Università degli Studi di Padova

Dipartimento di Fisica e Astronomia "G. Galilei"

Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Magistrale in Fisica

Direct Reconstruction of the Quantum Density Operator via Measurements of Arbitrary Strength

Author:
Giulio Foletto

Supervisor:
Prof. Giuseppe Vallone

September 2017

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "G. Galilei"
Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Magistrale in Fisica

Direct Reconstruction of the Quantum Density Operator via Measurements of Arbitrary Strength

by Giulio Foletto

Abstract

One of the fundamental problems of experimental quantum physics is the determination of the state of a system, which provides its accurate description and contains all the available information about it. The standard solution is *quantum state tomography*, a technique that employs a large number of measurements of complementary observables to reconstruct a general density operator. Recently, new protocols have been proposed which allow to obtain the matrix elements one by one with fewer and less complicated observations by exploiting so-called *weak measurements*, that perturb the system less than ideal projections at the cost of being less precise.

This thesis presents an evolution of these schemes that eschews the weakness of the measurement, thus gaining in the accuracy of the results and in practical feasibility. We realized an experiment that uses these methods to reconstruct the density operator of the polarization state of light in single photons regime.

Acknowledgements

I would like to thank first my supervisor Prof. Giuseppe "Pino" Vallone for having given me the opportunity to work on this project and for having provided direction and much needed help. Not only he made this work possible, that would be true of any thesis supervisor, but he made it *enjoyable* with his passion and energy. He and Prof. Paolo Villoresi are at the helm of a top notch research group in one of the most promising fields of physics and technology, but their merit is also the creation of a great team of people with whom my collaboration has been easy and fun.

On this note I would also like to thank everyone working at the LUXOR laboratory for the help and for having accepted me as part of the family. In particular, Luca Calderaro has been fundamental for the experimental part of this thesis and I thank him for having introduced me to the lab life and for having taught me most of what I now know about the technical aspects of quantum optics.

Finally I thank my family, friends and girlfriend for having put up with me in the last few months, during which most of my replies were "Sorry, I'm at the lab right now" or "Back to writing my thesis, bye!". Despite my love for physics, without you this work and the five years before it would have been much more difficult.

Contents

Chapter 1

Introduction

At the very foundation of every physical theory stands the concept of *state*. A state is often a mathematical entity that, together with a physical interpretation, can describe a system and allows the determination of (at least some of) its properties. It does so by restricting such properties "by as many conditions or data as theoretically possible" [1, p. 11]. For instance, thermodynamics can define the state of an ideal gas with the variables P (pressure), V (volume) and T (absolute temperature), whereas in classical mechanics the phase space of positions and momenta is the domain of the possible states of a particle. Given perfect knowledge of a state and of the laws governing its evolution (e.g. a sequence of transformations of the gas, or the Hamilton's equations of motion), many theories can find the entire future and past of a system.

Quantum physics is by nature not so deterministic. On account of the *superposition principle*, its states reside in a vector structure, precisely a Hilbert space [2, Ch. 2, Sec. 1], but are described by rays so that, intuitively, the sum of a state with itself does not change it. Unit vectors are conventionally chosen as representatives of such rays and contain all the information that can ideally be extracted from the particular configuration the system is in. Even with this knowledge, only the outcomes of some measurements can be predicted for sure, while others remain aleatory because of the *uncertainty principle*.

Moreover, a distinction has to be made between these cases (*pure* states) and the ones in which even less information is available (*mixed* ones). For the latter situations, more general descriptions such as that of *density operators*, must be introduced. Just as with rays in a Hilbert space, with this formalism it is possible to calculate the mean values and all other statistical moments of any measurement from a single mathematical object.

But then how does one find it? *Quantum state tomography (QST)* comes to the aid. It is a set of techniques that aim to reconstruct the state of a system by means of performing specific measurements on many identical copies of it. It has been used extensively for the purpose of characterizing devices that produce quantum systems in a predetermined state, but its importance has recently been growing in the promising fields of *quantum information* and *communication*. Here, states are used to encode and transmit a message, so that their precise knowledge can become relevant for the success of an exchange or the certification of the involved parties. Moreover, the *quantum computer* of the future will need verification of the states it uses as inputs for its calculations and validation for the reading of its results. In more fundamental research, QST can be used as the final step of *quantum simulations* to find and characterize the ground state of complicated Hamiltonians.

Up until recently, QST of d-dimensional states has involved $O(d^2)$ measurements and a linear reconstruction of the entire density operator. Yet, in 2011 Lundeen *et al.* proposed a scheme that can directly obtain each single element of a quantum wavefunction [3] or density operator [4] in a fixed number of measurements and a few simple calculations that in general do not depend on the nature of the system to be studied or the dimension d of its associated Hilbert space. However, this protocol is based on the concepts of *weak values* and weak couplings between the system and an *ancilla*, the measurement device that actually provides the needed information, but new theoretical work suggests that this weakness is not necessary and even counter-productive for tomography [5].

The main goal of this thesis is to expand on these ideas and show that the density operator can be directly measured within the framework of ancillae but without the use of weak couplings between systems and with better results in terms of accuracy and precision. After a brief review of the necessary formalism and of the procedures of standard QST, I will explain the fundamentals of weak values and measurements and their use in state reconstruction schemes. I shall then propose a new protocol based on strong couplings and present an experimental realization aimed at finding the polarization of a photon system.

1.1 Mixed States and Density Operators

In order for this thesis to be as self-contained as possible, the rest of this chapter will provide the rudiments of the formalism that will be used in the following dissertation. However, I shall not detail the mathematics of pure states more than I have already done and I will assume that the reader is familiar with the very basic principles of quantum mechanics, the Dirac notation and concepts such as *observables*, *projectors*, *unitary evolutions*, *mean values*, *ideal measurements*, *uncertainty* and *entanglement* [6]. Yet, since the quantum state is clearly fundamental for this work, I will focus on how it is generalized beyond the restrictions of purity and what are the main ways in which it is usually represented.

No realistic process is so accurate that it can prepare perfectly pure states, therefore all the quantum systems that can be encountered in practice should be considered *statistical mixtures*, better known as *mixed states*. In order to define them, consider the situation in which one knows that the system under study can be represented by one of n possible *kets* (rays) $|\phi_1\rangle..|\phi_n\rangle$, each with an associated probability $p_1..p_n$ such that $\sum_{k=1}^{n} = 1$. The most common description of this configuration is offered by the density operator

$$\hat{\varrho} \equiv \sum_{k=1}^{n} p_k |\phi_k\rangle\langle\phi_k| \tag{1.1}$$

which can be seen as a complex $d \times d$ matrix if d is the dimension of the Hilbert space of the system. The presence of a *bra* $\langle\phi_k|$ for each corresponding ket $|\phi_k\rangle$ makes this expression manifestly invariant for phase multiplication, another advantage of this formalism over that of Hilbert space vectors.

Its other fundamental properties are:

- $\hat{\varrho}$ is hermitian: $\hat{\varrho} = \hat{\varrho}^\dagger$. Consequently, like any observable, it can be spectrally decomposed and expressed as a real combination of projectors $\hat{\varrho}' = \sum_j \lambda_j \sum_r |\lambda_{j,r}\rangle\langle\lambda_{j,r}|$.

Although $\hat{\varrho}$ and $\hat{\varrho}'$ apparently refer to different mixtures, they represent the same mixed state.

- $\hat{\varrho}$ is positive semidefinite: $\langle\psi|\hat{\varrho}|\psi\rangle \geq 0 \ \forall\psi$ or equivalently $\lambda_j \geq 0 \ \forall j$.
- $\hat{\varrho}$ has unit trace: $\mathrm{Tr}(\hat{\varrho}) = 1$. This condition is equivalent to the normalization request for vectors that describe pure states and is deeply integrated into the formalism.

It is clear that, as a consequence of these characteristics, $\mathrm{Tr}(\hat{\varrho}^2) \leq 1$. This quantity is called the *purity* of the system and is equal to 1 if and only if the state described by $\hat{\varrho}$ is pure, whereas it takes its minimum value d^{-1} in the *completely mixed state*. Notice that $\hat{\varrho}$ has $d^2 - 1$ real degrees of freedom.

All the rules of quantum mechanics that are defined in the ket language can easily be extended to density operators. For example the evolution of a state according to unitary transformation U is

$$\hat{\varrho} \to \sum_{k=1}^{n} p_k U|\phi_k\rangle\langle\phi_k|U^\dagger = U\hat{\varrho}U^\dagger \tag{1.2}$$

whereas the m-th statistical moment of the measurement of observable $\hat{A}$ is simply

$$\langle\hat{A}^m\rangle_{\hat{\varrho}} = \sum_{k=1}^{n} p_k\langle\hat{A}^m\rangle_{\phi_k} = \sum_{k=1}^{n} p_k\langle\phi_k|\hat{A}^m|\phi_k\rangle = Tr(\hat{\varrho}\hat{A}^m) \tag{1.3}$$

This is reminiscent of the *Born rule*, according to which the probability of obtaining the j-th outcome (i.e. the j-th eigenvalue) of the measurement of $\hat{A}$ is

$$\mathrm{Pr}_j = \sum_k p_k\|\hat{\Pi}_j|\phi_k\rangle\|^2 = \mathrm{Tr}(\hat{\varrho}\hat{\Pi}_j) \tag{1.4}$$

and if this happens, the state collapses onto

$$\frac{\hat{\Pi}_j\hat{\varrho}\hat{\Pi}_j}{\mathrm{Pr}_j} \tag{1.5}$$

where the denominator is needed for renormalization of the trace. In what follows, I will mostly use the density operator formalism for the important results, and revert to unit vectors when simplicity is a priority.

There are also other mathematical entities that contain the same information of $\hat{\varrho}$, such as the *Dirac distribution*, which is of tangential interest for this work [7, 8]. Originally a way to represent operators using a set of non-commuting observables, it can be seen as the Fourier transform of the density matrix. Indeed if $\{|b_l\rangle \equiv \frac{1}{\sqrt{d}}\sum_j |a_j\rangle e^{i\frac{2\pi jl}{d}} \mid l = 0..d-1\}$ is the *Fourier basis* of $\{|a_j\rangle \mid j = 0..d-1\}$, element jl of the (discrete) Dirac distribution that describes the same state of $\hat{\varrho}$ coincides with

$$D_{jl} = \langle a_j|\hat{\varrho}|b_l\rangle\langle b_l|a_j\rangle = \frac{1}{d}\sum_{k=0}^{d-1}\langle a_j|\hat{\varrho}|a_k\rangle e^{i\frac{2\pi l(k-j)}{d}} \tag{1.6}$$

Other descriptions for quantum states are used in specific fields [9] such as the *Wigner distribution* [10, 11] or the *Moyal M-Function* [12] in statistical mechanics, the *Husimi*

Q-Function [13] or the *Glauber-Sudarshan P-Representation* [14, 15] in quantum optics. Formulations like these are rarely used because they do not share all the properties of a good probability distribution. Nonetheless, in some cases they can be an intermediate step to retrieve the density operator, as in one of the protocols that I shall detail in the next chapter, or in other methods [16, 17].

1.2 Purification of a Mixed State

In order to better understand the nature of mixed states, some new mathematical tools must be introduced. Given a tensor product Hilbert space $\mathscr{H}_A \otimes \mathscr{H}_B$ and a density operator $\hat{\varrho}$ on it, the *partial trace* of $\hat{\varrho}$ on $\mathscr{H}_B$ is defined as:

$$\mathrm{Tr}_{\mathscr{H}_B}(\hat{\varrho}) \equiv \sum_{l=1}^{d_B} \langle b_l|\hat{\varrho}|b_l\rangle = \hat{\varrho}_A \tag{1.7}$$

where $\{|b_l\rangle\}$ is a basis for $\mathscr{H}_B$. Any information that can be extracted by focusing only on subsystem $\mathscr{H}_A$ is contained in $\hat{\varrho}_A$. This means that the mean value of any observable that acts only on $\mathscr{H}_A$ (and therefore is written as an operator on $\mathscr{H}_A \otimes \mathscr{H}_B$ as $A \otimes \mathbb{1}$) can be calculated by ignoring (tracing out) $\mathscr{H}_B$:

$$\langle \hat{A} \otimes \mathbb{1} \rangle_{\hat{\varrho}} = \mathrm{Tr}(\hat{\varrho}_A \hat{A}) \tag{1.8}$$

I now show that any mixed state can be seen as the partial trace of a pure state in a larger Hilbert space. Indeed let $\hat{\varrho}_A$ be a density operator on $\mathscr{H}_A$ and $\hat{\varrho}_A = \sum_j \lambda_j \sum_r |\lambda_{j,r,A}\rangle\langle\lambda_{j,r,A}|$ be its spectral decomposition. One can define $|\psi_{AB}\rangle = \sum_j \sqrt{\lambda_j} \sum_r |\lambda_{j,r,A}\rangle \otimes |b_{j,r,B}\rangle$ in the tensor product space $\mathscr{H}_A \otimes \mathscr{H}_B$, with $\mathscr{H}_B$ being any d_A-dimensional Hilbert space and $\{|b_{j,r,B}\rangle\}$ its basis. Then

$$\hat{\varrho}_A = \mathrm{Tr}_{\mathscr{H}_B}(|\psi_{A,B}\rangle\langle\psi_{A,B}|) \tag{1.9}$$

This means that mixed states are a consequence of the fact that quantum systems interact with one another: if it was possible to truly isolate them, they would remain pure. This is one of the most important obstacles to the realization of a quantum computer, in which the environment can inconveniently change the results of calculations and cause loss of information or errors.

However, since every cloud has a silver lining, this fact can also be exploited to practically generate states of very low purity. If it is possible to produce states that are almost pure and entangled in a bipartite system, one can ignore all measurements in a subsystem and get a mixed state in the other. This technique has been effectively used in the experiment that will be described in Chapter 4.

It is interesting to observe that after the introduction of the principles of quantum mechanics, many generalizations had to be added to better describe the experimental evidence. Mixed states are one of them, others are for example *completely positive maps*, which expand the evolution of a system out of the bounds of unitary transformations, or *POVMs*, which provide a model of measurement that is more realistic than that of projections. However, in all these cases ideality can be restored by extending the system under study. The principles only seem incomplete when we are missing the bigger picture.

1.3 The Bidimensional Case: The Qubit

In order to simplify the understanding of this formalism and to introduce some important notations, I will now provide some examples in the easiest possible case: the *qubit*. This term is vastly used, especially in the field of quantum information, as a synonym for two-level quantum system, and therefore is applied to many different physical implementations, such as spin 1/2 of a particle, currents in superconductors and photon polarizations. Since the Hilbert space of a qubit is of course bidimensional ($d = 2$), one can define a basis with two orthonormal pure states $\{|0\rangle, |1\rangle\}$ and numerically represent them with the unit vectors of their coordinates in $\mathbb{C}^2$:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.10}$$

where an irrelevant phase factor has been neglected. Any complex linear combination which preserves the normalization request is still a pure state:

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \qquad \text{with} \qquad |\alpha|^2 + |\beta|^2 = 1 \tag{1.11}$$

From these, it is possible to write the corresponding density matrices:

$$\hat{\varrho}_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \hat{\varrho}_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \hat{\varrho}_\psi = |\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \alpha\beta^\star \\ \alpha^\star\beta & |\beta|^2 \end{pmatrix} \tag{1.12}$$

All the aforementioned properties hold true and in these cases $\mathrm{Tr}(\hat{\varrho}^2) = 1$ because the considered states are pure.

It is paramount to understand the profound difference between a linear combination of pure states and one of their density matrices. Let, for example, $|+\rangle$ be the pure state with $\alpha = \beta = \frac{1}{\sqrt{2}}$:

$$|+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \rightarrow \qquad \hat{\varrho}_+ = |+\rangle\langle +| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{1.13}$$

A measurement in the $\{|0\rangle, |1\rangle\}$ basis will result in either of the two outcomes with 50% probability, but there exists at least one observable the measurement of which is deterministically predictable, that is, of course, the one which has $|+\rangle$ and its orthogonal state $|-\rangle$ as eigenvectors. Consider instead a combination of $\hat{\varrho}_0$ and $\hat{\varrho}_1$ such as

$$\hat{\varrho}_M = \frac{1}{2}(\hat{\varrho}_0 + \hat{\varrho}_1) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.14}$$

There is no (non-trivial) measurement for which one outcome has unitary probability, and this state is completely mixed (also said *totally unpolarized*) and $\mathrm{Tr}(\hat{\varrho}_M^2) = \frac{1}{2}$. In order to exemplify the process of purification, I point out that the two-qubit pure state $|\psi_{A,B}\rangle = \frac{1}{\sqrt{2}}(|0_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle)$ is such that $\hat{\varrho}_M = \mathrm{Tr}_{\mathscr{H}_B}(|\psi_{A,B}\rangle\langle\psi_{A,B}|)$.

There is another vastly used notation for two-level states. Since all possible density operators of qubits are 2×2 complex hermitian matrices, they can be expressed as real

linear combinations of the three Pauli matrices $\hat{\sigma}_x,\ \hat{\sigma}_y,\ \hat{\sigma}_z$ and the identity $\mathbb{1}_2$.

$$\mathbb{1}_2 \equiv \hat{\sigma}_0 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \hat{\sigma}_x \equiv \hat{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{\sigma}_y \equiv \hat{\sigma}_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{\sigma}_z \equiv \hat{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.15}$$

The necessity that $\mathrm{Tr}(\hat{\varrho}) = 1$ fixes the coefficient of $\mathbb{1}_2$ at $\frac{1}{2}$:

$$\hat{\varrho} = \frac{1}{2}(\mathbb{1}_2 + \vec{r} \cdot \vec{\hat{\sigma}}) \tag{1.16}$$

where the request for positivity imposes $r^2 \leq 1$.

This shows that the space of states of two-level systems is in one-to-one correspondence with a ball in $\mathbb{R}^3$ called the *Bloch sphere*. From (1.16) it follows that $\mathrm{Tr}(\hat{\varrho}^2) = \frac{1}{2}(1+r^2)$, therefore $r = 1$ for pure states and $r = 0$ for completely mixed ones. The Bloch sphere visualization is useful not just for purity, but also to find eingenvalues and eigenvectors of a density operator because

$$\hat{\varrho} = \left(\frac{1+r}{2}\right)\frac{1}{2}\left(\mathbb{1}_2 + \frac{\vec{r}}{r} \cdot \vec{\hat{\sigma}}\right) + \left(\frac{1-r}{2}\right)\frac{1}{2}\left(\mathbb{1}_2 - \frac{\vec{r}}{r} \cdot \vec{\hat{\sigma}}\right) \tag{1.17}$$

where $|\lambda_\pm\rangle = \frac{1}{2}(\mathbb{1}_2 \pm \frac{\vec{r}}{r} \cdot \vec{\hat{\sigma}})$ are valid pure states and eigenvectors of $\hat{\varrho}$, while $\frac{1\pm r}{2}$ are the corresponding eigenvalues. Given a density operator as a point in the Bloch sphere, its eigenstates are the projections of its direction on the surface of the sphere, whereas in the case of $r = 0$ all the states are eigenvectors since $\hat{\varrho}$ is a scalar matrix.

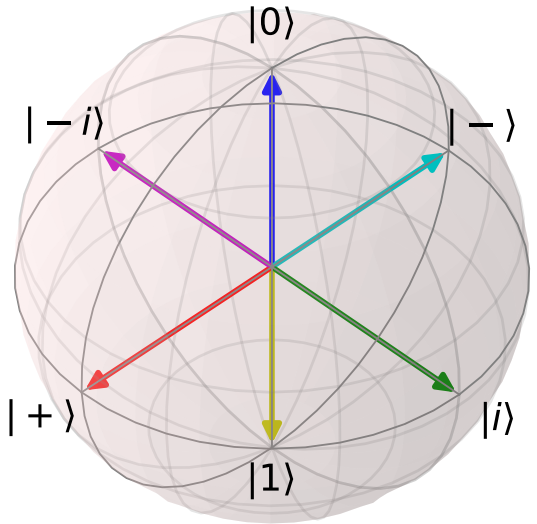


FIGURE 1.1: A visualization of the Bloch Sphere.

1.4 Quantum State Tomography

The central objective of this work is the exploration of new alternatives to quantum state tomography, hence I will now briefly explain the general ideas of its standard procedures [18]. Reconstructing a density matrix means finding its $d^2 - 1$ real parameters with appropriate measurements on the system. It is straightforward to imagine that at least the same amount of independent pieces of data is needed, and in order to compensate for experimental errors, measurements must be repeated many times. Since every external action changes or even destroys the system, tomography needs a large number of copies of it, so as to perform one measurement on each.

Provided one can produce these systems, they then need an experimental setup capable of projecting them on d^2-1 different states and extract the success probabilities of these operations. Moreover, any detector can provide only data that are proportional to such probabilities, therefore one needs also a normalization constant, bringing the total count of measurements to d^2.

There are many different choices of useful projection states, but a good starting point is a basis for the complex vector space of $d \times d$ matrices, which is, not by accident, d^2-dimensional. Let $\{\hat{\sigma}_k\}$ be such a basis, where the symbol $\hat{\sigma}$ is used because in the case of $d = 2$ one example is, again, the set of Pauli matrices plus identity $\{\mathbb{1}_2, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$. Let $\mathrm{Tr}(\hat{\sigma}_j\hat{\sigma}_k) = \alpha\delta_{j,k}$, with α a constant. Then for any $d \times d$ complex matrix $\hat{M}$:

$$\hat{M} = \sum_{k=1}^{d^2} \alpha^{-1}\hat{\sigma}_k \,\mathrm{Tr}(\hat{\sigma}_k\hat{M}) \tag{1.18}$$

After applying the measurement projectors $\{\hat{\Pi}_j \mid j = 1..d^2\}$ to systems all in the state described by density operator $\hat{\varrho}$, the observed data are:

$$n_k = N \,\mathrm{Tr}(\hat{\Pi}_j\hat{\varrho}) \tag{1.19}$$

where N is the aforementioned normalization constant. Substituting from (1.18) into (1.19), one obtains:

$$n_j = N \,\mathrm{Tr}\left(\hat{\Pi}_j \sum_{k=1}^{d^2} \alpha^{-1}\hat{\sigma}_k \,\mathrm{Tr}(\hat{\sigma}_k\hat{\varrho})\right) = N\alpha^{-1} \sum_{k=1}^{d^2} \mathrm{Tr}(\hat{\sigma}_k\hat{\varrho}) \,\mathrm{Tr}(\hat{\Pi}_j\hat{\sigma}_k) \tag{1.20}$$

By defining the $d^2 \times d^2$ matrix $\hat{B}$ such that $\hat{B}_{j,k} = \mathrm{Tr}(\hat{\Pi}_j\hat{\sigma}_k)$ and by inverting it, one gets:

$$\mathrm{Tr}(\hat{\sigma}_k\hat{\varrho}) = N^{-1}\alpha \sum_{j=1}^{d^2} (\hat{B}^{-1})_{k,j} n_j \tag{1.21}$$

From these values it is possible to reconstruct $\hat{\varrho}$ via (1.18). In order to find N^{-1} one simply has to divide the entire $\hat{\varrho}$ by its trace so that it becomes equal to 1. It is clear that projectors $\{\hat{\Pi}_j\}$ have to be linearly independent, so that $\hat{B}$ is invertible: only in this case the measurements can carry through the scheme and are said to be *tomographically complete*.

The example of the qubit can again facilitate comprehension. In 1852, long before the advent of quantum mechanics, George Stokes proposed a way to experimentally define

the polarization state of light [19]. A slightly simplified protocol uses the following set of projectors:

$$\hat{\Pi}_0 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \qquad \hat{\Pi}_1 = |+\rangle\langle +| \qquad \hat{\Pi}_2 = |i\rangle\langle i| \qquad \hat{\Pi}_3 = |0\rangle\langle 0| \tag{1.22}$$

where $|i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ while the other symbols are defined in Section 1.3. Observe that in this case each of the projectors is present in the spectral decomposition of exactly one of the basis matrices $\hat{\sigma}_k$: this is a sufficient but not necessary condition for a set to be tomographically complete. Moreover, with this choice the search for N is straightforward because for any valid density operator $\hat{\varrho}$

$$\mathrm{Tr}(\hat{\Pi}_0\hat{\varrho}) = 1 \qquad \rightarrow \qquad n_0 = N \tag{1.23}$$

In this basis $\alpha = 1$ and

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad \rightarrow \qquad B^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \tag{1.24}$$

so that the desired coefficients of the Bloch vector are simply

$$\mathrm{Tr}(\hat{\sigma}_k\hat{\varrho}) = \frac{n_k - n_0}{n_0} \tag{1.25}$$

In general, B can be much more complicated, especially when the choices of basis and projectors are limited by experimental feasibility. This means that it is usually necessary to perform the entire set of d^2 measurements even if one is interested in a single coefficient. Besides, for very large d, matrix inversion can become computationally prohibitive and the propagation of experimental errors through (1.21) can largely invalidate the results.

It is common, for example, that the reconstructed matrix is not positive semidefinite and $\mathrm{Tr}(\hat{\varrho}^2) > 1$. In this case one must use the experimental data in a further numerical minimization step based on χ^2 or likelihood functions that force the result to satisfy the properties of a density operator. This can be done through the definition of a general valid density matrix $\hat{\rho}$ which is a function of d^2 real parameters $t = t_0..t_{d^2-1}$. These can all be contained in a triangular matrix $\hat{T}(t)$ with real diagonal elements such as (in $d = 2$):

$$\hat{T}(t) = \begin{pmatrix} t_0 & t_1 + it_2 \\ 0 & t_3 \end{pmatrix} \tag{1.26}$$

from which it is possible to obtain any quantum state through expression

$$\hat{\rho}(t) \equiv \frac{\hat{T}(t)\hat{T}^\dagger(t)}{\mathrm{Tr}(\hat{T}(t)\hat{T}^\dagger(t))} \tag{1.27}$$

Then one can calculate the expectations values $\mathrm{Tr}(\hat{\Pi}_j\hat{\rho}(t))$ as functions of the parameters and minimize $\chi^2 \equiv \sum_j \frac{\left(n_j - N\,\mathrm{Tr}(\hat{\Pi}_j\hat{\rho}(t))\right)}{\mathrm{Var}\left(N\,\mathrm{Tr}(\hat{\Pi}_j\hat{\rho}(t))\right)}$ to find t^{opt} that better matches the measurements results and provides the best estimate for $\hat{\varrho} = \hat{\rho}(t^{opt})$.

Chapter 2

Weak Measurements

In a seminal paper published in 1988 [20], authors Y. Aharonov, D. Albert and L. Vaidman (later AAV) introduced a new way of performing a quantum measurement which produced unexpected, strange, yet very useful results. They proposed to weaken the perturbation on the system caused by its observation so as to prevent or at least reduce the collapse of its state. The cost of this operation is a greater imprecision in the results, which has to be compensated for with more repeated acquisitions. Since then, *weak measurements* have found many theoretical and experimental applications, some of which are clever new schemes of state reconstruction and are hence crucial for the central part of this thesis. This chapter is devoted to the detailed explanation of these ideas and to the presentation of the existing *direct state measurement* protocols that I wish to evolve.

2.1 Weak Values and the Ancilla Measurement Scheme

AAV started their dissertation by modeling the observation process as an interaction between the system under study (or *object*) and a measuring device, the *ancilla* (also *pointer*, *meter*). This visualization was introduced first by Von Neumann [2, Ch. 6] and is actually pretty similar to the common intuition of a classical measurement, in which the result is read off a tool that is separated from the variable of interest. In the quantum case, though, the important distinction is that between two Hilbert spaces, $\mathscr{H}_S$ and $\mathscr{H}_A$ respectively for object and ancilla, which in practice can also refer to different properties of the same physical entity.

Suppose that at the beginning the total system lies in the separable and (for simplicity) pure state:

$$|\psi\rangle = |\phi_S\rangle \otimes |I_A\rangle \tag{2.1}$$

and that one wants to measure observable $\hat{S}$ in the object system. In order to do so, they can *couple* it to an operator of the ancilla, so that a change in its state reflects the action of $\hat{S}$ on $|\phi_S\rangle$. Intuitively, the pointer, just like the needle of a scale, reacts to the observation of $\hat{S}$ and carries the result of the measurement. Without loss of generality, I assume $\mathscr{H}_A$ to be continuous, so that an example of coupled operator can be the canonical momentum $\hat{P}_A$ (conjugated to position $\hat{X}_A$). This interaction can be modeled by Hamiltonian

$$\hat{H}_{int}(t) = g(t)\hat{S} \otimes \hat{P}_A \tag{2.2}$$

where $g(t)$ describes its strength and duration. Assuming that the intrinsic Hamiltonians of each Hilbert space vanish, the system evolves according to unitary operator:

$$\hat{U} = e^{-\frac{i}{\hbar}\int \hat{H}_{int}(t)dt} = e^{-i\frac{g}{\hbar}\hat{S}\otimes\hat{P}_A} \tag{2.3}$$

where $g \equiv \int g(t)dt$. Since $\hat{P}_A$ generates translations for states expressed in the x_A basis, the outcome of the measurement of $\hat{S}$ can be read as a shift in the position of the pointer. For instance let the initial state of the ancilla be a real valued function of x_A with zero mean such as the gaussian $I_A(x_A) \equiv \langle x_A|I_A\rangle = (\sqrt{2\pi}\Delta_x)^{-\frac{1}{2}} e^{-\frac{x^2}{4\Delta_x^2}}$ and let $|\phi_S\rangle$ be written in the basis of eigenstates of $\hat{S}$ as $|\phi_S\rangle = \sum_j c_j|s_j\rangle$. Then, in the result of the evolution $|\psi'\rangle$, the pointer position is distributed in a comb of translated gaussians centered around different gs_j: if they do not meaningfully overlap ($\Delta_x \ll g|s_j - s_k|$), a measurement of $\hat{X}_A$ selects one of them and makes the object system collapse onto one particular $|s_j\rangle$. It is also clear that

$$\langle \hat{X}_A^{(g)}\rangle_{\psi'} = g\langle \hat{S}\rangle_{\phi_S} \tag{2.4}$$

in which $\langle \hat{X}_A\rangle_{I_A}$ would appear added to the right-hand side if it was not null as in the present example. In this *strong* case, this model closely represents the axiomatic idea of quantum measurement.

AAV asked themselves what happens when the strength of the interaction g is small compared to the precision of the meter ($\Delta_x \gg g|s_j - s_k|$). In symbols, this means that the evolution operator can be truncated to first order in g:

$$\hat{U} = e^{-i\frac{g}{\hbar}\hat{S}\otimes\hat{P}_A} \approx \mathbb{1} - i\frac{g}{\hbar}\hat{S}\otimes\hat{P}_A \tag{2.5}$$

so that the state of the joint system after its action is

$$|\psi'\rangle \approx |\phi_S\rangle\otimes|I_A\rangle - i\frac{g}{\hbar}\hat{S}|\phi_S\rangle\otimes\hat{P}_A|I_A\rangle \tag{2.6}$$

which is similar to the initial unperturbed configuration. Of course, now a measurement of $\hat{X}_A$ cannot directly provide a value for s_j, because the distribution of results is large even if $|\phi_S\rangle$ is an eigenstate of $\hat{S}$. However, the central moment of this distribution still coincides with the mean value of $g\hat{S}$ as in (2.4) and the precision of its determination can be improved by repeating the procedure many times on identical systems. I point out that the literature occasionally uses the term *weak average* for $\langle\hat{S}\rangle$ when it is measured in this way with small g.

More interesting features appear when another step is added to the recipe: *postselection*. This means performing a projective measurement on the object after the interaction and considering only those cases in which it collapses on one particular (postselected) state. If this chosen state is $|\Phi_S\rangle$, the pointer lies in

$$|F_{\Phi,A}\rangle = \frac{\langle\Phi_S|\psi\rangle}{\langle\Phi_S|\phi_S\rangle} \approx |I_A\rangle - i\frac{g}{\hbar}\langle\hat{S}^W\rangle_\phi^\Phi\hat{P}_A|I_A\rangle \tag{2.7}$$

where

$$\langle\hat{S}^W\rangle_\phi^\Phi \equiv \frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle} \tag{2.8}$$

is defined as the *weak value* of observable $\hat{S}$. The normalization of $|F_{\Phi,A}\rangle$ through division by $\langle\Phi_S|\phi_S\rangle$ also assumes that $\frac{g^2|\langle\hat{S}^W\rangle_\phi^\Phi|^2}{4\Delta_x^2} \ll 1$.

It is conspicuous that $\langle\hat{S}^W\rangle_\phi^\Phi$ is not necessarily an eigenvalue of $\hat{S}$ and is not limited by the bounds of its spectrum. In general it is not even a real number. Nonetheless, it is a measurable quantity and it is easy to show (Appendix A.1) that it can be obtained from the ancilla variables:

$$\begin{aligned} \Re(\langle\hat{S}^W\rangle_\phi^\Phi) &\approx \frac{1}{g}\langle\hat{X}_A\rangle_{F_{\Phi,A}} \\ \Im(\langle\hat{S}^W\rangle_\phi^\Phi) &\approx \frac{2\Delta_x^2}{\hbar g}\langle\hat{P}_A\rangle_{F_{\Phi,A}} \end{aligned} \tag{2.9}$$

Assumptions like the reality and null average of the initial meter wavefunction in both the x_A and p_A representations have been used here and will be used again in the course of this work, yet they are inessential and can be relaxed at the cost of complicating the formalism [21, 22].

For completeness and future reference I add that it is easy to generalize the weak value to the case of a mixed initial state [23]. In the density matrix formalism, if the object system begins in state $\hat{\varrho}_S$, the weak value is:

$$\langle\hat{S}^W\rangle_{\hat{\varrho}}^\Phi \equiv \frac{\langle\Phi_S|\hat{S}\hat{\varrho}_S|\Phi_S\rangle}{\langle\Phi_S|\hat{\varrho}_S|\Phi_S\rangle} \tag{2.10}$$

whereas the weak average, which is measured as the weak value but without postselection, coincides with the mean value of the observable:

$$\langle\hat{S}^W\rangle_{\hat{\varrho}} \equiv \mathrm{Tr}(\hat{S}\hat{\varrho}) \tag{2.11}$$

Even broader generalizations consider non-projective postselection [24], and mixed pointer states [25] but are of minor interest for this thesis.

2.2 The Qubit Pointer

For many applications, the nature of the ancilla is substantially irrelevant. Any property of the physical system under study or of the environment surrounding it can be treated as a pointer, provided that the coupling with the object observable $\hat{S}$ can take place. For this reason, it is often more convenient to use discrete and finite-dimensional ancillae that simplify the notation and are experimentally accessible. In particular, in the next sections and in the experimental part of this thesis, qubit pointers will be used extensively, therefore it is important to render the previous formalism into this new setting.

The natural meter operators in place of $\hat{P}_A$ and $\hat{X}_A$ are the Pauli matrices $\hat{\sigma}_{xA}$, $\hat{\sigma}_{yA}$, $\hat{\sigma}_{zA}$. Any of them can be used in the coupling but I adopt $\hat{\sigma}_{yA}$ for consistency with most of the literature; then complementary observable $\hat{\sigma}_{xA}$ can take the role of $\hat{X}_A$. The initial state of the pointer is chosen so that it has zero mean and is real in both the $\hat{\sigma}_{xA}$ and $\hat{\sigma}_{yA}$ representations, therefore the eigenstate $|0_A\rangle$ of $\hat{\sigma}_{zA}$ is a valid candidate.

