RAPID COMMUNICATIONS

PHYSICAL REVIEW A, VOLUME 63, 020101(R)

Maximum-likelihood estimation of quantum processes

Jaromír Fiurášek and Zdeněk Hradil
Department of Optics, Palacký University, 17. listopadu 50, 772 07 Olomouc, Czech Republic
(Received 29 September 2000; published 9 January 2001)

Maximum-likelihood estimation is applied to identification of an unknown quantum-mechanical process. In contrast to linear reconstruction schemes, the proposed approach always yields physically sensible results. Its feasibility is demonstrated by performing the Monte Carlo simulations for the two-level system (single qubit).

DOI: 10.1103/PhysRevA.63.020101 PACS number(s): 03.65.Ta, 03.67.−a

During recent years a great deal of attention has been devoted to the *measurement of the quantum state* of various simple quantum-mechanical systems. All proposed reconstruction techniques follow the common underlying strategy: a set of measurements is performed on many identically prepared copies of the quantum state that is then estimated from the collected data. Feasible reconstruction schemes were devised for a wide variety of systems including the modes of running electromagnetic field (optical homodyne tomography [1,2] and unbalanced homodyning [3]), cavity electromagnetic field [4,5], motional state of the ion in a Paul trap [6], vibrational state of the molecule [7], and spin [8].

These significant achievements stimulated the development of a new remarkable branch of the reconstruction techniques that allow for the experimental determination of the unknown *quantum-mechanical processes* [9–14]. This is of great practical importance because such a technique may be used to experimentally evaluate the performance of the two-bit quantum gate—a building block of quantum computers [9]. The usual setup considered also in this paper is shown in Fig. 1. The input state prepared by an experimentalist and characterized by a density matrix $\varrho_{\rm in}$ enters the ''black box'' where it is transformed into the output $\varrho_{\rm out}$. The task for the experimentalist is to retrieve information on the physical process hidden in the black box from the measurements on the output states $\varrho_{\rm out}$ obtained from various input states $\varrho_{\rm in}$. The assumption taken for granted here is that the mapping $\varrho_{\rm out}=\mathcal{G}\varrho_{\rm in}$ is *linear*, as dictated by the linearity of quantum mechanics,

$$\varrho_{{\rm out},ij}=\sum_{kl} \mathcal{G}_{ij}^{kl}\varrho_{{\rm in},kl}. \tag{1}$$

Here $\varrho_{ij}=\langle i|\varrho|j\rangle$ are density matrix elements in some complete orthogonal basis of states spanning the Hilbert space on which the density operator ϱ acts. As illustrated in Fig. 1, the system may be entangled with the environment and the transformation $\mathcal{G}$ need not preserve the purity of the state. The Green superoperator $\mathcal{G}$ can describe a diverse variety of the physical processes, such as unitary evolution, damping, and decoherence. From the reconstructed superoperator $\mathcal{G}$ one may further estimate the Liouville superoperator $\mathcal{L}$, which governs the evolution of the density matrix in the black box, $\dot{\varrho}=\mathcal{L}\varrho$. If the superoperator $\mathcal{L}$ exists, then $\mathcal{G}=\exp(\mathcal{L}\tau)$, where τ is the interaction time, and an inversion of this relation yields $\mathcal{L}$ [11,12].

The estimation of the elements $\mathcal{G}_{ij}^{kl}$ by means of linear algorithms has been addressed in several papers [9–11]. The appropriate quantum-state reconstruction technique is employed to estimate the output states $\varrho_{\rm out}^{(m)}$ corresponding to several different input states $\varrho_{\rm in}^{(m)}$, and the unknown parameters $\mathcal{G}_{ij}^{kl}$ in Eq. (1) are then obtained by solving the system of linear equations. This linear reconstruction procedure is simple and straightforward, but it suffers from one significant drawback. The elements $\mathcal{G}_{ij}^{kl}$ are estimated as a set of seemingly unrelated numbers. However, $\mathcal{G}_{ij}^{kl}$ cannot be arbitrary because the linear mapping $\mathcal{G}$ must preserve the positive semidefiniteness and trace of the density matrix. These conditions impose bounds on the allowed values of $\mathcal{G}_{ij}^{kl}$. In this Rapid Communication the superoperator $\mathcal{G}$ is reconstructed using maximum-likelihood (ML), which allows the natural incorporation of all the constraints of quantum theory. Since one can only collect a finite amount of data, the linear mapping cannot be determined exactly. In accordance with the probabilistic interpretation of the quantum theory, the ML estimation answers the question, ''*which process is most likely to yield the measured data?*''

