cta physica slovaca vol. 46 No. 3, 355 – 366 June 1996

QUANTUM INFERENCE FOR STATISTICAL MIXTURES

R. Derka, V. Bužek
Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovakia

G. Adam
Institut für Theoretische Physik, Technische Universität Wien Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

Received 31 May 1996, accepted 7 June 1996

We generalize the Bayesian scheme of quantum inference for a reconstruction of impure states of quantum systems. We have solved the problem of ambiguity in a definition of the invariant integration measure in a space of impure quantum-mechanical states, which has been the main obstacle in application of Bayesian methods for statistical mixtures [see K.R.W. Jones, *Ann. Phys. (N.Y.)* **207**, 140 (1991)]. As an illustration, we analyze in detail how the standard Bayesian inference can be applied for a reconstruction of a pure state of a spin-1/2 system. We also show how this scheme fails when the spin is prepared in an impure state. We apply our generalized Bayesian inference scheme for a consistent reconstruction of an impure state of the spin-1/2. In addition we show that in the limit of infinite number of measurements this reconstruction scheme gives the same result as the Jaynes principle of the maximum entropy.

1. Introduction

The concept of a state of a physical system represents one of the most important pillars of the paradigm of the quantum theory [1]. From the mathematical point of view a *pure* quantum-mechanical state is represented by a point in an abstract state space. Physical interpretation of a state is more trickier [1]: the state is understood as an ensemble of identically prepared quantum-mechanical systems. One of the most important problems in quantum mechanics is how to specify (i.e., how to reconstruct) the density operator describing an ensemble of systems obtained via the particular preparation procedure. If no *a priori* information about the state is available then a complete reconstruction of the density operator of the quantum-mechanical system can be performed providing *all* mean values of a complete set of physical observables (i.e. the *quorum* of observables [2] which corresponds to a reconstruction on the complete observation level [3]) associated with the given system are measured. When a complete set of measurements over the system is performed then a complete reconstruction of the state of this system can be realized. The particular reconstruction procedure on the complete observation level depends on the way in which the data are collected. Two well known examples are **(1)** the optical homodyne tomography [4] (here the data are obtained via the measurement of probability distributions of rotated quadratures [5]) **(2)** "filtering" of quantum states with quantum filters [6] (here the data are obtained in a process of a simultaneous measurement of conjugated observables [7]). Complete reconstruction schemes which are based on measurements of either mean values of physical observables or their probability distributions can be methodologically unified by the Jaynes principle of maximum entropy (*MaxEnt* principle) [8]. With the help of this principle a complete reconstruction of the quantum-mechanical state can be straightforwardly performed. The *MaxEnt* principle can be fruitfully applied also in the case when only a subset $\hat{G}_\nu$ $(\nu = 1, ..., n)$ of physical observables (the so called reduced observation level) is measured. In this case the density operator of the physical system can be reconstructed on the reduced observation level. This so called generalized canonical density operator fulfils several conditions. Firstly, its trace is equal to unity ($\mathrm{Tr}\hat{\rho} = 1$). Secondly, $\mathrm{Tr}\{\hat{\rho}\hat{G}_\nu\} = G_\nu$ $(\nu = 1, ..., n)$, which means that the reconstructed density operator provides us with the measured mean values of those observables which constitute the given observation level. Obviously, a large number of density operators can fulfill these two constraints. So one needs an additional criterion which would uniquely specify the generalized canonical density operator. According to Jaynes [8] this operator has to be the one with the largest value of the von Neumann entropy $S = -\mathrm{Tr}\{\hat{\rho}\ln\hat{\rho}\}$. This additional condition means that the *MaxEnt* principle is the most conservative assignment in the sense that it does not permit one to draw any conclusion not warranted by the experimental data.

The advantage of the *MaxEnt* principle is that no *a priori* information about the reconstructed state is needed. As soon as the mean values of a given set of physical observables are available (i.e., measured) then the generalized canonical density operator can be (in principle) found. Here we note, that the *exact* knowledge of any mean value of a physical observable implicitly assumes an *infinite* number of repeated measurements. In practice an observer can perform only a limited number of different measurements (which specifies a given observation level) on a limited number of elements of the given ensemble (i.e., only a finite number of measured events can be registered in a finite time). In this case mean values of the measured observables are not known exactly and consequently the Jaynes principle of the maximum entropy cannot be used. What is known from the measurement is a specific set of data indicating number of how many times eigenvalues of given observables have appeared (which in the limit of infinite number of measurements results in the corresponding quantum probability distributions). The question is, how to obtain the best *a posteriori* estimation of the density operator based on the measured data. Helstrom [9], Holevo [10] and Jones [11] have shown that the answer to this question can be given by the Bayesian inference method providing it is *a priori* known that the quantum-mechanical state which is going to be reconstructed is prepared in a pure (even though unknown) state. Once this condition is fulfilled, then the observer can systematically predict (i.e. reconstruct) an *a posteriori* density operator based upon an incomplete set of experimental data. This density operator is equal to the mean over all possible *pure* states weighted by a specific probability distribution in an abstract state space with the unique invariant integration measure. It is this probability distribution (conditioned by the assumed Bayesian prior) which characterizes observer's knowledge of the system at every moment during the measurement sequence. We note once again that the Bayesian inference has been developed for a reconstruction of *pure* quantum mechanical states and in this sense it corresponds to an averaging over a *microcanonical* ensembles. To illustrate this scheme we can imagine a spin-1/2. The parametric space of pure states of the spin-1/2 is represented by the Poincare sphere and the Bayesian reconstruction scheme corresponds to a specific averaging over points (states) *on* the sphere.

