The Lindley paradox in optical interferometry

Camillo Mauri [a], Matteo G.A. Paris [a,b,c,*]

[a] *Quantum Technology Lab, Dipartimento di Fisica, Università degli Studi di Milano, I-20133 Milano, Italy*
[b] *CNISM, Unità Milano Statale, I-20133 Milano, Italy*
[c] *INFN, Sezione di Milano, I-20133 Milano, Italy*

A B S T R A C T

The so-called *Lindley paradox* is a counterintuitive statistical effect where the Bayesian and frequentist approaches to hypothesis testing give radically different answers, depending on the choice of the prior distribution. In this paper we address the occurrence of the Lindley paradox in optical interferometry and discuss its implications for high-precision measurements. In particular, we focus on phase estimation by Mach–Zehnder interferometers and show how to mitigate the conflict between the two approaches by using suitable priors.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Interferometric setups are at the heart of several high-precision measurement schemes, ranging from gravitational wave-detectors to laser gyroscopes and clocks synchronization protocols. In particular, a direct detection of gravitational waves is supposed to come from laser-interferometric detectors [1–4], where they induce a measurable variation of the optical paths of the light beams traveling along the arms of an interferometric setup.

Interferometers are usually employed to detect small perturbations, so that detectors must be able to measure phase shifts with remarkable precision and any source of noise must be carefully removed. Indeed, recent technological advances in precision lasers, vacuum technology and optical systems had made it possible to greatly reduce classical noise. However, an unavoidable constraint to the precision with which optical signals can be measured arises from the quantum nature of the electromagnetic field. The so-called shot noise, results as a consequence of the uncertainty relations existing for the quantum field operators.

A way to partially circumvent the effects of quantum fluctuations and increase interferometric precision is to exploit the use of squeezed light [5–7] or other nonclassical states of light [8,9], also in the presence of inefficient detectors [10]. The implementation of squeezing-enhancement techniques in interferometric detectors

has been a great challenge: after a large number of experimental and theoretical studies it has now became possible to take advantage of squeezing [11,12] so that the next generation of interferometers will be equipped with quantum technologies.

In order to effectively exploit the potential improvements offered by quantum signals, the statistical analysis of the data should be refined as well. In this context, the key point is to improve the ability of discriminating whether a signal contains instrumental noise only or also a trace of a signal. This problem, which arises in the form of a hypothesis testing problem [13], might be conveniently studied by means of Bayesian inference [8,14–20]. However, a Bayesian approach may be challenging since it requires the full knowledge of the system under consideration in order to properly introduce a suitable prior, together with a high computational power. Indeed, in a Bayesian framework we have to provide explicit models of the signal induced by a perturbation to achieve, after specifying our prior knowledge, the odds ratio of the detection hypothesis over the null hypothesis, corresponding to absence of signal. In this context, it may happen that the lack of knowledge about the system increases the false alarm probability. In other words, the posterior odds may favor the alternative hypothesis, i.e. presence of perturbations, even if no signals are actually present in data. In particular, when a diffuse prior distribution is taken, a Bayesian approach to hypothesis testing leads to conflicting evidences compared to a frequentist approach. Such a counterintuitive situation was studied by Jeffrey [21] together with Lindley, who first referred to this disagreement as a statistical paradox [22].

---

* Corresponding author.
  E-mail address: matteo.paris@fisica.unimi.it (M.G.A. Paris).

The relevance of Bayesian analysis in interferometry has been recognized since a long time [8,23]. The interest raised for at least three reasons. On the one hand, common interferometric estimators are known to achieve optimal performances only at a specific working point [5,24], thus resulting in interferometric protocols that do not allow the measurement of arbitrary phase shifts. Besides, in order to estimate small fluctuations around the working point, the interferometer has to be actively stabilized by the addition of some feedback mechanism [25]. Finally, Bayesian estimators have been shown to achieve the asymptotic regime, where they saturate the Cramer–Rao bound [26], already with few measurements [12,16–18], thus representing a convenient choice in any setting where resources are limited or the samples involved in the interferometric setup are fragile.

In this work we address in details the occurrence of the *Lindley paradox* [27,28] in the analysis of data coming from optical interferometry. The importance of the topic seats in the fact that often, e.g. in the framework of gravitational antennas, each event is of great relevance and should be analyzed in the most refined way. More generally, it is often the case that besides the estimation of small fluctuations, interferometric measurements are involved in monitoring a given physical configuration, where the information coming from the experimental data is exploited to statistically discriminate between two hypothesis [29–33]. Here, we first consider homodyne detection, then classical-like Mach–Zehnder (MZ) interferometer, and finally we consider the case of squeezed-enhanced quantum interferometers. For the MZ interferometer we start from the analysis of the ideal case where the detectors are perfectly efficient, and a diffuse prior is assumed, and then proceed by assessing the realistic case where inefficient detectors and a concentrated prior distribution is employed. As we will see, the Lindley paradox may indeed arise in optical interferometry. Here we show how to mitigate the conflict between the Bayesian and frequentist approaches by using suitable priors.

The paper is organized as follows. In Sec. 2 we introduce the notation and illustrate how and when the paradox can arise. In Sec. 3 we discuss phase estimation by homodyne detection as a suitable example to illustrate the statistical analysis step by step. In Sec. 4 we analyze in details the occurrence of the paradox in optical interferometry, whereas Section 5 closes the paper with some concluding remarks.

