Contents

Introduction

It from bit

John A. Wheeler

David Deutsch's motto "Information is physical", utmost famous among quantum information theorists, has been rightly charged of semantic naturalism ([Ti04]), or, translating the philosophical jargon, of being an attempt to reduce meanings to facts. On the other hand, John A. Wheeler's "dual" programme, reassumed by the quotation in epigraphe, of deriving physical properties from information-theoretic notions (a sort of "physics is informational"), though appealing, also suffers from conceptual ambiguity.

Despite these difficulties, the discovery of the deep connection between Quantum Mechanics and Information Theory, and the subsequent institution of a new "Quantum Information Theory", has been growingly recognized as one of the major achievements in the understanding of the physical world, a *callida iunctura* capable of shedding light on both fields by putting them side by side, and opening new perspectives on many foundational issues. For example, the Bayesian interpretation of probability has received new life from Quantum Information [CFS02] [Fu02], as well as algorithmic complexity theory and cryptography, just to name a few.

A prominent example of the fruitfulness of this approach, has been the progressive - but not yet concluded - clarification of the long-standing interpretational problems of Heisenberg's Uncertainty Principle. Although the venerable phrase "Uncertainty Principle" is never mentioned in the seminal paper ([He27]) [1], Heisenberg hoped, at the beginning at least, that the "indeterminacy relation" (as he always calls it), would eventually serve as a foundational basis from which the formal structure of quantum mechanics could be derived, much like Einstein's theory is a straightforward consequence of the relativity principle.

In his article, Heisenberg presented formal arguments, along with a bunch of diverse physical examples, aiming at the quantification of a theoretical limit on the precision attainable in any measurement of conjugate observables, say the position q and the momentum p of a particle. He obtained the expression

$$\Delta q \Delta p \gtrsim h \tag{1}$$

where h is Planck's constant. According to Heisenberg, the physical interpretation of this inequality is the following: whenever we measure the position of a particle with the precision Δq, we perturb the momentum at least by $\Delta p \sim h/\Delta q$. The well-known γ-ray microscope *Gedankenexperiment* was intended to be emblematic of this interpretation: the position of a particle is reaveled by the scattering of a photon with increasing precision as the photon's energy increases. At the same time, the impact of higher-energy photons causes greater variations of particle's momentum.

However, as soon spotted by Bohr (see for example the comments in [WZ83]), Heisenberg's derivation of equation (1) was sloppy and relied on a semiclassical vision of particles and apparatuses. In the following years,

[1] Actually, it was coined by Ruark [Ru28].

the mathematical part of his proof was refined by Kennard [Ke28], who explicited the Fourier transform-based argument in Heisenberg's paper, and by Robertson [Ro29] [Ro34], who generalized the proof to every pair of non-commuting observables. Robertson's inequality

$$\Delta A_\psi \Delta B_\psi \geq \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle| \tag{2}$$

where A and B are observables and $|\psi\rangle$ is a quantum state, soon became, and nowadays continues to be, the standard mathematical reference of the Uncertainty Principle.

However, by this time the interpretation changed: since ΔA_ψ and ΔB_ψ are the standard deviations of the result of measurements performed on a system in the state ψ, equation (2) tells us that there is a limit in our ability to *predict* the outcome of an experiment: if we know that, on some system in the state $|\psi\rangle$ the result of the measurement of an observable A can be predicted with precision ΔA_ψ, the result of the measurement of a non-commuting observable B cannot be predicted with precision greater than ΔB_ψ. This perspectival shift towards a statistical interpretation was accepted by Heisenberg himself in his Chicago lectures [He30], but the original "information/disturbance" view, though incorrect, continued to be widespread.

It was not until the seminal paper by Fuchs and Peres [FP96] that this problem was correctly addressed, this time in the framework opened by Quantum Communication Theory. They considered the most elementary two-party communication protocol: Alice sends a message to Bob, codifying it on a quantum system. After receiving it, Bob can communicate with Alice and check if he effectively obtained the message sent by Alice. Fuchs and Peres showed that as an eavesdropper, "Eve", tries to read the message (to extract "information"), Alice and Bob can catch the intrusion, verifying that Alice's message was different from what Bob obtained. It turns out that the

more information Eve extracts, the more Alice and Bob are likely to become aware of the eavesdropping. In a word, the extraction of information perturbs the system.

Many more different scenarios were subsequently analyzed ([Fu98] [BD01] [Ba00] [BS06] just to name a few), varying the kind of states, the protocols, the number of parties, with the same qualitative result. Our work inserts in this research line: in chapter 4 we present a simple cryptographic-like protocol and derive the information/disturbace tradeoff curves.

Inspired by [Da03], in chapter 5 we also analyze the different context of probabilistic state transformation. Here we want to transform some states in a prescribed way, with good probability. We will show that in this case the probability of a transformation and its "quality" (in a sense precised in the text) again satisfy a tradeoff relation.

As side results, we obtain the isofidelity surfaces for a qubit (chapter 3), and a proof of the Jaeger-Shimony bound, somehow different from the original one (chapter 4).

Chapter 1

Formulation of QM

1.1 Quantum states

Every quantum mechanical system is described in terms of a separable complex Hilbert space $\mathcal{H}$, and of the operators acting on it. In the following the dimension of $\mathcal{H}$ is always assumed to be finite.

The *state* of the system is represented by a *density operator*.

Definition 1.1.1 (Density operator). *A* density operator *is trace-class operator ρ satisfying two conditions:*

1. *Positivity:* $\rho \geq 0$.

2. *Normalization:* $\mathrm{Tr}(\rho) = 1$.

The set of density operators is indicated with $\mathcal{S}(\mathcal{H})$.

These requirements are the quantum analogue of the conditions imposed on a classical probability distribution $p(x)$: $p(x) \geq 0$, $\sum_x p(x) = 1$.

Given a set of density operators $\{\rho_i\}$, the convex combination

$$\rho = \sum_i p_i \rho_i \tag{1.1}$$

with $p_i \geq 0$, $\sum_i p_i = 1$, is obviously positive and normalized. Density operators, thus, constitute a convex set. The extremal points are rank-one projectors of the form $\rho = |\psi\rangle\langle\psi|$ for some normalized $\psi \in \mathcal{H}$ and are called *pure states*. Non-extremal points are called *mixed states*.

A measurement performed on a quantum system is described by the probability space Ω of possible outcomes, and a mathematical object called Positive Operator Valued Measurement (POVM).

Definition 1.1.2 (POVM). *Let Ω be a probability space and $\sigma(\Omega)$ its σ-algebra of events. A* POVM *is a set of operators $\{\Pi_\Delta\}_{\Delta\in\sigma(\Omega)}$ such that*

1. *Π_Δ is positive for every $\Delta \in \sigma(\Omega)$.*
2. *$\Pi_{\cup_n \Delta_n} = \sum_n \Pi_{\Delta_n}$ for every countable set of pairwise disjoint events $\{\Delta_n\}$.*
3. *$\Pi_\emptyset = 0$ and $\Pi_\Omega = I$.*

The probability p_Δ of the event Δ, when the system is in the state ρ, is given by the *Born's statistical formula*:

$$p_\Delta = \mathrm{Tr}(\rho\Pi_\Delta). \tag{1.2}$$

The conditions imposed by definition 1.1.2 guarantee that $p_\Delta \geq 0$ and $p_\Omega = 1$.

When the probability space is discrete ($\Omega = \{1, 2, ...\}$) we can simply assign a positive operator Π_k for each atomic event $k \in \Omega$, with the constraint $\sum_k \Pi_k = I$. The whole POVM is then defined by extension through the second condition in definition 1.1.2.

The particular case in which the operators Π_k are orthogonal projectors, is called PVM (Projection Valued Measurement), and corresponds to the measurement of an *observable*. According to standard formulations of QM,

an observable is an Hermitian operator A, usually constructed from some classical quantity via the correspondence principle. Its spectral decomposition is $A = \sum_n a_n P_n$, where P_n is the projector on the eigenspace relative to the eigenvalue a_n. The spectrum of A gives the possible outcomes of a measurement, with the following probability distribution:

$$p(a_n) = \mathrm{Tr}(\rho P_n).$$

In this sense POVMs are a generalization of standard projective measurements and thus are also known simply as *generalized measurements*.

Born's formula helps in clarifying the operational meaning of convex combinations like (1.1). In fact, computing a probability for some operator Π over the state ρ we have

$$\mathrm{Tr}(\rho\Pi) = \mathrm{Tr}\left(\sum_i p_i \rho_i \Pi\right) = \sum_i p_i \,\mathrm{Tr}(\rho_i \Pi).$$

The last member gives the same quantity computed for a *mixture* of states ρ_i, sampled with probability p_i. The mixture is therefore represented by the convex combination (1.1).

1.1.1 Orthogonality and classicality

Quantum mechanics is an essentialy probabilistic theory and, to some extent, it can be regarded as a "non-commutative" generalization of classical probabilistic calculus, where operators have replaced probability distributions. Classical expressions can be recovered from quantum formalism choosing once and for all an orthonormal basis, and using only diagonal operators on that basis.

Let $\{\psi_n\}_{n=1\ldots N}$ be an o. n. basis for an N-dimensional system associated

to $\mathcal{H}$. The density operators diagonal on this basis are of the form

$$\rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|, \quad \sum_n p_n = 1.$$

They are in bijective correspondence with the N-simplex of classical probability distributions $\{p_n\}$ over the N pure states $|\psi_n\rangle\langle\psi_n|$. Let $A = \sum_n a_n |\psi_n\rangle\langle\psi_n|$ be another diagonal operator. From the trace formula (1.2) we have

$$\mathrm{Tr}(\rho A) = \sum_n p_n a_n,$$

i. e. the classical expectation value of the random variable $A(n) := a_n$. Of course, since there are many different o. n. bases for a given Hilbert space, there are correspondingly many ways to embed a classical probability space into a quantum one.

These examples, though simple, already allow one to grasp the deep relationship between orthogonality and classicality. In order to better understand this link, it is useful to introduce the concept of *distinguishability* (which will be discussed in more detail in chapter 4): given two orthogonal states $\rho_1 = |\psi_1\rangle\langle\psi_1|$ and $\rho_2 = |\psi_2\rangle\langle\psi_2|$, $\langle\psi_1|\psi_2\rangle = 0$ and choosing, for example, $\Pi = \rho_1 = |\psi_1\rangle\langle\psi_1|$, form equation (1.2) we have

$$\mathrm{Tr}(\rho_1 \Pi) = 1$$
$$\mathrm{Tr}(\rho_2 \Pi) = 0.$$

When the system is in the state ρ_1, the outcome of the "question" formalized with the operator Π is always positive. On the other hand, the outcome is negative when the state is ρ_2. An experimental setup can thus be designed, which allows to perfectly distinguish between the states ρ_1 and ρ_2.

In the classical simplex of probability distributions, extremal points are always orthogonal and distinguishable in principle. They constitute a discrete

set endowed with the structure of Boolean algebra which is indipendent of any probabilistic concern. So to speak, one can separate the "logic" structure away from the probability space. In a quantum system, on the other hand, pure states in general have a non-zero overlap: they always form a continuous manifold, with no indipendent logic underlying it. In this sense, quantum mechanics is essentialy probabilistic.

1.2 Quantum operations

Every physical evolution of a quantum system must be reflected in a transformation of the associated density operator. Formally, the transformation is realized by a superoperator μ:

$$\rho \longmapsto \mu(\rho).$$

However, only a subset of all the conceivable superoperators realizes a physical transformation. The elements of this set are called *quantum operations* (QOs). They include both the dynamical evolution typical of closed systems and the discontinuous state reduction subsequent to a measurement. In the following, we will present three different ways of characterizing QOs:

1. Axiomatizing the properties required for a superoperator in order to respect the structure of the set of quantum states $\mathcal{S}(\mathcal{H})$.

2. Isolating some fundamental operations (extension of the system, unitary evolution, projective measurements, partial traces) and combining them to build the whole set of possible transformations.

3. Exploiting the *Kraus representation*, more technical though extremely useful in practical application.

1.2.1 Axiomatic characterization

A Quantum Operation is a superoperator $\mu : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ satisfying the following conditions:

1. Affinity: $\mu\left(\sum_n p_n \rho_n\right) = \sum_n p_n \mu(\rho_n)$ for every convex combination of states $\sum_n p_n \rho_n$.

2. Complete positivity (CP): for any Hilbert space $\mathcal{K}$ the map $\mu \otimes \mathcal{I} : \mathcal{S}(\mathcal{H} \otimes \mathcal{K}) \to \mathcal{S}(\mathcal{H} \otimes \mathcal{K})$, where $\mathcal{I}$ is the identity on $\mathcal{S}(\mathcal{K})$, is a positive operator.

3. Trace decreasing (TD): for any operator ρ we have $\mathrm{Tr}[\mu(\rho)] \leqslant \mathrm{Tr}[\rho]$.

Affinity is needed because the QO must respect the convex structure of the set of states $\mathcal{S}(\mathcal{H})$. Since every affine map can be extended to a linear map relaxing the constraint $\sum_n p_n = 1$, we can equivalently impose the linearity.

The complete positivity condition is needed because density operators are positive, and positivity must be conserved after any evolution. The condition grants that for every Hilbert space $\mathcal{K}$ and every bipartite state $\sigma \in \mathcal{S}(\mathcal{H} \otimes \mathcal{K})$, the operator after the joint evolution $\sigma' = (\mu \otimes \mathcal{I})(\sigma)$ is still a valid density operator.

Since density operators have unit trace, the trace decreasing condition becomes

$$\mathrm{Tr}[\mu(\rho)] \leqslant 1, \tag{1.3}$$

allowing us to interpret the trace $\mathrm{Tr}[\mu(\rho)]$ as the probability that the quantum operation μ wll happen, given that the system is initially in the state ρ. After the evolution the system is left in the state

$$\rho' = \frac{\mu(\rho)}{\mathrm{Tr}[\mu(\rho)]}.$$

This rule is known as *state reduction*.

A quantum operation which is also trace preserving (TP), is called *quantum channel*, and obviously it occurs with certainty. It is the most general deterministic (but not necessarily reversible) transformation of an open system.

1.2.2 Ancillary extensions

There is a fundamental dilation theorem for CP maps proved by Stinespring [St55], which makes use of C^*-algebraic formalism. We present it in a form suited for our formalism (see for example [NC00]):

Theorem 1.2.1. *A map $\mu : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is a quantum operation if and only if there is an ancillary system $\mathcal{K}$ such that*

$$\mu(\rho) = \mathrm{Tr}_{\mathcal{K}} \left[U(\rho \otimes |0\rangle\langle 0|)U^\dagger P \right] \tag{1.4}$$

where $|0\rangle$ is some pure state on $\mathcal{K}$, U a unitary operator on $\mathcal{H} \otimes \mathcal{K}$, P a projector on $\mathcal{K}$, and $\mathrm{Tr}_{\mathcal{K}}$ the partial trace on $\mathcal{K}$.

We observe that for trace preserving maps the projector P becomes the identity I. The ancillary system $\mathcal{K}$ is sometimes referred to as the *environment*.

1.2.3 Kraus form

In order to express a quantum operation μ in a simple and explicit form, we can use the following theorem [Kr83]:

Theorem 1.2.2 (Kraus). *A map μ is CP if and only if there is a set of operators $\{M_n\}$ such that $\mu(\rho) = \sum_n M_n \rho M_n^\dagger$.*

The operators M_n (known as *Kraus operators*) are not uniquely determined. In fact, for every rectangular matrix T satisfying the condition $T^\dagger T = I$ the set $\{N_j\}$ given by $N_j = \sum_i T_{ji} M_i$ is also a valid Kraus representation of μ. The *rank* of the quantum operation is the minimum number of operators needed to construct a Kraus representation.

The trace decreasing condition becomes

$$\sum_n M_n^\dagger M_n \leq I. \tag{1.5}$$

and the equality is satisfied if and only if the map is trace preserving.

1.3 Instrument

A single QO represents one possible evolution of the density operator; but in a measurement process a system can undergo different evolutions, depending on the outcomes of the experiment. In order to describe these processes, we have to assign a CP map for every possible event. This is exactly what is done by the *instrument*. Given a probability space Ω and the σ-algebra of the events $\sigma(\Omega)$, an instrument is a map $\mathcal{E}$ defined on $\sigma(\Omega)$ and taking values in completely positive superoperators on $\mathcal{S}(\mathcal{H})$, satistfying the following conditions:

1. $\mathcal{E}_{\cup_n \Delta_n} = \sum_n \mathcal{E}_{\Delta_n}$ for every countable set of pairwise disjoint events $\Delta_n \in \sigma(\Omega)$.

2. $\mathcal{E}_\emptyset = 0$.

3. $\mathcal{E}_\Omega$ is trace perserving.

The interpretation follows from what we said about QOs: the probability of the event Δ when the state of the system is ρ is given by $p(\Delta|\rho) =$

$\mathrm{Tr}[\mathcal{E}_\Delta(\rho)]$, and the system is left in the state

$$\rho' = \frac{\mathcal{E}_\Delta(\rho)}{\mathrm{Tr}[\mathcal{E}_\Delta(\rho)]}. \tag{1.6}$$

We can also introduce the Kraus representation for an instrument, assigning a set of Kraus operators $\{M_{\Delta,n}\}_{n\in S_\Delta}$ to every CP map corrisponding to each event Δ:

$$\mathcal{E}_\Delta(\rho) = \sum_{n\in S_\Delta} M_{\Delta,n}\rho M_{\Delta,n}^\dagger. \tag{1.7}$$

1.4 QOs and POVMs

The formalism described in the previous sections allows us to determine the probability of occurrence of a QO and the state in which the system is left. The former is given by $\mathrm{Tr}[\mu(\rho)]$, that is, exploiting Kraus representation and the cyclic invariance of the trace,

$$p(\mu|\rho) = \mathrm{Tr}\left[\rho \sum_n M_n^\dagger M_n\right]. \tag{1.8}$$

If we are interested only in probabilities, we see that in order to calculate them we only need to know the operator $\Pi = \sum_n M_n^\dagger M_n$. This operator is positive by construction, since we have, for every $|\psi\rangle \in \mathcal{H}$

$$\langle\psi|\Pi|\psi\rangle = \langle\psi|\sum_n M_n^\dagger M_n|\psi\rangle = \sum_n \|M_n\psi\|^2 \geq 0. \tag{1.9}$$

Analogously, in the case of an instrument, the probability distribution can be rewritten as

$$p(\Delta|\rho) = \mathrm{Tr}(\rho\Pi_\Delta) \tag{1.10}$$

where $\Pi_\Delta = \sum_{n\in S_\Delta} M_{\Delta,n}^\dagger M_{\Delta,n}$.

