

## Adaptive quantum homodyne tomography

Giacomo M. D'Ariano and Matteo G. A. Paris

*Theoretical Quantum Optics Group, Istituto Nazionale di Fisica della Materia – Unità di Pavia,  
Dipartimento di Fisica “Alessandro Volta” dell’Università di Pavia, via Bassi 6, I-27100 Pavia, Italy*

(Received 14 December 1998)

An adaptive optimization technique to improve the precision of quantum homodyne tomography is presented. The method is based on the existence of so-called *null functions*, which have a zero average for an arbitrary state of radiation. The addition of null functions to the tomographic kernels does not affect their mean values, but changes statistical errors, which can then be reduced by an optimization method that “adapts” kernels to homodyne data. Applications to tomography of the density matrix and other relevant field observables are studied in detail. [S1050-2947(99)00707-6]

PACS number(s): 42.50.Dv

### I. INTRODUCTION

The possibility of measuring the quantum state of radiation has received increasing interest in the last years [1–3], as it opens perspectives for new kinds of experiments in quantum optics, with the possibility of measuring photon correlations on a subpicosecond time scale [4], of characterizing squeezing properties [5], photon statistics in parametric fluorescence [6], quantum correlations in down conversion [7], and nonclassicality of states [8], and of measuring Hamiltonians of nonlinear optical devices [9]. Among the many state reconstruction techniques suggested in the literature [10–19], quantum homodyne tomography (QHT) [11–13,18] of radiation fields has received much attention [1], being the only method which has been implemented in quantum optical experiments [4,5,11]. It has also recently been extended to estimate the expectation value of any operator of the field [18], which makes the method the first universal detectors for radiation.

On one hand, QHT takes advantage of amplification from the local oscillator in the homodyne detector, avoiding the need for single-photon resolving photodetectors, and hence created the possibility of achieving very high quantum efficiency using photodiodes [7]. On the other hand, the method of QHT is very efficient and statistically reliable, and can be implemented on line with the experiment.

In principle, a precise knowledge of the density matrix would require an infinite number of measurements on identical preparations of radiation. However, in real experiments one has only a finite number of data at ones disposal, and thus a statistical analysis and errors estimation are needed. The purpose of this paper is to analyze the possibility of improving the current QHT technique, in order to minimize statistical errors. We will present a method that “adapts” the tomographic estimators to a given finite set of data, improving the precision of the tomographic measurement.

Quantum tomography of a single-mode radiation field consists of 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  $\phi$ . The expectation value of a generic operator can be obtained by averaging a suitable kernel function  $R[\hat{O}](x, \phi)$  as follows [18]:

$$\langle \hat{O} \rangle \doteq \text{Tr}\{\hat{\rho}\hat{O}\} = \int_0^{\pi} \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx p_\eta(x, \phi) R_\eta[\hat{O}](x, \phi), \quad (1)$$

where  $p_\eta(x, \phi)$  denotes the probability distribution of the outcomes  $x$  for the quadrature  $\hat{x}_\phi$  with quantum efficiency  $\eta$ , and  $R_\eta[\hat{O}](x, \phi)$  is given by

$$R_\eta[\hat{O}](x, \phi) = \frac{1}{4} \int_0^\infty dr \exp\left[\frac{1-\eta}{8\eta}r\right] \text{Tr}\{\hat{O} \cos[\sqrt{r}(x - \hat{x}_\phi)]\}. \quad (2)$$

In the following we will focus attention only on the  $\eta=1$  case, and we will drop the subscript  $\eta$  in the notation. As will become clear in the following, the method also works equally well for nonunit quantum efficiency, and a detailed numerical analysis versus  $\eta$  will be given elsewhere. On the basis of identity (1), it follows that the ensemble average  $\langle O \rangle$  can be experimentally obtained by averaging  $R[\hat{O}](x, \phi)$  over the set of homodyne data, namely

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

$N$  being the total number of measurements of the sample. The statistical error 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, which assures that the partial average over a block of data is Gaussian distributed around the global average over all data. In this case, the error is evaluated by dividing the ensemble of data into subensembles, and calculating the rms deviation of each subensemble mean value with respect to the global average. The estimated value of such a confidence interval is given by

$$\delta O = \frac{1}{\sqrt{N}} \overline{\{\Delta R^2[\hat{O}]\}}^{1/2}, \quad (4)$$

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

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

Following this scheme, the tomographic precision in determining matrix elements of the density operator  $\hat{\rho}$  was discussed in Refs. [13,20,21], with asymptotic estimations in Ref. [22], whereas relevant observables  $\hat{O}$  were analyzed in Ref. [23], also in comparison with the corresponding ideal measurement.

The crucial point of the method presented in this paper is that the tomographic kernel  $R[\hat{O}](x, \phi)$  is not unique, since a large class of *null functions* [24,25]  $F(x, \phi)$  exists that have a zero tomographic average for arbitrary state, namely,

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

Therefore, the addition of null functions to a generic kernel gives a kernel with the same tomographic average, hence suitable for an estimation of the same ensemble average  $\langle \hat{O} \rangle$ . On the other hand, adding null functions would modify the kernel variance, and so the statistical error over data. The adaptive tomography method thus consists of optimizing the kernel in the equivalence class, in order to minimize the statistical errors. In this paper, we perform such optimization over relevant classes of null functions, rather than over the entire equivalence class.

The paper is structured as follows. In Sec. II we introduce the classes of null functions that will be used in the paper, and describe the adaptive optimization method in detail. In Sec. III we apply the adaptive method to the tomography of the density matrix in the photon number representation. In Sec. IV we analyze the improvement of precision in tomographic measurement of some relevant field observables. Section V briefly describes the effects of systematic errors on the effectiveness of the method. Finally, Sec. VI closes the paper by summarizing the main results.

## II. ADAPTIVE TOMOGRAPHY

The following functions have the vanishing tomographic expectation [Eq. (6)]

$$\begin{aligned} G_n^+(x, \phi) &= e^{i(1+n)2\phi} g_+(xe^{i\phi}), \\ G_n^-(x, \phi) &= e^{-i(1+n)2\phi} g_-(xe^{-i\phi}). \end{aligned} \quad (7)$$

In Eqs. (7),  $n \geq 0$  and  $g_{\pm}(z)$  are analytic functions of  $z$ . The set  $\mathcal{G}$  of null functions defined in Eqs. (7) forms a vector space over  $\mathbb{C}$ , and each class  $\mathcal{G}^{\pm} = \{G_n^{\pm}\}$  separately is closed under multiplication (without inverse).

In order to prove the vanishing expectation (6) for  $G_n^{\pm}(x, \phi)$ , we consider the Taylor expansion of functions  $g_{\pm}(xe^{i\phi})$ :

$$g_{\pm}(xe^{i\phi}) = \sum_{k=0}^{\infty} c_k^{\pm} x^k e^{\pm ik\phi}, \quad (8)$$

which allows us to write

$$\begin{aligned} \overline{G_n^{\pm}} &= \int_0^{\pi} \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx p(x, \phi) e^{\pm i(1+n)2\phi} g_{\pm}(xe^{i\phi}) \\ &= \sum_{k=0}^{\infty} c_k^{\pm} \int_0^{\pi} \frac{d\phi}{\pi} e^{\pm i(k+2+2n)\phi} \langle \hat{x}_{\phi}^k \rangle, \end{aligned} \quad (9)$$

where  $\langle \dots \rangle$  denotes the usual ensemble average. Using the Wilcox decomposition formula [26], one can write

$$\langle \hat{x}_{\phi}^k \rangle = \frac{k!}{2^k} \sum_{p=0}^{\lfloor k/2 \rfloor} \sum_{s=0}^{k-2p} \frac{\langle a^{\dagger s} a^{k-2p-s} \rangle}{2^p p! s! (k-2p-s)!} e^{i(2p+2s-k)\phi}, \quad (10)$$

where  $\lfloor [x] \rfloor$  denotes the integer part of  $x$ . Equation (10), together with the identity

$$\int_0^{\pi} \frac{d\phi}{\pi} e^{iq\phi} = \begin{cases} 0, & q \text{ even} \\ 1, & q=0 \\ \frac{2i}{\pi q}, & q \text{ odd}, \end{cases} \quad (11)$$

prove that

$$\int_0^{\pi} \frac{d\phi}{\pi} e^{\pm i(k+2+2n)\phi} \langle \hat{x}_{\phi}^k \rangle = 0, \quad n \geq 0, k \geq 0. \quad (12)$$

Hence

$$\int_0^{\pi} \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx p(x, \phi) G_n^{\pm}(x, \phi) = 0, \quad n \geq 0, \quad (13)$$

$G_n^{\pm}(x, \phi)$  are null functions for  $n \geq 0$ .

