

# Compensation of self-phase modulation in a QND measurement of photon number based on the optical Kerr effect

Hans Martens and Willem M de Muynck

Department of Theoretical Physics, Eindhoven University of Technology, PO Box 513  
5600 MB Eindhoven, The Netherlands

Received 12 December 1991, in final form 6 April 1992

**Abstract.** We study a quantum non-demolition (QND) set-up for the measurement of photon number, employing the optical Kerr effect. A method for the compensation of the negative effects of self-phase modulation is discussed. Losses are taken into account. It is seen that the performance of the device is limited in the first place by losses, rather than by self-phase modulation.

## 1. Introduction

The conventional way of counting photons, i.e. detecting them with a photomultiplier or some such device, results in their annihilation. In recent years, however, several proposals have been made to measure photon numbers without absorbing the light upon which the measurement is performed [1–8]. In such a *quantum non-demolition* (QND) scheme it is possible to perform a second measurement on the outgoing signal beam yielding information about phase [9]. This realises an analogue of Heisenberg's famous  $\gamma$ -microscope *Gedanken* experiment for the joint measurement of two incompatible observables, thus making the QND scheme also interesting from the point of view of the quantum theory of measurement [10].

Basic to the problem of the joint measurement of incompatible observables is the inaccuracy or non-ideality induced in the measurement results of one observable by adapting the measurement arrangement so as to yield also information on the second one. A theory of non-ideal quantum measurements was developed [11], allowing a characterization of the non-ideality thus induced. The existence was demonstrated of a theoretical lower bound to the measurement accuracies.

We first discuss in section 2 a basic set-up, based on the optical Kerr effect. This device can be seen as an application of the theory of the non-ideal measurement of photon number. Contributions to the non-ideality stem from: (i) self-phase modulation caused by an intensity-dependent index of refraction, and inducing additional noise in the photocurrent, (ii) losses due to absorption in the Kerr medium. After discussing in section 3 the non-ideality of this experiment in the idealized case in which self-phase modulation and losses can be neglected, in section 4 the influence of self-phase modulation is discussed. It is demonstrated that in the basic set-up this non-linear effect increases the inaccuracy of the measurement, and cannot be compensated

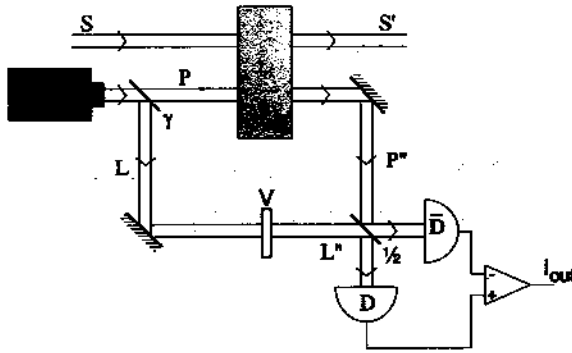


Figure 1. Basic set-up. The signal beam  $S$  is passed through a non-linear Kerr medium  $K$ , placed in a Mach-Zehnder interferometer. This interferometer is driven by a laser, emitting a field in coherent state  $|\beta\rangle$ . The transmittivity of the first beamsplitter is  $\gamma$ . The paper focuses on a  $\gamma = \frac{1}{2}$  scheme.  $V$  indicates a delay, which controls the reference beam phase;  $D$  and  $\bar{D}$  are two identical detectors. Primes label outgoing fields.

for completely. An alternative measurement set-up is proposed in which this effect can be compensated for completely in the lossless case. In section 5 the influence of losses on the functioning of the alternative device as a non-ideal photon meter is discussed. It is demonstrated that, although the losses reduce the effectiveness of the compensation, the alternative set-up is still preferable for small interaction times. Finally, in section 6 the dependence on losses of the alternative set-up as a QND measuring apparatus is investigated.

## 2. Basic set-up

One type of QND scheme involves an optical Kerr medium placed in a Mach-Zehnder interferometer [1, 2]. The basic set-up is shown in figure 1. A probe beam  $P$  and a signal beam  $S$  are passed through a medium  $K$  (length  $\ell$ ) with a third-order non-linearity. Assuming both signal and probe fields to be single mode, the effect of the medium can be described by the application of an interaction Hamiltonian<sup>†</sup>  $\hat{H}_i$  for a time  $\tau = \ell/c$  [1, 2],

$$\hat{H}_i = \chi^{(3)} \hat{N}_S \hat{N}_P + \frac{1}{4} \chi^{(3)} \hat{N}_P^2 + \frac{1}{4} \chi^{(3)} \hat{N}_S^2 \quad (1)$$

(carets denote operators;  $\hbar = 1$ ; indices indicate the fields the operators work on, corresponding to figure 1). Since  $\hat{H}_i$  commutes with the free field Hamiltonian  $\hat{H}_0 = (\hat{N}_S + \frac{1}{2})\omega_S + (\hat{N}_P + \frac{1}{2})\omega_P$ , we can use in the following the interaction picture having  $\hat{H}_i$  as a Hamiltonian. As the first term in (1) indicates, the probe beam acquires a phase shift proportional to the signal photon number  $\hat{N}_S$ . This phase is measured by mixing the probe beam and a reference beam  $L$  in a balanced detector: we measure the difference in the photocounts of detectors  $D$  and  $\bar{D}$ . Consequently our read-out observable, referred to as  $\hat{X}_{\text{meas}}$ , is

$$\hat{X}_{\text{meas}} = \hat{N}_D - \hat{N}_{\bar{D}} = i(\hat{a}_P^\dagger \hat{a}_L - \hat{a}_P \hat{a}_L^\dagger). \quad (2)$$

<sup>†</sup> The  $\chi$  for the three terms in (1) are in principle different, depending on frequency and polarization of the light fields. This has been neglected. An analysis with three different  $\chi$  would, however, be completely analogous to the one presented here.

We shall at first neglect the last two terms in (1). Then we can derive for the output of this scheme (Heisenberg picture)

$$\hat{X}_{\text{meas}} = i[\hat{a}_p^\dagger \exp(i\tau\chi^{(3)}\hat{N}_s + i\varphi)\hat{a}_L - \hat{a}_p \exp(-i\tau\chi^{(3)}\hat{N}_s - i\varphi)\hat{a}_L^\dagger]. \quad (3)$$

We have included a  $\pi/2$  phase jump for every reflection. The reference beam phase  $\varphi$  can be controlled via a phaseshifter (V in figure 1).