Again, the set of operators $\{\Pi_\Delta\}$ is enough to obtain the probabilities. Consistency is granted by the σ-additivity property, inherited by the operators $\{\Pi_\Delta\}$ from the definition of instrument. Precisely, we have that

$$\mathrm{Tr}\left[\rho\Pi_{\cup_n\Delta_n}\right] = \mathrm{Tr}\left[\mathcal{I}_{\cup_n\Delta_n}(\rho)\right] = \mathrm{Tr}\left[\sum_n \mathcal{I}_{\Delta_n}(\rho)\right] = \mathrm{Tr}\left[\rho\sum_n \Pi_{\Delta_n}\right] \tag{1.11}$$

and since this is valid for every density operator ρ it follows that $\Pi_{\cup_n\Delta_n} = \sum_n \Pi_{\Delta_n}$. In the same way, we have the normalization $\Pi_\Omega = I$.

This shows that QOs are compatible with POVMs. To some extent, the path from QOs to POVMs can be reverted. Taking for example the simple case of discrete probability space, we are given a POVM $\{\Pi_k\}$ and we want to find a corresponding QO for each operator Π_k. To this end, consider the operator

$$M_k = U_k\Pi_k^{1/2} \tag{1.12}$$

where U_k is any unitary operator. The square root of Π_k always exists thanks to positivity. We can easily see that $\Pi_k = M_k^\dagger M_k$ and therefore each M_k is a valid Kraus operator for the QO

$$\mu_k(\rho) = M_k\rho M_k^\dagger \tag{1.13}$$

from which we can obtain the state after the occurence of the outcome k. Obviously, the operators M_k are not uniquely determined, since we can freely change the unitary U_k without affecting the POVM. This is due to the fact that for a given POVM there are different possible realizations of the measuring apparatus, each one leaving the system in a different final state.

1.5 Ensembles of states

Given some system Q, we will indicate with

$$\mathsf{E} = \{p_i, \rho_i^Q\} \tag{1.14}$$

the ensemble of states ρ_i^Q, distributed with probabilities p_i. The corresponding (unique) density operator is given by $\rho_{\mathsf{E}}^Q = \sum_i p_i \rho_i$. We note that knowing ρ_{E}^Q is perfectly equivalent to the knowledge of the ensemble E, as far as we are interested only in prediction about the behavior of Q (as follows directly from what we said in sections 1.1 and 1.2). However, the correspondence

$$\mathsf{E} \longrightarrow \rho_{\mathsf{E}}^Q \tag{1.15}$$

is not injective, since each decomposition of ρ_{E}^Q as convex combination $\rho_{\mathsf{E}}^Q = \sum_n q_n \tau_n^Q$ gives an ensemble $\mathsf{F} = \{q_n, \tau_n^Q\}$ with the obvious property $\rho_{\mathsf{F}}^Q = \rho_{\mathsf{E}}^Q$.

Thus, ρ_{E}^Q is not enough to identify E. This ambiguity follows from the fact that an ensemble is an attempt to represent as *perfectly distinguishable* objects which actually are not so, insofar as an ensemble is an attribution of a classical probability distribution over a set of non classical objects. In other words, ensembles are a semiclassical attempt to treat quantum states in a classical way. To some extent, this amounts to think of "being in one of the states of E" as a *property* of the system; but in Quantum Mechanics properties have to do with projectors and orthogonality, i.e. with *observables*. Of course, this does not cause any trouble if everything is managed consistently, but it would be nice to describe ensembles without the necessity of notations like (1.14), which are not included in stardard quantum mechanical formalism.

In order to obtain some insights, it is useful to imagine *how* ensembles pop up in real lab experiments. One can easily convince oneself that in any experimental set up, an ensemble comes always from a (controlled) *preparation* of some system (in our example Q) by one party A ('Alice' or 'Ancilla'), which we assume to be a classical object. After the preparation, the compound system $A + Q$ is in a state of the following form:

$$\rho^{AQ} = \sum_i p_i P_i^A \otimes \rho_i^Q. \tag{1.16}$$

Here P_i^A is the state represented by the projector $P_i^A = |i\rangle\langle i|$, where $\{|i\rangle\}_i$ is a set of orthogonal states of A. Consistently, we have that

$$\mathrm{Tr}_A(\rho^{AQ}) = \rho_{\mathsf{E}}^Q, \tag{1.17}$$

so the state ρ^{AQ} is indeed an extension of ρ_{E}^Q. The classicality of A is reflected in the orthogonality of the her possible states P_i^A.

Moreover, the probability distribution for the outcomes of the projective measurement on A given by the orthogonal projectors P_i^A is exactly the ensemble probability distribution:

$$p_i = \mathrm{Tr}_A(P_i^A \rho^A), \tag{1.18}$$

where $\rho^A = \mathrm{Tr}_Q(\rho^{AQ})$. This can be interpreted as "Alice is in the state labelled with i" or "Alice decided to prepare Q in the state labelled with i".

The introduction of an ancillary system used to express the implied "orthogonality" of the ensemble may appear as a gratuituous formal trick and, actually, such an explicit characterization is seldom needed. Anyway, a simple but interesting phenomenon can be illustrated resorting to this representation.

Suppose we succeded in applying some quantum operation $\mathcal{E}$ on Q. It is tempting now to apply Bayes' theorem to 'upgrade' our 'knowledge' of the ensemble. This leads directly to the following formula:

$$p_i' = p(i|\mathcal{E}) = \frac{p(\mathcal{E}|i)p_i}{p(\mathcal{E})}. \tag{1.19}$$

Interestingly, this upgrade is reflected exactly in the probability distribution of the PVM $\{P_i^A\}$ on A after the state reduction:

$$\rho^{AQ'} = \frac{(\mathcal{I}\otimes\mathcal{E})(\rho^{AQ})}{\mathrm{Tr}[(\mathcal{I}\otimes\mathcal{E})(\rho^{AQ})]} = \frac{\sum_i p_i P_i^A \otimes \mathcal{E}(\rho_i^Q)}{\mathrm{Tr}[\mathcal{E}(\rho_{\mathsf{E}}^Q)]}. \tag{1.20}$$

In fact, posing as usual $\rho^{A'} = \mathrm{Tr}_Q(\rho^{AQ'})$, which we can easily obtain

$$\mathrm{Tr}(P_i^A \rho^{A'}) = \frac{\mathrm{Tr}[\mathcal{E}(\rho_i^Q)]p_i}{\mathrm{Tr}[\mathcal{E}(\rho_{\mathsf{E}}^Q)]} = \frac{p(\mathcal{E}|i)p_i}{p(\mathcal{E})}, \tag{1.21}$$

and, consistently with the naïve application of Bayes' theorem (1.19), $p_i' = \mathrm{Tr}(P_i^A \rho^{A'})$. This analysis shows that to some extent dealing with an ensemble on Q actually amounts to deal with properties of its preparer A.

It is worth noting that we needed to enlarge the system, including degrees of freedom from an ancillary system, which we suppose our initial system Q had interacted with. This appears to be a general fact: whenever we need to consider the whole *context* of a given system, we have to extend its state in order to take into account correlations which could have been generated in past interactions.

For example, this happens when we try to find out conditions guaranteeing the reversibility of some quantum channel. The solution to this problem ([SN96]) can be roughly stated as follows. A channel $\mathcal{T}$ acting on system Q is perfectly invertible on the support of some density operator ρ^Q if and only if

$$I(A' : E') = 0 \tag{1.22}$$

where A is a system introduced to purify ρ^Q and E is the environment exploited to realize the unitary dilation of $\mathcal{T}$ (see theorem 1.2.1). $I(A' : E')$ is the quantum mutual information (see for example [NC00]) between A and E *after* the action of $\mathcal{T}$. It measures the total degree of correlation between the two systems.

As before, we need to extend Q with an ancilla A, but this time we have to use the "strongest" extension represented by the purification of ρ^Q instead of the mixed state (1.16). This happens because error correction is required to be perfect on *all* the states in the support of ρ^Q. On the other hand, when

we are treating an ensemble E, we are concerned only with the particular states in the support of ρ_{E}^{Q} which compose E.

Chapter 2

Projective spaces and QM

In the classical formulation of QM, a state is defined to be a vector ψ in the Hilbert space, with unit norm

$$\|\psi\| = 1.$$

Normalization is required for a consistent probabilistic interpretation of the vector. Since every observable quantity is ultimately computed with an expression of the form

$$\langle\psi|A|\psi\rangle,$$

where A is an Hermitian operator (see Born's formula 1.2 for pure states), it follows that vectors differing only by a phase factor give the same predictions. This is usually phrased as "overall phases are unobservable".

The conditions imposed on elements of a Hilbert space (normalization, irrelevance in the choice of the phase) are elegantly accounted for by projective spaces [1].

[1] For a general reference on the topics discussed in this chapter see [BZ06].

2.1 Projective Hilbert spaces

The action of the multiplicative group $\mathbb{C}^* = \mathbb{C} - \{0\}$ on $\mathcal{H}$ defines an equivalence relation:

$$\psi \sim \phi \Longleftrightarrow \exists \lambda \in \mathbb{C}^*, \psi = \lambda\phi. \tag{2.1}$$

Definition 2.1.1. *The projective Hilbert space is the quotient space of* $\mathcal{H}$ *with respect to the relation* $\sim$*:*

$$\mathsf{P}\mathcal{H} := \frac{\mathcal{H}}{\sim} = \frac{\mathcal{H}}{\mathbb{C}^*}. \tag{2.2}$$

The quotient defines a natural projection $\pi_\sim$*:*

$$\begin{aligned} \pi_\sim : \mathcal{H} &\longrightarrow \mathsf{P}\mathcal{H} \\ \psi &\longmapsto [\psi]. \end{aligned} \tag{2.3}$$

The elements of $\mathsf{P}\mathcal{H}$ *are called* rays.

Equivalently, $\mathsf{P}\mathcal{H}$ is defined to be the set of one-dimensional subspaces of $\mathcal{H}$ [2].

Since subspaces of a vector space V can be put in bijective correspondence with projection operators

$$W \subseteq V \longleftrightarrow P_W,$$

rays in $\mathsf{P}\mathcal{H}$ can be naturally identified with one-dimensional projectors $\rho = |\psi\rangle\langle\psi|$, $\|\psi\| = 1$. The projective space $\mathsf{P}\mathcal{H}$ is thus embeddable in an operator space, whose linear structure is inherited by $\mathsf{P}\mathcal{H}$. This structure grants the

[2] This is the simplest example of more general constructions known as *Grassmannians*. The Grassmannian $\mathbb{G}(n,k)$ is the collection of k-dimensional subspaces of $\mathbb{C}^n$. Grassmannians are homogeneous spaces of unitary groups and thus have a natural structure of differentiable manifold.

possibility of forming convex combinations, thus allowing to recover the whole set of states.

A projective space can be endowed with a metric in the following way. Let $\mathcal{R}_\psi = [\psi]$ and $\mathcal{R}_\phi = [\phi]$ be rays in $\mathsf{P}\mathcal{H}$. Since the expression given by

$$\mathcal{R}_\psi \cdot \mathcal{R}_\phi := \frac{|\langle\psi|\phi\rangle|}{\|\psi\| \cdot \|\phi\|} \tag{2.4}$$

is indipendent of the choice of the vectors representing the rays, it yields a well-defined real-valued product on $\mathsf{P}\mathcal{H}$. From this product we can define the *Fubini-Study distance* D_{FS} as follows

$$D_{FS}(\mathcal{R}_\psi, \mathcal{R}_\phi) := \arccos \mathcal{R}_\psi \cdot \mathcal{R}_\phi. \tag{2.5}$$

Computing this expression with $|\phi\rangle = |\psi + \delta\psi\rangle$ and keeping only the leading terms yields the *Fubini-Study metric*

$$ds^2 = \frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\delta\psi|\psi\rangle\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle^2} \tag{2.6}$$

which endows $\mathsf{P}\mathcal{H}$ with the structure of Riemannian manifold.

2.2 Projective groups

In the passage from the Hilbert space to its projective version, the behavior of the operators acting on $\mathcal{H}$ is also affected. Let A be any such operator

$$A : \mathcal{H} \to \mathcal{H}. \tag{2.7}$$

Clearly $\psi \sim \phi \Rightarrow A\psi \sim A\phi$ and thus the action on $\mathsf{P}\mathcal{H}$

$$\begin{aligned} A : \mathsf{P}\mathcal{H} &\longrightarrow \mathsf{P}\mathcal{H} \\ [\psi] &\longmapsto [A\psi] \end{aligned} \tag{2.8}$$

is well-defined. However, two operators A, B, differing by a complex factor

$$A = \lambda B, \quad \lambda \in \mathbb{C}^* \tag{2.9}$$

induce the same transformation on $\mathsf{P}\mathcal{H}$

$$A[\psi] = [A\psi] = [\lambda B\psi] = [B\psi] = B[\psi], \quad \forall [\psi] \in \mathsf{P}\mathcal{H}. \tag{2.10}$$

To remove this ambiguity from the action of a group of operators, the action of $\mathbb{C}^*$ needs to be quotiented away. The most general case is the general linear group of $\mathcal{H}$

$$\mathcal{H} \overset{GL(\mathcal{H})}{\longrightarrow} \mathcal{H}$$

which is replaced by it projective version $PGL(\mathcal{H}) = GL(\mathcal{H})/\mathbb{C}^*$:

$$\mathsf{P}\mathcal{H} \overset{PGL(\mathcal{H})}{\longrightarrow} \mathsf{P}\mathcal{H}.$$

Since in QM we are especially interested in unitary groups, we have to specialize the previous considerations to the n-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^n$ and analyze the group $U(n) \subset GL(\mathbb{C}^n)$.

We premise the following result:

Lemma 2.2.1. *Let G be a group, H a subgroup of G and K a normal subgroup of G. Then the following diagram is commutative*

$$\begin{array}{ccc} H & \longrightarrow & G \\ \downarrow & & \downarrow \\ H/(H \cap K) & \longrightarrow & G/K. \end{array}$$

Proof. It is a consequence of the homomorphism theorem (see for example [Her75]) applied to the compound homomorphism

$$H \hookrightarrow G \to G/K,$$

whose kernel is easily seen to be $H \cap K$. In fact, an element of H is mapped to the identity of G/K if and only if it also belongs to K. □

This lemma shows how the quotient is reflected on a subgroup. Posing $G = GL(\mathbb{C}^n)$, $H = U(n)$ and $K = \mathbb{C}^*$ (identifying $\mathbb{C}^*$ with the scalar operators multiple of identity λI, $\lambda \in \mathbb{C}^*$) the lemma yields

$$\begin{array}{ccc} U(n) & \longrightarrow & GL(\mathbb{C}^n) \\ \downarrow & & \downarrow \\ U(n)/(U(n)\cap\mathbb{C}^*) & \longrightarrow & PGL(\mathbb{C}^n). \end{array}$$

Groups of the form

$$PU(n) := \frac{U(n)}{U(n)\cap\mathbb{C}^*} = \frac{U(n)}{U(1)} \cong \frac{SU(n)}{\mathbb{Z}_n} \tag{2.11}$$

are known as *projective unitary groups* and constitute the natural action on a projective space induced by unitary transformations. The quotient by the group of phases $U(1)$ is again interpreted as irrelevance of overall phases.

2.3 The projective qubit

The two-dimensional case $\mathcal{H} = \mathbb{C}^2$ deserves a particular attention because the projective objects introduced in previous sections become quite simple and allow a straightforward geometrical interpretation.

2.3.1 Bloch sphere representation

Let $\{|0\rangle, |1\rangle\}$ be a o. n. basis. Any unit vector $|\psi\rangle$ can be represented as

$$|\psi\rangle = e^{i\lambda}\left(\cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\right) \tag{2.12}$$

with $\phi \in [0, 2\pi)$ and $\theta \in [0, \pi]$. In the corresponding density operator the global phase factor disappears

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & 1-\cos\theta \end{pmatrix}. \tag{2.13}$$

Introducing real parameters $x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$, $z = \cos\theta$ we have

$$\rho = \frac{I + \boldsymbol{r}\cdot\boldsymbol{\sigma}}{2} \tag{2.14}$$

where $\boldsymbol{r} = (x, y, z)$ is a unit real vector and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.15}$$

The set of pure states is thus represented by the unit sphere $S^2 = \{\boldsymbol{r} \in \mathbb{R}^3, |\boldsymbol{r}| = 1\}$, which is known as *Bloch sphere*[3]. It is an explicit representation of the projective space

$$\mathsf{P}\mathbb{C}^2 = S^2. \tag{2.16}$$

Since the eigenvalues of ρ are

$$\lambda_\pm = 1 \pm |\boldsymbol{r}| \tag{2.17}$$

the positivity constraint $\rho \geq 0$ becomes $|\boldsymbol{r}| \leq 1$. Inner points of the sphere are thus valid density operators. They constitute the convex hull of the points in S^2 and represent mixed states.

We note that the determinant of ρ is

$$\det(\rho) = \frac{1 - |\boldsymbol{r}|^2}{4}. \tag{2.18}$$

The norm of $\boldsymbol{r}$ is thus a function of $\det(\rho)$

$$|\boldsymbol{r}| = \sqrt{1 - 4\det(\rho)}. \tag{2.19}$$

[3] *Riemann sphere* in mathematical literature.