In the following, we will focus attention on three particular sets of null functions. The type-I null functions are obtained from Eq. (7) by choosing  $n=0$  and  $g(xe^{i\phi}) \equiv x^k e^{ik\phi}$  for a given  $k$ , and will be denoted by  $F_k^I(x, \phi)$ , namely,

$$F_k^I(x, \phi) = x^k e^{i(k+2)\phi}, \quad k=0,1,\dots \quad (14)$$

The type-II null functions correspond to the simple choice  $g(xe^{i\phi}) \equiv 1$ , i.e.,

$$F_n^{II}(\phi) = e^{i(1+n)2\phi}, \quad n=0,1,\dots \quad (15)$$

Finally, the type-III null functions are a kind of intermediate choice between type I and type II classes, and are defined as follows:

$$F_l^{III}(x, \phi) = x^{k[l]} e^{i(k[l]+2+2n[l])\phi}, \quad l=0,1,\dots, \quad (16)$$

where  $k[l]$  and  $n[l]$  are given in Table I. In the following we will use the notation  $F_k(x, \phi)$ , dropping the type indexes I–III, when the identity under consideration holds for all three types. Let us consider a generic real kernel  $R[\hat{O}](x, \phi)$ . By adding  $M$  null functions, keeping the kernel as real, we have a new kernel  $K[\hat{O}](x, \phi)$ :

TABLE I. Representation table for indices of type-III null functions.

$l$	0	1	2	3	4	5	6	7	...
$k+n$	0	1	1	2	2	2	3	3	...
$k$	0	1	0	2	1	0	3	2	...
$n$	0	0	1	0	1	2	0	1	...

$$K[\hat{O}](x, \phi) = R[\hat{O}](x, \phi) + \sum_{k=0}^{M-1} \mu_k F_k(x, \phi) + \sum_{k=0}^{M-1} \mu_k^* F_k^*(x, \phi), \quad (17)$$

where  $F_k(x, \phi) \in \mathcal{G}^+$ ,  $F_k^*(x, \phi) \in \mathcal{G}^-$ , and  $\mu_k$  are complex coefficients. By definition we have  $K[\hat{O}] = R[\hat{O}]$ , whereas the variance of the new kernel  $K[\hat{O}](x, \phi)$  is given by

$$\overline{\Delta K^2[\hat{O}]} = \overline{\Delta R^2[\hat{O}]} + 2 \left\{ \sum_{k,l=0}^{M-1} \mu_k \mu_l^* \overline{F_k F_l^*} + \sum_{k=0}^{M-1} \mu_k \overline{R[\hat{O}] F_k} + \sum_{k=0}^{M-1} \mu_k^* \overline{R[\hat{O}] F_k^*} \right\}. \quad (18)$$

In deriving the above formula we used the fact that both  $\mathcal{G}^+$  and  $\mathcal{G}^-$  are closed under multiplication.

The variance of the modified kernel function in Eq. (18) can be minimized with respect to the coefficients  $\mu_k$ , leading to the linear set of equations

$$\sum_l \mu_l \overline{F_k F_l^*} = -\overline{R[\hat{O}] F_k^*}. \quad (19)$$

It is convenient to rewrite the optimization equation (19) in matrix form as follows:

$$\mathbf{A}\boldsymbol{\mu} = \mathbf{b}, \quad (20)$$

where  $\mathbf{A}$  is the Hermitian  $M \times M$  matrix

$$A_{kl} = \overline{F_k F_l^*} = \int_0^{\pi} \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx p(x, \phi) F_k(x, \phi) F_l^*(x, \phi), \quad (21)$$

and  $\mathbf{b}$  is the complex vector

$$b_k = -\overline{R[\hat{O}] F_k^*} = -\int_0^{\pi} \frac{d\phi}{\pi} \int_{-\infty}^{\infty} dx p(x, \phi) R[\hat{O}](x, \phi) F_k^*(x, \phi). \quad (22)$$

Notice that the vector  $\mathbf{b}$  depends on both the kernel  $R[\hat{O}]$  and the state  $\hat{\rho}$  under examination, whereas the matrix  $\mathbf{A}$  depends on the state only.

By substituting Eq. (19) into Eq. (18) and inverting Eq. (20), we obtain

$$\begin{aligned} \Delta^2[\hat{O}] &\doteq \overline{\Delta R^2[\hat{O}]} - \overline{\Delta K^2[\hat{O}]} = 2 \sum_{k,l=0}^{M-1} \mu_k A_{kl} \mu_l^* \\ &= 2 \sum_{k,l=0}^{M-1} b_k (A^{-1})_{kl} b_l^* \geq 0, \end{aligned} \quad (23)$$

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

Let us summarize the optimization procedure for the kernel  $R[\hat{O}](x, \phi)$ . After collecting an ensemble of  $N$  tomographic data, the quantities  $\mathbf{A}$  and  $\mathbf{b}$  are evaluated as tomographic experimental averages. Then, by solving the linear system (20), one obtains the coefficients  $\mu_k$ , 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, the experimental error is evaluated, whose square now is reduced by the quantity  $\Delta^2[\hat{O}]/N$ .

The actual precision improvement of the tomographic measurement depends both on the state under examination (which affects both  $\mathbf{b}$  and  $\mathbf{A}$ ) and on the operator  $\hat{O}$ , whose kernel enters only in the expression of  $\mathbf{b}$ . An explicit expression for  $A_{kl}$  can be obtained by means of Eq. (10), and generally depends on the type of null function that are involved. For type-II null functions it reduces to the identity matrix, independently of the state

$$A_{kl}^{\text{II}} = \delta_{kl} \quad \text{type-II null functions}, \quad (24)$$

$\delta_{kl}$  denoting the Kronecker delta. For type-I null functions, one has

$$A_{kl}^{\text{I}} = \frac{(k+l)!}{2^{k+l}} \sum_{p=0}^{\min(k,l)} \frac{\langle a^{\dagger l-p} a^{k-p} \rangle}{2^p p! (l-p)! (k-p)!} \quad \text{type-I null functions}. \quad (25)$$

The explicit expression for coherent and Fock states is

$$A_{kl}^{\text{I}} = \alpha^{k-l} \frac{(k+l)!}{k!} 2^{-k-2l} L_l^{k-l}(-2|\alpha|^2) \quad \text{coherent state } |\alpha\rangle \quad (k \geq l), \quad (26)$$

$$A_{kl}^{\text{I}} = \delta_{kl} \frac{2^{k-n+1}}{n! \sqrt{\pi}} \int_0^{\infty} dy e^{-y^2} y^{2k} H_n^2(y) \quad \text{Fock state } |n\rangle, \quad (27)$$

where  $H_n(x)$  denotes Hermite polynomials. Notice that for Fock states the matrix is diagonal (which is true also for type-II and -III null functions).

