J. Opt. B: Quantum Semiclass. Opt. **3** (2001) 337–340 PII: S1464-4266(01)26468-4

Optimal quantum estimation of the coupling between two bosonic modes

G Mauro D'Ariano[1], Matteo G A Paris[2] and Paolo Perinotti[3]

[1] Sezione INFN, Universitá di Pavia, via Bassi 6, I-27100 Pavia, Italy
[2] Unitá INFM, Universitá di Pavia, via Bassi 6, I-27100 Pavia, Italy
[3] Sezione INFN, Universitá di Milano, via Celoria 16, I-20133 Milan, Italy

Received 3 July 2001, in final form 24 September 2001
Published 26 September 2001
Online at stacks.iop.org/JOptB/3/337

Abstract
We address the problem of the optimal quantum estimation of the coupling parameter of a bilinear interaction, such as the transmittivity of a beamsplitter or the internal phase-shift of an interferometer. The optimal measurement scheme confirms Heisenberg scaling of precision versus the total energy as an unsurpassable bound, but with a largely reduced multiplicative constant.

Keywords: Quantum interferometry, quantum phase, quantum estimation theory

1. Introduction

How effectively may we estimate the strength of a simple interaction such as the transmittivity of a beamsplitter or the phase-shift imposed in the internal arms of an interferometer? The Hamiltonian describing the bilinear coupling between two bosonic modes has the form

$$H = \kappa(a^\dagger b + b^\dagger a), \tag{1}$$

where κ depends on the specific interaction under consideration. By using the Schwinger representation of the $SU(2)$ Lie algebra, with generators

$$J_x = \frac{1}{2}(a^\dagger b + b^\dagger a),$$
$$J_y = \frac{1}{2i}(a^\dagger b - b^\dagger a),$$
$$J_z = \frac{1}{2}(a^\dagger a - b^\dagger b), \tag{2}$$

we can rewrite the Hamiltonian as $H = 2\kappa J_x$ and the evolution operator as

$$U_\psi = \exp(-\mathrm{i} J_x \psi), \tag{3}$$

where the global coupling constant ψ is equal to $2\kappa\Delta t$, Δt being the effective interaction time. The evolution in equation (3) describes, for example, the interaction of two light modes in a beamsplitter with transmittivity $\tau = \cos^2\psi$ [1], or, apart from a fixed rotation, the evolution of the arm modes in a Mach-Zehnder interferometer, with ψ representing the phase-shift between arms [2, 3]. If the initial preparation of the two modes is described by the density matrix ρ_0 the evolved state in the interaction picture is given by

$$\rho_\psi = \exp(-\mathrm{i} J_x \psi)\rho_0 \exp(\mathrm{i} J_x \psi). \tag{4}$$

In this paper we devote our attention to the estimation of ψ through measurements performed on ρ_ψ. We denote the generic probability operator-valued measure (POVM) for the estimation process by $\mathrm{d}\nu(\phi)$, so that the results of the measurement are distributed according to

$$p(\phi|\psi)\mathrm{d}\phi = \mathrm{Tr}[\rho_\psi \mathrm{d}\nu(\phi)], \tag{5}$$

where $p(\phi|\psi)$ represents the conditional probability of registering the outcome ϕ when the true value of the parameter is ψ.

