

# Evaluation of Counterfactuality in Counterfactual Communication Protocols

D. R. M. Arvidsson-Shukur and C. H. W. Barnes

*Cavendish Laboratory, Department of Physics, University of Cambridge, CB3 0HE, Cambridge, United Kingdom*

A. N. O. Gottfries

*Faculty of Economics, University of Cambridge, CB3 9DD, Cambridge, United Kingdom*

(Dated: February 26, 2019)

We provide an in-depth investigation of parameter estimation in Nested Mach-Zehnder interferometers (NMZIs) using two information measures: the Fisher information and the Shannon mutual information. Protocols for counterfactual communication (CFC) have, so far, been based on two different definitions of counterfactuality. In particular, some schemes are based on NMZI devices, and have recently been subject to criticism. We provide a methodology for evaluating the counterfactuality of these protocols, based on an information theoretical framework. More specifically, we make the assumption that any realistic quantum channel in MZI structures will have some weak uncontrolled interaction. We then use the Fisher information of this interaction to measure counterfactual violations. The measure is used to evaluate the suggested counterfactual communication protocol of Salih et al. [1]. The protocol of Arvidsson-Shukur and Barnes [2], based on a different definition, is evaluated with a probability measure. Our results show that the definition of Arvidsson-Shukur and Barnes is satisfied by their scheme, whilst that of Salih et al. is only satisfied by perfect quantum channels. For realistic devices the latter protocol does not achieve its objective.

**PACS numbers:** 03.65.Ta, 03.67.Hk, 03.67.Ac

## I. INTRODUCTION

During the past one and a half centuries, the study of interferometers has resulted in some of the most profound discoveries in physics. From the Michelson–Morley experiment [3], which established the speed of light as a constant, to Hardy’s Paradox [4], which elegantly demonstrates the non-local behaviour of the fundamentals of quantum physics, interferometers have played a pivotal role. This is perhaps more evident today than ever before, considering the recent discoveries of gravitational waves made with two power-recycled Michelson interferometers.[5]

Studies with optical quantum interferometers have shown great promise, not only for the detection of novel physics, but also for external field detection and external parameter estimation.[6–16] A common example of this is phase estimation. By letting optical quantum states interfere with a medium inside an interferometer it can be easier to establish the nature of a phase-shift caused by the medium, than with a *direct* interaction.[10, 17–19]

A framework for studying phase estimation in interferometers has been developed by Bahder et al. [13, 18, 19]. This framework uses the Shannon mutual information and the classical Fisher information as measures of the phase estimating capacity of different interferometers and input states. The Shannon mutual information provides a measure of the suitability of a specific experiment for the estimation of a phase given some known or unknown phase probability distribution. The classical Fisher information, on the other hand, provides a measure of how well a specific—but unknown—phase-shift in the interferometer can be estimated from the outcome events of the specific interferometry experiment.

Another area of physics that has been developed en-

tirely via the study of interferometers, is that of counterfactual phenomena:

*Counterfactuals* – that is, things that might have happened, although they did not in fact happen.

– Roger Penrose in ‘Shadows of the Mind’ [20]

The counterfactual phenomenon of interaction-free measurements was originally discovered by Elitzur and Vaidman in their seminal paper on quantum bomb diffusal.[21] They showed how a Mach-Zehnder interferometer (MZI) could be used in order to query whether or not an absorbing object (e.g. a bomb) was or was not present in the lower interferometer arm. The novelty of their setup was that the photons propagating through the interferometer sometimes allowed for the answering of the query without interacting with the object in question, i.e. counterfactually.[20] Kwiat et al. then showed how the efficiency of this scheme could be taken arbitrarily close to unity by utilising a chain of several MZIs.[22, 23]

During the last decade there has been further investigations of counterfactual schemes. Many of these are based on the use of so called “nested” Mach-Zehnder interferometers (NMZIs). There have been suggestions that quantum computation [24], direct communication [1, 25] or transmission of quantum states [26] can be conducted without the interrogating particle ever entering the quantum computer in the former case or interacting with the information transmitter in the two latter. However, these schemes have been under intense debate.[1, 26–34] The criticism resulted in the development of another counterfactual communication (CFC) scheme.[2] The definition of counterfactuality in the former schemes does not allow any particles to travel be-

tween receiver and transmitter. Alternatively, the definition in the latter allows particles to travel from receiver to transmitter but not vice versa (i.e. particles are only allowed to propagate in the opposite direction to the message). The essence of the criticism of the former schemes is based on an investigation of the *weak trace* [27] that the interrogating particle leaves within the inner part of a NMZI. Essentially, the inner part of the NMZIs have to be inaccessible to the interrogating particle. However, if a weak interaction is present in those parts, the effect of that is of the same order or stronger than if the particle had freely propagated through the interaction.[29, 30] The common rebuttal is that this criticism is invalid as a weak interaction alters the perfect interference of the interferometer used in the suggested schemes.

Whilst the weak trace is an interesting concept, we feel that an argument based on information theoretical principles is desirable to bring clarity to the subtleties of the counterfactual protocols. Furthermore, owing to the intense discussion regarding NMZIs, we also see the need of a thorough investigation of parameter estimation in these structures.

In this paper we adapt the information measures of phase estimation such that they can be used for parameter estimations in Mach-Zehnder interferometer structures. We firstly give an outline of the theoretical framework of the paper. Secondly, we provide a detailed analysis of the wavefunction evolution through the NMZI devices. We observe how the classical Fisher information and Shannon mutual information changes between experiments with NMZIs depending on where in the interferometer an interacting medium is placed. Furthermore, we evaluate the two different definitions of CFC, with suitable operational (and interpretation independent) measures of their respective violations. The underlying argument of our work, is that absolutely lossless and pure spatial transmission of quantum particles is not attainable. Thus, a theory that relies on such perfect quantum channels is as valid (read: invalid) as a thermodynamic proof only valid at 0 K. We provide a model of realistic devices that contain a weak un-controlled polarization rotation. This serves the purpose of mimicking real quantum evolutions. Based on this, we can evaluate the counterfactuality of the communication schemes according to their respective definitions. We see that some “counterfactual” protocols violate their definition of counterfactuality more than a free-space propagation from a transmitter to a receiver. We can thus rule out the counterfactuality of these schemes.

## II. INFORMATION MEASURES

The basis of this work is the knowledge of how the single particle wavefunction (and thus the probability density distribution) evolves through the devices we wish to investigate. The wavefunction evolution is provided by the calculation of unitary operations on some ini-

tial quantum state:  $|\psi_{in}\rangle \rightarrow \hat{U}|\psi_{in}\rangle$ . Throughout the evolution, information will be encoded in the probability density distribution via the interactions that act on the wavefunction. Whilst the extension to multi-photon states is straightforward, we wish to keep this study in line with suggested counterfactual schemes [1, 2, 24, 26], and restrict our work to single-photon inputs.

It is often nonsensical to ask where a quantum particle has been present between two observations. However, along the evolution of the quantum state, from input to output, one can introduce an interaction that results in parts of the wavefunction occupying a quantum state that only is made available via this interaction. The wavefunction will carry some *information* about the nature of that interaction and it is possible to interpret parts of the probability density—that occupy states only made available via this interaction—as having had a *presence* at the area where the interaction was located. The probability outputs, at the end of the quantum evolution, allow for the estimation of the interaction parameters. The effectiveness of such an estimation, for a given quantum device, can be evaluated with the two information measures outlined in the following subsection.

### II.I. Shannon Mutual Information & Fisher Information

Consider an experiment, given an input state,  $\psi_{in}$ , and a parameter,  $\theta$ , that sets some interaction. We can calculate the Shannon mutual information,  $H(\theta : M)$ , between the parameter  $\theta$  and the measurement outcomes  $M = \{M_i\}$ , where  $M_i$  represents an event that occurred in the  $i^{\text{th}}$  detector location of the total spatial Hilbert space,  $\mathcal{H}$ :

$$H(\theta : M) = \sum_{i \in \mathcal{H}} \int_{\theta_{min}}^{\theta_{max}} d\theta P(M_i|\theta, \psi_{in})p(\theta) \times \log_2 \left[ \frac{P(M_i|\theta, \psi_{in})}{\bar{P}(M_i|\psi_{in})} \right], \quad (1)$$

where  $\bar{P}(M_i|\psi_{in}) = \int_{\theta_{min}}^{\theta_{max}} d\theta' P(M_i|\theta', \psi_{in})p(\theta')$ ,  $P(M_i|\theta, \psi_{in})$  is the probability of  $M_i$  for some specific  $\theta$  and  $\psi_{in}$ , and  $p(\theta)$  is the *a priori* probability distribution of the parameter  $\theta$  [18, 35]. The Shannon mutual information provides a measure of how much information about  $\theta$  that can be obtained through knowledge of the measurement outcomes of a specific experiment. A large value of  $H(\theta : M)$  indicates a good device for parameter estimations of an unknown parameter  $\theta$ .

The Shannon mutual information takes into account a prior distribution of  $\theta$ :  $p(\theta)$ . However, if the value of  $\theta$  is fixed but unknown, one might ask oneself how much information, on average, a single use of a specific interferometer yields about  $\theta$ . This quantity of information is

given by the classical Fisher information [13, 35]:

$$F(\theta) = \sum_{i \in \mathcal{H}} \frac{1}{P(M_i|\theta, \psi_{in})} \left[ \frac{\partial}{\partial \theta} P(M_i|\theta, \psi_{in}) \right]^2. \quad (2)$$

In the Cramér-Rao inequality  $F(\theta)$  sets a lower bound on the variance of the estimator of  $\theta$ ,  $\theta_e$ , obtained in a specific experiment:

$$\text{Var}(\theta_e) \geq \frac{1}{F(\theta)}. \quad (3)$$

### II.II. Fisher Information as a Measure of Presence in Optical Circuits

Many optical quantum interferometers do not involve polarization operations. However, any “real” experiment with single photons will naturally include some polarization operations—owing to, for example, material impurities and systematic errors in the experimental setup. We mimic the inevitable imperfect nature of *real* quantum channels by introducing single polarization interactions somewhere in the optical circuits.<sup>1</sup> This represents an interaction with Bob’s laboratory rather than a generic noise model. We call this interaction the “tagging” of the wavefunction. By introducing the polarization degree of freedom, we can use the Fisher information to estimate the parameters associated with the interaction. As has been described above, a tagged part of the wavefunction can be considered to have previously existed at the location of the polarization interaction. In this subsection we show that, given access to all the outcome possibilities, the classical Fisher information, in the interferometers studied in this work, is always proportional to the integrated probability density distribution that has evolved through the interaction in the Schrödinger picture.

Firstly, we define an optical input vector,  $\mathbf{a}$ , of length  $2n$ , which evolves into an output vector,  $\mathbf{b}$ . The  $2n$  levels correspond to the  $n$  different optical paths of the device, each of which can exist in one of the two polarization states. We choose the order of the vector elements so that the first  $n$  entries have the polarization of the initial input state, and so that the following  $n$  entries have orthogonal polarization.

We can describe the evolution of the input state,  $\mathbf{a}$ , through the interferometer by a scattering matrix,  $\mathcal{S}$ , in terms of three operations:

$$\mathcal{S}\mathbf{a} \equiv (\hat{V} \cdot \hat{f}^{(k)}(\theta) \cdot \hat{U})\mathbf{a} \equiv \mathbf{b}^{(k)}, \quad (4)$$

<sup>1</sup> Weak unwanted polarization interactions can occur in all parts of a realistic device. However, our setup can be considered as one where the experimentalist, Alice, is allowed to correct for unwanted polarization rotations from all parts of the device, except for those that occur in a location controlled by Bob (central rectangle in Fig. 1). Hence, we consider one unwanted polarization rotation, due to Bob’s laboratory.

