Discriminating between quantum states is a fundamental problem in quantum information protocols. The optimum approach saturates the Helstrom bound, which quantifies the unavoidable error probability of mistaking one state for another. Computing the error probability directly requires complete knowledge and diagonalization of the density matrices describing these states. Both of these fundamental requirements become impractically difficult to obtain as the dimensions of the states grow large. In this paper, we analyze quantum illumination as a quantum channel discrimination protocol and circumvent these issues by using the normalized Hilbert-Schmidt inner product as a measure of distinguishability. Using this measure, we show that the greatest advantage gained by quantum illumination over conventional illumination occurs when one uses a Bell state.

DOI: 10.1103/PhysRevA.100.012327

I. INTRODUCTION

One of the main limitations to sending classical information using quantum states is the receiver’s ability to distinguish the states carrying said information. If these states do not have orthogonal support, there is an unavoidable probability that the receiver will mistake one state for another; this creates error in the message. Therefore, it is necessary to have a measure that quantifies the probability of making an error, or a measure of distinguishability when analyzing which states are optimal for sending information.

In 1969, Helstrom’s work [1] on the problem of discriminating between states $\Phi_0$ and $\Phi_1$ that are, respectively, sent with probabilities $p_0$ and $p_1$ established the Helstrom bound

$$\min_{\{\Pi_i\}} p_E = \frac{1}{2} (1 - ||p_0 \Phi_0 - p_1 \Phi_1||_1)$$

as the standard for quantifying the unavoidable error of mistaking one state for another. Indeed, Eq. (1) is the minimization of the error probability

$$p_E = p_0 \text{Tr}[\Phi_0 \Pi_1] + p_1 \text{Tr}[\Phi_1 \Pi_0]$$

with respect to a set of positive operator value measures $\{\Pi_i \geq 0, i = 0, 1\}$ where $\Pi_0 = \hat{1} - \Pi_1$ and $\hat{1}$ is defined as the identity operator. In Eq. (1), the trace norm $||\bullet||_1$ is defined as

$$||\rho||_1 = \text{Tr}[\sqrt{\rho^\dagger \rho}]$$

where $\rho$ is an arbitrary operator and $\rho^\dagger$ is its Hermitian transpose. Because the Helstrom bound is the standard for quantifying unavoidable error, most quantum information protocols that have a distinguishing process need to compute the trace norm, which requires diagonalization in general. This can be difficult to work with when conducting an analysis especially as the dimension of the state becomes large. One such class of protocols that require diagonalization is quantum channel discrimination (QCD). The focus of this paper is on the optimization of a specific QCD protocol.

In QCD, one sends an input state $\Phi^{(in)}$ through a quantum channel which performs one of two operations on the state given by $\{\mathcal{E}_i, i = 0, 1\}$. They then receive the output state $\Phi^{(out)} = \mathcal{E}_i(\Phi^{(in)})$ which is used to determine which operator acted on $\Phi^{(in)}$. Of course, some input states will work better than others depending on the distinguishability of $\Phi_0^{(in)}$ and $\Phi_1^{(out)}$. Here, the probability of mistaking one operation for another is quantified by the Helstrom bound

$$p_E = \min_{\{\Pi_i\}} p_E = \frac{1}{2} (1 - ||p_0 \mathcal{E}_0(\Phi^{(in)}) - p_1 \mathcal{E}_1(\Phi^{(in)})||_1)$$

where it is assumed that an optimal measurement scheme is used. In this context, QCD can be understood as the problem of finding the input state that minimizes Eq. (4) over the space of all $\Phi^{(in)}$. Moreover, extending the space of input states to higher dimension (including joint entangled states $\Phi^{(in)}$) can further reduce the error probability [2,3]. If one partitions the joint system into a signal subsystem and an idler subsystem, where the signal subsystem is sent as a probe, and the idler system is held in a local memory, when the signal returns, a joint measurement can be made; this changes Eq. (4) to

$$p_E = \frac{1}{2} (1 - ||p_0 (E_0 \otimes \hat{I}_I) \Phi_q^{(in)} - p_1 (E_1 \otimes \hat{I}_I) \Phi_q^{(in)}||_1)$$

where $\hat{I}_I$ is the identity operator on the idler subsystem. In this paper, we analyze a postselected model of quantum illumination (QI) as a QCD protocol where Eq. (5) is minimized in the space of all $\Phi_q^{(in)}$. Recently, the fundamental limits of QCD were established [4]. This limit can be less than the minimum of Eq. (5) in the more general setting of adaptive discrimination protocols.