Due to its nonlinearity, the ML estimation is computationally a much more expensive task than the linear procedures. This is the prize for the physically sound result. ML estimation has been applied to various problems recently: to the measurements of the quantum phase shift [15], a coupling constant between atom and a cavity electromagnetic field [16], and the parameters of the quantum-optical Hamiltonian [17]. Reconstruction of the generic quantum state using the ML estimation and its interpretation as quantum measurement has been proposed in [18]. Subsequent Monte Carlo simulations performed for the quantum states of electromag-

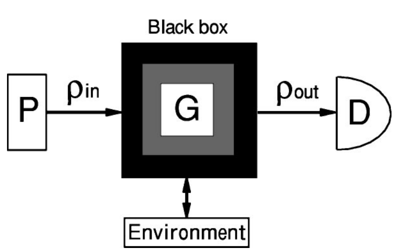

FIG. 1. Sketch of experimental setup for determination of the quantum-mechanical process. The input state $\varrho_{\rm in}$ is prepared in the preparator P and enters the black box where it is transformed to the output state $\varrho_{\rm out}=\mathcal{G}\varrho_{\rm in}$, which may be entangled with the environment. The detector D measures some observable of the output $\varrho_{\rm out}$.

netic field modes and spin [19–21] illustrated the feasibility of this technique. Here we shall demonstrate that the ML estimation is also suitable for determination of the generic quantum-mechanical processes.

The sought after superoperator $\mathcal{G}$ can be determined as the superoperator maximizing a likelihood function $\mathcal{L}[\mathcal{G}]$. The measurements performed on the output states $\varrho_{\rm out}^{(m)}$ can be described by positive-operator-valued measures (POVM) $\Pi^{(m)}$. If the experimental data contain n different combinations of detected POVMs $\Pi^{(m)}$ and corresponding input states $\varrho_{\rm in}^{(m)}$, $m=1,\ldots,n$, then $\mathcal{L}[\mathcal{G}]$ reads

$$\mathcal{L}[\mathcal{G}]=\prod_{m=1}^{n}\left(\mathrm{Tr}[\Pi^{(m)}\varrho_{\rm out}^{(m)}]\right)^{f_m}=\prod_{m=1}^{n}\left(\sum_{ijkl}\Pi_{ij}^{(m)}\mathcal{G}_{ji}^{kl}\varrho_{{\rm in},kl}^{(m)}\right)^{f_m},\tag{2}$$

where f_m is the (relative) frequency of the combination $(\varrho_{\rm in}^{(m)},\Pi^{(m)})$ in the data. The maximum of the likelihood function should be found in the domain of physically allowed superoperators $\mathcal{G}$, whose determination is crucial for the successful implementation of the ML estimation. The *linear positive map* (1) can be conveniently cast into the form that explicitly preserves the positive semidefinitness of the density matrix [10],

$$\varrho_{\rm out}=\sum_i A_i\varrho_{\rm in}A_i^{\dagger}.\tag{3}$$

It follows from the condition $\mathrm{Tr}\,\varrho_{\rm out}=1$ that

$$\sum_i A_i^{\dagger}A_i=I,\tag{4}$$

where I denotes the identity operator. Further we can expand A_i in some complete operator basis $\tilde{A}_j$,

$$A_i=\sum_j c_{ij}\tilde{A}_j.\tag{5}$$

If we deal with the N level system $|i\rangle$, $i=0,\ldots,N-1$, then it is natural to choose the N^2 basis operators as

$$\tilde{A}_{Ni+j}=|i\rangle\langle j|,\quad i,j=0,\ldots,N-1.\tag{6}$$

Inserting Eq. (5) into Eq. (3), we find that

$$\varrho_{\rm out}=\sum_{jk}\chi_{jk}\tilde{A}_j\varrho_{\rm in}\tilde{A}_k^{\dagger},\tag{7}$$