In a real situation one can never design a preparator such that it produces an ensemble of identical pure states. What usually happens is that the ensemble consists of a set of pure states each of which is represented in the ensemble with a certain probability. So now the question is how to apply Bayesian reconstruction scheme providing the quantum-mechanical system under consideration is in an impure state (i.e. statistical mixture). To apply the Bayesian inference scheme one has to define exactly three objects: **(1)** the abstract state space of the measured system; **(2)** the corresponding invariant integration measure of this space; and **(3)** the *prior* (i.e. *a priori* known probability distribution on the given parametric state space). It is relatively easy to specify the parametric state space. For instance, in the case of the spin-1/2 which is prepared in a statistical mixture of pure state this parametric space can be identified with all points inside the Poincare sphere. On the other hand, no unique prescription how to specify the invariant measure and the prior for impure states can be found in the literature [11,12].

The main purpose of the present paper is to show how to generalize the Bayesian quantum inference for a reconstruction of impure states. The main idea of our approach is based on an observation that a quantum-mechanical system which is prepared in a statistical mixture can be represented as a subsystem of a composite system which itself is in a pure state (for a simplicity we will denote the composite system as S, the subsystem of interest as P and the additional degrees of freedom will be called as a "reservoir" R). This means that the Bayesian inference scheme can be safely applied to a reconstruction of the composite systems. Finally, by tracing over the "irrelevant" degrees of freedom (i.e., over the reservoir R) one can obtain the reconstructed density operator of the subsystem of interest. The only problem in this generalized Bayesian reconstruction scheme is how to specify uniquely the composite system of which our system of interest is a subsystem. Here we apply the idea of the Schmidt decomposition [13] which says us that it is suffiecient that the dimension of a state space of the reservoir R is the same as the dimension of the state space of the subsystem P. We will illustrate our ideas on an example of a reconstruction of an impure state of a spin-1/2. We will show how the *a priori* assumption about the purity/impurity of the reconstructed scheme can change the *a posteriori* estimation of the density operator. We will also show that for specific sets of data the Bayesian reconstruction scheme based on an *a* *priori* assumption that the reconstructed system is in a pure state can completely fail.

In addition we show that our Bayesian scheme of quantum inference developed for a reconstruction of statistical mixtures corresponds to averaging over grand canonical ensemble. Moreover, in the limit of infinite number of measurements the reconstructed density operator is *equal* to the generalized canonical density operator obtained via the Jaynes principle of the maximum entropy.

The paper is organized as follows, in Section 2 we briefly review the Bayesian reconstruction scheme for pure states. In Section 3 we analyze quantum inference in the limit of infinite number of measurements. A simple example of a reconstruction of pure states of the spin-1/2 is presented in Section 4. General principles of the Bayesian reconstruction of impure states are discussed in Section 5. In Section 6 we analyze reconstruction of impure states of the spin-1/2.

2. Bayesian reconstruction scheme

The general idea of the Bayesian reconstruction scheme is based on manipulations with probability distributions in parametric state spaces. To understand this reconstruction scheme we introduce several definitions and concepts. Firstly, it is a space of states of the measured system. Quantum Bayesian method as discussed in the literature [9-11] is based on the assumption that the reconstructed system is in a pure state described by a state vector $|\Psi\rangle$ or equivalently by a pure-state density operator $\hat{\rho} = |\Psi\rangle\langle\Psi|$. The manifold of all pure states is a continuum which we denote as Ω. Secondly, it is the discrete space A of reading states of a measuring apparatus associated with the observable $\hat{O}$. These states are intrinsically related to the projectors $\hat{P}_{\lambda_i,\hat{O}}$, where λ_i are the eigenvalues of the observable $\hat{O}$.

The Bayesian reconstruction scheme is formulated as a three-step inversion process:

(1) As a result of the measurement a conditional probability

$$p(\hat{O},\lambda_i|\hat{\rho}) = \mathrm{Tr}\left(\hat{P}_{\lambda_i,\hat{O}}\hat{\rho}\right), \tag{2.1}$$

on the discrete space A is defined. This conditional probability distribution specifies a probability of finding the result λ_i if the measured system is in a particular state $\hat{\rho}$.

(2) To perform the second step of the inversion procedure one has to specify an *a priori* distribution $p_0(\hat{\rho})$ defined on the space Ω. This distribution describes our initial knowledge about the measured system. Using the conditional probability distribution $p(\hat{O},\lambda_i|\hat{\rho})$ and the *a priori* distribution $p_0(\hat{\rho})$ we can define the *joint* probability distribution $p(\hat{O},\lambda_i;\hat{\rho})$

$$p(\hat{O},\lambda_i;\hat{\rho}) = p(\hat{O},\lambda_i|\hat{\rho})p_0(\hat{\rho}), \tag{2.2}$$

on the space $\Omega\otimes A$. We note that if no initial information about the measured system is known then the prior $p_0(\hat{\rho})$ has to be assumed to be constant.