2. The Lindley paradox

Let us consider a random variable X distributed according to a Gaussian distribution of unknown mean ϕ and known variance σ^2. Suppose that we want to test a sharp null hypothesis H_0, corresponding to the prediction $\phi = \phi_0$, against the alternative hypothesis H_1, corresponding to the diffuse prediction $\phi \neq \phi_0$. In a Bayesian framework this is done assigning a priori a probability $p(H_i)$ and a prior distribution $\pi_i(\phi)$ to any hypothesis H_i taken into account. Let us denote the priors by $p(H_0) = z_0$ and $p(H_1) = 1 - z_0$, with $z_0 \in [0, 1]$. Besides, we take $\pi_0(\phi)$ to be a Dirac's delta function since the null hypothesis involves a single parameter value. As for the probability density $\pi_1(\phi)$, concerning the alternative hypothesis, we assume a normal distribution with variance τ^2. That being said, the posterior probability $\bar{z}_0$ that the outcome $X = x$ provides evidence of $\phi = \phi_0$, i.e. confirms the null hypothesis, is evaluated via Bayes' theorem, which yields (see Appendix A):

$$\bar{z}_0 = p(\phi = \phi_0 | x) = \left\{ 1 + \frac{1 - z_0}{z_0} \frac{p(x | \phi \neq \phi_0)}{p(x | \phi = \phi_0)} \right\}^{-1}. \tag{1}$$

In order to illustrate the occurrence of the Lindley paradox, let us consider the simple case where $\phi_0 = 0$ and $\sigma = 1$, where we have (see Appendix A for the general case)

$$\bar{z}_0 = p(\phi = \phi_0 | x)$$

$$= \left\{ 1 + \frac{1 - z_0}{z_0} \frac{1}{\sqrt{1 + \tau^2}} \exp\left[\frac{x^2 \tau^2}{2(1 + \tau^2)} \right] \right\}^{-1}, \tag{2}$$

which goes to 1 as τ goes to infinity, no matter the value of the outcome x and of the prior probability z_0. In other words, when the prior distribution for the alternative hypothesis becomes non-informative (i.e. approaches a flat distribution), a Bayesian approach awards high odds to the null hypothesis even if the observed value is several standard deviation away from ϕ_0. This is clearly in contrast with the predictions that any frequentist approach based on sampling theory may provide, not speaking of common sense.

The nature of the paradox, actually the lack of any paradox, has been extensively analyzed [27,28] and we are not going to discuss the situation from a statistical point of view, which would be beyond the scope of our analysis. We limit ourselves to notice that the disagreement between the two approaches is basically caused by the fact that the frequentist approach tests one hypothesis without reference to the other, whereas Bayesian analysis assess them as alternative one to each other, looking for the one in better agreement with the observations. We also point out that the paradox is of great generality since the key point for its appearance is the use of a diffuse prior distribution. In the following sections, we will exhibit three situation of physical interest where the data analysis naturally involves the assignment of a prior, such that a discussion about the occurrence of the paradox is worthwhile.

3. Phase estimation by homodyne detection

Homodyne detection is a measurement scheme where the observed mode of the field interferes with a reference mode, thus providing indirect information about its phase. The device is illustrated schematically in Fig. 1. The field to be measured, prepared in an arbitrary state ρ of a single-mode a, is mixed with the field produced by an intense laser, i.e. with a highly excited coherent state $|z\rangle$, with $z = |z| e^{i\phi} \in \mathbb{C}$. Two detectors measure the photocurrents i_c and i_d, while a differential amplifier determines the *homodyne photocurrent* H, i.e. the rescaled difference photocurrent $\hat{H} = (\hat{i}_c - \hat{i}_d)/|z|$. In the limit of large $|z|$ it turns out that homodyne detection provides the measurement of the quadrature operator $\hat{\chi}_\phi$ given by:

$$\hat{\chi}_\phi = \frac{\hat{a}^\dagger e^{i\phi} + \hat{a} e^{-i\phi}}{\sqrt{2}}. \tag{3}$$

We point out that the device works properly only if the incoming fields occupy the same spatial mode, have the same frequency ω and are co-polarized. Besides, note that our description refers to ideal photodetectors with quantum efficiency equal to 1. This being said, we address homodyne detection as a tool to establish whether the phase of a given initial field has been perturbed or not, e.g. due to interaction with the environment. In particular, we consider a squeezed state $|\alpha, \lambda\rangle$, with $\alpha = |\alpha| e^{i\theta}$, $\lambda = r e^{i\varphi} \in \mathbb{C}$,
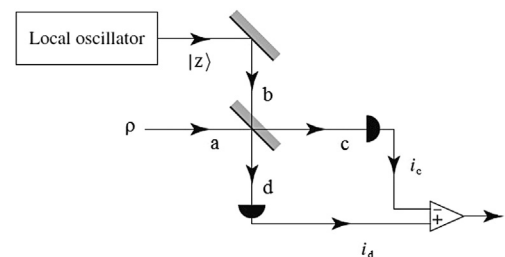

Fig. 1. Schematic diagram of homodyne detection scheme.

$r > 0$. That is, we take the impinging beam to be in a pure state given by:

$$|\psi_0\rangle = |\alpha, \lambda\rangle = \hat{D}(\alpha)\hat{S}(\lambda)|0\rangle, \tag{4}$$

where $\hat{S}(\lambda) = \exp(-\frac{\lambda}{2}\hat{a}^{\dagger 2} + \frac{\lambda^*}{2}\hat{a}^2)$ and $\hat{D}(\alpha) = \exp(-\alpha\hat{a}^\dagger + \alpha^*\hat{a})$ are unitary operators, namely the *squeezing operator* and the *displacement operator*, whose action in the phase space defined by the observables $\hat{\chi}_\phi$ and $\hat{\chi}_{\phi+\frac{\pi}{2}}$ is, respectively, a hyperbolic rotation and a translation. Besides, we assume that the interaction with the environment results in a phase shift expressed by the evolution operator $\hat{U}_I = \exp(-i\hat{a}^\dagger\hat{a}\beta)$, with $\beta \in [0, 2\pi]$. The final state $|\psi\rangle$, impinging on the semi-reflecting mirror, is readily obtained using the BCH formulas, which yield:

$$|\psi\rangle = \hat{U}_I|\psi_0\rangle = \exp(-i\hat{a}^\dagger\hat{a}\beta)|\alpha, \lambda\rangle = \left|\alpha e^{-i\beta}, \lambda e^{-2i\beta}\right\rangle. \tag{5}$$

