Optimization of direct-detection quantum illumination

Thomas Brougham and John Jeffers
Department of Physics, University of Strathclyde, John Anderson Building, 107 Rottenrow, Glasgow G4 0NG, United Kingdom

(Received 9 January 2024; accepted 22 April 2024; published 7 May 2024)

Quantum illumination uses quantum correlations to improve the detection of objects in the presence of background noise. A simple and common practical approach to quantum illumination is to use threshold detectors to measure the idler mode. We investigate the optimization of these *direct measurement* protocols. Surprisingly, we find that there can be an advantage to having a signal detector whose quantum efficiency is significantly less than perfect, one that does not vanish for low object reflectivities or signal strengths. We also show that decreasing the separation between pulses, while keeping the rate of photons transmitted per second fixed, can improve the performance, under appropriate conditions. We further show that postselecting on the idler detector firing is less efficient at object detection than not postselecting.

DOI: 10.1103/PhysRevA.109.052612

I. INTRODUCTION

Quantum correlations have long been of fundamental importance [1,2] but with the development of quantum information, quantum correlations form the basis of many applications such as sensing [3,4], quantum imaging [5], teleportation [6], and quantum key distribution [7,8]. One promising application is quantum illumination, which provides an advantage for detecting objects in the presence of background noise [9–12]. Normally, to detect an object against background light, the intensity of the signal must be increase to compensate. However, this is not always possible. For instance, the object of interest might be fragile, or laser safety might limit the signal intensities. Another possibility is that one wants to detect an object covertly, in which case the signal's intensity should be lower than that of the background [13–15]. In each of these examples, quantum illumination can be used to detect an object without requiring a high signal intensity.

Quantum illumination exploits correlations in intensity between two optical modes, often called the signal and idler. One can generate correlated twin beams by using nondegenerate spontaneous parametric down-conversion (SPDC) [16–18]. One of the beams, the signal, is transmitted to the target region, where an object may or may not be present, but from which emanates bright background light. If an object is present, then some of the signal light will be reflected to the detection system. Alternatively, if there is no object, then only background light is detected. These two cases can be distinguished more easily using the intensity correlations. In the early quantum illumination protocols [9,10], a joint measurement on the idler and signal modes was proposed to provide close to optimal detection [10]. However, this measurement is difficult to achieve in experiments [19,20]. This has motivated the exploration of experimentally simpler approaches. One common approach is to use threshold detectors to measure the idler mode and then transmit the signal mode to the target [14,21–26]. Any reflected light is measured by another threshold detector. This measurement is not optimal and does not yield the 6 dB quantum advantage over optimal classical schemes [10]. Nevertheless, using this approach has been shown to give an advantage compared with transmitting single-mode coherent states [15,26]. Other types of protocol exist, such as using multiple idler detectors [26], using random coherent states to mimic quantum illumination [15], and jamming resistant protocols [14,27].

The performance of direct detection protocols depends on environmental parameters, which we cannot control, together with system parameters, which we can change. Examples of system parameters are the mean number of signal photons, the efficiencies of the idler and signal detectors, and the spacing between pulses. We have examined the effects of changing each of these system parameters, with the aim of optimizing the performance of the protocol. One parameter that we can control easily is detector efficiency, the probability that a single photon causes a detector to fire. Better technology can improve efficiency, but we can always make it worse via attenuation. Surprisingly, we find that, under certain conditions, the optimal efficiency for the signal detector is less than one.

There are other ways in which one could try to improve the performance of the protocol. For instance, suppose we keep the rate of transmitted signal photons per second constant. One can then investigate whether we improve the performance of the protocol by using pulses of shorter duration and shorter separations, where each pulse has fewer photons per pulse. We find that using more frequent and shorter pulses with fewer photons per pulse does help to improve the performance of direct measurement illumination. Finally, we analyze protocols that postselect on the idler detector firing. We argue that postselection does not improve the performance even if one

neglects the increased time required to acquire postselected data.

The outline of the paper is as follows: In Sec. II we describe direct measurement illumination, before introducing an analytic criterion for its evaluation. This criterion is then used in Sec. III to show that, under appropriate conditions, using unequal detector efficiencies can improve quantum illumination. The size of this effect is determined both analytically for a single experimental shot and by performing Monte Carlo simulations for multiple shots. In Sec. IV we analyze quantum illumination protocols in terms of the number of photons transmitted. Using this approach, we find that using pulses with fewer photons, but with shorter separations, can improve the effectiveness of quantum illumination. We also investigate the effects of postselection. Finally, we discuss our results in Sec. V.

II. DIRECT MEASUREMENT QUANTUM ILLUMINATION

We start by outlining the particular quantum illumination scheme we are investigating. This protocol replaces the complex joint measurement of the idler and signal modes proposed in many schemes [10,11,20,28,29], with separate measurements on each mode. As such, we call it a *direct measurement* protocol. This type of protocol has been demonstrated experimentally [14,21,22]. The use of separate measurement is, however, nonoptimal. In particular, the protocol does not provide a quantum advantage over the all possible classical protocols. Nevertheless, direct measurement protocols have been shown to have an advantage over using coherent states [15,26].

Direct measurement protocols requires a state with quantum intensity correlations between two modes such that when one mode is brighter than average, the other one is also. An example of such as state is a two-mode squeezed vacuum (TMSV) state [30]

$$|\Psi\rangle_{I,S} = \frac{1}{\sqrt{1+\bar{n}}} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{1+\bar{n}}\right)^{n/2} |n\rangle_I |n\rangle_S, \qquad (1)$$

where $\bar{n}$ is the mean photon number in each mode, $|n\rangle$ is an n-photon Fock state, and the subscripts I and S respectively denote the idler and signal modes. This state has perfect photon number correlations and can be produced to a good approximation by a nondegenerate SPDC source [13,16–18]. The average unconditioned state in the signal (or idler) mode is provided by tracing over the idler (or signal) mode and is a thermal state with mean photon number $\bar{n}$ [31,32]. The interbeam correlations ensure that measurements of the idler mode will condition the state of the signal mode. The conditional state can have an enhanced mean photon number, depending on the measurement outcome. However, due to the no-signaling theorem [33], the averaged state of the signal mode is still a thermal state with mean photon number $\bar{n}$. This is important when covertness is required as it means that on average, the signal mode has the same photon statistics as the background [13].

The TMSV state is not the only two-mode state with perfect photon number correlation. An alternative is to use the mixed state $\hat{\rho}_{IS} = (1-\lambda)\sum_n \lambda^n |n,n\rangle_{IS}\langle n,n|$, where

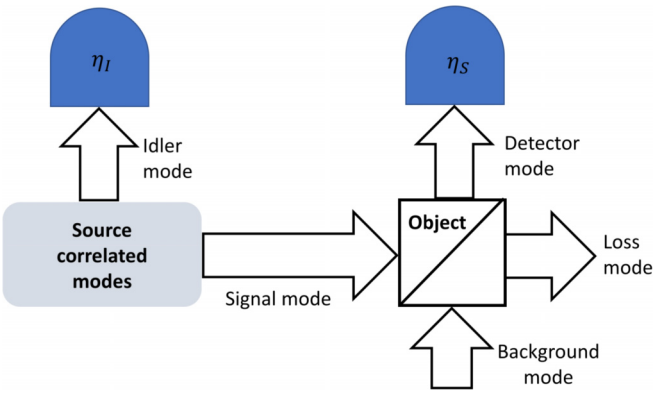

FIG. 1. A diagram illustrating direct detection illumination.

$\lambda = \bar{n}/(1+\bar{n})$. The reduced state for the signal mode is a thermal state with mean photon number $\bar{n}$, the same as for the TMSV state. Similarly, the conditional states are found to be the same as for the TMSV state, which are given in Eq. (A5) and (A8) of Appendix A. All of the results we present for the TMSV state would be the same if one used the mixed state $\hat{\rho}_{IS}$. For the sake of definiteness, we outline the protocol using the TMSV state (1). One interesting feature of both the TMSV and the mixed state is that the state of the signal mode conditioned on detecting photons in the idler mode, given in Eq. (A8) of Appendix A, can display sub-Poissonian photon statistics when $\bar{n}$ is low [26,34]. Furthermore, both the TMSV state and the mixed state have been shown to exhibit sub-Poissonian photon statistics in experiments [35]. This holds even if one uses single-photon detectors [36]. Additionally, the conditional state of the signal mode given in (A8) violates Klyshko's criterion for nonclassicality for all values of $\bar{n}$ [37]. Both the TMSV and the mixed state are thus nonclassical.

We consider the experimental setup shown in Fig. 1. The idler and signal detectors are inefficient threshold detectors with efficiency η_I and η_S, respectively. The target object is modeled by a beam splitter with reflectivity $\sqrt{\kappa}$. The thermal background is accounted for by injecting a thermal state with mean photon number $\bar{n}_B/(1-\kappa)$ into the other input port of the beam splitter so that no matter what its reflectivity, the background mean photon number that is emitted from the target is $\bar{n}_B$. In the absence of signal photons, this will be the mean number of photons incident on the signal detector. The threshold detectors are both described by POVMs, where for a detector with efficiency η the POVM element corresponding to the detector not firing is $\hat{\Pi}_0(\eta) = \sum_n (1-\eta)^n |n\rangle\langle n|$ and the POVM element for the detector to fire is $\hat{\Pi}_1(\eta) = \hat{1} - \hat{\Pi}_0(\eta)$. The no-fire POVM corresponds, in the Copenhagen interpretation, to collapse of the idler state onto the thermal state of mean photon number $(1-\eta)/\eta$ [38]. For unit efficiency this provides measurement in the vacuum state, but poorer detection efficiency provides collapse onto states of increasingly higher mean photon number. Dark counts within the signal detector can be taken account of within the background using the method outlined in Ref. [15]. In pulsed systems the idler detector can be gated so as to reduce dark counts to such an extent that they can be neglected.

The performance of quantum illumination for optimal measurements is determined using the Helstrom bound and the

quantum Chernoff bound [10,39,40]. For direct measurement illumination, the measurement data must be analyzed using an appropriate statistical method [23,41]. We consider a Bayesian approach because it allows one to easily incorporate any prior information that exists for whether an object is present within a given target location. However, here we consider only equal initial priors for all examples, i.e., $P(O) = P(\bar{O}) = 1/2$, where O denotes object present and $\bar{O}$ means no object. The reason for this is that we want to evaluate the performance of the protocols in the situation where we have no additional information.

In total, we transmit N copies of the TMSV state. We assume the copies are temporally separated into distinct time bins. This could be achieved experimentally by using a pulsed laser to pump a nonlinear crystal, but it is not strictly necessary because appropriate time bins can be defined for a cw source via an appropriate detection window at the idler detector. The parameters $\bar{n}$ and $\bar{n}_B$ are taken to be the mean number of photons per time bin. Within each time bin, we have measurement results for the idler and signal detectors. These detection probabilities are calculated in Appendix A. For the TMSV the probability for the idler detector to not register a click in one shot is

$$P_I(0) = \frac{1}{1+\eta_I \bar{n}}, \tag{2}$$

and the probability for it to fire is $P_I(1) = 1 - P_I(0)$. When there is no object present, the detection probability for the signal detector is independent of the idler detector. In this case, the probability for the signal detector to not fire is $P_S(0|\bar{O}) = 1/(1+\eta_S \bar{n}_B)$ and the probability to fire is $P_S(1|\bar{O}) = 1 - P_S(0|\bar{O})$. When an object is present, we denote the detection probabilities for the signal mode as $P_S(s|O)$, where $s \in \{0, 1\}$. In this case, the results for the signal mode are not independent of the idler measurement result and we must consider the joint probability $P_{IS}(i, s|O)$, where $i, s \in \{0, 1\}$. The form of the joint probability is rather long and thus we only state $P_{IS}(0, 0|O)$ and the marginal probability $P_S(0|O)$ (see Appendix A for more details),

$$\begin{aligned} P_S(0|O) &= \frac{1}{1+\eta_S(\kappa\bar{n}+\bar{n}_B)}, \\ P_{IS}(0,0|O) &= \frac{1}{1+\eta_I\bar{n}+\eta_S[\bar{n}_B+\eta_I\bar{n}_B\bar{n}+(1-\eta_I)\kappa\bar{n}]}. \end{aligned} \tag{3}$$

Let Ω_r denote the set of measurement outcomes after measurements on r time slots. After each measurement, we use Bayes' rule to update the posterior probability for the object to be present:

$$\begin{aligned} &P(O|\Omega_r) \\ &= \frac{P_{IS}(i_r, s_r|O)P(O|\Omega_{r-1})}{P_{IS}(i_r, s_r|O)P(O|\Omega_{r-1}) + P_{IS}(i_r, s_r|\bar{O})P(\bar{O}|\Omega_{r-1})}. \end{aligned} \tag{4}$$

To evaluate the performance of the protocol, we perform a Monte Carlo simulation. This entails generating a random set of measurement outcomes that on average satisfy the calculated measurement probabilities when either the target is present or not and then calculating the posterior probability for the object to be present. The results are then averaged over multiple runs; for more details see Refs. [15,26]. This approach shows the utility of the Bayesian approach, which becomes more apparent for multiple pulses for which the subsequent pulse priors are altered by earlier measurement results.

