Dynamic identification for bi-linear and non-linear systems in presence of uncertainties

Ying Fu

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Identification de dynamique pour des systèmes bilinéaires et non-linéaires en présence d'incertitudes

École Doctorale de Dauphine — ED 543

Spécialité **Sciences**

COMPOSITION DU JURY :

M. MAZYAR Mirrahimi
INRIA Paris-Rocquencourt
Rapporteur

M. SUGNY Dominique
Université de Bourgogne
Rapporteur

M. LE BRIS Claude
Ecole Nationale de Ponts et Chaussées
Président du jury

M. SALOMON Julien
Université Paris Dauphine
Membre du jury

M. TURINICI Gabriel
Université Paris Dauphine
Directeur de thèse

Soutenue le 09/12/2016
par FU Ying

Dirigée par **TURINICI Gabriel**

Université Paris Dauphine

DAUPHINE UNIVERSITÉ PARIS | PSL RESEARCH UNIVERSITY PARIS

Thèse effectuée au sein du **Laboratoire CEREMADE**

de l'Université Paris-Dauphine
Place du Maréchal de Lattre de Tassigny
75016 Paris
France

Résumé

Dans le cadre du contrôle quantique bilinéaire, cette thèse étudie la possibilité de retrouver l'Hamiltonien et/ou le moment dipolaire à l'aide de mesures d'observables pour un ensemble grand de contrôles. Si l'implémentation du contrôle fait intervenir des bruits alors les mesures prennent la forme de distributions de probabilité. Nous montrons qu'il y a toujours unicité (à des phases près) des Hamiltoniens de du moment dipolaire retrouvés. Plusieurs modèles de bruit sont étudiés: bruit discrète constant additif et multiplicatif ainsi qu'un modèle de bruit dans les phases sous forme de processus Gaussien. Les résultats théoriques sont illustrés par des implémentations numériques.

Mots-clé:

équation Schrödinger, système bilinéaire, contrôle quantique, identification Hamiltonien

Dynamic identification for bi-linear and non-linear systems in presence of uncertainties

Abstract

The problem of recovering the Hamiltonian and dipole moment, termed inversion, is considered in a bilinear quantum control framework. The process uses as inputs some measurable quantities (observables) for each admissible control. If the implementation of the control is noisy the data available is only in the form of probability laws of the measured observable. Nevertheless it is proved that the inversion process still has unique solutions (up to phase factors). Several models of noise are considered including the discrete noise model, the multiplicative amplitude noise model and a Gaussian process phase model. Both theoretical and numerical results are established.

Keywords:

Schrödinger equation, bi-linear system, quantum control, inversion Hamiltonian

Acknowledgements

This thesis is written based on a 3 years professional work, with the kind help and advises of a number of individuals.

First of all, let me extend my thanks to M. MAZYAR Mirrahimi and M. SUGNY Dominique, for writing a report to this work. I am also grateful to M. LE BRIS Claude and M. SALOMON Julien for their participations as members of examination committee.

My most sincere gratitude is to my advisor M. TURINICI Gabriel for his expertise, understanding, patient and generous guidance, encouragement and support. I am very glad to finish several works and this thesis with him.

I am grateful to Université Paris Dauphine and the laboratory CEREMADE for the financial support of this thesis. Thanks for giving me the chance to participate in several meetings and congress. I would also like to thank the members of CEREMADE for the weekly seminars and frequent discussions.

Special thanks to the project EMAQS and all the participants of this project for exchanging the newest results and technologies in quantum control theories. I am also highly indebted to M. Rabitz Herschel and his research group for sharing their knowledge in physics.

Last but not the least, thanks to my parents and my friends for their support.

Contents

Introduction

In this thesis, the bi-linear system we study is the Schrödinger equation. Quantum system inversion concerns learning the characteristics of the underlying Hamiltonian by measuring suitable observables from the responses of the system's interaction with members of a set of applied fields. Various aspects of inversion have been confirmed in theoretical, numerical and experimental works. The theoretical part of the thesis addresses the uniqueness, and the numerical part consists in the inversion and recovery problem. Various noise models are considered. The plan of the thesis is the following.

The first chapter of the thesis introduces the mathematical formulation of the problem. We start with some examples of applications of quantum mechanics. Then we introduce the basic elements of quantum theory: the state vectors, the observables, and the density operators. Once these notations are defined, we can write the Schrödinger equation. We continue by giving the noise model appearing in the interaction between the electromagnetic field and a quantum system. We go on with a brief introduction of the Lie group and Lie algebra, which are essential in the study of controllability of the Schrödinger equation. The last part of this chapter concerns the distances between the probability distributions and Gaussian process theory which is crucial in numerical simulations.

The second chapter reproduces the content of the accepted paper [18]. In this section, the noise is modeled as a discrete random variable. We

first improve the hypotheses of uniqueness result in [29]. Then we prove that in presence of a discrete noise, the Hamiltonian H and the dipole moment μ of the system are unique within some phases. Numerical simulations in the case of additive noise shows that the we can recovery the H and μ from experimental data.

The third chapter reproduces the content of the accepted paper [19]. In this chapter, we consider the amplitude noise model. The noise is supposed to be multiplicative and identical for all frequencies. In contrast to the previous chapter, the noise is supposed to be unknown. Theoretical results and numerical simulations show that not only the dipole moment but also the noise distribution are unique within some factors and phases and can be recovered.

The fourth chapter discuss the phase noise model. The noises related to different frequencies are correlated and are modeled by Gaussian process in the frequency space. Numerical tests are implemented for the square exponential covariance model, the Ornstein-Uhlenbeck process and the Brownian motion.

Résumé de la thèse

Ce chapitre est un résumé en français de cette thèse. Le plan et les principaux résultats de cette thèse sont présentés dans ce résumé. Chapitre 1 introduit les ingrédients pour la formulation mathématique du problème étudié. Chapitre 2, 3 et 4 présentent les résultats théoriques et numériques avec les modèles de bruit dans les trois cas: variable aléatoire discrète, multiplicative de loi inconnue sur les amplitudes et processus Gaussian sur les phases respectivement.

Nous présentons par la suite les principaux résultats de cette thèse.

0.1 Chapitre 1: formulation mathématique du problème

Le chapitre 1 consiste en une introduction aux formulations mathématiques des problèmes étudiés dans cette thèse. On commence par donner des exemples d'applications en mécanique quantique. Ensuite on introduit les ingrédients essentiels de la mécanique quantique: les états, les observables et les matrices de densité. On continue avec l'équation de Schrödinger et modélisations du bruit apparaissant lors de l'intéraction entre un champ électromagnétique et un système quantique. La présentation se poursuit avec une brève introduction aux algèbres et groupes de Lie qui seront invoqués lors de l'étude des résultats de contrôlabilité de l'équation de Schrödinger. A la fin on donne quelques outils concernant les distances entre des lois de probabilité, ainsi que des rudiments de la théorie des processus gaussiens, qui seront nécessaires pour la formulation de nos simulations numériques.

0.2 Notations

Les notations suivantes sont utilsées dans les Chapitres 2, 3 et 4:

- $\mathbb{L}_{M_1,M_2,\cdots,M_m}$ est l'algèbre de Lie généréé par les matrices $M_1, M_2, \cdots, M_m$;
- pour tout matrice ou vecteur X, on note par $\overline{X}$ son conjugué et X^* la matrice adjointe;
- $\mathfrak{H}$ est l'ensemble de matrices Hermitiennes. $\mathfrak{H} = \{X \in \mathbb{C}^{N\times N} | X^* = X\}$;
- $\mathcal{S}_N$ est la sphére unitaire de $\mathbb{C}^N$: $\mathcal{S}_N = \{v \in \mathbb{C}^N | \|v\| = 1\}$;
- $\Psi(t, H, u(\cdot), \mu, \Psi_0)$ est la solution de la équation (3), pour simplifier la notation, on la note $\Psi(t)$;
- $\lambda_k(X)$, $k = 1, ..., N$ sont les valeurs propres de $X \in \mathfrak{H}$ dans l'ordre croissant; on introduit aussi $\phi_k(X)$ les vecteurs propres de X pour $k = 1, ..., N$ (qui forment une base orthonormée de $\mathbb{C}^N$) correspondent aux valeurs propres $\lambda_k(X)$; le choix n'est pas unique;
- $SU(N)$ est le group special unitaire de degré N, qui est l'ensemble de matrices unitaires de tailles $N \times N$ avec determinant 1;
- $\mathfrak{su}(N)$ est l'algèbre de Lie du $SU(N)$;

0.3 Résultats de contrôlabilité

Dans la suite on introduit quelques résultats de contrôlabilité.

Soit $L \in \mathbb{N}^*$ et L groupes de Lie $G_1, \cdots, G_L$ de dimension finie, connexes et compacts avec élément neutre Id. Soient $A_\ell, B_\ell \in \mathfrak{g}_\ell$ pour tout $\ell = 1, \cdots, L$ avec $\mathfrak{g}_\ell$ L'algèbre de Lie associée à G_ℓ.

Definition 0.1 *Soient L systèmes bilinéaires:*

$$\begin{cases} \frac{dX_\ell(t)}{dt} = (A_\ell + u(t)B_\ell)X_\ell(t), \quad X_\ell \in G_\ell \\ X_\ell(0) = Id \in G_\ell. \end{cases} \tag{1}$$

Les systèmes sont dites simultanément contrôlables s'il existe $T_{A_1,\cdots,A_L,B_1,\cdots,B_L} > 0$ tel que pour tout $T \geq T_{A_1,\cdots,A_L,B_1,\cdots,B_L}$ et pour tout $V_\ell \in G_\ell$, $\ell = 1, \cdots, L$ arbitairement choisis, il existe un contrôle $u \in \mathrm{L}^1([0,T], \mathbb{R})$ avec $X_\ell(T) = V_\ell$, $\forall \ell = 1, \cdots, L$.

Soient $\mathcal{A} = A_1 \oplus \cdots \oplus A_L \in \oplus_{\ell=1}^L \mathfrak{g}_\ell$ et $\mathcal{B} = B_1 \oplus \cdots \oplus B_L \in \oplus_{\ell=1}^L \mathfrak{g}_\ell$.

Theorem 0.1 *La collection (1) de L systèmes bilinéaires est simultanément contrôlable si et seulement si $\mathbb{L}_{\mathcal{A},\mathcal{B}} = \oplus_{\ell=1}^L \mathfrak{g}_\ell$ (ou $dim_{\mathbb{R}}\mathbb{L}_{\mathcal{A},\mathcal{B}} = \sum_{\ell=1}^L dim_{\mathbb{R}}\mathfrak{g}_\ell$.*

Lemma 0.1 *On suppose que $\mathbb{L}_{A_\ell,B_\ell} = \mathfrak{g}_\ell$ pour tout $\ell = 1, \cdots, L$. Alors $\mathbb{L}_{\mathcal{A},\mathcal{B}} \neq \oplus_{\ell=1}^L \mathfrak{g}_\ell$ si et seulement si il existent $\ell, \ell' \in \{1, \cdots, L\}$, $\ell \neq \ell'$ et un isormorphisme $f : \mathfrak{g}_\ell \to \mathfrak{g}_{\ell'}$ tels que $f(A_\ell) = A_{\ell'}$ et $f(B_\ell) = B_{\ell'}$.*

Theorem 0.2 *Soit G un groupe de Lie de dimension finie, connexe, compact et simple et $\mathfrak{g}$ l'algèbre de Lie associée. Soient $A, B \in \mathfrak{g}$ tels que $\mathbb{L}_{A,B} = \mathfrak{g}$ et α_1,...,$\alpha_L \in \mathbb{R}$ sont les réels, $\alpha_i \neq \alpha_j$ $\forall i \neq j$. On considère la collection de systèmes en G:*

$$\begin{cases} \frac{dX_\ell(t)}{dt} = \{A + (u(t) + \alpha_\ell)B\}X_\ell(t), \\ X_\ell(0) = Id. \end{cases} \tag{2}$$

Alors la collection du systèmes (2) est simultanément contrôlable.

0.4 Un exemple de système à nombre fini de niveaux

Pour les simulations numériques intervenant dans cette thèse nous allons utiliser l'équation de Schrödinger:

$$\begin{cases} i\dot{\Psi}(t, H, u(\cdot), \mu, \Psi_0) = (H + u(t)\mu)\Psi(t, H, u(\cdot), \mu, \Psi_0) \\ \Psi(0, H, u(\cdot), \mu, \Psi_0) = \Psi_0, \end{cases} \tag{3}$$

Le choix d'opérateurs H et μ suivra l'exemple introduit dans [14]:

$$H_{réel} = \begin{pmatrix} 0.0833 & -0.0038 & -0.0087 & 0.0041 \\ -0.0038 & 0.0647 & 0.0083 & 0.0038 \\ -0.0087 & 0.0083 & 0.0036 & -0.0076 \\ 0.0041 & 0.0038 & -0.0076 & 0.0357 \end{pmatrix}, \tag{4}$$

$$\mu_{réel} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 7 \\ 0 & -1.5 & 7 & 0 \end{pmatrix}. \tag{5}$$

Pour rappel, sur un tel système il est généralement possible de mesurer des observables du type:

$$\langle O_j \rangle(t) = \langle \Psi(0, H, u(\cdot), \mu, \Psi_0) | O_j | \Psi(0, H, u(\cdot), \mu, \Psi_0) \rangle, \tag{6}$$

pour certains opérateurs $O_j \in \mathfrak{H}$, $j = 1, \cdots$.

0.5 Formulation d'un problème d'identification

Lorsque dans l'équation (3) nous connaissons H, μ et Ψ_0, nous pouvons calculer les observables $\langle O_j \rangle(t)$, $O_j \in \mathfrak{H}$, $j = 1, \cdots$. Le but de cette thèse est la formulation et l'étude de faisabilité théorique et numérique du problème, dit inverse, suivant: connaissant quelques $\langle O_j \rangle(t)$ pour certains $O_j \in \mathfrak{H}$ et $t \geq 0$, trouver les opérateurs μ et/ou H (et occassionement Ψ_0). Le cas sans bruit a été traité dans [29], nous allons ainsi nous focaliser sur le cas avec bruit.

0.6 Chapitre 2: contrôle perturbé par une variable aléatoire discrète

Dans ce chapitre, le contrôle est supposé perturbé par une variable aléatoire discrète, notée Y. On suppose que Y est indépendante du temps. Le contrôle est sous la forme de $u(t, \epsilon(\cdot), Y)$.

Le problème d'identification se pose dans deux cas:

- Cas **(S1)**: l'Hamiltonien H est connu et on veut identifier le moment dipolaire μ.
- Cas **(S2)**: ni l'Hamiltonien H ni le moment dipolaire μ ne sont connus.

On commence par rappeler la définition d'ensemble complet d'observables qui commutent (ECOC).
Un ensemble d'observables $\mathcal{O} = \{O_1, ..., O_K\}$ est dit *ensemble d'observables qui commutent* (EOC) si $[O_k, O_\ell] = 0$, $\forall k, \ell \in \{1, ..., K\}$.
Toutes les observables dans un EOC $\mathcal{O}$ sont co-diagonalisables. C'est-à-dire, il existe au moins une base orthonormée $\Phi = \{\phi_1, ..., \phi_N\}$ de $\mathbb{C}^N$ telle que tout $O \in \mathcal{O}$ est diagonale dans la base Φ. Un EOC est dit *ensemble complet d'observables qui commutent* (ECOC) si la base orthonormée qui diagonalise le EOC est unique à des phases et permutation près.

Le lemme suivant donne une propriété importante de ECOC.

Lemme 1 *Soit $\mathcal{O} = \{O_1, ..., O_K\}$ un EOC. Alors $\mathcal{O}$ est un ECOC si et seulement si il existent $\gamma_1, ..., \gamma_K \in \mathbb{R}$ tels que tous les valeurs propres de $\sum_{k=1}^{K} \gamma_k O_k$ soient de multiplicité 1.*

0.6.1 Inversion sans perturbation

Le problème est d'abord traité dans le cas sans perturbation, c'est-à-dire $Y = 0$. Nous améliorons ainsi les résultat dans [29].

Théorème 1 (Cas (S1)) *Soient H, μ_1, $\mu_2 \in \mathfrak{H}$, H diagonales, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$. Pour $a = 1, 2$ et $\epsilon \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$, on note $\Psi_a(t, \epsilon) = \Psi(t, H, \epsilon(\cdot), \mu_a, \Psi_0^a)$. Soit $\mathcal{O}$ un EOC non trivial. On suppose que $N \geq 3$ et*

- ***(H1):*** $\mathbb{L}_{iH, i\mu_1} = \mathbb{L}_{iH, i\mu_2} = \mathfrak{su}(N)$.
- ***(H2):*** $tr(H) = tr(\mu_1) = tr(\mu_2) = 0$.
- *les valeurs propres de H sont de multiplicité 1.*

Alors il existe $T > 0$ tel que si

$$\begin{aligned} &\langle O\Psi_1(T, \epsilon), \Psi_1(T, \epsilon)\rangle = \langle O\Psi_2(T, \epsilon), \Psi_2(T, \epsilon)\rangle \\ &\forall \epsilon \in L^1([0, T]; \mathbb{R}), \quad \forall O \in \mathcal{O}, \end{aligned} \tag{7}$$

alors il existent des phases $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$ telles que:

$$(\mu_1)_{jk} = e^{i(\alpha_j - \alpha_k)}(\mu_2)_{jk}, \ \forall j, k \leq N. \tag{8}$$

L'hypothèse **(H1)** est nécessaire pour la contrôlabilité simultanée. Par contre on peut poser l'hypothèse **(H2)** sans perte de généralité.
L'Hamiltonian H et le moment dipolaire μ ne sont identifiables qu'aux phases près. Pour contre-exemple, voir Remarque 2.3, page 79.

Quand les valeurs propres de H sont non dégénérées mais $\mathcal{O}$ est un ECOC, c'est un cas particulier du théorème suivant:

Théorème 2 (Cas (S2)) *Soient $\mu_1, \mu_2, H_1, H_2 \in \mathfrak{H}$, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$. Pour $a = 1, 2$ et $\epsilon \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$, on note $\Psi_a(t, \epsilon) = \Psi(t, H_a, \epsilon(\cdot), \mu_a, \Psi_0^a)$. Soit*

$\mathcal{O} = \{O_1, ..., O_K\}$ un ECOC et $\Phi = \{\phi_1, ..., \phi_N\}$ une base orthonormée qui diagonalise $\mathcal{O}$. On suppose que $N \geq 3$ et que les hypothèses suivants sont vraies:

(H1'): $\mathbb{L}_{iH_1, i\mu_1} = \mathbb{L}_{iH_2, i\mu_2} = \mathfrak{su}(N)$;

(H2'): $tr(H_1) = tr(H_2) = tr(\mu_1) = tr(\mu_2) = 0$;

Alors il existe $T > 0$ tel que si

$$\begin{aligned} &\langle O_k \Psi_1(T, \epsilon), \Psi_1(T, \epsilon)\rangle = \langle O_k \Psi_2(T, \epsilon), \Psi_2(T, \epsilon)\rangle \\ &\forall \epsilon \in L^1([0, T]; \mathbb{R}), \quad \forall k = 1, \cdots, K, \end{aligned} \tag{9}$$

alors il existent des phases $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$ et $\theta \in \mathbb{R}$ telles que pour tout $j, k \leq N$, soit

$$\begin{cases} \langle \mu_1 \phi_j, \phi_k\rangle = e^{i(\alpha_j - \alpha_k)} \langle \mu_2 \phi_j, \phi_k\rangle \\ \langle H_1 \phi_j, \phi_k\rangle = e^{i(\alpha_j - \alpha_k)} \langle H_2 \phi_j, \phi_k\rangle \\ \langle \Psi_0^1, \phi_j\rangle = e^{i(\theta - \alpha_j)} \langle \Psi_0^2, \phi_j\rangle, \end{cases} \tag{10}$$

soit

$$\begin{cases} \langle \mu_1 \phi_j, \phi_k\rangle = -e^{i(\alpha_j - \alpha_k)} \overline{\langle \mu_2 \phi_j, \phi_k\rangle} \\ \langle H_1 \phi_j, \phi_k\rangle = -e^{i(\alpha_j - \alpha_k)} \overline{\langle H_2 \phi_j, \phi_k\rangle} \\ \langle \Psi_0^1, \phi_j\rangle = e^{i(\theta - \alpha_j)} \overline{\langle \Psi_0^2, \phi_j\rangle}. \end{cases} \tag{11}$$

Quand $\mathcal{O}$ n'est pas un ECOC, la preuve permet d'obtenir l'existence d'un isomorphisme d'algèbres de Lie qui envoie iH_1 en iH_2 et $i\mu_1$ en $i\mu_2$.

0.6.2 Inversion en présence de perturbation de loi connue

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité discret, $\mathcal{V} = \{y_\ell \in \mathbb{R}^d | \ell \in \mathcal{I} \subset \mathbb{N}\}$ un ensemble inclus dans $\mathbb{R}^d$ (qui peut être infini). La perturbation est modélisée par une variable aléatoire discrète $Y : \Omega \to \mathcal{V}$.

On peut supposer que pour tout $y_\ell \in \mathcal{V}$, $\mathbb{P}(Y = y_\ell) > 0$ et $\mathcal{I} = \mathbb{N}^*$ ou $\mathcal{I} = \{1, \cdots, L_0\}$ avec $L_0 \in \mathbb{N}^*$. On note $\xi_k = \mathbb{P}(Y = y_k)$, $\forall k \in \mathcal{I}$. On suppose que $(\xi_\ell)_{\ell \geq 1}$ est une suite décroissante en permutant les indices. En particulier la loi de Y est complètement donnée par la connaissance de $\mathcal{V}$ et des $(\xi_\ell)_{\ell \geq 1}$ qui sont supposées connues.

Un lemme essentiel pour la démonstration du résultat d'identification dans ce cas est le suivant:

Lemme 2 *Soit $J_a : \mathbb{C}^{N \times N} \to \mathbb{R}$, $a = 1, 2$ et $h : \mathbb{R}^{d+1} \to \mathbb{R}$ des fonctions analytiques réelles avec J_a bornée. Soit $A_a, B_a \in \mathfrak{su}(N)$, $T > 0$, $\epsilon \in \mathrm{L}^1([0, T], \mathbb{R})$ et on note $X_a(t, y_\ell, \epsilon)$ la solution de*

$$\begin{cases} \frac{dX_a(t, y_\ell, \epsilon)}{dt} = (A_a + h(\epsilon(t), y_\ell) B_a) X_a(t, y_\ell, \epsilon) \\ X_a(0, y_\ell, \epsilon) = Id, \end{cases} \tag{12}$$

pour $a = 1, 2$ et tout $\ell \in \mathcal{I}$.

On suppose avoir l'égalité en loi suivante:

$$\mathcal{L}_Y(J_1(X_1(T, Y, \epsilon))) = \mathcal{L}_Y(J_2(X_2(T, Y, \epsilon))) \quad \forall \epsilon \in \mathrm{L}^1([0, T], \mathbb{R}). \tag{13}$$

Alors pour tout $\ell \in \mathcal{I}$, il existent $n_0(\ell, \xi_1, \cdots, \xi_n, \cdots)$ et un indice $\kappa(\ell) \in \mathcal{I}$, $\kappa(\ell) \leq n_0(\ell, \xi_1, \cdots, \xi_n, \cdots)$ tels que

$$J_1(X_1(T, y_\ell, \epsilon)) = J_2(X_2(T, y_{\kappa(\ell)}, \epsilon)) \quad \forall \epsilon \in \mathrm{L}^1([0, T], \mathbb{R}). \tag{14}$$

Dans la suite, on suppose que la perturbation est additive. C'est-à-dire $u(t) = \epsilon(t) + Y$.

Les résultats principaux de ce chapitre sont les théorèmes suivants:

Théorème 3 *On considère la même situation et on suppose les mêmes hypothèses que dans le théorème 2 sauf la relation (9). Alors il existe $T > 0$ tel que si:*

$$\mathcal{L}_Y\langle O_k\Psi_1(T,\epsilon+Y),\Psi_1(T,\epsilon+Y)\rangle = \mathcal{L}_Y\langle O_k\Psi_2(T,\epsilon+Y),\Psi_2(T,\epsilon+Y)\rangle \tag{15}$$

$$\forall \epsilon \in L^1([0,T];\mathbb{R}), \quad \forall k = 1,\cdots,K,$$

alors les conclusions (10) et (11) du théorème 2 restent vraies.

Ici, le temps T doit être assez grand. La preuve est adaptée pour tout temps T^* supérieur au temps T^*.

De même, pour le cas **(S1)**, on a:

Corollaire 1 *On considère la même situation et on suppose les mêmes hypothèses que dans le théorème 1 sauf la relation (21). Alors il existe $T > 0$ tel que si:*

$$\mathcal{L}_Y\langle O\Psi_1(T,\epsilon+Y),\Psi_1(T,\epsilon+Y)\rangle = \mathcal{L}_Y\langle O\Psi_2(T,\epsilon+Y),\Psi_2(T,\epsilon+Y)\rangle \tag{16}$$

$$\forall \epsilon \in L^1([0,T];\mathbb{R}), \quad \forall O \in \mathcal{O},$$

alors la conclusion (8) du théorème 1 reste vraie.

Maintenant on suppose que la perturbation est multiplicative. Le contrôle est sous la forme de $u(t) = Y \cdot \epsilon(t)$. On suppose en plus que la perturbation est positive: $\mathcal{V} \subset \mathbb{R}^+$.

Corollaire 2 *On considère la même situation et on suppose les mêmes hypothèses que dans le théorème 2 sauf la relation (9). Alors il existe $T > 0$ tel que si:*

$$\mathcal{L}_Y\langle O_k\Psi_1(T,\epsilon Y),\Psi_1(T,\epsilon Y)\rangle = \mathcal{L}_Y\langle O_k\Psi_2(T,\epsilon Y),\Psi_2(T,\epsilon Y)\rangle \quad (17)$$
$$\forall \epsilon \in L^1([0,T];\mathbb{R}), \quad \forall k = 1,\cdots,K,$$

alors les conclusions (10) et (11) du théorème 2 restent vraies.

0.6.3 Application numérique: perturbation additive

La simulation numérique est faite dans le cas **(C2)** pour le système défini par les matrices (4) et (5) (voir section 0.4).

On diagonalise la matrice $H_{réel}$ avec $H_{réel} = e^{\mathcal{P}_{réel}} D e^{-\mathcal{P}_{réel}}$,

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.0365 & 0 & 0 \\ 0 & 0 & 0.0651 & 0 \\ 0 & 0 & 0 & 0.0857 \end{pmatrix}, \mathcal{P}_{real} = \begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & -1 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{pmatrix}.$$

En réalité les valeurs propres de l'Hamiltonien sont connues avec grande précisions. Donc on peut supposer que la matrice D est connue. Alors l'identification de $H_{réel}$ revient à l'identification de la matrice de rotation antihermitienne $\mathcal{P}_{réel}$.

La distribution de la perturbation Y est donnée dans le tableau 2.1. On

utilise plusieurs contrôles sous la forme

$$\epsilon(t) = \exp\left(-40\left(\frac{t-T/2}{T}\right)^2\right) \sum_{1\leq i<j\leq N} A_{ij}\sin[(\lambda_j(H_{réel})-\lambda_i(H_{réel}))t+\theta_{ij}]. \tag{18}$$

Ici $\lambda_i(H_{réel})$ sont les valeurs propres de $H_{réel}$, $i \leq N$ et A_{ij}, θ_{ij} des paramètres de notre choix. Le temps total de simulation est $T = 3200$ qui correspond à 10 périodes de la plus petite transition $\lambda_4(H_{real}) - \lambda_3(H_{real})$.

Les observables choisies sont les populations $\mathcal{O} = \{e_k e_k^*, k \leq N\}$ associées à la base canonique $\{e_k; k \leq N\}$. On choisit $N_\epsilon = 36$ contrôles $\epsilon_1(t), ..., \epsilon_{N_\epsilon}(t)$ avec θ_{ij} arbitairement dans $[0, 2\pi]$ et A_{ij} arbitairement dans $[0, 0.0012]$ et la fonction à minimiser est:

$$\mathcal{J}(\mathcal{P},\mu) = \sum_{i=1}^{N_\epsilon}\sum_{j=1}^{N} d_{\mathcal{W}_1}(\mathcal{L}_Y(|\langle\Psi(T, e^{\mathcal{P}}De^{-\mathcal{P}}, \epsilon_i + Y, \mu, \Psi_1^0), e_j\rangle|^2, \mathcal{L}_Y(|\langle\Psi(T, H_{réel}, \epsilon_i + Y, \mu_{real}, \Psi_{réel}^0), e_j\rangle|^2). \tag{19}$$

La distance qu'on utilise ici est la 1-distance de Wasserstein.

On commence par une erreur relative de 10% sur μ et $\mathcal{P}$. Après 277 itérations, on obtient:

$$\mathcal{P}_{277} = \begin{pmatrix} 0 & 0.999 & -0.999 & 1.002 \\ -0.999 & 0 & 1 & 0.999 \\ 0.999 & -1 & 0 & -1.002 \\ -1.002 & -0.999 & 1.002 & 0 \end{pmatrix},$$

$$\mu_{277} = \begin{pmatrix} 0 & 4.999 & -0.998 & -0.003 \\ 4.999 & 0 & 6 & -1.5 \\ -0.998 & 6 & 0 & 7 \\ -0.003 & -1.5 & 7 & 0 \end{pmatrix}.$$

Ces résultats correspondent aux erreurs relatives de 0.003% sur μ et 0.001% sur $\mathcal{P}$. Les figures Figure 1 et Figure 2 montrent que l'algorithme converge bien vers le système réel.

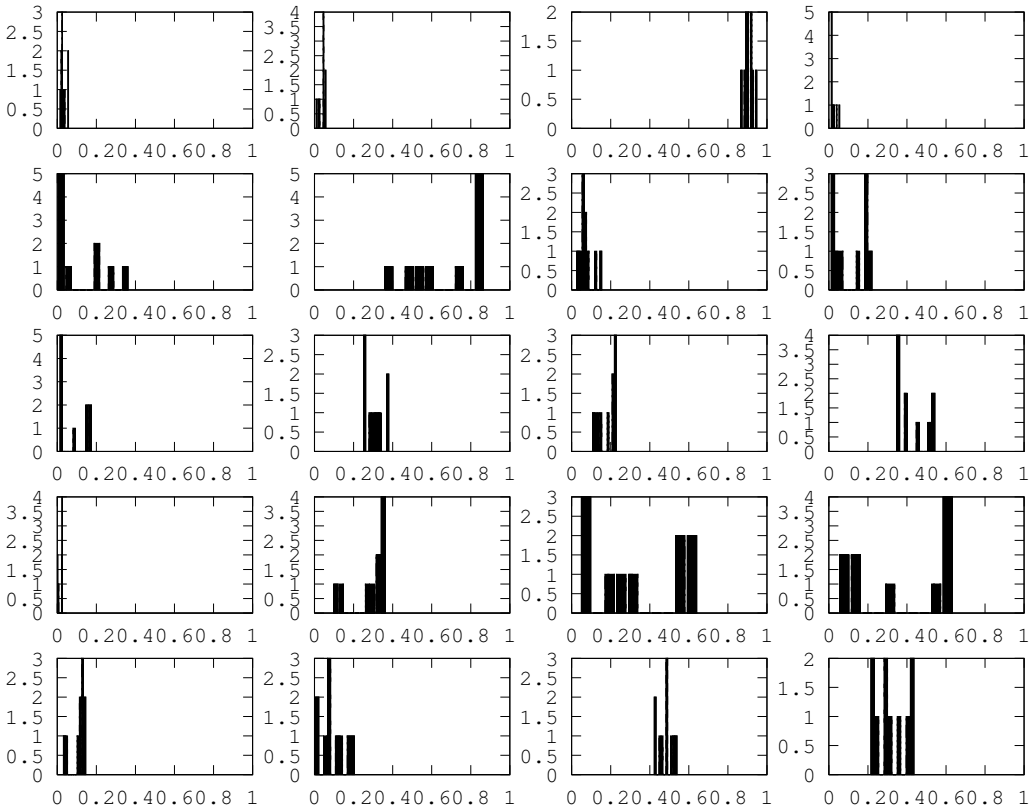

Figure 1: On figure l'histogramme des lois $\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}_{real}} D e^{-\mathcal{P}_{real}}, \epsilon_i + Y, \mu_{real}, \Psi_1^0), e_j \rangle|^2$ pour choix différents de $i = 1, ..., 5$ et $j = 1, 2, 3, 4$.

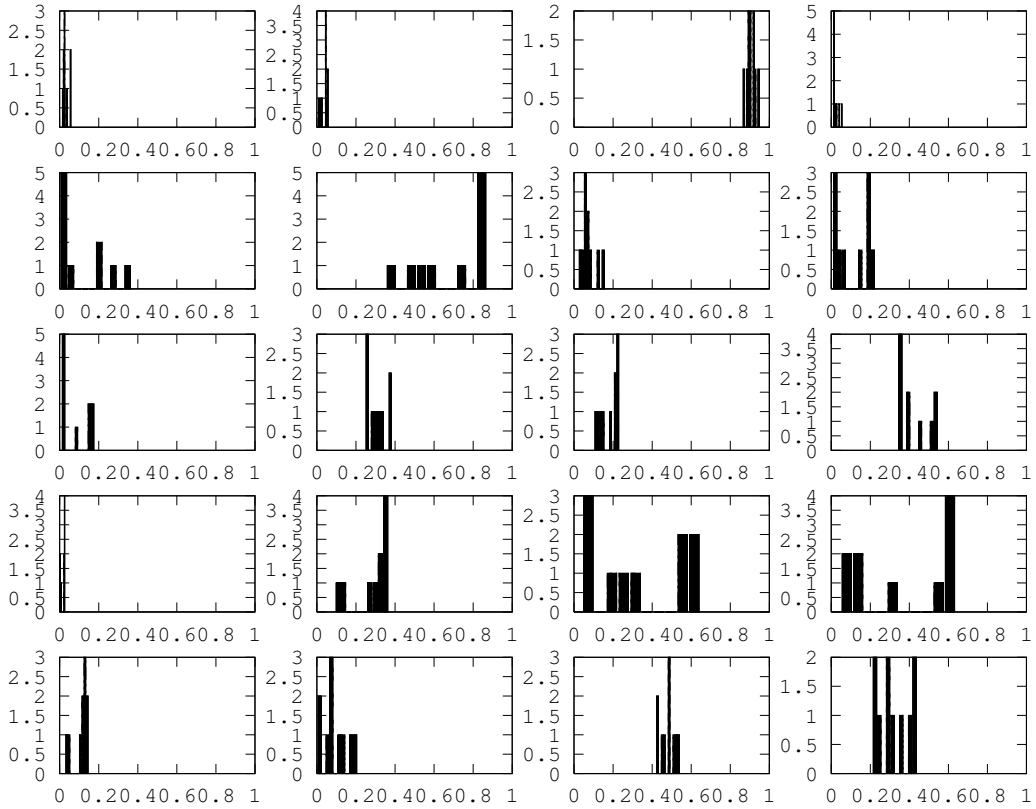

Figure 2: On figure l'histogramme des lois $\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}_{277}} D e^{-\mathcal{P}_{277}}, \epsilon_i + Y, \mu_{277}, \Psi_1^0), e_j \rangle|^2$ pour choix différents de $i = 1, ..., 5$ et $j = 1, 2, 3, 4$.

0.7 Chapitre 3: perturbation multiplicative de loi inconnue sur les amplitudes

Dans ce chapitre, les perturbations sont dans les amplitudes. On sait que dans le laboratoire, le contrôle est une superposition de plusieurs fréquences:

$$u(t) = \sigma(t) \sum_{\alpha \neq \beta} A_{\alpha\beta} sin(\omega_{\alpha\beta} t + \theta_{\alpha\beta}), \tag{20}$$

avec $\sigma(t)$ est une fonction Gaussian en temps et $\omega_{\alpha\beta} = E_\beta - E_\alpha$ les transitions liées aux valeurs propres de H E_α et E_β . Les amplitudes $A_{\alpha\beta}$ et les phases $\theta_{\alpha\beta}$ sont des paramètres à contrôler.