2.3.2 Unitary transformations

In this case the projective group acting on the Bloch sphere is

$$\frac{U(2)}{U(1)} \cong \frac{SU(2)}{\mathbb{Z}_2} \cong SO(3). \tag{2.20}$$

In the following we give an explicit representation of the action of $SU(2)$ on the Bloch sphere. The transformation

$$\rho \longmapsto \rho' = U\rho U^\dagger \tag{2.21}$$

preserves the determinant. By equation (2.19) it also follows that the norm of the Bloch vector is preserved. The map induced by U is therefore an orthogonal transformation $O(3)$ of the Bloch sphere, and it could be shown that actually it is a $SO(3)$ rotation. The explicit homomorphism

$$SU(2) \longrightarrow SO(3) \tag{2.22}$$

can be obtained as follows. $SU(2)$ matrices are in bijective correspondence with traceless self-adjoint matrices via the exponential relation

$$U = e^{iA}. \tag{2.23}$$

Since Pauli matrices are a basis for traceless self-adjoint matrices, we can write

$$U(\boldsymbol{\theta}) = e^{i\frac{\boldsymbol{\theta}}{2}\cdot\boldsymbol{\sigma}}, \quad \boldsymbol{\theta} \in \mathbb{R}^3. \tag{2.24}$$

The $SO(3)$ transformation associated to $U(\boldsymbol{\theta})$ is the rotation $R(\boldsymbol{\theta}) \in SO(3)$ around the axis directed along $\boldsymbol{\theta}$, with angle given by $|\boldsymbol{\theta}|$

$$U(\boldsymbol{\theta}) \longmapsto R(\boldsymbol{\theta}). \tag{2.25}$$

This is a surjective homomorphism. From equation (2.24) we have that the kernel of the homomorphism is given by $\mathbb{Z}_2 \cong \{\pm I\}$. Via homomorphism theorem, one obtains equation 2.20.

2.4 Wigner's theorem

We present here an important theorem, without proof, obtained by Wigner [Wi31]. We premise that a transformation $\mathcal{T} : \mathsf{P}\mathcal{H} \to \mathsf{P}\mathcal{H}$ which preserves the projective product

$$\mathcal{T}\mathcal{R}_\psi \cdot \mathcal{T}\mathcal{R}_\phi = \mathcal{R}_\psi \cdot \mathcal{R}_\phi \tag{2.26}$$

is called a *symmetry* because it does not change observed expectation values (again, in force of Born's rule). Since this is equivalent to the preservation of the Fubini-Study metric, it turns out that a quantum mechanical symmetry is precisely an isometry of $\mathsf{P}\mathcal{H}$ as a Riemannian manifold.

Theorem 2.4.1 (Wigner). *Any surjective symmetry $\mathcal{T} : P\mathcal{H} \to P\mathcal{H}$ admits an operator U which is unitary or antiunitary and makes the following diagram commutative:*

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{U} & \mathcal{H} \\ \downarrow & & \downarrow \\ P\mathcal{H} & \xrightarrow{\mathcal{T}} & P\mathcal{H} \end{array}$$

The lifting operator U is unique up to a phase factor.

Chapter 3

Distance measures between states

The notion of "closeness" between a pair of states is important in many areas of quantum information, but there is no unique formal definition of this intuitive concept. In fact, there are several ways to quantify the distance between two density operator, and each one has proved useful in some applications.

The most common measures employed in quantum information are the *trace distance* and the *fidelity* (which is related to the Fubini-Study distance). Since they will be used in following chapters, we present them along with their main properties.

Hereafter, complex vectors $\psi \in \mathcal{H}$ are always assumed to be normalized.

3.1 Trace distance

The trace distance is the the L_p-distance between operators, computed for $p = 1$, with an addictional factor $\frac{1}{2}$:

$$d_{tr}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\operatorname{Tr}|\rho - \sigma|. \tag{3.1}$$

It enjoys many interesting properties (see for example [NC00]):

1. $0 \leq d_{tr}(\rho, \sigma)$ with equality if and only if $\rho = \sigma$.

2. $d_{tr}(\rho, \sigma) \leq 1$ with equality if and only if ρ and σ have orthogonal supports.

3. Symmetry: $d_{tr}(\rho, \sigma) = d_{tr}(\sigma, \rho)$.

4. Triangle inequality: $d_{tr}(\rho, \sigma) \leq d_{tr}(\rho, \tau) + d_{tr}(\tau, \sigma)$.

5. Unitary invariance: $d_{tr}(\rho, \sigma) = d_{tr}(U\rho U^\dagger, U\sigma U^\dagger)$.

6. Monotonicity: $d_{tr}(\rho, \sigma) \geq d_{tr}(\mathcal{E}(\rho), \mathcal{E}(\sigma))$, where $\mathcal{E}$ is a trace preserving CP map.

The operational meaning of the trace distance is related to the problem of distinguishing two density operators with a single experiment. This is precisely the content of the following theorem. We remind that $d_1(x_i, y_i)$ is the trace distance of the vectors x_i and y_i:

$$d_1(x_i, y_i) := \frac{1}{2} \sum_i |x_i - y_i|. \tag{3.2}$$

Theorem 3.1.1. *Let $p_i = \mathrm{Tr}(E_i\rho)$ and $q_i = \mathrm{Tr}(E_i\sigma)$ for some POVM $\{E_i\}$. Then*

$$d_{tr}(\rho, \sigma) = \max_{\{E_i\}} d_1(p_i, q_i) \tag{3.3}$$

where we maximize over all POVMs.

Proof. We give the proof only for projective measurements. Indicating with N_+ and N_- the positive and the negative part of the operator $\rho - \sigma$, we observe that

$$d_{tr}(\rho, \sigma) = \frac{1}{2} \mathrm{Tr}(N_+ + N_-) = \mathrm{Tr}\, N_+ = \mathrm{Tr}\, N_- \tag{3.4}$$

because the difference $N_+ - N_-$ is a traceless operator. Then, if P is any projector,

$$\operatorname{Tr} P(\rho - \sigma) = \operatorname{Tr} P(N_+ - N_-) \leq \operatorname{Tr} PN_+ \leq \operatorname{Tr} N_+ = d_{tr}(\rho, \sigma). \tag{3.5}$$

The equality holds if and only if P is the projector on the support of N_+. We have that

$$\operatorname{Tr} |E_i(\rho - \sigma)| = \operatorname{Tr} |E_i(N_+ - N_-)| \leq \operatorname{Tr} E_i(N_+ + N_-) = \operatorname{Tr} E_i|\rho - \sigma|, \tag{3.6}$$

and therefore

$$d_1(p_i, q_i) = \frac{1}{2} \sum_i \operatorname{Tr} |E_i(\rho - \sigma)| \leq \frac{1}{2} \sum_i \operatorname{Tr} E_i|\rho - \sigma| = d_{tr}(\rho, \sigma) \tag{3.7}$$

and the equality holds if we choose as POVM the projective measurements $\{P_\pm\}$ consisting of projectors on the supports of N_+ and N_- respectively. □

We also give the proof of monotonicity.

Proof. From the first part of the proof of theorem (3.1.1) we can find a projector P such that

$$d_{tr}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = \operatorname{Tr} P(\mathcal{E}(\rho) - \mathcal{E}(\sigma)). \tag{3.8}$$

Since $\mathcal{E}$ is trace preserving we have that $\operatorname{Tr} \mathcal{E}(N_+) = \operatorname{Tr} \mathcal{E}(N_-)$ (where $N_\pm$ is the usual decomposition of $\rho - \sigma$ in positive and negative part). Then

$$\begin{aligned} d_{tr}(\rho, \sigma) = \frac{1}{2} \operatorname{Tr}(N_+ + N_-) = \frac{1}{2} \operatorname{Tr}(\mathcal{E}(N_+) + \mathcal{E}(N_-)) = \\ = \operatorname{Tr} \mathcal{E}(N_+) \geq \operatorname{Tr} P\mathcal{E}(N_+) \geq \operatorname{Tr} P(\mathcal{E}(N_+) - \mathcal{E}(N_-)) = \\ = \operatorname{Tr} P(\mathcal{E}(\rho) - \mathcal{E}(\sigma) = d_{tr}(\mathcal{E}(\rho), \mathcal{E}(\sigma)). \end{aligned} \tag{3.9}$$

□

3.2 Fidelity

The classical fidelity of two probability distributions p_i, and q_i is

$$F(p_i, q_i) = \sum_i \sqrt{p_i} \cdot \sqrt{q_i}. \tag{3.10}$$

The quantum generalization is defined as

$$F(\rho, \sigma) := \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \mathrm{Tr}\,|\sqrt{\rho}\sqrt{\sigma}|. \tag{3.11}$$

The fidelity was originally introduced by Uhlmann [Uh76] but its use became widespread in Quantum Information only after Jozsa reintroduced it in [Jo94]. In Jozsa's definition the quantum fidelity is the square of (3.11): $F_J = F^2$. For this reason the quantity (3.11) is sometimes referred to as *root fidelity* and is indicated with $\sqrt{F}$. Anyway, we stick to Uhlmann's definition.

We note that when one of the states is pure, say $\rho = |\psi\rangle\langle\psi|$, the fidelity becomes

$$F(|\psi\rangle, \sigma) = |\langle\psi|\sigma|\psi\rangle|^{\frac{1}{2}} \tag{3.12}$$

and, when also the other state is pure $\sigma = |\phi\rangle\langle\phi|$

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle| \tag{3.13}$$

Since this expression coincides with the product (2.4) introduced on $\mathsf{P}\mathcal{H}$, the Fubini-Study distance could also be defined starting from the fidelity

$$D_{FS}(|\psi\rangle, |\phi\rangle) = \arccos F(|\psi\rangle, |\phi\rangle). \tag{3.14}$$

Among the properties of the fidelity we have ([NC00]):

1. $0 \leq F(\rho, \sigma)$ with equality if and only if ρ and σ have orthogonal supports.

2. $F(\rho, \sigma) \leq 1$ with equality if and only if $\rho = \sigma$.

3. Symmetry: $F(\rho,\sigma) = F(\sigma,\rho)$.

4. Unitary invariance: $F(\rho,\sigma) = F(U\rho U^\dagger, U\sigma U^\dagger)$.

5. Monotonicity: $F(\rho,\sigma) \geq F(\mathcal{E}(\rho),\mathcal{E}(\sigma))$, where $\mathcal{E}$ is a trace preserving CP map.

The symmetry is not obvious from the definition, but it is an easy consequence of Uhlmann's theorem [Uh76], which gives a very useful characterization of fidelity in terms of purification of the density operators ρ and σ. We present the theorem and the proof, preceded by a simple lemma.

Lemma 3.2.1. *Let A be an operator and U a unitary operator. Then*

$$|\operatorname{Tr}(AU)| \leq \operatorname{Tr}|A|. \tag{3.15}$$

The equality is obtained choosing $U = V^\dagger$, where V is the unitary part of the polar decomposition $A = |A|V$.

Proof. We have

$$\begin{aligned} |\operatorname{Tr}(AU)| = |\operatorname{Tr}(|A|VU)| = |\operatorname{Tr}(|A|^{\frac{1}{2}}|A|^{\frac{1}{2}}VU)| \leq \\ \leq \sqrt{\operatorname{Tr}|A|}\sqrt{\operatorname{Tr}(U^\dagger V^\dagger |A| VU)} = \operatorname{Tr}|A| \end{aligned} \tag{3.16}$$

where we have applied the Cauchy-Schwartz inequality to the Hilbert-Schimdt scalar product. □

Theorem 3.2.2 (Uhlmann). *Let ρ, σ be density operators for some system Q. Then*

$$F(\rho,\sigma) = \max_{|\psi_\sigma\rangle} |\langle\psi_\rho|\psi_\sigma\rangle| \tag{3.17}$$

where $|\psi_\rho\rangle$ and $|\psi_\sigma\rangle$ are purification of ρ and σ respectively, on a common ancillary system R, identical to Q.

Proof. We indicate with $|\omega\rangle$ the (non normalized) maximally entangled bipartite state for the compound system $R+Q$:

$$|\omega\rangle = \sum_i |i\rangle_R \otimes |i\rangle_Q, \tag{3.18}$$

where $\{|i\rangle\}$ is a basis for R and Q. Then, a purification of ρ can be written as

$$|\psi_\rho\rangle = (\sqrt{\rho} \otimes I)|\omega\rangle. \tag{3.19}$$

Analogously, for σ one has

$$|\psi_\sigma\rangle = (\sqrt{\sigma} V_R \otimes V_Q)|\omega\rangle \tag{3.20}$$

for some other unitary operators V_R and V_Q. Taking the absolute value of the scalar product we have

$$\begin{aligned} |\langle\psi_\rho|\psi_\sigma\rangle| &= |\langle\omega|\sqrt{\rho}\sqrt{\sigma} V_R \otimes V_Q|\omega\rangle| = \\ &= |\operatorname{Tr}(\sqrt{\rho}\sqrt{\sigma} V_R V_Q^T)| \leq \operatorname{Tr}|\sqrt{\rho}\sqrt{\sigma}|, \end{aligned} \tag{3.21}$$

using the identity $\operatorname{Tr}(AB^T) = \langle\omega|A \otimes B|\omega\rangle$ and lemma (3.2.1). The equality is attained posing $V_Q = I$ and $V_R = V^\dagger$, where V is given by the polar decomposition $\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|V$. □

Monotonicity is also an immediate consequence of Uhlmann's theorem.

Proof. We can find purification $|\psi_\rho\rangle$ and $|\psi_\sigma\rangle$ such that $F(\rho,\sigma) = |\langle\psi_\rho|\psi_\sigma\rangle|$. Let U a unitary realization of the CP map $\mathcal{E}$, with ancillary system in some pure state $|0\rangle$. Then

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq |\langle\psi_\rho|\langle 0|U^\dagger U|\psi_\sigma\rangle|0\rangle| = |\langle\psi_\rho|\langle 0|\psi_\sigma\rangle|0\rangle| = F(\rho,\sigma) \tag{3.22}$$

where inequality follows from Uhlmann's theorem. □

For a two-dimensional system, we have the following explicit expression for the fidelity ([Hü92]):

$$F(\rho,\sigma) = \frac{\left(1 + \boldsymbol{r}_\rho \cdot \boldsymbol{r}_\sigma + \sqrt{(1-|\boldsymbol{r}_\rho|^2)(1-|\boldsymbol{r}_\sigma|^2)}\right)^{\frac{1}{2}}}{\sqrt{2}} \tag{3.23}$$

where $\boldsymbol{r}_\rho$ is the Bloch vector of the state ρ.

3.3 Isofidelity surfaces

We want to address the following question: given a mixed state ρ and a real number f, $0 \leq f \leq 1$, determine the set

$$\mathfrak{L}_{\rho,f} = \{\sigma, F(\rho,\sigma) = f\}.$$

In other words, for every ρ we want to identify the level surfaces of the real-valued function [1]

$$\begin{aligned} F_\rho(\sigma): \quad & \mathcal{S} \to \mathbb{R} \\ & \sigma \mapsto F(\rho,\sigma). \end{aligned}$$

This function is defined on the set $\mathcal{S}$ of density operators of a n-dimensional system, which is a (n^2-1)-dimensional manifold [2]. Its level surfaces, thus, are (n^2-2)-dimensional manifolds.

The explicit representation of $\mathfrak{L}_{\rho,f}$ is in general hard to find, because an explicit formula for the fidelity would be needed, which is known only in the qubit case. However, some insights about the shape of $\mathfrak{L}_{\rho,f}$ can be obtained simply by symmetry arguments.

[1] To avoid clumsiness of notation, in this section we pose $\mathcal{S} = \mathcal{S}(\mathcal{H})$.

[2] The space of self-adjoint operators on a n-dimensional complex vector space has dimension n^2. The condition $\mathrm{Tr}(\rho) = 1$ required for density operators diminishes the dimension by one.

Let G_ρ be the group of all the unitary transformations U which have ρ as a fixed point

$$U\rho U^\dagger = \rho \tag{3.24}$$

The action of G_ρ on $\mathcal{S}$ is given by

$$\begin{aligned} G_\rho \times \mathcal{S} &\to \mathcal{S} \\ (U, \sigma) &\mapsto U\sigma U^\dagger. \end{aligned} \tag{3.25}$$

Two unitary operators differing by a phase factor induce the same transformation on $\mathcal{S}$. The phase factors are exactly the scalar elements of G_ρ (i.e. the multiples of the identity), forming the central subgroup $U(1) \subset G_\rho$. In order to eliminate this ambiguity we can "quotient away" the center of G_ρ and obtain the effective action of $G_\rho/U(1)$ given by $([U], \sigma) \mapsto U\sigma U^\dagger$, which makes the following diagram commutative

$$\begin{array}{ccc} G_\rho \times \mathcal{S} & \longrightarrow & \mathcal{S} \\ \downarrow & \nearrow & \\ \frac{G_\rho}{U(1)} \times \mathcal{S} & & \end{array}$$

The action of G_ρ induces a quotient structure $\mathcal{S}/G_\rho$ (the set of the orbits of G_ρ), along with the natural projection $\pi : \mathcal{S} \to \mathcal{S}/G_\rho$, which takes each density operator to its orbit. We have the following

Proposition 3.3.1. *Each G_ρ-orbit is entirely contained in one of the level surfaces of F_ρ.*

Proof. Since the fidelity is invariant under any unitary transformation of both states,

$$F(\rho, \sigma) = F(U\rho U^\dagger, U\sigma U^\dagger), \tag{3.26}$$

we have that, for every $U \in G_\rho$,

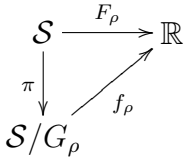
$$F_\rho(U\sigma U^\dagger) = F(U\sigma U^\dagger, \rho) = F(U\sigma U^\dagger, U\rho U^\dagger) = F(\sigma, \rho) = F_\rho(\sigma). \tag{3.27}$$

Therefore, under the action of the group G_ρ the points of $\mathcal{S}$ are transformed into points belonging to the same level surface. $\square$

This means that each level surface $\mathfrak{L}_{\rho,f}$ is obtained "gluing" together some of the orbits of G_ρ. Moreover, the proposition implies that the function $f_\rho : \mathcal{S}/G_\rho \to \mathbb{R}$ given by

$$f_\rho(\pi(\sigma)) := F_\rho(\sigma), \tag{3.28}$$

is well-defined, making the following diagram commutative



Thanks to the factorization $F_\rho = f_\rho \circ \pi$ the problem is reduced to the determination of the level sets $\mathfrak{l}_{\rho,f}$ of the function f_ρ, which is defined on the smaller space $\mathcal{S}/G_\rho$. After obtaining $\mathfrak{l}_{\rho,f}$, we can recover $\mathfrak{L}_{\rho,f}$ by taking the inverse image of $\mathfrak{l}_{\rho,f}$ under the projection map π

$$\mathfrak{L}_{\rho,f} = \pi^{-1}(\mathfrak{l}_{\rho,f}). \tag{3.29}$$

The defining condition $U\rho U^\dagger = \rho$ for the elements of G_ρ can be expressed as the commutation relation $[\rho, U] = 0$. The following proposition gives an explicit characterization of these elements in terms of the eigenspaces of ρ, thus providing the structure of G_ρ.