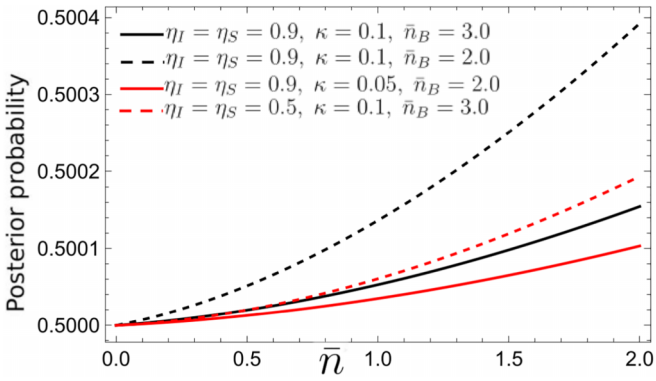

FIG. 2. A plot of the single copy averaged posterior probability for an object to be present, $\mathcal{P}(O)$, plotted against $\bar{n}$. The solid black line is for $\eta_I = \eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; the dashed black line is $\eta_I = \eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 2.0$; the solid red line (gray offline) is $\eta_I = \eta_S = 0.9$, $\kappa = 0.05$, and $\bar{n}_B = 2.0$; and the dashed red line (gray offline) is $\eta_I = \eta_S = 0.5$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$.

Suppose one is interested in the *relative* change in performance when a parameter is altered, i.e., does varying a parameter make the performance better or worse? We could evaluate this without a full Monte Carlo simulation; instead one can consider the average behavior of $P(O|i, s)$ for one copy of the TMSV state, $N = 1$. The averaged posterior probability is

$$\begin{aligned} \mathcal{P}(O) &= \sum_{i,s} P_{IS}(i, s|O)P(O|i, s) \\ &= \sum_{i,s} P_{IS}(i, s|O)\left[1 + \frac{P_S(s|\bar{O})}{P_{S|I}(s|i, O)}\right]^{-1}, \end{aligned} \tag{5}$$

where $P(O|i, s)$ is found from Eq. (4), $P_{IS}(i, s|O)$ is given in Eqs. (3) and (A10), and $P_{S|I}(s|i)$ is the conditional probability defined in Eq. (A7) of Appendix A. The quantity $\mathcal{P}(O)$ is a function of the parameters: $\bar{n}$, η_I, η_S, κ, and $\bar{n}_B$. However, the parameters κ and $\bar{n}_B$ are determined by the environment and are beyond our control. In contrast, we can control the source intensity, $\bar{n}$, and we should have some control over the efficiencies η_I and η_S. Equation (5) can be used to optimize the protocol by exploring the effect of changing these parameters. These effects will be small, as we are only working with one copy of the state. The changes in probability from the prior value of 0.5 will be tiny. However we see that they accrue with the number of copies (pulses) to provide measurable and significant differences.

To provide a basic understanding of how changing each parameter affects $\mathcal{P}(O)$, we plot Eq. (5) in Fig. 2 against $\bar{n}$ for various different values of κ, $\bar{n}_B$, η_I, and η_S. The solid black curve corresponds to $\eta_I = \eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; the dashed black curve is $\eta_I = \eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B =$

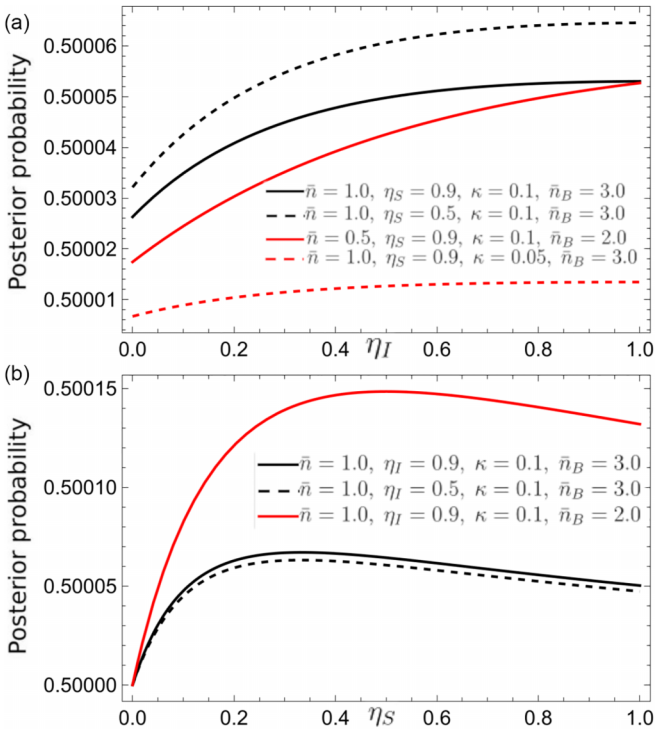

FIG. 3. A plot of the single copy averaged posterior probability for an object to be present, $\mathcal{P}(O)$, plotted against (a) η_I and (b) η_S.

2.0; the solid red (gray offline) curve is $\eta_I = \eta_S = 0.9$, $\kappa = 0.05$, and $\bar{n}_B = 2.0$; and the dashed red (gray offline) curve is $\eta_I = \eta_S = 0.5$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$. As expected, we see that increasing $\bar{n}$ improves the performance of the protocol. We also see that decreasing the detector efficiencies from 0.9 to 0.5 has an effect, but in relative terms, this is not as strong an effect as decreasing either $\bar{n}_B$ or κ. For example, the dashed black curve ($\bar{n}_B = 2.0$) reaches the value of $\mathcal{P}(O) \approx 0.50014$ at $\bar{n} = 1.0$, while the solid black curve ($\bar{n}_B = 3.0$) requires $\bar{n} = 1.85$, almost double the mean photon number. We also see that halving κ has an even stronger effect.

III. IMPROVING QUANTUM ILLUMINATION VIA DETECTOR MISMATCH

Intuitively, one would expect that the optimal performance of a quantum illumination system would occur for $\eta_I = \eta_S = 1$, i.e., for perfect detectors. This idea can be tested using Eq. (5). In Fig. 3(a) we plot $\mathcal{P}(O)$ as a function of the idler detector efficiency η_I. The solid black line corresponds to $\bar{n} = 1.0$, $\eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; the dashed black line is $\bar{n} = 1.0$, $\eta_S = 0.5$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; the solid red (gray offline) is $\bar{n} = 0.5$, $\eta_S = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 2.0$; and the dashed red (gray offline) is $\bar{n} = 1.0$, $\eta_S = 0.9$, $\kappa = 0.05$, and $\bar{n}_B = 3.0$. As expected, we see that increasing η_I always improves $\mathcal{P}(O)$. However, we sometimes see that the values tend to saturate, so that further improvement can be incremental. This is not always the case, as the solid red curve shows. However, when $\mathcal{P}(O)$ saturates it tells us that improving the efficiency η_I will provide only incremental gains. This is an important observation for the design of practical quantum illumination systems. We see that, under

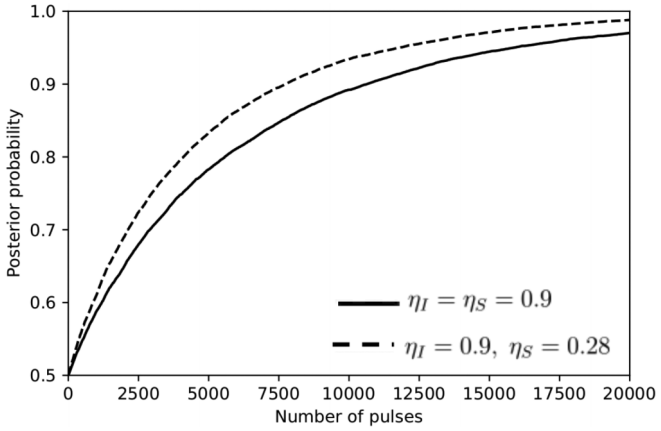

FIG. 4. A plot of the probability for the object to be present given we have transmitted N signals. All curves are for the case where an object is present and for the parameters: $\bar{n} = 2.0$, $\bar{n}_B = 3.5$, $\eta_I = 0.9$, $\kappa = 0.1$ and we average over 8000 runs. The solid curve is for $\eta_S = \eta_I = 0.9$, while the dashed curve is for $\eta_S = 0.28$.

certain circumstances, increasing η_I by using better detectors and improving the mode coupling into the detector, might not be worth the effort. For example, it is easy to show that for $\bar{n} = 1.0$, $\kappa = 0.1$, $\eta_S = 0.9$, and $\bar{n}_B = 3.0$, that the values of $\mathcal{P}(O)$ for $\eta_I = 0.8$ and $\eta_I = 1.0$ agree to six decimal places [approximately the linewidths in Fig. 3(a)]. A more extreme example occurs for $\bar{n} = 1.0$, $\kappa = 0.05$, $\eta_S = 0.9$, and $\bar{n}_B = 3.0$, where we again obtain agreement to six decimal places for $\mathcal{P}(O)$ with $\eta_I = 1.0$ and $\eta_I = 0.65$.

In Fig. 3(b), we plot $\mathcal{P}(O)$ as a function of η_S. The solid black line corresponds to $\bar{n} = 1.0$, $\eta_I = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; while the dashed black line is for $\bar{n} = 1.0$, $\eta_I = 0.5$, $\kappa = 0.1$, and $\bar{n}_B = 3.0$; and the solid red (gray in offline version) line is for $\bar{n} = 1.0$, $\eta_I = 0.9$, $\kappa = 0.1$, and $\bar{n}_B = 2.0$. Surprisingly, we see that, for each curve, there is a value for η_S that maximizes the averaged posterior probability. Beyond this maximum, increasing the efficiency will decrease $\mathcal{P}(O)$. We see that this behavior occurs for different values for η_I and $\bar{n}_B$. Numerical investigations also show that this behavior is also found for different values of κ. Nevertheless, there are parameters for which the maximum of $\mathcal{P}(O)$ occurs for $\eta_S = 1.0$. For the cases where a maximum does occur for $\eta_S < 1$, one can improve the performance of quantum illumination by using the optimal value for η_S. Heuristic approaches for understanding this effect will be presented later. For now, an important question to address is how great is this improvement in practice?