The probe beam and the reference beam L are obtained via a beam splitter from a single source, a laser that produces a coherent state  $|\beta\rangle$ . Accordingly, the probe mode is brought into the state  $|\sqrt{\gamma\beta}\rangle_P$ , the reference mode into  $|i\sqrt{1-\gamma\beta}\rangle_L$ , again including a  $\pi/2$  phase jump from the reflection. Here  $\gamma$  denotes the transmittivity of the first beamsplitter (see figure 1). In the present paper we focus on a set-up with  $\gamma = \frac{1}{2}$ . A second type of scheme involves a homodyne detection of the probe beam. This can be treated analogously to the following, if we let  $\gamma \rightarrow 0$ , while keeping  $\sqrt{\gamma\beta} = \tilde{\beta}$  constant.

### 3. Non-ideality and measurement characterization

As shown by (3), a read-out of  $\hat{X}_{\text{meas}}$  will yield information on the photon number in the incoming signal state,  $\hat{N}_s$ . In fact, the probability that the  $\hat{X}_{\text{meas}}$  readout will yield outcome  $x$ , can be written in the form

$$\text{prob}_{\hat{X}_{\text{meas}}}(x) = \sum_n \lambda(x|n) \text{prob}_{\hat{N}_s}(n) \quad \forall_{n,x} \lambda(x|n) \geq 0, \quad \forall_n \int \lambda(x|n) dx = 1. \quad (4)$$

From (4) we see two things. Firstly, the outcome probabilities are not generated by a *projection-valued measure*, as prescribed by von Neumann, but by a more general type of object: a *positive operator-valued measure* [12] (POVM). A discrete POVM  $\{\hat{M}_k\}$  is a set of operators that satisfies

$$\forall_k \hat{M}_k \geq \hat{0} \quad \sum_k \hat{M}_k = \hat{1}. \quad (5)$$

For an object state  $\rho$ , the probability of obtaining outcome  $k$  is then given by  $\text{Tr}(\rho \hat{M}_k)$ . We do not demand that  $\hat{M}_k^2 = \hat{M}_k$ , so that the  $\hat{M}_k$  are not necessarily projectors.

The second thing we note in (4) is the particular form of the outcome distribution. It is a 'smeared version' of the  $\hat{N}_s$  distribution. Even if the input photon number is sharp, the measurement outcome is not quite certain. Elsewhere [11] we have used the term *non-ideality* for this relationship between the measurement POVM and the observable it is intended to measure, here  $\hat{N}_s$ . Thus, although we do not realize a measurement of photon number in the sense of von Neumann, the Kerr device can still be seen as a non-ideal photon number meter.

The precise measurement statistics can be derived once probe and reference beam states are given. In the set-up of figure 1 these are the coherent states  $|\sqrt{\gamma\beta}\rangle_P$  and  $|i\sqrt{1-\gamma\beta}\rangle_L$ , as we saw in the previous section. We take  $\varphi = \pi/2$ , which is optimal in a  $\gamma = \frac{1}{2}$  scheme [6]. Taking expectation values over the probe and reference states, equation (3) gives

$$\langle \hat{X}_{\text{meas}} \rangle_{P+L, n_s} = |\beta|^2 \sin(\tau\chi^{(3)}n_s) \quad (6)$$

conditional on the S-photon number  $n_s$ . More generally the conditional probability distribution of the output variable  $\hat{X}_{\text{meas}}$  can be obtained from the characteristic function

$$\langle \exp(ic\hat{X}_{\text{meas}}) \rangle_{P+L, n_s} = \exp[|\beta|^2 (\cos(c) - 1) + i|\beta|^2 \sin(\tau\chi^{(3)}n_s)\sin(c)]. \quad (7)$$

If the probe beam is strong enough, (7) may be approximated by

$$\langle \exp(i c \hat{X}_{\text{meas}}) \rangle_{P+L, n_S} \approx \exp[-\frac{1}{2} |\beta|^2 c^2 + i |\beta|^2 \sin(\tau \chi^{(3)} n_S) c]. \quad (8)$$

This is the characteristic function of a Gaussian with mean  $\mu$  and standard deviation  $\sigma$  given by

$$\begin{aligned} \mu &= \langle \hat{X}_{\text{meas}} \rangle_{P+L, n_S} = |\beta|^2 \sin(\tau \chi^{(3)} n_S) \\ \sigma^2 &= \langle \Delta^2 \hat{X}_{\text{meas}} \rangle_{P+L, n_S} = |\beta|^2. \end{aligned} \quad (9)$$

Here  $\langle \Delta^2 \hat{X}_{\text{meas}} \rangle_{P+L, n_S} := \langle (\hat{X}_{\text{meas}} - \langle \hat{X}_{\text{meas}} \rangle)^2 \rangle_{P+L, n_S}$  denotes the read-out variance, conditional on the signal photon number  $n_S$ . The set-up is therefore characterized by the noise  $\sigma$  and the gain  $G$ ,  $G := |\partial \mu / \partial n_S|$ . These may be combined in a measure for the device's inaccuracy as an  $\hat{N}_S$  meter:

$$\delta_{N_S} := \sigma / G = \frac{1}{\tau \chi^{(3)} |\beta \cos(\tau \chi^{(3)} n_S)|}. \quad (10)$$

The device functions linearly in the low photon number regime ( $\tau \chi^{(3)} \langle \hat{N}_S \rangle \ll 1$ ), because there we may approximate (9) and (10) according to

$$\left. \begin{aligned} \mu &\approx \tau \chi^{(3)} |\beta|^2 n_S \\ \delta_{N_S} &\approx \delta_0 := \sigma / G|_{n_S=0} = \frac{1}{\tau \chi^{(3)} |\beta|} \end{aligned} \right\} \quad (11)$$

and hence the non-ideality function  $\lambda$  of (4) satisfies

$$\lambda(x|n) \approx (2\pi)^{-1/2} \sigma^{-1} \exp\left(-\frac{1}{2\sigma^2} (Gn - x)^2\right). \quad (12)$$

For the characterization of a non-destructive measurement device, such as this one, measurement inaccuracy is not the only parameter. In fact, four 'correlations' may be distinguished [5]:

