MEASURING QUANTUM STATES

G. M. D'ARIANO
Dipartimento di Fisica "A. Volta",
via Bassi 6, I-27100 Pavia, Italy

1. Introduction

The possibility of "measuring" quantum states has been already considered by W. Pauli [1], who wrote in a footnote: "The mathematical problem, as to whether for given functions $W(\mathbf{x})$ and $W(\mathbf{p})$, the wave function ψ, if such function exists, is always uniquely determined, has still not been investigated in all its generality" [by $W(\mathbf{x})$ and $W(\mathbf{p})$ Pauli denoted the probability distributions $W(\mathbf{x}) = |\langle \mathbf{x}|\psi\rangle|^2$ and $W(\mathbf{p}) = |\langle \mathbf{x}|\psi\rangle|^2$, whereas, for "if such function exists" he meant that $W(\mathbf{x})$ and $W(\mathbf{p})$ are compatible]. The answer to Pauli's question is clearly negative [2]. In fact, it is quite obvious that the probabilities $W(\mathbf{x})$ and $W(\mathbf{p})$ at least cannot determine the quantum correlation $\langle\psi|\mathbf{xp}|\psi\rangle$, which can be obtained, for example, from a *joint* measurement of $\mathbf{x}$ and $\mathbf{p}$. On the other hand, as shown in my lectures on "Quantum estimation theory and optical detection" [3], a joint measurement of two conjugated observables would exhibit an additional noise equivalent to an effective quantum efficiency $\eta = 1/2$: as we will see in Sect. 7.5, this is exactly the threshold below which the density matrix cannot be measured.

That more than two observables—actually a complete set of them—would be needed in order to determine the density matrix was clear from the earlier theoretical studies of Fano [4] and remarked in the book of d'Espagnat [5]. However, it is difficult to devise concretely measurable translational observables—other than position, momentum and energy: for this reason a Schrödinger-picture point of view was adopted, in which, instead of measuring varying operators, the state itself is changed in a controlled way, and eventually the energy is measured [6] [for "spin" degrees of freedom a first study has been published only very recently [7]]. Still, such a fundamental problem—measuring the quantum state—has remained at the level of mere speculation for many years [8], and entered the realm of experiments only less than three years ago, after the experiments by Raymer's group [9], in the domain of quantum optics.

What is so special with quantum optics? In quantum optics, differently from quantum mechanics of particles, there is the unique opportunity of measuring all possible linear combinations of position $\hat{q}$ and momentum $\hat{p}$ of a harmonic oscillator, here a mode of the electromagnetic field. As explained in my previous set of lectures [3], such measurement can be achieved by means of a balanced homodyne detector, which measures the quadrature $\hat{x}_\phi = \frac{1}{2}\left(a^\dagger e^{i\phi} + a e^{-i\phi}\right)$ of the field at any desired phase ϕ with respect to the local oscillator (LO) [as usual a denotes the annihilator of the field mode]. The first technique

T. Hakioğlu and A. S. Shumovsky (eds.), Quantum Optics and the Spectroscopy of Solids, 175–202.
© 1997 *Kluwer Academic Publishers. Printed in the Netherlands.*

to reconstruct the density matrix from homodyne measurements—so called *homodyne-tomography*—originated from the idea of Vogel and Risken [10] that the collection of probability distributions $\{p(x,\phi)\}_{\phi\in[0,\pi)}$ is just the Radon transform (or "tomography") of the Wigner function $W(\alpha,\overline{\alpha})$. Then, by inverting the Radon transform, one obtains the Wigner function, and from the knowledge of $W(\alpha,\overline{\alpha})$ one can recover the matrix elements of the density operator $\hat{\varrho}$. In short, this was the basis of the pioneer method used by Raymer's group.

The first experimental work [9]—which unequivocally demonstrated the feasibility of the method—was, however, affected by some uncontrollable approximations that produce serious systematic errors when the detected state $\hat{\varrho}$ is truly "nonclassical". In fact, as we will see in Sect. 5, strictly speaking the Wigner function cannot be measured, because the inversion of the Radon transform needs an *analytic knowledge* of the probability distributions $p(x,\phi)$. In practice, the inverse Radon transform is achieved by "filtering" data—as in the usual X-ray medical tomography—upon introducing a suitable cut-off parameter that sets a coarse resolution on $W(\alpha,\overline{\alpha})$. Such a cut-off does not correspond to any controllable approximation in terms of density matrix. Setting such resolution in advance makes the measured state "more classical", thus loosing the most interesting quantum features that the experimentalist is looking for! Hence, the density matrix cannot be measured by measuring $W(\alpha,\overline{\alpha})$ as an intermediate step. However, the main idea—namely, homodyning the density matrix—still remains valid: the problem is only related to the method of data-processing, i. e. the use of the inverse Radon transform. With Macchiavello and Paris [11] I derived an exact technique that produces the number-state matrix elements only by averaging functions of data. The rather involved original formulas were simplified later by myself, Leonhardt, and Paul [12]. In the new form, the algorithm was so simple that it was possible also to recognize the feasibility of homodyning the density matrix even for nonideal detector quantum efficiency $\eta < 1$, giving mathematically proved lower bounds for η that depend on the chosen representation for the matrix. After these first results, further theoretical progress has been made, understanding the mechanisms that underly the generation of statistical errors [13] and limit the sensitivity of the method when used for indirectly measuring a single generic observable [14]. The problem of determining the Wigner function was revisited [15], and it was lately recognized [16] that $W(\alpha,\overline{\alpha})$ can be measured at some fixed resolution in polar coordinates, and that fixing such resolution corresponds to truncating the dimension of the density matrix in the Hilbert space. It was also realized that the new method can be profitably used as a novel imaging algorithm—what I call *the quantum Radon transform*—for conventional tomography [16], particularly useful in the extreme situation of low signals and high experimental resolution. Recently, new progress by Richter [17] and Leonhardt *et al.* [18] have made possible to factorize the averaged functions (socalled "pattern functions" [19, 20]) speeding up the algorithm greatly for $\eta = 1$ [an analogous factorization for $\eta < 1$ is still lacking: very recently the factorization has been derived by algebraic methods [21], and we hope that this will eventually provide the ultimate technique]. Currently, the algorithm is so simple, fast, and low-memory, that I believe that after reading these notes the reader will certainly try to implement it on his PC! In the meanwhile there has been an explosion of interest on the subject of *measuring quantum states* from both theoreticians and experimentalists (I apologize for not quoting all the literature): the new exact method has been eventually adopted for real experiments [22, 23], providing a very efficient way for measuring the photon number distribution.

I will review the exact method for homodyning the state of radiation in Sections 6

up to 9, after a premise in Sect. 2 on the central limit theorem, which is at the basis of the method and is needed in order to understand what can be measured and what cannot (for example, the Wigner function cannot be measured: see Sect. 5). A short review on conventional tomographic imaging is given in Sect. 3, and the old homodyne-tomographic method for measuring the state is revisited in Sect. 4. The possibility of using the new method for conventional imaging and the coding of images into density matrices is discussed in Sect. 11. In Sect. 10 a recent debate on the possibility of measuring the density matrix of a *single* quantum system is reviewed. In fact, the notion of "measuring a state" is meaningful only if it refers to many repeated measurements performed on an *ensemble of equally prepared* systems: otherwise, the state of an individual system is unobservable, whatever sequence of measurements is performed on the system.

Appendices contain some relevant proofs and derivations, along with the recently derived factorization formulas. Notice: some parts of the present notes include material that entered my knowledge (or was discovered) after the end of the set of lectures given at Bilkent University. As the field is so rapidly growing, I decided to include this material here in order to make these notes more complete and up to date. In particular the new matter regards the factorization formulas (Refs. [17], [18], and [21]) , the new experiments in Refs. [22] and [23], and finally my recent work with Yuen on the impossibility of measuring the density matrix of a single quantum system [24].

2. Premise on the central limit theorem

The exact method for homodyning the density matrix of the radiation field is based on the fact that the matrix element $\langle\psi|\hat{\varrho}|\varphi\rangle$ can be written as the expectation of a function $f_{\psi\varphi}(I,\phi)$ of the random homodyne photocurrent I at random phase ϕ relative to the LO. The analytic form of the function $f_{\psi\varphi}$ is known in advance, and depends parametrically on both vectors ψ and φ of the matrix element: it will be explicitly given in Sect. 7 for some relevant representations. Hence, the problem of measuring a matrix element resorts to the estimation of a *generalized moment* of the homodyne probability $p(I,\phi)$, or, in other words, to the estimation of what I call the "mathematical expectation" of a given function of I and ϕ. To make things easier, let us consider the case that the random variable is scalar real-valued, denoted by x. Then, the *generalized f-moment* or *mathematical expectation* $\mathbf{E}f(x)$ of the function $f(x)$ is defined as follows

$$F = \mathbf{E}f(x) \equiv \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x) f(x)\,. \tag{1}$$

What we can say in advance on the possibility of "measuring" $\mathbf{E}f(x)$ experimentally? We would be tempted of measuring F by means of the "experimental average"

$$F_N = \frac{1}{N}\sum_{n=1}^{N} f(x_n)\,, \tag{2}$$

where $\{x_n\}_{n=1}^{N}$ denote the outcomes of x in an experiment that consists of N repeated measurements. Is this procedure correct? How reliable is such estimate F_N of F? The answer to these questions are provided by the *central limit theorem* (the proof is reported in Appendix 13.1. According to this theorem, for functions $f(x)$ having moments bounded up to the third order, the random variable F_N will exhibit Gaussian distribution asymptotically for large N, with average F and variance $\varepsilon_N^2 = \mathbf{E}[\Delta f^2]/N$. In such case, we can

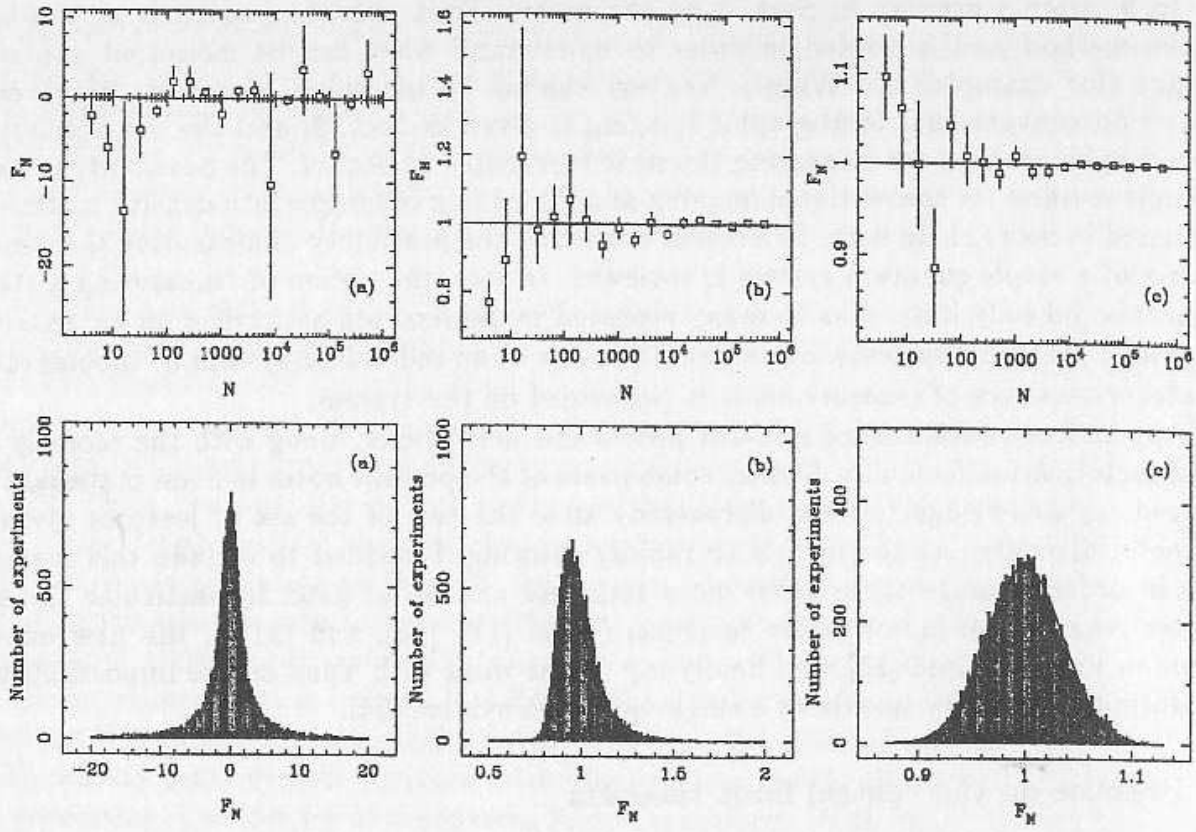

Figure 1. Numerical simulation of experiments for estimating $F = \mathrm{E}f(x)$ in Eq. (1) by means of F_N in Eq. (2). Here, we have uniform probability $p(x) = 1/2$ for $x \in [-1, 1]$, $p(x) = 0$ otherwise. Top figures give F_N within the confidence interval δ_N of Eq. (3) versus N. (a) $f(x) = \frac{1}{x}$: the function is not integrable; (b) $f(x) = \frac{1}{2|x|^{1/2}}$: the function is not square-integrable; (c) $f(x) = \frac{3}{2}|x|^{1/2}$: the function is square-integrable. The three histograms on the bottom give the distribution of F_N for different experiments; every histogram pertains the corresponding case on the top. There are N=20000 measurement for each experiment. For the case (*c*) the histogram is compared with a Gaussian with the same average and variance.