(3) The final step of the Bayesian reconstruction is based on the well known Bayes rule $p(x|y)p(y) = p(x;y) = p(y|x)p(x)$ with the help of which we find the conditional probability $p(\hat{\rho}|\hat{O},\lambda_i)$ on the state space Ω:

$$p(\hat{\rho}|\hat{O},\lambda_i) = \frac{p(\hat{O},\lambda_i,\hat{\rho})}{\int\limits_{\Omega} p(\hat{O},\lambda_i,\hat{\rho})d_\Omega}, \tag{2.3}$$

from which the reconstructed density operator can be obtained [see Eq.(2.4)].

In the case of the repeated N-trial measurement the reconstruction scheme consists of an iterative utilization of the three-step procedure as described above. After the N-th measurement we use as an input for the prior distribution the conditional probability distribution which is an output after the $(N-1)$st measurement. However, we can equivalently define the N-trial measurement conditional probability $p(\{\ \}_N|\hat{\rho}) = \prod_{i=1}^{N} p(\hat{O}_i,\lambda_j|\hat{\rho})$ and applying the three-step procedure (2.1-2.3) just once we get the reconstructed density operator

$$\hat{\rho}(\{\ \}_N) = \frac{\int\limits_{\Omega} p(\{\ \}_N|\hat{\rho})\hat{\rho}d_\Omega}{\int\limits_{\Omega} p(\{\ \}_N|\hat{\rho})d_\Omega}, \tag{2.4}$$

where $\hat{\rho}$ in the r.h.s. of Eq.(2.4) is a properly parameterized density operator in the state space Ω. Until now we have not mentioned one essential problem in the Bayesian reconstruction scheme, which is the determination of the integration measure d_Ω.[1] The integration measure has to be invariant under unitary transformations in the space Ω. This requirement uniquely determines the form of the measure. However, this is no longer valid when we extend Ω to the space of mixed states formed by all convex combinations of elements of the original pure state space Ω. Although the Bayesian procedure itself doesn't require any special conditions imposed on the space Ω, the ambiguity in determination of the integration measure prevents us to generalize this method straightforwardly for the case of impure quantum states.

3. Bayesian inference in the limit of infinite number of measurements

The explicit evaluation of the *a posteriori* estimation of the density operator $\hat{\rho}\{\ \}_N$ is significantly limited by technical difficulties when integration over parametric space is performed [see Eq.(2.4)]. Even for simplest quantum systems and a relatively small number of measurements the reconstruction procedure can be practically unrealizable problem.

On the other hand let us assume that the number of measurements approaches infinity (i.e. $N \to \infty$). It is clear that in this case mean values of all projectors $\langle\hat{P}_{\lambda_j,\hat{O}_i}\rangle$ associated with the observables $\hat{O}_i$ are *precisely* know (measured); i.e.

$$\langle\hat{P}_{\lambda_j,\hat{O}_i}\rangle = \alpha^i_j, \tag{3.1}$$

[1] Many authors (see, for instance, Ref.[11]) identify the prior distribution with the integration measure on the space Ω. However, the particular form of d_Ω is associated with the topology and the particular parameterization of the space Ω rather then with some prior information $p_0(\hat{\rho})$ about this system. We will distinguish between these two objects.

which can be considered as a probability distribution such that $\sum_j \alpha^i_j = 1$. In [14] we have shown[2] that in the limit $N \to \infty$ Eq.(2.4) reads

$$\hat{\rho}(\{\ \}_{N\to\infty}) = \frac{1}{N_0}\int_\Omega \prod_i \Big[\prod_{j=1}^{n_i} \delta\left(\mathrm{Tr}\left(\hat{P}_{\lambda_j,\hat{O}_i}\hat{\rho}\right) - \alpha^i_j\right)\Big]\hat{\rho}\, d_\Omega, \tag{3.2}$$

where N_0 is a normalization constant determined by the condition $\mathrm{Tr}\,[\hat{\rho}(\{\ \}_{N\to\infty})] = 1$.

The interpretation of Eq.(3.2) is straightforward. The reconstructed density operator is equal to the sum of equally-weighted pure-state density operators on the manifold Ω, which do satisfy the conditions given by Eq.(3.1). [these are guaranteed by the δ-functions in the r.h.s. of Eq.(3.2)]. In terms of statistical physics this is an averaging over the microcanonical ensemble of those pure states which satisfy the conditions on the mean values of the measured observables. Consequently, Eq.(3.2) represents the principle of the maximum entropy on the microcanonical ensemble under the constraint (3.1).

4. Bayesian inference for spin-1/2 pure states

In order to appreciate the beauty of the Bayesian inference for pure states and to understand the complexity of the reconstruction of impure states we present in this section a relatively simple example of the reconstruction of a pure state of the spin-1/2.