Proposition 3.3.2. *Let ρ be a self-adjoint operator whose (distinct) eigenvalues are $\lambda_1, \ldots, \lambda_n$, and the corrisponding eigenspaces are $V_{\lambda_1}, \ldots, V_{\lambda_n}$. Let U be a unitary operator. Then the following conditions are equivalent:*

(1) $[\rho, U] = 0$

(2) U decomposes as

$$U = U_1 \oplus \ldots \oplus U_n, \tag{3.30}$$

where U_i is a unitary operator acting on V_{λ_i}.

Proof. (1)⇒(2). If ψ is any λ_1-eigenvector then

$$\rho U\psi = U\rho\psi = \lambda_1 U\psi. \tag{3.31}$$

Therefore, also $U\psi$ is a λ_1-eigenvector, implying that V_{λ_1} is U-invariant

$$U(V_{\lambda_1}) \subseteq V_{\lambda_1}. \tag{3.32}$$

The restriction $U_{|V_{\lambda_1}}$ is a unitary operator acting on V_{λ_1} and thus there must be some U_1 such that

$$U_{|V_{\lambda_1}} = U_1. \tag{3.33}$$

Indicating with P the projector on V_{λ_1} and with Q the projector on the direct sum of the remaining eigenspaces, U can be written as

$$U = PUP \oplus QUQ = U_{|V_{\lambda_1}} \oplus QUQ = U_1 \oplus QUQ. \tag{3.34}$$

The same argument can be applied to the remaining eigenvalues to obtain the required decomposition.

(2)⇒(1). Follows readly from the representation

$$\rho = \lambda_1 I_1 \oplus \ldots \oplus \lambda_n I_n, \tag{3.35}$$

where I_i is the identity on V_{λ_i}. □

When G_ρ is known, the shape of the orbits can be easily obtained analysing the effective action of $G_\rho/U(1)$ on the space of states. In the subsequent sections we give some details for the qubit.

3.3.1 The qubit case

For a two-level system, the problem can be completely solved since we have the explicit formula for the fidelity (3.23). The sets $\mathfrak{L}_{\rho,f}$ are simply two-dimensional surfaces.

If ρ has equal eigenvalues (i.e. it is the maximally chaotic state $\rho = I/2$), G_ρ consists of all the unitary operators $G_\rho \cong U(2)$, and thus

$$\frac{G_\rho}{U(1)} \cong \frac{U(2)}{U(1)} \cong SO(3). \tag{3.36}$$

Each orbit is a sphere centered in the origin and, being a two-dimensional manifold, it exhausts one of the level surfaces $\mathfrak{L}_{\rho,f}$.

If the eigenvalues are different (i.e. the state is anything but the maximally chaotic) we have $G_\rho \cong U(1) \oplus U(1)$. Hence

$$\frac{G_\rho}{U(1)} \cong \frac{U(1) \oplus U(1)}{U(1)} \cong U(1) \cong SO(2). \tag{3.37}$$

and the orbits are circles. In this case we can exploit equation (3.23) and carry on the calculation of the sets $\mathfrak{l}_{\rho,f}$. The quotient space $\mathcal{S}/G_\rho$ can be identified with any half-disk whose diameter passes through $\boldsymbol{r}_\rho$. Each point $\boldsymbol{r}_\sigma$ of the half-disk is a representative of its own orbit $[\sigma]$. We can conveniently introduce real coordinates (x, y) chosen in such a way that the x-axis lies along the diameter and the y-axis lies along the ray orthogonal to the diameter. Thus, the coordinates of ρ are $\boldsymbol{r}_\rho = (\alpha, 0)$ for some α, $|\alpha| \leq 1$, whilst a generic point has coordinates $\boldsymbol{r}_\sigma = (x, y), y \geq 0, x^2 + y^2 \leq 1$. The fidelity formula (3.23) becomes

$$f_\rho([\sigma]) = \frac{\left(1 + \alpha x + \sqrt{(1-\alpha^2)(1-(x^2+y^2))}\right)^{\frac{1}{2}}}{\sqrt{2}}. \tag{3.38}$$

Imposing the condition $f_\rho = f$ and introducing the parameter $k = 2f^2 - 1$,

Figure 3.1: Some of the truncated ellipses, representing states which have the same fidelity with respect to ρ. Rotating the arcs around the x-axis we obtain the whole isofidelity surfaces for the Bloch ball.

after some rearrangements we obtain

$$\frac{(x-\alpha k)^2}{(1-\alpha^2)(1-k^2)} + \frac{y^2}{(1-k^2)} = 1, \quad \alpha x \leq k. \tag{3.39}$$

This equation represents the family $\mathcal{C}_{\rho,f}$ parameterized by k. It consists of arcs of ellipses (the truncation is given by the condition $\alpha x \leq k$). The isofidelity surfaces are partial ellipsoids obtained rotating the arcs around the diameter. When $\alpha = \pm 1$ ρ is pure and the ellipsoids degenerate into planes orthogonal to the diameter.

Chapter 4

Quantum State Discrimination

Having gathered the formal tools needed to describe in full generality quantum measurements, we can now face the specific issue of *quantum state discrimination*.

Strictly speaking, the state of a quantum system cannot be measured because it is not even a real *unknown* quantity. It simply represents a summary of everything we know about the system.

A *state discrimination* actually refers to the following two-party scenario: Alice prepares a system in some state, drawing arbitrarly, but with definite probabilities, from a given ensemble. Subsequently she sends the system to Bob, whose task is to perform a measurement on it in order to gain information about the preparation chosen by Alice.

There are two main strategies we can employ to accomplish this job. The first is the *quantum hypothesis testing* (QHT). Here, we are given an unknown state chosen from some finite ensemble, and we have to guess with the minimum error probability which of these states it is. In this approach, we are always required to choose, and we cannot say 'don't know'. But when such inconclusive answers are allowed, one can show that sometimes

another strategy is possible, which determines the correct state every time it does not end up with an inconclusive result, and therefore it is called *unambiguous state discrimination* (USD). In the following sections we will discuss these two strategies.

4.1 Quantum hypothesis testing

Suppose we know that the ensemble from which Alice picks up the state is represented by the set of N density operators $\{\rho_i\}_{i=1,\dots,N}$, whose *a priori* probability distribution is $\{p_i\}_{i=1,\dots,N}$. Since we are required to make a choice about what state was prepared by Alice, we perform an experiment described by a POVM $\{\Pi_i\}_{i=1,\dots,N}$, and interpret the outcome j as the detection of the state ρ_j. We want to minimize the average probability p_E of incorrect identification, given by

$$p_E = 1 - \sum_i p_i \operatorname{Tr}(\rho_i \Pi_i). \tag{4.1}$$

The necessary and sufficient conditions that the POVM must satisfy in order to achieve minimum p_E are [Ho73]

$$\Pi_j(p_j\rho_j - p_i\rho_i)\Pi_i = 0 \tag{4.2}$$

$$\sum_j p_j\Pi_j\rho_j - p_i\rho_i \geq 0 \tag{4.3}$$

but general solutions are hard to find. It is possible to show ([Ken73], [El03]) that for mixed linearly indipendent states, the optimal POVM is simply a PVM. In the following we will focus on the particular case of two mixed states, and then, further specializing to pure states, we will obtain the value of p_E originally found by Helstrom in [Hel76].

Let us indicate the two mixed states with $\rho_\pm$, and let $p_\pm$ be their respective probabilities. For a generic POVM $\{\Pi_\pm\}$ the error probability is

$$p_E = p_+ \operatorname{Tr}(\rho_+\Pi_-) + p_- \operatorname{Tr}(\rho_-\Pi_+). \tag{4.4}$$

Remembering that $\Pi_+ + \Pi_- = I$ we can rewrite it as

$$p_E = p_+ - \operatorname{Tr}[\Pi_+(p_+\rho_+ - p_-\rho_-)] = p_- + \operatorname{Tr}[\Pi_-(p_+\rho_+ - p_-\rho_-)]. \tag{4.5}$$

Taking the sum of these expressions and dividing by two we finally have

$$p_E = \frac{1}{2}\left\{1 - \operatorname{Tr}[(p_+\rho_+ - p_-\rho_-)(\Pi_+ - \Pi_-)]\right\}. \tag{4.6}$$

This expression attains the minimum when $\Pi_\pm$ are projectors on the support of, respectively, the positive and the negative part of the operator $p_+\rho_+ - p_-\rho_-$. Hence we have

$$p_E = \frac{1}{2}(1 - \|p_+\rho_+ - p_-\rho_-\|_1) \tag{4.7}$$

where $\|\cdot\|_1$ is the trace-norm of its argument.

When $\rho_\pm$ are pure they can be written as $\rho_\pm = |\psi_\pm\rangle\langle\psi_\pm|$ where $|\psi_\pm\rangle$ are suitable normalized vectors in $\mathcal{H}$. Choosing an orthonormal basis $|\pm\rangle$ for the space spanned by $|\psi_\pm\rangle$ we can find some angle θ such that

$$|\psi_\pm\rangle = \cos\theta|+\rangle \pm \sin\theta|-\rangle. \tag{4.8}$$

From the matrix representation of the operator $p_+\rho_+ - p_-\rho_-$ in the basis $|\pm\rangle$

$$\begin{pmatrix} \Delta\cos^2\theta & \frac{\sin 2\theta}{2} \\ \frac{\sin 2\theta}{2} & \Delta\sin^2\theta \end{pmatrix}, \tag{4.9}$$

where $\Delta = p_+ - p_-$, one can easily obtain the eigenvalues

$$\lambda_\pm = (\Delta \pm \sqrt{1 + (\Delta^2 - 1)\cos^2 2\theta})/2 \tag{4.10}$$

and verify that the (orthonormal) eigenvectors are

$$|\omega_\pm\rangle = \frac{1}{\sqrt{2}}\left(\sqrt{1\pm\xi}|+\rangle \pm \sqrt{1\mp\xi}|-\rangle\right), \tag{4.11}$$

where $\xi = \Delta\cos 2\theta/\sqrt{1+(\Delta^2-1)\cos^2 2\theta}$. Since $|\omega_\pm\rangle$ are relative to the positive and the negative eigenvalue, respectively, it follows that the projectors we are looking for are precisely

$$\Pi_\pm = |\omega_\pm\rangle\langle\omega_\pm|. \tag{4.12}$$

Inserting this PVM for example in equation (4.6) we obtain, after some algebra, the optimal error probability as found by Helstrom:

$$\begin{aligned} p_E = \frac{1}{2}\left(1-\sqrt{1-4p_+p_-|\langle\psi_+|\psi_-\rangle|^2}\right) = \\ \frac{1}{2}\left(1-\sqrt{1-4p_+p_-\cos^2 2\theta}\right). \end{aligned} \tag{4.13}$$

4.2 Unambiguous state discrimination

If we admit the possibility of getting *inconclusive* outcomes, i.e. outcomes that do not corrispond to any possible states, one can show that in some cases it is possible to design a measurement which correctly identifies the state every time it does not give such an inconclusive result.

Here we address the simple case of two pure states, $|\psi_\pm\rangle$, as defined in equation (4.8). Introducing two further states

$$|\psi_\pm^\perp\rangle = \sin\theta|+\rangle \pm \cos\theta|-\rangle, \tag{4.14}$$

which have the property $\langle\psi_+^\perp|\psi_-\rangle = 0$ and $\langle\psi_-^\perp|\psi_+\rangle = 0$, we can define the following POVM:

$$\begin{aligned} \Pi_\pm &= c_\pm|\psi_\pm^\perp\rangle\langle\psi_\pm^\perp| \\ \Pi_{inc} &= I - \Pi_+ - \Pi_- \end{aligned} \tag{4.15}$$

where the coefficients $c_\pm$ are constrained by the positivity conditions of the operators Π_+, Π_-, Π_{inc}.

Since it is obvious that

$$\langle\psi_-|\Pi_+|\psi_-\rangle = 0 \tag{4.16}$$

$$\langle\psi_+|\Pi_-|\psi_+\rangle = 0 \tag{4.17}$$

we will never get the outcome '+' when the state is $|\psi_-\rangle$ and vice versa. Thus, obtaining one of these two results permits us to tell what the initial state was. However, we pay this error-free identification with the fact the sometimes the result of the measurement will be the inconclusive outcome, whose total probability is

$$\begin{aligned} p_{inc} &= 1 - c_+ p_+ |\langle\psi_+^\perp|\psi_+\rangle|^2 - c_- p_- |\langle\psi_-^\perp|\psi_-\rangle|^2 \\ &= 1 - (c_+ p_+ + c_- p_-)\sin^2 2\theta \end{aligned} \tag{4.18}$$

where $p_\pm$ are, as usual, the *a priori* probabilities of $|\psi_\pm\rangle$.

One can ask how the coefficients $c_\pm$ should be tuned in order to minimize this probability, recalling that the operator Π_{inc} must remain positive. This problem has been addressed and solved for equal *a priori* probabilities $p_\pm = 1/2$ by Ivanovic [Iv87], Dieks [Di88] and Peres [Pe88], whose work established that the optimum value of p_{inc} is

$$p_{inc} = |\langle\psi_+|\psi_-\rangle| = \cos 2\theta. \tag{4.19}$$

The general solution for unbalanced probabilities is also known, though its expression is more involved (Jaeger and Shimony [JS95]). All these derivations try to reach the limit by building an explicit ancillary extension and subsequently performing a suited projective measurement. In the next section we provide an alternative proof, directly based on the POVM formalism, which does not require extra ancillas.

When we get an inconclusive outcome we do not gain any information about the initial state, but this does not mean that it is not affected by the measurement. In fact, quite the opposite is true: in general, as the probability p_{inc} decreases, the states after the measurement become more and more closer (in term of their overlap), and eventually become equal when p_{inc} reaches the Ivanovic-Dieks-Peres bound. Performing an unambiguous state discrimination is therefore a sort of gamble: when it succeeds, we achieve perfect discrimination; but if it fails, the states become less distinguishable than before. Since the success probability $p_s = 1 - p_{inc}$ can be thought of as an indicator of the average information gain due to the measurement, this is an example of the information/disturbance tradeoff.

4.2.1 The Jaeger-Shimony bound

We want to minimize the function

$$p_{inc}(c_+, c_-) = 1 - (c_+ p_+ + c_- p_-)(1 - \omega^2) \tag{4.20}$$

where $\omega = \cos 2\theta$, subject to the positivity constraints $\Pi_+ \geq 0$, $\Pi_- \geq 0$, $\Pi_{inc} \geq 0$. The first two conditions immediately give

$$\begin{array}{l} c_+ \geq 0 \\ c_- \geq 0 \end{array} . \tag{4.21}$$

The constraint on Π_{inc} can be expressed imposing the positivity of its minimum eigenvalue λ_-

$$\lambda_-(c_+, c_-) \geq 0. \tag{4.22}$$

In the basis $|\pm\rangle$ the operator Π_{inc} has the following matrix representation

$$\Pi_{inc} = \begin{pmatrix} 1 - (c_+ + c_-)\sin^2\theta & (c_- - c_+)\sin\theta\cos\theta \\ (c_- - c_+)\sin\theta\cos\theta & 1 - (c_+ + c_-)\cos^2\theta \end{pmatrix} \tag{4.23}$$

Figure 4.1: The values of c_+ and c_- inside the area delimited by black lines define a valid POVM. The limiting cases $\omega = 1$ (parallel states) and $\omega = 0$ (orthogonal states) are shown with dashed lines.

The eigenvalues are

$$\lambda_\pm(c_+, c_-) = \frac{1}{2}\left(2 - c_+ - c_- \pm \sqrt{c_+^2 + c_-^2 + (4\omega^2 - 2)c_+c_-}\right). \qquad (4.24)$$

The area defined by inequalities (4.21) and (4.22) is depicted in figure 4.1, delimited by black lines. The dashed lines represents the limiting cases $\omega = 1$ and $\omega = 0$.

The function that we are minimizing is linear, and the extremal points of a linear function can lie only on the frontier of the region in which it is defined. Since $p_{inc}(c_+, c_-)$ clearly diminishes along the segments $\overrightarrow{OA}$ and

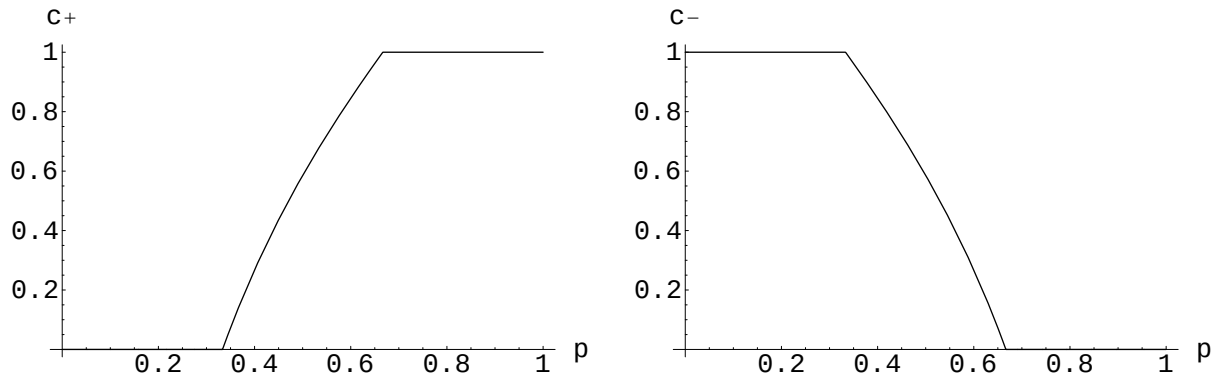


Figure 4.2: c_+ and c_- plotted for $2\theta = \pi/4$. Here $p = p_+$.