The evolution operator describing the coupling is:

$$\hat{U} = e^{-i\theta \hat{S} \otimes \hat{\sigma}_{yA}} \tag{2.12}$$

which, similarly to the translation of formula (2.3), implements a rotation of the pointer state around the y axis of the Bloch sphere by an angle that is the result of the measurement of $\hat{S}$ in the object system, multiplied by twice the strength coefficient θ. Repeating the same steps of the previous section, the weak value can be measured from the ancilla variables as:

$$\begin{aligned} \Re(\langle \hat{S}^W \rangle_\phi^\Phi) &\approx \frac{1}{2\theta} \langle \hat{\sigma}_{xA} \rangle_{F_{\Phi,A}} \\ \Im(\langle \hat{S}^W \rangle_\phi^\Phi) &\approx \frac{1}{2\theta} \langle \hat{\sigma}_{yA} \rangle_{F_{\Phi,A}} \end{aligned} \tag{2.13}$$

which are valid supposing that $\theta^2 |\langle \hat{S}^W \rangle_\phi^\Phi|^2 \ll 1$.

Summarizing, the reinterpretation of the formalism goes as follows:

$$\begin{aligned} \hat{P}_A &\to \hat{\sigma}_{yA} \\ \hat{X}_A &\to \hat{\sigma}_{xA} \\ g &\to \theta \\ |I_A\rangle &\to |0_A\rangle \\ e^{-i\frac{g}{\hbar}\hat{S}\otimes\hat{P}_A} &\to e^{-i\theta\hat{S}\otimes\hat{\sigma}_{yA}} \end{aligned} \tag{2.14}$$

The point where this analogy slightly falters is equation (2.4) which becomes

$$\langle \hat{\sigma}_{xA}^{(\theta)} \rangle_{\psi'} = \left\langle \sin\left(2\theta \hat{S}\right) \right\rangle_{\phi_S} \tag{2.15}$$

and loses much of its significance outside of the weak regime. The appearance of rotations and periodic functions such as $\sin(\cdot)$ is an indicator of the fundamental limitation of qubit ancillae, which lies in the compactness of the Bloch sphere. In the previous formalism of the continuous pointer, if the initial and translated gaussians are not separated enough to be unambiguously distinguishable, one can imagine to increase g to make them farther apart. In the qubit case this separation is achieved when the initial and rotated state are orthogonal, but if the measuring device is so imprecise that it still cannot demarcate them (e.g. a polarizer that transmits some light although it being orthogonally polarized to its axis), the only solution is enhancing the statistics. In other words, it is useless to amplify a measurement using an arbitrarily large value of θ, as the pointer state would simply rotate and come back to where it was before. However, in what follows it is assumed that the ancilla is not affected by this problem or the statistics is rich enough to neglect the imprecision in the discrimination of orthogonal states. Then, as described in Chapter 3, strong measurements with qubit pointers regain their usefulness. In particular, if $\hat{S} = \hat{\Pi}_S$ is a projector, the choice of $\theta = \frac{\pi}{2}$ rotates $|0_A\rangle$ to its orthogonal $|1_A\rangle$ when the projection yields a positive result. The mean value of $\hat{\Pi}_{1A}$ carries the information of $\langle \hat{\Pi}_S \rangle_{\phi_S}$ so that in place of (2.15) one can write:

$$\langle \hat{\Pi}_{1A}^{(\theta)} \rangle_{\psi'} = \sin^2(\theta) \langle \hat{\Pi}_S \rangle_{\phi_S} \tag{2.16}$$

2.3 Applications of the Weak Value

The possibility of observing a quantum system without perturbing it has elicited extensive research around weak values in the last 30 years. Some of it has brought clarity onto many fundamental questions of the theory, while some has made the weak measurement an important simplifying tool in experiments. This section has the purpose of listing some of these applications before moving on to those that are essential for this work.

2.3.1 Amplification through Postselection

The weak value was appreciated first as a source of amplification of feeble signals. This is mainly due to the fact that if the initial and postselected states almost do not overlap, the real and imaginary parts of $\langle \hat{S}^W \rangle_\phi^\Phi$ can reach great numerical values, making the pointer measurements achievable with equipment of limited sensitivity. From them, it is usually possible to extract interesting physical parameters such as the small coupling constant g. To clarify, the point is not to use an interaction so strong that even a small value of $\langle \hat{S} \rangle$ can be observed, rather the opposite: the coupling is extremely weak, but thanks to a clever postselection, the product $g|\langle \hat{S}^W \rangle_\phi^\Phi|$ becomes measurable. However, this amplification is not arbitrary, as the relations such as (2.9) or (2.13) that govern it are valid only if $\frac{g^2 |\langle \hat{S}^W \rangle_\phi^\Phi|^2}{4\Delta_x^2} \ll 1$, or $\theta^2 |\langle \hat{S}^W \rangle_\phi^\Phi|^2 \ll 1$ for qubit ancillae: the initial spread of the meter wavefunction is the fundamental limit of the process.

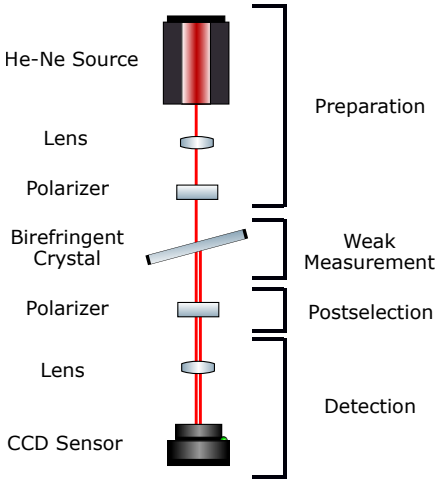


FIGURE 2.1: Scheme of the first experiment that exploited the amplifying effects of postselection [26].

This has not stopped experimentalists from achieving remarkable results, the first of which came in 1990 [26] with an optical setup [27] inspired by the particle analog originally proposed by AAV [20]. Laser light passes through a thin birefringent quartz plate which separates two orthogonal polarizations by a distance a that is much smaller than the beam waist (hence the weakness of the interaction). A subsequent polarizer postselects only a fraction of the emitted light and sends it to a CCD. In this case the space of polarizations is the object system and the role of $\hat{S}$ is played by the projector on the ket corresponding to the displaced beam. The measured meter observable is the position of the light on the sensor, represented by an intensity distribution over its pixels. The mean value of this distribution is proportional to small parameter a, but is greatly amplified by the weak value if the postselected state is chosen to be nearly orthogonal to the initial one. Therefore it can be observed and the value of a can be accurately ascertained even if the width of the pixels is large. In the cited experimental case, the precision of the measurement reached the tens of nms despite the use of 1.4 μm wide pixels.

In 2008 an even more extreme amplification was used to observe the *spin Hall effect of light*, a splitting of a beam into two polarized components due to a gradient in refractive index. Using clever choices of initial and final polarization states, and exploiting also a free evolution of the system before the measurement, experimenters could achieve an amplification factor of almost 10^4 and detected displacements close to the angstrom [28]. Other noteworthy examples reported observations of beam angular deflections in the hundreds of femtoradians [29] and studies of the Imbert-Fedorov effect with amplifications of at least two orders of magnitude [30]. For a more comprehensive list see Reference [31].

Of course these results come at a price. Since their origin is essentially the small overlap $\langle\Phi_S|\phi_S\rangle$, only a few events pass the postselection, forcing the acquisition of more data to achieve meaningful statistics. In the above example, this means that although the position distribution shows an appreciable displacement, its height and integral are lower unless more measurements are added to compensate.

2.3.2 The Meaning of the Weak Value and the Quantum Three-Box Paradox

Whether $\langle\hat{S}^W\rangle_\phi^\Phi$ is somehow related to the (mean) *value* of observable $\hat{S}$ has been the subject of a lengthy discussion that has shaken the bedrock of the concepts of measurement and state [32, 33]. Some light on this question can be shed by the study of quantum paradoxes such as Hardy's [34] or the three-box paradox [35]. Focusing for simplicity on the latter, suppose that a particle can be found in one of three boxes A, B or C and its state is initially described by $\phi = \frac{1}{\sqrt{3}}(|A\rangle + |B\rangle + |C\rangle)$. One of the three projectors $\hat{\Pi}_A$, $\hat{\Pi}_B$, $\hat{\Pi}_C$ is measured and then $\Phi = \frac{1}{\sqrt{3}}(|A\rangle + |B\rangle - |C\rangle)$ is postselected. Using Bayes' Theorem it is possible to calculate the probabilities of finding the particle in each box with any of the three intermediate measurements. For instance if the chosen observable is $\hat{\Pi}_A$, the probability of finding box A populated is

$$\Pr_A(\text{in A}) = \frac{|\langle\Phi|\hat{\Pi}_A|\phi\rangle|^2}{|\langle\Phi|\hat{\Pi}_A|\phi\rangle|^2 + |\langle\Phi|\mathbb{1} - \hat{\Pi}_A|\phi\rangle|^2} = 1 \tag{2.17}$$

Similarly, one can obtain

$$\Pr_B(\text{in B}) = 1 \qquad \Pr_C(\text{in C}) = \frac{1}{5} \tag{2.18}$$

which seem absurd because they affirm that both A and B are certainly occupied. However, one has to realize that these values are refer to different projective measurements which cannot be performed at the same time.

The problem can be circumvented by the use of AAV's recipe. For instance, imagining that the particle is charged, one can evaluate its presence in any box by sending a probe charge near it, effectively coupling the corresponding projector with the position of the probe [36]. A measurement of the pointer momentum followed by postselection of Φ will yield the weak values

$$\langle\hat{\Pi}_A^W\rangle_\phi^\Phi = 1 \qquad \langle\hat{\Pi}_B^W\rangle_\phi^\Phi = 1 \qquad \langle\hat{\Pi}_C^W\rangle_\phi^\Phi = -1 \tag{2.19}$$

The fact that $\langle \hat{\Pi}_C^W \rangle_\phi^\Phi < 0$ shows that it is incorrect to interpret these values as indications of the location of the particle. While a projective measurement of $\hat{\Pi}_C$ can only return 0 or 1, $\langle \hat{\Pi}_C^W \rangle_\phi^\Phi$ can take almost any value, for example if the postselected state is a more general $\Phi = \sin(\theta)\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle) + \cos(\theta)|C\rangle$ one finds:

$$\langle \hat{\Pi}_C^W \rangle_\phi^\Phi = \frac{1}{\sqrt{2}\tan(\theta) + 1} \tag{2.20}$$

which cannot be linked to the number of particles in box C. Then, the solution to the paradox is that although it is true (and experimentally verified [37]) that the weak values are those of (2.19), they do not represent actual probabilities of finding a box occupied.

2.3.3 Average Trajectories of Photons in a Double Slit Experiment

During the first lectures of any basic quantum physics course, students learn about Heisenberg's uncertainty principle and its consequences in the double slit experiment. Since $||[\hat{X}, \hat{P}]|| > 0$, it is not possible to observe an interference pattern (connected to momentum) and simultaneously know the trajectories of the photons (position). If one manages to ascertain through which slit a photon has passed, interference and any momentum information are lost.

But what if this measurement was carried out weakly? In 2011, this idea was put to the test [38]. Single photons are emitted from a quantum dot device and separated by an in-fiber 50:50 beam splitter, analog to a slit screen. Their polarization (a qubit pointer) is prepared in an initial linear ket and then weakly rotated towards an elliptical state by a calcite prism depending on the momenta of the photons (coupling). A *polarizing beam displacer (PBD)* separates the two circular polarizations along the y direction and sends the results to a cooled CCD. By observing one y-parallel column of pixels at a time (postselection on x) and comparing the intensities of the two ports of the displacer (pointer measurement on $\hat{\sigma}_z$), it is possible to obtain information on such momenta. Then this analysis can be repeated at different positions on the propagation direction z, thus allowing to map complete spatial trajectories that are correlated to the interference pattern on the sensor.

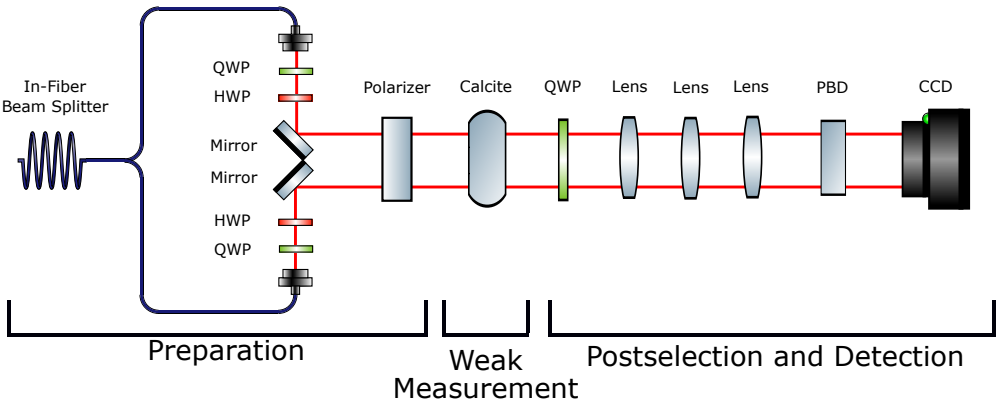


FIGURE 2.2: Scheme of the experimental setup [38].

It is necessary to underline that this result does not violate Heisenberg's principle, as these are only average trajectories and not relative to each photon. However, it shows that weak measurements can provide new insight into a fundamental experiment of quantum mechanics such as the double slit interferometer.

2.3.4 Monogamy of Bell's Inequalities Violations

Interactions of strength that is neither maximal nor close to 0 have been recently useful in the study of entangled systems [39]. Entanglement is possibly "the characteristic trait" of quantum mechanics [40] and its existence has been subject of discussion since the universally renowned paper by Einstein, Podolsky and Rosen *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* [41]. The fact that, in some cases, action on a subsystem can instantaneously trigger change on another was termed "spooky action at a distance". In 1964, J.S. Bell devised his famous inequality which can rule out semiclassical explanations such as local hidden variable models (LHVMs) [42]. If Alice and Bob, the two parts of a bipartite system, can measure two observables each (A, A', B, B'), then the CHSH inequality, a variation of Bell's, reads:

$$I_{CHSH} = \langle A \otimes B \rangle + \langle A \otimes B' \rangle + \langle A' \otimes B \rangle - \langle A' \otimes B' \rangle \leq 2 \tag{2.21}$$

A violation of such bound, which has been observed experimentally [43], excludes LHVMs but can be compatible with the predictions of quantum physics: for maximally entangled states and appropriate choices of observables, $I_{CHSH} = 2\sqrt{2}$.

Nonetheless, it is generally accepted that if there are three *non-signaling* subsystems, it is not possible to simultaneously violate the CHSH inequality in two different pairs of observers Alice-Bob1 and Alice-Bob2 (monogamy) [39, 44]. This is no longer true if the Bobs can communicate with each other, in particular Bob1 must measure his substate before sending it to Bob2 and has to do it weakly, so as not to perturb it much. Supposing that a projector is coupled to a qubit meter, for which the strength of the interaction between object and ancilla can be described by an angle $\theta \in]0, \frac{\pi}{2}]$, it can be shown that with the best choices of initial state and observables

$$I^{(1)}_{CHSH} = 2\sqrt{2}\sin^2(\theta) \qquad I^{(2)}_{CHSH} = \sqrt{2}(1 + \cos(\theta)) \tag{2.22}$$

There is a narrow range around $\theta = \frac{\pi}{3}$ (neither strong, $\theta = \frac{\pi}{2}$, nor completely weak, $\theta \to 0$) for which both inequalities are violated at the same time. Weak measurements allow entanglement to be bigamous.

2.4 Direct Weak Measurement of a Pure Quantum State

The most interesting application of weak values for this thesis is of course connected to the reconstruction of the state of a quantum system. Since this is described by complex numbers (the coordinates of a unit vector or the elements of a density matrix), it is intuitive that a complex value can be of aid. J. S. Lundeen and his group in Ottawa were the first to notice the connection that I will now report [3].

Suppose that one can prepare a quantum system in a pure state described by unknown unit vector $|\phi_S\rangle$ and they want to represent it in the basis $\{|a_j\rangle \mid j = 1..d\}$ of the

d-dimensional Hilbert space $\mathscr{H}_S$. The key idea is to weakly measure projector $\hat{\Pi}_{a_j} \equiv |a_j\rangle\langle a_j|$ on $|\phi_S\rangle$ and then postselect on $|b_0\rangle \equiv \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |a_j\rangle$ in order to obtain

$$\langle \hat{\Pi}_{a_j}^W \rangle_{\phi}^{b_0} \equiv \frac{\langle b_0|a_j\rangle\langle a_j|\phi_S\rangle}{\langle b_0|\phi_S\rangle} = \nu \langle a_j|\phi_S\rangle \tag{2.23}$$

Provided one can measure $\langle \hat{\Pi}_{a_j}^W \rangle_{\phi}^{b_0}$ as explained in Section 2.1, they can find the coordinates $\langle a_j|\phi_S\rangle$ by scanning over j and keeping $|b_0\rangle$ fixed. The choice of this postselection state is made so that the factor ν does not depend on j and can be eliminated by normalization of the vector after all its components have been determined.

As if the simplicity of this conclusion was not already stunning, I shall illustrate the example given by Lundeen *et al.* and that is often quoted in the literature. Suppose to be interested in the transverse spatial wavefunction of a photon $\phi_S(x)$. In order to perform a weak measurement, it is possible to couple $\hat{\Pi}_x \equiv |x\rangle\langle x|$ to another variable that acts as a pointer, for instance the photon polarization. Many copies of the system are prepared in the initial state

$$|\psi\rangle = |\phi_S\rangle \otimes |0_A\rangle \tag{2.24}$$

where $|0_A\rangle$ labels (for example) the horizontal polarization in the ancilla Hilbert space $\mathscr{H}_A$ and is an eigenstate of $\hat{\sigma}_z$. For each position x_j, the evolution operator that describes the interaction between object and meter is

$$\hat{U}(x_j) = e^{-i\theta \hat{\Pi}_{x_j} \otimes \hat{\sigma}_y} \approx \mathbb{1}_S \otimes \mathbb{1}_A - i\theta \hat{\Pi}_{x_j} \otimes \hat{\sigma}_y \tag{2.25}$$

in which θ is a small angle. This means that if the measurement of $\hat{\Pi}_{x_j}$ yields a positive result, the polarization is slightly rotated, otherwise it is left as it is.

The connection between these two degrees of freedom can be experimentally achieved using a small birefringent retarder, a *half-wave plate (HWP)*, that is only hit by those photons that are found in the appropriate position x_j, and rotates their linear polarization. Of course, a discretization of the theoretically infinite dimensional and continuous Hilbert space of transverse modes $\mathscr{H}_S$ is here carried out by the finite width of the plate, which interacts with a range of positions that is necessarily larger than point-like. Moreover, the number of sampled values of x is obviously a finite d, so that the final result is a d-long series of points $\phi_S(x_j)$ which is interpolated by the actual wavefunction. Recall that the objective is a numerical expression of $|\phi_S\rangle$ in the basis of positions and not a symbolic formula.

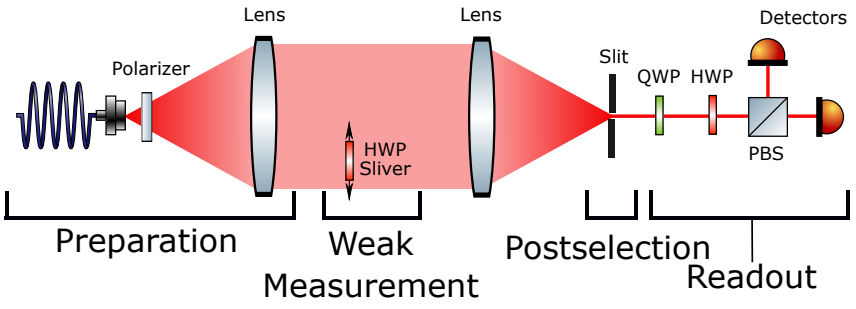


FIGURE 2.3: Scheme of Lundeen's experiment [3].

Postselection on complementary variable p_x, the transverse momentum, can be accomplished via a Fourier lens and a slit that blocks all photons except for those with $p_x = 0$. This means that the postselected state is indeed

$$|p_0\rangle \equiv \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |x_j\rangle \tag{2.26}$$

Finally, the pointer state is described by normalized vector

$$|F_{p_0,A}\rangle = \frac{\langle p_0|\hat{U}(x_j)|\psi\rangle}{\langle p_0|\phi_S\rangle} \approx |0_A\rangle + \theta \frac{\phi_S(x_j)}{\tilde{\phi}} |1_A\rangle \tag{2.27}$$

where the weak value appears as $\frac{\phi_S(x_j)}{\tilde{\phi}}$ and $\tilde{\phi} \equiv \sum_{j=1}^{d} \phi_S(x_j)$. By choosing the overall phase of ϕ_S so that $\tilde{\phi}$ is real and positive, one easily recognizes that

$$\begin{aligned} \Re(\phi_S(x_j)) &\approx \frac{\tilde{\phi}}{2\theta} \langle \hat{\sigma}_x^{(j)} \rangle_{F_{p_0,A}} \\ \Im(\phi_S(x_j)) &\approx \frac{\tilde{\phi}}{2\theta} \langle \hat{\sigma}_y^{(j)} \rangle_{F_{p_0,A}} \end{aligned} \tag{2.28}$$

After scanning the values of x_j, the factor $\frac{\tilde{\phi}}{2\theta}$ can be eliminated by normalization. In some experimental realizations, it may be possible to know this proportionality constant a priori or with a limited number of additional measurements, so that one can quickly find a single $\phi_S(x_j)$ if they are not interested in the entire vector (Section 2.6). This is what makes this technique *direct* and less prone to inaccuracies due to the propagation of experimental errors in the convoluted reconstruction of standard QST.

This method provides an universal recipe that can be applied to any kind of system provided one can implement the necessary couplings. Apart from the normalization problem, one can determine each coordinate of the state vector with a fixed number of measurements that are always the same regardless of the nature of the system or its dimension.

The very particular eventuality of $\tilde{\phi} = 0$ can be treated separately using a different $|b\rangle$ in place of $|b_0\rangle$. In this case factor $\langle b|a_j\rangle$ can vary with j but is known and can be eliminated manually from every measurement before normalization.

2.5 Direct Weak Measurement of a General Quantum State

The measurement scheme I have just presented is based on the formal connection (2.23) between the weak value and the unit vector representation, and is therefore applicable only to pure states. Nonetheless, there exist other protocols, mostly introduced by the same authors [4], that take care of the mixed case. Here, I am going to present two of them: the first provides a way to reconstruct the density operator element by element, whereas the second, aiming at the discrete Dirac distribution, renounces directness but gains in practical simplicity.

2.5.1 Direct Weak Reconstruction of the Density Operator

Suppose to be interested in the matrix elements of unknown density operator $\hat{\varrho}$ in basis $\{|a_j\rangle \mid j = 1..d\}$ and consider $\hat{\Pi}_{a_j a_k} \equiv \hat{\Pi}_{a_k}\hat{\Pi}_{b_0}\hat{\Pi}_{a_j}$, where again $|b_0\rangle \equiv \frac{1}{\sqrt{d}}\sum_{j=1}^{d}|a_j\rangle$. It is clear that the desired coefficients are expressed by

$$\langle a_j|\hat{\varrho}|a_k\rangle = d \cdot \mathrm{Tr}(\hat{\Pi}_{a_j a_k}\hat{\varrho}) \tag{2.29}$$

However, $\hat{\Pi}_{a_j a_k}$ contains the products of non-commuting projectors and is hence not hermitian nor an observable. This means that the right-hand side of (2.29) cannot be the mean value of a projective measurement as in (1.3), nor it can be obtained in the ancilla scheme, since regardless of the coupled meter operator, the evolution is not unitary and non-physical.

The solution is offered by coupling each of the three projectors separately to three different ancillae and performing three *sequential* weak measurements. In the example of qubit pointers $\mathscr{H}_A$, $\mathscr{H}_B$, $\mathscr{H}_C$ and coupled operators $\hat{\sigma}_{yA}$, $\hat{\sigma}_{yB}$, $\hat{\sigma}_{yC}$, the evolution is governed by:

$$\hat{U} = e^{-i\theta_C\hat{\Pi}_{a_k}\otimes\hat{\sigma}_{yC}}e^{-i\theta_B\hat{\Pi}_{b_0}\otimes\hat{\sigma}_{yB}}e^{-i\theta_A\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}} \tag{2.30}$$

Reference [45] introduces a clever way to calculate weak averages (or weak values if post-selection occurs) of products, which I extented to the case of qubit pointers (Appendix A.2). It is a general result that

$$\langle\hat{\Pi}^W_{a_j a_k}\rangle_{\hat{\varrho}} = \mathrm{Tr}(\hat{\Pi}_{a_j a_k}\hat{\varrho}) \approx \frac{1}{\theta_A\theta_B\theta_C}\langle\hat{\Sigma}_A\hat{\Sigma}_B\hat{\Sigma}_C\rangle_{F_{ABC}} \tag{2.31}$$

where $\hat{\Sigma}_l \equiv \frac{\hat{\sigma}_{xl}+i\hat{\sigma}_{yl}}{2}$. In practice, because $\hat{\Sigma}_l$ is not hermitian either, the products of Pauli operators that appear in $\langle\hat{\Sigma}_A\hat{\Sigma}_B\hat{\Sigma}_C\rangle_{F_{ABC}}$ have to be measured on different subensembles of identically prepared systems. Since these operators act on distinct Hilbert spaces, those of the three ancillae, they all commute with one another and their products are observables.

For this protocol to work properly, it is fundamental, according to the authors, that the couplings are indeed weak, so that a measurement does not affect the next one. However, as Lundeen himself explains, there is no reason for the last interaction not to be strong. In other words, one can consider the measurement of $\hat{\Pi}_{a_k}$ as a postselection on $|a_k\rangle$ and find the matrix element of interest through the weak value

$$\varrho_{jk} \equiv \langle a_j|\hat{\varrho}|a_k\rangle = d\,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})\langle(\hat{\Pi}_{b_0}\hat{\Pi}_{a_j})^W\rangle^{a_k}_{\hat{\varrho}} \tag{2.32}$$

This fact makes any practical realization much simpler by reducing to 2 the number of needed ancillae and has been used in the first verification of this protocol [46], and in the experimental part of this work. From (2.31) follows that

$$\begin{aligned}
\Re(\varrho_{jk}) &\approx \frac{d}{4\theta_A\theta_B}\,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})\left(\langle\hat{\sigma}^{(a_j)}_{xA}\hat{\sigma}^{(b_0)}_{xB}\rangle_{F_{a_k},AB} - \langle\hat{\sigma}^{(a_j)}_{yA}\hat{\sigma}^{(b_0)}_{yB}\rangle_{F_{a_k},AB}\right)\\
\Im(\varrho_{jk}) &\approx \frac{d}{4\theta_A\theta_B}\,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})\left(\langle\hat{\sigma}^{(a_j)}_{yA}\hat{\sigma}^{(b_0)}_{xB}\rangle_{F_{a_k},AB} + \langle\hat{\sigma}^{(a_j)}_{xA}\hat{\sigma}^{(b_0)}_{yB}\rangle_{F_{a_k},AB}\right)
\end{aligned} \tag{2.33}$$

where the mean values of the Pauli operators are measured on the final pointer state after postselection. Provided some information on the normalization of the results is available

(Section 2.6), this scheme allows to find every density matrix element independently and can be used with any kind of system of any dimension.

2.5.2 Direct Weak Reconstruction of the Dirac Distribution

Another measurement protocol was also proposed by Lundeen [47] and put to the test [48]. Due to its comparatively easier realization, its verification and extension to the case of strong couplings is a secondary interest of this thesis.

Let $\{|a_j\rangle \mid j = 0..d-1\}$ be the basis in which one wants to express $\hat{\varrho}$ and let $\{|b_l\rangle \mid l = 0..d-1\}$ be its Fourier basis. Consider the non-hermitian operator $\hat{D}_{jl} \equiv \hat{\Pi}_{b_l}\hat{\Pi}_{a_j}$, then

$$\mathrm{Tr}(\hat{D}_{jl}\hat{\varrho}) = \langle a_j|\hat{\varrho}|b_l\rangle\langle b_l|a_j\rangle \tag{2.34}$$

is the discrete Fourier transform of the density operator:

$$\langle a_j|\hat{\varrho}|a_k\rangle = \sum_{l=0}^{d-1} \mathrm{Tr}(\hat{D}_{jl}\hat{\varrho})e^{i\frac{2\pi l(j-k)}{d}} \tag{2.35}$$

and coincides with the jl element of the corresponding discrete Dirac distribution. Again, one can measure it by coupling sequentially $\hat{\Pi}_{a_j}$ and $\hat{\Pi}_{b_l}$ to two distinct ancillae and finding the weak average

$$\langle \hat{D}^W_{jl} \rangle_{\hat{\varrho}} = \mathrm{Tr}(\hat{D}_{jl}\hat{\varrho}) \approx \frac{1}{\theta_A\theta_B}\langle \Sigma_A\Sigma_B \rangle_{F_{AB}} \tag{2.36}$$

Since the last measurement does not have to be weak, it can be treated as a postselection:

$$\langle \hat{D}^W_{jl} \rangle_{\hat{\varrho}} = \mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho})\langle \hat{\Pi}^W_{a_j} \rangle^{b_l}_{\hat{\varrho}} \approx \frac{1}{2\theta_A}\mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho})\left(\langle \hat{\sigma}^{(a_j)}_{xA} \rangle_{F_{b_l,A}} + i\langle \hat{\sigma}^{(a_j)}_{yA} \rangle_{F_{b_l,A}}\right) \tag{2.37}$$

so that only one ancilla is needed.

This technique allows to extract the discrete Dirac distribution element by element, but cannot be considered direct as the previous one if the goal is the density operator. Indeed, equation (2.35) shows that at least $O(d)$ measurements are needed for each matrix element, although with the same data one can determine an entire row.

2.6 Normalization of the Results

In the previous section I have shown that it is possible to extract the properties of the state under study (e.g. its density matrix elements) directly from measurements on the pointers and known constants such as the dimension of the system and the strength of the coupling. Nonetheless, there seems to be a problem represented by the postselection probability (later p_{ps}), that appears as $\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})$ in (2.33) and as $\mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho})$ in (2.37). The solution is connected to the more fundamental procedure of normalization of experimental results.

In a general quantum physics experiment, especially in the case of photonics as in Chapter 4, the datum usually consists of counts that are proportional to the mean value of a

projector. Let $\hat{O}$ be an observable and let $\hat{O} = \sum_l \lambda_l \hat{\Pi}_l$ be its spectral decomposition: the system is projected sequentially using the various $\hat{\Pi}_l$ and counts

$$N_l = N\langle\hat{\Pi}_l\rangle \tag{2.38}$$

are recorded, where N is a constant factor representing the total power of the signal. It is clear that

$$\langle\hat{O}\rangle = \sum_l \lambda_l \langle\hat{\Pi}_l\rangle = \sum_l \lambda_l \frac{N_l}{N} \tag{2.39}$$

in which $\frac{N_l}{N}$ takes the role of the probability (in frequentist terms) of projection.

In the specific case of the protocols of Section 2.5, the probabilities that appear in the mean values of pointer observables are *conditioned* on successful postselection, that is

$$\langle\hat{O}\rangle_{F_{ps}} = \sum_l \lambda_l p_{l|ps} = \sum_l \lambda_l \frac{p_{l\cap ps}}{p_{ps}} \tag{2.40}$$

where $p_{l\cap ps}$ is the *joint* probability of projection on the l-th outcome of the ancilla measure and postselection on a specific state of the object. Since non-selected counts are discarded, $N_l = N p_{l\cap ps}$, so that

$$\sum_l \lambda_l \frac{N_l}{N} = \langle\hat{O}\rangle_{F_{ps}} \cdot p_{ps} \tag{2.41}$$

which is conveniently the term that appears in (2.33) and (2.37), and makes a separate determination of p_{ps} unnecessary.

Of course this is only a way of concealing the weak average notation of (2.31) and (2.36), but it is preferable to always consider the last coupling as a postselection because it makes the experimental realization and the calculations much simpler. To further highlight this and for ease of writing, in future sections I will often use the symbol

$$\langle\hat{O}\rangle^{ps} \equiv \langle\hat{\Pi}_{ps} \otimes \hat{O}\rangle_{\psi'} = \langle\hat{O}\rangle_{F_{ps}} \cdot p_{ps} \tag{2.42}$$

where $|\psi'\rangle$ is the joint state just before postselection and $\hat{\Pi}_{ps}$ is the postselecting projector. One can imagine that the mean value on the left-hand side is calculated on the *unnormalized* state after postselection.

However, this has merely shifted the problem from p_{ps} to N which is still unknown. The most trivial but lengthy way to find it is to renormalize the density matrix after it has been determined in its entirety. If this is not possible, then one can focus on the diagonal elements, reducing the number of needed measurements from d^2 to d. Finally, some experimental setups may provide an estimate of the signal power, and thus of N, with a single measure analog to that of $\hat{\Pi}_0$ in (1.22).

Unfortunately, this does not apply to the formulas of Section 2.4, since even with the knowledge of N, factor $\tilde{\phi}$ remains out of reach. In this case one either has to measure the complete state vector or must have some a priori information on one of its elements to calibrate the results.

Chapter 3

Direct State Reconstruction via Strong Measurements

The fundamental idea on which this thesis stands is that the state reconstruction protocols described in the previous chapter do not need weak couplings and can work in (almost) the same way with strong ones [5, 49]. The ancilla scheme is still necessary because it separates the interesting variables from those that are actually observed and provides an useful formalism with which one can find the needed measurements. A strong link between object system and pointer device is usually much easier to practically implement and allows to extract more information per measure, making the results precise without the need for a large number of acquisitions. Moreover, the final values are bound to the measured quantities without any approximation (such as that of (2.25)) and are hence less prone to inherent biases. My goal is to show that the exact version of Lundeen's formulas can be used for state reconstruction without critical formal or experimental complications, and the greater disturbance on the object system caused by strong measurements does not constitute a problem for these applications.

3.1 Evolution of the Basic Protocol to Couplings of Arbitrary Strength

Despite the fact that the main focus of this work is on evolving the schemes reported in Section 2.5, it is best to start from the basics and consider the extension of the most fundamental of Lundeen's protocols (Section 2.4) to the case of strong measurements. Suppose, again, that one wants to reconstruct the transverse spatial wavefunction of a photon and does so by coupling projector $\hat{\Pi}_{x_j}$ to the Pauli operator $\hat{\sigma}_y$ in the Hilbert space of polarizations that acts as a pointer. The evolution is the same as that of (2.25) but this time θ can take any value in $]0,\ \frac{\pi}{2}]$.

$$\hat{U}(x_j) = e^{-i\theta\hat{\Pi}_{x_j}\otimes\hat{\sigma}_y} = (\mathbb{1}_S - \hat{\Pi}_{x_j}) \otimes \mathbb{1}_A + \hat{\Pi}_{x_j} \otimes e^{-i\theta\hat{\sigma}_y} \tag{3.1}$$

which is now exact. The subsequent steps in the procedure are the same: $|p_0\rangle$ is postselected in the object system and the final normalized pointer state becomes

$$|F_{p_0,A}\rangle = \frac{\langle p_0|\hat{U}(x_j)|\psi\rangle}{\langle p_0|\phi_S\rangle} = \left(1 + (\cos(\theta) - 1)\frac{\phi_S(x_j)}{\tilde{\phi}}\right)|0_A\rangle + \sin(\theta)\frac{\phi_S(x_j)}{\tilde{\phi}}|1_A\rangle \tag{3.2}$$

Although the coupling is strong, the weak value $\frac{\phi_S(x_j)}{\tilde{\phi}}$ formally appears again here. Finally, projective measurements are carried out on the pointer to find that

$$\begin{aligned}\Re(\phi_S(x_j)) &= \frac{\tilde{\phi}}{2\sin(\theta)}\left(\langle\hat{\sigma}_x^{(a_j)}\rangle_{F_{p_0,A}} + 2\tan\left(\frac{\theta}{2}\right)\langle\hat{\Pi}_1^{(a_j)}\rangle_{F_{p_0,A}}\right)\\ \Im(\phi_S(x_j)) &= \frac{\tilde{\phi}}{2\sin(\theta)}\langle\hat{\sigma}_y^{(a_j)}\rangle_{F_{p_0,A}}\end{aligned} \tag{3.3}$$

The resemblance to (2.28) is conspicuous. However, one additional pointer measurement is needed, that of $\hat{\Pi}_1$, where $|1_A\rangle$ is defined as the orthogonal state of the initial $|0_A\rangle$. The entire cost of strong couplings is in this single further step. Again, if one knows the normalization factor $\frac{\tilde{\phi}}{2\sin(\theta)}$, they can extract each coordinate of the state vector in a fixed number of measurements regardless of the dimension of the object Hilbert Space.

