Journal of Statistical Planning and Inference 184 (2017) 62–76

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Pseudo-Bayesian quantum tomography with rank-adaptation

The Tien Mai *, Pierre Alquier

CREST, ENSAE, Université Paris Saclay, 3 av. Pierre Larousse, 92245 Malakoff CEDEX, France

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 10 October 2016
Accepted 24 November 2016
Available online 10 December 2016

Keywords:
Quantum statistics
Bayesian statistics
PAC-Bayesian bounds
Oracle inequalities
MCMC

ABSTRACT

Quantum state tomography, an important task in quantum information processing, aims at reconstructing a state from prepared measurement data. Bayesian methods are recognized to be one of the good and reliable choices in estimating quantum states (Blume-Kohout, 2010). Several numerical works showed that Bayesian estimations are comparable to, and even better than other methods in the problem of 1-qubit state recovery. However, the problem of choosing prior distribution in the general case of n qubits is not straightforward. More importantly, the statistical performance of Bayesian type estimators has not been studied from a theoretical perspective yet. In this paper, we propose a novel prior for quantum states (density matrices), and we define pseudo-Bayesian estimators of the density matrix. Then, using PAC-Bayesian theorems (Catoni, 2007), we derive rates of convergence for the posterior mean. The numerical performance of these estimators is tested on simulated and real datasets.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Playing a vital role in quantum information processing, as well as being fundamental for characterizing quantum objects, quantum state tomography focuses on reconstructing the (unknown) state of a physical quantum system (Paris and Řeháček, 2004), usually represented by the so-called density matrix ρ (the exact definition of a density matrix is given in Section 2). This task is done by using outcomes of measurements performed on many independent systems identically prepared in the same state.

The 'tomographic' method, also named as linear/direct inversion (Vogel and Risken, 1989; Řeháček et al., 2010), is the simplest and oldest estimation procedure. It is actually the analogous of the least-square estimator in the quantum setting. Although easy in computation and providing unbiased estimate (Schwemmer et al., 2015), it does not generate a physical density matrix as an output (Shang et al., 2014). Maximum likelihood estimation (Hradil et al., 2004) is the current procedure of choice. Unfortunately, it has some critical flaws detailed in Blume-Kohout (2010), including a huge computational complexity. Furthermore, both these methods are not adaptive to the case where a system is in a state ρ for which some additional information is available. Note especially that, physicists focus on so-called pure states, for which $\text{rank}(\rho) = 1$.

The problem of rank-adaptivity was tackled thanks to adequate penalization. Rank-penalized maximum likelihood (BIC) was introduced in Guţă et al. (2012) while a rank-penalized least-square estimator $\hat{\rho}_{\text{rank-pen}}$ was proposed in Alquier et al. (2013), together with a proof of its consistency. More specifically, when the density matrix of the system is ρ^0 with $r = \text{rank}(\rho^0)$, the authors of Alquier et al. (2013) proved that the Frobenius norm of the estimation error satisfies $\|\hat{\rho}_{\text{rank-pen}} - \rho^0\|_F^2 = \mathcal{O}(r4^n/N)$ where N is the number of quantum measurements. The rate was improved to $\mathcal{O}(r3^n/N)$ by

* Corresponding author.
E-mail addresses: thetien.mai@ensae.fr (T.T. Mai), pierre.alquier@ensae.fr (P. Alquier).

http://dx.doi.org/10.1016/j.jspi.2016.11.003
0378-3758/© 2016 Elsevier B.V. All rights reserved.

Butucea et al. (2015), using a thresholding method. Note that the rate $\mathcal{O}(r2^n/N)$ was first claimed in the paper, but in the Corrigendum (Butucea et al., 2016), the authors acknowledge that this is not the case. The paper however contains a proof that no method can reach a rate smaller than $r2^n/N$. So, the minimax-optimal rate is somewhere in between $r2^n/N$ and $r3^n/N$.

Note that all the aforementioned papers only cover the complete measurement case (the definition is given in Section 2, basically it means that we have observations for all the observables given by the Pauli basis). The statistical relationship between matrix completion and quantum tomography with incomplete measurements (in the Le Cam paradigm) has been investigated in Wang (2013). Thus compressed sensing ideas have been successfully proposed in estimating a density state from incomplete measurements (Gross et al., 2010; Gross, 2011; Flammia et al., 2012; Koltchinskii, 2011).

On the other hand, Bayesian estimation has been considered in this context. The papers (Bužek et al., 1998; Baier et al., 2007) compare Bayesian methods to other methods on simulated data. More recently, Kravtsov et al. (2013), Ferrie (2014), Kueng and Ferrie (2015) and Schmied (2016) discuss efficient algorithms for computing Bayesian estimators. Importantly, Blume-Kohout (2010) showed that Bayesian method comes with natural error bars and is the most accurate scheme w.r.t. the expected error (operational divergence) (even) with finite samples. However, there is no theoretical guarantee on the convergence of these estimators.