To determine the size of this effect for a repeated experiment, we perform a full Monte Carlo simulation for N pulses. In Fig. 4, we plot the probability for an object to be present against the number of pulses. In all curves, an object is present and we average the Monte Carlo simulation over 8000 runs. Both curves are for $\bar{n} = 2.0$, $\kappa = 0.1$, $\bar{n}_B = 3.5$, and $\eta_I = 0.9$. The solid curve is for $\eta_S = \eta_I = 0.9$, while the dashed curve is for $\eta_S = 0.28$. The figure shows that using $\eta_I \neq \eta_S$ can give a clear improvement in the posterior probability for an object to be present. For example, we reach a posterior probability of 0.9 for $N \approx 7700$ pulses, while $N \approx 10\,500$ pulses

are needed for $\eta_S = \eta_I = 0.9$. Using $\eta_S = \eta_I = 0.9$ required roughly 36% more pulses to achieve a posterior probability of 0.9 compared with using $\eta_S = 0.28$, a signal detector less than one third as efficient. The results show a that using the optimal value of η_S in Eq. (5) can significantly improve the performance of direct measurement quantum illumination.

The examples provided in Fig. 3 all showed that increasing η_I increases $\mathcal{P}(O)$. Of course, we found that sometimes the increase in $\mathcal{P}(O)$ would slow and start to saturate. For practical purposes, this is important because it allows one to use smaller values for η_I without a significant decreases in $\mathcal{P}(O)$. Nevertheless, we did not find any situation where increasing η_I would lead to a decrease in $\mathcal{P}(O)$. This was checked further by performing numerical investigations of the derivative of $\mathcal{P}(O)$ with respect to η_I, to see if we can find any stationary points within the interval $\eta_I \in (0, 1]$. For the range of parameters $\bar{n}$, $\bar{n}_B \in (0, 5]$, $\eta_S \in (0, 1]$, and $\kappa \in (0.001, 0.2]$, we did not find any stationary points. Furthermore, we found the derivative was always positive. This suggests that the optimal choice for η_I is always unity.

One can understand why choosing $\eta_I = 1$ is optimal by considering the effects of the idler measurement on the state of the signal mode. An outcome i for the idler measurement will condition the state of the signal mode to be $\hat{\rho}_{S|i}$ with probability $P_I(i)$, where $i = 0$ or 1. If there is no object, then detection probability at the signal is independent of the conditioning. However, if an object is present, then conditioning on the idler detector will affect the signal mode. For conditioning to have a strong effect, we want the outcomes to be as distinct as possible. One way of measuring the distinguishability is

$$\Delta_T = ||P_I(0)\hat{\rho}_{S|0} - P_I(1)\hat{\rho}_{S|1}||, \tag{6}$$

where we use the trace norm, i.e., $||\hat{A}|| = \frac{1}{2}\mathrm{Tr}[(\hat{A}^\dagger\hat{A})^{1/2}]$. This quantity is related to the minimum error for discriminating between the two conditional signal states [39]. In Appendix B we show that for $\bar{n}$ fixed, Δ_T is maximized when $\eta_I = 1$. We see that using $\eta_I = 1$ leads to signal states that are as distinct as possible. This optimal distinctness improves the contrast between states at the signal detector when an object is present and when we have only background, which should increase $\mathcal{P}(O)$.

The values of η_I and η_S that yield the optimal performance can be found from (5). The previous discussion suggests that η_I should equal one. With this assumption, it is possible to find the maximum of $\mathcal{P}(O)$ as a function of η_S analytically, with all other parameters treated as constants. This is achieved using elementary calculus, which after some algebra, yields the result

$$\eta_S^{(\mathrm{opt})} = \sqrt{\frac{2\bar{n}_B + \kappa(1+\bar{n})}{\bar{n}_B\left[2\bar{n}_B^2 + \kappa\bar{n}_B(3\bar{n}-1) + \kappa^2\bar{n}(\bar{n}-1)\right]}}. \tag{7}$$

This result gives the optimal value of η_S for use in the direct measurement protocol. As an example, consider the special case where $\bar{n} = 1$. We find that $\eta_S^{(\mathrm{opt})} = 1/\bar{n}_B$, which means that for $\bar{n}_B = 3$ we find that $\eta_S^{(\mathrm{opt})} = 1/3$.

For certain values of $\bar{n}$, κ, and $\bar{n}_B$, we find that (7) will either equal or exceed the value one. In these instances, the optimal value for η_S is one. We can thus use (7) as a condition for determining when we can obtain an improvement with $\eta_S < 1$. This is explored in Fig. 5, were we plot the minimum value of $\bar{n}_B$ for which $\eta_S^{(\mathrm{opt})} < 1$, against $\bar{n}$, for different values of κ. The solid black curve is for $\kappa = 0.1$, the dashed black curve is for $\kappa = 0.05$, and the solid red (gray offline) curve is for $\kappa = 0.01$. A point on or above these curves shows that range of values for $\bar{n}_B$ such that $\eta_S^{(\mathrm{opt})} < 1$. As such, the curves show the region in parameter space where an inefficient signal detector improves performance. We see that as κ decreases, the curves tend towards a constant value, which is independent of $\bar{n}$. This can be seen directly by taking the limit of (7) as $\kappa \to 0$, which yields the result that $\eta_S^{(\mathrm{opt})} = 1/\bar{n}_B$. All three curves show that, for sufficiently low numbers of background photons, the optimal performance occurs when $\eta_S = \eta_I = 1$. For $\kappa = 0.1$, i.e., the solid black curve, we see that increasing $\bar{n}$ decreases the minimum number of background photons needed for $\eta_S^{(\mathrm{opt})} < 1$. This suggests that the key feature in this phenomenon is the number of photons incident on the signal detector.

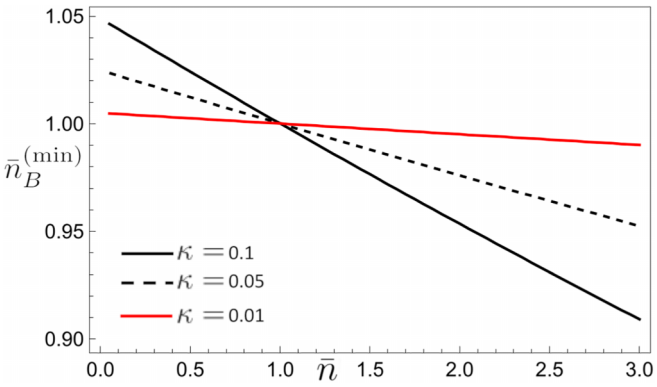

FIG. 5. A plot showing how $\bar{n}_B^{(\min)}$, the minimum value for $\bar{n}_B$ for which $\eta_S^{(\mathrm{opt})} < 1$, varies with $\bar{n}$ for different values of κ. All plots are for $\eta_I = 1$. The solid black curve is for $\kappa = 0.1$, the dashed black curve is for $\kappa = 0.05$ and the solid red (gray offline) curve is for $\kappa = 0.01$.

To understand why we sometimes have a maximum in $\mathcal{P}(O)$ for $\eta_S < 1$, we investigate further how the protocol works. An outcome $i \in \{0, 1\}$ for the idler detector conditions the signal mode to be in the state $\hat{\rho}_{S|i}$ with probability $P_I(i)$. If there were no object present, then the mode incident on the signal detector would be in a thermal state, $\hat{\sigma}_{\bar{n}_B}$, with mean photon number $\bar{n}_B$. Alternatively, when an object is present, then the state of the signal detector's mode is

$$\hat{\rho}_{S|i,O} = \hat{U}_\kappa\left(\hat{\rho}_{S|i} \otimes \hat{\sigma}_{\bar{n}_B/(1-\kappa)}\right)\hat{U}_\kappa^\dagger, \tag{8}$$

where $\hat{U}_\kappa$ is a unitary of the beam splitter that describes the object. To be able to determine if an object is present, we require $\hat{\sigma}_{\bar{n}_B} - \hat{\rho}_{S|i,O} \neq 0$ for at least one of the conditional states.

To quantify how difficult it is to determine if an object is present, we need a quantity that depends on $\hat{\sigma}_{\bar{n}_B} - \hat{\rho}_{S|i,O}$, such as the trace distance of this operator. The trace distance is related to minimum error state discrimination [39]. However, the signal detector does not perform an optimal minimum error measurement. As such, we propose a different figure of merit for quantifying the difficulty in determining if an object

is present

$$\Delta = \left| \sum_i \mathrm{Tr}\left[P_I(i) \left(\hat{\sigma}_{\bar{n}_B} - \hat{\rho}_{S|i,O} \right) [\hat{\Pi}_1(\eta_S) - \hat{\Pi}_0(\eta_S)] \right] \right|, \quad (9)$$

where the term $\hat{\Pi}_1(\eta_S) - \hat{\Pi}_0(\eta_S)$ takes account of the fact that the information gained from the measurement depends on how different the two measurement operators are. In Appendix C we explicitly calculate Δ and determine a necessary and sufficient condition for it to have a maximum for $\eta_S \in (0, 1)$. A maximum only occurs within this interval when $\bar{n}_B(\kappa\bar{n} + \bar{n}_B) > 1$, i.e., when the flux of photons is sufficiently great. In particular, there is always a maximum for $\eta_S \in (0, 1)$ when the background mean photon number is larger than unity, whatever the value of the object reflectivity or the signal field strength. When such a maximum exists, then it occurs at the value $\eta_S = \sqrt{1/(\bar{n}_B[\kappa\bar{n} + \bar{n}_B])}$. While this is not equal to (7), it is very close for typical values for $\bar{n}$ and $\bar{n}_B$ that satisfy the respective conditions for the existence of a maximum.

An alternative way of understanding why a maximum can occur for $\eta_S \in (0, 1)$ is found by considering the probabilities $P_{IS}(i, s|O)$ and $P_{IS}(i, s|\bar{O})$. If these two joint probabilities are too similar, then it is difficult to determine whether an object is present. This is because in Bayes' rule, the posterior probabilities $P(O|i, s)$ and $P(\bar{O}|i, s)$ depend respectively on the joint probabilities $P_{IS}(i, s|O)$ and $P_{IS}(i, s|\bar{O})$. One can thus get a qualitative feel for our ability to distinguish if an object is present by investigating the difference between the two joint probabilities. This difference can be measured using the $\mathcal{L}_1$ distance [42,43]. In Appendix C we calculate the $\mathcal{L}_1$ distance between $P_{IS}(i, s|O)$ and $P_{IS}(i, s|\bar{O})$. We find that it is the same as Δ in Eq. (9).

These two approaches both confirm the results shown in Fig. 5: that a maximum occurs for $\eta_S \in (0, 1)$ only when the flux of photons on the signal detector is sufficiently high. Intuitively, if the flux of photons on the signal detector is too great, then the signal detector will fire very frequently and, if a target is present, changes in light levels at the signal detector due to idler conditioning do not affect the signal detector firing probability very much. A decrease in η_s, while lowering the probability of detecting a possible signal photon, increases the information gain when the detector either fires or does not. The situation is analogous to when η_S is very small. We find then that the signal detector will fire infrequently, which makes it difficult to gain information on whether an object is present. Increasing η_S then improves the situation.

IV. OPTIMIZING PERFORMANCE PER TRANSMITTED PHOTON

To evaluate the performance of the protocol, we have previously examined how the posterior probability for an object to be present grows with the number of pulses. The importance of the number of pulses is that it is a proxy for the time taken to identify if an object is present. Let ΔT be the time between the center of each pulse and let τ be the pulse width. If N pulses are required to achieve a satisfactory level of confidence, then this corresponds to a time period of $N\Delta T$. One way of reducing the time is to use pulses with a shorter separation, which enables more pulses to be transmitted within the same time period. If the new pulses have the same mean photon number, then the number of photons transmitted will have increased. One would expect that transmitting more photons should improve the performance. A more interesting question is whether the performance can be improved without increasing the number of transmitted photons? One approach to this is to assume that the total number of photons per second is fixed. Using pulses with shorter separations would then necessitate a decrease in the mean photon number of each pulse. We assume that this is achieved by decreasing the pulse width such that the mean photon number per pulse is decreased. For example, if we halve the separation ΔT, then to keep the rate fixed, we also halve τ such that $\bar{n}/2$ signal photons on average are transmitted in each pulse. The situation is illustrated in Fig. 9 of Appendix D.