The distribution $P_{\alpha,\lambda}(\chi_\phi)$ of the outcomes of repeated measurements of $\hat{\chi}_\phi$ performed on radiation prepared in state $|\psi\rangle$ is a Gaussian with expectation value $\bar{\chi}_\phi$ and variance σ_χ^2 given by:

$$\bar{\chi}_\phi = \langle\psi|\hat{\chi}_\phi|\psi\rangle = \frac{1}{\sqrt{2}}\left(\alpha e^{-i(\phi+\beta)} + \alpha^* e^{i(\phi+\beta)}\right), \tag{6}$$

$$\sigma_\chi^2 = \langle\psi|\hat{\chi}_\phi^2|\psi\rangle - \langle\psi|\hat{\chi}_\phi|\psi\rangle^2 =$$
$$= \frac{1}{2}\left[e^{2r}\sin^2\left(\phi + \beta - \frac{\varphi}{2}\right) + e^{-2r}\cos^2\left(\phi + \beta - \frac{\varphi}{2}\right)\right]. \tag{7}$$

The above equations lead us to conclude that we shall check whether the outcomes of measurements by means of homodyne detection are consistent with the null hypothesis $\beta = \beta_0 = 0$ in order to establish whether the initial state $|\psi_0\rangle$ has been perturbed or not. Without loss of generality, we focus on a specific example and set $\theta = 0$, $\varphi = 0$, corresponding to an input amplitude-squeezed state. We may also fix the quadrature component to be measured: for the sake of simplicity we take $\phi = 0$. In these conditions, Eqs. (6), (7) become

$$\bar{\chi}_0(\beta) = |\alpha|\sqrt{2}\cos\beta = \bar{\chi}_0(\beta), \tag{8}$$

$$\sigma_\chi^2(\beta) = \frac{1}{2}[e^{2r}\sin^2\beta + e^{-2r}\cos^2\beta]. \tag{9}$$

Let us now suppose that a run returns an outcome $q = \bar{\chi}_0 + t \cdot \sigma_\chi$, with $t \in \mathbb{R}$. In agreement with a standard sampling-theory test based on the value of t we shall conclude that if q is several σ_χ from $\bar{\chi}_0$ the null hypothesis is to be rejected: we have to say that a phase shift occurred and the state has been indeed perturbed. In order to provide a Bayesian assessment for H_0 we set a value for the prior probability, say $z_0 = 0.9$, and distribute uniformly the remainder probability $1 - z_0$ over the interval $I = [-\pi, \pi]$, where the phase may take values. On one hand, we are allowed to set such a high value for z_0 under the assumption that we are leading an experiment in optimal conditions, any source of noise being taken into account. On the other hand, the choice of a flat prior should not be surprising since we have no idea about the dynamics of a possible interaction: we are equally (un)expected to measure a small or a large value for the phase shift. Eventually, the posterior probability $\bar{z}_0$ that $\beta = \beta_0 = 0$ is evaluated by the Bayes theorem:

$$\bar{z}_0 = \left\{1 + \frac{1-z_0}{z_0}\frac{p(q|\beta \neq \beta_0)}{p(q|\beta = \beta_0)}\right\}^{-1}, \tag{10}$$

where

$$p(q|\beta = \beta_0) = \frac{1}{\sqrt{2\pi\sigma_\chi^2(0)}}\exp\left[-\frac{q^2}{2\sigma_\chi^2(0)}\right], \tag{11}$$
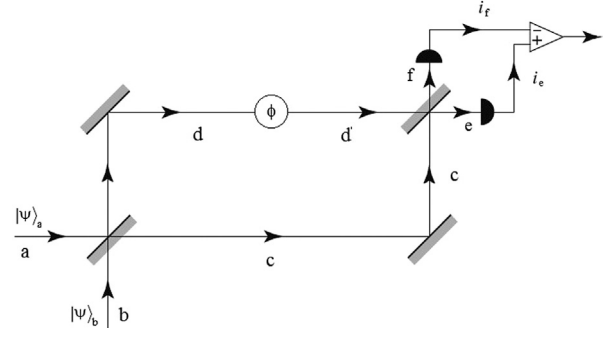

Fig. 2. Schematic diagram of a Mach–Zehnder interferometer.

$$p(q|\beta \neq \beta_0) = \int_I d\beta \frac{1}{\sqrt{2\pi\sigma_\chi^2(\beta)}}\exp\left[-\frac{\left(q - |\alpha|\sqrt{2}\cos\beta\right)^2}{2\sigma_\chi^2(\beta)}\right]. \tag{12}$$

Note that the value of $\bar{z}_0$ only depends on the value q of the outcome: we expect that the bigger is the difference $q - \bar{\chi}_0$, the smaller the posterior probability $\bar{z}_0$. In order to give a numerical estimate let us set $|\alpha| = 10$ and $r = 1$, i.e. we consider a squeezed state with one hundred photons, where only about 1% of the total energy is employed in squeezing. Recalling that the null hypothesis corresponds to the condition $\beta = \beta_0 = 0$ we get $\bar{\chi}_0(0) = 14.14$ and $\sigma_\chi^2(0) = 0.07$. By explicit calculations one can check that for $|q - \bar{\chi}_0(0)| \leq 4\sigma_\chi(0)$ the posterior probability is still over 50%: a Bayesian approach leads us to conclude that no phase perturbation actually occurred. That is, if $2\sigma_\chi(0) \leq |q - \bar{\chi}_0(0)| \leq 4\sigma_\chi(0)$ the Bayesian approach provides conflicting evidence with respect to the values obtained by a sampling test: we are in presence of the paradox.