More works on quantum state tomography in various settings include Audenaert and Scheel (2009), Carlen (2010), Rau (2011), Rau (2014) and Ferrie and Granade (2014).

In this paper, we consider a pseudo-Bayesian estimation, where the likelihood is replaced by pseudo-likelihoods based on various moments (two estimators, corresponding to two different pseudo-likelihood, are actually proposed). Using PAC-Bayesian theory (Shawe-Taylor and Williamson, 1997; McAllester, 1998; Catoni, 2004, 2007; Dalalyan and Tsybakov, 2008; Suzuki, 2012), we derive oracle inequalities for the pseudo-posterior mean. We obtain rates of convergence for these estimators in the complete measurement setting. One of them has a rate as good as the best known rate up to date $\mathcal{O}(\text{rank}(\rho^0)3^n/N)$ (still, the other one is interesting for computational reasons that are discussed in the paper).

The rest of the paper is organized as follows. We recall the standard notations and basics about quantum theory in Section 2. Then the definition of the prior and of the estimators are presented in Section 3. The statistical analysis of the estimators is given in Section 4, while all the proofs are delayed to the Appendix. Some numerical experiments on simulated and real datasets are given in Section 5.

2. Preliminaries

2.1. Notations

A very good introduction to the notations and problems of quantum statistics is given in Artiles et al. (2005). Here, we only provide the basic definitions required for the paper.

In quantum physics, all the information on the physical state of a system can be encoded in its *density matrix* ρ. Depending on the system in hand, this matrix can have a finite or infinite number of entries. A two-level system of n-qubits is represented by a $2^n \times 2^n$ density matrix ρ, with coefficients in $\mathbb{C}$. For the sake of simplicity, the notation $d = 2^n$ is used in Butucea et al. (2015), so note that ρ is a $d \times d$ matrix. This matrix is Hermitian $\rho^\dagger = \rho$ (i.e. self-adjoint), semidefinite positive $\rho \geq 0$ and has $\text{Trace}(\rho) = 1$. Additionally, it often makes sense to assume that the rank of ρ is small (Gross et al., 2010; Gross, 2011). In theory, the rank can be any integer between 1 and 2^n, but physicists are especially interested in pure states and a pure state ρ can be defined by $\text{rank}(\rho) = 1$.

The objective of quantum tomography is to estimate ρ on the basis of experimental observations of many independent and identical systems prepared in the state ρ by the same experimental device.

For each particle (qubit), one can measure one of the three Pauli observables $\sigma_x, \sigma_y, \sigma_z$. The outcome for each will be 1, or -1, randomly (the corresponding probability depends on the state ρ and will be given in (1)). Thus for a n-qubits system, we consider 3^n possible experimental observables. The set of all possible performed observables is

$$\{\sigma_{\mathbf{a}} = \sigma_{a_1} \otimes \cdots \otimes \sigma_{a_n};\ \mathbf{a} = (a_1, \ldots, a_n) \in \mathcal{E}^n := \{x, y, z\}^n\},$$

where vector $\mathbf{a}$ identifies the experiment. The outcome for each fixed observable setting will be a random vector $\mathbf{s} = (s_1, \ldots, s_n) \in \mathcal{R}^n := \{-1, 1\}^n$, thus there are 2^n outcomes in total.

Let us denote $R^{\mathbf{a}}$ a $\mathcal{R}^n$-valued random vector that is the outcome of an experiment indexed by $\mathbf{a}$. From the basic principles of quantum mechanics (Born's rule), its probability distribution is given by

$$\forall \mathbf{s} \in \mathcal{R}^n, \quad p_{\mathbf{a},\mathbf{s}} := \mathbb{P}(R^{\mathbf{a}} = \mathbf{s}) = \text{Trace}\left(\rho \cdot P_{\mathbf{s}}^{\mathbf{a}}\right), \tag{1}$$

where $P_{\mathbf{s}}^{\mathbf{a}} := P_{s_1}^{a_1} \otimes \cdots \otimes P_{s_n}^{a_n}$ and $P_{s_i}^{a_i}$ is the orthogonal projection associated to the eigenvalue s_i in the diagonalization of σ_{a_i} for $a_i \in \{x, y, z\}$ and $s_i \in \{-1, 1\}$ — that is $\sigma_{a_i} = -1P_{-1}^{a_i} + 1P_{+1}^{a_i}$.

The quantum state tomography problem is as follows: a physicist has access to an experimental device that produces n-qubits in a state ρ^0, and ρ^0 is assumed to be unknown. He/she can produce a large number of replications of the n-qubits and wants to infer ρ^0 from this.

Download English Version:

https://daneshyari.com/en/article/5129587

Download Persian Version:

https://daneshyari.com/article/5129587

Daneshyari.com