Annals of Physics **266**, 454–496 (1998)
Article No. PH985802

Reconstruction of Quantum States of Spin Systems: From Quantum Bayesian Inference to Quantum Tomography

V. Bužek

Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, England; and Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovakia

R. Derka

Department of Physics, Oxford University, Parks Road, Oxford OX1 3PU, England

G. Adam

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

and

P. L. Knight

Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, England

Received July 1, 1997; revised December 17, 1997

We study in detail the reconstruction of spin-1/2 states and analyze the connection between (1) quantum Bayesian inference, (2) reconstruction via the Jaynes principle of maximum entropy, and (3) complete reconstruction schemes such asdiscrete quantum tomography. We derive an expression for a density operator estimated via Bayesian quantum inference in the limit of an infinite number of measurements. This expression is derived under the assumption that the reconstructed system is in a pure state. In this case the estimation corresponds to averaging over a microcanonical ensemble of pure states satisfying a set of constraints imposed by the measured mean values of the observables under consideration. We show that via a "purification" ansatz, statistical mixtures can also be consistently reconstructed via the quantum Bayesian inference scheme. In this case the estimation corresponds to averaging over the generalized grand canonical ensemble of states satisfying the given constraints, and in the limit of large number of measurements this density operator is equal to the generalized canonical density operator, which can be obtained with the help of the Jaynes principle of the maximum entropy. We also discuss inseparability of reconstructed density operators of two spins-1/2. © 1998 Academic Press

0003-4916/98 $25.00

I. INTRODUCTION

The concept of a quantum state represents one of the most fundamental pillars of the paradigm of quantum theory [1–3]. Contrary to its mathematical elegance and convenience in calculations, the physical interpretation of a quantum state is not so transparent. The problem is that the quantum state (described either by a state vector, or density operator or a phase-space probability density distribution) does not have a well defined objective status, i.e., a state vector is not an *objective* property of a particle. According to Peres (see [1, p. 374]), "There is no physical evidence whatsoever that every physical system has at every instant a well defined state.... In strict interpretation of quantum theory these mathematical symbols [i.e., state vectors] represent *statistical information* enabling us to compute the probabilities of occurrence of specific events." Once this point of view is adopted then it becomes clear that any "measurement" or reconstruction of a density operator (or its mathematical equivalent) can be understood exclusively as an expression of our knowledge about the quantum mechanical state based on a certain set of measured data. To be more specific, any quantum-mechanical reconstruction scheme is nothing more than an *a posteriori* estimation of the density operator of a quantum-mechanical (microscopic) system based on data obtained with the help of a macroscopic measurement apparatus [3]. The recognition of quantum-state measurement and reconstruction schemes stems from potential applications of these schemes in atomic, molecular, and condensed-matter physics, as well as quantum-information processing [4].

The quality of the reconstruction depends on the "quality" of the measured data and the efficiency of the reconstruction procedure with the help of which the data analysis is performed. In particular, we can specify three different situations. First, in the case when all system observables are precisely measured, the complete reconstruction of an initially unknown state can be performed (we will call this the reconstruction on the complete observation level). Second, when just part of the system observables is precisely measured then one cannot perform a complete reconstruction of the measured state. Nevertheless, the reconstruction of the density operator of the quantum system under consideration can be performed in this case with the help of the Jaynes principle of *Maximum Entropy* (see below). This reconstructed density operator uniquely determines mean values of the measured observables and in addition it can provide us with nontrivial estimations of unmeasured observables (we will denote this type of scheme as reconstruction on incomplete observation levels). Finally, when measurement does not provide us with sufficient information to specify the exact mean values (or probability distributions) but only the frequencies of appearances of eigenstates of the measured observables, then one can perform an estimation (reconstruction based on quantum Bayesian inference) which is the "best" with respect to the given measured data and the *a priori* knowledge about the state of the measured system.

The main purpose of the present paper is to demonstrate the intrinsic connection between quantum Bayesian inference, incomplete quantum state reconstruction based on the *MaxEnt* principle and complete reconstruction of a quantum-mechanical state. We start the paper with a brief description of various reconstruction schemes and we set up the scene for our further discussion. In Sections III and IV we review the Jaynes principle of maximum entropy and the quantum Bayesian inference scheme, respectively. In Section V we present a proof that the standard Bayesian inference (developed for the reconstruction of pure states) in the limit of infinite number of measurements is equivalent to an averaging over generalized microcanonical ensembles under the given constraints. In Section VI we show that with the help of the purification ansatz the Bayesian inference scheme can be used for the reconstruction of impure states. In subsequent sections we concentrate our attention on the reconstruction for one and two spins-1/2 systems. In particular, in Section VII we briefly discuss a reconstruction of spin states via the Jaynes *MaxEnt* principle. In Section VIII we illustrate how Bayesian inference works for a single spin-1/2 system when it is *a priori* known that this system is prepared in a pure state. Section IX is devoted to a state reconstruction of two spins-1/2 via quantum Bayesian inference. In Section X we present a systematic analysis of Bayesian reconstruction of a single spin-1/2 under the *a priori* assumption that this system is prepared in a statistical mixture. We summarize our results in Section XI.

II. QUANTUM-STATE RECONSTRUCTION

A. *Complete Observation Level*

Provided that all system observables (i.e., the quorum [5, 6]) have been precisely measured, then the density operator of a quantum-mechanical system can be *completely* reconstructed (i.e., the density operator can be uniquely determined based on the available data). In principle, we can consider two different schemes for reconstruction of the density operator (or, equivalently, the Wigner function) of the given quantum-mechanical system. The difference between these two schemes is based on the way in which information about the quantum-mechanical system is obtained. The first type of measurement is such that on each element of the ensemble of the measured states only a *single* observable is measured. In the second type of measurement a *simultaneous* measurement of conjugate observables is assumed. We note that in both cases we will assume ideal, i.e., unit-efficiency, measurements.

1. *Quantum Tomography*

When the single-observable measurement is performed, a *distribution* $W_{|\Psi\rangle}(A)$ for a particular observable $\hat{A}$ of the state $|\psi\rangle$ is obtained in an unbiased way [7], i.e., $W_{|\Psi\rangle}(A) = |\langle\Phi_A|\Psi\rangle|^2$, where $|\Phi_A\rangle$ are eigenstates of the observable $\hat{A}$ such that $\sum_A |\Phi_A\rangle\langle\Phi_A| = \hat{1}$. Here a question arises: What is the *smallest* number of distributions $W_{|\Psi\rangle}(A)$ required to determine the state uniquely? If we consider the reconstruction of the state of a harmonic oscillator, then this question is directly related to the so-called Pauli problem [8] of the reconstruction of the wave-function from distributions $W_{|\Psi\rangle}(q)$ and $W_{|\Psi\rangle}(p)$ for the position and momentum of the state $|\Psi\rangle$. As shown by Gale *et al.* [9] the knowledge of $W_{|\Psi\rangle}(q)$ and $W_{|\Psi\rangle}(p)$ is not in general sufficient for a complete reconstruction of the wave function. In contrast, one can consider an *infinite* set of distributions $W_{|\Psi\rangle}(x_\theta)$ of the rotated quadrature $\hat{x}_\theta = \hat{q}\cos\theta + \hat{p}\sin\theta$. Each distribution $W_{|\Psi\rangle}(x_\theta)$ can be obtained from a measurement of a *single* observable $\hat{x}_\theta$, in which case a detector (filter) is prepared in an eigenstate $|x_\theta\rangle$ of this observable. It has been shown by Vogel and Risken [10, 11] that, from an infinite set (in the case of the harmonic oscillator) of the measured distributions $W_{|\Psi\rangle}(x_\theta)$ for all values of θ such that $[0 < \theta \leqslant \pi]$, the Wigner function can be reconstructed uniquely via the inverse Radon transformation. This scheme for reconstruction of the Wigner function (i.e., the *optical homodyne tomography*) has recently been realized experimentally by Raymer and his co-workers [12].

Quantum-state tomography can be applied not only to optical fields but for reconstruction of other physical systems. In particular, recently Janicke and Wilkens [13] have suggested that Wigner functions of atomic waves can be tomographically reconstructed. Kurtsiefer *et al.* [14] have performed experiments in which Wigner functions of matter wave packets have been reconstructed. Yet another example of the tomographic reconstruction is a reconstruction of Wigner functions of vibrational states of trapped atomic ions theoretically described by a number of groups [15] and experimentally measured by Leibfried *et al.* [16]. Vibrational motional states of molecules have also been reconstructed by this kind of quantum tomography by Dunn *et al.* [17].

Leonhardt [18] has recently developed a theory of quantum tomography of discrete Wigner functions describing states of quantum systems with finite-dimensional Hilbert spaces (for instance, angular momentum or spin). We note that the problem of reconstruction of states of finite-dimensional systems is closely related to various aspects of quantum information processing, such as reading of registers of quantum computers [19]. This problem also emerges when states of atoms are reconstructed (see, for instance, [20]).

Here we stress once again, that reconstruction on the complete observation level (such as quantum tomography) is a *deterministic* inversion procedure which helps us to "rewrite" measured data in the more convenient form of a density operator (Wigner function) of the measured state.

2. *Filtering with Quantum Rulers*

For the case of simultaneous measurement of two non-commuting observables (let us say $\hat{q}$ and $\hat{p}$), it is not possible to construct a joint eigenstate of these two operators, and therefore it is inevitable that the simultaneous measurement of two non-commuting observables introduces additional noise (of quantum origin) into measured data. This noise is associated with Heisenberg's uncertainty relation and it results in a specific "smoothing" (equivalent to a reduction of resolution) of the original Wigner function of the system under consideration (see Refs. [21] and [22]).

To describe the process of simultaneous measurement of two non-commuting observables, Wódkiewicz [23] has proposed a formalism based on an operational probability density distribution which explicitly takes into account the action of the measurement device modelled as a "filter" (quantum ruler). In particular, if the filter is considered to be in its vacuum state then the corresponding operational probability density distributions is equal to the Husimi (Q) function [21]. The Q function of optical fields has been experimentally measured using such an approach by Walker and Carroll [24]. The direct experimental measurement of the operational probability density distribution with the filter in an arbitrary state is feasible in an 8-port experimental setup of the type used by Noh *et al.* [25].

We note that propensities, and in particular Q-functions, can also be associated with discrete phase space and they can in principle be measured directly [26]. These discrete probability distributions contain complete information about density operators of measured systems. Consequently, these density operators can be uniquely determined from the discrete-phase space propensities.

B. *Reduced Observation Levels and MaxEnt Principle*

As we have already indicated, it is well understood that density operators can, in principle, be uniquely reconstructed using either the single observable measurements (optical homodyne tomography) or the simultaneous measurement of two non-commuting observables. The completely reconstructed density operator contains information about *all* independent moments of the system operators. For example, in the case of the quantum harmonic oscillator, the knowledge of the density operator is equivalent to the knowledge of all moments $\langle (\hat{a}^\dagger)^m \hat{a}^n \rangle$ of the creation ($\hat{a}^\dagger$) and annihilation ($\hat{a}$) operators.

In many cases it turns out that the state of a harmonic oscillator is characterized by an *infinite* number of independent moments $\langle (\hat{a}^\dagger)^m \hat{a}^n \rangle$ (for all m and n). Analogously, the state of a quantum system in a finite-dimensional Hilbert space can be characterized by a very large number of independent parameters. A *complete* measurement of these moments would take an infinite time to perform. This means that even though the density operator can in principle be reconstructed the collection of a complete set of experimental data points is (in principle) a never ending process. In addition the data processing and numerical reconstruction of the density operator are time consuming. Therefore experimental realization of the reconstruction of the density operators for many systems can be difficult.

In practice, it is possible tomeasure just a finite number of independent moments of the system operators, so that only a subset $\hat{G}_\nu$ ($\nu = 1, 2, ..., n$) of observables from the quorum (this subset constitutes the so-called observation level [27]) is measured. In this case, when the complete information about the system is not available, one needs an additional criterion which would help to reconstruct (or estimate) the density operator uniquely. Provided mean values of all observables on the given observation level are measured precisely, then the density operator of the system under consideration can be reconstructed with the help of the Jaynes principle of maximum entropy (the so-called *MaxEnt* principle) [27]. The MaxEnt principle provides us with a very efficient prescription to reconstruct density operators of quantum-mechanical systems providing the mean values of a given set of observables are known. It works perfectly well for systems with infinite Hilbert spaces (such as the quantum-mechanical harmonic oscillator) as well as for systems with finite-dimensional Hilbert spaces (such as spin systems). If the observation level is composed of the quorum of the observables (i.e., the complete observation level), then the MaxEnt principle represents an alternative to quantum tomography, i.e., both schemes are equally suitable for the analysis of the tomographic data (for details see [28]). To be specific, the observation level in this case is composed of all projectors associated with probability distributions of rotated quadratures. The power of the MaxEnt principle can be appreciated in analyses of incomplete tomographic data (equivalent to a reconstruction of the Wigner function in a discrete phase space). In particular, Wiedemann [29] has performed a numerical reconstruction of the Wigner function from incomplete tomographic data based on the MaxEnt principle as discussed by Bužek *et al.* [28]. Wiedemann has shown that in particular cases *MaxEnt* reconstruction from incomplete tomographic data can be several orders better than a standard tomographic inversion. This result suggests that the MaxEnt principle is the conceptual basis underlying incomplete tomographic reconstruction (irrespective whether this is employed in continuous or discrete phase spaces).

C. *Incomplete Measurement and Bayesian Inference*

It has to be stressed that the Jaynes principle of maximum entropy can be consistently applied only when *exact* mean values of the measured observables are available. This condition implicitly assumes that an infinite number of repeated measurements on different elements of the ensemble has to be performed to reveal the exact mean value of the given observable. In practice only a *finite* number of measurements can be performed. What is obtained from these measurements is a specific set of data indicating the number of times the eigenvalues of given observables have appeared (which in the limit of an 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 [30], Holevo [31], and Jones [32] 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 to be reconstructed is prepared in a pure (although unknown) state. When the purity condition is fulfilled, then the observer can systematically estimate an *a posteriori* probability distribution in an abstract state space of the measured system. It is this probability distribution (conditioned by the assumed Bayesian prior) which characterizes observer's knowledge of the system after the measurement is performed. Using this probability distribution one can derive a reconstructed density operator, which however is subject to certain ambiguity associated with the choice of the *cost* function (see Ref. [30, p. 25]). In general, depending on the choice of the cost function one obtains different estimators (i.e., different reconstructed density operators). In this paper we adopt the approach advocated by Jones [32] when the estimated density operator is equal to the mean over all possible *pure* states weighted by the estimated probability distribution [see Eq. (4.4)]. We note once again that the quantum Bayesian inference has been developed for a reconstruction of *pure* quantum mechanical states and in this sense it corresponds to an averaging over a generalized *microcanonical* ensemble. Obviously, the mean of pure states is in general impure state. The deviation of the reconstructed density operator from pure states (measured in terms of the von Neumann entropy) can then serve as a measure of the quality of the reconstruction.

In a real situation one can never design a state-preparation device to produce 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 (alternatively, we can say that the system under consideration is entangled with other quantum-mechanical systems). So now the question is how to use the Bayesian reconstruction scheme when the quantum-mechanical system under consideration is in an impure state (i.e., a statistical mixture). To apply the Bayesian inference scheme, one has to define precisely three concepts: (1) the abstract state space of the measured system; (2) the corresponding invariant integration measure of this space; and (3) the *prior* (i.e., the *a priori* known probability distribution on the given parametric state space). Once these objects are specified one can estimate an *a posteriori* probability distribution on the state space after each individual outcome of the measurement has been registered. Using this distribution the reconstructed density operator can be derived as an average over all state space.

III. JAYNES PRINCIPLE OF MAXIMUM ENTROPY

From a mathematical point of view any quantum state is an element $\hat{\rho}$ of a manifold $\{\hat{\rho} \in \mathbf{L}(\mathscr{H});\ \hat{\rho}^\dagger = \hat{\rho},\ \mathrm{Tr}(\hat{\rho}) = 1\}$ of Hermitian operators acting on the Hilbert space $\mathscr{H}$ with a trace equal to unity. This manifold is a convex space in which extreme points (i.e., pure states) play an exceptional role. Namely, each element of the manifold can be expressed as a convex linear combination of these extreme points. With each quantum state $\hat{\rho}$ an infinite number of macroscopic quantities $\langle \hat{O}_i \rangle$ (i.e., mean values of observables $\hat{O}_i$) are associated. These mean values are defined via a linear functional

$$\langle \hat{O}_i \rangle = \mathrm{Tr}(\hat{O}_i \hat{\rho}). \tag{3.1}$$

Pure states have also another exclusive property that, $\mathrm{Tr}(\hat{\rho}^k) = 1$, $k \in \mathbf{Z}$. Consequently, there exists an observable (at least as a Hermitian operator) which acquires in this state an exact mean value, with a zero dispersion. Purity of quantum states can also be quantified with the help of the von Neumann entropy [7]

$$S[\hat{\rho}] = -\mathrm{Tr}(\hat{\rho} \ln \hat{\rho}). \tag{3.2}$$

The von Neumann entropy is a positive functional which has many appealing properties such as concavity, continuity, additivity, and monotonicity (see [33]). This entropy provides us with an effective measure of the deviation of a quantum state from a pure state (for pure states $S[\hat{\rho}]=0$).