safely rely on the estimate F_N of F, knowing *a priori* that another similar experiment will produce a result within the *confidence interval* $[F_N - \varepsilon_N, F_N + \varepsilon_N]$ with probability 68%. An estimate of the confidence interval ε_N itself is provided by the well known root mean square deviation

$$\delta_N = \sqrt{\frac{\sum_{n=1}^{N}[f(x_n) - F_N]^2}{N(N-1)}}\,. \tag{3}$$

What happens when the conditions of the central limit theorem are not fulfilled? For example, if the function f is not integrable with respect to the probability measure $p(x)$, the estimate F_N will not approach any definite value for large N. On the other hand, if f is integrable but not square-integrable (i.e. it has unbounded second moment), then δ_N will not rescale correctly as $N^{-1/2}$, and the confidence interval will be utterly unreliable. We can gain some feeling on such pathological situations with the aid of a computer experiment. In the standard Fortran language the random number generator "ran" produces uniform deviates between 0 and 1: by means of translation and rescaling one can generate uniform deviates in any prechosen interval. Let us consider, for example, a uniform probability density: $p(x) = 1/2$ for $x \in [-1, 1]$, $p(x) = 0$ otherwise. In Fig. 1 the result from numerical

simulations of the estimate F_N in Eq. (2) with error-bar δ_N in Eq. (3) are given for three kinds of functions: (a) non integrable $f(x) = \frac{1}{x}$; (b) non square-integrable $f(x) = \frac{1}{2|x|^{1/2}}$; (c) square-integrable $f(x) = \frac{3}{2}|x|^{1/2}$. The behavior of F_N versus N is erratic in the non integrable case, whereas it converges nicely to the expected value F in the other two cases.

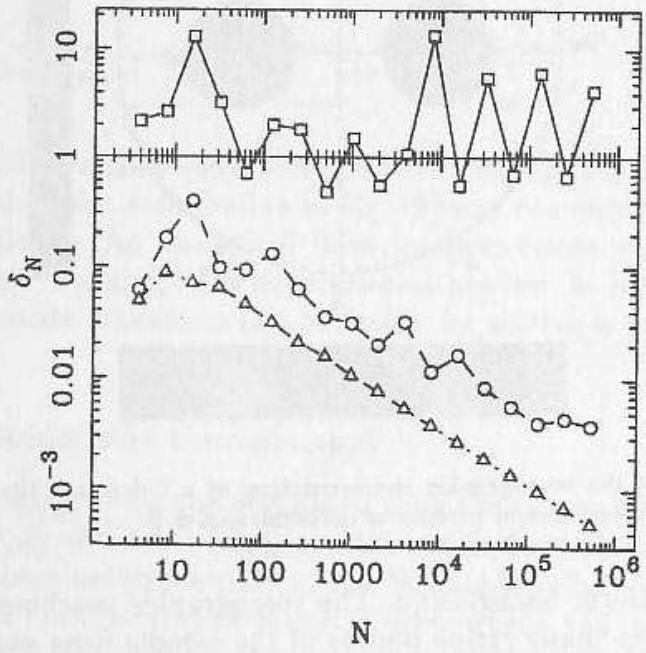

Figure 2. Behavior of the confidence interval δ_N in Eq. (3) versus the number of measurements N per experiment. The squares refer to the case (a) in Fig. 1 [non integrable function], the circles to case (b) [non square-integrable function] and the triangles to case (c) [square-integrable function].

On the other hand, from the same Fig. 1 it is evident that the histogram of the values of F_N from different experiments is Gaussian only in the square-integrable case, and from Fig. 2 we can see that the this is the only case leading to a $\delta_N \sim N^{-1/2}$.

What we can say if the probability $p(x)$ is completely unknown *a priori*—as it is the case in practice? Then, in order to assure in advance that $f(x)$ satisfies the conditions imposed by the central limit theorem, the function $f(x)$ must be square-integrable for any possible probability measure $p(x)$. It is easy to see that this requirement is fulfilled only if $f(x)$ is bounded by some constant $|f(x)| < C < \infty$. In fact, if $f(x)$ is unbounded, then it is always possible to find a probability measure $p(x)$ for which $f(x)$ is not even integrable. On the other hand, if $f(x)$ is bounded, then every moment of f is bounded—in particular the first three ones—and the central limit theorem holds. Thus we conclude that for all possible (a priori unknown) probabilities $p(x)$, the generalized f-moment in Eq. (1) can be estimated by means of the experimental average (2) if and only if the function $f(x)$ is bounded. In such case we will also say that F is *statistically sampled* by F_N, and Eqs. (1) and (2) will be referred to as *sampling formulas* for F.

3. What is a tomography?

The essential problem of tomographic imaging is to recover a distribution of mass $m(x, y)$ in a 2-d slab from a finite collection of one dimensional projections $r(s, \phi)$ at different angles ϕ. The situation is schematically sketched in Fig. 3, where $m(x, y)$ describes two

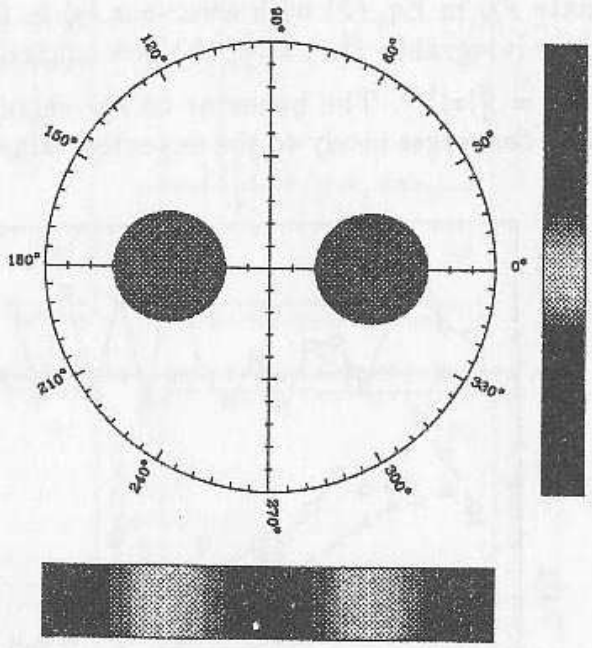

Figure 3. Illustration of the tomographic reconstruction of a 2-d image (here two holes in a uniform background) from its 1-d transmission profiles at different angles ϕ.

circular holes in a uniform background. The tomographic machine—for example, an X-ray equipment—collects many stripe photos of the sample from various directions ϕ, and then numerically performs a mathematical transform in order to reconstruct $m(x,y)$ from its radial profiles $r(s,\phi)$ at different ϕ's. The word "tomography" is customary to denote such imaging procedure starting from radial projections. Mathematically, the radial profiles $r(s,\phi)$ are just the integrals of the image $m(x,y)$ along different directions ϕ, namely

$$r(s,\phi) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{\pi}\, m\left(s\cos\phi - t\sin\phi, s\sin\phi + t\cos\phi\right) , \tag{4}$$

with s denoting the current coordinate along the direction orthogonal to the projection, and t the coordinate along the projection direction. The choice of constants in Eq. (4) gives $r(s,\phi)$ normalized to unit versus s for $m(x,y)$ also normalized to unit. In the following I will adopt the complex notation $m(x,y) \to m(\alpha,\bar{\alpha})$, with $\alpha = x + iy$. Then Eq. (4) is rewritten as follows

$$r(x,\phi) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}y}{\pi}\, m\left((x+iy)e^{i\phi}, (x-iy)e^{-i\phi}\right) , \tag{5}$$

The collection of all projections $r(x,\phi)$ at different ϕ's is called *Radon transform.*

The reconstruction of the "image" $m(\alpha,\bar{\alpha})$ from its "projections" $r(x,\phi)$—this reconstruction is also called "back-projection"—is given by the inverse Radon transform [25]

$$m(\alpha,\bar{\alpha}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k\,|k|}{4} \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\; r(x,\phi) \exp\left[ik\left(x - \alpha_\phi\right)\right] , \tag{6}$$

where $\alpha_\phi = \mathrm{Re}(\alpha e^{-i\phi})$. The derivation of (6) from (5) is given in Appendix 13.2. Upon exchanging integrals over k and ϕ, the image $m(\alpha,\bar{\alpha})$ is written in double integral form

$$m(\alpha,\bar{\alpha}) = \int^{\pi} \frac{d\phi}{\pi} \int^{+\infty} \mathrm{d}x\, r(x,\phi)\, K(x-\alpha_\phi) \tag{7}$$

with the kernel $K(z)$ given by

$$K(z) = -\frac{1}{2}\,\mathrm{P}\frac{1}{z^2} \equiv -\lim_{\varepsilon\to 0^+}\frac{1}{2}\mathrm{Re}\frac{1}{(z+i\varepsilon)^2}\,, \tag{8}$$

P denoting the Cauchy principal value. Integrating Eq. (7) by parts, one obtains the familiar "filtering" procedure

$$m(\alpha,\bar{\alpha}) = \int_0^{\pi}\frac{\mathrm{d}\phi}{2\pi}\mathrm{P}\int_{-\infty}^{+\infty}\mathrm{d}x\,\frac{\partial r(x,\phi)/\partial x}{x-\alpha_\phi}\,, \tag{9}$$

which is commonly used in conventional tomographic imaging [26]. Notice that because of the principal value and of the x-derivative in Eq. (9) one can recover $m(\alpha,\bar{\alpha})$ only when $r(x,\phi)$ is given analytically. An analytical form for projections $r(x,\phi)$ can be achieved from raw data either by "splining" the experimental profiles, or by filtering them with a high-k cutoff in the Fourier transform (6), or finally by setting a nonzero ε in the kernel (8).

4. Why quantum homodyne tomography?

The homodyne-tomography technique originated from the simple idea [10] that the probability distributions $p(x,\phi)$ of the outcomes x of the quadrature $\hat{x}_\phi$ at phase ϕ with respect to the LO is just the marginal probability of the Wigner function $W(\alpha,\overline{\alpha})$ along the line $\alpha = re^{i\phi}$, $r\in(-\infty,+\infty)$ in the complex plane: in other words, the collection of probability densities $p(x,\phi)$ for $0\le\phi\le\pi$ is the Radon transform of $W(\alpha,\overline{\alpha})$. Hence, $W(\alpha,\overline{\alpha})$ can be evaluated by tomographic imaging (inverse Radon transform) from $p(x,\phi)$, namely

$$W(\alpha,\overline{\alpha}) = \int_{-\infty}^{+\infty}\frac{\mathrm{d}k\,|k|}{4}\int_0^{\pi}\frac{\mathrm{d}\phi}{\pi}\int_{-\infty}^{+\infty}\mathrm{d}x\,p(x,\phi)\exp\left[ik(x-\alpha_\phi)\right]\,, \tag{10}$$

where $\alpha_\phi = \mathrm{Re}(\alpha e^{-i\phi})$. Eq. (10) was the basis of the first tomographic method for detecting the density matrix of the radiation field. By means of the filtering or splining procedures discussed in Section 3, an analytic form for $p(x,\phi)$ is recovered from histograms of many collected data, and the Wigner function is obtained. Then, the density matrix in the number representation is evaluated using the following integral transforms [9]

$$\langle x+x'|\hat{\varrho}|x-x'\rangle = \int_{-\infty}^{\infty}\frac{\mathrm{d}y}{\pi}\,e^{2ix'y}W(x+iy,x-iy)\,, \tag{11}$$

$$\langle n|\hat{\varrho}|m\rangle = \frac{1}{\sqrt{\pi 2^{n+m-1}n!m!}}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}x'\,e^{-(x^2+x'^2)}H_n(\sqrt{2}x)H_m(\sqrt{2}x')\,\langle x|\hat{\varrho}|x'\rangle\,, \tag{12}$$

where $H_n(x)$ denote Hermite polynomials. The choice of the k-cutoff or of the maximum degree of splining should be done carefully as a function of the number of data. However, both filtering parameters do not correspond to any meaningful kind of approximation in terms of the matrix elements $\langle n|\hat{\varrho}|m\rangle$: thus, this method is quantitatively unreliable, and can only give qualitative informations on the Wigner function for large numbers of data and nearly unit quantum efficiency. As noticed in Ref.[11], *over-smoothing* produces systematic errors equivalent to make the state "more classical", thus preventing from detecting the

most interesting quantum features, such as, for example, the oscillations of the photon number probability of a squeezed state [27]. On the other hand, *under-smoothing* would lead to systematic negative probabilities (i. e. diagonal matrix elements). In Section 6 I will derive the exact sampling formulas of Refs. [11] and [12] for measuring the density matrix without filtering.

