ADAPTIVE ESTIMATION OF THE DENSITY MATRIX IN QUANTUM HOMODYNE TOMOGRAPHY WITH NOISY DATA

PIERRE ALQUIER
SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY COLLEGE DUBLIN

KATIA MEZIANI
CEREMADE, UNIVERSITÉ PARIS-DAUPHINE

GABRIEL PEYRÉ
CEREMADE, UNIVERSITÉ PARIS-DAUPHINE

ABSTRACT. In the framework of noisy quantum homodyne tomography with efficiency parameter $1/2 < \eta \leq 1$, we propose an estimator of a quantum state whose density matrix elements $\rho_{m,n}$ decrease like $Ce^{-B(m+n)^{r/2}}$, for fixed $C \geq 1$, $B > 0$ and $0 < r \leq 2$. On the contrary to previous works, we focus on the case where r, C and B are unknown. The procedure estimates the matrix coefficients by a projection method on the pattern functions, following [ABM09], and then by soft-thresholding the coefficients. We compute the convergence rates of these estimators, in $\mathbb{L}_2$ risk. We obtain the same rate of convergence as if r, C and B were known.

CONTENTS

1. Introduction 1
1.1. Physical background 2
1.2. Statistical model 4
1.3. Outline of the results 6
2. Density matrix estimation 6
2.1. Adapted pattern functions 6
2.2. Estimation procedure 7
2.3. Main results 8
3. Simulation study 11
3.1. Examples 11
3.2. Pattern functions $f^{\eta}_{j,k}$ 11
3.3. Implementation of our procedure 11
3.4. Studies of the performance of our test procedure 12
4. Proof of Theorem 2.1 14
4.1. Some preliminary results 14
4.2. Proof of Theorem 2.1 17
References 18

1. INTRODUCTION

This paper deals with a *severely ill-posed inverse problem* come from quantum optics. Quantum optics is a branch of quantum mechanics which studies physical systems at the atomic and subatomic scales. As the language used by physicists differs from the one that used by statisticians, we start with general notions on quantum mechanics. "laws" or " random variables"...), The interested reader can get further acquaintance with quantum concepts through the textbooks or the review article [Hel76, Hol82, BNGJ03, Leo97].

1.1. **Physical background.**

1.1.1. *Quantum mechanics.*

In quantum mechanics, the quantum state of a system is a mathematical object which encompasses all the information about the system. The most common representation of a quantum state is an operator ρ on a complex Hilbert space $\mathcal{H}$ (called the space of states) such that ρ is

(1) Self adjoint: $\rho = \rho^*$, where ρ^* is the adjoint of ρ.
(2) Positive: $\rho \geq 0$, or equivalently $\langle \psi, \rho\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$.
(3) Trace one: $\mathrm{tr}(\rho) = 1$.

A quantum state ρ encodes the probabilities of the measurable properties, or "**observables**" (energy, position, ...) of the considered quantum system. Generally, in quantum mechanics the expected results of the measurements of an observable are not deterministic values but predictions about probability distributions, that is the probability of obtaining each of the possible outcomes when measuring an observable.

An observable $\mathbf{X}$ is described by a self adjoint operator on the space of states $\mathcal{H}$ and

$$\mathbf{X} = \sum_{a}^{dim\mathcal{H}} x_a \mathbf{P}_a,$$

where the eigenvalues $\{x_a\}_a$ of the observable $\mathbf{X}$ are real and $\mathbf{P}_a$ is the projection onto the one dimensional space generated by the eigenvector of $\mathbf{X}$. Then, when performing a measurement of the observable $\mathbf{X}$ of a quantum state ρ, the result is a random variable X with values in the set of the eigenvalues of the observable $\mathbf{X}$. For a quantum system prepared in state ρ, X has the following probability distribution and expectation function

$$\mathbb{P}_\rho(X = x_a) = Tr(\mathbf{P}_a \rho) \text{ et } \quad \mathbb{E}_\rho(X) = Tr(\mathbf{X}\rho).$$

An important element which affects the result of the measurement process is the purity of quantum states. A state is called pure if it cannot be represented as a mixture (convex combination) of other states, i.e., if it is an extreme point of the convex set of states. All other states are called mixed states. We give examples of states in Section 3.

1.1.2. *Quantum optics.*

In this paper, the quantum system we work with is a monochromatic light in a cavity described by a quantum harmonic oscillator. In the framework of quantum optics, the space of states is known to be the separable Hilbert space $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, i.e. the space of square integrable complex valued functions on the real line. A particular orthonormal basis $(\psi_j)_{j\in\mathbb{N}}$ comes with, called the Fock basis. This physically very meaningful basis is defined for all $j \in \mathbb{N}$ as follow

$$\psi_j(x) := \frac{1}{\sqrt{\sqrt{\pi}2^j j!}} H_j(x) e^{-x^2/2}, \tag{1}$$

where $H_j(x) := (-1)^j e^{x^2} \frac{d^j}{dx^j} e^{-x^2}$ is the j-th Hermite polynomial. In the Fock basis (1), a state is described by infinite density matrix $\rho = [\rho_{j,k}]_{j,k\in\mathbb{N}}$.

We may give an equivalent representation for a quantum state ρ: the associated Wigner function W_ρ (see [Wig32]). The Wigner function W_ρ is a real function of two variables and may be defined by its Fourier Transform $\mathcal{F}_2$ with respect to both variables

$$\widetilde{W}_\rho(u,v) \quad := \quad \mathcal{F}_2[W_\rho](u,v) = \mathrm{Tr}\left(\rho \exp(iu\mathbf{Q} + iv\mathbf{P})\right),$$

where $\mathbf{Q}$ and $\mathbf{P}$ are respectively the electric and magnetic fields. These two observables, we are concerned by, do not commute. As non-commuting observables, they may not be simultaneously measurable. Therefore, by performing measurements on $(\mathbf{Q},\mathbf{P})$, we cannot get a density probability of the result (Q,P). However, for $\phi \in [0,\pi]$ we can measure the quadrature observables $\mathbf{X}_\phi := \mathbf{Q}\cos\phi + \mathbf{P}\sin\phi$, and then the above Wigner function plays the role of a quasi-probability density. It does not satisfy all the properties of a conventional probability distribution but satisfies boundedness properties unavailable to classical distributions. For instance, the Wigner distribution can and normally does go negative for states which have no classical model. The Wigner function is such that

- $W_\rho : \mathbb{R}^2 \to \mathbb{R}$
- $\iint_{\mathbb{R}^2} W_\rho(q,p)dqdp = 1$,

However its Radon transform is always a probability density

$$p_\rho(x|\phi) := \mathcal{R}[W_\rho](x,\phi) = \int_{-\infty}^{\infty} W_\rho(x\cos\phi - t\sin\phi,\, x\sin\phi + t\cos\phi)dt, \tag{2}$$

with respect to $\frac{1}{\pi}\lambda$, λ being the Lebesgue measure on $\mathbb{R}\times[0,\pi]$.

Now we can explicit the links between the state ρ and the Radon transform of its associated Wigner function W_ρ, $p_\rho(x|\phi)$. In the Fock basis (1), a state is described by infinite density matrices $\rho = [\rho_{j,k}]_{j,k\in\mathbb{N}}$ with entries $\rho_{j,k}$ expressed as

$$\rho_{j,k} = \frac{1}{\pi}\int\int_0^\pi p_\rho(x|\phi) f_{j,k}(x) e^{-i(k-j)\phi} d\phi dx \tag{3}$$

for all $j,k\in\mathbb{N}$. The functions $f_{j,k} = f_{k,j}$, in the expression (3), are bounded real functions commonly called *pattern functions* in quantum homodyne literature. A concrete expression for their Fourier transform using Laguerre polynomials can be found in [Ric00]: for $j \geq k$,

$$\tilde{f}_{k,j}(t) \quad = \quad \pi(-i)^{j-k}\sqrt{\frac{2^{k-j}k!}{j!}}|t|t^{j-k}e^{-\frac{t^2}{4}}L_k^{j-k}(\frac{t^2}{2}). \tag{4}$$

where $\tilde{f}_{k,j}$ denotes the Fourier transform of the Pattern function $f_{k,j}$ and

$$L_n^\alpha(x) := (n!)^{-1}e^x x^{-\alpha}\frac{d^n}{dx^n}(e^{-x}x^{n+\alpha})$$

is the Laguerre polynomial of degree n and order α.

