PACS NUMBERS:03.65.-w

KEY WORDS: quantum estimation, incompatible observations, quantum state reconstruction, maximum likelihood

Quantum tomography as normalization of incompatible observations

Z. Hradil,[*] J. Summhammer, H. Rauch
Atominstitut der Österreichischen Universitäten,
Stadionallee 2,
A–1020 Wien, Austria
(Dated: December 15, 2002)

Synthesis of incompatible observations – a reconstruction of quantum state – is formulated consistently with quantum theory using maximum likelihood estimation. Solving the nonlinear multidimensional equation for the density matrix is equivalent to finding the probability operator measure, whose expectation values are measured frequencies, and which provides the decomposition of an identity operator.

Quantum theory handles observable events at the most fundamental level currently available and predicts the statistics of quantum phenomena. The randomness is hidden in the quantum state. Observing various faces of the same system enables us to reconstruct its quantum state. Due to the similarities with X-ray tomography, the state reconstruction is sometimes called quantum tomography. Although this problem may be traced back to the early days of quantum mechanics [1], it is quantum optics that opened a new era [2–4]. Nowadays, similar techniques are commonly used in various branches of contemporary physics [5]. An overview can be found in Ref. [6, 7]. Though these techniques may give a rough picture of the state, they do not provide its full quantum description. Since the available measurement is always limited as far as the amount and accuracy of data is concerned, problems with positivity of reconstructed density matrix arise. A proper statistical approach provides a remedy for these problems and several approaches have recently been formulated [8, 9]. A novel formulation relating the Maximum Likelihood (MaxLik) estimation to generalized measurements in quantum theory will be addressed in this paper.

The MaxLik estimation itself represents a standard tool of statistical analysis [10]. It has already been used even in quantum theory for quantum phase [11, 12] and quantum state estimations [13, 14]. However, in all these applications the MaxLik has been used in technical sense for optimization of several parameters restricted by some additional constraints. Without this simplification the analysis is considered as intractable due to the nonlinear multidimensional optimization.

In the following, the states $|y_i\rangle$ will denote general nonorthogonal states. Let us assume that a quantum measurement has been done n times yielding the relative frequencies of the events represented by the states $|y_i\rangle$ as $f_i > 0, \sum_i f_i = 1$. The states are assumed to be nonorthogonal. A measurement is sharp provided that it may be described by a projection into a pure normalized states $|y_i\rangle\langle y_i|$. On the contrary, an unsharp measurement is represented by a probability operator measure (POM) [15] representing a superposition in bins D_i of indistinguishable states $\hat{\Pi}_i = \sum_{j\in D_i} |y_j\rangle\langle y_j|$. The measurement is complete if all the POMs corresponding to the counted data yield the decomposition of identity operator. In the case that not all such values were detected, the measurement is incomplete. Quantum theory predicts the probability of the outcomes to be $\rho_{ii} = \langle y_i|\hat{\rho}|y_i\rangle$. Standard approaches straightforwards associate the probabilities with detected frequencies as

$$\rho_{ii} = f_i. \tag{1}$$

However, this may appear inconsistent with quantum theory. Problems of this kind are very often neglected in the existing literature. A typical result of standard state reconstruction is a "density matrix" with coefficients fluctuating within certain errors. Unfortunately, such an object does not represent any quantum state due to the necessary condition of semipositive definiteness. Consequences are crucial as may be illustrated by the simple but physically valuable example of 1/2 spin measurement. Consider the "quantum game" proposed by Massar and Popescu [16]. Here N realizations of the system prepared in an unknown pure quantum state

$|\varphi_{true}\rangle$ are measured by some device yielding data denoted for brevity by a parameter k. The unknown quantum state is estimated by an estimator $|\varphi_{est}(k)\rangle$ depending on the measured data. The quality of the estimate may be evaluated by the scalar product, the so–called fidelity $F = |\langle\varphi_{true}|\varphi_{est}\rangle|^2$. The total score of the proposed quantum game is given by the averaging the fidelity over all possible data k and all the possible initial quantum states $|\varphi_{true}\rangle$

$$S = \left\{ |\langle\varphi_{true}|\varphi_{est}(k)\rangle|^2 \right\}_{k,\varphi_{true}} \leq 1. \qquad (2)$$