The key point, in the present case as well as in the situation illustrated by Lindley, is that the prior is taken to be flat. Notice that despite this assumption is somehow consistent from a physical perspective, in practice it turns out that outcomes being several $\sigma_\chi(0)$ from the expectation value are associated to statistical fluctuations and the null hypothesis keeps on overwhelming any alternative. For this reason we shall rather conclude that the Bayesian approach is somehow misleading if a scarce information about the prior distributions is available, whereas a sampling-theory test better accounts what is really happening.

Since our main goal is to address the occurrence of the Lindley paradox in interferometry we will not go further in the discussion of this case. We only notice that here we have a physical example where the paradox unavoidably arises, the assumptions being consistent with the dynamics of the system. From a practical point of view, we notice that the characterization and the calibration of homodyne detectors is one of the building blocks for the robust implementation of quantum tomography of the radiation field [34].

4. Phase estimation with Mach–Zehnder interferometry

In the present section we consider a MZ interferometer to monitor a classical phase parameter via photon counting measurements. A schematic diagram of the detector is shown in Fig. 2.

Two modes of the radiation field enters the interferometer through a first beam splitter, then a phase difference ϕ is inserted between the beams and, finally, the two fields exit through a second beam splitter. The number of photons in each mode is detected by means of two photodetectors of quantum efficiency η that measure the photocurrents i_f and i_e; a differential amplifier eventually determines the difference photocurrent, whose possible
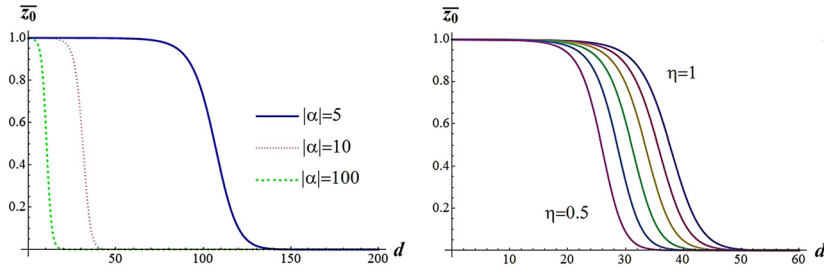
Fig. 3. Bayesian posterior probability $\bar{z}_0(d)$ in coherent light interferometry with diffuse prior. Left: we show $\bar{z}_0(d)$ as a function of d for different values of $|\alpha|$ and fixed quantum efficiency $\eta = 0.7$. Right: $\bar{z}_0(d)$ as a function of d for different values of η and fixed photon number $\bar{n} = 100$. Given the distance $t = (d - \bar{d}_{\frac{\pi}{2}})/\sigma_d$ between the outcome and the expectation value the probability $\bar{z}_0(d)$ varies slowly with η. (For interpretation of the colors in this figure, the reader is referred to the web version of this article.)

values correspond to the eigenvalues of the Hermitian operator

$$\hat{D}(\phi) = \eta \left(\hat{f}^\dagger \hat{f} - \hat{e}^\dagger \hat{e} \right) \tag{13}$$

which is a normalized version of the integrated photocurrent. It can be expressed in terms of the input fields as follows:

$$\hat{D}(\phi) = \eta \left[\left(\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b} \right) \cos\phi + i \left(\hat{b}^\dagger \hat{a} - \hat{a}^\dagger \hat{b} \right) \sin\phi \right]. \tag{14}$$

From the above equation it is clear that measurements of the difference photocurrent at the output allows one to assess variations of the phase difference ϕ. In particular, from error propagation one obtains the following expression for the sensitivity $\Delta\phi$ of the device:

$$\Delta\phi = \sqrt{\left\langle \hat{D}^2(\phi) \right\rangle - \left\langle \hat{D}(\phi) \right\rangle^2} \left| \frac{\partial}{\partial\phi} \left\langle \hat{D}(\phi) \right\rangle \right|_{\phi=\phi_0}^{-1}, \tag{15}$$

which depends from the initial value ϕ_0 and the expectation value $\left\langle \hat{D}(\phi) \right\rangle$. If any perturbations changes the length of the two arms of the interferometer, the optical paths covered by the light beams is altered. As a consequence, the phase difference ϕ will vary and so will do the outcomes of measurements of the operator $\hat{D}$: once the system has been set in a proper initial working point ϕ_0 an outcome $\phi \neq \phi_0$ shall be associated to a displacement. Hence, the data analysis consists of assessing the probability that an outcome d confirms the null hypothesis H_0, corresponding to no phase variation, or the alternative hypothesis H_1 that $\phi \neq \phi_0$. In this context, we discuss the possibility of coming across the Lindley paradox, paying particular attention to the conditions in which it can arise.

4.1. Coherent light interferometry

At first, we consider a classical-like MZ interferometer fed by an input state of the form $|\psi\rangle = |\alpha\rangle |0\rangle$ where $\alpha = |\alpha| e^{i\theta} \in \mathbb{C}$. In this case, the outcomes d of measurements of the operator $\hat{D}$ are distributed according to a Skellam distribution [35]:

$$p_\phi(d) = e^{-\eta|\alpha|^2} \left(\frac{1 + \cos\phi}{1 - \cos\phi} \right)^{\frac{d}{2}} I_{|d|} \left(\eta |\alpha|^2 |\sin\phi| \right), \tag{16}$$

where $I_k(z)$ is the modified Bessel function of the first kind. The mean value $\bar{d}_\phi$ and variance σ_d^2 of the above distribution are given by

$$\bar{d}_\phi = \left\langle \hat{D}(\phi) \right\rangle = \eta |\alpha|^2 \cos\phi, \tag{17}$$

$$\sigma_d^2 = \Delta^2 \hat{D} = \eta |\alpha|^2. \tag{18}$$

This result may be easily understood recalling that the number of photons contained in a coherent state follows a Poissonian statistics and that the Skellam distribution provides the probability density of the difference between two Poisson variates. It is easy to see

that for the given input state $|\psi\rangle$ the minimum value of $\Delta\phi$ is obtained for $\phi_0 = \frac{\pi}{2}$. Therefore, we assume the interferometer to be set in this optimal working point (corresponding to a null mean value $\bar{d}_{\frac{\pi}{2}} = 0$).

