acta physica slovaca vol. 48 No. 3, 1 – 8 **June 1998**

OPTIMIZED TOMOGRAPHY OF OBSERVABLES[1]

G. Mauro D'Ariano, Matteo G. A. Paris
Theoretical Quantum Optics Group[2]
Dipartimento di Fisica 'Alessandro Volta' dell' Universitá degli studi di Pavia
Istituto Nazionale di Fisica della Materia – Unitá di Pavia
via A. Bassi 6, I-27100 Pavia, Italy

Received May 12, 1998, accepted May 26, 1998

Tomographic measurement of observables is revisited, and an adaptive optimization of the kernel functions suggested. The method is based on the existence of a class of *null functions*, which have zero tomographic average for *any* state of radiation. The general procedure is illustrated, and application to relevant observables analyzed in details for coherent, squeezed and "cat" states.

1. Introduction

Quantum tomography of a single-mode radiation field consists in a set of repeated measurements of the field-quadrature $\hat{x}_\phi = \frac{1}{2}(ae^{-i\phi} + a^\dagger e^{i\phi})$ at different values of the reference phase ϕ. The expectation value of a generic operator can be expressed in terms of the tomographic probability $p(x,\phi)$ as

$$\langle \hat{O} \rangle \doteq \mathrm{Tr}\left\{ \hat{\varrho}\, \hat{O} \right\} = \int_0^\pi \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx\; p(x,\phi)\; R[\hat{O}](x,\phi) \;, \tag{1}$$

$R[\hat{O}](x,\phi)$ being the *tomographic kernel function* for the operator $\hat{O}$, whose definition reads as follows [1]

$$R[\hat{O}](x,\phi) = -\frac{1}{2}\mathrm{Tr}\left\{ \hat{O}\; \mathrm{P}\frac{1}{(x-\hat{x}_\phi)^2} \right\} \;, \tag{2}$$

P denoting the Cauchy principal value. Any physical property of the radiation field is the expectation value of some operator. By virtue of Eq. (1), the generic quantity of

[1] Special Issue on Quantum Optics and Quantum Information
[2] `http:// enterprise.pv.infn.it`

interest $\langle\hat{O}\rangle = \mathrm{Tr}\left\{\hat{\varrho}\,\hat{O}\right\}$ can be obtained as a simple average over the set of tomographic data [2]

$$\langle\hat{O}\rangle = \overline{R[\hat{O}]} = \frac{1}{N}\sum_{i=1}^{N} R[\hat{O}](x_i,\phi_i)\;, \tag{3}$$

N being the total number of measurements. The precision of the tomographic measurement in Eq. (3) can be easily evaluated provided that the corresponding kernel function satisfies the hypothesis of the central limit theorem [3], which assures that the partial average over a block of N_b data is always Gaussian distributed around the global average over all blocks. Thus, one evaluates the tomographic precision by dividing the ensemble of data into subensembles, and then calculates the r.m.s. deviation of subensemble averages with respect to the global one. The estimated value of such a confidence interval is given by

$$\delta O = \frac{1}{\sqrt{N}}\left\{\overline{\Delta R^2[\hat{O}]}\right\}^{1/2}\;, \tag{4}$$

where $\overline{\Delta R^2[\hat{O}]}$ is the variance of the kernel over the tomographic probability, that is

$$\overline{\Delta R^2[\hat{O}]} = \int_0^\pi \frac{d\phi}{\pi}\int_{-\infty}^{\infty} dx\; p(x,\phi)\; R^2[\hat{O}](x,\phi) - \left\{\int_0^\pi \frac{d\phi}{\pi}\int_{-\infty}^{\infty} dx\; p(x,\phi)\; R[\hat{O}](x,\phi)\right\}^2\;. \tag{5}$$

Following this scheme, the tomographic precision in determining matrix elements has been discussed in [2,4], whereas the measurements of relevant observables have been analyzed in [5], also in comparison with the corresponding direct detection.