3.2 Evolution of the Advanced Protocols

If the most essential model of direct state measurement can be generalized to strong couplings, it seems natural that the other protocols, those that take care of the mixed case, should too. In general, this extension is just a matter of applying result (3.1) to the joint evolution operators (such as that of equation (2.30)) and then finding the most useful measurements on the pointers.

3.2.1 The Density Operator Protocol

Let the initial object system be prepared in state $\hat{\varrho}_S$ of unknown matrix elements $\varrho_{jk} = \langle a_j|\hat{\varrho}_S|a_k\rangle$, with $\{|a_j\rangle \mid j = 1..d\}$ being an orthornormal basis for the Hilbert space $\mathscr{H}_S$. Two ancillae $\mathscr{H}_A$ and $\mathscr{H}_B$ are prepared in pure states $\hat{\varrho}_A = |0_A\rangle\langle 0_A|$ and $\hat{\varrho}_B = |0_B\rangle\langle 0_B|$, so that in the general $\mathscr{H}_{tot} = \mathscr{H}_S \otimes \mathscr{H}_A \otimes \mathscr{H}_B$, the configuration is described by the separable tensor product

$$\hat{\varrho}_{tot} = \hat{\varrho}_S \otimes \hat{\varrho}_A \otimes \hat{\varrho}_B \tag{3.4}$$

After the choice of a particular $|a_j\rangle$, a coupling is made between $\hat{\Pi}_{a_j}$ in $\mathscr{H}_S$ and $\hat{\sigma}_{yA}$ in $\mathscr{H}_A$: the system then evolves according to $\hat{U}_A(a_j) = e^{-i\theta_A\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}}$ in $\mathscr{H}_S \otimes \mathscr{H}_A$ and $\mathbb{1}_B$ in $\mathscr{H}_B$. Afterwards, a second coupling involves the other ancilla and $\hat{\Pi}_{b_0}$, with $|b_0\rangle \equiv \frac{1}{\sqrt{d}}\sum_{j=1}^{d}|a_j\rangle$ as usual. The corresponding evolution operator is $\hat{U}_B = e^{-i\theta_B\hat{\Pi}_{b_0}\otimes\hat{\sigma}_{yB}}$ in $\mathscr{H}_S \otimes \mathscr{H}_B$ whereas the first pointer remains unchanged. After both interactions, the system is in a state described by:

$$\hat{\varrho}'_{tot}(a_j) = \hat{U}_B\hat{U}_A(a_j)\hat{\varrho}_{tot}\hat{U}_A^\dagger(a_j)\hat{U}_B^\dagger \tag{3.5}$$

Subsequently, a postselection on $|a_k\rangle$ occurs on $\mathscr{H}_S$, leaving the pointers in a state that I call $\hat{\varrho}_{AB}$, since its derivation is quite lengthy, it is postponed to Appendix A.3.

At this point, ϱ_{jk} can be extracted through different measurement strategies on the ancillae. Some of them only work for particular values of d or of $\theta_{A,B}$, but others are general. It is certainly possible to extend Lundeen's formulas of (2.33) (which are still valid in the limit of $\theta_{A,B} \to 0$) to the case of arbitrary strength by adding some

measurements on $\hat{\varrho}_{AB}$:

$$\begin{aligned}\Re(\varrho_{jk}) &= \frac{d}{4\sin\theta_A\sin\theta_B}\Big(\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle^{a_k} - \langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k} \\ &\quad + 2\tan\frac{\theta_B}{2}\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle^{a_k} + 2\tan\frac{\theta_A}{2}\langle\hat{\Pi}_{1A}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle^{a_k} + 4\tan\frac{\theta_A}{2}\tan\frac{\theta_B}{2}\langle\hat{\Pi}_{1A}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle^{a_k}\Big) \\ \Im(\varrho_{jk}) &= \frac{d}{4\sin\theta_A\sin\theta_B}\left(\langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle^{a_k} + \langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k} + 2\tan\frac{\theta_B}{2}\langle\hat{\sigma}_{yA}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle^{a_k}\right)\end{aligned} \tag{3.6}$$

Yet much easier results exist:

$$\begin{aligned}\varrho_{jj} &= \frac{d^2}{\sin^2\theta_A\sin^2\theta_B} \qquad \langle\hat{\Pi}_{1A}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle^{a_k} \qquad \forall k \\ \Re(\varrho_{jk}) &= -\frac{d}{2\sin\theta_A\sin\theta_B} \qquad \langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k} \qquad j\neq k \\ \Im(\varrho_{jk}) &= \frac{d}{2\sin\theta_A\sin\theta_B} \qquad \langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k} \qquad j\neq k\end{aligned} \tag{3.7}$$

The abundance of superscripts $(^{a_k})$ means that postselection probability $\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})$ is left implicit. As previously explained in Section 2.6, this factor is not problematic as it can be absorbed in the measurement of the pointer observables if joint probabilities (outcome and postselection) are used in place of conditioned ones or if the density operator can be normalized after its determination.

It is clear that, again, the cost of strong couplings is represented by measurements of $\hat{\Pi}_1$ on the two ancillae. It is also noticeable that, as usual, $\hat{\sigma}_y$ appears in the determination of the imaginary parts of ϱ_{jk}, while its complementary observable $\hat{\sigma}_x$ is necessary for the real parts.

3.2.2 The Dirac Distribution Protocol

Albeit less direct, the second of Lundeen's schemes is still worthy of study because of its practical feasibility, improved by the fact that it needs only one ancilla. Just as in the previous case, a coupling has to be made between the pointer, in particular its $\hat{\sigma}_{yA}$ operator, and projector $\hat{\Pi}_{a_j}$ in the object system:

$$\hat{\varrho}_{tot} = \hat{\varrho}_S\otimes\hat{\varrho}_A \qquad \rightarrow \qquad \hat{\varrho}'_{tot} = e^{-i\theta_A\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}}\hat{\varrho}_{tot}e^{i\theta_A\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}} \tag{3.8}$$

Then, state $|b_l\rangle \equiv \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1}|a_j\rangle e^{i\frac{2\pi jl}{d}}$ is postselected, and the pointers associated to systems that pass are measured. I denote:

$$\begin{aligned}\rho_{A1}(jl) &\equiv \frac{1}{2\sin(\theta_A)}\left(\langle\hat{\sigma}_{xA}^{(a_j)}\rangle^{b_l} + i\langle\hat{\sigma}_{yA}^{(a_j)}\rangle^{b_l}\right) \\ \rho_{A2}(jl) &\equiv \frac{1}{\sin(\theta_A)}\langle\hat{\Pi}_{1A}^{(a_j)}\rangle^{b_l}\end{aligned} \tag{3.9}$$

where $\rho_{A2}(jl)$ is independent of l and superscripts $(^{b_l})$ imply that mean values are to be multiplied by $\mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho}_S)$. If the interest is actually in the Dirac distribution, its elements can be found as:

$$D_{jl} = \rho_{A1}(jl) + \tan\left(\frac{\theta_A}{2}\right)\rho_{A2}(jl) \tag{3.10}$$

However, after looping over all l and choosing one $k \in 0..d-1$

$$\varrho_{jk} = d \tan\left(\frac{\theta_A}{2}\right) \delta_{jk} \rho_{A2}(jl) + \sum_l e^{\frac{2\pi i l(j-k)}{d}} \rho_{A1}(jl) \tag{3.11}$$

This procedure can be repeated with different j or k in order to reconstruct the entire density operator.

3.2.3 A Summary of the two Protocols

From now on these two protocols will be termed *DRDO* (after *direct reconstruction of the density operator*) and *DRDD* (*direct reconstruction of the Dirac distribution*). The second will be considered an indirect way to obtain the density matrix, with the advantage of using one less ancilla compared to the first. Here is a brief summary of their steps.

DRDO	DRDD
1. Preparation of the pointer in state $\hat{\varrho}_A \otimes \hat{\varrho}_B$.	1. Preparation of the pointer in state $\hat{\varrho}_A$.
2. Choice of j, coupling between $\hat{\Pi}_{a_j}$ and $\hat{\sigma}_{yA}$ and subsequent evolution.	2. Choice of j, coupling between $\hat{\Pi}_{a_j}$ and $\hat{\sigma}_{yA}$ and subsequent evolution.
3. Coupling between $\hat{\Pi}_{b_0}$ and $\hat{\sigma}_{yB}$ and subsequent evolution.	3. Choice of l and postselection of state $\|b_l\rangle$ of the Fourier basis.
4. Choice of k and postselection of state $\|a_k\rangle$.	4. Extraction of $\rho_{A1}(jl)$ and $\rho_{A2}(jl)$ as in (3.9) through measurements of $\langle\hat{\sigma}_x\rangle^{b_l}$, $\langle\hat{\sigma}_y\rangle^{b_l}$ and $\langle\hat{\Pi}_1\rangle^{b_l}$ on the ancilla.
5. Extraction of ϱ_{jk} as in (3.7) through measurements of $\langle\hat{\sigma}_x\rangle^{a_k}$, $\langle\hat{\sigma}_y\rangle^{a_k}$ and $\langle\hat{\Pi}_1\rangle^{a_k}$ on the ancillae.	5. Repetition of steps 3-4 with all the different l.
6. Repetition of steps 2-5 with different j, k if more than one matrix element is needed.	6. Choice of k and extraction of ϱ_{jk} as in (3.11) from the results of step 5.
	7. Repetition of only step 6 with different k if more than one matrix element is needed in the j-th row.
	8. Repetition of steps 2-6 with different j, k if other matrix elements are needed.

TABLE 3.1: List of steps of the DRDO and DRDD protocols.

The step of normalization of the results is implicit in the determination of pointer mean values.

3.3 Comparison with Weak Reconstruction Protocols

As previously stated, strong measurements are usually easier to practically implement because they are in essence equivalent to standard projections. In the next chapter I will show that even if one wants to visualize the role of the pointer with the use of a quantum system that plays it, reducing the strength of the coupling is not always simple. However, this section is devoted to explaining the intrinsic advantages offered by a greater strength in the measurement schemes discussed so far, in terms of accuracy (lack of biases) and precision (small statistical errors). As usual, the focus is mainly on the DRDO protocol.

3.3.1 The Inherent Bias of Weak Measurements

Lundeen's direct reconstruction method is summarized by equation (2.33) that is here reported:

$$\begin{aligned}\Re(\hat{\varrho}_{jk}) &\approx \frac{d}{4\sin(\theta_A)\sin(\theta_B)}\left(\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle^{a_k} - \langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k}\right)\\ \Im(\hat{\varrho}_{jk}) &\approx \frac{d}{4\sin(\theta_A)\sin(\theta_B)}\left(\langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle^{a_k} + \langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k}\right)\end{aligned} \tag{3.12}$$

This is still a first order approximation in parameter $\sin(\theta_A)\sin(\theta_B)$ and has to be corrected as in equation (3.6) or replaced as in (3.7) when the coupling strength cannot be considered close to 0. It is important to study the importance of this bias for growing strength to understand the errors of a realization of this method. One way to do so is to compute the *trace distance* between the measured result and the actual state. Let these be called respectively $\hat{\varrho}^W$ and $\hat{\varrho}$, then this distance is defined as:

$$t(\hat{\varrho}^W, \hat{\varrho}) \equiv \frac{1}{2}\operatorname{Tr}\sqrt{(\hat{\varrho}^W - \hat{\varrho})^\dagger(\hat{\varrho}^W - \hat{\varrho})} \tag{3.13}$$

It is possible to find (Appendix A.3) the elements of $\hat{\varrho}^W - \hat{\varrho}$ as

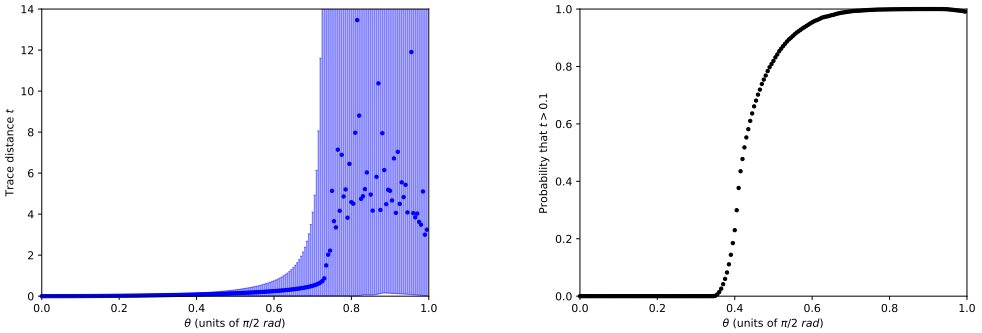
$$\varrho_{jk}^W - \varrho_{jk} = (\cos(\theta_A) - 1)\delta_{jk}\varrho_{jk} + \frac{\cos(\theta_B) - 1}{d}\sum_l \varrho_{jl} + \frac{(\cos(\theta_A) - 1)(\cos(\theta_B) - 1)}{d}\varrho_{jj} \tag{3.14}$$

and notice that they vanish for small $\theta_{A,B}$, thus validating Lundeen's results. Nonetheless, the presence of term $\sum_l \varrho_{jl}$ makes $\hat{\varrho}^W$ manifestly non-hermitian, against the very definition of density operator. In order to improve the result, one can take the hermitian part of $\hat{\varrho}^W$ and normalize its trace to 1:

$$\hat{\varrho}^{WHN} \equiv \frac{\frac{\hat{\varrho}^W + \hat{\varrho}^{W\dagger}}{2}}{\operatorname{Tr}\left(\frac{\hat{\varrho}^W + \hat{\varrho}^{W\dagger}}{2}\right)} \tag{3.15}$$

There is still no guarantee that this matrix is positive semidefinite, but since it has unit trace, the comparison offered by $t(\hat{\varrho}^{WHN}, \hat{\varrho})$ is more meaningful, moreover it is realistic to assume that an experimenter who measures the entire $\hat{\varrho}^W$ would convert it to $\hat{\varrho}^{WHN}$ also to make the procedure more robust to statistical and systematic errors of the implementation.

The analytic calculation of $t(\hat{\varrho}^{WHN}, \hat{\varrho})$ is not trivial, therefore I built a simulation (detailed in Appendix A.5.1) that randomly generates a valid state $\hat{\varrho}$ and computes $\hat{\varrho}^{WHN}$ and $t(\hat{\varrho}^{WHN}, \hat{\varrho})$ for different values of $\theta \equiv \theta_A = \theta_B$.



(A) Average trace distance. Vertical bars stretch to the minimum and maximum values.

(B) Probability that $t > 0.1$.

FIGURE 3.1: Results of the simulation for 10^4 different bidimensional states, each measured with 200 different values of $\theta \in]0, \frac{\pi}{2}[$.

As expected the trace distance grows with the strength. Figure 3.1b shows that this method is approximately sure to provide accurate results (i.e. $t(\hat{\varrho}^{WHN}, \hat{\varrho}) \leq 0.1$) if $\theta \lesssim 0.15\pi$, which is an experimentally viable value. However it is also clear that the method is off the mark for θ close to $\frac{\pi}{2}$ which precludes its use with more convenient strong measurements. In particular, values $t(\hat{\varrho}^{WHN}, \hat{\varrho}) > 1$ in Figure 3.1a mean that the produced matrix is non-physical and does not contain any information from which $\hat{\varrho}$ can be recovered.

3.3.2 The Role of Experimental Errors

When I introduced weak measurements in Section 2.1, I underlined that the distribution of results is large compared to its mean value and that many repetitions are needed to achieve meaningful precision. A similar problem affects the DRDO protocol if the strength of the coupling between object and ancilla is kept small. Equation (3.7) shows that for non-diagonal elements, the measured quantities are mean values in the form

$$\langle \hat{\sigma}_\mu^{(a_j)} \hat{\sigma}_\nu^{(b_0)} \rangle^{a_k} = \sum_{a,b=\pm 1} ab \langle \hat{\Pi}_{\mu a}^{(a_j)} \hat{\Pi}_{\nu b}^{(b_0)} \rangle^{a_k} \tag{3.16}$$

where $\mu, \nu \in \{x, y\}$, $\hat{\Pi}_{\mu a}$ and $\hat{\Pi}_{\nu b}$ are the elementary projectors that appear in the spectral decomposition of $\hat{\sigma}_\mu$ and $\hat{\sigma}_\nu$ and the mean values are calculated on the unnormalized (hence the superscript) pointer state after postselection $\langle a_k | \varrho'_{tot}(a_j) | a_k \rangle$.

As mentioned before, an experiment only measures counts $N_{\mu a, \nu b, jk} = N \cdot \langle \hat{\Pi}_{\mu a}^{(a_j)} \hat{\Pi}_{\nu b}^{(b_0)} \rangle^{a_k}$, with N representing the total power of the signal. These results are affected by Poissonian statistical error

$$\delta N_{\mu a, \nu b, jk} = \sqrt{N \langle \hat{\Pi}_{\mu a}^{(a_j)} \hat{\Pi}_{\nu b}^{(b_0)} \rangle^{a_k}} \tag{3.17}$$

so that, using equation 3.16, error $\delta_{\mu\nu,jk}$ on $\langle\hat{\sigma}_\mu\hat{\sigma}_\nu\rangle'_{j,k}$ is

$$\delta_{\mu\nu,jk} = \frac{1}{N}\sqrt{\sum_{a,b=\pm1}(\delta N_{\mu a,\nu b,jk})^2} = \sqrt{\frac{\mathrm{Tr}(\langle a_k|\varrho'_{tot}(a_j)|a_k\rangle)}{N}} \tag{3.18}$$

which does not depend on μ,ν but can change with j,k. Here I am not considering the standard deviation of N, which depends on how it is measured. Although it would be wrong to neglect it in practice, its contribution is usually similar to that of the other counts, so that the order of magnitude of the global error is the same with or without it. See Appendix A.3 for a quick proof of the last equation and the detailed expression of the trace.

It is interesting to observe that the value of $\delta_{\mu\nu,jk}$ does not exhibit any strong dependence on $\theta_{A,B}$ (Figure 3.2a) and it can even be shown that the propagation of this error on all the matrix elements ϱ_{jk} is

$$\sqrt{\sum_{j,k}\delta^2_{\mu\nu,jk}} = \sqrt{\frac{d}{N}} \tag{3.19}$$

regardless of the strength of the coupling or the object state. This seems to indicate that stronger measurements do not improve precision. Yet, this picture changes when one considers that the relevant quantities are not $\langle\hat{\sigma}_\mu\hat{\sigma}_\nu\rangle'_{j,k}$ but rather the density matrix elements, for which the error grows by factor $\frac{d}{2\sin(\theta_A)\sin(\theta_B)}$ and is clearly bigger in the weak case. The expression for the relative errors is:

$$\begin{aligned}\frac{\delta(\Re(\varrho_{jk}))}{|\Re(\varrho_{jk})|} &= \frac{d}{2\sin(\theta_A)\sin(\theta_B)}\sqrt{\frac{\mathrm{Tr}(\langle a_k|\varrho'_{tot}(a_j)|a_k\rangle)}{N}}\frac{1}{|\Re(\varrho_{jk})|}\\ \frac{\delta(\Im(\varrho_{jk}))}{|\Im(\varrho_{jk})|} &= \frac{d}{2\sin(\theta_A)\sin(\theta_B)}\sqrt{\frac{\mathrm{Tr}(\langle a_k|\varrho'_{tot}(a_j)|a_k\rangle)}{N}}\frac{1}{|\Im(\varrho_{jk})|}\end{aligned} \tag{3.20}$$

In Lundeen's method, two mean values of the form $\langle\hat{\sigma}_\mu\hat{\sigma}_\nu\rangle'_{j,k}$ are averaged, thus making these results smaller by coefficient $\frac{1}{\sqrt{2}}$.

The presence of factor $\frac{1}{\sin(\theta_A)\sin(\theta_B)}$ reflects the fact that the density matrix elements are calculated from linear combinations of counts that are small compared to the counts themselves. Inverting equation (3.7) (again for non-diagonal elements) shows that for instance

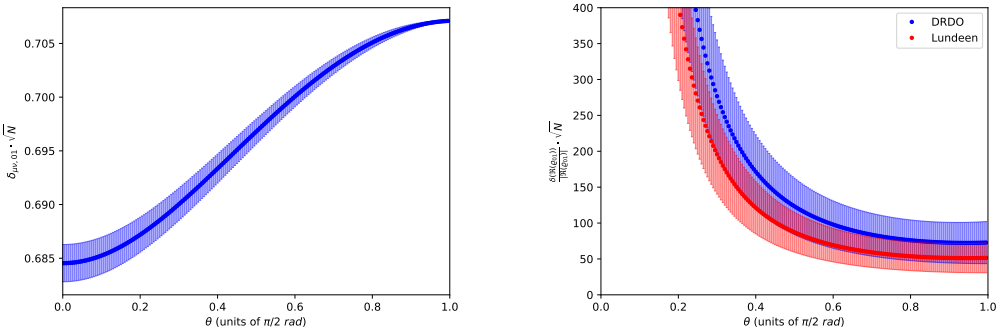
$$\left|\sum_{ab} abN_{ya,yb,jk}\right| = \frac{2\sin(\theta_A)\sin(\theta_B)}{d}N|\Re(\varrho_{jk})| \tag{3.21}$$

In the weak limit, this quantity is much smaller than N and therefore vulnerable to systematic experimental errors, as even a slight proportional bias on one of the counts $N_{\mu a,\nu b,jk}$ can radically change its value.

Different considerations apply to the case of diagonal elements, for which the dependence of the standard deviation on the strength is more evident, as only projector $\hat{\Pi}_{1A}\hat{\Pi}_{1B}$ has to be measured and

$$\frac{\delta(\varrho_{jj})}{\varrho_{jj}} = \frac{d}{\sin(\theta_A)\sin(\theta_B)}\frac{1}{\sqrt{N\varrho_{jj}}} \tag{3.22}$$

Here even the counts are small in the weak limit, thus making their relative errors more relevant.



(A) Standard deviation of a pointer observable needed for element 01.

(B) Relative error on the real part of density matrix element 01.

FIGURE 3.2: Results of the simulation for 10^4 different bidimensional states, each measured with 200 different values of $\theta \in]0, \frac{\pi}{2}[$. The choice of matrix element $\hat{\varrho}_{01}$ is not particularly relevant. The weak method shows better precision for equal strength only because it averages two pointer measurements. The high values on the vertical axis of 3.2b are due the fact that the contribution of $N^{-\frac{1}{2}}$ is neglected. Error bars reflect the standard deviation associated to the mean over the 10^4 samples.

The use of strong measurements, allowed by exact formulas such as (3.7), makes the DRDO protocol reach levels of precision (Figure 3.2b) that are unthinkable for weak couplings without building up the statistics. The only price to pay is, again, the need to implement projector $\hat{\Pi}_1$, which can cause an additional complication in the apparatus. Moreover, an accurate estimate of the coupling strength is needed in all the evolved proposals, while Lundeen's methods can ignore it if the entire density operator is measured, since the value of θ appears only as a multiplicative constant which can be eliminated by normalization. Nonetheless, the fact that the global number of needed pointer measurements is lower (one observable for each diagonal element, two for each non-diagonal element, against twice this amount in equation (2.33)), should make the experimental execution faster, an important advantage for increasing the strength.

3.4 Comparison with QST

Compared to the standard techniques of QST, the protocols described so far have the aforementioned advantage of directness: if a normalization is available, it is possible to find one matrix element with a very limited number of observations. For instance, the DRDO protocol can extract the real or imaginary part of a non-diagonal element of the density operator with only 4 different acquisitions, regardless of the dimension d of the object system.

Even when the entire reconstruction is needed, as is generally the case, QST can be inconvenient because it needs to implement a set of d^2 linearly independent projections. Experimentally, this is usually achieved with only one projecting device and d^2 unitary transformations that change the input state. Indeed if only $\hat{\Pi}_0$ is feasible in the set of $\{\hat{\Pi}_j \mid j = 0..d^2 - 1\}$ necessary projectors, one has to find $\{\hat{U}_j \mid j = 0..d^2 - 1\}$ such that $\hat{\Pi}_0 = \hat{U}_j \hat{\Pi}_j \hat{U}_j^\dagger$. In this way

$$\mathrm{Tr}(\hat{\Pi}_0 \hat{U}_j \hat{\varrho} \hat{U}_j^\dagger) = \mathrm{Tr}(\hat{\Pi}_j \hat{\varrho}) \tag{3.23}$$

This is not an easy task. One of the simplest possible cases is that of the polarization of a photon, which will also be treated in detail in the next chapter. To find it, one usually projects on the horizontal, diagonal and left-circular polarizations, and then calculates the normalization factor using a fourth direction, such as the vertical one, or with a raw measure of intensity (such as that of $\hat{\Pi}_0$ in (1.22)). If only one polarizer, for example on $|H\rangle$, is available, then a HWP is necessary to project on $|D\rangle$ and a *quarter-wave plate (QWP)* for $|L\rangle$. Every new linearly independent projection complicates the apparatus. The QST of this system is a standard procedure that can be carried out with modern equipment in a few seconds, but this may no longer be true if d becomes considerably larger than 2.

In the case of direct reconstruction protocols, this problem is reduced. The experimental setup must only extract from the system the d components corresponding to the states in the measurement basis, by implementing $\{\hat{\Pi}_{a_j} \mid j = 1..d\}$ for the first coupling and postselection, plus one additional projection on $|b_0\rangle$ for the second coupling. If the basis $\{|a_j\rangle\}$ is chosen cleverly, it is possible to find *one* unitary transformation $\hat{U}$ such that $\hat{\Pi}_0 = \hat{U}^j \hat{\Pi}_j (\hat{U}^\dagger)^j$. Consequently, regardless of d, the number of unrelated fundamental unitaries (that is, not counting powers of the same) is only 2, one is $\hat{U}$ and the other is needed for $\hat{\Pi}_{b_0}$.

The cost of this is represented by the two ancillae, which can sometimes complicate the system, thus making it more vulnerable to experimental errors. Moreover, they surely increase the number of required observations by a factor of 8 for non-diagonal elements (4 for the real part plus 4 for the imaginary part). However, only three different bases of the Hilbert space of each pointer are necessary, so that the number of transformations is again independent of d.

	QST	DRDO
Number of measurements for diagonal elements	d	d
Number of measurements for non-diagonal elements	$d^2 - d$	$8(d^2 - d)$
Total number of measurements	d^2	$8d^2 - 7d$
Different unitary transformations	d^2	$d + 1$ on the system 5 on each pointer
Unrelated unitary transformations	d^2	2 on the system 3 on each pointer

TABLE 3.2: Number of different measurements that are needed to reconstruct a density operator in a d-dimensional Hilbert space. The DRDO protocol is assumed to implement two bidimensional ancillae.

In conclusion, when d is large, the DRDO protocol allows experimenters to reconstruct the density operator with fewer different manipulations, so that the number and cost of components is reduced.

3.5 Measuring Products of Non-Commuting Observables with Strong Couplings

The extension of weak measurement methods to strong couplings is not necessarily limited to the field of state reconstruction. In this section I will briefly deviate from the main course of this thesis and show how the ancilla scheme can provide insight into the measurement of products of non-commuting observables.

Suppose to be interested in the mean value of the product of N observables on state $\hat{\varrho}_S$, that is:

$$\left\langle \prod_{j=1}^{N} \hat{O}_j \right\rangle_{\hat{\varrho}_S} = \mathrm{Tr}\left(\prod_{j=1}^{N} \hat{O}_j \hat{\varrho}_S \right) \tag{3.24}$$

If these observables do not commute with one another, their product is not hermitian. However, this mean value can still be measured by sequentially coupling the $\hat{O}_j$s (in reverse order) to different ancillae, as suggested by the DRDO protocol of Subsection 2.5.1. The main result exploiting weak measurement has already been mentioned in formula (2.31) and is detailed in Appendix A.2:

$$\left\langle \prod_{j=1}^{N} \hat{O}_j \right\rangle_{\hat{\varrho}_S} \approx \frac{1}{\prod_j \theta_j} \left\langle \prod_j \hat{\Sigma}_j \right\rangle_{F_N} \tag{3.25}$$

where $|F_N\rangle$ is the joint state of all N pointers after all the evolutions and $\hat{\Sigma}_j$ is the operator $\frac{\hat{\sigma}_x + i\hat{\sigma}_y}{2}$ on the j-th qubit ancilla Hilbert space.

As in the previous sections, this relation can easily be made exact and extended to arbitrary coupling strength θ when all the observables are projectors. It is sufficient to measure on each pointer the operator

$$\hat{E} \equiv \frac{\hat{\sigma}_x + i\hat{\sigma}_y}{2} + \tan\left(\frac{\theta}{2}\right) \hat{\Pi}_1 \tag{3.26}$$

then

$$\left\langle \prod_{j=1}^{N} \hat{\Pi}_j \right\rangle_{\hat{\varrho}_S} = \frac{1}{\prod_j \sin(\theta_j)} \left\langle \prod_j \hat{E}_j \right\rangle_{F_N} \tag{3.27}$$

Interestingly, it is also possible to find the mean value of each projector, simply by measuring $\hat{E}$ on the associated pointer and

$$\hat{E}' \equiv \hat{\Pi}_0 + \tan\left(\frac{\theta}{2}\right) \hat{\sigma}_x + \tan^2\left(\frac{\theta}{2}\right) \hat{\Pi}_1 \tag{3.28}$$

on all the others. Indeed

$$\langle \hat{\Pi}_k \rangle_{\hat{\varrho}_S} = \frac{1}{\prod_{j\neq k} \sin(\theta_j)} \left\langle \hat{E}_k \prod_{j\neq k} \hat{E}'_j \right\rangle_{F_N} \tag{3.29}$$

The study of non-commuting observables has always been a profound problem of quantum physics due to Heisemberg's uncertainty principle. Of course, this method does not violate the basic axioms of the theory, because the results are obtained statistically and

not from a single object. However, this is different from simply measuring each projector on disjoint subensembles of identical systems, because each acquisition contains information on all of them. For instance, the measurement of $\hat{\Pi}_1$ on the k-th ancilla appears in both $\hat{E}_k$ and $\hat{E}'_k$, therefore it contributes to $\langle\hat{\Pi}_k\rangle_{\hat{\varrho}_S}$ and to all $\langle\hat{\Pi}_{j\neq k}\rangle_{\hat{\varrho}_S}$. Similar procedures have been proposed and tested in the weak regime [50, 51], but the use of exact formulas such as these might provide more accurate results and reduce the statistical errors.

The derivation of these expressions is reported in Appendix A.4. As in the weak case, the fact that $\hat{E}$ and $\hat{E}'$ are not hermitian is not problematic because it is possible to measure the various pointer observables ($\hat{\sigma}_x$, $\hat{\sigma}_y$, $\hat{\Pi}_0$ and $\hat{\Pi}_1$) one by one and then sum the results.

I conclude adding that if the focus is on products of observables that are not projectors, one has to spectrally decompose them and apply this procedure to all the different combinations. Of course this means that the number of measurements grows with d^N, but this technique can still be convenient in some experimental cases. For instance if multiple apparati that can measure the various $\hat{O}_j$s are available, it may be better to exploit them than to handle the possibly much different product operator (or its hermitian and antihermitian parts if it is not an observable).

Chapter 4

The Experiment

The bulk of this work is the experimental realization of an apparatus that can reproduce the steps of the protocols reported in Section 3.2. Designed and built by Luca Calderaro and myself at the CNR-IFN Luxor laboratory in Padova, the setup that will be described in this chapter aims to reconstruct the polarization state of a photonic system using optical devices. In order to implement the ancilla scheme we used PBDs to couple the object to another qubit, the *path* taken by each photon in a Mach-Zehnder interferometer that acts as our pointer. This choice of degrees of freedom is dictated by the availability of components and by the expertise of our lab and research group. Of course, the fact that the object Hilbert space is only bidimensional puts our measurement schemes at a disadvantage with respect to standard QST, which can reconstruct any polarization density operator with only $d^2 = 4$ data points. However, this experiment is to be interpreted as a *proof of concept*, the goal of which is to show that direct reconstruction can be performed with strong measurements with improved results compared to those obtained in the weak case. For this purpose, our setup can change the strength of the interaction with the ancilla, making it possible to test our and Lundeen's protocols in both regimes.

Upstream of the measurement system proper we used a custom built source that exploits *spontaneous parametric down-conversion (SPDC)* to produce pairs of polarization-entangled photons of wavelength $\lambda \approx 809\ nm$. Compared to a continuous wave laser, this source has the disadvantage of lowering the power of the signal from some mWs to the thousands of photons per second (power $\sim 10^{-16}\ W$), making the use of *single photon avalanche diodes (SPADs)* necessary. These solid state detectors employ the impact ionization mechanism of reversely biased p-n junctions to induce avalanche currents every time they are hit by one photon. With the aid of two of these devices (Excelitas SPCM-AQRH-14-FC) and a fast time-tagger (qutools quTAU), it is possible to recognize the pairs from the difference in detection time: we used a 1 ns-wide window to define *coincident counts* that can be attributed to the entangled photons. Considering only these events, we could operate in true quantum conditions and we could easily produce mixed states by sending one member of each pair to the measurement setup and using the other only as a *herald*, so that its contribution to the joint state is effectively traced out.

After the detailed description of this source and of the principles on which it is based, this chapter will focus on the process of planning, building and calibrating the measurement apparatus.

4.1 The Source of Entangled Photons

Nowadays, entanglement is a fundamental piece of many quantum physics experiments because it is strikingly alien to the classical framework and permits to investigate the most mysterious aspects of the theory. It can exist between degrees of freedom of the same entity or in two separated ones, the latter case being the most interesting for quantum communication because it allows non-classical correlations to move between parties that are far apart.

Modern techniques can generate entangled states with atoms [52], ions [53] or superconductors [54], but the simplest and most widely used case is that of photons, in which correlated pairs are created through SPDC. However, correlation is not the only ingredient: the other is the superposition of (at least) two generation processes, so that it is not possible to determine from which one the photons have originated. I will now explain how our source achieves these two goals.

4.1.1 Correlation: The SPDC process

Spontaneous Parametric Down Conversion originates from the interaction between light and matter and exploits the non-linear optical properties of certain materials. Let $\vec{E}$ and $\vec{B}$ label the classical components of an EM field. Their effects on a medium are described by Maxwell's equations:

$$\begin{aligned} \vec{\nabla} \cdot \vec{D}(\vec{r},t) &= \rho(\vec{r},t) \\ \vec{\nabla} \cdot \vec{B}(\vec{r},t) &= 0 \\ \vec{\nabla} \times \vec{E}(\vec{r},t) + \frac{\partial \vec{B}(\vec{r},t)}{\partial t} &= 0 \\ \vec{\nabla} \times \vec{H}(\vec{r},t) - \frac{\partial \vec{D}(\vec{r},t)}{\partial t} &= \vec{J}(\vec{r},t) \end{aligned} \tag{4.1}$$

where $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$ and $\vec{H} = \frac{1}{\mu_0}\vec{B} - \vec{M}$ are respectively the displacement and magnetizing fields. Since most optical components are uncharged, non-conducting and non-magnetic, the density of free charges ρ, that of free currents $\vec{J}$ and the magnetization vector $\vec{M}$ are all null, so that the interaction is entirely captured by the polarization vector $\vec{P}$.