$\overrightarrow{OB}$, the minimum must lie on the curve defined by the equality in equation (4.22)

$$\lambda_-(c_+, c_-) = 0. \tag{4.25}$$

The constrained minimization problem can now be addressed employing the Lagrange multipliers method. The new equation to be solved is

$$\nabla p_{inc}(c_+, c_-) = k \nabla \lambda_-(c_+, c_-) \tag{4.26}$$

where k is the Lagrange multiplier. After some algebra we obtain the system

$$\begin{cases} (1-R)(2-c_+-c_-) = R[2c_+ + (4\omega^2-2)c_-] - [2c_- + (4\omega^2-2)c_+] \\ 1 - c_+ - c_- + (1-\omega^2)c_+c_- = 0 \end{cases}, \tag{4.27}$$

where we have posed $R = p_+/p_-$. The solutions $\tilde{c}_\pm$, taking into account conditions (4.21), are

$$\begin{cases} \tilde{c}_+ = 0, \quad \tilde{c}_- = 1 & R \leq \omega^2 \\ \tilde{c}_+ = \frac{1-\omega/\sqrt{R}}{1-\omega^2}, \quad \tilde{c}_- = \frac{1-\omega\sqrt{R}}{1-\omega^2} & \omega^2 \leq R \leq \frac{1}{\omega^2} \\ \tilde{c}_+ = 1, \quad \tilde{c}_- = 0 & R \geq \frac{1}{\omega^2} \end{cases} . \tag{4.28}$$

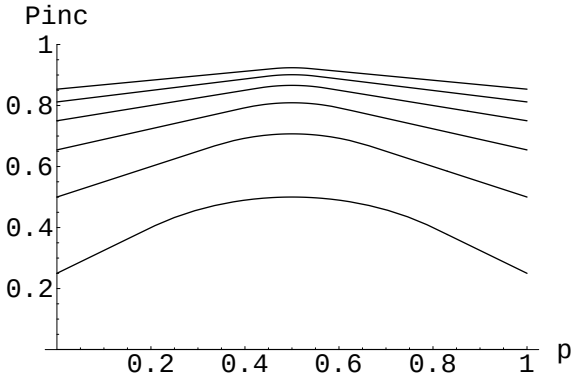


Figure 4.3: p_{inc} plotted for $2\theta = \pi/j$ with $j = 2 \ldots 8$. Here $p = p_+$. We note that for orthogonal states, $2\theta = \pi/2$, p_{inc} vanishes, while it approaches 1 as θ diminishes.

The corresponding optimal value for p_{inc} is

$$\begin{cases} \omega^2 + (1-\omega^2)p_+ & p_+ \leq \frac{\omega^2}{1+\omega^2} \\ 2\omega\sqrt{p_+ p_-} & \frac{\omega^2}{1+\omega^2} \leq p_+ \leq \frac{1}{1+\omega^2} \\ 1 - (1-\omega^2)p_+ & p_+ \geq \frac{1}{1+\omega^2} \end{cases} . \tag{4.29}$$

The optimal POVM corresponding to $\tilde{c}_\pm$ will be indicated with $\{\tilde{\Pi}_\pm, \tilde{\Pi}_{inc}\}$. We note that when odds ratio of *a priori* probability is greater than the treshold value $1/\omega^2$ or lesser than ω^2 the POVM becomes a Projective Measurement on the most probable state between $|\psi_\pm\rangle$ and its orthogonal subspace.

4.3 Information/disturbance tradeoff

In order to express the tradeoff quantitatively, we have to choose suitable variables representing information and disturbance in this specific setting. We have already seen that the total probability of obtaining a conclusive answer $p_s = 1 - p_{inc}$ is a good choice for the estimation of information. To quantify the disturbance, we imagine to be in the following scenario: Alice

prepares one of the two states $|\psi_\pm\rangle$ with *a priori* probabilities $p_\pm$, then she sends the system to Bob, who performs the optimal USD derived in the previous section.

$$\underset{|\psi_?\rangle}{\text{Alice}} \longrightarrow \underset{\{\tilde{\Pi}_\pm, \tilde{\Pi}_{inc}\}}{\text{Bob}}$$

If nothing happens during the transmission of the system we already know that Bob will identify the state with probability $p_s^{Bob} = 1 - p_{inc}^{Bob}$. Now, suppose that an eavesdropper, say Eve, intercepts the system and, after performing some kind of measurement, she resends it to Bob. Being unaware of the intrusion, he tries the same USD as before, but clearly this time the result he obtains could be *wrong*, because the eavesdropping attempt tampered with the state. Precisely, he could obtain as conclusive answer '+' even if the state was $|\psi_-\rangle$, and vice versa.

Suppose that Eve's measurement is a suboptimal USD, i.e. a measurement whose POVM Π is given by equation (4.15) where $c_\pm$ are constrained only by positivity.

$$\underset{|\psi_?\rangle}{\text{Alice}} \longrightarrow \underset{\{\Pi_\pm, \Pi_{inc}\}}{\text{Eve}} \longrightarrow \underset{\{\tilde{\Pi}_\pm, \tilde{\Pi}_{inc}\}}{\text{Bob}}$$

The information extracted by Eve can be quantified by her probability p_s^{Eve} of correct identification

$$I = p_s^{Eve} = 1 - p_{inc}^{Eve} = (c_+ p_+ + c_- p_-) \sin^2(2\theta) \tag{4.30}$$

which of course depends only on the POVM: $I = I(\Pi)$. The disturbance, on the other hand, will depend crucially on how the states are transformed after Eve's measurement, that is, on the instrument $\{\mathcal{E}_\pm, \mathcal{E}_{inc}\}$ realizing the

POVM: $D = D(\mathcal{E})$. As precise definition of disturbance, we can take the total probability p_e^{Bob} that Bob would wrongly identify Alice's preparation

$$D = p_e^{Bob} = p_+ \operatorname{Tr}(\tilde{\Pi}_- \mathcal{E}_{inc}(|\psi_+\rangle\langle\psi_+|)) + p_- \operatorname{Tr}(\tilde{\Pi}_+ \mathcal{E}_{inc}(|\psi_-\rangle\langle\psi_-|)). \quad (4.31)$$

Each instrument yields a point in the I/D plane, and the information-disturbance tradeoff is the optimal frontier which delimits the area corresponding to all the possible instruments.

Eve obtaines the maximum information I_{max} when she performs the optimal USD, and in force of (4.29) this maximum value is fixed by the separation between the states $|\langle\psi_+|\psi_-\rangle| = \cos 2\theta$ and the *a priori* probabilities $p_\pm$.

$$I_{max} = 1 - 2\cos(2\theta)\sqrt{p_+ p_-} \quad (4.32)$$

In order to derive the optimal frontier, for every fixed value $I \leq I_{max}$ we minimize the disturbance D over all the instruments realizing a POVM of the form (4.15) which gives the information I. In this way we obtain a family of tradeoff curves $D = D(I)$ parametrized by θ and $p_\pm$.

The minimization is a two-step process: for every POVM Π, we have to minimize D over the set $\mathfrak{I}$ of instruments $\mathcal{E}$ realizing Π; then, we minimize over the set $\mathfrak{P}$ of POVMs giving infomation I.

$$D(I) = \min_{\mathfrak{P}} \min_{\mathfrak{I}} D(\mathcal{E}) \quad (4.33)$$

Step 1 We observe that for the conclusive outcomes $\Pi_\pm$ the corresponding optimal quantum operations are always the pure contractions $M_\pm$ leaving the system in the initial state:

$$\mathcal{E}_\pm(\rho) = M_\pm \rho M_\pm^\dagger \quad (4.34)$$

with

$$M_\pm = \sqrt{c_\pm}|\psi_\pm\rangle\langle\psi_\pm^\perp|. \quad (4.35)$$

Thus, we only need to worry about the quantum operation corresponding to the inconclusive outcome Π_{inc}. Here we assume that, again, a single Kraus operation is optimal. The assumption is reasonable, since the classical mixing added by a many-Kraus QO always makes the states less distiguishable. Thus we have

$$\mathcal{E}_{inc}(\rho) = M_{inc}\rho M_{inc}^{\dagger} \tag{4.36}$$

with

$$M_{inc} = U\Pi_{inc}^{1/2} \tag{4.37}$$

where U is unitary. Equation (4.31) becomes

$$D = p_+\langle\psi_+|\Pi_{inc}^{1/2}U^{\dagger}\tilde{\Pi}_-U\Pi_{inc}^{1/2}|\psi_+\rangle + p_-\langle\psi_-|\Pi_{inc}^{1/2}U^{\dagger}\tilde{\Pi}_+U\Pi_{inc}^{1/2}|\psi_-\rangle \tag{4.38}$$

and finally, using the identities $\tilde{\Pi}_{\pm} = \tilde{c}_{\pm}I - \tilde{c}_{\pm}|\psi_{\mp}\rangle\langle\psi_{\mp}|$,

$$\begin{aligned} D = \tilde{c}_-p_+\langle\psi_+|\Pi_{inc}|\psi_+\rangle - \tilde{c}_-p_+|\langle\psi_+|U\Pi_{inc}^{1/2}|\psi_+\rangle|^2 + \\ +\tilde{c}_+p_-\langle\psi_-|\Pi_{inc}|\psi_-\rangle - \tilde{c}_+p_-|\langle\psi_-|U\Pi_{inc}^{1/2}|\psi_-\rangle|^2 = \\ = \tilde{c}_-p_+\langle\psi_+|\Pi_{inc}|\psi_+\rangle\left(1 - \left|\langle\psi_+|U\frac{\Pi_{inc}^{1/2}|\psi_+\rangle}{\|\Pi_{inc}^{1/2}|\psi_+\rangle\|}\right|^2\right) + \\ +\tilde{c}_+p_-\langle\psi_-|\Pi_{inc}|\psi_-\rangle\left(1 - \left|\langle\psi_-|U\frac{\Pi_{inc}^{1/2}|\psi_-\rangle}{\|\Pi_{inc}^{1/2}|\psi_-\rangle\|}\right|^2\right). \end{aligned} \tag{4.39}$$

Our task is to minimize this expression varying the unitary U. We observe that, in order to maximize the ovarlaps appearing in equation (4.39) we can search only among unitaries of the form $U = R(\beta)V$, where $R(\beta)$ is a rotation of angle β on the plane individuated by $|\psi_{\pm}\rangle$, while V is the unitary which moves the normalized pair $\Pi_{inc}^{1/2}|\psi_{\pm}\rangle/\|\Pi_{inc}^{1/2}|\psi_{\pm}\rangle\|$ to the configuration $|\xi_{\pm}\rangle$ complanar with respect to $|\psi_{\pm}\rangle$ and with the same symmetry axis. Thus, ignoring the part of equation (4.39) which does not depend on U, we obtain

the function

$$f(\beta) = -\left(\tilde{c}_- p_+ \langle\psi_+|\Pi_{inc}|\psi_+\rangle |\langle\psi_+|R(\beta)|\xi_+\rangle|^2 + \right. \\ \left. + \tilde{c}_+ p_- \langle\psi_-|\Pi_{inc}|\psi_-\rangle |\langle\psi_-|R(\beta)|\xi_-\rangle|^2\right) \tag{4.40}$$

to be minimized over β. The separation angle $2\theta'$ between $|\xi_\pm\rangle$ is given by

$$\cos(2\theta') = |\langle\xi_+|\xi_-\rangle| = \frac{|\langle\psi_+|\psi_-\rangle|}{\sqrt{\langle\psi_+|\Pi_{inc}|\psi_+\rangle\langle\psi_-|\Pi_{inc}|\psi_-\rangle}}. \tag{4.41}$$

Introducing the parameters $A_\pm = \tilde{c}_\mp p_\pm \langle\psi_\pm|\Pi_{inc}|\psi_\pm\rangle$ and $\Delta\theta = \theta - \theta'$, after some algebra we have

$$f(\beta) = -\frac{A_+ + A_-}{2} - \frac{1}{2}\left[(A_+ + A_-)\cos(2\Delta\theta)\cos(2\beta) + \right. \\ \left. + (A_+ - A_-)\sin(2\Delta\theta)\sin(2\beta)\right]. \tag{4.42}$$

whose minimum is attained when

$$\tan 2\beta = \frac{A_+ - A_-}{A_+ + A_-}\tan 2\Delta\theta. \tag{4.43}$$

In this way we have obtained the minimum disturbance for a fixed POVM:

$$D(\Pi) = D(c_+, c_-) = \min_{\mathcal{E}\Rightarrow\Pi} D(\mathcal{E}). \tag{4.44}$$

Step 2 The dependence of the disturbance on the coefficients $c_\pm$ is much more complicated and the minimization can be carried on numerically. The values of $c_\pm$, with $2\theta = \pi/4$, are plotted in figure 4.4 for various *a priori* probabilities. Each curve interpolates between the totally uninformative measurement $c_\pm = 0$ and the optimal POVM $\tilde{\Pi}$.

For uniform probabilities $p_\pm = 1/2$ we can provide the analytical solution. In this case, by symmetry of the configuration, the coefficients in Eve's POVM are equal $c_\pm = c$. Thus $A_+ = A_-$ and $\beta = 0$, and the disturbance becomes

$$D = \frac{1}{2}\left[1 - \cos(2\theta)\right]\left(1 - c + \sqrt{c^2\sin^2(2\theta) - 2c + 1}\right). \tag{4.45}$$

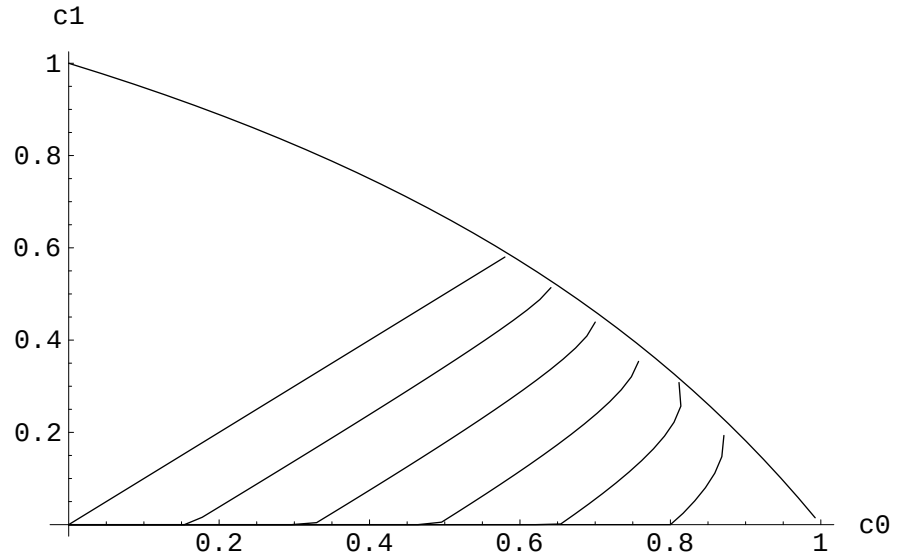


Figure 4.4: The coefficients $c_\pm$ of the minimum disturbing POVMs for $2\theta = \pi/4$. Each curve is obtained with fixed $R = p_+/p_-$ and the bisectrix corresponds to $R = 1$. When R reaches the threshold $1/\cos^2(2\theta) = 2$ the POVM becomes a projective measurement and the corresponding curve is represented by the segment connecting the origin with the point $(1, 0)$.

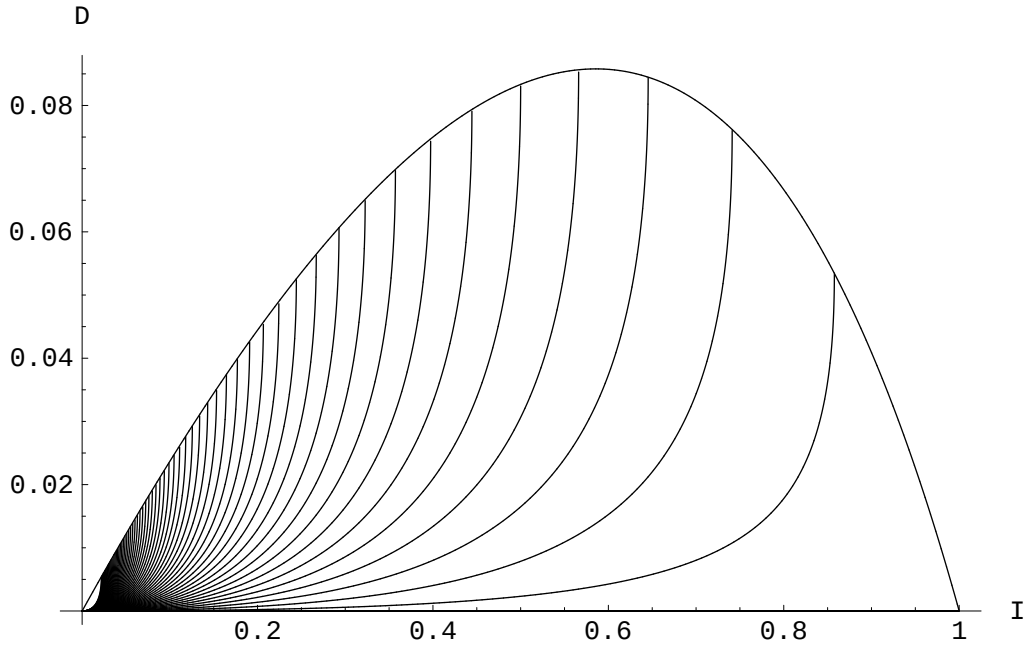


Figure 4.5: Balanced probabilities $p_{\pm} = 1/2$.

By equation (4.30) information is $I = c\sin^2(2\theta) \leq I_{max} = 1 - \cos(2\theta)$ and thus

$$D(I) = \frac{1}{2}\left[1 - \cos(2\theta)\right]\left(1 - \frac{I}{\sin^2(2\theta)} + \sqrt{\frac{I^2}{\sin^2(2\theta)} - 2\frac{I}{\sin^2(2\theta)} + 1}\right). \tag{4.46}$$

The curves are plotted in figure 4.5 for various angles θ. The rightmost curves correspond to nearly orthogonal states, while the leftmost to nearly parallel states.

We note that as the orthogonality increases, the tradeoff becomes more and more favourable to the eavesdropper, because she can extract a good amount of information while generating little disturbance. This gets along with the intuitive notion that orthogonal states are classical, and thus represent accessible, 'public' information. On the other hand, nearly parallel states are so poorly distinguishable that there is no way to extract from

them information about the choice of the preparer.

