Quantum estimation via the minimum Kullback entropy principle

Stefano Olivares*

Dipartimento di Fisica dell'Università di Milano, I-20133, Italy
and Dipartimento di Matematica e Fisica dell'Università dell'Insubria, I-22100, Como, Italy

Matteo G. A. Paris†

Dipartimento di Fisica dell'Università di Milano, I-20133, Italy
and Institute for Scientific Interchange Foundation, I-10133 Torino, Italy
(Received 8 August 2007; published 22 October 2007)

We address quantum estimation in situations where one has at disposal data from the measurement of an incomplete set of observables and some *a priori* information on the state itself. By expressing the *a priori* information in terms of a bias toward a given state, the problem may be faced by minimizing the quantum relative entropy (Kullback entropy) with the constraint of reproducing the data. We exploit the resulting minimum Kullback entropy principle for the estimation of a quantum state from the measurement of a single observable, either from the sole mean value or from the complete probability distribution, and apply it as a tool for the estimation of weak Hamiltonian processes. Qubit and harmonic oscillator systems are analyzed in some detail.

I. INTRODUCTION

Quantum estimation of states and operations is a relevant topic in the broad field of quantum-information science [1]. The subject has a fundamental interest of its own, since it concerns the characterization of the basic objects of the quantum description of physical systems. In addition, quantum estimation techniques have been receiving attention for their role in the characterization of gates and registers at the quantum level, which itself is a basic tool in the development of quantum-information technology.

In order to characterize a quantum system one may measure an observable, or a set of observables, on repeated preparations of the system. As a matter of fact, the set of observables is usually *incomplete*, i.e., it is not sufficient to give the complete quantum information about the system. In other words, it is not possible to deterministically reconstruct the full density matrix of the system from the measured data [2]. In these cases, the question is not that of finding the actual state of the system, but rather that of estimating the state that best represents the knowledge we have acquired about the system from the measured data [3]. If we assume we have no prior information about the system, and quantify this ignorance by entropy, then the best estimate may be found by the Jaynes maximum entropy principle (MaxEnt) [4], which includes the information obtained by measurements while not allowing one to draw any conclusions not warranted by the data themselves.

On the other hand, in most cases one has some *a priori* information about the state of system under investigation. This may come, e.g., from some energy constraint or from the consideration that the system has been weakly perturbed from an initial given preparation which is under control of the experimenter. A question naturally arises as to how this *a priori* information may be incorporated into the estimation procedure and whether this can be done together with the constraint of reproducing the observed data. The quantum-mechanical way to incorporate some *a priori* information is that of bias of the state to be estimated toward a given quantum state τ, which expresses the information held by the experimenter and which may contain one or more free parameters depending on the amount of *a priori* information available. Upon quantifying the bias toward τ through the relative (Kullback) entropy [5,6], an estimate may be found by minimizing this quantity with the constraint of reproducing the data. This minimum Kullback entropy (MKE) principle emerges naturally as a way to estimate the quantum state of a system from incomplete data, when some *a priori* information on the state is available, i.e., when a bias toward a known quantum state is present. We will exploit these ideas to estimate the quantum state of a system from the measurement of a single observable, either from the sole mean value or from the full probability distribution, and apply it to the estimation of *weak* Hamiltonian processes. In particular, the case of a weak, but otherwise generic, Hamiltonian is analyzed in some detail, with emphasis on qubit and harmonic oscillator systems.

The paper is structured as follows. In Sec. II we review the classical Kullback-Leibler divergence for the probability distribution and its quantum counterpart, the quantum relative Kullback entropy. In addition, we state the minimum Kullback entropy principle which is used in the subsequent sections. In Sec. III we exploit the MKE principle for state estimation. In particular, we analyze MKE estimation from the measurement of a single observable, either from the sole mean value or from the complete probability distribution. Qubit and harmonic oscillator systems are analyzed in some detail. In Sec. IV we address the MKE as a tool in order to estimate a weak Hamiltonian through suitable measurements performed on the evolved states. Section V closes the paper with some concluding remarks.

*stefano.olivares@mi.infn.it
†matteo.paris@unimi.it

II. KULLBACK-LEIBLER DIVERGENCE AND QUANTUM KULLBACK ENTROPY

Let us consider a *classical* system which can be in any of N states, and let p_k be the probability for the kth state. Suppose we want to estimate the probability distribution $\boldsymbol{p} = \{p_k\}$. If all that we know is the number of possible states, there is no way to do better than choosing $p_k = 1/N$, $\forall$ k. Fortunately, in general we have further information about our system, e.g., we have at disposal the value of certain functions of the state of the system. Thus, we can use these as *constraints* to estimate $\boldsymbol{p}$. The problem of finding the most likely distribution satisfying a set of given constraints was solved by Jaynes, who proposed a general method of inference known as the principle of maximum entropy [4]. According to the MaxEnt principle, the best estimate of the $\boldsymbol{p}$ is that *maximizing* the Shannon entropy

$$H(\boldsymbol{p}) = -\sum_k p_k \log p_k \qquad (1)$$

given the constraints. As described by Jaynes, this distribution is the one that agrees with what is known, but expresses maximum uncertainty with respect to all other matters [7].

