



**HAL**  
open science

# Estimations et tests non paramétriques en tomographie quantique homodyne

Katia Meziani

► **To cite this version:**

Katia Meziani. Estimations et tests non paramétriques en tomographie quantique homodyne. Mathématiques [math]. Université Paris-Diderot - Paris VII, 2008. Français. NNT : . tel-00351294

**HAL Id: tel-00351294**

**<https://theses.hal.science/tel-00351294>**

Submitted on 8 Jan 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS VII - DENIS DIDEROT  
UFR DE MATHÉMATIQUES

THÈSE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ PARIS VII

Spécialité : Mathématiques Appliquées

présentée par

**Katia MEZIANI**

---

ESTIMATIONS ET TESTS NON PARAMÉTRIQUES EN  
TOMOGRAPHIE QUANTIQUE HOMODYNE

---

Soutenue publiquement le 9 Décembre 2008, devant le jury composé de :

Mme	<b>Cristina</b>	<b>BUTUCEA</b>	Université Lille 1	(Directrice)
M.	<b>Laurent</b>	<b>CAVALIER</b>	Université Aix-Marseille 1	(Rapporteur)
M.	<b>Richard</b>	<b>GILL</b>	Université de Leiden	(Rapporteur)
M.	<b>Madalin</b>	<b>GUTA</b>	Université de Nottingham	
M.	<b>Marc</b>	<b>HOFFMANN</b>	ENSAE	
Mme	<b>Dominique</b>	<b>PICARD</b>	Université Paris 7	(Directrice)
M.	<b>Alexandre</b>	<b>TSYBAKOV</b>	CREST	(Président)



La partie *Remerciements* n'est pas une tâche facile, mais elle est néanmoins une des plus agréables. C'est l'occasion pour moi de remercier toutes celles et ceux qui m'ont soutenue de quelque manière que ce soit dans cette grande aventure qu'est la thèse.

Mes premiers remerciements vont tout naturellement à Cristina Butucea qui m'a encadrée pendant trois ans. Je voudrais lui témoigner ma gratitude. Elle a fait preuve de pédagogie, de patience et m'a offert beaucoup de son temps. Chercheuse brillante, j'ai énormément appris à ses côtés. Je lui dois mes premiers pas dans la recherche.

Je profite de cette préface pour exprimer toute l'admiration que j'ai pour Dominique Picard. Elle a toujours pris le temps de me conseiller et de m'écouter avec beaucoup de bienveillance. Je tiens également à remercier Alexandre Tsybakov avec qui j'ai découvert les statistiques en maîtrise et fait mon mémoire de DEA. J'ai pour lui le plus grand respect que ce soit pour ses qualités humaines ou pour la qualité de ses recherches. C'est un grand honneur pour moi aujourd'hui de les compter parmi les membres de mon jury.

Je remercie tout particulièrement Laurent Cavalier et Richard Gill pour avoir accepté de rapporter ma thèse et ce dans un délai très court. Je remercie également Marc Hoffman qui me fait l'honneur d'être parmi les membres du jury et Madalin Guta qui m'a permis d'avancer dans mes recherches par ses remarques et ses conseils judicieux.

Je remercie Jean-Marie Aubry avec qui j'ai eu le plus grand plaisir à collaborer. Tout au long de ma thèse, j'ai toujours pu compter sur Stéphane Boucheron et je le remercie ici pour son soutien. Je n'oublie pas celui qui parle à l'oreille des ordinateurs, Erwan Lepennec, merci pour ta bonne humeur. Lors de mes trois années de monitorat, j'ai travaillé en collaboration avec René Guitart et Gabrielle Viennet que je remercie sincèrement pour leur soutien et leur sympathie. Cela a été une expérience réellement enrichissante qui m'a conforté dans l'idée que je me faisais de l'enseignement.

Un atout essentiel pour un thésard, c'est l'équipe administrative ! Alors je n'oublie pas de remercier très chaleureusement Michèle Wasse qui est une vraie mère pour nous, Laure Elie pour tout le temps qu'elle nous consacre, Pascal Chiettini pour sa disponibilité, Isabelle Mariage et Valerie Juvé pour leur efficacité, Jacques Portes toujours prêt à nous rendre service, Hubert Andrique à la reprographie pour sa gentillesse, et tous les autres qui dans l'ombre nous aident à venir à bout de la machine administrative.

Je ne peux oublier de remercier sincèrement mes deux complices et amis, Arnak Dalalyan et Mohamed Hebiri, pour leur soutien sans faille dans les moments durs et les parties de franches rigolades. Un grand merci à Arnaud Siri-Jegousse, Vathana LyVath, Thomas Willer, Joseph Salmon, Sisi Yé et Adrien Saumard (Se Marre) qui en plus de leur amitiés, m'ont offert par certains de relire ce manuscrit. Mes pensées vont également aux rencontres faites à Chevaleret, en particulier, à Frédéric Guilloux, Christophe Chesneau et ceux du bureau 5B01, avec qui j'ai passé beaucoup de bons moments, et tous les autres qui se reconnaîtront. Une dédicasse à Kamel Assam et Zaid Harchaoui qui m'ont également soutenue ces derniers jours. Je n'oublie pas tous mes amis et il y en a beaucoup ! Alors, ne m'en voulez pas si je ne vous ai pas cité par vos prénoms. Mais, merci de m'avoir permis de décompresser quand j'en avais besoin...

Enfin, ma grande famille, je vous remercie tous pour l'amour dont vous faites preuve

chaque jour à mon égard, vous êtes mon équilibre et ma force. Un petit clin d'oeil à Mathys mon neveu, qui est arrivé parmi nous il n'y a que quelques heures, au moment où j'écris ces mots. Félicitation à Bibichat et Cédric, il est magnifique Mathys, on dirait un ange !  
Je finis par toi, *nounge*, toujours à mes cotés. Merci de ton soutien mais surtout de ta patience durant ces très nombreuses années.



*A Mathys, un ange qui vient de naître...*





# TABLE DES MATIÈRES

<b>1</b>	<b>Introduction</b>	<b>16</b>
1.1	Introduction à l'optique quantique . . . . .	16
1.1.1	Notation de Dirac . . . . .	16
1.1.2	Formalisme de la mécanique quantique . . . . .	19
1.1.3	Formalisme de l'optique quantique . . . . .	27
1.1.4	Exemples d'états quantiques . . . . .	39
1.2	Formalisme statistique . . . . .	43
1.2.1	Modèle statistique non-paramétrique . . . . .	43
1.2.2	Problèmes inverses . . . . .	47
1.2.3	Modèle statistique de la TQH . . . . .	49
1.3	Description du contenu de la thèse . . . . .	50
1.3.1	Estimation de la matrice de densité . . . . .	51
1.3.2	Estimation de la fonction de Wigner . . . . .	51
1.3.3	Estimation de la pureté . . . . .	52
1.3.4	Test d'adéquation . . . . .	53
1.3.5	Perspectives . . . . .	54
<b>2</b>	<b>Nonparametric estimation of the density matrix</b>	<b>58</b>
2.1	Introduction . . . . .	58
2.1.1	Physical background . . . . .	59
2.1.2	Statistical model . . . . .	60
2.1.3	Outline of the results . . . . .	61
2.2	Pattern functions . . . . .	62
2.3	Density matrix estimation . . . . .	63
2.3.1	Estimation procedure . . . . .	63

2.3.2	Main results . . . . .	65
2.4	Numerical results . . . . .	67
2.5	Proof . . . . .	70
2.5.1	Proof of Proposition 2.1 . . . . .	70
2.5.2	Proof of Proposition 2.2 . . . . .	70
2.5.3	Proof of Proposition 2.3 . . . . .	71
2.6	Proof of Lemmata . . . . .	73
<b>3</b>	<b>Nonparametric Wigner function estimation</b>	<b>76</b>
3.1	Introduction . . . . .	77
3.1.1	Physical background . . . . .	77
3.1.2	Statistical model . . . . .	79
3.1.3	Outline of the results . . . . .	80
3.2	Decrease and smoothness of the Wigner distribution . . . . .	81
3.3	Wigner function estimation . . . . .	82
3.3.1	Kernel estimator . . . . .	82
3.3.2	$\mathbb{L}_2$ risk estimation . . . . .	83
3.4	Other Proofs . . . . .	86
<b>4</b>	<b>Nonparametric estimation of the purity</b>	<b>92</b>
4.1	Introduction . . . . .	93
4.2	Physical and statistical context . . . . .	95
4.2.1	A short introduction to Quantum Mechanics . . . . .	95
4.2.2	Quantum Homodyne Tomography . . . . .	97
4.2.3	Examples . . . . .	98
4.2.4	Problem formulation . . . . .	99
4.3	Estimation procedure and main results . . . . .	100
4.3.1	Kernel estimator . . . . .	100
4.3.2	Bias-variance decomposition . . . . .	101
4.3.3	Main results . . . . .	101
4.4	Proof of Proposition 1 . . . . .	105
4.4.1	Proof of (4.9) . . . . .	105
4.4.2	Proof of (4.10) . . . . .	108

<b>5</b>	<b>Nonparametric goodness-of fit testing</b>	<b>114</b>
5.1	Physical background . . . . .	115
5.1.1	Short introduction to quantum optics . . . . .	115
5.1.2	Fock basis and <i>pattern</i> functions . . . . .	117
5.1.3	Examples of quantum states . . . . .	118
5.2	Problem formulation . . . . .	119
5.2.1	Statistical model . . . . .	120
5.2.2	Statistical tests . . . . .	121
5.2.3	Outline of results . . . . .	122
5.3	Testing procedure and results . . . . .	123
5.3.1	Pattern functions . . . . .	123
5.3.2	Testing procedure . . . . .	124
5.3.3	Properties of the estimator $M_n$ . . . . .	125
5.3.4	Main results . . . . .	127
5.4	Simulations . . . . .	129
5.4.1	Pattern functions $f_{j,k}^\eta$ . . . . .	130
5.4.2	Implementation of $M_n$ . . . . .	131
5.4.3	Studies of the performance of our test procedure $\Omega_n$ . . . . .	133
5.5	Proof . . . . .	136
5.5.1	Proof of the Theorems . . . . .	136
5.5.2	Proof of the Propositions . . . . .	138
5.6	Appendix . . . . .	140



# SOMMAIRE

Le monde quantique est riche et surprenant, si petit et pourtant plein de promesses! Grâce aux progrès technologiques des dernières décennies, le domaine de la mécanique quantique connaît actuellement un essor très important. La variété des problèmes théoriques posés amène de nombreux mathématiciens à s'y intéresser. Ainsi, on assiste aujourd'hui à une floraison de travaux liés à ce domaine, entre autres sur les probabilités non commutatives, la géométrie non commutative, l'informatique quantique et les statistiques quantiques.

En mécanique quantique, on décrit tous les aspects d'un système physique par la donnée de son **état quantique** et toute propriété ou grandeur physique (la position, l'énergie cinétique,...) est appelée **observable**. À l'inverse de la mécanique classique, l'état d'un système physique n'est en général jamais totalement mesurable et la mesure d'une observable d'un système dans un état donné connu peut donner aléatoirement une valeur prise dans un ensemble de résultats possibles. De plus, au cours de la mesure, le système quantique voit son état modifié. Ce phénomène est inhérent à la mesure et ne dépend pas du soin que l'expérimentateur prend à ne pas "déranger" le système.

Le système quantique considéré dans le cadre de ce manuscrit est un faisceau de lumière monochromatique. L'état quantique détermine la distribution de probabilité associée à chacune des observables du système et sa connaissance nous permet de faire des prévisions probabilistes sur les résultats de mesure. Mathématiquement, l'état quantique de la lumière peut être représenté par sa **matrice de densité** ou de façon équivalente par la **fonction de Wigner** associée. Statistiquement, cette thèse traite, d'un point de vue non paramétrique, de l'étude d'un double problème inverse posée par l'optique quantique. Aux observations aléatoires issues de la mesure tomographique se rajoutent des erreurs indépendantes, de loi normale de variance connue par les physiciens appelée efficacité.

Le Chapitre 1 pose le cadre physique et statistique relatif à ce travail de thèse. Dans le Chapitre 2, on se propose d'estimer, dans des classes naturelles en physique, la matrice de densité  $\rho$  par un estimateur de projection sur des fonctions dont la signification physique

est connue. Au Chapitre 3, les propriétés des fonctions de Wigner dans ces classes sont étudiées et un estimateur de type noyau de ces fonctions de Wigner est donné. L'estimation d'une fonctionnelle quadratique de la fonction de Wigner par une méthode de type noyau est traitée au Chapitre 4. Cette fonctionnelle quadratique permet de quantifier la pureté d'un état quantique. Dans le Chapitre 5, des tests d'adéquation non paramétrique à la matrice de densité  $\rho$  sont mis en place. Les notions de physique et/ou de statistiques, nécessaires à la compréhension de chacun des différents travaux présentés dans cette thèse, étant reprises et résumées en chaque début de chapitre, chacun d'eux peut être lu indépendamment des autres..







# Chapitre 1

## Introduction

Le Chapitre 1, introduction générale cette thèse, se divise en trois grandes sections. Une large Section 1.1 est consacrée à la présentation de notions touchant aux fondements de la mécanique quantique et plus précisément du cadre physique de ce manuscrit qu'est l'optique quantique. Quelques notions de statistique non-paramétrique sont également données à la Section 1.2. La Section 1.3 comporte une description et une discussion de ce travail de thèse ainsi que la présentation des perspectives découlant de ces travaux.

### 1.1 Introduction à l'optique quantique

Dans une première Section 1.1.1, nous développons le formalisme conçu par Paul Dirac et universellement utilisé en physique quantique. La Section 1.1 a pour but de familiariser le lecteur avec quelques notions touchant aux fondements de la mécanique quantique. Le cadre physique de cette thèse, détaillé dans la Section 1.1.3, est l'optique quantique qui constitue la branche commune de la mécanique quantique et de l'optique. Nous présentons dans la Section 1.1.4 quelques exemples d'états quantiques. Pour plus de détails, nous renvoyons les lecteurs intéressés par la statistique quantique aux livres de Helstrom [46], Holevo [47] et à l'article Barndorff-Nielsen, Gill, et Jupp [11] pour l'inférence quantique ; et pour l'optique quantique le livre de Leonhardt [65] et l'article de Artiles, Gill, et Guță [4].

#### 1.1.1 Notation de Dirac

La notation *bra-ket* a été introduite par Paul Dirac pour faciliter l'écriture des équations de la mécanique quantique. Nous verrons que cette notation présente d'énormes commodités.

### Vecteurs ket/bra

Soit  $\mathcal{H}$  un espace complexe de Hilbert séparable, on définit d'abord un élément de l'espace  $\mathcal{H}$ , le vecteur appelé vecteur *ket* et noté :  $|\psi\rangle$ . Puis, à chaque vecteur ket  $|\psi\rangle$  on fait correspondre son vecteur dual, appelé vecteur *bra* et noté :  $\langle\psi|$ .

Le *produit scalaire* de deux vecteurs kets  $|\psi\rangle$  et  $|\varphi\rangle$  associé à l'espace  $\mathcal{H}$  est noté :  $\langle\psi|\varphi\rangle$  et

$$\langle\psi|\varphi\rangle^* = \langle\varphi|\psi\rangle.$$

Comme tout produit scalaire complexe, ce produit est sesquilinéaire :

$$\begin{aligned} \text{pour } \lambda_1, \lambda_2 \in \mathbb{C} : \quad & \langle\varphi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1\langle\varphi|\psi_1\rangle + \lambda_2\langle\varphi|\psi_2\rangle, \\ & \langle\lambda_1\varphi_1 + \lambda_2\varphi_2|\psi\rangle = \lambda_1^*\langle\varphi_1|\psi\rangle + \lambda_2^*\langle\varphi_2|\psi\rangle. \end{aligned}$$

Le produit scalaire induit une norme sur  $\mathcal{H}$  :  $\|\psi\|^2 = \langle\psi|\psi\rangle$ .

### Base

Soit une base orthonormée complète  $\{|u_i\rangle\}_i$  de  $\mathcal{H}$ . Par définition, elle est telle que

- $\langle u_i|u_j\rangle = \delta_{ij}$ , où  $\delta_{ij}$  est le symbole de Kronecker ( $\delta_{ij} = 1$  si  $i = j$  et  $\delta_{ij} = 0$  sinon),
- chaque vecteur ket peut s'écrire de façon unique

$$|\psi\rangle = \sum_i a_i|u_i\rangle = \sum_i \langle u_i|\psi\rangle|u_i\rangle.$$

De la même façon  $\{\langle u_i|\}_i$  constitue une base orthonormée complète pour les bras.

Le symbole  $|u_i\rangle\langle u_i|$  dénote l'opérateur de projection sur  $|u_i\rangle$ , dans le sens où

$$|u_i\rangle\langle u_i|\psi\rangle = \langle u_i|\psi\rangle|u_i\rangle = a_i|u_i\rangle.$$

De plus,

$$\sum_i |u_i\rangle\langle u_i| = \mathbf{1}. \tag{1.1}$$

où  $\mathbf{1}$  est l'opérateur identité de  $\mathcal{H}$ .

Soient  $\{|u_i\rangle\}_i$  une base orthonormée complète de l'espace de Hilbert  $\mathcal{H}$  et  $|\psi\rangle$  un ket de  $\mathcal{H}$  tel que  $|\psi\rangle = \sum_i a_i|u_i\rangle$ , on dit que  $|\psi\rangle$  est normalisé s'il est de norme 1 :

$$\langle\psi|\psi\rangle = \sum_i a_i^* a_i = 1.$$

### Opérateurs/matrice

Soit  $\{u_i\}_i$  une base orthonormée complète de l'espace de Hilbert  $\mathcal{H}$ , les opérateurs linéaires agissant sur l'espace  $\mathcal{H}$  peuvent s'écrire sous la forme d'une combinaison linéaire d'opérateurs  $|u_i\rangle\langle u_j|$ . D'une façon générale, si  $A$  est un opérateur linéaire agissant sur  $|\psi\rangle = \sum_j a_j |u_j\rangle$ , un ket quelconque, alors par (1.1)

$$\begin{aligned} A|\psi\rangle &= \mathbf{1}A|\psi\rangle = \sum_{i,j} a_j |u_i\rangle \langle u_i|A|u_j\rangle = \sum_{i,j} a_j \langle u_i|A|u_j\rangle |u_i\rangle \\ &= \sum_{i,j} \langle u_i|A|u_j\rangle a_j |u_i\rangle = \sum_{i,j} A_{i,j} a_j |u_i\rangle. \end{aligned}$$

Donc l'opérateur  $A$  transforme le ket  $|\psi\rangle$  en un autre ket  $|\psi'\rangle = A|\psi\rangle = \sum_i b_i |u_i\rangle$ . Par identification  $b_i = \sum_j A_{i,j} a_j$ . L'ensemble des nombres  $A_{i,j} = \langle u_i|A|u_j\rangle$  constitue les éléments de la matrice représentative de l'opérateur  $A$ . Avec ce formalisme, si  $\sum_i |\langle u_i|A|u_i\rangle| < \infty$  la trace de  $A$  définie comme la somme des éléments diagonaux de sa matrice, s'écrit

$$\text{Tr}(A) = \sum_i \langle u_i|A|u_i\rangle \quad (1.2)$$

et ne dépend pas de la base choisie.

### Vecteurs propres/valeurs propres

Soit  $A$  un opérateur et  $|\alpha\rangle$  un ket tel que

$$A|\alpha\rangle = \alpha|\alpha\rangle,$$

où  $\alpha$  est un nombre complexe. On dit que  $|\alpha\rangle$  est un vecteur propre de l'opérateur  $A$  et  $\alpha$  la valeur propre correspondante.

#### **Définition 1.1.**

Notons  $A^\dagger$  l'adjoint de l'opérateur  $A$ , défini tel que

$$\langle u|Av\rangle = \langle A^\dagger u|v\rangle, \quad \forall u, v \in \mathcal{H}.$$

Dans une base orthonormée, si la matrice d'un opérateur  $A$  est telle que

$$A = A^\dagger,$$

alors l'opérateur  $A$  est un opérateur hermitien (auto-adjoint).

Un opérateur  $A$  est compact si l'adhérence de l'image de la boule unité par  $A$  est compact.

**Propriétés 1.1.**

Soit  $A$  un opérateur compact et hermitien alors,

1. les valeurs propres de  $A$  sont réelles,
2. les vecteurs propres de  $A$  associés à des valeurs propres distinctes sont orthogonaux,
3. il est toujours possible de trouver, dans le cas d'un espace de Hilbert de dimension  $d$ ,  $d$  vecteurs propres orthogonaux, correspondant chacun à une valeur propre.

**Définition 1.2.**

On définit le commutateur de deux opérateurs  $A$  et  $B$  par

$$[A, B] = AB - BA. \quad (1.3)$$

Si  $[A, B] = 0$ , on dit que  $A$  et  $B$  commutent.

**1.1.2 Formalisme de la mécanique quantique**

Un état est une description complète d'un système physique. En mécanique quantique, l'état d'un système physique peut être décrit par un vecteur  $|\psi\rangle$ , normalisé, d'un espace complexe de Hilbert séparable  $\mathcal{H}$ . L'espace  $\mathcal{H}$  est appelé espace des états et le vecteur  $|\psi\rangle$  normalisé est appelé état **pur**. On peut de façon équivalente décrire l'état d'un système physique par un opérateur auto-adjoint  $\rho$  vérifiant certaines propriétés et agissant sur l'espace  $\mathcal{H}$ . La représentation en terme d'opérateur permet la description d'un plus grand nombre de situations rencontrées en physique. Pour chacune de ces représentations nous définissons les notions d'observables et de mesure, qui sont indissociables. Dans cette partie, nous donnerons des exemples dans le cas fini dimensionnel, le cas infini dimensionnel (cadre de la thèse) étant détaillé dans la prochaine Section 1.1.3.

**A/ Etat quantique pur**

Dans le formalisme de la mécanique quantique, un système physique est dans un état pur s'il est décrit complètement par un vecteur d'état  $|\psi\rangle$  normalisé appartenant à un espace de Hilbert (complexe, séparable)  $\mathcal{H}$ . Plaçons nous dans le cas fini dimensionnel où  $\mathcal{H} = \mathbb{C}^d$  est un espace de Hilbert de dimension  $d$ .

Exemple du qubit :

Le plus petit espace de Hilbert non trivial de dimension deux est  $\mathbb{C}^2$  ( $d = 2$ ). Soit  $\{|0\rangle, |1\rangle\}$

une base orthonormée pour un tel espace, appelée *base standard*. Dans cette base, l'état quantique  $|\psi\rangle$  le plus général peut être exprimé comme :

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

où  $\alpha, \beta$  sont des nombres complexes tels que  $|\alpha|^2 + |\beta|^2 = 1$ . On appelle **qubit** n'importe quel état de cette forme.

### Observables

A toute propriété ou toute grandeur physique observable (position, énergie...) correspond un opérateur auto-adjoint linéaire agissant sur les vecteurs de l'espace des états  $\mathcal{H}$ . Cet opérateur est appelé **observable** et possède une représentation spectrale. Soit  $\mathbf{X}$  une telle observable alors de la Définition 1.1 et des propriétés données par la Propriété 1.1 découlent les propriétés suivantes pour l'observable  $\mathbf{X}$ .

#### **Propriétés 1.2.**

Soit  $\mathbf{X}$  une observable, alors

1. les valeurs propres de  $\mathbf{X}$  sont des nombres réels notés :  $\{x_a\}_a$ ,
2. les vecteurs propres correspondant à des valeurs propres distinctes sont orthogonaux,
3. il est toujours possible de trouver une base orthonormée complète de  $\mathcal{H}$  constituée de l'ensemble des vecteurs propres d'un tel opérateur.

Les vecteurs propres  $|x_a\rangle$  sont appelés **états propres**. Soit  $\mathbf{X}$  une observable, agissant sur  $\mathcal{H}$ , dont le spectre  $\{x_a\}_a$  est discret et les vecteurs propres  $\{|x_a\rangle\}_a$  forment une base orthonormée complète de  $\mathcal{H}$  :

$$\langle x_a | x_{a'} \rangle = \delta_{a'a}, \quad \sum_{a=1}^d |x_a\rangle \langle x_a| = \mathbf{1},$$

où  $\mathbf{1}$  est l'opérateur identité. Nous pouvons alors écrire

$$\mathbf{X} = \sum_{a=1}^d x_a |x_a\rangle \langle x_a| = \sum_{a=1}^d x_a \mathbf{P}_a,$$

où  $\mathbf{P}_a = |x_a\rangle \langle x_a|$  est le projecteur orthogonal sur l'espace engendré par le vecteur propre  $|x_a\rangle$ . Il vérifie les propriétés suivantes.

#### **Propriétés 1.3.**

Soit  $\mathbf{P}_a = |x_a\rangle \langle x_a|$  le projecteur associé à la valeur propre  $x_a$ , alors

- $\mathbf{P}_a = \mathbf{P}_a^\dagger = \mathbf{P}_a^2$ ,
- $\mathbf{1} - \mathbf{P}_a$  est encore un opérateur de projection,
- $\sum_{a=1}^d \mathbf{P}_a = \mathbf{1}$ .

D'autre part, la Propriété 1.2.(3) signifie qu'un état quantique normalisé  $|\psi\rangle$  de l'espace des états  $\mathcal{H}$  peut être décomposé de manière unique sur la base  $\{|x_a\rangle\}_a$  :

$$|\psi\rangle = \sum_{a=1}^d c_a |x_a\rangle,$$

avec  $c_a = \langle x_a | \psi \rangle$  vérifiant

$$\begin{aligned} \sum_{a=1}^d |c_a|^2 &= \sum_{a=1}^d |\langle x_a | \psi \rangle|^2 = \sum_{a=1}^d \langle \psi | \mathbf{P}_a | \psi \rangle \\ &= \langle \psi | \sum_{a=1}^d \mathbf{P}_a | \psi \rangle = \langle \psi | \mathbf{1} | \psi \rangle = \langle \psi | \psi \rangle = 1. \end{aligned} \quad (1.4)$$

### Mesure

En mécanique quantique, les résultats possibles de la **mesure** d'une grandeur physique représentée par l'observable  $\mathbf{X}$  sont les valeurs propres de  $\mathbf{X}$ , autrement dit le spectre  $\{x_a\}_a$  de  $\mathbf{X}$ . Les vecteurs propres  $\{|x_a\rangle\}_a$ , représentent l'état quantique du système lors de cette mesure.

La mesure de l'observable  $\mathbf{X}$ , effectuée sur l'état normalisé  $|\psi\rangle$  donne la valeur  $x_a$  avec la probabilité  $|c_a|^2 = \langle \psi | \mathbf{P}_a | \psi \rangle$ . En d'autres termes, le résultat d'une mesure de l'observable  $\mathbf{X}$  sur l'état  $|\psi\rangle$  fournit une variable aléatoire  $X$  à valeurs dans l'ensemble des valeurs propres  $\{x_a\}_{a=1}^d$ . On obtient le résultat  $x_a$ , avec la probabilité

$$P_\psi(X = x_a) = |c_a|^2 = \langle \psi | \mathbf{P}_a | \psi \rangle. \quad (1.5)$$

Il s'agit bien d'une mesure de probabilité, puisque les probabilités sont positives et d'après (1.4)

$$\sum_{a=1}^d P_\psi(X = x_a) = \sum_{a=1}^d |c_a|^2 = 1.$$

Il est intéressant de noter qu'une fois cette mesure effectuée, le système physique mesuré sera "projeté" de manière aléatoire dans l'un des états propres  $|x_a\rangle$  définis par l'observable  $\mathbf{X}$  et celui-ci doit rester le même si on applique de nouveau le même opérateur de mesure. La

probabilité de trouver, comme résultat d'une seconde application de l'opérateur, un autre vecteur propre  $|x_{a'}\rangle$  avec  $a \neq a'$  doit donc être nulle.

La valeur moyenne de la mesure d'une telle observable  $\mathbf{X}$  appliquée à cet état quantique  $|\psi\rangle$  est notée  $\langle\mathbf{X}\rangle$ , telle que

$$\langle\mathbf{X}\rangle = \langle\psi|\mathbf{X}|\psi\rangle = \langle\psi|\left(\sum_{a=1}^d x_a \mathbf{P}_a\right)|\psi\rangle = \sum_{a=1}^d x_a \langle\psi|\mathbf{P}_a|\psi\rangle = \sum_{a=1}^d x_a P_\psi(X = x_a). \quad (1.6)$$

Il s'agit de  $E_\psi[X]$ , l'espérance de la variable aléatoire  $X$  résultant de la mesure de l'observable  $\mathbf{X}$  sur l'état  $|\psi\rangle$ .

## B/ Etat quantique représenté par un opérateur $\rho$

La description d'un état par un ket normalisé  $|\psi\rangle$  permet de définir les états purs, mais elle est loin de recouvrir toutes les situations que l'on rencontre en physique. Les physiciens ont recours à la description en terme d'**opérateur de densité**  $\rho$ . On définit l'opérateur de densité  $\rho$  pour les états purs  $|\psi\rangle$  comme

$$\rho = |\psi\rangle\langle\psi|.$$

La représentation de cet opérateur de densité sur une base orthonormée  $\{|u_i\rangle\}_i$  est appelée **matrice de densité** et s'écrit

$$\rho = |\psi\rangle\langle\psi| = \sum_{i,j} c_j^* c_i |u_i\rangle\langle u_j|, \quad \text{avec} \quad \sum_i |c_i|^2 = 1.$$

Ainsi, les entrées de la matrice de densité d'un état pur s'écrivent sur la base  $\{|u_i\rangle\}_i$

$$\rho_{i,j} = \langle u_i|\rho|u_j\rangle = c_j^* c_i. \quad (1.7)$$

### Opérateur de densité

L'opérateur de densité ne se restreint pas à la représentation des états purs, il permet de décrire une classe plus large d'objets quantiques (les états mixtes). Il rend compte des imperfections lors de la génération des états quantiques qui peuvent préparer un mélange, une superposition statistique de différents états purs  $\{|\psi_a\rangle\}_{a \in \mathbb{N}^*}$  avec des probabilités associées  $\{p_a\}_{a \in \mathbb{N}^*}$ . Un tel système (superposition statistique d'états purs) peut être également défini en terme d'opérateur de densité

$$\rho = \sum_a p_a |\psi_a\rangle\langle\psi_a|, \quad (1.8)$$

où

- les vecteurs d'états purs  $|\psi_a\rangle$  sont normalisés,
- les probabilités  $p_a$  sont normalisées :  $\sum_a p_a = 1$ .

La représentation de cet opérateur sur la base  $\{|u_i\rangle\}_i$ , et sur une base de manière générale, est la matrice de densité

$$\rho = \sum_a p_a |\psi_a\rangle \langle \psi_a| = \sum_a \sum_{i,j} p_a c_{j,a}^* c_{i,a} |u_i\rangle \langle u_j|,$$

et ses entrées s'écrivent :

$$\rho_{i,j} = \langle u_i | \rho | u_j \rangle = \sum_a p_a c_{j,a}^* c_{i,a}. \quad (1.9)$$

#### Propriétés 1.4.

L'opérateur de densité  $\rho$  agissant sur l'espace de Hilbert  $\mathcal{H}$  et défini par l'équation (1.8) satisfait les propriétés suivantes :

- $\rho$  est autoadjoint :  $\rho^* = \rho$ ,
- $\rho$  est positif :  $\langle \psi | \rho | \psi \rangle = \sum_a p_a |\langle \psi | \psi_a \rangle|^2 \geq 0$ , pour  $|\psi\rangle$  quelconque,
- $\rho$  est de trace 1 :  $\text{Tr}(\rho) = \sum_i \sum_a p_a c_{i,a}^* c_{i,a} = \sum_n p_n \sum_i |c_{i,n}|^2 = \sum_a p_a = 1$ .

#### Définition 1.3.

1. Si on dénote par  $\mathcal{S}(\mathcal{H})$  l'ensemble des opérateurs de densité sur  $\mathcal{H}$ , alors  $\mathcal{S}(\mathcal{H})$  est convexe et les états purs correspondent aux points extrémaux de  $\mathcal{S}(\mathcal{H})$ .
2. Tous les autres états, combinaisons convexes d'états purs, sont les états mixtes.

L'opérateur de densité  $\rho$  défini par l'équation (1.8) est un opérateur auto-adjoint, il existe une base orthonormée  $\{|v_a\rangle\}_a$  de  $\mathcal{H}$  où  $\rho$  est diagonal, tel que

$$\rho = \sum_a p_a |v_a\rangle \langle v_a|,$$

où  $0 < p_a \leq 1$  et  $\sum_a p_a = 1$ . Dans cette base, si l'état n'est pas pur, il y a plus d'un terme dans cette somme et on parle de mélange statistique, tandis que pour un état pur cette somme se réduit à un unique élément. A partir de cela, on peut montrer qu'une matrice de densité  $\rho$  représente un état pur si et seulement si  $\rho^2 = \rho$ . En effet, dans la base  $\{|v_a\rangle\}_a$

$$\text{Tr}(\rho^2) \leq 1 = \sum_a p_a, \quad (1.10)$$

et la dernière inégalité devient une égalité si et seulement si  $\rho$  représente un état pur. En effet, si  $\sum_a p_a^2 = \sum_a p_a$  alors  $\sum_a p_a(1 - p_a) = 0$  et comme  $0 < p_a \leq 1$  on doit avoir  $p_a(1 - p_a) = 0$ ,  $\forall a$  et donc  $\exists ! a$   $p_a = 1$  et  $\forall a' \neq a$ ,  $p_{a'} = 0$ .



Exemple du qubit :

Revenons au cas de l'espace de Hilbert, de dimension 2, muni de la base orthonormée standard  $\{|0\rangle, |1\rangle\}$ . Soit  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  avec  $|\alpha|^2 + |\beta|^2 = 1$ , la matrice de densité dans cette base est

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \beta^*\alpha \\ \alpha^*\beta & |\beta|^2 \end{pmatrix}.$$

On vérifie aisément que cette matrice de densité satisfait bien les Propriétés 1.4.

**Observables et Mesure**

Si nous considérons à nouveau l'espace de Hilbert  $\mathcal{H} = \mathbb{C}^d$ , il a été établi précédemment qu'effectuer la mesure de l'observable  $\mathbf{X}$  sur l'état  $|\psi\rangle$  fournit une variable aléatoire  $X$  sur  $(\Omega, \Sigma)$ , où l'ensemble  $\Omega = \{x_a, a \in \mathbb{N}^*\}$  est l'ensemble des valeurs propres de  $\mathbf{X}$ . Si on note  $\{|\psi_a\rangle, a \in \mathbb{N}^*\}$  l'ensemble des vecteurs propres de  $\mathbf{X}$  associés aux valeurs propres  $x_a$ , alors si l'état quantique juste avant la mesure est  $|\psi\rangle$ , nous avons montré en (1.5) et (1.6) que

$$P_\psi(X = x_a) = \langle\psi|\mathbf{P}_a|\psi\rangle, \quad E_\psi(X) = \langle\psi|\mathbf{X}|\psi\rangle,$$

avec  $\mathbf{P}_a = |\psi_a\rangle\langle\psi_a|$ . De la même façon, l'opérateur de densité  $\rho = |\psi\rangle\langle\psi|$  sert de représentation pour l'état  $|\psi\rangle$ , il permet de rendre compte des mêmes probabilités suite à une mesure et donne la même valeur de l'espérance de la variable aléatoire  $X$

$$\begin{aligned} \text{Tr}(\rho\mathbf{P}_a) &= \text{Tr}(|\psi\rangle\langle\psi|\psi_a\rangle\langle\psi_a|) = \langle\psi|\psi_a\rangle\text{Tr}(|\psi\rangle\langle\psi_a|) = |\langle\psi|\psi_a\rangle|^2 = \langle\psi|\mathbf{P}_a|\psi\rangle, \\ \text{Tr}(\rho\mathbf{X}) &= \sum_a x_a \text{Tr}(\rho\mathbf{P}_a) = \sum_a x_a \langle\psi|\mathbf{P}_a|\psi\rangle = \langle\psi|\mathbf{X}|\psi\rangle. \end{aligned}$$

Et pour les états mixtes définis par (1.8), la probabilité d'obtenir le résultat  $X = x_a$  en effectuant une mesure de l'observable  $\mathbf{X}$  sur l'état  $\rho$  et l'espérance de la variable aléatoire  $X$  sont

$$P_\rho(X = x_a) = \sum_b p_b \langle\psi_b|\mathbf{P}_a|\psi_b\rangle, \quad E_\rho(X) = \sum_b p_b \langle\psi_b|\mathbf{X}|\psi_b\rangle.$$

L'opérateur de densité  $\rho$  a un intérêt certain pour les physiciens. Il décrit complètement l'état d'un système physique, le Chapitre 2 est consacré à son estimation.

Exemple du qubit :

Comme décrit précédemment, on peut faire une mesure qui *projette* le qubit dans la base standard. Nous obtiendrons le résultat  $|0\rangle$  avec la probabilité  $|\alpha|^2$  et le résultat  $|1\rangle$  avec la probabilité  $|\beta|^2$ .

On peut définir d'une manière plus aisément généralisable, une mesure projective  $\mathbf{M}$  qui associe  $\rho$  à une variable aléatoire  $X$  sur  $(\Omega, \Sigma)$ . Une telle mesure est appelée PVM (Projection Valued Measure). C'est-à-dire que l'on associe à tout  $O \in \Sigma$  un projecteur de  $\mathcal{H}$ , de telle sorte que

1.  $M(\Omega) = \mathbf{1}_{\mathcal{H}}$ ,
2.  $M(\cup_n O_n) = \sum_n M(O_n)$  pour toute suite  $\{O_n, n \in \mathbb{N}\}$  d'ensembles deux à deux disjoints.

Le résultat de la mesure sur  $\rho$  suit alors une loi de probabilité définie par :

$$P_\rho(X \in O) = \text{Tr}(\rho M(O)). \quad (1.11)$$

Le cas discret décrit précédemment rentre évidemment dans ce cadre :  $M(O) = \sum_{a: x_a \in O} \mathbf{P}_a$ . Et dans le cas d'un espace de Hilbert  $\mathcal{H}$  infini dimensionnel, les théorèmes spectraux assurent que l'on peut toujours décomposer une observable (auto-ajointe) en somme (intégrale) de projecteurs, c'est-à-dire :

$$\mathbf{X} = \int_{\Omega} \omega M(d\omega),$$

où  $M(d\omega)$  est une mesure spectrale sur  $\Omega$ , permettant de définir  $M(O) = \int_O M(d\omega)$ , et d'associer par ce biais une PVM à la mesure de l'observable  $\mathbf{X}$  sur l'état  $\rho$ .

### POVM

Cependant, la plupart du temps le système que l'on considère n'est pas isolé puisque celui-ci fait partie d'un système plus grand avec lequel il interagit. Supposons qu'une mesure soit effectuée sur le système global. Cette mesure est caractérisée par un ensemble de projecteurs orthogonaux associés aux différents résultats de la mesure. Au niveau du sous-système, l'effet de la mesure n'est pas la même ; il n'est en général plus possible d'y associer un ensemble de projecteurs. Il faut introduire la notion de mesure généralisée  $\mathbf{M}$ , appelée une POVM dans la littérature Holevo [47], pour positive operator-valued measure, i.e. mesure à valeurs dans les opérateurs positifs.

#### **Définition 1.4.**

*Soit  $(\Omega, \Sigma)$  un espace mesurable, on appelle  $\mathbf{M}$  mesure généralisée ou une POVM (positive operator-valued measure) une collection d'opérateurs auto-adjoints  $\mathbf{M} = \{M(A), A \in \Sigma\}$ , agissant sur l'espace de Hilbert  $\mathcal{H}$ , telle que*

1.  $M(\Omega) = \mathbf{1}_{\mathcal{H}}$ , où  $\mathbf{1}_{\mathcal{H}}$  est l'opérateur identité,

2.  $M(A)$  est positif pour tout  $A \in \Sigma$ ,
3.  $M(\cup_n O_n) = \sum_n M(O_n)$  pour toute suite  $\{O_n, n \in \mathbb{N}\}$  d'ensembles deux à deux dis-joints.

Exemple du qubit :

Reprenons l'espace de Hilbert, de dimension 2, muni de la base orthonormée standard  $\{|0\rangle, |1\rangle\}$ . Soit  $|\psi_0\rangle = |0\rangle$  et  $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  deux états quantiques dont les matrices de densité dans cette base sont respectivement

$$\rho_0 = |\psi_0\rangle\langle\psi_0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ et } \rho_1 = |\psi_1\rangle\langle\psi_1| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Comme  $\langle\psi_0|\psi_1\rangle = \frac{1}{\sqrt{2}}$ , il n'existe pas de mesure quantique permettant de distinguer ces deux états avec certitude. Considérons la POVM

$$\mathbf{M} = \left\{ M_1 = a|1\rangle\langle 1|, M_2 = \frac{a}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|), M_3 = \mathbf{1} - M_1 - M_2 \right\},$$

où  $\mathbf{1}$  est la matrice identité et  $a = 2 - \sqrt{2} > 0$ . On vérifie facilement que  $\mathbf{M}$  représente une POVM :

- les  $M_i$  sont hermitiens,
- leurs valeurs propres sont respectivement  $\{(a, 0), (a, 0), (\sqrt{2}a, 0)\}$  donc les  $M_i$  sont positifs,
- leur somme vaut bien  $\mathbf{1}$ .

Si l'on effectue cette mesure pour l'état  $\rho_0$ , on obtiendra après la mesure, les états  $\tau_1, \tau_2$  et  $\tau_3$ , avec respectivement les probabilités  $p_{\rho_0,1}, p_{\rho_0,2}$  et  $p_{\rho_0,3}$  suivantes :

$$\begin{aligned} p_{\rho_0,1} &= \text{Tr}(M_1\rho_0) = 0, \\ p_{\rho_0,2} &= \text{Tr}(M_2\rho_0) = a/2, \\ p_{\rho_0,3} &= \text{Tr}(M_3\rho_0) = 1 - a/2. \end{aligned}$$

Si l'on effectue cette mesure pour l'état  $\rho_1$ , on obtiendra après la mesure, les états  $\tau_1, \tau_2$  et  $\tau_3$ , avec respectivement les probabilités  $p_{\rho_1,1}, p_{\rho_1,2}$  et  $p_{\rho_1,3}$  suivantes :

$$\begin{aligned} p_{\rho_1,1} &= \text{Tr}(M_1\rho_1) = a/2, \\ p_{\rho_1,2} &= \text{Tr}(M_2\rho_1) = 0, \\ p_{\rho_1,3} &= \text{Tr}(M_3\rho_1) = 1 - a/2. \end{aligned}$$

On peut donc utiliser cette POVM pour distinguer l'état  $\rho_0$  de l'état  $\rho_1$  dans le sens où si on obtient  $\tau_1$ , on sait que l'état avant la mesure était  $\rho_1$  et si on obtient  $\tau_2$ , on sait que l'état avant la mesure était  $\rho_0$ . Si on obtient  $\tau_3$ , on ne peut rien conclure.

**Remarque 1.1.**

Les POVM auxquelles on s'intéresse plus particulièrement sont les POVM dominées par une mesure  $\sigma$ -fini  $\mu$  sur  $(\Omega, \Sigma)$  et qui s'écrivent

$$M(A) = \int_A m(x) d\mu(x), \forall A \in \Sigma,$$

où

- $m(x)$  est positif pour tout  $x$ ,
- $M(\Omega) = \int_{\Omega} m(x) d\mu(x) = \mathbf{1}_{\mathcal{H}}$ .

Dans le cadre de cette thèse, l'espace de Hilbert considéré est l'espace  $\mathbb{L}^2(\mathbb{R})$ , la mesure effectuée se fait par tomographie quantique homodyne, une technique que nous explicitons dans la Section 1.1.3.

**1.1.3 Formalisme de l'optique quantique**

Dans le cadre de l'optique quantique, la lumière est considérée comme constituée de **photons**, objets quantiques qui se comportent comme des corpuscules dans leurs interactions avec la matière et comme des ondes pour leur propagation. La lumière est une onde électromagnétique. L'*oscillateur harmonique* décrit, notamment un mode du champ électromagnétique : la **lumière monochromatique** qui est l'objet quantique étudié dans cette thèse. Dans ce cadre, nous considérons l'espace complexe de Hilbert séparable

$$\mathcal{H} = \mathbb{L}^2(\mathbb{R}),$$

qui est l'ensemble des fonctions à valeurs complexes  $\psi(x)$  définies sur  $\mathbb{R}$ , de carré intégrable. Cet espace est muni du produit scalaire

$$\langle \psi_1 | \psi_2 \rangle = \langle \psi_1, \psi_2 \rangle = \int_{\mathbb{R}} \psi_1^*(x) \psi_2(x) dx.$$

**Quadratures Q et P**

En mécanique classique, l'oscillateur harmonique est une particule assujettie à se déplacer le long d'un axe et soumise à une force de rappel proportionnelle à la distance à un point situé sur cet axe. En mécanique quantique, la notion de trajectoire est abandonnée au profit de la notion d'état. A un instant  $t_0$ , l'état d'un système est défini par la donnée d'un *ket*  $|\psi(t_0)\rangle$  appartenant à l'espace des états  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$ . L'oscillateur classique peut être décrit à tout moment par la donnée d'un point dans l'espace des phases  $(x, p_x)$ . Par analogie, à la

coordonnée position  $x$ , on fait correspondre l'opérateur de position  $\mathbf{Q}$  (le champs électrique), et à la quantité de mouvement  $p_x$ , on fait correspondre un opérateur de dérivée partielle par rapport à  $x$ , i.e. l'opérateur de quantité de mouvement  $\mathbf{P}$  (le champs magnétique). Ces opérateurs sur l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  sont tels que

$$(\mathbf{Q}\psi_1)(x) = x\psi_1(x), \quad (1.12)$$

$$(\mathbf{P}\psi_2)(x) = -i\frac{\partial\psi_2(x)}{\partial x}, \quad (1.13)$$

avec  $\psi_1, \psi_2$  des fonctions de  $\mathcal{H}$ . De plus,  $\mathbf{Q}$  et  $\mathbf{P}$  vérifient les relations de commutation de Heisenberg

$$[\mathbf{Q}, \mathbf{P}] = \mathbf{QP} - \mathbf{PQ} = i\mathbf{1}, \quad (1.14)$$

où  $\mathbf{1}$  est l'opérateur identité. Elles caractérisent le fait que l'on ne peut connaître à la fois la position et la quantité de mouvement d'une particule, car toute mesure de l'une entraîne une incertitude sur la valeur de l'autre : il n'existe pas de mesure spectrale positive.

### Equation de Schrödinger

Un opérateur hermitien particulièrement important est l'opérateur d'énergie totale, ou Hamiltonien<sup>1</sup>  $\mathbf{H}$ . Cet opérateur permet de calculer les valeurs propres de l'énergie. Dans le cas de l'oscillateur quantique, le Hamiltonien  $\mathbf{H}$  est indépendant du temps et s'écrit :

$$\mathbf{H} \equiv \frac{\mathbf{Q}^2}{2} + \frac{\mathbf{P}^2}{2}. \quad (1.15)$$

Il nous faut donc résoudre l'équation de Schrödinger indépendante du temps associée à cet Hamiltonien :

$$\mathbf{H}|\psi\rangle = E|\psi\rangle, \quad (1.16)$$

où  $E$  est l'énergie associée à l'état quantique  $|\psi\rangle$ .