We have already noticed that complete information about a pure state can, in principle, be obtained via the measurement of the single observable. However, in general, this observable is not known *a priori* (we note that for statistical mixtures it does not exist at all). Therefore a determination (equivalent to a *reconstruction* or an *a posteriori estimation*) of a quantum state is in general based on a measurement of the mean values [see Eq. (3.1)] of a specific set of observables $\hat{O}_i$, $i=1, ..., n$. This set of observables specifies the so called observation level $\mathcal{O}$ [27, 28]. If the chosen observation level is *incomplete*, that is, if the density operator cannot be reconstructed uniquely from the measured mean values, then one should expect that there will be a number of density operators which fulfill the constraints given by Eq. (3.1). In this case one needs an additional criterion which would help to determine the reconstructed density operator uniquely. We note that if the *pure* state is incompletely reconstructed on a given observation level then the corresponding von Neumann entropy of the reconstructed state is *larger* than zero. According to Jaynes [27] this density operator must be the one which fulfills the constraints (3.1) and which in addition maximizes the Von Neumann entropy Eq. (3.2) (this is the so-called *MaxEnt* principle). In other words, an *a posteriori* estimation (reconstruction) of the density operator based on a measurement of a given set of observables is the most conservative assignment in the sense that it does not permit one to draw any conclusions not warranted by the data [34].

With the help of the *MaxEnt* principle one can reconstruct the density operator on a specific observation level $\mathcal{O}$. Following Jaynes [27] one can perform a reconstruction procedure in three steps:

(1) Firstly, for given observation level the generalized partition function

$$Z_{\mathcal{O}} = \mathrm{Tr}\left[\exp\left(-\sum_i \lambda_i \hat{O}_i\right)\right], \tag{3.3}$$

as a function of Lagrange multipliers λ_i has to be specified.

(2) Then the system of algebraic equations for the unknown λ_i,

$$\langle \hat{O}_i \rangle = \mathrm{Tr}(\hat{O}_i \hat{\rho}_{\mathcal{O}}) = -\frac{\partial}{\partial \lambda_i} \ln Z_{\mathcal{O}}(\lambda_1, ..., \lambda_n), \tag{3.4}$$

has to be solved. This system of algebraic equations corresponds to the constraints imposed by the measured mean values Eq. (3.1)

(3) Finally, the so-called generalized canonical density operator (i.e., the reconstructed density operator on a given observation level)

$$\hat{\rho}_{\mathcal{O}} = \frac{1}{Z_{\mathcal{O}}}\left[\exp\left(-\sum_i \lambda_i \hat{O}_i\right)\right]. \tag{3.5}$$

can be obtained, where λ_i are the solutions of Eqs. (3.4).

It is important to understand that in the reconstruction scheme via the MaxEnt principle, no *a priori* information about the state of the measured system is assumed. In other words, the reconstruction is performed on the most general state space of both pure and impure states of the system.

We stress here that the *exact* mean value of an arbitrary observable can only be obtained when a very large (in principle, infinite) number of measurements on individual elements of an ensemble is performed. On the other hand, it is a very legitimate question to ask "What is the best *a posteriori* estimation of a quantum state when a measurement is performed on a *finite* (arbitrarily small) number of elements of the ensemble?" To estimate the state of the system based on an incomplete set of data, one has to utilize more powerful estimation schemes such as the quantum Bayesian inference.

IV. QUANTUM BAYESIAN INFERENCE

The general idea of the Bayesian reconstruction scheme is based on manipulations with probability distributions in parametric state spaces Ω and A of the measured system and the measuring apparatus, respectively. The quantum Bayesian method as discussed in the literature [30–32] 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. The state space A of reading states of a measuring apparatus will be for the purpose of this paper associated with a discrete set of states.

In our case, when the standard Von Neumann measurement is considered, the apparatus states are intrinsically related to the set of mutually orthogonal projectors $\hat{P}_{\lambda_i, \hat{O}}$ associated with the spectrum of a given observable $\hat{O}$ (where λ_i are the corresponding eigenvalues).

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

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

$$p(\hat{O}, \lambda_i \,|\, \hat{\rho}) = \mathrm{Tr}(\hat{P}_{\lambda_i, \hat{O}}\hat{\rho}), \tag{4.1}$$

on the discrete space A is obtained. This conditional probability distribution specifies the 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 we have to specify an *a priori* distribution $p_0(\hat{\rho})$ defined on the space Ω. This distribution describes our initial knowledge concerning the measured system. Using the conditional probability distribution $p(\hat{O}, \lambda_i \mid \hat{\rho})$ and the *a priori* distribution $p_0(\hat{\rho})$ we can define the *joint* probability distribution $p(\hat{O}, \lambda_i; \hat{\rho})$,

$$p(\hat{O}, \lambda_i; \hat{\rho}) = p(\hat{O}, \lambda_i \mid \hat{\rho})\, p_0(\hat{\rho}), \tag{4.2}$$

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 (this assumption is related to the Laplace principle of indifference [34]).

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

$$p(\hat{\rho} \mid \hat{O}, \lambda_i) = \frac{p(\hat{O}, \lambda_i, \hat{\rho})}{\int_\Omega p(\hat{O}, \lambda_i, \hat{\rho})\, d_\Omega}, \tag{4.3}$$

from which the reconstructed density operator can be obtained [see Eq. (4.4)].

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 Nth measurement we use as an input for the prior distribution the conditional probability distribution given by the output of the $(N-1)$st measurement. However, we can equivalently define the N-trial measurement conditional probability $p(\{\ \}_N \mid \hat{\rho}) = \prod_{i=1}^N p(\hat{O}_i, \lambda_j \mid \hat{\rho})$ and applying the three-step procedure just once to obtain the reconstructed density operator

$$\hat{\rho}(\{\ \}_N) = \frac{\int_\Omega p(\hat{\rho} \mid \{\ \}_N)\, \hat{\rho}\, d_\Omega}{\int_\Omega p(\hat{\rho} \mid \{\ \}_N)\, d_\Omega}, \tag{4.4}$$

where $\hat{\rho}$ in the r.h.s. of Eq. (4.4) is a properly parameterized density operator in the state space Ω.

At this point we should mention one essential problem in the Bayesian reconstruction scheme, which is the determination of the integration measure d_Ω. 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 Ω is considered to be a space of mixed states formed by all convex combinations of elements of the original pure state space Ω. Although the Bayesian procedure itself does not require any special conditions imposed on the space Ω, the ambiguity in determination of the integration measure is the main obstacle in generalization of the Bayesian inference scheme for reconstruction of *a priori* impure quantum states. We will show in Section VI that this problem can be solved with the help of a purification ansatz. We will also discuss in detail how to apply the quantum Bayesian inference for a reconstruction of states of a spin-1/2 when just a finite number of elements of an ensemble have been measured. Before we do this, in the following Section we will analyze the limit of a large number of measurements.

V. BAYESIAN INFERENCE IN LIMIT OF INFINITE NUMBER OF MEASUREMENTS

The explicit evaluation of an *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. (4.4)]. Even for the simplest quantum systems and for a relatively small number of measurements, the reconstruction procedure can present technically insurmountable problems.

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

$$\langle \hat{P}_{\lambda_j, \hat{O}_i} \rangle = \alpha_j^i, \tag{5.1}$$

where $\sum_j \alpha_j^i = 1$. In this case the integral on the right-hand side of Eq. (4.4) can be significantly simplified with the help of the following lemma:

LEMMA. *Let us define the integral expression*

$$I(\alpha_1, ..., \alpha_{n-1}) \equiv \int_0^1 dx_1 \int_0^{y_2} dx_2 \cdots \int_0^{y_{n-1}} dx_{n-1}\, F(x_1, ..., x_{n-1} | \alpha_1, ..., \alpha_{n-1}), \tag{5.2}$$

where

$$F(x_1, ..., x_{n-1} | \alpha_1, ..., \alpha_{n-1}) = \frac{1}{B} x_1^{\alpha_1 N} x_2^{\alpha_2 N} \cdots x_{n-1}^{\alpha_{n-1} N} (1 - x_1 \cdots - x_{n-1})^{\alpha_n N} \tag{5.3}$$

and α_i satisfy condition $\sum_i^n \alpha_i = 1$. The integration boundaries y_k are given by relations:

$$y_k = 1 - \sum_{j=1}^{k-1} x_j; \qquad k = 2, ..., n-1 \tag{5.4}$$

and B equals the product of Beta functions $B(x, y)$:

$$B \equiv B(a_n + 1, a_{n-1} + 1)\, B(a_n + a_{n-1} + 1, a_{n-2} + 2) \cdots B(a_n + a_{n-1} \cdots a_2 + 1, a_1 + n - 1). \tag{5.5}$$

(i) *The function* $F(x_1, ..., x_{n-1} | \alpha_1, ..., \alpha_{n-1})$ *in the integral* (5.2) *is a normalized probability distribution in the* $(n-1)$*-dimensional volume given by integration boundaries.*

(ii) *For* $N \to \infty$, *this probability distribution has the properties*

$$\langle x_i \rangle \to \alpha_i \qquad \langle x_i^2 \rangle \to \alpha_i^2 \qquad i = 1, 2, 3, ..., n-1; \tag{5.6}$$

i.e., this probability density tends to the product of delta functions:

$$\lim_{N \to \infty} F(x_1, ..., x_{n-1} | \alpha_1, ..., \alpha_{n-1}) = \delta(x_1 - \alpha_1)\, \delta(x_2 - \alpha_2) \cdots \delta(x_{n-1} - \alpha_{n-1}). \tag{5.7}$$

Proof. Statement (i) can be derived by the successive application of the equation [see for example [35, Eqs. (3.191)]]

$$\int_0^u x^{\nu-1}(u-x)^{\mu-1}\, dx = u^{\mu+\nu-1} B(\mu, \nu). \tag{5.8}$$

Statement (ii) can be obtained as a result of straightforward calculation of limits of certain expressions containing Beta functions with integer-number arguments. In our calculations we have used the identity

$$\frac{B(n+1, m)}{B(n, m)} = \frac{n}{n+m}, \tag{5.9}$$

which is satisfied by Beta functions with integer-number arguments.

A. *Conditional Density Distribution*

Let us start with the expression for conditional probability distribution $p(\{\ \}_N | \hat{\rho})$ for the N-trial measurement of a set of observables $\hat{O}_i$. If we assume that the number of measurements of each observable $\hat{O}_i$ goes to infinity then we can write

$$p(\{\ \}_{N \to \infty} | \hat{\rho}) = \lim_{N \to \infty} \prod_i \left[\prod_{j=1}^{n_i} \mathrm{Tr}(\hat{P}_{\lambda_j, \hat{O}_i} \hat{\rho})^{\alpha_j^i N} \right]. \tag{5.10}$$

The first product on the right-hand side (r.h.s.) of Eq. (5.10) is associated with each measured observable $\hat{O}_i$ on a given observation level. The second product runs over eigenvalues n_i of each observable $\hat{O}_i$.

In what follows we formally rewrite the r.h.s. of Eq. (5.10): we insert in it a set of functions and we perform the integration

$$p(\{\ \}_{N\to\infty}\,|\,\hat{\rho})=\prod_i\left\{\int_0^1 dx_1^i\int_0^{y_2^i} dx_2^i\cdots\int_0^{y_{n_i-1}^i} dx_{n_i-1}^i\,\delta[x_1^i-\mathrm{Tr}(\hat{P}_{\lambda_1,\hat{O}_i}\hat{\rho})]\cdots\right.$$
$$\left.\times\,\delta[x_{n_i-1}^i-\mathrm{Tr}(P_{\lambda_{n_i-1},\hat{O}_i}\rho)]\prod_{j=1}^{n_i-1}(x_j^i)^{\alpha_j^i N}(1-x_1^i\cdots x_{n_i-1}^i)^{\alpha_{n_i}^i N}\right\} \tag{5.11}$$

In Eq. (5.12) we perform an integration over a volume determined by the integration boundaries y_k^i [see Eq. (5.4)], i.e., due to the condition $\sum_{j=1}^{n_i}\mathrm{Tr}(\hat{P}_{\lambda_j,\hat{O}_i}\hat{\rho})=1$, there is no need to perform integration from $-\infty$ to ∞.

At this point we utilize our Lemma. To be specific, first we separate in Eq. (5.11) the term, which corresponds to the function I given by Eq. (5.2). Then we replace this term by its limit expression (5.7). After a straightforward integration over variables x_j^i we finally obtain an explicit expression for the conditional probability $p(\hat{\rho}\,|\,\{\ \}_{N\to\infty})$ which we insert into Eq. (4.4), from which we obtain the expression for an *a posteriori* estimation of the density operator $\hat{\rho}(\{\ \}_{N\to\infty})$ on the given observation level:

$$\hat{\rho}(\{\ \}_{N\to\infty})=\frac{1}{\mathcal{N}}\int_\Omega\prod_i\left\{\prod_{j=1}^{n_i-1}\delta[\mathrm{Tr}(\hat{P}_{\lambda_j,\hat{O}_i}\hat{\rho})-\alpha_j^i]\right\}\hat{\rho}\,d_\Omega. \tag{5.12}$$

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

The interpretation of Eq. (5.12) is straightforward. The reconstructed density operator is equal to the sum of equally weighted pure-state density operators on the manifold Ω, which satisfy the conditions given by Eq. (5.1) [these conditions are guaranteed by the presence of the functions in the r.h.s. of Eq. (5.12)]. In terms of statistical physics Eq. (5.12) can be interpreted as an averaging over the generalized microcanonical ensemble of those *pure* states which satisfy the conditions on the mean values of the measured observables. Consequently, Eq. (5.12) represents the principle of the "maximum entropy" associated with the generalized microcanonical ensemble which fullfills the constraint (5.1).

VI. BAYESIAN RECONSTRUCTION OF IMPURE STATES

In classical statistical physics a mixed state is interpreted as a statistical average over an ensemble in which any individual realizations is in a pure state. This is also true in quantum physics, but here a mixture can also be interpreted as a state of a quantum system, which cannot be completely described in terms of its own Hilbert space. That is the system under considerationis a nontrivial part of a larger quantum system. When we say nontrivial, we mean that the system under consideration is quantum-mechanically entangled [1] (see also [36]) with the other parts of the composite system. Due to the lack of information about other parts of this complex system, the description of the subsystem is possible only in terms of mixtures.

Let us assume that the quantum system P is entangled with another quantum system$R$ (a reservoir). Let us assume that the composed system S ($S = P \oplus R$) itself is in a pure state $|\Psi\rangle$. The density operator $\hat{\rho}_P$ of the subsystem P is then obtained via tracing over the reservoir degrees of freedom:

$$\hat{\rho}_P = \mathrm{Tr}_R[\hat{\rho}_S]; \qquad \hat{\rho}_S = |\Psi\rangle\langle\Psi|. \tag{6.1}$$

Once the system S is in a pure state, then we can determine an invariant integration measure on the state space of the composite system S and then we can safely apply the Bayesian reconstruction scheme as described in Section IV. 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* estimated density operator $\hat{\rho}_P$ for the system P (with no *a priori* constraint on the purity of the state of the system P). These arguments are intrinsically related to the "purification" ansatz as proposed by Uhlmann [37] (see also [38]).

To make our reconstruction scheme for impure states consistent, we have to chose the reservoir R uniquely. This can be done with the help of the Schmidt theorem (see Ref. [1, 39]) 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

$$|\Psi\rangle = \sum_{i=1}^{M} c_i \, |\alpha_i\rangle_P \otimes |\beta_i\rangle_R, \tag{6.2}$$

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. (6.2) 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 an M-dimensional "reservoir." In this case the dimensionality of the Hilbert space of the composite system is $2M$. Using the standard techniques (see Section VIII) we can then evaluate the invariant integration measure on the manifold of pure states and we can apply the quantum Bayesian inference as discussed above. We stress once again that using the purification procedure we have determined the *invariant* integration measure on the space of *pure* states of the composite system.

We conlude this section with three comments:

(1) First, we note that there also exists another approach to the problem of the integration measure on thespace of *impure* states. Namely, Braunstein and Caves [40] used statistical distinguishability between neighboring quantum states to define the Bures metric [41] on the space of all (pure and mixed) states of the original system S (see also recent work by Slater [42]). The two approaches differ conceptually in understanding what is an impure quantum-mechanical state. That is, in our approach we assume that impurity results as a consequence of the fact

that the system under consideration is entangled with some other system. The other approach accepts the possiblity that an isolated quantum system can be in a statistically mixed state (we will discuss consequences of these two conceptually different approaches elsewhere).

(2) From our previous discussion it follows that for ensembles large enough the two reconstruction schemes provide us with the same results. Consequently, because the Bayesian inference is technically difficult to perform it is useful to utilize the *MaxEnt* principle for the reconstruction in this case. On the other hand, if the ensemble is small, then *MaxEnt* reconstruction scheme cannot be applied and the quantum Bayesian inference has to be utilized.