1.1.3. *QHT.*

In this paper, we address the problem of reconstructing the density matrix ρ of a monochromatic light in a cavity. As the observables $\mathbf{Q}$ and $\mathbf{P}$ cannot be measured simultaneously, we measure the quadrature $\mathbf{X}_\phi := \mathbf{Q}\cos\phi + \mathbf{P}\sin\phi$, where $\phi \in [0,\pi]$. Each of these quadrature could be measured on a laser beam by a technique put in practice for the first time in [SBRF93] and called Quantum

Homodyne Tomography (QHT). The theoretical foundation of quantum homodyne tomography was outlined in [VR89].

The experimental set-up described in Figure 1 consists of mixing the cavity pulse prepared in state ρ with an additional laser of high intensity $|z| >> 1$ called the local oscillator. After the mixing, the beam is split again and each of the two emerging beams is measured by one of the two photodetectors which give integrated currents I_1, I_2 proportional to the number of photons. The result of the measurement is produced by taking the difference of the two currents and rescaling it by the intensity $|z|$. Just before the mixing the experimentalist may choose the phase Φ of the local oscillator, randomly, uniformly distributed on $[0, \pi]$. In the case of noiseless measurement and for a phase $\Phi = \phi$, the result $X_\phi = \frac{I_2 - I_1}{|z|}$ has density $p_\rho(x|\phi)$ corresponding to measuring $\mathbf{X}_\phi$.

In practice, a number of photons fails to be detected. These losses may be quantified by one single coefficient $\eta \in [0, 1]$, such that $\eta = 0$ when there is no detection and $\eta = 1$ corresponds to the ideal case. The physicists argue, that their machines actually have high detection efficiency, around 0.8/0.9. Thus, we suppose η known. As the detection process is inefficient, an independent gaussian noise interferes additively with the ideal data X_ϕ. Thus for $\Phi = \phi$, the effective result of the QHT measurement (Figure 1) is for a known efficiency $\eta \in]0.5, 1]$,

$$Y = \sqrt{\eta} X_\phi + \sqrt{(1-\eta)/2}\xi$$

where ξ is a standard Gaussian random variable, independent of X_ϕ.

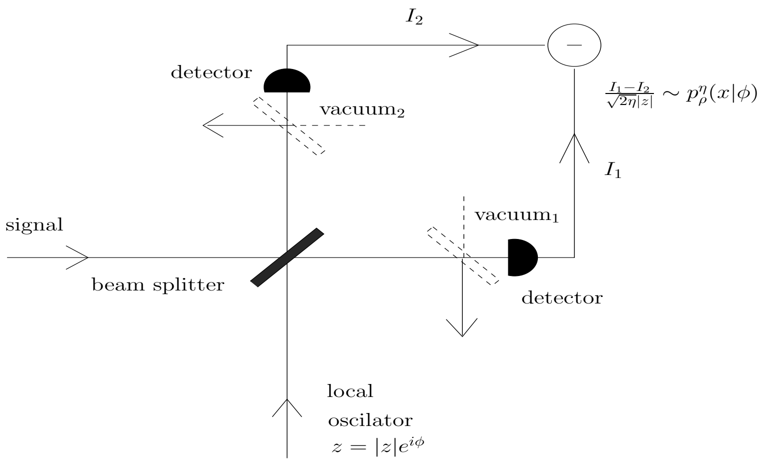


FIGURE 1. QHT measurement scheme

1.2. **Statistical model.**

This paper aims at reconstructing the density matrix of a monochromatic light in a cavity prepared in state ρ. As we cannot measure precisely the quantum state in a single experiment, we perform measurements on n independent identically prepared quantum systems. The measurement carried out on each of the n systems in state ρ is done by QHT as described in Section 1.1.3. In the ideal setting, the results of such experiments would be n independent identically distributed random variables $(X_1, \Phi_1), \dots, (X_n, \Phi_n)$ with values in $\mathbb{R} \times [0, \pi]$ and distribution P_ρ having density with respect to λ, λ being the Lebesgue measure on $\mathbb{R} \times [0, \pi]$ equals to

$$p_\rho(x, \phi) = \frac{1}{\pi} p_\rho(x|\phi) = \frac{1}{\pi} \mathcal{R}[W_\rho](x, \phi), \tag{5}$$

where $\mathcal{R}$ is the Radon transform defined in equation (2). As underlined in Section 1.1.3, we do not observe $(X_\ell, \Phi_\ell)_{\ell=1,\ldots n}$ but the noisy version $(Y_\ell, \Phi_\ell)_{\ell=1,\ldots n}$ where

$$Y_\ell := \sqrt{\eta} X_\ell + \sqrt{(1-\eta)/2}\ \xi_\ell. \tag{6}$$

Here ξ_ℓ are independent standard Gaussian random variables, independent of all (X_ℓ, Φ_ℓ), $\ell = 1, \ldots n$. The detection efficiency $\eta \in]0.5, 1]$ is a known parameter and $1 - \eta$ represents the proportion of photons which are not detected due to various losses in the measurement process.
Let us denote by $p^\eta_\rho(y, \phi)$ the density of (Y_ℓ, Φ_ℓ). Then, for $\Phi = \phi$, the density $p^\eta_\rho(\cdot|\phi)$ is the convolution of the density $\frac{1}{\sqrt{\eta}} p_\rho(\frac{\cdot}{\sqrt{\eta}}|\phi)$ of $\sqrt{\eta} X$ with N^η the density of a centered Gaussian density having variance $(1 - \eta)/2$ s.t.

$$p^\eta_\rho(y|\phi) = \left(\frac{1}{\sqrt{\eta}} p_\rho \left(\frac{\cdot}{\sqrt{\eta}} |\phi \right) * N^\eta \right)(y) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\eta}} p_\rho \left(\frac{y-x}{\sqrt{\eta}} |\phi \right) N^\eta(x) dx. \tag{7}$$

For $\Phi = \phi$, an useful equation in the Fourier domain, deduced by the previous relation (7) is

$$\mathcal{F}_1[\sqrt{\eta} p^\eta_\rho(\cdot\sqrt{\eta}|\phi)](t) = \mathcal{F}_1[p_\rho(\cdot|\phi)](t) \widetilde{N}^\gamma(t), \tag{8}$$

where $\mathcal{F}_1$ denotes the Fourier transform with respect to the first variable and $\widetilde{N}^\eta(t) = e - \frac{1-\eta}{4\eta}$ is the Fourier transform of $N^\gamma(x)$, the density of a centered Gaussian density having variance $(1 - \eta)/2\eta = \gamma$.

In order to estimate the element of the density matrix defined in (3) from the data $(Y_\ell, \Phi_\ell)_{\ell=1,\ldots n}$, we define a realistic class of quantum states $\mathcal{R}(C, B, r)$. For $C \geq 1$, $B > 0$ and $0 < r \leq 2$, the class $\mathcal{R}(C, B, r)$ is defined as follow

$$\mathcal{R}(C, B, r) := \{\rho \text{ quantum state} : |\rho_{m,n}| \leq C \exp(-B(m+n)^{r/2})\}. \tag{9}$$

Note that the class $\mathcal{R}(C, B, r)$ has been translated in terms of Wigner functions in [ABM09], where it has been proved that the fast decay of the elements of the density matrix implies both rapid decay of the Wigner function and of its Fourier transform.
An useful lemma is the following.

Lemma 1.1. *For $\rho \in \mathcal{R}(C, B, r)$, the set defined in (9), there exists a M_0 s.t. $\forall M \geq M_0$ implies*

$$\sum_{j+k>M} \rho^2_{j,k} \leq \mathcal{C} M^{2-\frac{r}{2}} e^{-2BM^{\frac{r}{2}}}, \tag{10}$$

where $\mathcal{C} = \frac{2C^2}{Br}$.

Proof. For $\rho \in \mathcal{R}(C, B, r)$, we have by the definition of the class $\mathcal{R}(C, B, r)$ and by Lemma 3 in [ABM09]

$$\sum_{j+k>M} \rho^2_{j,k} \leq C^2 \sum_{j+k>M} \exp(-2B(j+k)^{r/2}) \leq \frac{2C^2}{Br} M^{2-\frac{r}{2}} e^{-2BM^{\frac{r}{2}}}.$$

□

However, on the contrary to previous works, we do not assume here that the constants r, B and C are known. From now on we denote by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ the usual Euclidean scalar product and norm, while $C(\cdot)$ will denote positive constants depending on parameters given in the parentheses.

1.3. **Outline of the results.**

The goal of this paper is to define an estimators of the density matrix and to evaluate its performance in $\mathbb{L}_2$ risk.

