BAYESIAN QUBIT TOMOGRAPHY

Attila Magyar,
Katalin M. Hangos

Process Control Research Group
HAS Computer and Automation Research Institute
Budapest, Hungary

Abstract: In the paper the Bayesian and the least squares methods of quantum state estimation are compared for a single qubit. The quality of the estimates is compared by computer simulation when the true state is either mixed or pure. The fidelity is used to quantify the estimation error. Both methods are sensitive to the degree of the purity of the state to be estimated, that is, their performance can be quite bad near pure states.

Keywords: system identification, quantum systems, Bayesian estimation, least squares method

1. INTRODUCTION

Quantum mechanics is one of the most interesting fields in modern physics. In spite of its great importance related to quantum computers and quantum information theory, just a few persons have tried to apply the tools of advanced systems and control theory to quantum systems. Due to some similarities with X-ray tomography, the state reconstruction is often called quantum tomography.

The methods of *classical statistical estimation* are used to develop state estimation of quantum systems in the first group of papers (D'Ariano, *et al*., 2003; Řehaček, *et al.*, 2004). This approach suffers from the fact that the state estimation is usually based on a few types of measurements (observables) that are incompatible, thus there is no joint probability density function of the measurement results in the classical sense.

The other way of computing a point estimate of the state of a quantum system is to use *convex optimization* methods such as in (Kosut, *et al*., 2004). Here one can respect the constraints imposed on the components of the state but there is no information on the probability distribution of the estimate.

The aim of this paper is to work out a type of Bayesian quantum state estimation as a statistical method that respects the constraints on Bloch vectors and compare it with the least squares (LS) method as an optimization-based method by using simulation experiments.

2. BASIC NOTIONS

The simplest system in quantum mechanics is the two-level system that represents a *spin½ particle*, a quantum bit, or shortly a *qubit*. We illustrate the notions on this example, as the quantum state estimation methods will also be developed for this simple quantum system.

2.1 State representation of quantum systems

The state of a quantum system is described by *density matrices* which are statistical operators (positive operators of trace 1) acting on the underlying Hilbert space. The Hilbert space associated to a qubit is $\mathbb{C}^2$ with the usual inner product. The Pauli matrices

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1}$$

form a basis among the self-adjoint operators over $\mathbb{C}^2$. A density matrix written up in the above basis (1) gives

$$\rho = \frac{1}{2}\left(I + s_1\sigma_1 + s_2\sigma_2 + s_3\sigma_3\right), \tag{2}$$

$$s_1^2 + s_2^2 + s_3^2 \le 1. \tag{3}$$

Therefore, one can represent the state of this system by a 3-dimensional real vector (a so called *Bloch vector*):

$$s = \left[s_1, s_2, s_3\right]^{\mathrm{T}} \tag{4}$$

with length less than or equal to *1*. The algebraic constraint (3) is equivalent to the fact that ρ is positive semidefinite. The state space of this system is the unit sphere, which is also called *Bloch ball.*

2.2 Observables

To each dynamical variable one associates a self-adjoint operator $A = A^*$ of the underlying Hilbert space, that is called an *observable*. Because of its self-adjoint property, the operator admits a spectral decomposition:

$$A = \sum_{i=1}^{n} \lambda_i P_i \quad \lambda_j \ne \lambda_k,\ j \ne k, \tag{5}$$

where λ_i are the different eigenvalues and P_i are pairwise orthogonal eigenprojections. Measuring the above observable the possible outcomes are λ_i, $i = 1,\ldots,n$ with the following probabilities

$$P(\lambda_i) = \mathrm{Tr}\ \rho P_i \tag{6}$$

supposed that the system is in state ρ. It can be shown that

$$\sum_i P(\lambda_i) = 1, \tag{7}$$

so the different outcomes of the measurement form a probability distribution. The mean value of the measurement for the observable A is

$$\langle A \rangle = \sum_{i=1}^{n} \lambda_i Tr(\rho P_i) = Tr(\rho A). \tag{8}$$

3. QUANTUM STATE ESTIMATION

For the state estimation, we will consider *3n* identical copies of qubits being in the state ρ. We measure all three Pauli spin matrices $\{\sigma_1, \sigma_2, \sigma_3\}$ on all n copies. The possible outcomes for each of these single measurements, i.e. the eigenvalues of σ_i, are ± 1 and the corresponding spectral projections are given by

$$P_i^{\pm} = \frac{1}{2}(I \pm \sigma_i) \tag{9}$$

For the sake of definiteness, we assume that first σ_1 is measured n times, then σ_2 and then σ_3. The data set of the outcomes of this measurement scheme consists of three strings of length n with entries ± 1:

$$D_i^n = \{D_i^n(j) : j = 1,\ldots,n\} \quad i = 1,2,3. \tag{10}$$

The predicted probabilities of the outcomes depend on the true state ρ of the system and they are given by

$$\Pr ob(D_i^n(j) = 1) = Tr(\rho P_i^+) = \frac{1}{2}\left(1 + \langle \sigma_i \rangle_\rho\right) = \frac{1}{2}(1 + s_i). \tag{11}$$