En réalité, quand on rèpète la même expérience plusieurs fois, il y aura des perturbations dans les amplitudes. C'est-à-dire pour chaque expérience, il y a un facteur multiplicatif sur le contrôle. La perturbation est modélisée par une variable aléatoire Y, le contrôle est sous la forme multiplicative $Y \cdot u(t)$. On fait l'hypothèse que les perturbations sur toutes les amplitudes sont identiques. En plus on ne considère pas de perturbations sur les phases $\theta_{\alpha\beta}$.

Le support de la perturbation Y est un ensemble fini $\mathcal{V} = \{y_\ell, \ell \leq L\} \subset \mathbb{R}$. Ainsi la distribution de Y est complètement définie par les probabilités $\xi_\ell = \mathbb{P}(Y = y_\ell)$.

A la différence du chapitre précédant, la distribution $(\xi_\ell)_{\ell=1}^L$ fait partie des inconnues du problème. L'ensemble $\mathcal{V}$ est supposé donné (ce qui n'enlève en pratique rien à la généralité du problème, car on peut prendre $\mathcal{V}$ aussi grand que nécessaire).

Le théorème suivant montre que sous certaines hypothèses, si on obtient les même distributions pour tous les observables dans un EOC $\mathcal{O}$ et tous les contrôles, alors le moment dipolaire μ et les probabilités $(\xi_\ell)_{\ell=1}^L$ sont identifiables à des phases multiplicatives près.

Théorème 4 *Soient H, μ_1, $\mu_2 \in \mathfrak{H}$, H diagonale, $\mu_1 \neq 0$, $\mu_2 \neq 0$, Y_1, Y_2 deux variables aléatoires avec le même support $\mathcal{V}$ qui contient au moins un élement non nul, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$ des états initiaux et pour $a = 1, 2$ et $u \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$, on note: $\Psi_a(t,u) = \Psi(t, H, u(\cdot), \mu_a, \Psi_0^a)$. Soit $\mathcal{O}$ un EOC non trivial.*

On suppose que $N \geq 3$ et les hypothèses suivants:

1. $\mathbb{L}_{iH,i\mu_1} = \mathbb{L}_{iH,i\mu_2} = \mathfrak{su}(N)$;
2. $tr(H) = tr(\mu_1) = tr(\mu_2) = 0$;
3. *les valeurs propres de H sont tous de multiplicité* 1.
4. $|(\mu_1)_{k,\ell}|^2 = |(\mu_2)_{k,\ell}|^2 \neq 0$ *pour un couple (k,ℓ) fixé.* .

Le temps final d'observation est noté T et est supposé assez grand. Si on a l'égalité de distributions suivante:

$$\begin{aligned} &\mathcal{L}(\langle O\Psi_1(T, uY_1), \Psi_1(T, uY_1)\rangle) = \mathcal{L}(\langle O\Psi_2(T, uY_2), \Psi_2(T, uY_2)\rangle) \quad (21)\\ &\forall u \in L^1([0,T];\mathbb{R}), \quad \forall O \in \mathcal{O}, \end{aligned}$$

alors il existent des phases $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$ telles que

$$\begin{cases} (\mu_1)_{jk} = \pm e^{i(\alpha_j - \alpha_k)}(\mu_2)_{jk}, \ \forall j,k \leq N, \\ \mathbb{P}(Y_1 = y_\ell) = \mathbb{P}(Y_2 = \pm y_\ell) \qquad \forall \ell \leq L. \end{cases} \tag{22}$$

On remarque que si le couple (Y, μ) est une solution, alors toutes les couples $(Y/\lambda, \lambda\mu)$ sont des solutions.

Plusieurs simulations numériques sont faites avec le système défini par les matrices (4) et (5).

On suppose que l'Hamiltonien H est connu, donc les observables $\mathcal{O} = \{O_1, ..., O_K\}$ sont les projections $\{|e_1\rangle\langle e_1|, \ ... \ , |e_N\rangle\langle e_N|\}$ sur une base propre de H. Ici $|e_1\rangle$ est le j^e vecteur propre de H.

Les expériences sont répétés pour plusieurs contrôles u_1,...,u_{N_u} et on minimise le critère défini par

$$\mathcal{J}(\mu, (\xi_k)_{k=1}^{L}; (u_i)_{i=1}^{N_u}) = \log \left\{ \frac{1}{N_u} \sum_{i=1}^{N_u} \sum_{j=1}^{N} \mathcal{W}_1 \left[\sum_{k=1}^{L} \xi_k \delta_{|\langle \Psi(T,H,u_i \cdot y_k, \mu, \Psi_1^0), e_j \rangle|^2}, \right. \right.$$
$$\left. \left. \sum_{k=1}^{L} \xi_k^{real} \delta_{|\langle \Psi(T,H,u_i \cdot y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2} \right] \right\}. \tag{23}$$

On commence par un choix initial μ^0 et une distribution initiale ξ^0 qui est la loi uniforme. L'itération $n \geq 1$ consiste en plusieurs étapes:

1. On choisit arbitairement N_u contrôles u_i^n, $i = 1, ..., N_u$;

2. on minimise $\xi \to \mathcal{J}(\mu^{n-1}, \xi; (u_i)_{i=1}^{N_u})$ et on note par ξ^n le minimiseur;

3. on minimise $\mu \to \mathcal{J}(\mu, \xi^n; (u_i)_{i=1}^{N_u})$ and on note par μ^n un minimiseur.

Pour les illustrations numériques on prend $N_u = 36$; les amplitudes $(A_{\alpha\beta})$ sont uniformément prises dans $[0, 0.0012]$ et les phases $\theta_{\alpha\beta}$ dans $[0, 2\pi]$.

Les valeurs intiales de μ sont:

$$\mu_g^0 = \begin{pmatrix} 0 & 3.76 & -1.31 & 0 \\ 3.76 & 0 & 3.51 & -1.78 \\ -1.31 & 3.51 & 0 & 6.72 \\ 0 & -1.78 & 6.72 & 0 \end{pmatrix}, \tag{24}$$

$$\mu_e^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \tag{25}$$

$$\mu_b^0 = \begin{pmatrix} 0 & 7.48 & -0.51 & 0 \\ 7.48 & 0 & 8.83 & -0.87 \\ -0.51 & 8.83 & 0 & 5.87 \\ 0 & -0.87 & 5.87 & 0 \end{pmatrix}. \tag{26}$$

Les observables sont des projections $\{|e_1\rangle\langle e_1|, |e_2\rangle\langle e_2|, |e_3\rangle\langle e_3|, |e_4\rangle\langle e_4|\}$ sur les vecteurs propres:

$$\begin{aligned} |e_1\rangle &= \begin{pmatrix} 0.0845 & -0.1313 & 0.9651 & 0.2101 \end{pmatrix}^T, \\ |e_2\rangle &= \begin{pmatrix} -0.1305 & -0.0856 & -0.2103 & 0.9651 \end{pmatrix}^T, \\ |e_3\rangle &= \begin{pmatrix} 0.2118 & 0.9647 & 0.0838 & 0.1325 \end{pmatrix}^T, \\ |e_4\rangle &= \begin{pmatrix} 0.9649 & -0.2118 & -0.1314 & 0.0830 \end{pmatrix}^T. \end{aligned} \tag{27}$$

On pose le temps final $T = 3200$ et on suppose que $|(\mu_{real})_{12}|^2 = 25$ est connu.

Le support de la distribution Y est connu et on le note $[y_m, y_M]$. On prend $y_m = 0.5$ et $y_M = 1.5$. On discrétise le support avec $L = 51$ points equidistants $y_\ell = y_m + (\ell - 1) \cdot \frac{y_M - y_m}{L-1}$, $\ell = 1, ..., L$. Plusieurs distributions sont testées:

- $Y^{real} = Y^g$ la distribution Gaussienne centrée en 1 avec la variance 0.0025 (voir équation (3.17) dans section 3.3.2);

- $Y^{real} = Y^e$ la distribution exponentielle déplacée (voir équation (3.18) dans section 3.3.2);

- $Y^{real} = Y^b$ la distribution bi-modale qui est la somme des deux distributions Gaussiennes: la première est centrée en 0.8 avec la variance 0.0025 et la seconde est centrée en 1.2 avec la variance 0.0049 (voir équation (3.19) dans section 3.3.2).

Après 10 itérations, on obtient:

$$\|\mu_g^{10} - \mu_{real}\|_\infty = 5 \cdot 10^{-5},\ \mu_g^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99995 & -1.5 \\ -1 & 5.99995 & 0 & 6.99999 \\ 0 & -1.5 & 6.99999 & 0 \end{pmatrix}, \tag{28}$$

$$\|\mu_e^{10} - \mu_{real}\|_\infty = 10^{-4},\ \mu_e^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 6.9999 \\ 0 & -1.5 & 6.9999 & 0 \end{pmatrix}, \tag{29}$$

$$\|\mu_b^{10} - \mu_{real}\|_\infty = 6 \cdot 10^{-5},\ \mu_b^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99999 & -1.5 \\ -1 & 5.99999 & 0 & 6.99994 \\ 0 & -1.5 & 6.99994 & 0 \end{pmatrix}. \tag{30}$$

Les figures 3.1,3.2, 3.3 et les tableaux 3.1 et 3.2 présentent la convergence de l'algorithme.

La simulation est aussi faite dans le cas où on utilise juste une observable qui est la projection $|e_3\rangle\langle e_3|$. Le choix initial est

$$\mu_b^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \tag{31}$$

La distribution Y^{real} testée est la distribution bimodale. Après 5,10 et 15 itérations, le moment dipolaire obtenu ont respectivement les erreurs L^2 de 0.17246, 0.03736 et $4.5652 \cdot 10^{-4}$. La vitesse de convergence est moins rapide que dans le cas précédent.

0.8 Chapitre 4: perturbation sur les phases

Dans ce chapitre nous considérons un modèle de bruit de loi paramétrique (avec paramètre connu). Le bruit est cette fois-ci une fonction dépendante de la fréquence (en particulier donc pas constant).

On commence par écrire le contrôle comme une intégrale sur les fréquences:

$$u(t) = S(t) \int_{\omega \in \mathcal{D}} A(\omega) cos(\omega t + \theta(\omega)) d\omega. \tag{32}$$

Ici $\mathcal{D}$ est une partie bornée de $\mathbb{R}_+$ qui répresente l'ensemble des fréquences possibles. $S(t)$ est une enveloppe Gaussienne dans le temps. Les amplitudes $A(\omega) : \mathcal{D} \to \mathbb{R}_+$ et les phases $\theta(\omega) : \mathcal{D} \to [0, 2\pi]$ sont les paramètres nominaux du contrôle. Les amplitudes A doivent être intégrable sur $\mathcal{D}$.

Dans ce chapitre, on considère des perturbations dans les phases qui sont modélisées comme un processus Gaussien indexé par les frequences. On introduit le modèle suivant:

$$u(t) = S(t) \int_{\omega \in \mathcal{D}} A(\omega) cos(\omega t + \theta(\omega) + \delta\theta_\omega). \tag{33}$$

Les perturbations dans les phases sont $(\delta\theta_\omega)_{\omega \in \mathcal{D}}$.

En pratique, les physiciens utilisent un nombre fini de frequences pour construire le contrôle. Notons par N_ω ce nombre. Alors une discrétisation de l'équation (33) est:

$$u(t) = S(t) \sum_{l=0}^{N_\omega} A_l cos(\omega_l t + \theta_l + \delta\theta_l), \tag{34}$$

avec $A_l = A(\omega_l)$, $\theta_l = \theta(\omega_l)$ et $\delta\theta_l$ est la variable aléatoire $\delta\theta_{\omega_l}$.

L'espérance de $u(t)$ est $\mathbb{E}(u(t)) = \alpha u_0(t)$ avec $u_0(t) = S(t)\sum_{l=1}^{N_\omega} A_l cos(\omega_l t + \theta_l)$ le contrôle sans perturbation et $\alpha = \sum_{k=0}^{\infty}(-1)^k \frac{(2k)!}{2^k k!}\sigma^{2k}$.

Les observables qu'on utilise dans ce chapitre sont les projections sur une base propre de H: $\mathcal{O} = \{O_1, ..., O_N\} = \{|e_1\rangle\langle e_1|, \ ... \ , |e_N\rangle\langle e_N|\}$ avec $|e_i\rangle$ le i^e vecteur propre de H.

Les fréquences $(\omega_l)_{1\leq l\leq N_\omega}$ sont choisies comme les transitions des valeurs propres de H: $|\lambda_i - \lambda_j|$, avec $(\lambda_i)_{1\leq i\leq N}$ les valeurs propres de H. En particulier, $N_\omega = \frac{N(N-1)}{2}$.

Le processus Gaussien $(\delta\theta_\omega)_{\omega\in\mathcal{D}}$ est simulé par N_r réalisations aléatoires.

La distribution du contrôle simulé $u(t)$ est donc

$$\mathcal{L}_{u(t)} = \sum_{k=1}^{N_r} \frac{1}{N_r}\delta_{S(t)\sum_{l=1}^{N_\omega} A_l cos(\omega_l t + \theta_l + \delta\theta_{l,k})}, \tag{35}$$

où $(\delta\theta_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r} \in \mathbb{R}^{N_\omega \cdot N_r}$ sont N_r réalisations.

On minimise un critère construit avec des contrôles $(u^j)_{1\leq j\leq N_u}$ et $(\tilde{u}^j)_{1\leq j\leq N_u}$ defini par N_u couples d'amplitudes différentes $(A_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$ et phases différentes $(\theta_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$.

Chaque u^j est simulé avec N_r réalisations $(\delta\theta_{l,k}^j)_{1\leq l\leq N_\omega, 1\leq k\leq N_r} \in \mathbb{R}^{N_\omega \cdot N_r}$:

$$\mathcal{L}_{u^j(t)} = \sum_{k=1}^{N_r} \frac{1}{N_r}\delta_{S(t)\sum_{l=1}^{N_\omega} A_l^j cos(\omega_l t + \theta_l^j + \delta\theta_{l,k}^j)}. \tag{36}$$

Les $\tilde{u}^j(t)$ sont définis par:

$$\tilde{u}^j(t) = S(t) \sum_{l=0}^{N_\omega} A_l^j cos(\omega_l t + \theta_l^j + \delta\theta_l). \tag{37}$$

La critère à minimiser est:

$$\begin{aligned}\mathcal{J}(u^j, \mu, \tilde{u}^j, \mu_{real}) = &\sum_{i=1}^{N} \mathcal{W}_1(|\langle \Psi(T, H, u^j, \mu, \Psi_0), e_i\rangle|^2, \\ &|\langle \Psi(T, H, \tilde{u}^j, \mu_{real}, \Psi_0), e_i\rangle|^2),\end{aligned} \tag{38}$$

Sa moyenne est:

$$\tilde{\mathcal{J}}((u^j)_{1\leq j\leq N_u}, \mu, (\tilde{u}^j)_{1\leq j\leq N_u}, \mu_{real}) = \frac{1}{N_u} \sum_{j=1}^{N_u} \mathcal{J}(u^j, \mu, \tilde{u}^j, \mu_{real}). \tag{39}$$

Plusieurs simulations numériques sont faites avec le système défini par les matrices (4) et (5).

On choisit $N_u = 36$. Les amplitudes sont prises uniformément dans $[0, 0.0012]$ et les phases dans $[0, 2\pi]$. On commence par

$$\mu^0 = \begin{pmatrix} 0 & 5.1295 & -0.9762 & 0.0962 \\ 5.1295 & 0 & 5.5100 & -1.6434 \\ -0.9762 & 5.5100 & 0 & 7.6117 \\ 0.0962 & -1.6434 & 7.6117 & 0 \end{pmatrix}, \tag{40}$$

qui correspond à une erreur relative de 8.7%.

On choisit la covariance de double exponentiel (voir exemple 1.3) pour les perturbations:

$$\Sigma_{l,l'} = \sigma^2 e^{-\frac{(\omega_l - \omega_{l'})^2}{\beta}} \tag{41}$$

avec β un paramètre réel. On pose $\sigma = 0.1$.

La valeur β doit être bien calibrée pour que les perturbations soient corrélées mais pas trop corrélées. Les figures Figure 4.1, Figure 4.2 et Figure 4.3 montrent que $\beta = 0.1$ est un bon choix.

On pose $N_r = 1000$. Après 50 itérations, on obtient:

$$\mu^{50} = \begin{pmatrix} 0 & 4.9781 & -0.9307 & -0.0063 \\ 4.9781 & 0 & 5.9568 & -1.5189 \\ -0.9307 & 5.9568 & 0 & 7.0682 \\ -0.0063 & -1.5189 & 7.0682 & 0 \end{pmatrix},$$

qui correspond à une erreur relative de 1%.

Maintenant on utilise le modèle de matrice de covariance exponentielle (voir exemple 1.4):

$$\Sigma_{l,l'} = \sigma^2 e^{-\frac{|\omega_l - \omega_{l'}|}{\beta'}}. \tag{42}$$

Un bon candidat de β' est $\beta' = 2$ (voir Figure 4.8). Cette fois $N_r = 100$. Après 60 itérations, on obtient

$$\mu^{60} = \begin{pmatrix} 0 & 5.0641 & -1.0585 & 0.0068 \\ 5.0641 & 0 & 5.9642 & -1.5139 \\ -1.0585 & 5.9642 & 0 & 7.0294 \\ 0.0068 & -1.5139 & 7.0294 & 0 \end{pmatrix},$$

qui corresponds à une erreur relative de 0.91%.

Finalement on suppose que la perturbation est un mouvement Brownien dans l'espace des fréquences avec

$$\Sigma_{l,l'} = \sigma^2 \min(\omega_l, \omega_{l'}). \tag{43}$$

On choisit $N_r = 100$. Après 60 itérations, on obtient

$$\mu^{60} = \begin{pmatrix} 0 & 5.0043 & -1.0138 & 0 \\ 5.0043 & 0 & 6.0078 & -1.5015 \\ -1.0138 & 6.0078 & 0 & 6.9834 \\ 0 & -1.5015 & 6.9834 & 0 \end{pmatrix},$$

qui corresponds à une erreur relative de 0.24%.

Chapter 1

Mathematical formulation of quantum mechanics

The quantum theory is widely used in various domains, including the laser and NMR technology. To give a mathematical formulation of the quantum theory, we introduce the notions of states, observables and density matrices. The main equation we consider in this thesis is the Schrödinger equation. To study the controllability of a Schrödinger equation, the theory of Lie group and Lie algebra is necessary. Some tools concerning the distance between the probability distributions and several basic elements in the Gaussian process theory which are useful in the numerical simulation are also introduced.

1.1 Physical backgrounds

Today, classical physics is still used in much of modern science and technology. However classical physics can only explains energy and matter on a scale familiar to human experience. In the end of 19th century scientists discovered phenomena in both macro (the large) and micro (the small) worlds that classical physics could not explain. These limitations led to the development of quantum mechanics.

In contrast to the classical physics, quantum theory is invoked to describe all phenomena from the very small to the very large, covering about sixty orders of magnitude in dimensions. Between these extremes, we find all the objects of the world around us.

In the following, we give some examples of the applications of quantum theory. We cite the presentation in [23] and [33].

1.1.1 Applications of quantum mechanics

Quantum theory as a unifying description of Nature *As a proof of the success of quantum theory stands, the unification of three out of the four fundamental interactions – the electromagnetic, strong and weak forces – in the Standard Model, which reveals the deep symmetries of Nature. These are also the promising attempts to include gravitation in a unified theory of strings. And all physicists will certainly emphasize the universality of physics, a remarkable consequence of quantum laws. It is the quantum theory which explains the radiation spectrum of hydrogen in our laboratory lamps, but also in intergalactic space. Quantum chemistry applies to the reactions in laboratory test tubes, but also to the processes in the interstellar dust where molecules are formed and destroyed, producing radiation detected by our telescopes after traveling billions of years through space. Let us finally evoke cosmology and the remarkable link between the infinitely small and large scales, underlined*

by the quantum theory. It emphasizes the similarity between the phenomena which occurred at the origin of the Universe, in a medium of inconceivably large temperature and density, and those which happen in the violent collisions between particles studied by large accelerators on Earth.

The laser and the optical revolution *The laser is an example of another invention based on a quantum idea. Light had been forever made of random waves, difficult to direct, to focus or to force to oscillate at a well-defined frequency. The laser has changed this state of affairs and has allowed us to tame radiation by exploiting the properties of atomic stimulated emission, discovered by Einstein at the dawn of the quantum era. Lasers now have a huge variety of uses, from the very mundane to the most sophisticated. Laser light traveling through fibers can transport huge amounts of information over very long distances. Laser beams are used to print and read out information on compact disks, with applications for the reproduction of sounds, pictures and movies.*

In scientific research, the flexible properties of laser beams have countless applications, some of which are essential to realize the manipulations of single particles which will be our main topic here. Experiments exploit their high intensity to study non-linear optical processes. The extreme monochromaticity of lasers and their time coherence is used for high-resolution spectroscopy of atoms and molecules. Combining monochromaticity and high intensity has proved essential to trap and cool atoms to extremely low temperatures. Laser pulses of femtosecond duration probe very fast processes in molecules and solids and study chemical reactions in real time.

Nuclear magnetic resonance and medical diagnosis *Nuclear Magnetic Resonance (NMR) is another technology based on quantum science which nowadays plays a major role in scientific research and in medical diagno-*

sis. The nuclei of a variety of atoms carry magnetic moments attached to their intrinsic angular momentum or spin. The origin of this magnetism is fundamentally quantum. The quantization of the spin orientation in space, the fact that its projection on any direction can take only discrete values, has been one of the smoking guns of the quantum theory, forcing physicists to renounce their classical views. The evolution of spins in magnetic fields, both static and time-modulated, requires a quantum approach to be understood in depth. In a solid or liquid sample, the spins are affected by their magnetic environment and their evolution bears witness of their surroundings.

In a typical NMR experiment, one immerses the sample in a large magnetic field which superimposes its effects on the local microscopic field. One furthermore applies to the sample sequences of tailored radio-frequency pulses. These pulses set the nuclei in gyration around the magnetic field and the dance of the spins is detected through the magnetic flux they induce in pick-up coils surrounding the sample. A huge amount of information is gained on the medium, on the local density of spins and on their environment.

This simple idea has been carried further in quantum information physics. By performing complex NMR experiments on liquid samples made of organic molecules, simple quantum computing operations have been achieved. The spins of these molecules are manipulated by complex pulse sequences, according to techniques originally developed by chemists, biologists and medical doctors in NMR and MRI. In these experiments, a huge number of molecules contribute to the signal. The situation is thus very different from the manipulation of the simple quantum systems we will be considering. Some of these methods do however apply. We will see that the atoms and ions which are individually manipulated in the modern realizations of thought experiments are also two-level spin-like systems and that one applies to them sequences of pulses similar to those invented

for NMR physics.

1.2 Basic elements of quantum theory

In this section we present the basic elements of quantum theory for *finite-dimensional systems*. This section is adapted from [3] and [8].

A quantum system $\mathcal{Q}$ can be modeled by a complex Hilbert space $\mathcal{H}$ whose dimension is decided by the amount of variables of the system we try to describe which is in general infinite.

However, in order to avoid technical complications, infinite dimensional quantum systems are often approached by finite dimensional quantum systems $\mathcal{H} \cong \mathcal{H}^N \simeq \mathbb{C}^N$. The advantage of finite dimensional systems is that many results of Lie algebra are established in this case, which are essential for the theory of controllability.

1.2.1 State vectors

In quantum mechanics, we describe the state in a closed quantum system by a *state vector* (also called wave function) ψ in a Hilbert space $\mathcal{H}$, in which the Hermitian inner product between two vectors ϕ and ψ is defined as $\langle\phi|\psi\rangle$. Sometimes we use also the notation $|\psi\rangle$ to represent the state vector in order to distinguish with its dual vector $\langle\psi|$. The norm of a vector ψ is now given by $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$. The state vector should have norm 1.

Given an orthonormal basis of $\mathcal{H}$: $\{|e_1\rangle, \ldots, |e_N\rangle\}$, the state vector $|\psi\rangle$ of the system with respect to this basis can be written as $|\psi\rangle = \sum_{j=1}^{N} c_j|e_j\rangle$, where $c_1, \cdots, c_N$ are complex numbers, then

$$\langle\psi|\psi\rangle = \sum_{j=1}^{N} |c_j|^2 = 1, \tag{1.1}$$

and $|c_j|^2$ is called the population of the j-th eigenstate.

Geometrically, equation (1.1) means that the state vector ψ is living on the unit sphere $\mathcal{S}_N$ of dimension N of the Hilbert space $\mathcal{H}$ of dimension N: $\psi \in \mathcal{S}_N \subset \mathcal{H}$.

1.2.2 Observables

The measurable quantities (energy, position, spin,...), or *observables* of the closed physical system are formulated mathematically by Hermitian operators on the Hilbert space $\mathcal{H}$.

Let $\alpha_1, \cdots, \alpha_N \in \mathbb{R}$ be the eigenvalues of an observable O: $O = \sum_{j=1}^{N} \alpha_j \Pi_j$, where $\{\Pi_j\}$ are the corresponding orthogonal projectors. The eigenvalues $\{\alpha_j\}$ are real numbers as O is self-adjoint. They represent the *possible outcomes* of a measurement of O. The projections Π_j, which play the role of *quantum events*, form a resolution of the identity, which means $\Pi_k \Pi_j = \delta_{kj} \Pi_j$ and $\sum_j \Pi_j = I$.

1.2.3 Density operators

Density operator is introduced to describe the state of *statistical ensembles*. With a single matrix, all possible states of a quantum system are summarized.

For a *statistical set* of states $|\psi_i\rangle \in \mathcal{H}$, $i = 1, 2, \cdots, m$ with probabilities p_i respectively, the associated quantum *density operator* (also called *density matrix*) is defined by

$$\rho = \sum_{j=1}^{m} p_j \mathcal{P}_{\psi_j}, \quad \mathcal{P}_{\psi_j} = \frac{|\psi_j\rangle\langle\psi_j|}{\langle\psi_j|\psi_j\rangle}. \tag{1.2}$$

where $\mathcal{P}_{\psi_j}$ is the orthogonal projector on the state $|\psi_j\rangle$.

In the particular case of a single state $|\psi\rangle$, $\rho = \mathcal{P}_\psi$ is a projector, and the quantum system is called in a *pure state*.

By definition of the density operator, it is a self-adjoint and positive-definite matrix. Thus the eigenvalues are positive which is coherent with the fact that p_j are some probabilities. The conservation of the total probability takes the form

$$\mathrm{tr}\,(\rho) = \sum_{j=1}^{m} p_j \mathrm{tr}\,(\mathcal{P}_j) = \sum_{j=1}^{m} p_j = 1. \tag{1.3}$$

One of the most important property of density operators is the following inequality

$$\mathrm{tr}\left(\rho^2\right) \leq \mathrm{tr}\,(\rho) = 1. \tag{1.4}$$

The quantity $\mathrm{tr}\,(\rho^2)$ is called the *purity* of ρ. The equality sign holds if and only if the system is in a *pure state*. In this case the density matrix ρ does not change when we replace $|\psi\rangle$ by $e^{i\theta}|\psi\rangle$ with $\theta \in [0, 2\pi]$ an arbitrary phase. As a result, we can therefore say that the set of pure states is isomorphic to the projective Hilbert space, which is the set of rays in $\mathcal{H}$.

Alternatively, let us denote the set of all density matrices by $\mathcal{S}(\mathcal{H})$. This set is convex which means that for any two density matrices ρ_1 and ρ_2, the convex linear combination $\rho = \lambda\rho_1 + (1-\lambda)\rho 2$ with $\lambda \in [0, 1]$ is also a density matrix. More precisely, if ρ is a pure state, then $\rho_1 = \rho_2 = \rho$.

When $\mathrm{tr}\,(\rho) < 1$, the ensemble is called in a *mixed state*.

While every statistical set of states together with the associated probabilities corresponds a unique ρ, the same density operator can be obtained from different sets. For example, $\rho = \frac{1}{N} I$ can be obtained by any orthonormal basis of $\mathcal{H}$ with probabilities $p_j = \frac{1}{N}, j = 1, \cdots, m$.

1.2.4 Probabilities and expectations

Recall that the state of a system is described by a state vector $|\psi\rangle$. For some observable $O = \sum_{j=1}^{m} \alpha_j \Pi_j$, the probability of observing the eigenvalue α_j as an outcome of the observable is computed as

$$p(\alpha_j) = \langle\psi|\Pi_j|\psi\rangle. \tag{1.5}$$

Thus $p(\alpha_j)$ is real and positive, and such that

$$\sum_{j=1}^{m} p(\alpha_j) = \langle\psi| \sum_{j} \Pi_j |\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{1.6}$$

In the case of density operators, given a pure state ρ, the probability of obtaining α_j as an outcome of O is:

$$p(\alpha_j) = \mathrm{tr}(\rho\Pi_j); \tag{1.7}$$

In quantum theory, a state vector is changed after a measurement. It becomes

$$|\psi\rangle_j = \frac{\Pi_j|\psi\rangle}{\sqrt{\langle\psi|\Pi_j|\psi\rangle}}, \tag{1.8}$$

where $|\psi\rangle_{O=\alpha_j}$ denotes the transformed state vector on recording the outcome α_j in a measurement of the observable O.

For some density operators ρ in a pure state, the conditional density operator after the outcome α_i has been recorded becomes

$$\rho|_{O=\alpha_j} = \frac{\Pi_j \rho \Pi_j}{\mathrm{tr}(\rho \Pi_j)}. \tag{1.9}$$

On mixed states (1.7) still holds by linearity of the trace function, and the validity of (1.9) is extended to this generic case.

1.3 The Schrödinger equation

The time evolution of the state vector ψ of a closed system can be described by the autonomous linear ordinary differential equation:

$$\begin{cases} \hbar\dot{\psi} & = -iH_0\psi, \qquad \psi \in \mathcal{S}_N \\ \psi(0) & = \psi_0, \end{cases} \tag{1.10}$$

called the *Schrödinger equation for the state vector*, where H_0 is the Hamiltonian of the system. To simplify, fix the Plank constant $\hbar$ to 1.

An equivalent description is the von Neumann equation, which gives the time evolution of the density matrix rho:

$$\begin{cases} \hbar\dot{\rho} & = -i[H_0, \rho] \\ \rho(0) & = \rho_0, \end{cases} \tag{1.11}$$

The real eigenvalues e_j, $j = 1, \ldots, N$ of H_0 are called the *energy levels* of the quantum system

.

When the quantum system is coupled with one or more electromagnetic fields, it can be described by:

$$\begin{cases} \dot{\psi} & = -i(H_0 + \Sigma_j\, u_j H_j)\psi, \qquad \psi \in \mathcal{S}_N \\ \psi(0) & = \psi_0, \end{cases} \tag{1.12}$$

where the self-adjoint Hamiltonian operators H_j contain the couplings between the energy levels of the free Hamiltonian H_0, and u_j, the amplitude of the interactions, represent our control parameters, $u_j \subset \mathbb{R}$.

Such a model is called "bi-linear" because both the control and the state enter linearly but the control multiplies the state. Further nonlinearities

in the control can also be proposed, see for instance [12].

We shall see that one control function is generically sufficient to ensure controllability. Hence, we consider from now on the simple case:

$$\begin{cases} \dot{\psi} & = -i(H_0 + u\mu)\psi \\ \psi(0) & = \psi_0. \end{cases} \tag{1.13}$$

The operator μ is called the dipole moment of the system.

1.3.1 Unitary propagator

The complex sphere $\mathcal{S}_N$, representing pure states, is a homogeneous space of the Lie group $U(N) = \{U \in GL(N, \mathbb{C}) \mid UU^T = U^T U = I\}$ as well as of its proper subgroup $SU(N) = U(N)/U(1)$, in which the global phase factor has been eliminated.

The Schrödinger equation can therefore be lifted to the Lie group $SU(N)$, obtaining in correspondence of (1.13) the right invariant matrix ODE:

$$\begin{cases} \dot{U} & = -i(H_0 + u\mu)U, \qquad U \in SU(N) \\ U(0) & = I, \end{cases} \tag{1.14}$$

called the *Schrödinger equation for the unitary propagator*.

For the system (1.14), the total Hamiltonian $H(t, u) = H_0 + u(t)H_1$ is in general time-varying. The solution of (1.14) is therefore given by a formal, time-ordered exponential

$$U(t) = \mathcal{T}\exp\left(-i\int_0^t H(s, u)ds\right), \tag{1.15}$$

Consequently, for (1.13) we have: $\psi(t) = U(t)\psi_0 = \mathcal{T}\exp\left(-i\int_0^t H(s,u)ds\right)\psi_0$.

The Lie algebras of $U(N)$ and $SU(N)$ are, respectively, $\mathfrak{u}(N) = \{A \in \mathbb{C}^{N\times N} \mid A^* = -A\}$ and $\mathfrak{su}(N) = \{A \in \mathfrak{u}(N) \mid \mathrm{tr}(A) = 0\}$.

Recall that $\mathfrak{su}(N)$ are semi-simple compact Lie algebras, meaning that the corresponding Killing forms are negative definite (see Section 1.5).

1.4 Noise models

Noise in the laser pulse originates from different sources. A general form of the laser pulse in the presence of noise can be expressed as:

$$\varepsilon^{\delta A,\delta\omega,\delta\theta}(t) = S(t)\int_{\mathcal{D}} A(\omega)\cdot(1+\delta A(\omega))cos(\omega(1+\delta\omega)(t-T/2)+\theta(\omega)+\delta\theta(\omega))d\omega. \tag{1.16}$$

where $\mathcal{D}$ is the set of possible frequencies.

A discretization of equation (1.16) is given by

$$\varepsilon^{\delta A,\delta\omega,\delta\theta}(t) = S(t)\sum_{l=1}^{N} A_l\cdot(1+\delta A_l)cos(\omega_l(1+\delta\omega_l)(t-T/2)+\theta_l+\delta\theta_l). \tag{1.17}$$

Here $S(t)$ is a global overall time envelope (for instance of Gaussian form); A_l is the nominal amplitude and δA_l is the relative shift in the pulse amplitude at a given frequency; ω_l is the nominal pulse frequency and $\delta\omega_l$ is the relative shift in pulse frequency; θ_l is the nominal pulse phase and $\delta\theta_l$ is the absolute shift in pulse phase induced by noise.

For experimental results, please refer to [40].

1.5 Lie Algebras

In this section we introduce briefly the basic properties of the Lie algebras. More details are given in [22], Chapter 3.

Definition 1.1 *A finite-dimensional differentiable manifold G which is also a group is said to be a Lie group if the group operations of multiplication and inversion are smooth maps, which is equivalent to the requirement that the mapping from $G \times G$ to G given by $(x, y) \to x^{-1}y$ is a smooth mapping.*

Definition 1.2 *A finite-dimensional real or complex Lie algebra is a finite-dimensional real or complex vector space $\mathfrak{g}$, together with a binary operation $[\cdot, \cdot]$ from $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$, called the Lie bracket satisfying the following properties:*

1. *$[\cdot, \cdot]$ is bilinear.*
2. *$[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$.*
3. *$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.*

Although the two definitions seem at first unrelated, in fact to any Lie group one can associate a Lie algebra which is isomorphic to the tangent space of the manifold isomorphism.

Proposition 1.1 *The Lie algebra $\mathfrak{g}$ of a matrix Lie group G is a real Lie algebra.*

Definition 1.3 *Let $\mathfrak{g}$ be a Lie algebra.Then the linear mapping* ad : $\mathfrak{g} \to End(\mathfrak{g})$ *given by $X \to \mathrm{ad}X$ with*

$$\mathrm{ad}X(Y) = [X, Y].$$

is a representation of a Lie algebra and is called the adjoint representation of the algebra.

The Killing form of a Lie algebra is

$$K(X,Y) = \mathrm{tr}(ad(X)ad(Y)). \tag{1.18}$$

It allows to give a simple characterization of compact Lie groups: a compact Lie group corresponds to a negative definite Killing form.

1.6 Controllability

Recall that a quantum system can be described by a finite-dimensional bilinear ODE

$$\dot{\psi} = (A + Bu(t))\psi, \tag{1.19}$$

where ψ represents the state vector varying on the unit sphere $\mathcal{S}_N$. The matrices A, B are in the Lie algebra of *skew-Hermitian* matrices of dimension n.