The unbalanced case $p_+ \neq p_-$ can be treated numerically. In figure 4.6 we plot the curves for 'mutually unbiased' states $2\theta = \pi/4$, for various *a priori* probabilities. The uppermost curve represents again symmetric probabilities. Here we note that as one of the states becomes more probable, the information becomes more accessible (because, in a sense, it is already 'public') and the eavesdropper can extract it more easily. In the same way, the disturbance diminishes because the measurement performed by Bob is more and more 'focused' on the most probable state. Eventually, when the odds ratio of the *a priori* probabilities exceeds the threshold value derived in section 4.2.1 Bob's measurement can no longer detect with certainty an eavesdropping attempt and the tradeoff curves flatten down on the horizontal axis. The threshold odds ratio $R = 1/\cos^2(2\theta)$ corresponds to $p_+ = 1/(1+\cos^2(2\theta))$, which gives the maximal information

$$I_{max} = \frac{1-\cos^2(2\theta)}{1+\cos^2(2\theta)} \tag{4.47}$$

The bell-shaped curves corresponding to the extremal points of the tradeoff curves are plotted in figure 4.8, for various *a priori* probabilities. We observe that uniform probability distribution corresponds to higher disturbance, while unbalanced distributions yield curves that are more and more 'squeezed' towards the bottom-right corner of the plot (favourable to the eavesdropper).

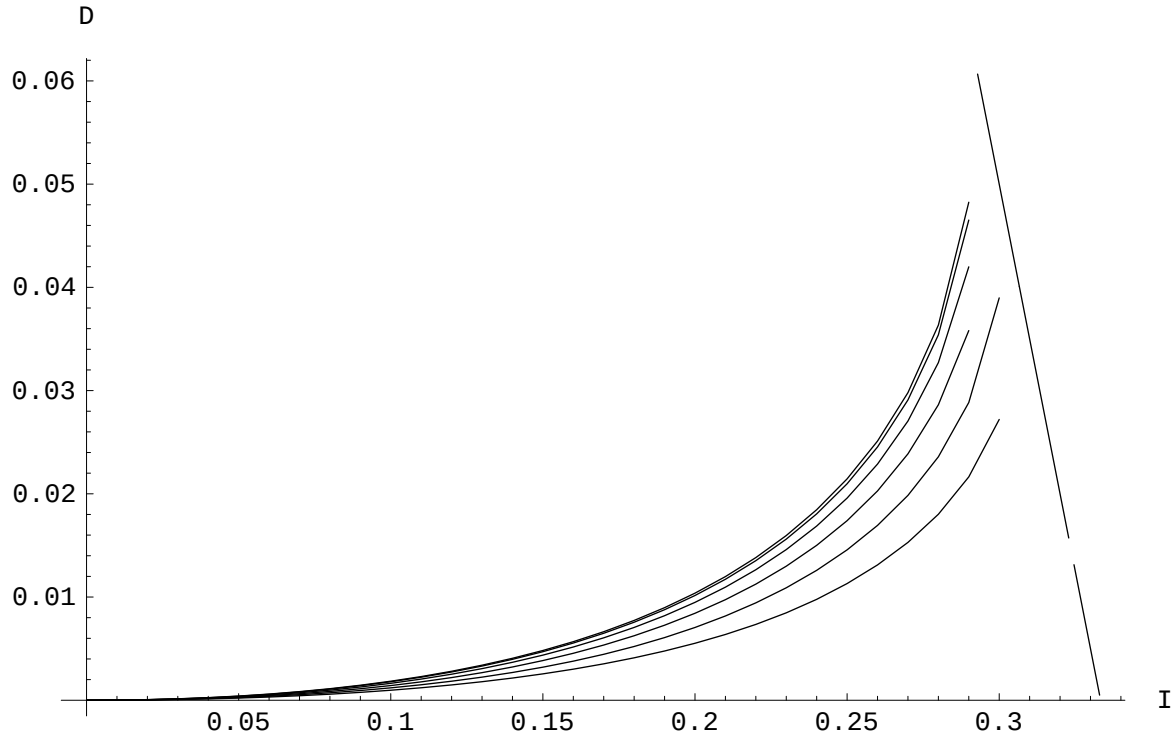


Figure 4.6: Unbalanced probabilities $p_+ \neq p_-$ for $2\theta = \pi/4$

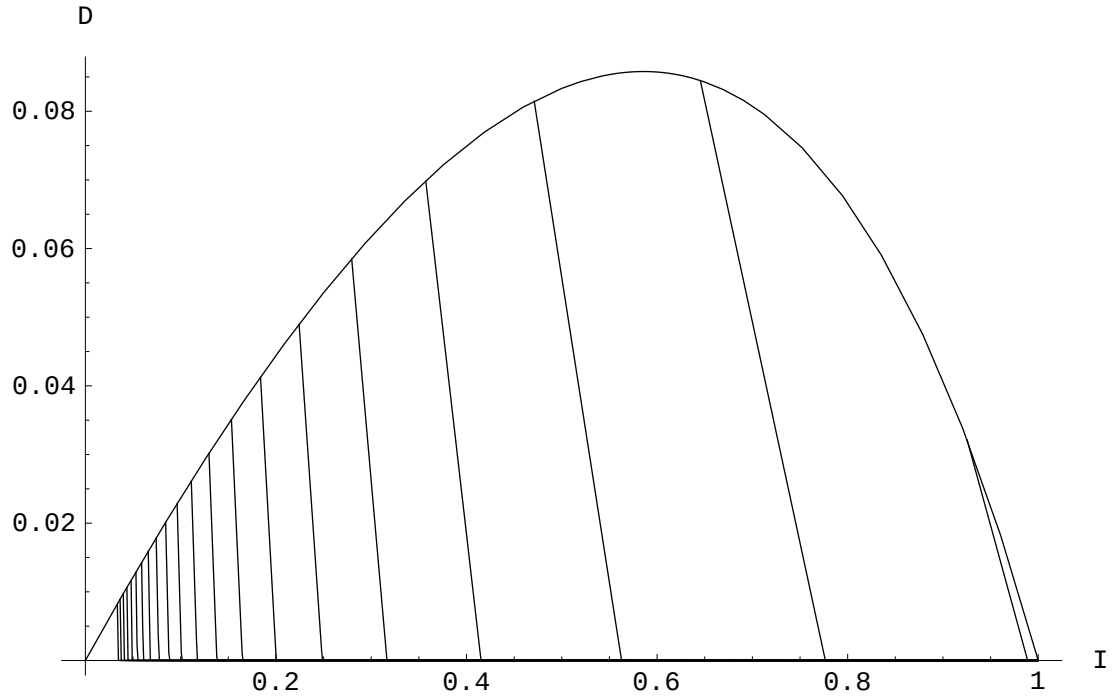


Figure 4.7: Extremal tradeoff points. Each curve is obtained plotting the extremal points of the tradeoff curves for a fixed separation angle 2θ and varying the ratio $R = p_+/p_-$. Rightmost curves correspond to nearly orthogonal states, while leftmost curves to nearly parallel states.

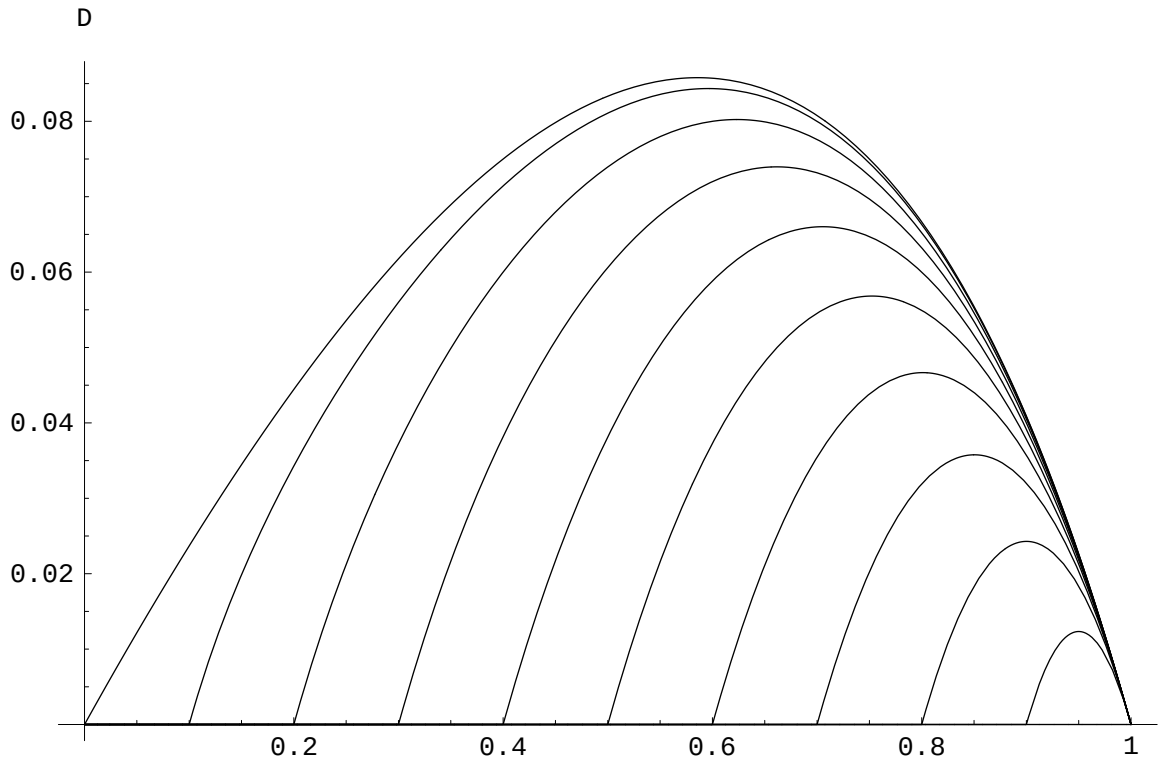


Figure 4.8: Extremal tradeoff points. Each curve is obtained plotting the extremal points of the tradeoff curves for a fixed $R = p_+/p_-$ and varying the separation angle 2θ. The uppermost curve corresponds to the uniform distribution $p_\pm = 1/2$.

Chapter 5

State transformations

5.1 Probabilistic transformation of a pair of states

We are given an ensemble $\mathsf{E} = \{q_\pm, |\psi_\pm\rangle\langle\psi_\pm|\}$ of two pure states $|\psi_\pm\rangle$ with uniform *a priori* probabilities $q_\pm = 1/2$, and a pair of (generally mixed) states $\rho_\pm$. We want to find a quantum operation which realizes the transformation

$$|\psi_\pm\rangle \longrightarrow \rho_\pm \tag{5.1}$$

maximizing the mean probability of success over the ensemble.

For pure final states $\rho_\pm = |\phi_\pm\rangle\langle\phi_\pm|$ the problem has been essentially solved in [CB98]

Proposition 5.1.1. *Under the aformentioned hypotheses, the maximum mean probability is*

$$p = \min\left\{\frac{1 - |\langle\psi_+|\psi_-\rangle|}{1 - |\langle\phi_+|\phi_-\rangle|}, 1\right\}. \tag{5.2}$$

Moreover, this probability is achieved with a balanced transformation, i.e. a transformation occuring with equal probability on both initial states.

Proof. If $|\langle\phi_+|\phi_-\rangle| \geq |\langle\psi_+|\psi_-\rangle|$ we can always extend the system with an ancilla whose initial state is some pure state $|0\rangle$, and find a pair of ancillary states $|\alpha_\pm\rangle$ such that

$$\langle\psi_+|\psi_-\rangle = \langle\phi_+|\phi_-\rangle\langle\alpha_+|\alpha_-\rangle. \tag{5.3}$$

Thus, the transformation

$$|\psi_\pm\rangle|0\rangle \longrightarrow |\phi_\pm\rangle|\alpha_\pm\rangle \tag{5.4}$$

is isometric and can be obtained with a unitary operation on the compound system. This realizes the required trace preserving quantum operation on the original system.

Now, suppose that $|\langle\phi_+|\phi_-\rangle| \leq |\langle\psi_+|\psi_-\rangle|$. A general quantum operation realizing this transformation can be described by a set of Kraus operators $\{A_k\}$ (see theorem 1.2.2) such that

$$A_k|\psi_\pm\rangle = \mu_{k\pm}|\phi_\pm\rangle \tag{5.5}$$

where $\mu_{k\pm}$ are complex numbers satisfying $\sum_k |\mu_{k\pm}|^2 \leq 1$. These are exactly the success probabilities $p_{s\pm} = \sum_k |\mu_{k\pm}|^2$. The mean probability is therefore

$$p_s = \frac{1}{2}\sum_k \left(|\mu_{k+}|^2 + |\mu_{k-}|^2\right). \tag{5.6}$$

A bound for p_s can be found imposing on the positive operator $\Pi_s = \sum_k A_k^\dagger A_k$ the condition $\Pi_s \leq I$, which can be expressed as a constraint on its eigenvalues $\lambda_{1,2}$:

$$\lambda_{1,2} \leq 1. \tag{5.7}$$

Consider any state which is a superposition of $|\psi_\pm\rangle$:

$$|\psi\rangle = \frac{1}{\sqrt{N}}\sum_{r=\pm} c_r|\psi_r\rangle \tag{5.8}$$

where the normalization factor is $N = \sum_{r,r'} c^*_{r'} c_r \langle \psi_{r'} | \psi_r \rangle$. The condition $\langle \psi | \Pi_s | \psi \rangle \leq 1$ can be written as

$$\begin{pmatrix} c^*_+ & c^*_- \end{pmatrix} \begin{pmatrix} p_{s+} & Q\beta - \alpha \\ Q^*\beta^* - \alpha^* & p_{s-} \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \leq 1, \tag{5.9}$$

where $Q = \sum_k \mu^*_{k+} \mu_{k-}$, $\alpha = \langle \psi_+ | \psi_- \rangle$ and $\beta = \langle \phi_+ | \phi_- \rangle$. The condition (5.7) becomes a condition on the eigenvalues of the matrix in equation (5.9)

$$(1 - p_{s+})(1 - p_{s-}) \geq |\alpha - Q\beta|^2. \tag{5.10}$$

We observe that

$$(1 - p_s)^2 \geq (1 - p_{s+})(1 - p_{s-}) \tag{5.11}$$

and the equality is satisfied only when $p_{s\pm} = p_s$. From the triangle inequality we have

$$|\alpha - Q\beta| \geq |\alpha| - |Q||\beta| \tag{5.12}$$

where the equality is obtained only when α and $Q\beta$ have the same phase. Moreover, the Cauchy-Schwartz inequality gives

$$|Q| \leq \sqrt{p_{s+} p_{s-}} \leq p_s \tag{5.13}$$

Here, the first inequality is satisfied when μ_{k+} is proportional to μ_{k-}. Toghether with (5.11) it gives $\mu_{k+} = \mu_{k-} e^{i\theta}$ for some angle θ. Combining this inequality with (5.12) we obtain

$$|\alpha| - |Q||\beta| \geq |\alpha| - p_s|\beta| \tag{5.14}$$

and finally

$$p_s \leq \frac{1 - |\langle \psi_+ | \psi_- \rangle|}{1 - |\langle \phi_+ | \phi_- \rangle|} \tag{5.15}$$

where the equality requires that $p_{s\pm} = p_s$ and that the phases of the overlaps between the initial states $|\psi_\pm\rangle$ and the final states $\mu_{k\pm}|\phi_\pm\rangle$ are equal. This condition can be easily satisfied since we can change the phase of the states without altering their physical meaning. □

Indeed, this formula can be extended also to final mixed states.

Proposition 5.1.2. *For generally mixed final states $\rho_\pm$ the maximum mean probability is*

$$p = \min\left\{\frac{1-|\langle\psi_+|\psi_-\rangle|}{1-F(\rho_+,\rho_-)}, 1\right\} \tag{5.16}$$

Moreover, the probability is achieved with a balanced transformation.

Proof. Suppose we have a quantum operation $\mathcal{E}$ realizing the transformation

$$|\psi_\pm\rangle \longrightarrow \rho_\pm, \tag{5.17}$$

with certain probabilities $p_\pm$. Exploiting the dilation theorem 1.2.1 we can represent this quantum operation in the following way

$$\mathcal{E}(\rho) = \mathrm{Tr}_2(U\rho\otimes|0\rangle\langle 0|U^\dagger P) \tag{5.18}$$

where U is unitary operator and P is a projector. Since unitaries and projectors cannot turn a pure state into a mixed one, the quantum operation $\mathcal{E}$, when applied the our initial states $|\psi_\pm\rangle$, will have the form

$$\mathcal{E}(|\psi_\pm\rangle\langle\psi_\pm|) = p_\pm\,\mathrm{Tr}_2(|\varphi_\pm\rangle\langle\varphi_\pm|) \tag{5.19}$$

where $|\varphi_\pm\rangle$ are suitable normalized states and $p_\pm$ are the success probabilities. We note that $|\varphi_\pm\rangle$ are actually purifications of the final states $\rho_\pm$.

In this way we have proved that every transformation $|\psi_\pm\rangle \to \rho_\pm$ can be realized with a corresponding trasformation between pure states $|\psi_\pm\rangle \to |\varphi_\pm\rangle$. Thus, in order to maximize the probability of $|\psi_\pm\rangle \to \rho_\pm$ it is not restrictive to scan only those transformations which take $|\psi_\pm\rangle$ into purifications of the final states $\rho_\pm$.

From Uhlmann's theorem 3.2.2 we have that

$$|\langle\varphi_+|\varphi_-\rangle| \le F(\rho_+,\rho_-), \tag{5.20}$$

for all the purifications of $\rho_\pm$, and thus

$$\frac{1-|\langle\psi_+|\psi_-\rangle|}{1-|\langle\varphi_+|\varphi_-\rangle|} \leq \frac{1-|\langle\psi_+|\psi_-\rangle|}{1-F(\rho_+,\rho_-)}. \tag{5.21}$$

From the previous proposition we already know that the maximum probability for $|\psi_\pm\rangle \to |\varphi_\pm\rangle$ is (5.2) and thus there is the upper bound

$$p \leq \min\left\{\frac{1-|\langle\psi_+|\psi_-\rangle|}{1-F(\rho_+,\rho_-)}, 1\right\}. \tag{5.22}$$

This bound can be achieved by choosing the purification which gives the equality in equation (5.20). Since, by the previous proposition, the transformation happens with the same probability on both states, the proposition is proved. □

5.2 Probability/fidelity tradeoff

Taking advantage of the results obtained in the previous sections, we can derive an interesting tradeoff for the transformation

$$|\psi_\pm\rangle \longrightarrow |\varphi_\pm\rangle, \quad |\langle\varphi_+|\varphi_-\rangle| \leq |\langle\psi_+|\psi_-\rangle| \tag{5.23}$$

if we allow it to be approximate, that is, if we admits quantum operations which transform $|\psi_\pm\rangle$ into some states $\rho_\pm$

$$|\psi_\pm\rangle \longrightarrow \rho_\pm = \frac{\mathcal{E}(|\psi_\pm\rangle\langle\psi_\pm|)}{p_\pm}, \quad p_\pm = \mathrm{Tr}(\mathcal{E}(|\psi_\pm\rangle\langle\psi_\pm|)) \tag{5.24}$$

close to $|\varphi_\pm\rangle$, though not exactly equal. In this case there are two incompatible parameters of quality: the probability of success and the fidelity between the target states and the states actually obtained. Intuitively, the more we try to 'stretch' the pair $|\psi_\pm\rangle$ towards the target states, the less the transformation is likely to happen.