2.1 Least Squares estimation

Let $\pi_i(\pm)$ be the relative frequency of ± 1 in the string D_i^n, then the difference

$$\pi_i := \pi_i(+) - \pi_i(-) \tag{12}$$

is an estimate of the *i*-th spin component s_i $(i = 1, 2, 3)$. As a measure of unfit (estimation error) we use the *Hilbert-Schmidt norm* of the difference between the empirical and the predicted data according to the least squares (LS) principal. (Note that in this case the Hilbert-Schmidt norm is simply the Euclidean distance in the 3-space.) Then the following loss function is defined:

$$L(\omega) = d^2(s,\pi) = \sum_{j=1}^{3}(s_j - \pi_j)^2 = \|s\|^2 + \|\pi\|^2 - 2s\pi, \tag{13}$$

where s is the Bloch vector of the density operator ω.

An estimate of the unknown parameter vector $s = [s_1, s_2, s_3]^T$ is obtained by solving the following constrained quadratic optimization problem:

$$\begin{aligned} &Minimize \quad L(\omega) \\ &s.t. \qquad\quad \|s\| \le 1. \end{aligned} \tag{14}$$

The above loss function is simple so we can solve the constrained minimization problem explicitly. In the unconstrained minimization, two cases are possible. First, $||\pi|| \le 1$, and in this case the constrained minimum is taken at $s = \pi$. When the unconstrained minimum is at π with $||\pi|| > 1$, then it is clear from the 3-dimensional geometry that the constrained minimum is taken at

$$s = \frac{\pi}{\|\pi\|}. \tag{15}$$

2.1 Bayesian quantum state estimation

In the Bayesian parameter estimation, the parameters θ to be estimated are considered as random variables. The probability $P(\theta \mid D^n)$ of a specific value of the parameters conditioned on the measured data D^n is evaluated. Afterwards, the mean value of this distribution can be used as the estimate. If the measured data is a sequence of outcomes, as in our case, it can be split into the latest outcome $D^n(n)$ of D^n and D^{n-1}, the preceding. Then the conditional distribution of the parameter becomes $P(\theta \mid D^n(n), D^{n-1})$ and the Bayes formula

$$P(a \mid b,c) = \frac{P(b \mid a,c)P(a \mid c)}{\int P(b \mid \nu,c)P(\nu \mid c)d\nu} \tag{16}$$

can be applied resulting in the following recursive formula for $P(\theta|D^n)$

$$P(\theta \mid D^n) = \frac{P(D^n(n) \mid D^{n-1},\theta)P(\theta \mid D^{n-1})}{\int P(D^n(n) \mid D^{n-1},\nu)P(\nu \mid D^{n-1})d\nu}. \tag{17}$$

In our state estimation, we have three data sets D_i^n, $i = 1, 2, 3$, corresponding to the three directions. The estimation is performed for the three directions independently (afterwards a conditioning has to be made).

The probabilities $P(D_i^n(n) \mid D_i^{n-1},\theta)$ have the form

$$P(D_i^n(n) \mid D_i^{n-1}, s_i) = P(\pm 1 \mid s_i) = \frac{1}{2}Tr\rho(1 \pm \sigma_i) = \frac{1}{2}(1 \pm s_i). \tag{18}$$

If we denote by $\ell(i)$ the number of $+1$'s in the data string D_i^n , then (17) becomes

$$P(s_i \mid D_i^n)(t) = \frac{\left(\frac{1}{2}(1+t)\right)^{\ell(i)} \cdot \left(\frac{1}{2}(1-t)\right)^{n-\ell(i)} \cdot P_i^0(t)}{\int \left(\frac{1}{2}(1+\nu)\right)^{\ell(i)} \cdot \left(\frac{1}{2}(1-\nu)\right)^{n-\ell(i)} \cdot P_i^0(\nu)d\nu}, \tag{19}$$

where $P_i^0(v)$ is an assumed prior distribution, with which the recursive estimation is started. For the sake of simplicity we assume that $P_i^0(v)$ has the same form with parameters κ and λ in place of n and ℓ, respectively.