The solution of (1.19) at time t, $\psi(t)$, with initial condition ψ_0, is given by:

$$\psi(t) = X(t)\psi_0, \tag{1.20}$$

where $X(t)$ is the solution at time t of the equation

$$\dot{X}(t) = (A + Bu(t))X(t), \tag{1.21}$$

with initial condition $X(0) = I_n$. The matrix $X(t)$ varies on the Lie group of special unitary matrices $SU(n)$ or the Lie group of unitary matrices $U(n)$ according to whether or not the matrices A and B have all zero trace.

The following three types of controllability are introduced in [2].

Definition 1.4 ***(Operator-Controllability)***
The system (1.21) is operator-controllable if every desired unitary (or special unitary) operation on the state can be performed using an appropriate control field. This means that for any $X_f \in U(n)$ (or $SU(n)$) there exists an admissible control to drive the state X in (1.21) from the Identity to X_f.

Definition 1.5 ***(Pure-State-Controllability)***
The system is pure-state-controllable if for every pair of initial and

final states, ψ_0 and ψ_1 in $S^{n-1}_{\mathbb{C}}$ there exists control function u and a time $t > 0$ such that the solution of (1.19) at time t, with initial condition ψ_0, is $\psi(t) = \psi_1$.

Definition 1.6 ***(Density-Matrix-Controllability)***

The system is density matrix controllable if, for each pair of unitarily equivalent density matrices ρ_1 and ρ_2 (there exists a matrix $U \in U(n)$ such that $U\rho_1 U^* = \rho_2$), there exists a control $u_1, u_2, ..., u_m$ and a time $t > 0$, such that the solution of (1.21) at time t, $X(t)$, satisfies

$$X(t)\rho_1 X^*(t) = \rho_2. \tag{1.22}$$

The controllability can be investigated with tools coming from the theory of Lie groupsm, we refer to Chapter 2 and Chpter 3 for details.

1.7 Distance between two distributions

Definition 1.7 *A Polish space X is a topological space which is separable and completely metrizable.*

Definition 1.8 *Let $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$ be two probability spaces. Coupling μ and ν means constructing two random variables X and Y on some probability space $(\Omega, \mathbb{P})$, such that law$(X) = \mu$, law$(Y) = \nu$. The couple (X, Y) is called a coupling of (μ, ν).*

In the formulation of measure theory, coupling μ and ν means constructing a measure π on $\mathcal{X} \times \mathcal{Y}$ such that π admits μ and ν as **marginals** on $\mathcal{X}$ and $\mathcal{Y}$ respectively. We denote by $\Pi(\mu, \nu)$ the set of all couplings π between μ and ν.

In the following, we will introduce several usual distances used in the measure theory.

Definition 1.9 ***(Wasserstein distances)***
Let (χ, d) be a Polish metric space, and let $p \in [1, \infty)$. Let μ and ν two probability measures on χ. The Wasserstein distance of order p between μ and ν is defined by

$$W_p(\mu, \nu) = (\inf_{\pi \in \Pi(\mu,\nu)} \int_{\chi} d(x, y)^p d\pi(x, y))^{1/p} \tag{1.23}$$
$$= inf\{[\mathbb{E}(d(X, Y)^p)]^{1/p}, law(X) = \mu, law(Y) = \nu\}.$$

Remark 1.1 ***(Kantorovich - Rubinstein distance)***
The Wasserstein distance of order 1 is also called the Kantorovich - Rubinstein distance.

Definition 1.10 ***(Kolmogorov distance)***

Let the sample space $\Omega = \mathbb{R}$. Let F, G be two distribution functions. The Kolmogorov distance between F and G is defined by

$$d_K(F, G) = \sup_{x \in \mathbb{R}} |F(x) - G(x)| \tag{1.24}$$

1.8 Gaussian process

In probability theory and statistics, a Gaussian process is a statistical model where observations occur in a continuous domain, for example, time or space. In a Gaussian process, every point in some continuous input space is associated with a Gaussian distributed random variable. Moreover, every finite collection of those random variables has a multivariate normal distribution. We introduce here only some notions relative to Gaussian processes and some examples of Gaussian process mentioned in [32].

Definition 1.11 ***(Multivariate Gaussian distribution)***

Let $X = (X_1, \cdots, X_N)$ be a random vector defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. X is said to have the multivariate Gaussian distribution if and only if every linear combination of its components is normally distributed. That is, for any constant vector $a \in \mathbb{R}^N$, the random variable $Y^a = \langle a, X \rangle = \sum_{k=1}^{N} a_k X_k$ has a Gaussian distribution.

An equivalent definition using the characteristic function is:

Proposition 1.2 *If a random vector $X = (X_1, \cdots, X_N)$ has the multivariate Gaussian distribution, then the characteristic function of X is:*

$$\phi_X : \xi \in \mathbb{R}^N \to e^{i\langle \xi, \mathbb{E}[X] \rangle - \frac{1}{2}\langle \xi, K_X \xi \rangle} \tag{1.25}$$

where K_X is the covariance matrix of X, which is a symmetric, non-negative definite $N \times N$ matrix.

Theorem 1.1 *Let $X = (X_1, \cdots, X_N)$ be a multivariate Gaussian random variable. Its components are independent if and only if they are not correlated.*

Definition 1.12 ***(Gaussian process)***

Let $(X_t)_{t\in T}$ be a family of random variables defined on the same probability space indexed by a set T. The process X is said to be Gaussian if for any finite subset $F \subset T$, the random vector $X_F = (X_t)_{t\in F}$ has a multivariate Gaussian distribution.

If X is a Gaussian process, then according to the Proposition 1.2, its law is completely characterized by its mean value function $e_X : t \to \mathbb{E}[X_t]$ and its covariance operator $K_X : (s,t) \to \mathrm{cov}(X_s, X_t)$.

1.8.1 Regularity of random process

The continuity and differentiability of random process is very important in practical applications. The continuity of a random process is related to the convergence of sequences $(X_{t_n})_{t_n\in T}$ of random variables. In the following, we will introduce three types of continuity.

Definition 1.13 ***(Sample path continuity)***

The random process $(X_t)_{t\in T}$ indexed by a set T has continuous sample paths if

$$\mathbb{P}(|X_{t_n} - X_t| \overset{t_n\to t}{\longrightarrow} 0, \forall t \in T) = 1.$$

Definition 1.14 ***(Almost sure continuity)***

The random process $(X_t)_{t\in T}$ indexed by a set T is called almost surely continuous at the point $t \in T$ if for every sequence $(t_n)_{n\in\mathbb{N}}$ for which $|t_n - t| \to 0$ as $n \to 0$,

$$\mathbb{P}(\lim_{n\to+\infty} X_{t_n} = X_t) = 1.$$

The random process $(X_t)_{t\in T}$ is called almost surely continuous on T if it is almost surely continuous at all $t \in T$.

Definition 1.15 ***(Mean square continuity)***

The random process $(X_t)_{t\in T}$ indexed by a set T is called continuous in mean square (m.s. continuous) at the point $t \in T$ if for every sequence $(t_n)_{n\in\mathbb{N}}$ for which $|t_n - t| \to 0$ as $n \to 0$,

$$\mathbb{E}(|X_{t_n} - X_t|^2) \overset{n\to+\infty}{\longrightarrow} 0.$$

The random process $(X_t)_{t\in T}$ is called continuous in mean square (m.s. continuous) on T if it is mean square continuous at all $t \in T$.

Sample path continuity means there are no discontinuities within the whole domain T, while almost sure continuity allows discontinuities. Obviously, sample path continuity is stronger than almost sure continuity. Generally, m.s. continuity does not imply sample path continuity. Nor does sample path continuity imply m.s. continuity. Mean square continuity implies the continuity in probability. Let us recall the definition.

Definition 1.16 ***(Continuity in probability)***

The random process $(X_t)_{t\in T}$ *indexed by a set* T *is called continuous in probability if for every sequence* $(t_n)_{n\in\mathbb{N}}$ *for which* $|t_n - t| \to 0$ *as* $n \to 0$ *and for all* $\epsilon > 0$,

$$\lim_{n\to+\infty} \mathbb{P}(|X_{t_n} - X_t| \geq \epsilon) = 0.$$

In fact, a random process $(X_t)_{t\in T}$ is continuous in mean square at t if and only if its covariance function K_X is continuous at the point (t, t).

Mean square differentiability is defined in the similar way as the following.

Definition 1.17 ***(Mean square differentiability)***

The random process $(X_t)_{t\in T}$ *indexed by a set* T *is called differentiable at the point* $t \in T$ *if there exists*

$$\dot{X}_t = \lim_{r\to 0} \frac{X_{t+r} - X_t}{r}.$$

The random process $(X_t)_{t\in T}$ *is called differentiable in the mean square sense (m.s. differentiable) if*

$$\lim_{r\to 0} \mathbb{E}\|\frac{X_{t+r} - X_t}{r} - \dot{X}_t\|^2 = 0. \tag{1.26}$$

The m.s. differentiability of a random process is also related to the differentiability of its covariance function. This definition can be extended to higher orders.

Further details including the demonstrations and examples are discussed in Chapter 2 of [1].

1.8.2 Covariance Functions

A covariance function $(s,t) \to K_X(\mathrm{s,t})$ is said to be stationary if it depends only on $t-s$. Thus it is invariant to transitions. If in addition the covariance functions is a function depends only on $r = |t-s|$, then it is called isotropic. Some usual isotropic covariance functions are introduced in Chapter 4 of [32].

Example 1.1 ***(The Matérn class covariance function)***

The Matérn class covariance function is defined by

$$K_{Matérn}(r) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\frac{\sqrt{2\nu}r}{\ell})^{\nu} k_{\nu}(\frac{\sqrt{2\nu}r}{\ell}), \tag{1.27}$$

where k_ν is a modified Bessel function and σ^2, ν, ℓ are positive parameters.

When $\nu = \frac{1}{2}$, we obtain the exponential covariance function; when $\nu \to +\infty$, we obtain the squared exponential covariance function. The process associated is n-times m.s. differentiable if and only if $\nu > k$.

Figure 1.1 and Figure 1.2 show the Matérn covariance function and the process associated with different parameter ν.

Example 1.2 ***(γ-Exponential Covariance Function)***

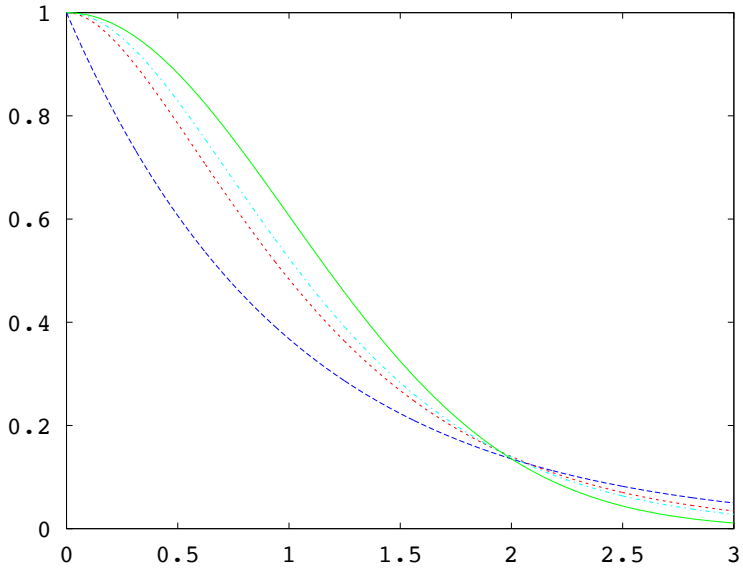

Figure 1.1: Matérn covariance functions for $\ell = 1$, $\sigma^2 = 1$ and different ν: blue for $\nu = \frac{1}{2}$, red for $\nu = \frac{3}{2}$, cyan for $\nu = \frac{5}{2}$ and green for $\nu \to +\infty$.

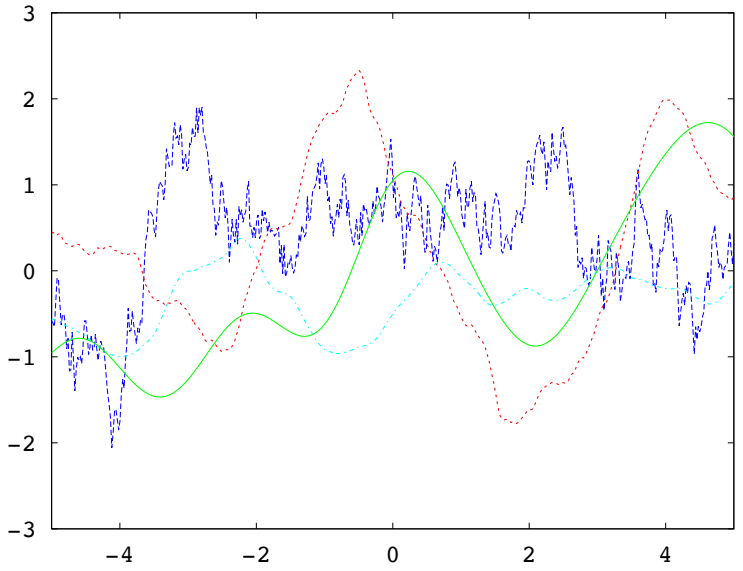

Figure 1.2: Realizations of the Gaussian process with Matérn covariance functions for $\ell = 1$, $\sigma^2 = 1$ and different ν: blue for $\nu = \frac{1}{2}$, red for $\nu = \frac{3}{2}$, cyan for $\nu = \frac{5}{2}$ and green for $\nu \to +\infty$.

The γ-exponential family covariance function is given by

$$K_\gamma(r) = \sigma^2 e^{-(\frac{r}{\ell})^\gamma} \qquad for \quad \gamma \in]0, 2] \tag{1.28}$$

where ℓ and σ^2 are positive parameters.

Although it is similar to the Matérn class, it is less flexible. The corresponding process is not m.s. differentiable except when $\gamma = 2$.

Figure 1.3 and Figure 1.4 show the γ-exponential covariance function and the process associated with different parameter γ.

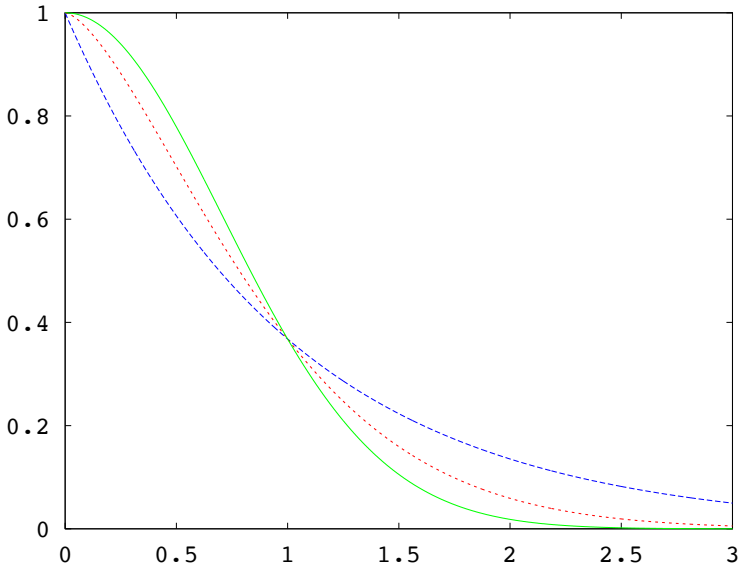

Figure 1.3: γ-exponential covariance functions for $\ell = 1$, $\sigma^2 = 1$ and different γ: blue for $\gamma = 1$, red for $\gamma = \frac{3}{2}$ and green for $\gamma = 2$.

Example 1.3 ***(Squared exponential covariance function)***

The squared exponential covariance function is defined by

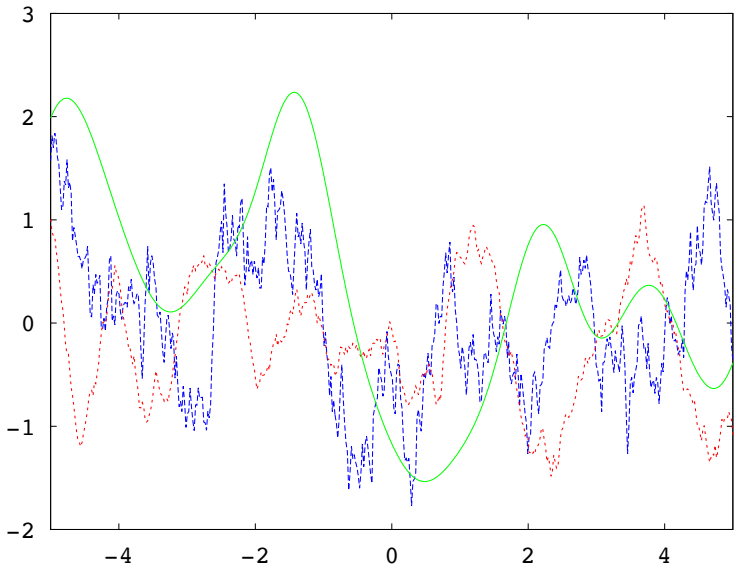

Figure 1.4: Realizations of the Gaussian process for γ-exponential covariance functions with $\ell = 1$, $\sigma^2 = 1$ and different γ: blue for $\gamma = 1$, red for $\gamma = \frac{3}{2}$ and green for $\gamma = 2$.

$$K_{SE}(r) = \sigma^2 e^{-\frac{r^2}{2\ell^2}} \tag{1.29}$$

with parameter $\ell > 0$ defining the characteristic length-scale, and σ^2 a positive parameter.

It is a particular case of the γ-exponential covariance function with $\gamma = 2$. It is also the limit for $\nu \to +\infty$ of the Matérn class covariance function. The corresponding process is infinitely mean square differentiable, which means that the Gaussian process with this covariance function has mean square derivatives of all orders and is thus very smooth. This is a model widely used in the kernel machines field.

Figure 1.5, Figure 1.6, Figure 1.7 and Figure 1.8 show the squared-exponential covariance function and the process associated with differ-

ent parameter ℓ.

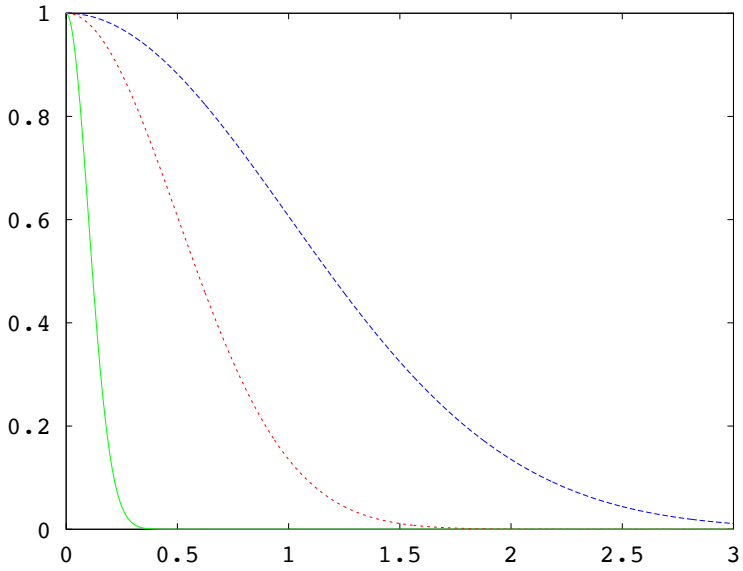

Figure 1.5: Squared exponential covariance functions with $\sigma^2 = 1$ and different ℓ: blue for $\ell = 1$, red for $\ell = 0.5$ and green for $\ell = 0.1$.

Example 1.4 ***(Exponential covariance function)***

The exponential covariance function is defined by

$$K_E(r) = \sigma^2 e^{-\frac{r}{\ell}} \tag{1.30}$$

with parameter ℓ and σ^2 positive.

It is a particular case of the γ-exponential covariance function with $\gamma = 1$. It is also in the Matérn class covariance family with parameter $\nu = \frac{1}{2}$.

The corresponding process is mean square continuous but not mean

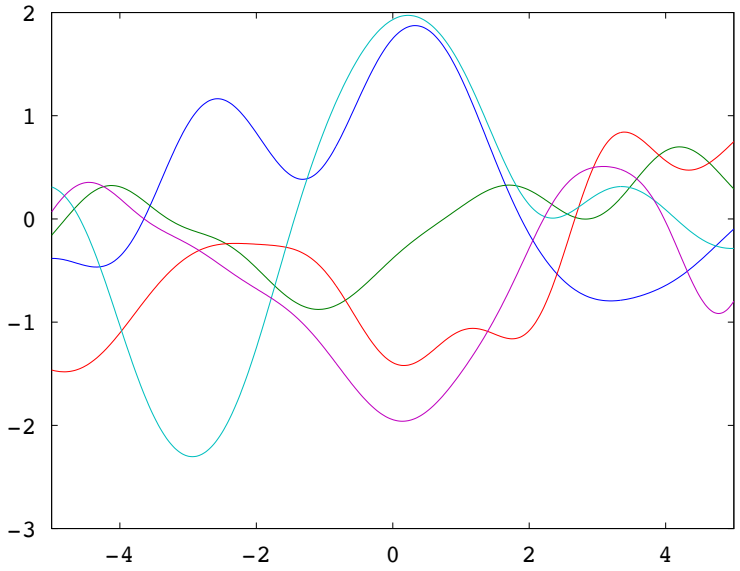

Figure 1.6: 5 realizations of the Gaussian process for squared exponential covariance functions with $\ell = 1$ and $\sigma^2 = 1$.

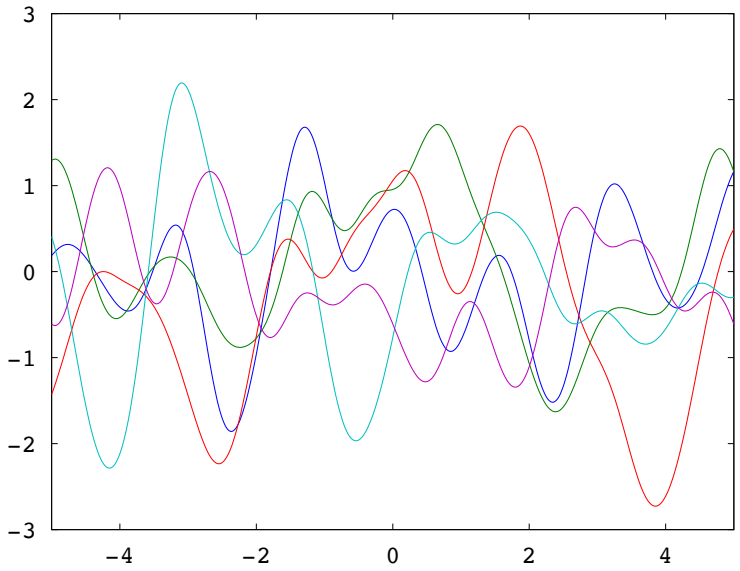

Figure 1.7: 5 realizations of the Gaussian process for squared exponential covariance functions with $\ell = 0.5$ and $\sigma^2 = 1$.

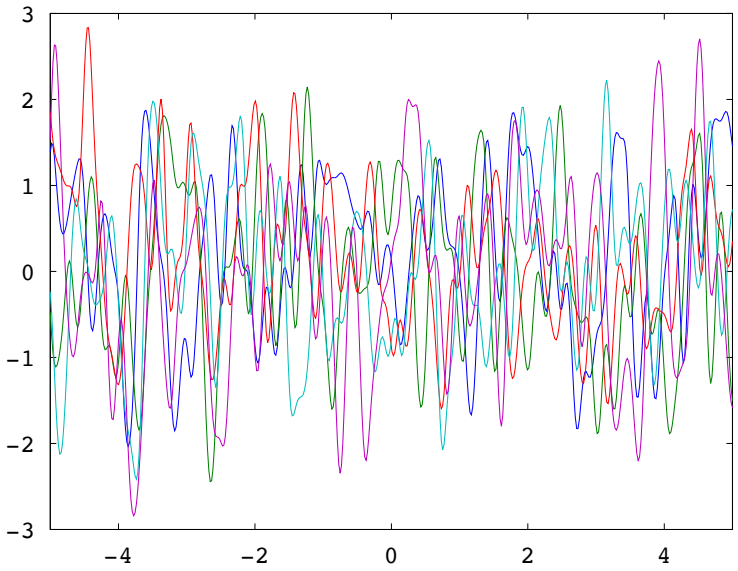

Figure 1.8: 5 realizations of the Gaussian process for squared exponential covariance functions with $\ell = 0.1$ and $\sigma^2 = 1$.

square differentiable.

Figure 1.9, Figure 1.10, Figure 1.11 and Figure 1.12 show the exponential covariance function and the process associated with different parameter ℓ.

Example 1.5 ***(Rational Quadratic Covariance Function)***

The rational quadratic (RQ) covariance function is given by

$$K_{RQ}(r) = \sigma^2(1 + \frac{r^2}{2\alpha\ell^2})^{-\alpha} \tag{1.31}$$

with $\alpha, \ell > 0$ and σ^2 a positive parameter.

It can be seen as a scale mixture (an infinite sum) of squared exponen-

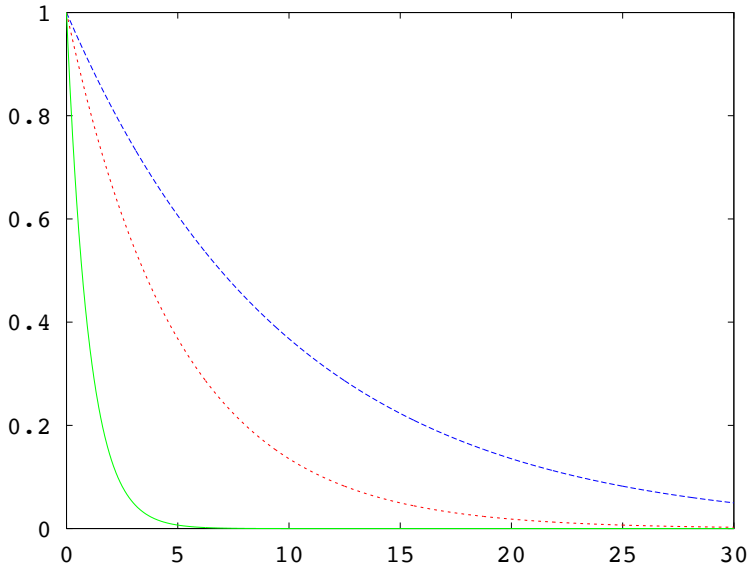

Figure 1.9: Exponential covariance functions with $\sigma^2 = 1$ and different ℓ: blue for $\ell = 10$, red for $\ell = 5$ and green for $\ell = 1$.

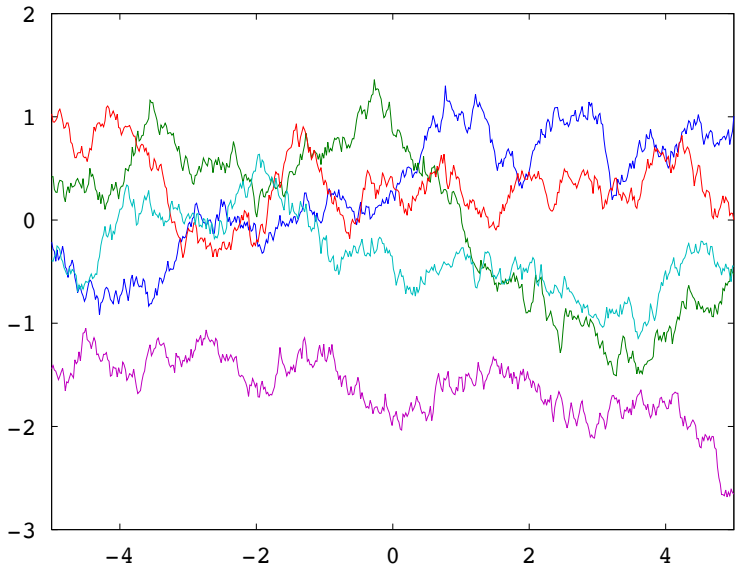

Figure 1.10: 5 realizations of the Gaussian process for exponential covariance functions with $\ell = 10$ and $\sigma^2 = 1$.

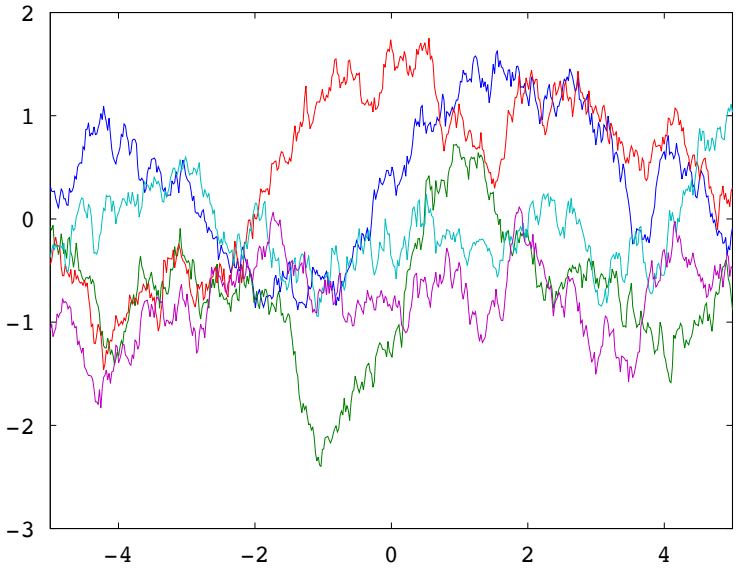

Figure 1.11: 5 realizations of the Gaussian process for exponential covariance functions with $\ell = 5$ and $\sigma^2 = 1$.

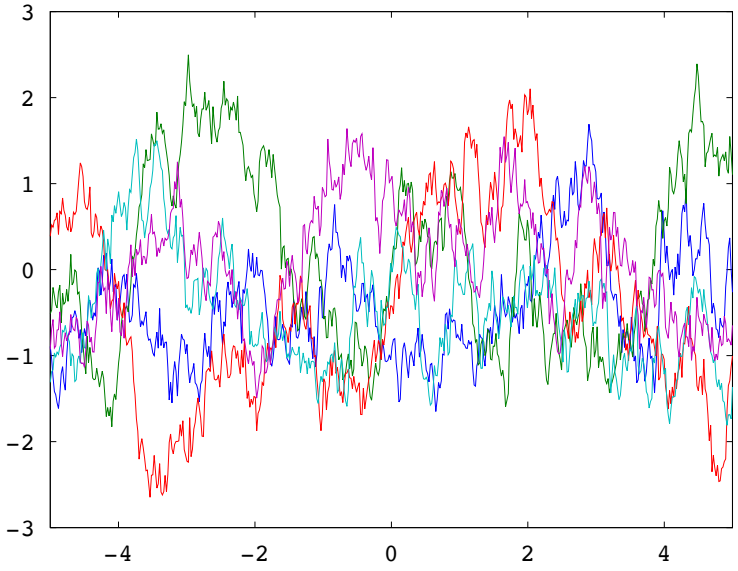

Figure 1.12: 5 realizations of the Gaussian process for exponential covariance functions with $\ell = 1$ and $\sigma^2 = 1$.

tial covariance function with different characteristic length-scales. The limit of the rational quadratic covariance for $\alpha \to \infty$ is the squared exponential covariance function with characteristic length-scale ℓ.

The Gaussian process defined by this kind of covariance function is infinitely mean square differentiable for every α in contrast to the Matérn covariance function.

1.8.3 Examples of Gaussian process

In the following, we will present some widely used Gaussian process.

Example 1.6 ***(Real value Brownian motion)***

The Brownian motion, also called the Wiener process, is a stochastic process $B = (B_t)_{t\geq 0}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ having the natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ associated to the process B such that:

- *B has continuous paths: with probability 1, B_t is continuous;*
- *B has independent increments: $\forall 0 \leq s \leq t$, the random variable $B_t - B_s$ is independent of $\mathcal{F}_s$;*
- *$\forall 0 \leq s \leq t$, the random variable $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$.*

We say that B is a standard Brownian motion if in addition $B_0 = 0$.

The standard Brownian motion B has the mean function $\mathbb{E}[B] = 0$ and the covariance function $K_B(s,t) = min(s,t)$.

Figure 1.13 shows some realizations of the Brownian motion.

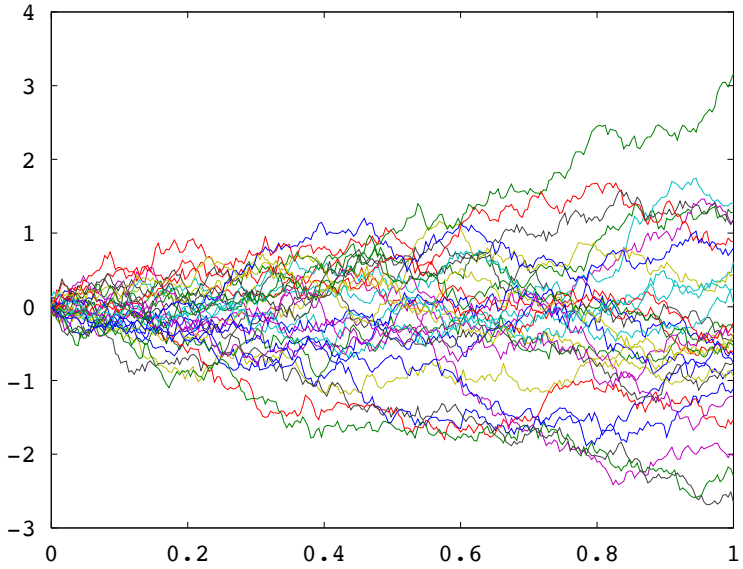

Figure 1.13: 30 realizations of the Brownian motion.

The Brownian motion plays an important role in many scientific domains. In pure mathematics, it is a key process in terms of which more complicated stochastic processes can be described. As such, it plays a crucial role in stochastic calculus, diffusion processes and even potential theory. In applied mathematics, the Brownian motion is used to represent the integral of a white noise Gaussian process, and used as a model of noise in electronics engineering, instrument errors in filtering theory and unknown forces in control theory. In physics, it is used to study the diffusion of particles suspended in fluid, and other types of diffusion via the Fokker–Planck and Langevin equations. It also forms the basis for the rigorous path integral formulation of quantum mechanics. In finance, it is also widely used, in particular in the Black–Scholes option pricing model.

Example 1.7 ***(Brownian bridge)***

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion and T>0. Then $(\Pi_t = B_t - \frac{t}{T}B_T)_{t\geq 0}$ is a Brownian bridge for $t \in [0, T]$. The Brownian bridge $(\Pi_t)_{t\in[0,1]}$ is a continuous-time Gaussian process having mean function $\mathbb{E}[\Pi_t] = 0$ and covariance function $K_\Pi(s,t) = min(s,t) - st$.

It is independent of B_T, in fact, the process is pinned both at $t = 0$ and $t = T$. The increments in a Brownian bridge are not independent.

Example 1.8 ***(Ornstein-Uhlenbeck process)***

There are several different representations of the Ornstein - Uhlenbeck process. A primary definition of an Ornstein - Uhlenbeck process is a stochastic process X_t satisfying the following stochastic differential equation:

$$dU_t = \theta(\mu - U_t)dt + \sigma dB_t \tag{1.32}$$

where $\theta > 0$, μ and $\sigma > 0$ are parameters and $(B_t)_{t\geq 0}$ denotes the Brownian motion.

Another possible and more convenient representation of the Ornstein-Uhlenbeck process is as a scaled time-transformed Brownian motion:

$$U_t = \mu + \frac{\sigma}{\sqrt{2\theta}} e^{-\theta t} B_{e^{2\theta t}} \tag{1.33}$$

where B is a Brownian motion.

For simplicity, let us choose $\mu = 0$, $\sigma = 1$ and $\theta = \frac{1}{2}$. It is simple to prove that $U_t = e^{-\frac{t}{2}} B_{e^t} \backsimeq \mathcal{N}(0,1)$, so the process is stationary. Its

covariance function is the exponential covariance function $K(s,t) = e^{-|t-s|/2}$ *(See Example 1.4).*

The Ornstein-Uhlenbeck process was introduced as a mathematical model of the velocity of a particle undergoing Brownian motion.

1.8.4 Gaussian process generator

Gaussian process can be thought of as generalizations of Gaussian random vectors. To generate a realization of a Gaussian process at $t_1, \cdots, t_N$ is to generate a Gaussian random vector with expectation function and covariance function at $t_1, \cdots, t_N$.