(3) From the results presented in Sections VI and VII it directly follows that as soon as the number of measurements becomes large then Bayesian inference scheme becomes equal to the reconstruction scheme based on the Jaynes principle of maximum entropy, i.e., in the limit of infinite number of measurements *a posteriori* estimated density operator fulfills the condition of the maximum entropy. Consequently, it is equal to the generalized canonical density operator. If the quorum of observables is measured, then the generalized canonical operator is equal to the "true" density operator of the system itself, i.e., a complete reconstruction via the MaxEnt principle is performed.

VII. RECONSTRUCTION OF SPIN-1/2 STATES VIA THE MaxEnt PRINCIPLE

In this section we present the reconstruction of the state of a single and two spin-1/2 particles. Even though the MaxEnt reconstruction scenario in the case of a single spin-1/2 is very simple we will use this example for a detailed illustration of calculation techniques which will be later used for the reconstruction of two-spin-1/2 states.

A. *Single Spin*-1/2

Let us assume that we want to reconstruct an unknown state of a single spin-1/2. This state is described by a density operator

$$\hat{\rho}(\theta, \phi) = \frac{1}{2}(\hat{1} + \vec{r} \cdot \hat{\vec{\sigma}}) \tag{7.1}$$

where $\vec{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$; $\phi \in (0, 2\pi)$, $\theta \in (0, \pi)$, and $\hat{1}$ is the unit operator. The Pauli spin operators $\hat{\vec{\sigma}}$ in the matrix representation in the basis $|0\rangle$, $|1\rangle$ of the eigenvectors of the operator $\hat{\sigma}_z$ are

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.2}$$

To determine completely the unknown state (7.1) of a single spin-1/2 particle we have to measure three linearly independent (e.g., orthogonal) projections of the spin. One possible choice of the complete set of observables (i.e., the quorum [5]) associated with the spin-1/2 are spin projections for three orthogonal directions represented by the Hermitian operators:

$$\hat{s}_i \equiv \frac{\hat{\sigma}_i}{2}, \qquad i = x,\, y,\, z \tag{7.3}$$

After the measurement of expectation values of each observable, a reconstruction of the generalized canonical density operator (3.5) according to the MaxEnt principle can be performed. In what follows we will consider three observation levels defined as $\mathcal{O}_A^{(1)} = \{\hat{s}_z\}$, $\mathcal{O}_B^{(1)} = \{\hat{s}_z, \hat{s}_x\}$ and $\mathcal{O}_C^{(1)} = \{\hat{s}_z, \hat{s}_x, \hat{s}_y\} \equiv \mathcal{O}_{comp}$ [superscript of the observation levels indicates number of spins-1/2 under consideration].

1. *Observation Level* $\mathcal{O}_A^{(1)} = \{\hat{s}_z\}$

On the observation level $\mathcal{O}_A^{(1)}$ only the mean value of the spin $\hat{s}_z$ is measured. This kind of measurement can be performed with the help of a single Stern–Gerlach apparatus (which fixes the z component of the spin). When the mean value of $\hat{s}_z$ is precisely known then the generalized canonical density operator (i.e., an *a posteriori* estimation of the density operator based on the Jaynes principle) can be written as

$$\hat{\rho} = \frac{1}{Z}\exp(-\lambda_z \hat{\sigma}_z) = \frac{1}{2}[\hat{1} - (\tanh \lambda_z)\,\hat{\sigma}_z], \tag{7.4}$$

where the partition function Z [see Eq. (3.3)] reads

$$Z = \mathrm{Tr}[\exp(-\lambda_z \hat{\sigma}_z)] = 2 \cosh \lambda_z. \tag{7.5}$$

The Lagrange multiplier λ_z is given by the algebraic equation (3.4)

$$\langle \hat{\sigma}_z \rangle = -\frac{1}{Z}\frac{\partial Z}{\partial \lambda_z} = -\tanh \lambda_z. \tag{7.6}$$

When we substitute the solution of the last equation into Eq. (7.4) we find the reconstructed density operator

$$\hat{\rho} = \tfrac{1}{2}(\hat{1} + \langle \hat{\sigma}_z \rangle \hat{\sigma}_z). \tag{7.7}$$

This is the result one would expect intuitively. That is, once we have a precise knowledge of the mean value of the observable $\hat{\sigma}_z$, then the reconstructed state has to be of the form given by Eq. (7.7). From this expression one does not obtain a nontrivial (i.e., nonzero) estimation of unmeasured observables. This is also true when we extend the observation level.

2. *Observation Level* $\mathcal{O}_B^{(1)} = \{\hat{s}_z, \hat{s}_x\}$

To be specific, let us extend the observation level $\mathcal{O}_A^{(1)}$ so that we assume a measurement of two spin projections $\hat{s}_z$ and $\hat{s}_x$. In this case the partition function (3.3) reads

$$Z = \mathrm{Tr}\left[\cosh\lambda\hat{1} - \frac{\sinh\lambda}{\lambda}(\lambda_z\hat{\sigma}_z + \lambda_x\hat{\sigma}_x)\right] = 2\cosh\lambda, \tag{7.8}$$

where $\lambda = (\lambda_z^2 + \lambda_x^2)^{1/2}$. The algebraic equations (3.4) for the Lagrange multipliers λ_z and λ_x can now be written as

$$\langle\hat{\sigma}_z\rangle = -\frac{\tanh\lambda}{\lambda}\lambda_z; \qquad \langle\hat{\sigma}_x\rangle = -\frac{\tanh\lambda}{\lambda}\lambda_x. \tag{7.9}$$

After we find the solutions of these equations, we can write the generalized canonical density operator on $\mathcal{O}_B^{(1)}$ as

$$\hat{\rho} = \tfrac{1}{2}(\hat{1} + \langle\hat{\sigma}_z\rangle\,\hat{\sigma}_z + \langle\hat{\sigma}_x\rangle\,\hat{\sigma}_x), \tag{7.10}$$

which again demonstrates that the reconstructed density operator does not provide us with a nontrivial estimation for the unmeasured observable $\hat{\sigma}_x$. This means that in the case of the single spin-1/2 particle a complete measurement (i.e., the observation level $\mathcal{O}_B^{(1)}$ has to be extended by the inclusion of the operator $\hat{s}_y$) has to be performed to determine a nontrivial mean value of $\langle\hat{s}_y\rangle$.

3. *Observation Level* $\mathcal{O}_C^{(1)} = \{\hat{s}_z, \hat{s}_x, \hat{s}_y\}$

Using the algebraic properties of the $\hat{\sigma}_v$ operators, the generalized canonical density operator on the observation level $\mathcal{O}_C^{(1)}$ can be expressed as

$$\hat{\rho} = \frac{1}{Z}\exp(-\vec{\lambda}.\hat{\vec{\sigma}}) = \frac{1}{Z}\left[\cosh|\lambda|\,\hat{1} - \sinh|\lambda|\,\frac{\vec{\lambda}.\hat{\vec{\sigma}}}{|\lambda|}\right], \tag{7.11}$$

where the partition function reads $Z = 2\cosh|\lambda|$, while $\vec{\lambda} = (\lambda_x, \lambda_y, \lambda_x)$ and $|\lambda|^2 = \lambda_x^2 + \lambda_y^2 + \lambda_z^2$. The algebraic equations for the Lagrange multipliers can be found straightforwardly and we find for the reconstructed density operator the expression

$$\hat{\rho} = \tfrac{1}{2}[\hat{1} + \langle\hat{\sigma}_z\rangle\,\hat{\sigma}_z + \langle\hat{\sigma}_x\rangle\,\hat{\sigma}_x + \langle\hat{\sigma}_y\rangle\,\hat{\sigma}_y]. \tag{7.12}$$

To complete this example we present an expression for the von Neumann entropy of the reconstructed density operators. This entropy can be written as

$$S_{\mathcal{O}} = -p_{\mathcal{O}}\ln p_{\mathcal{O}} - (1 - p_{\mathcal{O}})\ln(1 - p_{\mathcal{O}}), \tag{7.13}$$

where $p_{\mathcal{O}}$ is one eigenvalue of $\hat{\rho}_{\mathcal{O}}$ [the other eigenvalue is equal to $(1 - p_{\mathcal{O}})$] which reads as

$$p_A = \frac{1 + |\langle \hat{\sigma}_z \rangle|}{2}, \qquad p_B = \frac{1 + \sqrt{\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_z \rangle^2}}{2},$$
$$p_{\text{comp}} = \frac{1 + \sqrt{\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_y \rangle^2 + \langle \hat{\sigma}_z \rangle^2}}{2}. \tag{7.14}$$

We stress here that the reconstruction scheme based on the *MaxEnt* principle is perfectly well suitable for a reconstruction of both pure states and statistical mixtures; i.e., the entropy on the complete observation level can be larger then zero. That is, the measured mean values of spin observables may be such that

$$\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_y \rangle^2 + \langle \hat{\sigma}_z \rangle^2 < 1, \tag{7.15}$$

which is in a striking contrast with the Bayesian quantum inference scheme [32] in which it is *a priori* assumed that the reconstructed state has to be a pure one, which means that on the complete observation level the condition

$$\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_y \rangle^2 + \langle \hat{\sigma}_z \rangle^2 = 1 \tag{7.16}$$

has to be fulfilled, as otherwise the reconstruction scheme fails.

B. *Two Spins*-1/2

Let us now consider a system composed of two *distinguishable* spins-1/2. In general, any density operator of the system composed of two distinguishable spins-1/2 can be represented by a 4×4 Hermitian matrix and 15 independent numbers are required for its complete determination. The quorum of observables for a system of two spins-1/2 is given by 15 operators:

$$\hat{s}_i^{(1)} = \frac{\hat{\sigma}_i \otimes \hat{1}}{2}; \qquad \hat{s}_i^{(2)} = \frac{\hat{1} \otimes \hat{\sigma}_i}{2}; \qquad \hat{s}_i^{(1)} \hat{s}_j^{(2)} = \frac{\hat{\sigma}_i \otimes \hat{\sigma}_j}{4} \qquad i, j = x, y, z. \tag{7.17}$$

These operators together with the identity operator $\hat{1} \otimes \hat{1}$ form the basis of operator algebra in which any operator associated with the system under consideration can be expressed. In Eq. (7.17) we use the notation such that the operators which stand left (right) to the symbol $\otimes$ are associated with the first (second) spin-1/2.

Using the maximum-entropy principle we can (partially) reconstruct an unknown density operator $\hat{\rho}$ on various observation levels associated with observables given by Eq. (7.17). In what follows we will consider three observation levels.

1. *Observation Level* $\mathcal{O}_A^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z^{(2)}\}$

Let us assume a reconstruction of a quantum state of the two spins-1/2 system when the observation level is given by two observables [see Eq. (7.17)] $\hat{s}_z^{(1)}$ and $\hat{s}_z^{(2)}$

related to the first and to the second spin, respectively. Due to the fact that observables associated with the first spin commute with the observables of the second spin we can express the partition function (3.3) in a factorized form of two exponentials associated with each spin separately. Therefore the partition sum reads

$$Z = \mathrm{Tr}[\exp(-\lambda_1 \hat{\sigma}_z \otimes \hat{1} - \lambda_2 \hat{1} \otimes \hat{\sigma}_z)] = 4 \cosh \lambda_1 \cosh \lambda_2. \tag{7.18}$$

The set of algebraic equations (3.4) now separates, and the resulting density operator can be expressed as a product of two density operators

$$\begin{aligned}\hat{\rho} &= \frac{1}{Z}[(\cosh \lambda_1 \hat{1} - \sinh \lambda_1 \hat{\sigma}_z) \otimes \hat{1}][(\hat{1} \otimes (\cosh \lambda_2 \hat{1} - \sinh \lambda_2 \hat{\sigma}_z)] \\ &= \frac{1}{4}[\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z \otimes \hat{\sigma}_z] \\ &= \frac{1}{2}[\hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z] \otimes \frac{1}{2}[\hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z].\end{aligned} \tag{7.19}$$

Here we use the notation such that $\langle \hat{\sigma}_z^{(1)} \rangle$, $\langle \hat{\sigma}_z^{(2)} \rangle$, and $\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle$ describe the mean values of operators $\langle \hat{\sigma}_z \otimes \hat{1} \rangle$, $\langle \hat{1} \otimes \hat{\sigma}_z \rangle$, and $\langle \hat{\sigma}_z \otimes \hat{\sigma}_z \rangle$, respectively.

2. *Observation Level* $\mathcal{O}_B^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z \otimes \hat{s}_z\}$

Let us assume an observation level given by operators $\hat{s}_z^{(1)}$ and $\hat{s}_z^{(1)} \hat{s}_z^{(2)}$. These operators also commute and therefore the partition function Z can be expressed as

$$\begin{aligned}Z &= \mathrm{Tr}[([\cosh \lambda_1 \hat{1} - \sinh \lambda_1 \hat{\sigma}_z] \otimes \hat{1})(\cosh \lambda_{12} \hat{1} \otimes \hat{1} - \sinh \lambda_{12} \hat{\sigma}_z \otimes \hat{\sigma}_z)] \\ &= 4 \cosh \lambda_1 \cosh \lambda_{12}\end{aligned} \tag{7.20}$$

The corresponding Lagrange multipliers λ_1, λ_{12} can be found straightforwardly and we obtain the reconstructed density matrix

$$\hat{\rho} = \tfrac{1}{4}[\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z \otimes \hat{\sigma}_z]. \tag{7.21}$$

We point out that in this particular example the Jaynes principle of maximum entropy provides us with a nontrivial estimation for an unmeasured mean value of the observable $\hat{\sigma}_z^{(2)}$. The *a posteriori* estimation of $\langle \hat{\sigma}_z^{(2)} \rangle$ is equal to $\langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle$.

3. *Observation Level* $\mathcal{O}_C^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z^{(2)}, \hat{s}_z \otimes \hat{s}_z\}$

Finally we assume the observation level when the spin projections $\hat{s}_z^{(1)}$, $\hat{s}_z^{(2)}$ as well as the correlation $\hat{s}_z^{(1)} \hat{s}_z^{(2)}$ between them is measured. All these operators mutually commute and so the partition function can be expressed as a product of three independent terms

$$Z = \mathrm{Tr}[([\cosh\lambda_1\hat{1} - \sinh\lambda_1\hat{\sigma}_z]\otimes\hat{1})(\hat{1}\otimes[\cosh\lambda_2\hat{1} - \sinh\lambda_2\hat{\sigma}_z]) \times(\cosh\lambda_{12}\hat{1}\otimes\hat{1} - \sinh\lambda_{12}\hat{\sigma}_z\otimes\hat{\sigma}_z)], \tag{7.22}$$

from which we find

$$Z = 4[\cosh\lambda_1\cosh\lambda_2\cosh\lambda_{12} - \sinh\lambda_1\sinh\lambda_2\sinh\lambda_{12}]. \tag{7.23}$$

The Lagrange multipliers have to be found from the set of algebraic equations

$$\begin{aligned}\langle\hat{\sigma}_z^{(1)}\rangle &= 4[\cosh\lambda_1\sinh\lambda_2\sinh\lambda_{12} - \sinh\lambda_1\cosh\lambda_2\cosh\lambda_{12}]/Z\\ \langle\hat{\sigma}_z^{(2)}\rangle &= 4[\sinh\lambda_1\cosh\lambda_2\sinh\lambda_{12} - \cosh\lambda_1\sinh\lambda_2\cosh\lambda_{12}]/Z\\ \langle\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\rangle &= 4[\sinh\lambda_1\sinh\lambda_2\cosh\lambda_{12} - \cosh\lambda_1\cosh\lambda_2\sinh\lambda_{12}]/Z.\end{aligned} \tag{7.24}$$

Having solved these equations the reconstructed density operator reads

$$\hat{\rho} = \tfrac{1}{4}[\hat{1}\otimes\hat{1} + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{\sigma}_z\otimes\hat{1} + \langle\hat{\sigma}_z^{(2)}\rangle\,\hat{1}\otimes\hat{\sigma}_z + \langle\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\rangle\,\hat{\sigma}_z\otimes\hat{\sigma}_z]. \tag{7.25}$$

For further discussion on the reconstruction of spin-states via MaxEnt principle see Ref. [43].

VIII. SPIN-1/2 RECONSTRUCTION VIA BAYESIAN INFERENCE I

In this section we study Bayesian reconstruction of spin-1/2 states on various observation levels. That is, we investigate how the best *a posteriori* estimation of the density operator of the spin-1/2 system based on an incomplete set of data (in this case the exact mean values of the spin observables are not available) can be obtained. We have already stressed the fact that the Bayesian inference scheme as introduced by Jones [32] is suitable only for pure states. This means that the completely reconstructed density operator has to fulfill the purity condition (7.16).