The higher score, the better quantum state prediction. As the fundamental consequence of quantum theory, the score is limited by the upper limit $S = (N+1)/(N+2)$[16]. Let us now derive the score corresponding to the standard reconstruction. The standard measurement of the spin may be regarded as subsequent Stern–Gerlach projections of an unknown spin state into the three fixed orthogonal directions $\mathbf{x_i}, i = 1, 2, 3$ on the Poincaré sphere. A general projection represented by an unity vector $\mathbf{x}$ may be parametrized by means of matrix

$$|\mathbf{x}\rangle\langle\ \mathbf{x}| = \frac{1}{2}\big(1 + (\mathbf{x}\sigma)\big), \qquad (3)$$

where $(\mathbf{x}\sigma)$ denotes the scalar product of a unity vector and Pauli sigma matrices. Each measurement may be done separately with N_i particles, $N_1 + N_2 + N_3 = N$ yielding results k_i, respectively. Obviously, the relative frequencies k_i/N_i are estimating the orientation of spin directions

$$\frac{1}{2}(1 + \xi_i) = k_i/N_i. \qquad (4)$$

This is an analogy to the general deterministic relation (1). The estimates (ξ_1, ξ_2, ξ_3) of the unknown spin state fluctuates around its true value $\mathbf{n}$. Provided that $|\xi|^2 \leq 1$, the estimator is the well defined (mixed) state $\hat{\rho}_{est} = 1/2(1 + (\xi\sigma))$. However, for $|\xi|^2 > 1$ the vectors are outside the Poincaré sphere and do not correspond to any quantum state. The measured data k_i fluctuate independently according to the binomial distributions

$$P_i(k_i) = \binom{N_i}{k_i} p_i^{k_i}(1 - p_i)^{N_i - k_i},$$

where $\mathbf{p} \equiv (p_1, p_2, p_3) = (1 + \mathbf{n})/2$. Averaging the fidelity

$$F(\xi, \mathbf{n}) = \frac{1}{2}(1 + \xi_1 n_1 + \xi_2 n_2 + \xi_3 n_3) \qquad (5)$$

over the measured data $k = (k_1, k_2, k_3)$ and over the unknown spin state $\mathbf{n}$ regardless of the semipositive definiteness can be done. The exact result

$$S = \frac{1}{4\pi}\int d^2\Omega_n \sum_{k_1,k_2,k_3} P_1(k_1)P_2(k_2)P_3(k_3)F(k, \mathbf{n}) = \frac{1}{4\pi}\int d^2\Omega_n \frac{1}{2}(1 + |\mathbf{n}|^2) = 1 \qquad (6)$$

is independent of the total number of particles N. Does this example provide a better state prediction than the ideal quantum case? No. The standard approach neglects the quantum noises because the necessary premise of quantum theory, semipositiveness of the density matrix, is not fulfilled. The subtle distinction between the standard and statistical methods is then just at the level of quantum noise $O(1/N)$. This is a characteristic feature for any unbiased standard reconstruction of a pure state. If the estimates are fluctuating around the true projectors unbiasely, then

$$\left\{ |\varphi_{est}(k)\rangle\langle\varphi_{est}(k)| \right\}_k \to |\varphi_{true}\rangle\langle\varphi_{true}| \qquad (7)$$

and the score of noisy measurement may reach 1. The relation (7) is contradictory to quantum theory since superposition of projectors cannot be a projector again! Due to the improper manipulation with quantum noises, the standard approaches describe the state reconstruction in a classical way. A proper treatment should exhibit enhanced noise. These conclusions are obvious in the above example of the spin system, however, they may seem to be counterintuitive.

Let us formulate more adequate approach. Separate projections of the spin (4) for $i = 1, 2, 3$ are well established Stern–Gerlach measurements. These observations are incompatible, because the corresponding projectors do not commute among themselves. Each of the measurements estimates components of the spin ξ_1, ξ_2 and ξ_3 separately. However, since the data are fluctuating, each prediction is done with different error $\Delta\xi_i$. Proper method for state estimation should take these errors into account giving the higher credit to the more accurate observations. Unfortunately, errors depend on the deviations between the true spin state and projections. This closes the loop of the estimation procedure: Incompatible measurements of an unknown state provide certain data. To interpret correctly the observed data, one should know the state, which was observed. The sophisticated scheme for state estimation will be therefore nonlinear and results of incompatible observations must be normalized among themselves.