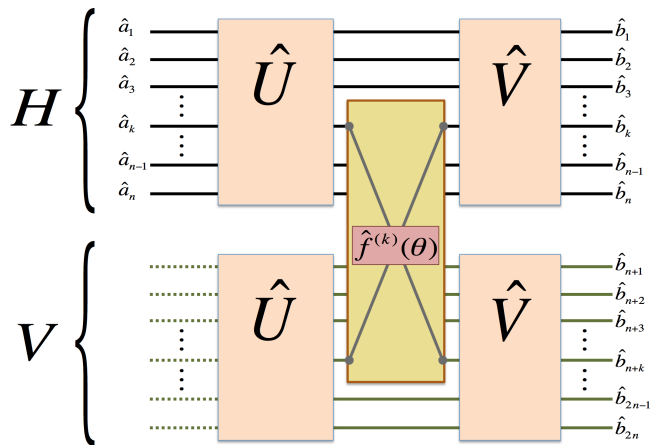


FIG. 1. (color online) Sketch of an optical circuit of the form of Eq. 4, which is described in the text.

where  $\hat{U}$  and  $\hat{V}$  are the unitary operators of the evolution up to and after the tagging polarization rotation respectively, and  $\hat{f}^{(k)}(\theta)$  is the unitary operator that describes the *single* polarization rotation at the specific spatial path of  $k$  (where  $1 \leq k \leq n$ ) by an angle  $\theta$ . A sketch of the optical circuit of  $\mathcal{S}$  is given in Fig. 1. Note that  $\hat{U}$  and  $\hat{V}$  acts solely on the spatial degree of freedom and do not manipulate the polarization of the wavefunction.

We can define the wavefunction after  $\hat{U}$  has been applied as  $\mathbf{c} \equiv \hat{U}\mathbf{a}$ , with:

$$c_i = \sum_{j=1}^{2n} U_{i,j} a_j. \quad (5)$$

The rotation matrix is then applied to this state. It rotates the quantum state between two of the vector levels,  $k$  and  $k'$ . It can be represented by the following real matrix:

$$\hat{f}^{(k)}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & f_{k,k}(\theta) & \cdots & f_{k,k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & f_{k',k}(\theta) & \cdots & f_{k',k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \quad (6)$$

where  $f_{k,k}(\theta) = \sqrt{1 - (f_{k,k'}(\theta))^2}$ ,  $f_{k',k}(\theta) = -f_{k,k'}(\theta)$  and  $f_{k',k'}(\theta) = f_{k,k}(\theta)$ .

We further arrange the vector entries such that  $l$  and  $l+n$  correspond to the same spatial location for  $1 \leq l \leq n$ . This means that  $k' = k+n$  in Eq. 6 and  $c_i = 0$  for  $i > n$  in Eq. 5. We express the quantum state after the polarization interaction as  $\mathbf{d}^{(k)}(\theta) \equiv \hat{f}^{(k)}(\theta)\mathbf{c}$ , with:

$$d_i^{(k)}(\theta) = \sum_{j=1}^{2n} f_{i,j}^{(k)}(\theta) c_j. \quad (7)$$

We note that the only components of  $\mathbf{d}^{(k)}$  that depend on  $\theta$  are  $d_{i=k}^{(k)} = f_{k,k}(\theta)c_k$  and  $d_{i=k+n}^{(k)} = f_{k+n,k}(\theta)c_k$ .

Finally, we can apply the last unitary evolution  $\hat{V}$ . Following the steps above, we express the output vector as  $\mathbf{b}^{(k)}(\theta) \equiv \hat{V}\mathbf{d}^{(k)}(\theta)$ , with:

$$b_i^{(k)}(\theta) \equiv \beta_i^{(k)} + b_{i,\theta}^{(k)}(\theta) = \sum_{j=1}^{2n} V_{i,j} d_j^{(k)}(\theta), \quad (8)$$

where the  $\theta$ -dependence of  $b_i^{(k)}(\theta)$  is encapsulated in  $b_{i,\theta}^{(k)}(\theta)$  and a corresponding term, independent of  $\theta$ , is defined as  $\beta_i^{(k)}$ .

The probability of measuring the single-photon in the  $i^{\text{th}}$  output port, is then given by:

$$P_i^{(k)}(\theta) = |\beta_i^{(k)} + b_{i,\theta}^{(k)}(\theta)|^2. \quad (9)$$

This can be re-expressed as

$$P_i^{(k)}(\theta) \equiv |\beta_i^{(k)} + b'_i{}^{(k)} f_{j,k}^{(k)}(\theta)|^2, \quad (10)$$

where  $j = k$  if  $i \leq n$  and  $j = k + n$  if  $i > n$ . For convenience we now drop the  $(k)$  superscript.

Using Eq. 2, the individual Fisher information components can be expressed as:

$$\begin{aligned} F_i &= \frac{1}{|\beta_i + b'_i f_{j,k}(\theta)|^2} \left[ \frac{\partial}{\partial \theta} |\beta_i + b'_i f_{j,k}(\theta)|^2 \right]^2 \\ &= \frac{1}{(\beta_i + b'_i f_{j,k}(\theta))(\beta_i^* + b_i'^* f_{j,k}(\theta))} \\ &\quad \times \left[ \frac{\partial}{\partial \theta} (\beta_i + b'_i f_{j,k}(\theta))(\beta_i^* + b_i'^* f_{j,k}(\theta)) \right]^2 \end{aligned} \quad (11)$$

This expression can be simplified by expressing the coefficients as  $\beta_i \equiv |\beta_i|e^{i\phi_i}$  and  $b'_i \equiv |b'_i|e^{i\phi_{i,\theta}}$  and defining  $\Phi_i \equiv \phi_i - \phi_{i,\theta}$ :

$$\begin{aligned} F_i &= \frac{(\cos(\Phi_i)|\beta_i| + |b'_i|f_{j,k}(\theta))^2}{(|\beta_i|^2 + |b'_i|^2 f_{j,k}^2(\theta) + 2\cos(\Phi_i)|\beta_i||b'_i|f_{j,k}(\theta))} \\ &\quad \times 4|b'_i|^2 \left( \frac{\partial}{\partial \theta} f_{j,k}(\theta) \right)^2. \end{aligned} \quad (12)$$

We notice that if the phase difference  $\Phi_i$  is a multiple of  $\pi$ , the expression simplifies to:

$$F_i = 4|b'_i|^2 \left( \frac{\partial}{\partial \theta} f_{j,k}(\theta) \right)^2. \quad (13)$$

This phase criterion is satisfied for all  $i$  if the phases of all the spatial components,  $i \leq n$ , of the input state are the same and  $\mathcal{S}$  is real (e.g. the optical setup only contains beam-splitters that can be represented by real operators).

It is also satisfied by all the optical setups considered in Refs. [1, 2, 21–23, 27].

For the quantum optical setups of interest in this paper, we can thus express the classical Fisher information (Eq. 2) as:

$$F(\theta) = \sum_{i=1}^{2n} 4 \left[ \frac{\partial}{\partial \theta} |b_{i,\theta}(\theta)| \right]^2. \quad (14)$$

This can be re-expressed as:

$$\begin{aligned} F(\theta) &= \sum_{i=1}^n 4 \left[ \frac{\partial}{\partial \theta} |V_{i,k} d_k(\theta)| \right]^2 \\ &\quad + \sum_{i=n+1}^{2n} 4 \left[ \frac{\partial}{\partial \theta} |V_{i,k+n} d_{k+n}(\theta)| \right]^2 \\ &= \sum_{i=1}^n 4 |V_{i,k}|^2 \left[ \frac{\partial}{\partial \theta} |d_k(\theta)| \right]^2 \\ &\quad + \sum_{i=n+1}^{2n} 4 |V_{i,k+n}|^2 \left[ \frac{\partial}{\partial \theta} |d_{k+n}(\theta)| \right]^2 \end{aligned} \quad (15)$$

As  $\hat{U}$  and  $\hat{V}$  do not manipulate the polarization of the wavefunction, the symmetry of the matrix  $\hat{V}$  is such that  $V_{i,k} = V_{i+n,k+n}$ . It also implies that  $V_{i,k} = 0$  for  $i > n$  and  $V_{i,k+n} = 0$  for  $i \leq n$ . By assuming a real  $\mathcal{S}$  and defining a suitable reference-point for the global input phase we simplify our expression further:

$$\begin{aligned} F(\theta) &= \sum_{i=1}^{2n} 4 |V_{i,k}|^2 \left[ \frac{\partial}{\partial \theta} |d_k(\theta)| \right]^2 \\ &\quad + \sum_{i=1}^{2n} 4 |V_{i,k}|^2 \left[ \frac{\partial}{\partial \theta} |d_{k+n}(\theta)| \right]^2. \end{aligned} \quad (16)$$

We sum the squared entries in the column of our unitary matrix to unity, and the expression simplifies to:

$$F(\theta) = 4 \left[ \frac{\partial}{\partial \theta} |d_k(\theta)| \right]^2 + 4 \left[ \frac{\partial}{\partial \theta} |d_{k+n}(\theta)| \right]^2. \quad (17)$$

At this stage we note that the total Fisher information of the device does not contain any dependence on the unitary operation  $\hat{V}$ .

We continue by substituting the expressions of  $d_k$  and  $d_{k+n}$  from above (Eq. 7) to obtain a final expression of the Fisher information:

$$\begin{aligned} F^{(k)}(\theta) &= 4|c_k|^2 \left( \left[ \frac{\partial}{\partial \theta} (f_{k,k}(\theta)) \right]^2 + \left[ \frac{\partial}{\partial \theta} (f_{k+n,k}(\theta)) \right]^2 \right) \\ &= 4|c_k|^2 \frac{\left( \frac{\partial}{\partial \theta} f_{k+n,k}(\theta) \right)^2}{(f_{k,k}(\theta))^2}, \end{aligned} \quad (18)$$

where we briefly re-introduce the  $(k)$  superscript.

To conclude this section, we make the observation that the Fisher information, Eq. 18, is proportional to  $|c_k|^2$ .

$|c_k|^2$  is the probability of observing the photon in the  $k^{\text{th}}$  spatial state if a detector had been placed at the location of the tagging interaction. Naïvely the extent to which the wavefunction spreads into a spatial location, according to the time-dependent Schrödinger equation, could be interpreted as a measure of how much the particle has been *present* there. However, owing to the nature of quantum mechanics, it is philosophically problematic to specify what the physical meaning of the wavefunction between measurements is.<sup>2</sup> Nevertheless, in the circuits studied above, the tagging mechanism, which rotates the initial polarization of the photon into a superposition state, is the only polarization component of the interferometer. Hence, it is arguably less naïve to consider an output photon in an altered polarization state to have had a past that has included the passage through the tagging part of the interferometer. Even if the introduced tagging polarization rotations are vanishingly small and do not affect the specific interferometer significantly, Eq. 18 shows that the information content, which travels from the polarization rotator to the output ports, is weighted by the square of the integrated wavefunction at the location of the interaction. Hence, the Fisher information is arguably a good measure of the *extent* to which a particle can be considered to previously have had a presence at the tagging location.

We finish this section with a note regarding an extension of the above theory, to include other degrees of freedom. If an interferometer contains polarization rotations (such as the Michelson-based device in Ref. [1]), these rotations can simply be included in  $\hat{U}$  and  $\hat{V}$ . The single rotation, that we calculate the Fisher information with respect to, should then be changed to an alternative degree of freedom. With the corresponding alterations of the measure,  $\hat{f}^{(k)}(\theta)$  can, for example, be taken to be a weak rotation of the photon's internal orbital angular momentum. The beauty of this analysis is that Eq. 18 will still be valid. However, as the interferometers considered in this paper are all ideally non-polarizing, we conduct our analysis using a weak disturbances on the polarization.

### III. MEASURES OF COUNTERFACTUAL VIOLATION

In this section we develop measures for the extent a process violates counterfactuality. These measures will be used in Section V and VI to investigate counterfactual violations of the CFC protocols proposed by Salih et al. [1] and Arvidsson-Shukur and Barnes [2].