On the other hand, it is not unlikely that the probability distribution we want to estimate is biased toward *a prior* distribution $\boldsymbol{q} = \{q_k\}$. In this case it is useful to introduce the Kullback entropy [8,9]:

$$K(\boldsymbol{p}|\boldsymbol{q}) = \sum_k p_k \log(p_k/q_k), \qquad (2)$$

which represents the relative entropy of *a posterior* distribution $\boldsymbol{p}$ relative to the prior $\boldsymbol{q}$. It is worth noting that in the absence of any knowledge on $\boldsymbol{q}$, i.e., $q_k = 1/N$ (uniform distribution), we have $K(\boldsymbol{p}|\boldsymbol{q}) = \log N - H(\boldsymbol{p})$: this is just the opposite of the Shannon entropy up to an additive constant. On the contrary, when $\boldsymbol{q}$ is accessible, the Kullback entropy (2) leads to a generalized method of inference: the most likely distribution $\boldsymbol{p}$ with a bias toward $\boldsymbol{q}$ is the one minimizing $K(\boldsymbol{p}|\boldsymbol{q})$ given the constraints. This is the classical principle of minimum Kullback entropy, also referred to as the principle of minimum relative entropy [8–10], which found applications in several branches of science [11–13]. Indeed, minimizing the relative entropy has all the important attributes of the maximum entropy approach, with the advantage that prior information may be easily included.

In the following we will deal with the quantum version of the MKE principle and its application to quantum estimation of states and operations. Before going to the main point, let us spend a few words on how this works in the classical case. Let us consider the problem of finding the posterior $\boldsymbol{p}$ given the prior $\boldsymbol{q}$ and the constraints $\Sigma_k p_k = 1$ (normalization) and $\Sigma_k p_k A_k = \langle A \rangle$, i.e., we assume we know the first moment of the quantity A. Using the Lagrange multiplier method, the problem reduces to the minimization of

$$F(\boldsymbol{p}, \lambda, \eta) = K(\boldsymbol{p}|\boldsymbol{q}) + \eta\left(\sum_k p_k - 1\right) + \lambda\left(\sum_k p_k A_k - \langle A \rangle\right),$$

$$(3)$$

η and λ being Lagrange multipliers. We have

$$0 = \log(p_k/q_k) + 1 + A_k\lambda + \eta \quad (k = 1, \ldots, N), \qquad (4)$$

$$1 = \sum_k p_k, \qquad (5)$$

$$\langle A \rangle = \sum_k p_k A_k. \qquad (6)$$

Solving the first equation with respect to p_k and using the condition (5), we obtain:

$$p_k(\lambda) = \frac{q_k e^{-A_k\lambda}}{\sum_s q_s e^{-A_s\lambda}}, \qquad (7)$$

where we explicitly wrote the dependence of p_k on the Lagrange multiplier λ. Note that for $\lambda = 0$ one has $p_k = q_k$. Differentiating Eq. (7) and using the constraint (6), we obtain the differential equation

$$\frac{dp_k}{d\lambda} = -(A_k - \langle A \rangle)p_k. \qquad (8)$$

Upon considering the distribution $\boldsymbol{p}$ as a point in the probability distribution simplex, then Eq. (8) can be seen as a trajectory in such a space. According to the Lagrange multiplier method, if we integrate the trajectory (8) assuming $\boldsymbol{p}(0) = \boldsymbol{q}$ for $\lambda = 0$, then the minimum of the Kullback entropy is achieved when the trajectory passes through the surface satisfying the constraint $\Sigma_k p_k A_k = \langle A \rangle$ [14].

A. Minimum Kullback entropy principle

Let us turn our attention to quantum mechanics. The problem is now to find the most appropriate density matrix ϱ describing a quantum system that should satisfy some given constraints, which, in turn express the results of an incomplete set of measurements. If there is no *a priori* information, then the optimal choice is given by the density matrix ϱ which maximizes the von Neumann entropy

$$H(\varrho) = -\operatorname{Tr}(\varrho \log \varrho), \qquad (9)$$

and satisfies the constraints, i.e., reproduces the observed data. This is the quantum counterpart of the MaxEnt principle introduced above [3]. On the other hand, when we have some *a priori* information about the system under investigation then the state to be estimated has a bias toward a prior one, τ. This is the case, for example, of a quantum system evolving according to an unknown but weak Hamiltonian starting from a known initial state. In order to build a proper estimation scheme for these situations, let us consider the quantum Kullback entropy, defined as follows [5,6]:

$$K(\varrho|\tau) = \operatorname{Tr}[\varrho(\log \varrho - \log \tau)], \qquad (10)$$

that is, the relative quantum entropy of the density matrix ϱ with respect to τ.