Let us formulate righorous treatment. Proper statistical approach must release the relation (1). For this purpose, the so–called likelihood functional

$$\mathcal{L}(\hat{\rho}) = \prod_i \langle y_i|\hat{\rho}|y_i\rangle^{nf_i} \qquad (8)$$

may be constructed as an unnormalized multidimensional probability. It is given by product of probability densities predicted by quantum theory $\langle y_i|\hat{\rho}|y_i\rangle$ corresponding to all detected data. In principle, many quantum states might yield the detected data. This is the viewpoint of Bayesian approaches which normalize the likelihood functional and consider it to be a posterior probability density conditioned by the detected data. Unfortunately, this is intractable due to the enormous problems with normalization over all density matrices. Considering only those states, which may yield the given data most likely, seems to be a reasonable compromise between the physical content and the mathematical complexity. Quantum state attributed to the data will be searched in the form of the density matrix $\hat{\rho}$ which maximizes the likelihood functional (8) (MaxLik estimation). As shown in [17], this problem make sense, because the likelihood functional is always limited by its upper bound following from the Gibbs inequality. Let us assume the diagonal representation of a density matrix in an orthogonal basis $|\phi_k\rangle$

$$\hat{\rho} = \sum_k r_k |\phi_k\rangle\langle\phi_k|. \tag{9}$$

The existence of parameters $r_k \geq 0, \sum_k r_k = 1$ is guaranteed by quantum theory. Normalized extreme states satisfy the relation

$$\frac{\partial}{\partial\langle\phi_k|}\left[\frac{1}{n}\ln\mathcal{L} - \Lambda\mathrm{Tr}(\rho)\right] = 0,$$

Λ being a Lagrange multiplier. This multidimensional nonlinear optimization gives a system of coupled equations

$$\hat{R}|\phi_k\rangle = |\phi_k\rangle, \tag{10}$$

$$\hat{R} = \sum_i \frac{f_i}{\rho_{ii}}|y_i\rangle\langle y_i|, \quad \rho_{ii} = \sum_k r_k |\langle\phi_k|y_i\rangle|^2.$$

The normalization condition $\mathrm{Tr}\hat{\rho} = 1$ implies $\Lambda = 1$. The relation (10) enables a correct statistical synthesis of observations in the form of the nonlinear equation for density matrix

$$\hat{R}(\hat{\rho})\hat{\rho} = \hat{\rho}. \tag{11}$$

The reconstruction is done in the subspace, where the operator $\hat{R}$ represents an identity operator. The mathematical problem of optimization is tightly related to the structure of quantum theory. The states in Hilbert space are given except for the multiplicative factors. The renormalized POM

$$|y_i\rangle\langle y_i| \rightarrow |y_i'\rangle\langle y_i'| = \frac{f_i}{\rho_{ii}}|y_i\rangle\langle y_i|$$

characterizes the synthesis of measurements analogously to the relation (1)

$$\langle y_i'|\hat{\rho}|y_i'\rangle \equiv f_i. \tag{12}$$

The MaxLik equation (10) may then be interpreted as the completeness relation in the appropriate subspace

$$\hat{R} \equiv \sum_i |y_i'\rangle\langle y_i'| = \hat{1}_\rho. \tag{13}$$

Because only the decomposition of identity matters, the formulation is common for both the sharp and unsharp observations. This decomposition of identity may be compared with formally similar one that is spanned by the linear combinations of all observed rays [18]

$$\sum_i \hat{G}^{-1/2}|y_i\rangle\langle y_i|\hat{G}^{-1/2} = \hat{1},$$

where $\hat{G} = \sum_i |y_i\rangle\langle y_i|$. As the significant difference, this identity encompasses a larger subspace than the decomposition (13). In general, the given observation is insufficient for a successful reconstruction here.