In Lloyd’s seminal paper [5] on QI, the experimenter uses a biphoton $d$-mode Bell state to enhance the detection of a potential surface in a noisy background (see Fig. 1 for a diagram). Formulating the problem with the simplest possible mathematical treatment, Lloyd assumes that a single photon is detected per trial if anything is detected at all. This detection may be due to a returning signal or surrounding noise. In our treatment of QI as a QCD problem, we denote the scenario of receiving a mixture between signal and noise by the operation $(E_0 \otimes \hat{I}_I) \Phi_q$, and the operation where the surface is not present and only noise is detected as $(E_1 \otimes \hat{I}_I) \Phi_q$. Here, $\Phi_q$ is
To simplify the analysis of other quantum information protocols, we introduced the HS distinguishability measure and its teleported counterpart, and we used it in [10] to avoid the trace norm when quantifying the average distance between two states. Although the HS inner product satisfies Josza’s axioms [11] of a fidelity measure, it does not increase monotonically under general quantum operations [11,12]. This is important, where the action of a quantum channel on a pair of quantum states cannot increase their distinguishability (or decrease their fidelity). Fortunately, for the class of states considered in the model of QI considered here, we show that the normalized HS inner product is monotonic with respect to its parametrization.

In this paper, we analyze a modified model of Lloyd’s original QI formulation. Not only do we seek the states $\Phi_q^{(in)}$ that minimize Eq. (5) for this model, we also show that the $d$-dimensional Bell state, defined as a maximally entangled state with equal-dimension subsystems, gives the greatest advantage of QI over conventional illumination (CI). Conventional illumination uses the same input signal as the entangled case, but there are no idlers held to increase its effective brightness; the advantage is defined as the difference in distinguishability between signal and noise as given by QI versus CI.

This paper is structured in the following way. In the next section, we present some background on QI and the mathematical framework used to conduct our analysis. After that, we introduce the HS distinguishability measure and show that it reduces the analysis of QI as a QCD protocol entirely in terms of dimensional arguments and the purity of the ancilla/idler subsystem. After that, we present the result that the $d$-dimensional Bell state gives the greatest advantage over CI for any other choice of $\Phi_q^{(in)}$. This agrees with the recent results of De Palma and Borregaard [13] where they used asymmetric hypothesis testing [14] to show that the two-mode squeezed state gives the greatest advantage of QI. These results are consistent since the two-mode squeezed state is the continuous-variable analogue to the $d$-dimensional Bell state. Finally, we conclude with a discussion on the advantages of using a Hilbert-Schmidt based measure to address the problem of discrimination and its possible applications to quantum information protocols beyond QI.

II. QUANTUM ILLUMINATION

In this section, we describe the model used to analyze QI. To do this, we present the original formulation of QI by Lloyd [5]. Then we discuss our model, which is the postselected model used by Weedbrook et al. [15].

In Lloyd’s original formulation of QI, only single-photon events are considered. In this setting, the signal consists of a single photon in $d_S$ possible modes. It also assumes that the detector can distinguish between $d_S$ modes and its detection window is set to only detect a single photon per trial. In this formulation, Lloyd chooses $\Phi_q$ to be the $d$-mode Bell state, which is given by $\Phi_q = |\Phi_{Bell}\rangle |\Phi_{Bell}\rangle$ where $|\Phi_{Bell}\rangle = d^{-1/2} \sum_{l=1}^{d_S} |l_1\rangle |l_1\rangle$. Here, $|l_1\rangle$ is the state with exactly one photon in the signal mode $k$ and no photons in the other $d_S - 1$ modes. A similar description is given for $|l_1\rangle$. Next we describe how noise is modeled in this setting.
When the signal is lost due to the target being absent, the remaining state is given by
\[
\rho^{(1)} = (\mathcal{E}_1 \otimes \mathbf{1}_t) \Phi_q = \left(1 - \lambda \right) |\text{vac}\rangle \langle \text{vac}| + \lambda \frac{1}{d_S} \right) \otimes \Phi_t \tag{6}
\]
where \(\lambda\) is the average number of noise photons received over many trials, \(|\text{vac}\rangle_S\) represents the vacuum where no photons are found in any of the signal modes, \(\mathbf{1}_S/d_s\) is the state representing the detection of a random mode from the surrounding noise, and \(\Phi_t = \text{Tr}_s[\Phi_q]\) is the idler subsystem held in local memory. If the surface is present, the returning signal is a mixture between signal and noise, which is given by
\[
\rho^{(0)} = (\mathcal{E}_0 \otimes \mathbf{1}_t) \Phi_q = \eta \Phi_q + (1 - \eta)\rho^{(1)} \tag{7}
\]
where \(\eta\), the average number of signal photons received over many trials, represents the degradation of the signal due to noise.