For $\hat{O}$ not of trace-class the crude definition (2) of kernel function leads to singular expressions, and needs a regularization procedure. Tomographic kernels for the matrix elements in the Fock basis have been firstly calculated in [2], whereas an extension to data coming from inefficient detectors has been presented in [6]. A factorization method has been subsequently suggested in [7]. Tomographic kernels for the normally ordered moments have been evaluated in [8,14], and application to intensity [9] and phase [10] detection has been analyzed.

Indeed, the tomographic kernel for a given operator $\hat{O}$ is not unique, as there exist a large class of *null functions* $F(x,\phi)$, which have zero tomographic average for *any* state of radiation field, in formula

$$\overline{F} = \int_0^\pi \frac{d\phi}{\pi}\int_{-\infty}^{\infty} dx\; p(x,\phi)\; F(x,\phi)\;\equiv\;0\;. \tag{6}$$

Adding any number of null functions to a generic kernel results in a new kernel which has the same tomographic average, and thus is equivalent in the reconstruction of the given expectation value.

In this paper we describe a method to improve precision in the tomographic measurement of observables, which is based on the existence of null functions. In fact, adding null functions to a tomographic kernel does not affect its average, but generally modifies its variance, which can be effectively reduced by an adaptive optimization method.

The general procedure will be illustrated in Sec. 2, whereas its application to noise reduction of tomographic detection of relevant observables will be analyzed in Sec. 3.

2. Adaptive Tomography

Let us consider the following class of functions

$$F_{nk}(x,\phi) = x^k \exp\{\pm\, i\,(k+2+2n)\,\phi\} \qquad n,k=0,1,\ldots\,. \tag{7}$$

The tomographic average of $F_{nk}(x,\phi)$ can be written as

$$\overline{F_{nk}} = \int_0^\pi \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx\; p(x,\phi)\; F_{nk}(x,\phi) = \int_0^\pi \frac{d\phi}{\pi}\; e^{\pm\, i\,(k+2+2n)\,\phi}\, \langle \hat{x}_\phi^k \rangle\,, \tag{8}$$

where again $\langle\cdot\rangle$ denotes the ensemble average $\mathrm{Tr}\{\hat{\varrho}\,\cdot\}$. By means of the Wilcox decomposition formula we can write

$$\langle \hat{x}_\phi^k \rangle = \frac{k!}{2^k} \sum_{p=0}^{[[k/2]]} \sum_{s=0}^{k-2p} \frac{\langle a^{\dagger s} a^{k-2p-s}\rangle}{2^p p!\, s!\,(k-2p-s)!}\, e^{i\,(2p+2s-k)\,\phi}\,, \tag{9}$$

where $[[x]]$ denotes the integer part of x. Together with the trivial integral

$$\int_0^\pi \frac{d\phi}{\pi}\, e^{i\,q\,\phi} = \left\{ \begin{array}{cl} 0 & \text{q even} \\ 1 & \text{q} = 0 \\ 2i/q & \text{q odd} \end{array} \right. \,, \tag{10}$$

Eqs. (8) and (9) prove that $\overline{F_{nk}} \equiv 0 \quad \forall\, n,k$, namely that $F_{nk}(x,\phi)$ in Eq. (7) are tomographic null functions.

In order to optimize the tomographic kernels corresponding to the observables of interest in this paper we don't need to consider the whole class of functions, but, as it will be clear in the following it is enough to consider the null functions for $n=0$, namely

$$F_k(x,\phi) = x^k \exp\{i\,(k+2)\,\phi\} \qquad k=0,1,\ldots\,. \tag{11}$$

Let us consider a generic real kernel $R[\hat{O}](x,\phi)$. By adding M null functions with the constraint of maintaining the kernel as real, we arrive to the new kernel

$$K[\hat{O}](x,\phi) = R[\hat{O}](x,\phi) + \sum_{k=0}^{M-1} \mu_k F_k(x,\phi) + \sum_{k=0}^{M-1} \mu_k^* F_k^*(x,\phi)\,, \tag{12}$$