Our objective is to find the best strategy, i.e. the POVM that provides the optimal estimation of the parameter ψ [5]. Since ψ is manifestly a phase-shift, we can use general results from the phase estimation theory, which provides the optimal POVM to estimate the phase-shift induced by a *phase generator*, i.e. a self-adjoint operator with a discrete, equally spaced, spectrum. The optimality criterion is given in terms of the minimization of the mean value of a cost function [5] that assesses the quality of the strategy, i.e. it weights the errors in the estimates. Since a phase-shift is a 2π-periodic parameter, the cost function must be a 2π-periodic even function of $(\psi - \phi)$, i.e. it has Fourier expansion with cosines only. The appropriate concavity properties of the cost function are reflected by expansion coefficients which are all negative apart from the irrelevant additive constant. A cost function with such Fourier expansion was first considered by Holevo [4], and for this reason it is usually referred to as belonging to Holevo's class. Notice that the optimal POVM for a given state ρ_0 is the same for every cost function in this class. A more general quantum estimation approach for different kinds of phase-shift and general quantum system is given in [6]. There, it is also shown that the kind of problem we are presently dealing with is, in a certain sense, the best situation, since the spectrum of our phase generator contains the whole set of integers $\mathbb{Z}$, including the negative ones. In this case, there is an optimal orthogonal projective POVM, which can be regarded as the spectral resolution of a self-adjoint phase operator. However, if the estimation is performed with the constraint of bounded or fixed energy, the optimal POVM and the optimal input state ρ_0 do not correspond to a canonical quantum observable scheme [7]. Moreover, in general, the optimal POVM depends on the input state and the optimal POVM of [4, 5] holds only for pure states, whereas a generalization to a class of mixed states, the so-called *phase-pure* states, have been considered in [6, 8].

The optimal POVM, in the sense described above, provides an unbiased estimation of ψ for preparation ρ_0 of the two modes, in formula

$$\langle\phi\rangle = \psi \ \text{ with } \langle\phi\rangle = \int_0^{2\pi} \phi \mathrm{Tr}[\rho_\psi \mathrm{d}\nu(\phi)] \ \ \forall \rho_\psi = U_\psi \rho_0 U_\psi^\dagger.$$

On the other hand, we also want to find the optimal state ρ_0 for the estimation of ψ according to the cost function, which quantifies the noise of the estimation. The customary root mean square is not a good choice for a cost function, since the function $(\phi - \psi)^2$ is not 2π-periodic. A good definition for the precision of the measurement is given by the average of the cost function $C(\phi - \psi) = 4\sin^2(\frac{\phi-\psi}{2})$, i.e. a 'periodicized variance', which obviously belongs to Holevo's class. If the estimates occur within a small interval around the true value of the parameter ψ, one has approximately $C(\phi-\psi) \simeq (\phi-\psi)^2$, where $\delta\psi = \sqrt{C}$ can be assumed as a reasonable measure of the precision of the measurement.

2. Optimal estimation of the coupling parameter

In order to solve our estimation problem, let us consider the following unitary transformation

$$\mathcal{U} = \exp\left\{-\frac{\pi}{4}(a^\dagger b - b^\dagger a)\right\} = \exp\left\{-\mathrm{i}\frac{\pi}{2}J_y\right\}. \tag{6}$$

Using equation (6) we may rewrite equation (5) in the more familiar form of rotation along the z-axis

$$\begin{aligned} p(\phi|\psi)\mathrm{d}\phi &= \mathrm{Tr}\left[\mathcal{U}\rho_\psi \mathcal{U}^\dagger \mathcal{U}\mathrm{d}\nu(\phi)\mathcal{U}^\dagger\right] \\ &= \mathrm{Tr}\left[\exp(-\mathrm{i}J_z\psi)\mathcal{U}\rho_0\mathcal{U}^\dagger \exp(\mathrm{i}J_z\psi)\mathcal{U}\mathrm{d}\nu(\phi)\mathcal{U}^\dagger\right] \\ &= \mathrm{Tr}\left[\exp(-\mathrm{i}J_z\psi)R_0\exp(\mathrm{i}J_z\psi)\,\mathrm{d}\mu(\phi)\right], \end{aligned} \tag{7}$$

where we used the identity $\mathcal{U}J_x\mathcal{U}^\dagger = J_z$. Equation (7) shows that the problem of estimating the shift generated by J_x on the state ρ_0 is equivalent to that of estimating the same shift generated by J_z on the rotated state $R_0 = \mathcal{U}\rho_0\mathcal{U}^\dagger$. In particular, any POVM $\mathrm{d}\nu(\phi)$ to estimate the J_x-induced shift can be written as $d\nu(\phi) = \mathcal{U}^\dagger \mathrm{d}\mu(\phi)\mathcal{U}$, where $\mathrm{d}\mu(\phi)$ is a POVM for the J_z-induced shift estimation.