In the paper by Weedbrook et al. [15], they use a simplified model of Lloyd’s original formulation that assumes the detector always receives a photon from either the signal or surrounding noise. They also complete their analysis using a Bell state, though they show that their result is independent of the state chosen in their Appendix. Their formulation of QI corresponds to a postselected model where \(\lambda = 1\) in Eq. (6); this simplifies the remaining state to be
\[
\rho^{(1)} = \frac{\mathbf{1}_S}{d_S} \otimes \Phi_t. \tag{8}
\]
Using this simplified model, they argue that quantum discord [16] explains the underlying advantage of QI. For computational clarity, we will also be working with this postselected model.

In the next section, we use the normalized HS inner product to show that the advantage of this postselected model can be understood in terms of the purity of the idler subsystem given by \(\text{Tr}[\Phi^2_t]\). In fact, a pure idler subsystem (i.e., \(\text{Tr}[\Phi^2_t] = 1\)) is necessarily uncorrelated from the signal, making the protocol using such states equivalent to CI. Any value of \(\text{Tr}[\Phi^2_t] < 1\) implies an advantage gained by using a QI protocol. Where the minimum value of the purity for any density operator is \(d^{-1}\), when \(\text{Tr}[\Phi^2_t] = d^{-1}\), the maximum advantage has been gained; this is equivalent to minimizing Eq. (5). Unlike Weedbrook et al. and Lloyd, we do not assume \(\Phi_q\) are the \(d\)-dimensional Bell state. Instead, we derive in Sec. IV that this state gives the greatest advantage for this model.

### III. Hilbert-Schmidt Distiguishability Measure

Between two arbitrary quantum states \(\rho\) and \(\sigma\), the normalized HS inner product is given by
\[
\langle \rho, \sigma \rangle = \frac{\text{Tr}[\rho^\dagger \sigma]}{\sqrt{\text{Tr}[\rho^\dagger \rho] \text{Tr}[\sigma^\dagger \sigma]}}. \tag{9}
\]
It has a lower extreme value of zero if and only if \(\rho\) and \(\sigma\) are states with orthogonal support [17]. It has an upper extreme value of unity if and only if \(\rho\) and \(\sigma\) are identical, and it is symmetric between them. The normalized HS inner product is invariant under unitary transformations, and it reduces to the ordinary inner product between quantum states when \(\rho\) and \(\sigma\) are pure. Moreover, we will show for the states \(\rho^{(0)}\) from Eq. (7) and the remaining state \(\rho^{(1)}\) from Eq. (8) that it is straightforwardly related to the physical parameters of QI. Now we will write Eq. (9) explicitly in terms of these physical parameters.

To simplify Eq. (9) and write it in terms of the physical parameters of QI, we replace \(\rho\) and \(\sigma\) with \(\rho^{(0)}\) and \(\rho^{(1)}\), respectively. This is computed explicitly in the Appendix. Defining \(\mathcal{H}_{01}\) as the normalized HS inner product between \(\rho^{(0)}\) and \(\rho^{(1)}\) to condense notation, our relations (from the Appendix) simplify \(\mathcal{H}_{01}\) to
\[
\mathcal{H}_{01} \equiv \langle \rho^{(0)}, \rho^{(1)} \rangle = \frac{1}{\sqrt{1 + \eta^2 (d_s K_I - 1)}} \tag{10}
\]
where \(K_I \equiv \text{Tr}[\Phi^2_t]^{-1}\) is the inverse of the purity of the idler state \(\Phi_t\). Here, the physical parameters that completely characterize QI for a fixed \(p_0\) are the relative signal fraction \(\eta\), the dimension of the signal subsystem \(d_s\), and the entanglement between signal and idler which is captured by \(K_I\). Next, we want to show that both the minimum error probability \(p_E^*\) and \(\mathcal{H}_{01}\) are extremized simultaneously with respect to these variables so that we can use the distinguishability measure \(\mathcal{H}_{01}\) to determine which states minimize the unavoidable error probability without diagonalization.