where μ_k are complex coefficients. ¿From the definition of null function, we have $\overline{K[\hat{O}]} = \overline{R[\hat{O}]}$, whereas the variance of the new kernel $K[\hat{O}]$ is given by

$$\overline{\Delta K^2[\hat{O}]} = \overline{\Delta R^2[\hat{O}]} + 2\left\{ \sum_{kl} \mu_k \mu_l^* \overline{F_k F_l^*} + \sum_k \mu_k \overline{R[\hat{O}] F_k} + \sum_k \mu_k^* \overline{R[\hat{O}] F_k^*} \right\}\,. \tag{13}$$

In deriving the above formula one considers the fact that the square of a null function is again a null function.

The variance (13) of the modified kernel can be minimized with respect to the coefficients μ_k, leading to the linear set of equations

$$\sum_l \mu_l \, \overline{F_k F_l^*} = -\overline{R[\hat{O}] F_k^*} \ . \tag{14}$$

The optimization equations (14) can be written in the matrix form

$$\mathbf{A}\, \mu = \mathbf{b} \ . \tag{15}$$

$\mathbf{A}$ being the Hermitean $M \times M$ matrix

$$A_{kl} = \overline{F_k F_l^*} \equiv \overline{x^{k+l} \exp\{i(k-l)\phi\}}$$

and $\mathbf{b}$ the complex vector

$$b_k = -\overline{R[\hat{O}] F_k^*}$$

The vector $\mathbf{b}$ depends both on the kernel to be optimized and on the state $\hat{\varrho}$ under examination, whereas the matrix $\mathbf{A}$ depends on the state only. The explicit expression of A_{kl} can be obtained by means of Eq.(9) as follows

$$A_{kl} = \frac{(k+l)!}{2^{k+l}} \sum_{p=0}^{min(k,l)} \frac{\langle a^{\dagger l-p}\, a^{k-p}\rangle}{2^p\, p!(l-p)!(k-p)!} \ . \tag{16}$$

By substituting Eq. (14) in Eq. (12) and inverting (15) we obtain

$$\Delta^2[\hat{O}] = \overline{\Delta R^2[\hat{O}]} - \overline{\Delta K^2[\hat{O}]} = 2 \sum_{kl}^{M-1} \mu_k \, A_{kl} \, \mu_l^* = 2 \sum_{kl}^{M-1} b_k \, \left(A^{-1}\right)_{kl} \, b_l^* \ , \tag{17}$$

which expresses the decrease of the variance in terms of $\mathbf{A}$ and $\mathbf{b}$.

In the case of a complex kernel $R[\hat{O}](x,\phi)$ (as for the detection of field amplitude) the equivalence class of kernel functions can be written as

$$K[\hat{O}](x,\phi) = R[\hat{O}](x,\phi) + \sum_{p=0}^{M-1} \mu_p F_p(x,\phi) + \sum_{p=0}^{M-1} \nu_k F_p^*(x,\phi) \ . \tag{18}$$

μ_p and ν_p being two independent sets of complex coefficients. The noise-figure that we want to optimized now is

$$\overline{\Delta_* K^2[\hat{O}]} = \frac{1}{2} \left\{ \overline{\left|K[\hat{O}]\right|^2} - \left|\overline{K^2[\hat{O}]}\right|^2 \right\} \ , \tag{19}$$

corresponding to the average of the noises for the real and imaginary part of $K[\hat{O}](x,\phi)$ respectively, namely the trace of the covariance matrix [5]. The optimization procedure is similar to the real case, and is reduced to solving the two linear systems

$$\mathbf{A}\, \mu = \mathbf{b} \qquad \mathbf{A}\, \nu = \mathbf{c} \ , \tag{20}$$

where $\mathbf{c}$ is given by

$$c_p = -\overline{R[\hat{O}]F_p} \; .$$

Finally, by inverting (20), we arrive to

$$\Delta^2_*[\hat{O}] = \overline{\Delta_* R^2[\hat{O}]} - \overline{\Delta_* K^2[\hat{O}]} = \sum_{p,q=0}^{M-1} \left[b_p \, \left(A^{-1}\right)_{qp} \, b_q^* + c_p \, \left(A^{-1}\right)_{pq} \, c_q^* \right] \; . \qquad (21)$$

