PHYSICAL REVIEW A **95**, 012105 (2017)

Quantum information theory of the Bell-state quantum eraser

Jennifer R. Glick* and Christoph Adami†
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA
(Received 26 September 2016; published 6 January 2017)

Quantum systems can display particle- or wavelike properties, depending on the type of measurement that is performed on them. The Bell-state quantum eraser is an experiment that brings the duality to the forefront, as a single measurement can retroactively be made to measure particlelike or wavelike properties (or anything in between). Here we develop a unitary information-theoretic description of this and several related quantum measurement situations that sheds light on the trade-off between the quantum and classical features of the measurement. In particular, we show that both the coherence of the quantum state and the classical information obtained from it can be described using only quantum-information-theoretic tools and that those two measures satisfy an equality on account of the chain rule for entropies. The coherence information and the which-path information have simple interpretations in terms of state preparation and state determination and suggest ways to account for the relationship between the classical and the quantum world.

DOI: 10.1103/PhysRevA.95.012105

I. INTRODUCTION

Wave-particle duality is an iconic feature of quantum mechanics, one not shared by classical systems in which an object cannot simultaneously have a wave and particle nature. Unraveling the mysteries behind wave-particle duality has occupied the better part of the last century, while significant advances in our understanding have come both from clever experimental approaches and from theoretical developments. Pivotal experiments were quantum optical implementations of Wheeler's [1] delayed choice experiment (see, e.g., [2], as well as [3,4], which were based on the theoretical proposal in [5]), which are equivalent in principle to delayed-choice quantum eraser experiments [6,7]. For a thorough review of delayed-choice experiments, see [8] and references therein. The theoretical advances have framed the discussion of the wave-particle duality in terms of a quantum-information-theoretic trade-off between the coherence of the quantum system and the information that one may attempt to obtain about the particle path in the interferometer or double-slit experiment (the "which-path" information) [9–13].

The delayed-choice experiments highlight an important feature of this new understanding of wave-particle duality: while as per Bohr's complementarity principle [14] it is the nature of the experiment that determines whether we shall observe wave- or particlelike behavior, it is clear that the nature of the experiment can be changed after it has already taken place. In other words, the same experiment can retroactively be made to measure wave or particle properties or anything in between [3,4,15]. Such a state of affairs is often greeted with skepticism, as the experiments seem (to some) to imply that the delayed choice of the measurement changes the quantum state retroactively, thus violating causality (see, e.g., the discussion in [15]). The natural interpretation of these results is that a quantum system has both particlelike and wavelike properties at the same time and that the results of measurements can reveal one or a mixture of those characteristics. Due to the classical nature of the measuring devices, however, they do not—in fact, cannot—reflect the true nature of the quantum state. In the following, we will make these arguments in a strictly information-theoretic setting.

We develop the framework of (possibly delayed) complementary measurements (which-path or which-phase) in terms of a quantum-information-theoretic description of the Bell-state quantum eraser, but the formalism is general and applies equally to any situation where measurements are made in parallel on two (and in an obvious extension to several) entangled quantum systems, such as the Garisto-Hardy entanglement eraser [16].

We describe first the ordinary double-slit experiment performed on one half of a Bell state, then the polarization-tagged version where which-path information can be extracted, followed by the quantum erasure procedure. In the next section we describe these steps in terms of classical and quantum information theory that results in an information-theoretic equality that mirrors (and is completely analogous to) the trade-off between distinguishability and visibility of Greenberger and Yasin [9] as well as Englert [10]. The equality involves the coherence of the quantum system and the information obtained about its path, just as Bagan *et al.* have recently shown [13], but we do not assume a specific form for the measure of coherence as it emerges naturally from the information-theoretic analysis. The "conservation law" between coherence and information appears simply as a consequence of the chain rule for entropies. We offer conclusions in which we suggest what information is actually encoded in the measurement devices, given that they cannot reflect information about the quantum state.

II. THE BELL-STATE QUANTUM ERASER

The Bell-state quantum eraser experiment [17], as illustrated in Fig. 1, proceeds as follows. By spontaneous parametric downconversion (SPDC) [6], a pair of photons A

*patte399@msu.edu
†adami@msu.edu

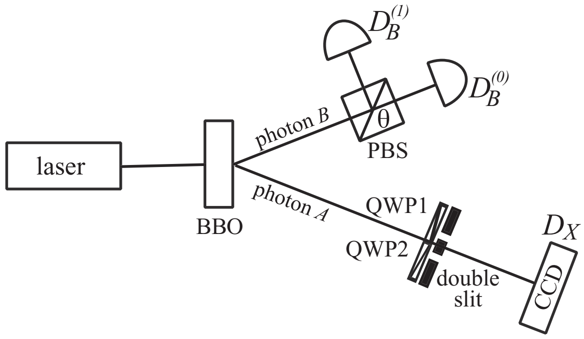

FIG. 1. Schematic of the double-slit Bell-state quantum eraser experiment [17]. After production of the Bell state (1) by type-II spontaneous parametric down-conversion (SPDC) in the β-barium borate (BBO) crystal, photon B travels along the upper branch to a polarizing beam splitter (PBS), with the optical axis oriented at angle θ relative to the $|h\rangle$, $|v\rangle$ basis. Its polarization is subsequently measured in the rotated basis by the photodetectors $D_B^{(0)}$, $D_B^{(1)}$. Photon A travels down to the quarter-wave plates (QWPs) and double slit and then to a CCD camera, denoted D_X, which plays the role of an interference screen.

and B are prepared in an entangled Bell state [18],

$$|\Psi\rangle_{AB} = \frac{|h\rangle|v\rangle + |v\rangle|h\rangle}{\sqrt{2}}, \tag{1}$$

where the first and second states refer to photons A and B, respectively, and $|h\rangle$,$|v\rangle$ are the horizontal and vertical linear polarization states. Photon B, called the "idler," travels along the upper path, where a polarizing beam splitter (PBS) with its optical axis oriented at an angle θ to the $|h\rangle$,$|v\rangle$ basis allows for polarization measurements in specific bases. Meanwhile, photon A, called the "signal," travels along the lower path towards a double-slit apparatus. Photon A will pass through the double slit to subsequently be detected by a CCD camera (denoted D_X), from which it is possible to construct an interference pattern. The pattern is erased by placing two quarter-wave plates (QWPs) in front of each slit. This tags the path of photon A and provides path information. See Fig. 1 for a schematic of the experiment.

A. Splitting the photon path

The full wave function describing the entangled pair of photons A and B is

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi\rangle_Q, \tag{2}$$

where the Hilbert space $\mathcal{H}_A = \mathcal{H}_P \otimes \mathcal{H}_Q$ of photon A is composed of polarization P and spatial Q degrees of freedom. The polarizations of photons A and B, entangled in a Bell state, are decoupled from the spatial state $|\psi\rangle$ of photon A. We drop the spatial states of photon B as they remain decoupled throughout. Sending photon A through the double slit transforms only the spatial states of A so that Eq. (2) evolves to

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes \frac{|\psi_1\rangle_Q + |\psi_2\rangle_Q}{\sqrt{2}}. \tag{3}$$

The states $|\psi_j\rangle$ denote the path of photon A corresponding to slit j [19]. The spatial degree of freedom of A, denoted by Q, is still independent from its polarization.