If χ is the non-linear electric susceptibility of the medium, then

$$P_i = \varepsilon_0 \left(\sum_j \chi^{(1)}_{ij} E_j + \sum_{jk} \chi^{(2)}_{ijk} E_j E_k + \sum_{jkl} \chi^{(3)}_{ijkl} E_j E_k E_l + \dots \right) \tag{4.2}$$

in which terms of order higher than 1 can give rise to EM fields at different frequencies. It is possible to write the Hamiltonian density as

$$\begin{aligned} \mathcal{H} &= \frac{1}{2}\left(\vec{E}\cdot\vec{D} + \vec{B}\cdot\vec{H}\right) \\ &= \frac{1}{2}\varepsilon_0 |\vec{E}|^2 + \frac{1}{2\mu_0}|\vec{B}|^2 + \frac{\varepsilon_0}{2}\sum_{ij}\chi^{(1)}_{ij}E_iE_j + \frac{\varepsilon_0}{2}\sum_{ijk}\chi^{(2)}_{ijk}E_iE_jE_k + \dots \\ &= \mathcal{H}_0 + \mathcal{H}_I \end{aligned} \tag{4.3}$$

in which $\mathcal{H}_I$ describes the interaction between fields and matter.

Without delving too much into the details of quantum field theory, suffice it to say that fields can be transformed into operators as prescribed by second quantization so that the interaction Hamiltonian density becomes

$$\begin{aligned}\hat{\mathcal{H}}_I \propto &\sum_{ij} \chi^{(1)}_{ij} \left(\hat{E}^{(+)}_i + \hat{E}^{(-)}_i\right)\left(\hat{E}^{(+)}_j + \hat{E}^{(i)}_j\right) \\ &+ \sum_{ijk} \chi^{(2)}_{ijk} \left(\hat{E}^{(+)}_i + \hat{E}^{(-)}_i\right)\left(\hat{E}^{(+)}_j + \hat{E}^{(i)}_j\right)\left(\hat{E}^{(+)}_k + \hat{E}^{(i)}_k\right) + \dots\end{aligned} \tag{4.4}$$

where $\hat{E}^{(+)}$ and $\hat{E}^{(-)}$ can respectively annihilate and create a field [55]. One particular term of the Hamiltonian is responsible for the pair generation:

$$\hat{H}_{SPDC} = \int d^3r \sum_{i_p j_s k_i} \chi^{(2)}_{i_p j_s k_i} \hat{E}^{(+)}_{i_p} \hat{E}^{(-)}_{j_s} \hat{E}^{(-)}_{k_i} \tag{4.5}$$

in which symbol p stands for the annihilated *pump* photon, whereas i and s label the created *idler* and *signal* photons.

In order for the flow of energy to steadily move from the pump to the produced photons during the propagation in the medium, the so-called *phase matching* conditions have to be valid:

$$\begin{aligned}\omega_p &= \omega_s + \omega_i \\ \vec{k}_p &= \vec{k}_s + \vec{k}_i\end{aligned} \tag{4.6}$$

If the second is not verified, which can happen because of dispersion, the SPDC process is canceled out by its opposite an the net efficiency stays low. There are various ways to achieve phase matching, some of which are connected to the polarization of the involved light:

- *Type I birefringent phase matching.* A birefringent crystal is used, the pump sees the extraordinary refractive index while idler and signal see the ordinary one (or viceversa). This means that the produced photons share the same polarization which is orthogonal to that of the pump. One way of verifying phase matching is:

  $$\omega_s = \omega_i \quad \text{and} \quad n_o(\omega_s) + n_o(\omega_i) = 2n_o(\omega_s) = 2n_e(\omega_p) \tag{4.7}$$

  Photons are emitted symmetrically in a cone the axis of which coincides with the propagation direction of the pump.

- *Type II birefringent phase matching.* A birefringent crystal is used, pump and signal share the same polarization, orthogonal to that of the idler. One way of verifying phase matching is:

  $$\omega_s = \omega_i \quad \text{and} \quad n_e(\omega_s) + n_o(\omega_i) = 2n_e(\omega_p) \tag{4.8}$$

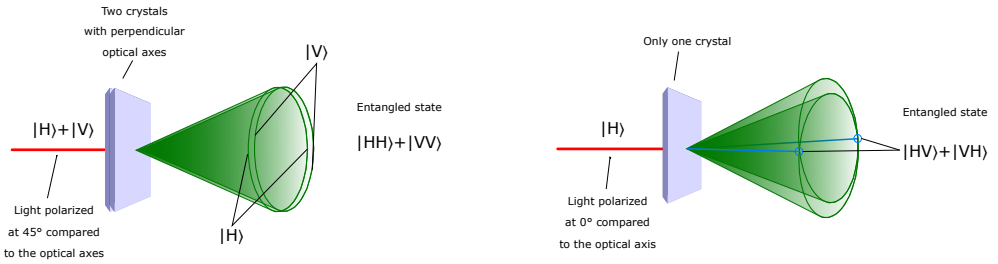
  Photons are emitted in two different cones.

- *Quasi phase matching.* The crystal is engineered so that the sign of $\chi^{(2)}$ changes periodically and the production contributions always sum constructively. This can be achieved also without birefringence and with only one polarization involved (a configuration known as *type 0 phase matching*).

In all cases the polarization states are correlated with one another.

4.1.2 Superposition: The Sagnac Interferometer

The polarization and direction of the emitted photons suggest different strategies to achieve the superposition necessary for reliable entanglement generation. For instance, two type I crystals can be paired with their optical axes orthogonal, so that the incoming pump light can interact with the same probability with each one [56]. A single crystal is sufficient for type II, provided photons are collected from the intersection of the two exit cones [57].



(A) Type I. (B) Type II.

FIGURE 4.1: Visualization of the pair production process for entanglement generation.

Our source, developed by Dr. Matteo Schiavon as part of his PhD project [58], employs a $KTiOPO_4$ crystal (RAICOL Crystals PPKTP) to achieve quasi phase matching, but the produced polarization states have the properties of type II SPDC (idler and signal are orthogonally polarized). However, superposition is realized in a way different from any of those of Figure 4.1.

The pump is provided by a CW laser diode (Ondax LM-405) of nominal wavelength 405 *nm* which enters a Sagnac interferometer through a *polarizing beam splitter (PBS)*. In this configuration the two arms share the same physical path and OPL, hence cannot introduce a phase difference. The polarization of the clockwise arm is soon switched from $|V\rangle$ to $|H\rangle$ by a HWP, so that both beams are horizontally polarized when they hit the crystal. Some of the photons from each arm trigger the SPDC process and are converted into pairs that are orthogonally polarized and have approximate wavelength 809 *nm*. Those of the counterclockwise path encounter the HWP (which is designed to work at both wavelengths) and exchange their polarizations. At the PBS, the two idler photons proceed towards the A exit and the two signals towards B, where they are collected into two single mode fibers and sent to the rest of the experiment. The superposition request is fulfilled by the fact that it is impossible to ascertain from which arm each of the signal (or idler) photons have come from. The pump light exiting the interferometer from the entrance port is deflected by dichroic mirrors and filtered out.

This source can generate maximally entangled states such as the singlet

$$|\Psi^-\rangle = \frac{|H_A V_B\rangle - |V_A H_B\rangle}{\sqrt{2}} \tag{4.9}$$

Indeed, just before the fibers, the state is described by

$$|\psi_{pre}\rangle = a|H_A V_B\rangle - e^{i\phi} b|V_A H_B\rangle \tag{4.10}$$

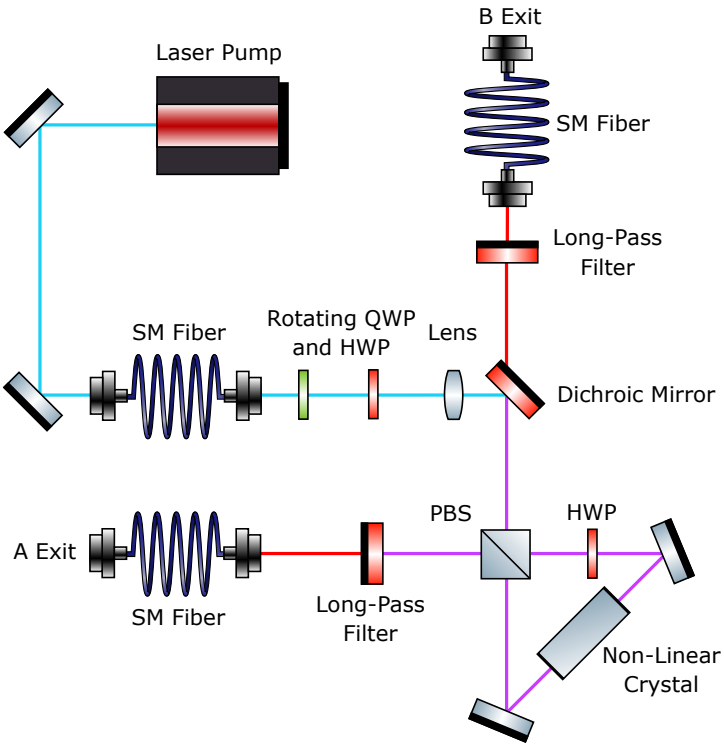


FIGURE 4.2: Scheme of the source. For reasons of mobility, the laser pump is separated from the rest of the setup and connected to it via a single mode (SM) fiber. Originally, a polarization maintaining fiber was used in its place, which allowed to keep the linear polarization of the pump. Unfortunately, it had to be replaced due to damage and a QWP was added to better tune the phase between the $|H\rangle$ and $|V\rangle$ components.

where a, b, ϕ are real parameters that depend on the polarization of the pump light entering the interferometer and $a^2 + b^2 = 1$. By rotating the HWP before the PBS, it is always possible to tune a and b so that $a = b = \frac{1}{\sqrt{2}}$, then it is only necessary to compensate for ϕ and for the effects of the two exit fibers. Supposing that these apply arbitrary unitary transformations $\hat{U}_A$ and $\hat{U}_B$ to the states of the photons that pass through them, the polarization after the fiber is

$$|\psi_{post}\rangle = (\hat{U}_A \otimes \hat{U}_B)\frac{|H_A V_B\rangle - e^{i\phi}|V_A H_B\rangle}{\sqrt{2}} \tag{4.11}$$

If further transformation $\hat{U} = \hat{U}_A \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \hat{U}_B^{-1}$ is applied to the light exiting the B fiber (for example with additional waveplates), the final state is

$$\begin{aligned} |\psi_{final}\rangle &= (\mathbb{1}_2 \otimes \hat{U})|\psi_{post}\rangle = e^{i\phi}\frac{\hat{U}_A|H_A\rangle \otimes \hat{U}_A|V_B\rangle - \hat{U}_A|V_A\rangle \otimes \hat{U}_A|H_B\rangle}{\sqrt{2}} \\ &= (\hat{U}_A \otimes \hat{U}_A)|\Psi^-\rangle = |\Psi^-\rangle \end{aligned} \tag{4.12}$$

where the last equation is justified by the fact that $(\hat{U}_A \otimes \hat{U}_B)|\Psi^-\rangle = |\Psi^-\rangle$ as long as $\hat{U}_B = \hat{U}_A$.

However, for our experiment, the production of such a precise state is not necessary, it is sufficient to obtain states in the form of equation (4.10) with some control over a and b offered by the HWP.

4.1.3 Performance and Polarization Stability

The number of produced pairs depends on many different factors, the most important of which is of course the power of the pump laser. Due to the small $\chi^{(2)}$ coefficient, only a tiny fraction of the photons that hit the crystal is converted. Of the results, many are lost in the fiber coupling and some are not detected because of the imperfect SPAD efficiency (nominally $\eta_{SPAD} \sim 45\%$).

We analyzed the production for different power values by counting photons just after the two exit fibers of the source. The ratio between coincident and total counts is $\eta_{coinc} \approx 7\%$, which suggests that the efficiency of the fiber coupling is around $\eta_{coupling} = \frac{\eta_{coinc}}{\eta_{SPAD}} \approx 16\%$. Such a small value is mostly due to the fact that the signal and idler beams are not collimated because of the lens that focuses the pump light at the center of the crystal. This compromise was made as it greatly improves the production efficiency of the material itself.

Driving Current (mA)	Pump power on the crystal (μW)	Photon kilo-counts/s	Coincident Photon kilo-counts/s
20	10.76 ± 0.01	2.8 ± 0.1	0.19 ± 0.01
25	343 ± 1	116 ± 1	10.3 ± 0.1
30	2302 ± 2	598 ± 1	44.3 ± 0.2
35	3331 ± 1	961 ± 5	67.2 ± 0.4
40	6840 ± 1	1580 ± 7	106.2 ± 0.3
45	9210 ± 2	2065 ± 2	138.2 ± 0.5
50	11050 ± 4	2509 ± 3	166.0 ± 0.5
55	12660 ± 4	2962 ± 4	196.6 ± 0.5

TABLE 4.1: Detected photons with different power settings. The pump power on the crystal does not coincide with the optical power emitted by the diode laser, much of which is lost at the first fiber coupling and some is reflected by the optical components before the interferometer. The third column lists the average tally of the two SPADs.

For the rest of the experiment we set the laser diode at a driving current of 50 mA, corresponding to around 11 mW of pump light entering the interferometer. With the aforementioned values of η_{SPAD} and $\eta_{coupling}$, we can estimate that approximately one in a billion pump photons is converted and around 3.5 million pairs are produced each second. Considering that the measurement apparatus normally loses $95 - 99$ % of the signal (depending on the settings, see Section 4.2), the rate of coincident counts used in the actual experiments is in the low thousands per second.

Another necessary condition for the success of the tomography process is that the measured state remains the same for all its duration. We could confirm this by means of checking the pump power in the two arms of the interferometer, which is connected to the a and b coefficients of the previous section.

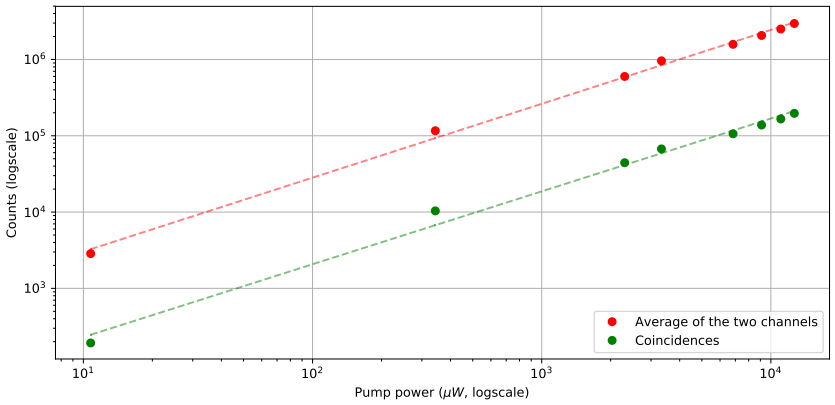


FIGURE 4.3: Counts as a function of power with linear fit of the logarithmic curve. Error bars are smaller than marker size. The linear law is confirmed by slopes ≈ 0.95.

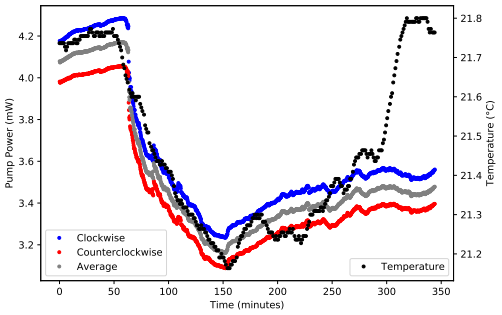


FIGURE 4.4: Power in the clockwise (blue) and counterclockwise (red) arms as a function of elapsed time from the start of the acquisition. Their average is shown in grey. The black curve is a concurrent measure of temperature.

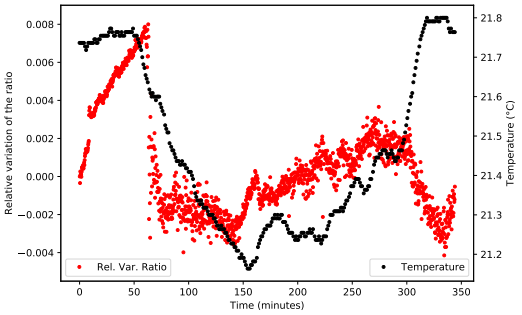


FIGURE 4.5: Variation of the ratio between powers in the two arms, relative to its initial value. Temperature in black.

Figure 4.5 shows that the ratio between the two values is approximately constant. This means that the SM fiber that brings the pump light to the Sagnac interferometer does not induce time dependent transformations on the polarization state (it could change the phase ϕ between the horizontal and vertical components, but these unitaries form a set of measure 0 in the space of all possibilities, therefore this behavior is extremely unlikely). Supposing that the exit fibers are similar and that all losses are polarization-independent or anyway also constant, we can conclude that the state entering the measurement apparatus is stable for long periods of time.

4.1.4 Temporal Coherence of the Photons

Temporal coherence is a measure of the phase correlations between two values of a wave that are separated in time and shows how strongly they can interfere with each other. Since interference is the foundation of our measurement setup it is important to understand if and how well the used light can display this phenomenon.

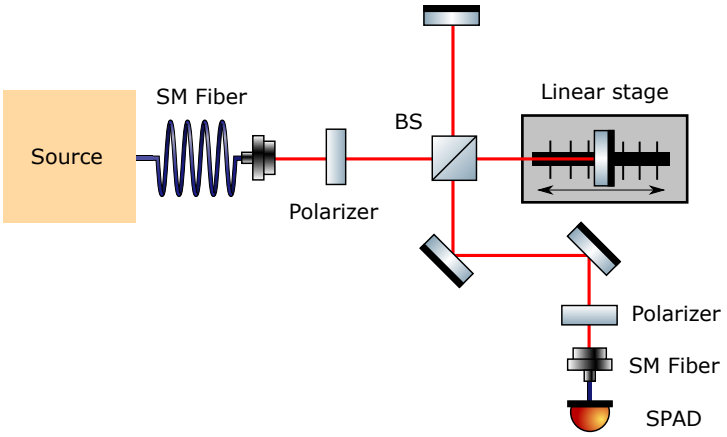


FIGURE 4.6: Scheme of the Michelson interferometer used to study temporal coherence.

In order to study this property we built a *Michelson interferometer* with a moving arm (Figure 4.6) and sent the single photons produced by our source through it. After taking one of the paths, photons reenter the beam splitter and reach a single mode fiber where the two spatial modes are projected onto one. In this way the strength of the interference does not depend on their spatial overlap but only on how well they are coupled into the fiber and on their optical path difference Δl. The intensity, measured as photon counts per second by the SPAD, can be expressed as:

$$I_{int}(\Delta l) = I_F + I_V(\Delta l) + \frac{2\sqrt{I_F I_V(\Delta l)}}{I_F + I_V(\Delta l)}\gamma \cos\big(\varphi(\Delta l)\big) \tag{4.13}$$

where I_F and $I_V(\Delta l)$ are the intensities of the beams taking the path with fixed or variable length, φ is the phase difference between the two and $\gamma \in [0, 1]$ is a coefficient representing the phase correlation. Due to the precision of the linear stage that moves the mirror, we could not observe the oscillating part of the pattern but only its envelope, which is, however, enough to infer meaningful information about temporal coherence.

We changed the length of the moving arm with steps of 1 μm, each corresponding to an actual variation of Δl of 2 μm because light travels its path twice before getting back to the beam splitter. After measuring I_F once and for all, we registered $I_{int}(\Delta l)$ and $I_V(\Delta l)$ at each position, considering only coincidence counts in the two SPADs, so as not to include background light entering the fiber. Then, we measured the *visibility* as

$$V \equiv \frac{I_{int,max} - I_{int,min}}{I_{int,max} + I_{int,min}} \tag{4.14}$$

where max and min refer to intervals of ten steps. If we can consider $I_V(\Delta l)$ constant in each 20 μm wide interval, a reasonable approximation, then $V \approx \frac{2\sqrt{I_F I_V(\Delta l)}}{I_F + I_V(\Delta l)}\gamma$, from which we can extract γ. By defining the coherence length Δl_c as the value of Δl for which the phase correlation is $1/e$, Figure 4.7 tells us that

$$\Delta l_c \approx (9 \pm 1) \cdot 10^{-5}\ m \tag{4.15}$$

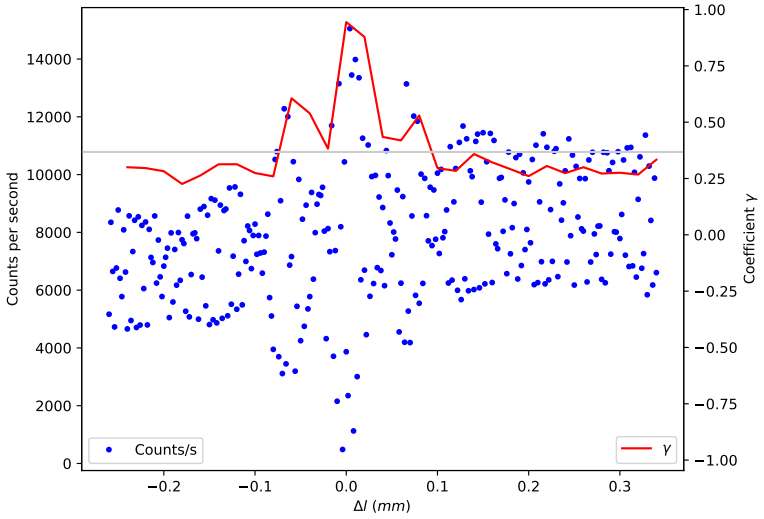


FIGURE 4.7: Results of the coherence analysis. Due to imperfect alignment, I_V changes with Δl, although it can be considered constant in each 20 μm wide interval. This causes the slight asymmetry in the power values (blue) between positive and negative OPL regions. This has been taken into account by correcting γ (red) as $\gamma = V \cdot \frac{I_F + I_V(\Delta l)}{2\sqrt{I_F I_V(\Delta l)}}$. The gray line corresponds to $\gamma = 1/e$.

which corresponds to a temporal coherence of $\Delta t_c = \frac{\Delta l_c}{c} \approx (30 \pm 3) \cdot 10^{-14}$ s. This small value shows that a strong interference is only available if paths are very carefully balanced. The presence of lateral lobes in the visibility graph suggests that the spectrum of the analyzed photons might be multi-modal.

4.2 The Measurement System

The center of this experiment is the measurement setup. As briefly mentioned at the beginning of this chapter, this apparatus has the purpose of faithfully implementing the direct measurement strategies of Section 3.2 to the polarization states of photons. Specifically, it was built with the DRDO scheme in mind, but can also easily perform all the steps of the simpler DRDD protocol. First we built a version that is only able to actualize strong couplings between object and ancilla, then we made some additions that allowed to adjust the weakness of the interaction, which will be described in Section 4.3.

4.2.1 Design

The setup was developed around three goals:

- Separation between object and ancilla.
- Measurements on the ancilla.
- Postselection.

Since we chose the path as our pointer, the first objective can be fulfilled by any optical device that can spatially separate a beam into two polarized components. One of the two states in the measurement basis $\{|H\rangle, |V\rangle\}$ is chosen as $|a_j\rangle$ and sent into path $|1_A\rangle$ while the other stays in $|0_A\rangle$. Considering for simplicity a pure initial state, the action of this piece can be summarized by

$$|\psi\rangle = (c_{j\perp}|a_{j\perp}\rangle + c_j|a_j\rangle) \otimes |0_A\rangle \;\; \rightarrow \;\; |\psi'\rangle = e^{-i\frac{\pi}{2}\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}}|\psi\rangle = c_{j\perp}|a_{j\perp}\rangle|0_A\rangle + c_j|a_j\rangle|1_A\rangle \tag{4.16}$$

Notice that $|0_A\rangle$ labels both the beam course before the displacement and one of the paths after it. The strength of the interaction is maximal ($\theta = \frac{\pi}{2}$), which means that the two polarizations are completely separated into two orthogonal (non-overlapping, perfectly distinguishable) pointer states and there is symmetry between object and ancilla.

From formulas (3.6), (3.7) and (3.9), it is clear that the setup must projectively measure $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $\hat{\Pi}_1$ on the ancilla, therefore the pointer state can collapse on one of six possibilities: $|\pm_A\rangle$ (eigenstates of $\hat{\sigma}_x$), $|\pm i_A\rangle$ (eigenstates of $\hat{\sigma}_y$), $|1_A\rangle$ or $|0_A\rangle$ (successful or unsuccessful projection on $|1_A\rangle$). Simultaneously, the object collapses too, so that the joint system lies in one of the six separable states of Table 4.2.

Measurement	Outcome	Probability	Final state
$\hat{\sigma}_x$	1	$\frac{1}{2}$	$(c_{j\perp}\|a_{j\perp}\rangle + c_j\|a_j\rangle) \otimes \|+_A\rangle$
$\hat{\sigma}_x$	-1	$\frac{1}{2}$	$(c_{j\perp}\|a_{j\perp}\rangle - c_j\|a_j\rangle) \otimes \|-_A\rangle$
$\hat{\sigma}_y$	1	$\frac{1}{2}$	$(c_{j\perp}\|a_{j\perp}\rangle - ic_j\|a_j\rangle) \otimes \|+i_A\rangle$
$\hat{\sigma}_y$	-1	$\frac{1}{2}$	$(c_{j\perp}\|a_{j\perp}\rangle + ic_j\|a_j\rangle) \otimes \|-i_A\rangle$
$\hat{\Pi}_1$	1	$\|c_j\|^2$	$\|a_j\rangle \otimes \|1_A\rangle$
$\hat{\Pi}_1$	0	$\|c_{j\perp}\|^2$	$\|a_{j\perp}\rangle \otimes \|0_A\rangle$

TABLE 4.2: The states after the useful pointer measurements. These results are also valid for the second coupling provided a_j and $a_{j\perp}$ are replaced by b_0 and b_1.

We implemented the measurement process by closing the interferometer in a way that can morph $|\psi'\rangle$ selectively into each of these states. In the first four cases a phase shift must be introduced between the two beams that have then to be recombined so that half the power is discarded, whereas for the measurements of $\hat{\Pi}_1$ one arm has to be blocked.

Subsequently, another identical block receives the photons, realizes the coupling with the second ancilla and performs the same measurements on it. The only difference with the first one is that the interacting polarization states are $|b_0\rangle = |D\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$ and $|b_1\rangle = |A\rangle = \frac{|H\rangle-|V\rangle}{\sqrt{2}}$ instead of being $|a_j\rangle, |a_{j\perp}\rangle \in \{|H\rangle, |V\rangle\}$.

Measurement	Outcome	Delay	Block
$\hat{\sigma}_x$	1	0 rad	/
$\hat{\sigma}_x$	-1	π rad	/
$\hat{\sigma}_y$	1	$\frac{3\pi}{2}$ rad	/
$\hat{\sigma}_y$	-1	$\frac{\pi}{2}$ rad	/
$\hat{\Pi}_1$	1	Any	$\|0_A\rangle$
$\hat{\Pi}_1$	0	Any	$\|1_A\rangle$

TABLE 4.3: Phase delays that must be applied to the $|1_A\rangle$ path compared to $|0_A\rangle$. Notice that a delay of $\frac{3\pi}{2}$ rad to a beam is equivalent to a delay of $\frac{\pi}{2}$ rad to the other.

Finally, postselection on $|a_k\rangle \in \{|H\rangle, |V\rangle\}$ is carried out on the object system and the results are sent to a detector that counts the photons for a fixed exposure time.

4.2.2 Practical Realization

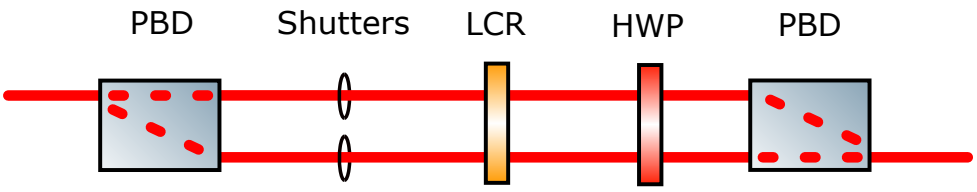


FIGURE 4.8: Scheme of each Mach-Zehnder interferometer.

The key building blocks of this apparatus are the two identical Mach-Zehnder interferometers that split and recombine the two polarizations and carry out the ancilla measurements (Figure 4.8). These are based on PBDs, provided by Thorlabs under model name BD40, which employ birefringent calcite crystals to separate light into two parallel beams at a distance of approximately (4.1 ± 0.1) mm at the operating wavelength of 809 nm. We chose to use PBDs instead of more conventional beam splitters because they made the alignment process much easier and allowed us to interact with both beams using the same optics, thus reducing the cost of the setup.

Various possibilities were considered for the application of the phase delays between the paths. The most straightforward one was the use of two glass plates, one of which could be tilted so as to tune the optical path length (OPL) of the photons passing through it. The other plate was necessary because of the short coherence time of our source, which did not allow for too big a difference in OPL. This option was discarded when we noticed that imperfections in the available glass plates caused a deflection of the beams which was difficult to compensate for, especially considering that the plates had to be frequently moved during each measurement.

Then we contemplated tiltable birefringent waveplates spanning over both beams. If aligned with their optical axes parallel or perpendicular to the polarization of the incoming light, a rotation around said axes could provide the desired phase delay. However the plates had to be tilted at an angle so great that the beams were clipped, thus forcing another change of strategy

Finally, our choice fell on *liquid crystal retarders (LCRs)*, devices which can change their birefringence with an applied voltage. Unfortunately, the only available model, Thorlabs LCC1221-A, is optimized for a shorter wavelength and can only provide a phase shift in range $[\sim 0, \sim 0.8\pi]$ rad at 809 nm, therefore we mounted the two pieces on motorized rotators (Thorlabs PRM1Z8) so as to change the orientation of their optical axes. In this way we effectively doubled the achievable delay range to $[\sim -0.8\pi, \sim 0.8\pi]$ rad. We introduced an additional fixed phase shift of $\sim 0.8\pi$ rad by slightly tilting the PBDs to cover the more useful interval $[\sim 0, \sim 1.6\pi]$ rad, as prescribed by Table 4.3.

After each LCR, a HWP with its axis oriented at $\frac{\pi}{8}$ *rad* compared to the horizontal has the purpose of changing the polarizations so that each of the exits of the subsequent PBD collects half the power of each of the incoming beams, as required by the measurements of $\hat{\sigma}_x$ and $\hat{\sigma}_y$. For those of $\hat{\Pi}_1$ we blocked the $|0\rangle$ path (which is the one taken by the $|a_{j\perp}\rangle$ and $|b_1\rangle$ polarizations) using manual shutters. In this case the halving is undesired, therefore we compensated for it by doubling the resulting counts in post-processing.

We chose to position the closing PBD of each interferometer in the same manner as the opening one, so that both arms have the same OPL. This, again, is necessary due to the short coherence time of photons, but introduces another problem: compared to those of Table 4.2, the final states are polarized symmetrically with respect to the D axis after the first interferometer and to the H axis in the second. We took this into account in the design of the second interferometer, which is a duplicate of the first one but is rotated by $\frac{\pi}{4}$ *rad* along the propagation direction, so as to distinguish between $|D\rangle$ and $|A\rangle$ polarizations. At the end of it, we added a HWP which not only provides the needed correction, but also chooses the postselected state that passes through the final polarizer and reaches the detector.

In order to implement the DRDD protocol, that does not need two ancillae, this last HWP can be oriented so as to disregard the contribution of one of the two paths in the second interferometer, that thus acts only as a glorified postselecting polarizer.

By considering all the useful combinations of measurement outcomes in the two pointers and possible choices of $|a_j\rangle$ and $|a_k\rangle$, we can measure photon counts that are proportional to the desired joint probabilities that appear in the mean values of equations (3.6), (3.7) and (3.9) (and include factor $\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho})$). For instance if $\langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle^{a_k}$ is needed to find element HV of the density operator, four acquisitions must take place, one for each projector of the spectral decomposition of the observable $\hat{\sigma}_{yA}\hat{\sigma}_{yB}$. According to Table 4.3, the first interferometer is set to transmit a phase delay of $\frac{3\pi}{2}$ to the $|H\rangle$ polarization, and the second does the same to the $|D\rangle$ polarization, so that after postselection on $|V\rangle$ count $N_{y(+1),y(+1),HV}$ is recorded. Similar combinations of phases are set for the other three counts. Then

$$\langle\hat{\sigma}_{yA}\hat{\sigma}_{yB}\rangle' \propto \sum_{a,b=\pm1} abN_{ya,yb,HV} \tag{4.17}$$

If normalization is needed, it can be carried out by setting the two interferometers at 0 delay and by measuring the $|H\rangle$ and $|V\rangle$ components of the exiting state. The sum of these two counts, corrected because of the unnecessary halving operated by the HWPs, is a measure of the total number of photons that can reach the detectors and is the proportionality constant between data and probabilities.

After the second interferometer, a QWP helps performing the measurements needed by the QST protocol, which was also put to the test. In this case the rest of the setup is set in such a way that the transformation applied to the state is known and can be taken into account during the analysis of the results.

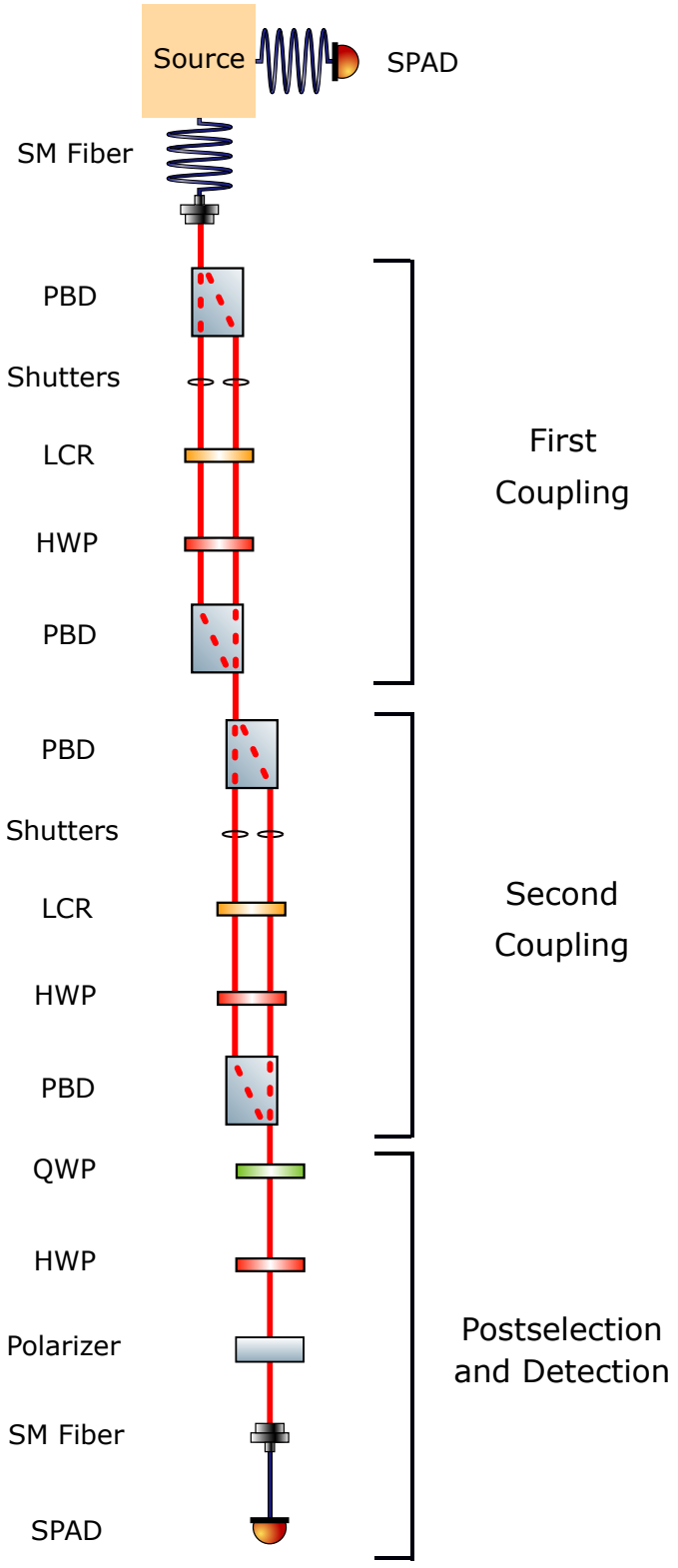


FIGURE 4.9: Scheme of the entire measurement system.