Suppose that we initially use pulses with mean photon number $\bar{n}$ and separation ΔT such that there are $\bar{n}_B$ background photons per time bin. We replace this with pulses of separation $\Delta T/m$ and mean photon number of $\bar{n}/m$. This is achieved by reducing the pulse width to τ/m. One can thus reduce the time that the signal detector is gated on. This reduction in time will reduce $\bar{n}_B$, the number background photons detected within the time bin. The number of background photons in each time bin is now taken to be $\bar{n}_B/m$. The following analysis does not hold if the pulse width is held constant, unless we are not gating the detectors. In that case the mean number of background photons per time bin depends only on the size of the time bin, which equals the spacing of the pulses. For more details of this analysis, see Appendix D.

To determine the effect of this, one can look at the posterior probability averaged over m consecutive time bins, which are of total length ΔT and so still contain on average $\bar{n}$ photons. For $m = 1$, this corresponds to (5), while for $m > 1$ the averaged posterior probability is

$$\mathcal{P}^{(m)}(O) = \sum_{\Omega_m} P(\Omega_m|O)P(O|\Omega_m), \quad (10)$$

where Ω_m is a possible set of m outcomes for the idler and signal detectors, and the summation is taken to be over all possible sets of m outcomes. The number of terms within the summation in (10) scales as 4^m. Evaluating the sum for large values of m is impractical but can easily be approximated using Monte Carlo simulations. However, for small values of m we can analytically evaluate the summation, as we did for the case of $m = 1$ in (5). Nevertheless, the number of terms soon grows and, as such, we use the *Mathematica* software package to aid us in this task [44]. Investigating how $\mathcal{P}^{(m)}$ depends on m allows us to determine whether there is an advantage to using pulses with both shorter widths and separations, even when the rate of transmitted photons per second is fixed.

In Fig. 6 we plot the averaged posterior probability against m for different values of $\bar{n}/m$ and $\bar{n}_B/m$. For all plots, $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, the black line is for the quantum illumination scheme and the red line (gray offline) is for when we transmit coherent states with mean photon number $\bar{n}/m$. In Fig. 6(a) $\bar{n} = 0.1$ and $\bar{n}_B = 0.1$, and in Fig. 6(b) $\bar{n} = 1.0$ and $\bar{n}_B = 3.0$. Both Figs. 6(a) and 6(b) show that, for direct measurement quantum illumination, increasing m improves the averaged posterior probability for the object to be present. This shows

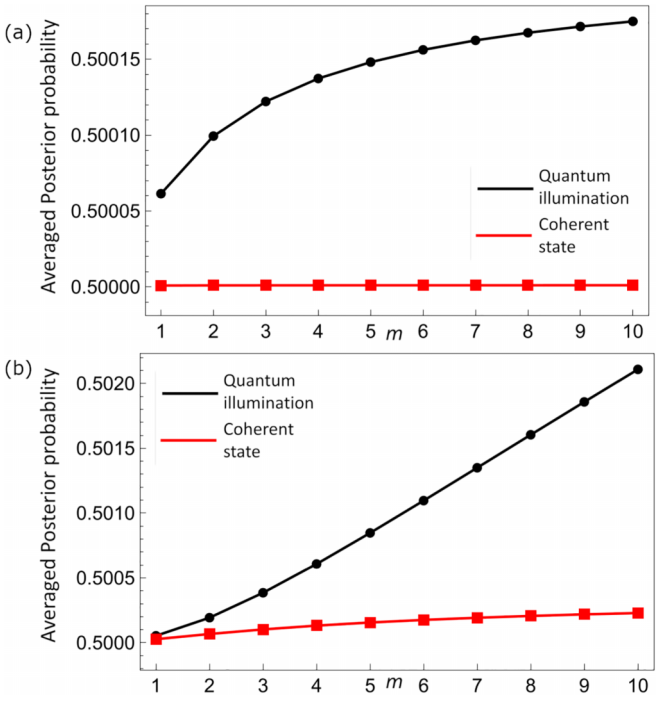

FIG. 6. A plot of the averaged posterior probability for an object to be present against m, the number of copies of the state. All plots are for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, and for the case when an object is present. Panel (a) is for $\bar{n} = 0.01$ and $\bar{n}_B = 0.1$, while panel (b) is for $\bar{n} = 1.0$ and $\bar{n}_B = 3.0$. In both plots, the black curves are for a direct detection quantum illumination, while the red (gray offline) are for transmitting coherent states.

that using more pulses gives an advantage, even if doing so decreases the mean photon number of each pulse. However, in Fig. 6(a) we see that the rate of increase for the posterior probability, decreases with m, which leads to $\mathcal{P}^{(m)}(O)$ starting to plateau. This suggests that for given values of the parameters, such as $\bar{n}$, $\bar{n}_B$, η_I, η_S, and κ, there is a limiting value for $\mathcal{P}^{(m)}(O)$. In practice, if the separation, $\Delta T/m$, is too small, detector jitter will become important and will start to limit $\mathcal{P}^{(m)}(O)$. A more sophisticated analysis is required to find the trade-off, one in which jitter is included [45].

In both Figs. 6(a) and 6(b), we compare the performance of quantum illumination against transmitting a coherent state. We make the same assumption about coherent pulses that both the width and separation are decreased and that the signal detector can be gated on for less time so as to decrease the background. Both plots confirm the previous findings [15,26] that direct measurement protocols can outperform coherent states under appropriate conditions. Furthermore, Figs. 6(a) and 6(b) also show how the posterior probability of coherent states scales with the mean photon number, $\bar{n}/m$, of each pulse. We see that increasing m has either no effect as in Fig. 6(a) or a very weak effect as shown in Fig. 6(b). This is in stark contrast to quantum illumination, where increasing m has a relatively strong effect. This should not be surprising as previous research has established that direct measurement quantum illumination schemes work well when the number of signal photons is much less than the background, but this is not true for coherent states [15,26]. Increasing m increases the number of pulses per second, but decreases the number of photons per pulse. For quantum illumination, this provides more pulses while keeping us within a regime for which each pulse still provides information on whether an object is there or not. In contrast, for coherent states, the increase in the number of pulses is partially balanced by the decrease in the information each pulse provides. This idea of information gained per photon can be made more rigorous, see Appendix E for a discussion on this topic.

Another way of understanding this is in terms of the signal-to-noise ratio (SNR). For coherent states, the mean number of signal photons per pulse is $\kappa\bar{n}/m$, while the mean number of background photons per time bin is $\bar{n}_B/m$, when we match the detectors temporal window to the pulse width. For this case, the SNR is $\kappa\bar{n}/\bar{n}_B$, which is independent of m. This explains why increasing m does not provide a benefit for coherent states.

In contrast, for direct detection illumination, the idler conditions the signal to be in one of two different states. The results of the idler measurement indicates which state was transmitted and thus we can partition the events into two different cases, each with different SNR. If we again match the detector's temporal widow to the pulse width, then in both cases, the mean number of background photons per time bin is again $\bar{n}_B/m$. The first case is when the idler fires. For $m = 1$, the conditional mean number of signal photons is $\bar{n}_1 = \bar{n} + (1 + \bar{n})/(1 + \eta_I\bar{n})$ [26]. In general, the SNR is $\kappa\bar{n}/\bar{n}_B + \kappa(m + \bar{n})/[\bar{n}_B(1 + \eta_I\bar{n}/m)]$, which increases with m. As m increases, $\bar{n}/m$ will eventually grow smaller, which makes the SNR approximately $\kappa(\bar{n} + m)/\bar{n}_B$. In the limit where $\bar{n}$ is very small, the SNR is approximately $\kappa m/\bar{n}_B$. In each instance we see a clear growth in the SNR with m. The other case is when the idler does not fire. The SNR is now $\kappa\bar{n}(1 - \eta_I)/[\bar{n}_B(1 + \eta_I\bar{n}/m)]$, which in the limit of $\bar{n}/m \ll 1$, is approximately $\kappa(1 - \eta_I)\bar{n}/\bar{n}_B$. The key point is that the SNR for when the idler fires, grows with m. One can average together the SNRs for each case. This gives an indication of the performance of the system if we did *not* use the idler information when calculating the posterior probability. When $\bar{n} \ll 1$, increasing m will eventually lead to a point where the total probability for the idler to fire becomes small. This will balance the increase in the SNR when the idler fires. The observation explains the behavior found in Fig. 6(a), where $n = 0.01$. We see a growth in the posterior probability in m, until the probability starts to flattens out.

Figure 6 demonstrates that increasing the number of pulses, while keeping the rate of photons transmitted per second fixed, can provide an advantage in direct measurement based quantum illumination. To quantify how great this advantage is we again perform Monte Carlo simulations. In Fig. 7 we show the averaged results of a Monte Carlo simulation with 8000 runs. In particular, we plot the averaged posterior probability against the mean of the total number of photons transmitted. All curves are for $\eta_I = \eta_S = 0.9$, $\kappa = 0.1$, $\bar{n} = 1.0$, and $\bar{n}_B = 3.0$. The solid black curve is for $m = 1$, the dashed curve is for $m = 2$ and the dotted curve is for $m = 3$. We see that increasing m has a significant effect. For example, to reach a posterior probability of 0.9 requires approximately 20,200 photons for $m = 1$, while only ≈5800 photons are required for $m = 2$, and ≈2900 photons are needed for $m = 3$. Recall, that for all three

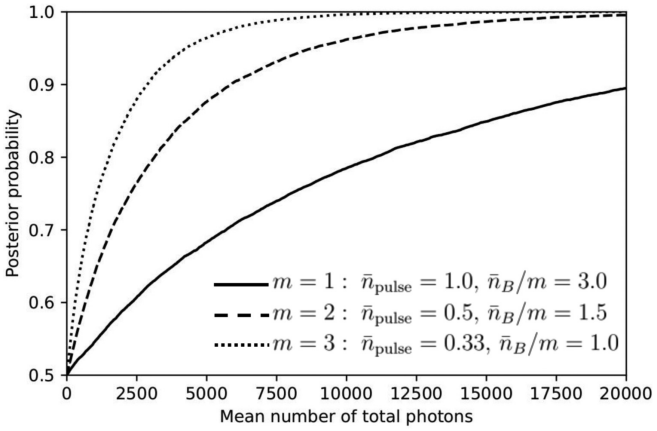

FIG. 7. A plot of the averaged posterior probability for an object to be present the total number of photons transmitted. All curves are for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$. The solid curve is for $m = 1$, $\bar{n} = 1.0$, and $\bar{n}_B = 3.0$. The dashed curve is for $m = 2$, $\bar{n}/2 = 0.5$, and $\bar{n}_B/2 = 1.5$, while the dotted curve is for $m = 3$, $\bar{n}/3 = 0.33$, and $\bar{n}_B/3 = 1.0$.

curves in Fig. 7, the flux is the same, so the horizontal axis is a proxy for time. If we set our confident detection threshold at 0.9 then for $m = 3$ we would detect an object in around $1/7$ of the time taken for $m = 1$. Thus using fewer photons translates into acquiring the same information faster. Using less intense pulses with shorter duration, in principle, provides a clear benefit. In practice, one must be careful not to make m so great such that detector jitter starts to become the dominant source of noise.

Why is there such a benefit to increasing the pulse rate while keeping the intensity constant? There are three effects at play here but one is dominant. First the probability of the idler detector firing per unit time increases so that slightly more signal pulses contain a postselection-increased number of photons. Against this, when the idler detector fires the increase in signal photon number is itself marginally smaller, so these two effects counteract (and both vanish at $\eta_I = 1$). By far the dominant effect is based on the constant rate of background photons, which means that the background mean photon number per pulse decreases to $\bar{n}_B/m$. The system now has the easier task of searching for roughly the same fraction of postselected signal photons as before, but against an enfeebled background. The overall result is to provide the better than $1/m$ decreases in time taken to reach any imposed object detection probability threshold in Fig. 7.