Tracing over the states of photon B, the density matrix describing the spatial modes of photon A is the pure state

$$\rho_Q = \tfrac{1}{2}(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2| + |\psi_1\rangle_Q\langle\psi_2| + |\psi_2\rangle_Q\langle\psi_1|). \tag{4}$$

The expectation value in the position basis $|x\rangle$ of the screen D_X yields the probability to observe photon A at the spatial location x,

$$\langle x|\rho_Q|x\rangle = p(x) = \tfrac{1}{2}|\psi_1(x) + \psi_2(x)|^2, \tag{5}$$

where we define the amplitudes $\psi_j(x) = \langle x|\psi_j\rangle$. This probability distribution is a coherent superposition, and the usual double-slit interference fringes will be observed on the screen. In the Appendix we show how the characteristic fringes can be derived from a von Neumann measurement of Q by the detector D_X.

B. Tagging the photon path

To extend this discussion to a quantum eraser experiment, a tagging operation is performed on the two branches of the double-slit apparatus in order to provide information about the path of photon A. In practice, this is implemented by placing a quarter-wave plate (QWP) in front of each slit. The Jones matrix for a general wave plate (WP) oriented at an angle β (the fast axis) to our coordinate system (in this case, $|h\rangle$ and $|v\rangle$) is [20,21]

$$U = \begin{pmatrix} \cos\left(\frac{\alpha}{2}\right) + i\sin\left(\frac{\alpha}{2}\right)\cos(2\beta) & i\sin\left(\frac{\alpha}{2}\right)\sin(2\beta) \\ i\sin\left(\frac{\alpha}{2}\right)\sin(2\beta) & \cos\left(\frac{\alpha}{2}\right) - i\sin\left(\frac{\alpha}{2}\right)\cos(2\beta) \end{pmatrix}, \tag{6}$$

where $\alpha = \pi/2$ for a QWP. Here, we orient the QWP in front of slit 1 (slit 2) with its fast axis at $\beta = 45^\circ(\beta = -45^\circ)$, which leads to

$$U_{QWP}^{(\pm)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \pm i \\ \pm i & 1 \end{pmatrix}, \tag{7}$$

where $U_{QWP}^{(+)} = U_{QWP}^{(1)}$ and $U_{QWP}^{(-)} = U_{QWP}^{(2)}$ are the matrices associated with slits 1 and 2, respectively. These transform linearly polarized states $|h\rangle$ and $|v\rangle$ into circularly polarized states according to

$$U_{QWP}^{(1)}|h\rangle = \frac{|h\rangle + i|v\rangle}{\sqrt{2}} = |L\rangle,$$

$$U_{QWP}^{(1)}|v\rangle = \frac{|v\rangle + i|h\rangle}{\sqrt{2}} = i|R\rangle,$$

$$U_{QWP}^{(2)}|h\rangle = \frac{|h\rangle - i|v\rangle}{\sqrt{2}} = |R\rangle,$$

$$U_{QWP}^{(2)}|v\rangle = \frac{|v\rangle - i|h\rangle}{\sqrt{2}} = -i|L\rangle,$$

where $|R\rangle$ ($|L\rangle$) denotes right-handed (left-handed) circular polarization.

When photon A passes through the QWPs and the double slit, its polarization becomes entangled with its spatial degree

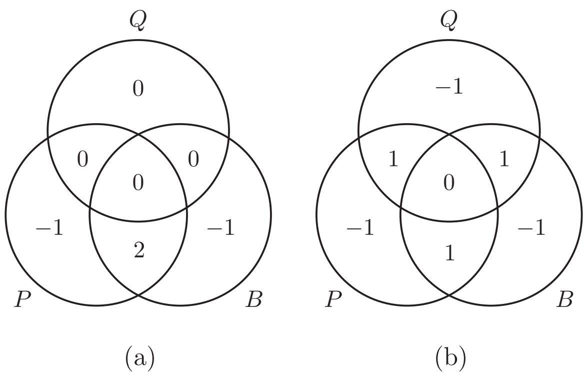

FIG. 2. Entropic Venn diagrams showing the effect of the tagging operation. (a) Before tagging, the polarization of photons A (denoted by variable P) and B are entangled in a Bell state. (b) After tagging, the spatial degree of freedom of photon A (denoted by variable Q) becomes entangled with the polarizations of A and B according to Eq. (9).

of freedom so that the wave function (2) evolves to

$$|\widetilde{\Psi}\rangle_{AB} = \frac{1}{\sqrt{2}}\bigg(\frac{|L\rangle_P|v\rangle_B + i\,|R\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi_1\rangle_Q + \frac{|R\rangle_P|v\rangle_B - i\,|L\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi_2\rangle_Q\bigg), \quad (8)$$

where the tilde indicates that the tagging operation has been performed. Grouping together the polarization states of photon B, we can equivalently express this state as

$$|\widetilde{\Psi}\rangle_{AB} = \frac{1}{\sqrt{2}}\bigg(\frac{|\psi_1\rangle_Q|L\rangle_P + |\psi_2\rangle_Q|R\rangle_P}{\sqrt{2}} \otimes |v\rangle_B + i\,\frac{|\psi_1\rangle_Q|R\rangle_P - |\psi_2\rangle_Q|L\rangle_P}{\sqrt{2}} \otimes |h\rangle_B\bigg). \quad (9)$$

The entanglement between the two degrees of freedom of photon A causes the spatial modes Q to become completely mixed,

$$\rho_Q = \tfrac{1}{2}(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2|), \quad (10)$$

so that interference is no longer observed on the screen. In Fig. 2 we show the entropic Venn diagrams [22] before [Fig. 2(a)] and after [Fig. 2(b)] the tagging operation with the QWPs. In these diagrams, the sum of all the entries in a circle adds up to the entropy of the subsystem, and the entropy shared between subsystems is indicated in the overlap between circles. Conditional entropies appear in unshared areas of the circle and can be negative in quantum mechanics [23] (they must be positive if they are classical Shannon entropies). Entropies shared between three systems (the center of the diagrams in Fig. 2) can be negative in both classical and quantum physics [22]. All of the von Neumann entropies $S(\rho) = -\mathrm{Tr}\rho\log_2\rho$ can be calculated in a straightforward manner from the density matrix $\rho_{QPB} = |\widetilde{\Psi}\rangle_{AB}\langle\widetilde{\Psi}|$ and the marginalized density matrices $\rho_Q = \mathrm{Tr}_{PB}(\rho_{QPB})$, $\rho_B = \mathrm{Tr}_{QP}(\rho_{QPB})$, etc.

Before tagging [see Eq. (2)], Q is completely decoupled from the polarization P of photon A and of photon B, which together are entangled in a Bell state. After tagging [see Eq. (9)], all three variables, Q, P, and B, are in a tripartite entangled state. Note that the ternary mutual entropy $S(Q:P:B)$ vanishes in both diagrams since the underlying density matrix is a pure state [22]. The expression for the ternary shared entropy in terms of subsystem entropies can be read off the Venn diagram in general as $S(Q:P:B) = S(Q) + S(P) + S(B) - S(QB) - S(QP) - S(PB) + S(QPB)$ and similarly for any pairwise shared entropies.

