Noise reduction in quantum tomography

J. Opt. B: Quantum Semiclass. Opt. **2** (2000) 113–117. Printed in the UK PII: S1464-4266(00)06597-6

Noise reduction in quantum tomography

Giacomo M D'Ariano and Matteo G A Paris

Theoretical Quantum Optics Group—INFM, Unità di Pavia, Dipartimento di Fisica 'Alessandro Volta', Università di Pavia, via Bassi 6, I-27100 Pavia, Italy

Received 2 August 1999, in final form 21 January 2000

Abstract. We review the 'adaptive tomography' technique, and study its application to the measurement of the field intensity and correlations, and to the reconstruction of the Wigner function of a single-mode radiation field. In both cases, this method strongly reduces statistical errors, thus making homodyne tomography a kind of low-noise universal detector.

Keywords: State reconstruction, quantum tomography

1. Introduction

Quantum homodyne tomography (QHT) is the most successful technique for the measurement of the quantum state of radiation. In QHT, in fact, the signal mode is amplified by the local oscillator, thus avoiding the need for single-photon resolving photodetectors, and hence providing the possibility of achieving very high quantum efficiency by linear photodiodes [4]. Moreover, QHT is efficient and statistically reliable, and can be used on-line with the experiment. Indeed, among other proposed state reconstruction methods, QHT is the only one which has been implemented in quantum optical experiments [1, 2].

Applications of QHT range from the measurement of photon correlations on a sub-picosecond timescale [1] to the characterization of squeezing properties [2], photon statistics in parametric fluorescence [3], quantum correlations in down-conversion [4], nonclassicality of states [5], and the measurement of Hamiltonians of nonlinear optical devices [6].

Actually, QHT can be used to estimate the expectation value of any operator of the field [7], which makes the method a kind of universal detector for radiation. On the other hand, the price to pay for such universality is the occurrence of large statistical errors, and indeed the tomographic determination of relevant field observables are generally more noisy than the corresponding direct detection [8]. However, in view of the wide range of applications of QHT, it is desirable to reduce the tomographic noise, in order to extract the maximum amount of information from the homodyne data sample.

Recently, we have developed an adaptive optimization technique to improve the precision of quantum homodyne tomography [9]. 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. In this paper, we review the adaptive tomography method and study its application to the measurement of the field intensity and correlations. In addition, as a further development in comparison with [9], we apply the adaptive method to the reconstruction of the Wigner function, which provides a clear picture of the quantum state as a whole. As we will see, this method strongly reduces the noise coming from statistical fluctuations, thus making homodyne tomography a low-noise universal detector. Moreover, in this paper, the evaluation of the density matrix and the reconstruction of the Wigner function have been improved in comparison with [9], since we also exploit unitary transformations of the state, such as displacement and squeezing, which reduces the average number of photons of the state, and take advantage of the adaptive method, which is more effective for low photon numbers.

In section 2 we review QHT and the adaptive method in order to reduce statistical fluctuations. In section 3 we apply adaptive tomography to the measurement of the field intensity and of intensity correlations, whereas in section 4 we consider the reconstruction of the Wigner function. Section 5 closes the paper with some concluding remarks.

2. Adaptive tomography

Quantum tomography of a single-mode radiation field consists of a set of repeated measurements of the field-quadrature $\hat{x}_\phi = \frac{1}{2}(a\mathrm{e}^{-\mathrm{i}\phi} + a^\dagger \mathrm{e}^{\mathrm{i}\phi})$ at different values of the reference phase ϕ. The expectation value of a generic operator can be obtained by averaging a suitable kernel function $R[\hat{O}](x,\phi)$ as follows [7]:

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

where $p(x,\phi)$ denotes the probability distribution of the outcomes x for the quadrature $\hat{x}_\phi$, and $R[\hat{O}](x,\phi)$ is given

by

$$R[\hat{O}](x,\phi) = \tfrac{1}{4}\int_0^\infty \mathrm{d}r\ \mathrm{Tr}\{\hat{O}\cos[\sqrt{r}(x-\hat{x}_\phi)]\}. \qquad (2)$$

From identity (1), it follows that the ensemble average $\langle\hat{O}\rangle$ can be experimentally obtained by averaging $R[\hat{O}](x,\phi)$ over the set of homodyne data, namely