We start our example with a definition of the parametric state space associated with the spin-1/2. The rigorous way to determine this parametric state space Ω is based on the diffeomorphism between Ω and the quotient space $^{SU(n)}|_{U(n-1)}$, where n is the dimensionality 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 on to the Poincaré sphere and the parameterized density operator (i.e., the point on the Poincaré sphere) is given by Eq. (7.1). The topology of the Poincaré sphere determines also the integration measure for which we have $d_\Omega = \sin\theta\, d\theta\, d\phi$ (for more details see Section IX).

The observables associated with the spin-1/2 are spin projections for three orthogonal directions represented by Hermitian operators $\hat{s}_j$ given by Eq. (7.3).

These observables have spectra equal to $\pm\frac{1}{2}$. In what follows we distinguish between these two possible measurement results by the sign, i.e., $s=\pm 1$. The projectors $\hat{P}_{s,\hat{s}i}$ on to the corresponding eigenvectors are

$$\hat{P}_{s,\hat{s}_i}=\frac{\hat{1}+s\hat{\sigma}_i}{2}; \qquad i=x,y,z \tag{8.1}$$

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

$$p(s,\hat{s}_i\,|\,\hat{\rho}(\theta,\phi))=\frac{1+sr_i}{2}; \qquad i=x,y,z. \tag{8.2}$$

Now using the procedure described in Section IV, we can construct an *a posteriori* estimation of the density operator $\hat{\rho}(\{\ \}_N)$ based on a given sequence of measurement outcomes on different observation levels.

A. *Estimation Based on Results of Fictitious Measurements*

In Table I we present results of an *a posteriori* estimation of density operators based on data obtained from "experiments" performed with three Stern–Gerlach devices oriented along the axes x, y, and z. We first discuss in detail reconstruction of a single spin-1/2 state under the *a priori* assumption that the system is in a pure state.

1. *Observation Level* $\mathcal{O}_A^{(1)}=\{\hat{s}_z\}$

The first five lines in Table I describe results of a fictitious measurement of the spin component $\hat{s}_z$ and the corresponding estimated density operators. In particular, let us assume that just one detection event (spin "up", i.e., $\uparrow$) is registered in the given Stern–Gerlach apparatus (associated with the measurement of $\hat{s}_z$). Taking into account the parameterization of the single spin-1/2 density operator expressed by Eq. (7.1) we find for the corresponding conditional probability distribution $p(s,\hat{s}_i\mid\hat{\rho}(\theta,\phi))$ (8.2) the expression

$$p(s,\hat{s}_i\,|\,\hat{\rho}(\theta,\phi))=\frac{1+\cos\theta}{2}. \tag{8.3}$$

Using Eq. (4.4) we can express the estimated density operator based on the registration of just one result (spin "up") as

$$\begin{aligned}\hat{\rho}&=\frac{1}{8\pi}\int_0^{\pi}\sin\theta\,d\theta\int_0^{2\pi}d\phi(1+\cos\theta)(\hat{1}+\sin\theta\cos\phi\hat{\sigma}_x+\sin\theta\sin\phi\hat{\sigma}_y+\cos\theta\sigma_z)\\&=\frac{1}{2}\left(\hat{1}+\frac{1}{3}\hat{\sigma}_z\right).\end{aligned} \tag{8.4}$$

TABLE I

Results of *a posteriori* Bayesian Estimation of Density Operators of the Spin-1/2

	$\hat{\sigma}_z$	$\hat{\sigma}_x$	$\hat{\sigma}_y$	$\hat{\rho}$ via pure-state reconstruction	S	$\hat{\rho}$ via mixture-state reconstruction	S
1.	$\uparrow$			$\frac{1}{2}[\hat{1} + \frac{1}{3}\hat{\sigma}_z]$	0.637	$\frac{1}{2}[\hat{1} + \frac{1}{5}\hat{\sigma}_z]$	0.673
2.	$\uparrow^4$			$\frac{1}{2}[\hat{1} + \frac{2}{3}\hat{\sigma}_z]$	0.451	$\frac{1}{2}[\hat{1} + \frac{1}{2}\hat{\sigma}_z]$	0.562
3.	$\uparrow^5\downarrow$			$\frac{1}{2}[\hat{1} + \frac{1}{2}\hat{\sigma}_z]$	0.562	$\frac{1}{2}[\hat{1} + \frac{2}{5}\hat{\sigma}_z]$	0.611
4.	$\uparrow^{10}\downarrow^2$			$\frac{1}{2}[\hat{1} + \frac{4}{7}\hat{\sigma}_z]$	0.520	$\frac{1}{2}[\hat{1} + \frac{1}{2}\hat{\sigma}_z]$	0.562
5.	$\uparrow^{15}\downarrow^3$			$\frac{1}{2}[\hat{1} + \frac{3}{5}\hat{\sigma}_z]$	0.501	$\frac{1}{2}[\hat{1} + \frac{6}{11}\hat{\sigma}_z]$	0.536
6.	$\uparrow$	$\downarrow$		$\frac{1}{2}[\hat{1} - \frac{1}{3}\hat{\sigma}_x + \frac{1}{3}\hat{\sigma}_z]$	0.578	$\frac{1}{2}[\hat{1} - \frac{1}{5}\hat{\sigma}_x + \frac{1}{5}\hat{\sigma}_z]$	0.653
7.	$\uparrow^4$	$\uparrow^3\downarrow$		$\frac{1}{2}[\hat{1} + \frac{10}{37}\hat{\sigma}_x + \frac{26}{37}\hat{\sigma}_z]$	0.374	$\frac{1}{2}[\hat{1} + \frac{68}{309}\hat{\sigma}_x + \frac{158}{309}\hat{\sigma}_z]$	0.529
8.	$\uparrow^5\downarrow$	$\uparrow^4\downarrow^2$		$\frac{1}{2}[\hat{1} + \frac{704}{2601}\hat{\sigma}_x + \frac{1460}{2601}\hat{\sigma}_z]$	0.484	$\frac{1}{2}[\hat{1} + \frac{218}{1105}\hat{\sigma}_x + \frac{464}{1105}\hat{\sigma}_z]$	0.581
9.	$\uparrow^{10}\downarrow^2$	$\uparrow^8\downarrow^4$		$\frac{1}{2}[\hat{1} + \frac{1599844}{5073971}\hat{\sigma}_x + \frac{3143928}{5073971}\hat{\sigma}_z]$	0.427	$\frac{1}{2}[\hat{1} + \frac{513984}{2093401}\hat{\sigma}_x + \frac{1083360}{2093401}\hat{\sigma}_z]$	0.519
10.	$\uparrow$	$\downarrow$	$\uparrow$	$\frac{1}{2}[\hat{1} - \frac{1}{3}\hat{\sigma}_x + \frac{1}{3}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.518	$\frac{1}{2}[\hat{1} - \frac{1}{5}\hat{\sigma}_x + \frac{1}{5}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.632
11.	$\uparrow^4$	$\uparrow^3\downarrow$	$\uparrow^4$	$\frac{1}{2}[\hat{1} + \frac{831}{3503}\hat{\sigma}_x + \frac{2026}{3503}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.264	$\frac{1}{2}[\hat{1} + \frac{1051}{5253}\hat{\sigma}_x + \frac{2382}{5253}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.446
12.	$\uparrow^5\downarrow$	$\uparrow^4\downarrow^2$	$\uparrow^5\downarrow$	$\frac{1}{2}[\hat{1}+ \frac{47109}{169636}\hat{\sigma}_x+ \frac{99310}{169636}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.236	$\frac{1}{2}[\hat{1}+ \frac{279193}{1446325}\hat{\sigma}_x+ \frac{593708}{1446325}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.492
13.	$\uparrow^{10}\downarrow^2$	$\uparrow^8\downarrow^4$	$\uparrow^{10}\downarrow^2$	$\frac{1}{2}[\hat{1}+ \frac{1222748838}{4026213681}\hat{\sigma}_x + \frac{2532792812}{4026213682}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.135	$\frac{1}{2}[\hat{1}+ \frac{250224710127}{1073523481830}\hat{\sigma}_x + \frac{531888078934}{1073523481830}(\hat{\sigma}_y + \hat{\sigma}_z)]$	0.388
14.	$\uparrow^3\downarrow$	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\frac{1}{2}[\hat{1} + \frac{101}{161}\hat{\sigma}_z]$	0.481	$\frac{1}{2}[\hat{1} + \frac{413}{1389}\hat{\sigma}_z]$	0.648
15.	$\uparrow^6\downarrow^2$	$\uparrow^4\downarrow^4$	$\uparrow^4\downarrow^4$	$\frac{1}{2}[\hat{1} + \frac{88}{117}\hat{\sigma}_z]$	0.374	$\frac{1}{2}[\hat{1} + \frac{3125918}{8023325}\hat{\sigma}_z]$	0.615
16.	$\uparrow^9\downarrow^3$	$\uparrow^6\downarrow^6$	$\uparrow^6\downarrow^6$	$\frac{1}{2}[\hat{1} + \frac{10642815}{13619371}\hat{\sigma}_z]$	0.345	$\frac{1}{2}[\hat{1} + \frac{57056845292}{134078568484}\hat{\sigma}_z]$	0.600
17.	$\uparrow^{12}\downarrow^4$	$\uparrow^8\downarrow^8$	$\uparrow^8\downarrow^8$	$\frac{1}{2}[\hat{1} + \frac{10875098376}{13696058161}\hat{\sigma}_z]$	0.332	$\frac{1}{2}[\hat{1} + \frac{3073000318516432}{6928263111521097}\hat{\sigma}_z]$	0.591

Note. Results 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. We also present values of von Neumann entropy [see Eq. (7.13)] associated with the given estimated density operator. In the case of a reconstruction of pure states, the value of the von Neumann entropy reflects the fidelity of the estimation.

We stress that we started our estimation procedure with an a priori assumption that the measured system is in a *pure* state, for which the von Neumann entropy S (3.2) has to be equal to zero. But the estimated density operator (8.4) describes a statistical mixture with the von Neumann entropy $S \simeq 0.637$ (see Table I). There is no contradiction here. In the reconstruction of pure states, a nonzero value of the von Neumann entropy of the estimated density operator reflects the fidelity with which the reconstruction is performed. That is, before any measurement is performed, the "estimated" density operator is $\hat{\rho} = \hat{1}/2$, for which the von Neumann entropy takes the maximal value $S = \ln 2 \simeq 0.693$. As soon as the first measurement is performed, some information about the state of the system is acquired, which is reflected by the decrease of the entropy and a better estimation of the density operator. The estimated density operator is expressed as a statistical mixture because it is equal to a specifically weighted sum of a set of *pure* states [see the reconstruction formula (4.4)] which also reflects our incomplete knowledge about the state of the measured system. Obviously, the more measurements we perform, the better the estimation can be performed (compare lines 2–5 in Table I). Nevertheless, we have to stress that the von Neumann entropy is not a monotonically decreasing function of a number of measurements. To be specific, in the case when just a small number of measurements is performed, the estimation is very sensitive with respect to the outcome of any additional measurement.

Comparing the lines 2 and 3 in Table I, we see that the entropy "locally" increases in spite of the fact that more measurements are performed. Nevertheless, in the limit of large number of measurements, the entropy approaches its minimum possible value associated with a given measurement. Providing the quorum of observables is measured, the entropy tends to zero and the state is completely reconstructed.

In general, increasing the number of measurements improves the *a posteriori* estimation of the density operator on the given observation level (see lines 2–5 in Table I). Using the general results of Section V we can evaluate the *a posteriori* estimation of the density operator of the spin-1/2 system on the observation level $\mathcal{O}_A^{(1)}$ in the limit of infinite number of measurements of the spin component $\hat{s}_z$. We note that in this case, when the observable has only two eigenvalues, the information obtained in the spectral distribution (5.1) is equivalently given only by the mean value of this observable. Once we know the spectral distribution Eq. (5.1) corresponding to the measurement of the spin projection $\hat{s}_z$ of single spin-1/2, then with the help of Eq. (5.12) we can express the reconstructed density operator as

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta \, \delta(\langle \hat{\sigma}_z \rangle - \cos\theta) \times (\hat{1} + \sin\theta \cos\phi \, \hat{\sigma}_x + \sin\theta \sin\phi \, \hat{\sigma}_y + \cos\theta \, \hat{\sigma}_z), \tag{8.5}$$

where $\mathcal{N}$ is the normalization constant such that $\mathrm{Tr}\,\hat{\rho} = 1$. Integration over the variable ϕ in Eq. (8.5) cancels all terms in front of the operators $\hat{\sigma}_x$ and $\hat{\sigma}_y$ and we obtain

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{\pi} \sin\theta \, d\theta \, \delta(\langle \hat{\sigma}_z \rangle - \cos\theta)(\hat{1} + \cos\theta \, \hat{\sigma}_z). \tag{8.6}$$

The right hand side of this equation suggests a simple geometrical interpretation of the quantum Bayesian inference in the limit of infinite number of measurements. Namely, the density operator (8.6) can be understood as an equally weighted average of all *pure* states with the same (i.e., measured) mean value of the operator $\hat{s}_z$. These states are represented as points on a circle on the Poincaré sphere. When we perform the integration over θ in Eq. (8.6) we obtain the final expression

$$\hat{\rho} = \tfrac{1}{2}(\hat{1} + \langle \hat{\sigma}_z \rangle \hat{\sigma}_z). \tag{8.7}$$

for the density operator on the given observation level. Formally this is the same density operator as that reconstructed with the help of the Jaynes principle [see Eq. (7.7)]. But there is a difference: the formula (8.7) is obtained as a result of averaging of the generalized *microcanonical* ensemble of *pure* states, while the reconstruction via the MaxEnt principle is based on an averaging over the generalized *grand canonical* ensemble of all states. The two reconstruction schemes differ by the *a priori* assumptions about the possible states of the measured system. As we will see later, these different assumptions result in different estimations (see below).

2. *Observation Level* $\mathcal{O}_B^{(1)} = \{\hat{s}_z, \hat{s}_x\}$

The results of a numerical reconstruction of the density operator of the spin-1/2 based on the measurement of two spin components $\hat{s}_z$ and $\hat{s}_x$ are presented in Table I (lines 6–9). Lines 1–4 and 6–9 describe estimations based on the same data for the $\hat{s}_z$ measurement, but they differ in the data for the $\hat{s}_x$ measurement. That is, lines 1–4 describe the situation for which no results for $\hat{s}_x$ are available, while lines 6–9 describe the situation with specific outcomes for the $\hat{s}_x$ measurements. Comparing these two cases (i.e., if we compare the values of the von Neumann entropy for pairs of lines $\{x, x+5\}$; $x=1, 2, 3, 4$) we see that any measurement performedon the additional observable ($\hat{s}_x$) can only improve our estimation based on the measurement of the original observable ($\hat{s}_z$).

In the limit of infinite number of measurements, when we have information about the spectral distribution corresponding to measurement of spin projections $\hat{s}_x, \hat{s}_z$ the particular form of Eq. (5.12) reads

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta \, \delta(\langle \hat{\sigma}_z \rangle - \cos\theta) \, \delta(\langle \hat{\sigma}_x \rangle - \sin\theta \cos\phi) \times (\hat{1} + \sin\theta \cos\phi \hat{\sigma}_x + \sin\theta \sin\phi \, \hat{\sigma}_y + \cos\theta \, \hat{\sigma}_z). \tag{8.8}$$

As seen from the right-hand side of Eq. (8.8) in this case the reconstructed density operator is represented by an equally weighted sum of points given by an intersection of two circles lying on the Poincaré sphere. These two circles are specified by the two equations $\langle \hat{\sigma}_z \rangle = \cos\theta$ and $\langle \hat{\sigma}_x \rangle = \sin\theta \cos\phi$.

With the help of the identity

$$\delta(f(x)) = \sum_{x_0, \, f(x_0)=0} \frac{\delta(x - x_0)}{|f'(x_0)|}, \tag{8.9}$$

we can perform the integration over ϕ in Eq. (8.8) and obtain

$$\rho = \frac{1}{\mathcal{N}} \int_{\mathcal{L}} d\theta \sum_{\phi_0} \frac{\sin\theta}{|\sin\theta \sin\phi_0|} \delta(\langle \hat{\sigma}_z \rangle - \cos\theta) \times (\hat{1} + \langle \hat{\sigma}_x \rangle \, \hat{\sigma}_x + \sin\theta \sin\phi_0 \, \hat{\sigma}_y + \cos\theta \, \hat{\sigma}_z). \tag{8.10}$$

The integration boundaries $\mathcal{L}$ on the right-hand side of Eq. (8.10) are defined as

$$\mathcal{L} := 0 \leqslant \theta \leqslant \pi \qquad \text{and} \qquad |\sin\theta| \geqslant |\langle \hat{\sigma}_x \rangle|. \tag{8.11}$$

The sum on the right-hand side of Eq. (8.10) refers to two values of the parameter ϕ_0 which fulfill the condition $\cos\phi_0 = \langle \hat{\sigma}_x \rangle / \sin\theta$. We note that the function in front of the operator $\hat{\sigma}_y$ disappears due to the fact that it is proportional to $\sin\phi_0 / |\sin\phi_0|$, which is an odd function of ϕ_0. After we perform the integration over θ we obtain

$$\hat{\rho} = \tfrac{1}{2}(\hat{1} + \langle \hat{\sigma}_x \rangle \, \hat{\sigma}_x + \langle \hat{\sigma}_z \rangle \hat{\sigma}_z). \tag{8.12}$$

What we see again is that in the limit of a large number of measurements the Bayesian inference formally gives us the same result as the Jaynes principle of maximum entropy [compare with Eq. (7.10)].