C. Erasing the photon path

As is by now well known [24], it is still possible to extract an interference pattern from the screen data, even when the system Q has been tagged, *if* we have additional information about the state of photon B. Suppose we perform a polarization measurement of B in a basis that is described by an angle θ relative to the $|h\rangle$, $|v\rangle$ basis. For a general change of basis, the polarization states of B are written as

$$|v\rangle_B = U_{00}|0\rangle_B + U_{01}|1\rangle_B, \quad (11)$$

$$|h\rangle_B = U_{10}|0\rangle_B + U_{11}|1\rangle_B. \quad (12)$$

For simplicity, we use the following parametrization for the matrix elements U_{ij} of the rotation operator that transforms $|h\rangle$,$|v\rangle$ to the new basis spanned by $|0\rangle$,$|1\rangle$ in terms of the single angle θ:

$$U = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}. \quad (13)$$

Rewriting the states of B in the basis $|0\rangle$,$|1\rangle$ and entangling B with a polarization detector D_B transforms Eq. (9) to [22]

$$|\widetilde{\Psi}\rangle_{ABD_B} = \frac{1}{2}\sum_{mk} i^m\,\big|\psi_m^k\big\rangle_Q \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \quad (14)$$

where $k = 0,1$ labels the polarization of B and D_B, while $m = 0$ ("L"), 1 ("R") denotes the polarization of A, and where we defined the spatial states $|\psi_m^k\rangle_Q$ of photon A,

$$\big|\psi_L^k\big\rangle_Q = U_{0k}|\psi_1\rangle_Q - iU_{1k}|\psi_2\rangle_Q, \quad (15)$$

$$\big|\psi_R^k\big\rangle_Q = U_{1k}|\psi_1\rangle_Q - iU_{0k}|\psi_2\rangle_Q. \quad (16)$$

These states describe the spatial state of photon A (the system Q) when it has a circular polarization m and is correlated with photon B that has polarization k.

Only the states for a given polarization m are orthonormal,

$$\big\langle\psi_m^{k'}\big|\psi_m^k\big\rangle = \delta_{kk'}, \quad (17)$$

$$\big\langle\psi_L^{k'}\big|\psi_R^k\big\rangle = U_{0k'}^*\,U_{1k} + U_{0k}\,U_{1k'}^*. \quad (18)$$

Of course, the state ρ_Q derived from (14) is still completely mixed as in (10), so that no interference can be observed on the screen. However, as long as the erasure angle θ is nonzero, it is now possible to extract an interference pattern given the outcome of the polarization measurement of photon B (even if that measurement occurs much later than the measurement of photon A).

From the wave function (14), we can compute the joint density matrix for photon $A \equiv QP$ (spatial Q and polarization P) and detector D_B. Tracing over B yields

$$\rho_{AD_B} = \frac{1}{2}\sum_k \rho_A^k \otimes |k\rangle_{D_B}\langle k|, \quad (19)$$

where the (orthonormal) states of photon A, conditional on state k of detector D_B, are $\rho_A^k = |\phi_k\rangle\langle\phi_k|$ and $|\phi_k\rangle = \frac{1}{\sqrt{2}}\sum_m i^m\,|\psi_m^k\rangle_Q \otimes |m\rangle_P$. The effect of the erasure is contained in the behavior of these states as the measurement angle θ is varied. For a measurement in the original basis ($\theta = 0$), detector D_B prepares photon A in one of the fully entangled states: $|\phi_0\rangle \propto |\psi_1\rangle_Q|L\rangle_P + |\psi_2\rangle_Q|R\rangle_P$ and $|\phi_1\rangle \propto |\psi_1\rangle_Q|R\rangle_P - |\psi_2\rangle_Q|L\rangle_P$. From these expressions we can infer, with a polarization measurement of photon A, its path. For instance, outcome $k = 0$, $m = L$ is associated with the spatial state $|\psi_1\rangle$ for slit 1. Therefore, polarization measurements of photon B at $\theta = 0$ yield full path information and no interference fringes.

On the other hand, for a measurement in the diagonal basis at $\theta = \pi/4$, detector D_B prepares photon A in one of the completely decoupled states: $|\phi_0\rangle \propto (|\psi_1\rangle_Q - i|\psi_2\rangle_Q) \otimes (|L\rangle_P + i|R\rangle_P)$ and $|\phi_1\rangle \propto (|\psi_1\rangle_Q + i|\psi_2\rangle_Q) \otimes (-|L\rangle_P + i|R\rangle_P)$. Now, a polarization measurement of A cannot reveal path information, and the coherently summed spatial modes lead to interference fringes. In the Appendix, we compute these interference patterns and show their dependence on the erasure angle θ. Regardless of the temporal order of the two polarization measurements, the measurement of B can be seen as state *preparation*, while the measurement of A is state *determination*, that is, extraction of which-path information.

III. INFORMATION THEORY

A. State preparation

We illustrate the quantum erasure mechanism by building on the entropic Venn diagrams in Fig. 2 and constructing the entropic relationships between the random variables Q, P, and D_B from the joint and marginal entropies associated with Eq. (19). In the basis $|\phi_k\rangle \otimes |k\rangle$ it is clear that the entropy of (19) is $S(QPD_B) = 1$, while the marginal entropies are all $S(Q) = S(P) = S(D_B) = 1$. Tracing over detector D_B, the total entropy of $A \equiv QP$ is also $S(QP) = 1$. The pairwise entropy $S(PD_B)$ is equal to $S(QB)$ by the Schmidt decomposition of (14), which in turn is equal to $S(QD_B)$ by the symmetry between the B and D_B states. The remaining pairwise entropy is computed from the density matrix

$$\rho_{QD_B} = \frac{1}{2}\sum_k \rho_Q^k \otimes |k\rangle_{D_B}\langle k|, \tag{20}$$

where the state of Q conditional on the outcome k of detector D_B is

$$\begin{aligned}\rho_Q^k &= \frac{1}{2}\sum_m |\psi_m^k\rangle_Q\langle\psi_m^k| \\ &= \frac{1}{2}(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2| \\ &\quad + i\,(-)^k \sin 2\theta[|\psi_1\rangle_Q\langle\psi_2| - |\psi_2\rangle_Q\langle\psi_1|]).\end{aligned} \tag{21}$$

From this expression, we see that for a given outcome k, the spatial degree of freedom of photon A is generally no longer mixed [as in Eq. (10)] and has a coherence that is controlled by the sine of the measurement angle. In turn, it is possible to extract interference fringes from the screen. With the Bloch vector $\vec{a} = -(-)^k \sin(2\theta)\,\hat{y}$, Eq. (21) can be expressed as $\rho_Q^k = \frac{1}{2}[\mathbb{1} - (-)^k \sin(2\theta)\,\sigma_y]$, where $\hat{y}$ is a unit vector, $\mathbb{1}$ is the identity matrix of dimension two, and σ_y is a Pauli matrix. This density matrix varies from a fully mixed state $|\vec{a}| = 0$ at $\theta = 0$ to a pure state $|\vec{a}| = 1$ at $\theta = \pi/4$.