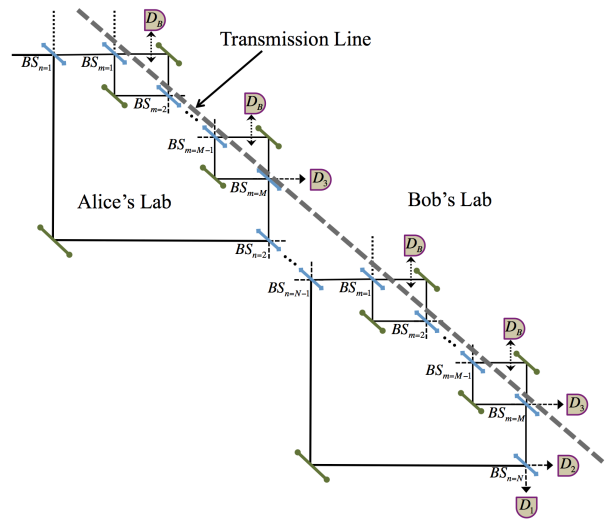


FIG. 2. (color online) The chained nested Mach-Zehnder interferometer suggested for CFC in Ref. [1]. Alice inputs a photon in the upper left path and Bob has the choice of introducing detectors in his part of the device. His choice governs the statistics of the final output detections at  $D_1$  and  $D_2$ .

#### III.I. Two CFC Definitions

There are two main schemes for CFC, based on different definitions of the concept. One was developed in 2013 by Salih et al.[1] (See Fig. 2). We refer to the definition of CFC in that protocol as the Type I definition. Another scheme was developed by Arvidsson-Shukur and Barnes in 2016 [2] (see Fig. 3) and we refer to its definitions as the Type II definition.

In both schemes Bob transmits a message to Alice. The schemes respective bit-transmissions are initiated by Alice sending a single particle into the upper left input path of the devices. Bob can then choose to transmit a 0 bit or a 1 bit by keeping his laboratory free or inserting absorbing detectors respectively. In the Type I protocol (Fig. 2), the quantum Zeno effect [37, 38] is used for both bit-processes, such that the particle ends up at Alice's detectors  $D_1$  (0 bit) or  $D_2$  (1 bit) without ever having crossed the transmission line between Alice and Bob. The Type II protocol (Fig. 3), on the other hand, only utilise the quantum Zeno effect for the 1 bit process. In this process the particle enters the transmission line and returns to Alice, but it never enters Bob's laboratory. In the 0 bit process the particle travels from Alice's laboratory into the transmission line. It then evolves into Bob's laboratory. The protocols are described in further detail in Section V and VI of this paper. The differences in the counterfactuality definitions are summarised in Table I.

The counterfactuality of an interaction-free process, naturally depends on the boundaries of the spatial extent of the respective "laboratories" of the participants in the protocol. Both the Type I and II definitions state that for a process in Bob's laboratory to be counterfactual

<sup>2</sup> This is discussed at length in Wheeler's "The 'Past' and the 'Delayed-Choice' Double-Slit Experiment", which has been reproduced in ref. [36]

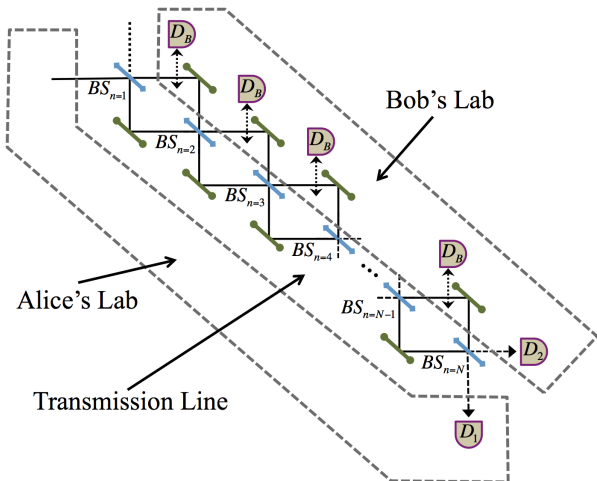


FIG. 3. (color online) The chained Mach-Zehnder interferometer as used in the CFC scheme of Ref. [2]. As in Fig. 2, Alice inputs a photon in the upper left path and Bob governs the statistics of the final output detections at  $D_1$  and  $D_2$  by inserting or not inserting detectors respectively.

TABLE I. The counterfactual nature of the Type I [1] and the Type II [2] communication schemes. In both schemes a message is transmitted from a transmitter, Bob, to a receiver, Alice.

Scheme \ Bit	Logical 0:	Logical 1:
Type I:	No particles cross the transmission line between Alice and Bob.	
Type II:	Particles propagate from Alice to Bob via the transmission line.	Particles propagate from Alice to the transmission line and back again.

with respect to Alice, it is essential that particles should never propagate from Bob to Alice (such that parts of a wavefunction that interacts with Bob's laboratory will have a vanishing probability to end up in Alice's).

As mentioned above, the Type I and II schemes utilise the quantum Zeno effect, triggered by absorbing detectors in Bob's laboratory, to produce their respective logical 1 bit-values in a counterfactual manner. We are not aware of any works disputing the counterfactual nature of these processes. That leaves us with the task of evaluating the Type I and II logical 0 processes.

### III.II. The Measures

The Type I definition should forbid particles to propagate from Alice to Bob and vice versa. Following the discussion in the section above, a good measure of the violation of such a process could be based on the Fisher information encoded at Bob's laboratory, in a particle originating from Alice.

In order to evaluate the "strength" of a violation of a Type I logical 0 process (caused by a tagging interac-

tion as discussed above) we need a Fisher information benchmark. In this work we benchmark with respect to the Fisher information of a free-space evolution,  $F_{free}$ , subject to the same tagging interaction as the circuits of interest. This can, for example, be the scenario of Bob directly sending a photon to Alice (see Fig. 4)—a clearly non-counterfactual scenario. We can now define our new measure for the violation of Type I counterfactuality. We call the measure the Type I counterfactual violation strength:

$$\mathcal{D} := \frac{F}{F_{free}}. \quad (19)$$

This quantity can effectively be thought of as the Fisher information encoded in a particle originating from Alice, owing to an interaction at Bob's laboratory, as a fraction of the Fisher information of a free-space interaction. A value of  $\mathcal{D} = 0$  corresponds to no wavefunction interacting with Bob's laboratory; and a value of  $\mathcal{D} \geq 1$  corresponds to an interaction stronger than or equal to that of a free-space interaction. Values of the order of unity or bigger are convicted of fully violating Type I counterfactuality.

When it comes to evaluating the logical 0 in Type II protocols, we need a different measure than Eq. 19. This is because this scheme allows Alice's particles to be encoded by Bob, as long as they do not return to Alice.[2] The probability of detection in Alice's laboratory in this process is null for perfect quantum channels ( $\theta = 0$ ). Thus, a reasonable measure of a counterfactual violation would be the total probability of a particle returning to Alice as a result of a non-collapsing interaction (i.e.  $\theta \neq 0$ ) in Bob's laboratory. Let  $M_{j \in \mathcal{H}_A}$  denote the measurement outcomes triggered in the Hilbert space of states within the spatial extent of  $\mathcal{A}$ . Additionally, let  $M_{j' \notin \mathcal{H}_A}$  denote the negative measurement that indicate all outcome states outside the spatial extent of  $\mathcal{A}$ . Our new Type II measure can then be expressed as:

$$P_A := \sum_{j \in \mathcal{H}_A} P(M_j | \theta, \psi_{in}) = 1 - P(M_{j' \notin \mathcal{H}_A} | \theta, \psi_{in}). \quad (20)$$

This measure can be interpreted in a way such that  $P_A = 1$  corresponds to a full counterfactual violation of the Type II logical 0 and  $P_A = 0$  corresponds to perfect counterfactuality.

Moreover, even though the probability to trigger a detection in Alice's laboratory can be very small ( $P_A \approx 0$ ), small probabilities can generate large Fisher informations. We thus introduce the spatially restricted Fisher information,  $F_A$ .  $F_A$  is a measure of the sum of the individual components of the classical Fisher information

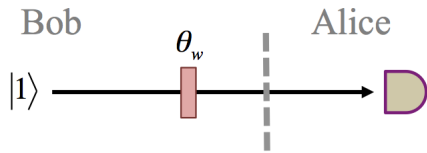


FIG. 4. (color online) A single-photon state is incident on a weak polarization rotator. The photon is then measured in its number state and polarization state.

(Eq. 2), as experienced by an observer  $\mathbf{A}$ . We define:

$$F_{\mathbf{A}}(\theta) := \sum_{j \in \mathcal{H}_{\mathbf{A}}} \frac{1}{P(M_j|\theta, \psi_{in})} \left[ \frac{\partial}{\partial \theta} P(M_j|\theta, \psi_{in}) \right]^2 + \frac{1}{1 - P_{\mathbf{A}}} \left[ \frac{\partial}{\partial \theta} P_{\mathbf{A}} \right]^2. \quad (21)$$

The first term corresponds to the summation of the Fisher information components of particle detections by  $\mathbf{A}$ , whilst the second term corresponds to the Fisher information component of negative measurements by  $\mathbf{A}$ , i.e. when no particle is detected by  $\mathbf{A}$ . Interestingly, in a Type II scheme, Alice can still obtain a large Fisher information of Bob's  $\theta$ , even if counterfactuality is only violated weakly by  $\theta$ .

We are now in the position to evaluate Type I and Type II CFC protocols. However, first we conduct an elaborate study of parameter estimation in NMZIs. This is done in the next section, where we utilise the information measures from Section II.I and extend the works of Bahder *et al.* [13, 18] to NMZI structures. In Section V and VI, we then expand the analysis in order to evaluate the counterfactuality of Type I and II CFC protocols.

## IV. INFORMATION IN NESTED MZIS

### IV.I. Free Space Interaction

As a reference scenario for the optical circuits discussed in following sections of this paper, we provide the simplest of examples of perturbations caused by a polarization rotator. We consider a single-photon state. It has a polarization degree of freedom, and propagates in a straight line. It interacts with a polarizing medium—the rotator—shortly after which it is measured. (See Fig. 4).

The quantum interaction of the wavepacket with the rotator results in a rotation of the polarization set by the parameter  $\theta = \theta_w$ . We can calculate the Fisher information,  $F(\theta_w)$ , and the Shannon mutual information,  $H(\theta_w : M)$ , of the parameter  $\theta_w$  and the measurement outcomes.

We introduce the creation operators  $\hat{a}_H^\dagger$  and  $\hat{a}_V^\dagger$ , which create a single photon in a horizontal and vertical state respectively. Our polarization axes are defined such that

the input state can be written as:

$$\hat{a}_H^\dagger |0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (22)$$

For the case of a single polarizing rotator, the scattering matrix (Eq. 4) takes the form:

$$S = \hat{f}^{(1)}(\theta_w) = \begin{pmatrix} \sqrt{1 - \theta_w^2} & \theta_w \\ -\theta_w & \sqrt{1 - \theta_w^2} \end{pmatrix}. \quad (23)$$

The polarization rotations of the different optical circuits studied in this paper will all be in the form of Eq. 23.

The detector in Fig. 4 measures in the basis:  $|n_H, n_V\rangle$ , where  $n_H$  and  $n_V$  are the respective photon numbers of horizontal and vertical polarization at the output. The output probabilities are given by:

$$P(n_H = 1|\theta_w) = 1 - \theta_w^2, \quad (24)$$

$$P(n_V = 1|\theta_w) = \theta_w^2, \quad (25)$$

where we have adopted a notation such that the statement  $n_s = 1$  implicitly assumes that all other possible measurement outcomes,  $t \neq s$ , satisfy  $\sum_{t \neq s} n_t = 0$ . As we continue to work with single-photon input states, we keep this notation throughout the paper.

