'_{\max}(\mathcal{E}, \Pi)$ is never smaller than the mutual information $I(\mathcal{E} : \Pi)$ of the first measurement, i.e., $I'_{\max}(\mathcal{E}, \Pi) \geq I(\mathcal{E} : \Pi)$. Equality holds if and only if the postmeasurement states are independent of the input state, i.e., when $I_{\mathrm{acc}}(\mathcal{E}^{(k)}) = 0$ for all k.

Information-theoretic characterization.—The efficiency of a measurement Π in attaining information can be quantified by the fraction of information extracted, defined as

$$\bar{E} \equiv \frac{I(\mathcal{E}:\Pi)}{I_{\mathrm{acc}}(\mathcal{E})}. \tag{6}$$

The amount of extracted information is normalized by the total accessible information $I_{\mathrm{acc}}(\mathcal{E})$ such that $0 \le \bar{E} \le 1$. When information is not fully extracted by the measurement, i.e., $\bar{E} < 1$, part of the information can still be preserved in the postmeasurement state and hence might be accessible via suitable subsequent measurements. Thus, we define the fraction of residual information that can potentially be extracted via subsequent measurements as

$$\bar{R} \equiv \frac{I'_{\mathrm{max}}(\mathcal{E},\Pi) - I(\mathcal{E}:\Pi)}{I_{\mathrm{acc}}(\mathcal{E})}. \tag{7}$$

The residual information is bounded by $0 \le \bar{R} \le 1 - \bar{E}$.

The maximal mutual information $I'_{\mathrm{max}}(\mathcal{E},\Pi)$ achievable by any multistep protocol with a particular first measurement Π performed by the receiver cannot exceed the accessible information $I_{\mathrm{acc}}(\mathcal{E})$ of the original signal states; thus, $I'_{\mathrm{max}}(\mathcal{E},\Pi) \le I_{\mathrm{acc}}(\mathcal{E})$. The first measurement does not destroy any information if and only if $I'_{\mathrm{max}}(\mathcal{E},\Pi) = I_{\mathrm{acc}}(\mathcal{E})$. Hence, we define the fraction of information destroyed as

$$\bar{D} \equiv \frac{I_{\mathrm{acc}}(\mathcal{E}) - I'_{\mathrm{max}}(\mathcal{E},\Pi)}{I_{\mathrm{acc}}(\mathcal{E})}, \tag{8}$$

which quantifies the reduction of accessible information due to the measurement Π. Combing the three parts—the fraction of extracted information $\bar{E}$, residual information $\bar{R}$, and destroyed information $\bar{D}$, respectively—we have conservation of total accessible information,

$$\bar{E} + \bar{R} + \bar{D} = 1. \tag{9}$$

Here, we choose to use the accessible information, which is computed in (3) via an optimization over all possible measurements, as the conserved quantity and to normalize other quantities by. When the accessible information is not known, we can replace $I_{\mathrm{acc}}(\mathcal{E})$ by the mutual information for a suboptimal measurement $\widetilde{\Pi}$ for the task at hand (such as the Helstrom measurement). Then the corresponding quantities are defined in relation to $\widetilde{\Pi}$.

Examples.—In the following, we illustrate the significance of characterizing a quantum measurement by $\bar{E}$, $\bar{R}$, and $\bar{D}$ in the scenario of binary coherent-state discrimination, an important example for classical-quantum optical communication. For the discrimination of binary coherent states $\{|\alpha\rangle, |-\alpha\rangle\}$ with prior probabilities $\{\eta_1, \eta_2\}$, the well-known Helstrom measurement is not only the measurement that minimizes the average error probability but also the measurement that maximizes the mutual information [14]. Therefore, for such a given set of signals, the amount of accessible information $I_{\mathrm{acc}}(\mathcal{E})$ is known, and it is less than unity owing to the intrinsic nonorthogonality among coherent states. Although the mathematical construction of the Helstrom measurement has been known for decades, it has not yet been experimentally realized due to limitations in both the experimental apparatus and receiver strategies [25].