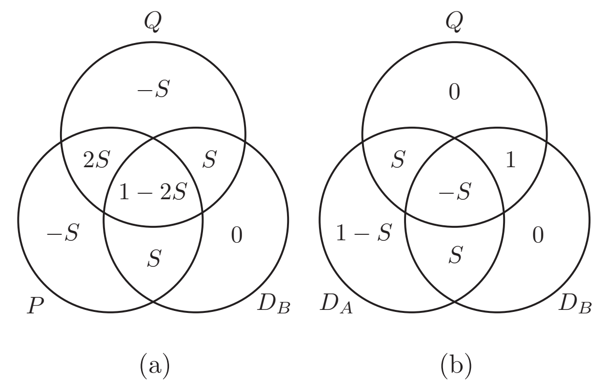

FIG. 3. Entropic Venn diagrams for (a) state preparation with detector D_B [see Eq. (14)] and (b) state determination with detector D_A [see Eq. (23)]. Here, S is the von Neumann entropy of the density matrix in Eq. (21).

To compute the entropy of the block-diagonal matrix in Eq. (20), we first find the entropy of ρ_Q^k. The eigenvalues of ρ_Q^k are $\lambda_\pm = \frac{1}{2}(1 \pm |\vec{a}|) = \frac{1}{2}[1 \pm \sin(2\theta)]$ and are independent of the index k, leading to equal entropies $S(\rho_Q^0) = S(\rho_Q^1)$. Therefore, the entropy of (20), which is also equal to the entropy $S(PD_B)$, is [25]

$$S(QD_B) = 1 + \frac{1}{2}\sum_k S(\rho_Q^k) = 1 + S, \tag{22}$$

where $S = -\lambda_+ \log_2 \lambda_+ - \lambda_- \log_2 \lambda_-$ is the entropy of (21) and varies from $S = 1$ at $\theta = 0$ to $S = 0$ at $\theta = \pi/4$.

The relationship between the variables Q, P, and D_B is summarized by the entropic Venn diagram in Fig. 3(a). As a result of tagging the path of photon A, the spatial and polarization modes of A are entangled, given the state of D_B, for $S > 0$ ($\theta < \pi/4$). The amount of entanglement S varies with the erasure angle θ and specifies the degree to which the polarization P can reveal information about the spatial mode Q. The nonzero ternary mutual entropy $S(Q:P:D_B) = 1 - 2S$ indicates that the mutual entropy of Q and P can be shared by detector D_B [26].

B. State determination

In order to reveal information about the path of photon A, its polarization is measured after it passes the double-slit apparatus using a detector D_A in the circular $|L\rangle,|R\rangle$ basis. After the measurement with D_A, the wave function (14) evolves to

$$|\Psi'\rangle_{AD_ABD_B} = \frac{1}{2}\sum_{mk} i^m\,|\psi_m^k\rangle_Q \otimes |mm\rangle_{PD_A} \otimes |kk\rangle_{BD_B}. \tag{23}$$

The entropic relations between the variables Q, D_A, and D_B are computed from their joint density matrix, which is found by tracing out the polarization states of photons A and B

from (23),

$$\rho_{QD_AD_B} = \frac{1}{4}\sum_{mk} |\psi_m^k\rangle_Q\langle\psi_m^k| \otimes |m\rangle_{D_A}\langle m| \otimes |k\rangle_{D_B}\langle k|. \quad (24)$$

For a set of outcomes k and m, the corresponding spatial state of photon A is $|\psi_m^k\rangle_Q$. It is straightforward to show that the entropy of Eq. (24) is $S(QD_AD_B) = 2$, while the marginal entropies are $S(Q) = S(D_A) = S(D_B) = 1$.

Tracing over the spatial states Q in Eq. (24), we find that the two polarization detectors are uncorrelated,

$$\rho_{D_AD_B} = \tfrac{1}{2}\mathbb{1}_{D_A} \otimes \tfrac{1}{2}\mathbb{1}_{D_B}, \quad (25)$$

with a joint entropy $S(D_AD_B) = 2$ and

$$S(D_A : D_B) = 0. \quad (26)$$

Thus, the measurement with D_A reveals no information about the state of D_B. This is not surprising as the QWPs act as a controlled-NOT gate on the polarization. Indeed, conditioning on the spatial states of photon A yields

$$S(D_A : D_B|Q) = S \geqslant 0. \quad (27)$$

In a sense, therefore, the states of Q encrypt the relationship between the state preparation with D_B and its readout with D_A.

Tracing out the states of the polarization detector D_A in Eq. (24), the joint density matrix ρ_{QD_B} is unchanged from Eq. (20) since the measurement with D_A does not affect the correlations between Q and D_B.

Tracing over the states of D_B in Eq. (24), the joint density matrix for Q and D_A in turn is

$$\rho_{QD_A} = \frac{1}{2}\sum_m \rho_Q^m \otimes |m\rangle_{D_A}\langle m| = \frac{1}{2}\mathbb{1}_Q \otimes \frac{1}{2}\mathbb{1}_{D_A}, \quad (28)$$

where the density matrix of Q, conditional on the polarization outcome m of detector D_A, is

$$\rho_Q^m = \frac{1}{2}\sum_k |\psi_m^k\rangle_Q\langle\psi_m^k| = \frac{1}{2}(|\psi_1\rangle_Q\langle\psi_1| + |\psi_2\rangle_Q\langle\psi_2|), \quad (29)$$

which is independent of the polarization index m and is equivalent to the full density matrix ρ_Q. That is, $\rho_Q^m = \rho_Q = \frac{1}{2}\mathbb{1}_Q$ is a completely mixed state. Finally, the joint entropy of Q and D_A is $S(QD_A) = 2$. The entropic relationships between the variables Q, D_A, and D_B are summarized by the Venn diagram in Fig. 3(b).

C. Information-theoretic origins of coherence and path information

From the marginal and joint entropies computed in the previous section, we can construct information-theoretic relationships between the variables Q, D_A, and D_B.

Coherence. The information shared between the preparation with detector D_B and the quantum system Q (the spatial state of photon A) is given by the mutual entropy

$$S(Q : D_B) = 1 - S \leqslant 1, \quad (30)$$

with S defined in Eq. (22). We can understand how this entropy depends on the erasure angle θ by considering two cases. First, from the joint density matrix ρ_{QD_B} in Eq. (20), a measurement of photon B at an angle $\theta = 0$ decouples Q from detector D_B,

$$\theta = 0: \ \rho_{QD_B} = \tfrac{1}{2}\mathbb{1}_Q \otimes \tfrac{1}{2}\mathbb{1}_{D_B}, \quad (31)$$

and the conditional state ρ_Q^k in Eq. (21) becomes a statistical mixture; that is, interference cannot be observed on the screen. In this case, $S(Q : D_B) = 0$, and there is no information shared between the two variables. However, increasing the erasure angle to $\theta = \pi/4$ leads to perfect correlation,

$$\theta = \frac{\pi}{4} : \rho_{QD_B} = \frac{1}{2}(|f\rangle_Q\langle f| \otimes |0\rangle_{D_B}\langle 0| + |a\rangle_Q\langle a| \otimes |1\rangle_{D_B}\langle 1|), \quad (32)$$

where $|f\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q - i|\psi_2\rangle_Q)$ corresponds to a fringe pattern and $|a\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q + i|\psi_2\rangle_Q)$ corresponds to an antifringe (phased-shifted) pattern. Now, $\rho_Q^0 = |f\rangle_Q\langle f|$ and $\rho_Q^1 = |a\rangle_Q\langle a|$ are coherent superpositions; that is, it is possible to extract interference on the screen. At this angle, $S(Q : D_B) = 1$. Therefore, the mutual entropy $S(Q : D_B)$ is related to the coherence of the conditional states ρ_Q^k and, in turn, to the visibility of interference fringes, as we will see below.