If we had decreased the mean photon number per pulse without also decreasing the pulse width, then $\bar{n}_B$ would not be reduced. In that case the results would be changed. For example, in Fig. 6(a), if $\bar{n}_B$ was unchanged, then black curve, which corresponds to the direct measurement illumination, would not increase with m and is flat. In contrast, the red (gray offline) curve, which corresponds to coherent states, would decrease with m. The improvement in performance is thus partially related to the decrease in background. The background can also be reduced by using mode filtering [29]. However, for direct detection illumination, we have the added complication that the idler measurement results in two possible signal states.

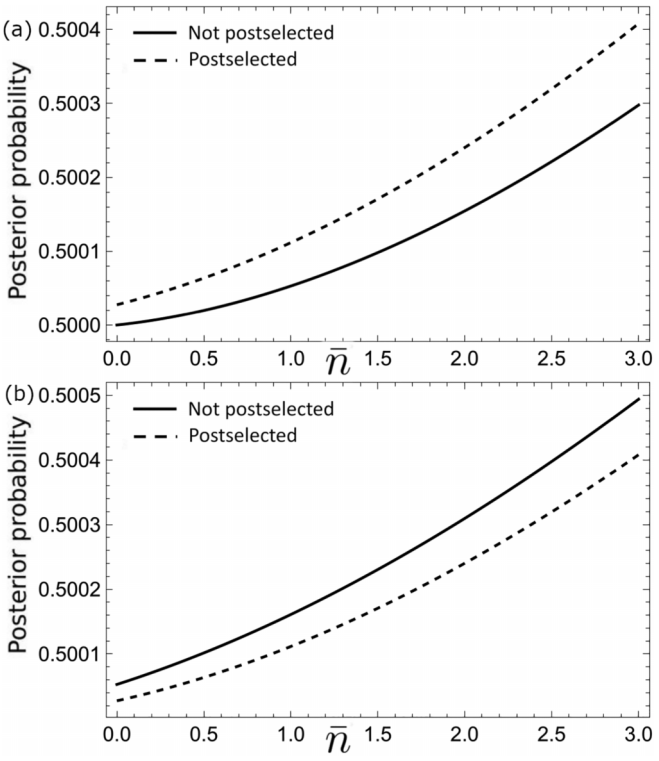

FIG. 8. A plot comparing the averaged posterior probability for the standard protocol and a protocol with postselection. For the postselected protocol, we neglect the failure probability. A more realistic assessment of postselection is found by multiplying the postselected curve by $P_I(1)$. This significantly decreases the postselected curves such that they are all below the curves for the standard protocol by such an amount that they would not be visible using the current axes. All curves in panels (a) and (b) are for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, $\bar{n}_B = 3.0$, and for the case where an object is present. In panel (a), both curves are for the same TMSV state with mean photon number $\bar{n}$. In panel (b), the postselected curve is for a TMSV state with mean photon number $\bar{n}$, while the standard protocol is for a TSMV state with mean photon number $\bar{n} + (1 + \bar{n})/(1 + \eta_I \bar{n})$.

The analysis in this and the previous section involves averaging over all possible measurement outcomes. However, it is common in many quantum optics experiments to postselect on a given outcome. For direct measurement illumination, one could postselect on the idler detector firing. To study the effects of this, we evaluate the average posterior probability for $N = 1$, with postselection. Let $\mathcal{P}(O|i = 1)$ denote the average posterior probability to detect an object given one is present, when we postselect on the idler detector firing. This is defined in Eq. (F1) of Appendix F. In Fig. 8(a), we compares the results of (5) against (F1) for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, and $\bar{n}_B = 3.0$ and for an object present. We see that that postselection does seem to improve the performance. However, Eq. (F1) neglects the fact that, on average, order $1/P_I(1)$ pulses are required for the idler to fire. To take account of the failure probability, we multiply (F1) by the probability for the idler detector to fire. The new averaged posterior probability $\mathcal{P}(O, i = 1)$ is defined in Eq. (F2) of Appendix F. By comparing (5) with (F2), we can verify that $\mathcal{P}(O, i = 1) \leqslant \mathcal{P}(O)$, i.e., postselection makes the performance worse. However,

even if one neglected the cost of postselection, the result in Fig. 8(a) would still be misleading as the mean photon number for the postselected protocol is no longer equal to $\bar{n}$, that of the TMSV state. Instead, it equals the mean photon number of the postselected conditional state (A8). In Fig. 8(b) we account for this discrepancy and find that the nonpostselected protocol is now superior to postselection, even when we ignore the postselection probability. For more details, see Appendix F.

V. CONCLUSIONS

We have investigated the optimization of direct measurement based illumination, in which the idler and signal modes are measured separately using threshold detectors. The performance of such a system depends on several system parameters such as the efficiency of the idler detector, the efficiency of the signal detector, the mean number of signal photons and the spacing between pulses. To optimize the protocol, we examined the effects of changing these system parameters. This was achieved by deriving a simple analytic expression, which came from the average performance of the protocol for a single set of idler and signal measurements.

Surprisingly, we found that, when the background count probability is high, it can be beneficial to use a signal detector with efficiency less than one. For example, an efficiency of 0.33 can be optimal for the signal detector when the mean number of background photons was three per time bin and the reflectance of the object was 0.1. The size of this effect was demonstrated by performing a full Monte Carlo simulation for N measurements. The results show that optimizing the signal detector efficiency could significantly reduce the number of pulses needed to determine if a target object is present. Note that, for a sufficiently bright background, the optimal signal detector efficiency is less than 1, no matter what the reflectivity of the target is, nor the mean photon number of the signal beam. This could be an important result for quantum lidar systems, for which the effective target reflectivity is often limited by geometric factors, and for which weak optical nonlinearity means that TMSV generating systems are inefficient. In contrast, we also found that increasing the idler detector efficiency always improved the performance of the protocol. However, sometimes the increase in performance would saturate. This meant that for all practical purposes, a lower value for the idler detector efficiency would be sufficient.

We also investigated the interplay between the number of photons vs number of pulses needed to determine if an object is present. In particular, we considered shortening both the duration and separation of the pulses while keeping the number of transmitted photons per second constant. This meant that if we transmitted the pulses more frequently, then the mean photon number of each pulse would be decreased. We found that increasing the frequency of pulses does improve the performance of direct measurement based illumination, provided the one also reduces the gating time of the detectors. This was in contrast with using coherent pulses, which were found to either not increase, or increase very slowly before saturating. The results show that using higher repetition rates for quantum illumination sources will provide an advantage even if the total number of photons transmitted is not increased.

Finally, we looked at the effect of postselecting on the idler detector firing. Provided one neglects the postselection probability, then postselection was found to improve the performance of the protocol. However, the improvement was achieved by effectively increasing the mean number of signal photons per time bin. When this effect was accounted for, then a comparison of postselected illumination with the standard protocol showed that the standard protocol performed better.

The results in this paper should help to improve experimental implementations of direct measurement based illumination. In particular, Eq. (5) can be used to determine reasonable values for the idler detector efficiency and the optimal choice for the signal detector's efficiency. Our analysis also points to an advantage in increasing repetition rates, without the need to increase power. The results thus provide a framework for the analysis and improvement of experimental implementations of direct measurement based illumination. Together with recent experimental such as [14], our work helps point the way towards the deployment of direct measurement based illumination.

ACKNOWLEDGMENTS

We acknowledge financial support from the UK Engineering and Physical Sciences Research Council for funding via the UK National Quantum Technology Programme and the QuantIC Imaging Hub (Grant No. EP/T00097X/1). We also acknowledge important discussions with N. Samantaray and H. Yang on quantum illumination.

APPENDIX A: CALCULATION OF DETECTION PROBABILITIES

In this Appendix we calculate the detection probabilities. The approach will use the results of Refs. [26] and [15]. A thermal state with mean photon number $\bar{m}$ has the form

$$\hat{\sigma}_{\bar{m}} = \frac{1}{1+\bar{m}} \sum_{n=0}^{\infty} \left(\frac{\bar{m}}{1+\bar{m}} \right)^n |n\rangle\langle n|, \tag{A1}$$

The probability for the idler detector to fire is $P_I(0) = \mathrm{Tr}_{IS}[|\Psi\rangle_{IS}\langle\Psi|\hat{\Pi}_0(\eta_I) \otimes \hat{1}]$. Taking the partial trace over the signal mode gives $P_I(0) = \mathrm{Tr}_I[\hat{\sigma}_{\bar{n}}\hat{\Pi}_0(\eta)]$, where $\hat{\sigma}_{\bar{n}}$ is a thermal state, which is formed by tracing over the signal mode of the state (1). A straightforward calculation yields Eq. (2).

Consider now the probability for the signal detector to not fire. If there is no object present, then we have only the thermal background incident on the signal detector. The probability to not register a click is $P_S(0|\bar{O}) = 1/(1+\eta_S\bar{n}_B)$. When an object is present, then the probability to not register a click is found using the reduced state for the signal mode, which is a thermal state with mean photon number $\bar{n}$. We now prove a useful result that we use repeatedly in this Appendix.

A general thermal state, with mean photon number $\bar{m}$, can be represented in terms of coherent states

$$\hat{\sigma}_{\bar{m}} = \frac{1}{\pi\bar{m}} \int e^{-|\alpha|^2/\bar{m}} |\alpha\rangle\langle\alpha| d^2\alpha, \tag{A2}$$

where $|\alpha\rangle$ is a coherent state. Suppose the mode that is incident on the object, as depicted in Fig. 1, is a thermal state with

mean photon number $\bar{m}$. Let $\hat{U}_\kappa$ be the unitary of the beam splitter that describes the object. The probability $P_S(0|O)$ is

$$P_S(0|O) = \mathrm{Tr}\big[\hat{U}_\kappa\big(\hat{\sigma}_{\bar{m}} \otimes \hat{\sigma}_{\bar{n}_B/(1-\kappa)}\big)\hat{U}_\kappa^\dagger \hat{1} \otimes \hat{\Pi}_0(\eta_S)\big] = \frac{1}{\pi\bar{m}} \int e^{-|\alpha|^2/\bar{m}} \mathrm{Tr}\big[\hat{U}_\kappa\big(|\alpha\rangle\langle\alpha| \otimes \hat{\sigma}_{\bar{n}_B/(1-\kappa)}\big)\hat{U}_\kappa^\dagger \hat{1} \otimes \hat{\Pi}_0(\eta_S)\big] d^2\alpha. \tag{A3}$$

The trace factor inside the integral is the probability for the signal detector to not fire, given the input was the coherent state $|\alpha\rangle$. This probability is calculated in Appendix A of Ref. [15]. Using the previous results we find that

$$P_S(0|O) = \frac{1}{\pi\bar{m}} \int e^{-|\alpha|^2/\bar{m}} \frac{1}{1+\eta\bar{n}_B} \exp\left(\frac{-\eta\kappa|\alpha|^2}{1+\eta\bar{n}_B}\right) d^2\alpha = \frac{1}{1+\eta_S[\kappa\bar{m}+\bar{n}_B]}, \tag{A4}$$

and $P_S(1|O) = 1 - P_S(0|O)$. A straightforward application of this result is to calculate $P_S(0|O)$ when our input state is the reduced state of a TMSV. From (A4) we find that $P(0|O) = 1/(1+\eta_S[\kappa\bar{n}+\bar{n}_B])$.

The joint probability $P_{IS}(i, s|O)$ will be calculated by looking at the case where both the idler and signal detectors do not fire. The state of the idler mode conditioned on the idler detector not firing is

$$\hat{\rho}_{S|0} = \frac{\mathrm{Tr}_I[|\Psi\rangle_{IS}\langle\Psi|\hat{\Pi}_0(\eta_I) \otimes \hat{1}]}{\mathrm{Tr}_{IS}[|\Psi\rangle_{IS}\langle\Psi|\hat{\Pi}_0(\eta_I) \otimes \hat{1}]} = \sigma_{\bar{m}_0}, \tag{A5}$$

where

$$\bar{m}_0 = \frac{\bar{n}(1-\eta_I)}{1+\eta_I\bar{n}}. \tag{A6}$$

The conditional state is a thermal state with mean photon number $\bar{m}_0$. The probability for the signal detector to not fire, conditioned on the idler not firing, is

$$P_{S|I}(0|0, O) = \mathrm{Tr}\big[\hat{U}_\kappa\big(\hat{\sigma}_{\bar{m}_0} \otimes \hat{\sigma}_{\bar{n}_B/(1-\kappa)}\big)\hat{U}_\kappa^\dagger \hat{1} \otimes \hat{\Pi}_0(\eta_S)\big] = \frac{1}{1+\eta_S[\kappa\bar{m}_0+\bar{n}_B]}, \tag{A7}$$

where the second line follows by the same argument that led to (A4). Using the definition $P_{IS}(0, 0|O) = P_{S|I}(0|0, O)P_I(0)$ and some straightforward algebra, we find that $P_{IS}(0, 0|O)$ is given by (3).