For discrimination schemes that use a two-element POVM $\{\Pi_1, \Pi_2\}$, including the Helstrom measurement, the measurement outcomes of Π_1 and Π_2 are associated with the signal states $|\alpha\rangle$ and $|-\alpha\rangle$, respectively. The success probabilities for identifying the states are $\{p_1 = \langle\alpha|\Pi_1|\alpha\rangle, p_2 = \langle-\alpha|\Pi_2|-\alpha\rangle\}$, and the error probabilities are $\{r_1 = \langle\alpha|\Pi_2|\alpha\rangle, r_2 = \langle-\alpha|\Pi_1|-\alpha\rangle\}$. The measurement outcome for Π_1 occurs with probability $P_{\Pi_1} = \eta_1 p_1 + \eta_2 r_2 = \eta_1(1 - r_1) + \eta_2 r_2$, and for Π_2 with probability $P_{\Pi_2} = 1 - P_{\Pi_1}$. The mutual information extracted by this measurement is

$$I(\mathcal{E}:\Pi) = H(\eta_1) - P_{\Pi_1} H\left(\frac{\eta_2 r_2}{P_{\Pi_1}}\right) - P_{\Pi_2} H\left(\frac{\eta_1 r_1}{P_{\Pi_2}}\right), \tag{10}$$

where $H(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ gives the Shannon entropy for a binary random variable with distribution $\{p, 1-p\}$. Thus, evaluating (10) at the minimum error probabilities r_1 and r_2 given by the Helstrom measurement yields the accessible information

$$I_{\mathrm{acc}}(\mathcal{E}) = I(\mathcal{E}:\Pi)\Big|_{r_{1,2}=\frac{1}{2}\left(1-\frac{1-2\eta_{2,1}|\langle\alpha|-\alpha\rangle|^2}{\sqrt{1-4\eta_1\eta_2|\langle\alpha|-\alpha\rangle|^2}}\right)}. \tag{11}$$

The maximal mutual information $I'_{\mathrm{max}}(\mathcal{E},\Pi)$ for the subsequent measurement can then be evaluated from a Helstrom measurement on the postmeasurement states.

The minimum error probability attainable by all Gaussian field detectors is achieved by the perfect homodyne receiver [26]. In the hard-decision scheme, a binary decision is made upon the sign of the measured quadrature. The corresponding POVM elements are $\Pi_1 = \int_0^\infty dx|x\rangle\langle x|$ and $\Pi_2 = \int_{-\infty}^0 dx|x\rangle\langle x|$, where $|x\rangle$ denotes the state with quadrature value x. The probability that a coherent state $|\alpha\rangle$ has a measured quadrature x using a balanced homodyne detector is $|\langle\alpha|x\rangle|^2 = \sqrt{2/\pi}e^{-2(x-|\alpha|)^2}$. The probabilities of error $r_1 = r_2 = [1 - \mathrm{erf}(\sqrt{2}|\alpha|)]/2$ are identical for the two signal states, and (10) gives the mutual information for this two-element POVM. The maximum mutual information attainable using a homodyne receiver is, however, only achieved by the soft-decision scheme regarding each measured quadrature value as a measurement outcome of the projector $\Pi_x = |x\rangle\langle x|$ in the continuous space of quadratures. The hard- and soft-decision schemes yield the same average error probability, but the amount of extracted information with the soft-decision scheme is significantly larger; see Fig. 1(a). Since there is no access to the postmeasurement state for such a homodyne receiver,