Path information. From the joint density matrix ρ_{QD_A} computed in Eq. (28), the polarization measurement with detector D_A reveals nothing about the spatial degree of freedom of A since the joint state is completely decoupled. It follows that the mutual information vanishes,

$$S(Q : D_A) = 0. \quad (33)$$

In other words, if we do not know the outcome of the polarization measurement of photon B, an attempt to measure the polarization of A after it traverses the double slit and QWPs will not reveal anything about the spatial state of A. On the other hand, if we do know the state of D_B, then the conditional mutual information is

$$S(Q : D_A|D_B) = S \geqslant 0 \quad (34)$$

and varies with the erasure angle θ.

To understand the behavior of this quantity as a function of θ, consider the state of Q and D_A, conditional on the outcome k of D_B. According to Eq. (24), this state is $\rho_{QD_A}^k = \frac{1}{2}\sum_m |\psi_m^k\rangle_Q\langle\psi_m^k| \otimes |m\rangle_{D_A}\langle m|$. For an outcome $k = 0$ of a polarization measurement of photon B at angle $\theta = 0$,

$$\theta = 0 : \rho_{QD_A}^0 = \tfrac{1}{2}\big[|\psi_1\rangle_Q\langle\psi_1| \otimes |L\rangle_{D_A}\langle L| + |\psi_2\rangle_Q\langle\psi_2| \otimes |R\rangle_{D_A}\langle R|\big]. \quad (35)$$

At this angle, the conditional mutual information is maximal, $S(Q : D_A|D_B) = 1$, and it is clear that the state of the polarization detector D_A is associated with one of the states $|\psi_j\rangle_Q$, so that the measurement can reveal information about the path of photon A. For instance, an outcome L corresponds to the state $|\psi_1\rangle_Q$. As the erasure angle θ increases from zero to $\pi/4$, the information we have about the spatial state of A is reduced to zero since the density matrix becomes decoupled:

$$\theta = \frac{\pi}{4} : \ \rho_{QD_A}^0 = |f\rangle_Q\langle f| \otimes \frac{1}{2}\mathbb{1}_{D_A}. \quad (36)$$

At this angle, $S(Q : D_A|D_B) = 0$. Therefore, the tagging operation with the QWPs only reveals information about the path of A as long as we have additional information about the state of photon B from its polarization measurement with detector D_B. Thus, Eq. (34) is the correct expression for which-path information. Note further that

$$S(Q:D_A D_B) = 1, \tag{37}$$

which implies that, given the outcomes of both polarization detectors D_A and D_B, it is possible to predict *with certainty* the outcome of a direct measurement of the path of photon A.

IV. DISCUSSION

The quantities considered in the previous section can be used to generalize the usual concepts of coherence and path information, allowing us to construct a more fundamental relationship that is derived from information-theoretic principles and a no-collapse model of quantum measurement [22].

Recapitulating the results from the previous section, we know that whether we extract fringes or antifringes from the screen is controlled by the state of D_B. The visibility of the fringes is, in turn, related to the coherence of the conditional state ρ_Q^k of Q in Eq. (21) and the mutual information $S(Q:D_B)$. As we have already seen, at angle $\theta = 0$ ($\theta = \pi/4$) the two variables Q and D_B are completely uncorrelated (correlated), the conditional state ρ_Q^k is a statistical mixture (coherent superposition), and we observe no interference (full interference).

On the other hand, information about the path of photon A is determined by the correlation between its polarization (via the state of detector D_A) and its spatial states. This correlation is computed from the conditional mutual information $S(Q:D_A|D_B)$ and must be conditioned on the state of D_B since Q and D_A are otherwise uncorrelated. When $\theta = 0$ ($\theta = \pi/4$), the variables Q and D_A are completely correlated (uncorrelated) given D_B, so that D_A can (cannot) reveal path information, and we extract no interference (full interference) from the screen.

These two information-theoretic quantities, namely, the coherence $S(Q:D_B)$ of system Q and its path information $S(Q:D_A|D_B)$, are fundamentally linked through the chain rule for entropies

$$S(Q:D_B) + S(Q:D_A|D_B) = S(Q:D_A D_B) = 1, \tag{38}$$

and their sum is conserved throughout the erasure process. This information-theoretic formulation of complementarity generalizes earlier attempts [12,13,27] by explicitly referencing the measurement devices. We note that the absence of correlations between detectors D_A and D_B, $S(D_A:D_B) = 0$, is crucial to enforce complementarity.

We show in the top of Fig. 4 the coherence and information in Eq. (38) as a function of the erasure angle θ. We can compare these expressions to two other measures that are commonly used to discuss the wave-particle duality, namely, the distinguishability D and the visibility V [9,10]. In general, $D^2 + V^2 \leqslant 1$ but becomes an equality when the detectors are prepared in pure states. In the particular case we are discussing, $D^2 = \cos^2(2\theta)$, while the fringe visibility $V^2 = \sin^2(2\theta)$. They are shown in the bottom part of Fig. 4 and exhibit a remarkably similar behavior when compared to the information-theoretic complementarity principle.

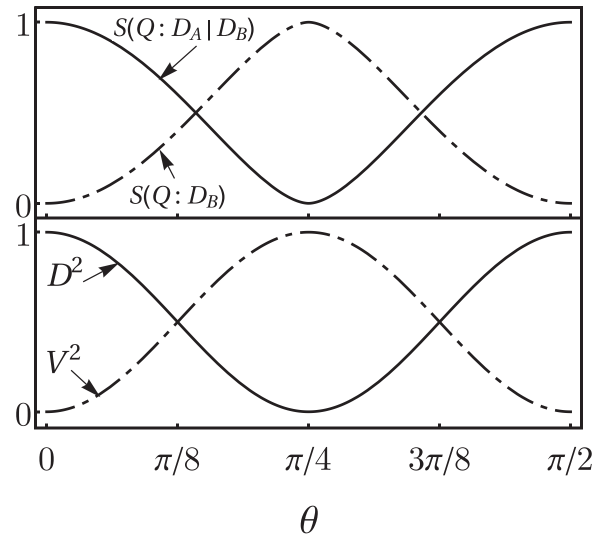

FIG. 4. Top: Relationship between coherence and path information in Eq. (38) for the Bell-state quantum eraser as a function of the erasure angle θ. Shown are the information-theoretic quantities for the path information $S(Q:D_A|D_B) = S$ (solid) and coherence $S(Q:D_B) = 1 - S$ (dashed). Bottom: The square of the distinguishability D (solid) and of the fringe visibility V (dashed), defined in text, as a function of the erasure angle θ. When $\theta = 0$ ($\theta = \pi/4$), there is full (no) path information and no (full) coherence.