As in the classical case a probability distribution can be seen as a point, in the quantum case the role of the "point" is played by the density matrix ϱ. We can also associate a family of surfaces $\operatorname{Tr}(A\varrho) = \langle A \rangle$ with each quantum observ-

able A. Moreover, it is still possible to define a suitable metric in the space of the density matrices (the Hilbert space) [15]: this allows us to introduce the infinitesimal increment $d\varrho$. In this way one can obtain the following equation for the quantum trajectory [14]:

$$\frac{d\varrho}{d\lambda} = -\frac{1}{2}\{\varrho, A - \langle A \rangle\}, \tag{11}$$

where $\{\hat{X}, \hat{Y}\} = \hat{X}\hat{Y} + \hat{Y}\hat{X}$ and λ is again a Lagrange multiplier. Equation (11) is the quantum counterpart of Eq. (8) and satisfies the constraints $\mathrm{Tr}(\varrho) = 1$ and $\mathrm{Tr}(A\varrho) = \langle A \rangle$, i.e., we are assuming that the expectation value $\langle A \rangle$ of A is known. By formal integration of Eq. (11), it is straightforward to obtain the solution

$$\varrho(\lambda) = \frac{e^{-A\lambda/2} \tau e^{-A\lambda/2}}{\mathrm{Tr}(\tau e^{-A\lambda})}, \tag{12}$$

where we assumed that the formal integration starts from $\varrho(0) = \tau$, τ being the prior density matrix.

There are cases in which the trajectory (11) yields the optimal state with respect to the quantum Kullback entropy, i.e., $K(\varrho(\lambda)|\tau)$ is minimized [15]. In this paper we focus our attention precisely on one of these cases, namely, when the prior τ and the posterior ϱ are *close* each other according to the Fisher information metric.

B. Remarks

As may be expected, when the prior information is very weak the MKE principle reduces to the MaxEnt one. On the other hand, some prior knowledge is often present and the MKE principle fully exploits the additional information to improve the estimation procedure. A question may arise about the choice of the measurements: as a general rule, those should be tuned in order to add information with respect to the prior, being otherwise useless for estimation purposes. In other words, the information coming from the measurements should not be subsumed by the prior one.

The Kullback-Leibler divergence and quantum Kullback entropy are involved in different aspects of quantum estimation of states and operations, including assessments of priors [16] in the context of the quantum Bayes rule [17]. Here we briefly mention a few relevant applications in order to show the analogies with and differences from our approach.

Let us first consider a situation in which we have no *a priori* information and want to estimate the state of a system from the measurement of a set of projector $A_k = |\varphi_k\rangle\langle\varphi_k|$. In this case, the maximum likelihood estimate of the state [18] is the density matrix ϱ that minimizes the Kullback-Leibler divergence $\Sigma_k p_k \log(p_k/\varrho_k)$ between the observed probability distribution and the quantum-mechanical prediction $\varrho_k = \langle\varphi_k|\varrho|\varphi_k\rangle$ [19]. In other words, maximum likelihood estimation leads to a state that fits the given data obtained by the given measurement without using prior information about the quantum state.

In Ref. [20], Kullback entropy is used for quantum estimation as a *loss* (cost) function in the search for the optimal predictive density matrix in a Bayesian (global) approach. In

other words, the best estimate is found by minimizing the average Kullback entropy with respect to the true state. In the same perspective, the Kullback entropy has also been used as a regularizing functional in seeking solutions to multivariable and multidimensional moment problems [21].

Finally, notice that a symmetrized version of the Kullback entropy has also been suggested [22] and may be employed to assess the distance between quantum states.

III. MINIMUM KULLBACK PRINCIPLE FOR STATE ESTIMATION

In this section, we exploit the MKE principle for estimation of the full density matrix ϱ from an incomplete set of measurements. We assume that the system under investigation has a bias toward the known prior τ, and first consider that we have access to the mean value of a single observable. Then the analysis is generalized to the case of N observables, with application to the measurement of the complete distribution of a single observable.

A. Measurement of a single observable

In this section we assume that only the mean value of the observable A can be measured. This is the simplest observation level [23] that one may devise for a quantum system, and is generally considered to provide only a small amount of information about the state under investigation. Upon applying the quantum MKE principle, the best estimate for the density matrix compatible with the bias is given by Eq. (12). By introducing the partition function $Z = \mathrm{Tr}(\tau e^{-A\lambda})$ the estimated density matrix reads as follows:

$$\varrho = \frac{1}{Z} e^{-A\lambda/2} \tau e^{-A\lambda/2}. \tag{13}$$

Moreover, using the spectral decomposition $A = \Sigma_k \alpha_k |\varphi_k\rangle\langle\varphi_k|$, $\{|\varphi_k\rangle\}$ being a complete orthonormal system of eigenvectors of the operator A, we can write: $\exp\{-\frac{1}{2}A\lambda\} = \Sigma_k \exp\{-\frac{1}{2}\alpha_k\lambda\}|\varphi_k\rangle\langle\varphi_k|$, so that

$$Z = \sum_k e^{-\alpha_k\lambda}\langle\varphi_k|\tau|\varphi_k\rangle, \tag{14}$$

$$\langle\varphi_m|\varrho|\varphi_n\rangle = \frac{1}{Z} e^{-(\alpha_m+\alpha_n)\lambda/2}\langle\varphi_m|\tau|\varphi_n\rangle. \tag{15}$$