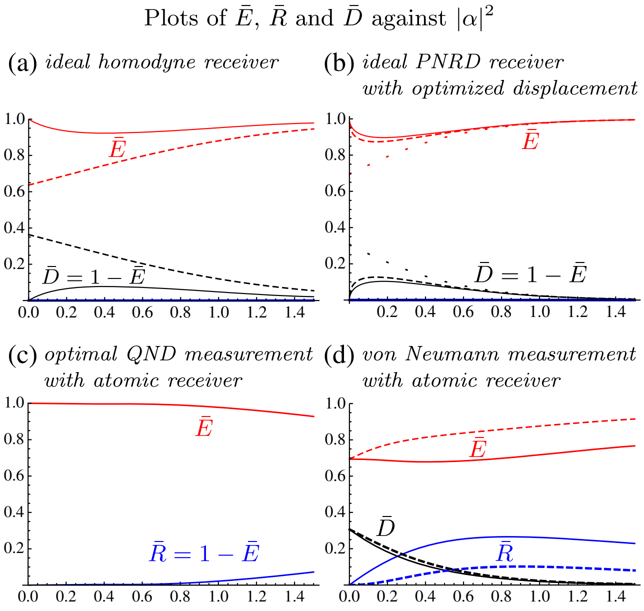

FIG. 1. The fraction of information extracted $\bar{E}$ (red), the fraction of information destroyed $\bar{D}$ (black), and the fraction of residual information $\bar{R}$ (blue) are plotted against the mean photon number $|\alpha|^2$ for the discrimination of binary coherent states $\{|\alpha\rangle, |-\alpha\rangle\}$ with equal priors. The hard-decision scheme for the homodyne receiver (a) and PNRD receiver (b) is illustrated by the dashed curves, and the soft-decision scheme is illustrated by the solid curves. In (b), the dotted curves are plotted for the Kennedy receiver with displacement operation $D(\beta = \alpha)$. The schemes of the nondestructive implementation using an atomic receiver described in Ref. [27] are illustrated by (c) and (d). The data for the optimal scheme where the average error probability is minimized are plotted in (c), and (d) shows the von Neumann measurement scheme that unambiguously discriminates the signal state $|\alpha\rangle$ with a single measurement (solid curves) and two sequential measurements (dashed curves).

the information that fails to be extracted is completely destroyed, i.e., $\bar{D} = 1 - \bar{E}$.

Another class of popular schemes uses photon-number-resolving-detection (PNRD) receivers that discriminate coherent states by their photon statistics, which is a non-Gaussian property of the field [28–31]. A displacement operator $D(\beta)$ displaces the signal states $\{|\alpha\rangle, |-\alpha\rangle\}$ to $\{|\alpha+\beta\rangle, |\beta-\alpha\rangle\}$ before the signal is sent to the photon-number-resolving detectors. In the hard-decision scheme, a binary decision is made upon whether photons are detected or not. In the Fock basis, the POVM elements corresponding to the two signal states are $\Pi_1 = \sum_{j=1}^{\infty} |j\rangle\langle j|$ and $\Pi_2 = |0\rangle\langle 0|$. The probabilities for wrongly identifying the states $|\alpha+\beta\rangle$ and $|\beta-\alpha\rangle$ are $r_1 = |\langle 0|\alpha+\beta\rangle|^2 = e^{-|\alpha+\beta|^2/2}$ and $r_2 = 1 - |\langle 0|\beta-\alpha\rangle|^2 = 1 - e^{-|\beta-\alpha|^2/2}$, respectively. The soft-decision scheme for the PNRD receiver fully takes into account each specific measurement outcome of the POVM given by projections onto all elements of the Fock basis: $\{\Pi_j = |j\rangle\langle j| : j = 0, 1, 2, \ldots\}$. The difference in the fraction of information extracted $\bar{E}$ between the hard- and soft-decision scheme is large for a very weak light field, and becomes smaller as the field amplitude $|\alpha|$ increases; see Fig. 1(b). Similar to the homodyne receiver, the residual light field is completely destroyed by the detector, and hence $\bar{D} = 1 - \bar{E}$.