- (i)  $\hat{X}_{\text{meas}} \leftrightarrow \hat{N}_S$
- (ii)  $\hat{X}_{\text{meas}} \leftrightarrow \hat{N}_{S'}$
- (iii)  $\hat{N}_S \leftrightarrow \hat{N}_{S'}$
- (iv)  $\hat{\varphi}_S \leftrightarrow \hat{\varphi}_{S'}$

the  $S'$  observables to be measured in the outgoing signal field  $S'$  (cf figure 1);  $\hat{\varphi}$  is a phase observable [9]. Firstly, it is of course important how good a measure of  $\hat{N}_S$  the read-out observable  $\hat{X}_{\text{meas}}$  is. Precisely this aspect has just been quantified by means of the device's measurement inaccuracy  $\delta_{N_S}$  (or  $\delta_0$ ). Secondly, in a non-destructive device it is possible to prepare an output signal state with an approximately known number of photons, conditional on the measurement outcome. Case (iii) indicates the losses in the device. The last correlation involves input and output signal phase. Its quality shows to what extent phase, the observable conjugate to photon number, is *disturbed* by the instrument. The phase disturbance is complementary to the measurement inaccuracy (i) [10]. Because this fourth aspect, and its complementary relationship with (i), has been treated in [9], and because the third is a straightforward function of the losses in the medium, we will focus in the present paper on the first two. In

addition to  $\delta_0$ , we therefore need a characterization of the quality of the correlation between  $\hat{X}_{\text{meas}}$  and  $\hat{N}_S$ . First define the correlation coefficient,

$$C(\hat{X}_{\text{meas}}, \hat{N}_S) := \frac{\langle \hat{X}_{\text{meas}} \hat{N}_S \rangle - \langle \hat{X}_{\text{meas}} \rangle \langle \hat{N}_S \rangle}{\langle \Delta^2 \hat{X}_{\text{meas}} \rangle^{1/2} \langle \Delta^2 \hat{N}_S \rangle^{1/2}}. \quad (13)$$

Here  $\langle \Delta^2 \hat{X}_{\text{meas}} \rangle := \langle (\hat{X}_{\text{meas}} - \langle \hat{X}_{\text{meas}} \rangle)^2 \rangle$  denotes the unconditional read-out variance. Now consider the  $\hat{N}_S$  distribution. As losses have been neglected so far,  $\hat{N}_S = \hat{N}_S$ . Hence the joint probability distribution of  $\hat{X}_{\text{meas}}$  and  $\hat{N}_S$  satisfies

$$\text{prob}_{\hat{X}_{\text{meas}}, \hat{N}_S}(x, n) = \lambda(x|n) \text{prob}_{\hat{N}_S}(n) \quad (14)$$

the function  $\lambda(x|n)$  being given, in the linear regime, by (12). If we use the fact that both noise and gain in (12) are independent of the signal photon number, the correlation coefficient can be calculated to be

$$\begin{aligned} C(\hat{X}_{\text{meas}}, \hat{N}_S) &= \frac{G \langle \Delta^2 \hat{N}_S \rangle}{(G^2 \langle \Delta^2 \hat{N}_S \rangle + \sigma^2)^{1/2} \langle \Delta^2 \hat{N}_S \rangle^{1/2}} \\ &= (\sigma^2 G^{-2} \langle \Delta^2 \hat{N}_S \rangle^{-1} + 1)^{-1/2}. \end{aligned} \quad (15)$$

Note that the joint distribution (14) can be seen as the unconditional  $\hat{N}_S$  distribution, multiplied by a narrow function of  $n$ . Thus, for the quality with which the  $\hat{X}_{\text{meas}}$  result determines the outgoing signal photon number, it is the  $n$ -width  $\varepsilon$  of this multiplying function which is important. As this function approximately equals the Gaussian (12), clearly  $\varepsilon \approx \delta_0$ . Then (15) suggests the quantity

$$\varepsilon^2 := \left( \frac{1}{C^2(\hat{X}_{\text{meas}}, \hat{N}_S)} - 1 \right) \langle \Delta^2 \hat{N}_S \rangle \quad (16)$$

as a suitable 'preparation inaccuracy' measure, representing the inaccuracy in an estimation of  $n_S$  from the  $\hat{X}_{\text{meas}}$  data obtained in this idealized set-up.

#### 4. Self-phase modulation

In both the homodyne and the  $\gamma = \frac{1}{2}$  schemes, however, the situation worsens as soon as the two  $\hat{N}^2$  terms in (1) are included. These become important in a non-resonant Kerr medium as probe and/or signal strength increase. Including the two terms, we get for the probe beam at time  $\tau$ :

$$\hat{a}_P = \hat{a}_P \exp[-i\tau\chi^{(3)}(\hat{N}_S + \frac{1}{2}\hat{N}_P - \frac{1}{4})]. \quad (17)$$

The  $\hat{X}_{\text{meas}}$  distribution will still be of the same form as (4). But the  $\hat{N}^2$  terms in  $\hat{H}_I$  cause a broadening of the probe states in the phase direction ('crescent squeezing') [1]. This effect is known as *self-phase modulation* (SPM). As it is precisely the phase of the probe beam that contains the  $\hat{N}_S$  information, the SPM spoils the quality with which this information can be detected. It increases the device's noise. Without the self-phase modulation effect it would follow from (10) that the inaccuracy can be reduced indefinitely by increasing  $|\beta|$ . An analysis including SPM [9], however, shows that this is not the case because the non-ideality function of the measurement changes into a convolution of (12) with a distribution function representing the extra noise, thus

increasing the quantity  $\delta_0$  from (11) according to

$$\delta_0 \approx [(\tau\chi^{(3)}|\beta|)^{-2} + \frac{1}{8}|\beta|^2]^{1/2} \geq (\tau\chi^{(3)})^{-1/2} 2^{-1/4}. \quad (18)$$

Thus, inclusion of self-phase modulation causes the measurement inaccuracy to have a non-zero lower bound. The preparation inaccuracy  $\varepsilon$  is similarly affected.