The rigorous way how to determine the parametric state space Ω is based on the diffeomorphism between Ω and quotient space ${}^{SU(n)}|_{U(n-1)}$, where n is the number of dimensions of the Hilbert space of the measured quantum system. In the particular case of the spin-1/2 we work with the commutative group $U(1)$ and the construction of Ω is very simple. The space Ω can be mapped into the so called Poincare sphere and the parameterized density operator (i.e. the point on the Poincare sphere) reads:

$$\hat{\rho}(\theta,\phi) = \frac{\hat{1} + \vec{r}\hat{\vec{\sigma}}}{2} = \frac{1}{2}(1 + \sin\theta\cos\phi\,\hat{\sigma}_x + \sin\theta\sin\phi\,\hat{\sigma}_y + \cos\theta\,\hat{\sigma}_z), \tag{4.1}$$

where $\phi \in (0, 2\pi)$, $\theta \in (0,\pi)$. The topology of the sphere determines also the integration measure for which we have $d_\Omega = \sin\theta\, d\theta d\phi$.

One possible choice of the complete set of observables (i.e., the quorum [2]) associated with the spin-1/2 are the spin projections for three orthogonal directions represented by Hermitian operators:

$$\hat{s}_i \equiv \frac{\hat{\sigma}_i}{2}, \qquad i = x, y, z \tag{4.2}$$

---

[2]The proof is based on the fact, that for the appropriate normalization factor B the integral $\frac{1}{B}\int_0^1 dx_1 \int_0^{(1-x_1)} dx_2 \ldots \int_0^{(1-x_1-x_2\ldots-x_{n-2})} dx_{n-1} x_1^{\alpha_1 N} x_2^{\alpha_2 N} \ldots x_{n-1}^{\alpha_{n-1}N}(1 - x_1 \ldots - x_{n-1})^{\alpha_n N}$, with coefficients α_i satisfying the condition $\sum_1^N \alpha_i = 1$, tends in the limit $N \to \infty$ to the integral of a product of delta functions: $\int_0^1 dx_1 \int_0^{(1-x_1)} dx_2 \ldots \int_0^{(1-x_1-x_2\ldots-x_{n-2})} dx_{n-1}\delta(x_1 - \alpha_1)\delta(x_2 - \alpha_2)\ldots\delta(x_{n-1} - \alpha_{n-1})$.

$\hat{\sigma}_z$	$\hat{\sigma}_x$	$\hat{\sigma}_y$	$\hat{\rho}$ via pure-state reconstruction	$\hat{\rho}$ via mixture-state reconstruction
$\uparrow$			$\hat{1} + \frac{1}{3}\hat{\sigma}_z$	$\hat{1} + \frac{1}{5}\hat{\sigma}_z$
$\uparrow^4$			$\hat{1} + \frac{2}{3}\hat{\sigma}_z$	$\hat{1} + \frac{1}{2}\hat{\sigma}_z$
$\uparrow^5\downarrow$			$\hat{1} + \frac{1}{2}\hat{\sigma}_z$	$\hat{1} + \frac{2}{5}\hat{\sigma}_z$
$\uparrow^{10}\downarrow^2$			$\hat{1} + \frac{4}{7}\hat{\sigma}_z$	$\hat{1} + \frac{1}{2}\hat{\sigma}_z$
$\uparrow$	$\downarrow$		$\hat{1} - \frac{1}{3}\hat{\sigma}_x + \frac{1}{3}\hat{\sigma}_z$	$\hat{1} - \frac{1}{5}\hat{\sigma}_x + \frac{1}{5}\hat{\sigma}_z$
$\uparrow^3\downarrow$	$\uparrow^4$		$\hat{1} + \frac{10}{37}\hat{\sigma}_x + \frac{26}{37}\hat{\sigma}_z$	$\hat{1} + \frac{68}{309}\hat{\sigma}_x + \frac{158}{309}\hat{\sigma}_z$
$\uparrow^4\downarrow^2$	$\uparrow^5\downarrow$		$\hat{1} + \frac{704}{2601}\hat{\sigma}_x + \frac{1460}{2601}\hat{\sigma}_z$	$\hat{1} + \frac{218}{1105}\hat{\sigma}_x + \frac{646}{1105}\hat{\sigma}_z$
$\downarrow$	$\uparrow$	$\uparrow$	$\hat{1} - \frac{1}{3}\hat{\sigma}_x + \frac{1}{3}\hat{\sigma}_y + \frac{1}{3}\hat{\sigma}_z$	$\hat{1} - \frac{1}{5}\hat{\sigma}_x + \frac{1}{5}\hat{\sigma}_y + \frac{1}{5}\hat{\sigma}_z$
$\uparrow^3\downarrow$	$\uparrow^4$	$\uparrow^4$	$\hat{1} + \frac{831}{3503}\hat{\sigma}_x + \frac{2026}{3503}\hat{\sigma}_y + \frac{2026}{3503}\hat{\sigma}_z$	$\hat{1} + \frac{1051}{5253}\hat{\sigma}_x + \frac{2382}{5253}\hat{\sigma}_y + \frac{2382}{5253}\hat{\sigma}_z$
$\uparrow^4\downarrow^2$	$\uparrow^5\downarrow$	$\uparrow^5\downarrow$	$\hat{1} + \frac{47109}{169636}\hat{\sigma}_x + \frac{99310}{169636}\hat{\sigma}_y + \frac{99310}{169636}\hat{\sigma}_z$	$\hat{1} + \frac{279193}{1446325}\hat{\sigma}_x + \frac{593708}{1446325}\hat{\sigma}_y + \frac{593708}{1446325}\hat{\sigma}_z$
$\uparrow^3\downarrow$	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\hat{1} + \frac{101}{161}\hat{\sigma}_z$	$\hat{1} + \frac{413}{1389}\hat{\sigma}_z$
$\uparrow^6\downarrow^2$	$\uparrow^4\downarrow^4$	$\uparrow^4\downarrow^4$	$\hat{1} + \frac{88}{117}\hat{\sigma}_z$	$\hat{1} + \frac{3125918}{8023325}\hat{\sigma}_z$