In order to quantify the tradeoff, we choose as precise parameters the minimum probability and the minimum fidelity over the two states (in other words, a *worst case* criterion):

$$\mathsf{p} = \min\{p_+, p_-\} \tag{5.25}$$

$$\mathsf{F} = \min\{F(|\varphi_+\rangle, \rho_+), F(|\varphi_-\rangle, \rho_-)\}\,. \tag{5.26}$$

An immediate consequence of this choice is that, looking for the optimum frontier, without loss of generality we can search among the operations whose final states have Bloch vectors $\boldsymbol{r}_{\rho_\pm}$ complanar with respect to the Bloch vectors of the target states $\boldsymbol{r}_{|\varphi_\pm\rangle}$, and with the same symmetry axis. In fact, if the initial states are taken to a position which is not complanar to the target states, the transformation is surely non-optimal, since one can always bring the states to be complanar (and thus closer to the target) performing a unitary operation which, being deterministic, does not diminish the probability.

Moreover, for each operation $\mathcal{E}$ realizing a certain transformation

$$\mathcal{E}(|\psi_\pm\rangle\langle\psi_\pm|) = p_\pm\rho_\pm, \tag{5.27}$$

where $\rho_\pm$ are complanar to $|\varphi_\pm\rangle$, we can construct an operation $\mathcal{E}'$ acting in the following way

$$\mathcal{E}'(|\psi_\pm\rangle\langle\psi_\pm|) = \frac{1}{2}(p_\pm\rho_\pm + p_\mp\sigma_z\rho_\mp\sigma_z), \tag{5.28}$$

where σ_z is the rotation around the symmetry axis of the pair $|\varphi_\pm\rangle$. The second term in r.h.s. is simply the "mirror image" of $\mathcal{E}$. This new quantum operation is symmetric since $\sigma_z\mathcal{E}'(|\psi_\pm\rangle\langle\psi_\pm|)\sigma_z = \mathcal{E}'(|\psi_\mp\rangle\langle\psi_\mp|) = \mathcal{E}'(\sigma_z|\psi_\pm\rangle\langle\psi_\pm|\sigma_z)$ and behaves better than the original one with respect to *both* the quality parameters:

$$\mathrm{Tr}(\mathcal{E}'(|\psi_\pm\rangle\langle\psi_\pm|)) = \frac{1}{2}(p_+ + p_-) \geq \min\{p_+, p_-\} \tag{5.29}$$

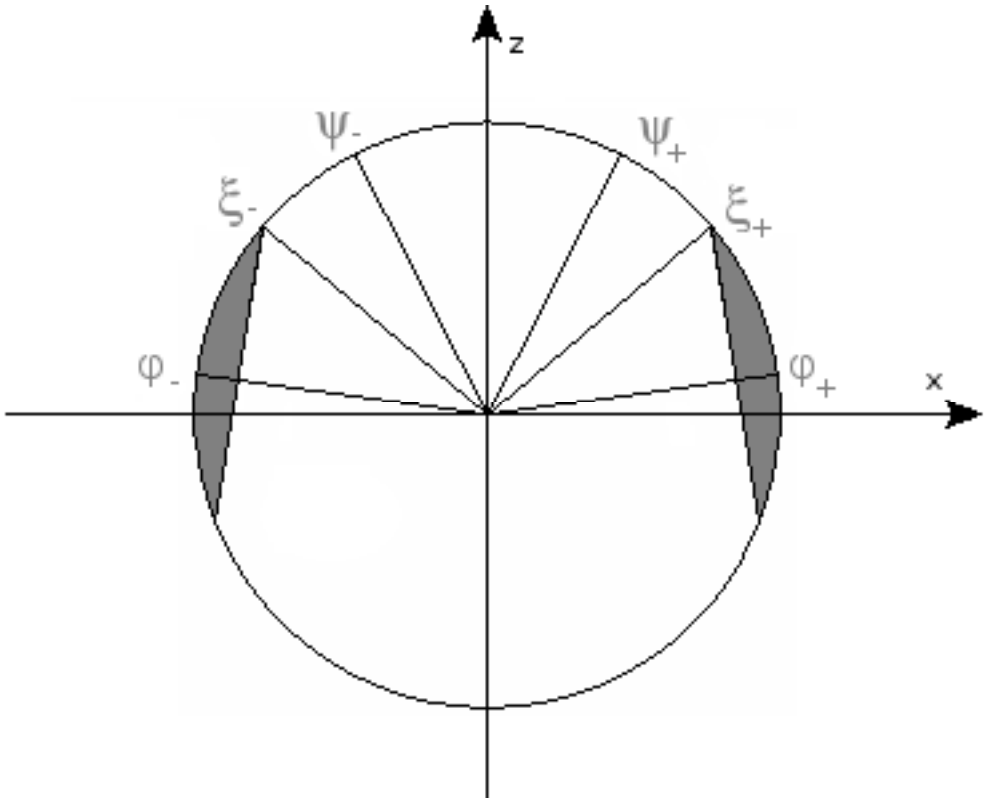


Figure 5.1: The section of the Bloch ball containing the initial pair $|\psi_\pm\rangle$ and the target pair $|\varphi_\pm\rangle$. The shadowed area embraces all the states $\rho_\pm$ with fidelity $F(\rho_\pm, |\varphi_\pm\rangle) \geq |\langle\xi_\pm|\varphi_\pm\rangle|$.

and

$$F(|\varphi_\pm\rangle, \mathcal{E}'(|\psi_\pm\rangle\langle\psi_\pm|)) \geq \min\{F(|\varphi_+\rangle, \rho_+), F(|\varphi_-\rangle, \rho_-)\}. \tag{5.30}$$

Having proved that the frontier is made up of symmetric transformations, we move on to its explicit calculation. We notice that we can assume the initial states $|\psi_\pm\rangle$ to be in the symmetric configuration complanar to $|\varphi_\pm\rangle$, because this position can always be reached via a unitary operation which, being deterministic, does not affect the probability. In the p/F plane this configuration corresponds to the point $(1, f_0)$, where $f_0 := |\langle\psi_+|\varphi_+\rangle| \equiv |\langle\psi_-|\varphi_-\rangle|$.

Now, let f, $f_0 \leq f \leq 1$ be the fidelity we want to achieve. Then, the set

of possible final states $\rho_\pm$ compatible with the constraint $F(\rho_\pm, |\varphi_\pm\rangle) \geq f$ is the shadowed area depicted in figure 5.1, where $|\xi_\pm\rangle$ are pure states such that $|\langle\xi_\pm|\varphi_\pm\rangle| = f$. This follows from the fact, proved in section 3.3, that the isofidelity surfaces for the pure state $|\varphi_+\rangle$ (or $|\varphi_-\rangle$) are planes orthogonal the Bloch vector $\boldsymbol{r}_{|\varphi_+\rangle}$ (respectively, $\boldsymbol{r}_{|\varphi_-\rangle}$). We claim that the most probable attainable configuration is the pair $|\xi_\pm\rangle$. We need to prove that the probability as obtained by proposition 5.1.2

$$p = \frac{1 - |\langle\psi_+|\psi_-\rangle|}{1 - F(\rho_+, \rho_-)} \tag{5.31}$$

where $\rho_\pm$ is any symetric pair inside the area, reaches the maximum for the pair $|\xi_\pm\rangle$. This would follow easily provided that the fidelity $F(\rho_+, \rho_-)$ has its maximum for the pair $|\xi_\pm\rangle$. To prove the claim, thus, we only need to compute the fidelity $F(\rho_+, \rho_-)$ for states of the form $\rho_\pm = \frac{1}{2}(I \pm \beta\sigma_x + \gamma\sigma_z)$; but, relying on Hübner's expression (3.23), one can immediately obtain

$$F(\rho_+, \rho_-) = \sqrt{1 - \beta^2}. \tag{5.32}$$

This clearly shows that the optimal states are those which minimize β, a task very well accomplished by the pair $|\xi_\pm\rangle$.

The remaining part of the optimal tradeoff curve can now be completed quite easily; we only need to sweep along the pure states comprised in the arc between $|\psi_\pm\rangle$ and $|\varphi_\pm\rangle$ to obtain the points connecting $(1, f_0)$ and $(p_0, 1)$, where $p_0 = (1 - |\langle\psi_+|\psi_-\rangle|)/(1 - |\langle\varphi_+|\varphi_-\rangle|)$. The explicit expression for this part of the curve is

$$F(p) = \cos\left(\frac{\arccos|\langle\varphi_+|\varphi_-\rangle| - \arccos\left(1 - \frac{1-|\langle\psi_+|\psi_-\rangle|}{p}\right)}{2}\right). \tag{5.33}$$

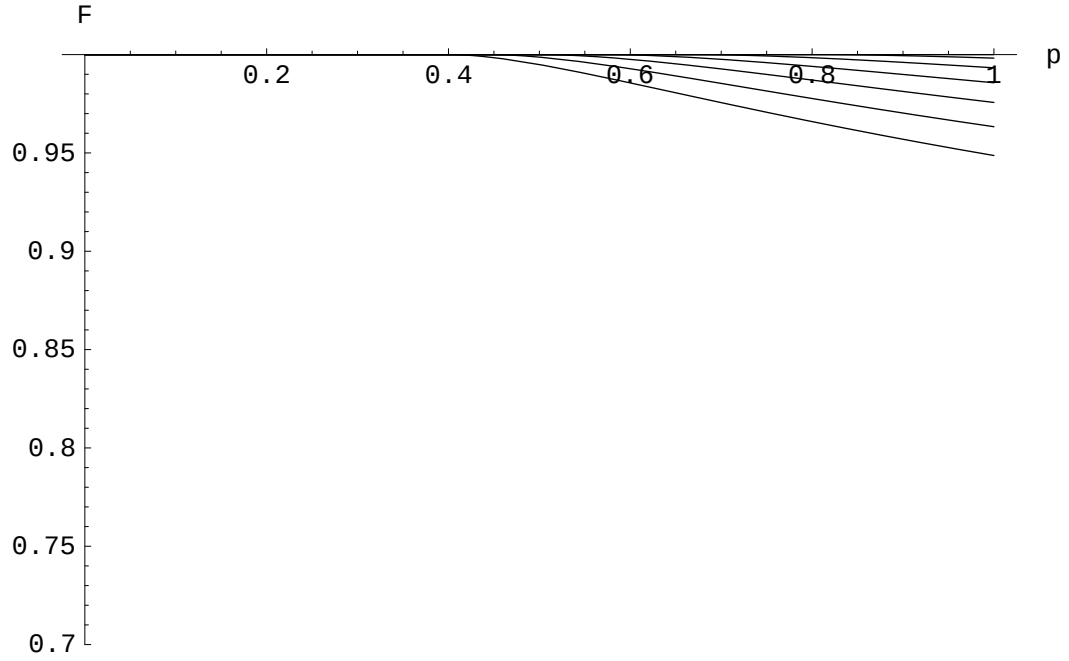


Figure 5.2: Tradeoff curves $F(p)$ for $|\langle\psi_+|\psi_-\rangle| = 0.6$ and for $|\langle\varphi_+|\varphi_-\rangle| = 0, \ldots, 0.6$, at intervals of 0.1

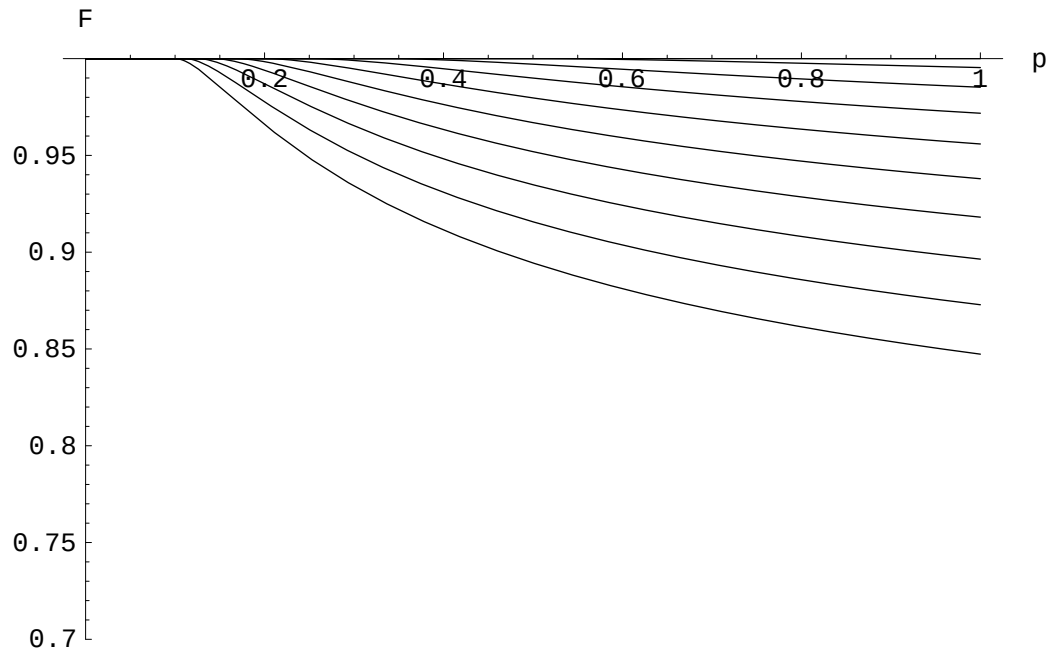


Figure 5.3: Tradeoff curves $F(p)$ for $|\langle\psi_+|\psi_-\rangle| = .9$ and for $|\langle\varphi_+|\varphi_-\rangle| = 0, \ldots, 0.9$, at intervals of 0.1

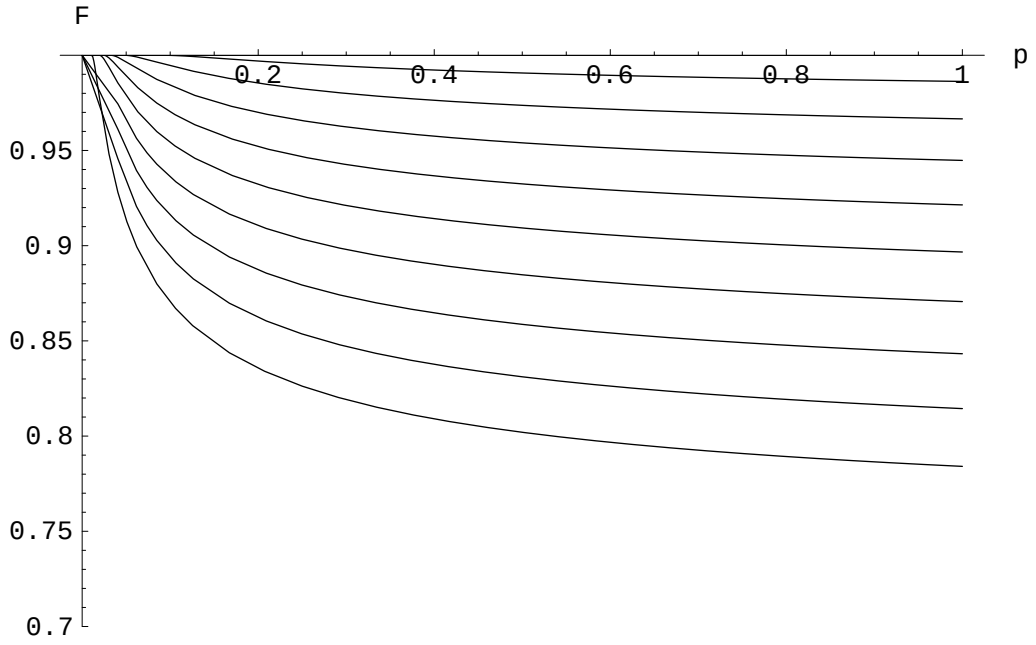


Figure 5.4: Tradeoff curves $F(p)$ for $|\langle\psi_+|\psi_-\rangle| = .99$ and for $|\langle\varphi_+|\varphi_-\rangle| = 0.09, \ldots, 0.99$, at intervals of 0.1

5.3 Tradeoff in the inversion of a contraction

In quantum information theory an important question is to find out whether a given transformation $\mathcal{E}$ can be inverted deterministically on some subspace $L \subseteq \mathcal{H}$, in other words whether there is a quantum channel $\mathcal{R}$ such that

$$\rho \longrightarrow \rho' = \frac{\mathcal{E}(\rho)}{\mathrm{Tr}(\mathcal{E}(\rho))} \longrightarrow \mathcal{R}(\rho') = \rho \tag{5.34}$$

for every ρ, such that $\mathrm{supp}(\rho) \subseteq L$.

Necessary and sufficient conditions for this inversion have been proved by Knill and Laflamme [KL96], while Schumacher and Nielsen [SN96] provided an equivalent condition based on information-theoretic quantities such as entropy and coherent information.

If the transformation is not invertible, or the inversion is not required to be perfect, it is still possible to perform an approximate correction which

brings ρ' *close* to ρ. This "closeness" have been quantified by Schumacher and Westmoreland [SW02], whenever $\mathcal{E}$ is a channel, and by Buscemi *et al.* [BHH07] for general quantum operations.

Approximate or not, all these inversions are deterministic; the operation $\mathcal{R}$ is always required to be a channel. If we expand the set of possible operations, including probabilistic ones, a wider and largely unexplored scenario emerges. For example, we could be interested in an *exact* inversion, even if we are not guaranteed that it would happen with certainty; or in any kind of inversion interpolating between the latter and the best approximate deterministic one.