3. *Observation Level* $\mathcal{O}_C^{(1)} = \{\hat{s}_z, \hat{s}_x, \hat{s}_y\}$

Further extension of the observation level $\mathcal{O}_B^{(1)}$ leads us to the complete observation level, when all three spin components $\hat{s}_x$, $\hat{s}_y$ and $\hat{s}_z$ of the spin-1/2 are measured. Results of the numerical reconstruction are presented in Table I (lines 10–13). Now we compare the *a posteriori* estimation of density operators based on data presented in lines 6–9. The "experimental data" in line 10 are equal to those presented in line 6 except that now some additional knowledge concerning the spin component $\hat{s}_y$ is available. We note that this additional information about $\hat{s}_y$ improves our estimation of the density operator which is clearly seen when we compare values of the von Neumann entropy presented in Table I.

Providing that we have information concerning the spectral distribution associated with the measurement of a complete set (i.e., the quorum) of operators $\hat{s}_x, \hat{s}_y, \hat{s}_z$ (i.e., after an infinite number of measurements of the three spin components have been performed), then we can express the estimated density operator as [see Eq. (5.12)]

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta \, \delta(\langle \hat{\sigma}_z \rangle - \cos\theta) \, \delta(\langle \hat{\sigma}_x \rangle - \sin\theta \cos\phi) \, \delta(\langle \hat{\sigma}_y \rangle - \sin\theta \sin\phi)$$
$$\times (\hat{1} + \sin\theta \cos\phi \hat{\sigma}_x + \sin\theta \sin\phi \hat{\sigma}_y + \cos\theta \hat{\sigma}_z). \tag{8.13}$$

The integral on the right-hand side of Eq. (8.13) can only be performed if the purity condition (7.16) is fulfilled, otherwise it simply does not exist. When the purity condition is fulfilled then from Eq. (8.13) we obtain

$$\hat{\rho} = \tfrac{1}{2}(\hat{1} + \langle \hat{\sigma}_x \rangle \, \hat{\sigma}_x + \langle \hat{\sigma}_y \rangle \, \hat{\sigma}_y + \langle \hat{\sigma}_z \rangle \, \hat{\sigma}_z). \tag{8.14}$$

Here we can again utilize a simple geometrical interpretation of the limit formula (8.13) for the Bayes inference. The three functions in Eq. (8.13) correspond to three specific orbits (circles) on the Poincaré sphere each of which is associated with a set of pure states which posses the measured value of a given observable $\hat{s}_i$. The reconstructed density operator then describes a point on the Poincaré sphere which coincides with an intersection of these three orbits. Consequently, if the three orbits have no intersection the reconstruction scheme fails, because there does not exist a *pure* state with the given mean values of the measured observables.

We illustrate this failure of the Bayesian inference scheme in lines 14–17 of Table I. Here we present a numerical simulation of the measurement in which all three observables are measured. It is assumed that the spin-1/2 is in the state with $\langle \hat{\sigma}_z \rangle = 1/2$ and $\langle \hat{\sigma}_x \rangle = \langle \hat{\sigma}_y \rangle = 0$, which apparently does not fulfill the purity condition (7.16). For a given set of measurement outcomes (line 14) the Bayesian inference scheme provides us with an *a posteriori* estimation such that $\langle \hat{\sigma}_z \rangle = 101/161$ which is above the expected mean value which is equal to 1/2. Moreover if we increase the number of measurements (lines 15–17) the a posteriori estimation deviates more and more from what would be a correct estimation (i.e., results presented in lines 14–17 correspond to the following sequence of mean values of $\hat{\sigma}_z$: 0.481; 0.375; 0.345; 0.332) but simultaneously the von Neumann entropy S decreases, which should indicate that our estimation is better and better. This clearly illustrates the intrinsic conflict in the estimation procedure.

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 (7.16) and so the Bayesian method *cannot* be applied safely in the present case. The larger the number of measurement the more clearly the inconsistency is seen and, as follows from Eq. (8.13), 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 generalized grand canonical ensemble, which means that all states (pure and impure) are taken into account.

In the present example the discrepancy between the *a posteriori* estimations of density operators based on the two different schemes has appeared only on the complete observation level. For more complex quantum-mechanical systems the difference between the density operator reconstructed with the help of the Jaynes principle of maximum entropy and the density operator obtained via the Bayesian inference scheme may differ even on incomplete observation levels. To see this we present in the following sections an example of reconstruction of density operators describing states of two spins-1/2.

IX. QUANTUM BAYESIAN INFERENCE OF TWO-SPINS-1/2 STATES

In order to apply the general formalism of quantum Bayesian inference as described in Section IV we have to properly parameterize the state space of the quantum system under consideration. Once this is done we have to find the invariant integration measure d_Ω associated with the state space and only then can we effectively use the reconstruction formula (4.3). We start this section with a description of how the state space of two spins-1/2 can be parameterized. We show how the corresponding integration measure can then be found.

A. *Parameterization of Two-Spins-1/2 State Space*

One way to determine the state space Ω of a given quantum-mechanical system is via a diffeomorphism $\Omega \equiv {}^{SU(n)}|_{U(n-1)}$. This directly provides us with information about the dimensionality of Ω, which is $(dim_{SU(n)} - dim_{U(n-1)}) = 2n - 2$. This means that in our case of two spins-1/2 which are prepared in a *pure* state we need 6 coordinates which parameterize Ω $(n = 4)$. Unfortunately, it is not very convenient to determine the state space via the given diffeomorphism because then we have to work with noncommutative groups.

It is much simpler to parameterize the state space Ω utilizing the idea of the Schmidt decomposition [1]. In this case we can represent any pure state $|\Psi\rangle$ describing two spins-1/2 as

$$|\Psi\rangle = A\,|\uparrow_1\rangle \otimes |\uparrow_2\rangle + B\,|\downarrow_1\rangle \otimes |\downarrow_2\rangle, \tag{9.1}$$

where $|\downarrow_j\rangle$, $|\uparrow_j\rangle$, are two general 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 then reads

$$\hat{\rho} = |A|^2\,|\uparrow_1\rangle\langle\uparrow_1| \otimes |\uparrow_2\rangle\langle\uparrow_2| + AB^*\,|\uparrow_1\rangle\langle\downarrow_1| \otimes |\uparrow_2\rangle\langle\downarrow_2|$$

$$+ A^*B\,|\downarrow_1\rangle\langle\uparrow_1| \otimes |\downarrow_2\rangle\langle\uparrow_2| + |B|^2\,|\downarrow_1\rangle\langle\downarrow_1| \otimes |\downarrow_2\rangle\langle\downarrow_2|. \tag{9.2}$$

The projectors $|\uparrow_j\rangle\langle\uparrow_j|$ and $|\downarrow_j\rangle\langle\downarrow_j|$ $(j = 1, 2)$ are given by $(\hat{1} + \vec{r}^{(j)}\hat{\vec{\sigma}}\,|\,A^{(j)})$ and $(\hat{1} - \vec{r}^{(j)}\hat{\vec{\sigma}}\,|\,A^{(j)})$, respectively [see Eq. (8.1)], where $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ are two arbitrary unity vectors. The operators $|\downarrow_j\rangle\langle\uparrow_j|$ and their Hermitian conjugates $|\uparrow_j\rangle\langle\downarrow_j|$ are determined as

$$|\downarrow_j\rangle\langle\uparrow_j|\,(\hat{1} + \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)})\,|\uparrow_j\rangle\langle\downarrow_j| = (\hat{1} - \vec{r}^{(j)}\hat{\vec{\sigma}}^{(j)}), \tag{9.3}$$

from which the relation

$$|\uparrow_j\rangle\langle\downarrow_j| = e^{i\psi_j}(\vec{k}^{(j)}\hat{\vec{\sigma}}^{(j)} + i\vec{l}^{(j)}\hat{\vec{\sigma}}^{(j)}) \tag{9.4}$$

follows. 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 $\vec{l}^{(j)} = \vec{r}^{(j)} \times \vec{k}^{(j)}$. A particular choice of vectors $\vec{k}_j$ is not important because phase factors $e^{i\psi_j}$ [$\psi_j \in (0, 2\pi)$] rotate them along all possible directions. We also note that the phase factors $e^{i\psi_j}$ can be always incorporated in the phase ψ of a complex number AB^*. Using the parameterization $|A| = \cos(\alpha/2)$ and $|B| = \sin(\alpha/2)$ we can parameterize $\hat{\rho}$ as

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

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

$$- \sin\alpha\sin\psi \left[\frac{\vec{k}^{(1)}\hat{\vec{\sigma}} \otimes \vec{l}^{(2)}\hat{\vec{\sigma}}}{4} + \frac{\vec{l}^{(1)}\hat{\vec{\sigma}} \otimes \vec{k}^{(2)}\hat{\vec{\sigma}}}{4}\right], \tag{9.5}$$

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

$$\begin{aligned}\vec{k}^{(j)} &= (\sin\phi_j, \ -\cos\phi_j, 0);\\ \vec{l}^{(j)} &= (\cos\theta_j \cos\phi_j, \cos\theta_j \sin\phi_j, \ -\sin\theta_j);\\ \vec{r}^{(j)} &= (\sin\theta_j \cos\phi_j, \sin\theta_j \sin\phi_j, \cos\theta_j).\end{aligned} \tag{9.6}$$

Once we have parameterized the state space Ω we can find the invariant integration measure d_Ω.

1. *Invariant Integration Measure*

In differential geometry the 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

$$d_\Omega = d_{U\Omega U^{-1}} \Leftrightarrow L_{Vi}\omega = 0 \qquad i = 1, ..., n^2-1, \tag{9.7}$$

that the Lie derivative of ω with respect to the fundamental field V_i of action of the group $SU(n)$ in the space Ω is zero. The vector fields

$$V_i = V_i^b(x_1, ..., x_{(2n-2)}) \frac{\partial}{\partial x_b}; \qquad b = 1, 2...(2n-2) \tag{9.8}$$

are defined via the actions of one-parametric subgroups $\exp(it\hat{S}_i) \subset SU(n)$, $t \in R$ (one action for each generator $\hat{S}_i$). On the other hand the elements of the space Ω [see Eq. (9.5)] have a structure

$$\hat{\rho}(x_1, ..., x_{(2n-2)}) = \frac{\hat{1}}{n} + f^i(x_1, ..., x_{(2n-2)}) \hat{S}_i, \tag{9.9}$$

where $\hat{S}_i$ are n^2-1 linearly independent, zero-trace, Hermitian, $n \times n$ matrixes; i.e., they are generators of the $SU(n)$ group. Due to this we can express the vector fields V_i,

$$V_i^b \frac{\partial}{\partial x_b} \hat{\rho} = \frac{\partial}{\partial t} [\exp(it\hat{S}_i)\, \hat{\rho} \exp(-it\hat{S}_i)] \Big|_{t=0}, \tag{9.10}$$

as the solutions of the equation

$$V_i^b \frac{\partial}{\partial x_b} f^k = i c_{ij}^k f^j. \tag{9.11}$$

The complex numbers c_{ij}^k are the coefficients in commutation relations $[\hat{S}_i, \hat{S}_j] = c_{ij}^k \hat{S}_k$. We note that Eq. (9.11) represents for each fixed index i an overdetermined system of n^2-1 linear equations for $2n-2$ unknown functions V_i^b (the fact that this system is consistent confirms the correctness of our parameterization of the state space Ω). Finally, we present an explicit coordinate form of Eq. (9.7), which determines the invariant volume form $\omega = m(x_1, ..., x_{(2n-2)}) \wedge dx_1 \wedge \cdots \wedge dx_{(2n-2)}$ as the solution of a system of partial differential equations:

$$\frac{\partial}{\partial x_b}(mV_i^b) = 0. \tag{9.12}$$

Here we note, that $m\vec{V}_i$ in Eq. (9.12) has the meaning of a "flow" of the density of states generated by unitary transformations associated with the ith generator. From the physical point of view Eq. (9.12) means that the divergence of this flow is zero, i.e., the number of states in each (confined) volume element is constant.

As an illustration of the above discussion we firstly evaluate the invariant measure for the state space of a single spin-1/2. Using Eq. (7.1) and the definition (9.11) we find the fundamental field of action V_i ($i = 1, 2, 3$) for the three generators (7.2) of the $SU(2)$ group:

$$V_1 = \cos(\phi)\cot(\theta)\,\partial_\phi + \sin(\phi)\,\partial_\theta \qquad V_2 = \sin(\phi)\cot(\theta)\,\partial_\phi - \cos(\phi)\,\partial_\theta \qquad V_3 = -\partial_\phi. \tag{9.13}$$

We substitute these generators into Eq. (9.12) and after some algebra we obtain the system of differential equations

$$\frac{\partial}{\partial \phi} m = 0 \qquad \frac{\partial}{\partial \theta} m = m\cot(\theta), \tag{9.14}$$

which can be easily solved,

$$m(\theta, \phi) = const\,\sin(\theta). \tag{9.15}$$

The multiplicative factor is given by the normalization condition. This is the route to derive the integration measure of the Poincaré sphere. Analogously we evaluate the invariant integration measure for a state space of two spins-1/2. The calculations are technically more involved, but the result is simple:

$$d_\Omega = \cos^2\alpha\,\sin\alpha\,\sin\theta_1\,\sin\theta_2\,d\alpha\,d\psi\,d\phi_1\,d\theta_1\,d\phi_2\,d\theta_2. \tag{9.16}$$

B. *Quantum Bayesian Inference of the State of Two-Spins*-1/2

To perform the Bayesian reconstruction of density operators of the two-spins-1/2 system we introduce a set of projectors associated with the observables (7.17)

$$\hat{P}_{s,\hat{s}_i^{(1)}} = \frac{(\hat{1} + s\hat{\sigma}_i)}{2} \otimes \hat{1}; \qquad \hat{P}_{s,\hat{s}_i^{(2)}} = \hat{1} \otimes \frac{(\hat{1} + s\hat{\sigma}_i)}{2};$$
$$\hat{P}_{s,\hat{s}_i^{(1)}\hat{s}_j^{(2)}} = \frac{\hat{1}\otimes\hat{1}}{2} + s\frac{\hat{\sigma}_i\otimes\hat{\sigma}_j}{2} \tag{9.17}$$

The corresponding conditional probabilities can be expressed as

$$p(s,\hat{s}_i^{(1)}\,|\,\hat{\rho}(\alpha\cdots))=\frac{1}{2}+s\,\frac{\cos(\alpha)}{2}\,r_i^{(1)};\qquad p(s,\hat{s}_i^{(2)}\,|\,\hat{\rho}(\alpha\cdots))=\frac{1}{2}+s\,\frac{\cos(\alpha)}{2}\,r_i^{(2)}$$

$$p(s,\hat{s}_i^{(1)}\hat{s}_j^{(2)}\,|\,\hat{\rho}(\alpha\cdots))=\frac{1}{2}+s\frac{r_i^{(1)}r_j^{(2)}}{2} + \frac{s\sin\alpha}{2}\left[\cos\psi(k_i^{(1)}k_j^{(2)}-l_i^{(1)}l_j^{(2)})-\sin\psi(k_i^{(1)}l_j^{(2)}+l_i^{(1)}k_j^{(2)})\right], \tag{9.18}$$

where s is the sign of the measured eigenvalue (i.e., the spectrum of observables (7.17) consists only from $\pm 1/2$). Here we comment briefly on the physical meaning of the projectors defined by Eq. (9.17). Namely, the single-particle projectors of the form $\hat{P}_{s,\hat{s}_i^{(1)}}$ are associated with a measurement of the spin component of the first particle in the i-direction ($i=x, y, z$). Obviously this spin component can have only two values, i.e., "up" ($s=1$) and "down" ($s=-1$). In Tables I and II we will denote outcomes of the measurements "up" and "down" as $\uparrow$ and $\downarrow$, respectively. The two-particle projectors $\hat{P}_{s,\hat{s}_i^{(1)}\hat{s}_j^{(2)}}$ are associated with measurements of correlations between the two spin. Namely, if $s=1$, the two spins are *correlated*, which means that they both are registered in the same, yet unspecified, state (that is, both spins are registered either in the state $|\uparrow_1\uparrow_2\rangle$ or $|\downarrow_1\downarrow_2\rangle$). In Tables I and II we will denote this outcome of the measurement as $\uparrow$. On the contrary, if the particles are registered as *anticorrelated*, that is, after the measurement they are in one of the two states $|\uparrow_1\downarrow_2\rangle$ or $|\downarrow_1\uparrow_2\rangle$, then $s=-1$. In Tables I and II we will denote the outcome of this measurement for $\hat{\sigma}_i\otimes\hat{\sigma}_j$ as $\downarrow$.