Dans le cadre de cette thèse, l'espace d'état étant  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$ , le ket  $|\psi\rangle$  est une fonction de  $\mathbb{L}^2(\mathbb{R})$  et dans un formalisme plus analytique l'équation de Schrödinger s'écrit

$$\left[ -\frac{1}{2} \frac{\partial^2}{\partial x^2} + \frac{x^2}{2} \right] \psi = E\psi, \quad E \in \mathbb{R}, \quad (1.17)$$

---

<sup>1</sup> $\hbar\mathbf{H}$  avec ici la constante de Planck  $\hbar = 1$ , la masse  $m = 1$  et à la fréquence  $\omega = 1$ .

### Opérateurs d'échelle et états de Fock

Résoudre l'équation (1.16), revient à déterminer le spectre de le Hamiltonien  $\mathbf{H}$ . Ceci se fait à l'aide des opérateurs d'échelle :

$$\mathbf{A} = 2^{-1/2}(\mathbf{Q} + i\mathbf{P}). \quad (1.18)$$

En tenant compte de la relation (1.14), on verifie aisément que

$$[\mathbf{A}, \mathbf{A}^\dagger] = \mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A} = \mathbf{1}, \quad (1.19)$$

donc ces opérateurs ne sont pas hermitiens et ne commutent pas. Le Hamiltonien s'écrit également

$$\mathbf{H} = \mathbf{N} + \frac{1}{2}, \quad (1.20)$$

avec  $\mathbf{N} \equiv \mathbf{A}^\dagger\mathbf{A}$ , appelé opérateur *nombre de photons*. Soit  $|n\rangle$  un vecteur propre de  $\mathbf{N}$  associé à la valeur propre  $n$  tel que

$$\mathbf{N}|n\rangle = n|n\rangle, \quad (1.21)$$

alors  $\mathbf{A}^\dagger|n\rangle$  est un vecteur propre de  $\mathbf{N}$  associé à la valeur propre  $(n+1)$ . En effet, en utilisant successivement les équations (1.19), (1.21) et la définition de  $\mathbf{N} \equiv \mathbf{A}^\dagger\mathbf{A}$ , on peut montrer que

$$\begin{aligned} \mathbf{A}\mathbf{A}^\dagger &= \mathbf{A}^\dagger\mathbf{A} + \mathbf{1} \\ \mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger &= \mathbf{A}^\dagger(\mathbf{A}^\dagger\mathbf{A} + \mathbf{1}) \\ \mathbf{N}\mathbf{A}^\dagger|n\rangle &= \mathbf{A}^\dagger(\mathbf{N} + \mathbf{1})|n\rangle = (n+1)\mathbf{A}^\dagger|n\rangle. \end{aligned}$$

De la même façon, on pourrait montrer que  $\mathbf{A}|n\rangle$  est un vecteur propre de  $\mathbf{N}$  associé à la valeur propre  $(n-1)$ . Les opérateurs  $\mathbf{A}$  et  $\mathbf{A}^\dagger$  ont la propriété de diminuer et d'augmenter respectivement la valeur propre de  $\mathbf{N}$ , et donc de  $\mathbf{H}$ , on nomme parfois  $\mathbf{A}$  et  $\mathbf{A}^\dagger$  respectivement opérateur d'*annihilation* et opérateur de *création* de photons. Alors pour

$$\mathbf{A}|n\rangle = \sqrt{n}|n-1\rangle, \quad n \neq 0 \quad \text{et} \quad \mathbf{A}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \quad (1.22)$$

$$\mathbf{A}|0\rangle = 0. \quad (1.23)$$

#### Remarque 1.2.

L'autre alternative à l'équation (1.23) est que 0 ne soit pas valeur propre de  $\mathbf{A}$ ,  $\mathbf{A}|0\rangle \neq 0$  et  $\mathbf{A}^\dagger(\mathbf{A}|0\rangle) = 0$  qui donne la seconde solution de l'équation de Schrödinger (1.17). Cette solution n'est pas normalisable.

On peut montrer par récurrence à partir de l'équation (1.19) et des équations (1.22) et (1.23) que  $n$  est entier,  $n \in \mathbb{N}$ . La dernière équation implique que le nombre moyen de photons dans l'état  $|0\rangle$  est nul,  $\langle \mathbf{N} \rangle = 0$ . Pour cette raison, l'état  $|0\rangle$  est appelé l'état *vide* et l'état fondamental de le Hamiltonien est donc

$$\mathbf{H}|0\rangle = \frac{1}{2}|0\rangle.$$

Il est intéressant de remarquer que l'énergie moyenne de l'oscillateur harmonique est alors minimal

$$\langle 0|\mathbf{H}|0\rangle = \frac{1}{2}.$$

On peut alors définir n'importe quel état propre  $|n\rangle$  comme une excitation obtenue à partir du vide. Les états excités s'obtiennent alors simplement en appliquant  $\mathbf{A}^\dagger$  à répétition :

$$|n\rangle = \frac{(\mathbf{A}^\dagger)^n}{\sqrt{n!}}|0\rangle, \quad (1.24)$$

et  $\mathbf{H}|n\rangle = (n + \frac{1}{2})|n\rangle$ . Les vecteurs propres  $|n\rangle$  sont appelés **états nombres** ou **états de Fock**. Ce sont les états pour lesquels le nombre de photons dans le mode est parfaitement défini et vaut  $n$  (si on mesure  $\mathbf{N}$  sur un état  $|n\rangle$  on obtient une variable aléatoire  $N = n$  p.s.).

### Base de Fock et fonction d'onde

En utilisant les équations (1.18), (1.12) et (1.13) dans l'équation (1.23), on obtient l'équation différentielle

$$\mathbf{A}\psi_0(x) = \frac{1}{\sqrt{2}}(x + \frac{\partial}{\partial x})\psi_0(x) = 0,$$

dont la solution normalisée  $\psi_0(x)$  est appelée **fonction d'onde** du vide et s'écrit

$$\psi_0(x) = \pi^{-1/4}e^{-x^2/2}. \quad (1.25)$$

De manière similaire, on trouve les fonctions d'onde  $\psi_n(x)$  des états de Fock  $|n\rangle$  par les équations (1.23), (1.12) et (1.13). On obtient la relation de récurrence pour  $n \geq 1$

$$\mathbf{A}^\dagger\psi_n(x) = \frac{1}{\sqrt{2}}(x - \frac{\partial}{\partial x})\psi_n(x) = \psi_{n+1}(x),$$

satisfaite par les fonctions d'onde  $\psi_n(x)$ ,  $n \in \mathbb{N}$

$$\psi_n(x) := \frac{1}{\sqrt{\pi 2^n n!}} H_n(x) e^{-x^2/2}. \quad (1.26)$$

Ici,  $H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$  dénote le  $n$ -ième polynôme de Hermite.

**Définition 1.5.**

La base  $\{\psi_n\}_{n \in \mathbb{N}}$  de l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  définie par l'expression (1.26) est appelée **base de Fock**.

**Remarque 1.3.**

L'équation de Schrödinger (1.17) est une équation différentielle du second ordre, elle admet deux solutions fondamentales, la première est la fonction d'onde  $\psi_n(x)$  et la seconde découle de la Remarque 1.2. On l'appelle la fonction d'onde irrégulière  $\varphi_n(x)$  car elle n'est pas normalisable. Ces deux solutions  $\psi_n(x)$  et  $\varphi_n(x)$  nous permettrons dans un prochain paragraphe, d'introduire des fonctions très importantes en physique : les *pattern functions*.

Exemple :

Considérons l'opérateur de "multiplication par  $x$ ",  $\mathbf{Q}$  défini par (1.12), la mesure associée  $\mathbf{M}$  est telle que  $M(O) = \mathbb{1}_O(x)$  et

$$P_{|n\rangle}(Q \in O) = \langle n | \mathbb{1}_O | n \rangle = \int_O |\psi_n(x)|^2 dx.$$

La variable aléatoire  $Q$  résultat de la mesure suit la loi de densité de probabilité  $|\psi_n(x)|^2$ .

En particulier, pour l'état fondamental  $|0\rangle$ ,  $|\psi_0(x)|^2 = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ , donc la variable aléatoire  $Q$  suit la loi Gaussienne de moyenne nulle et de variance 1/2.

D'une façon générale si  $\rho = \sum_n c_n |n\rangle \langle n|$  alors  $Q$  a une densité de probabilité

$$\sum_n c_n |\psi_n(x)|^2.$$

**Etats cohérents**

Les états de *Fock* sont difficiles à produire expérimentalement au delà de  $n = 2$  et apportent donc plutôt une commodité de représentation et de calcul. Afin de modéliser un champ monomode, on introduit une nouvelle famille d'états, les **états cohérents** très utile pour la théorie classique de l'oscillateur harmonique. Ces états sont définis par

$$|\alpha_{coh}\rangle = e^{-|\alpha|^2/2} e^{\alpha \mathbf{A}^\dagger} |0\rangle,$$

où  $\mathbf{A}^\dagger$  est l'opérateur de *création* et  $\alpha$  est un nombre complexe. Les états *cohérents* peuvent être vus comme des états *vides*  $|0\rangle$  "*déplacés*".



En se servant des états de Fock comme base de l'espace de Hilbert, un état *cohérent*  $|\alpha_{coh}\rangle$  n'est autre qu'une superposition de tous les états propres  $|n\rangle$ . En effet, en développant l'exponentielle en série et en utilisant la définition (1.24) des états de Fock  $|n\rangle$  :

$$|\alpha_{coh}\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n (\mathbf{A}^\dagger)^n}{\sqrt{n!}} |0\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (1.27)$$

On peut également souligner que le *vide*  $|0\rangle$  est un état *cohérent* et ces états sont des états purs.

L'équation (1.27) montre que le nombre de photons moyen d'un état *cohérent* suit une loi de Poisson de paramètre  $|\alpha|^2$ . Une propriété essentielle des états *cohérents* est qu'ils sont les vecteurs propres de l'opérateur d'*annihilation*

$$\mathbf{A}|\alpha_{coh}\rangle = \alpha|\alpha_{coh}\rangle. \quad (1.28)$$

*Démonstration.*

Une démonstration en 3 temps :

1. remarquons que  $[\mathbf{A}, (\mathbf{A}^\dagger)^n] = n(\mathbf{A}^\dagger)^{n-1}$  (simple récurrence),
2. notons que

$$\begin{aligned} [\mathbf{A}, e^{\alpha\mathbf{A}^\dagger}] &= \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} [\mathbf{A}, (\mathbf{A}^\dagger)^n] = \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{(n-1)!}} (\mathbf{A}^\dagger)^{n-1} \\ &= \alpha \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} (\mathbf{A}^\dagger)^n = \alpha e^{\alpha\mathbf{A}^\dagger}, \end{aligned}$$

3. et par l'équation (1.23), on déduit que  $|\alpha_{coh}\rangle$  est un vecteur propre de  $\mathbf{A}$  :

$$\begin{aligned} \mathbf{A}|\alpha_{coh}\rangle &= e^{-|\alpha|^2/2} \mathbf{A} e^{\alpha\mathbf{A}^\dagger} |0\rangle = e^{-|\alpha|^2/2} [\mathbf{A}, e^{\alpha\mathbf{A}^\dagger}] |0\rangle, \\ &= \alpha e^{-|\alpha|^2/2} e^{\alpha\mathbf{A}^\dagger} |0\rangle = \alpha|\alpha_{coh}\rangle. \end{aligned}$$

□

Etudions maintenant les valeurs moyennes et les variances des composantes de quadrature  $\mathbf{Q}$  et  $\mathbf{P}$  définies en (1.12) et (1.13) dans l'état *cohérent*  $|\alpha_{coh}\rangle$ . Compte tenu de l'équation (1.18), ils vérifient  $\mathbf{Q} = 2^{-1/2}(\mathbf{A}^\dagger + \mathbf{A})$  et  $\mathbf{P} = i2^{-1/2}(\mathbf{A}^\dagger - \mathbf{A})$ . On trouve facilement à partir de l'équation (1.28) les égalités suivantes :

$$\begin{aligned} \langle \alpha_{coh} | \mathbf{A} | \alpha_{coh} \rangle &= \alpha, & \langle \alpha_{coh} | \mathbf{A}^\dagger \mathbf{A} | \alpha_{coh} \rangle &= |\alpha|^2, \\ \langle \alpha_{coh} | \mathbf{A}^\dagger | \alpha_{coh} \rangle &= \alpha^*, & \langle \alpha_{coh} | \mathbf{A} \mathbf{A}^\dagger | \alpha_{coh} \rangle &= |\alpha|^2 + 1. \end{aligned}$$

Ainsi les valeurs moyennes de  $\mathbf{Q}$  et  $\mathbf{P}$  sont

$$\begin{aligned}\langle \mathbf{Q} \rangle &= \langle \alpha_{coh} | \mathbf{Q} | \alpha_{coh} \rangle = 2^{-1/2} \langle \alpha_{coh} | (\mathbf{A}^\dagger + \mathbf{A}) | \alpha_{coh} \rangle = \sqrt{2} \mathcal{R}e(\alpha), \\ \langle \mathbf{P} \rangle &= \langle \alpha_{coh} | \mathbf{P} | \alpha_{coh} \rangle = i 2^{-1/2} \langle \alpha_{coh} | (\mathbf{A}^\dagger - \mathbf{A}) | \alpha_{coh} \rangle = \sqrt{2} \mathcal{I}m(\alpha),\end{aligned}$$

où  $\mathcal{R}e(\alpha)$  et  $\mathcal{I}m(\alpha)$  dénotent respectivement la partie réelle et la partie imaginaire de  $\alpha$ . De plus,  $\langle \mathbf{Q}^2 \rangle = \langle \alpha_{coh} | \mathbf{Q}^2 | \alpha_{coh} \rangle = 2((\mathcal{R}e(\alpha))^2 + \frac{1}{4})$  et  $\langle \mathbf{P}^2 \rangle = \langle \alpha_{coh} | \mathbf{P}^2 | \alpha_{coh} \rangle = 2((\mathcal{I}m(\alpha))^2 + \frac{1}{4})$ . On démontre donc que les variances des composantes de quadrature  $\mathbf{Q}$  et  $\mathbf{P}$  sont

$$\begin{aligned}(\Delta \mathbf{Q})^2 &\equiv \langle \mathbf{Q}^2 \rangle - (\langle \mathbf{Q} \rangle)^2 = \frac{1}{2}, \\ (\Delta \mathbf{P})^2 &\equiv \langle \mathbf{P}^2 \rangle - (\langle \mathbf{P} \rangle)^2 = \frac{1}{2}.\end{aligned}$$

Notons que pour ces états, la relation d'incertitude est tout juste satisfaite :

$$(\Delta \mathbf{Q})(\Delta \mathbf{P}) = \frac{1}{2},$$

alors que dans l'état propre  $|n\rangle$  la relation devient

$$(\Delta \mathbf{Q})(\Delta \mathbf{P}) = (2n + 1) \frac{1}{2}.$$

Les états *cohérents* minimisent l'incertitude quantique. Dans ce sens, ce sont les états qui se rapprochent le plus des trajectoires classiques et on les nomme **états classiques**. Notons que les états *cohérents* permettent de modéliser au mieux le champ issu d'un laser mono-mode opérant à haute intensité (dépourvu de bruit technique) compte tenu de la relation d'Heisenberg. Pour un champs dans l'état *cohérent*  $|\alpha_{coh}\rangle$  :  $|\alpha|$  est la valeur du champs en amplitude et  $\arg(\alpha)$  est la phase.

### Chat de Schrödinger

Si la mécanique quantique prédit l'existence d'états du champs électromagnétique, elle augmente également toute superposition linéaire de tels états. On désigne une superposition linéaire de deux états *cohérents* par le nom de *chat de Schrödinger* :

$$|\alpha_{sch}\rangle \propto (|\alpha_{coh}\rangle + |-\alpha_{coh}\rangle),$$

où  $\propto$  signifie à une constante de normalisation près. Le *chat de Schrödinger* est une superposition quantique (non-classique) d'états, à l'inverse d'une superposition classique décrite en (1.8).

Nous étudierons plus en détails ces différents états quantiques (états *nombres*, *cohérents*) dans la Section 1.1.4.

### Tomographie quantique homodyne TQH

Dans le cadre de cette thèse, nous travaillons sur l'espace  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  muni de la base de Fock  $\{\psi_n\}_{n \in \mathbb{N}}$  donnée par la Définition 1.5 et l'état quantique étudié est un faisceau de lumière monochromatique dans une cavité décrit par l'état d'un oscillateur quantique. Les observables qui nous intéressent sont les champs électrique et magnétique  $\mathbf{Q}$  et  $\mathbf{P}$ , définies en (1.12) et (1.13), qui vérifient les relations de commutation (1.14). Ce qui implique qu'ils ne peuvent pas être mesurés simultanément. Une technique de mesure a été mise en place par Smithey, Beck, Raymer, et Faridani [82] appelée la tomographie quantique homodyne basée sur la théorie développée par Vogel et Risken [91]. Il s'agit de mesurer l'observable

$$\mathbf{X}_\phi = \cos(\phi)\mathbf{Q} + \sin(\phi)\mathbf{P}, \quad (1.29)$$

pour toutes phases  $\phi \in [0, \pi]$ . L'observable  $\mathbf{X}_\phi$  est appelée *quadrature de phase*  $\phi$ .

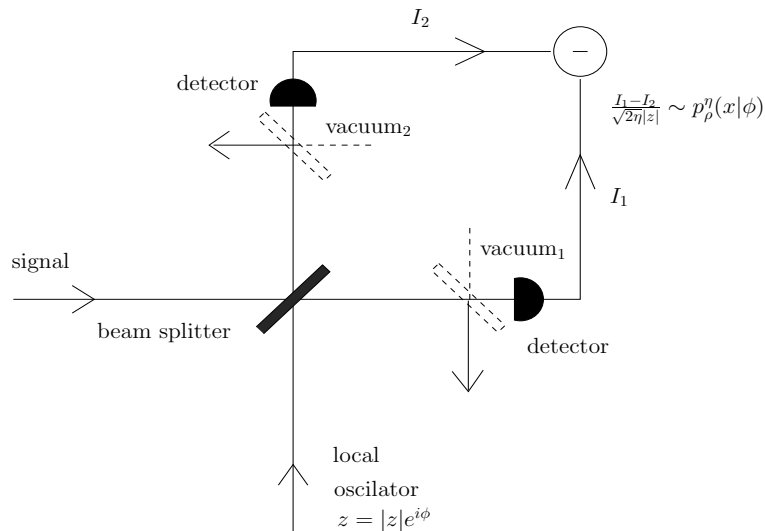


FIG. 1.1 – Schématisation de la mesure par TQH

Cette technique est illustrée sur la Figure 1.1.

- On prépare un laser monochromatique dans l'état  $\rho$  dans le mode  $\mathbf{A}$  défini en (1.18), on l'appelle *le faisceau signal*.
- On se dote d'un laser de référence d'amplitude complexe  $|z|$  de très haute intensité  $|z| \gg 1$  et de phase  $\arg z = \phi$ , on le nomme l'*oscillateur local*.
- On choisit au préalable la phase  $\Phi$  de l'oscillateur local de façon aléatoire et uniforme sur  $[0, \pi]$ .
- Le faisceau signal est mélangé avec l'oscillateur local sur une lame semi-réfléchissante (beam splitter); c'est-à-dire que les deux modes incidents  $\mathbf{A}$ ,  $|z|e^{i\phi}$  interfèrent sur la

lame pour produire deux modes émergents  $\mathbf{A}_1, \mathbf{A}_2$ . La relation suivante exprime le passage à travers la lame semi-réfléchissante

$$\mathbf{A}_1 = 2^{-1/2}(\mathbf{A} - z) \quad \text{et} \quad \mathbf{A}_2 = 2^{-1/2}(\mathbf{A} + z).$$

- Les photocourants  $I_1$  et  $I_2$  de chaque faisceau sont mesurés à l'aide d'une photodiode de haute efficacité quantique. Pour supprimer les intensités moyennes mesurées par chaque photodiode qui n'apportent aucune information utile, les photocourants sont soustraits pour ne conserver que les termes croisés d'interférence signal-oscillateur local. Dans le cadre d'une **détection idéale**, le signal utile en sortie  $X$  est alors directement proportionnel à la composante du champ entrant en phase avec l'oscillateur local. C'est-à-dire pour simplifier si on note

$$\mathbf{N}_1 = \mathbf{A}_1^\dagger \mathbf{A}_1 \quad \text{et} \quad \mathbf{N}_2 = \mathbf{A}_2^\dagger \mathbf{A}_2,$$

le nombre de photons des faisceaux arrivant à chaque photodiode, alors le signal utile en sortie est proportionnel à

$$\frac{1}{\sqrt{2|z|}}(\mathbf{N}_2 - \mathbf{N}_1).$$

Ainsi on remarque aisément par l'équation (1.18), que le signal utile en sortie est alors directement une mesure de

$$\mathbf{X}_\phi = \frac{1}{\sqrt{2|z|}}(\mathbf{N}_2 - \mathbf{N}_1).$$

La tomographie quantique homodyne revient donc à effectuer une mesure sur l'observable  $\mathbf{X}_\phi$  définie par l'équation (1.29). Cela fournit dans le cas dit idéal une variable aléatoire  $X$  (cf. (1.11)) qui admet une densité de probabilité conditionnelle à  $\Phi = \phi$ ,  $p_\rho(\cdot|\phi)$ . On peut écrire sa fonction caractéristique comme :

$$\mathcal{F}_1[p_\rho(\cdot|\phi)](t) = \text{Tr}(\rho e^{it\mathbf{X}_\phi}) = \text{Tr}(\rho e^{it(\cos(\phi)\mathbf{Q} + \sin(\phi)\mathbf{P})}), \quad (1.30)$$

où  $\mathcal{F}_1[p_\rho(\cdot|\phi)](t) := \int e^{itx} p_\rho(x|\phi)$  dénote la transformée de Fourier de  $p_\rho(\cdot|\phi)$ . De plus, si  $\Phi$  est choisie selon une loi uniforme, les observations  $(X, \Phi)$  sont de loi jointe

$$p_\rho(x, \phi) = \frac{1}{\pi} p_\rho(x|\phi) \mathbb{1}_{[0, \pi]}(\phi). \quad (1.31)$$

En pratique, il faut prendre en compte les erreurs qui s'ajoutent, du fait des appareils de mesure aux résultats d'une telle mesure. Heuristiquement, un état *vide* se mélange aux états pendant la mesure. Dans la littérature physique, il est établi que ces erreurs peuvent être

quantifiées en un seul coefficient  $\eta \in ]0, 1]$ , qui est évalué facilement par les physiciens. Les observations sont en fait  $(Y, \Phi)$  où

$$Y = \sqrt{\eta}X + \sqrt{(1-\eta)/2}\xi, \quad (1.32)$$

$\eta \in ]0, 1]$  est connu et  $\xi$  sont des erreurs de loi gaussienne standard, i.i.d. et indépendantes des  $X$ .

### Fonction de Wigner

Une représentation alternative à la représentation matricielle (cf. (1.8)) très utilisée est la fonction de Wigner, introduite pour la première fois par Wigner [92]. L'estimation de cette fonction est l'objet du Chapitre 3.

La fonction de Wigner  $W_\rho : \mathbb{R}^2 \rightarrow \mathbb{R}$  associée à l'état quantique  $\rho$  peut se définir à l'aide de sa transformée de Fourier

$$\widetilde{W}_\rho(u, v) := \mathcal{F}_2[W_\rho](u, v) = \text{Tr}(\rho e^{i(u\mathbf{Q}+v\mathbf{P})}), \quad (1.33)$$

où  $\mathcal{F}_2[W_\rho] = \iint e^{iuq+ivp}W_\rho(q, p)dqdp$ .

La définition (1.33) est à mettre en relation avec l'équation (1.30), en effet si on effectue le changement de variables  $(u, v) = (t \cos(\phi), t \sin(\phi))$ , on obtient :

$$\widetilde{W}_\rho(u, v) = \mathcal{F}_1[p_\rho(\cdot|\phi)](t) = \text{Tr}(\rho e^{it\mathbf{X}_\phi}). \quad (1.34)$$

La procédure décrite ci-dessus ((1.33) et (1.34)) est à rapprocher de la tomographie par émission de positrons, où la densité des observations est la transformée de Radon de la densité jointe sous-jacente

$$p_\rho(x|\phi) = \Re[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x \cos \phi + t \sin \phi, x \sin \phi - t \cos \phi) dt, \quad (1.35)$$

d'où l'origine de l'appellation *tomographie* quantique homodyne.

La propriété essentielle de la fonction de Wigner ainsi définie est qu'elle se comporte "comme" une densité de probabilité conjointe pour des variables aléatoires  $Q$  et  $P$  résultats de la mesure des observables  $\mathbf{Q}$  et  $\mathbf{P}$  si on pouvait effectuer une mesure simultanée des quadratures. Mais comme les observables  $\mathbf{Q}$  et  $\mathbf{P}$  ne commutent pas, on ne peut pas parler de densité jointe au sens probabiliste du terme. Il n'existe pas de POVM pour le couple d'observables  $(\mathbf{Q}, \mathbf{P})$ , on ne peut donc pas récolter des variables aléatoires  $(Q, P)$ . On parlera alors de *densité de probabilité jointe généralisée* puisque  $W_\rho$  peut prendre des valeurs négatives et

$\iint W_\rho = 1$ . En effet, par un changement de variables  $(q, p) = (x \cos \phi + t \sin \phi, x \sin \phi - t \cos \phi)$  et en substituant la définition (1.35)

$$\iint W_\rho(q, p) dq dp = \iint W_\rho(x \cos \phi + t \sin \phi, x \sin \phi - t \cos \phi) dt dx = \int p_\rho(x|\phi) dx = 1.$$

### Fonctions pattern

On souligne que sur l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  muni de la base de Fock  $\{\psi_n\}_{n \in \mathbb{N}}$  (Définition 1.5), la matrice de densité  $\rho$  a des entrées s'écrivant

- dans le formalisme physique :  $\rho_{j,k} = \langle j|\rho|k\rangle$ ,
- dans le formalisme mathématique :  $\rho_{j,k} = \langle \psi_j, \rho \psi_k \rangle$ .

La mesure de l'observable  $\mathbf{X}_\phi$  définie en (1.29), sur l'état  $\rho$  fournit dans le cadre d'une détection idéale une variable aléatoire  $X$  de densité de probabilité  $p_\rho$  définie en (1.31).

On écrit souvent la densité de probabilité  $p_\rho$  en fonction des entrées de la matrice  $\rho$  dans la base de Fock Leonhardt [65]

$$p_\rho(x, \phi) = \sum_{j,k=0}^{\infty} \rho_{j,k} \psi_j(x) \psi_k(x) e^{-i(j-k)\phi}. \quad (1.36)$$

Et inversement (D'Ariano, Macchiavello, et Paris [29], Leonhardt [65]), il est possible d'écrire les entrées de la matrice  $\rho$  dans la base de Fock à partir de la densité de probabilité  $p_\rho$  :

$$\rho_{j,k} = \int_0^\pi \int p_\rho(x, \phi) f_{j,k}(x) e^{i(j-k)\phi} dx d\phi, \quad (1.37)$$

où les fonctions  $f_{j,k} : \mathbb{R} \rightarrow \mathbb{R}$  introduites par Leonhardt, Paul, et D'Ariano [66] sont appelées les **fonctions "pattern"** (fonctions "dessin" en français). On peut les définir (Leonhardt, Munroe, Kiss, Richter, et Raymer [67], D'Ariano, Leonhardt, et Paul [30]) comme la dérivée du produit des deux solutions fondamentales de l'équation de Schrödinger (1.17)

$$f_{j,k}(x) = \frac{\partial}{\partial x} (\psi_j(x) \varphi_k(x)),$$

où la fonction  $\psi_k$  est la fonction d'onde régulière (base de Fock  $\int |\psi_k|^2 = 1$ ) et  $\varphi_k$  est la fonction d'onde irrégulière qui elle n'est pas de carré intégrable (cf Remarques 1.2 et 1.3).

Une expression plus explicite des fonctions *pattern* définies par leur transformée de Fourier  $\tilde{f}_{j,k}$ , est donnée par Richter [79] : pour  $j \geq k$

$$\tilde{f}_{j,k}(t) = \pi(-i)^{j-k} \sqrt{\frac{2^{k-j} k!}{j!}} |t| t^{j-k} e^{-\frac{t^2}{4}} L_k^{j-k}\left(\frac{t^2}{2}\right), \quad (1.38)$$

où  $L_k^\alpha(x)$  denote le polynôme de Laguerre généralisé. Les fonctions *pattern* dont la Figure 1.2 donne quelques exemples, sont telles que pour tout  $j, k \in \mathbb{N}$  :

- $f_{j,k}$  est une fonction bornée
- $f_{j,k}(x) = f_{k,j}(x)$ ,
- $f_{j,k}(-x) = (-1)^{j-k} f_{j,k}(x)$ , i.e. elles sont paires pour  $j - k$  pair et impaires pour  $j - k$  impair.

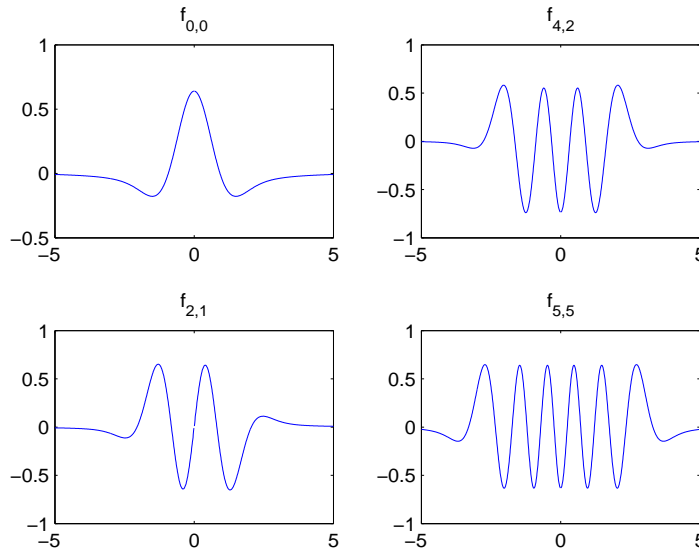


FIG. 1.2 – Exemple de fonctions *pattern*  $\pi^{-1} f_{j,k}(x)$ .

### Relation entre la fonction de Wigner et la matrice de densité

La matrice densité  $\rho$  et la fonction de Wigner  $W_\rho$  représentent complètement l'état quantique  $\rho$  associé. Ainsi, il est utile de souligner une des relations qui les lie. Il a été prouvé dans le livre Leonhardt [65] que la fonction de Wigner  $W_\rho$  est décrite en fonction des entrées de la matrice de densité  $\rho$  dans la base de Fock,  $\{\psi_n\}_{n \in \mathbb{N}}$  comme suit

$$W_\rho(q, p) = \sum_{j,k} \rho_{jk} W_{j,k}(q, p), \quad (1.39)$$

avec

$$W_{j,k}(q, p) = \frac{1}{\pi} \int e^{2ipx} \psi_j(q-x) \psi_k(q+x) dx. \quad (1.40)$$

Différentes expressions de la fonction  $W_{j,k}$  sont données par Leonhardt [65] (chap. 5) dont la suivante : pour  $j \geq k$

$$W_{j,k}(q, p) = \frac{(-1)^j}{\pi} \sqrt{\frac{k!}{j!}} e^{-q^2 - p^2} \left( \sqrt{2}(ip - q) \right)^{j-k} L_k^{j-k}(2q^2 + 2p^2), \quad (1.41)$$

où  $L_n^\alpha$  dénote le polynôme de Laguerre généralisé de degré  $n$  et d'ordre  $\alpha$  défini par

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

De plus,  $W_{j,k}(q, p) = W_{k,j}(q, -p)$ , pour tout  $j, k \in \mathbb{N}$  et une propriété intéressante est celle de sa transformée de Fourier

$$\widetilde{W}_{j,k}(q, p) = \frac{(-1)^{j+k}}{2} W_{j,k}\left(\frac{q}{2}, \frac{p}{2}\right).$$

Une relation très importante, dont la preuve peut être trouvée dans Leonhardt [65] est : soient  $\mathbf{F}_1, \mathbf{F}_2$  deux opérateurs arbitraires, alors

$$\text{Tr}(\mathbf{F}_1 \mathbf{F}_2) = 2\pi \iint W_1(q, p) W_2(q, p) dq dp,$$

où  $W_1, W_2$  les fonctions de Wigner de  $\mathbf{F}_1, \mathbf{F}_2$ . Ainsi, pour  $\mathbf{F}_1 \equiv \mathbf{F}_2 \equiv \rho$

$$\text{Tr}(\rho^2) = 2\pi \iint W_\rho^2(q, p) dq dp. \quad (1.42)$$

Ainsi de façon équivalente à (1.10), on a

$$\mathcal{P} := \text{Tr}(\rho^2) = 2\pi \iint W_\rho^2(q, p) dq dp \leq 1, \quad (1.43)$$

et un état quantique est pur si et seulement  $\mathcal{P} = 1$ . Au Chapitre 4, on se propose d'estimer la fonctionnelle  $\iint W_\rho^2$  comme estimation de la pureté d'un état quantique.

#### 1.1.4 Exemples d'états quantiques

Dans la section précédente, nous avons défini les états *nombres* (ou états de *Fock*), les états *cohérents*. Dans cette section, nous allons introduire le fameux *chat de Schrödinger* et donner les matrices de densité et les fonctions de Wigner de ces différents états.

##### A/ Description en terme de matrice de densité

Dans la base orthonormée complète constituée des états de *Fock*  $\{|n\rangle\}_{n \in \mathbb{N}}$  définie en (1.24), un état quantique peut être décomposé comme  $|\psi\rangle = \sum_n c_n |n\rangle$  et sa matrice de densité dans cette base vaut

$$\rho = |\psi\rangle\langle\psi| = \sum_{j,k} c_j^* c_k |j\rangle\langle k|.$$



### Les états de Fock

La matrice de densité la plus simple est la matrice de densité d'un état de *Fock*. Pour tout  $n \in \mathbb{N}$  un état de *Fock*  $|n\rangle$  (ou état  $n$ -photons) admet comme matrice de densité  $\rho_{|n\rangle} = |n\rangle\langle n|$  dont les entrées sont

$$\rho_{j,k} = \mathbb{1}_{\{j=k=n\}}, \quad (1.44)$$

où  $\mathbb{1}$  dénote l'indicatrice. La Figure 1.3 en donne quelques exemples.

### Les états cohérents

On a établi en (1.27) qu'un état *cohérent* n'était autre qu'une superposition de tous les états de *Fock*  $|n\rangle$ . Les entrées de la matrice de densité  $\rho_{|\alpha_{coh}\rangle}$  d'un état *cohérent*- $\alpha$  avec .

$$\rho_{j,k} = \langle j|\alpha_{coh}\rangle\langle\alpha_{coh}|k\rangle = e^{-|\alpha|^2} \frac{\alpha^j \alpha^{*k}}{\sqrt{j!k!}}. \quad (1.45)$$

Soit  $\alpha = (q_0 + ip_0)/\sqrt{2}$  avec  $q_0$  et  $p_0$  des nombres réels, la Figure 1.3 donne quelques exemples d'états *cohérents*.

### Le chat de Schrödinger

Le *chat de Schrödinger* est défini comme une superposition linéaire de deux états *cohérents* :

$$|\alpha_{sch}\rangle \propto (|\alpha_{coh}\rangle + |-\alpha_{coh}\rangle),$$

où  $\propto$  signifie à une constante près, telle que  $\langle\alpha_{sch}|\alpha_{sch}\rangle = 1$ . La matrice de densité est telle que

$$\begin{aligned} \rho_{|\alpha_{sch}\rangle} &\propto |\alpha_{sch}\rangle\langle\alpha_{sch}| \propto (|\alpha\rangle + |-\alpha_{coh}\rangle)(\langle\alpha_{coh}| + \langle-\alpha_{coh}|) \\ &\propto \rho_{|\alpha_{coh}\rangle} + \rho_{|-\alpha_{coh}\rangle} + (|\alpha_{coh}\rangle\langle-\alpha_{coh}| + |-\alpha_{coh}\rangle\langle\alpha_{coh}|), \end{aligned} \quad (1.46)$$

avec  $\rho_{|\alpha_{coh}\rangle}$  et  $\rho_{|-\alpha_{coh}\rangle}$  les matrices de densité, respectivement de l'état *cohérent*- $(\alpha)$  et de l'état *cohérent*- $(-\alpha)$ . C'est un mélange dit non classique car il apparaît des interférences. Soit  $\alpha = (q_0 + ip_0)/\sqrt{2}$ , avec  $q_0, p_0$  réels, les entrées de la matrice de densité  $\rho_{|\alpha_{sch}\rangle}$  d'un *chat de Schrödinger* sont pour  $p_0 = 0$

$$\rho_{j,k} = \langle j|\alpha_{sch}\rangle\langle\alpha_{sch}|k\rangle = \frac{2(q_0/\sqrt{2})^{j+k}}{\sqrt{j!k!}(e^{q_0^2/2} + e^{-q_0^2/2})} \mathbb{1}_{\{j \text{ et } k \text{ pair}\}}. \quad (1.47)$$

On peut souligner que le *chat de Schrödinger* est un état pur, sa matrice de densité vérifie  $\text{Tr}(\rho^2) = \text{Tr}(\rho) = 1$ . La Figure 1.3 représente un *chat de Schrödinger* défini comme la superposition d'un état *Coherent*- $\alpha$  avec  $q_0 = 3$  et  $p_0 = 0$  d'un état *Coherent*- $\alpha$  avec  $q_0 = -3$  et  $p_0 = 0$ .

### Représentations de quelques matrices de densité

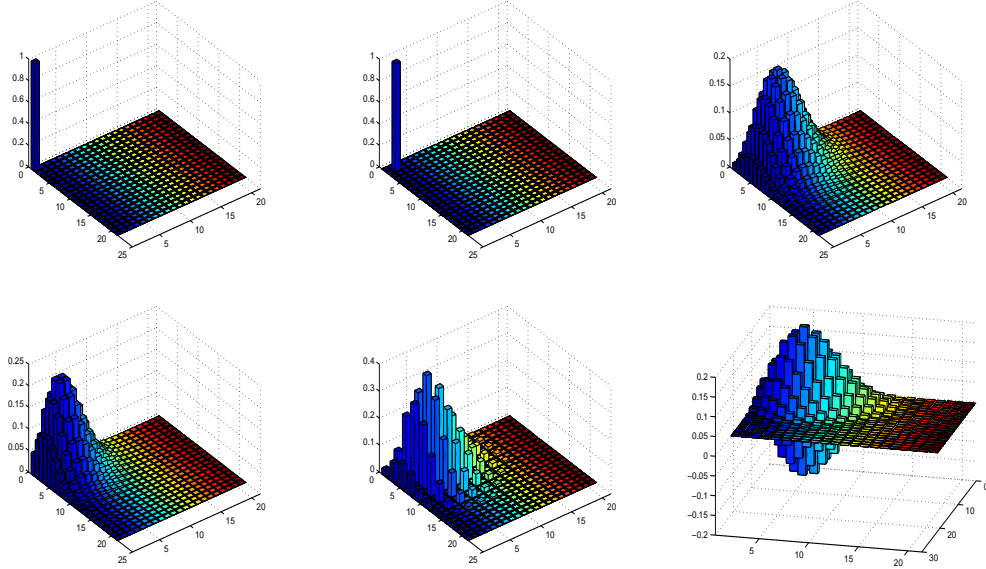


FIG. 1.3 – Représentation des matrices de densité du *vide* (1<sup>er</sup> en haut), de l'état *nombre 1-photon* (2<sup>nd</sup> en haut), de l'état *Coherent- $\alpha$*  avec  $q_0 = 3$  et  $p_0 = 0$  (3<sup>ème</sup> en haut), de l'état *Coherent- $\alpha$*  avec  $q_0 = \sqrt{6}$  et  $p_0 = 0$  (1<sup>er</sup> en bas), du *chat de Schrödinger*,  $q_0 = 3$  (2<sup>nd</sup> en bas) et de l'état *Coherent- $\alpha$*  avec  $q_0 = -3$  et  $p_0 = 0$  (3<sup>ème</sup> en bas).

### B/ Description en terme de fonction de Wigner

#### L'état *vide*

L'état fondamental est l'état *vide*  $|0\rangle$ , *vacuum state* en anglais, nous avons vu que c'est à la fois l'état *nombre 0-photon* et un état *cohérent*. La fonction d'onde associée à cet état a été donnée en (1.25) par

$$\psi_0(x) = \pi^{-1/4} e^{-x^2/2}, \quad \forall x \in \mathbb{R}.$$

D'après les définitions (1.39), (1.40) de la fonction Wigner et compte tenu des entrées de la matrice de densité du *vide* (1.44)

$$W_{|0\rangle}(q, p) = W_{0,0}(q, p) = \frac{1}{\pi} \int e^{2ipx} \psi_0(q-x) \psi_0(q+x) dx = \frac{1}{\pi} e^{-q^2-p^2}, \quad (1.48)$$

La fonction de Wigner associée au vide est donc une gaussienne centrée de variance  $1/2$  (cf Figure 1.4).

### Les états de Fock

Les états nombres en général, admettent comme fonction d'onde  $\psi_n(x)$  donnée par (1.26)

$$\psi_n(x) := \frac{1}{\sqrt{\sqrt{\pi}2^n n!}} H_n(x) e^{-x^2/2}, \quad \forall x \in \mathbb{R}, \quad n \in \mathbb{N},$$

où  $H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$  dénote le  $n$ -ième polynôme de Hermite.

D'après les définitions (1.39), (1.41) de la fonction Wigner et compte tenu des entrées de la matrice de densité du *vide* (1.44)

$$W_{|n\rangle}(q, p) = W_{n,n}(q, p) = \frac{(-1)^n}{\pi} e^{-q^2 - p^2} L_n(2q^2 + 2p^2),$$

où  $L_n(x)$  dénote le polynôme de Laguerre de degré  $n$ . Plus particulièrement la fonction de Wigner (cf Figure 1.4) associée à l'état nombre *nombre 1-photon* est

$$W_{|1\rangle}(q, p) = \frac{1}{\pi} (2q^2 + 2p^2 + 1) e^{-q^2 - p^2}. \quad (1.49)$$

### Les états cohérents

Il a été mentionné dans la section précédente que les états cohérents peuvent être vus comme un déplacement du *vide*. Dans le livre Leonhardt [65], il a été établi que ce déplacement se reflète sur la fonction de Wigner. Pour  $\alpha = (q_0 + ip_0)/\sqrt{2}$  la fonction de Wigner d'un état *Cohérent- $\alpha$*  est

$$W_{|\alpha_{coh}\rangle}(q, p) = W_{|0\rangle}(q - q_0, p - p_0) = \frac{1}{\pi} e^{-(q-q_0)^2 - (p-p_0)^2}. \quad (1.50)$$

La Figure 1.4 donne les représentations des fonctions de Wigner associées aux matrices de densité de la Figure 1.3.

### Le chat de Schrödinger

La définition d'un *chat de Schrödinger* comme la superposition linéaire d'un état *cohérent- $(\alpha)$*  et d'un état *cohérent- $(-\alpha)$* , a pour conséquence l'apparition d'interférence  $I(q, p)$  (1.46) qui se retrouvent dans la fonction de Wigner

$$W_{|\alpha_{sch}\rangle}(q, p) = \frac{1}{2} W_{|\alpha_{coh}\rangle}(q, p) + \frac{1}{2} W_{|-\alpha_{coh}\rangle}(q, p) + \frac{1}{2} I(q, p),$$

Pour  $\alpha = (q_0 + ip_0)/\sqrt{2}$ , avec  $p_0 = 0$ , la fonction de Wigner d'un *chat de Schrödinger* s'écrit

$$W_{|\alpha_{sch}\rangle}(q, p) = \frac{1}{2\pi} e^{-(q-q_0)^2 - p^2} + \frac{1}{2\pi} e^{-(q+q_0)^2 - p^2} + \frac{1}{\pi} e^{-q^2 - p^2} \cos(2q_0 p), \quad (1.51)$$

et la Figure 1.4 en est une représentation pour  $q_0 = 3$ .

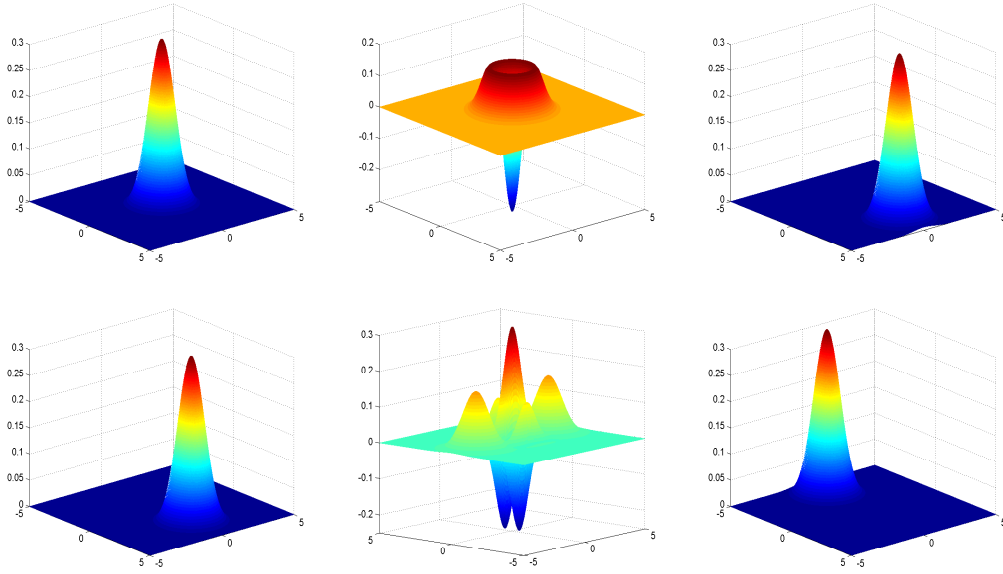


FIG. 1.4 – Représentation des fonctions de Wigner du *vide* (1<sup>er</sup> en haut), de l'état *nombre 1-photon* (2<sup>nd</sup> en haut), de l'état *Coherent- $\alpha$*  avec  $q_0 = 3$  et  $p_0 = 0$  (3<sup>ème</sup> en haut), de l'état *Coherent- $\alpha$*  avec  $q_0 = \sqrt{6}$  et  $p_0 = 0$  (1<sup>er</sup> en bas), du *chat de Schrödinger*,  $q_0 = 3$  (2<sup>nd</sup> en bas) et de l'état *Coherent- $\alpha$*  avec  $q_0 = -3$  et  $p_0 = 0$  (3<sup>ème</sup> en bas).

### Représentations de quelques fonctions de Wigner associées

## 1.2 Formalisme statistique

Cette section est consacrée à la présentation et à la modélisation statistique du problème posé par l'optique quantique. Quelques notions préliminaires liées à la théorie d'estimation et de tests minimax sont d'abord définies.

### 1.2.1 Modèle statistique non-paramétrique

Tout d'abord, rappelons la définition d'un modèle statistique.