5. On the impossibility of measuring the Wigner function

Upon exchanging the "averaging" integrals over ϕ and x with respect to the outer integral over k, Eq. (10) can be written in a form similar to Eq. (7) as expectation over the random couple (x, ϕ). One has

$$W(\alpha, \overline{\alpha}) = \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x, \phi)\, K(x - \alpha_\phi)\,, \tag{13}$$

where, as in Eq. (8), the kernel $K(x)$ is given by

$$K(x) = -\frac{1}{2}\mathrm{P}\frac{1}{x^2} \equiv -\lim_{\varepsilon \to 0^+} \frac{1}{2}\mathrm{Re}\frac{1}{(x + i\varepsilon)^2}\,. \tag{14}$$

From Eq. (14) it is apparent that the kernel $K(x)$ is unbounded, and thus, as discussed in Sect. 2, the Wigner function cannot be "measured" (i.e. statistically sampled), unless boundness is artificially introduced by fixing a nonzero value of ε.

6. Exact method for measuring the density matrix

We want to obtain a general matrix element $\langle\psi|\hat{\varrho}|\varphi\rangle$ in form of expectation, namely as in Eq. (13), but with $\langle\psi|\hat{\varrho}|\varphi\rangle$ in place of $W(\alpha, \overline{\alpha})$ and with a function $f_{\psi\varphi}(x, \phi)$ in place of the kernel $K(x - \alpha_\phi)$. Then, if and only if the function $f_{\psi\varphi}$ is bounded, every moment of the function will be bounded for all possible probability distributions $p(x, \phi)$ and, according to the central-limit theorem discussed in Sect. 2, the matrix element can be sampled over a sufficiently large set of data, and the average values for different experiments will be Gaussian distributed, allowing estimation of the confidence intervals (3).

In order to obtain the integral kernel $f_{\psi\varphi}$ for the matrix element $\langle\psi|\hat{\varrho}|\varphi\rangle$, one can start from the operator identity

$$\hat{\varrho} = \int \frac{\mathrm{d}^2\alpha}{\pi} \mathrm{Tr}(\hat{\varrho} e^{-\overline{\alpha}a + \alpha a^\dagger})\, e^{-\alpha a^\dagger + \overline{\alpha} a} \tag{15}$$

which, by changing to polar variables $\alpha = (i/2)ke^{i\phi}$, becomes

$$\hat{\varrho} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k\,|k|}{4} \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi} \mathrm{Tr}(\hat{\varrho} e^{ik\hat{x}_\phi})\, e^{-ik\hat{x}_\phi}\,. \tag{16}$$

Equation (15) is nothing but the operator form of the Fourier-transform relation between Wigner function and characteristic function: the reader can easily check it by himself. The trace-average in Eq. (16) can be evaluated in terms of $p(x, \phi)$, using the complete set $\{|x\rangle_\phi\}$ of eigenvectors of $\hat{x}_\phi$. After exchanging the integrals over x and ϕ with respect to the outer integral over k, one obtains the quantum equivalent of Eq. (13), namely

$$\hat{\varrho} = \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x, \phi) K(x - \hat{x}_\phi)\,, \tag{17}$$

where K is still given by Eq. (14), but now is a function of the quadrature operator $\hat{x}_\phi$. Taking matrix elements of both sides of Eq. (17) between vectors ψ and φ, we obtain the desired sampling formula

$$\langle\psi|\hat{\varrho}|\varphi\rangle = \int_0^\pi \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x,\phi)\langle\psi|K(x-\hat{x}_\phi)|\varphi\rangle\,. \tag{18}$$

Hence the averaged function is $f_{\psi\varphi}(x,\phi) \equiv \langle\psi|K(x-\hat{x}_\phi)|\varphi\rangle$. Now the reader could object that Eq. (18) is useless, because $K(x)$ is unbounded. But, he must carefully consider that for some vectors ψ and φ the matrix element $\langle\psi|K(x-\hat{x}_\phi)|\varphi\rangle$ can be bounded, although the function $K(x)$ is not! As an example, one can consider the expectation value $\langle\alpha|\delta(x-\hat{x}_\phi)|\alpha\rangle$ of $\delta(x-\hat{x}_\phi)$ between coherent states $|\alpha\rangle$, which is just the Gaussian $|\langle x|\alpha\rangle|^2 = \sqrt{2/\pi}\exp[-2(x-\alpha_\phi)^2]$, and hence is clearly bounded.

Before analyzing specific matrix representations, I generalize the sampling formula (18) to the case of nonunit quantum efficiency. As shown in my lectures on "Quantum estimation theory and optical detection" [3], low efficiency detectors in a homodyne scheme simply produce a probability $p_\eta(x,\phi)$ which is a Gaussian convolution of the ideal probability $p(x,\phi)$ for $\eta = 1$. In terms of the generating functions of the $\hat{x}_\phi$-moments one has

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, p_\eta(x,\phi)e^{ikx} = \exp\left(-\frac{1-\eta}{8\eta}k^2\right) \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x,\phi)e^{ikx}\,. \tag{19}$$

Upon substituting Eq. (19) into Eq. (16) and following the same lines that lead us to Eq. (17), one obtains the operator identity

$$\hat{\varrho} = \int_0^\pi \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\, p_\eta(x,\phi)K_\eta(x-\hat{x}_\phi)\,, \tag{20}$$

where now the kernel is given by

$$K_\eta(x) = \frac{1}{2}\mathrm{Re}\int_0^{+\infty} \mathrm{d}k\, k\, \exp\left(\frac{1-\eta}{8\eta}k^2 + ikx\right)\,. \tag{21}$$

Again the desired sampling formula for $\langle\psi|\hat{\varrho}|\varphi\rangle$ is obtained by taking matrix elements of both sides of Eq. (21). Notice that now the kernel $K_\eta(x)$ is generally not even a tempered distribution: however, as we will see in the Sect. 7, the matrix elements of $K_\eta(x-\hat{x}_\phi)$ are bounded for some representations, depending on the value of η.

7. Measurable representations and bounds for quantum efficiency

The matrix element $\langle\psi|K_\eta(x-\hat{x}_\phi)|\varphi\rangle$ is bounded if $\langle\psi|e^{-ir\hat{x}_\phi}|\varphi\rangle$ decays faster than $\exp\left[-(1-\eta)k^2/8\eta\right]$ in Eqs. (20) and (21). We can determine the balancing of Gaussian functions in Eqs. (20) and (21) upon introducing the concept of "resolution" $\varepsilon^2_\psi(\phi)$ of the vector ψ in the $\hat{x}_\phi$-representation. This is given by

$$|_\phi\langle x|\psi\rangle|^2 \simeq \exp\left[-\frac{x^2}{2\varepsilon^2_\psi(\phi)}\right] \qquad \text{(leading term)}\,. \tag{22}$$

The resolution $\varepsilon^2_\psi(\phi)$ represents the Gaussian decay length of the probability $|_\phi\langle x|\psi\rangle|^2$ versus x, with the rule that $\varepsilon^2_\psi(\phi) = 0$ and $\varepsilon^2_\psi(\phi) = \infty$ for functions decaying faster and

slower than the Gaussian one, respectively. Now, let us define the "reduced" resolution averaged on the two vectors ψ and φ

$$\frac{2}{\varepsilon^2(\phi)} \doteq \frac{1}{\varepsilon_\psi^2(\phi)} + \frac{1}{\varepsilon_\varphi^2(\phi)} \,. \tag{23}$$

In terms of $\varepsilon^2(\phi)$ it is easy to state the balancing between the decay of $\langle\psi|e^{-ir\hat{x}_\phi}|\varphi\rangle$ and the growing of the anti-Gaussian in Eq. (21). Using the resolution of the identity in terms of the eigenvectors $\{|x\rangle_\phi\}$ of $\hat{x}_\phi$ we obtain

$$\begin{aligned}\langle\psi|K_\eta(x-\hat{x}_\phi)|\varphi\rangle &= \frac{1}{2}\mathrm{Re}\int_0^{+\infty} \mathrm{d}k\, k\, e^{ikx} \int_{-\infty}^{+\infty} \mathrm{d}x'\, e^{-ikx'} \exp\left(\frac{1-\eta}{8\eta}k^2\right) \langle\psi|x'\rangle_{\phi\phi}\langle x'|\varphi\rangle \\ &= \frac{1}{2}\mathrm{Re}\int_0^{+\infty} \mathrm{d}k\, k\, e^{ikx} \exp\left[-\frac{1}{2}k^2\left(\varepsilon^2(\phi) - \frac{1-\eta}{4\eta}\right)\right] \qquad \text{(leading term)}\,, \end{aligned} \tag{24}$$

namely, we can readily assert that the matrix element $\langle\psi|K(x-\hat{x}_\phi)|\varphi\rangle$ is bounded if the following inequality is satisfied for $\phi \in [0,\pi)$

$$\eta > \frac{1}{1+4\varepsilon^2(\phi)}\,, \tag{25}$$

or, in terms of $\varepsilon(\phi)$,

$$\varepsilon(\phi) > \frac{1}{2}\sqrt{\eta^{-1}-1}\,. \tag{26}$$

Hence, we conclude that:

The matrix element $\langle\psi|\hat{\varrho}|\varphi\rangle$ can be measured if the minimum reduced resolution of vectors ψ and φ in the quadrature representations satisfies the bound

$$\varepsilon = \min_{0\le\phi\le\pi}\{\varepsilon(\phi)\} > \frac{1}{2}\sqrt{\eta^{-1}-1}\,. \tag{27}$$

We now consider some particular representations of interest for applications.

7.1. QUADRATURE REPRESENTATION

The scalar product of two quadrature eigenvectors at the same phase ϕ is a delta-function, and the square of the delta-function in Eq. (22) is not well defined. We can easily check that the kernel is unbounded, even for $\eta = 1$, where it is given by

$$_\phi\langle x_1|K(x-\hat{x}_\phi)|x_2\rangle = \frac{|x_1-x_2|}{\sin^2\phi}\exp\left\{-2i\frac{x_1-x_2}{\sin\phi}\left[x-\frac{1}{2}(x_1+x_2)\cos\phi\right]\right\}\,. \tag{28}$$

Hence, one concludes that it is not possible to measure the density matrix in the quadrature representation, even for $\eta = 1$.

7.2. COHERENT-STATE REPRESENTATION

The resolution is $\varepsilon = \frac{1}{2}$: according to Eq. (27) the density matrix can be measured for $\eta > \frac{1}{2}$. The integral kernel is given by

$$\langle\alpha|K_\eta(x-\hat{x}_\phi)|\beta\rangle = 2\kappa^2\langle\alpha e^{i\phi}|\beta e^{i\phi}\rangle e^{-2(\kappa x-w_\phi)^2}\,\Phi\left(-\frac{1}{2},\frac{1}{2};2(\kappa x-w_\phi)^2\right)\,, \tag{29}$$

with $\kappa = \sqrt{\eta/(2\eta-1)}$, $w_\phi = \frac{1}{2}(\beta e^{i\phi} + \overline{\alpha} e^{-i\phi})$, and $\Phi(\alpha,\beta;z)$ denoting the degenerate hypergeometric function.

7.3. NUMBER–STATE REPRESENTATION

The resolution is $\varepsilon = \frac{1}{2}$: the matrix element can be measured for $\eta > \frac{1}{2}$. The integral kernel is given by

$$\langle n|K_\eta(x-\hat{x}_\phi)|n+d\rangle = e^{-id\phi} 2\kappa^{d+2} \sqrt{\frac{n!}{(n+d)!}} e^{-\kappa^2 x^2} \times \sum_{\nu=0}^{n} \frac{(-)^\nu}{\nu!} \binom{n+d}{n-\nu} (2\nu+d+1)! \kappa^{2\nu} \mathrm{Re}\left\{(-i)^d D_{-(2\nu+d+2)}(-2i\kappa x)\right\} . \tag{30}$$

In Eq. (30) $D_\sigma(z)$ denotes the parabolic cylinder function. Notice that the analytic expression of the kernel does not contain any dimensional truncation of the Hilbert space.

7.4. SQUEEZED–STATE REPRESENTATION

The leading term in Eq. (22) comes only from the squeezed vacuum (the coherent part just shifts the quadrature). For a squeezed vacuum with squeezing parameter s, the probability of the quadrature $\hat{x}_\phi$ is the Gaussian $|_\phi\langle x|\psi\rangle|^2 = \sqrt{\frac{2s_\phi^2}{\pi}} \exp\left(-2s_\phi^2 x^2\right)$, where $s_\phi = |s^{\frac{1}{2}} \sin\phi - i s^{-\frac{1}{2}} \cos\phi|^{-2}$. Hence, the smallest resolution is $\varepsilon = \frac{1}{2}\min(s, s^{-1})$ and the matrix element can be measured for $\eta > [1 + \min(s, s^{-1})]^{-1}$.