where

$$\chi_{jk}=\sum_i c_{ij}c_{ik}^{*},\quad j,k=0,\ldots,N^2-1.\tag{8}$$

Thus χ is a positive semidefinite Hermitian matrix [10]. This is the desired condition revealing a domain of the allowed parameters $\mathcal{G}_{ij}^{kl}$ (or alternatively χ_{ij}). The matrix χ is parametrized by N^4 real numbers, but the condition (4) imposes N^2 real constraints so that the number of independent parameters reads N^4-N^2. The relation between χ and $\mathcal{G}$ can be found by comparing Eqs. (1) and (7), $\mathcal{G}_{ij}^{kl}=\chi_{iN+k,jN+l}$, and the constraints (4) can be equivalently written as

$$\sum_i \mathcal{G}_{ii}^{kl}=\delta_{kl}.\tag{9}$$

The minimal parametrization can easily be achieved if one makes use of Eq. (9) and expresses N^2 real parameters in terms of the remaining N^4-N^2 ones.

To provide an explicit example, let us consider a two-level system (single qubit). The matrix χ can be expressed in terms of $\mathcal{G}_{ij}^{kl}$ as follows:

$$\chi=\begin{pmatrix}\mathcal{G}_{00}^{00} & \mathcal{G}_{00}^{01} & \mathcal{G}_{01}^{00} & \mathcal{G}_{01}^{01}\\ \mathcal{G}_{00}^{10} & \mathcal{G}_{00}^{11} & \mathcal{G}_{01}^{10} & \mathcal{G}_{01}^{11}\\ \mathcal{G}_{10}^{00} & \mathcal{G}_{10}^{01} & \mathcal{G}_{11}^{00} & \mathcal{G}_{11}^{01}\\ \mathcal{G}_{10}^{10} & \mathcal{G}_{10}^{11} & \mathcal{G}_{11}^{10} & \mathcal{G}_{11}^{11}\end{pmatrix},\tag{10}$$

and the constraints (4) yield $\mathcal{G}_{11}^{kl}=\delta_{kl}-\mathcal{G}_{00}^{kl}$. Thus χ is parametrized by $16-4=12$ real parameters that can be collected in a vector

$$\vec{G}=(\mathcal{G}_{00}^{00},\mathcal{G}_{00}^{11},\mathrm{Re}\,\mathcal{G}_{00}^{01},\mathrm{Im}\,\mathcal{G}_{00}^{01},\mathrm{Re}\,\mathcal{G}_{01}^{00},\mathrm{Im}\,\mathcal{G}_{01}^{00},\mathrm{Re}\,\mathcal{G}_{01}^{10},\mathrm{Im}\,\mathcal{G}_{01}^{10},\mathrm{Re}\,\mathcal{G}_{01}^{01},\mathrm{Im}\,\mathcal{G}_{01}^{01},\mathrm{Re}\,\mathcal{G}_{01}^{11},\mathrm{Im}\,\mathcal{G}_{01}^{11}).\tag{11}$$

Note that $\mathcal{G}_{ij}^{kl}=(\mathcal{G}_{ji}^{lk})^{*}$ since χ is Hermitian. The positive semidefiniteness of the matrix (10) can be easily checked for each $\mathcal{G}$ where the likelihood function (2) is evaluated. If the matrix χ is not positive semidefinite, then one may simply put $\mathcal{L}[\mathcal{G}]=0$. The maximum of $\mathcal{L}$ can be found for example with the help of the downhill-simplex algorithm. In the case of a two-level system it is sufficient to search for the maximum in the finite volume subspace of 12-dimensional space.

Alternatively, one can find the maximum of $\mathcal{L}[\mathcal{G}]$ from the extremum condition. It is convenient to work with the log-likelihood function. The constraints (9) must be incorporated by introducing N^2 (complex) Lagrange multipliers $\lambda_{mn}=\lambda_{nm}^{*}$. Assume first that the maximum of $\mathcal{L}[\mathcal{G}]$ is located inside the domain of physical superoperators $\mathcal{G}$, hence all eigenvalues of the estimated χ are positive. The extremum conditions then read

$$\frac{\partial}{\partial\mathcal{G}_{ij}^{kl}}\left[\ln\mathcal{L}[\mathcal{G}]-\sum_{mn}\lambda_{mn}\sum_p\mathcal{G}_{pp}^{mn}\right]=0.\tag{12}$$