In the literature several proposals can be found for dealing with the effects of self-phase modulation by a clever choice of the reference beam phase  $\varphi$ . Thus, in [13] the effect of the shift in the pump wave is taken into account by a suitable interaction picture. In [1] it is demonstrated that, for the homodyne receiver (where the probe strength is given by  $\beta = \sqrt{\gamma\beta} \ll \beta$ ), the choice

$$\varphi = \cot^{-1}(\tau\chi^{(3)}|\beta|^2) \quad (19)$$

of the reference beam phase can cancel the contributions of SPM to the measured observable. In [9], however, it was shown numerically that this only holds up to a certain intensity of the probe beam, above which the effects of crescent squeezing on the initially coherent reference beam can no longer be neglected. From the numerical calculation it follows that, although for each value of  $|\beta|^2$  the measurement inaccuracy can be decreased by an optimal choice of  $\varphi$  (see also [3]), it is in general not possible to counter the effects of SPM completely. For large probe intensities  $|\beta|^2$  the optimal value of  $\varphi$  differs from (19) (see figure 2 of [9]). We found it possible, by choosing the optimal value of  $\varphi$ , to improve the lower bound (18) to

$$\delta_0 \geq \mathcal{O}[(\tau\chi^{(3)})^{-2/5}]. \quad (20)$$

In [2] two methods are proposed in order to avoid the self-modulation effect, namely (i) using a resonant  $\chi^{(3)}$  medium at a frequency  $\omega = \omega_s + \omega_p$ , (ii) cancelling the effect by means of a negative  $\chi^{(3)}$  medium. Here we propose an alternative scheme having the same effect. This scheme is based on the set-up of figure 2, which was first considered by Imoto and Saito [6] in their discussion of losses. Again  $\gamma = \frac{1}{2}$ . Three identical Kerr media are used. Then we can derive

$$\begin{aligned} \hat{a}_p &= \exp(-i\tau\chi^{(3)}\hat{N}_L - \frac{1}{2}i\tau\chi^{(3)}\hat{N}_p)\hat{a}_p \\ &= \exp(-i\tau\chi^{(3)}\hat{N}_L - i\tau\chi^{(3)}\hat{N}_s - i\tau\chi^{(3)}\hat{N}_p)\hat{a}_p \\ \hat{a}_L &= \exp(-i\tau\chi^{(3)}\hat{N}_p - \frac{1}{2}i\tau\chi^{(3)}\hat{N}_L + i\varphi)\hat{a}_L \\ &= \exp(-i\tau\chi^{(3)}\hat{N}_p - i\tau\chi^{(3)}\hat{N}_L + i\varphi)\hat{a}_L. \end{aligned} \quad (21)$$

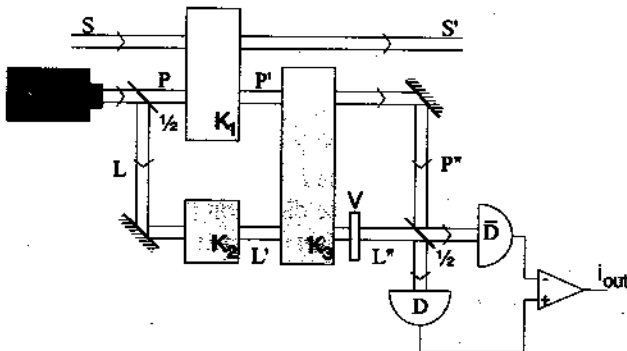


Figure 2. Set-up with compensation. The set-up of figure 1, with  $\gamma = \frac{1}{2}$ , is supplemented by two more Kerr media, identical to the first. These compensate for the effects of self-phase modulation on the probe beam P.

Substituting these into (2), we see that the SPM disappears out of  $\hat{X}_{\text{out}}$  altogether, yielding (3) once again. The SPM is completely compensated for†.

### 5. Losses and inaccuracy

So far, however, we have neglected the effect of losses in the media. As these are never completely transparent, losses may be expected to be the next performance limiting factor [6]. In a single mode optical field, the effects of losses may be described in the Schrödinger–Markov picture [15] by a quantum stochastic master equation ( $\hbar = 1$ ),

$$\frac{\partial \hat{\rho}}{\partial t} = -i[\hat{H}, \hat{\rho}]_- + \zeta(\hat{a}\hat{\rho}\hat{a}^\dagger - \frac{1}{2}[\hat{N}, \hat{\rho}]_+). \quad (22)$$

Equivalently, in the Heisenberg–Markov picture, we may use

$$\frac{\partial \langle \hat{A} \rangle_{\text{R}}}{\partial t} = \langle i[\hat{H}, \hat{A}]_- + \zeta(\hat{a}^\dagger \hat{A} \hat{a} - \frac{1}{2}[\hat{N}, \hat{A}]_+) \rangle_{\text{R}}. \quad (23)$$

Note that in (22) pure states do not in general remain pure. If  $\hat{H} = \omega \hat{a}^\dagger \hat{a}$ , coherent states form the only exception. In (23) the notation  $\langle \hat{A} \rangle_{\text{R}}$  indicates that the equation's derivation involves reservoir averaging [15] so that  $\langle \hat{A} \hat{B} \rangle_{\text{R}} \neq \langle \hat{A} \rangle_{\text{R}} \langle \hat{B} \rangle_{\text{R}}$  in general. Combining (23) with (1), we get the evolution equation for two modes with Kerr interaction and losses‡

$$\frac{\partial \langle \hat{A} \rangle_{\text{R}}}{\partial t} = \left\langle i[\hat{H}_i, \hat{A}]_- + \zeta \sum_{i=P,S} (\hat{a}_i^\dagger \hat{A} \hat{a}_i - \frac{1}{2}[\hat{N}_i, \hat{A}]_+) \right\rangle_{\text{R}}. \quad (24)$$

Equation (24) can be explicitly solved to give (see appendix)

$$\langle (\hat{a}_P^{i m_1} \hat{a}_P^{i n_1} \hat{a}_S^{i m_2} \hat{a}_S^{i n_2})(t) \rangle_{\text{R}} = \hat{a}_P^{i m_1} \hat{a}_S^{i m_2} \exp[C(t) + D(t)\hat{N}_P + E(t)\hat{N}_S] \hat{a}_P^{i n_1} \hat{a}_S^{i n_2} \quad (25)$$

with

$$\begin{aligned} C(t) &= -\frac{1}{2}(m_1 + n_1 + m_2 + n_2)t\zeta \\ &\quad + i[m_1 m_2 - n_1 n_2 + \frac{1}{4}(m_1^2 - n_1^2) + \frac{1}{4}(m_2^2 - n_2^2)]t\chi^{(3)} \\ D(t) &= \Xi^{(m_2 - n_2 + 1/2 m_1 - 1/2 n_1)}(t) \quad E(t) = \Xi^{(m_1 - n_1 + 1/2 m_2 - 1/2 n_2)}(t) \end{aligned}$$

and

$$\Xi^{(m)}(t) := \log \left( \frac{\zeta - i m \chi^{(3)} \exp(i m t \chi^{(3)} - \zeta t)}{\zeta - i m \chi^{(3)}} \right). \quad (26)$$

We may approximate  $\Xi$  for small  $t\chi^{(3)}$  by

$$\Xi^{(m)}(t) = \frac{1 - \exp(-t\zeta)}{\zeta} i m \chi^{(3)} - \frac{1 - 2t\zeta \exp(-t\zeta) - \exp(-2t\zeta)}{2\zeta^2} m^2 \chi^{(3)2} \quad (27)$$

† The compensation goes a little differently if the three  $\chi$  are not equal. Let  $\chi_{LP}$ ,  $\chi_{LL}$  and  $\chi_{PP}$  denote the  $\chi$  that L and P beams experience in the Kerr medium. Then, since L and P beams have the same frequency,  $\chi_{LL} = \chi_{PP}$ . The interaction time  $\tau_3$  for  $K_3$  must then be taken  $\tau_3 = \tau_{\chi_{LL}} / (2\chi_{LP} - \chi_{LL})$  to effect a complete compensation.