We now suppose that a trial results in the outcome d of the difference photocurrent, that we represent as $d = \bar{d}_{\frac{\pi}{2}} + t \cdot \sigma_d = t \cdot \sigma_d$, with $t \in \mathbb{R}$. According to a standard sampling-theory test d is to be associated to a displacement induced by a perturbation if its value is several σ_d from 0. In this event we shall reject the null hypothesis H_0. On the other side, in order to give a Bayesian assessment to H_0 we assign a probability z_0 to the null hypothesis and distribute the remainder over the interval $I = [0, \pi]$ where we expect the phase to vary. Assuming to deal with rare events we take $z_0 = 0.99$, i.e. we assume that almost only 1 event out of 100 is to be associated to an incoming perturbation rather than to a random error. With respect to the prior distribution for the alternative hypothesis we initially consider a situation where no information about the signal is available so that a flat prior results a proper choice. At this point we can compute the posterior probability $\bar{z}_0$ given $\phi = \phi_0 = \frac{\pi}{2}$ by Eq. (10), where β is to be substituted with ϕ and the probability densities take the following expressions:

$$p(d|\phi = \phi_0) = \int_I d\phi \, p_\phi(d) \, \delta(\phi - \phi_0)$$

$$= e^{-\eta|\alpha|^2} I_{|d|}(\eta |\alpha|^2), \tag{19}$$

$$p(d|\phi \neq \phi_0) = \int_I d\phi \, p_\phi(d). \tag{20}$$

The behavior of $\bar{z}_0(d)$ with respect to the number of photons injected and the quantum efficiency of the detectors is reported in Fig. 3. We note that the trend is the same for different $\bar{n}$ and η: the presence of non-unit quantum efficiency in the detection process does not change the shape of $\bar{z}_0(d)$ and the same goes for the light intensity. In facts, the overall effect is to reduce the posterior probabilities, i.e. given an outcome d and $\eta_1 < \eta_2$ we have $\bar{z}_{0,1}(d) < \bar{z}_{0,2}(d)$, i.e. we get a lower value of $\bar{z}_0(d)$ when we lower down the efficiency of the apparatus.

That being said, for a given number of photons impinging on the device and efficiency of the photocounters, we can identify three different regions corresponding to different values of the outcome. In the range $0 \leq d \leq 2\sigma_d$ the posterior probability is approximately 1, in agreement with a sampling-theory approach. In this case the phase shift is so little that its origin is much more likely to be a random fluctuation rather than an actual perturbation. If $d \geq 4\sigma_d$ the posterior probability is approximately 0: both the Bayesian and the Frequentist approach affirm that the phase shift has been induced by a perturbation. At last, $\bar{z}_0(d)$ falls from 1 to 0 as the value of the outcome increases within the interval $2\sigma_d \leq d \leq 4\sigma_d$. Since a frequentist analysis shall provide a value of
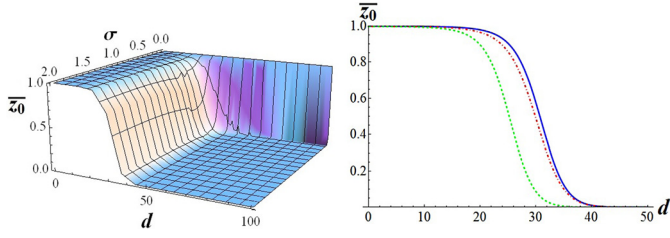
Fig. 4. Posterior probability $\bar{z}_0$ in coherent light interferometry with wrapped normal prior. Left: $\bar{z}_0(d)$ as function of d and σ with $\eta = 0.7$. The coherent amplitude is $\alpha = 10$ so that $\bar{n} = 100$ and $\sigma_d^2 = 70$. An actual disagreement with respect to a frequentist approach occurs for $0.2 \leq \sigma \cap 2\sigma_d \leq d \leq 4\sigma_d$. Right: $\bar{z}_0(d)$ as a function of d. The continuous blue line and the dot-dashed red line are obtained for $\eta = 0.7$ and $\sigma = 1$; $\sigma = 0.5$ respectively. The dashed green line refers to the case $\eta = 0.5$ and $\sigma = 0.5$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$\bar{z}_0$ below 0.05 we have that in this region we are in presence of the paradox.

On the other hand, one may find uneasy to say that an outcome four times σ_d far from the mean value is to be associated to statistical fluctuations. In facts, we can describe what is happening for $2\sigma_d \leq d \leq 4\sigma_d$ as follows. Our measurement d of the difference photocount gave a random value that was manifestly inconsistent with the expected value in absence of interaction; however, as a consequence of the prior assumed we are led to conclude that it would be even more improbable to observe that specific outcome were it the result of an interaction. That is, what makes the paradox arise is that the prior asks us to regard at every alternative possibility as an improbable coincidence. For this reason, in a real experiment an outcome in the range $2\sigma_d \leq d \leq 4\sigma_d$ should be interpreted as evidence of detection.