As such, the fundamental generation method is the same as for a multivariate normal random vector. We can sample a multivariate normal random vector $X = (X_1, \cdots, X_N)^T$ as follows:

Algorithm 1.1 ***(Gaussian process generator)***

1. *Construct the mean vector* $m = (\mathbb{E}_{t_1}, \cdots, \mathbb{E}_{t_N})^T$ *and covariance matrix* $\Sigma = (\Sigma_{ij})_{1 \le i,j \le N}$ *with* $\Sigma_{ij} = Cov(t_i, t_j)$*;*
2. *Calculate the square root of the covariance matrix by Cholesky decomposition* $\Sigma = AA^T$*;*
3. *Generate* n *i.i.d normal distributions* $Y_1, \cdots, Y_n \sim \mathcal{N}(0,1)$*. Let* $Y = (Y_1, \cdots, Y_n)^T$*.*
4. X *is obtained by* $X = m + AY$*.*

Chapter 2

Discrete time independent noise model with application to additive noise

The problem of recovering the Hamiltonian and dipole moment is considered in a bilinear quantum control framework. The process uses as inputs some measurable quantities (observables) for each admissible control. If the implementation of the control is noisy the data available is only in the form of probability laws of the measured observable. Nevertheless it is proved that the inversion process still has unique solutions (up to phase factors). Both additive and multiplicative noises are considered. Numerical illustrations support the theoretical results. This chapter reproduces the content of the accepted paper [18].

2.1 Introduction and motivation

Successful manipulation of quantum dynamics (see [9] and references therein for a recent review) leads to interesting perspectives among which is the possibility to identify the system through measurements of control-dependent observations. This technique, called quantum identification or quantum inversion, was documented both theoretically [4, 7, 29, 39] and numerically [14, 20, 31]. However although the numerical implementations show interesting robustness of the identification process with respect to noise, there is less theoretical guidance to explain this fact. Two fundamental questions concerning the well-posedness of this problem arise: the existence and the uniqueness of the Hamiltonian, and/or the dipole moment, compatible with the given measurements. In this work we only study the uniqueness.

More specifically we start from the setting in [29] which treats the case without noise. After some technical preliminaries in section 2.3 we address the noise-free case in section 2.4 and relax many of the assumptions used in the previous work. Then in section 2.5 we introduce the possibility that the control is subject at each time to unknown perturbations. We consider both additive and multiplicative noise. Since the actual control that acts on the system is unknown, only the probability laws of the observations are available. We explain which are the properties of the set of measurements required to determine uniquely (up to phase factors) the free Hamiltonian and dipole moment.

Then a numerical implementation is presented in section 2.6. Some closing remarks are the object of section 2.7.

2.1.1 Notations

We introduce the following notations

- $\mathbb{L}_{M_1,M_2,\cdots,M_m}$ is the Lie algebra spanned by the matrices $M_1, M_2, \cdots, M_m$;
- for any matrix or vector X we denote by $\overline{X}$ its conjugate (the matrix whose entries are the complex conjugates of the entries of X) and by X^* its adjoint (the transpose conjugate);
- $\mathfrak{H}$ is the set of all Hermitian matrices $\mathfrak{H} = \{X \in \mathbb{C}^{N\times N} | X^* = X\}$;
- $\mathcal{S}_N$ is the unit sphere of $\mathbb{C}^N$: $\mathcal{S}_N = \{v \in \mathbb{C}^N | \|v\| = 1\}$;
- $\Psi(t, H, u(\cdot), \mu, \Psi_0)$ is the solution of the equation (2.1) below; to simplify the notation, when there is no ambiguity, we denote it $\Psi(t)$;
- $\lambda_k(X)$, $k = 1, ..., N$ are the eigenvalues of $X \in \mathfrak{H}$ taken in increasing order; we also introduce $\phi_k(X)$ $k = 1, ..., N$ to be eigenvectors of X (forming an orthonormal basis of $\mathbb{C}^N$) corresponding to eigenvalues $\lambda_k(X)$; note that $Span\{\phi_k(X)\}$ may not be unique;
- $SU(N)$ is the special unitary group of degree N, which is the group of $N \times N$ unitary matrices with determinant 1;
- $\mathfrak{su}(N)$ is the Lie Algebra of skew-Hermitian matrices (the Lie algebra of $SU(N)$);

2.2 The model

We present the mathematical framework following closely the notations of the previous work [29].

Consider a controlled quantum system with time-dependent wave-function $\Psi(t)$ satisfying the Schrödinger equation:

$$\begin{cases} i\dot{\Psi}(t, H, u(\cdot), \mu, \Psi_0) = (H + u(t)\mu)\Psi(t, H, u(\cdot), \mu, \Psi_0) \\ \Psi(0, H, u(\cdot), \mu, \Psi_0) = \Psi_0, \end{cases} \tag{2.1}$$

where H is the internal ("free") Hamiltonian and μ the coupling operator between the control $u(t) \in \mathrm{L}^1_{loc}(\mathbb{R}_+; \mathbb{R})$ and the system. We work in a finite dimensional framework, therefore $H, \mu \in \mathfrak{H}$ for some $N \in \mathbb{N}^*$. The goal is to determine the matrix entries of H and μ from laboratory measurements of some observables depending on $\Psi(t)$. The control $u(t)$ can be changed in order to gather enough information on the system.

However, contrary to [29], we allow in this work some time independent perturbations to appear in the control $u(t)$. That is, when the control is implemented in practice the nominal control intensity required by the experimentalist, denoted $\epsilon(t)$, is perturbed by Y which means that $u = u(t, \epsilon(\cdot), Y)$; here Y is a discrete random variable with possible outcomes $y_1, y_2, ...$ We assume that the law of the random variable Y is time independent. A first example is the additive perturbation $u(t) = \epsilon(t) + Y$. Such perturbation models have already been used in the quantum computing literature under the name of "fixed systematic errors", see section VI.A. equation (40) of [27] or "systematic control error", see [28]. In [34] the authors use a noise model called "low frequency noise"(see section IV. C. of [24]): it is defined as the portion of

the (control) amplitude noise that has a correlation time that is long (up to 10^3 times) compared to the timescale of the dynamics therefore it can be considered as constant in time. Additional noise models (additive or multiplicative) are presented in [40].

The perturbation Y is unknown and thus $\Psi(t)$ is a random variable, as are all measurements depending on $\Psi(t)$. Repeating the control experiment several times the experimentalist will only learn the law of the measurements. From now on we will denote by $\mathcal{L}_Y Z$ the law of the random variable Z (that is measurable with respect to the sigma-algebra generated by Y).

Two different settings are considered depending on which parameters are to be identified and the nature of the information available:

- Setting **(S1)**: The Hamiltonian H is known and the goal is to identify the dipole moment μ.
- Setting **(S2)**: Both the Hamiltonian H and the dipole moment μ are unknown.

The measurements are of the form $\langle O\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ with $O \in \mathfrak{H}$ a member of a list of possible measurements. Often, the experimentalist only measures one observable in a list (but can repeat the experiment many times). This means that for general $O_1, O_2 \in \mathfrak{H}$ no information is available on the joint distribution of the values $\langle O_1\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ and $\langle O_2\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ of these two observables.

2.3 Some technical preliminaries

2.3.1 Complete sets of commuting observables

We recall in this section several facts about complete sets of commuting observables (hereafter abbreviated CSCO). We refer the reader to [10, page 146] for details.

First, recall that an observable is a self-adjoint operator on $\mathbb{C}^N$. Once a basis of $\mathbb{C}^N$ is chosen the observable can be represented as a matrix $O \in \mathfrak{H}$.

A set of observables $\mathcal{O} = \{O_1, ..., O_K\}$ is called *set of commuting observables* (named SCO hereafter) if $[O_k, O_\ell] = 0$, $\forall k, \ell \in \{1, ..., K\}$.

When all observables in the SCO are multiples of the identity operator the SCO is said to be trivial; unless specified otherwise, we only work with non-trivial SCO.

All observables in the SCO $\mathcal{O}$ can be diagonalized simultaneously i.e., there exists at least an orthonormal basis $\Phi = \{\phi_1, ..., \phi_N\}$ of $\mathbb{C}^N$ such that any $O \in \mathcal{O}$ is diagonal in the basis Φ. This means that in particular any ϕ_ℓ is an eigenvector of any observable $O \in \mathcal{O}$. In general the basis Φ is not unique because of possible degeneracies in the spectrum of the observables in $\mathcal{O}$. By definition a SCO is called a *complete set of commuting observables* (CSCO) if the orthonormal basis that diagonalizes the SCO is unique up to phase factors and permutations, i.e., if $\{\varphi_1, ..., \varphi_N\}$ is another orthonormal basis rendering all $O \in \mathcal{O}$ diagonal then there exists a permutation σ of $\{1, ..., N\}$ and phases $\beta_1, ..., \beta_N \in \mathbb{R}$ such that $\varphi_k = e^{i\beta_k}\phi_{\sigma(k)}$ for all $k = 1, ..., N$.

Examples:

1. Let H be a Hamiltonian with all eigenvalues $\lambda_\ell(H)$ of multiplicity 1. Then $\mathcal{O} = \{H\}$ is a CSCO.

2. Let $\{v_1, ..., v_N\}$ be an orthonormal basis of $\mathbb{C}^N$. Then defining P_k to be the projection on v_k (that is $P_k = v_k v_k^*$) the set $\mathcal{O} = \{P_k, 1 \leq k \leq N\}$ is a CSCO. In this case P_k are called populations of the states v_k.

3. Consider $N = 3$ and $\mathcal{O} = \{O_d\}$ with:

$$O_d = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}. \tag{2.2}$$

   Because the eigenspace corresponding to the eigenvalue $1/2$ is of dimension 2 $\mathcal{O}$ is not a CSCO. In this case both the canonical base of $\mathbb{C}^3$: $\{(1,0,0)^T, (0,1,0)^T, (0,0,1)^T\}$ and the orthonormal basis $\{(1,0,0)^T, (0,1/2,-\sqrt{3}/2)^T, (0,\sqrt{3}/2,1/2)^T\}$ render O_d diagonal.

4. Consider the truncated spin-less Hydrogen atom whose eigenstates can be labeled by a set of three indexes $\phi_{n,l,m}$ with $n = 1, 2, ..., N_t$, $l = 0, 1, ..., n-1$, $m = -l, -l+1, ..., l-1, l$. Here $N_t \in \mathbb{N}$ is a fixed truncation threshold. A CSCO is given by the operators H (Hamiltonian), L^2 (square of the angular momentum operator), L_z (the z component of the angular momentum operator) which act on the eigenstate $\phi_{n,l,m}$ as:

$$H\phi_{n,l,m} = \frac{C_H}{n^2}\phi_{n,l,m},\ L^2\phi_{n,l,m} = l(l+1)\hbar^2\phi_{n,l,m},\ L_z\phi_{n,l,m} = m\hbar\phi_{n,l,m}, \tag{2.3}$$

   with C_H an universal constant and $\hbar$ the Plank constant. Here n is called principal quantum number, l the angular momentum

quantum number and m the magnetic quantum number. Note that in this case $\{H, L^2\}$ is a SCO but not a CSCO.

Measuring simultaneously all observables in a CSCO is in principle possible as it is compatible with the Heisenberg uncertainty principle since all observables in a CSCO commute two by two; therefore the values of those observables may be simultaneously computed with infinite precision.

The following characterization of a CSCO will be used in the following sections:

Lemma 2.1 *Let $\mathcal{O} = \{O_1, ..., O_K\}$ be a SCO. Then $\mathcal{O}$ is a CSCO iff there exist $\gamma_1, ..., \gamma_K \in \mathbb{R}$ such that all eigenvalues of $\sum_{k=1}^K \gamma_k O_k$ have multiplicity one.*

Proof. We prove first the direct implication. Consider a basis $\Phi = \{\phi_1, ..., \phi_N\}$ of $\mathbb{C}^N$ that renders all O_k diagonal and denote $(O_k)_j$ the j-th eigenvalue of O_k, that is $O_k\phi_j = (O_k)_j\phi_j$. Suppose now by contradiction that for any $\gamma = (\gamma_1, ..., \gamma_K) \in \mathbb{R}^K$ there exists $i(\gamma) \neq j(\gamma) \leq N$, such that $\sum_{k=1}^K \gamma_k (O_k)_{i(\gamma)} = \sum_{k=1}^K \gamma_k (O_k)_{j(\gamma)}$. Define the functions $g^{\ell_1,\ell_2} : \mathbb{R}^K \to \mathbb{R}$ by $g^{\ell_1,\ell_2}(\gamma_1, ..., \gamma_K) = \sum_{k=1}^K \gamma_k[(O_k)_{\ell_1} - (O_k)_{\ell_2}]$ and let $A^{\ell_1,\ell_2} = \{\gamma \in \mathbb{R}^K; g^{\ell_1,\ell_2}(\gamma) = 0\}$. We obtain that $\cup_{1\leq\ell_1<\ell_2\leq N} A^{\ell_1,\ell_2} = \mathbb{R}^K$. By the Baire's theorem at least a couple (ℓ_1^*, ℓ_2^*) exists such that $A^{\ell_1^*,\ell_2^*}$ has non empty interior. Therefore the analytic function $g^{\ell_1^*,\ell_2^*}$ is null on a non empty open set hence it is null everywhere.

But this means that $(O_k)_{\ell_1^*} = (O_k)_{\ell_2^*}$ for all $k = 1, ..., K$. Therefore for all $O_k \in \mathcal{O}$ the ℓ_1^*-th eigenvalue is of multiplicity 2 (the ℓ_1^*-th eigenvector and the ℓ_2^*-th eigenvector are associated to the same eigenvalue) which contradicts the uniqueness of the basis that diagonalizes $\mathcal{O}$, and hence we obtain a contradiction with the definition of a CSCO.

The reverse implication is more straightforward. Any basis $\Phi = \{\phi_1, ..., \phi_N\}$

that renders all $O_k \in \mathcal{O}$ diagonal will also render $\sum_{k=1}^{K} \gamma_k O_k$ diagonal. But by hypothesis all eigenvalues of $\sum_{k=1}^{K} \gamma_k O_k$ are distinct and therefore the basis Φ is unique (up to permutation and phases) and hence $\mathcal{O}$ is a CSCO.

2.3.2 Background on controllability results

Let $L \in \mathbb{N}^*$ and $G_1, \cdots, G_L$ be L finite dimensional, connected, compact and simple Lie groups with the identity element Id. Let $A_\ell, B_\ell \in \mathfrak{g}_\ell$ for all $\ell = 1, \cdots, L$ where $\mathfrak{g}_\ell$ is the Lie algebra of G_ℓ.

Definition 2.1 *Consider L bilinear systems on the Lie groups G_ℓ:*

$$\begin{cases} \frac{dX_\ell(t)}{dt} = (A_\ell + u(t)B_\ell)X_\ell(t), \\ X_\ell(0) = Id. \end{cases} \tag{2.4}$$

The systems are called simultaneously controllable (or ensemble controllable) if there exists $T_{A_1,\cdots,A_L,B_1,\cdots,B_L} > 0$ such that for all $T \geq T_{A_1,\cdots,A_L,B_1,\cdots,B_L}$ and for all $V_\ell \in G_\ell$, $\ell = 1, \cdots, L$ arbitrary, there exists a control $u \in \mathrm{L}^1([0,T],\mathbb{R})$ with $X_\ell(T) = V_\ell$, $\forall \ell = 1, \cdots, L$.

Let $\mathcal{A} = A_1 \oplus \cdots \oplus A_L \in \oplus_{\ell=1}^{L} \mathfrak{g}_\ell$ and $\mathcal{B} = B_1 \oplus \cdots \oplus B_L \in \oplus_{\ell=1}^{L} \mathfrak{g}_\ell$. The following simultaneous controllability results are proved in [36, Theorems 1 & 2] and [6, Lemma 3 page 29].

Theorem 2.1 *The collection (2.4) of L bilinear systems is simultaneously controllable if and only if $\mathbb{L}_{\mathcal{A},\mathcal{B}} = \oplus_{\ell=1}^{L} \mathfrak{g}_\ell$ or equivalently $dim_{\mathbb{R}} \mathbb{L}_{\mathcal{A},\mathcal{B}} = \sum_{\ell=1}^{L} dim_{\mathbb{R}} \mathfrak{g}_\ell$.*

Lemma 2.2 *We suppose that $\mathbb{L}_{A_\ell,B_\ell} = \mathfrak{g}_\ell$, for all $\ell = 1, \cdots, L$. Then $\mathbb{L}_{\mathcal{A},\mathcal{B}} \neq \oplus_{\ell=1}^{L} \mathfrak{g}_\ell$ if and only if there exist $\ell, \ell' \in \{1, \cdots, L\}$, $\ell \neq \ell'$ and an isomorphism $f : \mathfrak{g}_\ell \to \mathfrak{g}_{\ell'}$ such that $f(A_\ell) = A_{\ell'}$ and $f(B_\ell) = B_{\ell'}$.*

Theorem 2.2 *Let G be a finite dimensional, connected, compact and simple Lie group and $\mathfrak{g}$ be its Lie algebra. Let $A, B \in \mathfrak{g}$ such that $\mathbb{L}_{A,B} = \mathfrak{g}$ and $\alpha_1,...,\alpha_L \in \mathbb{R}$ be real constants, $\alpha_i \neq \alpha_j \ \forall i \neq j$. Consider the collection of control systems on G:*

$$\begin{cases} \frac{dX_\ell(t)}{dt} = \{A + (u(t) + \alpha_\ell)B\}X_\ell(t), \\ X_\ell(0) = Id. \end{cases} \tag{2.5}$$

Then the collection of systems (2.5) is simultaneously controllable.

Remark 2.1 *Although the Theorems 2.1 and 2.2 are formulated on a Lie group, this is enough to obtain controllability for the wave-function; recall that if $X(t, H, u(\cdot), \mu) : \mathbb{R} \times \mathfrak{H} \times \mathrm{L}^1_{loc}(\mathbb{R}_+; \mathbb{R}) \times \mathfrak{H} \to SU(N)$ satisfies the following equation:*

$$\begin{cases} i\dot{X}(t, H, u(\cdot), \mu) = (H + u(t)\mu)X(t, H, u(\cdot), \mu), \\ X(0, H, u(\cdot), \mu) = Id, \end{cases} \tag{2.6}$$

then $\Psi(t, H, u(\cdot), \mu, \Psi_0) = X(t, H, u(\cdot), \mu)\Psi_0$ is the solution of (2.1). Since $SU(N)$ is transitive on the sphere $\mathcal{S}_N$ (see [13, page 88]), if the control system is controllable on the Lie group $SU(N)$ then it will also be controllable in the wave-function formulation.

2.4 Inversion without noise

Theorem 2.3 (Setting (S1)) *Let H, μ_1, $\mu_2 \in \mathfrak{H}$, H diagonal, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$ and denote for $a = 1, 2$ and $\epsilon \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$:*
$\Psi_a(t, \epsilon) = \Psi(t, H, \epsilon(\cdot), \mu_a, \Psi_0^a)$. *Let $\mathcal{O}$ be a (non-trivial) SCO. We suppose that $N \geq 3$ and:*

- ***(A1):*** $\mathbb{L}_{iH, i\mu_1} = \mathbb{L}_{iH, i\mu_2} = \mathfrak{su}(N)$.
- ***(A2):*** $tr(H) = tr(\mu_1) = tr(\mu_2) = 0$.
- *the eigenvalues of H are all of multiplicity one.*

Then there exists $T > 0$ such that if:

$$\langle O\Psi_1(T, \epsilon), \Psi_1(T, \epsilon)\rangle = \langle O\Psi_2(T, \epsilon), \Psi_2(T, \epsilon)\rangle \quad \forall \epsilon \in L^1([0, T]; \mathbb{R}), \forall O \in \mathcal{O}, \tag{2.7}$$

then for some $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$:

$$(\mu_1)_{jk} = e^{i(\alpha_j - \alpha_k)}(\mu_2)_{jk}, \ \forall j, k \leq N. \tag{2.8}$$

Remark 2.2 *1. Assumption **(A1)** is required for the simultaneous controllability, see theorem 2.1 and lemma 2.2.*

*2. The assumption **(A2)** can be made without loss of generality according to [29]. In fact, changing the Hamiltonian H and/ or dipole moment μ by adding a multiple of the identity operator Id, does not change the observations. In this case, the state $\Psi(t)$ is replaced by $e^{i\varphi}\Psi(t)$ with the phase $\varphi \in \mathbb{R}$ depending on $tr(H)$, $tr(\mu)$ and on the control u.*

Remark 2.3 *The proof also shows that the values of any additional observable commuting with H are identical for both systems, in particular all populations are always identical.*

Moreover, when μ_1 and μ_2 are matrices of dipole operators (i.e., have the form of real potentials) truncated to dimension N, then μ_1 and μ_2 are real symmetric matrices; the Theorem implies $(\mu_1)_{jk} = \pm(\mu_2)_{jk}$ for all j, k. In general this is not enough to conclude that $\mu_1 = \mu_2$ as it can be seen from the counter-example 1 from [29, page 381] where $N = 3$, $\Psi_0^1 = \Psi_0^2 = (1, 0, 0)^T$:

$$H = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_2 \end{pmatrix},$$

$$\mu_1 = \begin{pmatrix} 0 & -\mu_\alpha & 0 \\ -\mu_\alpha & 0 & \mu_\beta \\ 0 & \mu_\beta & 0 \end{pmatrix},$$

$$\mu_2 = \begin{pmatrix} 0 & \mu_\alpha & 0 \\ \mu_\alpha & 0 & \mu_\beta \\ 0 & \mu_\beta & 0 \end{pmatrix},$$

$$E_1, E_2, \mu_\alpha, \mu_\beta \in \mathbb{R} \text{ (arbitrary).}$$

In this case all control fields give rise to identical populations for both systems. This under-determination can be mitigated under additional hypothesis as in Remark 2.7.

Remark 2.4 *When eigenvalues of H are degenerate but $\mathcal{O}$ is a CSCO the theorem 2.4 below should be used instead.*

Proof. Consider the collection of two systems (H, μ_1) and (H, μ_2) seen as a control system on $SU(N) \oplus SU(N)$ with operators $iH \oplus iH$, $i\mu_1 \oplus i\mu_2 \in \mathfrak{su}(N) \oplus \mathfrak{su}(N)$. This collection can either be controllable or not. Denote $\mathcal{R}_t = \{(\Psi_1(t, \epsilon), \Psi_2(t, \epsilon)) | \epsilon \in L^1([0, t]; \mathbb{R})\}$, $\mathcal{R}_\infty = \cup_{t \geq 0} \mathcal{R}_t$. It is known (see [26, Theorem 6.5 item (ii) page 322]) that there exists T

such that $\mathcal{R}_T = \mathcal{R}_\infty$.

Since $\mathcal{O}$ is a non-trivial SCO it contains at least an observable, denoted O, that is not multiple of the identity. For this observable there exist $\Psi_x, \Psi_y \in \mathcal{S}_N$ such that $\langle O\Psi_x, \Psi_x\rangle \neq \langle O\Psi_y, \Psi_y\rangle$. But the condition (2.7) shows that no control ϵ exists that drives Ψ_0^1 to Ψ_x and Ψ_0^2 to Ψ_y; therefore the joint system $iH \oplus iH$, $i\mu_1 \oplus i\mu_2$ is not controllable simultaneously. Then it exists an automorphism of $\mathfrak{su}(N)$ that sends iH to iH and $i\mu_1$ to $i\mu_2$. But the automorphisms of $\mathfrak{su}(N)$ are of the form $X \in \mathfrak{su}(N) \mapsto WXW^{-1} \in \mathfrak{su}(N)$ or $X \in \mathfrak{su}(N) \mapsto W\overline{X}W^{-1} \in \mathfrak{su}(N)$ for some $W \in SU(N)$. Recall that the matrix H is real (because it is diagonal and in $\mathfrak{H}$). Consider first that one can find $W \in SU(N)$ such that $H = WHW^{-1}$ and $\mu_2 = W\mu_1 W^{-1}$. The first identity shows that $[H, W] = 0$ and therefore W is diagonal with the diagonal containing entries of the form $e^{i\alpha_\ell}$, $\ell \leq N$; the conclusion follows from the second identity. Consider now that there exists $W \in SU(N)$ such that $iH = W\overline{iH}W^{-1}$; then $[H^2, W] = 0$, thus W diagonal and therefore $H = -H$, impossible.

Remark 2.5 *The result is stronger than the Theorem 1 in [29, page 380] which requires:*

- *a stronger condition on the spectrum of H (the non-degenerate transition condition); recall that the transitions of H are called non-degenerate if the eigenvalues $\lambda_k(H)$ of H satisfy $\lambda_i(H) - \lambda_j(H) \neq \lambda_a(H) - \lambda_b(H)$ for all $(a, b) \neq (i, j)$. Here we only ask that the eigenvalues have multiplicity one.*

- *that observables in $\mathcal{O}$ are the populations (thus in particular $\mathcal{O}$ is a CSCO). Here a single non-trivial observable is enough.*

- *that the equality (2.7) take place at all times $T \geq 0$. Here only one time (large enough) is required.*

Theorem 2.4 (Setting (S2)) *Let $\mu_1, \mu_2, H_1, H_2 \in \mathfrak{H}$, $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$ and denote for $a = 1, 2$ and $\epsilon \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$: $\Psi_a(t, \epsilon) = \Psi(t, H_a, \epsilon(\cdot), \mu_a, \Psi_0^a)$. We suppose that $N \geq 3$ and the following assumptions hold true:*

(A1): *$\mathbb{L}_{iH_1, i\mu_1} = \mathbb{L}_{iH_2, i\mu_2} = \mathfrak{su}(N)$;*

(A2): *$tr(H_1) = tr(H_2) = tr(\mu_1) = tr(\mu_2) = 0$;*

Let $\mathcal{O} = \{O_1, ..., O_K\}$ be a CSCO and $\Phi = \{\phi_1, ..., \phi_N\}$ an orthonormal basis that diagonalizes $\mathcal{O}$.

Then there exists $T > 0$ such that if:

$$\begin{aligned} &\langle O_k \Psi_1(T, \epsilon), \Psi_1(T, \epsilon)\rangle = \langle O_k \Psi_2(T, \epsilon), \Psi_2(T, \epsilon)\rangle \\ &\quad \forall \epsilon \in L^1([0, T]; \mathbb{R}), \forall k = 1, \cdots, K, \end{aligned} \tag{2.9}$$

then there exist $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$ and $\theta \in \mathbb{R}$ such that for all $j, k \leq N$ either

$$\begin{cases} \langle \mu_1 \phi_j, \phi_k\rangle = e^{i(\alpha_j - \alpha_k)} \langle \mu_2 \phi_j, \phi_k\rangle, \\ \langle H_1 \phi_j, \phi_k\rangle = e^{i(\alpha_j - \alpha_k)} \langle H_2 \phi_j, \phi_k\rangle, \\ \langle \Psi_0^1, \phi_j\rangle = e^{i(\theta - \alpha_j)} \langle \Psi_0^2, \phi_j\rangle, \end{cases} \tag{2.10}$$

or

$$\begin{cases} \langle \mu_1 \phi_j, \phi_k\rangle = -e^{i(\alpha_j - \alpha_k)} \overline{\langle \mu_2 \phi_j, \phi_k\rangle}, \\ \langle H_1 \phi_j, \phi_k\rangle = -e^{i(\alpha_j - \alpha_k)} \overline{\langle H_2 \phi_j, \phi_k\rangle}, \\ \langle \Psi_0^1, \phi_j\rangle = e^{i(\theta - \alpha_j)} \overline{\langle \Psi_0^2, \phi_j\rangle}. \end{cases} \tag{2.11}$$

Remark 2.6 *When $\mathcal{O}$ is not a CSCO, the same proof allows only to obtain that an isomorphism of Lie algebras exists that sends iH_1 to iH_2 and $i\mu_1$ to $i\mu_2$. In general it is not possible to obtain more than a*

general isomorphism as shown by the following counter-example:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix},$$

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & \sqrt{3}/2 \\ 0 & -\sqrt{3}/2 & 1/2 \end{pmatrix},$$

$$\mu_1 = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix},$$

$$H_2 = W H_1 W^{-1} = H_1,$$

$$\mu_2 = W \mu_1 W^{-1} = \begin{pmatrix} 0 & \sqrt{3}+1/2 & 1-\sqrt{3}/2 \\ \sqrt{3}+1/2 & 0 & 0 \\ 1-\sqrt{3}/2 & 0 & 0 \end{pmatrix},$$

$$\Psi_0^2 = W \Psi_0^1,$$

$$\mathcal{O} = \{O_d\}.$$

It is immediate to see that $\langle O_d \psi, \psi \rangle = 1/2 - 3/2 |\langle \psi, (1,0,0)^T \rangle|^2$ *and that* $W(1,0,0)^T = (1,0,0)^T$. *When the control* ϵ *on the first system realizes the transformation* X *the observable is* $1/2 - 3/2 |\langle X\Psi_0^1, (1,0,0)^T \rangle|^2$; *at the same time the control realizes the transformation* WXW^{-1} *on the second system giving the observable* $1/2 - 3/2 |\langle WXW^{-1} W\Psi_0^1, (1,0,0)^T \rangle|^2$. *But* $\langle WXW^{-1} W\Psi_0^1, (1,0,0)^T \rangle = \langle X\Psi_0^1, W^{-1}(1,0,0)^T \rangle = \langle X\Psi_0^1, (1,0,0)^T \rangle$. *Therefore it is not possible to distinguish between the couple* (H_1, μ_1) *and* (H_2, μ_2) *(at least for this initial data).*

Remark 2.7 *If, for physical reasons, we know the initial state of the*

system, then $\Psi_0^1 = \Psi_0^2$*; when this initial state has non-zero components along every element of the basis, i.e.* $\langle \Psi_0^1, \phi_k \rangle \neq 0$*, for all* $k \geq 1$ *the equation (2.10) implies* $e^{i\alpha_j} = e^{i\theta}$ *for all* $j = 1, \cdots, N$ *which means* $H_1 = H_2$ *and* $\mu_1 = \mu_2$*. When some coefficients* $\langle \Psi_0^1, \phi_k \rangle$ *are zero, further symmetries may occur and one can have, for instance,* $\mu_1 \neq \mu_2$*: see counter-example 1 from [29, page 381] presented in Remark 2.3.*

On the other hand, in this case, the conclusion (2.11) can be written more conveniently in the adapted basis $\{v_1 = e^{-i\alpha_1/2}\phi_1, \cdots, v_N = e^{-i\alpha_N/2}\phi_k\}$*:*

$$\begin{cases} \langle \mu_1 v_j, v_k \rangle = -\overline{\langle \mu_2 v_j, v_k \rangle}, \\ \langle H_1 v_j, v_k \rangle = -\overline{\langle H_2 v_j, v_k \rangle}, \\ \Psi_0^1 = \Psi_0^2 = e^{i\theta/2} \sum_{\ell=1}^N \varsigma_\ell v_\ell, \ \varsigma_\ell \in \mathbb{R}. \end{cases} \tag{2.12}$$

Proof. Denote by T the time at which the couple of systems (H_1, μ_1), (H_2, μ_2), seen as a control system on $SU(N) \oplus SU(N)$ with operators $iH \oplus iH$, $i\mu_1 \oplus i\mu_2 \in \mathfrak{su}(N) \oplus \mathfrak{su}(N)$ reaches all attainable states. Since a CSCO is a non-trivial SCO it follows as in the theorem 2.3 that there exists an isomorphism of $f : \mathfrak{su}(N) \to \mathfrak{su}(N)$ such that $iH_2 = f(iH_1)$, $i\mu_2 = f(i\mu_1)$.

All isomorphisms of $\mathfrak{su}(N)$ are of the form $\mathfrak{X} \in \mathfrak{su}(N) \mapsto W\mathfrak{X}W^{-1} \in \mathfrak{su}(N)$ or $\mathfrak{X} \in \mathfrak{su}(N) \mapsto W\overline{\mathfrak{X}}W^{-1} \in \mathfrak{su}(N)$ for some $W \in SU(N)$. We only treat here the 'exotic' case $f(\mathfrak{X}) = W\overline{\mathfrak{X}}W^{-1}$ as the second alternative is similar. Thus $H_2 = -W\overline{H_1}W^{-1}$ and $\mu_2 = -W\overline{\mu_1}W^{-1}$. With the notations in the equation (2.6) we write:

$$\begin{aligned} X(t, H_2, u(\cdot), \mu_2) =& X(t, -W\overline{H_1}W^{-1}, u(\cdot), -W\overline{\mu_1}W^{-1}) \\ =& W\overline{X(t, H_1, u(\cdot), \mu_1)}W^{-1}. \end{aligned} \tag{2.13}$$

As the first system is controllable then every state $X \in SU(N)$ can be

reached by some control $u(\cdot)$ thus

$$\langle O_k X\Psi_0^1, X\Psi_0^1\rangle = \langle O_k W\overline{X}W^{-1}\Psi_0^2, W\overline{X}W^{-1}\Psi_0^2\rangle, \forall X \in SU(N), \forall k \leq K. \tag{2.14}$$

Note that (2.14) also holds for any linear combination of observables in $\mathcal{O}$. We invoke the lemma 2.1 and obtain the existence of an observable O, diagonal in the basis Φ and with all eigenvalues distinct, such that

$$\langle OX\Psi_0^1, X\Psi_0^1\rangle = \langle OW\overline{X}W^{-1}\Psi_0^2, W\overline{X}W^{-1}\Psi_0^2\rangle, \forall X \in SU(N). \tag{2.15}$$

The vectors ϕ_k are eigenvectors of O and denote as $\lambda_k(O)$ the corresponding eigenvalues. In particular $O = \sum_{k=1}^N \lambda_k(O)\phi_k\phi_k^*$. We can suppose that $\lambda_1(O) < \lambda_2(O) < ... < \lambda_N(O)$ (otherwise re-index the vectors).

Let us write $\overline{W^{-1}\Psi_0^2} = x\Psi_0^1 + yv$ with $x, y \in \mathbb{C}$, $|x|^2 + |y|^2 = 1$, $v \in \mathcal{S}_N$, $v \perp \Psi_0^1$.

Suppose $y \neq 0$; then there exists $X \in SU(N)$ such that $X\Psi_0^1 = \phi_N$ and $Xv \in Span\{\overline{W^{-1}\phi_N}, \phi_N\}^\perp$. Then $\langle OX\Psi_0^1, X\Psi_0^1\rangle = \lambda_N(O) = \langle OW\overline{X}W^{-1}\Psi_0^2, W\overline{X}W^{-1}\Psi_0^2\rangle$. Since $\lambda_N(O)$ is the maximum possible value for O and all eigenspaces of O are of dimension 1 it follows that $W\overline{X}W^{-1}\Psi_0^2 \in Span\{\phi_N\}$ hence $X\overline{W^{-1}\Psi_0^2} \in Span\{\overline{W^{-1}\phi_N}\}$. Then:

$$\begin{aligned} 1 &= |\langle X\overline{W^{-1}\Psi_0^2}, \overline{W^{-1}\phi_N}\rangle| = |\langle X(x\Psi_0^1 + yv), \overline{W^{-1}\phi_N}\rangle| \\ &= |\langle x\phi_N, \overline{W^{-1}\phi_N}\rangle| = |x||\langle \phi_N, \overline{W^{-1}\phi_N}\rangle|. \end{aligned} \tag{2.16}$$

It follows $y = 0$, $|x| = 1$; therefore $\overline{W^{-1}\Psi_0^2} \in Span\{\Psi_0^1\}$ which means that there exists $\theta \in \mathbb{R}$ such that $\Psi_0^2 = e^{i\theta}W\overline{\Psi_0^1}$; after trivial simplifications the equality (2.15) can be written

$$\langle OX\Psi_0^1, X\Psi_0^1\rangle = \langle OW\overline{X\Psi_0^1}, W\overline{X\Psi_0^1}\rangle, \forall X \in SU(N). \tag{2.17}$$

But $X \in SU(N)$ means that $X\Psi_0^1$ can be chosen arbitrary in $\mathcal{S}_N$; we

have therefore:

$$\forall w \in \mathcal{S}_N : \langle Ow, w\rangle = \langle OW\overline{w}, W\overline{w}\rangle = \langle W^*OW\overline{w}, \overline{w}\rangle = \langle \overline{W^*OW}w, w\rangle. \tag{2.18}$$

But this implies $O = \overline{W^*OW}$ and thus:

$$\begin{aligned}\sum_{k=1}^{N} \lambda_k(O)\phi_k\phi_k^* = O = \overline{W^*OW} &= \overline{W^*\left(\sum_{k=1}^{N} \lambda_k(O)\phi_k\phi_k^*\right)W} \\ &= \sum_{k=1}^{N} \lambda_k(O)(\overline{W}^*\overline{\phi_k})(\overline{W}^*\overline{\phi_k})^*.\end{aligned} \tag{2.19}$$

Since all eigenvalues of O are non-degenerate the representation $O = \sum_{k=1}^{N} \lambda_k(O)\phi_k\phi_k^*$ is unique up to phases. Therefore there exist $\alpha_k \in \mathbb{R}$ such that $\overline{W}^*\overline{\phi_k} = e^{-i\alpha_k}\phi_k$ or, equivalently, $W^*\phi_k = e^{i\alpha_k}\overline{\phi_k}$.