For pure states $R_0 = |\psi_0\rangle\rangle\langle\langle\psi_0|$ (in the following we use double brackets for two-mode vectors) the degeneration of the spectrum of $a^\dagger a - b^\dagger b$ can be treated using the technique introduced in [6], and the optimal POVM for cost functions in Holevo's class [4] is proved to be of the form

$$\mathrm{d}\mu(\phi) = \frac{\mathrm{d}\phi}{2\pi}\,|E_\phi\rangle\rangle\langle\langle E_\phi| \tag{8}$$

with the vectors $|E_\phi\rangle\rangle$ given by

$$|E_\phi\rangle\rangle = \sum_{d\in\mathbb{Z}} \mathrm{e}^{id\phi}|d\rangle\rangle. \tag{9}$$

The vectors $|d\rangle\rangle$ are certain eigenvectors of $D = a^\dagger a - b^\dagger b$ built by picking up, in every eigenspace $\mathcal{H}_d$ of the eigenvalue d, the normalized vector parallel to the projection of $|\psi_0\rangle\rangle$ on $\mathcal{H}_d$. In order to be more specific, let us consider an input state of the form

$$\begin{aligned} |\psi_0\rangle\rangle &= \sum_{n=0}^{\infty}\psi_n^{(0)}|n,0\rangle\rangle \\ &+\sum_{n=0}^{\infty}\sum_{d=1}^{\infty}\left[\psi_n^{(d)}|n,d\rangle\rangle + \psi_n^{(-d)}|n,-d\rangle\rangle\right], \end{aligned} \tag{10}$$

where $|n,d\rangle\rangle$ is given by

$$|n,d\rangle\rangle \equiv \begin{cases} |n+d\rangle_a|n\rangle_b & \text{if } d \geqslant 0 \\ |n\rangle_a|n-d\rangle_b & \text{if } d < 0 \end{cases}. \tag{11}$$

The projection of $|\psi_0\rangle\rangle$ in $\mathcal{H}_d$ is equal to

$$\sum_{n=0}^{\infty}\psi_n^{(d)}\,|n,d\rangle\rangle, \tag{12}$$

such that the eigenvector $|d\rangle\rangle$ reads as follows

$$|d\rangle\rangle = \frac{\sum_{n=0}^{\infty}\psi_n^{(d)}|n,d\rangle\rangle}{\sqrt{\sum_{n=0}^{\infty}\left|\psi_n^{(d)}\right|^2}}, \tag{13}$$

and the input state can be rewritten as

$$|\psi_0\rangle\rangle = \sum_{d\in\mathbb{Z}}\gamma_d|d\rangle\rangle \qquad \gamma_d = \sqrt{\sum_{n=0}^{\infty}\left|\psi_n^{(d)}\right|^2}.$$

Notice that the dependence of the POVM on the state $|\psi_0\rangle\rangle$ is contained in the vectors $|E_\phi\rangle\rangle$.

By adopting $C(\phi - \psi)$ as a cost function the average cost of the strategy corresponds to the expectation value of the *cost* operator $C = 2 - \mathrm{E}_+ - \mathrm{E}_-$, where the raising and lowering operators E_+ and E_- are given by

$$\mathrm{E}_+ = \sum_{d\in\mathbb{Z}}|d+1\rangle\rangle\langle\langle d| \qquad \mathrm{E}_- = \mathrm{E}_+^\dagger$$

(with the vectors $|d\rangle\rangle$ defined as above).

The optimization problem is that of minimizing the average cost of the strategy

$$\bar{C} = \int_0^{2\pi} \frac{d\psi}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi}\, C(\phi - \psi)\; p(\phi|\psi)$$
$$= \mathrm{Tr}\,[R_0\, C] \equiv \langle\langle \psi_0 | 2 - \mathrm{E}_+ - \mathrm{E}_- | \psi_0 \rangle\rangle, \qquad (14)$$

with the constraint that the solution is a normalized state. The Lagrange function is given by

$$\mathcal{L} = \bar{C} - \lambda \langle\langle \psi_0 | \psi_0 \rangle\rangle, \qquad (15)$$

with λ being the Lagrange multiplier for the normalization constraint. The solution of this problem is a state with infinite mean energy $N = \langle\langle \psi_0 | a^\dagger a + b^\dagger b | \psi_0 \rangle\rangle$. Indeed, these states are unitarily connected to the eigenstates of the relative-phase operator [9] through the transformation of equation (6). In order to find physical states, one must impose a constraint on N too, and the Lagrange function becomes

$$\mathcal{L} = \bar{C} - \mu \langle\langle \psi_0 | a^\dagger a + b^\dagger b | \psi_0 \rangle\rangle - \lambda \langle\langle \psi_0 | \psi_0 \rangle\rangle, \qquad (16)$$