7.5. ABSOLUTE BOUND FOR QUANTUM EFFICIENCY

We have seen that for the coherent and number-state representations the density matrix can be measured for quantum efficiency $\eta > 1/2$. On the other hand, for the squeezed state representation one has $\eta > [1 + \min(s, s^{-1})]^{-1}$, which is even greater than $1/2$. Therefore, one could wonder if $1/2$ is an absolute bound for η, as conjectured in Ref. [12]. This is clearly true for all representations with minimum uncertainty product

$$\varepsilon(\phi)\varepsilon(\phi+\pi/2) \geq \frac{1}{4} , \tag{31}$$

and either the number (complete) or the coherent/squeezed (over-complete) representations satisfy the Heisenberg inequality (31) with the equal sign. Now the question is if it is possible to find (over)-complete basis for the Fock space having uncertainty product larger than $\frac{1}{4}$: in my knowledge, there are no such "spread" basis. Notice that it is difficult to obtain "spread" basis as unitary transformations of either the number or the coherent basis. In fact, any unitary operator that is the exponential of a Hermitian bilinear polynomial in a and $a^\dagger$ preserves the uncertainty product: on the other hand, the exponential of a Hermitian polynomial of degree higher than 2 is generally not analytical on the Fock basis (apart from trivial functions of the number operator) [28].

8. Statistical errors

In this section we study in more detail the analytic form of the integral kernel, in order to understand the mechanisms producing statistical errors in the measured matrix elements.

We only analyze the case of the number representation, with the kernel given by Eq. (30): the same arguments can be easily extended to other representations.

In Fig. 4 the kernel $\langle n|K_\eta(x-\hat{x}_\phi)|n+d\rangle$ for $\eta = 1$ is plotted versus x for $\phi = 0$ at different values of n, and d. One can see that for $d = 0$—along the diagonal of the

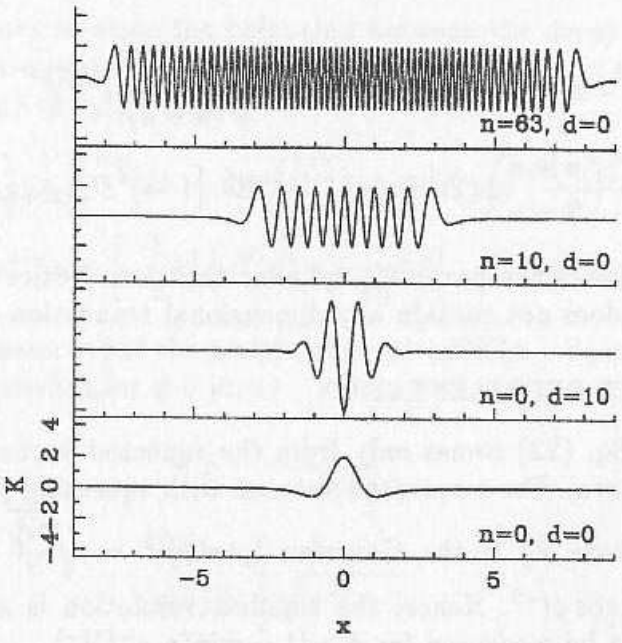

Figure 4. The kernel $\langle n|K_\eta(x-\hat{x}_\phi)|n+d\rangle$ of Eq. (30) for $\eta = 1$ (the y-scale is fixed).

matrix—the range of the kernel is bounded between -2 and 2, and increases slowly versus distance d from the diagonal. For increasing n and d the kernel oscillates fast, with an increasing number of nodes. Fast oscillations make the average of the kernel—hence the measured value $\langle n|\hat{\varrho}|m\rangle$—more sensitive to fluctuations of the random outcomes x of the quadrature, producing confidence intervals that increase versus n and d. On the other hand, the bounded range makes errors themselves bounded, so they saturate at large n's.

For $\eta < 1$ the behavior of the kernel changes dramatically, with the range increasing versus n more and more rapidly as η approaches the lower bound $\eta = 0.5$ (see Fig.5). In this case the resulting errors increase rapidly versus n, and more data will be needed to "clean out" the additional noise due to nonunit quantum efficiency.

The above mechanisms for errors are well illustrated in Figs. 6 and 7, where the photon number probability distribution is plotted from computer-simulated experiments. The behavior of the confidence intervals (represented by the gray-shaded thickness of horizontal lines) is qualitatively very different for $\eta = 1$ (Fig. 6) and for $\eta = .8$ (Figs. 7). Fig. 6 shows the Wheeler-Schleich oscillations [27] of a squeezed state. Fig. 7a gives the probability for a squeezed vacuum measured for $\eta = 0.8$, whereas Fig. 7b gives the probability obtained from the same data at $\eta = .8$, but using the kernel with $\eta = 1$: this shows the "smearing" effect of improper accounting for quantum efficiency. From both Figs. 6 and 7 it is evident that homodyne detection of the density matrix allows to recover delicate interference oscillations of nonclassical states.

The statistical reliability of both measured values and confidence intervals can be checked by means of Monte Carlo simulations. A typical check histogram for the devia-

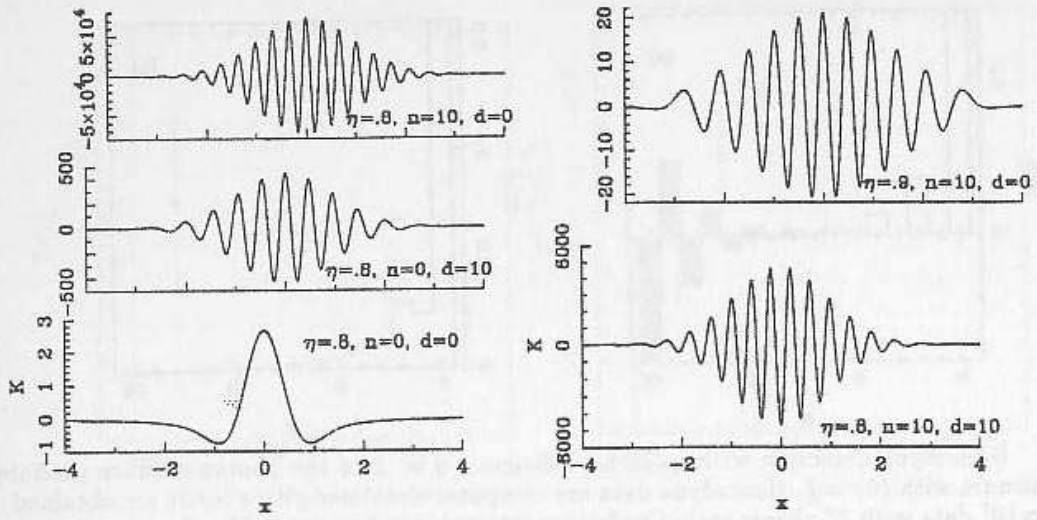

Figure 5. The kernel $\langle n|K_\eta(x-\hat{x}_\phi)|n+d\rangle$ in Eq. (30) for $\eta < 1$. Notice thevery different ranges.

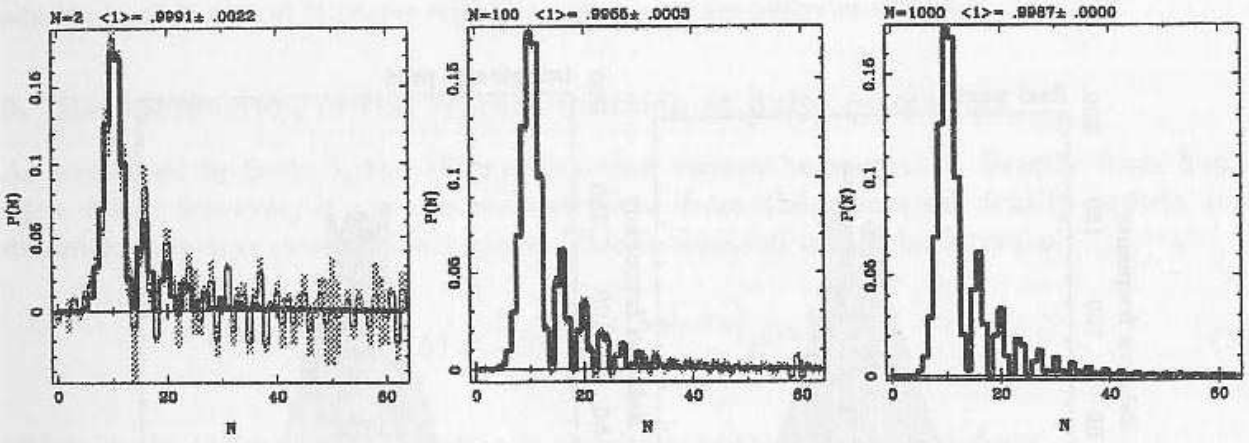

Figure 6. Tomographic reconstruction of the Wheeler-Schleich oscillations of a squeezed state with $\langle\hat{n}\rangle = 13.5125$ and 5.51 squeezing photons. Confidence intervals are represented by the gray-shaded thickness of horizontal lines. Detection efficiency is $\eta = 1$. Homodyne data are computer-simulated. From the left to the right we have N=2, 100 and 1000 blocks of data, respectively, with 5000 data each at totally random phases. On the top of each figure a normalization check is reported.

tions from theoretical values $\Delta\varrho_{n,m} \equiv \left(\varrho_{n,m}^{(exper)} - \varrho_{n,m}^{(theor)}\right)/\delta_{n,m}$ normalized by confidence intervals $\delta_{n,m}$ is given in Fig. 8 for all matrix elements of a 65×65 density matrix: both histograms for real and imaginary parts of $\varrho_{n,m}$ compare very well with a standard Gaussian.

Finally, in Fig. 9 a typical result for convergence of normalization and mixing is given for a coherent state.

In conclusion of this section a remark is in order on homodyning the density matrix as a tool for measuring any desired observable. In fact, from the measured density matrix one can evaluate the expected value of any operator (which is sufficiently well behaved as a function of a and $a^\dagger$). However, nonzero confidence intervals for the matrix elements correspond to additional noise for the indirectly detected single observable. In Ref. [14] it has been estimated that homodyning the density matrix adds several dB's of noise

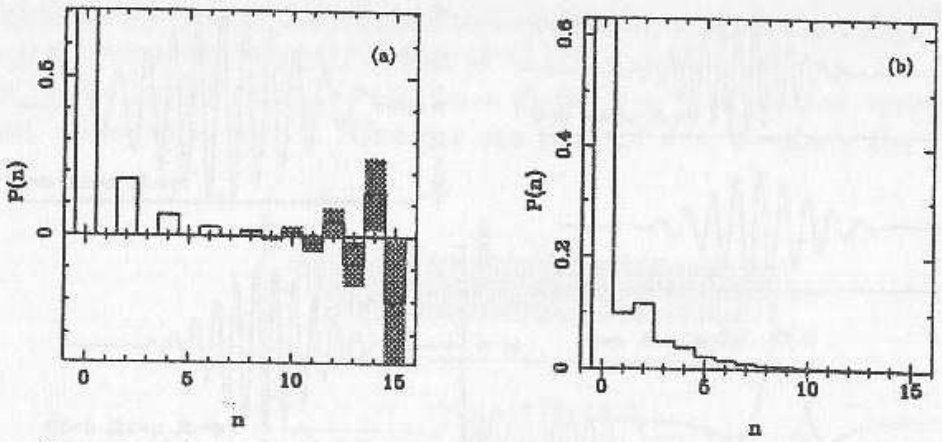

Figure 7. Homodyne detection with quantum efficiency $\eta = .8$ of the photon-number probability of a squeezed vacuum with $\langle \hat{n} \rangle = 1$. Homodyne data are computer-simulated. Here result are obtained from 200 blocks of 5×10^5 data with 27 phases each. Confidence intervals are represented by the gray-shaded thickness of horizontal lines. Fig. 7a: reconstruction using the kernel (30) with $\eta = .8$. Fig. 7b: reconstruction from the same data without taking into account quantum efficiency, namely using the kernel (30) with $\eta = 1$.