From Eq. (38), we can derive additional information-theoretic relations with conditional and mutual entropies. With the definition of conditional mutual information [22], $S(Q:D_A|D_B) = S(Q|D_B) - S(Q|D_A D_B) = S(Q|D_B)$, Eq. (38) becomes a relation between mutual entropy and conditional entropy:

$$S(Q:D_B) + S(Q|D_B) = 1. \tag{39}$$

Furthermore, $S(Q:D_A) - S(Q:D_A|D_B) = S(Q:D_A:D_B) \leqslant 0$, so that Eq. (38) becomes

$$0 \leqslant S(Q:D_A) + S(Q:D_B) \leqslant 1, \tag{40}$$

where the lower bound comes from the non-negativity of mutual entropies. This can be rewritten in terms of conditional entropies as

$$1 \leqslant S(Q|D_A) + S(Q|D_B) \leqslant 2, \tag{41}$$

where we used $S(Q) = 1$.

Bagan *et al.* recently constructed an entropic complementarity relation between coherence and path information in an interferometer [13] using an entropic measure for coherence. For path states with equal probability, orthogonal detector states, and orthogonal measurements, their relation can be written as an equality very similar to ours,

$$C_{\text{rel ent}}(\rho) + H(M:D) = 1, \tag{42}$$

where $C_{\text{rel ent}}(\rho) = 1 - S(\rho)$ is a measure of the coherence of the particle's state ρ in the interferometer and $H(M:D) = S(\rho)$ is the path information, which is the mutual entropy of the path detector states D and the results M of probing them with a measurement. The connection to our result (38) is immediately obvious, as $S(\rho)$ in Eq. (42) is indeed equivalent to our S, the entropy of the conditional state ρ_Q^k of photon A in Eq. (21). However, our measures of coherence and path information emerge naturally from a full information-theoretic analysis and yield more insight into the origins of their complementarity, in particular how the entropy of system Q is distributed among the detectors D_A and D_B (as summarized by Fig. 3).

We end this discussion by noting that if we prepare photons A and B in an imperfectly entangled state, e.g., the q-Bell state [26] $\sqrt{q}\,|h\rangle|v\rangle + \sqrt{1-q}\,|v\rangle|h\rangle$, the coherence and path information we derived are modified from their original forms. For an erasure angle $\theta = \pi/4$, we can write down a simple replacement for Eqs. (30) and (34) that appear in the complementarity relation of Eq. (38). That is,

$$\theta = \frac{\pi}{4}: \quad S(Q:D_B) = 1 - S(q), \tag{43}$$

$$S(Q:D_A|D_B) = S(q), \tag{44}$$

where $S(q)$ replaces S, the entropy of the state (21). The quantity $S(q)$ is defined by the eigenvalues $\lambda_\pm = 1/2 \pm \sqrt{q(1-q)}$. The parameter q controls the initial entanglement between photons A and B and allows us to extract nonzero path information, at the cost of reduced coherence, even when $\theta = \pi/4$. Setting $q = 1/2$ recovers the original result of zero path information and full coherence, $S(q = 1/2) = S = 0$, at this erasure angle.

V. CONCLUSIONS

We prefer to tread lightly when using our results to discuss aspects of quantum theory that have been discussed in a controversial manner since the discussions between Bohr and Einstein concerning these matters [28]. Nevertheless, we believe some statements can be made unequivocally. For example, it is now clear (and has been pointed out repeatedly before), that a quantum system not only carries *both* particle and wave attributes, but that these quantities are manifested in measurement devices in a fluid manner. In particular, the dynamics of the Bell-state quantum eraser, which allows us to give measurements different "meanings" depending on what state preparation we may choose *after* the state determination has taken place, cannot possibly be consistent with a picture of quantum measurement in which the quantum state is irreversibly projected so as to be consistent with the state of the measurement device. The actuality of not only information erasure but the production of alternative outcomes via the retroactive manipulation of state preparation confirms the picture that the wave function after measurement continues to carry amplitudes that are *not* consistent with the state of the measurement device.

That the quantum state can be inconsistent with the state of the measurement device should not come as a surprise to practitioners of quantum information science. After all, the idea of the classical measurement, in which the state of the system to be measured is copied onto the state of the measurement device, cannot carry over to quantum mechanics because of the no-cloning theorem [29,30]. Indeed, the central idea of classical measurement, in which the variation of the system is fully correlated with the variation in the device, cannot be realized for pure quantum states that carry no entropy whatsoever.

Of course, mixed quantum states (pure joint states with a reference state traced out) can carry entropy, and this entropy can be shared with classical measurement devices. This appears to be the case in the construction described here, as the entropy of the system Q is exactly one bit (in the ideal case whose extension is discussed in the note in [18]). If the classical device (here the device D_A) cannot carry information about the state of Q, what information does it reflect? In our view, a classical device's state must be consistent with the state of other classical measurement devices so as to ensure a causally consistent world. Here the information $S(Q:D_A|D_B)$ predicts the outcome of a measurement of the which-path information that would be obtained if a device were placed squarely in the path of the beam. Of course, such a device would record a random outcome (the photon would be found in state ψ_1 half the time), and D_A would perfectly predict this random outcome. Still, neither of these states predicts the state of Q, which, after all, is neither here nor there. We are thus forced to admit that our classical devices do not (and *cannot*) reveal to us the quantum reality underlying our classical world [31]. However, experimental (and theoretical) ingenuity has allowed us to be aware of our classical device's deceptions and has shown us the path to perhaps design even more clever schemes to lift the veil from the underlying quantum reality.

ACKNOWLEDGMENT

Financial support by a Michigan State University fellowship to J.R.G. is gratefully acknowledged.

APPENDIX: INTERFERENCE PATTERNS

Here we derive the spatial probability distribution for photons A that are incident on a screen D_X by modeling the interaction as a von Neumann measurement of the spatial location of photon A. Expanding the spatial states of A in terms of the position basis of the screen yields

$$|\psi_j\rangle_Q = \sum_{x=1}^{n} \psi_j(x)\,|x\rangle_Q, \tag{A1}$$

where $j = 0,1$ labels each slit. The states $|x\rangle$ can be discretized into n distinct locations according to

$$\begin{aligned} |x = 1\rangle &= |100\cdots 0\rangle, \\ |x = 2\rangle &= |010\cdots 0\rangle, \\ &\vdots \\ |x = n\rangle &= |0\cdots 001\rangle, \end{aligned}$$

which denote the location x at which a photon is detected by D_X. Inserting this basis into expression (14) and performing the measurement of Q with D_X (which starts in the initial state $|x = 0\rangle = |0\cdots 0\rangle$), we obtain

$$|\Psi'\rangle_{AD_XBD_B} = \frac{1}{2}\sum_{xmk} i^m \psi_m^k(x)|xx\rangle_{QD_X} \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \tag{A2}$$

where we define the amplitudes $\psi_m^k(x) = \langle x|\psi_m^k\rangle_Q$.