If  $\theta_w$  is fixed, the classical Fisher information of the free-space rotation is given by Eq. 2 or Eq. 18:

$$F_{free} = \frac{4}{1 - \theta_w^2}. \quad (26)$$

Eq. 26 will be used as the free-space benchmarking Fisher information in Eq. 19 when calculating counterfactual violation strengths further on in this paper.

If we instead assume no prior knowledge of  $\theta_w$  such that  $\theta_{min} = -1$ ,  $\theta_{max} = 1$  and  $p(\theta) = 1/2$  in Eq. 1, the mutual information is given by:

$$H(\theta_w : M) = \frac{\ln(108) - 4}{3 \ln(2)} \approx 0.328. \quad (27)$$

### IV.II. Nested MZI Interaction

We now investigate how the position of a polarization rotator in the nested Mach-Zehnder interferometer (see Fig. 5) changes the output probabilities, the Fisher information and the Shannon mutual information.

In general one can describe the normalised input and output vectors,  $\mathbf{a}$  and  $\mathbf{b}$ , of the NMZIs by:

$$\mathbf{a} = \frac{1}{\sqrt{L_a}} \begin{pmatrix} n_{1,H}^a \\ n_{2,H}^a \\ n_{3,H}^a \\ n_{1,V}^a \\ n_{2,V}^a \\ n_{3,V}^a \end{pmatrix}, \quad \mathbf{b} = \frac{1}{\sqrt{L_b}} \begin{pmatrix} n_{1,H}^b \\ n_{2,H}^b \\ n_{3,H}^b \\ n_{1,V}^b \\ n_{2,V}^b \\ n_{3,V}^b \end{pmatrix}, \quad (28)$$

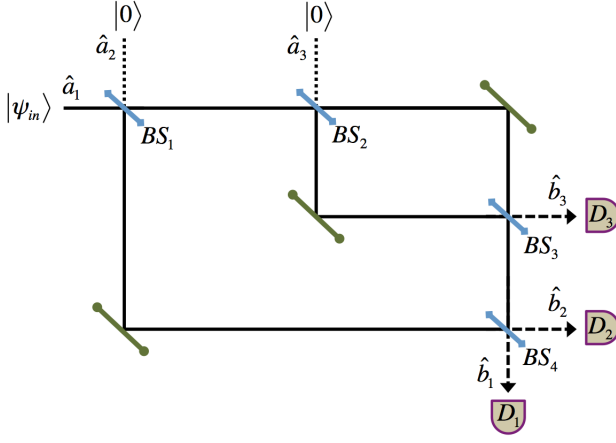


FIG. 5. (color online) A single-photon state is sent as the input to one of the ports of a NMZI. There are six possible detection outcomes (three spatial outcomes, each having a polarization degree of freedom). The green barred lines and blue arrowed lines represent mirrors and non-polarizing beam splitters respectively.

where  $L_a$  and  $L_b$  are some normalisation constants. The corresponding input creation operators are given by  $\hat{a}_i^\dagger$ . These are transformed into the output operators via the scattering matrix (see Eq. 4)  $\hat{b}_j^\dagger = \mathcal{S}_{i,j} \hat{a}_i^\dagger$ . In the following sections we restrict our input states to horizontal single-photon states, initially occupying the first spatial input port, such that  $|\psi_{in}\rangle = \hat{a}_{1,H}^\dagger |0\rangle$ . We drop the superscripts of the vector elements.

The beam splitters,  $BS_i$ , have reflection and transmission coefficients  $r_i$  and  $t_i$  respectively. In the NMZI device  $r_2 = t_2 = r_3 = t_3 = \frac{1}{\sqrt{2}}$ , such that the scattering matrix of the NMZI device, without polarization rotators (i.e.  $\hat{f}^{(k)}(\theta) = \hat{1}$  in Eq. 4), is given by:

$$\mathcal{S} = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 & 0 & 0 & 0 \\ -r_1 t_4 & -t_1 t_4 & r_4 & 0 & 0 & 0 \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_1 r_4 & t_1 r_4 & t_4 \\ 0 & 0 & 0 & -r_1 t_4 & -t_1 t_4 & r_4 \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}. \quad (29)$$

We proceed by evaluating the NMZI device for parameter estimation by considering five different scenarios with a polarization rotator placed in one out of five locations in the NMZI device. (See Fig. 6).

### 1. One

In our first scenario, we introduce a rotator in the lower arm of the Nested MZI. See position (1) in Fig. 6.

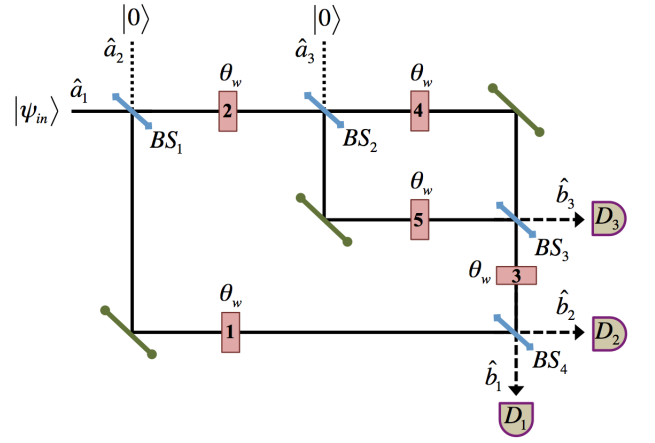


FIG. 6. (color online) A polarization rotation is added to one, but only one, of the positions, 1 - 5, in the Nested Mach-Zehnder interferometer from Fig. 5.

The scattering matrix of this device is given by:

$$\mathcal{S}_1 = \begin{pmatrix} r_1 r_4 2\bar{\theta}_w & t_1 r_4 2\bar{\theta}_w & t_4 & r_1 r_4 \theta_w & t_1 r_4 \theta_w & 0 \\ -r_1 t_4 2\bar{\theta}_w & -t_1 t_4 2\bar{\theta}_w & r_4 & -r_1 t_4 \theta_w & -t_1 t_4 \theta_w & 0 \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ -r_1 r_4 \theta_w & -t_1 r_4 \theta_w & 0 & r_1 r_4 2\bar{\theta}_w & t_1 r_4 2\bar{\theta}_w & t_4 \\ r_1 t_4 \theta_w & t_1 t_4 \theta_w & 0 & -r_1 t_4 2\bar{\theta}_w & -t_1 t_4 2\bar{\theta}_w & r_4 \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}, \quad (30)$$

where  $\bar{\theta}_w \equiv \sqrt{1 - \theta_w^2}/2$ .

Assuming the single-photon input state from above,  $|\psi_{in}\rangle$ , we get the following conditional probabilities for the possible output detections:

$$P(n_{1,H} = 1 | \theta_w) = r_1^2 r_4^2 (1 - \theta_w^2), \quad (31)$$

$$P(n_{1,V} = 1 | \theta_w) = r_1^2 r_4^2 \theta_w^2, \quad (32)$$

$$P(n_{2,H} = 1 | \theta_w) = r_1^2 t_4^2 (1 - \theta_w^2), \quad (33)$$

$$P(n_{2,V} = 1 | \theta_w) = r_1^2 t_4^2 \theta_w^2, \quad (34)$$

$$P(n_{3,H} = 1 | \theta_w) = t_1^2, \quad (35)$$

$$P(n_{3,V} = 1 | \theta_w) = 0. \quad (36)$$

The Fisher information of this device is given by:

$$F = \frac{4}{1 - \theta_w^2} r_1^2. \quad (37)$$

Furthermore, the Shannon mutual information is given by:

$$H(\theta_w : M) = \frac{\ln(108) - 4}{3 \ln(2)} r_1^2. \quad (38)$$

We see that the information content obtained from the measurement outcomes, as compared to the free-space scenario, is scaled by the square of the reflection coefficient of the first beam-splitter ( $BS_1$ ). The information is



reduced exactly by the square of the norm of the wavefunction that is scattered through parts of the interferometer not passing through the polarization rotator in the Schrödinger picture.

### 2. Two

In our next scenario, we consider the rotator to be placed in the upper interferometer path, before the second beam-splitter. The scattering matrix then takes the form of:

$$\mathcal{S}_2 = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 & 0 & 0 & 0 \\ -r_1 t_4 & -t_1 t_4 & r_4 & 0 & 0 & 0 \\ t_1 2\bar{\theta}_w & -r_1 2\bar{\theta}_w & 0 & t_1 \theta_w & -r_1 \theta_w & 0 \\ 0 & 0 & 0 & r_1 r_4 & t_1 r_4 & t_4 \\ 0 & 0 & 0 & -r_1 t_4 & -t_1 t_4 & r_4 \\ -t_1 \theta_w & r_1 \theta_w & 0 & t_1 2\bar{\theta}_w & -r_1 2\bar{\theta}_w & 0 \end{pmatrix}. \quad (39)$$

The conditional probabilities of this example are similar to those in Scenario One. However, the dependence on  $\theta_w$  is transferred to the third output port, from the first and second in the example above. The conditional probabilities are now given by:

$$P(n_1^H = 1|\theta_w) = r_1^2 r_4^2, \quad (40)$$

$$P(n_1^V = 1|\theta_w) = 0, \quad (41)$$

$$P(n_2^H = 1|\theta_w) = r_1^2 t_4^2, \quad (42)$$

$$P(n_2^V = 1|\theta_w) = 0, \quad (43)$$

$$P(n_3^H = 1|\theta_w) = t_1^2 (1 - \theta_w^2), \quad (44)$$

$$P(n_3^V = 1|\theta_w) = t_1^2 \theta_w^2. \quad (45)$$

The Fisher information is given by the expression:

$$F = \frac{4}{1 - \theta_w^2} t_1^2. \quad (46)$$

Additionally, the Shannon mutual information is scaled similarly, such that it is given by:

$$H(\theta_w : M) = \frac{\ln(108) - 4}{3 \ln(2)} t_1^2. \quad (47)$$

The  $r_1$  dependency of the previous scenario has, naturally, been transformed into a  $t_1$  dependency. Owing to the design of the NMZI device, the beam-splitters of the inner MZI have no effect on the information of  $\theta_w$  if the rotator is placed just after the first beam-splitter.

### 3. Three

In this scenario we investigate how the above studied properties change if the polarizing rotator is instead placed after the third—but before the fourth—beam-splitter. The scattering matrix is then given by:

$$\mathcal{S}_3 = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 2\bar{\theta}_w & 0 & 0 & t_4 \theta_w \\ -r_1 t_4 & -t_1 t_4 & r_4 2\bar{\theta}_w & 0 & 0 & r_4 \theta_w \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t_4 \theta_w & r_1 r_4 & t_1 r_4 & t_4 2\bar{\theta}_w \\ 0 & 0 & -r_4 \theta_w & -r_1 t_4 & -t_1 t_4 & r_4 2\bar{\theta}_w \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}. \quad (48)$$

The conditional probabilities are given by:

$$P(n_1^H = 1|\theta_w) = r_1^2 r_4^2, \quad (49)$$

$$P(n_1^V = 1|\theta_w) = 0, \quad (50)$$

$$P(n_2^H = 1|\theta_w) = r_1^2 t_4^2, \quad (51)$$

$$P(n_2^V = 1|\theta_w) = 0, \quad (52)$$

$$P(n_3^H = 1|\theta_w) = t_1^2, \quad (53)$$

$$P(n_3^V = 1|\theta_w) = 0. \quad (54)$$

The lack of dependence on  $\theta_w$  can simply be explained by the fact that the interference effects of the device prohibits any part of the wavepacket to evolve into the spatial location where the rotator is placed in this scenario.

The Fisher information is given by:

$$F = 0. \quad (55)$$

The absence of  $\theta_w$  in the conditional probabilities leads to a vanishing Fisher information. The same applies to the Shannon mutual information:

$$H(\theta_w : M) = 0. \quad (56)$$

In the investigation of this scenario, we make the observation of how the introduction of a polarizing rotator, (3) in Fig. 6, does not alter the output probabilities from the original device in Fig. 5. There is never any part of the wavefunction moving from the third to the fourth beam-splitter. Hence, an interaction in this region should not yield any information encoded in the particle.

