PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
PHYSICS FACULTY
PHYSICS INSTITUTE

Weak Value Amplification in an Optomechanical System with Mach-Zehnder Interferometer

by

Sergio José Carrasco Novoa

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Ph.D. in Physics

at

Pontificia Universidad Católica de Chile

Supervisor	:	Dr. Miguel Orszag (PUC Chile)
Committee	:	Dr. Jerónimo Maze (PUC Chile)
		Dr. Enrique Munoz (PUC Chile)
		Dr. Vitalie Eremeev (UDP Chile)

October, 2021
Santiago, Chile

To my parents.

Acknowledgments

I thank to the Physics Institute at Pontificia Universidad Católica de Chile for the support provided to me during all these years devoted to the obtention of my doctoral degree. I am grateful to the government of Chile for the financial support granted over four years by means of the Conicyt program. Special thanks to my thesis advisor, professor Miguel Orszag. I truthfully appreciate his commitment to my academic formation, his kindness and patience. His guidance has played a central role in my formation in the area of quantum measurement theory. Also, I would like to thank the support of my family and friends. Without them this project would simply not have been possible. Finally, I thank to God for his love and for giving me the possibility to study physics during all these years.

Abstract

In this thesis work an optomechanical system inside a Mach-Zehnder interferometer is studied from the perspective of the *weak value amplification effect*. The optomechanical system consists of a Fabry-Pérot cavity with a moving mirror in the middle. Single photons are post-selected in the detector in one of the output ports of the interferometer (dark port), which allows to enlarge the displacement caused by a single photon over the moving mirror of the cavity. Since the interaction between a single photon and the mirror is weak, the amplification factor of the displacement corresponds to a weak value. By making the initial and final states of the photon quasi-orthogonal, the weak value becomes large and the radiation pressure force exerted by the photon is increased, making a single photon behave as "many photons" will do. The amplification effect comes, however, at the cost of the lost of data. The usefulness of weak values for parameter estimation in our setup is analysed from the perspective of the Fisher information. Although the precision of the estimation does not change either by using weak values or by implementing measurements that do not rely on post-selection, in the first scenario all the information can be put in a small amount of post-selected events, which is a verification of a well known general result in the existing literature on the subject.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The concept of the *weak value of a quantum variable* was introduced in 1988 in a seminal article written by Aharonov, Albert and Vaidman [1] with the surprising tittle: "*How the result of a measurement of a component of the spin of a spin* 1/2 *particle can turn out to be 100?*". In this article the authors begin by considering a measurement of an observable $\hat{A}$ performed on a quantum system prepared in an initial pure state $|\psi_i\rangle$. Then, after the measurement of $\hat{A}$, a *second measurement* is performed. The results of the *first measurement* are taken into consideration only when the second measurement leaves the system in a pure final state $|\psi_f\rangle$. In any other case, the outcomes of the first measurement are simply ignored. The idea of the procedure is to examine the results of the measurement of $\hat{A}$ when both the initial and the final states of the system are known. This kind of ensembles are called pre- and post-selected quantum systems. The historical motivation to study measurements performed on these systems relates to the analysis of the suggestion that the process of measurement, or "reduction of the wave packet", introduces a time-asymmetric element into quantum theory [2]. However, the asymmetry may be actually caused *not* by the measurement itself but by the manner in which a conventional ensemble is constructed, i.e. by the specification of only the initial state of the system. By constructing a "time-symmetric ensemble" (specifying the initial and the final states), then the probability distribution of the outcomes of a measurement becomes time-symmetric [3]. This idea has even given rise to a new formalism of quantum mechanics, that has been

called the "two-state vector formalism of quantum mechanics" [4].

For a certain type of generalized measurements, called *weak measurements* and characterized by a "weak coupling" or a "weak interaction" between the system and the measurement device, the authors showed that the ensemble average of the outcomes of the first measurement was the weak value of $\hat{A}$, which was denoted by A_w and defined as

$$A_w = \frac{\langle \psi_f | \hat{A} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle}. \tag{1.1}$$

Consequently, the weak value was presented as the ensemble average of a weak measurement of an observable, performed on a pre- and post-selected quantum system. This is a *statistical* or *operational* definition of the weak value, analog to $\langle \psi_i | \hat{A} | \psi_i \rangle$, which is the ensemble average of a *projective measurement* of $\hat{A}$, performed on a quantum system prepared in the initial state $|\psi_i\rangle$. One can immediately notice that, unlike the standard expectation value, the weak value is not restricted to the interval $[\min\{a_i\}, \max\{a_i\}]$, where a_i denotes one of the eigenvalues of $\hat{A}$. Indeed, when the overlap between the initial an final states is small, the weak value becomes large and may exceed the range of eigenvalues. In this case, the weak value is said to be anomalous. In particular, in [1] the authors considered a pair of initial and final states such that a weak measurement of the spin along the z-direction (of a spin 1/2 particle) was equal to 100, namely, a value far larger than any of the two eigenvalues, ± 1. This feature has been used in precision metrology for the estimation of tiny interaction parameters, as will be seen along this work.

In my thesis I will focus on weak values of number operators. An anomalously large weak value of the number of particles will enlarge the effect of the particles on a second system. In particular, our aim consists to design and analyze an experiment in which the radiation pressure effect of a single photon can be enlarged using weak values. Indeed, a single photon inside an optical cavity exerts a tiny force over the walls of the cavity. The force displaces the walls by $4gx_0$, where g is a parameter that quantifies the coupling between a microscopic degree of freedom (the photon) and a "macroscopic" degree of freedom (the center of mass of the cavity wall) and x_0 are the zero-point fluctuations

of the wall. However, if the photon is pre- and post-selected, then it behaves as "many photons" will do. In fact, as it will be shown along this work, the displacement will be $4gx_0 \cdot F$, where F is the weak value of the number of photons. This result has been published in [5]. Our main motivation to analyze such a system is to contribute to the understanding of the meaning of a weak value. Currently, as was pointed out above, the weak value has been understood as an ensemble average, i.e. as an statistical quantity. However, it has also been argued that the weak value may be understood as a property of a quantum system. In this sense, it resembles more to an eigenvalue than to an expectation value.

A second task was to study whether the amplification effect may be helpful to estimate small optomechanical parameters in our experimental setup. In this work the precision of the estimation has been studied from the perspective of the quantum Fisher information, but also other figures of merit such as the mean squared error and the signal to noise ratio of an estimate have been taken into consideration. The amplification effect comes at a cost, which is the lost of data. Therefore, although one photon may behave as many photons, there will be at the end fewer photons available for the estimation. It turns out that the precision of the estimation does not change, either by using all the "standard" (pre-selected) photons or by considering the fewer pre- and post-selected photons. Consequently, from the perspective of parameter estimation, large weak values are helpful when the estimation needs to be performed from a small amount of data. These conclusions are also described in the preprint [6].

The structure of this thesis work is described below. Chapter 2 offers a short review on quantum measurement theory. Here the language of generalized measurements is introduced, which is based on POVM's, measurements operators, quantum operations and other related concepts. At the end of this chapter weak measurements of an observable are defined. The definition is made using the so called "von Neumann" model. The next chapter is devoted to the study of the theory of pre- and post-selected quantum measurements. In this chapter the weak value is formally introduced, together with its most popular applications. The different existing controversies regarding its interpretation as a property of a quantum system, or if there exists a classical analog to the phenomena of

large weak values, among others, are also put over the table. Chapter 4 is a long chapter in which the tools of parameter estimation are presented. In the first half of the chapter, from section 4.1 to section 4.8, the basic elements to deal with the problem of estimation are described (Fisher information, Cramér-Rao bound, among others). Throughout the second half of the chapter, in the section 4.9, the theory of parameter estimation is applied to weak measurements, both on only pre-selected ensembles and on pre- and post-selected quantum systems. In the final chapter, an experimental proposal to produce weak values of a photon number operator is presented (the variable is actually the difference of photons between two sides of a cavity with a moving mirror in the middle). The experiment consists of a Mach-Zehnder interferometer with an optomechanical system put along the arms of the interferometer. The setup allows to create optomechanical entanglement between a single photon and the center of mass of a moving wall (a large object formed by $\sim 10^{12}$ atoms, or more). Weak values are generated due to the detection of the photon in one of the detectors located at the output ports of the interferometer, a procedure that has been called dark port post-selection. Additionally, along this chapter it is explained that, when the measurement is not weak but strong, non-classical mechanical states of the measurement device can be generated via post-selection.

Chapter 2

Quantum measurement theory

This chapter may be considered as a short review on quantum measurement theory. In the first section, the theory of *generalized quantum measurements* is presented. The formalism introduced in this section will allow us to define later *weak measurements*, a kind of measurements that minimally disrupt the system under observation. In section 2.2 it will be shown how a generalized measurement can be implemented using an "ancilla system", and it is explained how Neumark's theorem ensures a one-to-one correspondence between a generalized measurement and an specific ancilla-system model. In the next section, we will consider a particular kind of interaction between the ancilla and the system, typically called the "von Neumann model". This model will enable us to define both projective and weak measurements of an observable, by strengthening or weakening the coupling between the ancilla and the system.

2.1 Generalized measurements

Quantum measurements affect the state of the system, i.e. they constitute a form of evolution of a quantum state. A general description of quantum measurements can be made using the language of *operations* and *effects* [8]. The formalism of quantum operations is not restricted to measurements, but it is also used to describe the transformation of a quantum system in a large class of physical processes, such as unitary evolution or the dynamics of an open quantum system.

Let R be the result of a measurement performed on a quantum system in an initial (normalized) state $\rho(t)$. The instant t represents the time at which the measurement process begins and R is a random variable that takes values in the space of results $\mathcal{R}$. Every observation of the random variable R, namely, every possible outcome of the measurement, will be denoted with the lowercase letter $r \in \mathcal{R}$. For every possible outcome r it is possible to define a *quantum operation*, denoted by $\mathcal{O}_r$, which is allows to describe the *selective evolution* of the system as follows

$$\tilde{\rho}_r(t+\tau) = \mathcal{O}_r \rho(t). \tag{2.1}$$

The state $\tilde{\rho}_r(t+\tau)$ corresponds to the *unnormalized* state of the system when the measurement has produced the outcome r. It is also called the (unnormalized) *conditioned state* of the system. The time τ corresponds to the duration of the measurement. It is clear from (2.1) that the quantum operation $\mathcal{O}_r$ is a *superoperator* or a *quantum map*, since it acts on an density operator and produces another (possible unnormalized). For any operation $\mathcal{O}_r$ to represent a physical process it has to fulfil three important properties, described below.

Every operation is a *convex linear map*. This means that, for probabilities $\{p_i\}$ and a set of density operators $\{\rho_i(t)\}$, the action of the quantum operation over the ensemble $\sum_i p_i \rho_i(t)$ is linear, namely,

$$\mathcal{O}_r \sum_i p_i \rho_i(t) = \sum_i p_i \mathcal{O}_r \rho_i(t). \tag{2.2}$$

Additionally, every quantum operation is a *trace preserving* or *trace decreasing* map, meaning that $0 \leqslant \mathrm{Tr}[\tilde{\rho}_r(t+\tau)] \leqslant 1$. As we will see shortly this property allows to assign probabilities to each possible outcome r. Finally, every operation is a *completely positive map*. This third property means that the "enlarged" superoperator $\mathbb{1} \otimes \mathcal{O}_r$ is a positive map, where $\mathbb{1}$ is the identity superoperator acting on a second Hilbert space of operators. In this way, the action of the quantum operation over a system that is entangled with another will preserve the positiveness of the *joint* quantum state.

The probability that the result of the measurement R takes the value r will be denoted

by $P(r)$ and defined as

$$P(r) = \mathrm{Tr}[\mathcal{O}_r \rho(t)]. \tag{2.3}$$

Consequently, the normalized *conditioned state* of the system is obtained dividing (2.1) by the probability (2.3),

$$\rho_r(t+\tau) = \frac{\mathcal{O}_r \rho(t)}{P(r)}. \tag{2.4}$$

This equation describes the *selective evolution* of the system under a measurement and, unlike equation (2.1), it is non linear due to the normalization factor in the denominator.

On the other hand, the *unconditioned state* of the system after the measurement is given by

$$\rho(t+\tau) = \sum_r P(r)\rho_r(t+\tau) = \sum_r \mathcal{O}_r \rho(t) \equiv \mathcal{O}\rho(t). \tag{2.5}$$

The state $\rho(t+\tau)$ corresponds to the state of the system after the measurement, *when all the results are ignored*, i.e. averaging over all the possible results. The operation $\mathcal{O} = \sum_r \mathcal{O}_r$ generates therefore the *non-selective evolution* of the system under a measurement. It is easy to show that the operation $\mathcal{O}$, defined in this way, is indeed a quantum operation, namely, it satisfies the three described properties. Notice that the sum should be replaced by an integral if the space of results was continuum.

Every quantum operation admits a *Krauss representation* or *operator-sum representation*, namely,

$$\mathcal{O}_r \rho(t) = \sum_k \hat{A}_{r,k} \rho(t) \hat{A}_{r,k}^\dagger, \qquad \sum_k \hat{A}_{r,k} \hat{A}_{r,k}^\dagger \leqslant \mathbb{1}. \tag{2.6}$$

The operators $\hat{A}_{r,k}$ are called the *elements* of the operation. The index k goes over a finite or countably infinite set. It is worth to emphasise that this representation is not unique.

From (2.3) it is clear that

$$P(r) = \mathrm{Tr}[\mathcal{O}_r\rho(t)] = \mathrm{Tr}\Big[\sum_k \hat{A}_{r,k}\rho(t)\hat{A}^\dagger_{r,k}\Big] = \mathrm{Tr}\Big[\Big(\sum_k \hat{A}^\dagger_{r,k}\hat{A}_{r,k}\Big)\rho(t)\Big]. \tag{2.7}$$

Consequently, the operator that defines the statistics of the measurement is given by

$$\hat{E}_r = \sum_k \hat{A}^\dagger_{r,k}\hat{A}_{r,k}. \tag{2.8}$$

The *positive* operator $\hat{E}_r$, associated to the outcome r, is called a *probability operator* or *effect operator* [8]. As it is clear from (2.7), the effect operators allow to compute the probability of the outcome r to occur as

$$P(r) = \mathrm{Tr}\Big[\hat{E}_r\rho(t)\Big]. \tag{2.9}$$

The effect operators should satisfy the completeness condition

$$\sum_r \hat{E}_r = \mathbb{1}, \tag{2.10}$$

which ensures that all the probabilities add up to one (again, the sum might be an integral if the space of results r was continuum). This restriction entails that the non-selective evolution generated by $\mathcal{O}$ is trace-preserving,

$$\mathrm{Tr}[\mathcal{O}\rho(t)] = \mathrm{Tr}\Big[\sum_r \mathcal{O}_r\rho(t)\Big] = \mathrm{Tr}\Big[\sum_{r,k} \hat{A}_{r,k}\rho(t)\hat{A}^\dagger_{r,k}\Big] = \mathrm{Tr}\Big[\sum_r \hat{E}_r\rho(t)\Big] = \mathrm{Tr}[\rho(t)]. \tag{2.11}$$

The set $\{\hat{E}_r, r \in \mathcal{R}\}$ defines a *probability operator valued measure* on the space of results (POVM) and characterises the statistics of the measurement, i.e. the *information-gathering process*. On the other hand, the set of operations $\{\mathcal{O}_r, r \in \mathcal{R}\}$ describes the evolution of the system during the measurement process, i.e. the disturbance experienced by the system due to the measurement.

The language of *operations* and *effects* allows to classify measurement into different

classes. This classification can be done in terms of the information gain (in which case, additional restrictions are imposed over the effects), the disturbance (imposing restrictions over the operations) or both. There is a large variety of classes. In the following, we will briefly describe the ones that are relevant to our work. These classes, and their different intersections, are drawn in figure 2.1.

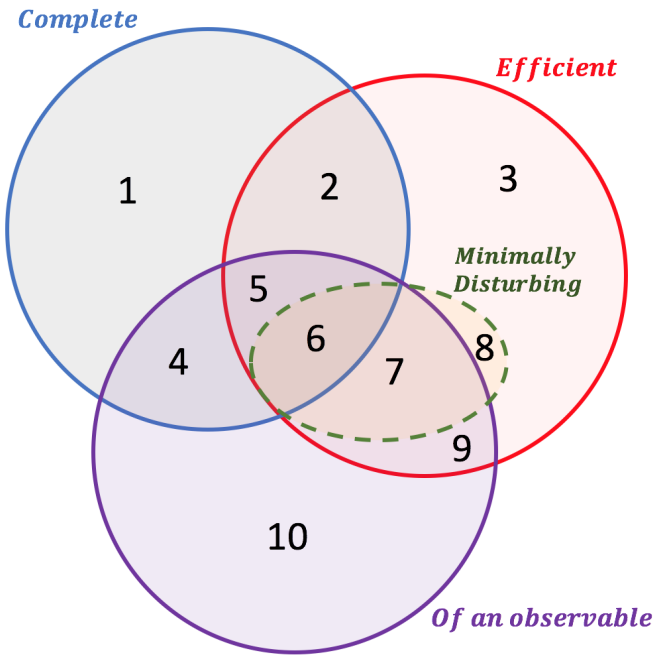


Figure 2.1: Different measurement classes: three main classes of measurements will be distinguished (complete, efficient and of an observable). Notice that minimally disturbing measurements (MDM) are a subset of efficient measurements. Projective measurements arise at the intersection of MDM and measurements of an observable (regions 6 and 7). Von Neumann measurements are the intersection of all groups (region 6).

2.1.1 Efficient measurements

A measurement is said to be *efficient* if, for every operation $\mathcal{O}_r$, there exists an operator $\hat{M}_r$ such that the action of the operation is described by $\mathcal{O}_r\rho(t) = \hat{M}_r\rho(t)\hat{M}_r^\dagger$. The operators $\hat{M}_r$ are called *measurement operators*. This kind of measurements are interesting because they do not introduce classical noise, i.e. any noise in the process can be regarded as quantum noise. In an efficient measurement, the conditional evolution transforms pure

states into pure states. For efficient measurements equation (2.3) tells that

$$P(r) = \mathrm{Tr}[\mathcal{O}_r \rho(t)] = \mathrm{Tr}\Big[\hat{M}_r \rho(t) \hat{M}_r^\dagger\Big] = \mathrm{Tr}\Big[\hat{M}_r^\dagger \hat{M}_r \rho(t)\Big]. \tag{2.12}$$

Therefore, $\hat{E}_r = \hat{M}_r^\dagger \hat{M}_r$ and the completeness condition (2.10) is $\sum_r \hat{M}_r^\dagger \hat{M}_r = \mathbb{1}$.

According to the polar decomposition theorem any operator can be written as the product of a positive operator and a unitary operator. Hence, in particular, a measurement operator may be written as $\hat{M}_r = \hat{U}_r \hat{P}_r$, where $\hat{U}_r$ is a unitary operator and $\hat{P}_r$ is a positive operator. As a consequence of the polar decomposition $\hat{E}_r = \hat{P}_r^2$ and, therefore, according to equation (2.9), the statistics of the measurement, i.e. the extraction of information, depends only on $\hat{P}_r$ and not on $\hat{U}_r$. Thus, the back-action produced by the operator $\hat{P}_r$ is the one responsible for the information gain. On the contrary, the unitary operator $\hat{U}_r$ provides an additional back-action on the system, which does not contribute to the gathering of information. For an efficient measurement, it is instructive to apply the polar decomposition theorem to equation (2.4) and rewrite it as

$$\rho_r(t+\tau) = \frac{\hat{M}_r \rho(t) \hat{M}_r^\dagger}{P(r)} = \hat{U}_r \Big(\frac{\hat{P}_r \rho(t) \hat{P}_r}{P(r)} \Big) U_r^\dagger, \tag{2.13}$$

which shows that an efficient measurement can be thought as the action of a measurement operator, represented solely by $\hat{P}_r$, followed by the unitary evolution generated by $\hat{U}_r$. This evolution can be viewed as an additional force applied to the system, when the measurement outcome is r. The application of forces, based on a certain measurement result, is called *feedback*. When the measurement adds no feedback, i.e. when $\hat{M}_r$ is a positive operator, the measurement is said to be *minimally disturbing*. In this case, the only disturbance of the system is due to the information gain.

2.1.2 Complete measurements

An important class of measurements are *complete measurements*. A measurement is complete when the conditioned state after the measurement is independent of the initial state. This means that $\rho_r(t+\tau)$ is not correlated with $\rho(t)$. This type of measurements

extract the maximum amount of information from the initial state.

When a *complete* measurement is *efficient*, then all the effects are rank one operators, i.e. $\hat{E}_r = |\psi_r\rangle\langle\psi_r|$, the set of operators $\{|\psi_r\rangle\langle\psi_r|, r\}$ being not necessarily orthonormal. In figure 2.1 efficient and complete measurement are represented by the union of the regions $2, 5$ and 6.

The complement of this class are *incomplete measurements*. In this case, the final state will depend on the initial state and more information can be gained with further measurements. Incomplete and efficient measurement have at least one effect whose rank is higher than one. In figure 2.1 incomplete and efficient measurements correspond to the groups $3, 7, 8$ and 9. Within this sub class of measurements (incomplete and efficient), when all the effects have rank higher than one, the measurement are said to be *weak measurements* [126]. This is a general definition of weak measurements. In this work, we will focus on weak measurements of an observable (they belong to the groups 7 or 9).

2.1.3 Measurement of an observable

A third important class of measurements are *measurements of observables*. A measurement belongs to this class if every effect operator $\hat{E}_r$ is a function of the observable. Thus, if $\hat{A}$ is an hermitian operator representing a physical variable, the measurement will be a measurement of $\hat{A}$ if $\forall r\ \hat{E}_r = E_r(\hat{A})$. Consequently, if the spectral decomposition of the observable is $\hat{A} = \sum_i a_i \hat{\Pi}_i$, then $\hat{E}_r = \sum_i \hat{\Pi}_i E_r(a_i)$ and the probability distribution of the outcomes is $P(r) = \sum_i E_r(a_i) \operatorname{Tr}\Big[\hat{\Pi}_i \rho(t)\Big]$.

2.1.4 Projective measurements

Let us consider an hermitian operator $\hat{A}$ with discrete eigenvalues (which may be degenerate) and whose spectral decomposition is given by $\hat{A} = \sum_i a_i \hat{\Pi}_i$, where a_i is a real eigenvalue and $\hat{\Pi}_i$ is a projector into the subspace of eigenvectors of $\hat{A}$ with eigenvalue a_i. The index number i labels each eigenvalue. The set of projectors $\hat{\Pi}_i$ is orthogonal, obeying $\hat{\Pi}_i \hat{\Pi}_j = \delta_{i,j} \hat{\Pi}_i$.

If a measurement of the observable $\hat{A}$ is minimally disturbing, then the measurement is said to be *projective*. In this case each possible outcome r can be "linked" to an eigenvalue

of $\hat{A}$. For a projective measurement of $\hat{A}$ the effects and measurement operators are equal, and correspond to projectors,

$$\hat{E}_r = \hat{M}_r = \hat{\Pi}_i. \tag{2.14}$$

Notice that the label i corresponds to the eigenvalue associated with the outcome r. Another way to write the effects operators is $\hat{E}_r = \delta_{\hat{A},\mathbb{1}\cdot i}$, which shows that the effect is a function of $\hat{A}$ and, therefore, according to the definition given in the previous section, the measurement corresponds to a measurement of $\hat{A}$. The projectors are obviously positive operators and the measurement is consequently a minimally disturbing measurement. Thus, the measurement described by (2.14) is indeed a projective measurement. For a projective measurement, equation (2.4) reduces to

$$\rho_r(t+\tau) = \frac{\mathcal{O}_r\rho(t)}{\mathrm{Tr}\Big[\hat{E}_r\rho(t)\Big]} = \frac{\hat{M}_r\rho(t)\hat{M}_r^\dagger}{\mathrm{Tr}\Big[\hat{M}_r^\dagger\hat{M}_r\rho(t)\Big]} = \frac{\hat{\Pi}_i\rho(t)\hat{\Pi}_i}{\mathrm{Tr}\Big[\hat{\Pi}_i\rho(t)\Big]}. \tag{2.15}$$

This is the *projection postulate*, or Lüder's rule [9], for updating the initial state after a measurement. Therefore, equation (2.4) can be understood as a generalization of the projection postulate. Also, it should be clear that for a projective measurement the probability to read the i-th eigenvalue is given by $P(i) = \mathrm{Tr}\Big[\hat{\Pi}_i\rho(t)\Big]$ (the normalization factor in the projection postulate). Notice that a projective measurement is not necessarily complete, since the projectors might not be rank one projectors, due to the possible degeneracy of the eigenvalues.

When all the projectors have rank one, i.e. when there is no degeneracy, the measurement corresponds to a *von Neumann measurement*. In this case, a projective measurement is complete and equation (2.15) becomes the projection postulate introduced by von Neumann, namely,

$$\rho_i(t+\tau) = \hat{\Pi}_i. \tag{2.16}$$

Therefore, the von Neumann measurement corresponds to an *efficient* (more specifically, to a minimally disturbing) and *complete* measurement of an *observable*, i.e. it lays at the

intersection of the first three classes introduced above.

2.2 Implementing operations with an *ancilla* system

An important fact about quantum operations is that they can be constructed from an *apparatus-system model*, i.e. using a second quantum system that is "attached" to the quantum system that we want to measure. This second system has received different names; *measurement device*, *ancilla*, *pointer*, *meter* [14] or *apparatus* [12]. The first system, the one which we want to measure, is called the *target system* or simply "the system". The apparatus-system model is represented in figure 2.2 and described below.

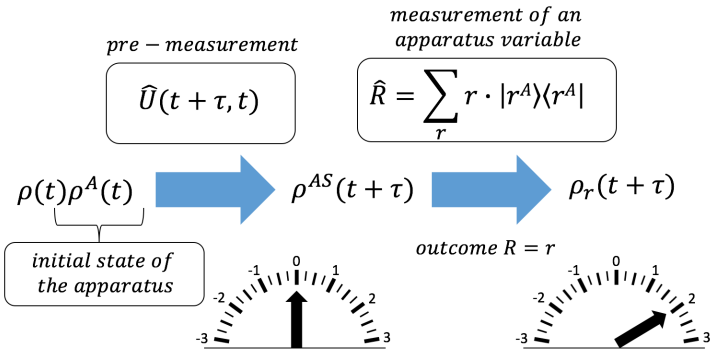


Figure 2.2: Representation of the measurement process. The first stage (pre-measurement) entangles the system with the measurement device. In the second stage the measurement device is measured by observing the variable $\hat{R}$.

We begin assuming that at an initial time t the joint state of the apparatus and the system is given by

$$\rho^{AS}(t) = \rho^A(t)\rho(t), \tag{2.17}$$

where $\rho^A(t)$ represents the initial state of the apparatus, $\rho(t)$ is the initial state of the system and $\rho^{AS}(t)$ the initial joint state of the apparatus and the system. In this work, we will use the superscript A to denote a state of the apparatus, AS to describe a joint state of the apparatus and the system, and no special notation for a system state. Notice that equation (2.17) assumes that initially the system is uncorrelated with the apparatus,

which is a reasonable assumption for a measurement.

The measurement is described as a "two-stage process". The first part is sometimes called *pre-measurement*. In this stage, the apparatus interacts with the system during a time τ, i.e. this part of the process has a certain duration. The interaction is described by a *unitary* operator $\hat{U}(t+\tau,t)$, which brings the initial joint state $\rho^{AS}(t)$ to the state $\rho^{AS}(t+\tau) = \hat{U}(t+\tau,t)\rho^{AS}(t)\hat{U}^{\dagger}(t+\tau,t)$, entangling the system with the apparatus. A specific form of $\hat{U}$ will be given in the next section, where the so called "von Neumann model" [14] is presented.

By introducing a (non unique) pure state decomposition for the initial state of the apparatus, $\rho^{A}(t) = \sum_k p_k \left|\psi_k^A\right\rangle\left\langle\psi_k^A\right|$, the joint state after the interaction can be written as

$$\rho^{AS}(t+\tau) = \sum_k p_k \hat{U}(t+\tau,t)\left|\psi_k^A\right\rangle \rho(t) \left\langle\psi_k^A\right| \hat{U}^{\dagger}(t+\tau,t). \tag{2.18}$$

The second part of the process corresponds to the *measurement of the apparatus*, in which a certain outcome is observed. The observation of the apparatus is modelled assuming that a *von Neumann measurement* of an apparatus variable $\hat{R}$ is made at an observation time $t_O > t+\tau$. Here we will assume that $t_O = t+\tau$, i.e. that apparatus is measured immediately after the interaction. This can be done without loss of generality, provided that no evolution takes place between $t+\tau$ and t_O.

For a von Neumann measurement, the measurement operators are orthogonal projectors, $\left|r^A\right\rangle\left\langle r^A\right|$, that project into the one-dimensional space spanned by the eigenvector $\left|r^A\right\rangle$ associated to the *r-th* eigenvalue of an hermitian operator $\hat{R}$ of the apparatus. Hence, applying the projection postulate, the *selective joint state* after the measurement is given by

$$\rho_r^{AS}(t+\tau) = \left|r^A\right\rangle\left\langle r^A\right| \otimes \frac{\left\langle r^A\right| \rho^{AS}(t+\tau) \left|r^A\right\rangle}{\mathrm{Tr}\left[\left\langle r^A\right| \rho^{AS}(t+\tau) \left|r^A\right\rangle\right]}. \tag{2.19}$$

Notice that the trace in the denominator is a partial trace over the system degrees of freedom. The state $\rho_r^{AS}(t+\tau)$ is a product state, i.e. after the measurement the apparatus is no longer correlated with the system. While the apparatus is left in a pure state $\left|r^A\right\rangle\left\langle r^A\right|$,

the conditioned system state is

$$\rho_r(t+\tau) = \frac{\left\langle r^A\right| \rho^{AS}(t+\tau) \left|r^A\right\rangle}{\mathrm{Tr}\left[\left\langle r^A\right| \rho^{AS}(t+\tau) \left|r^A\right\rangle \right]}. \tag{2.20}$$

The term in the numerator is the unnormalized conditioned state of the system after the measurement has produced the outcome r, while the denominator represents the probability for this result to occur, $P(r)$. In order to define a quantum operation $\mathcal{O}_r$ that describes the selective evolution of the system, let us consider the unnormalized conditioned state of the system,

$$\begin{aligned} \tilde{\rho}_r(t+\tau) &= \left\langle r^A\right| \rho^{AS}(t+\tau) \left|r^A\right\rangle \\ &= \sum_k p_k \left\langle r^A\right| \hat{U}(t+\tau,t) \left|\psi_k^A\right\rangle \rho(t) \left\langle \psi_k^A\right| \hat{U}^\dagger(t+\tau,t) \left|r^A\right\rangle . \end{aligned} \tag{2.21}$$

From this expression it is clear that the elements of the operation $\mathcal{O}_r$ can be defined as

$$\hat{A}_{r,k} \equiv \sqrt{p_k} \left\langle r^A\right| \hat{U}(t+\tau,t) \left|\psi_k^A\right\rangle , \tag{2.22}$$

that ensures that the action of the operation $\mathcal{O}_r\rho(t) = \sum_k \hat{A}_{r,k}\rho(t)\hat{A}_{r,k}^\dagger$ reproduces the result (2.21). Consequently, the effect operators are

$$\hat{E}_r = \sum_k \hat{A}_{r,k}^\dagger \hat{A}_{r,k} = \sum_k p_k \left\langle \psi_k^A\right| \hat{U}^\dagger(t+\tau,t) \left|r^A\right\rangle \left\langle r^A\right| \hat{U}(t+\tau,t) \left|\psi_k^A\right\rangle . \tag{2.23}$$

From (2.23) it is clear that $\sum_r \hat{E}_r = \mathbb{1}$, as is required from (2.10). The probability of getting the outcome r is given by $P(r) = \mathrm{Tr}\Big[\hat{E}_r\rho(t)\Big]$ that equals the term in the denominator of (2.20). The fact that the effects satisfy the completeness condition, ensures that the operations defined using the elements (2.22) are completely positive, trace decreasing and convex linear maps [15]. It can also be shown that the non-selective evolution of the system, generated by $\mathcal{O} = \sum_r \mathcal{O}_r$,

$$\rho(t+\tau) = \mathcal{O}\rho(t) = \sum_{r,k} \hat{A}_{r,k}\rho(t)\hat{A}_{r,k}^\dagger, \tag{2.24}$$

also satisfies these three conditions.

We have shown that operations and effects representing measurements can be constructed from an specific *apparatus-system model*, by using an apparatus (enlarging the Hilbert space) that starts in an initial state $\rho^A(t)$, couples to the system via a unitary operator $\hat{U}$ during a time τ, and whose variable $\hat{R}$ is finally measured (projectively) at a time t_O. An important observation is that the operations constructed in this way are not unique since the pure state ensemble decomposition used for $\rho^A(t)$ is not unique.

However, if the initial state of the apparatus is pure, $\rho^A(t) = \left|\psi^A\right\rangle\left\langle\psi^A\right|$, then each operation will have only one element. In this case the measurement is efficient and we can speak of measurement operators, which are $\hat{M}_r = \left\langle r^A\right| \hat{U}(t+\tau,t) \left|\psi^A\right\rangle$ and satisfy the completeness condition $\sum_r \hat{M}_r^\dagger \hat{M}_r = \mathbb{1}$. Therefore, from a unitary operator $\hat{U}$, a pure state $\left|\psi^A\right\rangle$ and an apparatus variable $\hat{R}$, the apparatus-system model generates a set of measurement operators $\{\hat{M}_r\}$ satisfying the completeness condition. The converse is also true. From a set of measurement operators $\{\hat{M}_r\}$ that satisfy $\sum_r \hat{M}_r^\dagger \hat{M}_r = \mathbb{1}$, an apparatus-system model may be constructed, i.e. $\hat{R}$, $\hat{U}$ and $\left|\psi^A\right\rangle$ may be found, such that $\hat{M}_r = \left\langle r^A\right| \hat{U}(t+\tau,t) \left|\psi^A\right\rangle$ and the completeness relation is fulfilled. This one-to-one correspondence between the set of measurement operators and an ancilla-system model is assured by *Neumark's theorem* [16].

2.3 The von Neumann model: strong and weak measurements of an observable

In the last chapter of the well known book [10], von Neumann described the general features that the interaction between the system and the meter should satisfy in order to have the character of a measurement of a physical variable. An interaction satisfying these features is called a "von Neumann model". Sometimes, it is also referred as the *von Neumann scheme* [17] or von *Neumann protocol* [14]. In this section we describe first these general characteristics. Afterwards, a particular form of interaction is presented from which *weak measurements of an observable* will be defined.

Let $\mathcal{B}_A = \{\left|m^A\right\rangle, m\}$ and $\mathcal{B} = \{|s\rangle, s\}$ be vector bases for the meter and the system,

respectively. Therefore, $\mathcal{B}_{AS} = \{\left|m^A\right\rangle |s\rangle, m, s\}$ is a basis of the Hilbert space of the meter and the system. Furthermore, assume that $\mathcal{B}_A$ is formed by the eigenstates of some meter operator $\hat{M}$, while $\mathcal{B}$ is the eigenbasis of a system operator $\widehat{S}$.

The operator that entangles the system with the meter (during the pre-measurement stage) is defined in terms of its action over the elements of $\mathcal{B}_{AS}$, as follows:

$$\hat{U}(t+\tau, t) \left|m^A\right\rangle |s\rangle = \left|U(m,s)^A\right\rangle |s\rangle, \tag{2.25}$$

where the states $\left|U(m,s)^A\right\rangle \in \mathcal{B}_A$. For a fixed m, each state $\left|U(m,s)^A\right\rangle$ should have no overlap with another state $\left|U(m,s')^A\right\rangle$, $s' \neq s$. This property will allow to distinguish the different eigenvectors of $\hat{S}$ by observing the variable $\hat{M}$ after the interaction. Consequently, we shall demand that $\left\langle U(m,s)^A \middle| U(m,s')^A\right\rangle = \delta_{s,s'}$.

On the other hand, for a fixed s, the function $U(m,s)$ should put in a one-to-one correspondence the basis element $\left|m^A\right\rangle$ with the basis element $\left|U(m,s)^A\right\rangle$. This condition is sufficient and necessary for $\hat{U}$ to be unitary.

The bases $\mathcal{B}_A$ and $\mathcal{B}$ are called *measurement basis* (for the apparatus or for the system). The eigenstates $\left|m^A\right\rangle$ are called *pointer states* and $\hat{M}$ is the *pointer variable*. Observing the pointer variable $\hat{M}$ shows on a certain scale the value of $\hat{S}$.

Let us assume that the apparatus is prepared in some state of $\mathcal{B}_A$ denoted by $\rho^A(t) = \left|0^A\right\rangle \left\langle 0^A\right|$. If the interaction between the apparatus and the meter satisfies the conditions described by von Neumann, and the apparatus is observed in the measurement basis (the pointer variable is observed), then the measurement operators are projectors,

$$\hat{M}_m = \left\langle m^A\right| \hat{U}(t+\tau) \left|0^A\right\rangle = \sum_{s'} \left\langle m^A \middle| U(0,s')^A\right\rangle \hat{\Pi}_{s'} = \hat{\Pi}_s. \tag{2.26}$$

Where $\hat{\Pi}_s$ projects into the space spanned by the eigenstates associated to the s-th eigenvalue of the system variable $\hat{S}$. In this way, if the apparatus is prepared in a "pointer state" and observed in the measurement basis, the interaction (2.25) defines a projective measurement of $\hat{S}$. It is important to highlight that the eigenvalue s is linked to the outcome m by the relation $m = U(0,s)$, i.e. given m and U, the eigenvalue s may be

obtained. Evidently, if such an eigenvalue of $\hat{S}$ does not exist, then the outcome m has zero probability to occur.

In [10] von Neumann provided an example of the unitary operator $\hat{U}$. It was derived from the interaction hamiltonian between the system and the meter $\hat{H} = \omega\hat{Q}\hat{P}$, where $\hat{Q}$ was the position of the system, $\hat{P}$ the momentum of the measurement device and ω some characteristic frequency describing the coupling. The masses of the system and the apparatus were assumed to be large, and therefore the kinetic energy (free hamiltonians of the system and the meter) was neglected.

This type of interaction is very common in the laboratory. In this work we will consider an interaction hamiltonian between the apparatus and the system of the form $\hat{H}(t) = g(t)\hat{A}\hat{P}$, where $\hat{A}$ is *discrete* variable of the system, with spectral decomposition given by $\sum_k a_k\hat{\Pi}_k$, and $\hat{P}$ is a *continuous* variable of the apparatus. The measurement starts at an initial time t and lasts for a time τ. The *instantaneous* coupling constant $g(t)$ between the apparatus and the system satisfies

$$\int_t^{t+\tau} g(t')dt' = g. \tag{2.27}$$

The constant g is a parameter that quantifies the *strength* of the measurement. Assuming that the free evolution of the system and the measurement device may be neglected during the interaction, the evolution operator for the measurement process is

$$\hat{U}(t+\tau,t) = \exp\Big\{-(i/\hbar)g\hat{A}\hat{P}\Big\} = \sum_k \exp\Big\{-(i/\hbar)ga_k\hat{P}\Big\}\hat{\Pi}_k. \tag{2.28}$$

Let $\hat{Q}$ be the conjugate variable to $\hat{P}$, namely, $[\hat{Q},\hat{P}] = i\hbar$. We will speak of $\hat{P}$ as the momentum of the apparatus and of $\hat{Q}$ as its position. Also, let $\{\left|q^A\right\rangle, q \in \mathbb{R}\}$ be the eigenbasis of $\hat{Q}$ and $\{|a_k\rangle, k\}$ the eigenbasis of $\hat{A}$. The action of the evolution operator (2.28) over the elements of these bases is given by

$$\hat{U}(t+\tau,t)\left|q^A\right\rangle|a_k\rangle = \left|(q+ga_k)^A\right\rangle|a_k\rangle. \tag{2.29}$$

Consequently, according to (2.25) the "position basis" and the basis formed by the eigenvectors of $\hat{A}$ constitute the measurement bases. In other words, the eigenvalues of $\hat{A}$ will be read in the final position of the meter. This basis is important, since more information is obtained when the meter is measured in this basis, as will be shown later.

We will assume that the system starts in a mixed state $\rho(t)$, uncorrelated with the meter. With regards to the apparatus it will be assumed that it is prepared in a pure state $\rho^A(t) = \left|\psi^A\right\rangle\left\langle\psi^A\right|$, where

$$\left|\psi^A\right\rangle \equiv \int_{-\infty}^{+\infty} \psi(q) \left|q^A\right\rangle dq = \int_{-\infty}^{+\infty} \psi(p) \left|p^A\right\rangle dp. \tag{2.30}$$

The basis $\{\left|p^A\right\rangle, p \in \mathbb{R}\}$ is the basis formed by the eigenvectors of the operator $\hat{P}$. The wave function, $\psi(q)$, is a assumed to be a *wave packet* that has a mean value of Q_0 and a standard deviation of σ_q. The momentum wave function, $\psi(p)$, is the Fourier transform of $\psi(q)$, and has mean momentum of P_0 and a spread $\sigma_p \geqslant \hbar/(2\sigma_q)$.

Since we are making the assumption that the initial state of the apparatus is pure, the measurement is efficient and each operation will have only one element. Consequently, it makes sense to speak of measurement operators for each possible outcome of the measurement. Notice that $\left|\psi^A\right\rangle$ is not a pointer state, since the measurement basis is continuous and thus the pointer states are not physically realizable.

Assume first that *the apparatus is observed in the measurement basis.* For the outcome q, the measurement and effect operators are:

$$\hat{M}_q = \psi(q - g\hat{A}) = \sum_k \psi(q - ga_k)\hat{\Pi}_k, \tag{2.31}$$

$$\hat{E}_q = |\psi(q - g\hat{A})|^2 = \sum_k |\psi(q - ga_k)|^2\hat{\Pi}_k. \tag{2.32}$$

The expression for $\hat{M}_q$ shows that each measurement operator is a weighted summation of the projectors, with weights equal to $\psi(q - ga_k)$, which simply corresponds to the initial wave function of the meter displaced by a distance ga_k. Also, it is clear that the measurement corresponds to a measurement of $\hat{A}$, because the effect operators are a function of this variable.

On the other hand, the probability that the measurement device is found at "exact" position q after the measurement is given by

$$P(q) = \mathrm{Tr}\Big[\hat{E}_q \rho(t)\Big] = \Big\langle |\psi(q - g\hat{A})|^2 \Big\rangle \tag{2.33}$$

$$= \sum_k |\psi(q - g a_k)|^2 \, \mathrm{Tr}\Big[\hat{\Pi}_k \rho(t)\Big]. \tag{2.34}$$

Obviously, $P(q)$ is a probability density function, because the position of the meter can not be measured with absolute precision. Notice also that expression (2.34) is a *mixture density*, since it corresponds to a convex combination of the initial probability density function, $|\psi(q)|^2$, displaced by a distance $g a_k$, with weights equal to $\mathrm{Tr}\Big[\hat{\Pi}_k \rho(t)\Big]$.

A *strong measurement* corresponds to a measurement in which the overlap between the different densities of the mixture is negligible, i.e. when $g\Delta_a \gg \sigma_q$, where Δ_a is the minimum separation between two consecutive eigenvalues of $\hat{A}$ (typically of the order of the unity). This situation is represented in figure 2.3, which shows that the mixture density consists of a series of peaks at different positions $Q_0 + g a_k$, which are clearly distinguishable. Each final "possible" position of the meter is in a one to one correspondence with the eigenvalues of the operator $\hat{A}$.

As the ratio g/σ_q decreases, the measurement becomes weaker and the overlap between the different densities of the mixture begins to increase. Figure 2.4 shows the case when $g/\sigma_q \sim 1$ and the initial wave function is Gaussian. In this case the initial probability density is normal, but it does not retain the "Gaussian shape" after the measurement. Indeed, the final density arises from a series of overlapped Gaussian densities. For a weaker measurement, however, the final distribution remains approximately normal (in the vast part of its domain except in the tails) and it is almost as if the density function was just shifted by an amount equal to $g\Big\langle \hat{A} \Big\rangle$. This fact is shown in figure 2.5.

By comparing figure 2.3 to figure 2.5 it is possible to see that for a weak measurement each outcome can not be "linked" to a particular eigenvalue, due to the significant overlap between the different wave packets. The weak measurement contains therefore additional *quantum noise*, because the meter has been prepared in a state with large σ_q or the coupling constant g is sufficiently small. This noise is quantum since it has been assumed that the

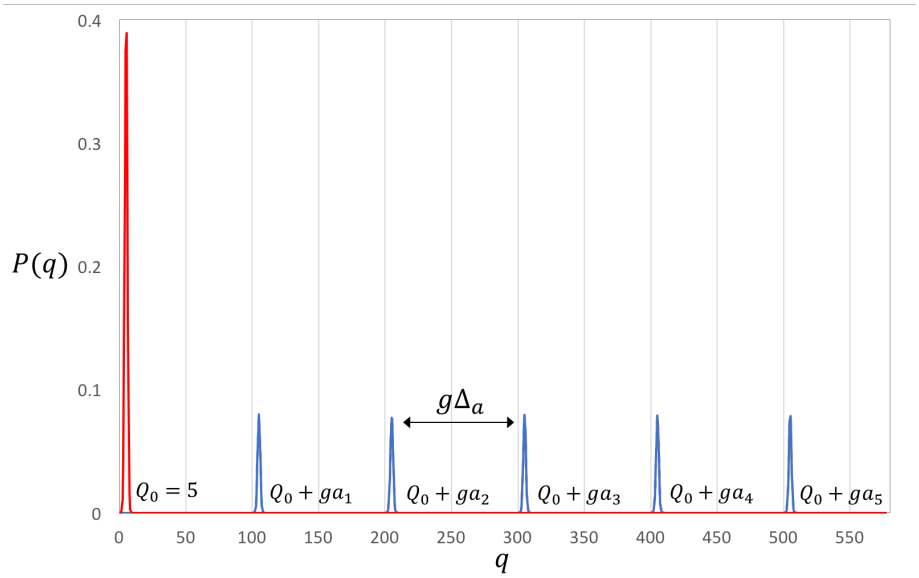


Figure 2.3: Strong measurement ($g = 100$) of an observable with 5 eigenvalues $a_i = i, i = 1...5$. The red curve shows the density function of the initial position of the meter, which is well defined and sharply peaked at some initial position, $Q_0 = 5$, with spread $\sigma_q = 1$. The density function for the final position of the meter is represented by blue curve, which is a series of distinguishable peaks.

meter started in a pure state.

The density (2.33) allows to compute the average position of the measurement device after the measurement,

$$Q_f = \int dqqP(q) = Q_0 + g\,\mathrm{Tr}\Big[\hat{A}\rho(t)\Big] = Q_0 + g\left\langle \hat{A} \right\rangle. \tag{2.35}$$

This shows that the displacement of the mean position is a linear function of g and is proportional to the expectation value of $\hat{A}$. This holds regardless of the value of g, i.e. both for a strong or a weak measurement. Analogously, the variance in the reading of the position is given by

$$Var(Q) = \int dq(q - Q_f)P(q) = \sigma_q^2 + g^2\left\langle \Delta\hat{A}^2 \right\rangle, \tag{2.36}$$

which shows that the dispersion of the position of the meter always increases after the

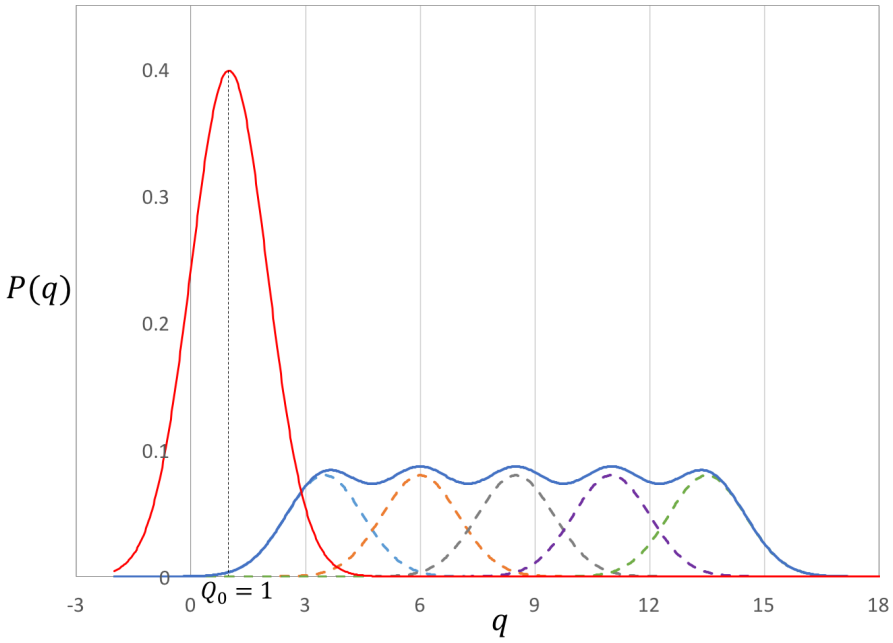


Figure 2.4: Weak measurement ($g = 2.5$) of an observable with 5 eigenvalues $a_i = i, i = 1...5$. The red curve shows the initial probability density function of the meter position, centered on $Q_0 = 1$ with spread $\sigma_q = 1$. The final density function (blue curve) consists of a series of well overlapped densities (dotted curves). A reading of the meter position gives less information than a strong measurement.

measurement [1].

It is instructive to see what happens with the system after the measurement. When the apparatus is found in position q, the conditioned state of the system is

$$\rho_q(t+\tau) = \frac{\sum_{k,k'} \psi(q-ga_k)\psi^*(q-ga_{k'})\hat{\Pi}_k\rho(t)\hat{\Pi}_{k'}}{\sum_k |\psi(q-ga_k)|^2 \,\mathrm{Tr}\Big\{\hat{\Pi}_k\rho(t)\Big\}}, \tag{2.37}$$

where ψ^* denotes complex conjugation. If the measurement is strong, then different wave packets do not overlap, namely, $\psi(q-ga_k)\psi^*(q-ga_{k'}) = \delta_{k,k'}|\psi(q-ga_k)|^2$, and the conditioned state of the system becomes

$$\rho_q(t+\tau) = \frac{\sum_k |\psi(q-ga_k)|^2\hat{\Pi}_k\rho(t)\hat{\Pi}_k}{\sum_k |\psi(q-ga_k)|^2 \,\mathrm{Tr}\Big\{\hat{\Pi}_k\rho(t)\Big\}}. \tag{2.38}$$

[1] $\Delta\hat{A} = \hat{A} - \Big\langle\hat{A}\Big\rangle$.