The problem of state reconstruction in quantum theory may be in alternative way reformulated as a problem of statistical distance. Counted frequencies f_i and ideal probabilities ρ_{ii} predicted by quantum theory represent points in multidimensional domain. To reconstruct a state means to find a probability distribution as close as possible to the given data. Natural measure on probability space is given by the Kullback–Leibler divergence [10]

$$K(\rho/f) = -\sum_i f_i \ln\frac{\rho_{ii}}{f_i} \geq 0. \tag{14}$$

This is nothing else than a normalized log likelihood function and its minimization reproduces the above mentioned results. One may consider other distances, such as the Euclidean distance [7] $K_E(\rho/f) = \sum_i(\rho_{ii} - f_i)^2$ or the Riemannian one [19] $K_R(\rho/f) = \sum_i(\rho_{ii} - f_i)^2/\rho_{ii}$. In these cases, however, the intuitive relations (12) and (13) characterizing the synthesis of incompatible observations as generalized measurement cannot be obtained any more. The formal relations linking the MaxLik optimization to quantum theory represent the key results of this paper. Solving the nonlinear multidimensional equation (11) for the density matrix is equivalent to finding the probability operator measure, whose expectation values are detected frequencies (12), and which provides the decomposition of an identity operator (13). This provides in certain sense a complementary formulation to the treatment in quantum estimation theory [15]. Usually, the optimal measurement, which minimizes the average cost, should be found for a given quantum state. The equation for optimum strategy is then linear. Our motivation was different: For a given measurement an optimum state matching the detected data, should be found. The resulting equation is nonlinear. The mathematical complexity of the latter formulation is compensated by the fact, that realistic measurements may be analyzed according this scheme. Particularly, a synthesis of any incompatible and incomplete observations is always complete somewhere.

The key role is played by the operator $\hat{R}$ concentrating our knowledge about the results of the measurement. To find its geometrical interpretation, let us consider the following simplified optimization associated with overlapping of the rays. Assume an orthogonal basis $|x_k\rangle$ spanning an orthogonal subspace $\mathcal{O}$. States may be decomposed as $|y_i'\rangle = \sum_k \langle x_k|y_i'\rangle|x_k\rangle + |Z_i\rangle$, where $|Z_i\rangle$ is orthogonal to the subspace $\mathcal{O}$, $\langle x_k|Z_i\rangle = 0$. The sum of scalar products $Z = \sum_i \langle Z_i|Z_i\rangle$ provides information about the part of the states outside the orthogonal subspace $\mathcal{O}$. Now the basis $|x_k\rangle$ will be chosen in order to minimize Z under the normalization condition. The optimum basis in $\mathcal{O}$ is spanned by the eigenstates which diagonalize the operator $\hat{R}$

$$\left[\sum_i |y_i'\rangle\langle y_i'|\right]|x_k\rangle = \lambda_k|x_k\rangle. \tag{15}$$

Hence the operator $\hat{R}$ characterizes the overlapping of nonorthogonal states and similarly does the operator $\hat{G}$. Of course, this is not fully equivalent to the above mentioned MaxLik estimation, where besides the optimum subspace the quantum state itself is searched. The MaxLik reconstruction may be accomplished in such subspace where the operator $\hat{R}$ become an identity operator. This is mutually related to finding an optimum state and optimum probability operator measure, which decomposes the subspace of reconstruction. This seems to be satisfactory from the intuitive viewpoint. Reconstruction techniques are frequently compared with the X-ray tomography in medicine, where only a certain field of view scanned by all the rays may be reconstructed. The same is true in quantum domain. The field of view corresponds to the subspace where operator $\hat{R}$ represents an identity operator. This part of Hilbert space is common to all observations.

A solution of the nonlinear eq. (11) represents the key point of the procedure. Formally it may be found in an iterative manner provided that the necessary conditions for convergence are fulfilled. Since for the exact solution the condition $\hat{R} = \hat{1}$ is true, there is an infinite number of formulations equivalent to the equation (11). As the main numerical difficulty, the different forms may appear as nonequivalent for iterations. This is why the particular cases of realistic measurements must be treated separately. This general theory provides a paradigm for the demanding program of quantum reinterpretation of nowadays existing standard reconstruction techniques. Several promising application are available. Theory considerably simplifies in the case of commuting observations. A numerical algorithm for MaxLik reconstruction of the diagonal elements of a density matrix using the homodyne detection with random phase [20] has been proposed by Banaszek [21]. Experimental data has been analysed in [22]. In general, the solution depends on the starting point of iterations. Consequently, the MaxLik estimation provides a family of extremum states not distinguished by the given measurement. This describes uncertainty of reconstruction in the space of allowed states. The MaxLik analysis of phase measurements in interferometry has been addressed in [23]. The case of spin estimation will be published in the forthcoming publication. The MaxLik estimation of the full homodyne detection is under consideration.