In approaching the most realistic description, we now discuss the occurrence of the paradox when a less diffuse prior is assumed. Here we consider a Gaussian-like distribution as a prior in order to account for our belief that a perturbation results in a small phase-shift around the expected value. In particular, we consider the so-called *wrapped normal* distribution, defined as follows:

$$g_{\sigma,\phi_0}(\phi) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{(\phi - \phi_0 + k\pi)^2}{2\sigma^2}\right]. \quad (21)$$

The wrapped distribution is normalized on the interval I and is characterized by two parameters, ϕ_0 and σ^2, which correspond to the expectation value and variance of a normal distribution. In recalling that the mean value of the outcomes is $\bar{d}_{\frac{\pi}{2}} = 0$ we impose $\phi_0 = \frac{\pi}{2}$ but we let the variance of the prior free to vary in order to assess the relevance of our knowledge about the system with respect to the statistics. Note that since $\bar{z}_0$ is inversely proportional to the ratio $p(d|\phi \neq \phi_0) / p(d|\phi = \phi_0)$, which increases as the variance of the prior assumed for the alternative hypothesis decreases, we expect that the sharper is the prior considered the higher will be the value of $\bar{z}_0$. For this reason we may expect the Lindley paradox not to occur when the variance σ of $g_{\sigma,\phi_0}(\phi)$ is taken to be little with respect to the one of the Skellam distribution, given by Eq. (17). We report in Fig. 4 the results obtained substituting equation (16) with the proper expression for the prior (21), so that:

$$p(d|\phi \neq \phi_0) = \int_I d\phi \, p_\phi(d) \, g_{\sigma,\frac{\pi}{2}}(\phi). \quad (22)$$

At first, we see that if the variance of the prior is taken too small then the posterior probability does not evolve (with reference to the plotted range for d): we get $\bar{z}_0 = z_0$ independently from the outcome d. From a mathematical perspective, this is due to the fact that the ratio $p(d|\phi = \phi_0) / p(d|\phi \neq \phi_0)$ appearing in

the expression of $\bar{z}_0$ tends to unity as $p(d|\phi \neq \phi_0)$ tends to a delta $\delta(\phi - \phi_0)$. In facts, we are assuming a sharp prior centered in the mean value both for the null hypothesis and the alternative hypothesis. That is, we are somehow stating that H_0 and H_1 are equivalent as both of them, according to our prior models, lead to a measurement of d to be associated to the outcome $d = \bar{d}_{\frac{\pi}{2}}$. Therefore, after a trial we cannot conclude anything else but that it is to be associated to a measurement of the mean value. Similarly, an evaluation of the posterior probability for the alternative hypothesis $p(d|\phi \neq \phi_0)$ yields $\overline{1 - z_0} = 1 - z_0$. This last remark makes it clear that we are providing too much prior information and Bayesian inference can only confirm what we already know, without the possibility of improving our knowledge about the system. Conversely, if the prior distribution is diffuse, i.e. $\sigma > 1$, we recover the case of flat prior examined above, where our knowledge is poor and, in turn, for $2\sigma_d \leq d \leq 4\sigma_d$ the paradox is again observed.

The most interesting region is $0.2 \leq \sigma \leq 1$. A comparison between the values of $\bar{z}_0$ obtained employing a widespread prior and a sharper one, see the right panel of Fig. 4, shows that the latter distribution awards less probability to the null hypothesis, as we were expected to. However, the weight of the prior is not sufficient to avoid the paradox, which still occurs if $2\sigma_d \leq d \leq 4\sigma_d$ even for small values of the variance of the wrapped normal distribution. Fig. 4 also suggests that if the interferometer has poor detectors then the odds in favor of the alternative hypothesis will be lower. Nevertheless, in recalling that σ_d decreases with η it is possible to see that the net effect remains unchanged.

4.2. Squeezed light interferometry

Let us now consider an input state of the form $|\psi\rangle = |\alpha\rangle |\lambda\rangle$, where $\alpha = |\alpha|e^{i\theta}$, $\lambda = re^{i\varphi} \in \mathbb{C}$, $r > 0$. That is, we replace the vacuum state $|0\rangle$ in mode b with a squeezed-vacuum state $|\lambda\rangle$, containing $\bar{n} = \sinh^2(r)$ photons. The advantages gained by squeezing enhancement have been widely studied in literature. Here, our aim is to assess the convenience of employing squeezed light from a statistical perspective. That is, we address the possibility of getting by the Lindley paradox when the injected beam is previously squeezed.

Similarly to the previous case we will evaluate the posterior probability using a flat prior, at first, and then a peaked distribution. At the same time, we will pay attention to the contribution of the quantum efficiency. The difference with respect to the classical scheme seats in the distribution of the outcomes $p_{\alpha,\lambda,\phi}(d)$, which turns out to be difficult to be computed. However, since the expectation value $\bar{d}_\phi$ and the variance $\sigma_d^2(\phi)$ may be calculated exactly we employ a Gaussian distribution for the outcomes as a first-order approximation:

$$p_{\alpha,\lambda,\phi}(d) = \frac{1}{\sqrt{2\pi\sigma_d^2(\phi)}} e^{-[d - \eta\cos\phi(|\alpha|^2 - \sinh^2 r)]^2 / 2\sigma_d^2(\phi)}, \quad (23)$$

where the variance is a function of the initial working point ϕ_0 and of the actual phase shift ϕ. Similarly to what happened in the previous case, σ_d^2 takes the minimum value at $\phi_0 = \pi/2$, namely:

$$\sigma_d^2(\phi) = \frac{1}{2}\eta(1 - \eta)(|\alpha|^2 + \sinh^2 r)$$

$$+ \eta^2 \left\{ |\alpha|^2 + 2\sinh^2 r \cosh^2 r \cos^2 \phi \right.$$

$$\left. + \sin^2 \phi \left[\sinh^2 r(1 + 2\alpha^2) - 2\alpha^2 \sinh r \cosh r\right] \right\} \quad (24)$$

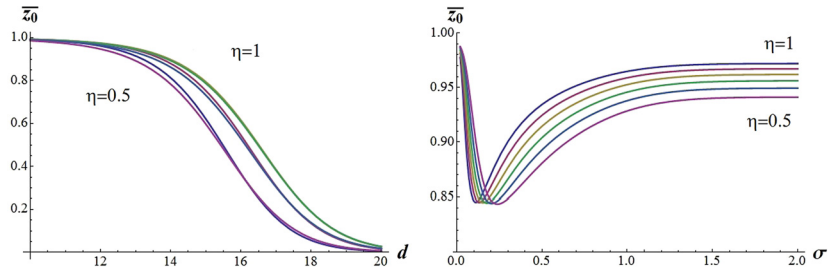whereas the expectation value results $\bar{d}_{\frac{\pi}{2}} = 0$.