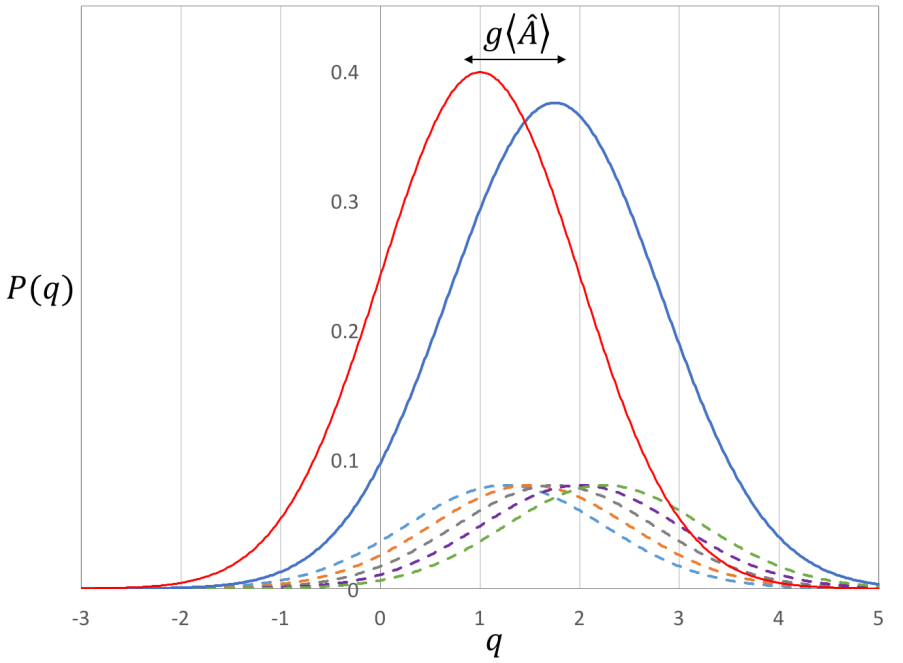


Figure 2.5: Weak measurement ($g = 0.25$) of an observable with 5 eigenvalues $a_i = i, i = 1...5$. The red curve shows the initial probability density function of the meter position. The curve in blue color is the final density function. As a consequence of the highly overlapped wave-packets (dotted curves), the final density is approximately the same initial Gaussian function, displaced by $g\left\langle \hat{A} \right\rangle$. One single outcome of the measurement gives a very small amount of information.

This expression does not correspond exactly to the projection postulate. However, every measurement should have in practise some classical noise, or uncertainty, ϵ. This means that the measurement device is found, not at some position q, but rather in an interval around that position. Let us denote this interval by $I_q = (q - \epsilon/2, q + \epsilon/2)$. If the noise is larger than the width of each wave packet, $\epsilon \gg \sigma_q$, but is much smaller than separation between the different wave packets, $\epsilon \ll g\Delta_a$, then the interval I_q may contain only one of the positions $Q_0 + ga_k$, or none. If I_q does not contain any of them, by integrating the density (2.33) in I_q, it is easy to see that $P(I_q) \approx 0$, i.e. with (almost) complete certainty the reading of the meter will contain one of the positions $Q_0 + ga_k$ among the noise ϵ. Let i be the union of all the intervals that contain the position $Q_0 + ga_i$. From (2.33), the probability to find the measurement device in this interval, i.e. to read one of the

eigenvalues of the operator $\hat{A}$, is given by

$$P(i) = \mathrm{Tr}\Big[\hat{\Pi}_i \rho(t)\Big], \tag{2.39}$$

and the conditioned system state is given by

$$\rho_i(t+\tau) = \frac{\hat{\Pi}_i \rho(t) \hat{\Pi}_i}{\mathrm{Tr}\Big[\hat{\Pi}_i \rho(t)\Big]}. \tag{2.40}$$

Thus, the projection postulate is recovered. The important point of the argument is that the projection postulate is recovered when the measurement is strong. Therefore, in this work, a strong measurement will have the same meaning as a projective measurement.

In this case one single measurement of the meter position tells the corresponding eigenvalue and that the system has "collapsed" into the state $\propto \hat{\Pi}_i \rho(t) \hat{\Pi}_i$. This state could be correlated with the initial state. However, if there was no degeneracy, then there is no correlation between the initial and final system states (because the final state is just $\hat{\Pi}_i$). Therefore, for a strong measurement, a large amount of information is gained with just one reading, and the initial state of the system is highly disturbed. Indeed, the unconditioned state of the system is

$$\rho(t+\tau) = \sum_i P(i) \hat{\Pi}_i \rho(t) \hat{\Pi}_i, \tag{2.41}$$

which is always a mixture, unless the initial state of the system is one of the eigenstates of the measured variable.

The opposite occurs when the measurement becomes weaker, i.e. the system is weakly perturbed and little information is obtained with one single measurement. In order to appreciate this, and assuming that wave function of the meter, $\psi(q)$, is a complex analytic function, it is possible to Taylor-expand the measurement and effect operators in powers of the coupling constant g, getting

$$\hat{M}_q = \sum_n \frac{(-g\hat{A})^n}{n!} \partial^n \psi(q) \quad , \quad \hat{E}_q = \sum_n \frac{(-g\hat{A})^n}{n!} \partial^n |\psi(q)|^2. \tag{2.42}$$

The expression $\partial^n f(q)$ denotes the *n-th* partial derivative of $f(q)$ with respect to q, and $\partial^0 f(q) = f(q)$. Using the expansion of the effect operators it is possible to express the probability density, $P(q) = \mathrm{Tr}\Big[\hat{E}_q \rho(t)\Big]$, in powers of g. By keeping the terms up to first order,

$$P(q) = \sum_n \frac{(-g)^n}{n!} \partial^n |\psi(q)|^2 \left\langle \hat{A}^n \right\rangle \approx |\psi(q)|^2 - g\partial |\psi(q)|^2 \left\langle \hat{A} \right\rangle \approx |\psi(q - g\left\langle \hat{A} \right\rangle)|^2. \quad (2.43)$$

The conditions under which the two approximations are valid depend on the initial wave function of the apparatus, $\psi(q)$, on the value of the coupling constant, g, and on the different moments of the measured variable, $\left\langle \hat{A}^n \right\rangle$[2]. This result shows that, at first orden in g, it is posible to gather information (provided that $\left\langle \hat{A} \right\rangle \neq 0$). With one single measurement, the extraction of information will be small, since the probability density will be noisy or wide, as it is shown in figure 2.5.

Next, in order to appreciate the disturbance on the system, it is possible to use the expansion of the measurement operators, and calculate the unconditioned state of the system after the measurement. Using (2.5), the state after the measurement is

$$\rho(t+\tau) = \int dq \hat{M}_q \rho(t) \hat{M}_q^\dagger \approx \rho(t) + \frac{ig}{\hbar} \left\langle \hat{P} \right\rangle [\rho(t), \hat{A}] + \mathcal{O}(g^2). \quad (2.44)$$

This expression shows that the linear contribution to the final state of the system depends on the initial expectation value of the momentum of the measurement device, $\left\langle \hat{P} \right\rangle = \left\langle \psi^A \right| \hat{P} \left| \psi^A \right\rangle$. Thus, by preparing the apparatus in a state with zero mean momentum, and assuming that the system initial expectation value of $\hat{A}$ is not zero, then *information can be extracted without perturbing the system*, for a sufficiently small value of g, i.e. for a sufficiently weak measurement of $\hat{A}$. The same occurs if the measured variable commutes with the initial state of the system, i.e. when $[\rho(t), \hat{A}]$ vanishes.

Let us consider now the case when the apparatus is not observed in the pointer basis. Moreover, let us assume that it is observed in the momentum basis. Recall that $\hat{P}$ is

[2]Recall that $\left\langle \hat{B} \right\rangle = \mathrm{Tr}\Big[\hat{B}\rho(t)\Big]$, where $\hat{B}$ is any system operator and $\rho(t)$ is the initial state of the system.

conjugate to the pointer variable $\hat{Q}$. In this case, the measurement operators are

$$\hat{M}_p = \psi(p) \exp\Big[-(i/\hbar) g p \hat{A}\Big] \quad , \quad \hat{E}_p = |\psi(p)|^2 \cdot \mathbb{1}. \tag{2.45}$$

Notice that the effect operators are proportional to the identity. Therefore, the probability density function is $P(q) = \mathrm{Tr}\Big[\hat{E}_p \rho(t)\Big] = |\psi(p)|^2$, namely, it remains unchanged after the measurement. Therefore, it is clear that no information is gained by observing the apparatus in the conjugate basis to $\hat{Q}$. On the other hand, by using the measurement operators, the conditioned system state is expressed as

$$\rho_p(t+\tau) = \frac{\hat{M}_p \rho(t) \hat{M}_p^\dagger}{P(p)} = \sum_{k,k'} e^{(-igp/\hbar)(a_k - a_{k'})} \hat{\Pi}_k \rho(t) \hat{\Pi}_{k'}, \tag{2.46}$$

which shows that the meter adds phase factors to the off-diagonal components of the system density operator $\rho(t)$.

Finally, it is interesting to generalize the conclusions of this section and assume that the apparatus is observed in the basis of some continuous variable $\hat{R} = \int r \hat{\Pi}_r dr$, where $\hat{\Pi}_r = |r\rangle\langle r|$ are rank one orthogonal projectors. Furthermore, let us make the assumption that this operator is hermitian and thus $r \in \mathbb{R}$. The measurement and effect operators can be expanded in powers of the coupling constant, as follows:

$$\hat{M}_r = \langle r| \exp\Big[-(ig/\hbar)\hat{A}\hat{P}\Big] |\psi\rangle = \sum_{n=0}^{\infty} \frac{(-ig\hat{A}/\hbar)^n}{n!} \langle r| \hat{P}^n |\psi\rangle, \tag{2.47}$$

$$\hat{E}_r = \hat{M}_r^\dagger \hat{M}_r = \sum_{n=0}^{\infty} \frac{(ig\hat{A}/\hbar)^n}{n!} \sum_{m=0}^{n} (-1)^m \binom{n}{m} \langle\psi| \hat{P}^{n-m} |r\rangle\langle r| \hat{P}^m |\psi\rangle. \tag{2.48}$$

The probability density function is given by $P(r) = \mathrm{Tr}\Big[\hat{E}_r \rho(t)\Big]$. The contribution of the linear term is proportional to the initial expectation value of $\hat{A}$,

$$P(r) = |\psi(r)|^2 + \frac{ig}{\hbar} \Big\langle \hat{A} \Big\rangle \Big\langle [\hat{P}, \hat{\Pi}_r] \Big\rangle + \mathcal{O}(g^2), \tag{2.49}$$

provided that the projector does not conmute with the momentum of the meter, $[\hat{P}, \hat{\Pi}_r] \neq 0$. This means that, for a sufficiently weak measurement, the expectation value of $\hat{R}$ will change by an amount proportional to $\left\langle \hat{A} \right\rangle$,

$$\left\langle \hat{R} \right\rangle_f = \int dr P(r) r = \left\langle \hat{R} \right\rangle + \frac{ig}{\hbar} \left\langle \hat{A} \right\rangle \left\langle [\hat{P}, \hat{R}] \right\rangle + \mathcal{O}(g^2). \tag{2.50}$$

This expression shows that the ensemble average of $\hat{R}$ will change as long as the momentum of the meter does not commute with the observed variable (the commutator should be computed over the initial system state).

On the other hand, the non-selective evolution of the density matrix is obtained by applying (2.5) to the initial system state. Using (2.47) the unconditioned state after the measurement may be expanded in powers of g as follows:

$$\rho(t+\tau) = \int dr \hat{M}_r \rho(t) \hat{M}_r^\dagger = \sum_{n=0}^{\infty} \frac{(ig/\hbar)^n}{n!} \left\langle \hat{P}^n \right\rangle \sum_{m=0}^{n} (-1)^m \binom{n}{m} \hat{A}^m \rho(t) \hat{A}^{n-m}. \tag{2.51}$$

Notice that the selective evolution (2.51) is independent of $\hat{R}$. To linear order in g, the system will remain unchanged when the initial mean momentum of the meter is zero. Indeed,

$$\rho(t+\tau) = \rho(t) + \frac{ig}{\hbar} \left\langle \hat{P} \right\rangle [\rho(t), \hat{A}] + \mathcal{O}(g^2). \tag{2.52}$$

Summarizing, equations (2.49) and (2.52) show that, in a weak measurement of an observable, information may be gained without perturbing the system, independent of the observation basis. The state of the system will be not perturbed when $\left\langle \hat{P} \right\rangle = 0$ or when the initial state of the system commutes with the measured variable. Information will be gained when $\left\langle \hat{A} \right\rangle \neq 0$ and $\left\langle [\hat{P}, \hat{R}] \right\rangle \neq 0$.

Chapter 3

Pre- and post-selected quantum measurements

In this chapter the theory of measurements performed on pre- and post-selected quantum systems is described. We begin with a brief description of the general theory, in which the initial and the final states are mixed states and the intermediate measurement may be of any kind. Then, we will focus on the particular case of a pre- and post-selected quantum system for which both, the initial and the final states, are pure and the intermediate measurement is projective. It worth to emphazise that section 3.1 is just an introduction to the subject. By no means it pretends to be an exhaustive review on generalized measurements on pre- and post-selected systems. However, all the elements presented in the first section will be useful to introduce the weak value in section 3.2. In section 3.3, the most important observations regarding the meaning of weak values are presented. In the last section we conclude the chapter with a summary of some of the most "well-known" experimental applications of weak values to the date.

3.1 General formulation of pre- and post-selected quantum measurements

The previous chapter described the basic measurement theory of a pre-selected quantum system, i.e. of a system that starts in an initial state $\rho(t)$. In this section, we will study the statistics of measurements performed on a quantum system whose *initial* and *final* states are partially or fully known. We will assume that the system starts in a state ρ_i at a time t_i and ends in a state ρ_f at a time $t_f > t_i$.

An *intermediate* measurement is performed between times t_i and t_f. We will ask for the probability distribution of the outcomes of this intermediate measurement, *given* that the initial and final states are (partially or fully) known. Recall that the measurement process has been divided into two stages; the first is the pre-measurement process and the second is the observation of the meter. The pre-measurement process begins at a time t and lasts for a time τ. The observation of the meter is an instantaneous step performed at a time $t_O > t + \tau$. The measurement is said to be *intermediate* because the pre-measurement process is performed between the times t_i and t_f, i.e. the interval $(t, t + \tau)$ is contained in the interval (t_i, t_f). The observation of the meter can be made indistinctly before or after the time t_f. For simplicity, in general, we will just say that the intermediate is performed at time t. However, the reader should keep in mind that this actually means that the pre-measurement process begins that time $t > t_i$ and ends at time $t + \tau < t_f$.

As has been explained, the quantum system starts in a state ρ_i at time t_i. Formally, we will say that the state ρ_i is *pre-selected*. Later, at a time t, the intermediate measurement is performed. Finally, at time t_f, a *second measurement* is performed, but only the results that produce the state ρ_f are taken into consideration. This "second measurement" is called *post-selection*, since it selects the final state ρ_f. Again, the second measurement will also have a pre-measurement stage and an observation stage, i.e. it has a certain duration. However, we will treat the second measurement as an "instantaneous" event, that occurs at time t_f. Otherwise, only irrelevant details will be introduced.

The post-selection is a probabilistic process, because there is a chance that it might fail. This failure occurs anytime the result obtained in the second measurement does not produce

the desired state ρ_f. When $t_O < t_f$, the results of the intermediate measurement are known before the post-selection. In this situation, the results of the intermediate measurement are taken into account *only when the post-selection is successful*. When it fails, the intermediate results are simply ignored and therefore not taken into consideration for the statistics of the measurement. When $t_O > t_f$ the intermediate results are known after the post-selection. In this case, the results of the intermediate measurement are acquired only when the (previous) post-selection process has been successful. Otherwise, when the post-selection has failed, it does has no sense to observe the results of the intermediate measurement.

The collection of identically pre- and post-selected quantum systems is called a *pre- and post-selected ensemble*. Similarly, the intermediate measurement performed on a pre- and post-selected ensemble is called a *pre- and post-selected measurement*.

As a summary: the system starts in a state ρ_i at the initial time t_i, an intermediate measurement is performed at time $t > t_i$, post-selection occurs at time t_f and the observation of the meter is realised at time t_O. We will assume that the system does not evolve in between those times. However, in a more general formulation, unitary or markovian evolution might be included [13]. Figure 3.1 shows a general representation of a pre- and post-selected measurement.

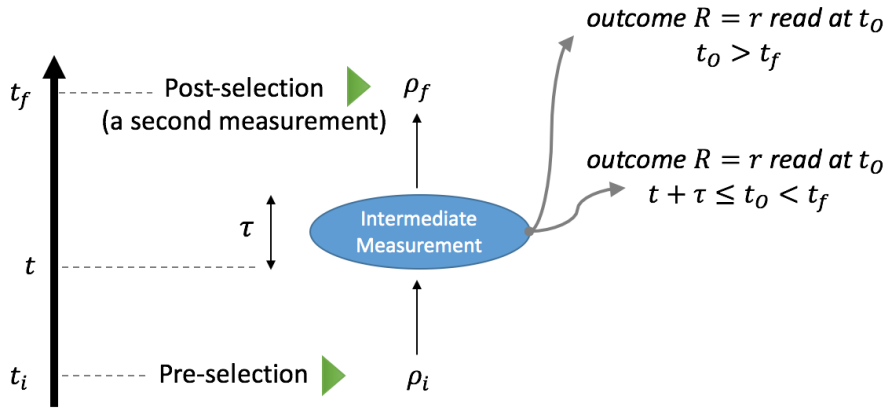


Figure 3.1: Representation of an pre- and post-selected measurement. An intermediate measurement is performed between an initial state ρ_i and a final state ρ_f. The outcome of the intermediate measurement may be read before or after the post-selection, which is a second measurement from which only the results that leave the system in the state ρ_f are kept.

Let us derive now an expression for the probability distribution of the outcomes of the

intermediate measurement, conditioning on the successful post-selection of the state ρ_f. If the outcome of the intermediate measurement is equal to r, then the conditioned state of the system will be $\rho_r = \mathcal{O}_r\rho_i/P(r)$, where $\mathcal{O}_r$ is the quantum operation associated with the outcome r and $P(r) = \mathrm{Tr}\{\mathcal{O}_r\rho_i\}$ is the probability of getting this result.

Then, a second measurement is performed at time t_f. As has been anticipated, no evolution is assumed to take place between the intermediate and the final measurement. Let $\hat{E}_f$ be the effect operator associated to the outcome f, which produces the final state ρ_f. This effect operator will be, in general, a function of the outcome of the intermediate result.

Consequently, the probability of getting the outcome f, *given* that the intermediate measurement produced the result r corresponds to

$$P(f|r) = \mathrm{Tr}\Big[\hat{E}_f\rho_r\Big]. \tag{3.1}$$

Employing *Bayes* theorem, the probability of obtaining the result r in the intermediate measurement, given that the final state is known to be ρ_f, may be expressed as

$$P(r|f) = \frac{P(f|r)\cdot P(r)}{P(f)} = \frac{\mathrm{Tr}\Big[\hat{E}_f\mathcal{O}_r\rho_i\Big]}{\mathrm{Tr}\Big[\hat{E}_f\mathcal{O}\rho_i\Big]}. \tag{3.2}$$

The term in the denominator, $P(f)$, represents the probability of successful post-selection. Notice that the quantum operation $\mathcal{O}$ corresponds to the non-selective evolution generated by the intermediate measurement $\mathcal{O} = \sum_r \mathcal{O}_r$. Expression (3.2) is a general expression that allows us to calculate the probability of the outcomes of a pre- and post-selected measurement.

Let us assume that we accept all the possible results of the final measurement .This would be equivalent of performing no post-selection. In this case, $\hat{E}_f = \mathbb{1}$ and therefore expression (3.2) becomes

$$P(r|f) = \frac{\mathrm{Tr}[\mathcal{O}_r\rho_i]}{\mathrm{Tr}[\mathcal{O}\rho_i]} = \mathrm{Tr}[\mathcal{O}_r\rho_i] = P(r), \tag{3.3}$$

namely, the probability distribution corresponds to the one of a standard quantum measurement. This type of ensemble is called an *only pre-selected ensemble*. One may consider the "opposite" the situation, in which the initial state is completely unknown, namely, $\rho_i = \mathbb{1}/D$ (D is the dimension of the Hilbert space of states). This situation corresponds to an *only post-selected ensemble*.

We will not study systematically a *mixed* pre- and post-selected ensemble, but consider a more restrictive scenario described by the following two conditions: i) the initial and final states are perfectly known, and ii) the intermediate measurement is efficient. The first condition entails that both states are pure, i.e. that the ensemble is a *pure* pre- and post-selected ensemble. Let $|\psi_i\rangle$ and $|\psi_f\rangle$ be those initial and final states, respectively. The fact that ρ_f is a pure state means that the final measurement is a von-Neumann type measurement and, consequently, the effect operator is a rank-one projector $\hat{E}_f = |\psi_f\rangle\langle\psi_f|$. The second condition implies that the action of the operation associated to the intermediate measurement may be written as $\mathcal{O}_r\rho_i = \hat{M}_r\rho_i\hat{M}_r^\dagger$, where $\hat{M}_r$ is the measurement operator associated to the r-th outcome. Thereby, expression (3.2) becomes

$$P(r|f) = \frac{|\langle\psi_f|\,\hat{M}_r\,|\psi_i\rangle|^2}{\sum_r |\langle\psi_f|\,\hat{M}_r\,|\psi_i\rangle|^2}. \tag{3.4}$$

Projective intermediate measurement

It is interesting to study expression (3.4) when the intermediate measurement is projective. Recall that every projective measurement is a minimally disturbing measurement of an observable. Let us denote this (possibly, degenerate) observable as $\hat{A}$, and let it spectral decomposition be $\hat{A} = \sum_r r \cdot \hat{\Pi}_r$. Since the intermediate measurement is projective each outcome is associated to one the eigenvalues of the observable. The measurement and effect operators are projectors, $\hat{M}_r = \hat{E}_r = \hat{\Pi}_r$. Therefore, expression (3.4) may be written as

$$P(r|f) = \frac{|\langle\psi_f|\,\hat{\Pi}_r\,|\psi_i\rangle|^2}{\sum_r |\langle\psi_f|\,\hat{\Pi}_r\,|\psi_i\rangle|^2}. \tag{3.5}$$

This formula allows to calculate the probabilities of each outcome of a projective measurement performed on a *pure* pre- and post-selected ensemble. Notice that it is symmetric under the exchange of the initial and final states.

One interesting application of this formula occurs in the so called *quantum box problem* [42, 43, 44, 45]. Consider a single particle that can be found in three boxes. Let $|i\rangle$ denote the quantum state corresponding to the spatial wave function of the particle localized in the i-th box ($i = 1, 2, 3$). These states are orthogonal and span the three dimensional Hilbert space of states. Assume that the experiment starts and ends with the particle in the pure states

$$|\psi_i\rangle = \frac{1}{\sqrt{3}}\Big(\,|0\rangle + |1\rangle + |2\rangle\,\Big) \quad \text{and} \quad |\psi_f\rangle = \frac{1}{\sqrt{3}}\Big(\,|0\rangle + |1\rangle - |2\rangle\,\Big). \tag{3.6}$$

Let us imagine that, between the initial and final times, we want to check whether the particle is in the k-th box or not. This can be regarded as an intermediate measurement of the observable $\hat{A} = |k\rangle\langle k|$ (a projector). Furthermore, in order to apply expression (3.5), assume that this measurement is projective (strong), namely, a measurement with two distinguishable outcomes, each one corresponding to one the two eigenvalues of the observable: $r = 1$ indicates that the particle is in the box, and $r = 0$ tells that the box is empty. In this situation, the POVM is given by the set $\{\hat{\Pi}_1, \hat{\Pi}_0\}$, where

$$\hat{\Pi}_1 = |k\rangle\langle k| \quad \text{and} \quad \hat{\Pi}_0 = \sum_{j\neq k} |j\rangle\langle j|\,. \tag{3.7}$$

Therefore, according to expression (3.5) if we check whether the particle is in box $k = 1$, then the probability distribution of the two possible outcomes is given by

$$P(1|f) = \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2}{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 1, \tag{3.8}$$

$$P(0|f) = \frac{|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2}{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 0. \tag{3.9}$$

Consequently, if box 1 is opened, but not the others, the particle will be found with certainty, which is a counterintuitive result. The same happens if we look for the particle

in box $k = 2$, but do not open the rest of the boxes. Indeed, in this case the probability distribution of the outcomes is

$$P(1|f) = \frac{|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2}{|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 1, \tag{3.10}$$

$$P(0|f) = \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2}{|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 0. \tag{3.11}$$

This case is represented in figure 3.2 (upper figure). Finally, if only box 3 was opened, the probability to find the particle is 1/5. Indeed,

$$P(1|f) = \frac{|\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2}{|\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2} = 1/5, \tag{3.12}$$

$$P(0|f) = \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2}{|\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\psi_i\rangle + \langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2} = 4/5. \tag{3.13}$$

Notice that the probabilities $P(1|f)$ and $P(0|f)$ depend on the whole measure $\{\hat{\Pi}_0, \hat{\Pi}_1\}$ and not only on the corresponding projector. This shows that the probabilities of a projective pre- and post-selected measurement are *context-dependent*. A discussion of a classical version the three box problem can be found in [46].

Let us consider now expression (3.4) when the intermediate measurement is a von Neumann type of measurement, i.e. a complete and minimally disturbing measurement of an observable $\hat{A}$. In this case, the effects and measurement operators are rank one projectors, $\hat{E}_r = \hat{M}_r = |a_r\rangle\langle a_r|$, where $\hat{A}\,|a_r\rangle = r\,|a_r\rangle$. Equation (3.4) becomes

$$P(r|f) = \frac{|\langle\psi_f|a_r\rangle|^2 \cdot |\langle a_r|\psi_i\rangle|^2}{\sum_r |\langle\psi_f|a_r\rangle|^2 \cdot |\langle a_r|\psi_i\rangle|^2}. \tag{3.14}$$

This expression is known as the Aharonov, Bergmann and Lebowitz (ABL) rule for calculating the probabilities of the outcomes of a von Neumann measurement performed on a pure pre- and post-selected ensemble [3]. Expression (3.5) can be regarded as an extension of the ABL rule for projective measurements.

Let us go back to the three-box problem and imagine that we open all boxes, instead of just one. This is an experiment with three possible outcomes. The effects and measurement operators are all rank one projectors, and the measure is given by the set

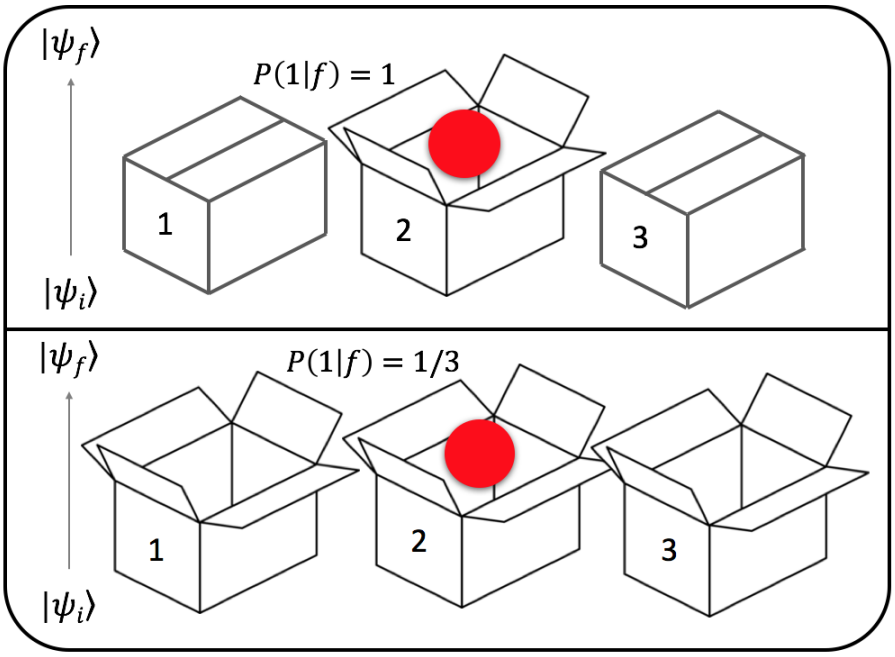


Figure 3.2: *Upper figure*: a strong measurement of the projector $|2\rangle\langle 2|$ is performed on a pre- and post-selected quantum system. The particle is always found in the box 2. *Lower figure*: a von Neumann measurement of the non-degenerate observable $\hat{A} = \sum i\,|i\rangle\langle i|$ on a pre- and post-selected system. The particle is found in box 2 with probability of 1/3.

$\{|1\rangle\langle 1|, |2\rangle\langle 2|, |3\rangle\langle 3|\}$. We can use the ABL rule (3.14) to obtain the probabilities of the three possible results,

$$P(1|f) = \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2}{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 + |\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 3|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 = 1/3$$

$$P(2|f) = \frac{|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2}{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 + |\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 3|\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 = 1/3$$

$$P(3|f) = \frac{|\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2}{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\psi_i\rangle|^2 + |\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2} = 3|\langle\psi_f|3\rangle\langle 3|\psi_i\rangle|^2 = 1/3$$

These results are now more intuitive, since the particle has equal probability of being found in each box. The expressions also show that the probabilities depend only on the corresponding projector, namely, von Neumann measurements performed on pure pre- and post-selected ensembles are *non-contextual*.

It is a well known fact that measurements of non commuting observables can not have definite values. For von Neumann measurements, performed on pure pre- and post-selected

ensembles, this is not the case. Assume that the initial state is $|\psi_i\rangle = |a\rangle$, where $\hat{A} = a\,|a\rangle$. Similarly, suppose that the final state is $|\psi_f\rangle = |b\rangle$ and $\hat{B} = b\,|b\rangle$. Thus, the system starts in a state with a well defined value of $\hat{A}$ and ends in a state in which the variable $\hat{B}$ is well defined. Furthermore, assume that the observables do not conmute, $[\hat{A}, \hat{B}] \neq 0$. If the intermediate measurement is a measurement of $\hat{A}$, according to (3.14), the probability of reading the eigenvalue a is equal to

$$P(a|f) = \frac{|\langle\psi_f|a\rangle|^2 \cdot |\langle a|\psi_i\rangle|^2}{\sum_{a'} |\langle\psi_f|a'\rangle|^2 \cdot |\langle a'|\psi_i\rangle|^2} = 1. \tag{3.15}$$

Analogously, if we measure $\hat{B}$, the odds to read the eigenvalue b are equal to $P(b|f) = 1$. Therefore, this pre- and post-selected system will have definite (dispersion-free) values $\hat{A}$ and $\hat{B}$, although these observables may not commute.

3.2 The weak value

Let us consider the formula (3.4), which allows to calculate the conditioned density function $P(r|f)$ of an efficient measurement. Assume that the meter is observed in the basis of some continuous variable $\hat{R} = \int r \cdot \hat{\Pi}_r dr$, where $\hat{\Pi}_r = |r\rangle\langle r|$. In this case, the conditioned probability density becomes

$$P(r|f) = \frac{|\langle\psi_f|\,\hat{M}_r\,|\psi_i\rangle|^2}{\int dr |\langle\psi_f|\,\hat{M}_r\,|\psi_i\rangle|^2}. \tag{3.16}$$

The operators $\hat{M}_r$ are the measurement operators associated to the outcome r, whose expansion in powers of the coupling constant is given in (2.47) for a von Neumann-type interaction. The term in the numerator of (3.16) is the *joint probability* of reading the outcome r *and* making a successful post-selection of the state $|\psi_f\rangle$. It may be written as

$$P(r, f) = |\langle\psi_f|\,\hat{M}_r\,|\psi_i\rangle|^2 = |\langle\psi_f|\psi_i\rangle|^2 \cdot |\psi(r)|^2 \cdot \sum_{n=0}^{\infty} g^n c_n(r). \tag{3.17}$$

Notice that the series is presented as a correction of the joint probability when no measurement takes place, in which case the joint probability is simply the multiplication of the

two independent probabilities. The coefficients of the series are given by

$$c_n(r) = \frac{(i/\hbar)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \frac{\langle \psi_f | \hat{A}^k | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \cdot \frac{\langle \psi_i | \hat{A}^{n-k} | \psi_f \rangle}{\langle \psi_i | \psi_f \rangle} \cdot \frac{\langle \psi | \hat{P}^{n-k} | r \rangle}{\langle \psi | r \rangle} \cdot \frac{\langle r | \hat{P}^k | \psi \rangle}{\langle r | \psi \rangle} \quad (3.18)$$

The different terms appearing in the sum are *weak values* of operators of the system and the apparatus, of different orders. The system starts in a pure state $|\psi_i\rangle$ and ends in a pure state $|\psi_f\rangle$. The n-th order weak value of the system variable $\hat{A}$, between those states, is defined as

$$A_w^n \equiv \frac{\langle \psi_f | \hat{A}^n | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle}. \quad (3.19)$$

The weak value is a complex number and diverges when the initial and final states are orthogonal. Unlike the expectation value of $\hat{A}$, the weak value is not restricted to lie in the range of eigenvalues, but can be outside this range when the overlap between the initial and final states is small.

From (3.18) it is clear that weak values of different orders completely characterize the correction to the joint detection probability $P(r, f)$ [18, 19, 20, 17]. The *weak measurement regime* corresponds to the linear regime, namely, when it is possible to keep the terms up to first order. Most of the experiments involving weak values has been done in this linear regime.

Analogously, the meter starts in a pure state $|\psi\rangle$ and ends in the eigenstate $|r\rangle$ when the outcome r is read. Therefore, the weak value of the momentum operator, between those states, corresponds to

$$P_w \equiv \frac{\langle r | \hat{P} | \psi \rangle}{\langle r | \psi \rangle}. \quad (3.20)$$

By integrating the joint density (3.17) over all the different possible readings of the meter, the probability of post-selection can be obtained. This is the term appearing in the

denominator of the conditioned density (3.16), which corresponds to

$$P(f) = \int dr P(r,f) = |\langle\psi_f|\psi_i\rangle|^2 \sum_n^\infty g^n C_n, \tag{3.21}$$

$$C_n = \left\langle \hat{P}^n \right\rangle \frac{(i/\hbar)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^k A_w^k \cdot \bar{A}_w^{n-k}. \tag{3.22}$$

The term $\bar{A}_w^{n-k}$ denotes the complex conjugate of A_w^{n-k}. The expansions of the joint density (3.17) and the post-selection probability (3.21) allow to expand the conditioned density (3.16). In the weak measurement regime, the conditioned probability is given by

$$P(r|f) \approx |\psi(r)|^2 - \frac{ig}{\hbar}\,\mathrm{Re}(A_w)\left\langle[\hat{\Pi}_r, \hat{P}]\right\rangle + \frac{g}{\hbar}\,\mathrm{Im}(A_w)\Big(\left\langle\{\hat{\Pi}_r, \hat{P}\}\right\rangle - 2\left\langle\hat{P}\right\rangle|\psi(r)|^2\Big) \tag{3.23}$$

This expression should be compared with (2.49), the probability density function in a pre-selected ensemble. Notice that the expectation value of $\hat{A}$ is now replaced with the real part of the weak value, $\mathrm{Re}(A_w)$. The third term, proportional to the imaginary part of the weak value, $\mathrm{Im}(A_w)$, does not have an equivalent term in (2.49). In this sense, a weak measurement performed on a pre- and post-selected system provides more information than a weak measurement made on a pre-selected system.

It is interesting to apply expression (3.16) to the case when the apparatus is observed in the pointer basis (the position). In this situation, the conditioned probability becomes

$$P(q|f) \approx |\psi(q)|^2 - g\,\mathrm{Re}(A_w)\partial|\psi(q)|^2 - \frac{2g}{\hbar}\,\mathrm{Im}(A_w)\Big\{\hbar\,\mathrm{Im}[\psi(q)\partial\psi^*(q)] + \left\langle\hat{P}\right\rangle|\psi(q)|^2\Big\} \tag{3.24}$$

Thus, if the initial wave function of the apparatus is real and the initial average momentum is zero, then the imaginary part of the weak value plays no role in the weak measurement regime. In this case or, equivalently, when the weak value is purely real, the probability density function is

$$P(q|f) \approx |\psi(q)|^2 - g\,\mathrm{Re}(A_w)\partial|\psi(q)|^2 \approx |\psi[q - g\,\mathrm{Re}(A_w)]|^2. \tag{3.25}$$

Again, this result resembles expression (2.43). For a (pre-selected) sufficiently weak mea-

surement the wave function of the meter is displaced by the expectation value of the measured observable. However, for a pre- and post-selected weak measurement the wave function is displaced by the real part of the weak value, which can be outside the range of the eigenvalues if the initial and final pure states are nearly orthogonal. This property has been used in precision metrology for parameter estimation.

On the other hand, using the general conditioned density (3.16), it is possible to compute the average value of the readings of the meter in the weak measurement regime,

$$\left\langle \hat{R} \right\rangle_f \approx \left\langle \hat{R} \right\rangle - \frac{ig}{\hbar} \operatorname{Re}(A_w) \left\langle [\hat{R}, \hat{P}] \right\rangle + \frac{g}{\hbar} \operatorname{Im}(A_w) \left(\left\langle \{\hat{R}, \hat{P}\} \right\rangle - 2 \left\langle \hat{P} \right\rangle \left\langle \hat{R} \right\rangle \right). \quad (3.26)$$

This expression was derived by Jozsa in [21]. It is instructive to apply this general formula to the cases when the apparatus is observed in the position and the momentum bases. The results are:

$$\langle P \rangle_f = \left\langle \hat{P} \right\rangle + \frac{2g}{\hbar} \operatorname{Im}(A_w) \left\langle \Delta \hat{P}^2 \right\rangle, \quad (3.27)$$

$$\langle Q \rangle_f = \left\langle \hat{Q} \right\rangle + g \operatorname{Re}(A_w) + \frac{g}{\hbar} \operatorname{Im}(A_w) \left(\left\langle \{\hat{Q}, \hat{P}\} \right\rangle - 2 \left\langle \hat{P} \right\rangle \left\langle \hat{R} \right\rangle \right). \quad (3.28)$$

Notice that the imaginary part of the weak value appears when the meter is observed in the momentum basis. This result is drastically different to the situation in which there is no post-selection. In that case, the observation of the meter's momentum provides no information, as was described in section 2.3.

3.3 Interpretation of weak values

3.3.1 Weak values as contextual values

From (3.26) it is clear that in a weak measurement the shift of the expectation value of the meter readings is not precisely equal to (1.1). This expression shows that this conditioned average depends on the system, i.e. on $|\psi_i\rangle$, and $|\psi_f\rangle$, on the measured observable

$\hat{A}$, but also on some features of the apparatus. If we make the assumption that $\left\langle \hat{R} \right\rangle = 0$, then expression (3.26) reduces to

$$\left\langle \hat{R} \right\rangle_f = g\left[\frac{2}{\hbar}\,\mathrm{Im}\left(A_w \left\langle \hat{R}\hat{P} \right\rangle\right)\right]. \tag{3.29}$$

This result clearly shows that the conditioned expectation of the meter readings in a weak measurement does not reflect a property of the quantum system, because it depends also on the expectation value of the product of $\hat{R}$ and $\hat{P}$, which are "features" of the measurement strategy. Therefore, the quantity $\left\langle \hat{R} \right\rangle_f / g$ can not be interpreted as the conditioned average of the observable $\hat{A}$. This argument is developed in [110], where the author defends the idea that weak values are not unique.

This issue was addressed in [109], where *contextual values* (CV) were introduced. CV of an observable $\hat{A} = \sum_k a_k \hat{\Pi}_k$ are a generalization of its eigenvalues. They are defined in terms of the observable but also in terms of the measurement strategy. Let $\{\hat{M}_r, r \in \mathcal{R}\}$ be the set of measurement operators, also called the *measurement context*. Recall that the set $\mathcal{R}$ represents the space of results, i.e. the set of all possible outcomes of the experiment. For simplicity, we will assume that there are N possible outcomes, $r = 1, ..., N$. This context generates the POVM $\{\hat{E}_r, r\}$, where $\hat{E}_r = \hat{M}_r^\dagger \hat{M}_r$. The set of N contextual values $\{\alpha_r\}$ is defined from the identity

$$\sum_{r=1}^{N} \alpha_r \hat{E}_r = \sum_{k=1}^{dim(S)} a_k \hat{\Pi}_k, \tag{3.30}$$

where $dim(S)$ corresponds to the dimension of the Hilbert space of the quantum system. This definition requires that $N \geqslant dim(S)$. If $N = dim(S)$ the solution of (3.30) is unique. If $N > dim(S)$, then a unique solution can be selected using the method described in [109]. Once the generalized eigenvalues are found, then the conditioned average of the observable $\hat{A}$ can be defined as

$$\left\langle \hat{A} \right\rangle_f \equiv \sum_r \alpha_r P(r|f). \tag{3.31}$$

The conditioned probability density is given by (3.2). In [109] the author shows that for a weak measurement, i.e. in the limit $g \to 0$, and when the measurement is minimally disturbing, the conditioned average (3.31) does not depend on the context and converges to a unique quantity. This quantity was called by the author a generalized weak value. It equals the real part of A_w when the initial and final states are pure. This argument provides an statistical definition of (the real part of) the weak value, as being the conditional expectation of contextual values in a weak and minimally disturbing measurement.

3.3.2 Weak values as a property of a quantum system

So far, the weak value has been presented as the conditioned average of the pointer readings in a weak measurement, namely, as a generalization of the standard expectation value when post-selection is included. From this perspective, the weak value is not a property of a single quantum system, but of a large collection of identical quantum systems. In this sense, the weak value is an statistical quantity. In the laboratory a large number of particles is actually needed in order to estimate a weak value as the average of the outcomes of weak measurements.

However, in [22] the authors argue that the weak value resembles more to an eigenvalue than to an expectation value. The ensemble is only needed because weak measurements extract little information in one single outcome. They considered a von Neumann model of measurement with strength g, and an apparatus initially prepared in a Gaussian state Φ_0 with spread σ_q. Regarding the system, three different scenarios were taken into account: i) the system starts in an eigenstate with eigenvalue equal to 1, ii) the system is prepared in a pure state $|\psi_i\rangle$ with post-selection of the state $|\psi_f\rangle$, in such a way that the weak value equals to 1, and iii) the system begins in a superposition of eigenstates, with expectation value equal to 1. In these three cases, the state of the apparatus after the measurement was computed and denoted as Φ_e, Φ_w and ρ_{ex}, respectively. Note that the final state of the meter in third case is a mixed state. Then, the Bures distance [131] between Φ_0 and the different final states of the apparatus was computed. The results for each case were

the following:

$$i)\ D(\Phi_0, \Phi_e) \sim \frac{g}{\sigma_q}, \quad ii)\ D(\Phi_e, \Phi_w) \sim \left(\frac{g}{\sigma_q}\right)^2, \quad iii)\ D(\Phi_e, \rho_{ex}) \sim \frac{g}{\sigma_q}. \tag{3.32}$$

The first result shows that the distance between the initial and final states of the meter is detectable at first order. The second result indicates that, at first order, there is no appreciable difference between the first and the second cases. Thus, preparing the system in an eigenstate with eigenvalue equal to 1 produces essentially the same effect over a measurement device as a pre- and post-selected quantum system with weak value equal to 1. Nevertheless, when the system starts in a superposition of eigenstates, the difference between the final states of the meter is appreciable. The conclusion of this analysis suggests that in a weak measurement procedure a quantum system with a certain weak value will affect a measurement device in the same way as a system prepared in an eigenstate with eigenvalue equal to the weak value. In this sense, the weak value can be rather understood as a "robust" property of a single quantum system than a purely statistical quantity.

3.3.3 Do weak values have classical analogs?

Ferrie and Combes [23] suggested that the measurement of anomalous weak values is not an inherently quantum phenomenon, but is rather an statistical feature that can be observed in a classical system, including pre- and post-selection, and a weak measurement with disturbance. In their article, published with the tittle: "*How the result of a single coin toss can turn out to be 100 heads*", the authors considered a weak measurement of $\hat{\sigma}_z$. The system (a qubit) starts in an initial state ρ_i. According to the weak measurement model employed by the authors, the (unnormalized) conditioned state of the system can be expressed as

$$\tilde{\rho}_s = \rho_i + \frac{s\lambda}{2}\Big[\rho_i\hat{\sigma}_z + \hat{\sigma}_z\rho_i\Big], \tag{3.33}$$

where $\tilde{\rho}_s$ is the unnormalized system state after the weak measurement has produced the result $s = \pm 1$, and λ is a parameter characterizing the weakness of the measurement. Next, Ferrie and Combes applied a "bit-flip channel" [15] to the state (3.33), in order to

describe the effect of a *classical disturbance*. This operation flips a qubit from $|0\rangle$ to $|1\rangle$ (and vice-verse) with probability p y leaves the qubit unchanged with probability $1-p$. In the article, the probability of "flipping" the qubit depends on the result of the weak measurement, namely, $p = p(s)$. Let $\mathcal{O}_{BF}(s)$ be the quantum operation describing the bit-flip channel. Notice that the operation depends on the outcome s. Finally, after the action of this operation, the state $|\psi_f\rangle$ is post-selected. In this way, given the initial state ρ_i, the joint probability of reading the outcome s in the weak intermediate measurement and post-selecting the desired pure state, is given by

$$P(f,s|i) = \langle\psi_f|\,\mathcal{O}_{BF}(s)\rho_s\,|\psi_f\rangle = [1-p(s)]\,\langle\psi_f|\,\tilde{\rho}_s\,|\psi_f\rangle + p(s)\,\langle\psi_f|\,\sigma_x\tilde{\rho}_s\sigma_x\,|\psi_f\rangle\,. \quad (3.34)$$

Hence, employing Bayes's Theorem, the conditional probability can be written as

$$P(s|f,i) = \frac{\langle\psi_f|\,\mathcal{O}_{BF}(s)\tilde{\rho}_s\,|\psi_f\rangle}{\sum_s \langle\psi_f|\,\mathcal{O}_{BF}(s)\tilde{\rho}_s\,|\psi_f\rangle}, \quad (3.35)$$

and, therefore, the expectation value of the weak intermediate measurement is given by

$$\mathbb{E}[s|f,i] = \frac{\langle\psi_f|\,\mathcal{O}_{BF}(1)\tilde{\rho}_1\,|\psi_f\rangle - \langle\psi_f|\,\mathcal{O}_{BF}(-1)\tilde{\rho}_{-1}\,|\psi_f\rangle}{\langle\psi_f|\,\mathcal{O}_{BF}(1)\tilde{\rho}_1\,|\psi_f\rangle + \langle\psi_f|\,\mathcal{O}_{BF}(-1)\tilde{\rho}_{-1}\,|\psi_f\rangle}. \quad (3.36)$$

In order to establish an analogy with the toss of a coin, the initial state was chosen to be pure, $\rho_i = |0\rangle\langle 0|$ (head), while the final state was $|\psi_f\rangle = |0\rangle$ (tail). These states are the eigenstates of $\hat{\sigma}_z$, namely, $\hat{\sigma}_z\,|0\rangle = |0\rangle$ and $\hat{\sigma}_z\,|1\rangle = -\,|1\rangle$. In this case, the general formula (3.36) reduces to a simply expression,

$$\mathbb{E}[s|f,i] = \frac{p(1)(1+\lambda) - p(-1)(1-\lambda)}{p(1)(1+\lambda) + p(-1)(1-\lambda)}. \quad (3.37)$$

Thus, the probabilities of flipping the coin may be chosen so that the expectation value equals an specific value. In [23] the probabilities were $p(1) = (1+\lambda-\delta)/(1+\lambda)$ and $p(-1) = (1-\lambda-\delta)/(1-\lambda)$. The disturbance parameter δ is constrained by $0 < \delta < 1-\lambda$, so that the probabilities are well defined. In this way, the expectation value of the weak

measurement is

$$\mathbb{E}[s|f,i] = \lambda\Big(\frac{1}{1-\delta}\Big). \tag{3.38}$$

The term in the parenthesis represents an amplification factor, that may be large. Indeed, the authors defined $\mathbb{E}[(s/\lambda)|f,i]$ as a generalized weak value. This argument was used to argue that the phenomenon of weak values could have an equivalence in a classical system. Obviously, the amplification factor obtained by Ferrie and Combes appears in a different setup: a weak measurement followed by a disturbance. In the original article by Aharonov, Albert and Vaidman [1] no disturbance is needed. Additionally, the amplification factor depends on the parameters associated to the disturbance, and on the initial and final states. Consequently, the nature of the amplification is essentially different. The article by Ferries was commented in [25] and their reply was presented in [24].

It is worth to mention that classical values, analog to weak values, can be measured in interference experiments of classical electromagnetic waves [26]. In this type of experiments, nevertheless, the apparatus and the system are different degrees of freedom of the same particle (a photon). For weak values, however, the apparatus and the system can be separate systems.

In [27] the authors used the Holevo quantity to distinguish between classical and quantum correlations, and showed that the phenomenon of weak value amplification is exclusively related to a small amount of quantum correlation. Other arguments defending the idea that weak values are a quantum phenomenon can be found in [28, 29, 30].

3.3.4 The three box problem revisited: weak values of projectors and negative probabilities.

Let us come back to the three box problem and consider weak measurements of the projector operators $\hat{\Pi}_k = |k\rangle\langle k|$, $k = 1, 2$ and 3. In this case, a single outcome of the measurement will not tell with enough clarity whether the particle is or not in the box we are looking for it. The average of such measurements, the weak value, will provide this information. The weak values of the projector operators for three box problem are the

following:

$$k = 1: \qquad \Pi_{1,w} = \frac{\langle\psi_i|1\rangle\langle 1|\psi_f\rangle}{\langle\psi_f|\psi_i\rangle} = 1, \tag{3.39}$$

$$k = 2: \qquad \Pi_{2,w} = \frac{\langle\psi_i|2\rangle\langle 2|\psi_f\rangle}{\langle\psi_f|\psi_i\rangle} = 1, \tag{3.40}$$

$$k = 3: \qquad \Pi_{3,w} = \frac{\langle\psi_i|3\rangle\langle 3|\psi_f\rangle}{\langle\psi_f|\psi_i\rangle} = -1. \tag{3.41}$$

The first two results show that, if we performed many weak measurements on identical pre- and post-selected particles, and took the average of all the outcomes, we would finally discover that the particle is actually in box 1 (or box 2). A theorem states that if the weak value of a dichotomous variable equals one of its eigenvalues, then the probability of measuring that eigenvalue in a strong measurement will be equal to 1 [42]. This can be verified by comparing the weak results (3.39) and (3.40) with the probabilities (3.8), (3.9), (3.10) and (3.11).

Imagine that the particle has positive charge and that the weak measurement of the projector is implemented by measuring the transverse deflection of a beam of electrons passing next to the corresponding box [50]. If the particle is in the box, then the beam will be deflected. According to the results (3.39) and (3.40), we would find that the average deflection of the beam was of one unit (in some scale) to the right, because particles of opposite charges attract each other, as it is shown in figure 3.3. However, the result for the weak value of $\hat{\Pi}_3$ is strange. In this case, the deflection is to the left, as if the particle in the box would have inverted the sign if its electrical charge. An optical experiment of the quantum box problem was made in [36].

Expectation values of projector operators can be thought as probabilities [35], because their values are between 0 and 1. Weak values of projectors, nevertheless, can take negative values, as its clear from the results above. The "negative probability" would then correspond to the probability for an "opposite" event to occur.