#### **Définition 1.6.**

On appelle modèle statistique  $\mathcal{M}$ , le triplet

$$\mathcal{M} = (\mathcal{X}, \Sigma, \{P_f, f \in \mathcal{F}\}),$$

où  $P_f$  est une loi de probabilité sur  $\mathcal{X}$  de paramètre  $f$  appartenant à la classe  $\mathcal{F}$ .

Nous nous intéressons ici au modèle d'une famille d'expériences statistiques et on note  $\mathcal{M}_n$  le modèle d'échantillonnage

$$\mathcal{M}_n = (\mathcal{X}^n, \Sigma^n, \{P_f^n, f \in \mathcal{F}\}), \quad (1.52)$$

où par exemple  $\Sigma^n = \Sigma^{\otimes n}$  et  $P_f^n = P_f^{\otimes n}$  dans le cas d'un échantillon i.i.d..

Le paramètre  $f$  que l'on considère dans cette thèse est soit fonctionnel, soit matriciel, et appartient à une classe  $\mathcal{F} = \mathcal{F}_\nu$  caractérisée par un indice de régularité  $\nu$ ; on se place donc dans un cadre **non-paramétrique**. Pour plus de détails sur l'estimation non-paramétrique nous référons au livre Tsybakov [89]. Nous considérons deux types de problèmes : le problème d'estimation et le problème de test. Nous introduisons brièvement les outils nécessaires à leur étude.

## A/ L'estimation non-paramétrique

L'estimation s'inscrit dans la théorie minimax asymptotique qui permet de mesurer la performance d'une procédure d'estimation uniformément sur une classe non-paramétrique de fonctions.

Le cadre : Soient  $\mathcal{M}_n$  le modèle d'échantillonnage (1.52) et  $\hat{f}$  un estimateur de  $f$  qui prend ses valeurs dans  $\mathcal{E} \supseteq \mathcal{F}_\nu$  et soit

$$d : \mathcal{F}_\nu \times \mathcal{E} \rightarrow \mathbb{R}_+,$$

la distance (ou semi-distance) entre  $\hat{f}$  et  $f$ . On appelle **fonction de perte**, la fonction croissante  $w : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  telle que  $w(0) = 0$  et pour une suite strictement positive  $\{\varphi_{n,\nu}\}_{n \geq 1}$  pour chaque paramètre  $\nu$ , on définit les notions suivantes. On note, dans la suite,  $\varphi_n$  pour  $\varphi_{n,\nu}$  afin d'alléger les notations.

### Définition 1.7.

*Dans le cadre décrit précédemment, le risque maximum de  $\hat{f}$  sur la classe  $\mathcal{F}_\nu$  et pour la suite  $\{\varphi_n\}$  est défini comme*

$$\mathbf{R}_n(\hat{f}, \mathcal{F}_\nu, \varphi_n) := \sup_{f \in \mathcal{F}_\nu} E_f[w(\varphi_n^{-1}d(f, \hat{f}))], \quad (1.53)$$

où  $E_f$  est l'espérance associée à la loi  $P_f^n$ .

**Définition 1.8.**

Dans ce même cadre, la suite  $\{\varphi_n\}$  est une vitesse de convergence sur la classe  $\mathcal{F}_\nu$ , et pour le modèle  $\mathcal{M}_n$  s'il existe une suite d'estimateurs  $\{f_n^*\}$  et une constante  $0 < C < \infty$  telles que

$$\limsup_{n \rightarrow \infty} \mathbf{R}_n(f_n^*, \mathcal{F}_\nu, \varphi_n) \leq C. \quad (1.54)$$

De plus, si pour  $c > 0$

$$\limsup_{n \rightarrow \infty} \inf_{T_n} \mathbf{R}_n(T_n, \mathcal{F}_\nu, \varphi_n) \geq c, \quad (1.55)$$

où  $\inf_{T_n}$  dénote l'infimum sur tous les estimateurs  $T_n$  de  $f$ , la suite  $\{\varphi_n\}$  est appelée vitesse minimax de convergence.

On appelle souvent la majoration (1.54) et la minoration (1.55) respectivement la borne supérieure et inférieure du risque minimax.

**Estimateur adaptatif**

Pour des raisons pratiques, on est amené à construire un estimateur optimal ne requérant pas la connaissance à priori d'un ou de plusieurs paramètres de régularité  $\nu$  dans un certain ensemble de valeurs  $\Theta$ . Cet estimateur est appelé estimateur adaptatif à la régularité.

**Définition 1.9.**

Un estimateur  $f_n^*$  est adaptatif optimal sur l'échelle de classes  $\mathcal{F}_\nu$ ,  $\nu \in \Theta$ , s'il existe une constante finie  $C_0 > 0$  telle que

$$\limsup_{n \rightarrow \infty} \sup_{\nu \in \Theta} \mathbf{R}_n(f_n^*, \mathcal{F}_\nu, \varphi_n) \leq C_0.$$

où  $\varphi_n$  est la vitesse minimax sur la classe  $\mathcal{F}_\nu$  et  $\Theta$  est un ensemble non vide de valeurs.

**Estimateur à noyau**

L'un des estimateurs sur lequel notre intérêt se porte plus particulièrement est l'estimateur à noyau (ou encore estimateur de *Parzen-Rosenblatt*). L'estimateur à noyau est une méthode non-paramétrique d'estimation de la densité de probabilité d'une variable aléatoire  $X$ . Soient  $(X_\ell)_{\ell=1, \dots, n}$  des variables aléatoires i.i.d. de densité de probabilité  $f$  par rapport à la mesure de Lebesgue sur  $\mathbb{R}$ , l'estimateur à noyau  $\hat{f}_{h,n}$  de  $f$  s'écrit

$$\hat{f}_{h,n}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right),$$

où  $K : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction intégrable telle que  $\int K(x)dx = 1$ . La fonction  $K$  est appelée *noyau* (*kernel* en anglais) et le paramètre  $h_n$  est nommé *fenêtre* (*bandwidth* en anglais), qui régit le degré de lissage d'estimation. Un exemple de noyau est le noyau gaussien

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Dans le cadre asymptotique qui nous intéresse  $h = h_n$  dépend de  $n$ , la taille de l'échantillon, avec  $(h_n)_n \geq 0$  une suite qui tend vers 0 quand  $n$  tend vers l'infini. On le notera  $h$  (sans l'indice  $n$ ) pour alléger les notations. Le choix de ce paramètre de lissage est très influent sur l'estimateur et son choix constitue une étape importante de la procédure d'estimation.

## B/ Test d'adéquation

En plus du problème d'estimation, nous considérons le problème non-paramétrique de test d'adéquation à une loi connue  $f_0 \in \mathcal{F}_\nu$  qui consiste à déterminer si un échantillon suit la loi  $f_0$  connue. L'hypothèse nulle  $H_0$  supposera l'adéquation de l'échantillon à cette loi et on note  $H_1(\mathcal{C}, \varphi_n)$  l'alternative :

$$\begin{aligned} H_0 : & \quad f = f_0, \quad f_0 \in \mathcal{F}_\nu, \\ H_1(\mathcal{C}, \varphi_n) : & \quad f \in \mathcal{F}_\nu \text{ telle que } d(f, f_0) \geq \mathcal{C}\varphi_n, \end{aligned}$$

pour la distance  $d$  et  $\{\varphi_n\}_{n \geq 1}$  une suite strictement positive telle que  $\varphi_n \rightarrow 0$  quand  $n \rightarrow \infty$ .

### Définition 1.10.

On appelle ici *procédure de test* et on note  $\Omega_n$  une fonction mesurable de  $(\mathcal{X}^n, \Sigma^n)$  vers  $\{0, 1\}$  tel que  $\Omega_n = 0$  signifie que l'on accepte  $H_0$  et  $\Omega_n = 1$  signifie que l'on accepte  $H_1$ . On définit l'erreur de première espèce et l'erreur de deuxième espèce respectivement par

$$\alpha_n := P_{f_0}^n[\Omega_n = 1] \quad \text{et} \quad \pi_n := \sup_{f \in H_1(\mathcal{C}, \varphi_n)} P_f^n[\Omega_n = 0].$$

### Définition 1.11.

La suite  $\{\varphi_n\}$  est une *vitesse de test* sur la classe  $\mathcal{F}_\nu$  et pour le modèle  $\mathcal{M}_n$  (1.52), si pour  $\lambda \in ]0, 1[$ , fixé, il existe une suite de statistiques de test  $\{\Omega_n^*\}$  et une constante  $0 < \mathcal{C}^* < \infty$  telles que

$$\limsup_{n \rightarrow \infty} \left( P_{f_0}^n[\Omega_n^* = 1] + \sup_{f \in H_1(\mathcal{C}, \varphi_n)} P_f^n[\Omega_n^* = 0] \right) \leq \lambda, \quad (1.56)$$

pour tout  $\mathcal{C} > \mathcal{C}^*$ . De plus, s'il existe pour  $\mathcal{C}_* > 0$

$$\limsup_{n \rightarrow \infty} \inf_{\Omega_n} \left( P_{f_0}^n[\Omega_n = 1] + \sup_{f \in H_1(\mathcal{C}, \varphi_n)} P_f^n[\Omega_n = 0] \right) \geq \lambda, \quad (1.57)$$

pour tout  $0 < \mathcal{C} < \mathcal{C}_*$ , où  $\inf_{\Omega_n}$  dénote l'infimum sur toutes les procédures de test  $\Omega_n$ , on appelle  $\{\varphi_n\}$  vitesse minimax de test.

On appelle souvent la majoration (1.56) et la minoration (1.57) respectivement la borne supérieure et inférieure du risque de test, avec le risque de test défini comme la somme des erreurs de première et deuxième espèce.

### 1.2.2 Problèmes inverses

Dans le formalisme mathématique, un problème inverse se modélise de la façon suivante : Soit  $\mathcal{H}$  et  $\mathcal{K}$  deux espaces de Hilbert et  $\mathbf{K} : \mathcal{H} \rightarrow \mathcal{K}$  un opérateur linéaire, un problème inverse consiste à reconstruire le signal source  $f$  à partir d'observations qui permettent d'estimer  $g = \mathbf{K}f$ . Si l'opérateur  $\mathbf{K}$  est bijectif et l'opérateur  $\mathbf{K}^{-1} : \mathcal{K} \rightarrow \mathcal{H}$  est continu, on dit que le problème inverse est **bien posé**. Dans le cas contraire, le problème inverse est dit **mal posé**.

Exemple : tomographie par émission de positons : (TEP) Un exemple de problème inverse mal posé est la tomographie par émission de positons, apparu dans les années 1970. Cette technique est couramment utilisée en imagerie médicale, par exemple. La TEP consiste à faire l'injection d'un radiotracteur et de faire une cartographie  $\mathfrak{R}[f]$  de sa distribution  $f$  pour chaque coupe suivant plusieurs angles dans chaque coupe :

$$\begin{aligned} \mathfrak{R}[f](x, \phi) &= \iint f(u, v) \delta(u \cos \phi + v \sin \phi - x) du dv \\ &= \int_{-\infty}^{\infty} f(x \cos \phi + t \sin \phi, x \sin \phi - t \cos \phi) dt, \end{aligned}$$

où l'opérateur compact  $\mathfrak{R}[f]$  est appelée transformée de Radon. Soit  $X_1, \dots, X_n$  i.i.d. de densité de probabilité  $\mathfrak{R}[f]$ , un estimateur pour la tomographie est donné pour  $h = h_n > 0$  par

$$\hat{f}_n = \frac{1}{n} \sum_{\ell=1}^n K_h([z, \Phi_\ell] - X_\ell),$$

où  $[z, \phi] := u \cos \phi + v \sin \phi$  avec  $z = (u, v) \in \mathbb{R}^2$  et le noyau

$$K_h(x) = \frac{1}{2\pi} \int e^{-itx} \frac{1}{2\pi} |t| \mathbb{1}_{|t| \leq 1/h} dt.$$



Le problème de la tomographie par émission de positons a été largement étudié en statistique par Korostelëv et Tsybakov [61]. Citons entre autres Johnstone et Silverman [53] qui ont obtenu des vitesses minimax pour l'estimation de densité de probabilité bivariée sur des classes de fonctions de régularité finie et Cavalier [22, 23] pour l'estimation de densité multidimensionnelle pour des fonctions super régulières.

Exemple : modèle de convolution : Un exemple important de problème inverse mal posé est le problème de déconvolution. Soit

$$Y_\ell = X_\ell + \xi_\ell, \quad \ell = 1, \dots, n$$

avec les  $X_\ell$  i.i.d. de densité de probabilité  $f$  et  $\xi_\ell$  des variables aléatoires i.i.d., indépendantes de tous les  $X_\ell$ , de densité de probabilité  $N$  connue. On dénote par  $\Phi_\xi$  la fonction caractéristique de  $\xi_\ell$  telle que  $\Phi_\xi(t) = E[e^{it\xi}]$  non nulle pour tout  $t$ . Dans ce contexte, on observe  $(Y_1, \dots, Y_n)$  i.i.d. de densité de probabilité  $\mathbf{K}f = f * N$ , où  $f * N$  dénote la convolution de  $f$  et  $N$ . Un estimateur à noyau connu de la densité de probabilité de  $f$  a été proposé par Carroll et Hall [21], Stefanski et Carroll [85] et s'écrit

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{\ell=1}^n K^n \left( \frac{x - Y_\ell}{h} \right)$$

où  $K^n$  est le noyau de déconvolution

$$K^n(y) = \frac{1}{2\pi} \int e^{-ity} \frac{\Phi_K(t)}{\Phi_\xi(t/h)} dt$$

avec  $K$  un noyau de transformée de Fourier  $\Phi_K(t) = \int e^{itx} K(x) dx$  et le paramètre  $h > 0$  la fenêtre .

De tels estimateurs à noyau ont été utilisés dans la littérature sur des classes de Sobolev par Fan [36] qui a établi des vitesses de convergence pour la norme  $\mathbb{L}_p$  ( $1 \leq p < \infty$ ). Des vitesses minimax exactes ont été obtenues par Butucea et Tsybakov [19] pour le risque ponctuel et le risque  $\mathbb{L}_2$ . Ils considèrent des fonctions super-régulières  $f$  dont la fonction caractéristique  $\Phi_f$  décroît de façon asymptotiquement exponentielle ( $\exp(-\alpha|u|^r)$ ) et est moins régulière que le bruit  $\Phi_\xi$  à décroissance asymptotique exponentielle ( $\exp(-\beta|u|^s)$ ); i.e.  $0 < r \leq s$  et pour  $\beta, \alpha > 0$ . Le cas  $r = s = 1$  a été traité par Golubev et Khasminskii [42], Tsybakov [88], Cavalier, Golubev, Lepski, et Tsybakov [25]. Des vitesses minimax pour le risque ponctuel ont été étudiées par Butucea [17] pour  $\Phi_\xi$  à décroissance asymptotique polynomiale et une classe de fonctions super-régulières avec  $0 < r \leq 2$ . Pour  $f$  dans une classe

de Sobolev, Goldenshluger [40] a donné des vitesses adaptatives optimales pour le risque ponctuel et pour un bruit super régulier ( $\Phi_\xi$  à décroissance asymptotique exponentielle). Donoho [34] a proposé une méthode appelée WVD (Wavelet-Vaguelet-Decomposition) qui permet de résoudre certains types de problèmes inverses comme la déconvolution, son idée est de trouver une base qui diagonalise simultanément l'opérateur  $\mathbf{K}$  et la classe de fonctions à priori. Johnstone, Kerkycharian, Picard, et Raimondo [54] ont étudié la déconvolution dans un modèle de bruit blanc gaussien. Ils proposent une procédure adaptative par seuillage d'ondelette et ont montré qu'elle est quasi optimale dans  $\mathbb{L}_p, \forall p > 1$ . Cette méthode consiste à ne sélectionner que les coefficients de plus forte contribution dans le développement en ondelettes.

### 1.2.3 Modèle statistique de la TQH

On rappelle que le domaine d'application de ce manuscrit est l'optique quantique et le système physique étudié est un laser monochromatique dans une cavité. On considère ici le modèle d'échantillonnage

$$\mathcal{M}_n = (\mathcal{X}^n, \Sigma^n, \{P_{\mathbf{K}f}^n, f \in \mathcal{F}_\nu\}).$$

Le paramètre inconnu  $f$ , dans ce contexte décrit l'état quantique de ce système physique. Dans l'approche matricielle,  $f = \rho$  et  $\rho$  est l'opérateur de densité (1.8) agissant sur l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  et vérifiant les Propriétés 1.4. Dans l'approche fonctionnelle  $f = W_\rho$  et  $W_\rho : \mathbb{R}^2 \rightarrow \mathbb{R}$  est la fonction de Wigner (1.33) associée. On se place bien dans un cadre non-paramétrique où  $f = \{(\rho_{i,j})_{i,j \geq 1}\}$  ou  $f = \{W_\rho\}$  est respectivement un paramètre matriciel ou fonctionnel appartenant à une classe  $\mathcal{F} = \mathcal{F}_\nu$  caractérisée par un indice de régularité  $\nu$ .

Pour obtenir des informations sur l'état quantique  $f = \{\rho, W_\rho\}$  du système physique, on aimerait dans l'idéal avoir accès à un couple de variables aléatoires  $(Q, P)$  résultant d'une mesure des quadratures  $(\mathbf{Q}, \mathbf{P})$ , qui désignent respectivement les champs électrique et magnétique. La mesure simultanée de ces deux observables étant impossible, il faut en mesurer une quadrature  $(X_\ell, \phi_\ell)$ .

#### Cadre idéal

Dans le cadre d'une **détection idéale** par tomographie quantique homodyne (TQH), on peut collecter des données  $(X_\ell, \phi_\ell)$  d'une quadrature  $\mathbf{X}_\phi$  de ces observables  $\mathbf{Q}$  et  $\mathbf{P}$ , définie en (1.29) par  $\mathbf{X}_\phi = \cos(\phi)\mathbf{Q} + \sin(\phi)\mathbf{P}$ , où  $\Phi$  est choisit de façon uniforme sur  $[0, \pi]$ . La

TQH étant à rapprocher de la TEP, le modèle est un problème inverse **mal posé**. Il a été montré, dans la Section 1.1.3 que les  $(X_\ell, \Phi_\ell)_\ell$  admettent comme densité de probabilité

$$p_\rho(x, \phi) = \frac{1}{\pi} \Re[W_\rho](x, \phi) \mathbb{1}_{[0, \pi]}(\phi), \quad (1.58)$$

où  $\Re[W_\rho]$  est la transformée de Radon de  $W_\rho$  définie par (1.35).

### Cadre pratique

On considère dans nos travaux un modèle plus réaliste. En effet, on a souligné en (1.32) que des erreurs gaussiennes se rajoutent aux observations, des erreurs dont la variance dépend du coefficient  $\eta \in ]0, 1]$ , paramètre connu des physiciens. On collecte alors des données  $(Y_\ell, \Phi_\ell)_\ell$

$$Y_\ell = \sqrt{\eta} X_\ell + \sqrt{(1-\eta)/2} \xi_\ell, \quad (1.59)$$

où les  $\xi_\ell$  sont des variables aléatoires gaussiennes de variance  $(1-\eta)/2$ , i.i.d. et indépendantes des  $X_\ell$ . Les  $(Y_\ell, \Phi_\ell)_\ell$  admettent comme densité de probabilité

$$p_\rho^\eta(y, \phi) = \frac{1}{\pi} \left( \frac{1}{\sqrt{\eta}} \Re[W_\rho] \left( \frac{\cdot}{\sqrt{\eta}}, \phi \right) \right) * N^\eta(y) \mathbb{1}_{[0, \pi]}(\phi), \quad (1.60)$$

où  $N$  est la densité de probabilité Gaussienne de la variable aléatoire  $\xi$  et  $f * g(y)$  dénote le produit de convolution  $f * g(y) = \int f(y-t)g(t)dt$ . On dit que ce problème inverse est **très mal posé** dû à la convolution avec une loi gaussienne.

Cette thèse s'appuie sur le modèle (1.59) donnée par l'optique quantique. Toutes les procédures d'estimation et de tests considérées dans cette thèse sont construites à partir de données bruitées  $(Y_\ell, \Phi_\ell)_\ell$  définies en (1.59) et obtenues par TQH (1.60). Les classes de régularité  $\mathcal{F}_\nu$  sont définies dans la prochaine section pour chaque problème considéré.

## 1.3 Description du contenu de la thèse

L'ensemble des travaux présentés ici s'inscrit dans le cadre asymptotique de la théorie minimax et de tests minimax. Le modèle utilisé est le problème inverse *très mal posé* (*severely ill-posed inverse problem* en anglais) décrit dans la Section 1.2.3. Les Chapitres 2 et 3 sont l'objet d'un article de Aubry, Butucea, et Meziani [6], le Chapitre 4 de l'article de Meziani [71] et le Chapitre 5 de l'article de Meziani [72]. Les différents chapitres peuvent être lus indépendamment les uns des autres, ainsi quelques répétitions dans les définitions sont inévitables.

Nous considérons toujours dans nos travaux le cas de  $n$  mesures identiques et indépendantes par TQH de systèmes quantiques préparés identiquement dans l'état  $\rho$ .

### 1.3.1 Estimation de la matrice de densité

Soit l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$  muni de la base de Fock (1.26). Notons que l'équation (1.37), des entrées de la matrice  $\rho$  sur cette base, s'écrit aussi sous la forme

$$\rho_{j,k} = E_\rho[f_{j,k}(X)e^{i(j-k)\Phi}],$$

avec  $(X, \Phi)$  le couple de variables aléatoires, défini en (1.58), de densité de probabilité  $p_\rho$  et  $f_{j,k}$  les fonctions *pattern* définies en (1.38). Dans le cadre idéal ( $\eta = 1$ ), un estimateur de projection de la mesure empirique sur les fonctions *pattern* est donné par Artiles, Gill, et Guță [4]. L'estimation de la matrice de densité a largement été considérée dans le cadre d'une détection idéale ( $\eta = 1$ ), entre autres par D'Ariano, Macchiavello, et Paris [29], D'Ariano, Leonhardt, et Paul [30], Leonhardt, Paul, et D'Ariano [66], D'Ariano et Paris [28]. L'adaptation par sélection de modèle dans la base de Fock et par ondelettes a été traitée par Kahn [55] dans le cas idéal où  $\eta = 1$ . Seul le cas bruité avec le paramètre d'efficacité  $\eta \in ]\frac{1}{2}, 1]$  a été traité dans la littérature, entre autres, par D'Ariano [26], D'Ariano, Maccone, et Sacchi [33], D'Ariano [27], Richter [78].

Dans le Chapitre 2 et dans Aubry, Butucea, et Meziani [6], on estime la matrice de densité  $\rho$  dans le cas où  $\eta \in ]0, 1]$  et des vitesses de convergences en risque  $\mathbb{L}_2$  sont obtenues. On suppose que  $\rho$  appartient à une classe naturelle pour les physiciens, la classe des matrices dont les éléments diagonaux décroissent de façon exponentielle  $\mathcal{R}(B, r, L)$ , avec  $L, B > 0$  et  $0 < r \leq 2$ ,

$$\mathcal{R}(B, r, L) := \{\rho \text{ état quantique} : |\rho_{m,n}| \leq L \exp(-B(m+n)^{r/2})\}. \quad (1.61)$$

Les procédures d'estimation que l'on considère utilisent des fonctions *pattern*  $f_{j,k}^\eta$  adaptées au cas pratique (1.59),  $\eta \in ]0, 1]$ . De plus, des résultats d'ordre analytique sont obtenus : on mène l'étude du comportement asymptotique des fonctions  $f_{j,k}$  et  $f_{j,k}^\eta$  en norme  $\mathbb{L}_\infty$  et en norme  $\mathbb{L}_2$ . L'étude de la borne supérieure du risque minimax (1.54) est faite en risque  $\mathbb{L}_2$ . Des résultats numériques sont présentés.

### 1.3.2 Estimation de la fonction de Wigner

La fonction de Wigner  $W_\rho$ , définie en (1.33), décrit de façon équivalente à  $\rho$ , l'état d'un système physique. Dans le Chapitre 3, on propose une procédure d'estimation de cette fonction  $W_\rho$ .

Un estimateur de la fonction de Wigner pour  $\eta = 1$  a été donné par Guță et Artiles [43]. Dans le modèle plus réaliste (1.60) où le paramètre d'efficacité  $\eta \in ]0, 1]$ , l'étude a été menée

par Butucea, Guță, et Artiles [20] et des vitesses minimax ont été obtenues pour le risque ponctuel avec la constante asymptotique exacte (quand  $r < 2$ ).

Dans ce cadre, nous étudions au Chapitre 3, l'estimation en risque  $\mathbb{L}_2$  de la fonction de Wigner  $W_\rho$  associée à la matrice de densité  $\rho$ . On suppose que  $\rho$  appartient à la classe matricielle  $\mathcal{R}(B, r, L)$  définie en (1.61). La condition de décroissance exponentielle des coefficients de la matrice de densité  $\rho$  est retranscrite sur la fonction de Wigner associée  $W_\rho$ , ainsi que sur sa transformée de Fourier  $\widetilde{W}_\rho$ . Il est prouvé que les comportements asymptotiques des fonctions  $W_\rho$  et  $\widetilde{W}_\rho$  sont de type exponentiel.

Nous donnons une procédure d'estimation de  $W_\rho$  basée sur une méthode de type noyau. Cette procédure est une version tronquée de l'estimateur en risque ponctuel considéré par Artiles, Gill, et Guță [4] et des vitesses équivalentes sont obtenues pour les bornes supérieures sur des classes différentes.

### 1.3.3 Estimation de la pureté

La question de la pureté d'un état quantique, est un problème qui préoccupe la communauté physique. C'est pourquoi dans le Chapitre 4, on se propose d'estimer la pureté  $\mathcal{P}$  définie en (1.43) d'un état quantique. Dans la Section 1.1.3, l'équation (1.43) permet de définir la pureté d'un état quantique  $\mathcal{P}$  comme

$$\mathcal{P} := \text{Tr}(\rho^2) = 2\pi \iint W_\rho^2(q, p) dq dp \leq 1,$$

et l'état quantique est pur si et seulement si  $\mathcal{P} = 1$ , et l'état quantique est mixte sinon. On se propose dans le Chapitre 4 d'estimer la fonctionnelle quadratique de la fonction de Wigner  $\iint W_\rho^2(q, p) dq dp$  comme estimation de la pureté.

On suppose que la fonction inconnue  $W_\rho$  appartient à une classe plus vaste que celle définie au Chapitre 3. Définissons d'abord la classe  $\mathcal{F}(\alpha, r, L)$  :

$$\mathcal{F}(\alpha, r, L) = \left\{ f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \int_{\mathbb{R}^2} |f(u, v)|^2 e^{2\alpha\|(u, v)\|^r} dudv \leq (2\pi)^2 L \right\},$$

où  $0 < r \leq 2$ ,  $\alpha > 0$ ,  $L > 0$  et  $\|(u, v)\| = \sqrt{u^2 + v^2}$  dénote la norme euclidienne. On suppose que  $W_\rho$  appartient à la classe  $\mathcal{A}(\alpha, r, L, L')$  des fonctions infiniment différentiables

$$\mathcal{A}(\alpha, r, L, L') = \left\{ \begin{array}{l} W_\rho : \mathbb{R}^2 \rightarrow \mathbb{R}, \text{ fonction de Wigner,} \\ W_\rho \in \mathcal{F}(2^r\alpha, r, L'), \widetilde{W}_\rho \in \mathcal{F}(\alpha, r, L), \end{array} \right. \quad (1.62)$$

pour  $\alpha, L, L' > 0$  et  $\widetilde{W}_\rho(u, v) = \iint W_\rho(q, p) e^{iuq+ivp} dq dp$  la transformée de Fourier de  $W_\rho$ .

L'étude d'une fonctionnelle quadratique d'une densité de probabilité a débuté avec Bickel et Ritov [12] où des bornes inférieures non-paramétriques ont été obtenues pour des fonctions dans des classes de Hölder de régularité inférieure à  $1/4$ . Puis Birgé et Massart [14] ont démontré des bornes inférieures pour des fonctionnelles quadratiques plus générales. Dans un cadre minimax Kerkyacharian et Picard [57] ont complété l'étude pour des fonctionnelles de type  $\int f^3$  et Birgé et Massart [14] pour des fonctionnelles de type  $\int T(f)$ . Des vitesses minimax ont été obtenues pour l'estimation de  $\|f\|_r$  pour le modèle de bruit blanc par Lepski, Nemirovski, et Spokoiny [68] et par Butucea [16] dans le modèle de convolution pour l'estimation de  $\int f^2$ .

Le problème de l'estimation adaptative dans le modèle de bruit blanc a été considéré par Efromovich et Low [35] pour  $\int_0^1 f^2$ , par Tribouley [87] pour  $\int T(f)$  avec  $T$  une fonctionnelle 4 fois continuellement différentiable et plus récemment par Klemelä [60] pour l'estimation adaptative avec la constante exacte pour une fonctionnelle quadratique.

Dans le Chapitre 4, on se propose d'estimer la fonctionnelle quadratique  $\iint W_\rho^2$  par une U-statistique d'ordre 2 basée sur un noyau et sur la classe  $\mathcal{A}(\alpha, r, L, L')$  définie en (1.62) pour  $r \in ]0, 2]$  et  $\alpha, L > 0$ . On donne également un estimateur adaptatif pour  $r \in ]0, 1[$  qui ne dépend pas des paramètres de régularité. On étudie les bornes supérieures (1.54) pour le risque  $\mathbb{L}_2$ .

### 1.3.4 Test d'adéquation

L'estimation de matrice de densité au Chapitre 2, nous a renvoyé vers un autre problème statistique, celui des tests d'adéquation à la matrice de densité  $\tau$  et  $\tau$  donnée.

Les physiciens aimeraient vérifier que le laser qu'ils ont créé est bien dans l'état  $\tau$  et pas dans un autre état. On peut pour cela faire un test d'adéquation à l'état  $\tau$  en distance  $\mathbb{L}_2$

$$\begin{aligned} H_0 : & \quad \rho = \tau, \quad \tau \in \mathcal{R}(B, r, L), \\ H_1(\mathcal{C}, \varphi_n) : & \quad \rho \in \mathcal{R}(B, r, L) \text{ telle que } \|\rho - \tau\|_2 \geq \mathcal{C}\varphi_n, \end{aligned} \quad (1.63)$$

où  $\varphi_n$  est une suite qui tend vers 0 avec  $n \rightarrow \infty$ ,

$$\|\rho - \tau\|_2^2 = \sum_{j,k \geq 0} |\rho_{j,k} - \tau_{j,k}|^2 \quad (1.64)$$

est la  $\mathbb{L}_2$ -distance et  $\mathcal{R}(B, r, L)$  est la classe matricielle définie par (1.61).

Le problème de test d'adéquation minimax à une loi donnée à été très largement traité dans la littérature statistique, nous citons entre autres les différents travaux de Ingster [49, 50, 51], Ingster et Suslina [52].

Des procédures adaptatives en distance  $\mathbb{L}_2$  puis en distance  $\mathbb{L}_p$  ont été considérées par Spokoiny [83, 84]. Le problème d'adéquation à une famille paramétrique de fonctions est considéré par Pouet [76], Gayraud et Pouet [38] dans un cadre minimax et par Gayraud et Pouet [39], Fromont et Laurent [37] dans un cadre adaptatif. Lepski et Tsybakov [69] ont établi une constante exacte de test pour le modèle de régression en risque ponctuel et  $\mathbb{L}_\infty$ . Dans le modèle de convolution (avec des observations indirectes), Butucea [16] a obtenu des vitesses minimax.

Dans un contexte quantique, le test à 2 hypothèses simples a été étudié par Kargin [56] sur  $\mathbb{C}^d$  d'un point de vue bayésien et Nussbaum et Szkola [74], Audenaert, Nussbaum, Szkola, et Verstraete [7] ont obtenu des bornes de Chernoff.

Nous étudions, au Chapitre 2, les bornes supérieures (1.56) du risque de test dans le cadre infini dimensionnel où  $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$  de notre procédure de test et nous appuyons nos résultats par des simulations numériques.

### 1.3.5 Perspectives

Les perspectives que nous offre le monde quantique sont nombreuses et en faire ici la liste, ne serait-ce que partielle, me serait impossible, je me limiterai donc à quelques axes de recherche envisageables dans l'immédiat ; dont certains découlent naturellement de ce travail de thèse.

Un des premiers problèmes serait de compléter notre étude en investiguant les bornes inférieures (1.57) et (1.55) dans le cadre de l'estimation de la matrice de densité  $\rho$  et des tests d'adéquation à  $\rho$ . En particulier, il serait intéressant de faire cette étude pour le paramètre de régularité  $r = 2$  défini en (1.61), qui correspond aux états créés actuellement en laboratoire, pour lequel Aubry, Butucea, et Meziani [6] ont obtenu des vitesses polynômiales.

Dans cette thèse, on s'est penché sur la question de savoir si des observations proviennent de l'état quantique souhaité théoriquement. Pour cela, on a proposé des tests d'adéquation à une matrice de densité  $\tau$ . Des tests d'adéquation à la fonction de Wigner associée  $W_\tau$  sont en cours dont la procédure de test est basée sur l'estimateur de la fonctionnelle quadratique  $\iint |W_\tau(q, p)|^2 dq dp$  proposé par Meziani [71].

Le support  $\mathcal{S}$  de la partie négative de la fonction de Wigner défini par

$$\mathcal{S} = \{(q, p) \in \mathbb{R}^2 : W_\rho(q, p) < 0\}$$

est non vide pour un état quantique non-classique. L'estimation du support  $\mathcal{S}$  et le problème de tester si  $\mathcal{S}$  est un ensemble non vide seraient donc un problème particulièrement intéressant d'un point de vue physique.

Un problème différent mais néanmoins tout aussi important, qui pourrait être traité par des tests d'homogénéité, est de déterminer si deux appareils produisent des états quantiques identiques ou différents. Récemment, Butucea et Tribouley [18] ont proposé des méthodes d'ondelettes minimax pour différents risques ( $\mathbb{L}_\infty$ ,  $\mathbb{L}_2$ , ponctuel et sur des boréliens) et ont obtenu des vitesses optimales.

La question de l'adaptivité des procédures d'estimation reste ouverte et on pourrait s'intéresser à une méthode de sélection de la fenêtre plus adaptée à l'échantillon que dans l'approche minimax. Des résultats de type inégalités oracle permettent d'ajuster les paramètres de manière adaptative à la fonction sous-jacente plutôt que d'une manière trop pessimiste comme le propose la théorie minimax (Voir des études dans ce sens, e.g. :Cavalier, Golubev, Lepski, et Tsybakov [25], Cavalier, Golubev, Picard, et Tsybakov [24], Rigollet [80]).

La sélection de modèle pour l'estimation des états quantiques a été considérée par Kahn [55] dans le cadre idéal, i.e. pour  $\eta = 1$ , avec la base de Fock et par ondelettes. Il serait intéressant également d'introduire des ondelettes pour tenter de résoudre les problèmes d'estimation et de tests abordés dans cette thèse. Ceci permettrait peut être d'obtenir des résultats théoriques complémentaires, voire plus généraux, en abordant l'estimation en norme  $\mathbb{L}_p$  et avec des hypothèses différentes sur les quantités à estimer. Nous avons été confrontés en particulier à des problèmes inverses de convolution ou de tomographie, et il faut savoir que des traitements par ondelettes ont été proposés pour ces deux types de problèmes. Concernant la déconvolution, de nombreuses procédures fondées sur des ondelettes ont été développées (voir par exemple Johnstone, Kerkyacharian, Picard, et Raimondo [54]). Des travaux récents ont introduit l'utilisation de nouvelles ondelettes, appelées "needlets". Kerkyacharian, Petrushev, Picard, et Willer [58] ont d'abord utilisé de telles ondelettes dans le cas unidimensionnel, pour traiter des problèmes inverses à "structure polynomiale", dans le



sens où les valeurs propres de l'opérateur ont une décroissance polynômiale. Par la suite, la construction de needlets sur  $\mathbb{R}^2$  a permis de traiter le problème inverse de la transformée de Radon, voir Kerkyacharian, Kyriazis, Le Pennec, Petrushev, et Picard [59]. On pourrait s'inspirer de ces méthodes pour chercher à les appliquer à des problèmes de statistique quantique, mais le bruit gaussien rend le problème plus difficile à traiter par des méthodes d'ondelettes.

Nos résultats sont obtenus pour le risque  $\mathbb{L}_2$ , mais pour leur interprétation physique, l'estimation en risques  $\mathbb{L}_\infty$  et  $\mathbb{L}_1$  de la fonction de Wigner serait également intéressante, mais l'espace  $\mathbb{L}_1$  ne possédant pas de base inconditionnelle, le bon cadre serait de considérer l'espace de Hardy  $H^1$  dont une base a été construite par Maurey [70]. L'espace de Hardy  $H^1$ , composé des fonctions holomorphes sur le disque unité, a une base inconditionnelle d'ondelettes. L'estimation en norme  $\mathbb{L}_1$  sur les opérateurs  $\rho$  seraient d'un grand intérêt en physique. Il s'agit dans ce cas d'évaluer le risque  $E[\|\rho_n - \rho\|_1]$ , où  $\|\rho\|_1 = \text{Tr}(\sqrt{\rho^2}) = \sum_i |\lambda_i|$ , avec  $(\lambda_i)_i$  les valeurs propres de  $\rho$ . De nouvelles méthodes pourraient être développées puisqu'il n'y a plus l'isométrie avec la fonction de Wigner.

D'autres problèmes sont actuellement étudiés dans la littérature quantique. Il a été considéré, récemment, d'un point de vue minimax et d'un point de vue bayésien, le problème de trouver la mesure physique optimale qui permet de différencier parmi deux états  $\rho_1$  et  $\rho_2$  (D'Ariano, Kahn, et Sacchi [32]) et deux canaux de Pauli (D'Ariano, Kahn, et Sacchi [31]). Des problèmes de type apprentissage pourrait être approfondis à la suite de ces travaux. Guță et Kahn [44] se sont intéressés à la normalité asymptotique locale de  $n$  qubits joints  $\rho^{\otimes n}$  et ont donné une équivalence asymptotique d'expériences statistiques en physique quantique. Un axe de recherche naturel sur le long terme serait de rechercher des mesures de  $n$  systèmes intriqués qui maximiseraient l'information contenue par la variable aléatoire résultante.



## Chapter 2

# Nonparametric estimation of the density matrix in quantum homodyne tomography with noisy data

*Abstract:* In the framework of noisy quantum homodyne tomography with efficiency parameter  $0 < \eta \leq 1$ , in this chapter an estimator of the density matrix of a quantum state  $\rho$  is proposed, whose density matrix elements  $\rho_{m,n}$  decrease like  $e^{-B(m+n)r/2}$ , for fixed known  $B > 0$  and  $0 < r \leq 2$ . The procedure estimates the matrix coefficients by a projection method on the pattern functions (that we introduce here for  $0 < \eta \leq 1/2$ ). We compute the convergence rates of these estimators, in  $\mathbb{L}_2$  risk. The proposed methodology is illustrated with simulated data sets.

### 2.1 Introduction

In quantum mechanics, it is difficult to observe physical objects in their full complexity. The compact hermitian operator  $\rho$ , however, catches all information relating to a physical object and describes completely a quantum state. As we cannot measure the quantum state in a single experiment, we perform measurements on identical quantum systems and we construct an estimator of  $\rho$  based on the measurement results.

### 2.1.1 Physical background

Let us briefly summarize the mathematical setting. Physicists are used to describing quantum states on a complex Hilbert space  $\mathcal{H}$  by a density operator  $\rho$  satisfying the following properties:

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

Here, the studied object is a monochromatic light in a cavity. In this framework, we deal with the separable space  $\mathcal{H} = L^2(\mathbb{R})$  and in the basis  $(\psi_j)_{j \in \mathbb{N}}$  the operator  $\rho$  is identified to a *density matrix*  $[\rho_{j,k}]_{j,k \in \mathbb{N}}$  and  $(\psi_j)_{j \in \mathbb{N}}$  is the orthonormal Fock basis

$$\psi_j(x) := \frac{1}{\sqrt{\sqrt{\pi} 2^j j!}} H_j(x) e^{-x^2/2}, \quad (2.1)$$

where  $H_j(x) := (-1)^j e^{x^2} \frac{d^j}{dx^j} e^{-x^2}$  is the  $j$ -th Hermite polynomial.

The matrix elements  $\rho_{j,k}$  of the state  $\rho$  in the Fock basis (2.1) can be expressed as kernel integrals: for all  $j, k \in \mathbb{N}$ ,

$$\rho_{j,k} = \int \int_0^\pi p_\rho(x, \phi) f_{j,k}(x) e^{-i(k-j)\phi} d\phi dx \quad (2.2)$$

where  $f_{j,k} = f_{k,j}$  are bounded real functions called *pattern functions* in quantum homodyne literature. A concrete expression for their Fourier transform using Laguerre polynomials was found in Richter [79]: for  $j \geq k$ ,

$$\begin{aligned} \tilde{f}_{k,j}(t) &= 2\pi^2 |t| \widetilde{W}_{j,k}(t, 0) \\ &= \pi (-i)^{j-k} \sqrt{\frac{2^{k-j} k!}{j!}} |t| t^{j-k} e^{-\frac{t^2}{4}} L_k^{j-k} \left( \frac{t^2}{2} \right). \end{aligned} \quad (2.3)$$

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

Let us state the lemmata which are used to prove upper bounds in Propositions 2.1, 2.2 and 2.3.

#### Lemma 2.1.

There exist constants  $C_2, C_\infty$  such that

$$\sum_{j+k=0}^N \|f_{k,j}\|_2^2 \leq C_2 N^{\frac{17}{6}} \text{ and } \sum_{j+k=0}^N \|f_{k,j}\|_\infty^2 \leq C_\infty N^{\frac{10}{3}}.$$

This is a slight improvement over Artiles, Gill, and Guță [4, Lemma 1] and the proof of this lemma is done Section 2.6.

An equivalent representation for a quantum state  $\rho$  is the associated Wigner function  $W_\rho$ , which can be defined by its Fourier Transform  $\mathcal{F}_2$  with respect to both variables

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

where  $\mathbf{Q}$  and  $\mathbf{P}$  are respectively the electric and magnetic fields. If one can measure simultaneously the quadrature  $\mathbf{Q}$  and  $\mathbf{P}$ , the Wigner function would be seen as the density probability of the result  $(Q, P)$  of the measure. As one cannot, we see it as a quasi-probability density such that

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

Although the Wigner function is in general not positive, its Radon transform is always a probability density

$$\mathcal{R}[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x \cos \phi - t \sin \phi, x \sin \phi + t \cos \phi) dt. \quad (2.4)$$

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.

### 2.1.2 Statistical model

Our aim is to estimate the element of the density matrix defined in (2.2) from data. We collect the data by a technique called quantum homodyne tomography (**QHT**) Smithey, Beck, Raymer, and Faridani [82]. In the ideal setting, we consider  $(X_1, \Phi_1), \dots, (X_n, \Phi_n)$  independent identically distributed random variables with values in  $\mathbb{R} \times [0, \pi]$  and distribution  $P_\rho$  having density

$$p_\rho(x, \phi) = \mathcal{R}[W_\rho](x, \phi), \quad (2.5)$$

where  $\mathcal{R}$  is the Radon transform defined in equation (2.4) with respect to  $\frac{1}{\pi}\lambda$ ,  $\lambda$  being the Lebesgue measure on  $\mathbb{R} \times [0, \pi]$ .

However, in practice, we observe a degraded version of  $(Y_\ell, \Phi_\ell)$ , as a Gaussian noise is added to the observations such that

$$Y_\ell := \sqrt{\eta}X_\ell + \sqrt{(1-\eta)/2} \xi_\ell, \quad (2.6)$$

with  $\xi_\ell$  a sequence of independent identically distributed standard Gaussians which are independent of all  $(X_j, \Phi_j)$  and the detection efficiency parameter  $0 < \eta \leq 1$  is known from the calibration of the apparatus. We denote by  $N^\eta$  the centered Gaussian density of variance  $(1-\eta)/2$  and  $\tilde{N}^\eta$  its Fourier transform. Thus, the density  $p_\rho^\eta$  of  $(Y_\ell, \Phi_\ell)$  can be written as the convolution of the density  $p_\rho(\cdot/\sqrt{\eta}, \phi)/\sqrt{\eta}$  with  $N^\eta$

$$\begin{aligned} p_\rho^\eta(y, \phi) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\eta}} p_\rho\left(\frac{y-x}{\sqrt{\eta}}, \phi\right) N^\eta(x) dx \\ &=: \left(\frac{1}{\sqrt{\eta}} p_\rho\left(\frac{\cdot}{\sqrt{\eta}}, \phi\right) * N^\eta\right)(y). \end{aligned}$$

In the Fourier domain this relation becomes

$$\mathcal{F}_1[p_\rho^\eta(\cdot, \phi)](t) = \mathcal{F}_1[p_\rho(\cdot, \phi)](t\sqrt{\eta})\tilde{N}^\eta(t), \quad (2.7)$$

where  $\mathcal{F}_1$  denotes the Fourier transform with respect to the first variable.

We assume that the unknown state belongs to the class  $\mathcal{R}(B, r)$  for  $B > 0$  and  $0 < r \leq 2$  defined by

$$\mathcal{R}(B, r) := \{\rho \text{ quantum state} : |\rho_{m,n}| \leq \exp(-B(m+n)^{r/2})\}. \quad (2.8)$$

For simplicity, we have chosen to express the results relative to a class which is the intersection of the (positive) ball of radius 1 in some Banach space with the hyperplane  $\text{tr}(\rho) = 1$ . Another radius for the class would only change the constant  $C$  in front of the asymptotic rates of convergence that we will find.

### 2.1.3 Outline of the results

The goal of this chapter is to define an estimator of the density matrix and to evaluate its performance in  $\mathbb{L}_2$  risk. In order to compute estimation risks and to tune the underlying parameters, we define a realistic class of quantum states  $\mathcal{R}(B, r)$ , depending on parameters  $B > 0$  and  $0 < r \leq 2$ , in which the elements of the density matrix decrease rapidly.

The theoretical foundation of quantum homodyne tomography was outlined in Vogel and Risken [91] and has inspired the first experiments determining the quantum state of a light field, initially with optical pulses in Smithey, Beck, Raymer, and Faridani [82], Smithey,

Beck, Cooper, and Raymer [81], Leonhardt [65]. The reconstruction of the density from averages of data has been discussed or studied in D’Ariano, Macchiavello, and Paris [29], D’Ariano, Leonhardt, and Paul [30], Leonhardt, Paul, and D’Ariano [66], Artiles, Gill, and Guță [4] for  $\eta = 1$  (no photon loss). Max-likelihood methods have been studied in Banaszek, D’Ariano, Paris, and Sacchi [9], Artiles, Gill, and Guță [4], D’Ariano, Maccone, and Sacchi [33], Guță [45] and procedure using adaptive tomographic kernels to minimize the variance has been proposed in D’Ariano and Paris [28]. 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’Ariano [26], D’Ariano, Maccone, and Sacchi [33], D’Ariano [27] and considered in Richter [78] via the pattern functions for the diagonal elements.