On inserting the explicit expression for the likelihood function (2) into Eq. (12) one obtains

$$\lambda_{kl}\delta_{ab}=\sum_m\frac{f_m}{p_m}\Pi_{ba}^{(m)}\varrho_{{\rm in},kl}^{(m)},\tag{13}$$

where we have introduced

$$p_m = \mathrm{Tr}\left[\sum_i A_i \varrho_{\mathrm{in}}^{(m)} A_i^\dagger \Pi^{(m)}\right] = \mathrm{Tr}(\Pi^{(m)} \mathcal{G} \varrho_{\mathrm{in}}^{(m)}). \quad (14)$$

As follows from Eq. (13), λ is a positive definite Hermitian matrix. Equation (12) may be rewritten to the form suitable for iterative solution. Multiplying Eq. (13) by $(\lambda^{-1})_{ln}\mathcal{G}_{ac}^{kp}$ and summing over a,k,l, one gets

$$\mathcal{G}_{bc}^{np} = \sum_m \frac{f_m}{p_m} \sum_{a,k,l} \Pi_{ba}^{(m)} \varrho_{\mathrm{in,kl}}^{(m)} (\lambda^{-1})_{ln} \mathcal{G}_{ac}^{kp}. \quad (15)$$

The convenient form of Lagrange multipliers λ_{mn} may be found by inserting Eq. (15) into Eq. (9),

$$\lambda_{ij} = \sum_m \frac{f_m}{p_m} \sum_{a,k,p} \Pi_{ka}^{(m)} \mathcal{G}_{ak}^{pi} \varrho_{\mathrm{in},pj}^{(m)}. \quad (16)$$

Notice that $\mathrm{Tr}\,\lambda \equiv \Sigma_j \lambda_{jj} = \Sigma_m f_m = 1$. The system of nonlinear equations (15) and (16) for the elements of $\mathcal{G}$ can be conveniently solved by repeated iterations, starting from $\mathcal{G}_{ij}^{kl} = N^{-1}\delta_{ij}\delta_{kl}$ and keeping only the Hermitian part of $\mathcal{G}$ at each iteration step.

Let us now formulate the theory in terms of the operators A_i, $A_i^\dagger$. It is helpful to define a Hermitian operator $\lambda = \Sigma_{mn}\lambda_{mn}|m\rangle\langle n|$. The maximum of log-likelihood function can be formally found as the relation

$$\frac{\partial}{\partial A_i^\dagger}\left(\ln \mathcal{L}[\{A_j\}] - \mathrm{Tr}\left[\lambda \sum_j A_j^\dagger A_j\right]\right) = 0. \quad (17)$$

On performing the differentiation with respect to $A_i^\dagger$, and solving for A_i, we obtain

$$A_i = \sum_m \frac{f_m}{p_m} \Pi^{(m)} A_i \varrho_{\mathrm{in}}^{(m)} \lambda^{-1}. \quad (18)$$

The validity range of this formula includes the cases when the maximum of $\mathcal{L}[\mathcal{G}]$ lies at the boundary of the domain of allowed superoperators $\mathcal{G}$, where several eigenvalues of χ are zero and the corresponding eigenvectors, i.e., operators A_i, vanish. If we multiply Eq. (18) from the left by operator $A_i^\dagger$, sum over i, and take into account the constraint (4), we find

$$\lambda = \sum_m \frac{f_m}{p_m} \sum_i A_i^\dagger \Pi^{(m)} A_i \varrho_{\mathrm{in}}^{(m)}, \quad (19)$$

which is equivalent to Eq. (16). Similarly, one can also derive Eq. (15) from Eq. (18). Since Eqs. (15) and (16) follow from Eq. (18), they always provide correct and reliable estimates.