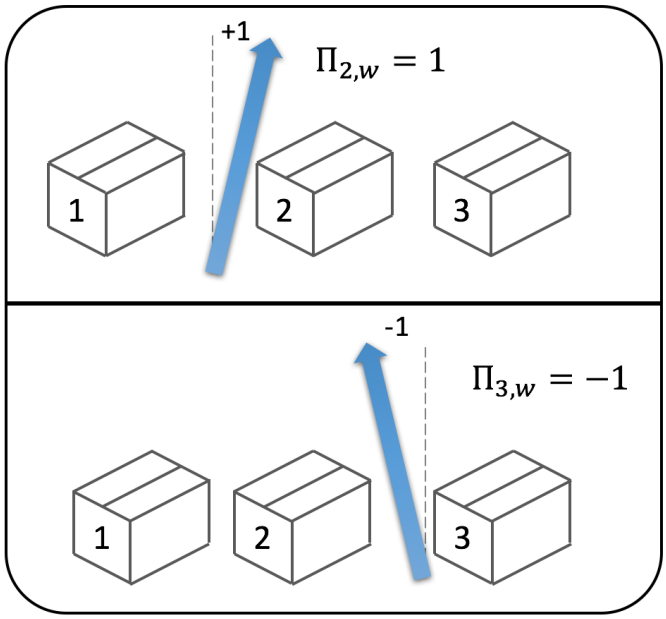


Figure 3.3: *Upper figure*: a weak measurement of the projector $|2\rangle\langle 2|$ is performed on a pre- and post-selected quantum system. The average of all the outcomes shows that the particle is in the box 2, since the average deflection is to the right. *Lower figure*: a weak measurement of the projector $|3\rangle\langle 3|$. The particle behaves as if it had negative charge, because the average deflection is to the left.

3.3.5 The quantum Cheshire cat

Alice in Wonderland replied to the Cheshire cat: "...and I wish you wouldn't keep appearing and vanishing so suddenly: you make one quite giddy!". Then, as the cat began slowly to disappear, while only its grin remained for some time after, Alice thought: "Well! I've often seen a cat without a grin... but a grin without a cat! It's the most curious thing I ever saw in all my life!" [11]. Pre- and post-selected quantum systems can affect measurement devices as if one of its properties was in a certain region of space, while some other property was located in another place. In this sense, they can behave as grins without cats. Consider a beam of photons going through the Mach-Zehnder interferometer shown in figure 3.4. The photon will play the role of the cat, while its polarization degree of freedom will be the grin.

Assume that a birefringent plate is placed in the arm A of the interferometer. The plate will spatially separate two orthogonal polarization states. The transverse (horizontal) position of the beam will be displaced to "the right" if the polarization is horizontal and

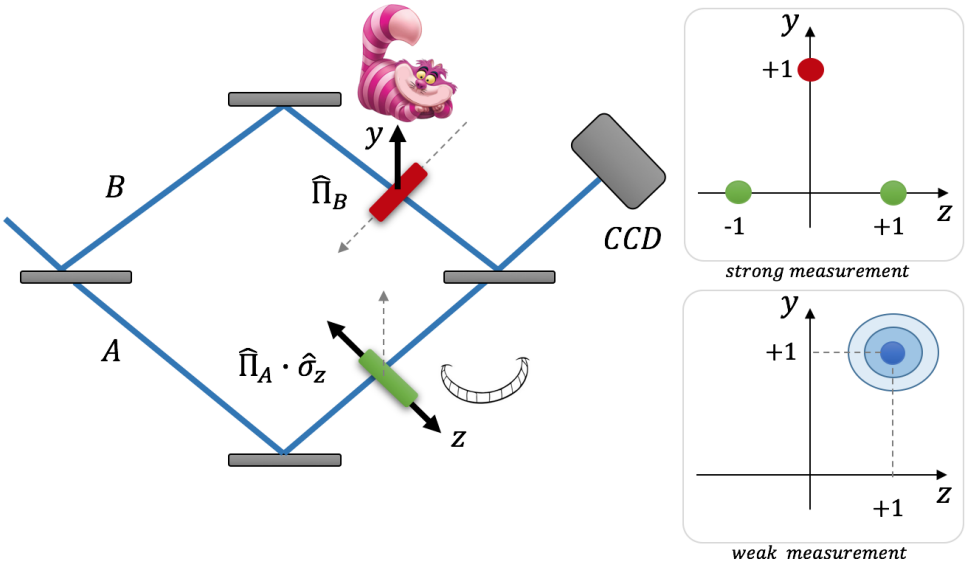


Figure 3.4: The grin of the cat travels through the arm A, while the cat goes through the arm B. The inset figures show the results of the intensity distribution for a strong (upper) and weak (lower) measurement.

to "the left" when light is vertically polarized. These polarization states will be denoted by $|H\rangle$ and $|V\rangle$, respectively, $\hat{\Pi}_H$ and $\hat{\Pi}_V$ being the corresponding projectors. Therefore, the plate will implement a measurement of the operator $\hat{\Pi}_A\hat{\sigma}_z$, where $\hat{\Pi}_A = |A\rangle\langle A|$ is the projector into the corresponding path state and $\hat{\sigma}_z = \hat{\Pi}_H - \hat{\Pi}_V$ is a spin-like polarization operator. The pointer variable will be the transverse (horizontal) position of the beam.

On the other hand, in path B a sheet glass is placed in order to deflect the beam "upwards". This element will implement therefore a measurement of the operator $\hat{\Pi}_B = |B\rangle\langle B|$, that projects into a state propagating across this path. In this case, the pointer variable will be the transverse (vertical) position of the beam.

In summary, the polarization in the arm A will be observed in the horizontal transverse position of the photon; a displacement to the left will indicate vertical polarization and to the right horizontal polarization. The vertical position is a which-path detector; when it is displaced upwards, we know that the photon travelled along path B. If it is not displaced, then it went trough the other arm. The intensity of the beam along the transverse plane will be scanned by a CCD camera placed at the exit of the interferometer. All the beamsplitters located at the entrance and the exit of the interferometer, together with the CCD camera,

are arranged in such a way that the states $|\psi_i\rangle$ and $|\psi_f\rangle$ are pre- and post-selected, where

$$|\psi_i\rangle = \frac{1}{\sqrt{2}}|+\rangle|A\rangle + \frac{1}{\sqrt{2}}|-\rangle|B\rangle \text{ and } |\psi_f\rangle = \frac{1}{\sqrt{2}}|-\rangle(|A\rangle + |B\rangle). \tag{3.42}$$

The states $|\pm\rangle$ are linear polarization states defined as $|\pm\rangle = (1/\sqrt{2})(|H\rangle \pm |V\rangle)$. Consider first the the case when both measurements are strong. The intensity profile will show three distinguishable peaks, as it is shown in figure (3.4). The upper peak, with coordinates $(y = +1, z = 0)$, will have a (normalized) intensity of 2/3, indicating that two of every three photons passed through the arm B. The coordinate $y = 1$ shows that the photon went through the arm B, while the coordinate $z = 0$ indicates that no polarization effect was detected in the arm A. The other two peaks will have each a normalized intensity of 1/6. This means that half of the photons moving along the arm A had horizontal polarization while the other half was vertically polarized.

However, when the measurements are weak, then only one "broad" peak will appear, located at $(y = +1, z = +1)$. The horizontal displacement corresponds to the weak value of $\hat{\Pi}_A\hat{\sigma}_z$, while the vertical coordinate is the weak value of $\hat{\Pi}_B$. This shows that, simultaneously, the polarization in arm A was horizontal, but the photon travelled along the other arm. Thus, the system behaves as if the polarization was spatially separated from the photon, in the sense of how it affects measurement devices in weak interactions. This is the "Cheshire cat effect" [47, 48]. It may allow to perform high-precision measurements over a spatial region, from which a certain unwanted property, that could cause disturbance, has been removed.

In this experiment the two degrees of freedom of the transverse position of the beam allowed to perform simultaneous measurements of two weak values, $(\Pi_A\sigma_z)_w$ and $(\Pi_B)_w$. In a more general setup, the following four weak values could all have been measured at the same time:

$$(\Pi_A\Pi_H)_w = 1/2\ ,\ (\Pi_A\Pi_V)_w = -1/2\ ,\ (\Pi_B\Pi_H)_w = 1/2\ ,\ (\Pi_B\Pi_V)_w = 1/2. \tag{3.43}$$

Where was the photon? The result $(\Pi_B)_w = (\Pi_B\Pi_H)_w + (\Pi_B\Pi_V)_w = 1$ indicates that it is was in arm 1. By virtue of the theorem mentioned above this means that the

probability of finding the photon in arm B using a strong measurement equals to one. So, "the cat" (the photon) was in arm B with certainty. *Did the photon had vertical or horizontal polarization?* Note that $(\sigma_z)_w = 1$, which indicates that the polarization was horizontal. Hence, and with certainty, the "cat was grinning".

May we conclude that the cat was grinning in the arm B of the interferometer? In other words, is it true that $(\Pi_B \sigma_z)_w = 1$? For standard expectation values the answer would be affirmative, because, when two commuting variables have simultaneously definite values, the expectation value of the product is the product of the expectation values. However, this is not the case for weak values or for average values of pre- and post-selected measurements. In particular, the weak value of the product of two operators is not the product of the weak values. In fact, it is easy to see that $(\Pi_B \sigma_z)_w = 0$, while $(\Pi_A \sigma_z)_w = (\Pi_A \Pi_H)_w - (\Pi_A \Pi_V)_w = 1$. The result $(\Pi_A \Pi_H)_w = 1/2$ shows that half of the photons that went through the arm A had horizontal polarization (displacing the position one unit to the right), while the other half had vertical polarization (not producing any displacement). The negative result $(\Pi_A \Pi_V)_w = -1/2$ is strange. One may interpret that half the photons were horizontally polarized, not producing any displacement at all. The other half were vertically polarized, but displaced the beam as if they were horizontally polarized (to the other direction). This analysis shows that a pre- and post-selected photon going through the arm A with vertical polarization can behave as a photon with horizontal polarization. Therefore, in a similar way to the three box problem, the negative weak value of a projector can be interpreted as if the photon has changed its polarization from vertical to horizontal.

3.3.6 Hardy's paradox

Hardy's paradox is a gedanken experiment described in [51] and is a demonstration of Bell's theorem [54] without the use of inequalities. It also presents a contradiction between quantum mechanics and a realistic theory, in the sense of Einstein, Podolsky and Rosen [55]. The gedanken experiment is based on the interaction-free measurement proposal of Elitzur and Vaidman [52]. In the latter, a photon is detected at the output ports of a Mach-Zehnder interferometer. The phases are set in such a way that the photon is always

detected at one port, which is called the *bright port*. The other port, called the *dark port*, will never click. Then, a "bomb" is inserted in one of the arms of the interferometer. If the photon goes through this arm, the bomb explodes with unit probability. In this situation, quantum theory predicts that there is a chance that the dark port clicks. Therefore, a single detection at this port will reveal the presence of the bomb, although the photon necessarily passed through the arm without the bomb or else an explosion would have occurred. The presence of the bomb was detected without having to make it explode!

Hardy's experiment consists of two Mach-Zehnder interferometers, one for a positron and the other for an electron, as it is shown in figure 3.5. If the interferometers are considered independently, then each particle will be detected at the port $C^{\pm}$ of the corresponding interferometer, where the plus sign labels the interferometer for the positrons and the minus signs labels the interferometer for the electrons. These ports are therefore the bright ports. However, if the interferometers overlap in one of its arms, then there is a probability that the particles are detected at the ports labeled as $D^{\pm}$. There is also a chance that the positron and the electron take both the overlapping path of each interferometer and meet at point P. In this case, they annihilate each other with unit probability.

When the detectors D^{+} and D^{-} click in coincidence the interference effect due to the overlap is revealed, in the same sense as the bomb in the free-interaction measurement was detected. In this situation, it is also known that the positron and the electron did not take to overlapping path simultaneously.

Conditioning on the simultaneous clicking of these detectors, the projectors on each path state (or occupation numbers) can be measured. Let $\widehat{O}_{+}$ and $\widehat{NO}_{+}$ be the projectors into the overlapping and non-overlapping paths for the positron, respectively, while $\widehat{O}_{-}$ and $\widehat{NO}_{-}$ are the equivalent projectors for the electron. A strong measurement of $\widehat{O}_{+}$ on the pre- and post-selected particles reveals that the positron travelled through this path with unity probability. This conclusion will be expressed as

$$\widehat{O}_{+} = 1. \tag{3.44}$$

Analogously, a strong measurement of $\widehat{O}_{-}$ indicates that the electron went through the

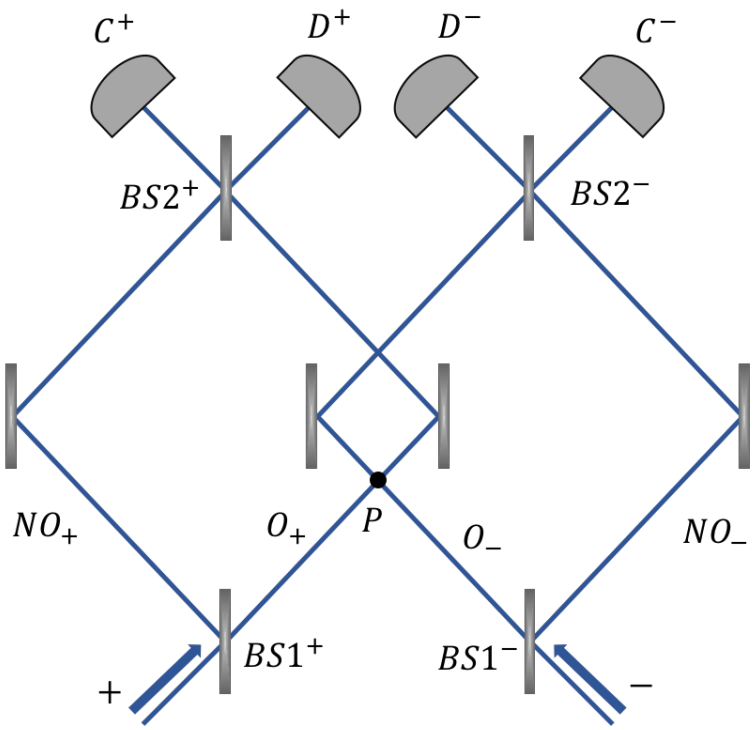


Figure 3.5: Two Mach-Zehnder interferometers. The plus signs labels the interferometer for the positrons and the minus sign labels the interferometer for the electrons. The annihilation occurs if the two particles take simultaneously the overlapping arms $O_{\pm}$ and meet at point P.

overlapping path. Thus, $\hat{O}_- = 1$. However, as was pointed out in the previous section, for a pre- and post-selected measurement, the average value of the product of two variables, each with a well-defined value, does not equal to the product of the individual values. Hence, if the trajectories of the positron and the electron are measured *jointly*, a strong measurement reveals that the probability that both particles have gone through the overlapping path is zero,

$$\hat{O}_+ \cdot \hat{O}_- = 0. \tag{3.45}$$

Thus, if the positron is searched in the overlapping path, it will be found with certainty, provided that the electron is not observed. The same happens with the electron. Nevertheless, if both particles are looked in the overlapping arms, the probability to find that both went along these arms is zero. This is an expected result, otherwise, an annihilation would have been produced and no detection could have been recorded.

On the other hand, strong measurements of the other three possible trajectories of the positron-electron pair also reveal reasonable results. In fact,

$$\widehat{NO}_+ \cdot \widehat{O}_- = \widehat{O}_+ \cdot \widehat{NO}_- = \widehat{NO}_+ \cdot \widehat{NO}_- = 1. \tag{3.46}$$

Hence, except for the "annihilation trajectory" (3.45), every strong measurement of the trajectories followed by the pair of particles, will reveal with certainty that the particles have travelled throughout the examined trajectory. Of course, these four measurement of joint operators can not be performed at the same time, because they disturb the state of the system. However, they can be made "in the same run" using weak measurements. In this case, the weak values are

$$(O_+O_-)_w = 0 \ , \ (NO_+O_-)_w = (O_+NO_-)_w = 1 \ , \text{ and } (NO_+NO_-)_w = -1. \tag{3.47}$$

These results show that the system behaves as if two pairs of particles went through the combination of an overlapping path with a non-overlapping one, while a "minus" electron-positron pair went through the non-overlapping paths of the interferometer. The effect of the "minus" pair of particle over a measurement device will be the opposite to the effect caused by a positron-electron pair. Also, notice that all the weak values add up to one.

An analysis of Hardy's paradox in terms of weak values can be found in [53]. The experiment was implemented using photons in [39, 41]. In these experiments the joint weak values were extracted from the correlation of the readings of two measurement devices [37].

3.3.7 Two slit experiments

Weak measurements of projectors have been useful to study the wave-particle duality in a two-slit experiment. In [58] projectors into momentum eigenstates, $|p_i\rangle$, were weakly measured on a pre- and post-selected ensemble of photons. The initial state of the photons was the two slit wave-function (a coherent superposition of the photon diffracted from both slits), denoted as $|\psi\rangle$, while the final state was a momentum eigenstate $|p_f\rangle$. For this experiment, the weak value corresponds to

$$(|p_i\rangle\langle p_i|)_w = \frac{\langle p_f|\,\widehat{U}\,|p_i\rangle\langle p_i|\psi\rangle}{\langle p_f|\,\widehat{U}\,|\psi\rangle}. \tag{3.48}$$

The unitary operator $\widehat{U}$ represents the system evolution that occurs after the weak measurement. In the experiment, the operator describes the *which way measurement* process that follows the weak measurement of momentum. Notice that if no which way information is gathered, i.e. $\widehat{U} = \mathbb{1}$, then the weak value equals 1 (if $p_f = p_i$) or zero (if $p_f \neq p_i$). Expression (3.48) corresponds to a weak value of a projector and can exceed the range $[0, 1]$. In order to distinguish it from a true probability, in [58] it is called a *weak valued probability*, that may be understood as the probability of a photon of having initially momentum equal to p_i, *given* that its final momentum was p_f. Certainly, it is not an actual probability, but a quasi-probability distribution (such as the Wigner function). It allows to compute the probability density of the momentum transfer in a two-slit experiment. The density function has zero variance and is therefore compatible with the claim of Scully [59] that interference may be destroyed without momentum transfer. However, the density function has a large width, which is also compatible with the claim of Storey [60] that any which way measurement will transfer momentum in accordance to the Heisenberg uncertainty principle.

In [61] weak measurements of the transverse momentum, along a sequence of planes, with post-selection of photons arriving at a certain position of a CCD camera, allowed the reconstruction of the trajectories followed by photons in a two-slit interferometer. In this experiment the weak measurement of the momentum was done using the polarization degree of freedom of the photons as the meter. The final post-selection on a certain position was performed with a CCD device.

3.4 Experimental applications of weak values

3.4.1 First experimental measurement of a weak value

In the same article in which the concept of the weak value was introduced [1], the authors proposed an experimental setup to measure an anomalously large weak value of a spin 1/2 component, using a sequence of three Stern-Gerlach magnets. Some technical corrections to the approximations presented in this article were made by Sudarshan in [31], where also an optical experiment was proposed in order to measure a large weak value of a polarization operator, which played the role of the original spin component. This experiment was realized by Ritchie and collaborators [32], and was the first experimental measurement of a weak value.

This experiment can be described using classical electromagnetic theory, because it used classical light. However, the formalism of quantum mechanics can be employed to describe a classical light beam in the paraxial approximation [33, 34]. Here, we will shortly describe the experiment using the formalism of quantum mechanics. In this way it is identical to the original proposal [1].

The output of a laser beam in a Gaussian TE mode, propagating in the z direction, passes through a first polarizer, which is oriented at an angle α with respect to the x axis, as it is shown in figure 3.6. This prepares the field in an initial state of polarization $|\psi_i\rangle = \cos(\alpha)\,|V\rangle + \sin(\alpha)\,|H\rangle$, where $|H\rangle$ and $|V\rangle$ are horizontal (y direction) and vertical (x direction) polarization states, respectively. Then, the light propagates through a birefringent plate that spatially separates the two orthogonal polarization components of the field by an small amount g. A second polarizer, located immediately after the plate and oriented at an angle β, performs the post-selection of the state $|\psi_f\rangle = \cos(\beta)\,|V\rangle + \sin(\beta)\,|H\rangle$. Finally, the intensity distribution of the field along the y direction is measured.

The transverse profile of the field, before passing through the plate, is described by a Gaussian wave function:

$$|\psi\rangle = \int \psi(x,y)\,|x,y\rangle\,dxdy, \quad \psi(x,y) = \Big(\frac{1}{\sqrt{2\pi}\sigma_{xy}^2}\Big)^{1/2} \exp\big[-(x^2+y^2)/(4\sigma_{xy}^2)\big]. \quad (3.49)$$

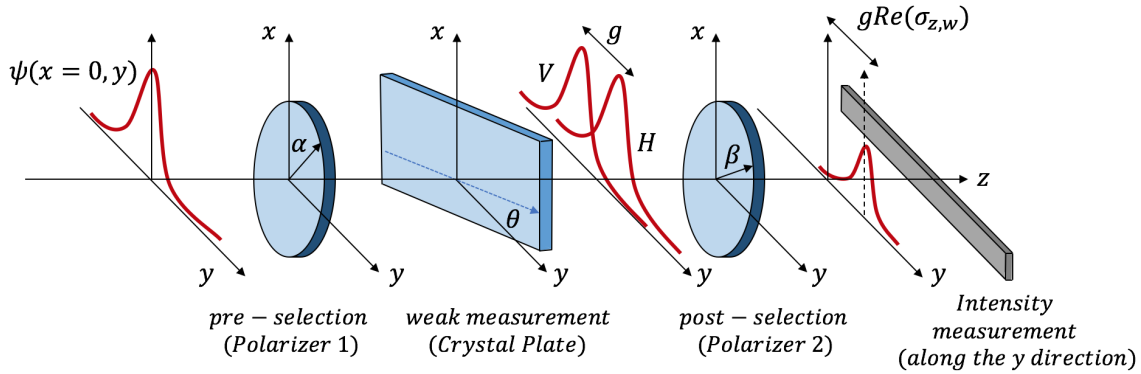


Figure 3.6: Schematic representation of the first experimental application of weak value amplification, employed to estimate a small birefringence-induced spatial separation of two orthogonal polarization components of a laser beam, of the order of micrometers.

The parameter σ_{xy} is the waist of the beam at the interface. The states $|x\rangle$ and $|y\rangle$ are eigenstates of the transverse position operator $\hat{X}$ and $\hat{Y}$, respectively. The conjugate variables to this set of operators are the transverse momenta, $\hat{P}_x$ and $\hat{P}_y$. The momentum in the transverse direction y will play the role of the apparatus variable.

Note that the wave function is normalized to one. Therefore, the actual intensity of the field at location (x, y) is given by $|E_0|^2 \cdot |\psi(x, y)|^2$, where $|E_0|^2$ is the intensity of the laser beam. In this sense, $|\psi(x, y)|^2$ represents the probability of detecting a photon at location (x, y), while $|E_0|^2$ is total number of photons in the beam.

On the other hand, the system variable is a polarization operator, $\hat{\sigma}_z = |H\rangle\langle H| - |V\rangle\langle V|$. The action of the plate can by described by a unitary operator that couples the transverse momentum of the field $\hat{P}_y$ with the polarization operator $\hat{\sigma}_z$ of the system:

$$\hat{U}(z_p + d, z_p) = \exp\Big[-(i/\hbar)(g/2)\hat{\sigma}_z\hat{P}_y\Big], \tag{3.50}$$

where z_p represents the position of the plate and d is its thickness. Notice that in this formalism time is replaced by the longitudinal coordinate. The parameter g depends on the incident angle of the field at the plate, the refraction index for the different polarization components and on the thickness of the plate.

After passing through the plate, the intensity of the field along the y axis is measured. This means that the transverse position $\hat{Y}$ of the apparatus is observed. Thus, expression

(3.24) allows to compute the probability of detecting a photon at position (x, y), which corresponds to

$$P(y|f) \approx |\psi(x,y)|^2 - g\,\mathrm{Re}(\sigma_{z,w})\partial_y|\psi(x,y)|^2 \approx |\psi[x, y - g\,\mathrm{Re}(\sigma_{z,w})]|^2. \tag{3.51}$$

Notice that the probability distribution has been approximately shifted in the y direction by g times the real part of the weak value of $\hat{\sigma}_z$. Also, note the imaginary part of the weak value plays no role in this experiment since the initial mean transverse momentum is zero and the wave function of the apparatus is real. Additionally, it is important to point out that the intensity of the beam has become weaker after passing through the second polarizer. The initial intensity $|E_0|^2$ has diminished to $|E_0|^2 \times P(f)$, where $P(f)$ is the probability of successful post-selection. This is the "cost" that must be paid as a trade off for the amplification of the displacement. In this experiment, the weak value of the polarization operator, between the initial and final polarization states, corresponds to

$$\sigma_{z,w} = \frac{\langle\psi_f|\,\hat{\sigma}_z\,|\psi_i\rangle}{\langle\psi_i|\psi_f\rangle} = -\frac{\cos(\beta+\alpha)}{\cos(\beta-\alpha)}. \tag{3.52}$$

Therefore, by making the difference between the orientations of the polarizers close to $\pi/2$, the weak value becomes large and the displacement can be amplified. Indeed, the displacement of the beam was approximately 20 times larger than g. In fact, when the initial and final polarization states were nearly the same, the displacement was $\sim g$ and unresolvable by the photo detector. When these states were orthogonal, then the amplification was larger, but the intensity of the beam was much weaker and a precise measurement of the intensity would have been required. In this sense, the weak value amplification protocol was the optimal measurement scheme that allowed to estimate the small birefringence-induced separation $\sim 0.6\ \mu m$.

3.4.2 Estimation of small differences of refraction indices

One of the most common uses of WVA has been the measurement of small differences of refraction indices [32, 80, 62, 63]. In these experiments, the system is typically the polarization degree of freedom of a photon and the measurement apparatus is its transverse

position. A birefringent element produces the coupling between the momentum of the meter and a polarization operator of the system, producing a spatial separation of the beam. Pre- and post-selection is performed using polarizing elements. In all these experiments the meter is prepared in a pure state, except for [63], where the effects of an incoherent apparatus are studied.

Hosten and Kwiat [64] were able to measure the spatial separation of two orthogonal polarization components produced by the spin Hall effect of light, allowing to check previous theoretical predictions [65]. In this experiment weak values were "enlarged" by taking advantage of the free evolution of the measuring device. Displacements of the order of one angstrom could be measured.

Angular rotations can be amplified using weak values. In [104] the orbital angular momentum of a beam is coupled with its polarization, via a spin-orbit interaction. Large weak values of the polarization operator produce an amplification of the angular rotations (due to the real part of the weak value) and a shift in the orbital angular momentum spectrum (due to the imaginary part).

3.4.3 Estimation of small angular displacements

On the other hand, weak value amplification has also been used to measure small angular displacements, of the order of femto radians [56, 66, 67, 81]. In these experiments the systems are which-path states of a photon in an interferometer, while the meter is its transverse momentum. The coupling between the system and the meter occurs due to a tilted mirror, which imparts a shift on the momentum wave function of the photons. Detection at the dark port of the interferometer corresponds to the post-selection of a certain which-path state. This setup has been adapted for frequency measurements [75]. In this case, the momentum kick, imparted by a dispersive prism, depends on frequency. Using this method, small changes in frequency were measured, achieving precision levels comparable to Fabry-Pérot interferometers. Analogously, this setup has been useful for phase amplification in interferometry [76], with similar sensitivities to balanced homodyne detection.

3.4.4 Weak values in interferometry

With regards to phase measurements in interferometry, in [26], the phase acquired by a beam travelling through one of the arms of an interferometer is amplified, using post-selection and standard techniques for phase measurements (the interferometer is slightly misaligned and the phase is extracted from the shift in the fringe pattern of the output beam). This experiment also admits a classical explanation. In [77], the weak value of a polarization operator was extracted from the difference of intensities of the output beams. In both experiments, the system is the polarization of a photon. In [78] a photon in one of the arms of a Hong-Ou-Mandel (HOM) interferometer acquires a small time delay, produced by a birefringent quartz plate. Pre- and post-selection is performed based on polarization. The HOM dip is shifted by a quantity proportional to the weak value of polarization projector.

It has been shown that imaginary weak values can outperform standard interferometry for the measurement of small time delays (longitudinal phase shifts). As it is explained in [105], an imaginary weak value of the polarization can enlarge the shift in the frequency spectrum of the pulse. This shift can improve the resolution limit of the interferometer.

3.4.5 Meter and system being two different particles

Unlike the previous experiments, in which the system and the apparatus are different degrees of freedom of the same particle, in [57, 74] the system and the meter are polarization states of two entangled photons, Moreover, in these experiments the meter variable is discrete (the polarization of a photon) and the coupling is produced by a non-deterministic optical controlled-sign gate [68]. The Leggett-Garg inequalities [69], which appeared, more or less, at the same time as weak values, and test the assumptions of macroscopic realism and non invasive measurements, were generalized for a system undergoing weak measurements in [70]. In [71, 72] it is shown that a violation of the generalized Leggett-Garg inequality is equivalent to an observation of an anomalous weak value. These violations and their correspondence with strange weak values were experimentally measured in solid-state devices [73], via deterministic coupling of two transmon qubits to a bus resonator,

and in optical setups where the polarizations of two photons became entangled through a non deterministic interaction in a controlled sign (CS) gate [74]. In [79] the system and the meter are two spins, coupled by heteronuclear coupling.

In [82] an experiment is proposed where a single photon enters an interferometer but the measurement of the photon number in one of the arms turns out to be larger than 1. This was accomplished through a photon-photon interaction in a non linear Kerr-type medium, where the phase shift given to a beam probe is proportional to the weak value of the number of photons in one of the arms of the interferometer. In [83] the amplification of the phase shift caused by a single photon was experimentally verified.

3.4.6 Direct measurement of a quantum state using weak values

In a finite dimensional Hilbert space (of dimension N), two bases $\{|n\rangle\}$ and $\{|m\rangle\}$ are said to be *mutually unbiased* (MU) if $|\langle n|m\rangle|^2 = 1/N$, for all n, m [84]. For an infinite dimensional Hilbert space, the square modulus of the overlap should be constant. For example, the position and momentum bases are MU because $|\langle q|p\rangle|^2 = 1/(2\pi\hbar)$, for all the eigenstates $|p\rangle$ and $|q\rangle$. This definition formalizes the concept of complementarity. Given a basis $\{|a\rangle\}$ of eigenstates of an observable $\hat{A}$, it is possible to pick an element of the unbiased basis, denoted by $|b_0\rangle$, for which the overlap $\langle b_0|a\rangle$ will be constant for all $|a\rangle$. Therefore, for any pure state $|\psi\rangle$, it is possible to define a factor $c = \langle b_0|a\rangle / \langle b_0|\psi\rangle$, that will be independent of $|a\rangle$. This constant factor allows to perform the following expansion:

$$c\,|\psi\rangle = c\sum_a \langle a|\psi\rangle\,|a\rangle = \sum_a \left(\frac{\langle b_0|a\rangle\langle a|\psi\rangle}{\langle b_0|\psi\rangle}\right)|a\rangle = \sum_a (\hat{\Pi}_a)_w\,|a\rangle\,. \tag{3.53}$$

Consequently, an unknown pure state $|\psi\rangle$ can be determined by measuring the weak values of the projectors $\hat{\Pi}_a = |a\rangle\langle a|$ between the initial state $|\psi\rangle$ and the final post-selected state $|b_0\rangle$, and then normalizing to eliminate the factor c. This procedure will generate the expansion coefficients of an unknown pure state $|\psi\rangle$ in the basis formed by the eigenstates of $\hat{A}$. This technique was proposed in [40], and constitutes a "direct" measurement of a quantum state, as opposed to quantum state tomography, in the sense that it requires

minimal post-processing and experimental error propagation [18]. It has been called direct state measurement (DSM) or weak state tomography [91].

In [40] the transverse wave function of a photon, $\psi(x)$, was reconstructed from the weak values of projectors into position eigenstates $\hat{\Pi}_x = |x\rangle\langle x|$, with post-selection of a state with zero transverse momentum, $|p = 0\rangle$. In this way, the weak values are proportional to the wave function of the photons. Indeed,

$$(\hat{\Pi}_x)_w = \frac{\langle p = 0|x\rangle\langle x|\psi\rangle}{\langle p = 0|\psi\rangle} = \psi(x)/\phi(0). \tag{3.54}$$

The factor $\phi(0)$ corresponds to the wave function in momentum space evaluated at $p = 0$, and can be removed by normalization after all the weak values are estimated. In this experiment, the meter was the polarization degree of freedom of the photons. The meter was observed in the linear polarization basis (to extract the real part of the weak value) and in the circular polarization basis (to extract the imaginary part of the weak value). Regarding the direct measurement of the spatial mode of photons an overview of the progress in this field is presented in [93].

Weak state tomography has been used for measurement of polarization states in [89, 90, 95]. In [87] this method allowed the measurement of a state vector of high dimensionality ($d = 27$), in the discrete basis of angular momentum. States with higher dimensionality were measured in [94].

The direct measurement of a quantum state has been also applied to mixed states [92, 89, 88, 95]. A comparison between direct state measurement and tomography is presented in [91]. In [96] the measurement of the density matrix of two entangled photons is investigated while other theoretical studies regarding DSM can be found in [85, 86].

3.4.7 Slow and fast light

In [97] two polarizers are put at the ends of an optical birefringent fiber (Lyot filter). The first pre-selects an initial polarization state and the second post-selects a certain final state. The fiber performs a weak measurement of a spin-like polarization operator,

introducing a small phase difference between the different polarization eigenstates. The mean arrival time of an optical pulse is amplified (or reduced), i.e. the group velocity of the pulse is modified, resulting in slow or superluminal propagation along the fiber. Similarly, in [98] slow and superluminal propagation of polarized microwaves was experimentally demonstrated using weak values of polarization operators. Importantly, in this article the relationship between weak values and the linear response function of a system was also described.

"Weak superluminal speed", namely, a large weak value of an internal degree of freedom of a particle (that plays the role of a velocity), was theoretically explained in [50]. The authors also showed that the electromagnetic field emitted by a charged particle, which moves with a (weak) superluminal speed, corresponds to Cherenkov radiation.

3.4.8 Weak values and tunneling times

The weak value of the "dwell time operator" [100] provides a description of the time spent by a particle tunnelling through a potential barrier. The dwell time operator is a projector, with eigenvalue equal to one when the particle is inside in the barrier region, and eigenvalue equal to zero if the the particle is outside the region. Its expectation value, divided by the incident flux, corresponds to the "dwell time" [99]. The weak value of the dwell time operator between an initial state associated to a wave packet incident from the left, $|i\rangle$, and a final state $|f\rangle$ that represents the state of a transmitted particle, divided by the incident flux, is a complex time, related to the Larmor times [99, 103] and to the complex time presented by Sokolovski and Connor using Feynman path-integrals [102]. The real part of the weak value is associated to the tunnelling time, while the imaginary part represents the measurement back-action. This time was measured in [101], using a Bose-Einstein condensate of Rb atoms, going through an optical potential barrier.

3.4.9 Superoscillations

One would expect that a Fourier series, with a finite number of frequencies ω_n, could not oscillate with a higher frequency than $\max\{\omega_n\}$. However, over certain region of their domain, some sequences can indeed oscillate faster than any of its Fourier components.

This behaviour has been called *superoscillation*.

This phenomenon appears in weak measurements performed on pre- and post-selected quantum systems. Consider a measurement of an observable $\hat{A}$ described by the von Neumann model, with an interaction of the form $\hat{P}\hat{A}$. Assume that the apparatus starts in a pure state, whose wave function in momentum space is $\psi(p)$. After an interaction with a system, pre-selected in a pure state $|\psi_i\rangle$ and post-selected in a pure state $|\psi_f\rangle$, the wave function of the apparatus in momentum space transforms according to:

$$\psi(p) \longrightarrow \frac{\langle\psi_f|\,\hat{M}_p\,|\psi_i\rangle}{\sqrt{\int dp|\,\langle\psi_f|\,\hat{M}_p\,|\psi_i\rangle\,|^2}}. \tag{3.55}$$

Recall from (2.45) that the measurement operators are $\hat{M}_p = \psi(p)e^{-igp\hat{A}/\hbar}$. The term in the denominator is the square root of the probability of post-selection (3.21), that will be denoted in this section simply as P. Thus, after the interaction the wave function may be expressed as $\psi(p)\cdot[\langle\psi_f|\exp\Big\{-igp\hat{A}/\hbar\Big\}|\psi_i\rangle/\sqrt{P}]$. The factor inside the brackets, that will be denoted as $S(p)$, represents the correction to the initial wave function of the apparatus due to the interaction with the pre- and post-selected system. This factor may be written as a Fourier series, namely,

$$S(p) = \frac{1}{\sqrt{P}}\langle\psi_f|\exp\Big\{-igp\hat{A}/\hbar\Big\}|\psi_i\rangle = \frac{1}{\sqrt{P}}\sum_{k=0}^{N-1}\langle\psi_f|\,\hat{\Pi}_k\,|\psi_i\rangle\exp[-igpa_k/\hbar]. \tag{3.56}$$

The value a_k is the k-th eigenvalue of $\hat{A}$ and $\hat{\Pi}_k = |a_k\rangle\langle a_k|$ is the projector into the space spanned by the corresponding eigenvector $|a_k\rangle$. The summation goes over all the eigenvalues, which are assumed to be finite ($k = 0, ..., N-1$). The frequencies are $g|a_k|/\hbar$ and, since the set of eigenvalues is finite, the sequence is band-limited. Consequently, one would expect that it will oscillate at some frequency within the minimum and the maximum frequency. Nevertheless, in order to show that this is not always the case, consider a second order expansion of $S(p)$:

$$S(p) \approx \frac{1}{\sqrt{P}}\Big[1 - \frac{igp}{\hbar}A_w + \frac{1}{2}\Big(-\frac{igp}{\hbar}\Big)^2 A_w^2\Big]. \tag{3.57}$$

For the simplicity of the argument we will assume that the weak value is real. Let us consider the difference: $\alpha = A_w^2 - (A_w)^2$. The square root of $|\alpha|$ corresponds to the "weak uncertainty" defined in [106]. Since the weak value is real, α is a real number that may be positive or negative. By using α to rewrite expression (3.57),

$$S(p) \approx \frac{1}{\sqrt{P}}\Big[1 - \frac{igp}{\hbar}A_w + \frac{1}{2}\Big(-\frac{igp}{\hbar}\Big)^2 (A_w)^2 - \frac{1}{2}\Big(\frac{gp}{\hbar}\Big)^2 \alpha\Big]. \tag{3.58}$$

If the the weak value is large enough so that its square overcomes the second order weak value, then α will be negative, and the last term of (3.58) will represent an "anti-Gaussian" behaviour of $S(p)$ away from the origin. This tells that near the origin $S(p)$ behaves like $\exp(-igpA_w/\hbar)$, namely, it oscillates with a frequency $g|A_w|/\hbar$, that is larger than any of the frequencies of its Fourier components. Therefore, the sequence (3.56) exhibits a superoscillatory behaviour near the origin (as long as the weak value is large). Away from the origin $S(p)$ will start to grow like p^2. In fact, studies of superoscillatory functions show that these functions are extremely small in the region where they superoscillate, which has consequences for information theory [107].

Recall that final wave function of the apparatus in momentum space is the product of the initial wave function $\psi(p)$ and the sequence $S(p)$. The initial wave function $\psi(p)$ should be wide enough to capture the superoscillatory behaviour near the origin. However, at the same time, we want that $\psi(q)$ has tails that decay fast enough to cancel the "anti-Gaussian" behaviour of $S(p)$ away from the origin. In this sense, $\psi(p)$ should not be too wide. Under these conditions, $\psi(p)S(p) \approx \psi(p)\exp(-igpA_w/\hbar)/\sqrt{P}$. The phase factor will produce a large shift of gA_w in coordinate space, namely, $\psi(q) \approx \psi(q - gA_w)$.

Superoscillations not only appear in quantum physics, but also in the study of radars, optical vortices, sub-wavelength microscopy, fractals and other fields [107]. A description of the mathematical properties of superoscillations and a survey of the existing literature is presented in [108].

Chapter 4

Parameter estimation

In this chapter the tools of estimation theory are applied to weak measurements, with and without post-selection. The first half of the chapter presents the basics tools for parameter estimation. From section 4.1 to section 4.5 the elements of classical estimation theory are introduced. In section 4.6 the quantum Fisher information and the quantum Cramér-Rao bound are presented. In section 4.7 the role of the quantum Fisher information in phase measurements and interferometry is analyzed and in section 4.8 the relationship between the quantum Fisher information and the Bures distance is briefly commented. In the last section the Fisher information in weak measurements is studied. Fist, weak measurements with and without post-selection are taken into consideration. Finally, the Fisher information in both types of measurements is studied when noise (white and correlated) affects the measurements.

4.1 Bayes' theorem

Let g be some unknown parameter that we want to estimate. Assume that we have some prior knowledge about g, which is expressed in the *prior probability distribution* denoted as $P(g)$. Suppose that we perform a measurement on a system in order to extract some information that will help us to estimate g. Let the result of the measurement R (a random variable) be equal to r (a number). Bayes' theorem [7] establishes that

$$P(g|r) = \frac{P(r|g)P(g)}{P(r)}. \tag{4.1}$$

The prior probability reflects our knowledge about the parameter before to the measurement. The probability distribution $P(g|r)$ is called the *posterior probability distribution*, and reflects our knowledge about g given that the result r was obtained. Expression (4.1) describes therefore how our knowledge about g is updated after a measurement.

The function $P(r|g)$ that appears in the right hand of (4.1) can be understood as a function of r or as a function of g. In the first case, it is a conditional probability distribution; the probability to obtain the result r given that the unknown parameter is equal to g. In the second case, viewed as function of g, it is called the *likelihood function*. It depends on the physical process employed to estimate g. Typically, the relationship between the measurement result R (a random variable) and the parameter g can be written as $R = f(g) + X$, where X is some random variable and f is a deterministic function. From this relationship, the likelihood function can be obtained.

4.2 Estimation theory: bayesian and of non random parameters

An estimator $\hat{g}$ is a function of the measurement result R, i.e. $\hat{g} = \hat{g}(R)$. In Bayesian estimation theory, in order to construct an estimator of the parameter of interest, first we need to choose a *cost function*. The cost function, $C = C(\hat{g}, g)$, is a function of two variables; it depends on the estimator and on the parameter. The expectation value of the cost function, calculated using the joint probability distribution $P(r, g)$ is called the *risk function* [119],

$$\mathfrak{R} = \int drdgC(\hat{g}, g)P(r, g) = \int drP(r) \int dgC(\hat{g}, g)P(g|r). \tag{4.2}$$

Notice that in the inner integral of the last identity the posterior probability distribution is employed. The Bayes estimator $\hat{g}$ is chosen then in order to minimize the risk. A common

cost function is the uniform cost, which is equal to zero when $|\hat{g} - g| \leqslant \epsilon$ and equal to one in any other case. The corresponding risk function is given by

$$\mathfrak{R} = \int dr dg C(\hat{g}, g) P(r, g) = \int dr P(r) \left[1 - \int_{\hat{g}-\epsilon}^{\hat{g}+\epsilon} dg P(g|r) \right]. \tag{4.3}$$

The estimator should therefore maximize the inner integral (in order to minimice the risk). Consequently, if ϵ is small, then the estimator should be equal to the value of g that maximizes the posterior probability distribution. This estimator is called the *maximum a posteriori* (MAP) estimator and is obtained by solving the equation:

$$\frac{\partial \ln P(g|r)}{\partial g} = 0. \tag{4.4}$$

The logarithm is taken in order to simplify the calculations. This equation can be formulated in terms of the likelihood function using Bayes' theorem (4.1),

$$\partial_g \ln P(r|g) + \partial_g \ln P(g) - \partial_g \ln P(r) = 0. \tag{4.5}$$

Notice that the probability $P(r)$ is independent of the parameter and hence the third term vanishes. The MAP estimator is therefore constructed by solving the equation $\partial_g \ln P(r|g) + \partial_g \ln P(g) = 0$. Assume also that we are "completely ignorant" about the parameter. This means that the prior distribution is uniform, i.e. $\partial_g P(g) = 0$. In this case, the optimization of the posterior distribution (4.5) is equivalent to the maximization of the likelihood function. Equation (4.5) reduces therefore to *likelihood equation*: $\partial_g \ln P(r|g) = 0$. This estimator is called the *maximum likelihood* (ML) estimator,

$$\hat{g}_{ML} = \arg\max_g \Big\{ P(r|g) \Big\}. \tag{4.6}$$

As a summary, when the cost function is uniform and $\epsilon \to 0$, then the Bayesian estimator reduces to the MAP estimator. Additionally, if the priori distribution is uniform, it becomes the ML estimator. Certainly, there are different cost functions, being the mean squared error the most common (see chapter 2 of [119]). When the parameter to be estimated is not

random, but a deterministic and unknown quantity, then the Bayesian approach, i.e. the minimization of the risk function (4.2), should be modified. For non random parameters, there are different measures of the quality of an estimator, which are introduced in the next section. These measures are defined in terms of the likelihood function and not in terms of the posteriori distribution.

4.3 Bias, variance and mean squared error

Let $\hat{g}$ be an estimator of g. The mean value of the estimator, averaging over all the possible measurement results, will be denoted as

$$\langle \hat{g} \rangle = \int \hat{g}(r) P(r|g) dr. \tag{4.7}$$

Notice that the mean value is a function of g, because the likelihood function depends on this parameter. The dependence on g is not explicit in the notation $\langle \hat{g} \rangle$, so the reader should keep in mind that it may depend on g. Analogously, the variance of the estimator is defined as

$$\left\langle \Delta \hat{g}^2 \right\rangle = \int [\hat{g}(r) - \langle \hat{g} \rangle]^2 P(r|g) dr. \tag{4.8}$$

On the other hand, the difference $\hat{g}(R) - g$ is the error of the estimator. The mean value of the error is called the *bias* of the estimator,

$$b(g) = \int [\hat{g}(r) - g] P(r|g) dr = \langle \hat{g} \rangle - g. \tag{4.9}$$

The bias may be a function of g. In this case, the estimator has unknown bias. It could be also a constant, in which case the estimator has a known bias. If $\forall g \; b(g) = 0$, the estimator is said to be *unbiased*. The mean value of the squared error (MSE), denoted as $\epsilon(g)$, is one of the simplest measures of the precision of an estimator. It is defined as

$$\epsilon(g) = \int [\hat{g}(r) - g]^2 P(r|g) dr. \tag{4.10}$$

The mean squared error can be decomposed into the variance and the bias of the estimator, as follows:

$$\epsilon(g) = \left\langle \Delta\hat{g}^2 \right\rangle + b^2(g). \tag{4.11}$$

This expression shows that for unbiased estimators the variance is equal to the mean squared error. For the construction of estimators of non random parameters, we will interested in estimators with small $\left\langle \Delta\hat{g}^2 \right\rangle$ or $\epsilon(g)$. In the next section, we introduce therefore the so called *Cramér-Rao Bound* [147], which constitutes a lower bound of the mean squared error (of any biased estimator) or of the variance (of any unbiased estimator). The quality of an estimator can be measured in relation with the achievement of this lower bound.

4.4 Cramér-Rao bound

As anticipated, a lower bound of $\epsilon(g)$ will be presented in this section. We start with the obvious fact that $\int(\hat{g} - \langle\hat{g}\rangle)P(r|g)dr = 0$. By differentiating this identity with respect to g the following expression expression is obtained:

$$\int(\hat{g} - \langle\hat{g}\rangle)\partial_g P(r|g)dr = \frac{d\langle\hat{g}\rangle}{dg}. \tag{4.12}$$

On the other hand, by differentiating the bias (4.9) with respect to g, it is clear that $d\langle\hat{g}\rangle/dg = b'(g) + 1$, where $b'(g)$ denotes differentiation with respect to g. It is also useful to note that $\partial_g P(r|g)dr = P(r|g)\partial_g \ln[P(r|g)]$. These considerations allow to rewrite expression (4.12) as

$$\int \left\{[\hat{g}(r) - \langle\hat{g}\rangle]\sqrt{P(r|g)}\right\} \cdot \left\{\sqrt{P(r|g)}\partial_g \ln[P(r|g)]\right\} dr = 1 + b'(g). \tag{4.13}$$

The integral form of the Cauchy-Schwarz inequality establishes that $\int |f(x)|^2 dx \cdot \int |g(x)|^2 dx \geqslant |\int f(x)h(x)dx|^2$. Applying this inequality to the last expression, is it clear that

$$\int [\hat{g}(r) - \langle \hat{g} \rangle]^2 P(r|g) dr \cdot \int \{\partial_g \ln[P(r|g)]\}^2 P(r|g) dr \geqslant [1 + b'(g)]^2. \tag{4.14}$$

The first term in the left hand side is the variance of the estimator (4.8). The second term corresponds to the *Fisher information*, denoted as $\mathcal{I}(g)$,

$$\mathcal{I}(g) = \int \{\partial_g \ln[P(r|g)]\}^2 P(r|g) dr = \int \frac{\dot{P}(r|g)^2}{P(r|g)} dr, \tag{4.15}$$

where $\dot{P}(r|g) = \partial_g P(r|g)$. Notice that the units of $\mathcal{I}(g)$ correspond to the units of $1/g^2$. From its definition the Fisher information can be understood as the expectation value of the square of $\partial_g \ln[P(r|g)]$, namely, of the gradient of the log-likelihood function. This function is called *score function*,

$$\mathcal{S}(g) = \partial_g \ln[P(r|g)] = \frac{\partial_g P(r|g)}{P(r|g)}, \tag{4.16}$$

and measures how sensitive is the likelihood function with respect to changes in the parameter. The average value of the score is zero [1]. Therefore, the Fisher information can also be viewed as the *variance of the score function*. For the toss of a coin, i.e. for a Bernoulli variable with probability of success equal to a parameter p, the Fisher information with respect to the parameter is $1/[p(1-p)]$. The information is larger near the extremes, $p = 1$ or $p = 0$, when the coin is highly unfair or unbalanced. The information is minimal when $p = 1/2$, i.e. when the coin is balanced.

Besides its definition as the variance of the score function (4.15), the Fisher information can also be defined as the average value of the (negative) curvature of the likelihood function. Indeed, note that if the likelihood function is twice differentiable, then

$$\partial_{gg} \ln[P(r|g)] = \frac{\partial_{gg} P(r|g)}{P(r|g)} - \{\partial_g \ln[P(r|g)]\}^2. \tag{4.17}$$

[1] This is true whenever the integral and the partial derivative can be exchanged, $\mathbb{E}[\mathcal{S}(g)] = \int \partial_g P(r|g) dr = (d/dg) \int P(r|g) dr = 0$.

When the partial derivatives operators can be interchanged with the integral operators, by taking the expectation value of (4.17), it is easy to see that (4.15) can be equivalently expressed as

$$\mathcal{I}(g) = -\mathbb{E}\{\partial_{gg} \ln[P(r|g)]\} = -\int \partial_{gg} \ln[P(r|g)] P(r|g) dr, \tag{4.18}$$

which shows that the Fisher information may be viewed as the expectation value of the negative curvature (second derivative with respect to the parameter) of the log-likelihood function. Typically this expression provides an easier way for the calculation of the Fisher information than the definition (4.15).