We derive in Section 2.3 analytic results on the pattern functions which are useful for the reconstruction of the elements of the density matrix. In Section 2.3, we give estimators of the density matrix  $\rho$ . The legend was somehow forged that no estimation of the matrix is possible when  $0 < \eta \leq 1/2$ . The physicists argue, that their machines actually have high detection efficiency, around 0.8/0.9; it is nevertheless satisfying to be able to solve this problem in any noise condition. We give here the so-called *pattern functions* to use for estimating the density matrix in the noisy case with *any* value of  $\eta$  between 0 and 1. These pattern functions allow us to solve an inverse problem which becomes (severely) ill-posed when  $0 < \eta \leq 1/2$ . In this case, we regularize the inverse problem and this introduces a smoothing parameter which we will choose in an optimal way. We compute the upper bounds for the rates achieved by our methods, with  $\mathbb{L}_2$  risk measure. A numerical implementation completes the study of our estimator in Section 2.4.

To conclude, we may infer that the performances of this estimator with the estimator of the Wigner function in Chapter 3 and in Aubry, Butucea, and Meziani [6] are comparable. We obtain nearly polynomial rates for the case  $r = 2$  and intermediate rates for  $0 < r < 2$  (faster than any logarithm, but slower than any polynomial). The estimator of the matrix  $\rho$  can be easily projected on the space of proper quantum states.

## 2.2 Pattern functions

We want to reconstruct the density matrix element from the noisy observations  $(Y_\ell, \Phi_\ell)$  by a projection type estimator on the *pattern* functions. In the presence of noise, it is necessary to adapt the pattern functions as follows. From now on, we shall use the notation

$$\gamma := \frac{1 - \eta}{4\eta}.$$

We consider the case  $\frac{1}{2} < \eta \leq 1$  from the case  $0 < \eta \leq \frac{1}{2}$ . When  $\frac{1}{2} < \eta \leq 1$ , we denote by  $f_{k,j}^\eta$  the function which has the following Fourier transform:

$$\tilde{f}_{k,j}^\eta(t) := \tilde{f}_{k,j}(t)e^{\gamma t^2}. \quad (2.9)$$

When  $0 < \eta \leq \frac{1}{2}$ , we introduce a cut-off parameter  $\delta > 0$  and define  $f_{k,j}^{\eta,\delta}$  via its Fourier transform:

$$\tilde{f}_{k,j}^{\eta,\delta}(t) := \tilde{f}_{k,j}(t)e^{\gamma t^2} \mathbb{I}\left(|t| \leq \frac{1}{\delta}\right). \quad (2.10)$$

Then we compute bounds on these pattern functions.

**Lemma 2.2.**

For  $1 > \eta > 1/2$ , there exist constants  $C_2^\eta$  and  $C_\infty^\eta$  such that

$$\sum_{j+k=0}^N \|f_{k,j}^\eta\|_2^2 \leq C_2^\eta N^{\frac{5}{6}} e^{8\gamma N},$$

and

$$\sum_{j+k=0}^N \|f_{k,j}^\eta\|_\infty^2 \leq C_\infty^\eta N^{\frac{1}{3}} e^{8\gamma N}.$$

The proof of the lemma is done Section 2.6

## 2.3 Density matrix estimation

The aim of this part is to estimate the density matrix  $\rho$  in the Fock basis directly from the data  $(Y_i, \Phi_i)_{i=1,\dots,n}$ . We show that for  $0 < \eta \leq 1/2$  it is still possible to estimate the density matrix with an error of estimation tending to 0 as  $n$  tends to infinity (Theorem 2.3). In both cases ( $\eta > \frac{1}{2}$  and  $\eta \leq \frac{1}{2}$ ), we construct an estimator of the density matrix  $(\rho_{j,k})_{j,k \leq N-1}$  from a sample of QHT data. We give theoretical results for our estimator when the quantum state  $\rho$  is in the class of density matrix with decreasing elements defined in (2.8).

### 2.3.1 Estimation procedure

For  $N := N(n) \rightarrow \infty$  and  $\delta := \delta(n) \rightarrow 0$ , let us define our estimator of  $\rho_{j,k}$  for  $0 \leq j+k \leq N-1$  by

$$\hat{\rho}_{j,k}^\eta := \frac{1}{n} \sum_{\ell=1}^n G_{j,k} \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right), \quad (2.11)$$

where

$$G_{j,k}(x, \phi) := \begin{cases} f_{j,k}^\eta(x) e^{-i(j-k)\phi} & \text{if } \frac{1}{2} < \eta \leq 1 \\ f_{j,k}^{\eta,\delta}(x) e^{-i(j-k)\phi} & \text{if } 0 < \eta \leq \frac{1}{2}. \end{cases}$$



using the pattern functions defined in (2.9) and (2.10). We assume that the density matrix  $\rho$  belongs to the class  $\mathcal{R}(B, r)$  defined in (2.8). In order to evaluate the performance of our estimators we take the  $\mathbb{L}_2$  distance on the space of density matrices  $\|\tau - \rho\|_2^2 := \text{tr}(|\tau - \rho|^2) = \sum_{j,k=0}^{\infty} |\tau_{j,k} - \rho_{j,k}|^2$ . We consider the mean integrated square error (MISE) and split it into a truncature bias term  $b_1^2(n)$ , a regularization bias terms  $b_2^2(n)$  and a variance term  $\sigma^2(n)$ .

$$\begin{aligned} E \left( \sum_{j,k=0}^{\infty} |\hat{\rho}_{j,k}^\eta - \rho_{j,k}|^2 \right) &= \sum_{j+k \geq N} |\rho_{j,k}|^2 + \sum_{j+k=0}^{N-1} |E[\hat{\rho}_{j,k}^\eta] - \rho_{j,k}|^2 \\ &\quad + \sum_{j+k=0}^{N-1} E |\hat{\rho}_{j,k}^\eta - E[\hat{\rho}_{j,k}^\eta]|^2 \\ &=: b_1^2(n) + b_2^2(n) + \sigma^2(n). \end{aligned}$$

The following propositions give upper bounds for  $b_1^2(n)$ ,  $b_2^2(n)$  and  $\sigma^2(n)$  in the different cases  $\eta = 1$ ,  $1/2 < \eta < 1$  or  $0 < \eta \leq 1/2$  and  $r = 2$  or  $0 < r < 2$ . Their proofs are deferred to Section 2.5.

**Proposition 2.1.**

Let  $\hat{\rho}_{j,k}^\eta$  be the estimator defined by (2.11), for  $0 < \eta \geq 1$ , with  $\delta \rightarrow 0$  and  $N \rightarrow \infty$  as  $n \rightarrow \infty$ , then for all  $B > 0$  and  $0 < r \leq 2$ ,

$$\sup_{\rho \in \mathcal{R}(B,r)} b_1^2(n) \leq c_1 N^{2-r/2} e^{-2BN^{r/2}} \quad (2.12)$$

where  $c_1$  is a positive constant depending on  $B$  and  $r$ .

**Proposition 2.2.**

Let  $\hat{\rho}_{j,k}^\eta$  be the estimator defined by (2.11), for  $0 < \eta \leq 1/2$ , with  $N \rightarrow \infty$  as  $n \rightarrow \infty$  and  $1/\delta \geq 2\sqrt{N}$ . In the case  $r = 2$ , for  $\beta := B/(1 + \sqrt{B})^2$  there exists  $c_2$ , while in the case  $0 < r < 2$ , for any  $\beta < B$  there exists  $c_2$  and  $n_0$  such that for  $n \geq n_0$ :

$$\sup_{\rho \in \mathcal{R}(B,r)} b_2^2(n) \leq c_2 N^2 \delta^{4r-12} e^{-\frac{2\beta}{(2\delta)^r} - \frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2}. \quad (2.13)$$

Note that for  $1/2 < \eta \leq 1$  we have  $b_2(n) = 0$  for all  $0 < r \leq 2$  ( $\hat{\rho}_{j,k}^\eta$  is unbiased).

**Proposition 2.3.**

For  $\hat{\rho}_{j,k}^\eta$  the estimator defined by (2.11),

$$\sup_{\rho \in \mathcal{R}(B,r)} \sigma^2(n) \leq c_3 \frac{\delta N^{17/6}}{n} e^{\frac{2\gamma}{\delta^2}} \quad \text{if } 0 < \eta \leq 1/2 \quad (2.14)$$

$$\sup_{\rho \in \mathcal{R}(B,r)} \sigma^2(n) \leq c'_3 \frac{N^{1/3}}{n} e^{8\gamma N} \quad \text{if } 1/2 < \eta < 1 \quad (2.15)$$

$$\sup_{\rho \in \mathcal{R}(B,r)} \sigma^2(n) \leq c''_3 \frac{N^{17/6}}{n} \quad \text{if } \eta = 1 \quad (2.16)$$

where  $c_3$ ,  $c'_3$  and  $c''_3$  are positive constants depending on  $\eta$ .

### 2.3.2 Main results

We measure the accuracy of  $\hat{\rho}_{j,k}^\eta$  by the maximal risk over the class  $\mathcal{R}(B, r)$

$$\limsup_{n \rightarrow \infty} \sup_{\rho \in \mathcal{R}(B,r)} \varphi_n^{-2} E \left( \sum_{j,k=0}^{\infty} |\hat{\rho}_{j,k}^\eta - \rho_{j,k}|^2 \right) \leq C_0. \quad (2.17)$$

where  $C_0$  is a positive constant and  $\varphi_n^2$  is a sequence which tends to 0 when  $n \rightarrow \infty$  and it is the rate of convergence. Cases  $\eta = 1$  (no noise),  $\frac{1}{2} < \eta < 1$  (weak noise) and  $0 < \eta \leq \frac{1}{2}$  (strong noise) are studied respectively in Theorems 2.1, 2.2 and 2.3.

#### Theorem 2.1.

When  $\eta = 1$ , the estimator defined in (2.11) for the model (2.6), where the unknown state belongs to the class  $\mathcal{R}(B, r)$ , satisfies the upper bound (2.17) with

$$\varphi_n^2 = \log(n)^{\frac{17}{3r}} n^{-1}$$

obtained by taking  $N(n) := \left( \frac{\log(n)}{2B} \right)^{\frac{2}{r}}$ .

#### Proof.

With the proposed  $N(n)$  one checks that the bias (2.12) is smaller than the variance (2.16) which is bounded by a constant times  $\log(n)^{\frac{17}{3r}} n^{-1}$ .  $\square$

#### Theorem 2.2.

When  $\frac{1}{2} < \eta < 1$ , the estimator defined in (2.11) for the model (2.6), where the unknown state belongs to the class  $\mathcal{R}(B, r)$ , satisfies the upper bound (2.17) with

- For  $r = 2$ ,

$$\varphi_n^2 = \log(n)^{\frac{12\gamma+B}{3(4\gamma+B)}} n^{-\frac{B}{4\gamma+B}}$$

with  $N(n) := \frac{\log(n)}{2(4\gamma+B)} \left( 1 + \frac{2}{3} \frac{\log(\log n)}{\log(n)} \right)$ .

- For  $0 < r < 2$ ,

$$\varphi_n^2 = \log(n)^{2-r/2} e^{-2BN(n)^{r/2}}$$

where  $N(n)$  is the solution of the equation  $8\gamma N + 2BN^{r/2} = \log(n)$ .

In that case we have  $N(n) = \frac{1}{8\gamma} \log(n) - \frac{2B}{(8\gamma)^{1+r/2}} \log(n)^{r/2} + o(\log(n)^{r/2})$ .

**Proof.**

When  $r = 2$ , the proposed  $N(n)$  ensures that the variance (2.15) is equivalent to the bias (2.12), which is bounded by a constant times  $\log(n)^{\frac{12\gamma+B}{3(4\gamma+B)}} n^{-\frac{B}{4\gamma+B}}$ .

When  $0 < r < 2$ , the proposed  $N(n)$  makes the variance (2.15) bounded by a constant times  $e^{-2BN(n)^{r/2}}$ , which is smaller than the bias, the latter being bounded by a constant times  $N(n)^{2-r/2} e^{-2BN(n)^{r/2}}$ .

The asymptotic expansion of  $N(n)$  is a standard consequence of its definition by the equation  $8\gamma N + 2BN^{r/2} = \log(n)$ .  $\square$

**Theorem 2.3.**

When  $0 < \eta \leq \frac{1}{2}$ , the estimator defined in (2.11) for the model (2.6), where the unknown state belongs to the class  $\mathcal{R}(B, r)$ , satisfies the upper bound (2.17) with

$$\varphi_n^2 = N^{2-r/2} e^{-2BN^{r/2}}$$

where  $N$  and  $\delta$  are solutions of the system

$$\begin{cases} \frac{2\beta}{(2\delta)^r} + \frac{1}{2} \left( \frac{1}{\delta} - 2\sqrt{N} \right)^2 + \frac{2\gamma}{\delta^2} = \log(n) \\ \frac{2\beta}{(2\delta)^r} + \frac{1}{2} \left( \frac{1}{\delta} - 2\sqrt{N} \right)^2 - 2BN^{r/2} = (\log \log(n))^2 \end{cases} \quad (2.18)$$

for arbitrary  $\beta < B$  in the case  $0 < r < 2$  or

$$\begin{cases} \frac{\beta+4\gamma}{2\delta^2} + \frac{1}{2} \left( \frac{1}{\delta} - 2\sqrt{N} \right)^2 - \frac{5}{3} \log(N) = \log(n) \\ \frac{\beta}{2\delta^2} + \frac{1}{2} \left( \frac{1}{\delta} - 2\sqrt{N} \right)^2 - 2BN - 3 \log(N) = 0 \end{cases} \quad (2.19)$$

with  $\beta := \frac{B}{(1+\sqrt{B})^2}$  in the case  $r = 2$ .

Theses bounds are optimal in the sense that (2.18) and (2.19) are obtained by minimizing the sum of the bounds (2.12), (2.13) and (2.14).

**Proof.**

We use the standard notations  $a(n) \sim b(n)$  if  $\frac{a(n)}{b(n)} \rightarrow 1$  and  $a(n) \approx b(n)$  if there exists a constant  $M < \infty$  such that  $\frac{1}{M} \leq \frac{a(n)}{b(n)} \leq M$  for all  $n$ .

Let us first examine the case  $0 < r < 2$ . Remark that the left-hand term of the second equation in (2.18) is strictly negative when  $1/\delta = 2\sqrt{N}$  and increases to  $\infty$  with  $1/\delta$ . This proves that the solution satisfies  $1/\delta > 2\sqrt{N}$  and that Proposition 2.2 applies. Furthermore, if we suppose that  $\frac{1/\delta}{\sqrt{N}}$  is unbounded when  $n \rightarrow \infty$ , then (up to taking a subsequence) by the first equation  $\frac{\frac{1}{\delta}}{\delta^2} \sim \log(n)$  whereas, by subtracting the two,  $\frac{2\gamma}{\delta^2} \sim \log(n)$ , which is contradictory. So  $1/\delta \approx \sqrt{N}$  and we deduce that  $N \approx \log(n)$ . Then (2.13) yields

$$\log \left( \frac{b_2^2(n)}{N^{2-r/2} e^{-2BN^{r/2}}} \right) \leq (4r - 12) \log(\delta) + \frac{r}{2} \log(N) - (\log \log(n))^2 \rightarrow -\infty$$

whereas (2.14) gives

$$\log \left( \frac{\sigma^2(n)}{N^{2-r/2} e^{-2BN^{r/2}}} \right) \leq \log(\delta) + \left( \frac{5}{6} + \frac{r}{2} \right) \log(N) - (\log \log(n))^2 \rightarrow -\infty.$$

We see that the dominant term is the bound (2.12) on  $b_1^2(n)$ , hence the result.

When  $r = 2$ , the same reasoning as above yields  $1/\delta > 2\sqrt{N}$ ,  $1/\delta \approx \sqrt{N}$  and  $N \approx \log(n)$ . Then the right-hand side of (2.13) and (2.14) are of the same order as  $N e^{-2BN}$ , which is the bound (2.12) on  $b_1^2(n)$ .  $\square$

We obtain for the matrix estimation nearly polynomial rates for the case  $\eta = 1$ , non-parametric polynomial rates for the case  $1/2 < \eta < 1$ ,  $r = 2$  and intermediate rates for  $0 < r < 2$  (faster than any logarithm, but slower than any polynomial). Our estimators require knowledge of class parameters  $B$  and  $r$  for matrix estimation. We note that non parametric adaptive estimators of the density matrix via model selection were proposed by Kahn [55] for the ideal case  $\eta = 1$ . It is therefore challenging to provide matrix estimator adaptive to parameters  $(B, R)$ .

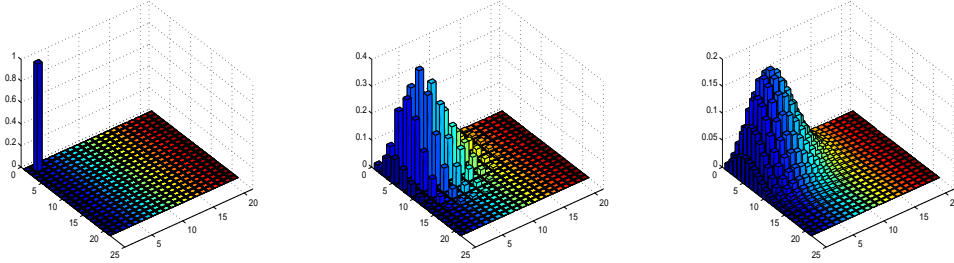
## 2.4 Numerical results

We study in this section the numerical performance of the matrix estimator. We consider successively 3 examples of quantum states, see Figure 2.1. They correspond to a one-photon state  $\rho^1$ , a Schrödinger cat state  $\rho^2$  with  $a = 3$  and a coherent state  $\rho^3$  with parameter  $a = 3$ , having

$$\rho_{m,n}^1 = I(m = n = 1), \quad \rho_{m,n}^2 = \frac{2(a/\sqrt{2})^{m+n} I(m, n \text{ pair})}{\sqrt{m!n!} (e^{-a^2/2} + e^{a^2/2})}, \quad \rho_{m,n}^3 = e^{-a^2/2} \frac{(a/\sqrt{2})^{m+n}}{\sqrt{m!n!}},$$

for  $m \geq 0$ ,  $n \geq 0$ .

Figure 2.1: Density matrices for the one-photon, 3-Schrödinger state and 3-coherent, respectively



For each state  $\rho^j$ ,  $j$  from 1 to 3, we produce a sample of size  $n$  which we bin into a smaller sample of size 500. Next we choose the parameter  $N$  and compute the matrix estimator  $\hat{\rho}_{m,n}^j$  for  $0 \leq m, n \leq N$  and the corresponding  $\mathbb{L}_2$  error  $\|\hat{\rho}^j - \rho^j\|_2^2$ . We repeat the procedure 1000 times and compute the mean error as shown in Table 2.4. The table shows very good performance and it is to be compared to the Wigner function estimator implemented in Butucea, Guță, and Artiles [20].

$\eta$	$n$	$N$	$\rho^1$	$\rho^2$	$\rho^3$
1	1000	8	0.1051	0.0726	0.0645
1	5000	9	0.0291	0.0336	0.0186
1	10000	10	0.0196	0.0149	0.0098
1	50000	11	0.0063	0.0077	0.0039
0.9	5000	8	0.0846	0.0600	0.0554
0.9	10000	9	0.0732	0.0565	0.0315
0.9	50000	10	0.0257	0.0172	0.0111
0.8	50000	8	0.0984	0.0660	0.0590
0.8	100000	8	0.0493	0.0406	0.0447

Table 2.1: Choice of parameter  $N$  and error risk estimators

Examples of matrix estimators are plotted in Figures 2.2 to 2.4. We note that the estimation is more difficult as  $\eta$  gets smaller, in particular for elements  $\rho_{m,n}^j$  of the matrix corresponding to large values of  $m$  and  $n$ . Nevertheless, the risk is getting smaller with the sample size  $n$ , for fixed parameter  $\eta$ .

Figure 2.2: Estimators for the one-photon state for  $(\eta = 1, n = 10000)$ ,  $(\eta = 0.9, n = 50000)$  and  $(\eta = 0.8, n = 100000)$ , respectively

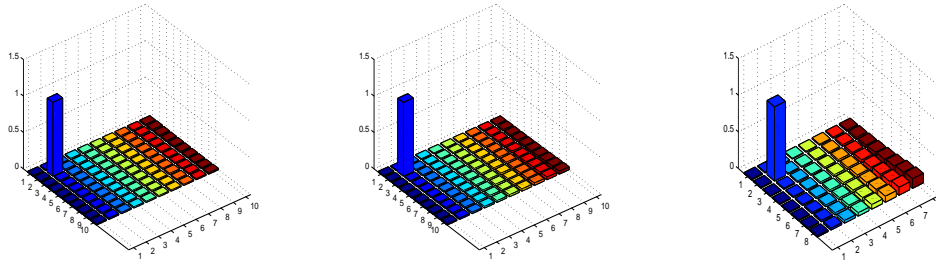


Figure 2.3: Estimators for the 3-Schrödinger state for  $(\eta = 1, n = 10000)$ ,  $(\eta = 0.9, n = 50000)$  and  $(\eta = 0.8, n = 100000)$ , respectively

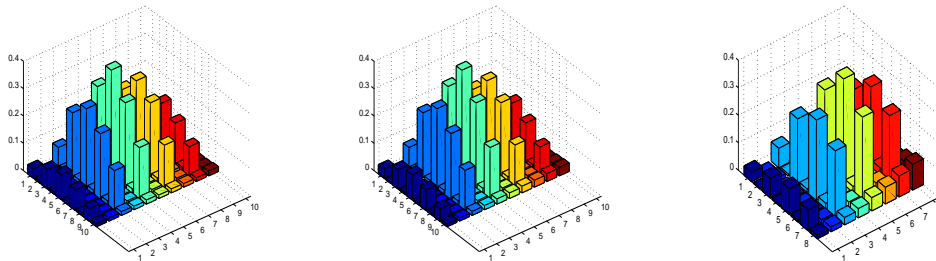
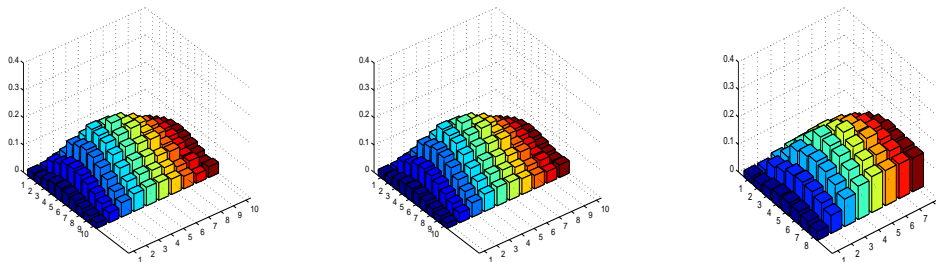


Figure 2.4: Estimators for the 3-coherent state for  $(\eta = 1, n = 10000)$ ,  $(\eta = 0.9, n = 50000)$  and  $(\eta = 0.8, n = 100000)$ , respectively



## 2.5 Proof

### 2.5.1 Proof of Proposition 2.1

By (2.8) the term  $b_1^2(n)$  can be bounded as follows

$$b_1^2(n) = \sum_{j+k \geq N} |\rho_{j,k}|^2 \leq \sum_{j+k \geq N} \exp(-2B(j+k)^{r/2}).$$

Compare to the double integral and change to polar coordinates to get

$$b_1^2(n) \leq c_1 N^{2-r/2} \exp(-2BN^{r/2}).$$

### 2.5.2 Proof of Proposition 2.2

To study the term  $b_2^2(n)$ , we denote

$$\mathcal{F}_1[p_\rho(\cdot|\phi)](t) := E_\rho[e^{itX} | \Phi = \phi] = \widetilde{W}_\rho(t \cos \phi, t \sin \phi),$$

the Fourier transform with respect to the first variable.

$$\begin{aligned} E[\hat{\rho}_{j,k}^\eta] &= E[G_{j,k}(\frac{Y}{\sqrt{\eta}}, \Phi)] = E[f_{j,k}^{\eta,\delta}(\frac{Y}{\sqrt{\eta}})e^{-i(j-k)\Phi}] \\ &= \frac{1}{\pi} \int_0^\pi e^{-i(j-k)\phi} \int f_{j,k}^{\eta,\delta}(y) \sqrt{\eta} p_\rho^\eta(y\sqrt{\eta}|\phi) dy d\phi \\ &= \frac{1}{\pi} \int_0^\pi e^{-i(j-k)\phi} \frac{1}{2\pi} \int \tilde{f}_{j,k}^{\eta,\delta}(t) \mathcal{F}_1[\sqrt{\eta} p_\rho^\eta(\cdot\sqrt{\eta}|\phi)](t) dt d\phi \\ &= \frac{1}{\pi} \int_0^\pi e^{-i(j-k)\phi} \frac{1}{2\pi} \int_{|t| \leq 1/\delta} \tilde{f}_{j,k}(t) e^{\gamma t^2} \mathcal{F}_1[p_\rho(\cdot|\phi)](t) \tilde{N}^\eta(t) dt d\phi. \end{aligned}$$

As  $\tilde{N}^\eta(t) = e^{-\gamma t^2}$  and by using the Cauchy-Schwarz inequality we have

$$\begin{aligned} |E[\hat{\rho}_{j,k}^\eta] - \rho_{j,k}|^2 &= \left| \frac{1}{\pi} \int_0^\pi e^{-i(j-k)\phi} \frac{1}{2\pi} \int_{|t| > 1/\delta} \tilde{f}_{j,k}(t) \mathcal{F}_1[p_\rho(\cdot|\phi)](t) dt d\phi \right|^2 \\ &\leq \frac{1}{\pi} \int_0^\pi \left( \frac{1}{2\pi} \int_{|t| > 1/\delta} |\tilde{f}_{j,k}(t) \widetilde{W}_\rho(t \cos \phi, t \sin \phi)| dt \right)^2 d\phi. \end{aligned}$$

If  $1/\delta \geq 2\sqrt{N} \geq 2s$  with  $s = \sqrt{j+k+1}$ , then whenever  $t \geq 1/\delta$  we get by Lemma 3.2 in Chapter 3 and in Aubry, Butucea, and Meziani [6]

$$\begin{aligned} |\tilde{f}_{j,k}(t)| &= \pi^2 |t| l_{j,k}(t/2) \\ &\leq \pi |t| e^{-\frac{1}{4}(|t|-2s)^2}. \end{aligned}$$

On the other hand, by Propositions 3.1 and 3.2 in Chapter 3 and in Aubry, Butucea, and Meziani [6], we have

$$|\widetilde{W}_\rho(t \cos \phi, t \sin \phi)| \leq A \left(\frac{|t|}{2}\right) e^{-\beta(\frac{|t|}{2})^r}$$

for  $\beta := \frac{B}{(1+\sqrt{B})^2}$  in the case  $r = 2$ , or for arbitrary  $\beta < B$  and  $t$  large enough in the case  $0 < r < 2$ . In both cases  $A$  is a polynomial of degree  $4 - r$ . We deduce the inequality

$$\begin{aligned} |E[\hat{\rho}_{j,k}^\eta] - \rho_{j,k}|^2 &\leq C \left( \int_{\frac{1}{\delta}}^{\infty} t^{5-r} e^{-\frac{1}{4}(t-2s)^2 - \beta 2^{-r} t^r} dt \right)^2 \\ &\leq C \left(\frac{1}{\delta}\right)^{12-4r} e^{-\frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 - \beta 2^{1-r}(\frac{1}{\delta})^r} \end{aligned}$$

by Lemma 8 in Butucea and Tsybakov [19], hence

$$b_2(n)^2 \leq C N^2 \left(\frac{1}{\delta}\right)^{12-4r} e^{-\frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 - \beta 2^{1-r}(\frac{1}{\delta})^r}$$

which covers both cases in the proposition.

### 2.5.3 Proof of Proposition 2.3

Let us write  $\sigma_{j,k}^2(n) := E |\hat{\rho}_{j,k}^\eta - E[\hat{\rho}_{j,k}^\eta]|^2$ . We bound it by

$$\begin{aligned} \sigma_{j,k}^2(n) &= E \left| \frac{1}{n} \sum_{\ell=1}^n \left( G_{j,k} \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) - E \left[ G_{j,k} \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) \right] \right) \right|^2 \\ &= \frac{1}{n} E \left| G_{j,k} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) - E \left[ G_{j,k} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right|^2 \\ &\leq \frac{1}{n} E \left| G_{j,k} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2. \end{aligned} \tag{2.20}$$

#### Proof of (2.14)

For  $0 < \eta \leq 1/2$ , let us denote by  $K_\delta$  the function with the following Fourier transform  $\widetilde{K}_\delta(t) = \mathbb{I}(|t| \leq \frac{1}{\delta}) e^{\gamma t^2}$ , then  $\widetilde{f}_{j,k}^{\eta,\delta} = \widetilde{f}_{j,k}(t) \widetilde{K}_\delta(t)$  and we have

$$\begin{aligned} \sigma_{j,k}^2(n) &\leq \frac{1}{n} E \left| f_{j,k}^{\eta,\delta} \left( \frac{Y}{\sqrt{\eta}} \right) e^{-i(j-k)\Phi} \right|^2 \\ &\leq \frac{1}{n} E \left| f_{j,k} * K_\delta \left( \frac{Y}{\sqrt{\eta}} \right) \right|^2 \\ &\leq \frac{1}{n} E \left| \int f_{j,k}(t) K_\delta \left( \frac{Y}{\sqrt{\eta}} - t \right) dt \right|^2. \end{aligned}$$



By using the Cauchy-Schwarz inequality

$$\begin{aligned}\sigma_{j,k}^2(n) &\leq \frac{1}{n} \int |f_{j,k}(t)|^2 dt E \int \left| K_\delta\left(\frac{Y}{\sqrt{\eta}} - t\right) \right|^2 dt \\ &\leq \frac{1}{n} \int |f_{j,k}(t)|^2 dt E \frac{1}{2\pi} \int \left| \tilde{K}_\delta(u) e^{-iu\frac{Y}{\sqrt{\eta}}} \right|^2 du \\ &\leq \frac{1}{n\pi} \|f_{j,k}\|_2^2 \int_0^{1/\delta} e^{2\gamma u^2} du.\end{aligned}$$

Then,

$$\sigma^2(n) \leq \frac{C}{n\pi} \sum_{j+k=0}^{N-1} \|f_{j,k}\|_2^2 \frac{\eta\delta}{1-\eta} e^{\frac{2\gamma}{\delta^2}}.$$

By Lemma 2.1 we have  $\sum_{j+k=0}^{N-1} \|f_{j,k}\|_2^2 \leq C_2 N^{17/6}$  thus

$$\sigma^2(n) \leq \frac{C_1 \eta \delta N^{17/6}}{n\pi(1-\eta)} e^{\frac{2\gamma}{\delta^2}}.$$

**Proof of (2.15) and (2.16)** By (2.11), for  $1/2 < \eta \leq 1$ ,

$$\begin{aligned}\sigma_{j,k}^2(n) &\leq \frac{1}{n} E \left| f_{j,k}^\eta\left(\frac{Y}{\sqrt{\eta}}\right) e^{-i(j-k)\Phi} \right|^2 \\ &\leq \frac{1}{n\pi} \int_0^\pi \int |f_{j,k}^\eta(y)|^2 \sqrt{\eta} p_\rho^\eta(\sqrt{\eta}y|\phi) dy d\phi \\ &\leq \frac{1}{n\pi} \|f_{j,k}^\eta\|_\infty^2\end{aligned}$$

For  $1/2 < \eta < 1$ , by Lemma 2.2,

$$\sigma^2(n) \leq \frac{C_\infty N^{1/3}}{n\pi} e^{8\gamma N}.$$

For  $\eta = 1$ , by Lemma 3.4 in Chapter 3 and in Aubry, Butucea, and Meziani [6]

$$\begin{aligned}\sigma_{j,k}^2(n) &\leq \frac{1}{n} \int_0^\pi \int |f_{j,k}(x)|^2 p_\rho(x, \phi) dx d\phi \\ &\leq \frac{C}{n} \|f_{j,k}\|_2^2\end{aligned}$$

hence by Lemma 2.1,

$$\sigma^2(n) \leq C \frac{C_2 N^{17/6}}{n}.$$

## 2.6 Proof of Lemmata

### *Proof of Lemma 2.1.*

By symmetry we can restrict the sum to  $j \geq k$ . For fixed  $k$  and  $j$  we have

$$\left\| \tilde{f}_{k,j} \right\|_2^2 = \int_{|t| \leq 2s} \left| \tilde{f}_{k,j}(t) \right|^2 dt + \int_{|t| > 2s} \left| \tilde{f}_{k,j}(t) \right|^2 dt$$

(with  $s = \sqrt{k+j+1}$ ). Because of Lemma 3.2 in Chapter 3 and in Aubry, Butucea, and Meziani [6], it is clear that the second integral is negligible in front of the first one, which we simply bound by  $4s \left\| \tilde{f}_{k,j} \right\|_\infty^2$ .

In view of (2.3), the main result in Krasikov [63] can be rewritten as follows: if  $k \geq 35$  and  $j - k \geq 24$ , then

$$\left\| \tilde{f}_{k,j} \right\|_\infty^2 \leq 2888\pi^2(j+1)^{\frac{1}{2}}k^{-\frac{1}{6}}. \quad (2.21)$$

In consequence, for these values of  $k$  and  $j$ ,

$$\left\| \tilde{f}_{k,j} \right\|_2^2 \leq C(jk^{-\frac{1}{6}} + j^{\frac{1}{2}}k^{\frac{1}{3}}). \quad (2.22)$$

On the other hand, a classical bound on Laguerre polynomials found in Szegő [86] yields that, for fixed values of  $j - k$ ,  $\left\| \tilde{f}_{k,j} \right\|_\infty^2 \leq Ck^{\frac{1}{3}}$ , hence for all  $k \geq 35$  and  $j - k < 24$ ,

$$\left\| \tilde{f}_{k,j} \right\|_2^2 \leq C(j^{\frac{1}{2}}k^{\frac{1}{3}} + k^{\frac{5}{6}}). \quad (2.23)$$

When  $k < 35$ , we can use another result in Krasikov [62] which gives  $\left\| \tilde{f}_{k,j} \right\|_\infty^2 \leq Ck^{\frac{1}{6}}j^{\frac{1}{2}}$  independently of  $j - k$ , thus

$$\left\| \tilde{f}_{k,j} \right\|_2^2 \leq Cj. \quad (2.24)$$

Comparing (2.22), (2.23) and (2.24) we see that when  $N$  is large enough, in the sum over  $0 \leq j, k \leq N$ , the terms  $k \geq 35$ ,  $j - k \geq 24$  dominate and (2.22) yields the first inequality.

The second inequality is obtained by doing a similar computation, starting with  $\|f_{j,k}\|_\infty \leq \left\| \tilde{f}_{j,k} \right\|_1$  and using (2.21) to bound

$$\left\| \tilde{f}_{j,k} \right\|_1^2 \leq C(j^{\frac{3}{2}}k^{-\frac{1}{6}} + j^{\frac{1}{2}}k^{\frac{5}{6}})$$

when  $k \geq 35$  and  $j - k \geq 24$ . □

### *Proof of Lemma 2.2.*

The proof is similar to the previous one and we skip some details. Once again we assume  $j \geq k$  and write

$$\left\| \tilde{f}_{k,j}^\eta \right\|_2^2 = \int_{|t| < 2s} \left| \tilde{f}_{k,j}(t) \right|^2 e^{2\gamma t^2} dt + \int_{|t| > 2s} \left| \tilde{f}_{k,j}(t) \right|^2 e^{2\gamma t^2} dt$$

(where  $s = \sqrt{k + j + 1}$ ). Because of Lemma 3.2 in Chapter 3 and in Aubry, Butucea, and Meziani [6], the second integral is of the same order as the first one, which we bound by

$$\left\| \tilde{f}_{k,j} \right\|_{\infty}^2 \int_{|t| < 2s} e^{2\gamma t^2} dt \leq C \left\| \tilde{f}_{k,j} \right\|_{\infty}^2 s^{-1} e^{8\gamma s^2}.$$

In the sum we are considering the terms  $k \geq 35$  and  $j - k \geq 24$  are dominant and, once again thanks to (2.21), remembering that  $s = \sqrt{j + k + 1}$ ,

$$\left\| \tilde{f}_{k,j}^{\eta} \right\|_2^2 \leq C k^{-\frac{1}{6}} e^{8\gamma(j+k)}$$

hence the first inequality.

The second inequality is, in the same fashion, based on

$$\begin{aligned} \left\| f_{k,j}^{\eta} \right\|_{\infty}^2 &\leq \left\| \tilde{f}_{k,j}^{\eta} \right\|_1^2 \leq C \left( j^{\frac{1}{4}} k^{-\frac{1}{12}} \int_{|t| < 2s} e^{\gamma t^2} dt \right)^2 \\ &\leq C j^{-\frac{1}{2}} k^{-\frac{1}{6}} e^{8\gamma(j+k)} \end{aligned}$$

when  $k \geq 35$  and  $j - k \geq 24$ , and the bound on the sum readily follows.  $\square$



## Chapter 3

# Nonparametric Wigner function estimation in quantum homodyne tomography with noisy data

*Abstract:* In the setting of quantum optic, we translate the exponential decrease condition of the coefficients of the density matrix  $\rho$  assumed in Chapter 2 and in Aubry, Butucea, and Meziani [6] on the corresponding Wigner function  $W_\rho$ , we want to estimate, and its Fourier transform  $W_\rho$ . In the framework of noisy quantum homodyne tomography with efficiency parameter  $0 < \eta \leq 1$ , we propose to estimate the Wigner function  $W_\rho$  by a kernel estimator procedure. We establish the convergence rates of this estimator, in  $\mathbb{L}_2$  risk.

## 3.1 Introduction

Quantum mechanics is a statistical description of small scale physics, where the relevant quantities are measured at the price of a perturbation of the system. Moreover, the result of the measurement is necessarily random. To take this fact into account, the *quantum state* of the system is modeled, not by a scalar or by a probability measure, but by an operator on some Hilbert space. Other operators called *observables* model both the measuring process and the perturbation caused by it. The resulting physics can be counter-intuitive and seemingly paradoxical, but its special properties have important real and potential applications (quantum cryptography, quantum computing).

One of the simplest ways to exhibit quantum behavior uses low intensity lasers and simple devices such as semitransparent mirrors. Experiments in quantum optics consist in creating, manipulating and measuring quantum states of light. We refer to [65] for a detailed exposition. The technique called quantum homodyne tomography allows to retrieve partial, noisy information from which the state is to be recovered: this is the subject of the present chapter.

### 3.1.1 Physical background

Mathematically, the main concepts of quantum mechanics are formulated in the language of selfadjoint operators acting on Hilbert spaces. To every quantum system one can associate a complex Hilbert space  $\mathcal{H}$  whose vectors represent the wave functions of the system. These vectors are identified to projection operators, or pure states. In general, a state is a mixture of pure states described by a compact operator  $\rho$  on  $\mathcal{H}$  having the following properties:

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

When  $\mathcal{H}$  is separable, endowed with a countable orthonormal basis, the operator  $\rho$  is identified to a *density matrix*  $[\rho_{m,n}]_{m,n \in \mathbb{N}}$ .

The positivity property implies that all the eigenvalues of  $\rho$  are nonnegative and by the trace property, they sum up to one. In the case of the finite dimensional Hilbert space  $\mathbb{C}^d$ , the density matrix is simply a positive semi-definite  $d \times d$  matrix of trace one. Our setup from now on will be  $\mathcal{H} = L^2(\mathbb{R})$ , in which case we employ the orthonormal Fock basis made

of the Hermite functions

$$h_m(x) := (2^m m! \sqrt{\pi})^{-\frac{1}{2}} H_m(x) e^{-\frac{x^2}{2}} \quad (3.1)$$

where  $H_m(x) := (-1)^m e^{x^2} \frac{d^m}{dx^m} e^{-x^2}$  is the  $m$ -th Hermite polynomial. Generalizations to higher dimensions are straightforward.

From now on, we denote by  $\langle \cdot, \cdot \rangle$  and  $\| \cdot \|$  the usual Euclidian scalar product and norm, while  $C(\cdot)$  will denote positive constants depending on parameters given in the parentheses.

To each state  $\rho$  corresponds a *Wigner distribution*  $W_\rho$ , which is defined via its Fourier transform in the way indicated by equation (3.2):

$$\widetilde{W}_\rho(u, v) := \iint e^{-i(uq+vp)} W_\rho(q, p) dq dp := \text{Tr}(\rho \exp(-iu\mathbf{Q} - iv\mathbf{P})) \quad (3.2)$$

where  $\mathbf{Q}$  and  $\mathbf{P}$  are canonically conjugate observables (e.g. electric and magnetic fields) satisfying the commutation relation  $[\mathbf{Q}, \mathbf{P}] = i\mathbf{1}$ , with  $\mathbf{1}$  the identity operator (we assume a choice of units such that  $\hbar = 1$ ). It is easily checked that  $W_\rho$  is real-valued, has integral  $\iint_{\mathbb{R}^2} W_\rho(q, p) dq dp = 1$  and uniform bound  $|W_\rho(q, p)| \leq \frac{1}{\pi}$ , for any state  $\rho$ .

For any  $\phi \in \mathbb{R}$ , the Wigner distribution allows one to easily recover the probability density  $x \mapsto p_\rho(x, \phi)$  of  $\mathbf{Q} \cos \phi + \mathbf{P} \sin \phi$  by

$$p_\rho(x, \phi) = \mathfrak{R}[W_\rho](x, \phi), \quad (3.3)$$

where  $\mathcal{R}$  is the Radon transform defined in equation (3.4)

$$\mathcal{R}[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x \cos \phi - t \sin \phi, x \sin \phi + t \cos \phi) dt. \quad (3.4)$$

Moreover, the correspondence between  $\rho$  and  $W_\rho$  is one to one and isometric with respect to the  $\mathbb{L}_2$  norms as in equation (3.5):

$$\|W_\rho\|_2^2 := \iint |W_\rho(q, p)|^2 dq dp = \frac{1}{2\pi} \|\rho\|_2^2 := \frac{1}{2\pi} \sum_{j,k=0}^{\infty} |\rho_{jk}|^2. \quad (3.5)$$

In the Fock basis, we can write  $W_\rho$  in terms of the density matrix  $[\rho_{m,n}]$  as follows (see Leonhardt Leonhardt [65] for the details).

$$W_\rho(q, p) = \sum_{m,n} \rho_{m,n} W_{m,n}(q, p)$$

where

$$W_{m,n}(q, p) = \frac{1}{\pi} \int e^{2ipx} h_m(q-x) h_n(q+x) dx. \quad (3.6)$$

It can be seen that  $W_{m,n}(q, p) = W_{n,m}(q, -p)$  and if  $m \geq n$ ,

$$\begin{aligned} W_{m,n}(q, p) &= \frac{(-1)^m}{\pi} \left( \frac{n!}{m!} \right)^{\frac{1}{2}} e^{-(q^2+p^2)} \\ &\quad \times \left( \sqrt{2}(ip - q) \right)^{m-n} L_n^{m-n}(2q^2 + 2p^2) \end{aligned} \quad (3.7)$$

thus, writing  $z := \sqrt{q^2 + p^2}$ ,

$$l_{m,n}(z) := |W_{m,n}(q, p)| = \frac{2^{\frac{m-n}{2}}}{\pi} \left( \frac{n!}{m!} \right)^{\frac{1}{2}} e^{-z^2} z^{m-n} |L_n^{m-n}(2z^2)| \quad (3.8)$$

where  $L_n^\alpha(x) := (n!)^{-1} e^x x^{-\alpha} \frac{d^n}{dx^n} (e^{-x} x^{n+\alpha})$  is the Laguerre polynomial of degree  $n$  and order  $\alpha$ . Concerning the Fourier transforms, we also recall that

$$\widetilde{W}_{m,n}(q, p) = \frac{(-i)^{m+n}}{2} W_{m,n}\left(\frac{q}{2}, \frac{p}{2}\right). \quad (3.9)$$

We suppose that the unknown state belongs to the class  $\mathcal{R}(B, r)$  for  $B > 0$  and  $0 < r \leq 2$  defined by

$$\mathcal{R}(B, r) := \{\rho \text{ quantum state} : |\rho_{m,n}| \leq \exp(-B(m+n)^{r/2})\}. \quad (3.10)$$

For simplicity, we have chosen to express the results relative to a class which is the intersection of the (positive) ball of radius 1 in some Banach space with the hyperplane  $\text{tr}(\rho) = 1$ . Another radius for the class would only change the constant  $C$  in front of the asymptotic rates of convergence that we will find.

As it will be made precise in Propositions 3.1 and 3.2, quantum states in the class given in (3.10) have fast decreasing and very smooth Wigner functions. From the physical point of view, the choice of such a class of Wigner functions seems to be quite reasonable considering that typical states  $\rho$  prepared in the laboratory do satisfy this type of condition.

### 3.1.2 Statistical model

Let us describe the statistical model. Consider  $(X_1, \Phi_1), \dots, (X_n, \Phi_n)$  independent identically distributed random variables with values in  $\mathbb{R} \times [0, \pi]$  and distribution  $P_\rho$  having density  $p_\rho(x, \phi)$  given by (3.3) with respect to  $\frac{1}{\pi} \lambda$ ,  $\lambda$  being the Lebesgue measure on  $\mathbb{R} \times [0, \pi]$ . The aim is to recover the Wigner function  $W_\rho$  from the observations.

However, there is a slight complication. What we observe are not the variables  $(X_\ell, \Phi_\ell)$  but the noisy ones  $(Y_\ell, \Phi_\ell)$ , where

$$Y_\ell := \sqrt{\eta} X_\ell + \sqrt{(1-\eta)/2} \xi_\ell, \quad (3.11)$$



with  $\xi_\ell$  a sequence of independent identically distributed standard Gaussians which are independent of all  $(X_j, \Phi_j)$ . The detection efficiency parameter  $0 < \eta \leq 1$  is known from the calibration of the apparatus and we denote by  $N^\eta$  the centered Gaussian density of variance  $(1 - \eta)/2$ , respectively by  $\tilde{N}^\eta$  its Fourier transform. Then the density  $p_\rho^\eta$  of  $(Y_\ell, \Phi_\ell)$  is given by the convolution of the density  $p_\rho(\cdot/\sqrt{\eta}, \phi)/\sqrt{\eta}$  with  $N^\eta$

$$\begin{aligned} p_\rho^\eta(y, \phi) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\eta}} p_\rho\left(\frac{y-x}{\sqrt{\eta}}, \phi\right) N^\eta(x) dx \\ &=: \left(\frac{1}{\sqrt{\eta}} p_\rho\left(\frac{\cdot}{\sqrt{\eta}}, \phi\right) * N^\eta\right)(y). \end{aligned}$$

In the Fourier domain this relation becomes

$$\mathcal{F}_1[p_\rho^\eta(\cdot, \phi)](t) = \mathcal{F}_1[p_\rho(\cdot, \phi)](t\sqrt{\eta})\tilde{N}^\eta(t), \quad (3.12)$$

where  $\mathcal{F}_1$  denotes the Fourier transform with respect to the first variable.

### 3.1.3 Outline of the results

The goal of this chapter is to define an estimator of the Wigner function and to compute its performance in  $\mathbb{L}_2$  risk. In order to compute estimation risks and to tune the underlying parameters, we defined in previous section a realistic class of quantum states  $\mathcal{R}(B, r)$ , depending on parameters  $B > 0$  and  $0 < r \leq 2$ , in which the elements of the density matrix decrease rapidly. In Section 3.2, we prove that the fast decay of the elements of the density matrix implies both rapid decay of the Wigner function and of its Fourier transform, allowing us to translate the classes  $\mathcal{R}(B, r)$  in terms of Wigner functions. This is true only when  $r \leq 2$ , and that is the reason for this restriction. The results in this section could not possibly be extended for  $r > 2$  because in that case (3.13) and (3.14) are only satisfied by the null function (a strong form of the uncertainty principle due to Hardy).