The reconstruction of the density from averages of data has been discussed or studied in [DMP94, DLP95, LPD95, AGG05] for $\eta = 1$ (no photon loss). Max-likelihood methods have been studied in [BDPS00, AGG05, DMS05, Guţ07] and procedure using adaptive tomographic kernels to minimize the variance has been proposed in [DP99]. The estimation of the density matrix of a quantum state of light in case of efficiency parameter $\frac{1}{2} < \eta \leq 1$ has been discussed in [D'A95, DMS05, D'A02] and considered in [Ric96] via the pattern functions for the diagonal elements. For the case $0 < \eta < 1/2$, see [ABM09]. However, the physicists argue, that their machines actually have high detection efficiency, around 0.8/0.9. So we do not deal here with values of η smaller than 1/2.

It is to be noted that the estimator proposed in [ABM09] depends on the knowledge of B and r. However, in practice, one will face situations where one wants to reconstruct a density matrix without assuming the knowledge of B and r. This is known in statistics as "adaptive estimation." In this paper, we tackle the problem of adaptive estimation over the class of quantum states $\mathcal{R}(C, B, r)$. Our estimator is actually thresholded version of the estimator in [ABM09]. Namely, in [ABM09], an estimator for each coefficient of the density matrix is proposed. Here, our estimator reaches adaptation by substituting a soft-thresholded version to this estimator. Coefficients thresholding is now a classical tool in statistics. It was introduced in a series of papers by Donoho and Johnstone [DJ94, DJ95, DJKP95] in the context of function estimation via wavelets coefficients. See [HKPT98] for a nice introduction to thresholding and wavelts, see also [Cai99, Alq08a]. These methods were extended to inverse problems [Don95, Joh99, Kol96, CGLT03], see [Cav11] for an introduction and a survey of the most recent results.

In Section 2 we define our thresholded estimator for the density matrix, and state our main result on adaptation. Namely, we prove that the estimator reaches the same rate as in [ABM09] without knowing B, C nor r. In Section 3 we provide some simulation study. Proofs are given in Section 4.

To conclude, we may infer that the performances of this estimator with the estimator of the Wigner function in Chapter 3 of [Mez08] and in [ABM09] are comparable. We obtain polynomial rates for the case $r = 2$ and intermediate rates for $0 < r < 2$ (faster than any logarithm, but slower than any polynomial).

2. Density matrix estimation

The aim of this part is to estimate the density matrix ρ in the Fock basis directly from the data $(Y_i, \Phi_i)_{i=1,\dots,n}$. We construct an estimator of the density matrix $[\rho_{j,k}]_{j,k}$ from a sample of QHT data. We give theoretical results for our estimator when the quantum state ρ is in $\mathcal{R}(C, B, r)$, the class of density matrix with decreasing elements defined in (9).

2.1. **Adapted pattern functions.**

In order to reconstruct the entries of the density matrix from the noisy observations (Y_ℓ, Φ_ℓ) by a projection type estimator on the *pattern* functions, we have to adapt the pattern functions as follows. From now on, we shall use the notation $\gamma = \gamma(\eta) := \frac{1-\eta}{4\eta}$. We denote by $f^{\eta}_{k,j}$ the function which has the following Fourier transform:

$$\tilde{f}^{\eta}_{k,j}(t) := \tilde{f}_{k,j}(t) e^{\gamma t^2}, \tag{11}$$

where $\tilde{f}_{k,j}$ are the pattern functions defined in expression (4).

Lemma 2.1. *For $\eta \in (1/2, 1)$, there exists a positive constant $C^{\eta}_{\infty} > 0$ s.t.*

$$\sum_{0 \le j+k \le M} \|f^{\eta}_{j,k}\|^2_{\infty} \le C^{\eta}_{\infty} M^{\frac{1}{3}} e^{8\gamma M}, \tag{12}$$

where $\gamma = (1-\eta)/(4\eta)$ and the $(f^{\eta}_{j,k})_{j,k}$ are the adapted pattern functions defined in expression (11).
There exists a positive constant $C_{\infty} > 0$ s.t.

$$\sum_{0 \le j+k \le M} \|f_{j,k}\|^2_{\infty} \le C_{\infty} M^{\frac{10}{3}}, \tag{13}$$

where the $(f_{j,k})_{j,k}$ are the pattern functions defined in expression (4).

Proof. For the proof of this lemma, we refer to Lemma 4 and Lemma 5 in [ABM09]. □

2.2. **Estimation procedure.**

For $N := N(n) \to \infty$, we define the set of indices $J(N) \subset \mathbb{N}^2$ as

$$J(N) := \{(j,k) \in \mathbb{N}^2, 0 \le j+k \le N-1\}. \tag{14}$$

For $N := N(n) \to \infty$, let $\hat{\rho}^{\eta}$ be a previous estimator of ρ defined by its entries

$$\hat{\rho}^{\eta}_{j,k} := \begin{cases} \frac{1}{n}\sum_{\ell=1}^{n} G_{j,k}\left(\frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell\right) & \forall (j,k) \in J(N), \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where $(G_{j,k})_{j,k}$ are the functions using the pattern functions in (11) and

$$G_{j,k}(x,\phi) := f^{\eta}_{j,k}(x) e^{-i(j-k)\phi}. \tag{16}$$

Note that this procedure estimates the matrix coefficients by a projection method on the pattern functions and has been introduced by [ABM09]. To define our final procedure estimation, let us introduce some notation. For any $x \in \mathbb{R}$, we denote by $(x)_+$ the following quantity

$$(x)_+ = \begin{cases} x & \text{, if } x \ge 0 \\ 0 & \text{, otherwise.} \end{cases} \tag{17}$$

We denote by $\operatorname{sgn}(\cdot)$ the sign function, given by, $\forall x \in \mathbb{R}$,

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{if } , x > 0, \\ 0 & \text{if } , x = 0, \\ -1 & \text{if } , x < 0. \end{cases} \tag{18}$$

From now, we denote by $\|.\|_\infty$ the supremum norm for functions, i.e. for any f,

$$\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|.$$

For $N := N(n) \to \infty$ and for some fixed confidence level $\varepsilon \in (0,1)$, we define the soft-thresholded estimator

$$\tilde{\rho}^{\eta}_{j,k} = \operatorname{sgn}(\hat{\rho}^{\eta}_{j,k}) \left(|\hat{\rho}^{\eta}_{j,k}| - t_{j,k}\right)_+, \tag{19}$$

where

$$t_{j,k} = \|f^{\eta}_{j,k}\|_{\infty}\sqrt{\frac{2\log\left(\frac{N(N+1)}{\varepsilon}\right)}{n}}. \tag{20}$$

Finally, our density matrix estimator is given by

$$\tilde{\rho}^{\eta} = [\tilde{\rho}^{\eta}_{j,k}]_{j,k}.$$

2.3. **Main results.**

For any density matrix $\nu = (\nu_{j,k})_{j,k\geq 0}$, we define the ℓ_2-norm of ν as

$$\|\nu\|_2 = \sqrt{\sum_{j,k\geq 0} \nu^2_{j,k}}.$$

We have the following oracle inequality that will allow to obtain rates of convergence on the classes $\mathcal{R}(C, B, r)$.

Theorem 2.1. *With probability larger than* $1 - \varepsilon$, *we have*

$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 \leq \inf_{I\subseteq J(N)} \left\{ 4 \sum_{(j,k)\in I} t^2_{j,k} + \sum_{(j,k)\notin I} \rho^2_{j,k} \right\},$$

where the set $J(N)$ *is defined in* (14).

The proof is given in Section 4.

Corollary 2.1. *Let us put* $r_0 \in (0, 2)$, $B_0 > 0$ *and let us choose*

$$N = N(n) := \left\lfloor \left(\frac{\log(n)}{2B_0}\right)^{\frac{2}{r_0}} \right\rfloor, \tag{21}$$

where $\lfloor x \rfloor$ *denote the integer part of* x *such that* $\lfloor x \rfloor \leq x \leq \lfloor x \rfloor + 1$. *Let us assume that we are in the case* $\rho \in \mathcal{R}(C, B, r)$, *for some unknown* $C \geq 1$, $B \geq B_0$, $r \in [r_0, 2]$. *Then, there are constants* $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3 > 0$ *such that with probability larger than* $1 - \varepsilon$, *we have*

- *For* $\eta = 1$ *and* $r \in [r_0, 2]$
$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 \leq \mathcal{C}_1 n^{-1} (\log(n))^{\frac{20}{3r}} \log(\log(n)/\varepsilon).$$

- *For* $\eta \in (\frac{1}{2}, 1)$ *and* $r = 2$
$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 \leq \mathcal{C}_2 n^{-\frac{B}{4\gamma+B}} \Big(\log(n) + (\log(n))^{1/3} \log(\log(n)/\varepsilon)\Big).$$

- *For* $\eta \in (\frac{1}{2}, 1)$ *and* $r \in (r_0, 2)$
$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 \leq \mathcal{C}_3 e^{-2BM(n)^{r/2}} \Big(\log(n)^{2-r/2} + \log(n)^{1/3} \log(\log(n)/\varepsilon)\Big),$$

  where $M(n)$ *satisfies* $8\gamma M(n) + 2BM(n)^{r/2} = \log(n)$. *In particular, note that*
$$M(n) = \frac{1}{8\gamma}\log(n) - \frac{2B}{(8\gamma)^{1+r/2}}\log(n)^{r/2} + o(\log(n)^{r/2}).$$

The rate of convergence in the case $1 > \eta > \frac{1}{2}$ and $r_0 < r < 2$ is slower than any polynomial rate but faster than any logarithmic rate. Observe that we obtain exactly the same rate than in [ABM09] and [Mez08], up to a $\log(\log(n)/\varepsilon)$ term. But on the other hand, this result is adaptive, in the sense that we do not have to know B or r to build the estimator.