If instead the idler detector fired, then the signal mode would be prepared in the state

$$\hat{\rho}_{S|1} = \frac{\mathrm{Tr}_I[|\Psi\rangle_{IS}\langle\Psi|(\hat{1} - \hat{\Pi}_0(\eta_I)) \otimes \hat{1}]}{P_I(1)} = \frac{1+\eta_I\bar{n}}{\eta_I\bar{n}}\sigma_{\bar{n}} - \frac{1}{\eta_I\bar{n}}\sigma_{\bar{m}_0}, \tag{A8}$$

where σ_x denotes a thermal state with mean photon number x. A straightforward application of (A4) yields the conditional probability

$$P_{S|I}(0|1, O) = \frac{1}{\eta_I\bar{n}}\left(\frac{1+\eta_I\bar{n}}{1+\eta_S[\kappa\bar{n}+\bar{n}_B]} - \frac{1}{1+\eta_S[\kappa\bar{m}_0+\bar{n}_B]}\right). \tag{A9}$$

The other conditional probabilities can be found using $P_{S|I}(1|0, O) = 1 - P_{S|I}(0|0, O)$ and $P_{S|I}(1|1, O) = 1 - P_{S|I}(0|1, O)$. Alternatively, one can use the definitions for the marginal probabilities: $P_I(i) = \sum_s P_{IS}(i, s|O)$ and $P_S(s|O) = \sum_i P_{IS}(i, s|O)$, to obtain

$$\begin{aligned} P_{IS}(0, 1|O) &= P_I(0|O) - P_{IS}(0, 0|O), \\ P_{IS}(1, 0|O) &= P_S(0|O) - P_{IS}(0, 0|O), \\ P_{IS}(1, 1|O) &= 1 - P_I(0) - P_S(0|O) + P_{IS}(0, 0|O), \end{aligned} \tag{A10}$$

where the last equation follows from $P_{IS}(1, 1) = 1 - P_{IS}(0, 0) - P_{IS}(0, 1) - P_{IS}(1, 0)$. The joint probabilities $P_{IS}(i, s|O)$ and $P_{IS}(i, s|\bar{O})$ can now all be evaluated.

APPENDIX B: MAXIMIZING THE DISTINGUISHABILITY OF THE CONDITIONAL SIGNAL STATES

In this section we calculate (6) and show that, for $\bar{n}$ fixed, Δ_T is maximum when $\eta_I = 1$. This shows that the conditional states of the signal mode are most distinguishable when the idler detector has perfect efficiency, i.e., $\eta_I = 1$. Using the results of Appendix A, in particular Eqs. (A5), (A7), (A8), and (A9), we find that

$$\begin{aligned} P_I(0)\hat{\rho}_{S|0} &= \frac{1}{1+\bar{n}} \sum_{k=0}^{\infty} (1-\eta_I)^k \lambda^k |k\rangle\langle k|, \\ P_I(1)\hat{\rho}_{S|1} &= \frac{1}{1+\bar{n}} \sum_{k=0}^{\infty} \lambda^k [1-(1-\eta_I)^k] |k\rangle\langle k|, \end{aligned} \tag{B1}$$

where $|k\rangle$ are Fock states and $\lambda = \bar{n}/(1+\bar{n})$. A straightforward calculation shows that

$$\Delta_T = ||P_I(0)\hat{\rho}_{S|0} - P_I(1)\hat{\rho}_{S|1}|| = \frac{1}{2(\bar{n}+1)} \sum_{k=0}^{\infty} \lambda^k |2(1-\eta_I)^k - 1|. \tag{B2}$$

We now consider $\bar{n}$ to be fixed and aim to maximize (B2) with respect to $\eta_I \in (0, 1]$. Note that if $\eta_I = 0$, then $P_I(1) = 0$ and thus we would only have one state prepared; for this reason, we exclude $\eta_I = 0$.

The only place η_I appears in (B2) is $|2(1-\eta_I)^k - 1|$. This vanishes when η_I is $x = 1 - (1/2)^{1/k}$. We find that, for $0 \leqslant \eta_I \leqslant x$, $|2(1-\eta_I)^k - 1|$ is strictly decreasing. For $x \leqslant \eta_I \leqslant 1$ we find that $|2(1-\eta_I)^k - 1|$ is strictly increasing. This means that the maximum of $|2(1-\eta_I)^k - 1|$ occurs at the endpoints of the interval $0 \leqslant \eta_I \leqslant 1$. We see that for η_I equal to zero or one, $|2(1-\eta_I)^k - 1| = 1$, however, $\eta_I = 0$ corresponds to $P_I(0)\hat{\rho}_{S|0} = 0$. To maximize the distinguishability, Δ_T, we want to make $|2(1-\eta_I)^k - 1|$ as large as possible, which means we set $\eta_I = 1$.

We have proven the desired result; we now examine the solution. A quick calculation shows that for $\eta_I = 1$, $\Delta_T = 1/2$

and the states $\hat{\rho}_{S|0}$ and $\hat{\rho}_{S|1}$ have disjoint support and are orthogonal, i.e., $\hat{\rho}_{S|0}\hat{\rho}_{S|1} = 0$. This confirms the mathematical result that the states are most distinguishable when $\eta_I = 1$.

APPENDIX C: CALCULATION OF EQUATION (9) AND $\mathcal{L}_1$ DISTANCE

In this Appendix we calculate equation (9) and the $\mathcal{L}_1$ distance between probability distribution $P_{IS}(i, s|O)$ and $P_{IS}(i, s|\bar{O})$. Both of these quantities help us gauge how difficult it is to distinguish between the situations where there is no object and when there is an object present.

We start by calculating (9) as a function of the parameters $\bar{n}$, $\bar{n}_B$, κ, η_S, and η_I. By expanding out (9) and evaluating the trace, we find

$$\begin{aligned}\Delta = &|[P_{IS}(0,0) - P_I(0)P_S(0|\bar{O})] \\ &+ [P_{IS}(1,0) - P_I(1)P_S(0|\bar{O})] \\ &= [P_I(0)P_S(1|\bar{O}) - P_{IS}(0,1)] \\ &+ [P_I(1)P_S(1|\bar{O}) - P_{IS}(1,1)]|, \end{aligned} \tag{C1}$$

where we have used the fact that $P_{IS}(i, s|O) = P_{S|I}(s|i, O)P_I(i)$. This expression can be further simplified using the fact that

$$P_{S|I}(0|i, O) - P_S(0|\bar{O}) = P_S(1|\bar{O}) - P_{S|I}(1|i). \tag{C2}$$

We thus find that

$$\begin{aligned}\Delta = &|2[P_{IS}(0,0) - P_I(0)P_S(0|\bar{O})] \\ &+ 2[P_{IS}(1,0) - P_I(1)P_S(0|\bar{O})]|. \end{aligned} \tag{C3}$$

Substituting in Eqs. (A7), (A9), and (A10) from Appendix A; one can show that

$$\Delta = \frac{2\kappa\bar{n}\eta_S}{(1+\eta_S\bar{n}_B)(1+\eta_S[\kappa\bar{n}+\bar{n}_B])}. \tag{C4}$$

Before studying this equation, we calculate the $\mathcal{L}_1$ distance.

To determine whether an object is present, we use the observation data, which is sampled from the probability distribution $P_{IS}(i, s|O)$ when an object is present or $P_{IS}(i, s|\bar{O})$ when there is no object. If these two joint probabilities are the same, then we cannot determine whether an object is present. The difference between them thus serves as a proxy for how easy it is to distinguish between these two alternatives. To quantify the difference between the probabilities, we use the $\mathcal{L}_1$ distance [42,43],

$$\mathcal{L}_1 = \sum_{i,s} |P(i, s|O) - P(i, s|\bar{O})|. \tag{C5}$$

This quantity can be simplified by recalling that $P_{IS}(i, s|O) = P_I(i)P_{S|I}(s|i, O)$ and $P_{IS}(i, s|\bar{O}) = P_I(i)P_S(s|\bar{O})$. We find that

$$\begin{aligned}\mathcal{L}_1 &= 2\sum_i P_I(i)|P_{S|I}(1|i, O) - P_S(1|\bar{O})| \\ &= 2\sum_i P_I(i)|P_{S|I}(0|i, O) - P_S(0|\bar{O})|, \end{aligned} \tag{C6}$$

where we have used the identity (C2).

The form given in Eq. (C6) is instructive. When $\bar{n}_B$ increases, both $P_{S|I}(0|i, O)$ and $P_S(0|\bar{O})$ will become small. Decreasing η_S from one causes both $P_{S|I}(0|i, O)$ and $P_S(0|\bar{O})$ to increase. However, when $\bar{n}_B$ is large, the probabilities increase in slightly different ways.

Note that Eq. (C6) equals (C3) when

$$\begin{aligned}P_{S|I}(0|0, O) &\leqslant P_S(0|\bar{O}), \\ P_{S|I}(0|1, O) &\leqslant P_S(0|\bar{O}). \end{aligned} \tag{C7}$$

Physically, these conditions say that adding photons to the signal mode makes us more likely to detect something. The first inequality follows immediately by comparing $P_S(0|\bar{O}) = 1/(1+\eta_S\bar{n}_B)$ to (A7). Using (A9) we find that the second inequality is equivalent to

$$Y = \eta_I\bar{n}BC - (1+\eta_I\bar{n})AC + AB \geqslant 0, \tag{C8}$$

where

$$\begin{aligned}A &= 1 + \eta_S\bar{n}_B, \\ B &= 1 = \eta_S[\kappa\bar{n} + \bar{n}_B], \\ C &= 1 + \eta_S[\kappa\bar{m}_0 + \bar{n}_B]. \end{aligned} \tag{C9}$$

We can prove this inequality by noting that $B \geqslant C$ and thus Y is lower bounded by

$$Y \geqslant \eta_I\bar{n}BC - (1+\eta_I\bar{n})AC + AC = C\eta_I\bar{n}^2\eta_S\kappa \geqslant 0. \tag{C10}$$

We have thus proven that $\mathcal{L}_1 = \Delta$, and thus the trace distance is given by Eq. (C4).

An interesting feature of Eq. (C4) is that it does not depend on η_I. However, from the numerical evaluation of (5) together with the argument around Eq. (6) and in Appendix B, we know that $\eta_I = 1$ is the optimal choice for the idler detector. Suppose we keep $\bar{n}$, κ, and $\bar{n}_B$ constant and treat Δ as a function of η_S. Using elementary calculus, we find that Δ has a maximum for $\eta_S < 1$ when $\bar{n}_B(\kappa\bar{n} + \bar{n}_B) > 1$. The maximum occurs at $\eta_S = \sqrt{1/(\bar{n}_B[\kappa\bar{n} + \bar{n}_B])}$. This confirms that when the total number of photons is high, we should use a value for η_S which is less than one.

APPENDIX D: EFFECTS OF CHANGING PULSE SEPARATION AND PULSE WIDTH ON SYSTEM PARAMETERS

In Sec. IV we looked at the effect of using shorter and more frequent pulses with lower mean photon numbers, but with the rate of photons per second fixed. In this section we elaborate on this and demonstrate that, while the mean number of background photons $\bar{n}_B$ can decrease, the rate of background photons is not affected.