In Section 3.3, we study a kernel estimator of the Wigner function in  $\mathbb{L}_2$  risk, over this class of Wigner functions. It is a truncated version of the estimator in Butucea, Guță, and Artiles [20] and tuned accordingly. We compute upper bounds for the rates of convergence of this estimator in  $\mathbb{L}_2$  risk. For the problem of estimating a probability density  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  directly from data  $(X_\ell, \Phi_\ell)$  with density  $\mathcal{R}[f]$  we refer to the literature on X-ray tomography and PET, studied by Vardi, Shepp, and Kaufman [90], Korostel'ev and Tsybakov [61], Johnstone and Silverman [53], Cavalier [22] and many other references therein. In the context of tomography of bounded objects with noisy observations Goldenshluger and Spokoiny [41]

solved the problem of estimating the borders of the object (the support). The estimation of a quadratic functional of the Wigner function has been treated in Meziani [71]. For the problem of Wigner function estimation when no noise is present, we mention the work by Guță and Artiles [43]. They use a kernel estimator and compute sharp minimax results over a class of Wigner functions characterised by their smoothness. In a more recent paper Butucea, Guță, and Artiles [20], Butucea, Guță and Artiles treated the noisy problem for the pointwise estimation of  $W_\rho$ ; however the functions needed to prove minimax optimality there do not belong to the class of Wigner functions that we consider here.

### 3.2 Decrease and smoothness of the Wigner distribution

In this chapter, the classes on which the estimator rate is evaluated will be defined in terms of decay of the density matrix coefficients. For the kernel estimator (Section 3.3) we need to translate these classes in terms of properties of the Wigner function. The next two propositions show how a decrease condition on the coefficients of the density matrix corresponds to a decrease on the corresponding Wigner function and its Fourier transform. First the case  $r < 2$ :

**Proposition 3.1.**

*Assume that  $0 < r < 2$  and that there exists  $B > 0$  such that, for all  $m \geq n$ ,*

$$|\rho_{m,n}| \leq e^{-B(m+n)^{r/2}}.$$

*Then for all  $\beta < B$ , there exists  $z_0$  (depending explicitly on  $r, B, \beta$ , see proof) such that  $z := \sqrt{q^2 + p^2} \geq z_0$  implies*

$$|W_\rho(q, p)| \leq A(z)e^{-\beta z^r} \tag{3.13}$$

*as well as*

$$\left| \widetilde{W}_\rho(q, p) \right| \leq A(z/2)e^{-\beta(z/2)^r} \tag{3.14}$$

*where  $A(z) := \frac{1}{\pi} \left( \sum_{m,n} e^{-B(m+n)^{r/2}} + \frac{4}{Br} z^{4-r} \right)$ .*

If  $r = 2$ , the result is a little different:

**Proposition 3.2.**

*Suppose that there exists  $B > 0$  such that, for all  $m \geq n$ ,*

$$|\rho_{m,n}| \leq e^{-B(m+n)}.$$

Then there exists  $z_0$  such that  $z := \sqrt{q^2 + p^2} \geq z_0$  implies

$$|W_\rho(q, p)| \leq A(z) e^{-\frac{B}{(1+\sqrt{B})^2} z^2} \quad (3.15)$$

as well as

$$\left| \widetilde{W}_\rho(q, p) \right| \leq A(z/2) e^{-\frac{B}{(1+\sqrt{B})^2} (z/2)^2} \quad (3.16)$$

for  $A(z) = \frac{1}{\pi} \left( \sum_{m,n} e^{-B(m+n)} + \frac{2e^B}{B(1+\sqrt{B})^2} z^2 \right)$ .

Note that  $\frac{B}{(1+\sqrt{B})^2} < \min(B, 1)$ . Even when  $B$  is very large, we cannot hope to obtain a faster decrease because  $e^{-z^2}$  is the decrease rate of the basis functions themselves (Lemma 3.2).

The proof of these propositions is deferred to Section 3.4. More general results and converses are studied in Aubry [5].

### 3.3 Wigner function estimation

#### 3.3.1 Kernel estimator

We describe now the estimation method for the Wigner function. In this chapter, as in Butucea, Guță, and Artiles [20], we modify the usual tomography kernel in order to take into account the additive noise on the observations and construct a kernel  $K_h^\eta$  which performs both deconvolution and inverse Radon transform on our data, asymptotically. We follow closely the paper [20] but for studying  $\mathbb{L}_2$  rates, instead of pointwise, an additional truncation of this estimator will be necessary. The class of Wigner functions considered here being slightly different from [20], we sketch the proof of the upper bounds of the risk. Let us define the estimator:

$$\widehat{W}_h^\eta(q, p) = \frac{1}{\pi n} \sum_{\ell=1}^n K_h^\eta \left( q \cos \Phi_\ell + p \sin \Phi_\ell - \frac{Y_\ell}{\sqrt{\eta}} \right), \quad (3.17)$$

where  $0 < \eta < 1$  is a fixed parameter, and the kernel is defined by

$$K_h^\eta(u) = \frac{1}{4\pi} \int_{-1/h}^{1/h} \frac{\exp(-iut)|t|}{\widetilde{N}^\eta(t/\sqrt{\eta})} dt, \quad \widetilde{K}_h^\eta(t) = \frac{1}{2} \frac{|t|}{\widetilde{N}^\eta(t/\sqrt{\eta})} I(|t| \leq 1/h), \quad (3.18)$$

and  $h > 0$  tends to 0 when  $n \rightarrow \infty$  in a proper way to be chosen later. For simplicity, let us denote  $z = (q, p)$  and  $[z, \phi] = q \cos \phi + p \sin \phi$ , then the estimator can be written:

$$\widehat{W}_h^\eta(z) = \frac{1}{\pi n} \sum_{\ell=1}^n K_h^\eta \left( [z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}} \right).$$

This is a one-step procedure for treating two successive inverse problems. The main difference with the noiseless problem treated by Guță and Artiles [43] is that the deconvolution is more ‘difficult’ than the inverse Radon transform. In the literature on inverse problems, this problem is qualified as severely ill-posed, meaning that the noise is dramatically (exponentially) smooth and makes the estimation problem much harder.

### 3.3.2 $\mathbb{L}_2$ risk estimation

We establish next the rates of estimation of  $W_\rho$  from i.i.d. observations  $(Y_\ell, \Phi_\ell)$ ,  $\ell = 1, \dots, n$  when the quality of estimation is measured in  $\mathbb{L}_2$  distance. In the literature,  $\mathbb{L}_2$  tomography is usually performed for boundedly supported functions, see Korostelëv and Tsybakov [61] and Johnstone and Silverman [53]. However, no Wigner function can have a bounded support! Instead, we use the fact that Wigner functions in the class  $\mathcal{R}(B, r)$  decrease very fast and show that a properly truncated estimator attains the rates we may expect from the statistical problem of deconvolution in presence of tomography. Thus, we modify the estimator by truncating it over a disc with increasing radius, as  $n \rightarrow \infty$ . Let us denote

$$D(s_n) = \{z = (q, p) \in \mathbb{R}_2 : \|z\| \leq s_n\},$$

where  $s_n \rightarrow \infty$  as  $n \rightarrow \infty$  will be defined in Theorem 3.1. Let now

$$\widehat{W}_{h,n}^{\eta,*}(z) = \widehat{W}_{h,n}^\eta(z) I_{D(s_n)}(z). \quad (3.19)$$

From now on, we will denote for any function  $f$ ,

$$\|f\|_{D(s_n)}^2 = \int_{D(s_n)} f^2(z) dz,$$

and by  $\overline{D}(s_n)$  the complementary set of  $D(s_n)$  in  $\mathbb{R}^2$ . Then,

$$\begin{aligned} E \left[ \left\| \widehat{W}_h^{\eta,*} - W_\rho \right\|_2^2 \right] &= E \left[ \left\| \widehat{W}_h^\eta - W_\rho \right\|_{D(s_n)}^2 \right] + \|W_\rho\|_{\overline{D}(s_n)}^2 \\ &= E \left[ \left\| \widehat{W}_h^\eta - E \left[ \widehat{W}_h^\eta \right] \right\|_{D(s_n)}^2 \right] + \left\| E \left[ \widehat{W}_h^\eta \right] - W_\rho \right\|_{D(s_n)}^2 \\ &\quad + \|W_\rho\|_{\overline{D}(s_n)}^2. \end{aligned}$$

When replacing the  $\mathbb{L}_2$  norm with the above restricted integral, the upper bound of the bias of the estimator is unchanged, whereas the variance part is infinitely larger than the deconvolution variance in Butucea and Tsybakov [19]. As the bias is dominating over the variance in this setup, we can still choose a suitable sequence  $s_n$  so that the same bandwidth

is optimal associated to the same optimal rate, provided that  $W_\rho$  decreases fast enough asymptotically. The following proposition gives upper bounds for the three components of the  $\mathbb{L}_2$  risk uniformly over the class  $\mathcal{R}(B, r)$ .

**Proposition 3.3.**

Let  $(Y_\ell, \Phi_\ell)$ ,  $\ell = 1, \dots, n$  be i.i.d. data coming from the model (3.11) and let  $\widehat{W}_h^\eta$  be an estimator (with  $h \rightarrow 0$  as  $n \rightarrow \infty$ ) of the underlying Wigner function  $W_\rho$ . We suppose  $W_\rho$  lies in the class  $\mathcal{R}(B, r)$ , with  $B > 0$  and  $0 < r \leq 2$ . Then, for  $s_n \rightarrow \infty$  as  $n \rightarrow \infty$  and  $n$  large enough,

$$\begin{aligned} \sup_{\rho \in \mathcal{R}(B, r)} \|W_\rho\|_{D(s_n)}^2 &\leq C_1 s_n^{10-3r} e^{-2\beta s_n^r}, \\ \sup_{\rho \in \mathcal{R}(B, r)} \left\| E[\widehat{W}_h^\eta] - W_\rho \right\|_{D(s_n)}^2 &\leq C_2 h^{3r-10} e^{-\frac{2^{1-r}\beta}{h^r}}, \\ \sup_{\rho \in \mathcal{R}(B, r)} E \left[ \left\| \widehat{W}_{h,n}^\eta - E[\widehat{W}_{h,n}^\eta] \right\|_{D(s_n)}^2 \right] &\leq C_3 \frac{s_n^2}{nh} \exp\left(\frac{2\gamma}{h^2}\right), \end{aligned}$$

where  $\beta < B$  is defined in Proposition 3.1 for  $0 < r < 2$  and  $\beta = B/(1 + \sqrt{B})^2$  for  $r = 2$ ,  $\gamma = (1 - \eta)/(4\eta) > 0$ ,  $C_1, C_2, C_3$  are positive constants,  $C_1, C_2$ , depending on  $\beta, B, r$  and  $C_3$  depending only on  $\eta$ .

We measure the accuracy of  $\widehat{W}_h^{\eta,*}$  by the maximal risk over the class  $\mathcal{R}(B, r)$

$$\limsup_{n \rightarrow \infty} \sup_{\rho \in \mathcal{R}(B, r)} E \left[ \left\| \widehat{W}_h^{\eta,*} - W_\rho \right\|^2 \right] \varphi_n^{-2}(\mathbb{L}_2) \leq C. \quad (3.20)$$

where  $C$  is a positive constant and  $\varphi_n^2$  is a sequence which tends to 0 when  $n \rightarrow \infty$  and it is the rate of convergence.

In the following Theorem we see the phenomenon which was noticed already: deconvolution with Gaussian type noise is a much harder problem than inverse Radon transform (the tomography part).

**Theorem 3.1.**

Let  $B > 0$ ,  $0 < r \leq 2$  and  $(Y_\ell, \Phi_\ell)$ ,  $\ell = 1, \dots, n$  be i.i.d. data coming from the model (3.11). Then  $\widehat{W}_h^{\eta,*}$  defined in (3.19) with kernel  $K_h^\eta$  in (3.18) satisfies the upper bound (3.20) with

- For  $r = 2$ , put  $\beta = B/(1 + \sqrt{B})^2$

$$\varphi_n^2 = (\log n)^{\frac{16\gamma+3\beta}{8\gamma+2\beta}} n^{-\frac{\beta}{4\gamma+\beta}},$$

$$\text{with } s_n = (h)^{-1} \text{ and } h = \left( \frac{2}{4\gamma+\beta} \log n + \frac{1}{4\gamma+\beta} \log(\log n) \right)^{-1/2}.$$

- For  $0 < r < 2$  and  $\beta < B$  as in Proposition 3.1,

$$\varphi_n^2 = h^{3r-10} \exp\left(-\frac{2^{1-r}\beta}{h^r}\right),$$

where  $s_n = (h)^{-1}$  and  $h$  is the solution of the equation

$$\frac{2^{1-r}\beta}{h^r} + \frac{2\gamma}{h^2} = \log n - (\log \log n)^2.$$

**Sketch of proof of the upper bounds.**

By Proposition 3.3, we get

$$\begin{aligned} \sup_{W_\rho \in \mathcal{R}(B,r)} E \left[ \left\| \widehat{W}_h^\eta - W_\rho \right\|^2 \right] &\leq C_1 s_n^{10-3r} e^{-2\beta s_n^r} + C_2 h^{3r-10} \exp\left(-\frac{2\beta}{(2h)^r}\right) \\ &\quad + \frac{C_3 s_n^2}{nh} \exp\left(\frac{2\gamma}{h^2}\right). \\ &=: A_1 + A_2 + A_3 \end{aligned}$$

For  $0 < r < 2$  and by taking derivatives with respect to  $h$  and  $s_n$ , we obtain that the optimal choice verifies the following equations:

$$\begin{aligned} 2\beta s_n^r + \frac{2\gamma}{h^2} &= \log(n) + \log(h s_n^{2(4-r)}) \\ \frac{2^{1-r}\beta}{h^r} + \frac{2\gamma}{h^2} &= \log(n) + \log(h^{2r-7} s_n^{-2}). \end{aligned}$$

We notice therefore that  $A_2$  is dominating over  $A_3$ , which is dominating over  $A_1$ . The proposed  $(s_n, h)$  ensure that the term  $A_2$  is still the dominating term and gives the rate of convergence.

The case  $r = 2$  is treated similarly, by taking derivatives we notice that the term  $A_2$  and the term  $A_3$  are of the same order and that the term  $A_1$  is smaller than the others.  $\square$

Compared to the density matrix estimation (Aubry, Butucea, and Meziani [6], Chapter 2), we obtain for the Wigner function estimation only polynomial rates for  $r = 2$  and intermediate rates elsewhere. This may be explained by the fact that the class  $\mathcal{R}(B, r)$  is smaller than the class of Wigner functions verifying decrease and smoothness conditions that we actually use for the upper bounds. We may infer that the performances of Wigner function and density matrix (Chapter 2 and Aubry, Butucea, and Meziani [6]) estimators are comparable. On the other hand, we may capture some features of the quantum states more easily on the Wigner function, for instance when this function has significant negative

parts, the fact that the quantum state is non classical.

For the estimation of the Wigner function in the real case ( $\eta < 1$ ), we can easily suggest, following Butucea, Guță, and Artiles [20], a slightly different kernel estimator adaptive to parameters  $\beta > 0$  and  $0 < r < 1$  and satisfying the same upper bounds.

### 3.4 Other Proofs

We prepare the proofs of Propositions 3.1 and 3.2 with a few general utility lemmata.

#### Lemma 3.1.

Let  $y$  and  $w$  be two  $C^2$  functions:  $[x_0, +\infty) \rightarrow (0, +\infty)$  such that  $y'(x) \rightarrow 0$ ,  $w$  is bounded, satisfying the differential equations

$$\begin{aligned} y''(x) &= \phi(x)y(x) \\ w''(x) &= \psi(x)w(x), \end{aligned}$$

with continuous  $\phi(x) \leq \psi(x)$ , and initial conditions  $y(x_0) = w(x_0)$ . Then for all  $x \geq x_0$ ,  $w(x) \leq y(x)$ .

#### Proof of Lemma 3.1.

Suppose that there exists  $x_1 \geq x_0$  where  $w(x_1) > y(x_1)$ . Then for some  $x_2 \in [x_0, x_1]$  we have  $w'(x_2) > y'(x_2)$  and  $w(x_2) \geq y(x_2)$ . Consequently, for all  $x \geq x_2$ ,  $w''(x) - y''(x) \geq 0$ , and  $w'(x) - y'(x) \geq w'(x_2) - y'(x_2)$ . When  $x \rightarrow \infty$ ,  $\liminf w'(x) \geq w'(x_2) - y'(x_2) > 0$ , which contradicts the boundedness of  $w$ .  $\square$

This lemma is used to prove a bound on the Laguerre functions.

#### Lemma 3.2.

For all  $m, n \in \mathbb{N}$  and  $s := \sqrt{m+n+1}$ , for all  $z \geq 0$ ,

$$l_{m,n}(z) \leq \frac{1}{\pi} \begin{cases} 1 & \text{if } 0 \leq z \leq s \\ e^{-(z-s)^2} & \text{if } z \geq s. \end{cases} \quad (3.21)$$

#### Proof of Lemma 3.2.

When  $z \leq s$ , the result follows from the uniform bound on Wigner functions obtained by applying the Cauchy-Schwarz inequality to (3.6).

When  $z \geq s$ ,  $L_n^\alpha(2z^2)$  doesn't vanish and keeps the same sign as  $L_n^\alpha(2s^2)$ . Now, as it can be seen from Szegő [86, 5.1.2], the function  $w(z) := \sqrt{z}l_{m,n}(z)$  satisfies the differential

equation  $w'' = (4(z^2 - s^2) + \frac{\alpha^2 - 1/4}{z^2})z$ . On the other hand,  $y(z) := \sqrt{s}l_{m,n}(s)e^{-(z-s)^2}$  satisfies  $y'' = (4(z-s)^2 - 2)y$ . When  $z \geq s$ ,

$$4(z-s)^2 - 2 < 4(z^2 - s^2) + \frac{\alpha^2 - 1/4}{z^2} \quad (3.22)$$

from which we conclude with Lemma 3.1 that  $w(z) \leq y(z)$ .  $\square$

Finally, a lemma to bound the tail of a series.

**Lemma 3.3.**

*If  $\nu > 0$  and  $C > 0$ , there exists a  $z_0$  such that  $z \geq z_0$  implies*

$$\sum_{m+n \geq z} e^{-C(m+n)^\nu} \leq \frac{2}{C\nu} z^{2-\nu} e^{-Cz^\nu}. \quad (3.23)$$

***Proof of Lemma 3.3.***

First notice that

$$\sum_{m+n \geq z} e^{-C(m+n)^\nu} = \sum_{t \geq z} (t+1)e^{-Ct^\nu} \leq \int_z^\infty (t+1)e^{-Ct^\nu} dt.$$

When  $t \geq z$  and  $z$  is large enough, we have

$$\begin{aligned} \int_z^\infty (t+1)e^{-Ct^\nu} dt &\leq \frac{2}{C\nu} \int_z^\infty (C\nu t - (2-\nu)t^{1-\nu})e^{-Ct^\nu} dt \\ &\leq \frac{2}{C\nu} z^{2-\nu} e^{-Cz^\nu} \end{aligned}$$

which is what we needed to prove.  $\square$

***Proof of the Proposition 3.1.***

Let  $\phi(z) := (z - \sqrt{\beta}z^{r/2})^2 - 1$ . Since  $r < 2$ , for  $z$  larger than a certain  $z_0$  (which depends only on  $\beta$ ,  $B$  and  $r$ ), it is true that  $\phi(z) \geq (\frac{\beta}{B})^{2/r} z^2$ . It follows that

$$e^{-B\phi(z)^{r/2}} \leq e^{-\beta z^r} \quad (3.24)$$

If  $m+n \leq \phi(z)$ , then  $s \leq \sqrt{1 + \phi(z)}$  and  $z-s \geq z - \sqrt{1 + \phi(z)} = \sqrt{\beta}z^{r/2}$ . By (3.21), this means that  $l_{m,n}(z) \leq \frac{1}{\pi} e^{-\beta z^r}$ . So

$$\sum_{m+n \leq \phi(z)} |\rho_{m,n}| l_{m,n}(z) \leq A e^{-\beta z^r} \quad (3.25)$$

for  $A := \frac{1}{\pi} \sum_{m,n} e^{-B(m+n)^{r/2}}$ .



On the other hand, using Lemma 3.3 with  $\nu := r/2$ , if  $\phi(z) \geq z_0$ ,

$$\begin{aligned} \sum_{m+n \geq \phi(z)} |\rho_{m,n}| l_{m,n}(z) &\leq \frac{4}{\pi B r} \phi(z)^{2-r/2} e^{-B\phi(z)^{r/2}} \\ &\leq \frac{4}{\pi B r} z^{4-r} e^{-\beta z^r} \end{aligned} \quad (3.26)$$

by (3.21) and (3.24). Combining (3.25) and (3.26) yields the announced result. The bound on  $\widetilde{W}_\rho$  is then a direct consequence of (3.9).  $\square$

**Proof of the Proposition 3.2.**

Let  $\phi(z) := \theta z^2 - 1$ , where  $\theta := \frac{1}{(1+\sqrt{B})^2}$  is the solution in  $(0, 1)$  of  $(1 - \sqrt{\theta})^2 = B\theta$ .

When  $m + n \leq \phi(z)$ , then  $s \leq \sqrt{\theta}z$  and  $z - s \geq z(1 - \sqrt{\theta}) = \sqrt{B\theta}z$ . By (3.21), this means that  $l_{m,n}(z) \leq \frac{1}{\pi} e^{-B\theta z^2}$ . So

$$\sum_{m+n \leq \phi(z)} |\rho_{m,n}| l_{m,n}(z) \leq A e^{-B\theta z^2} \quad (3.27)$$

for  $A := \frac{1}{\pi} \sum_{m,n} e^{-B(m+n)}$ .

On the other hand, by Lemma 3.3, if  $\phi(z) \geq z_0$ ,

$$\begin{aligned} \sum_{m+n \geq \phi(z)} |\rho_{m,n}| l_{m,n}(z) &\leq \frac{2}{\pi B} \phi(z) e^{-B\phi(z)} \\ &\leq \frac{2\theta e^B}{\pi B} z^2 e^{-B\theta z^2} \end{aligned} \quad (3.28)$$

by (3.21) and (3.24). Combining (3.27) and (3.28) yields the announced result. The bound on  $\widetilde{W}_\rho$  is then a direct consequence of (3.9).  $\square$

**Proof of the Proposition 3.3.**

It is easy to see that

$$\mathcal{F} \left[ E[\widehat{W}_h^\eta] \right] (w) = \widetilde{W}_\rho(w) I(\|w\| \leq 1/h).$$

We have, for  $n$  large enough  $s_n \geq z_0$  and by (3.13)

$$\begin{aligned} \|W_\rho\|_{D(s_n)}^2 &\leq C(B, r) \int_{\|z\| > s_n} \|z\|^{8-2r} \exp(-2\beta\|z\|^r) dz \\ &\leq C(B, r) \int_0^{2\pi} \int_{s_n}^\infty t^{9-2r} \exp(-2\beta t^r) dt d\phi \\ &\leq C_1 s_n^{10-3r} e^{-2\beta s_n^r}, \end{aligned}$$

where  $\beta < B$  and for  $n$  large enough in the case  $0 < r < 2$ , respectively  $\beta = B/(1 + \sqrt{B})^2$  in the case  $r = 2$ . Now we write for the  $\mathbb{L}_2$  bias of our estimator:

$$\begin{aligned} \|E[\widehat{W}_h^\eta] - W_\rho\|_{D(s_n)}^2 &\leq \|E[\widehat{W}_h^\eta] - W_\rho\|_2^2 = \frac{1}{(2\pi)^2} \|\mathcal{F} [E[\widehat{W}_h^\eta]] - \widetilde{W}_\rho\|_2^2 \\ &= \frac{1}{(2\pi)^2} \int \left| \widetilde{W}_\rho(w) \right|^2 I(\|w\| > 1/h) dw \\ &\leq \frac{C^2(B, r)}{(2\pi)^2} \int_{\|w\| > 1/h} \|w\|^{2(4-r)} e^{-2^{1-r}\beta\|w\|^r} dw \\ &\leq C_2 h^{3r-10} e^{-\frac{2^{1-r}\beta}{h^r}}, \end{aligned}$$

by the assumption on our class and (3.14), for  $0 < r < 2$ . The case  $r = 2$  is similar.

As for the variance of our estimator:

$$\begin{aligned} V[\widehat{W}_h^\eta] &= E \left[ \left\| \widehat{W}_h^\eta - E[\widehat{W}_h^\eta] \right\|_{D(s_n)}^2 \right] \\ &= \frac{1}{\pi^2 n} \left\{ E \left[ \left\| K_h^\eta \left( [\cdot, \Phi] - \frac{Y}{\sqrt{\eta}} \right) \right\|_{D(s_n)}^2 \right] \right. \\ &\quad \left. - \left\| E \left[ K_h^\eta \left( [\cdot, \Phi] - \frac{Y}{\sqrt{\eta}} \right) \right] \right\|_{D(s_n)}^2 \right\}. \end{aligned} \quad (3.29)$$

On the one hand, by using two-dimensional Plancherel formula and the Fourier transform shown above, we get:

$$\left\| E \left[ K_h^\eta \left( [\cdot, \Phi] - \frac{Y}{\sqrt{\eta}} \right) \right] \right\|_{D(s_n)}^2 \leq \pi^2 \int |W_\rho(w)|^2 dw \leq \pi^2. \quad (3.30)$$

In the last inequality we have used the fact that  $\|W_\rho\|_2^2 = \text{Tr}(\rho^2) \leq 1$  where  $\rho$  is the density matrix corresponding to the Wigner function  $W_\rho$ .

On the other hand, the dominant term in the variance will be given by

$$\begin{aligned} &E \left[ \left\| K_h^\eta \left( [\cdot, \Phi] - \frac{Y}{\sqrt{\eta}} \right) \right\|_{D(s_n)}^2 \right] \\ &= \int_0^\pi \int \int_{D(s_n)} (K_h^\eta([z, \phi] - y/\sqrt{\eta}))^2 dz p_\rho^n(y, \phi) dy d\phi \\ &= \int_0^\pi \int_{D(s_n)} \int (K_h^\eta(u))^2 \sqrt{\eta} p_\rho^n([z, \phi] - u) \sqrt{\eta}, \phi) dudz d\phi \\ &= \int (K_h^\eta(u))^2 \int_{D(s_n)} \int_0^\pi p_\rho(\cdot, \phi) * NN^\eta([z, \phi] - u) d\phi dz du \\ &\leq M(\eta) \pi s_n^2 \int (K_h^\eta(u))^2 du, \end{aligned}$$

using Lemma 3.4 below and the constant  $M(\eta) > 0$  depending only on  $\eta$ , defined therein. Indeed, let us note that  $\sqrt{\eta}p_\rho^\eta(\cdot, \sqrt{\eta}, \phi)$  is the density of  $Y/\sqrt{\eta} = X + \sqrt{(1-\eta)/(2\eta)}\varepsilon$  and let us call  $NN^\eta$  the Gaussian density of the noise as normalized in this last equation.

Let us first compute, by Plancherel formula,  $\|K_h^\eta\|_2^2$  and get

$$\begin{aligned} \|K_h^\eta\|_2^2 &= \frac{1}{2\pi} \int |\tilde{K}_h^\eta(t)|^2 dt = \frac{1}{2\pi} \int_{|t| \leq 1/h} \frac{t^2}{4\tilde{N}^2(t\sqrt{(1-\eta)/(2\eta)})} dt \\ &= \frac{1}{4\pi} \int_0^{1/h} t^2 \exp\left(t^2 \frac{1-\eta}{2\eta}\right) dt \\ &= \frac{1}{4\pi h} \frac{\eta}{1-\eta} \exp\left(\frac{1-\eta}{2\eta h^2}\right) (1 + o(1)), \text{ as } h \rightarrow 0. \end{aligned}$$

We replace in the second order moment, then as  $h \rightarrow 0$

$$E \left[ \left\| K_h^\eta \left( [\cdot, \Phi] - \frac{Y}{\sqrt{\eta}} \right) \right\|_{D(s_n)}^2 \right] \leq \frac{M(\eta)s_n^2}{16\gamma h} \exp\left(\frac{2\gamma}{h^2}\right) (1 + o(1)). \quad (3.31)$$

The result about the variance of the estimator is obtained from (3.29)-(3.31).  $\square$

**Lemma 3.4.**

For every  $\rho \in \mathcal{R}(B, r)$  and  $0 < \eta < 1$ , we have that the corresponding probability density  $p_\rho$  satisfies

$$\begin{aligned} 0 &\leq \int_0^\pi p_\rho(\cdot, \phi) * NN^\eta(x) d\phi \leq M(\eta), \\ 0 &\leq \int_0^\pi p_\rho(x, \phi) d\phi \leq C \end{aligned}$$

for all  $x \in \mathbb{R}$  eventually depending on  $\phi$ , where  $M(\eta) > 0$  is a constant depending only on fixed  $\eta$  and  $C > 0$ .

**Proof of Lemma 3.4.**

Indeed, using inverse Fourier transform and the fact that  $|\widetilde{W}_\rho(w)| \leq 1$  we get:

$$\begin{aligned} &\left| \int_0^\pi p_\rho(\cdot, \phi) * NN^\eta(x) d\phi \right| \\ &\leq \left| \int_0^\pi \frac{1}{2\pi} \int e^{-itx} \mathcal{F}_1[p_\rho(\cdot, \phi)](t) \cdot \widetilde{NN}^\eta(t) dt d\phi \right| \\ &\leq c(\eta) \int_0^\pi \int |\widetilde{W}_\rho(t \cos \phi, t \sin \phi)| \exp\left(-\frac{t^2(1-\eta)}{4\eta}\right) dt d\phi \\ &\leq c(\eta) \int \frac{1}{\|w\|} |\widetilde{W}_\rho(w)| \exp\left(-\frac{\|w\|^2(1-\eta)}{4\eta}\right) dw \leq M(\eta), \end{aligned}$$

where  $c(\eta)$ ,  $M(\eta)$  are positive constants depending only on  $\eta \in (0, 1)$ .  $\square$



## Chapter 4

# Nonparametric estimation of the purity of a quantum state in Quantum Homodyne Tomography with noisy data

*Abstract:* The aim of this chapter is to answer an important issue in quantum mechanics, namely to estimate the purity of a quantum state of a light beam. The estimation of the purity is based on results of quantum homodyne measurements performed on independent identically prepared quantum systems. The quantum state of the light is entirely characterized by the Wigner function which can take negative values and must verify certain constraints of positivity imposed by quantum physics. We estimate the integrated squared Wigner function by a kernel-based second order U-statistic. This quadratic functional is a physical measure of the purity of the state. We also give an adaptive estimator that does not depend on the smoothness parameters. We establish upper bounds of the minimax risk over a class of infinitely differentiable functions.

## 4.1 Introduction

In quantum mechanics, the quantum state of a system completely describes all aspects of the system. The instantaneous state of a quantum system encodes the probabilities of its measurable properties, or "observables" (examples of observables include energy, position, momentum and angular momentum). Generally, quantum mechanics does not assign deterministic values to observables. Instead, it makes predictions about probability distributions; that is, the probability of obtaining each of the possible outcomes when measuring an observable. In many applications of quantum information, one of the important elements which affect the result of quantum process, is the purity of quantum states produced or utilized. Hence, an interesting and important problem in quantum information is to estimate the purity of a quantum system Bagan, Ballester, R., Monras, and Munoz-Tapia [8], Zhang, Wang, and Ying [93]. This problem is also closely related to the estimation of the entanglement of multiparty systems Horodecki [48], Aolita [3]. 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. The measurement technique considered in this chapter is called Quantum Homodyne Tomography (QHT) and has been put in practice for the first time in Smithey, Beck, Raymer, and Faridani [82]. We will detail this technique in Section 4.2.2. A quantum state is represented through two mathematical objects: the density matrix  $\rho$  and the associated real function of two variables  $W_\rho$  called the Wigner function Wigner [92].

In this chapter we address the problem of estimating the quadratic functional  $d^2 = \int W_\rho^2$  of the Wigner function of a monochromatic light in a cavity prepared in the state  $\rho$  by using QHT data measurement performed on independent, identical systems. Our model takes into account the detection losses occurring in the measurement, leading to an additional additive Gaussian noise as we describe in Section 4.2.4. In the physics literature the purity is defined as  $Tr(\rho^2)$  (the trace) which is 1 for pure states. It is shown that  $d^2 = \int W_\rho^2(z)dz = \frac{1}{2\pi}Tr(\rho^2)$ , thus the quantity  $d^2 = \int W_\rho^2$  has an interest in itself as a physical measure of the purity of quantum state. It allows us to distinguish between pure state and mixed state since it always equals  $\frac{1}{2\pi}$  in case of pure states (see Section 4.2.1 for relation between this quantity and the notion of purity) and is smaller than  $\frac{1}{2\pi}$  if the state is mixed.

In general,  $W_\rho$  can be seen as a joint density of the electric and magnetic field of a light beam. However, since these quantities cannot be measured simultaneously, the Wigner function is not a probability density but may take negative values for certain "non-classical states". The

marginals of  $W_\rho$  along any direction in the plane are bona-fide probability densities, and correspond to the measurement of linear combinations of the electric and magnetic fields.

The problem of estimating quadratic functionals of a probability density was studied in details since Bickel and Ritov [12], where the problem of estimating the integral of the squared derivative of a probability density function was considered and nonparametric rates were obtained. These results have been extended to the estimation of general functionals of a density  $f$  of the type  $\int f^3$  in Kerkyacharian and Picard [57] and of the type  $\int T(f)$  in Birgé and Massart [14] where minimax rates have been established. Minimax rates have also been obtained in Lepski, Nemirovski, and Spokoiny [68] for the nonparametric estimation of  $\|f\|_r$  in the classical white noise model. More recently, the estimation of  $\int f^2$  in the convolution model have been treated in Butucea [16] and applied to the goodness-of-fit test in  $L_2$  distance.

The problem of adaptive estimation of general functionals in the white noise model has been considered in Efromovich and Low [35] for  $\int_0^1 f^2$ , in Tribouley [87] for  $\int T(f)$  for arbitrary 4 times continuously differentiable functionals  $T$  and more recently in Klemelä [60] for sharp adaptive estimation of quadratic functionals.

In a positron emission tomography (PET) perspective, the problem of estimating a probability density from tomographic data with minimax rates has been treated in Johnstone and Silverman [53] for bivariate density and in Cavalier [22] for multi-dimensional density. Some functional estimation problems in the image model, like estimating the area of an image, have been considered in Korostelev and Tsybakov [61].

Quantum statistical models are more recent, the estimation of the Wigner function  $W_\rho$  has been treated in Guță and Artiles [43] in the case of ideal detection, that is without noise. The estimation of the Wigner function in our noisy model has been studied in a nonparametric framework in Butucea, Guță, and Artiles [20] where the minimax rate was established for the pointwise risk and in Aubry, Butucea, and Meziani [6] where the upper bounds for the  $\mathbb{L}_2$  risk were given.

We emphasize that in our chapter we do not restrict ourselves to the parametric setting, but suppose that the Wigner function belongs to a nonparametric class  $\mathcal{A}(\alpha, r, L, L')$  described in Section 4.2.4. We refer the interested reader to Artiles, Gill, and Guță [4], Barndorff-Nielsen, Gill, and Jupp [11] for further details on the physical background.

In this chapter we propose a kernel estimator for the quantity  $d^2 = \int W_\rho^2$ . We investigate the rate of convergence of the procedure and compute the bandwidth leading to the optimal

bias-variance trade-off. This bandwidth depends on the parameter describing the functional class containing  $W_\rho$ . Therefore we propose an adaptive estimator based on a data-driven choice of the bandwidth. This adaptive estimator is shown to have the same rate of convergence. Let us briefly describe a possible application of our results to goodness-of-fit test in  $\mathbb{L}_2$ -norm in quantum statistics. The physical interpretation of such a test is to check whether the produced light pulse is in a known quantum state  $\rho_0$ , or not. This can be done via the Wigner functions as follows:

$$\begin{cases} H_0 : & W_\rho = W_{\rho_0}, \\ H_1 : & \sup_{W_\rho \in \mathcal{A}(\alpha, r, L, L')} \|W_\rho - W_{\rho_0}\|_2 \geq c \cdot \varphi_n \end{cases}$$

where  $\varphi_n$  is a sequence which tends to 0 when  $n \rightarrow \infty$  and it is the testing rate and  $\mathcal{A}(\alpha, r, L, L')$  is a class of smooth Wigner functions (see Section 4.2.4). We can devise a test statistic based on the estimator of  $d^2 = \int W_\rho^2$  constructed in this chapter. Similarly to Butucea [16] we conjecture that the testing rates are of the same order as the ones found in this chapter.

The rest of the chapter is organized as follows. In Sections 4.2.1 and 4.2.2, we give a short introduction to quantum mechanics. We formulate the statistical model in Section 4.2.4. In Section 4.3 we construct an estimator of the quadratic functional of the unknown Wigner function, and state a result on upper bound on the bias and the variance terms (proof in Section 4.4). Then we propose a choice of bandwidth independent of the smoothness parameters yielding the same rate of convergence. Our main theoretical results are presented in Section 4.3.3. We present some examples of quantum states in Section 4.2.3.

## 4.2 Physical and statistical context

### 4.2.1 A short introduction to Quantum Mechanics

In this subsection, we give a short introduction to quantum mechanics. For more physical background we refer the interested reader to the textbook Leonhardt [65] on quantum optics and quantum tomography, the paper Barndorff-Nielsen, Gill, and Jupp [11] on quantum statistical inference and the textbook Holevo [47] and the paper Helstrom [46] on quantum statistics. From a statistical point of view we refer to the paper Butucea, Guță, and Artiles [20] where the statistical problem of estimating the Wigner function of a beam of light given the same statistical model is studied. Quantum mechanics is a fundamental branch of theoretical physics, in the sense that it provides accurate and precise descriptions for



many phenomena on the atomic and subatomic level. Generally, quantum mechanics only makes predictions about probability distributions over possible outcomes from measuring an **observable**  $\mathbf{X}$  (position, number of photons...). Naturally, these probabilities will depend on the quantum state at the instant of the measurement. To every quantum state one can associate a complex Hilbert space  $\mathcal{H}$  with the inner product  $\langle \cdot, \cdot \rangle$  whose vectors represent pure state. Each state  $\rho$  is defined as a linear operator on the space  $\mathcal{H}$  having the following properties:

- (i) Self-adjoint (or Hermitian):  $\rho = \rho^*$ , where  $\rho^*$  is the adjoint of  $\rho$ .
- (ii) Positive:  $\rho \geq 0$ , or equivalently  $\langle \psi, \rho \psi \rangle \geq 0$  for all  $\psi \in \mathcal{H}$ .
- (iii) Trace one:  $\text{Tr}(\rho) = 1$ .

In this chapter, the studied quantum system is a monochromatic light in a cavity, whose state is described by **density matrices**  $\rho$  on the Hilbert space of complex valued square integrable functions on the line  $L_2(\mathbb{R})$  with  $(\psi_j)_{j \in \mathbb{N}}$  an orthonormal basis of the space. Notice that the convex combination of two density matrices is again a density matrix. 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. Thus, pure states are represented by one dimensional orthogonal projection operators i.e.  $\rho = \mathbf{P}_\psi$ . All other states are called mixed states and for a separable Hilbert space  $\mathcal{H}$ , they can be expressed as

$$\rho = \sum_i^{\dim \mathcal{H}} \rho_i \mathbf{P}_{\psi_i},$$

where  $\rho_i \geq 0$  are the eigenvalues of  $\rho$  and  $\mathbf{P}_{\psi_i}$  is the projection onto the one dimensional space generated by the eigenvector  $\psi_i \in \mathcal{H}$  of  $\rho$ . Due to the properties (i)-(iii) of  $\rho$  we have  $\rho_i \geq 0$  for all  $i = 1, \dots, n$  and  $\sum_i \rho_i = 1$ .

Mathematically, an observable is described by a selfadjoint operator  $\mathbf{X} = \mathbf{X}^*$  on the Hilbert space  $\mathcal{H}$ .

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

where  $x_a \in \mathbb{R}$  are the eigenvalues of  $\mathbf{X}$  and  $\mathbf{P}_a$  is the projection onto the one dimensional space generated by the eigenvector of  $\mathbf{X}$ . We denote by  $X$  the result of the measurement of  $\mathbf{X}$  where  $X$  is a random variable with values in the set  $\{x_1, x_2, \dots\}$ . For a quantum

system prepared in the state  $\rho$ ,  $X$  has the following probability distribution, expectation and characteristic function

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

In this chapter we deal with the measurement of two observables  $\mathbf{Q}$  and  $\mathbf{P}$  which are respectively the electric and magnetic fields. As one can not measure simultaneously  $\mathbf{Q}$  and  $\mathbf{P}$ , one measures only the quadrature observables  $\mathbf{X}_\phi := \mathbf{Q} \cos \phi + \mathbf{P} \sin \phi$  and the Wigner function  $W_\rho : \mathbb{R}^2 \rightarrow \mathbb{R}$  may be defined. In general,  $W_\rho$  is regarded as a generalized joint probability density of the two observables  $\mathbf{Q}$  and  $\mathbf{P}$ . The Wigner function may take negative values but it integrates to plus one over the whole plane. It also satisfies certain intrinsic positivity constraints stemming from the positivity of the density operator. Along any direction  $\phi \in [0, \pi]$  in the plane the Radon transform of  $W_\rho$

$$\Re[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x \cos \phi + t \sin \phi, x \sin \phi - t \cos \phi) dt \quad (4.1)$$

correspond to the measurement of  $\mathbf{X}_\phi$ . The Wigner function  $W_\rho$  can be defined by the property of its Fourier transform  $\mathcal{F}_2$  with respect to both variables

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

The Wigner function  $W_\rho$  describes completely the quantum state  $\rho$  as it is in one-to-one correspondence with the density matrix  $\rho$ . The important relation between  $\rho$  and  $W_\rho$  is the following one

$$2\pi \int_{\mathbb{R}^2} W_\rho^2(z) dz = \text{Tr}(\rho^2) = \text{Tr} \left( \left( \sum_i \rho_i P_{\psi_i} \right)^2 \right) = \text{Tr} \left( \sum_i \rho_i^2 P_{\psi_i} \right) = \sum_i \rho_i^2,$$

as  $P_{\psi_i} P_{\psi_j} = 0$  for  $i \neq j$ , since  $\psi_i$  are mutually orthogonal eigenvectors. Then if  $\sum_i \rho_i^2 = 1$ , then  $\rho_i = \delta_{ij}$  for all  $i \in \mathbb{N}$ , for some fixed  $j$ . Thus  $\rho = P_{\psi_j}$  is a pure state. If  $\sum_i \rho_i^2 < 1$  then at least 2 eigenvalues are non-zero and  $\rho$  is a mixed state. We propose in this chapter to study the quantity  $\int_{\mathbb{R}^2} W_\rho^2(z) dz$  as a physical measure of purity.

### 4.2.2 Quantum Homodyne Tomography

The theoretical foundation of quantum homodyne tomography was outlined in Vogel and Risken [91] and has inspired the first experiments determining the quantum state of a light field, initially with optical pulses in Smithey, Beck, Raymer, and Faridani [82], Smithey,

Beck, Cooper, and Raymer [81], Leonhardt [65]. The physicists prepare a monochromatic laser in state  $\rho$  in a cavity. In order to study it, one performs QHT measurements. The measurement procedure sketched in Figure 4.1 consists of mixing the light pulse in which we are interested with a reference laser of high intensity  $|z| \gg 1$  called local oscillator. Just before the mixing the experimenter chooses the phase  $\Phi$  of the local oscillator, randomly, uniformly distributed. After the mixing, the beam is split again. There are two emerging beams and each one is measured to give integrated currents  $I_1, I_2$  proportional to the intensities. In the case of noiseless measurement and for a phase  $\Phi$ , the result would be  $X = \frac{I_2 - I_1}{|z|}$ . In practice, the detection process is inefficient and an independent gaussian noise interferes additively with the ideal data  $X$ . Thus, the effective result of the QHT measurement (Figure 4.1) is  $Y = \sqrt{\eta}X + \sqrt{(1-\eta)/2}\xi$  for known efficiency  $\eta$  and  $\xi$  is a standard Gaussian random variable, independent of  $X$ .

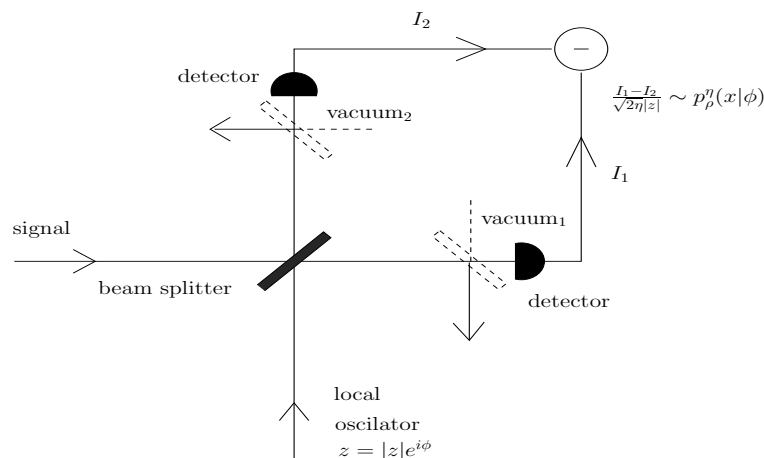


Figure 4.1: QHT measurement scheme

### 4.2.3 Examples

Table 4.1 follows Artiles, Gill, and Guță [4] and shows five examples of pure quantum states and one example of mixed state which can be created at this moment in laboratory. Among the pure states we consider the vacuum state which is the pure state of zero photons, the single photon state, the coherent state which characterizes the laser pulse with an average of  $N$  photons. The squeezed states (see e.g. Breitenbach, Schiller, and Mlynek [15]) have Gaussian Wigner functions whose variances in the two directions have a fixed product. The well-known Schrödinger Cat state is also a pure state described by a linear superposition of

two coherent vectors (see e.g. Ourjountsev, Tualle-Brouri, Laurat, and Grangier [75]). Note that for pure states  $d^2 = \frac{1}{2\pi}$  and for the thermal state which is a mixed state  $d^2 = \frac{\tanh(\beta/2)}{2\pi}$ .