The conclusion follows from the relationships $H_2 = -W\overline{H_1}W^*$, $\mu_2 = -W\overline{\mu_1}W^*$ and $\Psi_0^2 = e^{i\theta}W\overline{\Psi_0^1}$.

2.5 Inversion in presence of noise

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a discrete probability space, $\mathcal{V} = \{y_\ell \in \mathbb{R}^d | \ell \in \mathcal{I} \subset \mathbb{N}\}$ a set of values in $\mathbb{R}^d$ (possibly infinite) and let $Y : \Omega \to \mathcal{V}$ be a random variable. We can suppose that for all $y_\ell \in \mathcal{V}$, $\mathbb{P}(Y = y_\ell) > 0$ (otherwise we eliminate all y_ℓ such that $\mathbb{P}(Y = y_\ell) = 0$). Moreover after re-indexing $\mathcal{I}$ we can suppose that $\mathcal{I} = \mathbb{N}^*$ or $\mathcal{I} = \{1, \cdots, L_0\}$ for some $L_0 \in \mathbb{N}^*$. Denote $\xi_k = \mathbb{P}(Y = y_k)$, $\forall k \in \mathcal{I}$.

We can suppose that $(\xi_\ell)_{\ell \geq 1}$ is a decreasing sequence (re-indexing if necessary).

2.5.1 Technical preliminaries: a correspondence lemma

Let $J_a : \mathbb{C}^{N \times N} \to \mathbb{R}$, $a = 1, 2$ and $h : \mathbb{R}^{d+1} \to \mathbb{R}$ be real analytic functions with J_a bounded.

Lemma 2.3 *Let $A_a, B_a \in \mathfrak{su}(N)$, $T > 0$, $\epsilon \in \mathrm{L}^1([0, T], \mathbb{R})$ and denote by $X_a(t, y_\ell, \epsilon)$ the solution of*

$$\begin{cases} \frac{dX_a(t, y_\ell, \epsilon)}{dt} = (A_a + h(\epsilon(t), y_\ell) B_a) X_a(t, y_\ell, \epsilon) \\ X_a(0, y_\ell, \epsilon) = Id, \end{cases} \tag{2.20}$$

for $a = 1, 2$ and any $\ell \in \mathcal{I}$. Suppose that the following equality in law holds

$$\mathcal{L}_Y(J_1(X_1(T, Y, \epsilon))) = \mathcal{L}_Y(J_2(X_2(T, Y, \epsilon))) \quad \forall \epsilon \in \mathrm{L}^1([0, T], \mathbb{R}). \tag{2.21}$$

Then for any $\ell \in \mathcal{I}$, there exists $n_0(\ell, \xi_1, \cdots, \xi_n, \cdots)$ and $\kappa(\ell) \in \mathcal{I}$, $\kappa(\ell) \leq n_0(\ell, \xi_1, \cdots, \xi_n, \cdots)$ such that

$$J_1(X_1(T, y_\ell, \epsilon)) = J_2(X_2(T, y_{\kappa(\ell)}, \epsilon)) \quad \forall \epsilon \in \mathrm{L}^1([0, T], \mathbb{R}). \tag{2.22}$$

Proof. Let $\ell \in \mathcal{I}$. The proof is divided in several steps.

Step 1:

Fix a control ϵ. We introduce the notation:

$$v_a^k = J_a(X_a(T, y_k, \epsilon)), \qquad a = 1, 2 \quad \textit{and} \quad k \in \mathcal{I}.$$

According to the assumption (2.21) we know that $J_1(X_1(T, Y, \epsilon))$ and $J_2(X_2(T, Y, \epsilon))$ follow the same law. Thus $\mathbb{P}(J_1(X_1(T, Y, \epsilon)) = v_1^\ell) = \mathbb{P}(J_2(X_2(T, Y, \epsilon)) = v_1^\ell)$. Then

$$\sum_{k' \in \mathcal{I}/v_2^{k'} = v_1^\ell} \xi_{k'} = \sum_{k \in \mathcal{I}/v_1^k = v_1^\ell} \xi_k \geq \xi_\ell > 0. \tag{2.23}$$

Therefore $\{k' \in \mathcal{I}/v_2^{k'} = v_1^\ell\} \neq \emptyset$. In addition, there exists a $n_0(\ell) \in \mathcal{I}$ such that $\sum_{k > n_0(\ell), k \in \mathcal{I}} \xi_k < \xi_\ell$ and $\sum_{k > n_0(\ell)-1, k \in \mathcal{I}} \xi_k \geq \xi_\ell$ (by convention a sum over an empty set of indexes is zero). So we have $\{k' \in \mathcal{I}/k' \leq n_0(\ell), v_2^{k'} = v_1^\ell\} \neq \emptyset$. The index $n_0(\ell)$ depends only on the law of Y and the index ℓ.

Letting ϵ vary in $L^1([0, T]; \mathbb{R})$ we obtain a function $\kappa_1 : L^1([0, T]; \mathbb{R}) \to \{k \in \mathcal{I}, k \leq n_0(\ell)\}$ such that

$$J_1(X_1(T, y_\ell, \epsilon)) = J_2(X_2(T, y_{\kappa_1(\epsilon)}, \epsilon)). \tag{2.24}$$

Note: when the index $\kappa_1(\epsilon)$ with the property (2.24) is not unique, any compatible value in the set $\{k \in \mathcal{I}, k \leq n_0(\ell)\}$ can be chosen.

Step 2:

Let $n \in \mathbb{N}^*$. We consider the space $\mathcal{P}_n$ of piecewise constant controls $\mathcal{P}_n = \{f : [0, T] \to \mathbb{R} | \quad f = \alpha_1 \mathbf{1}_{[0, \frac{T}{n}]} + \alpha_2 \mathbf{1}_{]\frac{T}{n}, \frac{2T}{n}]} + \cdots + \alpha_n \mathbf{1}_{]\frac{(n-1)}{n}T, T]}, \alpha_1, \cdots, \alpha_n \in \mathbb{R}\}$. Denote $\alpha = (\alpha_1, \cdots, \alpha_n)$. Therefore for any $k \in \mathcal{I}$, we can define the functions g_k from $\mathbb{R}^n$ to $\mathbb{R}$ by

$$g_k(\alpha) = J_2(X_2(T, y_k, \epsilon_\alpha)) - J_1(X_1(T, y_\ell, \epsilon_\alpha)),$$

with $\epsilon_\alpha = \alpha_1 \mathbf{1}_{[0,\frac{T}{n}]} + \alpha_2 \mathbf{1}_{]\frac{T}{n},\frac{2T}{n}]} + \cdots + \alpha_n \mathbf{1}_{]\frac{(n-1)}{n}T,T]}$. We know that

$$X_a(T, y_k, \epsilon_\alpha) = e^{(A_a + h(\alpha_n, y_k)B_a)\frac{T}{n}} e^{(A_a + h(\alpha_{n-1}, y_k)B_a)\frac{T}{n}} \cdots e^{(A_a + h(\alpha_1, y_k)B_a)\frac{T}{n}} \tag{2.25}$$

for $a = 1, 2$. Therefore the functions X_a are analytic in α (recall that the function h is analytic in α), and since J_a are analytic, the functions g_k are analytic. We denote $A_k = \{\alpha \in \mathbb{R}^n / g_k(\alpha) = 0\}$. Each A_k is closed because g_k is continuous. In Step 1, it is proved that

$$\exists \kappa^{\mathcal{P}} : \mathbb{R}^n \to \{k \in \mathcal{I}, k \leq n_0(\ell)\} \quad \text{such} \quad \text{that} \quad \forall \alpha \in \mathbb{R}^n \quad g_{\kappa^{\mathcal{P}}(\alpha)}(\alpha) = 0. \tag{2.26}$$

So $\cup_{k \in \mathcal{I}, k \leq n_0(\ell)} A_k = \mathbb{R}^n$. By the Baire's theorem, it exists a k such that A_k has an interior point. This means that g_k is analytic and identically zero on a not empty open set. Therefore, $g_k \equiv 0$. So $\forall n$, $\exists \kappa_2(n) \in \{k \in \mathcal{I}, k \leq n_0(\ell)\}$ such that $g_{\kappa_2(n)}(\epsilon) = 0$, for any control $\epsilon \in \mathcal{P}_n$.

Step 3:
Take $q \in \mathbb{N}$ and denote $B_q = \{k \in \mathcal{I}, k \leq n_0(\ell)\} / g_k(\epsilon) = 0, \quad \forall \epsilon \in \mathcal{P}_{2^q}\}$. In Step 2 it is proved that for any $q \in \mathbb{N}$ the set B_q is not empty. Obviously $(B_q)_{q \in \mathbb{N}}$ is a decreasing sequence and B_q becomes constant from a certain term, thus $B_\infty = \cap_{q \geq 0} B_q \neq \emptyset$. This means that there exists $\kappa(\ell) \in \{k \in \mathcal{I}, k \leq n_0(\ell)\}$ such that $g_{\kappa(\ell)}(\epsilon) = 0$, $\forall \epsilon \in \mathcal{P}_{2^q}$ for all q. Yet, $\cup_{q=0}^{\infty} \mathcal{P}_{2^q}$ is dense in $L^1([0, T]; \mathbb{R})$. So we have $g_{\kappa(\ell)}(\epsilon) = 0$, for any control ϵ in $L^1([0, T]; \mathbb{R})$.

2.5.2 Main results

We set $d = 1$.

Theorem 2.5 *Consider the same setting and assumptions as in the theorem 2.4 with the exception of the relation (2.9). Then there exists*

$T > 0$ such that if:

$$\mathcal{L}_Y\langle O_k\Psi_1(T,\epsilon+Y),\Psi_1(T,\epsilon+Y)\rangle = \mathcal{L}_Y\langle O_k\Psi_2(T,\epsilon+Y),\Psi_2(T,\epsilon+Y)\rangle \tag{2.27}$$

$$\forall\epsilon\in L^1([0,T];\mathbb{R}),\quad \forall k=1,\cdots,K,$$

then either the conclusion (2.10) or the conclusion (2.11) of the theorem 2.4 holds (see also Remark 2.7).

Remark 2.8 *When $\mathcal{O}$ is not a CSCO, the same proof allows only to obtain that an isomorphism of Lie algebras exists that sends iH_1 to iH_2 and $i\mu_1$ to $i\mu_2$.*

Remark 2.9 *Relation (2.27) does not imply that for any $\gamma_k\in\mathbb{R}$:*

$$\begin{aligned}&\forall\epsilon\in L^1([0,T];\mathbb{R}):\\ &\mathcal{L}_Y\left\langle\left(\sum_{k=1}^K\gamma_kO_k\right)\Psi_1(T,\epsilon+Y),\Psi_1(T,\epsilon+Y)\right\rangle\\ &=\mathcal{L}_Y\left\langle\left(\sum_{k=1}^K\gamma_kO_k\right)\Psi_2(T,\epsilon+Y),\Psi_2(T,\epsilon+Y)\right\rangle,\end{aligned}\tag{2.28}$$

because the probability laws are not additive. This is in contrast with the situation in the theorem 2.4 (see the equations (2.14) and (2.15)). But the relation remains true for any operator of the form aO_k+bId, $a,b\in\mathbb{R}$.

Proof. Choose $m_0\in\mathbb{N}$ such that $m_0\xi_1>1$. Then there exists some $\eta>0$ small enough with $\sum_{m=0}^{m_0}(\xi_1-m\eta)>1$ and $\xi_1-m_0\eta>\eta$. As $\sum_{k\in\mathcal{I}}\xi_k=1$, there exists $n_0'\in\mathcal{I}$ such that $\sum_{k\in\mathcal{I},k>n_0'}\xi_k<\eta$ (by convention a sum over an empty set of indexes is zero). According to Definition 2.1, for all $\ell,\ell'\in\{1,\cdots,n_0'\}$, if the collection of 2 systems (2.4) for $A_1=-iH_1+y_\ell(-i\mu_1)\in\mathfrak{su}(N)$, $B_1=-i\mu_1\in\mathfrak{su}(N)$ and

$A_2 = -iH_2 + y_{\ell'}(-i\mu_2) \in \mathfrak{su}(N)$, $B_2 = -i\mu_2 \in \mathfrak{su}(N)$ is simultaneously controllable, then there exists $T_{H_1,H_2,\mu_1,\mu_2,y_\ell,y_{\ell'}} > 0$ such that the collection is simultaneously controllable at all times $T \geq T_{H_1,H_2,\mu_1,\mu_2,y_\ell,y_{\ell'}}$. If the collection is not controllable, we take $T_{H_1,H_2,\mu_1,\mu_2,y_\ell,y_{\ell'}}$ to be the time required to control one system (to any target). According to theorem 2.2, we know that the collection of n_0' systems (2.5) with $A = -iH_2$, $B = -i\mu_2$ and $(\alpha_1, \cdots, \alpha_{n_0'}) = (y_1, \cdots, y_{n_0'})$ is simultaneously controllable therefore there exists $T_{H_2,\mu_2,y_1,\cdots,y_{n_0'}}$ such that the collection is simultaneously controllable at any time $T \geq T_{H_2,\mu_2,y_1,\cdots,y_{n_0'}}$. Let $T = max_{1\leq\ell,\ell'\leq n_0'}(T_{H_2,\mu_2,y_1,\cdots,y_{n_0'}}, T_{H_1,H_2,\mu_1,\mu_2,y_\ell,y_{\ell'}})$. Suppose that the observations follow the same law at time T. Recall that $\Psi_a(T, \epsilon + y_\ell) = X_a(T, \epsilon + y_\ell)\Psi_0^a$ with $X_a(t, \epsilon + y_\ell)$ solutions of (2.20) where $A_a = -iH_a$, $B_a = -i\mu_a$, for $a = 1, 2$ respectively and $h(\epsilon(t), y_\ell) = \epsilon(t) + y_\ell$.

The second part of the remark 2.9 implies that we can suppose, without loss of generality, that any $\tilde{O} \in \mathcal{O}$ has the smallest eigenvalue equal to 0 and the largest one equal to 1. Fix now $\tilde{O} \in \mathcal{O}$. We apply the lemma 2.3 to $X \mapsto \langle \tilde{O}X\Psi_0^a, X\Psi_0^a\rangle$, $a = 1, 2$ which are obviously analytic with respect to X. Then for all $\ell \in \mathcal{I}$, $\exists\kappa(\ell)$ such that

$$\langle \tilde{O}\Psi_1(T, \epsilon + y_\ell), \Psi_1(T, \epsilon + y_\ell)\rangle = \langle \tilde{O}\Psi_2(T, \epsilon + y_{\kappa(\ell)}), \Psi_2(T, \epsilon + y_{\kappa(\ell)})\rangle. \tag{2.29}$$

Recall equation (2.23) in the lemma 2.3 (we use the same notations):

$$\sum_{k\in\mathcal{I}/v_1^k=v_1^\ell} \xi_k = \sum_{k'\in\mathcal{I}/v_2^{k'}=v_1^\ell} \xi_{k'},$$

for any control $\epsilon(t) \in \mathrm{L}^1([0, T], \mathbb{R})$. Now let us take $\ell = 1$. In the lemma 2.3 we proved that $\kappa(1) \leq n_0(1)$ with $\Sigma_{k>n_0(1)-1}\,\xi_k \geq \xi_1 > \xi_1 - m_0\eta > \eta > \Sigma_{k>n_0'}\,\xi_k$. Thus $n_0' \geq n_0(1) \geq \kappa(1)$. By simultaneous controllability, there exists a control ϵ such that $\langle \tilde{O}\Psi_2(T, \epsilon + y_{\kappa(1)}), \Psi_2(T, \epsilon + y_{\kappa(1)})\rangle = 1$ and $\langle \tilde{O}\Psi_2(T, \epsilon + y_j), \Psi_2(T, \epsilon + y_j)\rangle = 0$ for all $j \leq n_0'$ and $j \neq \kappa(1)$. In

addition, lemma 2.3 proves that for any control ϵ, $v_1^1 = v_2^{\kappa(1)} = 1$. So for this ϵ,

$$\xi_1 \leq \sum_{k\in\mathcal{I}/v_1^k=v_2^{\kappa(1)}} \xi_k = \sum_{k'\in\mathcal{I}/v_2^{k'}=v_2^{\kappa(1)}} \xi_{k'} \leq \xi_{\kappa(1)} + \sum_{k>n_0'} \xi_k \leq \xi_{\kappa(1)} + \eta. \quad (2.30)$$

We deduce that $\xi_{\kappa(1)} \geq \xi_1 - \eta$. With the same reasoning and by recurrence we demonstrate that $\xi_{\kappa^m(1)} \geq \xi_1 - m\eta$ for any $m \in \{1, \cdots, m_0\}$ thanks to the relationship $\sum_{k>n_0(\kappa^m(1))-1} \xi_k \geq \xi_{\kappa^m(1)} \geq \xi_1 - m\eta \geq \xi_1 - m_0\eta > \eta > \sum_{k>n_0'} \xi_k$. If $1, \kappa(1), \cdots, \kappa^{m_0}(1)$ are all distinct, then $1 = \sum_{k\in\mathcal{I}} \xi_k \geq \sum_{m=0}^{m_0} \xi_{\kappa^m(1)} > 1$, which leads to a contradiction. So at least two among the $1, \kappa(1), \cdots, \kappa^{m_0}(1)$ are equal.

On the other hand equation (2.29) implies that the collection of the two systems

$$\begin{cases} \frac{dX_1(t,\epsilon)}{dt} = [-i(H_1 + y_{\kappa^m(1)}\mu_1) + \epsilon(t)(-i\mu_1)]X_1(t,\epsilon) \\ X_1(0,\epsilon) = Id \end{cases} \quad (2.31)$$

and

$$\begin{cases} \frac{dX_2(t,\epsilon)}{dt} = [-i(H_2 + y_{\kappa^{m+1}(1)}\mu_2) + \epsilon(t)(-i\mu_2)]X_2(t,\epsilon) \\ X_2(0,\epsilon) = Id \end{cases} \quad (2.32)$$

is not ensemble controllable for all $m \in \{0, \cdots, m_0 - 1\}$. Applying the theorem 2.1 and the lemma 2.2 to $G = SU(N)$, $A_1 = -i(H_1 + y_{\kappa^m(1)}\mu_1)$, $A_2 = -i(H_2 + y_{\kappa^{m+1}(1)}\mu_2)$, $B_1 = -i\mu_1$ and $B_2 = -i\mu_2$ there exist f_m automorphisms of $\mathfrak{su}(N)$ such that $f_m(-i(H_1 + y_{\kappa^m(1)}\mu_1)) = -i(H_2 + y_{\kappa^{m+1}(1)}\mu_2)$ and $f_m(-i\mu_1) = -i\mu_2$. By linearity of f_1 and f_m, we obtain $(f_m^{-1} \circ f_1)(-iH_1) = -iH_1 + [(y_{\kappa(1)} - y_1) - (y_{\kappa^{m+1}(1)} - y_{\kappa^m(1)})](-i\mu_1)$ and $(f_m^{-1} \circ f_1)(-i\mu_1) = -i\mu_1$. Denote $f = f_m^{-1} \circ f_1$ and $\beta = (y_{\kappa(1)} - y_1) - (y_{\kappa^{m+1}(1)} - y_{\kappa^m(1)})$, then we have $-iH_1 = f(-iH_1) + i\beta\mu_1 = f(f(-iH_1) + i\beta\mu_1) + i\beta\mu_1 = f^2(-iH_1) + 2i\beta\mu_1$ and by recurrence $-iH_1 = f^p(-iH_1) + ip\beta\mu_2$ for all $p \in \mathbb{N}$. All automorphisms of $\mathfrak{su}(N)$ belong

to a compact set hence the set $\{f^p(-iH_1) \in \mathfrak{su}(N), p \geq 0\}$ is bounded for all m. Therefore the sequence $(ip\beta\mu_2)_{p\geq 0}$ is bounded which implies $\beta\mu_2 = 0$. According to assumption (**A1**), $\mu_2 \neq 0$. Thus $\beta = 0$. Denote $C = y_{\kappa(1)} - y_1$, then $y_{\kappa^{m+1}(1)} = y_{\kappa^m(1)} + C \ \forall m \in \{0, \cdots, m_0 - 1\}$. As $1, \kappa(1), \cdots, \kappa^{m_0}(1)$ are not all different, $C = 0$.

Since $\tilde{O} \in \mathcal{O}$ was arbitrary we proved so far that the systems without noise $(H_1 + y_1\mu_1, \mu_1)$ and $(H_2 + y_1\mu_2, \mu_2)$ give the same observations for the CSCO $\mathcal{O}$; the conclusion follows from the theorem 2.4.

Remark 2.10 *Here and in all similar results, the time T should be understood as 'if the time is large enough': the proof can be trivially adapted to treat the situation when the equality in law holds at some other final time T^* provided that T^* is larger than the time T given by the theorem.*

A similar reasoning allows to prove for the setting **(S1)** the following:

Corollary 2.1 *Consider the same setting and assumptions as in the theorem 2.3 with the exception of the relation (2.7). Then there exists $T > 0$ such that if:*

$$\mathcal{L}_Y\langle O\Psi_1(T, \epsilon + Y), \Psi_1(T, \epsilon + Y)\rangle = \mathcal{L}_Y\langle O\Psi_2(T, \epsilon + Y), \Psi_2(T, \epsilon + Y)\rangle \tag{2.33}$$

$$\forall \epsilon \in L^1([0, T]; \mathbb{R}), \quad \forall O \in \mathcal{O},$$

then the conclusion (2.8) of the theorem 2.3 holds.

2.5.3 The multiplicative perturbation case

In this section we consider the multiplicative perturbation, which means the control is in the form of $u = Y \cdot \epsilon$. We suppose moreover that this perturbation is positive: $\mathcal{V} \subset \mathbb{R}^+$.

Corollary 2.2 *Consider the same setting and assumptions as in the theorem 2.4 with the exception of the relation (2.9). Then there exists $T > 0$ such that if:*

$$\mathcal{L}_Y\langle O_k\Psi_1(T,\epsilon Y),\Psi_1(T,\epsilon Y)\rangle = \mathcal{L}_Y\langle O_k\Psi_2(T,\epsilon Y),\Psi_2(T,\epsilon Y)\rangle \quad (2.34)$$
$$\forall\epsilon\in L^1([0,T];\mathbb{R}),\quad \forall k=1,\cdots,K,$$

then either the conclusion (2.10) or the conclusion (2.11) of the theorem 2.4 holds (see also Remark 2.7).

Proof. The proof is similar with the exception that the simultaneous controllability result to be used is the corollary 5 page 25 in [6].

Remark 2.11 *When $\mathcal{V}$ also contains negative values, a similar result can be stated. The only difference is that one obtains:*

$$\begin{cases}\langle\mu_1\phi_j,\phi_k\rangle = \pm e^{i(\alpha_j-\alpha_k)}\langle\mu_2\phi_j,\phi_k\rangle,\\ \langle H_1\phi_j,\phi_k\rangle = \pm e^{i(\alpha_j-\alpha_k)}\langle H_2\phi_j,\phi_k\rangle,\\ \langle\Psi_0^1,\phi_j\rangle = \pm e^{i(\theta-\alpha_j)}\langle\Psi_0^2,\phi_j\rangle,\end{cases} \quad (2.35)$$

and a similar relation for the conjugate case. Furthermore, the polynomial situation $\mathcal{V}\subset\mathbb{R}^d$ with $d>1$, $u(t)=\sum_{a=0}^d Y_a\epsilon^a(t)$ can be studied. But although this case is also tractable with the controllability result in [35], the conclusion is very cumbersome to formulate and we leave it as an exercise for the reader.

2.6 Numerical application

Numerical tests are presented for the setting of the theorem 2.5. We consider the 4-level system ($N = 4$) in [14] and want to recover the Hamiltonian matrix H_{real} and the dipole moment matrix μ_{real}:

$$H_{real} = \begin{pmatrix} 0.0833 & -0.0038 & -0.0087 & 0.0041 \\ -0.0038 & 0.0647 & 0.0083 & 0.0038 \\ -0.0087 & 0.0083 & 0.0036 & -0.0076 \\ 0.0041 & 0.0038 & -0.0076 & 0.0357 \end{pmatrix},$$

$$\mu_{real} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 7 \\ 0 & -1.5 & 7 & 0 \end{pmatrix}.$$

Note that:

$$H_{real} = e^{\mathcal{P}_{real}} D e^{-\mathcal{P}_{real}},$$

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.0365 & 0 & 0 \\ 0 & 0 & 0.0651 & 0 \\ 0 & 0 & 0 & 0.0857 \end{pmatrix},$$

$$\mathcal{P}_{real} = \begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & -1 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{pmatrix}.$$

In practice the eigenvalues of the free Hamiltonian are measured by spectrometry and hence known with high precision, see also the discussion in [29, page 379 and Remark 7 page 384]. Accordingly, we suppose that the eigenvalues of H_{real} are known i.e., the matrix D is known. So

ℓ	1	2	3	4	5
y_ℓ	0.000400	0.000066	0.001025	0.000224	0.000816
ξ_ℓ	0.181810	0.163630	0.145450	0.127270	0.109090
ℓ	6	7	8	9	10
y_ℓ	0.000679	0.000740	0.000975	0.000211	0.000156
ξ_ℓ	0.090900	0.072720	0.054540	0.036360	0.018180

Table 2.1: Law of Y for the numerical example in section **??**. Here $L_0 = 10$; the second row presents the values y_ℓ, $\ell \leq L_0$ which have been chosen randomly (uniformly) in $[0, 0.1 * \frac{\|H\|_{l^\infty}}{\|\mu\|_{l^\infty}} = 0.0012]$. The third row displays the probabilities ξ_ℓ, $\ell \leq L_0$ which have been chosen at random, uniformly in $[0, 1]$, the sum rescaled to 1 and then ordered such that $(\xi_\ell)_{\ell \geq 1}$ is a decreasing sequence.

identifying H_{real} is equivalent to identifying the anti-Hermitian rotation matrix $\mathcal{P}_{real}$.

The law of the perturbation Y is given in table 2.1. We consider a finite set of test control fields of the form:

$$\epsilon(t) = \exp\left(-40(t - T/2)^2/T^2\right) \sum_{i<j,i,j=1}^{N} A_{i,j} \sin[(\lambda_j(H_{real}) - \lambda_i(H_{real}))t + \theta_{i,j}]. \tag{2.36}$$

Here $\lambda_i(H_{real})$ are eigenvalues of H_{real}, $i \leq N$ and $A_{i,j}$, $\theta_{i,j}$ are parameters to be chosen later. The total simulation time is $T = 3200$ which means about 10 periods of the smallest transition frequency $\lambda_4(H_{real}) - \lambda_3(H_{real})$.

Let $\{e_k; k \leq N\}$ be the canonical basis of $\mathbb{C}^N$ and $\mathcal{O} = \{e_k e_k^*, k \leq N\}$ (populations).

We choose $N_\epsilon = 36$ controls $\epsilon_1(t), ..., \epsilon_{N_\epsilon}(t)$ drawing θ_{ij} uniformly in

$[0, 2\pi]$ and A_{ij} uniformly in $[0, 0.0012]$ and we define the functional to be minimized:

$$\mathcal{J}(\mathcal{P}, \mu) = \sum_{i=1}^{N_\epsilon} \sum_{j=1}^{N} d_{\mathcal{W}_1}(\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}} D e^{-\mathcal{P}}, \epsilon_i + Y, \mu, \Psi_1^0), e_j \rangle|^2, \mathcal{L}_Y(|\langle \Psi(T, H_{real}, \epsilon_i + Y, \mu_{real}, \Psi_{real}^0), e_j \rangle|^2). \quad (2.37)$$

Here we use the 1-Wasserstein distance (see page 34-35 in [37]) $d_{\mathcal{W}_1}$ between two laws $L_Y Z_1$ and $L_Y Z_2$ defined as $d_{\mathcal{W}_1}(L_Y Z_1, L_Y Z_2) = \int_0^1 |F_{Z_1}^{-1}(x) - F_{Z_2}^{-1}(x)| dx$ with F_{Z_1}(respectively F_{Z_2}) the cumulative distribution function of Z_1(respectively Z_2)(see page 73-75 in [37] for details). We start with 10% relative error on μ and $\mathcal{P}$ and we use a classical unconstrained nonlinear optimization algorithm to minimize $\mathcal{J}(\mathcal{P}, \mu)$ (we used the Gnu Octave [16, 17] procedure "fminunc"). After 277 iterations, we find:

$$\mathcal{P}_{277} = \begin{pmatrix} 0 & 0.999 & -0.999 & 1.002 \\ -0.999 & 0 & 1 & 0.999 \\ 0.999 & -1 & 0 & -1.002 \\ -1.002 & -0.999 & 1.002 & 0 \end{pmatrix},$$

$$\mu_{277} = \begin{pmatrix} 0 & 4.999 & -0.998 & -0.003 \\ 4.999 & 0 & 6 & -1.5 \\ -0.998 & 6 & 0 & 7 \\ -0.003 & -1.5 & 7 & 0 \end{pmatrix}.$$

This corresponds to 0.003% relative error on μ and 0.001% relative error on $\mathcal{P}$. We note that the histograms for $(\mathcal{P}_{real}, \mu_{real})$ and $(\mathcal{P}_{277}, \mu_{277})$ are nearly the same. See figures 2.1, 2.2, 2.3 and 2.4 for details.

The optimization algorithm iterates starting from the initial guess $(\mathcal{P}_0, \mu_0)$ and constructs a sequence of estimations $(\mathcal{P}_k, \mu_k)$.

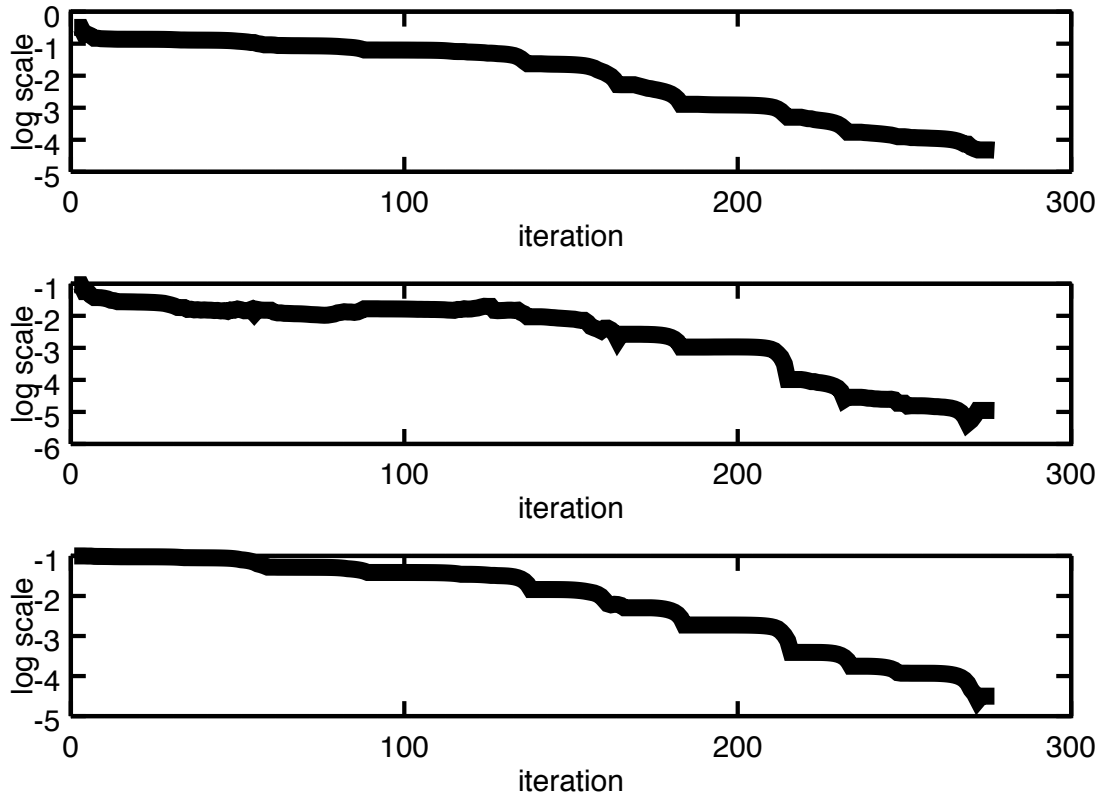

Figure 2.1: The (base 10) logarithm of $\mathcal{J}$ (upper plot), the (base 10) logarithm of the relative error on $\mathcal{P}$ (middle plot) and the (base 10) logarithm of the relative error on μ (lower plot) as a function of the iteration index.

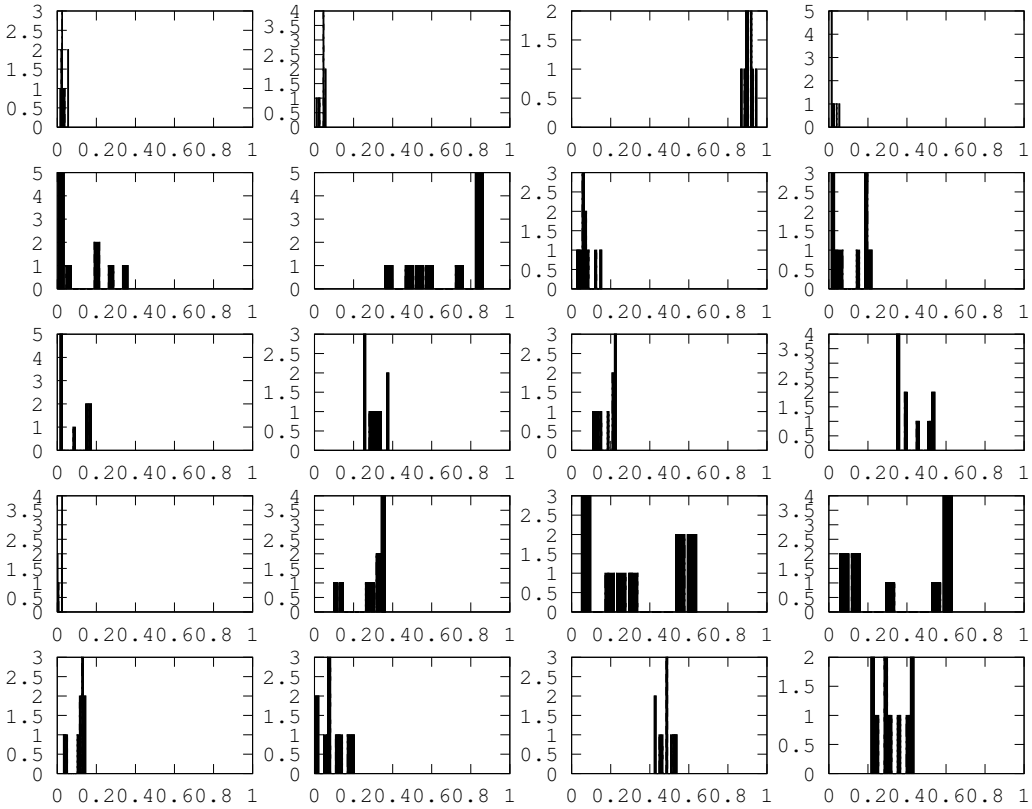

Figure 2.2: We plot the histograms of the laws $\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}_{real}} D e^{-\mathcal{P}_{real}}, \epsilon_i + Y, \mu_{real}, \Psi_1^0), e_j \rangle|^2$ for various choices of $i = 1, ..., 5$ and $j = 1, 2, 3, 4$.