The optimization procedure for a kernel $R[\hat{O}](x,\phi)$ can be summarized as follows: after collecting an ensemble of N tomographic data the quantities $\mathbf{A}$ and $\mathbf{b}$ (and $\mathbf{c}$ if needed) are evaluated as experimental averages. Then, by solving the linear systems (15) or (20) one obtains the coefficients which are used to build the optimized kernel $K[\hat{O}](x,\phi)$. At this point the same data set is used to average $K[\hat{O}](x,\phi)$ and, upon dividing the set into subensembles, to evaluate the experimental error, which now is reduced by the quantity $\Delta^2[\hat{O}]/N$ for a real kernel and $\Delta^2_*[\hat{O}]/N$ for a complex kernel. In the next section the noise reduction $\Delta^2[\hat{O}]$ or $\Delta^2_*[\hat{O}]$ pertaining the kernel for the tomographic detection of intensity, quadrature and field is evaluated. A Monte Carlo simulation of the whole optimization procedure is also reported for the case of intensity detection.

3. Noise reduction in tomographic measurements

As a starting point for the present optimization procedure we take the Richter form of the tomographic kernel for the normally ordered moment [8]

$$R[a^{\dagger n} a^m](x;\phi) = e^{i(m-n)\phi} \frac{H_{n+m}(\sqrt{2}\, x)}{\sqrt{2^{n+m}\binom{n+m}{n}}} \; , \qquad (22)$$

$H_n(x)$ being the Hermite polynomial of order n. In the case of tomographic detection of intensity Eq.(22) provides the kernel

$$R[a^\dagger a](x) = 2x^2 - \frac{1}{2} \; , \qquad (23)$$

whereas the vector $\mathbf{b}$ needed to optimize the kernel is given by

$$b_k = -\overline{R[a^\dagger a]F_k^*} = \overline{2x^{k+2} e^{-i(k+2)\phi}} = \frac{\langle a^{\dagger(k+2)} \rangle}{2^{1+k}} \; . \qquad (24)$$

We solved the optimization equations (15) analytically [11] for coherent states, squeezed vacuum and the "cat" superposition of coherent states $|\psi\rangle = [2(1+\exp\{-2|\alpha|^2)]^{-1/2}(|\alpha\rangle + |-\alpha\rangle)$ with M up to $M=10$ null functions. For all the states here considered, it turns out that only a single null function $F_0(\phi)$ is needed, namely one has

$$\mu_0 = b_0 \qquad \mu_k = 0 \, , \quad \forall \, k \geq 1 \; . \qquad (25)$$

The corresponding reduction of variance is given by

$$\Delta^2[a^\dagger a] = \frac{1}{2} \, \langle a^{\dagger 2} \rangle \, \langle a^2 \rangle \; , \qquad (26)$$

and can compensate the leading term of the variance of the original Richter kernel [5]

$$\overline{\Delta R^2[a^\dagger a]} = \langle\widehat{\Delta n^2}\rangle + \frac{1}{2}\left[\langle a^{\dagger 2}\, a^2\rangle + 2\langle a^\dagger a\rangle + 1\right] \ , \tag{27}$$

so that $\overline{\Delta K^2[a^\dagger a]}$ becomes much closer to the intrinsic intensity fluctuations $\langle\widehat{\Delta n^2}\rangle$ than the original noise $\overline{\Delta R^2[a^\dagger a]}$. In order to appreciate such noise reduction we introduce the two noise ratios

$$\delta n_R = \sqrt{\frac{\overline{\Delta R^2[a^\dagger a]}}{\langle\widehat{\Delta n^2}\rangle}} \qquad \delta n_K = \sqrt{\frac{\overline{\Delta K^2[a^\dagger a]}}{\langle\widehat{\Delta n^2}\rangle}} \ . \tag{28}$$