### III. ADAPTIVE TOMOGRAPHY OF THE DENSITY MATRIX

In this section we apply the adaptive method to the tomographic measurement of the density matrix in the photon number representation. We evaluate the variance reduction  $\Delta^2[\hat{P}_{nm}]$  in Eq. (23) for  $\hat{P}_{nm}^R = \frac{1}{2}[|n\rangle\langle m| + |m\rangle\langle n|]$  and

$\hat{P}_{nm}^I = (1/2i)[|n\rangle\langle m| - |m\rangle\langle n|]$ , corresponding to the tomographic measurement of the real and imaginary part of the matrix elements  $\varrho_{nm} = \langle m|\hat{\varrho}|n\rangle$ . We consider different types of null functions, and calculate  $\Delta^2[\hat{P}_{nm}^R]$  and  $\Delta^2[\hat{P}_{nm}^I]$  versus the number  $M$  of added null functions, for either coherent states, squeezed vacuum, Fock states, and the ‘‘Schrödinger-cat’’ like superposition of coherent states given by

$$|\psi\rangle = \frac{1}{2\sqrt{1+\exp(-2|\alpha|^2)}}[|\alpha\rangle + |-\alpha\rangle]. \quad (28)$$

In order to see our adaptive method at work, Monte Carlo simulated experiments are presented.

Tomographic kernels for the matrix elements in the Fock basis were first presented in Ref. [12], with extension to non-unit quantum efficiency in Ref. [13], and factorization identities for the kernel in Ref. [27]. However, none of the above methods allows for an explicit analytical evaluation of the vector  $\mathbf{b}$  in Eq. (22). For this reason, we compute  $\Delta^2[\hat{P}_{nm}]$  numerically, presenting results in terms of the relative variance reduction  $\gamma$ , defined as follows:

$$\gamma = 1 - \frac{\overline{\Delta K^2[\hat{O}]}}{\overline{\Delta R^2[\hat{O}]}} = \frac{\Delta^2[\hat{O}]}{\overline{\Delta R^2[\hat{O}]}}. \quad (29)$$

A complete removal of fluctuations would correspond to  $\gamma=1$ .

### A. Coherent states

The adaptive method leads to a significant error reduction for the detection of matrix elements  $\langle m|\hat{\varrho}|n\rangle$  of coherent states. Our results indicate that type-I null functions are the most effective, and that the larger the amplitude  $\alpha$  of the coherent state, the larger the noise reduction. In Fig. 1 numerical results are presented for diagonal elements  $\langle n|\hat{\varrho}|n\rangle$  for intensity  $|\alpha|^2=5$ . In Fig. 1(a) the noise reduction  $\gamma$  is given versus the number  $M$  of added type-I null functions. One can see that the noise reduction  $\gamma$  saturates for large  $M$ , and better levels  $\gamma$  of reduction are achieved for smaller  $n$ . In Fig. 1(b) the noise reduction is reported versus  $n$  for  $M=30$ . In Fig. 2 we report the results from a Monte Carlo experiment for  $|\alpha|^2=3$ , with optimization performed with  $M=6$  null functions. The reduction of statistical errors for low values of  $n$  is evident.

The noise reduction for the off-diagonal matrix elements behaves similarly to the diagonal ones, being more effective for low indices. In Fig. 3 the noise reduction  $\gamma$  versus  $n$  and  $m$  of the (real) matrix element  $\langle m|\hat{\varrho}|n\rangle$  is plotted for a coherent state with a real amplitude  $\alpha=\sqrt{10}$ , and for the three types of null functions. The type-I null functions are generally more effective, though not uniformly over all indices  $n$  and  $m$ .

### B. Squeezed states and Schrödinger-cat states

Results for squeezed states and ‘‘cat’’ superposition of coherent states are presented in the same subsection, since they behave similarly. This is due to the fact that both states

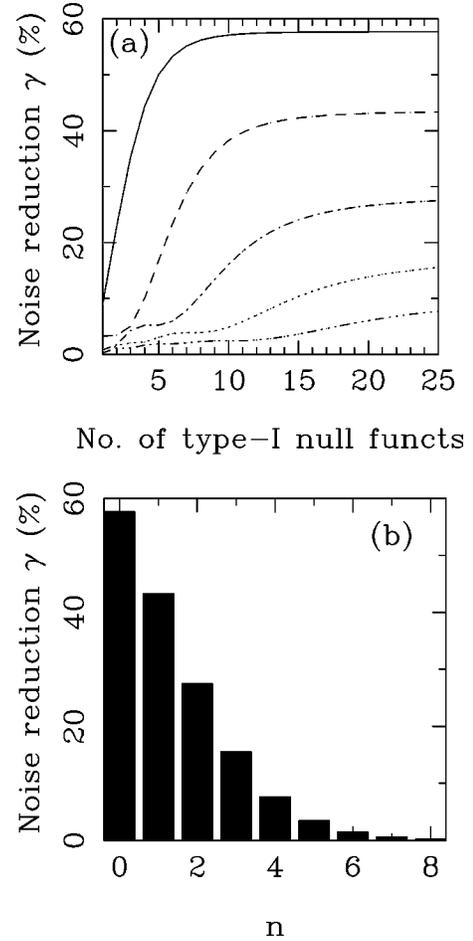


FIG. 1. Noise reduction in the tomographic measurement of the diagonal elements  $\langle n|\hat{\varrho}|n\rangle$  of the density matrix of a coherent state  $|\alpha\rangle$  with intensity  $|\alpha|^2=5$ . In (a), noise reduction  $\gamma$  vs the number of added type-I null functions: the full curve represents  $\langle 0|\varrho|0\rangle$ , the dashed curve  $\langle 1|\varrho|1\rangle$ , and so on, from the top to the bottom. In (b), noise reduction vs the index  $n$  of the diagonal matrix element for  $M=30$  added null functions.

have phase-dependent features reflected in a similar odd-even oscillation in the photon number probability distribution. In Figs. 4 and 5 the noise reduction for both cases is plotted for the three types of null functions, for  $M=10$ . From the plots it is apparent that type-II null functions are now the most effective ones, especially for off-diagonal matrix elements, though the same level of noise reduction for low  $n$  and  $m$  can also be obtained using type-I and -III null functions. In Fig. 6, results from a Monte Carlo simulated adaptive tomography on a squeezed vacuum are reported for  $\langle \hat{n} \rangle = 4$  and  $M=10$ . Matrix elements before and after optimization can be compared, showing the error reduction at work.

### C. Fock states

For Fock states the matrix  $\mathbf{A}$  is diagonal for all types of null functions, and therefore the optimization procedure just consists of the evaluation of the vector  $\mathbf{b}$ . The complex kernels for the matrix elements have the form  $R[|n\rangle\langle m|](x, \phi) = f_{n,m}(x) \exp(i(n-m)\phi)$ , where  $f_{n,m}(x)$

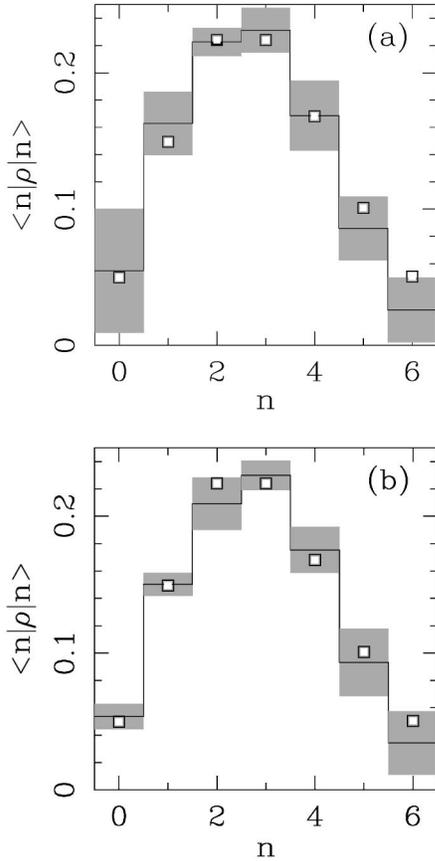


FIG. 2. Monte Carlo simulation of adaptive tomography of a coherent state with intensity  $|\alpha|^2 = 3$ . A sample of five blocks of 50 homodyne data is used for each of 25 phases (for a total number of measurements equal to 6250). The optimization has been performed by adding  $M = 6$  null functions. In (a) the measured diagonal matrix elements before optimization, and in (b) after optimization. The squares indicate theoretical values.

has the parity of  $n - m$  [2,27]. This fact, together with integral (11), makes it straightforward to show that

$$b_k^I \equiv b_k^{II} \equiv b_k^{III} \equiv 0, \quad \forall k; \quad (30)$$

that is, no improvement should be expected for the precision of quantum tomography on Fock states.