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

where N is the total number of measurements of the sample. The statistical error of the tomographic measurement in equation (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 obtained 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

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

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

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

The crucial point of the adaptive tomography method is that the tomographic kernel $R[\hat{O}](x,\phi)$ is not unique, since a large class of *null functions* [9,10] $F(x,\phi)$ exists that have a zero tomographic average for an arbitrary state, namely

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

Therefore, the addition of null functions to a generic kernel gives a new kernel with the same tomographic average, hence equivalent for the estimation of the same ensemble average $\langle\hat{O}\rangle$. On the other hand, adding null functions would modify the kernel variance, whence the statistical error is over data. The adaptive tomography method thus consists in optimizing the kernel in the equivalence class, in order to minimize the statistical errors.

A general expression for the null functions of (6) is given by

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

In equations (7) $n \geqslant 0$ and $g_\pm(z)$ are analytic functions of z. The set $\mathcal{G}$ of null functions defined in equations (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 the following we will focus attention on null functions obtained from equations (7) by choosing $n = 0$ and $g(z) \equiv z^k$ for a given k. This will be denoted by $F_k(x,\phi)$, namely

$$F_k(x,\phi) = x^k\mathrm{e}^{\mathrm{i}(k+2)\phi} \qquad k = 0, 1, \ldots. \qquad (8)$$

Let us now consider a generic real kernel $R[\hat{O}](x,\phi)$. By adding M null functions keeping the kernel as real, we have the new kernel $K[\hat{O}](x,\phi)$

$$K[\hat{O}](x,\phi) = R[\hat{O}](x,\phi) + \sum_{k=0}^{M-1}\mu_k F_k(x,\phi) + \sum_{k=0}^{M-1}\mu_k^* F_k^*(x,\phi), \qquad (9)$$

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

$$\overline{\Delta K^2[\hat{O}]} = \overline{\Delta R^2[\hat{O}]} + 2\left\{\sum_{k,l=0}^{M-1}\mu_k\mu_l^*\overline{F_kF_l^*} + \sum_{k=0}^{M-1}\mu_k\overline{R[\hat{O}]F_k} + \sum_{k=0}^{M-1}\mu_k^*\overline{R[\hat{O}]F_k^*}\right\}. \qquad (10)$$

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

The variance of the modified kernel function in equation (10) can be minimized with respect to the coefficients μ_k, leading to the linear set of equations

$$\sum_l \mu_l\overline{F_kF_l^*} = -\overline{R[\hat{O}]F_k^*}. \qquad (11)$$

It is convenient to rewrite the optimization equations (11) in matrix form as follows:

$$\boldsymbol{A}\mu = \boldsymbol{b} \qquad (12)$$

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

$$A_{kl} = \overline{F_kF_l^*} = \int_0^\pi \frac{\mathrm{d}\phi}{\pi}\int_{-\infty}^{\infty}\mathrm{d}x p(x,\phi)F_k(x,\phi)F_l^*(x,\phi), \qquad (13)$$

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

$$b_k = -\overline{R[\hat{O}]F_k^*} = -\int_0^\pi \frac{\mathrm{d}\phi}{\pi}\int_{-\infty}^{\infty}\mathrm{d}x p(x,\phi)R[\hat{O}](x,\phi)F_k^*(x,\phi). \qquad (14)$$

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

By substituting equation (11) in (10) and inverting (12) we obtain

$$\Delta^2[\hat{O}] \doteq \overline{\Delta R^2[\hat{O}]} - \overline{\Delta K^2[\hat{O}]} = 2\sum_{k,l=0}^{M-1}\mu_k A_{kl}\mu_l^* = 2\sum_{k,l=0}^{M-1} b_k(A^{-1})_{kl}b_l^* \geqslant 0, \qquad (15)$$