In this way, given the initial density matrix τ, it is possible to estimate the complete state ϱ as follows:

$$\varrho = \frac{1}{Z}\sum_{n,m}\langle\varphi_m|\tau|\varphi_n\rangle e^{-(\alpha_n+\alpha_m)\lambda/2}|\varphi_m\rangle\langle\varphi_n|, \tag{16}$$

where the value of the Lagrange multiplier λ is obtained as a (numerical) solution of the equation $\mathrm{Tr}(\varrho A) = \langle A \rangle$, i.e.,

$$\langle A \rangle = \frac{\sum_k \langle\varphi_k|\tau|\varphi_k\rangle e^{-\alpha_k\lambda}\alpha_k}{\sum_k \langle\varphi_k|\tau|\varphi_k\rangle e^{-\alpha_k\lambda}}, \tag{17}$$

which, of course, is equivalent to the relation $-\partial_\lambda \log Z = \langle A \rangle$. It is worth noting that the estimate (16) for the density

matrix has support in the same subspace of Hilbert space as does τ. Moreover, if $\tau = |\varphi_{\bar{n}}\rangle\langle\varphi_{\bar{n}}|$, i.e., if the prior density matrix is a projector onto the subspace generated by the eigenvector $|\varphi_{\bar{n}}\rangle$ of A, then $\langle\varphi_n|\tau|\varphi_n\rangle = \delta_{n\bar{n}}$, and the MKE principle reduces to the MaxEnt one, as in the classical case.

B. Measurement of N observables

Here we assume that the mean values of N different observables A_k, $k = 1, \ldots, N$, are experimentally accessible. This means that we have N constraints $\langle A_k\rangle$ to be considered. Upon writing the trajectory (11) for each constraint, we have a system of N differential equations which can be written in the following compact form:

$$\sum_{k=1}^{N} \frac{d\varrho}{d\lambda_k} = -\frac{1}{2}\left\{\varrho, \sum_{k=1}^{N}(A_k - \langle A_k\rangle)\right\}. \quad (18)$$

The solution can be written as

$$\varrho = \frac{1}{Z}\exp\left(-\frac{1}{2}\sum_k A_k\lambda_k\right)\tau\exp\left(-\frac{1}{2}\sum_k A_k\lambda_k\right), \quad (19)$$

where $\lambda_1, \ldots, \lambda_N$ are Lagrange multipliers, and we assumed $\varrho(0) = \tau$. The partition function Z reads

$$Z = \text{Tr}\left[\tau\exp\left(-\sum_k A_k\lambda_k\right)\right]. \quad (20)$$

Again, the values of the multipliers λ_k are obtained by solving the system of equations

$$\text{Tr}(\varrho A_k) = \langle A_k\rangle, \quad k = 1, \ldots, N. \quad (21)$$

The above analysis allows us to apply the MKE principle also starting from the measurement of the full distribution of an observable. In this case, the set of observables to be taken into account are the orthogonal (commuting) eigenprojectors $A_k = |\varphi_k\rangle\langle\varphi_k|$, $\langle\varphi_k|\varphi_s\rangle = \delta_{ks}$ of the measured observables. The constraints $\text{Tr}(\varrho A_k) = p_k$, correspond to the measured distribution. Equations (20) and (21) can be rewritten as

$$Z = \sum_k e^{-\lambda_k}\langle\varphi_k|\tau|\varphi_k\rangle, \quad (22)$$

$$p_k = \frac{1}{Z}e^{-\lambda_k}\langle\varphi_k|\tau|\varphi_k\rangle. \quad (23)$$

Finally, taking matrix elements of Eq. (19) and using Eqs. (22) and (23) it is possible to reconstruct the posterior state, given the initial density matrix τ and the measured probabilities p_k,

$$\varrho = \sum_{n,m} \frac{\langle\varphi_m|\tau|\varphi_n\rangle}{\sqrt{\langle\varphi_m|\tau|\varphi_m\rangle\langle\varphi_n|\tau|\varphi_n\rangle}}\sqrt{p_m p_n}|\varphi_m\rangle\langle\varphi_n|. \quad (24)$$

Notice that in this case we have been able to back-substitute the Lagrange multipliers, i.e., Eq. (24) no longer depends on the λ_k's.