#### IV. ADAPTIVE TOMOGRAPHIC MEASUREMENTS OF OBSERVABLES

The tomographic estimation of the ensemble average  $\langle \hat{O} \rangle$  of a radiation operator  $\hat{O}$  can be obtained by averaging the kernel  $R[\hat{O}](x, \phi)$  given in Eq. (2). However, Eq. (2) needs a procedure that exploits the null function equivalence, and this is given in Ref. [28]. For this reason, for simplicity here we use the Richter formula [29], which expresses the kernels for the normally ordered moments as follows:

$$R[a^{\dagger n} a^m](x; \phi) = e^{i(m-n)\phi} \frac{H_{n+m}(\sqrt{2}x)}{\sqrt{2^{n+m}} \binom{n+m}{n}}, \quad (31)$$

$H_n(x)$  being the Hermite polynomial of order  $n$ . We apply the adaptive method to the tomographic detection of the most

relevant observables: intensity, quadrature, and complex field amplitude. The optimization method is here particularly useful, as the tomographic detection of these observables using the Richter kernel is very noisy [23,30].

In contrast to the case of matrix elements given in Sec. III, here some analytical evaluations can be carried out. We consider measurements performed on coherent states, squeezed vacuum states, Fock states, and cat superposition of coherent states. It turns out that the addition of just a few null functions to the Richter kernels generally results in a large improvement of the tomographic precision, again with the exception of Fock states where no improvement can be obtained.

#### A. Intensity

The tomographic detection of intensity is obtained by averaging the kernel

$$R[a^{\dagger} a](x) = 2x^2 - \frac{1}{2}. \quad (32)$$

The vectors  $\mathbf{b}$  needed for the optimization procedure are given by

$$b_k^I = -\overline{R[a^{\dagger} a] F_k^{I*}} = -\overline{2x^{k+2} e^{-i(k+2)\phi}} = -\frac{\langle a^{\dagger(k+2)} \rangle}{2^{1+k}}, \quad (33)$$

$$b_n^{II} = -\overline{R[a^{\dagger} a] F_n^{II*}} = -\overline{2x^2 e^{-i(n+1)2\phi}} = -\begin{cases} \frac{\langle a^{\dagger 2} \rangle}{2}, & n=0 \\ 0, & n \neq 0, \end{cases} \quad (34)$$

$$b_l^{III} = -\overline{R[a^{\dagger} a] F_l^{III*}} = -\overline{2x^{k[l]+2} e^{-i(k[l]+2+n[l])\phi}} = -\begin{cases} \frac{\langle a^{\dagger 2} \rangle}{2}, & l=0 \\ 0, & l \neq 0. \end{cases} \quad (35)$$

From Eqs. (34) and (35) it follows that only  $F_0^{II}(\phi) \equiv F_0^{III}(\phi) \equiv \exp(2i\phi)$  are effective in reducing the variance. We analytically solved the optimization equations (20) for type-I null functions, and also in this case it turns out that for all the states here considered, only the single null function  $F_0^I(\phi)$  is needed; that is, one has

$$\mu_0 = b_0 \quad \mu_k = 0, \quad \forall k \geq 1. \quad (36)$$

The corresponding reduction of variance is easily obtained from Eq. (23), and is given by

$$\Delta^2[a^{\dagger} a] = \frac{1}{2} \langle a^{\dagger 2} \rangle \langle a^2 \rangle. \quad (37)$$

Actually,  $\Delta^2[a^{\dagger} a]$  can compensate for the leading term of the variance of the original Richter kernel [23], which, in turn, is given by

$$\overline{\Delta R^2[a^{\dagger} a]} = \langle \widehat{\Delta n^2} \rangle + \frac{1}{2} [\langle a^{\dagger 2} a^2 \rangle + 2\langle a^{\dagger} a \rangle + 1]. \quad (38)$$

This means that the variance of the optimized kernel  $\overline{\Delta K^2[a^{\dagger} a]}$  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 compare the two noise ratios

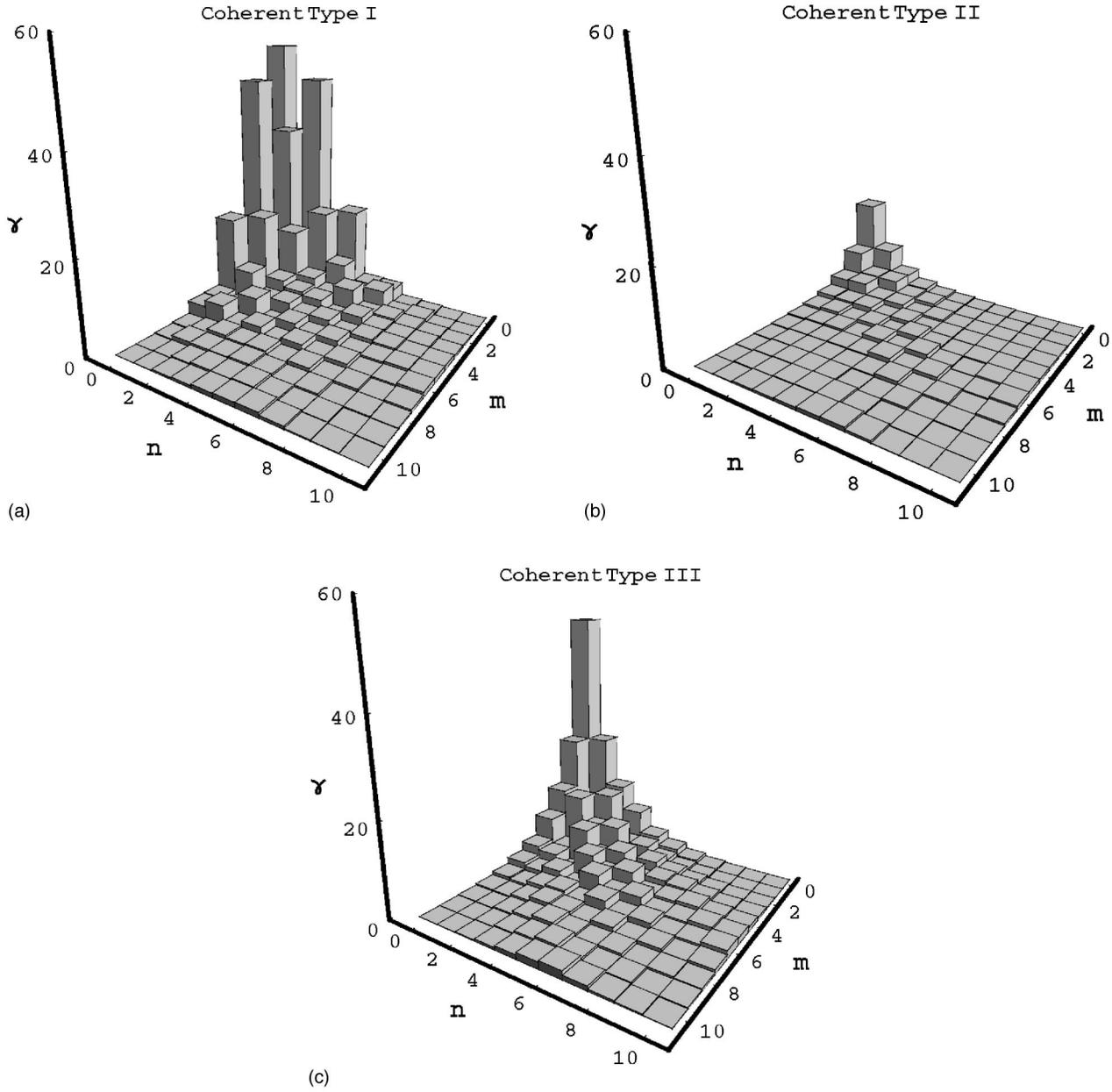


FIG. 3. Noise reduction  $\gamma$  vs indices  $n$  and  $m$  of the (real) matrix element  $\langle m|\hat{\rho}|n\rangle$  for a coherent state with intensity  $|\alpha|^2=5$ : (a) using only type-I null functions, (b) using only type-II null functions, and (c) using only type-III null functions. For all plots  $M=10$ , null functions have been used in the optimization procedure.