To show that both \(\mathcal{H}_{01}\) and \(p_E^*\) are extremized simultaneously, we must show that they are both monotonic with respect to parameters \(\eta\), \(d_s\), and \(K_I\). For a multivariate function, we take monotonicity to mean monotonic with respect to changes in each variable when all others are held constant. One can verify that \(\mathcal{H}_{01}\) is strictly monotonic by taking the gradient of Eq. (10) and showing that each term maintains the same sign over the intervals \(\eta \in [0, 1]\), \(d_s \in [2, \infty]\), and \(K_I \in [1, d_I]\). The interval for \(d_s\) is justified if one assumes a qubit is the smallest signal used and one is allowed to use an arbitrary number of modes.

From physical considerations, we can argue that \(p_E^*\) monotonically decreases with increasing \(\eta\), \(d_s\), and \(K_I\) given the possible values of the parameters. Holding \(d_s\) and \(K_I\) fixed, it is clear that the error probability strictly decreases with increasing \(\eta\) since it parametrizes the degradation of the signal due to noise. As the signal becomes less noisy, it becomes easier to distinguish it from noise, thus decreasing the chance of error. Given that \(d_s\) represents the possible modes the signal can be in as well as the number of modes distinguishable by the detector, increasing \(d_s\) only increases the dimension that is used to distinguish the known signal from surrounding noise. Therefore, when \(\eta\) and \(K_I\) are held fixed, increasing \(d_s\) strictly decreases the probability of mistaking signal with noise. Alternatively, lower-dimensional signals form a subset of higher-dimensional signals, and expanding the set of states one is minimizing over cannot produce a worse result. As in [10], \(K_I\) is the effective accessible dimension of the idler subsystem that expands the space of joint states obtainable through local manipulations of the signal subsystem (e.g., as in dense coding). Let \(d_s^* \equiv d_s K_I\) represent the effective
dimension of the signal subsystem. From here, we see that \( d_s^2 \in [d_S, d] \). When \( d_s^2 = d_s \), one can only reduce \( \mathcal{H}_{01} \) down to an amount limited by the dimension of the signal subsystem; this is equivalent to a CI protocol. When \( d_s^2 = d \), one has access to the entire dimension of the idler subsystem to minimize \( \mathcal{H}_{01} \). As \( K_I \) increases, the accessible dimension of the signal increases, thus decreasing the probability of mistaking signal from noise.

Where both \( \mathcal{H}_{01} \) and \( p_E^\prime \) decrease monotonically with respect to \( \eta, d_s, \) and \( K_I \), we can reach the minimum \( \mathcal{H}_{01} \) and \( p_E^\prime \) along parametric curves of increasing \( d_s, K_I, \) and \( \eta \). Along these trajectories, \( \mathcal{H}_{01} \) is monotonic with respect to \( p_E^\prime \). Because of this, the set of values of \( \eta, d_s, \) and \( K_I \) that minimizes \( \mathcal{H}_{01} \) also minimizes \( p_E^\prime \). Therefore, one only needs to consider \( \mathcal{H}_{01} \) when seeking to minimize Eq. (5).

Looking at Eq. (10), for a fixed \( \eta \) and composite dimension \( d \), it is clear that the minimum possible value of \( \mathcal{H}_{01} \) is taken when \( K_I = d_I \). Therefore, the states that minimize Eq. (5) are those the idler subsystems of which have minimum purity (and therefore maximum entanglement with the signal). This is equivalent to illumination protocols the remaining states of which, \( \rho^{(0)} \), are maximally mixed. Although all protocols for which \( K_I = d_I \) minimize the error probability for a fixed dimension \( d \), one must maximize \( d_I \) to maximize the advantage of QI. In the next section, we use the Schmidt decomposition to show that the \( d \)-dimensional Bell state is the only state that both has a remaining state that is maximally mixed and maximizes the idler dimension \( d_I \).

IV. PROOF THE BELL STATE GIVES THE MAXIMUM ADVANTAGE

In the previous section, we showed that the advantage of QI is quantified by \( \mathcal{H}_{01} \) and \( p_E^\prime \) and \( K_I \), we can reach the minimum \( \mathcal{H}_{01} \) and \( p_E^\prime \) along parametric curves of increasing \( d_s, K_I, \) and \( \eta \). Along these trajectories, \( \mathcal{H}_{01} \) is monotonic with respect to \( p_E^\prime \). Because of this, the set of values of \( \eta, d_s, \) and \( K_I \) that minimizes \( \mathcal{H}_{01} \) also minimizes \( p_E^\prime \). Therefore, one only needs to consider \( \mathcal{H}_{01} \) when seeking to minimize Eq. (5).