C. The qubit case

Here we address the estimation of a qubit state starting from the measurement of a single observable [26]. In order

to apply the MKE principle, we assume three is a bias toward the state τ and choose an observable to measure in the system. The measured quantity is the spin along direction $\vec{n}$, which is described by the operator

$$A = \vec{n} \cdot \boldsymbol{\sigma}, \quad (25)$$

where we defined the vector $\vec{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)$, $\boldsymbol{\sigma}_k$, $k = 1, 2, 3$, being the Pauli matrices. Upon writing the prior state in the Pauli basis

$$\tau = \frac{1}{2}(1 + \vec{\tau} \cdot \vec{\boldsymbol{\sigma}}), \quad |\vec{\tau}| \leq 1, \quad (26)$$

Eq. (13) reads

$$\varrho = \frac{1}{2}(1 + \vec{v} \cdot \vec{\boldsymbol{\sigma}}), \quad (27)$$

with

$$\vec{v} = \frac{\vec{\tau} + 2\sinh^2(\lambda/2)(\vec{n} \cdot \vec{\tau})\vec{n} - \sinh\lambda\vec{n}}{\cosh\lambda - \vec{\tau} \cdot \vec{n}\sinh\lambda}, \quad (28)$$

where we used $Z = \cosh\lambda - \vec{\tau} \cdot \vec{n}\sinh\lambda$. Now, thanks to the constraint

$$\text{Tr}(\varrho\vec{n} \cdot \vec{\boldsymbol{\sigma}}) = \langle\vec{n} \cdot \vec{\boldsymbol{\sigma}}\rangle, \quad (29)$$

we can calculate the value of the Lagrange multiplier λ, obtaining

$$\lambda = \text{arctanh}\frac{\vec{\tau} \cdot \vec{n} - \langle\vec{n} \cdot \vec{\boldsymbol{\sigma}}\rangle}{1 - \langle\vec{n} \cdot \vec{\boldsymbol{\sigma}}\rangle(\vec{\tau} \cdot \vec{n})}. \quad (30)$$

In order to have a compact expression of the result let us consider an operator basis composed of spin operators along three orthogonal directions $\vec{n}_1 \perp \vec{n}_2 \perp \vec{n}_3$, with $\vec{n}_1 \equiv \vec{n}$. In this way we can express the components of the vector $\vec{v}$ as follows:

$$\vec{v} \cdot \vec{n}_1 = \langle\vec{n} \cdot \vec{\boldsymbol{\sigma}}\rangle, \quad (31)$$

$$\vec{v} \cdot \vec{n}_k = \vec{\tau} \cdot \vec{n}_k\sqrt{\frac{1 - \langle\vec{n} \cdot \vec{\boldsymbol{\sigma}}\rangle^2}{1 - (\vec{\tau} \cdot \vec{n})^2}} \quad (k = 2, 3). \quad (32)$$

Equation (31) says that the estimated Bloch component in the direction of the measured observable is equal to the measured mean value, whereas the two other orthogonal components are obtained from the prior one by a common shrinking factor.

As an example, let us assume $\vec{n} = (1, 0, 0)$ and $\vec{\tau} = (0, 0, 1)$: this is the case of the measurement of $\boldsymbol{\sigma}_1$ (spin along the x direction) and bias toward the $+z$ direction. We have $x = -\langle\boldsymbol{\sigma}_1\rangle$ and thus

$$\vec{v} = (\langle\boldsymbol{\sigma}_1\rangle, 0, \sqrt{1 - \langle\boldsymbol{\sigma}_1\rangle^2}), \quad (33)$$

which satisfies Eqs. (31) and (32).

D. The harmonic oscillator case

Now we face the problem of estimating the state of a harmonic oscillator with a bias toward a coherent state τ

$= |\alpha\rangle\langle\alpha|$, with $\alpha \in \mathbb{C}$. Let us consider a photon number measurement, i.e., if a denotes the annihilation operator, $[a, a^{\dagger}] = 1$, the observable expressed by $A \equiv a^{\dagger}a = \Sigma_n n|n\rangle\langle n|$, $\{|n\rangle\}$ being the photon number basis. Since

$$\langle n|\tau|m\rangle = \frac{\alpha^n \bar{\alpha}^m e^{-|\alpha^2|}}{\sqrt{n!m!}}, \tag{34}$$

we have that the MKE estimated state is still a coherent state with amplitude

$$\beta = \alpha e^{-\lambda/2}. \tag{35}$$

Now, using the constraint

$$\text{Tr}(\varrho a^{\dagger}a) = N, \tag{36}$$

with N obtained from the experiment, we can evaluate the value of the Lagrange multiplier λ, namely,

$$\lambda = \log(|\alpha|^2/N), \tag{37}$$

also obtaining $Z = \exp\{N\}$. Finally, upon substituting in Eq. (16) we arrive at

$$\varrho = e^{-N}\sum_{nm} (N/|\alpha|^2)^{(n+m)/2}\frac{\alpha^n\bar{\alpha}^m}{\sqrt{n!m!}} \equiv |\sqrt{N}e^{i\phi}\rangle\langle\sqrt{N}e^{i\phi}|, \tag{38}$$

with $\phi = \arg \alpha$. In other words, the best estimate according to MKE is a coherent state with average number of photons equal to the measured one and phase equal to that of the prior coherent state. Notice that the best estimate obtained using the MaxEnt principle with the same constraint on the average number of photons, but without the bias, would have been a thermal state with N thermal photons.