‡ Note that the sub-unity quantum efficiency  $\eta$  of the detectors D and  $\hat{D}$  is neglected. When included, this simply leads to an increase of  $\delta_{N_S}$  by a factor  $\eta^{-1/2}$  (see e.g. [14]). We have also neglected that the  $\zeta$  may be different for the different modes. If we did include this, the reasoning would remain completely analogous.

for small  $\zeta$  by

$$\Xi^{(m)}(t) \approx imt\chi^{(3)}\exp(-\frac{1}{2}t\zeta) - \frac{1}{6}m^2t^3\zeta\chi^{(3)2}\exp(-t\zeta). \quad (28)$$

Thus the losses lead, amongst other effects, to an attenuation of the correlation between  $\hat{N}_S$  and the probe phase, and hence to an attenuation of the correlation between  $\hat{N}_S$  and  $\hat{X}_{\text{meas}}$ . Nevertheless, even with losses, we have a non-ideal photon meter in the sense of (4).

For the set-up of figure 2, we need to employ (25) three times, once for each Kerr medium. This leads to

$$\begin{aligned} \langle \hat{a}_P^\dagger \hat{a}_L \rangle_R &= \langle (\hat{a}_P^\dagger \hat{a}_L)(\tau) \exp(i\varphi) \rangle_R \\ &= \langle \hat{a}_P^\dagger \exp[-\tau\zeta + D(\tau)\hat{N}_P + E(\tau)\hat{N}_L + i\varphi] \hat{a}_L \rangle_R \\ &= a_P^\dagger \exp[-2\tau\zeta + \bar{D}(\tau; \tau)\hat{N}_P + \bar{E}(\tau; \tau)\hat{N}_L + \bar{F}(\tau)\hat{N}_S + i\varphi] \hat{a}_L \end{aligned} \quad (29)$$

with

$$\begin{aligned} \bar{D}(t_1; t_2) &:= \Xi^{(-1/2, 1/2)}(t_1; t_2) & \bar{E}(t_1; t_2) &:= \Xi^{(1/2, -1/2)}(t_1; t_2) \\ \bar{F}(t_2) &:= \Xi^{(1)}(t_2) \end{aligned}$$

and

$$\Xi^{(m; n)}(t_1; t_2) := \Xi^{(n)}(t_2 + t_0). \quad (30)$$

The parameter  $t_0$  is defined by

$$\Xi^{(n)}(t_0) := \Xi^{(m)}(t_1).$$

Note that  $\bar{E}(t_1; t_2) = \bar{D}^*(t_1; t_2)$ . Hence ( $\varphi = \pi/2$ )

$$\langle \hat{a}_P^\dagger \hat{a}_L \rangle_{P+L+R, n_S} = -\frac{1}{2}|\beta|^2 \exp(-2\tau\zeta + \text{Re}\{\exp[\bar{D}(\tau; \tau)] - 1\})|\beta|^2 + \bar{F}(\tau)n_S. \quad (31)$$

For small  $t\chi^{(3)}$ , the approximation

$$\begin{aligned} \exp(\Xi^{(m; -m)}(t; t)) - 1 &\approx \frac{[1 - \exp(-t\zeta)]^2}{\zeta} im\chi^{(3)} \\ &\quad - \frac{1 - 2t\zeta \exp(-t\zeta) - \exp(-2t\zeta)}{\zeta^2} m^2\chi^{(3)2} \end{aligned} \quad (32)$$

is useful. For small  $\zeta$  this further simplifies into

$$\exp(\Xi^{(m; -m)}(t; t)) - 1 \approx imt^2\zeta\chi^{(3)}\exp(-t\zeta) - \frac{1}{6}m^2t^3\zeta\chi^{(3)2}\exp(-\frac{3}{2}t\zeta). \quad (33)$$

With the aid of (29)–(31),  $\mu$  and  $G$  can be determined (see appendix). An analogous calculation yields  $\sigma$ . From these  $\delta_0$  follows, and is plotted in figure 3. For small  $t\chi^{(3)}$ , using (27) and (32)

$$\delta_0 \approx \frac{\zeta \exp(-\frac{1}{2}\tau\zeta)}{1 - \exp(-\tau\zeta)} \left( (\chi^{(3)}|\beta|)^{-2} \exp(3\tau\zeta) + |\beta|^2 \frac{\sinh(\tau\zeta) - \tau\zeta}{\zeta^2} \right)^{1/2} \quad (34)$$

$$\begin{aligned} &\approx [(\tau\chi^{(3)}|\beta|)^{-2} \exp(3\tau\zeta) + \frac{1}{6}\tau\zeta|\beta|^2]^{1/2} \\ &\geq \exp(\frac{3}{4}\tau\zeta) (\frac{2}{3}\zeta/\chi^{(3)})^{1/4} (\tau\chi^{(3)})^{-1/4}. \end{aligned} \quad (35)$$

The latter approximation holds for small  $\zeta$ , employing (33) and (28). As the set-up of figure 2 contains three media, compared to only one in figure 1, it is clear that the



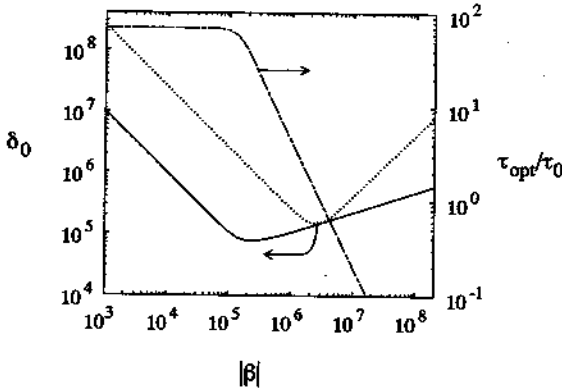


Figure 3. Performance of  $\gamma = \frac{1}{2}$  set-up with compensation (figure 2), for fixed length  $\tau = \tau_0$  (dotted curve) and for optimally chosen  $\tau = \tau_{\text{opt}}$  (full curve). The performance is calculated from (29)–(31). It is plotted against the local laser strength  $|\beta|$ .  $\tau_{\text{opt}}/\tau_0$  is also indicated (chain curve, right y-axis). ( $\tau_0\zeta = 10^{-2}$ ;  $\tau_0\chi^{(3)} = 4 \times 10^{-12}$ ).