Tracing over A and B in Eq. (A2), we arrive at the (classical) joint density matrix for detector D_B and the screen

$$\rho_{D_X D_B} = \frac{1}{4}\sum_{xmk} \left|\psi_m^k(x)\right|^2 |x\rangle_{D_X}\langle x| \otimes |k\rangle_{D_B}\langle k| = \frac{1}{2}\sum_k \rho_{D_X}^k \otimes |k\rangle_{D_B}\langle k|, \tag{A3}$$

where

$$\rho_{D_X}^k = \sum_x p_k(x)\,|x\rangle_{D_X}\langle x| \tag{A4}$$

are the conditional states of the screen D_X with the corresponding conditional probability distribution

$$p_k(x) = \frac{1}{2}\sum_m \left|\psi_m^k(x)\right|^2. \tag{A5}$$

Tracing out detector D_B from (A3) yields the full density matrix for D_X,

$$\rho_{D_X} = \frac{1}{2}\sum_k \rho_{D_X}^k = \sum_x p(x)\,|x\rangle_{D_X}\langle x|, \tag{A6}$$

where the total probability distribution of the screen is

$$p(x) = \frac{1}{2}\sum_k p_k(x). \tag{A7}$$

It is straightforward to show that the total probability distribution $p(x)$ for the screen is completely incoherent due to the cancellation of the cross terms of the two conditional probabilities. That is,

$$p(x) = \tfrac{1}{2}(|\psi_1(x)|^2 + |\psi_2(x)|^2). \tag{A8}$$

This distribution describes two intensity peaks on the screen corresponding to the pattern obtained from each slit individually. From the data as a whole (i.e., when we do not know the outcome of detector D_B) no interference is observed on the screen.

However, suppose we do know the outcome of the polarization measurement of photon B. For an outcome k, the conditional state of the screen D_X is given by Eq. (A4). To discern the type of interference pattern the probability distribution (A5) of this density matrix describes, we rewrite the conditional probability in terms of the original amplitudes $\psi_j(x) = \langle x|\psi_j\rangle_Q$, which leads to

$$p_k(x) = \tfrac{1}{2}\{|\psi_1(x)|^2 + |\psi_2(x)|^2 + i\,(-)^k \sin 2\theta \times [\psi_1(x)\,\psi_2(x)^* - \psi_1(x)^*\,\psi_2(x)]\}, \tag{A9}$$

where we used $U_{1k}\,U_{0k}^* + U_{1k}^*\,U_{0k} = (-)^k \sin 2\theta$. In general, this expression will describe interference fringes with a visibility that is controlled by the magnitude of the coefficient $\sin 2\theta$ in front of the cross terms. Let us consider two specific cases of the erasure angle θ.

First, $\theta = 0$ corresponds to a measurement of photon B in the linear $|h\rangle$, $|v\rangle$ basis. In this case, expression (A9) reduces

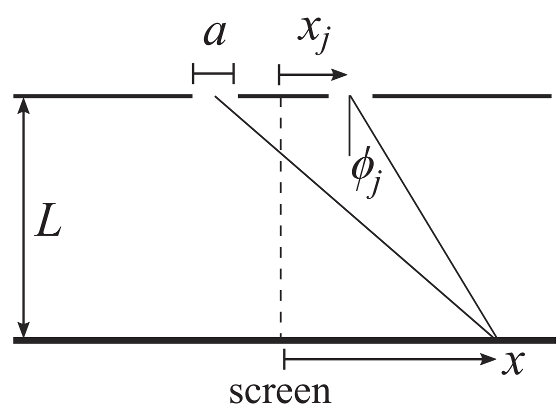

FIG. 5. Geometry of the double-slit apparatus. The slit width is a, the distance from the origin to the center of slit j is x_j, and the distance between slits is $d = |x_2 - x_1|$. The distance from the slits to the screen is L, while the angle from the center of slit j to point x on the screen is $\tan\phi_j = (x - x_j)/L$.

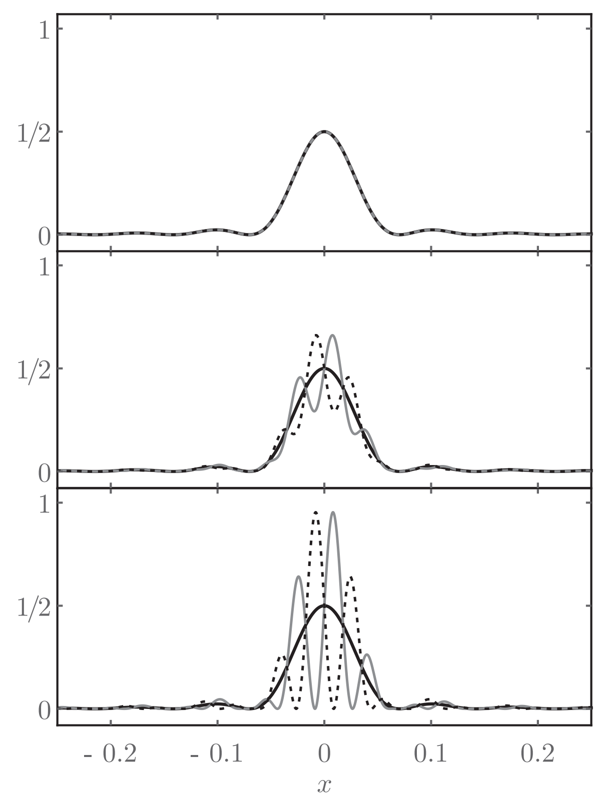

FIG. 6. Conditional probability distributions $p_k(x)$ plotted as a function of the position on the screen x (in meters) and normalized so that the maximum is at 1. The solid gray (black dotted) oscillations describe the interference pattern $p_0(x)$ [$p_1(x)$] of photon A conditional on outcome $k = 0$ ($k = 1$) of detector D_B. The solid black line is the total distribution $p(x)$ and is the single-slit diffraction result. Parameters for this specific case are $a = 10\,\mu$m, $d = 20\,\mu$m, $L = 1$ m, and $\lambda = 702$ nm. Top to bottom: probability distributions for three erasure angles, $\theta = 0$, $\theta = \pi/16$, $\theta = \pi/4$.

to an incoherent sum,

$$\theta = 0: \; p_k(x) = \tfrac{1}{2}(|\psi_1(x)|^2 + |\psi_2(x)|^2), \tag{A10}$$

and describes two intensity peaks on the screen with no interference. In turn, we have full information about the path of photon A. Second, $\theta = \pi/4$ describes a measurement of photon B in the diagonal $|\pm\rangle = \frac{1}{\sqrt{2}}(|h\rangle \pm |v\rangle)$ basis. In this case, expression (A9) becomes a perfectly coherent sum,

$$\theta = \frac{\pi}{4}: \; p_k(x) = \frac{1}{2}|\psi_1(x) - i\,(-)^k \psi_2(x)|^2. \tag{A11}$$

Thus, the effect of tagging has been erased since the standard double-slit diffraction pattern can be observed. In general, given an outcome k for detector D_B, the corresponding state of the screen is $\rho^k_{D_x}$ with probability distribution $p_k(x)$. This leads to fringes with a level of visibility that is determined by the erasure angle θ. The distribution for $k = 0$ is phase shifted relative to $k = 1$, so that depending on the state of D_B, one observes either fringes or antifringes. Therefore, measuring photon B in a basis characterized by the angle θ allows one to tune the visibility of the interference fringes from the standard two-slit diffraction to single-slit diffraction [4].