Proof of Corollary 2.1. For $r_0 \in (0,2)$, $B_0 > 0$ and N as in (21), let M be an integer s.t. $M < N$. We define the set

$$J(M) := \{(j,k) \in \mathbb{N}^2,\, 0 \le j+k \le M\}.$$

Then, for $\varepsilon \in (0,1)$ and by applying Theorem 2.1 to $I = J(M)$, with probability larger than $1-\varepsilon$, we obtain

$$\begin{aligned} \|\tilde{\rho}^\eta - \rho\|_2^2 &\le \inf_{0\le M\le N-1} \left\{ 4 \sum_{0\le j+k\le M} t_{j,k}^2 + \sum_{j+k>M} \rho_{j,k}^2 \right\} \\ &= \inf_{0\le M\le N-1} \left\{ \frac{8}{n} \sum_{0\le j+k\le M} \|f_{j,k}^\eta\|_\infty^2 \log\left(N(N+1)/\varepsilon\right) + \sum_{j+k>M} \rho_{j,k}^2 \right\}. \qquad (22) \end{aligned}$$

a) For $\eta = 1$ and $r \in [r_0, 2]$.

As $f_{j,k}^\eta = f_{j,k}$ for $\eta = 1$, we have by pluging (10) and (13) into (22)

$$\begin{aligned} \|\tilde{\rho}^\eta - \rho\|_2^2 &\le \inf_{0\le M\le N-1} \left\{ \frac{8}{n} \sum_{0\le j+k\le M} \|f_{j,k}\|_\infty^2 \log\left(N(N+1)/\varepsilon\right) + \sum_{j+k>M} \rho_{j,k}^2 \right\}. \\ &\le \inf_{0\le M\le N-1} \left\{ \frac{c_1}{n} M^{\frac{10}{3}} \log(N/\varepsilon) + \mathcal{C} M^{2-\frac{r}{2}} e^{-2BM^{\frac{r}{2}}} \right\}, \qquad (23) \end{aligned}$$

for some constant $c_1 > 0$.
For N such in (21) and by taking $M = (\log(n)/2B)^{2/r} < N$, it leads to

$$\|\tilde{\rho}^\eta - \rho\|_2^2 \le \mathcal{C}_1 \log\left(\log(n)/\varepsilon\right) \left(\log(n)\right)^{\frac{20}{3r}} n^{-1}$$

for some constant $\mathcal{C}_1 > 0$.

b) For $\eta \in (1/2, 1)$ and $r = 2$.

Next, we deal with the case $1 > \eta > 1/2$. We plug (10) and (12) into (22) to obtain in the case $r = 2$

$$\|\tilde{\rho}^\eta - \rho\|_2^2 \le \inf_{0\le M<N} \left\{ \frac{c_2}{n} \log\left(N/\varepsilon\right) M^{\frac{1}{3}} e^{8\gamma M} + \mathcal{C} M e^{-2BM} \right\},$$

for some constant $c_2 > 0$.
By taking $M = M(n)$ s.t.

$$M = \frac{\log(n)}{2(4\gamma + B)},$$

we obtain

$$\|\tilde{\rho}^\eta - \rho\|_2^2 \le \mathcal{C}_2 n^{-\frac{B}{4\gamma+B}} \left(\log\left(\log(n)/\varepsilon\right) \left(\log(n)\right)^{1/3} + \log(n) \right),$$

for some constant $\mathcal{C}_2 > 0$.

c) For $\eta \in (1/2,1)$ and $r \in (r_0,2)$.
Finally, in the case $\eta \in (1/2,1)$ and $r \in (r_0,2)$ and by plugging (10) and (12) into (22) we get:

$$\|\tilde{\rho}^\eta - \rho\|_2^2 \leq \inf_{0 \leq M < N} \left\{ \frac{c_3}{n} \log(N/\varepsilon) M^{\frac{1}{3}} e^{8\gamma M} + \mathcal{C} M^{2-\frac{r}{2}} e^{-2BM^{\frac{r}{2}}} \right\},$$

for some constant $c_3 > 0$.
For M a solution of the equation $8\gamma M + 2BM^{\frac{r}{2}} = \log(n)$ and for

$$M(n) = \frac{1}{8\gamma}\log(n) - \frac{2B}{(8\gamma)^{1+r/2}}\log(n)^{r/2} + o(\log(n)^{r/2})$$

in particular, we obtain

$$\|\tilde{\rho}^\eta - \rho\|_2^2 \leq \mathcal{C}_3 \exp^{-2BM^{r/2}} \left(\log(n)^{2-r/2} + \log(n)^{1/3} \log(N/\varepsilon) \right),$$

for some constant $\mathcal{C}_3 > 0$. □

Finally, let us remark that in the case where ρ is a pure state, we even have a faster rate of convergence.

Corollary 2.2. *Under the same choice for N than in Corollary 2.1,*

$$N = N(n) := \left\lfloor \left(\frac{\log(n)}{2B_0} \right)^{\frac{2}{r_0}} \right\rfloor,$$

if ρ is a pure state, i.e., if $\rho_{j_0,j_0} = 1$ and all the others $\rho_{j,k} = 0$, then we have, as soon as $N > \max(j_0,2)$, with probability larger than $1-\varepsilon$,

$$\|\tilde{\rho}^\eta - \rho\|_2^2 < \frac{32}{nr_0}\|f_{j_0,j_0}\|_\infty^2 \log\left(\frac{\log(n)}{B_0 \varepsilon} \right).$$

Proof of Corollary 2.2. We apply Theorem 2.1 for $I = \{(j_0,j_0)\}$. We obtain, with probability larger than $1-\varepsilon$,

$$\begin{aligned} \|\tilde{\rho}^\eta - \rho\|_2^2 &\leq \frac{8}{n} \sum_{(j,k)=(j_0,j_0)} \|f_{j,k}\|_\infty^2 \log(N(N+1)/\varepsilon) + \sum_{(j,k)\neq(j_0,j_0)} \rho_{j,k}^2 \\ &= \frac{8}{n}\|f_{j_0,j_0}\|_\infty^2 \log(N(N+1)/\varepsilon) + 0. \end{aligned}$$

For n large enough, $N = N(n) \geq 2$. Then, $(N+1) < 2N$ and

$$\begin{aligned} \|\tilde{\rho}^\eta - \rho\|_2^2 &< \frac{8}{n}\|f_{j_0,j_0}\|_\infty^2 \log\left(2N^2/\varepsilon\right) \\ &= \frac{8}{n}\|f_{j_0,j_0}\|_\infty^2 \left[2\log(N) + \log(2/\varepsilon)\right] \\ &\leq \frac{8}{n}\|f_{j_0,j_0}\|_\infty^2 \left[\frac{4}{r_0}\log\left(\frac{\log(n)}{2B_0}\right) + \log(2/\varepsilon) \right] \end{aligned}$$

where we replaced N by its definition. As $r_0 < 2$, $4/r_0 > 1$ and we have the following rough upper bound:

$$\begin{aligned} \|\tilde{\rho}^\eta - \rho\|_2^2 &< \frac{8}{n}\|f_{j_0,j_0}\|_\infty^2 \left[\frac{4}{r_0}\log\left(\frac{\log(n)}{2B_0}\right) + \frac{4}{r_0}\log(2/\varepsilon) \right] \\ &= \frac{32}{nr_0}\|f_{j_0,j_0}\|_\infty^2 \log\left(\frac{\log(n)}{B_0\varepsilon}\right). \end{aligned}$$