### 4. Four & Five

In Scenario Four and Five we place the polarizing rotator in the upper and lower path of the inner MZI respectively. The scattering matrices of these cases are given by:

$$\mathcal{S}_4 = \begin{pmatrix} r_1 r_4 - t_1 t_4 \vartheta_w^- & t_1 r_4 + r_1 t_4 \vartheta_w^- & t_4 \vartheta_w^+ & t_1 t_4 \bar{\vartheta}_w^- & t_1 t_4 \bar{\vartheta}_w^- \\ -r_1 t_4 - t_1 r_4 \vartheta_w^- & -t_1 t_4 + r_1 r_4 \vartheta_w^- & r_4 \vartheta_w^+ & t_1 r_4 \bar{\vartheta}_w^- & t_1 r_4 \bar{\vartheta}_w^- \\ t_1 \vartheta_w^+ & -r_1 \vartheta_w^+ & -\vartheta_w^- & t_1 \vartheta_w^- & t_1 \vartheta_w^- \\ -t_1 t_4 \bar{\vartheta}_w^- & r_1 t_4 \bar{\vartheta}_w^- & -t_4 \bar{\vartheta}_w^- & r_1 r_4 - t_1 t_4 \vartheta_w^- & r_1 r_4 - t_1 t_4 \vartheta_w^- \\ -t_1 r_4 \bar{\vartheta}_w^- & r_1 r_4 \bar{\vartheta}_w^- & -r_4 \bar{\vartheta}_w^- & -r_1 t_4 - t_1 r_4 \vartheta_w^- & -r_1 t_4 - t_1 r_4 \vartheta_w^- \\ -t_1 \vartheta_w^- & r_1 \vartheta_w^- & -\vartheta_w^- & t_1 \vartheta_w^- & t_1 \vartheta_w^- \end{pmatrix}. \quad (57)$$

$$S_5 = \begin{pmatrix} r_1 r_4 - t_1 t_4 \vartheta_w^- & t_1 r_4 + r_1 t_4 \vartheta_w^- & -t_4 \vartheta_w^+ & -t_1 t_4 \overline{\vartheta}_w \\ -r_1 t_4 - t_1 r_4 \vartheta_w^- & -t_1 t_4 + r_1 r_4 \vartheta_w^- & -r_4 \vartheta_w^+ & -t_1 r_4 \overline{\vartheta}_u \\ t_1 \vartheta_w^+ & -r_1 \vartheta_w^+ & \vartheta_w^- & -t_1 \overline{\vartheta}_w \\ -t_1 t_4 \overline{\vartheta}_w & r_1 t_4 \overline{\vartheta}_w & t_4 \overline{\vartheta}_w & r_1 r_4 - t_1 t_4 \\ -t_1 r_4 \overline{\vartheta}_w & r_1 r_4 \overline{\vartheta}_w & r_4 \overline{\vartheta}_w & -r_1 t_4 - t_1 r_4 \\ -t_1 \overline{\vartheta}_w & r_1 \overline{\vartheta}_w & \overline{\vartheta}_w & t_1 \overline{\vartheta}_w \end{pmatrix} \quad (58)$$

where  $\vartheta_w^\pm \equiv (1 \pm \vartheta_w)/2$  and where we temporarily make a superficial change of variables such that  $\vartheta_w \equiv \sqrt{1 - \theta_w^2}$ .

For the two scenarios of introducing the polarization rotator inside the nested part of the interferometer, the conditional probabilities take more complicated forms:

$$P(n_1^H = 1 | \vartheta_w) = \frac{1}{4} (2r_1 r_4 - t_1 t_4 (1 - \vartheta_w))^2, \quad (59)$$

$$P(n_1^V = 1 | \vartheta_w) = \frac{1}{4} t_1^2 t_4^2 (1 - \vartheta_w^2), \quad (60)$$

$$P(n_2^H = 1 | \vartheta_w) = \frac{1}{4} (2r_1 t_4 + t_1 r_4 (1 - \vartheta_w))^2, \quad (61)$$

$$P(n_2^V = 1 | \vartheta_w) = \frac{1}{4} t_1^2 r_4^2 (1 - \vartheta_w^2), \quad (62)$$

$$P(n_3^H = 1 | \vartheta_w) = \frac{1}{4} t_1^2 (1 + \vartheta_w)^2, \quad (63)$$

$$P(n_3^V = 1 | \vartheta_w) = \frac{1}{4} t_1^2 (1 - \vartheta_w)^2. \quad (64)$$

The corresponding Fisher information of the device in Fig. 6, with a polarization rotator in the nested part ((4) or (5)), is given by:

$$F(\theta_w) = \frac{2}{1 - \theta_w^2} t_1^2. \quad (65)$$

In accordance with Eq. 18, we see that the Fisher information is proportional to how much of the wavepacket that—in the Schrödinger picture—has passed through the rotator in the device. In the scenarios of this subsection, the part of the wavefunction that travelled through the rotator in Scenario Two, is halved by the second beam-splitter before it is allowed to interact with the rotator. Hence,  $F$  (in Eq. 65) is halved as compared to its value in Scenario Two (Eq. 46).

Whilst the Fisher information preserves the simple form of Eq. 18, the corresponding Shannon mutual information for these devices is complicated and not very informative. However, by making the assumption that  $t_4 = r_1$  and  $r_4 = t_1$ , we can simplify the expression of the mutual information such that:

TABLE II. Numerical constants in [6/4] Padé approximation of Eq. 66.

$i$	$a_i$	$b_i$
2	$\frac{-3 + \ln(2) + 3 \ln(3)}{3 \ln(2)}$	-1
4	$\frac{25 - 6 \ln(2) + 25 \ln(3)}{18 \ln(2)}$	$\frac{-1}{10(-7 + 3 \ln(3))}$
6	$\frac{254 - 3 \ln(2) - 429 \ln(3) + 180 \ln(3)^2}{90(-7 + 3 \ln(3)) \ln(3)}$	

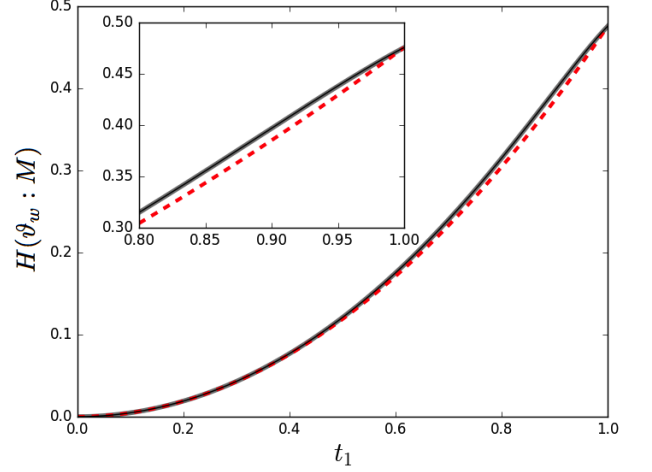


FIG. 7. (color online) The mutual information between the polarization rotation,  $\vartheta_w$ , and the measurement outcomes,  $\{M_i\}$ , as a function of beam-splitter transmission,  $t_1$ , as described in the text. The solid black line shows the true curve of Eq. 66, the thick grey line shows the Padé approximation (virtually indistinguishable from the true curve) and the red dashed line shows the second order Taylor expansion.

$$H(\vartheta_w : M) = \frac{1}{3 \ln(2) t_1^2} \left[ -2r_1^3 \ln(r_1^2) + t_1^2 \left( 3 \ln(3) + t_1^2 (\ln(2) - 1) - 2 - (3r_1^2 + t_1^4) \ln(3r_1^2 + t_1^4) \right) \right]. \quad (66)$$

Furthermore, we can approximate this expression. For scenarios where  $t_1 \approx 0$ , the Shannon mutual information is successfully modelled by a second order term:

$$H(\vartheta_w : M) \approx \frac{-3 + \ln(2) + 3 \ln(3)}{3 \ln(2)} t_1^2. \quad (67)$$

To obtain an even better model (for  $0 \leq t_1 \leq 1$ ) we can use a Padé approximant [39] of order [6/4]:

$$H(\vartheta_w : M) \approx \frac{a_2 t_1^2 + a_4 t_1^4 + a_6 t_1^6}{1 + b_2 t_1^2 + b_4 t_1^4}, \quad (68)$$

with constants  $a_i$  and  $b_j$  for  $i \in \{2, 4, 6\}$  and  $j \in \{2, 4\}$  given in Table II. Fig. 7 shows the mutual information from Eq. 66 and the two approximations as functions of  $t_1$ . The second order Taylor expansion (for  $t_1 \leq 0.4$ ) and the full Padé approximation, model the true curve within mean squared errors of  $3.1 \times 10^{-8}$  and  $2.8 \times 10^{-9}$  respectively.

After having established the bounds on parameter estimation with NMZI structures, we proceed to evaluate the counterfactuality of the Type I protocols, which are based on such devices.

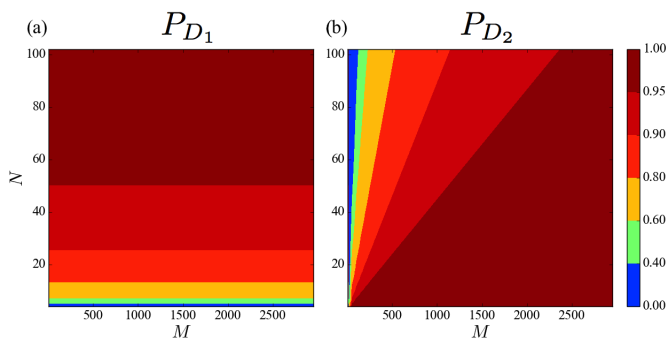


FIG. 8. (color online) Probability of detection at  $D_1$  (a) and  $D_2$  (b) if Bob unblocks and blocks his path in Fig. 2 respectively.

## V. EVALUATION OF TYPE I COUNTERFACTUAL COMMUNICATION

Fig. 2 shows a chained nested Mach-Zehnder interferometer. The first proposal of direct CFC [1] is based on such a device. The chained NMZI is divided such that the top right part of the individual NMZIs are in Bob’s laboratory. The communication scheme allows Alice to input a photon in the top left path (solid black line). If Bob wishes to transmit a logical 0, he leaves all paths open. If he, instead, wishes to transmit a logical 1, he blocks all paths with his detectors,  $D_B$ . Bob’s action, together with the number of inner and outer beam-splitters ( $M$  and  $N$  respectively), sets the output statistics in Alice’s laboratory (detections at  $D_1$  or  $D_2$ ). The outer and inner beam-splitters have their transmission coefficients set such that  $t_{n=1,\dots,N} = \sin(\pi/2N)$  and  $t_{m=1,\dots,M} = \sin(\pi/2M)$  respectively. In theory, for an infinitely large number of beam-splitters, the photon can be made to end up at  $D_2$  with  $P_{D_2} = 1$  (logical 1), or  $D_1$  with  $P_{D_1} = 1$  (logical 0), if Bob inserts or does not insert  $D_B$  in his laboratory, respectively [1].

Ref. [1] assumes that the evolution of the interrogating particle in the above described scheme can be modeled by perfectly unitary rotation matrices. For reference, we numerically calculate the detection probabilities of  $D_1$  and  $D_2$  detections in the scenarios of Bob keeping his laboratory open and blocked respectively. These probabilities are shown in Fig. 8 and are in accordance with those calculated in Ref. [1].

Fig. 8 suggests that for a communication scheme with low logical bit errors we need  $M \gg N$ . We see that success rates of 95% for the logical 0 (Fig. 8(a)) requires large  $N \approx 50$ . For such values of  $N$  we need  $M \approx 1200$  to keep the same success rates for the logical 1 (Fig. 8(b)). Hence, as discussed in Ref. [2], a success rate of about 95% requires a total of approximately 60000 beam-splitters to be used.