State	Fourier transform of Wigner function $\widetilde{W}_\rho(u, v)$	the Wigner function $W_\rho(p, q)$
Vacuum state	$\exp\left(\frac{-\ (u, v)\ _2^2}{4}\right)$	$\frac{1}{\pi} \exp(-q^2 - p^2)$
Single photon state	$\left(1 - \frac{\ (u, v)\ _2^2}{2}\right) \exp\left(\frac{-\ (u, v)\ _2^2}{4}\right)$	$\frac{1}{\pi} (2q^2 + 2p^2 - 1) \exp(-q^2 - p^2)$
Schrödinger Cat $X_0 > 0$	$e^{-\frac{\ (u, v)\ _2^2}{4}} (\cos(2uX_0) + e^{-X_0^2} \cosh(X_0v)) / (2(1 + e^{-X_0^2}))$	$e^{-p^2} \left( e^{-(q-X_0)^2} + e^{-(q+X_0)^2} + 2 \cos(2pX_0) e^{-q^2} \right) / (2\pi(1 + e^{-X_0^2}))$
Coherent state $N \in \mathbb{R}_+$	$\exp\left(\frac{-\ (u, v)\ _2^2}{4} + i\sqrt{N}v\right)$	$\frac{1}{\pi} \exp(-(q - \sqrt{N})^2 - p^2)$
Squeezed state $N \in \mathbb{R}_+, \xi \in \mathbb{R}$	$\exp\left(-\frac{u^2}{4} e^{2\xi} - \frac{v^2}{4} e^{-2\xi} + iv\alpha\right)$	$\frac{1}{\pi} \exp(-e^{2\xi}(q - \alpha)^2 - e^{-2\xi}p^2)$
Thermal state $\beta > 0$	$\exp\left(\frac{-\ (u, v)\ _2^2}{4(\tanh(\beta/2))^2}\right)$	$\frac{\tanh(\beta/2)}{\pi} \exp(-(q^2 + p^2) \tanh(\beta/2))$

Table 4.1: Examples of quantum states

#### 4.2.4 Problem formulation

In the present chapter we estimate the integral of the square of the Wigner function from data measurement performed on  $n$  identical quantum systems.

Our statistical problem can be formulated as follows: consider  $(X_1, \Phi_1) \dots (X_n, \Phi_n)$  independent identically distributed random variables with values in  $\mathbb{R} \times [0, \pi]$ . The probability density of  $(X, \Phi)$  equals the **Radon transform**  $\mathfrak{R}[W_\rho]$  of the Wigner function defined by (4.1) with respect to the measure  $\lambda/\pi$ , where  $\lambda$  is the Lebesgue measure on  $\mathbb{R} \times [0, \pi]$ . Thus

$$p_\rho(x|\phi) := \mathfrak{R}[W_\rho](x, \phi) \quad (4.2)$$

is the density of  $X$  given  $\Phi = \phi$ . As we explained in Section 4.2.3 we do not observe the ideal data  $(X_\ell, \Phi_\ell)$   $\ell = 1, \dots, n$  but a degraded noisy version  $(Y_1, \Phi_1) \dots (Y_n, \Phi_n)$ , where

$$Y_\ell := \sqrt{\eta}X_\ell + \sqrt{(1-\eta)/2}\xi_\ell. \quad (4.3)$$

Here  $\xi_\ell$  are independent standard Gaussian random variables, independent of all  $(X_k, \Phi_k)$ ,  $k = 1, \dots, n$ . The detection efficiency  $\eta$ ,  $0 < \eta < 1$  is a known parameter and  $1 - \eta$  represents

the proportion of photons which are not detected due to various losses in the measurement process. We denote  $p_\rho^\eta(x, \phi)$  the density of  $(Y_\ell, \Phi_\ell)$ . Thus,  $p_\rho^\eta(\cdot, \phi)$  is the convolution of the density  $\frac{1}{\sqrt{\eta}}p_\rho(\frac{\cdot}{\sqrt{\eta}}, \phi)$  of  $\sqrt{\eta}X$  with the density of a centered Gaussian density having variance  $(1 - \eta)/2$ . Let us define the following functional class  $\mathcal{F}(\alpha, r, L)$ :

$$\mathcal{F}(\alpha, r, L) = \left\{ f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \int_{\mathbb{R}^2} |f(u, v)|^2 e^{2\alpha\|(u, v)\|^r} dudv \leq (2\pi)^2 L \right\},$$

where  $0 < r \leq 2$ ,  $\alpha > 0$ ,  $L > 0$  and  $\|(u, v)\| = \sqrt{u^2 + v^2}$  is the euclidian norm.

In this chapter, we assume that the unknown Wigner function  $W_\rho$  belongs to the class  $\mathcal{A}(\alpha, r, L, L')$  of infinitely differentiable functions where

$$\mathcal{A}(\alpha, r, L, L') = \left\{ \begin{array}{l} W_\rho : \mathbb{R}^2 \rightarrow \mathbb{R}, \text{ Wigner function,} \\ W_\rho \in \mathcal{F}(2^r\alpha, r, L'), \widetilde{W}_\rho \in \mathcal{F}(\alpha, r, L) \end{array} \right. \quad (4.4)$$

for some  $\alpha, L, L' > 0$  where  $\widetilde{W}_\rho(u, v)$  denotes the Fourier transform of  $W_\rho$  w.r.t two variables.

**Remark 4.1.** All the typical states prepared in laboratory have density matrix with fastly decreasing diagonal elements: for some  $C > 0$ ,  $B > 0$  and  $r \in ]0, 2]$

$$|\rho_{m, \ell}| \leq C \exp(-B\alpha(m^{r/2} + \ell^{r/2})) \quad \forall m, \ell \in \mathbb{N}. \quad (4.5)$$

Moreover the density matrix  $\rho$  and the Wigner function  $W_\rho$  defined in Section 4.2.1 are in one to one correspondance. Recently, it was shown in Aubry, Butucea, and Meziani [6] that quantum states satisfying (4.5) have Wigner function  $W_\rho$  in a class  $\mathcal{A}(\alpha, r, L, L')$  for some  $\alpha > 0$ ,  $L, L' > 0$  and  $0 < r \leq 2$ . Thus, from a physical point of view, the choice of the class  $\mathcal{A}(\alpha, r, L, L')$  is quite reasonable.

### 4.3 Estimation procedure and main results

We are now able to define the estimation procedure of the quadratic functional  $d^2 = \int W_\rho^2$  of the unknown function  $W_\rho$  based on data  $(Y_\ell, \phi_\ell)$ . Next we state an upper bound of the maximal risk uniformly over all Wigner functions in the class  $\mathcal{A}(\alpha, r, L, L')$ .

#### 4.3.1 Kernel estimator

Let us define our estimator as a U-statistic of order 2.

**Definition 4.1.** For any  $\delta = \delta_n > 0$ , we define the estimator

$$d_n^2 = \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} K_\delta \left( [z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}} \right) K_\delta \left( [z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}} \right) dz, \quad (4.6)$$

where  $[z, \phi] := u \cos \phi + v \sin \phi$  with  $z = (u, v)$  and

$$K_\delta(x) = \frac{1}{(2\pi)^2} \int_{-1/\delta}^{1/\delta} |t| e^{-itx} e^{\frac{t^2}{4} \frac{1-\eta}{\eta}} dt = \frac{1}{(2\pi)^2} \int_{-1/\delta}^{1/\delta} |t| \cos(tx) e^{\frac{t^2}{4} \frac{1-\eta}{\eta}} dt. \quad (4.7)$$

Note that the Fourier transform of  $K_\delta$  is  $\tilde{K}_\delta(t) = \frac{1}{2\pi} |t| e^{\frac{t^2}{4} \frac{1-\eta}{\eta}} \mathbb{I}(|t| \leq 1/\delta)$ , where  $\mathbb{I}$  stands for the indicator function.

Let  $d_n^2$  be the estimator defined by (4.6), having bandwidth  $\delta > 0$ . We call the bias and the variance of the estimator, respectively:

$$B(d_n^2) := |E_\rho[d_n^2] - d^2| \quad \text{and} \quad \text{Var}(d_n^2) := E_\rho[|d_n^2 - E_\rho[d_n^2]|^2]. \quad (4.8)$$

### 4.3.2 Bias-variance decomposition

The following proposition plays an important role in the proof of the upper bound of the risk as we split it into the bias term and the variance term.

**Proposition 4.1.** *Let  $a = \frac{1-\eta}{2\eta}$  and  $d_n^2$  be the estimator defined by (4.6) with  $\delta \rightarrow 0$  such that  $e^{a/\delta^2}/(n\delta^2) \rightarrow 0$  as  $n \rightarrow \infty$ , then for all  $\alpha > 0$ ,  $L, L' > 0$  and  $0 < r \leq 2$*

$$\sup_{W_\rho \in \mathcal{A}(\alpha, r, L, L')} B^2(d_n^2) \leq L^2 e^{-\frac{4\alpha}{\delta^2}} (1 + o(1)), \quad (4.9)$$

$$\sup_{W_\rho \in \mathcal{A}(\alpha, r, L, L')} \text{Var}(d_n^2) \leq \frac{16}{na^2\delta^2} e^{\frac{a}{\delta^2}} (1 + o(1)). \quad (4.10)$$

The proof of this proposition is given in Section 4.4.

### 4.3.3 Main results

Let  $d_n^2$  be an estimator of  $d^2 = \int W_\rho^2$  defined by (4.6). We measure the accuracy of  $d_n^2$  by the maximal risk over the class  $\mathcal{A}(\alpha, r, L, L')$

$$\mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) = \sup_{W_\rho \in \mathcal{A}(\alpha, r, L, L')} E_\rho[|d_n^2 - d^2|^2].$$

Here  $E_\rho$  and  $P_\rho$  denote the expected value and the probability when the true underlying quantum state is  $\rho$ .

**Theorem 4.1.** *Let  $(Y_\ell, \phi_\ell), \ell = 1, \dots, n$  be i.i.d data coming from the model (4.3) where the underlying parameter is the Wigner function  $W_\rho$  lying in the class  $\mathcal{A}(\alpha, r, L, L')$ , with*

$0 < r < 2$ ,  $\alpha > 0$ ,  $L, L' > 0$ . Let  $a = \frac{1-\eta}{2\eta}$ , then  $d_n^2$  defined in (4.6) with bandwidth  $\delta := \delta_{opt}$  chosen as the solution of the equation

$$\frac{a}{\delta_{opt}^2} + \frac{4\alpha}{\delta_{opt}^r} = \log n - (\log \log n)^2, \quad (4.11)$$

satisfies the following upper bound

$$\limsup_{n \rightarrow \infty} \varphi_n^{-2} \mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) \leq L^2, \quad (4.12)$$

where the rate of convergence is  $\varphi_n^2 = e^{-4\alpha/\delta_{opt}^r}$ .

**Remark 4.2.** The previous theorem gives the upper bound of the risk. It is shown that the rate of convergence is given by the dominating term (bias term) at the selected bandwidth  $\delta := \delta_{opt}$ . Following the proof of the lower bound in Butucea, Guță, and Artiles [20], we can prove that similar lower bound holds in our setting when the Wigner function  $W_\rho$  belongs to the class  $\left\{ W_\rho : \widetilde{W}_\rho \in \mathcal{F}(\alpha, r, L) \right\}$  which is strictly larger than  $\mathcal{A}(\alpha, r, L)$ . Unfortunately, the Wigner functions constructed in Butucea, Guță, and Artiles [20] for proving the lower bound do not belong to the class  $\mathcal{F}(2^r\alpha, r, L')$ .

**Sketch of proof of Theorem 4.1** Using (4.9) and (4.10) with  $0 < r < 2$ , we select the bandwidth  $\delta^*$  as

$$\delta^* = \arg \min_{\delta > 0} \left\{ \frac{C_V}{n\delta^2} e^{\frac{a}{\delta^2}} + C_B e^{-\frac{4\alpha}{\delta^r}} \right\}.$$

By taking derivatives,  $\delta^*$  is a positive real number satisfying

$$\frac{a}{\delta^{*2}} + \frac{4\alpha}{\delta^{*r}} = \log(\delta^{*4-r}) + \log n + \text{const.}$$

We notice that the rate of convergence for the upper bound is given by the bias term of the estimator  $d_n^2$  with  $\delta = \delta^*$  i.e.  $\varphi_n^2 = B(d_n^2)(1 + o(1))$  as  $B(d_n^2) = C(\delta^*)^{r-2} \text{Var}(d_n^2)(1 + o(1))$ . On the other hand, by taking  $\delta := \delta_{opt}$  the unique solution of the equation

$$\frac{a}{\delta_{opt}^2} + \frac{4\alpha}{\delta_{opt}^r} = \log n - (\log \log n)^2,$$

the variance of the estimator  $d_n^2$  with  $\delta = \delta_{opt}$  is still smaller than its bias and its bias is of the same order as the bias of  $d_n^2$  with  $\delta = \delta^*$  (see Lemma 8 in Butucea and Tsybakov [19]). So, when replacing  $\delta^*$  with the slightly modified  $\delta_{opt}$  the upper bound of the minimax risk will remain asymptotically the same.  $\square$

**Remark 4.3.** If we consider the case  $r \in ]0, 2[$ , we can give a more explicit form for the bandwidth verifying (4.11) and thus, for the rate of convergence which is asymptotically equivalent to the bias term. Based on the results in Lacour [64], we make successive approximations starting with

$$\delta_0 := \left( \frac{\log n - (\log \log n)^2}{a} \right)^{-1/2},$$

and for all  $k \geq 1$ , if  $r \in I_k = ]\frac{2(k-1)}{k}, \frac{2k}{k+1}]$ , we get recursively  $\delta_k$  by plugging  $\delta_{k-1}$  into  $\delta_k = (\delta_0^{-2} - \frac{4\alpha}{a}\delta_{k-1}^{-r})^{-1/2}$ . Then by choosing  $\delta_{opt} = \delta_k$  and if  $r \in I_k$ , we obtain the following asymptotic equivalent of the rate of convergence

$$L^2 \exp \left( -4\alpha\delta_0^{-r} + C_1\delta_0^{2-r} - \dots + C_{k-1}\delta_0^{2(k-1)-kr} \right).$$

**Theorem 4.2.** Let  $(Y_\ell, \Phi_\ell), \ell = 1, \dots, n$  be i.i.d data coming from the model (4.3) where the underlying parameter is the Wigner function  $W_\rho$  lying in the class  $\mathcal{A}(\alpha, r, L, L')$ , with  $r = 2, \alpha > 0, L, L' > 0$ . Let  $a = \frac{1-\eta}{2n}$ , then  $d_n^2$  defined in (4.6) with bandwidth  $\delta = \delta^* = \left( \frac{\log n - \log(\log n / (4\alpha + a))}{4\alpha + a} \right)^{-1/2}$  satisfies the following upper bound

$$\limsup_{n \rightarrow \infty} \varphi_n^{-2} \mathcal{R}(d_n^2; \mathcal{A}(\alpha, r, L, L')) \leq C, \quad (4.13)$$

where the rate of convergence is  $\varphi_n^2 = \left( \frac{n}{\log n} \right)^{-\frac{4\alpha}{4\alpha+a}}$ , for some constant  $C > 0$  depending on parameters  $\alpha, r, L, L'$  and  $\eta$ .

**Sketch of proof of Theorem 4.2** For  $r = 2$  and by (4.9) and (4.10), we select the bandwidth  $\delta^*$  as

$$\tilde{\delta} = \arg \min_{\delta > 0} \left\{ \frac{C_V}{n\delta^2} e^{\frac{a}{\delta^2}} + C_B e^{-4\alpha/\delta^2} \right\},$$

by taking derivatives, we notice that  $B(d_n^2) \sim \text{Var}(d_n^2)$  for  $\delta = \tilde{\delta}$  and that the rate of convergence is  $\left( \frac{n}{\log n} \right)^{-\frac{4\alpha}{4\alpha+a}}$ . It is easy to check that if we choose  $\delta^*$  as bandwidth we get the same rate.  $\square$

In the previous theorems, the bandwidth  $\delta_{opt}$  depends on the parameters  $\alpha$  and  $r$  of the class  $\mathcal{A}(\alpha, r, L, L')$  which may be difficult to evaluate in practice. However, it is possible to construct an adaptive estimator which does not depend on these parameters and which has the same asymptotic behavior as in Theorem 4.1, provided that these parameters lie in a certain set. Note that the parameter  $\eta$  is considered to be known. Define the set of parameters

$$\Theta_1 = \{(\alpha, r, L, L') : \alpha > 0, L, L' > 0, 0 < r < 1\}.$$



**Theorem 4.3.** Let  $(Y_\ell, \phi_\ell), \ell = 1, \dots, n$  be i.i.d data coming from the model (4.3). Let  $d_{ad,n}^2$  be the estimator defined by

$$d_{ad,n}^2 = \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta_{ad}} K_{\delta_{ad}} \left( [z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}} \right) K_{\delta_{ad}} \left( [z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}} \right) dz,$$

with  $\delta = \delta_{ad} = \left( \frac{\log n}{a} - 2\sqrt{\frac{\log n}{a}} \right)^{-1/2}$ . Then, for all  $(\alpha, r, L, L') \in \Theta_1$ ,

$$\limsup_{n \rightarrow \infty} \sup_{W_\rho \in \mathcal{A}(\alpha, r, L, L')} E[|d_{\delta,n}^2 - d^2|^2] \varphi_n^{-2} \leq 1,$$

where  $\varphi_n$  is the rate defined in Theorem 4.1.

**Sketch of proof of Theorem 4.3** For  $\delta_{opt} = \delta_{opt}(n)$  defined by equation (4.11) we have

$$\frac{1}{\delta_{opt}(n)} = \left( \frac{\log n}{a} \right)^{1/2} (1 - \epsilon_1(n)),$$

where  $\epsilon_1(n) \rightarrow 0^+$  as  $n \rightarrow \infty$ . Then

$$\exp \left( -\frac{4\alpha}{(\delta_{opt}(n))^r} \right) = \exp \left( -4\alpha \left( \frac{\log n}{a} \right)^{r/2} (1 - \epsilon_1(n))^r \right) \geq \exp \left( -4\alpha \left( \frac{\log n}{a} \right)^{r/2} \right),$$

as  $(1 - \epsilon_1(n))^r \leq 1 - r\epsilon_1(n)$  for  $n$  large. Over the set  $\Theta_1$ , for  $0 < r < 1$

$$\begin{aligned} L^2 \exp \left( -\frac{4\alpha}{(\delta_{ad})^r} \right) &= L^2 \exp \left( -4\alpha \left( \frac{\log n}{a} - 2\sqrt{\frac{\log n}{a}} \right)^{r/2} \right) \\ &= L^2 \exp \left( -4\alpha \left( \frac{\log n}{a} \right)^{r/2} \left( 1 - 2 \left( \frac{\log n}{a} \right)^{-1/2} \right)^{r/2} \right) \\ &\leq L^2 \exp \left( -4\alpha \left( \frac{\log n}{a} \right)^{r/2} + c \left( \frac{\log n}{a} \right)^{r/2-1/2} \right) \\ &= L^2 \exp \left( -4\alpha \left( \frac{\log n}{a} \right)^{r/2} \right) (1 + \epsilon_2(n)) \leq L^2 \exp \left( -\frac{4\alpha}{(\delta_{opt}(n))^r} \right) (1 + \epsilon_2(n)), \end{aligned}$$

where  $c$  is a positive constant and  $\epsilon_2(n) \rightarrow 0^+$  as  $n \rightarrow \infty$ . Thus the upper bound of the bias of  $d_{ad,n}^2$  is not larger than the upper bound of the bias of  $d_n^2$  with  $\delta = \delta_{opt}$ . As  $0 < r/2 < 1/2$  it is easy to remark that for  $n$  large enough  $-\left(\frac{\log n}{a} - 2\sqrt{\frac{\log n}{a}}\right)^{r/2} > -\frac{a}{4\alpha} \sqrt{\frac{\log n}{a}}$  and  $\frac{1}{\delta_{ad}^2} \leq \frac{\log n}{a}$  and thus  $\exp \left( -a\sqrt{\frac{\log n}{a}} \right) \leq \exp \left( -\frac{4\alpha}{(\delta_{ad})^r} \right)$ . Then the dominating term in the variance found

in (4.10)

$$\begin{aligned}
\frac{1}{n\delta_{ad}^2} \exp\left(\frac{a}{\delta_{ad}^2}\right) &\leq \left(\frac{\log n}{a}\right) \exp\left(-2a\sqrt{\frac{\log n}{2a}}\right) \\
&= \left(\frac{\log n}{a}\right) \exp\left(-a\sqrt{\frac{\log n}{2a}}\right) \exp\left(-a\sqrt{\frac{\log n}{2a}}\right) \\
&\leq \left(\frac{\log n}{a}\right) \exp\left(-a\sqrt{\frac{\log n}{2a}}\right) \exp\left(-\frac{4\alpha}{\delta_{ad}^r}\right) \\
&\leq o(1) \exp\left(-\frac{4\alpha}{(\delta_{opt})^r}\right).
\end{aligned}$$

Thus,  $d_{ad,n}^2$  attains the rate  $\varphi_n^2$ . □

## 4.4 Proof of Proposition 1

Most of the proofs make extensive use of the following equations and properties of Wigner functions. A remarkable relation links the Fourier transform of the Wigner function to the Fourier transform of its Radon transform. If we denote

$$\widetilde{W}_\rho(u, v) := \mathcal{F}_2[W_\rho](u, v),$$

then

$$\widetilde{W}_\rho(t \cos \phi, t \sin \phi) := \mathcal{F}_1[p_\rho(\cdot|\phi)](t) = E_\rho[e^{itX} | \Phi = \phi], \quad (4.14)$$

where  $\mathcal{F}_2, \mathcal{F}_1$  denote the Fourier transform w.r.t two, respectively one variables. The Fourier transform w.r.t one variable of the density  $p_\rho^\eta(\cdot|\phi)$  of  $Y$  when  $\Phi = \phi$  is

$$\mathcal{F}_1[p_\rho^\eta(\cdot|\phi)](t) = \mathcal{F}_1\left[\frac{1}{\eta} p_\rho(\cdot|\phi)\right](t) \cdot \widetilde{N}^\eta(t) \quad (4.15)$$

$$= \mathcal{F}_1[p_\rho(\cdot|\phi)](\sqrt{\eta}t) \cdot \widetilde{N}^\eta(t), \quad (4.16)$$

where  $\widetilde{N}^\eta(t) = e^{-\frac{t^2}{4}(1-\eta)}$  denotes the characteristic function of the random variable  $\sqrt{(1-\eta)/2}\xi \sim \mathcal{N}(0; (1-\eta)/2)$ .

### 4.4.1 Proof of (4.9)

As  $(Y_k, \Phi_k)$  and  $(Y_\ell, \Phi_\ell)$  are i.i.d. for all  $k \neq \ell$ , we get

$$\begin{aligned}
E[d_n^2] &= \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} E\left[K_\delta\left([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}\right)\right] E\left[K_\delta\left([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}\right)\right] dz \\
&= \int_{\|z\| \leq 1/\delta} \left| E\left[K_\delta\left([z, \Phi] - \frac{Y}{\sqrt{\eta}}\right)\right] \right|^2 dz.
\end{aligned} \quad (4.17)$$

Moreover

$$\begin{aligned}
E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})] &= \int_{\mathbb{R}} \int_0^\pi K_\delta([z, \phi] - \frac{y}{\sqrt{\eta}}) p_\rho^\eta(y, \phi) dy d\phi \\
&= \int_0^\pi \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}_1[K_\delta * (\sqrt{\eta} p_\rho^\eta(\cdot, \sqrt{\eta}, \phi))](t) e^{-it[z, \phi]} dt d\phi \\
&= \int_0^\pi \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{K}_\delta(t) \mathcal{F}_1[\sqrt{\eta} p_\rho^\eta(\cdot, \sqrt{\eta}, \phi)](t) e^{-it[z, \phi]} dt d\phi.
\end{aligned}$$

Then by using the expressions in (4.14), (4.15), (4.16) and a change of variables  $(t, \phi)$  into  $w$  such that  $(t \cos \phi, t \sin \phi) = w$ , we get

$$\begin{aligned}
&E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})] \\
&= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| e^{\frac{t^2}{4} \frac{1-\eta}{\eta}} \mathbb{I}(|t| \leq 1/\delta) \mathcal{F}_1[p_\rho(\cdot, \phi)](t) \tilde{N}^\eta(t/\sqrt{\eta}) e^{-it[z, \phi]} dt d\phi \\
&= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| \mathbb{I}(|t| \leq 1/\delta) \mathcal{F}_1[p_\rho(\cdot, \phi)](t) e^{-it[z, \phi]} dt d\phi \\
&= \frac{1}{(2\pi)^2} \int_0^\pi \int_{\mathbb{R}} |t| \mathbb{I}(|t| \leq 1/\delta) \tilde{W}_\rho(t \cos \phi, t \sin \phi) e^{-it[z, \phi]} dt d\phi \\
&= \frac{1}{(2\pi)^2} \int \mathbb{I}(\|w\| \leq 1/\delta) \tilde{W}_\rho(w) e^{-i\langle z, w \rangle} dw. \tag{4.18}
\end{aligned}$$

Thus, the Fourier transform of  $E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})]$  as a function of  $z$  is  $\tilde{W}_\rho \cdot \mathbb{I}(\|\cdot\| \leq 1/\delta)$ . We write then

$$E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})] = [W_\rho * G_\delta](z), \tag{4.19}$$

where  $\tilde{G}_\delta(w) = \mathbb{I}(\|w\| \leq 1/\delta)$ . Let us study the bias term by (4.8). By (4.17) and (4.19)

$$\begin{aligned}
&B(d_n^2) \\
&= |d^2 - E[d_n^2]| = \left| \int_{\mathbb{R}^2} W_\rho^2(z) dz - \int_{\|z\| \leq 1/\delta} \left| E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})] \right|^2 dz \right| \\
&= \left| \int_{\mathbb{R}^2} W_\rho^2(z) dz - \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta]^2(z) dz \right| \\
&\leq \left| \int_{\mathbb{R}^2} (W_\rho^2(z) dz - [W_\rho * G_\delta]^2(z)) dz \right| + \left| \int_{\|z\| > 1/\delta} [W_\rho * G_\delta]^2(z) dz \right| \\
&\leq \left| \int_{\mathbb{R}^2} (W_\rho^2(z) - [W_\rho * G_\delta]^2(z)) dz \right| + 2 \left| \int_{\|z\| > 1/\delta} |[W_\rho * G_\delta](z) - W_\rho(z)|^2 dz \right| \\
&\quad + 2 \left| \int_{\|z\| > 1/\delta} |W_\rho(z)|^2 dz \right| \\
&= A_1 + 2A_2 + 2A_3, \tag{4.20}
\end{aligned}$$

where  $A_1$ ,  $2A_2$  and  $2A_3$  are respectively the first, the second and the third term of the previous sum (4.20). By the Plancherel formula and using the smoothness of  $W_\rho$

$$\begin{aligned} A_1 &= \frac{1}{(2\pi)^2} \left| \int_{\mathbb{R}^2} \left( |\widetilde{W}_\rho(w)|^2 - |\widetilde{W}_\rho(w)\mathbb{I}(\|w\| \leq 1/\delta)|^2 \right) dw \right| \\ &= \frac{1}{(2\pi)^2} \int_{\|w\| > 1/\delta} |\widetilde{W}_\rho(w)|^2 dw \leq Le^{-\frac{2\alpha}{\delta^r}}. \end{aligned} \quad (4.21)$$

Moreover, as

$$\begin{aligned} \int_{\mathbb{R}^2} |[W_\rho * G_\delta](z) - W_\rho(z)|^2 dz &= \frac{1}{(2\pi)^2} \int_{\|w\| > 1/\delta} |\widetilde{W}_\rho(w)|^2 dw \leq \|W_\rho\|^2 \\ &= \frac{\text{Tr}(\rho^2)}{2\pi} \leq \frac{1}{2\pi}. \end{aligned}$$

This implies as  $\delta \rightarrow 0$ ,  $n \rightarrow \infty$

$$\begin{aligned} A_2 &= o(1) \int_{\mathbb{R}^2} |[W_\rho * G_\delta](z) - W_\rho(z)|^2 dz \\ &= o(1) \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \left| \widetilde{W}_\rho(w)\mathbb{I}(\|w\| \leq 1/\delta) - \widetilde{W}_\rho(w) \right|^2 dw \\ &= o(1) \frac{1}{(2\pi)^2} \int_{\|w\| > 1/\delta} |\widetilde{W}_\rho(w)|^2 dw = o(1)Le^{-\frac{2\alpha}{\delta^r}}. \end{aligned} \quad (4.22)$$

Using now the asymptotic behaviour of  $W_\rho$

$$A_3 = \int_{\|z\| > 1/\delta} |W_\rho(z)|^2 dz \leq (2\pi)^2 Le^{-\frac{2\alpha 2^r}{\delta^r}} = o(1)Le^{-\frac{2\alpha}{\delta^r}}, \quad (4.23)$$

as  $\delta \rightarrow 0$ ,  $n \rightarrow \infty$  and  $2^r > 1 \forall r > 0$ . Then, by using (4.21), (4.22) and (4.23)

$$B(d_n^2) \leq Le^{-2\alpha/\delta^r} (1 + o(1)), \text{ as } \delta \rightarrow \infty.$$

□

### 4.4.2 Proof of (4.10)

We recall that  $E[K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}})] = [W_\rho * G_\delta](z)$ , then

$$\begin{aligned}
& n(n-1)(d_n^2 - E[d_n^2]) \\
&= \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} \left\{ K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta]^2(z) \right\} dz \\
&= \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} \left[ \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \\
&\quad \times \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \\
&\quad \left. + [W_\rho * G_\delta](z) \left\{ \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \left. \left. + \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right\} \right] dz.
\end{aligned}$$

Then,  $d_n^2 - E[d_n^2] = J_1 + J_2$  where we denote by  $J_1, J_2$  respectively

$$\begin{aligned}
J_1 &= \frac{1}{n(n-1)} \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \\
&\quad \times \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz,
\end{aligned}$$

and

$$J_2 = \frac{2}{n} \sum_{\ell} \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta](z) \left( K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz.$$

Then

$$\text{Var}(d_n^2) = E[(d_n^2 - E[d_n^2])^2] = E[J_1^2] + E[J_2^2] + 2E[J_1 J_2]. \quad (4.24)$$

For the third part of the previous sum we have:

$$\begin{aligned}
& n^2(n-1)E[J_1 J_2] \\
&= 2 \sum_{k \neq j} \sum_{\ell} E \left[ \left\{ \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \times \left. \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right\} \\
&\quad \times \left. \left\{ \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta](z) \left( K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right\} \right] \\
&= 0,
\end{aligned}$$

since  $E[K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * G_\delta](z)] = 0$  for all  $\ell = 1, \dots, n$  and because we always have either  $\ell \neq k$  and  $K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}), K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}})$  are independent or  $\ell \neq j$  and  $K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}), K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}})$  are independent. Now study

$$\begin{aligned}
& (n(n-1))^2 E[J_1^2] \\
&= E \left[ \left( \sum_{j \neq k} \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \left. \left. \times \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right)^2 \right] \\
&= \sum_{j_1 \neq k_1} \sum_{j_2 \neq k_2} E \left[ \left\{ \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_{j_1}] - \frac{Y_{j_1}}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \left. \left. \times \left( K_\delta([z, \Phi_{k_1}] - \frac{Y_{k_1}}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right\} \right. \\
&\quad \left. \times \left\{ \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_{j_2}] - \frac{Y_{j_2}}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \left. \left. \times \left( K_\delta([z, \Phi_{k_2}] - \frac{Y_{k_2}}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right\} \right].
\end{aligned}$$

Note that, as soon as either  $j_1$  is different from  $k_2$  and  $j_2$ , or  $k_1$  is different from  $k_2$  and  $j_2$  the expected value is 0. Thus,

$$\begin{aligned}
& (n(n-1))^2 E[J_1^2] \\
&= \sum_{j \neq k} E \left[ \left( \int_{\|z\| \leq 1/\delta} \left( K_\delta([z, \Phi_j] - \frac{Y_j}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right. \right. \\
&\quad \left. \left. \times \left( K_\delta([z, \Phi_k] - \frac{Y_k}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right)^2 \right].
\end{aligned}$$

By using the Cauchy-Schwarz inequality, as  $(Y_k, \Phi_k)$  and  $(Y_j, \Phi_j)$  are i.i.d. and by the definition (4.7) of  $K_\delta$ , we get

$$\begin{aligned}
& E[J_1^2] \\
& \leq \frac{2}{n^2} \left( \int_{\|z\| \leq 1/\delta} E \left[ \left| K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right|^2 \right] dz \right)^2 \\
& \leq \frac{2}{n^2} \left( \int_{\|z\| \leq 1/\delta} E \left[ \left| K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}}) \right|^2 \right] dz \right)^2 \\
& \leq \frac{2}{n^2} \left( \int_{\|z\| \leq 1/\delta} \|K_\delta\|_\infty^2 dz \right)^2 \leq \frac{2}{n^2} \left( \int_{\|z\| \leq 1/\delta} \left( \int |\tilde{K}_\delta(t) dt| \right)^2 dz \right)^2 \\
& = \frac{2}{n^2} \left( \int_{\|z\| \leq 1/\delta} \left( \int_{|t| \leq 1/\delta} |t| e^{t^2 \frac{1-\eta}{4\eta}} dt \right)^2 dz \right)^2 \\
& \leq \frac{1}{(2\pi)^2 n^2 \delta^4} \left( \frac{4\eta}{1-\eta} \right)^4 e^{\frac{1-\eta}{\eta \delta^2}}. \tag{4.25}
\end{aligned}$$

The term  $E[J_2^2]$  can be bounded as follows

$$\begin{aligned}
& E[J_2^2] \\
& = E \left[ \left( \frac{2}{n} \sum_\ell \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta](z) \left( K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) \right)^2 \right] \\
& = \frac{4}{n^2} \sum_\ell E \left[ \left( \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta](z) \left( K_\delta([z, \Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}) - [W_\rho * G_\delta](z) \right) dz \right)^2 \right] \\
& \leq \frac{4}{n} E \left[ \left( \int_{\|z\| \leq 1/\delta} [W_\rho * G_\delta](z) K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}}) dz \right)^2 \right].
\end{aligned}$$

By Cauchy-Schwarz inequality

$$\begin{aligned}
& E[J_2^2] \\
& \leq \frac{4}{n} \int_{\|z\| \leq 1/\delta} \left| E \left[ K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}}) \right] \right|^2 dz \int_{\|z\| \leq 1/\delta} E \left[ \left| K_\delta([z, \Phi] - \frac{Y}{\sqrt{\eta}}) \right|^2 \right] dz \\
& \leq \frac{4}{n} \int_{\|z\| \leq 1/\delta} |[W_\rho * G_\delta](z)|^2 dz \int_{\|z\| \leq 1/\delta} \left( \int_{|t| \leq 1/\delta} |t| e^{t^2 \frac{1-\eta}{4\eta}} dt \right)^2 dz \\
& \leq \frac{8\pi}{n\delta^2} \left( \frac{4\eta}{1-\eta} \right)^2 e^{\frac{1-\eta}{2\eta\delta^2}} (1 + o(1)) \int_{\mathbb{R}^2} |[W_\rho * G_\delta](z)|^2 dz.
\end{aligned}$$

Parseval's equality and using the fact that  $2\pi \|W_\rho\|^2 = \text{Tr}(\rho^2) \leq 1$

$$\begin{aligned}
E[J_2^2] &\leq \frac{8\pi}{n\delta^2} \left( \frac{4\eta}{1-\eta} \right)^2 e^{\frac{1-\eta}{2\eta\delta^2}} (1+o(1)) \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathbb{I}(\|w\| \leq 1/\delta) \left| \widetilde{W}_\rho(w) \right|^2 dw \\
&\leq \frac{8\pi}{n\delta^2} \left( \frac{4\eta}{1-\eta} \right)^2 e^{\frac{1-\eta}{2\eta\delta^2}} (1+o(1)) \|W_\rho\|^2 \\
&\leq \frac{4}{n\delta^2} \left( \frac{4\eta}{1-\eta} \right)^2 e^{\frac{1-\eta}{2\eta\delta^2}} (1+o(1)). \tag{4.26}
\end{aligned}$$

In view of (4.25), (4.26) as one chooses  $\delta$  such that  $n\delta^{-2}e^{\frac{1-\eta}{2\eta\delta^2}}$  tends to 0, then  $E[J_1^2] = o(1)E[J_2^2]$  and we conclude by using (4.24).  $\square$







## Chapter 5

# Nonparametric goodness-of fit testing in quantum homodyne tomography with noisy data

*Abstract:* In the framework of quantum optics, we study the problem of goodness-of-fit testing in a severely ill-posed inverse problem. A novel testing procedure is introduced and its rates of convergence are investigated under various smoothness assumptions. The procedure is derived from a projection-type estimator, where the projection is done in  $\mathbb{L}_2$  distance on some suitably chosen *pattern* functions. The proposed methodology is illustrated with simulated data sets.

In quantum mechanics, the results of physical measurements performed on a physical system are random. This fact provides a large variety of problems and mathematical and applied statistics offer many methods to answer them. For example, statisticians are asked to decide which measurement to perform in order to collect most informative observations or to rebuild the subjacent quantum object from the resulting observations. In particular, physicists have created laser in view of many applications (medical,...). An important question is to know whether they have generated the desired one. In order to answer this question, we propose a testing procedure based on the indirect observations they gathered by a measurement called quantum homodyne tomography (QHT). To describe our framework and to motivate our work, we present in Section 5.1 the needed physical background. In Section 5.2, we introduce our statistical model and define the nonparametric class containing the density matrices we are dealing with. To make clear to physicists the statistical part,

we recall some basic notions on the statistical tests in Section 5.2.2. The testing procedure and the theoretical results assessing its asymptotic behavior are presented in Section 5.3. Section 5.4 contains an illustration of the proposed methodology on simulated data sets while the proof are postponed to Section 5.5.

## 5.1 Physical background

Section 5.1.1 is a short introduction to quantum optics and we introduce in Section 5.1.2 the very useful and meaningful *pattern* functions. In Section 5.1.3, we present some examples of quantum states actually created in laboratory, on which our testing procedure may be applied.

### 5.1.1 Short introduction to quantum optics

In quantum mechanics, the quantum state fully describes all aspects of a physical system. Mathematically, the quantum state is described via **density operators**  $\rho$  on a complex Hilbert space  $\mathcal{H}$  such that

1.  $\rho$  is self-adjoint (or Hermitian):  $\rho = \rho^*$ , where  $\rho^*$  is the adjoint of  $\rho$ ,
2.  $\rho$  is positive:  $\rho \geq 0$ , or equivalently  $\langle \psi, \rho \psi \rangle \geq 0$  for all  $\psi \in \mathcal{H}$ ,
3.  $\rho$  has trace one:  $\text{Tr}(\rho) = 1$ .

Moreover, to each measurable physical property or quantity corresponds a self-adjoint operator, say  $\mathbf{X}$ , on the space of states  $\mathcal{H}$ . This operator is called an *observable*. Unlike in classical mechanics, when performing a certain measurement of an observable  $\mathbf{X}$  of a physical system prepared in a quantum state  $\rho$ , the result is in general random and is described by a probability distribution of a random variable  $X$ .

In this chapter, the studied quantum system is a monochromatic light in a cavity described by the state of a quantum harmonic oscillator. Hereafter, we consider the complex Hilbert space  $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ , the space of square integrable complex-valued functions on the real line. In quantum optics, the quantum harmonic oscillator can be represented by two operators on  $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ : the position operator  $\mathbf{Q}$  (or electric field) and the quantity of movement

operator  $\mathbf{P}$  (or magnetic field)

$$(\mathbf{Q}\psi_1)(x) = x\psi_1(x), \quad (5.1)$$

$$(\mathbf{P}\psi_2)(x) = -i\frac{\partial\psi_2(x)}{\partial x}, \quad (5.2)$$

with  $\psi_1, \psi_2$  functions of  $\mathcal{H}$ . The position  $\mathbf{Q}$  and momentum  $\mathbf{P}$  do not commute and satisfy the canonical commutation relation:

$$[\mathbf{Q}, \mathbf{P}] = \mathbf{Q}\mathbf{P} - \mathbf{P}\mathbf{Q} = i\mathbf{1},$$

where  $\mathbf{1}$  stands for the identity operator on  $\mathcal{H}$ . This relation may be understood as follows. The measurement of the position  $\mathbf{Q}$  necessarily disturbs the particle's momentum  $\mathbf{P}$ , and vice versa. They cannot be measured simultaneously. Thus, one cannot obtain a couple of random variables  $(Q, P)$ . However, a linear combination

$$\mathbf{X}_\phi = \cos(\phi)\mathbf{Q} + \sin(\phi)\mathbf{P}, \text{ for every phases } \phi \in [0, \pi],$$

can theoretically be measured by a technique put in practice for the first time in [82] and called quantum homodyne tomography (QHT<sup>1</sup>), where the phase  $\Phi$  is randomly and uniformly distributed on  $[0, \pi]$ .

In the *case of noiseless* measurement and for a phase  $\Phi = \phi$ , the result would be a random variable  $X$  with probability density noted  $p_\rho(\cdot|\phi)$ . Moreover, the characteristic function of  $X$  is

$$\mathcal{F}_1[p_\rho(\cdot|\phi)](t) = \text{Tr}(\rho e^{it\mathbf{X}_\phi}) = \text{Tr}(\rho e^{it(\cos(\phi)\mathbf{Q} + \sin(\phi)\mathbf{P})}).$$

Furthermore, to each state  $\rho$  corresponds a **Wigner function**  $W_\rho$ , which gives an equivalent representation of the quantum state  $\rho$ . The associated Wigner function  $W_\rho$  can be defined rigorously by its Fourier transform  $\mathcal{F}_2[W_\rho]$  with respect to both variables

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

The Wigner function is a mapping from  $\mathbb{R}^2$  to  $\mathbb{R}$  such that  $\iint W_\rho(q, p)dqdp = 1$ . Note that the Wigner function can take negative values. For this reason,  $W_\rho$  is a real-valued function regarded as a generalized joint probability density (quasi-probability density) of the two random variables  $Q$  and  $P$  that we would get if we could measure simultaneously

---

<sup>1</sup>We refer the interested reader to the book [65] or the paper [43] for further details on QHT.

the two observables  $\mathbf{Q}$  and  $\mathbf{P}$ , which are respectively given by equations (5.1) and (5.2). It is well-known that its Radon transform  $\mathfrak{R}[W_\rho]$  is such that

$$p_\rho(x|\phi) = \mathfrak{R}[W_\rho](x, \phi). \quad (5.3)$$

Here,  $\mathfrak{R}$  denotes the Radon transform, taking functions  $W_\rho(q, p)$  on  $\mathbb{R}^2$  into functions  $\mathfrak{R}[W_\rho](x, \phi)$  on  $\mathbb{R} \times [0, \pi]$  formed by integration along lines of direction  $\phi$  and distance  $x$  from the origin, expressed as

$$\mathfrak{R}[W_\rho](x, \phi) = \int_{-\infty}^{\infty} W_\rho(x \cos \phi - t \sin \phi, x \sin \phi + t \cos \phi) dt.$$

As  $\Phi$  is chosen uniformly and independently on  $[0, \pi]$ , we can define the probability density function  $p_\rho(x, \phi)$  of  $(X, \Phi)$  w.r.t. the measure  $\frac{1}{\pi}\lambda$ , where  $\lambda$  stands for the Lebesgue measure on  $\mathbb{R} \times [0, \pi]$  by

$$p_\rho(x, \phi) = \frac{1}{\pi} \mathfrak{R}[W_\rho](x, \phi). \quad (5.4)$$

### 5.1.2 Fock basis and *pattern* functions

An important equation in physics, especially in quantum mechanics, is the Schrödinger equation. The time-independent Schrödinger equation is written as

$$\left[ -\frac{1}{2} \frac{\partial^2}{\partial x^2} + \frac{x^2}{2} \right] \psi = w\psi, \quad w \in \mathbb{R}, \quad (5.5)$$

where  $w$  is the energy level. It turns out that the energies are "quantized", and may only take discrete values noted  $w_k = k + 1/2$ ,  $k \in \mathbb{N}$ . For a given frequency  $w_k$ , there are two fundamental solutions:  $\psi_k$  and  $\varphi_k$ . One is a normalized function, the function  $\psi_k$ , which is such that  $\int \psi_k^2 = 1$  and is called the regular wave function. The other one, the function  $\varphi_k$ , is called the irregular one as it cannot be normalizable as  $\psi_k$  is.

The functions  $\{\psi_k\}_{k \in \mathbb{N}}$  form an orthonormal basis of  $\mathbb{L}_2(\mathbb{R})$ . This particular basis, physically very meaningful, is called the **Fock** basis and is written as follows:

$$\psi_k(x) := \frac{1}{\sqrt{\sqrt{\pi} 2^k k!}} H_k(x) e^{-x^2/2}. \quad (5.6)$$

Here,  $H_k(x) := (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$  denotes the Hermite polynomial of degree  $k$ .

The density operator defined in the previous section corresponds to a **density matrix** under some orthonormal basis. In the Fock basis, the matrix entries  $\rho_{j,k}$  of the state  $\rho$

can be expressed as expected values of functions  $F_{j,k}(X_\ell, \Phi_\ell) = f_{j,k}(X_\ell)e^{-i(k-j)\Phi_\ell}$ , where  $f_{j,k} = f_{k,j}$  are bounded real functions called *pattern functions* [65]. For all  $j, k \in \mathbb{N}$ ,

$$\rho_{j,k} = \iint_0^\pi p_\rho(x, \phi) f_{j,k}(x) e^{-i(k-j)\phi} d\phi dx,$$

where  $p_\rho(x, \phi)$  is the joint probability density of  $(X, \Phi)$ . In other words

$$\rho_{j,k} = E_\rho [F_{j,k}(X, \Phi)]. \quad (5.7)$$

Equation (5.7) expresses the idea that one can reconstruct any density matrix element  $\rho_{j,k}$  using the *pattern functions*. First introduced in [66], the *pattern functions*  $f_{j,k}$  are well known in physics and are defined in [67] as the first derivatives of products of the two fundamental solutions  $\psi_k$  and  $\varphi_k$  of the Schrödinger equation given in (5.5) for  $j \geq k$

$$f_{j,k}(x) = \frac{\partial}{\partial x} (\psi_j(x)\varphi_k(x)).$$

A concrete expression for their Fourier transform using generalized Laguerre polynomials can be found in [79, 6]. For  $j \geq k$

$$\tilde{f}_{j,k}(t) = \pi(-i)^{j-k} \sqrt{\frac{2^{k-j}k!}{j!}} |t| t^{j-k} e^{-\frac{t^2}{4}} L_k^{j-k}\left(\frac{t^2}{2}\right), \quad (5.8)$$

where  $\tilde{f}_{j,k}$  denotes the Fourier transform of the *pattern function*  $f_{j,k}$  and  $L_k^\alpha(x)$  denotes the generalized Laguerre polynomial. We note that the *pattern functions*  $f_{j,k}(x)$  are even functions for even differences  $j - k$  and odd functions for odd ones

$$f_{j,k}(-x) = (-1)^{j-k} f_{j,k}(x).$$

### 5.1.3 Examples of quantum states

We present in Table 5.1 examples of pure quantum states, which can be created at this moment in laboratory and belong to the class  $\mathcal{R}(B, r, L)$  with  $r = 2$ . 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. This is equivalent to the density matrix being a one dimensional projector, i.e., of the form  $\rho = \mathbf{P}_\psi$  for some unit vector  $\psi$ . Equivalently, a state  $\rho$  is pure if  $\text{Tr}(\rho^2) = 1$ . All other states are called mixed states.