If by some means the complete photon distribution p_n is available, the reconstructed state, given by Eq. (24), reads as follows (we assumed that the bias is still toward $\tau = |\alpha\rangle\langle\alpha|$, $\alpha \in \mathbb{C}$):

$$\varrho = \sum_{n,m} \sqrt{p_n p_m} e^{i\phi(n-m)}|n\rangle\langle m|. \tag{39}$$

Remarkably, Eq. (39) no longer depends on the amplitude of the prior state, which enters only through the phase ϕ. This makes the above scheme quite promising though measuring the photon distribution is, in general, a challenging task. On the other hand, in the optical case it is possible to reconstruct the p_n by means of on-off photodetection and the maximum likelihood algorithm [27,28], a method that has been recently verified in the laboratory [29,30]. A multichannel fiber loop detector [31,32] may also be used.

Finally, we notice that a special case of bias is that toward a Gaussian state. In fact the Kullback relative entropy of a state ϱ with respect to a Gaussian state τ, with the same covariance matrix reduces to the difference of the von Neumann entropies [25]

$$K(\varrho|\tau) = H(\tau) - H(\varrho)$$

and thus the MKE principle reduces to the MaxEnt. Notice, however, that this is not in contrast with the results above, since in that case the results of the measurement do not impose the equality of the covariance matrices.

IV. MKE ESTIMATION OF WEAK HAMILTONIANS

In the previous section, we have shown how to fully reconstruct the density matrix of a quantum system from incomplete data. Here we will see how it is possible to estimate a weak Hamiltonian H by means of the MKE and suitable measurements on the evolved states. The idea behind this method is that, when considering weak Hamiltonian processes, the evolved state is not too different from the initial one, i.e., it has a natural bias toward the unperturbed state. This allows one to estimate the parameters (matrix elements) of the Hamiltonian from data obtained by an incomplete set of measurements on the evolved state, i.e., to use the MKE principle as an effective tool for process estimation.

A. The qubit case

The initial state τ and the evolved state of the qubit system under investigation are connected by the transformation

$$\varrho_t = e^{-iHt}\tau e^{iHt}. \tag{40}$$

In. Eq. (40) t is the time evolution. Using the Pauli basis, we can express the initial state and the Hamiltonian as

$$\tau = \frac{1}{2}(\mathbb{1} + \vec{\tau}\cdot\vec{\boldsymbol{\sigma}}), \quad |\vec{\tau}| \le 1, \tag{41}$$

and

$$H = \sum_{\nu=0}^{3} h_\nu\boldsymbol{\sigma}_\nu, \quad |\vec{h}| \ll 1, \tag{42}$$

respectively, where $\boldsymbol{\sigma}_\nu$ are the Pauli matrices with $\boldsymbol{\sigma}_0 = \mathbb{1}$, and $\vec{\boldsymbol{\sigma}}$ has been defined above. Expanding Eq. (40) at the first order in $\vec{h}$, we obtain

$$\varrho_t = \tau + it[\tau, H] + o(|\vec{h}|^2), \tag{43}$$

with

$$[\tau, H] = i\sum_{k,s=1}^{3} \tau_s h_k \varepsilon_{ksl}\boldsymbol{\sigma}_l, \tag{44}$$

ε_{ksl} being the totally antisymmetric tensor, $\varepsilon_{123} = 1$. In this way, the expansion (43) can be written as follows:

$$\varrho_t = \frac{1}{2}(\mathbb{1} + \vec{w}\cdot\vec{\boldsymbol{\sigma}}), \tag{45}$$

where the Bloch vector is given by

$$\vec{w} \equiv \vec{\tau} + 2\vec{h}\times\vec{\tau}. \tag{46}$$

Equation (46) represents a system of equations for the unknowns $\vec{h}$. The transfer matrix is singular, but the system may be inverted anyway using the Moore-Penrose generalized inverse [24], thus leading to the expression

$$\vec{h} = \frac{\vec{\tau}\times\vec{w}}{2|\vec{\tau}|^2}, \tag{47}$$

which provides an estimate for the Hamiltonian $\vec{h}$ once an estimate for $\vec{w}$ is given. The latter is obtained upon the mea-

surement of a spin observable described by the operator $\vec{n} \cdot \vec{\boldsymbol{\sigma}}$ on the evolved state ϱ_t. The best estimate for $\vec{w}$ according to the MKE principle is given by Eq. (28). By substituting in Eq. (47) and using Eq. (30), an estimate for $\vec{h}$ is achieved:

$$\vec{h} = \frac{1}{2|\vec{\tau}|^2} \frac{(1 - \sqrt{1 - \kappa^2})\vec{\tau} \cdot \vec{n} - \kappa}{1 - \vec{\tau} \cdot \vec{n}\kappa} \vec{\tau} \times \vec{n} \qquad (48)$$

where

$$\kappa = \frac{\vec{\tau} \cdot \vec{n} - \langle \vec{n} \cdot \vec{\boldsymbol{\sigma}} \rangle}{1 - \langle \vec{n} \cdot \vec{\boldsymbol{\sigma}} \rangle (\vec{\tau} \cdot \vec{n})}. \qquad (49)$$