After a parameter transformation, it appears to be a *β-distribution*

$$P(s_i \mid D_i^n)(u) = C\left(\frac{1+u}{2}\right)^{\ell(i)+\lambda}\left(\frac{1-u}{2}\right)^{n+\kappa-\ell(i)-\lambda}, \tag{20}$$

where C is the normalization constant and $u \in [0, 1]$. It is well known that the mean value of this distribution is

$$m_i = \frac{\ell(i)+1+\lambda}{n+\kappa+2}, \tag{21}$$

and its variance is

$$\frac{(\ell(i)+1+\lambda)(n-\ell(i)+1+\kappa-\lambda)}{(n+\kappa+2)^2(n+\kappa+3)}. \tag{22}$$

The above statistics (21, 22) is suited to construct an unbiased estimate for s_i, in the form

$$\hat{s}_i = 2\frac{\ell(i)+1+\lambda}{n+\kappa+2} - 1 \tag{23}$$

after the re-transformation of the variables.

Since the components of the Bloch vector are estimated independently, the constraint (3) has not been taken into account yet. Thus, a further step of conditioning is necessary. We simply condition $(\hat{s}_1, \hat{s}_2, \hat{s}_3)$ to (3):

$$\hat{m}_i = \frac{\iiint u_i f(u_1)f(u_2)f(u_3)du_1du_2du_3}{\iiint f(u_1)f(u_2)f(u_3)du_1du_2du_3}, \tag{24}$$

where both integrals are over the domain $\{(u_1,u_2,u_3): u_1^2+u_2^2+u_3^2 \leq 1\}$ and $f(u_i) = P(s_i|D_i^n)(u_i)$. Then the conditioned estimate of s_i will be

$$2\hat{m}_i - 1.$$

3. EXPERIMENTS

The aim of the following experiments is to compare the properties of the two methods described above. The base data of the estimation is obtained by measuring spin components σ_1, σ_2, and σ_3 of several qubits being in the same state (i.e. having the same Bloch vectors). The Bayesian method was applied with and without regularization to analyze its influence. The measurements were performed on a quantum simulator implemented in MATLAB for two level systems. An experiment setup consisted of a Bloch vector s to be estimated and a number (n) of spin measurements performed on the quantum system. Each experiment setup was repeated five times and the performance indicator quantity (the *fidelity* of the estimation error) was averaged.
The fidelity of two density operators ρ and ω is defined as

$$F(\rho,\omega) = Tr\sqrt{\rho^{1/2}\omega\rho^{1/2}}. \tag{25}$$

3.1 Number of measurements

The first set of experiments was to investigate the dependence between the fidelity (25) and the number of measurements n. It was expected that the fidelity goes to 1 when n goes to infinity. Fig. 1 shows the experimental results for estimating a pure state $s_{pure} = [0.5774, 0.5774, 0.5774]^T$. The result of the Bayesian estimation (dotted line) shows the weakest performance because of the conditioning feature of the method. The price of the validity of the Bayesian method with conditioning is the precision for (near) pure states. It is apparent that the least squares estimation does not have the above problem.

The situation is different for estimating mixed states ($s_{mixed} = [0.3, -0.4, 0.3]^T$). It can be seen in Fig. 2 that the two kinds of Bayesian estimation differ only for small n's. Least squares method also works a little bit better for mixed states than for pure states, at least for larger n's. It is apparent that pure states are a challenge for both methods but least squares handles this difficulty a bit better.