Unfortunately, these situations become soon quite awkward and very hard to manage, because of the mathematical complexity of the calculations involved. However, the prominent qualitative features are already visible in some simple cases, which can be worked out quite easily.

In the following we will focus on a two-level system, being acted upon by an operation $\mathcal{E}$ consisting of a single contraction M. Since every operator M can be decomposed as the product of a unitary operator and a positive one $M = UP$, and the unitary part can always be removed at no cost, we will assume M to coincide with its positive part, which has the following matrix representation:

$$M_\beta = \begin{pmatrix} 1 & 0 \\ 0 & \beta \end{pmatrix} \tag{5.35}$$

where β, $0 \leq \beta \leq 1$, is the smaller singular value. The largest singular value is fixed at 1 because every contraction is equivalent to M_β for some β up to a rescalement, which affects its output only by an overall factor, constant for all the states. This amounts to a global rescalement of the probabilities, which becomes unimportant, since we always take for granted that the operation

has happened.

We will consider two case-studies of the following form: there is some set of density operators $\mathcal{D}$ (not necessarily a subspace) consisting of all the possible initial states of the system. After the transformation represented by M_β has happened, we want to invert it choosing among a set $\mathcal{Q}$ of quantum operations. Suppose that the the system was initially in the state $\rho \in \mathcal{D}$. After the first transformation the state becomes

$$\rho' = \frac{M_\beta \rho M_\beta}{\mathrm{Tr}(\rho M_\beta^2)}. \tag{5.36}$$

A subsequent quantum operation $\mathcal{R} \in \mathcal{Q}$ leaves the system in the state

$$\rho'' = \frac{\mathcal{R}(\rho')}{\mathrm{Tr}(\mathcal{R}(\rho'))}. \tag{5.37}$$

The quality of this inversion is determined by two parameters: the probability of success

$$p(\mathcal{R};\rho) = \mathrm{Tr}(\mathcal{R}(\rho')) \tag{5.38}$$

and the fidelity between the initial state and the corrected one

$$f(\mathcal{R};\rho) = F(\rho,\rho''). \tag{5.39}$$

If we want the probability of success never to drop under some threshold value $\overline{p}$, we can consider only the subset $\mathcal{Q}' \subset \mathcal{Q}$ whose elements satisfy the constraint:

$$p(\mathcal{R};\rho) \geq \overline{p}, \quad \forall \rho \in \mathcal{D}. \tag{5.40}$$

In a worst case criterion we have to choose the inversion $\overline{\mathcal{R}} \in \mathcal{Q}'$ that maximizes the minimum fidelity over the set $\mathcal{D}$

$$\overline{\mathcal{R}} = \arg \max_{\mathcal{R}\in\mathcal{Q}'} \min_{\rho\in\mathcal{D}} f(\mathcal{R};\rho). \tag{5.41}$$

This gives the point $(\overline{p},\overline{F})$, with $\overline{F} = \min_{\rho\in\mathcal{D}} f(\overline{\mathcal{R}};\rho)$, in the $\mathsf{p/F}$ plane. The tradeoff curve is obtained varying $\overline{p}$ in the interval $[0,1]$.

In this way we obtain a curve $F = F(p)$ which, for every p, gives the minimum guaranteed fidelity over the set of possible initial states, achievable with a QO whose probability of success is at least p for all such states.

5.3.1 Semiclassical case

The first case is semiclassical. The set of states $\mathcal{D}$ consists of all the density operators jointly diagonal with the contraction

$$\rho_x = \begin{pmatrix} x & 0 \\ 0 & 1-x \end{pmatrix}, \quad 0 \leq x \leq 1. \tag{5.42}$$

In the same way the set of possible inversions $\mathcal{Q}$ is made up of the diagonal single-contraction operations

$$N_\gamma = \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix}, \quad \beta \leq \gamma \leq 1. \tag{5.43}$$

The extremal case N_β is the matrix inverse of M_β, rescaled in order to be a contraction, $N_\beta = M_\beta^{-1}/\|M_\beta^{-1}\|$. Obviously, it realizes the exact inversion.

Though defined by single-contraction elements, the set $\mathcal{Q}$ is more encompassing than it seems. Since its elements are bound to act only on diagonal states, it is actually closed under convex combinations

$$qN_\lambda \rho_x N_\lambda + (1-q) N_{\lambda'} \rho_x N_{\lambda'} = N_{\lambda''} \rho_x N_{\lambda''}, \tag{5.44}$$

with $\lambda'' = \sqrt{q\lambda^2 + (1-q)\lambda'^2}$, as well as under coherent sums

$$(qN_\lambda + (1-q)N_{\lambda'}) \rho_x (qN_\lambda + (1-q)N_{\lambda'}) = N_{\lambda''} \rho_x N_{\lambda''}, \tag{5.45}$$

with $\lambda'' = q\lambda + (1-q)\lambda'$.

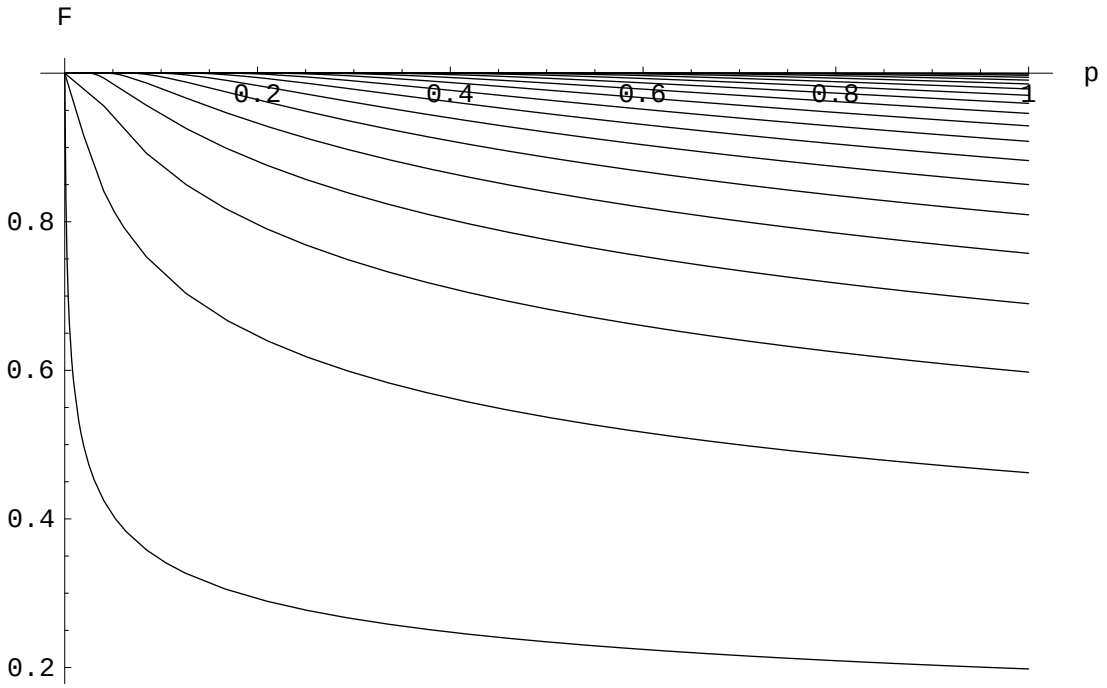


Figure 5.5: Tradeoff curves for the semiclassical case, far various β. Each curve gives the minimum guaranteed fidelity in the inversion of M_β, as a function of the minimum probability p of success over all the initial states. Upper curves are obtained for β close to 1, lower curve for β approaching 0.

The states ρ'_x and ρ''_x are easily computed

$$\rho'_x = \frac{1}{x + \beta^2(1-x)} \begin{pmatrix} x & 0 \\ 0 & \beta^2(1-x) \end{pmatrix}, \tag{5.46}$$

$$\rho''_x = \frac{1}{\gamma^2 x + \beta^2(1-x)} \begin{pmatrix} \gamma^2 x & 0 \\ 0 & \beta^2(1-x) \end{pmatrix}, \tag{5.47}$$

and so are the probability and the fidelity

$$p(N_\gamma; \rho_x) = \frac{\gamma^2 x + \beta^2(1-x)}{x + \beta^2(1-x)}, \tag{5.48}$$

$$f(N_\gamma; \rho_x) = \frac{\gamma x + \beta(1-x)}{\sqrt{\gamma^2 x + \beta^2(1-x)}}. \tag{5.49}$$

By inspection of these expressions one can see that the set $\mathcal{Q}'$ is

$$\mathcal{Q}' = \left\{ N_\gamma, \quad \gamma^2 \geq \overline{p} \right\} \tag{5.50}$$

and that

$$\arg \max_{N_\gamma \in \mathcal{Q}'} \min_{\rho_x} f(N_\gamma; \rho_x) = N_{\sqrt{\bar{p}}} \tag{5.51}$$

The corresponding tradeoff curves are plotted in figure 5.5 for various β. The uppermost curves are obtained when β approaches 1, i.e. when M_β is nearly the identity. Clearly, in this case there is almost no need of inversion and we have always a high fidelity. On the other hand, as β goes to zero M_β is closer and closer to be a projector which, in a worst case criterion, cannot be inverted with fidelity greater than zero.

5.3.2 Quantum case

In the second case we consider a set of two non-orthogonal states $\mathcal{D} = \{|\psi_\pm\rangle\}$, meanwhile allowing the inversion to be *any* quantum operation. The states after the first transformation are

$$|\psi'_\pm\rangle = \frac{M_\beta|\psi_\pm\rangle}{|M_\beta|\psi_\pm\rangle|}. \tag{5.52}$$

The required inversion is

$$|\psi'_\pm\rangle \longrightarrow |\psi_\pm\rangle \tag{5.53}$$

so this case is simply a rephrasing of the results obtained is sections 5.1 and 5.2.

Summary of results

In chapter 3 we addressed the problem of deriving the isofidelity surfaces of some state ρ. These surfaces are constituted by all the states σ having the same fidelity with respect to ρ. In the general case we have shown how to simplify the problem exploiting symmetry arguments, while for the qubit we have provided the analytical solution, resorting to Hübner's formula (3.23).

In chapter 4 we studied a communication protocol inspired by Quantum State Discrimination. We have discussed operational criteria to define the information and the disturbance in this setting, and computed the tradeoff curves in the symmetric case $p_+ = p_-$. The tradeoff curves in the asymmetric case have been plotted numerically, while the extremal points (corresponding to maximal information) are obtained analitycally.

In chapter 5, following an idea proposed in [Da03], we analyzed the probabilistic transformation of states. For two pure states we give a slight generalization of a known result ([CB98]), and apply it to derive a tradeoff curve between the probability and the fidelity of the transformation. We also studied the probabilistic inversion of a single-contraction QO, varying the set of possible initial states.

Concluding remarks

Information is a difference which makes a difference

Gregory Bateson

...and disturbance, too! Joking aside, in this work we decided to turn our attention to a QSD scenario and analyze the protocol discussed in section 4.3, not only because it is a paradigmatic example of the problem, but also because in this case there are easy-to-grasp operational definitions of information and disturbance, directly motivated by a communicational and cryptographical set. The presence of "agents", typical of these scenarios, helps in identifying the relevant operational notions for the problem under examination: "Eve" learns what is the state or not, "Alice" and "Bob" catch or don't catch her, etc. In other words, we can always tell the "difference".

However, there are other possible vistas on the information/disturbance problem, which are mainly concerned about the original question posed by Heisenberg: how the measurement of some observable influences non-commuting observables? This is related, for example, to the well-known problem of the Standard Quantum Limit, *vexata quaestio* of quantum measurement theory originated by the debate around the possibility of measuring small classical forces like gravitational waves.

Addressing these topics, difficulties often arise because of the lack of a clear-cut definition of disturbance. There are a couple of resolutive papers by Ozawa [Oz03] [Oz04], where these problems are carefully analyzed (although some isolated voices have recently questioned them, for example Busch [Bu07], where Ozawa's notion of "noise" is rejected as non-operational).

The main teaching that we want to underline is the following: the disturbance on some system Q provoked by some agent - say "Eve" - can be quantified only if we decide what kind of checks can or can not be performed on Q by someone else - "Alice". It would be an interesting research line to take advantadge of what we learned in the non-controversial QSD setting, and try to apply the same perspective in other cases.

Bibliography

[Ba00] H. Barnum, *Information-disturbance Tradeoff in Quantum Measurement on the Uniform Ensemble and on the Mutually Unbiased Bases.* Report University of Bristol (2000)

[BD01] K. Banaszek and I. Devetak, *Fidelity Trade-off for Finite Ensembles of Identically Prepared Qubits.*, Phys. Rev. A **64**, 052307 (2001)

[BHH07] F. Buscemi, M. Hayashi and M. Horodecki, *Information Gain and Approximate Reversibility of Quantum Measurements: an Entropic Approach.* arXiv:quant-ph/0702166 (2007)

[BS06] F. Buscemi and M. F. Sacchi, *Information-Disturbance Tradeoff in Quantum State Discrimination.* Phys. Rev. A **74**, 052320 (2006)

[Bu07] P. Busch, *"No Information without Disturbance": Quantum Limitations of Measurements.* arXiv:quant-ph/0706.3526 (2007)

[BZ06] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States.* Cambridge University Press, Cambridge 2006.

[CB98] A. Chefles and S. M. Barnett, *Quantum State Separation, Unambiguous Discrimination and Exact Cloning.* J. Phys. A **31** 10097-10103 (1998)

[CFS02] C. M. Caves, C. A. Fuchs and R. Schack, *Quantum Probabilities as Bayesian Probabilites.* Phys. Rev. A **65**, 022305 (2002)

[Da03] G. M. D'Ariano, *On the Heisenberg Principle, Namely on the Information-Disturbance Tradeoff in a Quantum Measurement.* Fortschr. Phys. **51** No. 4-5, 318-330 (2003)

[Di88] D. Dieks, Phys. Lett. A **126** 303 (1988)

[El03] Y. Eldar, *Von Neumann Measurement is Optimal for Detecting Linearly Indipendent Mixed States.* arXiv:quant-ph/0304077 (2003)

[FP96] C. A. Fuchs and A. Peres, *Quantum-state Disturbance Versus Information Gain: Uncertainty Relations for Quantum Information.* Phys. Rev. A **53**, 2038-2045 (1996)

[Fu98] C. A. Fuchs, *Information Gain vs. State Disturbance in Quantum Theory.* Fortschr. Phys. **46**, 535-565 (1998)

[Fu02] C. A. Fuchs, *Quantum Mechanics as Quantum Information (and Only a Little More).* arXiv:quant-ph/0205039 (2002)

[He27] W. Heisenberg, *Über den anschaulichen Inhalt der quantenteoretischen Kinematik und Mechanik.* Zeitschrift für Physik **43**, 172-198 (1927). English translation in [WZ83]

[He30] W. Heisenberg, *The Physical Principles of Quantum Theory.* University of Chicago Press, Chicago 1930.

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

[Her75] I. Herstein, *Topics in Algebra.* Wiley, 1975.

[Ho73] A. S. Holevo, *Statistical Detection Theory for Quantum Systems.* J. Multivar. Anal. **3**, 337 (1973)

[Hü92] M. Hübner, *Explicit Computation of the Bures Distance for Density Matrices.* Phys. Lett. A **163**, 239 (1992)

[Iv87] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987)

[Jo94] R. Jozsa, *Fidelity for Mixed Quantum States.* J. Mod. Opt. **41**, no. 12, 2314-2323 (1994)

[JS95] G. Jaeger and A. Shimony, *Optimal Distinction Between Two Non Orthogonal Quantum States.* Phys. Lett. A **197** 83 (1995)

[Ke28] E. H. Kennard, *Note on Heisenberg Indetermination Principle.* Phys. Rev. **31**, 344-348 (1928)

[Ken73] R. S. Kennedy, Mass. Inst. Tech. Res. Lab. Electron. Quart. Prog. Rep. No 110 142 (1973)

[KL96] E. Knill and R. Laflamme, *A Theory of Quantum Error-Correcting Codes.* arXiv:quant-ph/9604034 (1996)

[Kr83] K. Kraus, *States, Effects and Operations. Fundamental Notions of Quantum Theory.* Springer-Verlag, Berlin Heidelberg 1983.

[NC00] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge 2000.

[Oz03] M. Ozawa, *Universally Valid Reformulation of the Heisenberg Uncertainty Principle on Noise and Disturbance in Measurements.* Phys. Rev. A **67**, 042105 (2003)

[Oz04] M. Ozawa, *Uncertainty Relation for Noise and Disturbance in Generalized Quantum Measurements.* Ann. Phys. (N.Y.) **311** 350-416 (2004)

[Pe88] A. Peres, Phys. Lett. A **128**, 19 (1988)

[Ro29] H. P. Robertson, *The Uncertainty Principle.* Phys. Rev. **34**, 163-164 (1929)

[Ro34] H. P. Robertson, *An Indeterminacy Relation for Several Observables and Its Classical Interpretation.* Phys. Rev. **46**, 794-801 (1934)

[Ru28] A. E. Ruark, Phys. Rev. **31**, 311-312 (1928)

[St55] W. F. Stinespring, *Positive Functions on C^*-algebras.* Proc. Amer. Math. Soc. **6** 211 (1955)

[SN96] B. Schumacher and M. Nielsen, *Quantum Data Processing and Error Correction.* Phys. Rev. A **54** 2629 (1996)

[SW02] B. Schumacher and M. Westmoreland, *Approximate Quantum Error Correction.* Quantum Information Processing **1** No 1-2, 5-12 (2002)

[Ti04] C. G. Timpson, *Quantum Information Theory and the Foundations of Quantum Mechanics.* PhD thesis, Oxford University.

[Uh76] A. Uhlmann, *The Transition Probability in the State Space of a *-algebra.* Rep. Mat. Phys. **9**, 273 (1976)

[Wi31] E. P. Wigner, *Gruppentheorie.* Frederik Wieweg und Sohn, Braunschweig, Germany 1931. English translation: *Group theory.* Academic Press Inc., New York 1959.

[WZ83] J. A. Wheeler and W. H. Zurek, editors, *Quantum Theory and Measurement.* Princeton University Press, Princeton 1983.