Methods usually used for a deterministic quantum state reconstruction do not correctly describe the quantum noises. The MaxLik optimization provides a formulation, which can be interpreted in the language of quantum theory. Solving the nonlinear multidimensional equation for the density matrix is equivalent to finding the probability operator measure, whose expectation values are measured frequencies, and which provides the decomposition of an identity operator. The MaxLik approach focuses on the most likely interpretations of reality and its consequences are relevant to other statistical approaches as well.

This work was supported by TMR Network ERB FMRXCT 96-0057 “Perfect Crystal Neutron Optics” of the European Union and by the grant of Czech Ministry of Education VS 96028.

---

[*] Permanent address: Department of Optics, Palacký University, 17. listopadu 50, 772 07 Olomouc, Czech Republic, email:hradil@risc.upol.cz

[1] S. Weigert, Phys. Rev. **A 45**, 7688 (1992); *ibid.***A 53**, 2078 (1996).

[2] K. Vogel, H. Risken, Phys. Rev. **A 40**, 2847 (1989).

[3] D. T. Smithey, M. Beck, M. G. Raymer, A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).

[4] S. Schiller, G. Breitenbach, S. F. Pereira, T. Müller, J. Mlynek, Phys. Rev. Lett. **77**, 2933 (1996).

[5] S. Wallentowitz, W. Vogel, Phys. Rev. Lett. **75**, 2932 (1996); P. J. Bardoff, C. Leichtle, G. Schrade, W. P. Schleich, Phys. Rev. Lett. **77**, 2198 (1996); D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, D.J. Wineland, Phys. Rev. Lett. **77**, 4281 (1996).

[6] U. Leonhardt, *Measuring of the Quantum State of Light*, Cambridge Press, 1997.

[7] D.– G. Welsch, W. Vogel, T. Opatrný, “Homodyne Detection and Quantum State Reconstruction”, to appear in *Progress in Optics, 39* , ed. E. Wolf.

[8] K. R. W. Jones, Phys. Rev. **A 50**, 3682 (1994).

[9] V. Bužek, G. Adam, G. Drobný, Annals of Phys. (NY) **245**, 37 (1996).

[10] R. C. Rao, *Linear Statistical Inference and Its Applications* (New York, Wiley, 1973).

[11] J. H. Shapiro, S. R. Shepard, N. C. Wong, Phys. Rev. Lett. **62**, 2377 (1989).

[12] A. S. Lane, S. L. Braunstein, and C. M. Caves, Phys. Rev. **A 47**, 1667 (1993).

[13] D. Mogilevtsev, Z. Hradil, J. Peřina, Rev. Mod. Opt. **44**, 2261 (1997); D. Mogilevtsev, Z. Hradil, J. Peřina, Quantum Semiclass. Opt. **10**, 345 (1998).

[14] S. M. Tan, Rev. Mod. Opt. **44**, 2233 (1997).
[15] C. W. Helstrom, *Quantum Detection and Estimation Theory,* Academic Press, New York 1976.
[16] S. Massar, S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
[17] Z. Hradil, Phys. Rev. **A 55**, R1561 (1997).
[18] C. M. Caves, private communication.
[19] W. K. Wooters, Phys. Rev. **D 23** 357 (1981).
[20] M. Munroe, D. Boggavarapu, M. E. Anderson, M. G. Raymer, Phys. Rev. **A 52**, R924 (1995).
[21] K. Banaszek, Phys. Rev. **A 57**, 5013 (1998); K. Banaszek, acta phys. slov. **48**, 185 (1998).
[22] Z. Hradil, R. Myška, acta phys. slov. **48**, 199 (1998).
[23] J. Řeháček, Z. Hradil, M. Zawisky, S. Pascazio, H. Rauch, J. Peřina, to appear in Phys. Rev. A, July 1999.