4.2.3 Phase Calibration and Stability

The characterization of the LCRs is a fundamental preliminary phase of the experiment as it allows to achieve fine control over the phase delay between different polarization components. First we directly measured the retardation vs voltage curve using a simple setup: we placed each device between two polarizers and added a rotating HWP. For ease of explanation, suppose that the first polarizer selects $|A\rangle$ and the second $|D\rangle$, then if the slow axis of the LCR is horizontal and the HWP has its fast axis oriented at angle α with respect to it, the power at the detector is

$$P(\alpha) \propto 1 + \cos(\Gamma)\cos(4\alpha) \tag{4.18}$$

where Γ is the phase delay. For different voltage settings we scanned α and fitted this equation to the resulting power curve, thus finding $\cos(\Gamma)$ and ultimately Γ itself.

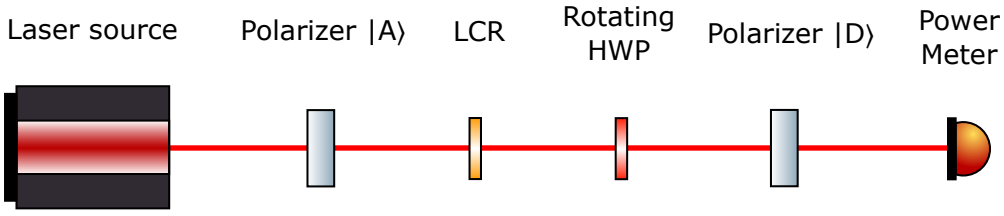


FIGURE 4.10: Scheme of this measurement setup.

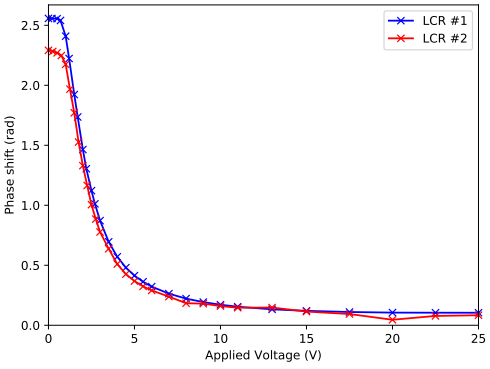


FIGURE 4.11: Retardation curve as a function of voltage. Measured with $\lambda \sim 808\ nm$ laser light.

We assumed Γ to be positive due to the denomination of fast and slow axis (which are indistinguishable with these measurements, therefore we had to rely on the markings on the device), however, a sign error is not particularly relevant as it can be corrected with an arbitrary redefinition of $|H\rangle$ and $|V\rangle$. The fact that neither of the two LCRs can transmit a delay of $\frac{3\pi}{2}$ (Figure 4.11) justified their aforementioned installation on rotators, which doubled their range by switching the roles of the axes.

We also studied the behavior of these devices in the true experimental conditions, that is, inside the interferometers. By sending light into each of them with a known linear polarization and by measuring the same component at the exit, we recorded power (counts per second when using single photons) at different voltage settings. Through parabolic regressions we found the voltages corresponding to maximum and minimum power, which indicate 0 and π *rad* delay respectively. Linear interpolations of the curves around the points near half the maximum power allowed to retrieve the

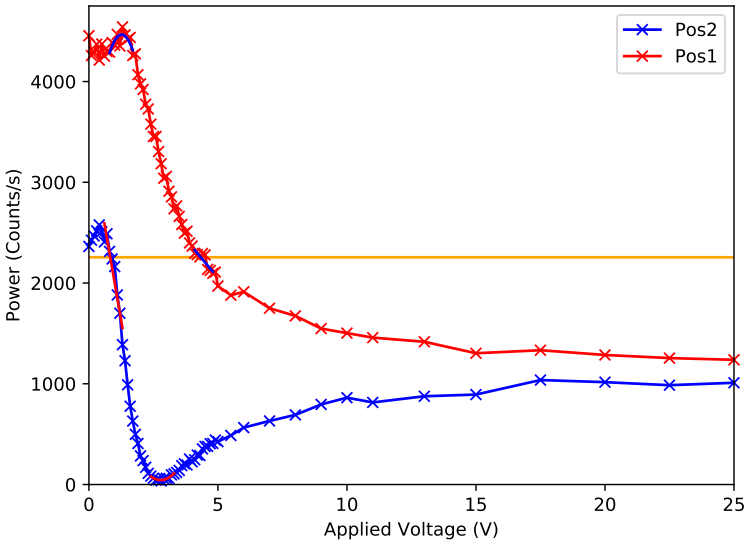


FIGURE 4.12: Power as a function of voltage. The red and blue curves refer to the two different orientations of the optical axis of the LCR (parallel to $|H\rangle$ or $|V\rangle$ in the first interferometer, to $|D\rangle$ or $|A\rangle$ in the second).

voltages needed to transmit $\frac{\pi}{2}$ or $\frac{3\pi}{2}$ *rad*. These values include not only the action of the LCR but also the phase due to OPL imbalances in the rest of the interferometer. Figure 4.12 also shows that visibility exceeds 98% which is optimal for the experiment.

This procedure had to be repeated before each measurement to find the precise value of the voltage, which can change over long periods of time. We studied this phenomenon by acquiring curves like 4.12 in sequence and measuring the variation of phase delay at fixed voltage. Figure 4.13 shows that the interferometer is stable for at least a couple of hours after calibration and its changes are correlated to the temperature of the room. The fact that the sign of the phase variation is the same in the two curves suggests that the culprit is not the LCR but rather the PBDs or more precisely their supports. Indeed the relative inclination of the two PBDs in each interferometer can change over time if temperature variations cause their mounts to move. Due to the macroscopic length of the birefringent material, even a small angle can induce relevant changes in the geometric path and therefore in the phase delay. This has justified the use of precision tilters (Thorlabs Polaris K1S5) which have demonstrated better stability compared to our initial trials with cheaper models.

In conclusion, provided the appropriate mounts are used and the room temperature does not change too much, this setup guarantees enough stability for the experiment.

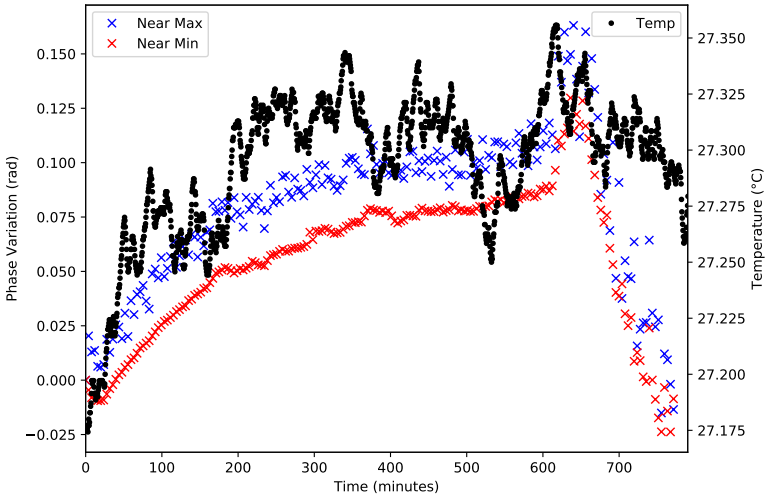


FIGURE 4.13: Variation of phase delay over time, relative to arbitrary initial 0. The black curve is a concurrent measure of temperature.

4.2.4 Analysis of the Statistical Error

In order to provide an estimate for the statistical error to attribute to each count, we took samples of repeated acquisitions in the same conditions and performed a simple analysis to extract mean and standard deviation. We used coincident counts across the two SPADs as in the actual experiment and covered a wide range of exposure times. We did this in four different situations, one in which both interferometers were *inactive* ("II"), meaning that one arm was blocked so that light could pass through the other but without any interference effect, one in which both were *active* ("AA"), meaning that interference took place after both of them, and two in which one was active and one was not ("IA" and "AI, where the order of symbols is that of the two interferometers).

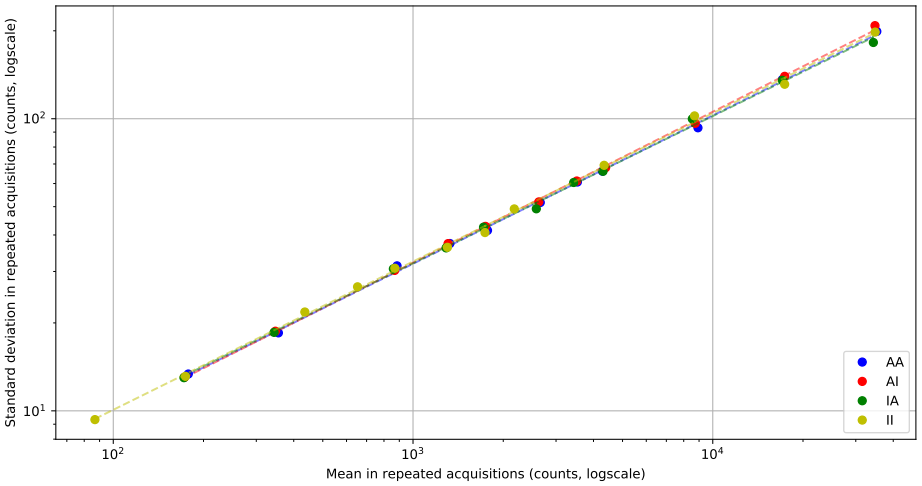


FIGURE 4.14: Relation between standard deviation and mean with linear fit of the logarithmic curve.

Figure 4.14 shows that the poissonian law is respected in all configurations. We verified this through linear regressions on the logarithmic curves which provided slopes that are very close to $\frac{1}{2}$ ($\min = 0.503 \pm 0.007$, $\max = 0.517 \pm 0.004$). In other words, the observed standard deviation is well approximated by the square root of the average.

4.3 Weakening the Coupling

In order to evaluate the protocols at different degrees of strength, we made a few modifications to the measurement setup. The most fundamental one was the addition of a HWP in one of the two arms of each interferometer: this new piece takes the role of implementing the coupling between object and ancilla, which is no longer properly the path taken by the photons.

Remaining true to the theoretical procedure, one of the two polarization components $|H\rangle$ or $|V\rangle$ is selected as $|a_j\rangle$ by the first PBD and encounters the subsequent HWP. Here it gets rotated by angle θ_A which represents the strength: when $\theta_A = \frac{\pi}{2}$, $|a_j\rangle$ becomes $|a_{j\perp}\rangle$ so that both arms are polarized in the same way. It may be said that inside the interferometer, the polarization acts as the pointer and the separation in two paths is necessary only because it allows to interact selectively with one component.

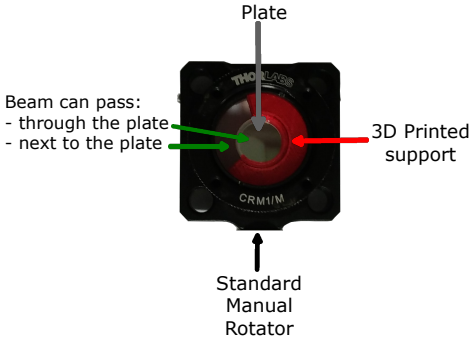


FIGURE 4.15: Photograph of the custom mount.

The HWP is about 2 mm thick and therefore introduces a difference in OPL that is much greater than the coherence length of our photons. In order to compensate, we also added a plate to the other arm, which was oriented in a way that did not change the incoming polarization. These two optics were mounted on custom 3D printed supports (Figure 4.15) designed by Luca Calderaro specifically for this experiment that allowed one beam to pass unperturbed next to the plate and could be placed inside standard 1″ manual rotators (Thorlabs CRM1). We used the Michelson interferometer of Subsection 4.1.4 to measure that the remaining OPL imbalance was approximately 10 μm, which we could correct by tilting the second PBD. Another HWP was inserted at the beginning of the setup to switch $|a_j\rangle$ between $|H\rangle$ and $|V\rangle$ without physically moving pieces in the two arms.

The ancilla measurements are performed by a rotating QWP and HWP placed inside the interferometer that are hit by both paths. At the exit, a LCR helps to tune the relative phase between the arms so that it is maintained at a known value during the entire experiment.

The second interferometer is a twin of the first, except for the fact that it is rotated by $\frac{\pi}{4}$ along the propagation direction and does not have any HWP before the PBD as $|b_0\rangle$ is always $|D\rangle$ and does not have to be switched to $|A\rangle$. Postselection is implemented by a HWP and a PBS just before the detector.

With an appropriate set of orientations of the plates, we were able to apply the necessary pointer projections as prescribed by the protocols. Again, the second interferometer was used as a polarizer (i.e. it had only one active arm) for the DRDD scheme, whereas QST was realized using the waveplates of the first interferometer at maximum strength.

$\|a_j\rangle$	Measurement	Outcome	QWP	HWP	Block
$\|H\rangle$	$\hat{\sigma}_x$	1	$\frac{3\pi}{4}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_x$	-1	$\frac{\pi}{4}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_y$	1	$\frac{\pi}{2}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_y$	-1	$\frac{\pi}{2}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\Pi}_1$	1	$0\ rad$	$0\ rad$	$\|0_A\rangle$
$\|H\rangle$	$\hat{\Pi}_1$	0	$0\ rad$	$0\ rad$	$\|1_A\rangle$
$\|V\rangle$	$\hat{\sigma}_x$	1	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_x$	-1	$\frac{3\pi}{4}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_y$	1	$\frac{\pi}{2}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_y$	-1	$\frac{\pi}{2}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\Pi}_1$	1	$0\ rad$	$0\ rad$	$\|0_A\rangle$
$\|V\rangle$	$\hat{\Pi}_1$	0	$0\ rad$	$0\ rad$	$\|1_A\rangle$

TABLE 4.4: Positions of the fast axes of the plates that carry out the measurements in the first interferometer. All angles are measured with respect to the $|H\rangle$ axis. These values can change if the relative phase between the paths is set differently.

$\|a_j\rangle$	Measurement	Outcome	QWP	HWP	Block
$\|H\rangle$	$\hat{\sigma}_x$	1	$\frac{\pi}{2}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_x$	-1	$0\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_y$	1	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\sigma}_y$	-1	$\frac{\pi}{4}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|H\rangle$	$\hat{\Pi}_1$	1	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{4}\ rad$	$\|0_A\rangle$
$\|H\rangle$	$\hat{\Pi}_1$	0	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{4}\ rad$	$\|1_A\rangle$
$\|V\rangle$	$\hat{\sigma}_x$	1	$0\ rad$	$\frac{\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_x$	-1	$\frac{\pi}{2}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_y$	1	$\frac{\pi}{4}\ rad$	$\frac{\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\sigma}_y$	-1	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{8}\ rad$	/
$\|V\rangle$	$\hat{\Pi}_1$	1	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{4}\ rad$	$\|0_A\rangle$
$\|V\rangle$	$\hat{\Pi}_1$	0	$\frac{\pi}{4}\ rad$	$\frac{3\pi}{4}\ rad$	$\|1_A\rangle$

TABLE 4.5: Positions of the fast axes of the plates that carry out the measurements in the second interferometer. All angles are measured with respect to the $|H\rangle$ axis. These values can change if the relative phase between the paths is set differently.

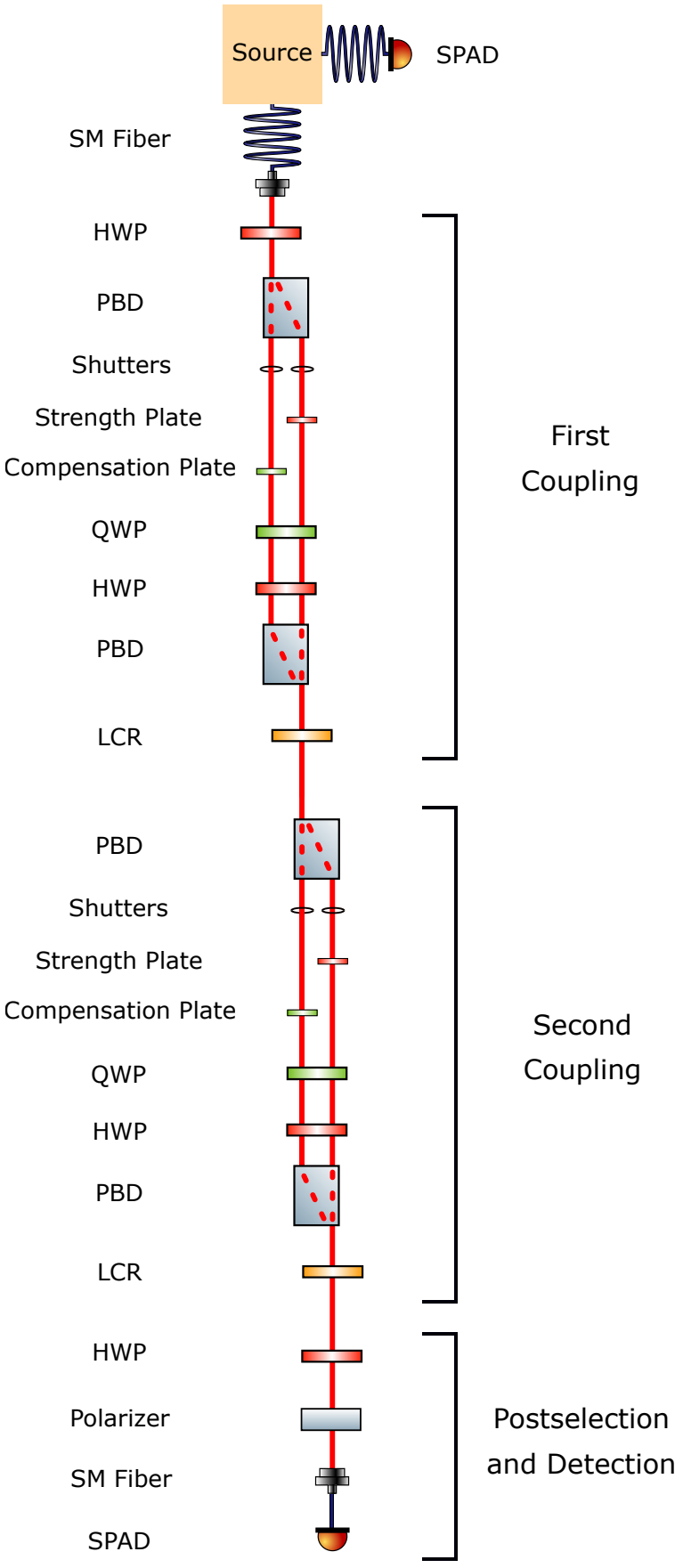


FIGURE 4.16: Scheme of the entire measurement system.

Chapter 5

Results

The main experiment consisted of two phases: the first was aimed at verifying the validity of our protocols at full strength and used the setup of Section 4.2, whereas the second focused on the role of the coupling strength and followed the scheme of Section 4.3. In both cases we tested all the methods at our disposal in order to compare them effectively. We calculated unnormalized density matrices $\hat{\varrho}^U$ from raw counts in the following different ways:

- DRDO protocol. From expression (3.7):

$$\begin{aligned} \varrho^U_{jj} &= \frac{d^2}{\sin^2\theta_A \sin^2\theta_B} \qquad N_{z(-1),z(-1),jk} \qquad \forall k \\ \Re(\varrho^U_{jk}) &= -\frac{d}{2\sin\theta_A \sin\theta_B} \qquad \sum_{a,b=\pm1} ab N_{ya,yb,jk} \qquad j \neq k \\ \Im(\varrho^U_{jk}) &= \frac{d}{2\sin\theta_A \sin\theta_B} \qquad \sum_{a,b} ab N_{xa,yb,jk} \qquad j \neq k \end{aligned} \tag{5.1}$$

  Despite it being unnecessary to measure both non-diagonal elements separately, we chose to do it as well in order to prove the ability of this scheme to pick any matrix element individually. Although these formulas are exact, $\hat{\varrho}^U$ is not guaranteed to be hermitian due to experimental errors. Its diagonal elements and trace are, however, necessarily real. This scheme employs 18 different measurements.

- DRDO Lundeen protocol. From expression (2.33):

$$\begin{aligned} \Re(\varrho^U_{jk}) &= \frac{d}{4\sin(\theta_A)\sin(\theta_B)} \sum_{a,b} ab(N_{xa,xb,jk} - N_{ya,yb,jk}) \\ \Im(\varrho^U_{jk}) &= \frac{d}{4\sin(\theta_A)\sin(\theta_B)} \sum_{a,b} ab(N_{ya,xb,jk} + N_{xa,yb,jk}) \end{aligned} \tag{5.2}$$

  We tested this method even when the approximations on which it is based are not valid (i.e. when the coupling is not weak) to better understand its limitations. Just as above, $\hat{\varrho}^U$ is not guaranteed to be hermitian but this time the diagonal elements can be complex. Even in absence of experimental errors, hermiticity is not a necessity unless $\theta_{A,B}$ is small. This scheme employs 64 different measurements.

- DRDO Lundeen protocol with corrections. From expression (3.6):

$$\begin{aligned}\Re(\varrho^U_{jk}) = \frac{d}{4\sin\theta_A\sin\theta_B}\Bigg(&\sum_{a,b} ab(N_{xa,xb,jk} - N_{ya,yb,jk}) \\ &+ 2\tan\frac{\theta_B}{2}\sum_a aN_{xa,z(-1),jk} + 2\tan\frac{\theta_A}{2}\sum_b bN_{z(-1),xb,jk} \\ &+ 4\tan\frac{\theta_A}{2}\tan\frac{\theta_B}{2}N_{z(-1),z(-1),jk}\Bigg) \\ \Im(\varrho^U_{jk}) = \frac{d}{4\sin\theta_A\sin\theta_B}&\left(\sum_{a,b} ab(N_{ya,xb,jk} + N_{xa,yb,jk}) + 2\tan\frac{\theta_B}{2}\sum_a aN_{ya,z(-1),jk}\right)\end{aligned} \tag{5.3}$$

  Again this matrix is not necessarily hermitian, although it would be in absence of experimental errors and in any conditions of strength. This scheme employs 92 different measurements.

- DRDD protocol. From expressions (3.9) and (3.11):

$$\begin{aligned}\rho^U_{A1}(jl) &= \frac{1}{2\sin(\theta_A)}\sum_a a(N_{xa,jl} + iN_{ya,jl}) \\ \rho^U_{A2}(jl) &= \frac{1}{\sin(\theta_A)}N_{z(-1),jl} \\ \varrho^U_{jk} &= d\tan\left(\frac{\theta_A}{2}\right)\delta_{jk}\rho^U_{A2}(jl) + \sum_l e^{\frac{2\pi il(j-k)}{d}}\rho^U_{A1}(jl)\end{aligned} \tag{5.4}$$

  No guarantee of hermiticity in presence of experimental errors. This scheme employs 20 different measurements.

- DRDD Lundeen protocol. From expressions (2.37) and (2.35):

$$\begin{aligned}\rho^U_{A1}(jl) &= \frac{1}{2\sin(\theta_A)}\sum_a a(N_{xa,jl} + iN_{ya,jl}) \\ \varrho^U_{jk} &= \sum_l e^{\frac{2\pi il(j-k)}{d}}\rho^U_{A1}(jl)\end{aligned} \tag{5.5}$$

  No guarantee of hermiticity in presence of experimental errors or in their absence if θ_A is not small enough. This scheme employs 16 different measurements.

- QST. Due to the small dimension of the system we could easily perform a standard tomography using the projectors on $|H\rangle$, $|V\rangle$, $|D\rangle$ and $|R\rangle$. From the corresponding counts we found

$$\begin{aligned}\varrho^U_{HH} &= N_H \\ \varrho^U_{VV} &= N_V \\ \Re(\varrho^U_{HV}) = \Re(\varrho^U_{VH}) &= N_D - \frac{1}{2}(N_H + N_V) \\ \Im(\varrho^U_{HV}) = -\Im(\varrho^U_{VH}) &= N_R - \frac{1}{2}(N_H + N_V)\end{aligned} \tag{5.6}$$

  This matrix is hermitian and is measured with only 4 acquisitions.

Because many of the needed projections are shared between the protocols, the total tally of acquisitions is 116. The reconstruction of a state took approximately 17 minutes in the first experiment and 34 in the second, an extension mostly due to the increased number of rotations of the plates. In both cases the exposure time was set at 2 seconds so that the value of all counts was in the thousands (except for those proportional to a $\langle \hat{\Pi}_{1A,B} \rangle$ which were much smaller in case of low coupling strength).

Although it would have been possible to normalize each matrix element for example using $N_H + N_V$ as the constant, we followed the steps that we imagine would be chosen in a tomography experiment: we took the hermitian part of each matrix and normalized it with its trace:

$$\hat{\varrho}^U \quad \rightarrow \quad \hat{\varrho}^H = \frac{\hat{\varrho}^U + \hat{\varrho}^{U\dagger}}{2} \quad \rightarrow \quad \hat{\varrho}^N = \frac{\hat{\varrho}^H}{\mathrm{Tr}(\hat{\varrho}^H)} \tag{5.7}$$

The resulting matrices are all hermitian and all have unit trace, although they are not necessarily positive semidefinite and therefore may not be physical density operators. However, in order to compare the methods in fair conditions, we did not add a likelihood maximization step and used $\hat{\varrho}^N$ as our starting points in the analysis I shall present case by case in the next sections.

5.1 Results at Maximum Strength

Our source is capable of sending to the measurement apparatus both mixed and (almost) pure states. In particular, the rotating HWP before the Sagnac interferometer can change the balance of pump light in the two arms so that if one of them is completely inactive, the entirety of the photons reaching one exit is $|H\rangle$ polarized while those that arrive at the other are $|V\rangle$ polarized. This means that despite the uncontrollable action of the two fibers, the state entering the measurement system is approximately pure. On the other hand, when the crystal is hit by both arms, each fiber will collect both $|H\rangle$ and $|V\rangle$ photons, so that the state at its exit is mixed (completely mixed when the two polarizations are perfectly balanced inside the fiber).

We used this technique to generate states with different degrees of purity, which we measured as the trace of the square of the matrix resulting from the QST method. The data confirm that we were able to produce a state very close to being completely mixed, indeed its purity is 0.501 ± 0.003, while the theoretical limit is $\mathrm{Tr}(\hat{\varrho}^2) = \frac{1}{d} = 0.5$. The maximum purity we observed is only 0.936 ± 0.006 due to the fact that even with the help of a QWP before the Sagnac interferometer, the pump light entering it was not linearly polarized. Moreover, imperfections in the PBS caused imbalanced extinction ratios in the two arms which explain the different heights of the two peaks in Figure 5.1.

Since the polarization of the pump light after the fiber did not change during the experiment, the measured states lie on a straight line in the Bloch sphere, as is shown in Figure 5.2 which plots the QST results.

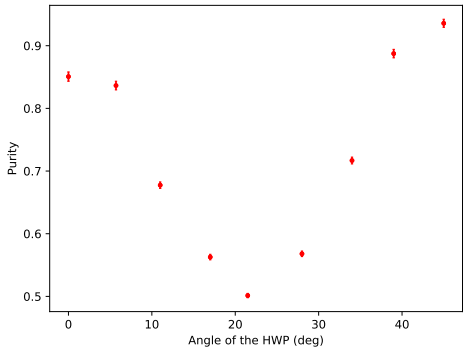


FIGURE 5.1: Purity as $\mathrm{Tr}((\hat{\varrho}^N_{QST})^2)$ as a function of the angle of the rotating HWP placed before the Sagnac interferometer, measured in respect of an arbitrary 0.

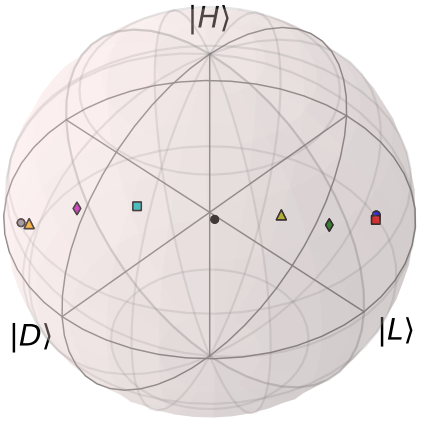
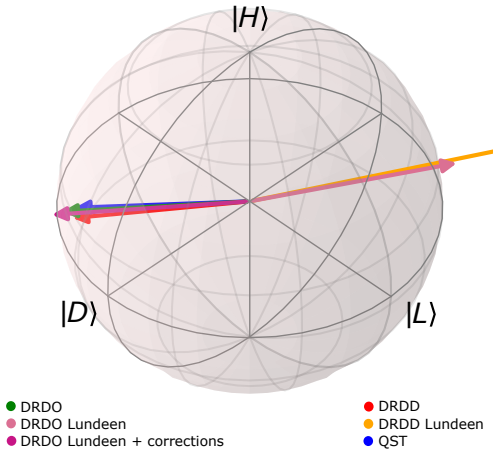
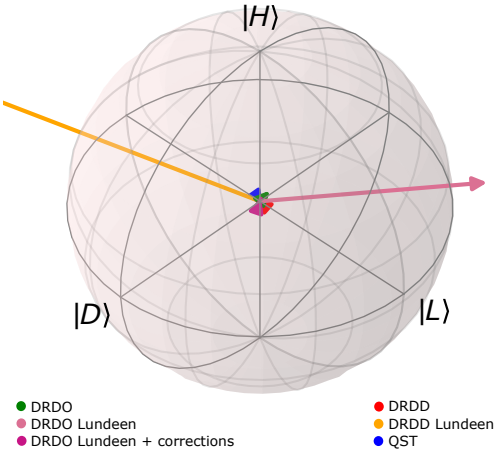


FIGURE 5.2: Visualization of the QST results on the Bloch sphere.

Then we calculated $\hat{\varrho}^N$ using the previously introduced formulas from the various methods and $\theta_A = \theta_B = \frac{\pi}{2}$. A first comparison can be done by simply observing that the estimated density operators are similar. From Figure 5.3 it is clear that all the measurement protocols are in agreement, only the methods that are based on invalid approximations (DRDD Lundeen and DRDO Lundeen) are off the mark.



(A) Purest state.



(B) Most mixed state.

FIGURE 5.3: Visualization of two measured states. Each arrow corresponds to the result of one of the protocols.

To express this in more quantitative terms, we measured the trace distance between the result of each scheme and that of the QST, treated as reference.

$$t(\hat{\varrho}^N, \hat{\varrho}^N_{QST}) = \frac{1}{2}\,\mathrm{Tr}\,\sqrt{(\hat{\varrho}^N - \hat{\varrho}^N_{QST})^\dagger(\hat{\varrho}^N - \hat{\varrho}^N_{QST})} = \frac{1}{2}\sqrt{\sum_{i=x,y,z}(r_i - r_{i,QST})^2} \tag{5.8}$$

in which r_i are the coefficients of the Bloch vector representing the density matrix.

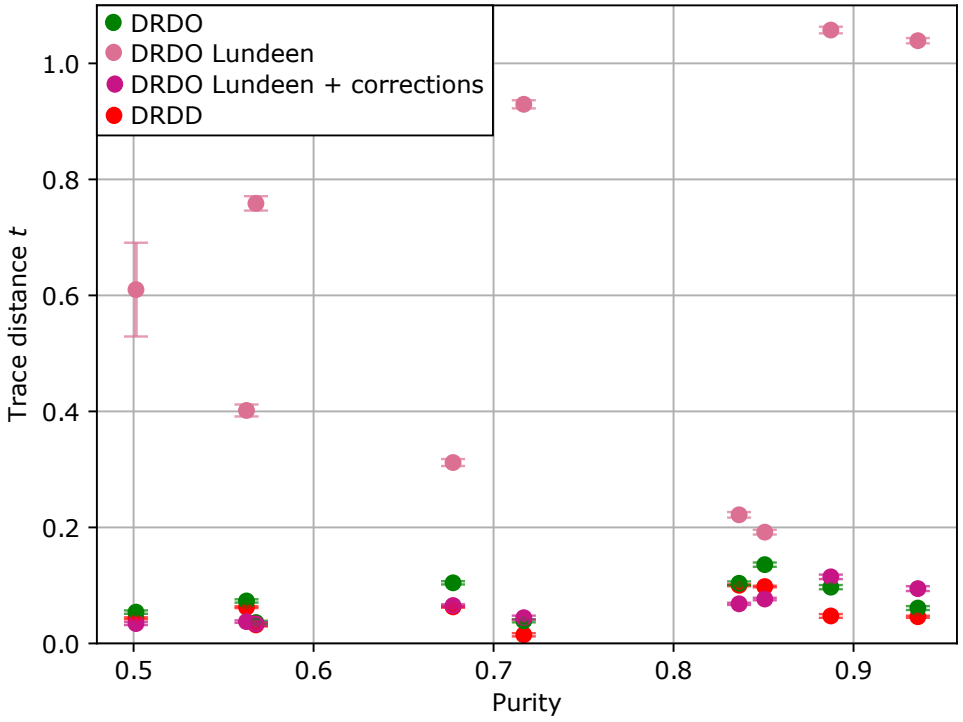
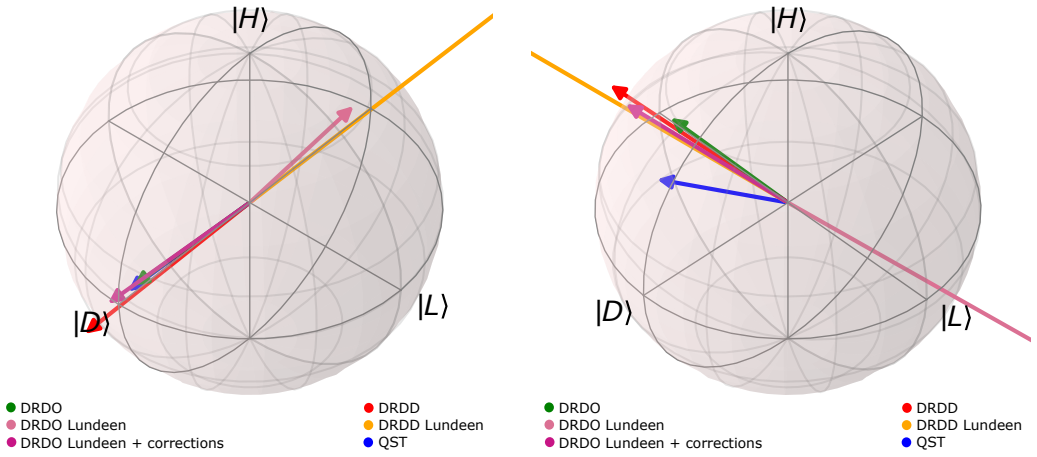


FIGURE 5.4: Trace distances between each method and QST. Error bars are calculated using a simulation of the experiment in which 10^4 samples were generated from a Poissonian distribution with the actual measurement results as mean (Appendix A.5.2).