Now we can apply general rules of Bayesian inference presented in Section IV for a two-spins-1/2 system. We will consider three specific incomplete observation levels and we will derive asymptotic expressions for the density operators in the limit of large number of measurements. We stress here that we assume the measured system to be prepared in a pure state. To be specific, let us suppose that the two spins are prepared in a state described by the state vector (obviously, this can be determined only after an infinite number of measurements on the complete observation level is performed)

$$|\Psi\rangle = A\,|\uparrow\rangle\otimes|\uparrow\rangle + B\,|\downarrow\rangle\otimes|\downarrow\rangle, \tag{9.19}$$

where $|\uparrow\rangle$ and $|\downarrow\rangle$ are eigenstates corresponding to the observable of the spin projection into the z-direction (i.e., $\langle\hat{\sigma}_z\otimes\hat{1}\rangle=\langle\hat{1}\otimes\hat{\sigma}_z\rangle=|A|^2-|B|^2$ and $\langle\hat{\sigma}_z\otimes\hat{\sigma}_z\rangle=1$). When we assume the coefficients $|A|$ and $B|$ to be real, then we can rewrite the density operator (9.19) in the form (9.5), i.e.,

$$\hat{\rho}=\frac{\hat{1}\otimes\hat{1}}{4}+\frac{\hat{\sigma}_z\otimes\hat{\sigma}_z}{4}+\frac{A^2-B^2}{4}(\hat{\sigma}_z\otimes\hat{1}+\hat{1}\otimes\sigma_z)+\frac{AB}{2}(\hat{\sigma}_x\otimes\hat{\sigma}_x-\hat{\sigma}_y\otimes\hat{\sigma}_y), \tag{9.20}$$

TABLE II

Results of *a posteriori* Bayesian Estimation of Density Operators of the Two-Spin-1/2 System

	$\hat{\sigma}_z \otimes \hat{1}$	$\hat{1} \otimes \hat{\sigma}_z$	$\hat{\sigma}_z \otimes \hat{\sigma}_z$	**reconstructed density operator** $\hat{\rho}$	S
1.	$\uparrow$			$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{5}\hat{\sigma}_z \otimes \hat{1}] = \frac{1}{2}[\hat{1} + \frac{1}{5}\hat{\sigma}_z] \otimes \frac{1}{2}\hat{1}$	1.366
2.	$\uparrow^4$			$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{2}\hat{\sigma}_z \otimes \hat{1}] = \frac{1}{2}[\hat{1} + \frac{1}{2}\hat{\sigma}_z] \otimes \frac{1}{2}\hat{1}$	1.255
3.	$\uparrow^5\downarrow$			$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{2}{5}\hat{\sigma}_z \otimes \hat{1}] = \frac{1}{2}[\hat{1} + \frac{2}{5}\hat{\sigma}_z] \otimes \frac{1}{2}\hat{1}$	1.304
4.	$\uparrow^{10}\downarrow^2$			$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{2}\hat{\sigma}_z \otimes \hat{1}] = \frac{1}{2}[\hat{1} + \frac{1}{2}\hat{\sigma}_z] \otimes \frac{1}{2}\hat{1}$	1.255
5.	$\uparrow$	$\uparrow$		$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{5}(\sigma_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{1}{15}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.346
6.	$\uparrow^4$	$\uparrow^4$		$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{799}{1506}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{594}{1506}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.078
7.	$\uparrow^5\downarrow$	$\uparrow^5\downarrow$		$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1175}{2882}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{781}{2882}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.207
8.	$\uparrow^{10}\downarrow^2$	$\uparrow^{10}\downarrow^2$		$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{17395923}{33863032}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{13818228}{33863032}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.091
9.			$\uparrow$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{5}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.366
10.			$\uparrow^4$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{2}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.255
11.			$\uparrow^6$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{3}{5}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.194
12.			$\uparrow^{12}$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{3}{4}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.070
13.	$\uparrow$		$\uparrow$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{5}\hat{\sigma}_z \otimes \hat{1} + \frac{1}{15}\hat{1} \otimes \hat{\sigma}_z + \frac{1}{5}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.346
14.	$\uparrow^4$		$\uparrow^4$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{799}{1506}\hat{\sigma}_z \otimes \hat{1} + \frac{594}{1506}\hat{1} \otimes \hat{\sigma}_z + \frac{799}{1506}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.078
15.	$\uparrow^5\downarrow^1$		$\uparrow^6$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{4509}{10278}\hat{\sigma}_z \otimes \hat{1} + \frac{3879}{10278}\hat{1} \otimes \hat{\sigma}_z + \frac{6221}{10278}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.079
16.	$\uparrow^{10}\downarrow^2$		$\uparrow^{12}$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{5787206}{10556539}\hat{\sigma}_z \otimes \hat{1} + \frac{5542104}{10556539}\hat{1} \otimes \hat{\sigma}_z + \frac{7953979}{10556539}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	0.889
17.	$\uparrow$	$\uparrow$	$\uparrow$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1}{4}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{1}{4}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.303
18.	$\uparrow^4$	$\uparrow^4$	$\uparrow^4$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{281501}{441004}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{281501}{441004}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	0.883
19.	$\uparrow^5\downarrow$	$\uparrow^5\downarrow$	$\uparrow^6$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{3126849}{6044314}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{3873174}{6044314}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	0.988
20.	$\uparrow^{10}\downarrow^2$	$\uparrow^{10}\downarrow^2$	$\uparrow^{12}$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{1372112265600137}{2298883143280046}(\hat{\sigma}_z \otimes \hat{1} + \hat{1} \otimes \hat{\sigma}_z) + \frac{1766795516375970}{2298883143280046}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	0.831

Note. We also present explicit values of the von Neumann entropy associated with given measured data.

with $\psi = 0$, $\phi_1 = \pi/2$, $\theta_1 = 0$, $\phi_2 = \pi/2$, $\theta_2 = 0$ and $\sin \alpha/2 = A$. In what follows we perform *a posteriori* estimation of the density operator based on incomplete data obtained from three different fictitious measurement sequences.

1. *Observation Level* $\mathcal{O}_A^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z^{(2)}\}$

In the first sequence of measurements we reconstructa density operator from data which refer to a measurement of the first spin-1/2 in the direction z, i.e., only the spin component $\hat{s}_z^{(1)}$ is measured (see lines 1–4 in Table II). We see that if only one spin is measured, then the reconstructed two-spin density operator can be factorized, while, as expected, the state of the unmeasured spin is estimated as $\rho = \hat{1}/2$. Obviously, in this kind of measurement, correlations between the two spins cannot be revealed, i.e., the estimated value of $\hat{\sigma}_z \otimes \hat{\sigma}_z$ is equal to zero. As in the case of the reconstruction of a single-spin-1/2 state, the reconstructed density operators describe statistical mixtures and the corresponding von Neumann entropy is directly related to the fidelity of the reconstruction. The maximum value of the von Neumann entropy is in the case of two-spins-1/2 equal to $S = \ln 4 \simeq 1.386$. This entropy is associated with the "total" mixture of the two-spin-1/2 system and in our case it reflects a complete lack of information about the state of the measured system (i.e., we have no knowledge about the state before a measurement is performed). As soon as the first measurement is performed, we gain some knowledge about the state of the system and the entropy of the estimated density operator is smaller than ln 4 (see line 1 in Table II).

Let us assume now that data from the measurement of the spin components $\hat{s}_z^{(1)}$ and $\hat{s}_z^{(2)}$ of the first and the second particle (spin-1/2), are available. In Table II (lines 5–8) we present results of a reconstruction procedure based on the given "measured" data. We see that though correlations between the two spins have not been measured directly our estimation procedure provides us with a nontrivial estimation for this observable (i.e., the density operator cannot be factorized). Obviously, this estimation is affected by the prior assumption about the purity of the reconstructed state. We see that with the increased number of detected spins the von Neumann entropy of the estimated density operator decreases (we note that it does not decrease monotonically as a function of the number of measurements).

In the limit of large (infinite) number of measurements spectral distributions Eq. (5.1) associated with observables on a given incomplete observation level are precisely determined by the measured data. Using the parameterization introduced earlier in this section [see Eqs. (9.5) and (9.16–9.18)] we can write down the expression (5.12) for the Bayesian *a posteriori* estimation of the density operator in the limit of large number of measurements. After we perform some trivial integrations and when the substitution $\cos\alpha = x$, $\cos\theta_1 = y$, $\cos\theta_2 = z$ is performed we can write the reconstructed density operator as

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{-1}^{1} x^2\, dx \int_{-1}^{1} dy \int_{-1}^{1} dz\, \delta(\langle \hat{\sigma}_z^{(1)} \rangle - xy)\, \delta(\langle \hat{\sigma}_z^{(2)} \rangle - xz)$$
$$\times (\hat{1} \otimes \hat{1} + xy\hat{\sigma}_z \otimes \hat{1} + xz\hat{1} \otimes \hat{\sigma}_z + yz\hat{\sigma}_z \otimes \hat{\sigma}_z). \tag{9.21}$$

The right-hand side of Eq. (9.21) can easily be integrated over the variables y and z so we can write

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathcal{L}} dx \left(\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \frac{\langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(2)} \rangle}{x^2}\, \hat{\sigma}_z \otimes \hat{\sigma}_z \right), \tag{9.22}$$

where the integration is performed over the interval $\mathcal{L}$,

$$\mathcal{L} := \{-1, 1\} \qquad \text{and} \qquad |x| \geqslant s_{max}, \tag{9.23}$$

with $s_{max} = \max\{|\langle \hat{\sigma}_z^{(1)} \rangle|, |\langle \hat{\sigma}_z^{(2)} \rangle|\}$. After we perform the integration over the variable x we find

$$\hat{\rho} = \frac{1}{4} \left(\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \frac{\langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(2)} \rangle}{s_{max}}\, \hat{\sigma}_z \otimes \hat{\sigma}_z \right). \tag{9.24}$$

Comparing Eqs. (7.19) and (9.24) we see that on the observation level $\mathcal{O}_B^{(2)}$ the quantum Bayesian inference and the Jaynes principle of maximum entropy provides us with the different *a posteriori* estimations of density operators. To be specific, the density operator (7.19) obtained with the help of the MaxEnt principle can be expressed in a factorized form while the density operator (9.24) cannot be factorized into a product of two density operators describing each spin separately [the only exception is when $s_{max} = 1$].

2. *Observation Level* $\mathcal{O}_B^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z^{(1)}\hat{s}_z^{(2)}\}$

Here we start our discussion with an assumption that only correlations between the particles are measured, while the state of each individual particle after the measurement is unknown (see lines 9–12 in Table II). In this case we are not able to make any nontrivial estimation for the mean values of the spin components of the individual particles. In order to have a better estimation we also have to measure at least one of the spin components of the first or the second spin.

Let us assume that the z-component of the first spin and the correlation $\hat{s}_z^{(1)}\hat{s}_z^{(2)}$ are measured. That is, the z-component of the second spin $\hat{s}_z^{(2)}$ is not directly observed. The question is, What is the estimation of the density operator on this observation level and in particular, what is the estimation for the mean value of the observable $\hat{s}_z^{(2)}$? In Table II (lines 13–16) we present numerical results for the *a posteriori* estimation of the density based on a finite set of "experimental" data. We see that the Bayesian scheme provides us with a nontrivial (i.e., nonzero) estimation of the mean value of $\hat{s}_z^{(2)}$. But the question is whether in the limit of a large number of measurements this is equal to the mean value estimated with the help of the Jaynes principle of maximum entropy. The expression for the *a posteriori* Bayes estimation of the density operator in the limit of infinite number of measurements on the given observation level [for technicalities see Appendix A] reads

$$\hat{\rho} = \frac{1}{4}\left[\hat{1}\otimes\hat{1} + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{\sigma}_z\otimes\hat{1} + \frac{\langle\hat{\sigma}_z^{(1)}\rangle\langle\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\rangle}{s_{max}}\,\hat{1}\otimes\hat{\sigma}_z + \langle\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\rangle\,\hat{\sigma}_z\otimes\hat{\sigma}_z\right], \tag{9.25}$$

where $s_{max} = \max\{|\langle\hat{\sigma}_z^{(1)}\rangle|, |\langle\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\rangle|\}$. Here again the Bayesian *a posteriori* estimation (9.25) is in general different from the estimation (7.21)) obtained with the help of the Jaynes *MaxEnt* principle. We see that these two results coincide only when $s_{max} = 1$. For instance, if $\langle\hat{\sigma}_z\otimes\hat{\sigma}_z\rangle = 1$, then s_{max} is equal to unity and the estimated density operators given by Eqs. (7.21) and (9.25) are equal and read

$$\hat{\rho} = \tfrac{1}{4}[\hat{1}\otimes\hat{1} + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{\sigma}_z\otimes\hat{1} + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{1}\otimes\hat{\sigma}_z + \hat{\sigma}_z\otimes\hat{\sigma}_z]. \tag{9.26}$$

In the case when $\langle\hat{\sigma}_z^{(1)}\rangle = 1$ the von Neumann entropy is equal to zero, i.e., the measured state is completely reconstructed, and is described by the state vector $|\Psi\rangle = |\uparrow_1\uparrow_2\rangle$.

3. *Observation Level* $\mathcal{O}_C^{(2)} = \{\hat{s}_z^{(1)}, \hat{s}_z^{(2)}, \hat{s}_z^{(1)}\hat{s}_z^{(2)}\}$

Finally, we will consider a measurement of both the spin projections $\hat{s}_z^{(1)}$, $\hat{s}_z^{(2)}$, as well as the correlation $\hat{s}_z^{(1)}\hat{s}_z^{(2)}$. Results of an estimation of the density operator based on a sequence of data associated with this observation level are given in Table II (lines 17–20). If an infinite number of measurements on the given observation level is performed then we can evaluate the *a posteriori* density operator analogously to that of the previous example [see Appendix A] and after some algebra we find

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathcal{L}''} \frac{x^2}{|x|}\, dx \int_{z_1}^{z_2} dz\, \frac{\delta(\langle \hat{\sigma}_z^{(2)} \rangle - xz)}{\sqrt{a + bz + cz^2}} \times [\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + xz\hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z \otimes \hat{\sigma}_z]. \tag{9.27}$$

Due to the presence of the δ-function the integration over the parameter z on the right-hand side of Eq. (9.27) is straightforward and we obtain

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathcal{L}''} \frac{dx}{\sqrt{a + bz_0 + cz_0^2}} \times [\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z \otimes \hat{\sigma}_z], \tag{9.28}$$

where $z_0 = \langle \hat{\sigma}_z^{(2)} \rangle / x$. From Eq. (9.28) we directly obtain the reconstructed density operator which reads

$$\hat{\rho} = \tfrac{1}{4}[\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(2)} \rangle\, \hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle\, \hat{\sigma}_z \otimes \hat{\sigma}_z]. \tag{9.29}$$

We see that on the present observation level the density operator (9.29) estimated via Bayesian inference is *equal* to the density operator (7.25) estimated with Jaynes principle of maximum entropy.

C. *Inseparability and Quantum Bayesian Inference*

We first recall that a density operator $\hat{\rho}_{ab}$ describing a system composed of two subsystems $\hat{\rho}_a$ and $\hat{\rho}_{ab}$ is inseparable if it *cannot* be written as the convex sum

$$\hat{\rho}_{ab} = \sum_m w^{(m)} \hat{\rho}_a^{(m)} \otimes \hat{\rho}_b^{(m)}. \tag{9.30}$$

Inseparability is one of the most fundamental quantum phenomena, which, in particular, may result in the violation of Bell's inequality (to be specific, a separable system always satisfy Bell's inequality, but the contrary is not necessarily true). Note that distant parties cannot prepare an inseparable state from a separable state if they only use local operations and classical communications.

In the case of two spins-1/2 we can effectively utilize the Peres–Horodecki theorem [38] which states that the positivity of the partial transposition of a state is *necessary* and *sufficient* for its separability. Before we proceed further we briefly described how to "use" this theorem: The density matrix associated with the density operator of two spins-1/2 can be written as

$$\rho_{m\mu, n\nu} = \langle e_m| \langle f_\mu| \hat{\rho} |e_n\rangle |f_\nu\rangle, \tag{9.31}$$

where $\{|e_m\rangle\}$ ($\{|f_\mu\rangle\}$) denotes an orthonormal basis in the Hilbert space of the first (second) spin-1/2 (for instance, $|e_0\rangle = |0\rangle_a$; $|e_1\rangle = |1\rangle_a$, and $|f_0\rangle = |0\rangle_b$; $|f_1\rangle = |1\rangle_b$). The partial transposition $\hat{\rho}^{T_2}$ of $\hat{\rho}$ is defined as

$$\rho^{T_2}_{m\mu, n\nu} = \rho_{m\nu, n\mu}. \tag{9.32}$$

Then the necessary and sufficient condition for the state $\hat{\rho}$ of two spins-1/2 to be inseparable is that at least one of the eigenvalues of the partially transposed operator (9.32) is negative.