The Neumark dilation theorem [12] enables the implementation of any two-element POVM by entangling the signal to a qubit ancilla and measuring the ancilla system. This process is described by

$$U|\alpha\rangle|i\rangle = \sqrt{p_1}|\varphi_1\rangle|1\rangle + \sqrt{r_1}|\phi_1\rangle|2\rangle,$$
$$U|-\alpha\rangle|i\rangle = \sqrt{r_2}|\varphi_2\rangle|1\rangle + \sqrt{p_2}|\phi_2\rangle|2\rangle, \qquad (12)$$

where U is the unitary entangling operation and $|i\rangle$ denotes the initial state of the ancilla qubit in the Hilbert space spanned by the orthogonal basis $\{|1\rangle, |2\rangle\}$. The ancilla state is measured using projections $\Pi_1 = |1\rangle\langle 1|$ and $\Pi_2 = |2\rangle\langle 2|$. The Helstrom measurement can be effectively implemented by optimizing the unitary operator U [32]. In practice, however, the set of implementable POVMs is limited by the choice of the physical ancilla and the available unitary operations or couplings between the field and the ancilla. Since the measurement is only on the ancilla, thus, nondestructive on the light state, additional information could be extracted by discriminating the postmeasurement states $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ when the measurement outcome is Π_1, or discriminating the postmeasurement states $\{|\phi_1\rangle, |\phi_2\rangle\}$ when the outcome is Π_2.

Reference [27] investigated the implementation of such nondestructive measurements with the Jaynes-Cummings interaction between the light signal and a two-level atomic ancilla. Effectively, the atom serves as a receiver where the information carried by the coherent state is transferred to and then measured. The optimal minimum-error discrimination strategy corresponds to initially preparing the atom in its ground state $|g\rangle$ and, after its interaction with the light field, projecting it onto the equal superposition states of the ground and excited states $\{|1\rangle = (|g\rangle - i|e\rangle)/\sqrt{2}, |2\rangle = (|g\rangle + i|e\rangle)/\sqrt{2}\}$. The minimum error probability for this scheme can be extremely close to the Helstrom bound for weak coherent signals; i.e., $\bar{E}$ is very close to unity when $|\alpha|^2$ is small. Moreover, this scheme also fully preserves in the postmeasurement states any information that has not yet been extracted, i.e., $I'_{\max}(\mathcal{E}, \Pi) = I_{\mathrm{acc}}(\mathcal{E})$; see Fig. 1(c). Hence, from the perspective of information theory, this discrimination scheme is completely nondestructive as $\bar{D} = 0$ and $\bar{R} = 1 - \bar{E}$.

The implementation for schemes of the Kennedy type [33], which unambiguously discriminate one of the signals, was also investigated in Ref. [27] using an atomic receiver. In order to unambiguously discriminate the state $|\alpha\rangle$, the signal set is displaced by $D(\alpha)$ to $\{|2\alpha\rangle, |0\rangle\}$. If the atom, initially prepared in its ground state $|g\rangle$, is detected in the excited state $|e\rangle$, the decision that the signal state is $|\alpha\rangle$ can be made with certainty and no sequential measurement is needed. If the atom is detected in $|g\rangle$, more information can be extracted by subsequent measurements on the postmeasurement state (for $|\alpha|^2 > 0$). However, in this scheme, part of the information is destroyed due to the atomic measurement, and the accessible information cannot be fully recovered through any subsequent measurement, i.e., $I'_{\max}(\mathcal{E},\Pi) < I_{\mathrm{acc}}(\mathcal{E})$ as long as $\bar{E} \neq 0$ and $|\alpha|^2 > 0$; see Fig. 1(d). The sequential measurement scheme for this unambiguous discrimination strategy has also been investigated, and a significant increment in the extracted information through subsequent measurements has been demonstrated as shown in Fig. 1(d).

Discussion.—For the problem of quantum state discrimination, and, in particular, quantum receivers, one aims at gaining maximal classical information from the quantum state. Our approach, based on mutual information, is not only directly linked to the capacity of the resulting classical communication channel, but allows moreover to quantify how much additional information could be obtained by subsequent measurements. It is the trade-off between the fraction of measured, residual, and destroyed information that well characterizes the performance of a quantum measurement for state discrimination. The method can, for example, be used to analyze sequential measurement schemes for any number of signal states, or a multipartite scenario with local measurements and classical communication of the measurement outcomes.