In terms of the introduced definitions (4.15) and (4.8), the inequality (4.14) can be written as $\left\langle \Delta\hat{g}^2 \right\rangle \mathcal{I}(g) \geqslant [1 + b'(g)]^2$. Therefore, the Fisher information can be employed to construct a lower bound on the variance of any estimator of g. This lower bound is the *Cramér-Rao bound* [119, 120]:

$$\left\langle \Delta\hat{g}^2 \right\rangle \geqslant \frac{[1 + b'(g)]^2}{\mathcal{I}(g)}, \tag{4.19}$$

and the inequality is called the Cramér-Rao inequality. The larger the Fisher information contained in the likelihood function, the smaller will be the lowest attainable variance of any estimator. Restricted to the class of unbiased estimators (i.e. estimators such that $\forall g$ $b(g) = 0$), the Cramér-Rao bound reduces to the inverse of the Fisher Information,

$$\left\langle \Delta\hat{g}^2 \right\rangle \geqslant \frac{1}{\mathcal{I}(g)}. \tag{4.20}$$

On the other hand, for biased estimators, by expressing the variance in terms of the MSE and the squared bias (4.11), the Cramér-Rao inequality can be formulated as

$$\epsilon(g) \geqslant \frac{[1 + b'(g)]^2}{\mathcal{I}(g)} + b^2(g). \tag{4.21}$$

Notice that for biased estimators the Cramér-Rao inequality actually establishes a lower bound on the mean squared error. If the bias is such that $|1 + b'(g)| < 1$, then the lower

bound can be inferior than the lower bound for unbiased estimators, $\mathcal{I}^{-1}(g)$.

An estimator that saturates the Cramér-Rao inequality (4.19) for all the possible values of the parameter is said to be an *efficient estimator*. An efficient estimator may depend on the parameter we wish to estimate and therefore require good prior knowledge about it, which is typically not the case. Therefore, an efficient estimator may not always be feasible. Moreover, in the class of unbiased estimators, an efficient estimator may not even exist, i.e. the inequality (4.20) might not be saturated by any unbiased estimator. In this class, the estimator that has minimum variance for all the possible values of the parameter is called the *minimum-variance unbiased estimator* (MVUE). Certainty, if an efficient unbiased estimator exists, then it will be also the MVUE.

As pointed out in the previous section, the variance (for unbiased estimators) and the mean squared error (for biased estimators) are the figures of merit to measure the performance of an estimator. Although these are the formal statistical tools for the estimation problem, sometimes the signal to noise ratio (SNR) is also taken into consideration in order to evaluate the quality of an estimator. In this work, we will employ the definition presented in [82, 121], where the SNR is defined as the ratio between the mean value of the estimator (the signal) and the standard deviation of the estimator (the noise). Hence,

$$SNR = \frac{\langle \hat{g} \rangle}{\sqrt{\langle \Delta \hat{g}^2 \rangle}}. \tag{4.22}$$

A "good" SNR should be greater than one, i.e the level of the signal should be greater than the level of the noise. Since the Cramér-Rao bound defines a lower bound on the variance, it constitutes an upper bound on the SNR, which is limited by $\langle \hat{g} \rangle \sqrt{\mathcal{I}(g)}/|1 + b'(g)|$.

4.5 The classical limit

In most cases there will be a collection of M measurement results $r_1, ..., r_M$ available for the estimation of g (M represents therefore the number of repetitions of the same experiment). In this case, the likelihood function is given by $P(\vec{r}|g) = P(r_1, ..., r_M|g)$. If

the observations are independent, then

$$P(\vec{r}|g) = \prod_{i=1}^{M} P(r_i|g). \tag{4.23}$$

The Fisher information of independent random variables is additive. Indeed, it is easy to check that the Fisher information in (4.23) becomes

$$\mathcal{I}(g) = \int d\vec{r} \Big\{ \partial_g \ln[P(\vec{r}|g)] \Big\}^2 P(\vec{r}|g) = \sum_{i=1}^{M} \mathcal{I}_i(g), \tag{4.24}$$

where $d\vec{r} = dr_1 ... dr_M$ and $\mathcal{I}_i(g)$ is the Fisher information associated to the i-th observation, namely,

$$\mathcal{I}_i(g) = \int dr_i \Big\{ \partial_g \ln[P(r_i|g)] \Big\}^2 P(r_i|g). \tag{4.25}$$

If all the observations are identically distributed, then $\mathcal{I}_i(g) = \mathcal{I}_1(g)$. In this case, the Cramér-Rao inequality for unbiased estimators becomes

$$\left\langle \Delta \hat{g}^2 \right\rangle \geqslant \frac{1}{M\mathcal{I}_1(g)}. \tag{4.26}$$

Therefore, the error (square root of the variance) of an efficient unbiased estimator scales as $1/\sqrt{M}$, where M is the number of independent observations. This type of scaling is considered to be classical, and arises as a consequence of adding independent information. The limit on the precision of any estimator $\sim 1/\sqrt{M}$ has been called the "classical limit" for metrology.

As an example, let us take into consideration the sample mean estimator (SME), which simply takes the average of all the measurement results,

$$\hat{g}_{SME}(M) = \frac{\sum_{i=1}^{M} r_i}{M}, \tag{4.27}$$

as an estimator of the population mean. If the elements of the sample are independent and identically distributed, with mean μ and variance σ^2, it is easy to verify that the SME is

an unbiased estimator of the population mean,

$$\mathbb{E}[\hat{g}_{SME}(M)] = \mu. \tag{4.28}$$

The variance of the sum of M independent variables corresponds to the sum of the variances and taking the average adds a $1/M^2$ factor. Therefore, the variance of the SME is given by

$$\left\langle [\Delta \hat{g}_{SME}(M)]^2 \right\rangle = \frac{\sigma^2}{M}, \tag{4.29}$$

which means that the error of the estimator presents the classical scaling $\sim 1/\sqrt{M}$, i.e. it achieves the classical limit for any finite sample. This does not mean that the estimator is necessary efficient, which occurs when the Cramér-Rao inequality is saturated. If each observation is taken from a normal population, then the estimator would be efficient, since $1/\sigma^2$ corresponds to the Fisher information with respect to the mean of a normal distribution. In this particular case, the SME estimator is also efficient for finite samples.

Typically, the efficiency can be achieved asymptotically, i.e. for a large number of observations. Let us consider the *maximum likelihood estimator* $\hat{g}_{ML}$, presented in section 4.2, and defined as

$$\hat{g}_{ML}(M) = \arg\max_{g} \Big\{ P(r_1, ..., r_M|g) \Big\}. \tag{4.30}$$

This estimator is *asymptotically efficient*, i.e. it saturates the Cramér-Rao inequality in the limit $M \to \infty$. For random samples (independent and identically distributed random variables), this means that $\sqrt{M}[\hat{g}_{MLE}(M) - g]$ *converges in distribution* to a random variable, normally distributed, with zero mean and variance equal to $\mathcal{I}^{-1}(g)$, where $\mathcal{I}^{-1}(g)$ is the inverse of the Fisher information contained in a single observation. In other words, for a sufficiently large sample,

$$\left\langle [\Delta \hat{g}_{MLE}(M)]^2 \right\rangle \approx \frac{1}{M\mathcal{I}(g)}. \tag{4.31}$$

This occurs as a consequence of the central limit theorem [148, 119].

4.6 Quantum Fisher information

For the quantum case, the likelihood function is $P(r|g) = \mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]$, where $\rho(g)$ is a quantum state that depends on the parameter g that we want to estimate and $\hat{E}_r$ is the effect operator associated to the outcome r of the measurement. Consequently, the classical Fisher information (4.15) can be expressed as

$$\mathcal{I}(g) = \int dr \frac{\Big\{\partial_g \,\mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]\Big\}^2}{\mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]}. \tag{4.32}$$

We want to maximice this expression over all possible POVM's. In order to derive an upper bound for the classical Fisher information let us define the *Symmetric Logarithmic Derivative* (SLD) as an hermitian operator L_g that satisfies the equation

$$\frac{L_g\rho(g) + \rho(g)L_g}{2} = \partial_g\rho(g). \tag{4.33}$$

Note that, in terms of the SLD, the classical fisher information (4.32) can be bounded in the following way:

$$\mathcal{I}(g) = \int dr \frac{\mathrm{Re}\Big\{\,\mathrm{Tr}\Big[\rho(g)\hat{E}_rL_g\Big]\Big\}^2}{\mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]} \leqslant \int dr \left|\frac{\mathrm{Tr}\Big[\rho(g)\hat{E}_rL_g\Big]}{\sqrt{\mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]}}\right|^2 \tag{4.34}$$

$$= \int dr \left|\,\mathrm{Tr}\left[\frac{\sqrt{\rho(g)}\sqrt{\hat{E}_r}}{\sqrt{\mathrm{Tr}\Big[\hat{E}_r\rho(g)\Big]}}\sqrt{\hat{E}_r}L_g\sqrt{\rho(g)}\right]\right|^2. \tag{4.35}$$

The inequality is achieved when the trace $\mathrm{Tr}\Big[\rho(g)\hat{E}_rL_g\Big]$ is real. Considering now the matrix form of the Cauchy-Schwarz inequality, $|\,\mathrm{Tr}\Big[\hat{A}^\dagger\hat{B}\Big]|^2 \leqslant \mathrm{Tr}\Big[\hat{A}^\dagger\hat{A}\Big]\,\mathrm{Tr}\Big[\hat{B}^\dagger\hat{B}\Big]$, the prior upper bound can be optimized over all over all possible POVM's, according to:

$$\mathcal{I}(g) \leqslant \int dr\, \mathrm{Tr}[L_g \rho(g) L_g] = \mathrm{Tr}\big[L_g^2 \rho(g)\big] \equiv \mathcal{I}_Q(g). \tag{4.36}$$

This upper bound is called the *quantum Fisher information* (QFI), and will be denoted as $\mathcal{I}_Q(g)$. It is easy to show that the expectation value of the SLD is zero. Therefore, the quantum Fisher information can be defined as the variance of the SLD. In this sense, the SLD plays the role of the score function. Using the QFI in (4.19) it is possible to derive the *quantum Cramér-Rao bound* [122, 123, 124],

$$\big\langle \Delta \hat{g}^2 \big\rangle \geqslant \frac{[1 + b'(g)]^2}{\mathcal{I}(g)} \geqslant \frac{[1 + b'(g)]^2}{\mathcal{I}_Q(g)}. \tag{4.37}$$

The first is *classical Cramér-Rao* inequality while the second corresponds to the *quantum Cramér-Rao* inequality.

The optimal POVM is the set of projectors $\{|L_i\rangle\langle L_i|\}$, where $|L_i\rangle$ is one of the eigenstates of L_g. This measurement may depend on the (unknown) parameter we wish to estimate and therefore it may be unfeasible to implement. It is important not to confuse the optimal POVM with the estimator. The first is given by the set of projectors $\{|L_i\rangle\langle L_i|\}$, while the second depends on how the acquired data is processed. In section (4.5) we pointed out that the ML estimator can be employed for these purposes because it saturates the Cramér-Rao inequality in the limit of a large number of observations (the classical inequality for any POVM and the quantum inequality if the optimal POVM is implemented).

4.6.1 Quantum Fisher information in mixed and pure states

Consider a mixed state with spectral decomposition $\rho(g) = \sum_{k=1}^{R} \lambda_k |\psi_k\rangle\langle\psi_k|$. The dimension of the support of $\rho(g)$ is $R \leqslant D$ (D being the dimension of the Hilbert space). Notice that both, the eigenvalue λ_k or the ket $|\psi_k\rangle$, may depend on the parameter. The partial derivatives of the eigenvalue and the eigenstate with respect to g will be denoted by $\dot{\lambda}_k$ and $\left|\dot{\psi}_k\right\rangle$, respectively. In this case, the solution of equation (4.33) may be expressed

as

$$L_g = \sum_{n=1}^{R} \left(\frac{\dot{\lambda}_n}{\lambda_n}\right) |\psi_n\rangle\langle\psi_n| + 2 \sum_{n\neq m} \left(\frac{\lambda_n - \lambda_m}{\lambda_n + \lambda_m}\right) \left\langle \psi_m \middle| \dot{\psi}_n \right\rangle |\psi_m\rangle\langle\psi_n| \,. \tag{4.38}$$

The summation in the second term goes over all n and m such that $\lambda_n + \lambda_m \neq 0$. Using this expression in (4.36), the quantum Fisher information can be expressed as

$$\mathcal{I}_Q(g) = \sum_{n=1}^{R} \left(\frac{\dot{\lambda}_n^2}{\lambda_n}\right) + 2 \sum_{n\neq m} \frac{(\lambda_n - \lambda_m)^2}{\lambda_n + \lambda_m} |\left\langle \psi_m \middle| \dot{\psi}_n \right\rangle|^2. \tag{4.39}$$

The first term corresponds to the information contained in the eigenvalues and represents the "classical contribution", while the second term reflects the information contained in the states and is properly the quantum contribution. For a pure state, i.e. $\rho(g) = |\psi\rangle\langle\psi|$, expression (4.39) reduces to

$$\mathcal{I}_Q(g) = 4\left\langle \dot{\psi} \middle| \dot{\psi} \right\rangle - 4\left| \left\langle \psi \middle| \dot{\psi} \right\rangle \right|^2. \tag{4.40}$$

4.6.2 Quantum Fisher information in unitary processes

Assume now that the state $\rho(g)$ is generated by a *unitary process*, namely, $\rho(g) = e^{-ig\hat{T}}\rho_0 e^{ig\hat{T}}$. The initial state ρ_0 is a mixed state with spectral decomposition $\rho_0 = \sum \lambda_n |\phi_n\rangle\langle\phi_n|$, and $\hat{T}$ is an hermitian operator called the *generator* of the process. The initial state does not depend on g, only the unitary process applied to it depends on the parameter[2]. In this case, the QFI (4.39) corresponds to

$$\mathcal{I}_Q(g) = 2 \sum_{n\neq m} \frac{(\lambda_n - \lambda_m)^2}{\lambda_n + \lambda_m} |\hat{T}_{m,n}|^2, \tag{4.41}$$

where $\hat{T}_{m,n} = \langle\phi_m| \hat{T} |\phi_n\rangle$. Since the generator $\hat{T}$ is independent of the parameter, the quantum Fisher information will also be independent of it. In fact, this occurs actually because the SLD operator (4.38) is independent of the parameter. Consequently, the optimal POVM will not depend on g and therefore it will possible in principle to implement

[2]We will also assume that the generator is independent of the parameter.

it, although we might be completely ignorant about the real value of the parameter. This type of optimal measurements, which are independent of the parameter, are called by Braunstein, Caves and Milburn [125] "global optimal measurements". When the generator has no degeneracies, such measurements can described in terms of effect operators $\hat{E}(x)$, where x is a real number that takes the same values as the parameter. If the generator is degenerated, then the optimal POVM requires more information than the one provided by a single real number x (see [125] for more details).

If the initial state is pure $\rho_0 = |\phi\rangle\langle\phi|$, then the quantum Fisher information (4.41) corresponds to

$$\mathcal{I}_Q(g) = 4\left\langle \Delta\hat{T}^2 \right\rangle_\phi = 4\langle\phi|\,\hat{T}^2\,|\phi\rangle - 4|\langle\phi|\,\hat{T}\,|\phi\rangle\,|^2, \tag{4.42}$$

namely, four times the variance of the generator. Notice that the quantum Fisher information of a pure state is greater than (4.41). Therefore, for initial pure states and any unbiased estimator of g, the quantum Cramér-Rao bound takes the form of an uncertainty principle,

$$\left\langle \Delta\hat{g}^2 \right\rangle \cdot \left\langle \Delta\hat{T}^2 \right\rangle \geqslant 1. \tag{4.43}$$

This is a *parameter-based uncertainty principle*, that resembles the standard uncertainty principle, with the difference that $\left\langle \Delta\hat{g}^2 \right\rangle$ refers to the variance of an unbiased estimator (which does not necessarily corresponds to an hermitian operator).

Finally, it is worth to mention that the quantum Fisher information (4.42) can be maximized over all possible initial states,

$$\max_\phi \left\{ 4\left\langle \Delta\hat{T}^2 \right\rangle_\phi \right\} = (\Delta t)^2, \tag{4.44}$$

where Δt is the difference between the maximum and minimum eigenvalues of the generator $\hat{T}$. The optimal initial state is an equal superposition of the eigenstates associated to the maximum and minimum eigenvalues of the generator [111].

4.7 Quantum Fisher information in phase measurements and interferometry

In quantum mechanics the scaling of the error of the type $1/\sqrt{M}$ appears in different situations. Consider, as an example, the measurement of the phase of an harmonic oscillator [126]. Assume that an harmonic oscillator is prepared in a coherent state $|\alpha\rangle$ and then evolves freely. For simplicity suppose that $\alpha \in \mathbb{R}$. After a time t the unitary evolution produced by $\exp\{-i\omega t\hat{n}\}$ will impart a phase of $g = \omega t$, and the final state will be the coherent state $\left|e^{-ig}\alpha\right\rangle$. Since the final state is generated by a unitary process, the quantum Fisher information is given by $4|\alpha|^2$, i.e. four times the mean energy (or the mean number of particles) of the oscillator. Therefore, the error of any estimator of the phase will be limited by a factor $\sim 1/\sqrt{N}$, where N denotes the mean number of particles.

This limit has been called the *standard quantum limit* (SQL) [149] for the measurement of the phase acquired by an harmonic oscillator. Although it is similar to the classical limit $1/\sqrt{M}$ (where M is the number of repetitions of the experiments) it has a different nature. The classical scaling arises as a consequence of adding independent information, while the scaling $1/\sqrt{N}$ simply appears because the initial state is a coherent state.

Indeed, assume that the harmonic oscillator, instead of being prepared in semi-classical state such as the coherent state, is prepared in a non-classical state that consists of a superposition of Fock states, $(1/\sqrt{2})(|0\rangle + |2N\rangle$. The quantum Fisher information of this state corresponds to $4N^2$, and therefore the scaling is improved when "quantum resources" are employed. The scaling of the error proportional to $1/N$ has been called the *Heisenberg limit* (HL) for metrology [150, 151, 152, 153]. Experimentally, the HL is harder to obtain than the SQL, since coherent states are easier to prepare.

For a coherent state, the SQL can be achieved implementing a POVM whose elements are $\hat{E}(\phi) = |\phi\rangle\langle\phi|/(2\pi)$, where $|\phi\rangle$ are the *Susskind-Glogower phase states* [127, 12], $-\pi \leqslant \phi \leqslant \pi$. These states can be expressed in terms of number states as $|\phi\rangle = \sum_{n=0}^{\infty} e^{-i\phi n}|n\rangle$. They are eigenstates of the phase operator $\hat{A} = \sum_{n=0}^{\infty} |n\rangle\langle n+1|$, which generates translations of the number operator, i.e. $\hat{A}^\dagger \hat{n} \hat{A} = \hat{n} + \mathbb{1}$. The exact probability density corresponds

to $P(\phi) = \mathrm{Tr}\Big[\hat{E}(\phi)\,\big|e^{-igt}\alpha\big\rangle\big\langle e^{-ig}\alpha\big|\Big]$,

$$\begin{aligned} P(\phi) &= \frac{1}{2\pi}\sum_{n,m}^{\infty}\Big(\frac{e^{-\frac{\alpha^2}{2}}\alpha^n}{\sqrt{n!}}\Big)\Big(\frac{e^{-\frac{\alpha^2}{2}}\alpha^m}{\sqrt{m!}}\Big)e^{i(\phi-g)(n-m)} \\ &\approx \frac{1}{2\pi}\sum_{n,m}^{\infty}\frac{1}{\sqrt{2\pi}\alpha}\exp\Big\{-\frac{1}{4\alpha^2}\Big[(\alpha^2-n)^2+(\alpha^2-m)^2\Big]+i(\phi-g)(n-m)\Big\} \end{aligned} \tag{4.45}$$

In the last step the approximation $e^{-\frac{\alpha^2}{2}}\alpha^n/\sqrt{n!} \approx (2\pi\alpha^2)^{-1/4}e^{-\frac{(\alpha^2-n)^2}{4\alpha^2}}$ was used, assuming that $\alpha \gg 1$. This probability density is normalized in the interval $[-\pi,\pi]$. Finally, by converting the summations into integrals, the density becomes

$$P(\phi) = \frac{1}{\sqrt{2\pi\Big(\frac{1}{4|\alpha|^2}\Big)}}\exp\left[-\frac{(\phi-g)^2}{2\Big(\frac{1}{4|\alpha|^2}\Big)}\right], \tag{4.46}$$

that is a Gaussian function with mean g and variance $1/(4\alpha^2)$, and is normalized over $[-\infty,\infty]$. The Fisher information in this distribution equals to $4\alpha^2$, which corresponds to the quantum Fisher information. This result shows that, when the harmonic oscillator is prepared in a coherent state, the POVM $\{\hat{E}(\phi)\}$ corresponds (approximately, for large α) to the optimal POVM.

Let us turn now to Mach-Zehnder interferometry. Figure 4.1 shows a Mach-Zehnder interferometer. The description of the interferometer requires two modes, which describe the field in the different arms of the interferometer. It is useful to employ the *Jordan-Schwinger map* [128, 129] to describe the action of the whole interferometer over the input states,

$$\begin{aligned} \hat{J}_x &= (\hat{a}^\dagger\hat{b}+\hat{a}\hat{b}^\dagger)/2, \\ \hat{J}_y &= (\hat{a}^\dagger\hat{b}-\hat{a}\hat{b}^\dagger)/(2i), \\ \hat{J}_z &= (\hat{a}^\dagger\hat{a}-\hat{b}^\dagger\hat{b})/2, \\ \hat{J}^2 &= \hat{J}_x^2+\hat{J}_y^2+\hat{J}_z^2 = \frac{\hat{N}}{2}\Big(\frac{\hat{N}}{2}+1\Big), \qquad \hat{N}=\hat{a}^\dagger\hat{a}+\hat{b}^\dagger\hat{b}, \end{aligned} \tag{4.47}$$

where $\hat{a}$ and $\hat{b}$ are the annihilation operators associated to the input fields, and $\hat{N}$ is

the total number of photons in the interferometer. The operators $\hat{J}_i$ satisfy the angular momentum algebra, i.e. $[\hat{J}_x, \hat{J}_y] = i\hat{J}_z$ and its cyclic permutations. As it is described in appendix (C) the action of a balanced beamsplitter can be described as a rotation of $\pm\pi/2$ around the x axis,

$$\hat{U}_{\pm} = \exp\Big[\pm i(\pi/2)\hat{J}_x\Big], \tag{4.48}$$

while the addition of a phase shift φ between both paths corresponds to a rotation of φ around the z-axis,

$$\hat{U}(\varphi) = \exp\Big[i\varphi\hat{J}_z\Big]. \tag{4.49}$$

Therefore, the action of the whole interferometer can be described by the unitary transformation

$$\hat{U}_{MZ} = \hat{U}_{-}\hat{U}_{\phi}\hat{U}_{+} = \exp\Big[-i\varphi\hat{J}_y\Big]. \tag{4.50}$$

Note that the interferometer (ideally) preserves the total number of photons because $[\hat{J}_y, \hat{N}] = 0$. Since the generator of the transformation is $\hat{J}_y$, the quantum Fisher information about the phase corresponds to

$$\mathcal{I}_Q(\varphi) = 4\Big\langle(\Delta\hat{J}_y)^2\Big\rangle. \tag{4.51}$$

If the input state is $|\alpha\rangle|0\rangle$, i.e. one input is a coherent beam $|\alpha\rangle$ and the second is the vacuum state $|0\rangle$, the quantum Fisher information will correspond to $|\alpha|^2$, the mean number of photons in the beam. The same result is obtained when the input state is a Fock state of the form $|N\rangle|0\rangle$, namely, $\mathcal{I}_Q(\varphi) = 4N$. Therefore, Fock or coherent states are the suitable states in order to achieve the SQL. This limit can be reached simply by counting the photons at the output ports of the interferometer and then taking the difference, as described below.

It is possible to work in the Schrödinger picture and evolve the input states along the in-

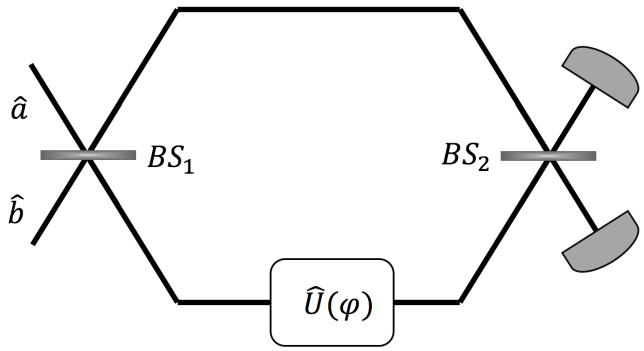


Figure 4.1: Mach-Zehnder interferometer. Two input modes ($\hat{a}$,$\hat{b}$) are mixed coherently in the first beam splitter. Inside the interferometer, a phase difference φ between both paths is introduced. The fields are merged coherently in the second beam splitter. Photons are counted at the detectors placed in the output ports. The SQL can be reached when a coherent (fock) state is sent through the input $\hat{a}$, while the second port is in the vacuum state.

terferometer using $\hat{U}_{MZ}$. Alternatively, we may work in the Heisenberg picture and evolve an input operator $\hat{O}$ using $\hat{U}^{\dagger}_{MZ}\hat{O}\hat{U}_{MZ}$. The operator $2\hat{J}_z = \hat{a}^{\dagger}\hat{a} - \hat{b}^{\dagger}\hat{b}$ corresponds the difference of photons at the input ports of the interferometer. The corresponding Heisenberg operator, $\hat{N}_{out}$, i.e. the difference of photons between the output ports, corresponds to

$$\hat{N}_{out} = 2\hat{U}^{\dagger}_{MZ}\hat{J}_z\hat{U}_{MZ} = 2\Big[\cos(\varphi)\hat{J}_z - \sin(\varphi)\hat{J}_x\Big]. \tag{4.52}$$

On the other hand, the variance of the difference of photons at the output is given by

$$(\Delta\hat{N}_{out})^2 = 4\Big[\cos^2(\varphi)(\Delta\hat{J}_z)^2 + \sin^2(\varphi)(\Delta\hat{J}_x)^2 - 2\cos(\varphi)\sin(\varphi)cov(\hat{J}_x,\hat{J}_y)\Big], \tag{4.53}$$

where $\Delta\hat{J}_i = \hat{J}_i - \Big\langle\hat{J}_i\Big\rangle$ and the covariance between $\hat{J}_x$ and $\hat{J}_y$ corresponds to $\{\hat{J}_x,\hat{J}_y\}/2 - \hat{J}_x\hat{J}_y$.

Both for a coherent state or a Fock state the mean value of $\hat{N}_{out}$ will be equal to $\cos(\varphi)N$. The bias of the estimator can be removed by normalizing $\hat{N}_{out}$ by N, and assuming that *the phase is known to be a small deviation around* $\pi/2$, namely, $\varphi = \pi/2 - \theta$, where $|\theta| \ll 1$.

For a coherent state, the variance of $\hat{N}_{out}$ corresponds to N, while for a Fock state it is $N\sin^2(\varphi)$. Taking into account the normalization factor of N and the fact that

$\sin(\varphi) \approx 1$ for small phase shifts around $\pi/2$, the variance of the estimator corresponds to $1/N$, both for coherent and Fock states. Consequently, when coherent or Fock states are employed, $\hat{N}_{out}/N$ is an unbiased and efficient estimator of small phase shifts around $\pi/2$, that achieves the SQL.

4.8 Relationship between QFI and the Bures distance

Consider an experiment with D possible outcomes and let $p_i = P(r_i|g)$ be the probability of the i-th outcome to occur, for $i = 1, ..., D$. For simplicity, we have chosen the probability to be discrete, but the argument that follows may be generalized to a continuous distribution.

If the parameter of the likelihood function $P(r|g)$ is changed infinitesimally, from g to $g + dg$, then the probabilities of the outcomes will be modified from p_i to $p_i + dp_i$, where $dp_i = \partial_g P(r_i|g)dg$. This change can be viewed as an infinitesimal displacement along a curve (parametrized by g) that "lives" in a statistical manifold of dimension $D-1$; we start from a point in the manifold corresponding to $P(r|g)$ and, by changing infinitesimally the parameter, we reach a second point that corresponds to $P(r|g + dg)$. Figure 4.2 illustrates this displacement for a distribution with three outcomes.

The classical Fisher information (4.15) is a metric on this manifold, and allows therefore to define the length of a line element along any curve, namely, $ds^2 = \mathcal{I}(g)dg^2$. Hence, the larger the Fisher information, the larger the distance between $P(r|g)$ and $P(r|g + dg)$. As explained in section (4.4), the Fisher information is associated to the precision of an estimator of g. Consequently, the notion of distance in a statistical manifold is associated to the notion of distinguishability, because the better the precision (the larger the Fisher information), the better the distributions $P(r|g)$ and $P(r|g + dg)$ can be resolved.

Now, consider two quantum states, $\rho(g)$ and $\rho(g+dg)$. Implementing a certain POVM, $\{\hat{E}_r(g)\}$ (notice the possible dependence on g), will map each state to a point on the statistical manifold, namely, to a probability distribution; $P(r|g) = \mathrm{Tr}\Big[\hat{E}_r(g)\rho(g)\Big]$ and $P(r|g + dg) = \mathrm{Tr}\Big[\hat{E}_r(g + dg)\rho(g + dg)\Big]$. The distance between both probabilities is defined by the classical Fisher information. Choosing the POVM that maximices the classical

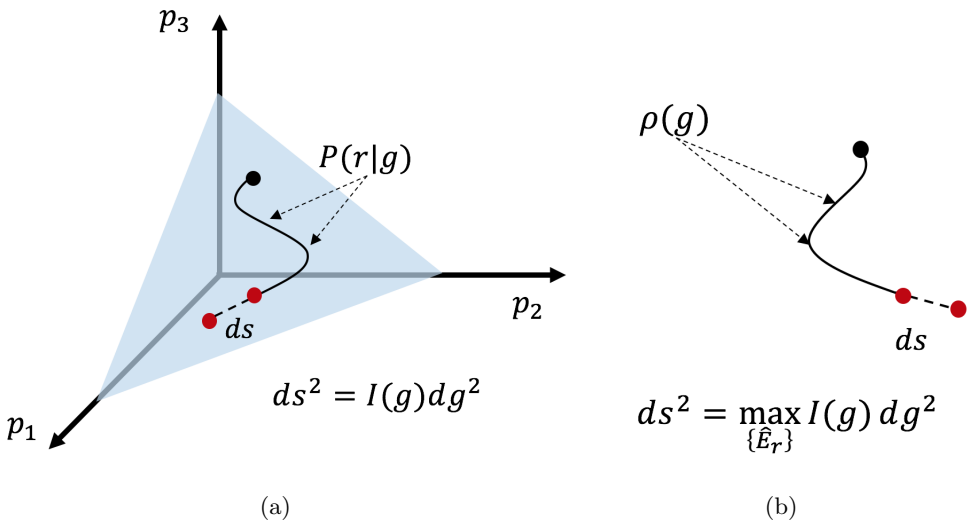


Figure 4.2: a) Infinitesimal displacement in the statistical manifold. The distance between two neighbouring distributions is obtained using the Fisher metric. b) Infinitesimal displacement in the manifold of density operators. The distance depends on the quantum Fisher metric, which is equal to the Bures metric (except for a factor of 1/4).

Fisher information, will maximice the distance (the distinguishability) between both quantum states. Therefore, the quantum Fisher information defines a metric in the manifold of density operators, namely, $ds^2 = \mathcal{I}_Q(g)dg^2$. The larger the quantum Fisher information, the larger the distance between $\rho(g)$ and $\rho(g+dg)$ and the better both states can be distinguished. See figure 4.2 for a representation of these concepts. The definition of a statistical distance was presented in [130] for pure states and in [124] for mixed states.

On the other hand, the *Bures distance* [131, 132, 133] between two quantum states ρ_1 and ρ_2 is defined as $d_B(\rho_1,\rho_2) = 2[1-\sqrt{F(\rho_1,\rho_2)}]$, where $F(\rho_1,\rho_2) = \mathrm{Tr}\big[\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\big]^2$ is the *quantum fidelity* between the states. When the states are parametrized by a single parameter g, then

$$ds^2 = d_B[\rho(g),\rho(g+dg)]^2 \tag{4.54}$$

$$= \left[\frac{1}{4}\mathcal{I}_Q(g)\right]dg^2, \tag{4.55}$$

which shows that, except for a factor of 1/4, the quantum Fisher metric equals the Bures metric on the manifold of quantum states [124, 134].

4.9 Fisher Information in weak measurements

In this section the Fisher information contained in weak measurements is analyzed. Two measurement strategies will be distinguished; with and without post-selection. See chapters 2 and 3 for a description of both measurement strategies. We consider weak measurements of an observable $\hat{A}$, according to the von Neumann model.

Before taking each strategy into consideration, it is important to point out that the maximum information about the parameter g corresponds to the quantum Fisher information. In the von Neumann model, the measurement is described by a unitary process $\hat{U}$ with generator equal to $\hat{A}\hat{P}/\hbar$. Therefore, for initial pure states states of the meter and the system, $|\psi\rangle\langle\psi|\otimes|\psi_i\rangle\langle\psi_i|$, the quantum Fisher information is obtained by a straightforward application of expression (4.42),

$$\begin{aligned}\mathcal{I}_Q(g) &= 4\left\langle \hat{A}^2\right\rangle\left\langle \hat{P}^2\right\rangle/\hbar^2 - 4\left\langle \hat{A}\right\rangle^2\left\langle \hat{P}\right\rangle^2/\hbar^2 \\ &= \frac{4}{\hbar^2}\Big[Var(\hat{A})\left\langle \hat{P}^2\right\rangle + \langle A\rangle^2 Var(\hat{P})\Big]. \end{aligned} \tag{4.56}$$

In the last expression, the term $Var[\hat{O}]$ denotes the variance of the operator $\hat{O}$ calculated in the initial state, and is also sometimes denoted by $\left\langle(\Delta\hat{O})^2\right\rangle$ along this text. Expression (4.56) represents the maximum precision that can be achieved, and hence the different measurement strategies will be compared with this limit.

4.9.1 Weak measurements without post-selection

When no post-selection is performed, i.e. when the apparatus variable $\hat{R}$ is measured and no second measurement is performed, the measurement and effect operators ($\hat{M}_r$ and $\hat{E}_r$, respectively) are given by (2.47). The probability density function $P(r) = \mathrm{Tr}\Big[\hat{E}_r\rho\Big]$ corresponds to

$$P(r) = \sum_{n=0}^{\infty}\frac{(ig/\hbar)^n}{n!}\left\langle \hat{A}^n\right\rangle\sum_{m=0}^{n}(-1)^m\binom{n}{m}\langle\psi|\,\hat{P}^{n-m}\hat{\Pi}_r\hat{P}^m\,|\psi\rangle\,, \tag{4.57}$$

where the measurement result r corresponds to one of the eigenvalues of the meter variable $\hat{R} = \int dr r \hat{\Pi}_r$. In the linear regime, the probability density (2.49) can be expressed as

$$P(r) = |\psi(r)|^2 \left\{ 1 + 2(g/\hbar) \left\langle \hat{A} \right\rangle \mathrm{Im}(P_w) \right\}, \tag{4.58}$$

where $P_w = \langle r| \hat{P} |\psi\rangle / (\langle r|\psi\rangle)$ is the weak value of the momentum operator of the meter. The regime in which this approximation can be made depend on the moments of $\hat{A}$ (thus, on the initial state of the system), and on the weak values $P_w^k = \langle r| \hat{P}^k |\psi\rangle / (\langle r|\psi\rangle)$ (which depend on $\hat{R}$ and on the initial state of the meter). In particular, later in this same section, the validity of the linear regime is analyzed for the case $\hat{R} = \hat{Q}$ ($\hat{Q}$ being the position operator), assuming that $\hat{A}$ is a dichotomic variable, and when the initial state of the meter is Gaussian.

On the basis of (4.58), and retaining the zero order contribution (since g is a small parameter), the *score function* $\mathcal{S}(g) = \partial_g \ln[P(r|g)]$ corresponds to

$$\mathcal{S}(g) = (2/\hbar) \left\langle \hat{A} \right\rangle \mathrm{Im}(P_w). \tag{4.59}$$

The variance of the score, calculated using (4.58), gives the Fisher information of a weak measurement (without post-selection),

$$\mathcal{I}(g) = (4/\hbar^2) \left\langle \hat{A} \right\rangle^2 \int \mathrm{Im}(P_w)^2 |\psi(r)|^2 dr. \tag{4.60}$$

This expression represents the maximum precision that can be achieved by observing the variable $\hat{R}$. The integral term can be understood as the variance of the imaginary part of the weak value of $\hat{P}$ (since the first moment is zero). An important result [135] regarding the variance of the imaginary and real parts of a weak value establishes that

$$\int |\psi(r)|^2 \Big[\mathrm{Re}(P_w) - \left\langle \hat{P} \right\rangle \Big]^2 dr + \int \mathrm{Im}(P_w)^2 |\psi(r)|^2 dr = Var(\hat{P}). \tag{4.61}$$

Consequently, the Fisher information (4.60) can be expressed as

$$\mathcal{I}(g) = \frac{4}{\hbar^2}\left\langle \hat{A}\right\rangle^2 \Big\{ Var(\hat{P}) - \int |\psi(r)|^2 \Big[\mathrm{Re}(P_w) - \left\langle \hat{P}\right\rangle \Big]^2 dr \Big\}, \tag{4.62}$$

which shows that, when the variance of the real part of the weak value of $\hat{P}$ is zero, the Fisher information achieves its maximum value, $(4/\hbar^2)\left\langle \hat{A}\right\rangle^2 Var(\hat{P})$. This value should be contrasted with the maximum precision (4.56). The difference corresponds to $Var(\hat{A})\left\langle \hat{P}^2\right\rangle$. Therefore, unless the the system starts in an eigenstate of $\hat{A}$, the quantum Fisher information can not be achieved with this strategy.

As an example, let us consider the case when the meter variable corresponds to the position, $\hat{R} = \hat{Q}$. Consider also the polar decomposition of the initial wave function of the meter, $\psi(q) = |\psi(q)| \exp\{i\Phi(q)/\hbar\}$. The weak value of the momentum operator of the meter is given by

$$P_w = \frac{\langle q|\, \hat{P}\, |\psi\rangle}{\langle q|\psi\rangle} = -i\hbar \frac{\partial_q \psi(q)}{\psi(q)} = -i\hbar \frac{\partial_q |\psi(q)|}{|\psi(q)|} + \partial_q \Phi(q), \tag{4.63}$$

which shows that the *phase gradient* corresponds to the real part of P_w, while its imaginary part is the logarithmic derivative with respect to q of the modulus of the initial wave function of the meter. Therefore, the maximum value of the Fisher information can be achieved when the gradient of the phase is equal to $\left\langle \hat{P}\right\rangle$. As an additional remark, it is worth to mention that the phase gradient has played a role in different formulations of quantum mechanics [136, 137, 138, 139].

Assume, for example, that the initial wave function of the meter corresponds to a complex Gaussian function with a quadratic phase,

$$\psi(q) = N_q \exp\left\{ (i\mathcal{F} - 1)\frac{(q-q_0)^2}{4\sigma_q^2} + i\left\langle \hat{P}\right\rangle q/\hbar \right\} \quad , \quad N_q = \left(\frac{1}{\sqrt{2\pi}\sigma_q} \right)^{1/2}. \tag{4.64}$$

For Gaussian states with quadratic phase the uncertainty principle is $\sigma_p^2 \sigma_p^2 = (\hbar^2/4)(1+\mathcal{F}^2)$.

Hence, the Fisher information can be expressed as

$$\mathcal{I}(g) = \frac{\left\langle \hat{A} \right\rangle^2}{\sigma_q^2} = \Big(\frac{4}{\hbar^2}\Big) \left\langle \hat{A} \right\rangle^2 Var(\hat{P}) \Big(\frac{1}{1+\mathcal{F}^2}\Big). \tag{4.65}$$

Expression (4.65) represents the Fisher information in a weak measurement, when no post-selection is performed, the meter starts in a state described by (4.64), and a projective measurement of the joint operator $\hat{Q} \otimes \mathbb{1}$ is performed. This result shows that, under these conditions, it is preferable that the quadratic phase factor disappears.

It is worth to mention that, when the momentum of the meter is observed ($\hat{R} = \hat{P}$), then the weak value of the momentum operator is purely real and therefore the information (4.60) is zero, as expected.

Linear regime for Gaussian states and $\hat{R} = \hat{Q}$

Finally, let us study the regime in which the approximation (4.58) holds. When the position of the meter is observed, the general expression of the density function (4.57) equals (2.33). For the family of wave functions described by (4.64), this probability density function can be expressed as

$$P(q) = |\psi(q)|^2 \left\langle \exp\left\{ z\Big(\frac{g}{\sigma_q}\Big)\hat{A} - \frac{1}{2}\Big(\frac{g}{\sigma}\Big)\hat{A}^2 \right\} \right\rangle, \tag{4.66}$$

where $z = (q - q_0)/\sigma_q$ is a standardized variable. Notice that the initial momentum of the meter and the quadratic phase factor do not appear in the density function. The expression inside the brackets can be expanded in terms of the *Hermite* polynomials [140],

$$\begin{aligned} P(q) &= |\psi(q)|^2 \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\frac{g}{\sigma_q}\Big)^n \left\langle \hat{A}^n \right\rangle He_n(z) && (4.67) \\ &= |\psi(q)|^2 \left[1 + \Big(\frac{g}{\sigma_q}\Big) \left\langle \hat{A} \right\rangle z + \frac{1}{2}\Big(\frac{g}{\sigma_q}\Big)^2 \left\langle \hat{A}^2 \right\rangle (z^2 - 1) + ... \right]. && (4.68) \end{aligned}$$

Note that the Hermite series $S(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\frac{g}{\sigma_q}\Big)^n \left\langle \hat{A}^n \right\rangle He_n(z)$ (from the first order term onwards) corresponds to a correction to $|\psi(q)|^2$, the probability of finding the measurement

device at position q in the absence of any interaction with the system. The first order approximation will be a "good" approximation when itself has greater magnitude than the sum of the neglected higher order terms, namely,

$$\left| \sum_{n=2}^{\infty} \frac{1}{n!} \Big(\frac{g}{\sigma_q} \Big)^n \left\langle \hat{A}^n \right\rangle He_n(z) \right| \ll \left| 1 + \frac{g}{\sigma_q} \left\langle \hat{A} \right\rangle z \right|. \tag{4.69}$$

For the values of z that satisfy this condition, the system will exhibit a *linear response*, i.e. the correction $S(z) - 1$ will be a linear function of g/σ_q. Since the initial wave function of the meter $|\psi(q)|^2$ is Gaussian with standard deviation σ_q, what is actually important is that the linear response dominates over the rest of the terms within some standard deviations away from the media.

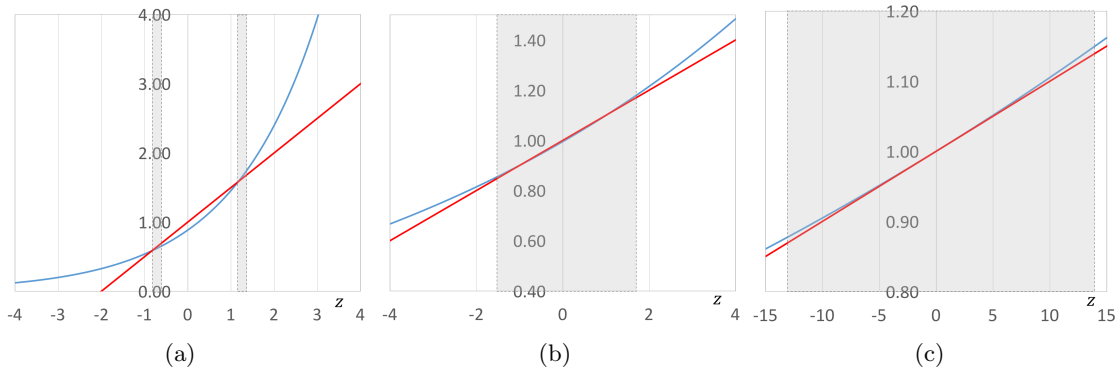
Consider, as an example, an operator that satisfies $\hat{A}^2 = \mathbb{1}$, i.e. that it is an *involution* and thus $-1 \leqslant \left\langle \hat{A} \right\rangle \leqslant 1$. In this case, the exact probability (4.67) becomes

$$P(q) = |\psi(q)|^2 \exp\left\{ -\frac{1}{2} \Big(\frac{g}{\sigma_q} \Big)^2 \right\} \Big\{ \cosh(gz/\sigma_q) + \left\langle \hat{A} \right\rangle \sinh(gz/\sigma_q) \Big\}, \tag{4.70}$$

that shows that the linear response holds when $|gz/\sigma_q| \ll 1$. The complete response $S(z)$ and the linear response are plotted in figure 4.3 for different values of g/σ_q. The exact probability and its first order approximation are presented in figure 4.4. Notice that, for $g/\sigma_q = 0.5$, the approximation, although normalized, takes negatives values in some part of the domain. This occurs because the region in which the approximation is valid (i.e. the relative error is less than 1%) is small. However, for $g/\sigma = 0.01$ the approximation is good in a large region of its domain. As a result, the approximation and the exact probability are almost indistinguishable.

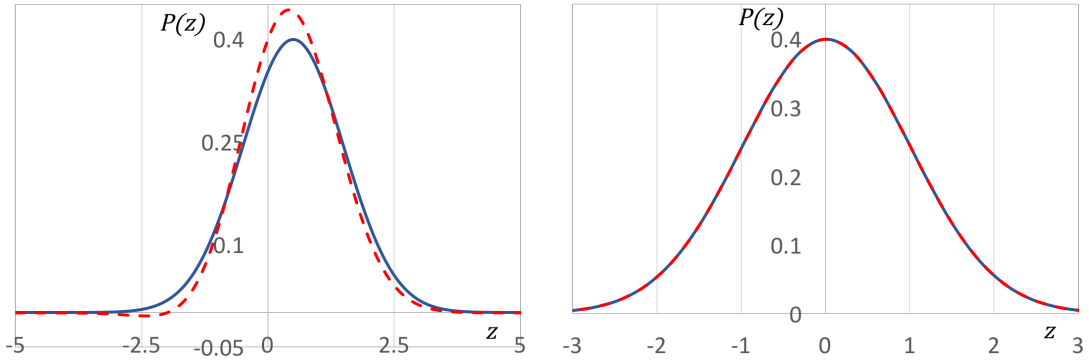
4.9.2 Fisher Information in weak measurements with post-selection

In section 3.1 we described pre- and post-selected measurements, where an intermediate generalized measurement of an observable $\hat{A}$ is performed between a pre-selected and a post-selected final state. The intermediate measurement was described by measurement operators $\hat{M}_r$ *acting on the system*, which result from the observation of the meter variable



(a) (b) (c)

Figure 4.3: The complete Hermite series $S(z)$ (blue) and the linear response (red) are plotted for different values of the parameter g/σ_q, assuming $\left\langle \hat{A} \right\rangle = 1$. The shadowed area represents the region where the relative error of the approximation is less than 1%. (a) $g/\sigma_q = 0.5$, (b) $g/\sigma_q = 0.1$, and (c) $g/\sigma_q = 0.01$.



(a) (b)

Figure 4.4: Exact probability (solid line) and the first order approximation (dashed line) for different values of the ratio g/σ_q: a) 0.5, (b) 0.01.

$\hat{R} = \int dr \hat{\Pi}_r$, a continuous variable *of the measurement device.*

The post-selection procedure corresponds to a second (projective) measurement of an hermitian operator $\hat{F}$, of the system, with eigenvectors and eigenvalues described by $\hat{F}\left|\psi_k\right\rangle = \lambda_k \left|\psi_k\right\rangle$, $k = 1..N$. When post-selection is performed not all results of the intermediate measurement are taken into consideration, but only those when the state $\left|\psi_f\right\rangle$ is detected in the second measurement. Therefore, the set of measurement operators that describe this measurement strategy [141] corresponds to

$$\left\{ \left|\psi_f\right\rangle\!\left\langle\psi_f\right| \otimes \hat{\Pi}_r, \left(\mathbb{1} - \left|\psi_f\right\rangle\!\left\langle\psi_f\right|\right) \otimes \mathbb{1}, r \in \mathbb{R} \right\}. \tag{4.71}$$

The probability to successfully select the state $\left|\psi_f\right\rangle$ in the second measurement *and* to read the outcome r in the intermediate measurement is given by

$$P(f,r) = \mathrm{Tr}\left[\left|\psi_f\right\rangle\!\left\langle\psi_f\right| \otimes \hat{\Pi}_r \hat{U} \left|\psi\right\rangle\!\left\langle\psi\right| \rho \hat{U}^\dagger \right] = \left\langle\psi_f\right| \hat{M}_r \rho \hat{M}_r^\dagger \left|\psi_f\right\rangle, \tag{4.72}$$

where $\left|\psi\right\rangle$ is the pure initial state of the meter and $\hat{U}$ is the unitary operator that describes the intermediate measurement. The probability of successful post-selection $P(f)$ is obtained by marginalizing (4.72) over r. The probability of failure corresponds to $1 - P(f)$. The classical Fisher information contained in the probability distribution (4.72) *and* $1 - P(f)$ is obtained by applying the definition (4.15),

$$\begin{aligned} \mathcal{I}(g) &= \int dr \frac{[\partial_g P(r,f)]^2}{P(r,f)} + \frac{\{\partial_g [1 - P(f)]\}^2}{1 - P(f)} \\ &= P(f) \int dr \frac{[\partial_g P(r|f)]^2}{P(r|f)} + \frac{[\partial_g P(f)]^2}{P(f)} + \frac{\{\partial_g [1 - P(f)]\}^2}{1 - P(f)}. \end{aligned} \tag{4.73}$$

The integral term in the last line represents the Fisher information *given* that the state $\left|\psi_f\right\rangle$ was post-selected, and will be denoted as $\mathcal{I}(g|f)$. The sum of the second and third terms correspond to the information contained in the partition into different sub-ensembles, according to to the result of the second measurement. This information will be denoted as

$\mathcal{I}_P$. Therefore, the Fisher information in the whole process can be expressed as

$$\mathcal{I}(g) = P(f)\mathcal{I}(g|f) + \mathcal{I}_P. \tag{4.74}$$

Therefore, the maximum amount of information provided by the post-selection strategy corresponds to

$$\mathcal{I}_{Max}(g) = P(f)\mathcal{I}_Q(g|f) + \mathcal{I}_P, \tag{4.75}$$

where $\mathcal{I}_Q(g|f)$ corresponds now to the *quantum Fisher information* in the final state of the measurement device, *given that the post-selection was successful*. If the initial state of the meter is pure, let us denote it as $|\psi\rangle$, then its final state, conditioned on the successful post-selection, corresponds to (3.55)[3]. To first order in g, the final state of the meter corresponds

$$|\phi\rangle = \frac{1}{\sqrt{N}}\Big[1 - (ig/\hbar)A_w\hat{P}\Big]\,|\psi\rangle\,, \tag{4.76}$$

where $N = 1 + 2gk_0\,\mathrm{Im}(A_w)$ is a normalization factor due to the post-selection. This is a pure state and therefore expression (4.40) can be employed to compute the quantum Fisher information,

$$\mathcal{I}_Q(g|f) = \Big(\frac{4}{\hbar^2}\Big)|A_w|^2 Var(\hat{P}). \tag{4.77}$$

On the other hand, from (3.21) the probability of post-selection corresponds to $P(f) = |\langle\psi_f|\psi_i\rangle|^2[1+2gk_0\,\mathrm{Im}(P_w)]$. It contains therefore information about g, as long as the weak value is not purely real and the meter has initial momentum. Thus, for a balanced meter (zero initial momentum) or for protocols based on real weak values, this term plays no role.