$$\delta n_R = \sqrt{\frac{\Delta R^2[a^\dagger a]}{\langle \Delta n^2 \rangle}}, \quad \delta n_K = \sqrt{\frac{\Delta K^2[a^\dagger a]}{\langle \Delta n^2 \rangle}}. \quad (39)$$

For coherent states  $|\alpha\rangle$ , we obtain

$$\delta n_R = \sqrt{2 + \frac{1}{2} \left( |\alpha|^2 + \frac{1}{|\alpha|^2} \right)}, \quad (40)$$

$$\delta n_K = \sqrt{2 + \frac{1}{2|\alpha|^2}};$$

that is, from an asymptotically linearly increasing function of  $|\alpha|$  the ratio becomes a constant  $\delta n_K \approx \sqrt{2}$ . Similar expres-

sions are obtained for other kinds of state: the noise ratio saturates at  $\delta n_K \approx \sqrt{3/2}$  for either squeezed vacuum or cat states.

In Fig. 7, results from a Monte Carlo simulation of the tomographic measurement of intensity on coherent states show the noise reduction obtained when using the optimized kernel. The noise reduction obtained by adding the single null function  $F_0(\phi)$  can also be easily evaluated for the generic diagonal moment  $\langle a^\dagger n a^n \rangle$ , using the formula

$$e^{i2\phi} R[a^\dagger n a^n](x) = \frac{n}{n+1} R[a^{\dagger(n+1)} a^{n-1}](x), \quad (41)$$

which leads to

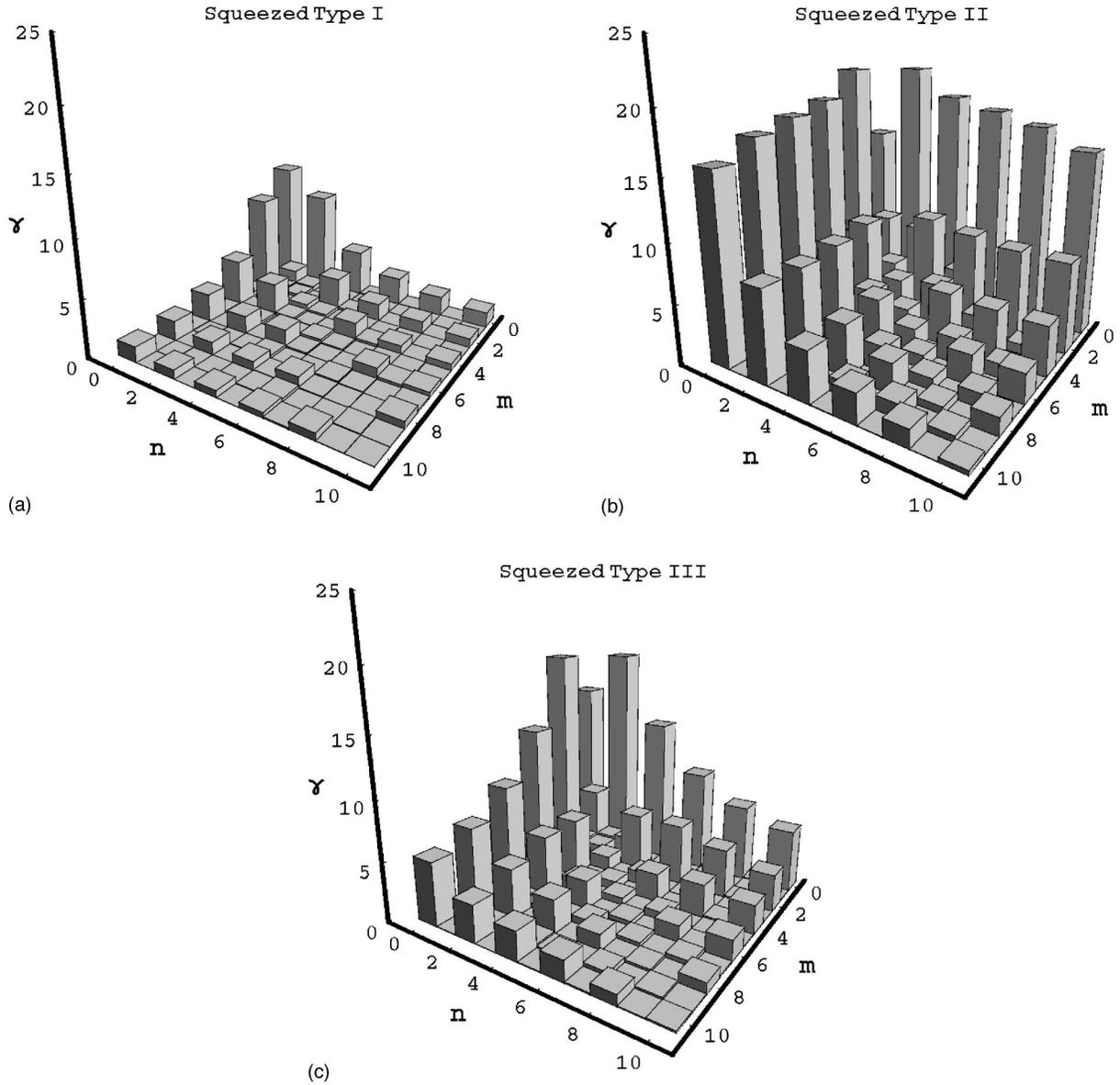


FIG. 4. Noise reduction for squeezed vacuum with  $\langle \hat{n} \rangle = 4$ : (a) using only type-I null functions, (b) using only type-II null functions, and (c) using only type-III null functions. For all plots  $M=10$ , null functions have been used in the optimization procedure.

$$b_0 = -\overline{R[a^\dagger n a^n] e^{i2\phi}} = -\frac{n}{n+1} \langle a^\dagger(n+1) a^{n-1} \rangle; \quad (42)$$

that is,  $\Delta^2[a^\dagger n a^n] = 2|b_0|^2$ . We just mention that optimizing the kernel  $R[a^\dagger a^2](x)$  is useful to improve detection of the second-order correlation function  $g^{(2)} = \langle a^\dagger a^2 \rangle / \langle a^\dagger a \rangle^2$ .