Let us discuss these few examples of quantum states. Among the pure states we consider the *single photon* state and 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

Table 5.1: Examples of quantum states

<b>Vacuum state</b>
<ul style="list-style-type: none"> <li>• <math>\rho_{0,0} = 1</math> rest zero,</li> <li>• <math>p_\rho(x \phi) = e^{-x^2}/\sqrt{\pi}</math>.</li> </ul>
<b>Single photon state</b>
<ul style="list-style-type: none"> <li>• <math>\rho_{1,1} = 1</math> rest zero,</li> <li>• <math>p_\rho(x \phi) = 2x^2 e^{-x^2}/\sqrt{\pi}</math>.</li> </ul>
<b>Coherent-<math>q_0</math> state</b> $q_0 \in \mathbb{R}$
<ul style="list-style-type: none"> <li>• <math>\rho_{j,k} = e^{- q_0/\sqrt{2} ^2} (q_0/\sqrt{2})^{j+k} / \sqrt{j!k!}</math>,</li> <li>• <math>p_\rho(x \phi) = \exp(-(x - q_0 \cos(\phi))^2) / \sqrt{\pi}</math>.</li> </ul>
<b>Squeezed state</b> $(M, \delta) \in \mathbb{R}_+^2, \xi \in \mathbb{R}$
<ul style="list-style-type: none"> <li>• <math>\rho_{j,k} = C(M, \xi) (\frac{1}{2} \tanh(\xi))^{k+j} H_j(\delta) H_k(\delta) / \sqrt{j!k!}</math>,</li> <li>• <math>p_\rho(x \phi) = \exp \left[ \frac{(\sin^2(\phi)(x e^{-2\xi} \cos(\phi) - (x \cos(\phi) - \alpha) e^{2\xi})^2)}{(e^{2\xi} \sin^2(\phi) + e^{-2\xi} \cos^2(\phi))} \right. \\ \left. \times -e^{-2\xi} (x \cos(\phi) - \alpha)^2 - e^{2\xi} x^2 \sin^2(\phi) \right] / \sqrt{\pi(e^{2\xi} \sin^2(\phi) + e^{-2\xi} \cos^2(\phi))}</math>.</li> </ul>
<b>Thermal state</b> $\beta > 0$
<ul style="list-style-type: none"> <li>• <math>\rho_{j,k} = (1 - e^{-\beta}) e^{-\beta k} \mathbb{1}_{j=k}</math>,</li> <li>• <math>p_\rho(x \phi) = \sqrt{\tanh(\beta/2)/\pi} \exp(-x^2 \tanh(\beta/2))</math>.</li> </ul>
<b>Schrödinger cat-<math>q_0</math></b> $q_0 > 0$
<ul style="list-style-type: none"> <li>• <math>\rho_{j,k} = 2(q_0/\sqrt{2})^{j+k} / (\sqrt{j!k!}(\exp(q_0^2/2) + \exp(-q_0^2/2)))</math>, for <math>j</math> and <math>k</math> even, rest zero,</li> <li>• <math>p_\rho(x \phi) = (\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))) / (2\sqrt{\pi}(1 + \exp(-q_0^2)))</math>.</li> </ul>

$p_\rho(x|\phi)$  via the ideal measurement of quantum homodyne tomography. 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. The *Squeezed* states (see e.g. [15]) have Gaussian Wigner functions whose variances in the two directions have a fixed product. The parameters  $M$  and  $\xi$  are such that  $M \geq \sinh^2(\xi)$ ,  $C(M, \xi)$  is a normalization constant,  $\alpha = ((M - \sinh^2(\xi))^{1/2})/(\cosh(\xi) - \sinh(\xi))$ , and  $\delta = (\alpha/(\sinh(2\xi)))^{1/2}$ . The *Schrödinger cat* state is described by a linear superposition of two *coherent* vectors (see e.g. [75]). Table 5.1 gives some explicit density matrix coefficients  $\rho_{j,k}$  and probability densities  $p_\rho(x|\phi)$ .

## 5.2 Problem formulation

In this chapter, the studied quantum system is a monochromatic light in a cavity, whose state is described by an infinite **density matrix**  $\rho$  on the Hilbert space  $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ . In this setting, a convenient representation of a quantum state can be obtained by the projection onto the orthonormal Fock basis  $(\psi_k)_{k \in \mathbb{N}}$  defined in (5.6). In quantum optics, physicists produce quantum state of light and via QHT measurements, they gather independent identically



distributed random variables containing information on the unknown, underlying quantum state  $\rho$ . In an ideal framework, the results of measurements would be  $(X_\ell, \Phi_\ell)_{\ell=1, \dots, n}$ , independent identically distributed random variables with values in  $\mathbb{R} \times [0, \pi]$ , with  $p_\rho(x, \phi)$  the probability density function of  $(X_1, \Phi_1)$  defined in (5.4).

### 5.2.1 Statistical model

In this chapter we consider a more realistic model in presence of an additional independent Gaussian noise. In practice from  $n$  identical, independent **QHT** measurements, we do not collect data  $(X_\ell, \Phi_\ell)$ , but we observe  $(Y_\ell, \Phi_\ell)_{\ell=1, \dots, n}$  independent identically distributed random variables, as error is add such that

$$Y_\ell := \sqrt{\eta}X_\ell + \sqrt{(1-\eta)/2} \xi_\ell, \quad (5.9)$$

where  $\xi_\ell$  is a sequence of independent identically distributed standard Gaussians, independent of all  $(X_\ell, \Phi_\ell)$ . The detection efficiency parameter  $\eta$ ,  $0 < \eta \leq 1$ , is a known parameter and  $1 - \eta$  represents the proportion of lost photons due to various losses in the measurement process.

First recall that for any functions  $f, g : \mathbb{R} \rightarrow \mathbb{R}$ , we denote by  $f * g$  the convolution product

$$f * g(y) = \int f(y-t)g(t)dt.$$

Due to (5.9), the density  $p_\rho^\eta$  of  $(Y_\ell, \Phi_\ell)$  is given by the convolution of the density  $p_\rho(\cdot/\sqrt{\eta}, \phi)/\sqrt{\eta}$  with  $N^\eta$  the centered Gaussian density of variance  $(1-\eta)/2$ . In other terms

$$p_\rho^\eta(y, \phi) = \left( \frac{1}{\sqrt{\eta}} p_\rho \left( \frac{\cdot}{\sqrt{\eta}}, \phi \right) * N^\eta \right) (y), \quad \forall y \in \mathbb{R}, \phi \in [0, \pi].$$

This corresponds to a severely ill-posed inverse problem as the additive noise is super-smooth Gaussian. In the Fourier domain, this relation becomes

$$\mathcal{F}_1[p_\rho^\eta(\cdot, \phi)](t) = \mathcal{F}_1[p_\rho(\cdot, \phi)](t\sqrt{\eta})\tilde{N}^\eta(t), \quad (5.10)$$

where  $\mathcal{F}_1[p_\rho^\eta(\cdot, \phi)]$  denotes the Fourier transform with respect to the first variable and  $\tilde{N}^\eta$  the Fourier transform of the Gaussian density  $N^\eta$ .

We suppose that the unknown state belongs to a natural class of states from the point of view of applications, the class  $\mathcal{R}(B, r, L)$  for  $B > 0$  and  $r \in ]0, 2]$ , defined by

$$\mathcal{R}(B, r, L) := \{\rho \text{ quantum state} : |\rho_{j,k}| \leq L \exp(-B(j+k)^{r/2})\}. \quad (5.11)$$

Let us note that all the states described in Table 5.1 belong to this class with  $r = 2$ .

### 5.2.2 Statistical tests

An important problem in quantum optics is to check whether the produced light pulse is in the desired known quantum state  $\tau$  or not. The purpose of this chapter is to answer this question via goodness-of-fit testing.

More precisely, we consider here the problem of nonparametric goodness-of-fit testing from the data  $(Y_\ell, \Phi_\ell)$  for  $\ell = 1, \dots, n$ ; i.e. given  $\tau \in \mathcal{R}(B, r, L)$ , we define the null hypothesis  $H_0$  and the alternative hypothesis  $H_1$  as follows:

$$\begin{cases} H_0 : & \rho = \tau, \\ H_1(\mathcal{C}, \varphi_n) : & \rho \in \mathcal{R}(B, r, L) \text{ s.t. that } \|\rho - \tau\|_2 \geq \mathcal{C} \cdot \varphi_n, \end{cases}$$

where  $\varphi_n$  is a sequence, which tends to 0 when  $n \rightarrow \infty$ . The physical interpretation of such a test is to check whether the produced light pulse is in a known quantum state  $\tau$  ( $H_0$  is accepted), or not ( $H_1$  is accepted). Here, the distance between the unknown state  $\rho$  and the presumed state  $\tau$  is measured by the squared- $\mathbb{L}_2$ -distance:

$$\|\rho - \tau\|_2^2 = \sum_{j,k \geq 0} |\rho_{j,k} - \tau_{j,k}|^2. \quad (5.12)$$

In nonparametric statistics, different tools are developed to evaluate the accuracy of a testing procedure. First, let us begin by reminding of some basic definitions. There are two important errors made in a statistical decision process:

- First-type error: the test will reject a correct null hypothesis,
- Second-type error: the test will accept a false null hypothesis.

Given a test procedure  $\Omega_n$  such that  $\Omega_n = 0$  when we accept  $H_0$  and  $\Omega_n = 1$  when we reject  $H_0$  and decide  $H_1$ , we denote by  $\alpha = P_\tau[\Omega_n = 1]$  and  $\beta = \beta(\rho) = P_\rho[\Omega_n = 0]$  the probabilities to make a first-type error and a second-type error respectively under  $\tau$  defined in  $H_0$  and  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ . We would like to control the sum of these two probabilities and we do it as described in Definition 5.1.

**Definition 5.1.** *For a given  $0 < \lambda < 1$ , a test procedure  $\Omega_n$  satisfies the upper bound (5.13) for the testing rate  $\varphi_n$  over the smoothness class  $\mathcal{R}(B, r, L)$  if there exists a constant  $\mathcal{C}^* > 0$  such that for all  $\mathcal{C} > \mathcal{C}^*$ :*

$$\limsup_{n \rightarrow \infty} \left\{ P_\tau[\Omega_n = 1] + \sup_{\rho \in H_1(\mathcal{C}, \varphi_n)} P_\rho[\Omega_n = 0] \right\} \leq \lambda, \quad (5.13)$$

where  $P_\tau$  denote the probability under  $\rho = \tau$  defined in  $H_0$ .

In the statistics, we may also compare the power of several testing procedures having first-type error less than some fixed  $\alpha$  and choose the most powerful procedure. The *power* of a statistical test is the probability that the test will reject a false null hypothesis (that it will not make a second-type error) and we denote it by  $\Pi$  such that

$$\Pi = 1 - \beta = P_\rho[\Omega_n = 1], \text{ under } \rho \text{ satisfying } H_1(\mathcal{C}, \varphi_n). \quad (5.14)$$

Remark that, the probability of a second-type error decreases as the *power* increases. In Section 5.4, to evaluate the performance of our testing procedure  $\Omega_n$ , we estimate empirically the *power* of our test.

### 5.2.3 Outline of results

The problem of reconstructing the quantum state of a light beam has been extensively studied in quantum statistics and physical literature. Methods for reconstructing a quantum state are based either on density matrix or on Wigner function estimation. The estimation of the density matrix from averages of data has been considered in the framework of ideal detection [29, 30, 66, 4] as well as in the more general case of an efficiency parameter  $\eta$  belonging to the interval  $]1/2, 1]$  (cf. [26, 33, 27, 78]). The case  $\eta \in ]0, 1]$  has been recently treated in [6]. The latter paper provides also rates of convergence in  $\mathbb{L}_2$  loss for an estimator of the Wigner function. The problem of pointwise estimation of the Wigner function has been previously studied in [43] for ideal data and in [20] for noisy data. It should be noted that the results of [43] and most part of the results in [20] are asymptotically minimax not only in the rate, but also in the constant.

In the present work, the goodness-of-fit problem in quantum statistics is considered. There is a large literature on nonparametric testing procedures for the goodness-of-fit of probability distributions. First of all, let us mention the family of test procedures built on certain distances between empirical cumulative distribution functions (c.d.f.), such as the Kolmogorov-Smirnov and the Cramer-Von Mises test statistics, for which extensive results in terms of asymptotic efficiency were established [73]. In order to be more sensitive to low-frequency components or narrow bumps for powerful discrimination, test procedures based on distances between densities were proposed, such as the Bickel-Rosenblatt test [13, 2] for the  $\mathbb{L}_2$ -distance and the test of [1] for the  $\mathbb{L}_1$ -distance. Theoretical results on such test statistics naturally stems from their nonparametric function estimation counterparts.

In order to compare nonparametric testing procedures, many approaches were proposed, as reviewed in [73, 52]. A common approach is to analyze the power against sequences of local

alternatives  $\{f_n\}_{n \geq 1}$  of the form  $f_n = f_0 + \varphi_n g$  where  $g$  is the direction defining the sequence of local alternatives, and where  $\varphi_n \rightarrow 0$  as  $n$  tends to infinity. Typically, while nonparametric test statistics achieve nontrivial power (i.e. the power of the test is strictly larger than the first type error) against directional alternatives  $f_n = f_0 + \varphi_n g$  when  $\varphi_n = Cn^{-1/2}$ , they achieve nontrivial power against nondirectional alternatives  $f_n = f_0 + \varphi_n g_n$  only for  $\varphi_n$  slower than  $n^{-1/2}$ . In other words, achieving nontrivial power uniformly against a large class of alternatives comes at the price of a slower rate than for parametric testing. The minimax distinguishability framework described in a non-asymptotic setting in [10, 37] and in an asymptotic setting in [52], allows to give precise statements about this phenomenon by characterizing the discrimination rate  $\varphi_n$  depending on a smoothness index of the class of alternatives one wish to discriminate from the null hypothesis. Sharp minimax results with pointwise and sup-norm distances have been established in [69] for the regression model and in [77] for the Gaussian white noise model for supersmooth functions. In [16], goodness-of-fit testing in the convolution model have been considered and minimax rates for testing in  $\mathbb{L}_2$ -norm from indirect observations have been established. The first testing procedure adaptive to smoothness of the alternative function was proposed by [83].

In quantum statistics framework, for the problem of discriminating between two different and fixed states we mention [74] among others. They have establish lower bound for the Bayesian error probability.

The remainder of the article is organized as follows. In Section 5.3, we present our testing procedure and our theoretical results. We study in Section 5.4 the numerical performance of our testing procedure. The proofs are deferred to Section 5.5.

## 5.3 Testing procedure and results

Our testing method is based on a projection method on pattern functions. We describe first the pattern functions adapted to our setting and derive some usefull properties. We define and study the test statistics and give the main results concerning the testing procedure.

### 5.3.1 Pattern functions

To take into account the detection losses described by the overall efficiency  $\eta$ , it is necessary to adapt the *pattern* functions. When  $\eta \in ]1/2, 1]$ , we denote by  $f_{j,k}^\eta$  the suitable functions introduced in [6] and defined by their Fourier transform as follows:

$$\tilde{f}_{j,k}^\eta(t) := \tilde{f}_{j,k}(t) e^{\frac{1-\eta}{4\eta} t^2}. \quad (5.15)$$

In this chapter we develop a procedure for  $\eta \in ]1/2, 1]$ , but it is possible to get quite similar results with the same procedure for  $0 < \eta \leq \frac{1}{2}$  by using modified *pattern* functions, those introduced in [6]. We restrict our study to the more interesting case  $\eta \in ]1/2, 1]$  as in practice  $\eta$  is around 0.9. The following lemma provides useful upper bounds on the  $\mathbb{L}_2$ -norm of the  $f_{j,k}$  and the  $\mathbb{L}_\infty$ -norm of the  $f_{j,k}^\eta$ . From now on, we denote  $\gamma := \frac{1-\eta}{4\eta}$ .

**Lemma 5.1** (Aubry *et al.* [6]). *There exists a constant  $\mathcal{C}_2$  such that for  $N$  large enough*

$$\sum_{j,k=0}^N \|f_{j,k}\|_2^2 \leq \mathcal{C}_2 N^{\frac{17}{6}}.$$

For  $\eta \in ]1/2, 1[$  and  $\gamma = \frac{1-\eta}{4\eta}$ , there exists constant  $\mathcal{C}_\infty^\eta$  such that for  $N$  large enough

$$\sum_{j,k=0}^N \|f_{j,k}^\eta\|_\infty^2 \leq \mathcal{C}_\infty^\eta N^{-\frac{2}{3}} e^{16\gamma N}.$$

For the  $\mathbb{L}_\infty$ -norm of the  $f_{j,k}^\eta$ , this lemma is slightly different from Lemmata 4 and 5 in [6] where the sum is over  $j+k=0, \dots, N$ . The proof remains similar.

### 5.3.2 Testing procedure

In this part, we propose a testing procedure which allow to choose among the hypothesis  $H_0$  and  $H_1$  defined in Section 5.2.2 by

$$\begin{cases} H_0 : & \rho = \tau, \\ H_1(\mathcal{C}, \varphi_n) : & \rho \in \mathcal{R}(B, r, L) \text{ s.t. } \|\rho - \tau\|_2 \geq \mathcal{C} \cdot \varphi_n, \end{cases}$$

As in the alternative  $H_1$ , the true value of  $\|\rho - \tau\|_2^2$  is unknown, we have to derive an estimator of this quantity. Then, in a second time, we provide our testing procedure  $\Omega_n$ .

For the known efficiency parameter  $\eta \in ]1/2, 1]$ , we define an estimator  $M_n$ , also called a test statistic of  $\|\rho - \tau\|_2^2$ . the estimator  $M_n$  is as an U-statistic of order 2 based on indirect observations  $(Y_i, \Phi_i)_{i=1, \dots, n}$ . For  $N := N(n) \rightarrow \infty$  as  $n \rightarrow \infty$ ,  $M_n$  is defined by

$$M_n := \frac{1}{n(n-1)} \sum_{j,k=0}^{N-1} \sum_{\ell \neq m=1}^n \left[ F_{j,k}^\eta \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) - \tau_{j,k} \right] \left[ \overline{F_{j,k}^\eta} \left( \frac{Y_m}{\sqrt{\eta}}, \Phi_m \right) - \overline{\tau_{j,k}} \right], \quad (5.16)$$

where  $F_{j,k}^\eta(x, \phi) := f_{j,k}^\eta(x) e^{-i(k-j)\phi}$  uses the *pattern* functions defined in (5.15) and  $\bar{a}$  denote the complex conjugate of  $a$ . Let us note that the density matrices are infinite, thus we truncate the sum to  $N$ . The parameter  $N$  is called the bandwidth and has to be optimized.

We study the properties of  $M_n$  in Section 5.3.3

Let us discuss the construction of our estimator  $M_n$ . First remark that we do not estimate

$$d = \sum_{j,k=0}^{\infty} |\rho_{j,k} - \tau_{j,k}|^2 = \sum_{j,k=0}^{N-1} |\rho_{j,k} - \tau_{j,k}|^2 + \sum_{j,k=N}^{\infty} |\rho_{j,k} - \tau_{j,k}|^2 := d_N + R_N$$

but the truncated part  $d_N$ . For an appropriate choice of  $N$ , the residual term  $R_N$  should be small due to the considered decrease condition (5.11). Moreover, we have underlined previously that  $\rho_{j,k} = E_\rho[F_{j,k}(X, \Phi)]$ , thus we can estimate  $\rho_{j,k}$  by the empirical version  $\hat{\rho}_{j,k}^N$ ,  $j, k < N$  (see [6]). For the case  $\eta \in ]0, 1[$  we shall use the *pattern* functions defined in (5.15). We estimate then  $d_N$  by an U-statistic of order 2 given by (5.16).

Now, for a constant  $\mathcal{C} > 0$  and some threshold  $t_n > 0$  defined later, we can define our testing procedure, based on the test statistic  $M_n$  defined in (5.16), as

$$\Omega_n = \mathbf{1}(|M_n| > \mathcal{C} \cdot t_n^2). \quad (5.17)$$

Here,  $\Omega_n = 0$  when  $\rho = \tau$  and  $\Omega_n = 1$  when the distance between the unknown state  $\rho$  and the presumed state  $\tau$  is larger than  $\mathcal{C}t_n^2$ . We study the upper bounds of our procedure  $\Omega_n$  in the sense of Definition 5.13 in Section 5.3.4.

### 5.3.3 Properties of the estimator $M_n$

We first remark that each element of the density matrix  $\rho_{j,k}$ , such that  $j, k < N$ , is estimated with no bias. Moreover, the estimator  $M_n$  is unbiased under  $\rho = \tau$  defined in  $H_0$  (Remark 5.1). Remark 5.2 derives useful tools for the proof of the upper bounds in the sense of Definition 5.1. From now on, we denote by  $E_\rho$  the expected value under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$  and  $E_\tau$  under  $\rho = \tau$  defined in  $H_0$ .

**Remark 5.1.** Note that for  $0 \leq k \leq j \leq N - 1$  and  $\eta \in ]1/2, 1[$

$$E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] = \rho_{j,k} \quad \text{and} \quad E_\tau \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] = \tau_{j,k}.$$

Indeed, from Plancherel formula

$$\begin{aligned} E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] &= \iint_0^\pi \int f_{j,k}^\eta \left( \frac{y}{\sqrt{\eta}} \right) e^{-i(k-j)\phi} p_\rho^\eta(y, \phi) d\phi dy \\ &= \iint_0^\pi f_{j,k}^\eta(y) e^{-i(k-j)\phi} \sqrt{\eta} p_\rho^\eta(\sqrt{\eta}y, \phi) d\phi dy \\ &= \iint_0^\pi e^{-i(k-j)\phi} \frac{1}{2\pi} \tilde{f}_{j,k}(t) e^{\frac{1-\eta}{4\eta}t^2} \mathcal{F}_1[\sqrt{\eta} p_\rho^\eta(\cdot, \sqrt{\eta}, \phi)](t) d\phi dt. \end{aligned}$$

From equation (5.10), and from Plancherel formula

$$\begin{aligned} E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] &= \iint_0^\pi \frac{e^{-i(k-j)\phi}}{2\pi} \tilde{f}_{j,k}(t) e^{\frac{1-\eta}{4\eta}t^2} \mathcal{F}_1[p_\rho(\cdot, \phi)](t) \tilde{N}^\eta(t/\sqrt{\eta}) d\phi dt \\ &= \iint_0^\pi e^{-i(k-j)\phi} f_{j,k}(x) p_\rho(x, \phi) d\phi dx = \rho_{j,k}. \end{aligned}$$

Thus,

$$E_\rho [M_n] = \sum_{j,k=0}^{N-1} (\rho_{j,k} - \tau_{j,k})^2 \quad \text{and} \quad E_\tau [M_n] = 0.$$

**Remark 5.2.** For  $N$  large enough and  $\rho$  belonging to the class  $\mathcal{R}(B, r, L)$

$$\sum_{j,k=0}^{N-1} E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] \leq \begin{cases} \mathcal{C}_\infty^\eta N^{-2/3} e^{16\gamma N} & \eta \in ]1/2, 1[, \\ C \cdot \mathcal{C}_2 N^{17/6} & \eta = 1, \end{cases}$$

where  $\mathcal{C}_\infty^\eta$  and  $\mathcal{C}_2$  are constants defined in Lemma 5.1 and  $C$  is a positive constant.

*Proof.* Since  $E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] = E_\rho \left[ \left| f_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right]$  For  $\eta \in ]1/2, 1[$

$$\sum_{j,k=0}^{N-1} E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] \leq \begin{cases} \sum_{j,k=0}^{N-1} \|f_{j,k}^\eta\|_\infty^2 & \eta \in ]1/2, 1[, \\ C \sum_{j,k=0}^{N-1} \|f_{j,k}^\eta\|_2^2 & \eta = 1. \end{cases}$$

For  $\eta = 1$ , we first apply Lemma 6 in [6] where it has been established that

$$\sup_{x \in \mathbb{R}} \int_0^\pi p_\rho(x, \phi) d\phi \leq C,$$

where  $C$  is a positive constant. Hence, the result is a direct consequence of Lemma 5.1.  $\square$

In order to choose the optimal the bandwidth  $N$ , we do the classical bias/variance trade-off of the risk  $E \left[ |M_n - \|\rho - \tau\|_2^2|^2 \right] = B^2(M_n) + V(M_n)$ , where the bias and the variance term are respectively defined as follows:

$$\begin{aligned} B(M_n) &:= |E[M_n] - \|\rho - \tau\|_2^2|, \\ V(M_n) &:= E[|M_n - E[M_n]|^2]. \end{aligned}$$

From now on, we denote by  $B_\rho, V_\rho$  the variance and the bias terms under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$  and  $V_\tau, B_\tau$  the variance and the bias terms under  $\rho = \tau$  defined in  $H_0$ . Note that under  $\rho = \tau$  defined in  $H_0$ , the bias term is written  $B_\tau(M_n) := |E_\tau[M_n]|$  and is equal to 0. Hence, the following Propositions 5.1 and 5.2 evaluate these quantities.

**Proposition 5.1.** For  $r \in ]0, 2]$ , and  $\eta \in ]1/2, 1]$  we have

$$\begin{aligned} B_\tau(M_n) &= 0, \\ B_\rho(M_n) &\leq C_B N^{2-r/2} e^{-2BN^{r/2}}, \text{ for } N \text{ large enough,} \end{aligned}$$

where  $C_B$  is a positive constant depending only on  $B$ ,  $r$  and  $L$ .

**Proposition 5.2.** For  $r \in ]0, 2]$ ,  $\eta = 1$  and for  $N$  large enough such that  $N^{17/6}/n \rightarrow 0$  as  $n \rightarrow \infty$ , we have

$$\begin{aligned} V_\rho(M_n) &\leq \frac{8C \cdot \mathcal{C}_2}{n} N^{17/6}, \\ V_\tau(M_n) &\leq \frac{(C \cdot \mathcal{C}_2)^2}{n^2} N^{17/3}. \end{aligned}$$

For  $r \in ]0, 2]$ ,  $\eta \in ]1/2, 1[$ ,  $\gamma = (1-\eta)/(4\eta)$  and for  $N$  large enough such that  $N^{-2/3} e^{16\gamma N}/n \rightarrow 0$  as  $n \rightarrow \infty$ , we have

$$\begin{aligned} V_\rho(M_n) &\leq \frac{8\mathcal{C}_\infty^\eta N^{-2/3}}{n} e^{16\gamma N}, \\ V_\tau(M_n) &\leq \frac{(\mathcal{C}_\infty^\eta)^2 N^{-4/3}}{n^2} e^{32\gamma N}. \end{aligned}$$

The constants  $\mathcal{C}_\infty^\eta$  and  $\mathcal{C}_2$  are defined in Lemma 5.1 and  $C$  is positive constant.

### 5.3.4 Main results

In the following Theorem, we establish upper bounds for the testing rates in the sense of Definition 5.1 where  $\rho$  is the unknown density matrix supposed to belong to the class  $\mathcal{R}(B, r, L)$  defined in (5.11). Theorem 5.1-(1) deals with the ideal detection case while Theorem 5.1-(2) and Theorem 5.2 take into account the Gaussian noise.

**Theorem 5.1.** The test procedure  $\Omega_n$  defined in (5.17) for the bandwidth  $N(n)$ , the threshold  $t_n$  and the constant  $\mathcal{C}^*$  satisfies the upper bound (5.13) for the rate  $\varphi_n$  such that

1. for  $r \in ]0, 2]$ ,  $\eta = 1$ , the bandwidth  $N(n) := N_1$  is equal to

$$N_1 := \left( \frac{\log n}{4B} + \frac{(\log \log n)^2}{4B} \right)^{2/r}, \quad (5.18)$$

and the rate

$$\varphi_n^2 = t_n^2 = n^{-1/2} (\log n)^{\frac{17}{6r}}, \quad (5.19)$$



2. for  $r = 2$ ,  $\eta \in ]1/2, 1[$  and  $\gamma := \frac{1-\eta}{4\eta}$ , the bandwidth  $N(n) := N_2$  is equal to

$$N_2 := \frac{\log(n)}{4(4\gamma + B)} \left( 1 + \frac{8 \log(\log n)}{3 \log(n)} \right), \quad (5.20)$$

and the rate

$$\varphi_n^2 = t_n^2 = \log(n)^{\frac{12\gamma-B}{3(4\gamma+B)}} n^{-\frac{B}{2(4\gamma+B)}}. \quad (5.21)$$

**Theorem 5.2.** For  $r \in ]0, 2[$ ,  $\eta \in ]1/2, 1[$  and  $\gamma = \frac{1-\eta}{4\eta}$ , the test procedure  $\Omega_n$  defined in (5.17) for the bandwidth  $N := N_3$  solution of the equation

$$16\gamma N_3 + 4BN_3^{r/2} = \log n, \quad (5.22)$$

the threshold  $t_n$  and the constant  $\mathcal{C}^*$  satisfies the upper bound (5.13) for the rate  $\varphi_n$

$$\varphi_n^2 = t_n^2 = N_3^{2-r/2} e^{-2BN_3^{r/2}}. \quad (5.23)$$

We can remark that

$$\limsup_{n \rightarrow \infty} \left\{ P_\tau[\Omega_n = 1] + \sup_{\rho \in H_1(\mathcal{C}, \varphi_n)} P_\rho[\Omega_n = 0] \right\} = 0.$$

Theorem 5.1 and 5.2 provide upper bounds for our testing procedure defined in equation (5.17) for the testing rates  $\varphi_n$ . We first remark that our procedure gives nearly parametric rate up to a logarithmic factor when we are in the framework of ideal detection (Theorem 5.1-(2): no noise) and supersmooth corresponding Wigner functions for all  $r \in ]0, 2[$ . In the setting of Theorem 5.1-(2) and Theorem 5.2, there are good reasons to believe that our testing procedure achieves optimal rates. This remark follows from a recent work of Butucea ([16]), where the author establishes minimax rates for testing in  $\mathbb{L}_2$ -norm from indirect observations. However, we do not attempt to go that far in this chapter.

In the framework of quantum statistics, one can investigate another testing approach based on kernel type estimator of the Wigner function describing the quantum state. Such a testing procedure can be directly derived from the kernel estimator in [71] of the quadratic functional  $\iint W_\rho^2$ , where  $W_\rho$  is the Wigner function associated to the quantum state  $\rho$ . Our test problem is equivalent to the following:

$$\begin{cases} H_0 : & W_\rho = W_{\rho_0}, \\ H_1 : & \sup_{\rho \in \mathcal{R}(B, r, L)} \|W_\rho - W_{\rho_0}\|_2 \geq \mathcal{C} \cdot \varphi_n \end{cases}$$

where  $\varphi_n$  is a sequence which tends to 0 when  $n \rightarrow \infty$ . We conjecture that its performances are comparable to those found in this chapter and we will leave this analysis for a separate work.

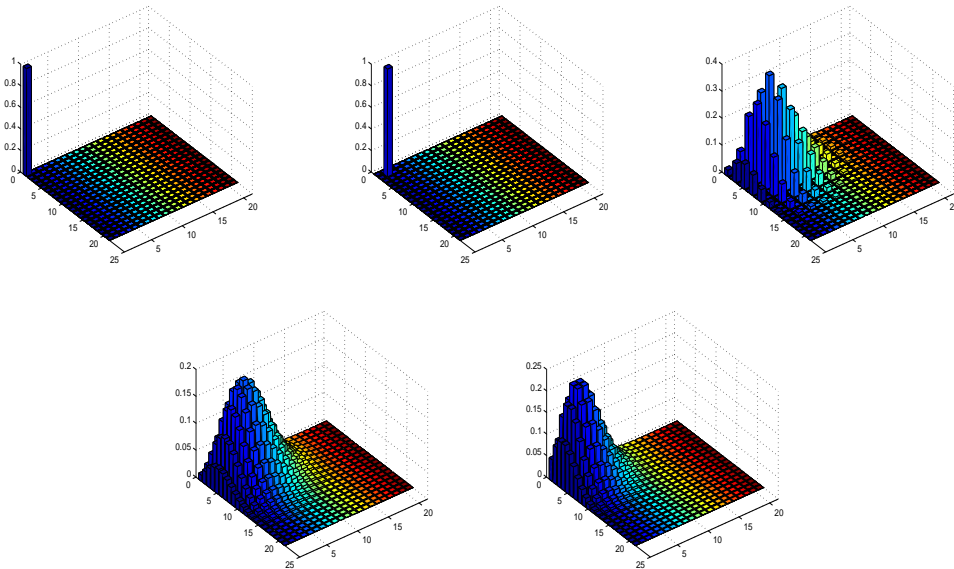
## 5.4 Simulations

From now on, we set  $r = 2$ . The purpose of this section is to implement our testing procedure and to investigate its numerical performances. Our motivation is, given a density matrix  $\tau \in \mathcal{R}(B, r)$ , to decide whether  $H_0$  or  $H_1$  is accepted

$$\begin{cases} H_0 : & \rho = \tau, \\ H_1(\mathcal{C}, \varphi_n) : & \rho \in \mathcal{R}(B, r, L) \text{ s.t. } \|\rho - \tau\|_2 \geq \mathcal{C} \cdot \varphi_n, \end{cases}$$

where  $\varphi_n$  is a sequence, which tends to 0 when  $n \rightarrow \infty$ .

Figure 5.1: Density matrices of the *vacuum* state, the *single photon* state, the *Schrödinger cat-3*, the *coherent-3* state and the *coherent- $\sqrt{6}$*  state respectively



We propose to simulate two different situations. In the first one, case **A**, we consider quantum states easily distinguishable, while in the second one, case **B**, we deal with quantum states, which are quite similar and it is difficult to differentiate between them. For  $\tau$  defined in  $H_0$ , we sample our procedure from different density matrices  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$  such that in

- the case **A**:  $\tau$  is the *vacuum* state, while
  - **a**)  $\rho$  is the *vacuum* state ( $\rho = \tau$ ),
  - **b**)  $\rho$  is the *single photon* state,
  - **c**)  $\rho$  is the *Schrödinger cat-3* state,

- the case **B**:  $\tau$  is the *coherent-3* state, while
  - a)  $\rho$  is the *coherent-3* state ( $\rho = \tau$ ),
  - b)  $\rho$  is the *coherent- $\sqrt{6}$*  state,
  - c)  $\rho$  is the *Schrödinger cat-3* state.

The latter case is a more complicated situation for two main reasons. The *Schrödinger cat- $q_0$*  state corresponds to a linear superposition of two *coherent* states ( $\pm q_0$ ). Moreover, the probability density of a *coherent- $q_0$*  state is Gaussian with a mean proportional to  $q_0$  (see Table 5.1). Figure 5.1 represents the density matrices of the states we consider in our simulations.

In each situation (cases **A** and **B**), we shall distinguish the case  $\eta = 1$ , which corresponds to the ideal detection, from the case  $\eta = 0.9$ , i.e. when noise is present. The latter case is the practical one in laboratory. From now on and for both situations

- when  $\eta = 1$ , we set  $N = 15$ ,
- when  $\eta = 0.9$ , we set  $N = 14$  and  $N = 13$ .

The values of  $N$  have been chosen after different simulations for different values of  $N$ . We notice that, when we are in presence of noise ( $\eta = 0.9$ ), we have to take  $N$  smaller than the one we choose in the ideal setting. It can be explained by the fact that the variance term of our estimator increase exponentially with  $N$  in the case  $\eta = 0.9$  (see Proposition 5.2).

To implement our procedure, we first compute our modified *pattern* function  $f_{j,k}^\eta(x)$  in Section 5.4.1. We implement in Section 5.4.2 the estimator  $M_n$  defined in (5.16) and finally study the performance of our test procedure  $\Omega_n$  defined in (5.17) in Section 5.4.3.

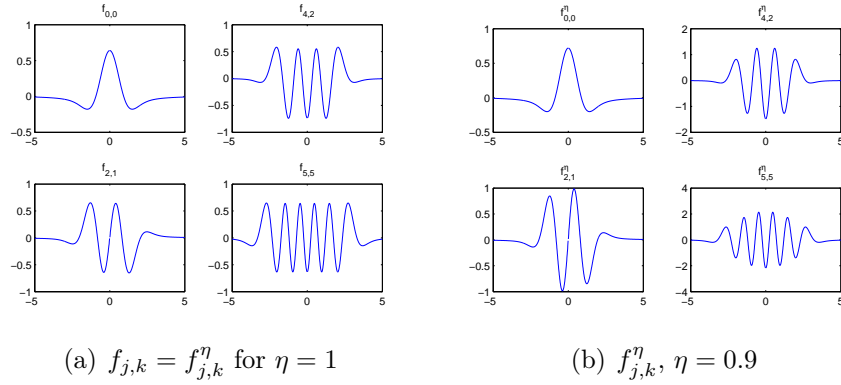
#### 5.4.1 Pattern functions $f_{j,k}^\eta$

To implement our procedure, we need the modified *pattern* function  $f_{j,k}^\eta(x)$  defined in (5.15) for all  $j \geq k$ . For this purpose, we compute here the inverse Fourier Transform  $f_{j,k}^\eta(x)$  of the  $\tilde{f}_{j,k}^\eta(t)$  given explicitly by (5.15) and the generalized Laguerre polynomials. Previous authors have used a different method to implement the *pattern* functions in [67, 4] via the following recurrence relation given in [65]:

$$f_{j,k}(x) = 2x\psi_k(x)\varphi_j(x) - \sqrt{2(k+1)}\psi_{k+1}(x)\varphi_j(x) - \sqrt{2(j+1)}\psi_k(x)\varphi_{j+1}(x),$$

for  $j \geq k$ , otherwise  $f_{j,k}(x) = f_{k,j}(x)$ . We display in Figure 5.2-(a) the corresponding graphical representations of the *pattern* functions up to a constant  $\pi$  as in physical literature

$\pi^{-1}f_{j,k}$  instead of  $f_{j,k}$  are often called *pattern* functions. Figure 5.2-(b) represents some modified *pattern* functions  $\pi^{-1}f_{j,k}^\eta$  for  $\eta = 0.9$ .

Figure 5.2: Examples of *pattern* functions

For some of  $f_{j,k}$  we expressed below their explicit form.

$$\begin{aligned}
 f_{0,0}(x) &= 2 - 2e^{-x^2} x \sqrt{\pi} \operatorname{Erfi}(x), \\
 f_{2,1}(x) &= e^{-x^2} \left[ -2e^{x^2} x(-3 + 2x^2) + \sqrt{\pi}(1 - 8x^2 + 4x^4) \operatorname{Erfi}(x) \right], \\
 f_{4,2}(x) &= \frac{e^{-x^2}}{2\sqrt{3}} \left[ 2e^{x^2} (-4 + 27x^2 - 24x^4 + 4x^6) \right. \\
 &\quad \left. + \sqrt{\pi} x (21 - 74x^2 + 52x^4 - 8x^6) \operatorname{Erfi}(x) \right], \\
 f_{5,5}(x) &= \frac{e^{-x^2}}{30} \left[ 2e^{x^2} (-30 + 435x^2 - 865x^4 + 526x^6 - 116x^8 + 8x^{10}) \right. \\
 &\quad \left. + \sqrt{\pi} x (225 - 1425x^2 + 2160x^4 - 1160x^6 + 240x^8 - 16x^{10}) \operatorname{Erfi}(x) \right].
 \end{aligned}$$

### 5.4.2 Implementation of $M_n$

The purpose of this section is to investigate the performance of the estimator  $M_n$  of  $\|\rho - \tau\|_2^2$  defined in (5.16) by

$$M_n = \frac{1}{n(n-1)} \sum_{j,k=0}^{N-1} \sum_{\ell \neq m=1}^n \left[ F_{j,k}^\eta \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) - \tau_{j,k} \right] \left[ \overline{F_{j,k}^\eta} \left( \frac{Y_m}{\sqrt{\eta}}, \Phi_m \right) - \overline{\tau_{j,k}} \right],$$

where  $F_{j,k}^\eta(x, \phi) = f_{j,k}^\eta(x) e^{-i(k-j)\phi}$  and  $\tau$  is defined in  $H_0$ . Table 5.2 gives the actual values of  $\|\rho - \tau\|_2^2$  in all considered setups.

From now on and for all considered cases (**A** and **B**), the procedure to simulate one value of  $M_n$  is designed as follows. We first sample  $n = 50\,000$  i.i.d. data from each of the three

Table 5.2: The values of  $\|\rho - \tau\|_2^2$  for  $\tau$  the *vacuum* state (case **A**) and  $\tau$  the *coherent-3* state (case **B**)

Case <b>A</b> :	<b>a)</b> $\rho$ <i>vacuum</i>	<b>b)</b> $\rho$ <i>single photon</i>	<b>c)</b> $\rho$ <i>Schrödinger cat-3</i>
	0	2	1.9556
Case <b>B</b> :	<b>a)</b> $\rho$ <i>coherent-3</i>	<b>b)</b> $\rho$ <i>coherent-<math>\sqrt{6}</math></i>	<b>c)</b> $\rho$ <i>Schrödinger cat-3</i>
	0	0.2812	0.9999

probability densities  $p_\rho$  defined in Table 5.1 for each considered  $\rho$  (Figure 5.1). The Gaussian noise component  $\xi$  is simulated independently with a variance equal to  $\sqrt{(1-\eta)/2}$ . Then, we run independently of the other runs, 1 000 values of  $M_n$  denoted by  $\{M_n^k\}_{k=1,\dots,1000}$ , where each value  $M_n^k$  is based on  $n = 50\,000$  i.i.d. noisy data  $(Y_\ell, \Phi_\ell)$ .

We recall that the choice of  $N = 15$  corresponds to the setting of ideal detection with  $\eta = 1$ , while  $N = 14$  and  $N = 13$  deal with the noisy detection with  $\eta = 0.9$ . Note that for  $\rho = \tau$  we expect the  $M_n^k$  to be close to 0, while for  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$  we expect the  $M_n^k$  to be close to  $\|\rho - \tau\|_2^2 > 0$ .

Table 5.3: The median values of  $M_n^k$  for  $\tau$  the *vacuum* state (case **A**) and  $\tau$  the *coherent-3* state (case **B**)

Case <b>A</b> :	<b>a)</b> $\rho$ <i>vacuum</i>	<b>b)</b> $\rho$ <i>single photon</i>	<b>c)</b> $\rho$ <i>Schrödinger cat-3</i>
$N = 15, \eta = 1$	0.0043	1.9972	1.9417
$N = 14, \eta = 0.9$	0.0256	1.9936	1.9360
$N = 13, \eta = 0.9$	0.0144	1.9943	1.9378
Case <b>B</b> :	<b>a)</b> $\rho$ <i>coherent-3</i>	<b>b)</b> $\rho$ <i>coherent-<math>\sqrt{6}</math></i>	<b>c)</b> $\rho$ <i>Schrödinger cat-3</i>
$N = 15, \eta = 1$	$-6.8721 \cdot 10^{-4}$	0.2688	0.8962
$N = 14, \eta = 0.9$	-0.0466	0.1704	0.7735
$N = 13, \eta = 0.9$	-0.0376	0.1847	0.7840

In frameworks **A** and **B**, we draw boxplots of the  $\{M_n^k\}_{k=1,\dots,1000}$  in Figures 5.3 to 5.8. For a clearer reading, these boxplots (Figures 5.3 to 5.8) have been moved in an appendix at the end of the chapter in Section 5.6. We summarize the median values in Table 5.3 and we can make the following remarks.

- In the setting **A**: the procedure  $M_n$  is a good estimation of  $\|\rho - \tau\|_2^2$  both for  $\eta = 1$

and  $\eta = 0.9$ .

- In the setting **B**: when  $\rho = \tau$ , the procedure  $M_n$  shows excellent results since the  $M_n^k$  are close to  $\|\rho - \tau\|_2^2 = 0$  when either  $\eta = 1$  or  $\eta = 0.9$ . In the ideal case and when  $\rho$  is the *coherent- $\sqrt{6}$*  state, the procedure  $M_n$  gives a good result since the median of the boxplot of the  $M_n^k$  is equal to 0.2688 and the true value  $\|\rho - \tau\|_2^2 = 0.2812$ . Otherwise, the procedure  $M_n$  under evaluates the distance  $\|\rho - \tau\|_2^2$ .

In order to evaluate the quality of our procedure  $M_n$  we estimate the mean square error  $\text{MSE} = E_\rho[|M_n - \|\rho - \tau\|_2^2|^2]$  as the average over the 1 000 independent runs of  $|M_n - \|\rho - \tau\|_2^2|^2$ . In other words the MSE is empirically assessed as

$$\frac{1}{1\,000} \sum_{k=1}^{1\,000} |M_n^k - \|\rho - \tau\|_2^2|^2,$$

with  $\|\rho - \tau\|_2^2$  given by Table 5.2. Table 5.4 summarizes the results. As we have already noticed the procedure  $M_n$  gives excellent results in every cases but it has a larger MSE when we evaluate the distance between the *coherent-3* state and the *Schrödinger cat-3* state: MSE is equal to 0.0531 and 0.0484.

Table 5.4: The empirical values of MSE for  $\tau$  the *vacuum* state (case **A**) and  $\tau$  the *coherent-3* state (case **B**)

Case <b>A</b> :	a) $\rho$ <i>vacuum</i>	b) $\rho$ <i>single photon</i>	c) $\rho$ <i>Schrödinger cat-3</i>
$N = 15, \eta = 1$	$2.3988 \cdot 10^{-5}$	$1.3175 \cdot 10^{-4}$	$4.7877 \cdot 10^{-4}$
$N = 14, \eta = 0.9$	0.0181	0.0056	0.0071
$N = 13, \eta = 0.9$	0.0069	0.0025	0.0051
Case <b>B</b> :	a) $\rho$ <i>coherent-3</i>	b) $\rho$ <i>coherent-<math>\sqrt{6}</math></i>	c) $\rho$ <i>Schrödinger cat-3</i>
$N = 15, \eta = 1$	$1.4303 \cdot 10^{-6}$	$1.9473 \cdot 10^{-4}$	0.0110
$N = 14, \eta = 0.9$	0.0044	0.0132	0.0531
$N = 13, \eta = 0.9$	0.0019	0.0093	0.0484

### 5.4.3 Studies of the performance of our test procedure $\Omega_n$

In this part, we would like to confirm the performance of our testing procedure  $\Omega_n$ . In order to appreciate it, we are interested in the *power* of our test  $\Pi$ , such that  $\Pi = P_\rho[\Omega_n = 1]$  under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ . In this view, we want to compare our estimator  $M_n$  with a threshold  $\nu_n$  (with  $\nu_n = \mathcal{C}^* t_n^2$  in (5.14)) and decide as follows:

- if  $|M_n| > \nu_n$ , we accept  $H_1$ ,
- otherwise we accept  $H_0$ .

In the statistical literature, the parameter  $\nu_n = \nu_n(\alpha)$  is also called the critical value associated to  $\alpha = P_\tau[\Omega_n = 1]$ , which is the probability error of first type (under  $\tau$  defined in  $H_0$ ) and defined in Section 5.2.2. In our framework, we set  $\alpha = 1\%$  and  $\alpha = 5\%$ . As we don't know the density probability of  $M_n$  under  $\rho = \tau$  defined in  $H_0$ , we shall evaluate empirically the threshold  $\nu_n$  at the testing level  $\alpha$  for our particular choices of  $\tau$ ,  $n$ ,  $\alpha$  and  $\eta$ . In this purpose and independently of the future runs, we first compute, for  $n = 50\,000$  and for all considered cases 1 000 independent values  $\{M_n^k\}_{k=1,\dots,1\,000}$  of the estimator  $M_n$  as described previously. We summarize our results in Table 5.5. We report that the obtained values  $\nu_n$  are larger when  $\eta = 0.9$  than when  $\eta = 1$ . It is due to the noise effect.