For coherent states $|\alpha\rangle$ we have

$$\delta n_R = \sqrt{2 + \frac{1}{2}\left(|\alpha|^2 + \frac{1}{|\alpha|^2}\right)} \qquad \delta n_K = \sqrt{2 + \frac{1}{2|\alpha|^2}} \ , \tag{29}$$

that is, from an asymptotically linearly increasing function $|\alpha|$ the ratio becomes a constant $\delta n_K \simeq \sqrt{2}$. Similar expressions are obtained for squeezed vacuum and cat states: in all cases the noise ratio saturates to a constant, which is $\delta n_K = \sqrt{3/2}$ for both the squeezed vacuum and the cat states.
In Fig. 1 we show the results from a Monte Carlo simulation of the tomographic detection of intensity on coherent states: the noise reduction obtained by using modified kernel is apparent.

We have applied the optimization procedure also to the quadrature $\hat{x} = \frac{1}{2}(a + a^\dagger)$, namely kernel $R[\hat{x}](x,\phi) = 2x\cos\phi$, and to the field amplitude kernel $R[a](x,\phi) = 2xe^{i\phi}$, where the optimization vectors $\mathbf{b}$ are given by $b_k = 2^{-k-1}\langle a^{1+k}\rangle$ and $b_k = 2^{-k}\langle a^{1+k}\rangle$ respectively. In both cases the optimization procedure shows that only the odd-index null functions $F_{2s+1}(x,\phi)$ give a contribution to noise reduction. Actually, the main contribution is obtained by adding the single null function $F_1(x,\phi)$, whereas null functions of higher order only improve the variances by a few percent. The variance reduction obtained by adding $F_1(x,\phi)$ is the same for both observables $\hat{x}$ and a and is given by

$$\Delta^2[\hat{x}] = \Delta^2_*[a] = \frac{1}{2\left(1 - |\langle a\rangle|^2 + 2\langle a^\dagger a\rangle\right)}\left[|\langle a\rangle|^2\left(\langle a^{\dagger 2}\rangle + \langle a^{\dagger 2}\rangle + \frac{1}{2} + \langle a^\dagger a\rangle\right) + |\langle a^2\rangle|^2\right] .$$

We calculated the noise ratios δx_K and δa_K, defined analogously in Eq.(28), for coherent states, the squeezed vacuum, and cat states, as a function of the mean photon number: in all cases the noise ratio saturates to a constant value, which is $\delta a_K = 1$, and $\delta x_K = \sqrt{2}$ for coherent states and $\delta a_K \simeq \sqrt{3/2}$ and $\delta x_K \simeq \sqrt{5/4}$ for both the squeezed vacuum and the cat states. Remarkably, for coherent states the heterodyne noise ration is reached.