□

Table 1. Examples of quantum states

***Vacuum* state** • $\rho_{0,0} = 1$ rest zero, • $p_\rho(x\|\phi) = e^{-x^2}/\sqrt{\pi}$.
Single photon state • $\rho_{1,1} = 1$ rest zero, • $p_\rho(x\|\phi) = 2x^2 e^{-x^2}/\sqrt{\pi}$.
***Coherent-q_0* state** $q_0 \in \mathbb{R}$ • $\rho_{j,k} = e^{-\|q_0\|^2}(q_0/\sqrt{2})^{j+k}/\sqrt{j!k!}$, • $p_\rho(x\|\phi) = \exp(-(x - q_0\cos(\phi))^2)/\sqrt{\pi}$.
Thermal state $\beta > 0$ • $\rho_{j,k} = \delta_k^j(1 - e^{-\beta})e^{-\beta k}$, • $p_\rho(x\|\phi) = \sqrt{\tanh(\beta/2)/\pi}\exp(-x^2\tanh(\beta/2))$.
Schrödinger cat $q_0 > 0$ • $\rho_{j,k} = 2(q_0/\sqrt{2})^{j+k}/\left(\sqrt{j!k!}(\exp(q_0^2/2) + \exp(-q_0^2/2))\right)$, for j and k even, rest zero, • $p_\rho(x\|\phi) = \big(\exp(-(x - q_0\cos(\phi))^2) + \exp(-(x + q_0\cos(\phi))^2)$ $+2\cos(2q_0 x\sin(\phi))\exp(-x^2 - q_0^2\cos^2(\phi))\big) / \left(2\sqrt{\pi}(1 + \exp(-q_0^2))\right)$.

3. Simulation study

3.1. **Examples.** We present in Table 1 examples of pure quantum states, which can be created at this moment in laboratory and belong to the class $\mathcal{R}(C, B, r)$ with $r = 2$. Table 1 gives also their explicit density matrix coefficients $\rho_{j,k}$ and probability densities $p_\rho(x|\phi)$.
Among the pure states we consider the *vacuum* state, which is the pure state of zero photons. Note that the *vacuum* state would provide a random variable of Gaussian probability density $p_\rho(x|\phi)$ via the ideal measurement of QHT (see Section 1.1.3). That explains the gaussian nature of the noise in the effective result of the QHT measurement.
The *single photon* state, which is the pure state of one photon. We consider also the *coherent-q_0* state, which characterizes the laser pulse with the number of photons Poisson distributed with an average of M photons. Remark that the well-known *Schrödinger cat* state is described by a linear superposition of two *coherent* vectors (see e.g. [OTBLG06]).

3.2. **Pattern functions $f_{j,k}^\eta$.** Since there is no closed-form expression for the pattern functions $f_{j,k}^\eta$, we evaluate them numerically on a 1-D regular grid of $Q = 4096$ points. We use expressions (4) and (11) to evaluate $\tilde{f}_{j,k}$ and $\tilde{f}_{j,k}^\eta$ on the 1-D frequency grid of Q discretized t points. The adapted pattern functions $f_{j,k}^\eta$ are computed on the 1-D spacial grid of Q discretized x points by applying to $\tilde{f}_{j,k}^\eta$ the inverse Fast Fourier Transform (FFT) in $O(Q\log(Q))$ operations. Figure 2 shows a few examples of pattern functions.

3.3. **Implementation of our procedure.** The deconvolved estimator $\hat{\rho}_{j,k}^\eta$ defined in (15) is computed by evaluating $G_{j,k}(x,\phi) = f_{j,k}^\eta(x)e^{-i(j-k)\phi}$ at point x using a cubic spline interpolation of the values of $f_{j,k}^\eta$ on the discrete grid of Q points.

In the following section, we evaluate the performance of the threshold estimator $\tilde{\rho}_{j,k}^\eta$. We perform this evaluation by creating noisy samples Y_ℓ as defined in (6).

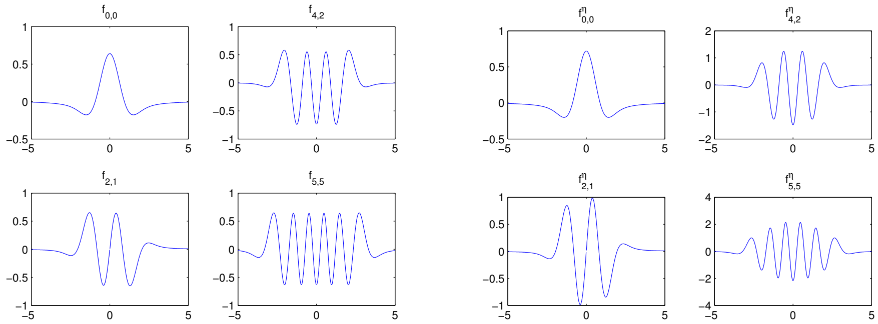


FIGURE 2. Examples of pattern functions $f_{j,k}$ (left) and adapted pattern functions $f^{\eta}_{j,k}$ (right).

The initial samples X_ℓ are drawn from the distribution $p_\rho(x|\phi)$ (see Table 1) using the rejection method.

The value of $N = N(n)$ is set following (21). We use $r_0 = 2$ and $B_0 = 1/2$ for all the numerical experiments.

A toolbox that implements this procedure and reproduces all the figures of this article is available online[1].

3.4. **Studies of the performance of our test procedure.** Figure 3 shows graphically the result of our estimator for three different pure quantum states, for several values of n. To further evaluate quantitatively the performance of the method, we estimate numerically the (relative) root mean square error (RMSE)

$$\mathrm{RMSE}(n) = \frac{\|\tilde{\rho}^{\eta} - \rho\|}{\|\rho\|}$$

of our soft thresholding estimator. More precisely, Figure 4 shows the evolution with n of the expected value of the RMSE. This expected value is evaluated by an empirical mean with Monte Carlo simulation using 50 samplings for each value of n. To evaluate the deviation with respect with this mean, we also display the confidence interval at ± 3 times the standard deviation of the RMSE.

The threshold values $t_{j,k}$ that are used in (19) to define our estimator are somehow conservative. In practice, smaller values offer better decay of the RMSE. Figure (4) display in dashed red (reap. dashed green) the decay of the RMSE obtained using thresholds $0.8t_{j,k}$ (resp. $0.5t_{j,k}$). We found on these three examples and for $\eta = 0.9$ that using $0.5t_{j,k}$ gives consistently the lowest RMSE among other choices of thresholds proportional to the $t_{j,k}$ values.

We found numerically that the decay of the RMSE with n almost perfectly fits a power-law, which (up to logarithmic factor) is in accordance with the upper-bounds of Corollary 2.1. Following this Corollary in the setting $\eta \in (\frac{1}{2}, 1)$ and $r = 2$, we fit a power law of the form

$$\mathbb{E}(\mathrm{RMSE}(n)) \approx n^{-\frac{\tilde{B}}{2(4\gamma + \tilde{B})}}.$$

We perform a linear regression in a log-log domain to estimate $\tilde{B}$. Table 2 reports the estimated value of $\tilde{B}$ we found using this procedure.