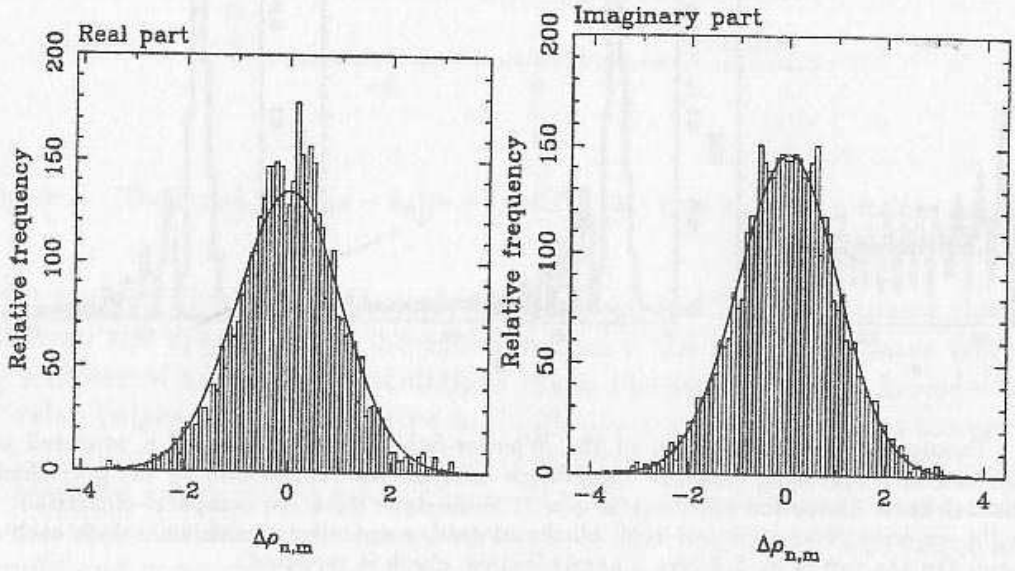

Figure 8. Distribution of the normalized deviations from the theoretical values $\Delta\varrho_{n,m} \equiv \left(\varrho_{n,m}^{(exper)} - \varrho_{n,m}^{(theor)}\right)/\delta_{n,m}$ for the first 65×65 matrix elements. Detection efficiency is $\eta = 1$. The quantum state is coherent with $\langle n \rangle = 4$. The histograms pertain the deviations of all matrix elements (real and imaginary parts separately) detected during 32 experiments with 26 scanning phases each, and 40 measurements for each phase. A standardized Gaussian curve is superimposed to each histogram.

with respect to the ideal measurement of the single observable (it is particularly noisy for detection of the phase of the field). Therefore, ideal detection of a single observable—which provides only a partial information on the quantum state—is always less noisy than homodyning the density matrix—which, however, gives the most complete knowable information.

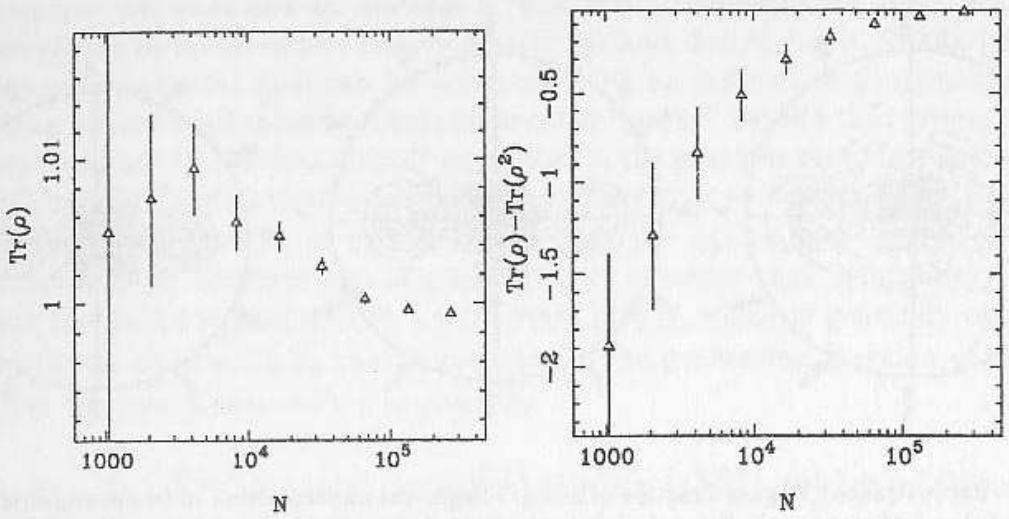

Figure 9. Normalization and mixing of the homodyne-detected density matrix versus the total number of data of the experiment (i.e. number of phases × number of data per phase) for a coherent state with $\langle n \rangle = 4$. Detection efficiency is $\eta = 1$. The matrix is 22×22. Error bars are estimated by grouping data into blocks of 20 data of 26 phases each. The experiments are computer-simulated.

9. Reconstruction of the Wigner function at finite resolution

As explained in Sect. 5, the Wigner function cannot be measured directly from homodyne data: however, it can be reconstructed from the measured density matrix in a dimensionally-truncated Hilbert space. This is achieved using the formula

$$W(\alpha, \bar{\alpha}) = \frac{2}{\pi} \mathrm{Tr} \left[\hat{\varrho} e^{2\left(\alpha a^\dagger - \bar{\alpha} a\right)} e^{i\pi a^\dagger a} \right] , \tag{32}$$

which in the number representation is rewritten as the Fourier transform

$$W(\alpha, \bar{\alpha}) = \mathrm{Re} \sum_{d=0}^{\infty} e^{id \arg(\alpha)} \sum_{n=0}^{\infty} \Lambda(n, d; |\alpha|^2) \varrho_{n,n+d} , \tag{33}$$

with $\varrho_{n,m} \doteq \langle n | \hat{\varrho} | m \rangle$, and

$$\Lambda(n, d; |\alpha|^2) = (-)^n 2(2 - \delta_{d0}) |2\alpha|^d \sqrt{\frac{n!}{(n+d)!}} e^{-2|\alpha|^2} L_n^d(|2\alpha|^2) , \tag{34}$$

$L_n^d(x)$ denoting Laguerre polynomials. The Hilbert space truncation of the sums in (33) sets the *resolution in polar coordinates* for the reconstructed Wigner function: this will be illustrated in more detail in Section 11. Laguerre polynomials in Eq. (34) produce radial oscillations, with the result that despite the truncated dimension of the matrix is relatively small, the natural quantum oscillations in the Wigner functions are easily recovered. In Fig. 10 the Wigner function of Schrödinger cat states are given as reconstructed from homodyne computer-simulated data. As compared to the exact method, the old filtered technique (revisited in Sect. 4) would have required very small cutoffs to recover such Wigner oscillations, hence needing much larger sets of data.

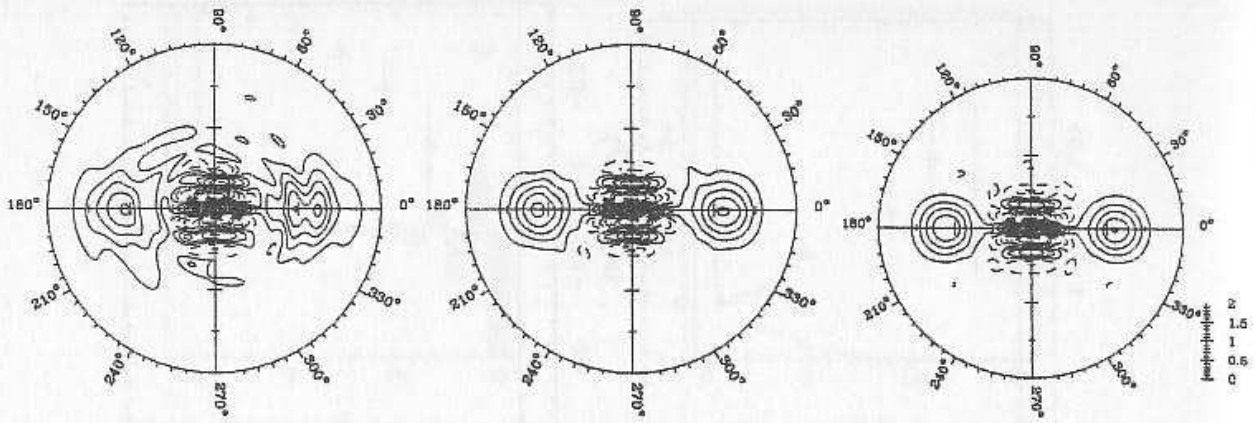

Figure 10. Reconstructed Wigner function of a Schrödinger-cat superposition of two symmetrical coherent states with $\langle n \rangle = 5$ photons each. From the left to the right: 1, 10, 100 blocks of data each with 52 phases and 1000 data for each phase. The dimension of the density matrix is truncated at 16 photons max. Detection efficiency is $\eta = 1$. Notice the oscillations near the origin (dashed lines represent negative values).

10. On the impossibility of measuring the density matrix of a single system

The notion of "measuring a state" is meaningful only if it refers to many repeated measurements performed on an *ensemble of equally prepared* systems: otherwise, the state of an individual system is unobservable, whatever sequence of measurements is performed on the system. In fact, a single measurement cannot yield sufficient information to reconstruct the state; on the other hand, it would uncontrollably disturb the state, thereby ruling out the possibility of extracting the remaining information from successive measurements.

Recently, the possibility of determining the wave function of a single quantum system has been reconsidered by several authors [29, 30, 31, 32, 33], exploring special measurement schemes, based essentially on vanishingly weak quantum nondemolition measurements [29, 30], or "reversible" measurements [31, 32, 33]. In each of these schemes the conclusion is that it is practically impossible to measure the wave function of a single system, either because the weakness of the measuring interaction prevents one from gaining information on the wave function [29], or because the scheme essentially measures only orthogonal states [30], or because quantum measurements can be physically reverted only with probability of success equal to 1/2 [33]. Only very recently Yuen and I[24] proved the general impossibility of determining the state of a single quantum system for arbitrary measuring schemes, including any kind of succession of measurements, which we proved to have the same probability distribution of a single joint measurement with output state independent on the input. We have derived this proof because it seems that, despite its fundamental relevance in the logical framework of quantum mechanics, it has been never proved in general. Here, from Ref. [24] I simply report a scheme for making repeated homodyne measurements on a single copy of the radiation field, in order to show in which way the system state is perturbed however weak the system-apparatus interaction is. This is due to the fact that in order to overcome the low effective quantum efficiency corresponding to the weak interaction, the apparatus must be prepared in a highly squeezed state, as also shown by Tombesi in his set of lectures [34] and in Ref. [35].

Insofar the tomographic scheme that we have considered is based on "second kind" measurements, which completely destroy the quantum state of the system. Many measurements are performed, but the system is prepared in the same state $\hat{\rho}$ before each

measurement. We want now to consider a "first kind" version of the above scheme that in principle allows to measure the density matrix without destroying it. Similarly to any first kind measurement, this goal can be achieved using an indirect-measurement, namely by performing second-kind measurements on another "probe" mode b that interacts with a via a unitary operator $\hat{U}$: this was already explained in the previous set of lectures on Quantum estimation theory and optical detection [3]. An easy way to achieve a first-kind measurement of the quadrature $\hat{a}_\phi \equiv \hat{x}_\phi$ is by measuring the quadrature $\hat{b}_\phi = \frac{1}{2}(b^\dagger e^{i\phi} + b e^{-i\phi})$ of another mode b. Without loss of generality, we consider that before every single measurement the probe is prepared in a pure state $|v(\phi)\rangle$, which is generally optimized as a function of the observable $\hat{b}_\phi$ that is measured. The generating function of the moments of $\hat{b}_\phi$ after the interaction with a is given by

$$X(\lambda,\phi) = \mathrm{Tr}\left(e^{i\lambda\hat{b}_\phi}\hat{U}\,\hat{\varrho}\otimes|v(\phi)\rangle\langle v(\phi)|\hat{U}^\dagger\right) = \mathrm{Tr}\left[\exp\left(i\lambda\hat{U}^\dagger\hat{b}_\phi\hat{U}\right)\hat{\varrho}\otimes|v(\phi)\rangle\langle v(\phi)|\right], \quad (35)$$

and is just the Fourier transform of the probability distribution of the experimental outcomes. We consider the following interaction

$$\hat{U} = \exp\left[\kappa\left(ab^\dagger - a^\dagger b\right)\right], \quad (36)$$

which describes the linear field transformation at a beam splitter [3]

$$\hat{U}^\dagger b\hat{U} = \sin\kappa\, a + \cos\kappa\, b \equiv \vartheta^{1/2}a + (1-\vartheta)^{1/2}b, \quad (37)$$

where ϑ is the mirror transmissivity, and b is the field mode of the unused port at the same frequency of a. Thus, the present scheme physically corresponds to let the field mode a shine over a long chain of low-transmissivity mirrors, detecting the quadrature of the weak transmitted field at each mirror. Due to the linearity of Eq. (37), the moment generating function factorizes in the following way

$$X(\lambda,\phi) = \chi_a(\vartheta^{1/2}\lambda,\phi)\chi_b((1-\vartheta)^{1/2}\lambda,\phi), \quad (38)$$

where $\chi_a(\lambda,\phi) = \mathrm{Tr}[\exp(i\lambda\hat{a}_\phi)\hat{\varrho}]$ is the generating function for the noninteracting mode a only, and, analogously, $\chi_b(\lambda,\phi)$ for the mode b. Using the operator identity (15) the density operator $\hat{\varrho}$ is written in terms of the generating function $\chi_a(\lambda,\phi)$

$$\hat{\varrho} = \int_0^\pi \frac{d\phi}{\pi}\int_{-\infty}^{+\infty}\frac{d\lambda\,|\lambda|}{4}e^{-i\lambda\hat{a}_\phi}\chi_a(\lambda,\phi), \quad (39)$$

and using Eq. (38) one obtains

$$\hat{\varrho} = \int_0^\pi \frac{d\phi}{\pi}\int_{-\infty}^{+\infty} dx\, p_\vartheta(x,\phi)\Xi_\vartheta(x-\hat{a}_\phi), \quad (40)$$

where $p_\vartheta(x,\phi)$ is the probability of the measured quadrature $\hat{b}_\phi$ rescaled by $\vartheta^{1/2}$ (as for customary homodyning, where the output photocurrent is rescaled by $\eta^{1/2}$[3])

$$p_\vartheta(x,\phi) = \vartheta^{1/2}\int_{-\infty}^{+\infty}\frac{d\lambda}{2\pi}e^{-i\lambda\vartheta^{1/2}x}X(\lambda,\phi). \quad (41)$$