Taking all the contributions into account, the maximum amount of information (4.75)

[3]This state corresponds to the case $\hat{R} = \hat{P}$. For the general case, $\hat{M}_p$ should be simply replaced by $\hat{M}_r$

that can be obtained using the post-selection strategy is given by

$$\mathcal{I}_{Max}(g) = \frac{4|\langle\psi_f|\psi_i\rangle|^2}{\hbar^2}\left\{|A_w|^2\left\langle\hat{P}^2\right\rangle - \left\langle\hat{P}\right\rangle^2\left[\mathrm{Re}(A_w)^2 - \frac{|\langle\psi_f|\psi_i\rangle|^2}{1-|\langle\psi_f|\psi_i\rangle|^2}\,\mathrm{Im}(A_w)^2\right]\right\} \quad (4.78)$$

Notice that, if the final state $|\psi_f\rangle$ is the same as the initial state $|\psi_i\rangle$, then $\mathrm{Im}(A_w) = 0$, $\mathrm{Re}(A_w) = \left\langle\hat{A}\right\rangle$ and the expression reduces to maximum information attainable without post-selection.

By comparing expression (4.78) to the quantum Fisher information (4.56), it is clear that, in order for $\mathcal{I}_{Max}(g)$ to reach the maximum precision, the following two conditions will be needed:

$$i) \qquad |\langle\psi_f|\psi_i\rangle|^2\,\mathrm{Re}(A_w)^2 + |\langle\psi_f|\psi_i\rangle|^2\,\mathrm{Im}(A_w)^2 = \left\langle\hat{A}^2\right\rangle. \quad (4.79)$$

$$ii) \qquad |\langle\psi_f|\psi_i\rangle|^2\,\mathrm{Re}(A_w)^2 - \left(\frac{|\langle\psi_f|\psi_i\rangle|^4}{1-|\langle\psi_f|\psi_i\rangle|^2}\right)\mathrm{Im}(A_w)^2 = \left\langle\hat{A}\right\rangle^2. \quad (4.80)$$

The first condition should be fulfilled for every initial pure state of the meter. Both equations need to be satisfied only for meter states with initial momentum, $\left\langle\hat{P}\right\rangle \neq 0$. It is more informative to express this set of equations as

$$|\langle\psi_f|\,\hat{A}\,|\psi_i\rangle|^2 = \left\langle\hat{A}^2\right\rangle, \quad (4.81)$$

$$|\langle\psi_f|\psi_i\rangle|^2\Big[\mathrm{Re}(A_w)^2 - |\langle\psi_f|\,\hat{A}\,|\psi_i\rangle|^2\Big] = \left\langle\hat{A}\right\rangle^2\Big(1-|\langle\psi_f|\psi_i\rangle|^2\Big). \quad (4.82)$$

First, if the meter is balanced $\left\langle\hat{P}\right\rangle = 0$, then we only have to worry about the first equation (4.81). In this situation, the post-selected state should be simply chosen in order to exactly satisfy (4.81). There is, however, a way to satisfy this equation in the limit $|\langle\psi_f|\psi_i\rangle|^2 \to 0$. A standard theorem of linear algebra (shown, for example, in [143]), establishes that

$$\hat{A}\,|\psi_i\rangle = \left\langle\hat{A}\right\rangle|\psi_i\rangle + \sqrt{\left\langle(\Delta\hat{A})^2\right\rangle}\,|\psi_i^{\perp}\rangle, \quad (4.83)$$

where $\left|\psi_i^{\perp}\right\rangle$ is any state orthogonal to the initial state. Assume that quasi-orthogonal state

$$|\psi_f\rangle = \left|\psi_i^{\perp}\right\rangle + \epsilon\,|\psi_i\rangle, \quad (4.84)$$

is post-selected, where ϵ is a small quantity, $\epsilon \ll 1$. By virtue of (4.83), and retaining up to first order terms of ϵ, condition (4.81) can be expressed as

$$\left\langle \hat{A} \right\rangle \left[\left\langle \hat{A} \right\rangle - 2\sqrt{\left\langle (\Delta \hat{A})^2 \right\rangle} \cdot \epsilon \right] = 0. \tag{4.85}$$

Consequently, when $\left\langle \hat{A} \right\rangle \sim \epsilon$ the left hand side will be of the order of ϵ^2 and the equation will be approximately satisfied. In simpler words, when $\left\langle \hat{A} \right\rangle$ is close to zero, the post-selection of quasi-orthogonal states will allow to satisfy the first condition in the limit $|\langle \psi_f | \psi_i \rangle|^2 \to 0$.

Assume now that the meter has initial momentum, in which case both equations should be fulfilled. When the weak value A_w is *purely real* or *purely imaginary*, it is easy to show that conditions (4.79) and (4.80) can not be satisfied, unless the system starts in an eigenstate of $\hat{A}$ (in which case both conditions are fulfilled and the weak value is real). Thus, besides the case in which the system starts in an eigenstate of the measured variable, the quantum Fisher information can not be exactly reached by using purely real or purely imaginary weak values.

Nevertheless, analogously to the previous case (balanced meter), both conditions can be met in the limit $|\langle \psi_f | \psi_i \rangle|^2 \to 0$. We have already shown that the first condition will be approximately satisfied when the initial and final states are quasi-orthogonal and the expectation value of $\hat{A}$ is close to zero (of the same order of overlap between the initial and final states). Let us study now what happens with the second condition when a state of the form (4.84) is post-selected. Using (4.83) it is easy to show that the second equation (4.80) reduces to

$$\epsilon^2 \cdot \left\{ \mathrm{Re}(A_w)^2 - \left\langle \hat{A}^2 \right\rangle + \left\langle \hat{A} \right\rangle \left[\left\langle \hat{A} \right\rangle - 2\sqrt{\left\langle (\Delta \hat{A})^2 \right\rangle} \cdot \epsilon \right] \right\} = \left\langle \hat{A} \right\rangle^2 (1 - \epsilon^2). \tag{4.86}$$

Notice that when $\left\langle \hat{A} \right\rangle \sim \epsilon$, then both sides of the equation will be of second order in ϵ only when the weak value is *purely imaginary* (notice that the real weak value in the left hand side of the equation is of the order of ϵ^{-2}). Therefore, both equations can be approximately satisfied when $\left\langle \hat{A} \right\rangle$ is close to zero, by using quasi-orthogonal states and

Table 4.1: The post-selection strategy may be optimal (in the sense that it achieves exactly the quantum Fisher information) using real or imaginary weak values. If the meter is prepared with zero initial momentum (balanced meter), then only the first condition (4.81) should be met. If the meter has momentum (not balanced meter), then the second condition (4.82) should be also taken into consideration. The strategies in which the states are close to orthogonality achieve the quantum information in the limit $|\langle\psi_f|\psi_i\rangle|^2 = \epsilon \to 0$. Recall, however, that the weak amplification effect is restricted by the validity of linear approximation (4.87) and the states can not become indefinitely close to orthogonality.

Strategy	Balanced	Not balanced
Exact achievement of QFI	$\lvert\langle\psi_f\rvert \hat{A} \lvert\psi_i\rangle\rvert^2 = \left\langle \hat{A}^2 \right\rangle$	The systems starts in in an eigenstate of $\hat{A}$
QFI is achieved in the limit $\lvert\langle\psi_f\rvert\psi_i\rangle\rvert^2 = \epsilon \to 0$	$\left\langle \hat{A} \right\rangle \sim \epsilon$	$\left\langle \hat{A} \right\rangle \sim \epsilon$ and imaginary weak values

imaginary weak values.

The previous analysis shows that, *when the meter is balanced*, then the quantum Fisher information can be *exactly* achieved by post-selecting a state such that $|\langle\psi_f| \hat{A} |\psi_i\rangle|^2 = \left\langle \hat{A}^2 \right\rangle$, keeping in mind the the final state should not be exactly orthogonal to the initial state. In the *general case*, either when the meter is balanced or not, then the quantum Fisher information can be achieved in the limit $|\langle\psi_f|\psi_i\rangle|^2\epsilon \to 0$, when $\left\langle \hat{A} \right\rangle \sim \epsilon$. If $\left\langle \hat{P} \right\rangle \neq 0$ then imaginary weak values are needed. The whole, already described, strategy is also summarized in table 4.1.

Fisher Information in the conditioned probability

Now, we will consider the Fisher information $\mathcal{I}(g|f)$ contained in the *conditioned* probability density function $P(r|f)$, i.e. assuming that the post-selection of the state $|\psi_f\rangle$ was successful. The exact density function was derived in (3.16). The first order expansion was presented in (3.23), and can be expressed as

$$P(r|f) = |\psi(r)|^2 \left\{ 1 + \frac{2g}{\hbar} \operatorname{Im}\left[A_w \cdot \left(P_w - \left\langle \hat{P} \right\rangle \right) \right] \right\}. \tag{4.87}$$

Now, the score function $\partial_g P(r|f)$ becomes

$$\mathcal{S}(g) = \frac{2}{\hbar}\,\mathrm{Im}\Big[A_w \cdot \Big(P_w - \Big\langle \hat{P} \Big\rangle\Big)\Big]. \tag{4.88}$$

Using (4.61), the Fisher information can be expressed as

$$\mathcal{I}(g|f) = \mathcal{I}_Q(g|f) - \Big(\frac{4}{\hbar^2}\Big)\int |\psi(r)|^2 \Bigg\{ \mathrm{Re}(A_w)\Bigg[\,\mathrm{Re}(P_w) - \Big\langle \hat{P} \Big\rangle\,\Bigg] - \mathrm{Im}(A_w)\,\mathrm{Im}(P_w)\Bigg\}^2 \tag{4.89}$$

The first term, $\mathcal{I}_Q(g|f)$, corresponds to the quantum Fisher information (4.77) and can be achieved by making the second term equal to zero.

What happens when the weak value A_w is purely real? In this case, it is clear that the second term of (4.89) will vanish when $\mathrm{Re}(P_w) = \Big\langle \hat{P} \Big\rangle$. Thus, for a balanced meter (zero initial momentum), when real weak values values are employed, *the weak value of the meter should be purely imaginary*, in order to achieve the quantum Fisher information of the conditioned state. As an example, consider the case when the position of the meter is observed. In this scenario, according to (4.63), the real and imaginary parts of $\hat{P}$ are given by

$$\mathrm{Re}(P_w) = \partial_q \phi(q), \qquad \mathrm{Im}(P_w) = -\hbar \frac{\partial_q |\psi(q)|}{|\psi(q)|}. \tag{4.90}$$

Hence, when the position of the meter is observed, preparing the meter in a state with a linear phase will allow to achieve the quantum information of the conditioned state (4.77).

What happens if the weak value A_w is purely imaginary? In this case, the second term of (4.89) will disappear when $\mathrm{Im}(P_w) = 0$ for all the possible values of r. It is easy to see that this occurs when the momentum of the meter is observed,

$$P_w = \frac{\langle p|\,\hat{P}\,|\psi\rangle}{\langle p|\psi\rangle} = p. \tag{4.91}$$

In this case, the weak value of the meter becomes real and the quantum Fisher information in the conditioned state can be achieved.

Finally, for *complex weak values*, the weak value associated with the apparatus variable

should also be complex. In fact, it is easy to see that the second term of (4.89) will vanish when

$$\frac{\mathrm{Re}(P_w) - \left\langle \hat{P} \right\rangle}{\mathrm{Im}(P_w)} = \frac{\mathrm{Im}(A_w)}{\mathrm{Re}(A_w)}, \tag{4.92}$$

for every possible value of r. As an example, let us consider again the case of a meter prepared in a Gaussian state with a quadratic phase (4.64). Physically, this state can be achieved when an harmonic oscillator is prepared in a *coherent squeezed state*, $|\psi\rangle = D(\alpha)S(\beta)\,|0\rangle$, where $D(\alpha)$ is a *Glauber displacement operator* and $S(\beta)$ is a *squeezing operator*. The displacement is given by $\alpha \in \mathbb{C}$ and the squeezing factor is $\beta = r\exp\{i\theta\}$. The coordinate representation of the wave function is given by

$$\psi(q) = \exp\left\{(i\mathcal{F}-1)\frac{[q-2x_0\,\mathrm{Re}(\alpha)]^2}{4\sigma_q^2} + \frac{i}{\hbar}\left\langle \hat{P} \right\rangle q\right\}, \tag{4.93}$$

where x_0 are the zero-point fluctuations of the harmonic oscillator and $\sigma_q^2 = x_0^2[e^{2r}\sin^2(\theta/2)+ e^{-2r}\cos^2(\theta/2)]$. The initial momentum of the meter $\left\langle \hat{P} \right\rangle$ corresponds to $(\hbar/x_0)\,\mathrm{Im}(\alpha)$. The adimensional parameter $\mathcal{F}$ (the coefficient associated to the quadratic phase) depends entirely on the squeezing factor and corresponds to $-\sinh(2r)\sin(\theta)$. This coefficient represents the covariance between the position and momentum of the meter (besides a factor of 1/2),

$$Cov(\hat{Q},\hat{P}) = \frac{1}{2}\left\langle \{\hat{Q},\hat{P}\} \right\rangle - \left\langle \hat{Q} \right\rangle\left\langle \hat{P} \right\rangle = \hbar\mathcal{F}/2. \tag{4.94}$$

When the position of the meter is observed, $\mathrm{Im}(P_w) = \hbar(q-q_0)/(2\sigma_q^2)$, and $\mathrm{Re}(P_w) - \left\langle \hat{P} \right\rangle = \mathcal{F}\,\mathrm{Im}(P_w)$. Therefore, the correlation between the position and the momentum of the meter (calculated in the initial state) should be equal to one half the tangent of the argument of the weak value of $\hat{A}$, namely,

$$\frac{1}{\hbar}Cov(\hat{Q},\hat{P}) = \left(\frac{1}{2}\right)\frac{\mathrm{Im}(A_w)}{\mathrm{Re}(A_w)}. \tag{4.95}$$

It should be noted that the effect of the squeezing establishes a correlation between the

position and the momentum of the meter. This correlation is necessary for protocols based on complex or imaginary weak values; otherwise, without applying squeezing, real weak values of $\hat{A}$ should be employed.

Finally, consider the case when the momentum of the meter is observed, $\hat{R} = \hat{P}$. The weak value of the momentum operator $P_w = \langle p| \hat{P} |\psi\rangle / (\langle p|\psi\rangle)$ is equal to p and the Fisher information $\mathcal{I}(g|f)$ becomes

$$\mathcal{I}(g|f) = (4/\hbar^2) \operatorname{Im}(A_w)^2 \sigma_p^2. \tag{4.96}$$

The momentum of the meter contains useful information about the parameter, unlike the situation without post-selection. In this case, as was previously explained, the weak value should be purely imaginary in order to achieve the quantum Fisher information (4.77).

Optimal post-selection strategy for a qubit and a Gaussian meter

As an example, let us analyse the case of a spin-like operator, $\hat{A} = \hat{\sigma}_z$. The meter is prepared in a Gaussian state described by (4.93) and the system is a two-level system, prepared in an initial pure state. The post-selected state is also a pure state. For a qubit, the initial and final states pure states can be described using the Bloch representation,

$$|\psi_i\rangle = \cos\left(\frac{\theta_i}{2}\right)|0\rangle + \sin\left(\frac{\theta_i}{2}\right) e^{i\phi_i} |1\rangle, \tag{4.97}$$

$$|\psi_f\rangle = \cos\left(\frac{\theta_f}{2}\right)|0\rangle + \sin\left(\frac{\theta_f}{2}\right) e^{i\phi_f} |1\rangle, \tag{4.98}$$

where $\hat{\sigma}_z |0\rangle = |0\rangle$ and $\hat{\sigma}_z |1\rangle = -|1\rangle$. The weak value of $\hat{\sigma}_z$ between the initial and final states is defined by

$$\sigma_{z,w} = \frac{\langle \psi_f| \hat{\sigma}_z |\psi_i\rangle}{\langle \psi_f|\psi_i\rangle}, \tag{4.99}$$

while the real and imaginary parts of the weak value, respectively, are given by

$$\mathrm{Re}(\sigma_{z,w}) = \frac{\cos\left(\frac{\theta_i+\theta_f}{2}\right)\cos\left(\frac{\theta_i-\theta_f}{2}\right)}{|\langle\psi_f|\psi_i\rangle|^2}, \tag{4.100}$$

$$\mathrm{Im}(\sigma_{z,w}) = \frac{1}{4}\frac{\sin(\Delta)\Big[\cos(\theta_i+\theta_f)-\cos(\theta_i-\theta_f)\Big]}{|\langle\psi_f|\psi_i\rangle|^2}. \tag{4.101}$$

In these expressions $\Delta = \phi_f - \phi_i$ and $|\langle\psi_f|\psi_i\rangle|^2$ corresponds to the transition probability between the initial and the final states, which corresponds to

$$|\langle\psi_f|\psi_i\rangle|^2 = \frac{1}{2}\left[1+\cos^2\left(\frac{\Delta}{2}\right)\cos(\theta_i-\theta_f)+\sin^2\left(\frac{\Delta}{2}\right)\cos(\theta_i+\theta_f)\right]. \tag{4.102}$$

According to (4.101), *real weak values* are obtained when: *i*) $\sin(\Delta) = 0$, or *ii*) $\cos(\theta_i - \theta_f) = \cos(\theta_i + \theta_f)$. Let us consider the first case, for which the transition probability $|\langle\psi_f|\psi_i\rangle|^2$ corresponds to $\cos^2(\frac{\theta_i+\theta_f}{2})$ and the first condition of optimality (4.81) becomes,

$$\cos^2\left(\frac{\theta_i-\theta_f}{2}\right) = 1 \tag{4.103}$$

Thus, by choosing $\theta_i = \theta_f$ the first condition will be exactly fulfilled. The transition probability, $|\langle\psi_f|\psi_i\rangle|^2 = \cos^2(\theta_i)$, will remain different from zero unless $\theta_i = \pi/2$. The weak value (4.99) becomes

$$\sigma_{w,z} = \frac{1}{\cos(\theta_i)}, \tag{4.104}$$

which is *always anomalous*, except when the initial state is an eigenstate of $\hat{\sigma}_z$. Consequently, and taking into consideration only the first condition of optimality (which is the case for meters without initial momentum), the quantum Fisher information can be achieved by post-selecting the state

$$|\psi_f\rangle = \cos\left(\frac{\theta_i}{2}\right)|0\rangle - \sin\left(\frac{\theta_i}{2}\right)e^{i\phi_i}|1\rangle. \tag{4.105}$$

As it has been pointed out, this protocol works for meters with zero initial momentum and for every initial state, except for states for which $\theta_i = \pi/2$. For these kind of states (eigenstates of the spin along any direction in the $x-y$ plane), the quantum Fisher information can be *approximately* achieved by post-selection of states with $\theta_f = \pi/2 - 2\epsilon$, where ϵ is a small deviation. In this case, the transition probability becomes $\sin^2(\epsilon) \approx \epsilon^2$ and the weak value (4.99) will correspond to

$$\sigma_{w,z} = \cot(\epsilon) \approx \frac{1}{\epsilon}. \tag{4.106}$$

Summarizing, for a balanced meter, post-selection of the state (4.105) will produce the real weak value (4.104) and the quantum Fisher information will be *exactly* achieved. When $\theta_i = \pi/2$, post-selection of states described by $\theta_f = \pi/2 - 2\epsilon$ will produce the real weak value (4.106) and reach the quantum Fisher information in the limit $\epsilon \to 0$. This strategy is represented in figure 4.5.

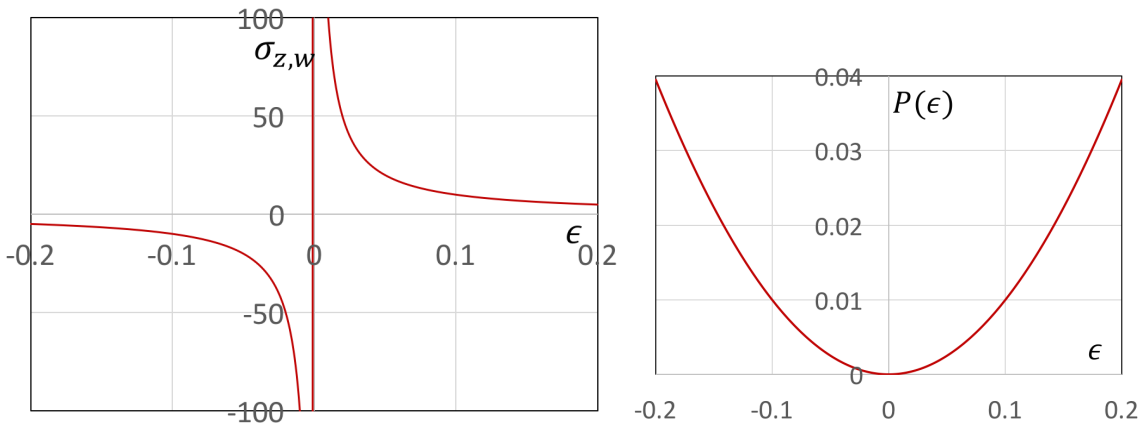
On the other hand, regarding the meter, since the weak value of the spin operator is real, $\mathrm{Re}(P_w)$ should be equal to the initial momentum of the meter, as was explained in the previous section. This can be achieved by preparing the meter in the state (4.93), without compressing (squeezing), and observing the position of the meter.

Finally, for real weak values obtained when $\cos(\theta_i - \theta_f) = \cos(\theta_i + \theta_f)$, the quantum Fisher information can be reached only when the initial state is an eigenstate of $\hat{\sigma}_z$ as well as the post-selected state.

Let us look now at the case of meters with initial momentum, for which *imaginary weak values* are needed. As it can be seen from (4.100), imaginary weak values will be generated when $\theta_i + \theta_f = \pi$. In this case, the transition probability $|\langle\psi_f|\psi_i\rangle|^2$ corresponds to $\cos^2(\Delta/2)\sin^2(\theta_i)$, the weak value is given by $-i\tan(\Delta/2)$ and the optimality conditions (4.81) and (4.82) become

$$\sin^2(\theta_i)\sin^2(\Delta/2) = 1, \tag{4.107}$$

$$\cos^2(\Delta/2)\sin^2(\theta_i)\Big[\cos^2(\theta_i) - \sin^2(\theta_i)\sin^2(\Delta/2)\Big] = \cos^2(\theta_i). \tag{4.108}$$



(a) (b)

Figure 4.5: *a)* The weak value obtained using the sub-optimal post-selection strategy $\theta_f = \pi/2 - 2\epsilon$ is plotted as a function of the parameter ϵ. *b)* The red curve shows transition probability $|\langle\psi_f|\psi_i\rangle|^2$ as a function of ϵ. Recall that the smaller ϵ, the larger the information about g and the larger the weak value.

Notice that when θ_i is close to $\pi/2$, i.e. $\theta_i = \pi/2 + \delta$, then, by making $\Delta = \pi - 2\epsilon$ (which makes the states quasi-orthogonal), both equations will be satisfied in the limit $\epsilon \to 0$. Consequently, states of the form

$$|\psi_f\rangle = \sin\left(\frac{\theta_i}{2}\right)|0\rangle - \cos\left(\frac{\theta_i}{2}\right)e^{i(\phi_i - 2\epsilon)}|1\rangle, \quad \theta_i = \pi/2 + \delta, \quad \delta, \epsilon \ll 1, \tag{4.109}$$

should be post-selected in order to reach the quantum Fisher information in the limit $\epsilon \to 0$. In this situation, the purely imaginary weak value will be equal to $-i\cot(\epsilon)$ (independent of θ_i) and the transition probability will be $\sin^2(\epsilon)\sin^2(\theta_i)$. This sub-optimal strategy is represented in figures 4.6 and 4.7. As explained in the previous section, for imaginary weak values, the momentum of the meter should be observed in order for P_w to be real. This result is valid for any pure initial state of the meter and not only for (4.93).

Finally, in order to study *complex weak values*, it is useful to look at the first condition of optimality, which, in its most general form, is given by

$$\cos^2\left(\frac{\Delta}{2}\right)\cos(\theta_i + \theta_f) + \sin^2\left(\frac{\Delta}{2}\right)\cos(\theta_i - \theta_f) = 1 \tag{4.110}$$

Excluding all the previous cases, of purely real weak values, i.e. assuming $\Delta \neq 0$ or

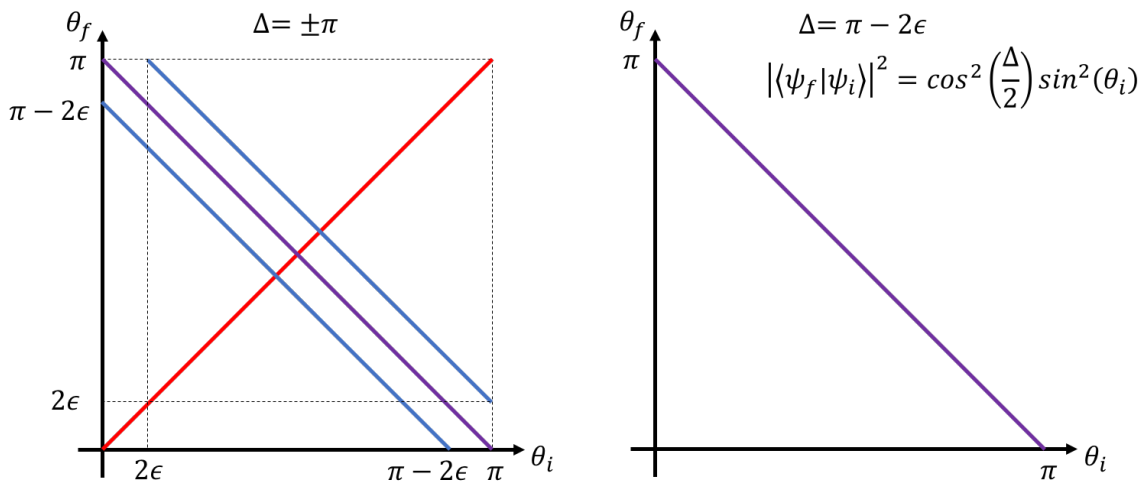
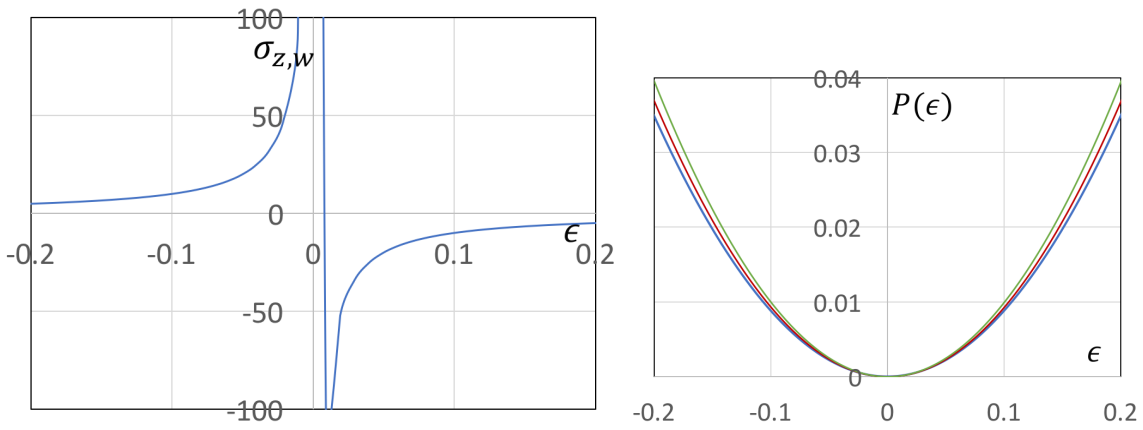


Figure 4.6: *Left Figure*: The red line, $\theta_i = \theta_f$, represents the optimal post-selection strategy for real weak values (for meters without initial momentum). Along the purple line the initial and final states are orthogonal. The two blue lines are sub-optimal strategies for meters with initial momentum, which work for $\theta_i = \pi/2$. *Right Figure*: sub-optimal strategy for imaginary weak values. This strategy is useful for all initial states and meters with momentum, and works for θ_i around $\pi/2$.

$\cos(\theta_i - \theta_f) \neq \cos(\theta_i + \theta_f)$, and purely imaginary weak values, namely, supposing that $\theta_i + \theta_f \neq \pi$, it is not possible to satisfy this condition, either in an exact or approximate manner. Indeed, any "way" to make the left hand side close to one, will rely on purely imaginary or real weak values. Consequently, for a two level system, the quantum Fisher information can be reached through post-selection of the state (4.105) (real weak value, balanced meter) and the state (4.109) (meter with initial momentum, imaginary weak value). The use of real or imaginary weak values will depend on the features of each experiment. For example, protocols based on real weak values may be used if it is experimentally easier to observe the position of the meter, or when the apparatus has no initial momentum (in which case the quantum Fisher information can be exactly achieved, except for the initial states described by $\theta_i = \pi/2$).

Linear regime for a qubit and a Gaussian meter

The weak value can not be increased indefinitely because the approximation (4.87) breaks down. We will study now the conditions under which the linear approximation is valid, that in turn depend on the initial wave function of the meter, and on the initial and



(a) (b)

Figure 4.7: *a)* The weak value, $-\cot(\epsilon)$, obtained using the sub-optimal post-selection strategy (4.109) is plotted as a function of the parameter ϵ. *b)* Transition probability $|\langle\psi_f|\psi_i\rangle|^2$ as a function of ϵ, for different initial system states; $\theta_i = \pi/2$ (green), $\theta_i = \pi/2 - 0.26$ (red) and $\theta_i = \pi/2 - 0.35$ (blue). As for real weak values, the smaller ϵ, the larger the information about g (the closest is the difference with respect to the quantum Fisher information) and the larger the weak value.

final states of system. The exact joint probability of selecting the state $|\psi_f\rangle$ and reading the outcome r was presented in (3.17). For convenience, we reproduce it below.

$$P(r,f) = |\langle\psi_f|\psi_i\rangle|^2 \cdot |\psi(r)|^2 \cdot \underbrace{\sum_{n=0}^{\infty} g^n c_n(r)}_{S(r)} . \tag{4.111}$$

The coefficients $c_n(r)$ are presented in (3.18), and contain weak values (of the system and the apparatus variables) of different orders. The series $S(r) = \sum_{n=0}^{\infty} g^n c_n(r)$ characterize the change of the "unperturbed" probability distribution $|\langle\psi_f|\psi_i\rangle|^2 \cdot |\psi(r)|^2$ due to the measurement process.

When the position of the meter is observed, namely $\hat{R} = \hat{Q}$, and the initial wave function of the apparatus belongs to the family of Gaussian functions described by (4.64), then the full series can be expressed in terms of the Hermite polynomials as

$$S(q) = \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\frac{\sqrt{\gamma} g}{2\sigma_q}\Big)^n \sum_{k=0}^{n} \binom{n}{k} A_w^k \bar{A}_w^{n-k} H_k\Big(\frac{\sqrt{\gamma} z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}} k_0\Big) H_{n-k}^*\Big(\frac{\sqrt{\gamma} z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}} k_0\Big) \tag{4.112}$$

where z is a standardized variable $z = (q - q_0)/\sigma_q$, $\gamma = 1 - i\mathcal{F}$ and $k_0 = p_0/\hbar$ is the initial momentum of the meter. See appendix (B) for the details regarding the derivation of this result.

Analogously to the condition (4.69), the linear approximation will dominate over the higher order terms when

$$i) \quad \left| \mathrm{Re}\left[\sum_{n=2}^{\infty} g^n c_n(q) \right] \right| \ll \left| 1 + \frac{gz}{2\sigma_q} \Big[\mathrm{Re}(A_w) + \mathcal{F}\,\mathrm{Im}(A_w) \Big] + g\,\mathrm{Im}(A_w) k_0 \right|, \tag{4.113}$$

$$ii) \quad \left| \mathrm{Im}\left[\sum_{n=2}^{\infty} g^n c_n(q) \right] \right| \ll \left| \frac{gz}{2\sigma_q} \Big[\mathrm{Im}(A_w) - \mathcal{F}\,\mathrm{Re}(A_w) \Big] - g\,\mathrm{Re}(A_w) k_0 \right|. \tag{4.114}$$

Since the function $S(r)$ is multiplied in (4.111) by a Gaussian function of width $\sim \sigma_q$, these conditions should be satisfied within a couple of standard deviations from the mean, in order to have a good approximation for at least 95% of the total probability.

As was previously indicated, these conditions are valid when the meter is prepared in a pure state of the form (4.64). Any further specification of them should specify the weak values of all orders, which will be done later in this chapter for a two-level system.

Before going into this last issue, let us study the *state of the measurement device in the linear regime*. Recall from (3.55) that, when the post-selection is successful, the wave function of the measurement device (in the coordinate representation) corresponds to

$$\int \frac{\langle \psi_f | \hat{M}_q | \psi_i \rangle}{\int dq |\langle \psi_f | \hat{M}_q | \psi_i \rangle|^2} |q\rangle\, dq \tag{4.115}$$

As can bee seen from appendix (B), when conditions (4.113) and (4.114) are satisfied, then the measurement operators can be safely expanded to first order, and the final wave function of the meter becomes

$$\int \left[\frac{1 + gA_w z\Big(\frac{\gamma}{2\sigma_q}\Big) - igk_0 A_w}{\sqrt{1 + 2gk_0\,\mathrm{Im}(A_w)}} \right] \psi(q)\, |q\rangle\, dq. \tag{4.116}$$

This state corresponds to the coordinate representation of (4.76). The denominator is a normalization factor, whose square is related to the probability of post-selection (it corresponds actually to the perturbation of the transition probability $|\langle \psi_f | \psi_i \rangle|^2$). The factor

can be expanded to first order by imposing an additional requirement: iii) $|gk_0 \operatorname{Im}(A_w)| \ll 1$. Then, the final state of the measurement device becomes

$$\int \left[1 + gA_w z \Big(\frac{\gamma}{2\sigma_q} \Big) - igk_0 \operatorname{Re}(A_w) \right] \psi(q) \, |q\rangle \, dq. \tag{4.117}$$

As an example, consider the case of a meter with zero quadratic phase, i.e. $\gamma = 1$, zero initial momentum, and a purely real weak value. In this situation, the state (4.116) corresponds to

$$\int \left[\psi(q) - gA_w \partial_q \psi(q) \right] |q\rangle \, dq \approx \int \psi(q - gA_w) \, |q\rangle \, dq, \tag{4.118}$$

namely, the (real) wave function of the meter is displaced by gA_w. On the other hand, the squared amplitude of the wave function (4.116) corresponds to the probability density of reading the output q given that the state $|\psi_f\rangle$ was successfully post-selected,

$$P(q|f) = |\psi(q)|^2 \left\{ 1 + \frac{g}{\sigma_q} \Big[\operatorname{Re}(A_w) + \mathcal{F} \operatorname{Im}(A_w) \Big] z \right\}, \tag{4.119}$$

which corresponds to the probability density (4.87) for the case $\hat{R} = \hat{Q}$.

Conditions (4.113), together with the additional third condition on $gk_0 \operatorname{Im}(A_w)$, can be easily analized when the operator $\hat{A}$ satisfies $\hat{A}^2 = \mathbb{1}$. In this scenario, the series (4.112) can be written in closed form as

$$\begin{aligned} S(q) = e^{-\frac{1}{2}(\frac{g}{\sigma})^2} \Big[\cosh\Big(\frac{gz}{\sigma}\Big) \Big(\frac{1 + |A_w|^2}{2} \Big) + \cos\Big(\frac{\mathcal{F} gz}{\sigma} + 2gk_0 \Big) \Big(\frac{1 - |A_w|^2}{2} \Big) \\ + \operatorname{Re}(A_w) \cos\Big(\frac{\mathcal{F} gz}{\sigma} + 2gk_0 \Big) \sinh\Big(\frac{gz}{\sigma}\Big) + \operatorname{Im}(A_w) \sin\Big(\frac{\mathcal{F} gz}{\sigma} + 2gk_0 \Big) \cosh\Big(\frac{gz}{\sigma}\Big) \Big]. \end{aligned} \tag{4.120}$$

From (4.111) it is clear that the probability distribution $P(q, f)$ corresponds to $|\langle \psi_f | \psi_i \rangle|^2 \cdot |\psi(q)|^2 S(q)$. Integration of this probability over q produces the "exact" probability of post-

selection, $P(f)$, which is given by

$$P(f) = |\langle\psi_f|\psi_i\rangle|^2 e^{-\frac{1}{2}(\frac{g}{\sigma})^2}\Bigg[e^{\frac{1}{2}(\frac{g}{\sigma})^2}\Big(\frac{1+|A_w|^2}{2}\Big) + \cos(2gk_0)e^{-\frac{1}{2}(\frac{\mathcal{F}g}{\sigma})^2}\Big(\frac{1-|A_w|^2}{2}\Big)$$
$$-\sin(2gk_0)e^{-\frac{1}{2}(\frac{g}{\sigma})^2(\mathcal{F}^2-1)}\cdot\Big\{\mathrm{Re}(A_w)\sin\Big[\Big(\frac{g}{\sigma}\Big)^2\mathcal{F}\Big] - \mathrm{Im}(A_w)\cos\Big[\Big(\frac{g}{\sigma}\Big)^2\mathcal{F}\Big]\Big\}\Bigg]. \quad (4.121)$$

The series $S(q)$ can be expanded to first order in the parameter g when the conditions (4.113) hold. In this case, these conditions will be fulfilled when the *coupling constant is small* and the *weak value is not too large*, in the sense that

$$i)\Big|\frac{g}{\sigma_q}\Big| \ll 1, \quad ii)\Big|\frac{\mathcal{F}g}{\sigma_q}\Big| \ll 1, \quad iii)\Big|gk_0\Big| \ll 1, \quad (4.122)$$

$$iv)\Big|A_w\frac{g}{\sigma_q}\Big| \ll 1, \quad v)\Big|A_w\Big(\frac{\mathcal{F}g}{\sigma_q}\Big)\Big| \ll 1, \quad vi)\Big|A_w gk_0\Big| \ll 1. \quad (4.123)$$

Under these conditions, and over a large region of the distribution (a region that accumulates at least 95% of the probability), the perturbation of the probability distribution will be a linear function of the parameter g, namely,

$$S(q) = 1 + \frac{g}{\sigma_q}\Big[\,\mathrm{Re}(A_w) + A\,\mathrm{Im}(A_w)\Big]z + 2gk_0\,\mathrm{Im}(A_w). \quad (4.124)$$

And the probability of post-selection will correspond to

$$P(f) = |\langle\psi_f|\psi_i\rangle|^2[1 + 2gk_0\,\mathrm{Im}(A_w)]. \quad (4.125)$$

Notice that, if the meter has no initial momentum, then the probability of post-selection is just the square of the transition amplitude of $\langle\psi_f|\psi_i\rangle$, i.e. as if no intermediate measurement had been performed. The conditions (4.123) define a limit on the maximum value that can be achieved by the weak value, which is given by

$$|A_w| \ll \min\left\{\Big|\frac{g}{\sigma_q}\Big|^{-1}, \Big|\frac{g}{\sigma_q}\Big|^{-1}\cdot\Big|\frac{1}{\mathcal{F}}\Big|, |gk_0|^{-1}\right\}. \quad (4.126)$$

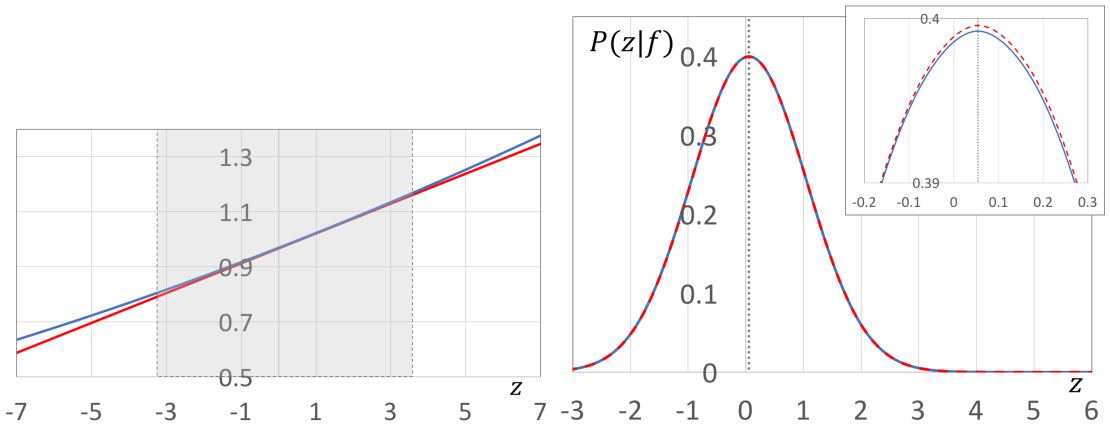


Figure 4.8: a) The blue curve corresponds to the full series, $S(q)$, and the red line is the linear response. The conditions (4.122) and (4.123) are satisfied over the grey area. The different parameters are: $g/\sigma_q = gk_0 = 0.005$, $\mathcal{F} = -0.5$, $\mathrm{Re}(A_w) = 9.18$ and $\mathrm{Im}(A_w) = -3.29$, b) The exact conditioned density function (blue curve) overlaps with its approximation (red dotted). Both functions are expressed in standardized units. The mean value corresponds to $(g/\sigma_q)[\mathrm{Re}(A_w) + \mathcal{F}\,\mathrm{Im}(A_w)] = 0.054$.

In the "standard case" of a real weak value and a meter with no complex phase factor ($\mathcal{F} = 0$) and zero initial momentum ($k_0 = 0$), the limit simply reduces to $|A_w| \ll (g/\sigma_q)^{-1}$. A more complex case is represented in figure 4.8, which shows the full series $S(q)$ as compared to the linear approximation, together with the exact and approximate probability distributions. In this example, the weak value is limited by $|A_w| \ll 200$.

On the other hand, figure 4.9 shows the conditioned probability density, $P(q|f)$, in three different regimes (the linear regime together with two other regimes in which the measurement is stronger).

Finally, it is worth to mention that the case of a meter initialized in a state of the form (4.64) with $\mathcal{F} = 0$, $q_0 = 0$ and $\left\langle \hat{P} \right\rangle = 0$, and a *general* variable $\hat{A}$ (not necessarily a dichotomous variable) is analyzed in [31]. In this case, the conditions that enable the first order approximation correspond to

$$i) \quad gA_w \ll \sigma_q \quad , \quad ii) \quad \frac{g}{\sigma_q} \ll \min_{n=2,3,\ldots} \left| \frac{\langle \psi_f | \hat{A} | \psi_i \rangle}{\langle \psi_f | \hat{A}^n | \psi_i \rangle} \right|^{1/(n-1)} . \tag{4.127}$$

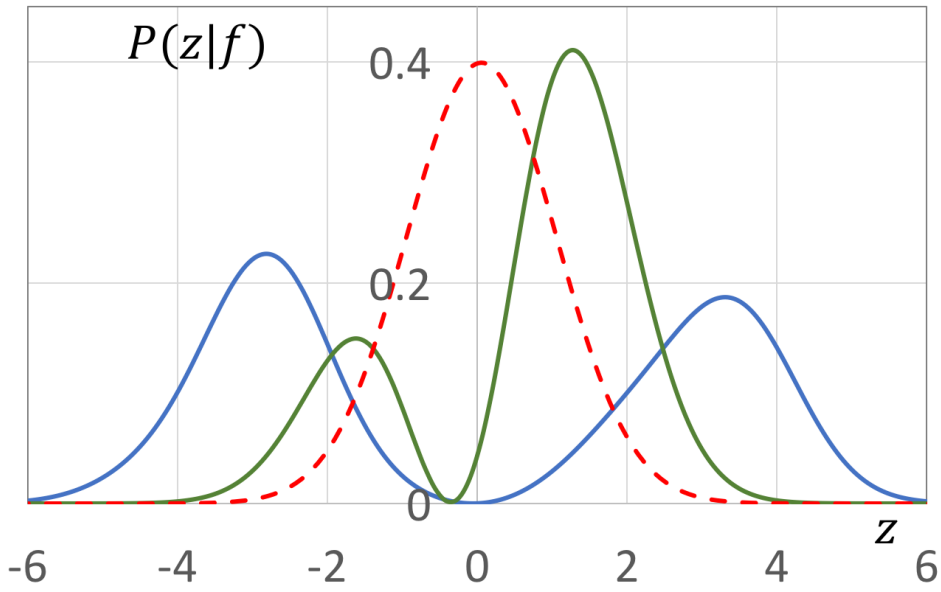


Figure 4.9: The conditioned density function (in standardized units) in different regimes. For all the curves the parameters are: $gk_0 = 0.005$, $\mathcal{F} = -0.5$, $\mathrm{Re}(A_w) = 9.18$ and $\mathrm{Im}(A_w) = -3.29$. The blue curve, with two peaks, corresponds to the case $g/\sigma_q = 3$. The measurement is weaker in the case described by the green curve ($g/\sigma_q = 0.5$), in which the peaks begin to overlap. The red dotted function corresponds to the linear regime, for which $g/\sigma_q = 0.05$. Notice that the function has a Gaussian shape, that arises as a large superposition of two peaks.

4.9.3 Fisher information in the presence of white noise

In this section we will consider a quantum measurement that is affected by the presence of classical noise. Let R be the result of an *efficient* quantum measurement performed on a system in a state ρ. The probability that the result R is equal to r is given by

$$P(R = r) = \mathrm{Tr}\Big[\hat{M}_r^\dagger \hat{M}_r \rho\Big], \tag{4.128}$$

where $\hat{M}_r$ is the measurement operator associated to the outcome r. Suppose, however, that we do not have access to the quantum measurement. Instead, a signal S is read, which consists of the sum of the quantum measurement $R \sim P(R = r)$ and a noise variable $\Upsilon \sim P(\Upsilon = \upsilon)$,

$$S = R + \Upsilon. \tag{4.129}$$

The noise variable Υ is assumed to be a continuous variable. Equation (4.129) describes the presence of *additive noise*. Since the signal arises from the addition of two random variables, the probability of reading the signal $S = s$ is given by the *convolution* between the probability density function of the quantum measurement and the noise,

$$\begin{aligned} P(S=s) &= \int P(\Upsilon=\upsilon)P(R=s-\upsilon)d\upsilon \\ &= \mathrm{Tr}\Big[\int \sqrt{P(\Upsilon=\upsilon)}\hat{M}_{s-\upsilon}\rho\hat{M}^{\dagger}_{s-\upsilon}\sqrt{P(\Upsilon=\upsilon)}d\upsilon\Big]. \end{aligned} \tag{4.130}$$

Hence, the measurement can be described by an operation $\mathcal{O}_s\rho = \int d\upsilon \hat{A}_{\upsilon,s}\rho\hat{A}^{\dagger}_{\upsilon,s}$, with elements (or Krauss operators) given by $\hat{A}_{\upsilon,s} = \sqrt{P(\Upsilon=\upsilon)}\hat{M}_{s-\upsilon}$. Thus, the *unnormalized conditioned* state of the system, after the result r is obtained, is given by $\tilde{\rho}_s = \mathcal{O}_s\rho$ and the probability (4.130) corresponds simply to $\mathrm{Tr}[\mathcal{O}_s\rho]$. When the measurement is weak, this probability reduces to

$$P(S=s) = \int d\upsilon P(\Upsilon=\upsilon)|\psi(s-\upsilon)|^2\Big\{1 + 2(g/\hbar)\Big\langle\hat{A}\Big\rangle \mathrm{Im}(P_w)\Big\}. \tag{4.131}$$

The Fisher information about g contained in this density function can be computed using (4.15) or (4.18), and the result should be compared to (4.60), in order to appreciate the effects of the noise. For simplicity, let us consider the "standard case", in which 1) the position of the measurement device is observed, 2) the meter starts in the complex Gaussian function (4.64), 3) the system is a qubit, and 4) the measurement is weak ($g/\sigma_q \ll 1$). As was shown in section (4.9.1), under these circumstances and in the absence of classical noise, the Fisher information corresponds to (4.65). If noise is taken into consideration, it is necessary to specify the type of noise that affects the measurement in order to solve the convolution (4.131) and compute the Fisher Information.

A common type of noise is the *Gaussian white noise*, described by a normal probability density with zero mean and variance σ^2_{noise},

$$P(\Upsilon=\upsilon) = \frac{1}{\sqrt{2\pi}\sigma_{noise}}\exp\Big\{-\frac{\upsilon^2}{2\sigma^2_{noise}}\Big\}. \tag{4.132}$$

In this case, the convolution becomes a Gaussian distribution with mean equal to q_0 and variance $\sigma_q^2 + \sigma_{noise}^2$, and the Fisher information corresponds to

$$\mathcal{I}(g) = \frac{\left\langle \hat{A} \right\rangle^2}{\sigma_q^2 + \sigma_{noise}^2}. \tag{4.133}$$

By comparing this result with (4.65) it is clear that the presence of additive Gaussian white noise reduces the information about the parameter we wish to estimate. The detrimental effects of the noise will be small as long as $\sigma_{noise}/\sigma_q < 1$.

Now, let us take a look at the scenario with post-selection. If post-selection is performed, the probability to read the signal $S = s$ in the intermediate result *and* to select the pure state $|\psi_f\rangle$ will be given by

$$P(s, f) = \langle \psi_f | \, \mathcal{O}_s \rho \, | \psi_f \rangle \, . \tag{4.134}$$

This result should be compared with (4.72). The probability of post-selection is obtained by integrating over s, and the conditioned probability is obtained from $P(s|f) = P(s, f)/P(f)$. In an analogous way to the result (4.74), now the Fisher information is

$$\mathcal{I}(g) = P(f)\mathcal{I}(g|f) + \mathcal{I}_P, \tag{4.135}$$

where $\mathcal{I}(g|f)$ corresponds to the Fisher information contained in $P(s|f)$ and $\mathcal{I}_P$ is the information in the partition of the whole ensemble into two sub-ensembles, according to the result of the post-selection (success or failure).

For simplicity, let us analyse the previously called "standard case". Assume also that the weak value is real and the noise is (4.132). In this simplified situation, the post-selection probability (4.125) is simply the the transition probability $|\langle \psi_f | \psi_f \rangle|^2$ (it contains no information about g) and the conditioned probability density becomes

$$P(s|f) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_q^2 + \sigma_{noise}^2}} \exp\left\{ -\frac{[s - g \, \mathrm{Re}(A_w)]^2}{2(\sigma_q^2 + \sigma_{noise}^2)} \right\}, \tag{4.136}$$

namely, the weak value acts as an amplification factor of the displacement, but the noise widens the probability density function. Thus, the Fisher information (4.135) corresponds to

$$\mathcal{I}(g|f) = |\langle\psi_i|\psi_f\rangle|^2 \cdot \frac{\mathrm{Re}(A_w)^2}{\sigma_q^2 + \sigma_{noise}^2}. \tag{4.137}$$

Consequently, the effect of Gaussian white noise is the same as in the case without post-selection, namely, it reduces the Fisher information by augmenting the variance. This occurs because the initial wave function of the meter is a Gaussian function and thus its convolution with another Gaussian function produces a final density function that is also Gaussian, whose variance and mean simply correspond to the addition of the variances and means of each function, respectively. As a final remark, notice that the Fisher information (4.133) has the same order of magnitude as (4.137). Thus, in the so called "standard case", the information remains approximately the same, with or without post-selection [142, 143]. However, in the last situation, all the information is contained in a small amount of post-selected events. This fact can be beneficial in the presence of detector saturation [144].