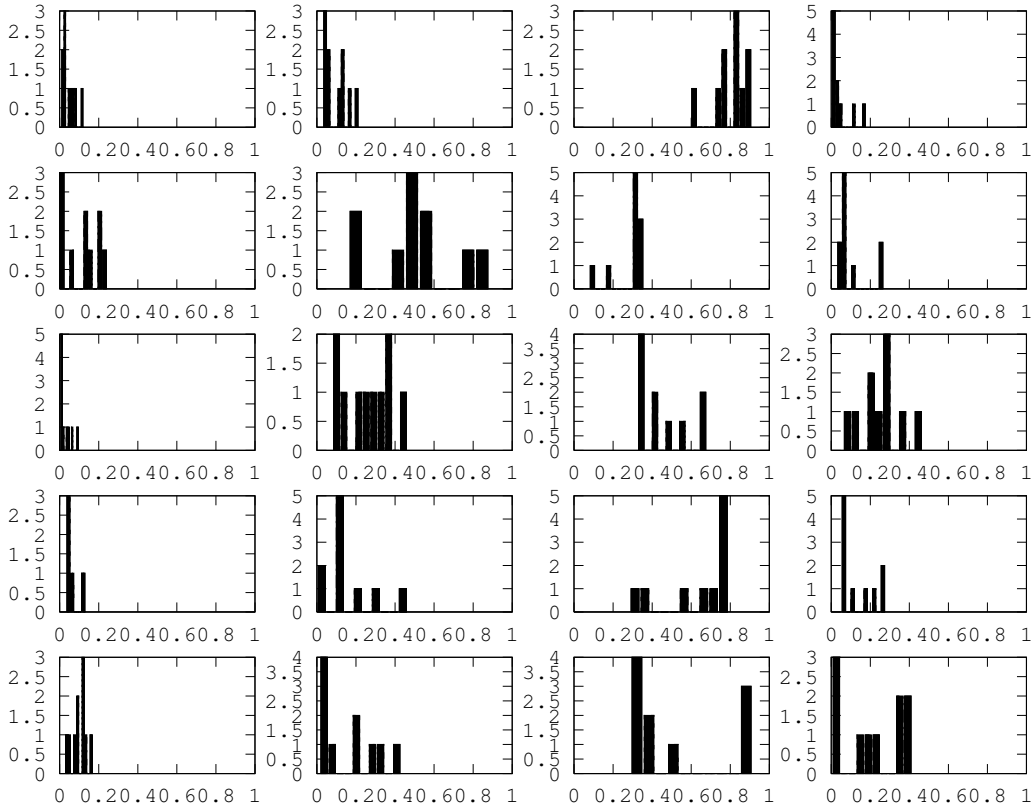

Figure 2.3: We plot the histograms of the laws $\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}_0} D e^{-\mathcal{P}_0}, \epsilon_i + Y, \mu_0, \Psi_1^0), e_j \rangle|^2$ for various choices of $i = 1, ..., 5$ and $j = 1, 2, 3, 4$.

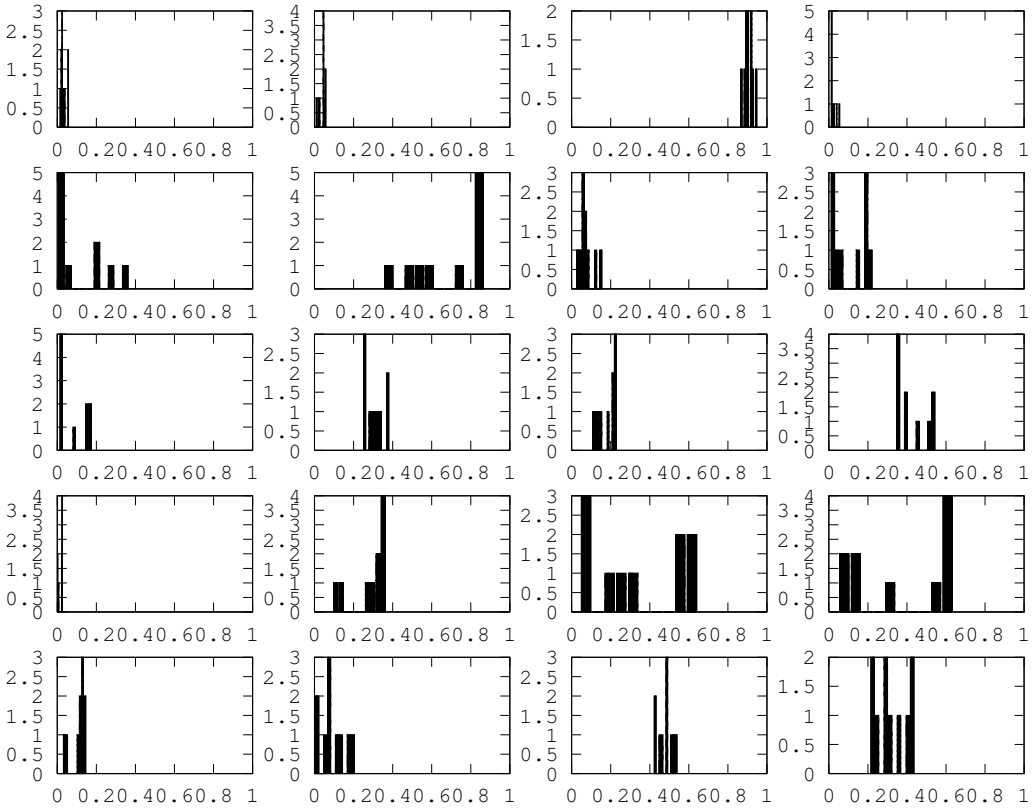

Figure 2.4: We plot the histograms of the laws $\mathcal{L}_Y(|\langle \Psi(T, e^{\mathcal{P}_{277}} D e^{-\mathcal{P}_{277}}, \epsilon_i + Y, \mu_{277}, \Psi_1^0), e_j \rangle|^2$ for various choices of $i = 1, ..., 5$ and $j = 1, 2, 3, 4$.

The optimization works well as there is an obvious match between the Figure 2.2 and the Figure 2.4 which represents the histograms for the real law and the final iteration respectively.

2.7 Perspectives and concluding remarks

Among the limitations of the present work is the requirement to consider only time-independent perturbations; it would be interesting to consider time-dependent perturbations and more elaborate noise models (beyond polynomial) and, of course, perturbations that can take values in an uncountable set (in the same spirit as in [5, 30]). Extension to infinite dimensional quantum systems can also be interesting; in all these cases one technical limitation is the absence of simultaneous controllability results analogue to theorems 2.1 and 2.2, still missing in general even for finite dimensional models as soon as the dimension is larger than 4.

A distinct extension, which seems attainable with the tools presented here, is to consider a framework that involves density matrices instead of wave-functions.

Chapter 3

Multiplicative amplitude noise model of unknown statistics

Quantum system inversion concerns learning the characteristics of the underlying Hamiltonian by measuring suitable observables from the responses of the system's interaction with members of a set of applied fields. Various aspects of inversion have been confirmed in theoretical, numerical and experimental works. Nevertheless, the presence of noise arising from the applied fields may contaminate the quality of the results. In this circumstance, the observables satisfy probability distributions, but often the noise statistics are unknown. Based on a proposed theoretical framework, we present a procedure to recover both the unknown parts of the Hamiltonian and the unknown noise distribution. The procedure is implemented numerically and seen to perform well for illustrative Gaussian, exponential and bi-modal noise distributions. This chapter reproduces the content of the accepted paper [19].

3.1 Introduction

The interaction between a quantum system and an applied field can be used to gather information about the system by measuring suitable observables for various incarnations of the field (see [4, 7, 9, 20, 31, 38, 39] for some related works). Such an inversion endeavor, seeking information about the system, has to take into account the possibility that the measurements are contaminated by noise, possibly from multiple sources, or that the model does not adequately describe all relevant characteristics of the system. Many considerations arise, including the following list:

S-1 The number of levels is unknown. Except for spins, the number of levels can generally take on any value, including the prospect of there being a continuum. The nature of the field bandwidth and intensity plays into which aspects of the system dynamics are accessible due to the field. In addition, unless the molecule is aligned (which is also possible) with the control field there will be vibration-rotation transitions. The circumstance does not change the key algorithmic principles set out in this work, but it can make the actual number of levels and dipole elements more involved.

S-2 Intense fields can be important in some problems where excited electronic states play a role, which can bring in non-resonant processes (i.e., virtual states, requiring adequate models). Moreover intense fields may lead to models including the nonlinear field coupling coefficients in the Hamiltonian.

S-3 If the system is in a thermal initial state, then the initial state will be a Boltzmann distribution instead of a single cold molecule in the ground or a pure state. The inversion will then need to deal with the distribution of molecules.

The performance of inversion needs to acknowledge the issues above. In addition, there will always be noise in the observations, but assuming

that it is random and symmetric (e.g. Gaussian) then signal averaging may reduce its influence. However, noise may accompany the control field with the following circumstances:

F-1 Different molecules (even if they are oriented) will see different fields as the field generally has a spatial (e.g., a Gaussian) lateral spread.

F-2 A so-called spatio-temporal "chirp" may occur, where there is some unknown complex spatio-temporal pattern (possibly fluctuating) in the field. When the pattern is not uniform it has to be factored into the model with, possibly nonlinear, additional terms, see [21].

F-3 Although the field may be known before entering the sample, optical distortion during propagation can lead to additional field uncertainty in the domain where the data is taken.

F-4 The variations in field are stochastic from shot-to-shot with respect to the frequency dependence of the amplitudes and phases. The origin of this behavior is mainly due to the laser source and possibly jitter in the optics. The implications for the control data due to this occurrence is likely more significant at higher intensities.

Collecting all of these items pose a daunting, but not an impossible challenge to treat. In this work we limit the focus on the randomness coming from the last item **F-4** above. These are related to so-called "fixed systematic errors", see [27, section VI.A. equations (38) and (40)] or "systematic control error", see [28] in the quantum computing literature. In [34] "low frequency noise" is used (cf. also [24, section IV. C.]): it is the portion of the (control) amplitude noise that has a correlation time up to 10^3 times longer than the timescale of the dynamics therefore it can be considered as constant in time. The understanding of laser noise and various associated models is an active area of research [15, 25, 40].

We present the theoretical background and analysis in Section 3.2. The numerical algorithm and simulation results are given in Section 3.3. Final remarks are made in Section 3.4.

3.2 Theoretical framework

3.2.1 The model

We introduce the following notation:

- $\mathbb{L}_{M_1,M_2,\cdots,M_m}$ is the Lie algebra spanned by the matrices $M_1, M_2, \cdots, M_m$;
- For any matrix or vector X we denote by X^* its adjoint (the transpose, complex conjugate);
- $\mathfrak{H}$ is the set of all Hermitian matrices $\mathfrak{H} = \{X \in \mathbb{C}^{N\times N} | X^* = X\}$;
- $\Psi(t, H, u(\cdot), \mu, \Psi_0)$ is the solution of the equation (4.1) below; to simplify the notation, when there is no ambiguity, we denote it as $\Psi(t)$;
- $SU(N)$ is the special unitary group of degree N, which is the group of $N \times N$ unitary matrices with determinant 1;
- $\mathfrak{su}(N)$ is the Lie Algebra of skew-Hermitian matrices (the Lie algebra of $SU(N)$);
- A set of commuting observables $\mathcal{O} = \{O_1, ..., O_K\}$ (named SCO hereafter) satisfies $[O_k, O_\ell] = 0$, $\forall k, \ell \in \{1, ..., K\}$;
- $\mathcal{L}(X)$ is the distribution of the random variable X.

Let us consider the following controlled quantum system with time-dependent wave-function $\Psi(t)$ satisfying the Schrödinger equation:

$$\begin{cases} i\dot{\Psi}(t, H, u(\cdot), \mu, \Psi_0) = (H - u(t)\mu)\Psi(t, H, u(\cdot), \mu, \Psi_0) \\ \Psi(0, H, u(\cdot), \mu, \Psi_0) = \Psi_0, \end{cases} \tag{3.1}$$

where H is the internal ("free") Hamiltonian and μ the coupling operator between the control $u(t) \in \mathrm{L}^1_{loc}(\mathbb{R}_+;\mathbb{R})$ and the system.

We work in a finite dimensional framework, therefore $H, \mu \in \mathfrak{H}$ for some $N \in \mathbb{N}^*$. The free Hamiltonian H is taken as known and the goal is to recover the matrix entries of μ from laboratory measurements of some observables depending on $\Psi(t)$. The control $u(t)$ can be changed in the laboratory to gather sufficient information on the system for the inversion to extract μ.

In the laboratory, the control field is produced by superposition of several frequencies, each with a specific amplitude and phase:

$$u(t) = \sigma(t) \sum_{\alpha \neq \beta} A_{\alpha\beta} sin(\omega_{\alpha\beta} t + \theta_{\alpha\beta}), \tag{3.2}$$

where $\sigma(t)$ is a Gaussian envelope in time and $\omega_{\alpha\beta} = E_\beta - E_\alpha$ is the transition frequency between the eigenvalues E_α and E_β of H. The amplitudes $A_{\alpha\beta}$ and the phases $\theta_{\alpha\beta}$ are the control parameters. In practice, the control fields cannot be produced perfectly. When we repeat nominally the same experiment, there can be a random shift in the amplitudes, which means for each experiment, there is a multiplicative noise factor on the amplitudes. Accordingly the perturbation is modeled by a random multiplicative factor Y acting on the control i.e., $u(t)$ is replaced by $Y \cdot u(t)$. This model has the assumption that each $A_{\alpha\beta}$ has the same shift; in addition we do not treat here possible noise appearing in the phases $\theta_{\alpha\beta}$.

In principle, the same methodology may apply to noise appearing in the phases; however at this time adapted ensemble controllability results are not available and on the other hand this requires to work with a high-dimensional probability distribution (one dimension for each inde-

pendent noise in a phase).

The perturbation Y takes values in $\mathcal{V} = \{y_\ell, \ell \leq L\} \subset \mathbb{R}$. We denote $\xi_\ell = \mathbb{P}(Y = y_\ell)$, where the column vector ξ with entries $\xi_1, \cdots, \xi_L$ is a probability distribution on $\mathcal{V}$. The numerical values of the possible perturbations y_ℓ are known but their occurrence probabilities ξ_ℓ are unknown and thus Ψ is a random variable, as are all measurements depending on Ψ. Repeating the control experiment several times will yield the distribution of the measurements.

The measurements are of the form $\langle O\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ with $O \in \mathfrak{H}$ being a member of a list of possible measurable operators. Often, only one observable operator is readily available, but the experiment can be repeated many times, including with distinct chosen fields. Generally for two observables $O_1, O_2 \in \mathfrak{H}$ no information is available on the joint distribution of the values $\langle O_1\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ and $\langle O_2\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$.

3.2.2 Theoretical result

The following theorem proves that under certain assumptions on the system and the SCO $\mathcal{O}$, if we obtain the same distributions for all observables in $\mathcal{O}$ and for all controls, then the dipole moment μ and the probabilities $(\xi_\ell)_{\ell=1}^L$ can be identified up to some multiplicative phases.

Remark 3.1 *As the noise (with unknown distribution) Y multiplies the unknown dipole μ, when the couple (Y, μ) is a solution, any couple $(Y/\lambda, \lambda\mu)$ is a solution too; thus it is only possible to obtain Y and μ up to a multiplicative factor (here λ). If, for instance, some additional information on Y or μ is known the constant λ can be set accordingly. As such, we will suppose from now on that for at least one transition*

an absorption intensity measurement can be performed which provides the value of $|\mu_{k,\ell}|^2$ *for some given* k, ℓ *(see [11, Chapter XIII, Section C.3.b Fermi's Golden Rule, p. 1299]). This is recalled in assumption* ***Hyp-D*** *below.*

Theorem 3.1 *Let* H, μ_1, $\mu_2 \in \mathfrak{H}$, H *diagonal,* $\mu_1 \neq 0$, $\mu_2 \neq 0$, Y_1, Y_2 *two random variables with values in the same set* $\mathcal{V}$, *at least one of which is non-null,* $\Psi_0^1, \Psi_0^2 \in \mathcal{S}_N$ *some initial states and denote for* $a = 1, 2$ *and* $u \in \mathrm{L}^1_{loc}(\mathbb{R}_+, \mathbb{R})$: $\Psi_a(t, u) = \Psi(t, H, u(\cdot), \mu_a, \Psi_0^a)$. *Let* $\mathcal{O}$ *be a (non-trivial) SCO. We suppose that* $N \geq 3$ *and:*

Hyp-A $\mathbb{L}_{iH, i\mu_1} = \mathbb{L}_{iH, i\mu_2} = \mathfrak{su}(N)$;

Hyp-B $tr(H) = tr(\mu_1) = tr(\mu_2) = 0$;

Hyp-C *the eigenvalues of* H *are all of multiplicity one.*

Hyp-D $|(\mu_1)_{k,\ell}|^2 = |(\mu_2)_{k,\ell}|^2 \neq 0$ *for some fixed* k, ℓ .

The final observation time is denoted T *(assumed large enough) and we suppose the following equality of distributions:*

$$\begin{aligned} &\mathcal{L}(\langle O\Psi_1(T, uY_1), \Psi_1(T, uY_1)\rangle) = \mathcal{L}(\langle O\Psi_2(T, uY_2), \Psi_2(T, uY_2)\rangle) \quad (3.3)\\ &\forall u \in L^1([0, T]; \mathbb{R}), \quad \forall O \in \mathcal{O}, \end{aligned}$$

then for some $(\alpha_i)_{i=1}^N \in \mathbb{R}^N$:

$$\begin{cases} (\mu_1)_{jk} = \pm e^{i(\alpha_j - \alpha_k)}(\mu_2)_{jk}, \ \forall j, k \leq N, \\ \mathbb{P}(Y_1 = y_\ell) = \mathbb{P}(Y_2 = \pm y_\ell) \qquad \forall \ell \leq L. \end{cases} \tag{3.4}$$

Remark 3.2 *The $\pm$ signs in equation (3.4) are due to the hypothesis* ***Hyp-D*** *which leaves the sign undetermined; it may also appear when the sign of the noise is ambigous i.e., when Y_1 and $-Y_1$ have the same distribution.*

Remark 3.3 *The SCO $\mathcal{O}$ may contain just one observable.*

Proof. The proof requires the tools introduced in [18] where the additive noise $u(\cdot) + Y$ was considered. In order to keep it simple we only give the main ideas and the modifications with respect to proof presented there. Using Lemma 5.1 in [18], for all $\ell \leq L$ and $O \in \mathcal{O}$, there exists $\kappa_O(\ell) \leq L$ such that

$$\langle O\Psi_1(T, u \cdot y_\ell), \Psi_1(T, u \cdot y_\ell)\rangle = \langle O\Psi_2(T, u \cdot y_{\kappa_O(\ell)}), \Psi_2(T, u \cdot y_{\kappa_O(\ell)})\rangle. \tag{3.5}$$

Reasoning as in the proof of theorem 4.1 of [18], this implies that there exists $W \in SU(N)$ diagonal such that $y_{\kappa_O(\ell)}\mu_2 = W y_\ell \mu_1 W^{-1}$. Since at least one of Y_1 or Y_2 is non-null, we can suppose, without loss of generality, that $y_\ell \neq 0$; we deduce the existence of some $\lambda \in \mathbb{R} \setminus \{0\}$ such that

$$\begin{cases} (\mu_1)_{jk} = \lambda e^{i(\alpha_j - \alpha_k)}(\mu_2)_{jk}, \ \forall j, k \leq N, \\ \mathbb{P}(Y_1 = y_\ell) = \mathbb{P}(Y_2 = \lambda y_\ell) \qquad \forall \ell \leq L. \end{cases} \tag{3.6}$$

However, since μ_1 and μ_2 are fixed for all O and ℓ we obtain that λ is independent of ℓ, which, using assumption **Hyp-D**, gives the conclusion.

Remark 3.4 *The hypothesis* ***Hyp-A*** *is required for identification while* ***Hyp-B*** *is rather a convention. Hypothesis* ***Hyp-C*** *can be relaxed (as in [18]) if the $\mathcal{O}$ is a Complete Set of Commuting Observables (CSCO).*

3.3 Numerical results

3.3.1 The algorithm

Several numerical simulations were performed in order to illustrate the theoretical result in Section 3.2. In all cases we simulate the "real" system with Hamiltonian H, dipole moment μ_{real}, noise Y^{real} with distribution $\mathcal{L}(Y^{real}) = \sum_{\ell=1}^{L} \xi_\ell^{real} \delta_{y_\ell}$ and observables $\mathcal{O} = \{O_1, ..., O_K\}$ that correspond to the Hamiltonian H, specified by projections $\{|e_1\rangle\langle e_1|, ... , |e_N\rangle\langle e_N|\}$. Here $|e_1\rangle$ is the j-th eigenstate of H. Note that here we take $\mathcal{O} = \{O_1, ..., O_K\}$ to be the entire set of projectors, but one would be enough for the theoretical result to hold, see the Remark 3.3.

The measurements provide the distributions of the observables for each control u:

$$\sum_{k=1}^{L} \xi_k^{real} \delta_{|\langle \Psi(T,H,u\cdot y_k,\mu_{real},\Psi_1^0),e_j\rangle|^2}). \tag{3.7}$$

The Hamiltonian H is assumed known to high accuracy. In contrast, we suppose that *a priori* we only know the order of magnitude of the dipole moment. This information is useful as an initial guess for the inversion procedure.

To find the dipole moment and the noise distribution of the control amplitudes we minimize the difference between

(i) the observed distribution calculated, at some suitably large time T, with the current dipole candidate μ and the current noise distribution candidate $\sum_{k=1}^{L} \xi_k \delta_{y_k}$ and

(ii) the real observable distribution calculated with the real dipole moment μ_{real} and the real distribution $\sum_{k=1}^{L} \xi_k^{real} \delta_{y_k}$.

The difference is summed over several controls $u_1,...,u_{N_u}$ and defined as:

$$\mathcal{J}(\mu,(\xi_k)_{k=1}^{L};(u_i)_{i=1}^{N_u}) = \log\left\{\frac{1}{N_u}\sum_{i=1}^{N_u}\sum_{j=1}^{N}\mathcal{W}_1\left[\sum_{k=1}^{L}\xi_k\delta_{|\langle\Psi(T,H,u_i\cdot y_k,\mu,\Psi_1^0),e_j\rangle|^2},\right.\right.$$
$$\left.\left.\sum_{k=1}^{L}\xi_k^{real}\delta_{|\langle\Psi(T,H,u_i\cdot y_k,\mu_{real},\Psi_1^0),e_j\rangle|^2}\right]\right\}. \tag{3.8}$$

Here $\mathcal{W}_1$ represents the 1-Wasserstein (also known as Kantorovich-Rubinstein) distance between two distributions (see page 34-35 in [37]); for two probability distributions Z_1, Z_2 having cumulative distribution functions F_{Z_1}(respectively F_{Z_2}) the distance is:

$$\mathcal{W}_1(Z_1,Z_2) = \int_0^1 |F_{Z_1}^{-1}(x) - F_{Z_2}^{-1}(x)|dx. \tag{3.9}$$

Other distances could also be used, e.g. $\mathcal{W}_2$.

We start the optimization with an initial guess μ^0 (see (3.15)); the distribution ξ^0 is initialized to be uniform. The iteration $n \geq 1$ consists in the following steps:

Algo 1 Randomly choose N_u controls u_i^n, $i = 1,...,N_u$;

Algo 2 minimize $\xi \to \mathcal{J}(\mu^{n-1},\xi;(u_i)_{i=1}^{N_u})$ and set ξ^n to be a minimizer (in practice a close approximation);

Algo 3 minimize $\mu \to \mathcal{J}(\mu,\xi^n;(u_i)_{i=1}^{N_u})$ and set μ^n to be a minimizer (in practice a close approximation);

For the step **Algo 2**, denote $a_k^{ij} = |\langle\Psi(T,H,u_i^n\cdot y_k,\mu^n,\Psi_1^0),e_j\rangle|^2$ and $b_k^{ij} = |\langle\Psi(T,H,u_i^n\cdot y_k,\mu_{real},\Psi_1^0),e_j\rangle|^2$. A part of the algorithm is to minimize the error:

$$\log\left[\frac{1}{N_u}\sum_{i=1}^{N_u}\sum_{j=1}^{N}\mathcal{W}_1(\sum_{k=1}^{L}\xi_k\delta_{a_k^{ij}},\sum_{k=1}^{L}\xi_k^{real}\delta_{b_k^{ij}})\right]. \tag{3.10}$$

The differential of the distance $\mathcal{W}_1$ is not trivial to compute (see [37] for a rigorous mathematical treatment); instead, in this step, we use the L^2 distance between the smoothed densities: each Dirac mass is replaced by a Gaussian distribution with small variance (here $\nu^2 = 10^{-4}$): $\delta_{a_k^{ij}}$ is replaced by the normal distribution $\mathcal{N}(a_k^{ij}, \nu^2)$ and $\delta_{b_k^{ij}}$ is replaced by $\mathcal{N}(b_k^{ij}, \nu^2)$). The use of parameter ν does not assume any particular distribution for ξ^{real} and does not bias towards one, it is only a rapid way to obtain a computable gradient. We compute the $L \times L$ matrix $\mathcal{M}^{ij}$ whose entries $\mathcal{M}^{ij}_{k,\ell} = \frac{1}{\nu\cdot\sqrt{2\pi}} e^{-\frac{(a_k^{ij} - a_\ell^{ij})^2}{2\cdot\nu^2}}$ account for the density of the distribution $\mathcal{N}(a_\ell^{ij}, \nu^2)$ at the point a_k^{ij} (and a similar matrix $\mathcal{M}'^{ij}$ with $\mathcal{M}'^{ij}_{k,\ell} = \frac{1}{\nu\cdot\sqrt{2\pi}} e^{-\frac{(b_k^{ij} - b_\ell^{ij})^2}{2\cdot\nu^2}}$ for the real distribution). When ν is small the minimizer of the term in (3.10) is close to the minimizer of $\sum_{i=1}^{N_u}\sum_{j=1}^{N} \|\mathcal{M}^{ij}\xi - \mathcal{M}'^{ij}\xi^{real}\|^2$ which is given by the formula:

$$\xi^{n,raw} = \left(\sum_{i=1}^{N_u}\sum_{j=1}^{N} (\mathcal{M}^{ij})^T \mathcal{M}^{ij}\right)^{-1} \left(\sum_{i=1}^{N_u}\sum_{j=1}^{N} (\mathcal{M}^{ij})^T \mathcal{M}'^{ij}\right) \xi^{real}. \tag{3.11}$$

The term $\xi^{n,raw}$ is corrected to be a probability distribution and set to ξ^n:

$$\xi^n = \frac{1}{\sum_{\ell=1}^{L} |\xi_\ell^{n,raw}|} \left(|\xi_1^{n,raw}|, ..., |\xi_L^{n,raw}|\right). \tag{3.12}$$

For the step **Algo 3** there is no explicit solution; a classical unconstrained nonlinear optimization algorithm is employed (we used the Gnu Octave procedure "fminunc").

Once the algorithm finished, the remaining overall multiplicative constant λ (see Remark 3.1)) is set consistent with the data (see below the examples).

3.3.2 Numerical tests: N observables

We consider the 4-level system ($N = 4$) in [14] having:

$$H = \begin{pmatrix} 0.0833 & -0.0038 & -0.0087 & 0.0041 \\ -0.0038 & 0.0647 & 0.0083 & 0.0038 \\ -0.0087 & 0.0083 & 0.0036 & -0.0076 \\ 0.0041 & 0.0038 & -0.0076 & 0.0357 \end{pmatrix}, \tag{3.13}$$

$$\mu_{real} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 7 \\ 0 & -1.5 & 7 & 0 \end{pmatrix}. \tag{3.14}$$

We set $N_u = 36$; the controls are defined by formula (3.2) with amplitudes $(A_{\alpha\beta})$ chosen at random, uniformly in $[0, 0.1 \cdot \frac{\|H\|_{l^\infty}}{\|\mu_{real}\|_{l^\infty}} = 0.0012]$ and the phases $\theta_{\alpha\beta}$ in $[0, 2\pi]$. For the values (3.13) and (3.14) the initial guesses are:

$$\mu_g^0 = \begin{pmatrix} 0 & 3.76 & -1.31 & 0 \\ 3.76 & 0 & 3.51 & -1.78 \\ -1.31 & 3.51 & 0 & 6.72 \\ 0 & -1.78 & 6.72 & 0 \end{pmatrix},$$

$$\mu_e^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \tag{3.15}$$

$$\mu_b^0 = \begin{pmatrix} 0 & 7.48 & -0.51 & 0 \\ 7.48 & 0 & 8.83 & -0.87 \\ -0.51 & 8.83 & 0 & 5.87 \\ 0 & -0.87 & 5.87 & 0 \end{pmatrix}.$$

The average relative errors of these initial guesses are 42%, 70% and 50%. Guesses μ_g^0 and μ_b^0 are obtained by multiplying element-wise μ_{real}

by uniform random variables in $[0.4, 1.6]$ (i.e., up to 60% average relative error); the term μ_e^0 is obtained by taking roughly the order of magnitude of the entries of μ_{real}: the values smaller or about 1 are taken to be 1 while the others are set to 10. The observables arising from the Hamiltonian H are projections $\{|e_1\rangle\langle e_1|,\ |e_2\rangle\langle e_2|,\ |e_3\rangle\langle e_3|,\ |e_4\rangle\langle e_4|\}$ to eigenstates:

$$
\begin{aligned}
|e_1\rangle &= \begin{pmatrix} 0.0845 & -0.1313 & 0.9651 & 0.2101 \end{pmatrix}^T, \\
|e_2\rangle &= \begin{pmatrix} -0.1305 & -0.0856 & -0.2103 & 0.9651 \end{pmatrix}^T, \\
|e_3\rangle &= \begin{pmatrix} 0.2118 & 0.9647 & 0.0838 & 0.1325 \end{pmatrix}^T, \\
|e_4\rangle &= \begin{pmatrix} 0.9649 & -0.2118 & -0.1314 & 0.0830 \end{pmatrix}^T. \qquad (3.16)
\end{aligned}
$$

In the basis $\{e_1, e_2, e_3, e_4\}$ the Hamiltonian H is diagonal with eigenvalues $E_1 = 0$, $E_2 = 0.0365$, $E_3 = 0.0651$, $E_4 = 0.0857$. We set the final time $T = 3200$ which is about 10 periods of the smallest transition frequency $2\pi/(E_4 - E_3) = 314$ in H.

With respect to Remark 3.1 we suppose that the value $|(\mu_{real})_{12}|^2 = 25$ is known.

The support for the distribution Y is known and denoted $[y_m, y_M]$; in the numerical tests we take $y_m = 0.5$ and $y_M = 1.5$. We discretize the set of possible values of the perturbation with $L = 51$ equidistant points $y_\ell = y_m + (\ell - 1) \cdot \frac{y_M - y_m}{L-1}$, $\ell = 1, ..., L$. The values y_ℓ are supposed known but not the probabilities ξ_ℓ^{real} that define the distribution $\mathcal{L}(Y^{real}) = \sum_{\ell=1}^{L} \xi_\ell^{real} \delta_{y_\ell}$ of the perturbation Y^{real}. Several distributions are tested; they are constructed by discretizing, truncating and re-normalizing several classical distributions:

- $Y^{real} = Y^g$ being a Gaussian distribution centered at 1 with vari-

ance equal to 0.0025:

$$\xi_\ell^{real,g} = \frac{f_\ell^g}{\sum_{k=1}^L f_k^g} \quad \ell = 1, \cdots, L \tag{3.17}$$
$$\text{with } f_k^g = \frac{1}{\sqrt{0.0025}\sqrt{2\pi}} e^{-\frac{(y_k-1)^2}{2\cdot 0.0025}},$$

- $Y^{real} = Y^e$ being a shifted exponential distribution form:

$$\xi_\ell^{real,e} = \frac{f_\ell^e}{\sum_{k=1}^L f_k^e} \quad \ell = 1, \cdots, Ł \tag{3.18}$$
$$\text{with } f_k^e = 5 \cdot e^{-5\cdot(y_k - y_m)},$$

- $Y^{real} = Y^b$ being the bi-modal distribution which is the sum of two Gaussian distributions. We choose the first one centered at 0.8 with variance equal to 0.0025 and the second one centered at 1.2 with variance equal to 0.0049:

$$\xi_\ell^{real,b} = \frac{f_\ell^b}{\sum_{k=1}^L f_k^b} \quad \ell = 1, \cdots, L \tag{3.19}$$
$$\text{with } f_k^b = \frac{1}{\sqrt{0.0025}\sqrt{2\pi}} e^{-\frac{(y_k-0.8)^2}{2\cdot 0.0025}} + \frac{1}{\sqrt{0.0049}\sqrt{2\pi}} e^{-\frac{(y_k-1.2)^2}{2\cdot 0.0049}}.$$

The dipole moments converge in 10 iterations. The numerical values are rescaled in order to use that $|(\mu_{real})_{12}|^2 = 25$; same is done for the noise distributions. We obtain μ_g^{10}, μ_e^{10} and μ_b^{10}:

$$\|\mu_g^{10} - \mu_{real}\|_\infty = 5 \cdot 10^{-5}, \ \mu_g^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99995 & -1.5 \\ -1 & 5.99995 & 0 & 6.99999 \\ 0 & -1.5 & 6.99999 & 0 \end{pmatrix}, \tag{3.20}$$

$$\|\mu_e^{10} - \mu_{real}\|_\infty = 10^{-4},\ \mu_e^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 6.9999 \\ 0 & -1.5 & 6.9999 & 0 \end{pmatrix}, \quad (3.21)$$

$$\|\mu_b^{10} - \mu_{real}\|_\infty = 6 \cdot 10^{-5},\ \mu_b^{10} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 5.99999 & -1.5 \\ -1 & 5.99999 & 0 & 6.99994 \\ 0 & -1.5 & 6.99994 & 0 \end{pmatrix}. \quad (3.22)$$

The error norm is $\|\cdot\|_\infty$ is the largest, in absolute value, of the error components. See figures 3.1,3.2, 3.3 for the results. The tables 3.1 and 3.2 present the match of the probability distributions of the observables.

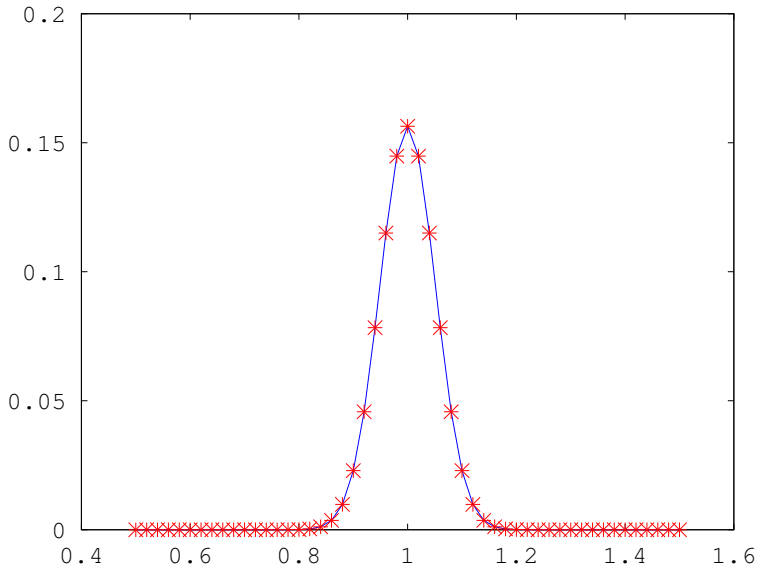

Figure 3.1: Identification of the Gaussian distribution (3.17). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is $5 \cdot 10^{-5}$, see equation (3.20).