Figure 5.4 shows that for our methods $t < 0.15$, while the DRDO Lundeen protocol always fails to produce accurate results. The trace distance relative to the DRDD Lundeen protocol is not plotted as it is always greater than 40. This is because at maximum strength all the information about the diagonal elements is recorded in ρ_{A2}, therefore an estimate that only considers ρ_{A1} necessarily returns values close to 0. When normalization occurs, non-diagonal elements skyrocket and the matrix is non-physical. It is important to remember than if $t > 1$, the matrix is necessarily out of the Bloch sphere and is often completely unrelated to the true state, indicating that the method has utterly failed to provide any meaningful result.

In order to prove the validity of the protocols at higher purities, we added a polarizer, a QWP and a HWP at the beginning of the measurement apparatus so as to prepare the six states $|A\rangle$, $|D\rangle$, $|H\rangle$, $|L\rangle$, $|R\rangle$, $|V\rangle$. All the purities obtained from the QST matrices lie in interval $[0.94, 0.99]$. This time we calculated the trace distances with respect to the theoretical density operators, and again all the methods that are expected to succeed report acceptable values, while Lundeen's protocols generally fare worse, with $t > 0.5$ for

the DRDO and $t > 5$ for the DRDD protocol. For this reason Figure 5.6 does not show any of the DRDD Lundeen results nor two of the DRDO ones which lie well outside the Bloch sphere.



(A) A linear polarization. (B) A circular polarization.

FIGURE 5.5: Visualization of two measured states. Each arrow corresponds to the result of one of the protocols.

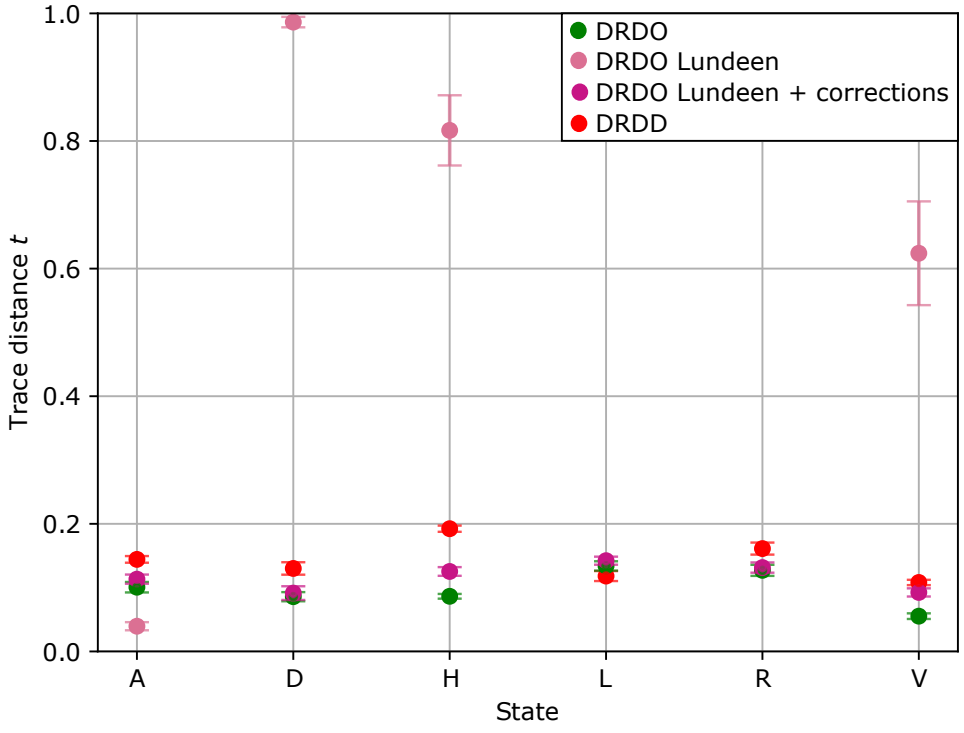


FIGURE 5.6: Trace distances between each method and the theoretical density matrix. Two points of the DRDO Lundeen protocol are not shown because they are above 5.

These results prove that the DRDO protocol is capable of measuring accurate density operators using maximally strong couplings. Trace distances are generally satisfactory regardless of the purity of the states. The fact that they are often incompatible with 0 is due to the complexity of the apparatus, which, despite the fine alignment, can still introduce systematic errors. It is also clear that the schemes that are based on weak measurements are not functional with strong ones, especially in the DRDD case.

5.2 Results for Variable Strength

After applying to the measurement setup the modifications described in Section 4.3, we proceeded with the second part of the experiment. Due to the increased number of rotations (both manual and motorized) required by the implementation, we evaluated only two states, one mixed and one (almost) pure, each with five different values of strength $\theta \equiv \theta_A = \theta_B$. We measured these values using the angles of the rotators that control the HWPs which take care of the coupling, or more precisely twice the difference between such angles and those corresponding to orientations of the plates that do not change the polarization of the incoming light. Although inaccurate, this estimate is surely sufficient to discriminate the trials between strong and weak regime.

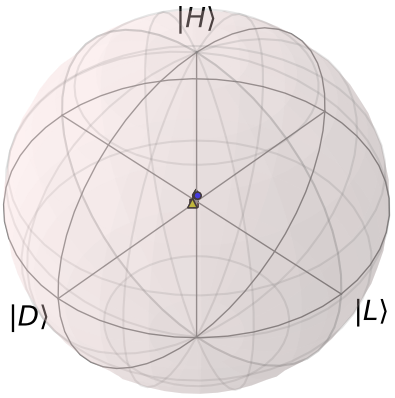
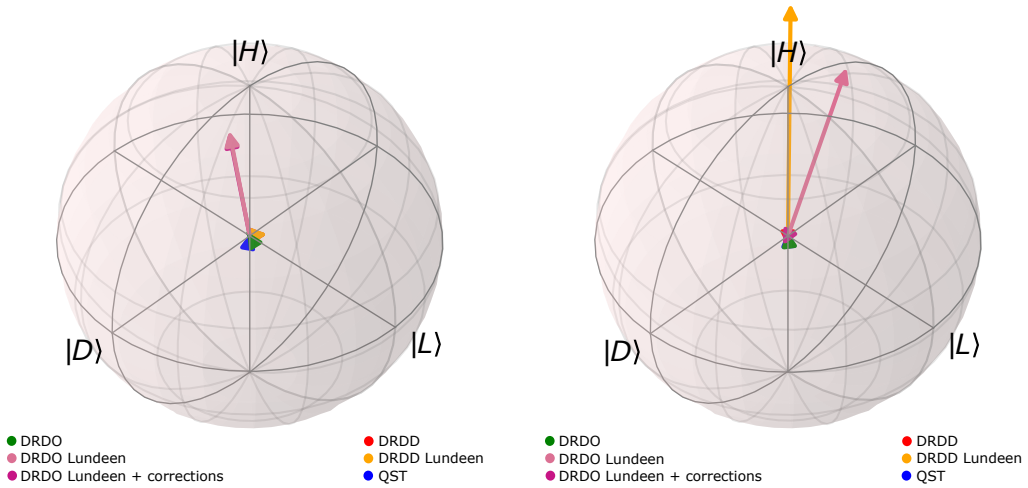


FIGURE 5.7: The five QST results for the mixed state.

In order to produce the mixed state we made sure that the two arms of the Sagnac interferometer of the source were approximately balanced in terms of pump power. Figure 5.7 shows the results of the QST, which we repeated after each trial although it is independent of the strength. This visualization confirms that during the entire experiment the incoming polarization did not change, with the highest trace distance between the resulting matrices being only 0.02. Moreover the state stayed close to being maximally mixed, with purities always smaller than 0.51.

Figure 5.8b again confirms that the methods that are based on the weak approximation do not work properly at higher values of strength. Figure 5.8a instead shows that for small θ the DRDD Lundeen scheme can produce accurate results. It is interesting to observe that the DRDO Lundeen protocol, with or without corrections, gives a higher trace distance than expected (Figure 5.9). We ascribe this to slight misalignments in the setup which can introduce systematic errors in the counts that, however small, become important at low strength (as explained in Subsection 3.3.2).



(A) Measurements with $\theta \approx 0.07$ *rad*. (B) Measurements with $\theta \approx \frac{\pi}{2}$ *rad*.

FIGURE 5.8: Visualization of the results of two measurements of the mixed state. Each arrow corresponds to one of the protocols.

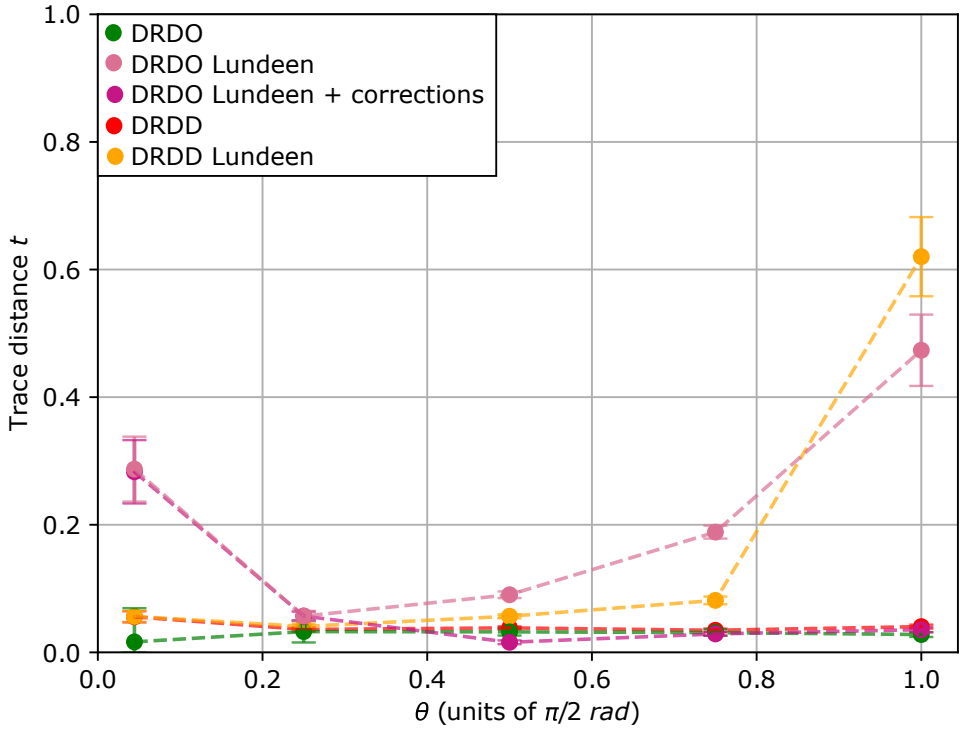


FIGURE 5.9: Trace distances between each method and QST.

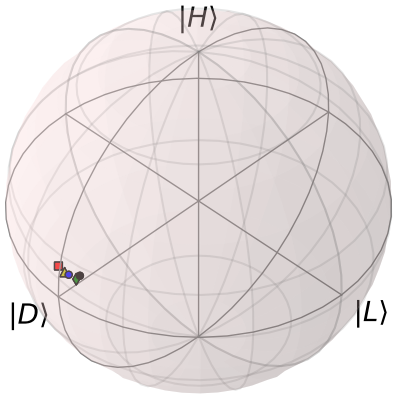
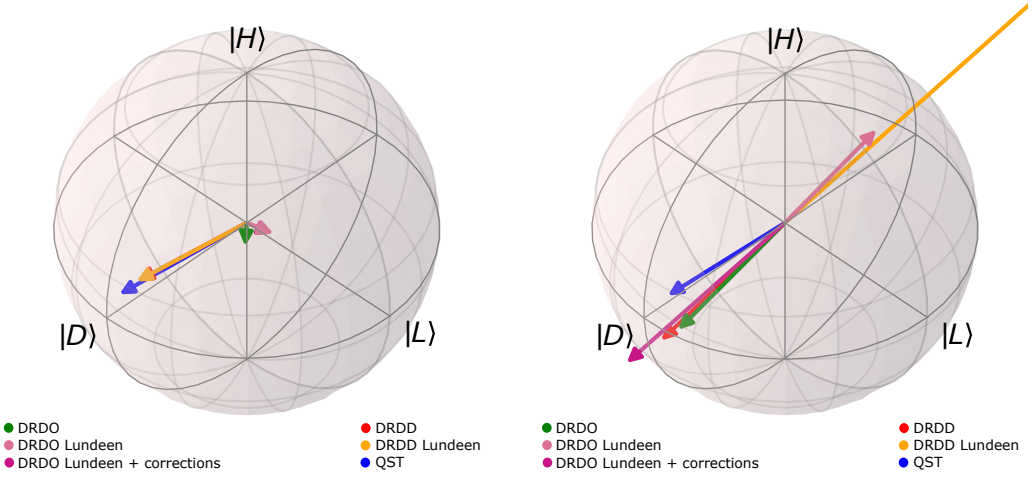


FIGURE 5.10: The five QST results for the pure state.

These problems become more evident in the case of the pure state, which we generated through the insertion of a linear polarizer at the beginning of the measurement setup. The QST results are close to each other (the maximum mutual trace distance is 0.06) but do not coincide with the expected state $|D\rangle$. Indeed a small rotation towards the $|H\rangle$ pole is noticeable in Figure 5.10, which suggests again the presence of imperfections in the setup.

The conclusions we can draw from these results are similar to the previous, although the errors of Lundeen's schemes are more relevant. Moreover, the vertical bars of the green plot of Figure 5.9 show the higher relative statistical uncertainties of the DRDO scheme, which are due to small factor $\sin^4(\theta)$ at low strength. The differences between Figures 5.12 and 5.9 also highlight the fact that the quality of the results can vary with the state. However it is clear that the choice of operating in weak conditions is generally inconvenient even if the appropriate measurement protocols are used.



(A) Measurements with $\theta \approx 0.07$ *rad*.

(B) Measurements with $\theta \approx \frac{\pi}{2}$ *rad*.

FIGURE 5.11: Visualization of the results of two measurements of the pure state. Each arrow corresponds to one of the protocols.

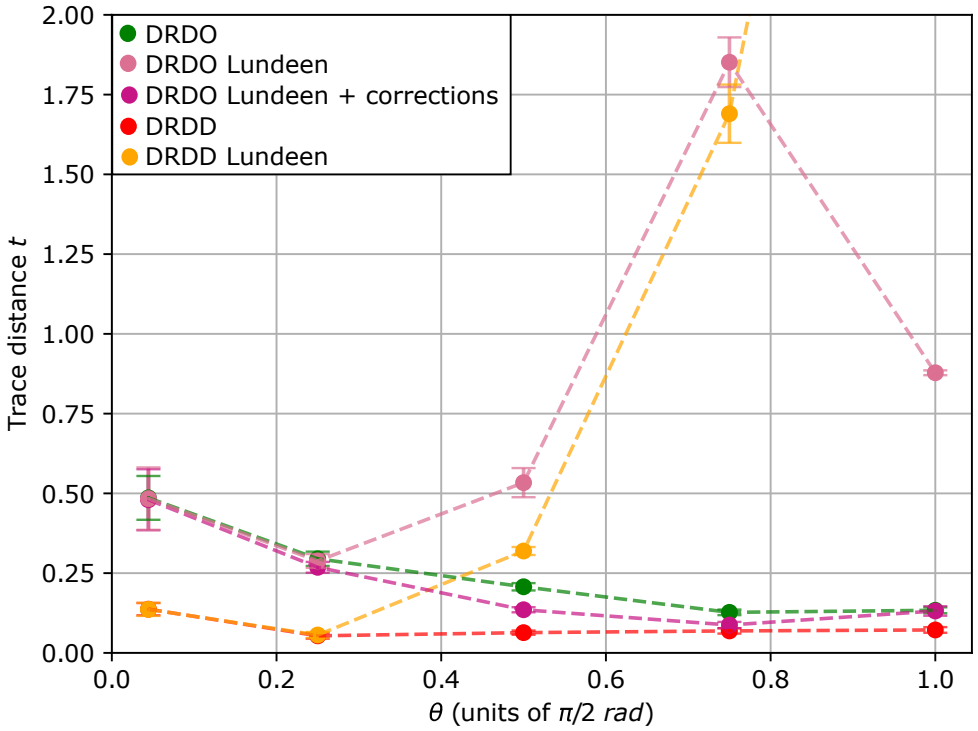


FIGURE 5.12: Trace distances between each method and the expected matrix $|D\rangle\langle D|$. The point of the DRDD Lundeen protocol at maximum strength is not plotted because it is too far apart from the others (4.9 ± 0.3).

Chapter 6

Conclusions

After the detailed study of the fundamentals of quantum mechanics and measurement theory, this thesis has investigated the new possibilities that are offered by weak values especially in the field of quantum state reconstruction (Chapter 2, Reference [3]). Despite the revolutionary contribution of the work of Aharonov, Albert and Vaidman [20] to quantum sensors and fundamental research (Section 2.3, References [26, 38, 39]), we recognized that weak measurements are not strictly necessary for efficient tomography protocols and therefore we evolved Lundeen's proposals [4] to exact formulas that do not require the weak approximation (Section 3.2).

We compared the results with the aid of simulations (Section 3.3) and proved that the use of strong couplings allows to obtain more precise estimates with lesser statistics, because of the $\sin^{-2}(\theta)$ law followed by the error, with θ representing the strength. The measured matrix elements are also more resistant to experimental systematic biases because they are extracted from linear combinations of counts that are proportional to $\sin^2(\theta)$ and hence are more vulnerable to small errors in the data when $\theta \to 0$. We also argued that our methods may be preferable to standard QST when the dimension d of the system is large. This is because only $d+1$ different projectors are needed (instead of d^2) and all but one of them correspond to states in the measurement basis (Section 3.4).

The bulk of this work has been the realization of a measurement setup that can apply all these reconstruction strategies to the polarization state of light. We employed an entangled photon source (Section 4.1) to work in true quantum regime and easily generate states of arbitrary purity. We first optimized the apparatus for maximum coupling strength (Section 4.2) and then evolved it to investigate the weak regime (Section 4.3). The experimental results confirm the validity of our proposals and their ability to produce estimates that are close to those of QST (trace distance $\lesssim 0.15$). We also showed the predictable failure of methods based on the weak approximation to be accurate at high values of strength (Figure 5.4, for $\theta = \frac{\pi}{2}$ results are mostly non-physical) and the vulnerability of most protocols to experimental errors at low θ (Figure 5.12).

In conclusion, this experiment can demonstrate that our evolutions to direct state reconstruction schemes are functional and represent a significant step forward in terms of quality of the results. We believe that equivalent transitions from weak to strong measurement regime might provide similar improvements also in other fields of application of the weak value, such as the study of incompatible observables (Section 3.5).

Appendix A

Proof of the Theoretical Results

This appendix has the purpose of proving the most important relations that are used in the main text. Although some of these results are already known, it is not easy to find an exhaustive and detailed derivation of all of them, therefore I hope the reader will find the following dissertation an useful supplement to the reading of Chapters 2 and 3.

The next section will focus on equation (2.9) that shows how to extract the weak values from pointer averages, whereas the subsequent one will derive the properties of weak measurements of products of more observables, thus justifying formula (2.31). Section A.3 will instead delve into mostly uncharted territory, as it will demonstrate the fundamental relations of the new measurement protocols that are the backbone of this thesis, while also confirming Lundeen's results in the weak limit. In Section A.4, I will show how I extended the procedure for the measurement of products of incompatible observables using the ancilla scheme to the case of arbitrary strength, thus proving equations (3.27) and (3.29). Finally, Section A.5 will provide some details on the simulations that were used in this work.

A.1 Measuring the Weak Value

Consider the situation described in Section 2.1. Suppose $|\phi_S\rangle$ to be the initial state of the object system and to prepare ad hoc a continuous uni-dimensional ancilla in which position and momentum operators $\hat{X}_A$ and $\hat{P}_A$ are naturally defined. It is convenient to choose its initial state $|I_A\rangle$ so that its representations in the x_A and p_A bases are real and have zero mean (see Reference [21] for other cases). An example is the gaussian state

$$I_A(x_A) \equiv \langle x_A|I_A\rangle = (\sqrt{2\pi}\Delta_x)^{-\frac{1}{2}} e^{-\frac{x^2}{4\Delta_x^2}} \tag{A.1}$$

Observable $\hat{S}$ of the object space is coupled to momentum $\hat{P}_A$ of the pointer with an interaction Hamiltonian in the form

$$\hat{H}_{int}(t) = g(t)\hat{S} \otimes \hat{P}_A \tag{A.2}$$

so that the joint evolution operator is

$$\hat{U} = e^{-\frac{i}{\hbar}\int \hat{H}_{int}(t)dt} = e^{-i\frac{g}{\hbar}\hat{S}\otimes\hat{P}_A} \tag{A.3}$$

It is easy to see that $\hat{U}$ defines a translation of the pointer the strength of which depends on g and on the action of $\hat{S}$ on $|\phi_S\rangle$. For example if $|\phi_S\rangle = \sum_j c_j|s_j\rangle$ where $c_j \in \mathbb{C}$ and $|s_j\rangle$ are the eigenstates of $\hat{S}$, the initial joint state $|\psi\rangle = |\phi_S\rangle \otimes |I_A\rangle$ evolves into

$$\psi'(x_A) \equiv \langle x_A|\hat{U}|\psi\rangle = \sum_j c_j \langle x_A|e^{-i\frac{g}{\hbar}s_j\hat{P}_A}\left(|s_j\rangle \otimes |I_A\rangle\right) = \sum_j c_j I_A(x_A - gs_j)|s_j\rangle \quad \text{(A.4)}$$

On this state:

$$\begin{aligned}
\langle \hat{X}_A^{(g)}\rangle_{\psi'} &= \int dx_A \sum_j \sum_k I_A(x_A - gs_j) I_A(x_A - gs_k) x_A \langle s_j|s_k\rangle c_j^\star c_k \\
&= \int dx_A \sum_j \sum_k I_A(x_A - gs_j) I_A(x_A - gs_k) x_A \delta_{jk} c_j^\star c_k \\
&= \int dx_A \sum_j I_A^2(x_A - gs_j) x_A |c_j|^2 \\
&= \sum_j |c_j|^2 (\langle \hat{X}_A\rangle_{I_A} + gs_j) = \sum_j |c_j|^2 gs_j = g\langle \hat{S}\rangle_{\phi_S}
\end{aligned} \quad \text{(A.5)}$$

where $\langle \hat{X}_A\rangle_{I_A} = 0$ has been used. This proves relation (2.4). Moreover

$$\langle (\hat{X}_A^{(g)})^2\rangle_{\psi'} = \int dx_A \sum_j I_A^2(x_A - gs_j) x_A^2 |c_j|^2 = \sum_j |c_j|^2 (\Delta_x^2 + g^2 s_j^2) = g^2 \langle \hat{S}^2\rangle_{\phi_S} + \Delta_x^2 \quad \text{(A.6)}$$

in which the property of gaussian integrals $\int dx_A I_A^2(x_A - x')x^2 = x'^2 + \Delta_x^2$ has been used. Then the variance of the measurement becomes

$$\mathrm{Var}(\hat{X}_A^{(g)})_{\psi'} = \langle (\hat{X}_A^{(g)})^2\rangle_{\psi'} - (\langle \hat{X}_A^{(g)}\rangle_{\psi'})^2 = g^2 \,\mathrm{Var}(\hat{S})_{\phi_S} + \Delta_x^2 \quad \text{(A.7)}$$

Apart from known constant g, the mean value of $\hat{X}_A$ coincides with that of $\hat{S}$, so that this model mimics the axiomatic quantum measurement. However, the variance of the distribution of results is also made larger by that of the pointer Δ_x^2.

If g is so small that $\Delta_x \gg g|s_j - s_k|$, it is possible to approximate the evolution operator as

$$\hat{U} \approx \mathbb{1} - i\frac{g}{\hbar}\hat{S} \otimes \hat{P}_A \quad \text{(A.8)}$$

then

$$|\psi'\rangle = \hat{U}|\psi\rangle \approx |\phi_S\rangle \otimes |I_A\rangle - i\frac{g}{\hbar}\hat{S}|\phi_S\rangle \otimes \hat{P}_A|I_A\rangle \quad \text{(A.9)}$$

On this state, neglecting terms of order higher than 1 in g:

$$\begin{aligned}
\langle \hat{X}_A^{(g)}\rangle_{\psi'} &\approx \langle \phi_S|\phi_S\rangle \otimes \langle \hat{X}_A\rangle_{I_A} + i\frac{g}{\hbar}\langle \hat{S}\rangle_{\phi_S}\langle \hat{P}_A\hat{X}_A - \hat{X}_A\hat{P}_A\rangle_{I_A} = g\langle \hat{S}\rangle_{\phi_S} \\
\langle (\hat{X}_A^{(g)})^2\rangle_{\psi'} &\approx \Delta_x^2 + i\frac{g}{\hbar}\langle \hat{S}\rangle_{\phi_S}\langle \hat{P}_A\hat{X}_A^2 - \hat{X}_A^2\hat{P}_A\rangle_{I_A} = \Delta_x^2 + 2g\langle \hat{S}\rangle_{\phi_S}\langle \hat{X}_A\rangle_{I_A} = \Delta_x^2
\end{aligned} \quad \text{(A.10)}$$

$$\mathrm{Var}(\hat{X}_A^{(g)})_{\psi'} \approx \Delta_x^2 - (g\langle \hat{S}\rangle_{\phi_S})^2 \approx \Delta_x^2$$

which are first order approximations of the previous results.

The weak value naturally arises if the object system is projected onto state $|\Phi_S\rangle$ (post-selection). Due to entanglement the joint state collapses onto

$$|\psi_{post}\rangle = |\Phi_S\rangle\langle\Phi_S|\psi'\rangle \approx |\Phi_S\rangle \otimes \Big(\langle\Phi_S|\phi_S\rangle|I_A\rangle - i\frac{g}{\hbar}\langle\Phi_S|\hat{S}|\phi_S\rangle\hat{P}_A|I_A\rangle\Big) \tag{A.11}$$

the norm of which is

$$|\langle\psi_{post}|\psi_{post}\rangle| = \sqrt{|\langle\Phi_S|\phi_S\rangle|^2 + \frac{g^2|\langle\Phi_S|\hat{S}|\phi_S\rangle|^2}{4\Delta_x^2}} \approx |\langle\Phi_S|\phi_S\rangle| \tag{A.12}$$

if it is imposed that

$$\frac{g^2}{4\Delta_x^2}\left|\frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle}\right|^2 \ll 1 \tag{A.13}$$

Using this normalization and ignoring the global phase of $\langle\Phi_S|\phi_S\rangle$, the pointer state can be written as

$$|F_{\Phi,A}\rangle = \frac{\langle\Phi_S|\psi'\rangle}{\langle\Phi_S|\phi_S\rangle} \approx |I_A\rangle - i\frac{g}{\hbar}\frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle}\hat{P}_A|I_A\rangle \tag{A.14}$$

where the weak value $\langle\hat{S}^W\rangle_\phi^\Phi \equiv \frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle}$ appears. Ancilla observables can be measured on this state, then:

$$\begin{aligned}\langle\hat{X}_A\rangle_{F_{\Phi,A}} &\approx i\frac{g}{\hbar}\Re(\langle\hat{S}^W\rangle_\phi^\Phi)\langle\hat{P}_A\hat{X}_A - \hat{X}_A\hat{P}_A\rangle_{I_A} + \frac{g}{\hbar}\Im(\langle\hat{S}^W\rangle_\phi^\Phi)\langle\hat{P}_A\hat{X}_A + \hat{X}_A\hat{P}_A\rangle_{I_A} \\ &= g\Re(\langle\hat{S}^W\rangle_\phi^\Phi) + \frac{g}{\hbar}\Im(\langle\hat{S}^W\rangle_\phi^\Phi)\langle 2\hat{X}_A\hat{P}_A - i\hbar\rangle_{I_A} \\ &= g\Re(\langle\hat{S}^W\rangle_\phi^\Phi)\end{aligned} \tag{A.15}$$

where $\langle 2\hat{X}_A\hat{P}_A\rangle_{I_A} = i\hbar$ has been used. Moreover:

$$\langle\hat{P}_A\rangle_{F_{\Phi,A}} \approx i\frac{g}{\hbar}\Re(\langle\hat{S}^W\rangle_\phi^\Phi)\langle\hat{P}_A^2 - \hat{P}_A^2\rangle_{I_A} + \frac{g}{\hbar}\Im(\langle\hat{S}^W\rangle_\phi^\Phi)\langle\hat{P}_A^2 + \hat{P}_A^2\rangle_{I_A} = \frac{g\hbar}{2\Delta_x^2}\Im(\langle\hat{S}^W\rangle_\phi^\Phi) \tag{A.16}$$

where $\langle\hat{P}_A^2\rangle_{I_A} = \frac{\hbar^2}{4\Delta_x^2}$ has been used. These two results prove relation (2.9) of the main text. Finally

$$\langle\hat{X}_A^2\rangle_{F_{\Phi,A}} \approx \langle\hat{X}_A^2\rangle_{I_A} = \Delta_x^2 \qquad \langle\hat{P}_A^2\rangle_{F_{\Phi,A}} \approx \langle\hat{P}_A^2\rangle_{I_A} = \frac{\hbar^2}{4\Delta_x^2} \tag{A.17}$$

I now show that similar expression are valid if the ancilla is a qubit. Let the initial pointer state be the eigenstate $|0_A\rangle$ of $\hat{\sigma}_{zA}$ and let the unitary evolution be

$$\hat{U} = e^{-i\theta\hat{S}\otimes\hat{\sigma}_{yA}} \approx \mathbb{1}_S \otimes \mathbb{1}_A - i\theta\hat{S}\otimes\hat{\sigma}_{yA} \tag{A.18}$$

where small angle θ plays the role of g. The joint state after postselection is

$$|\psi_{post}\rangle \approx |\Phi_S\rangle \otimes \Big(\langle\Phi_S|\phi_S\rangle|0_A\rangle + \theta\langle\Phi_S|\hat{S}|\phi_S\rangle|1_A\rangle\Big) \tag{A.19}$$

so that, assuming

$$\theta^2\left|\frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle}\right|^2 \ll 1 \tag{A.20}$$

the pointer state can be expressed as

$$|F_{\Phi,A}\rangle \approx |0_A\rangle + \theta\langle\hat{S}^W\rangle_\phi^\Phi|1_A\rangle \tag{A.21}$$

Therefore

$$\begin{aligned}\langle\hat{\sigma}_{xA}\rangle_{F_{\Phi,A}} &= \langle\hat{\sigma}_{xA}\rangle_{0_A} \\ &+ \theta\Re(\langle\hat{S}^W\rangle_\phi^\Phi)(\langle 1_A|\hat{\sigma}_{xA}|0_A\rangle + \langle 0_A|\hat{\sigma}_{xA}|1_A\rangle) \\ &+ i\theta\Im(\langle\hat{S}^W\rangle_\phi^\Phi)(\langle 0_A|\hat{\sigma}_{xA}|1_A\rangle - \langle 1_A|\hat{\sigma}_{xA}|0_A\rangle) \\ &= 2\theta\Re(\langle\hat{S}^W\rangle_\phi^\Phi)\end{aligned} \tag{A.22}$$

and

$$\begin{aligned}\langle\hat{\sigma}_{yA}\rangle_{F_{\Phi,A}} &= \langle\hat{\sigma}_{yA}\rangle_{0_A} \\ &+ \theta\Re(\langle\hat{S}^W\rangle_\phi^\Phi)(\langle 1_A|\hat{\sigma}_{yA}|0_A\rangle + \langle 0_A|\hat{\sigma}_{yA}|1_A\rangle) \\ &+ i\theta\Im(\langle\hat{S}^W\rangle_\phi^\Phi)(\langle 0_A|\hat{\sigma}_{yA}|1_A\rangle - \langle 1_A|\hat{\sigma}_{yA}|0_A\rangle) \\ &= 2\theta\Im(\langle\hat{S}^W\rangle_\phi^\Phi)\end{aligned} \tag{A.23}$$

where $\langle\hat{\sigma}_{xA}\rangle_{0_A} = \langle\hat{\sigma}_{yA}\rangle_{0_A} = 0$, $\hat{\sigma}_{xA}|0_A\rangle = |1_A\rangle$, $\hat{\sigma}_{xA}|1_A\rangle = |0_A\rangle$, $\hat{\sigma}_{yA}|0_A\rangle = i|1_A\rangle$ and $\hat{\sigma}_{yA}|1_A\rangle = -i|0_A\rangle$ have all been used. Finally, $\langle\hat{\sigma}_{xA}^2\rangle_{F_{\Phi,A}} = \langle\hat{\sigma}_{yA}^2\rangle_{F_{\Phi,A}} = \langle\hat{\sigma}_{xA}^2\rangle_{0_A} = \langle\hat{\sigma}_{yA}^2\rangle_{0_A} = 1$. Notice that this formalism is much simpler compared to that of the continuous pointer.

A.2 Weak Measurements of Products

Sometimes it may be interesting to find the weak value of the product of two or more observables. If they commute with one another, this product is hermitian and therefore an observable itself, so that it plays the role of $\hat{S}$ in the previous formalism and it can be treated the usual way as explained in Sections 2.1 and A.1. However, it may be more difficult to experimentally implement the product operator compared to the single factors, or they may not commute, like $\hat{\Pi}_{a_j}$ and $\hat{\Pi}_{b_0}$ in the direct weak reconstruction protocol proposed by Lundeen. For these cases there exists a different formulation that treats the various factors separately [45, 59].

I have shown in the previous section that

$$\Re\left(\langle\hat{S}^W\rangle_\phi^\Phi\right) \approx \frac{1}{g}\langle\hat{X}_A\rangle_{F_{\Phi,A}} \qquad \Im\left(\langle\hat{S}^W\rangle_\phi^\Phi\right) \approx \frac{2\Delta_x^2}{\hbar g}\langle\hat{P}_A\rangle_{F_{\Phi,A}} \tag{A.24}$$

I define the pointer operator $\hat{\Sigma}_A \equiv \frac{1}{2\Delta_x}\hat{X}_A + i\frac{\Delta_x}{\hbar}\hat{P}_A$ so that

$$\langle\hat{S}^W\rangle_\phi^\Phi = \frac{2\Delta_x}{g}\langle F_{\Phi,A}|\hat{\Sigma}_A|F_{\Phi,A}\rangle \tag{A.25}$$

Suppose that the initial state of the ancilla is the usual gaussian function of x_A but this time I call it $|0_A\rangle$ because of its resemblance to the ground state of an harmonic oscillator. It is also evident from its definition that $\hat{\Sigma}_A$ behaves as a lowering operator and $\hat{\Sigma}_A|0_A\rangle = 0$, while its transpose conjugate $\hat{\Sigma}_A^\dagger \equiv \frac{1}{2\Delta_x}\hat{X}_A - i\frac{\Delta_x}{\hbar}\hat{P}_A$ is a raising operator and $\hat{\Sigma}_A^\dagger|0_A\rangle = |1_A\rangle$. Notice moreover that $\hat{P}_A = i\frac{\hbar}{2\Delta_x}\left(\hat{\Sigma}_A^\dagger - \hat{\Sigma}_A\right)$.