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

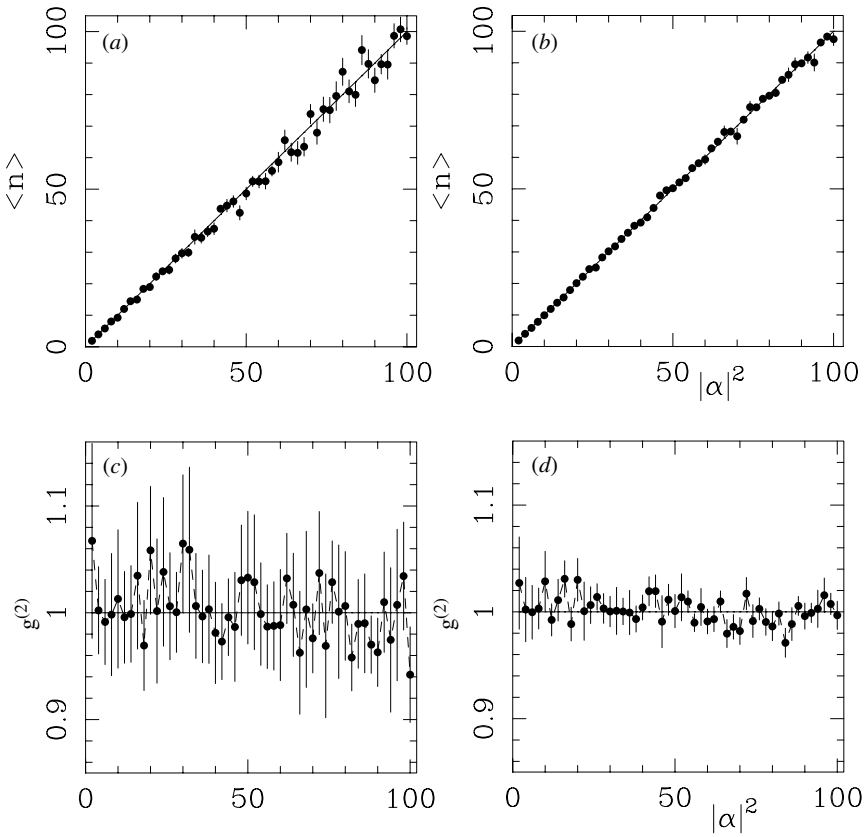

Figure 1. Monte Carlo simulation of tomographic measurement of the intensity $\langle a^\dagger a\rangle$ and of the second-order correlation function $g^{(2)} = \langle a^{\dagger 2}a^2\rangle/\langle a^\dagger a\rangle^2$ on coherent states. (*a*),(*c*) The determination by usual tomography as a function of the average photon number $|\alpha|^2$ of the coherent state. (*b*),(*d*) The determination after the adaptive optimization of the kernel functions. Simulated experiments have been performed using 5000 random-phase homodyne data, divided into ten blocks in order to evaluate statistical errors.

The optimization procedure for the kernel $R[\hat{O}](x,\phi)$ can be summarized as follows: after collecting an ensemble of N tomographic data, the quantities $\boldsymbol{A}$ and $\boldsymbol{b}$ are evaluated as tomographic experimental averages. Then, by solving the linear system (12) one obtains the coefficients μ_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 $\boldsymbol{b}$ and $\boldsymbol{A}$) and on the operator $\hat{O}$, whose kernel enters only in the expression of $\boldsymbol{b}$.

3. Field intensity and intensity correlations

In this section we apply the adaptive tomography method to the detection of the field intensity $I = \langle a^\dagger a\rangle$ and the second-order correlation function $g^{(2)} = \langle a^{\dagger 2}a^2\rangle/\langle a^\dagger a\rangle^2$. As a starting point for the present optimization procedure we take the Richter form of the tomographic kernel for the normally ordered moment [11]

$$R[a^{\dagger n}a^m](x;\phi) = \mathrm{e}^{\mathrm{i}(m-n)\phi}\frac{H_{n+m}(\sqrt{2}x)}{\sqrt{2^{n+m}\binom{n+m}{n}}}, \qquad (16)$$

where $H_n(x)$ is the Hermite polynomial of order n. In the case of tomographic detection of intensity, equation (16) provides the kernel

$$R[a^\dagger a](x) = 2x^2 - \tfrac{1}{2}, \qquad (17)$$

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

$$b_k = -\overline{R[a^\dagger a]F_k^*} = \overline{2x^{k+2}\mathrm{e}^{-\mathrm{i}(k+2)\phi}} = \frac{\langle a^{\dagger(k+2)}\rangle}{2^{1+k}}. \qquad (18)$$