The kernel $\Xi_\vartheta(x)$ in Eq. (40) is given by

$$\Xi_\vartheta(x) = \frac{1}{2}\mathrm{Re}\int_0^{+\infty} d\lambda\, \lambda e^{i\lambda x}[\chi_b(\lambda\sqrt{(1-\vartheta)/\vartheta},\phi)]^{-1}, \tag{42}$$

and generally depends on the coupling parameter ϑ ($\equiv \sin\kappa$) and on the probe state $|v(\phi)\rangle$. One can easily see that when the probe mode b is in the vacuum state the kernel (42) is identical to $K_\eta(x)$ in Eq. (21) with $\vartheta \equiv \eta$, namely, the transmissivity ϑ plays the role of the overall quantum efficiency of the indirect measurement. However, as noticed in Ref. [35], the effective quantum efficiency can be decreased at will by squeezing the probe mode b in the direction of the quadrature $\hat{b}_\phi$. More precisely, one prepares the probe in the squeezed vacuum

$$|v(\phi)\rangle \equiv \hat{S}_\phi|0\rangle\,, \tag{43}$$

where

$$\hat{S}_\phi = e^{ib^\dagger b\phi}\hat{S}_0 e^{-ib^\dagger b\phi}\,, \quad \hat{S}_0 = \exp\left[-\frac{r}{2}\left(b^{\dagger 2} - b^2\right)\right], \tag{44}$$

$r > 0$ denoting the squeezing parameter. One has

$$\hat{S}_0\hat{b}_0\hat{S}_0^\dagger = e^r\hat{b}_0\,, \qquad \hat{S}_0^\dagger|x\rangle = e^{r/2}|e^r x\rangle\,, \tag{45}$$

with $|x\rangle$ denoting the eigenvector of $\hat{b}_0$ for eigenvalue x: the rescaling (45) more generally holds for the quadrature $\hat{b}_\phi$ and its eigenvectors $e^{ib^\dagger b\phi}|x\rangle$ when using the rotated squeezing operator $\hat{S}_\phi$ in place of $\hat{S}_0$. With the help of transformations (44) and (45) it is easy to check that the kernel $\Xi_\vartheta(x)$ in Eq. (42) coincides with $K_\eta(x)$ in Eq. (21) with effective efficiency

$$\eta \equiv \frac{e^{2r}\vartheta}{e^{2r}\vartheta + 1 - \vartheta}\,, \tag{46}$$

as also shown in Refs. [34] and [35]. Therefore, by increasing the squeezing parameter r it is possible to enhance the effective quantum efficiency η beyond the allowed bound for measuring the density matrix ($\eta > 1/2$ for number and coherent states). At this point one may think that squeezing the vacuum of b allows one to consider weaker and weaker interactions with $\vartheta \to 0$, with the possibility of performing repeated measurements on the same system with vanishing perturbation at each measurement. However, as we will show immediately, this cannot be attained, because the squeezing needed to keep η as constant also amplifies the perturbation back to a finite extent. This can be seen upon analyzing the limiting behavior of the transition operator $\hat{\Omega}(x,\phi)$, that gives the state reduction after each measurement [3]

$$\hat{\varrho}' = \frac{\hat{\Omega}(x,\phi)\hat{\varrho}\hat{\Omega}^\dagger(x,\phi)}{\mathrm{Tr}[\hat{\varrho}\hat{\Omega}^\dagger(x,\phi)\hat{\Omega}(x,\phi)]}\,. \tag{47}$$

Here, the transition operator $\hat{\Omega}(x,\phi)$ is given by

$$\hat{\Omega}(x,\phi) = \vartheta^{1/4}\langle\vartheta^{1/2}x|e^{-ib^\dagger b\phi}e^{\kappa(ab^\dagger - a^\dagger b)}\hat{S}_\phi|0\rangle\,, \tag{48}$$

where the powers of ϑ account for quadrature rescaling, and one should keep in mind that the matrix element is evaluated between vectors $|\vartheta^{1/2}x\rangle$ and $|0\rangle$ in the Hilbert space of mode b only, so that $\hat{\Omega}(x,\phi)$ is an operator on the Hilbert space of mode a. One has

$$\hat{\Omega}(x,\phi) = \vartheta^{1/4}\langle\vartheta^{1/2}x|\hat{S}_0 e^{-ib^\dagger b\phi} e^{\kappa(a_r b^\dagger - a_r^\dagger b)}|0\rangle \,, \tag{49}$$

where

$$a_r = \cosh r\, a + e^{2i\phi}\sinh r\, a^\dagger \,. \tag{50}$$

Using Eq. (45) and normal ordering the interaction operator with respect to b (by the BCH formula for the $SU(2)$ group [36]), one obtains

$$\hat{\Omega}(x,\phi) = (e^{2r}\vartheta)^{1/4}\langle(e^{2r}\vartheta)^{1/2}x|e^{-ib^\dagger b\phi} e^{\tan\kappa a_r b^\dagger}|0\rangle|\cos\kappa|^{a_r^\dagger a_r} \,. \tag{51}$$

Eq. (51) can be arranged in the following way

$$\hat{\Omega}(x,\phi) = \exp\left[\tan\kappa\, a_r \frac{d}{d\lambda}\bigg|_{\lambda=0}\right] e^{\frac{1}{2}\lambda^2}(e^{2r}\vartheta)^{1/4}\langle(e^{2r}\vartheta)^{1/2}x|\lambda e^{-i\phi}\rangle|\cos\kappa|^{a_r^\dagger a_r} \,, \tag{52}$$

where λ is a real running parameter, and $|\lambda e^{-i\phi}\rangle$ is a coherent state. Upon evaluating the derivative with respect to λ one obtains the explicit operator form of $\hat{\Omega}(x,\phi)$

$$\begin{aligned}\hat{\Omega}(x,\phi) &= \left(\frac{2e^{2r}\vartheta}{\pi}\right)^{1/4} \exp\left[-\left(e^{2r}\vartheta x - \tan\kappa e^{-i\phi}a_r\right)^2\right] \\ &\times \exp\left[\frac{1}{2}\tan^2\kappa e^{-2i\phi}a_r^2\right]|\cos\kappa|^{a_r^\dagger a_r} \,.\end{aligned} \tag{53}$$

We can now finally evaluate the asymptotic form of the transition operator in Eq. (53) in the simultaneous limits of vanishing transmission coefficient $\vartheta \to 0$ and infinite squeezing parameter $r \to \infty$ keeping the effective quantum efficiency η as constant according to Eq. (46). In Eq. (53), for $\vartheta \to 0$ using Eq. (37) one obtains $e^{2r}\vartheta \to \Delta_\eta^{-2}$, where

$$\Delta_\eta^2 \equiv \frac{1-\eta}{\eta} \,. \tag{54}$$

Moreover, one has $e^{-r}e^{-i\phi}a_r \to \hat{a}_\phi$, and $|\cos\kappa|^{a_r^\dagger a_r} \to \exp[-\hat{a}_\phi^2/(2\Delta_\eta^2)]$. Therefore, the asymptotic form of the transition operator is simply given by

$$\hat{\Omega}(x,\phi) = \left(\frac{2}{\pi\Delta_\eta^2}\right)^{1/4} \exp\left[-\frac{(x-\hat{a}_\phi)^2}{\Delta_\eta^2}\right] \,. \tag{55}$$

Equation (55) is the typical form of a von Neumann "reduction" of the state (see my other set of lectures in this volume [3]). Therefore, one concludes that despite the interaction has been tuned as vanishingly small, the state is in fact "reduced" at each measuring step.

In conclusion of this section I want to emphasize that the need of a lower bound for quantum efficiency η is of fundamental relevance. In fact, if the density matrix could be

homodyne detected for vanishingly small quantum efficiency, then there would be no need for squeezing the vacuum of the measuring apparatus in order to enhance the efficiency beyond the bound, and the density matrix of a single system could be detected by vanishingly weak measurements.

11. Quantum and classical Radon transform

In this section I show how the new exact algorithm for homodyning the density matrix can be used also for ordinary imaging. The situation of interest is when the radial profiles are not well resolved digitalized functions, but actually represent the density distribution of random points—as if in our X-ray machine the beam is so weak that radial photos are just the collection of many small spots, each from a single X-ray photon (this situation is sketched in Fig. 11. It is obvious that this case can be reduced to the one of conventional imaging by counting all points falling in a predetermined 1-d mesh, and giving radial profiles in form of histograms (this is what actually happens in a real machine that uses arrays of photodetectors). However, we want to utilize the whole available information from each "event"—i.e. the exact 1-d location of each spot—in a way that is independent on any predetermined mesh. In practice, this extreme situation happens when the signal is so weak and the machine resolution is so high (i.e. the mesh-step is so tiny) that only zero or one photon at most can be collected in each channel.

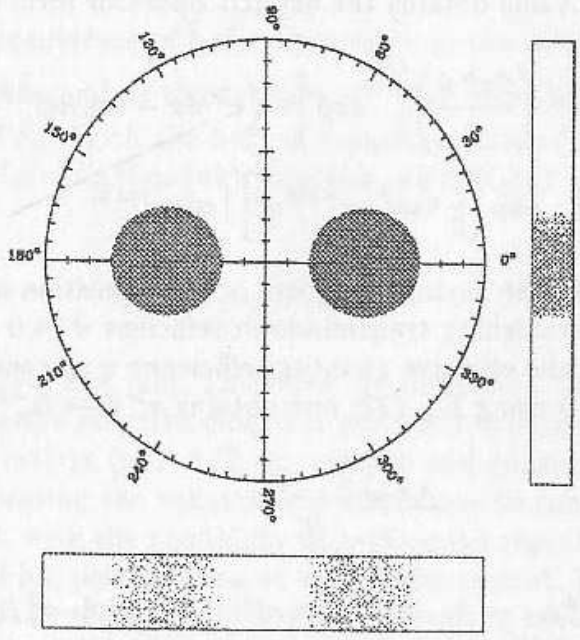

Figure 11. Illustration of imaging in presence of very weak signals. The image is the same as in Fig. 3, but here the transmission profiles are given in terms of random points on a photographic plate.

The above low-signal/high-resolution case naturally brings the imaging problem into the quantum domain. Here, we have at disposal our technique for measuring the density matrix of the field in terms of homodyne outcomes (the equivalent of our radial spots) for different phases of the local oscillator (the equivalent of our ϕ angle). In general an image $W(\alpha, \bar{\alpha})$ does not correspond to a Wigner function of a physical state, due to the fact that the Heisenberg relations unavoidably produce only smooth Wigner functions, whereas a

conventional image can have very sharp edges. However, if one allows the density matrix to be no longer positive definite (but still trace class), a correspondence between density matrices and images is obtained that holds in general, and a convenient truncation of the matrix dimension $d_{\mathcal{H}}$ can be chosen to set the imaging resolution. In summary, the connection between images and density matrix $\varrho_{n,m}$—what I call either the *quantum Radon transform* or the *quantum imaging*—is given by Eqs. (18), (30), (33) and (34) with the sums truncated at maximum photon number $d_{\mathcal{H}} - 1$. I emphasize again that the density matrix does not correspond to any physical state of radiation, and this is why the method is named *Fictitious Photons Tomography* [16].