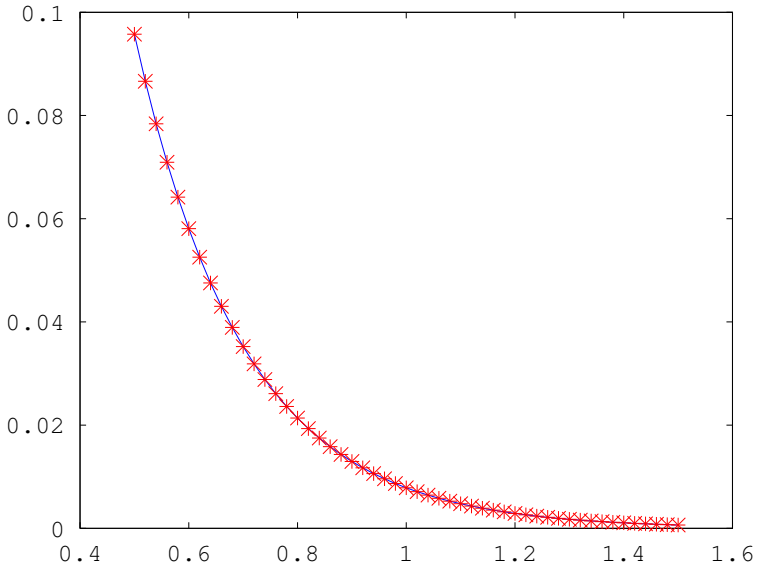

Figure 3.2: Identification of the exponential distribution (3.18). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is 10^{-4}, see equation (3.21).

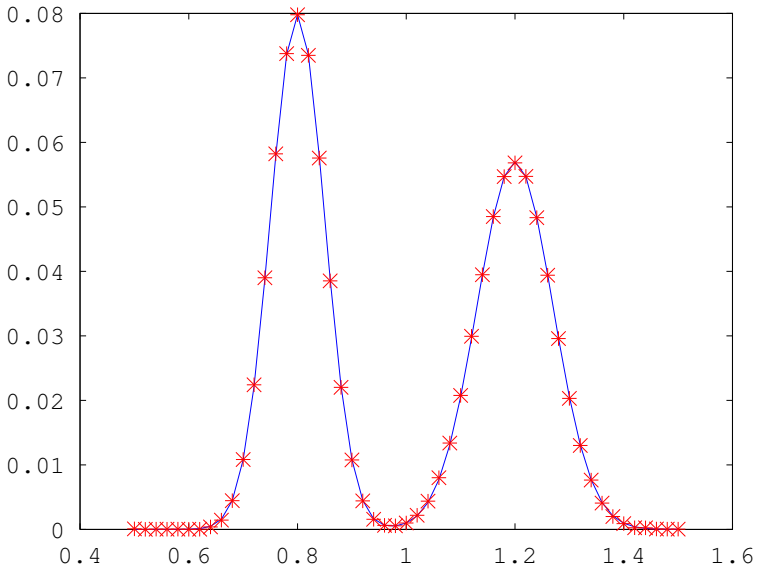

Figure 3.3: Identification of the bi-modal distribution (3.19). The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the error on the dipole moment is $6 \cdot 10^{-5}$, see equation (3.22).

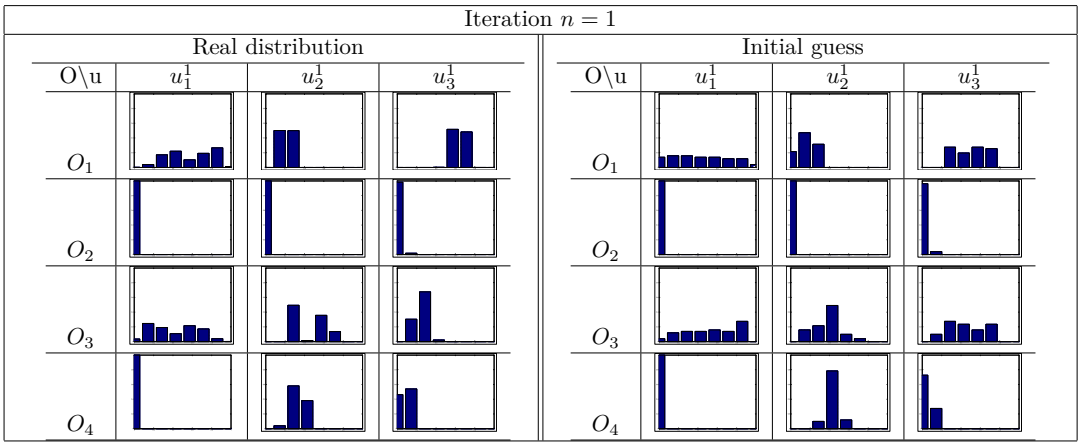

Table 3.1: The match of the observation distributions for the bi-modal distribution (3.19). We plot the histograms at the start of the inversion algorithm i.e., iteration step $n = 1$. In the left sub-table are the histograms of the observations $\sum_{k=1}^{L} \xi_k^{real} \delta_{|\langle \Psi(T,H,u_1^1 \cdot y_k, \mu_{real}, \Psi_1^0), e_j \rangle|^2}$ with the real distribution; in the right sub-table the histograms correspond to the initial guess which is the uniform distribution. Each column corresponds to a control field, here only the first 3 control fields u_1^1, u_2^1 and u_3^1 are shown. Each line corresponds to a specific observable in the SCO set $\mathcal{O}$. The initial guess is seen to be a poor approximation, as the histograms in left and right sub-tables differ substantially.

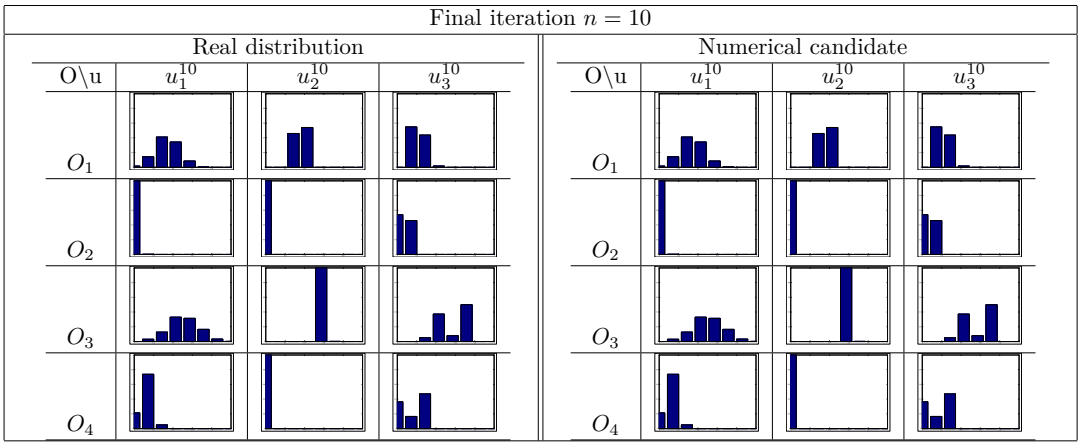

Table 3.2: Converged result (iteration $n = 10$) from initial guess Table 3.1. The identification works well as the left and right sub-tables match. As explained, the control fields are chosen randomly at each iteration and in particular the controls chosen at iteration $n = 1$ and $n = 10$ are not the same (otherwise the histograms corresponding to observations with the real noise distribution would be the same as in Table 3.1 left sub-table). Here the results for u_1^{10}, u_2^{10} and u_3^{10} are displayed.

3.3.3 A single measured observable

We use the same system as in the Section 3.3.2 except that here we consider the extreme case when only one observable is available as function of the control field. The observable is the projection $|e_3\rangle\langle e_3|$ to the third eigenstate.

The initial guess is:

$$\mu_b^0 = \begin{pmatrix} 0 & 10 & 1 & 1 \\ 10 & 0 & 10 & 1 \\ 1 & 10 & 0 & 10 \\ 1 & 1 & 10 & 0 \end{pmatrix}, \tag{3.23}$$

The coefficient $\mu_{2,1} = \mu_{1,2} = 5$ is fixed, as previously stated, but now it is treated as a constraint by the algorithm (which will thus only optimize the other coefficients). The algorithm minimizes the difference:

$$\mathcal{J}(\mu, (\xi_k)_{k=1}^L; (u_i)_{i=1}^{N_u}) = \log \left\{ \frac{1}{N_u} \sum_{i=1}^{N_u} \mathcal{W}_1 \left[\sum_{k=1}^{L} \xi_k \delta_{|\langle \Psi(T,H,u_i \cdot y_k, \mu, \Psi_1^0), e_3 \rangle|^2}, \right. \right.$$
$$\left. \left. \sum_{k=1}^{L} \xi_k^{real} \delta_{|\langle \Psi(T,H,u_i \cdot y_k, \mu_{real}, \Psi_1^0), e_3 \rangle|^2} \right] \right\}. \tag{3.24}$$

The distribution Y^{real} tested is the bi-modal model. The algorithm converges and after 5,10 and 15 iterations, the dipole moments we obtain respectively have an L^2 error of 0.17246, 0.03736 and $4.5652 \cdot 10^{-4}$ respectively. In order to test the robustness of the algorithm with respect to other error norms, we use the L^2 error norm $\|\cdot\|_{L^2}$ (the square root of the sum of squares of the components). The convergence is slower than in Section 3.3.2 when we treat $\mu_{2,1} = \mu_{1,2} = 5$ as a constraint.

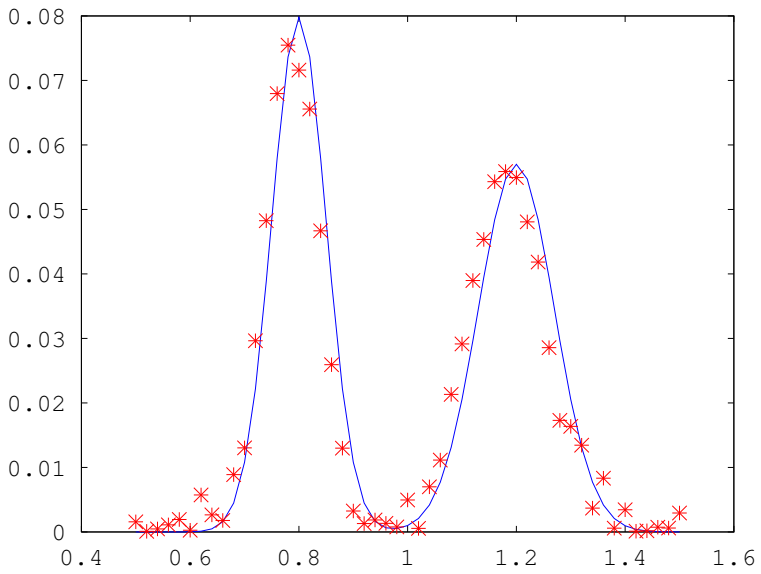

Figure 3.4: Identification of the bi-modal distribution (3.19) after 5 iterations. The real distribution is in blue, the numerical result in red. The noise distribution starts to have the same qualitative features as ξ^{real}; the L^2 error on the dipole moment is 0.17246.

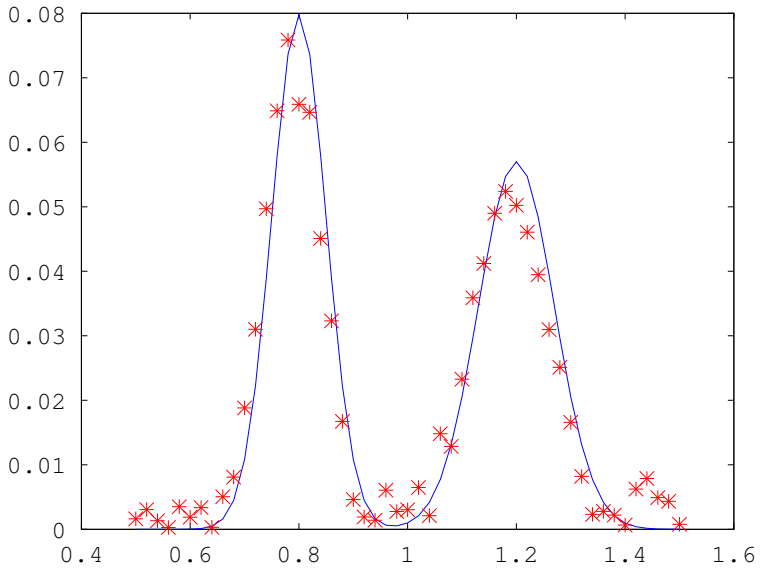

Figure 3.5: Identification of the bi-modal distribution (3.19) after 10 iterations. The real distribution is in blue, the numerical result in red. The noise distribution starts to converge; the L^2 error on the dipole moment is 0.03736.

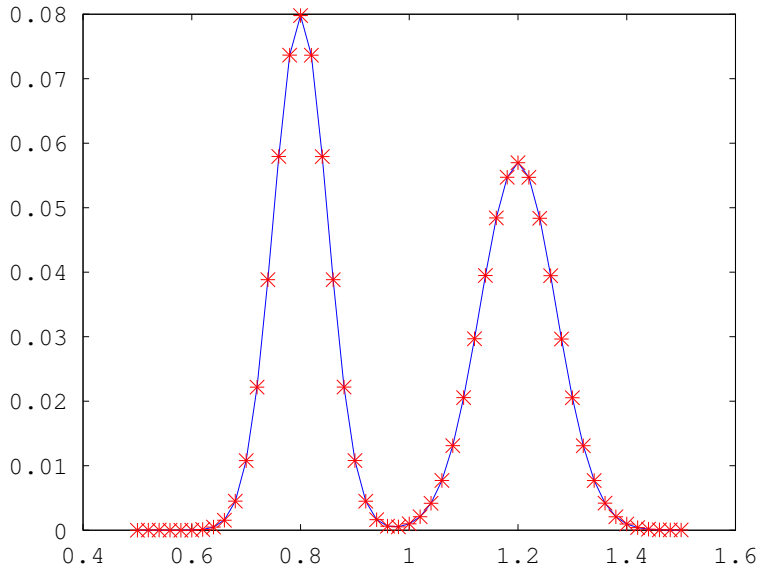

Figure 3.6: Identification of the bi-modal distribution (3.19) after 15 iterations. The real distribution is in blue, the numerical result in red. Good agreement with the unknown noise distribution ξ^{real} is obtained; the L^2 error on the dipole moment is $4.5652 \cdot 10^{-4}$.

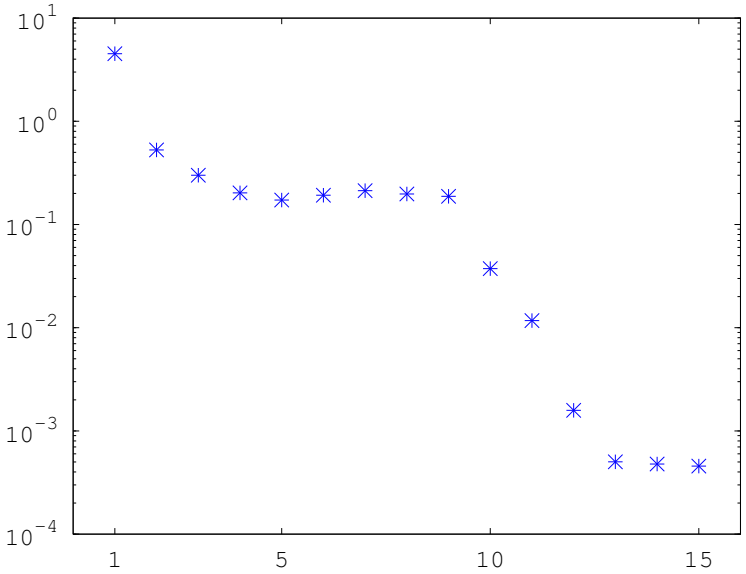

Figure 3.7: The L^2 error $\|\mu^k - \mu^{real}\|_{L^2}$ for $k = 1, ..., 15$ iterations.

Remark 3.5 *A legitimate question is related to the scaling of the identification with respect to the number N of levels. Note first that the computation in equation (3.24) only depends on the number of observables (and not on N, as equation (3.10) seemed to indicate). On the other hand, the computation of the $\mathcal{W}_1$ distance is independent on N and, for one dimensional laws, straightforward.*

Of course in order to compute the observables, numerical simulations are performed, and these do depend on N. But such simulations allow for trivial parallelization which could bring down the wall-clock time per iteration to that of a single numerical resolution of the N-level system. Such a time is a lower bound because one needs to check whether a candidate solution is indeed a good solution.

3.4 Conclusion

The inversion of the dipole moment has been considered in a model where the noise coming from the laser source is non-negligible and of unknown distribution. The model considered here has noise acting multiplicatively on the control intensity and is the same for all frequency components.

First, we proved theoretically that if one can measure repeatedly (at least) one observable for many control fields, the set of probability distributions of this observable is enough to recover both the dipole and the noise distribution. Then, a numerical algorithm based on the Wasserstein distance between the probability distributions was proposed and seen to perform well for several different, non-perturbative, noise distributions and initial guesses for the dipole. As subject for future study is the question of how to treat noise of possibly distinct character reflected in each of the amplitudes and phases or other different noise models as discussed in the Introduction. In addition, there are several further issues to consider in future work for creating a realistic algorithm for quantum system data inversion.

Chapter 4

Gaussian process phase noise model

In this section, we study the phase noise model. The noise in the phase are supposed dependent and modeled by a Gaussian process. The numerical tests are made in three cases: square exponential covariance model, Ornstein - Unlenbeck process and Brownian motion. In all cases, the recovery algorithm works well.

4.1 Introduction

Let us recall the controlled quantum system with time-dependent wavefunction $\Psi(t)$ satisfying the Schrödinger equation:

$$\begin{cases} i\dot{\Psi}(t, H, u(\cdot), \mu, \Psi_0) = (H - u(t)\mu)\Psi(t, H, u(\cdot), \mu, \Psi_0) \\ \Psi(0, H, u(\cdot), \mu, \Psi_0) = \Psi_0, \end{cases} \tag{4.1}$$

where H is the internal Hamiltonian and μ the coupling operator between the control $u(t) \in \mathrm{L}^1_{loc}(\mathbb{R}_+; \mathbb{R})$ and the system.

The space under study is finite dimensional, thus H and μ are $N \times N$ Hermitian matrices for some $N \in \mathbb{N}^*$. The free Hamiltonian H is supposed to be known. The physicists change the control $u(t)$ to gather as much information as they need on the system to recover the operator μ from laboratory measurements of some observables depending on $\Psi(t)$.

The control is parameterized as the superposition of lasers beams with different frequencies. Each has a specific amplitude and phase, a mathematical model of the control is:

$$u(t) = S(t) \int_{\omega \in \mathcal{D}} A(\omega) cos(\omega t + \theta(\omega)) d\omega. \tag{4.2}$$

$\mathcal{D}$ is a bounded part of $\mathbb{R}_+$ which is the set of all possible frequencies. $S(t)$ is a Gaussian function in time. The amplitudes $A(\omega) : \mathcal{D} \to \mathbb{R}_+$ and the phases $\theta(\omega) : \mathcal{D} \to [0, 2\pi]$ are the control parameters which can be adjusted during the experiments. The amplitude function A should be integrable on $\mathcal{D}$.

In practice, there are always some noises in the control. In this section,

we only consider noises in the phases. Let us introduce the phase noise model, which is

$$u(t) = S(t) \int_{\omega \in \mathcal{D}} A(\omega) cos(\omega t + \theta(\omega) + \delta\theta_\omega). \tag{4.3}$$

Denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the space of probability. The phase noises $(\delta\theta_\omega)_{\omega\in\mathcal{D}}$ are approximated as a Gaussian process indexed by frequencies. For all instants t, $u(t)$, the control at instant t, is now a random variable. As a result, the wave function $\Psi(t)$ also becomes a random variable.

The measurements are of the form $\langle O\Psi(T, H, u, \mu, \Psi_0), \Psi(T, H, u, \mu, \Psi_0)\rangle$ with O some $N \times N$ matrix being a member in a list of possible measurable operators. Unfortunately, often one experience can only provide the measurement with one operator. Although the experiments can be repeated as many times as we want, including with distinct chosen fields, there is no information available on the joint distribution of the measurements for two observables O_1, O_2 in general.

4.2 The noise model

A key fact of Gaussian processes is that they can be completely defined by their mean value functions and covariance functions. It is natural to assume that all Gaussian variables $\delta\theta(\omega)$ follow the same normal distribution $\mathcal{N}(0, \sigma^2)$ centered in 0.

In reality, the physicists use only a finite amount of laser frequencies to construct the control. Denote by N_ω the number of laser frequencies superposed. Then a discretization of equation 4.3 is:

$$u(t) = S(t) \sum_{l=0}^{N_\omega} A_l cos(\omega_l t + \theta_l + \delta\theta_l), \tag{4.4}$$

where $A_l = A(\omega_l)$, $\theta_l = \theta(\omega_l)$ and $\delta\theta_l$ is the random variable $\delta\theta_{\omega_l}$.

4.2.1 The expectation of $u(t)$

Denote by $f_{\delta\theta_l}$ the density function of the random variable $\delta\theta_l$. Let us calculate the expectation of $u(t)$:

$$\begin{aligned}
\mathbb{E}(u(t)) &= \int_{x\in\mathbb{R}} S(t) \sum_{l=1}^{N_\omega} A_l cos(\omega_l t + \theta_l + x) f_{\delta\theta_l}(x) dx \\
&= S(t) \sum_{l=1}^{N_\omega} A_l [cos(\omega_l t + \theta_l) \int_{x\in\mathbb{R}} cos(x) f_{\delta\theta_l}(x) dx) \\
&- sin(\omega_l t + \theta_l) \int_{x\in\mathbb{R}} sin(x) f_{\delta\theta_l}(x) dx].
\end{aligned}$$

As the $(\delta\theta_l)_{1\leq l\leq N_\omega}$ follow the identical centered normal distribution which is symmetric,

$$\int_{x\in\mathbb{R}} sin(x) f_{\delta\theta_l}(x) dx = 0,$$

and

$$\int_{x\in\mathbb{R}} cos(x) f_{\delta\theta_l}(x)dx = \int_{x\in\mathbb{R}} cos(x) f_{\delta\theta_{l'}}(x)dx$$

for all l and l'. Thus

$$\mathbb{E}(u(t)) = \alpha u_0(t),$$

where $u_0(t) = S(t)\sum_{l=1}^{N_\omega} A_l cos(\omega_l t + \theta_l)$ the control without noises and $\alpha = \int_{x\in\mathbb{R}} cos(x) f_{\delta\theta_l}(x)dx$ only depends on the variance σ^2.

We should note that when σ^2 is small, which means the Gaussian distribution is sharp, the parameter α is close to 1. On the contrary, when σ^2 is large, which corresponds to a uniform distribution, the parameter α tends to 0. In fact,

$$\begin{aligned}\alpha &= \sum_{k=0}^{\infty}(-1)^k \int_{x\in\mathbb{R}} x^{2k} f_{\delta\theta_l}(x)dx = \sum_{k=0}^{\infty}(-1)^k \mathbb{E}(\delta\theta_l^{2k}) \\ &= \sum_{k=0}^{\infty}(-1)^k \frac{(2k)!}{2^k k!}\sigma^{2k}.\end{aligned}$$

4.2.2 Correlations between the noises

The next question is that what kind of covariance function is suitable to model the correlation between the phase noises? First we remark that it is reasonable to make the assumption that the phase noises are isotopic processes, which means the covariance $K(\delta\theta_\omega, \delta\theta_{\omega'})$ depends only on the Euclidean distance between the frequencies $|\omega - \omega'|$. When the frequencies are close, the noises have more chance to be correlated, so the covariance function should be a decreasing function. The Matérn class of covariance functions (see Example 1.1) and the γ-exponential

covariance functions (see Example 1.2) seem to be good models.

4.3 Numerical simulations

The observables used here are the projections on the eigenstates of the real Hamiltonian H: $\mathcal{O} = \{O_1, ..., O_N\} = \{|e_1\rangle\langle e_1|, \ ... \ , |e_N\rangle\langle e_N|\}$. $|e_i\rangle$ is the i-th eigenstate of H.

In order to enable all the transformations, the frequencies $(\omega_l)_{1\leq l\leq N_\omega}$ are often chosen as the transitions of the eigenvalues of H: $|\lambda_i - \lambda_j|$, with $(\lambda_i)_{1\leq i\leq N}$ the eigenvalues of H. In particular, $N_\omega = \frac{N(N-1)}{2}$.

The Gaussian process $(\delta\theta_\omega)_{\omega\in\mathcal{D}}$ can be simulated by N_r random realizations. This number N_r should be large enough such that when we change the N_r realizations, the error occurred does not prevent the recovery of the dipole moment μ.

The simulation provide the distribution of the control $u(t)$ by

$$\mathcal{L}_{u(t)} = \sum_{k=1}^{N_r} \frac{1}{N_r} \delta_{S(t)\sum_{l=1}^{N_\omega} A_l cos(\omega_l t + \theta_l + \delta\theta_{l,k})}, \tag{4.5}$$

where $(\delta\theta_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r} \in \mathbb{R}^{N_\omega \cdot N_r}$ are N_r realizations of the correlated random variables $(\delta\theta_l)_{1\leq l\leq N_\omega}$.

The Hamiltonian H is assumed known to high accuracy. In contrast, we suppose that we only know μ_{real} with an error about 10%. We would like to know μ_{real} with more accuracy, for example, within 1% of error.

To find the dipole moment μ_{real} we minimize the difference between:

(i) the virtual distribution simulated, at some suitably large time T, with the current dipole candidate μ and the simulated control distributions

$u(t)$

and

(ii) the real observable distribution measured in the laboratory at T with the real dipole moment μ_{real} and the real control distributions $\tilde{u}(t)$.

The difference is summed over N_u control distributions $(u^j)_{1\leq j\leq N_u}$ and $(\tilde{u}^j)_{1\leq j\leq N_u}$ defined by the N_u set of different amplitudes $(A_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$ and phases $(\theta_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$.

Each control distribution u^j is simulated by N_r realizations of phase noises $(\delta\theta_{l,k}^j)_{1\leq l\leq N_\omega, 1\leq k\leq N_r} \in \mathbb{R}^{N_\omega \cdot N_r}$:

$$\mathcal{L}_{u^j(t)} = \sum_{k=1}^{N_r} \frac{1}{N_r} \delta_{S(t)\sum_{l=1}^{N_\omega} A_l^j cos(\omega_l t + \theta_l^j + \delta\theta_{l,k}^j)}. \tag{4.6}$$

And each control distribution $\tilde{u}^j(t)$ is defined as

$$\tilde{u}^j(t) = S(t) \sum_{l=0}^{N_\omega} A_l^j cos(\omega_l t + \theta_l^j + \delta\theta_l). \tag{4.7}$$

Now we define the difference for each $(u^j, \tilde{u}^j)$:

$$\begin{aligned} \mathcal{J}(u^j, \mu, \tilde{u}^j, \mu_{real}) = \sum_{i=1}^{N} \mathcal{W}_1(&|\langle \Psi(T, H, u^j, \mu, \Psi_0), e_i\rangle|^2, \\ &|\langle \Psi(T, H, \tilde{u}^j, \mu_{real}, \Psi_0), e_i\rangle|^2), \end{aligned} \tag{4.8}$$

where the $|\langle \Psi(T, H, \tilde{u}^j, \mu_{real}, \Psi_0), e_i\rangle|^2$ are obtained directly by measurements in the laboratory.

Then the average difference is obtained by:

$$\tilde{\mathcal{J}}((u^j)_{1\leq j\leq N_u}, \mu, (\tilde{u}^j)_{1\leq j\leq N_u}, \mu_{real}) = \frac{1}{N_u}\sum_{j=1}^{N_u} \mathcal{J}(u^j, \mu, \tilde{u}^j, \mu_{real}). \quad (4.9)$$

Here the distance we use is $\mathcal{W}_1$, the 1-Wasserstein. Other distances could also be used, for example $\mathcal{W}_2$.

4.3.1 Main algorithm

We start the optimization with an initial guess μ^0.

Algorithm 4.1 *The iteration $n \geq 1$ consists in the following steps:*

1. *Randomly choose N_u sets of amplitudes $(A_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$ and phases $(\theta_l^j)_{1\leq j\leq N_u, 1\leq l\leq N_\omega}$;*
2. *Use the Algorithm 1.1 to construct N_r random realizations of phase noises $(\delta\theta_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r}$;*
3. *Calculate $(u^j)_{1\leq j\leq N_u}$ via formula (4.6);*
4. *minimize $\mu \to \tilde{\mathcal{J}}((u^j)_{1\leq j\leq N_u}, \mu, (\tilde{u}^j)_{1\leq j\leq N_u}, \mu_{real})$ and set μ^n to be a minimizer (in practice a close approximation);*

We stop the algorithm when $\mathcal{J}$ is smaller than a target error.

4.4 Numerical tests

The numerical test use the 4-level system ($N = 4$) in [14] having:

$$H = \begin{pmatrix} 0.0833 & -0.0038 & -0.0087 & 0.0041 \\ -0.0038 & 0.0647 & 0.0083 & 0.0038 \\ -0.0087 & 0.0083 & 0.0036 & -0.0076 \\ 0.0041 & 0.0038 & -0.0076 & 0.0357 \end{pmatrix}$$

and

$$\mu_{real} = \begin{pmatrix} 0 & 5 & -1 & 0 \\ 5 & 0 & 6 & -1.5 \\ -1 & 6 & 0 & 7 \\ 0 & -1.5 & 7 & 0 \end{pmatrix}.$$

We set $N_u = 36$. The amplitudes are chosen at random, uniformly in $[0, 0.1 \cdot \frac{\|H\|_{l^\infty}}{\|\mu_{real}\|_{l^\infty}} = 0.0012]$ and the phases in $[0, 2\pi]$. We set the Gaussian envelope in time $S(t) = \exp\left(-40\left(\frac{t-T/2}{T}\right)^2\right)$.

The initial guess is:

$$\mu^0 = \begin{pmatrix} 0 & 5.1295 & -0.9762 & 0.0962 \\ 5.1295 & 0 & 5.5100 & -1.6434 \\ -0.9762 & 5.5100 & 0 & 7.6117 \\ 0.0962 & -1.6434 & 7.6117 & 0 \end{pmatrix}. \tag{4.10}$$

The relative error of this initial guess is about 8.7%. The initial guess μ^0 is obtained by multiplying element-wise μ_{real} by uniform random variables in $[0.9, 1.1]$ (i.e., up to 10% average relative error).

The observables arising from the Hamiltonian H are projections $\{|e_1\rangle\langle e_1|, |e_2\rangle\langle e_2|, |e_3\rangle\langle e_3|, |e_4\rangle\langle e_4|\}$ to eigenstates:

$$\begin{aligned}
|e_1\rangle &= \begin{pmatrix} 0.0845 & -0.1313 & 0.9651 & 0.2101 \end{pmatrix}^T, \\
|e_2\rangle &= \begin{pmatrix} -0.1305 & -0.0856 & -0.2103 & 0.9651 \end{pmatrix}^T, \\
|e_3\rangle &= \begin{pmatrix} 0.2118 & 0.9647 & 0.0838 & 0.1325 \end{pmatrix}^T, \\
|e_4\rangle &= \begin{pmatrix} 0.9649 & -0.2118 & -0.1314 & 0.0830 \end{pmatrix}^T. \qquad (4.11)
\end{aligned}$$

In the basis $\{e_1, e_2, e_3, e_4\}$ the Hamiltonian H is diagonal with eigenvalues $E_1 = 0$, $E_2 = 0.0365$, $E_3 = 0.0651$, $E_4 = 0.0857$. Thus $N_\omega = 6$, $\omega_1 = 0.0365$, $\omega_2 = 0.0651$, $\omega_3 = 0.0856$, $\omega_4 = 0.0286$, $\omega_5 = 0.0492$ and $\omega_6 = 0.0206$.

We set the final time $T = 3200$ which is about 10 periods of the smallest transition frequency $2\pi/(E_4 - E_3) = 314$ in H.

4.4.1 Choice of the correlation operator

We will use the square exponential covariance matrix (see Example 1.3 for the numerical tests). Then

$$\Sigma_{l,l'} = \sigma^2 e^{-\frac{(\omega_l - \omega_{l'})^2}{\beta}} \qquad (4.12)$$

with β a length-scale parameter. We set $\sigma = 0.1$, so that the noises are not negligible.

The value of β should be appropriate such that the noises are correlated but not too correlated. Figure 4.1, Figure 4.2 and Figure 4.3 shows the realizations for $\beta = 1$, $\beta = 0.1$ and $\beta = 0.01$.

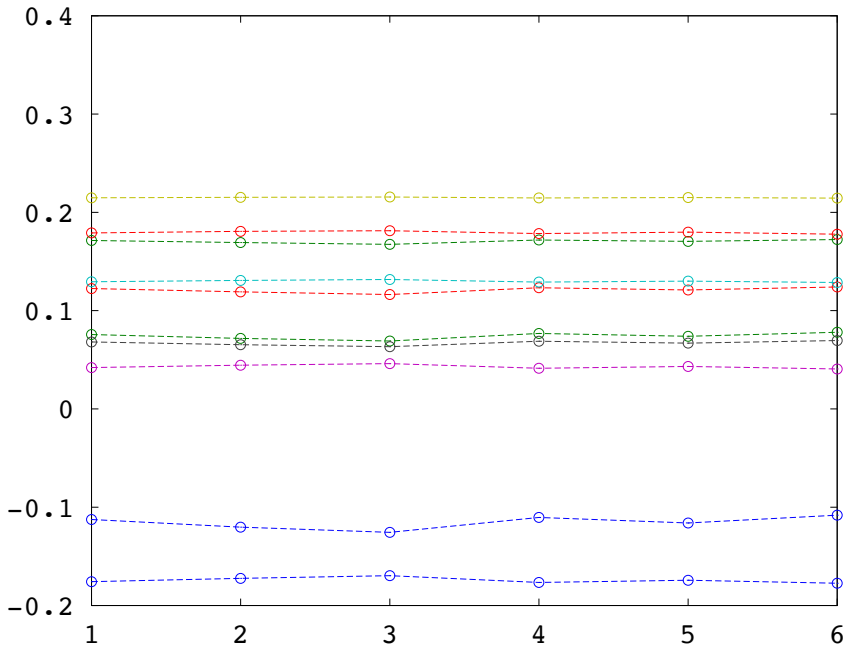

Figure 4.1: 10 realizations of the phase noises with $\beta = 1$. The x-axis represents the 6 frequencies $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6]$

We remark that when $\beta = 1$, it is too large that the noises are too correlated. The $\delta\theta_{\omega_l}$ take almost the same values.

The choice of $\beta = 0.1$ seems to be an adequate choice. The noises influenced each other.

When $\beta = 0.01$, the noises are not enough correlated.

4.4.2 Choice of the number of realizations

The number of realization N_r should be large enough such that the error due to the choice of realizations is negligible compared to the error due to the dipole moment μ.

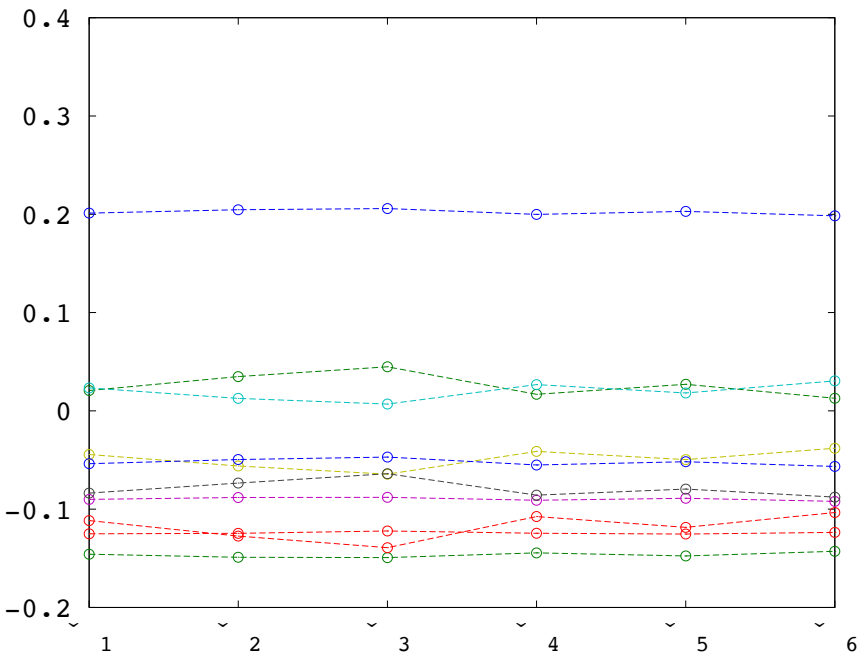

Figure 4.2: 10 realizations of the phase noises with $\beta = 0.1$. The x-axis represents the 6 frequencies $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6]$

For any amplitudes $(A_l)_{1\leq l\leq N_\omega}$ and phases $(\theta_l)_{1\leq l\leq N_\omega}$ randomly chosen, and for any dipole moment μ such that the relative error between μ and μ_{real} is about 1%, we simulate the distribution of the controls u_1 and u_2 for two different groups of realizations $(\delta\theta^1_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r}$ and $(\delta\theta^2_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r}$ respectively with the formula (4.5). The difference between the two laws: $\mathcal{J}(u_1, \mu, u_2, \mu)$ should be negligible compared to the difference between the simulation results with μ and the real measurements with μ_{real} for the same input of amplitudes and phases. This second difference is around 0.02.