μ being the Lagrange multiplier for the mean energy. It is useful, in order to calculate the solution of this equation, to write the generic state $|\psi_0\rangle\rangle$ in the following way

$$|\psi_0\rangle\rangle = \sum_{d\in\mathbb{Z}} \psi_d \sum_{n=0}^{\infty} c_{n,d} |n, d\rangle\rangle. \qquad (17)$$

The coefficients $c_{n,d}$ determine the normalized projection of $|\psi\rangle\rangle$ into the eigenspace $\mathcal{H}_d$, whereas the ψ_d's are the coefficients which combine those projections.

Using equation (17), the Lagrange function (16) explicitly shows terms accounting for the normalization of the vectors $\sum_{n=0}^{\infty} c_{n,d} |n, d\rangle\rangle$ in each $\mathcal{H}_d$, with $\nu^{(d)}$ denoting Lagrange multipliers for the normalization of projections, and rewrites as

$$\mathcal{L} = \sum_{d\in\mathbb{Z}} \Bigg\{ 2|\psi_d|^2 \sum_{n=0}^{\infty} |c_{n,d}|^2$$
$$- \left(\bar{\psi}_d \sum_{n=0}^{\infty} |c_{n,d}|^2 \right) \left(\psi_{d-1} \sum_{m=0}^{\infty} |c_{m,d-1}|^2 \right) +$$
$$- \left(\bar{\psi}_d \sum_{n=0}^{\infty} |c_{n,d}|^2 \right) \left(\psi_{d+1} \sum_{m=0}^{\infty} |c_{m,d+1}|^2 \right) +$$
$$-\mu |\psi_d|^2 \sum_{n=0}^{\infty} |c_{n,d}|^2 (2n + |d|)$$
$$-\lambda |\psi_d|^2 \sum_{n=0}^{\infty} |c_{n,d}|^2 - \nu^{(d)} \sum_{n=0}^{\infty} |c_{n,d}|^2 \Bigg\}. \qquad (18)$$

By taking derivatives of the Lagrange function with respect to $c_{n,d}^*$ and ψ_d^* with the constraints

$$\sum_{n=0}^{\infty} |c_{n,d}|^2 = 1 \qquad \sum_{d\in\mathbb{Z}} |\psi_d|^2 = 1$$
$$\sum_{d\in\mathbb{Z}} \sum_{n=0}^{\infty} |\psi_d|^2 |c_{n,d}|^2 (2n + |d|) = N, \qquad (19)$$

and by rephasing the $|d\rangle\rangle$'s we arrive at the system

$$\begin{cases} (2-\lambda)\psi_d - \psi_{d-1} - \psi_{d+1} \\ -\mu \sum_{n=0}^{\infty} (2n + |d|) |c_{n,d}|^2 \psi_d = 0 \\ \Big[(2-\lambda)\psi_d^2 - 2(\psi_d \psi_{d-1} + \psi_{d+1}\psi_d) \\ -\mu \psi_d^2 (2n + |d|) - \nu^{(d)} \Big] c_{n,d} = 0 \end{cases} \qquad (20)$$

The second equation in (20) implies that for a fixed d only one coefficient $c_{n,d}$ can be different from zero, say for the value $\bar{n}$, and in this case $|c_{\bar{n},d}| = 1$ [4].