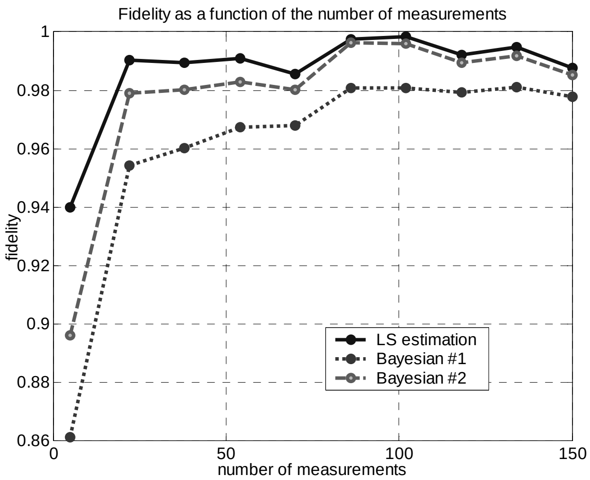

Figure 1: pure state

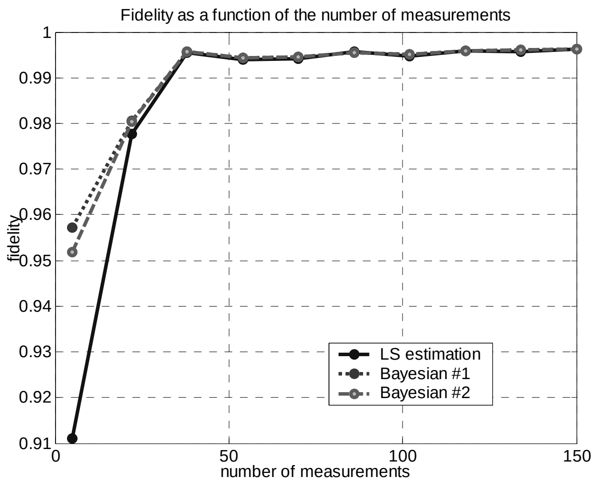

Figure 2: mixed state

3.2 Bloch vector length

During the second set of experiments, the length of the Bloch vector was varying. Its direction was $s = [0.5774,\ 0.5774,\ 0.5774]^T$. The expectation was that the fidelity would be relatively independent of the Bloch vector length $||s||$. The experiment results are plotted in Fig. 3 and Fig. 4. The first picture shows the case $n = 100$, where, in spite of the big variance, the conditioned Bayesian shows an increase near the pure state ($||s|| = 1$). At $n = 900$ (shown in Fig. 4) it is more apparent that LS and conditioned Bayesian methods (both have certain conditioning feature to avoid faulty estimates near $||s|| = 1$) have worse performance near pure states.

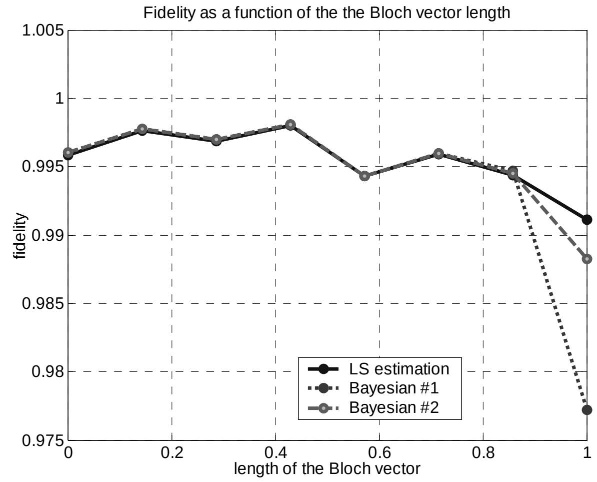

Figure 3: n=100

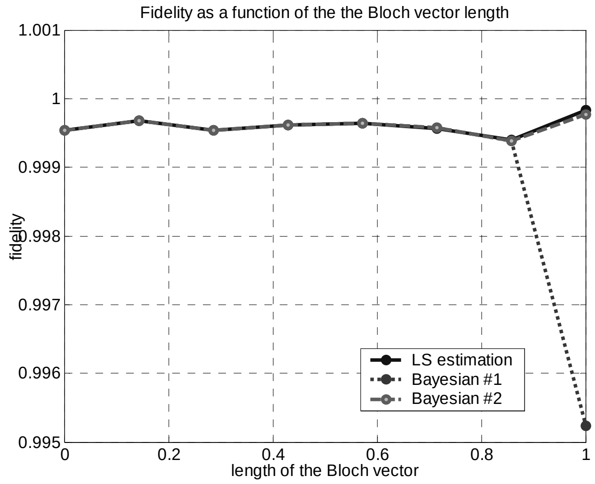

Figure 4: n=900

3. CONCLUSIONS

The performance of two state estimation methods, the Bayesian state estimation as a statistical method and the least squares (LS) method as an optimization-based method is investigated in this paper by using simulation experiments. The fidelity is used as a performance indicator quantity.

The investigated methods were found to be quite sensitive to the length of the Bloch vector, i.e. to the fact if a pure or mixed state was the one to be estimated. The methods that are not informed about the purity of the state can perform quite badly if they are used to estimate a pure state or a ”nearly pure” state. It seems that the way of conditioning is critical for the methods capable of estimating both pure and mixed states. The simple length constraint of the least squares method works rather effectively, thus a version of the Bayesian estimation method with LS-type constraining is a good candidate of an improved stochastic state estimation method.

REFERENCES

D’Ariano, G.M., M.G.A. Paris and M.F. Sacchi (2003). Quantum tomography. *arXiv* **quant-ph/0302028**, v1.

Kosut, R.L., I. Walmsley and H. Rabitz (2004), Optimal experiment design for quantum state and process tomography and hamiltionan parameter estimation. *arXiv* **quant-ph/0411093**, v1.

Řehaček, J., B.-G. Englert and D. Kraszlikowski (2004), Minimal qubit tomography. *arXiv* **quant-ph/0405084**, v2.