4.9.4 Fisher information in the presence of correlated noise

In this last section we will examine the Fisher information when the intermediate measurement is affected by classical noise that has correlations in time. For simplicity, we will restrict our analysis to the "standard case", that is characterized by 1) a meter initially prepared in the state (4.64), 2) the observation of the position of the measurement device, 3) a two-level system (qubit), and 4) a weak measurement. Without post-selection this last feature means that $g/\sigma_q \ll 1$, whereas with post-selection a weak measurement is described by conditions (4.122) and (4.123). Also, when post-selection is performed, it will be assumed that the weak value is real.

Consider a sample of N elements $\{S_1, ..., S_N\}$, or repetitions of the experiment. As in the previous section, assume that each observation has a "quantum component", denoted by R_i, and a "classical component", denoted by Υ_i. The first component appears due to a measurement performed on a quantum system, and the latter is a consequence of

the presence of classical noise affecting the measurement process. The variables R_i are uncorrelated, while the classical variables Υ_i may be correlated. If the noise is additive, then each observation, or signal, can be expressed as

$$S_i = R_i + \Upsilon_i, \qquad \forall i = 1, ..., N. \tag{4.138}$$

First, let us consider the case without post-selection. The joint probability to read $S_i = s_i$, $\forall i = 1, .., N$, will be denoted as $P(\vec{S}_N = \vec{s}_N)$, where $\vec{S}_N$ and $\vec{s}_N$ are N-dimensional vectors; the first contains N random variables and the second contains the N real values taken by them. The subscript N has been used explicitly to emphasize that the vector has dimension N (later, when the case with post-selection is analyzed, the dimension of the vectors will correspond to the number of post-selected events).

On the other hand, the noise process will be described by an N-dimensional Gaussian distribution,

$$P(\vec{\Upsilon}_N = \vec{\upsilon}_N) = \frac{1}{\sqrt{2\pi}^N |\boldsymbol{C}|^{N/2}} \exp\left\{-\frac{\vec{\upsilon}_N^T \boldsymbol{C}^{-1} \vec{\upsilon}_N}{2}\right\}. \tag{4.139}$$

The N dimensional vectors, $\vec{\Upsilon}_N$ and $\vec{\upsilon}_N$, contain N random variables or *noise variables* (Υ_i) and the N values taken by them (υ_i), respectively. Thus, $P(\vec{\Upsilon}_N = \vec{\upsilon}_N)$ denotes the joint probability that $\Upsilon_i = \upsilon_i$, $\forall\ i = 1, .., N$. The vector $\vec{\upsilon}_N^T$ is the transpose vector of $\vec{\upsilon}_N$. The covariance matrix $\boldsymbol{C}$ is an $N \times N$ matrix whose diagonal elements represent the variance of the noise, which will be assumed to be the same for all the noise variables, i.e. $C_{i,i} = \sigma^2_{noise}$. The off-diagonal terms $C_{i,j}$ $(i \neq j)$ represent the covariance between the different noise variables. In general, the covariances $C_{i,j}$ may take positive or negative values. The covariance matrix should be a *positive definite* matrix and the covariances satisfy the Cauchy-Schwarz inequality, i.e. $|C_{i,j}| \leqslant \sigma^2_{noise}$. The inverse of the covariance matrix is $\boldsymbol{C}^{-1}$ and $|\boldsymbol{C}|$ is its determinant.

Consequently, the joint probability of obtaining the outcomes $S_i = s_i$, in the presence

of additive noise of the type (4.139), is given by

$$P(\vec{S}_N = \vec{s}_N|g) = \int d\vec{\upsilon}_N P(\vec{\Upsilon}_N = \vec{\upsilon}_N) P(R_1 = s_1 - \upsilon_1, ..., R_N = s_i - \upsilon_N|g), \quad (4.140)$$

where $d\vec{\upsilon}_N = d\upsilon_1 \dots d\upsilon_N$. Notice that now it has been made explicit the fact that this distribution depends on the value of the parameter g, i.e. that it can be viewed as a likelihood function. Note also that the integral represents a convolution between two multidimensional probability distributions, of the noise and of the quantum measurements. Since the quantum measurements are independent of each other, then

$$P(R_1 = s_1 - \upsilon_1, ..., R_N = s_i - \upsilon_N|g) = \prod_{i=1}^{N} P(R_i = s_i - \upsilon_i). \quad (4.141)$$

In the standard case (described at the beginning of this section), the probability that each quantum measurement takes a certain value follows a Gaussian distribution with mean value equal to $g\left\langle \hat{A} \right\rangle$ and variance equal to σ_q^2. This fact can be seen from expression (2.43). Then, since the distribution (4.141) is the product of N Gaussian distributions,

$$P(R_1 = r_1, ..., R_N = r_N|g) = \frac{1}{\sqrt{2\pi}^N |\boldsymbol{D}|^{N/2}} \exp\left\{ -\frac{(\vec{r}_N - \vec{\mu}_N)^T \boldsymbol{D}^{-1} (\vec{r}_N - \vec{\mu}_N)}{2} \right\}. \quad (4.142)$$

The vector $\vec{r}_N$ has the values r_i as its elements. The covariance matrix is a *diagonal matrix* with elements $D_{ii} = \sigma_q^2$. The vector $\vec{\mu}_N$ contains the N means of the distribution, all of which are equal to $g\left\langle \hat{A} \right\rangle$,

$$\vec{\mu}_N = g\left\langle \hat{A} \right\rangle \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \quad (4.143)$$

The convolution of two Gaussian distributions is also a Gaussian distribution, with covariance matrix and vector of means equal to the sum of the corresponding covariance matrices and vector of means. Therefore, the distribution (4.140) will be a Gaussian with vector of

means given by (4.143) and a covariance matrix equal to

$$\boldsymbol{E} = \boldsymbol{C} + \boldsymbol{D} = \begin{pmatrix} \sigma_q^2 + \sigma_{noise}^2 & C_{1,2} & \cdots & C_{1,N} \\ C_{2,1} & \sigma_q^2 + \sigma_{noise}^2 & & C_{2,N} \\ \vdots & & \ddots & \vdots \\ C_{N,1} & C_{N,2} & & \sigma_q^2 + \sigma_{noise}^2 \end{pmatrix}. \tag{4.144}$$

From (4.15) or (4.18) it is a straightforward calculation [146, 145] to show that the Fisher information about the parameter g, contained in a multidimensional Gaussian distribution with covariance matrix $\boldsymbol{E}$ and vector of means $\vec{\mu}$, is given by

$$\mathcal{I}(g) = \left\langle \hat{A} \right\rangle^2 \sum_{i,j}^{N} E_{i,j}^{-1}, \tag{4.145}$$

where $E_{i,j}^{-1}$ denotes the (i,j) element of the inverse of the covariance matrix $\boldsymbol{E}^{-1}$. Notice that, if the noise is uncorrelated, i.e. $C_{i,j} = 0 \ \forall i \neq j$, then expression (4.145) reduces to (4.133). In the opposite situation, when the correlations do not decay and are constant [145],

$$C_{i,j} = \eta \in \mathbb{R} \quad \forall i, j \in \{1, ..., N\} \wedge i \neq j, \tag{4.146}$$

then it is easy to check that the elements of the inverse of the covariance matrix of $\boldsymbol{E}$ (4.144) are given by

$$E_{i,j}^{-1} = \frac{1}{\sigma_q^2 + \sigma_{noise}^2} \left(\delta_{ij} - \frac{\eta}{\sigma_q^2 + \sigma_{noise}^2 + N\eta} \right). \tag{4.147}$$

This scenario, in which the correlations of the noise process are constant, will be referred as the case affected by the presence of *colored noise*. In this model, in order for the covariance matrix to be positive definite, the covariances should satisfy $0 \leqslant \eta \leqslant \sigma_{noise}^2$. In this situation, the Fisher information (4.145) becomes

$$\mathcal{I}(g) = \left\langle \hat{A} \right\rangle^2 \frac{N}{\sigma_{noise}^2 + \sigma_q^2 + N\eta}. \tag{4.148}$$

First, notice that when the *sample is large*, in the sense that $N \gg (\sigma^2_{noise} + \sigma^2_q)/\eta$, the Fisher information will be *constant*. It will be independent of σ^2_q and σ^2_{noise}, and will only depend on the expectation value of $\hat{A}$ and on the covariance η. Indeed, the Fisher information will reach a constant value of

$$\mathcal{I}(g) = \left\langle \hat{A} \right\rangle^2 / \eta. \tag{4.149}$$

On the other hand, if the *sample is small*, $N \ll (\sigma^2_{noise} + \sigma^2_q)/\eta$, then $\mathcal{I}(g) \sim N$, i.e. the precision of the measurement will scale as the SQL (as long as N is small). In this regime the Fisher information becomes

$$\mathcal{I}(g) = \frac{\left\langle \hat{A} \right\rangle^2}{\sigma^2_q + \sigma^2_{noise}} \cdot N, \tag{4.150}$$

that corresponds to N times the information (4.133), i.e. the measurement behaves has N independent measurements. This scenario is feasible when η is much smaller than the variance, which explains that the measurements behave almost as if the were independent.

Let us consider now the case when post-selection is performed, which is a bit more involved. In this scenario, the experiment is repeated N times, but the post-selection may be successful in k cases, $0 \leqslant k \leqslant N$. Notice that the k successfully post-selected events may occur in different ways.

For example, if $N = 4$ and $k = 2$, then the two successful post-selections may occur in six different ways; in the first and second experiments, which is denoted by $\{1, 2\}$, or in the second and third experiments, which is similarly denoted by $\{2, 3\}$, and so on for the rest of the other four possibilities. Thus, in general, the set $\{i_1, i_2, ..., i_k\}$ means that the first successful post-selection occurred in the i_1-th repetition of the experiment, the second in the i_2-th repetition, and so on. Notice that $i_1 < i_2 < ... < i_k$ and that $1 \leqslant i_j \leqslant N$, $\forall$ $j = 1, ..., k$. For N repetitions and k successful post-selections, the number of possible sets equals $\binom{N}{k}$. Let us label each of these sets with an auxiliary variable $J = 1, ..., \binom{N}{k}$. For example, in the previously considered case ($N = 4$, $k = 2$), the set $J = 1$ may correspond to the occurrence of successful post-selections in the first and second repetitions, namely,

to the set $\{1, 2\}$.

Recall that every time the post-selection is successful the meter is observed and when it fails the results are ignored. Therefore, if k out of N events are successfully post-selected, and they occur according to the set $J = \{i_1, i_2, ..., i_k\}$, we will end with k observations of the measurement device (4.138),

$$\vec{S}_k^J = \begin{pmatrix} S_{i_1} \\ \vdots \\ S_{i_k} \end{pmatrix}. \tag{4.151}$$

Let us compute the probability that $\vec{S}_k^J = \vec{s}_k$, where $\vec{s}_k$ is a k-dimensional vector with real values. This probability will be denoted as $P_N(\vec{S}_k^J = \vec{s}_k)$ and is given by

$$P_N(\vec{S}_k^J = \vec{s}_k) = \left(1 - |\langle\psi_f|\psi_i\rangle|^2\right)^{N-k} \int d\vec{v}_k P(\vec{\Upsilon}_k^J = \vec{v}_k) P(\vec{R}_k^J = \vec{s}_k - \vec{v}_k, f), \tag{4.152}$$

where $d\vec{v}_k = dv_1 \cdots v_k$. In an analogous way to (4.151), the vector $\vec{\Upsilon}_k^J$ is formed by the k noise variables specified by the set J. Similarly, $\vec{R}_k^J$ is a vector of k random variables associated to the results of the quantum measurements, specified by J.

The probability $P(\vec{R}_k^J = \vec{s}_k - \vec{v}_k, f)$ corresponds to the probability that $R_{i_j} = s_j - v_j$ *and* that the state $|\psi_f\rangle$ is successfully post-selected, $\forall\ j = 1...k$. Since the results of the quantum measurements are not correlated, then

$$P(\vec{R}_k^J = \vec{q}_k, f) = \prod_{j=1}^{k} P(R_{i_j} = q_j, f), \tag{4.153}$$

which is the equivalent of (4.141) for the post-selection case. From (3.25) it is easy to see that, in the i_j-th repetition of the experiment, the probability that the quantum measurement R_j equals q_j *and* the state $|\psi_f\rangle$ is post-selected, is given by

$$P(R_{i_j} = q_j, f) = |\langle\psi_f|\psi_i\rangle|^2 \cdot |\psi(q_j - gA_w)|^2. \tag{4.154}$$

Therefore, the distribution (4.153) corresponds to the product of $|\langle\psi_f|\psi_i\rangle|^{2k}$ and a k-

dimensional Gaussian distribution with covariance matrix equal to $\boldsymbol{D}$ (see equation 4.142), which is a diagonal matrix with elements $D_{i,i} = \sigma_q^2$. The vector of means is given by

$$\vec{\mu}_k = gA_w \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{(k\times 1)} , \tag{4.155}$$

that is the equivalent result to (4.143) for the post-selection scenario. Hence, besides a factor of $|\langle\psi_f|\psi_i\rangle|^{2k}$, the integral term in (4.152) corresponds to the convolution between a Gaussian distribution (with vector of means $\vec{\mu}_k$ and covariance matrix $\boldsymbol{D}$) and the distribution $P(\vec{\Upsilon}_k^J = \vec{v}_k)$, which is obtained from (4.139), by marginalizing over the variables *not* included in the set J. So, it is also a Gaussian distribution, with vector of means equal to the null vector and covariance matrix given by

$$\boldsymbol{C}_k^J = \begin{pmatrix} \sigma_{noise}^2 & C_{i_1,i_2} & \cdots & C_{i_1,i_k} \\ C_{i_2,i_1} & \sigma_{noise}^2 & & C_{i_2,i_k} \\ \vdots & & \ddots & \vdots \\ C_{i_k,i_1} & C_{i_k,i_2} & & \sigma_{noise}^2 \end{pmatrix}_{(k\times k)} . \tag{4.156}$$

Notice that it contains the covariances between the noise variables, but only between those k variables associated to the cases when the post-selection is successful (which are specified by J). Therefore, besides the factor of $|\langle\psi_f|\psi_i\rangle|^{2k}$, the convolution in (4.152) is finally a Gaussian distribution with the vector of means given by (4.155) and covariance matrix $\boldsymbol{E}_k^J = \boldsymbol{C}_k^J + \boldsymbol{D}$, which is a similar result to (4.144).

The previous analysis shows that, *given* k successful post-selections out of N repetitions of the experiment, and assuming that the post-selected events occur in the order defined by the set J, the probability to read $\vec{s}_k$, i.e. the distribution (4.152), is simply the product between $(1-|\langle\psi_f|\psi_i\rangle|^2)^{N-k}\cdot|\langle\psi_f|\psi_i\rangle|^{2k}$ and a Gaussian distribution with vector of means $\vec{\mu}_k$ and covariance matrix $\boldsymbol{E}_k^J$. In an analogous way to (4.145), the Fisher information about

g contained in this distribution is given by

$$\mathcal{I}^J_{N,k}(g) = A_w^2 \cdot (1 - |\langle\psi_f|\psi_i\rangle|^2)^{N-k} \cdot |\langle\psi_f|\psi_i\rangle|^{2k} \cdot \sum_{i,j}^{k} (E^J_k)^{-1}_{i,j}. \tag{4.157}$$

This is a involved expression, which requires knowledge about the different correlations between the noise in the post-selected events. If the noise is "coloured" (4.146), namely, the correlations are constant, then the elements of the matrix $\boldsymbol{E}^J_k$ are independent of the order defined by the set J. In this scenario, the Fisher information becomes

$$\mathcal{I}^J_{N,k}(g) = A_w^2 \cdot (1 - |\langle\psi_f|\psi_i\rangle|^2)^{N-k} \cdot |\langle\psi_f|\psi_i\rangle|^{2k} \cdot \left[\frac{k}{\sigma_q^2 + \sigma^2_{noise} + k\eta}\right]. \tag{4.158}$$

Finally, the total Fisher information in the post-selection strategy is obtained simply by summing $\mathcal{I}^J_{N,k}(g)$ over all the possible values of k and J,

$$\mathcal{I}(g) = A_w^2 \sum_{k=0}^{N} \binom{N}{k} (1 - |\langle\psi_f|\psi_i\rangle|^2)^{N-k} \cdot |\langle\psi_f|\psi_i\rangle|^{2k} \cdot \left[\frac{k}{\sigma_q^2 + \sigma^2_{noise} + k\eta}\right]. \tag{4.159}$$

This expression should be compared with the Fisher information associated to the strategy that does not employ post-selection (4.148). If the sample is small, in the sense that $N \ll (\sigma_q^2 + \sigma^2_{noise})/\eta$, then the Fisher information (4.159) becomes

$$\mathcal{I}(g) = \Big(\frac{A_w^2}{\sigma_q^2 + \sigma^2_{noise}}\Big) |\langle\psi_f|\psi_i\rangle|^2 \cdot N, \tag{4.160}$$

that corresponds to N times the information (4.137), because the effect of the correlations are small as compared to the variance of the noise. On the other hand, if $N|\langle\psi_f|\psi_i\rangle|^2 \gg (\sigma^2_{noise} + \sigma_q^2)/\eta$, then the Fisher information (4.159) achieves a constant value equal to

$$\mathcal{I}(g) = \frac{A_w^2}{\eta}. \tag{4.161}$$

This value may be far larger than $\left\langle \hat{A} \right\rangle^2 / \eta$, the Fisher information without post-selection (4.149). In conclusion, in the presence of correlated noise and for large samples, the saturation value of the Fisher information (in the so called "standard case") is largely

increased (as compared to the non post-selection case) by employing the post-selection strategy and anomalous real weak values.

The Fisher information is a formal statistical tool to characterise the maximum precision that can be achieved by an estimation procedure. There are other criteria, such as the signal to noise ratio (SNR), under which the benefits of the post-selection strategy over the non post-selection scenario can be appreciated. Consider that the parameter is estimated from the sample mean estimator (4.27),

$$\hat{g}_{SME}(N) = \frac{\sum_{i=1}^{N} S_i}{N}, \tag{4.162}$$

where S_i are given by (4.138). In the non post-selection case the mean value of the estimator equals $g\left\langle \hat{A} \right\rangle$. If the noise variables have long correlation times (coloured noise), then the variance of the estimator corresponds to

$$Var[\hat{g}_{SME}(N)] = \frac{\sigma_{noise}^2 + \sigma_q^2 - \eta}{N} + \eta. \tag{4.163}$$

Thus, for a large sample, the SNR (4.22) of the estimator becomes $g\left\langle \hat{A} \right\rangle / \sqrt{\eta}$. When post-selection is employed there will be, on average, M successfully post-selected events, where $M = N|\langle \psi_f | \psi_i \rangle|^2$. The mean value of the sample mean estimator will be gA_w, while its variance will correspond to

$$Var[\hat{g}_{SME}(M)] = \frac{\sigma_{noise}^2 + \sigma_q^2 - \eta}{N|\langle \psi_f | \psi_i \rangle|^2} + \eta. \tag{4.164}$$

As in the previous case, when the sample is large, the SNR of the sample reaches the value $gA_w/\sqrt{\eta}$, which is larger than the SNR of the non post-selection case (as long as the weak value is anomalous). Therefore, as with the Fisher information, the SNR is also increased when post-selection is employed and the experiment is affected by noise with long correlation times. If the noise is uncorrelated, the SNR of the sample mean, with and

without post-selection, is given by

$$SNR = \frac{gA_w|\langle\psi_f|\psi_f\rangle|}{\sqrt{\sigma_q^2 + \sigma_{noise}^2}} \cdot \sqrt{N} \quad , \quad SNR = \frac{g\langle A\rangle}{\sqrt{\sigma_q^2 + \sigma_{noise}^2}} \cdot \sqrt{N}, \tag{4.165}$$

respectively. Notice that both values scale as $\sqrt{N}$ and have the same order of magnitude. Consequently, when the measurement is affected by white noise, the SNR does not change substantially by using one strategy or the other. We have already arrived to this conclusion in the previous section, by comparing the Fisher information (4.133) to (4.137).

4.9.5 Conclusions

As a brief summary of this last section, it should be pointed out that in 4.9.1 we have analyzed the Fisher information contained in a weak measurement without post-selection and showed, on the basis of the results presented in [135], that the quantum Fisher can be achieved when the initial state is an eigenstate of the measured variable. Then, we have provided an analysis of the validity of the so called "linear regime". Next, in 4.9.2 we have described the "optimal post-selection strategy", which presents some of the results given in [141] but mostly in [143]. The subsections "Fisher Information in the conditioned probability" and "Linear regime for a qubit and a Gaussian meter" are of our own analysis. The conclusions to which we have arrived in subsections 4.9.3 and 4.9.4 are also presented in [146, 145]. Expression (4.159), the Fisher information in the presence of correlated noise, is of our own analysis.

Chapter 5

Optomechanical system in Mach-Zehnder interferometer

Optomechanical systems [154, 155] provide an interface between light and mechanical motion. In its most simple configuration, an optomechanical system consists of an optical or microwave cavity with one of its end mirrors that vibrates as if it was attached to a spring. Typically, the cavity is driven by laser light. The estimation of the phase shift of the light reflected from (or transmitted through) the cavity allows to monitor continuously the motion of the mirror, achieving the SQL in position measurements [156, 157, 158, 159, 160, 161] or approaching the HL when quantum non demolition (QND) measurements [166, 160, 165] of a single mechanical quadrature [163, 162] or of the phonon occupation number [164] are implemented. Displacements in the order of attometers, and below, can be detected within one second of monitoring [167].

These hight sensitivities are particularly relevant for gravitational waves detection. In a gravitational wave interferometer [168], an optical cavity is put in each arm, the length of the cavity being in the order of km and the masses of the mirrors of several kg. A gravitational wave changes the length of the cavity, which modifies the frequencies of the cavity modes. In turn, the light emitted from the cavity and detected at the output of the interferometer is modified.

Quantum effects can be appreciated when all thermal quanta are removed from the

mechanical component, i.e. when it is put near to its motional ground state. Micro and nano mechanical devices, whose masses are of the order of 10^{-12} kg (or less), have been cooled down to states with mean occupation number less than one [169, 170, 171, 172]. Large objects, with masses of the order of kg, have been cooled down to states with nearly 200 phonons [173]. Recently, employing the suspended mirrors of the Laser Interferometer Gravitational-Wave Observatory (LIGO), the center of mass of a mechanical oscillator with a mass of 10 kg was cooled down to a state with nearly ten quanta [211].

The cooling of mechanical devices with laser light can be made in so called "strong coupling regime", in which the Heisenberg-Langevin equations for the optical and mechanical variables become linear. In this regime, not only energy can be extracted from the mechanical components into the optical modes, but also the inverse is possible, i.e. the conversion of photons into phonons.

The coherent control and manipulation of mechanical states offers the possibility to study fundamental aspects of quantum theory, such as the "measurement problem", with regards to the decoherence of superpositions of macroscopic objects into classical states. Different scaling rules for the decoherence rates (in terms of the number of particles) can in principle be tested by generating distinct superpositions of mechanical objects with large masses [179, 177, 180]. Also, the relation between gravity and quantum theory can be explored, by analysing the role of gravity in the decoherence of the superpositions [174, 175], or even by directly testing predictions of quantum gravity [176].

The control of mechanical states is also important for sensing applications, in areas such as microscopy[181], magnetometry [182] and acceleration sensing [183]. In particular, the generation of mechanical squeezed states [184], with reduced uncertainty in the position or momentum variables, can lead to higher sensitivities in different applications.

Mechanical elements can also be used to mediate the interaction between different systems, a configuration that has been referred as *hybrid systems*. Such systems can employ a mechanical device to couple cold atoms [185], spins [186, 187] or superconducting qubits [171, 188] to an optical mode, which has applications in quantum information processing [189]. The possibility to establish interactions between flying qubits and stationary (non optical) qubits has applications in communications [190]. Devices that act as quantum

memories [191] can be designed, based on the possibility to "storage" light inside optomechanical systems due to the *optomechanically induced transparency* effect [192, 193, 194].

In this work, we use an optomechanical system (a cavity with a vibrating mirror) inserted inside an interferometer to create optomechanical entanglement. Single photon detection at the outputs of the interferometer will produce weak value amplification of the effect of the photon on the mechanical oscillator, via post-selection. In the first section of this chapter we present the hamiltonian description of a cavity with a moving mirror. In section 5.2, the generation of optomechanical entanglement with an interferometer is described. Also, it is explained how the observation of the interferometric visibility allows to study the decoherence of a vibrating mirror. In the next section, our experimental proposal is described, which is based on a Mach-Zehnder interferometer. The weak value amplification effect, both for a mirror starting in its motional ground state or in a thermal state, is analyzed in section 5.4 from the perspective of the Fisher information and the signal to noise ratio. Finally, in section 5.5, the amplification effect in the presence of classical noise is studied.

5.1 Cavity optomechanics

A Fabry-Pérot resonator consists of two highly reflecting mirrors separated by a distance L. The electric field vanishes at the location of the mirrors and thus the wave number of the different (longitudinal) modes are $k_n = (\pi/L)n$, $n \in \mathbb{N}$. The dispersion relation in empty space is $\omega_n = ck_n$, where c is the speed of light. The difference between the angular frequencies of neighbouring modes is called the free spectral range (FSR) of the cavity, $\Delta\omega = (\pi/L)c$. For resonators with lengths in the order of centimetres, $\Delta\omega$ is of the order of GHz. Thus, the modes are well separated in frequency and are individually addressable. Let ω_{cav} be the angular frequency of one mode of interest. The energy of the mode is $\hbar\omega_{cav}\hat{a}^\dagger\hat{a}$, where $\hat{a}$ ($\hat{a}^\dagger$) is the boson annihilation (creation) operator of the mode into consideration.

Assume now that one of the mirrors is allowed to vibrate, as it is shown in figure 5.1. The vibrations of the center of mass of the mirror can be treated quantum-mechanically,

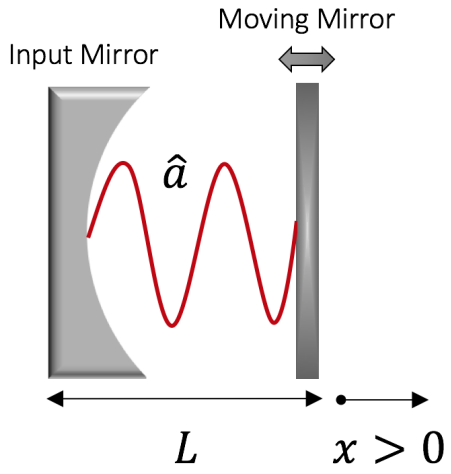


Figure 5.1: Optomechanical system (OM): A Fabry-Pérot with a movable end mirror. The center of mass of the mirror is treated as a quantum harmonic oscillator. The length of the cavity is modulated by the small vibrations of the mirror. The interaction between the radiation field and the mechanical oscillator occurs due to the radiation pressure force.

as an harmonic oscillator, with position and momentum operators denoted as $\hat{X}$ and $\hat{P}$. The corresponding creation and annihilation operators are

$$\hat{b} = \sqrt{\frac{M\Omega}{2\hbar}}\left(\hat{X} + i\frac{1}{M\Omega}\hat{P}\right) \quad , \quad \hat{b}^\dagger = \sqrt{\frac{M\Omega}{2\hbar}}\left(\hat{X} - i\frac{1}{M\Omega}\hat{P}\right), \tag{5.1}$$

where M is the total mass of the mirror and Ω is the natural frequency of the oscillator. For a micro-sized mechanical oscillator $M \sim 10^{-12}$ kg and $\Omega \sim 10^8$ Hz [167]. Consequently, the total energy inside the cavity is given by

$$\hat{H} = \hbar\omega_{cav}(\hat{X})\hat{a}^\dagger\hat{a} + \hbar\Omega\hat{b}^\dagger\hat{b}. \tag{5.2}$$

Notice that the length of the cavity (and therefore the frequency of the mode) depends on the position of the mechanical oscillator, i.e. the hamiltonian represents the energy of a *parametric oscillator*. In turn, the frequency of the mode may be Taylor expanded as

$$\omega_{cav}(\hat{X}) = \frac{nc\pi}{L + \hat{X}} = \underbrace{\left(\frac{nc\pi}{L}\right)}_{\omega_{cav}} - \frac{1}{L}\underbrace{\left(\frac{nc\pi}{L}\right)}_{\omega_{cav}}\hat{X} + ... \tag{5.3}$$

For most applications it suffices to keep the first order term of the Taylor expansion. The term $G \equiv -\partial\omega_{cav}(\hat{X})/\partial\hat{X}$ is called the *optomechanical frequency shift per displacement* or the *frequency pull parameter* [155]. For the system under consideration $G = \omega_{cav}/L$. Notice that the sign of G depends on the how the sign of the mechanical displacement is defined. If a positive displacement of the mirror increases the length of the cavity (as in the present case), then G has positive sign. In the opposite case, when a positive displacement of the mirror reduces the cavity length, producing an increase of the frequency, then G has negative sign. Therefore, (5.2) becomes

$$\hat{H} = \hbar\omega_{cav}\hat{a}^\dagger\hat{a} + \hbar\Omega\hat{b}^\dagger\hat{b} - \hbar G\hat{a}^\dagger\hat{a}\hat{X}. \tag{5.4}$$

The term that multiplies $\hat{X}$ in the hamiltonian is the *radiation pressure force*, $\hat{F}_{rad} = \hbar G\hat{a}^\dagger\hat{a}$, i.e. it is proportional to the number of photons inside the cavity. It is possible to arrive to a similar result by using a classical argument. A single photon that hits a wall transfers an amount of momentum equal to $2h/\lambda$, where λ is the wavelength of the photon. If the photon stays inside the cavity for a time equal to τ, then it will hit one the mirrors $\tau \cdot c/(2L)$ times, producing a total transfer of momentum Δp equal to $\tau \cdot (ch/\lambda) \cdot (1/L) = \tau \cdot \hbar \cdot \omega_{cav}/L = \tau \cdot \hbar \cdot G$. Consequently, the radiation pressure force exerted by an intracavity photon is $\Delta p/\tau = \hbar G$. For an optical cavity, with length $L \sim$ mm, $G \sim 10 - 1000$ Hz/μm.

Notice that the interaction between the vibrating mirror and the electromagnetic field inside the cavity is a non linear process, described by the product between three operators; the complex amplitude of the field, that appears twice (three-wave mixing process), and the position of the mirror. The interaction $\propto \hat{a}^\dagger\hat{a}\hat{X}$ is called the *optomechanical interaction*. Although it has been derived here for a Fabry-Pérot cavity with a vibrating mirror, it is expected to appear every time the boundary conditions of an electromagnetic cavity are modified. For the particular system under consideration, a derivation of the optomechanical interaction from first principles can be found in [195].

The position operator can be expressed as $\hat{X} = x_0(\hat{b} + \hat{b}^\dagger)$, where x_0 are the so called

zero-point fluctuations of the mechanical oscillator, which are defined as

$$x_0 = \sqrt{\frac{\hbar}{2M\Omega}}. \tag{5.5}$$

For a mirror of the size of the order of μm $x_0 \sim 10^{-15}$ m. In terms of the operators $\hat{b}$ and $\hat{b}^\dagger$ the hamiltonian (5.4) can be written as

$$\hat{H} = \hbar\omega_{cav}\hat{a}^\dagger\hat{a} + \hbar\Omega\hat{b}^\dagger\hat{b} - \hbar g_0\hat{a}^\dagger\hat{a}(\hat{b} + \hat{b}^\dagger), \tag{5.6}$$

where the parameter $g_0 = Gx_0 \sim 1 - 10^3$ Hz is the *vacuum optomechanical coupling strength* that characterises the magnitude of the coupling between a single photon and a single phonon.

Let $|n\rangle_m$ be a number state of the mechanical harmonic-oscillator (the subscript m has been used to indicate that it is a state of the mirror, which will act as the measurement device or apparatus) and let $|m\rangle$ be a number state of the cavity, where n and m in $\mathbb{N}_0$ (the number m inside the ket, that indicates the number of photons inside the cavity, should not be confused with the subscript m, which simply indicates that the state is a mechanical state). Next, let us define a *displaced number state* [196] as

$$|n(m)\rangle_m = \exp\Big\{m\Big(\frac{g_0}{\Omega}\Big)(\hat{b}^\dagger - \hat{b})\Big\}\,|n\rangle_m\,, \tag{5.7}$$

where $m \in \mathbb{N}_0$ represents the number of photons inside the cavity, i.e. the state $|n(m)\rangle_m$ is a mechanical state that arises from the displacement of a mechanical number state $|n\rangle_m$ by m photons. For example, if $n = 0$ (the mirror is in the ground state) and the cavity contains one photon ($m = 1$), then the state $|0(1)\rangle_m$ is simply a coherent state $|g_0/\Omega\rangle_m$. Also, if there are no photons inside the cavity ($m = 0$), then $|n(0)\rangle_m = |n\rangle_m$. The eigenstates of the hamiltonian (5.6) are *product states* $|n(m)\rangle_m\,|m\rangle$, with energy levels given by

$$\hat{H}\,|n(m)\rangle_m\,|m\rangle = E_{n,m}\,|n(m)\rangle_m\,|m\rangle \quad , \quad E_{n,m} = \hbar\Big[\omega_{cav}m - \Big(\frac{g_0^2}{\Omega}\Big)m^2 + \Omega n\Big]. \tag{5.8}$$

The energy-level structure of the optomechanical system is shown in figure 5.2. Addition-

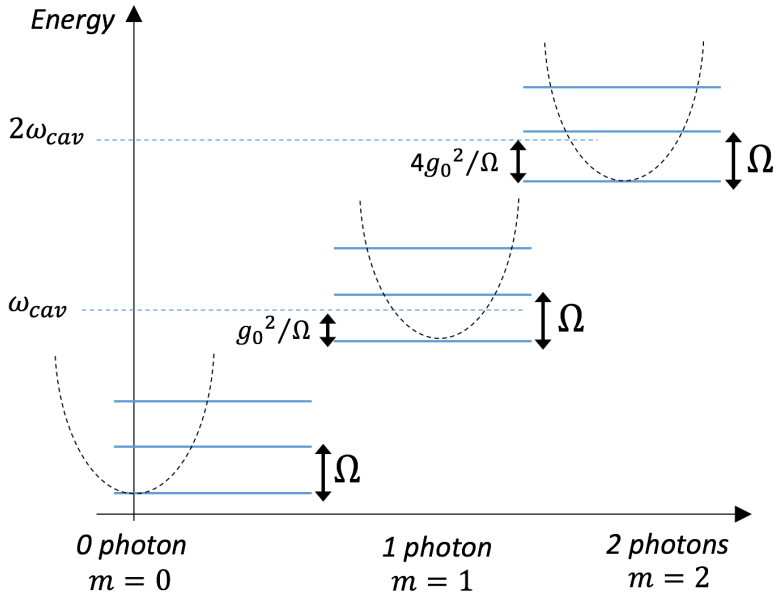


Figure 5.2: Energy levels of the optomechanical hamiltonian (5.6). For a fixed number of photons inside the cavity, the spacing between the energy levels is $\hbar\Omega$. The energy spectrum acquires an offset of $\omega_{cav}m - (g_0^2/\Omega)m^2$, where m denotes the number of photons inside the cavity. The non linear effect occurs because the frequency of the cavity depends on the number of photons, generating a Kerr-type non linearity or an effective *photon-photon interaction*.

ally, notice that the hamiltonian (5.6) preserves the total number of particles, $[\hat{H}, \hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}] = 0$.

In order to analyse the evolution generated by the optomechanical hamiltonian, notice that free energy of the electromagnetic field commutes with the other two terms in the hamiltonian (the energy of the mirror and the optomechanical interaction). Thus, the time evolution operator can be written as

$$\widehat{U}(t) = \exp\Big\{-(i/\hbar)\widehat{H}t\Big\} = \exp\{-i(\omega_{cav}t)\hat{a}^\dagger\hat{a}\}\exp\Big\{-i(\Omega t)\hat{b}^\dagger\hat{b} + i(g_0t)\hat{a}^\dagger\hat{a}(\hat{b} + \hat{b}^\dagger)\Big\}. \quad (5.9)$$

It is possible to employ the sometimes called *polaron transformation* [178] $\widehat{S} = \exp\Big\{(g_0/\Omega)(\hat{a}^\dagger\hat{a})(\hat{b}^\dagger - \hat{b})\Big\}$ in order to work out the previous expression. This transformation corresponds to a (mechanical) *displacement operator* that produces a displacement in phase space of $(g_0/\Omega)\hat{a}^\dagger\hat{a}$. The *real* parameter $g \equiv g_0/\Omega$ corresponds to the displacement produced by a single photon in the (mechanical) phase space, i.e. it is the displacement of the equilibrium position in units of x_0. For an oscillator with frequency $\Omega \sim 1 - 100$ MHz, g is typically a small

parameter with values in the range $10^{-5} - 10^{-1}$. Using the polaron transformation, the time evolution operator can be expressed as

$$\begin{aligned}
\widehat{U}(t) &= \widehat{S}\widehat{S}^{\dagger}\widehat{U}(t)\widehat{S}\widehat{S}^{\dagger} \\
&= \exp\{-i\omega_{cav}t\hat{a}^{\dagger}\hat{a}\}\exp\{i(g\hat{a}^{\dagger}\hat{a})^2\}\widehat{S}\exp\Big\{-i\Omega t\hat{b}^{\dagger}\hat{b}\Big\}\widehat{S}^{\dagger} \\
&= \exp\{-i\omega_{cav}t\hat{a}^{\dagger}\hat{a}\}\exp\{i(g\hat{a}^{\dagger}\hat{a})^2\}\widehat{S}\exp\Big\{-i\Omega t\hat{b}^{\dagger}\hat{b}\Big\}\widehat{S}^{\dagger}\exp\Big\{i\Omega t\hat{b}^{\dagger}\hat{b}\Big\}\exp\Big\{-i\Omega t\hat{b}^{\dagger}\hat{b}\Big\} \\
&= \exp\{-i\omega_{cav}t\hat{a}^{\dagger}\hat{a}\}\exp\{i(g\hat{a}^{\dagger}\hat{a})^2\}\widehat{S}\exp\Big\{g\hat{a}^{\dagger}\hat{a}(\hat{b}e^{i\Omega t} - \hat{b}^{\dagger}e^{-i\Omega t})\Big\}\exp\Big\{-i\Omega t\hat{b}^{\dagger}\hat{b}\Big\} \quad (5.10)
\end{aligned}$$

Notice that in this derivation the well known operator expansion for two non-commuting operators $\hat{A}$ and $\hat{B}$,

$$\exp\Big\{\alpha\hat{A}\Big\}\hat{B}\exp\Big\{-\alpha\hat{A}\Big\} = \hat{B} + \alpha[\hat{A},\hat{B}] + \frac{\alpha^2}{2!}[\hat{A},[\hat{A},\hat{B}]] + ..., \quad (5.11)$$

has been used, together with the property $\hat{T}f(\{\hat{Z}_i\})\hat{T}^{\dagger} = f(\{\hat{T}\hat{Z}_i\hat{T}^{\dagger}\})$, for any unitary transformation $\hat{T}$ and any analytic function f of a set of operators $\{\hat{Z}_i\}$ [197]. Using the property $\exp\Big\{\hat{A}+\hat{B}\Big\} = \exp\Big\{\hat{A}\Big\}\exp\Big\{\hat{B}\Big\}\exp\Big\{-[\hat{A},\hat{B}]/2\Big\}$, for any pair of operators $\hat{A},\hat{B}$ that satisfy $[\hat{A},[\hat{A},\hat{B}]] = [\hat{B},[\hat{A},\hat{B}]] = 0$, it is straightforward to show that

$$\widehat{S}\exp\Big\{g\hat{a}^{\dagger}\hat{a}(\hat{b}e^{i\Omega t} - \hat{b}^{\dagger}e^{-i\Omega t})\Big\} = \exp\Big\{\hat{a}^{\dagger}\hat{a}[\hat{b}^{\dagger}\varphi^*(t) - \hat{b}\varphi(t)]\Big\}\exp\{-i(g\hat{a}^{\dagger}\hat{a})^2\sin(\Omega t)\}, \quad (5.12)$$

where $\varphi(t) = g(1-e^{-i\Omega t})$ and $\varphi^*(t)$ is its complex conjugate. By using this result in (5.10) it is possible to express the time evolution operator as

$$\widehat{U}(t) = \exp\{-i\omega_{cav}t\hat{a}^{\dagger}\hat{a}\}\exp\{i\phi(t)(\hat{a}^{\dagger}\hat{a})^2\}\exp\Big\{\hat{a}^{\dagger}\hat{a}[\hat{b}^{\dagger}\varphi(t) - \hat{b}\varphi^*(t)]\Big\}\exp\Big\{-i\Omega t\hat{b}^{\dagger}\hat{b}\Big\}, \quad (5.13)$$

where $\phi(t) = g^2[\Omega t - \sin(\Omega t)]$. The first term (from left to right) corresponds to the free evolution of the radiation field. The second term adds a phase that depends quadratically on the number of photons, which shows that the optomechanical coupling generates an effective Kerr nonlinearity or *photon-photon* interaction. This occurs because the frequency of the cavity ω_{cav} depends on the position of the mechanical resonator, which in turn depends on the number of photons inside the cavity. Therefore, ω_{cav} depends finally on the

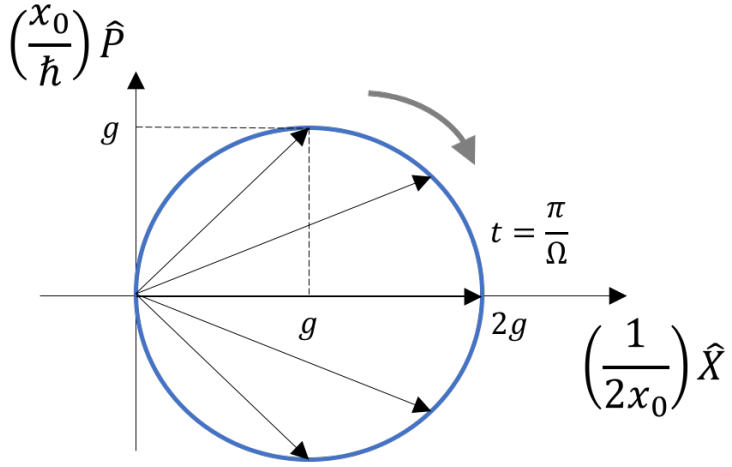


Figure 5.3: The (closed) evolution of the initial state $|1\rangle|0\rangle_m$ under the hamiltonian (5.6) is represented in phase space. Each point of the blue line (that represents a displaced harmonic oscillator) has the expectation value of the x-quadrature as the horizontal coordinate and the y-quadrature as the vertical coordinate, both calculated in the evolved state at time t. The x-quadrature is defined as $(\hat{a}^\dagger + \hat{a})/2$ and the y-quadrature corresponds to $i(\hat{a}^\dagger - \hat{a})/2$. At half the vibrational period the displacement of the average position is maximum and equal to $4x_0 g$.

number of photons, which generates the term $(\hat{a}^\dagger\hat{a})^2$. The third term in (5.13) entangles the photons with the mirror and the last term corresponds to the free evolution of the mirror.

As a simple example assume that cavity and the mirror start in the initial state $|1\rangle|0\rangle_m$ (the cavity contains one photon and the mirror is in the ground state). After a time t the initial state will evolve under (5.13) to $|1\rangle|\varphi(t)\rangle_m$, i.e. the mirror evolves into a coherent state with complex amplitude of $\varphi(t)$. This state is represented in figure 5.3, which shows that the equilibrium position of the ground state is displaced by $(2x_0)g$. It is possible to derive this result using a classical argument. The mirror is subjected to a constant driving force equal to $\hbar G$. Thus, the displacement of the equilibrium position will be equal to $\hbar G/(M\Omega^2) = (2x_0)g$. Additionally, notice that the maximum displacement of the position occurs at half the vibrational periods, i.e. at times that are odd multiples of π/Ω.

In this work we will consider a Fabry-Pérot cavity with a vibrating mirror placed in the middle of the cavity as shown in figure 5.4. This setup is sometimes called the "membrane in the middle" configuration [155]. The fields in both sides of the cavity couple to the vibrating mirror with a strength $g_0 = x_0 G$, i.e. the same as in the previous case of a

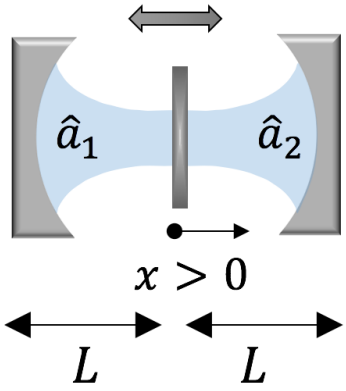


Figure 5.4: A vibrating mirror is placed in the middle of a Fabry-Pérot cavity. The fields interacts with the mirror via the radiation pressure force $\widehat{F} = \hbar G(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)$. In our model there is no photon hopping between both sides.

one-sided cavity. In this setup, typically, the fields interact with each other via photon hopping. However, in our model it will be assumed that this interaction is neglectable, i.e. that the mirror in the middle is perfectly reflecting. The hamiltonian of the system is therefore analogous to the previous hamiltonian (5.6),

$$\widehat{H} = \hbar\omega_{cav}(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2) + \hbar\Omega \hat{b}^\dagger \hat{b} - \hbar g_0(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)(\hat{b}^\dagger + \hat{b}). \tag{5.14}$$

The operators $\hat{a}_1$ ($\hat{a}_1^\dagger$) and $\hat{a}_2$ ($\hat{a}_2^\dagger$) are boson annihilation (creation) operators for the cavity modes of the left and right sides, respectively. Notice that the frequency of the left and right modes is the same, because the mirror is placed in the middle of the cavity. Note also that the radiation pressure force is now proportional to the *difference of photons* between both sides of the cavity, because the movement of the mirror in one direction shortens the length of one side of the cavity while enlarges the other.

The eigenstates of the two-sided optomechanical system are product states,

$$|n\rangle_1 |m\rangle_2 |k(n-m)\rangle_m , \tag{5.15}$$

where $|n\rangle_1$ and $|m\rangle_2$ are number states of the sides 1 and 2 of the cavity, respectively, and $|k(n-m)\rangle_m$ is a k mechanical number state, displaced by $m-n$ photons, according to (5.7). The energy levels are given by $E_{n,m,k} = \hbar\Big[\omega_{cav}(n+m) - \Big(\frac{g_0^2}{\Omega}\Big)(n-m)^2 + \Omega k\Big]$. The

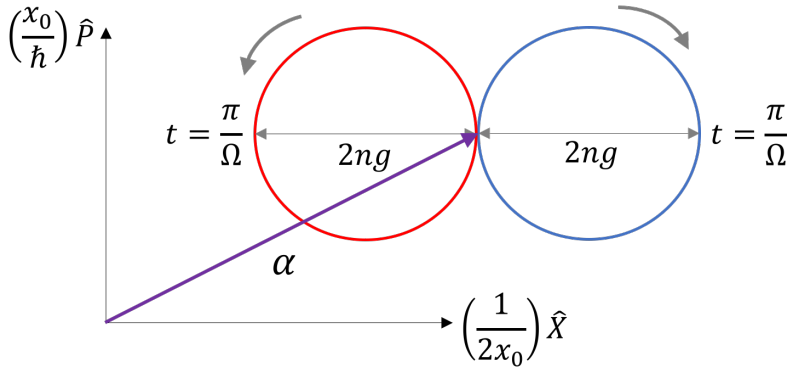


Figure 5.5: The evolution of the initial states $|n\rangle_1 |0\rangle_2 |\alpha\rangle_m$ (blue circle) and $|0\rangle_1 |n\rangle_2 |\alpha\rangle_m$ (red circle) under the hamiltonian (5.14) is represented in phase space. When n photons are in one side of the cavity, the mirror oscillates around a new *equilibrium position*. At half the vibrational period the displacement of the *average position* is maximum and equal to $4x_0 g$.

evolution generated by the hamiltonian (5.14) is analogous to (5.13), but now the phase associated to the Kerr nonlinearity depends quadratically on the difference of photons between both sides of the cavity,

$$\begin{aligned}\widehat{U}(t) = \exp\Big\{-i\omega_{cav}t(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2)\Big\}\exp\Big\{i\phi(t)(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)^2\Big\}\times \\ \exp\Big\{(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)\cdot[\hat{b}^\dagger\varphi(t) - \hat{b}\varphi^*(t)]\Big\}\exp\Big\{-i\Omega t\hat{b}^\dagger\hat{b}\Big\}.\end{aligned} \tag{5.16}$$

The evolution of states of the form $|n\rangle_1 |0\rangle_2 |\alpha\rangle_m$ or $|0\rangle_1 |n\rangle_2 |\alpha\rangle_m$, i.e. with n photons either in the right or the left side of the cavity and the mirror in a coherent state with amplitude α, is shown in figure 5.5.

5.2 Optomechanical system in an interferometer

The optomechanical hamiltonians (5.6) and (5.14) generate entanglement between "microscopic" degrees of freedom (the cavity modes) and "macroscopic" degrees of freedom (the center of mass of a vibrating mirror). The *optomechanical entanglement* can be achieved by putting the cavity in one of the arms of a Michelson interferometer [179], as shown in figure 5.6. The mirror needs to be cool down near to its ground state. Once the photon passes through the entrance beamsplitter it is put in a coherent superposition between both paths and therefore the initial state of the cavity and the mirror is described

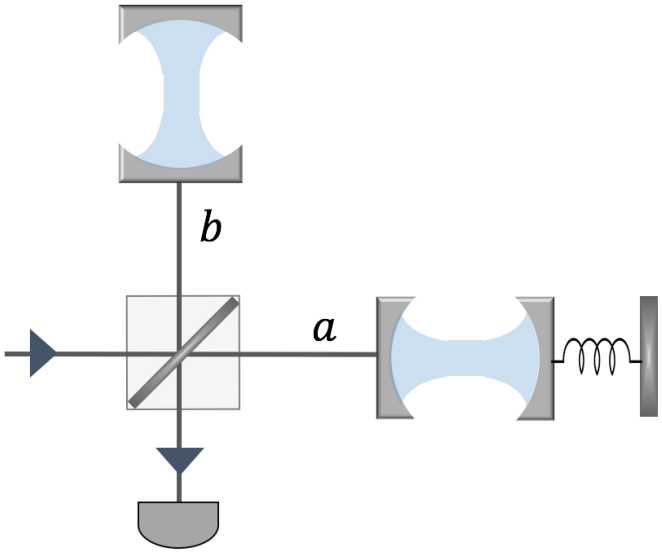


Figure 5.6: Michelson interferometer with an optomechanical system in one of the arms creates optomechanical entanglement.

by the state

$$\left(\frac{1}{\sqrt{2}}\,|1\rangle_a\,|0\rangle_b + \frac{1}{\sqrt{2}}\,|0\rangle_a\,|1\rangle_b\right)|0\rangle_m\,, \tag{5.17}$$

where the subscript a or b labels the corresponding path of the interferometer. Under the hamiltonian (5.6) the state evolves to

$$\frac{e^{i\phi(t)}}{\sqrt{2}}\,|1\rangle_a\,|0\rangle_b\,|\varphi(t)\rangle_m + \frac{1}{\sqrt{2}}\,|0\rangle_a\,|1\rangle_b\,|0\rangle_m\,, \tag{5.18}$$

which is an entangled state between the mirror and the photon. Nevertheless, after a full mechanical period, the mirror returns to the initial position and disentangles from the radiation field. By tracing over the mechanical degrees of freedom one obtains the reduced density matrix of the photon, that corresponds to a mixed state, except at multiples of the mechanical period (because at those instants the mirror is not entangled with the photon). Consequently, the interferometric visibility

$$\mathcal{V} = \exp\{-|\varphi(t)|^2/2\} = \exp\{-g^2[1-\cos(\Omega t)]\} \tag{5.19}$$

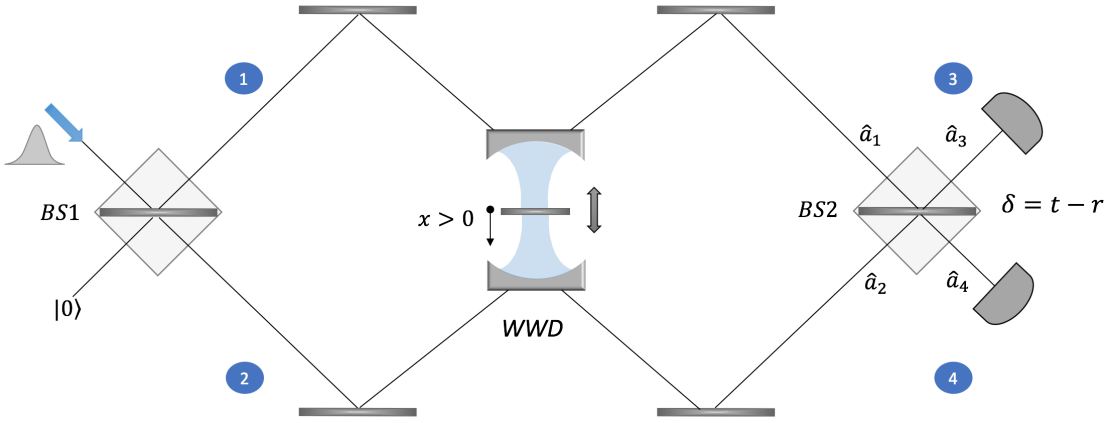


Figure 5.7: Mach-Zehnder interferometer with an optomechanical system in both arms, which acts as a which way detector (WWD). The second beam splitter is not balanced, i.e. the transmission and reflection coefficients are not necessarily equal. The level of unbalance is described by a parameter δ (see equation 5.25).

will exhibit *decays* and *revivals*. The visibility will be zero at odd multiples of half the mechanical period (when the entanglement is maximum) but the photon coherence will revive at multiples of the period, when the state of the photon is pure again. Therefore, the observation of the visibility, as a function of the time spent in the cavity, allows to study the entanglement between light and a vibrating mirror. The optomechanical entanglement was first studied in [178, 199]. Since the mirror is subject to decoherence and dissipation, the decays and revivals of the visibility will be eventually spoiled. Thus, by observing the visibility it is possible to test different models that describe the decoherence of a superposition state of the mirror [177, 179, 180], which is a "large" object formed by $\sim 10^{14}$ atoms.