[1] http://www.ceremade.dauphine.fr/~peyre/codes/

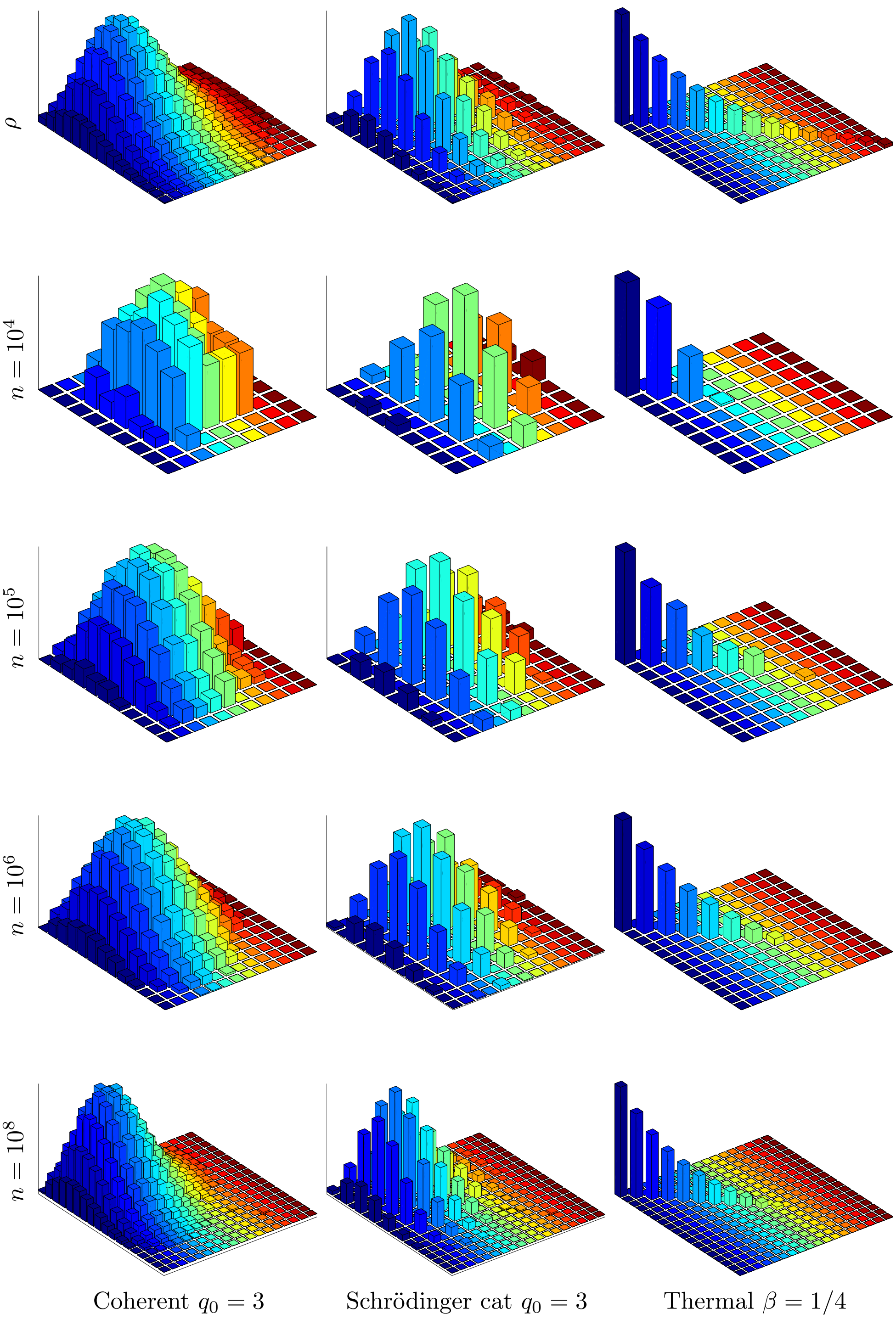
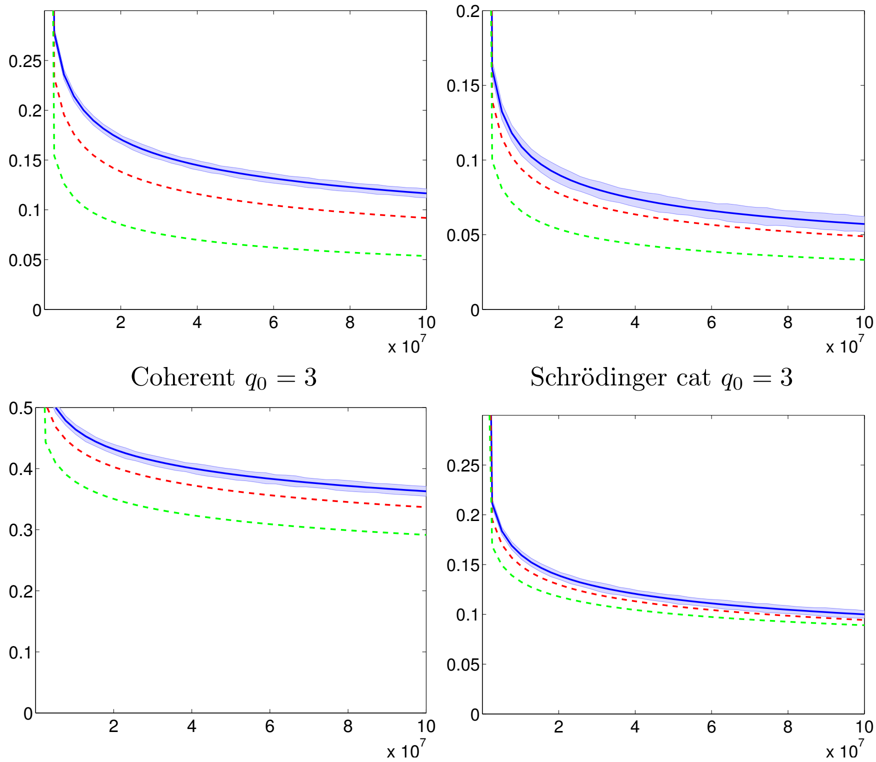


FIGURE 3. First row: ρ. Following rows: estimated $\tilde{\rho}^\eta$ for $B_0 = 0.5$, $\eta = 0.9$, $\varepsilon = 1$ and n respectively equal to 10^4 (row #2), 10^5 (row #3), 10^6 (row #4), 10^8 (row #5).



Coherent $q_0 = 3$ Schrödinger cat $q_0 = 3$

Thermal $\beta = 1/10$ Thermal $\beta = 1/4$

FIGURE 4. Blue curve: Evolution of $\mathbb{E}(\mathrm{RMSE}(n))$ as a function of n for $\eta = 0.9$, $\varepsilon = 1$. The blue shaded area represent the confidence interval at ± 3 times the standard deviation of $\mathrm{RMSE}(n)$. Red (reps. green) curve: evolution of $\mathbb{E}(\mathrm{RMSE}(n))$ obtained when replacing the threshold $t_{j,k}$ by $0.8 t_{j,k}$ (resp. $0.5 t_{j,k}$) in the estimator in (19).

TABLE 2. Estimated values of $\tilde{B}$ when using $\eta = 0.9$, $\varepsilon = 1$ and $N = 30$.

Coherent $q_0 = 3$	Schrödinger cat $q_0 = 3$	Thermal $\beta = 1/10$	Thermal $\beta = 1/4$
$\tilde{B} \approx 0.174$	$\tilde{B} \approx 0.227$	$\tilde{B} \approx 0.037$	$\tilde{B} \approx 0.082$

4. Proof of Theorem 2.1

The proof follow the main lines of [Alq08a, Alq08b]. First, we need a set of preliminary lemmas.

4.1. Some preliminary results.

Lemma 4.1. *For some fixed $\varepsilon \in (0,1)$, let us define the set*

$$\Omega_\varepsilon := \left\{ \forall (j,k) \in J(N), \quad \left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| \le t_{j,k} \right\},$$

where the $(t_{j,k})_{j,k}$ are defined in (20) *and the set $J(N)$ is defined in* (14). *Then we have*

$$P(\Omega_\varepsilon) \geq 1 - \varepsilon.$$

Proof. Lemma 4.1 is proved by using Hoeffding's inequality. In this aim, we have to first notice

$$E_\rho[\hat{\rho}^\eta_{j,k}] = \rho_{j,k}.$$

Indeed, by using (16) , (8), (11) and (3), we have

$$\begin{aligned}
E_\rho[\hat{\rho}^\eta_{j,k}] &= E_\rho[G_{j,k}(\frac{Y}{\sqrt{\eta}}, \Phi)] = E_\rho[f^\eta_{j,k}(\frac{Y}{\sqrt{\eta}})e^{-i(j-k)\Phi}] \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi} \int f^\eta_{j,k}(y)\sqrt{\eta}p^\eta_\rho(y\sqrt{\eta}|\phi)dyd\phi \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi}\frac{1}{2\pi}\int \widetilde{f}^\eta_{j,k}(t)\mathcal{F}_1[\sqrt{\eta}p^\eta_\rho(\cdot\sqrt{\eta}|\phi)](t)dtd\phi \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi}\frac{1}{2\pi}\int \widetilde{f}_{j,k}(t)e^{\gamma t^2}\mathcal{F}_1[p_\rho(\cdot|\phi)](t)\widetilde{N}^\eta(t)dtd\phi \\
&= \frac{1}{\pi}\int_0^\pi \int e^{-i(j-k)\phi} f_{j,k}(x)p_\rho(x|\phi)dxd\phi = \rho_{j,k}.
\end{aligned}$$

Moreover, we easily get from the definition of $G_{j,k}$ in (16) that for all $\ell = 1\cdots, n$ and $\forall (j,k) \in J(N)$

$$-\|f^\eta_{j,k}\|_\infty \leq G_{j,k}\left(\frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell\right) \leq \|f^\eta_{j,k}\|_\infty.$$