Tab.1. Results of *a posteriori* Bayesian estimation of density operators of the spin-1/2 are presented for two different cases: (1) when it is *a priori* assumed that the spin is in a pure state and (2) when no *a priori* constraint on the state is imposed. In this second case the generalized Bayesian scheme has been applied. The density operators are given up to normalization factor 1/2.

where $\hat{\sigma}_i$ are the Pauli spin operators. The observables $\hat{s}_i$ have the spectrum equal to $\pm\frac{1}{2}$. In what follows we will distinguish between these two possible measurement results only by the sign, i.e. $s = \pm 1$. The projectors $\hat{P}_{s,\hat{s}_i}$ onto the corresponding eigenvectors read

$$\hat{P}_{s,\hat{s}_i} = \frac{\hat{1} + s\hat{\sigma}_i}{2}; \qquad i = x, y, z \tag{4.2a}$$

and the conditional probabilities associated with this kind of the measurement can be written as

$$p(s, \hat{s}_i|\hat{\rho}(\theta,\phi)) = \frac{1 + s\, r_i}{2}; \qquad i = x, y, z. \tag{4.2b}$$

Now using the procedure described in Section 2 we can construct an *a posteriori* density operator $\hat{\rho}(\{\}_N)$ based on a given sequence of measurement outcomes. In Table 1 we present results of the Bayesian inference for the spin-1/2 based on the fictitious measurements of three spin components performed on three Stern-Gerlach apparatuses.

First let us assume that just one Stern-Gerlach apparatus measuring the spin $\hat{s}_z = \hat{\sigma}_z/2$ is used (i.e., this measurement setup fixes a specific observation level). If the first measurement of the spin $\hat{s}_z$ gives us the result $\uparrow$ (i.e. $s = +1$), then under the assumption that the spin-1/2 is in a pure state we can use the Bayesian inference scheme and we obtain the *a posteriori* estimation for the density operator presented in the first line of Tab.1. With the increase of the number of measurements we improve the *a posteriori* estimation of the density operator on the given observation level. In particular, let us assume that in twelve measurements we have detected ten spins up and two spins down (i.e. $\uparrow^{10}\downarrow^2$). The corresponding *a posteriori* density operator is presented in Tab.1. If we use the outcome of the twelve measurements then we can approximate the mean value of the operator $\hat{\sigma}_z$ to be equal to 2/3. If this would be a "true"mean value (i.e. obtained in an infinite sequence of measurements) of the operator $\hat{\sigma}_z$ then with help of the Jaynes principle of the maximum entropy we would find for the density operator the expression $\hat{\rho} = \frac{1}{2}(\hat{1} + \frac{2}{3}\hat{\sigma}_z)$.

We can extend the observation level and we can consider also the measurement of the spin components $\hat{s}_x$ and $\hat{s}_y$. In Tab.1 we present results of "numerical experiments" for a given set of outcomes. In particular, the third line of the table represents the simulation of measurements on the complete observation level when all three spin components of the spin-1/2 are measured. Here we also present a result of the Bayesian inference for the density operator. With the help of these results we can also approximately estimate the mean values of corresponding operators for which we find $\langle\hat{\sigma}_z\rangle = 1/3$ and $\langle\hat{\sigma}_x\rangle = \langle\hat{\sigma}_y\rangle = 2/3$. These mean values fulfill the "purity condition"

$$\langle\hat{\sigma}_x\rangle^2 + \langle\hat{\sigma}_y\rangle^2 + \langle\hat{\sigma}_z\rangle^2 = 1, \tag{4.3}$$

which means that the measured state is a pure state (providing the condition (4.3) is fulfilled also in the limit $N \to \infty$) Consequently, the Bayesian reconstruction scheme can be safely used in the limit $N \to \infty$ and the *a posteriori* density operator reads:

$$\hat{\rho} = \frac{1}{N_0}\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta\, \delta(\langle\hat{\sigma}_z\rangle - \cos\theta)\delta(\langle\hat{\sigma}_y\rangle - \sin\theta\sin\phi)\delta(\langle\hat{\sigma}_x\rangle - \sin\theta\cos\phi)\hat{\rho}(\theta,\phi). \tag{4.4}$$