1. *Inseparability of Estimated States*

To determine whether the data obtained from a incomplete measurement on a given observation level would allow us to draw any conclusion about nonclassical entanglement between two spins-1/2, we have to make sure that the used observation level is suitable for a "detection" for quantum entanglement. For instance, the density operators reconstructed on the observation levels $\mathcal{O}^{(2)}_A - \mathcal{O}^{(2)}_C$ in the basis of eigenstates of $\hat{\sigma}_z$ operators are *diagonal.* Consequently, the partially transposed density matrix is equal to the original density matrix. This means that all eigenvalues of the partially transposed matrix are *positive* irrespective of the number of measurements performed on these observation levels [see for instance, Eq. (9.29)]. From here it follows that the given observation levels are *not suitable* if we want to analyze whether the measured state is quantum-mechanically entangled. It is interesting to note that the estimation (9.29) has been made under the assumption that the two-particle system is in a pure state. Obviously, there are pure two-particle states which are disentangled (i.e., $|\Psi\rangle_{ab} = |\Psi\rangle_a |\Psi\rangle_b$) but all other *pure* two-particle state are inseparable. But because we have chosen "very bad" observation levels we are not able to reconstruct density operators which are inseparable.

To overcome this problem we have to extend our observation level. Namely, to detect any quantum entangled (which for instance can be seen through a violation of Bell inequalities [1]) we have to rotate the Stern–Gerlach apparatuses with which the measurement is performed. In particular, we can rotate them so that the projector $\hat{P}_{s, \hat{s}^{(1)}_x \hat{s}^{(2)}_x}$ (see Eq. (9.18) is measured. In this case the reconstructed density operator will contain nontrivial information about the quorum of operators, i.e., information about the mean values of the observables $\hat{\sigma}_x \otimes \hat{\sigma}_x$ and $\hat{\sigma}_y \otimes \hat{\sigma}_y$ which would allow us to check whether the measured state is separable.

To be specific, let us assume an extension of the observation level $\mathcal{O}^{(2)}_C = \{\hat{s}^{(1)}_z, \hat{s}^{(2)}_z, \hat{s}^{(1)}_z \hat{s}^{(2)}_z\}$ such that the operator $\hat{s}^{(1)}_x \hat{s}^{(2)}_x$ is included. We will study two estimations of the density operator after the measurement over N ($N = 5, ..., 8$) pairs, respectively,of two spins-1/2 is has been performed (we present results of the recontsruction in Table III). From the "measured" data we can estimate the

TABLE III

Results of *a posteriori* Bayesian Estimation of Density Operators of the Two-Spin-1/2 System

	$\hat{\sigma}_z \otimes \hat{1}$	$\hat{1} \otimes \hat{\sigma}_z$	$\hat{\sigma}_z \otimes \hat{\sigma}_z$	$\hat{\sigma}_x \otimes \hat{\sigma}_x$	**reconstructed density operator** $\hat{\rho}$	S
1.	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\uparrow^4$	$\uparrow$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{118283}{423147}\hat{\sigma}_x \otimes \hat{\sigma}_x - \frac{10699}{49782}\hat{\sigma}_y \otimes \hat{\sigma}_y + \frac{24233}{49782}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.219
2.	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\uparrow^4$	$\uparrow^2$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{118283}{270715}\hat{\sigma}_x \otimes \hat{\sigma}_x - \frac{181883}{541430}\hat{\sigma}_y \otimes \hat{\sigma}_y + \frac{27090}{54143}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.146
3.	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\uparrow^4$	$\uparrow^3$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{7884671}{14781924}\hat{\sigma}_x \otimes \hat{\sigma}_x - \frac{317261}{777996}\hat{\sigma}_y \otimes \hat{\sigma}_y + \frac{400739}{777996}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.081
4.	$\uparrow^2\downarrow^2$	$\uparrow^2\downarrow^2$	$\uparrow^4$	$\uparrow^4$	$\frac{1}{4}[\hat{1} \otimes \hat{1} + \frac{13559668}{22666595}\hat{\sigma}_x \otimes \hat{\sigma}_x - \frac{10288728}{22666595}\hat{\sigma}_y \otimes \hat{\sigma}_y + \frac{11973928}{22666595}\hat{\sigma}_z \otimes \hat{\sigma}_z]$	1.029

Note. We also present explicit values of the von Neumann entropy associated with given measured data.

two-spins-1/2 density operators as shown in the table. The estimated density operator in line 1 is separable, but the density operators in lines 2–4 are inseparable. That is, with the particular outcomes of the measurement, we can declare (with a certain degree of fidelity associated with the corresponding von Neumann entropy) that the system under consideration is inseparable.

X. RECONSTRUCTION OF IMPURE STATES VIA QUANTUM BAYESIAN INFERENCE

In this section we apply the purification ansatz as shown in Section VI for a reconstruction (estimation) of an impure state of a single spin-1/2. To do so, we apply the results of the previous section where we have discussed the Bayesian estimation of pure two-spins-1/2 states. In particular, in lines 1–4 of Table II we present results of the estimation of a two-spin density operator based on "results" of measurements of the $\hat{\sigma}_z$-component of just one spin-1/2. We see that in this case the two-spin density operator can be written in a factorized form, $\hat{\rho}_{ab} = \hat{\rho}_a \otimes \frac{1}{2}\hat{1}$. In this case we can easily trace over the unmeasured spin and we obtain the estimation for the density operator of the first spin (compare with lines 1–4 in Table I). This estimation is not based on the *a priori* purity assumption.

Comparing results of two estimations which differ by the *a priori* assumption about the purity of the reconstructed state we can conclude the following:

(1) In general, under the purity assumption the reconstruction procedure converges faster (simply compare the two columns in Table I) to a particular result. This is easy to understand, because in the case when the purity of measured states is *a priori* assumed, the state space of all possible states is much smaller compared to the state space of all possible (pure and impure) states.

(2) When the measured data are inconsistent with an a priori purity assumption, then estimations based on this assumption become incorrect. For instance, for the "measured" data presented in lines 14–17 of Table I we find that the estimated mean values of $\hat{\sigma}_z$ diverge from the expected mean value 1/2 (i.e., this is the mean value of $\hat{\sigma}_z$ when we detect in a sequence of $4N$ measurements $3N$ spins "up" and N spins "down"). As we have shown in Section VIII.A.3 in the limit $N \to \infty$ the reconstruction

can completely fail when the purity condition is imposed. On the other hand, if it is *a priori* assumed that the measured state can be in a statistical mixture, then the Bayesian quantum inference provides us with estimations which in the limit $N \to \infty$ coincide with estimations based on the Jaynes principle of maximum entropy.

A. *Estimation in the Limit of* $N \to \infty$ *Measurements*

1. *Observation Level* $\mathcal{O}_A^{(1)} = \{\hat{s}_z^{(1)}\}$

Using the techniques which have been demonstrated in Section IX we can express the estimated density operator on the given observation level in the limit $N \to \infty$ [see Eq. (5.12)]. We note that on the considered observation level, Eq. (5.12) contains many terms, which are odd functions of the corresponding integration variables. Therefore the integration over these parameters ($\theta_2, \phi_2, \psi, \phi_1$) is straightforward. Moreover, if we perform the trace over the "second" (reservoir) spin we can express the density operator of the spin-1/2 under consideration as

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{-1}^{1} y^2 \, dy \int_0^{\pi} \sin\theta_1 \, d\theta_1 \, \delta(\langle \hat{\sigma}_z^{(1)} \rangle - y \cos\theta_1)(\hat{1} + y \cos\theta_1 \hat{\sigma}_z), \tag{10.1}$$

where the variable α is substituted by $y = \cos\alpha$. When we perform integration over y we obtain the expression

$$\hat{\rho} = \frac{2}{\mathcal{N}} \int_{\mathcal{L}} d\theta_1 \frac{\sin\theta_1}{\cos^2\theta_1 \, |\cos\theta_1|} (\hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \, \hat{\sigma}_z), \tag{10.2}$$

with $\mathcal{L}$ defined as

$$\mathcal{L} := \{0, \pi\} \qquad \text{such that} \quad |\cos\theta_1| \geqslant |\langle \hat{\sigma}_z^{(1)} \rangle|. \tag{10.3}$$

After we perform the integration over θ_1 we obtain the expression for the density operator identical to that obtained via the Jaynes principle of maximum entropy [see Eq. (7.7)].

2. *Observation Level* $\mathcal{O}_B^{(1)} = \{\hat{s}_z^{(1)}, \hat{s}_x^{(1)}\}$

In the limit of infinite number of measurements one can express the Bayesian estimation of the density operator of the spin-1/2 on the given observation level as (here the trace over the "reservoir" spin has already been performed)

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{-1}^{1} y^2 \, dy \int_0^{\pi} \sin\theta_1 \, d\theta_1 \int_0^{2\pi} d\phi_1 \, \delta(\langle \hat{\sigma}_z^{(1)} \rangle - y \cos\theta_1) \, \delta(\langle \hat{\sigma}_x^{(1)} \rangle - y \sin\theta_1 \cos\phi_1)$$
$$\times (\hat{1} + y \sin\theta_1 \cos\phi_1 \hat{\sigma}_x + y \sin\theta_1 \sin\phi_1 \hat{\sigma}_y + y \cos\theta_1 \hat{\sigma}_z). \tag{10.4}$$

When we perform integration over the variable y we find

$$\hat{\rho}=\frac{1}{\mathcal{N}}\int_0^{2\pi} d\phi_1 \int_{\mathcal{L}'} d\theta_1 \frac{\sin\theta_1}{\cos^2\theta_1\,|\cos\theta_1|}\langle\hat{\sigma}_x^{(1)}\rangle-\tan\theta_1\cos\phi_1\langle\hat{\sigma}_z^{(1)}\rangle)$$
$$\times(\hat{1}+\langle\hat{\sigma}_z^{(1)}\rangle\tan\theta_1\cos\phi_1\hat{\sigma}_x+\langle\hat{\sigma}_z^{(1)}\rangle\tan\theta_1\sin\phi_1\hat{\sigma}_y+\langle\hat{\sigma}_z^{(1)}\rangle\hat{\sigma}_z). \tag{10.5}$$

The integration over the variable ϕ_1 in the right-hand side of Eq. (10.5) gives us

$$\hat{\rho}=\frac{1}{\mathcal{N}}\int_{\mathcal{L}''} d\theta_1 \sum_{j=1}^{2}\frac{1}{\cos^2\theta_1\,|\sin\phi_1^{(j)}|}$$
$$\times(\hat{1}+\langle\hat{\sigma}_x^{(1)}\rangle\hat{\sigma}_x+\langle\hat{\sigma}_z^{(1)}\rangle\tan\theta_1\sin\phi_1^{(j)}\hat{\sigma}_y+\langle\hat{\sigma}_z^{(1)}\rangle\hat{\sigma}_z), \tag{10.6}$$

where the integration is performed over the interval

$$\mathcal{L}'':=\{0,\pi\} \qquad \text{such that} \quad |\cos\theta_1|\geqslant|\langle\hat{\sigma}_z^{(1)}\rangle|, \qquad \text{and} \qquad |\tan\theta_1|\geqslant\left|\frac{\langle\hat{\sigma}_x^{(1)}\rangle}{\langle\hat{\sigma}_z^{(1)}\rangle}\right|. \tag{10.7}$$

The sum in Eq. (10.6) is performed over two values $\phi_1^{(j)}$ of the variable ϕ_1 which are equal to the two solutions of the equation

$$\cos\phi_1=\frac{\langle\hat{\sigma}_x^{(1)}\rangle}{\langle\hat{\sigma}_z^{(1)}\rangle\tan\theta_1}. \tag{10.8}$$

Due to the fact that the term in front of the operator $\hat{\sigma}_y^{(1)}$ is the odd function of $\phi_1^{(j)}$, we can straightforwardly perform in Eq. (10.6) the integration over θ_1 and we find the expression of the reconstructed density operator which again is exactly the same as if we perform the reconstruction with the help of the Jaynes principle [see Eq. (7.10)].

3. *Observation Level* $\mathcal{O}_C^{(1)}=\{\hat{s}_z^{(1)},\hat{s}_x^{(1)},\hat{s}_y^{(1)}\}$

On the complete observation level, the expression for the Bayesian estimation of the density operator of the spin-1/2 in the limit of infinite number of measurements can be expressed as (here again we have already traced over the "reservoir" degrees of freedoms) [see Eq. (10.4)]

$$\hat{\rho}=\frac{1}{\mathcal{N}}\int_{-1}^{1} y^2\,dy \int_0^{\pi}\sin\theta_1\,d\theta_1\int_0^{2\pi}d\phi_1$$
$$\times\delta(\langle\hat{\sigma}_z^{(1)}\rangle-y\cos\theta_1)\,\delta(\langle\hat{\sigma}_x^{(1)}\rangle-y\sin\theta_1\cos\phi_1)\,\delta(\langle\hat{\sigma}_y^{(1)}\rangle-y\sin\theta_1\sin\phi_1)$$
$$\times(\hat{1}+y\sin\theta_1\cos\phi_1\hat{\sigma}_x+y\sin\theta_1\sin\phi_1\hat{\sigma}_y+y\cos\theta_1\hat{\sigma}_z). \tag{10.9}$$

Performing similar calculations as in the previous subsection we can rewrite Eq. (10.9) as

$$\hat{\rho} \simeq \int_{\mathscr{L}''} d\theta_1 \sum_{j=1}^{2} \frac{1}{\cos^2\theta_1\,|\sin\phi_1^{(j)}|}\,\delta(\langle\hat{\sigma}_y^{(1)}\rangle - \tan\theta_1 \sin\phi_1^{(j)}\langle\hat{\sigma}_z^{(1)}\rangle)$$
$$\times(\hat{1} + \langle\hat{\sigma}_x^{(1)}\rangle\,\hat{\sigma}_x + \langle\hat{\sigma}_y^{(1)}\rangle \tan\theta_1 \sin\phi_1^{(j)}\hat{\sigma}_y + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{\sigma}_z)\otimes\hat{1}, \qquad (10.10)$$

where $\mathscr{L}''$ and $\phi_1^{(j)}$ are defined by Eqs. (10.7) and (10.8), respectively. Now the integration over θ_1 can be easily performed and for the density operator of the given spin-1/2 system we find

$$\hat{\rho} = \tfrac{1}{2}[\hat{1} + \langle\hat{\sigma}_x^{(1)}\rangle\,\hat{\sigma}_x + \langle\hat{\sigma}_y^{(1)}\rangle\,\hat{\sigma}_y + \langle\hat{\sigma}_z^{(1)}\rangle\,\hat{\sigma}_z], \qquad (10.11)$$

where the mean values $\langle\hat{\sigma}_j^{(1)}\rangle$ do not necessarily satisfy the purity condition (7.16).

XI. CONCLUSIONS

In the paper we have analyzed in detail the logical connection between three different reconstruction schemes: (1) If measurements over a finite number of elements of the ensemble are performed then one can obtain the *a posteriori* estimation of the density operator with the help of Bayesian inference. If nothing is known about the reconstructed state one has to assume a constant prior probability distribution on the parametric state space under the assumption that the system is in a statistical mixture. (2) As soon as the number of measurements becomes large the Bayesian inference scheme becomes equal to the reconstruction scheme based on the Jaynes principle of maximum entropy; i.e., in the limit of infinite number of measurements *a posteriori* estimated density operator fulfills the condition of the maximum entropy. Consequently, it is equal to the generalized canonical density operator. (3) If the quorum of observables is measured, then the generalized canonical operator is equal to the "true" density operator of the system itself, i.e., a complete reconstruction via the MaxEnt principle is performed. It is a matter of technical convenience which reconstruction scheme on the complete observation level is utilized (for instance, quantum tomography can be used), but all of these complete reconstruction schemes can be formulated as a maximization of the entropy under given constraints.