Table 5.5: Empirical values of  $\nu_n$  for  $\tau$  the *vacuum* state (case **A**) and  $\tau$  the *coherent-3* state (case **B**)

		$N = 15$	$N = 14$	$N = 13$
		$\eta = 1$	$\eta = 0.9$	$\eta = 0.9$
Case <b>A</b>	$\nu_n(1\%)$	0.0096	0.4605	0.2692
	$\nu_n(5\%)$	0.0079	0.2739	0.1666
Case <b>B</b>	$\nu_n(1\%)$	0.0028	0.1453	0.0798
	$\nu_n(5\%)$	0.0022	0.1100	0.0713

From now on, we fix the value of  $\nu_n$  as in Table 5.5. With the same protocol as above and for all considered cases, we compute 1 000 other independent values  $\{M_n^k\}_{k=1,\dots,1\,000}$  of the estimator  $M_n$ , for  $n = 50\,000$ . Hence, we evaluate the empirical power of our testing procedure and the empirical first type error. Tables 5.6 and 5.7 provide the empirical results obtained by our test procedure in the experiments we deal with. We see that our testing procedure provides very good results even in the framework **B**, we obtain powers  $\Pi$  equal to 1 since  $N = 13$  when we are in presence of noise. Otherwise, for  $N = 14$  and  $\eta = 0.9$  the powers of our testing procedure is a little bit degraded, but still remarkably good, since  $\Pi = 0.7160$  for  $\alpha = 1\%$  and  $\Pi = 0.9440$  for  $\alpha = 5\%$  in the framework **B** when  $\rho$  is the *coherent- $\sqrt{6}$*  state, the optimal  $N$  when we are in presence of noise is  $N = 13$  with powers of test equal to 1.

Table 5.6: Empirical values of the first type error  $\alpha$  and the power of the test  $\Pi$  over 1 000 runs for  $\nu_n$  given in Table 5.5 for  $\tau$  the *vacuum* state (case **A**)

		<b>a) <math>\rho</math> vacuum</b>	<b>b) <math>\rho</math> single ph.</b>	<b>c) <math>\rho</math> Schrödinger C.-3</b>
		$\alpha$	$\Pi$	$\Pi$
$N = 15, \eta = 1$	$\nu_n(1\%)$	0.0150	1.0000	1.0000
	$\nu_n(5\%)$	0.0570	1.0000	1.0000
$N = 14, \eta = 0.9$	$\nu_n(1\%)$	0.0070	1.0000	1.0000
	$\nu_n(5\%)$	0.0550	1.0000	1.0000
$N = 13, \eta = 0.9$	$\nu_n(1\%)$	0.0130	1.0000	1.0000
	$\nu_n(5\%)$	0.0680	1.0000	1.0000

Table 5.7: Empirical values of the first type error  $\alpha$  and the power of the test  $\Pi$  over 1 000 runs for  $\nu_n$  given in Table 5.5 for  $\tau$  the *coherent-3* state (case **B**)

		<b>a) <math>\rho</math> coherent-3</b>	<b>b) <math>\rho</math> coherent-<math>\sqrt{6}</math></b>	<b>c) <math>\rho</math> Schrödinger C.-3</b>
		$\alpha$	$\Pi$	$\Pi$
$N = 15, \eta = 1$	$\nu_n(1\%)$	0.0090	1.0000	1.0000
	$\nu_n(5\%)$	0.0530	1.0000	1.0000
$N = 14, \eta = 0.9$	$\nu_n(1\%)$	0.0120	0.7160	1.0000
	$\nu_n(5\%)$	0.0680	0.9440	1.0000
$N = 13, \eta = 0.9$	$\nu_n(1\%)$	0.0230	1.0000	1.0000
	$\nu_n(5\%)$	0.0620	1.0000	1.0000



## 5.5 Proof

In this section, we give the proofs of the Theorem 5.1 and 5.2 derived in Section 5.3.4 and the proofs of the Proposition 5.1 and 5.2 established in Section 5.3.3.

### 5.5.1 Proof of the Theorems

In the paragraphs below, we establish the results of the Theorem 5.1 and 5.2 derived in Section 5.3.4. We begin by Theorem 5.1-(1).

**Proof of Theorem 5.1-(1).** Take  $r \in ]0, 2]$ ,  $\eta = 1$ , the bandwidth  $N_1$  defined in the equation (5.18) and  $\varphi_n^2$  defined in (5.19), we have by Proposition 5.2-2 and Proposition 5.2-1

$$V_\tau(M_n) \leq \frac{(C \cdot \mathcal{C}_2)^2}{n^2} N_1^{17/3} \leq C_{V_\tau} \varphi_n^8, \quad (5.24)$$

$$V_\rho(M_n) \leq \frac{8C \cdot \mathcal{C}_2}{n} N_1^{17/6} \leq C_V \varphi_n^4, \quad (5.25)$$

where  $C_V$  and  $C_{V_\tau}$  are positive constants depending only on  $B$ ,  $r$  and  $L$ .

Under  $\rho = \tau$  defined in  $H_0$ , from Proposition 5.1, equations (5.24) and (5.19), we bound from above the first type error as follows:

$$P_\tau[\Omega_n = 1] = P_\tau[|M_n| \geq \mathcal{C}^* t_n^2] \leq \frac{E_\tau[|M_n|^2]}{\mathcal{C}^{*2} t_n^4} = \frac{V_\tau[M_n]}{\mathcal{C}^{*2} t_n^4} \leq \frac{C_{V_\tau} \varphi_n^4}{\mathcal{C}^{*2}} \rightarrow 0,$$

as  $n \rightarrow \infty$ .

On the other hand, under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , we have the second type error bounded as follows:

$$\begin{aligned} P_\rho[\Omega_n = 0] &= P_\rho[|M_n| < \mathcal{C}^* t_n^2] \\ &\leq P_\rho[|M_n - E_\rho[M_n]| \geq \|\rho - \tau\|^2 - \mathcal{C}^* t_n^2 - B_\rho[M_n]]. \end{aligned}$$

Moreover, under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , let  $\mathcal{C} = \mathcal{C}^*(1+\delta)$ ,  $\delta \in ]0, 1[$ , Proposition 5.1, equations (5.29) and (5.23) imply that  $B_\rho \leq \frac{\delta}{2} \mathcal{C}^* \varphi_n^2$  for  $n$  large enough and then

$$\begin{aligned} P_\rho[\Omega_n = 0] &\leq P_\rho \left[ \frac{|M_n - E_\rho[M_n]|}{\sqrt{V_\rho(M_n)}} \geq \frac{\|\rho - \tau\|^2 - \mathcal{C}^* t_n^2 - B_\rho[M_n]}{\sqrt{C_V \varphi_n^2}} \right] \\ &\leq P_\rho \left[ \frac{|M_n - E_\rho[M_n]|}{\sqrt{V_\rho(M_n)}} \geq \frac{\delta \mathcal{C}^*/2}{\sqrt{C_V}} \right] \\ &\leq \frac{4C_V}{(\delta \mathcal{C}^*)^2} \leq \frac{\lambda}{2}, \end{aligned}$$

for  $\mathcal{C}^*$  large enough. □

**Proof of Theorem 5.1-(2).** For  $r = 2$ ,  $\eta \in ]1/2, 1[$ , the bandwidth  $N_2$  defined in the equation (5.20) and  $\varphi_n^2$  defined in (5.23), we know by Proposition 5.2-2 and Proposition 5.2-1 that

$$V_\tau(M_n) \leq \frac{(\mathcal{C}_\infty^\eta)^2}{n^2} N_2^{-4/3} e^{32\gamma N_2} \leq C'_{V_\tau} \varphi_n^8, \quad (5.26)$$

$$V_\rho(M_n) \leq \frac{8\mathcal{C}_\infty^\eta}{n} N_2^{-2/3} e^{16\gamma N_2} \leq C'_V \varphi_n^4, \quad (5.27)$$

where  $C'_V$  and  $C'_{V_\tau}$  are positive constants depending only on  $B$ ,  $r$ ,  $L$  and  $\eta$ .

On one hand, under  $\rho = \tau$  defined in  $H_0$ , by Proposition 5.1, equations (5.26) and (5.21), we have the first type error bounded as follows:

$$\begin{aligned} P_\tau[\Omega_n = 1] &= P_\tau[|M_n| \geq \mathcal{C}^* t_n^2] \leq \frac{E_\tau[|M_n|^2]}{\mathcal{C}^{*2} t_n^4} \\ &= \frac{V_\tau[M_n]}{\mathcal{C}^{*2} t_n^4} \leq \frac{C'_{V_\tau} \varphi_n^4}{\mathcal{C}^{*2}} \rightarrow 0, \end{aligned}$$

as  $n \rightarrow \infty$ .

On the other hand, under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , the second type error is bounded as follows:

$$\begin{aligned} P_\rho[\Omega_n = 0] &= P_\rho[|M_n| < \mathcal{C}^* t_n^2] \\ &\leq P_\rho[|M_n - E_\rho[M_n]| \geq \|\rho - \tau\|^2 - \mathcal{C}^* t_n^2 - B_\rho[M_n]]. \end{aligned}$$

From Proposition 5.1, equations (5.27) and (5.21), since we are under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , we get

$$\begin{aligned} P_\rho[\Omega_n = 0] &\leq P_\rho \left[ \frac{|M_n - E_\rho[M_n]|}{\sqrt{V_\rho(M_n)}} \geq \frac{\|\rho - \tau\|^2 - \mathcal{C}^* t_n^2 - B_\rho[M_n]}{\sqrt{C'_V \varphi_n^2}} \right] \\ &\leq P_\rho \left[ \frac{|M_n - E_\rho[M_n]|}{\sqrt{V_\rho(M_n)}} \geq \frac{\mathcal{C} - \mathcal{C}^* - C_B}{\sqrt{C'_V}} \right] \\ &\leq \frac{C'_V}{(\mathcal{C} - \mathcal{C}^* - C_B)^2} \leq \frac{\lambda}{2}, \end{aligned}$$

for  $\mathcal{C} > \mathcal{C}^*$  and  $\mathcal{C}^*$  large enough.  $\square$

**Proof of Theorem 5.2.** In the case  $r \in ]0, 2[$  and  $\eta \in ]1/2, 1[$ , for the bandwidth  $N_3$  solution of the equation (5.22) and for  $\varphi_n^2$  defined in (5.23), Proposition 5.2-2 and Proposition 5.2-1 give that

$$V_\tau(M_n) \leq \frac{(\mathcal{C}_\infty^\eta)^2}{n^2} N_3^{-4/3} e^{32\gamma N_3} \leq C''_{V_\tau} \varphi_n^8 \cdot N_3^{\frac{2(3r-14)}{3}} \leq C'_{V_\tau} \varphi_n^8 \quad (5.28)$$

$$V_\rho(M_n) \leq \frac{8\mathcal{C}_\infty^\eta}{n} N_3^{-2/3} e^{16\gamma N_3} \leq C''_V \varphi_n^4 \cdot N_3^{\frac{3r-14}{3}} \quad (5.29)$$

where  $C_V''$  and  $C_{V_\tau}''$  are positive constants depending only on  $B$ ,  $r$ ,  $L$  and  $\eta$ .

Under  $\rho = \tau$  defined in  $H_0$ , from Proposition 5.1, equations (5.28) and (5.23), it follows that the first type error satisfies

$$P_\tau[\Omega_n = 1] = P_\tau [|M_n| \geq C^* t_n^2] \leq \frac{E_\tau[|M_n|^2]}{C^{*2} t_n^4} = \frac{V_\tau[M_n]}{C^{*2} t_n^4} \leq \frac{C_{V_\tau}'' \varphi_n^4}{C^{*2}} \rightarrow 0,$$

as  $n \rightarrow \infty$ .

Furthermore, under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , the second type error is such that

$$\begin{aligned} P_\rho[\Omega_n = 0] &= P_\rho [|M_n| < C^* t_n^2] \\ &\leq P_\rho [|M_n - E_\rho[M_n]| \geq \|\rho - \tau\|^2 - C^* t_n^2 - B_\rho[M_n]]. \end{aligned}$$

From Proposition 5.1, equations (5.29) and (5.23), since we are under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ , we deduce that

$$P_\rho[\Omega_n = 0] \leq P_\rho \left[ \frac{|M_n - E_\rho[M_n]|}{\sqrt{V_\rho(M_n)}} \geq \frac{\mathcal{C} - C^* - C_B}{\sqrt{C_V'' N_3^{\frac{3r-14}{6}}}} \right] \leq \frac{C_V'' N_3^{\frac{3r-14}{3}}}{(\mathcal{C} - C^* - C_B)^2} \rightarrow 0,$$

as  $n \rightarrow \infty$ , for  $\mathcal{C} > C^*$  and  $C^*$  large enough.  $\square$

## 5.5.2 Proof of the Propositions

In the following paragraphs, we give the proofs of Propositions 5.1 and 5.2 established in Section 5.3.3.

**Proof of Proposition 5.1.** By Remark 5.1, for  $r \in ]0, 2]$

$$B_\rho(M_n) = |E_\rho[M_n] - \|\rho - \tau\|_2^2| = \left| E_\rho[M_n] - \sum_{j,k \geq 0} |\rho_{j,k} - \tau_{j,k}|^2 \right|.$$

It is easy to see that under  $\rho = \tau$  defined in  $H_0$ , we have  $B_\tau(M_n) = 0$ . Under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$ ,

$$\begin{aligned} B_\rho(M_n) &= \left| \sum_{j,k=0}^{N-1} |\rho_{j,k} - \tau_{j,k}|^2 - \sum_{j,k \geq 0} |\rho_{j,k} - \tau_{j,k}|^2 \right| \\ &= 2 \sum_{j=N}^{\infty} \sum_{k=0}^{j-1} |\rho_{j,k} - \tau_{j,k}|^2 + \sum_{j=N}^{\infty} |\rho_{j,j} - \tau_{j,j}|^2 \\ &\leq 4 \sum_{j=N}^{\infty} \sum_{k=0}^{j-1} (|\rho_{j,k}|^2 + |\tau_{j,k}|^2) + 2 \sum_{j=N}^{\infty} (|\rho_{j,j}|^2 + |\tau_{j,j}|^2) \end{aligned}$$

Since  $\tau$  and  $\rho$  belong to the class  $\mathcal{R}(B, r, L)$  defined in (5.11), it implies

$$\begin{aligned} B_\rho(M_n) &\leq 8L^2 \sum_{j=N}^{\infty} \sum_{k=0}^{j-1} e^{-2B(j+k)^{r/2}} + 4L^2 \sum_{j=N}^{\infty} e^{-2B(j+j)^{r/2}} \\ &\leq 8L^2 \int_N^{\infty} (j-1) e^{-2Bu^{r/2}} du + 4L^2 \int_N^{\infty} e^{-2B(2u)^{r/2}} du \\ &\leq C_B N^{2-r/2} e^{-2BN^{r/2}} + C'_B N^{1-r/2} e^{-2B(2N)^{r/2}}, \end{aligned}$$

where  $C_B$  and  $C'_B$  denote positive constants depending only on  $B$ ,  $r$  and  $L$ . As  $2^{r/2} > 1$  for all  $r > 0$ ,

$$B_\rho(M_n) \leq C_B N^{2-r/2} e^{-2BN^{r/2}} (1 + o(1)),$$

as  $N \rightarrow \infty$ . □

**Proof of Proposition 5.2.** By centering variables, we have  $M_n - E_\rho[M_n] := F_1 + F_2$  where

$$\begin{aligned} F_1 &:= \frac{1}{n(n-1)} \sum_{j,k=0}^{N-1} \sum_{\ell \neq m=1}^n \left( F_{j,k}^\eta \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) - E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \\ &\quad \times \left( \overline{F_{j,k}^\eta} \left( \frac{Y_m}{\sqrt{\eta}}, \Phi_m \right) - E_\rho \left[ \overline{F_{j,k}^\eta} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \\ F_2 &:= \frac{2}{n} \sum_{j,k=0}^{N-1} \sum_{\ell=1}^n \operatorname{Re} \left( \left( F_{j,k}^\eta \left( \frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell \right) - E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \right. \\ &\quad \left. \times E_\rho \left[ \overline{F_{j,k}^\eta} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) - \overline{\tau_{j,k}} \right] \right) \end{aligned}$$

where  $\operatorname{Re}(z)$  stands for the real part of the complex number  $z$ .

Then,  $V_\rho(M_n) := E_\rho[|M_n - E_\rho[M_n]|^2] = E_\rho[|F_1|^2] + E_\rho[|F_2|^2]$ . First deal with the first term of the previous sum  $E_\rho[|F_1|^2]$

$$\begin{aligned} E_\rho[|F_1|^2] &= \frac{1}{n(n-1)} E_\rho \left[ \left| \sum_{j,k=0}^{N-1} \left( F_{j,k}^\eta \left( \frac{Y_1}{\sqrt{\eta}}, \Phi_1 \right) - E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \right. \right. \\ &\quad \left. \left. \left( \overline{F_{j,k}^\eta} \left( \frac{Y_2}{\sqrt{\eta}}, \Phi_2 \right) - E_\rho \left[ \overline{F_{j,k}^\eta} \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \right|^2 \right]. \end{aligned}$$

By the Cauchy-Schwarz inequality on the sum and as  $E[|X - E[X]|^2] \leq E[|X|^2]$ :

$$E_\rho[|F_1|^2] \leq \frac{1}{n^2} \left( \sum_{j,k=0}^{N-1} E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] \right)^2.$$

A direct consequence of Remark 5.2 is

$$E_\rho [|F_1|^2] \leq \begin{cases} \frac{(\mathcal{C}_\infty^\eta)^2}{n^2} N^{-4/3} e^{32\gamma N} & \text{for } \eta \in ]1/2, 1[, \\ \frac{(\mathcal{C} \cdot \mathcal{C}_2)^2}{n^2} N^{17/3} & \text{for } \eta = 1. \end{cases}$$

By noticing  $|\operatorname{Re}(z)| \leq |z|$ , the second term of the sum  $E_\rho [F_2^2]$  is such that

$$\begin{aligned} E_\rho [|F_2|^2] &\leq \frac{4}{n} E_\rho \left[ \left| \sum_{j,k=0}^{N-1} \left( F_{j,k}^\eta \left( \frac{Y_1}{\sqrt{\eta}}, \Phi_1 \right) - E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] \right) \right|^2 \right] \\ &\quad \times E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) - \tau_{j,k} \right|^2 \right] \end{aligned}$$

By the Cauchy-Schwarz inequality on the sum and as  $E[|X - E[X]|^2] \leq E[|X|^2]$ :

$$\begin{aligned} E_\rho [|F_2|^2] &\leq \frac{4}{n} \left( \sum_{j,k=0}^{N-1} E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] \right) \\ &\quad \times \left( \sum_{j,k=0}^{N-1} \left| E_\rho \left[ F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right] - \tau_{j,k} \right|^2 \right) \\ &\leq \frac{4}{n} \|\rho - \tau\|_2^2 \left( \sum_{j,k=0}^{N-1} E_\rho \left[ \left| F_{j,k}^\eta \left( \frac{Y}{\sqrt{\eta}}, \Phi \right) \right|^2 \right] \right) \end{aligned}$$

By Remark 5.2 and as  $\|\rho\|_2^2, \|\tau\|_2^2 \leq 1$  we obtain under  $\rho$  satisfying  $H_1(\mathcal{C}, \varphi_n)$

$$E_\rho [|F_2|^2] \leq \begin{cases} \frac{8\mathcal{C}_\infty^\eta}{n} N^{-2/3} e^{16\gamma N} & \text{for } \eta \in ]1/2, 1[, \\ \frac{8\mathcal{C} \cdot \mathcal{C}_2}{n} N^{17/6} & \text{for } \eta = 1. \end{cases}$$

If,  $N$  is such that the upper bound of  $E_\rho [|F_2|^2]$  tends to 0 with  $n$ , then this is the dominant term in the upper bound of  $V_\rho(M_n)$ . In addition, notice that under  $\rho = \tau$  defined in  $H_0$ , we have  $E_\tau [|F_2|^2] = 0$ .  $\square$

## 5.6 Appendix

The boxplots of the 1 000 values  $\{M_n^k\}_{k=1,\dots,1000}$  of the estimator  $M_n$  (for  $n = 50\,000$ ) in the case **A** and **B** are represented by Figures 5.3 to 5.5 and by Figures 5.6 to 5.8 respectively.

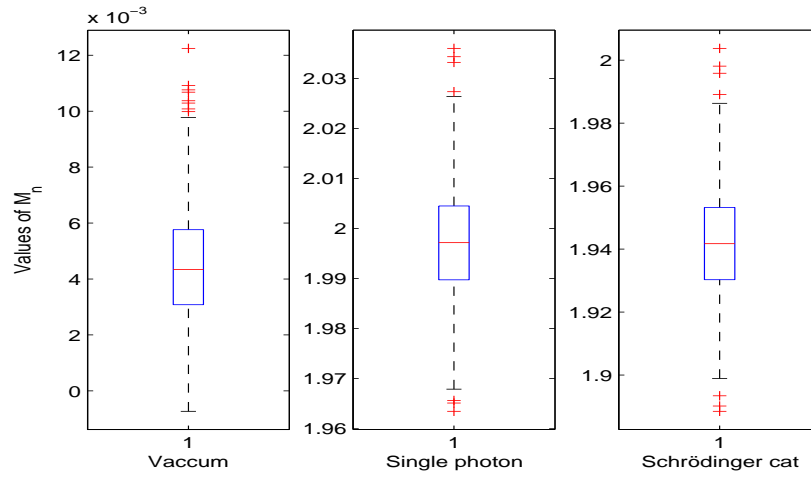


Figure 5.3: For  $\tau$  the *vacuum* state (case **A**): **a**)  $\rho$  the *vacuum* state, **b**)  $\rho$  the *single photon* state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 1$ ,  $N = 15$

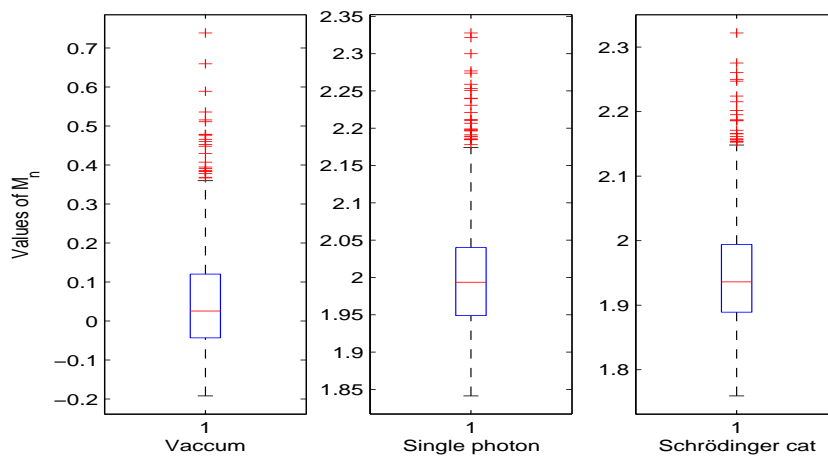


Figure 5.4: For  $\tau$  the *vacuum* state (case **A**): **a**)  $\rho$  the *vacuum* state, **b**)  $\rho$  the *single photon* state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 0.9$ ,  $N = 14$

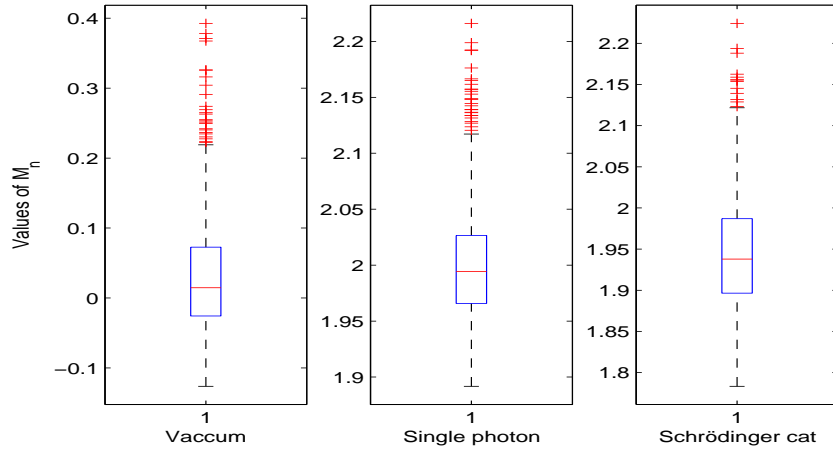


Figure 5.5: For  $\tau$  the *vacuum* state (case **A**): **a**)  $\rho$  the *vacuum* state, **b**)  $\rho$  the *single photon* state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 0.9$ ,  $N = 13$

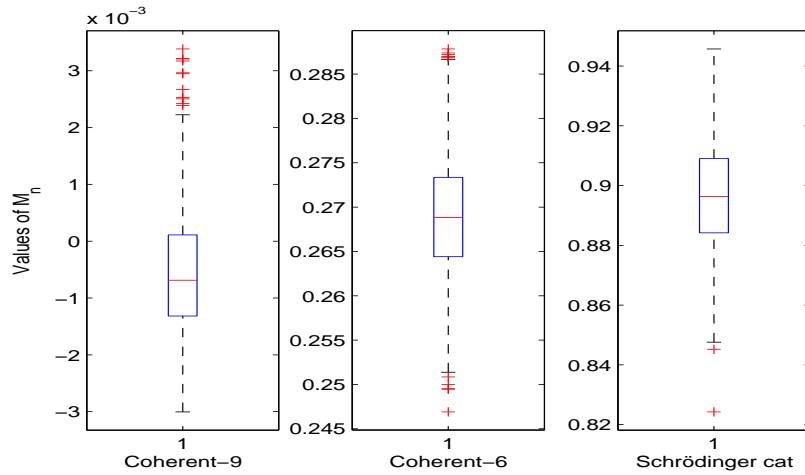


Figure 5.6: For  $\tau$  the *coherent-3* state (case **B**): **a**)  $\rho$  the *coherent-3* state, **b**)  $\rho$  the *coherent- $\sqrt{6}$*  state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 1$ ,  $N = 15$

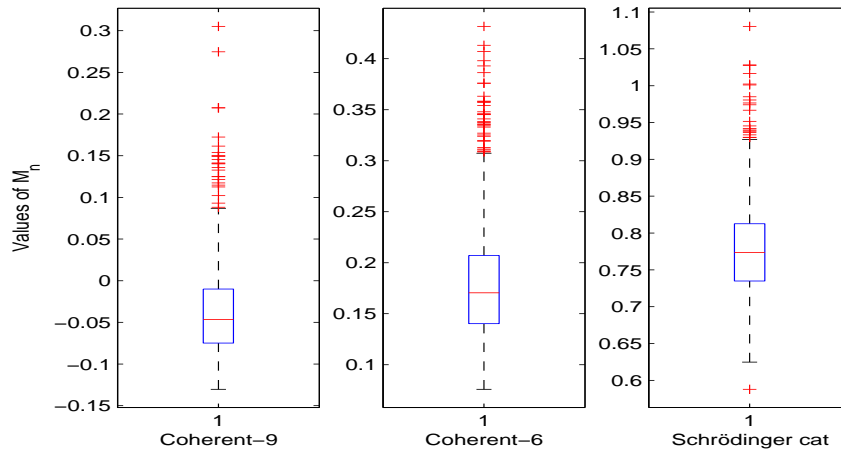


Figure 5.7: For  $\tau$  the *coherent-3* state (case **B**): **a**)  $\rho$  the *coherent-3* state, **b**)  $\rho$  the *coherent- $\sqrt{6}$*  state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 0.9$ ,  $N = 14$

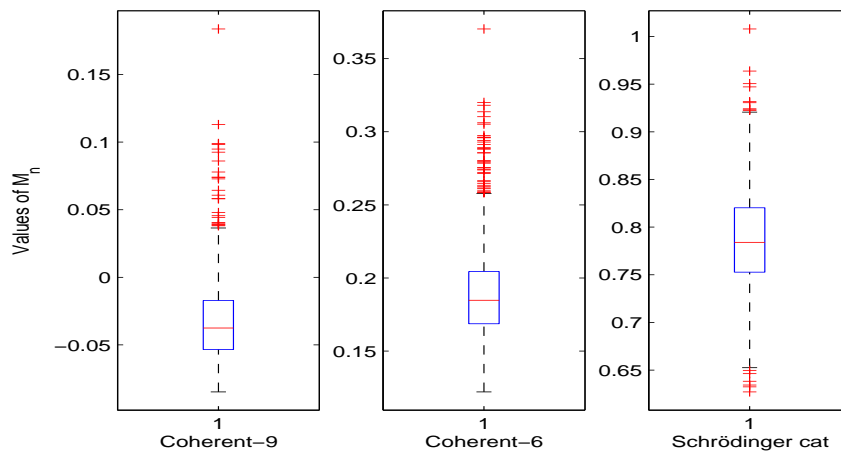


Figure 5.8: For  $\tau$  the *coherent-3* state (case **B**): **a**)  $\rho$  the *coherent-3* state, **b**)  $\rho$  the *coherent- $\sqrt{6}$*  state and **c**)  $\rho$  the *Schrödinger cat-3* state for  $\eta = 0.9$ ,  $N = 13$







# Bibliographie

- [1] D. L. Allen. Hypothesis testing using an  $L_1$ -distance bootstrap. *The American Statistician*, 51(2) :145–150, 1997. MR1450947.
- [2] N. H. Anderson, P. Hall, and D. M. Titterington. Two-sample test statistics for measuring discrepancies between two multivariate probability density functions using kernel-based density estimates. *Journal of Multivariate Analysis*, 50(1) :41–54, 1994. MR1292607.
- [3] L. Aolita. Measuring multipartite concurrence with a single factorizable observable. *Physical Review Letters*, 97(050501), 2006.
- [4] L. Artiles, R. Gill, and M. Guță. An invitation to quantum tomography. *J. Royal Statist. Soc. B (Methodological)*, 67 :109–134, 2005. MR2136642.
- [5] J. M. Aubry. Ultrarapidly decreasing ultradifferentiable functions, Wigner distributions and density matrices. URL <http://arxiv.org/abs/math.FA/0703884>. Submitted to J. London Math. Soc., 2008.
- [6] J. M. Aubry, C. Butucea, and K. Meziani. State estimation in quantum homodyne tomography with noisy data. *Inverse Problems* 25 (2009) 015003., <http://arxiv.org/abs/0804.2434>.
- [7] K. M. R. Audenaert, M. Nussbaum, A. Szkola, and F. Verstraete. Asymptotic error rates in quantum hypothesis testing. *Comm. Math. Phys.*, 279(1) :251–283, 2008. ISSN 0010-3616. MR2377635.
- [8] E. Bagan, M. Ballester, Gill R., A. Monras, and R. Munoz-Tapia. Optimal full estimation of qubit mixed states. *Physical Review*, A 73(032301), 2006.
- [9] 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.

- [10] Y. Baraud. Non asymptotic minimax rates of testing in signal detection. *Bernoulli*, 8, 2002. MR1935648.
- [11] 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.
- [12] P. J. Bickel and Y. Ritov. Estimating integrated squared density derivatives : sharp best order of convergence estimates. *Sankhyā (Statistics). The Indian Journal of Statistics. Series A*, 50(3) :381–393, 1988. ISSN 0581-572X.
- [13] P. J. Bickel and M. Rosenblatt. On some global measures of the deviations of density function estimates. *Annals of Statistics*, 1(6) :1071–1095, 1973. MR0348906.
- [14] L. Birgé and P. Massart. Estimation of integral functionals of a density. *The Annals of Statistics*, 23(1) :11–29, 1995. ISSN 0090-5364.
- [15] G. Breitenbach, S. Schiller, and J. Mlynek. Measurement of the quantum state of the squeezed light. *Nature*, 387 :471–475, 1997.
- [16] C. Butucea. Goodness-of-fit testing and quadratic functional estimation from indirect observations. *The Annals of Statistics*, 35(5) :1907–1930, 2007. ISSN 0090-5364. MR2363957.
- [17] C. Butucea. Deconvolution of supersmooth densities with smooth noise. *Canad. J. Statist.*, 32(2) :181–192, 2004. ISSN 0319-5724.
- [18] C. Butucea and K. Tribouley. Nonparametric homogeneity tests. *Journal of Statistical Planning and Inference*, 136(3) :597–639, 2006.
- [19] C. Butucea and A. B. Tsybakov. Sharp optimality for density deconvolution with dominating bias. i and ii. *Theory Probab. Appl.*, 2007.
- [20] C. Butucea, M. Guță, and L. Artiles. Minimax and adaptive estimation of the Wigner function in quantum homodyne tomography with noisy data. *Ann. Statist.*, 35(2) : 465–494, 2007. MR2336856.
- [21] R. J. Carroll and P. Hall. Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, 83(404) :1184–1186, 1988. ISSN 0162-1459.

- [22] L. Cavalier. Efficient estimation of a density in a problem of tomography. *Ann. Statist.*, 28 :630–647, 2000.
- [23] L. Cavalier. On the problem of local adaptive estimation in tomography. *Bernoulli*, 7 (1) :63–78, 2001. ISSN 1350-7265.
- [24] L. Cavalier, G. K. Golubev, D. Picard, and A. B. Tsybakov. Oracle inequalities for inverse problems. *Ann. Statist.*, 30(3) :843–874, 2002. ISSN 0090-5364. Dedicated to the memory of Lucien Le Cam.
- [25] L. Cavalier, Y. Golubev, O. Lepski, and A. Tsybakov. Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Teor. Veroyatnost. i Primenen.*, 48(3) :534–556, 2003. ISSN 0040-361X.
- [26] G. M. D’Ariano. Tomographic measurement of the density matrix of the radiation field. *Quantum Semiclass. Optics*, 7 :693–704, 1995.
- [27] G. M. D’Ariano. Tomographic methods for universal estimation in quantum optics. In *International School of Physics Enrico Fermi*, volume 148. IOS Press, 2002.
- [28] G. M. D’Ariano and M. G. A. Paris. Adaptive quantum tomography. *Physical Review A*, 60(518), 1999.
- [29] 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.
- [30] 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.
- [31] G. M. D’Ariano, J. Kahn, and M. F. Sacchi. Minimax discrimination of two pauli channels. *Phys. Rev. A*, 72 :052302, 2005.
- [32] G. M. D’Ariano, J. Kahn, and M. F. Sacchi. Minimax quantum state discrimination. *Phys. Rev. A*, 72 :032310, 2005.
- [33] G.M. D’Ariano, L. Maccone, and M. F. Sacchi. Homodyne tomography and the reconstruction of quantum states of light, 2005. URL [arXiv.org](http://arXiv.org) :quant-ph/0507078.

- [34] D. L. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.*, 2(2) :101–126, 1995. ISSN 1063-5203.
- [35] S. Efromovich and M. Low. On optimal adaptive estimation of a quadratic functional. *The Annals of Statistics*, 24(3) :1106–1125, 1996. ISSN 0090-5364.
- [36] J. Fan. Global behavior of deconvolution kernel estimates. *Statist. Sinica*, 1(2) :541–551, 1991. ISSN 1017-0405.
- [37] M. Fromont and B. Laurent. Adaptive goodness-of-fit tests in a density model. *Annals of Statistics*, 34(2), 2006. MR2281881.
- [38] G. Gayraud and C. Pouet. Minimax testing composite null hypotheses in the discrete regression scheme. *Math. Methods Statist.*, 10(4) :375–394 (2002), 2001. ISSN 1066-5307. Meeting on Mathematical Statistics (Marseille, 2000).
- [39] G. Gayraud and C. Pouet. Adaptive minimax testing in the discrete regression scheme. *Probab. Theory Related Fields*, 133(4) :531–558, 2005. ISSN 0178-8051.
- [40] A. Goldenshluger. On pointwise adaptive nonparametric deconvolution. *Bernoulli*, 5 (5) :907–925, 1999. ISSN 1350-7265.
- [41] A. Goldenshluger and V. Spokoiny. On the shape-from-moments problem and recovering edges from noisy Radon data. *Probab. Theory Related Fields*, 128(1) :123–140, 2004. ISSN 0178-8051.
- [42] G. Golubev and R. Khasminskii. A statistical approach to the Cauchy problem for the Laplace equation. In *State of the art in probability and statistics (Leiden, 1999)*, volume 36 of *IMS Lecture Notes Monogr. Ser.*, pages 419–433. Inst. Math. Statist., Beachwood, OH, 2001.
- [43] M. Guță and L. Artiles. Minimax estimation of the Wigner in quantum homodyne tomography with ideal detectors. *Math. Methods Statist.*, 16(1) :1–15, 2007. MR2319467.
- [44] M. Guță and J. Kahn. Local asymptotic normality for qubit states. *Phys. Rev. A (3)*, 73(5) :052108, 15, 2006. ISSN 1050-2947.
- [45] M. I. Guță. Maximum likelihood estimation of the density matrix through quantum tomography. Manuscript, 2007.

- [46] C. W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1 :231–252, 1969.
- [47] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland, 1982.
- [48] P. Horodecki. Measuring quantum entanglement without prior state reconstruction. *Physical Review Letters*, 90(167901), 2003.
- [49] Y. I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. I. *Math. Methods Statist.*, 2(2) :85–114, 1993. ISSN 1066-5307.
- [50] Y. I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. II. *Math. Methods Statist.*, 2(3) :171–189, 1993. ISSN 1066-5307.
- [51] Y. I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. III. *Math. Methods Statist.*, 2(4) :249–268, 1993. ISSN 1066-5307.
- [52] Y. I. Ingster and I. A. Suslina. *Nonparametric Goodness-of-fit Testing under Gaussian Models*. Lecture Notes in Statistics. Springer, 2003. MR1991446.
- [53] I. M. Johnstone and B. W. Silverman. Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.*, 18(1) :251–280, 1990.
- [54] I. M. Johnstone, G. Kerkycharian, D. Picard, and M. Raimondo. Wavelet deconvolution in a periodic setting. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(3) :547–573, 2004. ISSN 1369-7412.
- [55] J. Kahn. Model selection for quantum homodyne tomography. *to appear ESAIM : Probability and Statistics*, 2008. <http://arxiv.org/abs/0712.2912>.
- [56] V. Kargin. On the Chernoff bound for efficiency of quantum hypothesis testing. *Ann. Statist.*, 33(2) :959–976, 2005. ISSN 0090-5364.
- [57] G. Kerkycharian and D. Picard. Estimating nonquadratic functionals of a density using Haar wavelets. *The Annals of Statistics*, 24(2) :485–507, 1996. ISSN 0090-5364.
- [58] G. Kerkycharian, P. Petrushev, D. Picard, and T. Willer. Needlet algorithms for estimation in inverse problems. *Electron. J. Stat.*, 1 :30–76 (electronic), 2007. ISSN 1935-7524.

- [59] G. Kerkyacharian, G. Kyriazis, E. Le Pennec, P. Petrushev, and D. Picard. Inversion of noisy radon transform by svd based needlet. <http://arxiv.org/abs/0809.3332>, 2008.
- [60] J. Klemelä. Sharp adaptive estimation of quadratic functionals. *Probability Theory Related Fields*, 134(4) :539–564, 2006. ISSN 0178-8051.
- [61] A. P. Korostelëv and A. B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1993.
- [62] I. Krasikov. Inequalities for Laguerre polynomials. *East J. Approx.*, 11(3) :257–268, 2005.
- [63] I. Krasikov. Inequalities for orthonormal Laguerre polynomials. *J. Approx. Theory*, 144(1) :1–26, 2007. URL <http://dx.doi.org/10.1016/j.jat.2006.04.005>.
- [64] C. Lacour. Rates of convergence for nonparametric deconvolution. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 342(11) :877–882, 2006. ISSN 1631-073X.
- [65] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, 1997.
- [66] 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.
- [67] U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. G. Raymer. Sampling of photon statistics and density matrix using homodyne detection. *Opt. Commun.*, 127 :144–160, 1996.
- [68] O. Lepski, A. Nemirovski, and V. Spokoiny. On estimation of the  $L_r$  norm of a regression function. *Probability Theory Related Fields*, 113(2) :221–253, 1999. ISSN 0178-8051.
- [69] O. V. Lepski and A. B. Tsybakov. Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. *Probab. Theory Related Fields*, 117(1) :17–48, 2000. ISSN 0178-8051. MR1759508.
- [70] B. Maurey. Isomorphismes entre espaces  $H_1$ . *Acta Math.*, 145(1-2) :79–120, 1980. ISSN 0001-5962.
- [71] K. Meziani. Nonparametric estimation of the purity of a quantum state in quantum homodyne tomography with noisy data. *Mathematical Methods of Statistics*, 16(4) : 1–15, 2007. ISSN 1066-5307. MR2378280.



- [72] K. Meziani. Nonparametric goodness-of fit testing in quantum homodyne tomography with noisy data. *Electron. J. Stat.*, 2 :1195–1223, 2008.
- [73] Y. Nikitin. *Asymptotic Efficiency of Nonparametric Tests*. Cambridge University Press, 1995. MR1335235.
- [74] M. Nussbaum and A. Szkola. A lower bound of Chernoff type for symmetric quantum hypothesis testing. *arXiv :quant-ph/0607216v1*, 2008. to appear in *Ann. Statist.*
- [75] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier. Generating optical Schrödinger kittens for quantum information processing. *Science*, 312(5770) :83–86, 2006.
- [76] C. Pouet. An asymptotically optimal test for a parametric set of regression functions against a non-parametric alternative. *J. Statist. Plann. Inference*, 98(1-2) :177–189, 2001. ISSN 0378-3758.
- [77] C. Pouet. On testing nonparametric hypotheses for analytic regression functions in Gaussian noise. *Math. Methods Statist.*, 8(4) :536–549 (2000), 1999. ISSN 1066-5307. MR1755899.
- [78] T. Richter. Pattern functions used in tomographic reconstruction of photon statistics revisited. *Phys. Lett. A*, 211 :327–330, 1996.
- [79] T. Richter. Realistic pattern functions for optical homodyne tomography and determination of specific expectation values. *Physical Review A*, 61(063819), 2000.
- [80] P. Rigollet. Adaptive density estimation using the blockwise Stein method. *Bernoulli*, 12(2) :351–370, 2006. ISSN 1350-7265.
- [81] D. J. Smithey, M. Beck, M. J. Cooper, and M. G. Raymer. Experimental determination of number-phase uncertainty relations. *Optics Letters*, 18 :1259–1261, 1993.
- [82] 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.
- [83] V. G. Spokoiny. Adaptive hypothesis testing using wavelets. *Annals of Statistics*, 24 (6), 1996. MR1425962.

- [84] V. G. Spokoiny. Adaptive and spatially adaptive testing of a nonparametric hypothesis. *Math. Methods Statist.*, 7(3) :245–273, 1998. ISSN 1066-5307.
- [85] L. Stefanski and R. J. Carroll. Deconvoluting kernel density estimators. *Statistics*, 21(2) :169–184, 1990. ISSN 0233-1888.
- [86] G. Szegő. *Orthogonal polynomials*. American Mathematical Society Colloquium Publications, Vol. 23. Revised ed. American Mathematical Society, Providence, R.I., 1959.
- [87] K. Tribouley. Adaptive estimation of integrated functionals. *Mathematical Methods of Statistics*, 2000.
- [88] A. Tsybakov. On the best rate of adaptive estimation in some inverse problems. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(9) :835–840, 2000. ISSN 0764-4442.
- [89] Alexandre B. Tsybakov. *Introduction à l'estimation non-paramétrique*, volume 41 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2004. ISBN 3-540-40592-5.
- [90] Y. Vardi, L. A. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *J. Am. Stat. Assoc.*, 80 :8–37, 1985.
- [91] 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.
- [92] E. Wigner. On the quantum correction for thermodynamic equations. *Phys. Rev.*, 40 : 749–759, 1932.
- [93] C. Zhang, G. Wang, and M. Ying. Discrimination between pure states and mixed states. *Physical Review*, A 75(062306), 2007.





**RÉSUMÉ** : En optique quantique, la reconstruction de l'état quantique (fonction de Wigner ou matrice de densité infini-dimensionnelle) d'un faisceau de lumière correspond en statistique à un problème inverse très mal posé. Premièrement, nous proposons des estimateurs de la matrice de densité basés sur les fonctions *pattern* et des estimateurs à noyau de la fonction de Wigner. Nous faisons l'hypothèse que la matrice de densité inconnue appartient à une classe non paramétrique définie en accord avec les exemples étudiés par les physiciens. Nous en déduisons pour la fonction de Wigner associée à cette matrice des propriétés de décroissance rapide et de régularité. Deuxièmement, nous estimons une fonctionnelle quadratique de la fonction de Wigner par une U-statistique d'ordre deux sur une classe plus large. Cette fonctionnelle peut être vue comme une indication sur la pureté de l'état quantique considéré. Nous en déduisons un estimateur adaptatif aux paramètres de régularité de la fonction de Wigner. La dernière partie de ce manuscrit est consacrée au problème de test d'adéquation à la matrice de densité. Cette procédure est construite à partir d'un estimateur de type projection sur les fonctions *pattern*. Nous étudions les bornes supérieures de type minimax de toutes ces procédures. Les procédures d'estimation de la matrice de densité et de test d'adéquation à une matrice de densité sont implémentées et leurs performances numériques sont étudiées.

**MOTS-CLÉS** : Déconvolution, Estimation adaptative, Estimation de fonctionnelle quadratique, Estimation non paramétrique, Etats quantiques, Fonction de Wigner, Fonction pattern, Fonctions infiniment différentiables, Matrice de densité, Problème inverse, Risque minimax, Tests d'adéquation, Tomographie quantique homodyne, Transformée de Radon, Vitesses de convergence, Vitesses de test.

**DISCIPLINE** : MATHÉMATIQUES

**ABSTRACT** : In the setting of quantum optics, the reconstruction of the quantum state (Wigner function or infinite-dimensional density matrix) of a light beam can be seen as a statistical severely ill-posed inverse problem. First, we propose estimators of the density matrix and the Wigner function respectively, using *pattern* functions in the first case and kernel functions in the second. We assume that the unknown density matrix belongs to a nonparametric class which corresponds to typical states prepared in the laboratory. We translate these classes in terms of properties of the associated Wigner function. Second, we estimate the integrated squared Wigner function by a kernel-based second order U-statistic on a larger regularity class. This quadratic functional is a physical measure of the purity of the state. We deduce an adaptive estimator for the Wigner function that does not depend on the smoothness parameters. In the last part of the thesis, we are interested in the problem of goodness-of-fit testing. We give a testing procedure derived from a projection-type estimator on *pattern* functions. We study the upper bounds of the minimax risk for all our procedures. The density matrix estimation and the testing procedure are implemented and their numerical performances are studied.

**KEY WORDS** : Adaptive estimation, Deconvolution, Density matrix, Goodness-of-fit testing, Infinitely differentiable functions, Inverse problem, Minimax risk, Nonparametric estimation, Pattern function, Quadratic functional estimation, Quantum homodyne tomography, Quantum state, Radon transform, Rates of convergence, Testing rates, Wigner function.

**Laboratoire de Probabilités et Modèles Aléatoires,  
CNRS-UMR 7599, UFR de Mathématiques, case 7012  
Université Paris 7, Denis Diderot  
2, place Jussieu, 75251 Paris Cedex 05.**