This expression has an appealing geometrical interpretation: the three δ-functions correspond to three specific orbits on the Poincare sphere each of which is associated with a set of pure states which posses the measured value of a given observable $\hat{s}_i$. When we substitute the density operator (4.1) into the Eq.(4.4) we find the *a posteriori* density operator in the form

$$\hat{\rho} = \frac{1}{2}\left(\hat{1} + \langle\hat{\sigma}_x\rangle\hat{\sigma}_x + \langle\hat{\sigma}_y\rangle\hat{\sigma}_y + \langle\hat{\sigma}_z\rangle\hat{\sigma}_z\right), \tag{4.5}$$

where the mean values of the observables fulfill the condition (4.3). This density operator describes a point on the Poincare sphere which can be represented as an intersection of three "orbits" associated with three constraints described by δ-functions in Eq.(4.4). If the three orbits have no intersection reconstruction scheme fails. Consequently, there does not exist a *pure* state with the given mean values of the measured observables. In the second part of Tab.1 (below the double line) we present a numerical simulation of the measurement in which all three observables are measured. The first sequence of the measurement consists of four measurements over each spin. Based on an *a posteriori* estimation the mean values of the operators $\hat{\sigma}_x$ and $\hat{\sigma}_y$ are equal to zero (each of the spin components is measured "up" the same number of times as "down"). On the other hand from the measurement of the $\hat{\sigma}_z$ component one can make a simple estimation that its mean is equal to 1/2. But the Bayesian inference gives the result equal to 101/161 (i.e., this number is larger than 1/2). Moreover, with the increase of the number of the measurements Bayes estimation deviates even larger from what would be the estimation based on the Jaynes principle. The reason for this contradiction lies in the *a priori* assumption about the purity of the reconstructed state, i.e. the mean values of the spin components do not fulfill the condition (4.3) and so the Bayesian method *cannot* be applied safely in the present case. The larger the number of measurement more clear the inconsistency is seen and, as seen from Eq.(4.4), in the limit of infinite number of measurements the Bayesian method fails completely. On the other hand the Jaynes method can be applied safely in this case. The point is that this method is not based on an *a priori* assumption about the purity of the reconstructed state. The Jaynes principle is associated with maximization of entropy on the grand canonical ensemble which means that all states (pure and impure) are taken into account. This, obviously, is an advantage of the Jaynes reconstruction scheme. Nevertheless this method can be applied only when "exact" mean values of observables are known.

5. Bayesian reconstruction of impure states

If a quantum system P is in an impure state we can consider it as being entangled with some other quantum system R (reservoir). We assume a system S (composed of P and R) which itself is in a pure state $|\Psi\rangle$. The density operator $\hat{\rho}_P$ of the system P is then obtained via tracing over the reservoir degrees of freedom:

$$\hat{\rho}_P = \mathrm{Tr}_R\,[\hat{\rho}_S]; \qquad \hat{\rho}_S = |\Psi\rangle\langle\Psi|. \tag{5.1}$$

Once the system S is in a pure state, then we can safely apply the Bayesian reconstruction scheme as described in Section 2. The reconstruction itself is based only on data associated with measurements performed on the system P. When the density operator $\hat{\rho}_S$ is *a posteriori* estimated then by tracing over the reservoir degrees of freedom we obtain the *a posteriori* density operator $\hat{\rho}_P$ for the system P (with no *a priori* constraint on the purity of the state of the system P).

To make our reconstruction scheme for impure state selfconsistent we have to chose the reservoir R uniquely. This can be done with the help of the Schmidt theorem (see Ref. [13]) from which it follows that if the composite system S is in a pure state $|\Psi\rangle$ then its state vector can be written in the form:

$$\Psi = \sum_{i=1}^{M} c_i|\alpha_i\rangle_P \otimes |\beta_i\rangle_R, \tag{5.2}$$

where $|\alpha_i\rangle_P$ and $|\beta_i\rangle_R$ are elements from two specific orthonormalized bases associated with the subsystems P and R, respectively, and c_i are appropriate complex numbers satisfying the normalization condition $\sum|c_i|^2 = 1$. The maximal index of summation (M) in Eq.(5.2) is given by the dimensionality of the Hilbert space of the system P. In other words, when we apply the Bayesian method to the case of impure states of M-level system, it is sufficient to "couple" this system to another system, which has effectively the same number of levels. Due to the fact that we measure only observables of the first subsystem P particular form of states $|\beta_i\rangle_R$ of the second subsystem R does not affect results of the reconstruction.

6. Reconstruction of impure state of spin-1/2

We illustrate the Bayesian inference scheme for impure states on an example of the spin-1/2. Following the general idea described in Section 5 we have to consider two spins-1/2. One of these spins represents the reservoir. Here the state space $\Omega = ^{SU(n)}|_{U(n-1)}$ is parameterized by six coordinates (we note that $\dim_\Omega = \dim_{SU(n)} - \dim_{U(n-1)}$, for more details see Ref.[14]). In order to parameterize properly the state space Ω_S we use the Schmidt decomposition and we represent the state vector $|\Psi\rangle$ describing two spins-1/2 as:

$$|\Psi\rangle = A|1_1\rangle \otimes |1_2\rangle + B|0_1\rangle \otimes |0_2\rangle, \tag{6.1}$$

where $|0_i\rangle, |1_i\rangle$, are two appropriate orthonormalized bases in H^2 and A, B are two complex numbers satisfying the condition $|A|^2 + |B|^2 = 1$. The corresponding density operator of a pure state in Ω_S then reads

$$\hat{\rho} = |A|^2|1_1\rangle\langle 1_1| \otimes |1_2\rangle\langle 1_2| + AB^*|1_1\rangle\langle 0_1| \otimes |1_2\rangle\langle 0_2|$$
$$+A^*B|0_1\rangle\langle 1_1| \otimes |0_2\rangle\langle 1_2| + |B|^2|0_1\rangle\langle 0_1| \otimes |0_2\rangle\langle 0_2|. \tag{6.2}$$