APPENDIX

Bayesian Inference on $\mathcal{O}_B^{(2)}$ in the Limit of Infinite Number of Measurements

On the given observation level we can express the estimated density operator in the limit of infinite number of measurements as [see Eq. (5.12)]

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\psi \int_{-1}^{1} x^2\, dx \int_{-1}^{1} dy \int_{-1}^{1} dz$$
$$\times \delta(\langle \hat{\sigma}_z^{(1)} \rangle - xy)\, \delta(\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle - yz + [(1-x^2)(1-y^2)(1-z^2)]^{1/2} \cos\psi)$$
$$\times \{\hat{1} \otimes \hat{1} + xy\hat{\sigma}_z \otimes \hat{1} + xz\hat{1} \otimes \hat{\sigma}_z$$
$$+ [yz - ((1-x^2)(1-y^2)(1-z^2))^{1/2} \cos\psi]\, \hat{\sigma}_z \otimes \hat{\sigma}_z\}. \tag{A.1}$$

We integrate Eq. (A.1) over the variable y and we obtain

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\psi \int_{\mathcal{L}'} \frac{x^2}{|x|}\, dx \int_{-1}^{1} dz$$
$$\times \delta\left(\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle - \langle \hat{\sigma}_z^{(1)} \rangle \frac{z}{x} + \left[(1-x^2)(1-z^2) \left(1 - \frac{\langle \hat{\sigma}_z^{(1)} \rangle^2}{x^2}\right)\right]^{1/2} \cos\psi \right)$$
$$\times \left\{ \hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle\, \hat{\sigma}_z \otimes \hat{1} + xz\hat{1} \otimes \hat{\sigma}_z \right.$$
$$\left. + \left[\langle \hat{\sigma}_z^{(1)} \rangle \frac{z}{x} - \left((1-x^2)(1-z^2) \left(1 - \frac{\langle \hat{\sigma}_z^{(1)} \rangle^2}{x^2}\right)\right)^{1/2} \cos\psi \right) \right] \hat{\sigma}_z \otimes \hat{\sigma}_z \right\}, \tag{A.2}$$

where the integration boundaries are defined as

$$\mathcal{L}'' := \{-1, 1\} \qquad \text{and} \qquad |x| \geqslant |\langle \hat{\sigma}_z^{(1)} \rangle|. \tag{A.3}$$

Now we will integrate Eq. (A.2) over the variable ψ. There are two values $\psi_0^{(j)}$ $(j = 1, 2)$ of ψ, such that

$$\cos\psi_0 = \frac{\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle - \langle \hat{\sigma}_z^{(1)} \rangle\, z/x}{[(1-x^2)(1-z^2)(1-(\langle \hat{\sigma}_z^{(1)} \rangle / x)^2)]^{1/2}}, \tag{A.4}$$

providing that inequality

$$1 \geqslant \left| \frac{\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle - \langle \hat{\sigma}_z^{(1)} \rangle\, z/x}{[(1-x^2)(1-z^2)(1-(\langle \hat{\sigma}_z^{(1)} \rangle / x)^2)]^{1/2}} \right| \tag{A.5}$$

holds. The last relation can be rewritten as the condition $a + bz + cz^2 \geqslant 0$, where the explicit forms of the coefficients a, b, and c are

$$a = 1 - \langle \hat{\sigma}_z^{(1)} \rangle^2 / x^2 + \langle \hat{\sigma}_z^{(1)} \rangle^2 - \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle^2 - x^2;$$
$$b = 2\langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle / x; \qquad c = x^2 - \langle \hat{\sigma}_z^{(1)} \rangle^2 - 1. \tag{A.6}$$

The coefficient c is always negative, which means that we have a new condition for the parameter z, that is, $z \in \langle z_1, z_2 \rangle$, where z_1 and z_2 are two roots of the quadratic equation $a + bz + cz^2 = 0$. However, these roots exist only providing the discriminant $b^2 - 4ac \geqslant 0$ is nonnegative. Taking into account Eq. (A.6)we see that the last relation is a cubic equation with respect to the variable x^2, which imposes a new condition on the integration parameter x. That is, the interval $\mathscr{L}''$ through which the integration over x in Eq. (A.2) is performed is defined as

$$\mathscr{L}'' := \begin{cases} \{|\langle \hat{\sigma}_z^{(1)} \rangle|, 1\} & \text{for} \quad |\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle| \leqslant |\langle \hat{\sigma}_z^{(1)} \rangle|; \\ \{|\langle \hat{\sigma}_z^{(1)} \rangle|, \sqrt{1 + \langle \hat{\sigma}_z^{(1)} \rangle^2 - \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle^2}\} & \text{for} \quad |\langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle| \geqslant |\langle \hat{\sigma}_z^{(1)} \rangle|. \end{cases} \tag{A.7}$$

Taking into account all conditions imposed on parameters of integration we canrewrite Eq. (A.2) as

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathscr{L}''} \frac{x^2}{|x|} \, dx \int_{z_1}^{z_2} \frac{dz}{\sqrt{a + bz + cz^2}} \times (\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \, \hat{\sigma}_z \otimes \hat{1} + xz \hat{1} \otimes \hat{\sigma}_z + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle \, \hat{\sigma}_z \otimes \hat{\sigma}_z). \tag{A.8}$$

Using standard formulas [see, for example, [37, Eq. (2.261) and Eq. (2.264)]] the integration over parameter z in Eq. (A.8) can now be performed and we obtain

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathscr{L}''} dx \, \frac{x^2}{|x|} \, \frac{1}{(1 + \langle \hat{\sigma}_z^{(1)} \rangle^2 - x^2)^{1/2}} (\hat{1} \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \, \hat{\sigma}_z \otimes \hat{1} + \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle \, \hat{\sigma}_z \otimes \hat{\sigma}_z) + \int_{\mathscr{L}''} dx \, \frac{x^2}{|x|} \, \frac{\langle \hat{\sigma}_z^{(1)} \rangle \langle \hat{\sigma}_z^{(1)} \hat{\sigma}_z^{(2)} \rangle}{(1 + \langle \hat{\sigma}_z^{(1)} \rangle^2 - x^2)^{3/2}} (\hat{1} \otimes \hat{\sigma}_z). \tag{A.9}$$

After performing integration over x in Eq. (A.9) we obtain final the expression (9.25) for the *a posteriori* estimation of the density operator on the given observation level.

ACKNOWLEDGMENTS

We thank Jason Twamley, Sam Braustein, and Zdeněk Hradil for stimulating discussions. This work was supported in part by the United Kingdom Engineering and Physical Sciences Research Council, the European Community, and 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 Contract 45.367/6-IV/3a/95 of the Österreichisches Bundesministerium für Wissenschaft und Forschung.

REFERENCES

1. A. Peres, “Quantum Theory: Concepts and Methods,” Kluwer, Dordrecht, 1993.
2. R. Omnès, “The Interpretation of Quantum Mechanics,” Princeton Univ. Press, Princeton, 1994.
3. L. E. Ballentine, “Quantum Mechanics,” Prentice Hall, Englewood Cliffs, NJ, 1990.
4. See “Quantum State Preparation and Measurement,” special issue of *J. Mod. Opt.* **44**, No. 11/12 (1997), edited by W. P. Schleich and M. G. Raymer.
5. W. Band and J. L. Park, *Amer. J. Phys.* **47** (1979), 188; *Found. Phys.* **1** (1970), 133; *Found. Phys.* **1** (1971), 339; J. L. Park and W. Band, *Found. Phys.* **1** (1971), 211.
6. R. G. Newton and Bing-Lin Young, *Ann. Phys. (N.Y.)* **49** (1968), 393.

7. J. von Neumann, "Mathematical Foundations of Quantum Mechanics," Princeton Univ. Press, Princeton, 1955; see also J. A. Wheeler and W. H. Zurek, "Quantum Theory and Measurement," Princeton Univ. Press, Princeton, 1983.
8. W. Pauli, "General Principles of Quantum Mechanics," Springer-Verlag, Berlin, 1980; see also B. d'Espagnat, "Conceptual Foundations of Quantum Mechanics," 2nd ed., Benjamin, Reading, 1976.
9. W. Gale, E. Guth, and G. T. Trammel, *Phys. Rev.* **165** (1968), 1434; see also A. Orlowski and H. Paul, *Phys. Rev.* **50** (1994), R921.
10. K. Vogel and H. Risken, *Phys. Rev. A* **40** (1989), 2847; see also J. Bertrand and P. Bertrand, *Found. Phys.* **17** (1987), 397; M. Freyberger, K. Vogel, and W. P. Schleich, *Phys. Lett. A* **176** (1993), 41; H. Kühn, D.-G. Welsch, and W. Vogel, *J. Mod. Opt.* **41** (1994), 1607; G. S. Agarwal and S. Chaturvedi, *Phys. Rev. A* **49** (1994), R665; W. Vogel and D.-G. Welsch, *Acta Phys. Slov.* **45** (1995), 313; U. Leonhardt and H. Paul, *Phys. Lett. A* **193** (1994), 117; U. Leonhardt and H. Paul, *J. Mod. Opt.* **41** (1994), 1427; H. Kühn, D.-G. Welsch, and W. Vogel, *J. Mod. Opt.* **41** (1994), 1607.
11. G. M. D'Ariano, C. Machiavelo, and M. G. A. Paris, *Phys. Rev. A* **50** (1994), 4298; U. Leonhardt, H. Paul, and G. M. D'Ariano, *Phys. Rev. A* **52** (1995), 4899; G. M. D'Ariano, U. Leonhardt, and H. Paul, *Phys. Rev. A* **52** (1995), R1801; Th. Richter, *Phys. Lett. A* **211** (1996), 327; Th. Richter, *Phys. Rev. A* **53** (1996), 1197; U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. G. Raymer, *Opt. Commun.* **127** (1996), 144.
12. D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, *Phys. Rev. Lett.* **70** (1993), 1244; D. T. Smithey, M. Beck, J. Cooper, and M. G. Raymer, *Phys. Scr. T* **48** (1993), 35; M. Beck, M. G. Raymer, I. A. Walmsley, and V. Wong, *Opt. Lett.* **18** (1993), 2041; M. Beck, D. T. Smithey, and M. G. Raymer, *Phys. Rev. A* **48** (1993), R890; see also M. G. Raymer, M. Beck, and D. F. Mc Alister, *Phys. Rev. Lett.* **72** (1994), 1137; M. G. Raymer, D. T. Smithey, M. Beck, and J. Cooper, *Acta Phys. Pol.* **86** (1994), 71.
13. U. Janicke and M. Wilkens, *J. Mod. Opt.* **42** (1995), 2183.
14. Ch. Kurtsiefer, T. Pfau, and J. Mlynek, *in* "Technical Digest of the European Quantum Electronic Conference," p. 132, Hamburg, Germany, 1996; S. Schiller, G. Breitenbach, S. F. Pereira, T. Müller, and J. Mlynek, *Phys. Rev. Lett.* **77** (1996), 2933; Ch. Kurtsiefer, T. Pfau, and J. Mlynek, *Nature* **386** (1997), 150.
15. S. Wallentowitz and W. Vogel, *Phys. Rev. Lett.* **75** (1995), 2932; J. F. Poyatos, R. Walser, J. I. Cirac, P. Zoller, and R. Blatt, *Phys. Rev. A* **53** (1996), R1966; C. D'Helon and G. J. Milburn, *Phys. Rev. A* **53** (1996), R25.
16. D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland, *Phys. Rev. Lett.* **77** (1996), 4281; D. Wineland, C. Monroe, D. M. Meekhof, B. E. King, D. Liebfried, W. M. Itano, J. C. Bergquist, D. Berkeland, J. J. Bollinger, and J. Miller, "Quantum State Manipulation of Trapped Atomic Ions," Los Alamos e-print archive quant-ph/9705022 (1997).
17. T. J. Dunn, I. A. Walmsley, and S. Mukamel, *Phys. Rev. Lett.* **74** (1995), 884.
18. U. Leonhardt, *Phys. Rev. Lett.* **74** (1995), 4101; *Phys. Rev. A* **53** (1996), 2998; see also W. K. Wootters, *Found. Phys.* **16** (1986), 391; W. K. Wootters, *Ann. Phys. (N.Y.)* **175** (1987), 1; D. Galetti and A. F. R. De Toledo Piza, *Physica A* **149** (1988), 267.
19. A. Barenco and A. K. Ekert, *Acta Phys. Slovaca* **45** (1995), 205; A. Barenco, *Contemp. Phys.* **37** (1996), 359, and references therein.
20. R. Walser, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **77** (1996), 2658; see also A. S. Parkins, P. Marte, P. Zoller, O. Carnal, and H. J. Kimble, *Phys. Rev. A* **51** (1995), 1578; A. S. Parkins, P. Marte, P. Zoller, and H. J. Kimble, *Phys. Rev. Lett.* **71** (1993), 3095.
21. K. Husimi, *Proc. Phys. Math. Soc. Jpn.* **22** (1940), 264; Y. Kano, *J. Math. Phys.* **6** (1965), 1913.
22. E. Arthurs and J. L. Kelly, Jr., *Bell. Syst. Tech. J.* **44** (1965), 725; W. K. Wootters and W. H. Zurek, *Phys. Rev. D* **19** (1979), 473; Y. Lai and H. A. Haus, *Quantum Opt.* **1** (1989), 99; D. Lalović, D. M. Davidović, and N. Bijedić, *Phys. Rev. A* **46** (1992), 1206; D. M. Davidović and D. Lalović, *J. Phys. A* **26** (1993), 5099; S. Chaturvedi, G. S. Agarwal, and V. Srinivasan, *J. Phys. A* **27** (1994), L39; M. G. Raymer, *Amer. J. Phys.* **62** (1994), 986; U. Leonhardt and H. Paul, *Phys. Rev. A* **48** (1993),

4598; V. Bužek, C. H. Keitel, and P. L. Knight, *Phys. Rev. A* **51** (1995), 2575; see also the review articles S. Stenholm, *Ann. Phys. (N.Y.)* **218** (1992), 233; U. Leonhardt and H. Paul, *Quant. Electron.* **19** (1995), 89; A. Wünsche, and V. Bužek, *J. Quant. Semiclass. Opt.* **9** (1997), 631.
23. K. Wódkiewicz, *Phys. Rev. Lett.* **52** (1984), 1064; *Phys. Lett. A* **115** (1986), 304; *Phys. Lett. A* **129** (1988), 1.
24. N. G. Walker and J. E. Carroll, *Opt. Quant. Electron.* **18** (1986), 335; N. G. Walker, *J. Mod. Opt.* **34** (1987), 15.
25. J. W. Noh, A. Fougéres, and L. Mandel, *Phys. Rev. Lett.* **67** (1991), 1426; *Phys. Rev. A* **45** (1992), 424.
26. T. Opatrný, V. Bužek, J. Bajer, and G. Drobný, *Phys. Rev. A* **52** (1995), 2419; T. Opatrný, D.-G. Welsch, and V. Bužek, *Phys. Rev. A* **53** (1996), 3822.
27. E. T. Jaynes, *Phys. Rev.* **108** (1957), 171; 620; *Amer. J. Phys.* **31** (1963), 66; see also E. Fick and G. Sauermann, "The Quantum Statistics of Dynamic Processes," Springer-Verlag, Berlin, 1990; J. N. Kapur and H. K. Kesavan, "Entropic Optimization Principles with Applications," Academic Press, New York, 1992.
28. V. Bužek, G. Adam, and G. Drobný, *Ann. Phys. (N.Y.)* **245** (1996), 37; *Phys. Rev. A* **54** (1996), 801.
29. H. Wiedemann, "Quantum Tomography with the Maximum Entropy Principle," unpublished manuscript.
30. C. W. Helstrom, "Quantum Detection and Estimation Theory," Academic Press, New York, 1976.
31. A. S. Holevo, "Probabilistic and Statistical Aspects of Quantum Theory," North-Holland, Amsterdam, 1982.
32. K. R. W. Jones, *Ann. Phys. (N.Y.)* **207** (1991), 140; *Phys. Rev. A* **50** (1994), 3682; see also Z. Hradil, *Phys. Rev. A* **55** (1997), R1561.
33. A. Wehrl, *Rev. Mod. Phys.* **50** (1978), 221.
34. E. T. Jaynes, *in* "1962 Brandeis Lectures, Vol. 3," (K. W. Ford, Ed.), p. 181, Benjamin, New York, 1963. Here we note, that the *MaxEnt* principle has a very close relation to the Laplace principle of indifference [see, for instance, H. Jeffreys, "Theory of Probability," Oxford Univ. Press., Oxford, 1960] which states that where nothing is known one should choose a constant-valued function to reflect this ignorance. This obviously maximizes any uncertainty measure.
35. I. S. Gradstein and I. M. Ryzhik, "Table of Integrals, Series, and Products," Academic Press, New York, 1980.
36. A. Peres, *Phys. Rev. Lett.* **77** (1996), 1413; M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223** (1996), 1; *Phys. Rev. Lett.* **78** (1997), 574.
37. A. Uhlmann, *Rep. Math. Phys.* **9** (1976), 273; *Rep. Math. Phys.* **24** (1986), 229.
38. R. Derka, V. Bužek, and G. Adam, *Acta Phys. Slov.* **46** (1996), 355; see also R. Derka, V. Bužek, G. Adam, and P. L. Knight, *J. Fine Mech. Opt.* **11/12** (1996), 341.
39. A. K. Ekert and P. L. Knight, *Amer. J. Phys.* **63** (1995), 415.
40. S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72** (1994), 3439; see also S. L. Braunstein and C. M. Caves, *Ann. New York Acad. Sci.* **755** (1995), 798; S. L. Braunstein, *Phys. Lett. A* **219** (1996), 169.
41. D. Bures, *trans. Am. Math. Soc.* **135** (1969), 199; M. Hübner, *Phys. Lett. A* **163** (1992), 239; *Phys. Lett. A* **179** (1993), 226; J. Twamley, *J. Phys. A* **29** (1996), 3723.
42. P. B. Slater, *Physica A* **214** (1995), 584; P. B. Slater, "Noninformative Priors for Quantum Inference," Los Alamos e-print archive quant-ph/9703012.
43. V. Bužek, G. Drobný, G. Adam, R. Derka, and P. L. Knight, *J. Mod. Opt.* **44** (1997), 2607.