For suitably chosen dimension $d_{\mathcal{H}}$ the quantum Radon transform also provides a procedure for image compression, with $d_{\mathcal{H}}$ setting the resolution of imaging. The kind of resolution can be understood by studying the behavior of the kernels $\langle n|K(x-\hat{x}_\phi)|n+d\rangle$ in Eq. (30) which are averaged over the experimental data in order to obtain the matrix elements $\varrho_{n,n+d}$. As shown in Fig. 4, outside a region that is increasing versus the indices n and $n+d$, the functions $\langle n|K(x-\hat{x}_\phi)|n+d\rangle$ decrease as x^{-d-2}, whereas inside they oscillate with a number of nodes that increases with $2n+d$. This behavior produces the effects illustrated in Fig. 12 where for increasing dimension $d_{\mathcal{H}}$ I report the tomographic reconstruction of an image that describes a chess-board shaped step-function $W(\alpha,\bar{\alpha})$ in the plane. The plot is obtained by numerically integrating Eq. (18) with the kernel (30) from analytically given transmission profiles $p(x,\phi)$ and then using Eqs. (33) and (34). From Fig. 12 one can see that both radial and angular resolutions improve versus $d_{\mathcal{H}}$, making the details of the image sharper and sharper.

In Fig. 13 a Monte Carlo simulation is reported of the quantum imaging of the font "ψ" for weak signals. The reconstruction is quite accurate already for $5 * 10^6$ data with resolution $d_{\mathcal{H}} = 32$. The situation occurring for small numbers of data is given in the first plot, where the highly resolved image exhibits the natural statistical fluctuations due to the limited number of data. For larger samples the image starts appearing sharper from the random background. For comparison the analytical imaging (as in Fig. 12) is also given for resolution $d_{\mathcal{H}} = 128$.

12. Concluding remarks on experimental applications

I conclude these notes with some remarks on experimental applications of the method of homodyning the density matrix. It is important to realize that although photodetectors themselves in principle seem to provide the easiest way for measuring the photon number distribution, it is very difficult to achieve a photon counting that is ideal for a large range of photon numbers starting from the vacuum. In fact, it is well known that if the detector ideally has no threshold (as for an avalanche photo-diode) then it has a low quantum efficiency, or vice versa, if it has a good quantum efficiency, then it has a high sensitivity threshold (as for a linear photo-diode). Homodyning has the advantage of amplification from the LO, which is highly excited and coherent, such that the difference photocurrent at the output is very intense and can be detected by high-efficiency linear photo-diodes (the overall quantum efficiency of the homodyne is originated by incomplete mode-matching between the signal and the LO). In this way, homodyning the density matrix becomes a perfect method for measuring the photon statistics. In addition, as for the diagonal matrix elements the phase of the LO is averaged out [the kernel in Eq. (30) does not depend on ϕ for $d = 0$], it is sufficient to collect data without worrying for the phase, with the only caution of having negligible phase-correlation between signal and LO: this simplifies the

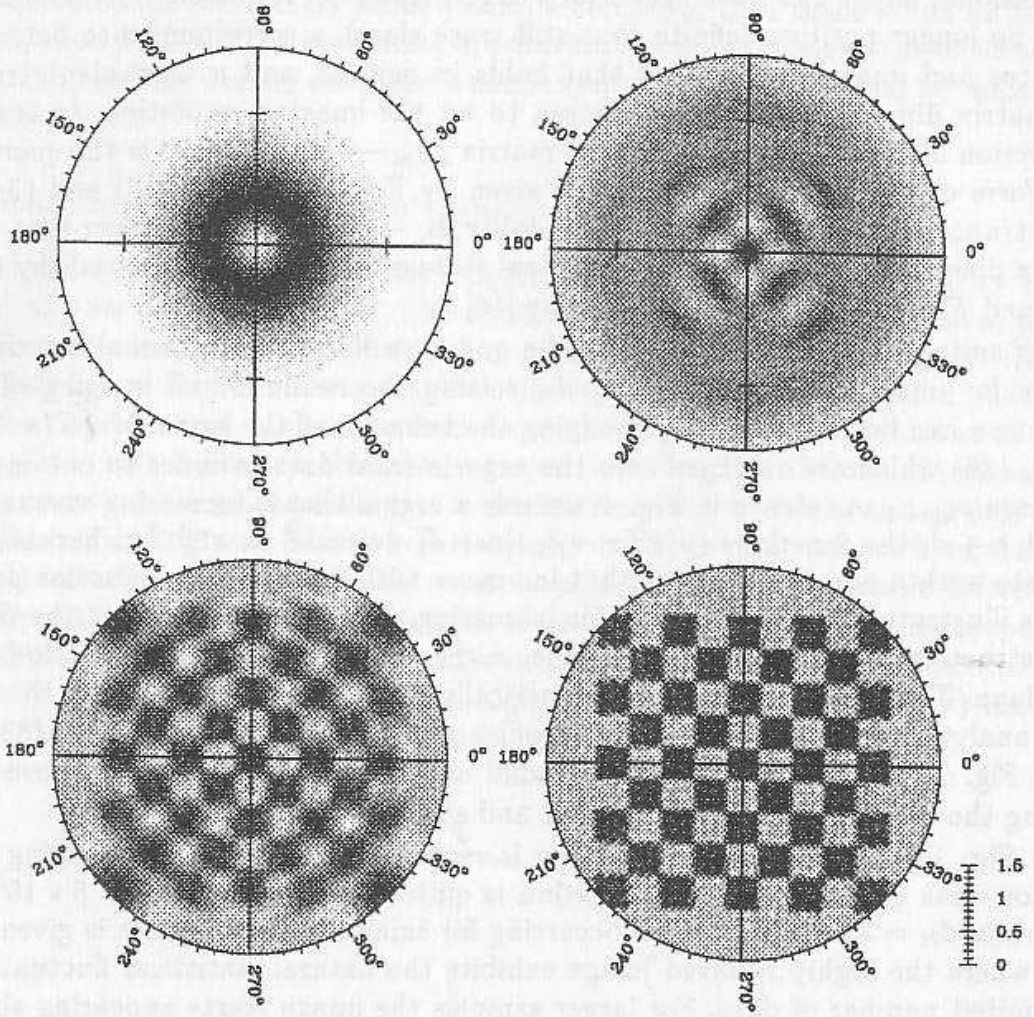

Figure 12. Tomographic reconstruction of a chess-board shaped step function $W(\alpha, \overline{\alpha})$ for increasing dimension of the truncated matrix, $d_{\mathcal{H}} = 2, 8, 32, 128$ from the top-left to the bottom-right. The plots are obtained by integrating numerically Eq. (18) with kernel (30), from analytically assigned transmission profiles $p(x, \phi)$, and then using Eqs. (33) and (34).

experimental method greatly. However, by no means such "homodyne photodetector" can be used to detect single events, but only statistics. That homodyning is a good tool to monitor delicate oscillations in photon statistics was already suggested in Refs. [12, 19, 20]. More recently, the new exact method has been adopted in real experiments: in Refs. [22, 23] it has been applied to measure ultrafast (subpicosecond) time-resolved photon statistics of weak fields.

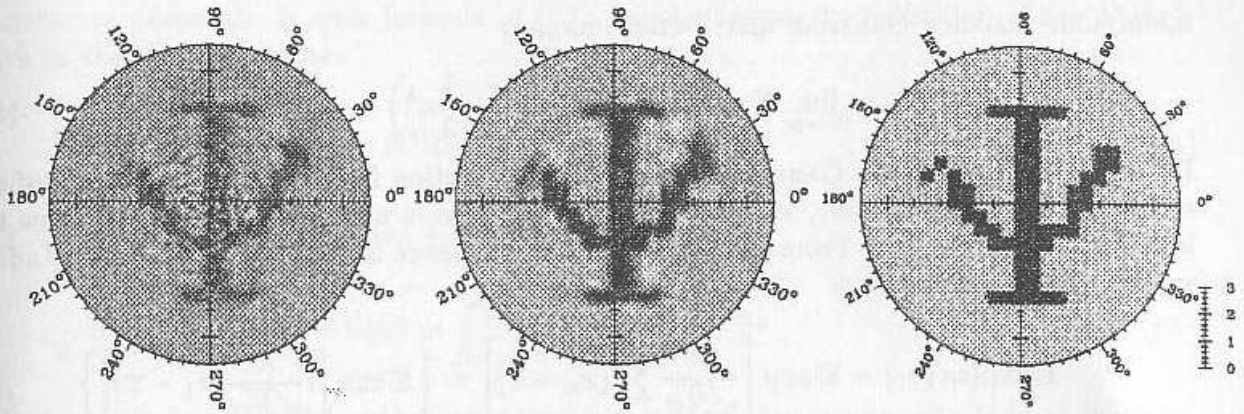

Figure 13. Monte Carlo simulation of a quantum imaging for weak signals of the font "ψ". The truncation dimension is fixed at $d_{\mathcal{H}} = 32$. The first plot on the left is the reconstruction from only 50000 data (50 phases and 1000 data each); the second plot is for $5 * 10^6$ data (100 blocks as before). The last plot is a reconstruction obtained by integrating numerically Eq. (18) with kernel (30) from analytically assigned transmission profiles $p(x, \phi)$, as in Fig. 12 (here $d_{\mathcal{H}} = 128$).

13. Appendices

13.1. PROOF OF THE CENTRAL LIMIT THEOREM

Generally speaking, the central limit theorem says that the probability distribution of the sum of a large number N of random variables $\{x_n\}_{n=1}^N$ approaches a Gaussian function. For application to the scheme of repeated measurements, we only need a simplified version of the theorem, with all variables x_n having the same probability distribution. In this case the proof of the theorem is relatively straightforward and is next given in detail.

Theorem: *Let $\{x_n\}_{n=1}^N$ be statistically independent random variables having the same probability distribution, with average $\overline{x}$, variance σ^2 and bounded third moment*

$$\mathbf{E}|x_n - \overline{x}|^3 < B < +\infty \,. \tag{56}$$

Then the random variable $X_N = \frac{1}{N}\sum_{n=1}^N x_n$ has average $\overline{x}$ and variance $\varepsilon_N^2 = \frac{\sigma^2}{N}$, and has a probability distribution approaching a Gaussian function asymptotically for $N \to \infty$.

Proof: The first part of the theorem is straightforward. One has that

$$\mathbf{E}X_N = \frac{1}{N}\sum_{n=1}^N \mathbf{E}x_n = \frac{1}{N} N\overline{x} \equiv \overline{x} \,. \tag{57}$$

Moreover, as a consequence of statistical independence of x_n one has

$$\begin{aligned} \mathbf{E}X_N^2 &= \frac{1}{N^2}\sum_{n=1}^N\sum_{m=1}^N \mathbf{E}[x_n x_m] = \frac{1}{N^2}\sum_{n\neq m=1}^N \mathbf{E}x_n\,\mathbf{E}x_m + \frac{1}{N^2}\sum_{n=1}^N \mathbf{E}x_n^2 \\ &= \frac{N-1}{N}\overline{x}^2 + \frac{1}{N}(\sigma^2 + \overline{x}^2) = \frac{1}{N}\sigma^2 + \overline{x}^2 \,, \end{aligned} \tag{58}$$

namely $\mathbf{E}[\Delta X_N^2] = \frac{\sigma^2}{N}$. The second part of the proof consists of showing that for $N \to \infty$

mean/unit-variance Gaussian distribution, namely

$$\lim_{N\to\infty} \mathbf{E}\exp(i\kappa\chi_N) = \exp\left(-\frac{1}{2}\kappa^2\right) . \tag{59}$$

In fact, the standardized Gaussian probability distribution is the Fourier transform of the characteristic function (59), and the Fourier transform is unique, and this will prove the last part of the theorem. From the statistical independence and equivalence of the random variables $\{x_n\}_{n=1}^N$ we have

$$\begin{aligned}\mathbf{E}\exp(i\kappa\chi_N) &= \mathbf{E}\exp\left[\frac{i\kappa}{\sqrt{N}\sigma}\sum_{n=1}^{N}(x_n-\overline{x})\right] = \left\{\mathbf{E}\exp\left[i\frac{\kappa}{\sqrt{N}\sigma}(x_1-\overline{x})\right]\right\}^N \\ &= \left\{\mathbf{E}\left[1+i\frac{\kappa}{\sqrt{N}\sigma}(x_1-\overline{x})+\frac{1}{2!}\left(i\frac{\kappa}{\sqrt{N}\sigma}\right)^2(x_1-\overline{x})^2+\frac{R_N}{N^{3/2}}\right]\right\}^N \\ &= \left\{1-\frac{\kappa^2}{2N}+\frac{\mathbf{E}R_N}{N^{3/2}}\right\}^N ,\end{aligned} \tag{60}$$

where by Taylor's formula $|\mathbf{E}R_N| \le C\frac{|\kappa|^3}{\sigma^3}\mathbf{E}|x_1-\overline{x}|^3$ for some constant C. Forming the natural logarithm we have

$$\log\mathbf{E}\exp(i\kappa\chi_N) = N\log\left(1-\frac{\kappa^2}{2N}+\frac{1}{N^{3/2}}\mathbf{E}R_N\right) = -\frac{\kappa^2}{2}+\mathcal{O}(N^{-1/2}) , \tag{61}$$

and hence Eq. (59). The theorem can be easily extended to more general situations, for example when the random variables $\{x_n\}_{n=1}^N$ are not statistically equivalent, but all of them have finite average and variance and satisfy the bound (56) for some common constant B. In this case for large N X_N has still Gaussian distribution, with mean value $\mathbf{E}X_N = \frac{1}{N}\sum_{n=1}^N \mathbf{E}x_n$ and variance $\mathbf{E}\Delta X_N^2 = \frac{1}{N}\sum_{n=1}^N \mathbf{E}[\Delta x_n^2]$. It is also easy to see that the theorem can be proven using, in place of the bound (56), the more general Lyapunov condition

$$\lim_{N\to\infty} \frac{\sum_{n=1}^N \mathbf{E}|x_n-\mathbf{E}x_n|^3}{\left\{\sum_{n=1}^N \mathbf{E}[\Delta x_n^2]\right\}^{3/2}} = 0 . \tag{62}$$

In fact, in this case the same asymptotic expansion (61) holds (see Ref. [37], pag. 78). Finally, the central limit theorem is known to apply to some cases of statistically dependent variables [38].