The first equation of (20) can therefore be rewritten as

$$(2-\lambda)\psi_d - \psi_{d-1} - \psi_{d+1} - \mu(2n(d) + |d|)\psi_d = 0, \qquad (21)$$

which allows us to obtain from the second one

$$\nu^{(d)} = -\psi_d(\psi_{d+1} + \psi_{d-1})\,. \qquad (22)$$

The solutions of equation (22) give local minima for the average cost $\bar{C}$, and one should solve the equation (21) for arbitrary choices of $n(d)$, looking for the optimal one. In the case $n(d) = 0$ we have

$$\frac{2(\lambda' + |d|)}{\frac{2}{\mu'}} \psi_d = \psi_{d-1} + \psi_{d+1}\,, \qquad (23)$$

where $\mu' = -\mu$ and $\lambda' = \frac{2-\lambda}{\mu'}$. Equation 23 is the recursion equation for Bessel functions, with solution given by

$$\psi_d = \mathcal{N}^{-1/2}(\mu', \lambda') J_{\lambda' + |d|} \left(\frac{2}{\mu'} \right), \qquad (24)$$

with $\mathcal{N}(\mu', \lambda') = \sum_{d\in\mathbb{Z}} J^2_{\lambda'+|d|} \left(\frac{2}{\mu'} \right)$ and with the boundary conditions $J'_{\lambda'}(2/\mu') = 0$, i.e. $J_{\lambda'+1} = J_{\lambda'-1}$. Finally, to obtain the optimal state ρ_0 one has to rotate $R_0 = |\psi_0\rangle\rangle\langle\langle\psi_0|$ by the unitary transformation (6).

In order to obtain the behaviour of the average cost versus the energy we numerically solved equation (21) with $n(d) = 0$. This problem can be rewritten as the eigenvalue problem $\mathbf{A}\psi = \lambda\psi$ for the matrix $\mathbf{A}$ with elements given by

$$(\mathbf{A})_{m,n} = (2 - \mu|m|)\delta_{m,n} - \delta_{m,n+1} - \delta_{m,n-1}. \qquad (25)$$

Numerical diagonalization gives the power-law $\bar{C} \simeq \frac{\gamma}{N^2}$ in the range $0 \lesssim N \lesssim 1000$ with $\gamma \simeq 0.1$. This behaviour is plotted in figure 1, where other solutions of equation (21) corresponding to local minima of the average cost are also shown. Since the phase distribution of the optimal state is singly peaked we may also write $\delta\psi \simeq \sqrt{\bar{C}} \simeq \gamma^{1/2}/N$, which means that the optimal states derived here are at the so-called Heisenberg limit of phase variance.

3. Conclusions

In conclusion, we dealt with the problem of estimating the coupling constant of a bilinear interaction. In our approach the coupling constant, which appears in the exponent of

[4] Indeed, let us suppose $c_{n,d} \neq 0$ for two values $n_1 = m$ and $n_2 = p$. Then we must have $(2-\lambda)\psi_d^2 - 2(\psi_d\psi_{d-1} + \psi_{d+1}\psi_d) - \mu\psi_d^2(2p + |d|) - \nu^{(d)} = 0$ which implies $2\mu(m - p)\psi_d = 0$. Since the case $\mu = 0$ is not interesting and $\psi_d = 0$ would imply that the choice of the $c_{n,d}$ is completely arbitrary and irrelevant, the only possibility is $m = p$.

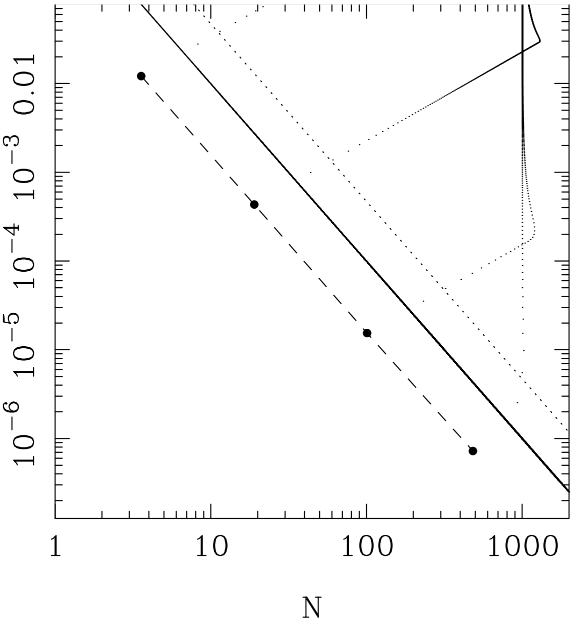

Figure 1. Average cost $\bar{C}$ as a function of the energy N for the optimal states (dashed line). The solid line is the function $1/N^2$. The points above the solid line are other solutions of equation (21) corresponding to local minima of the cost, distributed over lines characterized by a fixed value of the Lagrange parameter μ (a single dotted line is plotted, connecting points for different μ and increasing N).