To explicitly compute the interference patterns, we write the amplitudes of the jth slit for a photon of wavelength λ as [32]

$$\psi_j(x) = \frac{\sin\alpha}{\alpha}\; e^{-2i\,\alpha\, x_j/a}\,, \tag{A12}$$

where $\alpha = \pi a \sin\phi_j/\lambda$. The geometry of the double-slit apparatus (see Fig. 5) is described by the slit width a, the distance x_j to the center of the jth slit, the angle $\phi_j = \tan^{-1}[(x - x_j)/L]$ from the center of slit j to the position x on the screen, and the distance L from the slits to the screen. For a single slit at the origin, the probability to detect a photon at position x on the screen is

$$|\psi_j(x)|^2 = \left|\frac{\sin\alpha}{\alpha}\right|^2, \tag{A13}$$

which is the standard result for single-slit Fraunhofer diffraction. For two slits separated by a distance $d = |x_2 - x_1|$, the amplitudes for each slit are coherently added. In the far-field limit $L \gg d$, we can use the approximation $\phi_j = \phi = \tan^{-1}(x/L)$. Using the amplitudes $\psi_j(x)$ in expression (A9) for $p_k(x)$ leads to the interference patterns shown in Fig. 6. The patterns $p_0(x)$ and $p_1(x)$ are shifted relative to each other on the screen, and the envelope of each pattern is a single-slit diffraction pattern. We show the distributions for three erasure angles: $\theta = 0, \pi/16, \pi/4$. For a measurement in the linear $|h\rangle$, $|v\rangle$ basis ($\theta = 0$), there is no interference on the screen since there is full information about the path of photon A. As the erasure angle increases to $\pi/4$ (a measurement in the diagonal $|\pm\rangle$ basis), the oscillations increase to the level of the usual interference pattern for two-slit diffraction. The solid black line is the total distribution $p(x)$, which is the full data observed in the experiment, and shows no interference. Only by knowing the outcome k of detector D_B can one extract the associated conditional probability distribution $p_k(x)$ from the full distribution $p(x)$.

[1] J. A. Wheeler, The "past" and the "delayed-choice" double-slit experiment, in *Mathematical Foundations of Quantum Theory*, edited by A. R. Marlow (Academic, New York, 1978), pp. 9–48.

[2] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, Experimental realization of Wheeler's delayed-choice gedanken experiment, Science **315**, 966 (2007).

[3] A. Peruzzo, P. Shadbolt, N. Brunner, S. Popescu, and J. L. O'Brien, A quantum delayed-choice experiment, Science **338**, 634 (2012).

[4] F. Kaiser, T. Coudreau, P. Milman, D. B. Ostrowsky, and S. Tanzilli, Entanglement-enabled delayed-choice experiment, Science **338**, 637 (2012).

[5] R. Ionicioiu and D. R. Terno, Proposal for a Quantum Delayed-Choice Experiment, Phys. Rev. Lett. **107**, 230406 (2011).

[6] T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger, Complementarity and the Quantum Eraser, Phys. Rev. Lett. **75**, 3034 (1995).

[7] Y. H. Kim, R. Yu, S. P. Kulik, Y. H. Shih, and M. O. Scully, Delayed "Choice" Quantum Eraser, Phys. Rev. Lett. **84**, 1 (2000).

[8] X.-s. Ma, J. Kofler, and A. Zeilinger, Delayed-choice gedanken experiments and their realizations, Rev. Mod. Phys. **88**, 015005 (2016).

[9] D. M. Greenberger and A. Yasin, Simultaneous wave and particle knowledge in a neutron interferometer, Phys. Lett. A **128**, 391 (1988).

[10] B.-G. Englert, Fringe Visibility and Which-Way Information: An Inequality, Phys. Rev. Lett. **77**, 2154 (1996).

[11] P. J. Coles, J. Kaniewski, and S. Wehner, Equivalence of wave-particle duality to entropic uncertainty, Nat. Commun. **5**, 5814 (2014).

[12] R. M. Angelo and A. D. Ribeiro, Wave-particle duality: An information-based approach, Found. Phys. **45**, 1407 (2015).

[13] E. Bagan, J. A. Bergou, S. S. Cottrell, and M. Hillery, Relations Between Coherence and Path Information, Phys. Rev. Lett. **116**, 160406 (2016).

[14] N. Bohr, Discussions with Einstein on epistemological problems in atomic physics, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, N.J., 1984), pp. 9–49.

[15] Y. Aharonov and M. S. Zubairy, Time and the quantum: Erasing the past and impacting the future, Science **307**, 875 (2005).

[16] R. Garisto and L. Hardy, Entanglement of projection and a new class of quantum erasers, Phys. Rev. A **60**, 827 (1999).

[17] S. P. Walborn, M. O. Terra Cunha, S. Pádua, and C. H. Monken, Double-slit quantum eraser, Phys. Rev. A **65**, 033818 (2002).

[18] Note that an obvious extension of the present construction is to allow for arbitrary entangled photon pairs, such as the q-Bell states [26] in which a parameter q interpolates between product states ($q = 0$ or 1) and fully entangled states ($q = 1/2$). We mention briefly the effects of $q \neq 1/2$ at the end of Sec. IV.

[19] Note that the extension of this framework to allow for N-path devices is straightforward.

[20] R. C. Jones, A New Calculus for the Treatment of Optical Systems, I. Description and Discussion of the Calculus, J. Opt. Soc. Am. **31**, 488 (1941).

[21] E. Collett, *Field Guide to Polarization* (SPIE Press, Bellingham, WA, 2005).

[22] N. J. Cerf and C. Adami, Information theory of quantum entanglement and measurement, Phys. D (Amsterdam, Neth.) **120**, 62 (1998).

[23] N. J. Cerf and C. Adami, Negative Entropy and Information in Quantum Mechanics, Phys. Rev. Lett. **79**, 5194 (1997).

[24] M. O. Scully, B.-G. Englert, and H. Walther, Quantum optical tests of complementarity, Nature (London) **351**, 111 (1991).

[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, 2000).

[26] C. Adami and N. J. Cerf, von Neumann capacity of noisy quantum channels, Phys. Rev. A **56**, 3470 (1997).

[27] W. K. Wootters W. H. Zurek, Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle, Phys. Rev. D **19**, 473 (1979).

[28] J. A. Wheeler and W. H. Zurek (eds.), *Quantum Theory and Measurement* (Princeton University Press, Princeton, N.J., 1984).

[29] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature (London) **299**, 802 (1982).

[30] D. Dieks, Communication by EPR devices, Phys. Lett. A **92**, 271 (1982).

[31] C. Adami and N. J. Cerf, What information theory can tell us about quantum reality, in *Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science Vol. 1509 (Springer, Berlin, 1999), p. 258.

[32] D. Arsenović, M. Bozić, and L. Vusković, Is the classical law of the addition of probabilities violated in quantum interference? J. Opt. B **4**, S358 (2002).