Necessary conditions for the theorem's validity are difficult to state: however, it is obvious that the result of the theorem does not hold true at least when the first two moments of the random variables $\{x_n\}_{n=1}^N$ are not bounded.

13.2. DERIVATION OF THE INVERSE RADON TRANSFORM FORMULA

The Radon transform of a function $m(x,y)$ in the complex plane is given by the collection of functions $r(x,\phi)$ $\phi\in[0,\pi]$, which are the marginal integrals of $m(x,y)$ along the direction with polar angle ϕ, namely

$$r(x,\phi) = \int_{-\infty}^{+\infty}\frac{\mathrm{d}y}{\pi}\, m\left((x+iy)e^{i\phi},(x-iy)e^{-i\phi}\right) . \tag{63}$$

In order to derive the inverse formula of (63), we start from the definition of the Dirac's delta in the complex plane

$$\delta^{(2)}(\alpha) = \int \frac{\mathrm{d}^2\gamma}{\pi^2} e^{\gamma\bar{\alpha}-\bar{\gamma}\alpha} , \tag{64}$$

and change to polar coordinates $\gamma = -\frac{i}{2}e^{-i\phi}k$, obtaining

$$\delta^{(2)}(\alpha) = \int_0^\infty \frac{\mathrm{d}k\,k}{4} \int_0^{2\pi} \frac{\mathrm{d}\phi}{\pi^2} e^{-ik\alpha_\phi} , \tag{65}$$

where $\alpha_\phi = \mathrm{Re}(\alpha e^{-i\phi})$. Using the symmetry $\alpha_{\phi+\pi} = -\alpha_\phi$ we can rewrite (65) as follows

$$\delta^{(2)}(\alpha) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k\,|k|}{4} \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi^2} e^{-ik\alpha_\phi} . \tag{66}$$

The last formula is *in nuce* the inverse Radon transform. In fact, it is just sufficient to write $m(\alpha, \bar{\alpha})$ identically as follows

$$m(\alpha, \bar{\alpha}) = \int \mathrm{d}^2\beta\, \delta^{(2)}(\alpha - \beta)\, m(\beta, \overline{\beta}) . \tag{67}$$

Then, using the polar representation (66) of the Dirac delta and exchanging the integral over β with those defining $\delta^{(2)}(\alpha)$ we obtain

$$m(\alpha, \bar{\alpha}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k\,|k|}{4} \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi^2} \int \mathrm{d}^2\beta\, e^{-ik(\alpha_\phi - \beta_\phi)}\, m(\beta, \overline{\beta}) . \tag{68}$$

Finally, upon rewriting $\beta \doteq (x + iy)e^{i\phi}$ and using Eq. (63), we obtain the inverse Radon transform

$$m(\alpha, \bar{\alpha}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k\,|k|}{4} \int_0^{\pi} \frac{\mathrm{d}\phi}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}x\; r(x, \phi) \exp\left[ik\,(x - \alpha_\phi)\right] . \tag{69}$$

13.3. FACTORIZATION FORMULA FOR THE INTEGRAL KERNEL

The numerical evaluation of the integral kernel as is written in Eq. (30) involves the calculation of parabolic cylinder functions, which can be connected to each other iteratively. However, for sufficiently large n and d (around 10) the iteration becomes numerically unstable, and the cylinder functions must be evaluated by series summation, slowing down the numerical calculation considerably. On the other hand, the series over ν in Eq. (30) has terms with oscillating signs and absolute value that increases dramatically with n and d. Again for sufficiently large n and d (20) the calculation becomes unstable, and in order to reach values as $n = 40$ one is forced to use extended precision. Recently, elegant factorization formulas have been found by Richter [17], Leonhardt *et al.* [18], and more recently derived by D'Ariano *et al.* in a completely algebraic way [21]. Such formulas provide an algorithm that is fast, low-memory, and stable for large n and d. Here I only report the recipe for numerical calculation: for the derivation of formulas the reader is addressed to the above quoted references.

The kernel functions—also called "pattern functions" in Refs. [17, 19, 20] or "sampling functions" in Ref. [18]—can be factorized as follows

$$\begin{aligned}&\langle n|K(x-\hat{x})|m\rangle = \langle m|K(x-\hat{x})|n\rangle \\ &= 2x u_n(x)v_m(x) - \sqrt{n+1}u_{n+1}(x)v_m(x) - \sqrt{m+1}u_n(x)v_{m+1}(x)\,,\end{aligned} \tag{70}$$

for $m \geq n$. In Eq. (70) only the case of $\phi = 0$ ($\hat{x} \equiv \hat{x}_\phi$) has been considered, as for $\phi \neq 0$ the trivial rotation factor $\exp(-id\phi)$ is just needed [see Eq. (30)]. The functions $u_n(x)$ (energy eigenfunctions of the harmonic oscillator) are connected each other by the (stable) iteration

$$u_n(x) = \frac{1}{\sqrt{n}}[2x u_{n-1}(x) - \sqrt{n-1}u_{n-2}(x)]\,, \tag{71}$$

with starting values

$$u_{-1}(x) = 0\,, \qquad u_0(x) = \exp\left(-x^2\right)\,. \tag{72}$$

The functions $v_m(x)$ ("irregular" harmonic oscillator eigenfunctions [18]) also satisfy recursion (71), but now this would be unstable, because it depends too critically on accuracy of $v_0(x)$, which would involve an imaginary error function. Instead, Ref. [18] suggests using the backward recursion

$$v_{m-2}(x) = [2x v_{m-1}(x) - \sqrt{m}v_m(x)]/\sqrt{m-1}\,, \tag{73}$$

starting from the semiclassical approximation for large numbers

$$v_m(x) = \left(\frac{8\pi}{\alpha_m^2 \sin^2\kappa_m}\right)^{1/4} \sin\left\{\frac{1}{2}\alpha_m^2\left[\sin(2\kappa_m) - 2\kappa_m\right] + \frac{\pi}{4}\right\}\,, \tag{74}$$

where

$$\alpha_m = \sqrt{m+1/2}\,, \qquad \kappa_m = \arccos\frac{x}{\alpha_m}\,. \tag{75}$$

For a density matrix with maximal number of photons N it is suggested to start the iteration (73) from $m = 4N$ and $m = 4N-1$. The classically allowed region for such quantum number is $|x| < \alpha_{4N} - \frac{1}{2}\alpha_{4N}^{-1/3}$, otherwise one can use the asymptotic expression for $v_m(x)$ which are provided by the forward recursion

$$v_m(x) = \frac{\sqrt{m}}{2x}v_{m-1}(x)\,, \tag{76}$$

with the starting condition

$$v_0(x) = \exp\left(x^2\right)/x\,. \tag{77}$$

All tomographic reconstructions in the present paper for $\eta = 1$ have been obtained using the above numerical recipe. For $\eta < 1$ a factorization formula is not yet known, and the

results here presented have been obtained by summing the series (30). In Ref. [39] the following Bernoulli transformation has been suggested

$$\langle n|\hat{\varrho}|m\rangle = \eta^{-\frac{1}{2}(n+m)} \sum_{k=0}^{\infty} \langle n+k|\hat{\varrho}|m+k\rangle^{(\eta)} \left[\binom{n+k}{n} \binom{m+k}{m} \right]^{1/2} , \tag{78}$$

to connect the desired density matrix elements $\langle n|\hat{\varrho}|m\rangle$ to the "smeared" ones $\langle n|\hat{\varrho}|m\rangle^{(\eta)}$ measured by inefficient homodyne detection with the kernel for unit efficiency (such "smeared" matrix elements are Bernoulli convolutions of the true matrix elements: an example has been given in Fig. 7). However, in practice the deconvolution (78) can be used only for efficiency η very near to one, and for small indices n and m, because the convergence of the series in Eq. (78) is very slow (geometric convergence with finite convergence radius). Moreover, this method is not numerically efficient, because one needs to sample many smeared matrix elements $\langle n|\hat{\varrho}|m\rangle^{(\eta)}$ in order to obtain only few true elements $\langle n|\hat{\varrho}|m\rangle$.

References

1. W. Pauli, in *Encyclopedia of Physics* (Springer, Berlin, 1958), Vol. V, p. 17.
2. For example, the spherical-harmonic eigenstates $f(r)P_l^m(\theta)e^{im\phi}$ of the angular momentum (whatever is the radial function) have probabilities $W(\mathbf{x})$ and $W(\mathbf{p})$ that are independent on the sign of m.
3. G. M. D'Ariano, *Quantum estimation theory and optical detection*, in this volume.
4. U. Fano, Rev. Mod. Phys. **29**, 74 (1957)
5. B. d'Espagnat, *Conceptual Foundations of Quantum Mechanics*, W. A. Benjamin, Massachusetts (1976)
6. A. Royer, Found. Phys. **19**, 3 (1989)
7. U. Leonhardt, Phys. Rev. Lett. **74**, 4101 (1995)
8. For a historical review on theoretical work and proposals for measuring quantum states see Ref. [6]. For more recent references see also Refs. [7] and [20].
9. D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993). See also D. T. Smithey, M. Beck, J. Cooper, and M. G. Raymer, Phys. Rev. A **48**, 3159 (1993) and references therein.
10. K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989). This Radon-transform connection between homodyne probability distributions and Wigner function was lately generalized to Fourier-transform relations between different representations of the density matrix in: H. Kühn, D.–G. Welsch, W. Vogel, J. Mod. Optics **41**, 1607 (1994).
11. G. M. D'Ariano, C. Macchiavello and M. G. A. Paris, Phys. Rev. A**50**, 4298 (1994).
12. G. M. D'Ariano, U. Leonhardt and H. Paul, Phys. Rev. A **52**, R1801 (1995)
13. G. M. D'Ariano, Quantum Semiclass. Opt. **7**, 693 (1995)
14. G. M. D'Ariano, C. Macchiavello and M. G. A. Paris, Phys. Lett. A **195**, 31 (1994)
15. G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris, Nuovo Cimento **110** B, 237 (1995)
16. G. M. D'Ariano, C. Macchiavello and M. G. A. Paris, Opt. Commun. (1996 in press)
17. Th. Richter, Phys. Lett. A (in press)
18. U. Leonhardt, M. Munroe, T. Kiss, M. G. Raymer and Th. Richter, Opt. Commun. (1996 in press)
19. H. Paul, U. Leonhardt, and G. M. D'Ariano, Acta Phys. Slov. **45**, 261 (1995)
20. U. Leonhardt, H. Paul and G. M. D'Ariano, Phys. Rev. A **52** 4899 (1995)
21. G. M. D'Ariano, C. Macchiavello and N. Sterpi, unpublished.
22. M. Munroe, D. Boggavarapu, M. E. Anderson, and M. G. Raymer, Phys. Rev. A **52**, R924 (1995)
23. M. Munroe, D. Boggavarapu, M. E. Anderson, U. Leonhardt, and M. G. Raymer, in *Coherence and Quantum Optics VII*, ed. by J. Eberly, L. Mandel, and E. Wolf (in press)
24. G. M. D'Ariano and H. P. Yuen, Phys. Rev. Lett. **76** 2832 (1996)
25. F. Natterer, *The mathematics of computerized tomography*, (Wiley, 1986).
26. P. Mansfield and P. G. Morris, *NMR Imaging in Biomedicine*, (Academic Press, 1982).
27. W. Schleich and J. A. Wheeler, Nature **326**, 574 (1987).
28. R. A. Fisher, M. M. Nieto, and V. D. Sandberg, Phys. Rev. D**29**, 110 (1984)
29. O. Alter, and Y. Yamamoto, Phys. Rev. Lett. **74**, 4106 (1995).
30. Y. Aharonov, J. Anandan, L. Vaidman, Phys. Rev. A **47**, 4616 (1993); Y. Aharonov and L. Vaidman, Phys. Lett. A **178**, 38 (1993).