Let us test for $N_r = 1000$. We construct 10 different μ which have relative errors about 1% with μ_{real}. These are obtained by multiplying element-wise μ_{real} by uniform random variables in $[0.99, 1.01]$. For each μ, 10 couples of amplitudes $(A_l)_{1\leq l\leq N_\omega}$ and phases $(\theta_l)_{1\leq l\leq N_\omega}$ are randomly chosen in $[0, 0.0012]$ and $[0, 2\pi]$. And for each triple (dipole moment μ, amplitudes $(A_l)_{1\leq l\leq N_\omega}$ and phases $(\theta_k)_{1\leq k\leq N_\omega}$), we calcu-

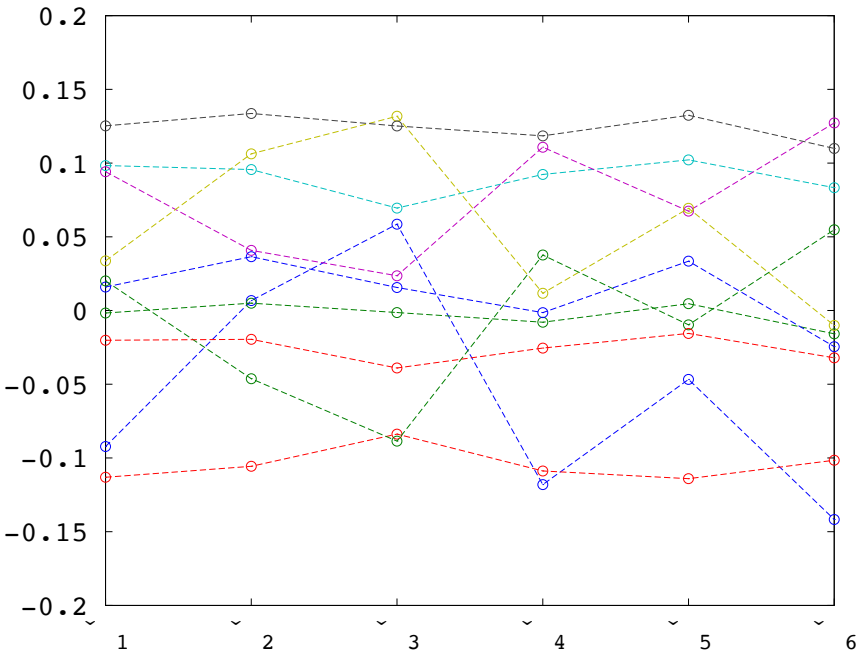

Figure 4.3: 10 realizations of the phase noises with $\beta = 0.01$. The x-axis represents the 6 frequencies $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6]$

late the difference $\mathcal{J}$ for 2 set of realizations $(\delta\theta^1_{l.k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r}$ and $(\delta\theta^2_{l,k})_{1\leq l\leq N_\omega, 1\leq k\leq N_r}$ randomly created. The 100 differences are mainly concentrated around 0.0017, which can be considered negligible compared to 0.02.

The Figure 4.4 shows how the 100 differences are distributed.

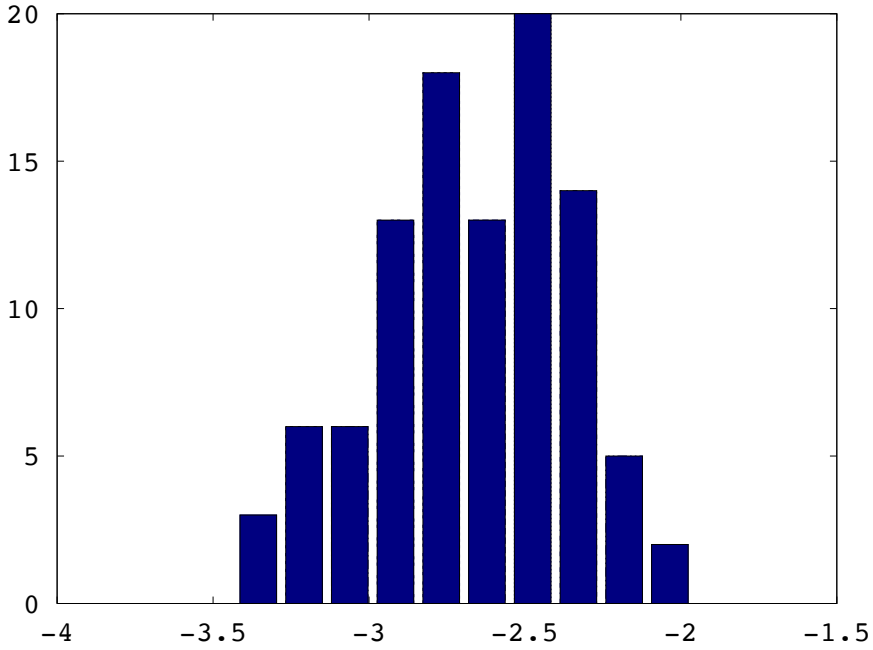

Figure 4.4: The histograms of the (base 10) logarithm of $\mathcal{J}$.

One of the μ chosen is

$$\mu_{sample} = \begin{pmatrix} 0 & 4.9861 & -1.0022 & 0.0099 \\ 4.9861 & 0 & 5.9995 & -1.4911 \\ -1.0022 & 5.9995 & 0 & 6.9376 \\ 0.0099 & -1.4911 & 6.9376 & 0 \end{pmatrix}.$$

The relative error between μ_{sample} and μ_{real} is 0.89%. The Figure 4.5 and Figure 4.6 shows the histograms of the distributions of measurements simulated by different realizations. The two figures are very similar,

which means the choice of realizations is not an important issue. The average difference $\tilde{\mathcal{J}}$ is about 0.0017.

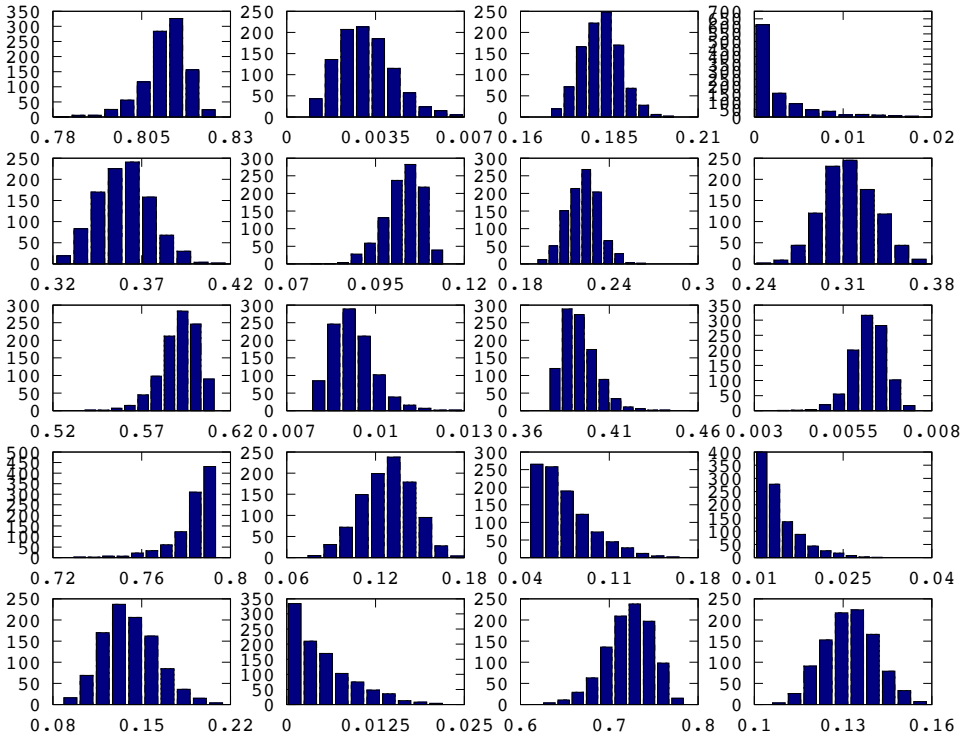

Figure 4.5: The 5×4 sized histograms of the simulated distribution of $\langle O\Psi, \Psi \rangle$. The 5 rows represents the 5 couples of amplitudes and phases chosen and the 4 columns represents the 4 observables.

We also test for other values. For $N_r = 100$, $N_r = 500$, and $N_r = 10000$ the average difference are 0.006, 0.0029 and 0.00061 respectively. This means if we want to recover the dipole moment μ_{real} within 1% of relative error, the choice of N_r should not be less than 500.

4.4.3 Numerical results

We set $N_r = 1000$. After 50 iterations, we obtain

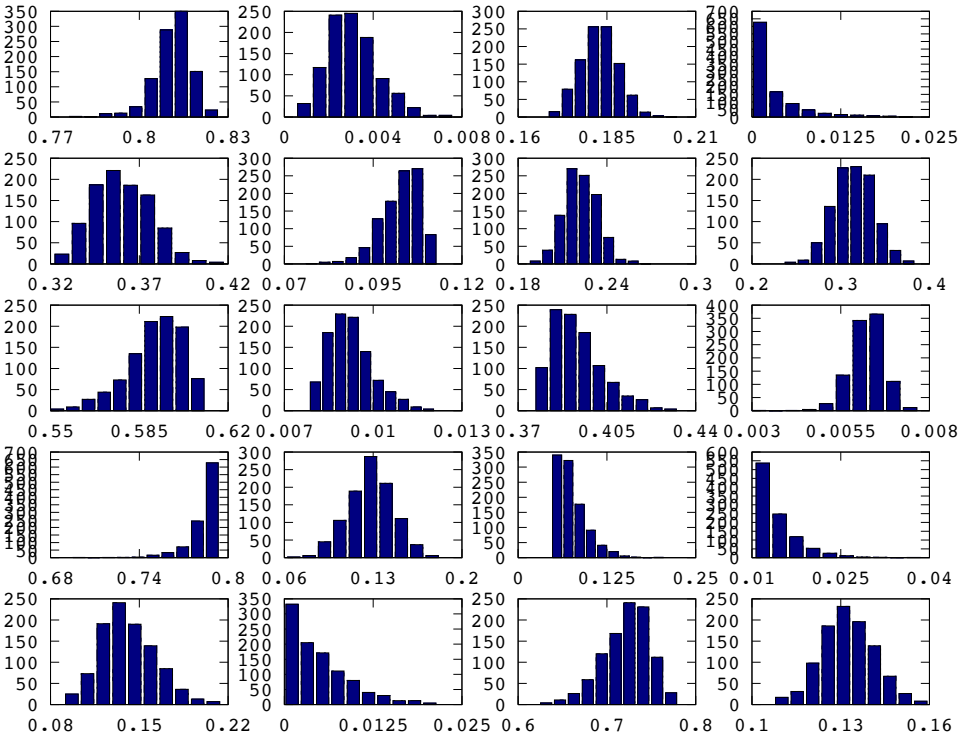

Figure 4.6: The 5×4 sized histograms of the simulated distribution of $\langle O\Psi, \Psi \rangle$. The 5 rows represents the 5 couples of amplitudes and phases chosen and the 4 columns represents the 4 observables. The amplitudes and the phases are the same as in the previous figure while the realizations are different.

$$\mu^{50} = \begin{pmatrix} 0 & 4.9781 & -0.9307 & -0.0063 \\ 4.9781 & 0 & 5.9568 & -1.5189 \\ -0.9307 & 5.9568 & 0 & 7.0682 \\ -0.0063 & -1.5189 & 7.0682 & 0 \end{pmatrix},$$

which corresponds to about 1% of relative error.

The Figure 4.7 shows the relative error $err = \frac{\|\mu^n - \mu_{real}\|_\infty}{\|\mu_{real}\|_\infty}$ for the 100 first iterations.

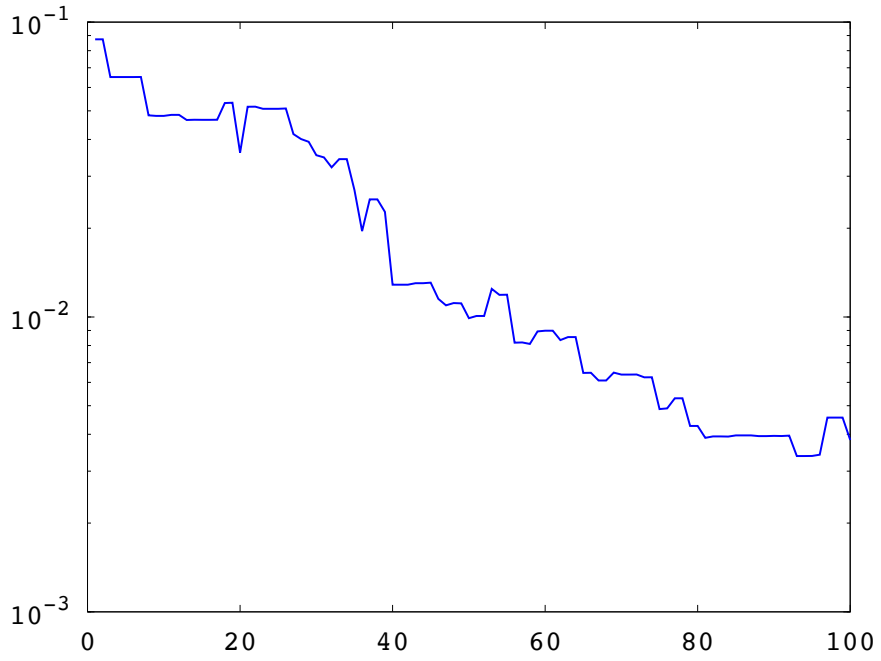

Figure 4.7: The relative error for 100 iterations using the square exponential covariance function.

In fact, after 80 iterations, we obtain

$$\mu^{80} = \begin{pmatrix} 0 & 4.9889 & -1.0180 & 0.0068 \\ 4.9889 & 0 & 6.0066 & -1.5029 \\ -1.0180 & 6.0066 & 0 & 7.0299 \\ 0.0068 & -1.5029 & 7.0299 & 0 \end{pmatrix},$$

which corresponds to an relative error of 0.42%. And the average difference $\tilde{\mathcal{J}}$ is about 0.007. This explains why the optimization algorithm is difficult to continue. Because the error from the choice of the distributions is not negligible now. If we want to have more accurate results, we should increase the number of realizations N_r.

4.4.4 Numerical results using the exponential covariance function

In this section we use the exponential covariance function, which corresponds to the Ornstein-Uhlenbeck process (see Example 1.8). Now

$$\Sigma_{l,l'} = \sigma^2 e^{-\frac{|\omega_l - \omega_{l'}|}{\beta'}} \tag{4.13}$$

with β' a parameter to be determinate and $\sigma = 0.1$.

A suitable value of β' for that the noises are enough correlated is $\beta' = 2$. Figure 4.8 shows the realizations in this case.

We start with the same μ^0 (see equation 4.10) and $N_r = 100$. After 60 iterations, we obtain

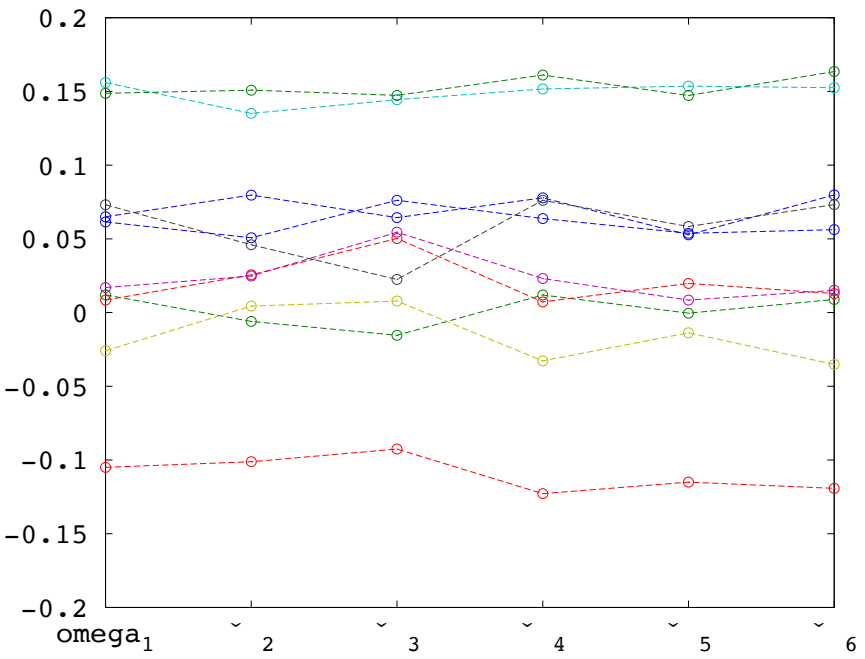

Figure 4.8: 10 realizations of the phase noises with $\beta' = 2$ in the equation (4.13). The x-axis represents the 6 frequencies $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6]$.

$$\mu^{60} = \begin{pmatrix} 0 & 5.0641 & -1.0585 & 0.0068 \\ 5.0641 & 0 & 5.9642 & -1.5139 \\ -1.0585 & 5.9642 & 0 & 7.0294 \\ 0.0068 & -1.5139 & 7.0294 & 0 \end{pmatrix},$$

which corresponds to an relative error of 0.91%. And the average difference $\tilde{\mathcal{J}}$ is about 0.014. The optimization algorithm can not continue any more because the error from the choice of the distributions becomes significant.

The Figure 4.9 shows the relative error $err = \frac{\|\mu^n - \mu_{real}\|_\infty}{\|\mu_{real}\|_\infty}$ for the 100 first iterations.

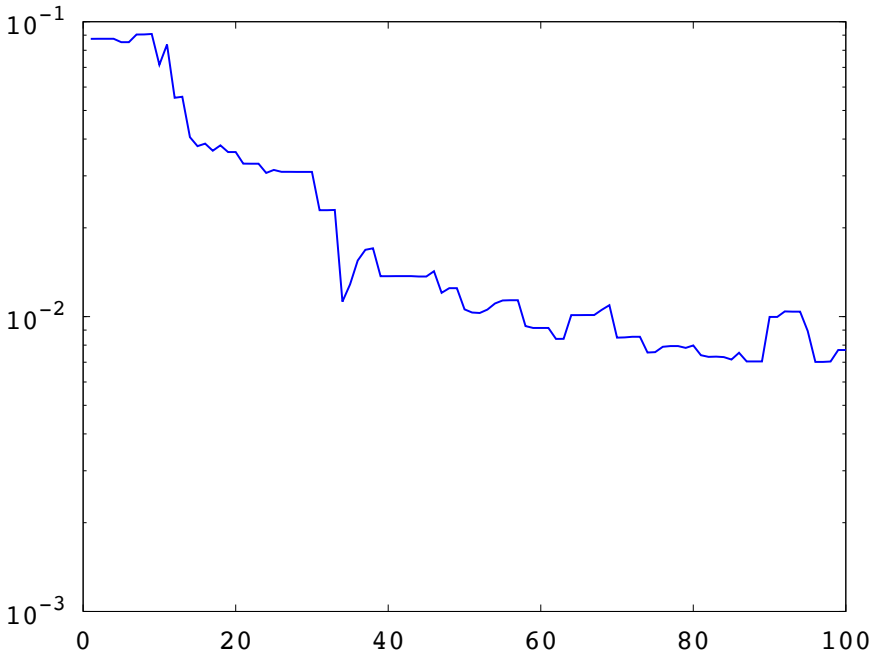

Figure 4.9: The relative error for 100 iterations using the exponential covariance function.

4.4.5 Numerical results using Brownian motion

In this section we use the Brownian motion (see Example 1.6). Now

$$\Sigma_{l,l'} = \sigma^2 \min(\omega_l, \omega_{l'}). \tag{4.14}$$

with $\sigma = 0.1$.

Figure 4.10 shows the realizations in this case.

We start with the same μ^0 (see equation 4.10) and $N_r = 100$. After 60 iterations, we obtain

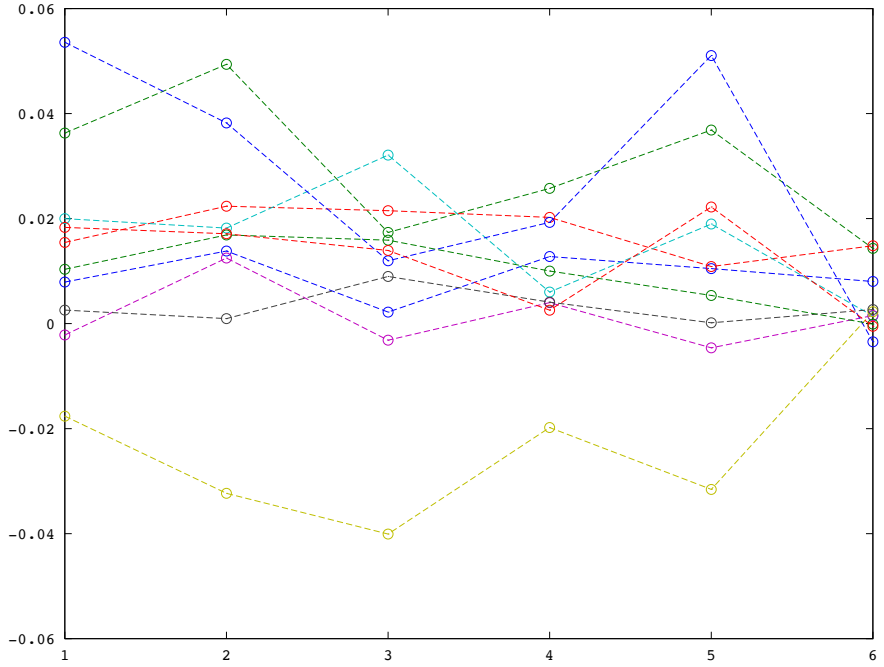

Figure 4.10: 10 realizations of the phase noises with Brownian motion (see equation 4.14, $\sigma = 0.1$). The x-axis represents the 6 frequencies $[\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6]$.

$$\mu^{60} = \begin{pmatrix} 0 & 5.0043 & -1.0138 & 0 \\ 5.0043 & 0 & 6.0078 & -1.5015 \\ -1.0138 & 6.0078 & 0 & 6.9834 \\ 0 & -1.5015 & 6.9834 & 0 \end{pmatrix},$$

which corresponds to an relative error of 0.24%.

The Figure 4.11 shows the relative error $err = \frac{\|\mu^n - \mu_{real}\|_\infty}{\|\mu_{real}\|_\infty}$ for the 100 first iterations.

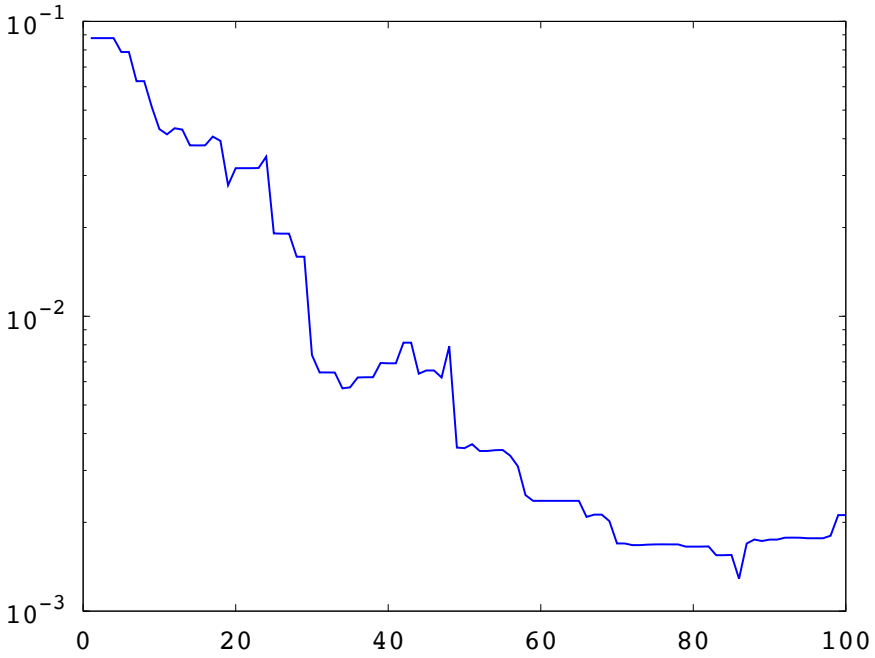

Figure 4.11: The relative error for 100 iterations using the Brownian motion.

Bibliography

[1] P. Abrahamsen and Norsk regnesentral. *A Review of Gaussian Random Fields and Correlation Functions*. Norsk Regnesentral/Norwegian Computing Center, 1997. ISBN 9788253904351. URL https://books.google.fr/books?id=InByAAAACAAJ.

[2] Francesca Albertini and Domenico D'Alessandro. Notions of controllability for quantum mechanical systems. In *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, volume 2, pages 1589–1594. IEEE, 2001.

[3] Claudio Altafini and Francesco Ticozzi. Modeling and control of quantum systems: An introduction. *IEEE Trans. Automat. Contr.*, 57:1898–1917, 2012.

[4] Lucie Baudouin and Alberto Mercado. An inverse problem for Schrodinger equations with discontinuous main coefficient. *Applicable Analysis*, 87(10-11):1145–1165, 2008. doi: 10.1080/00036810802140673. URL http://dx.doi.org/10.1080/00036810802140673.

[5] Karine Beauchard, Jean-Michel Coron, and Pierre Rouchon. Controllability issues for continuous-spectrum systems and ensemble controllability of Bloch equations. *Comm. Math. Phys.*, 296(2): 525–557, 2010. ISSN 0010-3616. doi: 10.1007/s00220-010-1008-9. URL http://dx.doi.org/10.1007/s00220-010-1008-9.

[6] Mohamed Belhadj, Julien Salomon, and Gabriel Turinici. Ensemble controllability and discrimination of perturbed bi-

linear control systems on connected, simple, compact lie groups. *European Journal of Control*, 22(0):23 – 29, 2015. ISSN 0947-3580. doi: http://dx.doi.org/10.1016/j.ejcon.2014.12.003. URL http://www.sciencedirect.com/science/article/pii/S094735801500014X.

[7] S. Bonnabel, M. Mirrahimi, and P. Rouchon. Observer-based Hamiltonian identification for quantum systems. *Automatica*, 45(5): 1144 – 1155, 2009. ISSN 0005-1098. doi: http://dx.doi.org/10.1016/j.automatica.2008.12.007. URL http://www.sciencedirect.com/science/article/pii/S0005109809000235.

[8] H.P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, 2002. ISBN 9780198520634. URL https://books.google.fr/books?id=0Yx5VzaMYm8C.

[9] Constantin Brif, Raj Chakrabarti, and Herschel Rabitz. Control of quantum phenomena: past, present and future. *New Journal of Physics*, 12(7):075008, 2010. URL http://stacks.iop.org/1367-2630/12/i=7/a=075008.

[10] Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloë. *Quantum Mechanics, vol 1*. Wiley, New-York, 1977.

[11] Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloë. *Quantum Mechanics, volume 2*. Wiley, VCH, reprint of the 2nd edition, 2005.

[12] Jean-Michel Coron, Andreea Grigoriu, Catalin Lefter, and Gabriel Turinici. Quantum control design by Lyapunov trajectory tracking for dipole and polarizability coupling. *New Journal of Physics*, 11: 105034 (22p), 2009. URL https://hal.archives-ouvertes.fr/hal-00410285.

[13] Domenico D'Alessandro. *Introduction to quantum control and dynamics*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2008. ISBN 978-1-58488-884-0; 1-58488-884-9.

[14] Ashley Donovan and Herschel Rabitz. Exploring the Hamiltonian inversion landscape. *Phys. Chem. Chem. Phys.*, 16:15615–15622, 2014. doi: 10.1039/C4CP02209B. URL http://dx.doi.org/10.1039/C4CP02209B.

[15] Pascal Dufour, Guy Rousseau, Michel Piché, and Nathalie McCarthy. Optical noise reduction in a femtosecond ti: sapphire laser pumped by a passively stabilized argon ion laser. *Optics communications*, 247(4):427–436, 2005.

[16] John W. Eaton, David Bateman, Soren Hauberg, and Rik Wehbring. *GNU Octave version 4.0.0 manual: a high-level interactive language for numerical computations.* http://www.gnu.org/software/octave/doc/interpreter, 2015. URL http://www.gnu.org/software/octave/doc/interpreter.

[17] John W. Eaton et al. *GNU Octave 4.0.0.* http://www.octave.org, 2015. URL http://www.octave.org.

[18] Ying Fu and Gabriel Turinici. Quantum Hamiltonian and dipole moment identification in presence of large control perturbations. doi: 10.1051/cocv/2016026 (accepted by ESAIM: Control, Optimisation and Calculus of Variations), September 2014. URL http://dx.doi.org/10.1051/cocv/2016026.

[19] Ying Fu, Herschel Rabitz, and Gabriel Turinici. Hamiltonian identification in presence of large control field perturbations. JPhysA-106056.R1 (accepted by Journal of Physics A: Mathematical and Theoretical), October 2016. URL https://hal.archives-ouvertes.fr/hal-01326060.

[20] J. M. Geremia and H. Rabitz. Optimal Hamiltonian identification: The synthesis of quantum optimal control and quantum inversion. *The Journal of Chemical Physics*, 118(12):5369–5382, 2003. doi: http://dx.doi.org/10.1063/

1.1538242. URL http://scitation.aip.org/content/aip/journal/jcp/118/12/10.1063/1.1538242.

[21] Xun Gu, Selcuk Akturk, and Rick Trebino. Spatial chirp in ultrafast optics. *Optics Communications*, 242(4-6):599 – 604, 2004. ISSN 0030-4018. doi: http://dx.doi.org/10.1016/j.optcom.2004.09.004. URL http://www.sciencedirect.com/science/article/pii/S0030401804008855.

[22] Brian C Hall. An elementary introduction to groups and representations. *arXiv preprint math-ph/0005032*, 2000.

[23] Serge Haroche and Jean Michel Raimond. *Exploring the Quantum: Atoms, Cavities, and Photons*. Oxford Univ. Press, Oxford, 2006. URL https://cds.cern.ch/record/993568.

[24] David Hocker, Constantin Brif, Matthew D. Grace, Ashley Donovan, Tak-San Ho, Katharine Moore Tibbetts, Rebing Wu, and Herschel Rabitz. Characterization of control noise effects in optimal quantum unitary dynamics. *Phys. Rev. A*, 90:062309, Dec 2014. doi: 10.1103/PhysRevA.90.062309. URL http://link.aps.org/doi/10.1103/PhysRevA.90.062309.

[25] Eugene N Ivanov, Scott A Diddams, and Leo Hollberg. Experimental study of noise properties of a ti: sapphire femtosecond laser. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 50(4):355–360, 2003.

[26] V. Jurdjevic and H. J. Sussmann. Control systems on Lie groups. *J. Differ. Equations*, 12:313–329, 1972.

[27] Kaveh Khodjasteh and Lorenza Viola. Dynamical quantum error correction of unitary operations with bounded controls. *Phys. Rev. A*, 80:032314, Sep 2009. doi: 10.1103/PhysRevA.80.032314. URL http://link.aps.org/doi/10.1103/PhysRevA.80.032314.

[28] Kaveh Khodjasteh and Lorenza Viola. Dynamically error-corrected gates for universal quantum computation. *Phys. Rev. Lett.*, 102: 080501, Feb 2009. doi: 10.1103/PhysRevLett.102.080501. URL http://link.aps.org/doi/10.1103/PhysRevLett.102.080501.

[29] Claude Le Bris, Mazyar Mirrahimi, Herschel Rabitz, and Gabriel Turinici. Hamiltonian identification for quantum systems: well-posedness and numerical approaches. *ESAIM Control Optim. Calc. Var.*, 13(2):378–395, 2007. ISSN 1292-8119.

[30] J.-S. Li and N. Khaneja. Control of inhomogeneous quantum ensembles. *Phys. Rev. A*, 73:030302, 2006.

[31] Y. Maday and J. Salomon. A greedy algorithm for the identification of quantum systems. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 375–379, Dec 2009. doi: 10.1109/CDC.2009.5400691.

[32] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. ISBN 026218253X. URL http://www.amazon.com/Gaussian-Processes-Learning-Adaptive-Computation/dp/026218253X.

[33] E. Schrödinger. Are there quantum jumps? *The British Journal for the Philosophy of Science*, 3(10):109 – 123, August 1952.

[34] Alexandre M. Souza, Gonzalo A. Álvarez, and Dieter Suter. Experimental protection of quantum gates against decoherence and control errors. *Phys. Rev. A*, 86:050301, Nov 2012. doi: 10.1103/PhysRevA.86.050301. URL http://link.aps.org/doi/10.1103/PhysRevA.86.050301.

[35] Gabriel Turinici. Beyond bilinear controllability : applications to quantum control. In *Control of coupled partial differential equa-*

tions, volume 155 of *Internat. Ser. Numer. Math.*, pages 293–309, Oberwolfach, Allemagne, 2007. Birkhauser.

[36] Gabriel Turinici, Viswanath Ramakhrishna, Baiqing Li, and Herschel Rabitz. Optimal discrimination of multiple quantum systems: controllability analysis. *Journal of Physics A: Mathematical and General*, 37(1):273, 2004. URL http://stacks.iop.org/0305-4470/37/i=1/a=019.

[37] Cédric Villani. *Topics in optimal transportation*. Graduate studies in mathematics. American Mathematical Society, cop., Providence (R.I.), 2003. ISBN 0-8218-3312-X. URL http://opac.inria.fr/record=b1122739.

[38] Jun Zhang and Mohan Sarovar. Quantum hamiltonian identification from measurement time traces. *Physical review letters*, 113(8): 080401, 2014.

[39] Wusheng Zhu and Herschel Rabitz. Potential surfaces from the inversion of time dependent probability density data. *The Journal of Chemical Physics*, 111(2):472–480, 1999. doi: http://dx.doi.org/10.1063/1.479328. URL http://scitation.aip.org/content/aip/journal/jcp/111/2/10.1063/1.479328.

[40] I. R. Zola and H. Rabitz. The influence of laser field noise on controlled quantum dynamics. *The Journal of Chemical Physics*, 120(19):9009 – 9016, 2004.

Résumé

Dans le cadre du contrôle quantique bilinéaire, cette thèse étudie la possibilité de retrouver l'Hamiltonien et/ou le moment dipolaire à l'aide de mesures d'observables pour un ensemble grand de contrôles. Si l'implémentation du contrôle fait intervenir des bruits alors les mesures prennent la forme de distributions de probabilité. Nous montrons qu'il y a toujours unicité (à des phases près) des Hamiltoniens de du moment dipolaire retrouvés. Plusieurs modèles de bruit sont étudiés: bruit discrète constant additif et multiplicatif ainsi qu'un modèle de bruit dans les phases sous forme de processus Gaussien. Les résultats théoriques sont illustrés par des implémentations numériques.

Abstract

The problem of recovering the Hamiltonian and dipole moment, termed inversion, is considered in a bilinear quantum control framework. The process uses as inputs some measurable quantities (observables) for each admissible control. If the implementation of the control is noisy the data available is only in the form of probability laws of the measured observable. Nevertheless it is proved that the inversion process still has unique solutions (up to phase factors). Several models of noise are considered including the discrete noise model, the multiplicative amplitude noise model and a Gaussian process phase model. Both theoretical and numerical results are established.

Mots Clés

équation Schrödinger, système bilinéaire, contrôle quantique, identification Hamiltonien

Keywords

Schrödinger equation, bi-linear system, quantum control, inversion Hamiltonian