[*]Present address: Centre for Quantum Technologies, National University of Singapore, 117543 Singapore.
han.rui@quantumlah.org

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[3] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[4] Y. C. Eldar and G. D. Forney, Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).
[5] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
[6] B.-G. Englert, Phys. Rev. Lett. **77**, 2154 (1996).
[7] F. Sciarrino, M. Ricci, F. De Martini, R. Filip, and L. Mišta, Phys. Rev. Lett. **96**, 020408 (2006).
[8] T. C. Ralph, S. D. Bartlett, J. L. O'Brien, G. J. Pryde, and H. M. Wiseman, Phys. Rev. A **73**, 012113 (2006).
[9] E. Nagali, S. Felicetti, P.-L. de Assis, V. D'Ambrosio, R. Filip, and F. Sciarrino, Sci. Rep. **2** (2012).
[10] S. M. Barnett, C. R. Gilson, and M. Sasaki, J. Phys. A **34**, 6755 (2001).
[11] H. Barnum, arXiv:quant-ph/0205155.
[12] M. Neumark, Izv. Akad. Nauk SSSR Ser. Mat. **4**, 53 (1940).
[13] A. S. Holevo, Prob. Peredachi Inf. **9**, 31 (1973) [Probl. Inf. Transm. **9**, 110 (1973)].
[14] L. B. Levitin, in *Quantum Communications and Measurement*, edited by V. P. Belavkin, O. Hirota, and R. L. Hudson (Springer, Boston, 1995), pp. 439–448.
[15] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, Phys. Rev. A **59**, 3325 (1999).
[16] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991).
[17] M. Osaki, O. Hirota, and M. Ban, J. Mod. Opt. **45**, 269 (1998).
[18] P. W. Shor, in *Quantum Communication, Computing, and Measurement 3*, edited by P. Tombesi and O. Hirota (Springer, Boston, 2002), pp. 107–114.
[19] J. Řeháček, B.-G. Englert, and D. Kaszlikowski, Phys. Rev. A **71**, 054303 (2005).
[20] A. S. Holevo, Prob. Peredachi Inf. **9**, 3 (1973) [Probl. Inf. Transm. **9**, 177 (1973)].
[21] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73**, 3047 (1994).
[22] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[23] For simplicity, we restrict ourselves to the construction of the POVM with single Kraus operators here, but the formalism directly generalizes to the case when $\Pi_k = \sum_\ell A^\dagger_{k\ell} A_{k\ell}$.
[24] K. Kraus, Ann. Phys. (N.Y.) **64**, 311 (1971).
[25] The Helstrom measurement would require the receiver to project the state onto a superposition of the two coherent signal states.
[26] M. Takeoka and M. Sasaki, Phys. Rev. A **78**, 022320 (2008).
[27] R. Han, J. A. Bergou, and G. Leuchs, New J. Phys. **20**, 043005 (2018).
[28] C. Wittmann, U. L. Andersen, M. Takeoka, D. Sych, and G. Leuchs, Phys. Rev. Lett. **104**, 100505 (2010).
[29] K. Tsujino, D. Fukuda, G. Fujii, S. Inoue, M. Fujiwara, M. Takeoka, and M. Sasaki, Phys. Rev. Lett. **106**, 250503 (2011).
[30] F. E. Becerra, J. Fan, and A. Migdall, Nat. Photonics **9**, 48 (2015).
[31] D. Sych and G. Leuchs, Phys. Rev. Lett. **117**, 200501 (2016).
[32] R. Han, G. Leuchs, and J. A. Bergou, arXiv:1710.09343.
[33] R. S. Kennedy, Research Laboratory of Electronics Quarterly Progress Report No. 108, MIT, 1973.