Suppose that one wants to find the weak value of

$$\hat{S} = \prod_{j=1}^{N} \hat{S}_j \tag{A.26}$$

which is not necessarily hermitian. They can couple each observable to a separate ancilla so that the interaction Hamiltonian is

$$\hat{H}_{int} = i\left(\frac{\hbar g(t)}{2\Delta_x}\right) \sum_j \hat{S}_j \left(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j\right) \tag{A.27}$$

where for simplicity the strength is assumed to be the same for all couplings, that is, $g_j(t) = g(t)\ \forall j$. The joint initial state is

$$|\psi\rangle = |\phi_S\rangle \otimes |0\rangle^{\otimes N} \tag{A.28}$$

where $|0\rangle^{\otimes N}$ is the separable products of N pointer ground states. The evolution operator is

$$\begin{aligned} \hat{U} = e^{-\frac{i}{\hbar}\int \hat{H}_{int} dt} &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{i}{\hbar} \int \hat{H}_{int} dt\right)^k \\ &= \sum_k \frac{1}{k!} \left(\frac{g}{2\Delta_x} \sum_j \hat{S}_j (\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j)\right)^k \\ &= 1 + \frac{g}{2\Delta_x} \sum_j \hat{S}_j \left(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j\right) + \dots \end{aligned} \tag{A.29}$$

Consider the term of this expansion corresponding to $k = N$, that is

$$\frac{1}{N!} \left(\frac{g}{2\Delta_x} \sum_j \hat{S}_j \left(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j\right)\right)^N \tag{A.30}$$

It is the first that contains operators that can raise all the N pointers to the first excited state. In particular the part that does this is

$$\frac{1}{N!} \left(\frac{g}{2\Delta_x}\right)^N \wp\left(\hat{S}_l \hat{\Sigma}_l^\dagger\right) \tag{A.31}$$

where $\wp\left(\hat{S}_l \hat{\Sigma}_l^\dagger\right)$ labels the sum of all the $N!$ possible orderings of products of the operators $\left\{\hat{S}_l \hat{\Sigma}_l^\dagger \mid l = 1..N\right\}$.

When $\hat{U}$ is applied to $|\psi\rangle$ and the object system is postselected on $|\Phi_S\rangle$, the final normalized pointer state is in the form

$$\begin{aligned} |F_{\Phi,N}\rangle = |0\rangle^{\otimes N} &+ \frac{g}{2\Delta_x} \sum_j \frac{\langle\Phi_S|\hat{S}|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle} \otimes |1_j\rangle + \dots \\ &+ \left(\frac{g}{2\Delta_x}\right)^N \frac{1}{N!} \frac{\langle\Phi_S|\wp(\hat{S}_l)|\phi_S\rangle}{\langle\Phi_S|\phi_S\rangle} \otimes |1\rangle^{\otimes N} + \dots \end{aligned} \tag{A.32}$$

where $|1_j\rangle$ is the state in which all pointers are in the ground state except for the j-th one, that is in the first excited level, whereas $|1\rangle^{\otimes N}$ labels the configuration in which all the pointers are excited. Now consider the object

$$\left\langle \prod_j \hat{\Sigma}_j \right\rangle_{F_{\Phi,N}} \equiv \langle F_{\Phi,N}| \prod_j \hat{\Sigma}_j |F_{\Phi,N}\rangle \tag{A.33}$$

When applying operator $\prod_j \hat{\Sigma}_j$ to $|F_{\Phi,N}\rangle$ all the terms that contain a pointer still in the ground state are annihilated:

$$\prod_j \hat{\Sigma}_j |F_{\Phi,N}\rangle = \left(\frac{g}{2\Delta_x}\right)^N \frac{1}{N!} \langle \wp(\hat{S}_j)^W \rangle_\phi^\Phi |0\rangle^{\otimes N} + \dots \tag{A.34}$$

If addends of order higher than N in g are neglected

$$\left\langle \prod_j \hat{\Sigma}_j \right\rangle_{F_{\Phi,N}} \approx \left(\frac{g}{2\Delta_x}\right)^N \frac{1}{N!} \langle \wp(\hat{S}_j)^W \rangle_\phi^\Phi \tag{A.35}$$

The term $\frac{1}{N!}\langle \wp(\hat{S}_j)^W \rangle_\phi^\Phi$ is the so-called *joint weak value* of observables $\{\hat{S}_j \mid j = 1..N\}$. If they commute with one another, one can write

$$\left\langle \left(\prod_j \hat{S}_j\right)^W \right\rangle_\phi^\Phi = \frac{1}{N!} \langle \wp(\hat{S}_j)^W \rangle_\phi^\Phi \approx \left(\frac{2\Delta_x}{g}\right)^N \left\langle \prod_j \hat{\Sigma}_j \right\rangle_{F_{\Phi,N}} \tag{A.36}$$

In order to measure $\langle \prod_j \hat{\Sigma}_j \rangle_{F_{\Phi,N}}$, which is not the mean value of an observable because the $\hat{\Sigma}_j$s are not hermitian, one has to expand the lowering operators in products of $\hat{X}_j$ and $\hat{P}_k$ and measure them on different subensembles of pointers. For instance for $N = 2$:

$$\begin{aligned} \langle (\hat{S}_1 \hat{S}_2)^W \rangle_\phi^\Phi &= \left(\frac{2\Delta_x}{g}\right)^2 \langle \hat{\Sigma}_1 \hat{\Sigma}_2 \rangle_{F_{\Phi,AB}} \\ &= \left(\frac{2\Delta_x}{g}\right)^2 \left\langle \left(\frac{1}{2\Delta_x}\hat{X}_1 + i\frac{\Delta_x}{\hbar}\hat{P}_1\right)\left(\frac{1}{2\Delta_x}\hat{X}_2 + i\frac{\Delta_x}{\hbar}\hat{P}_2\right)\right\rangle_{F_{\Phi,AB}} \end{aligned} \tag{A.37}$$

therefore

$$\begin{aligned} \Re\left(\langle (\hat{S}_1 \hat{S}_2)^W \rangle_\phi^\Phi\right) &= \left(\frac{1}{g}\right)^2 \left(\langle \hat{X}_1 \hat{X}_2 \rangle_{F_{\Phi,AB}} - \frac{4\Delta_x^4}{\hbar^2} \langle \hat{P}_1 \hat{P}_2 \rangle_{F_{\Phi,AB}}\right) \\ \Im\left(\langle (\hat{S}_1 \hat{S}_2)^W \rangle_\phi^\Phi\right) &= \frac{2\Delta_x^2}{\hbar}\left(\frac{1}{g}\right)^2 \left(\langle \hat{X}_1 \hat{P}_2 \rangle_{F_{\Phi,AB}} + \langle \hat{P}_1 \hat{X}_2 \rangle_{F_{\Phi,AB}}\right) \end{aligned} \tag{A.38}$$

If the $\hat{S}_j$s do not commute, these relations are still valid provided one specific order is chosen. In *sequential weak measurements* [60], the couplings between object and ancillae are done one after another so that the evolution operator is

$$\hat{U} = \prod_{j=1}^N e^{\frac{g}{2\Delta_x}\hat{S}_j(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j)} = \prod_j \sum_{k=0}^\infty \frac{1}{k!}\left(\frac{g}{2\Delta_x}\hat{S}_j(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j)\right)^k = \prod_j \sum_k \frac{B_j^k}{k!} \tag{A.39}$$

where for ease of notation I have defined $B_j^k \equiv \left(\frac{g}{2\Delta_x}\hat{S}_j(\hat{\Sigma}_j^\dagger - \hat{\Sigma}_j)\right)^k$. Just as above the joint system evolves according to this unitary, the object is postselected on $|\Phi_S\rangle$ and finally $\langle\prod_j \hat{\Sigma}_j\rangle_{F_{\Phi,N}}$ is measured on the ancillae. Terms in $|F_{\Phi,N}\rangle$ that have evolved from a B_j^0 contribute nothing to the result because of the action of the lowering operators on the ground state. The first non-zero term comes from $\prod_j B_j^1$ and in particular contains

$$\left(\frac{g}{2\Delta_x}\right)^N \prod_j \hat{S}_j\hat{\Sigma}_j^\dagger \tag{A.40}$$

which raises all the pointers to the first excited state. Following the above passages again

$$\left\langle\left(\prod_j \hat{S}_j\right)^W\right\rangle_\phi^\Phi \approx \left(\frac{2\Delta_x}{g}\right)^N \left\langle\prod_j \hat{\Sigma}_j\right\rangle_{F_{\Phi,N}} \tag{A.41}$$

This strategy is used in the direct weak reconstruction of the density operator where the goal is $\langle\hat{\Pi}_{b_0}\hat{\Pi}_{a_j}^W\rangle_\varrho^{a_j}$. The two couplings are actualized sequentially: first that of $\hat{\Pi}_{a_j}$ and then one for $\hat{\Pi}_{b_0}$. In general this is the only practically feasible procedure, if the joint weak value $\frac{1}{N!}\langle\wp(\hat{S}_j)^W\rangle_\phi^\Phi$ is needed, one either measures all the $N!$ possible sequential weak values with different orders and averages them, or uses Trotter's formula [22, 51]

$$e^{A+B} = \lim_{n\to\infty}\left(e^{\frac{A}{n}}e^{\frac{B}{n}}\right)^n \tag{A.42}$$

so that a long series of alternated couplings can approximate the joint result.

In order to get closer to the formalism of the main text I now show that it easy to replicate this discussion in the case of qubit pointers. If the coupled operator of each ancilla is $\hat{\sigma}_y$ and the initial state is the eigenstate $|0\rangle$ of $\hat{\sigma}_z$, then one can define the ladder operators:

$$\hat{\Sigma} \equiv \frac{\hat{\sigma}_x + i\hat{\sigma}_y}{2} \qquad \hat{\Sigma}^\dagger \equiv \frac{\hat{\sigma}_x - i\hat{\sigma}_y}{2} \tag{A.43}$$

so that $\hat{\Sigma}|0\rangle = 0$ and $\hat{\Sigma}^\dagger|0\rangle = |1\rangle$. If θ denotes the coupling strength, the joint weak value is

$$\frac{1}{N!}\langle\wp(\hat{S}_j)^W\rangle_\phi^\Phi \approx \frac{1}{\theta^N}\left\langle\prod_j \hat{\Sigma}_j\right\rangle_{F_{\Phi,N}} \tag{A.44}$$

and the sequential weak value for couplings made in a specific order is

$$\left\langle\left(\prod_j \hat{S}_j\right)^W\right\rangle_\phi^\Phi \approx \frac{1}{\theta^N}\left\langle\prod_j \hat{\Sigma}_j\right\rangle_{F_{\Phi,N}} \tag{A.45}$$

which coincides with equation (2.31) for $N = 3$.

A.3 Derivation of the Measurement Protocols

I shall now prove the relations that are used in the measurement protocols of Sections 2.5 and 3.2. I will focus initially on the DRDO scheme which is the most comprehensive.

Consider a d-dimensional Hilbert space $\mathscr{H}_S$ and suppose that a system is prepared in an unknown state described by density operator $\hat{\varrho}_S$. The goal of the following dissertation is to find the d^2 matrix elements $\varrho_{jk} = \langle a_j|\hat{\varrho}_S|a_k\rangle$ expressed in the orthonormal basis $\{|a_j\rangle \mid j = 1..d\}$ of $\mathscr{H}_S$. Two bidimensional ancillae $\mathscr{H}_{A,B}$ are prepared in the pure state $|0\rangle\langle 0|$ so that the initial joint state can be written as

$$\hat{\varrho}_{tot} = \hat{\varrho}_S \otimes |0_A\rangle\langle 0_A| \otimes |0_B\rangle\langle 0_B| \tag{A.46}$$

Since the pointers are built ad-hoc, it is not necessary to consider more general cases such as multi-dimensional Hilbert spaces or mixed initial states (although purity is always an approximation in practice) [24, 25]. For each choice of j, observables $\hat{\Pi}_{a_j} \equiv |a_j\rangle\langle a_j|$ and $\hat{\sigma}_{yA}$ are coupled so that the evolution operator is

$$\hat{U}_A(a_j) = e^{-i\theta_A \hat{\Pi}_{a_j} \otimes \hat{\sigma}_{yA}} = (\mathbb{1} + (c_A - 1)\hat{\Pi}_{a_j}) \otimes \mathbb{1}_A - i s_A \hat{\Pi}_{a_j} \otimes \hat{\sigma}_{yA} \tag{A.47}$$

which does not act on $\mathscr{H}_B$. (From now on I will use symbols $c_{A,B}$, $s_{A,B}$ and $t_{A,B}$ in place of $\cos(\theta_{A.B})$, $\sin(\theta_{A,B})$ and $\tan\left(\frac{\theta_{A,B}}{2}\right)$). Following the abstract working principle of any measuring device, this coupling is implemented using tools that respond to the j-th component of $\hat{\varrho}_S$ with a change of the state of the ancilla.

Subsequently a similar coupling is made between $\hat{\Pi}_{b_0} = |b_0\rangle\langle b_0|$ and $\hat{\sigma}_{yB}$, where $|b_0\rangle \equiv \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |a_j\rangle$. In this case the evolution operator is

$$\hat{U}_B = e^{-i\theta_B \hat{\Pi}_{b_0} \otimes \hat{\sigma}_{yB}} = (\mathbb{1} + (c_B - 1)\hat{\Pi}_{b_0}) \otimes \mathbb{1}_B - i s_B \hat{\Pi}_{b_0} \otimes \hat{\sigma}_{yB} \tag{A.48}$$

which does not act on $\mathscr{H}_A$. By defining the operators on $\mathscr{H}_S$:

$$\begin{aligned} \hat{\alpha}_0 &\equiv \mathbb{1} + (c_A - 1)\hat{\Pi}_{a_j} \\ \hat{\alpha}_1 &\equiv s_A \hat{\Pi}_{a_j} \\ \hat{\beta}_0 &\equiv \mathbb{1} + (c_B - 1)\hat{\Pi}_{b_0} \\ \hat{\beta}_1 &\equiv s_B \hat{\Pi}_{b_0} \end{aligned} \tag{A.49}$$

one can write

$$\hat{U}_A(a_j) = \sum_{n=0}^{1} (-i)^n \hat{\alpha}_n \hat{\sigma}_{yA}^n \qquad \hat{U}_B = \sum_{m=0}^{1} (-i)^m \hat{\beta}_m \hat{\sigma}_{yB}^m \tag{A.50}$$

Then the evolved state is

$$\begin{aligned} \hat{\varrho}'_{tot}(a_j) &\equiv \hat{U}_B \hat{U}_A(a_j) \hat{\varrho} \hat{U}_A^\dagger(a_j) \hat{U}_B^\dagger \\ &= \sum_{n,n',m,m'} i^{n'+m'-n-m} \hat{\beta}_m \hat{\alpha}_n \hat{\varrho}_S \hat{\alpha}_{n'} \hat{\beta}'_m \otimes \hat{\sigma}_{yA}^n |0_A\rangle\langle 0_A| \hat{\sigma}_{yA}^{n'} \otimes \hat{\sigma}_{yB}^m |0_B\rangle\langle 0_B| \hat{\sigma}_{yB}^{m'} \end{aligned} \tag{A.51}$$

The last step is a postselection on $|a_k\rangle$ so that the final state is

$$|a_k\rangle\langle a_k| \otimes \langle a_k|\hat{\varrho}'_{tot}(a_j)|a_k\rangle = |a_k\rangle\langle a_k| \otimes \frac{1}{4}\sum_{\mu,\nu=0}^{3} r_{\mu\nu}^{(jk)} \cdot \hat{\sigma}_{\mu A} \otimes \hat{\sigma}_{\nu B} \tag{A.52}$$

where $\hat{\sigma}_0 = \mathbb{1}$, $\hat{\sigma}_1 = \hat{\sigma}_x$, $\hat{\sigma}_2 = \hat{\sigma}_y$, $\hat{\sigma}_3 = \hat{\sigma}_z$ and

$$\begin{aligned} r_{\mu\nu}^{(jk)} &\equiv \mathrm{Tr}(\langle a_k|\hat{\varrho}'_{tot}(a_j)|a_k\rangle \otimes \hat{\sigma}_{\mu A} \otimes \hat{\sigma}_{\nu B}) \\ &= \sum_{n,n',m,m'} i^{n'+m'-n-m'} \langle a_k|\hat{\beta}_m \hat{\alpha}_n \hat{\varrho}_S \hat{\alpha}_{n'} \hat{\beta}_{m'}|a_k\rangle \cdot \langle 0_A|\hat{\sigma}_{yA}^{n'} \hat{\sigma}_{\mu A} \hat{\sigma}_{yA}^{n}|0_A\rangle \cdot \langle 0_B|\hat{\sigma}_{yB}^{m'} \hat{\sigma}_{\nu B} \hat{\sigma}_{yB}^{m}|0_B\rangle \\ &= \sum_{n,n',m,m'} i^{n'+m'-n-m'} \langle a_k|\hat{\beta}_m \hat{\alpha}_n \hat{\varrho}_S \hat{\alpha}_{n'} \hat{\beta}_{m'}|a_k\rangle \cdot G_{n,n'}^{\mu} G_{m,m'}^{\nu} \end{aligned} \tag{A.53}$$

in which

$$G_{n,n'}^{\mu} \equiv \delta_{\mu,0}\delta_{n,n'} + \delta_{\mu,1} i^{2n-1} \delta_{n,1-n'} + \delta_{\mu,2}\delta_{n,1-n'} + \delta_{\mu,3}(-1)^n \delta_{n,n'} \tag{A.54}$$

For simplicity I define:

$$\hat{A}_\mu \equiv \sum_{n,n'} i^{n'-n} \hat{\alpha}_n \hat{\varrho}_S \hat{\alpha}_{n'} G_{n,n'}^{\mu} \tag{A.55}$$

so that

$$\begin{aligned} \hat{A}_0 &= \hat{A}_0^\dagger = \hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_0 + \hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_1 = \hat{\varrho}_S + (c_A - 1)(\hat{\Pi}_{a_1}\hat{\varrho}_S + \hat{\varrho}_S\hat{\Pi}_{a_1} - 2\hat{\Pi}_{a_1}\hat{\varrho}_S\hat{\Pi}_{a_1}) \\ \hat{A}_1 &= \hat{A}_1^\dagger = \hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_1 + \hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_0 = s_A\left(\hat{\varrho}_S\hat{\Pi}_{a_1} + \hat{\Pi}_{a_1}\hat{\varrho}_S + 2(c_A - 1)\hat{\Pi}_{a_1}\hat{\varrho}_S\hat{\Pi}_{a_1}\right) \\ \hat{A}_2 &= \hat{A}_2^\dagger = i(\hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_1 - \hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_0) = is_A\left(\hat{\varrho}_S\hat{\Pi}_{a_1} - \hat{\Pi}_{a_1}\hat{\varrho}_S\right) \\ \hat{A}_3 &= \hat{A}_3^\dagger = \hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_0 - \hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_1 = \hat{\varrho}_S + (c_A - 1)(\hat{\Pi}_{a_1}\hat{\varrho}_S + \hat{\varrho}_S\hat{\Pi}_{a_1} + 2c_A\hat{\Pi}_{a_1}\hat{\varrho}_S\hat{\Pi}_{a_1}) \end{aligned} \tag{A.56}$$

Then (A.53) becomes

$$r_{\mu\nu}^{(jk)} = \sum_{m,m'} i^{m'-m} \langle a_k|\hat{\beta}_m \hat{A}_\mu \hat{\beta}_{m'}|a_k\rangle \cdot G_{m,m'}^{\nu} \tag{A.57}$$

In particular, considering that

$$\hat{\beta}_0|a_k\rangle = |a_k\rangle + \frac{c_B - 1}{d}\sum_{l=1}^{d}|a_l\rangle \qquad \hat{\beta}_1|a_k\rangle = \frac{s_B}{d}\sum_{l=1}^{d}|a_l\rangle \tag{A.58}$$

one can write:

$$\begin{aligned}
r_{\mu 0}^{(jk)} &= \langle a_k|\hat{\beta}_0\hat{A}_\mu\hat{\beta}_0|a_k\rangle + \langle a_k|\hat{\beta}_1\hat{A}_\mu\hat{\beta}_1|a_k\rangle \\
&= \langle a_k|\hat{A}_\mu|a_k\rangle + \frac{2(c_B-1)}{d}\left(\sum_l \Re(\langle a_k|\hat{A}_\mu|a_l\rangle) - \frac{1}{d}\sum_{l,l'}\langle a_{l'}|\hat{A}_\mu|a_l\rangle\right) \\
r_{\mu 1}^{(jk)} &= \langle a_k|\hat{\beta}_0\hat{A}_\mu\hat{\beta}_1|a_k\rangle + \langle a_k|\hat{\beta}_1\hat{A}_\mu\hat{\beta}_0|a_k\rangle \\
&= \frac{2s_B}{d}\left(\sum_l \Re(\langle a_k|\hat{A}_\mu|a_l\rangle) + \frac{c_B-1}{d}\sum_{l,l'}\langle a_{l'}|\hat{A}_\mu|a_l\rangle\right) \\
r_{\mu 2}^{(jk)} &= i\langle a_k|\hat{\beta}_0\hat{A}_\mu\hat{\beta}_1|a_k\rangle - i\langle a_k|\hat{\beta}_1\hat{A}_\mu\hat{\beta}_0|a_k\rangle \\
&= -\frac{2s_B}{d}\sum_l \Im(\langle a_k|\hat{A}_\mu|a_l\rangle) \\
r_{\mu 3}^{(jk)} &= \langle a_k|\hat{\beta}_0\hat{A}_\mu\hat{\beta}_0|a_k\rangle - \langle a_k|\hat{\beta}_1\hat{A}_\mu\hat{\beta}_1|a_k\rangle \\
&= \langle a_k|\hat{A}_\mu|a_k\rangle + \frac{2(c_B-1)}{d}\left(\sum_l \Re(\langle a_k|\hat{A}_\mu|a_l\rangle) + \frac{c_B}{d}\sum_{l,l'}\langle a_{l'}|\hat{A}_\mu|a_l\rangle\right)
\end{aligned} \tag{A.59}$$

Then it is possible to evaluate

$$\begin{aligned}
\langle a'_l|\hat{A}_0|a_l\rangle &= \varrho_{ll'} + (c_A-1)(\delta_{jl'}\varrho_{jl} + \delta_{jl}\varrho_{l'j} - 2\delta_{jl'}\delta_{jl}\varrho_{jj}) \\
\langle a'_l|\hat{A}_1|a_l\rangle &= s_A\left(\delta_{jl}\varrho_{l'j} + \delta_{jl'}\varrho_{jl} + 2(c_A-1)\delta_{jl'}\delta_{jl}\varrho_{jj}\right) \\
\langle a'_l|\hat{A}_2|a_l\rangle &= is_A\left(\delta_{jl}\varrho_{l'j} - \delta_{jl'}\varrho_{jl}\right) \\
\langle a'_l|\hat{A}_3|a_l\rangle &= \varrho_{l'l} + (c_A-1)(\delta_{jl'}\varrho_{jl} + \delta_{jl'}\varrho_{l'j} + 2c_A\delta_{jl'}\delta_{jl}\varrho_{jj})
\end{aligned} \tag{A.60}$$

and substitute into the previous equation in order to find a total of 16 quantities. Some of the most interesting ones are reported here:

$$\begin{aligned}
r_{11}^{(jk)} &= \frac{2s_As_B}{d}\left((1-\delta_{jk})\Re(\varrho_{jk}) + \delta_{jk}\sum_{l\neq j}\Re(\varrho_{jl}) + 2c_A\delta_{jk}\Re(\varrho_{jk})\right) + \\
&\quad + \frac{4s_As_B(c_B-1)}{d^2}\left(\sum_{l\neq j}\Re(\varrho_{jl}) + c_A\varrho_{jj}\right) \\
r_{12}^{(jk)} &= \frac{2s_As_B}{d}\left(\Im(\varrho_{jk}) - \delta_{jk}\sum_{l\neq j}\Im(\varrho_{jl})\right) \\
r_{21}^{(jk)} &= \frac{2s_As_B}{d}\Im\left(\varrho_{jk} + \delta_{jk}\sum_{l\neq j}\varrho_{jl} + \frac{2(c_B-1)}{d}\sum_l\varrho_{jl}\right) \\
r_{22}^{(jk)} &= \frac{2s_As_B}{d}\left(-\Re(\varrho_{jk}) + \delta_{jk}\sum_l\Re(\varrho_{jl})\right)
\end{aligned} \tag{A.61}$$

Considering that $r_{\mu\nu}^{(jk)} = \mathrm{Tr}(\langle a_k|\hat{\varrho}'_{tot}(a_j)|a_k\rangle \otimes \hat{\sigma}_{\mu A}^{(a_j)} \otimes \hat{\sigma}_{\nu B}^{(b_0)}) = \mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S)\langle \hat{\sigma}_{\mu A}^{(a_j)} \hat{\sigma}_{\nu B}^{(b_0)} \rangle_{F_{a_k},AB}$, one finds:

$$\begin{aligned}
&\langle \hat{\Pi}_{1A}^{(a_j)} \hat{\Pi}_{1B}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{00}^{(jk)} - r_{30}^{(jk)} - r_{03}^{(jk)} + r_{33}^{(jk)} = \frac{s_A^2 s_B^2}{d^2}\varrho_{jj} \qquad \forall k \\
&\langle \hat{\sigma}_{yA}^{(a_j)} \hat{\sigma}_{yB}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{22}^{(jk)} = -\frac{2s_A s_B}{d}\Re(\varrho_{jk}) \qquad j \neq k \\
&\langle \hat{\sigma}_{xA}^{(a_j)} \hat{\sigma}_{yB}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{12}^{(jk)} = \frac{2s_A s_B}{d}\Im(\varrho_{jk}) \qquad j \neq k
\end{aligned} \tag{A.62}$$

as stated in equation 3.7.

If θ_A and θ_B are small and terms of order higher than 1 in $\theta_A\theta_B$ are neglected, then also equation 2.33 is verified $\forall j, k$:

$$\begin{aligned}
&\left(\langle \hat{\sigma}_{xA}^{(a_j)} \hat{\sigma}_{xB}^{(b_0)} \rangle_{F_{a_k},AB} - \langle \hat{\sigma}_{yA}^{(a_j)} \hat{\sigma}_{yB}^{(b_0)} \rangle_{F_{a_k},AB}\right) \mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{11}^{(jk)} - r_{22}^{(jk)} \\
&= 4s_A s_B \left(\frac{(1-\delta_{jk})}{d}\Re(\varrho_{jk}) + \frac{c_B - 1}{d^2}\sum_{l\neq j}\Re(\varrho_{jl}) + \delta_{jk}\frac{c_A}{d}\Re(\varrho_{jk}) + \frac{c_A(c_B-1)}{d^2}\varrho_{jj} \right) \\
&\approx \frac{4\theta_A\theta_B}{d}\Re(\varrho_{jk}) \\
&\left(\langle \hat{\sigma}_{xA}^{(a_j)} \hat{\sigma}_{yB}^{(b_0)} \rangle_{F_{a_k},AB} + \langle \hat{\sigma}_{yA}^{(a_j)} \hat{\sigma}_{xB}^{(b_0)} \rangle_{F_{a_k},AB}\right) \mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{12}^{(jk)} + r_{21}^{(jk)} \\
&= 4s_A s_B \left(\frac{(1-\delta_{jk})}{d}\Im(\varrho_{jk}) + \frac{c_B - 1}{d^2}\sum_{l\neq j}\Im(\varrho_{jl}) \right) \\
&\approx \frac{4\theta_A\theta_B}{d}\Im(\varrho_{jk})
\end{aligned} \tag{A.63}$$

The exact equation, along with

$$\begin{aligned}
&\langle \hat{\Pi}_{1A}^{(a_j)} \hat{\sigma}_{xB}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{01}^{(jk)} - r_{31}^{(jk)} = \delta_{jk}\frac{2s_A^2 s_B}{d}\varrho_{jk} + \frac{2s_A^2 s_B(c_B-1)}{d^2}\varrho_{jj} \\
&\langle \hat{\sigma}_{xA}^{(a_j)} \hat{\Pi}_{1B}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{10}^{(jk)} - r_{13}^{(jk)} = \frac{2s_A s_B^2}{d^2}\sum_l \Re(\varrho_{jl}) + \frac{2s_A s_B^2(c_A-1)}{d^2}\varrho_{jj} \\
&\langle \hat{\sigma}_{yA}^{(a_j)} \hat{\Pi}_{1B}^{(b_0)} \rangle_{F_{a_k},AB} \,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = r_{20}^{(jk)} - r_{23}^{(jk)} = \frac{2s_A s_B^2}{d^2}\sum_l \Im(\varrho_{jl})
\end{aligned} \tag{A.64}$$

also proves formula 3.6. Indeed:

$$\begin{aligned}
&\Bigg(\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle_{F_{a_k},AB} - \langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle_{F_{a_k},AB} + 2t_B\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle_{F_{a_k},AB} + 2t_A\langle\hat{\Pi}_{1A}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle_{F_{a_k},AB} + \\
&+ 4t_At_B\langle\hat{\Pi}_{1A}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle_{j,k}\Bigg)\,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) \\
&= r_{11}^{(jk)} + r_{22}^{(jk)} + 2t_B\left(r_{10}^{(jk)} - r_{13}^{(jk)}\right) + 2t_A\left(r_{01}^{(jk)} - r_{31}^{(jk)}\right) + \\
&+ 4t_At_B\left(r_{00}^{(jk)} - r_{30}^{(jk)} - r_{03}^{(jk)} + r_{33}^{(jk)}\right) \\
&= \frac{4s_As_B}{d}\Re(\varrho_{jk}) \\
&\left(\langle\hat{\sigma}_{xA}^{(a_j)}\hat{\sigma}_{yB}^{(b_0)}\rangle_{F_{a_k},AB} + \langle\hat{\sigma}_{yA}^{(a_j)}\hat{\sigma}_{xB}^{(b_0)}\rangle_{F_{a_k},AB} + 2t_B\langle\hat{\sigma}_{yA}^{(a_j)}\hat{\Pi}_{1B}^{(b_0)}\rangle_{F_{a_k},AB}\right)\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) \\
&= r_{12}^{(jk)} + r_{21}^{(jk)} + 2t_B\left(r_{20}^{(jk)} - r_{23}^{(jk)}\right) \\
&= \frac{4s_As_B}{d}\Im(\varrho_{jk})
\end{aligned} \tag{A.65}$$

Inverting this shows that blindly using the approximation of formula (A.63) to find the matrix elements when θ_A and/or θ_B are not small enough causes a bias

$$\begin{aligned}
\Re(\varrho_{jk}^W - \varrho_{jk}) &= \frac{d}{4s_As_B}\cdot\Bigg(\left(r_{11}^{(jk)} + r_{22}^{(jk)}\right) \\
&- \left(r_{11}^{(jk)} + r_{22}^{(jk)} + 2t_B\left(r_{10}^{(jk)} - r_{13}^{(jk)}\right) + 2t_A\left(r_{01}^{(jk)} - r_{31}^{(jk)}\right) + \right. \\
&\left. + 4t_At_B\left(r_{00}^{(jk)} - r_{30}^{(jk)} - r_{03}^{(jk)} + r_{33}^{(jk)}\right)\right)\Bigg) \\
&= \delta_{jk}(c_A-1)\varrho_{jk} + \frac{c_B-1}{d}\sum_l\Re(\varrho_{jl}) + \frac{(c_A-1)(c_B-1)}{d}\varrho_{jj} \\
\Im(\varrho_{jk}^W - \varrho_{jk}) &= \frac{d}{4s_As_B}\cdot\left(\left(r_{12}^{(jk)} + r_{21}^{(jk)}\right) - \left(r_{12}^{(jk)} + r_{21}^{(jk)} + 2t_B\left(r_{20}^{(jk)} - r_{23}^{(jk)}\right)\right)\right) \\
&= \frac{c_B-1}{d}\sum_l\Im(\varrho_{jl})
\end{aligned} \tag{A.66}$$

which validates equation (3.14).

This formalism is also helpful to prove relation (3.18) of the main text. Supposing that the measurement results are only affected by Poissonian error, the variance associated to the counts that one measures in order to find the mean value of projector $\hat{\Pi}_{\mu a}\otimes\hat{\Pi}_{\nu b}$ is

$$\delta N_{\mu a,\nu b,jk}^2 = N\langle\hat{\Pi}_{\mu a}^{(a_j)}\hat{\Pi}_{\nu b}^{(b_0)}\rangle_{F_{a_k},AB}\,\mathrm{Tr}(\hat{\Pi}_{a_k}\hat{\varrho}_S) = \frac{N}{4}\left(r_{00}^{(jk)} + ar_{\mu 0}^{(jk)} + br_{0\nu}^{(jk)} + abr_{\mu\nu}^{(jk)}\right) \tag{A.67}$$

where N is the total power of the signal, $\hat{\Pi}_{\mu a} = \frac{1}{2}\left(\mathbb{1}_A + a\hat{\sigma}_{\mu A}\right)$, $\hat{\Pi}_{\nu b} = \frac{1}{2}\left(\mathbb{1}_B + b\hat{\sigma}_{\nu B}\right)$ and $a, b = \pm 1$. As in Subsection 3.3.2, the error of N is not considered because it depends on how it is calculated or measured. However, it usually does not change the order of

magnitude of the result. The variance of observable $\hat{\sigma}_{\mu A} \otimes \hat{\sigma}_{\nu B}$ becomes

$$\delta^2_{\mu\nu,jk} = \sum_{a,b} \frac{\delta N^2_{\mu a,\nu b,jk}}{N^2} = \frac{r^{jk}_{00}}{N} = \frac{\mathrm{Tr}(\langle a_k|\varrho'_{tot}(a_j)|a_k\rangle)}{N} \tag{A.68}$$

where

$$\begin{aligned} r^{(jk)}_{00} = \varrho_{kk} + \frac{2(c_B-1)}{d}\left(\sum_l \Re(\varrho_{lk}) - \frac{1}{d}\sum_{ll'}\varrho_{ll'}\right) + \\ + \frac{2(c_A-1)(c_B-1)}{d}\left(\Re(\varrho_{jk}) + \left(\delta_{jk}-\frac{2}{d}\right)\sum_l \Re(\varrho_{jl}) + 2\left(\frac{1}{d}-\delta_{jk}\right)\varrho_{jj}\right) \end{aligned} \tag{A.69}$$

From this expression, and also from the fact that $r^{(jk)}_{00}$ is the probability of successful postselection on state $|a_k\rangle$ after the two evolutions, it is clear that $\sum_k r^{(jk)}_{00} = 1$, which demonstrates formula (3.19).

The proof of the DRDD protocol is easier because only one coupling is involved. Similarly to the previous discussion, the starting point is

$$\hat{\varrho}_{tot} = \hat{\varrho}_S \otimes |0_A\rangle\langle 0_A| \tag{A.70}$$

and the only evolution operator is

$$\hat{U}_A(a_j) = e^{-i\theta_A \hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA}} = (\mathbb{1}+(c_A-1)\hat{\Pi}_{a_j})\otimes\mathbb{1}_A - is_A\hat{\Pi}_{a_j}\otimes\hat{\sigma}_{yA} = \sum_{n=0}^{1}(-i)^n\hat{\alpha}_n\hat{\sigma}^n_{yA} \tag{A.71}$$

again with

$$\hat{\alpha}_0 \equiv \mathbb{1} + (c_A-1)\hat{\Pi}_{a_j} \qquad \hat{\alpha}_1 \equiv s_A\hat{\Pi}_{a_j} \tag{A.72}$$

The result of the evolution is

$$\hat{\varrho}'_{tot}(a_j) \equiv \hat{U}_A(a_j)\hat{\varrho}\hat{U}^\dagger_A(a_j) = \sum_{n,n'} i^{n'-n}\hat{\alpha}_n\hat{\varrho}_S\hat{\alpha}_{n'}\otimes\hat{\sigma}^n_{yA}|0_A\rangle\langle 0_A|\hat{\sigma}^{n'}_{yA} \tag{A.73}$$

but this time postselection occurs on $|b_l\rangle \equiv \frac{1}{\sqrt{d}}\sum_{m=0}^{d-1}|a_m\rangle e^{i\frac{2\pi ml}{d}}$ with $l \in \{0..d-1\}$.