Looking at Eq. (10), for a fixed \( \eta \) and composite dimension \( d \), it is clear that the minimum possible value of \( \mathcal{H}_{01} \) is taken when \( K_I = d_I \). Therefore, the states that minimize Eq. (5) are those the idler subsystems of which have minimum purity (and therefore maximum entanglement with the signal). This is equivalent to illumination protocols the remaining states of which, \( \rho^{(0)} \), are maximally mixed. Although all protocols for which \( K_I = d_I \) minimize the error probability for a fixed dimension \( d \), one must maximize \( d_I \) to maximize the advantage of QI. In the next section, we use the Schmidt decomposition to show that the \( d \)-dimensional Bell state is the only state that both has a remaining state that is maximally mixed and maximizes the idler dimension \( d_I \).

Given an arbitrary entangled pure state \( \Phi_q = |\phi\rangle\langle\phi| \), its Schmidt decomposition is

\[
|\phi\rangle = \sum_{m=1}^{r_{\text{min}}} \sqrt{\lambda_m} |s_m\rangle |l_m\rangle, \quad \sum_m \lambda_m = 1. \tag{11}
\]

where \( r_{\text{min}} \) is the minimum rank between \( \Phi_S \) and \( \Phi_I \), \( |s_m\rangle \) and \( |l_m\rangle \) are orthonormal eigenbasis vectors for the signal and idler subspaces, respectively, and \( \sqrt{\lambda_m} \) are the real non-negative Schmidt coefficients. From here, we see that one must have \( d_S > d_I \) or \( d_S = d_I \) to get \( K_I = d_I \). Otherwise, its greatest value is restricted by the rank of the signal subsystem.

Assuming maximum idler rank \( K_I = d_I \), and the circumstances \( d_S > d_I \) or \( d_S = d_I \), the latter case achieves the largest possible effective signal dimension of \( d_s^2 = d_S^2 \) for a fixed signal, \( d_S \). Because the \( d \)-dimensional Bell state by definition is the only state with \( d_S = d_I \) and \( K_I = d_I \), it gives the greatest advantage of QI over CI for any other choice of \( \Phi_q \). Thus, we have found that the \( d \)-dimensional Bell state minimizes the error probability in the case of postselection on the biphoton section of the composite Hilbert space.

V. DISCUSSION

In this paper, we treated QI as a QCD protocol to determine which states minimize the error probability and give the greatest advantage of QI. Most approaches that address this problem require some diagonalization process such as when computing the trace norm or relative entropy. To avoid this problem we used the normalized HS inner product as a measure of distinguishability, which only requires the trace of the matrix product between density operators.

Using this HS distinguishability measure, we identified three parameters \( (\eta, d_s, K_I) \) in the postselected model of QI that completely determine the distinguishability between \( \rho^{(0)} \) and \( \rho^{(1)} \). The most important of these parameters is \( K_I = \text{Tr}(\Phi_I^2) \) since it quantifies the advantage of QI over CI. When \( K_I = d_I \), one gains the maximum advantage afforded by QI, and when \( K_I = 1, \Phi_S \) and \( \Phi_I \) share zero entanglement, which is equivalent to using a CI protocol.

Although our analysis was on QI, we believe that the HS inner product may have applications to other quantum information protocols. Similar analysis using the HS inner product may be possible for other protocols that use distributed entanglement among ancilla states to gain an advantage when sending or receiving information. It is our intention to extend this research by considering such applications.

ACKNOWLEDGMENTS

The authors wish to graciously acknowledge the anonymous referee for providing very insightful comments and suggestions. We would also like to thank the referee for a more direct derivation of terms in the normalized Hilbert-Schmidt inner products between \( \rho^{(0)} \) and \( \rho^{(1)} \); this derivation now appears in the Appendix. S.R. and J.S. would like to acknowledge support from the National Research Council Research Associateship Program. P.M.A., C.C.T., and J.S. would like to acknowledge support for this work from the Office of the Secretary of Defense Applied Research for Advanced Science and Technology, Quantum Science, and Engineering Program and the Defense Optical Channel Program. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Air Force Research Laboratory.