### B. Quadrature

The optimization procedure has also been tested on the kernel  $R[\hat{x}](x, \phi) = 2x \cos \phi$ , corresponding to the measurement of the quadrature operator  $\hat{x} = \frac{1}{2}(a + a^\dagger)$ . Similarly to the intensity case, the type-II and -III null functions do not play a role in improving precision, whereas type-I functions give  $b_k = -2^{-k-1} \langle a^\dagger(1+k) \rangle$  in Eq. (20). In this way the optimization procedure can also be carried out analytically in this case. The results indicate that for coherent states it is

enough to add the first null function  $F_0^I(\phi)$ , whereas for squeezed vacuum and cat states only the odd-index functions  $F_{2s+1}^I(x, \phi)$  contribute to noise reduction. In this case the main term is due to  $F_1^I(x, \phi)$ , whereas higher-order functions improve the variances only by a few percent. For coherent states the variance reduction from  $F_0^I(x, \phi)$  is given by

$$\Delta^2[\hat{x}] = \frac{1}{2} \langle a^\dagger \rangle \langle a \rangle = \frac{1}{2} |\alpha|^2, \quad (43)$$

which completely compensates for the leading term in the variance of the original Richter kernel [23]

$$\overline{\Delta R^2[\hat{x}]} = \langle \widehat{\Delta x^2} \rangle + \frac{1}{2} \langle a^\dagger a \rangle + \frac{1}{4}. \quad (44)$$

For squeezed vacuum and cat states the variance reduction due to  $F_1(x, \phi)$  is

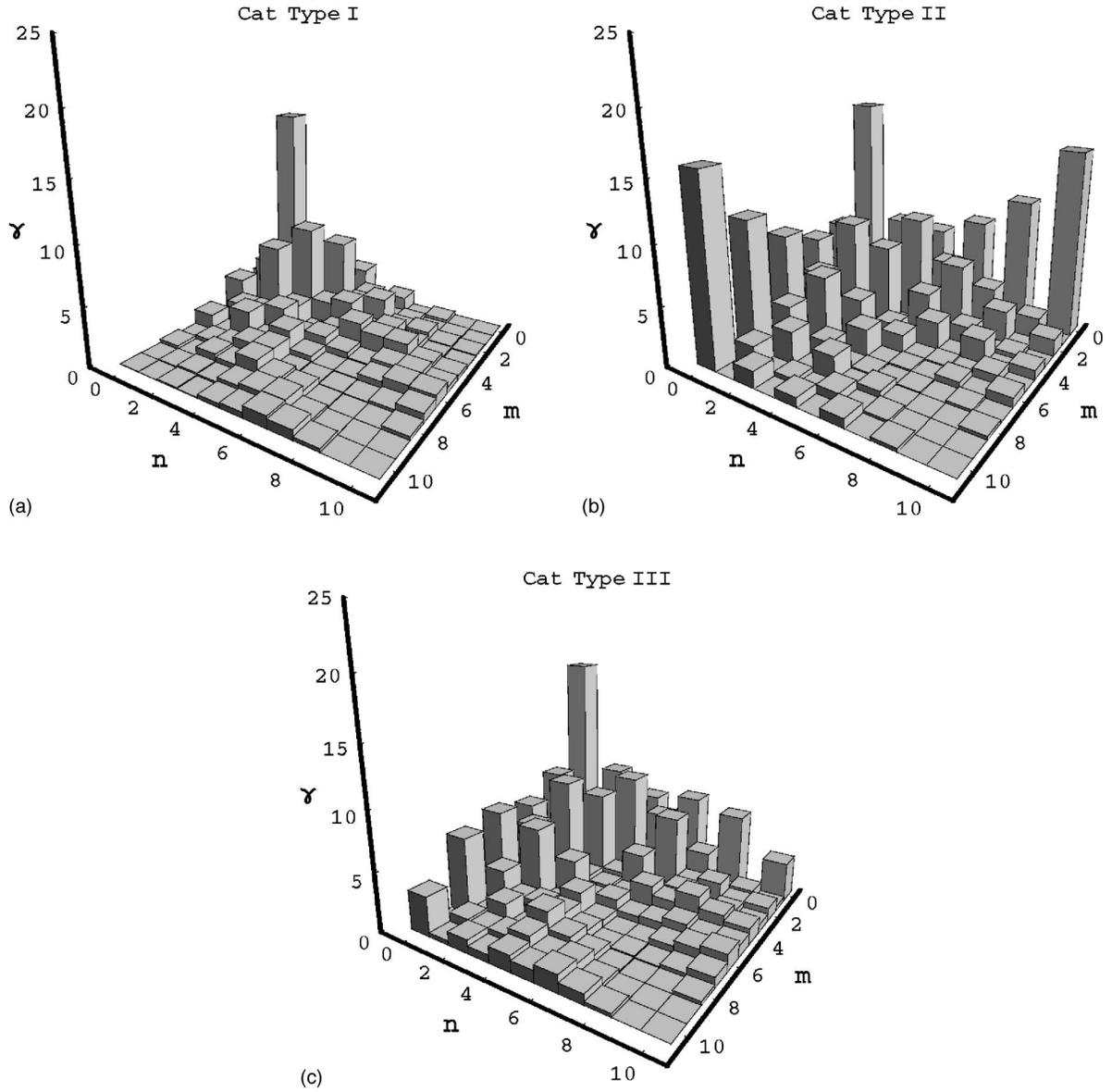


FIG. 5. Noise reduction for the catlike superposition of coherent states and for the three types of null functions in Eq. (28) with  $\alpha = \sqrt{3}$ : (a) using type-I null functions, (b) using type-II null functions, and (c) using type-III null functions. For all plots  $M=10$ , null functions have been used in the optimization procedure.

$$\Delta^2[\hat{x}] = \frac{1}{2(1 - |\langle a \rangle|^2 + 2\langle a^\dagger a \rangle)} \left[ |\langle a \rangle|^2 \left( \langle a^{\dagger 2} \rangle + \langle a^2 \rangle + \frac{1}{2} + \langle a^\dagger a \rangle \right) + |\langle a^2 \rangle|^2 \right]. \quad (45)$$

Upon defining the noise ratio  $\delta x_K$  in analogy to Eq. (39),

$$\delta x_K = \sqrt{\frac{\Delta K^2[a^\dagger a]}{\langle \Delta x^2 \rangle}}, \quad (46)$$

from Eqs. (43) and (45) we obtain the constant  $\delta x_K = \sqrt{2}$  for coherent states, independently of  $|\alpha|^2$ , whereas for squeezed vacuum and cat states the noise ratio saturates at  $\delta x_K \approx \sqrt{5/4}$ . In Fig. 8, results from simulated experiments of tomographic measurement of the quadrature on coherent

states are shown for  $|\alpha|^2 = 3$ . There the histograms of the original Richter kernel and of the optimized kernel are compared. The optimized kernel has a sharper distribution, which is peaked at the mean value  $\langle \hat{x} \rangle = \sqrt{3}$ . For this reason, it is quite obvious that the optimized kernel  $K[\hat{x}](x, \phi)$  gives a more precise determination of  $\langle \hat{x} \rangle$  than the original kernel  $R[\hat{x}](x, \phi)$ .

### C. Field amplitude

The tomographic kernel for the measurement of the complex field amplitude  $a$  is given by  $R[a](x, \phi) = 2x e^{i\phi}$ , and

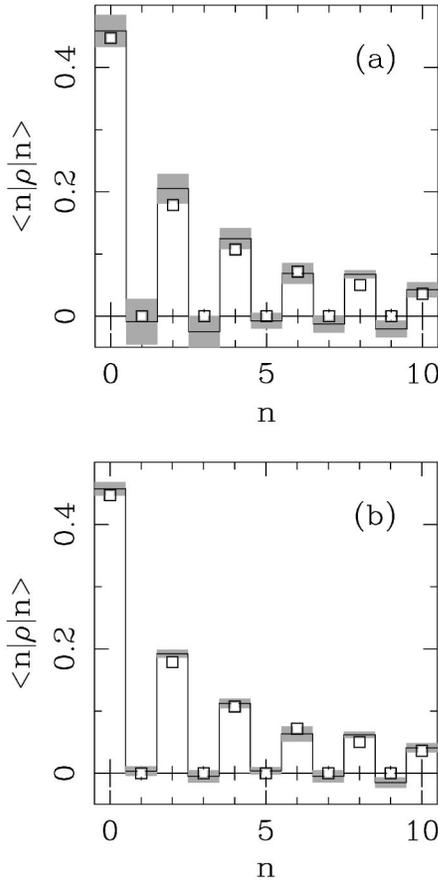


FIG. 6. Adaptive tomography of a squeezed vacuum with  $\langle \hat{n} \rangle = 4$ . The Monte Carlo sample includes five blocks of 100 data for each of 50 phases (for a total number of measurements equal to 25 000). The optimization has been performed by adding  $M=10$  type-II null functions. (a) Measured elements without optimization, and (b) with optimization. The squares indicate the theoretical values.