The final state is the separable product of $|b_l\rangle\langle b_l|$ in the object system and

$$\langle b_l|\hat{\varrho}'_{tot}(a_j)|b_l\rangle = \sum_{n,n'} i^{n'-n}\langle b_l|\hat{\alpha}_n\hat{\varrho}_S\hat{\alpha}_{n'}|b_l\rangle\otimes\hat{\sigma}^n_{yA}|0_A\rangle\langle 0_A|\hat{\sigma}^{n'}_{yA} = \frac{1}{2}\sum_{\mu=0}^{4} r^{(jl)}_\mu\hat{\sigma}_{\mu A} \tag{A.74}$$

for the ancilla. In this formalism

$$\begin{aligned} r^{(jl)}_\mu &\equiv \mathrm{Tr}(\langle b_l|\hat{\varrho}'_{tot}(a_j)|b_l\rangle\otimes\hat{\sigma}_{\mu A}) \\ &= \sum_{n,n'} i^{n'-n}\langle b_l|\hat{\alpha}_n\hat{\varrho}_S\hat{\alpha}_{n'}|b_l\rangle\cdot\langle 0_A|\hat{\sigma}^{n'}_{yA}\hat{\sigma}_{\mu A}\hat{\sigma}^n_{yA}|0_A\rangle \\ &= \sum_{n,n'} i^{n'-n}\langle b_l|\hat{\alpha}_n\hat{\varrho}_S\hat{\alpha}_{n'}|b_l\rangle\cdot G^\mu_{n,n'} \end{aligned} \tag{A.75}$$

and just as above

$$G^{\mu}_{n,n'} \equiv \delta_{\mu,0}\delta_{n,n'} + \delta_{\mu,1} i^{2n-1}\delta_{n,1-n'} + \delta_{\mu,2}\delta_{n,1-n'} + \delta_{\mu,3}(-1)^n\delta_{n,n'} \tag{A.76}$$

Considering that

$$\hat{\alpha}_0|b_l\rangle = \frac{1}{\sqrt{d}}\sum_{m=0}^{d-1} e^{i\frac{2\pi ml}{d}}|a_m\rangle + \frac{c_A-1}{\sqrt{d}} e^{i\frac{2\pi jl}{d}}|a_j\rangle \qquad \hat{\alpha}_1|b_l\rangle = \frac{s_A}{\sqrt{d}} e^{i\frac{2\pi jl}{d}}|a_j\rangle \tag{A.77}$$

one can write:

$$\begin{aligned}
r_0^{(jl)} &= \langle b_l|\hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_0|b_l\rangle + \langle b_l|\hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_1|b_l\rangle \\
&= \frac{1}{d}\sum_{m,m'} e^{i\frac{2\pi(m'-m)l}{d}}\varrho_{mm'} + \frac{2(c_A-1)}{d}\sum_m \Re\left(e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm}\right) - \frac{2(c_A-1)}{d}\varrho_{jj} \\
r_1^{(jl)} &= \langle b_l|\hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_1|b_l\rangle + \langle b_l|\hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_0|b_l\rangle \\
&= \frac{2s_A}{d}\sum_m \Re\left(e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm}\right) + \frac{2s_A(c_A-1)}{d}\varrho_{jj} \\
r_2^{(jl)} &= i\langle b_l|\hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_1|b_l\rangle - i\langle b_l|\hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_0|b_l\rangle \\
&= \frac{2s_A}{d}\sum_m \Im\left(e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm}\right) \\
r_3^{(jl)} &= \langle b_l|\hat{\alpha}_0\hat{\varrho}_S\hat{\alpha}_0|b_l\rangle - \langle b_l|\hat{\alpha}_1\hat{\varrho}_S\hat{\alpha}_1|b_l\rangle \\
&= \frac{1}{d}\sum_{m,m'} e^{i\frac{2\pi(m'-m)l}{d}}\varrho_{mm'} + \frac{2(c_A-1)}{d}\sum_m \Re\left(e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm}\right) + \frac{2c_A(c_A-1)}{d}\varrho_{jj}
\end{aligned} \tag{A.78}$$

According to formula (3.9)

$$\begin{aligned}
\rho_{A1}(jl) &= \frac{1}{2s_A}\mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho}_S)(\langle\hat{\sigma}^{(a_j)}_{xA}\rangle_{F_{b_l,A}} + i\langle\hat{\sigma}^{(a_j)}_{yA}\rangle_{F_{b_l,A}}) = \frac{1}{2s_A}\left(r_1^{(jl)} + ir_2^{(jl)}\right) \\
&= \frac{1}{d}\sum_m e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm} + \frac{c_A-1}{d}\varrho_{jj} \\
\rho_{A2}(jl) &= \frac{1}{s_A}\mathrm{Tr}(\hat{\Pi}_{b_l}\hat{\varrho}_S)\langle\hat{\Pi}^{(a_j)}_{1A}\rangle_{F_{b_l,A}} = \frac{1}{2s_A}\left(r_0^{(jl)} - ir_3^{(jl)}\right) = \frac{s_A}{d}\varrho_{jj}
\end{aligned} \tag{A.79}$$

so that $D_{jl} = \rho_{A1}(jl) + t_A\rho_{A2}(jl) = \frac{1}{d}\sum_m e^{i\frac{2\pi(m-j)l}{d}}\varrho_{jm}$ coincides with the jl-th element of the discrete Dirac distribution. Moreover, due to the fact that $\frac{1}{d}\sum_l e^{i\frac{2\pi(a-b)l}{d}} = \delta_{ab}$, one has

$$\begin{aligned}
&d\cdot t_A\delta_{jk}\rho_{A2}(jl) + \sum_l e^{\frac{2\pi il(j-k)}{d}}\rho_{A1}(jl) \\
&= (1-c_A)\delta_{jk}\varrho_{jj} + \sum_m\frac{1}{d}\sum_l e^{i\frac{2\pi(m-k)l}{d}}\varrho_{jm} + \frac{c_A-1}{d}\sum_l e^{i\frac{2\pi(j-k)l}{d}}\varrho_{jj} \\
&= \varrho_{jk}
\end{aligned} \tag{A.80}$$

which validates this protocol as a way to find the density operator. When $\theta_{A,B}$ is small, one can neglect the contribution of $\rho_{A2}(jl)$ and find Lundeen's proposal of (2.35).

A.4 Strong Measurements of Products

I now explain how to measure the mean value of an operator on the object system $\mathscr{H}_S$ that can be written as $\prod_{j=1}^{N}\hat{\Pi}_{j,S}$, using the ancilla scheme and without any approximation. Initially, the system is prepared in

$$\hat{\varrho}_S \otimes (|0\rangle\langle 0|)^{\otimes N} \tag{A.81}$$

in which the superscript $^{\otimes N}$ labels the tensor product over all the N pointers. Then, the various projectors $\hat{\Pi}_j$ are coupled sequentially and in reverse order to the $\hat{\sigma}_y$ observables of each ancilla so that the evolution operator of the joint system is

$$\hat{U} = \prod_{j=1}^{N}\hat{U}_j = \hat{U}_1..\hat{U}_N \tag{A.82}$$

where each $\hat{U}_j$ acts only on $\mathscr{H}_S$ and on the j-th meter as

$$\hat{U}_j = e^{-i\theta_j\hat{\Pi}_{j,S}\otimes\hat{\sigma}_{y,j}} = (\mathbb{1}_S + (s_j-1)\hat{\Pi}_{j,S})\otimes\mathbb{1}_j - is_A\hat{\Pi}_{j,S}\otimes\hat{\sigma}_{y,j} \tag{A.83}$$

(again $s_j \equiv \sin(\theta_j)$, $c_j \equiv \cos(\theta_j)$ and $t_j \equiv \tan\left(\frac{\theta_j}{2}\right)$).

The goal of this proof is to show that, as in relation (3.27), it is possible to define

$$\hat{E}_j \equiv \frac{\hat{\sigma}_{x,j} + i\hat{\sigma}_{y,j}}{2} + t_j\hat{\Pi}_{1,j} \tag{A.84}$$

and then to write

$$\left\langle \prod_{j=1}^{N}\hat{E}_j \right\rangle_{F_N} = \prod_{j=1}^{N} s_j \left\langle \prod_{j=1}^{N}\hat{\Pi}_{j,S} \right\rangle_{\hat{\varrho}_S} \tag{A.85}$$

First, notice the left-hand side can be expressed as

$$\begin{aligned}\left\langle \prod_{j=1}^{N}\hat{E}_j \right\rangle_{F_N} &= \mathrm{Tr}\left(\mathbb{1}_S\otimes\hat{E}^{\otimes N}\hat{U}_1..\hat{U}_N\hat{\varrho}_S\otimes(|0\rangle\langle 0|)^{\otimes N}\hat{U}_N^\dagger..\hat{U}_1^\dagger\right)\\ &= \mathrm{Tr}_S\left(\langle 0|^{\otimes N}\hat{U}_N^\dagger..\hat{U}_1^\dagger\mathbb{1}_S\otimes\hat{E}^{\otimes N}\hat{U}_1..\hat{U}_N|0\rangle^{\otimes N}\hat{\varrho}_S\right)\end{aligned} \tag{A.86}$$

so that (A.85) is equivalent to

$$\langle 0|^{\otimes N}\hat{U}_N^\dagger..\hat{U}_1^\dagger\mathbb{1}_S\otimes\hat{E}^{\otimes N}\hat{U}_1..\hat{U}_N|0\rangle^{\otimes N} = \prod_{j=1}^{N} s_j\hat{\Pi}_{j,S} \tag{A.87}$$

More generally, consider

$$\langle 0|^{\otimes N}\hat{U}_N^\dagger..\hat{U}_1^\dagger\hat{A}_S\otimes\hat{E}^{\otimes N}\hat{U}_1..\hat{U}_N|0\rangle^{\otimes N} = \hat{A}_S\prod_{j=1}^{N} s_j\hat{\Pi}_{j,S} \tag{A.88}$$

where $\hat{A}_S$ is an arbitrary operator on $\mathscr{H}_S$. It is easy to prove that this expression is valid for $N=1$. Indeed since $\hat{U}_1|0_1\rangle = (\mathbb{1}_S + (c_1-1)\hat{\Pi}_{1,S})|0_1\rangle + s_1\hat{\Pi}_{1,S}|1_1\rangle$ and

$\hat{E}_1\hat{U}_1|0_1\rangle = s_1\hat{\Pi}_{1,S}(|0_1\rangle + t_1|1_1\rangle)$, one arrives at

$$\begin{aligned}\langle 0_1|\hat{U}_1^\dagger \hat{A}_S \otimes \hat{E}_1\hat{U}_1|0_1\rangle &= s_1\hat{A}_S\hat{\Pi}_{1,S} + s_1(c_1-1)\hat{\Pi}_{1,S}\hat{A}_S\hat{\Pi}_S + s_1^2 t_1\hat{\Pi}_{1,S}\hat{A}_S\hat{\Pi}_S \\ &= s_1\hat{A}_S\hat{\Pi}_{j,S}\end{aligned} \tag{A.89}$$

Now suppose that (A.88) holds true for a given N, then it must also be valid for $N+1$. Indeed

$$\begin{aligned}&\langle 0|^{\otimes N+1}\hat{U}_{N+1}^\dagger..\hat{U}_1^\dagger\hat{A}_S\otimes\hat{E}^{\otimes N+1}\hat{U}_1..\hat{U}_{N+1}|0\rangle^{\otimes N+1} \\ &= \langle 0_{N+1}|\hat{U}_{N+1}^\dagger\langle 0|^{\otimes N}\hat{U}_N^\dagger..\hat{U}_1^\dagger\hat{A}_S\otimes\hat{E}^{\otimes N}\hat{U}_1..\hat{U}_N|0\rangle^{\otimes N}\otimes\hat{E}_{N+1}\hat{U}_{N+1}|0_{N+1}\rangle \\ &= \langle 0_{N+1}|\hat{U}_{N+1}^\dagger\left(\hat{A}_S\prod_{j=1}^N s_j\hat{\Pi}_{j,S}\right)\otimes\hat{E}_{N+1}\hat{U}_{N+1}|0_{N+1}\rangle\end{aligned} \tag{A.90}$$

Now, defining $\hat{A}_S\prod_{j=1}^N s_j\hat{\Pi}_{j,S} = \hat{A}'_S$ one can repeat the passages of (A.89) and get

$$\langle 0_{N+1}|\hat{U}_{N+1}^\dagger\hat{A}'_S\otimes\hat{E}_{N+1}\hat{U}_{N+1}|0_{N+1}\rangle = s_{N+1}\hat{A}'_S\hat{\Pi}_{N+1,S} \tag{A.91}$$

so that in conclusion

$$\langle 0|^{\otimes N+1}\hat{U}_{N+1}^\dagger..\hat{U}_1^\dagger\hat{A}_S\otimes\hat{E}^{\otimes N+1}\hat{U}_1..\hat{U}_{N+1}|0\rangle^{\otimes N+1} = \hat{A}_S\prod_{j=1}^{N+1}s_j\hat{\Pi}_{j,S} \tag{A.92}$$

Since (A.88) is true for $N=1$ and if it is for an N, it is also for $N+1$, it is necessarily valid for any finite N. In particular, when $\hat{A}_S = \mathbb{1}_S$, relation (3.27) is verified, but one can also choose $\hat{A}_S = \hat{\Pi}_{0,S}$, thus adding an additional projector to the products. In this way it is possible to measure $\langle\prod_{j=0}^N\hat{\Pi}_{j,S}\rangle$ using N couplings and one postselection on the object system.

I now show that it is possible to ignore the contribution of one of the projectors, for instance $\hat{\Pi}_l$, by measuring $\hat{E}'_l \equiv \hat{\Pi}_{0,l} + t_l\hat{\sigma}_{x,l} + t_l^2\hat{\Pi}_{1,l}$ instead of $\hat{E}_l$ on the associated pointer. Indeed, since $\hat{E}'_l\hat{U}_l|0_l\rangle = |0_l\rangle + t_l|1_l\rangle$, simply

$$\langle 0_l|\hat{U}_l^\dagger\hat{A}_S\otimes\hat{E}'_l\hat{U}_l|0_l\rangle = \hat{A}_S \tag{A.93}$$

This means that one can let the system evolve with the same unitary transformation but can choose to ignore one (or more) of the contributions with a simple change of the pointer measurements. Any subsequent observation can either concatenate a new $s_j\hat{\Pi}_{j,S}$ term to the previous $\hat{A}_S$ (when $\hat{E}_j$ is measured) or leave it as it is (when $\hat{E}'_j$ is measured). In conclusion a generalization of equation (3.29) is

$$\left\langle\prod_{j\in\Lambda}\hat{E}_j\prod_{j'\in\Lambda'}\hat{E}'_{j'}\right\rangle = \prod_{j\in\Lambda}s_j\left\langle\prod_{j\in\Lambda}\hat{\Pi}_{j,S}\right\rangle \tag{A.94}$$

Notice that at first order in small strength, $\hat{E}\approx\hat{\Sigma}$ (Section A.2), whereas $\langle 0_l|\hat{U}_l^\dagger\hat{A}_S\otimes\hat{E}'_l\hat{U}_l|0_l\rangle \approx \langle 0_l|\hat{U}_l^\dagger\hat{A}_S\otimes\hat{\Pi}_{0,l}\hat{U}_l|0_l\rangle \approx \langle 0_l|\hat{U}_l^\dagger\hat{A}_S\hat{U}_l|0_l\rangle \approx \hat{A}_S$, thus confirming the methods used in References [50, 51]. At maximum strength, instead, $\hat{E}' = \mathbb{1}+\hat{\sigma}_x = 2\hat{\Pi}_+$ can be measured with a single acquisition, as it is a multiple of a pointer projector.

A.5 Details about the Simulations

A.5.1 Generation of Random Density Matrices

In Section 3.3, the results of the various protocols are compared in different conditions of coupling strength. In order to access a large pool of initial states we randomly generated 10^4 bidimensional density operators and for each we calculated the predictions of Lundeen's DRDO method using formula (3.12) and the analytical relations reported in (A.61). We obtained the average, minimum and maximum trace distances between these results and the actual density matrices for 200 different settings of strength (Figure 3.1a) as well as the frequency of distances higher than 0.1 (Figure 3.1b). We also found the analytical results of formulas (3.18) and (3.20) (neglecting the contribution of $\frac{1}{\sqrt{N}}$) in order to estimate the standard deviation associated to each pointer mean value (Figure 3.2a) and the relative error on a matrix element (Figure 3.2b).

The only step which involves a random component is the initial generation of the true states. We followed the recipe introduced in References [61, 62] which is based on solid physical grounds. First, we imagined to couple the bidimensional ($d = 2$) Hilbert space $\mathscr{H}$ to another qubit system $\mathscr{H}_{aux}$. We generated 8 real numbers according to a normal distribution with zero mean and unit variance. They are the real and imaginary parts of the 4 coefficients that identify pure state $|\psi\rangle \in \mathscr{H} \otimes \mathscr{H}_{aux}$ expressed in a product basis:

$$|\psi\rangle = \sum_{j=1}^{d=2} \sum_{k=1}^{d_{aux}=2} c_{jk} |a_j\rangle \otimes |b_k\rangle \tag{A.95}$$

We arranged these c_{jk} in a $d \times d_{aux} = 2 \times 2$ matrix $\hat{X}$ so that $X_{jk} = c_{jk}$ which is called a *Ginibre matrix*. We then calculated the $d \times d$ operator

$$\hat{\varrho} = \hat{X}\hat{X}^\dagger \tag{A.96}$$

which coincides with the partial trace of $|\psi\rangle\langle\psi|$ on system $\mathscr{H}_{aux}$. Indeed:

$$\begin{aligned} \mathrm{Tr}_{\mathscr{H}_{aux}} \left(|\psi\rangle\langle\psi|\right) &= \sum_{j,j'=1}^{d=2} \sum_{k,k',l=1}^{d_{aux}=2} \langle b_l|c_{jk}|a_j b_k\rangle\langle a_{j'} b_{k'}|c^\star_{j'k'}|b_l\rangle = \sum_{j,j',k,k',l} \delta_{kl}\delta_{k'l} c_{jk} c^\star_{j'k'} |a_j\rangle\langle a_{j'}| \\ &= \sum_{jj'k} c_{jk} c^\star_{j'k} |a_j\rangle\langle a_{j'}| = \hat{X}\hat{X}^\dagger \end{aligned} \tag{A.97}$$

Finally we normalized $\hat{\varrho}$ with its own trace in order to end with a matrix that satisfies all the properties of a density operator.

This procedure based on the partial trace resembles the one we used in the actual experiment and assures the production of general states. The fact that $d_{aux} = d$ guarantees that the results are distributed according to the Hilbert-Schmidt metric, which induces the trace distance [63], but the choice of using qubit systems was only made to better match the experiment conditions and save computational time. Figure A.1 shows the distribution of 1000 samples in the Bloch Sphere:

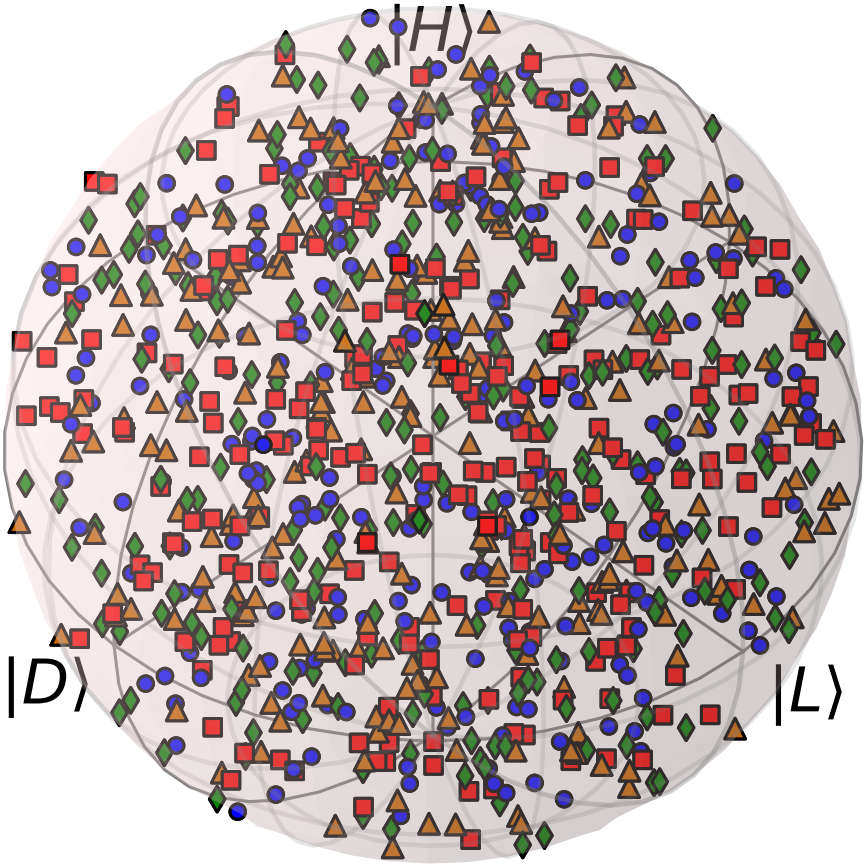


FIGURE A.1: 1000 randomly generated states.

A.5.2 Simulation of the Experiment

A completely unrelated simulation was used in connection with the experiment in order to estimate the errors associated to the trace distances of Figures 5.4, 5.6, 5.9 and 5.12. We generated 10^4 virtual repetitions of each experiment, using random counts distributed according to Poisson functions with parameters (mean and variance) coinciding with the actual measurement results. Of course this is not necessarily accurate as the outcome of a single measure is not the mean of repeated acquisitions, but is the best estimate with the available data. For each virtual dataset we performed the same analysis that we applied to the actual measurement results and extracted the trace distances between each protocol and a reference (the QST estimate or the expected pure state). The standard deviations of these distances over the 10^4 samples appear as error bars in the aforementioned figures.

Bibliography

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 1930.

[2] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics*, 1932.

[3] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart and C. Bamber, "Direct Measurement of the Quantum Wavefunction", in *Nature*, vol. 474, Jun 2011, doi: 10.1038/nature10120, Available online.

[4] J. S. Lundeen and C. Bamber, "Procedure for Direct Measurement of General Quantum States Using Weak Measurement", in *Phys. Rev. Lett.*, vol. 108, p. 070402, Feb 2012, doi:10.1103/PhysRevLett.108.070402, Available online.

[5] G. Vallone and D. Dequal, "Strong Measurements Give a Better Direct Measurement of the Quantum Wave Function", in *Phys. Rev. Lett.*, vol. 116, p. 040502, Jan 2016, doi:10.1103/PhysRevLett.116.040502, Available online.

[6] C. Cohen-Tannoudji, B. Diu and F. Laloe, *Quantum Mechanics*, Wiley-Interscience, Oct. 2006.

[7] P. A. M. Dirac, "On the Analogy Between Classical and Quantum Mechanics", in *Rev. Mod. Phys.*, vol. 17, pp. 195, Apr 1945, doi:10.1103/RevModPhys.17.195, Available online.

[8] S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda and R. Simon, "Wigner-Weyl Correspondence in Quantum Mechanics for Continuous and Discrete Systems. A Dirac–Inspired View", in *Journal of Physics A: Mathematical and General*, vol. 39, no. 6, pp. 1405, 2006, doi:10.1088/0305-4470/39/6/014, Available online.

[9] L. Cohen, "Generalized Phase-Space Distribution Functions", in *Journal of Mathematical Physics*, vol. 7, no. 5, pp. 781, 1966, doi:10.1063/1.1931206, Available online.

[10] E. Wigner, "On the Quantum Correction For Thermodynamic Equilibrium", in *Phys. Rev.*, vol. 40, pp. 749, Jun 1932, doi:10.1103/PhysRev.40.749, Available online.

[11] W. E. Brittin and W. R. Chappell, "The Wigner Distribution Function and Second Quantization in Phase Space", in *Rev. Mod. Phys.*, vol. 34, pp. 620, Oct 1962, doi:10.1103/RevModPhys.34.620, Available online.

[12] J. E. Moyal, "Quantum Mechanics as a Statistical Theory", in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 45, no. 1, p. 99124, 1949, doi:10.1017/S0305004100000487, Available online.

[13] K. Husimi, "Some Formal Properties of the Density Matrix", in *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, vol. 22, no. 4, pp. 264, 1940, doi:10.11429/ppmsj1919.22.4_264, Available online.

[14] E. C. G. Sudarshan, "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams", in *Phys. Rev. Lett.*, vol. 10, pp. 277, Apr 1963, doi:10.1103/PhysRevLett.10.277, Available online.

[15] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field", in *Phys. Rev.*, vol. 131, pp. 2766, Sep 1963, doi:10.1103/PhysRev.131.2766, Available online.

[16] A. Di Lorenzo, "Sequential Measurement of Conjugate Variables as an Alternative Quantum State Tomography", in *Phys. Rev. Lett.*, vol. 110, p. 010404, Jan 2013, doi:10.1103/PhysRevLett.110.010404, Available online.

[17] A. Di Lorenzo, "Quantum State Tomography from a Sequential Measurement of Two Variables in a Single Setup", in *Phys. Rev. A*, vol. 88, p. 042114, Oct 2013, doi:10.1103/PhysRevA.88.042114, Available online.

[18] D. F. V. James, P. G. Kwiat, W. J. Munro and A. G. White, "Measurement of Qubits", in *Phys. Rev. A*, vol. 64, p. 052312, Oct 2001, doi:10.1103/PhysRevA.64.052312, Available online.

[19] G. G. Stokes, "On the Composition and Resolution of Streams of Polarized Light from different Sources", in *Transactions of the Cambridge Philosophical Society*, vol. 9, pp. 399, 1852, Available online.

[20] 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", in *Phys. Rev. Lett.*, vol. 60, pp. 1351, Apr 1988, doi:10.1103/PhysRevLett.60.1351, Available online.

[21] R. Jozsa, "Complex Weak Values in Quantum Measurement", in *Phys. Rev. A*, vol. 76, p. 044103, Oct 2007, doi:10.1103/PhysRevA.76.044103, Available online.

[22] G. Mitchison, "Weak Measurement Takes a Simple Form for Cumulants", in *Phys. Rev. A*, vol. 77, p. 052102, May 2008, doi:10.1103/PhysRevA.77.052102, Available online.

[23] H. M. Wiseman, "Weak Values, Quantum Trajectories, and the Cavity-QED Experiment on Wave-Particle Correlation", in *Phys. Rev. A*, vol. 65, p. 032111, Feb 2002, doi:10.1103/PhysRevA.65.032111, Available online.

[24] S. Wu and K. Mlmer, "Weak Measurements with a Qubit Meter", in *Physics Letters A*, vol. 374, no. 1, pp. 34 , 2009, doi:10.1016/j.physleta.2009.10.026, Available online.

[25] L. M. Johansen, "Weak Measurements with Arbitrary Probe States", in *Phys. Rev. Lett.*, vol. 93, p. 120402, Sep 2004, doi:10.1103/PhysRevLett.93.120402, Available online.

[26] N. W. M. Ritchie, J. G. Story and R. G. Hulet, "Realization of a Measurement of a "Weak Value"", in *Phys. Rev. Lett.*, vol. 66, pp. 1107, Mar 1991, doi:10.1103/PhysRevLett.66.1107, Available online.

[27] 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", in *Phys. Rev. D*, vol. 40, pp. 2112, Sep 1989, doi:10.1103/PhysRevD.40.2112, Available online.

[28] O. Hosten and P. Kwiat, "Observation of the Spin Hall Effect of Light via Weak Measurements", in *Science*, 2008, doi:10.1126/science.1152697, Available online.

[29] P. B. Dixon, D. J. Starling, A. N. Jordan and J. C. Howell, "Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification", in *Phys. Rev. Lett.*, vol. 102, p. 173601, Apr 2009, doi:10.1103/PhysRevLett.102.173601, Available online.

[30] G. Jayaswal, G. Mistura and M. Merano, "Observation of the Imbert-Fedorov Effect via Weak Value Amplification", in *Opt. Lett.*, vol. 39, no. 8, pp. 2266, Apr 2014, doi:10.1364/OL.39.002266, Available online.

[31] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan and R. W. Boyd, "Understanding Quantum Weak Values: Basics and Applications", in *Rev. Mod. Phys.*, vol. 86, pp. 307, Mar 2014, doi:10.1103/RevModPhys.86.307, Available online.

[32] Y. Aharonov and L. Vaidman, "Properties of a Quantum System during the Time Interval between Two Measurements", in *Phys. Rev. A*, vol. 41, pp. 11, Jan 1990, doi:10.1103/PhysRevA.41.11, Available online.

[33] Y. Aharonov and L. Vaidman, *The Two-State Vector Formalism: An Updated Review*, pp. 399–447, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, doi: 10.1007/978-3-540-73473-4_13, Available online.

[34] L. Hardy, "Quantum Mechanics, Local Realistic Theories, and Lorentz-Invariant Realistic Theories", in *Phys. Rev. Lett.*, vol. 68, pp. 2981, May 1992, doi:10.1103/ PhysRevLett.68.2981, Available online.

[35] T. Ravon and L. Vaidman, "The Three-Box Paradox Revisited", in *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 11, p. 2873, 2007, Available online.

[36] B. Svensson, "Pedagogical Review of Quantum Measurement Theory with an Emphasis on Weak Measurements", in *Quanta*, vol. 2, no. 1, pp. 18, 2013, doi: 10.12743/quanta.v2i1.12, Available online.

[37] K. J. Resch, J. S. Lundeen and A. M. Steinberg, "Experimental Realization of the Quantum Box Problem", in *Physics Letters A*, vol. 324, no. 2, pp. 125 , 2004, doi:10.1016/j.physleta.2004.02.042, Available online.

[38] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm and A. M. Steinberg, "Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer", in *Science*, vol. 332, no. 6034, pp. 1170, 2011, doi:10.1126/ science.1202218, Available online.

[39] M. Schiavon, L. Calderaro, M. Pittaluga, G. Vallone and P. Villoresi, "Three-Observer Bell Inequality Violation on a Two-Qubit Entangled State", in *Quantum Science and Technology*, vol. 2, no. 1, p. 015010, 2017, Available online.

[40] E. Schroedinger, "Discussion of Probability Relations between Separated Systems", in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, no. 4, p. 555563, 1935, doi:10.1017/S0305004100013554, Available online.

[41] A. Einstein, B. Podolsky and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", in *Phys. Rev.*, vol. 47, pp. 777, May 1935, doi:10.1103/PhysRev.47.777, Available online.

[42] J. S. Bell, "On the Einstein-Podolsky-Rosen Paradox", in *Physics*, vol. 1, no. 3, pp. 195, 1964.

[43] A. Aspect, P. Grangier and G. Roger, "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities", in *Phys. Rev. Lett.*, vol. 49, pp. 91, Jul 1982, doi:10.1103/PhysRevLett.49.91, Available online.

[44] V. Coffman, J. Kundu and W. K. Wootters, "Distributed Entanglement", in *Phys. Rev. A*, vol. 61, p. 052306, Apr 2000, doi:10.1103/PhysRevA.61.052306, Available online.

[45] J. S. Lundeen and K. J. Resch, "Practical Measurement of Joint Weak Values and Their Connection to the Annihilation Operator", in *Physics Letters A*, vol. 334, no. 5, pp. 337 , 2005, doi:10.1016/j.physleta.2004.11.037, Available online.

[46] 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", in *Phys. Rev. Lett.*, vol. 117, p. 120401, Sep 2016, doi:10.1103/PhysRevLett.117.120401, Available online.

[47] J. S. Lundeen and C. Bamber, "Weak Measurement of the Dirac Distribution", , Oct 2011, Available online.

[48] J. Z. Salvail, M. Agnew, J. S. Allan, E. Bolduc, J. Leach and R. W. Boyd, "Full Characterization of Polarization States of Light via Direct Measurement", in *Nature Photonics*, vol. 7, pp. 316, Mar 2013, doi:10.1038/nphoton.2013.24, Available online.

[49] P. Zou, Z. Zhang and W. Song, "Direct Measurement of General Quantum States Using Strong Measurement", in *Phys. Rev. A*, vol. 91, p. 052109, May 2015, doi:10.1103/PhysRevA.91.052109, Available online.

[50] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa and M. Genovese, "Measuring Incompatible Observables by Exploiting Sequential Weak Values", in *Phys. Rev. Lett.*, vol. 117, p. 170402, Oct 2016, doi:10.1103/PhysRevLett.117.170402, Available online.

[51] G. Putz, T. Barnea, N. Gisin and A. Martin, "Experimental Weak Measurement of Two Non-Commuting Observables", , Oct 2016, Available online.

[52] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer and D. J. Wineland, "Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States", in *Science*, vol. 304, no. 5676, pp. 1476, 2004, doi:10.1126/science.1097576, Available online.

[53] G. Kirchmair, J. Benhelm, F. Zhringer, R. Gerritsma, C. F. Roos and R. Blatt, "Deterministic Entanglement of Ions in Thermal States of Motion", in *New Journal of Physics*, vol. 11, no. 2, p. 023002, 2009, Available online.

[54] A. J. Berkley, H. Xu, R. C. Ramos, M. A. Gubrud, F. W. Strauch, P. R. Johnson, J. R. Anderson, A. J. Dragt, C. J. Lobb and F. C. Wellstood, "Entangled Macroscopic Quantum States in Two Superconducting Qubits", in *Science*, vol. 300, no. 5625, pp. 1548, 2003, doi:10.1126/science.1084528, Available online.

[55] F. Mandl and G. Shaw, *Quantum Field Theory*, Wiley, 1984.

[56] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum and P. H. Eberhard, "Ultrabright Source of Polarization-Entangled Photons", in *Phys. Rev. A*, vol. 60, pp. R773, Aug 1999, doi:10.1103/PhysRevA.60.R773, Available online.

[57] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko and Y. Shih, "New High-Intensity Source of Polarization-Entangled Photon Pairs", in *Phys. Rev. Lett.*, vol. 75, pp. 4337, Dec 1995, doi:10.1103/PhysRevLett.75.4337, Available online.

[58] M. Schiavon, *Space Quantum Communication*, Ph.D. thesis, Università degli Studi di Padova, 2017.

[59] K. J. Resch and A. M. Steinberg, "Extracting Joint Weak Values with Local, Single-Particle Measurements", in *Phys. Rev. Lett.*, vol. 92, p. 130402, Mar 2004, doi:10.1103/PhysRevLett.92.130402, Available online.

[60] G. Mitchison, R. Jozsa and S. Popescu, "Sequential Weak Measurement", in *Phys. Rev. A*, vol. 76, p. 062105, Dec 2007, doi:10.1103/PhysRevA.76.062105, Available online.

[61] J. A. Miszczak, "Generating and Using Truly Random Quantum States in Mathematica", in *Computer Physics Communications*, vol. 183, no. 1, pp. 118 , 2012, doi:http://dx.doi.org/10.1016/j.cpc.2011.08.002, Available online.

[62] K. Zyczkowski, K. A. Penson, I. Nechita and B. Collins, "Generating Random Density Matrices", in *Journal of Mathematical Physics*, vol. 52, no. 6, p. 062201, 2011, doi:10.1063/1.3595693, Available online.

[63] K. Zyczkowski and H. J. Sommers, "Induced Measures in the Space of Mixed Quantum States", in *Journal of Physics A: Mathematical and General*, vol. 34, no. 35, p. 7111, 2001, Available online.