APPENDIX: SIMPLIFYING THE HILBERT-SCHMIDT INNER PRODUCT IN TERMS OF \( \eta, d_S \), AND \( K_I \)

We wish to compute \( \mathcal{H}_{01} \) from Eq. (10) as

\[
\mathcal{H}_{01} = \frac{\text{Tr}[\rho^{(0)}\rho^{(1)}]}{\sqrt{\text{Tr}[\rho^{(0)2}] \text{Tr}[\rho^{(1)2}]}}. \tag{A1}
\]

where

\[
\rho^{(0)} = \eta \Phi_q + (1 - \eta) \rho^{(1)} \quad \text{and} \quad \rho^{(1)} = \frac{1}{d_S} \otimes \Phi_I. \tag{A2}
\]

We tackle each of the HS inner products in the numerator and denominator of Eq. (A1) one at a time.
First, it will be useful to compute the inner product $\text{Tr}[\Phi \rho^{(1)}]$ directly:

$$\text{Tr}[\Phi \rho^{(1)}] = \sum_{m=1}^{d_f} \sum_{j=1}^{d_s} \langle s_j | s_j \rangle \left( \frac{\Phi^i}{d_s} \otimes \Phi^j \right) |s_j\rangle |s_j\rangle |s_m\rangle |s_m\rangle$$

where $|s_j\rangle$ and $|s_m\rangle$ are orthonormal bases of the signal and idler subspaces, respectively.

Next we compute the numerator $\text{Tr}[\rho^{(0)} \rho^{(1)}]$, which gives

$$\text{Tr}[\rho^{(0)} \rho^{(1)}] = \eta \text{Tr}[\Phi^i \rho^{(1)}] + (1 - \eta) \text{Tr}[\rho^{(12)}]$$

and

$$\text{Tr}[\rho^{(01)}] = \frac{\eta \text{Tr}[\Phi^i \rho^{(1)}]}{d_s} + (1 - \eta) \frac{\text{Tr}[\Phi^j \rho^{(1)}]}{d_s}$$

which is also needed in the denominator of Eq. (A1).

From Eqs. (A5) and (A8), we see that the inner product between $\Phi$ and $\rho^{(1)}$ is equal to the inner product between $\rho^{(1)}$ and itself; this implies that $\text{Tr}[\rho^{(0)} \rho^{(1)}] = \text{Tr}[\rho^{(12)}]$.

Lastly, computing $\text{Tr}[\rho^{(02)}]$ gives

$$\text{Tr}[\rho^{(02)}] = \eta^2 \text{Tr}[\Phi^2] + 2\eta(1 - \eta) \text{Tr}[\Phi \rho^{(1)}] + (1 - \eta^2) \text{Tr}[\rho^{(12)}]$$

Inserting Eqs. (A8) and (A12) into Eq. (A1) and simplifying terms gives Eq. (10).

---

[1] C. Helstrom, *J. Stat Phys* 1, 231 (1969).
[2] A. Y. Kitaev, *Russ. Math. Surv.* 52, 1191 (1997).
[3] M. F. Sacchi, *Phys. Rev. A* 71, 062340 (2005).
[4] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, *npj Quantum Inf.* 5, 3 (2019).
[5] S. Lloyd, *Science* 321, 1463 (2008).
[6] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, *Phys. Rev. Lett.* 101, 253601 (2008).
[7] S. Barzanjeh, S. Guha, C. Weedbrook, D. Vitali, J. H. Shapiro, and S. Pirandola, *Phys. Rev. Lett.* 114, 080503 (2015).
[8] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, *Nat. Photon.* 12, 724 (2018).
[9] J. Lee and M. S. Kim, *Phys. Rev. Lett.* 84, 4236 (2000).
[10] S. Popescu, A. J. Short, and A. Winter, *Nat. Phys.* 2, 754 (2006).
[11] Y.-C. Liang, Y.-H. Yeh, P. E. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond, *Rep. Prog. Phys.* 82, 076001 (2019).
[12] M. Ozawa, *Phys. Lett. A* 268, 158 (2000).
[13] G. De Palma and J. Borregaard, *Phys. Rev. A* 98, 012101 (2018).
[14] G. Spedalieri and S. L. Braunstein, *Phys. Rev. A* 90, 052307 (2014).
[15] C. Weedbrook, S. Pirandola, J. Thompson, V. Vedral, and M. Gu, *New J. Phys.* 18, 042027 (2016).
[16] H. Ollivier and W. H. Zurek, *Phys. Rev. Lett.* 88, 017901 (2001).
[17] $\rho^{(0)}$ and $\rho^{(1)}$ having orthogonal support is equivalent to saying the probability that a system prepared in state $\rho^{(0)}$ will be measured to have an outcome associated to any of the eigenstates of $\rho^{(1)}$ is zero.