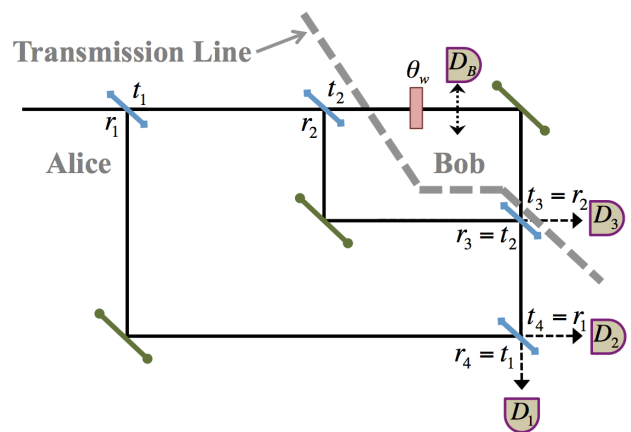


FIG. 9. (color online) The Nested Mach-Zehnder interferometer as used in the communication scheme presented in the text.

### V.I. Single NMZI

Ref. [30, 31, 40] suggests that the conceptual problem of the chained NMZI in Ref. [1] can be reduced to a study of a single NMZI device by considering pre- and post-selected events. Following this reduction, the mentioned references analytically analyse Type I “counterfactual” schemes based on single NMZI structures. However, to our knowledge there exist no rigorous proof that this reduction is an adequate representation of Salih’s scheme. Crucially, the single evaluation of the protocol that they consider does not allow for the transmission of logical bits. In this section we take an alternative approach.

#### 1. Analytical Analysis

Instead of treating the single NMZI device as a representation of Salih’s scheme, we evaluate a post-selected Type I protocol that actually allows for communication. Consider Fig. 9. We pre-select our states such that the input is the usual  $|\psi_{in}\rangle = a_{1,H}^\dagger |0\rangle$  from before. Bob has the option of introducing some absorbing object in his laboratory. Furthermore, we include a weak polarization rotation in Bob’s laboratory to mimic some disturbance in the device. We also post-select our states such that we exclude the events of absorption in Bob’s laboratory, by  $D_B$  or  $D_3$ .

If Bob wishes to transmit a logical 1 to Alice, he introduces the absorbing object in his laboratory. The renormalized output probabilities are then given by:

$$P_1(n_1^H = 1 | n_{3,B}^{H,V} = 0) = \mathcal{N}_1 \left( r_1 r_4 - r_2 t_1 t_2 t_4 \right)^2, \quad (69)$$

$$P_1(n_1^V = 1 | n_{3,B}^{H,V} = 0) = 0, \quad (70)$$

$$P_1(n_2^H = 1 | n_{3,B}^{H,V} = 0) = \mathcal{N}_1 \left( r_1 t_4 + r_2 r_4 t_1 t_2 \right)^2, \quad (71)$$

$$P_1(n_2^V = 1 | n_{3,B}^{H,V} = 0) = 0, \quad (72)$$

with  $\mathcal{N}_1 \equiv \left( 1 - t_1^2 (r_2^4 + t_2^2) \right)^{-1}$ .

We set the beam-splitter parameters according to Fig. 9, with  $r_2 = t_2 = 1/\sqrt{2}$ . This simplifies our expression such that:

$$P_1(n_1^H = 1 | n_{3,B}^{H,V} = 0) = \frac{r_1^2 t_1^2}{3r_1^2 + 1}, \quad (73)$$

$$P_1(n_1^V = 1 | n_{3,B}^{H,V} = 0) = 0, \quad (74)$$

$$P_1(n_2^H = 1 | n_{3,B}^{H,V} = 0) = \frac{(r_1^2 + 1)^2}{3r_1^2 + 1}, \quad (75)$$

$$P_1(n_2^V = 1 | n_{3,B}^{H,V} = 0) = 0. \quad (76)$$

The corresponding value for the conditioned Fisher information of this 1-bit protocol is:

$$F^1 = 0. \quad (77)$$

This is expected, as any part of the wavefunction that interacts with the rotator is then absorbed by  $D_B$  and can thus not reach Alice.

Now, consider the process in which Bob instead wishes to transmit a logical 0 to Alice. He then leaves his laboratory free, without any absorbing objects. The renormalized probabilities of detecting the particle in the following states are given by:

$$P_0(n_1^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}_0 \left( r_1 r_4 + r_2 t_1 t_2 t_4 \right. \\ \left. \times (2\bar{\theta}_w - 1) \right)^2, \quad (78)$$

$$P_0(n_1^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}_0 \left( r_2^2 t_1^2 t_2^2 t_4^2 \theta_w^2 \right), \quad (79)$$

$$P_0(n_2^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}_0 \left( r_1 t_4 - r_2 r_4 t_1 t_2 \right. \\ \left. \times (2\bar{\theta}_w - 1) \right)^2, \quad (80)$$

$$P_0(n_2^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}_0 \left( r_2^2 r_4^2 t_1^2 t_2^2 \theta_w^2 \right), \quad (81)$$

with  $\mathcal{N}_0 \equiv \left( 1 - t_1^2 (r_2^4 + t_2^4 + 4r_2^2 t_2^2 \bar{\theta}_w) \right)^{-1}$ . Again, we set the beam-splitter parameters according to Fig. 9, with  $r_2 = t_2 = 1/\sqrt{2}$ . The probabilities then simplify to:

$$P_0(n_1^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}'_0 r_1^2 t_1^2 (1 + 2\bar{\theta}_w)^2, \quad (82)$$

$$P_0(n_1^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}'_0 r_1^2 t_1^2 \theta_w^2, \quad (83)$$

$$P_0(n_2^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}'_0 (1 + r_1^2 - 2t_1^2 \bar{\theta}_w)^2, \quad (84)$$

$$P_0(n_2^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathcal{N}'_0 t_1^4 \theta_w^2, \quad (85)$$

where  $\mathcal{N}'_0 \equiv \left( 4 - 2t_1^2 (1 + 2\bar{\theta}_w) \right)^{-1}$ .

For the 0-bit scheme considered above, we then obtain:

$$F^0 = \frac{t_1^2}{1 - \theta_w^2}. \quad (86)$$

This can be compared to the 1-bit scheme, where the value is  $F^1 = 0$ , such that there is no information about the angle  $\theta_w$  given to Alice.

The two processes described in this section can be used in order to transmit information from Bob to Alice in a scenario where Bob only has access to the inner part of the NMZI (see Fig. 9). However, in order for our setup to be representative of the behaviour of the chained NMZI structure from the section above, we set  $\theta_w \ll t_1 \ll r_1$ . Hence, the polarization rotation will have a minute impact on the probability outputs. For  $t_1 \ll r_1$ , the probability distributions for the 0- and 1-bit processes are very similar. We see that in both processes of the scheme, Alice will detect the state  $|n_2^H = 1\rangle$  with high probability. Thus, in order for Alice to obtain Bob's choice of bit-value, with high probability, each logical bit has to be decoded from a larger bit string.

The communication scheme is as follows: Alice sends a number,  $n_\gamma$ , of single photons (excluding the particles that do not fulfil the post-selection criterion and are absorbed by either  $D_B$  or  $D_3$ ) into the device of Fig. 9, one after another. Depending on what logical bit Bob wishes to transmit, he either inserts detector  $D_B$  or leaves his laboratory open, for all the  $n_\gamma$  particles. Alice makes subsequent particle detections of the  $n_\gamma$  events:  $|n_1^H = 1\rangle$ ,  $|n_2^H = 1\rangle$ ,  $|n_1^V = 1\rangle$  or  $|n_2^V = 1\rangle$ . If she measures any event in  $|n_1^V = 1\rangle$  or  $|n_2^V = 1\rangle$ , she knows with certainty that a logical 0 was sent. However, owing to the fact that  $\theta_w$  is very small, the accumulative probability of these events is also small. Hence, Alice will, with high probability, have to use the statistics of  $|n_1^H = 1\rangle$  and  $|n_2^H = 1\rangle$  detections to infer the logical bit. From the number of particles,  $q$ , that Alice measure in the  $|n_1^H = 1\rangle$  state, she decides whether Bob sent a logical 0 or a logical 1.

The question of interest is: what number,  $n_\gamma$ , of single-photon evaluations of the device, allows for an effective communication scheme with a limited number of errors?

We re-define  $P_1 \equiv P_1(n_1^H = 1 | n_{3,B}^{H,V} = 0)$  and  $P_0 \equiv P_0(n_1^H = 1 | \theta_w, n_{3,B}^{H,V} = 0)$ . For small  $t_1$  and  $\theta_w \ll t_1$ , we see that  $P_1 < P_0$ . Alice will thus note down a 1 every time  $q < q'$ . In the limit of long message strings, Bob will produce logical 0s and 1s at the same rates, and the exact value of  $q'$  is given by:

$$q' = \left\lfloor \frac{n_\gamma \ln \left( \frac{1-P_1}{1-P_0} \right)}{\ln \left( \frac{P_0}{P_1} \right) - \ln \left( \frac{1-P_0}{1-P_1} \right)} \right\rfloor. \quad (87)$$

The probability for a non-faulty logical bit-transmission

is thus:

$$P_{succ.} = \frac{1}{2} \sum_{q=0}^{q'} \frac{n_\gamma! P_1^q (1 - P_1)^{n_\gamma - q}}{q!(n_\gamma - q)!} + \frac{1}{2} \sum_{q=q'+1}^{n_\gamma} \frac{n_\gamma! P_0^q (1 - P_0)^{n_\gamma - q}}{q!(n_\gamma - q)!}. \quad (88)$$

Eq. 88 can be used to numerically find an acceptable value of  $n_\gamma$ , given the parameters of the setup. However, in order to evaluate the setup discussed in Ref. [1, 27, 29–32], we need the transmission coefficient to be small ( $t_1 = \sin(\pi/2N) \ll 1$ , with  $N \gg 1$ ). As  $t_1$  is small, we can, by the central limit theorem, assume that  $n_\gamma$  has to be large and that the two bit-processes will generate normally distributed events. The two processes will each have a mean situated at  $P_1$  and  $P_0$  respectively. Their respective standard deviations will be given by:

$$\sigma_i = \sqrt{\frac{P_i(1 - P_i)}{n_\gamma}}, \quad (89)$$

where  $i = 0, 1$ , which decreases reciprocally with the square-root of  $n_\gamma$ . For Alice to be able to distinguish between the logical 0 and 1 bits correctly with probability  $1 - \epsilon$ , we require that:

$$n_\gamma \geq \left( \Phi^{-1}(\epsilon) \frac{\sqrt{P_0(1 - P_0)} + \sqrt{P_1(1 - P_1)}}{P_0 - P_1} \right)^2, \quad (90)$$

where  $\Phi^{-1}(\epsilon)$  is the inverse of the standard normal cumulative distribution function.