In the case when all eigenvalues of the estimated χ are nonzero, the procedure of ML estimation may be interpreted as a generalized measurement. To show this explicitly, let us calculate the trace of Eq. (13),

$$\mathrm{Tr}\,\lambda\,\delta_{ab} = \sum_m \frac{f_m}{p_m} \Pi_{ba}^{(m)}\ \mathrm{Tr}\,\varrho_{\mathrm{in}}^{(m)}. \quad (20)$$

Since all the traces are equal to 1, this relation reads in the operator form

$$\sum_m \frac{f_m}{p_m} \Pi^{(m)} = I, \quad (21)$$

which is the closure relation for renormalized positive-valued-operator measures $\Pi'^{(m)} = (f_m/p_m)\Pi^{(m)}$. Moreover, in spite of the fact that the exact agreement between theoretical probabilities p_m and measured relative frequencies f_m (assumed by standard reconstructions) cannot be achieved in general, the probabilities obtained from the renormalized POVMs $\Pi'^{(m)}$ are identical to f_m,

$$p'_m \equiv \mathrm{Tr}\left[\sum_i A_i \varrho_{\mathrm{in}}^{(m)} A_i^\dagger \Pi'^{(m)}\right] \equiv f_m. \quad (22)$$

This indicates the privileged role of ML estimation in analogy with the quantum state estimation [18]. ML estimation represents a genuine quantum measurement. Properties of a quantum black box are determined using the closure relation (21) for a POVM, expectation values of which are the registered data (22). Since the data are noisy, in general, this cannot be done using the linear algorithm of standard reconstruction schemes.

In the rest of the paper we demonstrate the feasibility of our approach by means of Monte Carlo simulations for two-level system (a single qubit). We shall consider the spin-1/2 system. The detector *D* shown in Fig. 1 is the Stern-Gerlach apparatus measuring the spin projections along one of three axes *x*, *y*, *z*. We further assume that ϱ_{in} is prepared in one of six eigenstates $|\uparrow_j\rangle$, $|\downarrow_j\rangle$ of the spin projectors (Pauli matrices) σ_j, $j = x,y,z$, $\sigma_j|\uparrow_j\rangle = |\uparrow_j\rangle$, and $\sigma_j|\downarrow_j\rangle = -|\downarrow_j\rangle$. We choose the basis $|0\rangle = |\downarrow_z\rangle$ and $|1\rangle = |\uparrow_z\rangle$. Each of the six input states is prepared $3\mathcal{N}$ times. At the output, one measures $\mathcal{N}$ times the spin along each of the three axes x,y,z. The corresponding six projectors read $\Pi_j = |j\rangle\langle j|$, $j \in \{\uparrow_x, \downarrow_x, \uparrow_y, \downarrow_y, \uparrow_z, \downarrow_z\}$. Let f_{jk} denote the relative fre-

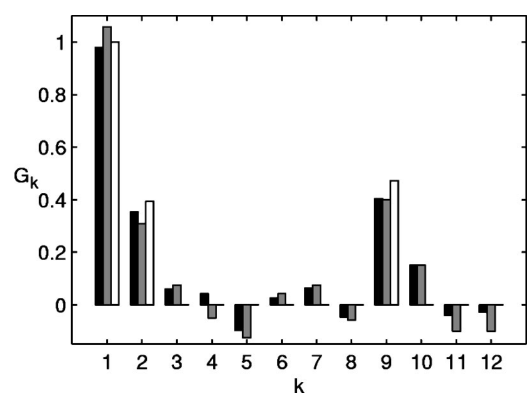

FIG. 2. Reconstructed dimensionless elements of the superoperator $\mathcal{G}$ plotted in the form of the vector $\vec{G}$. Bars correspond to the ML estimation (black), linear inversion (gray), and exact values (hollow). Missing hollow bars indicate the zero true values. The superoperator describes the process of damping, $\Gamma_{\|} = 0.5$ and $\Gamma_{\perp} = 0.75$, $\mathcal{N} = 20$.

quency of projections to the state $|k\rangle$ measured for the input state $|j\rangle$. The likelihood function can be expressed as product of 36 terms,

$$\mathcal{L}[\mathcal{G}]=\prod_{j,k} (\langle k|\mathcal{G}[|j\rangle\langle j|]|k\rangle)^{f_{jk}}, \tag{23}$$

where $j,k\in\{\uparrow_x,\downarrow_x,\uparrow_y,\downarrow_y,\uparrow_z,\downarrow_z\}$.