effects of the attenuation of probe and reference beams is larger in the set-up with compensation. Moreover, (34) shows that losses reduce the effectiveness of the compensation. Nevertheless, comparing (35) to (18), we see that the minimum inaccuracy has been reduced roughly by a factor  $\mathcal{O}[(\tau\zeta)^{1/4}]$ . Thus, for low  $\tau\zeta$ , the set-up with compensation is still preferable to the set-up without.

What we have not looked at so far is the interaction length  $\tau$ . This parameter may be varied, for given  $\zeta$  and  $\chi^{(3)}$ . Initially accuracy will increase with  $\tau$ , but as losses become more important, attenuating both probe and signal beams, the quality will eventually decrease with increasing  $\tau$ . Hence there is an optimal value  $\tau_{\text{opt}}$  for  $\tau$ . For small  $|\beta|$ , losses dominate and the second term (SPM) in the square root factor of (34) may be neglected, so that

$$\tau_{\text{opt}} = (1/\zeta)\log(2). \tag{36}$$

For large  $|\beta|$ , the influence of SPM dominates, and we may use (35), including  $\zeta$  only to lowest order, to get

$$\tau_{\text{opt}} \approx \left(\frac{1}{12}|\beta|^4\zeta\chi^{(3)^2}\right)^{-1/3}. \tag{37}$$

Combining the  $\tau_{\text{opt}}$  estimates for the two regimes with the expressions for the inaccuracy, we find

$$\delta_0 \approx 4\zeta(\chi^{(3)}|\beta|)^{-1} \tag{38}$$

$$\delta_0 \approx \left(\frac{1}{12}\zeta|\beta|/\chi^{(3)}\right)^{1/3}\sqrt{3} \tag{39}$$

for small and large  $|\beta|$ , respectively. Equating (38) and (39), we obtain the estimate

$$\min(\delta_0) \approx \mathcal{O}((\zeta/\chi^{(3)})^{1/2}) \tag{40}$$

for the best achievable inaccuracy. Both the optimum (40) and the two regimes are clearly distinguishable in figure 3.

## 6. Losses and output

As noted earlier, for larger values of  $\tau\zeta$ , the set-up of figure 2 is no better than that of figure 1. For low  $|\beta|$ , however,  $\tau_{\text{opt}}\zeta$  is large. Indeed an evaluation of  $\min(\delta_0)$  for figure

1 shows that it is slightly lower than (40), though of the same order. But because  $\hat{N}_s = \hat{N}_s \exp(-\tau\zeta)$ , an optimization of  $\tau$  leads to large signal losses. Then most of the advantages of using a QND set-up are lost. In other words, in most cases of interest we will prefer a situation with low  $\tau\zeta$ , rather than one with  $\tau = \tau_{\text{opt}}$ . Then the compensated scheme works best.

The correlation between  $\hat{N}_s$  and  $\hat{X}_{\text{out}}$  can be evaluated analogously to the measurement inaccuracy for the set-up of figure 2. This gives for small  $\tau\chi^{(3)}$ , using definition (16),

$$\begin{aligned} \varepsilon^2 &\approx (\tau\chi^{(3)}|\beta|)^{-2} \exp(2\tau\zeta) + \frac{1 - 2\tau\zeta \exp(-\tau\zeta) - \exp(-2\tau\zeta)}{2(\tau\zeta)^2} |\beta|^2 \\ &\quad + \left( 2 \frac{1 - \tau\zeta \exp(-\tau\zeta) - \exp(-\tau\zeta)}{(\tau\zeta)^2} - \exp(-\tau\zeta) \right) |\alpha|^2 \\ &= \left( \delta_0 \frac{1 - \exp(-\tau\zeta)}{\tau\zeta} \right)^2 \\ &\quad + \left( 2 \frac{1 - \tau\zeta \exp(-\tau\zeta) - \exp(-\tau\zeta)}{(\tau\zeta)^2} - \exp(-\tau\zeta) \right) |\alpha|^2 \\ &\approx \delta_0^2 \exp(-\tau\zeta) + \tau\zeta \exp(-\frac{3}{4}\tau\zeta) |\alpha|^2. \end{aligned} \quad (41)$$

The latter approximation is valid for small  $\tau\zeta$ . Hence also in the lossy case, the preparative quality is linked directly to the measurement inaccuracy, with deviations only for larger signal levels,  $|\alpha| \geq \mathcal{O}[|\beta|]$ .

## 7. Discussion and conclusions

The compensation method of figure 2 provides a perfect compensation for SPM, if losses are neglected. But if they are not, the intensities in the compensating medium  $K_3$  are different from those in the other two media. This has the consequence that the compensation becomes imperfect, as is evidenced by (34). If the set-up could be arranged so that compensation and P-S interaction were to occur simultaneously, this effect would be countered. A crude way of approximating this goal is sketched in figure 4. We have replaced the three media of length  $\tau$  by  $n$  triples of length  $\tau/n$ . In this

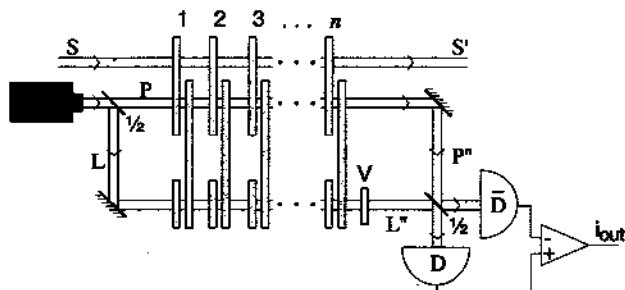


Figure 4. Set-up with distributed compensation ( $\gamma = \frac{1}{2}$ ). Instead of one triple of Kerr media, as in figure 2, we use  $n$  triples. This counters the adverse effects of losses on the compensation of the self-phase modulation.