Then, for

$$t_{j,k} = \|f^\eta_{j,k}\|_\infty \sqrt{\frac{2\log\left(\frac{N(N+1)}{\varepsilon}\right)}{n}}$$

and according to Hoeffding's inequality,

$$\begin{aligned}
&P\left(\left|\hat{\rho}^\eta_{j,k} - \rho_{j,k}\right| \geq t_{j,k}\right) \\
&\quad = P\left\{\left|\frac{1}{n}\sum_{\ell=1}^n\left[G_{j,k}\left(\frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell\right) - E\left(G_{j,k}\left(\frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell\right)\right)\right]\right| \geq t_{j,k}\right\} \\
&\qquad \leq 2\exp\left[-\frac{nt^2_{j,k}}{2\|f^\eta_{j,k}\|^2_\infty}\right] = \frac{2\varepsilon}{N(N+1)}.
\end{aligned}$$

By the classical union bound argument:

$$P\left(\Omega^c_\varepsilon\right) \leq \sum_{(j,k)\in J(N)} P\left(\left|\hat{\rho}^\eta_{j,k} - \rho_{j,k}\right| \geq t_{j,k}\right) \leq \sum_{(j,k)\in J(N)} \frac{2\varepsilon}{N(N+1)} \leq \varepsilon.$$

□

Lemma 4.2. *For some fixed $\varepsilon \in (0,1)$ and $\forall (j,k) \in J(N)$, with $J(N)$ defined in* (14), *we define the set*

$$R^\varepsilon_{j,k} := \left\{\nu \text{ density matrix}, |\nu_{j,k} - \hat{\rho}^\eta_{j,k}| \leq t_{j,k}\right\},$$

where the $(t_{j,k})_{j,k}$ are defined in (20). *Then, on the event Ω_ε defined in Lemma 4.1 and $\forall (j,k) \in J(N)$*

1. $\rho \in R^\varepsilon_{j,k}$.
2. $R^\varepsilon_{j,k}$ *is closed and convex set.*

(3) *For $\Pi^{\varepsilon}_{j,k}$ the orthogonal projection onto $R^{\varepsilon}_{j,k}$ and for any density matrix ν,*

$$\left\|\rho - \Pi^{\varepsilon}_{j,k}(\nu)\right\|_2^2 \leq \|\rho - \nu\|_2^2 . \tag{24}$$

Proof. The first point is just a consequence of Lemma 4.1. The second point comes from the definition of $R^{\varepsilon}_{j,k}$.
Moreover, it is well known that for any closed and convex set $\mathcal{C}$, if $\Pi_{\mathcal{C}}$ is the orthogonal projection on $\mathcal{C}$, the following property holds:

$$\forall x \in \mathcal{C}, \forall y, \quad \|\Pi_{\mathcal{C}}(y) - x\|_2 \leq \|y - x\|_2.$$

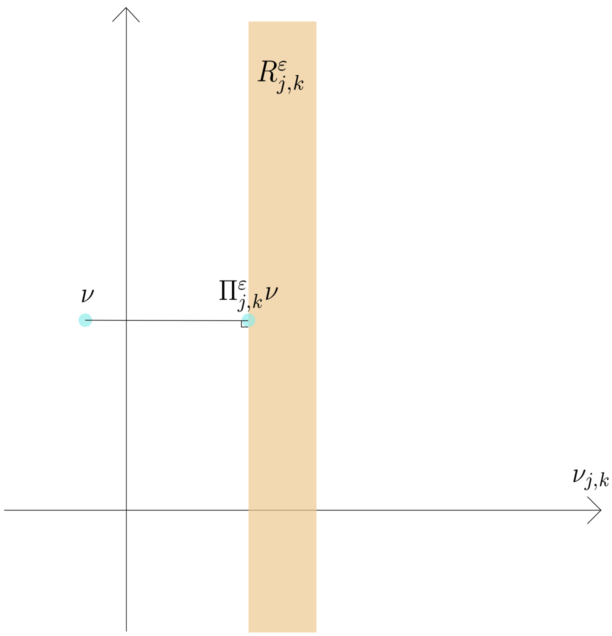
This conclude the proof of the third point. □

Lemma 4.3. *For $\varepsilon \in (0,1)$, any fixed $(j,k) \in J(N)$, with $J(N)$ defined in (14), and any density matrix ν, we denote by ν' the projection of ν into $R^{\varepsilon}_{j,k}$,*

$$\nu' := \Pi^{\varepsilon}_{j,k}(\nu) = [\nu'_{\ell,m}]_{\ell,m},$$

with $R^{\varepsilon}_{j,k}$ defined in Lemma 4.2. Then, the entries $\nu'_{\ell,m}$ of ν' are equal to

$$\nu'_{\ell,m} = \begin{cases} \nu_{j,k} + \operatorname{sgn}(\hat{\rho}^{\eta}_{j,k} - \nu_{j,k}) \left(\left| \hat{\rho}^{\eta}_{j,k} - \nu_{j,k} \right| - t_{j,k} \right)_+, & \textit{if } (\ell,m) = (j,k), \\ \nu_{\ell,m}, & \textit{otherwise.} \end{cases}$$



Proof. The projection ν' of ν into $R^{\varepsilon}_{j,k}$ satisfies

$$\nu' = \arg\min_{x \in R^{\varepsilon}_{j,k}} \|\nu - x\|_2^2 = \arg\min_{x \in R^{\varepsilon}_{j,k}} \sum_{\ell,m=0}^{\infty} |x_{\ell,m} - \nu_{\ell,m}|^2 .$$

As the constraint $x \in R^{\varepsilon}_{j,k}$ is only a constraint on $x_{j,k}$, it is clear that for $(\ell,m) \neq (j,k)$ the minimum is reached for $x_{j,k} = \nu_{j,k}$. Finally,

$$\nu'_{j,k} = \arg\min_{x_{j,k} : |x_{j,k} - \hat{\rho}^{\eta}_{j,k}| \leq t_{j,k}} (\nu_{j,k} - x_{j,k})^2 .$$

The solution $\nu'_{j,k}$ is obvious:

$$\nu'_{j,k} = \begin{cases} \hat{\rho}^{\eta}_{j,k} - t_{j,k} & \text{if } \nu_{j,k} < \hat{\rho}^{\eta}_{j,k} - t_{j,k}, \\ \nu_{j,k} & \text{if } \hat{\rho}^{\eta}_{j,k} - t_{j,k} \leq \nu_{j,k} \leq \hat{\rho}^{\eta}_{j,k} + t_{j,k}, \\ \hat{\rho}^{\eta}_{j,k} + t_{j,k} & \text{if } \nu_{j,k} > \hat{\rho}^{\eta}_{j,k} + t_{j,k} \end{cases}$$
$$= \nu_{j,k} + \operatorname{sgn}(\hat{\rho}^{\eta}_{j,k} - \nu_{j,k}) \left(\left| \hat{\rho}^{\eta}_{j,k} - \nu_{j,k} \right| - t_{j,k} \right)_{+}.$$

where the function $\operatorname{sgn}(\cdot)$ is as defined in (18) and $(\cdot)_{+}$ is defined in (17). □

Definition 4.1. *For $m > 0$ an integer, let*

$$I := \{(j_1, k_1), \ldots, (j_m, k_m)\} \subseteq J(N)$$

be a set of indices, where $J(N)$ is the set defined in (14). that $\forall \ell \neq i$, $(j_\ell, k_\ell) \neq (j_i, k_i)$. For $\varepsilon \in (0,1)$ and for any density matrix ν, we denote by $\Pi^{\varepsilon}_{I}(\nu)$ the successive projections of ν into spaces $\left(R^{\varepsilon}_{j_i,k_i} \right)_{j_i,k_i}$, i.e.

$$\Pi^{\varepsilon}_{I}(\nu) := \Pi^{\varepsilon}_{j_m,k_m} \Pi^{\varepsilon}_{j_{m-1},k_{m-1}} \ldots \Pi^{\varepsilon}_{j_2,k_2} \Pi^{\varepsilon}_{j_1,k_1}(\nu).$$

Note that for any set of indices I and from Lemma 4.3, the application of the successive projections Π^{ε}_{I} to a density matrix ν does not depend on the order of the successive projections.

Lemma 4.4. *For $\varepsilon \in (0,1)$, for $J(N)$ defined in (14) and for $\tilde{\rho}^{\eta}$ defined in (19), we have*

$$\tilde{\rho}^{\eta} = \Pi^{\varepsilon}_{J(N)}(\mathbf{0}),$$

where $\mathbf{0}$ is the zero-infinite matrix.