As it is apparent from Eq. (48) the MKE estimate for the Hamiltonian Bloch vector is orthogonal to both that of the prior state and the measurement direction. By a repeated randomized choice of the measurement an effective reconstruction may be achieved for any weak, qubit Hamiltonian. Generalizations to higher-dimensional systems may also be designed.

B. The harmonic oscillator case

Let us now assume that the expansion

$$\varrho(t) = \tau + it[\tau, H] + o(|H|^2) \qquad (50)$$

refers to the state of a harmonic oscillator evolving under the action of a weak Hamiltonian H that we want to estimate. We also assume that the full distribution of a single observable can be measured in the evolved state. Therefore, the evolved density matrix ϱ_t may be reconstructed by MKE principle, starting from the observation level $A_k = |\varphi_k\rangle\langle\varphi_k|$, as described in Sec. III B, and thus obtaining the state (24). Using the basis $\{|\varphi_k\rangle\}$, we can write Eq. (50) as follows:

$$\varrho_{mn}(t) = \tau_{mn} + it \sum_s (H_{ms}\tau_{sn} + \tau_{ms}H_{sn}), \qquad (51)$$

where we used $H_{mn} \equiv \langle\varphi_m|H|\varphi_n\rangle$. Using the MKE estimate (24) for the evolved density matrix, we obtain the following hierarchy of equations:

$$\sum_s (H_{ms}\tau_{sn} + \tau_{ms}H_{sn}) = \frac{i}{t}\tau_{mn}\left(1 - \sqrt{\frac{p_n p_m}{\langle\varphi_n|\tau|\varphi_n\rangle\langle\varphi_m|\tau|\varphi_m\rangle}}\right) \qquad (52)$$

where p_n are the measured probabilities (the constraints used for the MKE) and the matrix elements H_{nm} are the unknowns.

A relevant example, in which the Hamiltonian can be effectively estimated using the MKE, is that corresponding to $H = (ga + \text{H.c.})$, a being the annihilation operator starting from the sole measurement of the photon distributions. The evolution imposed by the Hamiltonian H corresponds to the unitary displacement operator $D(\beta) = \exp(\beta a^\dagger - \beta^* a)$, $\beta = gt$. The problem is then to estimate the displacement amplitude β from the measured photon distribution. We assume that the initial state is a coherent state $\varrho(\alpha) = |\alpha\rangle\langle\alpha|$. For the sake of simplicity we take α and β as real. Using the photon number basis the evolved state may be written as

$$\varrho(\beta) = e^{(\alpha + \beta)^2} \sum_{n,m} \frac{(\alpha + \beta)^{n+m}}{\sqrt{n!m!}} |n\rangle\langle m|. \qquad (53)$$

Assuming that the measurement of the photon number is made on the evolved state and that the MKE principle is used to estimate the density matrix we equate the above expression to that given in Eq. (39) for $\phi = 0$ (recall that α and β are taken as real). We thus obtain the following set of equations:

$$-(\alpha + \beta)^2 + (n + m)\ln(\alpha + \beta) = \ln \sqrt{n!m!p_n p_m}, \qquad (54)$$

to be solved for β. It is worth noting that in order to estimate β one can choose to measure a *finite* number of p_k, i.e., $k = 0, \ldots, N-1$. As a matter of fact, this choice also selects a subspace of the Hilbert space where the reconstructed state is defined. In turn, Eq. (54) corresponds to N^2 determinations of the same parameter β. Notice that without using the MKE principle the only way to exploit the information at disposal, i.e., the elements of the probability distribution $p_n = e^{-(\alpha + \beta)^2}(\alpha + \beta)^{2n}/n!$, $n = 0, \ldots, N-1$, is to invert those relations. In order to estimate β one should solve the set of equations

$$-(\alpha + \beta)^2 + 2n \ln(\alpha + \beta) = \ln(n! p_n), \qquad (55)$$

which provide only N determinations of β.