We can Taylor expand  $n_\gamma$  for small values of  $t_1$  such that:

$$n_\gamma \geq (\Phi^{-1}(\epsilon))^2 \frac{4}{t_1^2} + \mathcal{O}(t_1^{-1}). \quad (91)$$

As the Fisher information scales linearly with the number of evaluations of the channel,  $n_\gamma$ , our counterfactual violation strength for a Type I logical 0-bit is given by:

$$\mathcal{D} = n_\gamma \frac{F^0}{F_{free}} \gtrsim (\Phi^{-1}(\epsilon))^2. \quad (92)$$

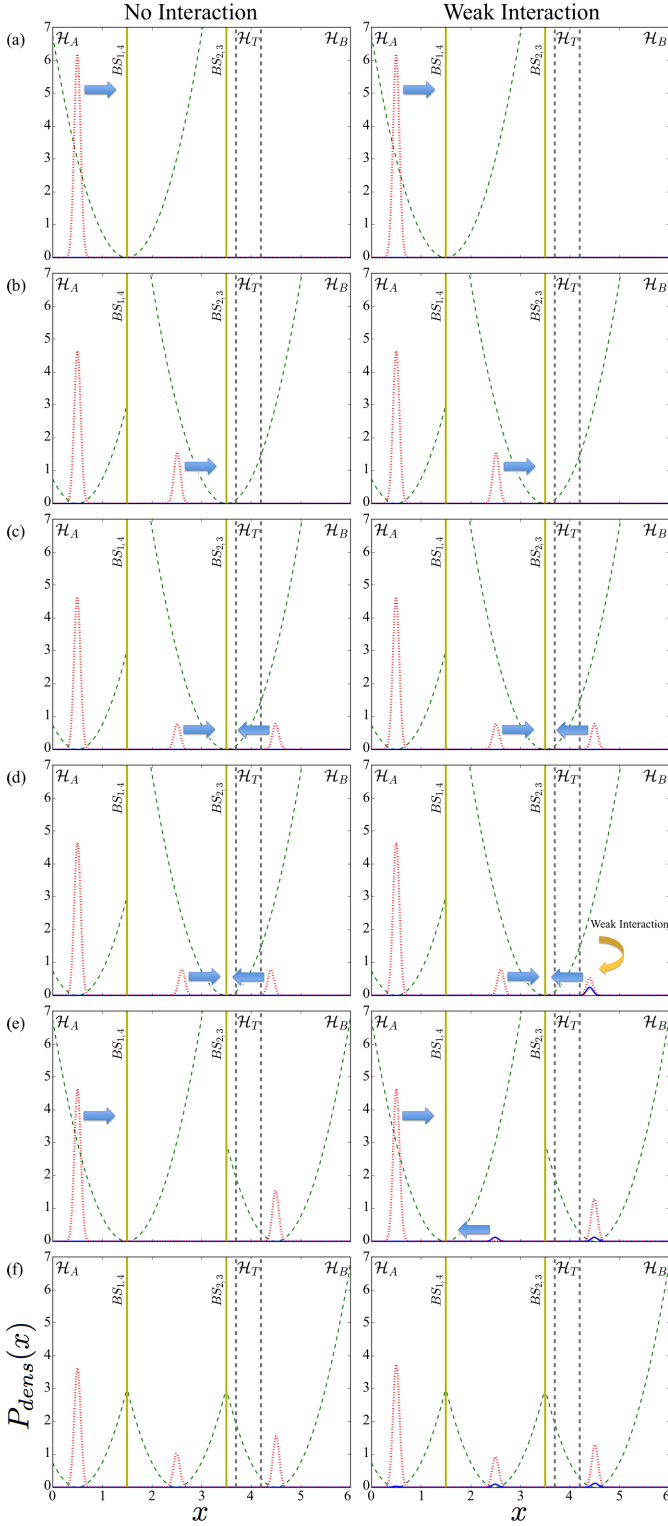
For a success rate of roughly 95%, we thus obtain a value of  $\mathcal{D} \approx 2.7$ , and we conclude that the CFC scheme described in this subsection is no more counterfactual than a free space evolution of particles between Alice and Bob.

## 2. Simulation of Quantum Evolution

In order to illustrate the origin of a counterfactual violation inside a NMZI device, we provide a numerical simulation of the time-dependent Schrödinger equation. We simulate a massive Gaussian spin- $\frac{1}{2}$  particle that propagates through a NMZI, which we have mapped onto a

linear 1D structure. The Hamiltonian to implement such an evolution can be tailored as in Ref. [2]. This allows us to design a toy model for the wavefunction evolution in a NMZI. The solution is calculated by a GPU-boosted version of the Staggered Leapfrog algorithm as in Refs. [41, 42]. Fig. V V.12 shows the evolution of the wavefunction.

The Hamiltonian has been tailored such that the beam-splitter parameters are given by:  $t_2 = r_2 = t_3 = r_3 = \frac{1}{\sqrt{2}}$ ,  $t_1 = t_4 = \frac{1}{2}$  and  $r_1 = r_4 = \frac{\sqrt{3}}{2}$ .



(color online) This figure shows the quantum evolution (time steps (a) to (f)) of the probability density distribution of a spin- $\frac{1}{2}$  particle in a nested Mach-Zehnder interferometer with and without a weak spin-rotation interaction (exaggerated for visibility) in Bob's laboratory. The dotted red and solid blue curves indicate spin up and spin down components of the wavefunction respectively. The dashed green curves show the potentials. Beam-splitters

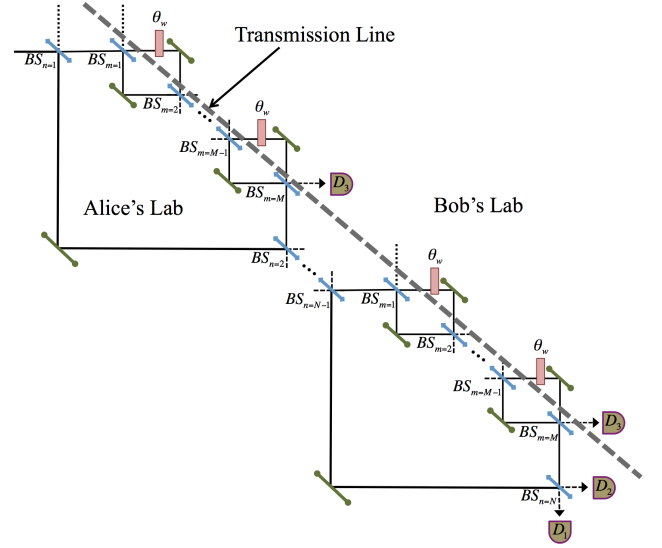


FIG. 10. (color online) The logical 0 process for the chained nested Mach-Zehnder interferometer suggested for CFC in Ref. [1]. The weak polarization interactions mimic realistic systematic errors in the quantum channels of Bob's laboratory.

are denoted with vertical yellow lines. The spatial components are indicated with the vertical dashed grey lines.

As can be seen from Fig. V V.I2, the effect of the weak interaction in Bob's laboratory (right frame in Fig. V V.I2(d)) is to distort the interaction on the beam-splitter (between Fig. V V.I2(d) and V V.I2(f)). In the scenario of no interaction, the second passage through the right beam-splitter causes the middle and right parts of the wavepacket to interfere constructively and propagate to Bob's laboratory, never to return to Alice. However, a weak interaction in Bob's laboratory allows for a fraction of the wavepacket (right frame in Fig. V V.I2(e)) to propagate back towards Alice's laboratory and interfere on the left beam-splitter. Hence, the probability density distributions around  $x = 0.5$  and  $x = 2.5$  in Alice's laboratory are different depending on whether or not a weak interaction took place. This is why the Type I counterfactuality is satisfied, only if absolutely pure quantum channels are present in the NMZI device.

## V.II. Chained NMZIs

In the previous subsection, we analysed a reduced NMZI CFC scheme. However, an extension of the analysis to the originally proposed chained NMZI protocol of Ref. [1] is not straightforward. (Previous attempts have been heavily criticised [1, 26–34, 40]). In general, complications arise from the multiple paths  $((N-1) \times (M-1))$  in and out of Bob's laboratory. In terms of our approach, this complicates the concept of “presence”. Nevertheless, the Fisher information with respect to the weak

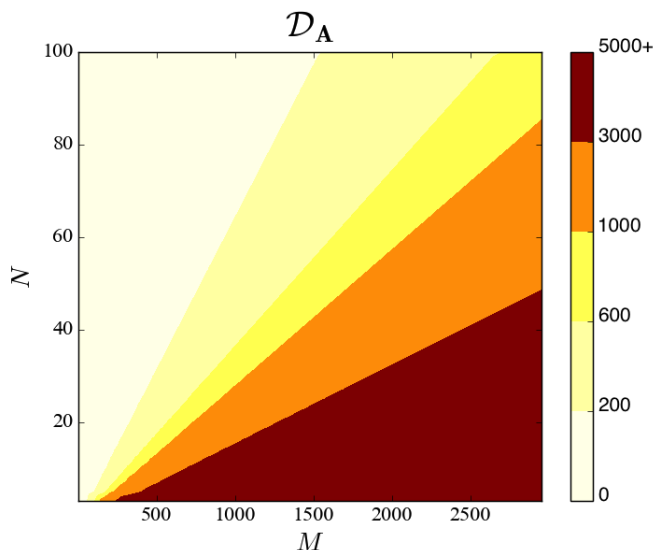


FIG. 11. (color online) The spatially conditioned Type I counterfactual violation strength as a function of the beam-splitter numbers  $N$  and  $M$  for the scenario of Bob not introducing his detectors in Fig. 2 or 10.

disturbance in Bob’s laboratory can still be calculated—and generally have significantly larger values than those calculated with Eq. 18. An analytical analysis of the Fisher information in the chained NMZI devices yields a complicated non-informative expression, even for small numbers of  $M$  and  $N$ . Instead, a numerical finite difference method allows for a comprehensive approximation of the Fisher information. This allows us to calculate the counterfactual violation strength (Eq. 19).

We now calculate the quantum evolution of the Type I logical 0 process (no  $D_B$ ), with a polarization rotation of  $\theta_w$  in every inner MZI in Bob’s laboratory. (See Fig. 10). In accordance with the previous sections of this paper, the weak rotations mimic disturbances of realistic quantum channels. Fig. 11 shows the spatially conditioned (i.e.  $F \rightarrow F_A$  in Eq. 19) Type I counterfactual violation strength,  $\mathcal{D}_A$ , as a function of  $N$  and  $M$ , assuming that Bob leaves his path open and that the polarization rotations are carried out with a weak polarization parameter:  $\theta_w = 10^{-6} \ll M^{-1}$ . A simple calculation shows that  $\mathcal{D} \geq \mathcal{D}_A$ , such that Fig. 11 can be used as a lower bound on the Type I counterfactual violation strength in the device of Fig. 10.

As stated before, large values of  $M$  and  $N$  are needed to carry out the direct communication scheme with high success probability. For such values, the counterfactual violation strength of the chained NMZI communication scheme is many orders of magnitude larger than unity. For realistic quantum channels, we can thus, based on our counterfactuality measure, conclude that the suggested communication scheme of Ref. [1] is, *de facto*, not counterfactual.

## VI. EVALUATION OF TYPE II COUNTERFACTUAL COMMUNICATION

We now consider the Type II protocol suggested by Arvidsson-Shukur and Barnes [2], which relies on a chained MZI (CMZI) device. Such a device is shown in Fig. 3.

In this protocol, Alice sends a single-photon state into the upper left input path of the device (solid black line). The photon then enters the transmission line, which shares a CMZI device with Bob’s laboratory. Bob has the possibility of inputting detectors in his path or leaving it open. If Bob leaves his path open, the quantum evolution of the photon in the device will lead it to emerge onto detector  $D_2$  in Bob’s laboratory, without the wavepacket ever spreading into Alice’s laboratory after it first left it. However, if Bob instead inputs detectors,  $D_B$ , after each beam-splitter, the wavepacket will either be absorbed by one of them or collapse onto a state in the lower part of the CMZI device that can re-enter Alice’s laboratory to be detected by detector  $D_1$ . In the limit of large  $N$  and inserted detectors, the quantum Zeno effect can make the probability of re-entering in Alice’s laboratory arbitrarily close to unity. Hence, Bob’s action of either leaving his path free or inserting detectors affects the detection probabilities in Alice’s laboratory and—in the limit of large  $N$ —allows her to deduce Bob’s action with high probability of success.<sup>3</sup>

This protocol is conceptually different from the one presented in Ref. [1]. That protocol suggests a scenario where the photon would never travel from Bob to Alice or vice versa. The protocol described in Ref. [2] indeed never sees the photon wavefunction travel from Bob to Alice. It does, however, allow for the wavefunction to propagate from Alice to Bob. The protocol is, nevertheless, counterfactual according to the Type II definition, which allows particles to travel in the opposite direction to the message.