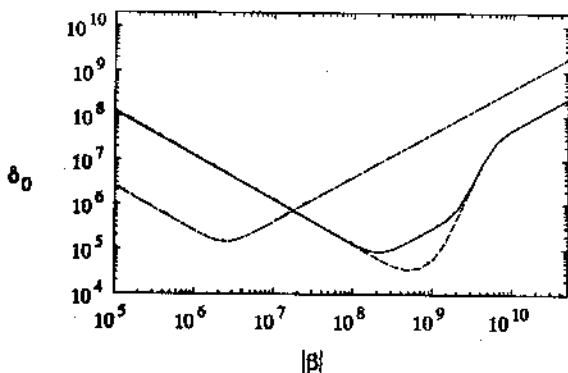


Figure 5. Performance of the  $\gamma = \frac{1}{2}$  scheme with the compensation of figure 2 (chain curve) and of a homodyne set-up ( $\gamma \ll 1$ ), plotted as a function of the local laser strength  $|\beta|$ . Indicated are homodyning variants using an optimal choice for  $\varphi$ , both with (full curve) and without (broken curve) losses. Note that for given  $\beta$  the probe field strength is different for homodyne and  $\gamma = \frac{1}{2}$  schemes, due to the difference in  $\gamma$ . ( $\tau\zeta = 10^{-2}$ ;  $\tau\chi^{(3)} = 4 \times 10^{-12}$ ; homodyne:  $\gamma = 10^{-4}$ .)

way the P-S interaction remains of the same strength, whereas the SPM term in (34), i.e. the second term in the square root factor, is reduced by a factor  $n^{-2}$ . As  $n \rightarrow \infty$ , this term disappears, and we get a system effectively described by an interaction Hamiltonian

$$\hat{H}_I = \chi^{(3)} \hat{N}_S \hat{N}_P + \chi^{(3)} \hat{N}_L \hat{N}_P + \frac{1}{2} \chi^{(3)} \hat{N}_P^2 + \frac{1}{2} \chi^{(3)} \hat{N}_L^2 + \frac{1}{4} \chi^{(3)} \hat{N}_S^2. \quad (42)$$

As can be straightforwardly checked, this Hamiltonian leads to a complete compensation of SPM, even in the presence of losses.

Comparing the compensation scheme of figure 2 to the homodyne scheme, where we counter SPM by a suitable choice of reference beam phase (cf (19)), we see that for typical values for  $\tau\chi^{(3)}$  and  $\tau\zeta$  they perform comparably (figure 5). For decreasing  $\tau\zeta$ , however, the quality of the homodyne scheme, unlike that of the compensation scheme, is limited. Thus for small enough  $\tau\zeta$  the compensation scheme will be better.

### Acknowledgment

One of the authors (HM) acknowledges the support of the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

### Appendix

In this appendix, we shall explicitly work out the evolution of a simple operator, for a lossy medium without SPM. The full derivation of (25) proceeds completely analogously. Next the compensated scheme of figure 2 is considered. Noise and gain are worked out for this configuration, leading to the expressions for inaccuracy presented in the text.

A solution to the Schrödinger equation with Hamiltonian (1), without SPM, is given by

$$\hat{a}_P(t) = \exp(-it\chi^{(3)}\hat{N}_S)\hat{a}_P. \quad (\text{A1})$$

Try therefore

$$\hat{a}_P(t) = \exp(f(t)\hat{N}_S + g(t))\hat{a}_P \quad (\text{A2})$$

as a solution for (24) without SPM. Substituting (A2) into (24), we get

$$\begin{aligned} & \left\langle \left( \frac{\partial f}{\partial t} N_S + \frac{\partial g}{\partial t} \right) \exp[f(t)\hat{N}_S + g(t)] \hat{a}_P \right\rangle_R \\ &= \langle -i\chi^{(3)}\hat{N}_S \exp(f(t)\hat{N}_S + g(t))\hat{a}_P \\ & \quad + \zeta \exp(f(t)\hat{N}_S + g(t))(\hat{a}_P^\dagger \hat{a}_P^2 - \frac{1}{2}[\hat{N}_P, \hat{a}_P]_+) \\ & \quad + \zeta \exp(g(t))\{\hat{a}_S^\dagger \exp(f(t)\hat{N}_S)\hat{a}_S - \frac{1}{2}[\hat{N}_S, \exp(f(t)\hat{N}_S)]_+\} \hat{a}_P \rangle_R \\ &= \langle (-i\chi^{(3)}\hat{N}_S - \frac{1}{2}\zeta + \zeta\hat{N}_S[\exp(-f(t)) - 1])\exp(f(t)\hat{N}_S + g(t))\hat{a}_P \rangle_R. \end{aligned}$$

Hence

$$\frac{\partial f}{\partial t} = -i\chi^{(3)} + \zeta[\exp(-f(t)) - 1] \quad \frac{\partial g}{\partial t} = -\frac{1}{2}\zeta \Rightarrow g(t) = g(0) - \frac{1}{2}\zeta t. \quad (\text{A3})$$

It can be verified that the solution to (A3) is given by  $f(t) = \Xi^{(-1)}(t + t_0)$ , for some fixed  $t_0$ . As  $f(0) = g(0) = 0$ , it follows that  $t_0 = 0$ . If we include SPM, and solve (24) for more complicated operators, equation (25) follows analogously to the above.

In the compensation scheme of figure 2, we first solve (24) for  $K_3$ , leading to a result analogous to (25). This is then used as an initial condition for the  $K_1/K_2$  equation. From this procedure we get (29). Now, using (31),  $\varphi = \pi/2$  and  $\gamma = \frac{1}{2}$ ,

$$\begin{aligned} & \langle \hat{X}_{\text{meas}} \rangle_{P+L+R, n_S} \\ &= -2\text{Im}(\langle \frac{1}{2}\beta\sqrt{2} | \otimes_L \langle \frac{1}{2}\beta\sqrt{2} | \otimes_S \langle n_S | \langle \hat{a}_P^\dagger \hat{a}_L \rangle_R | n_S \rangle_S \otimes | \frac{1}{2}\beta\sqrt{2} \rangle_L \otimes | \frac{1}{2}\beta\sqrt{2} \rangle_P) \\ &= |\beta|^2 \text{Im}(\exp[-2\tau\zeta + \text{Re}\{\exp(\bar{D}(\tau; \tau)) - 1\}]|\beta|^2 + \bar{F}(\tau)n_S). \end{aligned} \quad (\text{A4})$$

Thus

$$G|_{n_S=0} = |\beta|^2 \text{Im}[\bar{F}(\tau)] \exp\{-2\tau\zeta + \text{Re}\{\exp(\bar{D}(\tau; \tau)) - 1\}\}|\beta|^2. \quad (\text{A5})$$