Fig. 5. Bayesian posterior probability $\bar{z}_0(d)$ in squeezed light interferometry. Left: $\bar{z}_0(d)$ as a function of d for efficiency η from 0.5 to 1.0 and for $\alpha = 10$, $r = 1$, $\varphi = \pi$. Right: the posterior probability for given distance $|d - \bar{d}_\phi| = 3\sigma_d$ and the same values of the state parameters (while $\phi_0 = \pi/2$, $\phi = 0$ in σ_d).

In order to test the consistency of an outcome $d = \bar{d}_{\frac{\pi}{2}} + t \cdot \sigma_d = t \cdot \sigma_d$ with the null hypothesis H_0 we pose $z_0 = 0.99$ and assume that the remainder probability $1 - z_0$ is uniformly distributed over the interval $I = [0, \pi]$. Before computing the posterior probability we point out that the best sensitivity $\Delta\phi$ is achieved for $\varphi = \pi$ and $|\alpha|^2 \gg \sinh^2(r)$. The posterior probability $\bar{z}_0$ that $\phi = \phi_0$ is, then, obtained by means of equation (10), with β equal to ϕ, leading to

$$p(d|\phi = \phi_0) = \int_I d\phi \, p_{\alpha,\lambda,\phi}(d) \, \delta(\phi - \phi_0) \tag{25}$$

$$p(d|\phi \neq \phi_0) = \int_I d\phi \, p_{\alpha,\lambda,\phi}(d). \tag{26}$$

The behavior obtained for $\bar{z}_0$ is very close to that observed for its classical counterpart, i.e. the dashed green line in Fig. 3. Indeed, we can identify a region where $\bar{z}_0$ tends to unity (for small values of d), a region where it is approximately zero (to the right) and a "transition area", where the posterior probability varies between these two limits. What matters most is that we are again in presence of the Lindley paradox: if we test the hypothesis concerning the detection of a perturbation by means of the Bayes' theorem we will end up in a misleading analysis. It is possible to see from the plot of Fig. 5 that the situation does not change in presence of non-ideal photodetectors. In the same vein, when we consider input states containing a higher number of photons and such that the optimal conditions mentioned above are still satisfied (i.e. $\varphi = \pi$ and $|\alpha|^2 \gg \sinh^2 r$), the posterior probability approaches zero for lower values of d/σ_d. However, the contribution gained from the use of higher intensity light is not sufficient to make it disappear.

At last, we take into account the case of squeezed input light with a Gaussian-like prior distribution in the statistical description. The change in the shape of $p(d|\phi \neq \phi_0)$ has the usual effect of decreasing the value of $\bar{z}_0$, as it is illustrated in Fig. 5. Nevertheless, we observe that neither in this conditions we can obtain a consistent Bayesian statistical inference. That is, the value of the posterior probability does not decrease, not even taking an appreciably peaked prior, until the outcome is more than several σ_d far from the expectation value.

5. Conclusions

We have addressed the occurrence of the so-called Lindley paradox in the analysis of data coming from homodyne detection and optical interferometry. We found that the Lindley paradox may indeed occur. In particular, we have shown that the Bayesian approach is somehow misleading if a scarce information about the prior distributions is available, as it happens in the evaluation of the effect of an external perturbation, and sampling-theory should be preferred. Concerning MZ interferometers, we have shown that Lindley paradox appears both for coherent and squeezed signals and is present for any value of the quantum efficiency of the involved detectors. On the other hand, the disagreement between Bayesian and frequentist approach is less pronounced for increasing noise and may be softened by using a suitable, more localized, priors.

Our results, besides being of fundamental interest for interferometry, are of practical significance for quantum state reconstruction, where calibration of homodyne detectors represents a crucial step for the implementation of quantum tomography [34].

Appendix A. Posterior probabilities

Consider a physical system where the quantity X may be measured and assume that the outcomes from the measurement of X are distributed according to a normal distribution with an unknown mean value ϕ and a known variance σ^2. We perform a measurement of X and on the basis of the outcome x we want to test which of the two following hypothesis is true

$$H_0 : \text{the mean value is equal to a given value } \phi = \phi_0 \tag{A.1}$$

$$H_1 : \text{the mean value is not the value above } \phi \neq \phi_0. \tag{A.2}$$

We assume to have some prior knowledge on the system, which may be expressed in form of some a priori probabilities $p(H_0)$ and $p(H_1) = 1 - p(H_0)$ for the two hypothesis. The measurement of X is, in turn, intended to upgrade our knowledge of the system. Given the result x, Bayes theorem says that

$$p(H_k|x)p(x) = p(x|H_k)p(H_k),$$

for $k = 0, 1$, where $p(H_k)$ is the a priori probability introduced above, $p(x|H_k)$ is the conditional distribution of the outcomes given a hypothesis, $p(x) = p(x|H_0)p(H_0) + p(x|H_1)p(H_1)$ is the overall distribution of the outcomes, independently which hypothesis is actually true, and $p(H_k|x)$ is the a posteriori probability of the hypothesis H_k, i.e. the quantity of interest here. Using the Bayes theorem we may evaluate the a posteriori probabilities, e.g. $p(H_0|x)$ is given by

$$\begin{aligned}
p(H_0|x) &= \frac{p(x|H_0)p(H_0)}{p(x)} \\
&= \frac{p(x|H_0)p(H_0)}{p(x|H_0)p(H_0) + p(x|H_1)p(H_1)} \\
&= \left[1 + \frac{p(x|H_1)p(H_1)}{p(x|H_0)p(H_0)}\right]^{-1} \\
&= \left[1 + \frac{1 - p(H_0)}{p(H_0)} \frac{p(x|H_1)}{p(x|H_0)}\right]^{-1},
\end{aligned} \tag{A.3}$$

which, upon the substitutions $p(H_0) \to z_0$ and $p(H_0|x) \to \bar{z}_0$, coincides with Eq. (1).