We simulate the Type II counterfactual violation strength ( $P_A$  from Eq. 20) per photon transport through the CMZI with Bob’s path open, as a function of  $N$ . The weak polarization rotators are again inserted in each separate MZI of Bob’s laboratory. (See Fig. 12).  $P_A$  is dependent of  $\theta_w$  and we present results for various values of  $\theta_w$  in Fig. 13.

The CMZI CFC protocol can be carried out with a high success rate with less than 100 beam-splitters (the success rate of the 1-bit process when Bob inserts his detectors is above 95% if  $N \geq 50$ ). Hence, Fig. 13 validates the counterfactuality of the CMZI scheme for the values of  $\theta_w$  that we have considered. We see that for small values of  $\theta_w$  the value of  $P_A$  is kept well below the free-space interaction value of 1. Regardless of the small values

<sup>3</sup> The protocol presented in Ref. [2] shows how a clever logical bit-encoding scheme can take the probability of success close to unity, even for imperfect beam-splitters and  $N \leq 7$ .

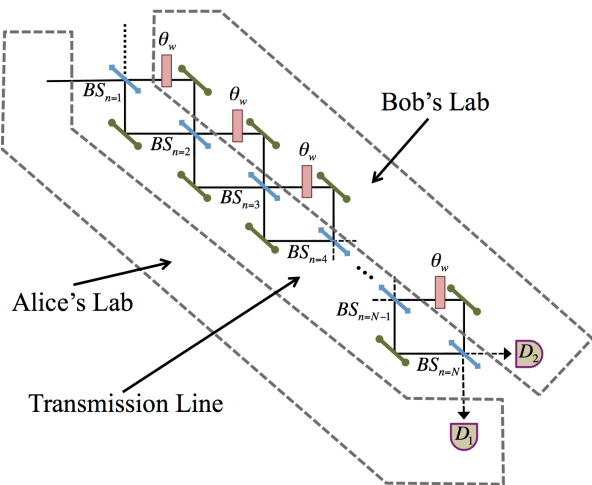


FIG. 12. (color online) The logical 0 process of the chained Mach-Zehnder interferometer CFC scheme of Ref. [2]. Again, the weak polarization interactions mimic realistic systematic errors in the quantum channels of Bob’s laboratory.

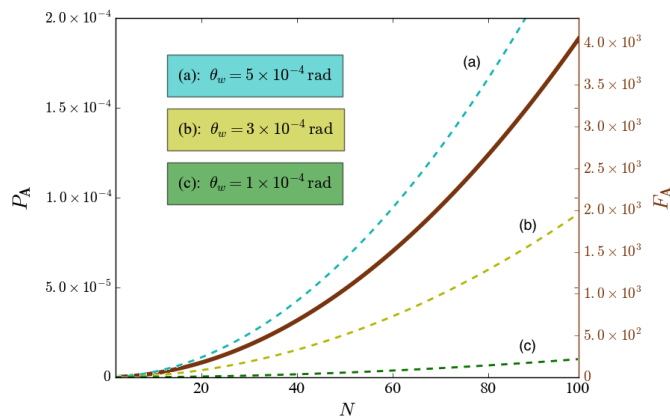


FIG. 13. (color online) The Type II logical 0 (see Ref. [2]) counterfactual violation strength,  $P_A$  (dashed lines read on the left  $y$ -axis), and the spatially conditioned Fisher information,  $F_A$  (solid line read on the right  $y$ -axis), as functions of the beam-splitter number,  $N$ .

of  $P_A$ , Fig. 13 shows that Alice obtains a large amount of Fisher information  $F_A(\theta_w)$  about Bob’s parameter  $\theta_w$ . The value of  $F_A(\theta)$  is independent of changes in the value of  $\theta$  for small  $\theta = \theta_w \approx 0$ . The value of  $P_A$  is not.

## VII. CONCLUDING REMARKS

In this paper we have carried out a thorough study of the classical Fisher information and the Shannon mutual information in nested Mach-Zehnder interferometers. We have calculated these information measures with respect to the measurement outcomes caused by polarization ro-

tations in different parts of the structure. We find that in an otherwise non-polarizing optical circuit with real beam-splitter matrices, the Fisher information caused by a single polarization rotation is always proportional to the integrated probability density distribution at the location of the interaction in the Schrödinger picture. The Fisher information—in such a scenario—can thus be thought of as a measure of inter-measurement “presence” at the rotator. Furthermore, we have developed interpretation independent measures for the strength of counterfactual violations in two different types of CFC schemes. Type I schemes do not allow particles to cross the transmission line between transmitter and receiver, whilst the Type II schemes allow particles to travel in to opposite direction to the message. The rudimentary assumption made in this paper, is that any real quantum channel will naturally have unwanted components associated with it. We introduce a small polarization rotation as a simple model of an unwanted quantum evolution in the devices. We find that a suitable measure of Type I counterfactual violations is based on the classical Fisher information of the parameter  $\theta_w$  that set the polarization rotation. It is scaled by the reciprocal Fisher information from a free-space interaction of the same strength as the one used in the device of interest. Hence, a value of  $\mathcal{D} = 1$  corresponds to the counterfactual violation strength of a free-space interaction inside the laboratory of study, and any value between 0 and 1 corresponds to a weaker-than-free-space counterfactual violation. For example, we can show that the suggested scheme of Salih et al. [1] strongly violates counterfactuality. Moreover, we provide an analytical study as well as a numerical simulation of the quantum evolution of a Gaussian particle in a single NMZI. The numerics show how the evolution of the quantum wavefunction leads to Fisher information of the interaction parameter (and ultimately the counterfactual violations) in NMZIs communication schemes. The analytical study supports the claim of the invalidity of counterfactual schemes based on NMZIs. The Type II scheme developed by Arvidsson-Shukur and Barnes [2] should be measured with another counterfactual violation measure. This measure,  $P_A$ , is based on how much probability density that weak non-collapsing interactions in Bob’s laboratory generate in Alice’s. We find that the Type II scheme of Ref. [2] keeps the value of  $P_A$  to a fraction of a percent for weak interactions, and thus satisfies its counterfactuality definition.

The authors would like to express their gratitude towards Thomas Bahder and Hugo Lepage for helpful discussions. This work was supported by the EPSRC, and the Cambridge Laboratory of Hitachi Limited via Project for Developing Innovation Systems of the MEXT in Japan. It has also received generous contributions from Helge Ax:son Johson’s Trust Fund, Anna Whitlock’s Trust Fund, Sixten Gemzeus’ Foundation, Långmanska kulturfonden’s Trust Fund, Tornspiran’s Trust Fund, Lars Hiertas’ Memorial Foundation and Gustaf Söderberg’s Trust Fund.



- 
- [1] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, *Phys. Rev. Lett.* **110**, 170502 (2013).
- [2] D. R. M. Arvidsson-Shukur and C. H. W. Barnes, *Phys. Rev. A* **94**, 062303 (2016).
- [3] A. A. Michelson and E. W. Morley, *American Journal of Science* **3**, **32**, 333 (1887).
- [4] L. Hardy, *Phys. Rev. Lett.* **68**, 2981 (1992).
- [5] B. P. Abbott and et al. (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [6] P. Hariharan, *Optical Interferometry*, 2nd ed. (Academic Press, Cambridge, Massachusetts, USA, 2003) p. 351.
- [7] G. Bertocchi, O. Alibart, D. B. Ostrowsky, S. Tanzilli, and P. Baldi, *Journal of Physics B: Atomic, Molecular and Optical Physics* **39**, 1011 (2006).
- [8] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglià, *Phys. Rev. D* **76**, 104044 (2007).
- [9] J. K. Stockton, X. Wu, and M. A. Kasevich, *Phys. Rev. A* **76**, 033613 (2007).
- [10] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, *Phys. Rev. A* **80**, 013825 (2009).
- [11] H. Cable and G. A. Durkin, *Phys. Rev. Lett.* **105**, 013603 (2010).
- [12] S. Vitale and M. Zanolin, *Phys. Rev. D* **82**, 124065 (2010).
- [13] T. B. Bahder, *Phys. Rev. A* **83**, 053601 (2011).
- [14] L. Rozema, A. Darabi, D. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, in *Frontiers in Optics 2012/Laser Science XXVIII* (Optical Society of America, 2012) p. FW4J.4.
- [15] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, *Phys. Rev. Lett.* **109**, 100404 (2012).
- [16] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, *Phys. Rev. Lett.* **111**, 183601 (2013).
- [17] Z. Hradil, *Phys. Rev. A* **51**, 1870 (1995).
- [18] T. B. Bahder and P. A. Lopata, *Phys. Rev. A* **74**, 051801 (2006).
- [19] D. S. Simon, A. V. Sergienko, and T. B. Bahder, *Phys. Rev. A* **78**, 053829 (2008).
- [20] R. Penrose, *Shadows of the Mind: A Search for the Missing Science of Consciousness*, 1st ed. (Oxford University Press, Inc., New York, NY, USA, 1994) p. 240.
- [21] A. C. Elitzur and L. Vaidman, *Foundations of Physics* **23**, 987 (1993).
- [22] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, *Phys. Rev. Lett.* **74**, 4763 (1995).
- [23] P. G. Kwiat, A. G. White, J. R. Mitchell, O. Nairz, G. Weihs, H. Weinfurter, and A. Zeilinger, *Phys. Rev. Lett.* **83**, 4725 (1999).
- [24] O. Hosten, M. T. Rakher, J. T. Barreiro, N. A. Peters, and P. G. Kwiat, *Nature* **439** (2006), <http://dx.doi.org/10.1038/nature04523>.
- [25] Y. Cao, Y.-H. Li, Z. Cao, J. Yin, Y.-A. Chen, H.-L. Yin, T.-Y. Chen, X. Ma, C.-Z. Peng, and J.-W. Pan, *Proceedings of the National Academy of Sciences* **114**, 4920 (2017).
- [26] Z.-H. Li, M. Al-Amri, and M. S. Zubairy, *Phys. Rev. A* **92**, 052315 (2015).
- [27] L. Vaidman, *Phys. Rev. A* **87**, 052104 (2013).
- [28] Z.-H. Li, M. Al-Amri, and M. S. Zubairy, *Phys. Rev. A* **88**, 046102 (2013).
- [29] L. Vaidman, *Phys. Rev. A* **88**, 046103 (2013).
- [30] L. Vaidman, *Phys. Rev. Lett.* **112**, 208901 (2014).
- [31] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, *Phys. Rev. Lett.* **112**, 208902 (2014).
- [32] L. Vaidman, *Phys. Rev. A* **93**, 066301 (2016).
- [33] Z.-H. Li, M. Al-Amri, and M. S. Zubairy, *Phys. Rev. A* **93**, 066302 (2016).
- [34] R. B. Griffiths, *Phys. Rev. A* **94**, 032115 (2016).
- [35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (John Wiley and Sons Inc., Hoboken, New Jersey, USA, 2006).
- [36] A. Marlow, *Mathematical Foundations of Quantum Theory*, 1st ed. (ACADEMIC PRESS, INC. (LONDON) LTD., London, UK, 1978).
- [37] A. Degasperis, L. Fonda, and G. C. Ghirardi, *Il Nuovo Cimento A (1965-1970)* **21**, 471 (1974).
- [38] B. Misra and E. C. G. Sudarshan, *Journal of Mathematical Physics* **18** (1977).
- [39] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed., Vol. 1 (Press Syndicate of the University of Cambridge, Cambridge, UK, 1992).
- [40] L. Vaidman, *Journal of Physics A: Mathematical and Theoretical* **48**, 465303 (2015).
- [41] E. T. Owen, M. C. Dean, and C. H. W. Barnes, *Phys. Rev. A* **89**, 032305 (2014).
- [42] D. R. M. Arvidsson-Shukur, H. V. Lepage, E. T. Owen, T. Ferrus, and C. H. W. Barnes, *Phys. Rev. A* **96**, 052305 (2017).