its fluctuations should be compared with those from the ideal measurement of  $a$ , which could be achieved by ideal eight-port [32–34] or six-port [35,36] homodyne detection. The optimization procedure depends on the choice for the definition of statistical error for a complex quantity. If one considers the real or imaginary part separately, the procedure coincides with the optimization of the precision in independent measurements of two conjugated quadratures. On the other hand, in order to take both noises into account jointly, we minimize the quantity

$$\overline{\Delta_* K^2[a]} = \frac{1}{2} \overline{\{|K[a]|^2 - |K^2[a]|^2\}} \quad (47)$$

corresponding to the average of noises for real and imaginary parts, namely, the trace of the noise covariance matrix. Now the equivalence class of kernel functions is written as

$$K[a](x, \phi) = R[a](x, \phi) + \sum_{p=0}^{M-1} \mu_p F_p(x, \phi) + \sum_{p=0}^{M-1} \nu_p F_p^*(x, \phi). \quad (48)$$

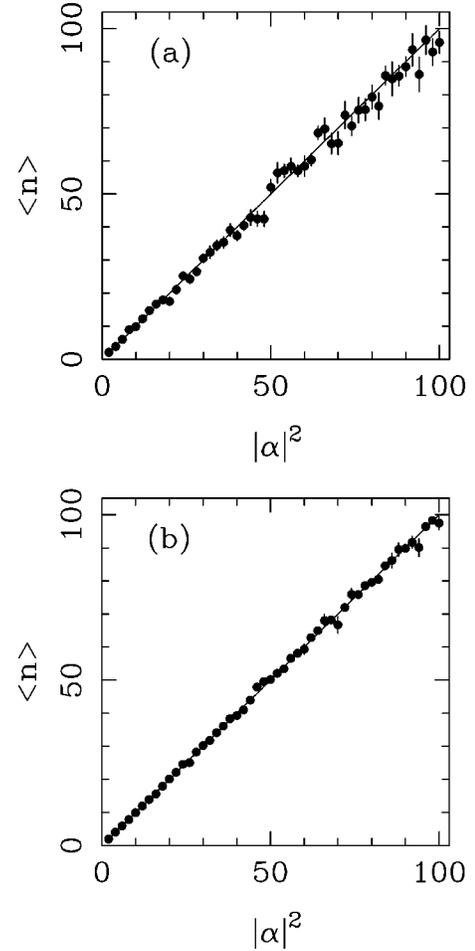


FIG. 7. Tomographic detection of the intensity on coherent states. The simulated experiment has been performed with 15 blocks of 15 data for 15 phases each (for a total number of  $N = 3375$  measurements). The tomographic result  $\langle \hat{n} \rangle$  is reported vs the theoretical values  $|\alpha|^2$ , (a) without and (b) with optimization.

$\mu_p$  and  $\nu_p$  are two independent sets of complex coefficients. The optimization procedure is similar to the real case, and is reduced to solving the two linear systems

$$\mathbf{A}\boldsymbol{\mu} = \mathbf{b}, \quad \mathbf{A}\boldsymbol{\nu} = \mathbf{c}, \quad (49)$$

where  $\mathbf{c}$  is given by

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

By inverting Eqs. (49), one obtains the noise reduction

$$\begin{aligned} \Delta_*^2[a] &= \overline{\Delta_* R^2[a]} - \overline{\Delta_* K^2[a]} \\ &= \sum_{p,q=0}^{M-1} [b_p (A^{-1})_{qp} b_q^* + c_p (A^{-1})_{pq} c_q^*]. \end{aligned} \quad (50)$$

Also in the present case it is sufficient to consider only type-I functions. The optimization vector  $\mathbf{b}$  is given by  $b_k = -2^{-k} \langle a^{\dagger(1+k)} \rangle$ . Similarly to the case of the quadrature, the optimization procedure shows that for coherent states only  $F_0^I(\phi)$  is needed, whereas for squeezed vacuum and cat states only the odd-index functions  $F_{2s+1}^I(x, \phi)$  contribute to

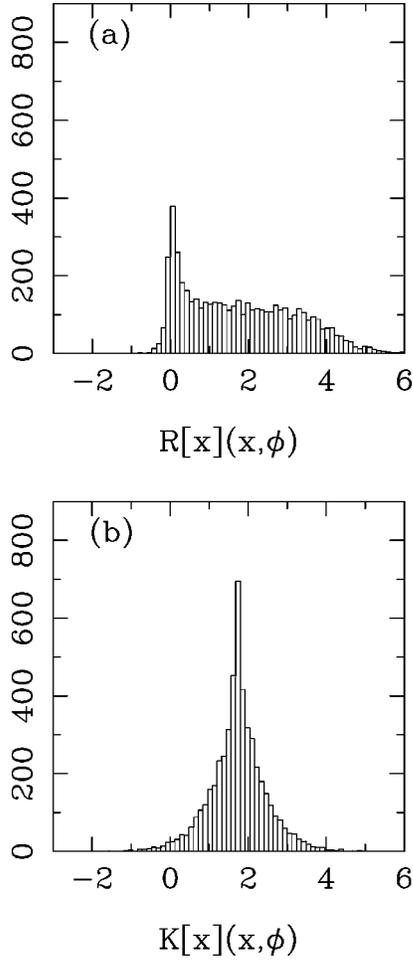


FIG. 8. Histograms of the kernel functions evaluated on the tomographic outcomes for a coherent state with  $|\alpha|^2=3$ . The sample has 50 phases with 100 data each. (a) Using the original Richter kernel  $R[\hat{x}](x, \phi)$ . (b) Using the optimized kernel  $K[\hat{x}](x, \phi)$ . The distribution for the optimized kernel is sharper and peaked near the theoretical value  $\langle \hat{x} \rangle = \sqrt{3}$ .

noise reduction, and the main term comes from  $F_1^I(x, \phi)$ . In this way, for coherent states one obtains

$$\Delta_*^2[a] = \frac{1}{2}|\alpha|^2, \quad (51)$$

whereas for squeezed vacuum and cat states one has

$$\Delta_*^2[a] = \frac{1}{2(1 - |\langle a \rangle|^2 + 2\langle a^\dagger a \rangle)} \times \left[ |\langle a \rangle|^2 \left( \langle a^{\dagger 2} \rangle + \langle a^2 \rangle + \frac{1}{2} + \langle a^\dagger a \rangle \right) + |\langle a^2 \rangle|^2 \right]. \quad (52)$$

Equations (51) and (52) should be compared with the noise figure of the original Richter kernel

$$\overline{\Delta_*^2 R[a]} = \frac{1}{2} [2\langle a^\dagger a \rangle + 1 - |\langle a \rangle|^2], \quad (53)$$

and with the intrinsic noise of a generalized measurement of the amplitude

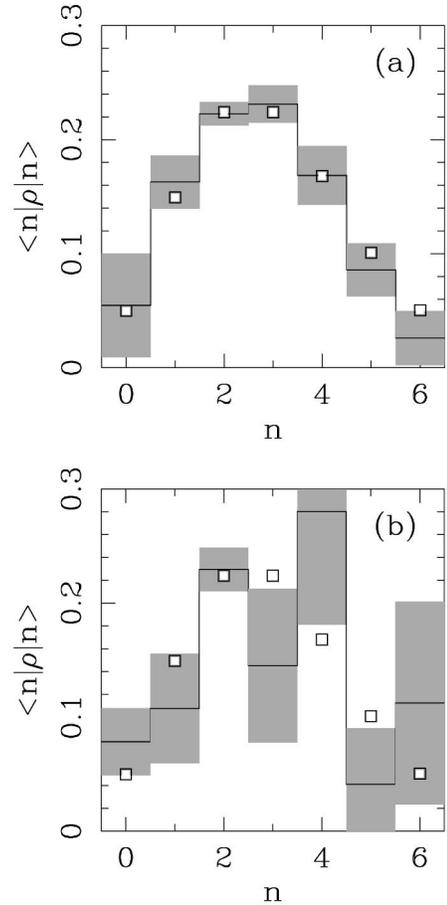


FIG. 9. Monte Carlo simulation of adaptive tomography with a bad choice of the number of added null functions. The state under examination is a coherent state with  $|\alpha|^2=3$ , and the simulated sample of homodyne data contains five blocks of 50 data for 25 phases each, for a total number of 6250 measurements (as in Fig. 2). Here the optimization has been performed by adding  $M=32$  null functions. Large fluctuations emerge instead of error reduction. The squares indicate theoretical values.