In our work we will employ a Mach-Zehnder interferometer to produce optomechanical entanglement. As will be explained in the following sections of this chapter, the small displacement caused by a single photon can be largely increased by using the so called "dark port post-selection" and the weak value amplification effect.

5.3 Description of the experiment

The Mach-Zehnder interferometer is shown in figure 4.1. The action of a balanced beamsplitter is described by the unitary transformation (4.48). Now, we will assume

that the second beam splitter, which sometimes is also called the *beam merger* [200], is unbalanced. Figure 5.7 shows a Mach-Zehnder with a second unbalanced beamsplitter. The relationship between the input and output modes is given by

$$\begin{aligned} \hat{a}_3 &= r\hat{a}_1 + t\hat{a}_2 \\ \hat{a}_4 &= t\hat{a}_1 - r\hat{a}_2, \end{aligned} \tag{5.20}$$

where t and r are (real and positive) transmittances an reflectances of the beamsplitter, respectively. The minus sign in the second equation is necessary in order for the transformation to be unitary. Physically, it means that the field reflected in lower side of the beamsplitter acquires a phase of π. In ideal conditions the beamsplitter preserves the energy, i.e. $t^2 + r^2 = 1$. Hence, $0 \leqslant r \leqslant 1$ and $0 \leqslant t \leqslant 1$. See appendix (C) for details regarding the quantum mechanical description of a beam splitter. The transformation of the fields (5.20) allows to transform states between the inner and outer paths of the interferometer, as it is described in the following example.

$$\begin{aligned} |0\rangle_3 |1\rangle_4 &= \hat{a}_4^\dagger |0\rangle_3 |0\rangle_4 \\ &= (t\hat{a}_1 - r\hat{a}_2) |0\rangle_1 |0\rangle_2 \\ &= t |1\rangle_1 |0\rangle_2 - r |0\rangle_1 |1\rangle_2 . \end{aligned} \tag{5.21}$$

The subindices of the states denote the corresponding path in the interferometer. Notice that in the second step of (5.21) the beam splitter transformation (5.20) was employed, while the vacuum states outside and inside the interferometer were identified. Consequently, counting a photon in the arm 4 is equivalent to select the state

$$|\psi_f\rangle = t |1\rangle_1 |0\rangle_2 - r |0\rangle_1 |1\rangle_2 \tag{5.22}$$

inside the interferometer. Analogously, counting a photon in the arm 3 is equivalent to the selection of the state

$$\left|\psi_f^\perp\right\rangle = r |1\rangle_1 |0\rangle_2 + t |0\rangle_1 |1\rangle_2 . \tag{5.23}$$

If one photon enters the interferometer through one of the input ports, while the second port is left in the vacuum state, the field will evolve along the first (balanced) beam splitter and enter the interferometer. The state of the field inside the interferometer will correspond to coherent a superposition of the two paths, described by

$$|\psi_i\rangle = \frac{|1\rangle_1 |0\rangle_2 + |0\rangle_1 |1\rangle_2}{\sqrt{2}}. \tag{5.24}$$

Therefore, in the absence of the optomechanical system, the probability to count the photon in the arm 4 is given by $|\langle\psi_i|\psi_f\rangle|^2 = (t-r)^2/2$, while the probability to detect the photon in the arm 3 corresponds to $|\left\langle\psi_i\middle|\psi_f^{\perp}\right\rangle|^2 = (r+t)^2/2 = 1-(t-r)^2/2$. It is useful to define a parameter δ that quantifies the level of unbalance of the beamsplitter as

$$\delta = t - r. \tag{5.25}$$

If $\delta = 0$ the beamsplitter will be balanced, as it is usual in most cases. In this situation, the photon will be detected with certainty in the arm 3, while no light will go through the other arm. This occurs due to the *wave behaviour* of the photon, i.e. because there is *constructive interference* in the arm 3 and *destructive interference* in the arm 4 (recall that the photon reflected in the lower side of the beamsplitter acquires a phase of π).

If δ is a small parameter, i.e. the beam splitter is slightly unbalanced, little light goes through the arm 4, while most light passes through the arm 3. For this reason, in this scenario, the arm 3 is typically called the *"bright port"* and the other outer arm is the *"dark port"*.

When $\delta = 1$ (the beam splitter transmits all the light) or $\delta = -1$ (the beamsplitter acts as perfectly reflecting mirror), there is a 50% probability to detect the photon in each arm. This configuration shows the *particle behaviour* of the photon. Indeed, when $\delta = \pm 1$ the second beam splitter plays no role, as if no second beam splitter was present. Thus, each photon counter will "click" every time the photon has gone through the corresponding path. Since the probabilities of "clicking" of both detectors are the same, the photon behaves like a particle that enters the interferometer and randomly takes one of the two paths.

In our setup a two-sided optomechanical cavity will be placed inside the interferometer, as shown in figure 5.7. As was explained in the previous section, a single photon will produce a maximum displacement of the average position of $4x_0 g$ at half the vibrational period (see figures 5.3 and 5.5). Placing the optomechanical system inside the interferometer will allow us to employ *weak value amplification* to enlarge the displacement caused by a single photon. This effect will put all the Fisher information about the parameter g in the post-selected events, which may be useful when we need to estimate g from a "small" amount of post-selected photons. As will be explained below a "post-selected photon" is simply a photon detected in the dark port. Additionally, if the experiment is performed in the presence of noise with long correlation times, then the Fisher information can be further increased as compared to the scenario without post-selection, i.e. when all the photons are used to estimate g, both those detected in the bright port and those detected in the dark port. The amplification of the displacement caused by a single photon on an optomechanical Fabry-Pérot cavity placed in one of the arms of a Mach-Zhender has also been studied in [201, 202, 203].

Before we turn into the description of our experimental proposal, it is worth mentioning that the optomechanical system acts as a *which way detector* [204], i.e. as a device that stores information about the path taken by the photon. It the photon has gone through the path 1, then we expect that the position of the mirror will be displaced along the positive direction of the x axis. On the contrary, if the photon goes through the path 2, the mirror should be displaced in the opposite direction. Analogously, if the photon behaves like a wave, then the position of the mirror should not be affected, since the photon in one side of the cavity "pushes" the mirror in one direction, while the photon in the other side exerts a force in the opposite direction.

The experimental proposal works as follows. The oscillator should be prepared in the ground state [171, 169, 205, 206] or in a thermal equilibrium state. The initial state of the mirror is denoted by ρ_m.

On the other hand, one single photon should be injected into the system through the input port, while the other port is left unused. After passing through the first balanced beamsplitter, the photon will be in a superposition of both paths and enter the cavity.

The absorption probability can be made close to unity when the shape of the photon is the "time reversed version" of a photon that is spontaneously emitted from the cavity [207, 208]. Therefore, once the photon has entered the cavity it will be in a superposition of both sides, and the initial state of the cavity and the mirror will be $|\psi_i\rangle \rho_m \langle\psi_i|$. This state will evolve under the time evolution operator (5.16). For NOON states [209], which are states of the form

$$\frac{|N\rangle_a |0\rangle_b + e^{iN\Theta} |0\rangle_a |N\rangle_b}{\sqrt{2}}, \quad N \in \mathbb{N}_0, \tag{5.26}$$

the free evolution and the "Kerr term" will simply add a global phase factor that can be disregarded. Consequently, after a time t the photon and the mirror will evolve into the state

$$\widehat{U}_e(t) |\psi_i\rangle \left(e^{-i\Omega t \hat{b}^\dagger \hat{b}} \rho_m e^{i\Omega t \hat{b}^\dagger \hat{b}} \right) \langle\psi_i| \widehat{U}_e^\dagger(t), \tag{5.27}$$

where $\widehat{U}_e(t) = \exp\Big\{(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2) \cdot [\hat{b}^\dagger \varphi(t) - \hat{b} \varphi^*(t)]\Big\}$ is the term in (5.16), responsible for the entanglement between the mirror and the photon. Note that the operator $\widehat{U}_e(t)$ displaces by $(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)\varphi(t)$, i.e. the displacement depends on the difference of photons between both sides of the cavity. From (4.41), the quantum Fisher information about the parameter g in the state (5.27) is given by

$$\mathcal{I}_Q(g) = 8[1 - \cos(\Omega t)]\Big(1 + 2\Big\langle \hat{b}^\dagger \hat{b} \Big\rangle\Big), \tag{5.28}$$

where the expectation value of the number of phonons is calculated over the initial state of the mirror ρ_m. Notice that the quantum Fisher information depends on time; at odd multiples of half the vibrational period (when the displacement of the average position is maximum) the quantum Fisher information achieves its maximum value, while at the vibrational periods (when the mirror and the photon are not entangled) the Fisher information is zero.

Once the photon leaks out from the cavity, it will be detected at one of the photon counters (the interferometer is consider to be ideal, i.e. without lost of photons). Successful

post-selection occurs when a photon is detected in the dark port. Only when post-selection is successful, the position of the oscillator should be observed [162, 184]. The cases when post-selection fails should be simply disregarded. In other words, the position of the mirror will be observed, conditioned on a photon detected at the dark port.

Under certain conditions, described in the next section, the displacement caused by a single photon will be proportional to the *weak value* of the difference of photons between both sides of the cavity. As will be shown in the next section, the weak value amplification effect will enlarge the displacement, which can be useful for the estimation of optomechanical parameters. In every repetition of the experiment, the state of the mirror should be reinitialized. The experiment is designed for single photons with frequency ω_{cav}. The photons are considered to be nearly monochromatic, i.e. with an spectral bandwidth ϵ much smaller than the cavity decay rate γ.

5.4 Weak value amplification

Assume that the photon is detected at the dark port at time t, i.e. the state (5.22) is post-selected. Then, the unnormalized state of the mirror is

$$\langle\psi_f|\,\widehat{U}_e(t)\,|\psi_i\rangle\left(e^{-i\Omega t\hat{b}^\dagger\hat{b}}\rho_m e^{i\Omega t\hat{b}^\dagger\hat{b}}\right)\langle\psi_i|\,\widehat{U}_e^\dagger(t)\,|\psi_f\rangle\,. \tag{5.29}$$

In order to employ *weak values* to enlarge the displacement produced by the photon, let us expand the "evolution operator" $\widehat{U}_e(t)$ to first order, namely,

$$\begin{aligned}\langle\psi_f|\,\widehat{U}_e(t)\,|\psi_i\rangle &\approx \langle\psi_f|\psi_i\rangle\Big\{1+(\hat{a}_1^\dagger\hat{a}_1-\hat{a}_2^\dagger\hat{a}_2)_w\cdot[\hat{b}^\dagger\varphi(t)-\hat{b}\varphi^*(t)]\Big\}, &(5.30)\\ &\approx \langle\psi_f|\psi_i\rangle\exp\Big\{(\hat{a}_1^\dagger\hat{a}_1-\hat{a}_2^\dagger\hat{a}_2)_w\cdot[\hat{b}^\dagger\varphi(t)-\hat{b}\varphi^*(t)]\Big\}. &(5.31)\end{aligned}$$

The term $(\hat{a}_1^\dagger\hat{a}_1-\hat{a}_2^\dagger\hat{a}_2)_w$ corresponds to the weak value of the difference of photons between the initial state $|\psi_i\rangle$ (5.24) and the final state $|\psi_f\rangle$ (5.22). It is given by

$$(\hat{a}_1^\dagger\hat{a}_1-\hat{a}_2^\dagger\hat{a}_2)_w=\frac{t+r}{t-r}=\frac{\sqrt{2-\delta^2}}{\delta}. \tag{5.32}$$

Since the transmission and reflection coefficients are real values, the weak value is also a real number. Both approximations, (5.30) and (5.31), can be made when $g|(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)|_w$ is small as compared to the inverse of the standard deviation of the rotated quadrature $i(1-e^{-i\Omega t})\hat{b}^\dagger - i(1-e^{i\Omega t})\hat{b}$, computed over the initial state of the mirror,

$$\begin{aligned} g|(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)|_w &\ll \left\{Var\Big[(1-e^{-i\Omega t})\hat{b}^\dagger - (1-e^{i\Omega t})\hat{b}\Big]\right\}^{-1/2} \\ &= \left\{2\Big[1-\cos(\Omega t)\Big]\Big(1+2\left\langle\hat{b}^\dagger\hat{b}\right\rangle\Big)\right\}^{-1/2}. \end{aligned} \tag{5.33}$$

Notice that near the mechanical periods, when there is little entanglement between the photons and the mirror, the bound on the right hand side gets larger. On the contrary, at half the vibrational periods, the bound reaches its minimum value. From the state (5.29) and the first order approximation (5.31), it is clear that the normalized state of the mirror is given by

$$\rho_m(t) = e^{(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w[\hat{b}^\dagger\varphi(t) - \hat{b}\varphi^*(t)]}\left(e^{-i\Omega t\hat{b}^\dagger\hat{b}}\rho_m e^{i\Omega t\hat{b}^\dagger\hat{b}}\right)e^{-(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w\cdot[\hat{b}^\dagger\varphi(t) - \hat{b}\varphi^*(t)]}. \tag{5.34}$$

While the probability of successful post-selection, $P(f)$, is $|\langle\psi_f|\psi_i\rangle|^2 = \delta^2/2$. Notice that the post-selection probability is the same as if no optomechanical system was employed. The *conditioned probability* (given that post-selection was successful) to read the position x of the mirror corresponds to

$$P(x|f) = \langle x|\,\rho_m(t)\,|x\rangle\,. \tag{5.35}$$

The Fisher information contained in this distribution depends on the initial state of the mirror (ρ_m). We will study two cases. First, we assume that the mirror starts in a pure state (the ground state). The results presented in the previous chapter (see section 4.9) will be applied to this situation. Then, we study the case of a mirror that starts in a thermal state.

5.4.1 Oscillator initially in the ground state

Let us consider first the case when the oscillator is cooled down to its ground state, i.e. $\rho_m = |0\rangle_m\langle 0|$. In this scenario, the conditioned density becomes a Gaussian function,

$$P(x|f) = \frac{1}{\sqrt{2\pi}x_0}\exp\left\{-\frac{\Big(x - 2x_0 g(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w[1-\cos(\Omega t)]\Big)^2}{2x_0^2}\right\}. \tag{5.36}$$

Using (4.18) it is easy to see that the Fisher Information about the parameter g contained in this distribution corresponds to

$$\mathcal{I}(g|f) = 4(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w^2[1-\cos(\Omega t)]^2. \tag{5.37}$$

Therefore, according to the result (4.74) presented in chapter 4, the Fisher information in the measurement protocol that includes post-selection is

$$\mathcal{I}(g) = P(f)\mathcal{I}(g|f) = 2[1-\cos(\Omega t)]^2(2-\delta^2). \tag{5.38}$$

The quantum Fisher information (5.28) is achieved at half the period and when $\delta \to 0$, i.e. when the weak value (5.32) is indefinitely large. Recall, nevertheless, that the weak value is restricted by the condition (5.33) and therefore can not become arbitrarily large.

On the other hand, notice that the evolution operator $\widehat{U}_e(t)$ couples the difference of photons with a rotated quadrature $i(1-e^{-i\Omega t})\hat{b}^\dagger - i(1-e^{i\Omega t})\hat{b}$ (an hermitian operator), with a coupling strength of g. The weak value of the quadrature is given by

$$\frac{\langle x|\, i(1-e^{-i\Omega t})\hat{b}^\dagger - i(1-e^{i\Omega t})\hat{b}\,|0\rangle}{\langle x|0\rangle} = i(1-e^{-i\Omega t})\frac{\psi_1(x)}{\psi_0(x)}, \tag{5.39}$$

where $\psi_0(x)$ and $\psi_1(x)$ are the wave functions of the ground state and the first exited state of the mirror, respectively. Both are real functions and therefore the weak value is always a complex value, except at half the period, when it becomes purely imaginary. In this case, the weak value of the system variable (the difference of photons) is a real value, while the weak value of the apparatus variable (the rotated quadrature) is purely imaginary.

Table 5.1: As the magnitude of the dark port post-selection parameter δ decreases, the magnitude of the weak value of the difference of photons is increased since the weak value $\propto \delta^{-1}$. At the same time the probability of post-selection decreases, because it is $\propto \delta^2$.

$\lvert\delta\rvert$	$\lvert(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)_w\rvert$	Post-selection Probability (%)
0.5	2.6	13
0.4	3.4	8
0.3	4.6	5
0.2	7.0	2
0.1	14.1	1
0.09	15.7	0.4

As was explained in the previous chapter (see section 4.9.2), this is a sufficient condition in order to extract the quantum Fisher information from the *conditioned state of the mirror* (5.34). Indeed, at half the period, the quadrature is proportional to momentum of the mirror and therefore, because the position (the conjugate variable) is observed, the weak value of the momentum will be imaginary, as long as the wave function of the meter is real. This fact can be seen from equation (4.63). Consequently, at an arbitrary instant of time, the conjugate quadrature of the meter should be observed in order to reach the quantum Fisher information (5.28).

Moreover, restricted to the space of single photons, i.e. for NOON states (5.26) with $N = 1$, the operator $\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2$ behaves as a $\hat{\sigma}_z$ operator, the states $|1\rangle_1 |0\rangle_2$ and $|0\rangle_1 |1\rangle_2$ being its eigenvectors. The initial state (5.24) is therefore an eigenstate of $\hat{\sigma}_x$, i.e. a qubit state (4.97) with $\theta_i = \pi/2$ and $\phi_i = 0$. When $\delta \ll 1$, the final state is a qubit with polar angle $\theta_f = \pi/2 - \sqrt{2}\delta$ and azimuthal angle $\phi_f = \pi$. The post-selected state is in accordance with the strategy described in section (4.9.2) for a qubit and a meter in a Gaussian (pure) state. Indeed, this strategy generates weak values of the form (4.106), with the identification $\epsilon = \delta/\sqrt{2}$, and achieves the quantum Fisher information in the limit $\delta \to 0$. The post-selection probability $\delta^2/2$ and the weak value of the difference of photons are summarized in table 5.1. Figure 5.8 shows the weak value and the amplification effect due to the post-selection for different values of δ.

If the experiment is repeated N times, from which there will be on average $M = N|\langle\psi_f|\psi_i\rangle|^2$ successfully post-selected events, the *maximum likelihood estimator* (4.6) of

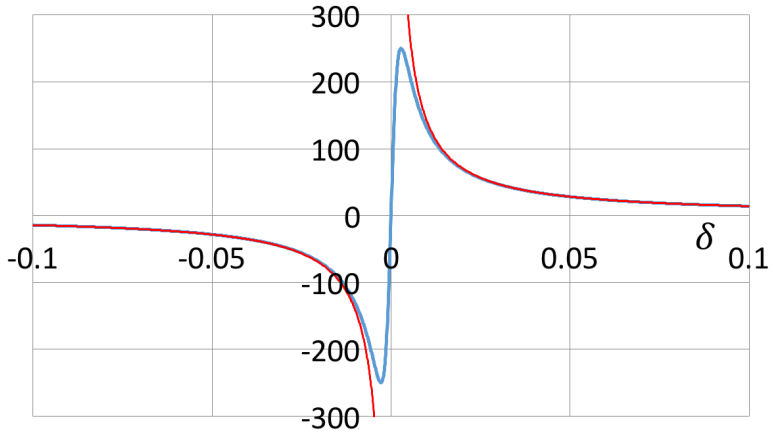


Figure 5.8: The weak value (red curve) and the amplification factor (blue curve) plotted against the post-selection parameter δ. The amplification factor corresponds to the displacement of the measurement device divided by $4gx_0$. The weak value equals the amplification factor when $g \ll \delta$ (weak measurement regime). The amplification factor achieves its maximum value outside this regime, when $g \sim \delta$. When δ is much smaller than g notice that there is no amplification affect. In this case $g = 10^{-3}$.

the parameter g and its variance (4.8) correspond to

$$\hat{g}_{ML} = \frac{1}{x_0[1-\cos(\Omega t)]\delta\sqrt{2-\delta^2}} \frac{\sum_{i=1}^{M} x_i}{N}, \tag{5.40}$$

$$\left\langle \Delta\hat{g}^2_{ML} \right\rangle = \left\{ \frac{1}{2[1-\cos(\Omega t)]^2(2-\delta^2)} \right\} \frac{1}{N}, \tag{5.41}$$

respectively. Notice that the estimator is the sample mean, corrected by a factor that depends on the prior knowledge of δ and x_0. The maximum likelihood estimator is an unbiased estimator of g. Notice that the variance corresponds to the inverse of the Fisher information (5.38), i.e. the estimator saturates the Crámer Rao inequality (4.19) and scales as N^{-1} (standard quantum limit). As was explained previously, the estimator will also saturate the quantum Crámer Rao inequality (4.37) at half the vibrational period, when the quantum Fisher information (5.28) is achieved. The signal to noise ratio (4.22) is

$$SNR_{0,w} = g[1-\cos(\Omega t)]\sqrt{2N(2-\delta^2)}. \tag{5.42}$$

The subscript 0 indicates that the initial state of the mirror is the ground state and that the measurement is weak, i.e. that it relies on the validity of the approximations (5.30)

and (5.31). Note that, in a single event, i.e. when $N = 1$, the $SNR_{0,w} < 1$. In order for the $SNR_{0,w}$ to be greater than one, the experiment should be repeated N times, where $N \gg g^{-2}$. When no post-selection is performed, the SNR is of the same order, i.e. the experiment should be repeated the same times, $N \gg g^2$, in order to estime g precisely. The difference between both strategies lies in the amount of photons employed for estimation; with post-selection fewer photons are needed to perform the estimation.

Finally, it is interesting to study measurements of any strength. For any value of g, it is easy to see that the state of the mirror, conditioned on the successful post-selection of the state $|\psi_f\rangle$, corresponds to

$$|\psi(t)\rangle = \frac{1}{\mathcal{N}}\left\{\left[\frac{1+(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w}{2}\right]|\varphi(t)\rangle + \left[\frac{1-(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w}{2}\right]|-\varphi(t)\rangle\right\}, \quad (5.43)$$

where $\mathcal{N}$ is a normalization factor and $|\pm\varphi(t)\rangle$ are coherent states. The mirror is therefore in a superposition of states that are not necessarily orthogonal. Since $\varphi(t) \ll 1$, a first-order expansion of the coherent states allows to express the previous state as

$$|\psi(t)\rangle = \frac{1}{\sqrt{1+|\varphi(t)|^2(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w^2}}\Big[\,|0\rangle - \varphi(t)(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w\,|1\rangle\,\Big]. \quad (5.44)$$

Consequently, this setup generates non classical mechanical states, which consist of a superposition of the ground state and the one-phonon state. In the weak measurement regime, defined by (5.33), this state becomes $|0\rangle + [\varphi(t)(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2)_w]\,|1\rangle$, i.e. a "weak" superposition between the ground state and the first excited state. The Wigner function of this state, shown in figure 5.9, has no negative part and thus can be understood as a classical state. As it is pointed out in [210], this slight superposition between the ground state and the one-phonon state can give rise to a large amplification effect, without changing the shape of initial wave function of the measuring device, but only translating it. On the other hand, outside the weak measurement regime, e.g when $\delta \sim g$, the state of the oscillator is $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, an equal superposition of the ground state and the first excited state (see figure 5.10). In this situation, the wave function of the oscillator is no longer Gaussian, due to the larger weight of the one-phonon state. The generation of non Gaussian states

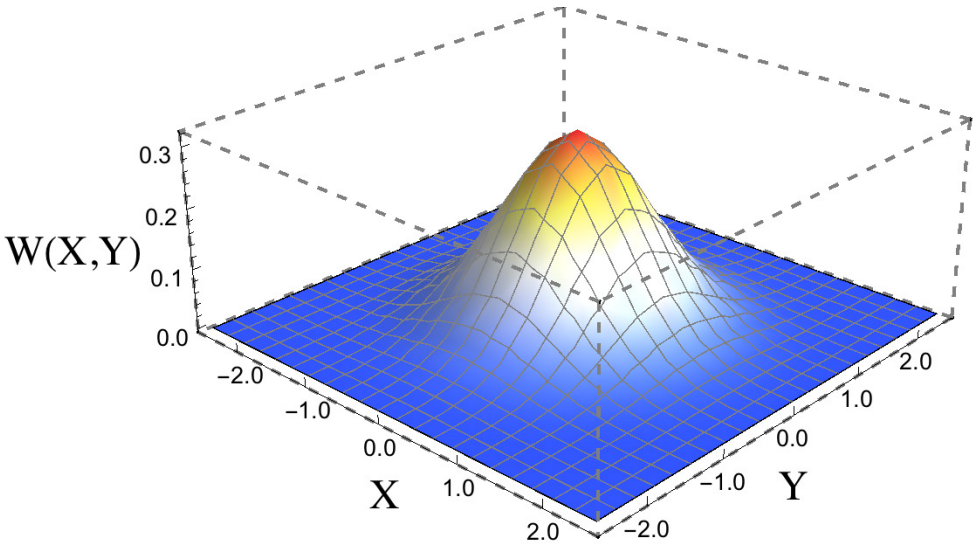


Figure 5.9: Wigner function of the final state of the measuring device (the oscillator) in the weak measurement regime . In this case, the optomechanical strength is $g = 10^{-3}$ and the post-selection parameter $\delta = 5 \cdot 10^{-2}$. Here (and in figure 5.10) $\hat{X} = (\hat{c} + \hat{c}^\dagger)/\sqrt{2}$ and $\hat{Y} = i(\hat{c}^\dagger - \hat{c})/\sqrt{2}$ are the mirror quadratures, satisfying $[\hat{X}, \hat{Y}] = i$.

occurs due to the post-selection.

The probability to read the position x, conditioned on a successful post-selection, can be computed from the state (5.43), i.e. $P(x|f) = |\langle x|\psi(t)\rangle|^2$. The distribution is presented in the appendix (A). For time $t = \pi/\Omega$, the result is a straightforward application of the expressions (4.111), (4.120) and (4.121) to the present case. The maximum likelihood estimator can be obtained, in principle, from this distribution (in an approximate manner). However, in general, it will depend on the parameter we wish to estimate and is therefore

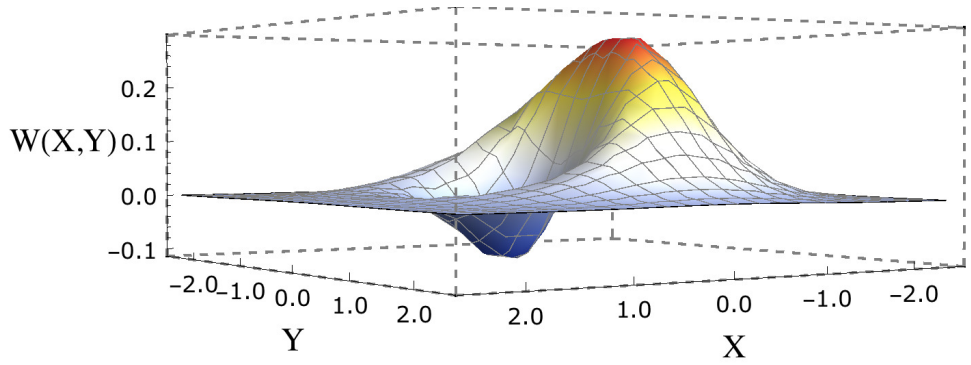


Figure 5.10: Wigner function of the final state of the oscillator *outside* the weak measurement regime, for the case $\delta \sim g$. The optomechanical strength g is set to be 10^{-3}.

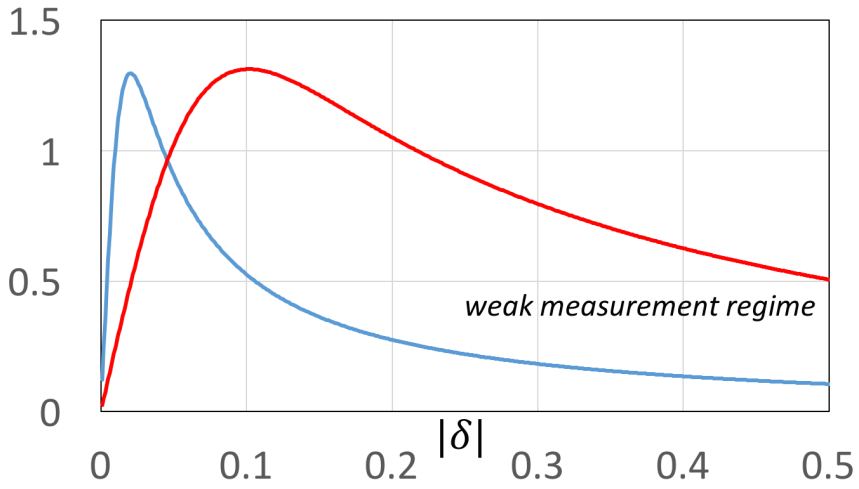


Figure 5.11: The average position of the oscillator, divided by its standard deviation, is plotted against the magnitude of the post-selection parameter. The optomechanical strengths are $g = 0.01$ (blue curve) and $g = 0.05$ (red curve). The weak measurement regime occurs when $g \ll \delta$. Note that the oscillator reaches the level of the fluctuations *outside* the weak measurement regime, when $\delta \sim g$.

unfeasible. Also, from this distribution it is possible to obtain the average position of the mirror and its standard deviation. Figure 5.11 shows the ratio between both quantities as a function of the parameter δ, illustrating the well known fact that, when the measurement is strong ($\delta \sim g$), the result is precise, while it is "noisy" when the measurement is weak ($g \ll \delta$).

5.4.2 Oscillator initially in thermal equilibrium

Let us assume now that the oscillator starts in a thermal state $\hat{\rho}_m = \sum p_n \, |n\rangle\langle n|$, where

$$p_n = \frac{e^{-E_n/k_B T}}{\sum_n^\infty e^{-E_n/k_B T}}, \quad E_n = n\hbar\Omega, \tag{5.45}$$

is the Boltzmann distribution. In this scenario, the state (5.34) will be a mixture of displaced number states,

$$\rho_m(t) = \sum_n p_n \cdot \hat{\mathcal{U}} \, |n\rangle\langle n| \, \hat{\mathcal{U}}^\dagger, \tag{5.46}$$

where we have defined the displacement operator $\widehat{\mathcal{U}} = \exp\Big\{(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)_w[\hat{b}^\dagger \varphi(t) - \hat{b}\varphi^*(t)]\Big\}$ just for the simplicity of the expression. The first and second moments of the position operator are

$$\begin{aligned} \mathrm{Tr}\{\hat{X}\rho_m(t)\} &= 2gx_0[1-\cos(\Omega t)](\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)_w, & (5.47)\\ \mathrm{Tr}\{\hat{X}^2\rho_m(t)\} &= 4g^2x_0^2[1-\cos(\Omega t)]^2(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)_w^2 + x_0^2(1+\langle n\rangle_{th}), & (5.48) \end{aligned}$$

where $\langle n\rangle_{th} = (e^{\hbar\Omega/k_BT}-1)^{-1}$ is the average number of thermal phonons. Therefore, we SNR of the estimator (5.40) corresponds to

$$SNR_{th,w} = \frac{SNR_{0,w}}{\sqrt{1+2\langle n\rangle_{th}}}. \tag{5.49}$$

From the expressions (5.47) and (5.48) it is clear that, although the shift of the average position remains the same as in the first case (when the mirror starts in the ground state), i.e. proportional to the weak value, the fluctuations are now increased with the mean number of phonons. Hence, the SNR will remain approximately the same when the average number of phonons is small, i.e. when the system is operated in the regime where $\hbar\Omega \approx k_BT$.

5.5 Weak value amplification in the presence of noise

In this last section we will study how white and coloured noise affect the Fisher information, by applying the results presented in the previous chapter, in section (4.9). We restrict our analysis to the case of a mirror initially prepared in the ground state. Up to this point, the time spent by photon inside the cavity has not been specified and therefore all the expressions for the Fisher information, SNR and other related quantities are time dependent. As it is shown in appendix (D), in a Markovian bath, the time t spent by the photon inside the cavity before it is emitted and detected at the output, follows an exponential distribution,

$$t \sim \gamma e^{-\gamma t}, \tag{5.50}$$

where γ is the cavity decay rate of the cavity. The mean lifetime of the photon inside the cavity is therefore γ^{-1}. Hence, in order for the photon to stay for half the mechanical period, $\gamma \sim \Omega$, which leads to a quality factor $Q = \omega_{cav}/\gamma = \omega_{cav}/\Omega$. Consequently, an optical cavity with a moving mirror with frequency of the order of MHz, should have a qualify factor $Q \sim 10^8$. With the time "fixed" around half the vibrational period, the Fisher information (5.38) and the SNR (5.42) become

$$\mathcal{I}(g) = 8(2-\delta^2) \quad , \quad SNR_{0,w} = 2g\sqrt{2N(2-\delta^2)}. \tag{5.51}$$

These values will be contrasted with the corresponding quantities in the presence of noise. Recall that the Fisher information $8(2-\delta^2)$ represents the information in a single event and scales as N, where N is the number of independent repetitions of the experiment. First, let us assume that the measurement of the position of the meter is affected by Gaussian white noise (4.132). In this scenario, in accordance with the result (4.136), the variance of the distribution (5.36) will be increased. The Fisher information becomes

$$\mathcal{I}_{white\ noise}(g) = \frac{8(2-\delta^2)}{1+\sigma^2_{noise}/x_0^2}, \tag{5.52}$$

which is a direct application of equation (4.137) to the present experiment. The maximum likelihood estimator is still (5.40), but now the variance of the estimator is increased. The SNR ratio of the estimator becomes

$$SNR_{white\ noise} = \frac{SNR_{0,w}}{\sqrt{1+\sigma^2_{noise}/x_0^2}}. \tag{5.53}$$

Both quantities will be affected by the noise when the classical fluctuations are much larger than the quantum fluctuations. Recall that (5.52) grows linearly with the number of repetitions of the experiment.

On the other hand, when the noise has long correlation times, the Fisher information

for the strategy without (4.149) and with post-selection (4.149) corresponds to

$$\mathcal{I}_{col.\ noise}(g) = 16\left(\frac{2-\delta^2}{\delta^2}\right)\frac{x_0^2}{\eta} \quad , \quad \mathcal{I}^{no\ post.}_{col.\ noise}(g) = 16\cdot\frac{x_0^2}{\eta}, \tag{5.54}$$

respectively. The benefits from the weak value amplification effect, as compared to a weak measurement without post-selection, are therefore evident when coloured noise is unavoidable. Let us look now into the maximum likelihood estimator. The probability distribution of the outcomes s_i will be a normal distribution with a covariance matrix $\boldsymbol{E}$ of the form (4.144), which has dimension $k \times k$, where k is the number of post-selected photons out of N trials (if post-selection is performed) or is simply equal to N (in the non post-selection scenario).

When post-selection is performed, instead of treating k as a random variable, we will replace it by its average value, i.e. by $N|\langle\psi_f|\psi_i\rangle|^2$. The maximum likelihood estimator corresponds to a weighted average of the observations s_i,

$$\hat{g}_{ML} = \frac{1}{4x_0\sqrt{2-\delta^2}/\delta}\sum_{i=0}^{k}\alpha_i s_i \quad , \quad \alpha_i = \frac{\sum_{j=1}^{k} E^{-1}_{i,j}}{\sum_{i,j=1}^{k} E^{-1}_{i,j}}. \tag{5.55}$$

Notice that the construction of the estimator requires precise knowledge of the correlations of the noise, besides x_0 and δ. When the correlations are constant over time, the weights α_i become equal to $1/k$, i.e. the sum in (5.55) reduces to the mean. The variance of the estimator corresponds to

$$Var[\hat{g}_{ML}] = \frac{1}{16(2-\delta^2)/\delta^2}\Big\{\frac{1+(\sigma^2_{noise}-\eta)/x_0^2}{k} + \frac{\eta}{x_0^2}\Big\}. \tag{5.56}$$

Consequently, when k is sufficiently large the variance of the estimator achieves a constant value, $(\eta/x_0^2)/[16(2-\delta^2)/\delta^2]$. When no post-selection is performed, the maximum likelihood estimator is analogous to (5.55) (the weak value does not appear in the denominator and $k = N$). In this case, the variance of the estimator reaches a constant value of $(\eta/x_0^2)/16$. Hence, when post-selection is implemented, the variance of the estimator will be reduced, being divided by the squared of the weak value $(2-\delta^2)/\delta^2$. Equivalently, the SNR will be amplified by the weak value when post-selection is performed.

5.6 Conclusions

Regarding our work, three main conclusions should be emphasized. First, we have presented an experiment in which the radiation pressure exerted by a single pre and post-selected photon can by amplified by anomalous weak values: an large momentum transfer is given by the photon to a mechanical oscillator when the photon makes a transition from a state that can produce constructive interference to a state that produces destructive interference. This is a quantum effect with no classical analog. Secondly, the use of the weak amplification effect achieves equal precision for parameter estimation, but *at a lower cost* since fewer observations of the measurement device are needed to construct an estimator. Finally, we have shown that in our experiment dark port post-selection can generate non classical states of a mechanical oscillator ("Schrödinger cat" states).

Appendix A

Conditioned probability distribution for measurements of any strength

Consider the case when the measurement device starts in the ground state,

$$\psi(x) = \frac{1}{\sqrt{2\pi}x_0} e^{-\frac{x^2}{4x_0^2}}, \tag{A.1}$$

and the system starts in the initial state (5.24). The state of the meter, when the final state (5.22) is post-selected, is given by (5.43). Then, if the position of the measurement device is observed, the probability to find it in small interval dx around the position x is $P(x|f)dx$, where the density function corresponds to

$$\begin{aligned} P(x|f) &= \frac{1}{\mathcal{N}}|\psi(x)|^2 e^{-2g^2[1-\cos(\Omega t)]^2} \left\{ \frac{1}{\delta^2} \cosh\Big[2g(1-\cos\{\Omega t\})x/x_0\Big] + \right. \\ &\left. \frac{\sqrt{2-\delta^2}}{\delta} \sinh\Big[2g(1-\cos\{\Omega t\})x/x_0\Big] - \frac{1-\delta^2}{\delta^2} \cos\Big[2g\sin\{\Omega t\}x/x_0\Big] \right\}, \end{aligned} \tag{A.2}$$

The normalization factor $\mathcal{N}$ depends both on g and δ,

$$\mathcal{N} = \frac{1}{\delta^2} - \frac{1-\delta^2}{\delta^2} e^{-4g^2[1-\cos(\Omega t)]}. \tag{A.3}$$

In the weak measurement regime (5.33), the probability density (A.2) reduces to (5.36), a Gaussian density equal to the original density function $|\psi(x)|^2$ displaced by a large amount (larger than the eigenvalues of the system variable).

Appendix B

Measurement operators

In this appendix we will present the measurement operators for the case of a measurement device that starts in the pure state (4.64). We consider measurements of an observable $\hat{A}$, which couples to an apparatus variable $\hat{P}$, according to the von Neumann model.

Due to (2.33) the measurement operators are given by

$$\begin{aligned}\widehat{M}_q &= N_q \exp\left\{(i\mathcal{F}-1)\frac{(q-q_0-g\hat{A})^2}{4\sigma_q^2} + ip_0(q-g\hat{A})/\hbar\right\} \\ &= \psi(q)\exp\left\{-(i\mathcal{F}-1)\Big(\frac{g}{2\sigma_q}\Big)\Big(\frac{q-q_0}{\sigma_q}\Big)\hat{A} - ig\Big(\frac{p_0}{\hbar}\Big)\hat{A} + (i\mathcal{F}-1)\Big(\frac{g}{2\sigma}\Big)^2\hat{A}^2\right\}, \qquad \text{(B.1)}\end{aligned}$$

that shows that the wave function of the measurement device is displaced by $g\hat{A}$. It is convenient to define the auxiliar (standardized) variable $z = (q-q_0)/\sigma_q$, the complex parameter $\gamma = 1 - i\mathcal{F}$ and $k_0 = p_0/\hbar$. Then, the measurement operator can be written as

$$\begin{aligned}\widehat{M}_q &= \psi(q)\exp\left\{2\Big(\sqrt{\gamma}\frac{g}{2\sigma_q}\hat{A}\Big)\Big(\frac{\sqrt{\gamma}z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}}k_0\Big) - \Big(\sqrt{\gamma}\frac{g}{2\sigma_q}\hat{A}\Big)^2\right\} \\ &= \psi(q)\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\frac{\sqrt{\gamma}g}{2\sigma_q}\hat{A}\Big)^n H_n\Big(\frac{\sqrt{\gamma}z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}}k_0\Big). \qquad \text{(B.2)}\end{aligned}$$

In the last line we have used the fact that $\exp\{2xt - t^2\}$ is the exponential generating

function of the Hermite sequence $H_n(x)$ [140], namely,

$$\exp\{2xt - t^2\} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x), \tag{B.3}$$

for any complex values t and x. Hence, the measurement operators associated to the selection of the state $|\psi_f\rangle$ in the second measurement and to the obtaining of the outcome q in the intermediate measurement correspond to $\widehat{M}_{f,q} = |\psi_f\rangle\langle\psi_f|\,\widehat{M}_q$. Their action over an initial pure state $|\psi_i\rangle$ is

$$|\psi_f\rangle\langle\psi_f|\,\widehat{M}_q\,|\psi_i\rangle = |\psi_f\rangle\langle\psi_f|\psi_i\rangle\,\psi(q)\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\frac{\sqrt{\gamma}g}{2\sigma_q}\Big)^n A_w^n H_n\Big(\frac{\sqrt{\gamma}z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}}k_0\Big). \tag{B.4}$$

The squared norm of this state corresponds to the probability $P(f,q)$, i.e. to the probability of selecting the desired state and reading the value q in the intermediate measurement,

$$\begin{aligned} P(q,f) &= |\langle\psi_f|\,\widehat{M}_q\,|\psi_i\rangle|^2 & \text{(B.5)}\\ &= |\psi(q)|^2|\langle\psi_f|\psi_i\rangle|^2\left|\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\frac{\sqrt{\gamma}g}{2\sigma_q}\Big)^n A_w^n H_n\Big(\frac{\sqrt{\gamma}z}{2} - i\frac{\sigma_q}{\sqrt{\gamma}}k_0\Big)\right|^2. & \text{(B.6)}\end{aligned}$$

Clearly, the probability can be expressed as $|\psi(q)|^2|\langle\psi_f|\psi_i\rangle|^2 S(q)$, where the series $S(q)$ is presented, expanded in powers of g, in equation (4.112).

Appendix C

Beam splitters

Classically, a beam splitter is a device that "splits" an incident beam of light into two components; a transmitted and a reflected component, as it is shown in the left side of figure (C.1). When the beam splitter is treated quantum mechanically, the second input should be necessarily taken into consideration, which is shown explicitly in the right side of (C.1).

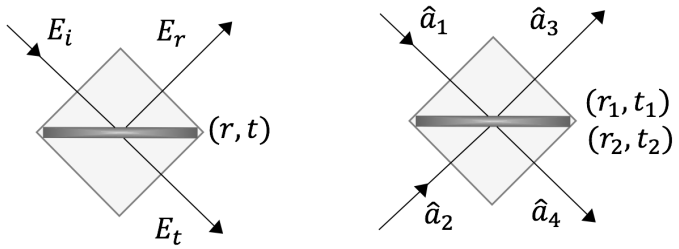


Figure C.1: *Left side figure*: the incident field is separated into a reflected component, $E_r = rE_i$, and a transmitted component, $E_t = tE_i$, where r and t are the reflectance and transmittance of the beam splitter, respectively. For a "quantum beam splitter" (*right side figure*) these equations are not valid and the second input must be necessarily taken into consideration, even when this port is left unused (in the vacuum state).

The relation between the input modes and the output modes is given by

$$\begin{aligned} \hat{a}_3 &= r_1\hat{a}_1 + t_2\hat{a}_2, \\ \hat{a}_4 &= t_1\hat{a}_1 + r_2\hat{a}_2, \end{aligned} \tag{C.1}$$

where r_1 and r_2 is the set of reflectances of the beam splitter, and t_1 and t_2 is the set

of transmittances. These quantities can be complex values and depend on how the beam splitter is constructed.

The mode operators in the input ports satisfy boson commutation relations, i.e. $[\hat{a}_1, \hat{a}_1^\dagger] = [\hat{a}_2, \hat{a}_2^\dagger] = 1$ and $[\hat{a}_1, \hat{a}_2] = [\hat{a}_1, \hat{a}_2^\dagger] = 0$. In order for the mode operators of the output ports to satisfy analogous boson commutation relations, the set of reflectances and transmittances should satisfy the *reciprocity relations* [198]:

$$|r_1| = |r_2| \ , \ |t_1| = |t_2| \ , \ |r_1|^2 + |t_1|^2 = 1 \ , \ r_1^* t_2 + r_2 t_1^* = r_1^* t_1 + r_2 t_2^* = 0. \tag{C.2}$$

When $|r_1| = |t_1|$ the beam splitter is said to be a balanced, or 50:50, beam splitter. Notice that the last equation relates the phase shifts between the output beams. The reciprocity relations (C.2) assure that the transformation (C.1) of the input modes $\hat{a}_1$ and $\hat{a}_2$ into the output modes $\hat{a}_3$ and $\hat{a}_4$ is *unitary*. Assume, for example, that all the transmittances and reflectances are real values, given by

$$r_1 = r \in [0,1] \quad , \quad r_2 = -r \quad , \quad t_1 = t_2 = t \in [0,1] \quad , \quad r^2 + t^2 = 1. \tag{C.3}$$

Allowing r to be distinct from t offers the possibility to design an unbalanced beam splitter. In this situation, the transformation (C.1) takes the form of (5.20). Notice that this transformation can be equivalently expressed as

$$\begin{aligned} \hat{a}_3 &= e^{(i/\hbar)\int dt \hat{H}_{BS2}} \cdot e^{(i/\hbar)\int dt \hat{H}_{BS1}} \cdot \hat{a}_1 \cdot e^{-(i/\hbar)\int dt \hat{H}_{BS1}} \cdot e^{-(i/\hbar)\int dt \hat{H}_{BS2}}, \\ \hat{a}_4 &= e^{(i/\hbar)\int dt \hat{H}_{BS2}} \cdot e^{(i/\hbar)\int dt \hat{H}_{BS}} \cdot \hat{a}_2 \cdot e^{-(i/\hbar)\int dt \hat{H}_{BS}} \cdot e^{-(i/\hbar)\int dt \hat{H}_{BS2}}, \end{aligned} \tag{C.4}$$

where the integrals are taken over the time the fields interact with each other inside the beam splitter. $\hat{H}_{BS1}$ are $\hat{H}_{BS2}$ are effective hamiltonians that describe the beam splitter. It is easy to show that the hamiltonians correspond to

$$\hat{H}_{BS1} = -i\hbar\theta\delta(t)(\hat{a}_2^\dagger \hat{a}_1 - \hat{a}_1^\dagger \hat{a}_2), \tag{C.5}$$

$$\hat{H}_{BS2} = \hbar\pi\delta(t)\hat{a}_2^\dagger \hat{a}_2, \tag{C.6}$$

where $\delta(t)$ is the Dirac delta function (which integrates to one during the time of interaction between the fields inside the beam splitter), $\cos(\theta) = r$ and $\sin(\theta) = t$. Physically, the first hamiltonian can be implement with a single layer of dielectric material. The second hamiltonian corresponds to a *phase shift gate*. The action of the beam splitter is described therefore by the consecutive evolution generated by $\hat{H}_{BS1}$ and $\hat{H}_{BS2}$. The output modes $\hat{a}_3$ and $\hat{a}_4$ are simply the evolution of the input modes in the *Heisenberg picture*.

It is also possible to work in the Schrödinger picture and evolve states instead of operators. Consider the input state $|0\rangle_1 |1\rangle_2$, that describes the input port 1 left unused and a single photon entering through the input port 2. The evolution of the state corresponds to

$$\begin{aligned} |0\rangle_1 |1\rangle_2 &= \hat{a}_2^\dagger |0\rangle_1 |0\rangle_2 \\ &= (t\hat{a}_3^\dagger - r\hat{a}_4^\dagger) |0\rangle_3 |0\rangle_4 \\ &= t |1\rangle_3 |0\rangle_4 - r |0\rangle_3 |1\rangle_4 . \end{aligned} \tag{C.7}$$

Note that in the second step the vacuum states in the input and the output have been identified, i.e. $|0\rangle_1 |0\rangle_2 = |0\rangle_3 |0\rangle_4$. The final state is a *path-entangled state*, namely, a coherent superposition of both paths. Single-photon interference can be studied by using second beam splitter in a Mach-Zehnder configuration (see figure 4.1) .