V. CONCLUSIONS

In this paper we have considered quantum estimation of states and weak Hamiltonian operations in situations where one has at disposal data from the measurement of an incomplete set of observables and, at the same time, some *a priori* information on the state itself. By expressing the *a priori* information in terms of a bias toward a given state, the best estimate is obtained using the principle of minimum Kullback entropy, i.e., by taking the state that reproduces the data while minimizing the relative entropy with respect to the bias. The MKE principle has been used to estimate the quantum state from the measurement of a single observable, either from the sole mean value or from the complete probability distribution. In particular, we have analyzed qubit and harmonic systems with some detail. We have also considered the problem of estimating a weak Hamiltonian process. In this case there is a natural bias of the evolved state toward the initial state, and the MKE principle can be used as a tool to estimate the Hamiltonian from an incomplete set of measurements.

Overall, the minimum Kullback entropy principle appears to be a convenient approach for quantum estimation inrealistic situations and a useful tool for the estimation of weak Hamiltonian processes.

ACKNOWLEDGMENTS

[1] M. G. A. Paris and J. Řeháček, in *Quantum State Estimation*, edited by M. G. A. Paris and J. Rehacek, Lecture Notes in Physics Vol. 649 (Springer, Berlin, 2004).

[2] G. M. D'Ariano, L. Maccone, and M. G. A. Paris, J. Phys. A **34**, 93 (2001).

[3] V. Buzek *et al.*, Chaos, Solitons Fractals **10**, 981 (1999); V. Bužek, in *Quantum State Estimation* (Ref. [1]), P. 189.

[4] E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).

[5] H. Umegaki, Kodai Math. Semin. Rep. **14**, 59 (1962); G. Linblad, Commun. Math. Phys. **33**, 305 (1973).

[6] V. Vedral, Rev. Mod. Phys. **74**, 197 (2002).

[7] E. T. Jaynes, IEEE Trans. Syst. Sci. Cybern. **SSC-4**, 227 (1968).

[8] S. Kullback, Ann. Math. Stat. **22**, 79 (1951).

[9] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).

[10] M. J. Donald, Commun. Math. Phys. **105**, 13 (1986); Found. Phys. **22**, 1111 (1992).

[11] G. L. Bilbro, W. E. Snyder, and R. C. Mann, J. Opt. Soc. Am. A **8**, 290 (1991).

[12] A. D. Woodbury and T. J. Ulrych, Water Resour. Res. **29**, 2847 (1993); **32**, 2671 (1996).

[13] G. U. Schreiber *et al.*, Phys. Lett. B **284**, 6 (1992); A. R. Plastino, H. G. Miller, and A. Plastino, Phys. Rev. E **56**, 3927 (1997).

[14] S. L. Braunstein, Phys. Lett. A **219**, 169 (1996).

[15] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[16] P. B. Slater, Phys. Lett. A **247**, 1 (1998).

[17] R. Schack, T. A. Brun, and C. M. Caves, Phys. Rev. A **64**, 014305 (2001).

[18] Z. Hradil, Phys. Rev. A **55**, R1561 (1997); Z. Hradil, J. Rehacek, J. Fiurasek, and M. Jezek, in *Quantum State Estimation* (Ref. [1]), p. 59.

[19] Z. Hradil and J. Summhammer, J. Phys. A **33**, 7607 (2000).

[20] F. Tanaka and F. Komaki, Phys. Rev. A **71**, 052323 (2005).

[21] T. T. Georgiou, IEEE Trans. Inf. Theory **52**, 1052 (2006).

[22] A. P. Majtey, P. W. Lamberti, and D. P. Prato, Phys. Rev. A **72**, 052310 (2005); H. Muhlenbein and R. Hons, Evol. Comput. **13**, 1 (2005).

[23] V. Bužek *et al.*, Ann. Phys. (N.Y.) **266**, 454 (1998).

[24] E. H. Moore, Bull. Am. Math. Soc. **26**, 394 (1920); R. Penrose, Proc. Cambridge Philos. Soc. **51**, 406 (1955).

[25] J. Eisert and M. M. Wolf, e-print arXiv:quant-ph/0505151.

[26] Since the qubit Hilbert space is two-dimensional, the measurements of the mean value and of the complete probability distribution give the same amount of information.

[27] D. Mogilevtsev, Opt. Commun. **156**, 307 (1998); Acta Phys. Slov. **49**, 743 (1999).

[28] A. R. Rossi, S. Olivares, and M. G. A. Paris, Phys. Rev. A **70**, 055801 (2004); A. R. Rossi and M. G. A. Paris, Eur. Phys. J. D **32**, 223 (2005).

[29] G. Zambra, A. Andreoni, M. Bondani, M. Gramegna, M. Genovese, G. Brida, A. Rossi, and M. G. A. Paris, Phys. Rev. Lett. **95**, 063602 (2005).

[30] M. Gramegna *et al.*, Laser Phys. **16**, 385 (2006).

[31] J. Řeháček, Z. Hradil, O. Haderka, J. Peřina, Jr., and M. Hamar, Phys. Rev. A **67**, 061801(R) (2003); O. Haderka, M. Hamar, and J. Peřina, Eur. Phys. J. D **28**, 149 (2004).

[32] K. Banaszek and I. A. Walmsley, Opt. Lett. **28**, 52 (2003).