$$\widehat{\langle \Delta_*^2 a^2 \rangle} = \frac{1}{2} [\langle a^\dagger a \rangle + 1 - |\langle a \rangle|^2]. \quad (54)$$

The noise ratio thus equals  $\delta a_K=1$  for coherent states, whereas it saturates at  $\delta a_K \approx \sqrt{3}/2$  for both squeezed vacuum and cat states. Remarkably, for coherent states heterodyne noise is reached, that is, the tomographic detection has ideal noise.

## V. EFFECTS OF SYSTEMATIC ERRORS

Throughout this paper tomographic kernels have been optimized by adding low-order null functions. Higher-order functions oscillate more rapidly. Since the method involves only the average of these functions on a small sample of data, fast oscillations in  $\phi$  and a higher power of  $x$  would introduce more noise, and including too many null functions would increase the error instead of reducing it. In Fig. 9 an example of such pathology is given.

Another point that should be mentioned is that in the tomographic detection here considered the phase  $\phi$  is a random parameter in  $[0, \pi]$ . A discrete scanning by equally spaced

phases would introduce systematic errors [21,31] that would mask the benefits from the optimization. Actually, for non-random uniform scanning, the null function  $F_0(\phi)$  has no effects when added to phase-independent kernels, whereas the other null functions have a considerably reduced effect, and obviously do not eliminate the systematic error due to the finite mesh of the deterministic scanning.

## VI. SUMMARY AND CONCLUSIONS

In this paper we have presented an adaptive method to optimize tomographic kernels, improving the precision of the tomographic measurement. The method has been analyzed in detail for coherent states, Fock states, squeezed vacuum, and ‘‘Schrödinger-cat’’ states. With the exception of Fock states, the method generally provides a sizable reduction of statistical errors. For coherent states the improvement mainly concerns the small-index matrix elements, whereas for squeezed

vacuum and cat states far off-diagonal elements are also improved.

The error reduction is much more significant for the measurement of intensity, quadrature and field amplitude, where for coherent states, squeezed vacuum, and cat states the ratio between the tomographic noise and uncertainty of the considered observable saturates for increasing energy. In this case, we can definitely assert that quantum tomography is a quasi-ideal measurement, as it adds only a small amount of noise as compared to ideal detection.

## ACKNOWLEDGMENTS

We would thank Dirk G. Welsch, Mohamed Dakna, and Nicoletta Sterpi for useful discussions. M.G.A.P. was partially supported by the ‘‘Francesco Somaini’’ Foundation. This work was part of INFM Contract No. PRA-97-CAT, and of MURST Contract No. COFINANZIAMENTO-97.

- 
- [1] J. Mod. Opt. **44** (11-12) (1997), special issue on quantum state preparation and measurement, edited by M. G. Raymer and W. Schleich.
  - [2] For a review, see G. M. D’Ariano, in *Quantum Optics and Spectroscopy of Solids*, edited by A. S. Shumowsky and T. Hakiouglu (Kluwer, Dordrecht, 1997), p. 175.
  - [3] For a more recent review, see D.-G. Welsch, W. Vogel, and T. Opatrny, Prog. Opt. (to be published).
  - [4] D. F. McAlister and M. G. Raymer, Phys. Rev. A **55**, R1609 (1997).
  - [5] G. Breitenbach, S. Schiller, and J. Mlynek, Nature (London) **387**, 471 (1997).
  - [6] M. Vasylyev, S.-K. Choi, P. Kumar, and G. M. D’Ariano, Opt. Lett. **23**, 1393 (1998).
  - [7] G. M. D’Ariano, M. Vasylyev, and P. Kumar, Phys. Rev. A **58**, 636 (1998).
  - [8] G. M. D’Ariano, M. F. Sacchi, and P. Kumar, Phys. Rev. A **59**, 826 (1999).
  - [9] G. M. D’Ariano and L. Maccone, Phys. Rev. Lett. **80**, 5465 (1998).
  - [10] K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989).
  - [11] M. Beck, D. T. Smithey, and M. G. Raymer, Phys. Rev. A **48**, R890 (1993).
  - [12] G. M. D’Ariano, C. Macchiavello, and M. G. A. Paris, Phys. Rev. A **50**, 4298 (1994).
  - [13] G. M. D’Ariano, U. Leonhardt, and H. Paul, Phys. Rev. A **52**, 1801 (1995).
  - [14] K. Banaszek and K. Wódkiewicz, Phys. Rev. Lett. **76**, 4344 (1996).
  - [15] V. Bužek, G. Adam, and G. Drobný, Ann. Phys. (N.Y.) **245**, 37 (1996).
  - [16] M. G. A. Paris, Phys. Rev. A **53**, 2658 (1996).
  - [17] H. Kühn, D.-G. Welsch, and W. Vogel, J. Mod. Opt. **41**, 1607 (1998).
  - [18] G. M. D’Ariano, in *Quantum Communication, Computing, and Measurement*, edited by O. Hirota, A. S. Holevo, and C. M. Caves (Plenum, New York, 1997), p. 253.
  - [19] T. Opatrny and D.-G. Welsch, Phys. Rev. A **55**, 1462 (1997).
  - [20] G. M. D’Ariano and C. Macchiavello, Phys. Rev. A **57**, 3131 (1998).
  - [21] G. M. D’Ariano, C. Macchiavello, and N. Sterpi, Quantum Semiclass. Opt. **9**, 929 (1997).
  - [22] G. M. D’Ariano, C. Macchiavello, and N. Sterpi, in *Quantum Communication, Computing, and Measurement*, edited by G. M. D’Ariano, P. Kumar, and O. Hirota (Plenum, New York, in press).
  - [23] G. M. D’Ariano and M. G. A. Paris, Phys. Lett. A **233**, 49 (1997).
  - [24] T. Opatrny, M. Dakna, and D.-G. Welsch, Phys. Rev. A **57**, 2129 (1998).
  - [25] G. M. D’Ariano and M. G. A. Paris, Acta Phys. Slov. **48**, 191 (1998).
  - [26] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).
  - [27] U. Leonhardt, M. Munroe, T. Kiss, Th. Richter, and M. G. Raymer, Opt. Commun. **127**, 144 (1996).
  - [28] G. M. D’Ariano, in *Quantum Communication, Computing, and Measurement* (Ref. [22]).
  - [29] Th. Richter, Phys. Lett. A **221**, 327 (1996).
  - [30] G. M. D’Ariano and M. G. A. Paris, Acta Phys. Slov. **47**, 281 (1997).
  - [31] U. Leonhardt and M. Munroe, Phys. Rev. A **54**, 3682 (1996).
  - [32] N. G. Walker, J. Mod. Opt. **34**, 15 (1987).
  - [33] Y. Lay and H. A. Haus, Quantum Opt. **1**, 99 (1989).
  - [34] G. M. D’Ariano and M. G. A. Paris, Phys. Rev. A **49**, 3022 (1994).
  - [35] A. Zucchetti, W. Vogel, and D.-G. Welsch, Phys. Rev. A **54**, 856 (1996).
  - [36] M. G. A. Paris, A. Chizhov, and O. Steuernagel, Opt. Commun. **134**, 117 (1997).