Projectors in the Hilbert space H^2 read $(\hat{1} - \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)})$ or $(\hat{1} + \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)})$, $[j = 1, 2]$, where $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ are two arbitrary unity vectors. The operators $|0_j\rangle\langle 1_j|$ and $|1_j\rangle\langle 0_j|$ are determined with the help of the identity $|0_j\rangle\langle 1_j|(\hat{1} + \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)})|1_j\rangle\langle 0_j| = (\hat{1} - \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)})$. This gives us the relation $|1_j\rangle\langle 0_j| = e^{i\psi_i}(\vec{k}^{(j)}\hat{\vec{\sigma}}^{(j)} + i\vec{l}^{(j)}\hat{\vec{\sigma}}^{(j)})$. Here the vectors $\vec{k}^{(j)}$ are two arbitrarily chosen unity vectors which satisfy the condition $\vec{k}^{(j)} \perp \vec{r}^{(j)}$ and $\vec{l}^{(j)}$ are equal to vector products $\vec{l}^{(j)} = \vec{r}^{(j)} \times \vec{k}^{(j)}$. Using the parameterization $|A| = \cos(\alpha/2)$ and $|B| = \sin(\alpha/2)$ we find $\hat{\rho}$ in the following form [14]:

$$\hat{\rho}(\alpha, \psi, \phi_1, \theta_1, \phi_2, \theta_2) = \frac{\hat{1} \otimes \hat{1}}{4} + \frac{\vec{r}^{(1)}\hat{\vec{\sigma}} \otimes \vec{r}^{(2)}\hat{\vec{\sigma}}}{4} + \cos(\alpha)\left[\frac{\vec{r}^{(1)}\hat{\vec{\sigma}} \otimes \hat{1}}{4} + \frac{\hat{1} \otimes \vec{r}^{(2)}\hat{\vec{\sigma}}}{4}\right] \tag{6.3}$$

$$+\sin(\alpha)\cos\psi\left[\frac{\vec{k}^{(1)}\hat{\vec{\sigma}} \otimes \vec{k}^{(2)}\hat{\vec{\sigma}}}{4} - \frac{\vec{l}^{(1)}\hat{\vec{\sigma}} \otimes \vec{l}^{(2)}\hat{\vec{\sigma}}}{4}\right] - \sin(\alpha)\sin\psi\left[\frac{\vec{k}^{(1)}\hat{\vec{\sigma}} \otimes \vec{l}^{(2)}\hat{\vec{\sigma}}}{4} + \frac{\vec{l}^{(1)}\hat{\vec{\sigma}} \otimes \vec{k}^{(2)}\hat{\vec{\sigma}}}{4}\right],$$

where $\psi, \phi_1, \phi_2 \in (0, 2\pi)$; $\alpha, \theta_1, \theta_2 \in (0, \pi)$ and

$$\vec{k}^{(j)} = (\sin\phi_j, -\cos\phi_j, 0), \qquad \vec{l}^{(j)} = (\cos\theta_j\cos\phi_j, \cos\theta_j\sin\phi_j, -\sin\theta_j)$$
$$\vec{r}^{(j)} = (\sin\theta_j\cos\phi_j, \sin\theta_j\sin\phi_j, \cos\theta_j). \tag{6.4}$$

Once we have parameterized the state space Ω_S and the corresponding density operator we have to find the invariant integration measure d_Ω. In differential geometry this integration measure is a global object - the so called invariant volume form ω. The condition that d_Ω is invariant under the action of each group element $U \in SU(n)$ is equivalent to the requirement that the Lie derivative of ω with respect to the fundamental field of action of the group $SU(n)$ in the space Ω is zero. This formulation of the problem how to find the integration measure leads to a system of linear differential

equations [14]. We note, that even for this relatively simple quantum system particular calculations are technically quite complicated [14]. However, the result is simple:

$$d_\Omega = \cos^2\alpha\sin\alpha\sin\theta_1\sin\theta_2\, d\alpha d\psi d\phi_1 d\theta_1 d\phi_2 d\theta_2. \tag{6.5}$$

The observables associated with the system P in analogy with Eq.(4.2) do read:

$$\hat{s}_i^{(1)} = \frac{\hat{\sigma}_i \otimes \hat{1}}{2}, \qquad i = x, y, z. \tag{6.6}$$

The projectors and conditional probabilities associated with the measurement of the system P are defined as:

$$\hat{P}_{s,\hat{s}_i^{(1)}} = \frac{(\hat{1} + s\hat{\sigma}_i)}{2} \otimes \hat{1} \quad \longrightarrow \quad p(s, \hat{s}_i^{(1)}|\hat{\rho}(\alpha\ldots)) = \frac{1}{2} + s\frac{\cos(\alpha)}{2}r_i^{(1)}. \tag{6.7}$$