Let $\bar{n}$ be the mean number of signal photons per pulse and let ΔT be the separation between pulses. We separate each pulse into its own time bin of width ΔT. The rate of signal photons is $R_S = \bar{n}/\Delta T$. Similarly, the rate of background photons is R_B. If we gate the detector on only for the duration of the pulse, then the mean number of background photons in each time bin, $\bar{n}_B$, is $R_B\tau$. Alternatively, if we do not gate the detector on, then the background mean photon umber would be $R_B\Delta T$. Suppose we decrease the width of the pulses such that the mean photon number per pulse is now $\bar{n}/2$. To keep R_S fixed, we must decrease the separation between pulses to $\Delta T/2$. The situation is illustrated in Fig. 9.

The effect of decreasing the pulse separation from ΔT to $\Delta T/2$ is to double the pulse repetition rate. To compensate

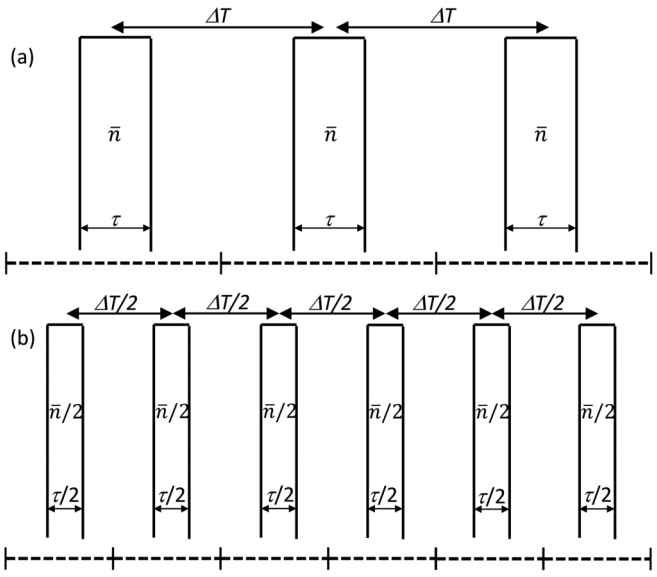

FIG. 9. Illustrating the setup for decreasing the width of the pulses such that the mean photon number per pulses is halved, but where the photon flux is kept fixed by increasing the rate of pulses. Panel (a) is the original case, where pulses have width τ, separation ΔT and mean photon number per pulses of $\bar{n}$. Panel (b) shows pulses with half the width, $\tau/2$ and half the separation, $\Delta T/2$. The mean photon per pulse is now $\bar{n}/2$, but the rate of photons per second is the same in panels (a) and (b). Notice that the size of the time bins (the dotted lines) in panel (b) are half the size of those in panel (a).

we halve the size of the time bins we use for each pulse to keep the rate of signal photons constant. We first analyze the situation without gating the signal detector on and off. In this case, halving the size of the time bin will result in halving of the mean number of background photons per time bin. If we originally had on average $\bar{n}_B$ background photons per time bin, then this will be reduced to $\bar{n}_B/2$. Note that the *rate* of background photons per second, R_B, is unchanged. Altering the pulse widths, their separation or the time-bin widths cannot affect the rate at which we receive background photons. However, changing the size of the time bin does affect the mean number of photons detected within each one.

In practice, one often gates the signal detector on and off to reduce the background. The time-duration for which the detector is gated on will depend on the pulse width τ, although there may be a minimum possible gating time set by detector jitter limits. If we had decreased the mean photon number per pulse while keeping the pulse width fixed, then the mean number of background photons would have been unchanged. However, we decrease the pulse width to $\tau/2$, so that the time the detector is gated on can also be decreased. It is for this reason that we can take the mean number of background photons per time bin as reduced to $\bar{n}_B/2$. Again note, the rate of background photons per second, R_B, has not changed. Instead, the background mean photon number has reduced due to the decrease in time for which the detector is active in each time bin.

In general, if we reduce the pulse width to τ/m and the mean photon number per pulse to $\bar{n}/m$, then to keep the rate of signal photons per second fixed, we must decrease the pulse separation to $\Delta T/m$, which reduces the width of the time bins to $\Delta T/m$. The time the signal detector is gated on can also be reduced by a factor of $1/m$ due to the decrease in pulse duration. The effect of this is to reduce the mean number of background photons per time bin to $\bar{n}_B/m$.

APPENDIX E: INFORMATION GAIN PER PHOTON

In this Appendix we use an information theoretic approach to examine our gain in information when we obtain measurement outcomes. The approach we take is the same as in Sec. IV. We transmit $\bar{n}$ signal photons spread over m time bins, where the width of each time bin is $\Delta T/m$, the width of each pulse is τ/m and the number of background photons per time bin is $\bar{n}_B/m$. For all positive integers m, we transmit the same number of photons in the time interval ΔT.

We are interested in quantifying how much information we gain after we have detected m pulses. Our knowledge of whether the object is present is encoded within the posterior probability $P(O|\Omega_m)$, where Ω_m is the set of m idler and signal detection events. Before the measurements, our knowledge was given by the prior probability. The increase in information after measurements is related to the change in the probability due to Bayesian updating.

A common way of quantifying the difference between two probabilities is to use the *relative entropy*, which is also known as the Kullback-Leibler divergence [42,43]. For a given set of measurement outcomes, Ω_m, the relative entropy is

$$S(\Omega_m) = \sum_{k\in\{O,\bar{O}\}} P(k|\Omega_m)\log_2\left(\frac{P(k|\Omega_m)}{P(k)}\right). \tag{E1}$$

The relative entropy has many different applications in physics and statistics [42,43,46]. For example, the relative entropy between a joint probability and its two marginal distributions is known as the mutual information, which plays a pivotal role in information theory [43]. A more pertinent application comes in the field of inference, where Eq. (E1) has been used to quantify the increase in our knowledge from Bayesian updating [47,48]. We take Eq. (E1) to be the increase in our knowledge given a specific set of measurement outcomes, Ω_m The average increase in our knowledge is found by averaging (E1) over all possible sets of measurement outcomes

$$\bar{S}_m = \sum_{\Omega_m} P(\Omega_m|O)S(\Omega_m), \tag{E2}$$

where the summation is over all possible sets of measurement outcomes for m time bins.

In Sec. IV we examined the averaged posterior probability for the object to be present given we had transmitted a given amount of signal photons. To determine the average gain per photon, we consider $\bar{S}_m$ divided by the mean number of photons incident on the signal detector during the time period ΔT. The quantity we consider is

$$I_m(O) = \frac{\bar{S}_m}{(\bar{n}+\bar{n}_B)}. \tag{E3}$$

The reasons for considering this quantity rather than $\bar{S}_m/\bar{n}$ is as follows. First, using $\bar{n}+\bar{n}_B$ in the denominator is more

pessimistic than using just $\bar{n}$. This becomes important when $\bar{n}$ is small and less than $\bar{n}_B$. In this regime there would be an incentive to decreasing $\bar{n}$ if we used $S_m/\bar{n}$. Related to this is the fact that in the limit of $\bar{n} \to 0$, the quantity $S_m/\bar{n}$ is not defined. We can verify this directly for $m = 1$. We first find that in the limit of $\bar{n} \to 0$

$$P(O|i=0,s=0) = P(O|i=0,s=1) = \frac{1}{2},$$
$$P(O|i=1,s=0) = \frac{1-\eta\kappa+\eta\bar{n}_B}{2-\eta\kappa+2\eta\bar{n}_B},$$
$$P(O|i=1,s=1) = \frac{\kappa+\bar{n}_B+\eta\bar{n}_B^2}{\kappa+2\bar{n}_B(1+\eta\bar{n}_B)}, \tag{E4}$$

where for simplicity, we have taken $\eta_I = \eta_S = \eta$. From this, we see that the limit of $\bar{S}_1$, as $\bar{n} \to 0$, is defined and nonzero. But this means that the limit of $\bar{S}_1/n$ will not be defined. An immediate consequence of this is that if we used $\bar{S}_1/n$, then we would find that making $\bar{n}$ arbitrarily small would yield an arbitrarily large value for S_1/n. These problems are all avoided by using (E3).

In addition to the average information gain per photon, we can also explore the information gain per photon for a specific set of measurement outcomes, which is given by

$$I_m(O|\Omega_m) = \frac{S(\Omega_m)}{(\bar{n}+\bar{n}_B)}. \tag{E5}$$

To understand the information measures we first look at the example of $m = 1$, where there is only a single outcome each for idler and signal detectors. Consider the case where the idler detector fires. If the signal detector does not fire, then we become less confident there is an object present. The posterior probability for the object to be present will thus decrease. In contrast, if the signal detector fires, the posterior probability will increase. Both of these outcomes provide information, which we quantify by $I_1(O|i=1,s=0)$ and $I_1(O|i=1,s=1)$. We find that the outcome that provides the most information depends on both the number of background photons, $\bar{n}_B$ and the number of signal photons $\bar{n}$. For high numbers of photons, we gain more information when the signal detector does not fire, while the opposite is true when the total number of photons is lower. This is illustrated in Fig. 10, where we plot the minimum number of background photons, $\bar{n}_B^{\min}$, such that $I_1(O|i=1,s=0) > I_1(O|i=1,s=1)$, against $\bar{n}$. Both curves are for $\eta_I = \eta_S = 0.9$, while the solid black curve is for $\kappa = 0.1$, while the dashed black curve is for $\kappa = 0.05$. We see that the value of $\bar{n}_B^{(\min)}$, for which $I_1(O|i=1,s=0) > I_1(O|i=1,s=1)$, becomes lower as $\bar{n}$ increases.

Intuitively, one can understand this behavior by recalling that when we condition on the idler detector firing, the number of photons in the signal mode is larger than $\bar{n}$. If an object is present, then we expect to detect light reflected from it, especially as the mean number of signal photons increases. If we do not detect photons, then it would be surprising if an object is present. This would instead imply that an object is not present. It is for this reason that $I_1(O|i=1,s=0)$ is greater than $I_1(O|i=1,s=1)$. In contrast, when we have few background photons, then we do not expect the signal detector to fire, unless the signal photons have been reflected from an object. In this case, we find that $I_1(O|i=1,s=1) > I_1(O|i=1,s=0)$.

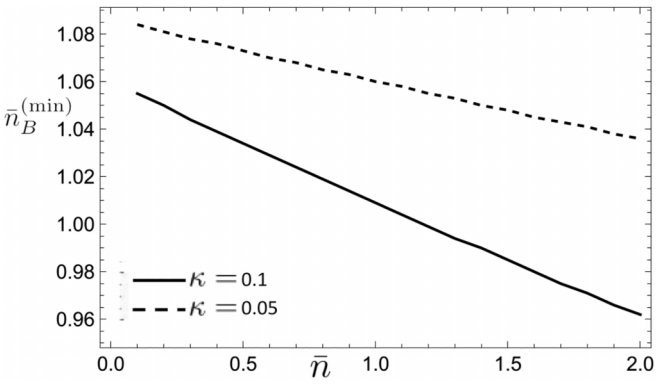

FIG. 10. A plot showing how $\bar{n}_B^{(\min)}$, the minimum value for $\bar{n}_B$ for which $I_1(O|i=1s=0) > I_1(O|i=1,s=1)$, varies with $\bar{n}$ for different values of κ. All plots are for $\eta_I = \eta_S = 0.9$. The solid black curve is for $\kappa = 0.1$, and the dashed black curve is for $\kappa = 0.05$.