In our simulations, the black box of Fig. 1 corresponds to the damping of $\varrho_{\rm in}$,

$$\varrho_{\rm out}=\begin{pmatrix} 1-\varrho_{\rm in,11}e^{-\Gamma_{\|}} & \varrho_{\rm in,01}e^{-\Gamma_{\perp}} \\ \varrho_{\rm in,10}e^{-\Gamma_{\perp}} & \varrho_{\rm in,11}e^{-\Gamma_{\|}} \end{pmatrix}. \tag{24}$$

Here $2\Gamma_{\perp}\geqslant\Gamma_{\|}\geqslant 0$ are transversal and longitudinal decay parameters. The elements of the reconstructed superoperator are depicted in Fig. 2. The solution was obtained by iterations of Eqs. (15) and (16). For the total amount of 360 measurements the ML estimate (black) is very close to the exact values $\mathcal{G}$ (hollow). Notice that the ML estimate always provides a physically sound result contrary to the linear inversion (gray).

Properties of transforming systems are of interest in any physical theory. The developed formalism shows how to identify a generic quantum-mechanical process. Quantum systems consisting of spins, two entangled or three entangled (GHZ) qubits are tractable due to their low dimensionality. However, a proper and full quantum description of possible transformations of such systems is more advanced, since they are characterized by 12, 240, or even 4032 parameters.

This work was supported by Grant No. LN00A015 of the Czech Ministry of Education. This paper is dedicated to the 65th birthday of Professor Jan Peřina.

[1] D.T. Smithey, M. Beck, M.G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993); S. Schiller, G. Breitenbach, S.F. Pereira, T. Müller, and J. Mlynek, *ibid.* **77**, 2933 (1996).
[2] M. Vasilyev, S.-K. Choi, P. Kumar, and G.M. D'Ariano, Phys. Rev. Lett. **84**, 2354 (2000).
[3] S. Wallentowitz and W. Vogel, Phys. Rev. A **53**, 4528 (1996); K. Banaszek and K. Wódkiewicz, Phys. Rev. Lett. **76**, 4344 (1996).
[4] L.G. Lutterbach and L. Davidovich, Phys. Rev. Lett. **78**, 2547 (1997).
[5] C.T. Bodendorf, G. Antesberger, M.S. Kim, and H. Walther, Phys. Rev. A **57**, 1371 (1998).
[6] S. Wallentowitz and W. Vogel, Phys. Rev. Lett. **75**, 2932 (1995); D. Leibfried, D.M. Meekhof, B.E. King, C. Monroe, W.M. Itano, and D.J. Wineland, *ibid.* **77**, 4281 (1996).
[7] T.J. Dunn, I.A. Walmsley, and S. Mukamel, Phys. Rev. Lett. **74**, 884 (1995); C. Leichtle, W.P. Schleich, I.Sh. Averbukh, and M. Shapiro, *ibid.* **80**, 1418 (1998).
[8] R.G. Newton and B. Young, Ann. Phys. (N.Y.) **49**, 393 (1968).
[9] J.F. Poyatos, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).
[10] I.L. Chuang and M.A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).
[11] G.M. D'Ariano and L. Maccone, Phys. Rev. Lett. **80**, 5465 (1998); Fortschr. Phys. **46**, 837 (1998).
[12] V. Bužek, Phys. Rev. A **58**, 1723 (1998).
[13] R. Gutzeit, S. Wallentowitz, and W. Vogel, Phys. Rev. A **61**, 062105 (2000).
[14] A. Luis and L.L. Sánchez-Soto, Phys. Lett. A **261**, 12 (1999).
[15] Z. Hradil, R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch, Phys. Rev. Lett. **76**, 4295 (1996).
[16] H. Mabuchi, Quantum Semiclassic. Opt. **8**, 1103 (1996).
[17] G.M. D'Ariano, M.G.A. Paris, and M.F. Sacchi, Phys. Rev. A **62**, 023815 (2000).
[18] Z. Hradil, Phys. Rev. A **55**, R1561 (1997); Z. Hradil, J. Summhammer, and H. Rauch, Phys. Lett. A **261**, 20 (1999).
[19] K. Banaszek, Phys. Rev. A **57**, 5013 (1998).
[20] K. Banaszek, G.M. D'Ariano, M.G.A. Paris, and M.F. Sacchi, Phys. Rev. A **61**, 010304(R) (1999).
[21] Z. Hradil, J. Summhammer, G. Badurek, and H. Rauch, Phys. Rev. A **62**, 014101 (2000).