Now, in order to evaluate explicitly the posterior distribution we assign some prior distribution $\pi_k(\phi) \equiv p(\phi|H_k)$ to the mean value under the two hypothesis. According to (A.1) and (A.2) we write

$$\pi_0(\phi) = \delta(\phi - \phi_0)$$

and assign the rest of the prior probability as a normal distribution with variance τ, i.e.

$$\pi_1(\phi) = (2\pi\tau^2)^{-1/2} \exp\{-(\phi - \phi_0)^2 / 2\tau^2\}.$$

It is now straightforward to evaluate the conditional probabilities as

$$p(x|H_k) = \int d\phi\, \pi_k(\phi)\, p(x|\phi),$$

which leads to

$$p(x|H_0) = (2\pi\sigma^2)^{-1/2} e^{-(\phi - \phi_0)^2 / 2\sigma^2} \tag{A.4}$$

$$p(x|H_1) = [2\pi(\sigma^2 + \tau^2)]^{-1/2} e^{-(\phi - \phi_0)^2 / 2(\sigma^2 + \tau^2)}. \tag{A.5}$$

Overall, Eq. (A.3) rewrites as

$$p(H_0|x) = \left[1 + \frac{1 - p(H_0)}{p(H_0)} \sqrt{\frac{\sigma^2}{\sigma^2 + \tau^2}} \right.$$
$$\left. \times \exp\left(\frac{\tau^2}{2\sigma^2} \frac{(x - \phi_0)^2}{\sigma^2 + \tau^2} \right) \right]^{-1}. \tag{A.6}$$

Upon substituting $\phi_0 = 0$ and $\sigma = 1$ we arrive at Eq. (2).

References

[1] K.S. Thorne, Rev. Mod. Phys. 52 (1980) 285.
[2] R.X. Adhikari, Rev. Mod. Phys. 86 (2014) 121.
[3] V.B. Braginsky, Astron. Lett. 34 (2008) 558.
[4] V.B. Braginsky, F.Ya. Khalili, K.S. Thorne, Quantum Measurement, Cambridge University Press, Cambridge, 1992.
[5] C.M. Caves, Phys. Rev. D 23 (1981) 1693.
[6] R.S. Bondurant, J.H. Shapiro, Phys. Rev. A 30 (1984) 2548.
[7] B. Yurke, S.L. McCall, J.R. Klauder, Phys. Rev. A 33 (1986) 4033.
[8] M.J. Holland, K. Burnett, Phys. Rev. Lett. 71 (1993) 1355.
[9] F.W. Sun, B.H. Liu, Y.X. Gong, Y.F. Huang, Z.Y. Ou, G.C. Guo, Europhys. Lett. 82 (2008) 24001.
[10] M.G.A. Paris, Phys. Lett. A 201 (1995) 132.
[11] The LIGO Scientific Collaboration, Nat. Photonics 7 (2013) 613.
[12] A.A. Berni, T. Gehring, B.M. Nielsen, V. Handchen, M.G.A. Paris, U.L. Andersen, Nat. Photonics 9 (2015) 577.
[13] M.G.A. Paris, Phys. Lett. A 225 (1997) 23.
[14] L. Pezzé, A. Smerzi, G. Khoury, J.F. Hodelin, D. Bouwmeester, Phys. Rev. Lett. 99 (2007) 223602.
[15] L. Pezzé, A. Smerzi, Phys. Rev. Lett. 100 (2008) 073601.
[16] S. Olivares, M.G.A. Paris, J. Phys. B 42 (2009) 055506.
[17] B. Teklu, S. Olivares, M.G.A. Paris, J. Phys. B 42 (2009) 035502.
[18] M.G. Genoni, S. Olivares, M.G.A. Paris, Phys. Rev. Lett. 106 (2011) 153603.
[19] M.G. Genoni, S. Olivares, D. Brivio, S. Cialdi, D. Cipriani, A. Santamato, S. Vezzoli, M.G.A. Paris, Phys. Rev. A 85 (2012) 043817.
[20] S. Olivares, S. Cialdi, F. Castelli, M.G.A. Paris, Phys. Rev. A 87 (2013) 050303(R).
[21] H. Jeffrey, Theory of Probability, Oxford Univ. Press, 1948.
[22] D.V. Lindley, Biometrika 44 (1957) 187.
[23] Z. Hradil, R. Myska, J. Perina, M. Zawisky, Y. Hasegawa, H. Rauch, Phys. Rev. Lett. 76 (1996) 4295.
[24] J.P. Dowling, Phys. Rev. A 57 (1998) 4736.
[25] G.M. D'Ariano, M.G.A. Paris, R. Seno, Phys. Rev. A 54 (1996) 4495.
[26] M.G.A. Paris, Int. J. Quantum Inf. 7 (2009) 125.
[27] D.V. Lindley, Biometrika 64 (1977) 207.
[28] G. Shafer, J. Am. Stat. Assoc. 77 (1982) 325.
[29] H.M. Wiseman, D.W. Berry, S.D. Bartlett, B.L. Higgins, G.J. Pryde, IEEE J. Sel. Top. Quantum Electron. 15 (2009) 1661.
[30] A. De Kumar, D. Roy, D. Goswami, Phys. Rev. A 83 (2011) 015402.
[31] T. Reisinger, G. Bracco, B. Holst, New J. Phys. 13 (2011) 065016.
[32] B.T. Kirby, J.D. Franson, Phys. Rev. A 89 (2014) 033861.
[33] X.M. Feng, G.R. Jin, W. Yang, Phys. Rev. A 90 (2014) 013807.
[34] M.G.A. Paris, J. Rehacek (Eds.), Quantum State Estimation, Lect. Notes Phys., vol. 649, Springer, Berlin, 2004.
[35] J.G. Skellam, J. R. Stat. Soc. 109 (1946) 296.