We can study the dependence of the averaged information gain, $I_m(O)$, on m, for fixed values of $\bar{n}$. In Fig. 11 we plot $I_m(O)$ against m for different values of $\bar{n}$ and $\bar{n}_B$. In both Figs. 11(a) and 11(b), $\kappa = 0.1$ and $\eta_I = \eta_S = 0.9$. In Fig. 11(a) we have $\bar{n} = 0.01/m$ and $\bar{n}_B = 0.1/m$, while for Fig. 11(b) we have $\bar{n} = 1/m$ and $\bar{n}_B = 3/m$. In both figures,

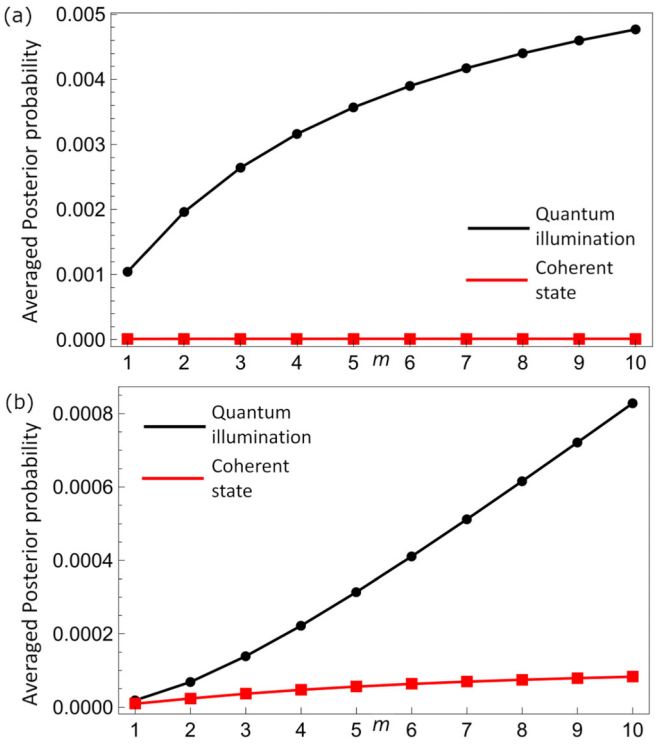

FIG. 11. A plot of the averaged information gain per photon, $I_m(O)$, when an object is present, against m, the number of copies of the state. All plots are for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$ and for the case when an object is present. Panel (a) is for $\bar{n} = 0.01$ and $\bar{n}_B = 0.1$, while panel (b) is for $\bar{n} = 1.0$ and $\bar{n}_B = 3.0$. In both plots, the black curves are for a direct detection illumination, while the red (gray offline) are for transmitting coherent states.

the black curve is for quantum illumination, while the red (gray offline) is for transmitting coherent states. The information gain per photon is seen to increase as m increases. This is consistent with the results shown in Fig. 6. The ratio background-to-signal photons, $\bar{n}_B/\bar{n}$, is ten in Fig. 11(a), but only three in Fig. 11(b). Yet despite the greater ratio of background-to-signal photons, the information gain per photon is greater in Fig. 11(a) than in Fig. 11(b). However, while the ratio $\bar{n}_B/\bar{n}$ is greater for Fig. 11(a) than Fig. 11(b), the number of background photons, $\bar{n}_B$, is greater in Fig. 11(b). It is thus the number of background photons, rather than the ratio of background to signal photons, which is important for the information gain per photon.

The functional dependence of $I_m(O)$ on $\bar{n}$ can be investigated. One interesting feature we find is that $I_m(O)$ can have a maximum for some value of $\bar{n}$. The position of this maximum will depend on η_I, η_S, $\bar{n}_B$, and κ. We often find that the value of the maximum occurs for large values of $\bar{n}$. For example, for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, and $\bar{n}_B = 3$, then the maximum of $I_1(O)$ occurs at $\bar{n} = 37.3$. If we keep κ, η_I, and η_S the same, but now $\bar{n}_B = 1.0$, then the maximum of $I_1(O)$ occurs at $\bar{n} = 18.1$. For large values of $\bar{n}$, we find that the state conditioned on the idler firing, $\hat{\rho}_{S|1}$, is approximately a thermal state [15,26]. Furthermore, the relative enhancement in mean photon number for the conditional state is negligible when $\bar{n}$ is large. This means that, for large values of $\bar{n}$, direct measurement based illumination does not perform as well as using a coherent states [15,26]. It is thus not advisable to run the setup at or near the value of $\bar{n}$ that optimizes the information gain per photon.

APPENDIX F: THE EFFECTS OF POST-SELECTING ON THE IDLER DETECTION FIRING

In this section we address two questions. First, does postselection improve the performance of direct detection illumination? Second, if we neglect the times we did not obtain the desired outcome, does postselection then improve the performance? Unsurprisingly, we find postselect does not improve direct detection illumination, unless we neglect the undesired outcome. This is surprising because normally postselection is associated with situations where we have a failure outcome. But in direct detection illumination, all outcomes yield information. It thus seems a paradox that discarding information can improve the performance. We resolve this issue in this Appendix.

We consider postselecting on the idler detector firing. We evaluate the average performance for $N = 1$, by modifying Eq. (5) to

$$\mathcal{P}(O|i=1) = \sum_s \frac{P_{IS}(1,s|O)P(O|1,s)}{\sum_s P_{IS}(1,s|O)}, \tag{F1}$$

where $\mathcal{P}(O|i=1)$ denotes the averaged posterior probability given we postselect on the idler detector firing. For a given TMSV state (1) with mean photon number of $\bar{n}$ in each mode and using the same values for κ, η_I, η_S, and $\bar{n}_B$, we find that $\mathcal{P}(O|i=1)$ is generally greater than $\mathcal{P}(O)$, i.e., postselection improves the performance. This is illustrated in Fig. 8(a). This compares the results of (5) against (F1) for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, $\bar{n}_B = 3.0$, and for an object present. However, Eq. (F1) neglects the fact that to obtain one postselected event requires of order $1/P_I(1)$ pulses. As such, postselection will not improve the time taken to be confident that an object is present or not. To take account of this, we multiply (F1) by the probability for the idler detector to fire. The new averaged posterior probability is

$$\mathcal{P}(O, i=1) = \sum_s P_{IS}(1,s|O)P(O|1,s). \tag{F2}$$

By comparing (5) with the above equation, we see that (5) is equal to $\mathcal{P}(O, i=1)$ plus additional non-negative terms, and thus $\mathcal{P}(O, i=1) \leqslant \mathcal{P}(O)$. Once we account for the postselection probability, the performance of postselection is found to be worse than not postselecting.

Postselection is thus not as effective as not postselecting. However, when we neglect the extra time taken to generate postselected results, we found that postselection can give an advantage. For example, Fig. 8 showed that (F1) can be greater than (5), but from a Bayesian perspective, the occasions when the idler detector does not fire should still provide some information. How then can discarding such events improve the performance?

When we do not postselect, then on average we transmit $\bar{n}$ photons per time bin. When the idler detector fires, the signal mode has a mean of $\bar{n} + (1+\bar{n})/(1+\eta_I\bar{n})$ photons in the time bin [26]. This is balanced by a reduction in the mean photon number when the idler detector does not fire. However, when we postselect, we neglect the situations where the mean photon number is reduced. In relative terms, postselection increases the mean number of photons transmitted in each pulse. To account for this we should compare postselection on a TMSV state with $\bar{n}$ to the previous protocol with a TMSV state with $\bar{n} + (1+\bar{n})/(1+\eta_I\bar{n})$. This is illustrated in Fig. 8(b), which is plotted for $\kappa = 0.1$, $\eta_I = \eta_S = 0.9$, $\bar{n}_B = 3.0$ and for an object present. We find that even when we neglect the postselection probability, the protocol described in Sec. II is better than postselection once we equalize the mean number of photons transmitted. This occurs because not postselecting retains information both for when the idler fires and when it does not. If there is an object present, then these two situations will lead to different detection probabilities for the signal detector. In contrast, when there is no object, then the background is fixed as is the signal detector's probability. The information discarded in postselection is thus useful in a Bayesian analysis.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[2] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 2004).

[3] S. Pirandola, B. R. Bardhan, T. Gehring *et al.*, Nat. Photon **12**, 724 (2018).

[4] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, Phys. Rev. Lett. **85**, 2733 (2000).

[5] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, Phys. Rev. A **52**, R3429 (1995).

[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[7] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).

[8] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[9] S. Lloyd, Science **321**, 1463 (2008).

[10] S. H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Phys. Rev. Lett. **101**, 253601 (2008).

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

[12] R. G. Torromé and S. Barzanjeh, Prog. Quantum Electron. **93**, 100497 (2024).

[13] S. Frick, A. McMillan, and J. Rarity, Opt. Express **28**, 37118 (2020).

[14] M. P. Mrozowski, R. J. Murchie, J. Jeffers, and J. D. Pritchard, Opt. Express **32**, 2916 (2024).

[15] T. Brougham, N. Samantaray, and J. Jeffers, Phys. Rev. A **108**, 052404 (2023).

[16] P. R. Tapster, S. F. Seward, and J. G. Rarity, Phys. Rev. A **44**, 3266 (1991).

[17] T. S. Iskhakov, V. C. Usenko, U. L. Andersen, R. Filip, M. V. Chekhova, and G. Leuchs, Opt. Lett. **41**, 2149 (2016).

[18] M. Mehmet, S. Ast, T. Eberle, S. Steinlechner, H. Vahlbruch, and R. Schnabel, Opt. Express **19**, 25763 (2011).

[19] S. Guha and B. I. Erkmen, Phys. Rev. A **80**, 052310 (2009).

[20] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Phys. Rev. Lett. **118**, 040801 (2017).

[21] E. D. Lopaeva, I. Ruo Berchera, I. P. Degiovanni, S. Olivares, G. Brida, and M. Genovese, Phys. Rev. Lett. **110**, 153603 (2013).

[22] D. G. England, B. Balaji, and B. J. Sussman, Phys. Rev. A **99**, 023828 (2019).

[23] R. J. Murchie, J. D. Pritchard, and J. Jeffers, Proc. SPIE **11835**, 118350G (2021).

[24] H. Yang, W. Roga, J. D. Pritchard, and J. Jeffers, Opt. Express **29**, 8199 (2021).

[25] H. Yang, W. Roga, J. D. Pritchard and J. Jeffers, Quantum Technol. **11347**, 113470I (2020).

[26] H. Yang, N. Samantaray, and J. Jeffers, Phys. Rev. Appl. **18**, 034021 (2022).

[27] W. Roga and J. Jeffers, Phys. Rev. A **94**, 032301 (2016).

[28] M. Reichert, Q. Zhuang, J. H. Shapiro and R. Di Candia, Phys. Rev. Appl. **20**, 014030 (2023).

[29] H. Shi, B. Zhang, J. H. Shapiro, Z. Zhang, and Q. Zhuang, Phys. Rev. Appl. **21**, 034004 (2024).

[30] S. M. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics* (Oxford University Press, Oxford, 1997).

[31] S. M. Barnett and P. L. Knight, J. Opt. Soc. Am. B **2**, 467 (1985).

[32] B. Yurke and M. Potasek, Phys. Rev. A **36**, 3464 (1987).

[33] C. G. Ghirardi, A. Rimni, and T. Weber, Lett. Nuovo Cimento **27**, 293 (1980).

[34] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).

[35] J. Peřina, O. Haderka, and V. Michálek, Opt. Express **21**, 19387 (2013).

[36] J. Sperling, T. J. Bartley, G. Donati, M. Barbieri, X.-M. Jin, A. Datta, W. Vogel, and I. A. Walmsley, Phys. Rev. Lett. **117**, 083601 (2016).

[37] D. N. Klyshko, Phys. Lett. A **213**, 7 (1996).

[38] S. M. Barnett, L. S. Phillips, and D. T. Pegg, Opt. Commun. **158**, 45 (1998).

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

[40] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).

[41] R. J. Murchie, J. D. Pritchard, and J. Jeffers, arXiv:2307.10785.

[42] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[43] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley and Sons, 1991).

[44] Wolfram Research, Inc., *Mathematica*, Version 13.0 (Champaign, 2023).

[45] T. Brougham, C. F. Wildfeuer, S. M. Barnett, and D. J. Gauthier, Eur. Phys. J. D **70**, 214 (2016).

[46] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

[47] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).

[48] P. A. Ortega and D. A. Braun, J. Artif. Int. Res. **38**, 475 (2010).