Furthermore,

$$\langle \hat{X}_{\text{meas}}^2 \rangle_{P+L+R, n_S} = \langle \hat{N}_P + \hat{N}_L + 2\hat{N}_P\hat{N}_L - \hat{a}_P^{\dagger 2}\hat{a}_L^2 - \hat{a}_P^2\hat{a}_L^{\dagger 2} \rangle_{P+L+R, n_S}. \quad (\text{A6})$$

We apply the analogue of (29) for  $\langle \hat{a}_P^{\dagger 2}\hat{a}_L^2 \rangle_R$  to (A6), to get

$$\begin{aligned} \langle \hat{X}_{\text{meas}}^2 \rangle_{P+L+R, n_S=0} &= |\beta|^2 \exp(-2\zeta\tau) + \frac{1}{2}|\beta|^4 \exp(-4\zeta\tau) \\ & \quad - \frac{1}{2}|\beta|^4 \text{Re}(\exp\{-4\tau\zeta + \text{Re}\{\exp(\bar{D}_2(\tau; \tau)) - 1\}\}|\beta|^2) \end{aligned}$$

where

$$\bar{D}_2(t_1; t_2) = \Xi^{(-1; 1)}(t_1; t_2) \quad \bar{F}_2(t_2) = \Xi^{(2)}(t_2).$$

This gives

$$\begin{aligned} \sigma|_{n_S=0} &= \langle \Delta^2 X_{\text{meas}} \rangle_{P+L+R, n_S=0} \\ &= |\beta|^2 \exp(-2\zeta\tau) \\ &\quad + \frac{1}{2} |\beta|^4 \exp(-4\tau\zeta) (1 - \exp[\text{Re}[\exp(\bar{D}_\chi(\tau; \tau)) - 1] |\beta|^2]). \end{aligned} \quad (\text{A7})$$

If we combine (A7) with (A5), we get  $\delta_0$ . Using approximations (27) and (32) shows us that for small  $\tau\chi^{(3)}$

$$\begin{aligned} G|_{n_S=0} &\approx \frac{1 - \exp(-\tau\zeta)}{\zeta} |\beta|^2 \chi^{(3)} \exp(-2\tau\zeta) \\ \sigma|_{n_S=0} &\approx |\beta|^2 \exp(-2\zeta\tau) \\ &\quad + \frac{1}{2} |\beta|^6 \exp(-4\tau\zeta) \frac{1 - 2\tau\zeta \exp(-\tau\zeta) - \exp(-2\tau\zeta)}{2\zeta^2} \chi^{(3)^2}. \end{aligned} \quad (\text{A8})$$

From these last two equations, equation (34) follows. We calculate  $\langle \hat{X}_{\text{meas}} \hat{N}_S \rangle_{S+P+L+R}$  in an analogous way, from which  $C(\hat{X}_{\text{meas}}, \hat{N}_S)$  follows. This gives (41).

If we consider the derivation of (A8) once again, we see that the real part of  $\exp(\Xi) - 1$  is important. As this is, in lowest order, proportional to  $\tau^3$ , it reduces by a factor  $n^{-3}$  if  $\tau$  is reduced by a factor  $1/n$ . Hence, combining the effect of  $n$  triples, as in figure 4,

$$\text{Re}[\exp(\Xi) - 1] \propto n(\tau\chi^{(3)}/n)^3 = n^{-2}(\tau\chi^{(3)})^3. \quad (\text{A9})$$

Similarly, the imaginary part of  $\exp(\Xi) - 1$  can be seen to be reduced by a factor  $1/n$ . But  $\hat{F}$  remains the same. Hence the effect of the set-up of figure 4 is a reduction of the second term in the square root factor of (34) by a factor  $n^{-2}$ . Moreover, if we let  $n \rightarrow \infty$ , the analogue of (29) for the set-up of figure 4 can be seen to yield an evolution compatible with (42):  $\bar{D}$  and  $\bar{E}$  disappear. For more general expressions, involving other powers of  $\hat{a}_P$  and  $\hat{a}_L$ , the same can be seen to hold. Hence we get a system effectively described by (42) and (24).

Note that in figures 3 and 5 the exact expressions are used, without using the  $\tau\chi^{(3)} \ll 1$  or  $\tau\zeta \ll 1$  approximations.

## References

- [1] Yamamoto Y, Machida S, Saito S, Imoto N, Yanagawa T, Kitagawa M and Björk G 1990 *Progress in Optics* vol 28 ed E Wolf (Amsterdam: North-Holland) p 87
- [2] Imoto N, Haus H and Yamamoto Y 1985 *Phys. Rev. A* **32** 2287
- [3] Shelby R, Levenson M, Perlmutter S, DeVoe R and Walls D 1986 *Phys. Rev. Lett.* **57** 691
- [4] Imoto N, Watkins S and Sasaki Y 1987 *Opt. Commun.* **61** 159
- [5] Bachor H-A, Levenson M, Walls S, Perlmutter S and Shelby R 1988 *Phys. Rev. A* **38** 180
- [6] Imoto N and Saito S 1989 *Phys. Rev. A* **39** 675
- [7] Holland M, Collett M, Walls D and Levenson M 1990 *Phys. Rev. A* **42** 2995
- [8] Dance M, Collett M and Walls D 1991 *Phys. Rev. Lett.* **66** 1115
- [9] Martens H and de Muynck W 1991 *Quantum Aspects of Optical Communication (Lecture Notes in Physics 378)* ed C Bendjaballah, O Hirota and S Reynaud (Berlin: Springer) p 41
- [10] Martens H 1991 The uncertainty principle, PhD Thesis, Eindhoven University of Technology unpublished
- [11] Martens H and de Muynck W 1990 *Found. Phys.* **20** 255,357 (also cf other references therein)

- [12] Davies E 1976 *Quantum Theory of Open Systems* (New York: Academic)
- [13] Milburn G, Levenson M, Shelby R, Perlmutter S, De Voe R and Walls D 1987 *J. Opt. Soc. Am. B* **4** 1476
- [14] de Muynck W and Martens H 1990 *Proc. 3rd Int. Symp. Foundations of Quantum Mechanics in the Light of New Technology* ed S Kobayashi, H Ezawa, Y Murayama and S Namura (Tokyo: Physical Society of Japan) p 171
- [15] Louisell W 1973 *Quantum Statistical Properties of Radiation* (New York: Wiley)