We solved the optimization equation (12) analytically for coherent states, squeezed vacuum and the 'cat' superposition of coherent states $|\psi\rangle = [2(1+\exp\{-2|\alpha|^2\})]^{-1/2}(|\alpha\rangle+|-\alpha\rangle)$ up to $M = 10$ null functions. For all the states considered here, it turns out that only the null function $F_0(\phi)$ is needed, namely one has

$$\mu_0 = b_0 \qquad \mu_k = 0, \qquad \forall k \geqslant 1. \qquad (19)$$

The corresponding reduction of variance is given by

$$\Delta^2[a^\dagger a] = \tfrac{1}{2}\langle a^{\dagger 2}\rangle\langle a^2\rangle, \qquad (20)$$

and can compensate the leading term of the variance of the original Richter kernel

$$\overline{\Delta R^2[a^\dagger a]} = \langle\widehat{\Delta n^2}\rangle + \tfrac{1}{2}[\langle a^{\dagger 2}a^2\rangle + 2\langle a^\dagger a\rangle + 1], \qquad (21)$$

so that $\overline{\Delta K^2[a^\dagger a]}$ becomes much closer to the intrinsic intensity fluctuations $\langle\widehat{\Delta n^2}\rangle$ than the original noise

$\overline{\Delta R^2[a^\dagger a]}$. 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

$$e^{i2\phi} R[a^{\dagger n} a^n](x) = \frac{n}{n+1} R[a^{\dagger(n+1)} a^{n-1}](x), \qquad (22)$$

which leads to

$$b_0 = -\overline{R[a^{\dagger n} a^n] e^{i2\phi}} = -\frac{n}{n+1} \langle a^{\dagger(n+1)} a^{n-1} \rangle, \qquad (23)$$

namely $\Delta^2[a^{\dagger n} a^n] = 2|b_0|^2$. Of course, the case $n = 2$ is of interest to measure the correlation function $g^{(2)}$. In figure 1 we show the results from a Monte Carlo simulation of the tomographic detection of intensity and correlations on coherent states: the noise reduction obtained by using the modified kernel is apparent.

4. Reconstruction of the Wigner function

Adaptive tomography can also be applied to the measurement of the density matrix element $\varrho_{nm} = \langle m|\hat{\varrho}|n\rangle$ of the quantum state of the field. Indeed, in [9] the method has been analysed in detail for coherent states, Fock states, squeezed vacuum, and 'Schrödinger-cat' states. With the exception of Fock states, the method generally provides a sizeable reduction of statistical errors. For coherent states the improvement mainly concerns matrix elements with small index, whereas for squeezed vacuum and cat states generic off-diagonal elements are also improved. In order to see the effectiveness of the method in the reconstruction of the quantum state as a whole, here we consider the application to the reconstruction of the Wigner function of the field, which is defined as follows:

$$W(z) = \frac{2}{\pi} \mathrm{Tr}\{\hat{\varrho} \hat{D}(2z)(-)^{a^\dagger a}\}, \qquad (24)$$

and can be expressed in terms of the matrix elements as

$$W(z) = \mathrm{Re} \sum_{d=0}^{\infty} e^{id\phi} \sum_{n=0}^{\infty} \Lambda(n, d; |z|^2) \rho_{n,n+d} \qquad (25)$$

where

$$\Lambda(n, d; |z|^2) = (-)^n 2(2 - \delta_{d0}) |2z|^d \sqrt{\frac{n!}{(n+d)!}} e^{-2|z|^2} L_n^d(|2z|^2), \qquad (26)$$

and where $L_n^d(x)$ denotes the Laguerre polynomials. Of course, the series in equation (25) should be truncated at some point, and therefore the Wigner function can be reconstructed only at some finite resolution. Remarkably, adaptive tomography largely improves the precision of the reconstruction at fixed truncation dimension. Since the adaptive method is more effective for low photon numbers, we can improve it through the following steps: (i) the homodyne data are first used to adaptively estimate the complex field amplitude and the squeezing parameter; (ii) these values are used to perform unitary transformations on the data sample, by unsqueezing and displacing data towards the origin of the phase space, where adaptive tomography is more effective; (iii) the inverse transformations (displacement and squeezing) are performed on the reconstructed state. We applied this procedure to a large variety of states, and the results indicate that the method allows us to reconstruct the Wigner function also starting from a small sample of homodyne data collected by low-efficiency detectors. In figure 2 we illustrate the application of adaptive tomography to the reconstruction of