Besides (C.3), another common design is a balanced beam splitter described by the following set of reflectances and transmittances:

$$r_1 = t_2 = \frac{1}{2} \quad , \quad r_2 = t_1 = \frac{i}{2}. \tag{C.8}$$

For this type of beam splitter, the transformation of the fields (C.1) can be written as

$$\hat{a}_3 = e^{(i/\hbar)\int dt \mathcal{H}_{BS}} \hat{a}_1 e^{-(i/\hbar)\int dt \mathcal{H}_{BS}}, \tag{C.9}$$

$$\hat{a}_4 = e^{(i/\hbar)\int dt \mathcal{H}_{BS}} \hat{a}_2 e^{-(i/\hbar)\int dt \mathcal{H}_{BS}}, \tag{C.10}$$

where $\mathcal{H}_{BS} = -\hbar(\pi/4)\delta(t)(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2)$ is an effective hamiltonian that describes the beam splitter. Physically, it can be implemented with a single layer of dielectric material. If the

operator $\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2$ is associated with a $\hat{J}_x$ operator, according to the Jordan-Schwinger map (4.47), then the evolution generated by the beam splitter corresponds to (4.48).

Appendix D

Photon detection

Let us consider a double sided cavity that interacts with two reservoirs, that are monitored with photon counters, as it is shown in figure D.1. Each reservoir (or bath) is modelled as a collection of harmonic oscillators, each one with a frequency ω_n. The hamiltonian that describes the *open quantum system* is

$$\begin{aligned}
\hat{H} &= \hbar\sum_n \omega_n \hat{c}^\dagger_{1,n}\hat{c}_{1,n} + \hbar\sum_n \omega_n \hat{c}^\dagger_{2,n}\hat{c}_{2,n} + && \text{(D.1)}\\
&\quad \hbar\left(\hat{a}^\dagger_1\sum_n g_{1,n}\hat{c}_{1,n} + \hat{a}_1\sum_n g_{1,n}\hat{c}^\dagger_{1,n}\right) + \hbar\left(\hat{a}^\dagger_2\sum_n g_{2,n}\hat{c}_{2,n} + \hat{a}_2\sum_n g_{2,n}\hat{c}^\dagger_{2,n}\right) + && \text{(D.2)}\\
&\quad \hbar\omega_{cav}(\hat{a}^\dagger_1\hat{a}_1 + \hat{a}^\dagger_2\hat{a}_2) + \hbar\Omega\hat{b}^\dagger\hat{b} - \hbar g_0(\hat{a}^\dagger_1\hat{a}_1 - \hat{a}^\dagger_2\hat{a}_2)(\hat{b}^\dagger + \hat{b}). && \text{(D.3)}
\end{aligned}$$

The operators $\hat{c}_{i,n}$ are boson annihilation operators of each reservoir ($i = 1, 2$). The subscript n labels a different mode. The modes can be thought as standing waves with a sinusoidal mode function that vanishes at the position of the mirrors. The terms in the first line (D.1) correspond the free energy of the reservoirs. The terms in the second line (D.2) describe the interaction between the external modes (of the reservoir) and the field inside the cavity. The constants $g_{i,n}$ are the different couplings between the cavity mode ($\hat{a}_i$) and the field outside the cavity ($\hat{c}_{i,n}$). This part of the hamiltonian describes the exchange of photons between the external field and the cavity that occurs because of the interaction between the fields inside the mirrors. Indeed, the mirrors of the cavity have actually some width and a small amount the electromagnetic field penetrates inside the mirrors, allowing

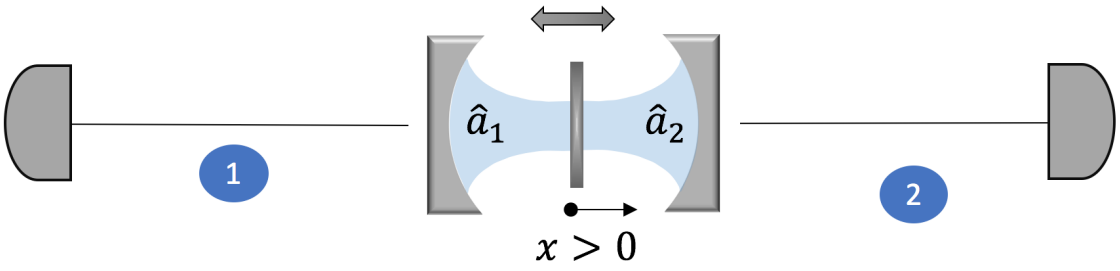


Figure D.1: An optomechanical system (a Fabry-Pérot cavity with a movable mirror placed in the center of the cavity) is monitored with photon counters. The optomechanical system interacts with the baths 1 and 2.

the exchange of photons. For a high Q cavity, the hamiltonian in the second line (D.2) is a good approximation to describe the absorption of photons from the bath or the emission of photons from the cavity into the bath. In our work, we will assume that the photon has been already absorbed. Therefore, this part of the hamiltonian will describe the emission or leakage of photons from the cavity into the reservoir. For more details regarding the hamiltonian that describes the exchange of photons between the cavity and the bath see [112, 113, 114, 115, 116, 117, 118]. Finally, the terms in the last line (D.3) correspond to the hamiltonian of the (closed) optomechanical system, which was described in equation (5.14).

First, we go to an *interaction picture* with respect to $\hbar\sum_n \omega_n \hat{c}^\dagger_{1,n}\hat{c}_{1,n} + \hbar\sum_n \omega_n \hat{c}^\dagger_{2,n}\hat{c}_{2,n} + \hbar\omega_{cav}(\hat{a}^\dagger_1\hat{a}_1 + \hat{a}^\dagger_2\hat{a}_2)$. In this frame, the hamiltonian becomes

$$\begin{aligned}\hat{H}_I/\hbar \;=\; & \hat{a}^\dagger_1\sum_n g_{1,n}e^{-i\delta_n t}\hat{c}_{1,n} + \hat{a}_1\sum_n g_{1,n}e^{i\delta_n t}\hat{c}^\dagger_{1,n} + \hat{a}^\dagger_2\sum_n g_{2,n}e^{-i\delta_n t}\hat{c}_{2,n} + \hat{a}_2\sum_n g_{2,n}e^{i\delta_n t}\hat{c}^\dagger_{2,n} \\ & +\Omega\hat{b}^\dagger\hat{b} - g_0(\hat{a}^\dagger_1\hat{a}_1 - \hat{a}^\dagger_2\hat{a}_2)(\hat{b}^\dagger + \hat{b}). \end{aligned} \tag{D.4}$$

Notice that the hamiltonian is time dependent (the subscript I indicates that it corresponds to the the hamiltonian in the interaction frame). The parameter $\delta_n = \omega_n - \omega_{cav}$ represents the detuning of each mode of the bath with respect to the frequency of the cavity. From now on, we base our analysis on the treatment of an open quantum described in section 3.11 of [12]. It is convenient to redefine the annihilation operators of the cavity as

$$\hat{a}_1 = -i\sqrt{\gamma_1}\hat{a}_1 \quad , \quad \hat{a}_2 = -i\sqrt{\gamma_2}\hat{a}_2, \tag{D.5}$$

where i is the imaginary unit and $\gamma_{1,2}$ are the dissipation rates of each side of the cavity. Additionally, it is useful to define the bath operators

$$\hat{c}_1(t) = \frac{1}{\sqrt{\gamma_1}} \sum_n g_{1,n} \hat{c}_{1,n} e^{i\delta_n t} \quad , \quad \hat{c}_2(t) = \frac{1}{\sqrt{\gamma_2}} \sum_n g_{2,n} \hat{c}_{2,n} e^{i\delta_n t}. \tag{D.6}$$

It is important to examine the commutation relations between the bath operators. Consider the commutator between the operators associated to the bath 1,

$$[\hat{c}_1(t), \hat{c}_1^\dagger(t')] = \frac{1}{\gamma_1} \sum_n g_{1,n}^2 e^{i\delta_n (t-t')} = \frac{1}{\gamma_1} \int_0^\infty g_1^2(\omega) \rho_1(\omega) e^{(\omega - \omega_{cav})(t-t')} d\omega, \tag{D.7}$$

where $\rho_1(\omega)$ is the density of modes of the bath 1 and $g_1(\omega)$ is the coupling between the cavity mode and the n-th mode of the bath, as a function of frequency. The density of modes is proportional to the volume V of the bath, but the coupling $g_1(\omega) \sim V^{-1/2}$. Consequently, the term $g_1^2(\omega)\rho(\omega)$ remains finite as V is increased. This allows to convert the summation into an integral. Additionally, if $g_1^2(\omega)\rho_1(\omega)$ is a smooth function of the frequency around ω_{cav}, then

$$\int_0^\infty g_1^2(\omega)\rho(\omega) e^{(\omega - \omega_{cav})(t-t')} d\omega \approx 2\pi g_1(\omega_{cav}) \rho(\omega_{cav})^2 \delta(t-t'). \tag{D.8}$$

The term $2\pi g_1(\omega_{cav})\rho(\omega_{cav})^2$ corresponds to the rate at which energy is dissipated from the cavity, through the mirror of the side 1. This fact can be checked when the equation that describes the evolution of the average photon number inside the cavity is derived, which can be made from the master equation for the cavity field in the *Born-Markov approximation*. Consequently, the bath operators are "delta-correlated",

$$[\hat{c}_1(t), \hat{c}_1^\dagger(t')] = \delta(t-t'). \tag{D.9}$$

The same result applies to the bath operators of the other side. Assuming that the decay rate through each mirror of the cavity is the same ($\gamma_1 = \gamma_2 = \gamma$), in terms of the bath

operators (D.6) and the cavity operators (D.5), the Hamiltonian (D.4) can be expressed as

$$\begin{aligned}\hat{H}_I/\hbar &= -i\Big[\hat{a}_1^\dagger \hat{c}_1(-t) - \hat{a}_1 \hat{c}_1^\dagger(-t)\Big] - i\Big[\hat{a}_2^\dagger \hat{c}_2(-t) - \hat{a}_2 \hat{c}_2^\dagger(-t)\Big] \\ &\quad + \Omega \hat{b}^\dagger \hat{b} - \frac{g_0}{\gamma}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)(\hat{b}^\dagger + \hat{b}). \end{aligned} \tag{D.10}$$

The term in the second line describes the optomechanical interaction. For simplicity, it will denoted by $\mathcal{H}$. The evolution generated by the hamiltonian (D.10) during an infinitesimal time interval dt is given by

$$\hat{U}(t+dt,t) = \exp\Big\{\hat{a}_1^\dagger d\hat{C}_1(t) + \hat{a}_1 d\hat{C}_1^\dagger(t) + \hat{a}_2^\dagger d\hat{C}_2(t) + \hat{a}_2 d\hat{C}_2^\dagger(t) - i\mathcal{H}dt\Big\}, \tag{D.11}$$

where we have introduced "new bath operators" $d\hat{C}_1(t) = \hat{c}_1(-t)dt$ and $d\hat{C}_2(t) = \hat{c}_2(-t)dt$. These operators are of order $\sqrt{dt}$. Indeed, from the commutation relation (D.9), it is clear that $[\hat{c}_1(t), \hat{c}_1^\dagger(t)] = \delta(0)$, which implies that $\hat{c}_1(t)$ is of order $1/\sqrt{dt}$ because the delta function at zero can be considered as $1/dt$. The commutation relation between these operators is described by

$$[d\hat{C}_1(t), d\hat{C}_1^\dagger(t)] = dt \quad , \quad [d\hat{C}_2(t), d\hat{C}_2^\dagger(t)] = dt. \tag{D.12}$$

Consequently, a first order expansion of (D.11), in terms of dt, should keep terms that involve products of $d\hat{C}_1(t)$ and $d\hat{C}_2(t)$ (and products between its hermitian conjugate variables). Therefore,

$$\begin{aligned}\hat{U}(t+dt,t) &\approx \mathbb{1} - i\mathcal{H}dt + \hat{a}_1 d\widehat{C}_1^\dagger(t) - \hat{a}_1^\dagger d\widehat{C}_1(t) + \hat{a}_2 d\widehat{C}_2^\dagger(t) - \hat{a}_2^\dagger d\widehat{C}_2(t) \\ &\quad - \frac{1}{2}\hat{a}_1^\dagger \hat{a}_1 dt - \frac{1}{2}\hat{a}_2^\dagger \hat{a}_2 dt - \frac{1}{2}\{\hat{a}_1^\dagger, \hat{a}_1\} d\widehat{C}_1^\dagger(t) d\widehat{C}_1(t) - \frac{1}{2}\{\hat{a}_2^\dagger, \hat{a}_2\} d\widehat{C}_2^\dagger(t) d\widehat{C}_2(t) \\ &\quad + \frac{[\hat{a}_1 d\widehat{C}_1(t)]^2}{2} + \frac{[\hat{a}_1 d\widehat{C}_1^\dagger(t)]^2}{2} + \frac{[\hat{a}_2 d\widehat{C}_2(t)]^2}{2} + \frac{[\hat{a}_2 d\widehat{C}_2^\dagger(t)]^2}{2}. \end{aligned} \tag{D.13}$$

We will assume that the cavity and the baths are initially uncorrelated. Also, let us suppose that the optomechanical system (the cavity and the vibrating mirror) starts in a pure state $|\Psi\rangle$ and both baths begin in the multi-mode vacuum state, denoted as $|0\rangle_1 |0\rangle_2$. Due to the fact that both baths are in the vacuum state all terms in (D.13) that involve $d\hat{C}_1(t)$, $d\hat{C}_2(t)$

and its hermitian conjugate variables, in *normal order*, can be disregarded. Consequently, the action of the evolution operator over the initial system state is given by

$$\begin{aligned}\hat{U}(t+dt,t)\,|\Psi\rangle\,|0\rangle_1\,|0\rangle_2 &= \left[\mathbb{1} - dt\Big(i\mathcal{H} + \frac{1}{2}\hat{a}_1^\dagger\hat{a}_1 + \frac{1}{2}\hat{a}_2^\dagger\hat{a}_2\Big)\right]|\Psi\rangle\,|0\rangle_1\,|0\rangle_2 \\ &\quad + \hat{a}_1\,|\Psi\rangle\,|0\rangle_2\,d\hat{C}_1^\dagger(t)\,|0\rangle_1 + \hat{a}_2\,|\Psi\rangle\,|0\rangle_1\,d\hat{C}_2^\dagger(t)\,|0\rangle_2\,. \quad \text{(D.14)}\end{aligned}$$

The terms in the second line create photons in the reservoirs. The probability to detect a photon in the first bath corresponds to norm of the single-photon state created in the this reservoir, namely,

$$P_1 = \langle\Psi|\,\hat{a}_1^\dagger\hat{a}\,|\Psi\rangle\langle 0|_1\,d\hat{C}_1(t)d\hat{C}_1^\dagger(t)\,|0\rangle_1\,. \quad \text{(D.15)}$$

The probability to detect the photon in the second bath is analogous. Finally, from the commutation relations (D.9) and using the actual mode operators of the cavity (D.5), during an infinitesimal time interval dt, the probabilities to detect a photon in the first bath, in the second bath, and to detect no photons at all, are given by

$$\begin{aligned}P_1 &= \gamma\left\langle\hat{a}_1^\dagger\hat{a}_1\right\rangle dt, & \text{(D.16)}\\ P_2 &= \gamma\left\langle\hat{a}_2^\dagger\hat{a}_2\right\rangle dt, & \text{(D.17)}\\ P_0 &= 1 - \gamma\left\langle\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_1^\dagger\hat{a}_1\right\rangle dt. & \text{(D.18)}\end{aligned}$$

Notice that the probabilities of detection are of order dt. The measurement operators associated to each one of the previous events correspond to

$$\widehat{M}_1 = -i\sqrt{\gamma dt}\hat{a}_1, \quad \widehat{M}_2 = -i\sqrt{\gamma dt}\hat{a}_2, \quad \widehat{M}_0 = \mathbb{1} - dt\Big(i\mathcal{H} + \gamma\frac{\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2}{2}\Big). \quad \text{(D.19)}$$

Indeed, it is easy to see that $P_0 = \langle\Psi|\,\widehat{M}_0^\dagger\widehat{M}_0\,|\Psi\rangle$, $P_1 = \langle\Psi|\,\widehat{M}_1^\dagger\widehat{M}_1\,|\Psi\rangle$ and $P_2 = \langle\Psi|\,\widehat{M}_2^\dagger\widehat{M}_2\,|\Psi\rangle$. The set of measurement operators (D.19) defines a *Poisson point process*. The probability

to detect no photons during an time interval t is given by

$$\begin{aligned} P(t) &= \mathrm{Tr}\left[\exp\left\{ -i\mathcal{H}t - \gamma t \frac{\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2}{2} \right\} |\Psi\rangle\langle\Psi| \exp\left\{ -i\mathcal{H}t - \gamma t \frac{\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2}{2} \right\} \right] \\ &= \exp\Big\{ -\gamma t \Big\langle \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2 \Big\rangle \Big\}. \end{aligned} \tag{D.20}$$

Consequently, the probability density $f(t)$ to detect a photon at time t corresponds to

$$f(t) = -\frac{dP(t)}{dt} = \gamma \Big\langle \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2 \Big\rangle \exp\Big\{ -\gamma t \Big\langle \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2 \Big\rangle \Big\}, \tag{D.21}$$

which corresponds to an exponential distribution with parameter $\gamma \Big\langle \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2 \Big\rangle$.

Bibliography

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, *How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100*, Phys. Rev. Lett. **60**, 1351 (1988).

[2] D. Bohm, *Quantum Theory* (Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1951).

[3] Y. Aharonov, P. G. Bergmann, and J. K. Lebowitz, *Time Symmetry in the Quantum Process of Measurement*, Phys. Rev. **134**, B1410 (1964).

[4] Y. Aharonov and L. Vaidman, in *Time in Quantum Mechanics. Lecture Notes in Physics.*, edited by J. Muga, R. S. Mayato, and Í. Egusquiza (Springer, Berlin, Heidelberg, 2008).

[5] S. Carrasco and M. Orszag, *Weak-value amplification of photon-number operators in the optomechanical interaction*, Phys. Rev. A **99**, 013801 (2019).

[6] S. Carrasco and M. Orszag, *Fisher Information, Weak Values and Correlated Noise in Interferometry*, arXiv:1902.09247v1.

[7] S. M. Stigler, *Thomas Bayes's Bayesian Inference*, J R Stat Soc Ser A Stat Soc **145**, 250 (1982).

[8] K. Kraus, *States, Effects and Operations* (Springer-Verlag, Berlin-Heidelberg, 1983).

[9] G. Lüders, *Über die Zustandsänderung durch den Messprozess*, Ann. Phys. **443**, 322 (1951).

[10] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).

[11] L. Carroll and H. Oxenbury, *Alice's Adventures in Wonderland* (Macmillan Co., New York-London, 1872).

[12] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).

[13] H. M. Wiseman, *Weak values, quantum trajectories, and the cavity-QED experiment on wave-particle correlation*, Phys. Rev. A **65**, 032111 (2002).

[14] B. E. Y. Svensson, *Pedagogical Review of Quantum Measurement Theory with an Emphasis on Weak Measurements*, Quanta **2**, 18 (2013).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[16] J. A. Bergou and M. Hillery, *Introduction to the Theory of Quantum Information Processing* (Springer, New York, 2013).

[17] A. G. Kofman, S. Ashhab, and F. Nori, *Nonperturbative theory of weak pre- and post-selected measurements*, Phys. Rep. **520**, 43 (2012).

[18] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, *Colloquium: Understanding quantum weak values: Basics and applications*, Rev. Mod. Phys. **86**, 307 (2014).

[19] J. Dressel and A. N. Jordan, *Weak Values are Universal in Von Neumann Measurements*, Phys. Rev. Lett. **109**, 230402 (2012).

[20] A. Di Lorenzo, *Full counting statistics of weak-value measurement*, Phys. Rev. A **85**, 032106 (2012).

[21] R. Jozsa, *Complex weak values in quantum measurement*, Phys. Rev. A **76**, 044103 (2007).

[22] L. Vaidman, A- Ben-Israel, J. Dziewior, L. Knips, M. Weil, J. Meinecke, C. Schwemmer, R. Ber, and H. Weinfurter, *Weak value beyond conditional expectation value of the pointer readings*, Phys. Rev. A **96**, 032114 (2017).

[23] C. Ferrie and J. Combes, *How the Result of a Single Coin Toss Can Turn Out to be 100 Heads*, Phys. Rev. Lett. **113**, 120404 (2014).

[24] C. Ferrie and J. Combes, *Ferrie and Combes Reply:*, Phys. Rev. Lett. **114**, 118902 (2015).

[25] A. Brodutch, *Comment on "How the Result of a Single Coin Toss Can Turn Out to be 100 Heads"*, Phys. Rev. Lett. **114**, 118901 (2015).

[26] D. Suter, *"Weak measurements" and the "quantum time-translation machine" in a classical system*, Phys. Rev. A **51**, 45 (1995).

[27] D. F. Mundarain and M. Orszag, *Quantumness of the anomalous weak measurement value*, Phys. Rev. A **93**, 032106 (2016).

[28] D. Sokolovski, *The meaning of "anomalous weak values" in quantum and classical theories*, Phys. Lett. A **114**, 1097 (2015).

[29] A. Romito, A. N. Jordan, Y. Aharonov, and Y. Gefen, *Weak values are quantum: you can bet on it*, Quantum Stud.: Math. Found. **3**, 1 (2016).

[30] L. Vaidman, *Weak value controversy*, Phil. Trans. R. Soc. A. **375**, 20160395 (2017).

[31] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, *The sense in which a "weak measurement" of a spin-1/2 particle's spin component yields a value 100*, Phys. Rev. D **40**, 2112 (1989).

[32] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, *Realization of a measurement of a "weak value"*, Phys. Rev. Lett. **66**, 1107 (1991).

[33] G. Nienhuis and L. Allen, *Paraxial wave optics and harmonic oscillators*, Phys. Rev. A. **48**, 656 (1993).

[34] D. Stoler, *Operator methods in physical optics*, J. Opt. Soc. Am **71**, 334 (1981).

[35] B. Tamir and E. Cohen, *Introduction to Weak Measurements and Weak Values*, Quanta **2**, 7 (2013).

[36] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, *Experimental Realization of the Quantum Box Problem*, Phys. Lett. A **324**, 125 (2004).

[37] K. J. Resch and A. M. Steinberg, *Extracting Joint Weak Values with Local, Single-Particle Measurements*, Phys. Rev. Lett. **92**, 130402 (2004).

[38] J. S. Lundeen, K. J. Resch, *Practical measurement of joint weak values and their connection to the annihilation operator*, Phys. Lett. A, **334**, 337 (2005).

[39] J. S. Lundeen and A. M. Steinberg, *Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox*, Phys. Rev. Lett. **102**, 020404 (2009).

[40] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, *Direct measurement of the quantum wavefunction*, Nature **474**, 188 (2011).

[41] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, *Direct observation of Hardy's paradox by joint weak measurement with an entangled photon pair*, New J. Phys. **11**, 033011 (2009).

[42] Y. Aharonov and L. Vaidman, *Complete description of a quantum system at a given time*, J. Phys. A: Math. Gen. **24**, 2315 (1991).

[43] L. Vaidman, *Weak-measurement elements of reality*, Found. Phys. **26**, 895 (1996).

[44] D. Z. Albert, Y. Aharonov, and S. D'Amato, *Curious New Statistical Prediction of Quantum Mechanics*, Phys. Rev. Lett. **54**, 5 (1985).

[45] T. Ravon and L. Vaidman, *The three-box paradox revisited*, J. Phys. A: Math. Theor. **40**, 2873 (2007).

[46] K. A. Kirkpatrick, *Classical three-box "paradox"*, J. Phys. A: Math. Gen. **36**, 4891 (2003).

[47] Y. Aharonov, S. Popescu, D. Rohrlich, and P. Skrzypczyk, *Quantum Cheshire Cats*, New J. Phys. **15**, 113015 (2013).

[48] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa, *Observation of a quantum Cheshire Cat in a matter-wave interferometer experiment*, Nat. Commun. **5**, 4492 (2014).

[49] A. Matzkin and A. K. Pan, *Three-box paradox and "Cheshire cat grin": the case of spin-1 atoms*, J. Phys. A: Math. Theor. **46**, 315307 (2013).

[50] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes. Quantum Theory for the Perplexed* (Wiley-VCH, Weinheim, 2005).

[51] L. Hardy, *Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories*, Phys. Rev. Lett. **68**, 2981 (1992).

[52] A.C. Elitzur and L. Vaidman, *Quantum mechanical interaction-free measurements*, Found. Phys. **23**, 987 (1993).

[53] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, *Revisiting Hardy's paradox: counterfactual statements, real measurements, entanglement and weak values*, Phys. Lett. A **301**, 130 (2002).

[54] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, Physics **1**, 195 (1964).

[55] A. Einstein, B. Podolsky, and R. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Phys. Rev. **47**, 777 (1935).

[56] P. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, *Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification*, Phys. Rev. Lett. **102**, 173601 (2009).

[57] G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, *Measurement of Quantum Weak Values of Photon Polarization*, Phys. Rev. Lett. **94**, 239904 (2005).

[58] R. Mir, J. S. Lundeen, M. W. Mitchell, A. M. Steinberg, J. L. Garretson, and H. M. Wiseman, *A double-slit "which-way" experiment on the complementarity?uncertainty debate*, New J. Phys. **9**, 287 (2007).

[59] M. O. Scully, B.-G. Englert, and H. Walther, *Quantum optical tests of complementarity*, Nature **351**, 111 (1991).

[60] E. P. Storey, S. M. Tan, M. J. Collet, and D. F. Walls, *Path detection and the uncertainty principle*, Nature **367**, 626 (1994).

[61] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. Krister Shalm, and A. M. Steinberg, *Observing the average trajectories of single photons in a two-slit interferometer*, Science **322**, 1170 (2011).

[62] A. D. Parks, D. W. Cullin, and D. C. Stoudt, *Observation and measurement of an optical Aharonov-Albert-Vaidman effect*, Proc. R. Soc. Lond. A **454** 2997 (1998).

[63] Y.-W. Cho, H.-T. Lim, Y.-S. Ra, and Y.-H. Kim, *Weak Value Measurement with an Incoherent Measuring Device*, New J. Phys. **12**, 023036 (2010).

[64] O. Hosten and P. Kwiat, *Observation of the Spin Hall Effect of Light via Weak Measurements*, Science **319**, 787 (2008).

[65] K. Y. Bliokh and Y. P. Bliokh, *Conservation of Angular Momentum, Transverse Shift, and Spin Hall Effect in Reflection and Refraction of an Electromagnetic Wave Packet*, Phys. Rev. Lett. **96**, 073903 (2006).

[66] J. C. Howell, D. J. Starling, P. B. Dixon, P. K. Vudyasetu, and A. N. Jordan, *Interferometric weak value deflections: Quantum and classical treatments*, Phys. Rev. A **81**, 033813 (2010).

[67] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, *Picoradian deflection measurement with an interferometric quasi-autocollimator using weak value amplification*, Opt. Lett. **36**, 1479 (2011).

[68] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White, *Measuring a Photonic Qubit without Destroying It*, Phys. Rev. Lett. **92**, 190402 (2004).

[69] A. Leggett and A. Garg, *Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?*, Phys. Rev. Lett. **54**, 857 (1985).

[70] A. N. Jordan, A. N. Korotkov, and M. Büttiker, *Leggett-Garg Inequality with a Kicked Quantum Pump*, Phys. Rev. Lett. **97**, 026805 (2006).

[71] N. S. Williams and A. N. Jordan, *Weak Values and the Leggett-Garg Inequality in Solid-State Qubits*, Phys. Rev. Lett. **100**, 026804 (2008).

[72] N. S. Williams and A. N. Jordan, *Erratum: Weak Values and the Leggett-Garg Inequality in Solid-State Qubits [Phys. Rev. Lett. 100, 026804 (2008)]*, Phys. Rev. Lett. **103**, 089902 (2009).

[73] J. P. Groen, D. Riste, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, *Partial-Measurement Backaction and Nonclassical Weak Values in a Superconducting Circuit*, Phys. Rev. Lett. **111**, 090506 (2013).

[74] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde, *Violation of the Leggett-Garg inequality with weak measurements of photons*, Proc. Natl. Acad. Sci. U.S.A. **108**, 1256 (2011).

[75] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, *Precision frequency measurements with interferometric weak values*, Phys. Rev. A **82**, 063822 (2010).

[76] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, *Continuous phase amplification with a Sagnac interferometer*, Phys. Rev. A **82**, 011802 (2010).

[77] M. Linuma, Y. Suzuki, G. Taguchi, Y. Kadoya, and H. F. Hofmann, *A method for weak measurement of photon polarization robust against experimental imperfections*, New J. Phys. **13**, 033041 (2011).

[78] Q. Wang, F.-W. Sun, Y.-S. Zhang, J.-Li, Y.-F. Huang, G.-C. Guo, *Experimental demonstration of a method to realize weak measurement of the arrival time of a single photon*, Phys. Rev. A **73**, 023814 (2006).

[79] D. Suter, M. Ernst, and R. R. Ernst, *Quantum time-translation machine, An experimental realization*, Mol. Phys. **78**, 95 (1993).

[80] J. M. Knight and L. Vaidman, *Weak measurement of photon polarization*, Phys. Lett. A **143**, 357 (1990).

[81] J. M. Hogan, J. Hammer, S. W. Chiow, S. Dickerson, D. M. Johnson, T. Kovachy, A. Sugarbaker, and M. A. Kasevich, *Precision angle sensor using an optical lever inside a Sagnac interferometer*, Opt. Lett. **36**, 1698 (2011).

[82] A. Feizpour, X. Xing and A.M. Steinberg, *Amplifying Single-Photon Nonlinearity Using Weak Measurements*, Phys. Rev. Lett. **107**, 133603 (2011).

[83] M. Hallaji, A. Feizpour, G. Dmochowski, J. Sinclair, and A. M. Steinberg, *Weak-value amplification of the nonlinear effect of a single photon*, Nat. Phys. **13**, 540 (2017).

[84] T. Durt, B.-G. Englert, I. Bengstsson, and K. Zyczkowski, *On mutually unbiased bases*, Int. J. Quantum Inf **8**, 535 (2010).

[85] J. Fischbach and M. Freyberger, *Quantum optical reconstruction scheme using weak values*, Phys. Rev. A **86**, 052110 (2012).

[86] S. Massar and S. Popescu, *Estimating preselected and postselected ensembles*, Phys. Rev. A **84**, 052106 (2011).

[87] M. Malik, M. Mirhosseini, M. P. J. Lavery, J. Leach, M. J. Padgett, and R. W. Boyd, *Direct measurement of a 27-dimensional orbital-angular-momentum state vector*, Nat. Commun. **5**, 3115 (2014).

[88] C. Bamber and J. S. Lundeen, *Observing Dirac's Classical Phase Space Analog to the Quantum State*, Phys. Rev. Lett. **122**, 070405 (2014).

[89] J. Z. Salvail, M. Agnew, A.S. Johnson, E. Bolduc, J. Leach, and R. W. Boyd, *Full characterization of polarization states of light via direct measurement*, Nat. Phot. **7**, 316 (2013).

[90] H. Kobayashi, K. Nonaka, and Y. Shikano, *Stereographical visualization of a polarization state using weak measurements with an optical-vortex beam*, Phys. Rev. A **89**, 053816 (2014).

[91] L. Maccone and C. C. Rusconi, *State estimation: A comparison between direct state measurement and tomography*, Phys. Rev. A **89**, 022122 (2014).

[92] S. Wu, *State tomography via weak measurements*, Sci. Rep. **3**, 1193 (2013).

[93] M. Mirhosseini, J. S. Lundeen, and R. W. Boyd, in *Quantum Photonics: Pioneering Advances and Emerging Applications*, edited by R. Boyd, S. Lukishova, and V. Zadkov (Springer Series in Optical Sciences, Switzerland, 2019).

[94] M. Mirhosseini, O. S. Magaña-Loaiza, S. M. H. Rafsanjani, and R. W. Boyd, *Compressive Direct Measurement of the Quantum Wave Function*, Phys. Rev. Lett. **113**, 090402 (2014).

[95] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen, *Direct Measurement of the Density Matrix of a Quantum System*, Phys. Rev. Lett. **117**, 120401 (2016).

[96] Y. Turek, *Direct measurement methods of density matrix of an entangled quantum state*, J. Phys. Commun. **4**, 075007 (2020).

[97] N. Brunner, V. Scarani, M. Wegmüller, M. Legré, and N. Gisin, *Direct Measurement of Superluminal Group Velocity and Signal Velocity in an Optical Fiber*, Phys. Rev. Lett. **93**, 203902 (2004).

[98] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, *Fast Light, Slow Light, and Phase Singularities: A Connection to Generalized Weak Values*, Phys. Rev. Lett. **92**, 043601 (2004).

[99] M. Büttiker, *Larmor precession and the traversal time for tunneling*, Phys. Rev. B **27**, 6178 (1983).

[100] A. M. Steinberg, *How Much Time Does a Tunneling Particle Spend in the Barrier Region?*, Phys. Rev. Lett. **74**, 2405 (1995).

[101] R. Ramos, D. Spierings, I. Racicot, and A. M. Steinberg, *Measurement of the time spent by a tunnelling atom within the barrier region*, Nature **583**, 529 (2020).

[102] D. Sokolovski and J. N. L. Connor, *Path-integral analysis of the time delay for wave-packet scattering and the status of complex tunneling times*, Phys. Rev. A **42**, 6512 (1990).

[103] H. G. Winful, *Tunneling time, the Hartman effect, and superluminality: A proposed resolution of an old paradox*, Phys. Rep. **436**, 1 (2006).

[104] O. S. Magaña-Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, *Detection of Time-Reversal Symmetry Breaking in the Noncentrosymmetric Superconductor Re6Zr Using Muon-Spin Spectroscopy*, Phys. Rev. Lett. **112**, 200401 (2014).

[105] N. Brunner and C. Simon, *Measuring Small Longitudinal Phase Shifts: Weak Measurements or Standard Interferometry?*, Phys. Rev. Lett. **105**, 010405 (2010).

[106] Y. Aharonov and L. Vaidman, *Properties of a quantum system during the time interval between two measurements*, Phys. Rev. A **41**, 11 (1990).

[107] M. Berry, N. Zheludev, Y. Aharonov, F. Colombo, I. Sabadini, D. C. Struppa, J. Tollaksen, E. T. F. Rogers, F. Qin, M. Hong *et al.*, *Roadmap on superoscillations*, J. Opt. **21**, 053002 (2019).

[108] Y. Aharonov, F. Colombo, I. Sabadini, D. C. Struppa, and J. Tollaksen, *Some mathematical properties of superoscillations*, J. Phys. A: Math: Theor. **44**, 065304 (2011).

[109] J. Dressel, S. Agarwal, and A. N. Jordan, *Contextual Values of Observables in Quantum Measurements*, Phys. Rev. Lett. **104**, 240401 (2010).

[110] S. Parrot, *Quantum weak values are not unique. What do they actually measure?*, arXiv:0909.0295; arXiv:0908.0035.

[111] D. R. M. Arvidsson-Shukur, N. Yunger Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, *Quantum advantage in postselected metrology*, Nat. Commun. **11**, 3775 (2020).

[112] I. R. Senitzky, *Dissipation in Quantum Mechanics. The Harmonic Oscillator*, Phys. Rev. **119**, 670 (1960).

[113] M. J. Collet and C. W. Gardiner, *Squeezing of intracavity and traveling-wave light fields produced in parametric amplification*, Phys. Rev. A. **30**, 1386 (1984).

[114] C. W. Gardiner and M. J. Collet, *Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation*, Phys. Rev. A **31**, 3761 (1985).

[115] R. Lang, M. O. Scully, and W. E. Lamb Jr., *Why is the Laser Line So Narrow? A Theory of Single-Quasimode Laser Operation*, Phys. Rev. A **7**, 1788 (1973).

[116] S. M. Barnett and P. M. Radmore, *Quantum theory of cavity quasimodes*, Opt. Commun. **68**, 364 (1988).

[117] B. J. Dalton, S. M. Barnett, and P. L. Knight, *Quasi mode theory of macroscopic canonical quantization in quantum optics and cavity quantum electrodynamics*, J. Mod. Opt. **46**, 1315 (1999).

[118] S. M. Dutra and G. Nienhuis, *A Hamiltonian for cavity decay*, Acta Phys. Slovaca **50**, 275 (2000).

[119] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* (John Wiley and Sons, Inc., New York, 2001).

[120] S. M. Kay, *Fundamentals of Statistical Signal Processing. Estimation Theory* (Prentice-Hall, New Jersey, 1993).

[121] M. G. A. Paris, *Quantum estimation for quantum technology*, Int. J. Quantum. Inform. **7**, 125 (2009).

[122] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

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

[124] S. L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states*, Phys. Rev. **72**, 3439 (1994).

[125] S. L. Braunstein, C. M. Caves and G.J. Milburn, *Generalized Uncertainty Relations: Theory, Examples, and Lorentz Invariance*, Ann. Phys. (N. Y.) **247**, 135 (1996).

[126] K. Jacobs, *Quantum Measurement Theory and its Applicactions* (Cambridge University Press, Cambridge, 2014).

[127] L. Susskind and J. Glogower, *Quantum mechanical phase and time operator*, Phys. Phys. Fiz **1**, 49 (1964).

[128] J. Schwinger, *Quantum Theory of Angular Momentum* (Academic Press, New York, 1965).

[129] B. Yurke, S. L. McCall, and J. R. Klauder, *SU(2) and SU(1,1) interferometers*, Phys. Rev. A **33**, 4033 (1986).

[130] W. K. Wooters, *Statistical distance and Hilbert space*, Phys. Rev. D **23**, 357 (1981).

[131] D. J. C. Bures, *An extension of Kakutani's Theorem on Infinite Product Measures to the Tensor Product of Semifinite w*-Algebras*, Trans. Am. Math. Soc. **135**, 199 (1969).

[132] A. Uhlmann, *The "transition probability" in the state space of a *-algebra*, Rep. Math. Phys. **9**, 273 (1976).

[133] M. Hübner, *Explicit computation of the Bures distance for density matrices*, Phys. Lett. A **163**, 239 (1992).

[134] H-J. Sommers and K. Zyczkowski, *Bures volume of the set of mixed quantum states*, J. Phys. A **36**, 10083 (2003).

[135] H. F. Hofmann, *Uncertainty limits for quantum metrology obtained from the statistics of weak measurements*, Phys. Rev. A **83**, 022106 (2011).

[136] D. Bohm, *A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I*, Phys. Rev. Lett. **85**, 166 (1952).

[137] D. Bohm, *A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II*, Phys. Rev. Lett. **85**, 180 (1952).

[138] E. Madelung, *Eine anschauliche Deutung der Gleichung von Schrödinger*, Naturwissenshaften **14**, 1004 (1926).

[139] E. Madelung, *Quantentheorie in hydrodynamischer Form*, Z. Phys. **40**, 322 (1927).

[140] G. Arfken, *Mathematical Methods For Physicists* (Academic Press, San Diego, 1985).

[141] L. Zhang, A. Datta and I. A. Walmsley, *Precision Metrology Using Weak Measurements*, Phys. Rev. Lett. **114**, 210801 (2015).

[142] G. C. Knee and E. M. Gauger, *When Amplification with Weak Values Fails to Suppress Technical Noise*, Phys. Rev. X **4**, 011032 (2014).

[143] G. B. Alves, B. M. Escher, R. L. de Matos Filho, N. Zagury, and L. Davidovich, *Weak-value amplification as an optimal metrological protocol*, Phys. Rev. A **91**, 062107 (2015).

[144] L. Xu, Z. Liu, A. Datta, G. C. Knee, J. S. Lundeen, Y-Q. Lu, and L. Zhang, *Approaching Quantum-Limited Metrology with Imperfect Detectors by Using Weak-Value Amplification*, Phys. Rev. Lett. **125**, 080501 (2020).

[145] J. Sinclair, M. Hallaji, A.M. Steinberg, J. Tollaksen, and A. N. Jordan, *Weak-value amplification and optimal parameter estimation in the presence of correlated noise*, Phys. Rev. A **96**, 052128 (2017).

[146] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, *Technical Advantages for Weak-Value Amplification: When Less Is More*, Phys. Rev. X **4**, 011031 (2014).

[147] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).

[148] R. A. Fisher, *Theory of Statistical Estimation. Mathematical Proceedings of the Cambridge Philosophical Society*, Proc. Camb. Soc. **22**, 700 (1925).

[149] V. B. Braginsky and Y. I. Vorontsov, *Quantum-mechanical limitations in macroscopic experiments and modern experimental technique*, Sov. Phys. Usp. **17**, 644 (1975).

[150] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, Science **306**, 1330 (2004).

[151] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, Phys. Rev. Lett. **96**, 010401 (2006).

[152] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in quantum metrology*, Nat. Photonics **5**, 222 (2011).

[153] M. Jarzyna and R. Demkowicz-Dobrzański, *True precision limits in quantum metrology*, New J. Phys. **17**, 013010 (2015).

[154] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Cavity Optomechanics. Nano- and Micromechanical Resonators Interacting with Light* (Springer-Verlag, Berlin Heidelberg, 2014).

[155] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Cavity optomechanics*, Rev. Mod. Phys. **86**, 1391 (2014).

[156] C. M. Caves, *Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer*, Phys. Rev. Lett. **45**, 75 (1980).

[157] C. M. Caves, *Quantum-mechanical noise in an interferometer*, Phys. Rev. D **23**, 1693 (1981).

[158] C. M. Caves, *Quantum limits on noise in linear amplifiers*, Phys. Rev. D **26**, 1817 (1982).

[159] V. B. Braginsky, V. P. Mitrofanov, and V. I. Panov, *Systems with small dissipation* (University of Chicago Press, Chicago, 1977).

[160] V. B. Braginsky and F. Y. Khalili, *Quantum Measurements* (Cambridge University Press, Cambridge, 1995).

[161] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, *Introduction to quantum noise, measurement, and amplification*, Rev. Mod. Phys. **82**, 1155 (2010).

[162] A. A. Clerk, F. Marquardt, and K. Jacobs, *Back-action evasion and squeezing of a mechanical resonator using a cavity detector*, New. J. Phys. **10**, 095010 (2008).

[163] C. M. Caves, K. S. Thorne, R. W. Drever, V. D. Sandberg, and M. Zimmermann, *On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle*, Rev. Mod. Phys. **52**, 341 (1980).

[164] A. M. Jayich, J. C. Sankey, B. M. Zwickl, C. Yang, J. D. Thompson, S. M. Girvin, A. A. Clerk, F. Marquardt, and J. G. E. Harris, *Dispersive optomechanics: a membrane inside a cavity*, New. J. Phys. **10**, 095008 (2008).

[165] V. B. Braginsky and F. Y. Khalili, *Quantum nondemolition measurements: the route from toys to tools*, Rev. Mod. Phys. **68**, 1 (1996).

[166] V. B. Braginsky, Y. I. Vorontsov, and K. Thorne, *Quantum Nondemolition Measurements*, Science **209**, 547 (1980).

[167] M. Aspelmeyer, P. Meystre, and K. Schwab, *Quantum optomechanics*, Phys. Today **65**, 29 (2012).

[168] P. R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors* (World Scientific Publishing, Singapore, 1994).

[169] J. D. Teufel, T. Donner, D. Li, J. W. Harlow, M. S. Allman, K. Cicak, A. J. Sirois, J. D. Whittaker, K. W. Lehnert, and R. W. Simmonds, *Sideband cooling of micromechanical motion to the quantum ground state*, Nature **475**, 359 (2011).

[170] J. Chan, T. P. Mayer Alegre, A. H. Safavi-Naeini, J. T. Hill, A. Krause, S. Gröblacher, M. Aspelmeyer, and O. Painter, *Laser cooling of a nanomechanical oscillator into its quantum ground state*, Nature **478**, 89 (2011).

[171] A. D. O'Connell, M. Hofheinz, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, J. M. Martinis, and A. N. Cleland, *Quantum ground state and single-phonon control of a mechanical resonator*, Nature **464**, 697 (2010).

[172] A. H. Safavi-Naeini, J. Chan, J. T. Hill, T. P. Mayer Alegre, A. Krause, and O. Painter, *Observation of Quantum Motion of a Nanomechanical Resonator*, Phys. Rev. Lett. **108**, 033602 (2012).

[173] B. P. Abbott, *Observation of a kilogram-scale oscillator near its quantum ground state*, New J. Phys. **11**, 073021 (2009).

[174] R. Penrose, *On Gravity's role in Quantum State Reduction*, Gen. Relativ. Gravit. **28**, 581 (1996).

[175] R. Penrose, *Quantum [un]speakables: from Bell to quantum information* (Springer, New York, 2000).

[176] I. Pikovski, M. R. Vanner, M. Aspelmeyer, M. S. Kim, and C. Brukner, *Probing Planck-scale physics with quantum optics*, Nat. Phys. **8**, 393 (2012).

[177] S. Bose, K. Jacobs, and P. L. Knight, *Scheme to probe the decoherence of a macroscopic object*, Phys. Rev. A **59**, 3204 (1999).

[178] S. Bose, K. Jacobs, and P. L. Knight, *Preparation of nonclassical states in cavities with a moving mirror*, Phys. Rev. A **56**, 4175 (1997).

[179] W. Marshall, C. Simon, R. Penrose, and D. Bouwmeester, *Towards Quantum Superpositions of a Mirror*, Phys. Rev. Lett. **91**, 130401 (2003); Erratum Phys. Rev. Lett. **91**, 159903 (2003).

[180] B. Pepper, E. Jeffrey, R. Ghobadi, C. Simon, and D. Bouwmeester, *Macroscopic superpositions via nested interferometry: finite temperature and decoherence considerations*, New J. Phys. **14**, 115025 (2012).

[181] F. He, J. Liu, and K-D. Zhu, *Optomechanical atomic force microscope*, Nanotechnology **32**, 085505 (2021).

[182] S. Forstner, S. Prams, J. Knittel, E. D. van Ooijen, J. D. Swaim, G. I. Harris, A. Szorkovszky, W. P. Bowen, and H. Rubinsztein-Dunlop, *Cavity Optomechanical Magnetometer*, Phys. Rev. Lett. **108**, 120801 (2012).

[183] A. G. Krausse, M. Winger, T. D. Blasius, Q. Li, and O. Painter, *A high-resolution microchip optomechanical accelerometer*, Nat. Photonics **6**, 768 (2012).

[184] M. R. Vanner, I. Pikovski, G. D. Cole, M. S. Kim, C. Brukner, K. Hammerer, G. J. Milburn, and M. Aspelmeyer, *Pulsed quantum optomechanics*, Proc. Natl. Acad. Sci. U.S.A. **108**, 16182 (2011).

[185] D. Hunger, S. Camerer, M. Korppi, A. Jöckel, T. Hänsch, and P. Treutlein, *Coupling ultracold atoms to mechanical oscillators*, C. R. Phys. **12**, 871 (2011).

[186] K. Hammerer, M. Aspelmeyer, E. Polzik, and P. Zoller, *Establishing Einstein-Poldosky-Rosen Channels between Nanomechanics and Atomic Ensembles*, Phys. Rev. Lett. **102**, 020501 (2009).

[187] K. Hammerer, A. S. Sorensen, and E. Polzik, *Quantum interface between light and atomic ensembles*, Rev. Mod. Phys. **82**, 1041 (2010).

[188] K. H. Lee, T. G. McRae, G. I. Harris, J. Knittel, and W. P. Bowen, *Cooling and Control of a Cavity Optoelectromechanical System*, Phys. Rev. Lett. **104**, 123604 (2010).

[189] M. Wallquist, K. Hammerer, P. Rabl, M. Lukin, and P. Zoller, *Hybrid quantum devices and quantum engineering*, Phys. Scr. **T137**, 014001 (2009).

[190] C. Dong, Y. Wang, and H. Wang, *Optomechanical interfaces for hybrid quantum networks*, Natl. Sci. Rev. **2**, 510 (2015).

[191] D. E. Chang, A. H. Safavi-Naeini, M. Hafezi, and O. Painter, *Slowing and stopping light using an optomechanical crystal array*, New J. Phys. **13**, 023003 (2011).

[192] A. Schliesser, *Cavity Optomechanics and Optical Frequency Comb Generation with Silica Whispering-Gallery-Mode Microresonators*, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2009.

[193] G. S. Agarwal and S. Huang, *Electromagnetically induced transparency in mechanical effects of light*, Phys. Rev. A **81**, 041803 (2010).

[194] S. Weis, R. Riviere, S. Deleglise, E. Gavartin, O. Arcizet, A. Schliesser, and T. Kippenberg, *Optomechanically Induced Transparency*, Science **330**, 1520 (2010).

[195] C. K. Law, *Interaction between a moving mirror and radiation pressure: A Hamiltonian formulation*, Phys. Rev. A **51**, 2537 (1995).

[196] M. M. Nieto, *Displaced and squeezed number states*, Phys. Lett. A, **229**, 135 (1997).

[197] M. Orszag, *Quantum Optics. Including Noise Reduction, Trapped Ions, Quantum Trajectories, and Decoherence* (Springer-Verlag, Berlin Heidelberg, 2008).

[198] C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2005).

[199] S. Mancini, V. I. Man'ko, and P. Tombesi, *Ponderomotive control of quantum macroscopic coherence*, Phys. Rev. A **55**, 3042 (1997).

[200] B-G. Englert, *Fringe Visibility and Which-Way Information: An Inequality*, Phys. Rev. Lett. **77**, 2154 (1996).

[201] G. Li, T. Wang, and H-S. Song, *Amplification effects in optomechanics via weak measurements*, Phys. Rev. A **90**, 013827 (2014).

[202] G. Li, L-B. Chen, X-M. Lin, and H-S. Song, *Weak measurement amplification in optomechanics via a squeezed coherent state pointer*, J. Phys. B: At. Mol. Opt. Phys. **48**, 165504 (2015).

[203] G. Li, T. Wang, M-Y. Ye, and H-S. Song, *Weak measurement combined with quantum delayed-choice experiment and implementation in optomechanical system*, Eur. Phys. J. D **69**, 266 (2015).

[204] M. O. Scully and K. Drühl, *Quantum eraser: A proposed photon correlation experiment concerning observation and "delayed choice" in quantum mechanics*, Phys. Rev. A **25**, 2208 (1982).

[205] D. D. Bhaktasala Rao, S. A. Momenzadeh, and J. Wrachtrup, *Heralded Control of Mechanical Motion by Single Spins*, Phys. Rev. Lett., **117**, 077203 (2016).

[206] R. W. Peterson, T. P. Purdy, N. S. Kampel, R. W. Andrews, P.-L. Yu, K. W. Lehnert, and C. A. Regal, *Laser Cooling of a Micromechanical Membrane to the Quantum Backaction Limit*, Phys. Rev. Lett. **116**, 063601 (2016).

[207] M. Bader, S. Heugel, A. L. Chekhov, M. Sondermann, and G. Leuchs, *Efficient coupling to an optical resonator by exploiting time-reversal symmetry*, New J. Phys., **15**, 123008 (2013).

[208] M. Sondermann, R. Maiwald, H. Konermann, N. Lindlein, U. Peschel and G. Leuchs, *Design of a mode converter for efficient light-atom coupling in free space*, Appl. Phys. B **89**, 489 (2007).

[209] B. C. Sanders, *Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement*, Phys. Rev. A **40**, 2417 (1989).

[210] C. Simon and E. Polzik, *Fock-state view of weak-value measurements and implementation with photons and atomic ensembles*, Phys. Rev. A **83**, 040101 (2011).

[211] C. Whittle, E. D. Hall, S. Dwyer, N. Mavalvala, V. Sudhir, R. Abbott, A. Ananyeva, C. Austin, L. Barsotti, J. Betzwieser *et al.*, *Approaching the motional ground state of a 10-kg object*, Science **372**, 1333 (2021).