the time evolution operator, has been treated as a phase parameter. The optimal POVM has been derived according to the theory of quantum phase estimation [4–6, 8]. As noted in [2] this resorts to an $SU(2)$ estimation problem with the Schwinger two-mode boson realization. However, the representation is not irreducible, and a more complicated problem is faced. In this sense our results generalize those of [2], where the estimation of $SU(2)$ phase-shifts has been analysed in irreducible subspaces: indeed we found an improved scaling of phase variance versus the total energy. The degeneracy of the spectrum of the the Hamiltonian can be treated using the technique of [6] and, in this way, the problem is reduced to a nondegenerate one with spectrum $\mathbb{Z}$ for the phase-shift operator. The $\delta\phi \propto N^{-1}$ scaling in interferometry has also been found for specific classes of input states such as optimized squeezed states [10] and number states [11]. Our analysis confirms such an unsurpassable bound, and provides the optimal states which, compared to states of [10, 11], show a largely reduced multiplicative constant.

It is important to remark that the optimal POVM depends on the preparation state. This is true for every kind of phase estimation problem, but in the presence of degeneracy the dependence is crucial. In fact, one has to define the optimal POVM as a block-diagonal operator, where the invariant subspaces are spanned by projections of the input state into the eigenspaces of the generator. From a practical point of view this means that optimal estimation of the phase-shift imposed to a state needs a measuring device which is adapted to the shifted state. We then optimized the input state by minimizing the average cost for fixed input energy and found a power law $\delta\psi \simeq \gamma/N$ in a range $0 \lesssim N \lesssim 1000$.

Notice that the law N^{-1} is the same as in the optimal phase estimation with only one mode [12], but with a much smaller constant γ ($\gamma \simeq 0.1$ instead of $\gamma \simeq 1.36$). We think that this phenomenon of improvement of phase sensitivity by increasing the number of modes is the same as that considered in [13], where an exponential improvement versus the number of modes has been estimated when increasing the number of modes and the number of photons per mode, jointly in the same proportion.

Acknowledgment

This work has been co-sponsored by MURST under the project 'Quantum Information Transmission and Processing' and by the INFM under the project PAIS-TWIN. The authors thank the Dipartimento di Fisica 'Alessandro Volta' for partial support.

References

[1] Mandel L and Wolf E 1995 *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press)
[2] Sanders B C and Milburn G J 1995 *Phys. Rev. Lett.* **75** 2944
[3] Paris M G A 1999 *Phys. Rev.* A **59** 1615
[4] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[5] Helstrom C W 1974 *Found. Phys.* **4** 453
Helstrom C W 1974 *Int. J. Theor. Phys.* **11** 357
Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[6] D'Ariano G M, Macchiavello C and Sacchi M F 1998 *Phys. Lett.* A **248** 103
[7] Paris M G A 1996 *Nuovo Cimento* B **111** 1151
[8] D'Ariano G M, Macchiavello C, Perinotti P and Sacchi M F 2000 *Phys. Lett.* A **268** 241
[9] Luis A and Sànchez-Soto L L 1993 *Phys. Rev.* A **48** 4708
[10] Paris M G A 1995 *Phys. Lett* A **201** 132
[11] Holland M J and Burnett K 1993 *Phys. Rev. Lett.* **71** 1355
[12] D'Ariano G M and Paris M G A 1994 *Phys. Rev.* A **49** 3022
[13] Yuen H P 1992 *Squeezed States and Uncertainty Relations* ed D Han *et al* (NASA CP 3135) p 13