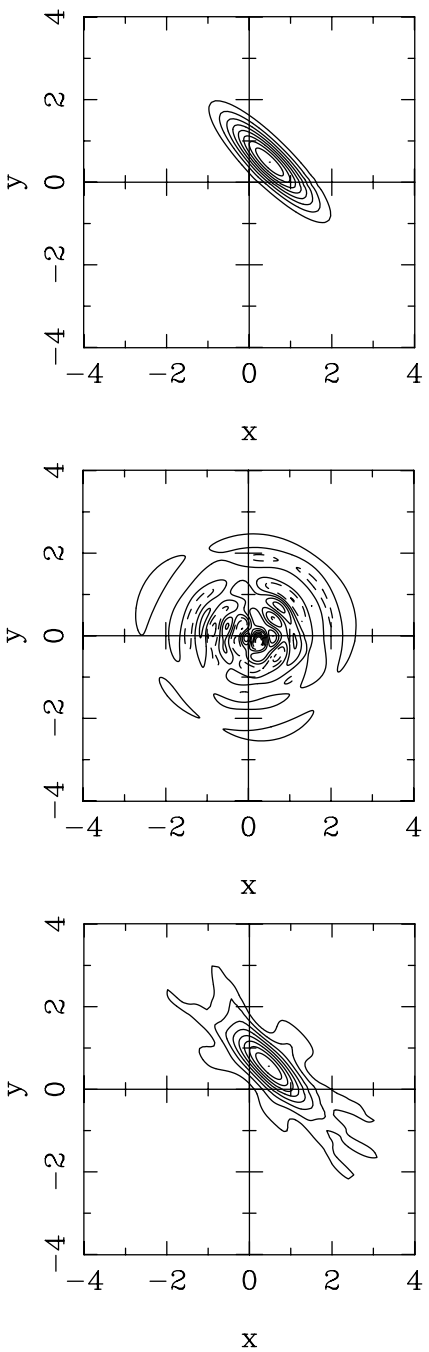

Figure 2. Monte Carlo simulation of the tomographic reconstruction of the Wigner function of a squeezed state. From top to bottom: the ideal Wigner function, the reconstruction by usual tomography, and the reconstruction obtained by the adaptive tomography method assisted by unitary displacement and unsqueezing of homodyne data (see text). Simulated experiments have been performed using ten blocks of 100 phases and 50 data each (for a total number of data equal to 50 000).

the Wigner function of a squeezed state $|\alpha, r\rangle$ with average photon number $\langle a^\dagger a\rangle = |\alpha| + \sinh^2 r = 1$, and squeezing fraction $\sinh^2 r/\langle a^\dagger a\rangle = \frac{1}{2}$. The simulated sample of figure 2 consists of 50 000 data coming from detectors with quantum efficiency $\eta = 75\%$. The improvement in the reconstruction is apparent.

5. Conclusion

In this paper, we have applied the adaptive tomography technique to the measurement of the field intensity and correlations, and to the reconstruction of the Wigner function of a single-mode of the radiation field. In both cases, the new technique strongly reduces statistical errors, thus providing a precise quantum state measurement. The method works equally well for nonunit quantum efficiency, and for a wide class of field observables [10], thus making homodyne tomography a low-noise universal detector.

References

[1] McAlister D F and Raymer M G 1997 *Phys. Rev.* A **55** R1609
[2] Breitenbach G, Schiller S and Mlynek J 1997 *Nature* **387** 471
[3] Vasylìev M, Choi S-K, Kumar P and D'Ariano G M 1998 *Opt. Lett.* **23** 1393
[4] D'Ariano G M, Vasylìev M and Kumar P 1998 *Phys. Rev.* A **58** 636
[5] D'Ariano G M, Sacchi M F and Kumar P 1999 *Phys. Rev.* A **59** 826
[6] D'Ariano G M and Maccone L 1998 *Phys. Rev. Lett.* **80** 5465
[7] D'Ariano G M 1997 *Quantum Communication, Computing, and Measurement* ed O Hirota, A S Holevo and C M Caves (New York: Plenum) p 253
[8] D'Ariano G M and Paris M G A 1997 *Phys. Lett.* A **233** 49
[9] D'Ariano G M and Paris M G A 1999 *Phys. Rev.* A **60** 518
[10] D'Ariano G M and Paris M G A 1998 *Acta Phys. Slov.* **48** 191
[11] Richter Th 1996 *Phys. Lett.* A **221** 327