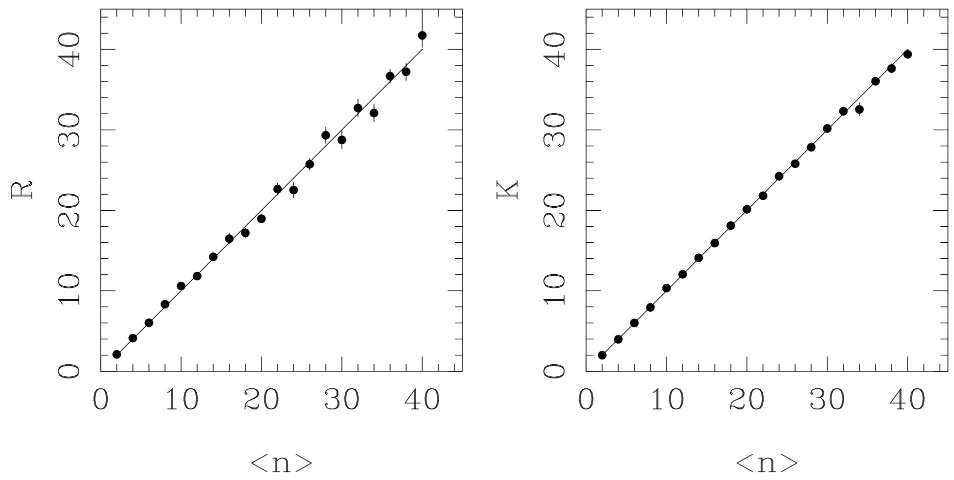

Fig. 1. Tomographic detection of the intensity on coherent states. The simulated experiment has been performed using 20 blocks of 20 data for 20 phases each, for a total number of $N = 8 \cdot 10^3$ measurements. We report the tomographic determination as a function of the actual mean number of photons of the coherent state. On the left we show $\overline{R[a^\dagger a]}$, whereas on the right we show $\overline{K[a^\dagger a]}$. The noise reduction obtained by using the optimized kernel is apparent.

4. Conclusions

In this paper we have presented an optimization procedure for the kernel functions used in tomographic detection of field observables. The application to intensity detection, quadrature detection, and field amplitude detection results in a large reduction of tomographic noise. The behaviour of noise for coherent states, squeezed vacuum, and cat states have been analyzed in details, and the conclusion is that the ratio between tomographic noise and intrinsic fluctuations of the considered observable saturates to constant values for increasing energy. We can *now* say that quantum tomography adds only a small amount of noise in comparison to the direct detection of a given observable. Notice that in this paper we have considered the tomographic detection without systematic errors, namely with the reference phase ϕ as a random parameter in $[0, \pi]$. As a matter of fact, a discrete scanning by equally spaced phases will introduce systematic errors [12,13]. For such uniform scanning, the null function $F_0(\phi)$ would have no effects on all kernels, whereas the other null functions would have reduced effect, without eliminating the systematic errors.

A systematic application of the present method to the detection of the density matrix and to generic kernel functions is under study and will be published elsewhere.

Acknowledgements We would thank Dirk -G. Welsch and Mohamed Dakna for valuable discussions and INFM (contract PRA97-CAT) for financial support.

References

[1] G. M. D'Ariano: in *Quantum Optics and Spectroscopy of Solids*, Eds. A.S. Shumowsky et al. (Kluwer, Amsterdam, 1997) pp.175-202

[2] G. M. D'Ariano, C. Macchiavello, M. G. A. Paris: *Phys. Rev. A* **50** (1994) 4298

[3] Central limit theorem applies to functions having bounded moments up to third order.

[4] G. M. D'Ariano, C. Macchiavello: *Phys. Rev. A* **57** (1998) 3131

[5] G. M. D'Ariano, M. G. A. Paris: *Phys. Lett. A* **233** (1997) 49

[6] G. M. D'Ariano, U. Leonhardt, H. Paul: *Phys. Rev. A* **52** (1995) 1801

[7] U. Leonhardt, M. Munroe, T. Kiss, Th. Richter, M. G. Raymer: *Opt. Comm.* **127** (1996) 144

[8] Th. Richter: *Phys. Lett. A* **221** (1996) 327

[9] G. M. D'Ariano, M. G. A. Paris: *Acta Phys. Slovaca* **47** (1997) 281

[10] T. Opatrny, M. Dakna, D.-G. Welsch: *Phys. Rev. A* **57** (1998) 2129

[11] We solved the systems by hand for $M = 2, 3$ and by MATHEMATICA up to $M = 10$.

[12] G. M. D'Ariano, C. Macchiavello, N. Sterpi: *Quantum Opt.* **9** (1997) 929

[13] U. Leonhardt, M.Munroe: *Phys. Rev. A* **54** (1996) 3682

[14] G. M. D'Ariano: in *Quantum Communication, Computing, and Measurement*, Eds. O. Hirota, A. S. Holevo, C. M. Caves (Plenum Publishing, New York, 1997)