Proof. This is obvious from the definition of $\tilde{\rho}^{\eta}$ and from Lemma 4.3 applied to $\nu = \mathbf{0}$. □

4.2. **Proof of Theorem 2.1.**

Proof. For $J(N)$ the set of indices defined in (14), let I be a subset of $J(N)$, $I \subseteq J(N)$. For a fixed $\varepsilon \in (0,1)$, we have by Lemma 4.4 and by successive applications of the inequality(24) to all pair of indices $(j,k) \notin I$

$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 = \|\Pi^{\varepsilon}_{J}(\mathbf{0}) - \rho\|_2^2 \leq \|\Pi^{\varepsilon}_{I}(\mathbf{0}) - \rho\|_2^2. \tag{25}$$

Moreover, from Lemma 4.3 applied to $\nu = \mathbf{0}$, we get

$$(\Pi^{\varepsilon}_{I}(\mathbf{0}))_{j,k} = \begin{cases} \operatorname{sgn}(\hat{\rho}^{\eta}_{j,k}) \left(|\hat{\rho}^{\eta}_{j,k}| - t_{j,k} \right)_{+} & \text{if } (j,k) \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, from (25) we get

$$\begin{aligned} \|\tilde{\rho}^{\eta} - \rho\|_2^2 &\leq \sum_{j,k=0}^{\infty} \left[\rho_{j,k} - (\Pi^{\varepsilon}_{I}(\mathbf{0}))_{j,k} \right]^2 \\ &= \sum_{(j,k)\in I} \left[\rho_{j,k} - \operatorname{sgn}(\hat{\rho}^{\eta}_{j,k}) \left(|\hat{\rho}^{\eta}_{j,k}| - t_{j,k} \right)_{+} \right]^2 + \sum_{(j,k)\notin I} \rho^2_{j,k} \\ &:= \sum_{(j,k)\in I} \left[A_{j,k} \right]^2 + \sum_{(j,k)\notin I} \rho^2_{j,k}, \end{aligned} \tag{26}$$

where

$$\begin{aligned} A_{j,k} &= \rho_{j,k} - \mathrm{sgn}(\hat{\rho}^{\eta}_{j,k}) \left(|\hat{\rho}^{\eta}_{j,k}| - t_{j,k} \right)_{+} \\ &= \begin{cases} \rho_{j,k}, & \text{if } |\hat{\rho}^{\eta}_{j,k}| \leq t_{j,k}, \\ \rho_{j,k} - \hat{\rho}^{\eta}_{j,k} + t_{j,k}, & \text{if } \hat{\rho}^{\eta}_{j,k} > t_{j,k}, \\ \rho_{j,k} - \hat{\rho}^{\eta}_{j,k} - t_{j,k}, & \text{if } \hat{\rho}^{\eta}_{j,k} < t_{j,k}. \end{cases} \end{aligned}$$

Moreover

$$\begin{aligned} A_{j,k} \leq |A_{j,k}| &\leq \begin{cases} \left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| + |\hat{\rho}^{\eta}_{j,k}|, & \text{if } |\hat{\rho}^{\eta}_{j,k}| \leq t_{j,k}, \\ \left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| + t_{j,k}, & \text{if } \hat{\rho}^{\eta}_{j,k} > t_{j,k}, \\ \left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| + t_{j,k}, & \text{if } \hat{\rho}^{\eta}_{j,k} < t_{j,k}, \end{cases} \\ &\leq \left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| + t_{j,k}. \end{aligned}$$

For any $(j,k) \in I$ and on the event Ω_{ε} defined in Lemma 4.1, it holds

$$\left| \hat{\rho}^{\eta}_{j,k} - \rho_{j,k} \right| \leq t_{j,k}.$$

Therefore from (26)

$$\|\tilde{\rho}^{\eta} - \rho\|_2^2 \leq \sum_{(j,k) \in I} (2t_{j,k})^2 + \sum_{(j,k) \notin I} \rho^2_{j,k}.$$

We conclude the proof by taking the infimum over the set $I \subseteq J(N)$. □

References

[ABM09] J.-M. Aubry, C. Butucea, and K. Meziani. State estimation in quantum homodyne tomography with noisy data. *Inverse Problems*, 25(1), 2009.

[AGG05] L. Artiles, R. Gill, and M. Guţă. An invitation to quantum tomography. *J. Royal Statist. Soc. B (Methodological)*, 67:109–134, 2005. MR2136642.

[Alq08a] P. Alquier. Iterative feature selection in least square regression estimation. *Annales de l'IHP - B (Probablity and Statistics)*, 44(1):47–88, 2008.

[Alq08b] P. Alquier. Lasso, iterative feature selection and the correlation selector: Oracle inequalities and numerical performances. *Electronic Journal of Statistics*, 2:1129–1152, 2008.

[BDPS00] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi. Maximum-likelihood estimation of the density matrix. *Physical Review A*, 61(R10304), 2000.

[BNGJ03] O. E. Barndorff-Nielsen, R. Gill, and P. E. Jupp. On quantum statistical inference (with discussion). *J. Royal Stat. Soc. B*, 65:775–816, 2003.

[Cai99] T. Cai. Adaptive wavelet estimation: a block thresholding and oracle inequality approach. *The Annals of Statistics*, 27:2607–2625, 1999.

[Cav11] L. Cavalier. Inverse problems in statistics. In P. Alquier, E. Gautier, and G. Stoltz, editors, *Inverse Problems and High-Dimensional Estimation*. Springer Lecture Notes in Statistics, 2011.

[CGLT03] L. Cavalier, G. Golubev, O. Lepski, and A. Tsybakov. Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Theory of Probability and Applications*, 48:426–446, 2003.

[D'A95] G. M. D'Ariano. Tomographic measurement of the density matrix of the radiation field. *Quantum Semiclass. Optics*, 7:693–704, 1995.

[D'A02] G. M. D'Ariano. Tomographic methods for universal estimation in quantum optics. In *International School of Physics Enrico Fermi*, volume 148. IOS Press, 2002.

[DJ94] D. Donoho and I. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[DJ95] D. Donoho and I. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90:1200–1224, 1995.

[DJKP95] D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: asymptotia? *Journal of the Royal Statistical Society B*, 57:301–369, 1995.

[DLP95] G. M. D'Ariano, U. Leonhardt, and H. Paul. Homodyne detection of the density matrix of the radiation field. *Phys. Rev. A*, 52:R1801–R1804, 1995.

[DMP94] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris. Detection of the density matrix through optical homodyne tomography without filtered back projection. *Phys. Rev. A*, 50:4298–4302, 1994.

[DMS05] G.M. D'Ariano, L. Maccone, and M. F. Sacchi. Homodyne tomography and the reconstruction of quantum states of light, 2005.

[Don95] D. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Applied Computationnal Harmonic Analysis*, 2:101–126, 1995.

[DP99] G. M. D'Ariano and M. G. A. Paris. Adaptive quantum tomography. *Physical Review A*, 60(518), 1999.

[Guţ07] M. I. Guţă. Maximum likelihood estimation of the density matrix through quantum tomography. Manuscript, 2007.

[Hel76] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.

[HKPT98] W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov. *Wavelets, approximation, and statistical applications*. Springer, 1998.

[Hol82] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland, 1982.

[Joh99] I. Johnstone. Wavelet shrinkage for correlated data and inverse problems: adaptivity results. *Statistica Sinica*, 9:51–83, 1999.

[Kol96] E. Kolaczyk. A wavelet shrinkage approach to tomographic image reconstruction. *Journal of the American Statistical Association*, 91:1079–1090, 1996.

[Leo97] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, 1997.

[LPD95] U. Leonhardt, H. Paul, and G. M. D'Ariano. Tomographic reconstruction of the density matrix via pattern functions. *Phys. Rev. A*, 52:4899–4907, 1995.

[Mez08] K. Meziani. Estimation non-paramétrique en tomographie quantique homodyne. PhD thesis, Université Paris 7 (in English), 2008.

[OTBLG06] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier. Generating optical SchrÃűdinger kittens for quantum information processing. *Science*, 312(5770):83–86, 2006.

[Ric96] T. Richter. Pattern functions used in tomographic reconstruction of photon statistics revisited. *Phys. Lett. A*, 211:327–330, 1996.

[Ric00] T. Richter. Realistic pattern functions for optical homodyne tomography and determination of specific expectation values. *Physical Review A*, 61(063819), 2000.

[SBRF93] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.*, 70:1244–1247, 1993.

[VR89] K. Vogel and H. Risken. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 40:2847–2849, 1989.

[Wig32] E. Wigner. On the quantum correction for thermodynamic equations. *Phys. Rev.*, 40:749–759, 1932.