Now we can apply general rules of Bayesian inference presented in Section 2 and we can evaluate the *a posteriori* estimation for the density operator $\hat{\rho}_S$ from which we find the density operator of the system P. In Tab.1 we present results of numerical reconstruction of the density operator of the spin-1/2. This reconstruction is based on exactly the same set of data as discussed is Section 4. The only difference consists in different *a priori* assumptions. In the first case we have assumed that the spin-1/2 is in a pure state, while in the second case we have lifted this constraint. We see that the results of the *a posteriori* estimation do differ, which is caused by a different topology of state spaces on which the estimation (reconstruction) is performed. If the measured data are consistent with the *a priori* assumption that the reconstructed system is in a pure state then both reconstruction procedures work well (except the convergence is slower in the case of the second method, i.e. more data are needed, because less *a priori* information is available). Nevertheless the power of the second (generalized Bayesian inference) is seen when the reconstructed system is in an impure state and a large number of measurements has been performed. In this case the standard Bayesian reconstruction fails (see Section 4) while the generalized Bayesian inference gives us a consistent result. In particular, in the limit $N \to \infty$ we find for the density operator of the system P the following expression

$$\hat{\rho} = \frac{1}{N_0}\int_{-1}^{1} y^2 dy \int_0^{2\pi} d\phi_1 \int_0^{\pi} \sin\theta_1 d\theta_1\, \delta(\langle\hat{\sigma}_z\rangle - y\cos\theta)\delta(\langle\hat{\sigma}_y\rangle - y\sin\theta\sin\phi)$$
$$\times\delta(\langle\hat{\sigma}_x\rangle - y\sin\theta\cos\phi)(\hat{1} + y\sin\theta_1\cos\phi_1\hat{\sigma}_x + y\sin\theta_1\sin\phi_1\hat{\sigma}_y + y\cos\theta_1\hat{\sigma}_z), \tag{6.8}$$

where $y = \cos\alpha$. Straightforward calculation shows us that the equation (6.8) leads to exactly the same result as the standard Bayesian estimation given by Eq.(4.5) except there are no restrictions on the mean values of spin operators. Therefore this operator can describe pure as well as impure states.

7. Conclusions

In the present paper we have generalized the standard Bayesian scheme of quantum inference for a reconstruction of impure states of quantum systems. We have solved the problem of ambiguity in a definition of the invariant integration measure in a space of impure quantum-mechanical states which has been the main obstacle in application of Bayesian methods for statistical mixtures [11]. As an illustration, we analyzed the reconstruction of states of the spin-1/2.

Finally we note, that the form of the integral in Eq.(6.8) indicates that the equal-weighted averaging in the generalized Bayesian scheme is performed on the grand canonical ensemble (i.e., in the case of the spin-1/2 all points inside the Poincare sphere are taken into account). This means that the expression (6.8) for the density operator maximizes the entropy under the given constraints on the measured mean values. Consequently, the generalized Bayesian inference in the limit of large number of measurements is equal to the Jaynes method of reconstruction based of the *MaxEnt* principle.

Acknowledgements This work was in part supported by the Grant agency VEGA of the Slovak Academy of Sciences. We acknowledge the support by the East-West Program of the Austrian Academy of Sciences under the contract No. 45.367/6-IV/3a/95 of the Österreichisches Bundesministerium für Wissenschaft und Forschung.

References

[1] R. Omnès: *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1994);

[2] W. Band, J.L. Park: *Am. J. Phys.* **47** (1979) 188;

[3] V. Bužek, G. Adam, G. Drobný: *Ann. Phys. (N.Y.)* **245** (1996) 36; and references therein;

[4] K. Vogel, H. Risken: *Phys. Rev. A* **40** (2847) 1989; for the review see U. Leonhardt, H. Paul: *Quant. Electron.* **19** (89) 1995;

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

[6] K. Wódkiewicz: *Phys. Rev. Lett.* **52** (1064) 1984; see also V. Bužek, C.H. Keitel, P.L. Knight: *Phys. Rev. A* **51** (1995) 2575;

[7] N.G. Walker, J.E. Carroll: *Opt. Quant. Electron.* **18** (335) 1986;

[8] E.T. Jaynes: *Phys. Rev.* **108** (1957) 171; *ibid.* **108** (1957) 620; see also E.T. Jaynes: *Am. J. Phys.* **31** (1963) 66;

[9] C.W. Helstrom: *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976);

[10] A.S. Holevo: *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982);

[11] K.R.W. Jones: *Ann. Phys. (N.Y.)* **207** (1991) 140;

[12] E.G. Larson, P.R. Dukes: in *Maximum Entropy and Bayesian Methods*, eds. W.T. Grandy, Jr. and L.H. Schick (Kluwer Academic Press, 1991), p. 181;

[13] A.K. Ekert: *Correlations in Quantum Optics*, Ph.D. thesis (University of Oxford, 1991); and references therein;

[14] R. Derka, V. Bužek, G. Adam: *Bayesian methods vs Jaynes principle*, unpublished.